Convergence rates for graph-based learning featuring singular PDEs

Leon Bungert

Institute of Mathematics & CAIDAS, University of Würzburg

March 19, 2024

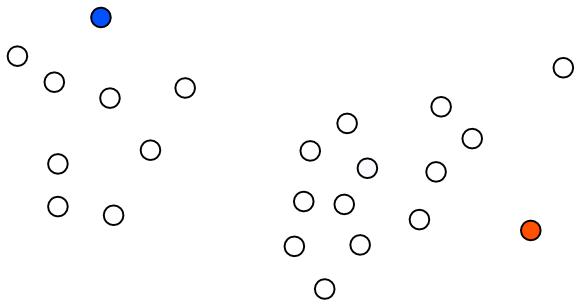
CNA seminar @ CMU Pittsburgh

Outline

Graph-Based Semi-Supervised Learning

Given:

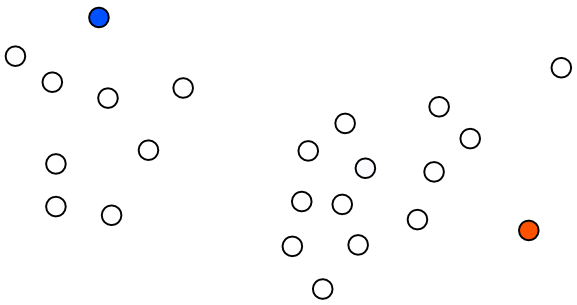
- a data set $\Omega_n \subset \mathbb{R}^d$,
- with labels $g : \mathcal{O}_n \subset \Omega_n \to \mathbb{R}$.



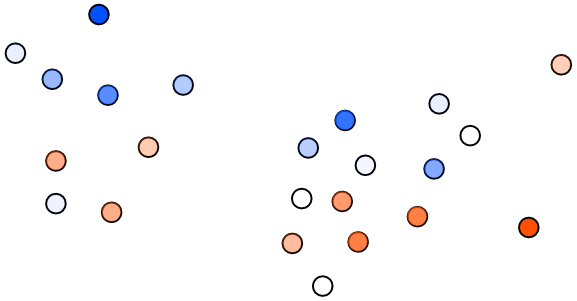
Graph-Based Semi-Supervised Learning

Given:

- a data set $\Omega_n \subset \mathbb{R}^d$,
- with labels $g : \mathcal{O}_n \subset \Omega_n \to \mathbb{R}$.



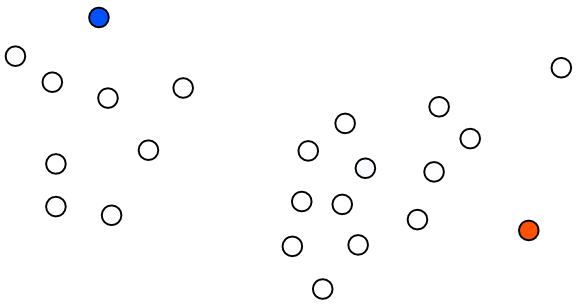
Goal:
a "smooth" function $\mathbf{u} : \Omega_n \to \mathbb{R}$
such that $\mathbf{u} = g$ on $\mathcal{O}_n$.



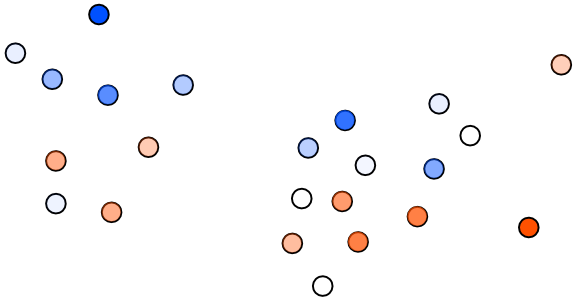
Graph-Based Semi-Supervised Learning

Given:

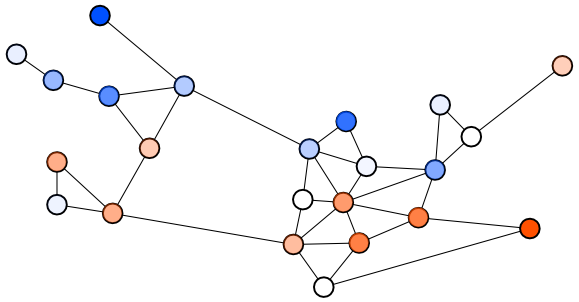
- a data set $\Omega_n \subset \mathbb{R}^d$,
- with labels $g : \mathcal{O}_n \subset \Omega_n \to \mathbb{R}$.



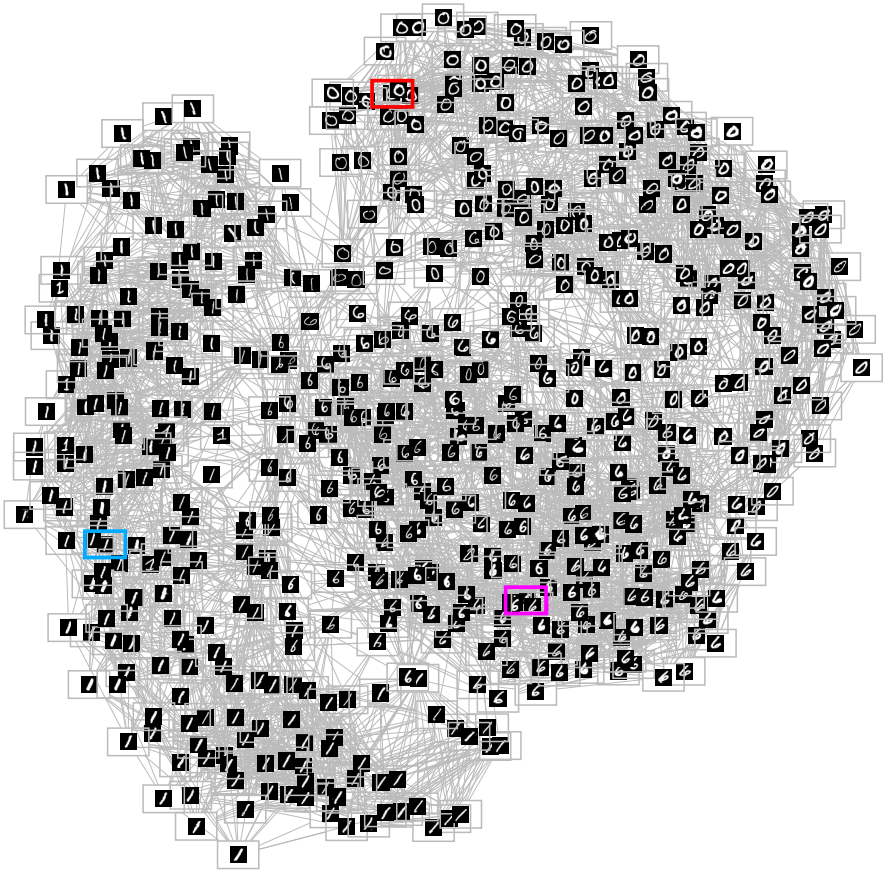
Goal:
a "smooth" function $\mathbf{u} : \Omega_n \to \mathbb{R}$
such that $\mathbf{u} = g$ on $\mathcal{O}_n$.



Model:
construct a weighted graph $G_n = (\Omega_n, \omega_n)$
with edge weights $\omega_n : \Omega_n \times \Omega_n \to \mathbb{R}$.

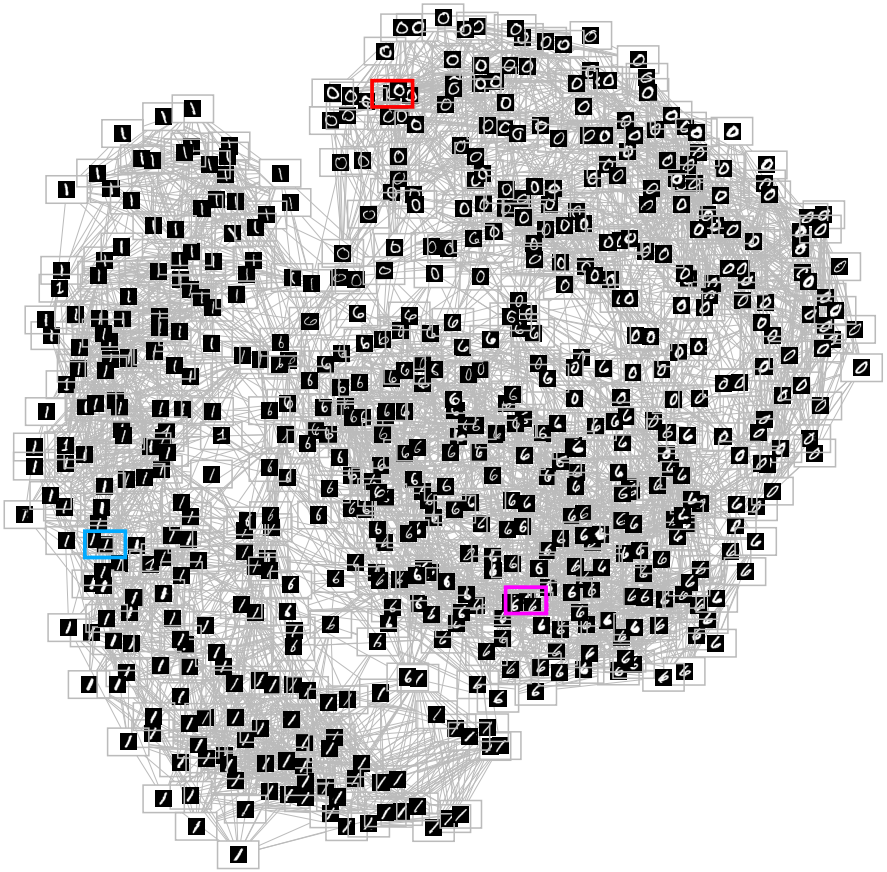


Graph-Based Semi-Supervised Learning

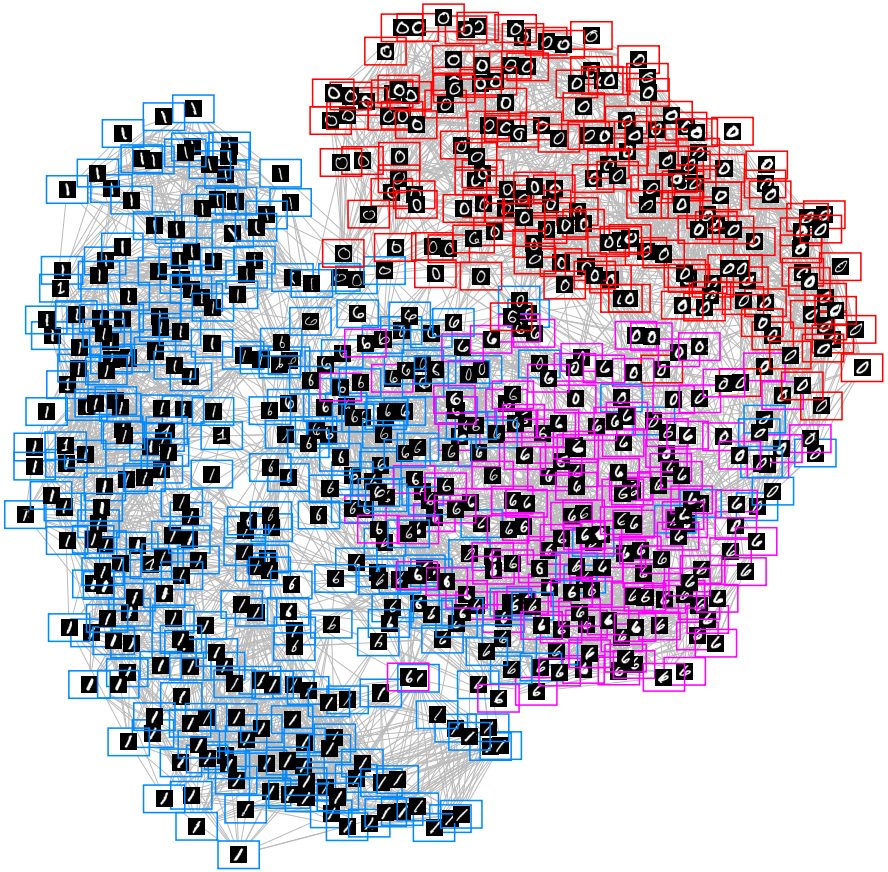


Represent data as weighted graph

Graph-Based Semi-Supervised Learning



Represent data as weighted graph



Extend labels solving a graph PDE

Laplacian Learning

Laplacian Learning

In Laplacian learning [ZGL03] one computes a harmonic extension of the labels via

Laplacian Learning

In Laplacian learning [ZGL03] one computes a harmonic extension of the labels via

$$\min_{\mathbf{u}} \ \mathcal{E}_2^n(\mathbf{u}),$$
$$\text{subject to } \mathbf{u} = g \text{ on } \mathcal{O}_n.$$

Laplacian Learning

In Laplacian learning [ZGL03] one computes a harmonic extension of the labels via

$$\min_{\mathbf{u}} \ \mathcal{E}_2^n(\mathbf{u}), \quad \text{subject to } \mathbf{u} = g \text{ on } \mathcal{O}_n. \qquad \Longleftrightarrow \qquad \mathcal{L}_2^n \mathbf{u} = 0 \text{ in } \Omega_n \setminus \mathcal{O}_n, \quad \mathbf{u} = g \text{ on } \mathcal{O}_n,$$

Laplacian Learning

In Laplacian learning [ZGL03] one computes a harmonic extension of the labels via

$$\min_{\mathbf{u}} \ \mathcal{E}_2^n(\mathbf{u}), \quad \text{subject to } \mathbf{u} = g \text{ on } \mathcal{O}_n. \qquad \Longleftrightarrow \qquad \mathcal{L}_2^n \mathbf{u} = 0 \text{ in } \Omega_n \setminus \mathcal{O}_n, \quad \mathbf{u} = g \text{ on } \mathcal{O}_n,$$

