On the Sparsity of Wavelet Coefficients for Signals on Graphs

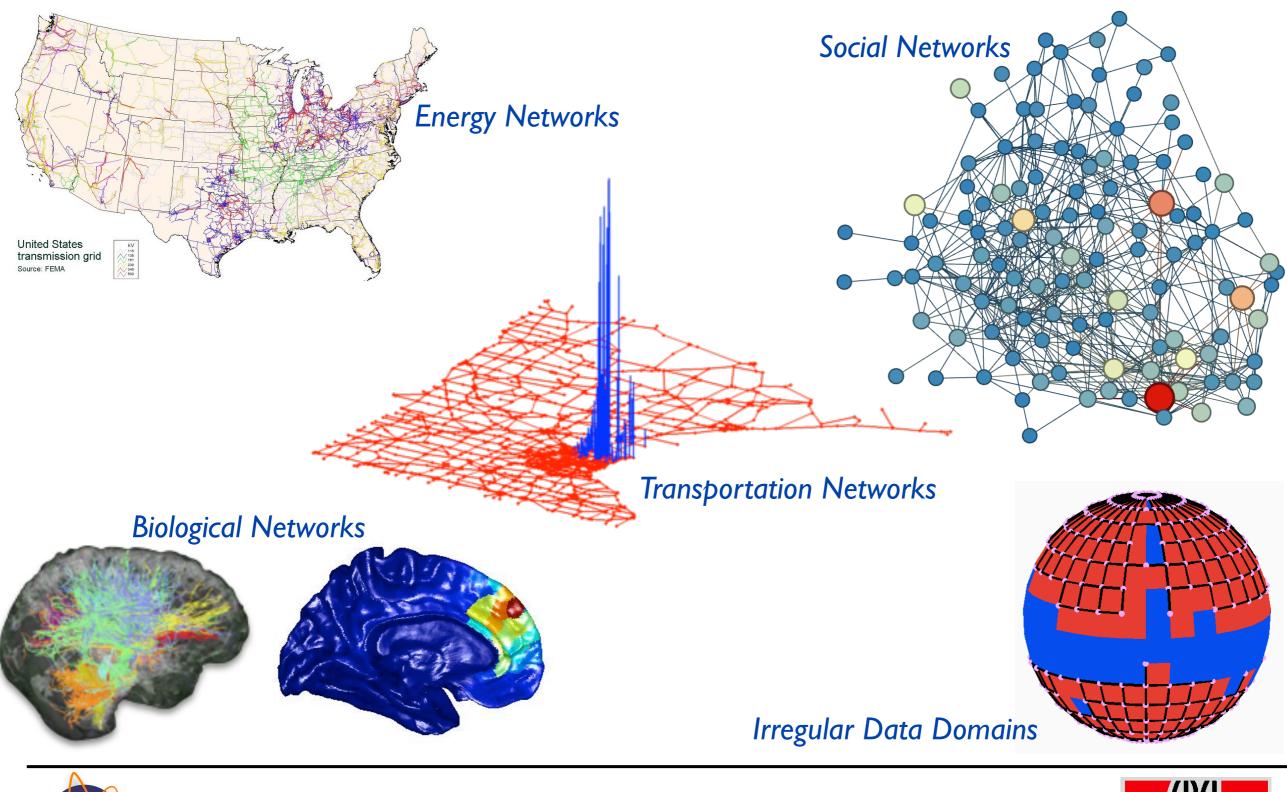
Benjamin Ricaud, David Shuman, and Pierre Vandergheynst

