

Geometry, Conditioning, and Parity in One-Sided Directional Curvature Recovery, with a Controlled Comparison of Estimators

RamaKrishna Pasupuleti

workisfun415@gmail.com

Kakatiya University <https://orcid.org/0009-0008-8418-1430>

Research Article

Keywords: Hessian recovery, one-sided finite differences, directional derivatives, spherical harmonics, condition number, spherical caps, derivative-free optimisation, step-size selection, benchmark methodology

Posted Date: August 5th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10592133/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Geometry, Conditioning, and Parity in One-Sided Directional Curvature Recovery, with a Controlled Comparison of Estimators

Ramakrishna Pasupuleti

Independent Researcher, Suryapet, Telangana, India

ORCID: 0009-0008-8418-1430 Email: workisfun415@gmail.com

Abstract

We study the recovery of a Hessian from directional curvature measurements taken along rays that must remain inside a prescribed feasible set, a situation that arises near domain boundaries, under active constraints in derivative-free optimisation, and wherever the complement of a domain cannot be probed. Four limits are established. First, an identifiability bound: without smoothness priors, no method using only point evaluations can determine a Hessian in $\mathbb{R}^3$ from fewer than ten evaluations at points in general position, and the correct uniqueness criterion for directional measurements is that the sampled directions do not lie on a common quadric cone. Second, a conditioning law: when the admissible directions are confined to a spherical cap of half-angle θ , the recovery operator has singular values of orders $\{1, \theta, \theta, \theta^2, \theta^2, \theta^2\}$, so its condition number grows as θ^{-2} ; the constants are derived exactly and depend on the sampling measure while the exponents do not. Third, a parity decomposition: the one-sided probe splits into an even part carrying the quadratic directional curvature and an odd part carrying the leading truncation term, which yields a measurement of the third-order directional derivative that is independent of the Hessian and accurate to $O(h^2)$. Fourth, an exact characterisation of attainable order: second-order accuracy is available if and only if the feasible direction set is antipodally symmetric, independently of how many directions are used. We then report a controlled comparison of eight estimators under a fairness protocol in which every method receives its own oracle-optimal sampling radius, and document the nine experimental artifacts that had to be removed before that comparison became trustworthy. All results concern any method that estimates curvature from one-sided directional measurements; no particular estimator is privileged.

Keywords: Hessian recovery; one-sided finite differences; directional derivatives; spherical harmonics; condition number; spherical caps; derivative-free optimisation; step-size selection; benchmark methodology.

MSC 2020: 65D25; 65F35; 41A25; 90C56; 65Y20.

Epistemic conventions.

Statements labelled Theorem, Proposition, Lemma or Corollary are proved. Statements labelled Numerical Observation are supported by computation only and are flagged as such wherever they are used; we assert no proofs for them. Open positioning questions — places where a result may turn out to be an instance of an existing theory — are stated explicitly in Section the corresponding section, and open mathematical questions in Section the corresponding section. We have tried to keep these three categories visibly distinct throughout, including in the abstract.

PART: Introduction and context

Introduction

The problem

Recovering second-order information from function values sampled along rays emanating from a single point arises in derivative-free optimisation, in finite-difference approximation near boundaries, in tensor reconstruction from directional encodings, and in inverse problems where the forward map is available only as a black box. In each of these settings the sampling is constrained. The evaluation point may sit within one step of a boundary, so that a symmetric stencil cannot be formed. It may sit at a point where several inequality constraints are active, so that only an acute cone of directions leads back into the feasible region. It may sit inside a thin structure, so that the usable span depends strongly on direction. In each case the admissible set of directions, and the admissible span along each of them, is dictated by geometry rather than chosen.

For *unrestricted* directions the design question is essentially settled. The optimal placement of six directions for recovering a symmetric second-order tensor is known from the diffusion tensor imaging literature: the icosahedral scheme is optimal, its condition number is exactly $1/2 \sqrt{10} \approx 1.58$, and several apparently different optimality criteria coincide with it [9, 10, 15]. Much less is known about what happens when the admissible directions are confined to a cone: which quantities remain identifiable at all, how fast conditioning degrades with the aperture, what order of accuracy remains attainable, and what auxiliary information the measurement still carries. This paper studies those limits.

What is and is not claimed

Our results are statements about the measurement geometry rather than about any particular estimator. They therefore bound one-sided finite-difference stencils, divided-difference schemes, directional probes inside trust-region methods, and directional variants of local polynomial regression alike. We do not propose a new estimator, and we do not claim that any estimator is superior; indeed Part the corresponding section reports the opposite for the directional schemes we tested. The value of the results, if they have any, is that they constrain what any such method can achieve, and that they identify one piece of structure in the measurement that is not usually exploited.

Contributions

1. **Identifiability** (Theorem 4, Proposition 5). Without priors, ten evaluations at points in general position are necessary and sufficient in $\mathbb{R}^3$. For directional measurements, the sharp uniqueness criterion is that the sampled directions do not lie on a common quadric cone; spanning $\mathbb{R}^3$ is neither necessary in form nor sufficient in substance.

2. **Conditioning under restricted access** (Theorem 10). For a cap of half-angle θ the singular values of the recovery operator are of orders $\{1, \theta, \theta, \theta^2, \theta^2, \theta^2\}$, and $\kappa(\theta) \asymp \theta^{-2}$ with constants computed exactly for two natural sampling measures.
3. **Parity decomposition** (Propositions 13 and 14, Corollary 15). The one-sided probe expands into even and odd spherical-harmonic content whose leading terms are the quadratic directional curvature and the third-order directional derivative respectively. The odd projection therefore measures the truncation term without knowledge of the Hessian, to $O(h^2)$.
4. **Attainable order** (Proposition 17). Second-order accuracy in the aggregated estimate is attainable if and only if the feasible direction set is antipodally symmetric, independently of the number of directions, the weighting scheme, or the direction design.
5. **Noise** (Propositions 19 and 21, Corollary 22). Exact variance for the probe and for the debiased truncation estimator, and the resulting lower bound on a usable pilot span, which scales as $\sigma^{1/3}$ in the same way as the optimal span itself.
6. **A structural summary** (Corollary 23). Geometry, truncation and noise limit accuracy through three mutually independent mechanisms that cannot be traded against one another.
7. **A controlled comparison and its methodology** (Part the corresponding section). Eight estimators under a fairness protocol, with the nine experimental artifacts that had to be removed before the comparison could be trusted, grouped by cause.

Organisation

Part the corresponding section treats geometry and conditioning, Part the corresponding section the parity decomposition, Part the corresponding section noise, Part the corresponding section algorithms and design rules, Part the corresponding section numerical verification, and Part the corresponding section the estimator comparison and the methodological lessons. Section the corresponding section positions the results against existing literature and states where they may not be new. Section the corresponding section states what the theorems do not say, and Section the corresponding section lists open questions.

Standing assumptions.

Throughout, $\Omega \subseteq \mathbb{R}^3$ is open, $x \in \Omega$ is fixed, and $f \in C^5(\mathcal{N})$ on a neighbourhood $\mathcal{N} \subseteq \Omega$ of x containing all sample points used. The C^5 requirement is needed only from Part the corresponding section onwards; C^3 suffices before that, and Theorem 4 requires no smoothness at all beyond the statement being about quadratics. Function values are exact except where a noise model is stated, in which case each evaluation carries an independent $N(0, \sigma^2)$ error, with σ quoted relative to $|f(x)| + 1$. Projections onto spherical-harmonic subspaces are taken in $L^2(S^2)$ with the normalised surface measure unless a cap C_θ is

named, and every statement about C_θ specifies the measure used, since constants — though not exponents — depend on it.

Relationship to existing work

The results below concern the measurement geometry rather than an estimator, so they intersect several literatures. We state the relationship explicitly, including where the present results are *not* new.

