



HAL
open science

A model for French voters

Antoine Vendeville

► **To cite this version:**

| Antoine Vendeville. A model for French voters. 2025. hal-04849938v2

HAL Id: hal-04849938

<https://hal.science/hal-04849938v2>

Preprint submitted on 24 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

A model for French voters

Antoine Vendeville^{1,2,3,*}

¹*médialab, Sciences Po, 75007 Paris, France*

²*Complex Systems Institute of Paris Île-de-France (ISC-PIF) CNRS, 75013 Paris, France*

³*Learning Planet Institute, Learning Transitions unit, CY Cergy Paris University, Paris, France*

(Dated: January 22, 2025)

Models of opinion dynamics describe how opinions are shaped in various environments. While these models are able to replicate macroscopical opinion distributions observed in real-world scenarios, their capacity to align with data at the microscopical level remains mostly untested. We evaluate the capacity of the multi-state voter model with zealots to capture individual opinions in a fine-grained Twitter dataset collected during the 2017 French Presidential elections. Our findings reveal a strong correspondence between individual opinion distributions in the equilibrium state of the model and ground-truth political leanings of the users. Additionally, we demonstrate that discord probabilities accurately identify pairs of like-minded users. These results emphasize the validity of the voter model in complex settings, and advocate for further empirical evaluations of opinion dynamics models at the microscopical level.

Introduction. Models of opinion dynamics provide potent tools to unveil mechanisms driving phenomena such as polarization [1–7] or echo chambers [8, 9]. These models usually comprise a network of agents endowed with opinions that evolve through interactions with one another. Social, psychological and technological features may govern these interactions—e.g. negative influence [3, 6, 7, 10], stubbornness [11–13] or recommendation algorithms [5, 14, 15]. Exploring the effect of such features and combinations thereof, can help us understand how ideological landscapes are shaped in various circumstances. This is particularly useful in the current context of raising concerns about the role of online social platforms in political debates and democratic processes [16–18].

The capacity of the models to replicate real-world observations has garnered increasing attention in recent years. Recurrent sources of data employed towards this goal include surveys [2, 19–21], election results [22–26], and interactions extracted from online social platforms [1, 4, 27–29]. The celebrated voter model has been the subject of several studies pertaining to this question. The seminal paper of Fernández-Gracia et al. [26] demonstrated the model’s ability to capture statistical features of votes shares in U.S. elections. These results were confirmed and enriched in subsequent works [22–25, 30]. We refer the interested reader to [31, 32] for in-depth reviews of the research about the empirical validity of opinion dynamics models.

These models are built to emulate global phenomena—e.g. polarization, echo chambers—and most do not pretend to truthfully describe individual processes of opinion formation. Therefore, research about the empirical validity of the models usually stands from a macroscopical point of view. For example, Valensise et al. [28] investigate the capacity of four different opinion dynamics models to reproduce global opinion distributions observed in various online social platforms. The capacity of the models to accurately identify opinions at the individual level remains mostly untested.

We evaluate the capacity of the multi-state voter model with zealots [11] to capture individual opinions in a fine-grained real-life dataset. We leverage a directed, weighted network of retweets between Twitter (now X) users, collected during the campaign of the 2017 French Presidential Election [33]. We uncover a strong correspondence between individual opinions in the equilibrium state of the model, and ground-truth party affiliations explicitly stated by the users in their publications and self-descriptions. Users are well separated along party lines in the opinion space of the model, and the model correctly identifies ground-truth party affiliations in 92.5% of cases. We also show that discord probabilities allow us to deduce with high accuracy whether or not two users support the same party. Neither the undirected or unweighted counterparts of the retweet network, nor the follow and mention networks produce comparable results. Our findings highlight the necessity for a fine-grained modelling approach, and contribute to the growing literature on the empirical validity of opinion dynamics models. Details of the computations, networks statistics, and additional figures can be found in the Supplemental Material [34].

