Estimating density-dependent elliptic energies and testing effective conductivity from spatial samples

Scott Schubr
Luminary Oddities

November 25, 2026 (edition 2)

Abstract

Effective conductivity and the inverse diffusion metric are elliptic energies in which both the operator $-\operatorname{div}(f\nabla)$ and the source depend on the sampling density f. We estimate such energies from n independent draws, over a bounded Hölder class of densities of smoothness s. For sources depending polynomially on f with a nondegenerate local symbol, the minimax root mean squared error on the flat torus is of order $n^{-\min(1/2,4s/(4s+d))}$, as for quadratic density functionals; the upper bound also holds on compact manifolds for source fields tangent to the boundary. The estimator corrects a pilot by quadratic and cubic influence terms that keep both mixed frequency blocks. Above $s = d/4$ it is asymptotically linear with the canonical influence function, while on the torus at $s = d/4$ no estimator has uniform centered Gaussian calibration with bounded scale, even near the uniform density. These conclusions hold for a diagonal entry of the inverse diffusion metric on every smooth compact connected manifold, with or without reflecting boundary, and for every nonzero symmetric contrast of the effective tensor of a periodic scalar conductor sampled proportionally to its conductivity. In a fixed direction the densities with no conductivity loss are exactly those constant along that direction, which are the zeros of the influence variance. Detecting a loss is Ingster's goodness-of-fit problem for a direction-weighted L^2 seminorm with this composite null; its separation rate $n^{-4s/(4s+d)}$, for every $s > 0$, is attained by a weighted Fourier statistic, adaptively at an iterated-logarithm cost, with exact calibration by Haar averaging.

Keywords. Effective conductivity, homogenization, inverse diffusion metric, minimax rates, higher-order influence functions, goodness-of-fit testing, nonparametric density model.
MSC 2020. Primary 62G05, 62G10, 62G20. Secondary 62C20, 35B27, 35J15.

1 Introduction

Effective conductivity describes the macroscopic response of a heterogeneous medium, and the inverse diffusion metric describes the energy of a density-weighted Laplacian. Both are energies of the form $\langle L, A_f^{-1} L\rangle$ in which the elliptic operator $A_f = -\operatorname{div}(f\nabla)$ and the source L both depend on the sampling density f that generates the data. This article determines how well such energies can be estimated, and how well a loss of effective conductivity can be detected, from n independent draws with density f.

Edition note. This second edition incorporates the Archive article *Sharp minimax rates for inverse diffusion metric inference* [18] into a general theorem for density-dependent elliptic energies (Part I), writes out the manifold and reflecting-boundary extension in full (Part II), derives the effective-conductivity results of edition 1 from the general theorem (Part III), and places the testing results in Ingster's minimax testing framework.

The questions. For quadratic integral functionals of a density the minimax rate over smoothness-s classes is $n^{-1/2}$ for $s \geq d/4$ and $n^{-4s/(4s+d)}$ below: Bickel and Ritov [4] found the threshold for integrated squared derivatives, Birgé and Massart [5] gave the general lower bound for smooth integral functionals, and Laurent [11] constructed efficient estimators above the threshold. The energies studied here are nonlocal, because the inverse operator A_f^{-1} depends on the density being estimated, and their expansion about a pilot contains cubic and higher-order terms (Sections 4 and 6). The comparison with that classical theory leads to three questions. (i) What is the minimax rate for such an energy, uniformly over a Hölder class, and above which smoothness is it estimated efficiently? (ii) At the critical smoothness $s = d/4$, can a root-n estimator be calibrated uniformly by a Gaussian limit? (iii) In a fixed direction e the effective conductivity $e^\top K_f e$ never exceeds $|e|^2$; which densities have no loss, and how small a loss can be detected? The third question is the minimax goodness-of-fit problem of Ingster [9, 10], with the L^2 norm replaced by a direction-weighted L^2 seminorm and the simple null replaced by a composite invariant null.

Answers. For every energy whose local symbol (Section 2) is nonzero in some direction, Theorems 1 and 2 answer (i) on the flat torus: for every $s > 0$ the minimax root mean squared error is $n^{-\min(1/2, 4s/(4s+d))}$ up to constants, and above $s = d/4$ an estimator is asymptotically linear with the canonical influence function, and asymptotically normal where the influence variance is positive. The upper bound and the expansion hold on compact manifolds for source fields tangent to the boundary. At $s = d/4$ the answer to (ii) is negative: no estimator with a uniformly bounded scale has a uniform centered Gaussian limit, already on every fixed L^∞ neighborhood of the uniform density. Part II (Theorems 4 and 5) gives both answers for the inverse diffusion metric on every smooth compact connected manifold, with or without reflecting boundary, and Part III (Theorem 6) gives them for the effective tensor of a periodic scalar conductor and every nonzero symmetric contrast. Theorem 7 answers (iii): the densities with no loss in direction e are exactly those constant along e, which are also exactly the zeros of the influence variance, and the minimax separation rate for detecting a loss is $n^{-4s/(4s+d)}$ for every $s > 0$. Theorem 8 attains it adaptively in s at an iterated-logarithm cost, and Part VI gives exact finite-sample calibration by Haar averaging. For instance, in dimension three at smoothness $s = 3$ the effective conductivity is estimated at the root-n rate, while a loss of order $n^{-4/5}$ in a fixed direction is already detectable, because the no-loss null consists of zeros of the influence variance.

Method. Section 4 gives an exact expansion of $\Phi(f) - \Phi(p)$ about a pilot p, using the uniform elliptic estimates of Section 3. The estimator of Sections 5 and 6 corrects the pilot by U-statistics for the quadratic and cubic terms: the quadratic kernel keeps both mixed frequency blocks, and the cubic kernel uses one unprojected slot. The lower bounds of Section 7 perturb the density by oscillating templates whose first variation cancels, so that the energy moves by the local symbol times the squared perturbation norm while a sign mixture of the alternatives stays close to the null in total variation.

Relation to the literature. Robins et al. developed higher-order influence functions [13] and the quadratic von Mises calculus [14]. The cubic correction of Section 6 is an instance of the higher-order template of [13]. In their accounting an order-m correction with a pilot converging at n^{-v}, $v = x/(1+2x)$, $x = s/d$, reaches the rate $n^{-\gamma}$, $\gamma = 4x/(1+4x)$, once $m+1 \geq \gamma/v = 4(1+2x)/(1+4x)$. This ratio is smaller than 4 for every $x > 0$, so third order suffices at every smoothness. Our contribution is the operator that represents the quadratic and cubic variation of the nonlocal energy and the oscillation construction that gives a signed energy increment for every nondegenerate

symbol. For comparison, Tchetgen et al. [22] use distinct projection levels for a power of a density; here one unprojected slot controls both tails of the nonlocal trilinear term with only two resolutions.

Related Archive articles. The inverse diffusion metric of Part II is the entry G^0_{uu} of [15, (2.1)]. Edition 3 of that article computes a shrinking-bandwidth heat-kernel Sinkhorn estimator of this population target at a fixed smooth positive density and gives its Gaussian limit, covariance and bootstrap with near-linear cost. Theorem 4 is the statistical complement: it gives the minimax rate for the same target uniformly over finite-smoothness classes, and the efficient influence function to which a root-n estimator such as that of [15] must conform above $s = d/4$. The Darcy articles [19, 20, 16, 17] study linear functionals of a conductivity in the Darcy inverse problem, observed through noisy solutions of the elliptic equation. Their rates are governed by the transport and exit-time geometry of the reference flow. The present model observes the conductivity through density sampling, and the rates are governed by the local symbol of the source.

Part I
A general theorem for density-dependent elliptic energies

This part proves the upper bound, the lower-bound template and the critical calibration failure once, for a general family of sources. The later parts check its hypotheses.

2 Setting

Let $\mathcal{X}$ be either the flat torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ or a smooth compact connected d-dimensional Riemannian manifold with smooth, possibly empty, boundary. Normalize the volume to one and write $\int$ for integration against it. Fix $0 < m < 1 < M$ and $B > \|1\|_{C^s}$, and let $\mathcal{F}_s$ be the set of densities f with $m \le f \le M$ and $\|f\|_{C^s} \le B$. At noninteger s the norm is the classical Hölder norm, and at integer s it is the supremum of derivatives through order s; on a manifold both are taken in a fixed finite smooth atlas up to the boundary. The uniform density lies in the interior of $\mathcal{F}_s$. The data are n independent draws with density f.

Write $P_f v = \int fv$. For pilot densities p the operator $A_p = -\operatorname{div}(p\nabla)$ acts variationally on the space $H^1_\circ$ of H^1 functions with zero volume mean: for a bounded linear functional r on H^1 vanishing on constants, $A_p^{-1} r$ is the unique $w \in H^1_\circ$ with $\int p\nabla w \cdot \nabla\zeta = r(\zeta)$ for all $\zeta \in H^1$. On a manifold with boundary this is the Neumann inverse. We write $\langle r, w\rangle = r(w)$.

