Opinion dynamics on non-sparse networks with community structure

Panagiotis Andreou

University of North Carolina at Chapel Hill

pandreou@email.unc.edu

Joint work with Prof. Mariana Olvera-Cravioto

Graduate Student Probability Conference

September 1, 2024

Table of contents

Motivation for Opinion Dynamics

1. Political Science:
   - Understand polarization in modern societies
   - Influence of the media in opinion shaping
   - Debunk myths about political personas

Motivation for Opinion Dynamics

1. Political Science:
   - Understand polarization in modern societies
   - Influence of the media in opinion shaping
   - Debunk myths about political personas
2. Probability Theory:
   - Stochastic Processes on Networks
   - Influence maximization in Social Networks
   - Community detection and clustering

Goals

1. Study the behavior of the system under varying **density** regimes and check for phase transitions.

Goals

1. Study the behavior of the system under varying **density** regimes and check for phase transitions.

2. Understand how the opinion process is affected by the passing of **time** and the change of the **network size**.

Goals

1. Study the behavior of the system under varying **density** regimes and check for phase transitions.
2. Understand how the opinion process is affected by the passing of **time** and the change of the **network size**.
3. Study the **typical** stationary opinion on an inhomogeneous network.

Modeling via Random Graphs

- In practice, we have a specified social network G.

Modeling via Random Graphs

- In practice, we have a specified social network G.
- *Idea:* think of G as a realization of a random graph model.

Modeling via Random Graphs

- In practice, we have a specified social network G.
- *Idea:* think of G as a realization of a random graph model.
- *Insight:* Even though the math of random graphs is harder, this idea allows us to talk about the *typical* stationary opinion and get way more general results that don't depend on the specific G.

Random Graphs

- Graphs where each edge is present with some probability.

- Useful for modeling first-order properties:

  1. Degree distribution
  2. Connectivity
  3. Community structure
  4. Average distances (small-world phenomenon)

Classification of Random Graphs

1. Static:
   - Snapshots of large networks
   - $G(V_n, E_n)$ and $G(V_{n+1}, E_{n+1})$ can be quite different
   - *Examples:* Erdös-Rényi, Stochastic Block Model, Configuration Model

Classification of Random Graphs

1. Static:
   - Snapshots of large networks
   - $G(V_n, E_n)$ and $G(V_{n+1}, E_{n+1})$ can be quite different
   - *Examples:* Erdös-Rényi, Stochastic Block Model, Configuration Model
2. Dynamic:
   - Addition of new vertices to the existing network
   - $G(V_n, E_n)$ and $G(V_{n+1}, E_{n+1})$ share most edges
   - *Examples:* Barabási-Albert model, Preferential Attachment networks

Classification of Random Graphs

1. Static:
   - Snapshots of large networks
   - $G(V_n, E_n)$ and $G(V_{n+1}, E_{n+1})$ can be quite different
   - *Examples:* Erdös-Rényi, Stochastic Block Model, Configuration Model
2. Dynamic:
   - Addition of new vertices to the existing network
   - $G(V_n, E_n)$ and $G(V_{n+1}, E_{n+1})$ share most edges
   - *Examples:* Barabási-Albert model, Preferential Attachment networks

Our opinion process is evolving on a **static** random graph.

Erdös - Rényi (static)

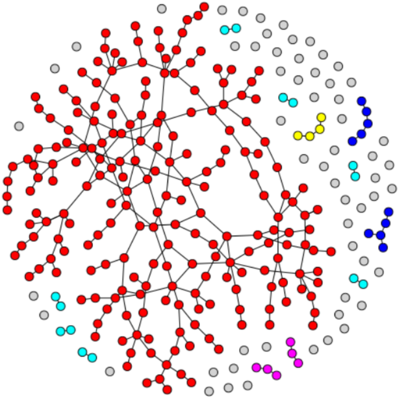

Figure: Different colors for different connected components (source: Fluid Limits and Random Graphs)

Stochastic Block Model (static)

