Opinion Dynamics on Networks with Community Structure

Panagiotis Andreou

University of North Carolina at Chapel Hill

Joint work with Mariana Olvera-Cravioto

University Research Week

October 24, 2024

Networks and Random Graphs

Networks:

1. Nodes and Edges
2. Nodes represent entities, and edges represent connections between them.

Random Graphs:

1. Graphs where each edge is present with some probability.
2. Useful for modeling first-order properties, like degree distribution, community structure, etc.

Erdös - Rényi

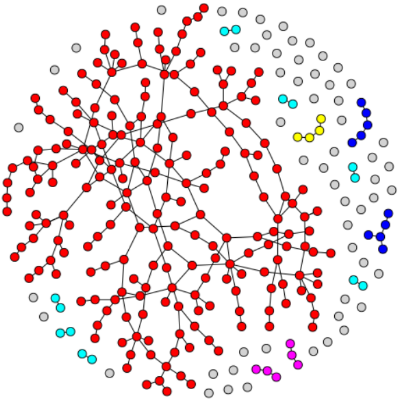

Figure: Different colors for different connected components (source: Fluid Limits and Random Graphs)

Stochastic Block Model

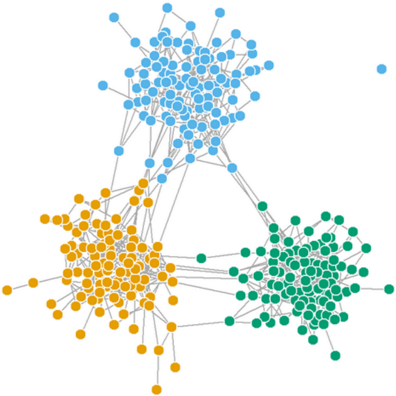

Figure: SBM with 3 communities
(source: Mathematics sin Fronteras)

Preferential Attachment

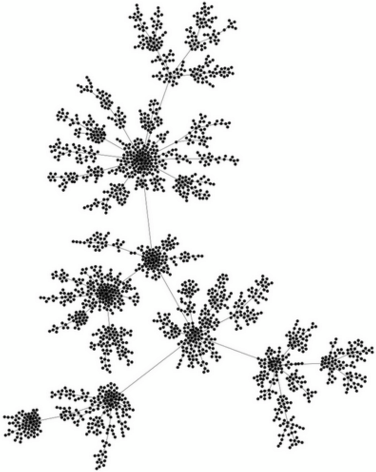

Figure: PA model - "the rich get richer" (source: ResearchGate)

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.

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 $0 < c + d \leq 1$.

Mean-Field Approximation

- *Issue:* our opinion process is too complicated because of the underlying network.

Mean-Field Approximation

- *Issue:* our opinion process is too complicated because of the underlying network.

- *Idea:* replace all the interactions in a complex system by an *average* interaction. That's the essence of mean-field approximation!

Mean-Field Approximation

- *Issue:* our opinion process is too complicated because of the underlying network.

- *Idea:* replace all the interactions in a complex system by an *average* interaction. That's the essence of mean-field approximation!

- *Intuition:* the presence of many particles should reduce the effect of each particle on the entire system.

Mean-Field Approximation

- *Issue:* our opinion process is too complicated because of the underlying network.
- *Idea:* replace all the interactions in a complex system by an *average* interaction. That's the essence of mean-field approximation!
- *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.

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 $a_{s,t} = \binom{t}{s}(1-c-d)^{t-s} c^s$.

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 $a_{s,t} = \binom{t}{s}(1-c-d)^{t-s}c^s$.

- *Key-property:* independent trajectories!

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 $n \to \infty$*.*

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.*

$$\begin{array}{ccc} \mathbf{R}_{I_n}^{(k)} & \xrightarrow{k\to\infty} & \mathbf{R}_{I_n} \\ \big\downarrow{\scriptstyle n\to\infty} & & \big\downarrow{\scriptstyle n\to\infty} \\ \mathcal{R}_\emptyset^{(k)} & \xrightarrow{k\to\infty} & \mathcal{R}_\emptyset \end{array}$$

Simulations

- $n = 2000$ individuals, grouped into $K = 2$ communities. The first half belong to community 1 and the other half to community 2.
- The kernel is

$$\kappa = \begin{pmatrix} 10 & 1 \\ 1 & 2 \end{pmatrix},$$

so the network structure plays an important role.
- Uniform media on topic 1, targeted media on topic 2.
- The initial opinions are uniform on both topics, regardless of community belongings.
- Consider different levels of densities θ_n, to see the effect of edge-density in the accuracy of the mean-field approximation.

