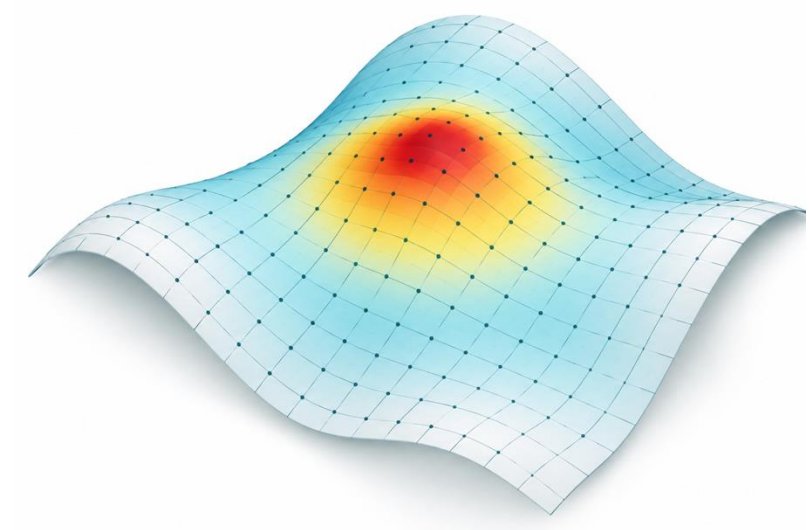


LECTURE 14

Manifold filters and manifold neural networks

Continuous geometric signal processing and graph limits

- 01 Diffuse and filter on a manifold
- 02 Build manifold neural networks
- 03 Explain deformation stability and graph transfer



local heat → geodesic diffusion → continuous filters

Geometry, filtering, and stable learning

01**WHAT IS THE DOMAIN?**

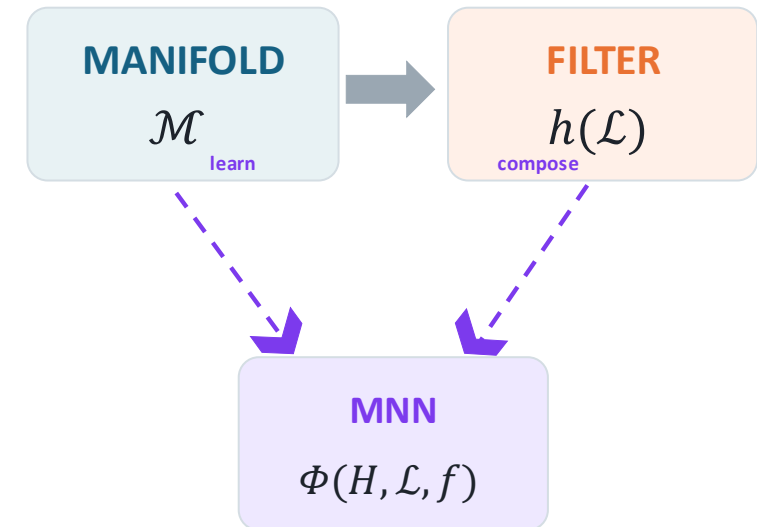
A smooth manifold models the geometry underlying a sampled point cloud.

02**HOW DO WE FILTER?**

Heat diffusion and the positive Laplace–Beltrami operator define propagation.

03**WHEN DOES IT GENERALIZE?**

Spectral regularity yields transfer across sampling and stability to deformation.



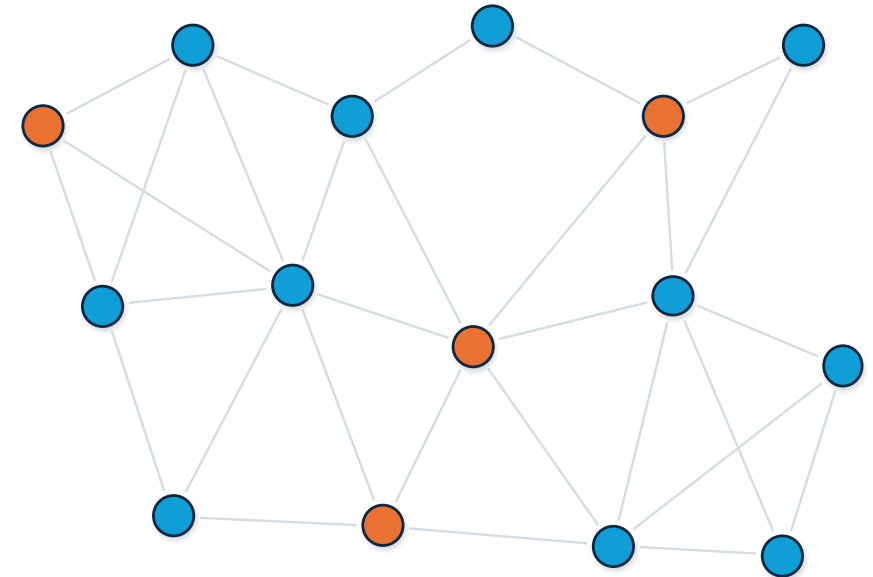
The continuous model connects geometry, filtering, and learning.

Large geometric graphs are samples of physical spaces

- Wireless devices, robots, and point clouds form graphs
- Interactions depend on distance or local geometry

Challenge: the number and placement of nodes change from deployment to deployment – shared limit model

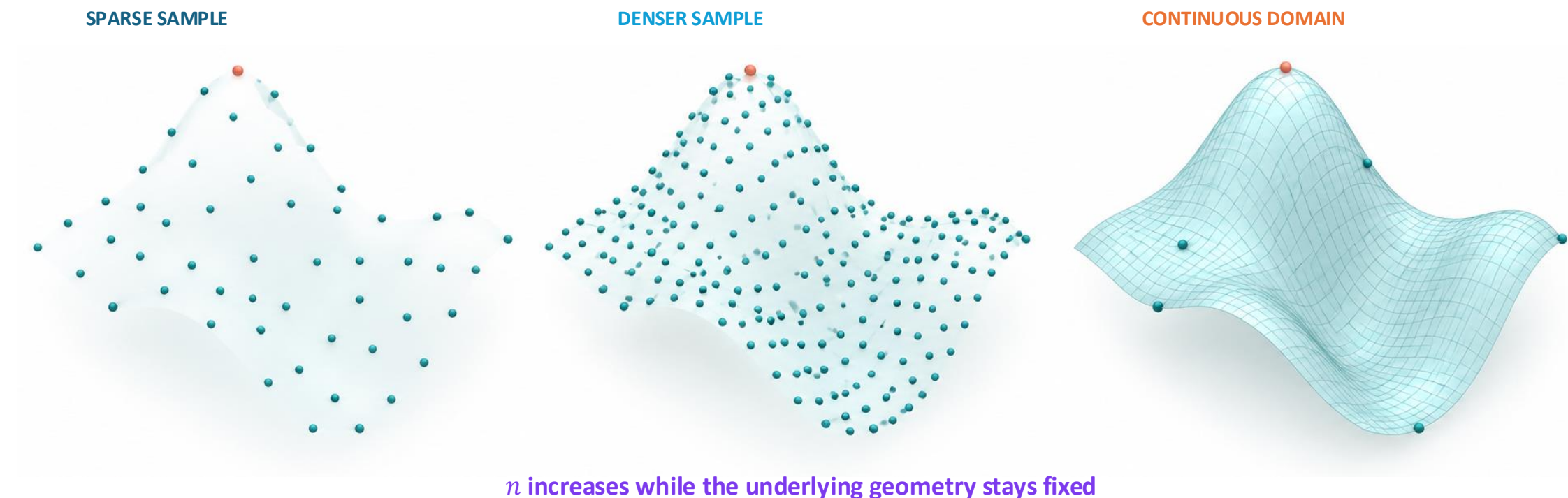
- A manifold is a smooth space that looks flat locally
- the graph is a finite sample of it



distance-based neighborhood graph

Model the geometry once; allow the sampling density to change.

Dense samples converge to a continuous manifold



Convergence requires compatible sampling, bandwidth, and Laplacian normalization.

Dense geometric graphs reveal a smooth domain rather than an unrelated larger graph.

References: Belkin & Niyogi (2008), Towards a theory of Laplacian-based manifold methods, J. Comput. Syst. Sci. 74(8);
Dunson, Wu & Wu (2021), Spectral convergence of graph Laplacian and heat kernel reconstruction in L^∞ from random samples, Appl. Comput. Harmon. Anal. 55;
Calder & García Trillos (2022), Improved spectral convergence rates for graph Laplacians on ϵ -graphs and k-NN graphs, Appl. Comput. Harmon. Anal. 60.