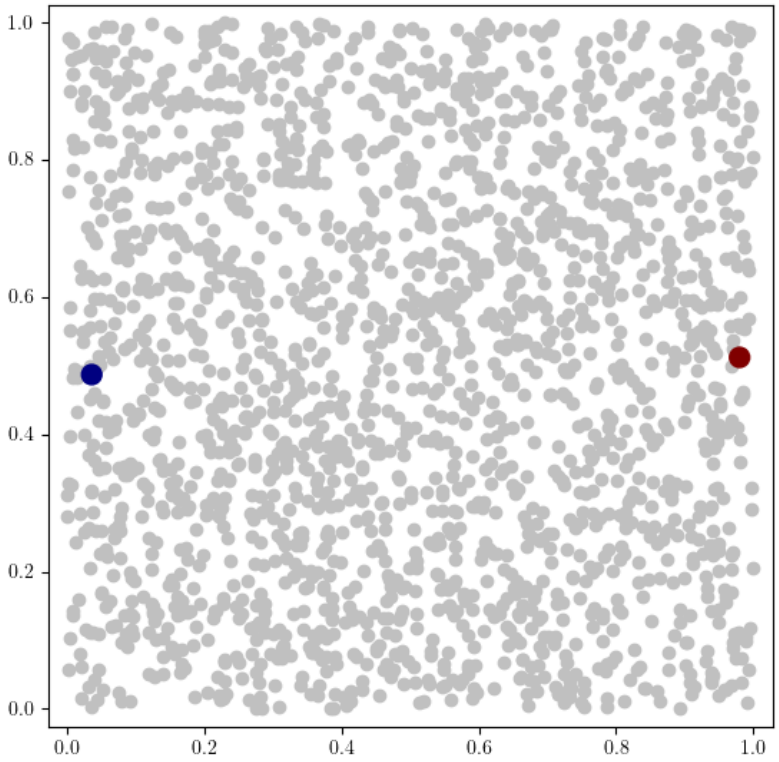
$$\mathcal{E}_2^n(\mathbf{u}) := \sum_{x,y \in \Omega_n} \omega_n(x,y) \left|\mathbf{u}(x) - \mathbf{u}(y)\right|^2$$ denotes the graph Dirichlet-energy,

$$\mathcal{L}_2^n \mathbf{u}(x) := \sum_{y \in \Omega_n} \omega_n(x,y)(\mathbf{u}(y) - \mathbf{u}(x))$$ denotes the graph Laplacian.

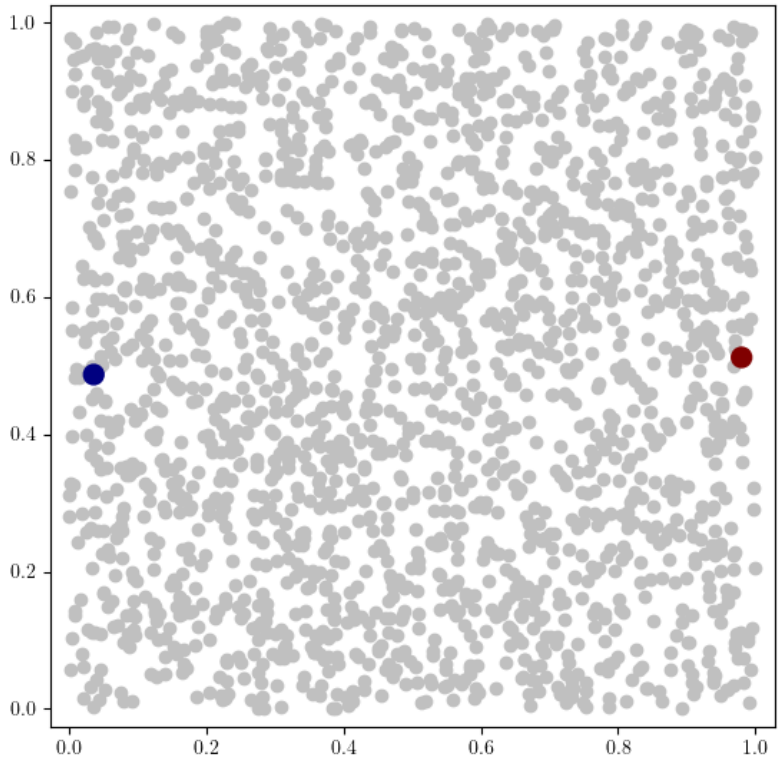
What are the challenges in semi-supervised learning?

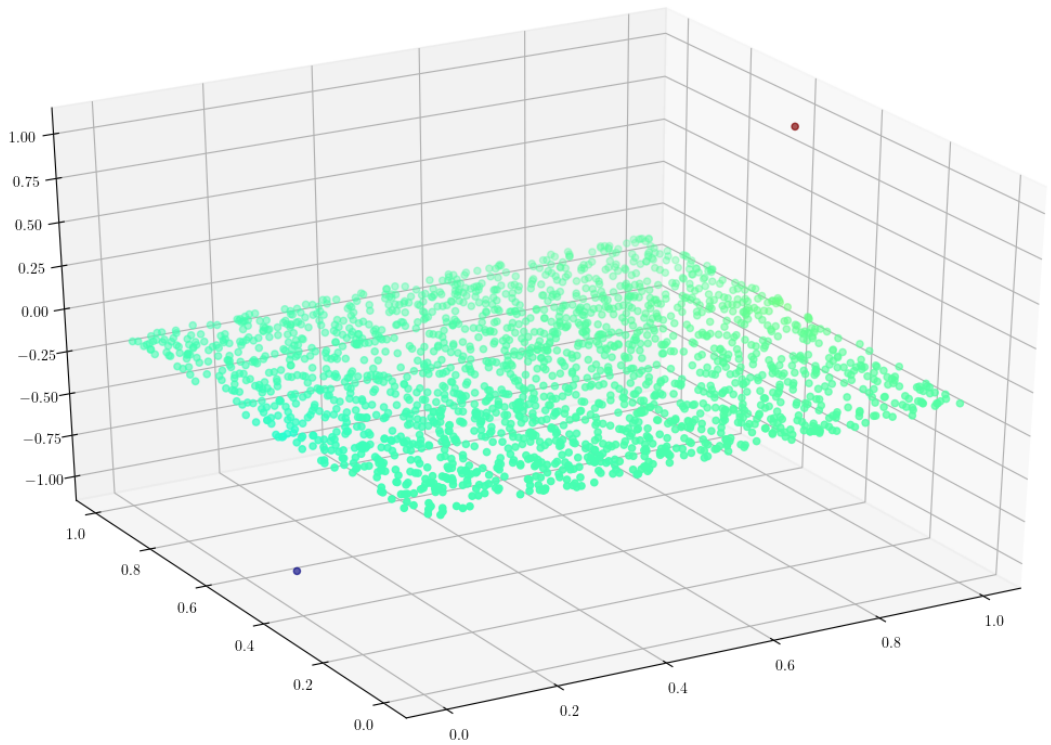
Spiking in Laplacian Learning

Spiking in Laplacian Learning



Spiking in Laplacian Learning





Graph harmonic extension of two labels [ZGL03]

$$\mathcal{L}_2^n \mathbf{u}_n = 0 \text{ on } \Omega_n \setminus \mathcal{O}_n \quad \text{subject to } \mathbf{u}_n = g \text{ on } \mathcal{O}_n.$$

Continuum limits of Graph-Based SSL

Graph-Based Semi-Supervised Learning

- **Given:** Weighted graph $G_n = (\Omega_n, \omega_n)$ with labels $g : \mathcal{O}_n \subset \Omega_n \to \mathbb{R}^k$.
- **Goal:** $\mathbf{u}_n : \Omega_n \to \mathbb{R}$ such that $\mathbf{u}_n = g$ on $\mathcal{O}_n$
- **Tool:** graph PDE

Continuum limits of Graph-Based SSL

Graph-Based Semi-Supervised Learning

- **Given:** Weighted graph $G_n = (\Omega_n, \omega_n)$ with labels $g : \mathcal{O}_n \subset \Omega_n \to \mathbb{R}^k$.
- **Goal:** $\mathbf{u}_n : \Omega_n \to \mathbb{R}$ such that $\mathbf{u}_n = g$ on $\mathcal{O}_n$
- **Tool:** graph PDE

Driving question:
What happens with more and more data, $|\Omega_n| \to \infty$ (and potentially $|\mathcal{O}_n| \to \infty$)?

Continuum limits of Graph-Based SSL

Graph-Based Semi-Supervised Learning

- **Given:** Weighted graph $G_n = (\Omega_n, \omega_n)$ with labels $g : \mathcal{O}_n \subset \Omega_n \to \mathbb{R}^k$.
- **Goal:** $\mathbf{u}_n : \Omega_n \to \mathbb{R}$ such that $\mathbf{u}_n = g$ on $\mathcal{O}_n$
- **Tool:** graph PDE

Driving question:
What happens with more and more data, $|\Omega_n| \to \infty$ (and potentially $|\mathcal{O}_n| \to \infty$)?

graph PDE $\to$ continuum PDE, as $n \to \infty$.

Continuum limits of Graph-Based SSL

Graph-Based Semi-Supervised Learning

- **Given:** Weighted graph $G_n = (\Omega_n, \omega_n)$ with labels $g : \mathcal{O}_n \subset \Omega_n \to \mathbb{R}^k$.
- **Goal:** $\mathbf{u}_n : \Omega_n \to \mathbb{R}$ such that $\mathbf{u}_n = g$ on $\mathcal{O}_n$
- **Tool:** graph PDE

Driving question:
What happens with more and more data, $|\Omega_n| \to \infty$ (and potentially $|\mathcal{O}_n| \to \infty$)?

graph PDE $\to$ continuum PDE, as $n \to \infty$.

Rates of convergence?

Four Fixes to the Spiking Issue

1. sufficiently large labeling rates $\frac{|\mathcal{O}_n|}{|\Omega_n|}$ [CST23] or reweighting [CS20];

Four Fixes to the Spiking Issue

1. sufficiently large labeling rates $\frac{|\mathcal{O}_n|}{|\Omega_n|}$ [CST23] or reweighting [CS20];
2. p-Laplace learning with $p > d$ [El +16], well-posed continuum limit [ST19] via Γ-convergence (no convergence rates);

Four Fixes to the Spiking Issue

1. sufficiently large labeling rates $\frac{|\mathcal{O}_n|}{|\Omega_n|}$ [CST23] or reweighting [CS20];
2. p-Laplace learning with $p > d$ [El +16], well-posed continuum limit [ST19] via Γ-convergence (no convergence rates);
3. Lipschitz learning (i.e., $p = \infty$), well-posed continuum limit [Cal19; RB23] with rates [BCR23; BCR24];

Four Fixes to the Spiking Issue

1. sufficiently large labeling rates $\frac{|\mathcal{O}_n|}{|\Omega_n|}$ [CST23] or reweighting [CS20];
2. p-Laplace learning with $p > d$ [El +16], well-posed continuum limit [ST19] via Γ-convergence (no convergence rates);
3. Lipschitz learning (i.e., $p = \infty$), well-posed continuum limit [Cal19; RB23] with rates [BCR23; BCR24];
4. Poisson learning [Cal+20], little empirical degradation for small label rates, well-posed continuum limit [B+24]

What is Lipschitz learning?

Lipschitz Learning

Lipschitz Learning

Find $\mathbf{u}_n : \Omega_n \to \mathbb{R}$ such that

$$\begin{cases} \mathcal{L}_\infty^n \mathbf{u}_n = 0, & \text{in } \Omega_n \setminus \mathcal{O}_n, \\ \mathbf{u}_n = g, & \text{in } \mathcal{O}_n. \end{cases} \tag{LL}$$

Lipschitz Learning

Lipschitz Learning

Find $\mathbf{u}_n : \Omega_n \to \mathbb{R}$ such that

$$\begin{cases} \mathcal{L}_\infty^n \mathbf{u}_n = 0, & \text{in } \Omega_n \setminus \mathcal{O}_n, \\ \mathbf{u}_n = g, & \text{in } \mathcal{O}_n. \end{cases} \tag{LL}$$

Continuum Problem

Find $u \in C(\Omega)$, viscosity solution of

$$\begin{cases} \Delta_\infty u = 0, & \text{in } \Omega \setminus \mathcal{O}, \\ u = g, & \text{in } \mathcal{O}. \end{cases} \tag{IL}$$

Lipschitz Learning

Lipschitz Learning

Find $\mathbf{u}_n : \Omega_n \to \mathbb{R}$ such that

$$\begin{cases} \mathcal{L}_\infty^n \mathbf{u}_n = 0, & \text{in } \Omega_n \setminus \mathcal{O}_n, \\ \mathbf{u}_n = g, & \text{in } \mathcal{O}_n. \end{cases} \tag{LL}$$

Continuum Problem

Find $u \in C(\Omega)$, viscosity solution of

$$\begin{cases} \Delta_\infty u = 0, & \text{in } \Omega \setminus \mathcal{O}, \\ u = g, & \text{in } \mathcal{O}. \end{cases} \tag{IL}$$

$$\mathcal{L}_\infty^n \mathbf{u}_n(x) := \max_{y \in \Omega_n} \omega_n(x,y)(\mathbf{u}_n(y) - \mathbf{u}_n(x)) + \min_{y \in \Omega_n} \omega_n(x,y)(\mathbf{u}_n(y) - \mathbf{u}_n(x))$$

$$\Delta_\infty u(x) := \langle \nabla u(x), D^2 u(x) \nabla u(x) \rangle$$

Lipschitz Learning

Lipschitz Learning

Find $\mathbf{u}_n : \Omega_n \to \mathbb{R}$ such that

$$\begin{cases} \mathcal{L}_\infty^n \mathbf{u}_n = 0, & \text{in } \Omega_n \setminus \mathcal{O}_n, \\ \mathbf{u}_n = g, & \text{in } \mathcal{O}_n. \end{cases} \tag{LL}$$

Continuum Problem

Find $u \in C(\Omega)$, viscosity solution of

$$\begin{cases} \Delta_\infty u = 0, & \text{in } \Omega \setminus \mathcal{O}, \\ u = g, & \text{in } \mathcal{O}. \end{cases} \tag{IL}$$

Previous work:

Lipschitz Learning

Lipschitz Learning

Find $\mathbf{u}_n : \Omega_n \to \mathbb{R}$ such that

$$\begin{cases} \mathcal{L}_\infty^n \mathbf{u}_n = 0, & \text{in } \Omega_n \setminus \mathcal{O}_n, \\ \mathbf{u}_n = g, & \text{in } \mathcal{O}_n. \end{cases} \tag{LL}$$

