

Opinion Dynamics with Differentiated Messaging: Conformity and Polarization

Andrew Harkins

October 2026

University of Warwick

Introduction

Naive Learning and Opinion Dynamics in Networks

Setting: Large number of agents share opinions in a network. Non-strategic, non-Bayesian (aka naive learning) - DeGroot (1974), DeMarzo et al. (2003)

Limitations of DeGroot

1. Ad-hoc updating rule.
2. Mechanical 'echo chamber' effect.
3. No long-run opinion polarization.

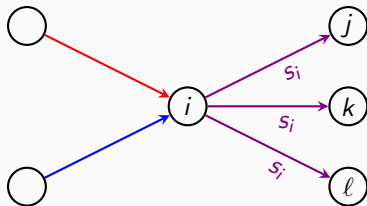
Much prior work which addresses these criticisms: e.g. Molavi et al. (2018), Cerreia-Vioglio et al. (2024); Banerjee and Compte (2024); Acemoglu et al. (2013), Büchel et al. (2015), Della Lena et al. (2026).

Key feature of the prior literature:

- ▶ Each agent broadcasts the same opinion to every neighbour.

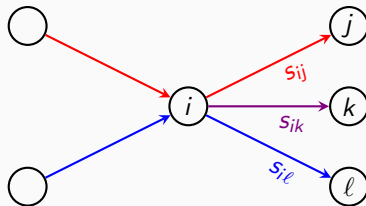
Differentiated Messaging

DeGroot



$$s_i = \mathbf{w}_i^T \mathbf{r}_i$$

This model



$$s_{ij} = \mathbf{w}_{ij}^T \mathbf{r}_i$$

$$s_{ik} = \mathbf{w}_{ik}^T \mathbf{r}_i$$

$$s_{i\ell} = \mathbf{w}_{i\ell}^T \mathbf{r}_i$$

Switch focus from belief dynamics (nodes) to message dynamics (edges).

- Study a richer class of opinion dynamics ($m > n$).
- Maintain tractability of linear non-strategic approach.

Messaging Rules

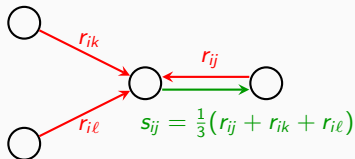
The vector of messages sent by i at $t + 1$ can be written as

$$\mathbf{s}_i(t + 1) = \mathbf{W}_i \mathbf{r}_i(t),$$

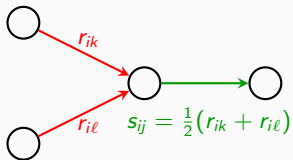
where $\mathbf{W}_i$ is called i 's *messaging rule*.

Examples:

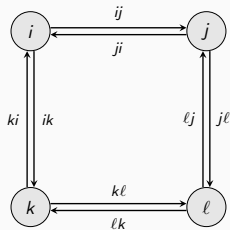
$$\mathbf{W}_i^U = \begin{matrix} & \begin{matrix} ji & ki & li \end{matrix} \\ \begin{matrix} ij \\ ik \\ il \end{matrix} & \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix} \end{matrix}$$



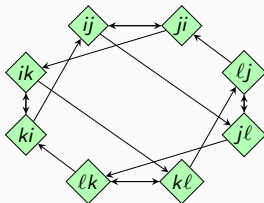
$$\mathbf{W}_i^{\text{NB}} = \begin{matrix} & \begin{matrix} ji & ki & li \end{matrix} \\ \begin{matrix} ij \\ ik \\ il \end{matrix} & \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{bmatrix} \end{matrix}$$



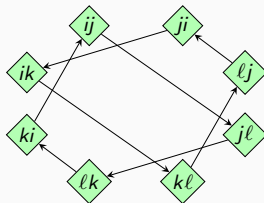
Influence in the Agent Graph vs. Message Graph



(a) Agent graph



(b) Message graph
Uniform rule



(c) Message graph
Non-backtracking rule

The message on edge ij can potentially influence:

- One step ($ij \rightarrow ji$), ($ij \rightarrow jl$)
- Two steps ($ij \rightarrow ji \rightarrow ij$), ($ij \rightarrow ji \rightarrow ik$), ($ij \rightarrow jl \rightarrow lj$), etc.

Consensus on Edges

Let $\mathbf{M}$ denote the $m \times m$ influence matrix encoding direct influence from one message to another.

- $\mathbf{M}$ is row-stochastic, non-negative.

Consensus on Edges

Let $\mathbf{M}$ denote the $m \times m$ influence matrix encoding direct influence from one message to another.

- $\mathbf{M}$ is row-stochastic, non-negative.

