

Notes for Week 1/23/20 - 1/30/20

Alden Green

March 9, 2020

Let N be an integer, $n = N^d$, and

$$\bar{X} := \left\{ \frac{1}{N}(j_1 - 1/2, \dots, j_d - 1/2) : j \in [N]^d \right\}$$

consist of n total evenly spaced grid points on $[0, 1]^d$. The lattice graph formed over these grid points is

$$\bar{G} := (\bar{X}, \bar{E}), \quad \bar{E} := \left\{ (\bar{x}, \bar{x}') : \|\bar{x} - \bar{x}'\|_1 = \frac{1}{t} \right\}.$$

We observe n samples according to regression model

$$y_j = f(\bar{x}_j) + \varepsilon_j, \quad \varepsilon_j \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1) \quad \text{for each } j \in [n], \quad (1)$$

and our goal is to perform the goodness-of-fit test

$$\mathbb{H}_0 : f = f_0 := 0 \quad \text{vs.} \quad \mathbb{H}_a : f \neq f_0.$$

For a given test $\phi : \mathbb{R}^n \rightarrow \{0, 1\}$ and function class \mathcal{F} , we evaluate ϕ through the lens of worst-case risk:

$$\mathcal{R}(\phi; \mathcal{F}) := \mathbb{E}_{f_0}(\phi) + \sup_{f \neq f_0 \in \mathcal{F}} \mathbb{E}_f(1 - \phi)$$

Unless there exists some separation between functions $f \in \mathcal{F}$ and f_0 , the type II error can be quite close to 1. We enforce this separation by removing a ball of radius ϵ centered at f_0 from the function class \mathcal{F} , and let

$$\mathcal{F}_\epsilon := \{f \in \mathcal{F} : \|f\|_{\mathcal{L}^2} \geq \epsilon\}.$$

A classical way to study the properties of ϕ over \mathcal{F} is through the critical radius: the smallest value of ϵ such that ϕ has small worst-case risk over the remaining function class $\mathcal{R}(\phi; \mathcal{F}_\epsilon)$, which we shall write as $\mathcal{R}_\epsilon(\phi; \mathcal{F})$.

1 Graph Laplacian Eigenvector Projection Test

For an arbitrary graph G on n nodes with graph Laplacian $L_G = D_G - A_G$, let $\Lambda(G)$ consist of the eigenvalues $0 = \lambda_1(G) \leq \lambda_2(G) \leq \dots \leq \lambda_n(G)$ of L_G , with $v_k(G) = (v_{k,1}(G), \dots, v_{k,n}(G)) \in \mathbb{R}^n$ denoting the eigenvector corresponding to the k th eigenvalues $\lambda_k(G)$. The graph Laplacian eigenvector projection test is a linear projection test; letting κ be some integer between 1 and n , we define $\Pi_{\kappa, G}(\theta) := \frac{1}{\sqrt{\kappa}} \sum_{k=1}^{\kappa} \langle v_k, \theta \rangle v_k$ to be a suitable rescaling of the projection of θ onto the span of $v_1(G), \dots, v_\kappa(G)$.

Our test statistic

$$T_{\text{spec}}(G) := \|\Pi_{\kappa, G}(Y)\|_n^2 - \frac{\kappa}{n}$$

is a simple function of the projection of the data Y onto the span of the first κ eigenvectors of the graph Laplacian $L(G)$ – it is a standard calculation to show that

$$\mathbb{E}_{Y|X} \left[\|\Pi_{\kappa,G}(Y)\|_n^2 \right] = \|\Pi_{\kappa,G}(f)\|_n^2 + \frac{\kappa}{n}$$

and hence our test statistic is an unbiased estimate of the norm $\|\Pi_{\kappa,G}(f)\|_n^2$. The graph Laplacian eigenvector test $\phi_{\text{spec}}(G) := \mathbf{1}\{T_{\text{spec}}(G) \geq \tau(b)\}$ rejects the null hypothesis when the test statistic $T_{\text{spec}}(G)$ is greater than a pre-specified cutoff $\tau(b)$, where the number b —or more precisely, $1/b$ —encodes the level of error the user is willing to tolerate.

The number of eigenvectors κ and cutoff $\tau(b)$ are user-specified hyperparameters. When they are appropriately chosen, the test $\phi_{\text{spec}}(\bar{G})$ is minimax optimal over the Holder ball $C^1(\mathcal{X}; R)$ when $d \leq 4$.

Theorem 1. *Let $L > 0$ be a fixed constant, and let $d \leq 4$. Suppose we observe data according to the model (1). Then when the test $\phi_{\text{spec}}(\bar{G})$ is performed with the parameter choices*

$$\kappa = n^{2d/(4+d)}, \quad \tau(b) = b\sqrt{\frac{2\kappa}{n^2}}$$

for some $b \geq 1$, the following statement is true: there exists a constant c such that for

$$\epsilon^2 \geq cL^2b^2n^{-4/(4+d)} \quad (2)$$

the worst-case risk of $\phi_{\text{spec}}(\bar{G})$ on $C^1(\mathcal{X}; L)$ is upper bounded

$$\mathcal{R}_\epsilon \left(\phi_{\text{spec}}(\bar{G}); C^1(\mathcal{X}; L) \right) \leq \frac{1}{b}. \quad (3)$$

The minimax rate for this problem is established by (Abramovich, De Feis, Sapatinas), who consider the more general question of testing over Besov spaces. The test they use to establish the upper bound is related to our own, however, they use a continuum wavelet basis orthogonalized in $L^2([0, 1]^d)$ rather than the grid Laplacian eigenvectors we consider.

As was the case with the neighborhood graph in the random design setting, the grid Sobolev norm fails to track the higher order derivatives of f near the boundary of \mathcal{X} . We therefore must insist that f be compactly supported in order for the grid eigenvector projection test to achieve minimax optimal testing rates over higher-order Holder classes.

