Linking Theory and Empirics: A General Framework to Model Opinion Formation Processes

Ivan V. Kozitsin (kozitsin.ivan@mail.ru)
V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences

Research Article

Keywords: opinion dynamics models, model identification, polarization, online social networks, personalization algorithms

Posted Date: November 11th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1049437/v1

Linking theory and empirics: a general framework to model opinion formation processes

Ivan V. Kozitsin

Laboratory of Active Systems, V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, 65 Profsoyuznaya street, Moscow, 117997, Russian Federation

Department of Higher Mathematics, Moscow Institute of Physics and Technology, 9 Institutskiy per., Dolgoprudny, Moscow Region, 141701, Russian Federation

Email: kozitsin.ivan@mail.ru

We introduce a minimal opinion formation model, which is quite flexible and can reproduce a broad variety of the existing micro-influence assumptions and models. At the same time, the model can be easily calibrated on real data, upon which it imposes only a few requirements. From this perspective, our model can be considered as a bridge, connecting theoretical studies on opinion formation models and empirical research on social dynamics. We investigate the model analytically by using mean-field approximation and numerically. Our analysis is exemplified by recently reported empirical data drawn from an online social network. We demonstrate that the model calibrated on these data may reproduce fragmented and polarizing social systems. Furthermore, we manage to generate an artificial society that features properties quantitatively and qualitatively similar to those observed empirically at the macro scale. This ability became possible after we had advanced the model with two important communication features: selectivity and personalization algorithms.

Keywords: opinion dynamics models, model identification, polarization, online social networks, personalization algorithms

1. Introduction

Models of opinion dynamics (aka social-influence models) concern how individuals change their opinions as a response to information their receive from social environments. Understanding the processes of opinion dynamics is important due to its applications in many fields, including policy-making, business, and marketing. This branch of modeling is naturally interdisciplinary, attracting scholars from different fields, including social psychology, control theory, and physics. Despite the theoretical side of these models having been seriously advanced, the problem of its applications to describing real social processes remains an important question[1–5]. This issue is rooted in the complex nature of the social systems. To be more precise, it is extremely difficult to calibrate parameters of the underlying social-influence models, which operate with hardly formalizable entities. One prominent example is that of opinions themselves, which are intrinsic components of such models[5].

The proliferation of online social networks (OSNs) has made it possible to identify the dynamics of users' opinions on a large scale by applying machine learning techniques[6,7]. Combining these methods with tools elaborated in the field of social network analysis, one can obtain both dynamics of opinions and information on social connections between individuals[8]. Further, recent research has proposed a methodology to identify not only the structure of ties but also their weights, which describe how well these ties conduct social influence (aka influence networks)[9]. This information may be effectively integrated into the existing social-influence models, allowing to calibrate them, validate, and, further, make necessary predictions.

However, the step involving the integration of information gathered from OSN into the models is hampered because there is a substantial scope of the opinion dynamics models, and each model may require a specific data format. Hence, empirical data gathered through an experiment (most likely an expensive and time-consuming one) may be useful in one case but, unfortunately, unacceptable or requiring a lot of extra work with the data in other situations. Thus, each new dataset on opinion dynamics will potentially have a limited area of application in the sense that it could be investigated only by a restricted number of opinion formation models.

Therefore, we propose a quite general and, nonetheless, minimal model of opinion formation. On the one hand, the model is extremely flexible and can approximate a broad variety of the existing micro-influence assumptions and models. On the other hand, our model can be easily calibrated on empirical data, upon which it imposes a relatively small number of requirements. From this perspective, this model can serve as a bridge, connecting theoretical studies on opinion formation models on the one hand and empirical research on the other. We investigate the model analytically and numerically and exemplify our analysis with recently reported empirical data drawn from an online social network.

2. Literature

Social-influence models are numerous, and it is extremely difficult to classify all of them correctly. However, most of them can be grouped into main classes to some extent. Within this paper, we concentrate on the micro-level models, whereby the opinions of every individual are initialized and can both change and cause changes. In such models, the modeler analyzes how influence processes at the micro scale affect the resulting state of the social system at the macro scale. In contrast, so-called macro-level models describe the behavior of macroscopic variables, such as populations of individuals espousing particular opinions[10].

The literature emphasizes three main classification criteria. (1) Time in the model: continuous[11] or discrete[12]; (2) opinions: continuous[13] or discrete[14,15]; and (3) micro-assumption of social influence (see below). Because the empirical data are typically gathered in discrete time, we will focus hereafter on discrete-time models. Besides, we will assume that opinions are represented by scalar quantities describing individuals' positions on a single issue. More complex situations could arise when several topics (logically connected or independent) are analyzed at once[16,17]. However, gathering data on individuals' opinions on two or more topics simultaneously is a challenging and costly task. On this basis, we will focus on scalar opinions.

There are three main micro-assumptions regarding social influence, built upon a continuous opinion space[2]. In that space, in the case of one-dimensional opinions, one can say that an individual's opinion is affected by positive (aka assimilative) influence from a different opinion if the former moves towards the source of influence[18]. However, the literature on social psychology stipulates that if opinions are too distant, then the positive influence may not be accepted. Hence, the concept of bounded confidence has been introduced whereby only individuals espousing sufficiently close opinions may communicate[19,20]. In turn, if individuals' opinions become more distant following communication, then such a mechanism is termed negative (aka dissimilative) influence[21,22]. Note that in the case of a discrete opinion space, these assumptions are rather meaningless unless opinion values are ordered.

The positive influence mechanism explains elegantly how individuals reach agreement, and bounded confidence can model a situation when a social system is characterized by persistent disagreement (opinion fragmentation), whereas negative influence is one of the possible mechanisms explaining opinion polarization—the process in which individuals' opinions are stretched to the polar positions of the opinion space[23]. Thus, two camps of diametrically opposite opinions appear, a state of the social system that may have potentially dangerous consequences because it prevents democracy processes in general and consensus reaching in particular[24]. An important challenge in the field of opinion dynamics models is to determine the settings under which the model would be able to generate stable opinion polarization. Possible solutions here, apart from negative influence, are mass media[24], social feedback processes[23], social influence structure[25], arguments exchange[26,27], social identity[28], external events[29], or mistrust[30].

3. Methods

Model. We consider the system of N agents that are connected by a social network G. Each agent's opinion may take one of m values from the set $X = \{x_1, \dots, x_m\}$ that represents a discrete opinion space, a construction that is extensively studied in the sociophysics literature[1]. In some situations, one may assume that a binary relation is predetermined on the opinion space whereby types x_1 and x_m stand for the most radical and polar position in that space: $x_1 < x_2 < \cdots < x_m$. Depending on the context, we will endow these quantities with different values.

In our model, the time is discrete; we denote the opinion of agent i at time t by $o_i(t) \in \{x_1, \dots, x_m\}$. The population of agents having opinion x_k at time t is described by the quantity $Y_k(t) \in \{0,1, \dots, N\}$. Note that we assume that the system is "conservative" (agents do not leave the system, and there are no incoming agents): $\sum_{k=1}^{m} Y_k(t) = N$ for any t.

A key element of the model is a 3-D matrix $P = [p_{s,l,k}] \in \mathbb{R}^{m\times m\times m}$ where $s, l, k \in \{1, \dots, m\}$, which governs opinion dynamics that unfolds on the social network. This matrix, which we will refer to as the *transition matrix* hereafter, prescribes probabilities of opinion shifts. Each opinion shift is a move in the opinion space that is a result of peer influence processes. The first index in $p_{s,l,k}$ stands for an agent's current opinion x_s, the second index describes the opinion x_l of an influence source, and the last index represents the *potential* opinion x_k of the target agent at the next time point. In other words, $p_{s,l,k}$ is the probability that an agent with opinion x_s will switch their opinion to x_k after being influenced by an agent holding opinion x_l:

$$p_{s,l,k} = \Pr[o_i(t+1) = x_k \mid o_i(t) = x_s, o_{i\leftarrow}(t) = x_l],$$

where $o_{i\leftarrow}$ denotes the opinion of the source of influence. As such, we should require $\sum_{k=1}^{m} p_{s,l,k} = 1$ for any s and l. Note that $p_{s,l,s}$ represents the probability of staying at the current position after interaction with opinion x_l. In the following, it will be convenient to represent different transition matrices by considering their slices over the first index. We will denote these slices, which are row-stochastic 2-D matrices, by $P_{s,:,:} \in \mathbb{R}^{m\times m}$. In brief, the matrix $P_{s,:,:}$ outlines the behavior of an agent who has opinion x_s. Its rows indicate the opinion of an influence source, and its columns represent potential opinion options (see Example 1 in Appendix for details):

$$P_{s,:,:} = \begin{bmatrix} p_{s,1,1} & \cdots & p_{s,1,m} \\ \cdots & \cdots & \cdots \\ p_{s,m,1} & \cdots & p_{s,m,m} \end{bmatrix}.$$

The transition matrix encodes the influence processes in the model and may reflect a broad set of effects including persuasiveness of individuals as a function of their current positions and general trends in opinion change caused by influential external events[29]. Further, the transition matrix should be sensitive to such events and transform as a response to them. The number of parameters in the transition matrix depends only on the number of possible opinion values rather than on the total number of agents.