If $\mathbf{M}$ is irreducible and aperiodic then by Perron-Frobenius:

$$\mathbf{s}(t) = \mathbf{M}^t \mathbf{s}(0) \rightarrow \mathbf{1} c, \text{ as } t \rightarrow \infty$$

Sufficient condition for $\mathbf{M}$ irreducible and aperiodic:

1. The underlying agent graph $\mathcal{G}$ is strongly connected and aperiodic.
2. Each messaging rule $\mathbf{W}_i$ is strictly positive (any incoming message can influence any outgoing message).

Conformity

Conformity

Why send differentiated messages?

Conformity: Büchel et al. (2015), Anufriev et al. (2023).

Audience Tuning: Higgins and Rholes (1978).

"Pairwise Conformity":

$$s_{ij}(t+1) = \gamma r_{ij}(t) + (1 - \gamma)\bar{r}_i(t).$$

$\gamma > 0$ (Conformity),

$\gamma < 0$ (Anticonformity)

Conformity

Why send differentiated messages?

Conformity: Büchel et al. (2015), Anufriev et al. (2023).

Audience Tuning: Higgins and Rholes (1978).

"Pairwise Conformity":

$$s_{ij}(t+1) = \gamma r_{ij}(t) + (1-\gamma)\bar{r}_i(t).$$

$\gamma > 0$ (Conformity),

$\gamma < 0$ (Anticonformity)

$$\mathbf{W}_i^{\text{PC}} = \begin{bmatrix} \gamma + \frac{1-\gamma}{3} & \frac{1-\gamma}{3} & \frac{1-\gamma}{3} \\ \frac{1-\gamma}{3} & \gamma + \frac{1-\gamma}{3} & \frac{1-\gamma}{3} \\ \frac{1-\gamma}{3} & \frac{1-\gamma}{3} & \gamma + \frac{1-\gamma}{3} \end{bmatrix}$$

- ▶ If each i uses $\mathbf{W}_i^{\text{PC}}$ for $\gamma \in (-1, 1)$ then the consensus message is a uniform aggregation of all initial messages.

Conformity

Why send differentiated messages?

Conformity: Büchel et al. (2015), Anufriev et al. (2023).

Audience Tuning: Higgins and Rholes (1978).

"Pairwise Conformity":

$$s_{ij}(t+1) = \gamma r_{ij}(t) + (1-\gamma)\bar{r}_i(t).$$

$\gamma > 0$ (Conformity),

$\gamma < 0$ (Anticonformity)

$$\mathbf{W}_i^{\text{PC}} = \begin{bmatrix} \gamma + \frac{1-\gamma}{3} & \frac{1-\gamma}{3} & \frac{1-\gamma}{3} \\ \frac{1-\gamma}{3} & \gamma + \frac{1-\gamma}{3} & \frac{1-\gamma}{3} \\ \frac{1-\gamma}{3} & \frac{1-\gamma}{3} & \gamma + \frac{1-\gamma}{3} \end{bmatrix}$$

- ▶ If each i uses $\mathbf{W}_i^{\text{PC}}$ for $\gamma \in (-1, 1)$ then the consensus message is a uniform aggregation of all initial messages.
- ▶ Each agent has influence proportional to messages sent at $t = 0$, i.e. d_i .

Conformity

Why send differentiated messages?

Conformity: Büchel et al. (2015), Anufriev et al. (2023).

Audience Tuning: Higgins and Rholes (1978).

"Pairwise Conformity":

$$s_{ij}(t+1) = \gamma r_{ij}(t) + (1-\gamma)\bar{r}_i(t).$$

$\gamma > 0$ (Conformity),
 $\gamma < 0$ (Anticonformity)

$$\mathbf{W}_i^{\text{PC}} = \begin{bmatrix} \gamma + \frac{1-\gamma}{3} & \frac{1-\gamma}{3} & \frac{1-\gamma}{3} \\ \frac{1-\gamma}{3} & \gamma + \frac{1-\gamma}{3} & \frac{1-\gamma}{3} \\ \frac{1-\gamma}{3} & \frac{1-\gamma}{3} & \gamma + \frac{1-\gamma}{3} \end{bmatrix}$$

- ▶ If each i uses $\mathbf{W}_i^{\text{PC}}$ for $\gamma \in (-1, 1)$ then the consensus message is a uniform aggregation of all initial messages.
- ▶ Each agent has influence proportional to messages sent at $t = 0$, i.e. d_i .
- ▶ But, convergence speed differs according to γ .

Convergence Speed

- Convergence speed depends on the second largest eigenvalue modulus (SLEM) $\mu_* = \max\{|\mu_2(\mathbf{M})|, |\mu_m(\mathbf{M})|\}$.
- μ_* is a function of both network structure and the messaging rule.

Convergence Speed

- Convergence speed depends on the second largest eigenvalue modulus (SLEM) $\mu_* = \max\{|\mu_2(\mathbf{M})|, |\mu_m(\mathbf{M})|\}$.
- μ_* is a function of both network structure and the messaging rule.

