Community detection with the non-backtracking operator

Marc Lelarge [1]
Charles Bordenave[2] Laurent Massoulié[3]

[1]INRIA-ENS

[2]CNRS Université de Toulouse

[3]INRIA-Microsoft Research Joint Centre

Simons Institute, Berkeley, May 2016

Adjacency and non-backtracking matrices

Adjacency matrix

Take a finite, simple, non-oriented graph $G = (V, E)$.

Adjacency matrix : symmetric, indexed on vertices, for $u, v \in V$,

$$A_{uv} = \mathbf{1}(\{u, v\} \in E).$$

Perron eigenvalue

If $|V| = n$, the (real) eigenvalues of A are

$$\mu_1 \geqslant \mu_2 \geqslant \ldots \geqslant \mu_n$$

From Perron-Frobenius Theorem : if G is connected, then

$$\mu_1 > \mu_2 \quad \text{and} \quad \mu_1 \geqslant -\mu_n.$$

Moreover, $\mu_1 = -\mu_n$ is equivalent to G bipartite.

Regular graphs

Assume $\deg(v) = d$ for all $v \in V$.

Then

$$\mu_1 = d.$$

Spectral gap

Largest non-trivial eigenvalue

$$\mu = \max\{|\mu_k| : |\mu_k| \neq d\}.$$

Theorem (Alon-Boppana (1991))

$$\mu \geqslant 2\sqrt{d-1} - \frac{c_d}{\log n}.$$

Ramanujan graphs

A d-regular graph is Ramanujan if

$$\mu \leqslant 2\sqrt{d-1}$$

Existence of infinite sequence of Ramanujan graphs

- $d = p^k + 1$, p prime : *Lubotzky, Phillips & Sarnak (1988)*, *Margulis (1988)*, *Morgenstern (1994)*,
- any $d \geqslant 3$: *Marcus, Spielman, Srivastava (2013)*.

Spectral gap and Diameter

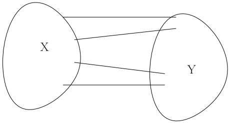
Recall $\mu = \max\{|\mu_k| : |\mu_k| \neq d\}$.

Theorem (Chung (1989))

$$\mathrm{diam}(G) \leqslant \frac{\log(n-1)}{\log d - \log \mu} + 2.$$

For $X, Y \subset V$, define

$$E(X,Y) = \sum_{x \in X, y \in Y} \mathbf{1}(\{u,v\} \in E).$$



Isoperimetric constant :

$$h(G) = \min_{X \subset V} \frac{E(X, X^c)}{\min(|X|, |X^c|)}.$$

Theorem (Cheeger's Inequality)

$$\frac{h(G)^2}{2d} \leqslant d - \mu_2 \leqslant 2h(G).$$

Random regular graph

Theorem (Friedman (2004))

Fix integer $d \geqslant 3$. Let G_n is a sequence of uniformly distributed d-regular graphs on n vertices, then with high probability,

$$\mu(G_n) = 2\sqrt{d-1} + o(1).$$

Most regular graphs are nearly Ramanujan !!

Non-regular graphs

It is not straightforward to extend the previous notions to non-regular graphs. *Lubotzky (1995), Hoory (2005)*.

Eigenvectors of extremal eigenvalues tend to localize on large degree vertices.

Non-regular graphs

It is not straightforward to extend the previous notions to non-regular graphs. *Lubotzky (1995), Hoory (2005)*.

Eigenvectors of extremal eigenvalues tend to localize on large degree vertices.

For example, if G is an Erdős-Rényi graph with parameter α/n, for any fixed $k \geqslant 1$, with high probability,

$$\mu_k \sim \sqrt{\max_{v \in V}{}^{[k]} \deg(v)} \sim \sqrt{\frac{\log n}{\log \log n}},$$

Sudakov & Krivelevich (2003).