Now let us introduce how an opinion dynamics protocol is organized. At each time point t, a randomly chosen agent i is influenced by one of their neighbors j in the social network (the neighbor is also chosen at random). Hence, the opinion of the focal agent $o_i(t)$ changes (or remains the same) according to the probability distribution established in the transition matrix. This influence mechanism is asymmetric: the opinion of agent j does not change following the interaction.

This model is extremely general and can capture a broad set of micro-influence mechanisms and models introduced in the literature, including assimilative influence, bounded confidence, and antagonistic interactions (see Appendix 2 for details). However, the flexibility of our model is not without limitations. One can notice that the model assumes that agents with similar opinions should act equally on average in similar situations (that is, being exposed to comparable influence opinions), an assumption that significantly reduces the model's predictive power because not all ties transmit influence on an equal basis. In other words, our model can reproduce only the average patterns of opinion formation processes and approximate only anonymized forms of continuous opinion models (whereby all influence weights are equal). This issue makes the model less flexible than, for example, the DeGroot model, which allows individuals to allocate different influence weights to their peers. On the other hand, our model can easily explain the situation when an agent acts differently (by choosing positive or negative shift) as a response to the same influence opinion. This ability may be (i) due to the stochasticity of the model and (ii) because the model allows encoding of different opinion-changing strategies, depending on the current opinion of the focal individual.

Model identification. To calibrate the elaborated model, one needs to know (1) the trajectories of individual opinions and (2) the history of the individuals' communications. To be more precise, for a given individual i, one must know their opinion y_i before communication, the opinion of the influence source $y_{i\leftarrow}$, and the focal agent's opinion z_i after communication. One should first apply the procedure of discretization on experimental opinions if these opinions are initially continuous. Here, one should find the most appropriate discretization step, which should be a sort of compromise: a step that is too large could lead to losing potentially useful information on individuals' opinion trajectories, whereas too small a step results in a sharp increase in the number of transition matrix elements, which are now highly difficult to interpret and gives way to unnecessary data fluctuations. The resulting discrete opinions (for convenience, we denote them similarly) can be used to estimate the transition matrix elements:

$$p_{s,l,k} = \frac{\#\{i \mid (y_i = x_s)\&(y_{i\leftarrow} = x_l)\&(z_i = x_k)\}}{\#\{i \in I \mid (y_i = x_s)\&(y_{i\leftarrow} = x_l)\}},$$

where $\#\{\dots\}$ denotes the cardinal number of the set. To put it simply, $p_{s,l,k}$ is computed as the fraction of individuals who made opinion change $x_s \to x_k$ among those whose opinion is x_s and are influenced by opinion x_l. To compute all the transition matrix components, one needs to be provided with a substantial opinion diversity, which ensures that all combinations of x_s and x_l are represented in the data. Otherwise, the available statistics would be insufficient to calibrate the transition matrix. Individuals' opinions should be represented at least twice in the data: before and after interactions. However, longer opinion trajectories will be useful because they provide more room for analysis.

Note that we do not impose any requirements on the nature of empirical opinions. They could be discrete—in this case, we will consider low-dimensional transition matrices, or continuous—then, we firstly discretize opinions. In principle, the history of individual's communications (with whom they talked), which is highly difficult to retrieve in non-laboratory settings, can be replaced by more simple forms of structures of social connections, such as a friendship network, which could be relatively easily retrieved from the Web. Of course, more detailed information on how individuals interact with each other will make the model more precise, but in below, we will demonstrate that even a simple friendship network may serve as a good approximation of the actual communication network in the sense that it could simulate artificial social systems consistent with empirics at the macro scale.

Mean-field analysis. The model elaborated above can be studied using mean-field approximation. We assume that the social network is a complete graph whereby each agent can communicate with each one.

Establishing scaled time $\tau = \frac{t}{N}$, scaled time step $\delta\tau = \frac{1}{N}$, and introducing the normalized quantity $y_s(\tau) = \frac{Y_s(\tau)}{N}$, for large N we get the nonlinear autonomous system of differential equations (see Appendix 3):

$$\frac{dy_f(\tau)}{d\tau} = \sum_{s=1}^{m}\sum_{l=1}^{m}\sum_{k=1}^{m} y_s(\tau)y_l(\tau)p_{s,l,k}\left(\delta_{k,f} - \delta_{s,f}\right), \qquad f \in \{1, \dots, m\}. \qquad (1)$$

Let us now exemplify the elaborated analytical results with the low-dimension cases $m = 2$ and $m = 3$, enriching them with data describing real opinion dynamics processes. The corresponding dataset was recently reported in Refs.[31,32], where the author analyzed longitudinal data representing the dynamics of (continuous) opinions of a large-scale sample of OSN VKontakte users (hereafter – Dataset). Using the first two opinion snapshots from Dataset (see Appendix 4,5), we obtain the following transition matrix:

$$P_{1,:,:} = \begin{bmatrix} 0.975 & 0.025 \\ 0.952 & 0.048 \end{bmatrix}, P_{2,:,:} = \begin{bmatrix} 0.066 & 0.934 \\ 0.049 & 0.951 \end{bmatrix}. \qquad (2)$$

Note that Dataset includes three opinion snapshots overall. All results presented below remain virtually the same if one uses different snapshot sequences (e.g., first and third or second and third). One can observe that transition matrix (2) is very different from the one that represents the voter model (see Example 2 in Appendix); in the current case, agents rarely change their opinions if they are exposed to challenging positions. In turn, they have a nonzero chance of their opinion changing after being exposed to the same position—a phenomenon that was referred to in Ref.[33] as *anticonformity*. Nonetheless, the likelihood of an opinion shift slightly increases if two agents with opposite positions communicate, compared to when they have similar opinions. System (1) in this case (see formula (A3) in Appendix) has only one meaningful (located in the unit square) equilibrium point $y_1^* \approx 0.644, y_2^* \approx 0.356$, which is asymptotically stable.

Let us now consider triple opinion space $X = \{x_1, x_2, x_3\}$ $(m = 3)$. In this configuration, opinions x_1 and x_3 stand for antagonistic positions, whereas x_2 is somewhat neutral, located in the center. For brevity, we do not describe how system (1) appears in this case. Instead, we immediately obtain the transition matrix from Dataset (similarly to how we calculated transition matrix (2)):

$$P_{1,:,:} = \begin{bmatrix} 0.96 & 0.04 & 0 \\ 0.942 & 0.057 & 0.001 \\ 0.907 & 0.091 & 0.002 \end{bmatrix}, P_{2,:,:} = \begin{bmatrix} 0.039 & 0.952 & 0.008 \\ 0.021 & 0.969 & 0.01 \\ 0.02 & 0.944 & 0.036 \end{bmatrix}, P_{3,:,:} = \begin{bmatrix} 0.001 & 0.082 & 0.916 \\ 0.001 & 0.07 & 0.929 \\ 0.001 & 0.054 & 0.945 \end{bmatrix}. \qquad (3)$$

The slices of the transition matrix presented above reflect several remarkable features. First, larger values of opinion difference between individuals increase the rate of positive influence. One can observe this trend by inspecting slices $P_{1,:,:}$ and $P_{3,:,:}$: for example, the second and third columns in matrix $P_{1,:,:}$ are established so that their values increase as the index of the columns also rises. Further, individuals located in the middle of the opinion space may make negative shifts (pointing in the direction opposite to the influence source): $p_{2,1,3} > 0, p_{2,3,1} > 0$. However, according to (3), positive shifts (directing towards the source of influence) are more likely than negative ones: $p_{2,1,3} < p_{2,1,1}, p_{2,3,1} < p_{2,3,3}$. System (1) augmented with transition probabilities (3) has only one meaningful (located in the unit square) equilibrium point that can be obtained graphically. We find that $y_1^* \approx 0.269$, $y_2^* \approx 0.611$ (and $y_3^* \approx 0.12$ correspondingly). The Jacobian matrix at the equilibrium point has two different negative eigenvalues; hence, we can identify the behavior of the phase curves near the equilibrium point, which is an asymptotically stable nodal sink (see Figure 1 and Figure A1 in Appendix).

