Trivial fixed point, noise sensitivity, and non-backtracking walks.

Recall BP for SBM:

\[ U_{i \rightarrow j} = \sum_{\ell \in \delta ij} \text{arctanh} \left( \tanh(B) \tanh(\beta U_{i \rightarrow j}) \right) \]

\[ = \sum_{\ell \in \delta ij} f \left( U_{i \rightarrow j}^{(t)} \right) \quad \text{(\(*\))} \]

where

\[ f(x) = \frac{1}{2} \text{arctanh} \left( \tanh(\beta) \tanh(\beta x) \right) \]

\[ \beta = \frac{1}{8\log 2} \]

Note that

\[ \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \]

\[ f(0) = 0 \]

\[ f'(x) = \frac{1}{1 - \tanh^2(\beta x) \tanh^2(\beta x)} \Rightarrow f'(0) = \tanh(\beta) \]

Thus \[ U_{i \rightarrow j}^{(t)} = 0 \] is a trivial fixed point of \( (*) \).

This corresponds to the case where every node is equally likely to be \( S = 1, 3 \).

If BP gets stuck at this trivial fixed pt, then BP does not do better than random guessing.
Question: If we initially perturb the BP away from the trivial fixed point, should BP fly away from it—hopefully toward the truth—or fall back in?

In other words, is the trivial fixed pt stable or unstable?

Answer: It turns out that the trivial fixed pt is unstable if and only if

\[ \sum \frac{(a_i - b_i)^2}{2(a_i + b_i)} > 1 \]

\[ \text{the correlated recovery threshold}. \]

\[ \text{Noise sensitivity of BP.} \]

Let’s assume \[ U_{\text{eq}}^{(t)} = \varepsilon_{\text{eq}} \] for some small \( \varepsilon_{\text{eq}} \)

such that \[ \sum \varepsilon_{\text{eq}} = 0 \]

so that the terms and constants are balanced.

Then \[ U_{\text{eq}}^{(t+1)} = \sum \text{Rea}_{ij} f \left( U_{\text{eq}}^{(t)} \right) \]

Since \( U_{\text{eq}}^{(t)} \) is small, we can apply Taylor expansion and approximate

\[ f(U_{\text{eq}}^{(t)}) \approx f(0) + f'(0) U_{\text{eq}}^{(t)} \]

Thus \[ U_{\text{eq}}^{(t+1)} \approx \text{tanh} \beta \sum \text{Rea}_{ij} U_{\text{eq}}^{(t)} \]

\[ \text{Linearized BP} \]

In matrix notation,

\[ \begin{align*}
U^{(t+1)} & \approx \text{tanh} \beta \cdot B^T U^{(t)} \\
( B^T U^{(t)} )_{ij} & = \sum B_{ij} U^{(t)}_{eq}
\end{align*} \]
where $B \in \mathcal{E}$ is called the non-backtracking matrix

where $\mathcal{E} = \{ (u,v) : \exists w \in \mathcal{V} 3 \in \mathcal{E} \}$ denote the set of directed edges

and for any $e = (e_1, e_2), f = (f_1, f_2) \in \mathcal{E}$

$Bef = 1 \text{ if } e_2 = f_1 \text{ and } e_1 \neq f_2$

Note that: $(B^T)_e = Bef$

Now, the stability of the trivial fixed point reduces to the study of eigenvalues of $B$.

Row sum:

$\forall e = (u,v) \sum_{f \in \mathcal{E}} Bef = d \forall e$.

Remark: In $SBM(\frac{a}{b}, \frac{1}{n}, \frac{1}{n})$, $d \forall e \approx \frac{ab}{b} \approx d$.

or $G(\frac{a}{b}, \frac{1}{n})$

It turns out that $\lambda_1(B) = d + o(1)$.

The corresponding eigenvector is asymptotically aligned with the all-1 vector.

Remark: In $SBM(\frac{a}{b}, \frac{1}{n}, \frac{1}{n})$, the second largest eigenvalue of $B$

$\lambda(B) = \max \{ \frac{a-b}{2} \} + o(1)$

Recall we assume $\sum_{(e_1) \in \mathcal{E}_1} U^{(e_1)} = 0$, thus $U^{(e)}$ is orthogonal to the all-1 vector.

