Supplementary Material

Gabriella Dantas Franco[1*], Marta G. Couto[1],
Vítor V. Vasconcelos[1], Fernando P. Santos[1]

[1]Informatics Institute, University of Amsterdam, Street, City, 100190, Netherlands.

*Corresponding author(s). E-mail(s): g.dantasfranco@uva.nl;
Contributing authors: m.gomesdacunhacouto@uva.nl;
v.v.vasconcelos@uva.nl; f.p.santos@uva.nl;

A Relationship between Polarization and Radicalization

Radicalization and polarization values can vary independently. However, both metrics are not as uncorrelated as they may seem. To better understand the behavior of polarization and radicalization, we examine how both metrics relate analytically. We partition the opinion space into two groups: $+$, encompassing all individuals i such that $x_i > 0$; and in $-$, the complement set, meaning nodes with $x_i \leq 0$. Let the number of users on group $+$ be denoted by n_+ and on group $-$, by n_-. Let the respective average opinions on each partition be $\mu_+ > 0$ and $\mu_- \leq 0$. Therefore, we can rewrite radicalization as $R = (n_+\mu_+ - n_-\mu_-)/N$. Similarly, we can also rewrite polarization.

Theorem 1. *Let $Var(\mathbf{X}|+) = \sigma_+^2$ and $Var(\mathbf{X}|-) = \sigma_-^2$ define the variance of the $+$ and $-$ partition respectively, then*

$$Polarization^2(\mathbf{X}) = \frac{n_-}{N}\sigma_-^2 + \frac{n_+}{N}\sigma_+^2 + \frac{n_+ n_-}{N^2}(\mu_+ - \mu_-)^2. \tag{1}$$

Proof Given the law of total variance, we know that for any conditioning variable G,

$$Var(\mathbf{X}) = \mathbb{E}[Var(\mathbf{X}|G)] + Var(\mathbb{E}[\mathbf{X}|G]).$$

Assuming $G = \{+, -\}$ we then have that $\mathbb{E}[Var(\mathbf{X}|G)] = (n_+\sigma_+^2 + n_-\sigma_-^2)/N$, accounting for intra-group dissent. The second term represents inter-group polarization, and is given by $Var(\mathbb{E}[\mathbf{X}|G]) = \left(n_+(\mu_+ - \mu)^2 + n_-(\mu_- - \mu)^2\right)/N$, where μ is the opinion average over the whole population.

By construction the averages relate via $\mu = (n_-\mu_- + n_+\mu_+)/N$, and thus:

$$\text{Polarization}^2(\mathbf{X}) = Var(\mathbf{X}) = \frac{n_-}{N}\sigma_-^2 + \frac{n_+}{N}\sigma_+^2 + \frac{n_+ n_-}{N^2}(\mu_+ - \mu_-)^2.$$

□

Interestingly, the last term on the right-hand side of this equation, which captures variability between group means, can also be seen as a measure of extremism, as it considers how distant the groups are from each other $(\mu_+ - \mu_-)$. In particular, radicalization takes a similar form to this term, apart from an weighted arithmetic mean. This highlights how polarization depends not only on dispersion within opposing groups $(\sigma_\pm^2)$ and on the relative group sizes $(n_\pm)$, but also on the extremeness of group stances $(\mu_\pm)$.

Additionally, in the special case where there is complete consensus within each group (i.e. $\sigma_\pm^2 = 0$), then

$$\text{Polarization}(X) = \frac{(\mu_+ - \mu_-)}{N}\sqrt{n_+ n_-}\,.$$

This shows that, for fixed group average opinions, polarization is maximal when both groups have equal sizes (i.e. $n_+ = n_- = N/2$), a case in which radicalization is equal to two times the polarization value.

B Analytic approach

For the purpose of this work, in the main text we rely heavily on numerical simulations. Individuals are described by a deterministic equation at time t with their neighbors, and connections are updated stochastically according to the principles of link-recommendation algorithms, based on opinion and structural similarity between users. This implies that we simulate, for each parameter combination, many independent experiments of such co-evolutionary system until a predetermined time step. This is computationally costly, and might be impractical when increasing the complexity of the opinion dynamics or population size.