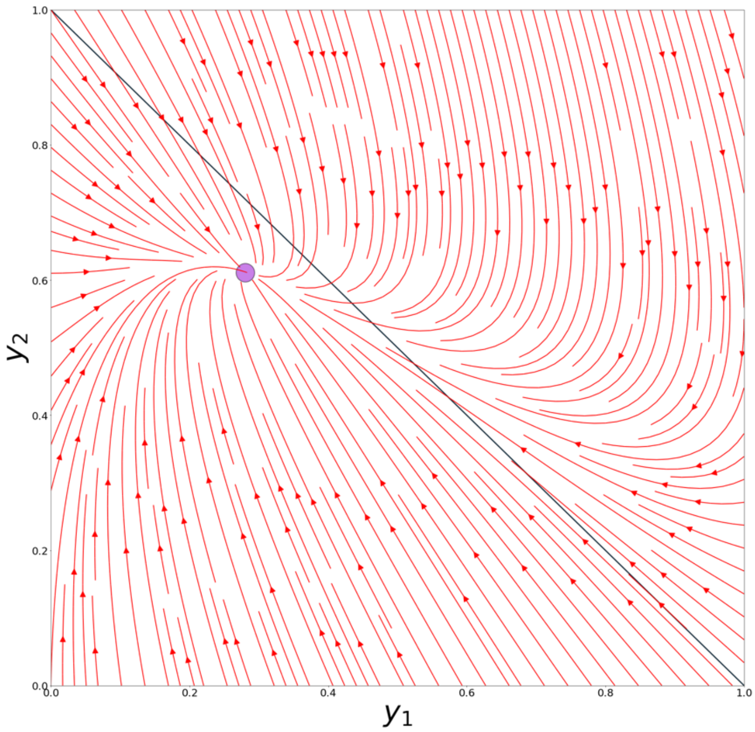

Figure 1. The violet circle marks the equilibrium point. The black solid line plots $y_1 + y_2 = 1$. We are interested in the area $y_1 + y_2 \leq 1$ beneath this line. The phase portrait of the system demonstrates that the equilibrium point is an asymptotically stable nodal sink.

Numerical experiments. We perform extensive numerical experiments to investigate the behavior of our model and, more specifically, support the analytical results derived above. Besides, we would like to understand whether our model can generate artificial social systems close (in some respects) to those observed empirically. We recognize that it is unlikely that we will be able to predict the opinion trajectories of specific individuals. However, we hope to predict the system's dynamics at the macroscopic level.

We build our further analysis on the investigation into how the model works on synthetic random networks by employing the following macroscopic metrics.

M1. The fraction of individuals y_i who have opinion x_s for $s \in \{1, \dots, m\}$. The combination of variables $y_1(t), \dots, y_m(t)$ represents public opinion at time t. In what follows, we will refer to them as public opinion variables.

M2. *Assortativity coefficient,* which measures whether the system at hand is homophilic (connected nodes tend to have similar opinions). To put it simply, the assortativity coefficient measures how similar the neighboring opinions are, compared with the configuration in which edges are placed at random (see Appendix 7 for details)[34]. For homophilic networks (most empirically observed social networks are homophilic), the assortativity coefficient takes positive values.

M3. The *dissimilarity coefficient* D, which effectively assesses the current level of polarization[35]. This measure is defined as the standard deviation of all pairwise opinion distances and takes values in the range between $D = 0$ (no polarization – all opinions are equal) and $D = 1$ (the highest level of polarization – individuals are divided into two equally sized camps located on the edges of the opinion space), provided that opinions lie in the interval $[-1, 1]$.

In experiments, we consider $N = 2000$ agents, who are endowed with randomly generated initial opinions. We tested different initial opinion configurations; however, we found that they have no influence on the asymptotic behavior of the model. Therefore, unless otherwise stated, opinions are initialized from the generalized Bernoulli distribution in which we set the uniform vector of probabilities ($y_i(0) = 1/m$ for $i \in \{1, \dots, m\}$). For large m, the initial opinion configuration is characterized by $C \approx 0, D \approx 0.5$. In each experiment, a new network is generated, as well as new initial opinion values. Apart from the complete graph model, we employ four synthetic graph models that are widely used in the social simulations literature[36]: (i) Erdős–Rényi network, (ii) random geometric network, (iii) Watts–Strogatz networks (WS1, WS2, and WS3), and (iv) Barabási–Albert network. Detailed information on network configurations and properties is presented in Table A2 (Appendix). Experiments typically lasted no more than one million iterations, a time interval that is sufficiently large to inspect the model's behavior. We repeated each experiment 20 times to gain more precise estimations.

We estimate the transition matrix using information from Dataset and concentrate on the first two opinion snapshots. We analyze cases $m = 2$, $m = 3$, and $m = 10$. The first two opinion space configurations require only a small number of variables to be parametrized, and this ability is useful in interpretations and for demonstrative purposes. Instead, the tenfold opinion space provides a more precise approximation of the underlying social processes. Further increase in m may lead to unnecessary fluctuations in the data and a sharp increase in the number of transition matric elements. Transition matrices for binary and triple opinion spaces have already been introduced in (2) and (3). The tenfold transition matrix is partially presented (and discussed) in Appendix (Tables A3-A5); its full representation can be found in Online Supplementary Materials.

The immediate observation that could be made from the estimated transition matrices is that the system has no stable states at the microscopic level: for every opinion vector, there is a nonzero probability that a randomly chosen agent will change their opinion at the next time point (even after exposure to the same opinion). However, in the following, we will demonstrate that the system has a stable converging tendency from the perspective of the macroscopic metrics.

Refs.[31,32] reported that the social system under consideration is homophilic with an assortativity coefficient of approximately 0.14, which may be considered as a not particularly strong (bot noticeable) rate of homophily. He also observed that users' opinions tend to stretch out to the edges of the opinion space in such a way that the fraction of individuals having middle-located (moderate) opinions decreases, individuals disposed on the left edge grow in number, and right-opinion users tend to keep their number or slightly decrease. To gain a more systematic understanding of the system, we calculate the metrics **M1**–**M3** using different discretization strategies for all three opinion snapshots from Dataset (see Table 1). We observe the following dynamical patterns: (i) growth of the population of individuals espousing left-side opinions, decrease of those who hold a middle-side opinion, and a relatively small decrease of right-side opinion persons; (ii) extremely small increasing trend in the homophily level; (iii) extremely small increasing trend in the polarization rate. As such, we hypothesize that the model calibrated on the same data should be able to achieve similar metric values (hereafter – reference values) at some point of its evolution and, further, demonstrate the same dynamics patterns near this point.

Table 1
Values of metrics **M1**–**M3** drawn from Dataset

Metric	Discretization step		
	$m = 2$	$m = 3$	$m = 10$
Public opinion variables	y_1: 0.568, 0.576, 0.582 y_2: 0.432, 0.424, 0.418	y_1: 0.161, 0.168, 0.174 y_2: 0.702, 0.697, 0.692 y_3: 0.137, 0.135, 0.134	
Assortativity coefficient	0.109, 0.109, 0.109	0.106, 0.107, 0.108	0.141, 0.141, 0.141
Dissimilarity coefficient		0.575, 0.585, 0.585	0.356, 0.366, 0.357

Note: the dissimilarity coefficient in the case of the binary opinion space is not useful and, therefore, we do not calculate it. The dynamics of public opinion variables in the case of the tenfold opinion space are too massive—one can find this information in Ref.[32] if necessary.

