[1]Technical University of Berlin, Institute of Mathematics

[2]University of Minnesota, School of Mathematics

[3]Friedrich-Alexander-Universität Erlangen-Nürnberg, Department Mathematik

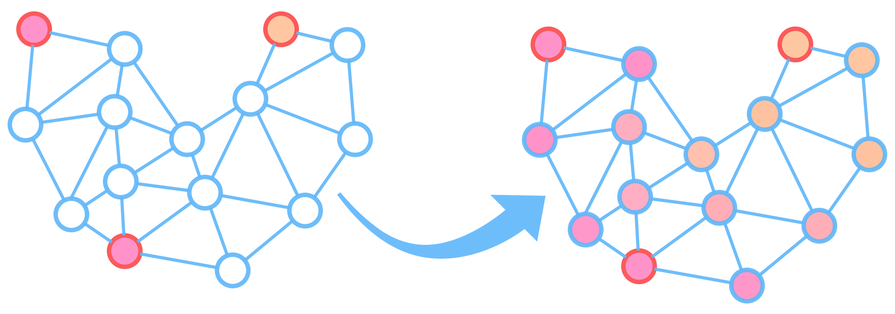
Convergence rates for Lipschitz learning on very sparse graphs

Leon Bungert[1] Jeff Calder[2] Tim Roith[3]

Semi-supervised Learning

Given a finite set of n points $\Omega_n \subset \Omega \subset \mathbb{R}^d$ with labels $g : \mathcal{O}_n \subset \Omega_n \to \mathbb{R}$, find a function

$$\mathbf{u} : \Omega_n \to \mathbb{R}, \quad \text{s.t. } \mathbf{u} = g \text{ on } \mathcal{O}_n.$$



Weighted Graphs

We model the data as a **weighted graph** (Ω_n, w_n) with edge weights given by

$$w_n(x,y) := \eta(|x-y|/\varepsilon_n), \qquad x,y \in \Omega_n.$$

Here $\varepsilon_n > 0$ is a **scaling parameter** and $\eta : [0,\infty) \to [0,\infty)$ is a kernel function.

Laplacian Learning

Laplacian learning for $p < \infty$ involves the energy

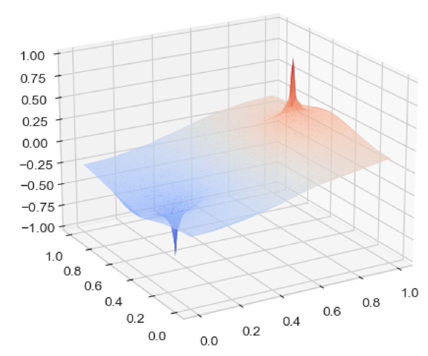
$$\mathbf{E}_p^{w_n}(\mathbf{u}) = \sum_{x,y\in\Omega_n} w_n(x,y)^p |\mathbf{u}(y) - \mathbf{u}(x)|^p$$

and the graph Laplacian

$$\Delta_p^{w_n}\mathbf{u}(x) := \sum_{y\in\Omega_n} w_n(x,y)^p \, |\mathbf{u}(y) - \mathbf{u}(x)|^{p-2} (\mathbf{u}(y) - \mathbf{u}(x))$$

which yields the equivalent problems:

$$\min_{\mathbf{u}} \mathbf{E}_p^{w_n}(\mathbf{u}), \ \text{s.t. } \mathbf{u} = g \text{ on } \mathcal{O}_n. \quad \Leftrightarrow \quad \Delta_p^{w_n}\mathbf{u} = 0 \text{ in } \Omega_n \setminus \mathcal{O}_n, \ \mathbf{u} = g \text{ on } \mathcal{O}_n.$$



Drawback: In the infinite data limit this problem is only well-posed if $p > d$.

Lipschitz Learning

In the limit $p \to \infty$ we obtain the energy

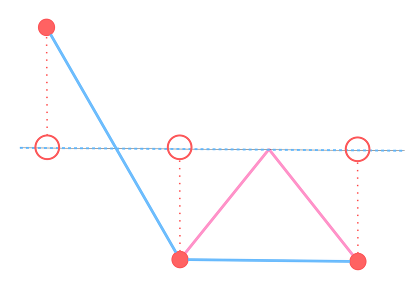
$$\mathbf{E}_\infty^{w_n}(\mathbf{u}) = \max_{x,y\in\Omega_n} w_n(x,y)|\mathbf{u}(y) - \mathbf{u}(x)|$$

and the graph infinity Laplacian

$$\Delta_\infty^{w_n}\mathbf{u}(x) := \max_{y\in\Omega_n} w_n(x,y)\,(\mathbf{u}(y) - \mathbf{u}(x)) + \min_{y\in\Omega_n} w_n(x,y)\,(\mathbf{u}(y) - \mathbf{u}(x)).$$