Summary table

Existing approach		Covered	Relation to the present results
One-sided differences	finite	Yes	Theorem 10 supplies conditioning when the usable directions are restricted; classical analyses treat a fixed stencil rather than the admissible direction set.
Divided differences		Yes	Equation (1) is a second divided difference; no novelty is claimed for the estimator. Proposition 14 concerns the structure of the measurement across directions, not the stencil.
Directional Hessian recovery		Yes	Proposition 5 replaces a spanning condition by the quadric-cone criterion; Theorem 4 gives the evaluation floor.
DTI reconstruction	tensor	<i>Prior art</i>	The full-sphere direction-design problem is solved here [9, 10, 15, 1]. We adopt these results and claim nothing.
Spherical theory	Slepian	<i>Positioning risk</i>	Restricted harmonic analysis on polar caps is developed [14, 11]. Observation 32 is plausibly an instance of it; see Section the corresponding section.
Quadratic regression, MLS		Partially	Theorems 4 and 10 bound every method using point evaluations, these included [13, 3]; the parity results apply only to schemes sampling along rays from one centre.
Step-size selection		Partially	Estimating the truncation constant is classical, but by comparing estimates across step sizes [16, 7, 8]. Corollary 15 obtains it at a <i>single</i> step size from angular structure.
Derivative-free optimisation		Partially	Δ -poisedness and model error bounds are established [4, 5, 6, 12, 2]; we adopt them as the fairness standard rather than proposing an alternative.

Direction design is prior art

For unrestricted directions the design question is settled. Hasan, Parker and Alexander showed that the icosahedral six-direction scheme is optimal for second-order tensor encoding and that minimum-variance, minimum-force and minimum-energy criteria all reduce to it [9]; the associated condition number $1/2\sqrt{10}$ is standard, and K -optimal design has since been extended to higher-order tensors by semidefinite relaxation [1]. The electrostatic-repulsion construction of Jones et al. [10] and the condition-number analysis of Skare et al. [15] belong to the same literature, as does the subsequent debate over whether six directions suffice or twenty to thirty are needed for robust anisotropy and orientation estimates. Nothing in the present paper improves on this, and the direction sets used in our experiments are taken from it.

The spherical-cap results and Slepian theory

Observation 32 states that the even and odd polynomial spans become numerically indistinguishable on a small cap. This is the regime studied by the spherical spatio-spectral concentration problem of Simons, Dahlen and Wieczorek [14], which characterises how many bandlimited functions can be well concentrated in a region — the Shannon number — and how the remaining concentration eigenvalues decay. The polar cap is the case with the most complete theory, since the concentration operator then commutes with a Sturm–Liouville operator. Related eigenvalue-distribution results go back to Landau and Widom [11].

We regard it as likely that Observation 32 is a corollary of that theory rather than an independent finding, and we do not claim otherwise. Two questions should be settled before any priority is asserted: whether the empirically observed θ^9 decay follows from known concentration-eigenvalue asymptotics for a band containing degrees 1 to 4, and whether the regularised estimator of Section the corresponding section differs materially from a Slepian-basis truncation. If it does not, the correct claim is the application of concentration theory to truncation-error estimation, which is narrower but still, as far as we are aware, not previously made.

Estimating the truncation constant

Practical step-size selection has long estimated the truncation term empirically. Curtis and Reid [7], Dumontet and Vignes [8] and Stepleman and Winarsky [16] all compare estimates computed at two or more step sizes and infer the leading coefficient from how they differ; modern numerical-differentiation packages implement variants of these procedures. Corollary 15 is an alternative mechanism rather than an alternative objective: the coefficient is read from the angular parity structure at a single step size. Whether this is preferable in practice depends on the relative cost of extra directions versus extra step sizes, and Corollary 22 shows the directional route is not cheap.

Derivative-free optimisation

The governing notion in this literature is Λ -poisedness of a sample set, introduced by Conn, Scheinberg and Vicente [4, 5], together with the fully-linear and fully-quadratic model error bounds and the geometry-improving algorithms that maintain them [6]; the survey of Larson, Menickelly and Wild [12] gives the current state, and comparisons of gradient-approximation schemes under a common budget have been carried out by Berahas, Cao, Choromanski and Scheinberg [2]. Theorem 4 is consistent with, and can be read as a sharpened statement of, the counting condition underlying poisedness for quadratics; the contribution here is the explicit consequence for directional schemes and the quadric-cone criterion of Proposition 5. We adopt Λ -poisedness as the fairness standard for the comparison baselines in Part the corresponding section.

Boundary effects in local quadratic fitting

Chen and Wu [3] have shown that the local quadratic-fitting estimator of the Hessian operator converges asymptotically under mild conditions on a manifold with boundary, with non-uniform sampling and curvature effects negligible even near boundaries. This is directly relevant: it establishes that the natural competitor to a directional scheme does not degrade near a boundary in the asymptotic regime, so any claim that one-sided directional sampling is preferable near boundaries must be made in the finite-sample, finite-noise regime rather than asymptotically. Our numerical results in Part the corresponding section are consistent with their conclusion.

Summary of novelty claims

We have not found, in the literature surveyed, results establishing the following; we list them in decreasing order of our confidence that they are new, and we make no priority claim where a positioning question above is unresolved. (i) the evaluation floor of Theorem 4 stated for directional schemes, together with Proposition 5; (ii) the cap-restricted conditioning law of Theorem 10 with explicit constants; (iii) the parity decomposition and the resulting Hessian-free measurement of third-order structure, Corollary 15; (iv) the antipodal-availability characterisation of attainable order, Proposition 17. Claim (ii) should be checked against the Slepian literature and against results on restricted Veronese embeddings in algebraic geometry; claim (iii) against the step-size-selection literature. We claim nothing regarding direction design. Where a claim's status depends on unresolved overlap, we prefer to state the overlap as an open positioning question rather than assert priority.

PART: Geometry and conditioning

Setting

Let $H = \nabla^2 f(x) \in \text{Sym}_3(\mathbb{R})$ and, for a unit vector $u \in S^2$,

$$D^k f(u) = \sum_{i_1, \dots, i_k} \partial_{i_1 \dots i_k} f(x) u_{i_1} \cdots u_{i_k}.$$

Definition 1 (One-sided directional probe). For a span $h > 0$, an interior node parameter $R \in (0,1)$ and a direction $u \in S^2$, the probe uses the three evaluations $\{x, x + Rhu, x + hu\}$, all on the segment $[x, x + hu]$, and forms

$$q(u, h) = \frac{2}{R(1-R)h^2} [(1-R)f(x) + Rf(x + hu) - f(x + Rhu)].$$

Remark 2 (No novelty in the estimator). Expression (1) equals $2f[x, x + Rhu, x + hu]$, twice the second divided difference on the non-uniform one-sided node set; this is verified in Appendix the corresponding section. We claim no novelty for it. All results below concern the geometry of the measurement process and apply to any scheme producing estimates of $u^\top Hu$ from rays emanating from x .

Definition (Admissible set and span). The admissible direction set is $U(x) = \{u \in S^2 : x + tu \in \Omega \text{ for all } t \in [0, h_{\max}(x, u)]\}$ with $h_{\max}(x, u) = \sup\{t > 0 : x + su \in \Omega \ \forall s \in [0, t]\}$. A cap of half-angle θ about an axis a is $C_\theta = \{u \in S^2 : u \cdot a \geq \cos\theta\}$.

Identifiability

Theorem 4 (Identifiability floor). Let $y_1, \dots, y_n \in \mathbb{R}^3$ be sample points not contained in a common plane. If $n \leq 9$ there exist quadratic polynomials $p \neq p'$ with $p(y_i) = p'(y_i)$ for all i and $\nabla^2 p \neq \nabla^2 p'$; consequently no map from the data $\{f(y_i)\}$ to $\text{Sym}_3(\mathbb{R})$ returns $\nabla^2 f$ correctly for every quadratic f . If $n \geq 10$ and the points are poised for quadratics, $\nabla^2 f$ is determined.

Proof. The space $\mathcal{P}_2$ of real quadratics on $\mathbb{R}^3$ has dimension 10. The evaluation map $L: \mathcal{P}_2 \rightarrow \mathbb{R}^n$, $Lp = (p(y_1), \dots, p(y_n))$, satisfies $\dim \ker L \geq 10 - n$. Suppose $n \leq 9$, so that $\ker L \neq \{0\}$, and pick $0 \neq q \in \ker L$. If $\nabla^2 q = 0$ then q is affine and vanishes at every y_i , forcing the points into the plane $\{q = 0\}$, contrary to hypothesis. Hence $\nabla^2 q \neq 0$, and p and $p' = p + q$ are indistinguishable in the data while having distinct Hessians. For $n \geq 10$ and a poised set, L is injective and ∇^2 factors through L . $\square$

Proposition 5 (Directional uniqueness). Let $U = \{u_i\} \subset S^2$ and suppose $q_i = u_i^\top Hu_i$ is observed. Then H is uniquely determined if and only if $\{u_i u_i^\top\}$ spans $\text{Sym}_3(\mathbb{R})$; equivalently, if and only if there is no $0 \neq M \in \text{Sym}_3(\mathbb{R})$ with $u_i^\top M u_i = 0$ for every i , that is, the directions do not all lie on a single quadric cone.

