

**Social influence with recurrent mobility and multiple options**

Jérôme Michaud\*

*Department of Physics and Astronomy, University of Uppsala, 752 37 Uppsala, Sweden  
and Department of Sociology, University of Uppsala, 751 20 Uppsala, Sweden*

Attila Szilva

*Department of Physics and Astronomy, University of Uppsala, 752 37 Uppsala, Sweden*

(Received 30 January 2018; revised manuscript received 4 May 2018; published 22 June 2018)

In this paper, we discuss the possible generalizations of the social influence with recurrent mobility (SIRM) model [*Phys. Rev. Lett.* **112**, 158701 (2014)]. Although the SIRM model worked approximately satisfying when U.S. election was modeled, it has its limits: it has been developed only for two-party systems and can lead to unphysical behavior when one of the parties has extreme vote share close to 0 or 1. We propose here generalizations to the SIRM model by its extension for multiparty systems that are mathematically well-posed in case of extreme vote shares, too, by handling the noise term in a different way. In addition, we show that our method opens alternative applications for the study of elections by using an alternative calibration procedure and makes it possible to analyze the influence of the “free will” (creating a new party) and other local effects for different commuting network topologies.

DOI: [10.1103/PhysRevE.97.062313](https://doi.org/10.1103/PhysRevE.97.062313)**I. INTRODUCTION**

There is an extensive literature on the applications of statistical physics in social modeling [1]. One of the subdomains in which statistical physics models have been used is opinion dynamics and voting behavior. In this context, the voter model (VM) introduced by Ref. [2] and named “voter model” in Ref. [3] had a huge influence and many scholars have explored its dynamics in various settings. The VM has been studied on regular lattices [4,5], where an exact solution to the fixation probability exists in any dimension, or random networks [6–9], and various modifications of it have been proposed to make it more realistic [10–15]. However, in most of these works, the VM has been analyzed from a mathematical point of view and the results are supported by computer simulations based on modeled data in the best cases. The validation of these models against real (election) data is usually not provided; see, for example, Refs. [16,17].

A shining exception is represented by Ref. [18] (“Is the voter model a model for voters?”) where a variation of the VM has been implemented and tested on voting data from U.S. elections. In this paper, the authors proposed a modification of the VM that they called *social influence with recurrent mobility*, abbreviated as SIRM. The results of Ref. [18] show potential practical applications of the VM that have not been shown before, which brings the VM in the focus of basic research, again. The limitations of the applicability of the SIRM model call for further scientific investigations with challenges in analytical derivations and numerical simulations.

The question “Is the voter model a model for voters?” still needs to be addressed, especially in extreme cases when the

vote share is close to 0 or 1. In the SIRM model, the network underlying the VM is inferred from the commuting pattern of voters. A commuter being an individual living in one place and working in another. Using census data, it is possible to construct the commuting network of the population, which is defined at the community level rather than at the individual level, assuming that all members of a community are connected together. In the application to the US election [18], counties were chosen as the relevant communities and two counties are linked by an edge if at least one individual lives in one the counties and works in the other. The edge is weighted by the total number of commuters. Furthermore, two parameters have been added to the standard VM. The first parameter,  $\alpha$ , controls the probability that an individual interacts where she lives or where she works. The second parameter,  $D$ , is a noise parameter accounting for any other factors that can influence the change of opinion, such as free will or mass media. Using this model, it was possible to recover some statistical features of U.S. presidential elections, such as the shape of the distribution of vote shares around the average and the logarithmic decay of the spatial correlations of vote shares [18].

As will be shown, it turns out that the SIRM model is not mathematically well-posed and that it can predict unphysical behavior of the population if one of the parties has a vote share that is either close to 0 or to 1. In the case of the US election, the importance of this issue is small, since the two parties considered share the votes almost equally. However, if we would like to generalize the SIRM model to an arbitrary number of parties, this issue becomes critical to avoid negative number of voters.

In this paper, we study the SIRM model and its extension from a mathematical point of view. The purpose of this paper is to provide a mathematically sound generalization to the SIRM model in the case of multiparty elections and where

\*jerome.michaud@soc.uu.se

the corresponding vote share can become as small or large as possible (close to 0 or 1). In the SIRM model, probabilities to change opinion are stochastic quantities and the main challenge consists in constructing a perturbed probability distribution from a given one. This will be done by using the Dirichlet distribution. The procedure outlined here is not restricted to the Voter Model and could be applied to various (socioeconomic) systems where a quantity defined on a simplex should be perturbed without leaving the simplex. To our knowledge, the method proposed in this paper is the first to address this issue.

We show that the original formulation of the SIRM model can be recovered as a limit of our model under specific assumptions that are valid in the case of U.S. elections. Furthermore, to illustrate how the model can be applied to real election data, we propose a calibration procedure to set the values of the parameters. The calibration procedure will be applied to a synthetic network and to an actual commuting network for illustration. We would like to underline the fact that VM-like models have fairly strong assumptions with respect to the behavior of the voters. Some of these assumptions are unlikely to be true, but we nevertheless expect to be able to capture some statistical regularities in elections using this model. Such regularities have been exemplified in Refs. [19–35]. The calibration procedure applied on the actual commuting network demonstrates that the stationarity of some properties of the vote share distribution can be recovered by the model. The analysis of predictions of our model and of spatial correlations in vote shares is outside the scope of this paper.

The remainder of this paper is organized as follows. In Sec. II, we present general requirements for the definition of a VM-like model based on a commuting network. In Sec. III, we recall the SIRM model as presented in [18]. In Sec. IV, we present our generalized version of the SIRM model and show that the SIRM model can be recovered as a limit of our model. In Sec. V, we illustrate our model with numerical experiments on synthetic input data and on a real situation, namely elections in the county of Stockholm in Sweden. We also illustrate a possible calibration procedure to set the parameters of our model. A concluding discussion is provided in Sec. VI.

## II. GENERALIZED VOTER MODEL WITH RECURRENT MOBILITY

In this section, we discuss the general form that a VM-like model on the commuting network should have to be mathematically well-posed. To do so, we start by specifying what we understand as the commuting network and then introduce the relevant opinion space. Finally, we specify the dynamics that any *generalized voter model with recurrent mobility* (GVMRM) model should have.

### A. Commuting network

Consider a population of  $N$  individuals divided into  $M$  regions, that we call municipalities. This regions usually are electoral regions, such as counties, municipalities or states. Let  $N_{ij}$  be the number of commuters between municipality  $i$  and  $j$ , that is the number of individuals living in municipality  $i$  and working in municipality  $j$ . The commuting network is constructed by taking the  $M$  municipalities as nodes and

creating weighted edges of weight  $N_{ij}$ . The commuting network is a directed weighted network with self-loops, since there are people who live and work in the same place. In the rest of the paper, we will refer to the sub-population  $N_{ij}$  as the commuting cell  $ij$ .

From the  $N_{ij}$  quantities, one can construct the number  $N_i$  of people living in municipality  $i$  by summing over the second index and the number  $N'_j$  of people working in municipality  $j$  by summing over the first index. We therefore defined

$$N_i := \sum_j N_{ij} \quad \text{and} \quad N'_j := \sum_i N_{ij}. \quad (1)$$

The total population  $N = \sum_i N_i = \sum_j N'_j$ . The primed quantities refer to the working population.

