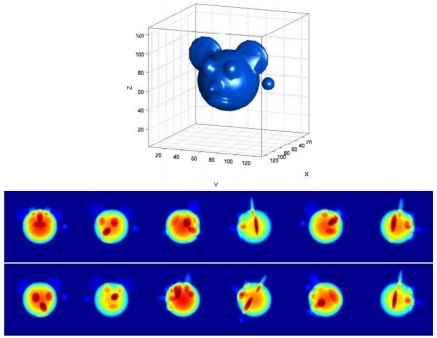
Message-passing algorithms for synchronization problems

Amelia Perry (MIT Mathematics)
with Afonso Bandeira, Ankur Moitra, and Alex Wein

Motivation: cryo-EM



Given many noisy 2D images of molecules, each with a different, unknown 3D rotation $g_u \in SO(3)$

Figure: courtesy of Amit Singer and Yoel Shkolnisky

[SS11] A. Singer and Y. Shkolnisky. Three-dimensional structure determination from common lines in Cryo-EM by eigenvectors and semidefinite programming. *SIAM J. Imaging Sciences*, 4(2):543–572, 2011.

Motivation: cryo-EM

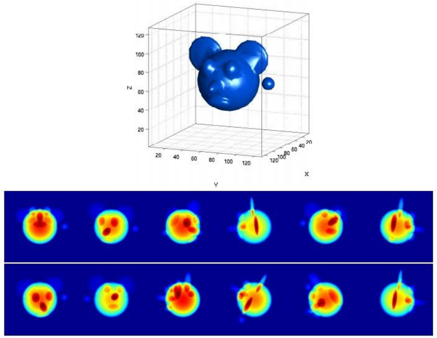


Figure: courtesy of Amit Singer and Yoel Shkolnisky

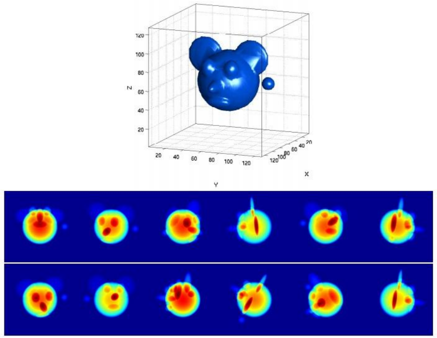
Given many noisy 2D images of molecules, each with a different, unknown 3D rotation $g_u \in SO(3)$

Comparing images u, v, we can learn a little about $g_u g_v^{-1}$

(relative alignment)

[SS11] A. Singer and Y. Shkolnisky. Three-dimensional structure determination from common lines in Cryo-EM by eigenvectors and semidefinite programming. *SIAM J. Imaging Sciences*, 4(2):543–572, 2011.

Motivation: cryo-EM



Given many noisy 2D images of molecules, each with a different, unknown 3D rotation $g_u \in SO(3)$

Comparing images u, v, we can learn a little about $g_u g_v^{-1}$

(relative alignment)

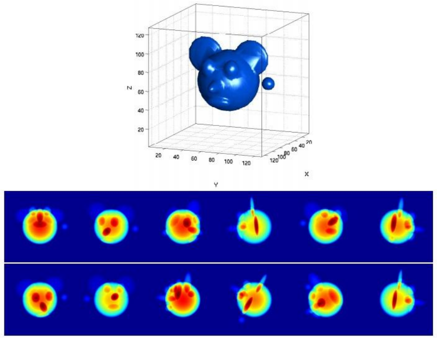
Q: how to synthesize into accurate estimation of all g_u?

(to reconstruct the molecule)

Figure: courtesy of Amit Singer and Yoel Shkolnisky

[SS11] A. Singer and Y. Shkolnisky. Three-dimensional structure determination from common lines in Cryo-EM by eigenvectors and semidefinite programming. *SIAM J. Imaging Sciences*, 4(2):543–572, 2011.

Motivation: cryo-EM



Given many noisy 2D images of molecules, each with a different, unknown 3D rotation $g_u \in SO(3)$

Comparing images u, v, we can learn a little about $g_u g_v^{-1}$
(relative alignment)

Q: how to synthesize into accurate estimation of all g_u?

(to reconstruct the molecule)

One answer: spectral methods (PCA) [CSSS10]

Figure: courtesy of Amit Singer and Yoel Shkolnisky

[SS11] A. Singer and Y. Shkolnisky. Three-dimensional structure determination from common lines in Cryo-EM by eigenvectors and semidefinite programming. *SIAM J. Imaging Sciences*, 4(2):543–572, 2011.

[CSSS10] R. R. Coifman, Y. Shkolnisky, F. J. Sigworth, A. Singer, "Reference free structure determination through eigenvectors of center of mass operators". Applied and Computational Harmonic Analysis, Volume 28, Issue 3 (2010).

$$\begin{pmatrix} g_1 g_1^{-1} & g_1 g_2^{-1} & g_1 g_3^{-1} & \\ g_2 g_1^{-1} & g_2 g_2^{-1} & g_2 g_3^{-1} & \\ g_3 g_1^{-1} & g_3 g_2^{-1} & g_3 g_3^{-1} & \\ & & & \ddots \end{pmatrix}$$

Motivation: cryo-EM

Trouble:

- PCA ignores the constraint to valid group elements.

$$\begin{pmatrix} g_1 g_1^{-1} & g_1 g_2^{-1} & g_1 g_3^{-1} & \\ g_2 g_1^{-1} & g_2 g_2^{-1} & g_2 g_3^{-1} & \\ g_3 g_1^{-1} & g_3 g_2^{-1} & g_3 g_3^{-1} & \\ & & & \ddots \end{pmatrix}$$

Motivation: cryo-EM

Trouble:

- PCA ignores the constraint to valid group elements.
- PCA effectively linearizes the observations, losing much of the signal.

$$\begin{pmatrix} g_1 g_1^{-1} & g_1 g_2^{-1} & g_1 g_3^{-1} & \\ g_2 g_1^{-1} & g_2 g_2^{-1} & g_2 g_3^{-1} & \\ g_3 g_1^{-1} & g_3 g_2^{-1} & g_3 g_3^{-1} & \\ & & & \ddots \end{pmatrix}$$

Motivation: cryo-EM

Challenge:

- PCA ignores the constraint to valid group elements. How do we make better use of this structure?
- PCA effectively linearizes the observations, losing much of the signal. How do we fully exploit our observations?

Motivation: cryo-EM

$$\begin{pmatrix} g_1 g_1^{-1} & g_1 g_2^{-1} & g_1 g_3^{-1} & \\ g_2 g_1^{-1} & g_2 g_2^{-1} & g_2 g_3^{-1} & \\ g_3 g_1^{-1} & g_3 g_2^{-1} & g_3 g_3^{-1} & \\ & & & \ddots \end{pmatrix}$$

Challenge:

- PCA ignores the constraint to valid group elements. How do we make better use of this structure?
- PCA effectively linearizes the observations, losing much of the signal. How do we fully exploit our observations?

We apply Approximate Message Passing, an existing framework for structured linear problems.

Motivation: cryo-EM

$$\begin{pmatrix} g_1 g_1^{-1} & g_1 g_2^{-1} & g_1 g_3^{-1} & \\ g_2 g_1^{-1} & g_2 g_2^{-1} & g_2 g_3^{-1} & \\ g_3 g_1^{-1} & g_3 g_2^{-1} & g_3 g_3^{-1} & \\ & & & \ddots \end{pmatrix}$$

Challenge:

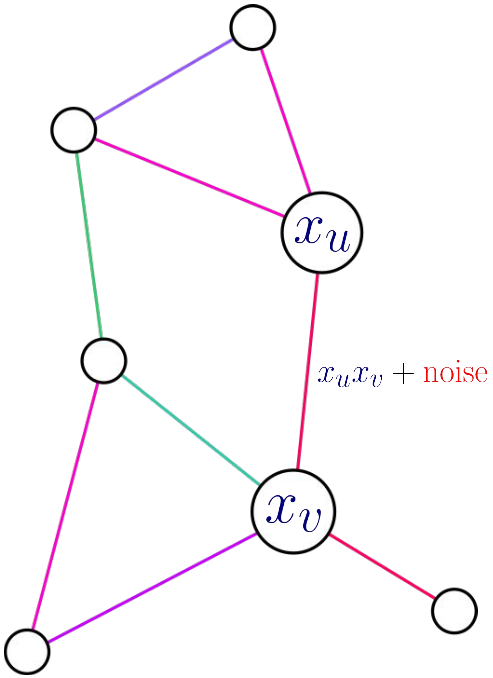
- PCA ignores the constraint to valid group elements. How do we make better use of this structure?
- PCA effectively linearizes the observations, losing much of the signal. How do we fully exploit our observations?