Modeling. Fraasier et al. [33] collected retweets and mentions between 22,853 Twitter accounts in the last six months leading to the 2017 French Presidential Election. Follow relationships between the accounts were collected in another study [35]. The accounts were selected based on the presence of political keywords in their tweets, and include about 2,000 non-individual political entities, e.g. official party accounts, activist groups, etc. On the basis of the content of their tweets and profile description, accounts were manually labelled by the creators of the dataset to indicate the party they support in the election: FI (France Insoumise, far-left), PS (Parti Socialiste, traditional left), EM (En Marche, center), LR (Les Républicains, traditional right), FN (Front National, far-right). These were the five main competing parties in the elections. We discard unlabelled accounts and accounts with multiple labels (resp. 3,401 and 803 accounts).

The multi-state voter model with zealots for complex networks unfolds as follows [11]. Consider a directed, weighted network of users labelled $1, \dots, N$ with weight w_{ij} on the edge $j \rightarrow i$. Users are initially endowed with a random opin-

* antoine.vendeville@sciencespo.fr

ion chosen in a finite discrete set \mathcal{S} . In iterated steps, a user i is chosen at random and adopts the opinion of another user chosen at random according to the probabilities $\{w_{ij}\}_{j=1,\dots,N}$ (edge weights are normalized so that each user has in-degree one). Let us suppose that users who never change opinion, called zealots [11, 13, 23, 36], are present in the system. Let \mathcal{N} denote the set of non-zealot users, and z_i^s the aggregated weight of edges from zealots with opinion s toward user $i \in \mathcal{N}$. Under mild topological conditions, the system converges towards a unique state of equilibrium, wherein the opinion of i fluctuates according to an individual probability distribution x_i such that, for any $s \in \mathcal{S}$ [11],

$$x_i^s = \sum_{j \in \mathcal{N}} w_{ij} x_j^s + z_i^s. \quad (1)$$

We apply this model on the #Elysée2017fr dataset. The opinion space is formed of the different parties that appear in the data: $\mathcal{S} = \{FI, PS, EM, LR, FN\}$. Each account is endowed with a ground-truth opinion $y_i \in \mathcal{S}$, indicating the party they support as identified by the creators of the dataset. We call supporters of party s all users with ground-truth opinion $y_i = s$. Retweet interactions induce a directed, weighted network so that the weight w_{ij} of the edge from j to i is defined as the normalized number of times that i has retweeted j . If this number is positive, we say that j is a leader of i . We let \mathcal{L}_i denote the set of leaders of i . We treat accounts of political entities as zealots and fix their opinions to their ground-truth opinions, which correspond to the party they officially represent. The value z_i^s is thus the number of times that i has retweeted a political entity that represents party s . We compute individual equilibrium opinion distributions as per Eq. 1. The value of x_i^s can be interpreted as the likeliness attributed by the model to the possibility that i supports party s . Crucially, the ground-truth opinion y_i of i does not bear involvement in the computation of x_i^s . This justifies our analysis, where we compare x_i and y_i in several fashions to assess the validity of the model.

A central aspect of our methodology is to consider political entities as zealots. It seems natural to assume that party affiliations of political entities are immutable: the Parisian branch of FI for example, is and will always be—by definition—in support of FI. As these affiliations are public knowledge, it is not such a restrictive hypothesis to assume that they are known. This choice is also motivated by a technical reason: zealots guarantee the existence of a unique equilibrium state, at least for the regions of the network that they can reach [12, Thm. 2.1]. Otherwise, there may exist a multitude of different equilibrium states, the potential diversity of which entails a difficulty in their interpretation. Zealots also act as reference points, ensuring a unique correspondence between opinions in the model and ground-truth opinions. Hereafter we focus our analysis on non-zealot users, which we call users for the sake of simplicity.