Proof. The measurement map is $H \mapsto (\langle u_i u_i^\top, H \rangle)_i$ with the trace inner product on $\text{Sym}_3(\mathbb{R})$. It is injective exactly when no nonzero M is orthogonal to every $u_i u_i^\top$, which is the stated condition. Since $\dim \text{Sym}_3(\mathbb{R}) = 6$, at least six directions are required. $\square$

Remark . Spanning $\mathbb{R}^3$ is not sufficient: $\{e_1, e_2, e_3\}$ spans $\mathbb{R}^3$ and determines only the diagonal of H . Nor is a set of six directions automatically adequate: any six coplanar

directions lie on a degenerate quadric cone and fail. A numerical instance is recorded in Observation 31.

Corollary 7 (Evaluation cost of directional schemes). A directional scheme using m directions and two new evaluations per direction costs $1 + 2m$ evaluations. Since $m \geq 6$ by Proposition 5, such a scheme costs at least 13 evaluations, exceeding the floor of 10 in Theorem 4.

Remark 8. Corollary 7 compares evaluation counts for exact recovery in the absence of priors. It does *not* assert that directional schemes are suboptimal in practice: smoothness priors, noise, feasibility constraints, and the target functional — a full Hessian, a curvature along one direction, a descent direction — each change what optimality means. Part the corresponding section gives one concrete setting in which the directional measurement carries information that a minimal evaluation set does not.

Remark (Three regimes). Theorem 4 concerns exact recovery with no prior. Under a smoothness prior, for instance $\|D^3f\|_\infty \leq M_3$, estimates from $n < 10$ evaluations have error controlled by M_3 rather than being undetermined; and under noise the operative question is variance rather than identifiability. Exact identifiability, practical identifiability under noise, and recovery under smoothness priors should not be conflated.

Conditioning under a restricted cone

Let $\mathcal{C}_\theta$ carry a rotationally symmetric probability measure μ_θ and let

$$\mathcal{A}_\theta: \text{Sym}_3(\mathbb{R}) \rightarrow L^2(\mathcal{C}_\theta, \mu_\theta), \quad (\mathcal{A}_\theta H)(u) = u^\top H u.$$

We use the Frobenius-consistent coordinates $(H_{11}, H_{22}, H_{33}, \sqrt{2}H_{12}, \sqrt{2}H_{13}, \sqrt{2}H_{23})$, in which the coordinate 2-norm equals $\|H\|_F$.

Theorem 10 (Cone conditioning). As $\theta \rightarrow 0$ the six singular values of $\mathcal{A}_\theta$ satisfy

$$s_1 \asymp 1, \quad s_2, s_3 \asymp \theta, \quad s_4, s_5, s_6 \asymp \theta^2,$$

so that $c_1\theta^{-2} \leq \kappa(\mathcal{A}_\theta) = s_1/s_6 \leq c_2\theta^{-2}$ for absolute constants $0 < c_1 \leq c_2$. The constants are explicit and depend on the measure. For the normalised surface measure,

$$s_1 \rightarrow 1, \quad \frac{s_2}{\theta} \rightarrow \frac{1}{\sqrt{2}}, \quad \frac{s_6}{\theta^2} \rightarrow \frac{1}{\sqrt{24}}, \quad \kappa(\mathcal{A}_\theta) \theta^2 \rightarrow \sqrt{24} \approx 4.899;$$

for the measure uniform in the polar angle φ , $s_2/\theta \rightarrow 1/\sqrt{3}$, $s_6/\theta^2 \rightarrow \sqrt{2/45}$ and $\kappa\theta^2 \rightarrow \sqrt{45/2} \approx 4.743$. The exponents are measure-independent; the constants are not.

Proof. Parametrise $u = (\varepsilon \cos \psi, \varepsilon \sin \psi, \sqrt{1 - \varepsilon^2})$ with $\varepsilon = \sin \varphi \in [0, \sin \theta]$ and $\psi \in [0, 2\pi)$, and write $n = (\cos \psi, \sin \psi)$. Block-decompose

$$H = \begin{bmatrix} T & b \\ b^\top & c \end{bmatrix}, \quad T \in \text{Sym}_2(\mathbb{R}), \quad b \in \mathbb{R}^2, \quad c \in \mathbb{R},$$

and split $T = \tau I_2 + T_0$ with $\text{tr} T_0 = 0$. Expanding,

$$u^\top H u = \underbrace{c}_{(i)} + 2\varepsilon \underbrace{\sqrt{1-\varepsilon^2}}_{(ii)} (b^\top n) + \varepsilon^2 \underbrace{(\tau - c)}_{(iii)} + \varepsilon^2 \underbrace{(n^\top T_0 n)}_{(iv)}.$$

The four groups carry distinct ψ -harmonic content: (i) and (iii) are ψ -independent, (ii) is a combination of $\cos\psi$ and $\sin\psi$, and (iv) of $\cos 2\psi$ and $\sin 2\psi$. Since $\{1, \cos\psi, \sin\psi, \cos 2\psi, \sin 2\psi\}$ are orthogonal in $L^2([0, 2\pi))$ and μ_θ is a product of a φ -measure and the uniform ψ -measure, the groups (i)+(iii), (ii) and (iv) are mutually orthogonal in $L^2(C_\theta, \mu_\theta)$; within the first group the ε^0 and ε^2 parts are linearly independent for $\theta > 0$. The parameter counts are $1 + 2 + 1 + 2 = 6 = \dim \text{Sym}_3(\mathbb{R})$, so (2) exhibits an orthogonal decomposition of the domain adapted to $\mathcal{A}_\theta$.

For the normalised surface measure the density of φ on C_θ is $\sin\varphi/(1 - \cos\theta)$; writing $r = \varphi/\theta$ this tends to $2r$ on $[0, 1]$ as $\theta \rightarrow 0$, whence

$$\langle \varepsilon^2 \rangle \rightarrow \frac{\theta^2}{2}, \quad \langle \varepsilon^4 \rangle \rightarrow \frac{\theta^4}{3}.$$

Since $\int \varepsilon^{2k} d\mu_\theta \asymp \theta^{2k}$, the orders follow: $\|\mathcal{A}_\theta H\| \asymp |c|$ in group (i), $\asymp \theta|b|$ in group (ii), and $\asymp \theta^2 \|\cdot\|$ in groups (iii) and (iv); orthogonality prevents cancellation between groups.

For the constants, normalise $\|H\|_F = 1$. In group (ii) only b is nonzero, so $\|H\|_F^2 = 2|b|^2$ and

$$\|\mathcal{A}_\theta H\|^2 = 4\langle \varepsilon^2 \rangle \langle (b^\top n)^2 \rangle_\psi = 4 \cdot \frac{\theta^2}{2} \cdot \frac{|b|^2}{2} = \theta^2 |b|^2,$$

giving $s_2 = s_3 = \theta/\sqrt{2}$. In group (iv), writing $n^\top T_0 n = a \cos 2\psi + b_0 \sin 2\psi$ with $\|T_0\|_F^2 = 2(a^2 + b_0^2)$,

$$\|\mathcal{A}_\theta H\|^2 = \langle \varepsilon^4 \rangle \cdot \frac{\|T_0\|_F^2}{4} = \frac{\theta^4}{12},$$

giving $\theta^2/\sqrt{12}$. In group (iii), setting $d = \tau - c$,

$$\|\mathcal{A}_\theta H\|^2 = c^2 + cd\theta^2 + \frac{d^2\theta^4}{3} = \left(c + \frac{d\theta^2}{2}\right)^2 + \frac{d^2\theta^4}{12},$$

while $\|H\|_F^2 = 2\tau^2 + c^2 \rightarrow 2d^2$ as $\theta \rightarrow 0$; minimising over unit H gives $\theta^4/24$, so $s_6 = \theta^2/\sqrt{24}$, which is smaller than the group (iv) value $\theta^2/\sqrt{12}$. Since $s_1 \rightarrow 1$ the stated limits follow. The polar-uniform case is identical with $\langle \varepsilon^2 \rangle \rightarrow \theta^2/3$ and $\langle \varepsilon^4 \rangle \rightarrow \theta^4/5$, giving $s_2 = \theta/\sqrt{3}$ and $s_6 = \theta^2\sqrt{2/45}$. $\square$

Remark (Interpretation). The three groups correspond to normal curvature $c = a^\top H a$, mixed curvature b , and tangential curvature T . A narrow cone measures normal curvature well, mixed curvature at first order in θ , and tangential curvature only at second order. This is intuitive: to see curvature in a direction one must move in that direction, and a narrow cone barely does.