We apply Approximate Message Passing, an existing framework for structured linear problems.

We will build up towards cryo-EM via simpler problems.

Warm-up: $\mathbb{Z}/2$ synchronization

e.g. [HLL77], [Sin11], [ABBS14]



Learn $x \in \{\pm 1\}^n$

from noisy pairwise measurements...

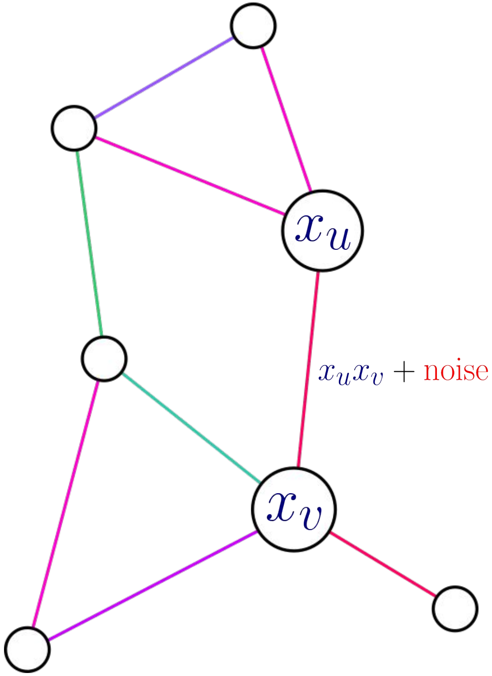
[HLL77] P. W. Holland, K. B. Laskey, and S. Leinhardt. "Stochastic blockmodels: First steps." *Social networks* 5.2 (1983): 109-137.

[Sin11] A. Singer. "Angular synchronization by eigenvectors and semidefinite programming." *Applied and computational harmonic analysis* 30.1 (2011).

[ABBS14] E. Abbe, A. S. Bandeira, A. Bracher, A, Singer. "Decoding binary node labels from censored edge measurements: Phase transition and efficient recovery." *IEEE Trans. Network Sci. Eng.* 1.1 (2014).

Warm-up: $\mathbb{Z}/2$ synchronization

e.g. [HLL77], [Sin11], [ABBS14]



Learn $x \in \{\pm 1\}^n$

from a matrix of noisy pairwise measurements:

$$\underset{(n \times n)}{Y} = \underbrace{\frac{\lambda}{n} x x^\top}_{\text{signal}} + \underbrace{\frac{1}{\sqrt{n}} W}_{\text{noise}}$$

λ: signal-to-noise ratio, W: Gaussian noise (GOE)

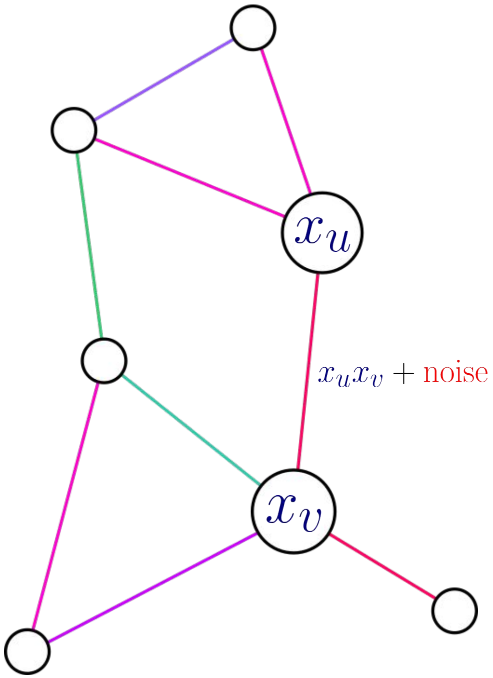
[HLL77] P. W. Holland, K. B. Laskey, and S. Leinhardt. "Stochastic blockmodels: First steps." *Social networks* 5.2 (1983): 109-137.

[Sin11] A. Singer. "Angular synchronization by eigenvectors and semidefinite programming." *Applied and computational harmonic analysis* 30.1 (2011).

[ABBS14] E. Abbe, A. S. Bandeira, A. Bracher, A, Singer. "Decoding binary node labels from censored edge measurements: Phase transition and efficient recovery." *IEEE Trans. Network Sci. Eng.* 1.1 (2014).

Warm-up: $\mathbb{Z}/2$ synchronization

e.g. [HLL77], [Sin11], [ABBS14]



Learn $x \in \{\pm 1\}^n$ (up to a global flip)

from a matrix of noisy pairwise measurements:

$$\underset{(n \times n)}{Y} = \underbrace{\frac{\lambda}{n} x x^\top}_{\text{signal}} + \underbrace{\frac{1}{\sqrt{n}} W}_{\text{noise}}$$

λ: signal-to-noise ratio, W: Gaussian noise (GOE)

[HLL77] P. W. Holland, K. B. Laskey, and S. Leinhardt. "Stochastic blockmodels: First steps." *Social networks* 5.2 (1983): 109-137.

[Sin11] A. Singer. "Angular synchronization by eigenvectors and semidefinite programming." *Applied and computational harmonic analysis* 30.1 (2011).

[ABBS14] E. Abbe, A. S. Bandeira, A. Bracher, A, Singer. "Decoding binary node labels from censored edge measurements: Phase transition and efficient recovery." *IEEE Trans. Network Sci. Eng.* 1.1 (2014).

$\mathbb{Z}/2$: some prior methods

$$\begin{pmatrix} 1 & x_1x_2 & x_1x_3 & \\ x_2x_1 & 1 & x_2x_3 & \\ x_3x_1 & x_3x_2 & 1 & \\ & & & \ddots \end{pmatrix}$$

$$Y = \frac{\lambda}{n}xx^\top + \frac{1}{\sqrt{n}}W$$

PCA: top eigenvector of Y [Sin11]

Power iteration: $v \leftarrow Yv$

[Sin11] A. Singer. "Angular synchronization by eigenvectors and semidefinite programming." *Applied and computational harmonic analysis* 30.1 (2011).

$\mathbb{Z}/2$: some prior methods

$$\begin{pmatrix} 1 & x_1x_2 & x_1x_3 & \\ x_2x_1 & 1 & x_2x_3 & \\ x_3x_1 & x_3x_2 & 1 & \\ & & & \ddots \end{pmatrix}$$

$$\boxed{Y = \frac{\lambda}{n} x x^\top + \frac{1}{\sqrt{n}} W}$$

PCA: top eigenvector of Y [Sin11]

Power iteration: $v \leftarrow Yv$

Projected power iteration ("majority dynamics") [Bou16]

$$v \leftarrow \operatorname{sgn}(Yv)$$

[Sin11] A. Singer. "Angular synchronization by eigenvectors and semidefinite programming." *Applied and computational harmonic analysis* 30.1 (2011).

[Bou16] N. Boumal, "Nonconvex phase synchronization". *arXiv:1601.06114* (2016).

$\mathbb{Z}/2$: some prior methods

$$\begin{pmatrix} 1 & x_1x_2 & x_1x_3 & \\ x_2x_1 & 1 & x_2x_3 & \\ x_3x_1 & x_3x_2 & 1 & \\ & & & \ddots \end{pmatrix}$$

$$\boxed{Y = \frac{\lambda}{n} x x^\top + \frac{1}{\sqrt{n}} W}$$

PCA: top eigenvector of Y [Sin11]

Power iteration: $v \leftarrow Yv$

Projected power iteration (“majority dynamics”) [Bou16]

$$v \leftarrow \mathrm{sgn}(Yv)$$

Semidefinite programming [Sin11, BCS15]

[Sin11] A. Singer. “Angular synchronization by eigenvectors and semidefinite programming.” *Applied and computational harmonic analysis* 30.1 (2011).

[Bou16] N. Boumal, “Nonconvex phase synchronization”. *arXiv:1601.06114* (2016).

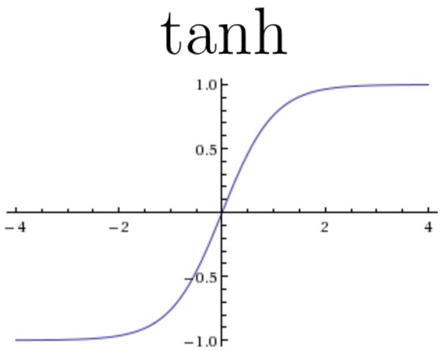
[BCS15] A. S. Bandeira, Y. Chen, and A. Singer. "Non-unique games over compact groups and orientation estimation in cryo-EM." *arXiv:1505.03840* (2015).

$\mathbb{Z}/2$: try soft thresholding?

Soft thresholding: $v \leftarrow Y f(v)$

(f is applied entry-wise to v)

$\mathbb{Z}/2$: try soft thresholding?