Results. The model is able to identify supporters of each party with very high accuracy. The most likely opinion of i according to the model, i.e. $\operatorname{argmax}_{s \in \mathcal{S}} x_i^s$, matches the ground-truth y_i for 92.5% of users. Party-wise accuracy values are

shown in Fig. 1a (plain bars). FN supporters are the easiest to identify (accuracy 0.95), while PS supporters are the hardest (accuracy 0.81). The worse performance for PS is a pattern that will repeat throughout our analysis. It might be due to the dire situation of the PS at the time, which was considerably weakened after five years of difficult Hollande presidency, and its voting base torn between the rise of FI on the left and EM on the right.

Not only does the model attributes a higher probability to ground-truth opinions than to others, but the difference between the probabilities is not negligible. Let \mathcal{N}_s denote the set of supporters of party s . In Fig. 1 we compare the distributions of

$$x^s = \{x_i^s : i \in \mathcal{N}_s\}, \text{ and} \quad (2)$$

$$x^{-s} = \{x_i^t : i \in \mathcal{N}_s, t \neq s\}. \quad (3)$$

The two are concentrated as opposite sides of the unit interval. The former is located towards 1, with means ranging from $\langle x^s \rangle = 0.71$ for PS to $\langle x^s \rangle = 0.86$ for EM, while the latter is strongly skewed towards 0, with means ranging from $\langle x^{-s} \rangle = 0.03$ for EM to $\langle x^{-s} \rangle = 0.07$ for PS. Note that we find consistently low probabilities for x^{-s} (standard deviation ranging from 0.09 for FN supporters to 0.15 for PS), while the magnitude of x^s is much more varied (standard deviation between 0.16 for FN supporters and 0.31 for PS).

An interpretation of the striking difference between the distributions of x^s and x^{-s} , is that the supporters of each party occupy a specific region in the five-dimensional space that contains the x_i . Figures 1c shows that in this space, users are significantly closer to those who support the same party than to others. Furthest away are the supporters of EM and FN, and closest to one another are those of EM and PS. As an illustrative example, we show in Fig. 2 the bidimensional distributions of $(x_i^{\text{EM}}, x_i^{\text{PS}})$ for EM and PS supporters, and $(x_i^{\text{EM}}, x_i^{\text{FN}})$ for EM and FN supporters. In the latter case, the densities are concentrated at opposite sides of the unit square. The former case however presents a different picture, as supporters of PS populate a larger and more central area of the subspace (Fig. 2, top plots). To confirm these findings, we find the best separating hyperplane between supporters of the different parties, by fitting a basic Support Vector Classifier (SVC) to the (x_i, y_i) pairs. The result is very accurate, as the SVC correctly estimates the opinion of 93% of the users. Party-wise accuracy scores are shown in Fig. 1a (hatched bars).

In summary, through the individual equilibrium distributions x_i , the voter model embeds the users in a space where supporters of each parties are located in a specific region, separated from the others, and therefore clearly identifiable.

Discord. Individual equilibrium opinion distributions x_i describe the stabilization of the system at the user level. It is often useful to also describe stabilization at the level of edges, or more generally user pairs, via the notion of active links or discord probabilities [11, 36, 37]. The discord probability ρ_{ij} is the probability to find i and j holding two different opinions. Under the same conditions as for x_i , this quantity

