The t-spacing test for a spiked tensor

Jean-Marc Azaïs[1], Federico Dalmao[2], and Yohann De Castro*[3,4]

[1]Institut de Mathématiques de Toulouse, Université de Toulouse, France.
[2]DMEL, CENUR Litoral Norte, Universidad de la República, Salto, Uruguay.
[3]Institut Camille Jordan, CNRS UMR 5208, École Centrale Lyon, France.
[4]Institut Universitaire de France (IUF)

November 27, 2026

Abstract

We observe one symmetric tensor $Y = \lambda e^{\otimes k} + \sigma W$ of odd order k on $\mathbb{R}^d$, with Gaussian noise W, and test $\lambda = 0$. The observation does not determine the noise level σ, so a test that uses it needs external information. The t-spacing test of Azaïs, Dalmao, and De Castro needs none and is exact at every σ, k and d, in closed form. We study its power. At every λ we bound the probability that the test misses a signal of amplitude λ by a noncentral Student distribution function. At fixed (k, d) this probability decays exponentially in λ^2 as λ grows, and we identify the rate. The rate is below the lower bound we prove for the rate of the known-variance spacing test. It tends to this bound as d grows. As d grows at fixed k, the power tends to one when λ/σ is a large enough multiple of $\sqrt{d}$. The test also returns an estimator of the direction e, the maximiser of the profile likelihood. It is asymptotically minimax as λ grows, with the sharp constant $(d-1)/(k\lambda^2)$ for the risk $\mathbb{E}\sin^2\angle(\hat{e}, e)$: a Bayesian Cramér–Rao lower bound valid at every λ, and an upper bound with the same leading term.

Keywords. Spiked tensor model; tensor PCA; spacing test; second maximum; Kac–Rice formula; minimax estimation; Gaussian orthogonal ensemble

MSC2020 subject classifications. Primary 62F03, 60G60; secondary 62C20, 15B52

1 Introduction

Let $Y = \lambda e^{\otimes k} + \sigma W$ be an observed symmetric tensor of odd order $k \geq 3$ on $\mathbb{R}^d$, where $e \in \mathbb{S}^{d-1}$ is an unknown direction, $\lambda \geq 0$ an unknown amplitude, $\sigma > 0$ an unknown noise level, and W a Gaussian symmetric tensor normalised so that $X(t) = \langle W, t^{\otimes k}\rangle$ is the Kostlan–Shub–Smale field, $\mathbb{E}X(s)X(t) = \langle s, t\rangle^k$. Testing $\lambda = 0$ against $\lambda > 0$ is the detection problem of tensor principal component analysis (Richard and Montanari, 2014; Ben Arous, Mei, Montanari, and Nica, 2019); the maximiser of the profile likelihood $Z(t) = \langle Y, t^{\otimes k}\rangle$ is the associated estimator of e.

Exact tests are hard to come by here. The null law of $\sup Z$ has no closed form. A scale-free statistic, such as $\sup Z$ divided by an estimate of σ, can be calibrated by simulation, but each replicate costs one full optimisation of a non-convex landscape, for every pair (k, d). Comparing it with the quantile of $\sup Z/\sigma$ is not valid: at $(k, d) = (3, 8)$ this test rejects too often (Figure 1(a)). Asymptotic calibrations are of little use at the dimensions where the model is interesting. The spacing test and the t-spacing test of Azaïs, Dalmao, and De Castro (2024), related to the Kac–Rice test of Taylor, Loftus, and Tibshirani (2016), solve the calibration problem in closed form: they are uniformly distributed on $(0, 1)$ under the null, at

*Corresponding author. E-mail: yohann.de-castro@ec-lyon.fr. ORCID: 0000-0002-9008-7474.

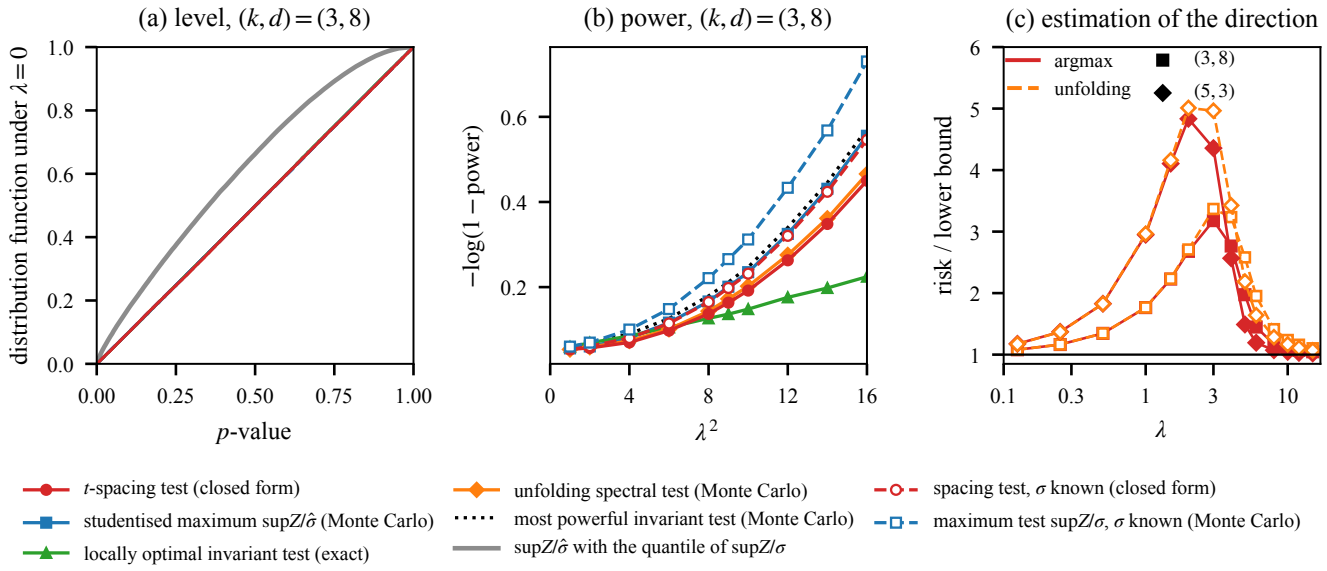

Figure 1: The t-spacing test against the competitors of Section 6, 10^5 replicates per point. (a) Under $\lambda = 0$, the distribution function of each p-value. Only $\sup Z/\hat\sigma$ with the quantile of $\sup Z/\sigma$ is not uniform: it rejects with probability 0.097 at the nominal level 0.05. (b) Power at level 0.05, on a logarithmic scale. The t-spacing test is below the simulated most powerful invariant test by at most 0.08. (c) Risk $\mathbb{E}\sin^2\angle(\hat e, e)$ of two estimators of the direction, divided by the lower bound of Theorem 4. The ratio tends to one at both ends. For the maximiser of the profile likelihood it is within 10% of one at $\lambda = 10$.

every dimension and every degree, and the second one does not use σ. They need no simulation, and the maximiser of Z that they compute is the estimator of e. The power of the spacing test was studied for least-angle regression by Azaïs, De Castro, and Mourareau (2018); for these tests in the tensor model nothing was known on the power. Figure 1 shows the picture.

Why the t-test, and not the known-variance one. In tensor principal component analysis the noise level is often unknown in practice (Xia, Zhang, and Zhou, 2022). It is tempting to regard the t-spacing test as the fallback one uses when σ is unavailable. In this model it is the other way round. Section 2.3 shows that the kernel $\langle s,t\rangle^k$ has finite rank $m = \binom{d+k-1}{k}$: the whole path of σX determines exactly m Gaussian coordinates and nothing more, so σ is not a function of the observation and the known-variance test requires information from outside the data. The t-spacing test needs no such information. It estimates the noise level with $\kappa = m - d$ degrees of freedom. In simulations its power is below that of the known-variance test, by an amount that decreases as κ grows (Sections 5 and 6).

Contributions. The test is invariant under rotations and the estimator is equivariant. So the power of the test and the risk of the estimator are the same at every direction e.

1. *Power.* At every λ, Corollary 4 bounds the probability that the t-spacing test misses a signal of amplitude λ by a noncentral Student distribution function. This probability decays exponentially in λ^2. We identify the rate E at every (k,d) and every level (Theorem 2). Its constant is a maximum over the spectra of the normalised Hessian, computed for $d = 3$ and bracketed for $d \geq 4$ in Remark 1. The lower bound on the rate follows from Corollary 4. As d grows at fixed k, the power tends to one when λ/σ is a large enough multiple of $\sqrt{d}$ (Theorem 3). The bound in closed form, Corollary 4, is informative only on the larger scale $\lambda/\sigma \asymp d^{k/2}$ (Remark 3).

2. *Minimax estimation.* The test also returns an estimator of the direction e, the argmax of the profile likelihood. At fixed (k,d) it is asymptotically minimax as $\lambda \to \infty$ for the squared sine of the angle to e, with the sharp constant $(d-1)/(k\lambda^2)$ (Theorem 4, Corollary 7). Theorem 4 is non-asymptotic:

a lower bound valid at every λ, and an upper bound with an explicit remainder which controls the probability that the argmax of a non-convex landscape with many critical points is far from e. The constant is the classical one: at fixed (k, d) the model is a curved exponential family in e with Fisher information $k\lambda^2 \,\mathrm{Id}$ on the tangent space, and the small-noise theory of the maximum likelihood estimator gives this constant for its limit law. The difficulty is not the local limit but the global landscape.

3. *Benchmarks.* With σ known, Corollary 5 bounds the probability of a miss for the spacing test, and gives the lower bound $k/(2(3k-2))$ on its rate. The rate E of the t-spacing test is below this bound at every (k, d) (Remark 2) and tends to it as $d \to \infty$ (Remark 3). With σ known, the test on the maximum of the profile likelihood has the rate $1/2$. At high signal the Bayes-optimal test, which averages the likelihood over directions, differs from it by a level shift of order $d \log \lambda/\lambda$ and has the same rate (Proposition 6, Remark 2). At vanishing signal the locally optimal test uses quadratic forms, not the t-test. At $(k, d) = (3, 3)$, with σ known, it is above the maximum test only where both have power below 0.19 at level 0.05. It gives no estimator of e (Section 5).

Growing dimension. In tensor principal component analysis the order k is fixed and the dimension d is large. The noise is then of order $\sqrt{d}$: $\sup X/\sqrt{d}$ tends almost surely to a constant, about 1.657 for $k = 3$ (Subag, 2017; Jagannath, Lopatto, and Miolane, 2020). On this scale, with $\lambda/\sigma = \bar{\lambda}\sqrt{d}$, the t-spacing test has power tending to one when $\bar{\lambda}$ is above a constant C_k, with $C_3 = 7.70$ (Theorem 3). For every odd k, $C_k \le 1.72 \cdot 2^{k/2}\sqrt{k}$, and $C_k/(2^{k/2}\sqrt{k}) \to 1$ as $k \to \infty$. So this constant grows exponentially with the order. It is a sufficient value of $\bar{\lambda}$, not a threshold, and for $k = 3$ it is far above the two values quoted below. We know of no other theorem on the power of spacing tests for a non-convex likelihood when the dimension grows. For $k = 3$ and a fixed $\bar{\lambda} < 1.12$, no test of level α has power larger than $\alpha + o(1)$ at every e (Perry, Wein, and Bandeira, 2020, Theorem 3.2). The maximum test, which rejects for large $\sup Z/\sigma$ with σ known, has power tending to one when $\bar{\lambda} > 1.21$ (Jagannath et al., 2020). We do not know the threshold of the t-spacing test (Remark 3).

Related work. Tensor methods recast several latent variable models as the low-rank decomposition of an empirical moment tensor of order $k \ge 3$, which removes the identifiability obstructions of the matrix case (Anandkumar, Ge, Hsu, Kakade, and Telgarsky, 2014); spiked tensor PCA, introduced by Richard and Montanari (2014), isolates the low-rank-plus-noise structure of that problem. The matrix analogue is the Baik–Ben Arous–Péché transition (Baik, Ben Arous, and Péché, 2005). Zhang and Xia (2018) give the minimax rates of low-rank tensor estimation, and their computational limits, as the dimension grows; Theorem 4 is at fixed (k, d) and gives the constant. On the Bayesian side, Lesieur, Miolane, Lelarge, Krzakala, and Zdeborová (2017) compute the asymptotic mutual information and the minimum mean square error for a spike with independent entries drawn from a known prior, at fixed signal-to-noise ratio as $d \to \infty$. Lesieur et al. (2017) and the works recalled in the previous paragraph normalise the noise differently from us; Appendix E of the Supplementary Material translates their thresholds into our variables. The critical points of the profile likelihood are counted in the spin-glass literature, by Auffinger, Ben Arous, and Černý (2013) for the pure noise field and by Ben Arous et al. (2019) for the spiked one. The noise field X is studied in Azaïs, Dalmao, and De Castro (2026), from which the present paper takes, restated without proof in the Supplementary Material, the layer decomposition of the helix process and a concentration bound for the norm of the Gaussian Orthogonal Ensemble. The high-dimensional dynamics of Langevin and online stochastic gradient descent for the multi-spiked model are characterised by Ben Arous, Gerbelot, and Piccolo (2024a,b); these address the computational side of the problem, complementary to the statistical guarantees here.

Notation. Throughout the paper $k \ge 3$ is odd and $d \ge 3$. $\mathbb{S}^{d-1}$ is the unit sphere of $\mathbb{R}^d$, $M = \mathbb{S}^{d-1}$ the parameter manifold, $\dim M = d - 1 =: N$. We write $\rho^2 = k/(2(k-1))$, φ and Φ for the standard Gaussian density and distribution function, $\lambda_{\max}$ and $\lambda_{\min}$ for the largest and smallest eigenvalues, $\|r\|_F$ and $\|r\|_{\mathrm{op}}$ for the Frobenius norm and the largest absolute eigenvalue of a symmetric matrix r, and $\mathrm{GOE}(n)$ for

the Gaussian Orthogonal Ensemble normalised by $\mathrm{Var}(v^\top G v) = 1$ for unit v. Multi-indices $\mathbf{a} \in \mathbb{N}^d$ with $|\mathbf{a}| = k$ index the monomials $t^{\mathbf{a}}$, and $\binom{k}{\mathbf{a}} = k!/\mathbf{a}!$. For a function on M, ∇ and ∇^2 denote its gradient and its Hessian on M.

2 The model, and why the noise level must be estimated

2.1 The spiked tensor and its profile likelihood

Write $\psi_t = t^{\otimes k}$, so that

$$\langle \psi_s, \psi_t \rangle = \langle s, t \rangle^k, \qquad \|\psi_t\| = 1,$$

and $Z(t) = \langle Y, \psi_t \rangle = \lambda \langle e, t \rangle^k + \sigma X(t)$. The model is the Gaussian shift $Y = \lambda \psi_e + \sigma W$ with a shift of norm exactly λ, so the signal-to-noise ratio is λ/σ.

2.2 The second maximum and the normalised Hessian

Take $\sigma = 1$ in this subsection; the general case follows by homogeneity, λ_1, λ_2 and Ω below being of degree one in (λ, σ). The test is built on three statistics of the profile likelihood: its maximum and the maximiser, which is the estimator of e; the second maximum, defined through a regressed field; and the normalised Hessian at the maximiser:

$$\lambda_1 = \max_{t \in M} Z(t) = Z(\hat{t}), \qquad \hat{t} = \arg\max_{t \in M} Z(t), \tag{1a}$$

$$Z^{|t}(s) = \frac{Z(s) - \langle s, t \rangle^k Z(t) - \langle s, t \rangle^{k-1} \langle s, \nabla Z(t) \rangle}{1 - \langle s, t \rangle^k}, \qquad s \neq t, \tag{1b}$$

$$\lambda_2 = \sup_{s \neq \hat{t}} Z^{|\hat{t}}(s), \tag{1c}$$

$$\Omega = \Omega_{\hat{t}}, \qquad \Omega_t = \frac{\nabla^2 Z(t) + k\, Z(t)\, \mathrm{Id}}{k} \quad \text{on the tangent space at } t. \tag{1d}$$

The maximiser is almost surely unique (Appendix A of the Supplementary Material). The second statistic looks at what remains of Z once the information carried by the maximiser is removed: for a fixed t, (1b) regresses $Z(s)$ on the value $Z(t)$ and the gradient $\nabla Z(t)$ and divides the residual by $1 - \langle s, t \rangle^k$. This is the *helix process* of Azaïs et al. (2024); under the null it is independent of $(Z(t), \nabla Z(t))$. The *second maximum* (1c) is its largest value at the maximiser. The *normalised Hessian* (1d) of Azaïs et al. (2024) is the Hessian of Z at the maximiser corrected by its value.

At a critical point t of Z with value $x = Z(t)$ and $c = \langle e, t \rangle$, the Hessian is $\nabla^2 Z(t) = \sqrt{2k(k-1)}\,(G_\lambda - \rho x\, \mathrm{Id})$ with $G \sim \mathrm{GOE}(N)$ and

$$G_\lambda = G + \beta_2(c)\, e_\perp e_\perp^\top, \qquad \beta_2(c) = \lambda \binom{k}{2}^{1/2} c^{k-2}(1 - c^2), \qquad e_\perp = \frac{e - ct}{\sqrt{1 - c^2}}, \tag{2a}$$

a rank-one spiked Gaussian Orthogonal Ensemble (Appendix B); $\beta_2(c)$ is the spike of the Hessian at t. The shift ρx cancels in Ω_t, and

$$\Omega_t = \frac{G_\lambda}{\rho}, \qquad \det\left(u\, \mathrm{Id} - \Omega_t\right) = \rho^{-N} \det\left(\rho u\, \mathrm{Id} - G_\lambda\right). \tag{2b}$$

The determinant in the spacing statistic is the determinant in the Kac–Rice integrand, the formula that counts the critical points of Z (Appendix A). Two order relations follow. At a local maximum $\nabla^2 Z \prec 0$, so $\lambda_{\max}(G_\lambda) < \rho x$, that is $\lambda_{\max}(\Omega) < \lambda_1$; and the limit of the helix as $s \to \hat{t}$ along a unit tangent direction

v is $v^\top \Omega v$, so $\lambda_2 \geq \lambda_{\max}(\Omega)$. Since $\nabla Z(\hat{t}) = 0$, (1b) gives $Z(s) - \lambda_1 = (1 - \langle s, \hat{t}\rangle^k)(Z^{|\hat{t}}(s) - \lambda_1)$, with $1 - \langle s, \hat{t}\rangle^k > 0$ for $s \neq \hat{t}$ as k is odd. As $Z(s) \leq \lambda_1$, this gives $\lambda_2 \leq \lambda_1$. Hence

$$\lambda_{\max}(\Omega) \;\leq\; \lambda_2 \;\leq\; \lambda_1. \tag{3}$$

For $t \in M$ let $M_t = \mathbb{S}^{d-1} \setminus \{\pm t\}$ and let ζ^t be the helix (1b) of the noise X. The second maximum at t is

$$\lambda_2^t = \sup_{M_t} Z^{|t}.$$

As $Z^{|t}$ is continuous at $-t$, it is also $\sup_{s \neq t} Z^{|t}(s)$, so $\lambda_2^{\hat{t}} = \lambda_2$. The helix process is made precise in Appendix B of the Supplementary Material, where statements and equations are numbered by appendix (Lemma B.1, equation (B.1a)).

2.3 The noise level is not determined by the observation

The kernel of X has finite rank. Write $f_{\mathbf{a}}(t) = \binom{k}{\mathbf{a}}^{1/2} t^{\mathbf{a}}$ for the multi-indices $|\mathbf{a}| = k$. The multinomial theorem gives

$$\langle s, t\rangle^k = \sum_{|\mathbf{a}|=k} f_{\mathbf{a}}(s) f_{\mathbf{a}}(t), \qquad \sum_{|\mathbf{a}|=k} f_{\mathbf{a}}(t)^2 = 1, \tag{4a}$$

$$X = \sum_{|\mathbf{a}|=k} \zeta_{\mathbf{a}} f_{\mathbf{a}}, \qquad (\zeta_{\mathbf{a}}) \text{ independent standard Gaussian variables}, \tag{4b}$$

and the $m = \binom{d+k-1}{k}$ functions $f_{\mathbf{a}}$ are linearly independent on $\mathbb{S}^{d-1}$, since a homogeneous polynomial of degree k that vanishes on the sphere vanishes everywhere, $p(x) = |x|^k p(x/|x|)$. So the observation $(\sigma X(t))_{t \in M}$ determines the m Gaussian variables $\sigma\zeta_{\mathbf{a}}$ and nothing more. We identify the symmetric tensors of order k, a Euclidean space of dimension m for $\langle \cdot, \cdot \rangle$, with $\mathbb{R}^m$ through an orthonormal basis; by (4b), W is a standard Gaussian vector of this space, and the model is the finite-dimensional regression

$$Y \sim N_m(\lambda \psi_e, \sigma^2 \mathrm{Id}_m), \tag{5}$$

whose mean lies on the cone $\{\lambda\psi_t : \lambda \geq 0,\ t \in \mathbb{S}^{d-1}\}$, of dimension d. The parameter σ is identifiable, the laws $N_m(0, \sigma^2 \mathrm{Id}_m)$ being distinct for distinct σ, but it is not a function of the observation: these laws are mutually absolutely continuous, so no measurable function of Y equals σ almost surely, and no estimator of σ is exact.

The estimator of Azaïs et al. (2024) is built on the regression of Y on $(Z(t), \nabla Z(t))$ at a fixed t. The regression spends d of the m coordinates, and the regressed field $Z^{|t}$ keeps the other κ coordinates $(\bar{\zeta}_i)_{i \leq \kappa}$ of its Karhunen–Loève expansion; the estimator is their mean square. It is independent in this sense: under the null, it is independent of the pair $(Z(t), \nabla Z(t))$ on which the regression is made. It also has a closed form in terms of Y.

Proposition 1 (The residual estimator of the noise level)**.** *Fix $t \in M$. Then*

$$\kappa = m - d, \tag{6a}$$

$$\hat{\sigma}_t^2 = \frac{1}{\kappa} \sum_{i \leq \kappa} \bar{\zeta}_i^2 \;\sim\; \sigma^2 \chi_\kappa^2/\kappa \quad \textit{under the null, independent of } (Z(t), \nabla Z(t)), \tag{6b}$$

$$\kappa\, \hat{\sigma}_t^2 = \|Y\|^2 - Z(t)^2 - \frac{1}{k} \|\nabla Z(t)\|^2 \quad \textit{under the null and under every alternative.} \tag{6c}$$

The proof is in Appendix D of the Supplementary Material: (6a) and (6b) are taken from Azaïs et al. (2024), and (6c) follows from the orthonormality of ψ_t and of the derivatives of $t \mapsto \psi_t$.

Since κ is an integer fixed by (k, d), the known-variance test is not a limiting case of the t-test *within* the model: the Student law tends to the Gaussian only as $m \to \infty$, that is across models. The two tests are two different procedures, and only the second is exact without external information.

2.4 The t-spacing test in closed form

Let $T_i = \lambda_i/\hat{\sigma}$ with $\hat{\sigma} = \hat{\sigma}_{\hat{\iota}}$, which is positive almost surely since $\kappa\hat{\sigma}^2 = \|Y\|^2 - \lambda_1^2 > 0$ by (6c), as $d < m$. For a symmetric matrix r,

$$H_r(\ell) = \int_\ell^\infty \det\left(u\,\mathrm{Id} - r\right) f_{m-1}\Big(u\sqrt{\tfrac{m-1}{\kappa}}\Big)\,\mathrm{d}u, \qquad \mathbf{T} = \frac{H_{\Omega/\hat{\sigma}}(T_1)}{H_{\Omega/\hat{\sigma}}(T_2)}, \tag{7}$$

f_ν being the Student density with ν degrees of freedom. By Azaïs et al. (2024, Theorem 3), $\mathbf{T}$ is uniform on $(0,1)$ under the null, whatever σ; the test of level α rejects when $\mathbf{T} \le \alpha$. With φ in place of the Student density, the known-variance statistic is

$$G_r(\ell) = \int_\ell^\infty \det\left(u\,\mathrm{Id} - r\right) \varphi(u)\,\mathrm{d}u, \qquad \mathbf{S} = \frac{G_{\Omega/\sigma}(\lambda_1/\sigma)}{G_{\Omega/\sigma}(\lambda_2/\sigma)}; \tag{8}$$

it is likewise uniform on $(0,1)$ under the null when σ is known (Azaïs et al., 2024, Theorem 2). Both G_r and H_r are finite sums of incomplete moments of the Gaussian and of the Student densities, computed in closed form in Appendix F of the Supplementary Material. Both are positive and strictly decreasing on $(\lambda_{\max}(r), \infty)$, with limit 0 at infinity, the integrand being positive there.

Recall that k is odd. For k even, $Z(-t) = Z(t)$ and $\langle \psi_s, \psi_t \rangle = 1$ at $s = -t$, so the assumptions of Azaïs et al. (2024) fail on $\mathbb{S}^{d-1}$. We do not treat this case.

The statistic $\mathbf{T}$ is invariant under $Y \mapsto aY$, $a > 0$, and $\mathbf{S}$ under $(Y, \sigma) \mapsto (aY, a\sigma)$, so the laws of both depend on (λ, σ) only through λ/σ; from Section 3 on we take $\sigma = 1$ and call λ the signal-to-noise ratio. Both statistics are also invariant under the orthogonal group. For $Q \in O(d)$, the field of $Q^{\otimes k}Y$ is $Z(Q^\top \cdot)$ and $\|Q^{\otimes k}Y\| = \|Y\|$. The two maxima, the spectrum of the normalised Hessian at the maximiser and the noise estimate do not change, so both statistics take the same value at Y and at $Q^{\otimes k}Y$. Also $Q^{\otimes k}Y = \lambda\psi_{Qe} + \sigma Q^{\otimes k}W$, and $Q^{\otimes k}W$ has the law of W, since $X(Q^\top \cdot)$ has the covariance $\langle s, t\rangle^k$ of X. So the law of each statistic is the same at e and at Qe. The group acts transitively on the sphere, so these laws do not depend on e. We write $\mathbb{P}_\lambda$ for the law of the observation at signal λ and any fixed direction e.