Thus, the recursion (**) becomes unstable when:

$tanh(\beta \cdot \frac{a-b}{2}) > 1 \iff \frac{a-b}{2} > 1 \iff \frac{ab}{b} > 1 \iff 2 > 1$. 

Some basic properties of NB matrix. Let $|V| = n$ and $|E| = m$.

1. $B$ is a $2m \times 2m$ matrix and can be partitioned into $4$ $m \times m$ blocks.

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

where $B_{11} = B_{22}^T$, $B_{12} = B_{21}^T$.

This is because given $e = (e_1, e_2)$, let $e^T = (e_2, e_1)$ denote its reversal.

Then $(B^T)_e f = B e^T f^{-1}$

Note: $B$ is not symmetric: $B e f \neq B f e$. The eigenvalues of $B$ are complex-valued.

2. In matrix notation, let $P_{ef} = 1$ if $e = f^{-1}$ denote the involution that maps a vector $(Xe : e e E)$ to $(Xe^T : e e E^T)$ such that $P^T = P$ and $P^2 = I$.

Then $B^T = P B P$.

In other words, $(B P)^T = B P$

and consequently $(B^T)^k = (P B P)^k = P \cdot B^k P$

$\Rightarrow (B^k P)^T = B^k P$.

3. The singular values of $B$ are given by

$$\sum_{d \neq 1 \in \text{rev } E} u \beta_1^2.$$
To see this, recall that $BP$ is symmetric. Denote the eigenvalue decomposition of $BP$ as

$$BP = \sum_{j=1}^{2m} S_j X_j X_j^T$$

Then

$$B = \sum_{j=1}^{2m} S_j X_j \{ (PX_j)^T \} = \sum_{j=1}^{2m} S_j X_j \{ (y_j)^T \} = \sum_{j=1}^{2m} \| S_j \| \; X_j \{ (\text{sgn}(S_j) \cdot y_j)^T \}$$

This is the singular value decomposition of $B$. Thus, it suffices to find the eigenvalues of $BP$.

Let's think about the eigenvalues and eigenvectors of $BP$

$$(BP^3) e = \sum_{f} (BP) e_f e_f^3 f$$

$$= \sum_{f} \sum_{e} \text{Be} \cdot \text{Pe} f e_f^3 f$$

$$= \sum_{f} \sum_{e} \text{Be}^1 e^1 e^1 e = f^{1} 3 f$$

$$= \sum_{f} \text{Be} f^{-1} 3 f.$$
If $3e = 1$, $3f = 1$ for all $f$ s.t. $Bf^{-1} = 1$, and $3e' = 0$ otherwise.

Then $(BP3)e = \sum_f Bef^{-1} 3f$

$= \sum_f Bef^{-1} = (dv - 1) = (dv - 1)3e$

Thus $3$ is an eigenvector of $BP$ with eigenvalue $dv - 1$.

There are $\sum_v (dv - 1) = n$ such eigenvectors.

If $3e = 1$ and $3f = -1$ for some $f$ s.t. $Bf^{-1} = 1$ and $3e' = 0$ otherwise.

Then $(BP3)e = \sum_f Bef^{-1} 3f$

$= -1 = (-1)3e$

Then $3$ is an eigenvector of $BP$ with eigenvalue $(-1)$.

There are $\sum_v (dv - 1) = 2M - n$ such eigenvectors.
Alternatively, we can also try to find eigenvalues of $BB^T$.

Note that

$$(BB^T)_{ef} = \sum_{e'} B_{ee'} (B')_{ef}$$

$$= \sum_{e'} B_{ee'} B_{fe'}$$

Thus

$$(BB^T) e = \sum_f (BB^T)_{ef} 3f$$

$$= \sum_f \sum_{e'} B_{ee'} B_{fe'} 3f$$

- If $3e=1$ for all incoming edges to node $v$.

Then

$$(BB^T) e = (dv-1)^2 = (dv-1)^2 3e.$$  

Thus $3$ is an eigenvector of $BB^T$ with eigenvalue $(dv-1)^2$.

- If $3e=1$ and $3f=-1$ for two different incoming edges to node $v$.