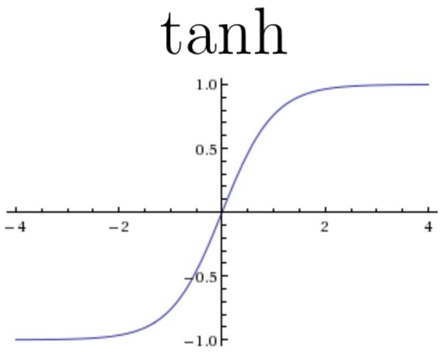


Soft thresholding: $v \leftarrow Y f(v)$

(f is applied entry-wise to v)

Optimal: $f(v) = \tanh(\lambda v)$

$\mathbb{Z}/2$: try soft thresholding?



Soft thresholding: $v \leftarrow Y f(v)$

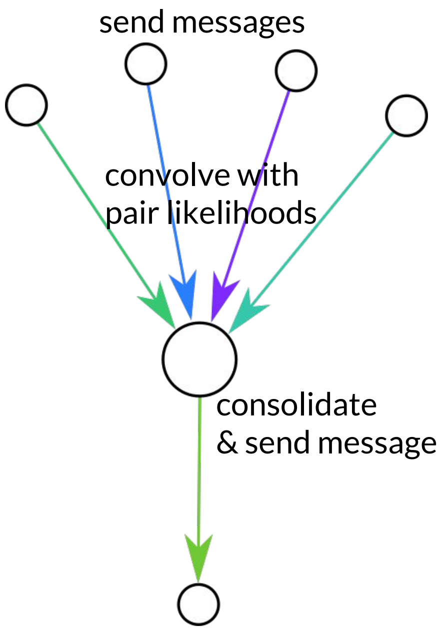
(f is applied entry-wise to v)

Optimal: $f(v) = \tanh(\lambda v)$

Outputs in $[-1, 1]$ capture “confidence” of estimates.

So this iterative algorithm passes around distributions...

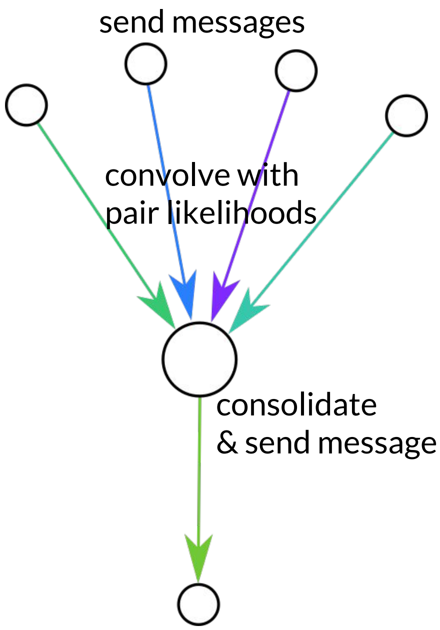
Belief Propagation (BP)



In each iteration, nodes send each other 'messages': their posterior **distributions** given the previous iteration.

[Pea86] Pearl, Judea. "Fusion, propagation, and structuring in belief networks." *Artificial intelligence* 29.3 (1986): 241-288.

[MPV86] M. Mézard, G. Parisi, and M. A. Virasoro. "SK model: The replica solution without replicas." *Europhys. Lett* 1.2 (1986): 77-82.



Belief Propagation (BP)

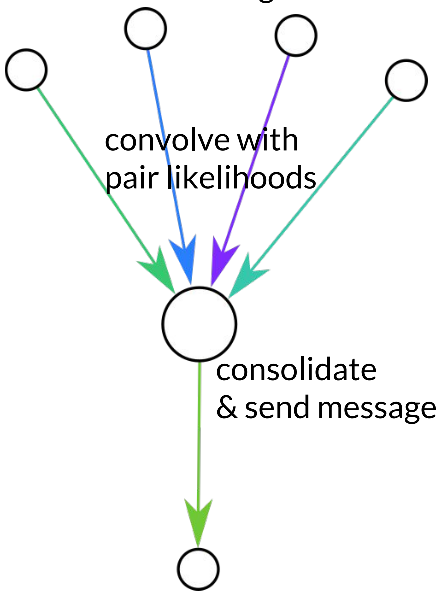
In each iteration, nodes send each other 'messages': their posterior **distributions** given the previous iteration.

Caveat: no backtracking!

[Pea86] Pearl, Judea. "Fusion, propagation, and structuring in belief networks." *Artificial intelligence* 29.3 (1986): 241-288.

[MPV86] M. Mézard, G. Parisi, and M. A. Virasoro. "SK model: The replica solution without replicas." *Europhys. Lett* 1.2 (1986): 77-82.

Belief Propagation (BP)



In each iteration, nodes send each other ‘messages’: their posterior **distributions** given the previous iteration.

Caveat: no backtracking!

Arose simultaneously as ‘cavity equations’ in physics.

Not rigorously well-understood.
(e.g. random SAT)

[Pea86] Pearl, Judea. "Fusion, propagation, and structuring in belief networks." *Artificial intelligence* 29.3 (1986): 241-288.

[MPV86] M. Mézard, G. Parisi, and M. A. Virasoro. "SK model: The replica solution without replicas." *Europhys. Lett* 1.2 (1986): 77-82.

Approximate Message Passing (AMP)

Simplifies belief propagation

- Exploits central limit theorems for dense graphs
- Encodes messages (distributions) in a few parameters

Approximate Message Passing (AMP)

Simplifies belief propagation

- Exploits central limit theorems for dense graphs
- Encodes messages (distributions) in a few parameters

Frequently yields state-of-the-art statistical performance.

- Compressed sensing [DMM09]
- Sparse PCA [DM14], non-negative / cone PCA [DMR14]

[DMM09] D. L. Donoho., A. Maleki, and A. Montanari. "Message-passing algorithms for compressed sensing." *P. Natl. Acad. Sci. USA* 106.45 (2009).

[DM14] Y. Deshpande and A. Montanari. Information-theoretically optimal sparse PCA." *IEEE ISIT*, 2014.

[DMR14] Y. Deshpande, A. Montanari, and E. Richard. "Cone-constrained Principal Component Analysis." *NIPS*, 2014.

Approximate Message Passing (AMP)

Simplifies belief propagation

- Exploits central limit theorems for dense graphs
- Encodes messages (distributions) in a few parameters

Frequently yields state-of-the-art statistical performance.

- Compressed sensing [DMM09]
- Sparse PCA [DM14], non-negative / cone PCA [DMR14]

Rigorous proof framework [BM11]

[BM11] M. Bayati and A. Montanari. "The dynamics of message passing on dense graphs, with applications to compressed sensing." *IEEE T. Inform. Theory* 57.2 (2011).

[DMM09] D. L. Donoho., A. Maleki, and A. Montanari. "Message-passing algorithms for compressed sensing." *P. Natl. Acad. Sci. USA* 106.45 (2009).

[DM14] Y. Deshpande and A. Montanari. Information-theoretically optimal sparse PCA." *IEEE ISIT*, 2014.

[DMR14] Y. Deshpande, A. Montanari, and E. Richard. "Cone-constrained Principal Component Analysis." *NIPS*, 2014.

AMP for $\mathbb{Z}/2$ synchronization

[DAM15]

$$c^t = \lambda Y v^{t-1} - \lambda^2(1 - \langle v^{t-1} \rangle^2) v^{t-2}$$

—Onsager correction—

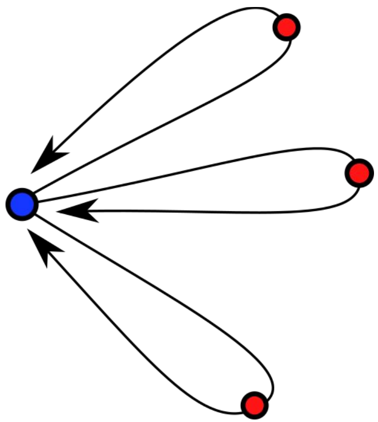
$$v^t = \tanh(c^t)$$

—soft thresholding—

[DAM15] Y. Deshpande, E. Abbe, and A. Montanari. "Asymptotic mutual information for the two-groups stochastic block model." *arXiv:1507.08685* (2015).

AMP for $\mathbb{Z}/2$ synchronization

[DAM15]



$$c^t = \lambda Y v^{t-1} - \lambda^2 (1 - \langle (v^{t-1})^2 \rangle) v^{t-2}$$

—Onsager correction—

$$v^t = \tanh(c^t)$$

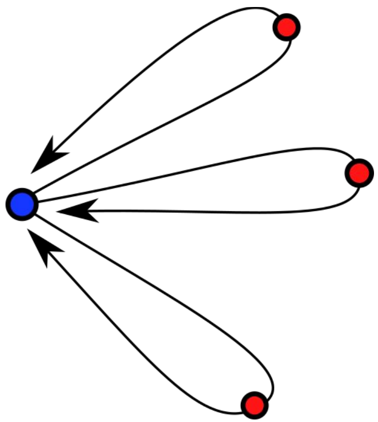
—soft thresholding—

Onsager term corrects for backtracking, to leading order.

