Large data limit of the MBO scheme for data clustering

Jona Lelmi

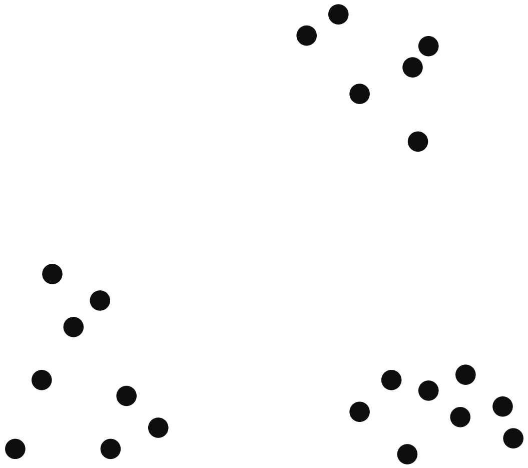
Hausdorff Center for Mathematics - Institute of Applied Mathematics

Math Machine Learning seminar MPI MiS + UCLA

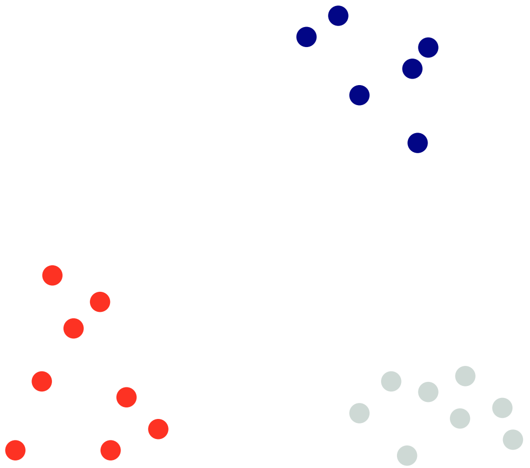
Dec 16, 2022

Joint work with Tim Laux (U Bonn)

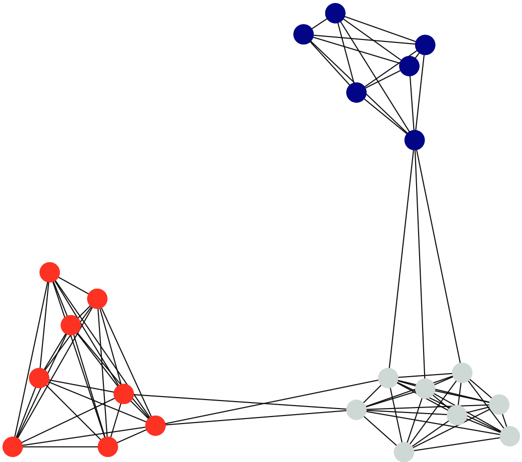
Data Clustering



Data Clustering



Data Clustering



Similarity Graph

$V = \{x_1, \ldots, x_n\} \subset \mathbf{R}^d$ data points, $G = (V, W)$ weighted graph.

Choice of *adjacency matrix* W:

- $W_{ii} = 0$ for all $1 \leq i \leq n$;
- $W_{ij} = \eta(|x_i - x_j|)$ for $i \neq j$. The function $\eta : \mathbf{R}_+ \to \mathbf{R}_+$ respects the principle

$$W_{ij} \text{ large} \Leftrightarrow x_i \text{ and } x_j \text{ similar} \Leftrightarrow |x_i - x_j| \text{ small.}$$

Similarity Graph

$V = \{x_1, \dots, x_n\} \subset \mathbf{R}^d$ data points, $G = (V, W)$ weighted graph.

Choice of *adjacency matrix* W:

- $W_{ii} = 0$ for all $1 \leq i \leq n$;
- $W_{ij} = \eta(|x_i - x_j|)$ for $i \neq j$. The function $\eta : \mathbf{R}_+ \to \mathbf{R}_+$ respects the principle

$$W_{ij} \text{ large} \Leftrightarrow x_i \text{ and } x_j \text{ similar} \Leftrightarrow |x_i - x_j| \text{ small.}$$

Possible choices for η

$$\eta(t) = exp(-t^2), \quad \eta = \mathbf{1}_{[0,1]}$$

Graph Laplacian

Define the graph Laplacian

$$\Delta_n = I_n - \frac{1}{n} D^{-1} W.$$

Here, $D = \operatorname{diag}(d_1, \ldots, d_n)$, where

$$d_i = \frac{1}{n} \sum_{j=1}^{n} W_{ij}$$

is the degree of the vertex x_i.