To better understand the numerical results obtained, we can approach our analysis from a simplified, analytic perspective. This allows us to have a broader perspective on parameter combinations that would allow a globally connected network on a large population and check whether these observations hold under a less computationally-demanding approach than agent-based simulations.

For this purpose, we can take the phenomenological approach introduced by Asikainen et al. (2020) and compare it to our results. Their model assumes that the number of individuals in each of the two possible groups $(+, -)$ is constant, and models the rewiring dynamics on a mean-field spirit. Although in our case there are dynamic opinions, the number of individuals in each group tends to be constant after the transition regime. This indicates their mean-field model might be applicable to our case as well.

As in (Asikainen et al. 2020) let us assume the model matrix:

$$M_{+-} = [(1-\rho)(T^2)_{+-} + \rho\, n_+](1-s_+)$$
$$M_{++} = [(1-\rho)(T^2)_{++} + \rho\, n_+]s_+\,,$$

where $(T^2) = (P + \mathrm{diag}(P))/\sum_i P_i + \mathrm{diag}(P)$, and P is the matrix of fraction of links. $P_{++} + P_{+-} + P_{--} = 1$ and $P_{+-} = P_{-+}$. $\sum_i P_i$ designates the sum over columns of P.

We modify their model as to assume that homophily does not interfere with triadic closure type of attachment, and include non-linearity by adding an exponent to this structural term. Leading to:

$$M_{+-} = (1-\rho)(T^2)^{\omega}_{+-} + \rho\, n_+(1-s_+)$$
$$M_{++} = (1-\rho)(T^2)^{\omega}_{++} + \rho\, n_+ s_+$$

Subsequently, we solve the group-wise master equation of the dynamics for the fraction of links within (F_{++}, F_{--}) and between groups (F_{+-}). We present the stationary results for the fraction of links between groups in Fig. S1.

This approach shows that, under these aggregated hypothesis, there is a threshold on the exponent ω that leads to group segregation. Similarly to what we observed, there are no connections between groups if transitivity is too strong, and echo chambers prevail (Cinelli et al. 2021; Baumann et al. 2020; Cinus et al. 2022). Nonetheless, if structural similarity is present but weak (low ω) the system allows for a globally connected network ($F_{+-} > 0$). Additionally, the transition value of ω will depend on how much recommendations prioritize opinion similarity: the more they consider opinion similarity, the lower the critical ω.

C Simulation Pseudo-code and parameters explored

Algorithm 1: Co-evolutionary Dynamics

```
Input   : nRuns, N, x_max, T_max
Output  : Result files
for r ← 1 to nRuns do
    Generate network G;
    Sample initial opinions in [−x_max, x_max];
    t ← 0;
    while t ≤ T_max do
        foreach node i ∈ G do
            Compute H_ij and S_ij;
            Select a node j to connect;
            Choose a connected node u uniformly at random to break link;
        Update opinions for all nodes: X[t + 1] ← f(X[t]);
        t ← t + 1;
return output files;
```

Table 1 Summary of parameters and notation

Parameter/Variable	Meaning and details	Domain	Values explored
η	Network similarity strength (transitivity)	$\eta \in \mathbb{R}$	$[0, 5]$
β	Opinion similarity strength (homophily)	$\beta \in \mathbb{R}$	$[0, 5]$
ρ	Relevance of homophily on recommender	$\rho \in [0, 1]$	$[0, 1]$
$\mathbf{X} = (x_1, x_2, \cdots, x_N)$	Population opinion vector	$\mathbf{X} \in \mathbb{R}^N$	
A	Adjacency matrix	$A_{ij} \in \{0, 1\}$	
N_i	Neighborhood of node i		
k_i	Degree of node i, $\sum_j A_{ij}$	$k_i \in \mathbb{N}$	
γ	Decay rate of opinion stance	$\gamma < 1, \gamma \in \mathbb{R}$	0.99
K	Interaction factor	$K \in \mathbb{R}$	0.1
α	Strength of interaction?	$\alpha \in \mathbb{R}$	0.3
N	Population size	$N \in \mathbb{N}$	$\{100, 300\}$
ϵ	Algorithmic noise	$\epsilon \in \mathbb{R}$	$\{0.0001, 0.001, 0.01, 0.1\}$

