Signal Processing on Graphs:

Extending High-Dimensional Data Analysis to Networks and Other Irregular Domains

David Shuman

Ecole Polytechnique Fédérale de Lausanne (EPFL)
david.shuman@epfl.ch

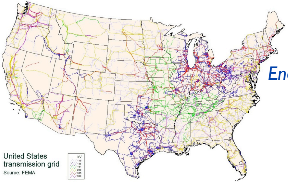
Macalester College
February 11, 2013

Special thanks and acknowledgment to my collaborators:

Xiaowen Dong, Mohammad Javad Faraji, Pascal Frossard, Jason McEwen, Daniel Kressner, Sunil Narang, Antonio Ortega, Javier Pérez-Trufero, Nathanaël Perraudin, Benjamin Ricaud, Dorina Thanou, and Pierre Vandergheynst

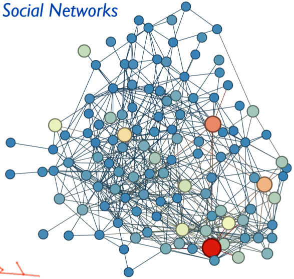
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

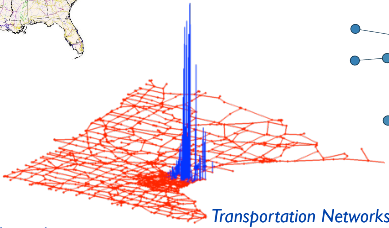
Signal Processing on Graphs



Energy Networks

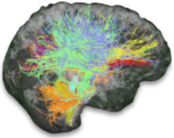
Social Networks

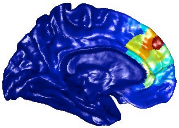




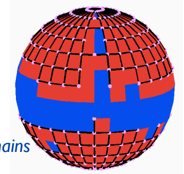
Transportation Networks

Biological Networks





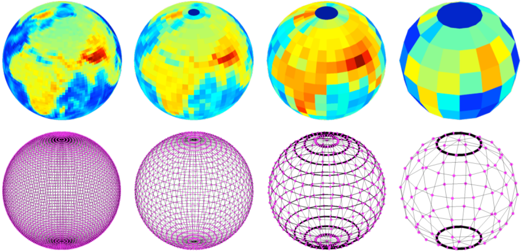
Irregular Data Domains



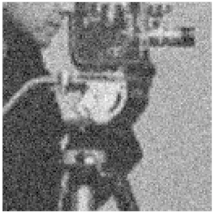
Weighted graphs are a flexible tool to represent a wide variety of topologically-complicated data domains