Graph Laplacian

Define the graph Laplacian

$$\Delta_n = I_n - \frac{1}{n} D^{-1} W.$$

Here, $D = \mathrm{diag}(d_1, \ldots, d_n)$, where

$$d_i = \frac{1}{n} \sum_{j=1}^{n} W_{ij}$$

is the degree of the vertex x_i.

The Laplacian Δ_n is a non-negative self-adjoint operator w.r.t. the scalar product on $\mathcal{V} = \{u : V \to \mathbf{R}\}$:

$$\langle u, v \rangle_{\mathcal{V}} = \frac{1}{n} \sum_{i=1}^{n} d_i u_i v_i.$$

Graph Based Learning

GENERAL IDEA use the graph structure to detect underlying geometry in the data.

Examples: Label Propagation/Laplace learning, diffusion maps, p-Laplace learning, Poisson learning, MBO

Graph Based Learning

GENERAL IDEA use the graph structure to detect underlying geometry in the data.

Examples: Label Propagation/Laplace learning, diffusion maps, p-Laplace learning, Poisson learning, **MBO**

MBO scheme

- Introduced by Merriman, Bence and Osher to approximate mean curvature flow;
- Adapted by Bertozzi et al. for data clustering.

MBO Scheme

Algorithm (MBO)

Choose $h > 0$ *step-size and let* $\chi : V_n \to \{0, 1\}$ *a proposed clustering. Update* χ *in the following way:*

1. DIFFUSION *set* $u := e^{-h\Delta_n}\chi$
2. THRESHOLDING *for* $x \in V_n$*, set* $\chi(x) = 1 \Leftrightarrow u(x) \geq \frac{1}{2}$

MBO Scheme

Algorithm (MBO)

Choose $h > 0$ *step-size and let* $\chi : V_n \to \{0, 1\}$ *a proposed clustering. Update* χ *in the following way:*

1. DIFFUSION *set* $u := e^{-h\Delta_n}\chi$
2. THRESHOLDING *for* $x \in V_n$, *set* $\chi(x) = 1 \Leftrightarrow u(x) \geq \frac{1}{2}$

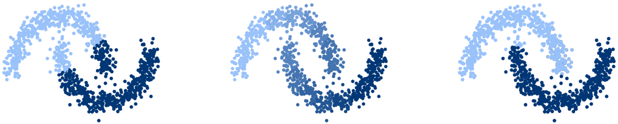


Figure: One iteration of the MBO scheme.

MBO Scheme

Algorithm (Semi-supervised MBO)

Choose $h > 0$ step-size. Let $f : V_n \to \mathbf{R}$ a forcing term and let $\chi : V_n \to \{0, 1\}$ a proposed clustering. Update χ in the following way:

1. DIFFUSION *set* $u := e^{-h\Delta_n}\chi$
2. THRESHOLDING *for* $x \in V_n$, *set* $\chi(x) = 1 \Leftrightarrow u(x) \geq \frac{1}{2} - \sqrt{h} f(x)$

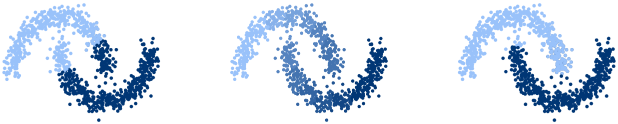


Figure: One iteration of the MBO scheme.

Minimizing Movements

Iterating the scheme $N \in \mathbf{N}$ times gives a sequence $\{\chi^l\}_{l=0}^N$ of partitions.

Esedoḡlu and Otto ('15) proved that

$$\chi^{l+1} = \operatorname{argmin}\left\{ E_h^n(u) + \frac{1}{2h} {d_h}^2(u, \chi^l) \right\},$$

where E_h^n is the thresholding energy on $G_n = (V_n, W_n)$, defined for $u : V_n \to [0,1]$ as

$$E_h^n(u) := \frac{1}{\sqrt{h}} \langle (1-u), e^{-h\Delta_n} u \rangle_{\mathcal{V}_n},$$

and d_h is a suitable distance.