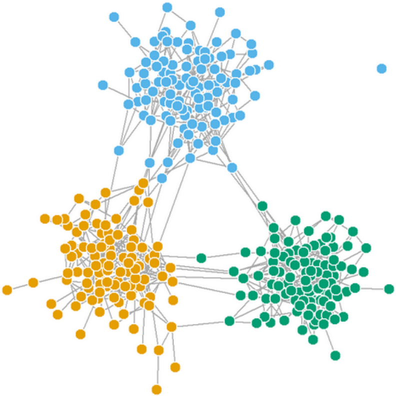

Figure: SBM with 3 communities
(source: Mathematics sin Fronteras)

Preferential Attachment (dynamic)

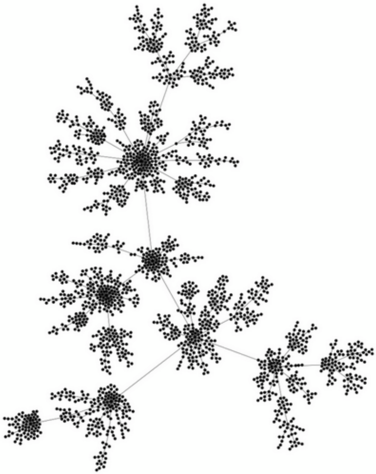

Figure: PA model - “the rich get richer” (source: ResearchGate)

dSBM

- Marked directed random graph $G(V_n, E_n; \mathscr{A}_n)$.

dSBM

- Marked directed random graph $G(V_n, E_n; \mathscr{A}_n)$.

- Each vertex $i \in V_n$ has a **community label** $J_i \in [K]$.

dSBM

- Marked directed random graph $G(V_n, E_n; \mathscr{A}_n)$.
- Each vertex $i \in V_n$ has a **community label** $J_i \in [K]$.
- Two nodes $i, j \in V_n$ are connected with an edge with probability

$$p_{ij}^{(n)} = \frac{\kappa(J_i, J_j)\theta_n}{n} \wedge 1,$$

where $\kappa \in \mathbb{R}_+^{K \times K}$ and θ_n is a **density** parameter.

Density regimes

- The expected degree of a vertex is of order θ_n.
- We call the graph **sparse** if $\theta_n = O(1)$.
- We call the graph **semi-sparse** if $\theta_n \to \infty$ and $\theta_n = O(\log n)$.
- We call the graph **dense** if $\frac{\theta_n}{\log n} \to \infty$ as $n \to \infty$.
- Our work covers the entire spectrum of sequences satisfying $\theta_n \to \infty$ as $n \to \infty$.

Our Opinion Process

- Individuals are represented by nodes on a directed SBM.
- An edge from j to i means "*i listens to j*".

Our Opinion Process

- Individuals are represented by nodes on a directed SBM.
- An edge from j to i means "*i listens to j*".
- $\mathbf{R}_i^{(k)} \in [-1, 1]^\ell$: the opinion that node i holds at time k on ℓ topics.
- $\mathbf{W}_i^{(k)} \in [-d, d]^\ell$: media signals that node i receives at time k on ℓ topics.
- $C_{ij} \in [0, 1]$: the weight that i puts in j's opinion.

Our Opinion Process

- Individuals are represented by nodes on a directed SBM.
- An edge from j to i means "*i listens to j*".
- $\mathbf{R}_i^{(k)} \in [-1,1]^\ell$: the opinion that node i holds at time k on ℓ topics.
- $\mathbf{W}_i^{(k)} \in [-d,d]^\ell$: media signals that node i receives at time k on ℓ topics.
- $C_{ij} \in [0,1]$: the weight that i puts in j's opinion.
- Update opinions according to the recursion

$$\mathbf{R}_i^{(k)} = c\sum_{j=1}^{n} C_{ij}\mathbf{R}_j^{(k-1)} + \mathbf{W}_i^{(k)} + (1-c-d)\mathbf{R}_i^{(k-1)}, \tag{1}$$

where $\{\mathbf{W}_i^{(k)} : k \geq 0\}$ are i.i.d. and $0 < c+d \leq 1$.

The Weights

- Define the weight C_{ij} that i puts on j's opinion as

$$C_{ij} = \frac{B_{ij} 1(j \to i)}{\sum_{r=1}^{n} B_{ir} 1(r \to i)} 1(D_i^- > 0, i \neq j),$$

where $D_i^- := \sum_{r=1}^{n} 1(r \to i)$ is the in-degree of i.

