On a Voter Model with Context-Dependent Opinion Adoption

Luca Becchetti
Sapienza University of Rome
Rome, Italy
becchetti@diag.uniroma1.it

Vincenzo Bonifaci
Roma Tre University
Rome, Italy
vincenzo.bonifaci@uniroma3.it

Emilio Cruciani
Paris-Lodron University of Salzburg
Salzburg, Austria
emilio.cruciani@plus.ac.at

Francesco Pasquale
University of Rome "Tor Vergata"
Rome, Italy
pasquale@mat.uniroma2.it

May 12, 2023

Abstract

Opinion diffusion is a crucial phenomenon in social networks, often underlying the way in which a collective of agents develops a consensus on relevant decisions. The voter model is a well-known theoretical model to study opinion spreading in social networks and structured populations. Its simplest version assumes that an updating agent will adopt the opinion of a neighboring agent chosen at random. The model allows us to study, for example, the probability that a certain opinion will fixate into a consensus opinion, as well as the expected time it takes for a consensus opinion to emerge.

Standard voter models are oblivious to the opinions held by the agents involved in the opinion adoption process. We propose and study a context-dependent opinion spreading process on an arbitrary social graph, in which the probability that an agent abandons opinion a in favor of opinion b depends on both a and b. We discuss the relations of the model with existing voter models and then derive theoretical results for both the fixation probability and the expected consensus time for two opinions, for both the synchronous and the asynchronous update models.

1 Introduction

The *voter model* is a well-studied stochastic process defined on a graph to model the spread of opinions (or genetic mutations, beliefs, practices, etc.) in a population [23, 15]. In a voter model, each node maintains a state, and when a node requires updating it will import its state from a randomly chosen neighbor. Updates can be asynchronous, with one node activating per step [23], or synchronous, with all nodes activating in parallel [15]. While the voter model on a graph has been introduced in the 1970s to model opinion dynamics, the case of a complete graph is also very well-known in population genetics where, in fact, it was introduced even earlier, to study the spread of mutations in a population [13, 26].

Mathematically, among the main quantities of interest in the study of voter models, there are the *fixation probability* of an opinion—the probability of reaching a configuration in which each node adopts such opinion—and the expected *consensus* (or *absorption*) *time*—the expected number of steps before all nodes agree on an opinion. Such quantities could in principle be computed for any n-node graph by defining a Markov chain on a set of C^n configurations, where C is the number of opinions, but such an approach is computationally infeasible even for moderate values of n. Therefore, a theoretical analysis of a voter process will often focus on obtaining upper and lower bounds for these quantities, still drawing heavily on the theory of Markov chains [1, 21], but with somewhat different approaches and tools for the synchronous and asynchronous cases.

A limitation of the standard voter process is that the dynamics is oblivious to the states of both the agent u that is updating and of the neighbor that u copies its state from, and the copying always occurs. One could easily imagine a situation (for example, in politics) where an agent holding opinion a is more willing to adopt the opinion b of a neighbor rather than to adopt opinion c; in general, the probability of abandoning opinion a in favor of opinion b might depend on both a and b. This motivates the study of *biased* voter models [6, 31] and in particular motivates us to introduce a voter model with an opinion adoption probability that depends on the context, that is, on the opinions of *both* agents involved in an opinion spreading step.

We define and study extensions of the voter model that allow the opinion adoption probability to depend on the pair of opinions involved in an update step. We consider both an asynchronous variant and a synchronous variant of a context-dependent voter model with two opinions, 0 and 1. We assume that an agent holding opinion $c \in \{0, 1\}$ is willing to copy the opinion of an agent holding opinion $c' \in \{0, 1\}$ with some probability $\alpha_{c,c'}$, which models the bias in the update. We study both the fixation probabilities and the expected consensus time.

1.1 Our Findings

In general, a seemingly minor feature as the form of bias we consider has a profound impact on the analytical tractability of the resulting model. While the unbiased case[1] can still be connected to a variant of the voter model and analyzed accordingly with some extra work, the same is not possible for the biased case. Specifically, in Section 3 we prove that a lazy variant of the voter model is equivalent (i.e., it produces the same distribution over possible system's configurations) to the unbiased variant of the model we consider. The proof, given in Lemma D.1 for the synchronous case[2], uses a coupling between the Markov chains that describe the two models. For the asynchronous case, this allows us to directly leverage known connections between the asynchronous voter model and random walks (Proposition 1). For the case of the clique, this general result can be improved, providing explicit, tight bounds on expected consensus time (Theorem 1). In the synchronous case, the above connection is not immediate (i.e., Proposition 1 does not apply) and analyzing expected time to consensus requires adapting arguments that have been used for continuos Markov chains to the synchronous, discrete setting (Theorem 2).

The biased case is considerably harder to analyze, the main reason being that it is no longer possible to collapse the Markov chain describing the system (whose state space is in general the exponentially large set of all possible configurations) to a "simpler" chain, e.g., a random walk on the underlying network, not even in the case of the clique. Despite these challenges, some trends emerge from specific cases. Interestingly, it is possible to derive the exact fixation probabilities for the class of regular networks, highlighting a non-linear dependence from the bias (Theorem 3), while an asymptotically tight analysis for the clique (Theorem 4) suggests that the presence of a bias may have a positive impact on achieving faster consensus in dense networks. Though seemingly intuitive, this last aspect is not a shared property of biased opinion models in general [24, 2]. The behavior of the model is considerably more complex and technically challenging in the synchronous, biased case. In particular, the preliminary results we obtain highlight a

[1]The model is unbiased when $\alpha_{0,1} = \alpha_{1,0}$.