Large Data Limit

N large $\Rightarrow \chi^N$ is close to a (local) minimizer of E_h^n

LARGE DATA LIMIT What is the asymptotic behavior of these (local) minimizers as the number of data points increases?

Large Data Limit

N large $\Rightarrow \chi^N$ is close to a (local) minimizer of E_h^n

LARGE DATA LIMIT What is the asymptotic behavior of these (local) minimizers as the number of data points increases?

Hot topic in theoretical data science: spectral clustering (Audibert, García Trillos, Gerlach, Hein, Slepčev, Von Luxburg), Lipschitz learning (Bungert, Calder, Roith), Laplace learning (Calder, García Trillos, Lewicka, Slepčev, Thorpe)...

Gives insights on *parameters tuning*, *regimes of validity of an algorithm*, *design of new algorithms*.

Γ-convergence

$$(X, d) \text{ metric space, } F_n, F : X \to \overline{\mathbf{R}}.$$

WANT

$$\underset{x\in X}{\operatorname{argmin}} F_n(x) \to \underset{x\in X}{\operatorname{argmin}} F(x), \quad \text{as } n \to +\infty.$$

Γ-convergence

$$(X, d) \text{ metric space}, \ F_n, F : X \to \overline{\mathbf{R}}.$$

WANT

$$\underset{x\in X}{\operatorname{argmin}} F_n(x) \to \underset{x\in X}{\operatorname{argmin}} F(x), \quad \text{as } n \to +\infty.$$

Γ-convergence

We say that $F = \Gamma - \lim_{n\to+\infty} F_n$ if:

1. For every $x \in X$ and every sequence $x_n \to x$, it holds that $\liminf_{n\to+\infty} F_n(x_n) \geq F(x)$;
2. For every $x \in X$ there exists a sequence $x_n \to x$ such that $\limsup_{n\to+\infty} F_n(x_n) \leq F(x)$.

Mathematical Framework

We work under the *manifold assumption*. The data points $\{x_i\}_{i=1}^{+\infty}$ are such that:

- $M \subset \mathbf{R}^d$ is closed, k-dimensional Riemannian submanifold.
- There exists $\nu := \rho \, \mathrm{Vol}_M \in \mathcal{P}(M)$ such that $\rho \in C^\infty(M)$, $\rho > 0$.
- They are *i.i.d.* with $x_i \sim \nu$.

Mathematical Framework

We work under the *manifold assumption*. The data points $\{x_i\}_{i=1}^{+\infty}$ are such that:

- $M \subset \mathbf{R}^d$ is closed, k-dimensional Riemannian submanifold.
- There exists $\nu := \rho \, \mathrm{Vol}_M \in \mathcal{P}(M)$ such that $\rho \in C^\infty(M)$, $\rho > 0$.
- They are *i.i.d.* with $x_i \sim \nu$.

For every n, $G_n = (V_n, W_n)$ is the weighted graph constructed on $V_n = \{x_1, ..., x_n\}$, with

$$(W_n)_{ij} = \frac{1}{\epsilon_n^k} \eta \left(\frac{|x_i - x_j|}{\epsilon_n} \right),$$

for some infinitesimal sequence $\epsilon_n \gg \left(\frac{\log n}{n} \right)^{1/k}$. Then

$$\Delta_{n,\epsilon_n} := \frac{1}{\epsilon_n^2} \Delta_n \xrightarrow{n \to +\infty} \kappa(\eta) \Delta_{\rho^2}$$

Mathematical Framework

We work under the *manifold assumption*. The data points $\{x_i\}_{i=1}^{+\infty}$ are such that:

- $M \subset \mathbf{R}^d$ is closed, k-dimensional Riemannian submanifold.
- There exists $\nu := \rho \, \mathrm{Vol}_M \in \mathcal{P}(M)$ such that $\rho \in C^\infty(M)$, $\rho > 0$.
- They are *i.i.d.* with $x_i \sim \nu$.