The same design pattern appears in time, graphs, and manifolds

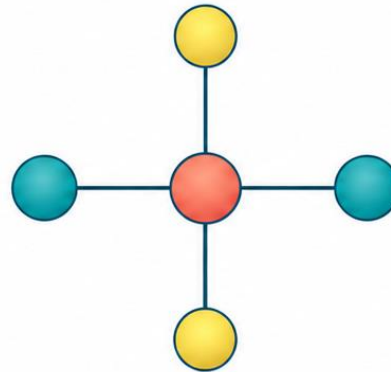
TIME



shift

$$x(t - \tau)$$

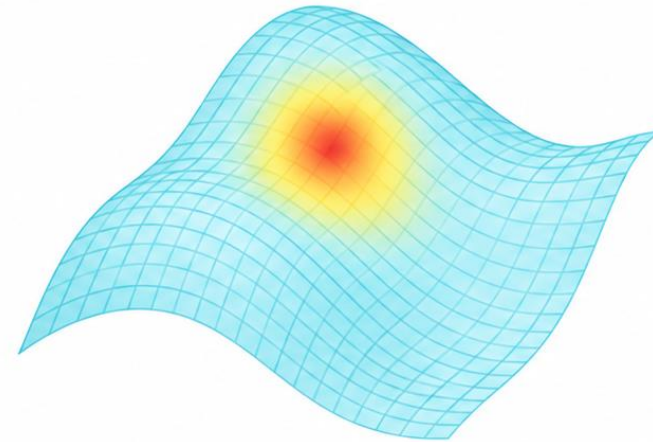
GRAPH



graph shift

$$S^k x$$

MANIFOLD



heat diffusion

$$e^{-t\mathcal{L}} f$$

A filter is a learned linear combination of repeated local propagation.

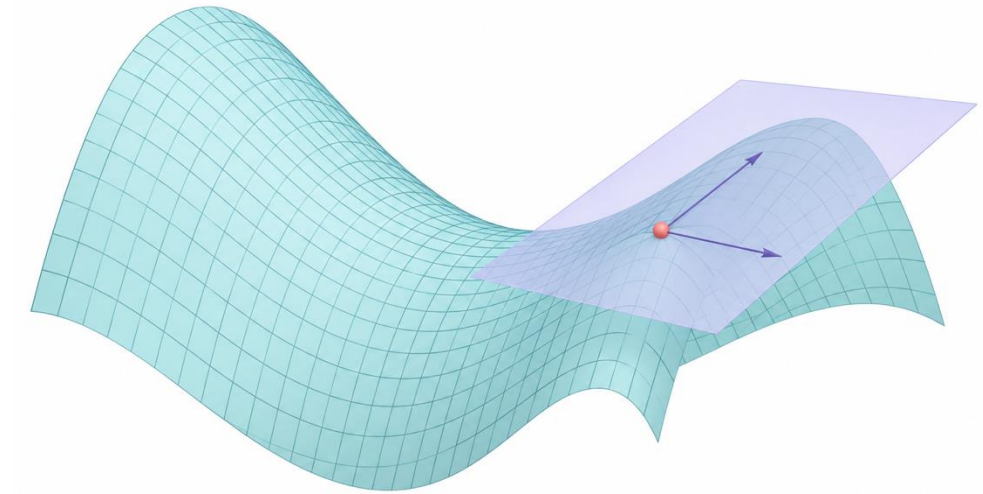
A manifold is locally flat but globally curved

LOCAL LINEAR MODEL AT x

A d -dimensional manifold looks like $\mathbb{R}^d$ locally. Its tangent space contains the first-order directions available at x .

Riemannian metric: lengths and angles along the surface
Geodesic: shortest path along the surface

Assume a smooth, compact, connected manifold without boundary.



Tangent space: a local linear approximation

Local neighborhoods support familiar calculus even when the whole domain bends.

Signals assign features across the manifold

SIGNAL VALUE AT EVERY POINT

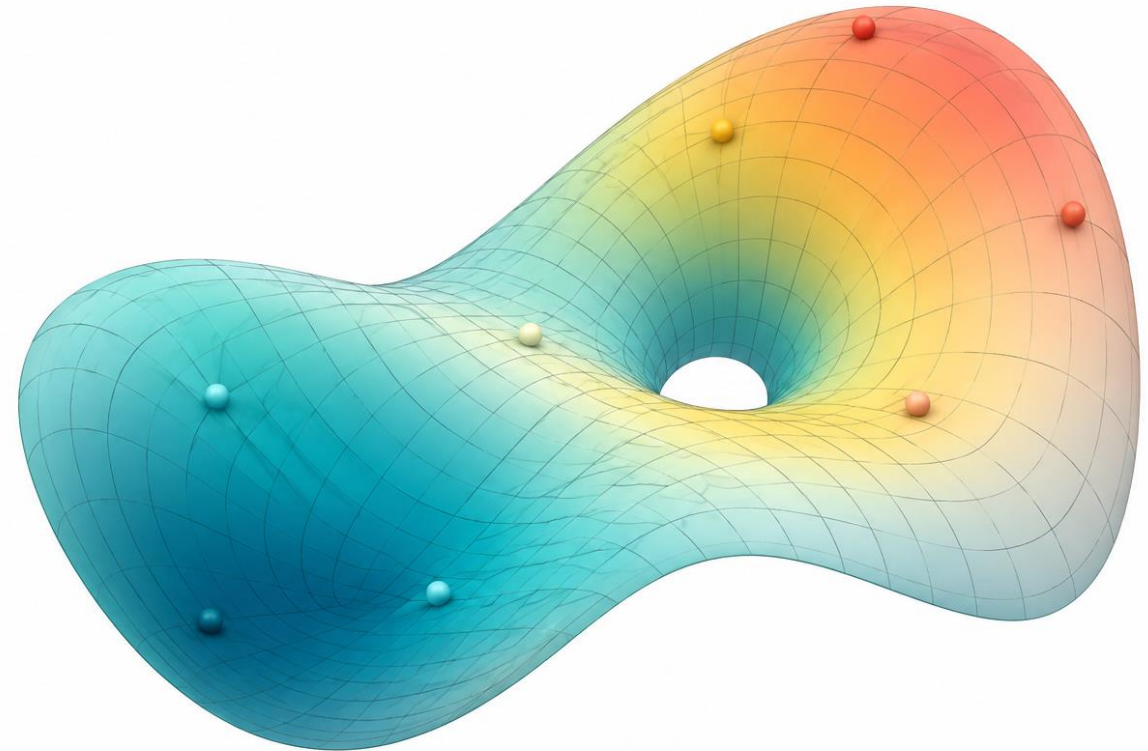
One channel gives one value at each point. Multiple channels store several measurements or learned features.

$$f: \mathcal{M} \rightarrow \mathbb{R}^C$$

color = signal amplitude
markers = sampled observations

$$\|f\|_2^2 = \int_{\mathcal{M}} |f(x)|^2 d\mu(x)$$

Finite energy uses the manifold volume measure μ .



temperature · concentration · learned feature

The geometry tells us which variations are local; the signal tells us what varies.

The Laplace–Beltrami operator measures local variation

GEOMETRIC ANALOGUE OF A LAPLACIAN

The **intrinsic** gradient measures change along the surface. Divergence measures how a vector field spreads.

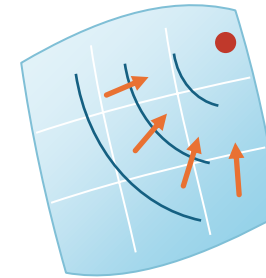
$$\mathcal{L}f = -\operatorname{div}_{\mathcal{M}}(\nabla_{\mathcal{M}}f)$$

Positive semidefinite operator

$$\langle f, \mathcal{L}f \rangle = \int_{\mathcal{M}} f \cdot \mathcal{L}f d\mu(x) = \int_{\mathcal{M}} |\nabla_{\mathcal{M}}f|^2 d\mu(x) \geq 0$$

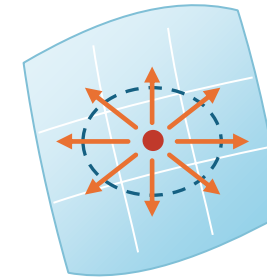
Constant signals have zero Laplacian

GRADIENT



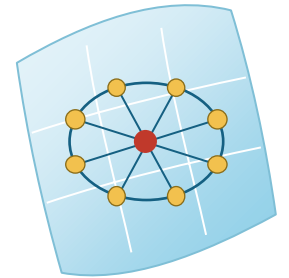
$\nabla_{\mathcal{M}}f$: steepest ascent along the surface