3 The rejection region and the power

The level $\alpha \in (0,1)$ is fixed. The rejection region $\{\mathbf{T} \le \alpha\}$ is sandwiched between two ratio tests on λ_1/λ_2 (Proposition 2), whose exponent κ is the number of degrees of freedom of $\hat{\sigma}$. On the event where the test fails, λ_1 is at most a constant times the noise estimate $\hat{\sigma}$ (Corollary 3). The first fact bounds the power from above (Theorem 1). The second bounds the deficiency $1 - \mathbb{P}_\lambda\{\mathbf{T} \le \alpha\}$ from above at every λ (Corollary 4), hence its exponent from below. One configuration of the normalised Hessian and of the noise estimate gives the matching bound, so the exponent is identified (Theorem 2). The last subsection lets the dimension grow (Theorem 3). All these statements are scale-free, as the statistic itself must be. For the known-variance statistic $\mathbf{S}$ the region lies between a shift of the squared level by c_α and a shift by $2\log(1/\alpha) - 2\log(\lambda_1/\lambda_2)$. Almost surely it contains $\{\lambda_1^2 \ge \lambda_2^2 + c_\alpha\}$, with c_α the $(1-\alpha)$-quantile of the χ^2_{2d} law (Corollary 5), and it is contained in $\{\lambda_1^2 \ge \lambda_2^2 + 2\log(1/\alpha) - 2\log(\lambda_1/\lambda_2)\}$ (Theorem 1). The difference between a dilation and a shift is the price of not knowing σ.

The reason is the tail of the reference density. In H_r the weight $\det(u\,\mathrm{Id} - r)$ has degree N and the Student density with $\nu = m-1$ degrees of freedom decays like $u^{-(\nu+1)}$, so the integrand decays like $u^{-\kappa-1}$ and H_r like $u^{-\kappa}$, since $\nu - N = (m-1) - (d-1) = \kappa$: multiplying the level by ϑ divides H_r by ϑ^κ. In G_r the Gaussian factor turns a dilation of the level into a shift of half its square. The next lemma makes both statements non-asymptotic. Write $\lambda_{\max}(r)^+ = \max(0, \lambda_{\max}(r))$ and $\lambda_{\min}(r)^- = \max(0, -\lambda_{\min}(r))$, and recall from (7) that the reference density in H_r is $f_\nu(u/\gamma)$ with the scale $\gamma = \sqrt{\kappa/\nu}$, so it is proportional to $(\kappa + u^2)^{-(\nu+1)/2}$, and $\gamma^2\nu = \kappa$ is exactly the number of degrees of freedom of $\hat{\sigma}$.

Lemma 1 (Dilation bounds). *Let r be a symmetric matrix of size N and let $\lambda_{\max}(r)^+ < \ell_2 \le \ell_1$; put $\vartheta = \ell_1/\ell_2$ and $g = \frac{1}{2}(\ell_1^2 - \ell_2^2)$. Then*

$$\vartheta^{-\kappa}\Big(1 + \frac{\lambda_{\min}(r)^-}{\ell_2}\Big)^{-N} \;\le\; \frac{H_r(\ell_1)}{H_r(\ell_2)} \;\le\; \vartheta^{N+1}\Big(1 - \frac{\lambda_{\max}(r)^+}{\ell_2}\Big)^{-N}\Big(\frac{\kappa+\ell_2^2}{\kappa+\ell_1^2}\Big)^{\frac{\nu+1}{2}}, \tag{9a}$$

$$\frac{\ell_2}{\ell_1}\,e^{-g} \;\le\; \frac{G_r(\ell_1)}{G_r(\ell_2)} \;\le\; e^{-g}\Big(1 + \frac{g}{\ell_2\big(\ell_2 - \lambda_{\max}(r)^+\big)}\Big)^{N}. \tag{9b}$$

The two bounds in (9a) differ by a factor tending to 1 as $\ell_2 \to \infty$ at fixed ϑ and fixed r, and likewise in (9b) at fixed g.

The proof is in Appendix D of the Supplementary Material.

Lemma 1 is applied below at $r = \Omega/\hat{\sigma}$ and $\ell_2 = T_2$, which requires $T_2 > \lambda_{\max}(\Omega/\hat{\sigma})^+$. This is a property of the field at the maximiser $\hat{t}$, a random point. The next lemma reduces such a property to the same property at each fixed point t, conditionally on $\nabla Z(t) = 0$, because $\hat{t}$ is a critical point of Z and the Kac–Rice formula of Appendix A counts them.

Lemma 2 (Transfer to the argmax). *Let $\lambda \ge 0$ and let $f : M \times \mathbb{R}^m \to \mathbb{R}$ be continuous; write $f_t = f(t, Y)$. For $t \in M$, let Π_t be the orthogonal projection of $\mathbb{R}^m$ onto the kernel of the linear map $Y \mapsto \nabla Z(t)$. By Gaussian regression, the law of Y given $\nabla Z(t) = 0$ is the law of $\Pi_t Y$, that is*

$$\mathbb{P}_\lambda\{f_t = 0 \mid \nabla Z(t) = 0\} \;=\; \mathbb{P}_\lambda\{f(t, \Pi_t Y) = 0\}.$$

If this probability is 0 for every $t \in M$, then $f_{\hat{t}} \neq 0$ almost surely under $\mathbb{P}_\lambda$.

Proof. The maximiser $\hat{t}$ is a critical point of Z. So $\mathbb{P}_\lambda\{f_{\hat{t}} = 0\}$ is at most the expected number of critical points t of Z with $f_t = 0$. By (A.1) for the closed condition $\{f_t = 0\}$, with $B = M$, and by monotone convergence as $u \to -\infty$, this expected number is

$$\int_M \mathbb{E}\Big[\big|\det \nabla^2 Z(t)\big|\,\mathbf{1}\{f_t = 0\} \;\Big|\; \nabla Z(t) = 0\Big]\, p_{\nabla Z(t)}(0)\,\mathrm{d}t \;=\; 0,$$

since $\mathbf{1}\{f_t = 0\}$ vanishes almost surely given $\nabla Z(t) = 0$. □

Lemma 3 (The second maximum is positive and exceeds the Hessian). *Let k be odd. For every Y and every $t \in M$, $Z(-t) = -Z(t)$ and $Z^{|t}(-t) = 0$. For every $\lambda \ge 0$, $\lambda_2 > \max\big(\lambda_{\max}(\Omega), 0\big)$ almost surely under $\mathbb{P}_\lambda$.*

Proof. As k is odd, $\psi_{-t} = -\psi_t$, so $Z(-t) = \langle Y, \psi_{-t}\rangle = -Z(t)$. At $s = -t$ in (1b), $\langle s, t\rangle^k = -1$ and $\langle s, \nabla Z(t)\rangle = 0$, as $\nabla Z(t)$ is tangent to M at t. So $Z^{|t}(-t) = \big(Z(-t) + Z(t)\big)/2 = 0$, and $\lambda_2^t \ge 0$ for every t.

Let $f_t = f(t, Y) = \lambda_2^t - \lambda_{\max}(\Omega_t)^+$. The function f is continuous on $M \times \mathbb{R}^m$ by Lemma B.3. By (3) and $\lambda_2 \ge 0$, $f_{\hat{t}} = \lambda_2 - \lambda_{\max}(\Omega)^+ \ge 0$. So the claim is that $f_{\hat{t}} \neq 0$ almost surely under $\mathbb{P}_\lambda$. The event $\{f_{\hat{t}} = 0\}$ is determined by Y, and the laws of Y under $\mathbb{P}_\lambda$ and under $\mathbb{P}_0$ are equivalent (Appendix A). So it suffices to take $\lambda = 0$.

Fix $t \in M$. Under $\mathbb{P}_0$, $Z = \sigma X$, so $Z^{|t} = \sigma\zeta^t$ and $\Omega_t = \sigma G/\rho$, with G the matrix (B.1b). By Lemma B.2(ii), ζ^t extends to an almost surely continuous field $\tilde{\zeta}$ on the blow-up $[0,\pi] \times \mathbb{S}^{d-2}$ of the sphere at $\pm t$, with $\tilde{\zeta}(0, v) = v^\top G v/\rho$ and $\tilde{\zeta}(\pi, \cdot) \equiv 0$. So the maxima of $\sigma\tilde{\zeta}$ on the two pole fibres $\theta = 0$ and $\theta = \pi$ are $\lambda_{\max}(\Omega_t)$ and 0. The blow-up is compact, so the supremum of $\tilde{\zeta}$ is attained. By Azaïs et al. (2026, Proposition 4.1 of arXiv:2611.09459v1), almost surely it is not attained on a pole fibre. It is then attained at a point of M_t, where $\sigma\tilde{\zeta} = Z^{|t}$. So λ_2^t is the supremum of $\sigma\tilde{\zeta}$ and it exceeds the maxima on the two pole fibres. Hence $\mathbb{P}_0\{f_t = 0\} = 0$. By Lemma 2 of Azaïs et al. (2024), under $\mathbb{P}_0$ the helix ζ^t and the matrix G are independent of $(X(t), \nabla X(t))$, as recalled in Appendix B at (B.1b) and after

the layer decomposition. All these variables are jointly Gaussian, so the pair (ζ^t, G) is independent of $(X(t), \nabla X(t))$, and the regression on $\nabla Z(t)$ does not change its law. As f_t is a function of (ζ^t, G) and $\nabla Z(t) = \sigma \nabla X(t)$, $\mathbb{P}_0\{f_t = 0 \mid \nabla Z(t) = 0\} = \mathbb{P}_0\{f_t = 0\} = 0$ for every t. Lemma 2 above, with $\lambda = 0$, gives $f_{\hat{t}} \neq 0$ almost surely under $\mathbb{P}_0$. □

Proposition 2 (The rejection region is sandwiched between two ratio tests)**.** *Almost surely, with* $\vartheta = \lambda_1/\lambda_2 \geq 1$,

$$\vartheta^{-\kappa}\Big(1 + \frac{\lambda_{\min}(\Omega)^-}{\lambda_2}\Big)^{-N} \;\leq\; \mathbf{T} \;\leq\; \vartheta^{-\kappa}\Big(1 - \frac{\lambda_{\max}(\Omega)^+}{\lambda_2}\Big)^{-N}\Big(1 + \frac{\kappa\hat{\sigma}^2}{\lambda_2^2}\Big)^{\frac{\nu+1}{2}}. \tag{10}$$

The left-hand side involves neither $\hat{\sigma}$ *nor the scale of the observation; the right-hand side involves* $\hat{\sigma}$ *only through* $\kappa\hat{\sigma}^2/\lambda_2^2$.

Proof. Apply Lemma 1 with $r = \Omega/\hat{\sigma}$ and $\ell_i = T_i = \lambda_i/\hat{\sigma}$. Its hypothesis $\lambda_{\max}(r)^+ < \ell_2$ holds almost surely: by Lemma 3, $T_2 > 0$ and $\lambda_{\max}(\Omega/\hat{\sigma}) < T_2$, as $\hat{\sigma} > 0$. All three quantities ℓ_1/ℓ_2, $\lambda_{\min}(r)^-/\ell_2$ and $\lambda_{\max}(r)^+/\ell_2$ are unchanged when r and the ℓ_i are divided by $\hat{\sigma}$, which gives the $\hat{\sigma}$-free form of ϑ and of the two Hessian corrections. For the last factor,

$$\Big(\frac{\kappa + \ell_2^2}{\kappa + \ell_1^2}\Big)^{\frac{\nu+1}{2}} \leq \Big(\frac{\kappa + \ell_2^2}{\ell_1^2}\Big)^{\frac{\nu+1}{2}} = \vartheta^{-(\nu+1)}\Big(1 + \frac{\kappa}{\ell_2^2}\Big)^{\frac{\nu+1}{2}},$$

and $\vartheta^{N+1}\vartheta^{-(\nu+1)} = \vartheta^{-\kappa}$, while $\kappa/\ell_2^2 = \kappa\hat{\sigma}^2/\lambda_2^2$. □

3.1 An upper bound on the power

The left-hand inequality of (10) costs nothing and yields an upper bound on the power.

Theorem 1 (Upper bound on the power)**.** *Let* $\lambda \geq 0$, $\sigma > 0$ *and* $\alpha \in (0, 1)$. *Then*

$$\mathbb{P}_\lambda\{\mathbf{T} \leq \alpha\} \;\leq\; \mathbb{P}_\lambda\Big\{\lambda_2 \;\leq\; \alpha^{1/\kappa}\,\lambda_1\Big(1 + \frac{\lambda_{\min}(\Omega)^-}{\lambda_2}\Big)^{N/\kappa}\Big\}, \tag{11a}$$

equivalently

$$1 - \mathbb{P}_\lambda\{\mathbf{T} \leq \alpha\} \;\geq\; \mathbb{P}_\lambda\Big\{\lambda_2 \;>\; \alpha^{1/\kappa}\,\lambda_1\Big(1 + \frac{\lambda_{\min}(\Omega)^-}{\lambda_2}\Big)^{N/\kappa}\Big\}. \tag{11b}$$

For the known-variance test, $\mathbb{P}_\lambda\{\mathbf{S} \leq \alpha\} \leq \mathbb{P}_\lambda\{(\lambda_1^2 - \lambda_2^2)/\sigma^2 \geq 2\log(1/\alpha) - 2\log(\lambda_1/\lambda_2)\}$. *Both bounds hold under the null as well, where the left-hand side is exactly* α.

Proof. On $\{\mathbf{T} \leq \alpha\}$ the left-hand inequality of (10) gives $\vartheta^{-\kappa}(1 + \lambda_{\min}(\Omega)^-/\lambda_2)^{-N} \leq \alpha$; raising to the power $1/\kappa$ and multiplying by $\lambda_1(1 + \lambda_{\min}(\Omega)^-/\lambda_2)^{N/\kappa}$, which turns $\vartheta^{-1}\lambda_1$ into λ_2, gives the event on the right of (11a), and the complement is the second display. For $\mathbf{S}$, use the left-hand inequality of (9b) in the same way. □

Under the null the bound is not vacuous and can be read as a statement about the null law of the triple: since $\mathbb{P}_0\{\mathbf{T} \leq \alpha\} = \alpha$ exactly, (11a) says that $\lambda_1/\lambda_2 \geq \alpha^{-1/\kappa}\big(1 + \lambda_{\min}(\Omega)^-/\lambda_2\big)^{-N/\kappa}$ with probability at least α. The correction lowers the threshold, so it does not by itself bound $\mathbb{P}_0\{\lambda_1/\lambda_2 \geq \alpha^{-1/\kappa}\}$. At $(k, d) = (3, 3)$ and $\alpha = 0.05$ one has $\kappa = 7$ and $\alpha^{-1/\kappa} = 1.534$; at $(k, d) = (5, 5)$, $\kappa = 121$ and $\alpha^{-1/\kappa} = 1.025$. The exponent $1/\kappa$ is what makes the t-version demanding when m is small: with few degrees of freedom the test must see a large *ratio* of the two maxima, not merely a large gap.

3.2 A lower bound on the power

A lower bound on the power needs an upper bound on $\mathbf{T}$. It comes from Lemmas 5 and 6, through Corollary 3 below. They rest on a decomposition of $P_t^{\perp}Y$ into layers, one for each power of the tangent direction.

Lemma 4 (Layers at a point)**.** *Fix $t \in M$ and identify the tangent space at t with $t^{\perp}$. For a unit vector $v \perp t$ and $0 \le j \le k$ let $\Psi_j(v) = \binom{k}{j}^{1/2} \operatorname{Sym}\big(t^{\otimes(k-j)} \otimes v^{\otimes j}\big)$, where* Sym *averages over the permutations of the k factors, let $\mathcal{V}_j$ be the span of the $\Psi_j(v)$, and let $w_j(\theta) = \binom{k}{j}^{1/2} \cos^{k-j}\theta \, \sin^j\theta/(1-\cos^k\theta)$, the weights of Appendix B. Then $\langle \Psi_j(v), \Psi_j(v')\rangle = \langle v, v'\rangle^j$, $\mathcal{V}_j \perp \mathcal{V}_{j'}$ for $j \ne j'$, the range of $P_t^{\perp}$ is $\mathcal{V}_2 \oplus \cdots \oplus \mathcal{V}_k$, and the map $vv^{\top} \mapsto \Psi_2(v)$ extends to a linear isometry ι from the symmetric matrices on $t^{\perp}$, with the Frobenius norm, onto $\mathcal{V}_2$. Let $P_{\mathcal{V}_j}$ be the orthogonal projection onto $\mathcal{V}_j$. For every Y and every $s = \cos\theta\, t + \sin\theta\, v \in M_t$,*

$$Z^{|t}(s) = \sum_{j=2}^{k} w_j(\theta)\, \langle Y, \Psi_j(v)\rangle, \qquad v^{\top}\Omega_t v = \rho^{-1}\langle Y, \Psi_2(v)\rangle,$$
$$\kappa\hat{\sigma}_t^2 = \rho^2\|\Omega_t\|_F^2 + \textstyle\sum_{j=3}^{k} \|P_{\mathcal{V}_j}Y\|^2.$$

The proof is in Appendix D of the Supplementary Material.

Lemma 5 (The second maximum and the spectrum of the Hessian)**.** *For every Y and every t, $\rho^2(\lambda_2^t)^2 + \rho^2\big(\|\Omega_t\|_F^2 - \|\Omega_t\|_{\mathrm{op}}^2\big) \le \kappa\hat{\sigma}_t^2$. In particular, with $T_{\max} = \sqrt{\kappa}/\rho$ and $r = \Omega/\hat{\sigma}$, almost surely $T_2^2 + \|r\|_F^2 - \|r\|_{\mathrm{op}}^2 \le T_{\max}^2$ and $T_2 \le T_{\max}$.*

Proof. Let $\Delta_t^2 = \kappa\hat{\sigma}_t^2 - \rho^2\|\Omega_t\|_F^2 = \sum_{j\ge 3}\|P_{\mathcal{V}_j}Y\|^2$. By Lemma 4, $|\langle Y, \Psi_2(v)\rangle| \le \rho\|\Omega_t\|_{\mathrm{op}}$, and $|\langle Y, \Psi_j(v)\rangle| \le \|P_{\mathcal{V}_j}Y\|$ for $j \ge 3$, as $\|\Psi_j(v)\| = 1$. The Cauchy–Schwarz inequality, used twice, gives for $s \in M_t$

$$|Z^{|t}(s)| \;\le\; |w_2(\theta)|\,\rho\|\Omega_t\|_{\mathrm{op}} + \Big(\sum_{j\ge 3} w_j(\theta)^2\Big)^{1/2}\Delta_t \;\le\; V(\theta)^{1/2}\big(\rho^2\|\Omega_t\|_{\mathrm{op}}^2 + \Delta_t^2\big)^{1/2},$$

with $V = \sum_{j\ge 2} w_j^2 \le \rho^{-2}$ by Lemma B.1. Take the supremum over M_t and square. At $t = \hat{t}$ divide by $\rho^2\hat{\sigma}^2$, which is positive almost surely, and use $\|r\|_F \ge \|r\|_{\mathrm{op}}$. □

The bound $T_{\max}$ on T_2 is 3.06 at $(k, d) = (3, 3)$, 6.32 at $(3, 5)$ and 13.9 at $(5, 5)$. On the simulations of Section 6 it is never exceeded, and it is approached on the replicates where the test fails. With the next lemma, Lemma 5 describes the deficiency event without reference to λ_2 or to the Hessian.

Lemma 6 (Monotonicity in the spectrum)**.** *Let r be a symmetric matrix of size N and $\lambda_{\max}(r) \le \ell_2 \le \ell_1$. The ratio $H_r(\ell_1)/H_r(\ell_2)$ is nondecreasing in ℓ_2, and in each eigenvalue of r on $(-\infty, \ell_2]$, the other eigenvalues being fixed.*

The proof is in Appendix D of the Supplementary Material.

For a symmetric matrix r of size N and $v \ge \lambda_{\max}(r)$, the function H_r is positive and continuous on $[v, \infty)$, strictly decreasing, with limit 0. So there is a unique $u_\alpha(r, v) > v$ with

$$H_r\big(u_\alpha(r, v)\big) = \alpha\, H_r(v), \tag{12a}$$

and, when $T_2 = v$ and $\Omega/\hat{\sigma} = r$, the test rejects if and only if $T_1 \ge u_\alpha(r, v)$. Let

$$A = \frac{1}{T_{\max}} \max\Big\{u_\alpha\big(r, \lambda_{\max}(r)\big) :\; r \succeq 0,\; \|r\|_F = T_{\max}\Big\}. \tag{12b}$$

The maximum exists: $H_r(u) - \alpha H_r(\lambda_{\max}(r))$ is continuous in (r, u) and strictly decreasing in u, so $u_\alpha(r, \lambda_{\max}(r))$ is a continuous function of the eigenvalues of r, which range over a compact set. A matrix of rank one in this set has $\lambda_{\max}(r) = T_{\max}$, so $A > 1$. The values of A are in Remark 1.

Corollary 3 (The deficiency event)**.** *Let A be the constant* (12b)*. Almost surely,*

$$\{\mathbf{T} > \alpha\} \;\subseteq\; \Big\{\lambda_1 < \frac{A\sqrt{\kappa}}{\rho}\,\hat{\sigma}\Big\}.$$

Proof. Almost surely the maximiser is unique and $\hat{\sigma} > 0$. Take such an observation with $\mathbf{T} > \alpha$, let $r = \Omega/\hat{\sigma}$ have the eigenvalues $r_1 \geq \cdots \geq r_N$, and let $r_i^+ = \max(r_i, 0)$. The sum $\|r\|_F^2 - \|r\|_{\mathrm{op}}^2$ contains all the r_i^2 but one of largest modulus, so it is at least $\sum_{i\geq 2}(r_i^+)^2$. Lemma 5 gives

$$T_2 \;\leq\; v' := \Big(T_{\max}^2 - \sum_{i\geq 2}(r_i^+)^2\Big)^{1/2},$$

and $r_i^+ \leq T_2 \leq T_1$ by (3) and Lemma 3. Let r'' have the eigenvalues $v', r_2^+, \ldots, r_N^+$: it is positive semidefinite, $\|r''\|_F = T_{\max}$ and $\lambda_{\max}(r'') = v'$. If $T_1 \leq v'$, then $T_1 < u_\alpha(r'', v')$. Otherwise Lemma 6, applied to the negative eigenvalues raised to 0, then to ℓ_2 raised from T_2 to v', then to r_1^+ raised to v', gives $\alpha < \mathbf{T} \leq H_{r''}(T_1)/H_{r''}(v')$, hence $T_1 < u_\alpha(r'', v')$, as $H_{r''}$ decreases. In both cases $T_1 < u_\alpha(r'', \lambda_{\max}(r'')) \leq AT_{\max}$, that is $\lambda_1 < AT_{\max}\hat{\sigma}$. □

Corollary 4 (A power bound in closed form)**.** *Let $\lambda \geq 0$, let A be the constant of Corollary 3, and let $\xi \sim N(0,1)$ and χ_{m-1} be independent, where χ_{m-1} is the norm of a standard Gaussian vector of $\mathbb{R}^{m-1}$. Then*

$$1 - \mathbb{P}_\lambda\{\mathbf{T} \leq \alpha\} \;\leq\; \mathbb{P}\Big\{\lambda + \xi < \frac{A}{\rho}\,\chi_{m-1}\Big\}, \tag{13a}$$

and, if $\rho\lambda \geq A\sqrt{m-1}$,

$$1 - \mathbb{P}_\lambda\{\mathbf{T} \leq \alpha\} \;\leq\; \exp\Big(-\frac{\big(\rho\lambda - A\sqrt{m-1}\big)^2}{2(A^2+\rho^2)}\Big). \tag{13b}$$

Proof of Corollary 4. In the coordinates of (5), with $\sigma = 1$, $Y = \lambda\psi_e + W$ and W is standard Gaussian in $\mathbb{R}^m$. The vectors collinear with some ψ_t form a set of dimension $d < m$, so almost surely Y is not one of them, and $|Z(t)| < \|Y\|$ for every t by the Cauchy–Schwarz inequality. As $\nabla Z(\hat{t}) = 0$, (6c) at $t = \hat{t}$ gives $\kappa\hat{\sigma}^2 = \|Y\|^2 - \lambda_1^2$, and Corollary 3 gives $\{\mathbf{T} > \alpha\} \subseteq \{\rho\lambda_1 < A(\|Y\|^2 - \lambda_1^2)^{1/2}\}$. Let $\xi = \langle W, \psi_e\rangle$ and $\chi = \|W - \xi\psi_e\|$. Then $Z(e) = \lambda + \xi$ and $Y - Z(e)\psi_e = W - \xi\psi_e$, so $\|Y\|^2 = Z(e)^2 + \chi^2$. The map $x \mapsto x(\|Y\|^2 - x^2)^{-1/2}$ is increasing on $(-\|Y\|, \|Y\|)$, and $Z(e) \leq \lambda_1$. Hence, on $\{\mathbf{T} > \alpha\}$, $\rho Z(e) < A(\|Y\|^2 - Z(e)^2)^{1/2} = A\chi$, that is $\rho(\lambda + \xi) < A\chi$, whatever the sign of $\lambda + \xi$. The variables ξ and $W - \xi\psi_e$ are the projections of W onto ψ_e and onto its orthogonal complement. So $\xi \sim N(0,1)$ is independent of χ, and χ has the law of χ_{m-1}. This proves (13a). For (13b), $A\chi - \rho\xi$ is the supremum of $\langle Au - \rho\psi_e, W\rangle$ over the unit vectors u orthogonal to ψ_e. This is a centred Gaussian process with variance $A^2 + \rho^2$, and the mean of its supremum is $A\,\mathbb{E}\chi \leq A(\mathbb{E}\chi^2)^{1/2} = A\sqrt{m-1}$. The concentration inequality for the supremum of a Gaussian process (Azaïs and Wschebor, 2009) gives $\mathbb{P}\{A\chi - \rho\xi > \rho\lambda\} \leq \exp\big(-(\rho\lambda - A\sqrt{m-1})^2/(2(A^2+\rho^2))\big)$ when $\rho\lambda \geq A\sqrt{m-1}$. □

The right-hand side of (13a) is the distribution function of the noncentral Student law with $m-1$ degrees of freedom and noncentrality λ, at the point $A\sqrt{m-1}/\rho$. At $\alpha = 0.05$ it is below 0.05 from about $\lambda = 11.0$ at $(k, d) = (3, 5)$, and from $\lambda = 11.4$ at $(3, 3)$.

The same route gives a power bound for the known-variance statistic $\mathbf{S}$, once the upper bound of (9b) is replaced by a bound free of the spectrum of r.