For every n, $G_n = (V_n, W_n)$ is the weighted graph constructed on $V_n = \{x_1, ..., x_n\}$, with

$$(W_n)_{ij} = \frac{1}{\epsilon_n^k} \eta \left(\frac{|x_i - x_j|}{\epsilon_n} \right),$$

for some infinitesimal sequence $\epsilon_n \gg \left(\frac{\log n}{n} \right)^{1/k}$. Then

$$\Delta_{n,\epsilon_n} := \frac{1}{\epsilon_n^2} \Delta_n \xrightarrow{n \to +\infty} \kappa(\eta) \Delta_{\rho^2}, \quad \Delta_{\rho^2} f := -\frac{1}{\rho^2} \operatorname{div} \left(\rho^2 \nabla f \right)$$

Energies: discrete-to-nonlocal

Theorem (Laux, L. '21)

In the regime $\left(\frac{\log(n)}{n}\right)^{\frac{1}{k+2}} \ll \epsilon_n \ll 1$ *it holds almost surely*

$$E_h^n \overset{\Gamma(\textit{weak-}TL^2)}{\longrightarrow} E_h,$$

where

$$E_h(u) = \frac{1}{\sqrt{h}} \int_M (1-u) e^{-h\Delta_{\rho^2}} u \, \rho^2 d \, \mathrm{Vol}_M$$

for $u : M \to [0,1]$*; and* $E_h(u) = +\infty$ *otherwise.*

TL^2-convergence

$E_h^n \to E_h$ requires measuring distance of $u_n : V_n \to \mathbf{R}$ and $u \in L^2(M)$, but $u|_{V_n}$ is not well-defined.

TL^2-convergence

$E_h^n \to E_h$ requires measuring distance of $u_n : V_n \to \mathbf{R}$ and $u \in L^2(M)$, but $u|_{V_n}$ is not well-defined.

NEW TOOL (García Trillos, Slepčev '16)

$$u_n \to u \text{ in } TL^2(M) \Leftrightarrow u_n \circ T_n \to u \text{ in } L^2(\nu),$$

for a sequence of transport maps T_n from ν to $\nu_n := \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$ such that

$$\lim_{n\to+\infty} \int_M d_M^2(x, T_n(x))d\nu(x) = 0.$$

TL^2-convergence

$E_h^n \to E_h$ requires measuring distance of $u_n : V_n \to \mathbf{R}$ and $u \in L^2(M)$, but $u|_{V_n}$ is not well-defined.

NEW TOOL (García Trillos, Slepčev '16)

$$u_n \to u \text{ weakly in } TL^2(M) \Leftrightarrow u_n \circ T_n \rightharpoonup u \text{ weakly in } L^2(\nu),$$

for a sequence of transport maps T_n from ν to $\nu_n := \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$ such that

$$\lim_{n\to+\infty} \int_M d_M^2(x, T_n(x))d\nu(x) = 0.$$

Energies: nonlocal-to-local

Theorem (Laux, L. '21)

As $h \to 0$ it holds

$$E_h \overset{\Gamma(\text{strong-}L^1)}{\longrightarrow} E,$$

where

$$E(\chi) = \frac{1}{\sqrt{\pi}} \int_{\partial^* \Omega} \rho^2 \, d\mathcal{H}^{k-1}$$

for $\chi = \mathbf{1}_\Omega \in BV(M)$; and $E(\chi) = +\infty$ otherwise.

Energies: nonlocal-to-local

Theorem (Laux, L. '21)

As $h \to 0$ it holds

$$E_h \overset{\Gamma(strong\text{-}L^1)}{\longrightarrow} E,$$

where

$$E(\chi) = \frac{1}{\sqrt{\pi}} \int_{\partial^* \Omega} \rho^2 d\mathcal{H}^{k-1}$$

for $\chi = \mathbf{1}_\Omega \in BV(M)$; and $E(\chi) = +\infty$ otherwise.

Euclidean case with constant unit density:

- [Alberti, Bellettini '98]
- [Miranda, Pallara, Paronetto, Preunkert '07a]
- [Esedoğlu, Otto '15]

Energies: extensions and open questions

GENERALITY OF RESULTS

- arbitrary number of labels;
- external forcing/drift f, additional term in energy $-\langle f, \chi\rangle_{\nu_n}$;
- other variants of graph Laplacians and data dependent weights.

