# Towards Multi-Scale Signal Processing on Graphs

Pierre Vandergheynst Signal Processing Lab (IEL-LTS2)







#### **Signal Processing on Graphs**



http://lts2.epfl.ch

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE It seems hard to formulate a linear shift-invariant systems theory (LTI) for graphs. But we can try to get close.

The (combinatorial) Laplacian will be our main building block

$$\mathcal{L} = \mathbf{D} - \mathbf{W} \qquad \{(\lambda_{\ell}, \mathbf{u}_{\ell})\}_{\ell=0,1,\dots,N-1}$$

That particular ortho basis will play the role of the Fourier basis

$$\hat{f}(\lambda_{\ell}) := \langle \mathbf{f}, \mathbf{u}_{\ell} \rangle = \sum_{i=1}^{N} f(i) u_{\ell}^{*}(i)$$
$$\mu := \max_{\ell, i} |\langle \mathbf{u}_{\ell}, \delta_{i} \rangle| \in \left[\frac{1}{\sqrt{N}}, 1\right[$$





 $\operatorname{argmin}_{f}\left\{||f-y||_{2}^{2}+\gamma f^{T}\mathcal{L}f\right\}$ 



















## Kernels, Convolutions and Translations

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

Inherits a lot of properties of the usual convolution

associativity, distributivity, diagonalized by GFT  $g_0(n) := \sum_{\ell=0}^{N-1} u_\ell(n) \qquad f * g_0 = f$   $\mathcal{L}(f * g) = (\mathcal{L}f) * g = f * (\mathcal{L}g)$ 

Use convolution to induce translations

LTS

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



Given a spectral kernel g, construct the family of features:

$$\phi_n(m) = (T_n g)(m)$$
  $\phi_n(m) = \sqrt{N} \sum_{\ell=0}^{N-1} \hat{g}(\ell) u_\ell^*(m) u_\ell(n)$ 

Are these features localized ?





Given a spectral kernel g, construct the family of features:

$$\phi_n(m) = (T_n g)(m)$$
  $\phi_n(m) = \sqrt{N} \sum_{\ell=0}^{N-1} \hat{g}(\ell) u_\ell^*(m) u_\ell(n)$ 

Are these features localized ?

Suppose the GFT of the kernel is smooth enough (K+1 different.):

$$B = \sup_{x} |\hat{g}^{(K+1)}(x)|$$





Given a spectral kernel g, construct the family of features:

$$\phi_n(m) = (T_n g)(m)$$
  $\phi_n(m) = \sqrt{N} \sum_{\ell=0}^{N-1} \hat{g}(\ell) u_\ell^*(m) u_\ell(n)$ 

Are these features localized ?

Suppose the GFT of the kernel is smooth enough (K+1 different.):

$$B = \sup_{x} \left| \hat{g}^{(K+1)}(x) \right|$$

Construct an order K polynomial approximation:

$$\sup_{\ell} |\hat{g}(x) - P_K(x)| \le \frac{B}{2^K (K+1)!}$$





Given a spectral kernel g, construct the family of features:

$$\phi_n(m) = (T_n g)(m)$$
  $\phi_n(m) = \sqrt{N} \sum_{\ell=0}^{N-1} \hat{g}(\ell) u_\ell^*(m) u_\ell(n)$ 

Are these features localized ?

Suppose the GFT of the kernel is smooth enough (K+1 different.):

$$B = \sup_{x} |\hat{g}^{(K+1)}(x)|$$

Construct an order K polynomial approximation:

$$\sup_{\ell} |\hat{g}(x) - P_K(x)| \le \frac{B}{2^K (K+1)!}$$

 $\phi'_n(m) = \langle \delta_m, P_K(\mathcal{L}) \delta_n \rangle$  Exactly localized in a K-ball around n





## **Polynomial Localization - Extended**

f is (K+1)-times differentiable:

$$\inf_{q_{K}} \{ \|f - q_{K}\|_{\infty} \} \leq \frac{\left\lfloor \frac{b-a}{2} \right\rfloor^{n+1}}{(K+1)! 2^{K}} \|f^{(K+1)}\|_{\infty}$$
  
Let  $K_{in} := d(i, n) - 1$   
 $|(T_{i}g)(n)| \leq \sqrt{N} \inf_{\widehat{p_{K_{in}}}} \left\{ \sup_{\lambda \in [0, \lambda_{\max}]} |\hat{g}(\lambda) - \widehat{p_{K_{in}}}(\lambda)| \right\} = \sqrt{N} \inf_{\widehat{p_{K_{in}}}} \{ \|\hat{g} - \widehat{p_{K_{in}}}\|_{\infty} \}$ 

r r K+1

**Regular Kernels are Localized** 

If the kernel is d(i, n)-times differentiable:

$$|(T_ig)(n)| \le \left[\frac{2\sqrt{N}}{d_{in}!} \left(\frac{\lambda_{\max}}{4}\right)^{d_{in}} \sup_{\lambda \in [0,\lambda_{\max}]} |\hat{g}^{(d_{in})}(\lambda)|\right]$$