### B. Opinion structure and vote shares

Let us assume that the individual can choose between  $K$  different opinions. We denote by  $V_{ij}^k$  the number of people in commuting cell  $ij$  that have opinion  $k$ . For consistency, we must have

$$\sum_{k=1}^K V_{ij}^k = N_{ij}. \quad (2)$$

The state of the model is fully defined by the quantities  $V_{ij}^k$ . For convenience, we also introduce the vote shares

$$v_{ij}^k := \frac{V_{ij}^k}{N_{ij}}. \quad (3)$$

With this definition one can easily show that

$$v_i^k = \sum_j \frac{N_{ij}}{N_i} v_{ij}^k \quad \text{and} \quad v_j^k = \sum_i \frac{N_{ij}}{N'_j} v_{ij}^k \quad (4)$$

are the vote shares of opinion  $k$  for the population living in municipality  $i$  and for the population working in municipality  $j$ .

### C. Dynamics

Let us now define the transition operators  $R_{ij}^{kk'}$

$$R_{ij}^{kk'} = P[(V_{ij}^k, V_{ij}^{k'}) \rightarrow (V_{ij}^k - 1, V_{ij}^{k'} + 1)] = \frac{N_{ij}}{N} v_{ij}^k p_{ij}^{k \rightarrow k'}, \quad (5)$$

that defines the probability that an individual changes from opinion  $k$  to opinion  $k'$  in the commuting cell  $ij$ . For the left most term, the first factor is the probability to choose an individual in the commuting cell  $ij$ ; the second factor is the probability that this individual holds opinion  $k$  and the third factor is the probability that she changes opinion to  $k'$ .

In general, any variation of the voter model defined on a commuting network can be defined by rates of the form

$$R_{ij}^{kk'} = \frac{N_{ij}}{N} v_{ij}^k p_{ij}^{k \rightarrow k'}, \quad (6)$$

where  $\mathbf{p}_{ij}^{k \rightarrow} = (p_{ij}^{k \rightarrow 1}, \dots, p_{ij}^{k \rightarrow K})$  is the probability distribution to transition from any opinion  $k$  to other opinions in commuting cell  $ij$  [36].

In principle, any GVMRM model can be defined through Eq. (6) by specifying  $\mathbf{p}_{ij}^{k \rightarrow k'}$ .

### III. THE SIRM MODEL

In this section, we reformulate the SIRM model in terms of (transition, raising, and lowering) operators as done in Ref. [9] and recall the analysis that has been performed in Ref. [18].

#### A. Formulation

The SIRM model has been defined for two opinions. Let the two opinions be referred to as  $\mathbf{0}$  and  $\mathbf{1}$ . The fundamental quantities of the models are given by

$$V_{ij} := V_{ij}^{\mathbf{1}} \quad \text{and} \quad N_{ij} - V_{ij} := V_{ij}^{\mathbf{0}} \quad (7)$$

and the corresponding vote shares

$$v_{ij} := v_{ij}^{\mathbf{1}} \quad \text{and} \quad 1 - v_{ij} := v_{ij}^{\mathbf{0}}. \quad (8)$$

The vote share at home and at work can be computed through the expressions

$$v_i = \sum_j \frac{N_{ij}}{N_i} v_{ij} \quad \text{and} \quad v'_j = \sum_i \frac{N_{ij}}{N'_j} v_{ij}. \quad (9)$$

To fully define the SIRM model we need to specify the dynamics, that is the four operators  $R_{ij}^{kk'}$ , for  $k, k' \in \{\mathbf{0}, \mathbf{1}\}$ . Since the operators  $R_{ij}^{kk'}$  gives the probability to change opinion from  $k$  to  $k'$ , we must have

$$R_{ij}^{\mathbf{0},\mathbf{0}} = 1 - R_{ij}^{\mathbf{0},\mathbf{1}} \quad \text{and} \quad R_{ij}^{\mathbf{1},\mathbf{1}} = 1 - R_{ij}^{\mathbf{1},\mathbf{0}}. \quad (10)$$

Therefore, it is sufficient to only specify  $R_{ij}^{\mathbf{0},\mathbf{1}}$  and  $R_{ij}^{\mathbf{1},\mathbf{0}}$  to fully define the model. In [18], they called  $R_{ij} := R_{ij}^{\mathbf{0},\mathbf{1}}$  the raising operator and  $L_{ij} := R_{ij}^{\mathbf{1},\mathbf{0}}$  the lowering operator. With this notation, the SIRM model is defined by

$$\begin{aligned} R_{ij} &= \frac{N_{ij}}{N} \left\{ (1 - v_{ij}) [\alpha v_i + (1 - \alpha) v'_j] + \frac{D}{2} \eta_{ij}^+ \right\}, \\ L_{ij} &= \frac{N_{ij}}{N} \left\{ v_{ij} [\alpha(1 - v_i) + (1 - \alpha)(1 - v'_j)] + \frac{D}{2} \eta_{ij}^- \right\}, \end{aligned} \quad (11)$$

where  $\eta_{ij}^{+,-}$  are independent Gaussian white noises. The parameter  $\alpha$  controls the ratio of interaction at home and at work and  $D$  controls the magnitude of the noise.

By the definition of the raising and lowering operators, it is clear that for a well-posed model we need

$$0 \leq R_{ij}, L_{ij} \leq 1. \quad (12)$$

For extreme values of  $v_{ij}$  it is obvious that Eq. (11) does not satisfy Eq. (12). For example, if  $v_{ij} = 0$  and  $\eta_{ij}^- < 0$ , then  $L_{ij} < 0$ . We can conclude that the SIRM model is ill-posed and the issue originates in the noise term used. Remark that if  $D = 0$ , then Eq. (11) is of the form of Eq. (6) and the model is well-posed.

#### B. Analysis

In this subsection, we consider the case of non-extreme vote shares as it has been developed for the U.S. election case and recall the analysis performed in Ref. [18]. In this case, the problem with the noise mentioned before can be ignored.

For clarity, we rewrite the raising and lowering operators as functions of  $v_{ij}$ , that is

$$R(v_{ij}) := R_{ij} \quad \text{and} \quad L(v_{ij}) := L_{ij}. \quad (13)$$

Using these operators, we can write the corresponding balance equation or master equation

$$\begin{aligned} p(v_{ij}, t + \delta t) &= R(v_{ij} - \delta v_{ij}) p(v_{ij} - \delta v_{ij}, t) \\ &\quad + L(v_{ij} + \delta v_{ij}) p(v_{ij} + \delta v_{ij}, t) \\ &\quad + [1 - R(v_{ij}) - L(v_{ij})] p(v_{ij}, t), \end{aligned} \quad (14)$$

where  $\delta t = \frac{1}{N}$  and  $\delta v_{ij} = \frac{1}{N_{ij}}$ .