Lemma 7 (A shift bound)**.** *Let r be a symmetric matrix of size N, let $\lambda_{\max}(r) \leq \ell_2$ and $0 \leq \ell_2 \leq \ell_1$, and put $g = \frac{1}{2}(\ell_1^2 - \ell_2^2)$. Then*

$$\frac{G_r(\ell_1)}{G_r(\ell_2)} \;\leq\; e^{-g}\sum_{j=0}^{N}\frac{g^j}{j!}.$$

Proof. For $g = 0$ both sides are 1; let $g > 0$. Write $D(u) = \det(u\,\mathrm{Id} - r) = \prod_i (u - r_i)$, the r_i being the eigenvalues of r. It is positive for $u > \ell_2$, since $r_i \le \ell_2$. As $\ell_2 \ge 0$, the substitution $u = (\ell_2^2 + 2s)^{1/2}$ maps $s \in (0,\infty)$ onto $u \in (\ell_2, \infty)$ and $s = g$ to $u = \ell_1$, with $\mathrm{d}u = \mathrm{d}s/u$ and $\varphi(u) = e^{-s}\varphi(\ell_2)$. Let $g_N(s) = s^N e^{-s}/N!$. Then, for $h \in \{0, g\}$,

$$G_r\big((\ell_2^2 + 2h)^{1/2}\big) = N!\,\varphi(\ell_2) \int_h^\infty g_N(s)\, q(s)\, \mathrm{d}s, \qquad q(s) = \frac{1}{u} \prod_i \frac{u - r_i}{s}. \tag{14a}$$

The function q is positive and nonincreasing: $1/u$ decreases, and s increases with u, with

$$\frac{\mathrm{d}}{\mathrm{d}u} \log \frac{u - r_i}{s} = \frac{1}{u - r_i} - \frac{2u}{(u - \ell_2)(u + \ell_2)} \;\le\; \frac{1}{u - \ell_2} - \frac{1}{u - \ell_2} = 0, \tag{14b}$$

because $r_i \le \ell_2$ and $2u \ge u + \ell_2 > 0$. Put $a = \int_g^\infty g_N q$, $b = \int_0^g g_N q$ and $\bar\Gamma = \int_g^\infty g_N$. As q is nonincreasing, $a \le q(g)\bar\Gamma$ and $b \ge q(g)(1 - \bar\Gamma)$. The ratio $a/(a+b)$ increases in a and decreases in b, so

$$\frac{G_r(\ell_1)}{G_r(\ell_2)} = \frac{a}{a+b} \;\le\; \frac{q(g)\bar\Gamma}{q(g)\bar\Gamma + q(g)(1 - \bar\Gamma)} = \bar\Gamma, \tag{14c}$$

as the factor $q(g) > 0$ cancels. Integration by parts gives $\bar\Gamma = e^{-g} \sum_{j \le N} g^j/j!$. □

Corollary 5 (A power bound for the known-variance test)**.** *Let $\sigma = 1$ and $\lambda \ge 0$, let c_α be the $(1-\alpha)$-quantile of the χ^2_{2d} law, and let $\xi \sim N(0,1)$ and χ_{m-1} be independent, as in Corollary 4. Then, almost surely, $\{\mathbf{S} > \alpha\} \subseteq \{\lambda_1^2 < \lambda_2^2 + c_\alpha\}$, and*

$$1 - \mathbb{P}_\lambda\{\mathbf{S} \le \alpha\} \;\le\; \mathbb{P}\big\{\rho^2(\lambda + \xi)^2 < \chi^2_{m-1} + \rho^2 c_\alpha\big\}, \tag{15a}$$

and, if $\rho\lambda \ge \sqrt{m-1} + \rho\sqrt{c_\alpha}$,

$$1 - \mathbb{P}_\lambda\{\mathbf{S} \le \alpha\} \;\le\; \exp\Big(-\frac{\big(\rho\lambda - \sqrt{m-1} - \rho\sqrt{c_\alpha}\big)^2}{2(1+\rho^2)}\Big). \tag{15b}$$

Proof. The inclusion. With $\sigma = 1$, $\mathbf{S} = G_\Omega(\lambda_1)/G_\Omega(\lambda_2)$. By (3), $\lambda_{\max}(\Omega) \le \lambda_2 \le \lambda_1$, and $\lambda_2 \ge 0$ almost surely by Lemma 3. Lemma 7 with $r = \Omega$ and $\ell_i = \lambda_i$ gives $\mathbf{S} \le e^{-g} \sum_{j \le N} g^j/j!$ with $g = \frac{1}{2}(\lambda_1^2 - \lambda_2^2)$. As $N + 1 = d$, the density g_N of its proof is the density of half a χ^2_{2d} variable, so the right-hand side is $\mathbb{P}\{\chi^2_{2d} > 2g\}$. It is strictly decreasing in g and equals α at $2g = c_\alpha$. So $\mathbf{S} > \alpha$ forces $\lambda_1^2 - \lambda_2^2 < c_\alpha$.

The bound. Lemma 3 and Lemma 5 at $t = \hat t$ give $0 \le \rho\lambda_2 \le \sqrt{\kappa}\,\hat\sigma$, as $\|\Omega\|_F \ge \|\Omega\|_{\mathrm{op}}$, and (6c) at $\hat t$, where $\nabla Z(\hat t) = 0$, gives $\kappa\hat\sigma^2 = \|Y\|^2 - \lambda_1^2$. With the inclusion, on $\{\mathbf{S} > \alpha\}$,

$$\rho^2\lambda_1^2 \;<\; \rho^2\lambda_2^2 + \rho^2 c_\alpha \;\le\; \|Y\|^2 - \lambda_1^2 + \rho^2 c_\alpha.$$

Let ξ and χ be as in the proof of Corollary 4, so that $Z(e) = \lambda + \xi$, $\|Y\|^2 = Z(e)^2 + \chi^2$, $\xi \sim N(0,1)$ is independent of χ, and χ has the law of χ_{m-1}. Also $\lambda_1 \ge Z(e)$ and $\lambda_1 \ge Z(-e) = -Z(e)$ by Lemma 3, so $Z(e)^2 \le \lambda_1^2$. Hence, on $\{\mathbf{S} > \alpha\}$, $(1+\rho^2)Z(e)^2 < Z(e)^2 + \chi^2 + \rho^2 c_\alpha$, that is $\rho^2(\lambda+\xi)^2 < \chi^2 + \rho^2 c_\alpha$. This proves (15a). On this event $\rho(\lambda + \xi) \le \rho|\lambda + \xi| < \chi + \rho\sqrt{c_\alpha}$, that is $\chi - \rho\xi > \rho\lambda - \rho\sqrt{c_\alpha}$. The variable $\chi - \rho\xi$ is the supremum of $\langle u - \rho\psi_e, W\rangle$ over the unit vectors u orthogonal to ψ_e. This is a centred Gaussian process with variance $1 + \rho^2$, and the mean of its supremum is $\mathbb{E}\chi \le \sqrt{m-1}$. The concentration inequality used for Corollary 4 gives (15b) when $\rho\lambda - \rho\sqrt{c_\alpha} \ge \sqrt{m-1}$. □

At $\alpha = 0.05$, $c_\alpha = 12.6$ for $d = 3$ and 18.3 for $d = 5$. The bound (15a) is below 0.05 from about $\lambda = 6.8$ at $(k, d) = (3, 3)$ and from $\lambda = 10.0$ at $(3, 5)$, against 11.4 and 11.0 for $\mathbf{T}$ in Corollary 4. By (15b), $\liminf_{\lambda\to\infty} -\lambda^{-2} \log\big(1 - \mathbb{P}_\lambda\{\mathbf{S} \le \alpha\}\big) \ge \rho^2/(2(1+\rho^2)) = k/(2(3k-2))$, the value in the column $\mathbf{S}$ of Remark 1; that the exponent of $\mathbf{S}$ is not larger is not proved.

3.3 The deficiency exponent

Corollary 4 bounds the exponential rate of the deficiency from below. A configuration of the normalised Hessian and of the noise estimate at which the test fails bounds it from above.

Theorem 2 (Deficiency exponent)**.** *Let $\alpha \in (0,1)$ and let A be the constant (12b). Then*

$$\lim_{\lambda\to\infty} \; -\frac{1}{\lambda^2}\log\big(1-\mathbb{P}_\lambda\{\mathbf{T}\le\alpha\}\big) \;=\; E := \frac{\rho^2}{2(A^2+\rho^2)}.$$

Proof. For $\rho\lambda \ge A\sqrt{m-1}$, (13b) gives

$$-\frac{1}{\lambda^2}\log\big(1-\mathbb{P}_\lambda\{\mathbf{T}\le\alpha\}\big) \;\ge\; \frac{\big(\rho - A\sqrt{m-1}/\lambda\big)^2}{2(A^2+\rho^2)},$$

and the right-hand side tends to E as $\lambda\to\infty$. So the lim inf is at least E. The matching bound, lim sup $\le E$, is proved in Appendix D of the Supplementary Material, through one configuration of the normalised Hessian and of the noise estimate at which the test fails, and an open cone around it. □

Remark 1 (The constant A and the exponent)**.** The decimal values of A below are computed by a maximisation over the spectra of r. For $d=3$ the spectrum is (a,b) with $a^2+b^2 = T_{\max}^2$, and the maximum is in one variable; at $(3,3)$ it is attained at $(a,b) = (0.995, 0.103)\,T_{\max}$. For $d\ge 4$ the column A gives the maximum over the spectra $(a,b,\dots,b)$, which is at most A. When the top eigenvalue is a, the others are at most $\min\{a, (T_{\max}^2 - a^2)^{1/2}\}$, so Lemma 6 bounds A from above by a maximum in one variable. This upper bound exceeds the column by $1.3\cdot 10^{-3}$ at $(3,4)$, $4.6\cdot 10^{-4}$ at $(3,5)$, $1.9\cdot 10^{-6}$ at $(5,5)$ and $3.6\cdot 10^{-8}$ at $(7,5)$, and the column E keeps the digits that both values give. A search over all spectra from random starting points returns the value of the column. At $\alpha = 0.05$, with the lower bound $k/(2(3k-2))$ on the exponent of the known-variance test $\mathbf{S}$ from Corollary 5:

(k,d)	κ	A	E	$\mathbf{S}$, lower bound
$(3,3)$	7	2.312	0.0615	0.2143
$(3,4)$	16	1.543	0.120	0.2143
$(3,5)$	30	1.287	0.156	0.2143
$(5,3)$	18	1.440	0.1158	0.1923
$(5,5)$	121	1.064	0.1778	0.1923
$(7,5)$	325	1.023	0.1789	0.1842

The slopes measured in Section 6 lie between 10% below and 7% above E.

3.4 Benchmarks at high signal

The benchmarks at high signal are the maximum test and the Bayes-optimal test. Under the uniform prior on e, which is least favourable by $O(d)$-invariance, the Bayes-optimal test statistic is the mixture likelihood ratio

$$\bar{L}_\lambda(Y) = e^{-\lambda^2/2}\int_{\mathbb{S}^{d-1}} e^{\lambda Z(e')}\,\mathrm{d}e',$$

with $\mathrm{d}e'$ the uniform probability measure on the sphere ($|B|$ denotes the measure of $B\subset\mathbb{S}^{d-1}$), a Laplace transform of the profile likelihood at inverse temperature λ. Since $\frac{1}{\lambda}\log\int e^{\lambda Z}\mathrm{d}e' \to \sup_M Z$ as $\lambda\to\infty$, the maximum test is the zero-temperature limit of the Bayes-optimal test. It is not available here: its threshold needs the noise level, which the observation does not determine (Section 2.3), and the null law of $\sup Z$, which has no closed form. The following estimate shows that the two tests differ by a level shift of order $d\log\lambda/\lambda$.

Proposition 6 (Zero-temperature estimate)**.** *Let* $\mathcal{L} = \frac{1}{\lambda}\log\int_{\mathbb{S}^{d-1}} e^{\lambda Z}\mathrm{d}e'$ *and* $K = \sup_M \|\nabla^2 Z\|_{\mathrm{op}}$*. Then, on the event* $\{\lambda K \geq 4(d-1)/\pi^2\}$*,*

$$0 \;\leq\; \sup_M Z - \mathcal{L} \;\leq\; \frac{d-1}{2\lambda}\Big(1 + \log\frac{K\lambda}{d-1}\Big) + \frac{(d-2)\log(\pi/2) + \log\big(\pi(d-1)\big)}{\lambda}.$$

Proof. The upper bound $\mathcal{L} \leq \sup Z$ is immediate. For the lower one, let $\mathcal{C}_r$ be the geodesic cap of radius $r \leq \pi/2$ around $\hat{t}$. Since $\nabla Z(\hat{t}) = 0$ and $\|\nabla^2 Z\|_{\mathrm{op}} \leq K$ on M, Taylor's inequality along the geodesic gives $Z \geq \sup Z - \frac{1}{2}Kr^2$ on $\mathcal{C}_r$, so $\int e^{\lambda Z}\mathrm{d}e' \geq |\mathcal{C}_r|\, e^{\lambda(\sup Z - Kr^2/2)}$ and

$$\sup Z - \mathcal{L} \;\leq\; \frac{Kr^2}{2} + \frac{\log\big(1/|\mathcal{C}_r|\big)}{\lambda}.$$

From $\sin\theta \geq 2\theta/\pi$ on $[0,\pi/2]$ and $\int_0^\pi \sin^{d-2} \leq \pi$ one gets $|\mathcal{C}_r| \geq (2/\pi)^{d-2} r^{d-1}/(\pi(d-1))$, so the second term is at most $\big[(d-1)\log(1/r) + (d-2)\log(\pi/2) + \log(\pi(d-1))\big]/\lambda$. Minimising over r gives $r^2 = (d-1)/(K\lambda)$, at most $\pi^2/4$ on this event, and the stated bound. □

Remark 2 (What exactness costs)**.** Take $\sigma = 1$ known. An oracle that knew e would test $N(0,1)$ against $N(\lambda,1)$ in one dimension, because $\|\psi_e\| = 1$. Its deficiency is $\Phi(z_\alpha - \lambda) = e^{-\lambda^2(1+o(1))/2}$, with z_α the $(1-\alpha)$-quantile of $N(0,1)$, and by the Neyman–Pearson lemma no test of level α has a smaller deficiency at e. The maximum test has deficiency at most $\mathbb{P}\{Z(e) \leq u\} = \Phi(u-\lambda)$ at its null threshold u, so its exponent is $1/2$. The Bayes-optimal test maximises the power averaged over e among the tests of level α. By invariance, its power and that of the maximum test do not depend on e. So its deficiency is also at most $\Phi(u-\lambda)$, and its exponent is $1/2$ as well. For the t-spacing test the exponent is $E = \rho^2/(2(A^2+\rho^2))$ (Theorem 2), below $1/2$.

With σ known, the spacing test rejects on a shift of the squared level. The liminf of $-\lambda^{-2}\log(1-\mathbb{P}_\lambda\{\mathbf{S} \leq \alpha\})$ is at least $\rho^2/(2(1+\rho^2)) = k/(2(3k-2))$, that is $k/(3k-2)$ of $1/2$ (Corollary 5); that the exponent of $\mathbf{S}$ is below $1/2$ is not proved. This lower bound exceeds the exponent E of $\mathbf{T}$, since $A > 1$. So the t-test pays more than the known-variance test on the exponential scale: the exponent of $\mathbf{T}$ is below the liminf for $\mathbf{S}$, at every (k,d) and every level.

3.5 Growing dimension

Theorem 3 (Detection at the scale $\sqrt{d}$)**.** *Let* $\alpha \in (0,1)$ *and* $\tau, \delta > 0$*, and put*

$$\begin{aligned} \Upsilon_k &= \sup_{0<\theta<\pi} \frac{(|\cos\theta| + \sin\theta)^k - |\cos\theta|^k - k|\cos\theta|^{k-1}\sin\theta}{1-\cos^k\theta}, \\ \bar{\mathcal{X}} &= \sqrt{kd} + \tau, \qquad \bar{v} = (\Upsilon_k + k - 1)\bar{\mathcal{X}} + (k-1)\tau, \\ \bar{\ell} &= (k-1)(2\bar{\mathcal{X}} + \tau), \qquad \bar{n}^2 = (\lambda+\tau)^2 + m - 1 + \delta. \end{aligned} \tag{16a}$$

If $\lambda \geq \bar{v} + \tau$*,* $m(\lambda-\tau)^2 \geq d\,\bar{n}^2$ *and*

$$d\log\frac{\lambda-\tau}{\bar{v}} + N\log\frac{\bar{v}}{\bar{v}-\bar{\ell}} + \frac{m}{2}\log\Big(1 - \frac{(\lambda-\tau)^2 - \bar{v}^2}{\bar{n}^2}\Big) \;\leq\; \log\alpha, \tag{16b}$$

then

$$1 - \mathbb{P}_\lambda\{\mathbf{T} \leq \alpha\} \;\leq\; e^{-\tau^2/2} + 2\big(1-\Phi(\tau)\big) + \mathbb{P}\{\chi^2_{m-1} > m - 1 + \delta\}. \tag{16c}$$

In particular, let $v_k = (\Upsilon_k + k - 1)\sqrt{k}$ *and* $\ell_k = 2(k-1)\sqrt{k}$*, and let* $C_k > v_k$ *solve*

$$\tfrac{1}{2}\big(C_k^2 - v_k^2\big) - \log\frac{C_k}{v_k} \;=\; \log\frac{v_k}{v_k - \ell_k}. \tag{16d}$$

If k *is fixed and* $\lambda_d/\sqrt{d} \to \bar{\lambda} \in (C_k, \infty)$ *as* $d \to \infty$*, then* $\mathbb{P}_{\lambda_d}\{\mathbf{T} \leq \alpha\} \to 1$*. Moreover*

$$C_k \;\leq\; 1.72 \cdot 2^{k/2}\sqrt{k}. \tag{16e}$$

In the proof, $\bar{\mathcal{X}}$ bounds $\sup X$, $\bar{v}$ bounds λ_2, $\bar{\ell}$ bounds $\lambda_{\max}(\Omega)$ and $\bar{n}$ bounds $\|Y\|$, outside an event whose probability is at most the bound of the theorem. The proof is in Appendix D of the Supplementary Material; the bound on C_k is in Lemma D.5 there. Numerically $\Upsilon_3 = 2.214$, and $C_3 = 7.70$, $C_5 = 21.74$ and $C_7 = 47.06$. As the order k grows, the constant C_k grows like $2^{k/2}\sqrt{k}$: $C_k/(2^{k/2}\sqrt{k}) \to 1$ as $k \to \infty$ along the odd integers (Lemma D.5 of the Supplementary Material). Figure 2 shows this ratio for the odd orders $k \le 999$. The ratio is largest at $k = 5$, where it is 1.72. It decreases from $k = 5$ to $k = 999$, and it is below 1.02 for $k \ge 21$.

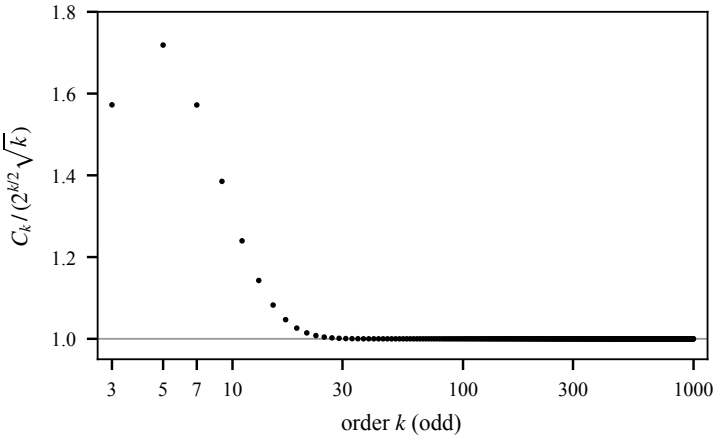

Figure 2: The ratio $C_k/(2^{k/2}\sqrt{k})$ for the odd orders $k = 3, 5, \ldots, 999$, where C_k is defined by (16d) in Theorem 3. The axis of the orders is logarithmic. The ratio tends to one (Lemma D.5 of the Supplementary Material).

Remark 3 (Growing dimension)**.** Let k and α be fixed and $d \to \infty$. The constant A tends to 1 (Lemma D.4 of the Supplementary Material). So the exponent of Theorem 2, which is a limit as $\lambda \to \infty$ at fixed d, tends to $k/(2(3k-2))$. This is the lower bound for $\mathbf{S}$ of Corollary 5. As $A > 1$, the bound (13a) is at least $\mathbb{P}\{\lambda + \xi < \chi_{m-1}/\rho\}$, which tends to one if $\lambda/\sqrt{m}$ tends to a constant below $1/\rho$, since $\chi_{m-1}/\sqrt{m} \to 1$ in probability. So the bound is informative only when λ is at least of order $\sqrt{m}$, that is $d^{k/2}$. This is a property of the bound, not of the test: by Theorem 3 the power tends to one on the scale $\sqrt{d}$. The power of a test at the worst e is at most its power averaged over a uniform e. For $\lambda = \bar{\lambda}\sqrt{d}$ with a fixed $\bar{\lambda}$ below a constant that depends on k, 1.12 for $k = 3$, the total variation distance between the null law and this average tends to zero (Perry et al., 2020, Theorem 3.2). So no test of level α has power larger than $\alpha + o(1)$ at every e, and the power of $\mathbf{T}$, which does not depend on e, is at most $\alpha + o(1)$. We do not know the constant of the threshold of the t-spacing test.

4 Minimax estimation of the direction

4.1 Notation

For $\lambda > 0$ and $b \in (0, \lambda^2]$, the cap of size b around e is

$$\mathcal{B}_b = \{t \in M : \langle e, t\rangle \ge \sqrt{1 - b/\lambda^2}\}.$$

A point t is in $\mathcal{B}_b$ if and only if $\langle e, t\rangle \ge 0$ and $\sin^2 \angle(t, e) \le b/\lambda^2$.

Let $\delta^\lambda(u; B)$ be the expected number of critical points of Z in $B = \{t : \langle e, t\rangle \in I\}$, for an interval $I \subset [-1, 1]$, with value at least u. By the Kac–Rice formula,

$$\delta^\lambda(u; B) = \omega_{d-2}\Big(\frac{k-1}{\pi}\Big)^{\frac{N}{2}} \int_I (1-c^2)^{\frac{d-3}{2}} e^{-\frac{k\lambda^2}{2}c^{2k-2}(1-c^2)} \int_u^\infty A_\lambda(\rho x; c)\varphi(x - \lambda c^k)\,\mathrm{d}x\,\mathrm{d}c,$$

where

$$\omega_{d-2} = 2\pi^{(d-1)/2}/\Gamma(\tfrac{d-1}{2}) \text{ is the area of } \mathbb{S}^{d-2};$$
$$A_\lambda(z;c) = \mathbb{E}\big|\det(z\,\mathrm{Id} - G_\lambda)\big| \text{ is the mean absolute determinant.}$$

4.2 Bounds on the risk

The test returns the argmax $\hat{t}$. At fixed (k,d) the classical theory gives, for the maximum likelihood estimator, a limit law with covariance $(k\lambda^2)^{-1}$ on the tangent space; the bounds below hold at finite λ, and the upper ones contain the control of the far critical points of Z, a global statement which the local theory does not give. Let the loss be $\ell(\hat{e},e) = 1 - \langle \hat{e}, e\rangle^2 = \sin^2 \angle(\hat{e},e)$. For an estimator $\hat{e} = \hat{e}(Y)$ with values in $\mathbb{S}^{d-1}$, let

$$R_\lambda(\hat{e}) = \mathbb{E}\big[\sin^2 \angle(\hat{e},e)\big], \tag{17a}$$
$$R(\lambda) = \inf_{\hat{e}}\ R_\lambda(\hat{e}), \tag{17b}$$

the expectation being under the joint law of (e,Y) with e uniform on $\mathbb{S}^{d-1}$, and the infimum over all estimators. So $R_\lambda(\hat{e})$ is the Bayes risk of $\hat{e}$ under the uniform prior, and $R(\lambda)$ is the Bayes risk. The minimax risk is at least $R(\lambda)$, a supremum over e being at least an average. The risk of $\hat{t}$, which is $O(d)$-equivariant, does not depend on e, so it equals $R_\lambda(\hat{t})$. A bracket on $R(\lambda)$ whose upper side is attained by $\hat{t}$ is therefore a statement about the minimax risk.

Part (iii) of the theorem holds beyond a threshold $\lambda_0(k,d)$, which is defined by explicit inequalities. Let $m_N = \sqrt{2N} + \frac{5}{6}\sqrt{\max(2N,8)}$, $K_N = 2\sqrt{2}\,(\frac{5}{4} + m_N/N)^N$ and $\omega_{d-1} = 2\pi^{d/2}/\Gamma(\frac{d}{2})$, the area of $\mathbb{S}^{d-1}$. Let $\lambda_0(k,d)$ be the least number such that the five conditions

$$\lambda^2 \geq 128 \log\lambda, \tag{18a}$$
$$\lambda > \frac{4N}{3\rho}, \qquad \lambda \geq \frac{16\sqrt{2N}}{9\rho}, \tag{18b}$$
$$g_1 := \tfrac{87}{128}\rho\lambda - \sqrt{2N} \geq 0, \qquad K_N \lambda^2 e^{-g_1^2/4} \leq 1, \tag{18c}$$
$$g_3 := \tfrac{9}{16}\rho\lambda - \sqrt{2N} \geq 0, \qquad \Big(\frac{5}{4} + K_N e^{-g_3^2/4}\Big) e^{\frac{128N(N+1)}{169\lambda^2}}\,\frac{4k-5}{3k-4} \leq 2, \tag{18d}$$
$$\lambda \geq 4\sqrt{d+3}, \qquad \tfrac{10(4k-1)}{3}\,\omega_{d-1}\big(\tfrac{k-1}{\pi}\big)^{N/2}\big(\rho\lambda + \rho\sqrt{N} + m_N\big)^N \lambda^4 e^{-\lambda^2/32} \leq 1 \tag{18e}$$

hold at every $\lambda > \lambda_0(k,d)$. This number is finite, and each condition, once true at some $\lambda \geq e$, stays true for larger λ (Appendix C of the Supplementary Material). Conditions (18a) and (18b) make the choice of θ_1 and u of part (iii) admissible in part (ii), and (18c), (18d) and (18e) bound its three error terms.