Following Golub and Jackson (2012), define

Consensus distance:

$$\text{CD}(t; \mathbf{M}, \mathbf{s}(0)) = \|\mathbf{M}^t \mathbf{s}(0) - c\mathbf{1}\|_2.$$

Consensus time:

$$\text{CT}(\varepsilon, \mathbf{M}) = \max_{\|\mathbf{s}(0) - c\mathbf{1}\|_2=1} \min\{t : \text{CD}(t) \leq \varepsilon\}.$$

Convergence Speed under Pairwise Conformity

Results:

1. For $\gamma \in (0, 1)$: $\text{CT}(\varepsilon, \mathbf{M}) \approx \left\lceil \frac{\log(1/\varepsilon)}{\log(1/\mu_*)} \right\rceil$.
2. μ_* is strictly increasing in γ for $\gamma \in (\bar{\gamma}, 1)$, where $-1 < \bar{\gamma} \leq 0$.
 - Stronger conformity leads to longer consensus time if $\gamma \in (0, 1)$.
3. For $\gamma \in (-1, 0)$: μ_* characterises asymptotic convergence speed as $t \rightarrow \infty$.
 - Lower μ_* means faster asymptotic convergence.
 - Some level of anticonformity ($\gamma < 0$) accelerates asymptotic convergence in (almost) any graph.

Group Conformity

Group Conformity

Suppose that each node i has a publicly observable type $\theta_i \in \{A, B\}$.

Let $\bar{r}_i^{\theta_i}(t)$ denote the average opinion of i 's in-group.

Group Conformity:

$$s_{ij}(t+1) = \begin{cases} \gamma^{\text{in}} \bar{r}_i^{\theta_i}(t) + (1 - \gamma^{\text{in}}) \bar{r}_i(t), & \theta_j = \theta_i, \\ \gamma^{\text{out}} \bar{r}_i^{\theta_j}(t) + (1 - \gamma^{\text{out}}) \bar{r}_i(t), & \theta_j \neq \theta_i \end{cases}$$

Group Conformity

Suppose that each node i has a publicly observable type $\theta_i \in \{A, B\}$.

Let $\bar{r}_i^{\theta_i}(t)$ denote the average opinion of i 's in-group.

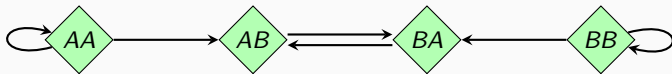
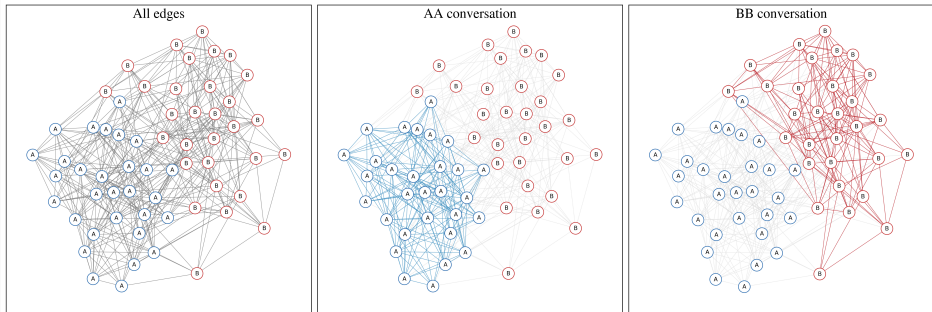
Group Conformity:

$$s_{ij}(t+1) = \begin{cases} \gamma^{\text{in}} \bar{r}_i^{\theta_i}(t) + (1 - \gamma^{\text{in}}) \bar{r}_i(t), & \theta_j = \theta_i, \\ \gamma^{\text{out}} \bar{r}_i^{\theta_i}(t) + (1 - \gamma^{\text{out}}) \bar{r}_i(t), & \theta_j \neq \theta_i \end{cases}$$

Results:

