Spread of opinions and proportional voting

Gonzalo Travieso* and Luciano da Fontoura Costa†
Instituto de Física de São Carlos, Universidade de São Paulo, Av. do Trabalhador São-Carlense, 400, 13566-590 São Carlos, SP, Brazil

(Received 7 March 2006; revised manuscript received 12 July 2006; published 22 September 2006)

Election results are determined by numerous social factors that affect the formation of opinion of the voters, including the network of interactions between them and the dynamics of opinion influence. In this work we study the result of proportional elections using an opinion dynamics model similar to simple opinion spreading over a complex network. Erdös-Rényi, Barabási-Albert, regular lattices, and randomly augmented lattices are considered as models of the underlying social networks. The model reproduces the power law behavior of a number of candidates with a given number of votes found in real elections with the correct slope, a cutoff for a larger number of votes, and a plateau for a small number of votes. It is found that the small world property of the underlying network is fundamental for the emergence of the power law regime.

DOI: 10.1103/PhysRevE.74.036112 PACS number(s): 89.75.−k, 87.23.Ge

I. INTRODUCTION

There has been a growing interest in the study of social phenomena through the use of tools from statistical physics [1–4]. This trend has been in part stimulated by developments in complex networks [5–8], which have uncovered properties of the structures underlying the interactions between agents in many natural, technological, and social systems. Social processes can be simulated through the use of complex network models over which a dynamical interaction between the agents represented by the nodes is defined, yielding results that can be compared with the macroscopic results found in real social networks.

Election of representatives are important social processes in democracies, where a large number of people take part and that represent the result of many social factors. It was found [9] that the number of candidates with a given number of votes in the 1998 Brazilian elections follows a power law with slope −1 for some orders of magnitude, or a generalized Zipf's law [10].

Elections depend on the process of opinion formation by the voters. Each voter chooses one candidate based on its beliefs and through interaction with other voters. Many works have been carried out on opinion formation while considering several types of dynamics and underlying network topologies. Bernades *et al.* [11] and González *et al.* [12] succeeded in reproducing the general −1 slope of candidates with a given number of votes in Brazilian election results by using the Sznajd [13] opinion formation model adapted to complex networks.

In the Sznajd model, two neighbors that happen to have the same opinion may convince their other neighbors. In this paper, we adopt a simpler model, where each single voter tries to convince its neighbors, regardless of their previous opinion. The obtained results exhibited a substantial agreement with real election results for some network models.

The paper is organized as follows. Firts we describe the network (Sec. II A) and opinion (Sec. II B) models used in the simulations. Then, in Sec. III we present and discuss the simulation results and study the effect of the model parameters. Finally, the conclusions are summarized in Sec. IV.

II. OPINION AND NETWORK MODELS USED

As done in other related works, we assume that the opinion formation for the voting process occurs as interactions between agents connected through a complex network. The result is thus determined by two factors: (i) the structure of the network that specifies the possible interactions between agents, and (ii) the dynamics of opinion formation between interacting agents. The following subsections describe the models used in this work.

A. Network models

The voters and their social interactions are represented as a network, so that the individuals are represented by nodes in the network and every social interaction between pairs of voters is represented by a link between the two corresponding nodes. The number of links attached to a node is called the *degree* of the node; the social distance between voters is given by the geodesic distance in the network, defined as the minimum number of links that must be traversed in order to reach one of the nodes starting from the other. Two important network properties [7] are the degree distribution and the average distance between pairs of nodes.

For the simulation of the opinion formation model we adopted the Erdös-Rényi and the Barabási-Albert [5,14] models of complex networks. For comparison, simulations were also performed in two-dimensional lattices, two-dimensional lattices with random connections added between its nodes, and a generalized model of Barabási and Albert [15]. The Erdös-Rényi networks are characterized by a Poisson degree distribution and the presence of the "small world" property: the average distance between nodes grows slowly with the number of nodes in the network. The Barabási-Albert model also has the small world property, but its degree distribution follows a power law, resembling in that sense many social networks. The regular lattice was chosen

*Electronic address: gonzalo@ifsc.usp.br
†Electronic address: luciano@ifsc.usp.br

as an example of a network without the small world property, while the addition of random connections enables a controlled introduction of this property (see Ref. [16]). The generalized Barabási-Albert model is characterized by an adjustable slope of the power law degree distribution and is used to assess possible influences of this slope on the results.