- The random variables B_{ij} are bounded and their distributions depend only on the communities J_i, J_j.

The Weights

- Define the weight C_{ij} that i puts on j's opinion as

$$C_{ij} = \frac{B_{ij}1(j \to i)}{\sum_{r=1}^{n} B_{ir}1(r \to i)}1(D_i^- > 0, i \neq j),$$

where $D_i^- := \sum_{r=1}^{n} 1(r \to i)$ is the in-degree of i.

- The random variables B_{ij} are bounded and their distributions depend only on the communities J_i, J_j.

- Note that $\sum_{j=1}^{n} C_{ij} = 1$ for every $i \in V_n$, i.e., C is stochastic.

The Weights

- Define the weight C_{ij} that i puts on j's opinion as

$$C_{ij} = \frac{B_{ij} 1(j \to i)}{\sum_{r=1}^{n} B_{ir} 1(r \to i)} 1(D_i^- > 0, i \neq j),$$

  where $D_i^- := \sum_{r=1}^{n} 1(r \to i)$ is the in-degree of i.

- The random variables B_{ij} are bounded and their distributions depend only on the communities J_i, J_j.

- Note that $\sum_{j=1}^{n} C_{ij} = 1$ for every $i \in V_n$, i.e., C is stochastic.

- Assume that the external media signals $\{\mathbf{W}_i^{(k)} : k \geq 0, i \in V_n\}$ are independent.

Simulations

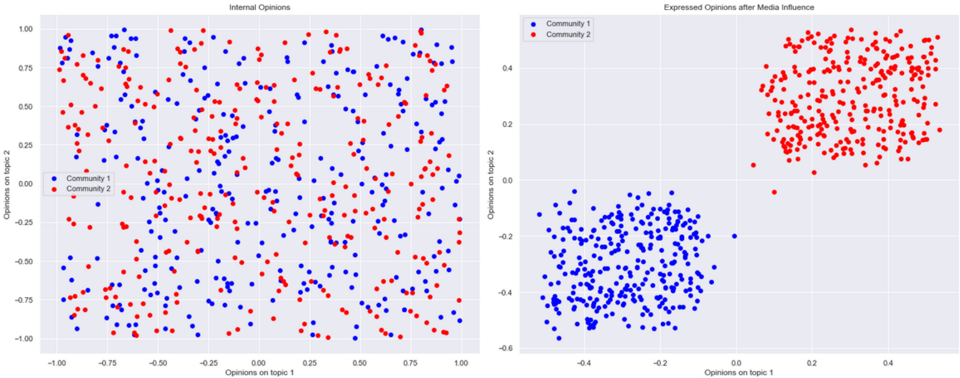

Figure: Internal vs. Expressed opinions. Initially, peoples' internal opinions are mixed. After they get **targeted media signals** based on their community belongings, they cluster into two separate opinion groups (polarization).

Simulations

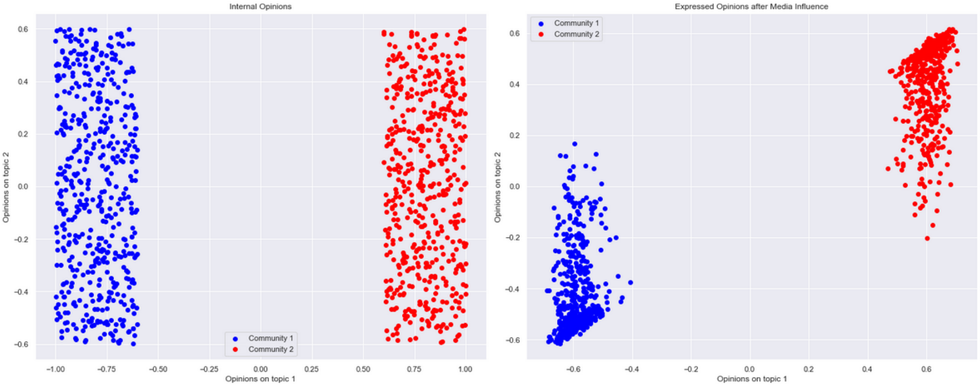

Figure: Initially, people feel more strongly about topic 1 and are indifferent about topic 2. After they are exposed to media signals that are **positively correlated between the two topics**, they start feeling more strongly about topic 2 as well. We say that *political personas are created*.