Fix $r \ge 1$, a real symmetric $r \times r$ matrix $\Gamma \ne 0$, smooth vector fields $V_1, \dots, V_r$ tangent to $\partial\mathcal{X}$, and smooth functions $w_1, \dots, w_r$. The sources are

$$\ell_i(f)(\zeta) = -\int fV_i \cdot \nabla\zeta + \int f\,(w_i - P_f w_i)\,\zeta, \tag{1}$$

that is $\ell_i(f) = \operatorname{div}(fV_i) + f(w_i - P_f w_i)$ in weak form. Each vanishes on constants because $\int f = 1$. The energy is

$$\Phi(f) = \sum_{i,j=1}^{r} \Gamma_{ij} \langle \ell_i(f), A_f^{-1} \ell_j(f)\rangle.$$

Put $\phi_i^p = A_p^{-1}\ell_i(p)$, $\phi_i^{p,\circ} = \phi_i^p - P_p\phi_i^p$,

$$\tilde{\psi}_p = \sum_{i,j}\Gamma_{ij}\big[-2V_i\cdot\nabla\phi_j^p + 2(w_i - P_pw_i)\phi_j^{p,\circ} - \nabla\phi_i^p\cdot\nabla\phi_j^p\big], \qquad \psi_p = \tilde{\psi}_p - P_p\tilde{\psi}_p.$$

Local symbol. At the uniform density put $G_i = V_i + \nabla\phi_i^1$. For an interior point x_0 and a unit vector $e \in T_{x_0}\mathcal{X}$ set

$$\sigma(x_0, e) = \sum_{i,j}\Gamma_{ij}\,(e\cdot G_i(x_0))(e\cdot G_j(x_0)).$$

The energy is *nondegenerate* if $\sigma(x_0, e) \neq 0$ for some interior x_0 and some e.

Theorem 1 (upper bound)**.** *On $\mathcal{X}$ as above, the minimax risk $\mathcal{R}_n(s) = \inf_{\widehat{\Phi}}\sup_{f\in\mathcal{F}_s}(\mathbb{E}_f|\widehat{\Phi} - \Phi(f)|^2)^{1/2}$ satisfies $\mathcal{R}_n(s) \le Cn^{-1/2}$ for $s \ge d/4$ and $\mathcal{R}_n(s) \le Cn^{-4s/(4s+d)}$ for $0 < s < d/4$. For $s > d/4$ there is an estimator with*

$$\widehat{\Phi} - \Phi(f) = \frac{1}{n}\sum_{i=1}^n \psi_f(X_i) + o_{\mathbb{P}_f}(n^{-1/2})$$

uniformly over $\mathcal{F}_s$. At each interior density ψ_f is the canonical gradient of Φ in the nonparametric density model, the estimator is regular, and at densities with $\operatorname{Var}_f\psi_f > 0$ it is asymptotically normal with variance $\operatorname{Var}_f\psi_f$, which is estimated consistently.

Theorem 2 (lower bound and critical calibration)**.** *Let $\mathcal{X} = \mathbb{T}^d$ and let the energy be nondegenerate. Then $\mathcal{R}_n(s) \ge cn^{-\min(1/2, 4s/(4s+d))}$ for every $s > 0$. At $s = d/4$ there is no sequence of estimators and positive scales $\sigma_n(f)$ bounded above by a constant on $\mathcal{F}_s$ such that $\sqrt{n}(\widehat{\Phi} - \Phi(f))/\sigma_n(f)$ converges to the standard Gaussian in Kolmogorov distance uniformly over $\mathcal{F}_s$. The same holds on every fixed L^∞ neighborhood of 1 in $\mathcal{F}_s$, and for positive random scales with a uniform deterministic upper bound.*

Theorem 5 gives the lower bound on manifolds for the inverse diffusion metric.

3 Elliptic estimates

Lemma 3. *Fix $0 < \alpha < \min(s, 1)$. There is C depending only on $\mathcal{X}$, d, m, M, B, α, V_i and w_i such that for all $q \in \mathcal{F}_s$, all vector fields V, all scalar g with $\int g = 0$ and $t \in \{2, 3, 4\}$,*

$$\|\phi_i^q\|_{W^{1,\infty}} \le C, \qquad \|\nabla A_q^{-1}\operatorname{div} V\|_t \le C\|V\|_t, \qquad \|\nabla A_q^{-1}g\|_t \le C\|g\|_t. \tag{2}$$

Here $\operatorname{div} V$ means the functional $\zeta \mapsto -\int V\cdot\nabla\zeta$.

Proof. Charts and reflection. Cover $\mathcal{X}$ by finitely many interior coordinate balls and boundary collars with boundary-normal coordinates (y', y_d), $y_d \ge 0$, in which the metric is $dy_d^2 + h_{y_d}(y')$. In coordinates the weak equation $\int q\nabla w\cdot\nabla\zeta = -\int V\cdot\nabla\zeta + \int g\zeta$ has coefficient matrix $a = q\mathcal{J}(g^{kl})$ relative to coordinate volume, where $\mathcal{J}$ is the volume Jacobian; the entries a^{kd}, $k < d$, vanish. Let $R(y', y_d) = (y', -y_d)$. Extend w and g evenly, extend a^{dd} and the tangential block evenly, keep the mixed entries zero, and extend the tangential components of V evenly and its normal component oddly. For a test function ζ on the doubled chart, the folded function $\zeta + \zeta\circ R$ restricted to $y_d \ge 0$ is a test function for the Neumann problem, and substituting it shows that the extensions solve the weak equation in the doubled chart. The extended coefficient is continuous, uniformly elliptic with bounds depending on m, M and the metric, and its C^α seminorm is at most twice the original one.

Hence on all charts the coefficients have a common modulus of continuity determined by B, α and the metric; in particular a common VMO modulus.

L^t gradient bounds. For $t > 2$ the interior $W^{1,t}$ estimate for divergence-form equations with VMO coefficients [3, §3] gives, on concentric balls $B \subset 2B$ in each (doubled) chart, $\|\nabla w\|_{L^t(B)} \le C(\|V\|_{L^t(2B)} + \|w\|_{L^t(2B)})$, with C depending only on d, t, the ellipticity bounds and the VMO modulus, hence only on m, M, B, α and the metric. Summing over a finite cover, $\|\nabla w\|_t \le C(\|V\|_t + \|w\|_t)$. Since $W^{1,t}$ embeds compactly in L^t and L^t continuously in L^2 on the compact $\mathcal{X}$, Ehrling's lemma gives $\|w\|_t \le \epsilon\|\nabla w\|_t + C_\epsilon\|w\|_2$, with constants depending only on $\mathcal{X}$. Poincaré's inequality and the energy estimate give $\|w\|_2 \le C\|\nabla w\|_2 \le Cm^{-1}\|V\|_2 \le C\|V\|_t$. Choosing ϵ small and absorbing proves the bound for divergence data. The a priori finiteness of $\|\nabla w\|_t$ used in the absorption holds for smooth coefficients and data; for general $q \in \mathcal{F}_s$ mollify q in the atlas, which preserves the ellipticity and the C^α bound up to a factor, and pass to the limit using H^1 convergence of the solutions and weak lower semicontinuity of the L^t norm. For scalar data, let ζ_g be the mean-zero Neumann solution of $\Delta\zeta_g = g$ for the background Laplacian. The $W^{2,t}$ estimate for this smooth Neumann problem [1] gives $\|\nabla\zeta_g\|_t \le C\|g\|_t$, and $\int g\zeta = -\int \nabla\zeta_g\cdot\nabla\zeta$ for all $\zeta \in H^1$, so g is divergence data with field $\nabla\zeta_g$. The case $t = 2$ is the energy estimate.

Gradient supremum. The source $\ell_i(q)$ has vector datum qV_i, which is C^α and, since V_i is tangent to the boundary, has normal component vanishing on $\partial\mathcal{X}$; its odd extension is therefore C^α. The scalar datum $q(w_i - P_q w_i)$ is bounded. Mollify q and the data in the atlas as in the previous paragraph; the mollified problems have smooth solutions ϕ_k, which converge to ϕ_i^q in H^1. For each k the interior divergence-form Schauder estimate [7, Theorem 8.32] in each doubled chart bounds $\nabla\phi_k$ in C^α by the supremum of ϕ_k and the data norms, and the De Giorgi–Nash–Moser bound [7, Theorem 8.17] controls the supremum by the L^2 norm, which is bounded by the energy estimate. The constants depend on the ellipticity and the C^α bound of the coefficient only, so they are uniform in k. By the Arzelà–Ascoli theorem a subsequence of $(\phi_k, \nabla\phi_k)$ converges uniformly; the H^1 convergence identifies the limit as $(\phi_i^q, \nabla\phi_i^q)$, which therefore obeys the same $C^{1,\alpha}$ bound. □

4 Exact expansion

For $p, f \in \mathcal{F}_s$ put $a = f - p$, so $\int a = 0$. The source splits as $\ell_i(f) = \ell_i(p) + \ell_{i,1}(a) + \ell_{i,2}(a)$ with

$$\ell_{i,1}(a) = \operatorname{div}(aV_i) + a(w_i - P_p w_i) - p\int a w_i, \qquad \ell_{i,2}(a) = -a\int a w_i.$$