In the Barabási-Albert model, the network starts with $m+1$ completely connected nodes and grows by the successive addition of single nodes with m connections established with the older nodes, chosen according to the preferential attachment rule [14]. The growth stops when the desired number of nodes N is reached. The generalized Barabási-Albert model is parametrized by two probabilities: the rewiring probability ρ and the linking probability λ, such that $\rho+\lambda<1$. At each step, one of the following actions is taken: With probability $1-\rho-\lambda$, a new node is inserted (as for the basic Barabási-Albert model); with probability ρ, m links are rewired by chosing a node at random, disconnecting it from one of its neighbors, and reconnecting to another node with preferential attachment; with probability λ, m links are added chosing a node at random and a new neighbor for it with preferential attachment (see Ref. [15]). The tail of the resulting degree distribution follows a power law as $k^{-\gamma}$, with

$$\gamma = 1 + \frac{2m(1-\rho) + 1 - \lambda - \rho}{m}. \tag{1}$$

To generate the Erdös-Rényi network, we start with N isolated nodes and insert L links connecting pairs of nodes chosen with uniform probability, avoiding self- and duplicate connections; for comparison with the Barabási-Albert model, we choose L so that $m=L/N$ is the same as the m values used for the Barabási-Albert model.

For the two-dimensional lattices, the N nodes are distributed in a square and the connections are established between neighboring nodes in the lattice. Afterwards, additional connections can be incorporated between uniformly random chosen pairs of nodes until a desired number of average additional links per node is included. This kind of randomly augmented regular network is similar to that used in the Newman and Watts small-world model [17].

B. Opinion model

For a given network with N voters (nodes), we start by distributing the C candidates among randomly chosen nodes (with uniform probability), that is, each candidate is assigned to just one node in the network (this reflects the fact that the candidates are also voters). The remaining voters start as "undecided," meaning that they have no favorite candidate yet. The following process is subsequently repeated a total of SN times: choose at random a voter i that already has an associated candidate c_i; for *all* neighbors of voter i, if they have no associated candidate (i.e., are as yet undecided), they are associated with candidate c_i, otherwise they change to candidate c_i with a given *switching probability* p. The constant S introduced above is henceforth called the *number of steps* of the algorithm (average number of interactions of each node). This opinion model is motivated by the following assumptions: (i) undecided voters are passive, in the sense that they do not spread their lack of opinion to other voters; (ii) undecided voters are easily convinced by interaction with someone that already has a formed opinion; (iii) the flexibility to change opinions due to an interaction, quantified by the parameter p, is the same for all voters. Despite the many limitations which can be identified in these hypotheses, they seem to constitute a good first approximation and can be easily generalized in future works.

This model is similar to a simple spreading to unoccupied sites, and can be reduced to an asynchronous spreading if the switching probability is zero. In spite of its simplicity, the model yields interesting results, as discussed below.

III. RESULTS

In the following, we present and discuss the histograms expressing the number of candidates with a given number of