The associated problems are not equivalent.

$$\min_{\mathbf{u}} \mathbf{E}_\infty^{w_n}(\mathbf{u}), \ \text{s.t. } \mathbf{u} = g \text{ on } \mathcal{O}_n. \quad \Leftarrow \quad \Delta_\infty^{w_n}\mathbf{u} = 0 \text{ in } \Omega_n \setminus \mathcal{O}_n, \ \mathbf{u} = g \text{ on } \mathcal{O}_n.$$



Minimizing $\mathbf{E}_\infty^{w_n}$ does not admit unique solutions: the blue line shows the **AMLE**.

Graph Resolution

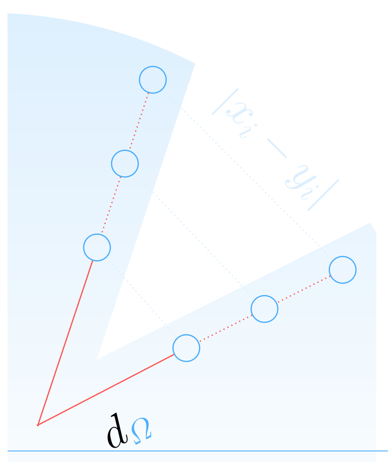
Using the Hausdorff distance

$$d_H(A,B) = \sup_{x\in A} \operatorname{dist}(x,B) \vee \sup_{x\in B} \operatorname{dist}(x,A)$$

we consider the **graph resolution**:

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

Local Convexity

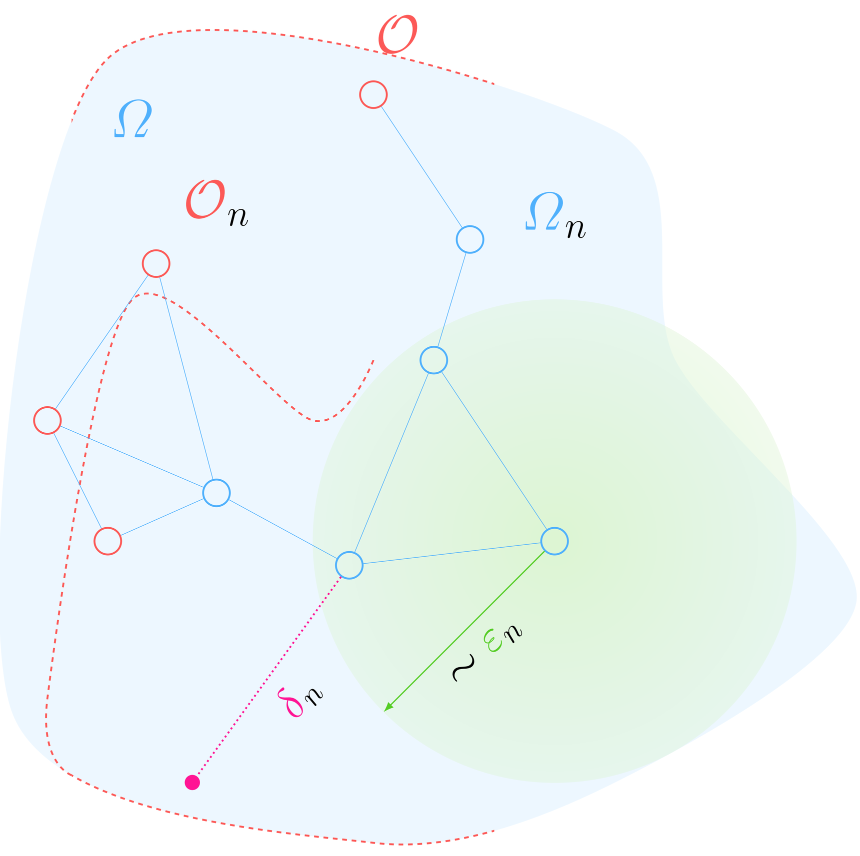


We assume that Ω is **locally convex**, i.e., $\forall x,y \in \Omega$:

$$d_\Omega(x,y) \le |x-y| + \phi(|x-y|)$$

where $\phi(\varepsilon) \ll \varepsilon$ as $\varepsilon \to 0$.

$\Rightarrow$ No sharp internal corners.