Theorem 4 (Estimation of the direction: lower and upper bounds)**.** (i) *For every estimator $\hat{e}$,*

$$R_\lambda(\hat{e}) \ \geq\ \frac{\big((d-1) - d\,R_\lambda(\hat{e})\big)^2}{(d-1)\,k\,\lambda^2}. \tag{19a}$$

(ii) *For the argmax $\hat{t}$, let $\theta_1 \in (0,\pi/2)$ and let u satisfy*

$$\frac{d-1}{\rho} < u < \lambda\cos^k\theta_1,$$
$$u \ \geq\ \max\Big\{4(k-1)\lambda\sin^2\theta_1,\ (k-1)\lambda\sin^2\theta_1 + \frac{\sqrt{2(d-1)}}{\rho}\Big\}.$$

Then

$$R_\lambda(\hat{t}) \ \leq\ \frac{d-1}{k\lambda^2}\cdot\frac{1+\epsilon_A}{(1-\varrho)^{\frac{d+1}{2}}} \ +\ \mathrm{Rem}_{\mathrm{far}} \ +\ \Phi(u-\lambda), \tag{19b}$$

with $\varrho = 1 - \inf_{0<\theta\le\theta_1} \cos^{2k-2}\theta\,(\sin\theta/\theta)^2$, ϵ_A *the relative error of Lemma C.2 and* $\mathrm{Rem}_{\mathrm{far}} = \delta^\lambda(u;\{\theta>\theta_1\})$, θ *the angle between* t *and* e.

(iii) *For the argmax* $\hat t$ *and* $\lambda > \lambda_0(k,d)$, *the threshold defined by* (18),

$$R_\lambda(\hat t) \;\le\; \frac{d-1}{k\lambda^2}\exp\Big(\frac{c_1\log\lambda + c_1'}{\lambda^2}\Big) + 2\,\mathbb{P}\big\{\chi^2_{d-1} > 8\log\lambda\big\} + \frac{1}{4\lambda^4}, \tag{19c}$$

with $c_1 = \frac{27(3k-2)(d+1)}{16k} + \frac{12(k-1)}{k}$ *and* $c_1' = \frac{8}{9}\big((d-1)^2+2\big)$.

Since $R(\lambda) \le R_\lambda(\hat t)$, (i) at a Bayes rule and (ii) or (iii) bracket $R(\lambda)$.

The three error terms of (ii) have simple bounds, proved in Lemmas C.3 and C.4 of the Supplementary Material. For every $\theta_1 \in (0,\pi/2)$, $\varrho \le (k-\frac{2}{3})\theta_1^2$; ϱ is the relative loss in the exponent of the Gaussian weight inside the cap $\theta \le \theta_1$. Assume now the hypotheses of (ii). If also $u \ge 4\sqrt{2(d-1)}/\rho$, then

$$\epsilon_A \;\le\; \frac{(k-1)\lambda\sin^2\theta_1}{u} + \frac{3(d-1)^2}{2\lambda^2\cos^{2k}\theta_1} + 3e^{\frac{5}{4}(d-1)-\frac{\rho^2u^2}{16}}, \tag{20a}$$

the three terms coming from the spike of the Hessian, from the Gaussian moments of the value and from the negative part of the determinant. If $3\rho u \ge 4\sqrt{2(d-1)}$, then

$$\mathrm{Rem}_{\mathrm{far}} \;\le\; C_M\,\mathbb{P}\big\{\chi^2_{d-1} > k\lambda^2\cos^{2k-2}\theta_1\sin^2\theta_1\big\} + F_{\mathrm{far}}, \tag{20b}$$

the chi-square tail counting the critical points near e beyond the angle θ_1, and F_{far} those far from e. The factor C_M and the remainder F_{far} are explicit functions of (k,d,λ,u). For $\lambda > \lambda_0(k,d)$ and the choice of θ_1 and u below, $C_M \le 2$ and $F_{\mathrm{far}} \le \lambda^{-4}/10$.

In (iii) the additive term is $O(\lambda^{-4}(\log\lambda)^{(d-3)/2})$: for $\log\lambda \ge (d-3)/2$, $2\,\mathbb{P}\{\chi^2_{d-1} > 8\log\lambda\} \le 8(4\log\lambda)^{(d-3)/2}\lambda^{-4}/(3\,\Gamma(\frac{d-1}{2}))$ (Appendix C). Rounded up, $\lambda_0(k,d) = 28.83$, 33.11, 29.50, 38.52 and 34.52 at $(k,d) = (3,3)$, $(3,5)$, $(5,3)$, $(3,8)$ and $(7,5)$. Part (iii) is (ii) at $\theta_1 = 3\lambda^{-1}\sqrt{\log\lambda/k}$ and $u = \lambda - 2\sqrt{2\log\lambda}$, with the bounds above.

Proof. (i) Write $R = R_\lambda(\hat e)$. Let $v_1,\dots,v_{d-1}$ be an orthonormal frame of $T_e\mathbb{S}^{d-1}$ and

$$S_i = \partial_{v_i}\log\big(\pi(e)f(Y\mid e)\big)$$

the score of the joint law of (e,Y). The prior being uniform, π is constant and $S_i = \lambda\langle Y, \partial_{v_i}\psi_e\rangle$; since $\langle\psi_e, \partial_v\psi_e\rangle = \frac{1}{2}\partial_v\|\psi_e\|^2 = 0$ and $Y = \lambda\psi_e + W$, $S_i = \lambda\langle W, \partial_{v_i}\psi_e\rangle$. Differentiating $\langle\psi_e,\psi_{e'}\rangle = \langle e,e'\rangle^k$ twice gives $\langle\partial_{v_i}\psi_e, \partial_{v_j}\psi_e\rangle = k\delta_{ij}$, so $\mathbb{E}\|S\|^2 = (d-1)k\lambda^2$.

Write $a = \langle\hat e, e\rangle$ and $\ell = 1-a^2$, and take the tangent field

$$g(e) = a\,\big(\hat e - a\,e\big), \qquad \|g\|^2 = a^2\ell. \tag{21a}$$

The factor a is what makes the argument insensitive to the sign of a, which the loss does not see and which no estimator can control: the field $\hat e - a\,e$ alone would produce the factor $(\mathbb{E}a)^2$, and $\mathbb{E}a$ vanishes when the data are uninformative. Moving e along a unit tangent v, with $\partial_v a = \langle\hat e, v\rangle$ and $\langle e, v\rangle = 0$,

$$\partial_v g = \langle\hat e, v\rangle\hat e - 2a\langle\hat e, v\rangle e - a^2 v, \qquad \langle\partial_v g, v\rangle = \langle\hat e, v\rangle^2 - a^2, \tag{21b}$$

and summing over the frame, with $\sum_i\langle\hat e, v_i\rangle^2 = \|\hat e - a\,e\|^2 = \ell$,

$$\mathrm{div}_e\, g = \ell - (d-1)a^2 = \ell - (d-1)(1-\ell). \tag{21c}$$

For fixed Y the field $e \mapsto g(e)$ is a polynomial in e, $\hat e = \hat e(Y)$ being a fixed unit vector, so the divergence theorem on the closed sphere gives $\int\langle g, \nabla_e(\pi f(Y\mid e))\rangle\,\mathrm{d}e = -\int \mathrm{div}_e\, g\,\pi f(Y\mid e)\,\mathrm{d}e$ for each Y; integrating in Y (Fubini, the integrand being integrable) and using $S\,\pi f = \nabla_e(\pi f)$, $\mathbb{E}\langle g, S\rangle = -\mathbb{E}[\mathrm{div}_e\, g] = (d-1) -$

$d\,R$. No regularity of $Y \mapsto \hat{e}(Y)$ beyond measurability is used. Cauchy–Schwarz with $\mathbb{E}\|g\|^2 = \mathbb{E}[a^2\ell] \le R$ gives $\big((d-1) - dR\big)^2 \le R\,(d-1)k\lambda^2$.

(ii) The hypotheses require $\lambda > (d-1)/(\rho\cos^k\theta_1)$; the hypotheses other than $u > (d-1)/\rho$ are used only inside Lemma C.2. On $\{\sin^2\angle(\hat{t}, e) > b/\lambda^2\}$ the argmax lies outside $\mathcal{B}_b$. Either $\sup Z \le u$, of probability at most $\mathbb{P}\{Z(e) \le u\} = \Phi(u-\lambda)$ since $Z(e) \sim N(\lambda, 1)$, or $\hat{t}$ is a critical point outside $\mathcal{B}_b$ with value above u, of probability at most $\delta^\lambda(u; \mathcal{B}_b^c)$ by Markov. Lemma C.2 bounds that count by

$$\frac{1+\epsilon_A}{(1-\varrho)^{\frac{d-1}{2}}}\,\mathbb{P}\big\{\chi^2_{d-1} > kb(1-\varrho)\big\} + \mathrm{Rem}_{\mathrm{far}}. \tag{22a}$$

Integrating the tail of the loss,

$$\begin{aligned} R_\lambda(\hat{t}) &= \frac{1}{\lambda^2}\int_0^{\lambda^2} \mathbb{P}\Big\{\sin^2\angle > \frac{b}{\lambda^2}\Big\}\,\mathrm{d}b \\ &\le \frac{1}{\lambda^2}\int_0^{\infty} \frac{(1+\epsilon_A)\,\mathbb{P}\{\chi^2_{d-1} > kb(1-\varrho)\}}{(1-\varrho)^{\frac{d-1}{2}}}\,\mathrm{d}b + \mathrm{Rem}_{\mathrm{far}} + \Phi(u-\lambda), \end{aligned} \tag{22b}$$

and $\int_0^\infty \mathbb{P}\{\chi^2_{d-1} > kb(1-\varrho)\}\,\mathrm{d}b = \mathbb{E}\chi^2_{d-1}/(k(1-\varrho)) = (d-1)/(k(1-\varrho))$.

Part (iii) is proved in Appendix C of the Supplementary Material. □

Corollary 7 (Asymptotic minimaxity with the sharp constant)**.** *As $\lambda \to \infty$,*

$$\begin{aligned} \frac{d-1}{k\lambda^2}\Big(1 - \frac{d\,R(\lambda)}{d-1}\Big)^2 \;&\le\; R(\lambda) \;\le\; R_\lambda(\hat{t}) \\ &\le\; \frac{d-1}{k\lambda^2}\Big(1 + O\Big(\frac{d^2 + d\log\lambda}{\lambda^2}\Big)\Big) + O\big(\lambda^{-4}(\log\lambda)^{\frac{d-3}{2}}\big), \end{aligned}$$

so that $\lambda^2 R_\lambda(\hat{t}) \to (d-1)/k$ and $\lambda^2 R(\lambda) \to (d-1)/k$: the argmax is asymptotically minimax for the direction, with the sharp constant.

Proof. The upper bound is Theorem 4(iii), with $e^x - 1 \le xe^x$, $c_1 \le \frac{81}{16}(d+1) + 12$, $c_1' \le \frac{8}{9}(d^2+2)$ and the bound on the chi-square tail given after the theorem. It forces $R(\lambda) \to 0$, and the lower bound is Theorem 4(i) applied to the Bayes rule, a top eigenvector of $\mathbb{E}[ee^\top \mid Y]$, whose risk is $R(\lambda)$. □

Remark 4. The localisation radius is never chosen: it is integrated out, and the constant $(d-1)/k$ is the whole chi-square tail. This is why the two sides of Theorem 4 carry the same constant. Solving the inequality of part (i) gives, at $(k, d) = (3, 5)$, the values 0.0472, 0.01291, 0.003306 and 0.000832 for $\lambda = 5, 10, 20, 40$, which are 88.54%, 96.80%, 99.18% and 99.79% of $(d-1)/(k\lambda^2)$.

5 Scope: what happens at vanishing signal

At high signal the maximum test is the zero-temperature limit of the Bayes-optimal test (Proposition 6). At vanishing signal this picture reverses. The mixture likelihood ratio has no first-order term, because k is odd and $\int e'^{\otimes k}\,\mathrm{d}e' = 0$ by antipodal symmetry, so

$$\bar{L}_\lambda = 1 + \frac{\lambda^2}{2}\,(T_{\mathrm{q}} - 1) + O(\lambda^3), \qquad T_{\mathrm{q}} = \int_{\mathbb{S}^{d-1}} Z(e')^2\,\mathrm{d}e' = \langle Y, \Sigma Y\rangle, \quad \Sigma = \int \psi_{e'}\psi_{e'}^\top\,\mathrm{d}e',$$

a quadratic form in the observed tensor. Expanding $\mathbb{E}_\lambda\phi = \mathbb{E}_0[\phi\bar{L}_\lambda]$ to second order, every test ϕ of size α satisfies $\mathbb{E}_\lambda[\phi] = \alpha + \frac{\lambda^2}{2}\mathrm{Cov}_0(\phi, T_{\mathrm{q}}) + o(\lambda^2)$, a linear functional of ϕ under the constraint $\mathbb{E}_0\phi = \alpha$, hence maximised by the upper-tail test on T_{q}. *At vanishing signal the optimal procedure is the quadratic test, not this one.*

The null law of T_{q} is explicit: by Funk–Hecke the kernel $\langle e, e' \rangle^k$ has eigenvalue

$$\mu_j = \int_{-1}^{1} t^k \, \frac{C_j^{((d-2)/2)}(t)}{C_j^{((d-2)/2)}(1)} \, \frac{\Gamma(\frac{d}{2})}{\sqrt{\pi}\,\Gamma(\frac{d-1}{2})} (1-t^2)^{\frac{d-3}{2}} \, \mathrm{d}t,$$

with $C_j^{((d-2)/2)}$ the Gegenbauer polynomial, on the spherical harmonics of degree $j \le k$ with $j \equiv k$ modulo 2, of dimension $n_j = \binom{d+j-1}{j} - \binom{d+j-3}{j-2}$, and all other eigenvalues vanish, so $T_{\mathrm{q}} \overset{d}{=} \sum_j \mu_j \chi^2_{n_j}$ with independent factors. One checks $\sum_j \mu_j n_j = \mathbb{E} X^2 = 1$ and $\sum_j \mu_j^2 n_j = \mathbb{E}\, T_{\mathrm{cor}}^{2k}$, where $T_{\mathrm{cor}} = \langle e, e' \rangle$ for independent uniform directions, so the mean shift under the alternative is $\lambda^2 \mathbb{E} T_{\mathrm{cor}}^{2k}$ and the null variance is $2\mathbb{E} T_{\mathrm{cor}}^{2k}$, with $\mathbb{E} T_{\mathrm{cor}}^{2k} = (2k-1)!!/\prod_{i<k}(d+2i)$.

How much does a test built on the maximum lose? With q_α and t_α the $(1-\alpha)$-quantiles of $\sup X$ and of T_{q} under the null, its local coefficient is

$$\tfrac{1}{2}\mathrm{Cov}_0(\mathbf{1}\{\sup X > q_\alpha\}, T_{\mathrm{q}}), \tag{23a}$$

and the optimal one is

$$\tfrac{1}{2}\mathrm{Cov}_0(\mathbf{1}\{T_{\mathrm{q}} > t_\alpha\}, T_{\mathrm{q}}) = \tfrac{1}{2}(\mathbb{E}[T_{\mathrm{q}}; T_{\mathrm{q}} > t_\alpha] - \alpha); \tag{23b}$$

since T_{q} is a chi-square mixture and not Gaussian, the optimal coefficient has to be computed from its own upper tail. At $(k,d) = (3,3)$, where $T_{\mathrm{q}} = \frac{1}{5}\chi^2_3 + \frac{2}{35}\chi^2_7$, simulation gives a local efficiency of 0.810 at $\alpha = 0.05$ and 0.777 at $\alpha = 0.01$.

When σ is unknown the quadratic test is not available. Among tests invariant under $Y \mapsto aY$, $a > 0$, and under the orthogonal group, the locally most powerful one rejects for large $\tilde{T} = \sum_j \mu_j Q_j / \sum_j Q_j$, $Q_j = \|P_j Y\|^2$ the squared norm of the projection on the harmonics of degree j: the likelihood ratio of $Y/\|Y\|$, averaged over the orbit of e, is $1 + \frac{\lambda^2}{2}(m\tilde{T} - 1) + O(\lambda^3)$. At $(k,d) = (3,3)$, $\tilde{T}$ is increasing in Q_1/Q_3 and $\frac{7}{3}Q_1/Q_3$ is $F_{3,7}$ under the null; the local coefficient of this F-test is $1.54 \cdot 10^{-2}$ at $\alpha = 0.05$, 42% of the quadratic test's and 52% of the maximum test's, and its power is 0.100 at $\lambda = 2$ and 0.191 at $\lambda = 5$.

The window in which this matters is narrow. Comparing the two powers at fixed (k,d) rather than their local coefficients is a finite computation: under the alternative the quadratic statistic is the noncentral mixture

$$T_{\mathrm{q}} \overset{d}{=} \sum_j \mu_j \, \chi^2_{n_j}(\lambda^2 \mu_j n_j),$$

because the projection of $\lambda \psi_e$ on the harmonics of degree j has squared norm $\lambda^2 \mu_j n_j$, while the power of the maximum test is obtained by simulating the field. At $(k,d) = (3,3)$:

	$\lambda = 1.4$	1.6	1.8	2.0	3.0	5.0
quadratic test, $\alpha = 0.05$	0.136	0.166	0.203	0.246	0.534	0.963
maximum test, $\alpha = 0.05$	0.131	0.165	0.205	0.254	0.584	0.982
S, $\alpha = 0.05$	0.112	0.138	0.173	0.213	0.510	0.962
T, $\alpha = 0.05$	0.069	0.078	0.090	0.105	0.233	0.650

The first two cross at $\lambda \approx 1.7$ for $\alpha = 0.05$ and at $\lambda \approx 1.46$ for $\alpha = 0.01$, where the common power is 0.19 and 0.045 respectively. *The quadratic test is preferable only where no test has useful power*, and the maximum test leads everywhere else. The spacing tests, built on the same maximum, pay in addition the price of exactness of Remark 2: **S** is below the maximum test at every λ and within 10^{-3} of the quadratic test at $\lambda = 5$, and **T**, which also estimates σ with $\kappa = 7$ degrees of freedom, has power 0.65 at $\lambda = 5$. The rows **S** and **T** come from $5 \cdot 10^5$ replicates each of the Karhunen–Loève simulation of the

field used for the measured exponent of Section 6, with standard error below 10^{-3}; both statistics pass a Kolmogorov–Smirnov test of uniformity on 10^6 null replicates, $p = 0.28$ and 0.99.

No single procedure is therefore optimal at both ends, and the end at which the maximum test is not optimal is the end at which nothing works. What the t-spacing test has and the quadratic test has not is an exact level at every λ, no need for the noise level, and the maximiser as a by-product, which Theorem 4 and Corollary 7 show to be an asymptotically minimax estimator of the direction.

6 Numerical illustration

The level maps and the power bound are computed from the closed forms of Appendix F of the Supplementary Material, which were checked against numerical quadrature to ten significant digits. The estimation sandwich also uses a Kac–Rice integral under the alternative, evaluated numerically. The simulations are described in each paragraph.

The model constants. The Karhunen–Loève order, the degrees of freedom of $\hat{\sigma}^2$, and the degrees of freedom of the reference Student density.

(k, d)	$(3, 3)$	$(3, 5)$	$(3, 8)$	$(5, 6)$
$m = \binom{d+k-1}{k}$	10	35	120	252
$\kappa = m - d$	7	30	112	246
$\nu = m - 1$	9	34	119	251

The price of knowing σ. The level map $u_\alpha(r, v)$ of (12a) gives the threshold of the t-test on T_1 at $T_2 = v$ and $\Omega/\hat{\sigma} = r$. The same definition with G_r gives the threshold of $\mathbf{S}$. Compare these two level maps at the same α and the same second-maximum threshold v, for one draw of $\Omega = G/\rho$, $G \sim \mathrm{GOE}(N)$, at each (k, d). The draw satisfies $v > \lambda_{\max}(\Omega)$, and we take $\hat{\sigma} = \sigma = 1$, so that both maps are evaluated at $r = \Omega$. The t-test asks for a higher first maximum, and that gap is the price of not having σ.

(k, d)	α	v	u_α, known variance	u_α, t-test	extra level
$(3, 5)$	0.05	4	4.897	5.486	+0.589
$(3, 5)$	0.01	4	5.276	6.171	+0.896
$(3, 8)$	0.05	4	5.036	5.202	+0.166
$(5, 6)$	0.05	4	4.862	4.923	+0.061

The draw at $(3, 5)$ is on the high side: over 3000 draws the median extra level is 0.505 and this one exceeds 90% of them; at $(3, 8)$ and $(5, 6)$ the draws are close to the median. The gap is not the inflation of a single quantile: with the scale γ of H_r, the 0.95-quantile of the reference Student law is 3.4% below the Gaussian one at $(3, 5)$, and its tail is heavier only beyond $u = 2.7$. The level map compares two tail integrals of a polynomial of degree N against the two densities, so the heavier Student tail is penalised N times over. It falls below 0.1 in level as soon as m is a few hundred.

The power bound. At $(k, d) = (3, 5)$ and $\alpha = 0.05$, the column of Remark 1 and its upper bound give $1.2873 \le A \le 1.2878$, and $m - 1 = 34$. The bound (13a) of Corollary 4 is 0.05 at $\lambda = 11.0$ and 0.01 at $\lambda = 12.0$. Its Gaussian form (13b) reaches these values at $\lambda = 13.1$ and 14.1. In the table of the next paragraph, the deficiency is 0.05 near $\lambda = 6.2$ and 0.01 near $\lambda = 7.2$, by log-linear interpolation. The bound therefore needs a signal 1.7 to 1.8 times larger.

The true exponent, measured. The deficiency $1 - \mathbb{P}_\lambda\{\mathbf{T} \le \alpha\}$ can be simulated directly, the field being generated from its Karhunen–Loève expansion and the two maxima, the Hessian and $\hat{\sigma}$ computed from the realised path. The maximum λ_1 is found by gradient ascent on the sphere from 41 starting points, 40 random ones and e. The supremum (1c) is found in the same way, from 40 random points, the directions

of the eigenvectors of Ω at $\hat{t}$ and the other local maxima of Z, and λ_2 is the larger of this value and its limit $\lambda_{\max}(\Omega)$ at $\hat{t}$. On 2% of the replicates, run again with four times as many starting points, λ_1 never changed, λ_2 rose 3 times in 279 000, and no decision changed. Before measuring anything one checks the calibration: with 10^6 null replicates at each pair the statistic passes a Kolmogorov–Smirnov test of uniformity, with $p = 0.18$, 0.31, 0.56 and 0.88 at $(3,3)$, $(3,4)$, $(3,5)$ and $(5,3)$, and the order relation (3) holds in every replicate under the alternative. Under the alternatives the number of replicates grows with λ, from $2 \cdot 10^4$ to $2 \cdot 10^6$ per point and $1.0 \cdot 10^7$ in all, so that at each point the test fails to reject at least 78 times. At $\alpha = 0.05$:

λ	5	6	7	8	9	10	12
$(3,3)$, $\kappa = 7$		0.183	0.0791	$2.90 \cdot 10^{-2}$	$1.00 \cdot 10^{-2}$	$2.84 \cdot 10^{-3}$	$1.57 \cdot 10^{-4}$
$(3,4)$, $\kappa = 16$	0.251	$8.65 \cdot 10^{-2}$	$2.21 \cdot 10^{-2}$	$4.01 \cdot 10^{-3}$	$5.75 \cdot 10^{-4}$	$6.10 \cdot 10^{-5}$	
$(3,5)$, $\kappa = 30$	0.238	$7.40 \cdot 10^{-2}$	$1.40 \cdot 10^{-2}$	$1.88 \cdot 10^{-3}$	$1.20 \cdot 10^{-4}$		
$(5,3)$, $\kappa = 18$	0.175	$5.05 \cdot 10^{-2}$	$1.02 \cdot 10^{-2}$	$1.53 \cdot 10^{-3}$	$1.95 \cdot 10^{-4}$		

Regressing $-\log$ of these on λ^2 gives a slope 0.065 ± 0.002 at $(3,3)$, flat to within 4% over $\lambda \in [6,12]$, and 0.115 ± 0.012 at $(3,4)$, where the slope is still climbing and the quoted uncertainty is dominated by that drift. At $(5,3)$ the slope is 0.124 ± 0.003, flat over $\lambda \in [6,9]$. At $(3,5)$ the slope is not stable over the window. Weighted fits give 0.136, 0.143 and 0.153 over $\lambda \in [6, 9.5]$, $[7, 9.5]$ and $[8, 9.5]$, and we quote 0.14 ± 0.01. The local slopes over successive intervals are 0.106, 0.128, 0.134, 0.162 and, on $[9, 9.5]$, 0.12 ± 0.02; the deficiency is $3.9 \cdot 10^{-5}$ at $\lambda = 9.5$, from 78 failures in $2 \cdot 10^6$ replicates. Two readings follow. The exponent depends on the dimension, as E does through κ and N: a d-free exponent cannot describe the three rows with $k = 3$. And the measured values are close to the exponent of Theorem 2:

(k,d)	measured	Theorem 2, E
$(3,3)$	0.065 ± 0.002	0.0615
$(3,4)$	0.115 ± 0.012	0.120
$(3,5)$	0.14 ± 0.01	0.156
$(5,3)$	0.124 ± 0.003	0.1158

A slope over a finite window is not the limit, and the theorem gives the limit only: the measured slope is 6% above E at $(3,3)$, 7% above at $(5,3)$, where the difference is 2.7 times the quoted uncertainty, within its error at $(3,4)$, and 10% below at $(3,5)$. At every λ of the table above, $-\lambda^{-2}$ log of the deficiency is below E: the deficiency exceeds $e^{-E\lambda^2}$, by a factor 1.7 at $\lambda = 6$ and 1.1 at $\lambda = 12$ at $(3,3)$.

The estimation sandwich. Parts (i) and (ii) of Theorem 4 bracketed numerically at $(k,d) = (3,5)$, the middle column being the common target $(d-1)/(k\lambda^2)$.