Some Typical Processing Problems

Visualization / Compression

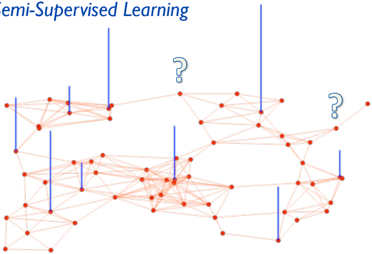


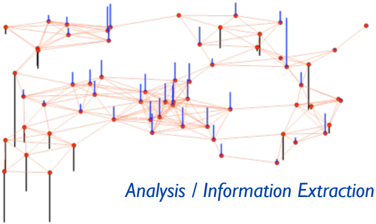
Earth data source: Frederik Simons



Denoising

Semi-Supervised Learning

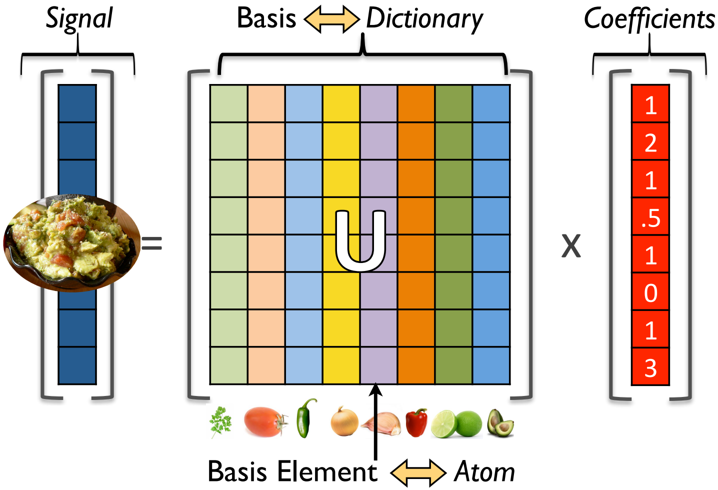




Analysis / Information Extraction

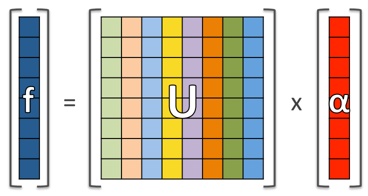
Outline

Orthonormal Dictionaries

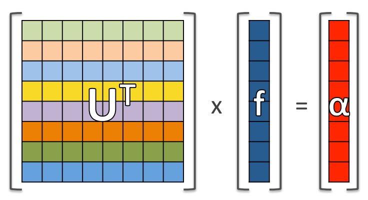


Orthonormal Dictionaries (cont.)

Synthesis



Analysis



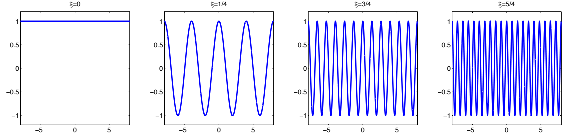
$$f = \sum_{\ell} \alpha_{\ell} u_{\ell} = \sum_{\ell} \langle f, u_{\ell} \rangle u_{\ell}$$

Overcomplete Dictionaries and Sparsity

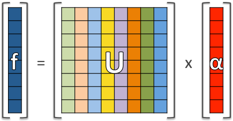
- Given an overcomplete Φ, there are infinitely many choices of α that lead to the same f
- Criteria for useful dictionaries?
  - Ability to *sparsely* represent signals – few non-zeros coefficients in α
  - Ability to capture the relevant characteristics of the signal
  - Computationally efficient to apply Φ and Φ^{T}
- Which signals? Need different dictionaries for different mathematical models of data with different structural patterns and for different processing tasks
  - E.g., different ingredient bundles for different cuisines

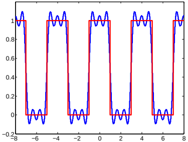
The Fourier Transform

- Collection of atoms: $\left\{e^{2\pi i \xi t}\right\}_{\xi \in \mathbb{R}} = \left\{\cos(2\pi\xi t) + i \sin(2\pi \xi t)\right\}_{\xi \in \mathbb{R}}$

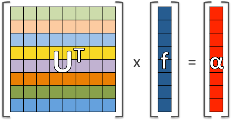


- Synthesis example:





- Fourier transform (analysis): $\hat{f}(\xi) = \langle f, e^{2\pi i \xi t}\rangle = \int_{\mathbb{R}} f(t) e^{-2\pi i \xi t}\, dt$



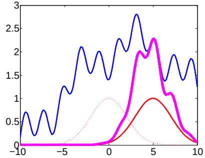
- e.g., if f is an audio signal, $\hat{f}$ tells us which frequencies are present
- A music expert could identify rhythmical patterns from f and the key from dominant frequencies of $\hat{f}$

Time-Frequency Dictionaries

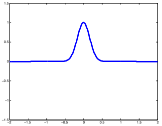
- Localized Fourier analysis – joint descriptions of signals' temporal and spectral behavior
  - Localized oscillations appear frequently in audio processing, vibration analysis, radar detection, etc.
  - e.g., identify musical notes and melody at different times

- Windowed (short-time) Fourier transform of $f \in L^2(\mathbb{R})$:

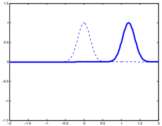
$$Sf(s,\xi) := \langle f, g_{s,\xi} \rangle = \int_{-\infty}^{\infty} f(t)\overline{g(t-s)}e^{-2\pi i \xi t}dt$$



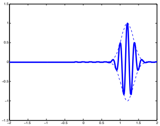
- The atoms $g_{s,\xi}$ are localized in time and frequency:



Translation T_s $\Longrightarrow$



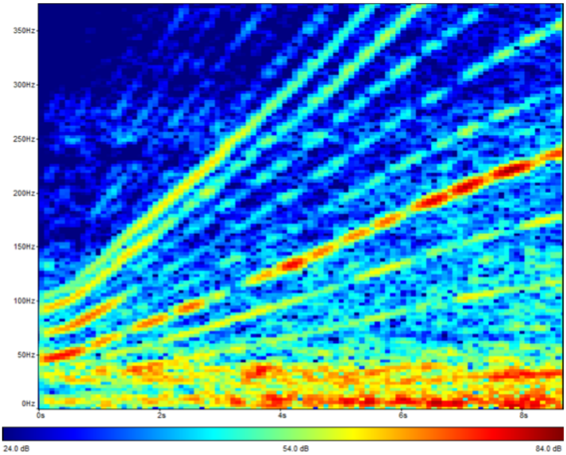
Modulation M_ξ $\Longrightarrow$



Spectrogram Example

$$|Sf(s,\xi)|^2$$

Frequency



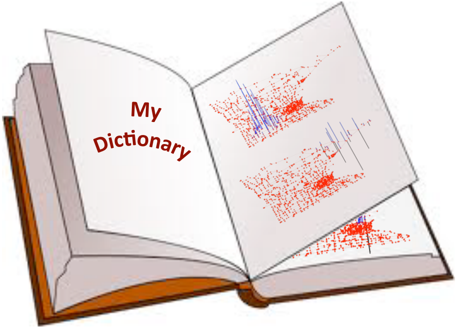
Time

Source: Genesis

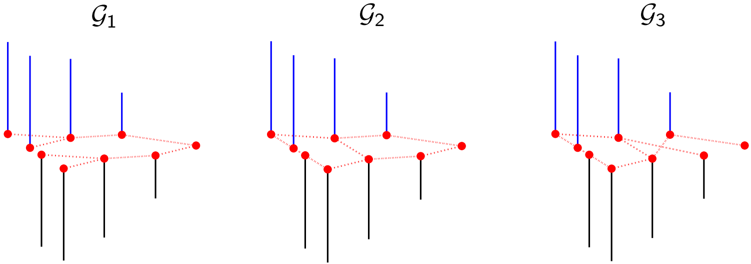
Dictionary Design for Signals on Graphs

Main Problem

Design dictionaries to (i) reveal relevant structural properties of signals on weighted, undirected graphs, and/or (ii) sparsely represent different classes of signals on graphs

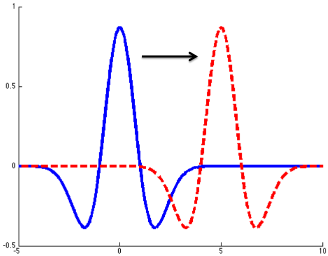


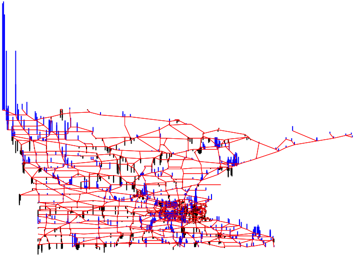
Why Do We Need New Dictionaries?



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

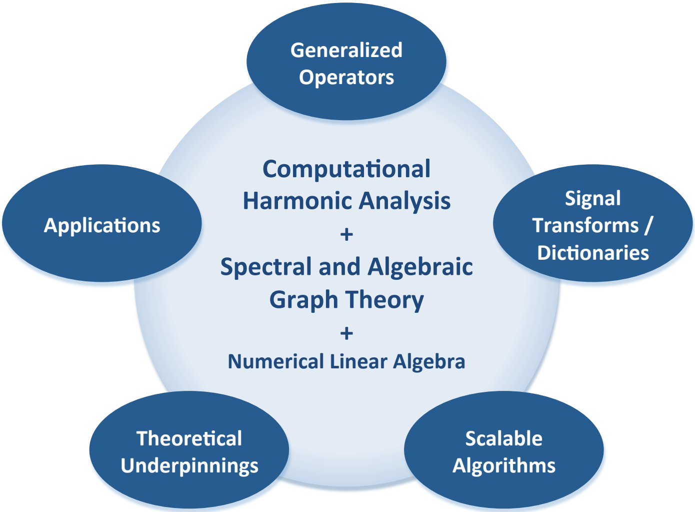
The Essence of the Problem





- Weighted graphs are irregular structures that lack a shift-invariant notion of translation
- Many simple yet fundamental concepts that underlie classical signal processing techniques become significantly more challenging in the graph setting

Approach


Generalized Operators
Applications
Computational Harmonic Analysis
+
Spectral and Algebraic Graph Theory
+
Numerical Linear Algebra
Signal Transforms / Dictionaries
Theoretical Underpinnings
Scalable Algorithms