Continuum Problem

Find $u \in C(\Omega)$, viscosity solution of

$$\begin{cases} \Delta_\infty u = 0, & \text{in } \Omega \setminus \mathcal{O}, \\ u = g, & \text{in } \mathcal{O}. \end{cases} \tag{IL}$$

Previous work:

- [RB23]: Γ-convergence of associated energies (just necessary conditions)

Lipschitz Learning

Lipschitz Learning

Find $\mathbf{u}_n : \Omega_n \to \mathbb{R}$ such that

$$\begin{cases} \mathcal{L}_\infty^n \mathbf{u}_n = 0, & \text{in } \Omega_n \setminus \mathcal{O}_n, \\ \mathbf{u}_n = g, & \text{in } \mathcal{O}_n. \end{cases} \tag{LL}$$

Continuum Problem

Find $u \in C(\Omega)$, viscosity solution of

$$\begin{cases} \Delta_\infty u = 0, & \text{in } \Omega \setminus \mathcal{O}, \\ u = g, & \text{in } \mathcal{O}. \end{cases} \tag{IL}$$

Previous work:

- [RB23]: Γ-convergence of associated energies (just necessary conditions)
- [Sma10; Cal19]: Convergence under strong assumptions (i.a., dense graphs), no rates

Lipschitz Learning

Lipschitz Learning

Find $\mathbf{u}_n : \Omega_n \to \mathbb{R}$ such that

$$\begin{cases} \mathcal{L}_\infty^n \mathbf{u}_n = 0, & \text{in } \Omega_n \setminus \mathcal{O}_n, \\ \mathbf{u}_n = g, & \text{in } \mathcal{O}_n. \end{cases} \tag{LL}$$

Continuum Problem

Find $u \in C(\Omega)$, viscosity solution of

$$\begin{cases} \Delta_\infty u = 0, & \text{in } \Omega \setminus \mathcal{O}, \\ u = g, & \text{in } \mathcal{O}. \end{cases} \tag{IL}$$

Previous work:

- [RB23]: Γ-convergence of associated energies (just necessary conditions)
- [Sma10; Cal19]: Convergence under strong assumptions (i.a., dense graphs), no rates

Aims: convergence rates for sparse graphs

Graph Bandwidth and Resolution

Graph Bandwidth

For a bandwidth $\varepsilon_n > 0$ and a function $\eta : (0, \infty) \to [0, \infty)$ we define

$$\omega_n(x, y) := \eta(|x - y| / \varepsilon_n), \quad x, y \in \Omega_n,\ x \neq y.$$

Graph Bandwidth and Resolution

Graph Bandwidth

For a bandwidth $\varepsilon_n > 0$ and a function $\eta : (0, \infty) \to [0, \infty)$ we define

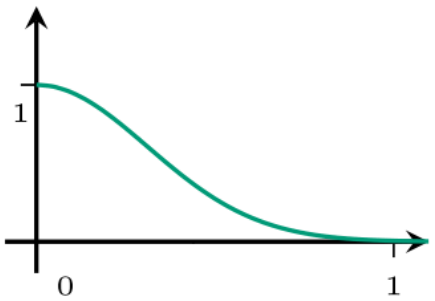
$$\omega_n(x, y) := \eta(|x - y| / \varepsilon_n), \quad x, y \in \Omega_n, \; x \neq y.$$

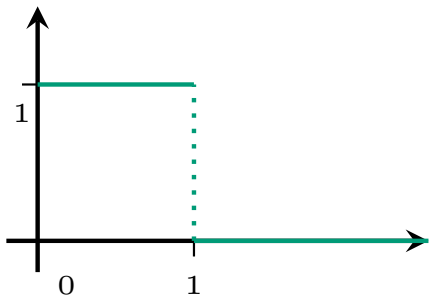
Assumptions on the Kernel

(K1) $\operatorname{supp}(\eta) \subset [0, 1]$

(K2) η is non-increasing

(K3) $\sigma_\eta := \operatorname{ess\,sup}_{t>0} t\eta(t) = 1$





Graph Bandwidth and Resolution

Graph Bandwidth

For a bandwidth $\varepsilon_n > 0$ and a function $\eta : (0, \infty) \to [0, \infty)$ we define

$$\omega_n(x, y) := \eta(|x - y| / \varepsilon_n), \quad x, y \in \Omega_n, \ x \neq y.$$

Graph Resolution

We define the graph resolution as

$$\delta_n := d_H(\Omega_n, \Omega) \vee d_H(\mathcal{O}_n, \mathcal{O}),$$

where we use the Hausdorff distance

$$d_H(A, B) = \sup_{x \in A} \inf_{y \in B} |x - y| \vee \sup_{x \in B} \inf_{y \in A} |x - y|, \quad A, B \subset \mathbb{R}^d.$$

Graph Bandwidth and Resolution

Graph Bandwidth

For a bandwidth $\varepsilon_n > 0$ and a function $\eta : (0, \infty) \to [0, \infty)$ we define

$$\omega_n(x, y) := \eta(|x - y| / \varepsilon_n), \quad x, y \in \Omega_n,\ x \neq y.$$

Graph Resolution

We define the graph resolution as

$$\delta_n := d_H(\Omega_n, \Omega) \vee d_H(\mathcal{O}_n, \mathcal{O}),$$

where we use the Hausdorff distance

$$d_H(A, B) = \sup_{x \in A} \inf_{y \in B} |x - y| \vee \sup_{x \in B} \inf_{y \in A} |x - y|, \quad A, B \subset \mathbb{R}^d.$$

Important relations:

$\frac{\delta_n}{\varepsilon_n} < \frac{1}{4\sqrt{d}}$: connected graph

$\delta_n \sim \left(\frac{\log n}{n}\right)^{\frac{1}{d}}$ for a random graph

General Convergence Statement

Theorem ([BCR23])

Let $\Omega \subset \mathbb{R}^d$ be a locally convex domain, $\Omega_n \subset \Omega$ be an arbitrary set of points, let $\mathbf{u}_n : \Omega_n \to \mathbb{R}$ solve (LL) *and $u : \Omega \to \mathbb{R}$ solve* (IL)*, and let $\tau > 0$ be arbitrary.*

If $\delta_n \lesssim \varepsilon_n \lesssim \tau$ then it holds

$$\max_{\Omega_n} |\mathbf{u}_n - u| \lesssim \tau + \sqrt[3]{\frac{\delta_n}{\varepsilon_n \tau} + \frac{\varepsilon_n}{\tau^2}}.$$

General Convergence Statement

Theorem ([BCR23])

Let $\Omega \subset \mathbb{R}^d$ be a locally convex domain, $\Omega_n \subset \Omega$ be an arbitrary set of points, let $\mathbf{u}_n : \Omega_n \to \mathbb{R}$ solve (LL) *and $u : \Omega \to \mathbb{R}$ solve* (IL)*, and let $\tau > 0$ be arbitrary.*

If $\delta_n \lesssim \varepsilon_n \lesssim \tau$ then it holds

$$\max_{\Omega_n} |\mathbf{u}_n - u| \lesssim \tau + \sqrt[3]{\frac{\delta_n}{\varepsilon_n \tau} + \frac{\varepsilon_n}{\tau^2}}.$$

Corollary

Optimizing over $\tau = \tau_n$ one gets:

- *(Sparse regime): If $\delta_n \lesssim \varepsilon_n \lesssim \delta_n^{\frac{5}{9}}$ the rate is $\left(\frac{\delta_n}{\varepsilon_n}\right)^{\frac{1}{4}}$.*

General Convergence Statement

Theorem ([BCR23])

Let $\Omega \subset \mathbb{R}^d$ be a locally convex domain, $\Omega_n \subset \Omega$ be an arbitrary set of points, let $\mathbf{u}_n : \Omega_n \to \mathbb{R}$ solve (LL) *and $u : \Omega \to \mathbb{R}$ solve* (IL)*, and let $\tau > 0$ be arbitrary.*

If $\delta_n \lesssim \varepsilon_n \lesssim \tau$ then it holds

$$\max_{\Omega_n} |\mathbf{u}_n - u| \lesssim \tau + \sqrt[3]{\frac{\delta_n}{\varepsilon_n \tau} + \frac{\varepsilon_n}{\tau^2}}.$$

Corollary

Optimizing over $\tau = \tau_n$ one gets:

- *(Sparse regime): If $\delta_n \lesssim \varepsilon_n \lesssim \delta_n^{\frac{5}{9}}$ the rate is $\left(\frac{\delta_n}{\varepsilon_n}\right)^{\frac{1}{4}}$.*
- *(Dense regime): If $\varepsilon_n \gtrsim \delta_n^{\frac{5}{9}}$ the rate is $\varepsilon_n^{\frac{1}{5}}$.*

Improved Convergence on Random Graphs

Theorem ([BCR24])

Assume that Ω_n is a uniform i.i.d. sample or a homogeneous Poisson point process, let $\eta(t) = \frac{1}{t} 1_{t \leq 1}$, et ceteris paribus.

If $(\log n/n)^{1/d} \lesssim \varepsilon_n \lesssim \tau$ then it holds with high probability:

$$\max_{x \in \Omega_n} |\mathbf{u}_n - u| \lesssim \tau + \sqrt[3]{\log n \frac{(\log n/n)^{1/d}}{\sqrt{\tau^3 \varepsilon_n}} + \frac{\varepsilon_n}{\tau^2}}$$

Improved Convergence on Random Graphs

Theorem ([BCR24])

Assume that Ω_n is a uniform i.i.d. sample or a homogeneous Poisson point process, let $\eta(t) = \frac{1}{t} 1_{t \leq 1}$, et ceteris paribus.

If $(\log n/n)^{1/d} \lesssim \varepsilon_n \lesssim \tau$ then it holds with high probability:

$$\max_{x \in \Omega_n} |\mathbf{u}_n - u| \lesssim \tau + \sqrt[3]{\log n \frac{(\log n/n)^{1/d}}{\sqrt{\tau^3 \varepsilon_n}} + \frac{\varepsilon_n}{\tau^2}}$$

Corollary

Optimizing over $\tau = \tau_n$ and choosing $\varepsilon_n \sim (\log n/n)^{1/d}$ yields almost surely:

$$\max_{x \in \Omega_n} |\mathbf{u}_n - u| \lesssim (\log n)^{\frac{2}{9}} \left(\frac{\log n}{n} \right)^{\frac{1}{9d}}$$

Proof Strategy, Step 1

Introduce non-local operator with larger bandwidth $\tau \gg \varepsilon_n$

Proof Strategy, Step 1

Introduce non-local operator with larger bandwidth $\tau \gg \varepsilon_n$

$$\Delta_\infty^\tau u(x) := \frac{1}{\tau^2}\left(\sup_{y\in B(x;\tau)}(u(y)-u(x)) + \inf_{y\in B(x;\tau)}(u(y)-u(x))\right),$$

Proof Strategy, Step 1

Introduce non-local operator with larger bandwidth $\tau \gg \varepsilon_n$

$$\Delta_\infty^\tau u(x) := \frac{1}{\tau^2}\left(\sup_{y\in B(x;\tau)}(u(y)-u(x)) + \inf_{y\in B(x;\tau)}(u(y)-u(x))\right),$$

and define the infimal convolution

$$u_\tau(x) := \min_{y\in B(x;\tau)} u(y),$$

Using comparison with cones [AS10] proved

$$-\Delta_\infty u \geq 0 \implies -\Delta_\infty^\tau u_\tau \geq 0,$$

Proof Strategy, Step 1

Introduce non-local operator with larger bandwidth $\tau \gg \varepsilon_n$

$$\Delta_\infty^\tau u(x) := \frac{1}{\tau^2}\left(\sup_{y\in B(x;\tau)}(u(y)-u(x)) + \inf_{y\in B(x;\tau)}(u(y)-u(x))\right),$$

and define the infimal and discrete-to-continuum supremal convolution

$$u_\tau(x) := \min_{y\in B(x;\tau)} u(y), \qquad u_n^\tau(x) := \max_{y\in B(x;\tau)\cap\Omega_n} \mathbf{u}_n(y).$$

Using comparison with cones on graphs we proved in [BCR23]

$$-\Delta_\infty u \geq 0 \implies -\Delta_\infty^\tau u_\tau \geq 0,$$
$$-\mathcal{L}_\infty^n \mathbf{u}_n \leq 0 \implies -\Delta_\infty^\tau u_n^\tau \lesssim \frac{r_\tau}{\tau} + \frac{\varepsilon_n}{\tau^2},$$

Proof Strategy, Step 1

Introduce non-local operator with larger bandwidth $\tau \gg \varepsilon_n$

$$\Delta_\infty^\tau u(x) := \frac{1}{\tau^2}\left(\sup_{y\in B(x;\tau)}(u(y)-u(x)) + \inf_{y\in B(x;\tau)}(u(y)-u(x))\right),$$

and define the infimal and discrete-to-continuum supremal convolution

$$u_\tau(x) := \min_{y\in B(x;\tau)} u(y), \qquad u_n^\tau(x) := \max_{y\in B(x;\tau)\cap\Omega_n} \mathbf{u}_n(y).$$

Using comparison with cones on graphs we proved in [BCR23]

$$-\Delta_\infty u \geq 0 \implies -\Delta_\infty^\tau u_\tau \geq 0,$$
$$-\mathcal{L}_\infty^n \mathbf{u}_n \leq 0 \implies -\Delta_\infty^\tau u_n^\tau \lesssim \frac{r_\tau}{\tau} + \frac{\varepsilon_n}{\tau^2},$$

where we define the ratio of graph distances as

$$r_\tau(x) := \frac{\sup_{y\in\overline{B}(x,\tau)\cap\Omega_n} d_n(x,y)}{\inf_{y\in\Omega_n\setminus\overline{B}(x,2\tau-\varepsilon_n)} d_n(x,y)} - \frac{1}{2}.$$

Proof Strategy, Step 2

Use perturbation and comparison principle for $-\Delta_\infty^\tau$ to show

$$\sup_{\Omega_n}(\mathbf{u}_n - u) \lesssim \tau + \sqrt[3]{\frac{r_\tau}{\tau} + \frac{\varepsilon_n}{\tau^2}}.$$

Proof Strategy, Step 2

Use perturbation and comparison principle for $-\Delta_\infty^\tau$ to show

$$\sup_{\Omega_n}(\mathbf{u}_n - u) \lesssim \tau + \sqrt[3]{\frac{r_\tau}{\tau} + \frac{\varepsilon_n}{\tau^2}}.$$

Repeat everything using $-\mathcal{L}_\infty^n \mathbf{u}_n \geq 0$.