August 29, 2013 SPIE Wavelets and Sparsity XV San Diego, CA





Signal Processing on Graphs

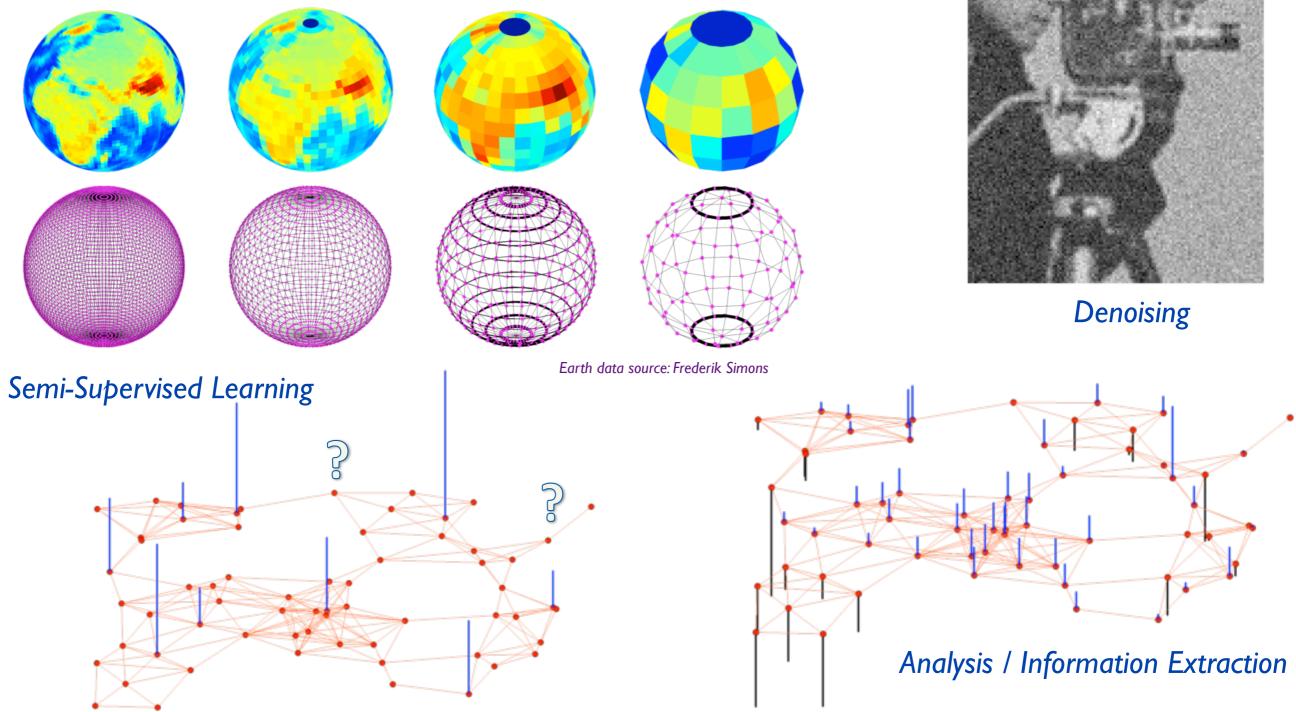






Some Typical Processing Problems

Compression / Visualization

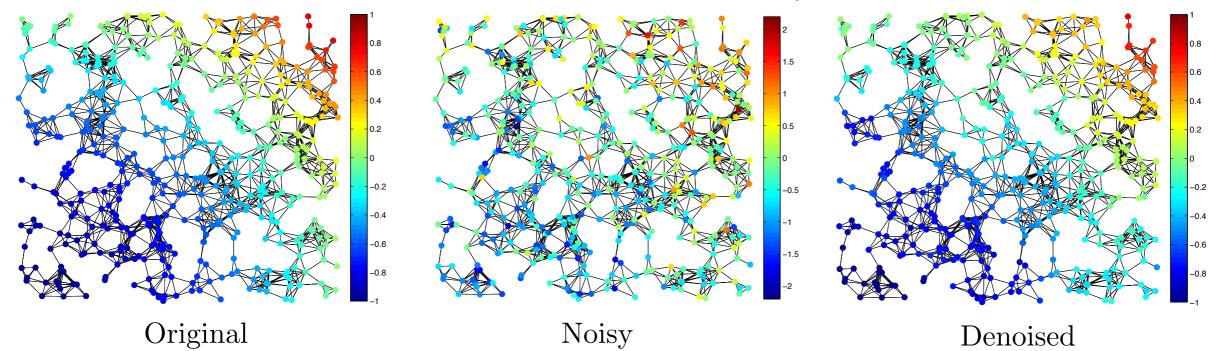






Simple Motivating Examples

• Tikhonov regularization for denoising: $\operatorname{argmin}_{f} \left\{ ||f - y||_{2}^{2} + \gamma f^{T} \mathcal{L} f \right\}$