[2]The proof for the asynchronous case is essentially the same up to technical details and is omitted for the sake of space.

general dependence of fixation probabilities (Proposition 2) and, notably, expected consensus time (Theorem 5) from both the bias and the initial configuration.

1.2 Related work

For the sake of space, we mostly discuss results that are most closely related to the setting we consider.

Voter and voter-like models. Due to its versatility, the voter model has been defined multiple times across different disciplines and has a vast literature. As mentioned in the introduction, the special case of the voter model on a complete graph was first introduced in mathematical genetics, being closely related to the so-called Wright-Fisher and Moran processes [25, 18, 13]. The first asynchronous formulation of a voter model on a connected graph has been proposed in the probability and statistics community in the 1970s [23, 11], while Hassin and Peleg [15] were the first to study this model in a synchronous setting. A classic result of these papers is that the fixation probability of an opinion c is equal to the weighted fraction of nodes holding opinion c, where the weight of a node is given by its degree [15, 31].

The expected consensus time of the voter model is much more challenging to derive exactly, even for highly structured graphs. In the asynchronous case, it has often been studied by approximating the process with a continuous diffusion partial differential equation [13, 31, 5]. For two opinions on the complete graph, this yields the approximation

$$T(n) \approx n^2 h(k/n) \tag{1}$$

where $T(n)$ is the expected consensus time on the n-clique, k is the number of nodes initially holding the first opinion and $h(p) = -p \ln p - (1-p)\ln(1-p)$ [31, 13]. To the best of our knowledge, however, no error bound was known for such diffusion approximations[3]. Another approach is to use the duality between voter model and coalescing random walks [23, 11], which involves no approximations, but the resulting formulas are hard to interpret, and to paraphrase Donnelly and Welsh [11], "an exact evaluation of the expected absorption time for a general regular graph is a horrendous computation". As for approximations, expected consensus time for the voter model can be bounded by $\mathcal{O}((d_{\text{avg}}/d_{\text{min}})(n/\Phi))$, where d_{avg} and d_{min} are, respectively, the average and minimum degrees [6]. Most relevant to our discussion is the biased voter model considered by Berenbrink et al. [6], in which the probability of adoption of an alternative opinion c' depends on c' (and only on c'). While our model is different if we consider more than 2 opinions, there are several other differences with respect to [6] even in the binary case. In particular, Berenbrink et al. only consider the synchronous setting, they assume there is a "preferred" opinion that is never rejected and that there is a constant gap between the adoption probabilities of the preferred opinion and of the non-preferred one. Finally, they only consider the case where the number k of nodes initially holding the preferred opinion is at least $\Omega(\log n)$. Thus, for example, their results do not apply in the neutral case ($\alpha_{01} = \alpha_{10}$) or when k is, say, $\sqrt{\log n}$. Our results for the biased, synchronous case (Theorem 5) are complementary to those in [6]. While their results are stronger when the above assumptions hold, ours address the general and challenging case of an arbitrary initial configuration.

Pull vs push. We only reviewed here models where nodes "pull" the opinion from their neighbor, since both the standard voter model and our generalization follow this rule, but we remark that "push" models, also known as invasion processes, have also been defined and studied on connected graphs [22, 10]. The asynchronous push model is sometimes called the (generalized) Moran process [10, 26]. We remark that while the pull and push models are interchangeable on regular graphs, on irregular graphs their behavior can be markedly different.

[3]For the n-clique, we show that (1) is correct within an additive $\mathcal{O}(n)$ term. See Theorem 1 (Section 3.1).

```
Sample v ∈ N(u)
c ← x_u, c' ← x_v
Sample θ ∈ [0,1]
if θ < α_{c,c'} then
    x_u ← x_v
    return accept
return reject
```

Algorithm 1: `Update`(u)

Other biased opinion dynamics. We are aware of only a few analytically rigorous studies of biased opinion dynamics, including biased variants of the voter model [31, 6, 2, 9, 12], sometimes framed within an evolutionary game setting [24]. In general, these contributions address different models, be it because of the way in which bias in incorporated within the voting rule, the opinion dynamics itself or the temporal evolution of the process (e.g., synchronous vs asynchronous). We remark that all these aspects can deeply affect the overall behavior of the resulting dynamics. Specifically, as observed in a number of more or less recent contributions [2, 9, 7, 16], even minor changes in the model that would intuitively produce consistent results with a given baseline can actually induce fundamental differences in the overall behavior, so that it is in general hard to predict if and when results for one model more or less straightforwardly carry over to another model, even qualitatively. Less related to the spirit of this work, a large body of research addresses biased opinion dynamics using different approaches, based on approximations and/or numerical simulations. Examples include numerical simulations for large and more complex scenarios [16], mean-field or higher-order [28] and/or continuous approximations [4]. While these approaches can afford investigation of richer and more complex evolutionary game settings (e.g., [28]), they typically require strong simplifying assumptions to ensure tractability, so that it is harder (if not impossible) to derive rigorous results.

