STATISTICAL

Voter Model Dynamics on Networks with Social Features

Gabriel G. Piva[1,2] · **Fabiano L. Ribeiro**[1] · **Angelica S. da Mata**[1]

Abstract

The study of collective human behaviour in theoretical and real social systems is fundamental to understand the role of the social influence in human-to-human interaction. The voter model has been extensively studied in this context because of its straightforward approach and feasible theoretical treatment. In this regard, we aim to investigate the collective behaviour based on the influence of different factors associated with micro-level social processes. For example, in the voter model, the behaviour of the average time needed to a complete consensus depends on the network structure. Then, we investigated, numerically and analytically, how social network topology can affect the evolution of the voter model dynamics. We considered the social features fitness, homophily and Euclidean distance between nodes as preferential attachment rules in evolving networks. We show that the fact that these social attributes change the topological structure of the network generates impacts on the behaviour of the voter model dynamics running on top of these substrates. However, our simulations aim to interesting findings. Surprisingly, despite the social features and geographic properties present in the investigated networks, the standard heterogeneous mean-field theory can accurately describe the voter model in these investigated networks. Our results show, on the one hand, a strong correlation between the consensus time calculated on these network models and the consensus time obtained for real *s*ocial networks. It is also verified, on the other hand, an absence of correlation when we compared the synthetic networks with *n*on-social real networks. This finding suggests that the characteristics of network models, such as fitness, homophily and Euclidean distance between nodes, artificially imposed by the preferential attachment rules of the models, can indeed play the role of real social features.

Keywords Social features · Voter model · Preferential attachment networks · Real networks

1 Introduction

Social and external factors strongly influence the behaviour of people and their decision making. For instance, the influence that comes from peers, as friends, co-workers and family, and also the influence that comes from political biases

or exposure to the mass media [1–4]. In this context, many studies have investigated the role of social networks in collective dynamics as voting behaviour, for example [5–7]. In this scenario, the agreement is a relevant aspect of social dynamics since we frequently face situations requiring decision-making. Even simple ones as buying something or not, choose restaurant A or B for lunch; or until more complex decisions as to choose between candidate C or D for president. Of course, we know that consensus states are not observed in many real situations, such as political elections. However, basic models as the voter model can be helpful to investigate interacting particle systems a priori as a first attempt to explore complex factors that can affect such social spreading dynamics.

The simplest version of the voter model [8–11] starts with a population of N individuals, represented by nodes of a given network. Each node can be only in two states typified as a binary variable $s = \pm 1$. At each time step, a given node

Gabriel G. Piva, Fabiano L. Ribeiro, and Angelica S. da Mata contributed equally to this work.

✉ Angelica S. da Mata
angelica.mata@ufla.br

1 Departamento de Física, Universidade Federal de Lavras, Caixa Posta 3037, Lavras 37200-900, Minas Gerais, Brazil

2 Departamento de Física, Pontifícia Universidade Católica do Rio de Janeiro, Terminal da PUC - Gávea, Rio de Janeiro 22451-900, Rio de Janeiro, Brazil

copies the state (opinion) of a randomly chosen neighbour. The consensus is reached when all N nodes converge to the same state. The opinion dynamic studies using the voter model is already extensively explored in regular and complex networks from an analytical and numerical perspectives [11–23]. A widely investigated quantity in these works is how the *mean consensus time T* - the mean time need to arrive at a full consensus - changes with the network topology. The regime of time consensus depends on the number of agents N, and it was already investigated in regular lattices and generic uncorrelated heterogeneous graphs. The analytic solution for regular lattices is exact [12], but in uncorrelated heterogeneous networks, the solution is obtained through mean-field theory [15].

Despite its simplicity, the voter model and versions based on it can describe opinion dynamics in several situations. For example, the authors in Ref. [6] used the voter model as their basic dynamical system to investigate the collective dynamics of voting in US presidential elections. Vendeville et al. [7] based on the voter model with stubborn nodes to predict the result of elections. Another application of the model can be found in Ref. [24], where the authors used the voter model to analyse the language competition problem. However, in a more realistic context is necessary to add ingredients with social characteristics because community behaviour is affected by factors involving social, economic and cultural issues. For example, a social experiment proof that socioeconomic factors, like immigration, deindustrialization, demand for public services, among others, can interfere in the globalization of extremist ideologies [25].

In this context, we model this dynamical process on networks grown according to social characteristics that change their topology. We implemented the node's fitness, the Euclidean distance and the homophily between nodes as social features presented in the network. The main point of this work is to show how topological properties of networks, that can represent social features, affect dynamical processes. The choice of a simple opinion model is reasonable in order to focus on the effects of several social features embedded in the different topologies. Our results show that these characteristics provide a more realistic description of real systems and impact on how the consensus is reached. However, unexpectedly the standard heterogeneous mean-field theory proposed by Sood and Redner [15] to investigate the voter model dynamics seems to be quite reasonable to predict the behaviour of the mean consensus time in function of the network size.

Furthermore, it has been shown that social networks differ from other types of networks as biological and technological ones [26, 27]. In this context, we compared our network models with real systems—social and non-social ones. As a result, we show that the synthetic networks with social features embedded in their topology are able to estimate the consensus time of the voter model dynamics in real social systems. This finding shows us that such networks with social attributes are robust for modelling dynamic processes in more realistic situations and that the characteristics of network models, such as fitness, homophily and Euclidean distance between nodes are able to represent real social features.

This manuscript is divided as follows: we first described the network models with social features and the voter model in the methodology section. Subsequently, we presented our results and discussions about social parameters' influence on networks on the voter model dynamics, including comparisons with real systems. A detailed description of this set of real networks can be found in Appendix A. Finally, in the last section, we draw our conclusions.

2 Methodology

We used four preferential attachment network models to investigate the behaviour of the voter model. We aim to study the relation between the social features incorporated in the networks through these preferential connection mechanisms and the time needed for the dynamic to reach the consensus. Next, we briefly present such network models and contextualise the voter model's main results already consolidated in the literature.