DIVERGENCE



$\operatorname{div}_{\mathcal{M}}F$: net outflow through a small patch

LAPLACE–BELTRAMI



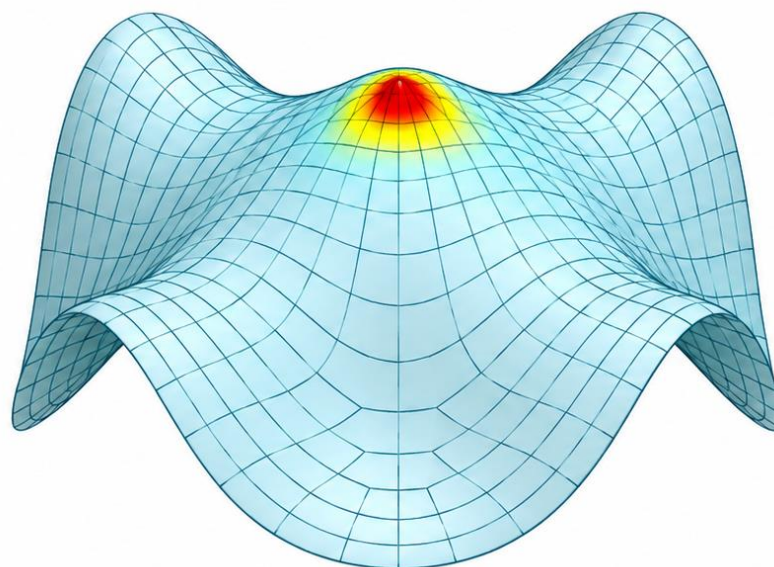
$\mathcal{L}f$: center minus neighbor average

all three depend only on the metric of $\mathcal{M}$, so they are intrinsic

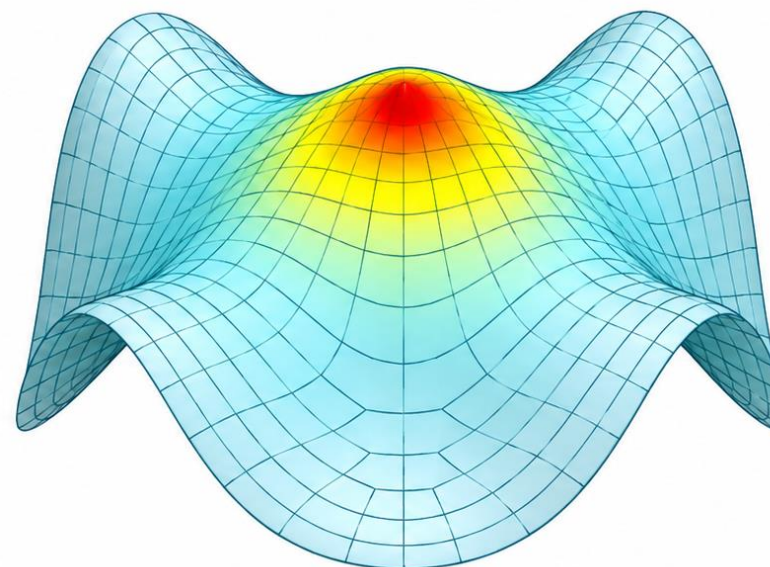
$\mathcal{L}$ encodes the intrinsic geometry needed for filtering.

Heat diffusion is the manifold shift

INITIAL SIGNAL



AFTER DIFFUSION



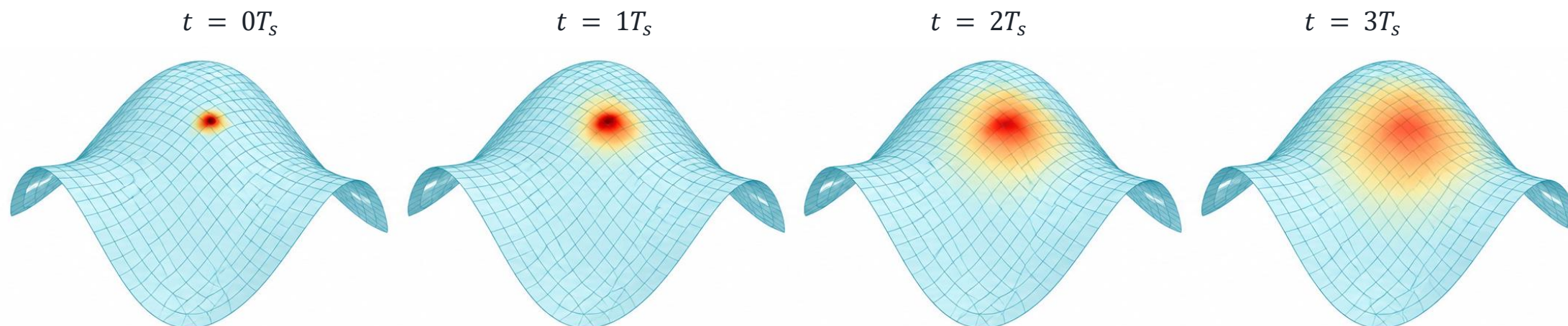
$$\frac{\partial u}{\partial t} = -\mathcal{L}u, \quad u(x, 0) = f(x), \quad \Rightarrow \quad u(x, t) = (e^{-t\mathcal{L}}f)(x)$$

Each mode decays by $e^{-t\lambda}$. Higher frequencies decay faster.

Larger diffusion time aggregates information from a wider geodesic neighborhood.

A manifold filter mixes every diffusion scale

$$g(x) = \int_0^\infty \tilde{h}(t) u(x, t) dt = \int_0^\infty \tilde{h}(t) (e^{-t\mathcal{L}} f)(x) dt = (h(\mathcal{L})f)(x)$$



The diffusion-time weights combine narrow and broad spatial context.

$\tilde{h}(t)$: time weighting $h(\mathcal{L})$: filter operator $\hat{h}(\lambda)$: frequency response

Filter design means choosing weights across diffusion time.

Sampling diffusion time produces a practical finite filter

$$g \approx \sum_{k=0}^{K-1} h_k e^{-kT_s \mathcal{L}} f, \quad h_k \approx T_s \tilde{h}(kT_s)$$

h_k can be learned directly.

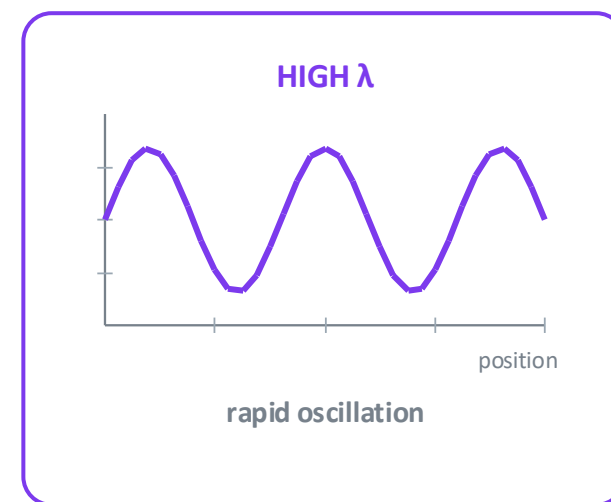
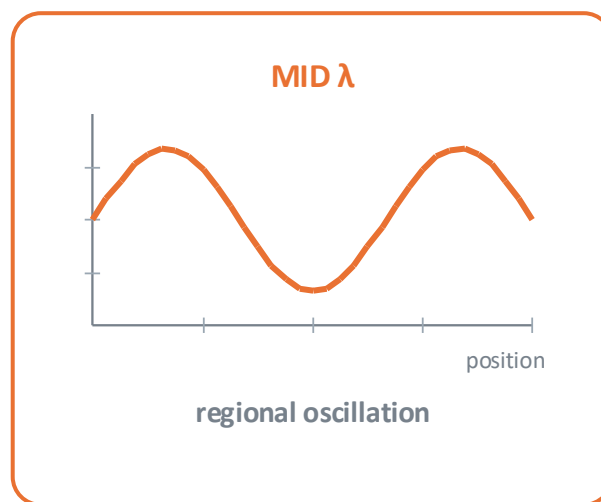


K taps trade computation for spatial reach

The manifold filter is implementable as repeated diffusion plus weighted summation.

Manifold frequencies come from eigenfunctions

$$\mathcal{L}\varphi_i = \lambda_i\varphi_i, \quad 0 = \lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \dots$$



$$\hat{f}_i = \langle f, \varphi_i \rangle, \quad f = \sum_{i=0}^{\infty} \hat{f}_i \varphi_i$$

The zero mode is constant. Curves illustrate frequency.