2 Model formulation

Notation. For a natural number k, let $[k] := \{0, 1, 2, \ldots, k-1\}$. If $G = (V, E)$ is a graph, we write $N_G(u)$ (or simply $N(u)$ if G can be inferred from the context) for the set of neighbors of node u in G. We write d_u for the degree of node u.

Model. We define an opinion dynamics model on networks. The parameters of the model are: i) an underlying *topology*, given by a graph G on n nodes, with symmetric adjacency matrix $A = (a_{uv})_{u,v\in[n]}$; ii) a number of *opinions* (or *colors*) $C \geqslant 2$; iii) an *opinion acceptance matrix* $(\alpha_{c,c'})_{c,c'\in[C]}$. The initial opinion of each agent (node) u is encoded by some $x_u^{(0)} \in [C]$.

For any node $u \in [n]$, we define an update process `Update`(u) consisting of the following steps (summarized in Algorithm 1):

1. **Sample**: Sample a neighbor v of u uniformly at random, i.e., according to the distribution $(a_{u1}/d_u, \ldots, a_{un}/d_u)$ where $a_{uv} = 1$ if u and v are adjacent, $a_{uv} = 0$ otherwise. Here $d_u = |N(u)| = \sum_{v\in[n]} a_{uv}$ is the degree of node u.
2. **Compare**: Compare u's opinion $c = x_u$ with v's opinion $c' = x_v$.
3. **Accept/reject**: With probability $\alpha_{c,c'}$, set $x_u \leftarrow x_v$; in this case we say u *accepts* v's opinion. Otherwise, we say u *rejects* v's opinion.

We consider two variants of the model, differing in how the updates are scheduled. In one iteration of the *asynchronous* variant, $u \in [n]$ is sampled at random and `Update`(u) is applied. In one iteration of the *synchronous* variant, each node $u \in [n]$ applies `Update`(u) in parallel. We denote by $x_u^{(t)}$ the random variable encoding the opinion of node u after t iterations of either the synchronous or the asynchronous dynamics (depending on the context).

The acceptance probabilities $\alpha_{c,c'}$ are parameters of the model. We note that the parameters $\alpha_{c,c'}$ with $c = c'$ are irrelevant for the dynamics, since a node sampling a neighbor of identical opinion will not change opinion, irrespective of whether it accepts the neighbor's opinion or not. Hence, to specify the opinion acceptance matrix $C(C-1)$ parameters are sufficient; we can assume that the diagonal entries equal, say, 1. In particular, when $C = 2$ it is enough to specify α_{01} and α_{10}. When $\alpha_{01} = \alpha_{10} = 1$, the model boils down to the standard voter model [15, 23].

In the rest of this work we assume $C = 2$. In this case, we say that the model is *unbiased* if the opinion acceptance matrix is symmetric, i.e., $\alpha_{01} = \alpha_{10}$, and *biased* otherwise.

Quantities of interest. The *fixation probability* of opinion 1 is the probability that there exists an iteration t such that $x_u^{(t)} = 1$ for all $u \in [n]$. The *consensus time* is the index of the first iteration t such that $x_u^{(t)} = x_v^{(t)}$ for all $u, v \in [n]$.

3 The unbiased setting

Before embarking on the biased case, which is substantially more complex, in this section we review or prove directly results for the unbiased setting ($\alpha_{01} = \alpha_{10}$). We consider the asynchronous and the synchronous variants separately. In all formulas of this section, $\alpha = \alpha_{01} = \alpha_{10}$.

3.1 Asynchronous variant

The main, intuitive observation about the unbiased asynchronous variant of our model is that the model can equivalently be described by a suitable, "lazy" voter model, where each iteration is either an idle iteration (with probability $1-\alpha$) or an iteration of the standard asynchronous voter model (with probability α).

This in turn implies that, for the fixation probability one can simply disregard the idle iterations and therefore obtain the same fixation probability as for the standard asynchronous voter model. In an arbitrary topology, this was derived by Sood et al. [31]: if we call ϕ^{avoter} the fixation probability of the asynchronous voter model, then

$$\phi^{\text{avoter}} = \frac{\sum_{u\in[n]} d_u x_u^{(0)}}{\sum_{u\in[n]} d_u}, \tag{2}$$

where $x_u^{(0)}$ and d_u are respectively the initial opinion and the degree of node u. Since $x_u^{(0)} \in \{0,1\}$, the fixation probability ϕ^{avoter} is proportional to the volume of nodes initially holding opinion 1[4].

In the analysis of the expected consensus time, instead, one cannot ignore the idle iterations, but since they occur with probability $1-\alpha$ independently of other random choices, their effect is simply that of slowing down the standard asynchronous voter process by a factor $1/\alpha$. This intuitive argument can be formalized through a standard Markov chain coupling argument (see Appendix B).

Proposition 1. *In the unbiased asynchronous case, the fixation probability is the same as for the standard asynchronous voter model. The expected consensus time is T^{avoter}/α, where T^{avoter} is the expected consensus time of the standard asynchronous voter model.*

[4] We note incidentally that the fixation probability can also be computed by suitably relating the asynchronous model to the transition matrix of the lazy random walk we discuss in Section 3.2. This connection is only mentioned here and made rigorous in Appendix A of the Supplementary Material for the sake of space.