Since $A_f\phi_i^p = \ell_i(p) - \operatorname{div}(a\nabla\phi_i^p)$ we have $\ell_i(f) = A_f\phi_i^p + \rho_i$ with $\rho_i = \rho_{i,1} + \rho_{i,2}$,

$$\rho_{i,1} = \ell_{i,1}(a) + \operatorname{div}(a\nabla\phi_i^p), \qquad \rho_{i,2} = \ell_{i,2}(a).$$

Expanding $\langle A_f\phi_i^p + \rho_i, \phi_j^p + A_f^{-1}\rho_j\rangle$ and using $\langle A_f\phi_i^p, \phi_j^p\rangle = \langle \ell_i(p), \phi_j^p\rangle + \int a\nabla\phi_i^p\cdot\nabla\phi_j^p$ gives the exact identity

$$\Phi(f) - \Phi(p) = \int a\psi_p + 2\sum_{i,j}\Gamma_{ij}\langle\rho_{i,2}, \phi_j^p\rangle + \sum_{i,j}\Gamma_{ij}\langle\rho_i, A_f^{-1}\rho_j\rangle. \tag{3}$$

Indeed, the linear part is $\sum\Gamma_{ij}[2\langle\ell_{i,1}(a), \phi_j^p\rangle - \int a\nabla\phi_i^p\cdot\nabla\phi_j^p]$, and $\langle\ell_{i,1}(a), \phi_j^p\rangle = \int a[-V_i\cdot\nabla\phi_j^p + (w_i - P_p w_i)\phi_j^{p,\circ}]$, since $\int a = 0$; adding a constant to $\tilde\psi_p$ does not change $\int a\tilde\psi_p$. Testing (3) along $p + \tau a$ shows $\int \psi_p\, a$ is the derivative of Φ at p in direction a.

By Lemma 3, $\|\rho_{i,1}\|_{H^{-1}} \le C\|a\|_2$, $\|\rho_{i,2}\|_{H^{-1}} \le C\|a\|_2^2$, and for $t \in \{2, 3, 4\}$,

$$\|\nabla A_q^{-1}\rho_{i,1}\|_t \le C\|a\|_t, \qquad \|\nabla A_q^{-1}\rho_{i,2}\|_t \le C\|a\|_t\|a\|_2 \qquad (q \in \mathcal{F}_s), \tag{4}$$

using $|\int a w_i| \le C\|a\|_2$ and the bounded fields $V_i, \nabla\phi_i^p$. Put

$$Q_p(a) = \sum_{i,j} \Gamma_{ij}\,[2\langle \rho_{i,2}, \phi_j^p\rangle + \langle \rho_{i,1}, A_p^{-1}\rho_{j,1}\rangle].$$

The terms of (3) involving $\rho_{i,2}$ inside the resolvent pairing are bounded by $C\|a\|_3^3$. The remaining difference is

$$\langle \rho_{i,1}, (A_f^{-1} - A_p^{-1})\rho_{j,1}\rangle = -\int a\,\nabla A_f^{-1}\rho_{i,1}\cdot\nabla A_p^{-1}\rho_{j,1},$$

bounded by $C\|a\|_3^3$ by (4) with $t = 3$. Hence

$$\Big|\Phi(f) - \Phi(p) - \int a\psi_p - Q_p(a)\Big| \le C\|a\|_3^3. \tag{5}$$

The form Q_p is bounded by $C\|a\|_2^2$, so polarization and the Riesz theorem give a self-adjoint operator H_p on mean-zero L^2, extended by zero on constants, with $\|H_p\| \le C$ and $Q_p(a) = \langle a, H_p a\rangle$. Finally,

$$\|\psi_f - \psi_p\|_2 \le C\|a\|_2, \tag{6}$$

because $\phi_i^f - \phi_i^p = A_f^{-1}\rho_i$ has H^1 norm at most $C\|a\|_2$, the gradients are bounded by (2), and the centering constants differ by at most $C\|a\|_2$.

5 Quadratic correction and the upper bound for $s \ge d/4$

Approximation operators. We use a self-adjoint finite-rank operator $S = S_J$ of resolution J with a representation $Sg = \sum_{\ell=1}^K e_\ell\langle e_\ell, g\rangle$, $K \asymp J^d$, such that

$$\sum_\ell |\langle e_\ell, g\rangle|^2 \le \|g\|_2^2, \quad \sum_\ell |e_\ell(x)|^2 \le CK, \quad \|S\|_{L^t\to L^t} \le C, \quad \|(I-S)g\|_\infty \le CK^{-s/d}\|g\|_{C^s}, \tag{7}$$

for $t \in \{2, 4\}$. On $\mathbb{T}^d$ take the Fourier multiplier with symbol $\chi(k/J)$, where $\chi \in [0,1]$ is smooth, compactly supported and equal to one near zero, and even, and let e_ℓ be the real trigonometric basis functions 1, $\sqrt{2}\cos(2\pi k\cdot x)$ and $\sqrt{2}\sin(2\pi k\cdot x)$, with one representative k of each pair $\pm k \ne 0$, times $\chi(k/J)^{1/2}$. All functions and pairings are then real. Its convolution kernel is the periodization of the dilate at scale J^{-1} of a Schwartz function with vanishing moments of all positive orders, so it has bounded L^1 norm and Taylor's formula gives the last bound. Section 9.1 constructs S on manifolds. S need not be a projection.

Pilot. On an independent subsample of size proportional to n, form $\bar f = n^{-1}\sum_i k_J(\cdot, X_i)$ with the kernel $k_J(x,y) = \sum_\ell e_\ell(x)e_\ell(y)$ of S_J and $J \asymp n^{1/(2s+d)}$. Its mean is $S_J f$, so its bias is $O(J^{-s})$ in supremum norm. By (7), $|k_J| \le CJ^d$ and $\int k_J(x,y)^2 dy \le CJ^d$, so Rosenthal's inequality gives

$$\mathbb{E}|\bar f(x) - S_J f(x)|^{2t} \le C\{(nJ^{-d})^{-t} + (nJ^{-d})^{-(2t-1)}\}, \qquad t = 3, 4.$$

Let p be the first element of a fixed countable L^t-dense subset of $\mathcal{F}_s$ whose L^t distance to $\bar f$ is within n^{-1} of the infimum. The selection is measurable, takes countably many values, and $\|p - f\|_t \le 2\|\bar f - f\|_t + n^{-1}$. Since the volume is one,

$$\sup_{f\in\mathcal{F}_s} \mathbb{E}_f\|p - f\|_3^6 \le Cn^{-6s/(2s+d)}, \qquad \sup_{f\in\mathcal{F}_s} \mathbb{E}_f\|p - f\|_4^8 \le Cn^{-8s/(2s+d)}. \tag{8}$$

Quadratic kernel. Define

$$B_{p,J} = H_pS + SH_p - SH_pS, \qquad H_p - B_{p,J} = (I-S)H_p(I-S).$$

The identity is algebraic and holds for every S. The operator $B_{p,J}$ has rank at most $2K$ and is represented by the explicit kernel

$$B(x,y) = \sum_{\ell} \left[(H_pe_\ell)(x)e_\ell(y) + e_\ell(x)(H_pe_\ell)(y)\right] - \sum_{\ell,k} e_\ell(x)\langle e_\ell, H_pe_k\rangle e_k(y),$$

built from the representation of S and not from an orthonormal basis. Its Hilbert–Schmidt norm is at most $3\|H_p\|\|S\|_{\mathrm{HS}}$ and $\|S\|_{\mathrm{HS}}^2 \le \int \sum_\ell |e_\ell|^2 \le CK$, by the synthesis bound that follows from Bessel's inequality in (7). The bias bound uses the class, not merely L^2: both f and p lie in $\mathcal{F}_s$, so $\|a\|_{C^s} \le 2B$ and (7) gives $\|(I-S)a\|_2 \le CK^{-s/d}$; hence

$$|Q_p(a) - \langle a, B_{p,J}a\rangle| \le \|H_p\|\|(I-S)a\|_2^2 \le CK^{-2s/d}. \tag{9}$$

Both mixed frequency blocks are thus estimated as well as the low-frequency block. Let $B^{[p]}$ be the centering of B in each variable against p. Because p takes countably many values, one measurable kernel is fixed for each value. On $N \asymp n$ new observations set

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

The quadratic term has conditional mean $\langle a, B_{p,J}a\rangle$, first Hoeffding projection $B_{p,J}a$ centered under f, with $L^2(f)$ norm at most $C\|a\|_2$, and second projection the double f-centering of B, with squared $L^2(f\otimes f)$ norm at most CK. With (5), (8) and bounded ψ_p,

$$\sup_f \left(\mathbb{E}_f|\widehat{\Phi}_p - \Phi(f)|^2\right)^{1/2} \le C\{n^{-1/2} + \sqrt{K}/n + K^{-2s/d} + n^{-3s/(2s+d)}\}. \tag{10}$$

For $s \ge d/4$ take $K \asymp n$: every term is $O(n^{-1/2})$.