λ	u	lower bound, (i)	$(d-1)/(k\lambda^2)$	upper bound, Kac–Rice	ratio
20	16	$3.306 \cdot 10^{-3}$	$3.333 \cdot 10^{-3}$	$3.390 \cdot 10^{-3}$	1.017
20	18	$3.306 \cdot 10^{-3}$	$3.333 \cdot 10^{-3}$	$2.604 \cdot 10^{-2}$	7.81
40	34	$8.316 \cdot 10^{-4}$	$8.333 \cdot 10^{-4}$	$8.349 \cdot 10^{-4}$	1.0019
40	36	$8.316 \cdot 10^{-4}$	$8.333 \cdot 10^{-4}$	$8.666 \cdot 10^{-4}$	1.0399

The upper column is the bound $\int \sin^2\theta\, \delta^\lambda(u; \mathrm{d}\theta) + \Phi(u-\lambda)$ reached in the proof of (ii) before Lemma C.2 is applied, the Kac–Rice integral being evaluated numerically. The right-hand side of (ii), with ϵ_A bounded as in Lemma C.2, $\mathrm{Rem}_{\mathrm{far}}$ evaluated numerically, and optimised over θ_1, is looser: $4.261 \cdot 10^{-3}$, $2.695 \cdot 10^{-2}$, $8.934 \cdot 10^{-4}$ and $9.244 \cdot 10^{-4}$ in the four rows, ratios 1.278, 8.09, 1.072 and 1.109. Optimised over (θ_1, u), the explicit bound of (ii) is within a factor 2 of the target from $\lambda \approx 11$, 13, 13 and 16 at $(3,3)$, $(5,3)$, $(3,5)$ and $(3,8)$, and within 10% from $\lambda \approx 26$, 28, 34 and 45 at the same pairs. At $\lambda = 10$ it is 4 to 46 times the target: the remainder $\varepsilon_{\mathrm{det}}$ of Lemma C.2, the gap term $\Phi(u-\lambda)$ and $\mathrm{Rem}_{\mathrm{far}}$ are then all of the order of the target. The Kac–Rice bound from which it is derived is within 10% of the lower bound from

$\lambda \approx 9$ (10.4 at $(3,8)$). With the Kac–Rice column the two sides bracket the Bayes risk within 2.6% at $\lambda = 20$ and within 0.40% at $\lambda = 40$. The rows with the larger u show the role of the gap $\lambda - u$: the term $\Phi(u-\lambda)$ enters the risk additively, so it must be pushed below the target, which asks for $\lambda - u \gtrsim \sqrt{2\log\lambda^2}$; at $\lambda = 40$, $\lambda - u = 4$ costs 3.8% and $\lambda - u = 6$ costs nothing.

The figure of the introduction. Figure 1 is computed on the Karhunen–Loève simulation above at $(k,d) = (3,8)$, $\kappa = 112$, with 10^5 replicates per point and the same replicates for every test. Besides $\mathbf{T}$, $\mathbf{S}$ and the maximum test, it shows four competitors. The studentised maximum $\sup Z/\hat\sigma$ is scale-free; its null law has no closed form and is simulated once per (k,d), here on 10^6 replicates. Compared instead with the quantile of $\sup Z/\sigma$, as if $\hat\sigma$ were exact, it rejects with probability 0.097 at nominal level 0.05 and 0.026 at 0.01. The locally optimal invariant test of Section 5 rejects for large Q_1/Q_3, with $(n_3/n_1)Q_1/Q_3 \sim F_{8,112}$ under the null; its power is 0.20 at $\lambda = 4$ and 0.94 at $\lambda = 12$. The unfolding test rejects for large $s_1^2/\|Y\|^2$, s_1 the largest singular value of the $d \times d^{k-1}$ unfolding of Y (Richard and Montanari, 2014), calibrated by simulation; its power is above that of $\mathbf{T}$ by at most 0.011. The most powerful test invariant under rotations and scalings, at a given λ/σ, rejects for large

$$\tilde L_\lambda(Y) = e^{-\lambda^2/2} \int_{\mathbb{S}^{d-1}} \frac{I(\lambda Z(e')/\|Y\|)}{I(0)}\,\mathrm{d}e', \qquad I(c) = \int_0^\infty r^{m-1} e^{-r^2/2+cr}\,\mathrm{d}r,$$

the likelihood ratio of the maximal invariant (Wijsman, 1967); the group being amenable, this test is maximin over (e,σ) at fixed λ/σ among all tests of level α, by the Hunt–Stein theorem (Lehmann and Romano, 2005, Section 8.5). The integral over the sphere is estimated by importance sampling with 2000 points per replicate and the threshold is calibrated on the same statistic, so that the level is exact and the power is a lower bound for that of the ideal test; on 3000 replicates a cloud four times larger changes the power by at most 0.002. Its power exceeds that of the studentised maximum by at most 0.011 at every λ, the finite-λ form of Proposition 6: the maximum carries almost all the information of the likelihood ratio. The t-spacing test is below both by at most 0.08, at $\lambda = 5$, and the tests with σ known are above it by at most 0.16. Panel (c) shows $(3,8)$ and $(5,3)$; the same simulation was also run at $(3,3)$ and $(3,5)$. The risk does not depend on e, by invariance, so it is the Bayes risk of Theorem 4. At $\lambda = 10$ the risk of the argmax is within 10% of the lower bound at the four pairs and the unfolding estimator is above it by 7 to 24%; at $\lambda = 20$ and $(3,8)$ the two figures are 2.5% and 5.5%.

Local efficiency at vanishing signal. At $(k,d) = (3,3)$, the local coefficients of Section 5, the field being simulated on 40 000 points of $\mathbb{S}^2$ with $2 \cdot 10^5$ replicates and the optimal coefficient cross-checked against the exact mixture $\frac{1}{5}\chi^2_3 + \frac{2}{35}\chi^2_7$.

α	q_α	t_α	$C_{\max}$	C_{opt}	efficiency
0.05	3.181	2.022	$2.959 \cdot 10^{-2}$	$3.655 \cdot 10^{-2}$	0.810
0.01	3.678	2.732	$8.398 \cdot 10^{-3}$	$1.081 \cdot 10^{-2}$	0.777

Acknowledgments

Yohann De Castro acknowledges the support of the Institut Universitaire de France (IUF).

Statements and Declarations

Competing interests. The authors have no competing interests to declare that are relevant to the content of this article.

Data availability. The scripts that produce every number and table of this article are available at `https://github.com/ydecastro/TENSOR-KSS`, folder `paper2`. No other data were generated.

Disclosure of AI use

In line with the Leiden Declaration on Artificial Intelligence and Mathematics (2026), we disclose the use of AI agents (Claude Code, Anthropic) in the preparation of this paper. These agents were used to audit the statements, the proofs and the numerical tables, through repeated adversarial verification passes recorded in a change-log, and to help with editorial tasks: consistency of notation, bibliography formatting, and typesetting. The authors produced and checked every result and every proof themselves and take sole responsibility for the correctness of the paper and for its citations.

Supplementary Material

The Supplementary Material contains the Kac–Rice formula with marks (Appendix A), the layer decomposition of the helix process and its drift, restated from Azaïs et al. (2026), and the continuity of the second maximum at a point (Appendix B), the lemmas on the Gaussian orthogonal ensemble used in the proofs and the proof of Theorem 4(iii) of the main paper (Appendix C), the proofs of Proposition 1, Lemmas 1, 4 and 6 and Theorem 3 of the main paper, the matching bound of Theorem 2, the limit of the constant A as the dimension grows and the growth of the constant C_k of Theorem 3 with the order k (Appendix D), the normalisations of the literature translated into our variables (Appendix E), and the closed forms of the functions G_r and H_r of the two statistics (Appendix F). Statements and equations are numbered by appendix. References to the main paper are marked as such.

A The Kac–Rice formula with marks

Let $\lambda \geq 0$. Laws and expectations are under $\mathbb{P}_\lambda$. Let $f : M \times \mathbb{R}^m \to \mathbb{R}$ be continuous and write $f_t = f(t, Y)$. A mark attached to a critical point t is an event $\mathcal{M}_t$ of the form $\{f_t \neq 0\}$, an open condition, or $\{f_t = 0\}$, a closed condition. With $f_t = 1$ the first form is the mark that always holds. Let $B \subseteq M$ be Borel and $u \in \mathbb{R}$. We use the formula

$$\begin{aligned} &\mathbb{E}\,\#\big\{t \in B : \nabla Z(t) = 0,\ Z(t) \geq u,\ \mathcal{M}_t\big\} \\ &\qquad = \int_B \mathbb{E}\Big[\big|\det \nabla^2 Z(t)\big| \mathbf{1}\{Z(t) \geq u\} \mathbf{1}_{\mathcal{M}_t} \;\Big|\; \nabla Z(t) = 0\Big]\, p_{\nabla Z(t)}(0)\, \mathrm{d}t. \end{aligned} \tag{A.1}$$

With $\{Z(t) > u\}$ in place of $\{Z(t) \geq u\}$ and a mark of the first form, this is the Kac–Rice formula with a lower semi-continuous weight of Armentano, Azaïs, and León (2025, Theorem 6.1), an extension by monotone convergence of the weighted formula (Azaïs and Wschebor, 2009, Theorem 6.4). For a Gaussian field it requires only C^2 paths of Z and a density of $\nabla Z(t)$ continuous and bounded on compacts, both immediate here, and it counts all zeros, degenerate ones included. We apply it on the charts of a finite atlas, to the gradient of Z read in the chart. Given $\nabla Z(t) = 0$, the integrand of the theorem in a chart is $|\det \nabla^2 Z(t)|\, p_{\nabla Z(t)}(0)\, \mathrm{d}t$, whatever the chart: the Jacobian of the chart cancels between the determinant, the density and the volume. The pair $(\nabla Z(t), Z)$ is degenerate, so we check the condition of the theorem on the conditional law directly. Let Π_t be the orthogonal projection onto the kernel of the linear map $Y \mapsto \nabla Z(t)$. The covariance of Y is $\sigma^2 \mathrm{Id}_m$. So, by Gaussian regression, the law of Y given $\nabla Z(t) = x$ is the law of $\Pi_t Y + y_x$, where y_x is linear in x. This law is weakly continuous in x. The path of Z is linear in Y, so its conditional law is weakly continuous in x too, as the theorem asks. At $x = 0$ it is the conditional law of (A.1) and of Lemma 2 of the main paper. The weight $\mathbf{1}\{Z(t) > u\}\mathbf{1}\{f_t \neq 0\}$ is the indicator of an open subset of $M \times \mathbb{R}^m$. A homogeneous polynomial that vanishes on an open subset of the sphere vanishes on an open cone, hence everywhere. So, as in Section 2.3 of the main paper, Y is a linear function of the values of Z at m points of the chart. Hence this weight is lower semi-continuous in t and in the path for the weak topology of the continuous functions on the chart domain, for which point evaluations are continuous.

The closed conditions are reached by dominated convergence. Let N_Z be the number of critical points of Z on M. The theorem with the weight 1 gives $\mathbb{E}N_Z = \int_M D(t)\,\mathrm{d}t$, where $D(t) = \mathbb{E}\big[|\det\nabla^2 Z(t)| \mid \nabla Z(t) = 0\big]\, p_{\nabla Z(t)}(0)$. The function D is bounded on M. Indeed, for a unit vector $v \perp t$, two derivatives of $\theta \mapsto Z(\cos\theta\, t + \sin\theta\, v)$ at 0 give $v^\top \nabla^2 Z(t) v = -kZ(t) + k(k-1)\langle Y, t^{\otimes(k-2)} \otimes v \otimes v\rangle$. So $|v^\top \nabla^2 Z(t) v| \le k^2\|Y\|$ and $|\det\nabla^2 Z(t)| \le k^{2(d-1)}\|Y\|^{d-1}$. Given $\nabla Z(t) = 0$, Y has the law of $\Pi_t Y$, and $\|\Pi_t Y\| \le \|Y\|$. So $\mathbb{E}\big[\|Y\|^{d-1} \mid \nabla Z(t) = 0\big] \le \mathbb{E}\|Y\|^{d-1}$. Also $p_{\nabla Z(t)}(0) \le (2\pi k\sigma^2)^{-(d-1)/2}$, as $\nabla Z(t)$ has the covariance $k\sigma^2 \mathrm{Id}_{d-1}$. Hence, for every $t \in M$,

$$D(t) \;\le\; k^{2(d-1)}\, \mathbb{E}\|Y\|^{d-1}\, (2\pi k\sigma^2)^{-(d-1)/2},$$

and $\mathbb{E}N_Z < \infty$. Let $\epsilon > 0$. Put $\mathcal{M}_t^\epsilon = \{|f_t| < \epsilon\}$ if $\mathcal{M}_t = \{f_t = 0\}$, and $\mathcal{M}_t^\epsilon = \mathcal{M}_t$ if $\mathcal{M}_t = \{f_t \neq 0\}$. In both cases $\mathcal{M}_t^\epsilon$ is a mark of the first form, as $\{|f_t| < \epsilon\}$ is the event that $(\epsilon - |f_t|)^+ \neq 0$. So the theorem gives (A.1) with $\{Z(t) > u - \epsilon\}$ and $\mathcal{M}_t^\epsilon$ in place of $\{Z(t) \ge u\}$ and $\mathcal{M}_t$. As ϵ decreases to 0, these two conditions decrease to $\{Z(t) \ge u\}$ and $\mathcal{M}_t$. On the left-hand side, the set of critical points that is counted decreases to the set counted in (A.1). It is finite when N_Z is finite, so the count converges almost surely, and it is at most N_Z, which is integrable. On the right-hand side, the variable under the conditional expectation decreases to that of (A.1). It is at most $|\det\nabla^2 Z(t)|$, which is integrable given $\nabla Z(t) = 0$, and the integrand in t is at most $D(t)$. Both sides converge by dominated convergence. So (A.1) holds for the two forms of marks. Lemma 2 of the main paper uses it with a closed condition.

In our model the maximum of Z is almost surely attained at a single point. Indeed the drift $\lambda\langle e, \cdot\rangle^k = \lambda\mathrm{Cov}(X(\cdot), X(e))$ is λ times the reproducing kernel of X at e, hence an element of its Cameron–Martin space, so the law of Z is equivalent to that of σX; and for k odd $\mathrm{Var}(X(s) - X(t)) = 2(1 - \langle s, t\rangle^k)$ vanishes only at $s = t$, so the maximum of X is attained at a single point (Lifshits, 1983, Theorem 3). Every almost sure property of X transfers to Z, in particular the non-degeneracy of the critical points and $\mathbb{P}\{\sup Z = u\} = 0$.

B The layer decomposition and the drift

Fix t and write $s = \cos\theta\, t + \sin\theta\, v$ with $v \in \mathbb{S}^{d-2} \subset t^\perp$ and $r = \cos\theta$, and set $M_t = \mathbb{S}^{d-1} \setminus \{\pm t\}$. Standard computations give $\mathrm{Cov}(X(s), X(t)) = r^k$, $\mathrm{Cov}(X(s), \partial_a X(t)) = k\, r^{k-1} \sin\theta\, v_a$, $\mathrm{Cov}(X(t), \nabla X(t)) = 0$ and $\mathrm{Cov}(\nabla X(t)) = k\,\mathrm{Id}_{d-1}$, so the Gaussian regression of $X(s)$ on $(X(t), \nabla X(t))$ reads

$$X(s) = r^k X(t) + r^{k-1} \sin\theta\, \langle v, \nabla X(t)\rangle + R^t(s), \tag{B.1a}$$

and the helix process is $\zeta^t(s) = R^t(s)/(1 - r^k)$ on M_t. The matrix

$$G := \frac{\nabla^2 X(t) + k\, X(t)\, \mathrm{Id}_{d-1}}{\sqrt{2k(k-1)}} \;\sim\; \mathrm{GOE}(d-1) \tag{B.1b}$$

is independent of $(X(t), \nabla X(t))$ and of the layers of index $j \ge 3$ below. The binomial expansion of $\langle s, \cdot\rangle^k$ and the regression on $(Z(t), \nabla Z(t))$ give the *layer decomposition* of the helix: with $w_j(\theta) = \binom{k}{j}^{1/2} r^{k-j} \sin^j\theta/(1 - r^k)$,

$$\zeta^t(s) = \sum_{j=2}^{k} w_j(\theta)\, Z_j(v), \tag{B.1c}$$

where $Z_2(v) = v^\top G v$ with $G \sim \mathrm{GOE}(d-1)$ and, for $j \ge 3$, Z_j is an independent unit-variance Kostlan–Shub–Smale field of degree j on $\mathbb{S}^{d-2}$; all of them are independent of $(Z(t), \nabla Z(t))$.

Under the alternative the same computation, applied to $\lambda\langle e, \cdot\rangle^k$ with $c = \langle e, t\rangle$, replaces each layer by a shifted one: the terms $j = 0$ and $j = 1$ of the expansion are exactly what the regression removes, and

$$\mu_t(s) = \sum_{j=2}^{k} w_j(\theta)\, \beta_j(c)\, \langle e_\perp, v\rangle^j, \qquad \beta_j(c) = \lambda \binom{k}{j}^{1/2} c^{k-j}(1 - c^2)^{j/2}, \tag{B.2a}$$

an identity for every k. Cauchy–Schwarz, $\sum_j w_j^2 \le \rho^{-2} = 2(k-1)/k$, $\binom{k}{j} \le \binom{k}{2}\binom{k-2}{j-2}$ for $j \ge 2$ and $\binom{k}{2}^{1/2}/\rho = k-1$ give the sharp drift bound

$$\sup_{M_t} |\mu_t| \;\le\; (k-1)\,\lambda\,(1-c^2). \tag{B.2b}$$

At the pole the drift is $\beta_2(c)\langle e_\perp, v\rangle^2/\rho$, whence $G_\lambda = G + \beta_2(c) e_\perp e_\perp^\top$ in (2b) of the main paper. The layer decomposition is proved in Azaïs et al. (2026); Lemmas B.1 and B.2 below are quoted from there without proof.

Let $\overline{M} = [0,\pi] \times \mathbb{S}^{d-2}$ be the blow-up of the sphere at $\pm t$, whose boundary components $\{0\} \times \mathbb{S}^{d-2}$ and $\{\pi\} \times \mathbb{S}^{d-2}$ are the pole fibres.

Lemma B.1 (Variance; see Azaïs et al. (2026)). *$V(\theta) := \mathrm{Var}\zeta^t(s) = \|w(\theta)\|^2 = \dfrac{1 - r^{2k} - k\,r^{2k-2}(1-r^2)}{(1-r^k)^2} \le \rho^{-2} = \dfrac{2(k-1)}{k}$ on $(0,\pi)$, with equality only in the limit $\theta \to 0$ when k is odd.*

Lemma B.2 (Poles; see Azaïs et al. (2026)). *(i) Each w_j is C^∞ on $[0,\pi)$, and as $\theta \to 0$*

$$w_j(\theta) = \tfrac{2}{k}\binom{k}{j}^{1/2}\theta^{j-2}\big(1 + O(\theta^2)\big), \qquad w_2(0) = \rho^{-1}.$$

(ii) ζ^t extends to an almost surely continuous field $\tilde\zeta(\theta, v) = \sum_j w_j(\theta) Z_j(v)$ on $\overline{M}$, with $\tilde\zeta(0,v) = v^\top G v/\rho$, and $\tilde\zeta(\pi,\cdot) \equiv 0$ for k odd.

The last lemma of this appendix is not taken from Azaïs et al. (2026). The proof of Lemma 3 of the main paper uses it. Its own proof uses Lemma 4 of the main paper, which is stated after Lemma 3. The proof of Lemma 4, in Appendix D, uses neither Lemma 2 nor Lemma 3 of the main paper. For a general field, Azaïs et al. (2024, Appendix C) write that the regularity of λ_2^t as a function of t is difficult to control. Here the model has finite rank. Lemma B.3 obtains the continuity of λ_2^t in (t, Y) from the finite sum of Lemma 4 of the main paper.

Lemma B.3 (Continuity of the second maximum at a point). *Let k be odd. The maps $(t, Y) \mapsto \lambda_2^t$ and $(t, Y) \mapsto \lambda_{\max}(\Omega_t)$ are continuous on $M \times \mathbb{R}^m$.*

Proof. For $t \in M$, $v \in \mathbb{R}^d$ and $2 \le j \le k$ let $\Psi_j^t(v) = \binom{k}{j}^{1/2} \mathrm{Sym}\left(t^{\otimes(k-j)} \otimes v^{\otimes j}\right)$. For a unit vector $v \perp t$ this is the tensor $\Psi_j(v)$ of Lemma 4 of the main paper at the point t. Its coordinates are polynomials in (t, v). Let S_t be the set of the unit vectors $v \perp t$. For $t \in M$, $\theta \in [0,\pi]$, $v \in \mathbb{R}^d$ and $Y \in \mathbb{R}^m$ let

$$F(t, \theta, v, Y) = \sum_{j=2}^{k} w_j(\theta)\,\big\langle Y, \Psi_j^t(v)\big\rangle. \tag{B.3a}$$

The weights w_j are continuous on $[0,\pi]$, with $w_2(0) = \rho^{-1}$, $w_j(0) = 0$ for $j \ge 3$ and $w_j(\pi) = 0$. On $[0,\pi)$ this is Lemma B.2(i). At π the denominator $1 - \cos^k\theta$ tends to 2, as k is odd, and $\sin^j\theta$ tends to 0. So F is continuous in (t, θ, v, Y). By Lemma 4 of the main paper, for every Y, $F(t, \theta, v, Y) = Z^{|t}(\cos\theta\, t + \sin\theta\, v)$ for $(\theta, v) \in (0,\pi) \times S_t$, and $F(t, 0, v, Y) = v^\top \Omega_t v$ for $v \in S_t$. The points $\cos\theta\, t + \sin\theta\, v$ with $(\theta, v) \in (0,\pi) \times S_t$ are the points of M_t, and $(0,\pi) \times S_t$ is dense in the compact set $[0,\pi] \times S_t$. So

$$\lambda_2^t = \max_{(\theta,v) \in [0,\pi] \times S_t} F(t, \theta, v, Y), \qquad \lambda_{\max}(\Omega_t) = \max_{v \in S_t} F(t, 0, v, Y). \tag{B.3b}$$

The sets in (B.3b) depend on t. Fix $t_0 \in M$ and let $U = \{t \in M : \langle t_0, t\rangle > -1\}$, an open neighbourhood of t_0. For $t \in U$ let

$$Q_t = \Big(\mathrm{Id} - \frac{(t_0 + t)(t_0 + t)^\top}{1 + \langle t_0, t\rangle}\Big)\big(\mathrm{Id} - 2\, t_0 t_0^\top\big). \tag{B.3c}$$

Both factors are reflections, the first one because $\|t_0+t\|^2 = 2(1+\langle t_0, t\rangle)$. The second maps t_0 to $-t_0$, and the first maps $-t_0$ to t. So Q_t is orthogonal, $Q_t t_0 = t$, and Q_t maps S_{t_0} onto S_t. It is continuous in $t \in U$. By (B.3b), for $t \in U$,

$$\lambda_2^t = \max_{(\theta,v)\in[0,\pi]\times S_{t_0}} F(t,\theta,Q_t v, Y), \qquad \lambda_{\max}(\Omega_t) = \max_{v\in S_{t_0}} F(t,0,Q_t v, Y). \tag{B.3d}$$

The function $g(t,Y,\theta,v) = F(t,\theta,Q_t v,Y)$ is continuous on $U\times\mathbb{R}^m\times K$, where $K=[0,\pi]\times S_{t_0}$ is compact and does not depend on (t,Y). Let $(t_n,Y_n)\to(t,Y)$ in $U\times\mathbb{R}^m$. The set $\big(\{(t_n,Y_n): n\ge 1\}\cup\{(t,Y)\}\big)\times K$ is compact, so g is uniformly continuous on it, and $g(t_n,Y_n,\cdot)\to g(t,Y,\cdot)$ uniformly on K. The maxima of two functions over a set differ by at most the supremum of their difference. Hence the two maxima in (B.3d) at (t_n,Y_n) converge to the maxima at (t,Y). So both maps are continuous on $U\times\mathbb{R}^m$, and t_0 is arbitrary. □

C The Gaussian Orthogonal Ensemble, and the count outside a cap

Lemma C.1 (Determinant of a rank-one spiked GOE). *Let $G\sim \mathrm{GOE}(N)$, let q be a unit vector of $\mathbb{R}^N$ and $\beta\in\mathbb{R}$, and write $D_n(z) = \mathbb{E}\det(z\,\mathrm{Id}_n - G_n)$ for $G_n\sim\mathrm{GOE}(n)$, $D_0=1$. Then*

$$\mathbb{E}\det\big(z\,\mathrm{Id}_N - G - \beta\, qq^\top\big) = D_N(z) - \beta\, D_{N-1}(z) \qquad \text{for every } z\in\mathbb{R}, \tag{C.1a}$$

and $\frac34 z^n \le D_n(z) \le z^n$ for $z\ge n$, so that

$$\mathbb{E}\big|\det(z\,\mathrm{Id}_N - G - \beta qq^\top)\big| \;\ge\; \tfrac12 z^N \qquad \text{for } z\ge N,\ |\beta|\le z/4. \tag{C.1b}$$

Proof. The law of G is invariant under $G\mapsto R^\top G R$, $R\in O(N)$. Picking R with $Re_1 = q$ gives $R^\top(qq^\top)R = e_1e_1^\top$, so one may take $q=e_1$. The matrix determinant lemma $\det(M-\beta e_1e_1^\top) = \det M - \beta\, e_1^\top \mathrm{adj}(M) e_1$ with $M = z\,\mathrm{Id}-G$ gives $\det(z\,\mathrm{Id}-G-\beta e_1e_1^\top) = \det(z\,\mathrm{Id}-G) - \beta\det(z\,\mathrm{Id}_{N-1}-G')$, the $(1,1)$ entry of the adjugate being the corresponding principal minor; $G' = (G_{ij})_{2\le i,j\le N}$ is a $\mathrm{GOE}(N-1)$ with the same normalisation, whence the identity. For the bracket, $\mathbb{E}\det G' = \sum_\pi \mathrm{sgn}(\pi)\mathbb{E}\prod_i G'_{i\pi(i)}$ survives only for π a product of disjoint transpositions, which gives $D_n(z) = \sum_{i\le n/2}(-1)^i t_i$ with $t_i = \frac{n!}{i!(n-2i)!4^i} z^{n-2i} \ge 0$ and $t_{i+1}/t_i = \frac{(n-2i)(n-2i-1)}{4(i+1)z^2} \le \frac{n^2}{4z^2} \le \frac14$ for $z\ge n$; an alternating series with decreasing terms is bracketed by its first two partial sums. Hence $D_N - \beta D_{N-1} \ge \frac34 z^N - \frac{z}{4}z^{N-1} = \frac12 z^N$ for $0\le\beta\le z/4$, while for $\beta\le 0$ the subtracted term is non-positive and the bound is $\frac34 z^N$; and $\mathbb{E}|\det| \ge \mathbb{E}\det$. □

We also use, from Azaïs et al. (2026), the Gaussian concentration of the GOE norm, $\mathbb{P}\{\|G\| > \sqrt{2N}+s\} \le 2e^{-s^2/2}$ for $s\ge 0$. It gives the moment bound $\|\,\|G\|\,\|_q \le \sqrt{2N} + \frac56\sqrt{q}$ for $q\ge 8$: integrating the tail, $\mathbb{E}(\|G\|-\sqrt{2N})_+^q \le 2q\int_0^\infty s^{q-1}e^{-s^2/2}\,\mathrm{d}s = q\,2^{q/2}\Gamma(q/2)$, whose q-th root is at most $\frac56\sqrt{q}$ for $q\ge 8$ (2.29 against 2.36 at $q=8$, and the ratio of the two sides decreases in q), and Minkowski's inequality in L^q gives the bound.