Expanding this equation to second order gives a Fokker-Planck equation,

$$\begin{aligned} \frac{\partial}{\partial t} p(v_{ij}, t) &= -\frac{\partial}{\partial v_{ij}} [\hat{d}(v_{ij}) p(v_{ij}, t)] \\ &\quad + \frac{1}{2} \frac{\partial^2}{\partial v_{ij}^2} [\hat{D}(v_{ij}) p(v_{ij}, t)], \end{aligned} \quad (15)$$

where the *drift* coefficient,  $\hat{d}(v_{ij})$ , and the *diffusion* coefficient,  $\hat{D}(v_{ij})$ , are given by

$$\begin{aligned} \hat{d}(v_{ij}) &= \frac{\delta v_{ij}}{\delta t} [R(v_{ij}) - L(v_{ij})], \\ \hat{D}(v_{ij}) &= \frac{\delta v_{ij}^2}{\delta t} [R(v_{ij}) + L(v_{ij})]. \end{aligned} \quad (16)$$

Equivalently, we can obtain a stochastic differential equation (SDE)/Langevin equation of the form

$$dv_{ij} = \hat{d}(v_{ij}) dt + \sqrt{\hat{D}(v_{ij})} dW_{ij}^*(t), \quad (17)$$

where  $dW_{ij}(t)$  is a white noise. There is a simple correspondence between  $dW_{ij}$  and  $\eta_{ij}$  given by

$$dW_{ij}^\circ = \eta_{ij}^\circ dt, \quad (18)$$

where  $\circ$  could be  $*, ', \dots$ ; see Ref. [37, p. 90].

For the raising and lowering operators of the SIRM model, we obtain

$$\begin{aligned} \hat{d}(v_{ij}) &= \alpha v_i + (1 - \alpha) v'_j - v_{ij} + \frac{D}{\sqrt{2}} \eta_{ij}, \\ \hat{D}(v_{ij}) &= \frac{1}{N_{ij}} \left[ (1 - 2v_{ij}) [\alpha v_i + (1 - \alpha) v'_j] + v_{ij} + \frac{D}{\sqrt{2}} \eta'_{ij} \right], \end{aligned} \quad (19)$$

where

$$\begin{aligned} \frac{1}{\sqrt{2}} \eta_{ij} &= \frac{1}{2} (\eta_{ij}^+ - \eta_{ij}^-), \\ \frac{1}{\sqrt{2}} \eta'_{ij} &= \frac{1}{2} (\eta_{ij}^+ + \eta_{ij}^-), \end{aligned} \quad (20)$$

since Gaussian white noises add as standard deviations of normally distributed variables.

We would like to stress that both the drift coefficient,  $\hat{d}(v_{ij})$ , and the diffusion coefficient,  $\hat{D}(v_{ij})$ , are stochastic functions because of the presence of  $\eta$  terms. This is a consequence of working with stochastic rate of change.

**IV. THE GENERALIZED SIRM MODEL**

In this section, we provide a generalization of the SIRM model (as a possible realization of the GVMMR model) that (i) can model any fixed number of opinions and (ii) has well-posed stochastic rates  $R_{ij}^{kk'}$ . We will show that the SIRM model can be recovered as a limit of our model when the vote share is far from the extreme cases (0 or 1).

**A. Generalization to  $K$  opinions**

Let us start the discussion of the extension to  $K$  opinions by considering deterministic rates. Under this assumption, we can define  $r_{ij}^{k \rightarrow k'}$  to be the probability that an individual holding opinion  $k$  will change to opinion  $k'$  in the commuting cell  $ij$ . The natural generalization to  $K$  opinions is given by defining

$$r_{ij}^{k \rightarrow k'} := \alpha v_i^{k'} + (1 - \alpha) v_j^{k'}. \tag{21}$$

The probabilities  $r_{ij}^{k \rightarrow k'}$  are a specific instance of  $p_{ij}^{k \rightarrow k'}$  in Eq. (6). This probability does not depend on the current option  $k$ , since the copying process is unconditional. We keep the  $k$  index, because the noise term that will be added will be  $k$ -dependent. For consistency, we have

$$\sum_{k'} r_{ij}^{k \rightarrow k'} = 1, \quad \forall i, j, k, \tag{22}$$

since  $r_{ij}^{k \rightarrow k'}$  forms a probability distribution.

**B. Noise handling**

We now discuss the different possible options to add noise on the rates  $r_{ij}^{k \rightarrow k'}$ , while keeping the probability distribution property. We discuss various ways to modify the rates both in a deterministic way and in a stochastic way to construct general noisy rates. The way noise has been added to the VM mainly relies on a deterministic modification of the rates [38,39].

**1. Adding free will**

One possibility to add a kind of noise is to introduce “free will.” Free will can be encoded in a fixed probability to randomly change state. Let us assume that there are  $K$  options, then one can choose the uniform probability to change opinion.

One can add free will to the SIRM model without noise by redefining  $R^{kk'}(\mathbf{v}_{ij})$  as

$$R^{kk'}(\mathbf{v}_{ij}) = \frac{N_{ij}}{N} v_{ij}^k \left[ (1 - \gamma) r_{ij}^{k \rightarrow k'} + \gamma \frac{1}{K} \right], \tag{23}$$

where  $\gamma$  controls the strength of the free will term. In the general notation introduced in Eq. (6), we have

$$p_{ij}^{k \rightarrow k'} = (1 - \gamma) r_{ij}^{k \rightarrow k'} + \gamma \frac{1}{K}. \tag{24}$$

This modification is interesting, because it allows for opinions that went extinct to reappear in the system. This can also be used to simulate the creation of a new party. Without such a term, the number of party is strictly decreasing. To simulate real life elections, such a noise (not necessarily uniform) would be needed.

**2. Adding intra-commuting-cell influence**

In the context of GVMMR models, the commuting pattern is important. One could assume that agents living in the same commuting cell interact more often than people who do not. According to Refs. [40,41], people sharing more loci (place of living, place of work, etc.) have more influence on one another. We can take this effect into account by increasing the weight of other individuals in the same commuting cell. Similarly to what has been done for free will, we can introduce an additional probability to change from opinion  $k$  to  $k'$  in commuting cell  $ij$  proportional to  $v_{ij}^{k'}$ . The model can now be modified through

$$R^{kk'}(\mathbf{v}_{ij}) = \frac{N_{ij}}{N} v_{ij}^k [(1 - \beta) r_{ij}^{k \rightarrow k'} + \beta v_{ij}^{k'}], \tag{25}$$

where  $\beta$  controls the strength of the local interaction term.

In principle, any noise originating in a probability distribution over different opinions can be added in this way. Note that most of the time, the form of the probability distribution does not depend on the initial opinion  $k$  but only on the final opinion  $k'$  and might even be fully independent of opinions as in the case of free will. Furthermore, all these modifications are deterministic in the sense that they the transition rates and deterministic. This is a major difference between the original SIRM model and our formulation. The next step is to obtain in a consistent way stochastic versions of the rates  $R^{kk'}(\mathbf{v}_{ij})$ .

**3. Putting noise on a probability distribution:**

*The Dirichlet solution*

To construct stochastic rates in a consistent manner, we need to find a way to transform a probability distribution into another that is stochastically perturbed from the initial one. To do so, we rely on probability distributions over the simplex. We have a few available choices, such as the Dirichlet distribution, the multinomial distribution or the lognormal distribution. The multinomial distribution has been chosen in Ref. [42] in the context of language evolution to model finite length conversation and it turns out that the structure of the covariance matrix of the multinomial distribution is proportional to that of the Dirichlet distribution.

