Planted Cliques, Iterative Thresholding and Message Passing Algorithms

Yash Deshpande and Andrea Montanari

Stanford University

November 5, 2013

Problem Definition

Given distributions Q_0, Q_1,

A Set $\to S \subset [n]$

Data $\to A_{ij} \sim \begin{cases} Q_1 & \text{if } i, j \in S \\ Q_0 & \text{otherwise.} \end{cases}$

$A_{ij} = A_{ji}$

Problem: Given A, identify S

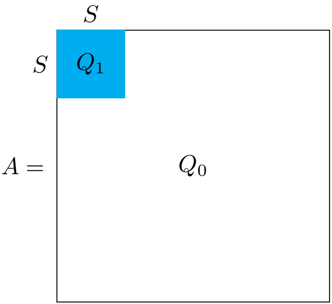
Problem Definition

Given distributions Q_0, Q_1,

A Set $\rightarrow S \subset [n]$

$$\text{Data} \rightarrow A_{ij} \sim \begin{cases} Q_1 & \text{if } i,j \in S \\ Q_0 & \text{otherwise.} \end{cases}$$

$$A_{ij} = A_{ji}$$



Problem: Given A, identify S

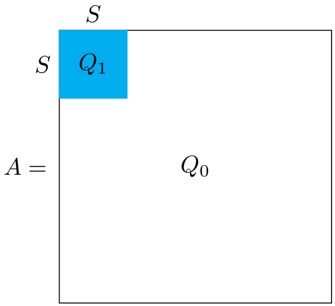
Problem Definition

Given distributions Q_0, Q_1,

A Set $\rightarrow S \subset [n]$

$$\text{Data} \rightarrow A_{ij} \sim \begin{cases} Q_1 & \text{if } i, j \in S \\ Q_0 & \text{otherwise.} \end{cases}$$

$$A_{ij} = A_{ji}$$

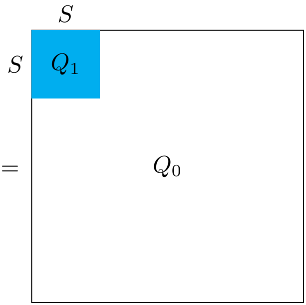


Problem: Given A, identify S

An Example

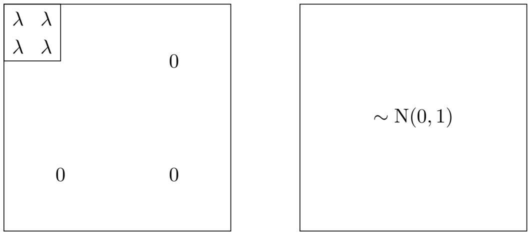
$$Q_1 = \mathrm{N}(\lambda, 1)$$
$$Q_0 = \mathrm{N}(0, 1).$$

$A =$



An Example

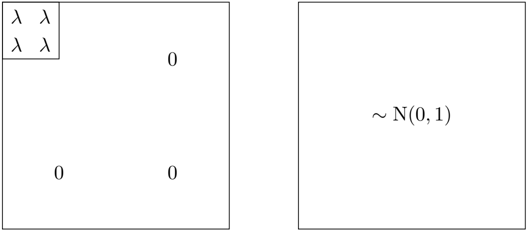
$$A = \qquad \lambda u_S u_S^{\mathsf{T}} \qquad + \qquad Z$$



Data = Sparse, Low Rank + noise

An Example

$$A = \qquad \lambda u_S u_S^{\mathsf{T}} \qquad + \qquad Z$$



Data = Sparse, Low Rank + noise

Much work in statistics

Denoising:

$$y = x + \text{noise}$$

[Donoho, Jin 2004], [Arias-Castro, Candes, Durand 2011]
[Arias-Castro, Bubeck, Lugosi 2012]

Sparse signal recovery:

$$y = Ax + \text{noise}$$

[Candes, Romberg, Tao 2004], [Chen, Donoho 1998]

The simple example combines different structures

Much work in statistics

Denoising:

$$y = x + \text{noise}$$

[Donoho, Jin 2004], [Arias-Castro, Candes, Durand 2011]
[Arias-Castro, Bubeck, Lugosi 2012]

Sparse signal recovery:

$$y = Ax + \text{noise}$$

[Candes, Romberg, Tao 2004], [Chen, Donoho 1998]

The simple example combines different structures

Much work in statistics

Denoising:

$$y = x + \text{noise}$$

[Donoho, Jin 2004], [Arias-Castro, Candes, Durand 2011]
[Arias-Castro, Bubeck, Lugosi 2012]

Sparse signal recovery:

$$y = Ax + \text{noise}$$

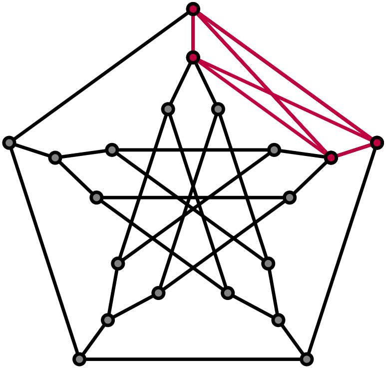
[Candes, Romberg, Tao 2004], [Chen, Donoho 1998]

The simple example combines different structures

Our Running Example: Planted Cliques

A is "adjacency" matrix

$$Q_1 = \delta_{+1}$$
$$Q_0 = \frac{1}{2}\delta_{+1} + \frac{1}{2}\delta_{-1}.$$



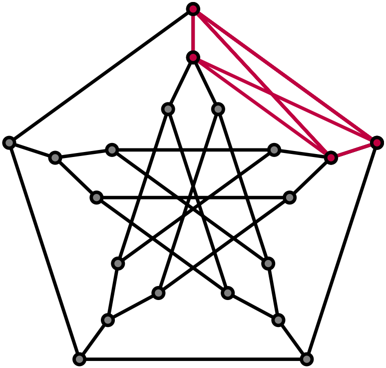
S forms a clique in the graph

[Jerrum, 1992]

Our Running Example: Planted Cliques

A is "adjacency" matrix

$$Q_1 = \delta_{+1}$$
$$Q_0 = \frac{1}{2}\delta_{+1} + \frac{1}{2}\delta_{-1}.$$

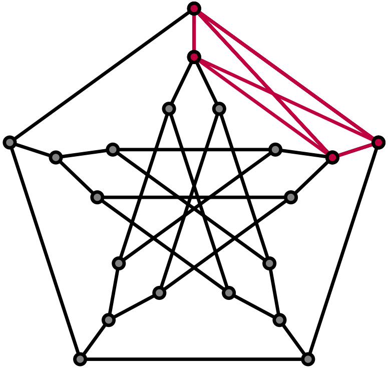


S forms a clique in the graph

[Jerrum, 1992]

Our Running Example: Planted Cliques

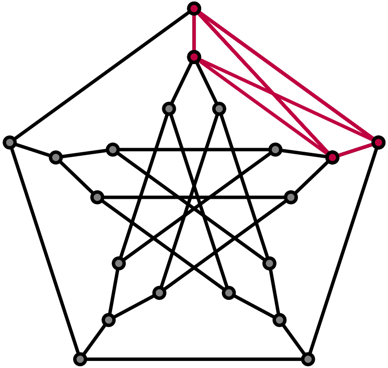
- Average case version of MAX-CLIQUE



- Communities in networks

Our Running Example: Planted Cliques

- Average case version of MAX-CLIQUE
- Communities in networks



How large should S be?

Let $|S| = k$

Size of largest clique in $G\left(n, \frac{1}{2}\right)$

Second moment calculation $\Rightarrow \quad k > 2\log_2 n$.

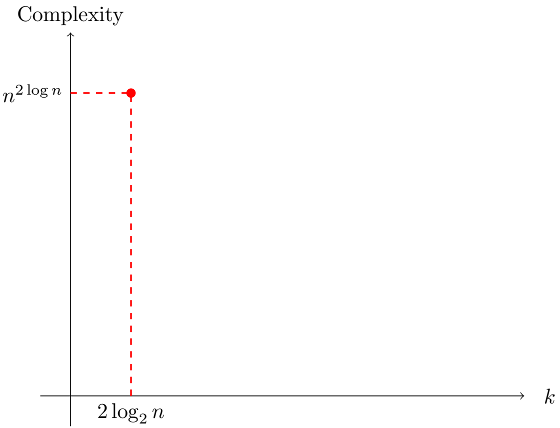
How large should S be?

