Sharp minimax rates for inverse diffusion metric inference

Scott Schubr
Luminary Oddities

November 23, 2026

Abstract

We determine when a fixed diagonal entry of the inverse diffusion metric has uniform root-sample risk. The observations have an unknown strictly positive density in a bounded Hölder class on a compact manifold, including with reflecting boundary. For every fixed smooth nonconstant score, the sharp threshold is $s = d/4$, including equality. Quadratic and cubic corrections attain the minimax rate at every smoothness. The quadratic correction estimates the Hessian with a finite-rank kernel that retains both mixed frequency blocks. Above the threshold the estimator is asymptotically linear with the canonical influence function. At equality, no estimator admits uniform centered Gaussian calibration with bounded scale, even near the uniform density. The lower bounds use local oscillations that annihilate the score and the first variation exactly and produce a positive inverse-energy increment.

1 Model and results

A diagonal entry of the inverse diffusion metric is an elliptic energy depending on the unknown sampling density. At the critical smoothness $s = d/4$, this energy can be estimated with root-sample accuracy, while uniform centered Gaussian calibration with bounded scale is impossible. Above the threshold an efficient estimator exists. The same separation holds on smooth domains with reflecting boundary.

The difficulty beyond standard quadratic-functional estimation is the density-dependent inverse operator. The upper bound uses second- and third-order corrections that preserve mixed frequency contributions. The lower bound cancels the first derivative exactly and controls the remaining positive energy increment.

Give $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ unit volume and fix a smooth real nonconstant score u. Let $\mathcal{F}_s$ be the set of densities f with $m \le f \le M$ and $\|f\|_{C^s} \le B$, where $0 < m < 1 < M$ and $B > \|1\|_{C^s}$. We use the classical Hölder norm at nonintegers and the supremum norm of derivatives through order s at integers. The observations are n independent draws with density f. Write $P_f v = \int f v$ and $u_f = u - P_f u$. Define

$$A_f = -\operatorname{div}(f\nabla), \qquad b(f) = f u_f, \qquad F(f) = \langle b(f), A_f^{-1} b(f)\rangle.$$

The inverse is variational on the fixed space of Lebesgue-mean-zero H^1 functions. If $\phi_f = A_f^{-1} b(f)$ and $\phi_f^\circ = \phi_f - P_f\phi_f$, then $F(f) = P_f|\nabla\phi_f|^2$. This is G^0_{uu} in Schubr [3, (2.1)]. That paper gives computational inference for a shrinking-bandwidth empirical heat-kernel Sinkhorn estimator at a fixed smooth positive density. Here we estimate the same population target over a bounded finite-smoothness class.

Theorem 1. *Define the minimax root mean squared error by*

$$\mathcal{R}_n(s) = \inf_{\widehat{F}} \sup_{f \in \mathcal{F}_s} \left(\mathbb{E}_f |\widehat{F} - F(f)|^2\right)^{1/2}.$$

For $s \geq d/4$, $\mathcal{R}_n(s) \asymp n^{-1/2}$. For $0 < s < d/4$,

$$\mathcal{R}_n(s) \asymp n^{-4s/(4s+d)}.$$

For $s > d/4$ there is an estimator which, uniformly over the class, has the expansion

$$\widehat{F} - F(f) = \frac{1}{n} \sum_{i=1}^{n} \psi_f(X_i) + o_{\mathbb{P}_f}(n^{-1/2}), \qquad \psi_f = 2u_f \phi_f^\circ - |\nabla \phi_f|^2 - F(f).$$

At each interior density, ψ_f is the canonical gradient in the density model. Its variance is positive, and the asymptotic variance can be estimated consistently.

Theorem 2. *Set $s = d/4$. There is no sequence of estimators and positive deterministic scales $\sigma_n(f)$ bounded above on $\mathcal{F}_s$ for which the distributions of $\sqrt{n}(\widehat{F} - F(f))/\sigma_n(f)$ converge uniformly over $\mathcal{F}_s$ to the standard Gaussian in Kolmogorov distance. This is already impossible on every fixed L^∞ neighborhood of the uniform density inside $\mathcal{F}_s$. It remains impossible for positive random studentizers with a uniform deterministic upper bound.*

The exponent $d/4$ is classical for quadratic density functionals. Quadratic influence corrections and the bias and variance balance of finite-rank kernels are developed in [2]. The target here contains a density-dependent inverse elliptic operator. We identify an operator on L^2 representing its quadratic variation, and construct oscillations whose inverse-energy increment has both a fixed sign and the required size for every nonconstant score. Our statements concern statistical risk for the population functional.