Then

$$(BB^T) e = \sum_f \sum_{e'} B_{ee'} B_{fe'} 3f$$
\[ = -\frac{\epsilon}{\epsilon'} B e' B e' + \frac{\epsilon}{\epsilon'} B e' \]

\[ = - (d v - 2) + (d v - 1) \]

\[ = 1. = 1. 3e. \]

Thus, \( 3 \) is an eigenvector of \( BB^T \) with eigenvalue 1.

\[ \text{Remarks} \]

For a d-regular graph with \( D = d I \). Then

\[ \lambda \text{ is an eigenvector of } B \iff \lambda = 1, -1, \text{ or } \frac{\lambda^2 + d - 1}{\lambda} \text{ is an eigenvector of } A. \]
Non-backtracking spectrum of random graphs.

![Diagram](image)

**Fig. 1.** Left: eigenvalues of $B$ for a realization of an Erdős–Rényi graph with parameters $(n, \alpha/n)$ with $n = 500$, $\alpha = 4$. Right: eigenvalues of $B$ for Example 2 with $n = 500$, $r = 2$, $a = 7$, $b = 1$.

BLM 18

Achieving correlated recovery threshold via non-backtrack matrix $B$.

Thm [BLM(18)]: Let $S = \frac{a + b}{2}$ and $d = \frac{a + b}{2}$. Let $U_2 = U_2(B)$ be the eigenvector of $B$ corresponding to the second-largest eigenvalue. Define

$$\hat{\sigma}_V = \text{sign} \left( \sum_{e = (e_1, e_2) \in E} (U_2)_e e_2 = V \right)$$

Then $\hat{\sigma}$ achieves correlated recovery, i.e., $|\langle \hat{\sigma}, \sigma \rangle| = \sqrt{n}$,

whp, provided that $S^2 > d$.

Remark: The proof is very sophisticated and beyond the scope here.
we explain some intuitions and provide some proof sketches.

**Question:** why the spectrum of $B$ is not hindered by high-degree vertices?

Let's consider the ER graph to illustrate the idea.

Recall that for $G(n, \frac{d}{n})$, the outlier eigenvalues of $A$ exist due to high-degree vertices.

This no longer occurs for $B$. To explain the intuition, we apply the trace method to $B^k (B^T)^k$ for some $k$.

We claim that for each oriented edge $e$:

$$(B^k (B^T)^k)_{ee} \tag{*}$$

$$= \# of NB walks starting with the $k$-th steps then reversing the last step and retracing $e$ in $k$ steps

Indeed, using the symmetry property

$$(B^k (B^T)^k)_{ee} = \sum_{e_2 \cdots e_k} B_{e_1} e_2 e_3 \cdots B_{e_k+1} B_{e_k+2} \cdots B_{e_k+1} e_{k+1}$$

$e_1 = e$, $e_{k+1} = e$
To simplify the computation of the number of NB closed walks, we will crucially exploit the locally tree-like structure of random graphs: for each vertex $u$, its $K$-hop neighborhood $N_k(u)$ is a tree, provided that $K$ is not too big, e.g., $K = o(\log n)$. If $N_k(u)$ is a tree, then for each summand in (i), the NB closed walk must reverse itself, otherwise there will be a cycle. Thus, on the event that the locally-tree-like structure holds, we have

\[
K \sum (B^k)^k \leq \begin{cases} \text{ate} & \text{if } K \text{ is not too big, e.g., } K = o(\log n) \end{cases}
\]

even if the degree of $v$ can be as large as $\frac{\log n}{(\log \log n)^2}$.

To justify the last step,

- For $G(n, d/n)$, the local neighborhood behaves as a Galton-Watson tree with offspring distribution $\text{pois}(a)$. 
For SBM \((n, \frac{a}{n}, \frac{b}{n})\), the local neighborhood behaves as a two-type Galton-Watson tree, where the total offspring distribution is still \(\text{pois}(d)\), and each node has \(\text{pois}(\frac{a}{2})\) children of type \(+\) and \(\text{pois}(\frac{b}{2})\) children of type \(-\). This can be encoded into the following matrix:

\[
\begin{bmatrix}
\frac{a}{2} & \frac{b}{2} \\
\frac{b}{2} & \frac{a}{2}
\end{bmatrix}
\]

Basic results in branching processes states that the total number of \(k\)-th generation children grows exponentially as \(d^k\).