On the n-clique topology. As discussed in the introduction, a very natural question is: how large is the expected consensus time on the n-clique as a function of n? Despite this question having been studied multiple times before, in the literature there are either diffusion approximations with unknown error [13], or exact formulas involving multiple partial summations that are hard to interpret asymptotically [14]. By further analyzing a result of Glaz [14], we derive here an explicit formula with bounded error that is easy to interpret, which in fact agrees with the diffusion approximation up to lower order terms, thus also showing that at least in this case, the diffusion approximation (1) yields a correct estimate.

In the unbiased asynchronous model, the expected consensus time on the clique is the same as the mean absorption time of the underlying birth-death process, the state of which is summarized by the number of nodes holding opinion 1. Call $T_k(n)$ the expected consensus time when starting from a configuration with k nodes holding opinion 1, and the remaining $n-k$ holding opinion 0. We prove that $T_k(n) = \mathcal{O}(n^2/\alpha)$, more precisely:

Theorem 1. *If $\alpha_{01} = \alpha_{10} = \alpha$ (for some $\alpha > 0$), then for each $k = 1, \dots, n-1$,*

$$T_k(n) = \frac{1}{\alpha} n^2 h(k/n) + \mathcal{O}(n/\alpha),$$

where $h(p) := -p \ln p - (1-p)\ln(1-p) \leqslant \ln 2$.

Proof. On an n-clique, the process is equivalent to a birth-and-death chain [21] on $n+1$ states $0, 1, 2, \dots, n$ (representing the number of nodes with opinion, say, 1). Let us define the following quantities:

- $p_k = \alpha_{01} k(n-k)/n(n-1)$ is the probability that the number of nodes holding opinion 1 increases from k to $k+1$ when $0 \leqslant k < n$,
- $q_k = \alpha_{10} k(n-k)/n(n-1)$ is the probability that the number of nodes holding opinion 1 decreases from k to $k-1$ when $0 < k \leqslant n$.

Note that $p_k = q_k$ for all k due to the assumption $\alpha_{01} = \alpha_{10}$. Define the vector $T(n)$ as $T(n) = (T_1(n), \dots, T_{n-1}(n))^\top$ and consider the matrix

$$B = \begin{pmatrix} p_1 + q_1 & -p_1 & 0 & \dots & 0 \\ -q_2 & p_2 + q_2 & -p_2 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & -q_{n-1} & p_{n-1} + q_{n-1} \end{pmatrix}.$$

The matrix B is constructed so that $BT(n) = \mathbf{1}$, where $\mathbf{1}$ is the all-1 vector. This holds because of the recurrence

$$T_k(n) = 1 + (1 - p_k - q_k) T_k(n) + q_k T_{k-1}(n) + p_k T_{k+1}(n)$$

for the mean consensus times. Therefore, $T(n) = B^{-1}\mathbf{1}$. The matrix B can be explicitly inverted thanks to its tridiagonal structure; an explicit computation (see Appendix C) yields

$$T_k(n) = \frac{n-1}{\alpha} \left((n-k)(H_{n-1} - H_{n-k}) + k(H_{n-1} - H_{k-1}) \right),$$

where H_k is the k-th harmonic number, $H_k = \sum_{j=1}^{k} 1/j$. Recalling the asymptotic expansion $H_n = \ln n + \gamma + \mathcal{O}(1/n)$, where γ is the Euler-Mascheroni constant,

$$\begin{aligned} T_k(n) &= \frac{n-1}{\alpha}((n-k)(H_n - H_{n-k}) + k(H_n - H_k)) + \mathcal{O}\left(\frac{n}{\alpha}\right) \\ &= \frac{n(n-1)}{\alpha}\left(\left(1 - \frac{k}{n}\right)\ln\frac{n}{n-k} + \frac{k}{n}\ln\frac{n}{k}\right) + \mathcal{O}(n/\alpha) \\ &= \frac{n^2}{\alpha} h(k/n) + \mathcal{O}(n/\alpha), \end{aligned}$$

where $h(p) = -p \ln p - (1-p)\ln(1-p)$, which is such that $0 \leqslant h(p) \leqslant \ln 2$ for every $p \in [0,1]$. □

3.2 Synchronous variant

The analysis of the synchronous variant in the unbiased setting relies on the tight connection between the unbiased case of the opinion dynamics we consider and (lazy) random walks on networks.

Connections to lazy random walks. We next provide an equivalent formulation of our model, which reveals an interesting and useful connection to lazy random walks. To this purpose, consider the following, alternative dynamics, in which the behavior of the generic node u at each iteration is the following:

- Node u independently tosses a coin with probability of "heads" equal to α;
- If "heads", u samples a neighbor v u.a.r. and copies v's opinion; otherwise u does nothing and keeps her opinion.