4. Results

Macroscopic behavior of the model. Our experiments reveal that regardless of the network topology, the behavior of public opinion variables remains the same. The evolution of the model can be decomposed into two periods (see Figure 2, panels A, B). In the first one, populations of camps $y_1(t), \dots, y_m(t)$ nearly monotonically converge to the theoretical predictions $y_1^*, \dots, y_m^*$. In the next period, the system fluctuates around these limiting values. At the beginning, the system is characterized by the almost zero assortativity because opinions are endowed at random. After a simulation has been initiated, the system rapidly becomes homophilic and then features fluctuations in a positive area demonstrating a relatively low rate of homophily (see Figure 2, panel C). For example, for the binary opinion space, the maximal assortativity rate observed is 0.05 (under WS1 topology), a value that is far from the empirical reference value (0.109). A similar difference in assortativity values was discerned in the case of the tenfold opinion space (0.07 in simulations against reference value 0.141). Further, we found that more clustered networks (random geometric, WS1, and WS2) tend to produce more homophilic systems (see Figure 2, panel D). The typical behavior of the polarization coefficient can be easily predicted because we know how public opinion evolves: starting from some point (that is determined by the initial opinion distribution), the polarization coefficient should firstly drift to the limiting value that characterizes the polarization of the stationary state opinion distribution $\{y_1^*, \dots, y_m^*\}$. Depending on the initial opinion configuration, this stage of evolution may feature growth (if the initial polarization level is lower than the limiting value) or decrease (if the initial polarization rate exceeds the limiting one). After the limiting value is achieved, the polarization coefficient should fluctuate around it. Further, this limiting value should not depend on the network topology because the latter does not affect the stationary state opinion distribution. Our numerical experiments (see Figure 2, panels E, F) confirm this proposal. We do not observe any relation between the network topology and the system's asymptotic polarization.

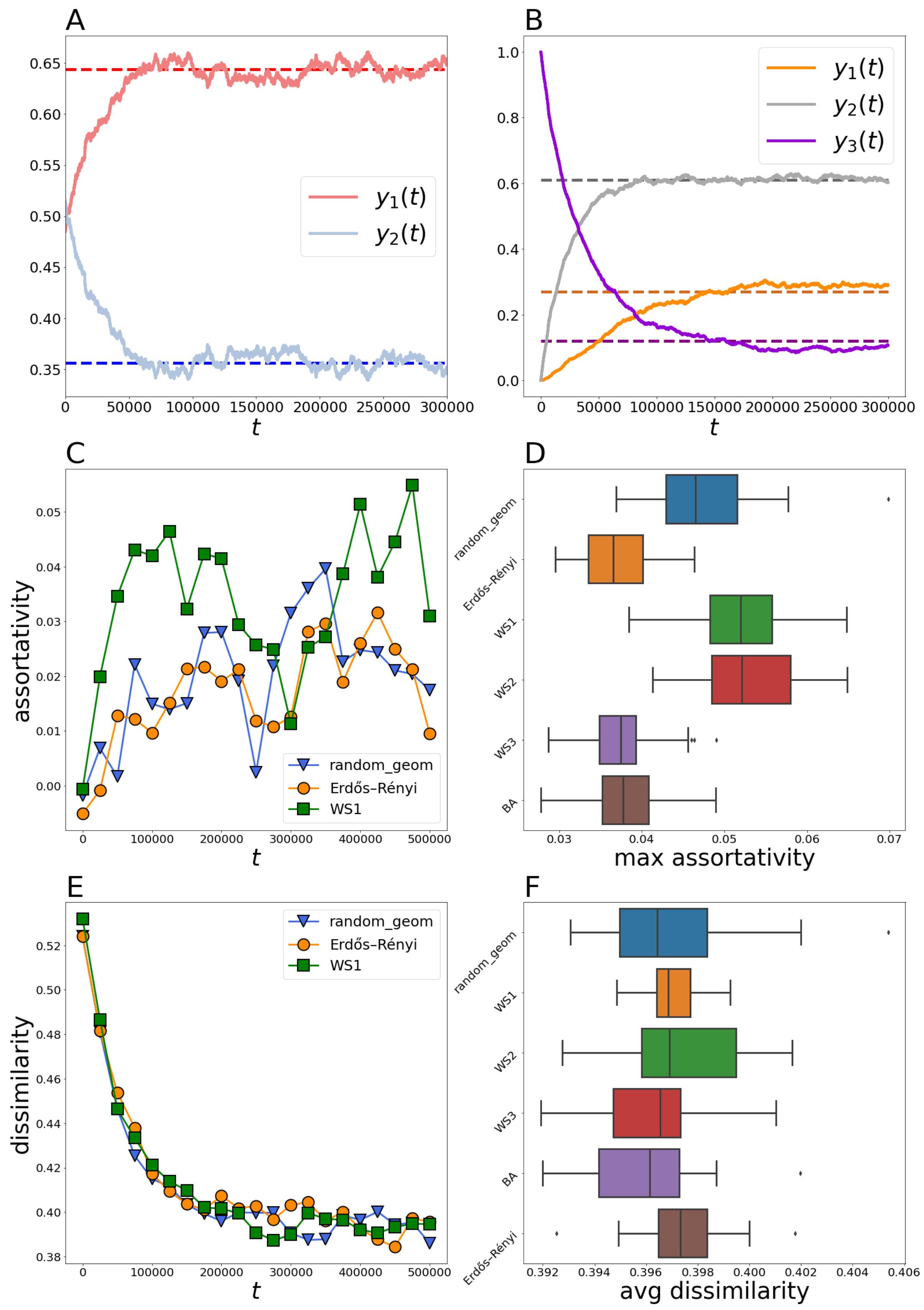

Figure 2. Ideal-typical evolution of public opinion variables in the model in the case of binary (panel A, initial opinions are drawn from the uniform distribution) and triple (panel B, initial opinions are equal to x_3) opinion spaces. Dashed lines plot theoretically predicted equilibrium values. (Panel C). Ideal-typical behavior of the assortativity coefficient in experiments ($m = 10$). The network topology with no clustering (Erdős–Rényi) exhibits lower assortativity values than clustered ones (random geometric and WS1). (Panel D). The

panel plots how the maximal value of assortativity during a single experiment varies with the network topology ($m = 10$). (Panel E). Ideal-typical dynamics of the polarization coefficient for different network topologies ($m = 10$). (Panel F). The average (over time) value of the dissimilarity coefficient after the system has been stabilized ($t \geq 250{,}000$) during a single experiment as a function of the network topology ($m = 10$).

The presented results indicate that the elaborated model demonstrates a stable converging tendency: its macroscopic parameters tend to some limiting values at first and then feature oscillations around them. Depending on a particular macroscopic metric, corresponding limiting values may (**M2**) or may not (**M1**, **M3**) be affected by the network topology. However, all of them are not sensitive to the initial opinion configuration. These findings contradict Refs.[23,37], who reported that the community organization of the network is one of the key factors of polarization. Instead, we found that this organization makes the network more homophilic. In the limit $t \to \infty$, the system features opinion fragmentation, which is characterized by persistent disagreement between agents. Further, one can also notice opinion polarization if the system begins from a densely concentrated opinion distribution. For example, if opinions are concentrated near the center of the opinion space, in this case, at the initial stage of the system's evolution, opinions will be prone to antagonistic or anticonformity-based interactions (in Ref.[33], this phenomenon was attributed to the striving for uniqueness[38]) and, thus, will move towards the extreme opinion values x_1 and x_m. However, increasing pairwise distances between agents' opinions will give way to assimilative interactions that are less likely to occur if opinions are too close. This process will continue until a sort of balance between assimilative and antagonistic interactions is reached.

The model exhibits good predictability from the perspective of the public opinion dynamics: the mean-field approximation is plausible not only for complete graphs (under the assumption of which this approximation was obtained) but also in other settings whereby social connections may have a quite complex structure. Next, the model can reproduce positive assortativity (i.e., can create a homophilic social system). However, the observed homophily rates are far lower than the empirical reference values. Further, dissimilarity demonstrated by the system after stabilization (which is slightly more than the corresponding reference values) provides a clue that the system can reach the reference polarization level if its initial degree of polarization is lower than the asymptotic one—this situation could arise if initial opinions are densely concentrated (for example, near the center of the opinion space).