2.1 Network Models

In this work, we used well-known network models that are based on two properties: the growth of the network, which means the nodes are added to the substrate; and the preferential attachment rule, that governs how these newly added nodes will connect to the ones already present in the network. Thus, the algorithm of the network models consists essentially of three stages: (a) the system starts with m_0 nodes connected to each other; (b) at each time step, a new node j enters on the network; (c) this new node chooses a node i, from all nodes already on the network, to connect. The general way to describe this last aspect is writing the probability $\Pi(i|j)$ that a node j chooses a node i to connect, which is given by

$$\Pi(i|j) = \frac{1}{Z} \frac{\eta_i(1 - A_{ij})}{r_{ij}^{\alpha_A}} k_i \tag{1}$$

where,

- k_i: is the degree of the i-th node;
- η_i: is the fitness of the i-th node. It is an intrinsic node property that gives its attractiveness, in the sense that as bigger it is, the more chance the node has to be chosen;
- $A_{ij} = |a_i - a_j|$: is a term that measures the similarity or affinity between the pair i and j, since a_i and a_j represent traits of the nodes i and j, respectively. The more similar

they are, the smaller A_{ij} is, and, consequently, they are more likely to connect.

- r_{ij}: is the euclidean distance between the pair i and j;
- α_A: is a parameter that controls the influence of the spatial distance in the connection between a pair of nodes;
- Z: is the normalization term, obtained by $\sum_i \Pi(i|j) = 1$.

Step (c) is repeated m times for each new node; that is, each new node j makes m connections when it is added to the network. That means the minimum degree of each node is m. The preferential attachment rule of the network models that we investigated can be reached as a particular case of the general formulation given by Eq. (1), as described below and summarized in Table (1).

For example, in the traditional Barabási-Albert (BA) model [28], new nodes preferably connected to other ones that already have more links. To represent this, we should consider $\alpha_A = 0$ (space is not considered), $A_{ij} = 0$ for any pair of nodes (absence of affinities), and $\eta_i = 1$ for all i (everybody has the same attractiveness) in the general formulation. The application of this preferential attachment rule generates a network with a power-law degree distribution $P(k) \sim k^{-\gamma}$, with $\gamma = 3$ in the thermodynamic limit.

The Bianconi and Barabàsi [29] model is obtained from this general framework when $\alpha_A = 0$ and $A_{ij} = 0$ for all pair of nodes. The probability that a new node j connect to a pre-existing node i dependent now on its degree (k_i) and its fitness (η_i). The choice of $\eta_i \in [0, 1]$ is usually given by a uniform distribution $\rho(\eta_i)$ [29]. In short, the node's fitness describing its ability to make links with other nodes. This quality can be associated with an individual's social skills to make friends, the content of a scientific publication, or even how enjoyable a YouTuber can be. This algorithm also generates a network with a power-law degree distribution but with $\gamma = 2.25$. A lower value of the γ exponent in comparison with BA network model is because the $r$ich-gets-richer phenomenon happens combined with the $f$it-gets-richer phenomenon. It means that even recently added nodes can obtain many links if it has a significant fitness.

Another relevant model within the network group of preferential attachment is the one proposed by de Almeida et al. [30] that includes the homophilic term. This term is able to increase the contact between similar nodes. This characteristic is related, for example, to the fact that people make friends easier if they share common features (mathematically represented by $A_{ij} = |a_i - a_j|$ in Eq. 1) such as musical taste, religion, soccer team, or professional environment. Then, the parameter $a_i \in [0, 1]$ represents an intrinsic property value of each node and it is given by a uniform distribution. It is reached from the general approach when $\alpha_A = 0$ and $\eta_i = 1$ for all i. This recipe also generates a network with a power-law degree distribution with $\gamma = 2.75$. This exponent value is smaller than the one of the BA model but higher than the one

of the fitness model. It happens because the $f$it-gets-richer phenomenon is not as pronounced as in the fitness model since similar values of a_i and a_j are necessary so that nodes i and j are more likely to connect.

Finally, the geographical proximity between the nodes is taken into account as a preferential attachment rule, beyond the degree of each node, in the network model proposed by Soares et al. [31]. It came from the general approach when $\eta_i = 1$ and $A_{ij} = 0$ for all pair. The following rule generates the spatial structure of the nodes distribution. It starts with just one node located in an arbitrary origin in a $2d$−dimensional substrate. Then, the second node is included in the network, and it is automatically connected to the node $i = 1$, and its localization is randomly chosen at a distance r from the first node. This distance r is distributed according to:

$$P_G(r) \propto \frac{1}{r^{2+\alpha_G}}, \tag{2}$$

where $\alpha_G > 0$ characterizes the growth pattern of the network, therefore, sub-index G. The new centre of mass (origin) is calculated, considering that each node has mass $\mu = 1$. A third node is including obeying the same spatial distribution, but it will be connected to node $i = 1$ or $i = 2$ according to the probability given by Eq. (1)—modified for this particular case—and so on.

The degree distribution for this type of network, independently of the values of α_G, is given by $P(k) \propto e_q^{-k/\kappa}$, where $e_q^x \equiv [1 + (1 - q)x]^{1/1-q}$ is the q−exponential function, where $e_1^x = e^x$. This function naturally emerges within non-extensive statistical mechanics [32], and q is a function of α_A and κ. Here, $\kappa > 0$ can be interpreted as the characteristic number of links. In this model, the degree distribution changes depending on how the distance interferes in the connection (controlled by the parameter α_A), as will be explored in Sect. 3. However, there is no change in the distribution of connectivity when we change the values of α_G, since this parameter does not interfere with the preferential attachment rule.

For a detailed explanation about all these network models, see a recent review in Ref. [33].