2 Variational expansion

We use two elliptic bounds, uniform for $q \in \mathcal{F}_s$:

$$\|\phi_q\|_{W^{1,\infty}} \leq C, \qquad \|\nabla A_q^{-1} \operatorname{div} V\|_3 \leq C\|V\|_3. \tag{1}$$

For the first, choose $0 < \alpha < \min(s, 1)$, write $b(q)$ as the divergence of a uniformly C^α vector field by the periodic Poisson equation, and apply the interior divergence-form Schauder estimate in periodic charts. For the second, the $W^{1,3}$ interior estimate for uniformly elliptic coefficients with a common modulus of continuity applies on a finite cover. For clarity, the local estimate bounds the L^3 gradient on a smaller ball by the L^3 data and the L^3 solution on a larger ball [1, §3]. Summing the estimates and using $\|v\|_3 \leq \epsilon\|\nabla v\|_3 + C_\epsilon\|v\|_2$ on the torus gives the claimed bound by absorption and the energy estimate. Approximation by smooth coefficients and data justifies the argument for weak solutions. The same bound with $\|g\|_3$ on the right holds for mean-zero scalar data g, since $\|\nabla \Delta^{-1} g\|_3 \leq C\|g\|_3$.

For pilot $p \in \mathcal{F}_s$ put $a = f - p$, $c_a = \int au$, and set

$$b_1 = au_p - pc_a, \qquad b_2 = -ac_a, \qquad r_1 = b_1 + \operatorname{div}(a\nabla\phi_p), \qquad r_2 = b_2.$$

The two r_i have mean zero for mean-zero a. Since $b(f) = A_f\phi_p + r_1 + r_2$, expansion of the energy gives the exact identity

$$F(f) - F(p) = \int a\psi_p + 2\langle b_2, \phi_p\rangle + \langle r_1 + r_2, A_f^{-1}(r_1 + r_2)\rangle. \tag{2}$$

In fact the linear term before centering is $2\langle b_1, \phi_p\rangle - \int a|\nabla\phi_p|^2$. The equality $P_p\psi_p = 0$ follows from $P_p(u_p\phi_p^\circ) = F(p)$.

Put $Q_p(a) = 2\langle b_2, \phi_p\rangle + \langle r_1, A_p^{-1}r_1\rangle$. The energy estimate and (1) give $\|r_1\|_{H^{-1}} \le C\|a\|_2$ and $\|r_2\|_{H^{-1}} \le C\|a\|_2^2$. The mixed terms with r_2 and its square in (2) are bounded by $C\|a\|_3^3$, using the fixed bound on $\|a\|_\infty$. The remaining difference is

$$\langle r_1, (A_f^{-1} - A_p^{-1})r_1\rangle = -\int a\,\nabla A_f^{-1}r_1 \cdot \nabla A_p^{-1}r_1.$$

Both gradients have L^3 norm at most $C\|a\|_3$ by (1). Consequently

$$\left|F(f) - F(p) - \int a\psi_p - Q_p(a)\right| \le C\|a\|_3^3. \tag{3}$$

The quadratic form Q_p is bounded by $C\|a\|_2^2$. Polarization and the Riesz representation theorem give a self-adjoint operator H_p on mean-zero L^2, with uniformly bounded operator norm, such that $Q_p(a) = \langle a, H_p a\rangle$. Extend it by zero on constants.

We also have the stability estimate

$$\|\psi_f - \psi_p\|_2 \le C\|a\|_2. \tag{4}$$

Indeed the difference of the variational equations has data $r_1 + r_2$ and its solution has H^1 norm at most $C\|a\|_2$. Centering changes the constant by at most $C\|a\|_2$. Use the gradient bound in (1) to take the difference of the squares, and (2) to control $F(f) - F(p)$.