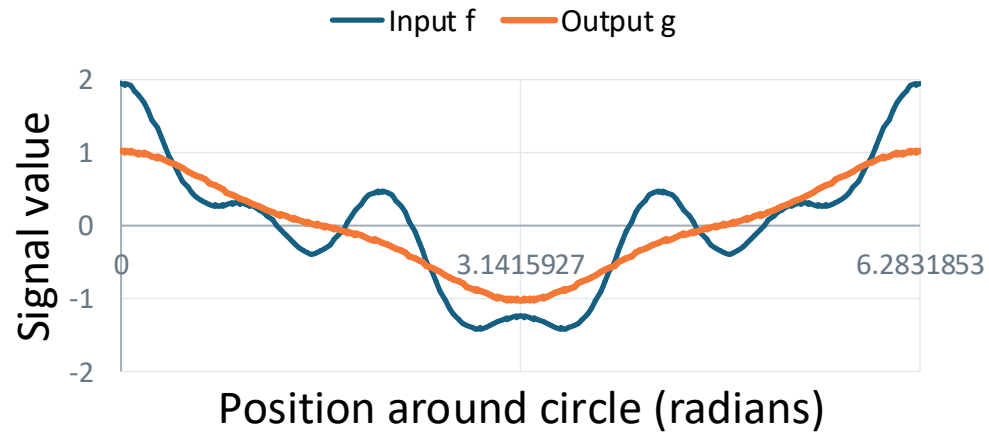
Frequency measures variation over geometry.

Manifold filtering: spatial and spectral views

SPATIAL VIEW

$$g = h(\mathcal{L})f$$

Signal around a unit circle

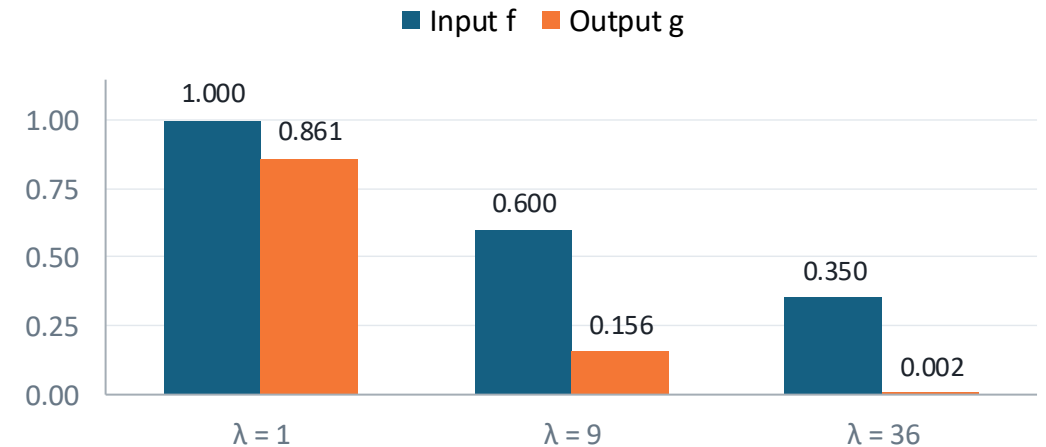


Fine oscillations shrink; broad structure remains.

SPECTRAL VIEW

$$\hat{g}_i = \hat{h}(\lambda_i) \hat{f}_i$$

Input and output mode amplitudes



Gains: 0.861

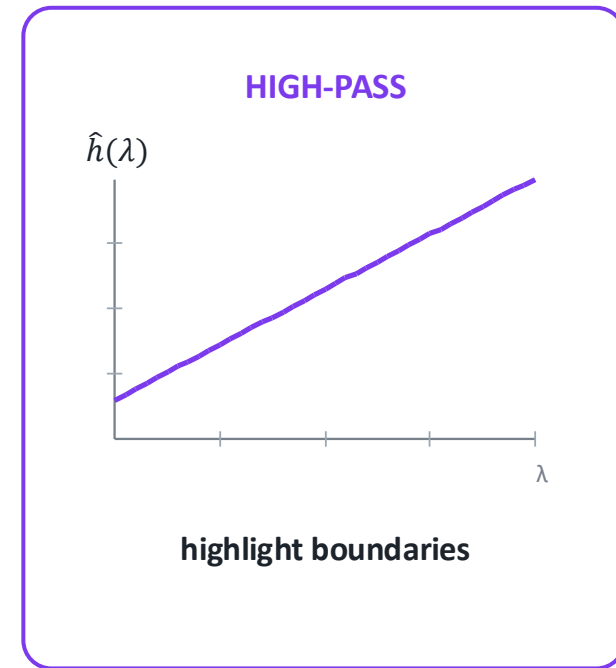
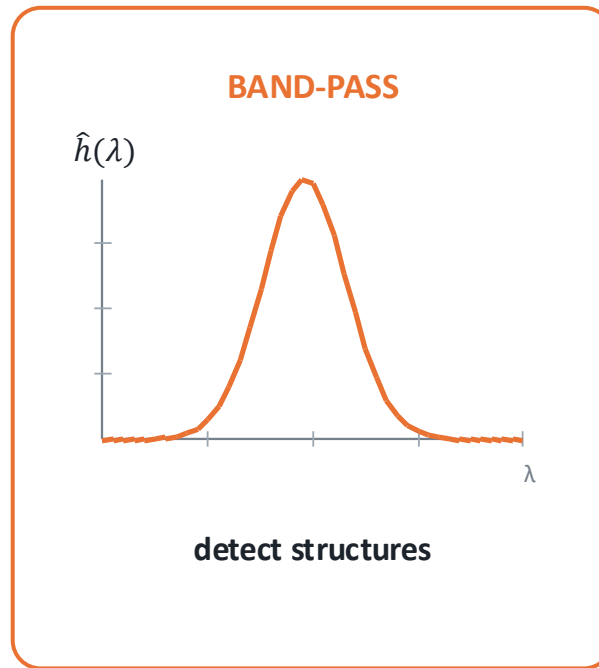
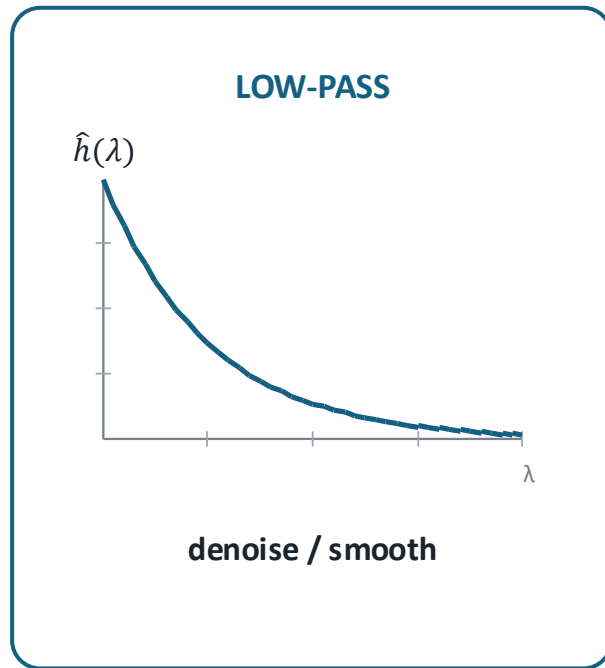
0.259

0.0045

$$\hat{h}(\lambda) = \int_0^\infty \tilde{h}(t) e^{-t\lambda} dt$$

Heat filter at time 0.15: higher frequencies receive smaller gains.

Filter design selects geometric scales

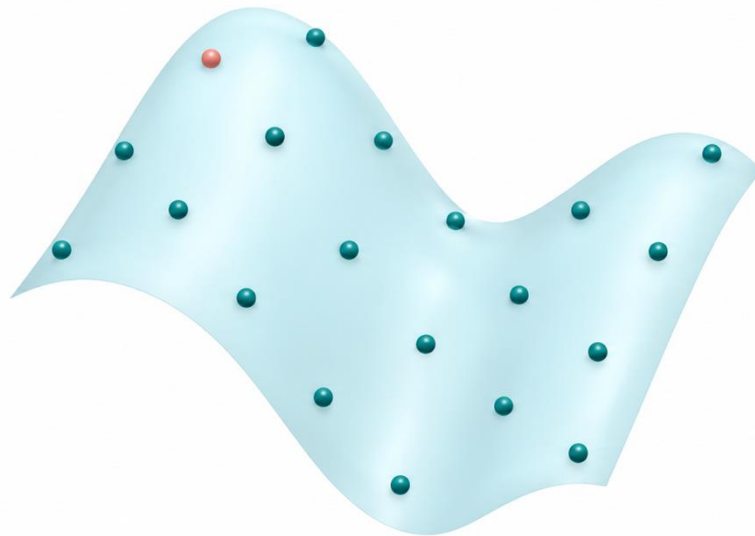


Conceptual responses over a finite spectral range. Realizability depends on the filter family.