Theorem 2. *Let $L > 0$ be fixed constants, $s \geq 2$ be a fixed integer, and let $d \leq 4s$. Suppose we observe data according to the model (1). Then when the test $\phi_{\text{spec}}(\bar{G})$ is performed with the parameter choices*

$$\kappa = n^{2d/(4s+d)}, \quad \tau(b) = b\sqrt{\frac{2\kappa}{n^2}}$$

for some $b \geq 1$, the following statement is true: there exists a constant c such that for any

$$\epsilon^2 \geq cL^2b^2n^{-4s/(4s+d)},$$

the worst-case risk of $\phi_{\text{spec}}(\bar{G})$ is upper bounded

$$\mathcal{R}_\epsilon \left(\phi_{\text{spec}}(\bar{G}); C_c^s(\mathcal{X}; L) \right) \leq \frac{1}{b}.$$

In order to prove Theorems 1 and 2, we decompose

$$T_{\text{spec}}(\bar{G}) - \|f\|_{\mathcal{L}^2}^2 = \left\{ \|\Pi_{\kappa,\bar{G}}Y\|_n^2 - \frac{\kappa}{n} - \|\Pi_{\kappa,\bar{G}}f\|_n^2 \right\} + \left\{ \|\Pi_{\kappa,\bar{G}}f\|_n^2 - \|f\|_n^2 \right\} + \left\{ \|f\|_n^2 - \|f\|_{\mathcal{L}^2}^2 \right\},$$

where we refer to the three terms on the right hand side as the estimation error, approximation error, and the discretization error, respectively.

Our technical work consists of upper bounding these three sources of error. When $4s \geq d$, the discretization error is negligible regardless of our choice of κ , and we focus our attention on minimizing the sum of approximation and estimation errors. This is the standard tradeoff in nonparametric testing problems, and the usual choice of projecting our data $Y \in \mathbb{R}^n$ onto a subspace of dimension $\kappa = n^{2d/(4s+d)}$ balances the contribution of these two terms.

1.1 Low-smoothness regime

In the low-smoothness regime $d > 4s$, the coupling between empirical norm over \bar{X} and the continuum \mathcal{L}^2 norm is sufficiently weak that the discretization error can become the dominant source of error, as we show in Lemma 1.

Lemma 1. *For any $L > 0$ and any integers s and d , there exists a function $f \in C^s(\mathcal{X}; L)$ such that*

$$\|f\|_{\mathcal{L}^2}^2 \geq \frac{1}{2^d} L^2 n^{-2s/d}$$

but $f(\bar{x}) = 0$ for all $\bar{x} \in \bar{X}$.

Lemma 1 implies the critical radius over the Holder classes $C^s(\mathcal{X}; L)$ must be at least on the order of $\epsilon^*(C^s(\mathcal{X}; L)) \gtrsim n^{-s/d}$. When $4s \geq d$, this term is negligible relative to $n^{-2s/(4s+d)}$, and we arrive at the “usual” rates of minimax testing we see in Theorems 1 and 2. On the other hand when $4s < d$, this becomes the dominant term, and we can no longer achieve the typical rate $\epsilon^2 \asymp n^{-4s/(4s+d)}$.

In this setting, when we enforce the wider radius $\epsilon^2 \asymp n^{-2s/d}$, Proposition 1 shows that a very simply test has small worst-case risk \mathcal{R}_ϵ .

Proposition 1. *Let $L > 0$ be a fixed constant, $s \geq 1$ a fixed integer, and let $d \geq 4s$. Suppose we observe data according to the grid design regression model (1). Then when the test $\phi_{\text{spec}}(\bar{G})$ is performed with the parameter choices*

$$\kappa = n, \quad \tau(b) = b\sqrt{\frac{2}{n}}$$

for some $b \geq 1$, the following statement is true: there exists a constant c such that for any

$$\epsilon^2 \geq cb^2 L^2 n^{-2s/d}$$

the worst case risk of $\phi_{\text{spec}}(\bar{G})$ is upper bounded

$$\mathcal{R}_\epsilon(\phi_{\text{spec}}(\bar{G}); C^s(\mathcal{X}; L)) \leq \frac{1}{b}.$$

- The test statistic $T_{\text{spec}}(\bar{G})$ is simply the empirical norm of Y .
- Together, Lemma 1 and Proposition (1) characterize the minimax rate of the grid design regression testing problem when $4s > d$.
- This minimax rate does not match the upper bound established in Proposition (Holder testing random design)—the fixed design problem is harder.
- Note that the usual minimax rate of estimation error over Holder classes is always $n^{-2s/(2s+d)}$, regardless of the relation between s and d , and regardless of whether loss is measured in \mathcal{L}^2 -norm or empirical norm. The explanation for this is that the rate $n^{-2s/(2s+d)}$ is always larger than $n^{-2s/d}$, and so the discretization error is never the dominant source of error in estimation. This reveals an interesting distinction between the testing and estimation problems, as testing is easy enough in a statistical sense that error incurred by the discrete nature of the problem may be the bottleneck.

1.2 Continuum Sobolev class.

Now suppose we assume $f \in H^s(\mathcal{X}; L)$ belongs to a relevant Sobolev ball, rather than a Holder ball, and we again ask about the worst case risk $\mathcal{R}_\epsilon(\phi; H^s(\mathcal{X}; L))$.

When $2s < d$, the situation is hopeless. The Sobolev class $H^s(\mathcal{X}; L)$ includes bump functions of arbitrarily small radius and arbitrarily large height—in the limit, delta functions. By placing these bump functions at design points \bar{x} , we can construct a function f which satisfies $f \equiv L$ almost everywhere on \mathcal{X} , while $f(\bar{x}) = 0$ for all $\bar{x} \in \bar{X}$. No test based on data observed according to model (1) can have more than trivial power against f , implying that for all tests ϕ the worst case risk $\mathcal{R}_\epsilon(\phi; H^s(\mathcal{X}; L))$ is 1 unless $\epsilon = \Omega(1)$.

When $2s > d$, we are not sure how the situation changes. The Sobolev Embedding Theorem gives us that the critical radius

$$\epsilon^*(H^s(\mathcal{X}; L)) \leq c\epsilon^*(C^{s'}(\mathcal{X}; L)), \quad \text{for } s' = s - d/2.$$

where we recall that the critical radius over Holder functions is

$$\epsilon^*(C^s(\mathcal{X}; L)) \asymp \max\{n^{-2s/(4s+d)}, n^{-s/d}\}$$

We might hope to do better, but it is hard to say.