Outline

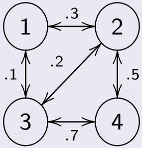
Spectral Graph Theory Notation

- Connected, undirected, weighted graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, W\}$
- Degree matrix D: zeros except diagonals, which are sums of weights of edges incident to corresponding node
- Non-normalized graph Laplacian: $\mathcal{L} := D - W$
- Complete set of orthonormal eigenvectors and associated real, non-negative eigenvalues:

$$\mathcal{L}u_\ell = \lambda_\ell u_\ell,$$

ordered w.l.o.g. s.t.

$$0 = \lambda_0 < \lambda_1 \leq \lambda_2 ... \leq \lambda_{N-1} := \lambda_{\max}$$



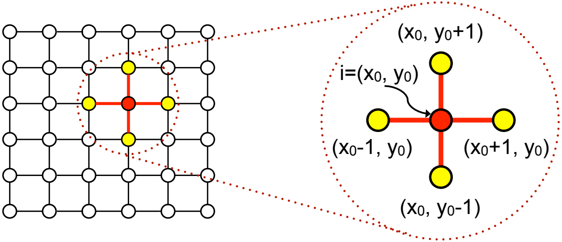
$$W = \begin{bmatrix} 0 & .3 & .1 & 0 \\ .3 & 0 & .2 & .5 \\ .1 & .2 & 0 & .7 \\ 0 & .5 & .7 & 0 \end{bmatrix}$$

$$D = \begin{bmatrix} .4 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1.2 \end{bmatrix}$$

The Graph Laplacian

- Discrete difference operator: $(\mathcal{L}f)(i) = \sum_{j \in \mathcal{N}_i} W_{i,j}[f(i) - f(j)]$

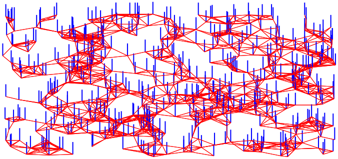
EXAMPLE: UNWEIGHTED GRID GRAPH



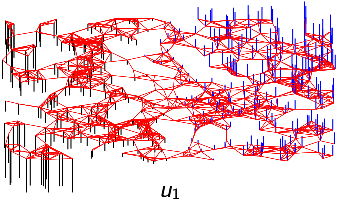
$$-\mathcal{L}f(i) = \Big[f(x_0+1,y_0) - f(x_0,y_0)\Big] - \Big[f(x_0,y_0) - f(x_0-1,y_0)\Big]$$
$$+ \Big[f(x_0,y_0+1) - f(x_0,y_0)\Big] - \Big[f(x_0,y_0) - f(x_0,y_0-1)\Big]$$
$$\sim \frac{\partial^2 f}{\partial x^2}(x_0,y_0) + \frac{\partial^2 f}{\partial y^2}(x_0,y_0) = (\Delta f)(x_0,y_0)$$

Graph Laplacian Eigenvectors

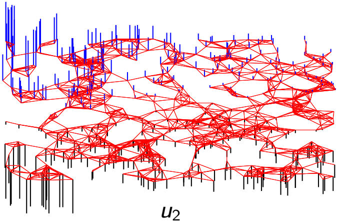
- Values of eigenvectors associated with lower frequencies (low λ_ℓ) change less rapidly across connected vertices



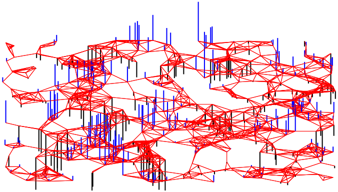
u_0



u_1



u_2



u_{50}

Graph Fourier Transform

- Fourier transform: expansion of f in terms of the eigenfunctions of the Laplacian / graph Laplacian

Functions on the Real Line

Fourier Transform

$$\hat{f}(\xi) = \langle f, e^{2\pi i \xi t} \rangle = \int_{\mathbb{R}} f(t) e^{-2\pi i \xi t} \, dt$$

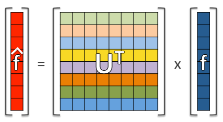
Inverse Fourier Transform

$$f(t) = \int_{\mathbb{R}} \hat{f}(\xi) e^{2\pi i \xi t} \, d\xi$$

Functions on the Vertices of a Graph

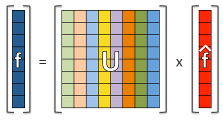
Graph Fourier Transform (Analysis)

$$\hat{f}(\lambda_\ell) = \langle f, u_\ell \rangle = \sum_{n=1}^{N} f(n) u_\ell^*(n)$$

