On discord in the voter model for complex networks: Supplementary material

Antoine Vendeville* and Shi Zhou

Department of Computer Science, University College London, UK

Benjamin Guedj

Department of Computer Science, University College London, UK and

Inria, France

(Dated: December 18, 2023)

I. UNICITY OF A SOLUTION TO EQ. 16

We demonstrate that the spectral radius of V is strictly less than 1, which implies that Eq. 16 has a unique solution. Let us first introduce a technical lemma from [1].

Lemma 1 (Azimzadeh, 2018, Lemma 2.1 [1]). *Let A be the adjacency matrix of a graph $\mathcal{F}$ so that a_{ij} is the weight of the edge $j \to i$. The spectral radius of A is strictly less than 1 if and only if for every row i, one of the following holds:*

- *row i sums to strictly less than 1, or*

- *there is a path $k \to \ldots \to i$ in $\mathcal{F}$ and row k sums to strictly less than 1.*

The matrix V can be seen as the adjacency matrix of a new graph $\mathcal{F}$. Its nodes correspond to agent pairs, and there is an edge from node $i'j'$ to node ij if and only if one of i' or j' is a leader of i or j in the original agent graph $\mathcal{G}$. Let ν_{ij} denote the sum of the row of V that corresponds to node ij. We have

$$\nu_{ij} = \frac{1}{2}\left(\sum_{k \in \mathcal{L}_i} w_{ik} + \sum_{k \in \mathcal{L}_j} w_{jk}\right). \tag{1}$$

Lemma 1 tells us that it suffices to prove, for every ij: either $\nu_{ij} < 1$, or there exists another node $i'j'$ with $\nu_{i'j'} < 1$ and a path from $i'j'$ to ij in $\mathcal{F}$. If $\nu_{ij} = 1$, assuming every agent can be influenced by a zealot, there exists an agent k such that $z_k^s > 0$ for some s and a path from k to i. Hence there is path from ik to ij in $\mathcal{F}$, and $\nu_{ik} < 1$ as shown by Eq. 2 in the main text.

II. CONDITIONS FOR INDEPENDENCE

We now prove that the opinions of i and j are independent if one of the following holds:

1. σ_i or σ_j is constant, or

2. there is no path from i to j nor from j to i, and i and j have no common ancestor.

---
* Corresponding author: a.vendeville@ucl.ac.uk

The first comes from the fact that a constant is independent from any other random variable. It applies in the case one of i and j can be reached by only one zealot, and thus never change opinion. For the second, recall that the opinion distribution of agent i evolves according to

$$\frac{dx_i^s}{dt} = \sum_{k \in \mathcal{L}_i} w_{ik} x_k^s + z_i^s - x_i^s, \quad s = 1, \dots, S. \tag{2}$$

Thus, this evolution is determined by the connections between i and zealots, and the opinion distributions of the leaders of i. In turn, the opinion distribution of a leader k of i is function of the connections between k and zealots, and the opinion distributions of the leaders of k. Iterating this reasoning, eventually all ancestors of i—and only them—intervene. Thus, if j is an ancestor of i, the opinions of the two are not independent. Otherwise, if there is no path between them but they have a common ancestor, their opinions are both function of the opinion of that ancestor, meaning they are not independent. If neither i nor j is an ancestor of the other, and if they have no common ancestor, their opinions are two independent random processes.

III.   THE DATASETS

In Sect. V A we use four datasets, described in Table I. In each of them, a node belongs to a single community, given by the creators of the dataset. We only keep the largest weakly connected component of each network. The in-degree and out-degree distributions we obtain are presented in Fig. 1.

The datasets are initially unweighted. We first set w_{ij} to 1 if there is an edge from j to i, and zero otherwise. Once all such values are set, we define values of zealousness as follows: for any agent i who belongs to community s, we attribute a random uniform value between 0 and 1 to z_i^s, and set $z_i^r = 0$ for $r \neq s$. We then proceed to multiply each weight w_{ij} by