Hashimoto's non-backtracking matrix

Oriented edge set :

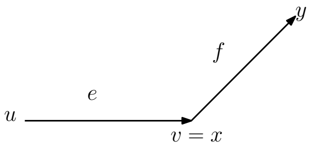
$$\vec{E} = \{(u, v) : \{u, v\} \in E\},$$

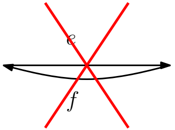
hence, $m = |\vec{E}| = 2|E|$.

If $e = uv, f = xy$ are in $\vec{E}$,

$$B_{ef} = \mathbf{1}(v = x)\mathbf{1}(u \neq y),$$

defines a $|\vec{E}| \times |\vec{E}|$ matrix on the oriented edges.





Perron eigenvalue

Eigenvalues, $m = 2|E|$,

$$\lambda_1 \geqslant |\lambda_2| \geqslant \cdots \geqslant |\lambda_m|.$$

A non-backtracking path $p = (v_1 \dots v_n)$ is a path such that $v_{i-1} \neq v_{i+1}$. If $e = uv$,

$$\|B^\ell \delta_e\|_1 = \text{nb of NB paths starting with } vu \text{ of length } \ell + 1.$$

Perron eigenvalue

Eigenvalues, $m = 2|E|$,

$$\lambda_1 \geqslant |\lambda_2| \geqslant \cdots \geqslant |\lambda_m|.$$

A non-backtracking path $p = (v_1 \dots v_n)$ is a path such that $v_{i-1} \neq v_{i+1}$. If $e = uv$,

$$\|B^\ell \delta_e\|_1 = \text{nb of NB paths starting with } vu \text{ of length } \ell + 1.$$

If G is 2-connected (any vertex or pair of vertices is part of a cycle) then B is irreducible and

$$\lambda_1 = \lim_{\ell \to \infty} \|B^\ell \delta_e\|_1^{1/\ell} = \text{growth rate of the universal cover of } B.$$

Ihara-Bass' Identity

Let Q the diagonal matrix with $Q_{vv} = \deg(v) - 1$. We have

$$\det(z - B) \quad = \quad (z^2 - 1)^{|E|-|V|} \det(z^2 - Az + Q)$$

Ihara-Bass' Identity

Let Q the diagonal matrix with $Q_{vv} = \deg(v) - 1$. We have

$$\det(z - B) \quad = \quad (z^2 - 1)^{|E|-|V|} \det(z^2 - Az + Q)$$

If G is d-regular, then $Q = (d-1)I$ and

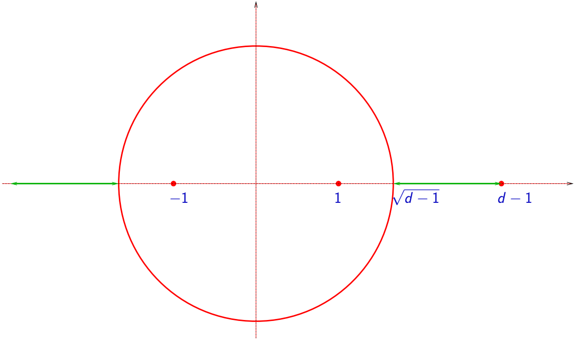
$$\sigma(B) = \{\pm 1\} \cup \left\{ \lambda : \lambda^2 - \lambda\mu + (d-1) = 0 \text{ with } \mu \in \sigma(A) \right\}.$$

Angel, Friedman, Hoory (2007), Terras (2011), ...

Non-Backtracking matrix of regular graphs

For a d-regular graph, $\lambda_1 = d - 1$,

- ⋆ Alon-Boppana bound : $\max_{k \neq 1} \mathfrak{Re}(\lambda_k) \geqslant \sqrt{\lambda_1} - o(1)$.
- ⋆ Ramanujan (non bipartite) : $|\lambda_2| = \sqrt{\lambda_1}$
- ⋆ Friedman's thm : $|\lambda_2| \leqslant \sqrt{\lambda_1} + o(1)$ if G random uniform.