We choose to use the Dirichlet distribution, because (i) it is a continuous distribution (not discrete like the multinomial); (ii) the correspondence between the initial probability distribution and the parameter of the Dirichlet distribution is easy to define; (iii) it has well-defined moments that can be used in the analysis; and (iv) it has the aggregation property:

$$\begin{aligned} \text{If } (X_1, \dots, X_K) &\sim \mathcal{D}(\alpha_1, \dots, \alpha_K), \\ \text{then } (X_1, \dots, X_i + X_j, \dots, X_K) \\ &\sim \mathcal{D}(\alpha_1, \dots, \alpha_i + \alpha_j, \dots, \alpha_K). \end{aligned} \tag{26}$$

This means that the model is compatible with aggregation of opinions. Aggregation of parties leads to a consistent formulation, which would not be true when using other distributions, such as the lognormal distribution. These four properties ensures that the resulting model is consistent, well-posed and robust with respect to opinion aggregation.

Let us assume that a sample  $\mathbf{X}$  from the Dirichlet distribution of parameter  $\boldsymbol{\alpha}$  is given by  $\mathbf{X} \sim \mathcal{D}(\boldsymbol{\alpha})$ . If we define

$\bar{\alpha} = \sum_k \alpha_k$ , then we can write the expected value

$$\mathbb{E}(X_k) = \frac{\alpha_k}{\bar{\alpha}} \quad (27)$$

and the covariance matrix

$$\mathbf{C} = \frac{1}{\bar{\alpha}^2(1 + \bar{\alpha})}(\text{diag}(\boldsymbol{\alpha}) - \boldsymbol{\alpha}\boldsymbol{\alpha}^T). \quad (28)$$

Let us now assume that we want to put noise on a probability distribution  $\mathbf{p} = (p_1, \dots, p_K)$  with the consistency relation  $\sum_k p_k = 1$  and  $p_k \geq 0, \forall k$ . Then we can define a noisy probability distribution  $\tilde{\mathbf{p}}_{\tilde{D}} \sim \mathcal{D}(\mathbf{p}/\tilde{D})$ , where  $\mathcal{D}$  stands for the Dirichlet distribution. It is straightforward to verify that

$$\mathbb{E}(\tilde{p}_{\tilde{D},k}) = p_k \quad \text{and} \quad \mathbf{C} = \frac{\tilde{D}}{\tilde{D} + 1}(\text{diag}(\mathbf{p}) - \mathbf{p}\mathbf{p}^T). \quad (29)$$

When  $\tilde{D}$  is small the variance is small. When  $\tilde{D} \rightarrow \infty$ , then the prefactor tends to 1, which is the largest possible variance on the simplex.

This procedure to get a stochastic version of a probability distribution can be used to add stochasticity to the rates of the GVMRM. In the rest of this paper, we consider the following model:

$$p_{ij}^{k \rightarrow k'} = (1 - \beta - \gamma)r_{ij}^{k \rightarrow k'} + \beta\tilde{v}_{ij,\tilde{D}}^{k'} + \gamma/K, \quad (30)$$

where

$$\tilde{\mathbf{v}}_{ij,\tilde{D}} = \mathcal{D}(\mathbf{v}_{ij}/\tilde{D}). \quad (31)$$

We also require  $\beta + \gamma \leq 1$  and  $\beta, \gamma \geq 0$ . The parameter  $\beta$  represents the strength of the intra-commuting-cell influence as defined in Eq. (25), the parameter  $\gamma$  controls the intensity of free will as defined in Eq. (24) and the parameter  $\tilde{D}$  controls the level of noise associated with the intra-commuting-cell influence. One parameter,  $\beta$  for example, can be eliminated by rescaling the time; see Ref. [43].

We stress that the model defined through Eq. (30) includes both a free will term and an intra-commuting-cell term. Furthermore, we add some stochasticity to the intra-commuting-cell term to introduce stochastic rates. The idea to only add noise in the intra-commuting-cell term is motivated by the fact that such a noise will be local (uncorrelated between regions) and will not generate a drift term in the diffusion approximation. Those properties are shared by the original version of the SIRM model [18]. Other choices could have been made and we will not investigate them in this paper.

### C. The SIRM model as a limit

In the case of two variants with the transition probability defined in Eq. (30), we can apply the same procedure in Sec. III B and compute the drift and diffusion coefficients,  $\hat{d}(v_{ij})$  and  $\hat{D}(v_{ij})$ , respectively. We obtain the following terms:

$$\begin{aligned} \hat{d}(v_{ij}) = & (1 - \beta - \gamma)[r_{ij}^{0 \rightarrow 1} - v_{ij}] \\ & + \beta(\tilde{v}_{ij,\tilde{D}} - v_{ij}) \\ & + \gamma\left(\frac{1}{2} - v_{ij}\right) \end{aligned}$$

$$\begin{aligned} \hat{D}(v_{ij}) = & \frac{1}{N_{ij}} \left[ (1 - 2v_{ij})[(1 - \beta - \gamma)r_{ij}^{0 \rightarrow 1} + \beta\tilde{v}_{ij,\tilde{D}}] \right. \\ & \left. + (1 - \gamma)v_{ij} + \frac{\gamma}{2} \right], \end{aligned} \quad (32)$$

where

$$r_{ij}^{0 \rightarrow 1} = \alpha v_i + (1 - \alpha)v_j'. \quad (33)$$

Equations (32) are generalizations of Eqs. (19).

It is important to note that the term  $\tilde{v}_{ij,\tilde{D}}$  is a random variable, since it is sampled from a Dirichlet (in this case  $\beta$ ) distribution. In the following discussion, we will only consider the  $\hat{d}(v_{ij})$  term because the diffusion term,  $\hat{D}$ , in the original version of the SIRM model was neglected, since it is proportional to  $\frac{1}{N_{ij}}$  (and  $N_{ij}$  is usually large). To get our formulation in a similar form as Eq. (19), we can use the normal approximation of the Dirichlet distribution and approximate

$$\tilde{v}_{ij,\tilde{D}} \approx v_{ij} + \sqrt{\frac{\tilde{D}}{\tilde{D} + 1}} \sqrt{v_{ij}(1 - v_{ij})} \eta_{ij}, \quad (34)$$

where  $\eta_{ij}$  is a  $\delta$ -correlated white noise. Under this approximation, the second term in Eq. (32) turns into a pure noise term, and we have

$$\begin{aligned} \hat{d}(v_{ij}) = & (1 - \beta - \gamma)[r_{ij}^{0 \rightarrow 1} - v_{ij}] \\ & + \beta \sqrt{\frac{\tilde{D}}{\tilde{D} + 1}} \sqrt{v_{ij}(1 - v_{ij})} \eta_{ij} + \gamma\left(\frac{1}{2} - v_{ij}\right). \end{aligned} \quad (35)$$

If we now assume that  $\gamma = 0$  and  $v_{ij} \approx \frac{1}{2}$  so that the square root  $\sqrt{v_{ij}(1 - v_{ij})} \approx \frac{1}{2}$ , we get

$$\hat{d}(v_{ij}) = (1 - \beta)[r_{ij}^{0 \rightarrow 1} - v_{ij}] + \frac{\beta}{2} \sqrt{\frac{\tilde{D}}{\tilde{D} + 1}} \eta_{ij}. \quad (36)$$

If we then set  $\beta = \frac{1}{2}$ , we get

$$2\hat{d}(v_{ij}) = [r_{ij}^{0 \rightarrow 1} - v_{ij}] + \frac{1}{2} \sqrt{\frac{\tilde{D}}{\tilde{D} + 1}} \eta_{ij}, \quad (37)$$

which, up to factor 2 that can be eliminated by a rescale of time, we have a very similar expression as Eq. (19). Furthermore, we have the approximate correspondence:

$$D \approx \frac{1}{\sqrt{2}} \sqrt{\frac{\tilde{D}}{1 + \tilde{D}}} \quad \text{and} \quad \tilde{D} \approx \frac{2D^2}{1 - 2D^2}. \quad (38)$$

With the adequate correspondence of the noise factor, the parameter  $\beta$  only acts as a timescale, and since the time correspondence between the model and the real data has to be calibrated for, its influence does not impact the general result (unless  $\beta = 1$  in which case the time scale factor become infinite).