2 Proofs

2.1 Proof of Theorem 1

In Lemma 2, we upper bound the Type I and Type II error of the test $\phi_{\text{spec}}(G)$ over an arbitrary graph G on n vertices, where we observe data sampled according to the model

$$y_i = \beta_i + \varepsilon_i, \quad \varepsilon_i \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1) \tag{4}$$

Our bound on the Type II error will be stated as a function of $\beta^T L^s \beta$ —a measure of the smoothness the signal β displays over the graph G —as well as the κ th eigenvalue of the Laplacian $\lambda_\kappa(G)$.

Lemma 2. *Let $b \geq 1$ be a fixed constant. Suppose we observe data Y according to the model (4), and perform the test $\phi_{\text{spec}}(G)$ for some $1 \leq \kappa \leq n$ and*

$$\tau = b\sqrt{\frac{2\kappa}{n^2}}.$$

Then the following statements are true:

1. **Type I error:** *Under the null hypothesis $\beta = \beta_0 = 0$, the Type I error of $\phi_{\text{spec}}(G)$ is upper bounded*

$$\mathbb{E}_{\beta_0}(\phi_{\text{spec}}(G)) \leq \frac{1}{b^2}. \tag{5}$$

2. **Type II error:** *For any $\beta \in \mathbb{R}^n$ such that*

$$\frac{1}{n} \sum_{i=1}^n \beta_i^2 \geq 2b\sqrt{\frac{2\kappa}{n^2}} + \frac{\beta^T (L_G)^s \beta}{n[\lambda_\kappa(G)]^s} \tag{6}$$

the Type II error of ϕ_{spec} is upper bounded,

$$\mathbb{E}_\beta(1 - \phi_{\text{spec}}(G)) \leq \frac{3}{b}. \tag{7}$$

The proof of Lemma 2 can be seen in the *graph_testing* document. The eigenvalues of the lattice \bar{G} are known exactly, and standard manipulations [Sadhanala](#) lead to the lower bound

$$\lambda_k(\bar{G}) \geq cn^{-2/d}k^{2/d} \text{ for } k = 1, \dots, n. \quad (8)$$

In light of (8) and Lemma 2, in order to prove Theorem 1 it will be sufficient to show the following two inequalities:

(E1) For every $f \in C^1(\mathcal{X}; R)$,

$$f^T(L_{\bar{G}})f \leq 4R^2n^{1-2/d} \quad (9)$$

(E2) When $d \leq 4$, for every $f \in C^1(\mathcal{X}; R)$ satisfying $\|f\|_{\mathcal{L}^2(\mathcal{X})} \geq R^2n^{-2/(4+d)}$, additionally

$$\frac{1}{n} \sum_{i=1}^n f(x_i)^2 \geq \frac{1}{4} \|f\|_{\mathcal{L}^2(\mathcal{X})}^2 \quad (10)$$

Once we have shown these inequalities, some basic algebra yields that for the specific choice $\kappa = (Rn)^{2d/(4s+d)}$ and a sufficiently large choice of c in (2),

$$2b\sqrt{\frac{2\kappa}{n^2}} + \frac{f^T(L_{\bar{G}})^s f}{n[\lambda_\kappa(\bar{G})]^s} \leq \frac{1}{n} \sum_{i=1}^n [f(x_i)]^2$$

and (3) follows from Lemma 2.

Proof of (9): The upper bound (9) follows straightforwardly from the Holder property,

$$\begin{aligned} f^T(L_{\bar{G}})f &= \sum_{i \in [t]^d} \sum_{j \in [t]^d} (f(\bar{x}_i) - f(\bar{x}_j))^2 \mathbf{1}\{\|i - j\|_1 = t\} \\ &\leq R^2 \sum_{i \in [t]^d} \sum_{j \in [t]^d} \|\bar{x}_i - \bar{x}_j\|_2^2 \cdot \mathbf{1}\{\|i - j\|_1 = t\} \\ &= R^2 n^{-2/d} \sum_{i \in [t]^d} \sum_{j \in [t]^d} \mathbf{1}\{\|i - j\|_1 = t\} \\ &\leq 4R^2 n^{1-2/d}. \end{aligned}$$

Proof of (10): For $x \in \mathbb{R}^d$, let $Q(x)$ be the d -dimensional cube of side length $1/t$ with x at its corner,

$$Q(x) = [x_1 - 1/t, x_1] \otimes \dots \otimes [x_d - 1/t, x_d]$$

The difference between $f(\bar{x})$ and the average of f over the cube $Q(\bar{x})$ can be bounded using the Holder property,

$$\int_{Q(\bar{x})} (f(x) - f(\bar{x}))^2 dx \leq \frac{1}{n} R^2 n^{-2/d}.$$

Dividing \mathcal{X} into cubes and summing over the cubes, we have

$$\begin{aligned}
\|f\|_{\mathcal{L}^2(\mathcal{X})}^2 &= \sum_{j \in [t]^d} \int_{Q(\bar{x}_j)} [f(x)]^2 dx \\
&\leq 2 \sum_{j \in [t]^d} \frac{1}{n} \left\{ [f(\bar{x}_j)]^2 + \int_{Q(\bar{x}_j)} [f(x) - f(\bar{x}_j)]^2 dx \right\} \\
&\leq \frac{2}{n} \sum_{j \in [t]^d} [f(\bar{x}_j)]^2 + R^2 n^{-2/d} \\
&\leq \frac{2}{n} \sum_{j \in [t]^d} [f(\bar{x}_j)]^2 + \frac{1}{2} \|f\|_{\mathcal{L}^2(\mathcal{X})}^2
\end{aligned}$$

where the last inequality follows by the assumption $\|f\|_{\mathcal{L}^2(\mathcal{X})}^2 \geq R^2 n^{-4/(4+d)}$ along with the fact that $4/(4+d) \leq 2/d$ when $d \leq 4$.