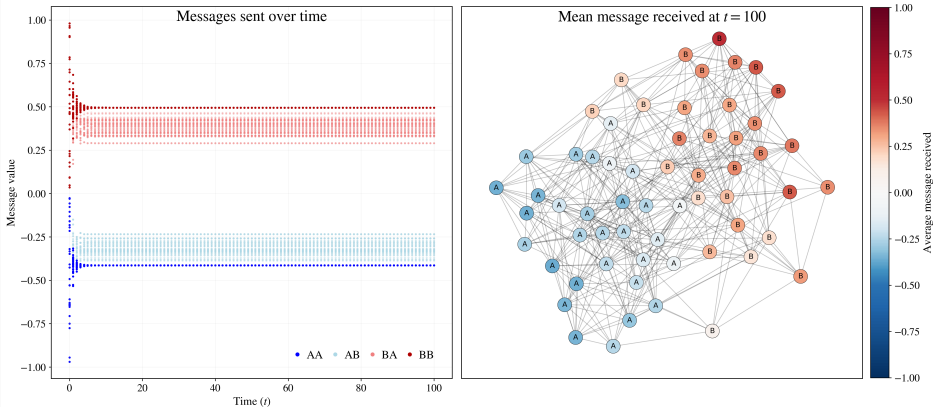
- For $\gamma^{\text{in}} \in (0, 1)$ and $\gamma^{\text{out}} \in (0, 1)$, then there is convergence to a consensus message.
- If we have $\gamma^{\text{in}} = 1$, then there is persistent disagreement (polarization).

Persistent Disagreement ($\gamma^{\text{in}} = 1, \gamma^{\text{out}} \in (0, 1)$)



- Underlying graph $\mathcal{G}$ is strongly connected, but the message graph is not.

Group Conformity and Polarization



Assume that the subgraphs $\mathcal{G}_A$ and $\mathcal{G}_B$ are strongly connected and aperiodic.

- Within-type messages converge to separate consensus c_{AA} and c_{BB} .
- The long-run messages on cross-type edges are convex combinations of c_{AA} and c_{BB} .
- Diversity in message content both across and within groups.

Polarization from Differentiated Messaging

- In the group conformity setting:
 - Within-type messages converge to one limit (distinct for As and Bs).
 - Cross-type messages converge to a different limit.
 - Key mechanism: Disagreement in messages emerges from exposure to multiple different 'conversations'.
- This yields long-run disagreement between As and Bs even when:
 - The underlying network is strongly connected.
 - There is no homophily between groups.
- Message content varies across type and within type, based on exposure to the out-group.

Appendix

Doubly Stochastic Messaging Rules

The rows of each $\mathbf{W}_i$ sum to 1, suppose that columns also sum to 1.

Economic interpretation:

- **Row sums = 1:** Truthfulness / no fabrication of messages.
- **Column sums = 1:** No source is over/under-represented.

$$\mathbf{W}_i^U = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix} \quad \mathbf{W}_i^{PC} = \begin{bmatrix} \gamma + \frac{1-\gamma}{3} & \frac{1-\gamma}{3} & \frac{1-\gamma}{3} \\ \frac{1-\gamma}{3} & \gamma + \frac{1-\gamma}{3} & \frac{1-\gamma}{3} \\ \frac{1-\gamma}{3} & \frac{1-\gamma}{3} & \gamma + \frac{1-\gamma}{3} \end{bmatrix}$$

Proposition

Assume that $\mathbf{M}$ is irreducible and aperiodic. Then $c = \frac{1}{m} \mathbf{1}^T \mathbf{s}(0)$ if and only if each $\mathbf{W}_i$ is doubly stochastic.

- Implication: Each node i creates d_i copies of $s_i(0)$. Node i 's influence is proportional to d_i , **as in DeGroot**.

Asymptotic Convergence Speed

Let $CD^*(t; \mathbf{M})$ be the worst-case distance from consensus, and $\mathbf{s}_t^*$ an initial condition attaining it at horizon t .

$$\rho_{\text{asym}}(\mathbf{M}) := - \lim_{t \rightarrow \infty} \frac{1}{t} \log(CD^*(t; \mathbf{M})).$$

- Intuition: average rate of exponential decay for worst case consensus distance as $t \rightarrow \infty$.

Polarization under Tribalists

Permuting $\mathbf{M}$

$$\hat{\mathbf{M}} = \left(\begin{array}{c|c} \mathbf{Q} & \Phi \\ \hline 0 & \Sigma \end{array} \right)$$

Messages as $t \rightarrow \infty$ are therefore

$$\lim_{t \rightarrow \infty} \hat{\mathbf{M}}^t \mathbf{s}(0) := \begin{pmatrix} \mathbf{s}_{AB}^\infty \\ \mathbf{s}_{BA}^\infty \\ \mathbf{s}_{AA}^\infty \\ \mathbf{s}_{BB}^\infty \end{pmatrix} = \begin{pmatrix} (\mathbf{I} - \mathbf{Q})^{-1} \Phi \begin{pmatrix} \mathbf{s}_{AA}(\infty) \\ \mathbf{s}_{BB}(\infty) \end{pmatrix} \\ \mathbf{s}_{AA}(\infty) \\ \mathbf{s}_{BB}(\infty) \end{pmatrix}$$

Observations:

1. Recurrent classes AA and BB reach consensus within class.
2. These AA and BB consensus values will generically **not** coincide.
3. Messages on transient edges (in AB or BA) converge to a convex combination of AA and BB 'consensus' messages.