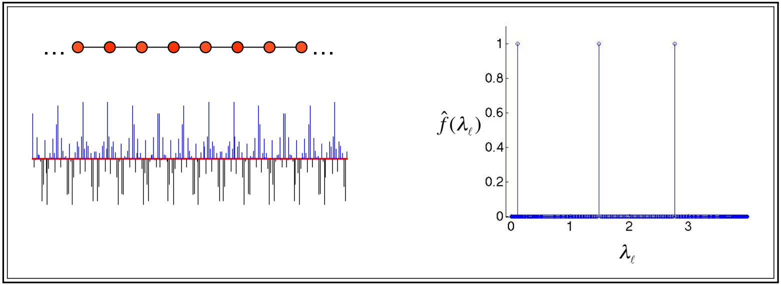


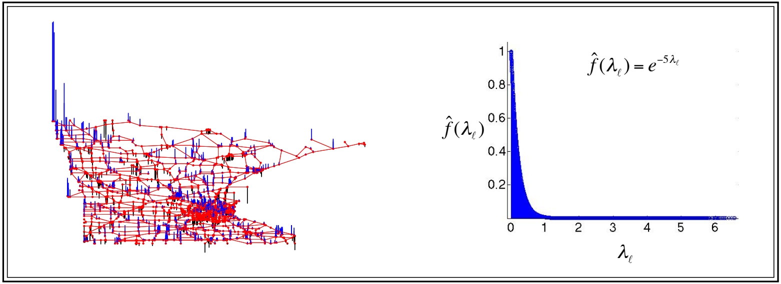
Inverse Graph Fourier Transform (Synthesis)

$$f(n) = \sum_{\ell=0}^{N-1} \hat{f}(\lambda_\ell) u_\ell(n)$$

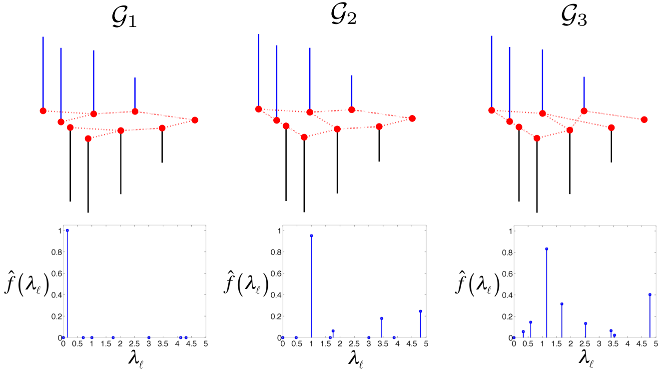


Signals on Graphs in Two Domains





Incorporation of the Underlying Graph Connectivity



- Recall, a signal is smooth *with respect to the intrinsic structure of its underlying graph*
- Similarly, the graph spectral content also depends on the underlying graph

Summary So Far

- {Complex exponentials } $\Longleftrightarrow$ { graph Laplacian eigenvectors }
  - Dictionaries that provide notions of frequency
- Analysis with these dictionaries $\Longleftrightarrow$ Fourier / graph Fourier transform
- Synthesis with these dictionaries $\Longleftrightarrow$ inverse Fourier / graph Fourier transform
- Transforms and their inverses provide two different ways to represent the same signal in two different domains
- Next step: generate overcomplete dictionaries whose atoms are localized in the vertex domain (time) and the graph spectral domain (frequency)

Outline

Generalized Translation on Graphs

- Define a generalized convolution by imposing that convolution in the vertex domain is multiplication in the graph spectral domain
- Define generalized translation via generalized convolution with a delta

Functions on the Real Line

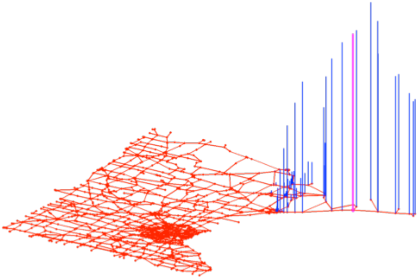
For $f \in L^2(\mathbb{R})$, in the weak sense

$$(T_s f)(t) := f(t-s) = (f * \delta_s)(t) = \int_{\mathbb{R}} \hat{f}(\xi) e^{-2\pi i \xi s} e^{2\pi i \xi t} d\xi$$