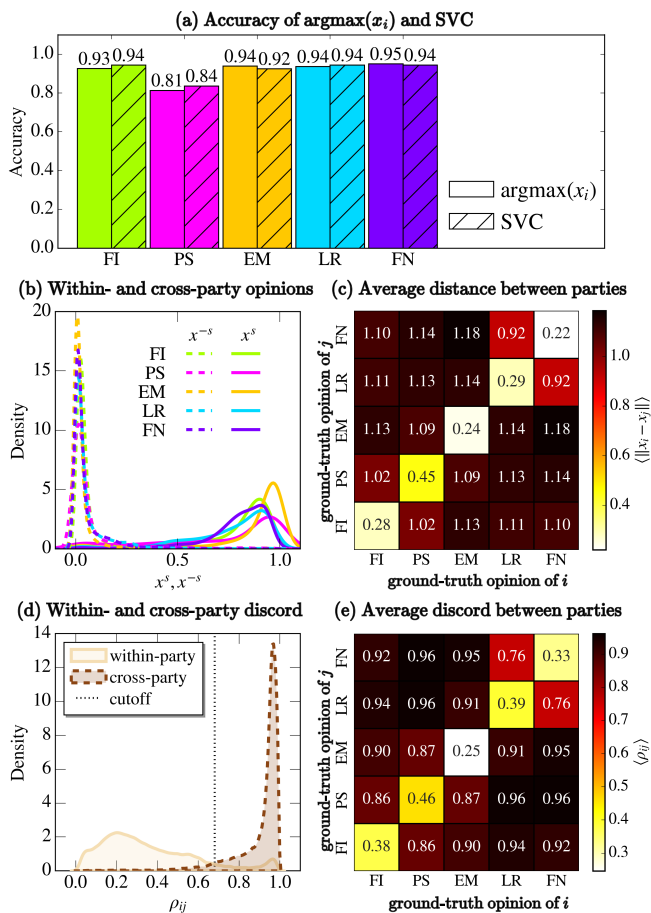


FIG. 1: Correspondence between theoretical opinion distribution x_i and ground-truth opinions y_i . **(a)** Party-wise accuracy of $\text{argmax}(x_i)$ (plain bars) and SVC (hatched bars) for the retrieval of ground-truth opinions y_i from x_i . **(b)** Distribution of within-party (x^s) and cross-party (x^{-s}) opinions among supporters of each party s , distinguished by color. **(c)** Average euclidean distance between supporters of the same party (diagonal cells) and supporters of different parties (off-diagonal cells). **(d)** Distribution of discord probabilities ρ_{ij} between supporters of the same party (blue) and supporters of different parties (red, dashed). The vertical dotted line indicates the cutoff of the logistic regression model: if ρ_{ij} is on the left of it, the logistic regression predicts that i and j support the same party, otherwise it predicts that they support different parties. **(e)** Average discord probability between supporters of the same party (diagonal cells) and supporters of different parties (off-diagonal cells). Smoothed distributions are computed via kernel density estimation.

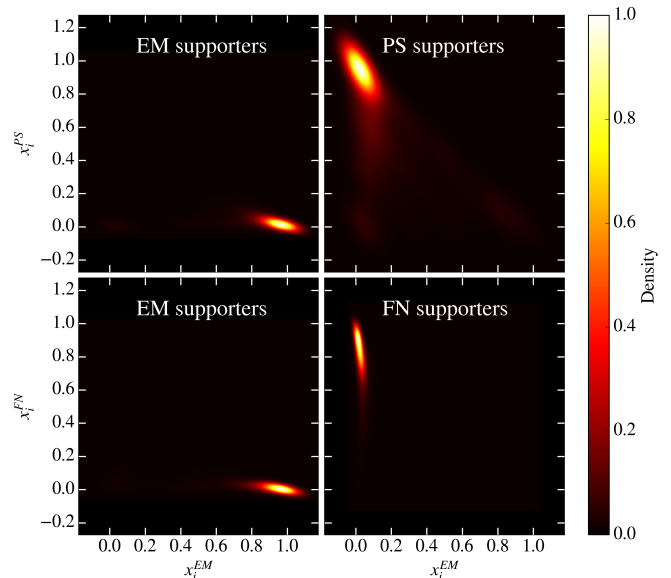


FIG. 2: **(Top)** Distributions of theoretical opinions towards EM and PS, among EM supporters (left) and PS supporters (right). **(Bottom)** Distributions of theoretical opinions towards EM and FN, among EM supporters (left) and FN supporters (right).