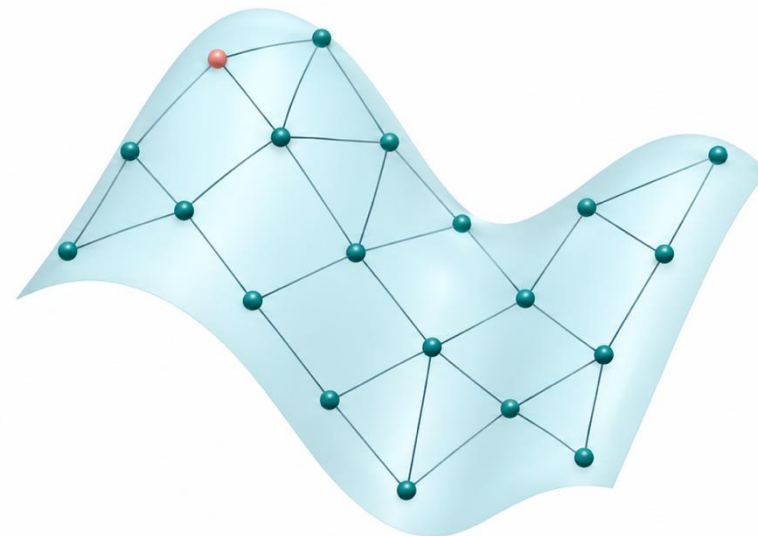
Design $\hat{h}(\lambda)$ as a function; the manifold supplies the eigenvalues where it is evaluated.

Sample a manifold, then connect nearby points with a kernel

1 · SAMPLE POSITIONS x_i



2 · CONNECT LOCAL NEIGHBORS



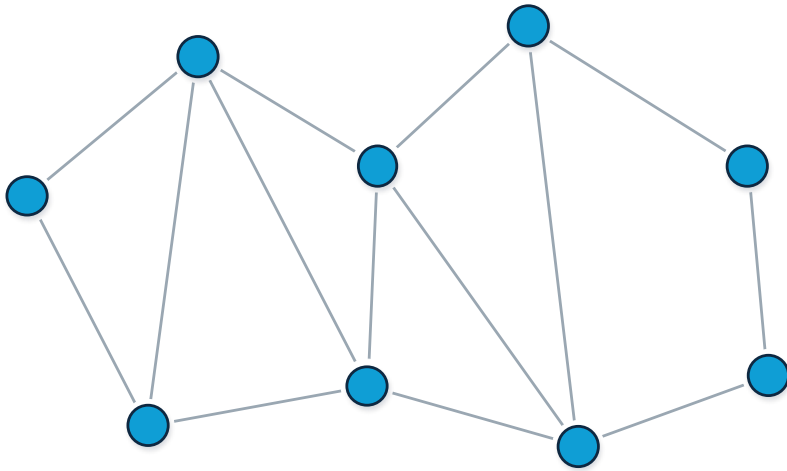
$$w_{ij} = K\left(\frac{\|x_i - x_j\|^2}{\xi^2}\right)$$

ξ sets the neighborhood scale. Convergence also needs scaling and density normalization.

Sampling creates the finite graph; geometry determines its edges and weights.

One response function can be shared by graph and manifold filters

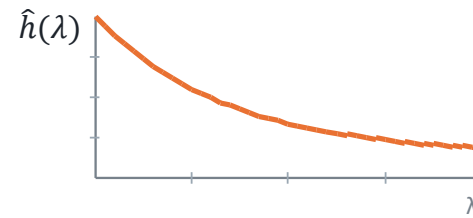
GRAPH REALIZATION



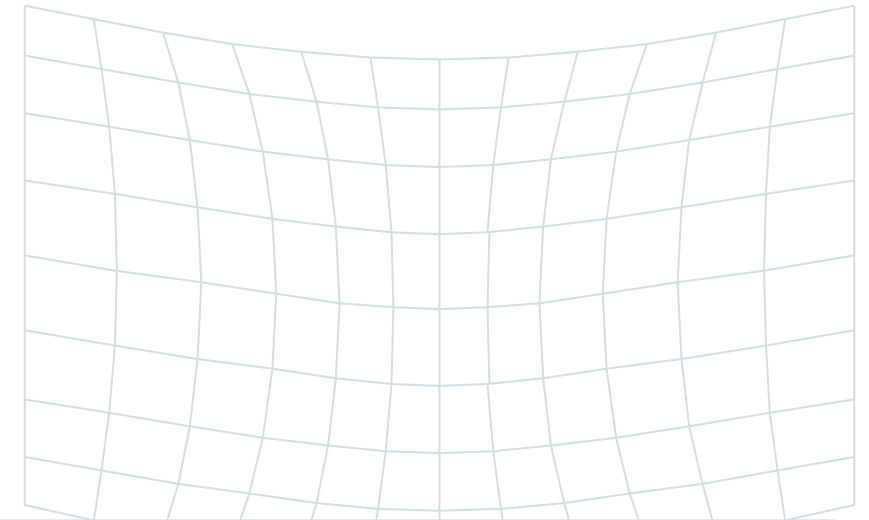
$$h(L_n)f_n = \sum_{i=1}^n \hat{h}(\lambda_{in}) \hat{f}_{in} \phi_{in}$$

MANIFOLD REALIZATION

SHARED RESPONSE $\hat{h}(\lambda)$



n grows

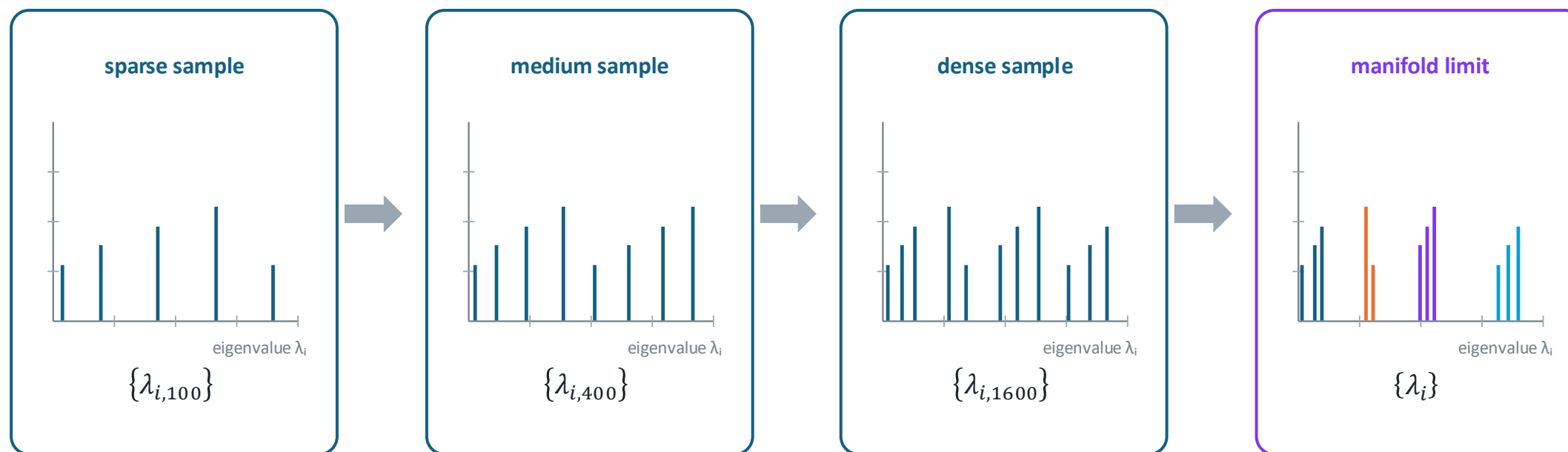


$$h(\mathcal{L})f = \sum_{i=1}^{\infty} \hat{h}(\lambda_i) \hat{f}_i \varphi_i$$

Compatible Laplacian normalization gives shared diffusion times the same meaning.

Transfer becomes possible because parameters describe a function, not a fixed matrix size.

Denser samples recover the manifold spectrum



operator convergence $\Rightarrow$ spectrum convergence $\Rightarrow$ filter convergence

Schematic convergence for fixed modes or controlled spectral bands under suitable assumptions.

A geometric graph filter approximates a fixed continuous computation.