Remark 12 (A smooth boundary is not a narrow cone). At a smooth boundary the admissible set is essentially a hemisphere, $\theta \approx 90^\circ$, where $\kappa \approx 2$ and there is no conditioning problem. Narrow cones arise at corners and cusps, in thin structures, and at points where several inequality constraints are simultaneously active. This should be kept in mind when reading Theorem 10 as a statement about boundaries.

PART: Parity

The parity decomposition

Proposition 13 (Probe expansion). Let $f \in C^5$ near x and $R \in (0,1)$. Then

$$q(u, h) = D^2 f(u) + \frac{(1+R)}{3} h D^3 f(u) + \frac{(1+R+R^2)}{12} h^2 D^4 f(u) + O(h^3).$$

Proof. Substituting the Taylor expansions of $f(x+hu)$ and $f(x+Rhu)$ into (1), the coefficient of $f(x)$ in the bracket is $(1-R) + R - 1 = 0$ and that of $D^1 f(u)$ is $Rh - Rh = 0$. For $k \geq 2$ the coefficient of $D^k f(u)$ in the bracket is $h^k(R - R^k)/k!$, and $R - R^k = R(1-R)(1+R+\dots+R^{k-2})$. Dividing by $R(1-R)h^2/2$ gives $2/k! h^{k-2}(1+R+\dots+R^{k-2})$, which is 1 for $k=2$, $(1+R)h/3$ for $k=3$ and $(1+R+R^2)h^2/12$ for $k=4$. $\square$

Proposition 14 (Parity separation). $D^k f(\cdot)$ is a homogeneous polynomial of degree k on S^2 , so its expansion in spherical harmonics contains only degrees $\ell \equiv k \pmod{2}$ with $\ell \leq k$. Consequently, with P_{odd} and P_{even} the $L^2(S^2)$ -projections onto odd and even harmonics,

$$P_{\text{odd}} q(\cdot, h) = \frac{(1+R)}{3} h D^3 f(\cdot) + O(h^3), \quad P_{\text{even}} q(\cdot, h) = D^2 f(\cdot) + O(h^2).$$

Proof. Immediate from Proposition 13 together with the parity of the terms: $D^2 f$ and $D^4 f$ are even and $D^3 f$ is odd. The first neglected odd contribution is the $h^3 D^5 f$ term, whence the $O(h^3)$ remainder in the odd projection; the first neglected even contribution is $h^2 D^4 f$, whence $O(h^2)$ in the even projection. $\square$

Corollary 15 (Hessian-free measurement of third-order structure). Define

$$\hat{C}_3(h) = \frac{3}{(1+R)h} \| P_{\text{odd}} q(\cdot, h) \|_{L^2(S^2)}.$$

Then $\hat{C}_3(h) = \| D^3 f \|_{L^2(S^2)} + O(h^2)$, independently of H .

Remark . The even projection should not be described as isolating the Hessian: it recovers the quadratic directional curvature to $O(h^2)$, the gain of one order over the raw probe being exactly what parity buys. Symmetrically, the odd projection recovers the third-order term to $O(h^2)$ after division by h .

Antipodal availability and attainable order

Proposition 17 (Order characterisation). Let $U \subset S^2$ be finite with $U = -U$, and let $\hat{H}$ solve $\min_H \sum_{u \in U} w_u (u^\top H u - q(u, h))^2$ with $w_{-u} = w_u$. Then the $O(h)$ term of Proposition 13 contributes nothing to $\hat{H}$ and $\hat{H} = H + O(h^2)$. If U contains no antipodal pair the $O(h)$ term is not annihilated and the estimate is in general $H + O(h)$, independently of $|U|$.

Proof. The normal equations read $(\sum_u w_u \text{vec}(uu^\top) \text{vec}(uu^\top)^\top) \hat{h} = \sum_u w_u \text{vec}(uu^\top) q(u, h)$. Writing $q(u, h) = u^\top H u + (1 + R)h/3 D^3 f(u) + O(h^2)$ and noting that $u \mapsto uu^\top$ and $u \mapsto w_u$ are even while $u \mapsto D^3 f(u)$ is odd, the contribution of the $O(h)$ term to the right-hand side is $(1 + R)h/3 \sum_u w_u \text{vec}(uu^\top) D^3 f(u)$, in which the terms for u and $-u$ cancel exactly. The remaining error is the even $O(h^2)$ term. Conversely, if no antipodal pairs are present the sum does not cancel and is generically nonzero. $\square$

Remark (What this does and does not give). Antipodal availability is precisely the situation in which a symmetric stencil could have been formed, so the second order recovered here is the classical second order of central differences, obtained in a different arrangement. Stated as a limitation the proposition is sharper: under genuinely one-sided access, no increase in the number of directions, no reweighting, and no direction design recovers the lost order.

PART: Noise

Noise propagation

Assume each evaluation carries an independent $N(0, \sigma^2)$ error and set $v^2 = (1 - R)^2 + R^2 + 1$.

Proposition 19 (Variance of the probe). The noise standard deviation of $q(u, h)$ is $s_q = \frac{2\sigma v}{R(1-R)h^2}$.

Proof. The bracket in (1) is a linear combination of three independent evaluations with coefficients $(1 - R)$, R and -1 , of squared norm v^2 ; multiplying by $2/(R(1 - R)h^2)$ gives the stated value. $\square$

Remark 20 (Choice of R). The noise factor $1/(R(1 - R))$ is minimised at $R = 1/2$, while the truncation coefficient $(1 + R)/3$ of Proposition 13 decreases as $R \rightarrow 0$. The optimum therefore moves towards $R = 1/2$ as noise grows and towards smaller R as noise vanishes; Observation 37 quantifies this.

Proposition 21 (Debiased truncation estimator). Let $p = \dim \mathcal{H}_1 + \dim \mathcal{H}_3 = 3 + 7 = 10$ and let the projection be computed from m sampled directions. With $T = \|P_{\text{odd}} q\|^2$ and $\kappa_h = 3/((1 + R)h)$, the estimator

$$\hat{C}_3^2 = \kappa_h^2 \max(T - p s_q^2/m, 0)$$

is unbiased for C_3^2 before truncation at zero, and

$$\text{Var}(\hat{C}_3) \approx \frac{\kappa_h^2 s_q^2}{m} \left(1 + \frac{p s_q^2}{2m \|P_{\text{odd}} q\|^2} \right).$$

Proof. T is a scaled non-central χ_p^2 variable with non-centrality proportional to the true odd energy, so $\mathbb{E}[T] = \|P_{\text{odd}} q\|^2 + p s_q^2/m$ and $\text{Var}[T] = 2p s_q^4/m^2 + 4s_q^2 \|P_{\text{odd}} q\|^2/m$. Subtracting the bias gives unbiasedness in the square; applying the delta method to $t \mapsto \sqrt{t}$ gives the stated variance. $\square$

Corollary 22 (Pilot design conditions). Since $\kappa_h s_q = \frac{6\sigma v}{(1+R)R(1-R)h^3}$, a usable estimate requires

$$h \gtrsim \left[\frac{6\sigma v \sqrt{p/m}}{(1+R)R(1-R) C_3} \right]^{1/3},$$

the same cube-root scaling in σ as the optimal span itself. Moreover the number of sampled directions must exceed the dimension of the basis used for the projection, and increasing m improves the estimate only as $m^{-1/6}$.

Proof. The condition is $\kappa_h s_q \sqrt{p/m} \lesssim C_3$, which rearranges as stated. The rank condition is required for the projection to be determined; the $m^{-1/6}$ rate follows from $\text{Var}(\hat{C}_3) \propto m^{-1}$ inside the cube root of the span rule. $\square$

Corollary 23 (Three independent sources of difficulty). For a scheme sampling along rays from a single centre, recovery accuracy is limited independently by

1. *geometric conditioning*, governed by the admissible aperture through Theorem 10, independent of h and σ ;
2. *truncation*, governed by the odd-order term of Proposition 13, of order h , independent of σ and of the direction set;
3. *measurement noise*, of order σh^{-2} by Proposition 19, independent of the aperture.