Let $|S| = k$

Size of largest clique in $G\left(n, \frac{1}{2}\right)$

Second moment calculation $\Rightarrow \quad k > 2\log_2 n.$

Progress(0)



Exhaustive search

A Naive Algorithm

Pick k largest degree vertices of G as $\widehat{S}$.

A Naive Algorithm

Pick k largest degree vertices of G as $\widehat{S}$.

Analysis of NAIVE

If $i \notin S$:

$$\deg(i) = \text{Binomial}\left(n-1, \frac{1}{2}\right)$$

$$\Rightarrow \max_{i \notin S} \deg(i) \leq \frac{n}{2} + O(\sqrt{n \log n})$$

If $i \in S$:

$$\deg(i) = k - 1 + \text{Binomial}\left(n-k+1, \frac{1}{2}\right)$$

$$\Rightarrow \min_{i \in S} \deg(i) \geq \frac{k-1}{2} + \frac{n}{2} - O(\sqrt{n \log n})$$

NAIVE works if: $k \geq O(\sqrt{n \log n})$

Analysis of NAIVE

If $i \notin S$:

$$\deg(i) = \text{Binomial}\left(n-1, \frac{1}{2}\right)$$
$$\Rightarrow \max_{i \notin S} \deg(i) \le \frac{n}{2} + O(\sqrt{n \log n})$$

If $i \in S$:

$$\deg(i) = k - 1 + \text{Binomial}\left(n-k+1, \frac{1}{2}\right)$$
$$\Rightarrow \min_{i \in S} \deg(i) \ge \frac{k-1}{2} + \frac{n}{2} - O(\sqrt{n \log n})$$

NAIVE works if: $k \ge O(\sqrt{n \log n})$

Analysis of NAIVE

If $i \notin S$:

$$\deg(i) = \text{Binomial}\left(n-1, \frac{1}{2}\right)$$

$$\Rightarrow \max_{i \notin S} \deg(i) \leq \frac{n}{2} + O(\sqrt{n \log n})$$

If $i \in S$:

$$\deg(i) = k - 1 + \text{Binomial}\left(n-k+1, \frac{1}{2}\right)$$

$$\Rightarrow \min_{i \in S} \deg(i) \geq \frac{k-1}{2} + \frac{n}{2} - O(\sqrt{n \log n})$$

NAIVE works if: $k \geq O(\sqrt{n \log n})$

Analysis of NAIVE

If $i \notin S$:

$$\deg(i) = \text{Binomial}\left(n - 1, \frac{1}{2}\right)$$

$$\Rightarrow \max_{i \notin S} \deg(i) \le \frac{n}{2} + O(\sqrt{n \log n})$$

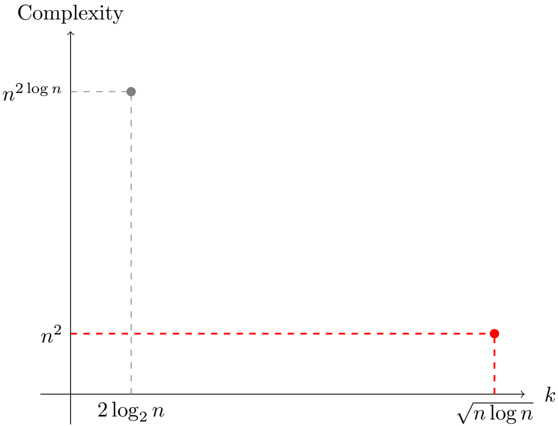
If $i \in S$:

$$\deg(i) = k - 1 + \text{Binomial}\left(n - k + 1, \frac{1}{2}\right)$$

$$\Rightarrow \min_{i \in S} \deg(i) \ge \frac{k-1}{2} + \frac{n}{2} - O(\sqrt{n \log n})$$

NAIVE works if: $k \ge O(\sqrt{n \log n})$

Progress(1)



[Kučera, 1995]

Spectral Method

$$A = \qquad u_S u_S^{\mathsf{T}} \qquad + \qquad Z$$

$$\left[\begin{array}{c|c} \boxed{\begin{matrix} 1 & 1 \\ 1 & 1 \end{matrix}} & 0 \\ 0 & 0 \end{array}\right] \qquad \left[\begin{array}{c|c} \boxed{0} & \pm 1 \\ \pm 1 & \pm 1 \end{array}\right]$$

Hopefully $v_1(A) \approx u_S$

Spectral Method

$$A = \underbrace{\begin{bmatrix} \begin{matrix} 1 & 1 \\ 1 & 1 \end{matrix} & 0 \\ 0 & 0 \end{bmatrix}}_{u_S u_S^{\mathsf{T}}} + \underbrace{\begin{bmatrix} 0 & \pm 1 \\ \pm 1 & \pm 1 \end{bmatrix}}_{Z}$$

Hopefully $v_1(A) \approx u_S$

Analysis of SPECTRAL

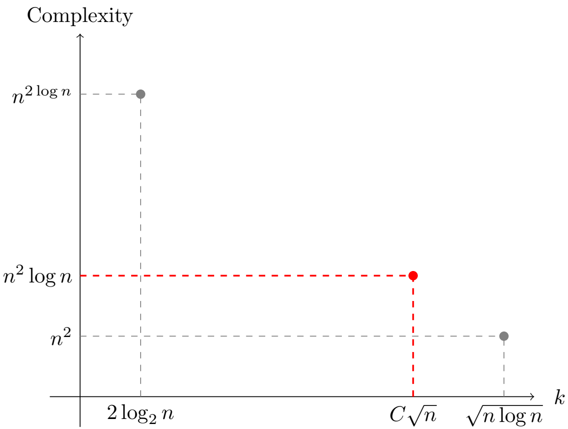
$$\frac{1}{\sqrt{n}} A = \underbrace{\left(\frac{k}{\sqrt{n}}\right)}_{\kappa} e_S e_S^{\mathsf{T}} + \frac{Z}{\sqrt{n}}.$$

By standard linear algebra:

$$\kappa \gg \left\| \frac{Z}{\sqrt{n}} \right\|_2 \approx 2 \Longrightarrow \langle v_1(A), e_S \rangle \geq 1 - \delta(\kappa)$$

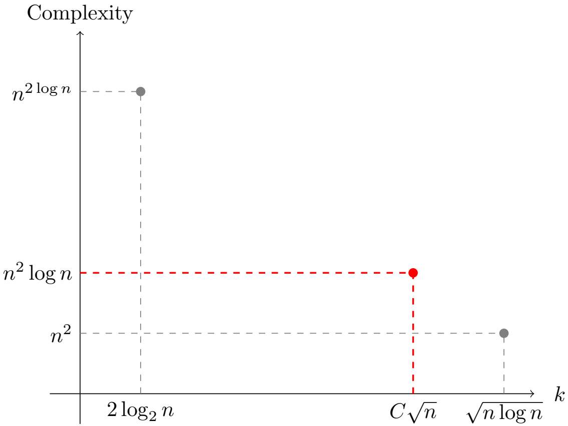
SPECTRAL works if $k \geq C\sqrt{n}$.

