Angular Synchronization and its application in Phase Retrieval

Afonso S. Bandeira PACM, Princeton University

joint work with Amit Singer (Princeton), Daniel A. Spielman (Yale), Boris Alexeev (Princeton), Matthew Fickus (AFIT), and Dustin G. Mixon (AFIT)

> OSL 2013, Les Houches. January 11, 2013

http//www.math.princeton.edu/~ajsb

Spectral Clustering – Cheeger Inequality

$$G = (V, E, (W)_{ij} = w_{ij})$$



Cheeger Constant:

$$h_G = \min_{S \subset V} h_G(S)$$

$$h_G(S) = \frac{\operatorname{cut}(S, S^c)}{\min\{\operatorname{vol}(S), \operatorname{vol}(S^c)\}}$$

Spectral Clustering – Cheeger Inequality

$$G = (V, E, (W)_{ij} = w_{ij})$$



Cheeger Constant:

$$h_G = \min_{S \subset V} h_G(S)$$

$$h_G(S) = \frac{\operatorname{cut}(S, S^c)}{\min\{\operatorname{vol}(S), \operatorname{vol}(S^c)\}}$$

Graph Laplacian

$$D = diag(d_i)$$

$$L_{0} = D - W \text{ and } L_{0} = I - D^{-1/2} W D^{-1/2}$$
$$\frac{x^{T} L_{0} x}{x^{T} D x} = \frac{1}{2} \frac{\sum_{ij} w_{ij} |x_{i} - x_{j}|^{2}}{\sum_{i} d_{i} x_{i}^{2}}$$

Spectral Clustering – Cheeger Inequality

$$G = (V, E, (W)_{ij} = w_{ij})$$



Cheeger Constant:

$$h_G = \min_{S \subset V} h_G(S)$$

$$h_G(S) = \frac{\operatorname{cut}(S, S^c)}{\min\{\operatorname{vol}(S), \operatorname{vol}(S^c)\}}$$

Graph Laplacian

$$D = diag(d_i)$$

$$\begin{split} L_0 &= D - W \quad \text{and} \quad \begin{array}{l} \mathcal{L}_0 &= I - D^{-1/2} W D^{-1/2} \\ \frac{x^T L_0 x}{x^T D x} &= \frac{1}{2} \frac{\sum_{ij} w_{ij} |x_i - x_j|^2}{\sum_i d_i x_i^2} \end{split}$$

Theorem (Cheeger Inequality (Alon 86))

$$\frac{1}{2}\lambda_2(\mathcal{L}_0) \le h_G \le \sqrt{2\lambda_2(\mathcal{L}_0)}$$



- f a function that takes values in [0, 1].
 - Want to minimize it over a (discrete) set "comb".



- f a function that takes values in [0, 1].
 - Want to minimize it over a (discrete) set "comb".
 - Relax the problem to a continuous set "relax" that contains "comb" and on which minimizing f is easier.

Problem Relaxation



- f a function that takes values in [0, 1].
 - Want to minimize it over a (discrete) set "comb".
 - Relax the problem to a continuous set "relax" that contains "comb" and on which minimizing *f* is easier.
 - "rounding" procedure (that takes elements in "relax" and sends them to "comb") on which, say, the value of *f* never more than doubles.

Problem Relaxation



- f a function that takes values in [0, 1].
 - Want to minimize it over a (discrete) set "comb".
 - Relax the problem to a continuous set "relax" that contains "comb" and on which minimizing *f* is easier.
 - "rounding" procedure (that takes elements in "relax" and sends them to "comb") on which, say, the value of f never more than doubles.

opt relax \leq opt comb ≤ 2 (opt relax)

Problem

Determine a potential on the set V of vertices of a graph, with values on a group ${\mathcal G}$

$$\begin{array}{rrrr} g:V & \to & \mathcal{G} \\ i & \to & g_i \end{array}$$

given a few, possibly noisy, of the pairwise offset measurements (corresponding to the edges E of the graph)

$$\begin{array}{rccc} \rho: E & \to & \mathcal{G} \\ (i,j) & \to & \rho_{ij} \approx g_i g_j^{-1}. \end{array}$$



Examples... $\mathcal{G} = O(1) = \mathbb{Z}_2$



Examples... $\mathcal{G} = O(1) = \mathbb{Z}_2$



Examples... $\mathcal{G} = O(1) = \mathbb{Z}_2$



When all edges are red this is essentially Max-Cut

Orientation of a Manifold.