No choice of direction set removes (ii) or (iii); no choice of span removes (i) or alters the trade-off between (ii) and (iii); no reduction of σ removes (i) or (ii).

Remark . Corollary 23 is the organising principle of this paper: Part the corresponding section bounds mechanism (i), Part the corresponding section characterises and measures (ii), Part the corresponding section quantifies (iii). It also explains the shape of the results in Part the corresponding section: a method that improves the handling of one mechanism cannot thereby overcome a deficit in another.

PART: Algorithms and design rules

Restricted caps and regularised parity inversion

On a cap the parity projection is no longer orthogonal, since $u \in C_\theta$ does not imply $-u \in C_\theta$. Separation of even from odd then rests on linear independence of the restricted polynomial bases, which degenerates as $\theta \rightarrow 0$ (Observation 32).

Let B be the design matrix of a real polynomial basis of degree ≤ 3 evaluated at the sampled directions, and q the vector of probe values.

Definition 25 (Regularised parity inversion). For $\lambda \geq 0$ set $\hat{c} = (B^\top B + \lambda \|B^\top B\|_2 I)^{-1} B^\top q$, extract the coefficients of odd degree, and form $\hat{C}_3$ from the resulting odd energy by Proposition 21.

Algorithm 26 (Span selection with a harmonic pilot). *Input:* evaluation point x , admissible set $U(x)$, noise level σ , pilot span h_0 , pilot direction count $m_0 > \dim(\text{basis})$, relative ridge parameter λ_{rel} .

1. Sample m_0 directions inside $U(x)$ and evaluate the probe at span h_0 .
2. Form $\hat{C}_3$ by Definition 25 with the noise-floor subtraction of Proposition 21.
3. For each working direction u , set $h(u) = \min(h_{\text{max}}(x, u), (2C_\sigma/C_3)^{1/3})$ with $C_\sigma = 2\nu\sigma/(R(1-R))$ and $C_3 = (1+R)\hat{C}_3/3$.
4. Recover H by weighted least squares with weights $w_u = (C_3^2 h(u)^2 + s_q(u)^2)^{-1}$.

Cost: $1 + 2m_0$ evaluations for the pilot and $2m$ for the working directions.

Remark (The pilot is not cheap). Corollary 22 and Observation 36 together imply a pilot of at least about 18 directions, hence about 37 evaluations, against the 13 needed for the estimate it calibrates. Algorithm 26 is therefore worthwhile only when the pilot can be amortised — gridded data where C_3 is estimated once per region, or trust-region iterations in which C_3 drifts slowly.

Practical design rules

Collected for reference; each is supported by a numbered statement or observation.

1. **Node placement.** Use $R = 1/2$ at high noise and $R \approx 0.3$ at low noise (Remark 20, Observation 37). The endpoints $R \leq 0.1$ and $R \geq 0.9$ cost a factor of two to three.
2. **Direction sets.** On the full sphere use the icosahedral six-axis scheme; use Fibonacci sets only for $m \geq 12$, since at $m = 6$ they are rank deficient (Observation 31).

3. **Under constraints, never use a fixed set.** A half-space removes half of an antipodally symmetric set, dropping it below the six directions required by Proposition 5. Construct the set inside the feasible region, for instance by sequential D -optimal selection (Observation 38).
4. **Aperture.** $\kappa \approx 4.9\theta^{-2}$; a hemisphere is benign, a 20° cone costs a factor of about 40 in conditioning relative to a hemisphere (Theorem 10).
5. **Order.** Expect first order under one-sided access and second order only where antipodal partners are available (Proposition 17, Observation 30).
6. **Span.** Balance $C_3 h$ against σh^{-2} ; the optimum scales as $\sigma^{1/3}$ and must be clipped at $h_{\max}(x, u)$, which makes it direction-dependent near a boundary.
7. **Pilot.** Use at least 18 directions and a degree-3 basis; apply the noise-floor subtraction; take any $\lambda_{\text{rel}} \in [10^{-5}, 10^{-2}]$ (Observation 33).

PART: Numerical verification

Conventions

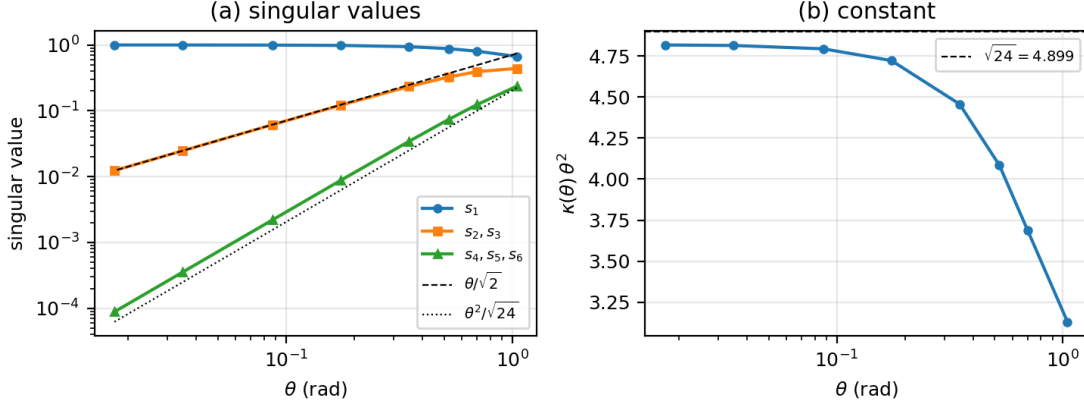
$\text{Sym}_3(\mathbb{R})$ carries the Frobenius-consistent coordinates of Section the corresponding section. Caps are sampled on a tensor grid in (φ, ψ) ; *area-uniform* means $\cos\varphi$ equispaced and *polar-uniform* means φ equispaced. The default probe uses $R = 1/2$. The evaluation point is $x = (0.10, -0.05, 0.20)$ throughout. Reported digits are those stable under a doubling of the grid in both coordinates unless stated otherwise. Test functions are

$$\begin{aligned}
 f_1 &= \exp(-\|x - c\|^2 / 2s^2), \quad c = (0.3, -0.2, 0.15), \quad s = 0.8; \\
 f_2 &= \exp(x + y + z); \\
 f_3 &= x^4 + 2x^2y^2 + y^3z + z^4 + xyz; \\
 f_4 &= \sin(x + 2y) + \cos(2y - z) + \sin(xz).
 \end{aligned}$$

Analytic Hessians are verified against fourth-order central differences at random points to a maximum discrepancy of 8.5×10^{-6} before any experiment runs; this check is part of the test suite.

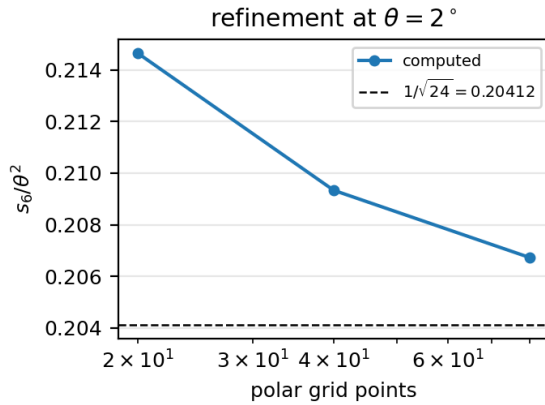
Verification of the theorems

Each figure verifies exactly one statement.