Progress(2)



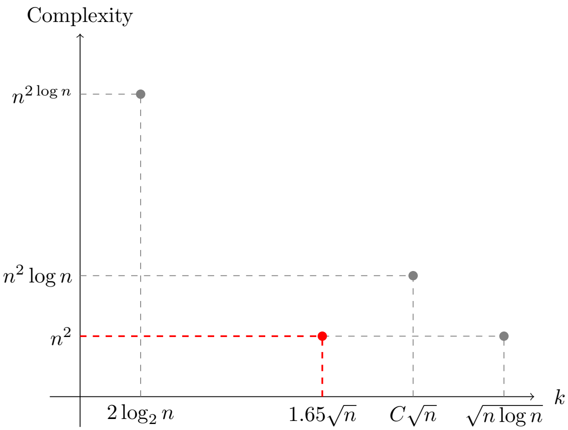
[Alon, Krivelevich and Sudakov, 1998]

Progress(2)



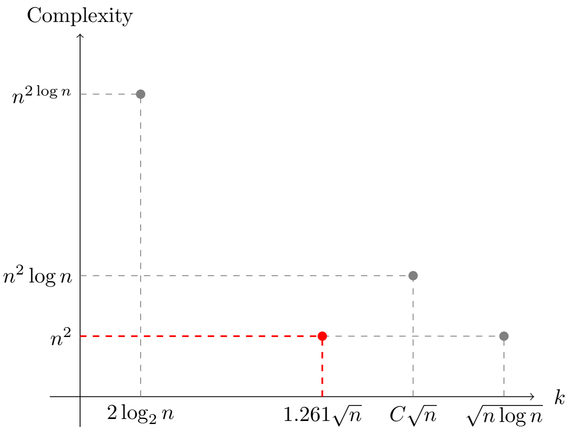
[Ames, Vavasis 2009], [Feige, Krauthgamer 2000]

Progress(3)



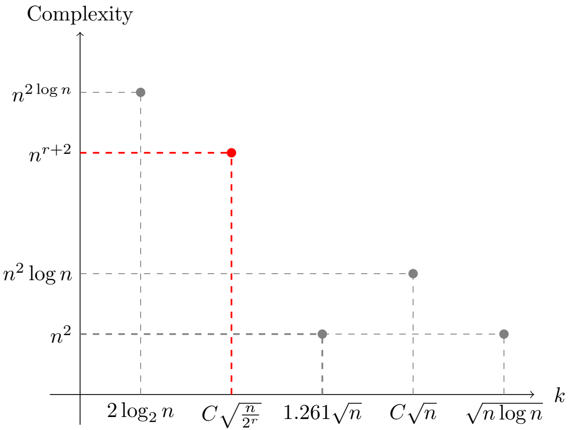
[Dekel, Gurel-Gurevich and Peres, 2010]

Progress(3)



[Dekel, Gurel-Gurevich and Peres, 2010]

Progress(3)

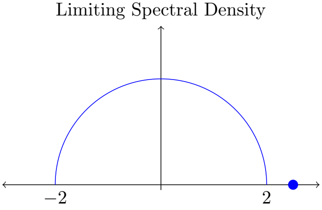


[Alon, Krivelevich and Sudakov, 1998]

Phase transition in spectrum of A

Phase transition in spectrum of A

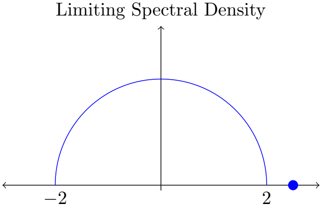
If $k \geq (1+\varepsilon)\sqrt{n}$



$$\langle v_1(A), e_S \rangle \geq \delta(\varepsilon)$$

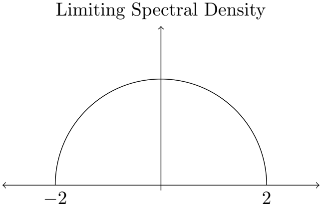
Phase transition in spectrum of A

If $k \geq (1+\varepsilon)\sqrt{n}$



$$\langle v_1(A), e_S \rangle \geq \delta(\varepsilon)$$

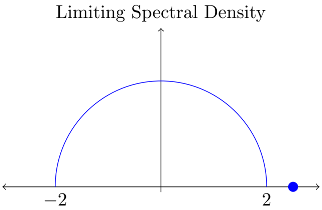
If $k \leq (1-\varepsilon)\sqrt{n}$



$$\langle v_1(A), e_S \rangle \leq n^{-1/2+\delta'(\varepsilon)}$$

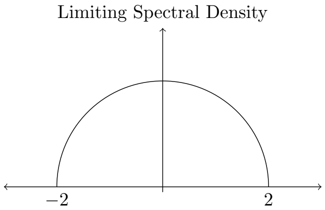
Phase transition in spectrum of A

If $k \geq (1+\varepsilon)\sqrt{n}$



$$\langle v_1(A), e_S \rangle \geq \delta(\varepsilon)$$

If $k \leq (1-\varepsilon)\sqrt{n}$



$$\langle v_1(A), e_S \rangle \leq n^{-1/2+\delta'(\varepsilon)}$$

[Knowles, Yin, 2011]

Seems much harder than it looks!

- "Statistical algorithms" fail if $k = n^{1/2-\delta}$: [Feldman et al., 2012]
- r-Lovász-Schrijver fails for $k \leq \sqrt{n/2^r}$: [Feige, Krauthgamer, 2002]
- r-Lasserre fails for $k \leq \frac{\sqrt{n}}{(\log n)^{r^2}}$: [Widgerson and Meka, 2013]

Seems much harder than it looks!

- "Statistical algorithms" fail if $k = n^{1/2-\delta}$: [Feldman et al., 2012]
- r-Lovász-Schrijver fails for $k \leq \sqrt{n/2^r}$: [Feige, Krauthgamer, 2002]
- r-Lasserre fails for $k \leq \frac{\sqrt{n}}{(\log n)^{r^2}}$: [Widgerson and Meka, 2013]

Seems much harder than it looks!

- "Statistical algorithms" fail if $k = n^{1/2-\delta}$: [Feldman et al., 2012]
- r-Lovász-Schrijver fails for $k \leq \sqrt{n/2^r}$: [Feige, Krauthgamer, 2002]
- r-Lasserre fails for $k \leq \frac{\sqrt{n}}{(\log n)^{r^2}}$: [Widgerson and Meka, 2013]

Our result

Theorem (Deshpande, Montanari, 2013)

If $|S| = k \geq (1+\varepsilon)\sqrt{n/e}$, there exists an $O(n^2 \log n)$ time algorithm that identifies S with high probability.

I will present:

1. A (wrong) heuristic analysis
2. How to fix the heuristic
3. Lower bounds

Our result

Theorem (Deshpande, Montanari, 2013)

If $|S| = k \geq (1+\varepsilon)\sqrt{n/e}$, there exists an $O(n^2 \log n)$ time algorithm that identifies S with high probability.

I will present:

1. A (wrong) heuristic analysis
2. How to fix the heuristic
3. Lower bounds

Our result

Theorem (Deshpande, Montanari, 2013)

If $|S| = k \geq (1+\varepsilon)\sqrt{n/e}$*, there exists an* $O(n^2 \log n)$ *time algorithm that identifies* S *with high probability.*

I will present:

1. A (wrong) heuristic analysis
2. How to fix the heuristic
3. Lower bounds

Iterative Thresholding

The power iteration:

$$v^{t+1} = A\ v^t.$$

Improvement:

$$v^{t+1} = AF_t(v^t).$$

where $F_t(v) = (f_t(v_1), f_t(v_2), \cdots, f_t(v_n))^{\mathsf{T}}$

Choose $f_t(\cdot)$ to exploit sparsity of e_S

Iterative Thresholding

The power iteration:

$$v^{t+1} = A \; v^t.$$

Improvement:

$$v^{t+1} = A F_t(v^t).$$

where $F_t(v) = (f_t(v_1), f_t(v_2), \cdots, f_t(v_n))^{\mathsf{T}}$

Choose $f_t(\cdot)$ to exploit sparsity of e_S

Iterative Thresholding

The power iteration:

$$v^{t+1} = A\ v^t.$$

Improvement:

$$v^{t+1} = AF_t(v^t).$$

where $F_t(v) = (f_t(v_1), f_t(v_2), \cdots, f_t(v_n))^{\mathsf{T}}$

Choose $f_t(\cdot)$ to exploit sparsity of e_S

Iterative Thresholding

The power iteration:

$$v^{t+1} = A \ v^t.$$

Improvement:

$$v^{t+1} = A F_t(v^t).$$

where $F_t(v) = (f_t(v_1), f_t(v_2), \cdots, f_t(v_n))^{\mathsf{T}}$

Choose $f_t(\cdot)$ to exploit sparsity of e_S

Iterative Thresholding

The power iteration:

$$v^{t+1} = A \; v^t.$$

Improvement:

$$v^{t+1} = AF_t(v^t).$$

where $F_t(v) = (f_t(v_1), f_t(v_2), \cdots, f_t(v_n))^{\mathsf{T}}$

Choose $f_t(\cdot)$ to exploit sparsity of e_S

(Wrong) Analysis

$$v_i^{t+1} = \frac{1}{\sqrt{n}} \sum_j A_{ij} f_t(v_j^t).$$

A_{ij} are random ± 1 r.v.