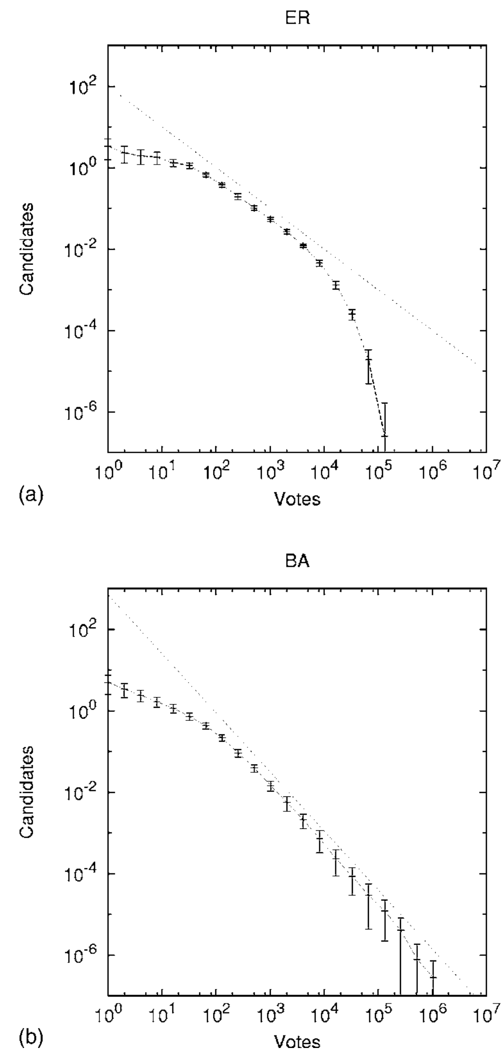

FIG. 1. Distribution of candidates with a given number of votes after 30 steps for networks with 2 000 000 voters, 1000 candidates, 5 links per node, and a switching probability of 0.1, on the left-hand side for Erdös-Rényi and on the right-hand side for Barabási-Albert networks. Error bars show one standard deviation; the lines show power laws of slope −1 (in the ER plot) and −1.45 (in the BA plot).

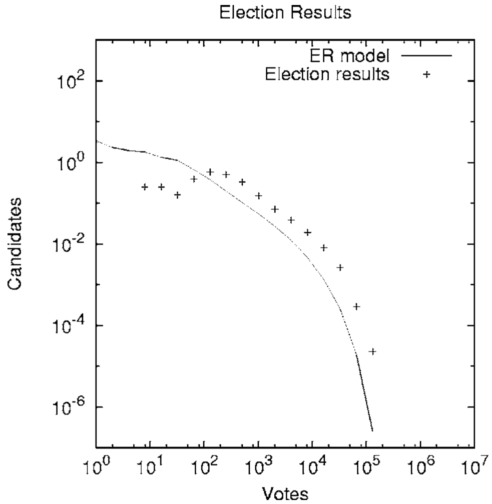

FIG. 2. Comparison of model (using ER networks) and real election results. Experimental data refer to state deputy elections in the São Paulo state, Brazil, in 1998, with 23 321 034 voters. Model results are for the same parameters as in Fig. 1.

nodes. The plots are in logarithmic scale, and the bin size doubles from one point to the next in order to provide uniformity. The number of candidates in a bin are normalized by the bin size. All results correspond to mean values obtained after 30 different realizations of the model with the given parameters.

As becomes clear from an analysis of the following figures, larger values of N/C tend to lead to more interesting results, motivating the adoption of large N and small C. The use of too large values of N implies a high computational and memory cost; the use of too small values of C leads to poor statistics implied by the large variations in the number of candidates inside the bins. The standard values of $N=2000\,000$ and $C=1000$ adopted in the following represent a good compromise considering our computational resources.

Figure 1 shows the results of the simulation for Erdös-Rényi and Barabási-Albert networks after 30 steps and with a switching probability of 0.1. The result for the Erdös-Rényi network is similar to results of real elections [9]. There is a power law regime for an intermediate number of votes, a plateau for a small number of votes, and a cutoff for a large number of votes; the power law regime has an exponent of -1, which is almost the same as that obtained for real elections [9]. The large variability on the plateau region is also consistent with the differences found at this part of the curves when considering different election outcomes (see, for example, the data in Ref. [10]). A direct comparison of the model with experimental election results is shown in Fig. 2, which presents the distribution of the number of votes for the 1998 Brazilian elections for state deputies in the São Paulo state [18]. Model parameters are the same as for Fig. 1. Other election results display a similar behavior (see, e.g., figures in Ref. [10]).

For the Barabási-Albert model, although two power law regimes with different exponents can be identified, neither corresponds to the experimental value of -1; the tail of the curve follows a power law with slope ≈ -1.45.