Non-Backtracking matrix of arbitrary graph

"In general graphs, the condition $|\lambda_2| \leqslant \sqrt{\lambda_1}$ *is one of the possible analogs of a Ramanujan property".*
Stark & Terras

BUT, due to the non-normality of B,

- ⋆ No Alon-Boppana lower bound.
- ⋆ No Cheeger-type isoperimetric inequality.
- ⋆ No Chung-type diameter inequality.

Only weak versions of these bounds in terms of the singular values of B^k are available.

Oriented path symmetry

If $\check{x}(e) = x(e^{-1})$ and $Px = \check{x}$, then $P^2 = I$ and we have

$$BP = PB^*.$$

In other words, $B^k P$ is symmetric.

This type of symmetry is refered in physics as PT invariance.

Oriented path symmetry

If $\check{x}(e) = x(e^{-1})$ and $Px = \check{x}$, then $P^2 = I$ and we have

$$BP = PB^*.$$

In other words, $B^k P$ is symmetric.

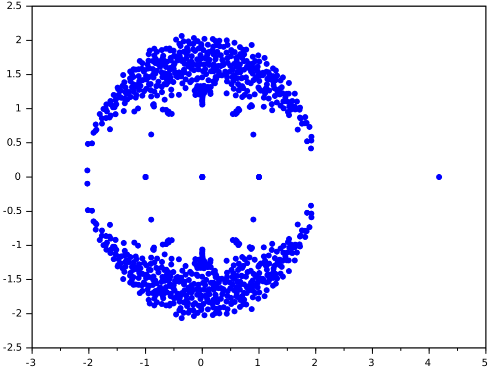
This type of symmetry is refered in physics as PT invariance.

From the spectral decomposition of $B^k P$, it is possible to get some analogs of Alon-Boppana, Chung or Cheeger type inequalities...

Non-backtracking spectrum of random graphs

Simulation for Erdős-Rényi Graph

Eigenvalues of B for an Erdős-Rényi graph $\mathcal{G}(n, \alpha/n)$ with $n = 500$ and $\alpha = 4$.



Erdős-Rényi Graph

$$\lambda_1 \geqslant |\lambda_2| \geqslant \ldots .$$

Theorem

Let $\alpha > 1$ *and* G *with distribution* $\mathcal{G}(n, \alpha/n)$. *With high probability,*

$$\begin{aligned} \lambda_1 &= \alpha + o(1) \\ |\lambda_2| &\leqslant \sqrt{\alpha} + o(1). \end{aligned}$$

Some ideas of proofs

Perron eigenvalue

We zoom and consider the matrix B^{ℓ} where for some well chosen $0 < \kappa < 1/2$,

$$\ell \sim \kappa \log_{\alpha} n.$$

In the Erdős-Rényi case, the graph spanned by vertices at distance ℓ from a fixed vertex is close in total variation to a Galton-Watson tree with $\mathrm{Poi}(\alpha)$ distribution

We will study the singular value decomposition of B^{ℓ}.

Perron eigenvalue

Assume that we can find unit vectors ζ, ψ and $s > 0$ such that

$$B^\ell = s\zeta\psi^* + C,$$

with $\|C\| = o(s/\ell)$ and $\langle \zeta, \psi \rangle \geqslant \kappa > 0$. Then, from Bauer-Fike Theorem,

$$|\lambda_1 - s^{1/\ell}| = O\left(\frac{s^{1/\ell}}{\ell}\right) \quad \text{and} \quad |\lambda_2| = O\Big(\|C\|^{1/\ell}\Big).$$

Perron eigenvalue

Assume that we can find unit vectors ζ, ψ and $s > 0$ such that

$$B^\ell = s\zeta\psi^* + C,$$

with $\|C\| = o(s/\ell)$ and $\langle \zeta, \psi \rangle \geqslant \kappa > 0$. Then, from Bauer-Fike Theorem,

$$|\lambda_1 - s^{1/\ell}| = O\left(\frac{s^{1/\ell}}{\ell}\right) \quad \text{and} \quad |\lambda_2| = O\Big(\|C\|^{1/\ell}\Big).$$