Mean-Field Theory

- Originally from Statistical Mechanics (P. Curie & P. Weiss, early 1900s).

Mean-Field Theory

- Originally from Statistical Mechanics (P. Curie & P. Weiss, early 1900s).
- *Idea:* replace all the interactions in a complex system by an average interaction.

Mean-Field Theory

- Originally from Statistical Mechanics (P. Curie & P. Weiss, early 1900s).
- *Idea:* replace all the interactions in a complex system by an average interaction.
- *Intuition:* the presence of many particles should reduce the effect of each particle on the entire system.

Mean-Field Theory

- Originally from Statistical Mechanics (P. Curie & P. Weiss, early 1900s).
- *Idea:* replace all the interactions in a complex system by an average interaction.
- *Intuition:* the presence of many particles should reduce the effect of each particle on the entire system.
- *Practicality:* reduce the initial high-dimensional problem of a stochastic process on a network to one of much lower dimension.

Notation

- Define $\mathscr{A}_n := \{J_i : i \in V_n\}$, $\mathscr{F}_n := \sigma(\mathscr{A}_n)$, $\mathbb{E}_n[\cdot] := E[\cdot|\mathscr{F}_n]$.
- $\pi_r^{(n)} := \frac{1}{n}\sum_{i=1}^{n} 1(J_i = r)$, the proportion of vertices having community $r \in [K]$.
- *Assumption:* $\pi_r^{(n)} \xrightarrow{P} \pi_r$, where $\pi_1 + \cdots + \pi_K = 1$.
- Define the matrix $M \in [0,1]^{K\times K}$ by

$$m_{rs} = \frac{\pi_s \beta_{r,s}\kappa(s,r)}{\pi_1\beta_{r1}\kappa(1,r) + \cdots + \pi_K\beta_{rK}\kappa(K,r)},$$

where $\beta_{r,s} = E[B_{ij}|J_i = r, J_j = s]$.

- Let $a_{l,s} = \binom{s}{l}(1-c-d)^{s-l}$, for $0 \le l \le s$.

The mean-field limit

- Approximate the original process $\{R^{(k)}\}_{k\geq 0}$ by another process $\{\mathcal{R}^{(k)}\}_{k\geq 0}$ whose main characteristic is that its rows $\{\boldsymbol{\mathcal{R}}_i^{(k)} : k \geq 0, i \in V_n\}$ are conditionally independent of each other given the community labels.

The mean-field limit

- Approximate the original process $\{R^{(k)}\}_{k\geq 0}$ by another process $\{\mathcal{R}^{(k)}\}_{k\geq 0}$ whose main characteristic is that its rows $\{\boldsymbol{\mathcal{R}}_i^{(k)} : k \geq 0, i \in V_n\}$ are conditionally independent of each other given the community labels.
- The approximating process is given by: $\mathcal{R}^{(0)} = R^{(0)}$ and

$$\boldsymbol{\mathcal{R}}_i^{(k)} = \sum_{t=0}^{k-1} (1-c-d)^t \mathbf{W}_i^{(k-t)} + 1(k \geq 2) \sum_{t=1}^{k-1} \sum_{s=1}^{t} a_{s,t} (M^s \bar{W})_{J_i \bullet} + \sum_{s=1}^{k} a_{s,k} (M^s \bar{R})_{J_i \bullet} + (1-c-d)^k \mathbf{R}_i^{(0)},$$

for $k \geq 1$ and $i \in V_n$, where $a_{s,t} = \binom{t}{s}(1-c-d)^{t-s} c^s$.