To understand whether the simulation system can demonstrate similar behavior to the empirically observed one (see Table 1), we use the triple opinion space because it requires only a few macroscopic metric values to be analyzed (for example, in the case $m = 10$, we need to inspect the behavior of 12 variables). Quite close matching (see Figure 3) between simulations and empirics can be achieved if one begins a simulation run from opinion distribution $y_1(0) = 0.07, y_2(0) = 0.79, y_3(0) = 0.14$ (in Discussion, we will explain where this configuration of initial opinions may stem from). In this case, at moment $t \approx 20{,}000$, we notice that metrics **M1** and **M3** match the reference values as well as their local dynamic patterns (see Table 1). An exception is the assortativity coefficient, the simulation values of which are below the reference one. These findings lead us to the following question: *what modifications should one add to the model to make it possible to reproduce empirically observed homophily?*

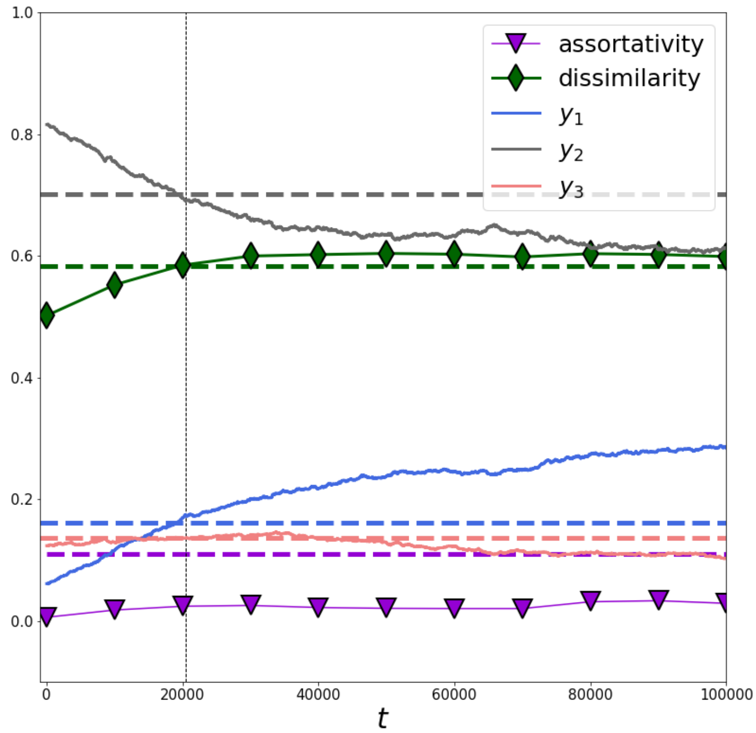

Figure 3. Ideal-typical behavior of the system in the case of the triple opinion space if initial opinions are drawn from the distribution $y_1(0) = 0.07, y_2(0) = 0.79, y_3(0) = 0.14$. Dashed horizontal lines plot reference values from Table 1 (we take values from the first opinion snapshot). Common colors indicate the same metrics. The vertical dashed line marks the point of time at which all macroscopic metrics except **M2** tend to coincide with the empirical reference values. One can observe a sharp discord in assortativity values between empirics (violet dashed line) and simulations (violet curve with triangle markers).

Possible modifications. Let us present some possible explanations of the observed discord in the values of the assortativity coefficient between simulations and empirics and feasible avenues for resolving this conflict.

Methodological errors in Dataset. A discerned divergence between numerical experiments and empirics may stem from methodological errors in obtaining the underlying empirical data. Because these data were derived through a natural experiment in which individuals' opinions were estimated by using some heuristics (more precisely, it was assumed that users' opinions are reflected by the information sources they are subscribed to), it could mean that the empirical reference values we strive to equal are incorrect. Unfortunately, we cannot fix this problem; hence, we leave this situation beyond the scope of this article and assume that the reference values are identified correctly.

Noise in the estimated transition matrices. A slightly different idea is to suppose that the reference values of macroscopic metrics *are correct* (i.e., individuals' opinions were estimated faithfully), but the transition matrix is identified *with errors*. The point is that errors in opinion identification naturally lead to mistakes in the estimated transition matrix (this is precisely what the previous paragraph discusses); however, to identify the transition matrix accurately, apart from the opinion values knowledge, we should also be able to determine the influence individuals are exposed to. This problem requires uncovering the influence network[9], which is also a challenging task. Note that in Dataset, all influence weights are assumed to be equal (the influence opinion directed on a user is the average of opinions of the user's friends), an assumption which is unlikely true: some ties may be more successful in social influence transmission[39]. Further, an influence system retrieved from OSN is likely incomplete because it neglects to consider the influence beyond the online world (more precisely, beyond a particular OSN). One more effect stems from our algorithm of the transition matrix identification and the nature of data: opinion dynamics presented in Dataset are identified under the assumption of many-to-one interactions, whereas the model assumes one-to-one interactions. Besides, such factors as selectivity or personalization algorithms (see below) may dominate. As a result, the real transition matrix may differ from the estimated one.

To model mistakes in the estimated transition matrix, we consider the binary opinion space, in which these mistakes may be parametrized by two variables:

$$P_{1,:,:} = \begin{bmatrix} 0.975+\alpha & 0.025-\alpha \\ 0.952-\beta & 0.048+\beta \end{bmatrix}, P_{2,:,:} = \begin{bmatrix} 0.066+\beta & 0.934-\beta \\ 0.049-\alpha & 0.951+\alpha \end{bmatrix}. \quad (4)$$

In (4), variables α and β represent small perturbations that stand for the difference between the estimated transition matrix and the real one. In the general case, we should use four different variables to parametrize noise. However, due to technical reasons, we have decided to consider "symmetric" disturbances. Using this approach, we analyze how the values of α and β affect the system's assortativity. We investigate domain $\{-0.02 \leq \alpha \leq 0.02, -0.02 \leq \beta \leq 0.02\}$ with step 0.005 across both dimensions. Assuming $\alpha = 0$ and $\beta = 0$, we return to the earlier-obtained transition matrix (2).

Selectivity is a well-documented tendency of social actors that we could not ignore, and it creates ties with those having similar opinions and breaks connections that promote uncomfortable information[40,41]. Along with (assimilative) social influence, selectivity is considered to be a main driver that makes social networks homophilic[42,43]. As such, we may hypothesize that by adding selectivity into the model, we will increase the level of homophily, which is one of our purposes.

There is a line of research devoted to modeling coevolutionary processes whereby social influence mechanisms are combined with the dynamics of social graphs driven by selectivity[44,45]. We incorporate selectivity into our model by introducing parameter $\gamma \in [0,1]$ (*selectivity rate*), which operates as follows. At each time point τ, we select an agent i (with opinion $o_i(t)$) and one of their neighbors j (having opinion $o_j(t)$) at random. If their opinions are not too distant ($|o_i(t) - o_i(t)| \leq \Delta o$), then they follow the standard opinion dynamics protocol: agent i changes their opinion in accordance with the transition matrix. Otherwise (if $|o_i(t) - o_i(t)| > \Delta o$), with probability γ tie (i,j) is deleted and a new tie appears: agent i creates a connection with a random (not neighboring) vertex k whose opinion lies within interval $[o_i(t) - \Delta o, o_i(t) + \Delta o]$. If there are no such vertices (it may be, for example, if the focal agent is the last representative of the opinion camp), then tie (i,j) does not disappear (nothing happens). In the contrary case (with probability $1-\gamma$), the standard opinion dynamics protocol is implemented. Thus, at each iteration, either an opinion changes, or the network evolves, or nothing happens. Note that the number of ties in the evolving social network remains the same.

Personalization systems are a signature of the contemporary social networking services. Due to our inability to capture all the information from these sites that we would like to obtain, personalization algorithms sort the online content and suggest to us those that will be most important (to us), from the algorithms' point of view, of course[46]. Personalization algorithms may significantly influence the macroscopic behavior of opinion dynamics processes because they control information flows between individuals. Relatively recently, scholars began to combine ideas of personalization algorithms with opinion formation models[36,47]. Their results indicate that personalization may both amplify and reduce polarization, subject to the underlying opinion formation model. Further, scholars argue that personalization algorithms may amplify the formation of echo chambers and, thus, increase the level of homophily. Besides, it is unlikely that the opinion dynamics presented in Dataset were not affected by personalization algorithms. On this basis, we need to incorporate personalization in our model. For our purposes, we employ one of the simplest approaches whereby communications between individuals may be declined. More precisely, selected agents i and j do not communicate (and the system goes to the next time step) with probability δ (*personalization rate*) if their opinions differ for more than Δo.