## V. NUMERICAL RESULTS

In this section, we study the behavior of the generalized SIRM model defined by Eqs. (6) and (30). We study our model both on a synthetic network with randomly generated initial conditions and on a real example. The data of our example corresponds to commuting and election data of the county of

Stockholm in Sweden. The commuting data are obtained from Refs. [44–49] and the initial condition corresponds to the year 1985.

Hereafter, we set  $\alpha$  and  $\beta$  to  $\frac{1}{2}$  for simplicity. In Ref. [18], the actual value of  $\alpha$  has been shown not to have a strong influence on the dynamics and  $\beta$  mainly acts as a timescale, as discussed in Sec. IV C. Thus, it is sufficient to study the influence of the two other parameters added to the model:  $\tilde{D}$  and  $\gamma$ . To run the model, we need to specify initial conditions, that is, we need to specify the commuting network and the initial vote share distribution. When applying the model to real election data, the initial conditions are inferred from the data. The focus of our numerical experiments is to study the influence of the two types of noise (free will and noisy intra-commuting-cell influence) added to the recurrent mobility component.

The commuting network we choose for numerical experiments is based on a fully connected directed graph with self-loops. The nodes represent living and working places and we associate with every link of that network a number of commuters  $N_{ij}$ , randomly chosen between 0 and 100. We then multiplied the  $N_{ij}$  quantities by a factor 100 to account for the fact that at the granularity level of the county or municipality the majority of people live and work in the same place. The simulated commuting network has 25 nodes and a total population of  $N = 152\,583$  [50].

The initial voting distribution is constructed by partitioning the commuting populations  $N_{ij}$  into the number  $K$  of possible options (chosen to be  $K = 10$  for this paper). The choice of 10 parties is not arbitrary but motivated by the fact that in many countries, the number of parties with representation in the government is roughly 10. For example, Iceland has 8 elected parties and Sweden has 9. Even in the U.S., where two parties dominate, if we look at the parties represented at

the state level, the number of parties is also 8. Furthermore, this number of parties does not seem to depend on the size of the population, as demonstrated by the chosen examples. For instance, populations range from the order of  $10^5$  for Iceland to the order of  $10^8$  for the U.S., covering 3 orders of magnitude. The partitioning is done by sequentially partitioning the  $N_{ij}$  into the different options  $V_{ij}^k$ :  $V_{ij}^1$  is uniformly sampled between 0 and  $N_{ij}$ , then the number of partisans of the second party  $V_{ij}^2$  is uniformly sampled from the remaining population  $N_{ij} - V_{ij}^1$ . This procedure is repeated until all parties but one have been assigned and the remaining population is assigned to the last party. This procedure generates a party distribution characterized by 1 or 2 large parties and many small parties, similar to actual party distributions [51].

In the case of Stockholm data, the commuting network is obtained from census data [44–49] and the initial voting distribution corresponds to the year 1985. The county of Stockholm contains 25 municipalities and a total voting population  $N = 1\,139\,658$ , which is one order of magnitude larger than the population synthetically generated.

Since there are multiple parties, we need to use an adapted visualisation technique. We choose to plot the standard deviation  $\sigma$  of a party vote share as a function of its mean  $\mu$ . This type of representation provides a good picture of the distribution of parties. In many previous studies [18,22,23], it has been observed that the distribution of party size is roughly stationary over multiple elections. For our data for Stockholm, this is also the case. The stationarity of the distribution will be used to calibrate our model.

In the rest of this section, we discuss the influence of  $\tilde{D}$  when  $\gamma = 0$ , which is the closest case to the SIRM model. This leads to a calibration procedure for  $\tilde{D}$ . We then investigate the influence of free will on the dynamics of the system [52].

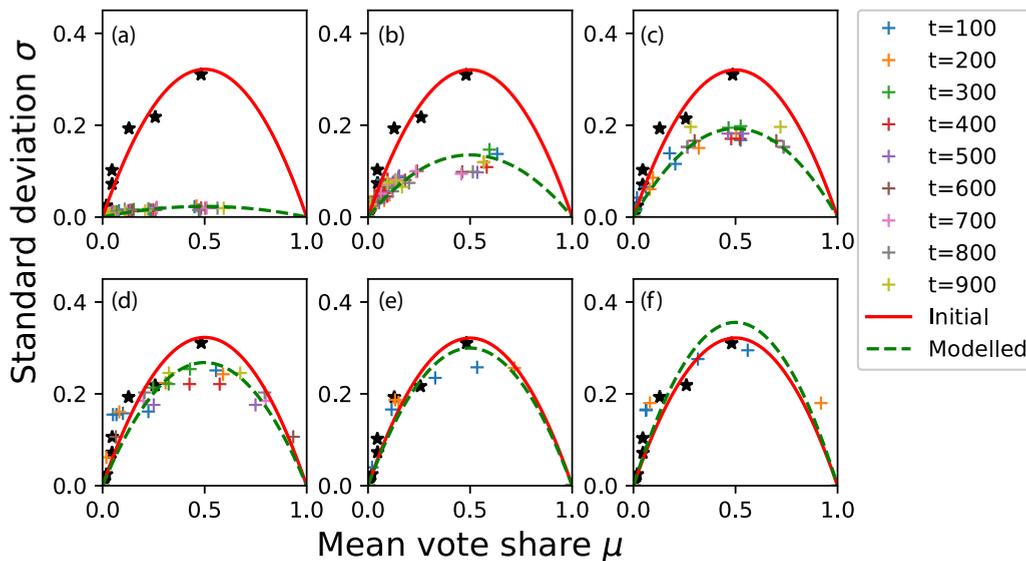


FIG. 1. Dependence of the results for the synthetic case as a function of  $\tilde{D}$ . We display the results for  $\tilde{D} = 0, 0.1, 0.3, 1.0, 5.0$ , and  $10.0$  in panels (a), (b), (c), (d), (e), and (f), respectively. Black stars represent the initial party distribution and the red continuous line represents the function  $f(x) = cx(1-x)$  fitted to these values (the fitted value for the initial condition is  $c^*$ ). The pluses represent the state of the model at different times and the dashed green line represents the fit of the function  $f(x)$  to these values. The graphs have been generated for a single realization of the model.