2.2 Proof of Theorem 2

Referring again to Lemma 2, to prove Theorem 2 it will be sufficient to prove the following inequalities

(E1) For any $f \in C_c^s(\mathcal{X}; L)$,

$$f^T \bar{L}^s f \leq c L^2 n^{1-2s/d}. \quad (11)$$

(E2) There exists a sufficiently small constant c_1 and a sufficiently large constant c_2 such that for any $f \in C^s(\mathcal{X}; L)$,

$$c_1 \|f\|_n^2 \geq \|f\|_{\mathcal{L}^2(\mathcal{X})}^2 - c_2 L n^{-2s/d} \quad (12)$$

for all n sufficiently large.

The proof of (12) is the content of Lemma 4. Here, we focus our attention on proving (11).

To obtain this upper bound, first observe that by letting $D_{e_i^2} f := f(\bar{x} + h e_i) + f(\bar{x} - h e_i) - 2f(\bar{x})$ and taking a 2nd order Taylor expansion, for all points $\bar{x} \in \mathcal{X}_h \cap \bar{X} =: \bar{X}_h$ we can express the grid Laplacian as

$$\begin{aligned}
\bar{L} f(\bar{x}) &= \sum_{i=1}^d D_{e_i^2} f(\bar{x}) \\
&= \sum_{i=1}^d \left[\frac{\partial}{\partial e_i} f(\bar{x})(e_i h - e_i h) + \frac{h^2}{2} \int_{-h}^h \frac{\partial^2}{\partial e_i^2} f(\bar{x} + e_i t) dt \right] \\
&= \frac{h^2}{2} \int_{-h}^h \sum_{i=1}^d \frac{\partial^2}{\partial e_i^2} f(\bar{x} + e_i t) dt
\end{aligned} \quad (13)$$

as a second-order differential operator which roughly approximates the continuum Laplace operator $\Delta = -\sum_{i=1}^d \frac{\partial^2}{\partial e_i^2}$. Motivated by this, in Lemma 3 we provide estimates on the absolute difference between the action of the iterated grid Laplacian and the iterated continuum Laplace operator. Specifically, we show that at all sufficiently interior points $\bar{x} \in \bar{X} \cap \mathcal{X}_{qh}$

$$\left\{ \begin{aligned} &|\bar{L}^{(s-2)/2} f(\bar{x}) - h^{(s-2)/2} \Delta^{(s-2)/2} f(\bar{x})|, & \text{when } s \text{ is even} \\ &|\bar{L}^{(s-1)/2} f(\bar{x}) - h^{(s-1)/2} \Delta^{(s-1)/2} f(\bar{x})|, & \text{when } s \text{ is odd} \end{aligned} \right\} \leq c L h^s$$

For those remaining points $\bar{x} \in \bar{X} \cap \mathcal{X}_{qh}$ sufficiently near the boundary of \mathcal{X} , since f is compactly supported in \mathcal{X} there exists a point x_0 and some $\delta > 0$ such that $\text{dist}(\bar{x}, x_0) < qh$, and $f(z) = 0$ for all $z \in B(x_0, r)$.

Therefore the partial derivatives $D^q f(x_0) = 0$ for each $q = 0, \dots, s-1$, and by taking a Taylor expansion of $f(x)$ around $f(x_0)$ we obtain

$$|f(x)| \leq cLh^s$$

where we have additionally used the Holder property of f . A crude bound of the form

$$|\bar{L}^s f(\bar{x})| \leq 2^s \cdot \max_{\bar{z} \in Q_{2qh}(\bar{x})} \{|f(\bar{z})|\} \leq cLh^s$$

will suffice for our purposes. When s is even, the graph Sobolev seminorm is then

$$\begin{aligned} f^T \bar{L}^s f &\leq \sum_{\bar{x} \in \bar{X}_{qh}} \left\{ \sum_{i=1}^d D_{e_i} \bar{L}^{(s-2)/2} f(\bar{x}) \right\}^2 + cL^2 h^{2s} n^{1-1/d} \\ &\leq h^{2s-4} \sum_{\bar{x} \in \bar{X}_{qh}} \left\{ \sum_{i=1}^d D_{e_i} \Delta^{(s-2)/2} f(\bar{x}) \right\}^2 + cL^2 h^{2s} n^{1-1/d} \\ &\leq \frac{h^{2s}}{4} \sum_{\bar{x} \in \bar{X}_{qh}} \left\{ \sum_{i=1}^d \int_{-h}^h \left(\frac{\partial^2}{\partial e_i^2} \Delta^{(s-2)/2} f \right)(\bar{x} + te_i) dt \right\}^2 + cL^2 h^{2s} n^{1-1/d} \\ &\leq d^2 h^{2s} L^2 n + cL^2 h^{2s} n^{1-1/d} \end{aligned}$$

where we have used, in order, our bounds on the absolute value of $\bar{L}^s f(\bar{x})$ and $|f|(\bar{x})$ at points \bar{x} near the boundary, our estimate on the approximation error between the grid Laplacian and continuum Laplace operator, the representation (13), and finally the Holder property

$$\|\Delta^{(s-2)/2} f\|_{C^2(\mathcal{X})} \leq \|f\|_{C^s(\mathcal{X})} \leq L.$$

Similar manipulations supply an equivalent bound when s is odd

$$\begin{aligned} f^T \bar{L}^s f &\leq \sum_{\bar{x} \in \bar{X}_{qh}} \left\{ \sum_{i=1}^d \bar{L}^{(s-1)/2} f(\bar{x} + he_i) - \bar{L}^{(s-1)/2} f(\bar{x}) \right\}^2 + cL^2 h^{2s} n^{1-1/d} \\ &\leq h^{2s-2} \sum_{\bar{x} \in \bar{X}_{qh}} \left\{ \sum_{i=1}^d \bar{\Delta}^{(s-1)/2} f(\bar{x} + he_i) - \bar{\Delta}^{(s-1)/2} f(\bar{x}) \right\}^2 + cL^2 h^{2s} n^{1-1/d} \\ &\leq d^2 h^{2s} L^2 n + cL^2 h^{2s} n^{1-1/d} \end{aligned}$$

where this time we have invoked the Holder property $\|\Delta^{(s-1)/2} f\|_{C^1(\mathcal{X})} \leq \|f\|_{C^s(\mathcal{X})} \leq L$. This establishes (11) and concludes the proof of Theorem 2.