Functions on the Vertices of a Graph

For $f \in \mathbb{R}^N$, we define

$$(T_i f)(n) := \sqrt{N}(f * \delta_i)(n) = \sqrt{N} \sum_{\ell=0}^{N-1} \hat{f}(\lambda_\ell) u_\ell^*(i) u_\ell(n)$$



Properties of Generalized Translation Operators on Graphs

- Warning 1: Do not have the group structure of classical translation:

$$T_i T_j \neq T_{i+j}$$

- Warning 2: Unlike the classical case, generalized translation operators are not unitary, so $\|T_i g\|_2 \neq \|g\|_2$ in general

- However, the mean is preserved: $\sum_n (T_i g)(n) = \sum_n g(n)$

Theorem (Smoothness of $\hat{g}$ leads to localization of $T_i g$ around vertex i)

Let $\hat{g} : [0, \lambda_{\max}] \to \mathbb{R}$ be a kernel and define $d_{in} := d_{\mathcal{G}}(i, n)$. Then

$$|(T_i g)(n)| \leq \sqrt{N} B_{\hat{g}}(d_{in} - 1),$$

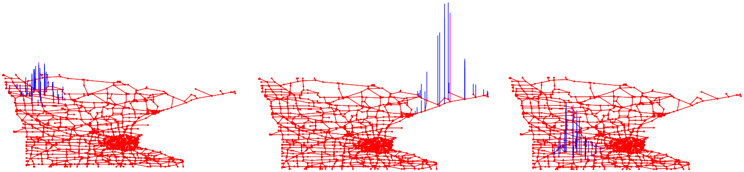
where $B_{\hat{g}}(K)$ is the minimax polynomial approximation error over all polynomials of degree K:

$$B_{\hat{g}}(K) := \inf_{\widehat{p_K}} \left\{ \sup_{\lambda \in [0, \lambda_{\max}]} |\hat{g}(\lambda) - \widehat{p_K}(\lambda)| \right\}.$$

Outline

A Windowed Graph Fourier Transform

1. Translate a window g to each vertex of the graph



$T_{100}g$ $T_{200}g$ $T_{2000}g$

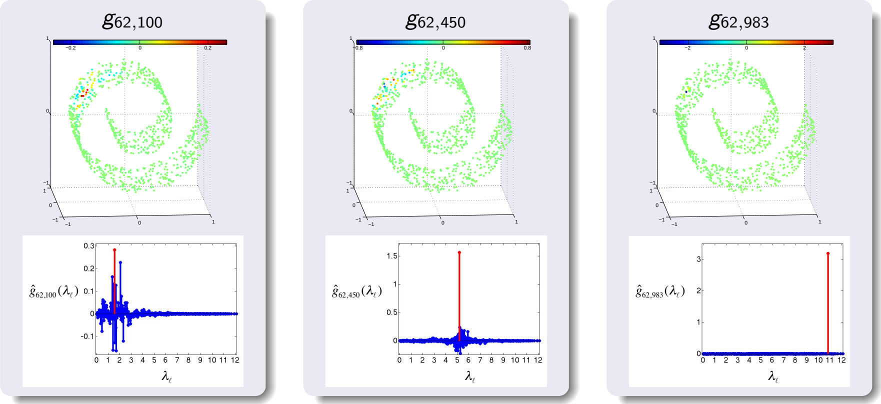
2. Multiply each component of the graph signal f of interest by the corresponding component of the translated window $T_i g$

3. Take the graph Fourier transform of $f.*T_i g$ (recall analysis)

A Windowed Graph Fourier Transform (cont.)