[DAM15] Y. Deshpande, E. Abbe, and A. Montanari. "Asymptotic mutual information for the two-groups stochastic block model." *arXiv:1507.08685* (2015).

AMP for $\mathbb{Z}/2$ synchronization

[DAM15]



$$c^t = \lambda Y v^{t-1} - \lambda^2(1 - \langle (v^{t-1})^2 \rangle) v^{t-2}$$

—Onsager correction—

$$v^t = \tanh(c^t)$$

—soft thresholding—

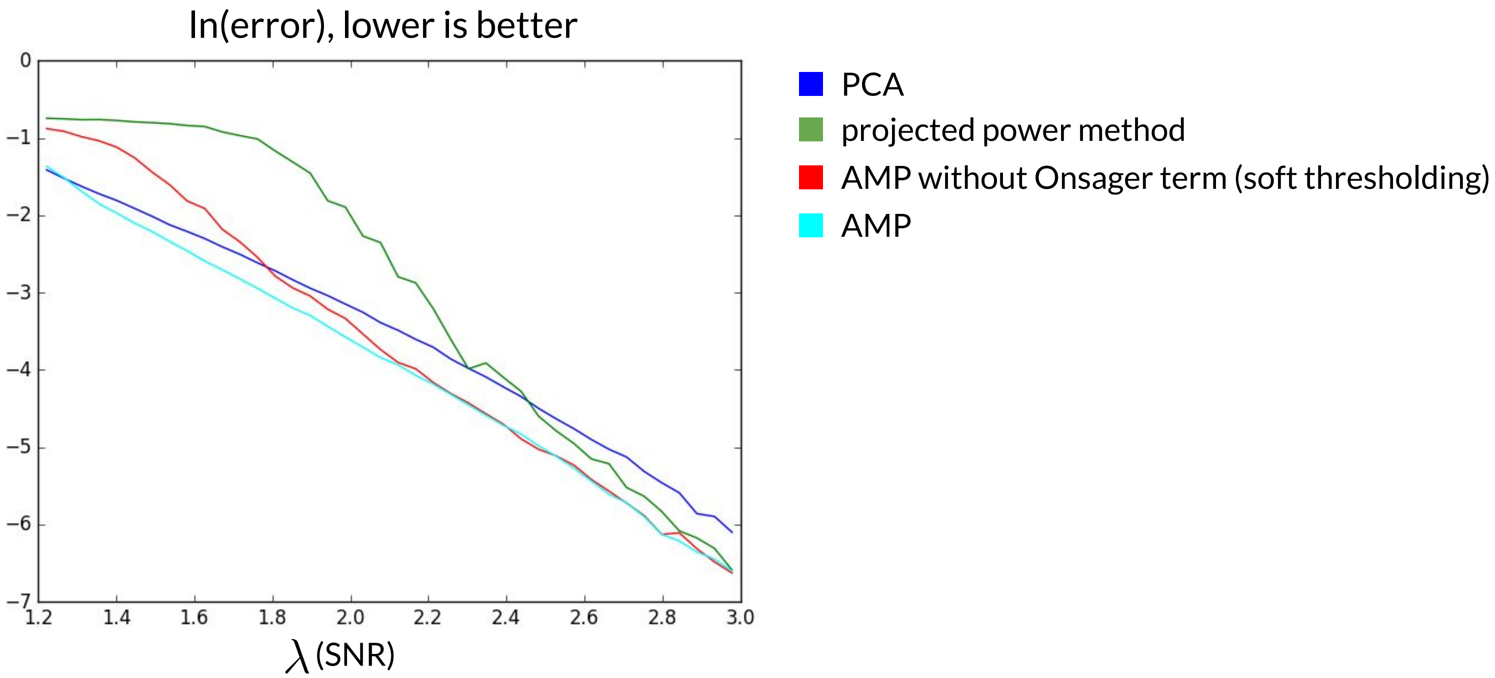
Onsager term corrects for backtracking, to leading order.

Each entry of v^t encodes a distribution over $\{\pm 1\}$.

(as the expectation)

[DAM15] Y. Deshpande, E. Abbe, and A. Montanari. "Asymptotic mutual information for the two-groups stochastic block model." *arXiv:1507.08685* (2015).

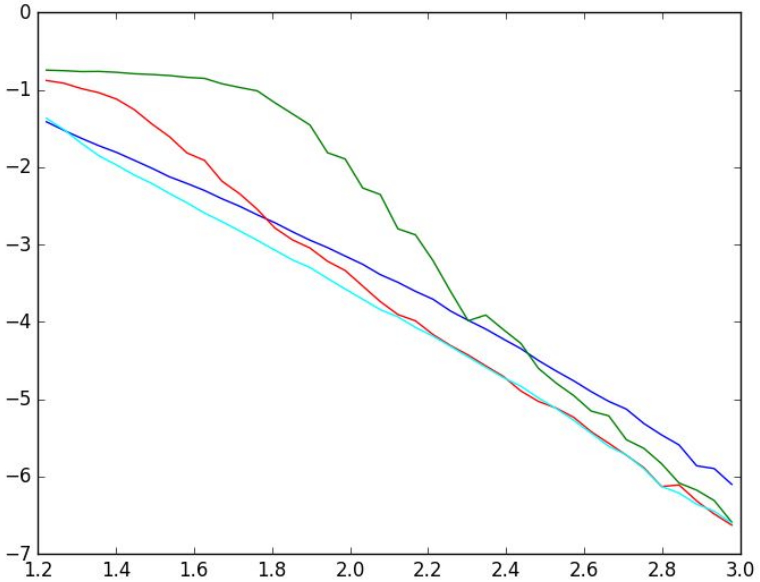
Comparison of Methods



[DAM15] Y. Deshpande, E. Abbe, and A. Montanari. "Asymptotic mutual information for the two-groups stochastic block model." *arXiv:1507.08685* (2015).

Comparison of Methods

ln(error), lower is better



λ (SNR)

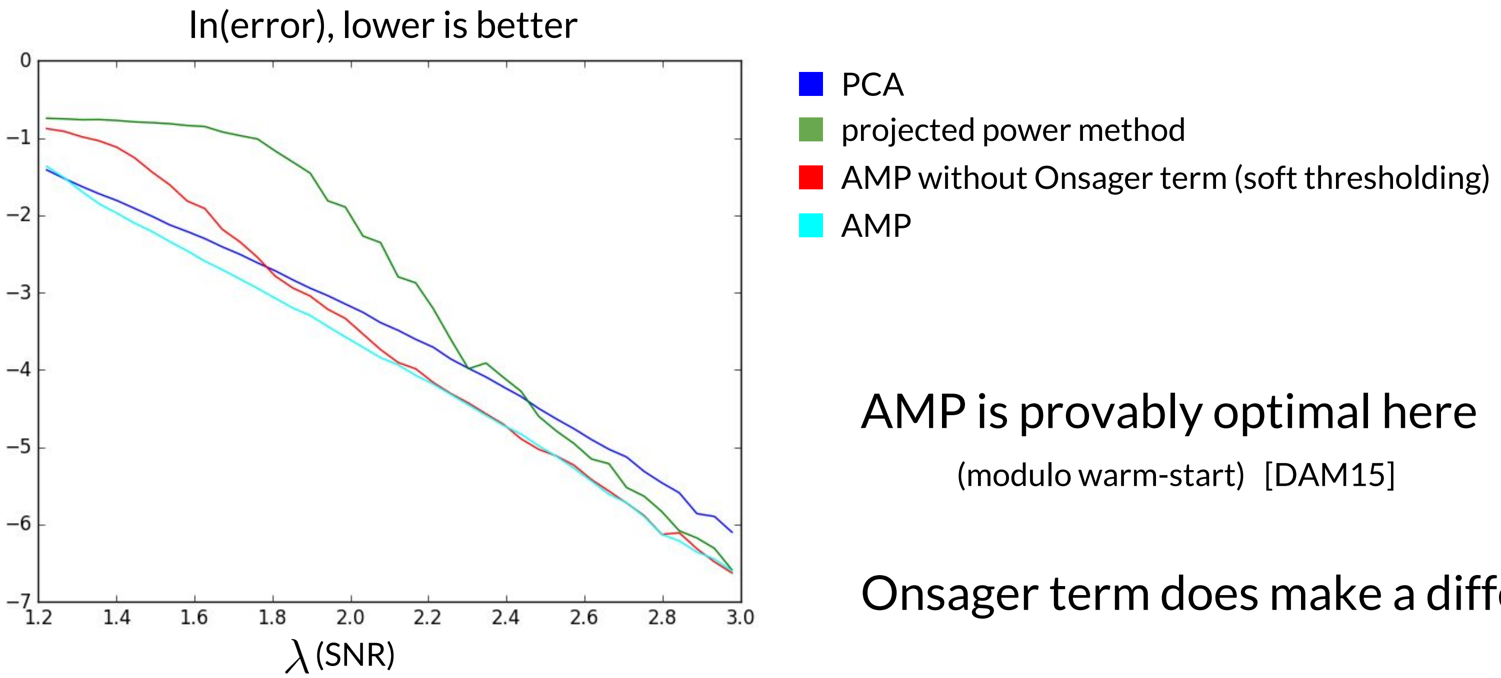
- PCA
- projected power method
- AMP without Onsager term (soft thresholding)
- AMP

AMP is provably optimal here

(modulo warm-start) [DAM15]

[DAM15] Y. Deshpande, E. Abbe, and A. Montanari. "Asymptotic mutual information for the two-groups stochastic block model." *arXiv:1507.08685* (2015).