converges towards an equilibrium state

$$\rho_{ij} = \frac{1}{2} \left[\sum_{k \in \mathcal{L}_i} w_{ik} \rho_{jk} + \sum_{k \in \mathcal{L}_j} w_{jk} \rho_{ik} + \sum_{s \in \mathcal{S}} z_i^s (1 - x_j^s) + \sum_{s \in \mathcal{S}} z_j^s (1 - x_i^s) \right]. \quad (4)$$

Certain conditions of independence between the opinions of i and j allow to use simply $\rho_{ij} = \sum_{s \in \mathcal{S}} x_i^s (1 - x_j^s)$. The derivation of these formulas is detailed in ref. [11]. As long as the existence and unicity of the x_i vectors are guaranteed, the same goes for the ρ_{ij} values. The value ρ_{ij} can be interpreted as the likeliness attributed by the model to the possibility that i and j do not support the same party.

Averages $\langle \rho_{ij} \rangle$ within and between parties are shown in Fig. 1e. EM supporters exhibit the lowest within-party discord, while PS exhibits the highest. The lowest cross-party discord is observed between LR and FN, the highest between PS and both LR, FN. In Fig. 1d, we plot separately the within-party and cross-party distributions of ρ_{ij} . The two differ greatly, the latter being located towards 1 and the former towards 0. Users with the same ground-truth opinion are much more likely to hold the same opinion in the model than those who do not. In fact, a simple logistic regression model is able to discriminate between within-party and cross-party pairs, based on the values ρ_{ij} , with 93% accuracy. This highlights the capacity of the model to reliably distinguish friends from foes. Because discord spans a narrower range cross-party than within-party, foes are easier to identify than friends (accuracy 93.4% versus 90.2%).

Estimating opinions directly from zealots. The value of x_i^s corresponds to the probability that a backward random walk initiated in i hits a zealot with opinion s before a zealot with any other opinion [12, Prop. 3.2]. Therefore zealots—here political entities—play an crucial role in the equilibrium state of the model. One may then wonder, are the encouraging results we just exposed a mere consequence of the fact that people are preferentially connected with political entities of the party they support? The question is particularly relevant, given that retweet networks are known to exhibit political homophily [38–40]. Let $z_i = (z_i^s)_{s \in \mathcal{S}}$ denote the vector containing the weight of all direct connections from zealots with opinion s toward user i . This vector describes the tendency of i to retweet political entities affiliated to the different parties. We expect the vectors z_i to provide good indicators of ground-truth opinions; but if they were to be as good or even better than x_i , one may question the relevance of using Eq. 1.

Because 27% of users do not have direct connections with zealots, one cannot infer anything from z_i for about a quarter of the considered population. Therefore the use of the vectors x_i is already beneficial in that it extends the realm of the analysis. Let us now focus on the 73% of users who exhibit direct connections with zealots. Via the z_i vectors, supporters of the same party are positioned closer to one another (average Euclidean distance $\langle \|x_i - x_j\| \rangle = 0.190$) and further from supporters of other parties ($\langle \|x_i - x_j\| \rangle = 1.339$) than they are via the x_i vectors ($\langle \|x_i - x_j\| \rangle = 0.219, 1.146$ respectively). However, we find that x_i is more reliable than z_i to identify ground-truth opinions. $\text{Argmax}(x_i)$ matches the ground-truth opinion of i for 96.2% of users (95.1% for $\text{argmax}(z_i)$) and a Support Vector Classifier fitted on the (x_i, y_i) pairs correctly categorizes 96.4% percent of the users (94.8% when fitted on the (z_i, y_i) pairs). Therefore, while direct connections with zealots separate more clearly between supporters of different parties, the voter model allows for a more accurate identification of individual opinions.