 $\rho_{ij} = \det(O_{ij})$

















Problem

Determine an angular potential on the set V of vertices of a graph,

$$egin{array}{rcl} eta_{\cdot}:V&\to&[0,2\pi) & e^{i heta_{\cdot}}=v:V&\to&\mathbb{T}\subset\mathbb{C}\ i&\to& heta_{i} & & i&\to&v_{i} \end{array}$$

given a few, possibly noisy, of the relative angle measurements (corresponding to the edges E of the graph)

Problem

Determine an angular potential on the set V of vertices of a graph,

given a few, possibly noisy, of the relative angle measurements (corresponding to the edges E of the graph)

Problem

6

Determine an angular potential on the set V of vertices of a graph,

given a few, possibly noisy, of the relative angle measurements (corresponding to the edges E of the graph)

Minimize:

$$\eta(v) = \frac{\sum_{ij} w_{ij} |v_i - \rho_{ij} v_j|^2}{\sum_i d_i |v_i|^2} = \frac{1}{\operatorname{vol}(G)} \sum_{ij} w_{ij} |v_i - \rho_{ij} v_j|^2.$$

Problem

6

Determine an angular potential on the set V of vertices of a graph,

given a few, possibly noisy, of the relative angle measurements (corresponding to the edges E of the graph)

The Frustration Constant:

$$\eta_G = \min_{v: V \to \mathbb{T}} \eta(v) = \frac{\sum_{ij} w_{ij} |v_i - \rho_{ij} v_j|^2}{\sum_i d_i |v_i|^2} = \frac{1}{\operatorname{vol}(G)} \sum_{ij} w_{ij} |v_i - \rho_{ij} v_j|^2.$$

The Graph Connection Laplacian

 $W_1 \in \mathbb{C}^{n \times n} \qquad (W_1)_{ij} = w_{ij} \rho_{ij} \in \mathbb{C}.$

The Graph Connection Laplacian is $L_1 \in \mathbb{C}^{n \times n}$ $L_1 = D - W_1$

The Graph Connection Laplacian

 $W_1 \in \mathbb{C}^{n \times n} \qquad (W_1)_{ij} = w_{ij} \rho_{ij} \in \mathbb{C}.$

The Graph Connection Laplacian is $L_1 \in \mathbb{C}^{n \times n}$ $L_1 = D - W_1$

The Normalized Graph Connection Laplacian is $\mathcal{L}_1 \in \mathbb{C}^{n imes n}$

$$\mathcal{L}_1 = D^{-1/2} L_1 D^{-1/2} = I_n - D^{-1/2} W_1 D^{-1/2}$$

• Under certain conditions \mathcal{L}_1 converges to the Connection Laplacian in Riemannian Geometry.

The Graph Connection Laplacian

 $W_1 \in \mathbb{C}^{n \times n} \qquad (W_1)_{ij} = w_{ij} \rho_{ij} \in \mathbb{C}.$

The Graph Connection Laplacian is $L_1 \in \mathbb{C}^{n \times n}$ $L_1 = D - W_1$

The Normalized Graph Connection Laplacian is $\mathcal{L}_1 \in \mathbb{C}^{n imes n}$

$$\mathcal{L}_1 = D^{-1/2} L_1 D^{-1/2} = I_n - D^{-1/2} W_1 D^{-1/2}$$

• Under certain conditions \mathcal{L}_1 converges to the Connection Laplacian in Riemannian Geometry.