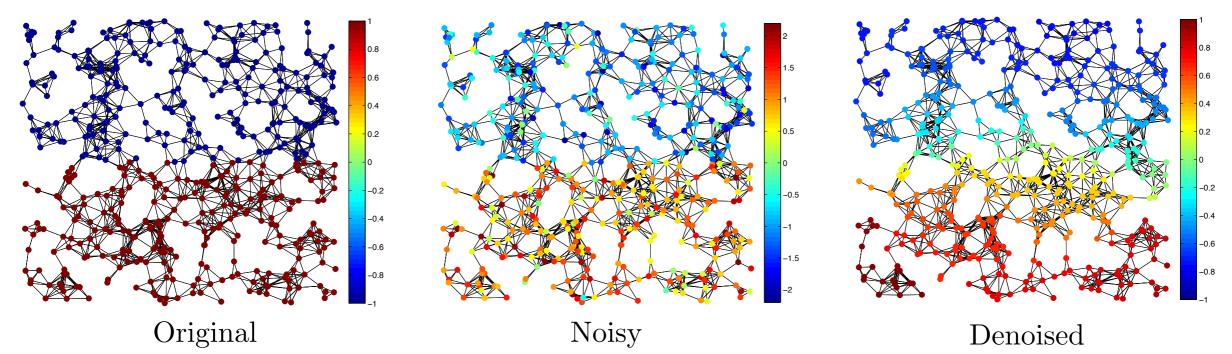




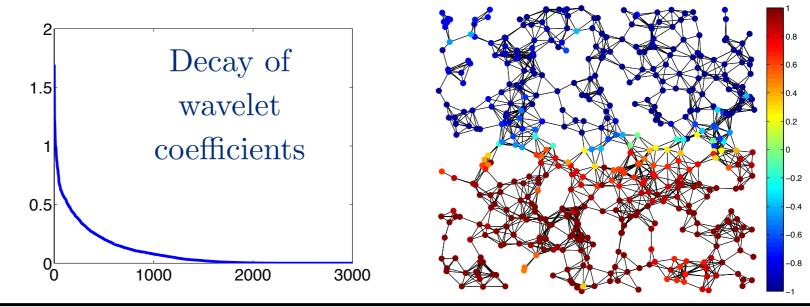


Simple Motivating Examples

• Tikhonov regularization for denoising: $\operatorname{argmin}_{f} \left\{ ||f - y||_{2}^{2} + \gamma f^{T} \mathcal{L} f \right\}$



• Wavelet denoising: $\operatorname{argmin}_{a} \left\{ ||f - W^*a||_2^2 + \gamma ||a||_{1,\mu} \right\}$







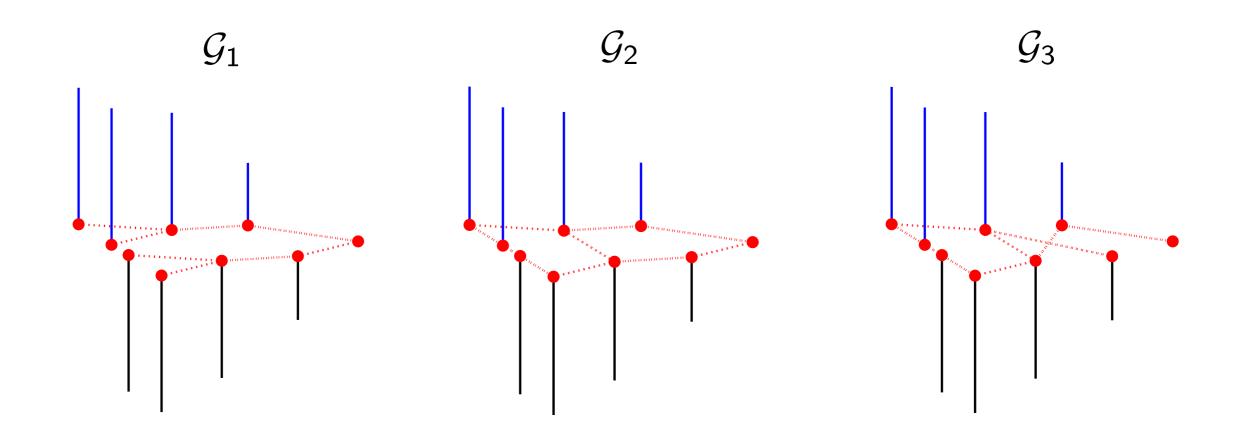
Research Questions

- For signals on Euclidean data domains, we have results characterizing classes of signals that are well-approximated by different transforms
 - e.g., piecewise-smooth 1D signals by wavelets, 2D cartoons with curvilinear discontinuities by curvelets/shearlets
- Which multiscale transforms for signals on graphs are wellsuited for which signal processing tasks, which classes of signals, and which types of graphs?
- Connections between properties of graph signals and the decay of their wavelet coefficients?





Smoothness of Graph Signals



To identify and exploit structure in the data, we need to account for the intrinsic geometric structure of the underlying graph data domain