Comparison of Methods



AMP is provably optimal here

(modulo warm-start) [DAM15]

Onsager term does make a difference!

[DAM15] Y. Deshpande, E. Abbe, and A. Montanari. "Asymptotic mutual information for the two-groups stochastic block model." *arXiv:1507.08685* (2015).

Motivation: multireference alignment

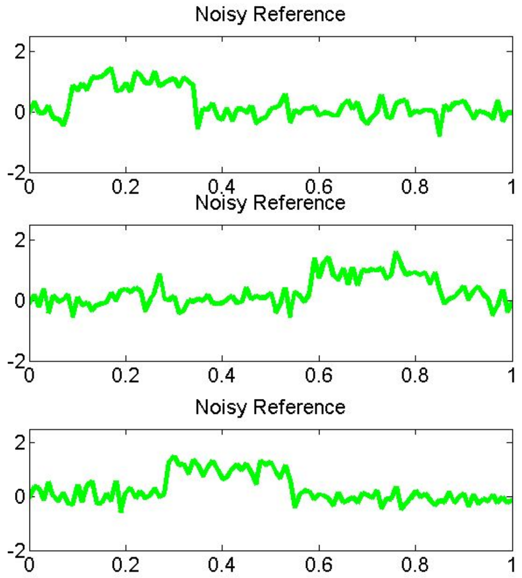
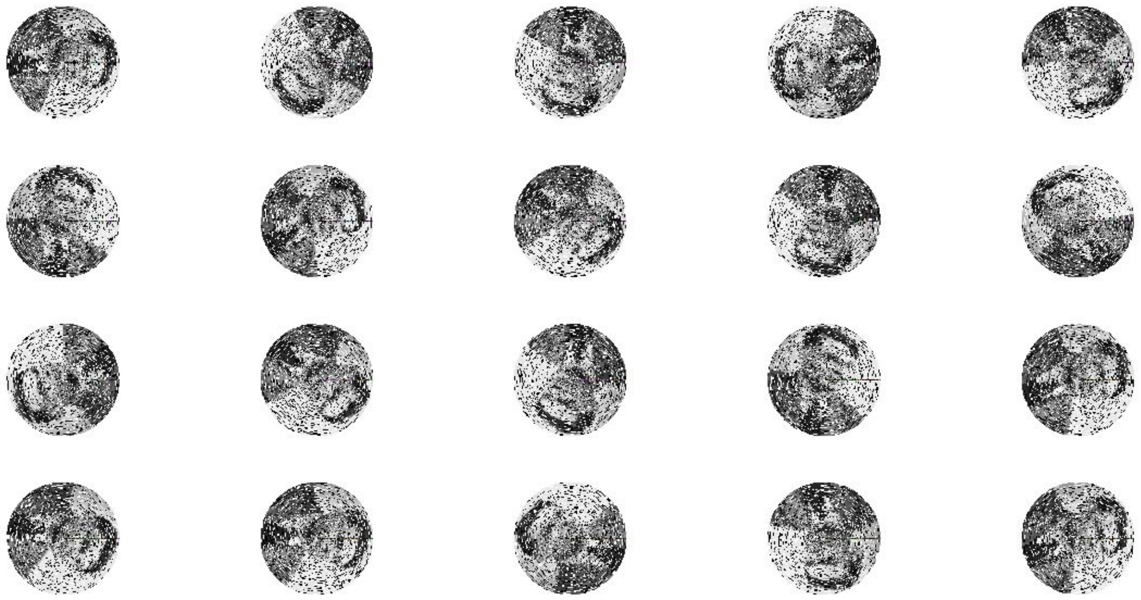


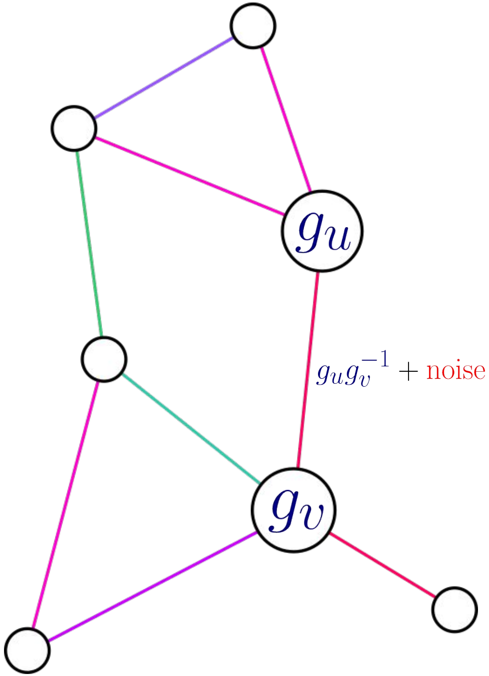
Figure: A. S. Bandeira, M. Charikar, A. Singer, and A. Zhu. Multireference alignment using semidefinite programming. *5th Innovations in Theoretical Computer Science (ITCS 2014)*, 2014.

Motivation: angular synchronization



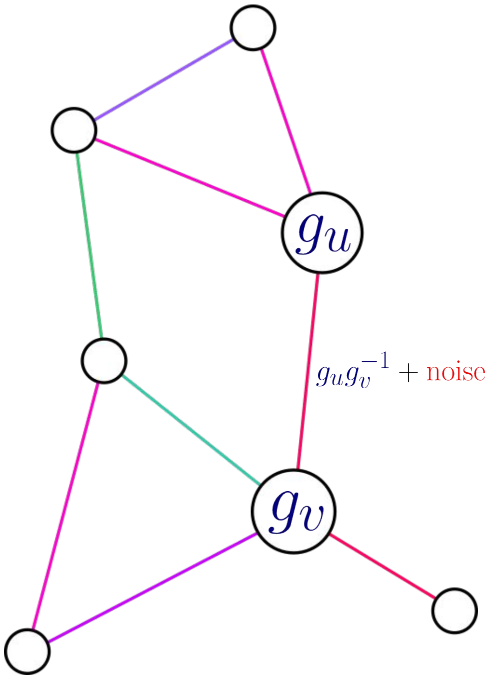
Synchronization over any group

[BCS15]



Learn a vector g of group elements
from noisy observations of $g_u g_v^{-1}$.
(up to global right-multiplication by a group element)

[BCS15] A. S. Bandeira, Y. Chen, and A. Singer. "Non-unique games over compact groups and orientation estimation in cryo-EM." *arXiv:1505.03840* (2015).



Synchronization over any group

[BCS15]

Learn a vector g of group elements
from noisy observations of $g_u g_v^{-1}$.
(up to global right-multiplication by a group element)

Our contribution: AMP for synchronization over any* group, with any* noise model

(e.g. $\mathbb{Z}/L,\ U(1),\ SO(3),$ compact Lie groups)

[BCS15] A. S. Bandeira, Y. Chen, and A. Singer. "Non-unique games over compact groups and orientation estimation in cryo-EM." *arXiv:1505.03840* (2015).

U(1) synchronization

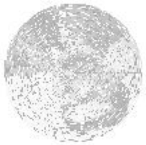
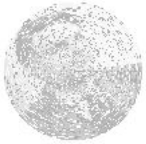
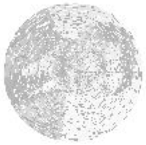
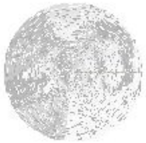
Observe $Y^{(1)} = \frac{\lambda}{n} x x^* + \frac{1}{\sqrt{n}} W^{(1)}$

—signal— —noise—

SDP is tight [BNS14]

[BNS14] A. S. Bandeira, N. Boumal, and A. Singer. "Tightness of the maximum likelihood semidefinite relaxation for angular synchronization." *arXiv:1411.3272* (2014).