Proof Strategy, Step 2

Use perturbation and comparison principle for $-\Delta_\infty^\tau$ to show

$$\sup_{\Omega_n}(\mathbf{u}_n - u) \lesssim \tau + \sqrt[3]{\frac{r_\tau}{\tau} + \frac{\varepsilon_n}{\tau^2}}.$$

Repeat everything using $-\mathcal{L}_\infty^n \mathbf{u}_n \geq 0$.
Estimate the ratio r_τ using the general estimate [BCR23]

$$|x-y| \leq d_n(x,y) \leq \left(1 + C\frac{\delta_n}{\varepsilon_n}\right)|x-y| + \tau_\eta \varepsilon_n$$

Proof Strategy, Step 2

Use perturbation and comparison principle for $-\Delta_\infty^\tau$ to show

$$\sup_{\Omega_n}(\mathbf{u}_n - u) \lesssim \tau + \sqrt[3]{\frac{r_\tau}{\tau} + \frac{\varepsilon_n}{\tau^2}}.$$

Repeat everything using $-\mathcal{L}_\infty^n \mathbf{u}_n \geq 0$.
Estimate the ratio r_τ using the general estimate [BCR23]

$$|x-y| \leq d_n(x,y) \leq \left(1 + C\frac{\delta_n}{\varepsilon_n}\right)|x-y| + \tau_\eta \varepsilon_n$$

or the percolation estimate [BCR24], valid for $\varepsilon_n \sim \delta_n$,

$$r_\tau \lesssim \left(\frac{\log n}{n}\right)^{\frac{1}{d}} \frac{\log n}{\sqrt{\tau\varepsilon}}$$

which is based on homogenization of the graph distance.

Numerical Results

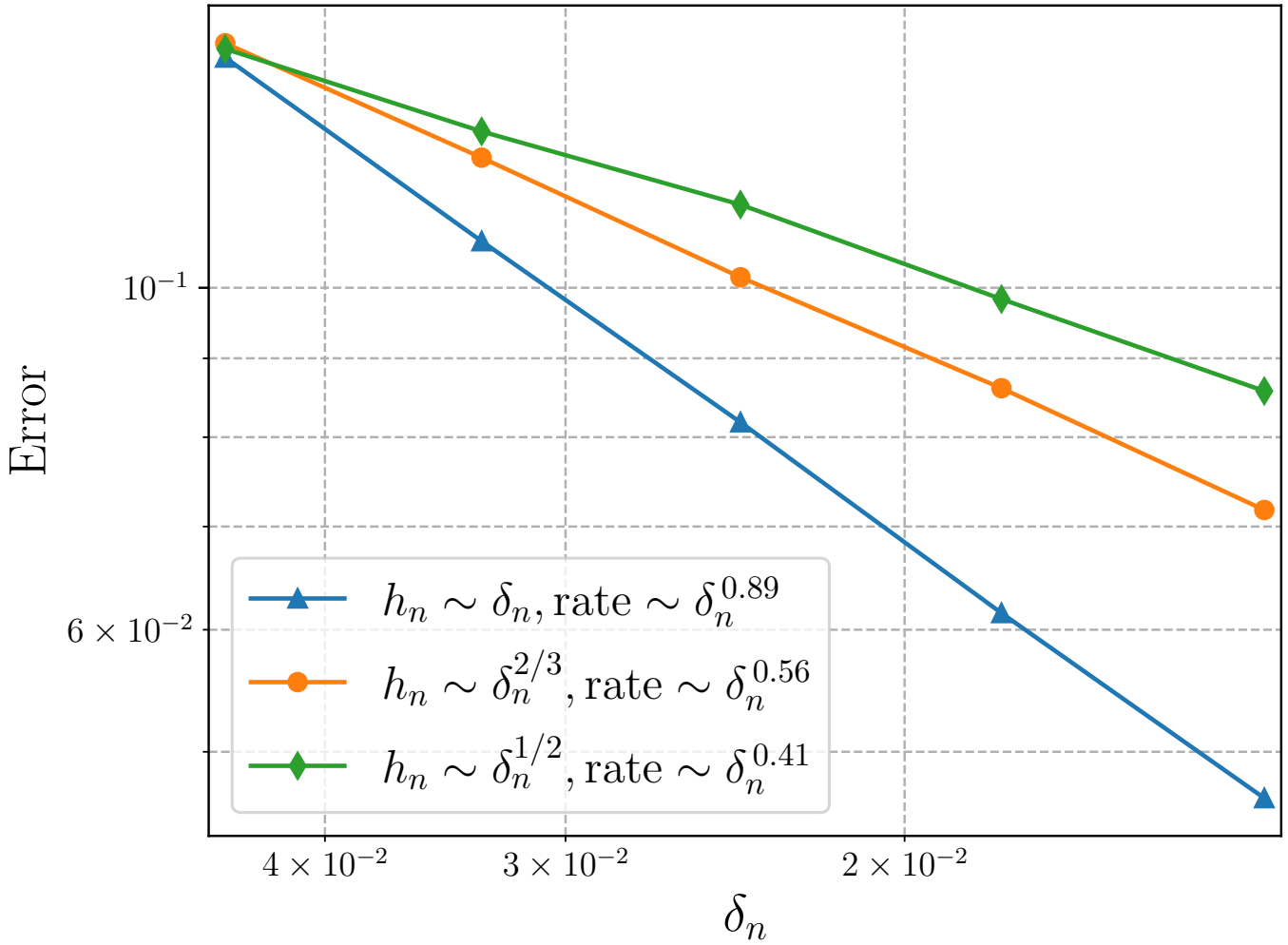


Figure: Experimental rates for constant weights on star domain.

Numerical Results

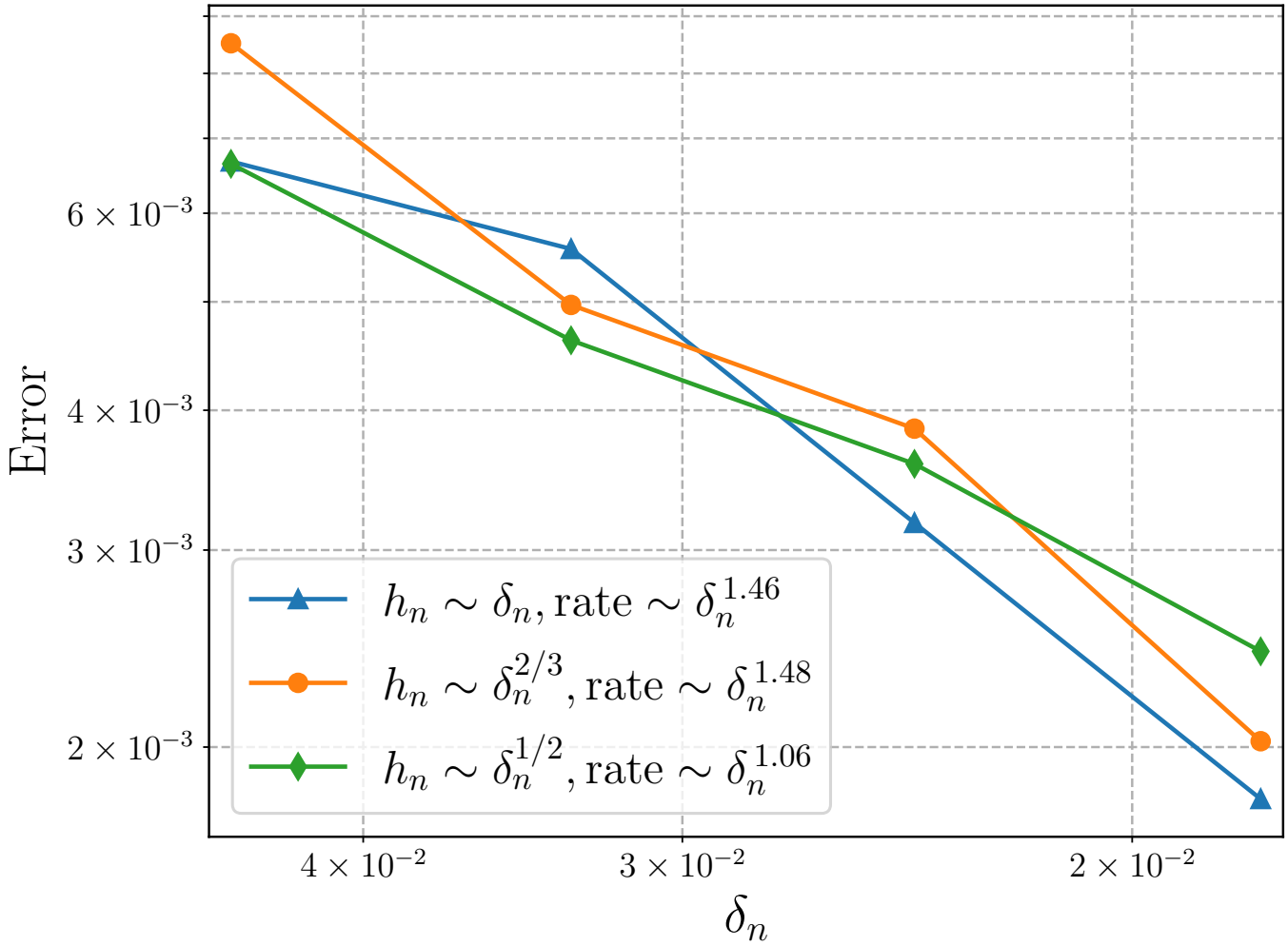


Figure: Experimental rates for singular weights on star domain.

What is Poisson learning?

Poisson Learning

Poisson Learning

Find $\mathbf{u}_n : \Omega_n \to \mathbb{R}$ such that

$$-\mathcal{L}_2^n \mathbf{u}_n = \sum_{x \in \mathcal{O}} (g(x) - \overline{g}) \delta_x,$$

subject to $\sum_{x \in \Omega_n} d_n(x) u(x) = 0$.

Poisson Learning

Poisson Learning

Find $\mathbf{u}_n : \Omega_n \to \mathbb{R}$ such that

$$-\mathcal{L}_2^n \mathbf{u}_n = \sum_{x \in \mathcal{O}} (g(x) - \overline{g}) \delta_x,$$

subject to $\sum_{x \in \Omega_n} d_n(x) u(x) = 0$.

Continuum Problem

Find $u \in W^{1,1}(\Omega)$, distributional solution of

$$-\operatorname{div}(\rho^2 \nabla u) = \sum_{x \in \mathcal{O}} (g(x) - \overline{g}) \delta_x,$$

subject to $\frac{\partial u}{\partial \nu} = 0$ and $\int_\Omega u \rho^2 \, \mathrm{d}x = 0$.

Poisson Learning

Poisson Learning

Find $\mathbf{u}_n : \Omega_n \to \mathbb{R}$ such that

$$-\mathcal{L}_2^n \mathbf{u}_n = \sum_{x \in \mathcal{O}} (g(x) - \overline{g}) \delta_x,$$

subject to $\sum_{x \in \Omega_n} d_n(x) u(x) = 0$.

Continuum Problem

Find $u \in W^{1,1}(\Omega)$, distributional solution of

$$-\operatorname{div}(\rho^2 \nabla u) = \sum_{x \in \mathcal{O}} (g(x) - \overline{g}) \delta_x,$$

subject to $\frac{\partial u}{\partial \nu} = 0$ and $\int_\Omega u \rho^2 \, \mathrm{d}x = 0$.

Here,

$$d_n(x) := \sum_{y \in \Omega_n} \omega_n(x, y), \qquad \overline{g} := \frac{1}{|\mathcal{O}|} \sum_{x \in \mathcal{O}} g(x),$$

and we assume that $\Omega_n = \{x_i\}_{i=1}^n$ with *i.i.d.* samples x_i with $\mathrm{Law}(x_i) = \rho$ and that $\mathcal{O}$ is a finite set.

Poisson Learning

Poisson Learning

Find $\mathbf{u}_n : \Omega_n \to \mathbb{R}$ such that

$$-\mathcal{L}_2^n \mathbf{u}_n = \sum_{x \in \mathcal{O}} (g(x) - \overline{g}) \delta_x,$$

subject to $\sum_{x \in \Omega_n} d_n(x) u(x) = 0$.

Continuum Problem

Find $u \in W^{1,1}(\Omega)$, distributional solution of

$$-\operatorname{div}(\rho^2 \nabla u) = \sum_{x \in \mathcal{O}} (g(x) - \overline{g}) \delta_x,$$

subject to $\frac{\partial u}{\partial \nu} = 0$ and $\int_\Omega u \rho^2 \, \mathrm{d}x = 0$.

Remark:

Poisson Learning

Poisson Learning

Find $\mathbf{u}_n : \Omega_n \to \mathbb{R}$ such that

$$-\mathcal{L}_2^n \mathbf{u}_n = \sum_{x \in \mathcal{O}} (g(x) - \overline{g}) \delta_x,$$

subject to $\sum_{x \in \Omega_n} d_n(x) u(x) = 0$.

Continuum Problem

Find $u \in W^{1,1}(\Omega)$, distributional solution of

$$-\operatorname{div}(\rho^2 \nabla u) = \sum_{x \in \mathcal{O}} (g(x) - \overline{g}) \delta_x,$$

subject to $\frac{\partial u}{\partial \nu} = 0$ and $\int_\Omega u \rho^2 \, \mathrm{d}x = 0$.

Remark:

- Proposed in [Cal+20], superior experimental results

Poisson Learning

Poisson Learning

Find $\mathbf{u}_n : \Omega_n \to \mathbb{R}$ such that

$$-\mathcal{L}_2^n \mathbf{u}_n = \sum_{x \in \mathcal{O}} (g(x) - \overline{g}) \delta_x,$$

subject to $\sum_{x \in \Omega_n} d_n(x) u(x) = 0$.