2.2 Voter Model Dynamics

Sood and Redner [15] evaluated analytically the consensus time for a generic uncorrelated heterogeneous graph using mean-field theory. Their analytical predictions for the voter model on complex networks show that the consensus time depends sublinearly on the network size N. They found that the time to reach consensus for uncorrelated degree networks with any degree distribution is a function of N given by

$$T(N) = -N \frac{\langle k \rangle^2}{\langle k^2 \rangle} [(1 - \omega) \ln(1 - \omega) + \omega \ln \omega], \tag{3}$$

where $\langle k^n \rangle$ is the n-th moment of the degree distribution and ω is the initial density of nodes with opinion $+1$, that is $\omega = \frac{1}{\langle k \rangle N} \sum_{i|s_i=+1} k_i$.

However, in this work, we will restrict the analysis using the initial configuration composed by half of the network with $+1$ opinion and the other half -1, which implies that the term in brackets in Eq. (3) is always equal to $-\ln(2)$. The authors also compared the mean-field results with numerical simulations, showing that they agree even for correlated degree networks, indicating that their theory is robust. For heterogeneous networks with power-law degree distribution, the asymptotic behaviour of consensus times is [15, 17]

$$T(N) \sim \begin{cases} N, & \gamma > 3 \\ N/\ln N, & \gamma = 3 \\ N^{2(\gamma-2)/(\gamma-1)}, & 2 < \gamma < 3 \\ (\ln N)^2, & \gamma = 2 \\ \mathcal{O}(1), & \gamma < 2 \end{cases} \qquad (4)$$

For Barabási-Albert network, for example, the numerical value found in simulations is $T(N) \sim N^{0.88}$ that agrees very well with this analytical forecast [21]. In the next section, we present the results for the voter model dynamics running on top of different network topologies as presented above: BA, Fitness (Bianconi-Baràbasi) and Homophilic networks, and the model with Euclidean distance. We also compare our simulation results with the mean-field theory proposed by Sood and Redner [15]. Interestingly, our results showed that this mean-field approach is robust to predict the consensus time behaviour even for heterogeneous networks with characteristics that generate degree social features and geographic properties (see Ref. [33] for more details about these network models).

Another way to implement the model consists of a link-update rule in which, instead of choosing a node and after, its neighbour, we randomly choose a link, and both nodes at the end of this link are randomly assigned to the same opinion if they are opposites [20]. Although these two approaches are equivalent in regular and homogeneous networks, they are different in heterogeneous scale-free networks. Both methods conserve the average magnetization in homogeneous graphs, but in heterogeneous networks this quantity is conserved only for the link-update rule. The node-update rule conserves the average magnetization weighted by the degree of the node [23].

3 Results

Our results are subdivided into three topics. First, we investigated the influence of the fitness and the homophilic between nodes in the behaviour of the mean consensus time of the voter model. Then, we compared these networks with the traditional BA network. After, we showed the influence of the Euclidean distance between the nodes in the voter model dynamics. At last, we compared our synthetic networks with real ones to investigate the behaviour of voter model in social systems.

3.1 The Influence of Fitness and Homophilic Parameters

The power-law degree exponent γ is, in some way, a measurement of the heterogeneity of the network. More specifically, networks with lower γ are more heterogeneous and, consequently, have more hubs, it means, more nodes that have a significantly larger number of links (neighbours) in comparison with most nodes in the network. Therefore, the time to reach the consensus is decreased as the power-law exponent is reduced [23] since the presence of many hubs favours order (the neighbours of hubs tend to copy their state (opinion)). This behaviour can be observed in Fig. 1a, where we show the mean consensus time as a function of the network size for BA, Homophilic and Fitness networks that have $\gamma = 3.0, 2.75$ and 2.25, respectively[1]. In this first analysis, we are not considering the Euclidean distance between the nodes, this means that, for now, we are keeping $\alpha_A = 0$. The Fitness model has a lower γ exponent, which means that it has more hubs that favour order and, consequently, the time to reach consensus is shorter compared to other networks of the same size. In this figure, we also show that Eq. (3) fits our numerical results very well, revealing that the Sood and Redner theory [15] is actually quite reasonable to predict the behaviour of the mean consensus time for all these network models.

3.2 The Relevance of Euclidean Distance

Recent studies [34–39] have been emphasized the importance to consider the euclidean distance among the nodes. For example, Goldenberg and Levy [34] investigated the role of geographic distance in Facebook's network of friends. They analysed $100,000$ Facebook users and $1,297$ linked pairs with reported zip codes and concluded that the probability of a friendship link between two people decreases with distance, and most Facebook connections are between people within a geographical distance of up to 100 miles. Similar behaviour was also observed for electronic communications. The same authors asked subjects to report the locations of the receivers of their last 50 email messages and

---

[1] As we observed in the illustrative scheme of Table (1), the BA network means $A_{ij} = 0$ for any pair of nodes and $\eta_i = 1$ for all node i. In the fitness model, $A_{ij} = 0$ and $\eta_i \in [0, 1]$, and in the homophilic model, $\eta_i = 1$ but $A_{ij} = |a_i - a_j|$, with $a_i \in [0, 1]$.
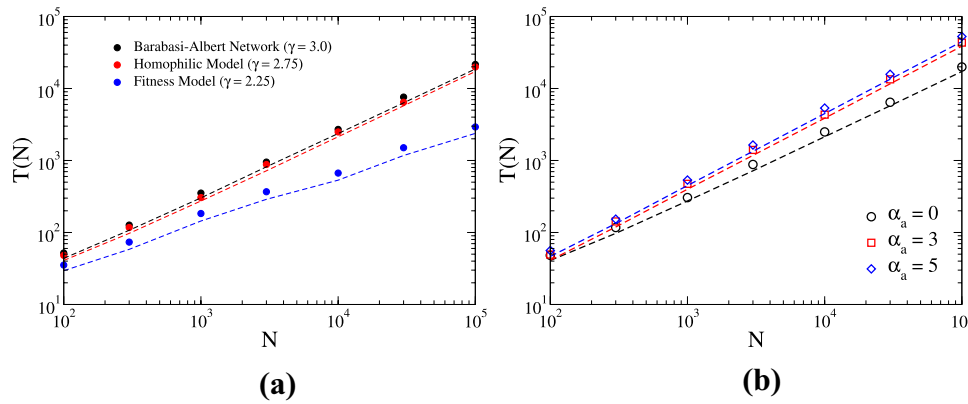
(a)   **(b)**