We will take for a well chosen ψ,

$$s\zeta = B^\ell \psi.$$

Then

$$\|C\| = \sup_{x : \langle x, \psi \rangle = 0} \frac{\|B^\ell x\|_2}{\|x\|_2}.$$

Perron eigenvalue

If $e = uv \in \vec{E}$ and $\chi(f) = 1$ for all $f \in \vec{E}$,

$\langle \delta_e, B^\ell \chi \rangle =$ nb of NB paths of length $\ell + 1$ starting with uv in G

is close to the population Z_ℓ at generation ℓ in a Galton-Watson process with $\mathrm{Poi}(\alpha)$ distribution.

Perron eigenvalue

Seneta-Heyde thm, conditionned on non-extinction, a.s.

$$\frac{Z_\ell}{\alpha^\ell} \to M \in (0, \infty).$$

Hence, conditionned on non-extinction, a.s.

$$\frac{Z_{2\ell}}{\alpha^\ell Z_\ell} \to 1.$$

Perron eigenvalue

Seneta-Heyde thm, conditionned on non-extinction, a.s.

$$\frac{Z_\ell}{\alpha^\ell} \to M \in (0, \infty).$$

Hence, conditionned on non-extinction, a.s.

$$\frac{Z_{2\ell}}{\alpha^\ell Z_\ell} \to 1.$$

The vector

$$\varphi = \frac{B^\ell \chi}{\|B^\ell \chi\|_2}$$

should be close to an eigenvector of B^ℓ associated to α^ℓ.

Seneta-Heyde thm, conditionned on non-extinction, a.s.

$$\frac{Z_\ell}{\alpha^\ell} \to M \in (0, \infty).$$

Hence, conditionned on non-extinction, a.s.

$$\frac{Z_{2\ell}}{\alpha^\ell Z_\ell} \to 1.$$

The vector

$$\varphi = \frac{B^\ell \chi}{\|B^\ell \chi\|_2}$$

should be close to an eigenvector of B^ℓ associated to α^ℓ.

But we have not found a way to upper bound $\|B^\ell x\|_2$ *for* $x \in \varphi^\perp$.

Strategy of proof

Recall $\check{x}(e) = x(e^{-1})$, we set

$$\zeta = \frac{B^\ell \check{\varphi}}{\|B^\ell \check{\varphi}\|_2} = \frac{B^\ell B^{*\ell} \chi}{\|B^\ell B^{*\ell} \chi\|_2} \quad \text{and} \quad s = \|B^\ell \check{\varphi}\|_2.$$

$$B^\ell = B^\ell \check{\varphi} \check{\varphi}^* + C = s\,\zeta \check{\varphi}^* + C.$$

Our claim: $\lambda_1 = \alpha + o(1)$ and $|\lambda_2| \leqslant \sqrt{\alpha} + o(1)$ is implied by

Strategy of proof

Recall $\check{x}(e) = x(e^{-1})$, we set

$$\zeta = \frac{B^\ell \check{\varphi}}{\|B^\ell \check{\varphi}\|_2} = \frac{B^\ell B^{*\ell}\chi}{\|B^\ell B^{*\ell}\chi\|_2} \quad \text{and} \quad s = \|B^\ell \check{\varphi}\|_2.$$

$$B^\ell = B^\ell \check{\varphi}\check{\varphi}^* + C = s\,\zeta\check{\varphi}^* + C.$$

Our claim: $\lambda_1 = \alpha + o(1)$ and $|\lambda_2| \leqslant \sqrt{\alpha} + o(1)$ is implied by

Proposition (Near eigenvector)

With high probability,

$$\langle \zeta, \check{\varphi}\rangle > c_0 \quad \textit{and} \quad c_0\alpha^\ell < s < c_1\alpha^\ell.$$

Proposition (Small norm in the complement)