Lemma C.2 (Upper bound on the count outside a cap). *Let $\theta_1\in(0,\pi/2)$, let $\beta_\star := \lambda\binom{k}{2}^{1/2}\sin^2\theta_1 = \rho(k-1)\lambda\sin^2\theta_1$, let u satisfy $(d-1)/\rho < u < \lambda\cos^k\theta_1$ and $\rho u \ge \max\{4\beta_\star,\ \beta_\star+\sqrt{2N}\}$, and set*

$$\varrho = 1 - \inf_{0<\theta\le\theta_1} \cos^{2k-2}\theta\Big(\frac{\sin\theta}{\theta}\Big)^2. \tag{C.2a}$$

Then, for every $b\in(0,\lambda^2]$,

$$\delta^\lambda\big(u;\mathcal{B}_b^c\big) \;\le\; \frac{1+\epsilon_A}{(1-\varrho)^{\frac{d-1}{2}}}\,\mathbb{P}\big\{\chi^2_{d-1} > k\,b\,(1-\varrho)\big\} \;+\; \mathrm{Rem}_{\mathrm{far}}, \tag{C.2b}$$

where $\mathrm{Rem}_{\mathrm{far}} = \delta^{\lambda}(u;\{\theta>\theta_1\})$ *and* $1+\epsilon_A = \big(1+\beta_\star/(\rho u)+\varepsilon_{\det}\big)(1+r_N)$*, the three remainders being, with* $q_\star = \max(2N,8)$ *and* $\ell_1 = \lambda\cos^k\theta_1$*,*

$$r_N \le e^{N^2/(2\ell_1^2)}-1, \qquad \varepsilon_{\det} \le 2\sqrt{2}\Big(\tfrac{5}{4}+\frac{\sqrt{2N}+\frac{5}{6}\sqrt{q_\star}}{N}\Big)^N e^{-\frac14(\rho u-\beta_\star-\sqrt{2N})^2}. \tag{C.2c}$$

Proof. Three steps. *The determinant.* On the event that all eigenvalues of G_λ lie below z the determinant $\det(z\,\mathrm{Id}-G_\lambda)$ is positive, so $A_\lambda(z;c) = \mathbb{E}\det(z\,\mathrm{Id}-G_\lambda)+2\mathbb{E}[(\det)^-]$, and $\mathbb{E}\det \le D_N(z)+|\beta_2(c)|D_{N-1}(z) \le z^N(1+|\beta_2(c)|/z)$ by Lemma C.1. For the negative part, Cauchy–Schwarz and $\|\det(z\,\mathrm{Id}-G_\lambda)\|_2 \le (z+|\beta_2(c)|+\|\,\|G\|\,\|_{q_\star})^N$ with $\|\,\|G\|\,\|_{q_\star} \le \sqrt{2N}+\frac56\sqrt{q_\star}$ give

$$2\mathbb{E}[(\det)^-] \le 2\big(z+\tfrac{z}{4}+\sqrt{2N}+\tfrac56\sqrt{q_\star}\big)^N\sqrt{2}\,e^{-\frac14(z-|\beta_2(c)|-\sqrt{2N})^2} \le \varepsilon_{\det}\,z^N,$$

using $|\beta_2(c)| \le z/4$ and $z \ge N$. *The Gaussian integral.* Extending the range to $\mathbb{R}$ and expanding, $\int_{\mathbb{R}} x^N\varphi(x-\ell_0)\,\mathrm{d}x = \mathbb{E}(\ell_0+\xi)^N = \ell_0^N(1+r_N)$ with $\ell_0 = \lambda c^k$, the odd moments vanishing and

$$r_N = \sum_{i\ge1}\binom{N}{2i}(2i-1)!!\,\ell_0^{-2i} \le \sum_{i\ge1}\frac{1}{i!}\Big(\frac{N^2}{2\ell_0^2}\Big)^i = e^{N^2/(2\ell_0^2)}-1, \tag{C.3a}$$

since $\binom{N}{2i}(2i-1)!! \le N^{2i}/(2i)!! = N^{2i}/(2^i i!)$; hence

$$\int_u^\infty A_\lambda(\rho x;c)\varphi(x-\ell_0)\,\mathrm{d}x \le (\rho\ell_0)^N(1+\epsilon_A). \tag{C.3b}$$

The extension does not decrease the integral. For N even the integrand is non-negative. For N odd the contribution of $x<0$ is at least $-\varphi(\ell_0)N!/\ell_0^{N+1}$, since $\varphi(x+\ell_0) \le \varphi(\ell_0)e^{-x\ell_0}$ for $x\ge0$, while that of $(0,u)$ is at least $\varphi(\ell_0)u^{N+1}/(N+1)$, since $\ell_0 x - x^2/2 \ge 0$ for $0\le x\le u\le\ell_0$; the latter dominates because $u\ell_0 \ge u^2 > N^2$ by $\rho u > N$ and $\rho<1$, whence $(u\ell_0)^{N+1} \ge N^{2N+2} \ge (N+1)!$. *The Laplace step.* For $b \ge \lambda^2\sin^2\theta_1$, $\mathcal{B}_b^c \subseteq \{\theta>\theta_1\}$ and the bound holds with $\mathrm{Rem}_{\mathrm{far}}$ alone. For $b<\lambda^2\sin^2\theta_1$, in the variable θ, on $(\theta_b,\theta_1]$ with $\sin\theta_b = \sqrt{b}/\lambda$, use $\sin^{d-2}\theta \le \theta^{d-2}$, $\cos^{kN}\theta\le1$ and $\frac{k\lambda^2}{2}\cos^{2k-2}\theta\sin^2\theta = \frac{w^2}{2}q(\theta)$ with $q\ge1-\varrho$ and $w=\lambda\sqrt{k}\,\theta$, so that

$$\int_{\theta_b}^{\theta_1} \le \frac{1}{(\lambda\sqrt{k})^{d-1}}\int_{\sqrt{kb}}^\infty w^{d-2}e^{-\frac{w^2}{2}(1-\varrho)}\,\mathrm{d}w. \tag{C.4a}$$

The prefactor normalises exactly:

$$\omega_{d-2}\big(\tfrac{k-1}{\pi}\big)^{\frac{d-1}{2}}\rho^{d-1}k^{-\frac{d-1}{2}}\int_0^\infty w^{d-2}e^{-w^2/2}\,\mathrm{d}w = 2^{\frac{d-1}{2}}\rho^{d-1}\big(\tfrac{k-1}{k}\big)^{\frac{d-1}{2}} = 1, \tag{C.4b}$$

because $\rho^2 = k/(2(k-1))$. Rescaling $w\mapsto w\sqrt{1-\varrho}$ turns the truncated integral into $(1-\varrho)^{-(d-1)/2}\mathbb{P}\{\chi^2_{d-1}>kb(1-\varrho)\}$; the range $\theta>\theta_1$ is $\mathrm{Rem}_{\mathrm{far}}$. □

Lemma C.3 (Simple bounds)**.** *For* $\theta_1\in(0,\pi/2)$*, the quantity* ϱ *of Lemma C.2 satisfies*

$$\varrho = 1-\cos^{2k-2}\theta_1\Big(\frac{\sin\theta_1}{\theta_1}\Big)^2 \;\le\; \Big(k-\frac23\Big)\theta_1^2. \tag{C.5a}$$

Under the hypotheses of Lemma C.2,

$$1+\epsilon_A \le \Big(1+\frac{(k-1)\lambda\sin^2\theta_1}{u}+\varepsilon_{\det}\Big)e^{\frac{N^2}{2\ell_1^2}}, \qquad \varepsilon_{\det} \le 3\,e^{\frac54 N-\frac14(\rho u-\beta_\star-\sqrt{2N})^2}. \tag{C.5b}$$

If also $\rho u \ge 4\sqrt{2N}$*, then* $\varepsilon_{\det} \le 3e^{\frac54 N-\frac{\rho^2u^2}{16}} \le 3e^{-\frac34 N}$ *and*

$$\epsilon_A \le \frac{(k-1)\lambda\sin^2\theta_1}{u}+\frac{3N^2}{2\ell_1^2}+3e^{\frac54 N-\frac{\rho^2u^2}{16}}. \tag{C.5c}$$

Proof. On $(0,\pi/2)$ the functions $\cos^{2k-2}\theta$ and $(\sin\theta/\theta)^2$ are positive and decreasing, so the infimum that defines ϱ is attained at θ_1. Next, $\sin x \geq x\cos^{1/3} x$ on $[0,\pi/2)$. Indeed, with $c=\cos x$ and $f(x) = c^{-1/3}\sin x - x$, one has $f(0)=0$ and $f'(x) = \frac{2}{3}c^{2/3} + \frac{1}{3}c^{-4/3} - 1 \geq 0$, by the weighted arithmetic-geometric inequality and $(c^{2/3})^{2/3}(c^{-4/3})^{1/3} = 1$. Hence $1-\varrho \geq \cos^p\theta_1$ with $p = 2k - \frac{4}{3} \geq 1$. With $1-c^p \leq p(1-c)$ for $c\in[0,1]$ and $1-\cos\theta_1 \leq \theta_1^2/2$, this gives $\varrho \leq (k-\frac{2}{3})\theta_1^2$.

By Lemma C.2, $1+\epsilon_A = (1+a+\varepsilon_{\mathrm{det}})(1+r_N)$ with $a = \beta_\star/(\rho u) = (k-1)\lambda\sin^2\theta_1/u$ and $1+r_N \leq e^{N^2/(2\ell_1^2)}$. The constant in the bound of Lemma C.2 on $\varepsilon_{\mathrm{det}}$ is $2\sqrt{2}B_N^N$ with $B_N = \frac{5}{4} + m_N/N$ and $m_N = \sqrt{2N} + \frac{5}{6}\sqrt{q_\star}$. One has $B_2 = \frac{9}{4} + \frac{5}{6}\sqrt{2} < 3.43$, $B_3 < 2.86$ and, for $N \geq 4$, where $q_\star = 2N$, $B_N = \frac{5}{4} + \frac{11}{6}\sqrt{2/N} < 2.55$. Since $e^{5/4} > 3.49$, $2\sqrt{2}B_N^N \leq 3e^{5N/4}$ for every $N \geq 2$. This proves the second display.

Let now $\rho u \geq 4\sqrt{2N}$. With $\rho u \geq 4\beta_\star$, the gap $g = \rho u - \beta_\star - \sqrt{2N}$ is at least $\rho u/2$, so $g^2/4 \geq \rho^2 u^2/16 \geq 2N$. This gives the two bounds on $\varepsilon_{\mathrm{det}}$, and $\varepsilon_{\mathrm{det}} \leq 3e^{-3/2} < 0.67$. Put $x = N^2/(2\ell_1^2)$. Since $\ell_1 > u > N/\rho$ and $\rho^2 \leq \frac{3}{4}$, $x < \rho^2/2 \leq \frac{3}{8}$, so $r_N \leq e^x - 1 \leq xe^x \leq e^{3/8}x < 0.73\, N^2/\ell_1^2$. Also $a \leq \frac{1}{4}$, so $1+a+\varepsilon_{\mathrm{det}} < 1.92$, and $\epsilon_A = a + \varepsilon_{\mathrm{det}} + r_N(1+a+\varepsilon_{\mathrm{det}}) \leq a + \varepsilon_{\mathrm{det}} + 1.41\, N^2/\ell_1^2$, which gives the last display. □

Lemma C.4 (The far zone)**.** *Let θ_1 and u satisfy the hypotheses of Lemma C.2, and let $3\rho u \geq 4\sqrt{2N}$. Define $\theta_3 \in (0,\pi/2)$ by $\sin^2\theta_3 = u/(4(k-1)\lambda)$, and let $\ell_3 = \lambda\cos^k\theta_3$, $h(\theta) = \cos^{2k-2}\theta\sin^2\theta$, $m_N = \sqrt{2N} + \frac{5}{6}\sqrt{q_\star}$, $K_N = 2\sqrt{2}\,(\frac{5}{4} + m_N/N)^N$, and $\omega_{d-1} = 2\pi^{d/2}/\Gamma(\frac{d}{2})$, the area of $\mathbb{S}^{d-1}$. Put*

$$C_M = \Big(\frac{5}{4} + K_N e^{-\frac{1}{4}(\frac{3}{4}\rho u - \sqrt{2N})^2}\Big)\frac{e^{N(N+1)/(2\ell_3^2)}}{1-(k-1)\tan^2\theta_3}, \tag{C.6a}$$

$$F_{\mathrm{far}} = \omega_{d-1}\Big(\frac{k-1}{\pi}\Big)^{\frac{N}{2}}\Big(1+\frac{(k-1)\lambda}{u}\Big)\big(\rho\lambda + \rho\sqrt{N} + m_N\big)^N e^{-E_{\mathrm{far}}}, \tag{C.6b}$$

$$E_{\mathrm{far}} = \sup_{\theta_3 \leq \theta_s \leq \pi/2} \min\Big\{\frac{k\lambda^2}{2}\min\big(h(\theta_3), h(\theta_s)\big),\ \frac{1}{2}\big(u - \lambda\cos^k\theta_s\big)_+^2\Big\}. \tag{C.6c}$$

Then $\theta_1 \leq \theta_3$, $(k-1)\tan^2\theta_3 \leq (k-1)/(4k-5)$, and

$$\mathrm{Rem}_{\mathrm{far}} = \delta^\lambda\big(u;\{\theta>\theta_1\}\big) \ \leq\ C_M\,\mathbb{P}\big\{\chi_N^2 > k\lambda^2 h(\theta_1)\big\} + F_{\mathrm{far}}. \tag{C.6d}$$

If also $u \geq 3\lambda/4$, then $E_{\mathrm{far}} \geq \lambda^2/32$ and $\cos^k\theta_3 \geq 13/16$.

Proof. The hypothesis $\rho u \geq 4\beta_\star$ of Lemma C.2 reads $\sin^2\theta_1 \leq \sin^2\theta_3$, so $\theta_1 \leq \theta_3$. Since $u < \lambda$, $s_3 := \sin^2\theta_3 < 1/(4(k-1))$, and $(k-1)\tan^2\theta_3 = (k-1)s_3/(1-s_3) \leq (k-1)/(4k-5) < 1$.

First, for $\ell > 0$ and $\xi \sim N(0,1)$, $\mathbb{E}|\ell+\xi|^N \leq \ell^N e^{N(N+1)/(2\ell^2)}$. For N even, $\mathbb{E}(\ell+\xi)^N = \ell^N(1+r_N) \leq \ell^N e^{N^2/(2\ell^2)}$ by the computation in the proof of Lemma C.2, which holds for every $\ell > 0$. For N odd, Hölder's inequality gives $\mathbb{E}|\ell+\xi|^N \leq (\mathbb{E}(\ell+\xi)^{N+1})^{N/(N+1)} \leq \ell^N e^{N(N+1)/(2\ell^2)}$.

Second, for $z > 0$ and every c, $A_\lambda(z;c) \leq (z+m_N)^N(1+|\beta_2(c)|/z)$. The matrix determinant lemma in the proof of Lemma C.1 gives

$$\det(z\,\mathrm{Id} - G_\lambda) = \det(z\,\mathrm{Id} - G) - \beta_2(c)\det(z\,\mathrm{Id}_{N-1} - G'), \tag{C.7a}$$

with G' a principal submatrix of G, so that $\|G'\| \leq \|G\|$. For $n \in \{N-1, N\}$ and G_n the matrix of size n among G' and G, Minkowski's inequality gives $\mathbb{E}|\det(z\,\mathrm{Id}_n - G_n)| \leq (z+m_N)^n$, since $n \leq q_\star$ and $\|\,\|G\|\,\|_{q_\star} \leq m_N$. Add the two terms and use $(z+m_N)^{N-1} \leq (z+m_N)^N/z$.

Third, let $(k-1)\tan^2\theta_{\max} = 1$. On $(0,\theta_{\max})$ the map $w(\theta) = \lambda\sqrt{k}\cos^{k-1}\theta\sin\theta$ is increasing, with $w'(\theta) = \lambda\sqrt{k}\cos^k\theta\,(1-(k-1)\tan^2\theta)$ and $w^2 = k\lambda^2 h(\theta)$. Also $\sin^{N-1}\theta\cos^{kN}\theta = (w/(\lambda\sqrt{k}))^{N-1}\cos^{N+k-1}\theta$, and the normalisation in the proof of Lemma C.2 reads $\omega_{d-2}((k-1)/\pi)^{N/2}\rho^N k^{-N/2} = 1/(2^{N/2-1}\Gamma(N/2))$. This gives the exact identity

$$\omega_{d-2}\Big(\frac{k-1}{\pi}\Big)^{\frac{N}{2}}\sin^{N-1}\theta\,(\rho\lambda\cos^k\theta)^N e^{-\frac{k\lambda^2}{2}h(\theta)}\,\mathrm{d}\theta = f_{\chi_N}(w)\,\frac{\cos^{N-1}\theta}{1-(k-1)\tan^2\theta}\,\mathrm{d}w, \tag{C.7b}$$

where f_{χ_N} is the density of the norm of a standard Gaussian vector of $\mathbb{R}^N$.

In the variable θ, with $c = \cos\theta$, the density of δ^λ is

$$\omega_{d-2}\Big(\frac{k-1}{\pi}\Big)^{\frac{N}{2}} \sin^{N-1}\theta\, e^{-\frac{k\lambda^2}{2}h(\theta)} \int_u^\infty A_\lambda(\rho x; c)\varphi(x-\lambda c^k)\,\mathrm{d}x.$$

Split $\{\theta > \theta_1\}$ into $(\theta_1, \theta_3]$ and $(\theta_3, \pi]$.

On $(\theta_1, \theta_3]$, $c \in (0,1)$ and $|\beta_2(c)| \le \rho(k-1)\lambda\sin^2\theta_3 = \rho u/4$. For $x \ge u$, $z = \rho x \ge \rho u > N$, $|\beta_2(c)| \le z/4$ and $z - |\beta_2(c)| - \sqrt{2N} \ge \frac{3}{4}\rho u - \sqrt{2N} \ge 0$. The determinant step of the proof of Lemma C.2, run with these bounds, gives $A_\lambda(\rho x; c) \le (\rho x)^N(\frac{5}{4} + \varepsilon_3)$, with $\varepsilon_3 = K_N e^{-\frac{1}{4}(\frac{3}{4}\rho u - \sqrt{2N})^2}$. The first fact with $\ell = \lambda c^k \ge \ell_3$ gives $\int_u^\infty A_\lambda(\rho x; c)\varphi(x-\lambda c^k)\,\mathrm{d}x \le (\frac{5}{4} + \varepsilon_3)e^{N(N+1)/(2\ell_3^2)}(\rho\lambda\cos^k\theta)^N$. The interval lies in $(0, \theta_{\max})$, and there the factor $\cos^{N-1}\theta/(1-(k-1)\tan^2\theta)$ of the third fact is at most $1/(1-(k-1)\tan^2\theta_3)$. By the third fact, this zone contributes at most $C_M \int_{w(\theta_1)}^\infty f_{\chi_N} = C_M\,\mathbb{P}\{\chi_N^2 > k\lambda^2 h(\theta_1)\}$.

On $(\theta_3, \pi]$, the second fact with $|\beta_2(c)| \le \rho(k-1)\lambda$ and $z = \rho x \ge \rho u$ gives $A_\lambda(\rho x; c) \le (1 + (k-1)\lambda/u)(\rho x + m_N)^N$ for $x \ge u$. Let $\ell = \lambda c^k \le \lambda$. Since $\mathbb{E}\xi^{2j} = (2j-1)!! \le (2j-1)^j$, Lyapunov's inequality with $2j \in \{N, N+1\}$ gives $\|\xi\|_N \le \sqrt{N}$. If $\ell \ge u$, Minkowski's inequality gives $\int_u^\infty (\rho x + m_N)^N\varphi(x-\ell)\,\mathrm{d}x \le \mathbb{E}(\rho|\ell+\xi| + m_N)^N \le (\rho\lambda + \rho\sqrt{N} + m_N)^N$. If $\ell < u$, put $x = u + y$ and $g = u - \ell$. Then $\varphi(g+y) \le e^{-g^2/2}\varphi(y)$ for $y \ge 0$, and the integral is at most $e^{-g^2/2}\mathbb{E}(\rho u + m_N + \rho|\xi|)^N \le e^{-g^2/2}(\rho\lambda + \rho\sqrt{N} + m_N)^N$. So the integrand of δ^λ is at most $\omega_{d-2}((k-1)/\pi)^{N/2}(1 + (k-1)\lambda/u)(\rho\lambda + \rho\sqrt{N} + m_N)^N \sin^{N-1}\theta\, e^{-E(\theta)}$, with $E(\theta) = \frac{k\lambda^2}{2}h(\theta) + \frac{1}{2}(u - \lambda\cos^k\theta)_+^2$. Fix $\theta_s \in [\theta_3, \pi/2]$. On $(0, \pi/2)$ the sign of h' is that of $\cos^2\theta - (k-1)\sin^2\theta$, so h increases and then decreases, and $h \ge \min(h(\theta_3), h(\theta_s))$ on $[\theta_3, \theta_s]$. For $\theta > \theta_s$, $\cos^k\theta < \cos^k\theta_s$, k being odd, so $(u - \lambda\cos^k\theta)_+ \ge (u - \lambda\cos^k\theta_s)_+$. Hence $E(\theta) \ge E_{\mathrm{far}}$ on $(\theta_3, \pi]$, and $\omega_{d-2}\int_0^\pi \sin^{N-1}\theta\,\mathrm{d}\theta = \omega_{d-1}$ gives the bound F_{far} for this zone.

Let now $u \ge 3\lambda/4$. Then $s_3 \ge 3/(16(k-1))$, and $\cos^k\theta_3 = (1-s_3)^{k/2} \ge 1 - \frac{k}{2}s_3 \ge 1 - \frac{k}{8(k-1)} \ge \frac{13}{16}$. Also $h(\theta_3) = (1-s_3)^{k-1}s_3 \ge (1-(k-1)s_3)s_3 \ge \frac{3}{4}s_3 \ge 9/(64(k-1))$, so $\frac{k\lambda^2}{2}h(\theta_3) \ge 9\lambda^2/128$. Take θ_s with $\cos^k\theta_s = \frac{1}{2}$; it lies in $[\theta_3, \pi/2]$, since $\cos^k\theta_3 \ge \frac{13}{16}$. Then $u - \lambda/2 \ge \lambda/4$, and the second term in E_{far} is at least $\lambda^2/32$. With $x = 2\log 2/k \le 0.47$, $\sin^2\theta_s = 1 - e^{-x} \ge x(1-x/2) \ge 0.76x$ and $\cos^{2k-2}\theta_s = 2^{-2(k-1)/k} \ge \frac{1}{4}$, so $\frac{k}{2}h(\theta_s) \ge 0.19\log 2 > \frac{1}{32}$. Since also $\frac{9}{128} > \frac{1}{32}$, $E_{\mathrm{far}} \ge \lambda^2/32$. □

The threshold $\lambda_0(k,d)$ of Theorem 4(iii) is defined in the main paper: it is the least number such that the five conditions (18a) to (18e) of the main paper hold at every $\lambda > \lambda_0(k,d)$. The constants K_N, m_N and ω_{d-1} are those of Lemma C.4. In the rest of this appendix, (18a) to (18e) are these five conditions, and $L = \log\lambda$. Condition (18a) fails at $\lambda = e$, so every number with this property is at least e; the least one exists, since the set of these numbers contains its infimum. For $\lambda \ge e$ each condition, once true, stays true for larger λ. Indeed $\lambda^2/\log\lambda$ increases for $\lambda > \sqrt{e}$. In (18c), g_1 increases, and the logarithm of $K_N\lambda^2 e^{-g_1^2/4}$ has derivative $2/\lambda - \frac{87}{256}\rho g_1$, which is negative wherever (18c) holds, because there $\frac{87}{128}\rho\lambda g_1 \ge g_1^2 \ge 4\log(K_N\lambda^2) > 4$. The left side of the second inequality of (18d) decreases once $g_3 \ge 0$. In (18e) the derivative of the logarithm of the left side is at most $(N+4)/\lambda - \lambda/16 < 0$ for $\lambda > 4\sqrt{d+3}$. Each condition holds for λ large, the left side of the second inequality of (18d) tending to $\frac{5}{4}\cdot\frac{4k-5}{3k-4} \le \frac{7}{4}$. So $\lambda_0(k,d)$ is finite for every odd $k \ge 3$ and $d \ge 3$, and it is the largest of the five points beyond which the conditions hold. The values of λ_0 given after the theorem in the main paper are computed numerically from (18) of the main paper; at these pairs (18e) is the last condition to hold.

Proof of Theorem 4(iii) *of the main paper.* Let $\lambda > \lambda_0(k,d)$, so that $L > 1$. We apply part (ii) with $\theta_1 = 3\lambda^{-1}\sqrt{L/k}$ and $u = \lambda - 2\sqrt{2L}$. By (18a), $\sqrt{8L} \le \lambda/4$, so $u \ge 3\lambda/4$, $\lambda/u \le 4/3$ and $L/\lambda \le \lambda/128$. Also $k\lambda^2\theta_1^2 = 9L$, $\theta_1 \le 3/\sqrt{128k} < 1$, and $\bar\varrho := (k - \frac{2}{3})\theta_1^2 = 3(3k-2)L/(k\lambda^2) \le 9L/\lambda^2 \le 9/128$.

The hypotheses of (ii) hold. By (18b), $u \ge 3\lambda/4 > N/\rho$. Since $1 - \cos^k\theta \le k(1-\cos\theta) \le k\theta^2/2$, $\lambda - \lambda\cos^k\theta_1 \le 9L/(2\lambda) < \sqrt{8L}$, the last step because $\lambda \ge \sqrt{128L}$; so $u < \lambda\cos^k\theta_1$. Next $4(k-1)\lambda\sin^2\theta_1 \le 4(k-1)\lambda\theta_1^2 \le 36L/\lambda \le 36\lambda/128 < u$, and, by (18b), $(k-1)\lambda\sin^2\theta_1 + \sqrt{2N}/\rho \le 9\lambda/128 + 9\lambda/16 < 3\lambda/4 \le u$. These are also the hypotheses of Lemma C.2. Finally $\frac{3}{4}\rho u \ge \frac{9}{16}\rho\lambda \ge \sqrt{2N}$ by (18b), so Lemma C.4 applies.