Γ-Convergence

Considering the functionals

$$E_\infty^{w_n}(\mathbf{u}) := \frac{1}{\varepsilon_n}\mathbf{E}_\infty^{w_n}(\mathbf{u}) \qquad + \textbf{constraint},$$

$$\mathcal{E}_\infty(u) := \|\nabla u\|_{L^\infty(\Omega)} \qquad + \textbf{constraint},$$

we prove in [LIP-I]:

- $E_\infty^{w_n} \xrightarrow{\Gamma} \sigma_\eta \, \mathcal{E}_\infty$ in a suitable topology,
- $\sup_{n\in\mathbb{N}} E_\infty^{w_n}(\mathbf{u}_n) < \infty$ and boundedness imply that $(\mathbf{u}_n)_{n\in\mathbb{N}}$ is relatively compact.

Here we assume the **weakest scaling**

$$\delta_n \ll \varepsilon_n \ll 1.$$

The ∞-Laplacian and AMLEs

For $\Delta_\infty u(x) := \langle \nabla u(x), \nabla^2 u(x) \nabla u(x)\rangle$ we consider the problems:

∞-Laplacian	AMLE	CDF
$\Delta_\infty u = 0$ in Ω, $u = g$ on $\mathcal{O}$, $\frac{\partial u}{\partial \nu} = 0$ on $\partial\Omega \setminus \mathcal{O}$.	$\operatorname{Lip}_{d_\Omega}(u;\overline{V}) = \operatorname{Lip}_{d_\Omega}(u,\partial V)$, $u = g$ on $\mathcal{O}$, for all open and connected sets $V \subset \overline{\Omega} \setminus \mathcal{O}$.	$h_\pm := \pm u - a\,\lvert\cdot - y\rvert$ satisfies $\max_{\overline{V}} h_\pm = \max_{\partial V} h_\pm$, for all such sets $V \subset \overline{\Omega} \setminus \mathcal{O}$, $y \in \overline{\Omega} \setminus V$, and $a \ge 0$.

- Regular $\partial\Omega$: ∞-harmonic $\Leftrightarrow$ AMLE $\Leftrightarrow$ CDF.
- AMLE and CDF can be generalized to the graph, using the distance

$$d_n(x,y) := \inf\left\{\sum_{i=1}^{m-1} \frac{1}{\omega_n(p_{i+1},p_i)} \ : \ m \in \mathbb{N}, \, p \in \Omega_n^m, \, p_1 = x, \, p_m = y\right\}.$$

[LIP-II]: $\Delta_\infty^{w_n}\mathbf{u} = 0 \ \Rightarrow \ \mathbf{u}$ is graph AMLE and fulfills graph CDF.

Meeting in the Middle

Larger length scale $\tau > 0$: consider **homogenized operator**

$$\tau^2 \Delta_\infty^\tau u(x) := \sup_{y\in\overline{B}(x;\tau)} (u(y) - u(x)) + \inf_{y\in\overline{B}(x;\tau)} (u(y) - u(x)).$$

Local to Non-Local: CDF implies [AS10] for

$$u_\tau(x) = \inf_{\overline{B}(x;\tau)} u$$

that

$$-\Delta_\infty u \ge 0 \Rightarrow -\Delta_\infty^\tau u_\tau \ge 0.$$

Discrete to Non-Local: CDF on the graph implies [LIP-II] for $u_n^\tau(x) := \max_{y\in\overline{B}(x;\tau)\cap\Omega_n} \mathbf{u}_n(y)$ that

$$-\Delta_\infty^{w_n}\mathbf{u}_n \le 0 \Rightarrow -\Delta_\infty^\tau u_n^\tau \lesssim \frac{r_\tau}{\tau} + \frac{\varepsilon_n}{\tau^2},$$

where

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

Non-Local: Using **strictness perturbations** and Lipschitzness, and swapping signs we obtain

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

Key Insight from Equation (rate):

"Ratio convergence rates for the distance function imply convergence rates for AMLEs."

In [LIP-II] we prove:

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

where $\sigma_\eta := \sup_{t>0} \eta(t)t$ and $\tau_\eta := \sup_{t>0} \sigma_\eta \eta(t)^{-1} - t$.

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

Dense regime: If $\varepsilon_n \gtrsim \delta_n^{\frac{5}{9}}$ the rate is $\varepsilon_n^{\frac{1}{5}}$.

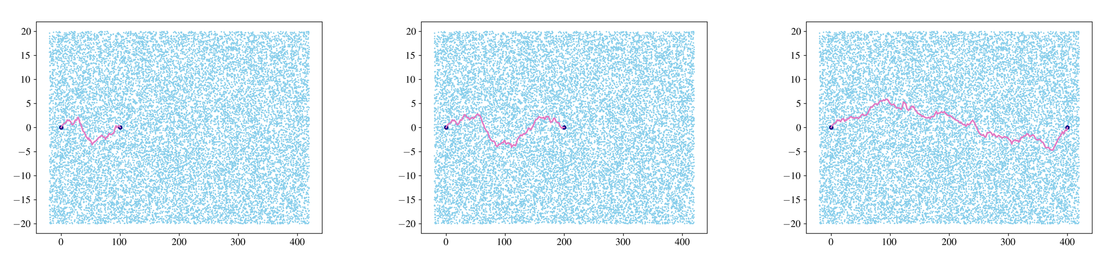
What about $\varepsilon_n \sim \delta_n$? Percolation!