3 Quadratic correction and upper bounds

Take a pilot using an independent subsample of size proportional to n. A periodic kernel density estimator $\bar f$ with bandwidth $h_0 \asymp n^{-1/(2s+d)}$ and a bounded compactly supported kernel with sufficiently many vanishing moments has supremum bias $O(h_0^s)$. The pointwise sixth central moment is bounded by

$$C\{(nh_0^d)^{-3} + (nh_0^d)^{-5}\}.$$

This follows from the sixth moment inequality for an independent centered sum, the per-observation second moment $O(h_0^{-d})$ and sixth moment $O(h_0^{-5d})$. Since the torus has volume one, $\mathbb{E}\|\bar f - f\|_3^6 \le \int \mathbb{E}|\bar f - f|^6$. Choose the first element p in a fixed countable L^3-dense subset of $\mathcal{F}_s$ whose distance to $\bar f$ is within n^{-1} of the infimum. The selection is measurable and

$$\sup_{f\in\mathcal{F}_s} \mathbb{E}_f\|p - f\|_3^6 \le Cn^{-6s/(2s+d)}. \tag{5}$$

Let P_J be the orthogonal Fourier projection at maximal coordinate frequency J, with rank $K = (2J+1)^d$. The projection error on $\mathcal{F}_s$ is at most $CK^{-s/d}$ in L^2. For example, convolve the periodic extension with the dilate at scale J^{-1} of a Schwartz function whose Fourier transform has compact support and is identically one near zero. Its higher moments vanish, Taylor's formula bounds the supremum error by CJ^{-s}, and orthogonal projection can only decrease the L^2 error. Define a self-adjoint correction operator by

$$B_{p,J} = H_pP_J + P_JH_p - P_JH_pP_J.$$

It has rank at most $2K$, and operator norm at most $3\|H_p\|$. The omitted operator is $(I-P_J)H_p(I-P_J)$, so

$$|Q_p(a) - \langle a, B_{p,J}a\rangle| \le CK^{-2s/d}. \tag{6}$$

In particular the two mixed frequency blocks are estimated as well as the low-frequency block.

Let $B(x, y)$ be an L^2 kernel of $B_{p,J}$ and $B^{[p]}$ its centering in each variable against p. The operator is finite rank, so a kernel exists and has squared L^2 norm $O(K)$. The pilot takes countably many values, which allows one fixed measurable representative for each kernel used. Given the pilot, on N new observations use

$$\widehat{F}_p = F(p) + \frac{1}{N}\sum_i \psi_p(X_i) + \frac{1}{N(N-1)}\sum_{i\neq j} B^{[p]}(X_i, X_j).$$

The conditional mean of the quadratic term is $\langle a, B_{p,J}a\rangle$. Its first Hoeffding projection is the f-centered function $B_{p,J}a$, whose $L^2(f)$ norm is at most $C\|a\|_2$. Its second projection is the double centering of B under f and has squared $L^2(f\otimes f)$ norm at most CK. Therefore the conditional variance of the quadratic term is bounded by $C(\|a\|_2^2/N + K/N^2)$. With $N \asymp n$, (3)–(6) and bounded ψ_p give

$$\sup_f \left(\mathbb{E}_f|\widehat{F}_p - F(f)|^2\right)^{1/2} \le C\{n^{-1/2} + \sqrt{K}/n + K^{-2s/d} + n^{-3s/(2s+d)}\}. \tag{7}$$

Choose $K \asymp n$ at and above equality. The next section supplies the cubic correction below equality.

Above equality, use $n^{d/(4s)} \ll K \ll n$. The bias and the centered quadratic term are then $o(n^{-1/2})$ in root mean square. By (4)–(5) the centered first empirical term can use ψ_f in place of ψ_p at the same error scale. Split into two nearly equal folds and use each for validation with the other for the pilot. Weighting the two estimates by validation sample size gives the full-sample expansion. Influence variance estimation follows by averaging squared pilot influences in the validation folds, using boundedness and (4).

For every f the influence variance is positive. The continuous mean-zero ϕ_f° has a zero, where $\psi_f \le -F(f) < 0$. Since $P_f\psi_f = 0$ and f is bounded below, the variance is positive. At an interior density, the scores δ/f for smooth mean-zero additive directions δ span a dense subspace of $L_0^2(f)$. Equation (3) gives the derivative $\int \psi_f\,\delta$. Thus ψ_f is the canonical gradient. The uniform remainder holds along local paths through an interior point, and (4) and the triangular-array central limit theorem give regularity and the canonical variance along these paths.