Global Regularity of Graph Signals

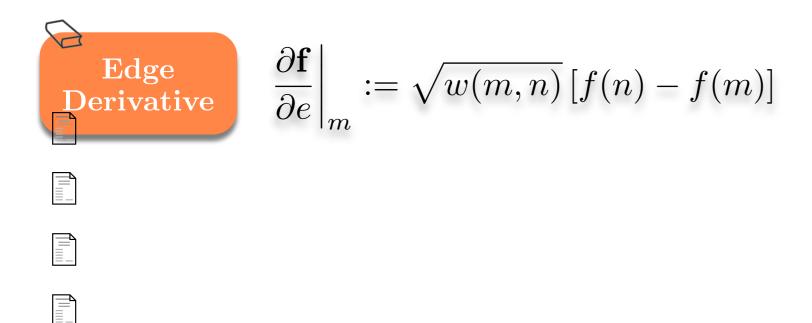


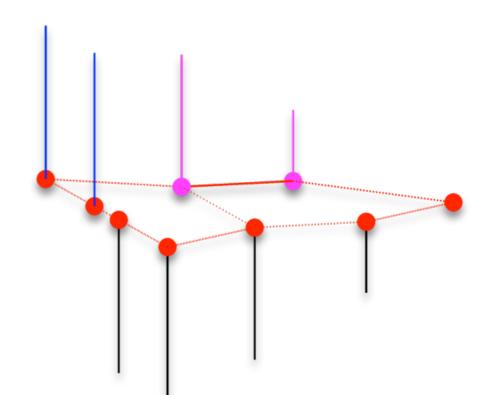


7

Notions of Global Regularity for Graph Signals

Discrete Calculus, Grady and Polimeni, 2010







Notions of Global Regularity for Graph Signals

Discrete Calculus, Grady and Polimeni, 2010

Edge
Derivative
$$\frac{\partial \mathbf{f}}{\partial e}\Big|_{m} := \sqrt{w(m,n)} [f(n) - f(m)]$$
Graph
Gradient $\nabla_{m} \mathbf{f} := \left[\left\{\frac{\partial \mathbf{f}}{\partial e}\Big|_{m}\right\}_{e \in \mathcal{E} \text{ s.t. } e=(m,n)}\right]$ Local
Variation $||\nabla_{m} \mathbf{f}||_{2} = \left[\sum_{n \in \mathcal{N}_{m}} w(m,n) [f(n) - f(m)]^{2}\right]^{\frac{1}{2}}$



 \equiv _

8

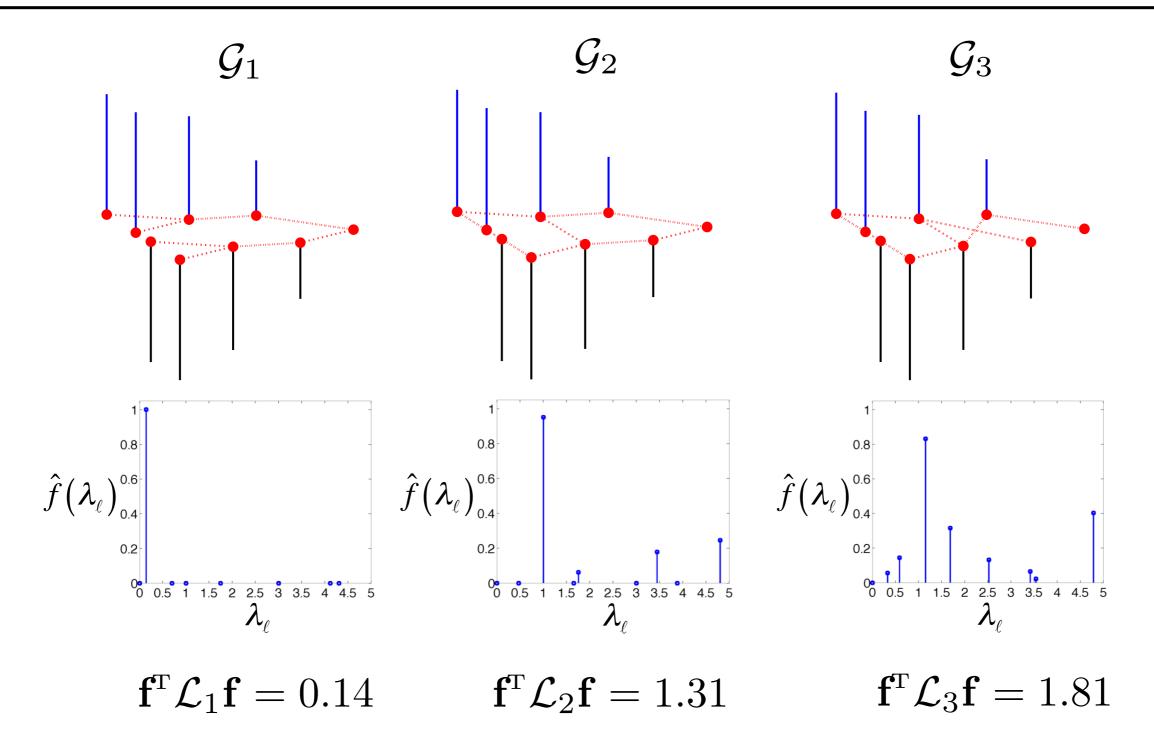
Notions of Global Regularity for Graph Signals