**A. Influence of  $\tilde{D}$  ( $\gamma = 0$ )**

Let us start by simulating the synthetic system for various values of  $\tilde{D}$ . Results are displayed in Fig. 1. The initial party distribution is displayed by stars and party distribution at later times by pluses. We observe that for small values of  $\tilde{D}$ , the standard deviation is quickly reduced, indicating a convergence over the population to the given averaged values of the party vote shares. When the value of  $\tilde{D}$  is increased, the standard deviation increases as well, indicating more heterogeneity in the system. This is expected, since  $\tilde{D}$  controls the importance of the noise and more noise leads to larger standard deviations.

At some critical value  $\tilde{D}^*$  of the diffusion coefficient, the evolution predicts a fairly stationary dynamics and the standard deviation stays at the same level as that of the initial conditions. The stationarity criterion can be used to calibrate the model. For instance, we can fit the initial data to a specific function and repeat the fit at later times. If we use a function with a single parameter, we can run the model for different values of  $\tilde{D}$  and choose the specific value  $\tilde{D}^*$  when the coefficients of the fit match. We assume that the functional form of the standard deviation dependence on the mean follows a function  $f(x) = cx(1 - x)$ . In this case, the only parameter of this function is  $c$ . We denote by  $c^*$  the fitted value for the initial condition. In Fig. 2, we report the calibration curve for  $\tilde{D}$ , which leads to a critical value of  $\tilde{D}^* \approx 1.41$ .

The calibration procedure was also performed on Stockholm data and leads to  $\tilde{D}_S^* \approx 0.08$ ; see Fig. 4. The value for Stockholm data is much smaller than for the randomly generated network, illustrating the fact that there is some ordered structure in real data compared to randomly generated data.

The critical value  $\tilde{D}^*$  can be related to the value of  $D$  used in the original SIRM model [18] using Eq. (38). In our synthetic example, we obtain a critical value of  $D^* \approx 0.54$ , which is one order of magnitude larger than the value of  $D_{SIRM}^* = 0.03$  obtained in Ref. [18], more precisely, 18 times larger. This indicates that the random assignment of vote shares used in our numerical experiments leads to a larger noise value than actual vote share distribution. What this result means is that

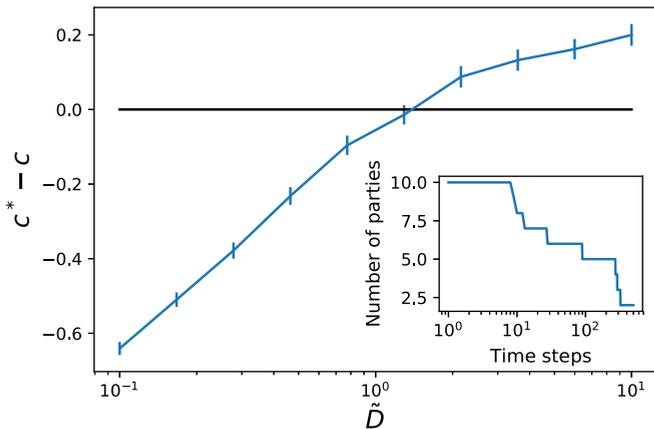


FIG. 2. Calibration curve for  $\tilde{D}^*$  for  $\gamma = 0$  in the synthetic case. The curve has been generated by averaging 50 runs of the model. Each run is computed for 100 time steps. Here,  $c^*$  is the fitted value of the function  $f(x) = cx(1 - x)$  on the initial value and  $c$  is the corresponding value for modeled data. Inset: Single realization of the decrease of the number of parties computed for the critical value  $\tilde{D}^*$ .

to conserve the stationarity of random initial data, the noise level should be significantly higher than to conserve real initial conditions.

For the Stockholm example, Eq. (38) leads to a value of  $D_S^* \approx 0.19$ , which is more than twice smaller than the synthetic case value. This value is also larger than the value obtained for the U.S. election case ( $D_{SIRM}^* = 0.03$  [18]) by a factor of roughly 6.

One of the drawbacks of using a model without a free will term ( $\gamma = 0$ ) is that the number of parties decreases during the evolution of the system, see the inset of Fig. 2, where the decrease of the number of parties is displayed ( $\tilde{D} = \tilde{D}^*$ ). As can be seen, many parties go extinct fairly quickly and the actual number of remaining parties turns out to be quite small (2 in this case). In real election systems, new parties can be created. Without a free will term, our model is unable to reproduce the emergence of new parties.

**B. Influence of free will**

For the study of the influence of free will, we consider the following settings. We set  $\alpha = 0.5$ ,  $\beta = 0.5 - \gamma$  and modify  $\tilde{D}$  for different values of  $\gamma$ . The presence of free will allows extinct parties to re-enter the dynamics, simulating the emergence of a new party. Results of calibration curves are displayed in Fig. 3 for synthetic data and in Fig. 4 for the Stockholm case. We observe that the critical value of  $\tilde{D}^*(\gamma)$  increases as a function of  $\gamma$ , see the inset in Fig. 3, and when  $\gamma$  exceeds a threshold, the calibration procedure fails, i.e., no intersection between the curves (not shown). This is due to free will dominating the dynamics and leading to a well-mixed state.

To fully calibrate the model, the intensity of free will should be specified. To fix the free-will intensity, we can use the relationship between  $\tilde{D}^*(\gamma)$  and  $\gamma$  and run the model for various values of  $\gamma$ . The calibration value  $\gamma^*$  is the value for which the expected size of the largest party is stable over time. In Fig. 5, we display the calibration curve for  $\gamma$ . To obtain this curve, we run the model 50 times for 100 time steps for each values of  $\gamma$  and averaged the largest party sizes. The error bars display the standard deviation of the largest

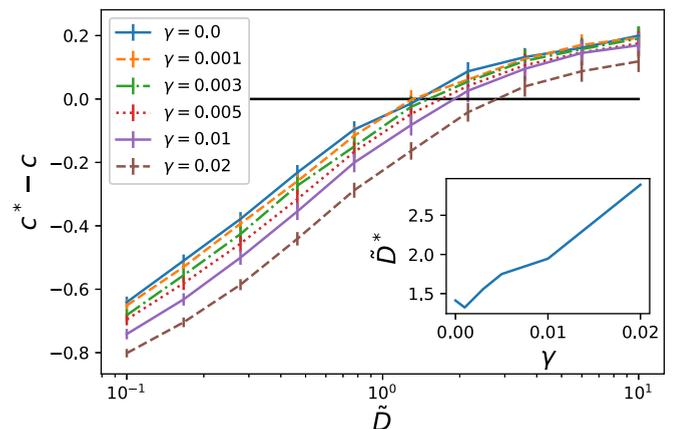


FIG. 3. Effect of  $\gamma$  on the calibration curves of  $\tilde{D}^*$  in the synthetic case. Here,  $c^*$  is the fitted value of the function  $f(x) = cx(1 - x)$  on the initial value and  $c$  is the corresponding value for modeled data. Inset: Dependence of  $\tilde{D}^*$  on  $\gamma$ .