4 Cubic correction below the threshold

Use the $W^{1,4}$ version of (1), obtained by the same argument, and let $A_a = -\operatorname{div}(a\nabla)$. The exact resolvent identity is

$$A_f^{-1} = A_p^{-1} - A_p^{-1}A_aA_p^{-1} + A_p^{-1}A_aA_f^{-1}A_aA_p^{-1}.$$

For $w_a = A_p^{-1}r_1$, the cubic term of (2) is

$$C_p(a,a,a) = 2\langle r_2, w_a\rangle - \int a|\nabla w_a|^2.$$

Here C_p denotes its symmetric trilinear polarization, extended by first removing the spatial mean of each input. The exact residual after including this term is bounded by $C\|a\|_4^4$. Indeed, $\|\nabla A_p^{-1}r_1\|_4 \le C\|a\|_4$ and $\|\nabla A_p^{-1}r_2\|_4 \le C\|a\|_4^2$. The final resolvent term is bounded by $C\|a\nabla A_p^{-1}(r_1+r_2)\|_2^2 \le C\|a\|_4^4$. In the other terms of degree at least four, apply Hölder's inequality with these gradient bounds and $\|a\|_4 \le C$.

Each summand of $C_p(a_1,a_2,a_3)$ has the form

$$\int (L_1a_1)(L_2a_2)(L_3a_3),$$

where all three L_i are uniformly bounded on L^2 and L^4. The possible maps are mean removal, a scalar score moment as a constant function, the mean-zero solution map $a \mapsto w_a$, and its gradient components. Scalar sources are handled as in (1). In particular we may bound any two inputs in L^4 and the remaining one in L^2.

Use a self-adjoint finite-rank smoothing S of resolution J that has a representation $Sg = \sum_{\ell=1}^{K} e_\ell \langle e_\ell, g\rangle$ with $K \asymp J^d$ and satisfies

$$\sum_\ell |\langle e_\ell, g\rangle|^2 \le \|g\|_2^2, \quad \sum_\ell |e_\ell(x)|^2 \le CK, \quad \|S\|_{L^q \to L^q} \le C, \quad \|(I-S)a\|_q \le CK^{-s/d}, \qquad q = 2, 4. \tag{8}$$

On the torus take a real Fourier multiplier with a smooth cutoff in $[0,1]$ equal to one near zero. The e_ℓ are the Fourier basis elements multiplied by square roots of the multiplier. The periodic convolution bounds follow from its uniformly bounded kernel L^1 norm and the same moment argument used for P_J. For the cubic correction retain the all-S term and each term with one $I - S$ argument. Equivalently, sum the three terms with two S arguments and one identity argument and subtract twice the all-S term. The omitted terms have two high-frequency inputs in L^4 and the third in L^2. Thus their bias is bounded by $CK^{-2s/d}\|a\|_2$. This uses the uniformly bounded L^2 norms of S and $I - S$.

We give the variance bound for the cubic correction. In each summand write the represented S arguments as functions of their coefficients $\langle e_\ell, g\rangle$. If j arguments are left random and the others are evaluated at a, the resulting kernel has squared L^2 norm $O(K^{j-1})$. To see this, use the L^2 adjoint of L_i in one live slot, choosing the identity slot if it is live, and replace its coefficient sum by Bessel's inequality if it carries S. Bound the remaining coefficient sums using

$$\left\| \left(\sum_\ell |Le_\ell|^2 \right)^{1/2} \right\|_4 \le C \left\| \left(\sum_\ell |e_\ell|^2 \right)^{1/2} \right\|_4 \le C\sqrt{K}.$$

The first inequality follows by taking fourth moments of independent Rademacher sums before and after applying the uniformly L^4-bounded L. For example, with three live slots the sum is bounded by $C \int (\sum_\ell |L_1 e_\ell|^2)(\sum_k |L_2 e_k|^2) \le CK^2$. With two live slots replace one square function by $|L_i a|$ or $|L_i S a|$, bounded in L^4. With one live slot use the L^4 bounds on both fixed factors. This argument covers every summand and also constructs an L^2 kernel when one slot uses the identity.