2.3 Proof of Lemma 1

TODO:

2.4 Proof of Proposition 1

TODO:

3 Proofs of Technical Results

3.1 Grid Laplacian Approximation Error

The following result follows by standard Taylor expansion arguments.

Lemma 3. Let $f \in C^s(\mathcal{X}; L)$ for $s \geq 3$. Then, for all grid points $\bar{x} \in \bar{X}_{qh}$, and for either $q = (s-1)/2$ or $q = (s-2)/2$,

$$|\bar{L}^q f(\bar{x}) - h^{2q} \Delta^q f(\bar{x})| \leq cLh^s.$$

Proof. We prove by induction on q . When $q = 1$ and $\bar{x} \in \bar{X}_h$, by taking Taylor expansions of $f(\bar{x} + e_i h)$ and $f(\bar{x} - e_i h)$ around $f(\bar{x})$ for $i = 1, \dots, d$, we obtain

$$\begin{aligned} \bar{L}f(\bar{x}) &= \sum_{i=1}^d D_{e_i^2} f(\bar{x}) \\ &= \sum_{i=1}^d \left[\sum_{\alpha=1}^{s-1} \frac{1}{\alpha!} \frac{\partial^\alpha}{\partial e_i^\alpha} f(\bar{x}) (h^\alpha + (-h)^\alpha) + R(\bar{x}) \right] \\ &= \sum_{i=1}^d \left[\sum_{\beta=1}^{\lfloor (s-1)/2 \rfloor} \frac{2}{(2\beta)!} \frac{\partial^{2\beta}}{\partial e_i^{2\beta}} f(\bar{x}) h^{2\beta} + R(\bar{x}) \right] \end{aligned} \quad (14)$$

where the remainder term is upper bounded

$$|R(\bar{x})| = \left| \frac{1}{s!} \sum_{i=1}^d \int_{-h}^h \frac{\partial^s}{\partial e_i^s} f(\bar{x} + e_i t) t^{s-1} dt \right| \leq \frac{d}{(s+1)!} Lh^s$$

When $s = 3$ or $s = 4$, the first sum includes only second-order derivatives,

$$\sum_{i=1}^d \sum_{\beta=1}^{\lfloor (s-1)/2 \rfloor} \frac{2}{(2\beta)!} \frac{\partial^{2\beta}}{\partial e_i^{2\beta}} f(\bar{x}) h^{2\beta} = h^2 \sum_{i=1}^d \frac{\partial^{2\beta}}{\partial e_i^{2\beta}} f(\bar{x}) = h^2 \Delta f(\bar{x}),$$

and the claim is shown.

When $q \geq 2$, we use the prior analysis to rewrite the q th-order iterated Laplacian

$$\begin{aligned} \bar{L}^q f(\bar{x}) &= \bar{L}^{q-1} \left(\sum_{i=1}^d \sum_{\beta=1}^{\lfloor (s-1)/2 \rfloor} \frac{2}{(2\beta)!} \frac{\partial^{2\beta}}{\partial e_i^{2\beta}} f h^{2\beta} + R \right) (\bar{x}) \\ &= \sum_{\beta=1}^{\lfloor (s-1)/2 \rfloor} \frac{2h^{2\beta}}{(2\beta)!} \bar{L}^{q-1} \left(\sum_{i=1}^d \frac{\partial^{2\beta}}{\partial e_i^{2\beta}} f \right) (\bar{x}) + \bar{L}^{q-1} R(\bar{x}) \end{aligned} \quad (15)$$

We deal with the remainder term first. Noting that $Q_{2(q-1)h}(\bar{x}) \subset \mathcal{X}_h$, we have

$$\begin{aligned} |\bar{L}^{q-1} R(\bar{x})| &\leq (2(q-1)d)^{q-1} \cdot \max_{\bar{z} \in Q_{2(q-1)h}(\bar{x})} \{|R(\bar{z})|\} \\ &\leq \frac{(2(q-1)d)^{q-1} d}{(s+1)!} Lh^s \end{aligned} \quad (16)$$

Next we deal with the higher-order terms in the first sum, meaning the terms where $\beta \geq 2$. For convenience, denote $f_\beta := \sum_{i=1}^d \frac{\partial^{2\beta}}{\partial e_i^{2\beta}} f$, and note that $f_\beta \in C^{s-2\beta}(\mathcal{X}; L)$.

We will show that for any $\bar{x} \in \bar{X}_{qh}$,

$$|\bar{L}^{q-1} f_\beta(\bar{x})| \leq cLh^{2-\beta}. \quad (17)$$

Let us assume $s - 2\beta$ is odd. Set $s' = (s - 2\beta - 1)/2$, and $s'' = q - 1 - (s' + 1)$, and note that s'' is a non-negative integer. By hypothesis, for any $\bar{z} \in \bar{X}_{hs'}$

$$\left| \bar{L}^{s'} f_\beta(\bar{z}) - h^{2s'} \Delta^{s'} f_\beta(\bar{z}) \right| \leq cLh^{s-2\beta}$$

and consequently,

$$\begin{aligned}
|\bar{L}^{q-1} f_\beta(\bar{x})| &\leq (2s''d)^{s''} \cdot \max_{\bar{z} \in Q_{2s''h}(\bar{x})} \left\{ |\bar{L}^{s'+1} f(\bar{z})| \right\} \\
&\leq (2s''d)^{s''} \cdot \max_{\bar{z} \in \bar{X}_{(s'+1)h}} \left\{ |\bar{L}^{s'+1} f(\bar{z})| \right\} \\
&\leq (2s''d)^{s''} \left(h^{2s} \cdot \max_{\bar{z} \in \bar{X}_{(s'+1)h}} \left\{ |\bar{L} \Delta^{s'} f_\beta(\bar{z})| \right\} + cLh^{s-2\beta} \right) \\
&\leq (2s''d)^{s''} \left(2dLh^{2s'+1} + cLh^{s-2\beta} \right)
\end{aligned}$$

where the last line follows since $\Delta^{s'} f_\beta \in C^1(\bar{X}; L)$. Since $2s' + 1 = s - 2\beta$, we have shown the desired result (17) when $s - 2\beta$ is odd, and similar manipulations prove the result when $s - 2\beta$ is even.