Let us call $\mathcal{M}_1$ the synchronous model described in Section 2 and $\mathcal{M}_2$ the dynamics described above. Then $\mathcal{M}_1$ and $\mathcal{M}_2$ are equivalent in the sense that, if they start from the same initial state, they generate the same probability distribution over all possible configurations of the system at any iteration t. Intuitively speaking this is true since in $\mathcal{M}_1$ each node first samples a neighbor and then it decides whether or not to copy its opinion according to the outcome of a coin toss, while in $\mathcal{M}_2$ each node first tosses a coin to decide whether or not to copy the opinion of one of the neighbors and then it samples the neighbor. Since the outcome of the coin toss and the choice of the neighbor are independent random variables, they produce the same distribution on the new opinion of the node when commuted. In Appendix D we formalize the above equivalence and we prove it by appropriately coupling the two processes using an inductive argument.

Model $\mathcal{M}_2$ is interesting, since it describes a (lazy) voter model. As such (and as we explicitly show in the proof of Theorem 2), it is equivalent, in a probabilistic sense, to n lazy, coalescing random walks on the underlying network. This connection allows us to extrapolate the probability of consensus to a particular opinion and to adapt techniques that have been used to analyze the consensus time of the standard voter model [15, 1].

Theorem 2. *Assume model $\mathcal{M}_2$ starts in a configuration in which all nodes of a subset $W \subset V$ have opinion 1 and all other nodes have opinion 0. Let ϕ and T^{cons} denote* fixation probability *(of opinion 1) and* time to consensus, *resp. Then:* (I) $\phi = (\sum_{u \in W} d_u)/(\sum_{u \in V} d_u)$, (II) $\mathbf{E}\left[T^{\text{cons}}\right] \leqslant \beta_n T^{\text{hit}}$*, where T^{hit} is the maximum expected hitting time associated with the graph and $\beta_n = \mathcal{O}(1)$ when $\alpha \leqslant 1/2$, while $\beta_n = \ln n + 3$ when $\alpha > 1/2$.*

Sketch of the proof. We here only give a short idea of the proof and we defer a full-detailed proof to Appendix E.

The proof of (I) follows from the observation that, if we call $\mathbf{p}(t)$ the vector $\mathbf{p}(t) = (p_1(t), \ldots, p_n(t))$ where $p_i(t)$ is the probability that node i has opinion 1 at round t conditional on the configuration at the previous round $\mathbf{x}^{(t-1)}$, then for every round t it holds that $\mathbf{E}\left[\mathbf{x}^{(t+1)} \mid \mathbf{x}^{(t)}\right] = \mathbf{p}(t+1) = P\mathbf{x}^{(t)}$ where P is the transition matrix of a lazy simple random walk on the underlying graph. Iterating the above equality we have that $\lim_{t\to\infty} \mathbf{E}\left[\mathbf{x}^{(t)} \mid \mathbf{x}^{(0)}\right] = \pi^{\intercal}\mathbf{x}^{(0)}\mathbf{1}$, where π is the stationary distribution of the random walk. Finally, the formula for ϕ follows from the fact that, for each node i, $\lim_{t\to\infty} \mathbf{E}\left[x_i^{(t)} \mid \mathbf{x}^{(0)}\right]$ equals the probability that node i ends up with opinion 1 and from the fact that the stationary probability of a simple random walk being on a node i is proportional to the degree of i.

The proof of (II) is an adaptation to our (discrete) case of the proof strategy for the continuous case described in [1, Section 14.3.2]: we leverage on the relation between the convergence time of $\mathcal{M}_2$ and the maximum *meeting time* of two lazy random walks and, by using an appropriate martingale, we show that the maximum meeting time is upper bounded by the maximum hitting time (see Lemma E.1). □

4 The biased setting

Without loss of generality, in the rest of this section we assume $\alpha_{01} \neq \alpha_{10}$ and we let $r = \alpha_{01}/\alpha_{10}$. In general, the biased setting is considerably harder to address, since the connection between our model and lazy random walks no longer applies in this setting, nor does it seem easy to track the evolution of the expected behavior of the model in a way that is mathematically useful.

4.1 Asynchronous variant

In the asynchronous case, we give a result for the fixation probability holding for regular graphs, thanks to an equivalence with the fixation probability for the n-clique (see Appendix F). We also bound the expected consensus time in the specific case of the n-clique. The asynchronous variant of the process on the n-clique is equivalent to a birth-and-death chain on $n+1$ states $0, 1, 2, \ldots, n$ representing the number of nodes with opinion 1. Now, however, the transition probabilities will not be symmetric. In fact, the transition probabilities can be specified by $\{p_k, q_k, r_k\}_{k=0}^{n}$ where $p_k + q_k + r_k = 1$, and:

- p_k is the probability that the number of nodes holding opinion 1 increases from k to $k+1$ when $0 \leqslant k < n$,
- q_k is the probability that the number of nodes holding opinion 1 decreases from k to $k-1$ when $0 < k \leqslant n$,
- r_k is the probability that the number of nodes holding opinion 1 remains k when $0 \leqslant k \leqslant n$.

