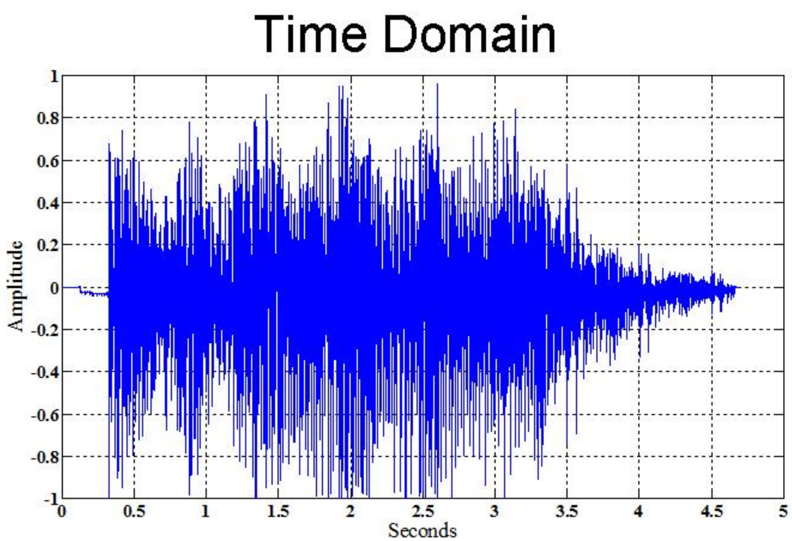
Processing Signals Supported on Graphs

Michael Rabbat

McGill

Traditional Signal Processing

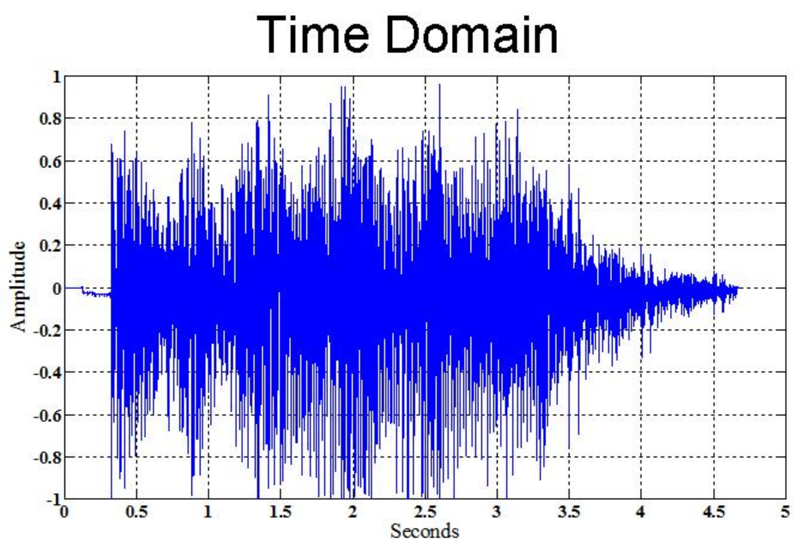


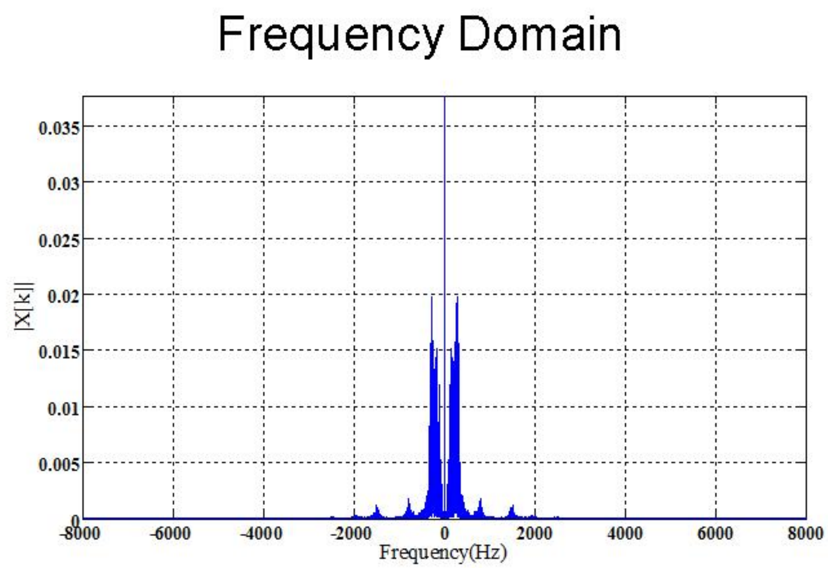
1-D (e.g., audio)

2-D (e.g., images)