Use Central Limit Theorem for v_i^{t+1}

(Wrong) Analysis

$$v_i^{t+1} = \frac{1}{\sqrt{n}} \sum_j A_{ij} f_t(v_j^t).$$

A_{ij} are random ± 1 r.v.

Use Central Limit Theorem for v_i^{t+1}

(Wrong) Analysis

If $i \notin S$:

$$v_i^{t+1} = \frac{1}{\sqrt{n}} \sum_j A_{ij} f_t(v_j^t)$$

$$\approx \mathrm{N}\left(0, \frac{1}{n} \sum_j f_t(v_j^t)^2\right)$$

Letting $v_i^t \approx \mathrm{N}(0, \sigma_t^2)$...

$$\sigma_{t+1}^2 = \frac{1}{n} \sum_j f_t(v_j^t)^2$$

$$= \mathbb{E}\{f_t(\sigma_t \xi)^2\},$$

$$\xi \sim \mathrm{N}(0,1)$$

(Wrong) Analysis

If $i \notin S$:

$$v_i^{t+1} = \frac{1}{\sqrt{n}} \sum_j A_{ij} f_t(v_j^t)$$
$$\approx \mathrm{N}\Big(0, \frac{1}{n} \sum_j f_t(v_j^t)^2\Big)$$

Letting $v_i^t \approx \mathrm{N}(0, \sigma_t^2)$...

$$\sigma_{t+1}^2 = \frac{1}{n} \sum_j f_t(v_j^t)^2$$
$$= \mathbb{E}\{f_t(\sigma_t \xi)^2\},$$
$$\xi \sim \mathrm{N}(0,1)$$

(Wrong) Analysis

If $i \notin S$:

$$v_i^{t+1} = \frac{1}{\sqrt{n}} \sum_j A_{ij} f_t(v_j^t)$$
$$\approx \mathrm{N}\left(0, \frac{1}{n} \sum_j f_t(v_j^t)^2\right)$$

Letting $v_i^t \approx \mathrm{N}(0, \sigma_t^2)\dots$

$$\sigma_{t+1}^2 = \frac{1}{n} \sum_j f_t(v_j^t)^2$$
$$= \mathbb{E}\{f_t(\sigma_t \xi)^2\},$$
$$\xi \sim \mathrm{N}(0,1)$$

(Wrong) Analysis

If $i \notin S$:

$$v_i^{t+1} = \frac{1}{\sqrt{n}} \sum_j A_{ij} f_t(v_j^t)$$
$$\approx \mathrm{N}\Big(0, \frac{1}{n} \sum_j f_t(v_j^t)^2\Big)$$

Letting $v_i^t \approx \mathrm{N}(0, \sigma_t^2)$...

$$\sigma_{t+1}^2 = \frac{1}{n} \sum_j f_t(v_j^t)^2$$
$$= \mathbb{E}\{f_t(\sigma_t \xi)^2\},$$
$$\xi \sim \mathrm{N}(0,1)$$

(Wrong) Analysis

If $i \in S$:

$$v_i^{t+1} = \underbrace{\frac{1}{\sqrt{n}} \sum_{j \in S} f_t(v_j^t)}_{\mu_{t+1}} + \underbrace{\frac{1}{\sqrt{n}} \sum_{j \notin S} A_{ij} f_t(v_j^t)}_{\approx \mathrm{N}(0, \sigma_{t+1}^2)}$$

where

$$\begin{aligned} \mu_{t+1} &= \frac{1}{\sqrt{n}} \sum_{j \in S} f_t(v_j^t) \\ &= \left(\frac{k}{\sqrt{n}}\right) \left(\frac{1}{k} \sum_{j \in S} f_t(v_j^t)\right) \\ &= \kappa \, \mathbb{E}\{f_t(\mu_t + \sigma_t \xi)\}, \\ \xi &\sim \mathrm{N}(0,1) \end{aligned}$$

(Wrong) Analysis

If $i \in S$:

$$v_i^{t+1} = \underbrace{\frac{1}{\sqrt{n}} \sum_{j \in S} f_t(v_j^t)}_{\mu_{t+1}} + \underbrace{\frac{1}{\sqrt{n}} \sum_{j \notin S} A_{ij} f_t(v_j^t)}_{\approx \mathrm{N}(0, \sigma_{t+1}^2)}$$

where

$$\mu_{t+1} = \frac{1}{\sqrt{n}} \sum_{j \in S} f_t(v_j^t)$$

$$= \left(\frac{k}{\sqrt{n}}\right)\left(\frac{1}{k}\sum_{j \in S} f_t(v_j^t)\right)$$

$$= \kappa\, \mathbb{E}\{f_t(\mu_t + \sigma_t \xi)\},$$

$$\xi \sim \mathrm{N}(0,1)$$

(Wrong) Analysis

If $i \in S$:

$$v_i^{t+1} = \underbrace{\frac{1}{\sqrt{n}} \sum_{j \in S} f_t(v_j^t)}_{\mu_{t+1}} + \underbrace{\frac{1}{\sqrt{n}} \sum_{j \notin S} A_{ij} f_t(v_j^t)}_{\approx \mathrm{N}(0, \sigma_{t+1}^2)}$$

where

$$\begin{aligned} \mu_{t+1} &= \frac{1}{\sqrt{n}} \sum_{j \in S} f_t(v_j^t) \\ &= \left(\frac{k}{\sqrt{n}}\right) \left(\frac{1}{k} \sum_{j \in S} f_t(v_j^t)\right) \\ &= \kappa\, \mathbb{E}\{f_t(\mu_t + \sigma_t \xi)\}, \end{aligned}$$

$$\xi \sim \mathrm{N}(0,1)$$

(Wrong) Analysis

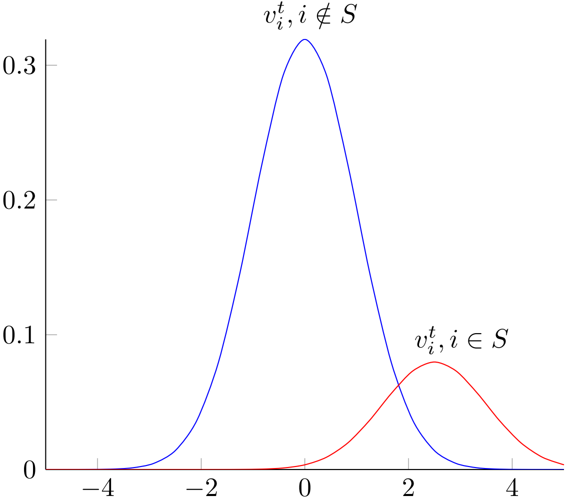
If $i \in S$:

$$v_i^{t+1} = \underbrace{\frac{1}{\sqrt{n}} \sum_{j \in S} f_t(v_j^t)}_{\mu_{t+1}} + \underbrace{\frac{1}{\sqrt{n}} \sum_{j \notin S} A_{ij} f_t(v_j^t)}_{\approx \mathrm{N}(0, \sigma_{t+1}^2)}$$

where

$$\begin{aligned} \mu_{t+1} &= \frac{1}{\sqrt{n}} \sum_{j \in S} f_t(v_j^t) \\ &= \left(\frac{k}{\sqrt{n}}\right) \left(\frac{1}{k} \sum_{j \in S} f_t(v_j^t)\right) \\ &= \kappa\, \mathbb{E}\{f_t(\mu_t + \sigma_t \xi)\}, \\ \xi &\sim \mathrm{N}(0,1) \end{aligned}$$

Summarizing . . .