Discrete Calculus, Grady and Polimeni, 2010 $\left. \frac{\partial \mathbf{f}}{\partial e} \right|_{m} := \sqrt{w(m,n)} \left[f(n) - f(m) \right]$ Edge Derivative $\nabla_m \mathbf{f} := \left| \left\{ \left. \frac{\partial \mathbf{f}}{\partial e} \right|_m \right\}_{e \in \mathcal{E} \text{ s.t. } e = (m, n)} \right|$ Graph Gradient $\left\| \nabla_m \mathbf{f} \right\|_2 = \left\| \sum_{n \in \mathcal{N}_m} w(m, n) \left[f(n) - f(m) \right]^2 \right\|^{\frac{1}{2}}$ Local Variation $\frac{1}{2}\sum_{m\in V} ||\nabla_m \mathbf{f}||_2^2 = \sum_{(m,n)\in\mathcal{E}} w(m,n) \left[f(n) - f(m)\right]^2 = \mathbf{f}^{\mathrm{T}} \mathcal{L} \mathbf{f}$ Quadratic Form $||\mathbf{f}||_{\mathcal{L}} := ||\mathcal{L}^{\frac{1}{2}}\mathbf{f}||_2 = \sqrt{\mathbf{f}^{\mathrm{T}}\mathcal{L}\mathbf{f}}$





Smoothness of Graph Signals Revisited







Notions of Global Regularity for Graph Signals^{*} *Generalizations*

p-Dirichlet Form (Elmoataz et al., 2008)

$$\frac{1}{p}\sum_{m\in V} ||\nabla_m \mathbf{f}||_2^p = \frac{1}{p}\sum_{m\in \mathcal{V}} \left[\sum_{n\in\mathcal{N}_m} w(m,n) \left[f(n) - f(m)\right]^2\right]^{\frac{p}{2}}$$

Discrete Sobolev Semi-Norm

$$\|f\|_{\mathcal{H}^p} := \|\mathcal{L}^p f\|_2 = \|\widehat{\mathcal{L}^p f}\|_2 = \sqrt{\sum_{\ell} |\lambda_{\ell}|^{2p} |\widehat{f}(\ell)|^2}$$

• In the continuous setting, the space $\mathbb{W}^p(\mathbb{R})$ of *p*-times differentiable Sobolev functions are those satisfying

$$\int_{-\infty}^{\infty} |\omega|^{2p} |\hat{f}(\omega)|^2 d\omega < \infty$$



• In the graph setting,

$$\frac{\|f\|_{\mathcal{H}^p}}{\|f\|_2} \le \lambda_{\max}^p \text{ for all } f \in \mathbb{R}^N$$



EPFL – Signal Processing Laboratory (LTS2) http://lts2.epfl.ch



 \boldsymbol{n}

Wavelet Coefficient Decay of Globally Regular Graph Signals

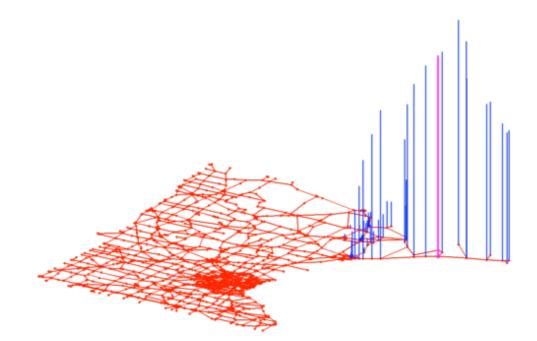




Hammond et al., Wavelets on graphs via spectral graph theory, ACHA, 2011

• Generalized translation

$$\begin{array}{l} \searrow \text{ Classical setting: } (T_s g)(t) = g(t-s) = \int_{\mathbb{R}} \hat{g}(\xi) e^{-2\pi i \xi s} e^{2\pi i \xi t} d\xi \\ & \searrow \end{array} \\ \text{Graph setting: } (T_n g)(i) := \sum_{\ell=0}^{N-1} \hat{g}(\lambda_\ell) u_\ell^*(n) u_\ell(i) \end{aligned}$$



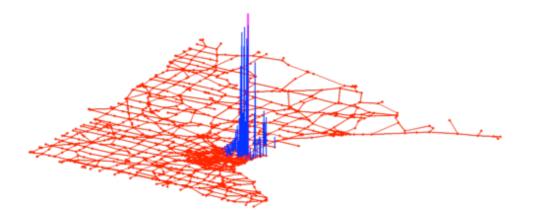




Hammond et al., Wavelets on graphs via spectral graph theory, ACHA, 2011

• Generalized translation

$$\begin{array}{l} \searrow \text{ Classical setting: } (T_s g)(t) = g(t-s) = \int_{\mathbb{R}} \hat{g}(\xi) e^{-2\pi i \xi s} e^{2\pi i \xi t} d\xi \\ & \searrow \end{array} \\ \text{Graph setting: } (T_n g)(i) := \sum_{\ell=0}^{N-1} \hat{g}(\lambda_\ell) u_\ell^*(n) u_\ell(i) \end{aligned}$$