Continuum Problem

Find $u \in W^{1,1}(\Omega)$, distributional solution of

$$-\operatorname{div}(\rho^2 \nabla u) = \sum_{x \in \mathcal{O}} (g(x) - \overline{g}) \delta_x,$$

subject to $\frac{\partial u}{\partial \nu} = 0$ and $\int_\Omega u \rho^2 \, \mathrm{d}x = 0$.

Remark:

- Proposed in [Cal+20], superior experimental results
- No variational interpretation of continuum problem $\implies$ continuum limit requires PDE techniques

Poisson Learning

Poisson Learning

Find $\mathbf{u}_n : \Omega_n \to \mathbb{R}$ such that

$$-\mathcal{L}_2^n \mathbf{u}_n = \sum_{x \in \mathcal{O}} (g(x) - \overline{g}) \delta_x,$$

subject to $\sum_{x \in \Omega_n} d_n(x) u(x) = 0$.

Continuum Problem

Find $u \in W^{1,1}(\Omega)$, distributional solution of

$$-\operatorname{div}(\rho^2 \nabla u) = \sum_{x \in \mathcal{O}} (g(x) - \overline{g}) \delta_x,$$

subject to $\frac{\partial u}{\partial \nu} = 0$ and $\int_\Omega u \rho^2 \, \mathrm{d}x = 0$.

Remark:

- Proposed in [Cal+20], superior experimental results
- No variational interpretation of continuum problem $\implies$ continuum limit requires PDE techniques

Aims: convergence rates

Formal Main Result

Theorem (Formal, [B+24] (forthcoming))

Under reasonable assumptions and for

$$1 \gg \varepsilon_n \gg \begin{cases} \left(\frac{\log n}{n}\right)^{\frac{1}{3d-2}} & \text{if } d \geq 4 \\ \left(\frac{\log n}{n}\right)^{\frac{1}{2d+2}} & \text{if } d < 4 \end{cases}$$

Formal Main Result

Theorem (Formal, [B+24] (forthcoming))

Under reasonable assumptions and for

$$1 \gg \varepsilon_n \gg \begin{cases} \left(\frac{\log n}{n}\right)^{\frac{1}{3d-2}} & \text{if } d \geq 4 \\ \left(\frac{\log n}{n}\right)^{\frac{1}{2d+2}} & \text{if } d < 4 \end{cases}$$

it holds with high probability that

$$\|\mathbf{u}_n - u\|_{\ell^1(\Omega_n)} \lesssim \varepsilon_n^{\frac{1}{d+2}}.$$

Formal Main Result

Theorem (Formal, [B+24] (forthcoming))

Under reasonable assumptions and for

$$1 \gg \varepsilon_n \gg \begin{cases} \left(\frac{\log n}{n}\right)^{\frac{1}{3d-2}} & \text{if } d \geq 4 \\ \left(\frac{\log n}{n}\right)^{\frac{1}{2d+2}} & \text{if } d < 4 \end{cases}$$

it holds with high probability that

$$\|\mathbf{u}_n - u\|_{\ell^1(\Omega_n)} \lesssim \varepsilon_n^{\frac{1}{d+2}}.$$

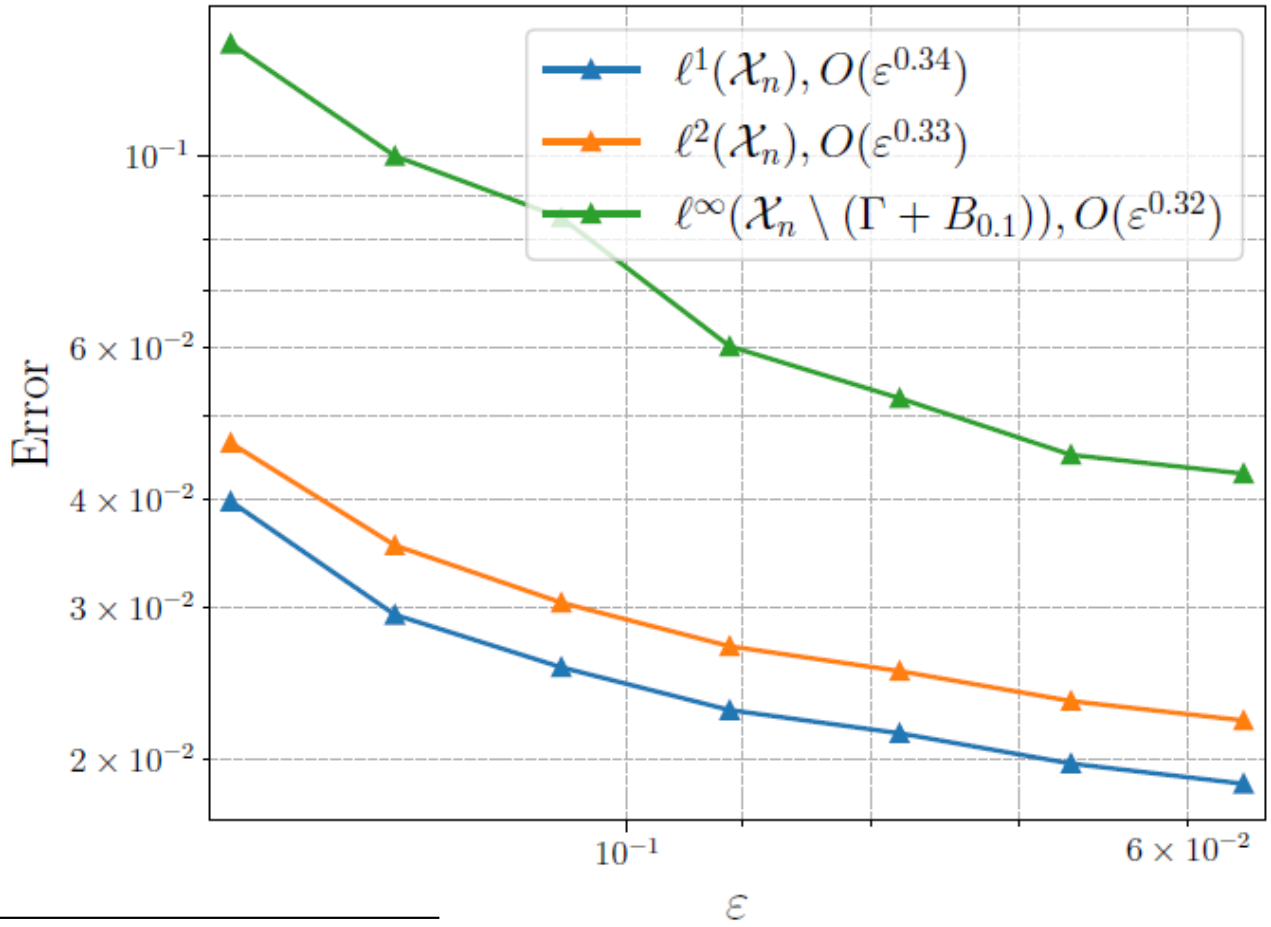
If $\rho \equiv const$ *and* $\varepsilon_n \gg \left(\frac{\log n}{n}\right)^{\frac{1}{3d-2(d+2)/(d+4)}}$ *this can be improved to*

$$\|\mathbf{u}_n - u\|_{\ell^1(\Omega_n)} \lesssim \varepsilon_n^{\frac{2}{d+4}}.$$

Numerical Results

We consider the problem $\Delta u = \Delta_z - \delta_{-z}$ for $z \in \Omega = B(0,1) \subset \mathbb{R}^2$ which has an explicit solution.

We choose $\varepsilon = 2\left(\frac{\log n}{n}\right)^{\frac{1}{d+2}}$ for $n = 2^{10}, \ldots, 2^{16}$ and get the following rates[1]:



[1] Figure courtesy of Jeff Calder

Three Levels of Approximation

Three Levels of Approximation

Continuum mollification

Replace continuum data $\sum_{x\in\mathcal{O}}(g(x)-\overline{g})\delta_x$ by $\sum_{x\in\mathcal{O}}(g(x)-\overline{g})\varphi_x$ with $\operatorname{supp}\varphi_x \subset B(x,R)$ obtain quantitative L^1-estimates in R.

Three Levels of Approximation

Continuum mollification

Replace continuum data $\sum_{x\in\mathcal{O}}(g(x)-\overline{g})\delta_x$ by $\sum_{x\in\mathcal{O}}(g(x)-\overline{g})\varphi_x$ with $\operatorname{supp}\varphi_x \subset B(x,R)$ obtain quantitative L^1-estimates in R.

Discrete mollification

Mollify discrete data $\sum_{x\in\mathcal{O}}(g(x)-\overline{g})\delta_x$ with k steps of the graph heat equation and obtain quantitative estimates in k and ε.

Three Levels of Approximation

Continuum mollification

Replace continuum data $\sum_{x\in\mathcal{O}}(g(x)-\overline{g})\delta_x$ by $\sum_{x\in\mathcal{O}}(g(x)-\overline{g})\varphi_x$ with $\operatorname{supp}\varphi_x \subset B(x,R)$ obtain quantitative L^1-estimates in R.

Discrete mollification

Mollify discrete data $\sum_{x\in\mathcal{O}}(g(x)-\overline{g})\delta_x$ with k steps of the graph heat equation and obtain quantitative estimates in k and ε.

Regular continuum limit

Prove discrete-to-continuum convergence rates for bounded right hand sides using variational methods (strong convexity).

Three Levels of Approximation

Continuum mollification

Replace continuum data $\sum_{x\in\mathcal{O}}(g(x)-\overline{g})\delta_x$ by $\sum_{x\in\mathcal{O}}(g(x)-\overline{g})\varphi_x$ with $\operatorname{supp}\varphi_x \subset B(x,R)$ obtain quantitative L^1-estimates in R.

Discrete mollification

Mollify discrete data $\sum_{x\in\mathcal{O}}(g(x)-\overline{g})\delta_x$ with k steps of the graph heat equation and obtain quantitative estimates in k and ε.

Regular continuum limit

Prove discrete-to-continuum convergence rates for bounded right hand sides using variational methods (strong convexity).

NB: Keeping track of all constants (which blow up) and optimizing over all parameters we obtain the final rate.

Continuum Mollification

We regularize the continuum equation by approximating the Dirac deltas:

$$-\operatorname{div}(\rho^2 \nabla u) = \sum_{x \in \Gamma} a_x \delta_x \qquad \text{and} \qquad -\operatorname{div}(\rho^2 \nabla u_R) = \sum_{x \in \Gamma} a_x \varphi_x,$$

where $\operatorname{supp} \varphi_x \subset B(x, R)$, $\varphi_x \geq 0$, and $\int_{B(x,R)} \varphi_x(y) \, \mathrm{d}y = 1$.

Continuum Mollification

We regularize the continuum equation by approximating the Dirac deltas:

$$-\operatorname{div}(\rho^2 \nabla u) = \sum_{x \in \Gamma} a_x \delta_x \qquad \text{and} \qquad -\operatorname{div}(\rho^2 \nabla u_R) = \sum_{x \in \Gamma} a_x \varphi_x,$$

where $\operatorname{supp} \varphi_x \subset B(x, R)$, $\varphi_x \geq 0$, and $\int_{B(x,R)} \varphi_x(y) \, \mathrm{d}y = 1$.

Theorem ([B+24])

If $\operatorname{dist}(\Gamma, \partial\Omega) > R$ *then*

$$\|u - u_R\|_{L^1(\Omega)} \lesssim \sum_{x \in \Gamma} |a_x| \, R.$$

Continuum Mollification

We regularize the continuum equation by approximating the Dirac deltas:

$$-\operatorname{div}(\rho^2 \nabla u) = \sum_{x \in \Gamma} a_x \delta_x \qquad \text{and} \qquad -\operatorname{div}(\rho^2 \nabla u_R) = \sum_{x \in \Gamma} a_x \varphi_x,$$

where $\operatorname{supp} \varphi_x \subset B(x, R)$, $\varphi_x \geq 0$, and $\int_{B(x,R)} \varphi_x(y)\, \mathrm{d}y = 1$.

Theorem ([B+24])

If $\operatorname{dist}(\Gamma, \partial\Omega) > R$ *then*

$$\|u - u_R\|_{L^1(\Omega)} \lesssim \sum_{x \in \Gamma} |a_x|\, R.$$

If in addition $\rho \equiv const$ *and* $\varphi_x(y) = R^{-d} \psi\left(|y - x| / R\right)$, *then*

$$\|u - u_R\|_{L^1(\Omega)} \lesssim \sum_{x \in \Gamma} |a_x|\, R^2.$$

Discrete Mollification

We define the random walk graph Laplacian:

$$\mathcal{L}_{rw}\mathbf{u}_n(x) := \frac{1}{\varepsilon_n^2 d_n(x)}\mathcal{L}_2^n\mathbf{u}_n(x) = \frac{1}{\varepsilon_n^2}\left(\frac{1}{d_n(x)}\sum_{y\in\Omega_n}\omega_n(x,y)\mathbf{u}_n(y) - \mathbf{u}_n(x)\right).$$

Discrete Mollification

We define the random walk graph Laplacian:

$$\mathcal{L}_{rw}\mathbf{u}_n(x) := \frac{1}{\varepsilon_n^2 d_n(x)}\mathcal{L}_2^n\mathbf{u}_n(x) = \frac{1}{\varepsilon_n^2}\left(\frac{1}{d_n(x)}\sum_{y\in\Omega_n}\omega_n(x,y)\mathbf{u}_n(y) - \mathbf{u}_n(x)\right).$$

The graph heat kernel $\mathcal{H}_k^x$ is the solution of the heat equation, starting with δ_x:

$$\mathcal{H}_{k+1}^x = \mathcal{H}_k^x - \varepsilon_n^2\mathcal{L}_{rw}^T\mathcal{H}_k^x, \qquad \mathcal{H}_0^x = n\delta_x.$$

We define the convolution $\mathcal{H}_k * \mathbf{u}_n(x) := \frac{1}{n}\sum_{y\in\Omega_n}\mathcal{H}_k^x(y)\mathbf{u}_n(y)$.