Comparison with other networks. Multiple types of user networks can be extracted from online social interactions. To put our results into perspective, we investigate whether different data collection and processing methods impact the empirical validity of the model. We consider five other networks: the unweighted (Unw.) counterpart of the retweet network, the undirected (Und.) counterpart of the retweet network, the undirected and unweighted counterpart of the retweet network (UU), the Follow network (where $w_{ij} = 1$ if i follows j and 0 otherwise), and finally the Mention network (where w_{ij} is the number of times that i mentioned j in a tweet). In Table I we show that the retweet network performs much better than the others according to all the metrics we have used in the above analysis. The poorer performances of the follow and mention networks are not too surprising. Indeed, while retweets are often considered to be markers of endorsement, people may follow a broader range of the political spectrum [38, 39], and mentions are often used to express hostility towards outgroup members [41, 42]. In fact, only 2% of FN supporters are correctly identified by $\text{argmax}(x_i)$ in the mention network. The unweighted counterpart of the retweet network exhibits slightly better performance than the undirected

	Retweet	Unw.	Und.	UU	Follow	Mention
Acc. $\text{argmax}(x_i)$	92.5	91.6	91.9	74.9	84.2	64.4
Acc. SVC	93.0	92.8	92.8	92.8	89.4	83.9
$\langle x^s \rangle$	0.781	0.672	0.636	0.507	0.471	0.337
$\langle x^{-s} \rangle$	0.055	0.082	0.091	0.123	0.132	0.166
$\langle \ x_i - x_j\ \rangle_{\text{within}}$	0.296	0.266	0.235	0.188	0.205	0.164
$\langle \ x_i - x_j\ \rangle_{\text{cross}}$	1.097	0.904	0.840	0.617	0.545	0.301
$\langle \rho_{ij} \rangle_{\text{within}}$	0.355	0.486	0.499	0.591	0.638	0.682
$\langle \rho_{ij} \rangle_{\text{cross}}$	0.902	0.852	0.835	0.779	0.795	0.730
Acc. logistic(ρ_{ij})	92.9	89.8	87.2	86.1	85.3	77.8

TABLE I: Comparison between the results of the model applied on the directed weighted retweet network, its unweighted counterpart (Unw.), its undirected counterpart (Und.), its undirected unweighted counterpart (UU), the follow network and the mention network. Subscripts precise whether the averages are within- or cross-party. $\langle \|x_i - x_j\| \rangle$ is the average Euclidean distance in the opinion space of the model for the considered user pairs, and $\langle \rho_{ij} \rangle$ the average discord probability. logistic(ρ_{ij}) denotes a logistic regression model trained to distinguish between supporters of the same party and supporters of different parties on the basis of the ρ_{ij} values. “Acc.” stands for accuracy, given in percentages.

counterpart: discarding edge weights appears to be less penalizing that discarding directionality. The combination of discarding both weights and directionality (UU network) affects the performance even further—only 6% of FN supporters are correctly identified by $\text{argmax}(x_i)$ in the UU network. This is an especially relevant finding, given that a large part of the literature on the voter model focuses on undirected unweighted networks.

Discussion. The multi-state voter model with zealots can accurately estimate individual opinions in a large, fine-grained, heterogeneous online population. The separation between users with different ground-truth opinions is clear in the opinion space of the model, and the model correctly identifies the ground-truth for 92.5% of the users. In addition, discord probabilities act as very good proxies to distinguish user pairs with the same ground-truth opinion from others. These findings are particularly surprising, as they result from a microscopical analysis while the voter model was originally suited for macroscopical analyses. Our study advocates for the pursuit of more theoretical works dedicated to directed, weighted, heterogeneous networks with more than two opinions.

The presented methodology is parameter-free and circumvents the thorny issue of parameter inference. Empirical applications of opinion dynamics models often rely on sweeps of the parameter space to find parameter values for which the model best reproduces the observed data. To fully demonstrate the empirical validity of the models, these analyses should ideally be complemented with evaluations of the models thus calibrated on different datasets. This is no easy task, given the difficulty to obtain fine-grained data, the heterogeneity of datasets, and the absence of universal inference methods for the parameters of these models.