Fig. 1 (**a**) Average consensus time against network size for BA, Fitness and Homophilic networks. We used $\alpha_g = 2$ and $\alpha_A = 0$, which means we not consider the euclidean distance and just compare the homophily and fitness parameters. The time to reach the consensus decreases as the power-law exponent reduces. (**b**) The same data for the Homophilic model with Euclidean distance. The parameter $\alpha_g = 2$ is kept fixed, and the value of α_A ranges between 0 and

5. For $\alpha_A > 2$, the pattern of the consensus time in function of N is pretty much the same (see Fig. 2b for more data). In both figures, the dashed lines are the prediction of the mean-field theory proposed by Sood and Redner [15] according to Eq. (3). Both results—simulation and prediction of Eq. (3)—were obtained using 1000 realizations of the voter model in each kind of network and therefore are subject to stochastic fluctuations

their city of residence, collecting data for 4,455 email messages. They showed that message exchanges also decrease with distance, and 30% of all messages were sent to colleagues who are at most 100 miles away.

Latané [40] also described how an individual can be influenced by another in a social group. The convincing power of individuals and the spatial proximity between them can influence personal relationships and also decision-making.

So, considering that physical proximity is an essential factor in social tie formation [41–43] even in virtual social networks, we can consider geographical proximity using the model proposed by Soares et al. [31]. Until now, we are verifying the influence of the Euclidean distance between nodes, so we are going to use just the BA network, without considering the fitness and homophily effects, this means, $A_{ij} = 0$ for any pair of nodes and $\eta_i = 1$ for all node i (see Table 1). As we mentioned before, the parameter α_G does not affect the behaviour of the connectivity distribution $P(k)$ of the network (for details see Ref. [33]). This parameter refers just to the distance distribution in relation to the centre of mass and does not impact the preferential attachment rules. Consequently, the behaviour of the voter model is not affected.

On the other hand, as α_A increases, the connectivity distribution changes (see Fig. 2a). A topological transition appears close to $\alpha_A \approx 2$ when the network changes from heterogeneous connectivity to an increasingly homogeneous topology. In the original work about the Euclidean distance model, Soares et al. [31] showed numerically this topological phase transition using the q-exponential distribution to fit the degree distribution. Recently, Piva et al. [33] extended this demonstration using other complex network's measures, such as degree correlation and shortest path length. Finally,

Cinardi et al. [44] and collaborators generalised the model for asymptotically-scale-free geographical networks. In our work, we will focus on investigating how this topological phase transition affects the consensus time in the voter model dynamics.

Numerical simulation supports the power-law behaviour of the time of consensus in function of N, since $T(N) \sim N^{-\nu}$, where ν is the exponent that depends on the model and can be numerically determined. As the network becomes more homogeneous as α_A increases, the time to reach the consensus also increases; consequently, the exponent ν becomes more prominent, as we show in Fig. 2b. This exponent changes sharply when $\alpha_A \approx 2$. For $\alpha_A = 0$ and 1, we obtained $\nu \approx 0.88(2)$, for $\alpha_A = 2$, we measured $\nu = 0.92(2)$, and, finally, for $\alpha_A > 2$ the exponent stabilized around 0.98(2). The numbers in parentheses indicate uncertainties in the last digit. The inset presented in Fig. 2b enlightened this crossover.

This transition can be understood as a consequence of the nodes interaction range that changes from a *long-range regime* (α_A sufficiently small) to a *short-range regime* (α_A sufficiently large). A naive calculus to support this idea can be done considering that $P(r) \sim 1/r^{\alpha_A}$ is the probability of two nodes being connected based on the preferential rule related to geographic proximity[2]. Then, the average distance $\langle r \rangle$ between two nodes can be estimated by the integral

---

[2] Note that we did not consider the effects of the preferential attachment according to the degree of the node. That is, it is just a mere naive calculus to give the idea of a behaviour change according to α_A values.

Table 1 A schema illustrating the parameters that are consider (green) or not (red) in each network model

Model	k_i	A_{ij}	η_i	r_{ij}
Barabási-Albert	✓	✗	✗	✗
Bianconi-Barabási (Fitness Model)	✓	✗	✓	✗
Almeida *et al.* (Homophilic Model)	✓	✓	✗	✗
Soares *et al.* (Euclidean Distance Model)	✓	✗	✗	✓

$$\langle r \rangle = \int_0^\infty r P(r)dr \sim r^{2-\alpha_A}\Big|_0^\infty. \tag{5}$$

That is, $\langle r \rangle \to \infty$ (long-range interaction regime) when $\alpha_A \leq 2$; and $\langle r \rangle = $ cte (short-range regime) when $\alpha_A > 2$. This is one more evidence that occurs a topological phase transition around $\alpha_A \approx 2$.

In order to show how this topological transition affects the dynamics of the voter model, we plot in Fig. 3a the number of hubs presented in the network as a function of the attachment parameter α_A. From $\alpha_A \approx 2$, the number of hubs starts to decrease; that is, the network becomes less and less heterogeneous. It affects the consensus time, which increases with α_A and becomes practically constant for $\alpha_A > 2$, as shown in Fig. 3b. These scenarios show how the network topology, influenced by the preferential attachment according to the geographic distance between the nodes, impacts the dynamic

process under discussion. As a criterion, we consider hubs all nodes with degree $k > 50$ since in a network with a size $N = 10^4$, the degree cutoff is around $k_c \sim \sqrt{N} \approx 100$ [45]. However, our results do not change qualitatively as a function of the fixed threshold.