Due to our definition of the opinion dynamics, we have $p_0 = q_n = 0$ and, for $0 < k < n$,

$$\begin{aligned} p_k &= \left(1 - \frac{k}{n}\right) \cdot \frac{k}{n-1} \cdot \alpha_{01} = \alpha_{01} \frac{k(n-k)}{n(n-1)}, \\ q_k &= \frac{k}{n} \cdot \frac{n-k}{n-1} \cdot \alpha_{10} = \alpha_{10} \frac{k(n-k)}{n(n-1)}, \\ r_k &= 1 - (\alpha_{01} + \alpha_{10}) \frac{k(n-k)}{n(n-1)}. \end{aligned}$$

Theorem 3. *Let $r = \alpha_{01}/\alpha_{10}$ and let ϕ_k be the fixation probability of opinion 1 on a regular n-nodes graph starting from a state in which k nodes hold opinion 1. Then, for $r \notin \{0, 1\}$,*

$$\phi_k = \frac{1 - r^{-k}}{1 - r^{-n}}. \tag{3}$$

Proof. Thanks to the equivalence of the fixation probability between a regular n-nodes graph and an n-clique (Supplemental Material, Appendix F), and by the analysis of a general birth-death process (see for example [26, Section 6.2]), we get

$$\phi_k = \frac{1 + \sum_{i=1}^{k-1} \prod_{j=1}^{i} \gamma_j}{1 + \sum_{i=1}^{n-1} \prod_{j=1}^{i} \gamma_j},$$

where $\gamma_j = q_j/p_j = \alpha_{10}/\alpha_{01} = 1/r$ for all j. Hence

$$\phi_k = \frac{1 + \sum_{i=1}^{k-1} r^{-i}}{1 + \sum_{i=1}^{n-1} r^{-i}} = \frac{1 + \frac{1-r^{-k}}{1-r^{-1}} - 1}{1 + \frac{1-r^{-n}}{1-r^{-1}} - 1} = \frac{1 - r^{-k}}{1 - r^{-n}}.$$

Note that when $r \to 1$, we can evaluate ϕ_k by applying L'Hôpital's rule to (3) and get $\phi_k = k/n$, which is consistent with the results for the unbiased setting. When $r = 0$, nodes can only switch from opinion 1 to opinion 0, so clearly $\phi_k = 0$; this is also consistent with (3) when $r \to 0$. $\square$

It is interesting to note that the expression from (3) coincides with the fixation probability in the standard Moran process [26, Chapter 6] when mutants (say, nodes with opinion 1) have a *relative fitness* equal to α_{01}/α_{10}, and in the initial configuration there are k mutants out of n nodes. In other words, on an n-nodes regular graph the ratio α_{01}/α_{10} can be interpreted as a *fitness* of sorts, even though there is no notion of fitness or selection built in our model (recall that nodes are activated uniformly at random).

For the n-clique we are also able to bound the expected consensus time. While the logic of the proof is similar to the one of Theorem 1, the proof itself is considerably more involved in the asymmetric setting, leading to qualitatively different results—namely, an $O(n \log n)$ instead of an $O(n^2)$ worst case bound. The reader is deferred to the Supplementary Material (Appendix G) for the full details.

Theorem 4. *If $T_k(n)$ is the expected consensus time in the n-clique when starting from a state with k nodes holding opinion 1, and α_{01}, α_{10} are constants, then for each $k = 1, \ldots, n-1$,*

$$T_k(n) = O(n \log n),$$

and for some values of k the above bound is tight.

4.2 Synchronous variant

In order to bound the fixation probabilities, denote by $\mathbf{x}^{(t)} \in \{0,1\}^n$ the state of the system at time t. Conditioned on the state vector $\mathbf{x}^{(t)}$, the probability that $x_u^{(t+1)} = 1$ can be expressed as follows:

$$\begin{aligned} &\mathbf{P}\left(x_u^{(t+1)} = 1 \,|\, \mathbf{x}^{(t)}\right) \\ &= \begin{cases} 1 - \alpha_{10}\left(1 - \frac{\sum_{v \in V} a_{uv} x_v^{(t)}}{d_u}\right) & \text{if } x_u^{(t)} = 1 \\ \alpha_{01} \frac{\sum_{v \in V} a_{uv} x_v^{(t)}}{d_u} & \text{if } x_u^{(t)} = 0. \end{cases} \end{aligned}$$

This follows since the probability that node u samples a neighbor with opinion 1 is $\sum_v a_{uv} x_v^{(t)}/d_u$ and:

- when $x_u^{(t)} = 1$, then $x_u^{(t+1)} = 1$ iff either u samples a neighbor with opinion 1, or u samples a neighbor with opinion 0 and does not accept its opinion (these two events are disjoint);
- when $x_u^{(t)} = 0$, then $x_u^{(t+1)} = 1$ iff u samples a neighbor with opinion 1 and accepts its opinion.

We did not exploit the graph topology so far. In the case of the n-clique (with loops, to simplify some expressions), let $k^{(t)}$ be the number of nodes with opinion 1 at time t. Specializing the formulas derived above we get

$$\mathbf{P}\left(x_u^{(t+1)} = 1 \,|\, \mathbf{x}^{(t)}\right) = \begin{cases} 1 - \alpha_{10}\left(1 - \frac{k^{(t)}}{n}\right) & \text{if } x_u^{(t)} = 1 \\ \alpha_{01} \frac{k^{(t)}}{n} & \text{if } x_u^{(t)} = 0. \end{cases}$$