Smoothness

Example: Audio signal

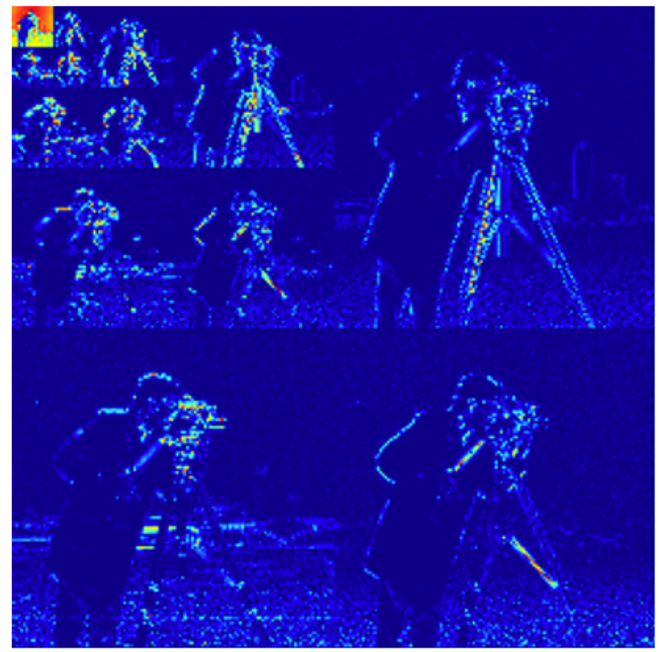




Smooth = (mostly) low frequency

Sparsity

Example: 2D Image and its Wavelet Transform



Sparsity = most wavelet coefficients are (nearly) zero
(Note: zero = blue)

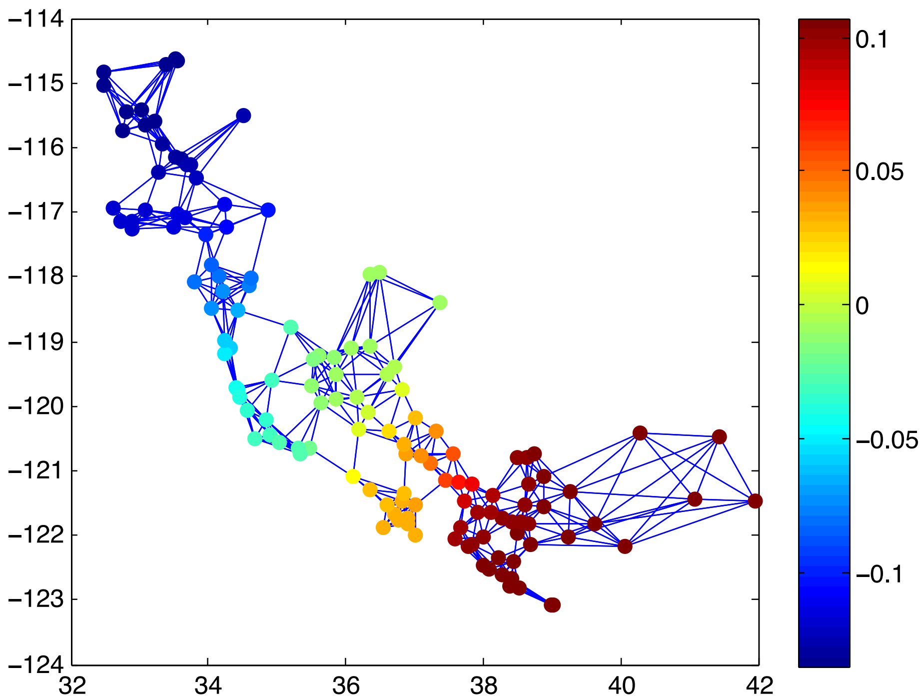
Implications of Smoothness & Sparsity

- Signal processing tasks
  - Signal measurement, acquisition → Estimation
  - Signal storage, communication → Compression

- Approximation Theory

  When/how can one signal be approximated well by another?
  - Other signal is "cleaner" or "simpler" than the other
  - Smoothness (focus on low frequency)
  - Sparsity (focus on few high-energy coefficients)

Signals Supported on Graphs



Many applications:

- Sensor networks
- Smart grid
- Social networks
- Transportation
- Internet monitoring
- Economic networks
- …

Data Source: California Irrigation Management Information System
http://www.cimis.water.ca.gov/

Questions

- When and how can we approximate signals on graphs?
- What is a "smooth" signal on a graph?
- What is a "Fourier" transform for signals on a graph?
- Which graphs have meaningful "Fourier" transforms?
- Which graphs have interesting smooth signals?
- When and how can smooth signals be helpful?

Outline

- Introduction and motivation
- Approximating signal supported on graphs
  - Classical approximation theory
  - Approximation theory for graphs
- Field estimation in sensor networks

- Based on joint work with Xiaofan Zhu
  - X. Zhu and M. Rabbat, "Approximating signals supported on graphs," *ICASSP* 2012
  - X. Zhu and M. Rabbat, "Graph spectral compressed sensing," *ICASSP* 2012