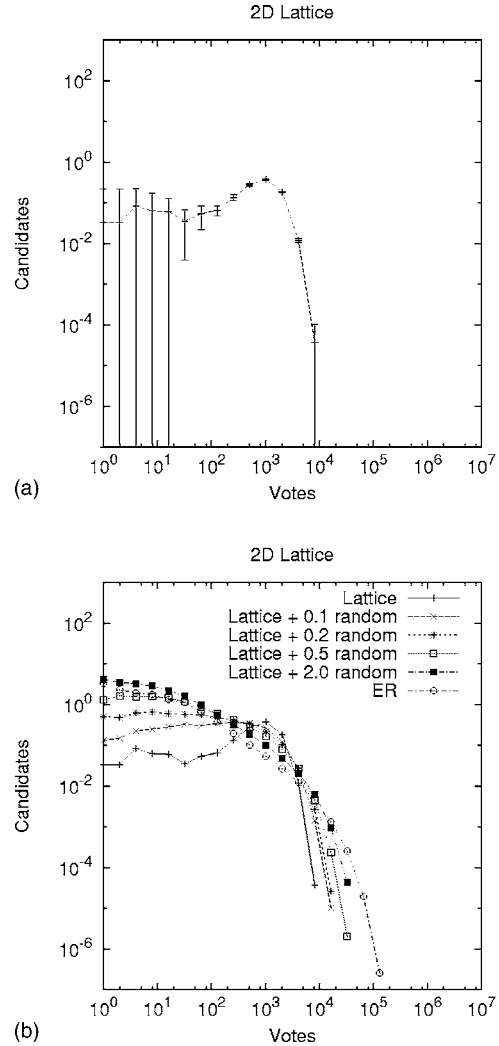

FIG. 3. Distribution of candidates with a given number of votes after 30 steps for two-dimensional lattices with 2 000 000 voters, 1000 candidates, 5 links per node, and a switching probability of 0.1, on the left-hand side for a pure lattice (error bars show one standard deviation) and on the right-hand side for lattices with the addition of the given average number of shortcut links per node between randomly selected nodes. The result for the Erdös-Rényi network is also shown for comparison.

The left-hand side of Fig. 3 shows the result for the simulation on a two-dimensional lattice. There is no sign of a power law regime and a clear peak around 1000 votes can be noted, in disagreement with the scale-free nature of the experimental results. On the right-hand side of the same figure, the effect of adding random connections to the lattice can be easily visualized. It is remarkable that the addition of just a small number of new links (about half the number of nodes) is enough to get a result similar to the one of the Erdös-Rényi model. It is a known fact [16] that a small number of random links in a regular network are enough for the emergence of the "small world" phenomenon. By enabling a candidate to reach the whole network of voters in a small number of steps, this phenomenon increases the chance of a candidate getting a very large number of votes, therefore broadening the distribution.

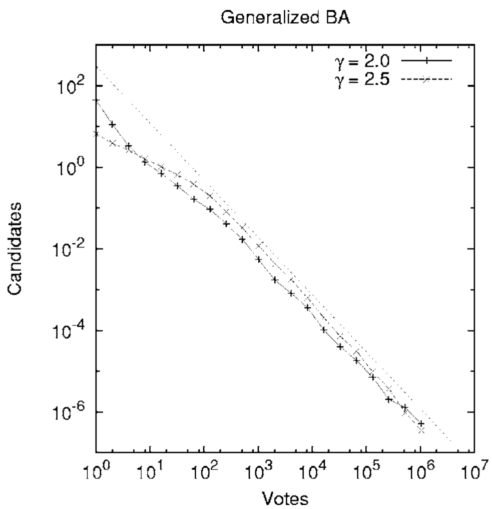

FIG. 4. Distribution of candidates with a given number of votes after 30 steps for the generalized Barabási-Albert model with slopes of the degree distribution of −2 and −2.5, with 2 000 000 voters, 1000 candidates, an average of 5 links per node, and a switching probability of 0.1. The straight line shows a power law with slope −1.4.