State Evolution

$$\mu_{t+1} = \kappa\,\mathbb{E}\left\{f_t(\mu_t + \sigma_t \xi)\right\}$$
$$\sigma_{t+1}^2 = \mathbb{E}\{f_t(\sigma_t \xi)^2\}.$$

Using the optimal function $f_t(x) = e^{\mu_t x - \mu_t^2}$

$$\mu_{t+1} = \kappa e^{\mu_t^2/2}$$
$$\sigma_{t+1}^2 = 1$$

State Evolution

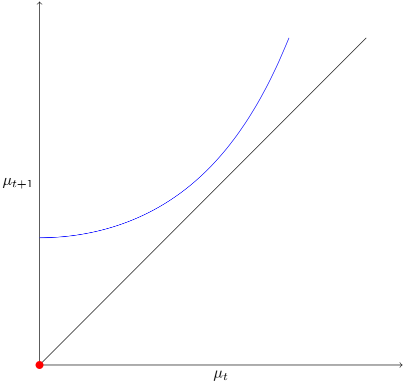
$$\mu_{t+1} = \kappa\, \mathbb{E}\left\{f_t(\mu_t + \sigma_t \xi)\right\}$$
$$\sigma_{t+1}^2 = \mathbb{E}\{f_t(\sigma_t \xi)^2\}.$$

Using the optimal function $f_t(x) = e^{\mu_t x - \mu_t^2}$

$$\mu_{t+1} = \kappa e^{\mu_t^2/2}$$
$$\sigma_{t+1}^2 = 1$$

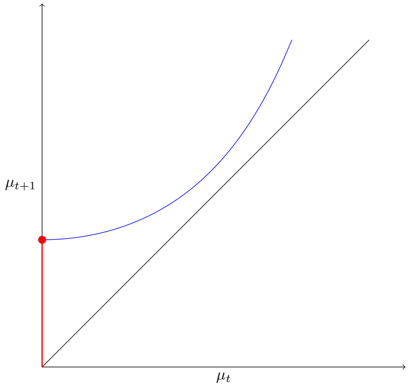
Fixed points develop below threshold!

If $\kappa > \frac{1}{\sqrt{e}}$



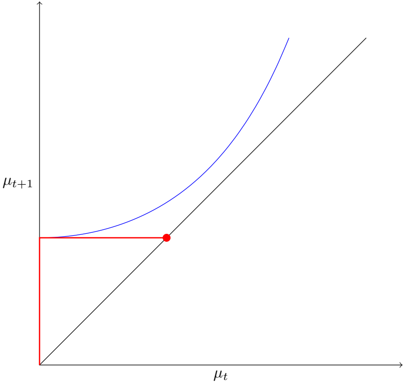
Fixed points develop below threshold!

If $\kappa > \frac{1}{\sqrt{e}}$



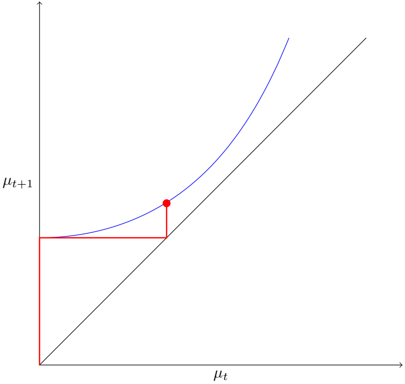
Fixed points develop below threshold!

If $\kappa > \frac{1}{\sqrt{e}}$



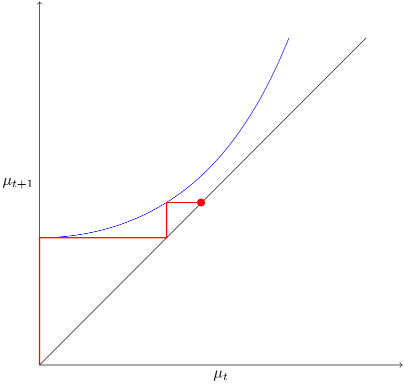
Fixed points develop below threshold!

If $\kappa > \frac{1}{\sqrt{e}}$



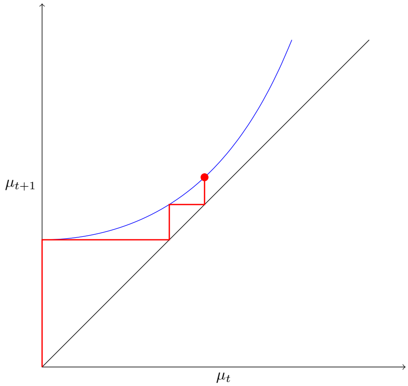
Fixed points develop below threshold!

If $\kappa > \frac{1}{\sqrt{e}}$



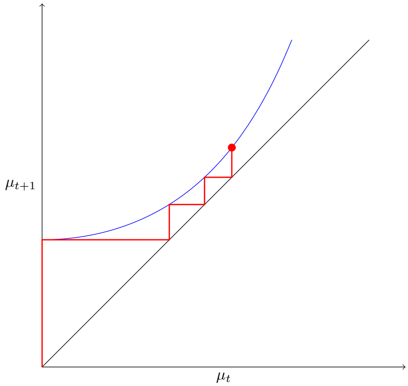
Fixed points develop below threshold!

If $\kappa > \frac{1}{\sqrt{e}}$



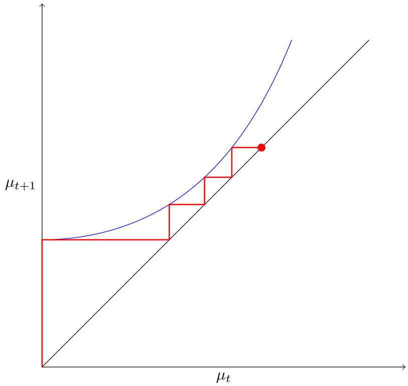
Fixed points develop below threshold!

If $\kappa > \frac{1}{\sqrt{e}}$



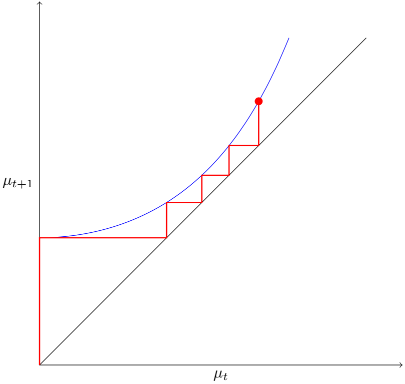
Fixed points develop below threshold!

If $\kappa > \frac{1}{\sqrt{e}}$



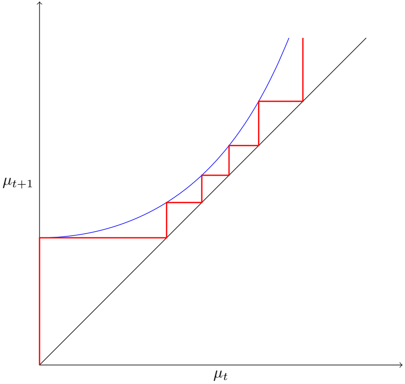
Fixed points develop below threshold!

If $\kappa > \frac{1}{\sqrt{e}}$



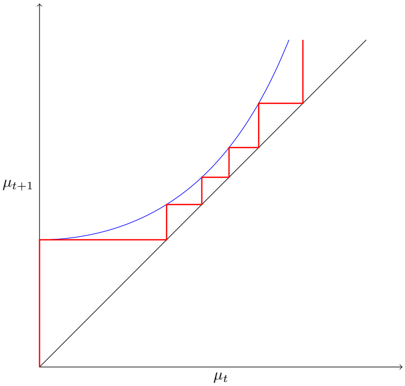
Fixed points develop below threshold!

If $\kappa > \frac{1}{\sqrt{e}}$

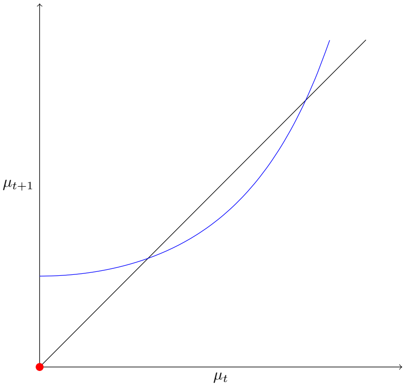


Fixed points develop below threshold!