$$\frac{x^T L_1 x}{x^T D x} = \frac{1}{2} \frac{\sum_{ij} w_{ij} |x_i - \rho_{ij} x_j|^2}{\sum_i d_i |x_i|^2} = \eta(x)$$

$$\frac{x^T L_1 x}{x^T D x} = \frac{1}{2} \frac{\sum_{ij} w_{ij} |x_i - \rho_{ij} x_j|^2}{\sum_i d_i |x_i|^2} = \eta(x)$$

$$\frac{x^T L_1 x}{x^T D x} = \frac{1}{2} \frac{\sum_{ij} w_{ij} |x_i - \rho_{ij} x_j|^2}{\sum_i d_i |x_i|^2} = \eta(x)$$
$$\lambda_1(\mathcal{L}_1) = \min_{x \in \mathbb{C}^n} \eta(x) = \min_{x: V \to \mathbb{C}} \eta(x)$$
$$\eta_G = \min_{v: V \to \mathbb{T}} \eta(v)$$

Question

Can we relate η_G to $\lambda_1(\mathcal{L}_1)$?

$$\frac{x^T L_1 x}{x^T D x} = \frac{1}{2} \frac{\sum_{ij} w_{ij} |x_i - \rho_{ij} x_j|^2}{\sum_i d_i |x_i|^2} = \eta(x)$$
$$\lambda_1(\mathcal{L}_1) = \min_{x \in \mathbb{C}^n} \eta(x) = \min_{x: V \to \mathbb{C}} \eta(x)$$
$$\eta_G = \min_{v: V \to \mathbb{T}} \eta(v)$$

Question

Can we relate η_G to $\lambda_1(\mathcal{L}_1)$?

NO!



$$\frac{x^T L_1 x}{x^T D x} = \frac{1}{2} \frac{\sum_{ij} w_{ij} |x_i - \rho_{ij} x_j|^2}{\sum_i d_i |x_i|^2} = \eta(x)$$
$$\lambda_1(\mathcal{L}_1) = \min_{x \in \mathbb{C}^n} \eta(x) = \min_{x: V \to \mathbb{C}} \eta(x)$$
$$\eta_G = \min_{v: V \to \mathbb{T}} \eta(v)$$

Question

Can we relate η_G to $\lambda_1(\mathcal{L}_1)$?

NO!



Fix - Consider instead:

$$\eta_G^* = \min_{v: V \to \mathbb{T} \cup \{0\}} \eta(v).$$

$$\frac{x^T L_1 x}{x^T D x} = \frac{1}{2} \frac{\sum_{ij} w_{ij} |x_i - \rho_{ij} x_j|^2}{\sum_i d_i |x_i|^2} = \eta(x)$$
$$\lambda_1(\mathcal{L}_1) = \min_{x \in \mathbb{C}^n} \eta(x) = \min_{x: V \to \mathbb{C}} \eta(x)$$
$$\eta_G = \min_{v: V \to \mathbb{T}} \eta(v)$$

Question

Can we relate η_G to $\lambda_1(\mathcal{L}_1)$?

NO!



Fix – Consider instead:

$$\eta_G^* = \min_{v: V \to \mathbb{T} \cup \{0\}} \eta(v).$$

Theorem

$$\lambda_1(\mathcal{L}_1) \le \eta_G^* \le \sqrt{10\lambda_1(\mathcal{L}_1)}$$

Problematic case:





If G has a large spectral gap $\lambda_2(\mathcal{L}_0)$ (or, equivalently a large Cheeger Constant), this should not be a problem.



If G has a large spectral gap $\lambda_2(\mathcal{L}_0)$ (or, equivalently a large Cheeger Constant), this should not be a problem.

$$\begin{split} \eta(v) &= \frac{1}{2} \frac{\sum_{ij} w_{ij} |v_i - \rho_{ij} v_j|^2}{\sum_i d_i |v_i|^2} \\ &\geq \frac{1}{2} \frac{\sum_{ij} w_{ij} (|v_i| - |v_j|)^2}{\sum_i d_i |v_i|^2} \end{split}$$



If G has a large spectral gap $\lambda_2(\mathcal{L}_0)$ (or, equivalently a large Cheeger Constant), this should not be a problem.

$$\begin{split} \eta(v) &= \frac{1}{2} \frac{\sum_{ij} w_{ij} |v_i - \rho_{ij} v_j|^2}{\sum_i d_i |v_i|^2} \\ &\geq \frac{1}{2} \frac{\sum_{ij} w_{ij} (|v_i| - |v_j|)^2}{\sum_i d_i |v_i|^2} \end{split}$$

Theorem

$$\lambda_1(\mathcal{L}_1) \leq \eta_G \leq rac{1}{\lambda_2(\mathcal{L}_0)} \, \mathcal{O}\Big(\lambda_1(\mathcal{L}_1)\Big).$$

What about beyond $\mathbb{Z}/2\mathbb{Z} = O(1)$ and SO(2) Synchronization?