One MNN layer filters, mixes channels, then rectifies

$$z_q^{(\ell)}(x) = \sum_{p=1}^{F_{\ell-1}} h_{qp}^{(\ell)}(\mathcal{L}) f_p^{(\ell-1)}(x)$$

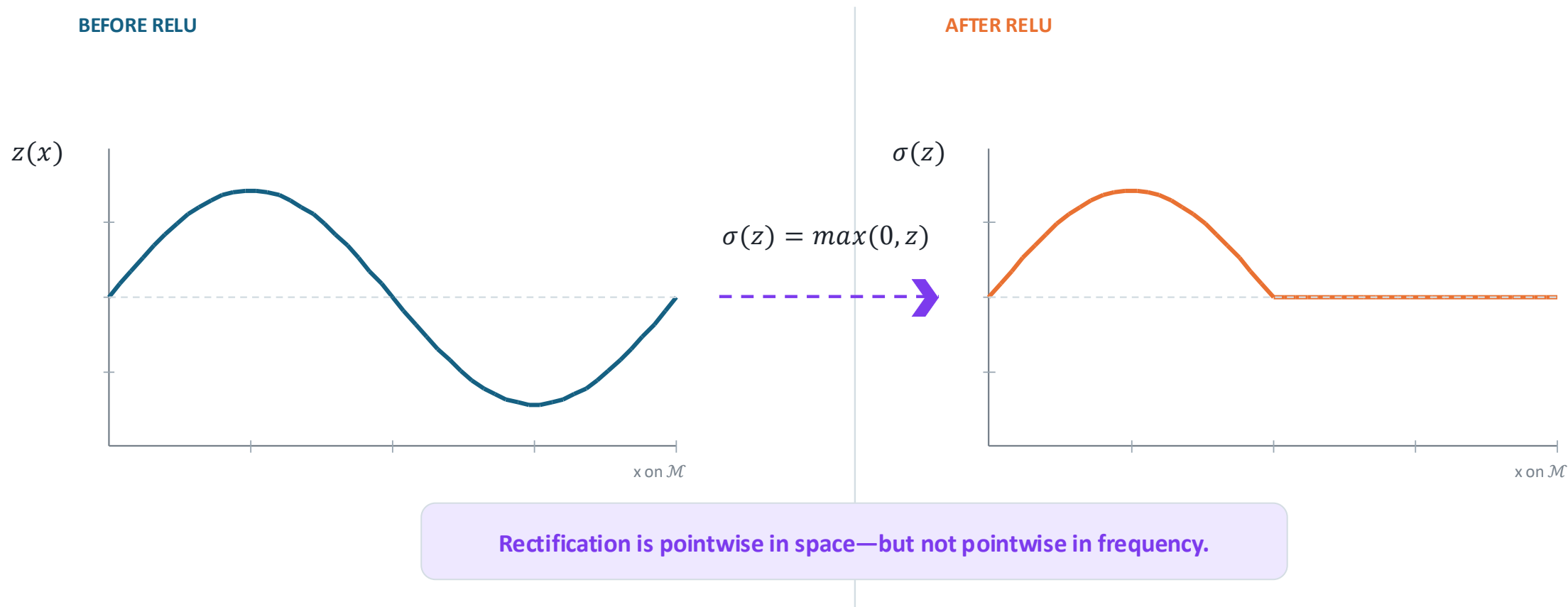
p: input channel q: output channel ℓ : layer σ : pointwise activation



The number of filters does not depend on the number of sampled nodes.

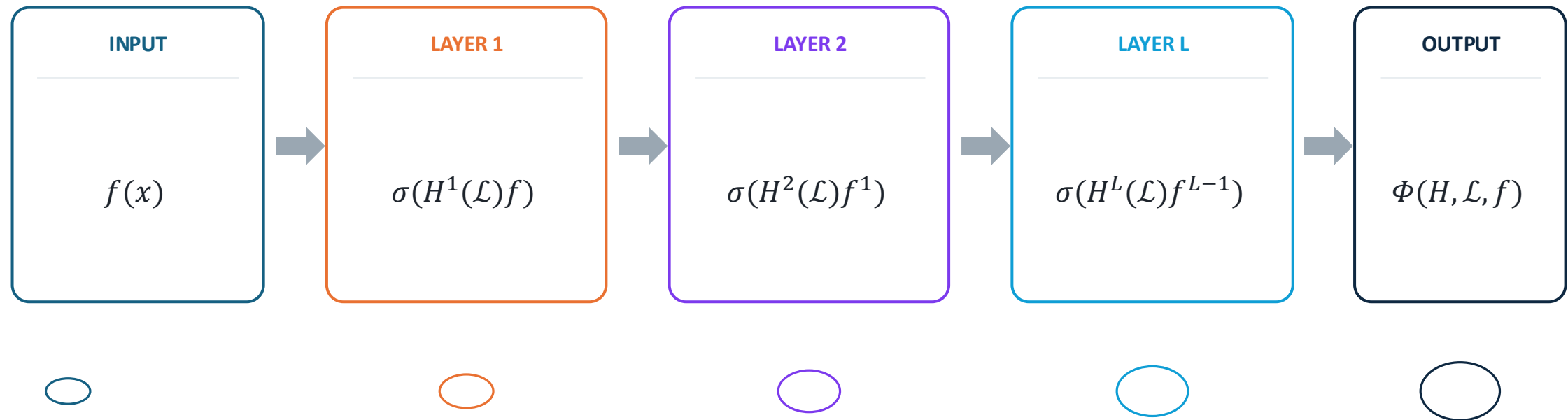
Geometry is handled by $h(\mathcal{L})$; feature interactions are handled by channel mixing.

ReLU creates new spectral mixtures



Nonlinearity lets the network recombine frequencies that a single linear filter cannot.

An MNN composes local geometric computations over depth



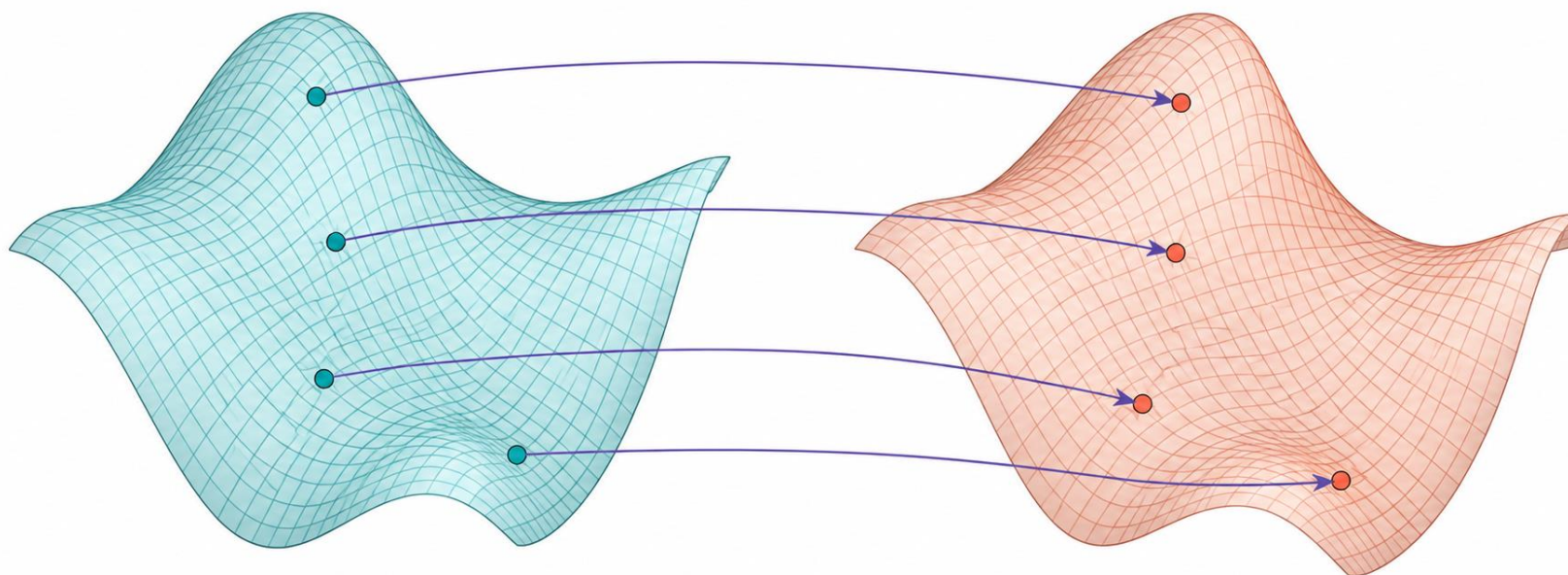
effective receptive field grows with diffusion time and layer depth

Depth builds multiscale receptive fields while parameters remain geometry-aware.

Stability controls output sensitivity to geometry

ORIGINAL DOMAIN $\mathcal{M}$

DEFORMED DOMAIN $\mathcal{M}'$



small deformation $\tau \Rightarrow$ small output change?

A stable model should follow the object, not overreact to its sampling or pose.

Deformation changes positions and geometry

A smooth correspondence τ maps each original point to its perturbed location.