D Extra results and robustness checks

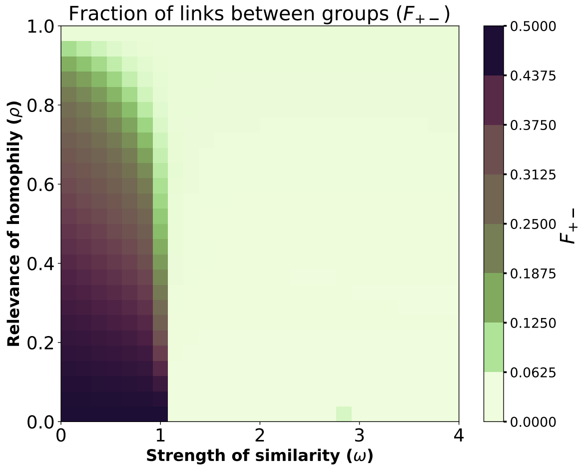

Fig. S1 Fraction of links between + and - groups when varying relevance of homophily (y-axis, ρ) and transitivity strength (x-axis, ω) for the phenomenological model. For values of transitivity strength that are above 1, the fraction of inter-group links drops to 0 for any value of ρ. The stronger the homophilic setting is, the more sensitive to transitivity strength the system becomes.

s

References

Asikainen A, Iñiguez G, Ureña-Carrión J, et al (2020) Cumulative effects of triadic closure and homophily in social networks. Science Advances 6(19):eaax7310

Baumann F, Lorenz-Spreen P, Sokolov IM, et al (2020) Modeling echo chambers and polarization dynamics in social networks. Physical Review Letters 124(4):048301

Cinelli M, De Francisci Morales G, Galeazzi A, et al (2021) The echo chamber effect on social media. Proceedings of the national academy of sciences 118(9):e2023301118

Cinus F, Minici M, Monti C, et al (2022) The effect of people recommenders on echo chambers and polarization. In: Proceedings of the International AAAI Conference

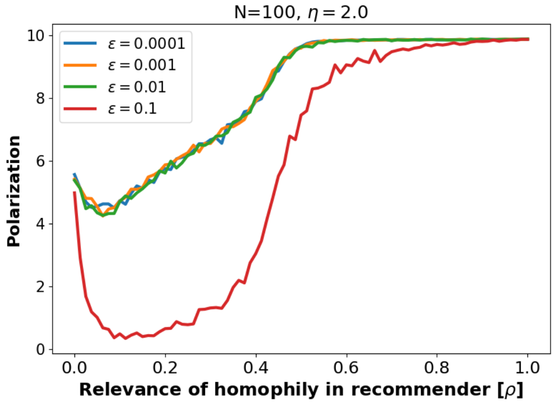

Fig. S2 Effect of varying noise parameter (ϵ) for $\eta = 2$ and $\beta = \eta$. The effect is negligible for $\epsilon = \{10^{-2}, 10^{-3}, 10^{-4}\}$, and the polarization curves for different values for the noise parameter remain virtually the same. For $\epsilon = 10^{-1}$ the polarization behavior changes significantly, as recommendation become more random.

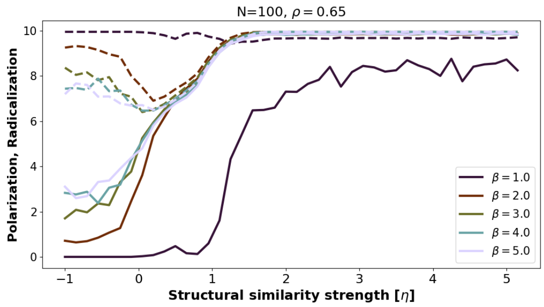

Fig. S3 Polarization (continuous) and radicalization (dashed) curves for a fixed value of $\rho = 0.65$ for the weight of homophily on recommender. As shown by Fig. 4 in the main text, this value of homophily would display values of radicalization lower than maximum for $\beta = 4$. Here we test whether for $\beta = \{1, 2, 3, 5\}$ there are values of parameter η that would allow that minimum to appear. We note that $\beta = 1$ is the only value that does not display a clear minimum on the radicalization curve, most likely because it is too close to the polarization transition to show any significant effect. All other values of β show a minimum value for the radicalization curve for a non-negative η.