We can also add the Euclidean distance between nodes as another preferential attachment rule in the Fitness and Homophilic models. It makes the network even more similar to a real one. In Fig. 1b, we observe the evolution of the consensus time in function of the network size for the Homophilic model with this additional attachment rule. We also observe a change in the behaviour of $T(N)$ versus N for α_A greater than or less than 2. The same results can be obtained for the fitness model with Euclidean distance. These findings show that the distance between the nodes as a preferential attachment rule is a robust feature that affects the voter model dynamics. We also show that the forecast of
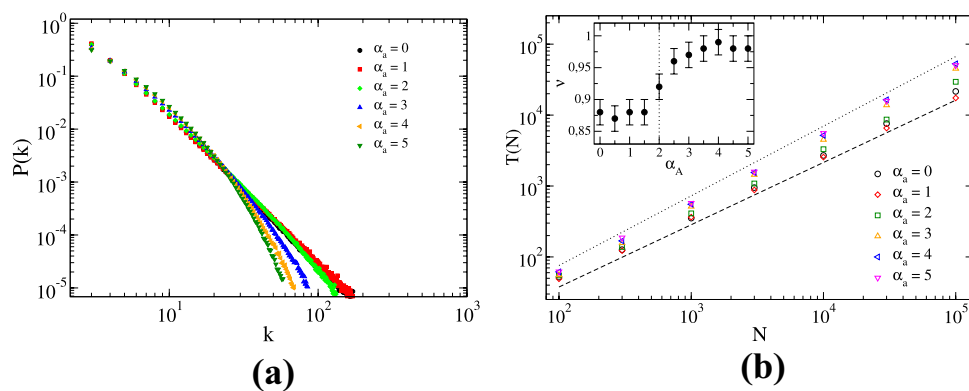

(a)

(b)

Fig. 2 (**a**) Connectivity distribution to different values of α_A with fixed $\alpha_G = 2$ on a network of size $N = 10^4$. As α_A increases, the connectivity distribution changes from heterogeneous connectivity to an increasingly homogeneous topology. The graph is on the log-log scale. Results obtained with an average of 1000 samples. (**b**) Consensus time against network size for the voter model in the Euclidean Distance model. The value of $\alpha_G = 2$ is fixed and the values of α_A range from 0 to 5. The straight slope $T(N) \sim N^{-\nu}$ changes with the variation of the parameter α_A. The dashed and dotted lines have $\nu = 0.88$ and 0.98, respectively, and it serves as a guide the to eyes. The points are obtained over 1000 realizations. In the inset, we show the variation of the exponent ν as a function of the parameter α_A. The dotted line enlightened the sharp change that occurs for ν values near $\alpha_A = 2$

Fig. 3 (**a**) Number of hubs as a function of the α_A parameter. The data are obtained over 5 network samples. (**b**) Mean consensus time against the attachment parameter. The points are obtained over 1000 realizations of the voter model. In both, we fixed the network size $N = 10^4$ and the parameter $\alpha_g = 2$. For $\alpha_A \gtrsim 2$, the quantity of hubs decreases, a characteristic of homogeneous networks, and the mean consensus time tends to stabilize at a higher value
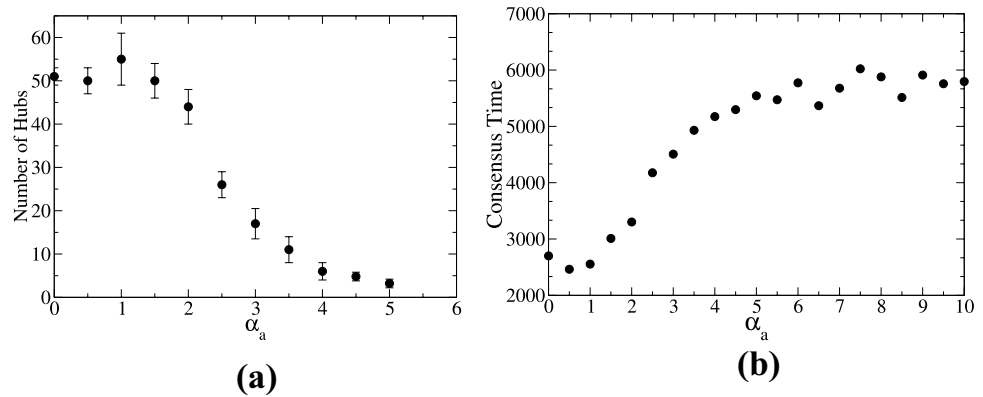
the heterogeneous mean-field theory given by Eq. (3) works very well, showing that the Sood and Redner theory [15] is sufficient to predict the consensus time in different networks.

3.3 Voter Model in Real Networks

In order to investigate the impact of the social features that we presented in synthetic networks, we calculated the mean consensus time using the voter model dynamics running on top of real substrates. From now on, we compared the time needed to reach consensus in both synthetic and real networks with the same size and analysed if there is a correlation between these values. If there is a positive correlation, we can infer that the artificial networks with preferential attachment based on social features are feasible models to predict outcomes related to real dynamics.

We selected two sets of systems—social and non-social one—with size ranging from $N \approx 100$ to $N \approx 10000$. We investigated eight social networks and seven non-social

ones. The classification of them was made based on the description of each network, as we showed in the Appendix A (Table 2) [46].

In Fig. 4, we showed the behaviour of the mean consensus time—obtained using the voter model dynamic—for real networks in comparison with the Barabási-Albert, Fitness and Homophilic synthetic networks, all of them also including euclidean distance ($\alpha_A = 2$ and $\alpha_g = 2$). In the panel (a), we plotted the results of the predicted consensus time given for synthetic networks versus the value of consensus time for real social networks of the same size. We observed a high correlation between them ($R = 0.99$); however, the same plot for non-social systems presented a distinct scenario, and there is no correlation between the data ($R = -0.34$).