Discrete Mollification

We define the random walk graph Laplacian:

$$\mathcal{L}_{rw}\mathbf{u}_n(x) := \frac{1}{\varepsilon_n^2 d_n(x)}\mathcal{L}_2^n\mathbf{u}_n(x) = \frac{1}{\varepsilon_n^2}\left(\frac{1}{d_n(x)}\sum_{y\in\Omega_n}\omega_n(x,y)\mathbf{u}_n(y) - \mathbf{u}_n(x)\right).$$

The graph heat kernel $\mathcal{H}_k^x$ is the solution of the heat equation, starting with δ_x:

$$\mathcal{H}_{k+1}^x = \mathcal{H}_k^x - \varepsilon_n^2\mathcal{L}_{rw}^T\mathcal{H}_k^x, \qquad \mathcal{H}_0^x = n\delta_x.$$

We define the convolution $\mathcal{H}_k * \mathbf{u}_n(x) := \frac{1}{n}\sum_{y\in\Omega_n}\mathcal{H}_k^x(y)\mathbf{u}_n(y)$.

Theorem ([B+24])

It holds $\mathcal{H}_k * (\mathcal{L}_{rw}\mathbf{u}_n) = \mathcal{L}_{rw}(\mathcal{H}_k * \mathbf{u}_n)$.

Estimates on Mollified Problem

Let $\mathbf{u}_n$ and $\mathbf{u}_n^{(k)} := \mathcal{H}_k^{\cdot} * \mathbf{u}_n$ solve

$$-\mathcal{L}\mathbf{u}_n = \sum_{x\in\mathcal{O}} a_x \delta_x \qquad \text{and} \qquad -\mathcal{L}\mathbf{u}_n^{(k)} = \sum_{x\in\mathcal{O}} a_x \mathcal{H}_k^x.$$

Then it holds

$$\mathbf{u}_n - \mathbf{u}_n^{(k)} = \frac{n\varepsilon_n^2}{d_n} \sum_{x\in\mathcal{O}} a_x \sum_{j=0}^{k-1} \mathcal{H}_j^x.$$

Estimates on Mollified Problem

Let $\mathbf{u}_n$ and $\mathbf{u}_n^{(k)} := \mathcal{H}_k^{\cdot} * \mathbf{u}_n$ solve

$$-\mathcal{L}\mathbf{u}_n = \sum_{x\in\mathcal{O}} a_x \delta_x \qquad \text{and} \qquad -\mathcal{L}\mathbf{u}_n^{(k)} = \sum_{x\in\mathcal{O}} a_x \mathcal{H}_k^x.$$

Then it holds

$$\mathbf{u}_n - \mathbf{u}_n^{(k)} = \frac{n\varepsilon_n^2}{d_n} \sum_{x\in\mathcal{O}} a_x \sum_{j=0}^{k-1} \mathcal{H}_j^x.$$

Using that $d_n \sim n$ and $\left\|\mathcal{H}_j^x\right\|_{\ell^1(\Omega_n)} = 1$ we get the estimate

$$\left\|\mathbf{u}_n - \mathbf{u}_n^{(k)}\right\|_{\ell^1(\Omega_n)} \lesssim k\varepsilon_n^2 \sum_{x\in\mathcal{O}} |a_x|$$

and will need to choose $1 \ll k \ll \frac{1}{\varepsilon_n^2}$.

Continuum Limit for Bounded Data

We consider Poisson equations with bounded right hand side:

$$-\operatorname{div}(\rho^2 \nabla u) = f \qquad \text{and} \qquad -\mathcal{L}_2^n \mathbf{u}_n = f_n$$

Continuum Limit for Bounded Data

We consider Poisson equations with bounded right hand side:

$$-\operatorname{div}(\rho^2 \nabla u) = f \qquad \text{and} \qquad -\mathcal{L}_2^n \mathbf{u}_n = f_n$$

Theorem ([B+24])

For all $R, \lambda_1, \lambda_2, \varepsilon_n, \delta > 0$ sufficiently small, and $q > \frac{d}{2}$ we have with high probability:

$$\begin{aligned}\|u - \mathbf{u}_n\|_{H^1(\mathcal{X}_n)}^2 \lesssim & \left(\|f_n - f\|_{\ell^1(\Omega_n)} + \|\operatorname{osc}_{\Omega \cap B(\cdot,\delta)} f\|_{L^1(\Omega)}\right)\left(\|f\|_{L^q(\Omega)} + \|f_n\|_{\ell^q(\Omega_n)}\right) \\ & + \|f\|_{L^\infty(\Omega)}^2 \lambda_1 + \|f\|_{L^\infty(\partial_{4\varepsilon_n}\Omega)}^2 \varepsilon_n + \|f_n\|_{\ell^2(\Omega_n)} \|f_n\|_{\ell^2(\Omega_n \cap \partial_{2R}\Omega)} \\ & + \left(\|f_n\|_{\ell^2(\Omega_n)}^2 + \|f\|_{L^2(\Omega)}^2\right)\left(\frac{\delta}{\varepsilon_n} + \varepsilon_n + \lambda_1^2 + \lambda_2\right).\end{aligned}$$

Continuum Limit for Bounded Data

We consider Poisson equations with bounded right hand side:

$$-\operatorname{div}(\rho^2 \nabla u) = f \qquad \text{and} \qquad -\mathcal{L}_2^n \mathbf{u}_n = f_n$$

Theorem ([B+24])

For all $R, \lambda_1, \lambda_2, \varepsilon_n, \delta > 0$ sufficiently small, and $q > \frac{d}{2}$ we have with high probability:

$$\begin{aligned}\|u - \mathbf{u}_n\|_{H^1(\mathcal{X}_n)}^2 \lesssim & \left(\|f_n - f\|_{\ell^1(\Omega_n)} + \|\operatorname{osc}_{\Omega \cap B(\cdot,\delta)} f\|_{L^1(\Omega)}\right)\left(\|f\|_{L^q(\Omega)} + \|f_n\|_{\ell^q(\Omega_n)}\right) \\ & + \|f\|_{L^\infty(\Omega)}^2 \lambda_1 + \|f\|_{L^\infty(\partial_{4\varepsilon_n}\Omega)}^2 \varepsilon_n + \|f_n\|_{\ell^2(\Omega_n)} \|f_n\|_{\ell^2(\Omega_n \cap \partial_{2R}\Omega)} \\ & + \left(\|f_n\|_{\ell^2(\Omega_n)}^2 + \|f\|_{L^2(\Omega)}^2\right)\left(\frac{\delta}{\varepsilon_n} + \varepsilon_n + \lambda_1^2 + \lambda_2\right).\end{aligned}$$

We choose $f_n = \sum_{x \in \mathcal{O}} a_x \mathcal{H}_k^x$ and prove discrete-to-continuum convergence rates towards a k-fold convolution $f := \sum_{x \in \mathcal{O}} a_x \rho(x)^{-1} \mathcal{M}_\varepsilon^k(\delta_x)$.

Conclusion

What we discussed today:

Conclusion

What we discussed today:

Part 1 (Lipschitz Learning):

- Lipschitz learning is asymptotically well-posed
- Uniform discrete-to-continuum convergence rates down to smallest length scales
- Novel "homogenized" proof technique

Conclusion

What we discussed today:

Part 1 (Lipschitz Learning):

- Lipschitz learning is asymptotically well-posed
- Uniform discrete-to-continuum convergence rates down to smallest length scales
- Novel "homogenized" proof technique

Part 2 (Poisson Learning):

- Poisson learning is asymptotically well-posed
- L^1-discrete-to-continuum convergence rates down for relatively large length scales
- Three-scale proof technique

Conclusion

What we discussed today:

Part 1 (Lipschitz Learning):

- Lipschitz learning is asymptotically well-posed
- Uniform discrete-to-continuum convergence rates down to smallest length scales
- Novel "homogenized" proof technique

Part 2 (Poisson Learning):

- Poisson learning is asymptotically well-posed
- L^1-discrete-to-continuum convergence rates down for relatively large length scales
- Three-scale proof technique

Future work:

- Convergence rates for Lipschitz Learning with density-drift
- Percolation rates for less rigid assumptions
- Sharpness of Poisson learning rates

Thank you for your attention!

Questions?

L. B, J. Calder, and T. Roith. "Uniform convergence rates for Lipschitz learning on graphs". In: *IMA Journal of Numerical Analysis* 43.4 (2023), pp. 2445–2495

L. B, J. Calder, and T. Roith. "Ratio convergence rates for Euclidean first-passage percolation: Applications to the graph infinity Laplacian". In: *Annals of Applied Probability* (2024). In press

L. B, J. Calder, F. Hoffmann, K. Houssou, M. Mihailescu, and A. Yuan. *Convergence rates for Poisson Learning to a Poisson equation with measure data*. 2024. In preparation

References I

[ZGL03] X. Zhu, Z. Ghahramani, and J. D. Lafferty. "Semi-supervised learning using gaussian fields and harmonic functions". In: *Proceedings of the 20th International conference on Machine learning (ICML-03)*. 2003, pp. 912–919.

[AS10] S. N. Armstrong and C. K. Smart. "An easy proof of Jensen's theorem on the uniqueness of infinity harmonic functions". In: *Calculus of Variations and Partial Differential Equations* 37.3 (2010), pp. 381–384.

[Sma10] C. K. Smart. "On the infinity Laplacian and Hrushovski's fusion". PhD thesis. UC Berkeley, 2010.

[GS15] N. García Trillos and D. Slepčev. "Continuum Limit of Total Variation on Point Clouds". In: *Archive for Rational Mechanics and Analysis* 220.1 (2015), pp. 193–241. DOI: 10.1007/s00205-015-0929-z.

[El +16] A. El Alaoui, X. Cheng, A. Ramdas, M. J. Wainwright, and M. I. Jordan. "Asymptotic behavior of ℓ_p-based Laplacian regularization in semi-supervised learning". In: *Conference on Learning Theory*. PMLR. 2016, pp. 879–906.

References II

[Cal18] J. Calder. "The game theoretic p-Laplacian and semi-supervised learning with few labels". In: *Nonlinearity* 32.1 (2018), p. 301.

[Cal19] J. Calder. "Consistency of Lipschitz learning with infinite unlabeled data and finite labeled data". In: *SIAM Journal on Mathematics of Data Science* 1.4 (2019), pp. 780–812.

[ST19] D. Slepčev and M. Thorpe. "Analysis of p-Laplacian regularization in semisupervised learning". In: *SIAM Journal on Mathematical Analysis* 51.3 (2019), pp. 2085–2120.

[Cal+20] J. Calder, B. Cook, M. Thorpe, and D. Slepcev. "Poisson learning: Graph based semi-supervised learning at very low label rates". In: *International Conference on Machine Learning*. PMLR. 2020, pp. 1306–1316.

[CS20] J. Calder and D. Slepčev. "Properly-weighted graph Laplacian for semi-supervised learning". In: *Applied mathematics & optimization* 82 (2020), pp. 1111–1159.

References III

[BCR23] L. B, J. Calder, and T. Roith. "Uniform convergence rates for Lipschitz learning on graphs". In: *IMA Journal of Numerical Analysis* 43.4 (2023), pp. 2445–2495.

[CST23] J. Calder, D. Slepčev, and M. Thorpe. "Rates of convergence for Laplacian semi-supervised learning with low labeling rates". In: *Research in the Mathematical Sciences* 10.1 (2023), p. 10.

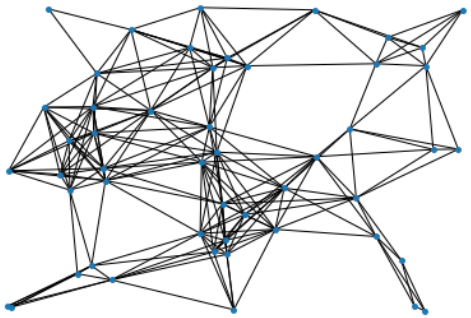
[RB23] T. Roith and L. B. "Continuum limit of Lipschitz learning on graphs". In: *Foundations of Computational Mathematics* 23.2 (2023), pp. 393–431.

[B+24] L. B, J. Calder, F. Hoffmann, K. Houssou, M. Mihailescu, and A. Yuan. *Convergence rates for Poisson Learning to a Poisson equation with measure data*. 2024. In preparation.

[BCR24] L. B, J. Calder, and T. Roith. "Ratio convergence rates for Euclidean first-passage percolation: Applications to the graph infinity Laplacian". In: *Annals of Applied Probability* (2024). In press.