To verify if the slope of the distribution of candidates with a given number of votes is influenced by the slope of the degree distribution for power law networks, we used the generalized Barabási-Albert model. The original Barabási-Albert model has a slope of −3. We choose therefore two sets of parameters that give slopes of −2 and −2.5, respectively, and at the same time have identical average degrees. The values of the parameters are $m=1$, $\rho=6/11$, $\lambda=4/11$ giving $\gamma=2$ and $m=4$, $\rho=7/22$, $\lambda=3/22$ giving $\gamma=2.5$. As can be seen from Fig. 4, the slope of the power law of the distribution of votes is almost unaffected by the slope of the degree distribution. In fact, fitting for a power law in the region of the number of votes between 100 and 200 000, we get for the original Barabási-Albert model $(\gamma=3)$ a slope of −1.45, for the generalized model with $\gamma=2.5$ a slope of −1.44, and for $\gamma=2$ a slope of −1.40.

Now we turn our attention to the influence of the parameters of the model. In Fig. 5 the effect of changing the number of candidates while keeping the other parameters fixed is shown. For the Erdös-Rényi model, the effect of increasing the number of candidates translates itself as an upward shift of the curve while, at the same time, the cutoff is shifted to the left. This is an expected result: as the number of candidates grows with a fixed number of voters, the candidates are initially distributed closer to one another in the network, and have therefore fewer opportunities to spread influence before hitting a voter already with an opinion; this leads to a cutoff in a smaller number of votes and in an increase in the number of candidates with fewer votes than the cutoff. In the Barabási-Albert model, the behavior for a small number of votes is similar: the curve is shifted up; but for the power law regime of a large number of votes, the curve decays more steeply as more candidates are added.

Changing the number of voters has an impact limited almost exclusively to the tail of the curves, as seen in Fig. 6.

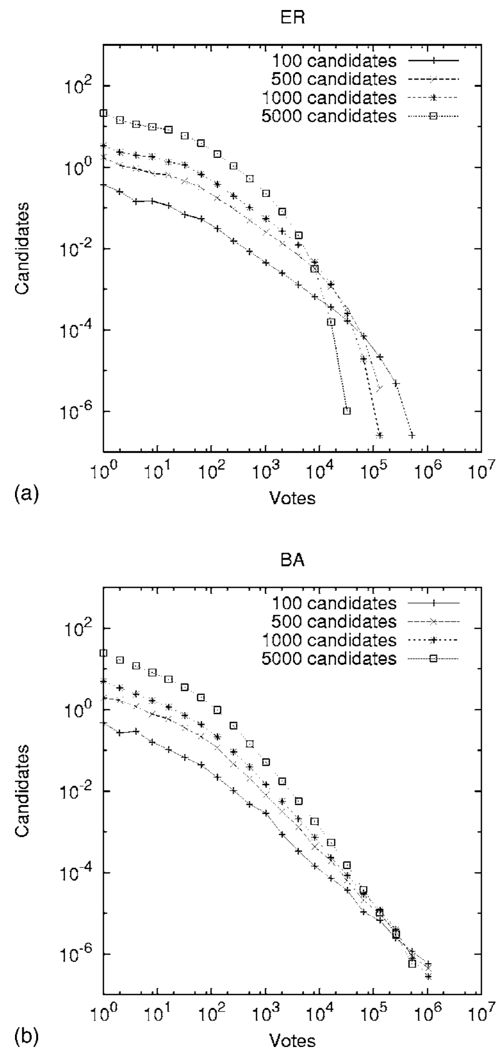

FIG. 5. Effect of the number of candidates. Distributions after 30 steps for networks with 2 000 000 voters, 5 links per node, a switching probability of 0.1, and different numbers of candidates, on the left-hand side for Erdös-Rényi and on the right-hand side for Barabási-Albert networks.

When the number of voters is increased, in the Erdös-Rényi model, the cutoff is shifted to the left and the power law regime is correspondingly increased. In the Barabási-Albert model, the maximum number of votes is shifted and the inclination of the second power law regime is changed to accommodate this displacement. Comparing with Fig. 5, we see that the tail of the curve for the Barabási-Albert model adapts its inclination according to the relation between the number of voters and candidates, i.e., a larger value of N/C implies a flatter tail.