APPROXIMATING SIGNALS SUPPORTED ON GRAPHS

Classical Approximation Theory

Let $f \in L^2([0,1])$

Fourier transform $\widehat{f}(\omega) = \int_0^1 f(t) e^{-i\omega t} dt$

Total variation $\|f\|_V = \int_0^1 |f'(t)| dt$

Proposition: $|\widehat{f}(\omega)| \leq \frac{\|f\|_V}{|\omega|}$

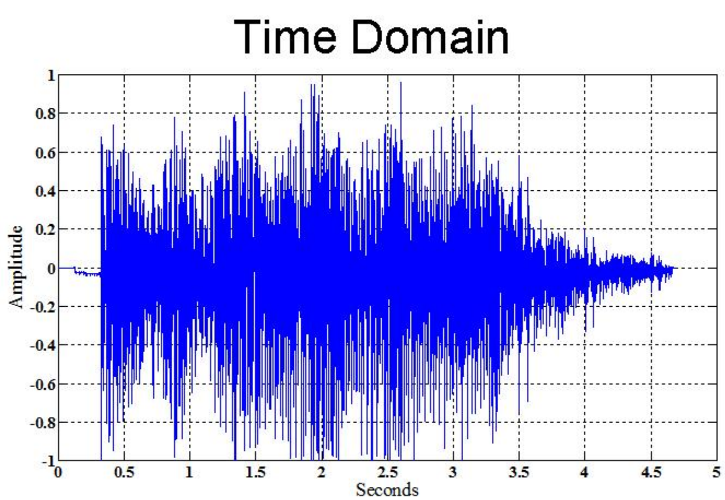
Small TV → energy mainly in low frequencies

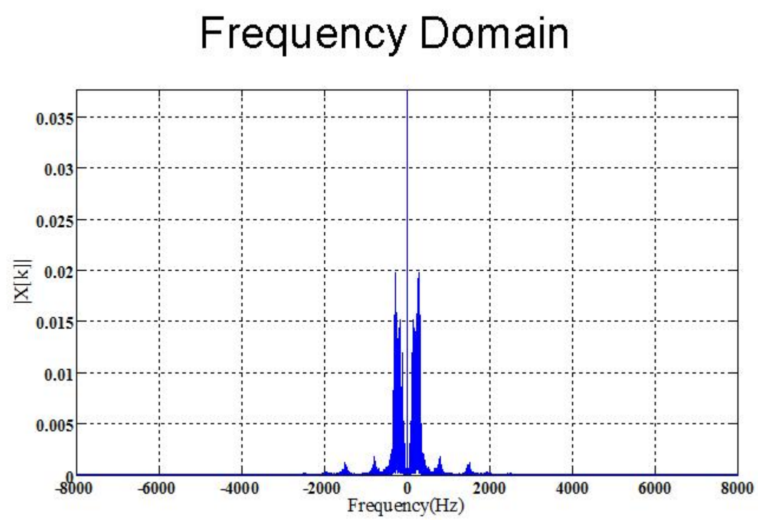
S. Mallat, *A Wavelet Tour of Signal Processing*, 3rd ed, Academic Press, 2009.

Fourier Approximation

Fourier coefficient $$\langle f(u), e^{i2\pi mu} \rangle = \int_0^1 f(u) e^{-i2\pi mu} du$$

Fourier expansion $$f(t) = \sum_{m=-\infty}^{\infty} \langle f(u), e^{i2\pi mu} \rangle e^{i2\pi mt}$$





M-term Linear Approximation

Only keep M lowest frequency coefficients
(Force others to zero)

M-term linear approximation:

$$f_M(t) = \sum_{m:|m|<M/2} \langle f(u), e^{i2\pi mu} \rangle e^{i2\pi mt}$$

M-term linear approximation error:

$$\begin{aligned}\epsilon_l(M, f) &= \|f - f_M\|^2 \\ &= \sum_{m:|m|>M/2} |\langle f(u), e^{i2\pi mu} \rangle|^2\end{aligned}$$