## **Polynomial Localization - Extended**

Example: for the heat kernel 
$$\hat{g}(\lambda) = e^{-\tau\lambda}$$
  
$$\frac{|(T_ig)(n)|}{||T_ig||_2} \leq \frac{2\sqrt{N}}{d_{in}!} \left(\frac{\tau\lambda_{\max}}{4}\right)^{d_{in}} \leq \sqrt{\frac{2N}{d_{in}\pi}} e^{-\frac{1}{12d_{in}+1}} \left(\frac{\tau\lambda_{\max}e}{4d_{in}}\right)^{d_{in}}$$

We can estimate an explicit measure of spread in terms of the degrees:







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







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



OOO

**4**||



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

• Generalized translation

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



**4**||

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



**4**||

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

• Generalized translation

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







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

• Generalized translation

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





**~** 



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





**~** 

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





**~** 



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

• Generalized translation

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







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

• Generalized translation

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





**~** 



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





**~** 

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





**~** 























































## **Simple De-Noising with Wavelets**







## **Simple De-Noising with Wavelets**



Original

Noisy

Denoised

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





## **Simple De-Noising with Wavelets**







## **Remark on Implementation**

40

Not necessary to compute spectral decomposition for filtering

Polynomial approximation :

λ

$$g(t\omega) \simeq \sum_{k=0}^{K-1} a_k(t) p_k(\omega)$$
ex: Chebysł

ex: Chebyshev, minimax

$$\tilde{W}_f(t_n, j) = \left(\frac{1}{2}c_{n,0}f^\# + \sum_{k=1}^{M_n} c_{n,k}\overline{T}_k(\mathcal{L})f^\#\right)_j$$

$$\overline{T}_k(\mathcal{L})f = \frac{2}{a_1}(\mathcal{L} - a_2I)\left(\overline{T}_{k-1}(\mathcal{L})f\right) - \overline{T}_{k-2}(\mathcal{L})f$$

Computational cost dominated by matrix-vector multiply with (sparse) Laplacian matrix. In particular  $O(\sum_{n=1}^{J} M_n |E|)$ 

http://wiki.epfl.ch/sgwt



0



#### Analysis operator





Shuman, Faraji, VDG, A framework for multiscale transforms on graphs, 2013





#### Analysis operator







#### Analysis operator







$$\mathcal{V}_1 = \mathcal{V}_+ := \{i \in \mathcal{V} : u_{\max}(i) \ge 0\}$$

Relaxed solution to 2-coloring for regular graphs

- Exact for bipartite graphs
- Connections with nodal domains theory for laplacian eigenvectors



(e)



































Analysis operator

$$\underbrace{\begin{pmatrix} y_0 \\ y_1 \end{pmatrix}}_{y} = \underbrace{\begin{pmatrix} \mathbf{H_m} \\ \mathbf{I} - \mathbf{GH_m} \end{pmatrix}}_{\mathbf{T_a}} x,$$







Analysis operator

$$\underbrace{\begin{pmatrix} y_0 \\ y_1 \end{pmatrix}}_{y} = \underbrace{\begin{pmatrix} \mathbf{H_m} \\ \mathbf{I} - \mathbf{GH_m} \end{pmatrix}}_{\mathbf{T_a}} x,$$

Simple (traditional) left inverse

$$\hat{x} = \underbrace{\left(\begin{array}{cc} \mathbf{G} & \mathbf{I} \end{array}\right)}_{\mathbf{T_s}} \underbrace{\left(\begin{array}{c} y_0 \\ y_1 \end{array}\right)}_{y}$$

 $\mathbf{T_sT_a} = \mathbf{I} \qquad \qquad \text{with no conditions on } \mathbf{H} \text{ or } \mathbf{G}$ 

Do, Vetterli, Framing Pyramids, IEEE TSP, 2003





## **Coarsening by Kron Reduction**

In order to iterate the construction, we need to construct a graph on the reduced vertex set.

$$\mathbf{A}_{\mathbf{r}} = \mathbf{A}[\alpha, \alpha] - \mathbf{A}[\alpha, \alpha) \mathbf{A}(\alpha, \alpha)^{-1} \mathbf{A}(\alpha, \alpha]$$
  

$$\mathbf{A}_{\mathbf{r}} = \begin{bmatrix} \mathbf{A}[\alpha, \alpha] & \mathbf{A}[\alpha, \alpha) \\ \mathbf{A}(\alpha, \alpha] & \mathbf{A}(\alpha, \alpha) \end{bmatrix}$$
  
Schur complement  

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}[\alpha, \alpha] & \mathbf{A}[\alpha, \alpha) \\ \mathbf{A}(\alpha, \alpha] & \mathbf{A}(\alpha, \alpha) \end{bmatrix}$$









Idea: Optimize interpolation for reduction:

$$y[u] = \sum_{v \in V_1} \alpha[v] \varphi^v[u]$$
 Shifted Green's function of  $L$  at vertex  $v$   
 $y[v'] = \sum_{v \in V_1} \alpha[v] \varphi^v[v'] = x[v'] \quad \forall v' \in V_1$ 