From Fig. 7 we can see that the behavior that is being discussed appears only if the network is sufficiently connected: for $m=1$ there is no power law regime for the Erdös-Rényi model and the behavior for the Barabási-Albert model is complex, with three different regions and a peak for a small number of votes. Also for this latter model, the inclination of the tail of the curve appears to be slightly influ-

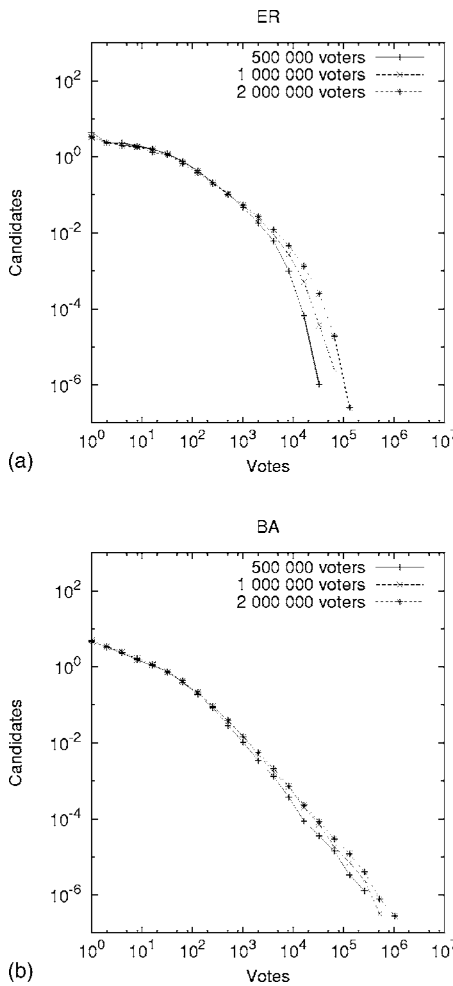

FIG. 6. Effect of the number of voters. Distributions after 30 steps for networks with 1000 candidates, 5 links per node, a switching probability of 0.1, and different numbers of voters, on the left-hand side for Erdös-Rényi and on the right-hand side for Barabási-Albert networks.

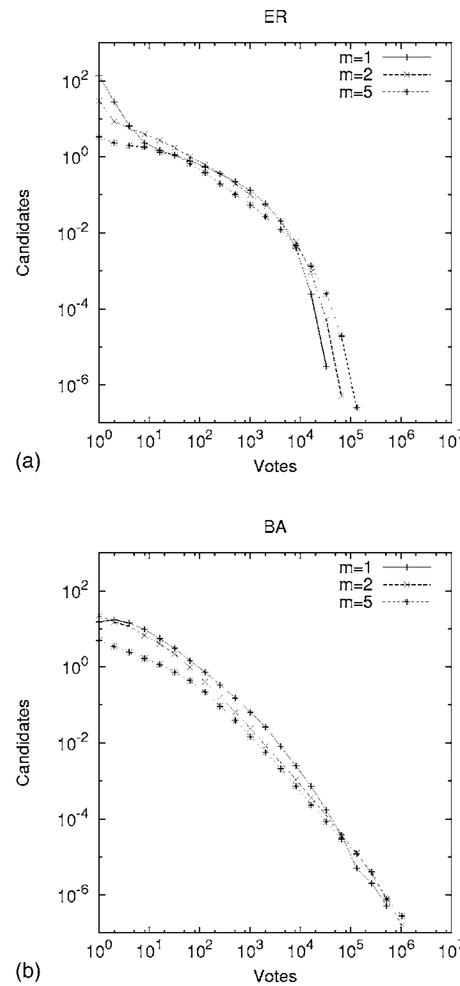

FIG. 7. Effect of the number of links. Distributions after 30 steps for networks with 2 000 000 voters, 1000 candidates, a switching probability of 0.1, and different numbers of links per node, on the left-hand side for Erdös-Rényi and on the right-hand side for Barabási-Albert networks.

enced by the average connectivity, with steeper tails for smaller connectivities.