Approximation Error Scaling

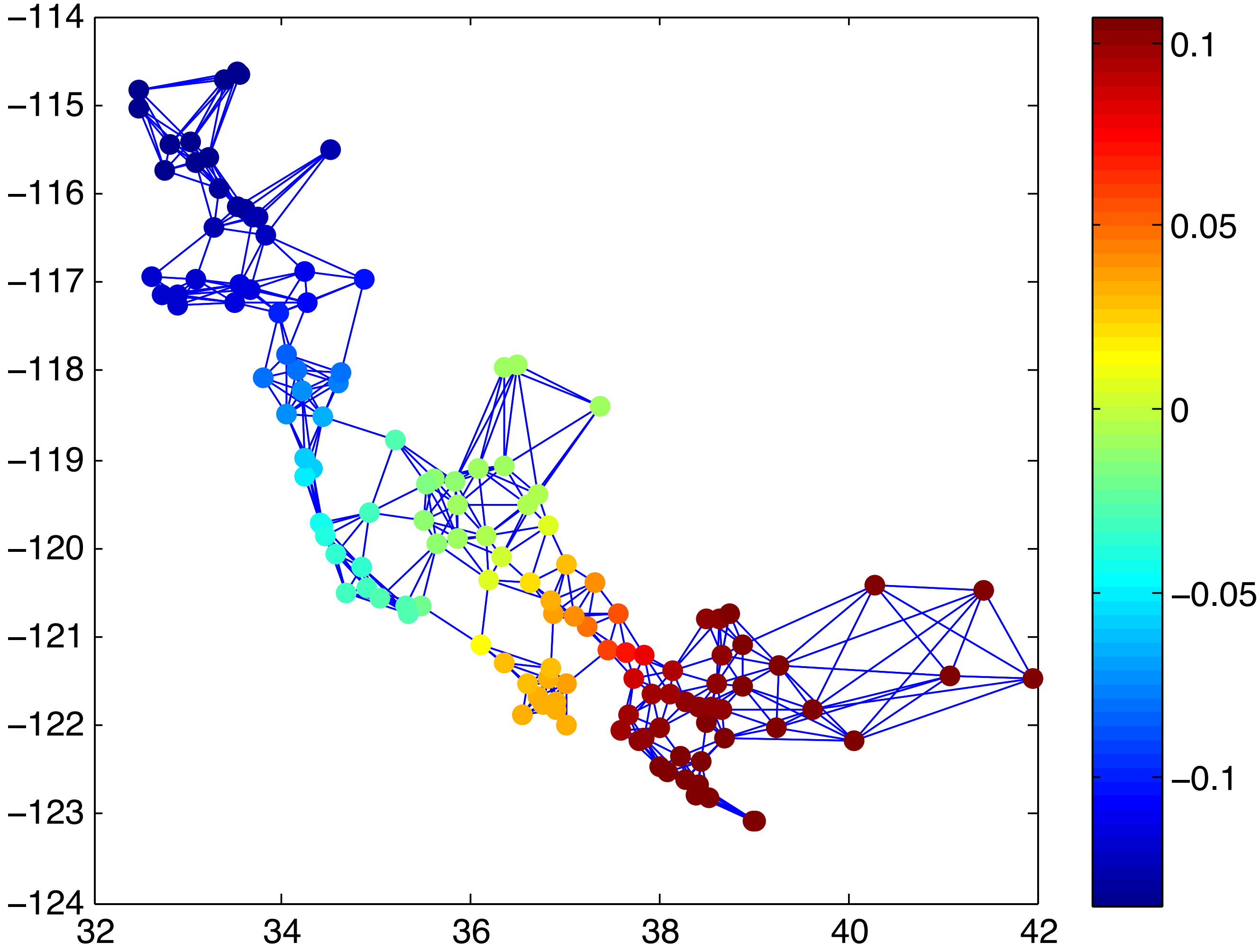
Theorem: If $\|f\|_V < \infty$ then $\epsilon_l(M, f) = O\left(\frac{\|f\|_V}{M^{-1}}\right)$

Theorem: For any $s > 1/2$, if

$$\sum_{m=0}^{\infty} |m|^{2s} \, |\langle f, e^{i2\pi m u}\rangle|^2 < \infty$$

then $\epsilon_l(M, f) = o(M^{-2s})$.

Signals on Graphs?



Quick Intro to Spectral Graph Theory

- Set representation of a graph $G = (V, E, w)$
- Adjacency Matrix A with entries

$$A_{u,v} = \begin{cases} w_{u,v} & \text{if } (u,v) \in E \\ 0 & \text{otherwise} \end{cases}$$

- Degree of node u: $d(u) = \sum_{v \in V} w_{u,v}$

- Degree matrix D is diagonal with entries $D_{u,u} = d(u)$

Smoothness and the Graph Laplacian

- Signal $x \in \mathbb{R}^{|V|}$ defined on vertices of G

  where x_v is the value at node v

- The graph Laplacian is $L = D - A$

- Define graph variation $\|x\|_G$ so that

$$\begin{aligned} \|x\|_G^2 &= x^T L x \\ &= \sum_{(u,v) \in E} w_{i,j} (x_u - x_v)^2 \end{aligned}$$

Graph Fourier Transform (GFT)

Consider eigenvalue decomposition of L

$$L = U \Lambda U^{-1}$$

with eigenvalues

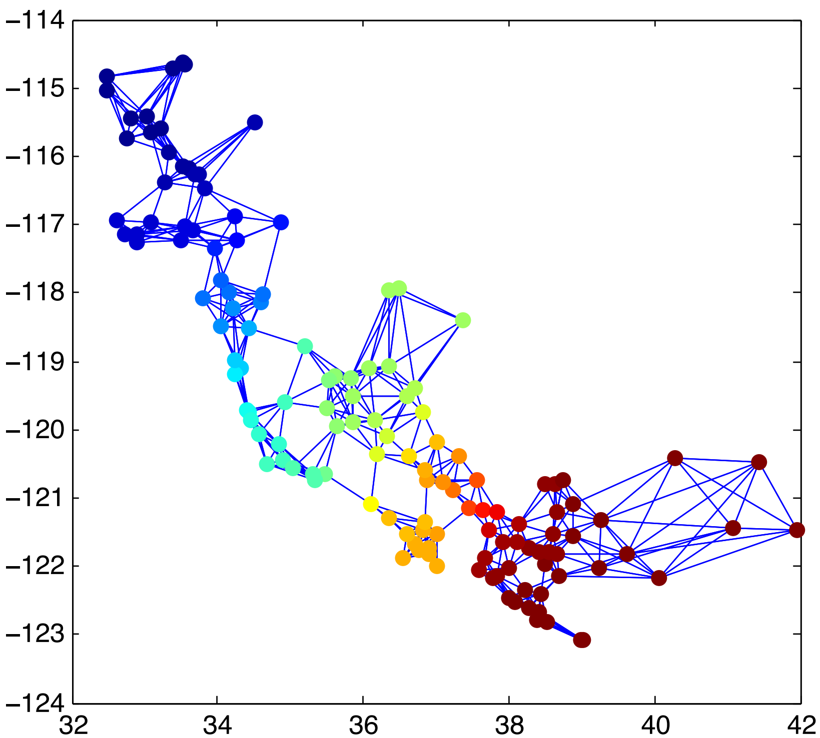
$$0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n \qquad\qquad n = |V|$$