To elucidate this scenario, we calculated and compared essential metrics for synthetic networks and both types of real networks, such as connectivity distribution, shortest path length, cluster coefficient, Pearson coefficient, the average degree of network nodes, the density of connections on
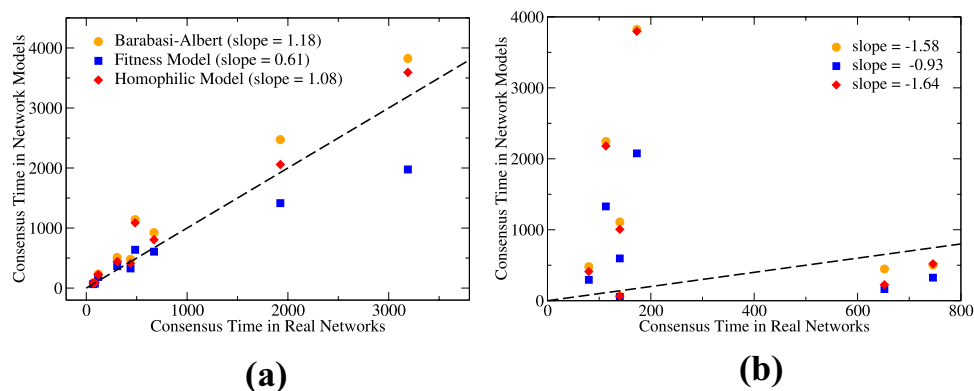

Fig. 4 Predicted consensus time in: synthetic networks and real ones: (**a**) social and (**b**) non-social networks. In both panels, the synthetic networks models are represented by the same symbols, and the dashed line represents a line $y = x$ that works as a guide to the eyes. In panel (**a**), we observed substantial linearity between data, mainly for Barabási-Albert and Homophilic models. The correlation coefficient between predicted and real data is around 0.99 for scenario

(**a**) and –0.34 for scenario (**b**). For social networks, we also found angular coefficient $a = 1.18$ for Barabási-Albert model, $a = 1.08$ for Homophilic model, and $a = 0.61$ for the Fitness Model. While for non-social networks, there is no linear relation between the data. Here, the Euclidean distance between nodes was included as a preferential attachment rule in all models. We used $\alpha_A = \alpha_G = 2$.

the network, etc. [47]. No similarity and/or significant difference was found between synthetic networks and both types of real networks concerning these characteristics. However, our finding related to the dynamic of the voter model shows that the network models we investigated help describe social systems since there is significant evidence that the fitness, homophily and Euclidean distance between nodes, included in the models through the preferential attachment rules, can capture essential features that seem to emerge spontaneously in the real social networks.

4 Concluding Remarks

In this work, we have investigated how social features can affect the evolution of the voter model dynamics on networks. We introduced network models that contemplate essential characteristics of social systems such as fitness, homophily and the concern about geographic distance between their elements. Our results illustrated that euclidean distance, fitness and homophily proved to be pertinent characteristics to highlight the difference between social and nonsocial systems. The fitness represents the capability of somebody to convince someone to change or not his/her opinion; the spatial proximity is related to personal relationships; and the homophily represents the cultural, political or religious affinity among people, for example. All these elements represent the influence of social groups in individual attitudes [40].

Within the assumptions and limitations of our model, it is essential to highlight that our results show that these social features change the topological structure of the network and consequently affect the voter model's consensus time. Therefore, we consider the choice of a simple opinion model to

explore the effects of different substrates is quite reasonable. Maybe, a more complex dynamical model could mask the main findings related to the structure of the network models. However, now that we have taken a step forward in understanding a dynamical process running on top of these kinds of substrate, some model extensions could be considered to improve its feasibility and become more plausible to describe empirical voting data. Interestingly, even with such topological changes in the network, due to the different preferential attachment rules, the heterogeneous mean-field theory [15] was able to describe the voter model dynamics well.

Going beyond this context and showing reasonably practical applications of these types of networks, we can cite the work of Paiva et al. [48]. They strongly suggest that real networks originating from the interaction of termites have preferential attachment features responsible for generating specific movement patterns in these social insects. Although more detailed studies are needed, this research shows that our findings are plausible, in order to that our results suggest that the characteristics of synthetic networks placed by the preferential attachment rules—homophily, fitness and Euclidean distance between the elements of the network—can generate the social interactions between the individuals of the system.

Finally, despite individual opinions being influenced by many social and psychological factors, taking all of them into account besides being impracticable, it is not necessary to understand the large-scale behaviour of the system. As a result, this study represents a step toward providing a background for understanding more realistic dynamics on more complex topologies. We hope that our analysis can be applied to any dynamic model running on top of a network, so our work could be helpful as a template for future works.

Appendix

A. Real Networks Information

In the table below, we describe the main data about the set
of real networks that we used in our analysis.

Table 2 Descriptions of the real networks. Here, we presented their names, sizes (N), descriptions, and what type of network they can classify—social or non-social networks. In the last column, we included the references from which we obtained the data

Network	N	Description	Type	Reference
E. coli transcription	97	A directed network of regulated gene.	Non-Social	[49]
s208	122	Network of electronic circuits.	Non-Social	[50]
Radoslaw Email	167	An internal email communication network. The network is directed and links between them represent individual emails.	Social	[51]
Little Rock Lake*	183	A food web of Little Rock Lake, Wisconsin (USA). Nodes represent autotrophs, herbivores, carnivores and decomposers and edges represent food sources.	Social	[52]
US Air transportation**	500	A weighted network consists of 500 United State airports with the largest amount of traffic. Nodes are airports and edges represent air travel connections among them.	Social	[53]
s838	512	Network of electronic circuits.	Non-Social	[50]
E-mail URV Network	1133	List of links of e-mail interchanges between members of the Univeristy Rovira i Virgili (Tarragona).	Social	[54]
Political Blogs	1222	A directed network of edges between weblogs on United State politics, recorded in 2005 by Adamic and Glance.	Social	[55]
Air traffic	1266	Network composed by airports and service centers whose edges are created from strings of preferred routes recommended by the National Flight Data Center of United States of America.	Non-Social	[56]
Japanese	2698	A language system represented by word adjacency network.	Non-Social	[50]
Open Flights**	2905	A directed network representing flights between airports of the world.	Social	[57]
Gnutella	6299	A file sharing network mainly used for exchanging music, movies and software. Nodes represent hosts in the Gnutella network and links occur between the Gnutella hosts. This network used here is a sequence of snapshots of the Gnutella file sharing from August 2002.	Social	[58, 59]
JDK	6434	A directed network in which nodes are classes and links between them represent a dependency between two classes. It is a software class dependency network of the JDK 1.6.0.7 framework.	Non-social	[56]
PGP	10680	The links of the giant component of a network of users of the Pretty-Good-Privacy algorithm for secure information interchange	Social	[27]
Spanish	11554	A language system represented by word adjacency network.	Non-Social	[50]