$$\frac{1 - z_i^s}{\sum_{k \in \mathcal{L}_i} w_{ik}}, \tag{3}$$

with s the community of user i, so that Eq. 2 holds. To highlight the absence of approximation in our results, we study the difference between empirical values of discord obtained in simulation, and theoretical ones from Eqs. 10 and 15. For the simulation we proceed as per Algorithm 1, with $T = 10^5 N$ steps and a burn time of $T_b = 10N$ steps. This means the first $10N$ steps are not taken into account when computing discord probabilities, in order to reduce the influence of the initial state.
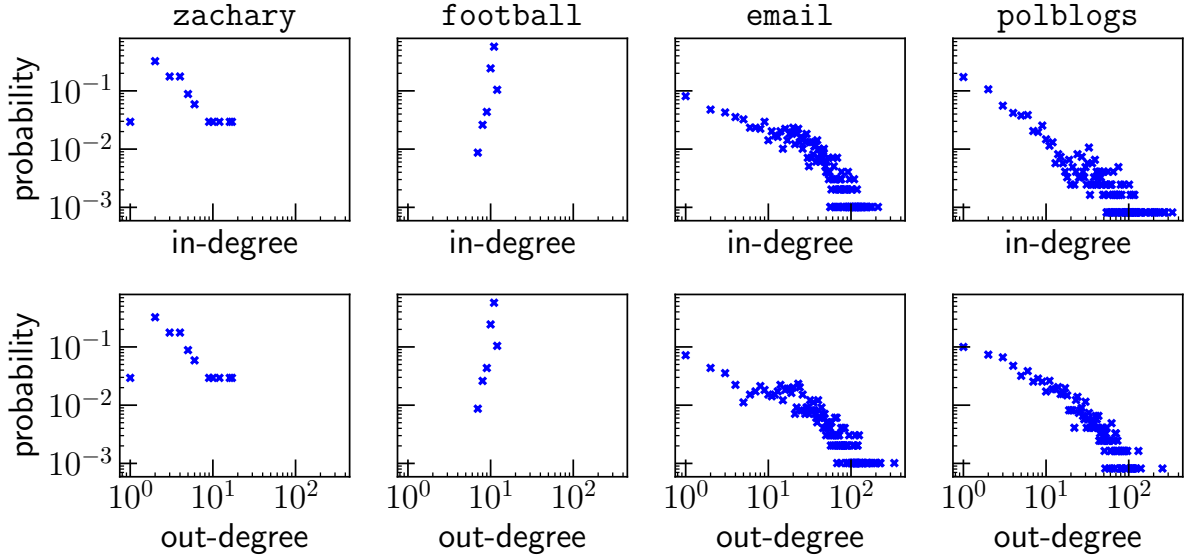
FIG. 1: Unweighted degree distributions for each toy dataset, logarithmic scale. `Zachary` and `Football` are undirected so both distributions are the same. **Top:** in-degrees. **Bottom:** out-degrees.

	Zachary [2]	Football [3]	Email [4]	Polblogs [5]
Number of agents	34	115	986	**1,222**
Number of edges	156	1,226	**25,552**	19,021
Density	**0.14**	0.09	0.03	0.01
Directed	No	No	Yes	Yes
Self-loops	No	No	Yes	No
Mean degree	4.59	10.66	**25.91**	15.57
Number of communities	2	12	**42**	2
Modularity	0.36	**0.53**	0.31	0.41
Average clustering	**0.57**	0.40	0.41	0.32
Independent agent pairs	0%	0%	4%	**40%**

TABLE I: Basic statistics for the datasets. Modularity and average clustering are computed on the undirected, unweighted networks. See Sect. IV B for a definition of independent agent pairs.

Simulated values are compared with their theoretical counterparts in Figure 2. For each pair (i, j) we compute the absolute and relative differences, respectively $|\rho_{ij}^{\text{theo}} - \rho_{ij}^{\text{simu}}|$ and $|\rho_{ij}^{\text{theo}} - \rho_{ij}^{\text{simu}}|/\rho_{ij}^{\text{theo}}$. We then average these values over all agent pairs. As the simulation time increases, we observe a tighter and tighter fit between the values, with relative differences

reaching between 10^{-3} and 10^{-2} for all four datasets—except for `polblogs`, with a difference slightly below 10^{-1}. This highlights the exactness of the equations we derived.