Our results add to the existing body of evidence for the empirical validity of the voter model. Other datasets, and extensions such as the partisan voter model [43] or the nonlinear voter model [44], could be the ground of further testing. Because we analyzed the equilibrium state of the model, our findings pertain to the long-term behavior of the model. Future works should strive to integrate temporal aspects. Overall, we call for more research dedicated to the evaluation of

opinion dynamics models on the microscopical level.

ACKNOWLEDGMENTS

I thank Fernando Diaz-Diaz for his careful reading and precious advices on this manuscript. I also thank Guillaume Cabanac and Ophélie Fraïsier-Vannier for reuploading the *Elysée2017fr* dataset, Effrosyni Papanastasiou for sharing the follower network with me, and Lucía S. Ramirez for insightful discussions on the voter model.

-
- [1] A. F. Peralta, P. Ramaciotti, J. Kertész, and G. Iñiguez, *Phys. Rev. Res.* **6**, 013170 (2024).
- [2] J. Ojer, M. Starnini, and R. Pastor-Satorras, *Phys. Rev. Lett.* **130**, 207401 (2023).
- [3] T. M. Pham, J. Korbelt, R. Hanel, and S. Thurner, *Proceedings of the National Academy of Sciences* **119**, e2121103119 (2022).
- [4] F. Baumann, P. Lorenz-Spreen, I. M. Sokolov, and M. Starnini, *Phys. Rev. X* **11**, 011012 (2021).
- [5] A. F. Peralta, M. Neri, J. Kertész, and G. Iñiguez, *Phys. Rev. E* **104**, 044312 (2021).
- [6] R. Axelrod, J. J. Daymude, and S. Forrest, *Proc Natl Acad Sci USA* **118**, e2102139118 (2021).
- [7] M. W. Macy, M. Ma, D. R. Tabin, J. Gao, and B. K. Szymanski, *Proc Natl Acad Sci USA* **118**, e2102144118 (2021).
- [8] A. Vendeville and F. Diaz-Diaz, *Echo chamber effects in signed networks* (2024), 2406.17435 [physics].
- [9] A. Vendeville, A. Giovanidis, E. Papanastasiou, and B. Guedj, in *Complex Networks and Their Applications XI*, Vol. 1077, edited by H. Cherifi, R. N. Mantegna, L. M. Rocha, C. Cherifi, and S. Micciché (Springer International Publishing, Cham, 2023) pp. 74–85.
- [10] Y. Li, W. Chen, Y. Wang, and Z.-L. Zhang, *Internet Mathematics* **11**, 93 (2015).
- [11] A. Vendeville, S. Zhou, and B. Guedj, *Phys. Rev. E* **109**, 024312 (2024).
- [12] M. E. Yildiz, A. Ozdaglar, D. Acemoglu, A. Saberi, and A. Scaglione, *Trans. Econ. Comput.* **1**, 10.1145/2538508 (2013).
- [13] M. Mobilia, A. Petersen, and S. Redner, *J. Stat. Mech.* **2007**, P08029 (2007).
- [14] G. Iannelli, G. De Marzo, and C. Castellano, *Chaos: An Interdisciplinary Journal of Nonlinear Science* **32**, 043103 (2022).
- [15] P. Ramaciotti Morales and J.-P. Cointet, in *RecSys '21: 15th ACM Conference on Recommender Systems* (Amsterdam, Netherlands, 2021).
- [16] P. S. Vishnuprasad, G. Nogara, F. Cardoso, S. Cresci, S. Giordano, and L. Luceri, *ICWSM* **18**, 1557 (2024).
- [17] P. Lorenz-Spreen, L. Oswald, S. Lewandowsky, and R. Hertwig, *Nat Hum Behav* **7**, 74 (2023).