With high probability,

$$\sup_{x:\langle x, \check{\varphi}\rangle = 0} \|B^\ell x\|_2 \leqslant (\log n)^c \alpha^{\ell/2}\|x\|_2.$$

Proposition (Near eigenvector)
With high probability,

$$\langle \zeta, \check{\varphi} \rangle > c_0 \quad \textit{and} \quad c_0 \alpha^{\ell} < s < c_1 \alpha^{\ell}.$$

It requires to prove convergence of expressions of the form

$$\alpha^{-2\ell} \langle \delta_e, B^{2\ell} B^{*\ell} \chi \rangle$$

toward a limit random variable.

Small norm in the complement

Proposition (Small norm in the complement)

With high probability,

$$\sup_{x : \langle x, \check{\varphi} \rangle = 0} \|B^\ell x\|_2 \leqslant (\log n)^c \alpha^{\ell/2} \|x\|_2 .$$

Small norm in the complement

Proposition (Small norm in the complement)
With high probability,

$$\sup_{x:\langle x,\check{\varphi}\rangle=0} \|B^{\ell}x\|_2 \leqslant (\log n)^c \alpha^{\ell/2}\|x\|_2.$$

Standard caveat : the graph contains a clique of size m with probability larger than $n^{-m^2/2}$,

$$\mathbb{E}(B^{\ell})_{ee} \geqslant (m-1)^{\ell} n^{-m^2/2} = e^{(\kappa \log(m-1) - m^2/2)\log n}.$$

Polynomially small event may have a big influence in expectation.

Small norm in the complement

With high probability, the graph is ℓ-tangle free that is : no vertex has more than two distinct cycles in its ℓ neighborhood.

Small norm in the complement

With high probability, the graph is ℓ-tangle free that is : no vertex has more than two distinct cycles in its ℓ neighborhood.

We may replace B^{ℓ} by

$$
\begin{aligned}
(B^{(\ell)})_{ef} &= \text{nb of NB tangle free paths } \gamma \text{ of length } \ell \text{ from } e \text{ to } f \\
&= \sum_{\gamma} \prod_{s=0}^{\ell} A_{\gamma_s, \gamma_{s+1}},
\end{aligned}
$$

where the sum is over NB tangle free paths of length ℓ from e to f in the complete graph.

Friedman (2004), Neeman-Sly-Mossel (2013), ...

Small norm in the complement

Consider the centered matrix

$$\Delta_{ef}^{(\ell)} = \sum_{\gamma} \prod_{s=0}^{\ell} \left(A_{\gamma_s,\gamma_{s+1}} - \frac{\alpha}{n} \right),$$

where the sum is over NB tangle free paths of length ℓ from e to f.

Consider the centered matrix

$$\Delta_{ef}^{(\ell)} = \sum_{\gamma} \prod_{s=0}^{\ell} \left(A_{\gamma_s,\gamma_{s+1}} - \frac{\alpha}{n} \right),$$

where the sum is over NB tangle free paths of length ℓ from e to f.

After a tricky decomposition,

$$\|B^{\ell} x\|_2 \leqslant \|\Delta^{(\ell)}\| + \frac{\alpha}{n} \sum_{t=1}^{\ell-1} \|\Delta^{(t-1)} \chi\|_2 \left| \langle (B^*)^{\ell-t-1} \chi, x \rangle \right| + \dots$$

which we should estimate over $\langle \check{\varphi}, x \rangle = \langle (B^*)^{\ell} \chi, x \rangle = 0$.