and corresponding *i*th eigenvector u_i

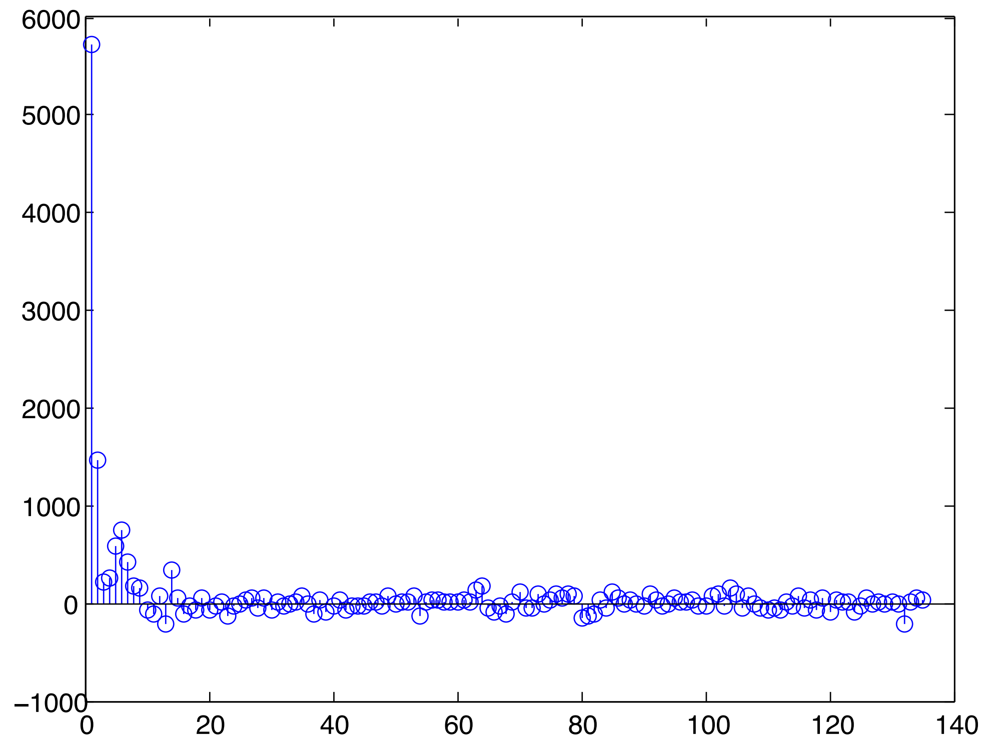
We'll call $\widehat{x}(\lambda_i) = \langle x, u_i \rangle$ the *i*th graph Fourier coefficient. Clearly,

$$x = \sum_{i=1}^{n} \widehat{x}(\lambda_i)\, u_i$$

GFT Example



x and G



$\widehat{x}(\lambda_k)$ vs k

Many Other Applications Using GFT

- Machine Learning
  - J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2000.
  - M. Belkin and P. Nyogi, "Using manifold structure for partially labeled classification," *NIPS*, 2002.
  - X. Zhu, J. Kandola, J. Lafferty, and Z. Gharamani, "Nonparametric transforms of graph kernels for semi-supervised learning," *NIPS*, 2005.
  - A. Smola and R. Kondor, "Kernels and regularization on graphs," *COLT*, 2003.
- Computer graphics
  - Z. Karni and C. Gotsman, "Spectral compression of mesh geometry," *ACM Conf. on Computer Graphics and Interactive Techniques*, 2000.

Why the Graph Laplacian Eigenbasis?

- Consider a ring graph on *n* vertices
  - Its Laplacian is circulant
  - Circulant matrices diagonalized by DFT matrix

$$U_{j,k} = e^{2\pi ijk/n}$$

  - Eigenvalues $\lambda_k = 2 - 2\cos(2\pi k/n)$
$$\approx (2\pi k/n)^2$$

Does this always make sense?

- Consider a complete graph on *n* vertices
  - Its Laplacian is circulant
  - Circulant matrices diagonalized by DFT matrix

$$U_{j,k} = e^{2\pi ijk/n}$$

  - Eigenvalues $\lambda_1 = 0$

$$\lambda_k = n \quad k \geq 2$$

- What does it mean to have a "smooth" signal on the complete graph?