- [18] E. Kubin and C. Von Sikorski, *Annals of the International Communication Association* **45**, 188 (2021).
- [19] H. Pérez-Martínez, F. Bauzá Minguez, D. Soriano-Paños, J. Gómez-Gardeñes, and L. Floría, *Chaos, Solitons & Fractals* **175**, 113917 (2023).
- [20] F. Baumann, P. Lorenz-Spreen, I. M. Sokolov, and M. Starnini, *Phys. Rev. Lett.* **124**, 048301 (2020).
- [21] O. Sikder, R. E. Smith, P. Vivo, and G. Livan, *Sci Rep* **10**, 5493 (2020).
- [22] M. Gsänger, V. Hösel, C. Mohamad-Klotzbach, and J. Müller, *Entropy* **26**, 212 (2024).
- [23] P. G. Meyer and R. Metzler, *New J. Phys.* **10.1088/1367-2630/ad27bc** (2024).
- [24] A. Vendeville, B. Guedj, and S. Zhou, *Applied Network Science* **6**, 1 (2021).
- [25] S. Mori, M. Hisakado, and K. Nakayama, *Phys. Rev. E* **99**, 052307 (2019).
- [26] J. Fernández-Gracia, K. Suchecki, J. J. Ramasco, M. San Miguel, and V. M. Eguíluz, *Phys. Rev. Lett.* **112**, 158701 (2014).
- [27] J. Liu, S. Huang, N. M. Aden, N. F. Johnson, and C. Song, *Phys. Rev. Lett.* **130**, 037401 (2023).
- [28] C. M. Valensise, M. Cinelli, and W. Quattrociocchi, *Information Sciences* **642**, 119152 (2023).
- [29] K. Sasahara, W. Chen, H. Peng, G. L. Ciampaglia, A. Flammini, and F. Menczer, *J. Comput. Soc. Sc.* **10.1007/s42001-020-00084-7** (2020).
- [30] D. Braha and M. A. M. de Aguiar, *PLOS ONE* **12**, 1 (2017).
- [31] A. F. Peralta, J. Kertész, and G. Iñiguez, *Opinion dynamics in social networks: From models to data* (2022), arXiv:2201.01322 [physics.soc-ph].
- [32] P. Sobkowicz, *Frontiers in Physics* **8**, 10.3389/fphy.2020.587009 (2020).
- [33] O. Fraïsier, G. Cabanac, Y. Pitarch, R. Besançon, and M. Boughanem, **10.5281/zenodo.5535333** (2018).
- [34] See Supplemental Material at [URL will inserted by publisher] for details of the computations, networks statistics, and additional figures.
- [35] E. Papanastasiou and A. Giovanidis, *J Soc. Netw. Anal. Min.* **2301.10646** (2023).
- [36] L. Ramirez, M. San Miguel, and T. Galla, *Phys. Rev. E* **106**, 054307 (2022).
- [37] F. Vazquez and V. M. Eguíluz, *New J. Phys.* **10**, 063011 (2008).
- [38] V. R. K. Garimella and I. Weber, in *Proceedings of the 11th International Conference on Web and Social Media (ICWSM 2017)* (AAAI Press, 2017) pp. 528–531.
- [39] Y. Halberstam and B. Knight, *J Public Econ* **143**, 73 (2016).
- [40] M. D. Conover, J. Ratkiewicz, M. Francisco, B. Goncalves, F. Menczer, and A. Flammini, in *ICWSM* (2011).
- [41] J. Tacchi, C. Boldrini, A. Passarella, and M. Conti, in *Proceedings of the 2022 IEEE International Conference on Big Data* (2022) pp. 6030–6038.
- [42] H. T. Williams, J. R. McMurray, T. Kurz, and F. Hugo Lambert, *Glob Environ Change* **32**, 126 (2015).
- [43] J. Llabrés, M. San Miguel, and R. Toral, *Phys. Rev. E* **108**,

054106 (2023).

Rev. E **109**, 034307 (2024).

[44] L. S. Ramirez, F. Vazquez, M. San Miguel, and T. Galla, *Phys.*