Finally, we consider the second order term, where $\beta = 2$ and $f_\beta = \Delta f \in C^{s-2}(\mathcal{X}; L)$. Then, since either $q - 1 = (s - 3)/2$ or $q - 1 = (s - 4)/2$, by hypothesis for any $\bar{x} \in \bar{X}_{qh} \subset \bar{X}_{(q-1)h}$,

$$h^2 \left| \bar{L}^{q-1} \Delta f(\bar{x}) - h^{2(q-1)} \Delta^q f(\bar{x}) \right| \leq cLh^2 h^{s-2} = cLh^s \quad (18)$$

The claimed result follows upon plugging (16), (17), and (18) back into (15). \square

3.2 Discretization Error

We prove the following lower bound on the empirical norm of f . The proof will be similar in spirit and at times directly follow the proof of Lemma 3 of (Arias-Castro), the key difference being that we wish to lower bound the empirical norm of f rather than a Riemann sum.

Lemma 4. *Let $f \in C^s(\mathcal{X}, L)$. Then, there exist a constant c such that*

$$\|f\|_n^2 \geq c\|f\|_{\mathcal{L}^2([0,1]^d)}^2 - cLn^{-2s/d}$$

for all n sufficiently large.

Proof. In the proof of this Lemma, we take η to be the tensorization of a univariate $(s - 1)$ -order kernel ν . More specifically, we let $\nu : \mathbb{R} \rightarrow \mathbb{R}$ satisfy

$$\int_{\mathbb{R}} \nu(x) dx = 1, \quad \int_{\mathbb{R}} x^\ell \nu(x) dx = 0 \quad \text{for } \ell = 1, \dots, s - 1, \quad |\nu(x)| \leq 1 \text{ for all } x \in [0, 1^d]$$

additionally be compactly supported on $[-1/2, 1/2]$ but otherwise arbitrary, and define $\eta : \mathbb{R}^d \rightarrow \mathbb{R}$ to be $\eta(z) = \prod_{i=1}^d \nu(x_i)$ for $z \in \mathbb{R}^d$. We denote $\eta_h(x) = \frac{1}{h^d} \eta(x/h)$, and note that for any polynomial u of degree at most $s - 1$, $(u * \eta_h)(x) = u(x)$.

Fix an odd integer $m \geq 1$ to be chosen large enough, but fixed, later on, and assume without loss of generality that $N/m =: M \in \mathbb{N}$. For $i \in [n]^d$, let u_i be the $(s - 1)$ -order Taylor expansion of f around \bar{x}_i , and further let

$$u(x) = \sum_{i \in [M]^d} u_{mi-1}(x) \cdot \mathbf{1}\{x \in Q_{hm}(\bar{x}_{mi-1})\}, \quad \tilde{u}(x) = \sum_{i \in [M]^d} \sum_{\bar{x} \in Q_{hm}(\bar{x}_{mi-1})} (u_{mi-1} * \eta_h)(\bar{x}) \cdot \mathbf{1}\{x \in Q_h(\bar{x})\}.$$

be piecewise polynomial and piecewise constant functions, respectively.

Standard properties of Taylor expansions imply that $|f(x) - u_{mi-1}(\bar{x})| \leq (L\sqrt{dm}h)^s$ for all $x \in Q_{mh}(\bar{x}_{mi-1})$, and as a result

$$\|u - f\|_{\mathcal{L}^2([0,1]^d)}^2, \quad \|u - f\|_n^2 \leq ch^{2s}$$

Since u_{mi-1} is an $(s-1)$ -order polynomial, we have that $(u_{mi-1} * \eta_h)(\bar{x}) = u_{mi-1}(\bar{x})$ for each $\bar{x} \in \bar{X}$; therefore $\|\tilde{u}\|_n = \|u\|_n$. Additionally \tilde{u} is piecewise constant over the grid cells $Q_h(\bar{x})$, and so we have

$$\|\tilde{u}\|_{\mathcal{L}^2([0,1]^d)}^2 = \sum_{\bar{x} \in \bar{X}} h^d \cdot [\tilde{u}(\bar{x})]^2 = \|\tilde{u}\|_n^2.$$

It remains only to show that

$$c\|u\|_{\mathcal{L}^2([0,1]^d)} \leq \|\tilde{u}\|_{\mathcal{L}^2([0,1]^d)},$$

for a constant c which does not depend on \tilde{u} . In Lemma 5, we show that the statement holds true if u is a polynomial on $[0,1]^d$; and the width of the partition used to form \tilde{u} is taken to be a sufficiently small constant. By translating and rescaling the cubes $Q_{mi-1}(\bar{x}_{mi-1})$ into $[0,1]^d$ as in (Arias-Castro), we have that the same result holds for the piecewise polynomial u when m is taken to be a sufficiently large constant. Putting the pieces together, we have that whenever n is great enough such that $M \in \mathbb{N}$,

$$\begin{aligned} \|f\|_{\mathcal{L}^2([0,1]^d)}^2 &\leq cn^{-2s/d} + 2\|u\|_{\mathcal{L}^2([0,1]^d)}^2 \\ &\leq cn^{-2s/d} + c\|\tilde{u}\|_{\mathcal{L}^2([0,1]^d)}^2 \\ &= cn^{-2s/d} + c\|u\|_n^2 \\ &\leq cn^{-2s/d} + c\|f\|_n^2. \end{aligned}$$

which completes the proof of Lemma 4. \square

To simplify the statement of the following Lemma, for $j \in \mathbb{N}$ let

$$\mathcal{K}_j(u) := \sum_{[j]^d} (u * \eta_h)(\bar{x}) \cdot \mathbf{1}\{x \in Q_{1/j}(\bar{x})\}.$$

where η_h is taken as in the proof of Lemma 4. Denote the class of polynomials on \mathbb{R}^d of degree at most m by \mathcal{P}_m^d .