Smooth Signals on Graphs

Intuitively x smooth on G if $\|x\|_G = x^T Lx$ small

Theorem: Let $\widehat{x}(\lambda_i) = \langle x, u_i \rangle$ where u_i is the *i*th eigenvector of L. Then

$$|\widehat{x}(\lambda_k)| \leq \frac{\|x\|_G}{\sqrt{\lambda_k}}$$

Approximating Signals on Graphs

Define M-term linear approximation of x on G as

$$x_M = \sum_{k=0}^{M} \widehat{x}(\lambda_k)\, u_k$$

M-term linear approximation error

$$\epsilon_l(M, x) = \sum_{k=M+1}^{n} |\widehat{x}(\lambda_k)|^2$$

Theorem: $\epsilon_l(M, x) \leq \|x\|_G^2 \lambda_M^{-1}$

Asymptotics

Let G be a graph with $|V| = \infty$

If $\displaystyle\sum_{k=0}^{\infty} k\lambda_k |\widehat{x}(\lambda_k)|^2 \leq \infty$

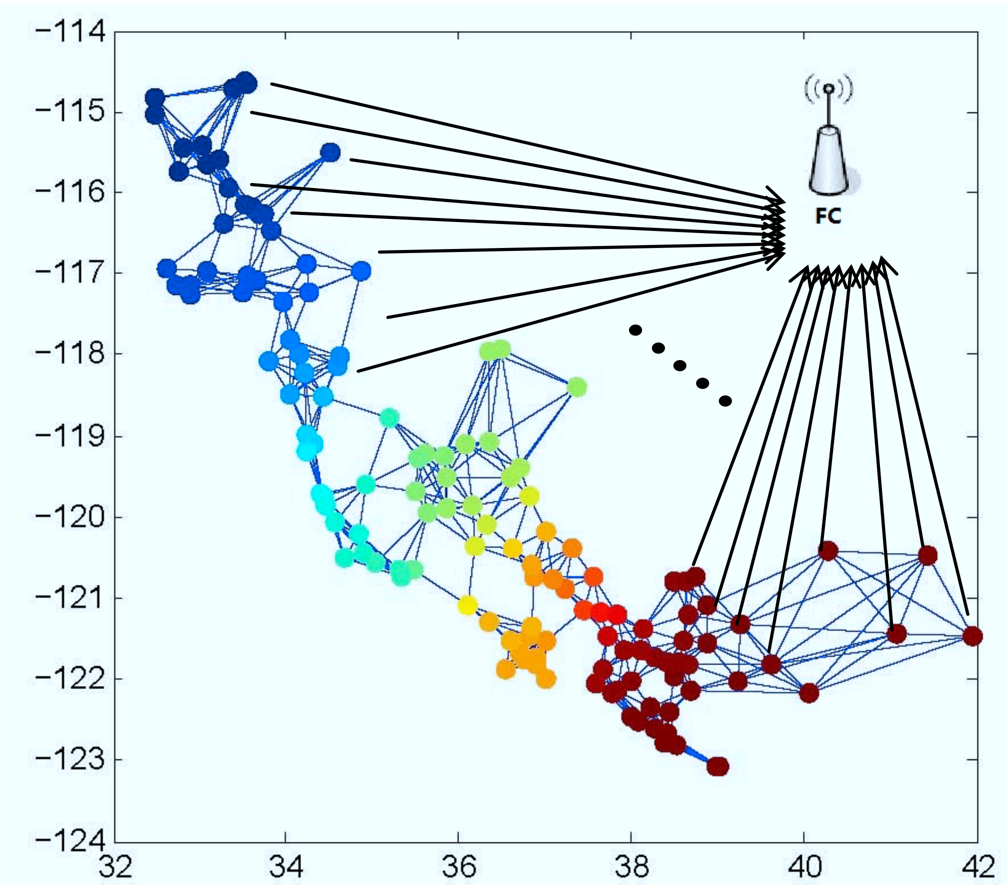
then $\epsilon_l(M, x) = o\left(\frac{1}{M\lambda_{M/2}}\right)$ as $M \to \infty$

Summary

- GFT has many similarities to the Fourier transform
  - Notion of smoothness
  - Linear approximation error
- Not all graphs support meaningful "smooth" signals
  - Laplacian eigenvalues should grow
- Can be used for "fitting" a graph to a signal or sequence of signals