Massoulié (2013)

Small norm in the complement

$$\|B^{\ell}x\|_2 \leqslant \|\Delta^{(\ell)}\| + \frac{\alpha}{n}\sum_{t=1}^{\ell-1}\|\Delta^{(t-1)}\chi\|_2\left|\langle (B^*)^{\ell-t-1}\chi, x\rangle\right| + \cdots$$

From the Galton-Watson tree coupling
$\langle (B^*)^{\ell}\chi, \delta_e\rangle \simeq \alpha^{\ell-t}\langle (B^*)^{t}\chi, \delta_e\rangle$,

$$\max_{\langle (B^*)^{\ell}\chi, x\rangle = 0}\left|\langle (B^*)^{t}\chi, x\rangle\right| \leqslant (\log n)^c\sqrt{n}\,\alpha^{t/2}\|x\|_2.$$

Small norm in the complement

$$\|B^\ell x\|_2 \leqslant \|\Delta^{(\ell)}\| + \frac{\alpha}{n}\sum_{t=1}^{\ell-1}\|\Delta^{(t-1)}\chi\|_2\left|\langle (B^*)^{\ell-t-1}\chi, x\rangle\right| + \cdots$$

From the Galton-Watson tree coupling
$\langle (B^*)^\ell\chi, \delta_e\rangle \simeq \alpha^{\ell-t}\langle (B^*)^t\chi, \delta_e\rangle$,

$$\max_{\langle (B^*)^\ell\chi, x\rangle = 0}\left|\langle (B^*)^t\chi, x\rangle\right| \leqslant (\log n)^c\sqrt{n}\,\alpha^{t/2}\|x\|_2.$$

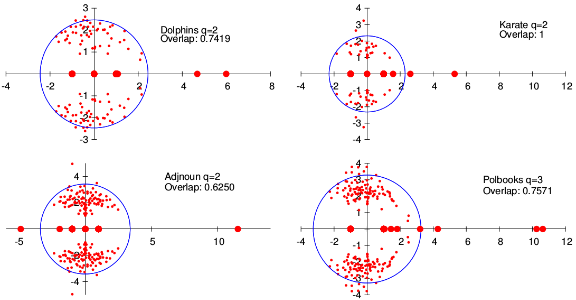
By the method of moments, with $m \simeq \log n/\log\log n$,

$$\|\Delta^{(t)}\| \leqslant \left(\mathrm{Tr}\left(\Delta^{(t)}\Delta^{(t)*}\right)^m\right)^{1/m} \leqslant (\log n)^c\alpha^{t/2}$$

$$\|\Delta^{(t)}\chi\|_2 \leqslant (\log n)^c\sqrt{n}\,\alpha^{t/2}.$$

Community detection

"Eigenvalues/eigenvectors such that $|\lambda_k| > \sqrt{\lambda_1}$ *should contain relevant global information on the graph".*



Krzakala/Moore/Mossel/Neeman/Sly/Zdeborová/Zhang (2013)

Stochastic Block Model

Consider a set of types $\{1, \cdots, r\}$ and assign type $\sigma_n(v)$ to vertex v. We assume that

$$\pi_n(i) = \frac{1}{n} \sum_{v=1}^{n} \mathbf{1}(\sigma_n(v) = i) = \pi(i) + O(n^{-\gamma}),$$

for some probability vector π.

Stochastic Block Model

Consider a set of types $\{1, \cdots, r\}$ and assign type $\sigma_n(v)$ to vertex v. We assume that

$$\pi_n(i) = \frac{1}{n} \sum_{v=1}^{n} \mathbf{1}(\sigma_n(v) = i) = \pi(i) + O(n^{-\gamma}),$$

for some probability vector π.

If $\sigma(u) = i, \sigma(v) = j$, the edge $\{u, v\}$ is present independently with probability

$$\frac{W_{ij}}{n} \wedge 1,$$

where W is a symmetric matrix.

(Inhomogeneous random graph, Chung-Lu random graph, ...)