*The Little Rock Lake represents a food chain network but can be characterized as a social network because it represents a community of species that related to each other; **Altough US Air transportation and Open Flights networks can be considered structural network as air traffic network, we believe that it is more reasonable to classify both as social systems since they are related to the flow of people and not to the structure of the network itself. This is different, for example, from the more technological description of the Air Traffic network

Declarations

Conflict of Interest The authors declare no competing interests.

References

1. R. Huckfeldt, J. Sprague, Discussant effects on vote choice: Intimacy, structure, and interdependence. J. Politics **53**(1), 122–158 (1991)
2. C.B. Kenny, Political participation and effects from the social environment. Am. J. Pol. Sci. **36**(1), 259–267 (1992)
3. S. Galam, F. Jacobs, The role of inflexible minorities in the breaking of democratic opinion dynamics. Phys. A: Stat. Mech. App. **381**, 366–376 (2007)
4. M. Gilens, L. Vavreck, M. Cohen, The mass media and the public's assessments of presidential candidates, 1952–2000. J. Politics **69**(4), 1160–1175 (2007)

5. P.A. Beck, R.J. Dalton, S. Greene, R. Huckfeldt, The social calculus of voting: Interpersonal, media, and organizational influences on presidential choices. Am. Pol. Sci. Rev. **96**(1), 57–73 (2002)

6. D. Braha, M.A. de Aguiar, Voting contagion: Modeling and analysis of a century of US presidential elections. PLoS One **12**(5), e0177970 (2017)

7. A. Vendeville, B. Guedj, S. Zhou, Forecasting elections results via the voter model with stubborn nodes. Appl. Netw. Sci. **6**(1) (2021)

8. M.S. Miguel, V.M. Eguiluz, R. Toral, K. Klemm, Binary and multivariate stochastic models of consensus formation. Comput. Sci. Eng. **7**(6), 67–73 (2005)

9. P. Clifford, A. Sudbury, A model for spatial conflict. Biometrika **60**(3), 581–588 (1973)

10. P.L. Krapivsky, Kinetics of monomer-monomer surface catalytic reactions. Phys. Rev. A **45**, 1067–1072 (1992)

11. L. Frachebourg, P. Krapivsky, Exact results for kinetics of catalytic reactions. Phys. Rev. E **53**(4), R3009 (1996)

12. J.T. Cox. Coalescing random walks and voter model consensus times on the torus in Zd. Ann. Probab .1333–1366 (1989)

13. R.A. Holley, T.M. Liggett. Ergodic theorems for weakly interacting infinite systems and the voter model. Ann. Probab. 643–663 (1975)

14. E. Ben-Naim, L. Frachebourg, P.L. Krapivsky, Coarsening and persistence in the voter model. Phys. Rev. E **53**(4), 3078 (1996)

15. V. Sood, S. Redner, Voter model on heterogeneous graphs. Phys. Rev. Lett. **94**, 178701 (2005)

16. C. Castellano, S. Fortunato, V. Loreto, Statistical physics of social dynamics. Rev. Modern Phys. **81**(2), 591 (2009)

17. V. Sood, T. Antal, S. Redner, Voter models on heterogeneous networks. Phys. Rev. E **77**(4), 041121 (2008)

18. C. Castellano, D. Vilone, A. Vespignani, Incomplete ordering of the voter model on small-world networks. EPL (Europhys. Lett.) **63**(1), 153 (2003)

19. F. Vazquez, V.M. Eguíluz, Analytical solution of the voter model on uncorrelated networks. New J. Phys. **10**(6), 063011 (2008)

20. K. Suchecki, V.M. Eguiluz, Miguel M. San, Conservation laws for the voter model in complex networks. EPL (Europhys. Lett.) **69**(2), 228 (2004)

21. C. Castellano, V. Loreto, A. Barrat, F. Cecconi, D. Parisi, Comparison of voter and Glauber ordering dynamics on networks. Phys. Rev. E **71**(6), 066107 (2005)

22. E. Pugliese, C. Castellano, Heterogeneous pair approximation for voter models on networks. EPL (Europhys. Lett.) **88**(5), 58004 (2009)

23. C. Castellano. Effect of network topology on the ordering dynamics of voter models. In: AIP Conference Proceedings, vol. 779 (AIP, 2005), pp. 114–120

24. I. Caridi, F. Nemiña, J.P. Pinasco, P. Schiaffino, Schelling-voter model: aAn application to language competition. Chaos Soliton. Fract. **56**, 216–221 (2013)

25. J. Essletzbichler, F. Disslbacher, M. Moser, The victims of neoliberal globalisation and the rise of the populist vote: a comparative analysis of three recent electoral decisions. Cambridge J. Reg. Econ. Soc. **11**(1), 73–94 (2018)

26. M.E.J. Newman, J. Park, Why social networks are different from other types of networks. Phys. Rev. E **68**, 036122 (2003)