Simple linear model:

$$\begin{aligned} f_{\text{interp}}(i) &= \sum_{j \in \mathcal{V}_r} \alpha[j]\varphi_j(i) & f_{\text{interp}} = \Phi\alpha \\ \text{With:} \quad \varphi_j(i) &= \left(T_j\varphi\right)(i) & \Phi[i,j] = \varphi_i(j) \end{aligned}$$





Simple linear model:

$$f_{\text{interp}}(i) = \sum_{j \in \mathcal{V}_r} \alpha[j]\varphi_j(i) \qquad f_{\text{interp}} = \Phi\alpha$$
  
With:  $\varphi_j(i) = (T_j\varphi)(i) \qquad \Phi[i,j] = \varphi_i(j)$ 

#### Interpolation condition:

On the known vertices:  $f_r = \Phi_{\mathcal{V}_r} \alpha$ 

Solution depends on efficient, robust inversion of:  $\alpha = \Phi_{\mathcal{V}_r}^{-1} f_r$ 

#### Those weights can be computed using only filtering !





Regularized Laplacian:

$$\tilde{\mathcal{L}} = \mu^{-1}\mathcal{L} + \mathbf{I}_{|\mathcal{V}|}$$

Stable pseudo-inverse:

$$\tilde{\mathcal{L}}^{-1}[i,j] = \sum_{\ell=0}^{|\mathcal{V}|-1} \frac{1}{1+\mu^{-1}\lambda_{\ell}} u_{\ell}(i)u_{\ell}(j)$$





Regularized Laplacian:

$$\tilde{\mathcal{L}} = \mu^{-1}\mathcal{L} + \mathbf{I}_{|\mathcal{V}|}$$

Stable pseudo-inverse:

$$\tilde{\mathcal{L}}^{-1}[i,j] = \sum_{\ell=0}^{|\mathcal{V}|-1} \frac{1}{1+\mu^{-1}\lambda_{\ell}} u_{\ell}(i)u_{\ell}(j)$$

$$\varphi_j(i) = \sum_{\ell=0}^{|\mathcal{V}|-1} \frac{1}{1+\mu^{-1}\lambda_\ell} u_\ell(i) u_\ell(j)$$

Shifted Green's functions





Regularized Laplacian:

$$\tilde{\mathcal{L}} = \mu^{-1}\mathcal{L} + \mathbf{I}_{|\mathcal{V}|}$$

Stable pseudo-inverse:

$$\tilde{\mathcal{L}}^{-1}[i,j] = \sum_{\ell=0}^{|\mathcal{V}|-1} \frac{1}{1+\mu^{-1}\lambda_{\ell}} u_{\ell}(i)u_{\ell}(j)$$

$$\varphi_j(i) = \sum_{\ell=0}^{|\mathcal{V}|-1} \frac{1}{1+\mu^{-1}\lambda_\ell} u_\ell(i) u_\ell(j)$$

Shifted Green's functions

$$\tilde{\mathcal{L}}_{r}f_{\text{interp}}(i) = \tilde{\mathcal{L}}_{r}f_{r}(i), \forall i \in \mathcal{V}_{r} \quad \text{Note:} \quad \tilde{\mathcal{L}}\varphi_{j}(i) = \tilde{\mathcal{L}}\tilde{\mathcal{L}}^{-1}\delta_{j}(i) \\
= \sum_{j \in \mathcal{V}_{r}} \alpha[j](\tilde{\mathcal{L}}_{r}\varphi_{j})(i) \quad \sum_{j \in \mathcal{V}_{r}} \alpha[j](\tilde{\mathcal{L}}_{r}\varphi_{j})(i) = \delta_{j}(i)$$

Does this property carry over to the Kron reduced Laplacian?





**Lemma:** Inversion/Reduction commute for the (regularized) Laplacian

$$\left(\tilde{\mathcal{L}}^{-1}\right)_{\mathcal{V}_r} = \left(\tilde{\mathcal{L}}_r\right)^{-1}$$

This implies invariance of the Green's functions via reduction and therefore

$$\alpha = \tilde{\mathcal{L}}_r f_r \qquad f_{\text{interp}} = \mathbf{\Phi} \alpha$$





 ${\bf Lemma:}\ Inversion/Reduction\ commute\ for\ the\ (regularized)\ Laplacian$ 

$$\left(\tilde{\mathcal{L}}^{-1}
ight)_{\mathcal{V}_r} = \left(\tilde{\mathcal{L}}_r
ight)^{-1}$$

This implies invariance of the Green's functions via reduction and therefore

$$\alpha = \tilde{\mathcal{L}}_r f_r \qquad f_{\text{interp}} = \mathbf{\Phi} \alpha$$

Algorithm: Reduce graph

Apply reduced Laplacian to vertex data

Replace old data with newly calculated coefficients

Filter with Green's kernel









#### Exact Representation











## 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







LTS

EPFL





