Energies: extensions and open questions

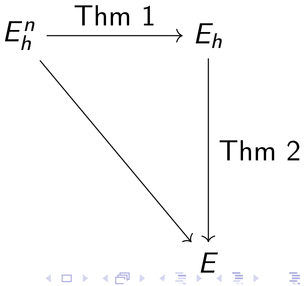
GENERALITY OF RESULTS

- arbitrary number of labels;
- external forcing/drift f, additional term in energy $-\langle f, \chi \rangle_{\nu_n}$;
- other variants of graph Laplacians and data dependent weights.

MAIN OPEN PROBLEM

$$\Gamma - \lim_{n \to +\infty} E^n_{h_n} = ?$$

Conjecture: diagonal arrow holds if $h_n \gg \epsilon_n^2$. Related to the *pinning* of the algorithm.



Dynamics of MBO

$\{\chi_n^l\}_{l=1}^N$ successive iterations of MBO $\Rightarrow \chi_n^N$ is close to a local minimizer of E_h^n.

Selection of local minimizer depends on dynamics of gradient descent (choice of h).

Dynamics of MBO

$\{\chi_n^l\}_{l=1}^N$ successive iterations of MBO $\Rightarrow \chi_n^N$ is close to a local minimizer of E_h^n.

Selection of local minimizer depends on dynamics of gradient descent (choice of h).

Take $h_n \downarrow 0$, can we identify a limit for $\{\chi_{h_n}^l\}_{l=1}^{+\infty}$ as the number of data points increases?

- If $h_n \ll \epsilon_n$ the scheme is *pinned*.
- If h_n is too large the dynamic is trivial.

Identification of continuum limit $\Rightarrow$ insight on the choice of *step size*.

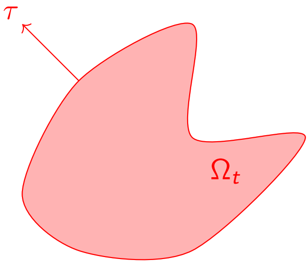
Convergence of the dynamics in the two-class case is possible by viscosity solutions method (approach not possible for more classes)

Mean Curvature Flow on (M, g, ρ^2)

Mean curvature flow on (M, g, ρ^2) is the trajectory of steepest descent for the area functional.

Given initial value $\Omega_0 \subset M$, it is the evolution $t \mapsto \Omega_t$ such that

$$V = -H - \tau \cdot \nabla \log \rho^2.$$



This PDE is driven by minimizing two quantities: AREA and DENSITY.

Satisfies *comparison principle*: if $\Omega_0 \subset \tilde{\Omega}_0$, then $\Omega_t \subset \tilde{\Omega}_t$ for every $t \geq 0$.

$$\downarrow$$

Viscosity solutions machinery.

Comparison principle for MBO

Also MBO satisfied a comparison principle, it is used to prove convergence of the scheme:

- Euclidean MBO $\rightarrow$ Euclidean MCF (Barles, Georgelin '95);
- Regular grid MBO ($\epsilon_n \mathbf{Z}^2$) $\rightarrow$ Euclidean MCF (Misiats, Yip '16).

Comparison principle for MBO

Also MBO satisfied a comparison principle, it is used to prove convergence of the scheme:

- Euclidean MBO $\rightarrow$ Euclidean MCF (Barles, Georgelin '95);
- Regular grid MBO ($\epsilon_n \mathbf{Z}^2$) $\rightarrow$ Euclidean MCF (Misiats, Yip '16).
- Random graph MBO $\rightarrow$ Manifold MCF (Laux, L. '22).