27. M. Boguñá, R. Pastor-Satorras, A. Díaz-Guilera, A. Arenas, Models of social networks based on social distance attachment. Phys. Rev. E **70**, 056122 (2004)

28. A.L. Barabási, R. Albert, Emergence of scaling in random networks. Science **286**(5439), 509–512 (1999)

29. G. Bianconi, A.L. Barabási, Competition and multiscaling in evolving networks. EPL (Europhys. Lett.) **54**(4), 436 (2001)

30. M.L. de Almeida, G.A. Mendes, G. Madras Viswanathan, L.R. da Silva, Scale-free homophilic network. Eur. Phys. J. B **86**(3), 38 (2013)

31. D.J. Soares, C. Tsallis, A.M. Mariz, L.R. da Silva, Preferential attachment growth model and nonextensive statistical mechanics. EPL (Europhys. Lett.) **70**(1), 70 (2005)

32. C. Tsallis, Nonextensive statistics: Theoretical, experimental and computational evidences and connections. Braz. J. Phys. **29**(1), 1–35 (1999)

33. G.G. Piva, F.L. Ribeiro, A.S. Mata, Networks with growth and preferential attachment: Modelling and applications. J. Complex Netw. **9**(1) (2021)

34. J. Goldenberg, M. Levy, Distance is not dead: Social interaction and geographical distance in the internet era (2009). Preprint at arXiv:09063202

35. Y. Xu, A. Belyi, I. Bojic, C. Ratti, How friends share urban space: an exploratory spatiotemporal analysis using mobile phone data. Trans. GIS **21**(3), 468–487 (2017)

36. B. Lengyel, A. Varga, B. Ságvári, Á. Jakobi, J. Kertész, Geographies of an Online Social Network. PLoS One **10**(9), 1–13 (2015)

37. D. Laniado, Y. Volkovich, S. Scellato, C. Mascolo, A. Kaltenbrunner, The impact of geographic distance on online social interactions. Info. Syst. Front. (2017)

38. L. Liu, B. Chen, C. Ai, L. He, Y. Wang, X. Qiu et al., ISPRS Int. J. Geo. Inf. **7**, 189 (2018)

39. X. Liu, Y. Xu, X. Ye, Outlook and next steps: Integrating social network and spatial analyses for urban research in the new data environment. Cities as spatial and social networks. (Springer, Cham, 2019), pp. 227–238

40. B. Latané, The psychology of social impact. Am. Psychol. **36**(4), 343–356 (1981)

41. F.L. Ribeiro, J. Meirelles, F.F. Ferreira, C.R. Neto, A model of urban scaling laws based on distance dependent interactions. Royal Soc. Open Sci. **4**(3), 160926 (2017)

42. V.M. Netto, J. Meirelles, F.L. Ribeiro, Social interaction and the city: the effect of space on the reduction of entropy. Complexity **2017** (2017)

43. V.M. Netto, J. Meirelles, F.L. Ribeiro, Cidade e interação: O papel do espaço urbano na organização social. Revista Brasileira de Gestão Urbana **10**(2), 249–267 (2018)

44. N. Cinardi, A. Rapisarda, C. Tsallis, A generalised model for asymptotically-scale-free geographical networks. J. Stat. Mech. **2020**(4), 043404 (2020)

45. A. Barrat, M. Barthelemy, A. Vespignani, Dynamical Processes on Complex Networks. (Cambridge: Cambridge University Press, 2008). https://doi.org/10.1017/CBO9780511791383

46. F. Radicchi, C. Castellano, Breaking of the site-bond percolation universality in networks. Nat. Commun. **6**(2041–1723), 10196 (2015)

47. R. Albert, A.L. Barabási, Statistical mechanics of complex networks. Rev. Modern Phys. **74**(1), 47 (2002)

48. L.R. Paiva, A. Marins, P.F. Cristaldo, D.M. Ribeiro, S.G. Alves, A.M. Reynolds, et al., Scale-free movement patterns in termites emerge from social interactions and preferential attachments. Proc. Nat. Acad. Sci. **118**(20) (2021)

49. S. Mangan, U. Alon, Structure and function of the feed-forward loop network motif. Proc. Nat. Acad. Sci. USA. **100**(21), 11980–11985 (2003)

50. R. Milo, S. Itzkovitz, N. Kashtan, R. Levitt, S. Shen-Orr, I. Ayzenshtat et al., Superfamilies of evolved and designed. Networks **303**(5663), 1538–1542 (2004)

51. R.A. Rossi, N.K. Ahmed, In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 29 (2015)

52. N.D. Martinez. Artifacts or attributes? Effects of resolution on the Little Rock Lake Food Web. Ecol. Monogr. **61**(4), 367–392 (1991)

53. V. Colizza, R. Pastor-Satorras, A. Vespignani, Nat. Phys. **3**(4), 276–282 (2007)

54. R. Guimerà, L. Danon, A. Díaz-Guilera, F. Giralt, A. Arenas, Phys. Rev. E **68**, 065103 (2003)

55. L.A. Adamic, N. Glance, In: Proceedings of the 3rd International Workshop on Link Discovery. LinkKDD '05. (New York, NY, USA: Association for Computing Machinery, 2005), pp. 36–43

56. J. Kunegis. KONECT: The Koblenz Network Collection. In: Proceedings of the 22nd International Conference on World Wide

Web. WWW '13 Companion. (New York, NY, USA: Association for Computing Machinery, 2013), pp. 1343–1350

57. T. Opsahl, F. Agneessens, J. Skvoretz, Node centrality in weighted networks: Generalizing degree and shortest paths. Soc. Netw. **32**(3), 245–251 (2010)

58. J. Leskovec, J. Kleinberg, C. Faloutsos, Graph evolution: Densification and shrinking diameters. ACM Trans. Knowl. Discov. Data **1**(1), 2–es (2007)

59. M. Ripeanu, A. Iamnitchi, I. Foster, Mapping the Gnutella Network. IEEE Internet Comput. **6**(1), 50–57 (2002)