Algorithm 1: Estimation of discord probabilities via simulation

---

Data: user set $\mathcal{N}$, number of opinions S, adjacency matrix W, zealousness matrix Z,

number of steps T, burn time T_b

Result: estimated discord probabilities ρ_{ij} for all user pairs (i, j)

begin

$\quad x_i \sim \mathcal{U}_{\{1,\ldots,S\}}$ for all $i \in \mathcal{N}$ `// initialize users opinions at random`

$\quad \rho_{ij} \longleftarrow 0$ for all $i, j \in \mathcal{N}$ `// initialize probabilities to zero`

$\quad \tau_{ij} \longleftarrow T_b + 1$ for all $i, j \in \mathcal{N}$ `// time of last update for each pair`

$\quad$ **for** $1 \le t \le T$ **do**

$\quad\quad i \sim \mathcal{U}_{\{1,\ldots,N\}}$ `// select a random user`

$\quad\quad x_i^{\text{old}} \longleftarrow x_i$ `// store current opinion`

$\quad\quad x_i \longleftarrow x_j$ with probability w_{ij}, or s with probability z_i^s `// draw new opinion`

$\quad\quad$ **if** $t > T_b$ *and* $x_i \ne x_i^{old}$ **then**

$\quad\quad\quad$ `// if i changed opinion, for users j who were disagreeing with i`

$\quad\quad\quad\quad$ `until now, we add the duration of the disagreement to` ρ_{ij}

$\quad\quad\quad$ **for** $j \in \mathcal{N}\backslash\{i\}$ **do**

$\quad\quad\quad\quad$ **if** $x_j \ne x_i^{old}$ **then**

$\quad\quad\quad\quad\quad \rho_{ij} \longleftarrow \rho_{ij} + (t - \tau_{ij})$ `// update discord probability`

$\quad\quad\quad\quad\quad \tau_{ij} \longleftarrow t$ `// set time of last update to current time`

$\quad \rho_{ij} \longleftarrow \rho_{ij}/(T - T_b)$ for all $i, j \in \mathcal{N}$ `// normalize probabilities`

---

[1] P. Azimzadeh, Math. Comp. **88**, 783–800 (2018).

[2] W. W. Zachary, J. Anthropol. Res. **33**, 452 (1977).

[3] M. Girvan and M. E. Newman, Proc. Natl. Acad. Sci **99**, 7821 (2002).

[4] J. Leskovec, J. Kleinberg, and C. Faloutsos, ACM Trans. Knowl. Discov. Data **1**, 2–es (2007).

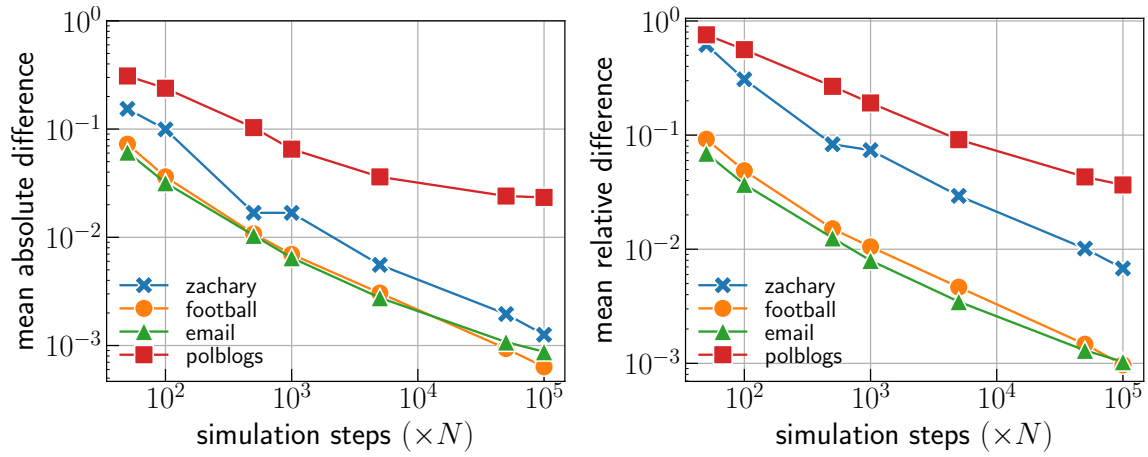FIG. 2: Difference between simulated and theoretical values of discord.

[5] L. A. Adamic and N. Glance, in *Proc. Int. Workshop Link Discov.*, LinkKDD '05 (Association for Computing Machinery, New York, NY, USA, 2005) p. 36–43.