Finally,

\[
\sum_{k=1}^{2m} |\lambda_k(B)|^{2k} = 11B^2l^2 = \text{Tr}(B^k(B^k)^T) = 2m d^k
\]

which implies that the "bulk" of the eigenvalues of \(B\) belong to the disk of radius \(10l\), i.e., all but a small fraction of eigenvalues of \(B\) should be in the disk \(|Z| < 10l + \epsilon\) for large \(n\).

Note: Later we will see that all but \((1+\epsilon)\text{ER}, 2n\text{SBM})\) eigenvalues are in the disk.
Question: Why is the second largest eigenvector of B informative under SBM?

Let \( 3 \in \mathbb{R}^E \) denote the 2nd eigenvector of B.

Let \( 3^* \in \mathbb{R}^E \) be defined as \( 3^*_e = \sigma(e)^2 \), where \( e = (e_1, e_2) \) and \( \sigma \) denotes the underlying community partition of SBM.

For each node, we estimate its Community label \( \sigma(u) \) by

\[
\hat{\sigma}_v = \text{sign} \left( \sum_{e : e \in V} 3_e \right)
\]

To gain some insight, let's proceed with the following wishful thinking:

Suppose we can apply a proper method to study the behavior of eigenvectors.

Since \( 3^* \) is almost orthogonal to the all-one vector, which is approximately the 1st eigenvector of B, let's hope to gain some insight about the 2nd eigenvector \( 3^* \) by studying \( B^k 3^* \) for some large \( k \).

Fix \( e = (u, v) \)

\[
(B^k 3^*)_e = \sum_f (B^k)_{ef} 3^*_f
\]

\[
= \sum_{f : \sigma(f_1) = +} (B^k)_{ef} - \sum_{f : \sigma(f_2) = -} (B^k)_{ef}
\]
The celebrated result of Kesten–Stigum [KS66] says that the behavior of \( Z^+ - Z^- \) is governed by the matrix \( \mathbf{M} \), whose eigenvalues are \( \lambda_1 = \delta \) and \( \lambda_2 = \lambda \).

If \( \lambda^2 > \lambda_1 \), then \( \frac{Z^+ - Z^-}{\lambda^2} \xrightarrow{L^2} X \)

where \( X \) is correlated with the label of the root \( V \).

This means that \( (B^k Z^*)_e \) has no-trivial correlation with \( V \), and correlated recovery can be achieved by majority vote.

Now, let's intuitively verify that \( (B^k Z^*)_e \) is approximately an eigenvector of \( B \) with eigenvalue \( \lambda_2 = S = \frac{\alpha - \beta}{2} \).

Indeed,

\[
B \left( \frac{B^k Z^*}{\lambda^2} \right) = \lambda_2 \cdot \left( \frac{B^{k+1} Z^*}{\lambda_2^2} \right)
\]

Now, if \( \lambda^2 > \lambda_1 \), \( \frac{B^k Z^*}{\lambda^2} \) converges as \( k \to \infty \).
Thus, \[
\frac{B^k \mathbf{x}^*}{\lambda_k^k} \sim \frac{B^{k+1} \mathbf{x}^*}{\lambda_{k+1}^{k+1}}
\]

Thus \[
B \left( \frac{B^k \mathbf{x}^*}{\lambda_k^k} \right) \sim \lambda_2 \cdot \left( \frac{B^k \mathbf{x}^*}{\lambda_k^k} \right).
\]

Similarly, we can argue that \[
\frac{B^k \mathbf{1}}{\lambda_k^k}
\]

is approximately the eigenvector of \(B\) with eigenvalue \(\lambda_1 = d = \frac{a+b}{2}\):

\[
B \cdot \left( \frac{B^k \mathbf{1}}{\lambda_1^k} \right) = \lambda_1 \cdot \left( \frac{B^{k+1} \mathbf{1}}{\lambda_1^{k+1}} \right) \sim \lambda_1 \cdot \left( \frac{B^k \mathbf{1}}{\lambda_k^k} \right)
\]