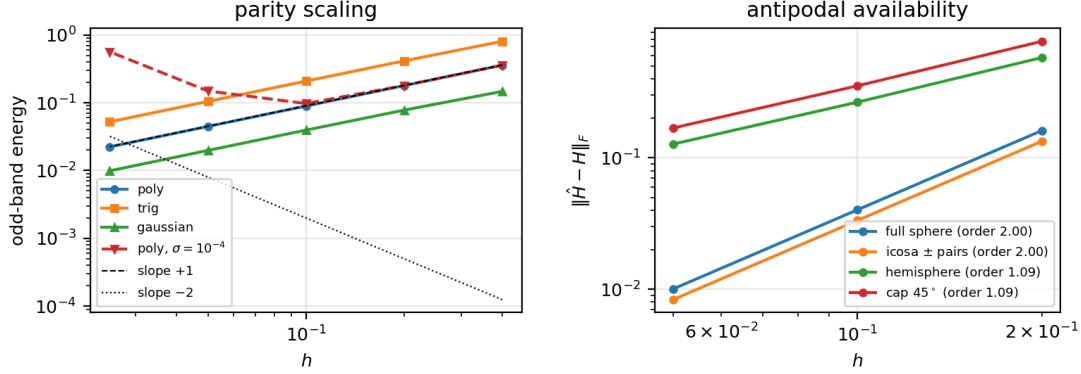


Verification of Theorem 10. (a) The six singular values of $\mathcal{A}_\theta$ against θ with the predicted envelopes $\theta/\sqrt{2}$ and $\theta^2/\sqrt{24}$. (b) $\kappa(\theta)\theta^2$ approaching $\sqrt{24} = 4.899$. Area-uniform sampling on a 60×96 grid.

Numerical Observation 28 (Constants and their uncertainty). On an area-uniform grid at $\theta = 2^\circ$, refining from 20×32 to 40×64 to 80×128 gives $s_6/\theta^2 = 0.2146, 0.2093, 0.2067$ with successive differences halving; Richardson extrapolation gives 0.2041 ± 0.0004 against the exact $1/\sqrt{24} = 0.204124$ of Theorem 10. The corresponding $\kappa\theta^2$ values are 4.656, 4.774, 4.834, extrapolating to 4.90 ± 0.01 against $\sqrt{24} = 4.8990$. The value s_2/θ is grid-insensitive at 0.70671 against $1/\sqrt{2} = 0.707107$. Under polar-uniform sampling the same quantities approach $\sqrt{2/45} = 0.21082$ and $1/\sqrt{3} = 0.57735$. Constants quoted from a single fixed grid without refinement are unreliable at the third digit; an earlier draft of this work reported 0.217 on that basis.



Observation 28: grid refinement of s_6/θ^2 at $\theta = 2^\circ$ against the exact value $1/\sqrt{24}$.



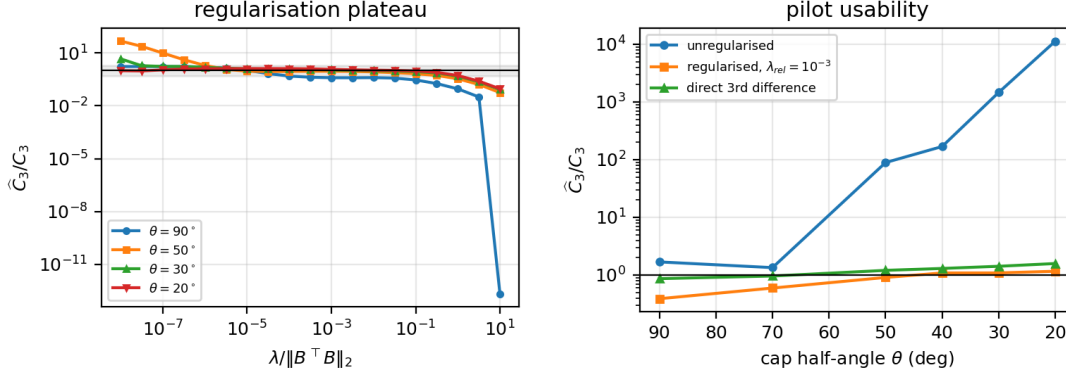
Left: Proposition 14 and Corollary 15. Noiseless odd-band energy is proportional to h for three test functions; with noise it reverts to the h^{-2} law of Proposition 19. Right: Proposition 17. Antipodally symmetric sets attain order 2; a hemisphere and a 45° cap do not, with 120 and 400 directions respectively.

Numerical Observation 29 (Parity scaling). Noiseless, with 120 Fibonacci directions, the ratio of odd-band energy to h is constant to three or four digits over a sixteenfold range of h : 0.893 for f_3 at every $h \in \{0.4, 0.2, 0.1, 0.05, 0.025\}$; 2.007 to 2.088 for f_4 ; 0.367 to 0.396 for f_1 . With $\sigma = 10^{-4}$ the measured exponent is -2.00 , matching Proposition 19.

Numerical Observation 30 (Attainable order). Measured convergence orders: full sphere 2.00; icosahedral set with antipodal partners 2.00; slab geometry with spans small relative to the slab thickness 2.00; hemisphere 1.03–1.09; cap of 45° 1.08–1.10; cap of 20° 1.11–1.16. The governing condition is antipodal *availability*, not proximity to a boundary as such.

Numerical Observation 31 (Rank deficiency of small Fibonacci sets). Design-matrix condition numbers at $m = 6$: icosahedral 1.581; the set $\{e_i\} \cup \{(e_i + e_j)/\sqrt{2}\}$ 3.226; Fibonacci 2.3×10^{15} . The Fibonacci spiral at $m = 6$ produces near-antipodal pairs, and since $u^T H u$ is even in u these are duplicate measurements. For $m = 12, 20, 50$ the Fibonacci condition numbers are 1.838, 1.637, 1.606.

Numerical Observation 32 (Parity separation on caps). The smallest principal angle between the restricted spans of even harmonics of degree ≤ 4 and odd harmonics of degree ≤ 3 decays far faster than the conditioning of Theorem 10: 3.8×10^{-3} at 90° , 3.6×10^{-4} at 70° , 1.9×10^{-5} at 50° , 2.9×10^{-6} at 40° , 2.4×10^{-7} at 30° , and below double precision at 20° — empirically of order θ^9 rather than θ^2 . Hessian conditioning and parity separation are therefore related degeneracy mechanisms with different asymptotic exponents, not a single law.



Left: Observation 33; $\hat{C}_3/C_3$ against the relative ridge parameter, shaded band a factor of two, plateau widening as θ decreases. Right: Observation 32 in operational form; the unregularised pilot diverges below roughly 70° while regularised parity inversion and a direct third difference remain within a small factor to 20° . $\sigma = 10^{-4}$ relative.

Numerical Observation 33 (Regularisation plateau). With λ measured relative to $\|B^\top B\|_2$, $\hat{C}_3(\lambda)$ remains within a factor of two of the true value over an interval spanning three to eight orders of magnitude, and the interval *widens* as θ decreases: measured widths 3×10^3 at 90° , 3×10^5 at 50° , 10^7 at 30° , 10^8 at 20° . Any $\lambda_{\text{rel}} \in [10^{-5}, 10^{-2}]$ performed acceptably at every aperture and on every test function considered. The estimate drifts low as λ grows, which is the benign direction under the cube-root span rule.

Remark 34 (What does not explain Observation 33). The natural spectral-filtering argument — that a large gap in the spectrum of $B^\top B$ renders λ irrelevant across the gap — does not account for the behaviour. Measured largest gap ratios are 1.2×10^{14} at 90° falling to 8×10^9 at 20° , predicting plateau widths some twenty orders of magnitude too large and the opposite trend in θ . The plateau concerns a single functional of the regularised solution rather than the solution itself. We have no proof.

Numerical Observation 35 (The regularised estimate tracks C_3). Scaling the cubic part of the test function by $A \in \{0.5, 1, 3, 10\}$ and regressing $\log \hat{C}_3$ on $\log A$ gives slopes 1.09, 1.12 and 1.06 at $\theta = 50^\circ, 30^\circ, 20^\circ$. A slope of one is what a working estimator must give; shrinkage alone would give zero.

Numerical Observation 36 (Pilot size). With a degree-3 basis (16 columns), pilots of $m_0 = 18, 24, 30, 40, 60$ directions give indistinguishable $\hat{C}_3$ at $\sigma = 10^{-4}$; at $\sigma = 10^{-2}$ no pilot size is stable, consistent with the h^{-3} variance of Corollary 22. A pilot with $m_0 = 24$ and a degree-4 basis (35 columns) is rank deficient and returned $\hat{C}_3$ inflated by a factor of about 800.

Numerical Observation 37 (Node placement). Error of the anisotropic-span directional estimator against R , for f_3 : at $\theta = 45^\circ$, $\sigma = 10^{-2}$, the values for $R = 0.1, 0.2, 0.3, 0.5, 0.7, 0.9$ are 6.17, 3.50, 2.76, **2.53**, 3.07, 6.73; at $\theta = 20^\circ$, $\sigma = 10^{-2}$, 27.3, 15.0, 11.3, **9.55**, 12.0, 30.0; at $\theta = 45^\circ$, $\sigma = 10^{-4}$ the minimum moves to $R \approx 0.3$. This matches Remark 20.