What about beyond $\mathbb{Z}/2\mathbb{Z} = O(1)$ and SO(2) Synchronization?





What about beyond $\mathbb{Z}/2\mathbb{Z} = O(1)$ and SO(2) Synchronization?





Higher-Order Rotation Groups

We want to globally estimate $O: V \to O(d)$ such that $O_i \approx \rho_{ij}O_j$. Minimize:

$$\nu(O) = \frac{1}{\text{vol}(G)} \sum_{ij} w_{ij} \| O_i - \rho_{ij} O_j \|_F^2.$$

Higher-Order Rotation Groups

We want to globally estimate $O: V \to O(d)$ such that $O_i \approx \rho_{ij}O_j$. Minimize:

$$\nu(O) = \frac{1}{\text{vol}(G)} \sum_{ij} w_{ij} \| O_i - \rho_{ij} O_j \|_F^2.$$

Many eigenvalues/eigenvectors are needed

Theorem

Let $\lambda_i(\mathcal{L}_1)$ and $\lambda_i(\mathcal{L}_0)$ denote the *i*-th smallest eigenvalue of, respectively, the normalized Connection Laplacian \mathcal{L}_1 and the normalized graph Laplacian \mathcal{L}_0 . Let ν_G denote the O(d) frustration constant of G. Then,

$$rac{1}{d}\sum_{i=1}^d\lambda_i(\mathcal{L}_1)\leq
u_G\leq \mathrm{poly}(d)rac{1}{\lambda_2(\mathcal{L}_0)}\sum_{i=1}^d\lambda_i(\mathcal{L}_1).$$

The proof is constructive – the Algorithm achieves this!





Let opt be the minimum fraction of edges the coloring gets wrong.



Let opt be the minimum fraction of edges the coloring gets wrong.

Conjecture (U.G.C.)

For every $\epsilon \sim 0$ and $\delta \sim 1$ there exists k and an assignment of the edges (with k colors) such that deciding whether

$$opt < \epsilon$$
 or $opt > \delta$

is NP-hard.



Let opt be the minimum fraction of edges the coloring gets wrong.

Conjecture (U.G.C.)

For every $\epsilon \sim 0$ and $\delta \sim 1$ there exists k and an assignement of the edges (with k colors) such that deciding whether

opt
$$< \epsilon$$
 or opt $> \delta$
is NP-hard.

• Corresponds to localization in S_k .

One can represent S_k as permutation matrices in O(k).



Let opt be the minimum fraction of edges the coloring gets wrong.

Conjecture (U.G.C.)

For every $\epsilon \sim 0$ and $\delta \sim 1$ there exists k and an assignement of the edges (with k colors) such that deciding whether

opt
$$< \epsilon$$
 or opt $> \delta$
is NP-hard.

• Corresponds to localization in S_k .

One can represent S_k as permutation matrices in O(k).

• There seems to be NO good "rounding procedure". e.g.: all-ones vector is a perfect localization for relaxed problem

PART II:

Reconstruction without phase

Reconstruction without phase

• A signal $x \in \mathbb{C}^M$ is measured using a linear system but only the absolute value of the measurements is obtained

 $|\langle x, \varphi_n \rangle|, \quad n = 1, \dots, N$

Reconstruction without phase

• A signal $x \in \mathbb{C}^M$ is measured using a linear system but only the absolute value of the measurements is obtained

 $|\langle x, \varphi_n \rangle|, \quad n = 1, \dots, N$



Motivation: X-ray Crystallography and inversion of spectrograms.

Reconstruction without phase

• A signal $x \in \mathbb{C}^M$ is measured using a linear system but only the absolute value of the measurements is obtained

 $|\langle x, \varphi_n \rangle|, \quad n = 1, \dots, N$



Motivation: X-ray Crystallography and inversion of spectrograms.