GRAPH SPECTRAL COMPRESSED SENSING

Field Estimation in Sensor Networks

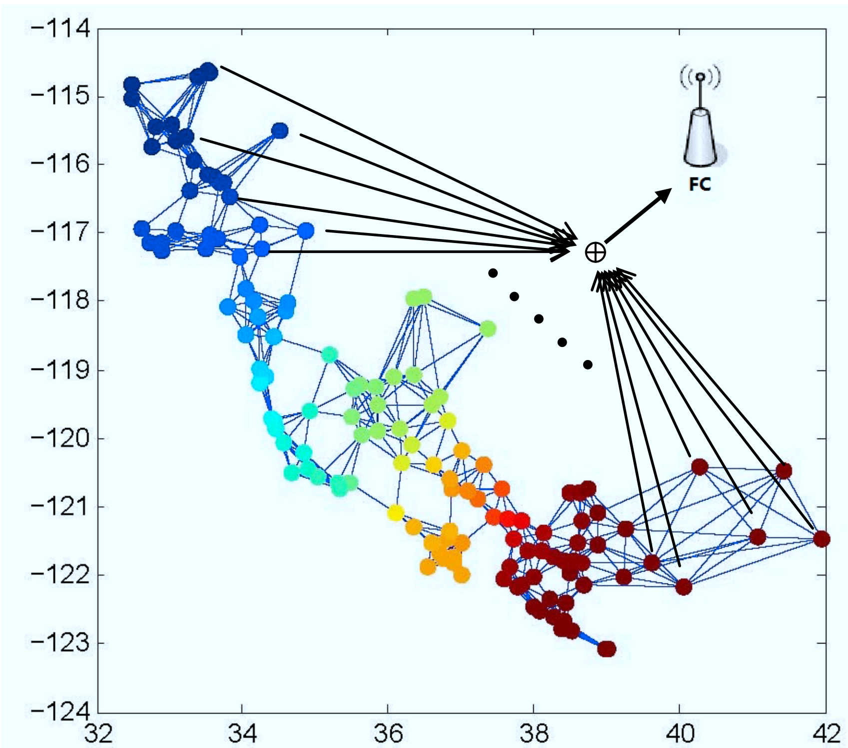


Estimate sensor measurements at fusion center (FC)

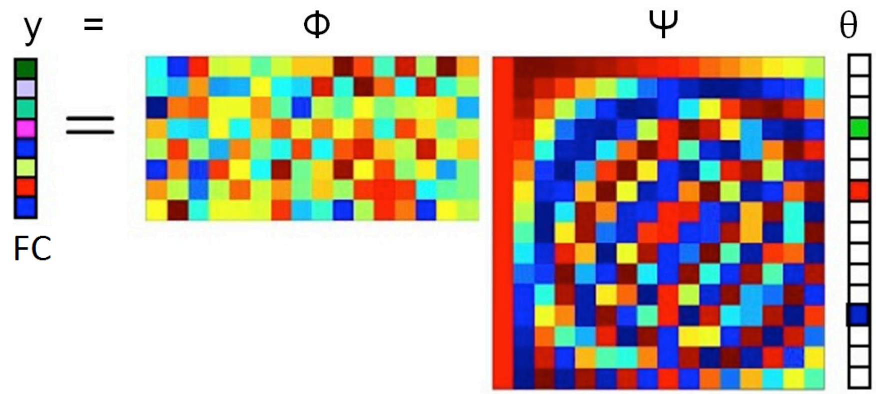
Performance metrics

- Distortion, MSE
- Bandwidth usage
- Energy usage

Compressed Sensing



- Assume signal is sparse
- Measure few random linear combinations



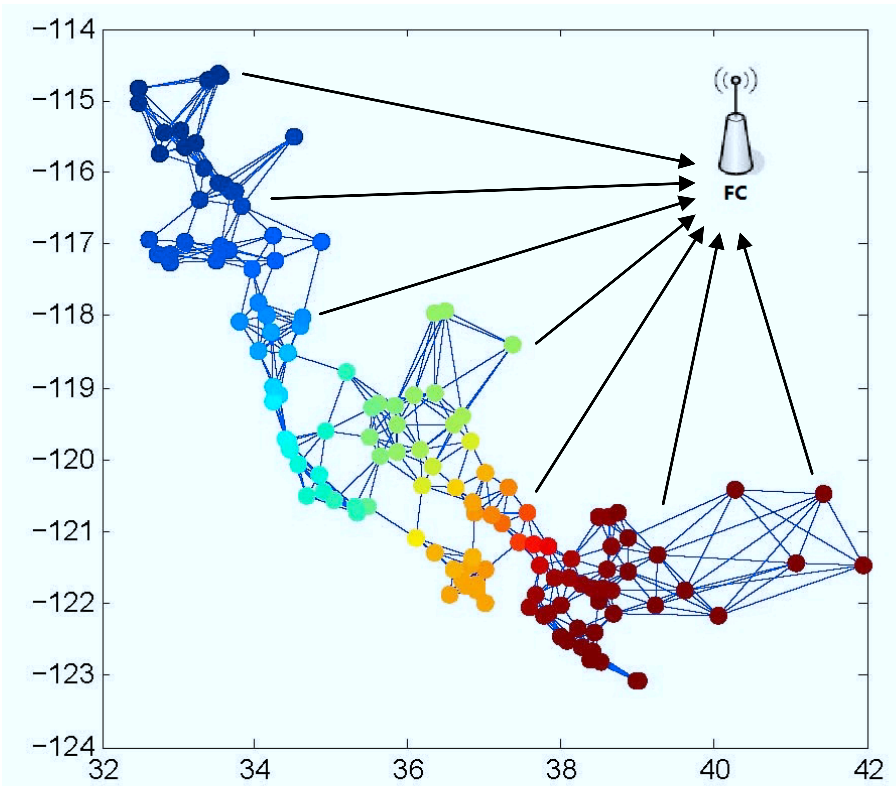
Candes & Tao, "Near-optimal signal recovery from random projections," *IEEE Trans Info Theory*, 2006.
D. Donoho, "Compressed sensing," *IEEE Trans Info Theory*, 2006.