Numerical Observation 38 (*D*-optimal direction selection). Sequential *D*-optimal selection inside the feasible cone reduces error relative to a fixed set compressed into the cone: at $\theta = 30^\circ$, $m = 12$, $\sigma = 10^{-4}$, from 8.76×10^{-1} to 6.84×10^{-1} with condition number falling from 16.1 to 10.7; at $\theta = 10^\circ$, from 7.04 to 4.80 with condition number 165.6 to 124.4. At $m = 6$ the compressed fixed set is numerically singular.

Numerical Observation 39 (Identifiability, numerically). Over 200 random point configurations per n , the number of Hessian degrees of freedom left unresolved by n evaluations at points in general position is 4,3,2,1,0,0,0 for $n = 6, \dots, 12$, confirming Theorem 4.

PART: A controlled comparison of estimators

Protocol

The theory above bounds every method; it does not say which to use. This part reports a comparison under a fairness protocol and, equally importantly, the experimental artifacts that had to be removed before that comparison could be trusted. Nothing in Parts the corresponding section– the corresponding section depends on the outcome.

Fairness principles

1. **Matched budget, not matched stencil.** Methods are compared at equal numbers of function evaluations. A directional scheme with $m = 6$ costs 13 evaluations, so every baseline receives 13.
2. **Every method receives its own oracle-optimal radius.** For each cell each method's sampling radius or span is chosen from a common logarithmic grid of at least 15 points so as to minimise that method's own error. Tuning one method while leaving another at a default is the single most effective way to manufacture a false result.
3. **Baselines receive proper geometry management.** The interpolation baseline is built on a Δ -poised set constructed inside the feasible region by a pivotal algorithm [6], not on an arbitrary sample.
4. **Shared noise realisations.** Within a cell all methods see the same pseudo-random draws.
5. **Infeasibility is reported, not penalised.** A method whose stencil does not fit inside the feasible set is recorded as inapplicable rather than assigned a large error.
6. **A decision rule fixed in advance.** A directional scheme was to be judged advantageous only if it beat *every* feasible baseline by more than two Monte-Carlo standard errors in at least one cell at matched budget.

Methods compared

Quadratic regression on a Λ -poised set augmented to the budget; the Λ -poised interpolation model itself; moving least squares with Gaussian distance weights [13]; the 19-point central-difference stencil; classical one-sided second differences on the $\{e_i\} \cup \{(e_i + e_j)/\sqrt{2}\}$ directions; and three directional variants — isotropic span, anisotropic span with variance weighting, and anisotropic span on D -optimal directions.

Results

Mean $\|\hat{H} - H\|_F$ over 25 replicates for f_3 at a matched budget of 13 evaluations, each method at its own oracle-optimal radius. “—” marks an infeasible stencil. The remaining test functions behave the same way.

geometry	σ	quad. reg.	Λ -poised	MLS	central FD	one-sided FD	directional
free	10^{-4}	1.31e-1	1.49e-1	1.28e-1	5.95e-2	4.75e-1	3.35e-1
free	10^{-2}	7.71e-1	8.05e-1	8.03e-1	8.96e-1	2.57e0	1.82e0
cap 90°	10^{-4}	2.94e-1	3.06e-1	2.90e-1	—	—	4.25e-1
cap 45°	10^{-4}	3.62e-1	3.64e-1	3.66e-1	—	7.76e-1	5.22e-1
cap 20°	10^{-2}	2.92e0	3.19e0	3.45e0	—	3.03e1	1.18e1
slab	10^{-2}	1.44e0	1.41e0	1.49e0	2.16e1	2.70e0	2.35e1
wedge	10^{-2}	1.35e0	1.51e0	1.49e0	2.33e1	2.83e0	3.08e1

Redundant directions at matched budget, f_3 . Error falls monotonically with m for the directional scheme but the gap never closes; quadratic regression on the same points uses cross-direction information that per-ray differencing discards.

geometry	σ	m	evaluations	quad. regression	directional
cap 30°	10^{-4}	6	13	4.39e-1	1.12e0
cap 30°	10^{-4}	20	41	3.94e-1	6.85e-1
cap 30°	10^{-4}	50	101	3.46e-1	5.71e-1
cap 10°	10^{-2}	50	101	5.47e0	1.94e1
slab	10^{-4}	50	101	1.51e-1	1.93e-1

Findings

Central differences are the most accurate method wherever the symmetric stencil fits, and are inapplicable in every cone. Where they do not fit, quadratic regression, the Λ -poised interpolation model and moving least squares are indistinguishable from one another to within a few per cent, and all three match or outperform every directional scheme at matched budget. One-sided finite differences track the directional estimator closely, as Remark 2 predicts. Under a sensitivity study over eight random base points and eight random cone axes at $\theta = 35^\circ$, the quadratic family won in eight cases out of eight.

Two further findings are worth recording. First, central differences are not merely infeasible near a boundary but actively dangerous where they barely fit: in the slab and wedge geometries at $\sigma = 10^{-2}$ the feasibility constraint forces a span at which noise amplification dominates, and the error exceeds that of quadratic regression by more than an order of magnitude. Second, fixed antipodally symmetric direction sets collapse under feasibility constraints, as recorded in Section the corresponding section.

Interpretation

The pre-registered decision rule was not met in any cell, and we report that outcome as such. It is consistent with Corollary 23: the directional schemes tested differ from the baselines in how they handle mechanism (ii), truncation, but inherit the same limits from mechanisms (i) and (iii), and by Proposition 17 they cannot recover the order lost to one-sided access. It is also consistent with the asymptotic result of Chen and Wu [3], which shows local quadratic fitting does not degrade near a boundary.

Experimental pitfalls

Over the course of this study nine distinct experimental artifacts produced results that appeared publishable and were false. We group them by cause, since the pattern matters more than the incidents.

Parameter propagation.

A parameter that a method can silently override must be asserted, not assumed. In our code the node parameter R was not passed through to the probe, so a sweep over R measured only the span rule's dependence on it; separately, an estimator that computed its own span internally ignored the span supplied by the harness, so its error ratio against itself was identically 1.00 in every cell.

Geometry generation.

A map used to compress a direction set into a cone distorted the set whenever the aperture exceeded 90° , penalising one method in the unconstrained case; a cache keyed on cone aperture but not on cone axis returned directions for the wrong axis; and a fixed direction set silently lost half its members at a boundary, falling below the rank required by Proposition 5.

Benchmark fairness.

Baselines were initially sampled at the full trust radius while the method under test received tuned spans; the tuning grid for the baselines was coarser than the continuous optimisation available to the method under test; and a decision rule omitted one baseline that was in fact feasible and three orders of magnitude more accurate in precisely the cells where the method appeared to win.

Estimation and inference.

A pilot fit used 35 basis columns on 24 sample directions, so the minimum-norm solution absorbed noise into the band energies and returned a truncation constant inflated by a factor of about 800; a radius parameter was normalised differently in two runs, making the fitted constants incomparable; and a calibration constant fitted across apertures produced a power law $\alpha \propto \theta^{-2.49}$ with $R^2 = 0.983$ for one method and $\theta^{-2.58}$ with $R^2 = 0.979$ for another. Agreement between independent methods to within four per cent over a 229-fold range looked conclusive. Decomposing the ratio showed otherwise: the optimal radius varied by a factor of about two across the sweep while the estimated truncation constant varied by five orders of magnitude, so the apparent law described the degradation of the pilot estimator and nothing else.

Lesson.

The last of these is the most instructive. A clean power law, reproduced by two independent estimator families, is the kind of evidence that normally ends an investigation rather than starting one. The check that caught it was cheap: decompose the fitted quantity into its factors and ask which one is moving. We recommend this as routine practice whenever an empirical scaling law appears in a ratio of estimated quantities.