If $\kappa > \frac{1}{\sqrt{e}}$

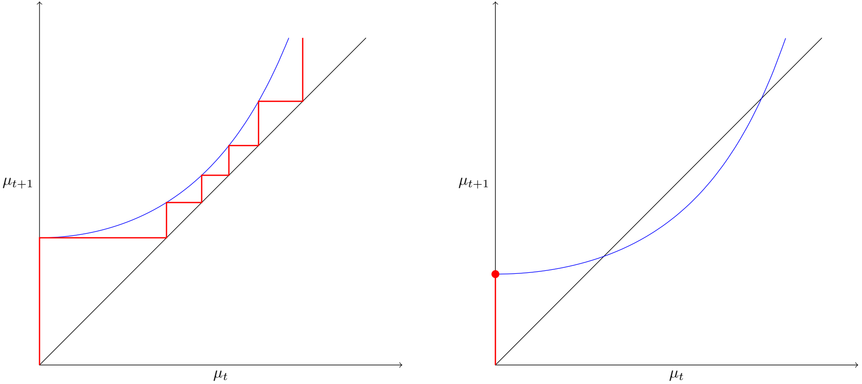


If $\kappa < \frac{1}{\sqrt{e}}$



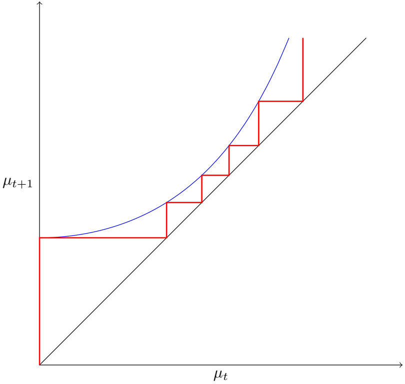
Fixed points develop below threshold!

If $\kappa > \frac{1}{\sqrt{e}}$ | If $\kappa < \frac{1}{\sqrt{e}}$

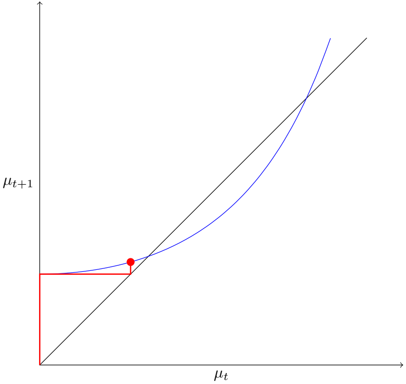


Fixed points develop below threshold!

If $\kappa > \frac{1}{\sqrt{e}}$

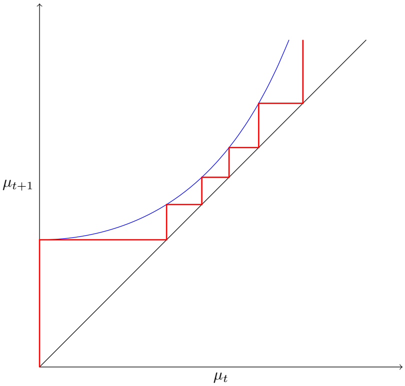


If $\kappa < \frac{1}{\sqrt{e}}$

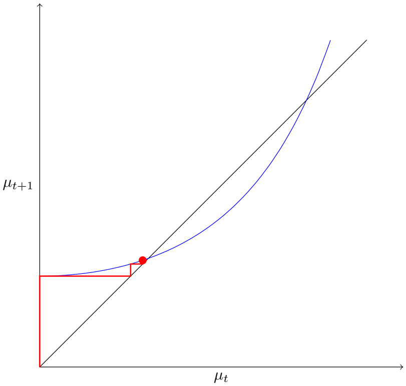


Fixed points develop below threshold!

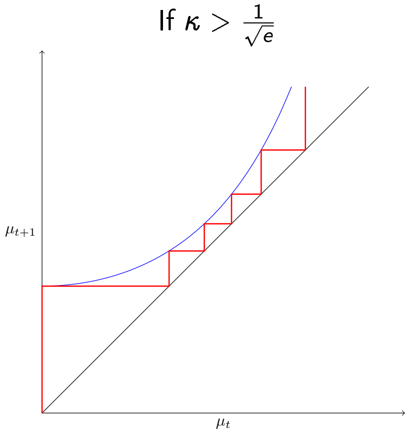
If $\kappa > \frac{1}{\sqrt{e}}$



If $\kappa < \frac{1}{\sqrt{e}}$



Fixed points develop below threshold!



Analysis is wrong but...

Theorem (Deshpande, Montanari, 2013)

If $|S| = k \geq (1+\varepsilon)\sqrt{n/e}$, there exists an $O(n^2 \log n)$ time algorithm that identifies S with high probability.

...so we modify the algorithm.

What algorithm?

Slight modification to iterative scheme:

$$(v_i^t)_{i\in[n]} \to (v_{i\to j}^t)_{i,j\in[n]}$$

$$v_{i\to j}^{t+1} = \frac{1}{\sqrt{n}} \sum_{\ell\neq i,j} A_{i\ell}\, f_t(v_{\ell\to i}^t).$$

Analysis is *exact* as $n \to \infty$.

What algorithm?

Slight modification to iterative scheme:

$$(v_i^t)_{i\in[n]} \to (v_{i\to j}^t)_{i,j\in[n]}$$

$$v_{i\to j}^{t+1} = \frac{1}{\sqrt{n}} \sum_{\ell\neq i,j} A_{i\ell}\, f_t(v_{\ell\to i}^t).$$

Analysis is *exact* as $n \to \infty$.

What algorithm?

Slight modification to iterative scheme:

$$(v_i^t)_{i \in [n]} \to (v_{i\to j}^t)_{i,j \in [n]}$$

$$v_{i\to j}^{t+1} = \frac{1}{\sqrt{n}} \sum_{\ell \neq i,j} A_{i\ell}\, f_t(v_{\ell \to i}^t).$$

Analysis is *exact* as $n \to \infty$.

What algorithm?

Slight modification to iterative scheme:

$$(v_i^t)_{i\in[n]} \to (v_{i\to j}^t)_{i,j\in[n]}$$

$$v_{i\to j}^{t+1} = \frac{1}{\sqrt{n}} \sum_{\ell \neq i,j} A_{i\ell}\, f_t(v_{\ell\to i}^t).$$

Analysis is *exact* as $n \to \infty$.

Fixing the heuristic

Lemma

Let $(f_t(z))_{t\geq 0}$ be a sequence of polynomials. Then, for every fixed t, and bounded, continuous function $\psi : \mathbb{R} \to \mathbb{R}$ the following limit holds in probability:

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{i\in S} \psi(v^t_{i\to j}) = \kappa\, \mathbb{E}\{\psi(\mu_t + \sigma_t \xi)\},$$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i\in [n]\setminus S} \psi(v^t_{i\to j}) = \mathbb{E}\{\psi(\sigma_t \xi)\},$$

where $\xi \sim \mathrm{N}(0,1)$.

Proof Technique

Key ideas:

Expand $v_{i\to j}^{t}$ for polynomial $f_t(\cdot)$

Wrong analysis works if $A \to A^t$

Proof Technique

Key ideas:

Expand $v^t_{i \to j}$ for polynomial $f_t(\cdot)$

Wrong analysis works if $A \to A^t$

Proof Technique

Key ideas:

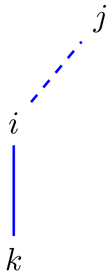
Expand $v_{i \to j}^t$ for polynomial $f_t(\cdot)$