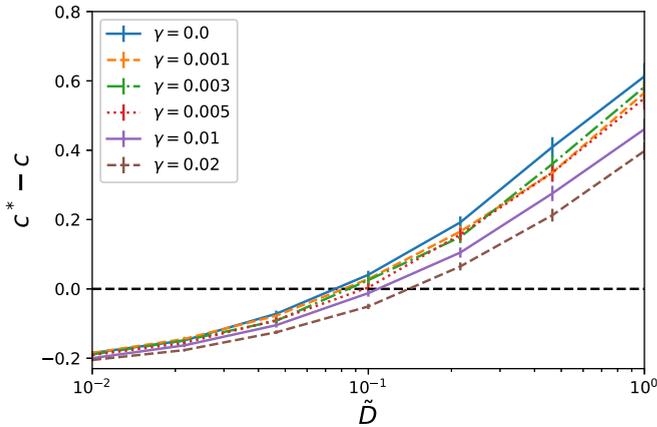


FIG. 4. Effect of  $\gamma$  on the calibration curves of  $\tilde{D}_S^*$  for the Stockholm case. Here,  $c^*$  is the fitted value of the function  $f(x) = cx(1 - x)$  on the initial value and  $c$  is the corresponding value for modeled data.

party sizes. We observe that the maximum party size is a decreasing function of  $\gamma$ , which justifies its use as a calibration criterion, but the standard deviation for low values of  $\gamma$  is rather large and leads to a big uncertainty on the value of  $\gamma^*$ . The calibrated value of the synthetic case is  $\gamma^* \approx 2.3 \times 10^{-3}$ , see Fig. 5, whereas the corresponding value in the Stockholm case is  $\gamma_S^* \approx 4.3 \times 10^{-4}$  (figure not shown). Given the large uncertainties, the calibration procedure suggests that the free will term might not be needed to model the Stockholm case.

The large uncertainty on the critical value for  $\gamma^*$  can be reduced by averaging over more realisations of the model (we used 50 runs) or by using longer simulations (we used 100 time steps). Doing both would be optimal. In this paper, the aim is to show that our generalization of the SIRM model can be used to study both synthetic and real situations. Since the study of the predictions of our model is outside the scope of this paper, there is no need to increase the precision of the calibration procedure.

In this section, we have provided a calibration procedure for  $\tilde{D}$  and  $\gamma$ , we did not explore in details the effect of  $\alpha$  and

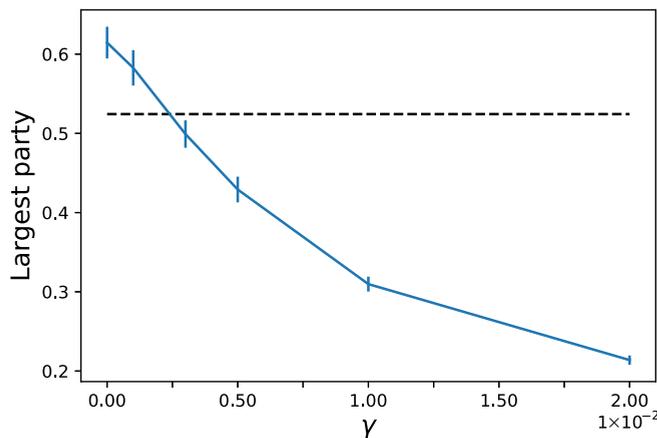


FIG. 5. Calibration curve of  $\gamma^*$  based on the stationarity of the largest party size for the synthetic case. For each values of  $\gamma$ , the model was run 50 times for 100 time steps and the reported results are the difference between the initial largest party size and the averaged largest party after 100 time steps.

$\beta$ . This is left for another study. According to Ref. [18], the effect of  $\alpha$  on the dynamics is negligible and  $\beta$  mainly acts as a timescale; see Sec. IV C. As was shown, the calibration procedure can be applied both to synthetic and real data.

### VI. CONCLUSION

In this paper, we have provided a generalized formulation for the voter model with recurrent mobility and illustrated the issues with its definition in Ref. [18]. We then developed a general procedure to add two types of noise ( $\tilde{D}$  and  $\gamma$ ) in this model and provided a generalization to the SIRM model. We stress the importance of using two different types of noise in the system: the free will term ( $\gamma$ ), modeling the emergence of new parties, and noisy intra-commuting-cell influence ( $\tilde{D}$ ) that models local variation and fluctuations of the probability to change opinion.

We have illustrated how the SIRM model could be recovered as a limit of our generalized model and exemplified a possible calibration procedure for  $\tilde{D}$  and  $\gamma$  through numerical experiments. Note that the diffusion constant ( $\tilde{D}^*$ ) calculated when  $\gamma = 0$  for the synthetic case is about 18 times larger than the corresponding value ( $D_{SIRM}^*$ ) obtained for the SIRM model [18], whereas the corresponding value for the Stockholm case ( $D_S^*$ ) is about six times larger. This indicates that a calibration procedure based on stationarity requires more noise to conserve random initial conditions than to conserve real initial conditions and that the Stockholm case is more noisy than the U.S. case. We have shown that our model can be applied to the Swedish electoral system, but other electoral systems are outside the scope of this model; see Ref. [53]. In this paper, we focused on the technical aspect of the model and did not fully analyze the prediction of our model, but only demonstrated how it can be used in real situations. We are now one step closer to answering the question “Is the voter model a model for voters?”—and to potential predictive practical applications—for a multiparty situation and where vote shares can be extreme (close to 0 or 1).

Free will is crucial to get a dynamics that can be related to real opinion dynamics, since new opinions do arise. However, the assumption that transition rates between different opinions are uniform is likely to be too strong; see Ref. [54]. One natural extension of this work would be to investigate the influence of structure in the opinion space by introducing constraints on the interactions, as done in Ref. [13], or by adapting the bounded confidence model of Hegselmann and Krause [55] to the voter model. One can also imagine applying this type of model in situations where opinions are represented in higher dimensional opinion spaces. The Axelrod model of cultural evolution [56] is an example of how this could be done.

This paper hints to many other possible developments of this work, since the rates can be modified in many ways. One could, for example, use the Dirichlet distribution to add noise on other components of the rate and include time dependent or spatially dependent rates to account for varying and heterogeneous socioeconomic factors that might have an influence on the dynamics of the system. Furthermore, the influence of the commuting network can be studied through numerical experiments. The synthetic network used in this

paper is just one among many possibilities and it would be interesting to understand the influence of the topology of this weighted directed graph on the dynamics of the system.

### ACKNOWLEDGMENTS

Earlier versions of this work have been presented at seminars at the University of Uppsala and at the University of

Örebro. We thank the audience of these seminars for interesting comments and discussions. In particular, we thank Prof. Ilkka Henrik Mäkinen and Prof. Olle Eriksson for support and advice. We also thank Dr. Juan Fernández-Gracia for advice and comments. Finally, we thank Emil Frisk, who helped in gathering the Stockholm data. This work was supported by the Swedish Strategic Research Programme eSENCE.