For an explicit estimator use three mutually independent validation groups, each of size of order n and independent of the pilot. Evaluate the cubic kernel on the product of the three empirical signed measures centered at $p\,dx$. Its mean is the truncated cubic term. Expand each measure as $a\,dx$ plus the corresponding empirical error around $f\,dx$. Terms with distinct nonempty sets of random slots are conditionally orthogonal. For a set of size j, the kernel variance is bounded by $Cn^{-j}K^{j-1}$ by the preceding bound, since centering is an $L^2(f^{\otimes j})$ contraction and f is bounded. Thus the cubic standard deviation is bounded by $C(n^{-1/2} + \sqrt{K}/n + K/n^{3/2})$. The quadratic and linear terms can use the same validation data, by the triangle inequality for root mean square error.

Below equality select the pilot in L^4 instead of L^3. The same kernel proof with eighth moments gives $\mathbb{E}\|a\|_4^8 \le Cn^{-8s/(2s+d)}$ (use the eighth moment bound $C[(nh_0^d)^{-4} + (nh_0^d)^{-7}]$ in place of the sixth). Put $x = s/d$, $v = x/(1+2x)$ and $\gamma = 4x/(1+4x)$. Take quadratic rank of order $n^{2/(1+4x)}$ and cubic rank of order n^κ, where

$$\kappa = \frac{2}{1+4x} - \frac{1}{2(1+2x)}.$$

The quadratic bias and standard deviation at this rank are both $O(n^{-\gamma})$. The cubic bias has root mean square $O(n^{-2x\kappa - v}) = O(n^{-\gamma})$, and the fourth-order pilot remainder is $O(n^{-4v}) = o(n^{-\gamma})$.

If $\kappa \le 1$ the cubic variance is $O(n^{-1})$. If $\kappa > 1$ its dominant standard deviation is $O(n^{\kappa-3/2}) = O(n^{-\gamma})$, because

$$\frac{3}{2} - \kappa - \gamma = \frac{1}{2} - \frac{1}{1+4x} + \frac{1}{2(1+2x)} > 0 \qquad (x > 0).$$

This proves the claimed subcritical upper bound. For comparison, Tchetgen et al. [4] use distinct projection levels for estimating a power of a density. Here keeping one slot unprojected controls the two-tail bias for the nonlocal trilinear term and requires only two operator resolutions.

5 Lower bounds and failure of critical Gaussian calibration

At $f = 1$, write ϕ_0 for the solution, ψ_0 for its influence, and $u_0 = u - \int u$. The smooth nonconstant potential has nonzero gradient at an interior point. Choose a small Euclidean cube, rotating its axes so that $\partial_1 \phi_0 \ge c > 0$ and the transverse gradient is as small as needed. Inside it place $m_h \asymp h^{-d}$ cubes of side h separated by distances of order h. In the interior of the rescaled unit cube fix three smooth translated bumps $\zeta_1, \zeta_2, \zeta_3$ of disjoint support and with nonzero first-coordinate derivative. In cell j choose a Euclidean unit coefficient vector in the kernel of the two linear constraints

$$\int u\, \partial_1 v_j = 0, \qquad \int \psi_0\, \partial_1 v_j = 0, \qquad v_j(x) = \rho h^{s+1} \sum_{\ell=1}^{3} c_{j\ell} \zeta_\ell((x - x_j)/h).$$

The linear constraints always have a nontrivial kernel. Put $a_j = \partial_1 v_j$; then $\int a_j = 0$. Disjoint bump supports give, uniformly in the unit coefficient vector,

$$\|a_j\|_2^2 \asymp \rho^2 h^{2s+d}, \qquad \|v_j\|_2 \le Ch\|a_j\|_2, \qquad \|\nabla v_j\|_2 \le C\|a_j\|_2.$$

All $f_\epsilon = 1 + \sum_j \epsilon_j a_j$ for signs $\epsilon_j \in \{-1, 1\}$ belong to $\mathcal{F}_s$ for a sufficiently small fixed $\rho > 0$. This follows by scaling the derivatives in each cell and, at noninteger s, from the separation of cells when the Hölder difference uses different supports.