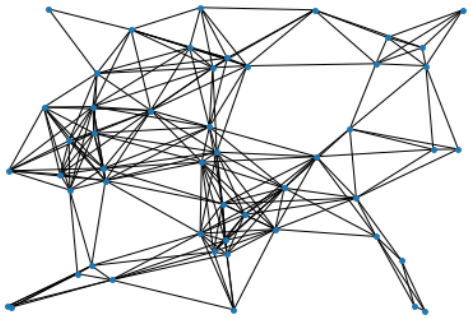
Length Scale Restrictions

Length Scale Restrictions

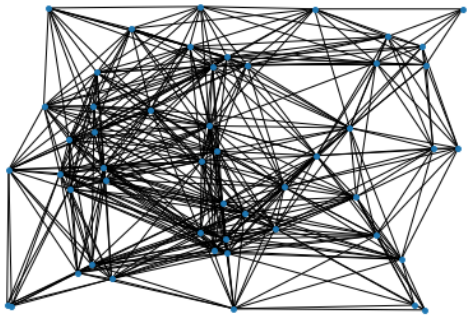


Sparse graph: bandwidth $\gg$ connectivity threshold,
Lipschitz learning (today, [BCR23])

Length Scale Restrictions

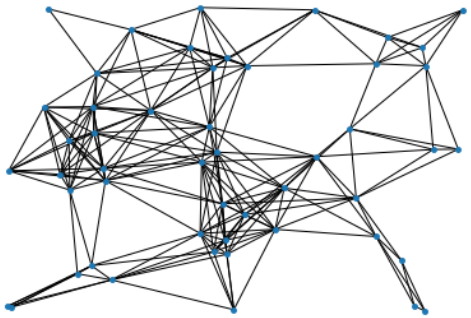


Sparse graph: bandwidth $\gg$ connectivity threshold,
Lipschitz learning (today, [BCR23])

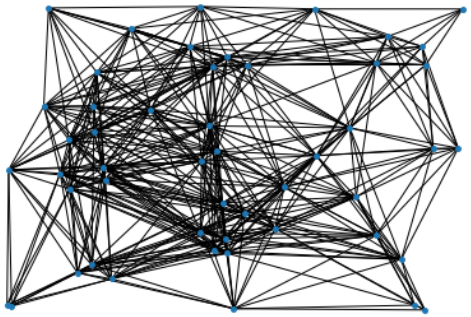


Medium graph: bandwidth $\gg$ (connectivity threshold)$^{\frac{2}{3}}$,
Lipschitz learning (best so far, [Cal19])

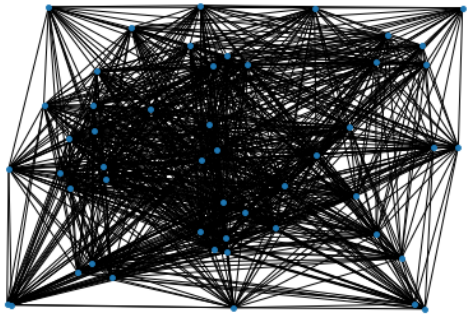
Length Scale Restrictions



Sparse graph: bandwidth $\gg$ connectivity threshold,
Lipschitz learning (today, [BCR23])



Medium graph: bandwidth $\gg$ (connectivity threshold)$^{\frac{2}{3}}$,
Lipschitz learning (best so far, [Cal19])



Dense graph: bandwidth $\gg$ (connectivity threshold)$^{\frac{1}{4}}$,
Laplacian learning [Cal18]

Experimental Convergence Rates

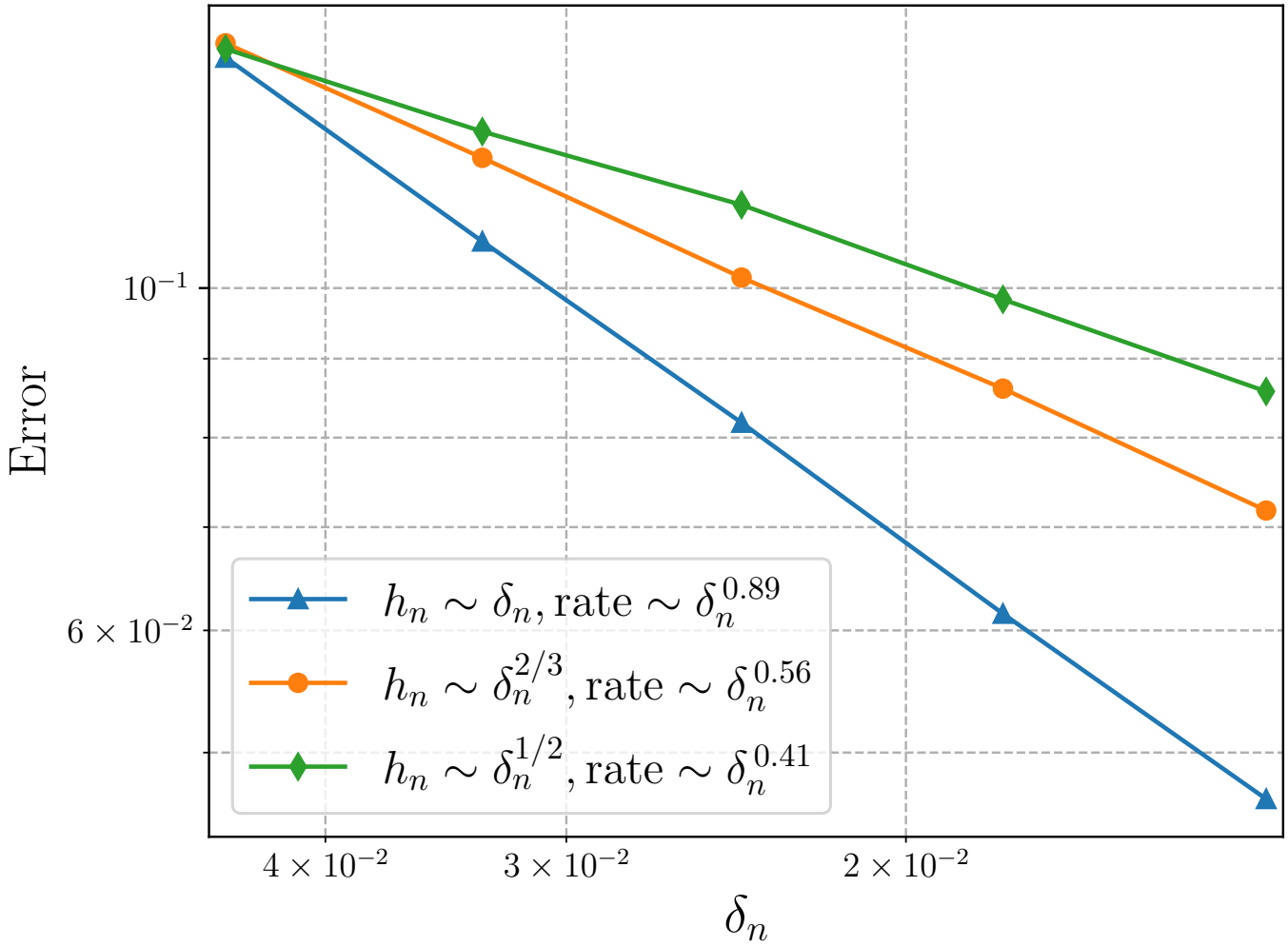


Figure: Experimental rates for constant weights on star domain.

Experimental Convergence Rates

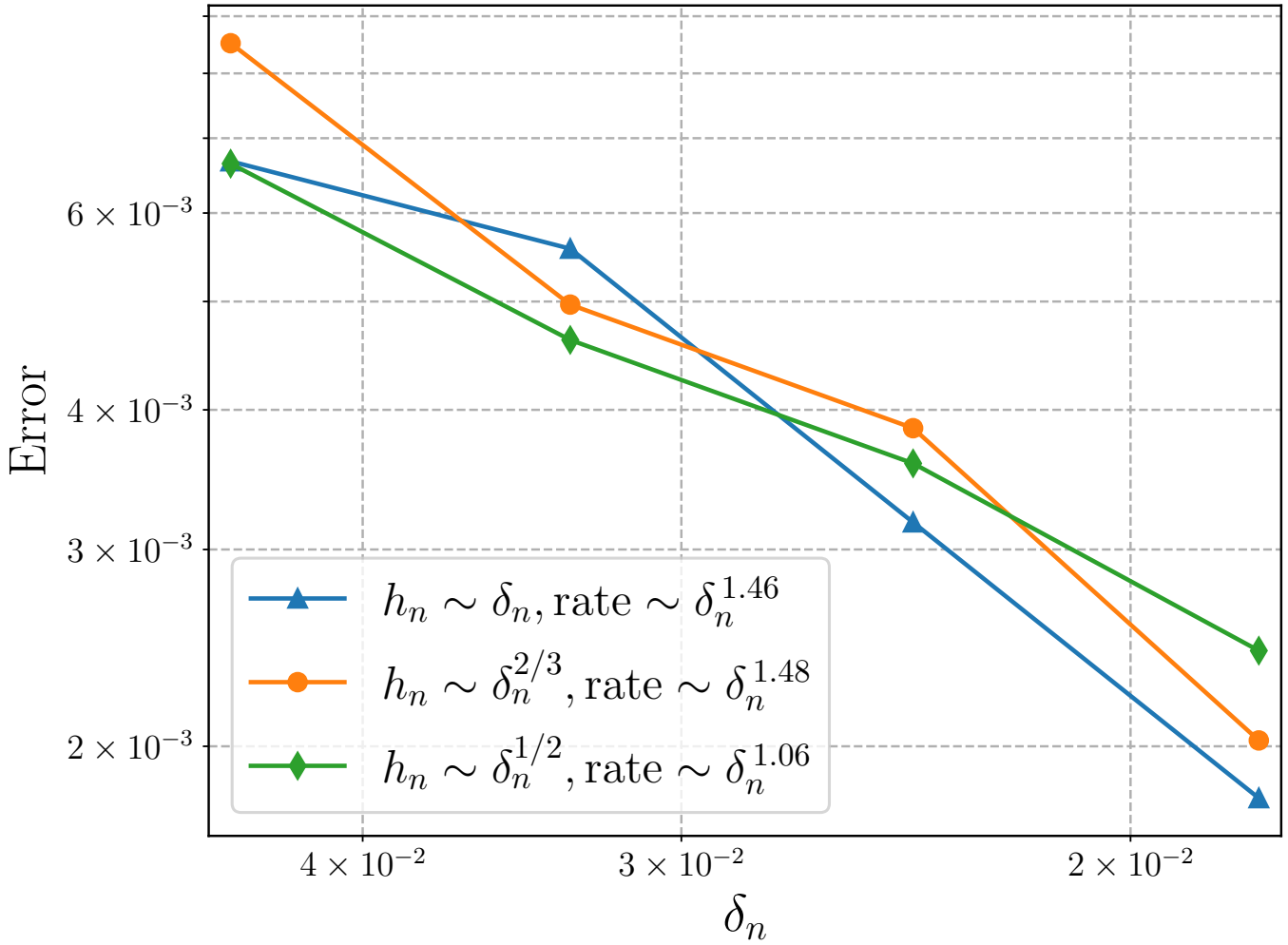


Figure: Experimental rates for singular weights on star domain.

Experimental Convergence Rates

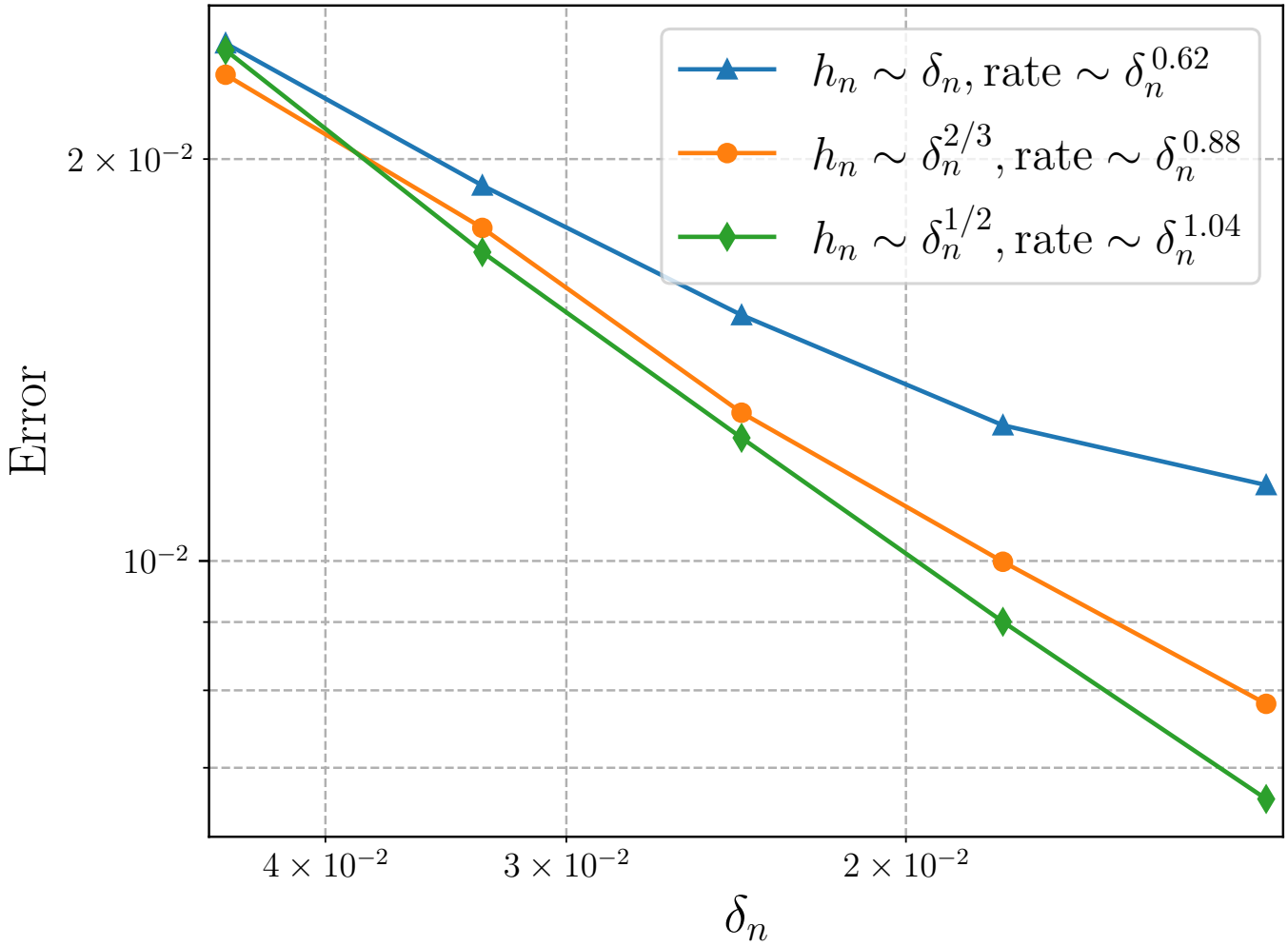
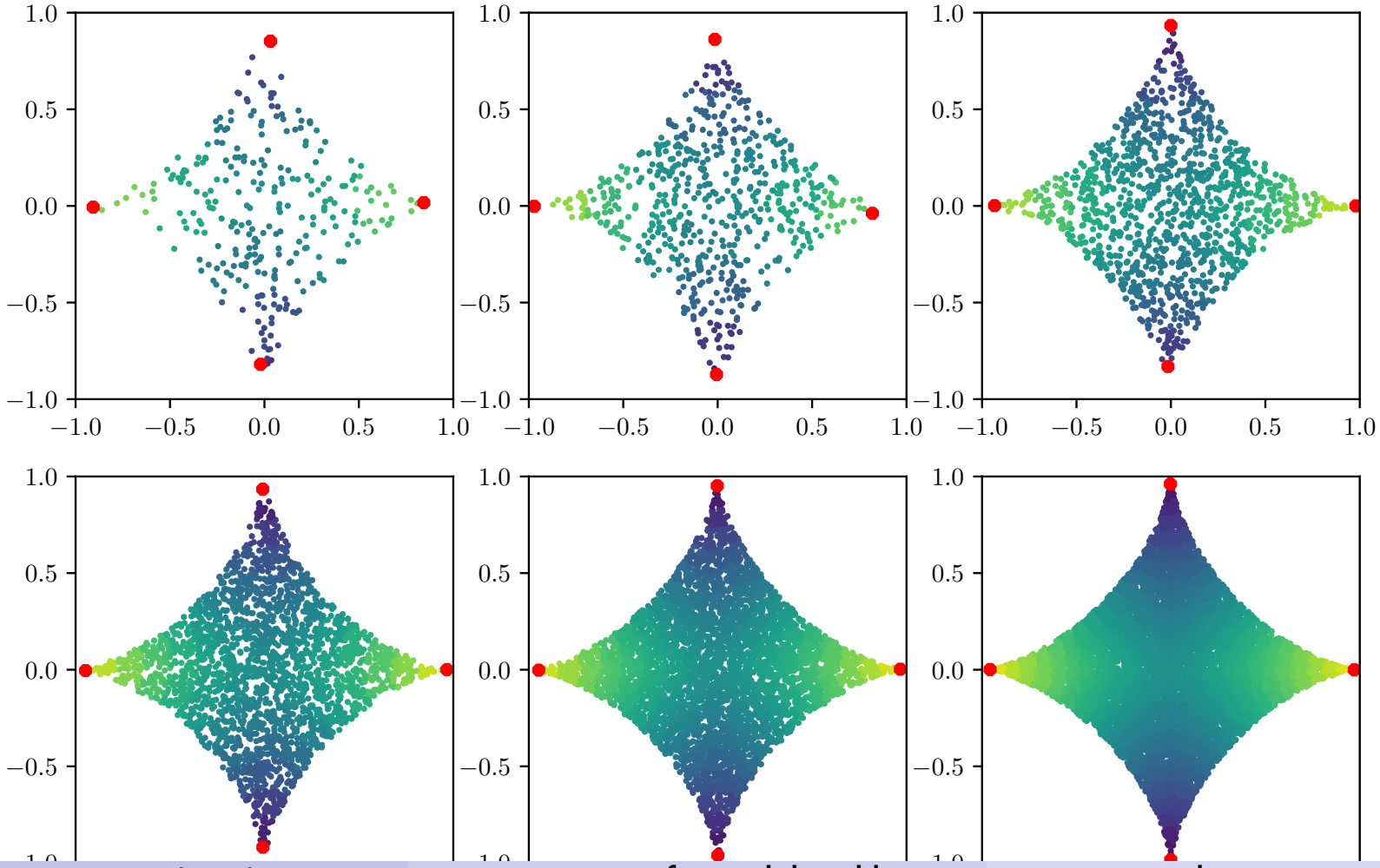