The switching probability has an effect only on the first part of the curve, as can be seen from Fig. 8. In both models, this part of the curve is shifted down as the probability increases and its range is extended until it touches the original (for zero probability) curve. Note that the inclination of the Barabási-Albert curve corresponding to a small number of votes is maintained for the different values of switching probability (but is different for zero probability).

A similar effect has been obtained while changing the number of steps (Fig. 9). As the number of steps is increased, the curve remains unchanged for a large number of votes, but is downshifted for a small number of votes. The similarity between an increase in the number of steps and an increase in switching probability is easily explained: after all voters have a candidate, changes occur only by switching candidates. In other words, increasing the number of steps increases the number of times a switching is tried, resulting in a similar effect as increasing the switching probability.

IV. CONCLUSIONS

We suggested and studied a simple voting model based on the spreading of opinions through the links of a network. The results of the simulation of the model show a remarkable qualitative agreement with experimental results for proportional voting in Brazilian elections [9] when the network model used is of Erdös-Rényi type or a lattice with sufficient random shortcuts added. In these networks, the model results in a power law distribution with an exponent of -1, but with a shortcut for a large number of votes and a plateau for a small number of votes, as observed in real elections. The "small world" effect appears to be of central importance in this result, as the result for a lattice without shortcuts is very different, without any power law regime.

Interestingly, the Barabási-Albert network model gives results that are not consistent with the considered real elections. For the BA model, two power law regimes have been identified, while the overall curve presents no clear cutoff. The second (and dominant) power law regime is not universal, depending on the number of links per node in the net-

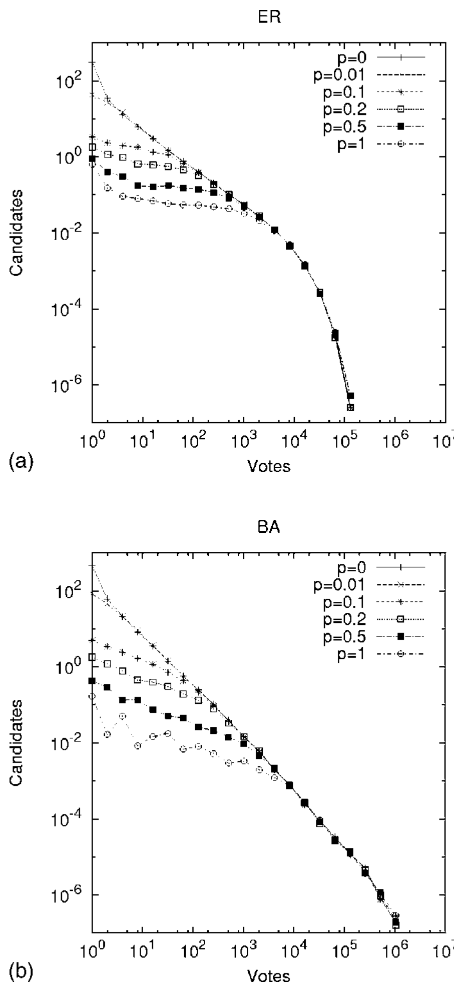

FIG. 8. Effect of the swtiching probability. Distributions after 30 steps for networks with 2 000 000 voters, 1000 candidates, 5 links per node, and different values for the switching probability, on the left-hand side for Erdös-Rényi and on the right-hand side for Barabási-Albert networks.