Graph MBO $\to$ Manifold MCF: heuristics

Algorithm (Abstract MBO)

Choose $h_n > 0$ *step-size and let* $\chi_n : V_n \to \{0, 1\}$ *a proposed clustering. Update* χ *in the following way:*

1. DIFFUSION *set* $u_n := S_n(h_n)\chi_n$
2. THRESHOLDING *for* $x \in V_n$, *set* $\chi_n(x) = 1 \Leftrightarrow u_n(x) \geq \frac{1}{2}$

Graph MBO → Manifold MCF: heuristics

Algorithm (Abstract MBO)

Choose $h_n > 0$ step-size and let $\chi_n : V_n \to \{0, 1\}$ a proposed clustering. Update χ in the following way:

1. DIFFUSION *set* $u_n := S_n(h_n)\chi_n$
2. THRESHOLDING *for* $x \in V_n$, *set* $\chi_n(x) = 1 \Leftrightarrow u_n(x) \geq \frac{1}{2}$

We show:

1. $\boxed{S_n(h_n) \sim e^{-h_n \Delta_{\rho^2}}}$ ⇒ MBO scheme converges to Manifold MCF;

Graph MBO → Manifold MCF: heuristics

Algorithm (Abstract MBO)

Choose $h_n > 0$ step-size and let $\chi_n : V_n \to \{0, 1\}$ a proposed clustering. Update χ in the following way:

1. DIFFUSION *set* $u_n := S_n(h_n)\chi_n$
2. THRESHOLDING *for* $x \in V_n$, *set* $\chi_n(x) = 1 \Leftrightarrow u_n(x) \geq \frac{1}{2}$

We show:

1. $\boxed{S_n(h_n) \sim e^{-h_n \Delta_{\rho^2}}}$ ⇒ MBO scheme converges to Manifold MCF;
2. On random geometric graphs

$$\begin{cases} S_n(h_n) & = e^{-h_n \Delta_{n,\epsilon_n}} \\ S_n(h_n) & = e^{-h_n \Delta_{n,\epsilon_n}} P_K \end{cases} \Rightarrow \boxed{S_n \sim e^{-h_n \Delta_{\rho^2}}}$$

Graph MBO $\to$ Manifold MCF

Theorem (Abstract convergence result, Laux, L. '22)

Suppose that

1. *For all* $u, v \in \mathcal{V}_n$

$$u \leq v \Rightarrow S_n(h_n)u \leq S_n(h_n)v + \|(u,v)\|_{V_n,\infty} O(h_n^{\frac{3}{2}}).$$

2. *For all* $f \in C^\infty(M)$

$$\|S_n(h_n)f - e^{-h_n \Delta_{\rho^2}} f\|_{\infty, V_n} = \|f\|_\infty o(h_n^{\frac{1}{2}}) + \|\nabla f\|_\infty O(h_n^{\frac{3}{2}}).$$

3. $\|S_n(h_n)1 - 1\|_{\infty, V_n} = O(h_n^{\frac{3}{2}})$.

Then the abstract MBO scheme converges to the unique viscosity solution of MCF on (M, g, ρ^2).

Graph MBO → Manifold MCF

Theorem (Laux, L. '22)

Let G_n be a random geometric graph. Assume that $q, \alpha, \beta > 0$ are suitably chosen and

1. $h_n \gg (\log(n))^{-\alpha}$;
2. $\left(\frac{\log(n)}{n^{1/k}}\right)^{\frac{k}{k+4}} \lesssim \epsilon_n \ll (\log(n))^{-\beta}$;
3. $K_n \geq (\log(n))^q$;
4. *The eigenvalues of* Δ_{ρ^2} *satisfy* $\inf_{i\in\mathbf{N}}(\lambda_i - \lambda_{i-1}) > 0$.

Then the operators $e^{-t\Delta_n}$ and $e^{-t\Delta_n}P_{K_n}$ satisfy conditions of the abstract result with probability greater than

$$1 - C\epsilon_n^{-6k}\exp\left(-\frac{n\epsilon_n^{k+4}}{C}\right) - Cn\exp\left(-\frac{n}{C(\log(n))^{2q}}\right).$$

Graph MBO → Manifold MCF

Corollary (Laux, L. '22)

Under the assumptions of the previous theorem, it holds almost surely

$$\textit{Graph MBO} \to \textit{Manifold MCF}.$$

Graph MBO $\to$ Manifold MCF

Corollary (Laux, L. '22)

Under the assumptions of the previous theorem, it holds almost surely

$$\textit{Graph MBO} \to \textit{Manifold MCF}.$$

Insight on:

- choice of number of eigenvalues K_n.
- choice of step size h_n.