State of the art

• Balan et al., 2006: For a generic system, phaseless measurements are injective whenever $N \geq 4M-2$

The right injectivity bound is believed to be 4M-4

• Phaselift (Candès et al., 2011) and Phasecut (Waldspurger et al., 2012): For a random system, stable recovery by Semi-Definite Programming for $N = \tilde{O}(M)$.

State of the art

• Balan et al., 2006: For a generic system, phaseless measurements are injective whenever $N \geq 4M-2$

The right injectivity bound is believed to be 4M-4

• Phaselift (Candès et al., 2011) and Phasecut (Waldspurger et al., 2012): For a random system, stable recovery by Semi-Definite Programming for $N = \tilde{O}(M)$.

Question

Can we design a measurement matrix such that it is possible to efficiently and stably recovery from only $N = \tilde{O}(M)$ measurements avoiding the SDP computational cost?

Polarization

• Synchronization allows to recover the phases of the measurements from the relative phases

$$\omega_{ij} := \left(\frac{\langle x, \varphi_i \rangle}{|\langle x, \varphi_i \rangle|}\right)^{-1} \frac{\langle x, \varphi_j \rangle}{|\langle x, \varphi_j \rangle|} = \frac{\overline{\langle x, \varphi_i \rangle} \langle x, \varphi_j \rangle}{|\langle x, \varphi_i \rangle||\langle x, \varphi_j \rangle|}$$

Polarization

• Synchronization allows to recover the phases of the measurements from the relative phases

$$\omega_{ij} := \left(\frac{\langle x, \varphi_i \rangle}{|\langle x, \varphi_i \rangle|}\right)^{-1} \frac{\langle x, \varphi_j \rangle}{|\langle x, \varphi_j \rangle|} = \frac{\overline{\langle x, \varphi_i \rangle} \langle x, \varphi_j \rangle}{|\langle x, \varphi_i \rangle||\langle x, \varphi_j \rangle|}$$

• We can determine ω_{ij} from other phaseless measurements:

$$\overline{\langle x,\varphi_i\rangle}\langle x,\varphi_j\rangle = \frac{1}{4}\sum_{k=1}^4 \mathbf{i}^k |\langle x,\varphi_i + \mathbf{i}^k\varphi_j\rangle|^2$$

• Synchronization allows to recover the phases of the measurements from the relative phases

$$\omega_{ij} := \left(\frac{\langle x, \varphi_i \rangle}{|\langle x, \varphi_i \rangle|}\right)^{-1} \frac{\langle x, \varphi_j \rangle}{|\langle x, \varphi_j \rangle|} = \frac{\overline{\langle x, \varphi_i \rangle} \langle x, \varphi_j \rangle}{|\langle x, \varphi_i \rangle||\langle x, \varphi_j \rangle|}$$

• We can determine ω_{ij} from other phaseless measurements:

$$\overline{\langle x,\varphi_i\rangle}\langle x,\varphi_j\rangle = \frac{1}{4}\sum_{k=1}^4 \mathbf{i}^k |\langle x,\varphi_i + \mathbf{i}^k\varphi_j\rangle|^2$$



- each φ_i corresponds to vertex i
- each set $\{\varphi_i + i^k \varphi_j\}_{k=1}^4$ to an edge between i and j.

• Synchronization allows to recover the phases of the measurements from the relative phases

$$\omega_{ij} := \left(\frac{\langle x, \varphi_i \rangle}{|\langle x, \varphi_i \rangle|}\right)^{-1} \frac{\langle x, \varphi_j \rangle}{|\langle x, \varphi_j \rangle|} = \frac{\overline{\langle x, \varphi_i \rangle} \langle x, \varphi_j \rangle}{|\langle x, \varphi_i \rangle||\langle x, \varphi_j \rangle|}$$

 $\bullet\,$ We can determine ω_{ij} from other phaseless measurements:

$$\overline{\langle x,\varphi_i\rangle}\langle x,\varphi_j\rangle = \frac{1}{4}\sum_{k=1}^4 \mathbf{i}^k |\langle x,\varphi_i + \mathbf{i}^k\varphi_j\rangle|^2$$



- each φ_i corresponds to vertex i
- each set $\{\varphi_i + i^k \varphi_j\}_{k=1}^4$ to an edge between i and j.
- We need a sparse graph!