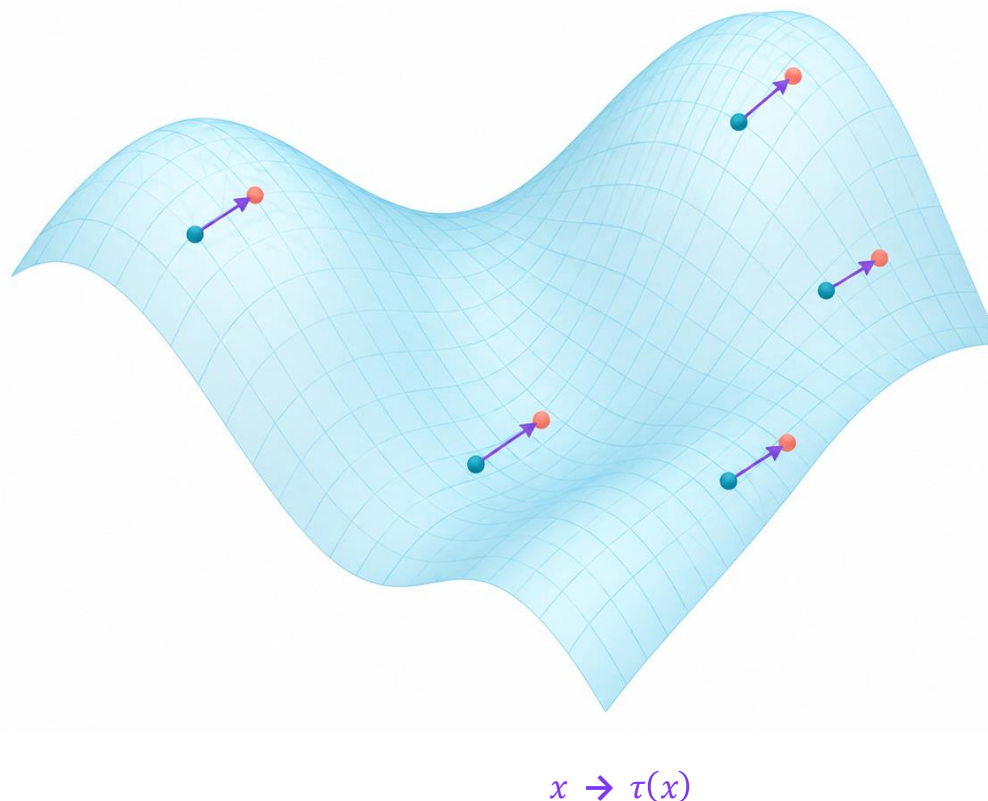
ε summarizes displacement and metric distortion

Small displacement alone is insufficient.
Control local metric distortion as well.

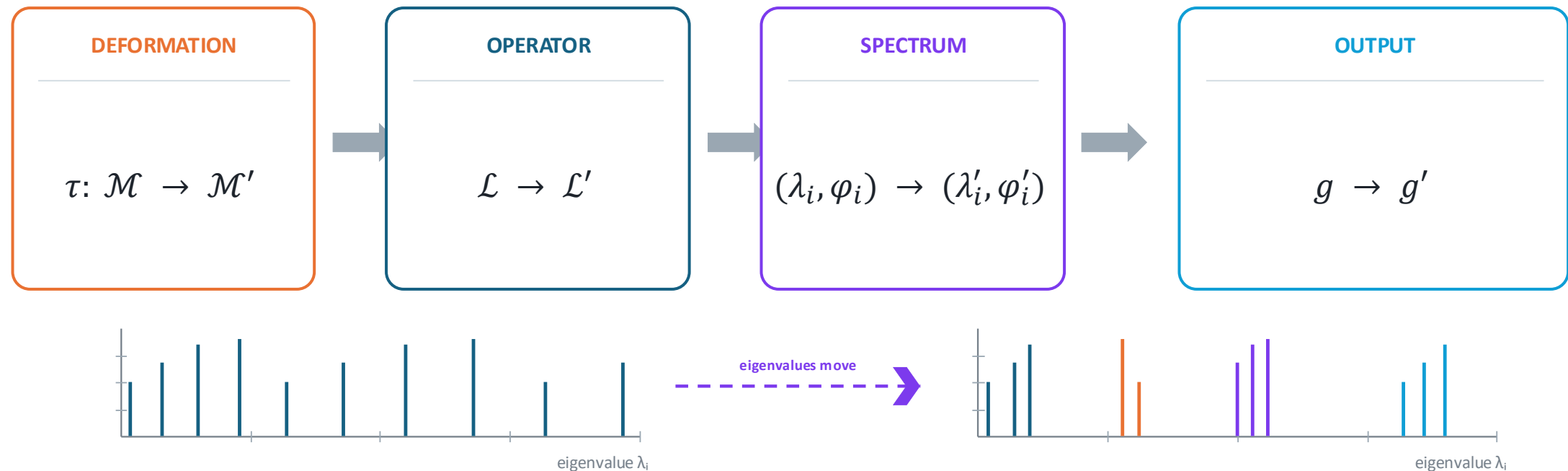
$$P_{\tau}g'(x) = g'(\tau(x))$$

Pullback puts both signals on the original domain.

Stability is meaningful only after defining how the domains correspond.



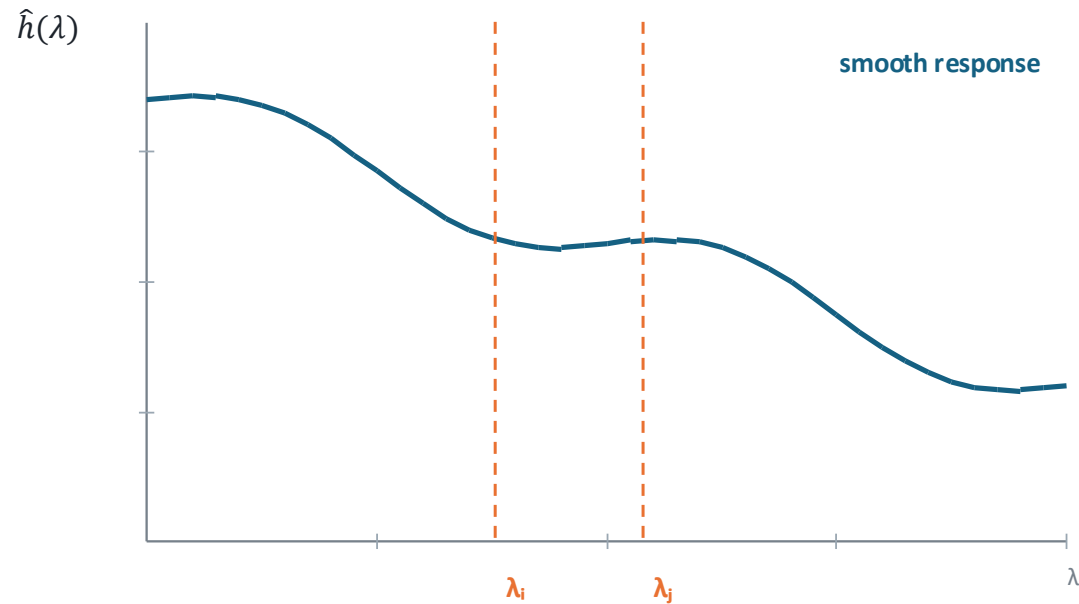
Geometry changes the Laplace–Beltrami operator and its spectrum



Both eigenvalues and eigenspaces change. Absolute and relative perturbations need different controls.

The proof challenge is to control how a learned spectral response reacts when frequencies move.

Lipschitz responses limit spectral sensitivity



$$|\hat{h}(\lambda_i) - \hat{h}(\lambda_j)| \leq A_h |\lambda_i - \lambda_j|$$

A_h is the maximum allowed slope.

Smaller A_h means less sensitivity when eigenvalues move.

This controls gain changes. A full stability result must also control eigenspace changes.

Smooth spectral responses convert small frequency perturbations into small gain perturbations.