Graph MBO → Manifold MCF

Corollary (Laux, L. '22)

Under the assumptions of the previous theorem, it holds almost surely

$$\textit{Graph MBO} \to \textit{Manifold MCF}.$$

Insight on:

- choice of number of eigenvalues K_n.
- choice of step size h_n.

Main technical ingredient: new heat kernel estimate which extends previous work [Dunson, Wu, Wu: ACHA '21].

Main Ingredient: new heat kernel estimate

More precisely, the operator $e^{-t\Delta_{n,\epsilon_n}} P_{K_n}$ corresponds to replacing the heat kernel by its truncated version

$$H_n^{K_n}(t,x,y) = \sum_{i=1}^{K_n} e^{-t\lambda_i^n} \psi_i^n(x) \psi_i^n(y) \frac{d_n(y)}{n}.$$

Main Ingredient: new heat kernel estimate

More precisely, the operator $e^{-t\Delta_{n,\epsilon_n}} P_{K_n}$ corresponds to replacing the heat kernel by its truncated version

$$H_n^{K_n}(t,x,y) = \sum_{i=1}^{K_n} e^{-t\lambda_i^n} \psi_i^n(x)\psi_i^n(y)\frac{d_n(y)}{n}.$$

Lemma (Laux, L. '22)

In the setting of the previous theorem, for $n \gg 1$ we have

$$\max_{x,y\in V_n} \left| H_n^{K_n}(h_n,x,y) - \frac{\rho(y)}{n} H(h_n,x,y)\right| = o\left(\frac{\sqrt{h_n}}{n}\right),$$

with probability greater than $1 - C\epsilon_n^{-6k}\exp(-\frac{n\epsilon_n^{k+4}}{C}) - Cn\exp(-\frac{n}{C(\log(n))^{2q}})$.

Summary

SUMMARY

- Outcome of several iterations of MBO $\sim$ local minimizers of graph thresholding energy E_h^n.
  - Local minimizers of thresholding energy E_h^n are qualitatively close to minimal surfaces on (M, ρ^2).
- Selection of local minimizers depends on the dynamics of MBO.
  - Dynamics of MBO converge to mean curvature flow on (M, ρ^2) in the viscosity sense.

Summary

SUMMARY

- Outcome of several iterations of MBO $\sim$ local minimizers of graph thresholding energy E_h^n.
  - Local minimizers of thresholding energy E_h^n are qualitatively close to minimal surfaces on (M, ρ^2).
- Selection of local minimizers depends on the dynamics of MBO.
  - Dynamics of MBO converge to mean curvature flow on (M, ρ^2) in the viscosity sense.

FUTURE WORK AND OPEN QUESTIONS

- Joint Γ-convergence.
  - Convergence of the dynamics in the multiphase case.
- Test parameters regime numerically: can we avoid parameter tuning by knowing sample size n and true dimension k?

Summary

- Tim Laux, J. L.: Large data limit of the MBO scheme for data clustering: Γ-convergence of the thresholding energies arXiv:2112.06737
- Tim Laux, J. L.: Large data limit of the MBO scheme for data clustering: convergence of the dynamics arXiv:2209.05837

THANK YOU FOR YOUR ATTENTION!

Discrete-to-nonlocal: main idea

From the energy-dissipation inequality for the gradient flow for the heat operator we obtain:

Theorem (Laux, L. '21)

In the setting of the previous theorem, let $u_n \in \mathcal{V}_n$ be a sequence of functions converging weakly to $u \in L^2(M)$ in TL^2, then for every $t > 0$ we have

$$\lim_{n\to+\infty} e^{-t\Delta_{n,\epsilon_n}} u_n = e^{-t\Delta_{\rho^2}} u \text{ strongly in } TL^2.$$

COROLLARIES for fixed $h > 0$

$$E_h^n(u_n) \xrightarrow{n\to+\infty} E_h(u),$$

$$\mathrm{MBO}^h(\Omega \cap G_n, G_n) \xrightarrow{n\to+\infty} \mathrm{MBO}^h(\Omega, M).$$