Figure: Experimental rates for constant weights on square domain.

Experimental Convergence Rates

Figure: Experimental rates for weights on domain.



Graph Functional

Graph Lipschitz Learning

Find $u : \Omega_n \to \mathbb{R}$ such that

$$u \in \arg\min \max_{x,y \in \Omega_n} \omega_n(x,y)|u(x) - u(y)|,$$

and $u = g$ on $\mathcal{O}_n$.

Lipschitz Learning

Find $u \in W^{1,\infty}(\Omega)$ such that

$$u \in \arg\min \operatorname{ess\,sup}_{x \in \Omega} |\nabla u(x)|,$$

and $u = g$ in $\mathcal{O}$.

- The graph functional has the form

$$E_n(u) = \frac{1}{\varepsilon_n} \max_{x,y \in \Omega_n} \eta_{\varepsilon_n}(|x-y|)\,|u(x) - u(y)|\,.$$

Graph Functional

Graph Lipschitz Learning

Find $u : \Omega_n \to \mathbb{R}$ such that

$$u \in \arg\min \max_{x,y \in \Omega_n} \omega_n(x,y)|u(x) - u(y)|,$$

and $u = g$ on $\mathcal{O}_n$.

Lipschitz Learning

Find $u \in W^{1,\infty}(\Omega)$ such that

$$u \in \arg\min \operatorname{ess\,sup}_{x \in \Omega} |\nabla u(x)|,$$

and $u = g$ in $\mathcal{O}$.

- The graph functional has the form

$$E_n(u) = \frac{1}{\varepsilon_n} \max_{x,y \in \Omega_n} \eta_{\varepsilon_n}(|x-y|)\,|u(x) - u(y)|\,.$$

- The respective continuum functional for $u \in W^{1,\infty}(\Omega)$ reads

$$\mathcal{E}(u) := \operatorname{ess\,sup}_{x \in \Omega} |\nabla u(x)|\,.$$

Graph Functional

Graph Lipschitz Learning

Find $u : \Omega_n \to \mathbb{R}$ such that

$$u \in \arg\min \max_{x,y \in \Omega_n} \omega_n(x,y)|u(x) - u(y)|,$$

and $u = g$ on $\mathcal{O}_n$.

Lipschitz Learning

Find $u \in W^{1,\infty}(\Omega)$ such that

$$u \in \arg\min \operatorname{ess\,sup}_{x \in \Omega} |\nabla u(x)|,$$

and $u = g$ in $\mathcal{O}$.

- The graph functional has the form

$$E_n(u) = \frac{1}{\varepsilon_n} \max_{x,y \in \Omega_n} \eta_{\varepsilon_n}(|x-y|)\,|u(x) - u(y)|\,.$$

- The respective continuum functional for $u \in W^{1,\infty}(\Omega)$ reads

$$\mathcal{E}(u) := \operatorname{ess\,sup}_{x \in \Omega} |\nabla u(x)|\,.$$

- We want to prove Γ-convergence and compactness for convergence of minimizers.

Domain Regularity

The key estimate in our arguments is

$$|u(x) - u(y)| \leq \|\nabla u\|_{L^\infty} d_\Omega(x, y),$$

where $u : W^{1,\infty}(\Omega) \to \mathbb{R}$ and

$$d_\Omega(x, y) := \inf \{\mathrm{len}(\gamma) \ : \ \gamma : [0, 1] \to \Omega \text{ is a curve with } \gamma(0) = x, \gamma(1) = y\}$$

denotes the geodesic distance in Ω.

Domain Regularity

The key estimate in our arguments is

$$|u(x) - u(y)| \leq \|\nabla u\|_{L^\infty} d_\Omega(x, y),$$

where $u : W^{1,\infty}(\Omega) \to \mathbb{R}$ and

$$d_\Omega(x, y) := \inf \{\mathrm{len}(\gamma) \ : \ \gamma : [0, 1] \to \Omega \text{ is a curve with } \gamma(0) = x,\ \gamma(1) = y\}$$

denotes the geodesic distance in Ω.

Our results only hold true for "locally convex" domains Ω which satisfy

Local Convexity Condition

$$\lim_{\delta \searrow 0} \sup \left\{ \frac{d_\Omega(x, y)}{|x - y|} \ : \ x, y \in \Omega,\ |x - y| < \delta \right\} = 1.$$

Non-Local Auxiliary Functional

Similar to the proof in [GS15] it is convenient to first establish a convergence result for a non-local continuum functional.

Non-Local Auxiliary Functional

Similar to the proof in [GS15] it is convenient to first establish a convergence result for a non-local continuum functional.

Non-Local to Local Convergence [RB23]

Let Ω be locally convex. For $u \in L^\infty(\Omega)$ and $h > 0$ we define

$$\mathcal{E}_h(u) := \frac{1}{h} \operatorname*{ess\,sup}_{x,y \in \Omega} \{\eta_s(|x-y|)\,|u(x)-u(y)|\}\,.$$

For any null sequence ε_n we have that

$$\mathcal{E}_{\varepsilon_n} \xrightarrow{\Gamma} \sigma_\eta\, \mathcal{E}.$$

Non-Local Auxiliary Functional

Similar to the proof in [GS15] it is convenient to first establish a convergence result for a non-local continuum functional.

Non-Local to Local Convergence [RB23]

Let Ω be locally convex. For $u \in L^\infty(\Omega)$ and $h > 0$ we define

$$\mathcal{E}_h(u) := \frac{1}{h} \operatorname*{ess\,sup}_{x,y \in \Omega} \{\eta_s(|x-y|)\,|u(x)-u(y)|\}\,.$$

For any null sequence ε_n we have that

$$\mathcal{E}_{\varepsilon_n} \xrightarrow{\Gamma} \sigma_\eta\, \mathcal{E}.$$

- The limit functional is defined as

$$\mathcal{E}(u) := \operatorname*{ess\,sup}_{\Omega} |\nabla u|.$$

Non-Local Auxiliary Functional

Similar to the proof in [GS15] it is convenient to first establish a convergence result for a non-local continuum functional.

Non-Local to Local Convergence [RB23]

Let Ω be locally convex. For $u \in L^\infty(\Omega)$ and $h > 0$ we define

$$\mathcal{E}_h(u) := \frac{1}{h} \operatorname*{ess\,sup}_{x,y\in\Omega} \left\{ \eta_s(|x-y|)\, |u(x)-u(y)| \right\}.$$

For any null sequence ε_n we have that

$$\mathcal{E}_{\varepsilon_n} \xrightarrow{\Gamma} \sigma_\eta\, \mathcal{E}.$$

- The limit functional is defined as

$$\mathcal{E}(u) := \operatorname*{ess\,sup}_{\Omega} |\nabla u|.$$

- The value σ_η is defined as

$$\sigma_\eta := \operatorname*{ess\,sup}_{x>0} \left\{ \eta(x)\, |x| \right\}.$$

How do we establish Γ-convergence for a sequence of discrete functionals with varying domain, namely the space of functions on the graph Ω_n?

Extension of the Discrete Functional

- In our case we associate graph functions with piecewise constant L^∞ functions.

Extension of the Discrete Functional

- In our case we associate graph functions with piecewise constant L^∞ functions.
- Consider a closest-point projection $p_n : \Omega \to \Omega_n$ such that

$$p_n(x) \in \arg\min_{y \in \Omega_n} |x - y| \,.$$

Extension of the Discrete Functional

- In our case we associate graph functions with piecewise constant L^∞ functions.
- Consider a closest-point projection $p_n : \Omega \to \Omega_n$ such that

$$p_n(x) \in \arg\min_{y \in \Omega_n} |x - y| \,.$$

- We extend the functional E_n onto L^∞ by defining for $u \in L^\infty(\Omega)$:

$$E_n(u) = \begin{cases} \frac{1}{\varepsilon_n} \max_{x,y \in \Omega_n} \eta_{\varepsilon_n}(|x-y|)|\bar{u}(x) - \bar{u}(y)|, & \text{if } u = \bar{u} \circ p_n, \\ +\infty, & \text{else.} \end{cases}$$

Discrete to Continuum Convergence

We can also define the constrained functionals

$$E_{n,\mathrm{cons}}(u) = \begin{cases} E_n(u), & \text{if } u = g \text{ on } \mathcal{O}_n, \\ \infty, & \text{else,} \end{cases} \qquad \mathcal{E}_{\mathrm{cons}}(u) = \begin{cases} \mathcal{E}(u), & \text{if } u = g \text{ on } \mathcal{O}, \\ \infty, & \text{else} \end{cases}$$

and obtain

Discrete to Continuum Convergence

We can also define the constrained functionals

$$E_{n,\mathrm{cons}}(u) = \begin{cases} E_n(u), & \text{if } u = g \text{ on } \mathcal{O}_n, \\ \infty, & \text{else,} \end{cases} \qquad \mathcal{E}_{\mathrm{cons}}(u) = \begin{cases} \mathcal{E}(u), & \text{if } u = g \text{ on } \mathcal{O}, \\ \infty, & \text{else} \end{cases}$$

and obtain

Discrete to Continuum Convergence [RB23]

Let Ω be locally convex. For any null sequence ε_n such that

$$d_H(\Omega_n, \Omega) \ll \varepsilon_n,$$
$$d_H(\mathcal{O}_n, \mathcal{O}) \ll \varepsilon_n.$$

we have that

$$E_{n,\mathsf{cons}} \xrightarrow{\Gamma} \sigma_\eta \, \mathcal{E}_{\mathsf{cons}}.$$

Discrete to Continuum Convergence

We can also define the constrained functionals

$$E_{n,\mathrm{cons}}(u) = \begin{cases} E_n(u), & \text{if } u = g \text{ on } \mathcal{O}_n, \\ \infty, & \text{else,} \end{cases} \qquad \mathcal{E}_{\mathrm{cons}}(u) = \begin{cases} \mathcal{E}(u), & \text{if } u = g \text{ on } \mathcal{O}, \\ \infty, & \text{else} \end{cases}$$

and obtain

Discrete to Continuum Convergence [RB23]

Let Ω be locally convex. For any null sequence ε_n such that

$$d_H(\Omega_n, \Omega) \ll \varepsilon_n,$$
$$d_H(\mathcal{O}_n, \mathcal{O}) \ll \varepsilon_n.$$

we have that

$$E_{n,\mathsf{cons}} \xrightarrow{\Gamma} \sigma_\eta \, \mathcal{E}_{\mathsf{cons}}.$$

Convergence of minimizers?

Compactness

A sequence of functionals $F_n : X \to \mathbb{R}$ is called compact if for any sequence $(x_n)_{n\in\mathbb{N}}$ the property

$$\sup_{n\in\mathbb{N}} F_n(x_n) < \infty$$

implies that $(x_n)_{n\in\mathbb{N}}$ is relatively compact.

Compactness

A sequence of functionals $F_n : X \to \mathbb{R}$ is called compact if for any sequence $(x_n)_{n\in\mathbb{N}}$ the property

$$\sup_{n\in\mathbb{N}} F_n(x_n) < \infty$$

implies that $(x_n)_{n\in\mathbb{N}}$ is relatively compact.

Compactness Result [RB23]

Let ε_n be a null sequence such that

$$d_H(\Omega_n, \Omega) \ll \varepsilon_n,$$
$$d_H(\mathcal{O}_n, \mathcal{O}) \ll \varepsilon_n.$$

then $E_{n,\text{cons}}$ is a compact sequence of functionals. Therefore, every sequence of minimizers for $E_{n,\text{cons}}$ has a cluster point which is a minimizer of $\mathcal{E}_{\text{cons}}$.