on Web and Social Media, pp 90–101

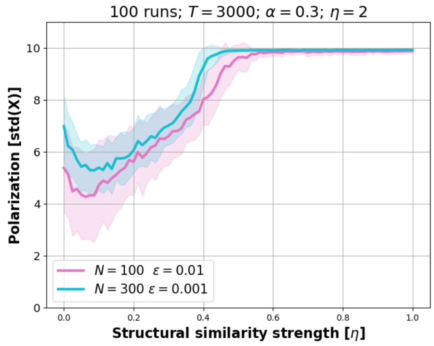

Fig. S4 Polarization curve as a function of homophily relevance on recommender ρ for $\eta = \beta$. We see that the qualitative behavior does not change significantly when we increase population size to $N = 300$.

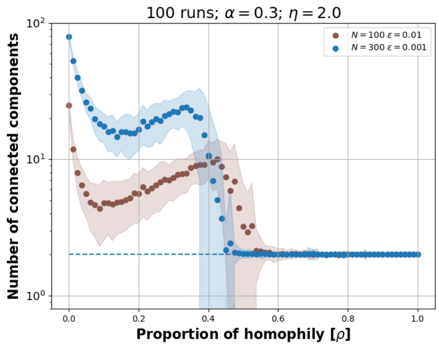

Fig. S5 Number of connected components as a function of the weight of homophily on the recommender ρ for two different network sizes. We note that, despite the absolute values varying between the two curves, the qualitative behavior remains the same. There is a global maximum of the number of connected components at $\rho = 0$ and a local one at an intermediate value of ρ. For high values of ρ the population converges to a stable state of 2 connected components.

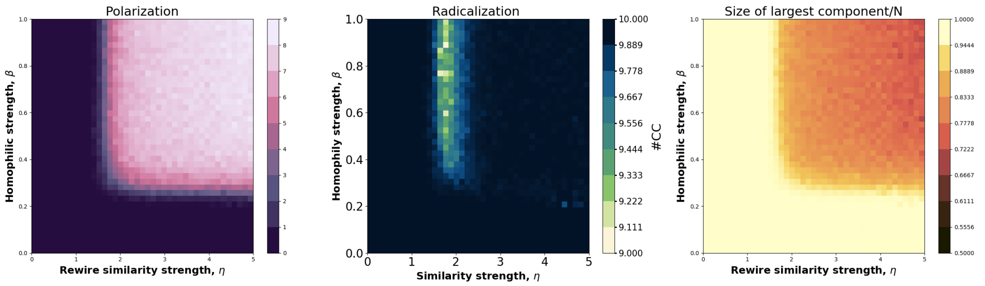

Fig. S6 (Left) Polarization, (center) radicalization and (right) normalized size of largest component on parameter space of similarity strengths (η, β) for opinion similarity weight of $\rho = 0.25$, averaged over 50 independent simulations. In the main text we show that a minimum appears on radicalization for intermediate values of opinion similarity weight. Here we show there is also a minimum for $\rho = 0.25$ corresponding to values of $\eta \approx 2$ and β above the polarization transition. We also note that, contrary to the case shown on the main text, this case does not correspond to a giant connected component, as the normalized size of the largest component is less than 1.

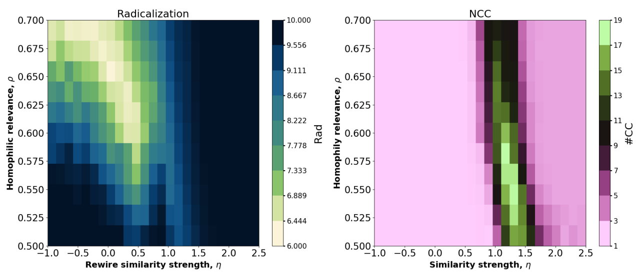

Fig. S7 Average radicalization (left) and number of connected components (right) over 50 independent simulations for $\beta = 4$. For $\rho > 0.65$ the minimum value of radicalization appears only when similarity strength is negative, indicating it is not always the case that useful recommendations via triadic closure dampens extremism on homophilic settings. For $\eta > 1$ we observe a transition on the number of connected components, indicating fragmentation. Such fragmentation criticality hinders the possibility of dampening extremism through generating a network as a single connected component.