Write $a = \sum_j \epsilon_j a_j$ and $V = \sum_j \epsilon_j v_j$. The score moment and first variation vanish exactly. In (2), $r_2 = 0$ and $r = r_1 = au_0 + \operatorname{div}(a\nabla\phi_0)$. Consequently

$$F(f_\epsilon) - F(1) = \langle r, A_{f_\epsilon}^{-1} r \rangle.$$

Testing r against $-V$ after subtracting its mean gives $\langle r, -V \rangle \ge c\|a\|_2^2$. In detail, the longitudinal term is $\int a^2 \partial_1 \phi_0$, the transverse term is bounded by $C\|\nabla_\perp \phi_0\|_\infty \|a\|_2^2$, and the source term is bounded by $Ch\|a\|_2^2$. The test's energy is at most $C\|a\|_2^2$. By variational duality,

$$c\rho^2 h^{2s} \le F(f_\epsilon) - F(1) \le C\rho^2 h^{2s}. \tag{9}$$

The upper bound follows directly from $\|r\|_{H^{-1}} \le C\|a\|_2$.

Let Q be the uniform sign mixture of the product laws of the alternatives and P_0 the product uniform law. Disjoint supports and unit masses give

$$\begin{aligned} 1 + \chi^2(\mathsf{Q}, \mathsf{P}_0) &= \mathbb{E}_{\epsilon,\epsilon'} \left(1 + \sum_j \epsilon_j \epsilon_j' \|a_j\|_2^2 \right)^n \\ &\le \prod_j \cosh(n\|a_j\|_2^2) \le \exp(Cn^2\rho^4 h^{4s+d}). \end{aligned} \tag{10}$$

For $h = tn^{-2/(4s+d)}$ and fixed small $t > 0$ the distance in total variation is bounded away from one. Thresholding an estimate halfway through the smallest gap gives the two-experiment testing bound and yields minimax risk at least $cn^{-4s/(4s+d)}$. For the parametric lower bound, ψ_0 is nonzero and smooth. Choose a smooth additive mean-zero direction with nonzero first variation at 1, and compare it at sizes 0 and $\eta/\sqrt{n}$ for fixed small η. The parameter gap is of order $n^{-1/2}$ by (3), while the product chi-square distance is bounded and can be made arbitrarily small by η. This yields the remaining lower bound.

At $s = d/4$ use $h = tn^{-1/d}$. Every alternative shifts the root-n parameter to the right by at least $\delta_t = c\rho^2 t^{d/2}$, whereas (10) bounds total variation by $C\rho^2 t^d$. If the asserted uniform centered normal approximation held with scales bounded by M_0, the null would give probability $1/2 + o(1)$ to $\{\widehat{F} \le F(1)\}$. Each alternative would give at most $\Phi(-\delta_t/M_0) + o(1)$, uniformly in signs. This also holds for random scales bounded by M_0, since the event implies the studentized variable is at most $-\delta_t/M_0$. Thus the probability gap would be at least $c'\rho^2 t^{d/2} + o(1)$ for fixed small t, exceeding the total variation bound for a sufficiently small choice of t. The alternatives converge uniformly to 1, proving the local version of Theorem 2. The same comparison of means, by Cauchy–Schwarz against the mixture likelihood ratio, rules out uniform $o(n^{-1/2})$ bias coupled with bounded variance of $\sqrt{n}\,\widehat{F}$ at the uniform density.

6 Smooth compact manifolds with reflecting boundary

Theorem 3. *Theorems 1 and 2 hold on any fixed smooth compact connected d-dimensional Riemannian manifold with smooth boundary, including an empty boundary. Use normalized volume, intrinsic gradient and divergence, and the variational Neumann inverse in the definition of F. The Hölder norm is taken in a fixed smooth atlas up to the boundary, with the uniform density in the interior of the class.*

All variational identities and moment calculations carry over using normalized volume. We verify the required analytic inputs.

For the elliptic bounds choose interior charts and boundary-normal collars. In a boundary chart the operator has coefficient matrix $fJ_g(g^{ij})$ relative to coordinate volume, where J_g is the volume Jacobian. The normal–tangential entries of the metric vanish in the collar. Reflect a weak Neumann solution evenly, reflect the coefficient matrix by conjugation by the coordinate reflection, and reflect vector data by the same coordinate reflection. Scalar data are reflected evenly. Folding a test function into the half-chart gives the reflected weak equation. The coefficient extends with a uniform C^α bound for $0 < \alpha < \min(s, 1)$. Thus the interior $W^{1,q}$ estimates for $q = 3, 4$ and the energy absorption argument from (1) apply on finitely many patches. For a mean-zero scalar source, use the Neumann Poisson solution of the smooth background Laplacian to write the source as divergence data with normal component zero and L^q norm bounded by the L^q norm of the source. This uses the classical $W^{2,q}$ estimate for that smooth Neumann problem. For the gradient supremum of ϕ_f, apply interior divergence-form Schauder in the reflected patches. The scalar datum in each patch is uniformly C^α and can be written locally as divergence of a C^α field by a Euclidean Poisson solve after a cutoff in the extended chart. The energy bound supplies the low norm in the local Schauder estimate. This gives the same uniform bounds and continuity of gradients up to the physical boundary. No condition is imposed on the normal derivative of f.