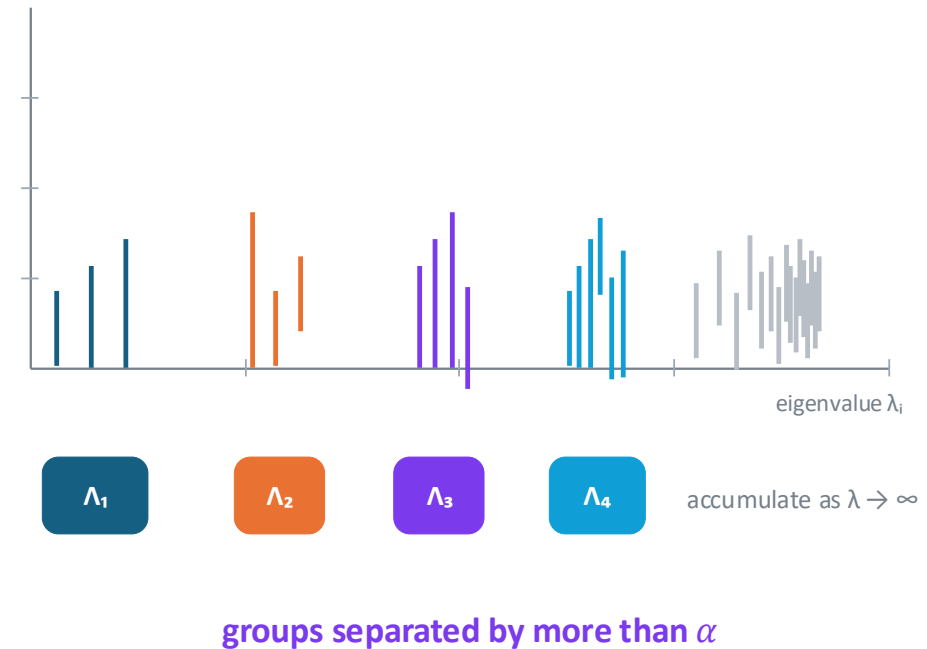
Spectral clusters make individual modes sensitive

Weyl's law describes spectral growth.
Gap behavior depends on dimension and multiplicity.

Problem: tracking every individual eigenvalue is fragile.

$$\lambda_i \sim C i^{\frac{2}{d}}$$

Repeated eigenvalues and tight clusters make individual eigenvectors unstable.



Group nearby frequencies; separate only the bands that geometry can resolve reliably.

An α -FDT filter is nearly constant inside each spectral group

$$|\hat{h}(\lambda_i) - \hat{h}(\lambda_j)| \leq \delta^D, \quad \lambda_i, \lambda_j \in \Lambda_k(\alpha)$$



FDT = frequency difference threshold. This condition supports absolute-perturbation stability.

FDT filters sacrifice fine spectral discrimination where the spectrum is unstable.

Stability needs alignment and regularity assumptions

$$\|\Phi(H, \mathcal{L}, f) - P_\tau \Phi(H, \mathcal{L}', f')\|_2 \leq LF^{L-1} \left(A_h + \pi \frac{N_\alpha}{\alpha} \right) \varepsilon \|f\|_2$$

Schematic bound: controlled perturbations, bounded gains, and Lipschitz activations

 ε

deformation size

 A_h

response slope

 α

spectral separation

 F, L

depth, width, gains

The input term vanishes for exact signal transport. General deformations require controls for both absolute and relative effects.

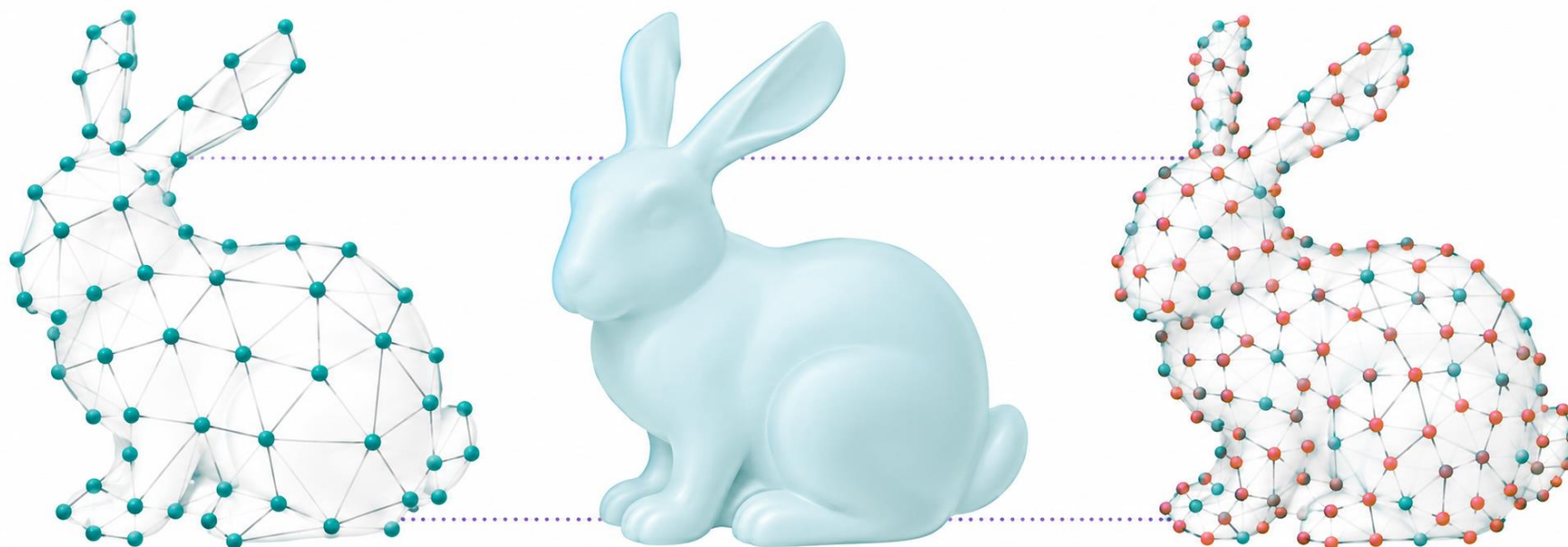
Stability controls sensitivity. It does not guarantee prediction accuracy.

Transfer across graph sizes uses one geometric reference

TRAIN · COARSE G_n

CONTINUOUS $\mathcal{M}$

DEPLOY · DENSE G_m



same H · same $\hat{h}(\lambda)$ · different sampling density

Compare through the continuous model after consistent sampling or interpolation.

Convergence to the same continuous computation explains cross-resolution transfer.

Point-cloud classification tests resolution transfer

MODELNET10 · TEST GRAPH HAS N = 1000 POINTS

Training n	Baseline GNN	Graph filter	GNN	Lipschitz GNN
300	16.95 ± 5.42	21.97 ± 4.17	10.10 ± 1.40	8.60 ± 2.95
500	13.11 ± 4.97	19.83 ± 5.94	7.74 ± 4.05	7.68 ± 3.75
700	10.02 ± 3.87	16.62 ± 2.38	7.92 ± 3.14	8.02 ± 2.77
900	6.83 ± 3.96	13.85 ± 3.81	7.45 ± 4.03	7.44 ± 3.30

Classification error (%) after training at n and testing at n = 1000 · lower is better

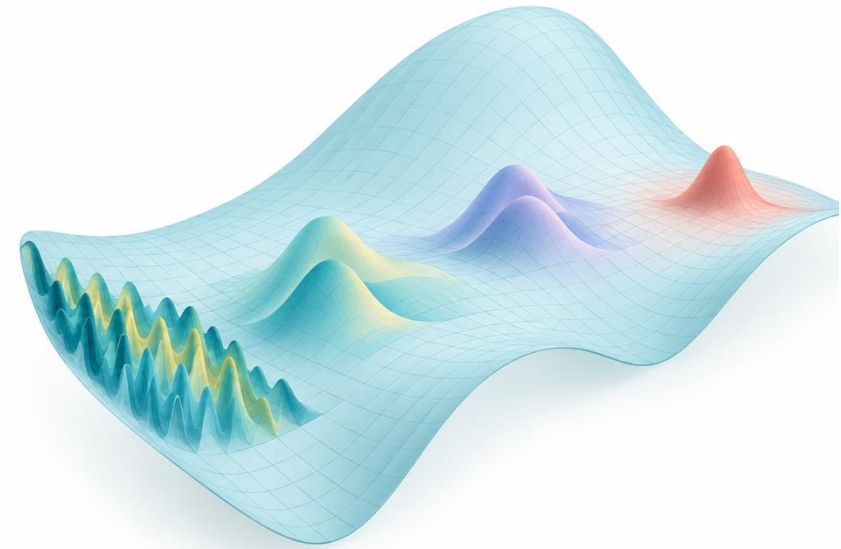
Variability is reproduced from the source. These entries alone do not establish statistical significance.

Regularized spectral responses transfer competitively even from substantially smaller graphs.

TAKEAWAYS

One geometric model. Many sampled graphs.

- 01** **DIFFUSE** Use $e^{-t}\mathcal{L}$ as the manifold shift.
- 02** **FILTER** Design the scalar response $\hat{h}(\lambda)$.
- 03** **LEARN** Compose filters, channels, and nonlinearities.
- 04** **CONTROL** Regularize spectral variation for stable transfer.



$$\mathcal{M} \rightarrow h(\mathcal{L}) \rightarrow \Phi(H, \mathcal{L}, f)$$

Continuous limits explain both scalability and stability.