Stochastic Block Model

If $\sigma(v) = j$, mean number of type i neighbors is

$$\pi(i)W_{ij} + O(1/n).$$

Mean progeny matrix

$$M = \mathrm{diag}(\pi)W.$$

Stochastic Block Model

If $\sigma(v) = j$, mean number of type i neighbors is

$$\pi(i) W_{ij} + O(1/n).$$

Mean progeny matrix

$$M = \mathrm{diag}(\pi) W.$$

We assume that the average degree is homogeneous, for all $1 \leqslant j \leqslant r$,

$$\sum_{i=1}^{r} M_{ij} = \alpha > 1.$$

Assume that M is strongly irreducible and we order its real eigenvalues

$$\alpha = \mu_1 > |\mu_2| \geqslant \cdots \geqslant |\mu_r|.$$

Stochastic Block Model

Model used in community detection. Notably for $r = 2$,

$$\pi = \left(\frac{1}{2}, \frac{1}{2}\right)$$

and, with $a > b$,

$$W = \begin{pmatrix} a & b \\ b & a \end{pmatrix}.$$

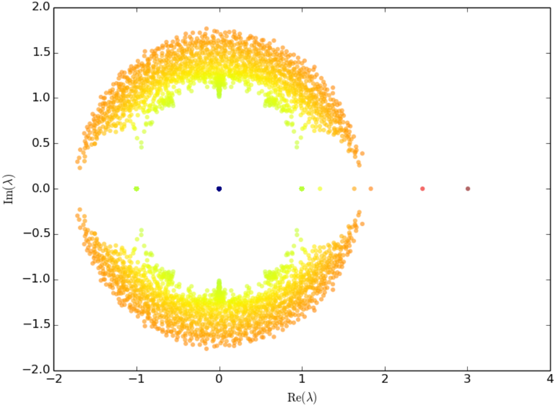
Then

$$\mu_1 = \frac{a+b}{2} \quad \text{and} \quad \mu_2 = \frac{a-b}{2}.$$

Decelle, Krzakala, Moore, Mossel, Neeman, Sly, Zdeborová, Zhang

Stochastic Block Model

Eigenvalues of B for a Stochastic Block Model with $n = 2000$, mean degree $\alpha = \frac{a+b}{2} = 3$ and $\frac{a-b}{2} = 2.45$



Stochastic Block Model

Let $1 \leqslant r_0 \leqslant r$ such that

$$\alpha = \mu_1 > |\mu_2| \geqslant \cdots \geqslant |\mu_{r_0}| > \sqrt{\mu}_1 \geqslant |\mu_{r_0+1}| \geqslant \cdots \geqslant |\mu_r|.$$

Theorem

Let $\alpha > 1$ and G a stochastic block model as above. With high probability, up to reordering the eigenvalues of B,

$$\begin{array}{rcll} \lambda_k & = & \mu_k + o(1) & \textit{if } 1 \leqslant k \leqslant r_0 \\ |\lambda_k| & \leqslant & \sqrt{\alpha} + o(1) & \textit{if } k > r_0. \end{array}$$

Stochastic Block Model

Let $1 \leqslant r_0 \leqslant r$ such that

$$\alpha = \mu_1 > |\mu_2| \geqslant \cdots \geqslant |\mu_{r_0}| > \sqrt{\mu}_1 \geqslant |\mu_{r_0+1}| \geqslant \cdots \geqslant |\mu_r|.$$

Theorem

Let $\alpha > 1$ and G a stochastic block model as above. With high probability, up to reordering the eigenvalues of B,

$$\begin{aligned} \lambda_k &= \mu_k + o(1) && \text{if } 1 \leqslant k \leqslant r_0 \\ |\lambda_k| &\leqslant \sqrt{\alpha} + o(1) && \text{if } k > r_0. \end{aligned}$$

(+ a description of the eigenvectors of λ_k, $1 \leqslant k \leqslant r_0$, if the μ_k are distinct, In particular, they are asymptotically orthogonal).

Stochastic Block Model

Assume

$$\pi = \left(\frac{1}{2}, \frac{1}{2}\right) \quad \text{and} \quad W = \begin{pmatrix} a & b \\ b & a \end{pmatrix}.$$

If $(a-b)^2 > 2(a+b)$, with high probability, we may reconstruct correctly a proportion larger than $1/2 + \varepsilon$ of the types from the second largest eigenvector of B.