For the pilot and the correcting kernels, use chartwise piecewise polynomials. Take a finite smooth square partition of unity $\sum_i \theta_i^2 = 1$ subordinate to rectangular interior or half-collar charts. Such a partition is obtained by dividing a subordinate partition by the square root of the sum of its squares. Rescale chart rectangles and subdivide into cubes of side comparable to J^{-1}, with the

boundary face on the grid. Let P_i be the orthogonal projection on coordinate polynomials of fixed degree at least $\lceil s \rceil$ in each cell, with the volume Jacobian included. Extend it by zero outside the chart and set

$$S_J g = \sum_i \theta_i P_i(\theta_i g).$$

In each chart cell the Gram matrix of fixed rescaled monomials, after division by cell volume, is uniformly positive and bounded because the smooth Jacobian is bounded above and below. The local projection has kernel bounded by CJ^d in its cell. It is bounded in all fixed L^q norms and in supremum norm. Taylor polynomials and local reproduction therefore give $\|(I - S_J)g\|_\infty \leq CJ^{-s}\|g\|_{C^s}$, by applying this to each $\theta_i g$ and using the square partition.

The functions e_ℓ of (8) may now be taken as the chartwise orthonormal polynomial basis on each cell multiplied by θ_i, with ℓ ranging over all charts and cells. The contraction bound follows from the square partition and Bessel's inequality in each chart. The pointwise sum of squares is $O(J^d)$ by the rescaled Gram bound and finite overlap. Thus (8) holds with $K \asymp J^d$. Use S_J in place of P_J for the quadratic correction as well. The rank and missing-block identities require only self-adjointness, bounded norm, and the same L^2 approximation, so the proof continues to apply. For the pilot use the kernel of S_J with $J \asymp n^{1/(2s+d)}$ to smooth the empirical measure, then use the same countable class selection. Its kernel has supremum norm $O(J^d)$ by the bound on the pointwise sum of squares. Its squared L^2 norm in the sampled argument, uniformly in the evaluation argument, is $O(J^d)$ by Bessel's inequality. These give the same sixth and eighth moment bounds as those used for the flat pilot. The bias follows from the supremum approximation just proved.

The lower bounds use a small interior coordinate cube. Choose the first coordinate along the gradient of ϕ_0 at its center. With J_g the Jacobian for normalized volume, take $a_j = J_g^{-1}\partial_1 v_j$ and impose the same two moment constraints, now with normalized volume. The mass constraint is automatic by cancellation of J_g. The estimates for the scaled templates and their norms are unchanged up to constants. The main term is the longitudinal component of $\nabla\phi_0 \cdot \nabla V$ in coordinates. Its coefficient is bounded below and its transverse components can be made arbitrarily small by shrinking the cube. Thus the same signed test gives (9). Disjoint supports give exactly the same product chi-square calculation. The local two-point parametric lower bound and influence nondegeneracy use connectedness as before. This proves Theorem 3.

References

[1] P. Auscher and M. Qafsaoui, *Observations on $W^{1,p}$ estimates for divergence elliptic equations with VMO coefficients*, Boll. Unione Mat. Ital. B (8) 5 (2002), 487–509. Article.

[2] J. Robins, L. Li, E. Tchetgen and A. W. van der Vaart, *Quadratic semiparametric von Mises calculus*, Metrika 69 (2009), 227–247. DOI.

[3] S. Schubr, *Fourier computation of empirical diffusion metrics*, Luminary Research Archive, edition 3 (2026). Edition 3.

[4] E. Tchetgen, L. Li, J. Robins and A. W. van der Vaart, *Minimax estimation of the integral of a power of a density*, Statist. Probab. Lett. 78 (2008), 3307–3311. DOI.