The main term. By Lemma C.3, $\varrho \le \bar\varrho$, and with $-\log(1-y) \le y/(1-y)$ and $1/(1-\bar\varrho) \le 128/119 < 9/8$,

$$(1-\varrho)^{-\frac{d+1}{2}} \le \exp\Big(\frac{9}{16}(d+1)\bar\varrho\Big) = \exp\Big(\frac{27(3k-2)(d+1)L}{16k\lambda^2}\Big).$$

Next, $(k-1)\lambda\sin^2\theta_1/u \le 9(k-1)L/(k\lambda u) \le 12(k-1)L/(k\lambda^2)$. Also $\beta_\star \le 9\rho L/\lambda \le 9\rho\lambda/128$, so $\rho u - \beta_\star - \sqrt{2N} \ge g_1 \ge 0$, and Lemma C.2 and (18c) give $\varepsilon_{\det} \le K_N e^{-g_1^2/4} \le \lambda^{-2} \le u^{-2}$. With $\ell_1 > u$, Lemma C.3 gives $1+\epsilon_A \le \exp(12(k-1)L/(k\lambda^2) + (N^2+2)/(2u^2))$, and $(N^2+2)/(2u^2) \le \frac{8}{9}(N^2+2)/\lambda^2$. So the main term of (ii) is at most $\frac{d-1}{k\lambda^2}\exp((c_1 L + c_1')/\lambda^2)$.

The far zone. In Lemma C.4, $\ell_3 \ge 13\lambda/16$ and $\frac{3}{4}\rho u - \sqrt{2N} \ge g_3 \ge 0$, so $C_M \le (\frac{5}{4} + K_N e^{-g_3^2/4}) e^{128N(N+1)/(169\lambda^2)}(4k-5)/(3k-4) \le 2$ by (18d). Also $E_{\text{far}} \ge \lambda^2/32$ and $1+(k-1)\lambda/u \le 1+4(k-1)/3 = (4k-1)/3$, so $F_{\text{far}} \le \lambda^{-4}/10$ by (18e). Since $k\lambda^2 h(\theta_1) = k\lambda^2\theta_1^2(1-\varrho) \ge 9L(1-9/128) \ge 8L$, $\text{Rem}_{\text{far}} \le 2\,\mathbb{P}\{\chi^2_N > 8L\} + \lambda^{-4}/10$.

The gap. $\Phi(u-\lambda) = \Phi(-\sqrt{8L}) \le \varphi(\sqrt{8L})/\sqrt{8L} = \lambda^{-4}/(4\sqrt{\pi L}) \le 0.142\,\lambda^{-4}$, as $L > 1$. Since $\lambda^{-4}/10 + 0.142\,\lambda^{-4} \le \lambda^{-4}/4$, part (ii) gives the bound of (iii).

The chi-square tail. Let $p = N/2 - 1 \ge 0$ and $y > N-2$. Integrating by parts once,

$$\int_y^\infty x^p e^{-x/2}\,\mathrm{d}x \;\le\; 2y^p e^{-y/2} + \frac{2p}{y}\int_y^\infty x^p e^{-x/2}\,\mathrm{d}x,$$

so $\mathbb{P}\{\chi^2_N > y\} \le (y/2)^p e^{-y/2}/(\Gamma(N/2)(1-(N-2)/y))$. At $y = 8L$ with $L \ge (d-3)/2$, $(N-2)/y \le \frac{1}{4}$. This gives $2\,\mathbb{P}\{\chi^2_{d-1} > 8\log\lambda\} \le 8(4\log\lambda)^{(d-3)/2}\lambda^{-4}/(3\,\Gamma(\frac{d-1}{2}))$, the bound stated after the theorem in the main paper. □

Note that $\beta_\star$ is the supremum of $|\beta_2|$ over the *near* zone $\theta \le \theta_1$, not over the whole sphere, where it would be of order λ and the bound would be useless. At the moderate values of λ of Section 6 of the main paper, θ_1 is optimised numerically in part (ii).

D Proofs for Sections 2 and 3 of the main paper

Proof of Proposition 1 of the main paper. The count (6a) and the law and the independence in (6b) are Azaïs et al. (2024, Assumption (A5) and Proposition 6), applied to X on $M = \mathbb{S}^{d-1}$, of dimension $d-1$, whose Karhunen–Loève expansion (4b) of the main paper has order m, with $Z = \sigma X$ under the null. The integer κ is also the codimension of the tangent space of the cone at $\lambda\psi_t$, $\lambda > 0$. For (6c), write $D_u\psi_t$ for the derivative of $t \mapsto \psi_t$ in a tangent direction u. Differentiating $\langle\psi_s, \psi_t\rangle = \langle s,t\rangle^k$ once in s, then once more in t, at $s = t$, shows that ψ_t and the vectors $k^{-1/2}D_u\psi_t$, for u in an orthonormal basis of the tangent space at t, are orthonormal in $\mathbb{R}^m$. Also $\langle Y, D_u\psi_t\rangle = \langle u, \nabla Z(t)\rangle$, as $Z(t) = \langle Y, \psi_t\rangle$. Let $P_t^\perp$ be the orthogonal projection onto the orthogonal complement of the span of ψ_t and the $k^{-1/2}D_u\psi_t$. The numerator of (1b) of the main paper is $\langle P_t^\perp Y, \psi_s\rangle$, and the ψ_s span $\mathbb{R}^m$. So, up to an orthogonal change of coordinates, which leaves $\sum_i \bar\zeta_i^2$ unchanged, the $\bar\zeta_i$ are the coordinates of $P_t^\perp Y$ in an orthonormal basis of its range, and $\kappa\hat\sigma_t^2 = \|P_t^\perp Y\|^2$, which is (6c). □

Proof of Lemma 1 of the main paper. Write $D(u) = \det(u\,\mathrm{Id} - r) = \prod_i (u - r_i)$, the r_i being the eigenvalues of r, and $h(u) = (\kappa + u^2)^{-(\nu+1)/2}$ with $\nu = m-1$, so that $H_r(\ell) = c\int_\ell^\infty Dh$ for a constant $c > 0$ that cancels; D is positive and increasing on $(\lambda_{\max}(r), \infty)$, and the integrals converge because $N < \nu$.

The Student bounds. Substituting $u = \vartheta w$, which maps $w = \ell_2$ to $u = \ell_1$,

$$H_r(\ell_1) = \vartheta\int_{\ell_2}^\infty D(\vartheta w)\,h(\vartheta w)\,\mathrm{d}w. \tag{D.1a}$$

For the density, $\kappa + \vartheta^2 w^2 \le \vartheta^2(\kappa + w^2)$ because $\vartheta \ge 1$, so $h(\vartheta w) \ge \vartheta^{-(\nu+1)}h(w)$; and $w \mapsto (\kappa + w^2)/(\kappa + \vartheta^2 w^2)$ has derivative of the sign of $\kappa(1-\vartheta^2) \le 0$, so for $w \ge \ell_2$ it is at most its value at ℓ_2 and

$h(\vartheta w) \le \big((\kappa+\ell_2^2)/(\kappa+\ell_1^2)\big)^{(\nu+1)/2} h(w)$. For the determinant, fix i and $w \ge \ell_2 > \lambda_{\max}(r)^+$. If $r_i \ge 0$ then $\vartheta w - r_i \ge \vartheta(w - r_i)$; if $r_i < 0$ then, with $b = -r_i \le \lambda_{\min}(r)^-$,

$$\frac{\vartheta w + b}{w + b} = \vartheta\,\frac{w + b/\vartheta}{w + b} \;\ge\; \vartheta\,\frac{w}{w+b} \;\ge\; \vartheta\,\frac{\ell_2}{\ell_2 + \lambda_{\min}(r)^-}, \tag{D.1b}$$

the last step because $w \mapsto w/(w+b)$ increases. Hence

$$D(\vartheta w) \ge \vartheta^N \big(1 + \lambda_{\min}(r)^-/\ell_2\big)^{-N} D(w). \tag{D.1c}$$

In the other direction

$$\frac{\vartheta w - r_i}{w - r_i} = \vartheta + \frac{(\vartheta-1) r_i}{w - r_i} \;\le\; \vartheta + \frac{(\vartheta-1)\lambda_{\max}(r)^+}{\ell_2 - \lambda_{\max}(r)^+} \;\le\; \vartheta\,\frac{\ell_2}{\ell_2 - \lambda_{\max}(r)^+}, \tag{D.1d}$$

since $z \mapsto z/(w-z)$ increases and $w \ge \ell_2$, so $D(\vartheta w) \le \vartheta^N \big(1 - \lambda_{\max}(r)^+/\ell_2\big)^{-N} D(w)$. Multiplying the two pairs of bounds and using $\vartheta \cdot \vartheta^N \cdot \vartheta^{-(\nu+1)} = \vartheta^{N-\nu} = \vartheta^{-\kappa}$ gives (9a) of the main paper.

The Gaussian bounds. Substituting $u = \sqrt{w^2 + 2g}$, which also maps $w = \ell_2$ to $u = \ell_1$ and has $\mathrm{d}u = (w/u)\,\mathrm{d}w$ and $\varphi(u) = e^{-g}\varphi(w)$,

$$G_r(\ell_1) = e^{-g} \int_{\ell_2}^{\infty} D(u)\,\frac{w}{u}\,\varphi(w)\,\mathrm{d}w. \tag{D.2a}$$

Here $u \ge w$, so $D(u) \ge D(w)$, and $w \mapsto w/\sqrt{w^2+2g}$ increases, so $w/u \ge \ell_2/\ell_1$; this is the left inequality. For the right one, $w/u \le 1$ and $u - w = 2g/(u+w) \le g/\ell_2$, so

$$\frac{u - r_i}{w - r_i} = 1 + \frac{u - w}{w - r_i} \;\le\; 1 + \frac{g}{\ell_2\big(\ell_2 - \lambda_{\max}(r)^+\big)}. \tag{D.2b}$$

The comparison. In (9a) the ratio of the two bounds is $\vartheta^{N+1+\kappa}\big((\kappa+\ell_2^2)/(\kappa+\ell_1^2)\big)^{(\nu+1)/2}$ times two factors that tend to 1; since $N + 1 + \kappa = \nu + 1$ and $(\kappa+\ell_2^2)/(\kappa+\ell_1^2) \to \vartheta^{-2}$ as $\ell_2 \to \infty$, that ratio tends to $\vartheta^{\nu+1}\vartheta^{-(\nu+1)} = 1$. In (9b) it is (ℓ_1/ℓ_2) times a factor tending to 1 at fixed g. □

Proof of Lemma 4 of the main paper. The tensor $\mathrm{Sym}(t^{\otimes(k-j)} \otimes v^{\otimes j})$ is the mean of the $\binom{k}{j}$ tensors with v at j positions and t elsewhere. So the expansion of $s^{\otimes k}$ reads $\psi_s = \sum_j \binom{k}{j}^{1/2} \cos^{k-j}\theta\,\sin^j\theta\,\Psi_j(v)$. Two of these tensors, built with v and v', have the inner product $\langle v, v'\rangle^j$ if the positions agree, and 0 otherwise because v and v' are orthogonal to t. This gives the inner products and the orthogonality of the $\mathcal{V}_j$. Each ψ_s lies in $\mathcal{V}_0 \oplus \cdots \oplus \mathcal{V}_k$ and the ψ_s span $\mathbb{R}^m$ (Section 2.3 of the main paper), so this sum is $\mathbb{R}^m$. As $\Psi_0 = \psi_t$ and $\Psi_1(u) = k^{-1/2} D_u \psi_t$, the range of $P_t^\perp$ is $\mathcal{V}_2 \oplus \cdots \oplus \mathcal{V}_k$. The numerator of (1b) of the main paper is $\langle P_t^\perp Y, \psi_s\rangle$, which gives the first formula. Write $Y[x_1, \dots, x_k] = \langle Y, x_1 \otimes \cdots \otimes x_k\rangle$. Two derivatives of $\theta \mapsto Z(\cos\theta\, t + \sin\theta\, v)$ at 0 give $v^\top \nabla^2 Z(t) v = -kZ(t) + k(k-1)Y[t, \dots, t, v, v]$, and $Y[t, \dots, t, v, v] = \binom{k}{2}^{-1/2}\langle Y, \Psi_2(v)\rangle$ as Y is symmetric. So $v^\top \Omega_t v = (k-1)\binom{k}{2}^{-1/2}\langle Y, \Psi_2(v)\rangle$, and $(k-1)\binom{k}{2}^{-1/2} = 1/\rho$. Since $\langle \Psi_2(v), \Psi_2(v')\rangle = \langle v, v'\rangle^2 = \langle vv^\top, v'v'^\top\rangle_F$ and the $vv^\top$ span the symmetric matrices, the map $vv^\top \mapsto \Psi_2(v)$ extends to the isometry ι. Then $P_{\mathcal{V}_2} Y = \iota(\rho\,\Omega_t)$, as both have the inner product $\rho\, v^\top \Omega_t v$ with every $\Psi_2(v)$, and the last formula follows from $\kappa\hat{\sigma}_t^2 = \|P_t^\perp Y\|^2$. □

Proof of Lemma 6 of the main paper. Write $D(u) = \prod_i (u - r_i)$ and $h(u) = (\kappa + u^2)^{-(\nu+1)/2}$, as in the proof of Lemma 1 above, so that $H_r(\ell) = c\int_\ell^\infty Dh$. As $D \ge 0$ on $[\ell_2, \infty)$, $H_r(\ell_2)$ is nonincreasing in ℓ_2, and the ratio is nondecreasing in ℓ_2. Fix i with $r_i < \ell_2$ and let U have the density proportional to Dh on (ℓ_2, ∞). Differentiating under the integral, which is licit since $|\partial_{r_i}(Dh)| \le Dh/(\ell_2 - r_i)$ there,

$$\partial_{r_i} \log \frac{H_r(\ell_1)}{H_r(\ell_2)} = \mathbb{E}\Big[\frac{1}{U - r_i}\Big] - \mathbb{E}\Big[\frac{1}{U - r_i} \;\Big|\; U > \ell_1\Big].$$

The function $u \mapsto 1/(u - r_i)$ is positive and decreasing on (ℓ_2, ∞), so its conditional mean on $\{U > \ell_1\}$ is at most $1/(\ell_1 - r_i)$, which is at most its conditional mean on $\{U \le \ell_1\}$. The mean $\mathbb{E}[1/(U - r_i)]$ lies between the two, so the derivative is nonnegative. The case $r_i = \ell_2$ follows by continuity of H_r in r. □

Proof of the matching bound in Theorem 2 of the main paper. With $\sigma = 1$, let $E = \rho^2/(2(A^2 + \rho^2))$ with A the constant (12b) of the main paper. We show that the lim sup of $-\lambda^{-2} \log(1 - \mathbb{P}_\lambda\{\mathbf{T} \le \alpha\})$ is at most E.

A configuration. Fix a positive semidefinite r of size N with $\|r\|_F = T_{\max}$, and $a \in \big(\lambda_{\max}(r), u_\alpha(r, \lambda_{\max}(r))\big)$. With $t = e$ in Lemma 4 of the main paper, let $y_0 = a\,\psi_e + \iota(\rho r)$. By the same lemma, its field at $s = \cos\theta\, e + \sin\theta\, v$ is $z(s) = a\cos^k\theta + (1 - \cos^k\theta)\,\rho\, w_2(\theta)\, v^\top r v$. By its definition, $\rho w_2 = \frac{k}{2}\cos^{k-2}\theta \sin^2\theta/(1 - \cos^k\theta)$. For $x \in [0, 1)$, $f(x) = 1 - x^k - \frac{k}{2}x^{k-2}(1 - x^2)$ has $f(1) = 0$ and $f'(x) = \frac{k}{2}(k-2)x^{k-3}(x^2 - 1) \le 0$, so $f \ge 0$: with $x = \cos\theta$, $\rho w_2 \le 1$ on $(0, \pi/2]$. On $[\pi/2, \pi)$, $w_2 \le 0$ as k is odd. As r is positive semidefinite, $\rho\, w_2\, v^\top r v \le \lambda_{\max}(r) < a$. So $z(s) < a = z(e)$ for $s \ne e$, the value at $-e$ being $-a$. At y_0 the maximiser is e and $\lambda_1 = a$; $\nabla z(e) = 0$, as y_0 has no component in $\mathcal{V}_1$; $\Omega = r$, so $\nabla^2 z(e) = k(r - a\,\mathrm{Id})$ is negative definite; $\kappa\hat\sigma^2 = \|\iota(\rho r)\|^2 = \rho^2 T_{\max}^2 = \kappa$; and $\lambda_2 = \sup \rho\, w_2\, v^\top r v = \lambda_{\max}(r)$, the value on the fibre $\theta = 0$. Hence $\mathbf{T}(y_0) = H_r(a)/H_r(\lambda_{\max}(r)) > \alpha$.

Continuity at y_0. The field $z_y(s) = \langle y, \psi_s\rangle$ is linear in y, so $z_y \to z_{y_0}$ in C^2 as $y \to y_0$. Take a geodesic ball B around e on which the Hessian of z_{y_0} is negative definite. As e is the only maximiser of z_{y_0}, there is $\varsigma > 0$ with $z_{y_0} \le a - \varsigma$ outside B. For y close to y_0, $z_y < a - \varsigma/2 < z_y(e)$ outside B and the Hessian of z_y is negative definite on B. So z_y has one maximiser $\hat t(y)$, and it lies in B. As B can be taken small, $\hat t(y) \to e$. Then λ_1, $\hat\sigma^2 = (\|y\|^2 - \lambda_1^2)/\kappa$ and the eigenvalues of Ω are continuous at y_0. By Lemma B.3, the second maximum λ_2^t of the field z_y at a point t is continuous in (t, y). As $\lambda_2(y)$ is its value at $t = \hat t(y)$ and $\hat t(y) \to e$, λ_2 is continuous at y_0. The map $(\ell_1, \ell_2, r) \mapsto H_r(\ell_1)/H_r(\ell_2)$ is continuous on $\{\lambda_{\max}(r) \le \ell_2 \le \ell_1\}$, where $H_r(\ell_2) > 0$, and $(T_1, T_2, \Omega/\hat\sigma)$ lies in this set by (3) of the main paper. So there is $\eta_0 > 0$ with $\mathbf{T} > \alpha$ on the ball $\mathbb{B}(y_0, \eta_0)$ of $\mathbb{R}^m$.

The upper bound. $\mathbf{T}$ is invariant under $y \mapsto cy$, $c > 0$ (Section 2.4 of the main paper), so $\mathbf{T} > \alpha$ on the open cone $\mathcal{U} = \{cy : c > 0,\ y \in \mathbb{B}(y_0, \eta_0)\}$. Let $y^* = c^* y_0$ with $c^* = \langle y_0, \psi_e\rangle/\|y_0\|^2 = a/(a^2 + \kappa)$, the point of the ray closest to ψ_e, and let $\eta > 0$ with $\mathbb{B}(y^*, \eta) \subset \mathcal{U}$. For $\lambda \ge 1$, $\mathbb{B}(\lambda y^*, \eta) \subset \lambda\,\mathbb{B}(y^*, \eta) \subset \mathcal{U}$. As $Y = \lambda\psi_e + W$ with W standard Gaussian in $\mathbb{R}^m$,

$$1 - \mathbb{P}_\lambda\{\mathbf{T} \le \alpha\} \;\ge\; \mathbb{P}\big\{W \in \mathbb{B}(\lambda(y^* - \psi_e), \eta)\big\} \;\ge\; \frac{|\mathbb{B}(0,\eta)|}{(2\pi)^{m/2}} \exp\Big(-\tfrac{1}{2}\big(\lambda\|y^* - \psi_e\| + \eta\big)^2\Big),$$

with $|\mathbb{B}(0,\eta)|$ the volume of the ball. Here $\|y^* - \psi_e\|^2 = 1 - \langle y_0, \psi_e\rangle^2/\|y_0\|^2 = \kappa/(\kappa + a^2) = \rho^2/(\rho^2 + (a/T_{\max})^2)$. So the lim sup is at most $\rho^2/(2(\rho^2 + (a/T_{\max})^2))$. Let a tend to $u_\alpha(r, \lambda_{\max}(r))$ and take r at the maximum in (12b) of the main paper: the lim sup is at most E. □

The proof of Theorem 3 of the main paper bounds the supremum of the noise field, then the second maximum and the Hessian through it. With $\sigma = 1$, write $\mathcal{X} = \sup_M X$, $\xi = X(e)$ and $\hat c = \langle e, \hat t\rangle$. As $X(-t) = -X(t)$ for all t (Lemma 3 of the main paper), $\mathcal{X} = \sup_M |X| \ge 0$.

Lemma D.1 (Supremum of the noise field)**.** *$\mathbb{E}\mathcal{X} \le \sqrt{kd}$, and $\mathbb{P}\{\mathcal{X} > \sqrt{kd} + \tau\} \le e^{-\tau^2/2}$ for $\tau \ge 0$.*

Proof. For $s, t \in M$, $\mathbb{E}(X(s) - X(t))^2 = 2(1 - \langle s, t\rangle^k)$. For $x \in [-1, 1)$, $1 - x^k = (1 - x)\sum_{i<k} x^i$, and the sum lies in $[0, k]$, since it equals $(1 - x^k)/(1 - x) \ge 0$ and each term is at most 1 in absolute value. So $\mathbb{E}(X(s) - X(t))^2 \le 2k(1 - \langle s, t\rangle) = k\|s - t\|^2 = \mathbb{E}(X'(s) - X'(t))^2$, with $X'(t) = \sqrt{k}\langle g, t\rangle$ and g standard Gaussian in $\mathbb{R}^d$. The Sudakov–Fernique inequality (Azaïs and Wschebor, 2009, Chapter 2) gives $\mathbb{E}\mathcal{X} \le \mathbb{E}\sup_M X' = \sqrt{k}\,\mathbb{E}\|g\| \le \sqrt{kd}$. The Borell–TIS inequality (Azaïs and Wschebor, 2009, Chapter 2), with $\sup_t \mathrm{Var}X(t) = 1$, gives $\mathbb{P}\{\mathcal{X} > \mathbb{E}\mathcal{X} + \tau\} \le e^{-\tau^2/2}$. □

Lemma D.2 (Pathwise bounds through the supremum)**.** *For $\theta \in (0, \pi)$ let*

$$g_k(\theta) = \frac{(|\cos\theta| + \sin\theta)^k - |\cos\theta|^k - k|\cos\theta|^{k-1}\sin\theta}{1 - \cos^k\theta},$$

so that $\Upsilon_k = \sup g_k$ *in Theorem 3 of the main paper. Then* $k-1 < \Upsilon_k < \infty$. *Almost surely, for every* $t \in M$, *every unit* $v \perp t$, *every* $2 \le j \le k$ *and* $s = \cos\theta\, t + \sin\theta\, v \in M_t$, *with* $\Psi_j(v)$ *of Lemma 4 of the main paper,* (a) $|\langle W, \Psi_j(v)\rangle| \le \binom{k}{j}^{1/2}\mathcal{X}$, *and* (b) $|\zeta^t(s)| \le g_k(\theta)\,\mathcal{X} \le \Upsilon_k\mathcal{X}$.

Proof. The form $S_W(x_1,\dots,x_k) = \langle W, x_1\otimes\cdots\otimes x_k\rangle$ is symmetric and k-linear on $\mathbb{R}^d$, and $S_W(x,\dots,x) = X(x)$. Banach's theorem on symmetric multilinear forms (Banach, 1938), stated for tensors in Friedland and Lim (2018, Section 5), gives

$$\sup\big\{|S_W(x_1,\dots,x_k)| :\ \|x_1\| = \cdots = \|x_k\| = 1\big\} = \sup_M |X| = \mathcal{X}.$$

As W is symmetric, $\langle W, \Psi_j(v)\rangle$ is $\binom{k}{j}^{1/2}$ times the value of S_W at $k-j$ copies of t and j copies of v, which gives (a). Lemma 4 of the main paper with W in place of Y gives $\zeta^t(s) = \sum_{j\ge2} w_j(\theta)\langle W, \Psi_j(v)\rangle$, so by (a) and the binomial formula, $|\zeta^t(s)| \le \mathcal{X}\sum_{j\ge2}\binom{k}{j}|\cos\theta|^{k-j}\sin^j\theta/(1-\cos^k\theta) = g_k(\theta)\mathcal{X}$. Finally $g_k(\theta)\to 0$ as $\theta\to\pi$, and $g_k(\theta) = k-1+\frac{2}{k}\binom{k}{3}\theta + O(\theta^2)$ as $\theta\to 0$, so g_k is bounded and $\Upsilon_k > k-1$. □

Lemma D.3 (The argmax under the alternative)**.** *On the event* $\{\lambda > \mathcal{X} - \xi\}$, $\hat c > 0$, $1-\hat c^2 \le (\mathcal{X}-\xi)/\lambda$, *and*

$$\lambda_2 \ \le\ \Upsilon_k\mathcal{X} + (k-1)(\mathcal{X}-\xi), \qquad \lambda_{\max}(\Omega) \ \le\ (k-1)(2\mathcal{X}-\xi).$$

Proof. $\lambda_1 = Z(\hat t) \ge Z(e) = \lambda+\xi$ and $Z(\hat t) = \lambda\hat c^k + X(\hat t) \le \lambda\hat c^k + \mathcal{X}$, so $\lambda\hat c^k \ge \lambda+\xi-\mathcal{X} > 0$. Hence $\hat c > 0$, $\hat c^2 \ge \hat c^k$ and $1-\hat c^2 \le 1-\hat c^k \le (\mathcal{X}-\xi)/\lambda$. On $M_{\hat t}$, $Z^{|\hat t} = \zeta^{\hat t} + \mu_{\hat t}$, with $\zeta^{\hat t} \le \Upsilon_k\mathcal{X}$ by Lemma D.2(b) and $|\mu_{\hat t}| \le (k-1)\lambda(1-\hat c^2) \le (k-1)(\mathcal{X}-\xi)$ by (B.2b); and $Z^{|\hat t}(-\hat t) = 0$. This bounds λ_2. For a unit $v \perp \hat t$, Lemma 4 of the main paper and Lemma D.2(a) give

$$\rho\, v^\top\Omega v = \langle W, \Psi_2(v)\rangle + \lambda\langle\psi_e, \Psi_2(v)\rangle \le \tbinom{k}{2}^{1/2}\big(\mathcal{X} + \lambda\hat c^{k-2}\langle e, v\rangle^2\big) \le \tbinom{k}{2}^{1/2}\big(\mathcal{X} + \lambda(1-\hat c^2)\big),$$

since $0 < \hat c \le 1$ and $\langle e, v\rangle^2 \le 1-\hat c^2$. As $\binom{k}{2}^{1/2}/\rho = k-1$, this gives the bound on $\lambda_{\max}(\Omega)$. □

Proof of Theorem 3 of the main paper. Let $\chi^2 = \|W - \xi\psi_e\|^2$, independent of ξ with the law χ^2_{m-1}, and $\mathcal{G} = \{\mathcal{X} \le \bar{\mathcal{X}}\}\cap\{|\xi| \le \tau\}\cap\{\chi^2 \le m-1+\delta\}$. By Lemma D.1, $\mathbb{P}(\mathcal{G}^c)$ is at most the bound of the theorem. We show that $\mathbf{T} \le \alpha$ almost surely on $\mathcal{G}$.