If $(a-b)^2 < 2(a+b)$, no algorithm can perform better than random guess *(Neeman, Mossel & Sly (2012))*.

Non-symmetric Stochastic Block Model

Assume

$$\pi = (p, 1-p) \quad \text{and} \quad W = \begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

$$\alpha = pa + (1-p)b = (1-p)c + pb$$

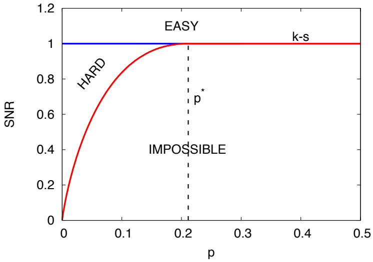
If $\left(\frac{p}{1-p}\right)^2 \frac{(a-\alpha)^2}{\alpha} > 1$, with high probability, we may reconstruct a positively correlated partition from the second largest eigenvector of B.

However if $\left(\frac{p}{1-p}\right)^2 \frac{(a-\alpha)^2}{\alpha} < 1$, exhaustive search might still work *(Neeman & Netrapalli (2014))*.

Non-symmetric Stochastic Block Model

Consider the case where

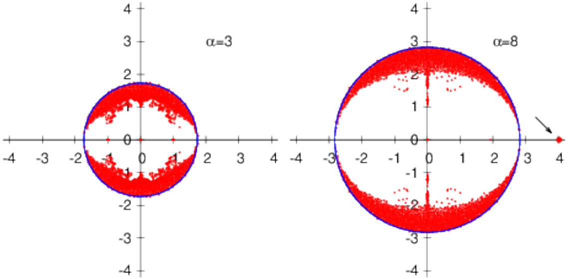
$$\alpha \to \infty \quad \text{while} \quad p \quad \text{and} \quad SNR = \left(\frac{p}{1-p}\right)^2 \frac{(a-\alpha)^2}{\alpha} \text{ are fixed.}$$



Kesten-Stigum bound obtained by vanilla spectral algorithm
Benaych-Georges, Couillet, Lelarge (2016)
$p^* = \frac{1}{2} - \frac{1}{2\sqrt{3}}$, ongoing work with *Caltagirone & Miolane*

Some extensions

For the labeled stochastic block model, we also conjecture a phase transition. We have partial results and an optimal spectral algorithm.



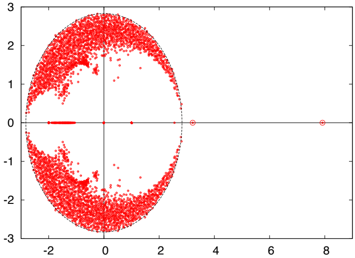
Saade, Krzakala, Lelarge, Zdeborovà, (2015,2016)

Some extensions

The non-backtracking matrix is also working for the degree-corrected SBM.
ongoing work with Gulikers and Massoulié.

Some extensions

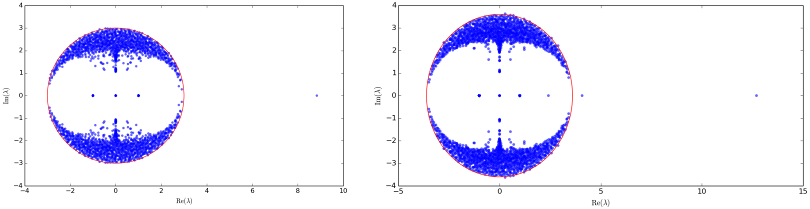
The non-backtracking matrix is also working for the degree-corrected SBM.
ongoing work with Gulikers and Massoulié.
We can adapt the non-backtracking matrix to deal with small cliques.



ongoing work with Caltagirone.

Some extensions

SBM with no noise $b = 0$ but with overlap.
Spectrum of the non-backtracking operator with $n = 1200$, $sn = 400$ and $a = 9$ and 13. The circle has radius $\sqrt{a(2-3s)}$ in each case.



Kaufmann, Bonald, Lelarge (2015)

Thank you for your attention !