1

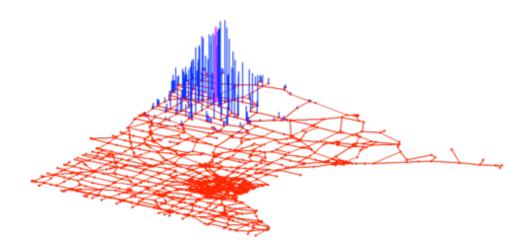
-



Hammond et al., Wavelets on graphs via spectral graph theory, ACHA, 2011

• Generalized translation

$$\begin{array}{l} \searrow \text{ Classical setting: } (T_s g)(t) = g(t-s) = \int_{\mathbb{R}} \hat{g}(\xi) e^{-2\pi i \xi s} e^{2\pi i \xi t} d\xi \\ & \searrow \end{array} \\ \text{Graph setting: } (T_n g)(i) := \sum_{\ell=0}^{N-1} \hat{g}(\lambda_\ell) u_\ell^*(n) u_\ell(i) \end{aligned}$$





1

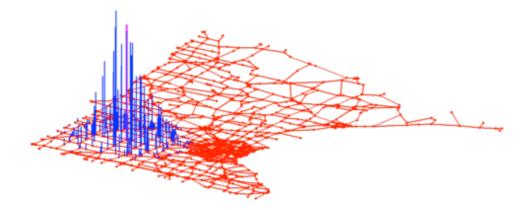
-



Hammond et al., Wavelets on graphs via spectral graph theory, ACHA, 2011

• Generalized translation

$$\begin{array}{l} \searrow \text{ Classical setting: } (T_s g)(t) = g(t-s) = \int_{\mathbb{R}} \hat{g}(\xi) e^{-2\pi i \xi s} e^{2\pi i \xi t} d\xi \\ & \searrow \end{array} \\ \text{Graph setting: } (T_n g)(i) := \sum_{\ell=0}^{N-1} \hat{g}(\lambda_\ell) u_\ell^*(n) u_\ell(i) \end{aligned}$$





1

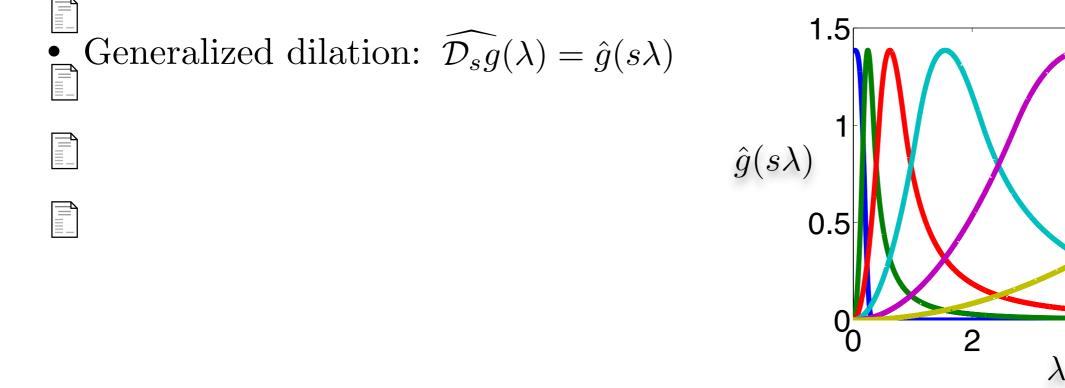


Hammond et al., Wavelets on graphs via spectral graph theory, ACHA, 2011

• Generalized translation

Classical setting:
$$(T_s g)(t) = g(t - s) = \int_{\mathbb{R}} \hat{g}(\xi) e^{-2\pi i \xi s} e^{2\pi i \xi t} d\xi$$

Graph setting: $(T_n g)(i) := \sum_{\ell=0}^{N-1} \hat{g}(\lambda_\ell) u_\ell^*(n) u_\ell(i)$





EPFL – Signal Processing Laboratory (LTS2) http://lts2.epfl.ch



4

6

the underlying graph, we include a regularization term of the form problem 14

$\operatorname*{argmin}_{\mathbf{f}} \left\{ \|\mathbf{f} - \mathbf{y}\|_2^2 + \gamma \right.$

Spectral Graph Wavelets

The first-order optimality conditions of the convex objective function [?, Proposition 1]) the optimal reconstruction is given by

Hammond et al., Wavelets on graphs via spectral graph theory, $ACH_{f_*(i)} = 2 \sum_{\ell=0}^{N-1} \left[\frac{1}{1+\gamma\lambda_\ell} \right] \hat{y}$

• Generalized translation

Classical setting:
$$(T_s g)(t) = g(t-s) = \int_{\mathbb{R}} \hat{g}(\xi) e_{me}^{G_{\ell}}$$

Graph setting: $(T_n g)(i) := \sum_{\ell=0}^{N-1} \hat{g}(\lambda_{\ell}) u_{\ell}^*(n) u_{\ell}(i)$