W. Bajwa, J. Haupt, A. Sayeed, and R. Nowak, "Joint source-channel communication for distributed estimation in sensor networks," *IEEE Trans Info Theory*, 2007

Using CS for Field Estimation

- Pros:
  - Require fewer overall measurements
  - Each measurement is equally important
  - Distortion performance nearly optimal
- Cons:
  - Requires synchronization across network
  - Fewer total measurements, but every node transmits for every measurement

Graph Spectral Compressed Sensing



- Randomly sample a few sensors
- Interpolate remaining values wrt GFT basis

Reconstruction Guarantee

Suppose there are constants s and S such that

$$\epsilon_l(M, x) \leq SM^{-s}$$

If the number of measurements *m* obeys

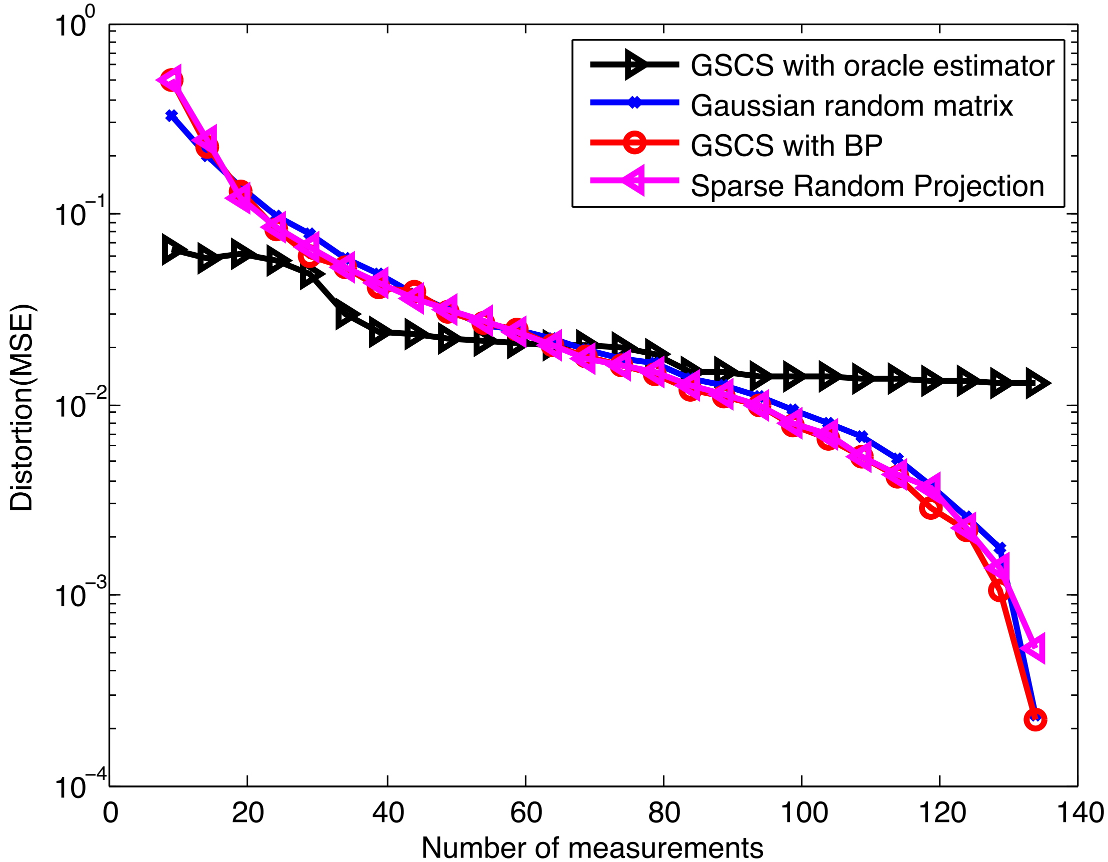
$$m \geq C_1 M \log(M/\delta)$$

then, with probability $1 - \delta$,

$$\|x - \tilde{x}\|_2 \leq \|x - x_M\|_2 + C_2 S M^{-s} \log\lceil n/M \rceil$$

where $\tilde{x} = \Phi_M^{\dagger} y$

Performance Example



- Using CIMIS data
- Comparing with
  - Gaussian random matrix: Bajwa, Haupt, Sayeed, and Nowak 2007
  - Sparse random projections: Wang, Garofalakis, Ramchandran 2007

Summary

- Graph structure can be useful for interpolation
  - When signal is smooth
  - (Graph should have interesting smooth signals)

- Potential implications for
  - Distributed measurement systems
  - Network design
  - Semi-supervised learning

Discussion and Directions

- From smoothness to sparsity

- Connection to random walks
  - Either G has interesting smooth signals
  - Or it has a rapidly mixing Markov chain

- Connection to gossip and network diffusion
  - Stop early, randomly sample a few nodes, and interpolate?

- Uncertainty principles for signals on graphs