Efficiency above $s = d/4$. Take $n^{d/(4s)} \ll K \ll n$. The bias and the degenerate quadratic term are then $o(n^{-1/2})$ in root mean square, and by (6)–(8) the centered empirical mean of ψ_p may be replaced by that of ψ_f. Split the sample into two nearly equal folds, use each as pilot for the other, and weight the two estimates by validation size; this gives the full-sample expansion. Averaging squared pilot influences in the validation folds estimates $\mathrm{Var}_f\,\psi_f$ consistently, by boundedness and (6). At an interior density, scores δ/f of smooth mean-zero additive directions δ are dense in $L_0^2(f)$ and (5) gives the pathwise derivative $\int \psi_f\delta$, so ψ_f is the canonical gradient. The uniform remainder along local paths, (6) and the triangular-array central limit theorem give regularity and the stated normal limit.

6 Cubic correction below the threshold

Let $A_a = -\operatorname{div}(a\nabla)$. The resolvent identity

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

and $w_i = A_p^{-1}\rho_{i,1}$ give the cubic term of (3),

$$C_p(a,a,a) = \sum_{i,j}\Gamma_{ij}\Big[2\langle\rho_{i,2}, w_j\rangle - \int a\,\nabla w_i\cdot\nabla w_j\Big],$$

and C_p denotes its symmetric trilinear polarization, applied after removing the volume mean of each input. The residual after this term is at most $C\|a\|_4^4$: the last resolvent term is bounded by $C\|a\nabla A_p^{-1}\rho_i\|_2\|a\nabla A_p^{-1}\rho_j\|_2 \le C\|a\|_4^4$ by (4), and the other terms of degree at least four follow from Hölder's inequality, (4) and $\|a\|_\infty \le C$.

Each summand of $C_p(a_1,a_2,a_3)$ is $\int (L_1a_1)(L_2a_2)(L_3a_3)$, where each L_k is mean removal, a moment $a \mapsto \int aw_i$ viewed as a constant function, the solution map $a \mapsto w_i$, or a gradient component of it; all are bounded on L^2 and L^4 by Lemma 3. Retain the all-S term and each term with exactly 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 at least two $(I-S)a$ inputs; placing these in L^4 and the third in L^2 bounds the bias by $CK^{-2s/d}\|a\|_2$.

Variance. In each summand write the S arguments through their coefficients $\langle e_\ell, g\rangle$. If j arguments are random and the others equal a, the kernel has squared L^2 norm $O(K^{j-1})$. To see this, take the L^2 adjoint of L_k in one live slot, choosing the identity slot when it is live, and use Bessel's inequality from (7) if it carries S. Bound each other live slot by its square function, using

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

which follows from Khintchine's inequality applied to Rademacher sums $\sum_\ell \epsilon_\ell e_\ell$ before and after the L^4-bounded map L. Fixed slots are bounded in L^4. Use three mutually independent validation groups of size of order n, independent of the pilot, and evaluate the kernel on the product of the three empirical measures centered at $p\,dx$. Expanding each about $a\,dx$, terms with different sets of live slots are conditionally orthogonal, and a set of size j contributes variance $O(n^{-j}K^{j-1})$, since centering is a contraction in $L^2(f^{\otimes j})$ and $f \le M$. The cubic standard deviation is thus $O(n^{-1/2} + \sqrt{K}/n + K/n^{3/2})$. The linear and quadratic terms may use the same groups, by the triangle inequality in root mean square.

Rates. Put $x = s/d < 1/4$, $v = x/(1+2x)$ and $\gamma = 4x/(1+4x)$. Use the L^4 pilot of (8), quadratic rank of order $n^{2/(1+4x)}$, and cubic rank of order n^κ with $\kappa = 2/(1+4x) - 1/(2(1+2x))$. The quadratic bias and standard deviation are $O(n^{-\gamma})$. The cubic bias is $O(n^{-2x\kappa-v}) = O(n^{-\gamma})$ in root mean square, and the quartic pilot remainder is $O(n^{-4v}) = o(n^{-\gamma})$ because $4v-\gamma = 8x^2/((1+2x)(1+4x)) > 0$. If $\kappa \le 1$ the cubic variance is $O(n^{-1})$; otherwise its dominant standard deviation is $O(n^{\kappa-3/2}) = O(n^{-\gamma})$, because $3/2-\kappa-\gamma = 1/2-(1+4x)^{-1}+[2(1+2x)]^{-1} > 0$. Since $\gamma < 1/2$ for $x < 1/4$, this proves Theorem 1.

7 Oscillation lower bounds and critical calibration

Let $\mathcal{X} = \mathbb{T}^d$, and fix x_0 and a unit e with $\sigma := \sigma(x_0,e) \neq 0$. Write ψ_1 for the influence at $f=1$ and use coordinates rotated so that e is the first axis; Fourier modes are those of the original torus.

Template, fixed first. Let $N_0 = r+2$ and $\epsilon_0 > 0$, to be chosen in terms of σ and Γ. In the unit cube fix N_0 smooth bumps $\beta_1,\dots,\beta_{N_0}$ with disjoint supports in its interior and set $\zeta_\ell = \beta_\ell \sin(Lx_1)$. The transverse gradient norms of ζ_ℓ are bounded independently of L and $\|\partial_1\zeta_\ell\|_2 \ge cL$; fix L so that $\|\nabla_\perp\zeta_\ell\|_2 \le \epsilon_0\|\partial_1\zeta_\ell\|_2$ for all ℓ. All template constants below, written C_ζ, are now fixed.

Cube, chosen second. Choose a cube Q of side δ centered at x_0, with δ so small that $\sup_Q |G_i - G_i(x_0)| \le \epsilon_0$ for all i. Inside it place $m_h \asymp \delta^d h^{-d}$ cells of side h at mutual distances of order h. In cell j with corner x_j choose a Euclidean unit vector $c_j \in \mathbb{R}^{N_0}$ in the kernel of the $r+1$ linear conditions

$$\int \psi_1\,\partial_1 v_j = 0, \qquad \int w_i\,\partial_1 v_j = 0 \ (i \le r), \qquad v_j = \rho h^{s+1}\sum_{\ell=1}^{N_0} c_{j\ell}\,\zeta_\ell((x-x_j)/h),$$

and put $a_j = \partial_1 v_j$, $a_\epsilon = \sum_j \epsilon_j a_j$, $V = \sum_j \epsilon_j v_j$ for signs $\epsilon_j = \pm 1$. Then $\int a_j = 0$, and by disjointness, uniformly over unit c_j,

$$\|a_j\|_2^2 \asymp \rho^2 h^{2s+d}, \quad \|a_j\|_\infty \le C_\zeta \rho h^s, \quad \|V\|_2 \le C_\zeta h \|a_\epsilon\|_2, \quad \|\nabla_\perp V\|_2 \le \epsilon_0 \|a_\epsilon\|_2.$$

For a sufficiently small fixed ρ every $f_\epsilon = 1 + a_\epsilon$ lies in $\mathcal{F}_s$, by scaling the derivatives in each cell and, at noninteger s, by the separation of the cells.

Signed gap. The conditions give $\int a_\epsilon \psi_1 = 0$ and $\int a_\epsilon w_i = 0$, so $\rho_{i,2} = 0$ and $\rho_i = \operatorname{div}(a_\epsilon G_i) + a_\epsilon(w_i - \int w_i)$ in (3), which reduces to

$$\Phi(f_\epsilon) - \Phi(1) = \sum_{i,j} \Gamma_{ij} \langle \rho_i, A_{f_\epsilon}^{-1} \rho_j \rangle.$$

Replacing $A_{f_\epsilon}^{-1}$ by $A_1^{-1} = (-\Delta)^{-1}$ costs $C\|a_\epsilon\|_\infty \|a_\epsilon\|_2^2 \le C_\zeta \rho h^s \|a_\epsilon\|_2^2$. The scalar part has $\|a_\epsilon(w_i - \int w_i)\|_{H^{-1}} \le C\|V\|_2 \le C_\zeta h \|a_\epsilon\|_2$, by integrating ∂_1 by parts onto the test function, and $\|\operatorname{div}(a_\epsilon(G_i - G_i(x_0)))\|_{H^{-1}} \le \epsilon_0 \|a_\epsilon\|_2$. For the constant vectors $g_i = G_i(x_0)$, Parseval gives

$$\sum_{i,j} \Gamma_{ij} \langle \operatorname{div}(a g_i), (-\Delta)^{-1} \operatorname{div}(a g_j) \rangle = \sum_{k \ne 0} \frac{\sum_{i,j} \Gamma_{ij} (g_i \cdot k)(g_j \cdot k)}{|k|^2} |\widehat{a}(k)|^2. \tag{11}$$