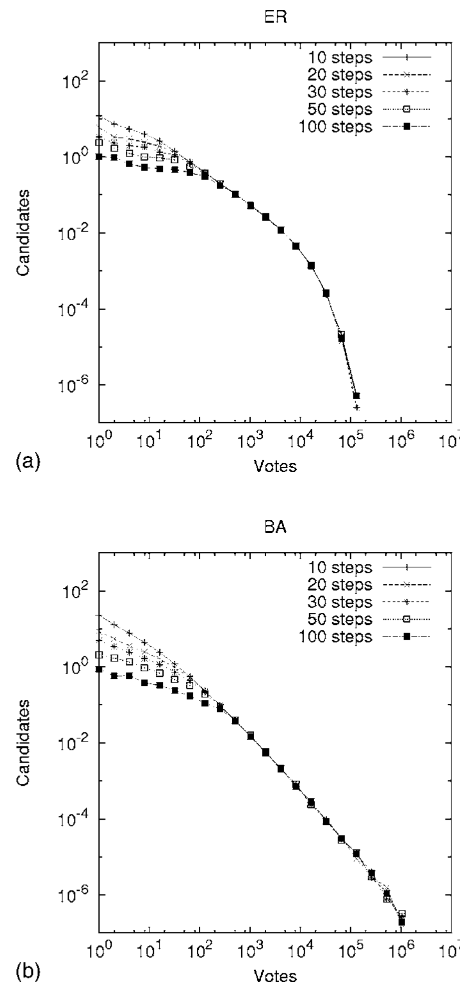

FIG. 9. Effect of the number of steps. Distributions for networks with 2 000 000 voters, 1000 candidates, 5 links per node, a switching probability of 0.1, and different total numbers of steps, on the left-hand side for Erdös-Rényi and on the right-hand side for Barabási-Albert networks.

work and the relation between the number of voters and the number of candidates. Changes on the slope of the power law of the degree distribution of the network (in the interesting region between −2 and −3) have no influence in the slope of the second power law regime. Also the first power law regime cannot be characterized by the experimental value of −1. This is somewhat puzzling, as many social networks have power law degree distributions [7] and are in this respect more closely related to the Barabási-Albert model than to the other two investigated models. Nevertheless, there is no reason to suppose that the contact networks underlying the dynamics of opinion formation should be characterized as a power law. Instead, the results presented suggest that these networks are characterized by a homogeneous distribution of connectivity among the nodes. Clustering and communities are frequent properties of social networks, but are not present in the studied network models. The influence of these factors in the election results deserves further investigation.

ACKNOWLEDGMENTS

L. da F. Costa is grateful to CNPq (Grant No. 308231/03-1) for financial assistance.

[1] R. Forsythe, T. Rietz, R. Myerson, and R. Weber, Int. J. Game Theory **25**, 355 (1996).
[2] H. E. Stanley, L. A. N. Amaral, D. Cuning, P. Gopikrishnan, Y. Lee, and Y. Liu, Physica A **269**, 156 (1999).
[3] D. Stauffer, Int. J. Mod. Phys. C **11**, 1081 (2000).
[4] D. Stauffer, Comput. Sci. Eng. **5**, 71 (2003).
[5] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[6] S. N. Dorogovtsev and J. F. F. Mendes, Adv. Phys. **51**, 1079

(2002).

[7] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).

[8] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, Phys. Rep. **424**, 175 (2006).

[9] R. N. Costa Filho, M. P. Almeida, J. S. Andrade, Jr., and J. E. Moreira, Phys. Rev. E **60**, 1067 (1999).

[10] M. L. Lyra, U. M. S. Costa, R. N. Costa Filho, and J. S. Andrade, Jr., Europhys. Lett. **62**, 131 (2003).

[11] A. T. Bernardes, D. Stauffer, and J. Kertesz, Eur. Phys. J. B **25**, 123 (2002).

[12] M. C. Gonzalez, A. O. Sousa, and H. J. Herrmann, Int. J. Mod. Phys. C **15**, 45 (2004).

[13] K. Sznajd-Weron and J. Sznajd, Int. J. Mod. Phys. C **11**, 1157 (2000).

[14] A.-L. Barabási and R. Albert, Science **286**, 509 (1997).

[15] R. Albert and A.-L. Barabási, Phys. Rev. Lett. **85**, 5234 (2000).

[16] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 409 (1999).

[17] M. E. J. Newman and D. J. Watts, Phys. Rev. E **60**, 7332 (1999).

[18] The data used to compute this results are available through the download section of the site http://www.tse.gov.br/