Main Theorem

Theorem (A., Olvera-Cravioto '24)

Suppose $\theta_n \geq (6H\Lambda_n)^2 \Delta_n \log n$*. Then, there exists a constant* $\Gamma < \infty$ *such that*

$$\sup_{k\geq 0} \mathbb{E}_n \left[\left\| R^{(k)} - \mathcal{R}^{(k)} \right\|_\infty \right] \leq \Gamma \left(\sqrt{\frac{\log n}{\theta_n}} + \mathcal{E}_n \right), \tag{2}$$

where $\mathcal{E}_n := \max_{1\leq r,s\leq K} \left| \frac{\pi_s^{(n)}\pi_r - \pi_s \pi_r^{(n)}}{\pi_r^{(n)}\pi_s} \right|$*. Moreover, for any sequence* θ_n *satisfying* $\theta_n \to \infty$ *as* $n \to \infty$,

$$\sup_{k\geq 0} \max_{i\in V_n} \mathbb{E}_n \left[\left\| \mathbf{R}_i^{(k)} - \boldsymbol{\mathcal{R}}_i^{(k)} \right\|_1 \right] \xrightarrow{P} 0, \tag{3}$$

as $n \to \infty$.

Remarks

- Since $\max_{i\in V_n}\mathbb{E}_n\left[\left\|\mathbf{R}_i^{(k)}-\boldsymbol{\mathcal{R}}_i^{(k)}\right\|_1\right]\le\mathbb{E}_n\left[\left\|R^{(k)}-\mathcal{R}^{(k)}\right\|_\infty\right]$, Theorem 1 shows that the approximation is stronger when $\theta_n/\log n\to\infty$, and it gradually weakens as the rate at which θ_n grows drops below the critical rate $\log n$.

Remarks

- Since $\max_{i\in V_n} \mathbb{E}_n\left[\left\|\mathbf{R}_i^{(k)} - \boldsymbol{\mathcal{R}}_i^{(k)}\right\|_1\right] \leq \mathbb{E}_n\left[\left\|R^{(k)} - \mathcal{R}^{(k)}\right\|_\infty\right]$, Theorem 1 shows that the approximation is stronger when $\theta_n/\log n \to \infty$, and it gradually weakens as the rate at which θ_n grows drops below the critical rate $\log n$.

- *Intuition*: the average number of neighbors that any vertex has grows with θ_n. The larger the number of neighbors, the more their aggregate contributions behave as the average opinion.

Remarks

- Since $\max_{i \in V_n} \mathbb{E}_n \left[\left\| \mathbf{R}_i^{(k)} - \boldsymbol{\mathcal{R}}_i^{(k)} \right\|_1 \right] \leq \mathbb{E}_n \left[\left\| R^{(k)} - \mathcal{R}^{(k)} \right\|_\infty \right]$, Theorem 1 shows that the approximation is stronger when $\theta_n / \log n \to \infty$, and it gradually weakens as the rate at which θ_n grows drops below the critical rate $\log n$.

- *Intuition*: the average number of neighbors that any vertex has grows with θ_n. The larger the number of neighbors, the more their aggregate contributions behave as the average opinion.

- The weakest result is valid for any $\theta_n \to \infty$, regardless of how slow the growth is.

Remarks

- Since $\max_{i\in V_n} \mathbb{E}_n\left[\left\|\mathbf{R}_i^{(k)} - \boldsymbol{\mathcal{R}}_i^{(k)}\right\|_1\right] \leq \mathbb{E}_n\left[\left\|R^{(k)} - \mathcal{R}^{(k)}\right\|_\infty\right]$, Theorem 1 shows that the approximation is stronger when $\theta_n/\log n \to \infty$, and it gradually weakens as the rate at which θ_n grows drops below the critical rate $\log n$.

- *Intuition*: the average number of neighbors that any vertex has grows with θ_n. The larger the number of neighbors, the more their aggregate contributions behave as the average opinion.

- The weakest result is valid for any $\theta_n \to \infty$, regardless of how slow the growth is.

- Since the rows in the limiting process $\{\mathcal{R}^{(k)} : k \geq 1\}$ are independent of each other, Theorem 1 yields that the trajectories of the process $\{R^{(k)} : k \geq 0\}$ are asymptotically independent, i.e., the system exhibits *propagation of chaos*.

Proof steps

1. First, write the opinion recursion in matrix form:

$$R^{(k)} = AR^{(k-1)} + W^{(k)},$$

where $A_{ij} = c\, C_{ij} 1(i \neq j) + (1 - c - d) 1(i = j)$.

Proof steps

1. First, write the opinion recursion in matrix form:

$$R^{(k)} = AR^{(k-1)} + W^{(k)},$$

where $A_{ij} = c\,C_{ij}1(i \neq j) + (1 - c - d)1(i = j)$.