Implementation details. We implement assortativity and personalization into the model only in the cases of the triple and tenfold opinion spaces. Importantly, we combine them: if selectivity and personalization rates are positive at once, then agents can both change their connections and be affected by the personalization algorithm, as occurs in the following fashion. First, the personalization algorithm checks whether $|o_i(\tau) - o_i(\tau)|$ is greater than Δo or not. If the latter is true, then communication between agents is allowed, and i changes their opinion as usual. Otherwise (if opinions are too distant), the personalization algorithm activates and prohibits the communication with probability δ, and nothing happens (the system goes to the next time step). In contrast, with probability $1 - \delta$ communication between agents i and j is allowed. This permitted communication can go in two different ways. In the first one (that occurs with probability γ) agent i decides to renew their social environment by replacing tie (i, j) because it makes agent i uncomfortable. The second direction implies that agent i accepts influence from agent j and follows the standard opinion dynamics protocol (whatever it leads to). Note that if we set $\gamma = 0$ and $\delta = 0$ in the resulting model (Model 2), we return to the previously elaborated model (Model 1).

We employ the idea of noise in the transition matrix only for the binary opinion space because it requires (under the symmetrical assumption on noise) only two parameters to be used. Hence, we do not investigate how the data noisiness affects polarization patterns and concentrate only on the behavior of the assortativity coefficient. Model 2 is investigated only under the empirically calibrated transition matrix derived for triple and tenfold opinion spaces. This approach gives us the opportunity to analyze the effects of selectivity and personalization factors on the dynamics of homophily and polarization. In Model 2, we use threshold $\Delta o = 0$ if $m = 3$ (which is applied to opinion values $x_1 = 0, x_2 = 1, x_3 = 2$) and threshold $\Delta o = 3$ if $m = 10$ (which is applied to opinion values $x_1 = 0, \ldots, x_{10} = 9$).

We recognize that the implementation of selectivity and personalization affects not only the opinion dynamics itself but also influences the transition matrix that we observe and estimate from the side (for example, transition matrices estimated from Dataset). From this perspective, the most faithful approach would be to use the (ideal) transition matrix, which, combined with selectivity and personalization, will produce (in simulations) social dynamics by estimating which we will come to that transition matrix, as estimated from Dataset. However, for the sake of simplicity, we build our analysis upon the estimated transition matrix, assuming that this matrix does not depend on selectivity and personalization factors.

Data noisiness. The ideal-typical behavior of the system under the presence of noise in the transition matrix remains the same. Thus, we concentrate on how α and β affect the limiting behavior of the assortativity coefficient.

Our analysis reveals that for a given network topology, the maximal value of assortativity depends positively on both β and α (see Figure A2 in Appendix). This result is intuitively clear. On the one hand, by increasing α, we reduce the likelihood that like-minded agents will have different opinions after an interaction. On the other hand, higher values of β amplify the probability of opinion adoption; thus, neighboring agents are more likely to espouse similar positions after interaction. The minimal disturbance (in the Euclidean metric) we should make with the transition matrix to achieve the acceptable value of the assortativity coefficient is $\alpha = 0.02$ and $\beta = 0$ (in the case of highly clustered networks). The resulting transition matrix

$$P_{1,:,:} = \begin{bmatrix} 0.995 & 0.005 \\ 0.952 & 0.048 \end{bmatrix}, P_{2,:,:} = \begin{bmatrix} 0.066 & 0.934 \\ 0.029 & 0.931 \end{bmatrix}$$

will be able to produce a sufficiently homophilic social system. However, this estimation does not work for low-clustered networks, for which one should take a more disturbed transition matrix (by establishing higher values of α and β).

Macroscopic behavior of Model 2. The presence of selectivity and personalization does not alter the qualitative behavior of the system (such behavior is inherited from Model 1). Besides, we have observed no situations when the social network becomes disconnected. Nonetheless, we notice that selectivity and personalization affect the limiting values of the macroscopic metrics (see Figure 4). Recall that our purpose is to determine the combination of γ and δ that would increase the assortativity coefficient up to 0.14 at some point. Figure 4 indicates that higher selectivity rate values lead to more homophilic systems, as expected. However, personalization has the opposite effect on assortativity. Note that for large personalization rate values, the effect of selectivity is reduced. In limiting case $\delta \to 1$, we obtain the system characterized to be the zero level of assortativity regardless of the selectivity rate. In this case, only sufficiently similar opinions can interact; thus, all they can do is make antagonistic interactions. The most homophilic system is obtained if $\gamma = 1$ and there is no personalization; then, we get $C \approx 0.5$. Interestingly, the effect of topology observed for Model 1 (more clustered networks produce more homophilic systems) disappears when we increase the selectivity or personalization rates (see Figure A3 in Appendix). Both selectivity and personalization have a positive effect on the system's polarization. However, in contrast to the assortativity coefficient, the dissimilarity coefficient varies in relatively small intervals $\approx [0.4, 0.51]$. As for assortativity, the effect of selectivity on polarization varies with the level of personalization: large values of the personalization rate discount the selectivity factor. In limiting case $\delta \to 1$, the dissimilarity coefficient fluctuates around the value $D \approx 0.51$ and does not depend on the selectivity rate.

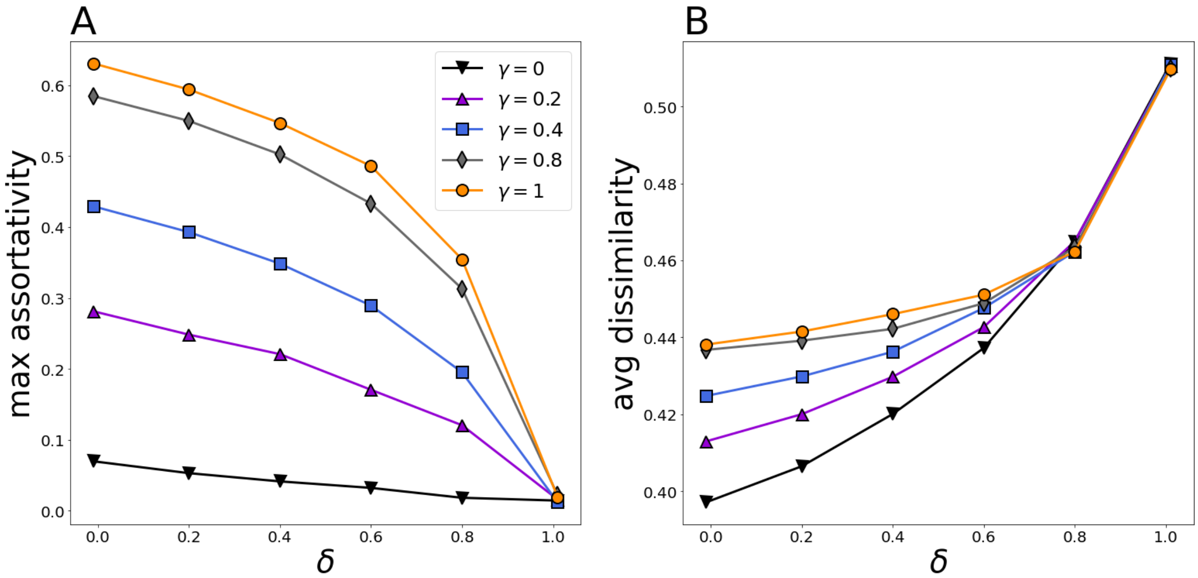

Figure 4. Maximal value of the assortativity coefficient and average (over time, starting from the moment the system had stabilized) value of the dissimilarity coefficient as functions of the personalization level, separated by the selectivity rate. The presented results are obtained on random geometric networks.

Based on these observations, we tune parameters γ and δ, attempting to get close to the empirical reference values in our simulations. The problem is that we need not only to achieve reference values but, what is important, do that *simultaneously*. We manage to resolve this problem in the triple opinion space by establishing $\gamma^* = 0.2, \delta^* = 0.1$. We begin again from opinion distribution $y_1(0) = 0.07, y_2(0) = 0.79, y_3(0) = 0.14$. In this case, the simulation system may reach the desired reference values at one point (see Figure 5). Furthermore, the local behavior of the macroscopic metrics near the reference values largely coincides with what we have noticed in the empirical data: the assortativity and dissimilarity increase, as does the fraction of individuals with the left-side opinions, whereas populations of individuals holding other (particularly middle-side) positions decrease. The asymptotic values of public opinion variables in this case are given by $y_1^* \approx 0.309, y_2^* \approx 0.573, y_3^* \approx 0.118$. They slightly differ from those obtained without selectivity and personalization. We do not state that there are no other combinations of γ, δ, and configurations of initial opinions, which could also generate a system that is consistent with the empirics. For example, a little level of asynchrony is observed by selecting $\gamma = 0.2, \delta = 0$ or $\gamma = 0.2, \delta = 0.2$ (i.e., by varying the level of personalization). However, we find that changes in the level of assortativity will lead to an apparent discord (see Figures A4-A7 in Appendix for details).