Write $k = (e \cdot k)e + k_\perp$. The symbol equals $\sigma(1 - |k_\perp|^2/|k|^2)$ plus terms bounded by $C|k_\perp|/|k|$, and since $a_\epsilon = \partial_1 V$, $\sum_k (|k_\perp|^2/|k|^2) |\widehat{a}_\epsilon(k)|^2 \le \|\nabla_\perp V\|_2^2 \le \epsilon_0^2 \|a_\epsilon\|_2^2$. By Cauchy–Schwarz the right side of (11) is $\sigma \|a_\epsilon\|_2^2 + O(\epsilon_0 \|a_\epsilon\|_2^2)$. Collecting terms,

$$\Phi(f_\epsilon) - \Phi(1) = \sigma \|a_\epsilon\|_2^2 + O\big((\epsilon_0 + C_\zeta h + C_\zeta \rho h^s) \|a_\epsilon\|_2^2\big),$$

where the O constant depends only on Γ and the fixed fields. Choosing ϵ_0 small in terms of σ and then h small, every alternative satisfies

$$\operatorname{sign}(\sigma)\,(\Phi(f_\epsilon) - \Phi(1)) \ge \tfrac{1}{2} |\sigma| \|a_\epsilon\|_2^2 \ge c \rho^2 h^{2s}, \qquad |\Phi(f_\epsilon) - \Phi(1)| \le C \rho^2 h^{2s}. \tag{12}$$

Mixture. Let Q be the uniform sign mixture of the n-fold product laws and P_0 the uniform product law. Disjoint supports and $\int a_j = 0$ give

$$1 + \chi^2(\mathsf{Q}, \mathsf{P}_0) = \mathbb{E}_{\epsilon, \epsilon'} \Big(1 + \sum_j \epsilon_j \epsilon_j' \|a_j\|_2^2\Big)^n \le \prod_j \cosh(n \|a_j\|_2^2) \le \exp(C_1 n^2 \rho^4 h^{4s+d}). \tag{13}$$

With $h = t n^{-2/(4s+d)}$ and small fixed t, total variation is bounded away from one; thresholding an estimator halfway into the gap (12) gives $\mathcal{R}_n(s) \ge c n^{-4s/(4s+d)}$.

Parametric bound. Fix one alternative $f_\epsilon = 1 + a$; it satisfies the defining inequalities of the class strictly along $1 + \tau a$, $0 \le \tau \le 1$. Since $\Phi(1 + a) \ne \Phi(1)$, the mean value theorem and the derivative formula of Section 4 give τ_* with $\int \psi_{1+\tau_* a}\, a \ne 0$. The pair $1 + \tau_* a$, $1 + (\tau_* + b/\sqrt{n})a$ has parameter gap of order $n^{-1/2}$ by (5) and product chi-square at most $\exp(C_2 b^2) - 1$, which gives $\mathcal{R}_n(s) \ge c n^{-1/2}$.

Critical calibration. At $s = d/4$ use $h = t n^{-1/d}$ and replace Φ and the estimator by their product with $\operatorname{sign}\sigma$. Every alternative shifts the root-n parameter up by at least $\delta_t = c \rho^2 t^{d/2}$, while (13) bounds total variation by $C_3 \rho^2 t^d$. Under uniform centered Gaussian calibration with scales bounded by M_0, the null gives the event $\{\widehat{\Phi} \le \Phi(1)\}$ probability $1/2 + o(1)$, and each alternative gives it at most $\Phi_{\mathrm{N}}(-\delta_t/M_0) + o(1)$, where Φ_{N} is the standard normal distribution function; this also holds for random scales bounded by M_0, since the event forces the studentized variable below $-\delta_t/M_0$. The

probability gap $c'\rho^2 t^{d/2} + o(1)$ exceeds the total variation bound for small fixed t. The alternatives converge uniformly to 1, which gives the local form. This proves Theorem 2.

Part II
Instance A: the inverse diffusion metric

8 Statement

Fix a smooth real nonconstant score u on $\mathcal{X}$ and put $u_f = u - P_f u$, $b(f) = f u_f$ and

$$F(f) = \langle b(f), A_f^{-1} b(f) \rangle = P_f |\nabla \phi_f|^2, \qquad \phi_f = A_f^{-1} b(f).$$

This is the energy of Part I with $r = 1$, $\Gamma = 1$, $V_1 = 0$ and $w_1 = u$, and the influence is

$$\psi_f = 2 u_f \phi_f^\circ - |\nabla \phi_f|^2 - F(f), \qquad \phi_f^\circ = \phi_f - P_f \phi_f.$$

On $\mathbb{T}^d$ it is the entry G^0_{uu} of [15, (2.1)].

Theorem 4. *On $\mathbb{T}^d$, the minimax risk for F is, up to constants, $\mathcal{R}_n(s) \asymp n^{-1/2}$ for $s \geq d/4$ and $\mathcal{R}_n(s) \asymp n^{-4s/(4s+d)}$ for $0 < s < d/4$. For $s > d/4$ an estimator has the uniform expansion of Theorem 1 with the influence ψ_f above, which is the canonical gradient at each interior density; its variance is positive at every $f \in \mathcal{F}_s$ and is estimated consistently. At $s = d/4$ the critical calibration failure of Theorem 2 holds, including on every fixed L^∞ neighborhood of the uniform density and for positive random studentizers with a uniform deterministic upper bound.*

Proof. The upper bounds and the expansion are Theorem 1. At $f = 1$, $A_1 = -\Delta$ and $\ell_1(1) = u_0 = u - \int u$, so $-\Delta \phi_0 = u_0 \neq 0$ for $\phi_0 = \phi_1^1$; the smooth potential ϕ_0 is therefore nonconstant and has nonzero gradient at some point x_0. With $G_1 = \nabla \phi_0$, $\sigma(x_0, e) = (e \cdot \nabla \phi_0(x_0))^2 > 0$ for $e = \nabla \phi_0(x_0)/|\nabla \phi_0(x_0)|$, so Theorem 2 applies. For the variance: $F(f) > 0$, since $b(f) \neq 0$ for nonconstant u. The function ϕ_f° is continuous with P_f-mean zero on the connected $\mathcal{X}$, so it vanishes at some x_*, where $\psi_f(x_*) \leq -F(f) < 0$. A continuous function with $P_f \psi_f = 0$ that is negative somewhere is nonzero in $L^2(f)$, and $f \geq m$. □

9 Manifolds and reflecting boundary

Theorem 5. *Theorem 4 holds on every smooth compact connected Riemannian manifold $\mathcal{X}$ with smooth, possibly empty, boundary, with normalized volume, intrinsic gradient and divergence, and the variational Neumann inverse in the definition of F.*

The proof uses Lemma 3, which is already proved on such $\mathcal{X}$, and the identities of Sections 4–6, which use only normalized volume. Two further steps complete the proof: the operators S_J of (7) and the lower bound.

9.1 Approximation operators

Cover $\mathcal{X}$ by finitely many coordinate charts U_i whose images are cubes (interior charts) or half-cubes with the boundary face on $\{y_d = 0\}$ (collar charts). Take smooth $\vartheta_i \geq 0$ subordinate to the cover with $\sum_i \vartheta_i > 0$ everywhere and put $\theta_i = \vartheta_i (\sum_k \vartheta_k^2)^{-1/2}$, so $\sum_i \theta_i^2 = 1$. Subdivide each chart

image into cubes of side comparable to J^{-1}, with the boundary face on the grid. On each cell Q let $P_{i,Q}$ be the orthogonal projection in $L^2(Q, \mathcal{J}_i\,dy)$, with $\mathcal{J}_i$ the volume Jacobian, onto polynomials of degree at most $\lceil s \rceil$. Let $P_i = \sum_Q P_{i,Q}$, extended by zero outside the chart, and

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

Kernel and representation. Rescaled to the unit cube, the Gram matrix of the monomials in $L^2(Q, \mathcal{J}_i\,dy)$, divided by the cell volume, is uniformly positive definite and bounded, because $\mathcal{J}_i$ is bounded above and below. Hence each cell has an orthonormal basis $\{\pi_{Q,\beta}\}$ of fixed size with $|\pi_{Q,\beta}| \le CJ^{d/2}$, and $P_{i,Q}$ has kernel bounded by CJ^d on $Q \times Q$. Put $e_{i,Q,\beta} = \theta_i \pi_{Q,\beta}$. Then $S_J g = \sum e_{i,Q,\beta}\langle e_{i,Q,\beta}, g\rangle$, S_J is self-adjoint in $L^2(\mathcal{X})$, and $K \asymp J^d$. Bessel's inequality in each chart and $\sum \theta_i^2 = 1$ give $\sum |\langle e_\ell, g\rangle|^2 \le \sum_i \|\theta_i g\|_2^2 = \|g\|_2^2$. Finite overlap and the pointwise basis bound give $\sum_\ell |e_\ell(x)|^2 \le CJ^d$. The kernel bound CJ^d on cells of volume $\asymp J^{-d}$ gives boundedness on every L^t, $1 \le t \le \infty$.