Note that the expression above depends only on $k^{(t)}$ and $x_u^{(t)}$, and not on the entire state $\mathbf{x}^{(t)}$. The process is thus equivalent to sampling, at each step t, $k^{(t)}$ Bernoulli random variables (r.v.) with parameter $\beta_k := 1 - \alpha_{10}(1 - k^{(t)}/n)$, and $n - k^{(t)}$ bernoulli r.v. with parameter $\gamma_k := \alpha_{01} k^{(t)}/n$. Collectively, the outcomes of these r.v. constitute the new state $\mathbf{x}(t+1)$. Then,

$$\mathbf{E}\left[k^{(t+1)} \,|\, k^{(t)}\right] = (n - k^{(t)})\alpha_{01}\frac{k^{(t)}}{n} + k^{(t)}\left(1 - \alpha_{10}(1 - \frac{k^{(t)}}{n})\right)$$

which, posing $y^{(t)} = k^{(t)}/n$, can be written as

$$\mathbf{E}\left[y^{(t+1)} \,|\, y^{(t)}\right] = y^{(t)} + (\alpha_{01} - \alpha_{10})y^{(t)}(1 - y^{(t)}). \tag{4}$$

Proposition 2. *Assume $\alpha_{01} \leqslant \alpha_{10}$. Then the fixation probability of opinion 0 is at least the fraction of agents holding opinion 0.*

Proof. Under the assumption $\alpha_{01} \leqslant \alpha_{10}$,

$$\begin{aligned}\mathbf{E}\left[y^{(t+1)}\right] &= \mathbf{E}\left[\mathbf{E}\left[y^{(t+1)} \mid y^{(t)}\right]\right] \\ &= \mathbf{E}\left[y^{(t)}\right] + (\alpha_{01} - \alpha_{10})\mathbf{E}\left[y^{(t)}(1 - y^{(t)})\right] \leqslant \mathbf{E}\left[y^{(t)}\right].\end{aligned}$$

Hence, the succession $(\mathbf{E}\left[y^{(t)}\right])_t$ is monotone and bounded and attains a limit. This limit must coincide with the fixation probability, because $y^{(t)}$ converges in distribution to a bernoulli random variable $y^{(\infty)}$ and $\mathbf{E}\left[y^{(\infty)}\right] = \mathbf{P}\left(y^{(\infty)} = 1\right) = \mathbf{P}\left(\exists t : y^{(t)} = 1\right)$ equals the fixation probability. Since $\mathbf{E}\left[y^{(0)}\right] = y^{(0)} = k^{(0)}/n$, the fixation probability of opinion 1 must be at most $k^{(0)}/n$, so that of opinion 0 is at least $1 - k^{(0)}/n$. □

Regarding the expected consensus time, we show the following by using the technique of *drift analysis* [19].

Theorem 5. *If $T_k(n)$ is the expected consensus time in the n-clique when starting from a configuration with k nodes holding opinion 1, and $\alpha_{01} = \alpha_{10} - \epsilon$, then*

$$T_k(n) \leqslant \frac{nk}{\epsilon(n-1)}.$$

In particular, $T_k(n) \leqslant \min(2k/\epsilon, n/\epsilon)$ for each $k = 1, 2, \ldots, n-1$.

Proof. We adapt a proof of [19, Theorem 2.1] to our setting, since their result is not suitable for systems with more than one absorbing state. In the remainder, we assume $\alpha_{01} < \alpha_{10}$, we let $\epsilon = \alpha_{10} - \alpha_{01}$, and we let $z^{(t)} = k^{(t)}/n$, i.e., $z^{(t)}$ is the fraction of agents with opinion 1 at time t. We begin by defining the following stopping time:

$$T := \inf\{t \geqslant 0 : z^{(t)} \in \{0, 1\}\}.$$

This definition is akin to the one given in [19, Theorem 2.1], but it accounts for the presence of two absorbing states in the Markov chain defined by $z^{(t)}$. Moreover, $z^{(t)} = z^{(t-1)}$ for every $t > T$, since $z^{(t)}$ does not change after absorption (regardless of the absorbing state). We next note that for every t, $z^{(t)} \in \mathcal{S} = \left\{0, \frac{1}{n}, \ldots, 1 - \frac{1}{n}, 1\right\}$. Moreover, for every $s \in \mathcal{S}$ we have $\mathbf{E}\left[z^{(t+1)} \mid z^{(t)} = s\right] = s - \epsilon s(1-s)$, whence:

$$\mathbf{E}\left[z^{(t)} - z^{(t+1)} \mid z^{(t)} = s\right] = \epsilon s(1-s),$$

with the last quantity at least $\epsilon \frac{1}{n}\left(1 - \frac{1}{n}\right)$ for $s \in \mathcal{S} \setminus \{0, 1\}$. Next:

$$\begin{aligned}&\mathbf{E}\left[z^{(t+1)} \mid T > t\right] \\ &= \sum_{s=1}^{n-1} \mathbf{E}\left[z^{(t+1)} \mid z^{(t)} = \frac{s}{n}\right] \cdot \mathbf{P}\left(z^{(t)} = \frac{s}{n} \mid T > t\right),\end{aligned}$$

where the first equality follows since, for $s \notin \{0, 1\}$, $z^{(t)} = s$ implies $T > t$. Similarly to the proof of [19, Theorem 2.1], the equality above implies

$$\mathbf{E}\left[z^{(t)} - z^{(t+1)} \mid T > t\right] \geqslant \epsilon \frac{1}{n}\left(1 - \frac{1}{n}\right). \tag{5}$$