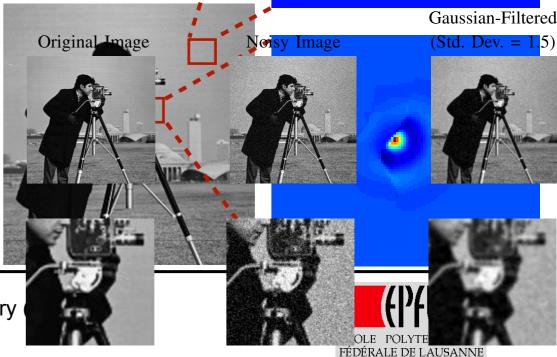
Generalized dilation:
$$\widehat{\mathcal{D}_s g}(\lambda) = \widehat{g}(s\lambda)$$

Spectral graph wavelet at scale s, centered at vertex n:

$$\psi_{s,n}(i) := (T_n D_s g)(i) = \sum_{\ell=0}^{N=1} \hat{g}(s\lambda_\ell) u_\ell^*(n) u_\ell(i)$$

or, equivalently, $\mathbf{f} = \hat{h}(\mathcal{L})\mathbf{y}$, where $\hat{h}(\lambda) := \frac{1}{1+\gamma\lambda}$ can be viewed as As an example, in the figure below, we take the 512 x 512 can GQuasign no Quarwith mean zero and standard deviation 0.1 to get a methods to denoise the signal. In the first method, we apply a sym size 72 x 72 with two different standard deviations: 1.5 and 3.5. In the pixels by connecting each pixel to its horizontal, vertical, and de (?) between two neighboring pixels according to the similarity of edges of the semi-local graph are independent of the noisy image, b@urai-fhoreigh@riag.pixel values in the noisy image. For the Ge We then perform the low-pass graph filtering (??) to reconstruct the anisotropic diffusion image smoothing method of [?].

In all image displays, we threshold the values to the [0,1] in zoomed versions of the top row of images. Comparing the results smooth sufficiently in smoother areas of the image, the classical G The graph spectral filtering method does not smooth as much across image is encoded in the graph Laplacian via the noisy image.





Wavelet Coefficient Decay of Globally Regular ¹⁵ Graph Signals

Proposition 1 Let $p \ge 1$, and assume that $C_p := \int_0^\infty |\hat{g}(s)|^2 / s^{2p} ds < \infty$. Then $\int_0^\infty s^{-2p} \sum_n |\langle f, \psi_{s,n} \rangle|^2 ds = C_p ||f||_{\mathcal{H}^{(2p-1)/2}}.$



Assume that $\hat{g}(\lambda) = \sum_{k=p}^{q} a_k \lambda^k$ for some $p \ge 1$ (implying $\hat{g} = 0$) Then $|\Psi f(s, n)| = |\langle f, \psi_{s, n} \rangle| \le \sum_{k=p}^{q} |a_k| s^k ||f||_{\mathcal{H}^k}.$

k = p





Ongoing Work: Local Regularity and Wavelet Coefficient Decay of Locally Regular Graph Signals





Notions of Local Regularity

Local Variation

$$|\nabla_m \mathbf{f}||_2 = \left[\sum_{n \in \mathcal{N}_m} w(m, n) \left[f(n) - f(m)\right]^2\right]^{\frac{1}{2}}$$

Hölder Regularity A graph signal f is (C, α, r) -Hölder regular with respect to the graph \mathcal{G} at vertex $n \in \mathcal{V}$ if

$$|f(n) - f(m)| \le C[d_{\mathcal{G}}(m,n)]^{\alpha}, \ \forall m \in \mathcal{N}(n,r)$$

Gavish et al. ICML, 2010

9

Laplacian as Derivative $(\mathcal{L}^k f)(n)$ as a measure of local regularity of f in a neighborhood of radius k around vertex $n \searrow$