Approximation. Since $g = \sum_i \theta_i(\theta_i g)$, $(I - S_J)g = \sum_i \theta_i(I - P_i)(\theta_i g)$. Each $P_{i,Q}$ reproduces polynomials of degree at most $\lceil s \rceil \ge \lfloor s \rfloor$ and is bounded in supremum norm, so $\|(I - P_{i,Q})h\|_{L^\infty(Q)} \le C \inf_\pi \|h - \pi\|_{L^\infty(Q)} \le CJ^{-s}\|h\|_{C^s}$, by Taylor's formula with Hölder remainder at noninteger s and with the order-s derivative bound at integer s. Applied to $h = \theta_i g$, this gives the last bound of (7).

Use in the estimators. The operator S_J is not a projection. The quadratic correction of Section 5 uses only the algebraic identity $H_p - B_{p,J} = (I - S_J)H_p(I - S_J)$, the explicit kernel built from the non-orthogonal representation $\{e_\ell\}$, the bound $\|S_J\|_{\mathrm{HS}}^2 \le CK$, and the bias bound $\|(I - S_J)a\|_2 \le CK^{-s/d}$, which requires $a = f - p$ with $f, p \in \mathcal{F}_s$. All of these hold. The cubic correction uses (7) only. The pilot uses the kernel k_J of S_J, whose supremum is $O(J^d)$ by the pointwise sum of squares and whose squared L^2 norm in the sampled variable is $O(J^d)$ by Bessel's inequality; this gives the moment bounds and (8).

9.2 Lower bound on manifolds

Since ϕ_0 is nonconstant there is an interior point x_0 with $\nabla\phi_0(x_0) \ne 0$ (if $\nabla\phi_0$ vanished on the interior, ϕ_0 would be constant on the connected $\mathcal{X}$). Use normal coordinates at x_0 rotated so that $\nabla\phi_0(x_0)$ points along ∂_1. In coordinates let $\mathcal{J}$ be the volume Jacobian and (g^{kl}) the inverse metric, and write $\gamma^l = g^{kl}\partial_k\phi_0$ for the components of $\nabla\phi_0$.

Template first. In the unit cube fix three smooth bumps $\zeta_1, \zeta_2, \zeta_3$ with disjoint interior supports and $\partial_1\zeta_\ell \ne 0$. Let C_ζ bound $\|\nabla v\|_2/\|\partial_1 v\|_2$ and $\|v\|_2/\|\partial_1 v\|_2$ over unit-coefficient combinations $v = \sum c_\ell\zeta_\ell$; it is finite by compactness of the unit sphere and disjointness.

Cube second. Choose a coordinate cube Q around x_0 so small that $\gamma^1 \ge c_0 > 0$ and $|\gamma^l| \le c_0/(4\sqrt{d}\,C'_\zeta)$ for $l \ge 2$ on Q, where C'_ζ is C_ζ times the bounds on $\mathcal{J}^{\pm 1}$. In Q place $m_h \asymp h^{-d}$ separated cells of side h, and in cell j choose a unit $c_j \in \mathbb{R}^3$ satisfying the two conditions $\int u\,a_j = 0$ and $\int \psi_0\,a_j = 0$, where

$$v_j = \rho h^{s+1} \sum_{\ell=1}^{3} c_{j\ell}\zeta_\ell((y - y_j)/h), \qquad a_j = \mathcal{J}^{-1}\partial_1 v_j.$$

Then $\int a_j = \int \partial_1 v_j\,dy = 0$, and $\|a_j\|_2^2 \asymp \rho^2 h^{2s+d}$, $\|v_j\|_2 \le Ch\|a_j\|_2$, $\|\nabla v_j\|_2 \le C\|a_j\|_2$. For small fixed ρ all $f_\epsilon = 1 + \sum_j \epsilon_j a_j$ lie in $\mathcal{F}_s$.

Signed gap by duality. With $a = \sum \epsilon_j a_j$, $V = \sum \epsilon_j v_j$, the conditions give $\rho_{1,2} = 0$ and $\int a\psi_0 = 0$, so $F(f_\epsilon) - F(1) = \langle \rho, A_{f_\epsilon}^{-1}\rho\rangle$ with $\rho = au_0 + \mathrm{div}(a\nabla\phi_0)$. Test against $-V$ after subtracting its mean

(which ρ annihilates). In coordinates,

$$\langle \operatorname{div}(a\nabla\phi_0), -V\rangle = \int a\,\gamma^l \partial_l V\,\mathcal{J}\,dy = \int (\partial_1 V)^2 \gamma^1 dy + \sum_{l\geq 2}\int \partial_1 V\,\gamma^l \partial_l V\,dy.$$

The first term is at least $c_0\|\partial_1 V\|^2_{L^2(dy)}$; the second is at most $\frac{1}{4}c_0\|\partial_1 V\|^2_{L^2(dy)}$ by the choice of Q; and $|\langle au_0, V\rangle| \leq Ch\|a\|_2^2$. Hence $\langle \rho, -V\rangle \geq c\|a\|_2^2$ for small h, while the test function has energy at most $C\|a\|_2^2$. Variational duality $\langle \rho, A^{-1}\rho\rangle \geq \langle \rho, \zeta\rangle^2/\langle A\zeta, \zeta\rangle$ gives the lower half of (12), and $\|\rho\|_{H^{-1}} \leq C\|a\|_2$ the upper half. The chi-square bound (13), the parametric two-point bound and the critical comparison of Section 7 use only disjoint supports, zero masses and (5), and carry over verbatim. The variance positivity in Theorem 4 used only connectedness. This proves Theorem 5.

Part III
Instance B: effective conductivity

10 Cell problem and estimation

On $\mathbb{T}^d$ let the local scalar conductivity f be normalized by its known mean; the data are n independent locations with density f. If the known mean is $\mu \neq 1$, the physical effective tensor is μ times the tensor below. For $e \in \mathbb{R}^d$ define K_f by

$$e^\top K_f e = \min_{\phi \in H^1(\mathbb{T}^d)} \int_{\mathbb{T}^d} f|e + \nabla\phi|^2. \tag{14}$$

The minimizer $\chi_{f,e}$ is unique with zero mean and solves $\operatorname{div}(f(e+\nabla\chi_{f,e})) = 0$, the scalar symmetric instance of the periodic corrector equation [12, §2.1]. The fixed-cell model complements the random-realization setting of [2].

Completing the square in (14) gives $K_f = I - (\langle \partial_i f, A_f^{-1}\partial_j f\rangle)_{ij}$ with $\chi_{f,e_i} = A_f^{-1}\partial_i f$. For a real symmetric contrast $\Gamma \neq 0$ put $T_\Gamma(f) = \operatorname{tr}(\Gamma K_f)$. Then

$$T_\Gamma(f) = \operatorname{tr}\Gamma - \Phi(f), \qquad \ell_i(f) = \operatorname{div}(fe_i) = \partial_i f,$$

which is the energy of Part I with $V_i = e_i$ and $w_i = 0$. Here $\phi_i^f = \chi_{f,e_i}$, and with $q_{f,i} = e_i + \nabla\chi_{f,e_i}$ the influence of T_Γ is

$$\psi_{f,\Gamma} = \sum_{i,j} \Gamma_{ij}\, q_{f,i}\cdot q_{f,j} - T_\Gamma(f),$$

since $q_{f,i}\cdot q_{f,j} = \delta_{ij} + e_i\cdot\nabla\chi_j + e_j\cdot\nabla\chi_i + \nabla\chi_i\cdot\nabla\chi_j$ and $P_f(q_{f,i}\cdot q_{f,j}) = (K_f)_{ij}$. The identity (3) for this source,

$$T_\Gamma(f) - T_\Gamma(p) = \int a\psi_{p,\Gamma} - \sum_{i,j}\Gamma_{ij}\langle \rho_i, A_f^{-1}\rho_j\rangle, \qquad \rho_i = \operatorname{div}(aq_{p,i}), \tag{15}$$

is the polarized completion of the square used in edition 1.

Theorem 6. *For every fixed nonzero real symmetric $d\times d$ matrix Γ, the minimax root mean squared error for T_Γ over $\mathcal{F}_s$ is $n^{-\min(1/2, 4s/(4s+d))}$ up to constants. If $s > d/4$ an estimator is uniformly asymptotically linear with influence $\psi_{f,\Gamma}$, which is the canonical gradient at interior densities; at densities with nonzero influence variance it is asymptotically normal with the canonical variance. If $s = d/4$, no estimator has a uniform centered standard Gaussian limit after division by a positive uniformly bounded scale, even on a fixed L^∞ neighborhood of 1 in the class, and the scale may be random with a uniform deterministic upper bound.*

Proof. Theorems 1 and 2 apply to Φ and hence to $T_\Gamma = \operatorname{tr}\Gamma - \Phi$. At $f = 1$ the correctors vanish, so $G_i = e_i$ and $\psi_1 = 0$, and $\sigma(x_0, e) = e^\top \Gamma e$. Taking e a unit eigenvector of Γ with nonzero eigenvalue λ gives $\sigma = \lambda \neq 0$ at every x_0. □

In this instance the lower-bound template of Section 7 needs no moment conditions (since $\psi_1 = 0$ and $w_i = 0$), the symbol (11) is $k^\top \Gamma k/|k|^2$ exactly because G_i is constant, and the gap in (12) has the sign of λ for every sign vector.