2. Iterate the recursion:

$$R^{(k)} = \sum_{t=0}^{k-1} A^t W^{(k-t)} + A^k R^{(0)}.$$

3. Note that

$$A^t = (c\,C + (1-c-d)I)^t = \sum_{s=0}^{t} \binom{t}{s} (1-c-d)^{t-s} c^s C^s = \sum_{s=0}^{t} a_{s,t} C^s.$$

3. Note that

$$A^t = (c\,C + (1-c-d)I)^t = \sum_{s=0}^{t} \binom{t}{s} (1-c-d)^{t-s} c^s C^s = \sum_{s=0}^{t} a_{s,t} C^s.$$

4. Thus, the recursion becomes

$$R^{(k)} = \sum_{t=0}^{k-1} \sum_{s=0}^{t} a_{s,t} C^s W^{(k-t)} + \sum_{s=0}^{k} a_{s,k} C^s R^{(0)}.$$

3. Define the *approximate* mean $\tilde{M}$ of the matrix C:

$$\tilde{M}_{ij} = \frac{\beta_{J_i,J_j}\kappa(J_j, J_i)}{n\left(\beta_{J_i,1}\pi_1^{(n)}\kappa(1, J_i) + \cdots + \beta_{J_i,K}\pi_K^{(n)}\kappa(K, J_i)\right)} 1(i \neq j).$$

3. Define the *approximate* mean $\tilde{M}$ of the matrix C:

$$\tilde{M}_{ij} = \frac{\beta_{J_i,J_j}\kappa(J_j, J_i)}{n\left(\beta_{J_i,1}\pi_1^{(n)}\kappa(1, J_i) + \cdots + \beta_{J_i,K}\pi_K^{(n)}\kappa(K, J_i)\right)} 1(i \neq j).$$

4. Approximate meaning that

$$\tilde{M}_{ij} = \frac{\mathbb{E}_n[B_{ij}1(j \to i)]}{\mathbb{E}_n\left[\sum_{i=1}^{n} B_{ir}1(r \to i)\right]} \approx \mathbb{E}_n[C_{ij}].$$

6. *Key idea:* Define the **intermediate** process

$$\tilde{R}^{(k)} = \sum_{t=0}^{k-1}(1-c-d)^t W^{(k-t)} + 1(k \geq 2)\sum_{t=1}^{k-1}\sum_{s=1}^{t} a_{s,t}\tilde{M}^s \breve{W}$$
$$+ \sum_{s=1}^{k} a_{s,k}\tilde{M}^s \breve{R} + (1-c-d)^k R^{(0)}, \quad k \geq 1, \qquad \tilde{R}^{(0)} = R^{(0)},$$

where $\breve{W} := \mathbb{E}_n\left[W^{(0)}\right]$ and $\breve{R} := \mathbb{E}_n\left[R^{(0)}\right]$.

6. *Key idea:* Define the **intermediate** process

$$\tilde{R}^{(k)} = \sum_{t=0}^{k-1}(1-c-d)^t W^{(k-t)} + 1(k \geq 2)\sum_{t=1}^{k-1}\sum_{s=1}^{t} a_{s,t}\tilde{M}^s \breve{W} + \sum_{s=1}^{k} a_{s,k}\tilde{M}^s \breve{R} + (1-c-d)^k R^{(0)}, \quad k \geq 1, \qquad \tilde{R}^{(0)} = R^{(0)},$$

where $\breve{W} := \mathbb{E}_n\left[W^{(0)}\right]$ and $\breve{R} := \mathbb{E}_n\left[R^{(0)}\right]$.

7. *Intuition:* Intuitively, $\tilde{R}^{(k)}$ replaces all neighbor contributions with their approximate means, i.e., every term of the form $C^s X$ with $s \geq 1$ and X a random matrix is replaced with $\tilde{M}^s \mathbb{E}_n[X]$. That's the essence of mean-field approximation!