U(1) with two frequencies

Observe
$$Y^{(1)} = \frac{\lambda}{n} x x^* + \frac{1}{\sqrt{n}} W^{(1)}$$
$$Y^{(2)} = \frac{\lambda}{n} x^2 (x^2)^* + \frac{1}{\sqrt{n}} W^{(2)}$$

—signal— —noise—

Multiple channels of pairwise information.

U(1) with multiple frequencies

Observe $Y^{(1)} = \frac{\lambda}{n} x x^* + \frac{1}{\sqrt{n}} W^{(1)}$

$$Y^{(2)} = \frac{\lambda}{n} x^2 (x^2)^* + \frac{1}{\sqrt{n}} W^{(2)}$$

...

$$Y^{(k)} = \underbrace{\frac{\lambda}{n} x^k (x^k)^*}_{\text{signal}} + \underbrace{\frac{1}{\sqrt{n}} W^{(k)}}_{\text{noise}}$$

—signal— —noise—

Multiple channels of pairwise information.

U(1) with multiple frequencies

Observe
$$Y^{(1)} = \frac{\lambda}{n} x x^* + \frac{1}{\sqrt{n}} W^{(1)}$$
$$Y^{(2)} = \frac{\lambda}{n} x^2 (x^2)^* + \frac{1}{\sqrt{n}} W^{(2)}$$
$$\cdots$$
$$Y^{(k)} = \frac{\lambda}{n} x^k (x^k)^* + \frac{1}{\sqrt{n}} W^{(k)}$$

—signal— —noise—

Multiple channels of pairwise information.

Multiple frequencies corresponds to nonlinear observations.

No clear PCA approach that couples them.

U(1): AMP algorithm

Represent distributions by discretizations?

U(1): AMP algorithm

Represent distributions by discretizations?

Discretizing $SO(3)$ is awkward: impossible without breaking symmetry.

Rotating a discretized function is lossy.

U(1): AMP algorithm

Represent distributions by Fourier coeffs of...

density?

$$\frac{d\mathbb{P}(g_u)}{d\theta} = \sum_k v_u^{(k)} e^{ik\theta}$$

U(1): AMP algorithm

Represent distributions by Fourier coeffs of...

density?

$$\frac{d\mathbb{P}(g_u)}{d\theta} = \sum_k v_u^{(k)} e^{ik\theta}$$

log-likelihood?

$$\log \frac{d\mathbb{P}(g_u)}{d\theta} + \text{const} = \sum_k c_u^{(k)} e^{ik\theta}$$

U(1): AMP algorithm

Represent distributions by Fourier coeffs of...

density?

$$\frac{d\mathbb{P}(g_u)}{d\theta} = \sum_k v_u^{(k)} e^{ik\theta}$$

log-likelihood?

$$\log \frac{d\mathbb{P}(g_u)}{d\theta} + \text{const} = \sum_k c_u^{(k)} e^{ik\theta}$$

Iteration:

$$x^{(k)} \leftarrow \lambda Y^{(k)} v^{(k)} + \text{onsager}$$ (messaging)

$$v_u^{(\bullet)} \leftarrow f(x_u^{(\bullet)})$$ (consolidation)

U(1): AMP algorithm

Represent distributions by Fourier coeffs of...

density?

$$\frac{d\mathbb{P}(g_u)}{d\theta} = \sum_k v_u^{(k)} e^{ik\theta}$$

log-likelihood?

$$\log \frac{d\mathbb{P}(g_u)}{d\theta} + \text{const} = \sum_k c_u^{(k)} e^{ik\theta}$$

Iteration: $c^{(k)} \leftarrow \lambda Y^{(k)} v^{(k)} + \text{onsager}$ (messaging)

$v_u^{(\bullet)} \leftarrow f(c_u^{(\bullet)})$ (consolidation)

f is the transformation from $c_u^{(\bullet)}$ to $v_u^{(\bullet)}$!

U(1): the nonlinear transformation

f converts Fourier coefficients of $g : U(1) \to \mathbb{R}$
into Fourier coefficients of $\exp(g)$, and then normalizes.

This couples Fourier components $Y^{(k)}$ of the measurements.

U(1): the nonlinear transformation

f converts Fourier coefficients of $g : U(1) \to \mathbb{R}$
into Fourier coefficients of $\exp(g)$, and then normalizes!

This couples Fourier components $Y^{(k)}$ of the measurements.

$\mathbb{Z}/2$: Only "Fourier coefficient" is $g(1) - g(-1)$.

U(1): the nonlinear transformation

f converts Fourier coefficients of $g : U(1) \to \mathbb{R}$
into Fourier coefficients of $\exp(g)$, and then normalizes!

This couples Fourier components $Y^{(k)}$ of the measurements.

$\mathbb{Z}/2$: Only "Fourier coefficient" is $g(1) - g(-1)$.

Then,

c $-c$

U(1): the nonlinear transformation

f converts Fourier coefficients of $g : U(1) \to \mathbb{R}$
into Fourier coefficients of $\exp(g)$, and then normalizes!

This couples Fourier components $Y^{(k)}$ of the measurements.

$\mathbb{Z}/2$: Only "Fourier coefficient" is $g(1) - g(-1)$.

Then,

$$e^c \quad e^{-c}$$

U(1): the nonlinear transformation

f converts Fourier coefficients of $g : U(1) \to \mathbb{R}$
into Fourier coefficients of $\exp(g)$, and then normalizes!

This couples Fourier components $Y^{(k)}$ of the measurements.

$\mathbb{Z}/2$: Only "Fourier coefficient" is $g(1) - g(-1)$.

Then,

$$e^c - e^{-c}$$

U(1): the nonlinear transformation

f converts Fourier coefficients of $g : U(1) \to \mathbb{R}$
into Fourier coefficients of $\exp(g)$, and then normalizes!

This couples Fourier components $Y^{(k)}$ of the measurements.

$\mathbb{Z}/2$: Only "Fourier coefficient" is $g(1) - g(-1)$.

Then,

$$f(c) = \frac{e^c - e^{-c}}{e^c + e^{-c}}$$

U(1): the nonlinear transformation

f converts Fourier coefficients of $g : U(1) \to \mathbb{R}$
into Fourier coefficients of $\exp(g)$, and then normalizes!

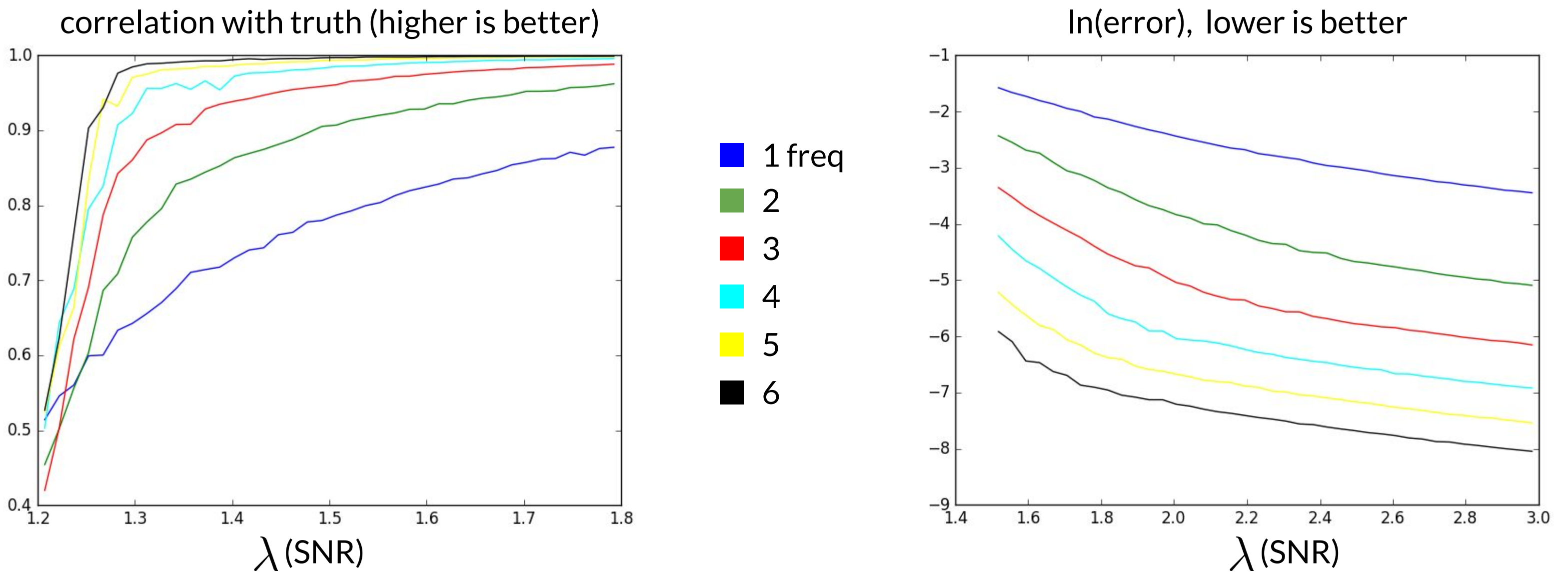
This couples Fourier components $Y^{(k)}$ of the measurements.