Lemma 5. *There exists a constant c such that for any $j > 1/c$,*

$$\|\mathcal{K}_j(u)\|_{\mathcal{L}^2([0,1]^d)} \geq c\|u\|_{\mathcal{L}^2([0,1]^d)}$$

for all polynomials $u \in \mathcal{P}_{s-1}^d$.

Proof. We proceed by contradiction, following the proof of (Arias-Castro) closely and making necessary adjustments. Suppose there exists a sequence $(u_j) \subset \mathcal{P}_{s-1}^d$ such that $\|\mathcal{K}_j(u)\|_{\mathcal{L}^2([0,1]^d)} \leq 1/j\|u_j\|_{\mathcal{L}^2([0,1]^d)}$ for each $j \in \mathbb{N}$. Assume without loss of generality that $\|u_j\|_{\mathcal{L}^2([0,1]^d)} = 1$; then there exists an accumulation point u_∞ of (u_j) , and $\|u_\infty\|_{\mathcal{L}^2([0,1]^d)} = 1$. However, by Jensen's inequality, and the boundedness and compact support of η , we have

$$\begin{aligned} \left[(u * \eta_h)(\bar{x})\right]^2 &= \left[\int u(y)\eta_h(y - \bar{x}) dy\right]^2 \\ &= \left[\frac{1}{h^d} \int u(y)\eta\left(\frac{y - \bar{x}}{h}\right) dy\right]^2 \\ &\leq \frac{1}{h^d} \left[u(y)\eta\left(\frac{y - \bar{x}}{h}\right)\right]^2 dy \\ &\leq \frac{1}{h^d} \int_{Q_h(\bar{x})} [u(y)]^2 dy \end{aligned}$$

which in turn gives

$$\begin{aligned}\|\mathcal{K}_j(u)\|_{\mathcal{L}^2([0,1]^d)}^2 &\leq \sum_{\bar{x} \in \bar{X}} h^d \left[(u * \eta_h)(\bar{x}) \right]^2 \\ &\leq \sum_{\bar{x} \in \bar{X}} \int_{Q_h(\bar{x})} [u(y)]^2 \\ &= \|u\|_{\mathcal{L}^2([0,1]^d)}^2.\end{aligned}$$

As a result,

$$\begin{aligned}\|\mathcal{K}_j(u_\infty)\|_{\mathcal{L}^2([0,1]^d)} &\leq \|\mathcal{K}_j(u_\infty - u_j)\|_{\mathcal{L}^2([0,1]^d)} + \|\mathcal{K}_j(u_j)\|_{\mathcal{L}^2([0,1]^d)} \\ &\leq \|u_\infty - u_j\|_{\mathcal{L}^2([0,1]^d)} + 1/j \rightarrow 0\end{aligned}$$

as $j \rightarrow \infty$. On the other hand, $\mathcal{K}_j(u_\infty) \rightarrow u_\infty$ as $j \rightarrow \infty$. In light of the boundedness of η and the boundedness of u_∞ on $[0,1]^d$, we may apply the dominated convergence theorem and obtain that $\|\mathcal{K}_j(u_\infty)\|_{\mathcal{L}^2([0,1]^d)} \rightarrow \|u_\infty\|_{\mathcal{L}^2([0,1]^d)}$. This establishes a contradiction, thus proving the claim. \square

3.3 Grid Sobolev seminorm

Actually, what I said before was wrong. Let me prove it. For notational convenience let $h = t^{-1}$. We write $\Delta = -\sum_{i=1}^d \frac{\partial^2}{\partial e_i^2}$ for the continuum Laplace operator on \mathbb{R}^d .

Lemma 6. *Let $f \in C_c^s(\mathcal{X})$.*

1. *If s is even, let $q = s/2$. We have*

$$|\bar{L}^q f(\bar{x})| \leq c \|f\|_{C^s(\mathcal{X})} h^s \quad (19)$$

for every $\bar{x} \in \bar{X}$.

2. *If s is odd, let $q = (s-1)/2$. Then*

(a) *If $\bar{x} \in \mathcal{X}_{q/t}$, then*

$$|\bar{L}^q f(\bar{x}) - h^{s-1} \Delta^q f(\bar{x})| \leq c \|f\|_{C^s(\mathcal{X})} h^s \quad (20)$$

(b) *Otherwise if $\bar{x} \in \partial_h(\mathcal{X})$, then*

$$|\bar{L}^q f(\bar{x})| \leq c \|f\|_{C^s(\mathcal{X})} h^s$$

Proof. We split our analysis into cases.

Proof of 1. Evaluated at an interior point $\bar{x} \in \mathcal{X}_r$, we have that

$$\bar{L}f(\bar{x}) = \sum_{i=1}^d \bar{D}_{e_i^2} f(\bar{x}) \quad (21)$$

where we use the notation $\bar{D}_{e_i^2} f(\bar{x}) := f(\bar{x} + he_i) + f(\bar{x} - he_i) - 2f(\bar{x})$. The second difference operator $\bar{D}_{e_i^2}$ zeros out the odd terms in a Taylor expansion as follows,

$$\begin{aligned}\bar{D}_{e_i^2} f(\bar{x}) &= \sum_{\alpha=1}^{s-1} \frac{\partial^\alpha}{\partial e_i^\alpha} f(\bar{x}) \left[(he_i)^\alpha + (-he_i)^\alpha \right] + R(\bar{x}, i) \\ &= 2 \sum_{\beta=1}^{q-1} h^{2\beta} \frac{\partial^{2\beta}}{\partial e_i^{2\beta}} f(\bar{x}) + R(\bar{x}, i).\end{aligned}$$

Here $R(\bar{x}, i)$ is the remainder term which satisfies

$$|R(\bar{x}, i)| = \left| \int_{-h}^h \frac{\partial^s}{\partial e_i^s} f(\bar{x} + ye_i) y^{s-1} dy \right| \leq 2h^s \|f\|_{C^s(\mathcal{X})}$$

by the Holder property of f . Plugging back in to (21), we obtain

$$\bar{L}f(\bar{x}) = 2 \sum_{i=1}^d \left\{ \sum_{\beta=1}^{q-1} h^{2\beta} \frac{\partial^{2\beta}}{\partial e_i^{2\beta}} f(\bar{x}) + R(\bar{x}, i) \right\}$$