Initial vs. Expressed Opinions

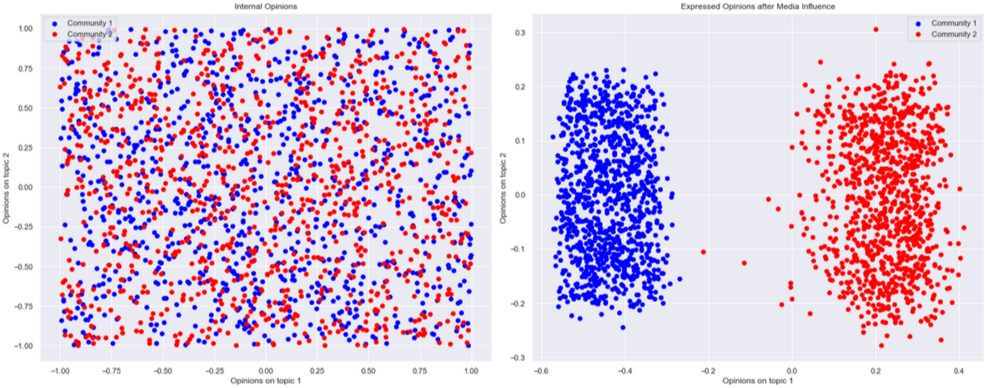

Figure: Initial and expressed opinions, with targeted media on the first topic and uniform media on the 2nd topic. The result is a separation in the expressed opinions across the first topic and a mixture of opinions on the second topic. We observe that for the first topic the opinions of community 1 individuals are more concentrated compared to those of community 2. This is justified by the fact that $\kappa(1,1) = 10 \gg 2 = \kappa(2,2)$ and at edge-density $\theta_n = \log n$ the network effect is still sufficiently present.

Accuracy of MFA

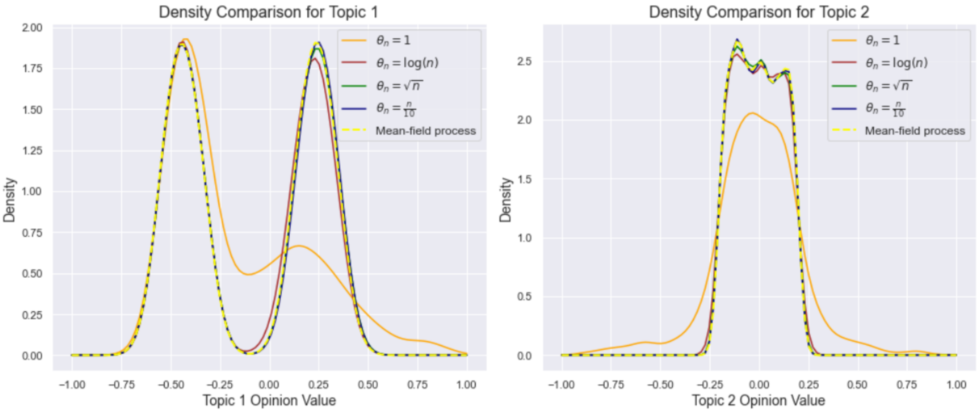

Figure: Distribution of opinions for both topics, for densities $\theta_n = 1, \log n, \sqrt{n}, \frac{n}{10}$. We observe what the theory suggests, namely, that the approximation is very tight on both topics for the dense regimes, $\theta_n = \frac{n}{10}$ and $\theta_n = \sqrt{n}$, with the densest one ($\theta_n = \frac{n}{10}$) being slightly more accurate. The approximation remains sufficiently accurate for the semi-sparse ($\theta_n = \log n$) regime, while it deteriorates in the sparse ($\theta_n = 1$) regime.

Political Personas

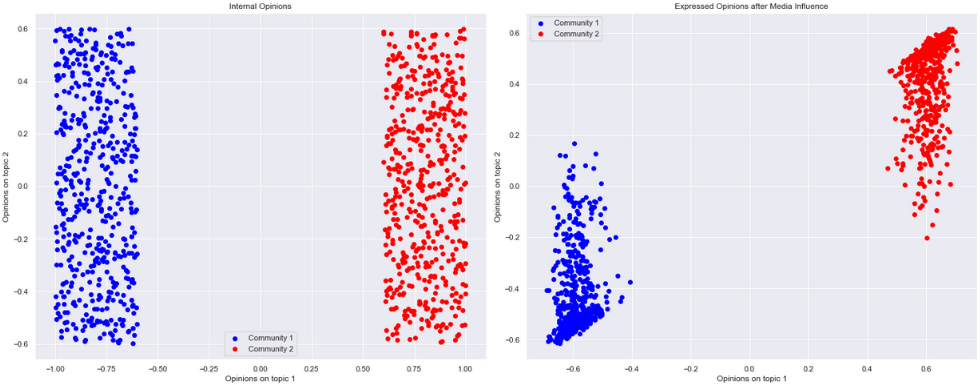

Figure: Initially, people feel more strongly about topic 1 and are indifferent about topic 2. After they are exposed to media signals that are **positively correlated between the two topics**, they start feeling more strongly about topic 2 as well. We say that *political personas are created*.

Key takeaways

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.