- 
- [1] C. Castellano, S. Fortunato, and V. Loreto, *Rev. Mod. Phys.* **81**, 591 (2009).
- [2] P. Clifford and A. Sudbury, *Biometrika* **60**, 581 (1973).
- [3] R. A. Holley and T. M. Liggett, *Ann. Probab.* **3**, 643 (1975).
- [4] L. Frachebourg and P. L. Krapivsky, *Phys. Rev. E* **53**, R3009 (1996).
- [5] P. L. Krapivsky, *Phys. Rev. A* **45**, 1067 (1992).
- [6] F. Wu, B. A. Huberman, L. A. Adamic, and J. R. Tyler, *Physica A: Stat. Mech. Appl.* **337**, 327 (2004).
- [7] K. Suchecki, V. M. Eguíluz, and M. S. Miguel, *Phys. Rev. E* **72**, 036132 (2005).
- [8] V. Sood and S. Redner, *Phys. Rev. Lett.* **94**, 178701 (2005).
- [9] V. Sood, T. Antal, and S. Redner, *Phys. Rev. E* **77**, 041121 (2008).
- [10] C. Sire and S. N. Majumdar, *Phys. Rev. E* **52**, 244 (1995).
- [11] M. Mobilia, *Phys. Rev. Lett.* **91**, 028701 (2003).
- [12] M. Mobilia, A. Petersen, and S. Redner, *J. Stat. Mech.: Theory Exp.* **2007**, P08029 (2007).
- [13] F. Vazquez, P. L. Krapivsky, and S. Redner, *J. Phys. A: Math. Gen.* **36**, L61 (2003).
- [14] R. Lambiotte and S. Redner, *J. Stat. Mech.: Theory Exp.* **2007**, L10001 (2007).
- [15] L. Dall'Asta and C. Castellano, *Europhys. Lett.* **77**, 60005 (2007).
- [16] A. Jędrzejewski, *Phys. Rev. E* **95**, 012307 (2017).
- [17] N. Khalil, M. San Miguel, and R. Toral, *Phys. Rev. E* **97**, 012310 (2018).
- [18] J. Fernández-Gracia, K. Suchecki, J. J. Ramasco, M. San Miguel, and V. M. Eguíluz, *Phys. Rev. Lett.* **112**, 158701 (2014).
- [19] S. Fortunato and C. Castellano, *Phys. Rev. Lett.* **99**, 138701 (2007).
- [20] A. Chatterjee, M. Mitrović, and S. Fortunato, *Sci. Rep.* **3**, 1049 (2013).
- [21] P. Klimek, Y. Yegorov, R. Hanel, and S. Thurner, *Proc. Natl. Acad. Sci. USA* **109**, 16469 (2012).
- [22] C. Borghesi and J.-P. Bouchaud, *Eur. Phys. J. B* **75**, 395 (2010).
- [23] C. Borghesi, J.-C. Raynal, and J.-P. Bouchaud, *PloS One* **7**, e36289 (2012).
- [24] R. Enikolopov, V. Korovkin, M. Petrova, K. Sonin, and A. Zakharov, *Proc. Natl. Acad. Sci. USA* **110**, 448 (2013).
- [25] R. N. Costa Filho, M. P. Almeida, J. S. Andrade, and J. E. Moreira, *Phys. Rev. E* **60**, 1067 (1999).
- [26] L. Araripe and R. C. Filho, *Physica A: Stat. Mech. Appl.* **388**, 4167 (2009).
- [27] A. Bernardes, D. Stauffer, and J. Kertész, *Eur. Phys. J. B* **25**, 123 (2002).
- [28] M. Lyra, U. Costa, R. Costa Filho, and Andrade, Jr., *Europhys. Lett.* **62**, 131 (2003).
- [29] G. Travieso and L. da Fontoura Costa, *Phys. Rev. E* **74**, 036112 (2006).
- [30] L. E. Araripe, R. N. Costa Filho, H. J. Herrmann, and J. S. Andrade, Jr., *Int. J. Mod. Phys. C* **17**, 1809 (2006).
- [31] M. Gonzalez, A. Sousa, and H. Herrmann, *Int. J. Mod. Phys. C* **15**, 45 (2004).
- [32] C. A. Andresen, H. F. Hansen, A. Hansen, G. L. Vasconcelos, and J. S. Andrade, Jr., *Int. J. Mod. Phys. C* **19**, 1647 (2008).
- [33] H. Hernández-Saldaña, *Physica A: Stat. Mech. Appl.* **388**, 2699 (2009).
- [34] N. A. Araújo, J. S. Andrade, Jr., and H. J. Herrmann, *PLoS One* **5**, e12446 (2010).
- [35] J. Kim, E. Elliott, and D.-M. Wang, *Electoral Studies* **22**, 741 (2003).
- [36] Here, we introduce the vector notation  $\mathbf{p}_{ij}^{k \rightarrow}$  for consistency of the presentation and to underline the fact that it represents a probability distribution.
- [37] C. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences*, 4th ed. (Springer, Berlin, 2009).
- [38] M. Scheucher and H. Spohn, *J. Stat. Phys.* **53**, 279 (1988).
- [39] B. L. Granovsky and N. Madras, *Stoch. Process. Appl.* **55**, 23 (1995).
- [40] S. L. Feld, *Amer. J. Sociol.* **86**, 1015 (1981).
- [41] S. Feld and B. Grofman, in *The Oxford Handbook of Analytical Sociology*, edited by P. Hedström and P. Bearman (Oxford University Press, Oxford, 2009), pp. 521–543.
- [42] J. Michaud, *Phys. Rev. E* **95**, 022308 (2017).
- [43] Note that our definition of  $R_{ij}^{kk'}$  is a probability and, therefore, a normalized rate of change. This implies that we can eliminate one parameter by rescaling the time. For example, in our model Eq. (30) we can scale the time by a factor  $(1 - \beta - \gamma)$  [best seen from Eq. (35)], this allows to eliminate  $\beta$ . The new rate corresponding to Eq. (30) reads  $p_{ij}^{k \rightarrow k'} = r_{ij}^{k \rightarrow k'} + \tilde{v}_{ij}^{k'}$ , where  $\hat{\gamma} = \frac{\gamma}{1 - \beta - \gamma}$  and  $\tilde{D} = (1 - \hat{\beta}^2 \frac{\tilde{D}}{\tilde{D} + 1})^{-1} - 1$ .
- [44] Statistics Sweden, Swedish National Data Service. Version 1.0 (1989).
- [45] Statistics Sweden, Swedish National Data Service. Version 1.0 (1989).
- [46] Statistics Sweden, Swedish National Data Service. Version 1.0 (1992).
- [47] Statistics Sweden, Swedish National Data Service. Version 1.0 (1996).
- [48] Statistics Sweden, Swedish National Data Service. Version 1.0 (1999).
- [49] Statistics Sweden, Swedish National Data Service (2017).

- [50] This number is specific to the particular run of the model, different runs lead different values of  $N$ .
- [51] S. Hart and M. Kurz, *Econometrica: J. Econ. Soc.* **51**, 1047 (1983).
- [52] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevE.97.062313> for the implementation of our model in Python and details about the computations.
- [53] In real election systems, like proportional elections, the voting dynamics is more complex and further work should be done to adapt our model to that case; see, for example, Ref. [19].
- [54] Opinions and parties are not fully independent of one another and opinion structures should be introduced at some point. In any VM-like model the opinion structure is ignored and opinion change only depends on demographics.
- [55] R. Hegselmann and U. Krause, *J. Artific. Soc. Soc. Simul.* **5** (2002).
- [56] R. Axelrod, *J. Conflict Resolut.* **41**, 203 (1997).