$\mathbb{Z}/2$: Only "Fourier coefficient" is $g(1) - g(-1)$.

Then,

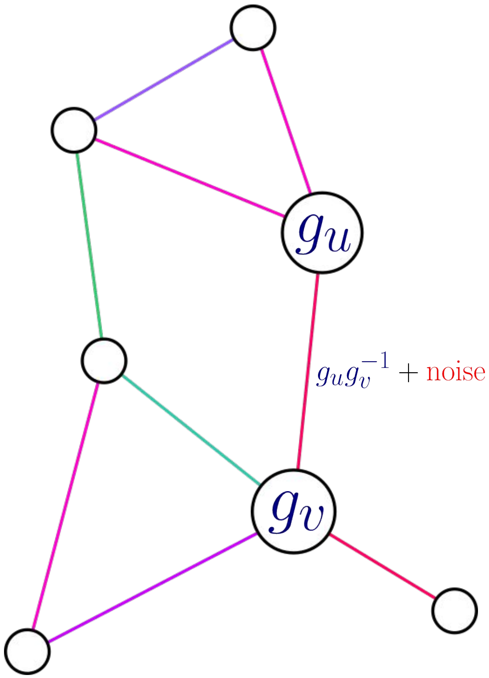
$$f(c) = \frac{e^c - e^{-c}}{e^c + e^{-c}} = \tanh(c)$$

U(1): empirical results

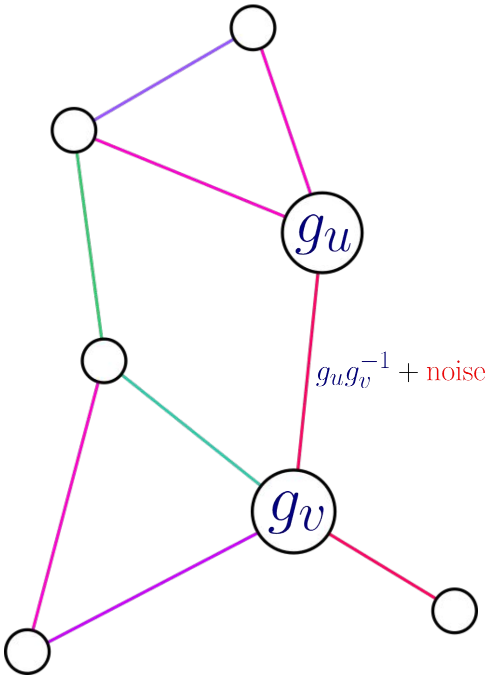


AMP can synthesize information across multiple frequencies.

Synchronization over any* group



Fourier theory becomes **representation theory**.

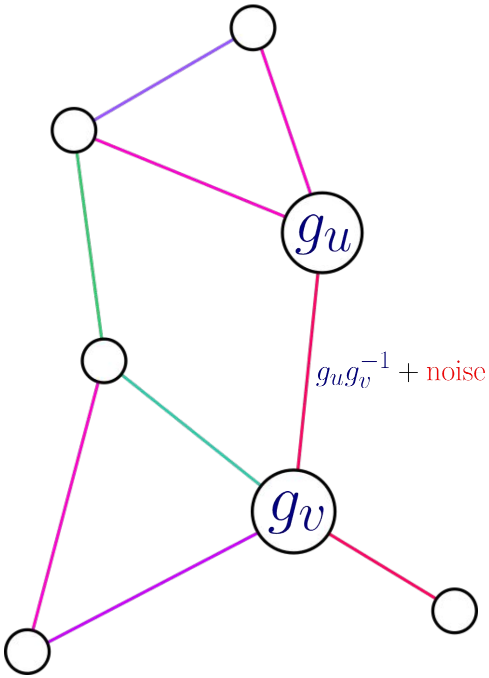


Synchronization over any* group

Fourier theory becomes **representation theory**.

Peter–Weyl theorem: any* $f : G \to \mathbb{C}$ decomposes into normal modes:

$$f(g) = \sum_{\text{irreps } \rho} \left\langle C^{(\rho)}, \rho(g) \right\rangle$$



Synchronization over any* group

Fourier theory becomes **representation theory**.

Peter–Weyl theorem: any* $f : G \to \mathbb{C}$ decomposes into normal modes:

$$f(g) = \sum_{\text{irreps } \rho} \left\langle C^{(\rho)}, \rho(g) \right\rangle$$

Apply this to distributions to describe the AMP iterations.

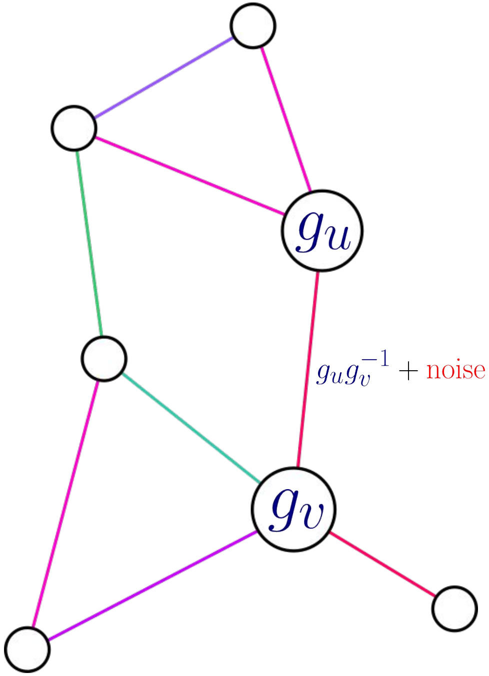
$$C^{(\rho)} \leftarrow Y^{(\rho)} V^{(\rho)} + \text{onsager}$$

(messaging)

$$V_u^{(\bullet)} \leftarrow f(C_u^{(\bullet)})$$

(consolidation: exp & normalize)

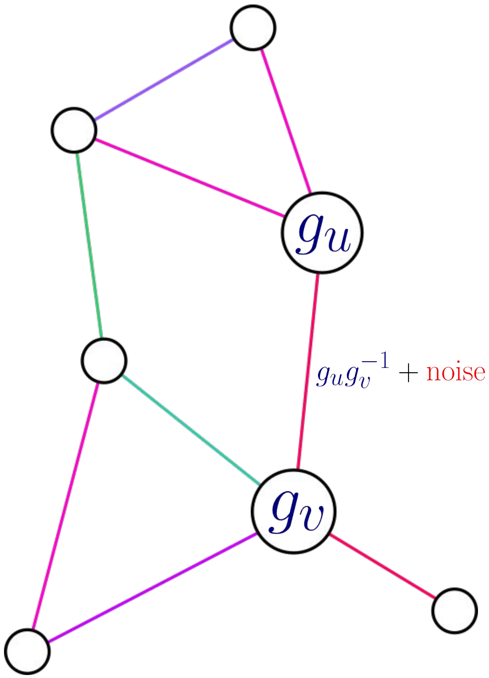
Noise models & non-unique games



What sort of noise?

[BCS15] A. S. Bandeira, Y. Chen, and A. Singer. "Non-unique games over compact groups and orientation estimation in cryo-EM." *arXiv:1505.03840* (2015).

Noise models & non-unique games

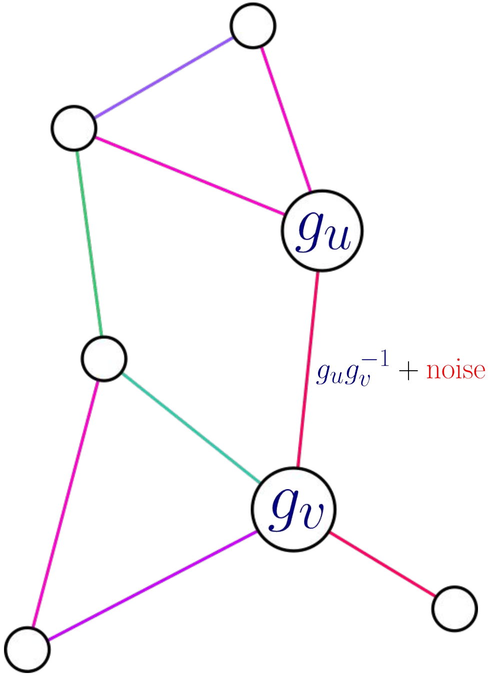


We assume pair measurements have independent noise.