- The measurements are noisy $(|\langle x, \varphi_i \rangle| + \epsilon_i)$.
- If x is nearly orthogonal to φ_i the noise in the relative phase blows-up

$$\omega_{ij} = \frac{\overline{\langle x, \varphi_i \rangle} \langle x, \varphi_j \rangle + \epsilon_{ij}}{\left| \overline{\langle x, \varphi_i \rangle} \langle x, \varphi_j \rangle + \epsilon_{ij} \right|}$$

- The measurements are noisy $(|\langle x, \varphi_i \rangle| + \epsilon_i)$.
- $\bullet~$ If x is nearly orthogonal to φ_i the noise in the relative phase blows-up

$$\omega_{ij} = \frac{\overline{\langle x, \varphi_i \rangle} \langle x, \varphi_j \rangle + \epsilon_{ij}}{\left| \overline{\langle x, \varphi_i \rangle} \langle x, \varphi_j \rangle + \epsilon_{ij} \right|}$$

• Vertices *i* for which $\langle \varphi_i, x \rangle \sim 0$ should be removed.

- The measurements are noisy $(|\langle x, \varphi_i \rangle| + \epsilon_i)$.
- If x is nearly orthogonal to φ_i the noise in the relative phase blows-up

$$\omega_{ij} = \frac{\overline{\langle x, \varphi_i \rangle} \langle x, \varphi_j \rangle + \epsilon_{ij}}{\left| \overline{\langle x, \varphi_i \rangle} \langle x, \varphi_j \rangle + \epsilon_{ij} \right|}$$

- Vertices i for which $\langle \varphi_i, x \rangle \sim 0$ should be removed.
- To succeed...
 - FRAME DESIGN do not delete too many vertices
 - **Q GRAPH DESIGN** still be able to perform synchronization



- The measurements are noisy $(|\langle x, \varphi_i \rangle| + \epsilon_i)$.
- If x is nearly orthogonal to φ_i the noise in the relative phase blows-up

$$\omega_{ij} = \frac{\overline{\langle x, \varphi_i \rangle} \langle x, \varphi_j \rangle + \epsilon_{ij}}{\left| \overline{\langle x, \varphi_i \rangle} \langle x, \varphi_j \rangle + \epsilon_{ij} \right|}$$

- Vertices i for which $\langle \varphi_i, x \rangle \sim 0$ should be removed.
- To succeed...
 - **FRAME DESIGN** do not delete too many vertices
 - **Q GRAPH DESIGN** still be able to perform synchronization



Theorem (Alexeev-Bandeira-Fickus-M, 2012)

Take $N \sim CM \log M$ with C sufficiently large. Then the following holds for all $x \in \mathbb{C}^M$ with overwhelming probability:

Given noisy intensity measurements

$$z_{\ell} := |\langle x, \varphi_{\ell} \rangle|^2 + \nu_{\ell},$$

if the noise-to-signal ratio satisfies $\mathrm{SNR}:=\frac{\|x\|_2^2}{\|\nu\|_2}\geq \frac{\sqrt{M}}{C'}$, then our phase retrieval procedure produces \tilde{x} with squared relative error

$$\frac{\|\tilde{x} - \mathrm{e}^{\mathrm{i}\theta}x\|_2^2}{\|x\|_2^2} \le K \sqrt{\frac{M}{\log M}} \,\,\mathrm{SNR}^{-1},$$

for some phase $\theta \in [0, 2\pi)$.

Polarization with Fourier Masks - Ongoing (with D. Mixon and Y. Chen)



We were able to design $O(\log M)$ Fourier Masks providing measurements that allow for reconstruction with the polarization algorithm,

both the vertex and edge measurements are contained in those $\mathcal{O}(\log M)$ designed Fourier Masks.

Thank You

A. S. Bandeira, A. Singer and D. A. Spielman, "A Cheeger Inequality for the Graph Connection Laplacian" arXiv:1204.3873

B. Alexeev, A. S. Bandeira, D. G. Mixon, and M. Fickus, "Phase retrieval with polarization" arXiv:1210.7752