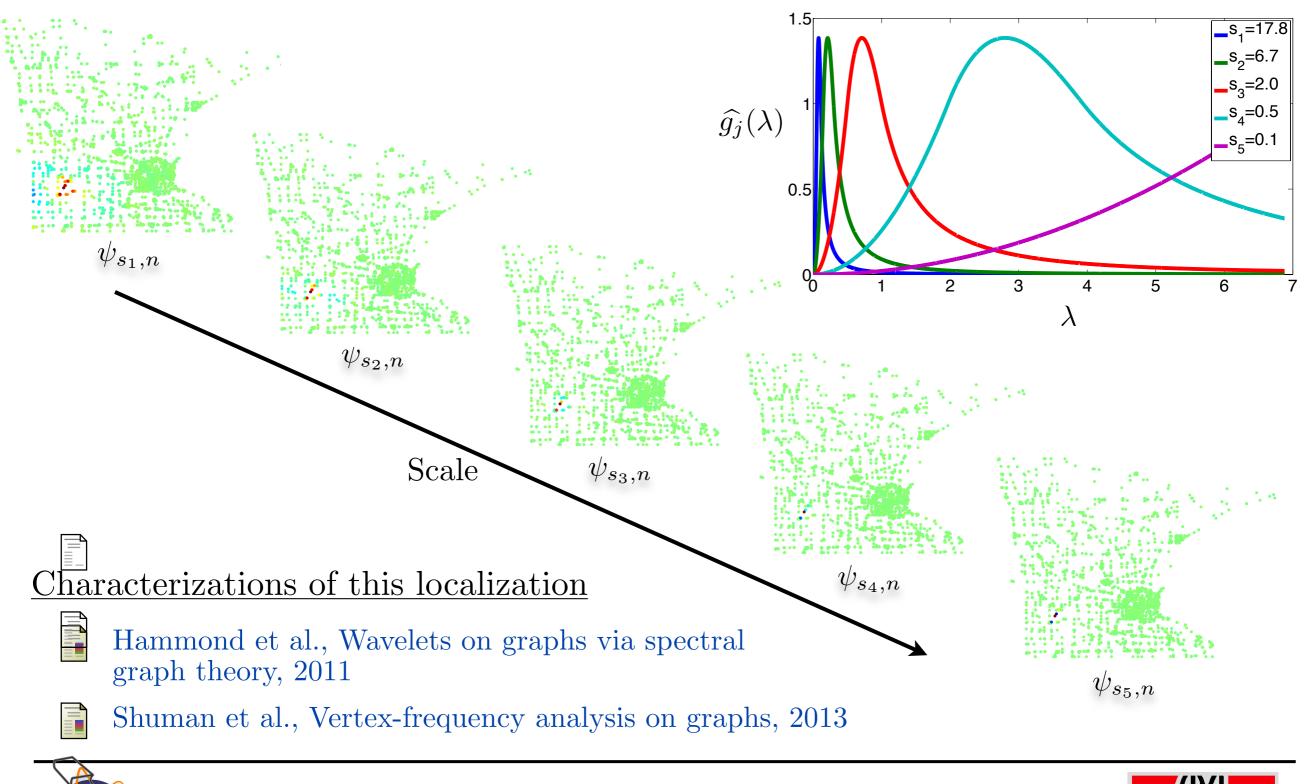
• For polynomial kernel:

$$\Psi f(s,n) = \sum_{k=p}^{q} a_k s^k (\mathcal{L}^k f)(n)$$





Spectral Graph Wavelet Localization







Wavelet Coefficient Decay of Locally Regular Graph Signals

High-level intuition

- Far away from vertex n, for small scales s, $|\Psi f(s, n)|$ is small because $\psi_{s,n}$ is highly localized around n
- Close to vertex $n, |\Psi f(s, n)|$ is small because f is locally regular

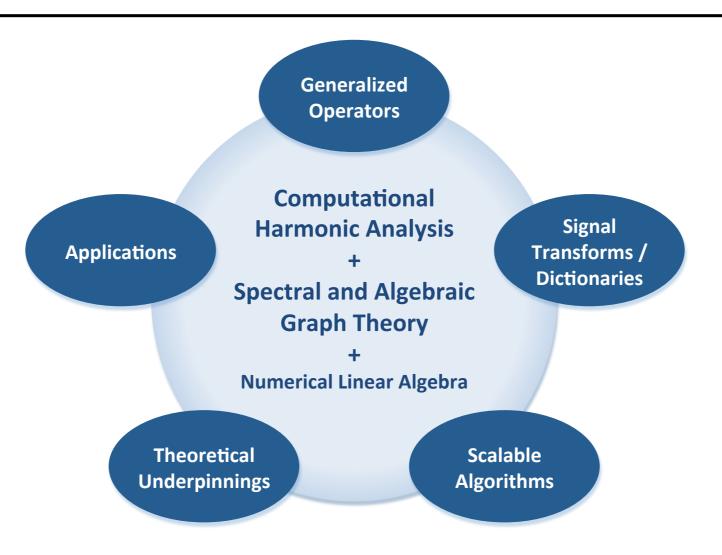
Proposition 3

Assume that f is (C, α, r) -Hölder regular for some $r \ge 1$, and let $\hat{g}(\lambda) = \sum_{k=r}^{q} a_k \lambda^k$ for some coefficients $\{a_k\}_{k=r,r+1,\ldots,q}$. Then there exist constants C_2 and \bar{s} such that for all $s < \bar{s}$, we have

$$\Psi f(s,n) \leq Cr^{\alpha} \sum_{m \in \mathcal{N}(n,r)} |\psi_{s,n}(m)| + C_2 s^{r+1} \sum_{m \notin \mathcal{N}(n,r)} |f(m) - f(n)|.$$



Outlook



- Application of graph signal processing techniques to real science and engineering problems is in its infancy
- Theoretical connections between classes of graph signals, the underlying graph structure, and sparsity of transform coefficients