Repeatedly applying this estimate along with the triangle inequality gives that for any $\bar{x} \in \mathcal{X}_{qr}$,

$$\left| \bar{L}^q f(\bar{x}) - 2 \sum_{i \in [d]^q} \sum_{\beta \in [q-1]^q} h^{2|\beta|} \frac{\partial^{2|\beta|}}{\partial e_{i_1}^{2\beta_1} \dots \partial e_{i_q}^{2\beta_q}} f(\bar{x}) \right| \leq ch^s \|f\|_{C^s(\mathcal{X})} \quad (22)$$

where the interior sum over all multi-indices β such that $|\beta| \leq s$. Since $|\beta|$ is at least q for all $\beta \in [q]^q$, using the Holder property of f for a second time we have

$$h^{2|\beta|} \left| \frac{\partial^{2|\beta|}}{\partial e_{i_1}^{2\beta_1} \dots \partial e_{i_q}^{2\beta_q}} f(\bar{x}) \right| \leq h^s \|f\|_{C^s(\mathcal{X})}$$

which along with (22) implies (19) for each $\bar{x} \in \mathcal{X}_{qh}$.

If instead $\bar{x} \in \partial_{qh}(\mathcal{X})$, there exists some boundary point $x_0 \in \partial(\mathcal{X})$ such that $\text{dist}(\bar{x}, x_0) \leq qh$. The order α -terms, for $\alpha = 1, \dots, s-1$, in the Taylor expansion of $f(\bar{x})$ about $f(x_0)$ are therefore zero, and

$$|f(\bar{x})| \leq (2qh)^s \|f\|_{C^s(\mathcal{X})}.$$

Applying similar logic to all $\bar{z} \in Q_{2qh}(\bar{z})$, we have that

$$|\bar{L}f(\bar{x})| \leq (2q)^d \cdot \max_{\bar{z} \in Q_{2qh}(\bar{z})} \left\{ |f(\bar{z})| \right\} \leq (2q)^{d+s} h^s \|f\|_{C^s(\mathcal{X})}.$$

Therefore (19) holds for all $\bar{x} \in \bar{X}$, regardless of their distance to the boundary.

Proof of 2a. We now deal with the case where s is odd, and take $q = (s-1)/2$. Similar reasoning to that of the preceding section implies that for any $\bar{x} \in \mathcal{X}_r$,

$$\bar{L}f(\bar{x}) = 2 \sum_{i=1}^d \left\{ \sum_{\beta=1}^q h^{2\beta} \frac{\partial^{2\beta}}{\partial e_i^{2\beta}} f(\bar{x}) + R(\bar{x}, i) \right\}$$

and additionally

$$\left| \bar{L}^q f(\bar{x}) - 2 \sum_{i \in [d]^q} \sum_{\beta \in [q]^q} h^{2|\beta|} \frac{\partial^{2|\beta|}}{\partial e_{i_1}^{2\beta_1} \dots \partial e_{i_q}^{2\beta_q}} f(\bar{x}) \right| \leq ch^s \|f\|_{C^s(\mathcal{X})} \quad (23)$$

where as before $|\beta| < s$ in the interior sum. Now, for the specific choice of $\beta_\star = (1, \dots, 1)$, we have

$$\sum_{i \in [d]^q} h^{2|\beta_\star|} \frac{\partial^{2q}}{\partial e_{i_1}^2 \dots \partial e_{i_q}^2} f(\bar{x}) = h^{s-1} \Delta^q f(\bar{x}_i).$$

For all other $\beta \neq \beta_\star \in [q]^q$ satisfying $|\beta| < s$ we have the same estimate as before,

$$h^{2|\beta|} \left| \frac{\partial^{2|\beta|}}{\partial e_{i_1}^{2\beta_1} \dots \partial e_{i_q}^{2\beta_q}} f(\bar{x}) \right| \leq h^s \|f\|_{C^s(\mathcal{X})}$$

and plugging back in to (23) establishes (20).

Proof of 2b. The proof of 2b. follows along the same lines as the boundary case when s was even. \square

4 Old Stuff.

To do so, we rely on Lemma 6, which estimates on the absolute value of $L^q f$. When s is even, the result follows immediately from (11) in that Lemma,

$$f^T \bar{L}^s f = \sum_{\bar{x} \in \bar{X}} (\bar{L}^q f(\bar{x}))^2 \leq cn \|f\|_{C^s(\mathcal{X})}^2 t^{-2s}.$$

When s is even, the result follows by distinguishing between points close to the boundary and points sufficiently on the interior,

$$\begin{aligned} f^T \bar{L}^s f &= \sum_{\bar{x}_i, \bar{x}_j \in \bar{X}} (\bar{L}^q f(\bar{x}_i) - \bar{L}^q f(\bar{x}_j))^2 \mathbf{1}\{\|\bar{x}_i - \bar{x}_j\|_1 \leq h\} \\ &\leq cnt^{-2s} \|f\|_{C^s(\mathcal{X})}^2 + \sum_{\bar{x}_i, \bar{x}_j \in \bar{X} \cap \mathcal{X}_{qh}} (\bar{L}^q f(\bar{x}_i) - \bar{L}^q f(\bar{x}_j))^2 \mathbf{1}\{\|\bar{x}_i - \bar{x}_j\|_1 \leq h\} \\ &\leq cnt^{-2s} \|f\|_{C^s(\mathcal{X})}^2 + h^{2(s-1)} \sum_{\bar{x}_i, \bar{x}_j \in \bar{X} \cap \mathcal{X}_{qh}} (\Delta^q f(\bar{x}_i) - \Delta^q f(\bar{x}_j))^2 \mathbf{1}\{\|\bar{x}_i - \bar{x}_j\|_1 \leq h\} \\ &\leq cnt^{-2s} \|f\|_{C^s(\mathcal{X})}^2 \end{aligned}$$

where the last inequality follows from applying the Holder condition to $\Delta^q f$.