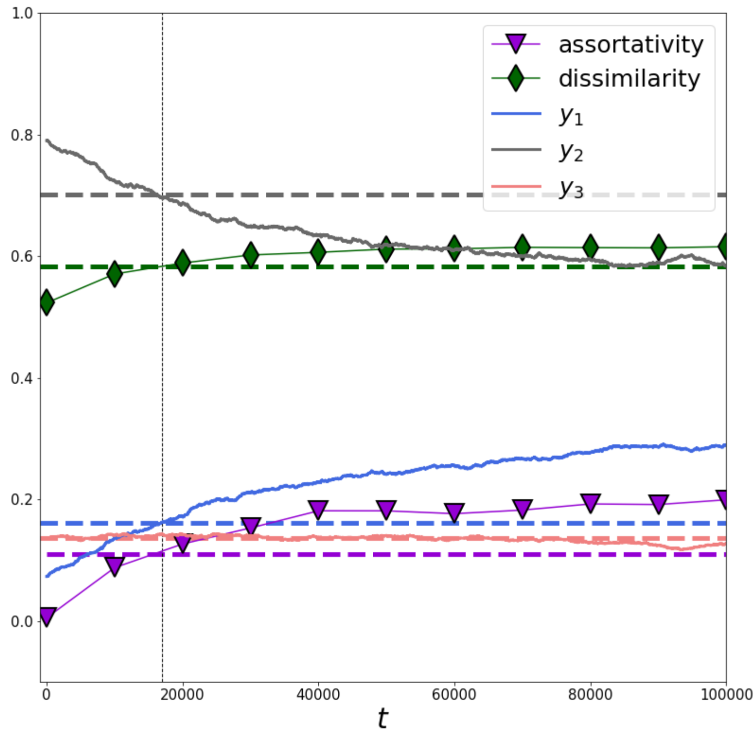

Figure 5. A simulation run for Model 2 if selectivity and personalization are $\gamma = 0.2, \delta = 0.1$ Initial opinions are drawn from distribution $y_1(0) = 0.07, y_2(0) = 0.79, y_3(0) = 0.14$. Dashed horizontal lines plot reference values from Table 1 (we take values from the first opinion snapshot). Common colors indicate the same metrics. The vertical dashed line marks the time point ($t \approx 17000$) at which all macroscopic metrics nearly coincide with the empirical reference values.

5. Discussion and future work

Selectivity and personalization, which we implemented into the model, sufficiently advanced it and made it more realistic. Hence, we managed to simulate an artificial society that demonstrates properties similar to those observed empirically at the macro scale (at the particular point). Our finding on the (possibly) most appropriate settings that could generate empirically acceptable systems may be employed in several ways. On the one hand, these settings may offer an opportunity to understand how strong personalization and particularly selectivity are in the referenced OSN, from which the empirical data were gathered. More precisely, our results indicate that without selectivity, our artificial systems cannot achieve the desired empirics under the assumption that the transition matrix is estimated correctly.

This knowledge can be further used in other studies in which this OSN is involved. However, this idea may be wrong in the case when the transition matrix is estimated with errors. Our current results do not answer the question of whether changes in the transition matrix could lead to *full* coincidence between simulations and empirics, but they at least hint that they *could do*. This problem requires additional analysis.

Next, knowing the current state of the system, we can predict its future evolution at the macro scale. However, only short-term predictions are meaningful because the transition matrix likely changes during long time intervals, reflecting events that occur both in this very OSN or beyond it. To be more specific, we calculate the transition matrix employing the second and third opinion snapshots from Dataset (recall that previous matrices were computed using the first two ones). We obtain the transition matrix

$$P_{1,:,:} = \begin{bmatrix} 0.976 & 0.024 & 0 \\ 0.967 & 0.033 & 0.000 \\ 0.94 & 0.058 & 0.002 \end{bmatrix}, P_{2,:,:} = \begin{bmatrix} 0.027 & 0.966 & 0.006 \\ 0.016 & 0.977 & 0.008 \\ 0.014 & 0.96 & 0.026 \end{bmatrix}, P_{3,:,:} = \begin{bmatrix} 0.002 & 0.065 & 0.933 \\ 0.001 & 0.048 & 0.951 \\ 0 & 0.036 & 0.963 \end{bmatrix}, \qquad (5)$$

which slightly differs from its previous version (3). This matrix, implemented in Model 2 (with parameters $\gamma^* = 0.2, \delta^* = 0.1$ tuned above), marks new the equilibrium point $y_1^* \approx 0.338, y_2^* \approx 0.534, y_3^* \approx 0.128$, which is a more polarized opinion configuration compared to previous one ($y_1^* = 0.309, y_2^* = 0.573, y_3^* = 0.118$) obtained under transition matrix (3). From this perspective, the opinion dynamics of our model can be understood in terms of the evolution of the transition matrix. Continuing this idea, opinion distribution $y_1(0) = 0.07, y_2(0) = 0.79, y_3(0) = 0.14$, used above as a starting point in simulations, may be understood as a state, to which the empirical system came under a different transition matrix. After that, the transition matrix was transforming as a response to some influential events and consistently took forms (3) and (5).

The model can be further advanced in several directions. First, it would be interesting to add in the model stubborn agents representing mass media or partisans. In its current form, the model cannot simulate a social network in which a few constantly stubborn agents are embedded. Further, one may inspect behavior of the system for a broader set of transition matrices focusing possibly on analytically accessible low-dimensional cases. A third possible direction concerns empirically analyzing the transition matrix's dynamics.

6. Data availability

All of the data, codes, and other support information can be found at https://dataverse.harvard.edu/privateurl.xhtml?token=a176d0d0-04c7-4608-b028-89be7dbd242a (Online Supplementary Materials). If the paper will be accepted for publication, then the online dataset will be published and this private URL will be replaced with the dataset doi.

7. Author Contributions

I.K. conceived and designed the research, analyzed the results, and wrote and revised the manuscript.

8. Competing interests

The author declares no competing interests.

9. Literature

1. Castellano, C., Fortunato, S. & Loreto, V. Statistical physics of social dynamics. *Rev. Mod. Phys.* **81**, 591–646 (2009).