Likelihood factors over edges: $$\log \mathcal{L}(g) = \sum_{u,v} \ell_{u,v}(g_u g_v^{-1})$$

[BCS15] A. S. Bandeira, Y. Chen, and A. Singer. "Non-unique games over compact groups and orientation estimation in cryo-EM." *arXiv:1505.03840* (2015).

Noise models & non-unique games



We assume pair measurements have independent noise.

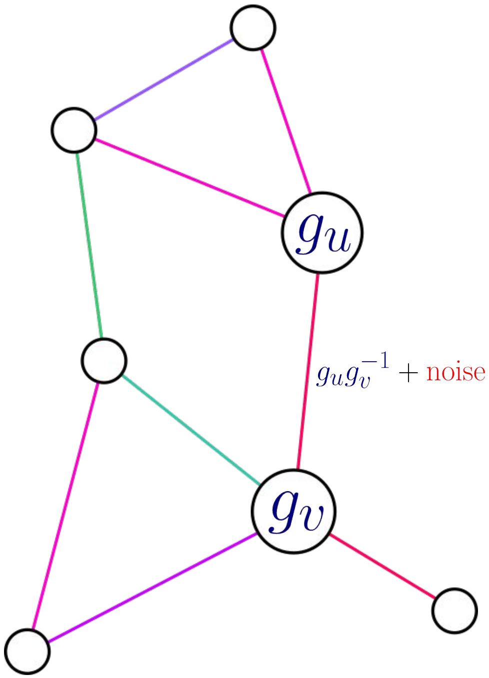
Likelihood factors over edges: $\log \mathcal{L}(g) = \sum_{u,v} \ell_{u,v}(g_u g_v^{-1})$

Assemble matrix coefficients of $\ell_{u,v}$ into matrices $Y^{(\rho)}$.

$$Y^{(\rho)} = \begin{pmatrix} \hat{\ell}_{1,1}(\rho) & \hat{\ell}_{1,2}(\rho) & \hat{\ell}_{1,3}(\rho) & \\ \hat{\ell}_{2,1}(\rho) & \hat{\ell}_{2,2}(\rho) & \hat{\ell}_{2,3}(\rho) & \\ \hat{\ell}_{3,1}(\rho) & \hat{\ell}_{3,2}(\rho) & \hat{\ell}_{3,3}(\rho) & \\ & & & \ddots \end{pmatrix}$$

[BCS15] A. S. Bandeira, Y. Chen, and A. Singer. "Non-unique games over compact groups and orientation estimation in cryo-EM." *arXiv:1505.03840* (2015).

Noise models & non-unique games



We assume pair measurements have independent noise.

Likelihood factors over edges: $\log \mathcal{L}(g) = \sum_{u,v} \ell_{u,v}(g_u g_v^{-1})$

Assemble matrix coefficients of $\ell_{u,v}$ into matrices $Y^{(\rho)}$.

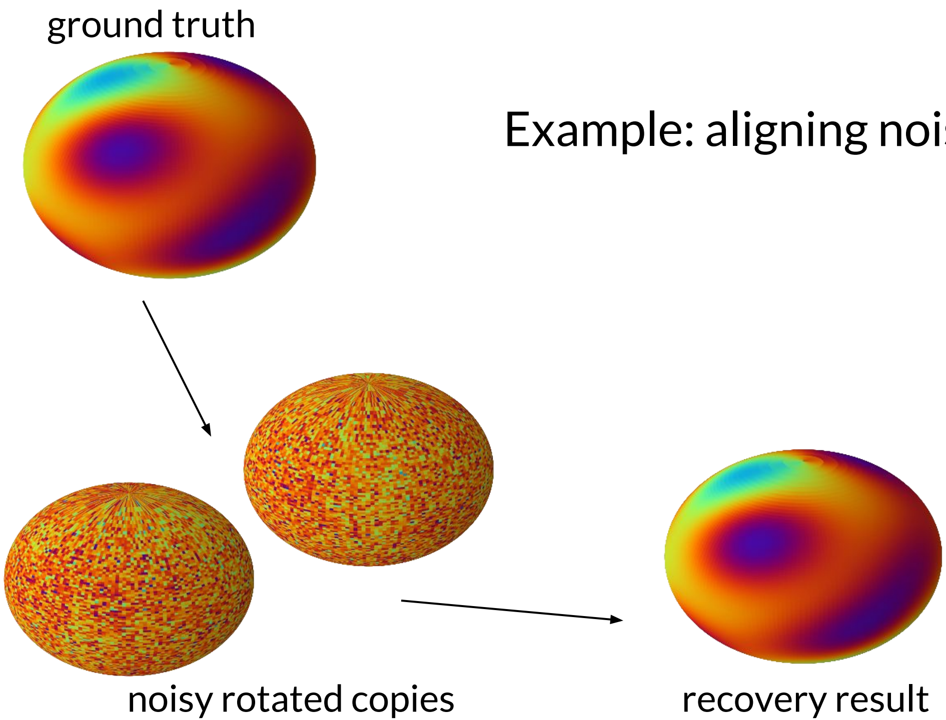
$$Y^{(\rho)} = \begin{pmatrix} \hat{\ell}_{1,1}(\rho) & \hat{\ell}_{1,2}(\rho) & \hat{\ell}_{1,3}(\rho) & \\ \hat{\ell}_{2,1}(\rho) & \hat{\ell}_{2,2}(\rho) & \hat{\ell}_{2,3}(\rho) & \\ \hat{\ell}_{3,1}(\rho) & \hat{\ell}_{3,2}(\rho) & \hat{\ell}_{3,3}(\rho) & \\ & & & \ddots \end{pmatrix}$$

$$C^{(\rho)} \leftarrow Y^{(\rho)} V^{(\rho)} + \text{onsager}$$
$$V_u^{(\bullet)} \leftarrow f(C_u^{(\bullet)})$$

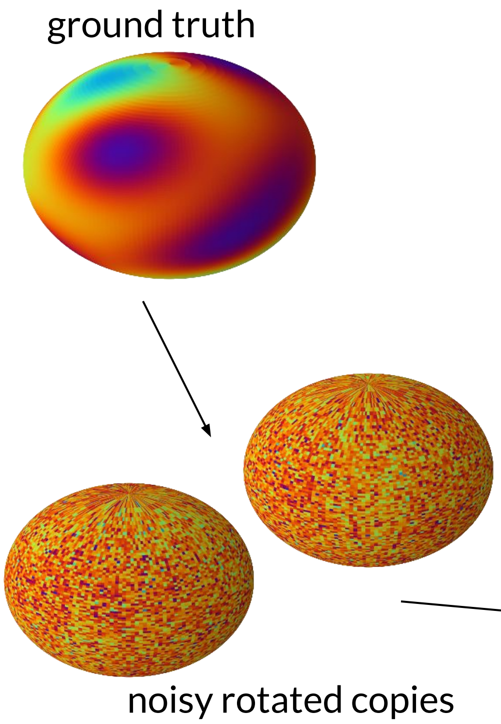
[BCS15] A. S. Bandeira, Y. Chen, and A. Singer. "Non-unique games over compact groups and orientation estimation in cryo-EM." *arXiv:1505.03840* (2015).

AMP for SO(3) synchronization

Example: aligning noisy copies of images on the sphere.

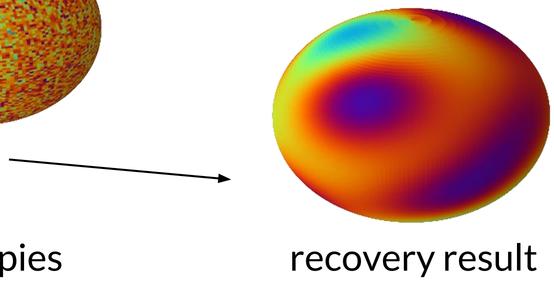


AMP for SO(3) synchronization



Example: aligning noisy copies of images on the sphere.

To form $Y^{(\rho)}$: decompose images into spherical harmonics.
j^{th} representation compares the degree j harmonics.



Ongoing work:

Correct AMP for per-vertex noise

- Cryo-EM and other problems have noise on each observation, not on each pair comparison.

Ongoing work:

Correct AMP for per-vertex noise

Cryo-EM and other problems have noise on each observation, not on each pair comparison.

We can derive correct AMP for each stochastic model—but can we make AMP tune itself?

More robust to uncertain noise models?

Ongoing work:

Correct AMP for per-vertex noise

Cryo-EM and other problems have noise on each observation, not on each pair comparison.

We can derive correct AMP for each stochastic model—but can we make AMP tune itself?

More robust to uncertain noise models?

What are the information limits of synchronization problems?

Does AMP match them?

Thanks!

Any questions?