PART: Discussion

What the theorems do not say

- Theorem 10 bounds the conditioning of recovery from a restricted direction set. It does not assert that any particular estimator is preferable near a boundary, and it applies equally to methods that do not sample along rays, through the same restricted geometry.
- Theorem 4 concerns exact recovery without priors. It does not imply that interpolation-based or regression-based methods are inferior; under smoothness priors, and in the presence of noise, the comparison is governed by different quantities, and local quadratic fitting is known to converge asymptotically even on domains with boundary [3].
- Proposition 14 uses only the parity of homogeneous polynomials on the sphere and the standard harmonic decomposition. Neither is new; the contribution claimed is the use of the odd component as a Hessian-free measurement of the truncation term, and that claim is subject to the positioning questions of Section the corresponding section.
- Corollary 15 measures $\|D^3 f\|$, not the error of any particular estimate. Using it for span selection requires a further constant that we do not determine, and Corollary 22 shows the measurement is not cheap.

- Observation 33 is numerical. We give no proof, and as recorded in Remark 34 the natural spectral-filtering argument does not account for it.
- Part the corresponding section compares particular implementations under one protocol. A negative result there is evidence about those implementations at those budgets, not a theorem.

Open questions

1. A proof of Observation 33, or a counterexample delimiting it.
2. The exact asymptotic exponent of the parity separation angle on C_θ (Observation 32) and its dependence on the polynomial degrees involved; we expect this to follow from concentration theory [14, 11].
3. How to estimate C_3 reliably from a restricted cone. Regularised parity inversion is one answer; a principled choice of λ is not yet available, and the plateau of Observation 33 suggests the question may matter less than it appears.
4. Whether the optimal sampling radius admits a geometry-independent rule once C_3 is estimated correctly. The measurements are consistent with the hypothesis that the observed geometry dependence of the empirical constant reflects the failure of the unregularised pilot on caps rather than a property of the radius, but this is a correlation, not a demonstrated cause. The decisive test is to re-measure the optimal radius across apertures using the regularised pilot and ask whether the constant flattens.
5. Whether Theorem 10 has a coordinate-free formulation in terms of restricted Veronese embeddings, and whether such a result already exists.

Conclusions

Recovery of a Hessian from one-sided directional measurements is limited by three independent mechanisms: the geometry of the admissible directions, which controls conditioning and degrades as θ^{-2} ; truncation, which is of first order under genuinely one-sided access and cannot be improved by any amount of redundancy or reweighting; and noise, which enters as σh^{-2} and fixes the $\sigma^{1/3}$ scaling of the optimal span. Below ten evaluations in three dimensions the problem is not identifiable at all without priors. The measurement does carry one piece of structure not usually exploited — its odd spherical-harmonic content measures the truncation term without knowledge of the Hessian — though extracting it on a restricted cap requires regularisation, and the pilot costs more evaluations than the estimate it calibrates.

Under a fairness protocol in which every method receives its own oracle-optimal sampling radius, the directional schemes we tested did not outperform quadratic regression on the same feasible points at matched budget in any configuration examined. We report that

outcome, and the nine experimental artifacts that had to be removed before we could trust it, in the belief that both are more useful to subsequent work than a favourable comparison would have been.

Reproducibility

Monte-Carlo quantities use fixed seeds derived deterministically from the cell index, so that every method within a cell sees the same noise realisation. Grid-refinement studies double both grid coordinates and report only digits stable under refinement, with Richardson extrapolation where the sequence is monotone. Scripts: `kr3d_killtest.py` (geometries, estimators, evaluator, analytic Hessians and their verification), `kr3d_harmonics.py` (parity projection and $\hat{C}_3$), `capproof.py` (singular values and parity separation on caps), `kr3d_full_study.py` (method inventory, R sweep, sensitivity), `kr3d_niche_probe.py` (identifiability check, eigenpair estimation, D -optimal directions), `make_figures.py` (all figures). Seeds and grid parameters are literals in the scripts; no figure depends on an unrecorded choice.

Declarations

Funding. None.

Conflicts of interest. The author declares no competing interests.

Data availability. All data are generated by the scripts listed above.

Ethics and consent to participate. Not applicable.

Author contributions. The sole author is responsible for all aspects of the work.

The divided-difference identity

With nodes $0, Rh, h$ relative to x , the second divided difference is

$$f[x, x + Rhu, x + hu] = \frac{f_0}{(0 - Rh)(0 - h)} + \frac{f_1}{(Rh)(Rh - h)} + \frac{f_2}{h(h - Rh)},$$

where $f_0 = f(x)$, $f_1 = f(x + Rhu)$, $f_2 = f(x + hu)$. Multiplying by $R(1 - R)h^2$ gives $(1 - R)f_0 - f_1 + Rf_2$, which is the bracket of (1) with $K = 1 - R$. Hence (1) equals $2f[x, x + Rhu, x + hu]$, and since $f[x_0, x_1, x_2] = 1/2 f''(\xi)$ for some ξ in the convex hull, the probe is a second-derivative estimate along u using three points strictly inside $[x, x + hu]$.

Coefficients of the probe expansion

For completeness, the bracket in (1) applied to $t \mapsto t^k$ along u contributes $h^k(R - R^k)/k!$, so after division by $R(1 - R)h^2/2$ the coefficient of $D^k f(u)$ is

$$\frac{2}{k!} h^{k-2} \frac{R - R^k}{R(1 - R)} = \frac{2}{k!} h^{k-2} \sum_{j=0}^{k-2} R^j,$$

giving 1 , $(1 + R)h/3$, $(1 + R + R^2)h^2/12$ and $(1 + R + R^2 + R^3)h^3/60$ for $k = 2, 3, 4, 5$. The $k = 5$ term is odd and is the first neglected contribution to the odd projection, which is why Corollary 15 is accurate to $O(h^2)$ after division by h .

References

1. M. Alipoor and I. Y. H. Gustafsson, *K-optimal gradient encoding scheme for fourth-order tensor-based diffusion profile imaging*, BioMed Research International, 2015.
2. A. S. Berahas, L. Cao, K. Choromanski and K. Scheinberg, *A theoretical and empirical comparison of gradient approximations in derivative-free optimization*, Foundations of Computational Mathematics **22** (2022), 507–560.
3. C.-W. Chen and H.-T. Wu, *Convergence of Hessian estimator from random samples on a manifold with boundary*, arXiv:2303.12547.
4. A. R. Conn, K. Scheinberg and L. N. Vicente, *Geometry of interpolation sets in derivative free optimization*, Mathematical Programming **111** (2008), 141–172.
5. A. R. Conn, K. Scheinberg and L. N. Vicente, *Geometry of sample sets in derivative-free optimization: polynomial regression and underdetermined interpolation*, IMA Journal of Numerical Analysis **28** (2008), 721–748.
6. A. R. Conn, K. Scheinberg and L. N. Vicente, *Introduction to Derivative-Free Optimization*, MPS–SIAM Series on Optimization, SIAM, Philadelphia, 2009.
7. A. R. Curtis and J. K. Reid, *The choice of step lengths when using differences to approximate Jacobian matrices*, Journal of the Institute of Mathematics and its Applications **13** (1974), 121–126.
8. J. Dumontet and J. Vignes, *Détermination du pas optimal dans le calcul des dérivées sur ordinateur*, RAIRO Analyse Numérique **11** (1977), 13–25.
9. K. M. Hasan, D. L. Parker and A. L. Alexander, *Comparison of gradient encoding schemes for diffusion-tensor MRI*, Journal of Magnetic Resonance Imaging **13** (2001), 769–780.
10. D. K. Jones, M. A. Horsfield and A. Simmons, *Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging*, Magnetic Resonance in Medicine **42** (1999), 515–525.
11. H. J. Landau and H. Widom, *Eigenvalue distribution of time and frequency limiting*, Journal of Mathematical Analysis and Applications **77** (1980), 469–481.

12. J. Larson, M. Menickelly and S. M. Wild, *Derivative-free optimization methods*, Acta Numerica **28** (2019), 287–404.
13. D. Levin, *The approximation power of moving least-squares*, Mathematics of Computation **67** (1998), 1517–1531.
14. F. J. Simons, F. A. Dahlen and M. A. Wiecezorek, *Spatiospectral concentration on a sphere*, SIAM Review **48** (2006), 504–536.
15. S. Skare, M. Hedehus, M. E. Moseley and T.-Q. Li, *Condition number as a measure of noise performance of diffusion tensor data acquisition schemes with MRI*, Journal of Magnetic Resonance **147** (2000), 340–352.
16. R. S. Stepleman and N. D. Winarsky, *Adaptive numerical differentiation*, Mathematics of Computation **33** (1979), 1257–1264.