6. *Key idea:* Define the **intermediate** process

$$\tilde{R}^{(k)} = \sum_{t=0}^{k-1} (1-c-d)^t W^{(k-t)} + 1(k \geq 2) \sum_{t=1}^{k-1} \sum_{s=1}^{t} a_{s,t} \tilde{M}^s \breve{W} + \sum_{s=1}^{k} a_{s,k} \tilde{M}^s \breve{R} + (1-c-d)^k R^{(0)}, \quad k \geq 1, \qquad \tilde{R}^{(0)} = R^{(0)},$$

where $\breve{W} := \mathbb{E}_n \left[W^{(0)} \right]$ and $\breve{R} := \mathbb{E}_n \left[R^{(0)} \right]$.

7. *Intuition:* Intuitively, $\tilde{R}^{(k)}$ replaces all neighbor contributions with their approximate means, i.e., every term of the form $C^s X$ with $s \geq 1$ and X a random matrix is replaced with $\tilde{M}^s \mathbb{E}_n[X]$. That's the essence of mean-field approximation!

8. *Key fact:* The components of $\tilde{R}^{(k)}$ are independent, since the only randomness comes from the media signals.

7. *Goal:* bound $\mathbb{E}_n \left[\left\| R^{(k)} - \tilde{R}^{(k)} \right\|_p \right]$ and $\mathbb{E}_n \left[\left\| \tilde{R}^{(k)} - \mathcal{R}^{(k)} \right\|_p \right]$, for $p \geq 1$.

7. *Goal:* bound $\mathbb{E}_n\left[\left\|R^{(k)} - \tilde{R}^{(k)}\right\|_p\right]$ and $\mathbb{E}_n\left[\left\|\tilde{R}^{(k)} - \mathcal{R}^{(k)}\right\|_p\right]$, for $p \geq 1$.

8. Bound these terms for different ranges of θ_n:

   - If $\theta_n / \log n \to \infty$, use concentration inequalities *(Chernoff bounds for ratios of random binomial sums)*.

   - If $\theta_n / \log n \to 0$, use local weak convergence *(conditional independence of branching processes)*.

Time and Network Size

Theorem (A., Olvera-Cravioto '24)

There exists a random variable $\mathcal{R}_\emptyset$ such that $\mathbf{R}_{I_n} \Rightarrow \mathcal{R}_\emptyset$ as $n \to \infty$, and $\mathcal{R}_\emptyset^{(k)} \Rightarrow \mathcal{R}_\emptyset$ as $k \to \infty$. Hence, the following diagram commutes.

$$
\begin{tikzcd}
\mathbf{R}_{I_n}^{(k)} \arrow[r, "k\to\infty"] \arrow[d, "n\to\infty"] & \mathbf{R}_{I_n} \arrow[d, "n\to\infty"] \\
\mathcal{R}_\emptyset^{(k)} \arrow[r, "k\to\infty"] & \mathcal{R}_\emptyset
\end{tikzcd}
$$

Key takeaways

- Mathematical insights:
  1. When the graph is sufficiently dense, the high number of neighbors allows us to use Chernoff bounds.
  2. As the graph gets sparser, we need to get independence from somewhere else. That's where branching processes help.

Key takeaways

- Mathematical insights:
  1. When the graph is sufficiently dense, the high number of neighbors allows us to use Chernoff bounds.
  2. As the graph gets sparser, we need to get independence from somewhere else. That's where branching processes help.
- Practical implications:
  1. When the network is sparse, individual opinions matter significantly.
  2. As the network gets denser, individuals essentially don't interact but rather update based on the "average" opinion.

References

- Andreou, P. and Olvera-Cravioto, M. (2024). "Opinion dynamics on non-sparse networks with community structure". *arXiv:2401.04598*
- Avrachenkov, K., Kadavankandy, A., and Litvak, N. (2018). "Mean field analysis of personalized PageRank with implications for local graph clustering." *Journal of statistical physics*, 173:895–916.
- Fraiman, N., Lin, T.-C., and Olvera-Cravioto, M. (2024). "Opinion dynamics on directed complex networks." *Mathematics of Operations Research (To appear)*.
- Olvera-Cravioto, M. (2022). "Strong couplings for static locally tree-like random graphs." *Journal of Applied Probability*, 59(4):1261–1285.
- Lee, J. and Olvera-Cravioto, M. (2020). "PageRank on inhomogeneous random digraphs." *Stochastic Processes and their Applications*, 130(4):2312–2348.