Wrong analysis works if $A \to A^t$

Proof Technique - Expanding v^t

Let $f_t(x) = x^2, v^0_{i\to j} = 1$

$$v^1_{i\to j} = \sum_{k\neq j} A_{ik}.$$

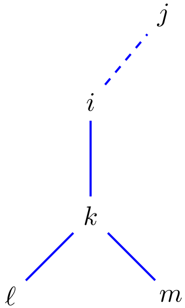


Proof Technique - Expanding v^t

$$v^2_{i\to j} = \sum_{k\neq j} A_{ik}(v^1_{k\to i})^2$$

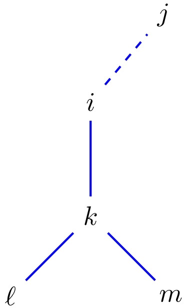
$$= \sum_{k\neq j} A_{ik}\Big(\sum_{\ell\neq i} A_{k\ell}\Big)\Big(\sum_{m\neq i} A_{km}\Big)$$

$$= \sum_{k\neq j}\sum_{\ell\neq i}\sum_{m\neq i} A_{ik}A_{k\ell}A_{km}.$$



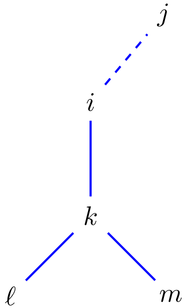
Proof Technique - Expanding v^t

$$
\begin{aligned}
v^2_{i\to j} &= \sum_{k\neq j} A_{ik}(v^1_{k\to i})^2 \\
&= \sum_{k\neq j} A_{ik}\Big(\sum_{\ell\neq i} A_{k\ell}\Big)\Big(\sum_{m\neq i} A_{km}\Big) \\
&= \sum_{k\neq j}\sum_{\ell\neq i}\sum_{m\neq i} A_{ik}A_{k\ell}A_{km}.
\end{aligned}
$$



Proof Technique - Expanding v^t

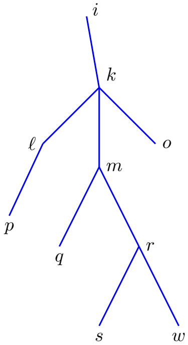
$$
\begin{aligned}
v^2_{i\to j} &= \sum_{k\neq j} A_{ik} (v^1_{k\to i})^2 \\
&= \sum_{k\neq j} A_{ik} \Big(\sum_{\ell \neq i} A_{k\ell} \Big) \Big(\sum_{m\neq i} A_{km} \Big) \\
&= \sum_{k\neq j} \sum_{\ell\neq i} \sum_{m\neq i} A_{ik} A_{k\ell} A_{km}.
\end{aligned}
$$

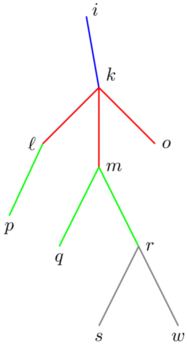


Proof Technique

$$v_{i\to j}^{t+1} = \sum_{k\neq i} A_{ik} f_t(v_{k\to i}^t).$$

$$\xi_{i\to j}^{t+1} = \sum_{k\neq i} A_{ik}^t f_t(\xi_{k\to i}^t).$$





Proof Technique - a Combinatorial Lemma

Lemma

$$v^t_{i\to j} = \sum_{T \in \mathcal{T}^t_{i\to j}} A(T)\Gamma(T)v^0(T)$$

where $\mathcal{T}^t_{i\to j}$ consists rooted, labeled trees that:

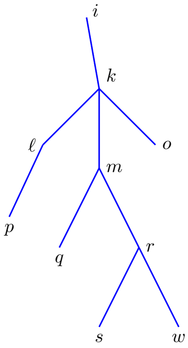
1. *have maximum depth t.*
2. *do not backtrack.*

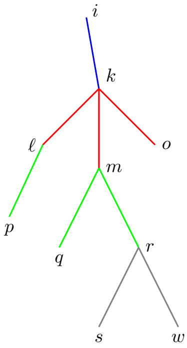
(Similarly for the $\xi^t_{i\to j}$)

Proof Technique - Moment Method

$$v_{i\to j}^{t+1} = \sum_{k\neq i} A_{ik} f_t(v_{k\to i}^t).$$

$$\xi_{i\to j}^{t+1} = \sum_{k\neq i} A_{ik}^t f_t(\xi_{k\to i}^t).$$



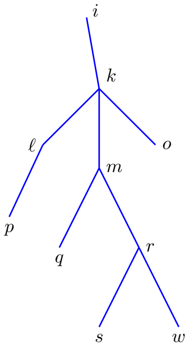


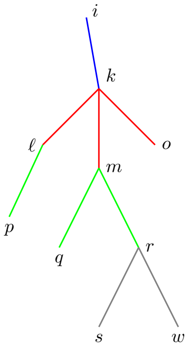
$\lim_{n\to\infty}$ Moments of v^{t+1} = Moments of ξ^{t+1}.

Proof Technique - Moment Method

$$v_{i\to j}^{t+1} = \sum_{k\neq i} A_{ik} f_t(v_{k\to i}^t).$$

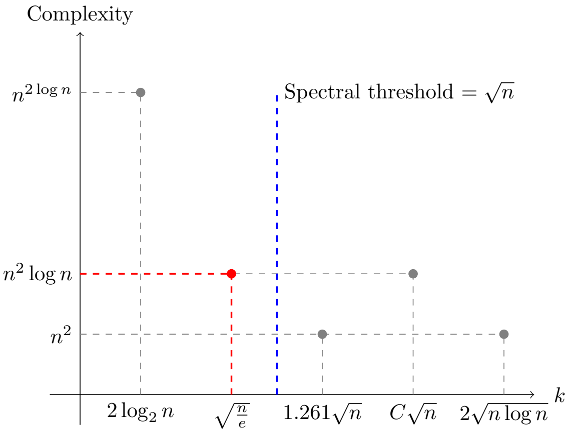
$$\xi_{i\to j}^{t+1} = \sum_{k\neq i} A_{ik}^t f_t(\xi_{k\to i}^t).$$





$\lim_{n\to\infty}$ Moments of v^{t+1} = Moments of ξ^{t+1}.

Progress(4)



Complexity
$n^{2\log n}$
$n^2 \log n$
n^2
Spectral threshold $= \sqrt{n}$
$2\log_2 n$
$\sqrt{\frac{n}{e}}$
$1.261\sqrt{n}$
$C\sqrt{n}$
$2\sqrt{n \log n}$
k

Is this threshold fundamental?

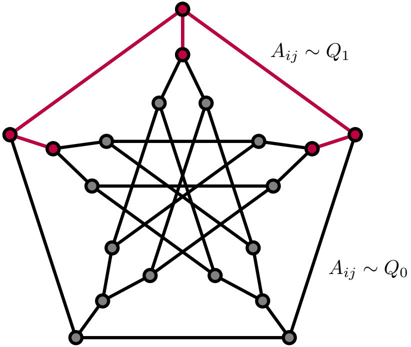
Rest of the talk: perhaps

The "Hidden Set" Problem

Given $G_n = ([n], E_n)$

A Set $\rightarrow S \subset [n]$

Data $\rightarrow A_{ij} \sim \begin{cases} Q_1 & \text{if } i,j \in S, \\ Q_0 & \text{otherwise.} \end{cases}$



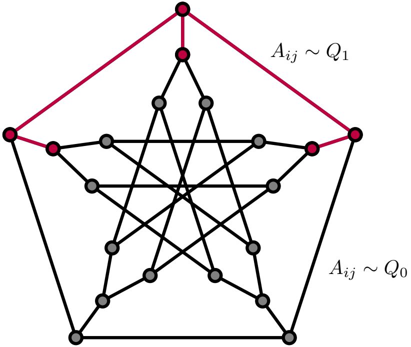
Problem: Given edge labels $(A_{ij})_{(i,j) \in E_n}$, identify S

The "Hidden Set" Problem

Given $G_n = ([n], E_n)$

A Set $\rightarrow S \subset [n]$

Data $\rightarrow A_{ij} \sim \begin{cases} Q_1 & \text{if } i,j \in S, \\ Q_0 & \text{otherwise.} \end{cases}$



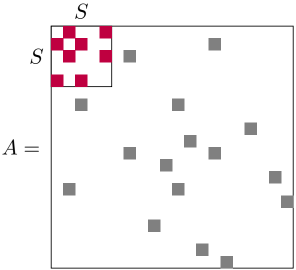
Problem: Given edge labels $(A_{ij})_{(i,j) \in E_n}$, identify S

The "Hidden Set" Problem

Given $G_n = ([n], E_n)$

A Set $\to S \subset [n]$

Data $\to A_{ij} \sim \begin{cases} Q_1 & \text{if } i,j \in S, \\ Q_0 & \text{otherwise.} \end{cases}$