Part IV
No-loss geometry and optimal tests

11 The no-loss null and the testing rate

For a fixed nonzero direction e let $D_e(f) = |e|^2 - e^\top K_f e$, which is nonnegative by testing (14) at the constant potential.

Theorem 7. *For $f \in \mathcal{F}_s$ the following are equivalent:*

$$D_e(f) = 0, \qquad \partial_e f = 0 \text{ in distributions}, \qquad \operatorname{Var}_f(\psi_{f,ee^\top}) = 0.$$

For testing $D_e(f) = 0$ against $D_e(f) \geq \delta$ over $\mathcal{F}_s$, the minimax separation order is $r_n = n^{-4s/(4s+d)}$ for every $s > 0$. Precisely, given $0 < \eta < 1/2$ there are A_η and tests at $\delta = A_\eta r_n$ whose worst type I and type II errors are each at most η for all large n, and for a sufficiently small $a_\eta > 0$ every test at $\delta = a_\eta r_n$ has sum of worst errors at least $1 - \eta$ for all large n.

Interpretation. By (16) below, D_e is comparable to the squared seminorm $Q_e(f) = \sum_{k \neq 0} (e \cdot k)^2 |k|^{-2} |\widehat{f}(k)|^2$, which weights each frequency by the squared cosine of its angle with e; it is a direction-weighted L^2 seminorm, and it equals $\|\partial_e f\|^2_{\dot{H}^{-1}} = \langle \partial_e f, (-\Delta)^{-1} \partial_e f \rangle$. The problem is therefore Ingster's minimax goodness-of-fit problem [9, 10], whose squared-L^2 separation over smoothness-s alternatives is $n^{-4s/(4s+d)}$, posed for this direction-weighted seminorm and for the composite null of densities invariant along e. The rate holds for every $s > 0$, also where estimation of D_e is root-n, because the null consists of zeros of the influence variance: there the first-order term vanishes and the problem is governed by the degenerate quadratic part, as in goodness-of-fit testing. The null contains nonconstant conductivities whenever there are nonzero integer frequencies orthogonal to e, and the rate is the same for every direction.

Proof of the equivalence. With $p = 1$ in (15) the exact deficit is $D_e(f) = \langle \partial_e f, A_f^{-1} \partial_e f \rangle$, which vanishes exactly when $\partial_e f = 0$. Then the corrector and the influence vanish. Conversely, zero influence variance and $f \geq m$ make $|e + \nabla \chi_{f,e}|^2$ equal to a constant; at a maximum of the C^1 periodic corrector its gradient vanishes, so the constant is $|e|^2$. Integrating against unweighted volume and using that the mean gradient vanishes gives $\int |\nabla \chi_{f,e}|^2 = 0$, so the corrector vanishes and the corrector equation gives $\partial_e f = 0$. □

Comparison. The order bounds $mA_1 \leq A_f \leq MA_1$ give

$$M^{-1} Q_e(f) \leq D_e(f) \leq m^{-1} Q_e(f). \tag{16}$$

Test. Let $Q_{e,J}$ be the truncation of Q_e to $0 < |k|_\infty \le J$; the Fourier tail bound gives $0 \le Q_e - Q_{e,J} \le CJ^{-2s}$ on the class. Estimate $Q_{e,J}$ by the off-diagonal U-statistic $\widehat{Q}_{e,J}$ with the real kernel

$$B_{e,J}(x,y) = \sum_{0<|k|_\infty \le J} \frac{(e\cdot k)^2}{|k|^2} \exp\big(2\pi i k\cdot(x-y)\big).$$

Its mean is $Q_{e,J}$, its squared kernel norm in unweighted volume is $O(J^d)$, and its first Hoeffding projection has squared $L^2(f)$ norm at most $CQ_{e,J}$, by Parseval, the bounded weights and $f \le M$. Thus

$$\mathrm{Var}_f(\widehat{Q}_{e,J}) \le C(Q_{e,J}/n + J^d/n^2). \tag{17}$$

Under the null $\widehat{f}(k) = 0$ unless $e\cdot k = 0$, where the weight vanishes, so $Q_{e,J} = 0$ and the kernel is canonical. With $J \asymp n^{2/(4s+d)}$, $J^{-2s} \asymp J^{d/2}/n \asymp r_n$. Reject when $\widehat{Q}_{e,J} > L_\eta r_n$; by (17) L_η can be chosen so that the type I error is at most η. Under $D_e(f) \ge A_\eta r_n$, (16) and the tail bound give $Q_{e,J} \ge (mA_\eta - C)r_n$, which exceeds twice the threshold for large A_η, and Chebyshev's inequality with (17) and $nr_n \to \infty$ bounds the type II error by η.

Converse. Apply the construction of Section 7 with $\Gamma = ee^\top$, $\sigma = |e|^2$. Every alternative has deficit at least $c\rho^2 h^{2s}$, and with $h = tn^{-2/(4s+d)}$ for small fixed t the mixture is within total variation η of the uniform product law. The uniform density belongs to the null, so every test at a small multiple of r_n has worst-error sum at least $1-\eta$.

Part V
Adaptive testing

12 Adaptation to the smoothness

Fix $0 < s_- < s_+ < \infty$ and let the smoothness $s \in [s_-, s_+]$ be free. The classes $\mathcal{F}_s$ share the positivity and Hölder bounds. For large n put $\ell_n = \log\log n$ and $r_n^{\mathrm{ad}}(s) = (\sqrt{\ell_n}/n)^{4s/(4s+d)}$.

Theorem 8. *There is a single test with worst type I error at most η over the union of the no-loss nulls for $s \in [s_-, s_+]$, and worst type II error at most η over $f \in \mathcal{F}_s$ with $D_e(f) \ge A_\eta r_n^{\mathrm{ad}}(s)$, simultaneously for all s in the interval. For a sufficiently small $a > 0$, the sum of the worst errors of any test over this null and the alternatives with $D_e(f) \ge a r_n^{\mathrm{ad}}(s)$ is at least $1 - o(1)$.*

Multi-resolution testing with the iterated-logarithm price is classical [21, 6]. Here the threshold must hold over the whole directional null, and the lower bound is in the physical loss, which gives small weight to frequencies nearly orthogonal to e.

Upper bound. Use $B_{e,J}$ on two independent halves, averaging over pairs across the halves; the mean and (17) are unchanged. Under every null the kernel is canonical. With $R = J^d$, its $L^2(f\otimes f)$ norm and maximal row norm in $L^2(f)$ are at most $C\sqrt{R}$, its supremum is at most CR, and its operator norm on $L^2(f)$ is bounded, because under the isometry $L^2(f) \to L^2$, $g \mapsto \sqrt{f}g$, it becomes a Fourier multiplier with weights in $[0,1]$ conjugated by multiplication by $\sqrt{f} \le \sqrt{M}$. These bounds hold with the same constants for every null density in $\mathcal{F}_s$, $s \in [s_-, s_+]$, since they use only $m \le f \le M$; this is what makes the type I bound uniform over the union of nulls. The exponential inequality of Giné, Latała and Zinn for decoupled canonical U-statistics of order two [8, Theorem 3.3 and Corollary 3.4], applied to two equal halves, has the four parameters of orders

$n\sqrt{R}$, n, $\sqrt{nR}$ and R, and gives tail at most Ce^{-t} at the threshold

$$C\Big(\frac{\sqrt{Rt}}{n}+\frac{t}{n}+\frac{\sqrt{R}\,t^{3/2}}{n^{3/2}}+\frac{Rt^2}{n^2}\Big).$$

Take dyadic J between n^{c_0} and n^{c_1} with $0<c_0<2/(4s_+{+}d)$ and $2/(4s_-{+}d)<c_1<2/d$, so $O(\log n)$ resolutions. With t a large constant times ℓ_n, the first term dominates the other three uniformly in this range, because $R\le n^{c_1 d}$ with $c_1 d<2$. A union bound gives type I control at thresholds of order $\sqrt{\ell_n R}/n$. For each s the grid contains J within a constant factor of $(n^2/\ell_n)^{1/(4s+d)}$, where tail and threshold are both $O(r_n^{\mathrm{ad}}(s))$ uniformly in s; the type II bound follows as for Theorem 7, since $nr_n^{\mathrm{ad}}(s)\to\infty$ uniformly.