2. Flache, A. *et al.* Models of Social Influence: Towards the Next Frontiers. *J. Artif. Soc. Soc. Simul.* **20**, (2017).
3. Proskurnikov, A. V. & Tempo, R. A tutorial on modeling and analysis of dynamic social networks. Part I. *Annu. Rev. Control* **43**, 65–79 (2017).
4. Proskurnikov, A. V. & Tempo, R. A tutorial on modeling and analysis of dynamic social networks. Part II. *Annu. Rev. Control* (2018).
5. Mäs, M. Challenges to Simulation Validation in the Social Sciences. A Critical Rationalist Perspective. in *Computer Simulation Validation: Fundamental Concepts, Methodological Frameworks, and Philosophical Perspectives* (eds. Beisbart, C. & Saam, N. J.) 857–879 (Springer International Publishing, 2019). doi:10.1007/978-3-319-70766-2_35.
6. Conover, M. D., Gonçalves, B., Ratkiewicz, J., Flammini, A. & Menczer, F. Predicting the political alignment of twitter users. in 192–199 (IEEE, 2011).
7. Barberá, P. Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data. *Polit. Anal.* **23**, 76–91 (2014).
8. Newman, M. Network structure from rich but noisy data. *Nat. Phys.* **14**, 542 (2018).
9. Ravazzi, C., Dabbene, F., Lagoa, C. & Proskurnikov, A. V. Learning Hidden Influences in Large-Scale Dynamical Social Networks: A Data-Driven Sparsity-Based Approach, in Memory of Roberto Tempo. *IEEE Control Syst. Mag.* **41**, 61–103 (2021).
10. Rashevsky, N. Studies in mathematical theory of human relations. *Psychometrika* **4**, 221–239 (1939).
11. Abelson, R. P. Mathematical models of the distribution of attitudes under controversy. *Contrib. Math. Psychol.* **14**, 1–160 (1964).
12. DeGroot, M. H. Reaching a consensus. *J. Am. Stat. Assoc.* **69**, 118–121 (1974).
13. Friedkin, N. E. & Johnsen, E. C. Social influence and opinions. *J. Math. Sociol.* **15**, 193–206 (1990).
14. Clifford, P. & Sudbury, A. A model for spatial conflict. *Biometrika* **60**, 581–588 (1973).
15. Sznajd-Weron, K. & Sznajd, J. Opinion evolution in closed community. *Int. J. Mod. Phys. C* **11**, 1157–1165 (2000).
16. Parsegov, S. E., Proskurnikov, A. V., Tempo, R. & Friedkin, N. E. Novel multidimensional models of opinion dynamics in social networks. *IEEE Trans. Autom. Control* **62**, 2270–2285 (2017).
17. Pham, T. M., Alexander, A. C., Korbel, J., Hanel, R. & Thurner, S. Balance and fragmentation in societies with homophily and social balance. *Sci. Rep.* **11**, 17188 (2021).
18. Takács, K., Flache, A. & Mäs, M. Discrepancy and Disliking Do Not Induce Negative Opinion Shifts. *PLOS ONE* **11**, e0157948 (2016).
19. Deffuant, G., Neau, D., Amblard, F. & Weisbuch, G. Mixing beliefs among interacting agents. *Adv. Complex Syst.* **3**, 87–98 (2000).
20. Hegselmann, R. & Krause, U. Opinion dynamics and bounded confidence models, analysis, and simulation. *J. Artif. Soc. Soc. Simul.* **5**, (2002).
21. Macy, M. W., Kitts, J. A., Flache, A. & Benard, S. Polarization in dynamic networks: A Hopfield model of emergent structure. (2003).
22. Altafini, C. Consensus problems on networks with antagonistic interactions. *IEEE Trans. Autom. Control* **58**, 935–946 (2013).
23. Banisch, S. & Olbrich, E. Opinion polarization by learning from social feedback. *J. Math. Sociol.* **43**, 76–103 (2019).
24. Prasetya, H. A. & Murata, T. A model of opinion and propagation structure polarization in social media. *Comput. Soc. Netw.* **7**, 2 (2020).
25. Friedkin, N. E. Choice Shift and Group Polarization. *Am. Sociol. Rev.* **64**, 856–875 (1999).
26. Mäs, M. & Flache, A. Differentiation without distancing. Explaining bi-polarization of opinions without negative influence. *PloS One* **8**, e74516 (2013).

27. Banisch, S. & Olbrich, E. An Argument Communication Model of Polarization and Ideological Alignment. *J. Artif. Soc. Soc. Simul.* **24**, 1 (2021).
28. Törnberg, P., Andersson, C., Lindgren, K. & Banisch, S. Modeling the emergence of affective polarization in the social media society. *PLOS ONE* **16**, e0258259 (2021).
29. Condie, S. A. & Condie, C. M. Stochastic events can explain sustained clustering and polarisation of opinions in social networks. *Sci. Rep.* **11**, 1355 (2021).
30. Adams, J., White, G. & Araujo, R. The Role of Mistrust in the Modelling of Opinion Adoption. *J. Artif. Soc. Soc. Simul.* **24**, 4 (2021).
31. Kozitsin, I. V. Formal models of opinion formation and their application to real data: evidence from online social networks. *J. Math. Sociol.* **0**, 1–28 (2020).
32. Kozitsin, I. V. Opinion dynamics of online social network users: a micro-level analysis. *J. Math. Sociol.* **0**, 1–41 (2021).
33. Krueger, T., Szwabiński, J. & Weron, T. Conformity, Anticonformity and Polarization of Opinions: Insights from a Mathematical Model of Opinion Dynamics. *Entropy* **19**, 371 (2017).
34. Newman, M. E. J. Mixing patterns in networks. *Phys. Rev. E* **67**, 026126 (2003).
35. FLACHE, A. & MACY, M. W. Small Worlds and Cultural Polarization. *J. Math. Sociol.* **35**, 146–176 (2011).
36. Perra, N. & Rocha, L. E. Modelling opinion dynamics in the age of algorithmic personalisation. *Sci. Rep.* **9**, 1–11 (2019).
37. Stern, S. & Livan, G. The impact of noise and topology on opinion dynamics in social networks. *R. Soc. Open Sci.* **8**, (2021).
38. Mäs, M., Flache, A. & Helbing, D. Individualization as driving force of clustering phenomena in humans. *PLoS Comput. Biol.* **6**, (2010).
39. Bond, R. M. *et al.* A 61-million-person experiment in social influence and political mobilization. *Nature* **489**, 295 (2012).
40. Neubaum, G., Cargnino, M., Winter, S. & Dvir-Gvirsman, S. “You’re still worth it”: The moral and relational context of politically motivated unfriending decisions in online networks. *PLOS ONE* **16**, e0243049 (2021).
41. Sasahara, K. *et al.* Social influence and unfollowing accelerate the emergence of echo chambers. *J. Comput. Soc. Sci.* **4**, 381–402 (2021).
42. Lewis, K., Gonzalez, M. & Kaufman, J. Social selection and peer influence in an online social network. *Proc. Natl. Acad. Sci.* **109**, 68–72 (2012).
43. Holme, P. & Newman, M. E. Nonequilibrium phase transition in the coevolution of networks and opinions. *Phys. Rev. E* **74**, 056108 (2006).
44. Frasca, P., Tarbouriech, S. & Zaccarian, L. Hybrid models of opinion dynamics with opinion-dependent connectivity. *Automatica* **100**, 153–161 (2019).
45. Steglich, C., Snijders, T. A. B. & Pearson, M. Dynamic Networks and Behavior: Separating Selection from Influence. *Sociol. Methodol.* **40**, 329–393 (2010).
46. Bakshy, E., Messing, S. & Adamic, L. A. Exposure to ideologically diverse news and opinion on Facebook. *Science* **348**, 1130–1132 (2015).
47. Maes, M. & Bischofberger, L. Will the Personalization of Online Social Networks Foster Opinion Polarization? *Available SSRN 2553436* (2015).

Legends

Figure 1. The violet circle marks the equilibrium point. The black solid line plots $y_1 + y_2 = 1$. We are interested in the area $y_1 + y_2 \leq 1$ beneath this line. The phase portrait of the system demonstrates that the equilibrium point is an asymptotically stable nodal sink.

Table 1
Values of metrics **M1**–**M3** drawn from Dataset

Figure 2. Ideal-typical evolution of public opinion variables in the model in the case of binary (panel A, initial opinions are drawn from the uniform distribution) and triple (panel B, initial opinions are equal to x_3) opinion spaces. Dashed lines plot theoretically predicted equilibrium values. (Panel C). Ideal-typical behavior of the assortativity coefficient in experiments ($m = 10$). The network topology with no clustering (Erdős–Rényi) exhibits lower assortativity values than clustered ones (random geometric and WS1). (Panel D). The panel plots how the maximal value of assortativity during a single experiment varies with the network topology ($m = 10$). (Panel E). Ideal-typical dynamics of the polarization coefficient for different network topologies ($m = 10$). (Panel F). The average (over time) value of the dissimilarity coefficient after the system has been stabilized ($t \geq 250{,}000$) during a single experiment as a function of the network topology ($m = 10$).

Figure 3. Ideal-typical behavior of the system in the case of the triple opinion space if initial opinions are drawn from the distribution $y_1(0) = 0.07, y_2(0) = 0.79, y_3(0) = 0.14$. Dashed horizontal lines plot reference values from Table 1 (we take values from the first opinion snapshot). Common colors indicate the same metrics. The vertical dashed line marks the point of time at which all macroscopic metrics except **M2** tend to coincide with the empirical reference values. One can observe a sharp discord in assortativity values between empirics (violet dashed line) and simulations (violet curve with triangle markers).

Figure 4. Maximal value of the assortativity coefficient and average (over time, starting from the moment the system had stabilized) value of the dissimilarity coefficient as functions of the personalization level, separated by the selectivity rate. The presented results are obtained on random geometric networks.

Figure 5. A simulation run for Model 2 if selectivity and personalization are $\gamma = 0.2, \delta = 0.1$ Initial opinions are drawn from distribution $y_1(0) = 0.07, y_2(0) = 0.79, y_3(0) = 0.14$. Dashed horizontal lines plot reference values from Table 1 (we take values from the first opinion snapshot). Common colors indicate the same metrics. The vertical dashed line marks the time point ($t \approx 17000$) at which all macroscopic metrics nearly coincide with the empirical reference values.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SI.docx