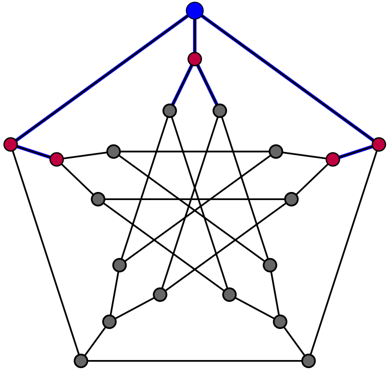
Problem: Given edge labels $(A_{ij})_{(i,j)\in E_n}$, identify S

"Local" Algorithms

A t-local algorithm computes:

Estimate at i:

$$\widehat{u}(i) = \mathsf{F}(A_{\mathrm{Ball}(i,t)})$$



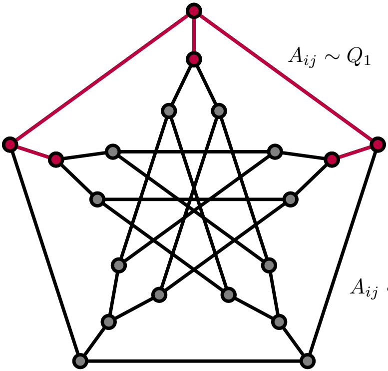
The Sparse Graph Analogue

$G_n = ([n], E_n)$, $n \geq 1$ satisfies:

- locally tree-like
- regular degree Δ

Further

- $Q_1 = \delta_{+1}$, $Q_0 = \frac{1}{2}\delta_{+1} + \frac{1}{2}\delta_{-1}$



What can local algorithms achieve?

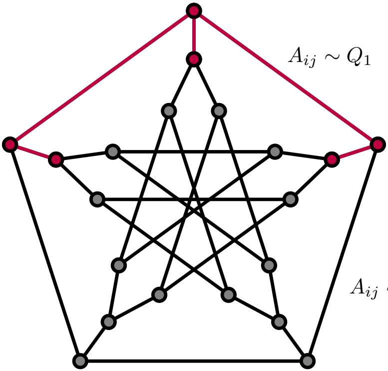
The Sparse Graph Analogue

$G_n = ([n], E_n), n \geq 1$ satisfies:

- locally tree-like
- regular degree Δ

Further

- $Q_1 = \delta_{+1}, \; Q_0 = \frac{1}{2}\delta_{+1} + \frac{1}{2}\delta_{-1}$



What can local algorithms achieve?

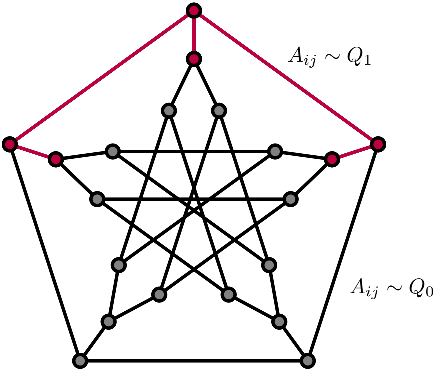
The Sparse Graph Analogue

$G_n = ([n], E_n), n \geq 1$ satisfies:

- locally tree-like
- regular degree Δ

Further

- $Q_1 = \delta_{+1}, \; Q_0 = \frac{1}{2}\delta_{+1} + \frac{1}{2}\delta_{-1}$



What can local algorithms achieve?

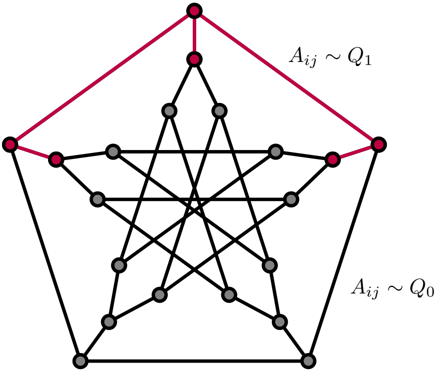
The Sparse Graph Analogue

$G_n = ([n], E_n)$, $n \geq 1$ satisfies:

- locally tree-like
- regular degree Δ

Further

- $Q_1 = \delta_{+1}$, $Q_0 = \frac{1}{2}\delta_{+1} + \frac{1}{2}\delta_{-1}$



What can local algorithms achieve?

What can we hope for?

If $|S| = C\frac{n}{\sqrt{\Delta}}$:

$\widehat{S}_{\sf naive}$ = Random set of size $|S|$

$$\Rightarrow \frac{1}{n}\mathbb{E}\{\widehat{S}_{\sf naive}\triangle S\} = \Theta\left(\frac{1}{\sqrt{\Delta}}\right).$$

What can we hope for?

If $|S| = C\frac{n}{\sqrt{\Delta}}$:

Poisson bound $\Rightarrow$ for any local algorithm:

$$\frac{1}{n}\mathbb{E}\{\widehat{S}\triangle S\} \geq e^{-C'\sqrt{\Delta}}.$$

A result for local algorithms. . .

Theorem (Deshpande, Montanari, 2013)

Let G_n converge locally to $\Delta-$regular tree:
If $|S| \geq (1+\varepsilon)\frac{n}{\sqrt{e\Delta}}$ there exists a local algorithm achieving

$$\frac{1}{n}\mathbb{E}\{S \triangle \widehat{S}\} \leq e^{-\Theta(\sqrt{\Delta})}.$$

Conversely, if $|S| \leq (1-\varepsilon)\frac{n}{\sqrt{e\Delta}}$ every local algorithm suffers

$$\frac{1}{n}\mathbb{E}\{S \triangle \widehat{S}\} \geq \Theta\left(\frac{1}{\sqrt{\Delta}}\right)$$

With $\Delta = n-1$ we recover the complete graph result!

A result for local algorithms...

Theorem (Deshpande, Montanari, 2013)

Let G_n converge locally to $\Delta-$regular tree:
If $|S| \geq (1+\varepsilon)\frac{n}{\sqrt{e\Delta}}$ there exists a local algorithm achieving

$$\frac{1}{n}\mathbb{E}\{S \triangle \widehat{S}\} \leq e^{-\Theta(\sqrt{\Delta})}.$$

Conversely, if $|S| \leq (1-\varepsilon)\frac{n}{\sqrt{e\Delta}}$ every local algorithm suffers

$$\frac{1}{n}\mathbb{E}\{S \triangle \widehat{S}\} \geq \Theta\left(\frac{1}{\sqrt{\Delta}}\right)$$

With $\Delta = n - 1$ we recover the complete graph result!

A result for local algorithms...

Theorem (Deshpande, Montanari, 2013)

Let G_n converge locally to $\Delta-$regular tree:
If $|S| \geq (1+\varepsilon)\frac{n}{\sqrt{e\Delta}}$ there exists a local algorithm achieving

$$\frac{1}{n}\mathbb{E}\{S \triangle \widehat{S}\} \leq e^{-\Theta(\sqrt{\Delta})}.$$

Conversely, if $|S| \leq (1-\varepsilon)\frac{n}{\sqrt{e\Delta}}$ every local algorithm suffers

$$\frac{1}{n}\mathbb{E}\{S \triangle \widehat{S}\} \geq \Theta\left(\frac{1}{\sqrt{\Delta}}\right)$$

With $\Delta = n-1$ we recover the complete graph result!

To conclude. . .

- Message-passing algorithm performs "weighted" counts of non-reversing trees

- Such structures have been used elsewhere:
  - Clustering sparse networks: [Krzakala et al. 2013]
  - Compressed sensing: [Bayati, Lelarge, Montanari 2013]

To conclude. . .

- Message-passing algorithm performs "weighted" counts of non-reversing trees

- Such structures have been used elsewhere:
    - Clustering sparse networks: [Krzakala et al. 2013]
    - Compressed sensing: [Bayati, Lelarge, Montanari 2013]

To conclude. . .

- Message-passing algorithm performs "weighted" counts of non-reversing trees

- Such structures have been used elsewhere:
    - Clustering sparse networks: [Krzakala et al. 2013]
    - Compressed sensing: [Bayati, Lelarge, Montanari 2013]

To conclude. . .

- What is the "dense" analogue for local algorithms?

- What about other structural properties?

Thank you!

To conclude. . .

- What is the “dense” analogue for local algorithms?
- What about other structural properties?

Thank you!

To conclude. . .

- What is the “dense” analogue for local algorithms?

- What about other structural properties?

Thank you!