Lower bound. Use the templates of Section 7 for $\Gamma=ee^\top$ with the same ρ and template at every resolution. Select $L_n\asymp \log n/\log\ell_n$ scales with $h_j^{4s_j+d}=\ell_n/n^2$, $s_j\in[s_-,s_+]$, spacing $\log(1/h_j)$ at increments $(4/d)\log\ell_n$, so successive scale ratios are at least $\ell_n^{4/d}$. Let Z_j be the likelihood ratio of the level-j sign mixture. Every level-j alternative has $D_e\ge c\rho^2 r_n^{\mathrm{ad}}(s_j)$, and (13) gives $\mathbb{E}_0 Z_j^2\le \exp(C\rho^4\ell_n)$. For $j\ne k$ let U_{jk} have entries $n\langle a_{j,b},a_{k,c}\rangle$; with within-level norms g_j,g_k and cell counts m_j,m_k, orthogonality within levels gives $\|U_{jk}\|_{\mathrm{op}}\le ng_jg_k$ and $\|U_{jk}\|_{\mathrm{HS}}^2\le n^2g_j^2g_k^2\min(m_j,m_k)\le C\rho^4\ell_n(\min(m_j,m_k)/\max(m_j,m_k))^{1/2}=O(\rho^4/\ell_n)$, while the operator norm tends to zero since $n^2g_j^4m_j\lesssim\rho^4\ell_n$ and the cell counts grow polynomially. Then $\mathbb{E}_0Z_jZ_k\le\mathbb{E}_{\epsilon,\epsilon'}\exp(\epsilon^\top U_{jk}\epsilon')\le\det(I-U_{jk}^\top U_{jk})^{-1/2}\le\exp(C\|U_{jk}\|_{\mathrm{HS}}^2)=1+o(1)$, by $\cosh x\le e^{x^2/2}$ in one sign vector, a Gaussian representation of the resulting exponential of a squared norm, and the same inequality in the other. The equal mixture over levels has chi-square distance at most $L_n^{-1}\exp(C\rho^4\ell_n)+o(1)=o(1)$ for small fixed ρ, and the uniform density is a common null.

Part VI
Exact calibration and positive semidefinite contrasts

13 Haar calibration

Let H_e be the closure of $\{te \bmod \mathbb{Z}^d : t\in\mathbb{R}\}$, a compact subgroup. A continuous density with $\partial_e f=0$ is H_e-invariant, by its Fourier coefficients. Conditional on the quotient locations in $\mathbb{T}^d/H_e$, null samples are independent and Haar-uniform on their cosets, by Haar disintegration, since an invariant density is constant on each coset. Hence the conditional null law of a statistic is obtained by adding an independent Haar element of H_e to each point. For the statistic of Theorem 7 or the maximum of the threshold-normalized statistics of Theorem 8, the randomized upper-tail probability gives an exact level-α test for every $0<\alpha<1$.

Both separation rates hold for this calibration. Under any alternative the shifted sample is independent with density $\bar f$, the Haar average of f, which lies in the null with the same density bounds. Choose a deterministic threshold b with unconditional null exceedance at most α_0 for every null. By Markov's inequality the conditional shifted exceedance at b is at least α with probability at most α_0/α, and whenever the observed statistic exceeds b and that probability is below α, the randomized tail at the observed value is below α. So the conditional test's type II error exceeds that of the deterministic test by at most α_0/α, and α_0 is arbitrary.

14 Positive semidefinite contrasts

All no-loss results hold for every fixed nonzero positive semidefinite $\Gamma = \sum_\ell \lambda_\ell v_\ell v_\ell^\top$. Put $D_\Gamma = \operatorname{tr}\Gamma - T_\Gamma$ and use the weights $k^\top \Gamma k/|k|^2$. Summing the directional inequalities with nonnegative weights gives (16), and the null is $\partial_{v_\ell} f = 0$ for all $\lambda_\ell > 0$. If the influence variance vanishes, the weighted sum $\sum_\ell \lambda_\ell |v_\ell + \nabla \chi_{f,v_\ell}|^2$ is a constant c_*. Its unweighted mean is $\operatorname{tr}\Gamma + \sum_\ell \lambda_\ell \int |\nabla \chi_{f,v_\ell}|^2 \geq \operatorname{tr}\Gamma$, because each correction has mean gradient zero, and its f-weighted mean is $T_\Gamma(f) \leq \operatorname{tr}\Gamma$. Both equal c_*, so every positive-weight corrector vanishes. The estimation and testing bounds use the bounded nonnegative symbol in place of the directional one, the mixture lower bound uses any positive eigenvalue, and exact calibration uses the closure of the image of range(Γ) in the torus.

Statements and Declarations

Funding. This research was self-funded.

Competing interests. The author owns Luminary Oddities and has no other relevant competing interests.

AI assistance. OpenAI Codex assisted with derivations, proof development, source checks and independent internal review of each part. Anthropic Claude edited, consolidated and cross-referenced the parts and checked computations. Scott Schubr directed the research and is responsible for the content.

References

[1] S. Agmon, A. Douglis and L. Nirenberg, *Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions I*, Comm. Pure Appl. Math. 12 (1959), 623–727. https://doi.org/10.1002/cpa.3160120405.

[2] A. Alexanderian, M. Rathinam and R. Rostamian, *Homogenization, symmetry, and periodization in diffusive random media*, Acta Math. Sci. Ser. B 32 (2012), 129–154. https://doi.org/10.1016/S0252-9602(12)60008-3.

[3] 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. https://www.bdim.eu/item?id=BUMI_2002_8_5B_2_487_0.

[4] P. J. Bickel and Y. Ritov, *Estimating integrated squared density derivatives: sharp best order of convergence estimates*, Sankhyā Ser. A 50 (1988), 381–393.

[5] L. Birgé and P. Massart, *Estimation of integral functionals of a density*, Ann. Statist. 23 (1995), 11–29. https://doi.org/10.1214/aos/1176324452.

[6] M. Fromont and B. Laurent, *Adaptive goodness-of-fit tests in a density model*, Ann. Statist. 34 (2006), 680–720. https://doi.org/10.1214/009053606000000119.

[7] D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Classics in Mathematics, Springer, Berlin, 2001 (reprint of the 1998 edition). https://doi.org/10.1007/978-3-642-61798-0.

[8] E. Giné, R. Latała and J. Zinn, *Exponential and moment inequalities for U-statistics*, in High Dimensional Probability II, Birkhäuser, Boston, 2000, 13–38; theorem numbers as in arXiv:math/0003228. https://doi.org/10.1007/978-1-4612-1358-1_2.

[9] Yu. I. Ingster, *Minimax testing of nonparametric hypotheses on a distribution density in the L_p metrics*, Theory Probab. Appl. 31 (1987), 333–337. https://doi.org/10.1137/1131042.

[10] Yu. I. Ingster and I. A. Suslina, *Nonparametric Goodness-of-Fit Testing Under Gaussian Models*, Lecture Notes in Statistics 169, Springer, 2003. https://doi.org/10.1007/978-0-387-21580-8.

[11] B. Laurent, *Efficient estimation of integral functionals of a density*, Ann. Statist. 24 (1996), 659–681. https://doi.org/10.1214/aos/1032894458.

[12] S. Neukamm, *An introduction to the qualitative and quantitative theory of homogenization*, Interdiscip. Inform. Sci. 24 (2018), 1–48. https://doi.org/10.4036/iis.2018.A.01.

[13] J. Robins, L. Li, E. Tchetgen and A. W. van der Vaart, *Higher order influence functions and minimax estimation of nonlinear functionals*, Probability and Statistics: Essays in Honor of David A. Freedman, IMS Collections 2 (2008), 335–421. https://doi.org/10.1214/193940307000000527.

[14] J. Robins, L. Li, E. Tchetgen and A. W. van der Vaart, *Quadratic semiparametric von Mises calculus*, Metrika 69 (2009), 227–247. https://doi.org/10.1007/s00184-008-0214-3.

[15] S. Schubr, *Fourier computation of empirical diffusion metrics*, Luminary Research Archive, edition 3 (2026), eq. (2.1). https://doi.org/10.5281/zenodo.24472464.

[16] S. Schubr, *Sharp inference rates at a power-law degenerate Darcy minimum*, Luminary Research Archive, edition 1 (2026). https://doi.org/10.5281/zenodo.24484658.

[17] S. Schubr, *Sharp linearized Darcy inference at a two-block saddle*, Luminary Research Archive, edition 1 (2026). https://doi.org/10.5281/zenodo.24484660.

[18] S. Schubr, *Sharp minimax rates for inverse diffusion metric inference*, Luminary Research Archive, edition 1 (2026); its content forms Parts I and II of this edition. https://doi.org/10.5281/zenodo.24484648.

[19] S. Schubr, *The nonlinear minimax exponent for Darcy inference at an anisotropic minimum*, Luminary Research Archive, edition 1 (2026). https://doi.org/10.5281/zenodo.24450833.

[20] S. Schubr, *The root-sample threshold for interior perturbations in the Darcy inverse problem*, Luminary Research Archive, edition 1 (2026). https://doi.org/10.5281/zenodo.24484656.

[21] V. G. Spokoiny, *Adaptive hypothesis testing using wavelets*, Ann. Statist. 24 (1996), 2477–2498. https://doi.org/10.1214/aos/1032181163.

[22] 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. https://doi.org/10.1016/j.spl.2008.07.001.