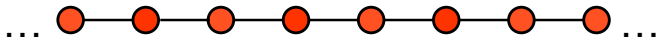
- Windowed graph Fourier atoms: $g_{i,k} := M_k T_i g$
- Windowed graph Fourier transform: $Sf(i,k) := \langle f, g_{i,k} \rangle$

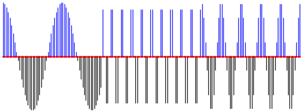
EXAMPLE: THREE DIFFERENT ATOMS ON THE SWISS ROLL GRAPH



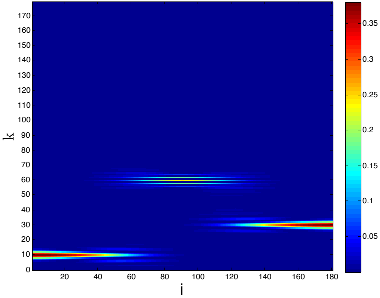
Example 1: The Path Graph

- Signal f on the path graph comprised of three different graph Laplacian eigenvectors restricted to three different segments of the graph:

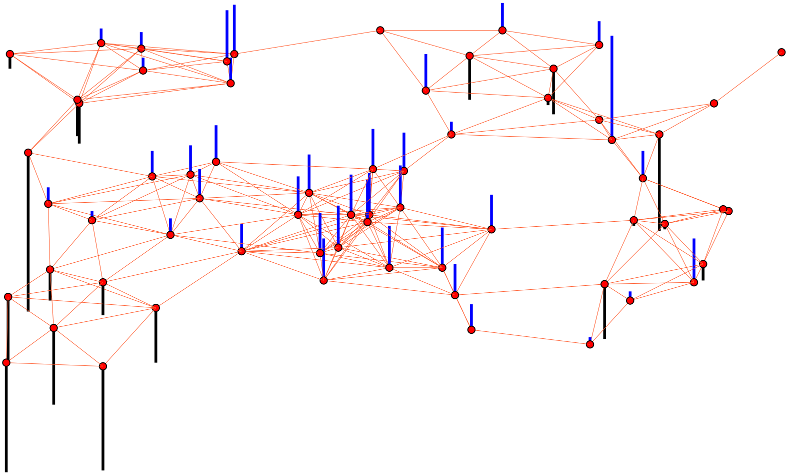




- "Spectrogram" of f showing $|Sf(i,k)|^2$



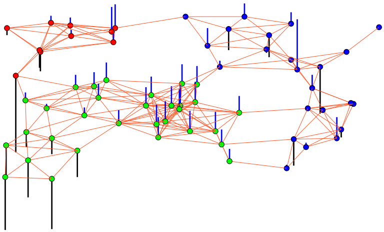
Example 2: A Signal on a Random Sensor Network

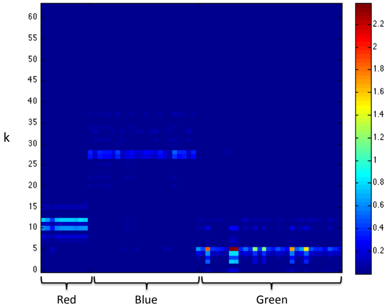


Any structure?

Example 2: A Signal on a Random Sensor Network

- Signal f comprised of three different graph Laplacian eigenvectors (u_{10}, u_{27}, u_5) restricted to the three different clusters of vertices

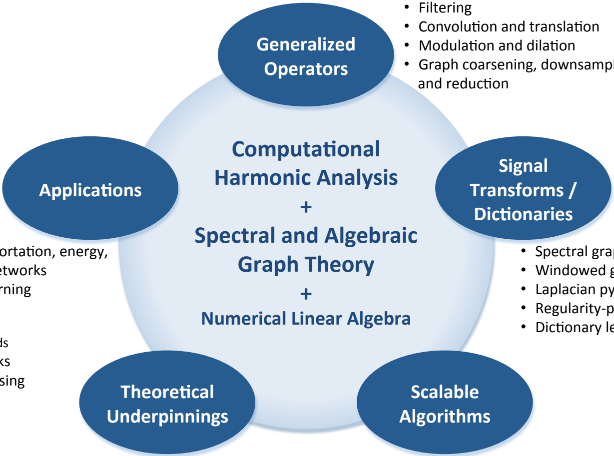




Summary

- Weighted graphs are a flexible tool to represent a wide variety of topologically-complicated data domains
- To identify and exploit structure in the data, we need to design dictionaries that incorporate the intrinsic geometric structure of the underlying data domain
- Try to leverage intuition from computational harmonic analysis of signals on Euclidean domains
  - Some ideas generalize relatively straightforwardly (e.g., notion of frequency)
  - However, signal and transforms on graphs can have surprising properties due to the irregularity of the data domains (e.g., uncertainty)
- Field is *emerging*
  - Requires more connections/iterations between dictionary design, theory, algorithms, and applications
  - Application of these techniques to real science and engineering problems is in its infancy