On a unit intensity Poisson Point Process $X \subset \mathbb{R}^d$ and for $h > 0$ let $\Pi_h(x,y)$ be the set of admissible paths and

$$T_s := d_{h_s}(0, se_1) := \inf\left\{\sum_{i=1}^{\operatorname{len}(p)-1} |p_{i+1} - p_i| \ : \ p \in \Pi_{h_s}(0, se_1)\right\}.$$

Important: Replace T_s by distance T_s' on an **enriched process** $\mathcal{X}_s$.

tim.roith@fau.de



$s \mapsto h_s$ are growing step sizes with $h_s \gtrsim \log(s)^{\frac{1}{d}}$.

Approximate sub- and superadditivity:

$$\mathbb{E}[T_{s+t}'] \le \mathbb{E}[T_s'] + \mathbb{E}[T_t'] + g(s+t) \quad ✔$$

$$\mathbb{E}[T_{2s}'] \ge 2\mathbb{E}[T_s'] - g(s) \quad ???$$

For **ratio convergence** it suffices to freeze h_s:

$$\mathbb{E}[d_{h_s,\mathcal{X}_s}(0, 2se_1)] \ge 2\mathbb{E}[d_{h_s,\mathcal{X}_s}(0,s)] - g(s).$$

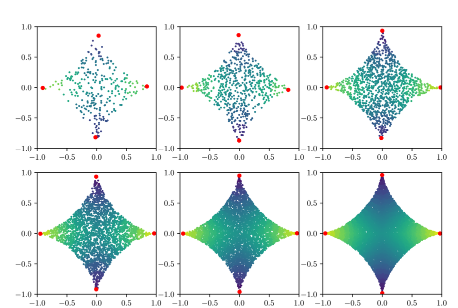
With **concentration of measure** we get [LIP-III]

$$\left|\frac{\mathbb{E}[d_{h_s,\mathcal{X}_s}(0,se_1)]}{\mathbb{E}[d_{h_s,\mathcal{X}_s}(0,2se_1)]} - \frac{1}{2}\right| \lesssim \sqrt{\frac{\log(s)^{2/d}}{h_s}\frac{\log(s)}{\sqrt{s}}}$$

$$\Rightarrow \max_{\Omega_n} |\mathbf{u}_n - u| \lesssim \log(n)^{2/9}\delta_n^{1/9}.$$

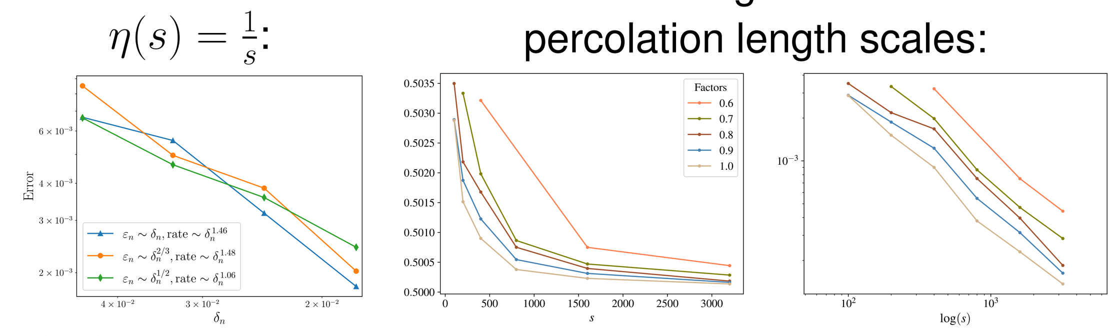
Numerical Examples

Infinity harmonic function $u(x_1,x_2) = |x_1|^{4/3} - |x_2|^{4/3}$ on $\Omega := \{|x_1|^{2/3} + |x_2|^{2/3} \le 1\}$.



AMLE rates for $\eta(s) = \frac{1}{s}$:

Ratio convergence rates for percolation length scales:



References

[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 PDEs* 37.3 (2010), pp. 381–384.

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

[LIP-II] L. Bungert, J. Calder, and T. Roith. "Uniform convergence rates for Lipschitz learning on graphs". In: *IMA Journal of Numerical Analysis* (2022).

[LIP-III] L. Bungert, J. Calder, and T. Roith. *Ratio convergence rates for Euclidean first-passage percolation: Applications to the graph infinity Laplacian*. 2022. arXiv: 2210.09023 [math.PR].