On $\mathcal{G}$, $\lambda \ge \bar v + \tau > \bar{\mathcal{X}} + \tau \ge \mathcal{X} - \xi$, so Lemma D.3 applies: $\lambda_2 \le \Upsilon_k\mathcal{X} + (k-1)(\mathcal{X}-\xi) \le \bar v$ and $\lambda_{\max}(\Omega)^+ \le (k-1)(2\mathcal{X}-\xi) \le \bar\ell$, using $2\mathcal{X}-\xi \ge \mathcal{X} \ge 0$. Also $\bar\ell < \bar v$, because $\bar v - \bar\ell = (\Upsilon_k - k + 1)\bar{\mathcal{X}}$ and $\Upsilon_k > k-1$, and $\lambda_1 \ge \lambda+\xi \ge \lambda-\tau \ge \bar v$.

Put $r = \Omega/\hat\sigma$ and $v = \bar v/\hat\sigma$. Then $\lambda_{\max}(r)^+ < v \le T_1$, and $\lambda_{\max}(r) \le T_2 \le v$ by (3) of the main paper. As H_r decreases on $[\lambda_{\max}(r), \infty)$, $\mathbf{T} \le H_r(T_1)/H_r(v)$. The right-hand inequality of (9a) of the main paper, with $\ell_2 = v$, $\ell_1 = T_1$, $\lambda_{\max}(r)^+/v \le \bar\ell/\bar v$ and $\nu+1 = m$, gives

$$\mathbf{T} \ \le\ \Big(\frac{\lambda_1}{\bar v}\Big)^{N+1}\Big(1-\frac{\bar\ell}{\bar v}\Big)^{-N}\Big(\frac{\kappa\hat\sigma^2+\bar v^2}{\kappa\hat\sigma^2+\lambda_1^2}\Big)^{m/2}.$$

By (6c) of the main paper at $\hat t$, $\kappa\hat\sigma^2 = \|Y\|^2 - \lambda_1^2$. Hence $\log\mathbf{T} \le F(\lambda_1) + N\log(\bar v/(\bar v-\bar\ell))$, with $F(x) = d\log(x/\bar v) + \frac{m}{2}\log\big(1-(x^2-\bar v^2)/\|Y\|^2\big)$ on $[\bar v, \|Y\|)$.

On $\mathcal{G}$, $\|Y\|^2 = (\lambda+\xi)^2 + \chi^2 \le \bar n^2$. For $x \ge \lambda-\tau$, $F'(x) = d/x - mx/(\|Y\|^2 - x^2 + \bar v^2) \le d/x - mx/\|Y\|^2 \le 0$, by the second hypothesis. As $\lambda-\tau \le \lambda_1 < \|Y\|$ (proof of Corollary 4 of the main paper), $F(\lambda_1) \le F(\lambda-\tau)$. Finally $F(\lambda-\tau)$ is nondecreasing in $\|Y\|^2$, since $(\lambda-\tau)^2 \ge \bar v^2$, so it is at most its value with $\bar n$ in place of $\|Y\|$, and (16b) of the main paper gives $\log\mathbf{T} \le \log\alpha$. □

Proof of the last part of Theorem 3 of the main paper. Put $\tau = d^{1/4}$ and $\delta = m^{3/4}$ in the theorem. By the Chebyshev inequality the last term of its bound is at most $2(m-1)/\delta^2$, so the bound tends to 0. As $d\to\infty$, $\bar v/\sqrt d \to v_k$, $\bar\ell/\sqrt d \to \ell_k$, $(\lambda_d - \tau)/\sqrt d \to \bar\lambda$ and $\bar n^2/m \to 1$, because $\lambda_d^2 = O(d)$ and $m \ge d^k/k!$.

The first hypothesis holds for d large, as $\bar\lambda > C_k > v_k$, and the second because $\bar\lambda > v_k > 1$. With $\log(1-y) \le -y$, the left-hand side of (16b) of the main paper is at most

$$d\Big[\log\frac{\bar\lambda}{v_k} + \log\frac{v_k}{v_k-\ell_k} - \frac12\big(\bar\lambda^2 - v_k^2\big) + o(1)\Big].$$

The map $x \mapsto \frac12(x^2-v_k^2) - \log(x/v_k)$ increases on (v_k,∞), as $v_k > 1$, and equals $\log(v_k/(v_k-\ell_k)) > 0$ at $x = C_k$. So the bracket has a negative limit when $\bar\lambda > C_k$, and (16b) holds for d large. □

Lemma D.4 (The constant in growing dimension)**.** *Let k and α be fixed. The constant A of (12b) of the main paper tends to 1 as $d \to \infty$.*

Proof. $A > 1$ (main paper, after (12b)). Let r be positive semidefinite with $\|r\|_F = T_{\max}$, and $u = u_\alpha(r, \lambda_{\max}(r))$. Every eigenvalue of r is at most $T_{\max}$. If $u > T_{\max}$, Lemma 6 of the main paper, applied to ℓ_2 raised from $\lambda_{\max}(r)$ to $T_{\max}$, then to every eigenvalue raised to $T_{\max}$, gives $\alpha = H_r(u)/H_r(\lambda_{\max}(r)) \le P(u)$, with $P(u) = H_{T_{\max}\mathrm{Id}}(u)/H_{T_{\max}\mathrm{Id}}(T_{\max})$. With $u = T_{\max}(1+x)$ and $\kappa = \rho^2 T_{\max}^2$, $P(T_{\max}(1+x)) = \int_x^\infty g / \int_0^\infty g$ with $g(x) = x^N(\rho^2 + (1+x)^2)^{-m/2}$, and $P(T_{\max}(1+x))$ decreases in x. So it suffices to show, for every $\epsilon \in (0,1]$, that $P(T_{\max}(1+\epsilon)) < \alpha$ for d large: then $u \le T_{\max}(1+\epsilon)$ for every such r, and $A \le 1+\epsilon$. As $\rho < 1$, $\rho^2 + (1+x)^2 \le 2(1+x)^2$ and $(\log g)'(x) \le N/x - m/(2(1+x)) = \big(N - x(m/2-N)\big)/(x(1+x))$. Let $p = (m/2-N)/2$, which tends to infinity with d, and take d so large that $p > 1$ and $\epsilon p \ge 4N$. Then $(\log g)' \le -p/(1+x)$ on $[\epsilon/4, \infty)$, so g decreases there and $g(x) \le g(\epsilon/2)\big((1+\epsilon/2)/(1+x)\big)^p$ for $x \ge \epsilon/2$. Hence

$$P\big(T_{\max}(1+\epsilon)\big) \;\le\; \frac{\int_\epsilon^\infty g}{\int_{\epsilon/4}^{\epsilon/2} g} \;\le\; \frac4\epsilon \frac{1+\epsilon}{p-1}\Big(\frac{1+\epsilon/2}{1+\epsilon}\Big)^p,$$

which tends to 0 as $d \to \infty$, since $N/m \to 0$. □

The last lemma of this appendix gives the growth, in the order k, of the constant C_k of Theorem 3 of the main paper.

Lemma D.5 (Growth of the constant C_k)**.** *Let $k \ge 3$ be odd, and let Υ_k, v_k, ℓ_k and C_k be as in Theorem 3 of the main paper. Then*

$$2^{k/2}\big(1-(k+1)2^{-k}\big) \;\le\; \Upsilon_k \;\le\; \max\Big\{\frac{2^{k/2}}{1-\cos^k(1/4)},\; k\Big(1+\frac{k}{32}\Big)e^{k/4}\Big\}, \tag{D.3a}$$

$$v_k \;<\; C_k \;\le\; v_k + \frac{v_k}{v_k^2-1}\log\frac{v_k}{v_k-\ell_k}. \tag{D.3b}$$

In particular $C_k/(2^{k/2}\sqrt{k}) \to 1$ as $k \to \infty$ along the odd integers. Moreover

$$C_k \le 1.72\cdot 2^{k/2}\sqrt{k} \quad \text{for } k \ge 3, \qquad C_k \le 1.02\cdot 2^{k/2}\sqrt{k} \quad \text{for } k \ge 21. \tag{D.3c}$$

Proof. Let g_k be the function of Lemma D.2, so that $\Upsilon_k = \sup_{0<\theta<\pi} g_k(\theta)$. Write $a = |\cos\theta|$ and $b = \sin\theta$. The numerator of $g_k(\theta)$ is $n_k(\theta) = (a+b)^k - a^k - ka^{k-1}b = \sum_{j=2}^k \binom{k}{j} a^{k-j}b^j$. It is non-negative, and it is at most $(a+b)^k \le 2^{k/2}$. At $\theta = \pi/4$, $a = b = 2^{-1/2}$ and $g_k(\pi/4) = \big(2^{k/2} - (k+1)2^{-k/2}\big)/\big(1-2^{-k/2}\big) \ge 2^{k/2} - (k+1)2^{-k/2}$, because the numerator is positive, as $2^k > k+1$, and the denominator is in $(0,1)$. This is the lower bound of (D.3a).

For the upper bound we split $(0,\pi)$ into three ranges. If $\pi/2 \le \theta < \pi$, then $\cos^k\theta \le 0$, as k is odd, so the denominator $1-\cos^k\theta$ is at least 1 and $g_k(\theta) \le 2^{k/2}$. If $1/4 \le \theta \le \pi/2$, then $0 \le \cos^k\theta \le \cos^k(1/4)$ and $g_k(\theta) \le 2^{k/2}/(1-\cos^k(1/4))$. Let now $0 < \theta < 1/4$. As $a \le 1$, $\binom{k}{j} \le k^j/j!$ and $b \le \theta$,

$$n_k(\theta) \;\le\; \sum_{j\ge2}\frac{(k\theta)^j}{j!} \;\le\; \frac{(k\theta)^2}{2}\,e^{k\theta}.$$

The second derivative of $\log\cos$ is $-1/\cos^2 \le -1$ on $[0,\pi/2)$, so $\cos\theta \le e^{-\theta^2/2}$. With $y = k\theta^2/2$ and $1-e^{-y} \ge y/(1+y)$, the denominator $1-\cos^k\theta \ge 1-e^{-y}$ is at least $y/(1+y)$. Hence

$$g_k(\theta) \;\le\; \frac{(k\theta)^2}{2}\, e^{k\theta}\, \frac{1+y}{y} \;=\; k\, e^{k\theta}\Big(1+\frac{k\theta^2}{2}\Big) \;\le\; k\Big(1+\frac{k}{32}\Big)e^{k/4}.$$

The three bounds give the upper bound of (D.3a), as $2^{k/2} \le 2^{k/2}/(1-\cos^k(1/4))$.

By Lemma D.2, $\Upsilon_k > k-1$, so $v_k > \ell_k$. Also $\ell_k \ge 4\sqrt{3} > 1$, as $k \ge 3$. So $L_k := \log\big(v_k/(v_k-\ell_k)\big) > 0$. Let $F_k(x) = \frac12(x^2 - v_k^2) - \log(x/v_k)$. By (16d) of the main paper, $C_k > v_k$ and $F_k(C_k) = L_k$. For $x > v_k$, $\log(x/v_k) \le (x-v_k)/v_k$, so

$$F_k(x) \;\ge\; (x-v_k)\Big(\frac{x+v_k}{2} - \frac{1}{v_k}\Big) \;\ge\; (x-v_k)\Big(v_k - \frac{1}{v_k}\Big).$$

At $x = C_k$ this gives $L_k \ge (C_k - v_k)(v_k - 1/v_k)$, which is (D.3b), as $v_k > 1$.

Let $k \to \infty$ along the odd integers. The lower bound of (D.3a), divided by $2^{k/2}$, is $1-(k+1)2^{-k}$, and it tends to 1. The upper bound, divided by $2^{k/2}$, is the maximum of $1/(1-\cos^k(1/4))$, which tends to 1 as $\cos(1/4) < 1$, and of $k(1+k/32)(e^{1/4}/\sqrt{2})^k$, which tends to 0 as $e^{1/4} < \sqrt{2}$, that is $e < 4$. So $\Upsilon_k/2^{k/2} \to 1$. As $(k-1)2^{-k/2} \to 0$, $v_k/(2^{k/2}\sqrt{k}) = (\Upsilon_k + k - 1)/2^{k/2} \to 1$, and $v_k \to \infty$. Also $\ell_k/v_k = 2(k-1)/(\Upsilon_k + k - 1) \to 0$, so $L_k \to 0$. By (D.3b), $C_k - v_k \to 0$, and the claim follows.

It remains to prove the two bounds on $C_k/(2^{k/2}\sqrt{k})$. Let first $k \ge 125$ be odd. Then $\cos^k(1/4) \le \cos^{125}(1/4) < 0.0194$. The function $k \mapsto k(1+k/32)(e^{1/4}/\sqrt{2})^k$ decreases on $[125,\infty)$, as the derivative of its logarithm is $1/k + 1/(32+k) + \frac14 - \frac12\log 2 < 0$ there, and it is below 0.01 at $k = 125$. So (D.3a) gives $\Upsilon_k \le 2^{k/2}/0.9806$, and $v_k/(2^{k/2}\sqrt{k}) = \Upsilon_k/2^{k/2} + (k-1)2^{-k/2} < 1.0198$, as $(k-1)2^{-k/2} < 10^{-16}$. By the lower bound of (D.3a), $\Upsilon_k \ge 3(k-1)$, so $\ell_k/v_k \le 1/2$ and $L_k \le \log 2$. As $v_k \ge \ell_k \ge 248\sqrt{125}$, (D.3b) gives $C_k \le v_k\big(1+\log 2/(v_k^2-1)\big) \le (1+10^{-7})\,v_k$. Hence $C_k \le 1.02\cdot 2^{k/2}\sqrt{k}$.

For the odd k from 3 to 123 the bounds are checked by interval arithmetic. On $(0,\pi/2]$, $1-\cos\theta = \sin^2\theta/(1+\cos\theta)$ gives

$$g_k(\theta) = (1+a)\,\frac{\sum_{j=2}^{k}\binom{k}{j}a^{k-j}b^{j-2}}{\sum_{i=0}^{k-1}a^i}, \qquad a=\cos\theta,\ b=\sin\theta,$$

a quotient of two polynomials with non-negative coefficients. Also $g_k(\pi-\theta) \le g_k(\theta)$ on $(0,\pi/2]$: the numerator is the same and the denominator $1+\cos^k\theta$ is larger, as k is odd. So Υ_k is the supremum of g_k over $(0,\pi/2]$. On a cell $[\theta_0,\theta_1] \subset [0,\pi/2]$, $\cos\theta_1 \le a \le \cos\theta_0$ and $b \le \sin\theta_1$. So g_k is at most the value of the quotient with $a = \cos\theta_0$ and $b = \sin\theta_1$ in the factor $1+a$ and in the numerator, and $a = \cos\theta_1$ in the denominator. The largest of these values over a subdivision of $[0,\pi/2]$ is an upper bound Υ^+ of Υ_k, and each value of g_k is a lower bound Υ^-. For $v > \ell_k$ let $C(v) > v$ solve (16d) of the main paper with v in place of v_k, so that $C_k = C(v_k)$. The equation is $\frac12(C^2-v^2) - \log C + \log(v-\ell_k) = 0$. Its left side increases in $C > 1$, and it decreases in v when $v(v-\ell_k) > 1$. So $C(v)$ increases on the half-line of the $v > \ell_k$ with $v(v-\ell_k) > 1$. The proof of (D.3b) uses only $v_k > \ell_k > 1$, so (D.3b) holds for v and $C(v)$. Let $v^\pm = (\Upsilon^\pm + k - 1)\sqrt{k}$. If $v^-(v^- - \ell_k) > 1$, then

$$C_k \;\le\; C(v^+) \;\le\; v^+ + \frac{v^+}{(v^+)^2-1}\,\log\frac{v^+}{v^+-\ell_k}.$$

The script `numerics/constant_Ck_certified.py`, in the folder `paper2` of the repository cited in the main paper, does this computation with outward rounding. It checks that $v^-(v^- - \ell_k) > 1$, and it finds $C_k/(2^{k/2}\sqrt{k}) \le 1.72$ for every odd k from 3 to 123, and at most 1.02 for every odd k from 21 to 123. The largest value is at $k = 5$, in $[1.71865, 1.71869]$. □

E Normalisations in the literature

The noise tensor of Perry et al. (2020) is $\sqrt{2/d}\,W$ in law, so their signal is our $\lambda/(\sigma\sqrt{d/2})$. At fixed k, their Theorem 3.2 makes weak detection impossible below the largest signal λ_{PWB} with $\frac{1}{2}\lambda_{\mathrm{PWB}}^2 t^k/(1+t^k) \le -\frac{1}{2}\log(1-t^2)$ for all $t \in [0,1)$, which is 1.5847 for $k=3$, that is $\lambda/\sigma = 1.1205\sqrt{d}$. Jagannath et al. (2020) observe $\lambda\sqrt{d}\,x^{\otimes k} + W'$, with x uniform on the sphere and W' a tensor of independent standard Gaussian entries. The projection of W' on the symmetric tensors has the law of W, and the rest is independent of it and does not depend on the signal, so their λ is our $\lambda/(\sigma\sqrt{d})$, and their critical value, 1.2066 for $k=3$, reads $\lambda/\sigma = 1.2066\sqrt{d}$. Lesieur et al. (2017) work at fixed signal-to-noise ratio as $d \to \infty$, with a spike whose entries are independent draws from a known prior, which is $\lambda \asymp \sqrt{d}$ in our normalisation.

F Closed forms of G_r and H_r

The functions G_r and H_r of Section 2.4 of the main paper, with the notation of that paper, are finite sums of incomplete moments.

Proposition F.1. *Write $\gamma = \sqrt{\kappa/(m-1)}$ and $\nu = m-1$, and let $e_j(r)$ be the elementary symmetric functions of the eigenvalues of r. Then*

$$G_r(\ell) = \sum_{j=0}^{N} (-1)^j e_j(r)\, I_{N-j}(\ell), \qquad I_0 = 1-\Phi, \quad I_1 = \varphi, \quad I_p(\ell) = \ell^{p-1}\varphi(\ell) + (p-1) I_{p-2}(\ell), \tag{F.1a}$$

and

$$H_r(\ell) = \gamma \sum_{j=0}^{N} (-1)^j e_j(r)\, \gamma^{N-j}\, K_{N-j}(\ell/\gamma), \qquad K_p(a) = \frac{c_\nu \nu^{\frac{p+1}{2}}}{2} B\Big(\tfrac{\nu-p}{2}, \tfrac{p+1}{2}\Big) I_{w_a}\Big(\tfrac{\nu-p}{2}, \tfrac{p+1}{2}\Big) \tag{F.1b}$$

for $a \ge 0$, with $w_a = \nu/(\nu+a^2)$, $c_\nu = \Gamma(\frac{\nu+1}{2})/(\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2}))$ and I_w the regularised incomplete Beta function, the integrals converging because $N < \nu$. Both G_r and H_r are positive and strictly decreasing on $(\lambda_{\max}(r), \infty)$, with limit 0 at $+\infty$. The prefactor γ and the constant c_ν cancel in $\mathbf{T}$.

Proof. Expanding $\det(u\,\mathrm{Id} - r) = \sum_j (-1)^j e_j(r) u^{N-j}$ reduces both statements to incomplete moments. For I_p, integrate by parts: $\int_\ell^\infty u^{p-1}\,(u\varphi(u))\,\mathrm{d}u = \ell^{p-1}\varphi(\ell) + (p-1)I_{p-2}(\ell)$. For K_p, substitute $u = \gamma y$ and then $w = \nu/(\nu+y^2)$ in $\int_a^\infty y^p (1+y^2/\nu)^{-(\nu+1)/2}\,\mathrm{d}y$, which becomes $\frac{1}{2}\nu^{(p+1)/2}\int_0^{w_a} w^{\frac{\nu-p}{2}-1}(1-w)^{\frac{p+1}{2}-1}\,\mathrm{d}w$, finite for $p < \nu$. Monotonicity and positivity hold because the integrand $\det(u\,\mathrm{Id} - r)$ is positive for $u > \lambda_{\max}(r)$. □

References

Animashree Anandkumar, Rong Ge, Daniel Hsu, Sham M. Kakade, and Matus Telgarsky. Tensor decompositions for learning latent variable models. *Journal of Machine Learning Research*, 15(1):2773–2832, 2014.

Diego Armentano, Jean-Marc Azaïs, and José R. León. On a general Kac–Rice formula for the measure of a level set. *The Annals of Applied Probability*, 35(3):1828–1851, 2025. https://doi.org/10.1214/25-AAP2158.

Antonio Auffinger, Gérard Ben Arous, and Jiří Černý. Random matrices and complexity of spin glasses. *Communications on Pure and Applied Mathematics*, 66(2):165–201, 2013. https://doi.org/10.1002/cpa.21422. arXiv:1003.1129v2.

Jean-Marc Azaïs and Mario Wschebor. *Level sets and extrema of random processes and fields*. John Wiley & Sons, 2009. https://doi.org/10.1002/9780470434642.

Jean-Marc Azaïs, Yohann De Castro, and Stéphane Mourareau. Power of the spacing test for least-angle regression. *Bernoulli*, 24(1):465–492, 2018. https://doi.org/10.3150/16-BEJ885.

Jean-Marc Azaïs, Federico Dalmao, and Yohann De Castro. Second maximum of a Gaussian random field and exact (t-)spacing test. *arXiv preprint arXiv:2406.18397v3*, 2024.

Jean-Marc Azaïs, Federico Dalmao, and Yohann De Castro. Non-asymptotic first- and second-order tail brackets for the Kostlan–Shub–Smale field and spherical k-spin complexity. *arXiv preprint arXiv:2611.09459*, 2026.

Jinho Baik, Gérard Ben Arous, and Sandrine Péché. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *The Annals of Probability*, 33(5):1643–1697, 2005. https://doi.org/10.1214/009117905000000233.

S. Banach. Über homogene Polynome in (L^2). *Studia Mathematica*, 7(1):36–44, 1938. https://doi.org/10.4064/sm-7-1-36-44.

Gérard Ben Arous, Song Mei, Andrea Montanari, and Mihai Nica. The landscape of the spiked tensor model. *Communications on Pure and Applied Mathematics*, 72(11):2282–2330, 2019. https://doi.org/10.1002/cpa.21861. arXiv:1711.05424v2.

Gérard Ben Arous, Cédric Gerbelot, and Vanessa Piccolo. Langevin dynamics for high-dimensional optimization: the case of multi-spiked tensor PCA. *arXiv preprint arXiv:2408.06401*, 2024a.

Gérard Ben Arous, Cédric Gerbelot, and Vanessa Piccolo. Stochastic gradient descent in high dimensions for multi-spiked tensor PCA. *arXiv preprint arXiv:2410.18162*, 2024b.

Shmuel Friedland and Lek-Heng Lim. Nuclear norm of higher-order tensors. *Mathematics of Computation*, 87(311):1255–1281, 2018. https://doi.org/10.1090/mcom/3239. arXiv:1410.6072v3.

Aukosh Jagannath, Patrick Lopatto, and Léo Miolane. Statistical thresholds for tensor PCA. *The Annals of Applied Probability*, 30(4):1910–1933, 2020. https://doi.org/10.1214/19-AAP1547. arXiv:1812.03403v2.

Erich L. Lehmann and Joseph P. Romano. *Testing Statistical Hypotheses*. Springer, 3rd edition, 2005.

Leiden Declaration. The Leiden Declaration on Artificial Intelligence and Mathematics. https://leidendeclaration.ai/, 2026. Endorsed by the International Mathematical Union.

Thibault Lesieur, Léo Miolane, Marc Lelarge, Florent Krzakala, and Lenka Zdeborová. Statistical and computational phase transitions in spiked tensor estimation. *Proceedings of the 2017 IEEE International Symposium on Information Theory (ISIT)*, pages 511–515, 2017. https://doi.org/10.1109/ISIT.2017.8006580.

M. A. Lifshits. On the absolute continuity of distributions of functionals of random processes. *Theory of Probability & Its Applications*, 27(3):600–607, 1983. https://doi.org/10.1137/1127066.

Amelia Perry, Alexander S. Wein, and Afonso S. Bandeira. Statistical limits of spiked tensor models. *Annales de l'Institut Henri Poincaré. Probabilités et Statistiques*, 56(1):230–264, 2020. https://doi.org/10.1214/19-AIHP960.

Emile Richard and Andrea Montanari. A statistical model for tensor PCA. In *Advances in Neural Information Processing Systems 27*, pages 2897–2905, 2014.

Eliran Subag. The complexity of spherical p-spin models—a second moment approach. *The Annals of Probability*, 45(5):3385–3450, 2017. https://doi.org/10.1214/16-AOP1139.

Jonathan E. Taylor, Joshua R. Loftus, and Ryan J. Tibshirani. Inference in adaptive regression via the Kac–Rice formula. *The Annals of Statistics*, 44(2):743–770, 2016. https://doi.org/10.1214/15-AOS1386.

Robert A. Wijsman. Cross-sections of orbits and their application to densities of maximal invariants. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 389–400. University of California Press, 1967.

Dong Xia, Anru R. Zhang, and Yuchen Zhou. Inference for low-rank tensors—no need to debias. *The Annals of Statistics*, 50(2):1220–1245, 2022. https://doi.org/10.1214/21-AOS2146.

Anru Zhang and Dong Xia. Tensor SVD: statistical and computational limits. *IEEE Transactions on Information Theory*, 64(11):7311–7338, 2018. https://doi.org/10.1109/TIT.2018.2841377.