Ongoing and Future Work



Generalized Operators

- Filtering
- Convolution and translation
- Modulation and dilation
- Graph coarsening, downsampling, and reduction

Applications

- Social, transportation, energy, and sensor networks
- Statistical learning
- Astrophysics
  - CMB
  - Asteroids
- Brain networks
- Image processing
- …

Signal Transforms / Dictionaries

- Spectral graph wavelets
- Windowed graph Fourier transform
- Laplacian pyramid and graph filter banks
- Regularity-preserving transforms
- Dictionary learning

Theoretical Underpinnings

- Connections between sparsity and smoothness
- Uncertainty principles
- Graph Laplacian eigenvector localization

Scalable Algorithms

- Goal: Avoid full eigen-decomposition
- Polynomial approximation algorithms
  - Centralized (incl. GPU)
  - Distributed

Further Reading

Tutorial Overviews

- D. I Shuman, S. K. Narang, P. Frossard, A. Ortega, P. Vandergheynst, "Signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *Signal Process. Mag.*, to appear May 2013.
- R. Rubinstein, A. M. Bruckstein, and M. Elad, "Dictionaries for sparse representation modeling," *Proc. IEEE*, vol. 98, no. 6, pp. 1045–1057, Jun. 2010.

Spectral Graph Theory and Graph Laplacian Eigenvectors

- F. K. Chung, *Spectral Graph Theory*, vol. 92 of the CBMS Reg. Conf. Ser. Math., AMS Bokstore, 1997.
- T. Bıyıkoğlu, J. Leydold, and P. F. Stadler, *Laplacian Eigenvectors of Graphs*, Springer, 2007.
- D. Spielman, "Spectral graph theory" in *Combinatorial Scientific Computing*, Chapman and Hall, 2012.

Dictionaries for Signals on Graphs

- R. R. Coifman and M. Maggioni, "Diffusion wavelets," *Appl. Comput. Harmon. Anal.*, vol. 21, no. 1, pp. 53–94, Jul. 2006.
- D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Appl. Comput. Harmon. Anal.*, vol. 30, no. 2, pp. 129–150, Mar. 2011.
- S. K. Narang and A. Ortega, "Perfect reconstruction two-channel wavelet filter banks for graph structured data," *IEEE Trans. Signal Process.*, vol. 60, pp. 2786–2799, Jun. 2012.
- D. I Shuman, B. Ricaud, and P. Vandergheynst, "A windowed graph Fourier transform,"in *Proc. IEEE Stat. Signal Process. Wrkshp.*, Ann Arbor, MI, Aug. 2012.

Further Reading (cont.)

Scalable Algorithms

D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Appl. Comput. Harmon. Anal.*, vol. 30, no. 2, pp. 129–150, Mar. 2011.

D. I Shuman, P. Vandergheynst, and P. Frossard, "Chebyshev polynomial approximation for distributed signal processing," in *Proc. Int. Conf. Distr. Comput. Sensor Sys. (DCOSS)*, Barcelona, Spain, Jun. 2011.

Applications

N. Leonardi and D. Van De Ville, "Wavelet frames on graphs defined by FMRI functional connectivity," *Proc. IEEE Int. Symp. Biomed. Imag.* Mar. 2011, Chicago, IL, pp. 2136-2139.

D. I Shuman, M. J. Faraji, and P. Vandergheynst, "Semi-supervised learning with spectral graph wavelets," in *Proc. Int. Conf. Sampling Theory and Appl.*, Singapore, May 2011.

W. H. Kim *et al.*, "Wavelet based multi-scale shape features on arbitrary surfaces for cortical thickness discrimination," in *Proc. NIPS*, Lake Tahoe, NV, Dec. 2012.

N. Tremblay and P. Borgnat, "Multiscale community mining in networks using spectral graph wavelets," *ArXiv*, Dec. 2012.

Theoretical Underpinnings

A. Agaskar and Y. M. Lu, "Uncertainty principles for signals defined on graphs: Bounds and characterizations," in *Proc. IEEE Int. Conf. Acc., Speech, and Signal Process., Kyoto, Japan, Mar. 2012, pp. 3493–3496.*

N. Perraudin, "On localisation and uncertainty measures on graphs," Master's thesis, EPFL, Jun. 2012.