We let $\delta = \epsilon\frac{1}{n}\left(1 - \frac{1}{n}\right)$ for conciseness. We next have:

$$\begin{aligned}\mathbf{E}\left[z^{(t)}\right] &\stackrel{(a)}{=} \mathbf{E}\left[z^{(t)} \mid T > t\right] \mathbf{P}\left(T > t\right) + \\ &\quad + \mathbf{P}\left(z^{(t)} = 1 \mid T \leqslant t\right) \cdot \mathbf{P}\left(T \leqslant t\right)\end{aligned}$$

and therefore

$$\begin{aligned}\mathbf{E}\left[z^{(t+1)}\right] &\stackrel{(a)}{=} \mathbf{E}\left[z^{(t+1)} \mid T > t\right] \cdot \mathbf{P}\left(T > t\right) + \\ &\quad + \mathbf{P}\left(z^{(t+1)} = 1 \mid T \leqslant t\right) \cdot \mathbf{P}\left(T \leqslant t\right) \\ &\stackrel{(b)}{\leqslant} \left(\mathbf{E}\left[z^{(t)} \mid T > t\right] - \delta\right) \cdot \mathbf{P}\left(T > t\right) + \\ &\quad + \mathbf{P}\left(z^{(t+1)} = 1 \mid T \leqslant t\right) \cdot \mathbf{P}\left(T \leqslant t\right) \\ &\stackrel{(c)}{=} \mathbf{E}\left[z^{(t)}\right] - \mathbf{P}\left(z^{(t)} = 1 \mid T \leqslant t\right) \cdot \mathbf{P}\left(T \leqslant t\right) + \\ &\quad + \mathbf{P}\left(z^{(t+1)} = 1 \mid T \leqslant t\right) \cdot \mathbf{P}\left(T \leqslant t\right) - \delta \cdot \mathbf{P}\left(T > t\right).\end{aligned}$$

In the derivations above, (a) simply follows from the law of total probability, considering that $T \leqslant t$ implies $z^{(t)} \in \{0, 1\}$, (b) follows from (5), while (c) follows by replacing the equation of $\mathbf{E}\left[z^{(t)}\right]$ into the last step of the derivation. Next, we note that

$$\mathbf{P}\left(z^{(t+1)} = 1 \mid T \leqslant t\right) = \mathbf{P}\left(z^{(t)} = 1 \mid T \leqslant t\right)$$

by definition of the $z^{(t)}$, whence we obtain:

$$\delta \cdot \mathbf{P}\left(T > t\right) \leqslant \mathbf{E}\left[z^{(t)}\right] - \mathbf{E}\left[z^{(t+1)}\right]. \tag{6}$$

Now, observe that (6) is exactly [19, (2.4) in Theorem 2.1]. From this point, the proof proceeds exactly as in [19, (2.4) in Theorem 2.1], so that we finally have:

$$\mathbf{E}\left[T\right] \leqslant \frac{z^{(0)}}{\delta} = \frac{nk}{\epsilon(n-1)},$$

if at time $t = 0$ we have k agents with opinion 1. □

5 Conclusions and Outlook

Natural directions for future work include considering more opinions and general topologies.

More opinions. The case of more opinions presents no major challenges in the unbiased case, both in its asynchronous and synchronous variants, something we did not discuss for the sake of space. In this case, one can simply focus on one opinion at a time, collapsing the remaining opinions into an "other" class. Proceeding this way, it is easy to extend the results we presented in Section 3 to the general case: for $k > 2$ opinion, the fixation probability for opinion i is $\frac{\sum_{u \in W} d_u}{\sum_u d_u}$, where W is the subset of nodes with opinion i in the initial configuration. The biased case is considerably harder and the technical barriers are twofold: one is the general difficulty of characterizing the expected change of the global state in the biased setting even in the case of 2 opinions (see next paragraph). The other is the possible presence of rock-paper-scissors like dynamics that may arise depending on the distribution of the opinion biases.

General topologies. As also suggested by previous work, albeit for different models [24, 2, 20], we believe the biased case might give rise to diverse and possibly counterintuitive behaviors. In general, a crucial technical challenge is characterizing the evolution of the global state across consecutive steps, since this in general depends on the current configuration in a way that is highly topology-dependent and hard to analyze. Some recent results [29, 30] proposed techniques relying on variants of the expander mixing lemma to investigate quasi-majority dynamics on expanders. Unfortunately, these techniques do not obviously extend to the biased voter models we consider. Indeed and interestingly, the class of dynamics these techniques apply to does not even include the standard voter model as a special case.

In general, we believe that extending and/or improving our results for the biased setting might require refining important techniques, such as those of [29, 30] or the ones discussed in [19].

Acknowledgements

- Partially supported by the ERC Advanced Grant 788893 AMDROMA "Algorithmic and Mechanism Design Research in Online Markets", the EC H2020RIA project "SoBigData++" (871042), and the MIUR PRIN project ALGADIMAR "Algorithms, Games, and Digital Markets.
- Work of the second author was carried out in association with Istituto di Analisi dei Sistemi ed Informatica, Consiglio Nazionale delle Ricerche, Italy.
- Supported by the Austrian Science Fund (FWF): P 32863-N.
- This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 947702).

References

[1] David Aldous and James Fill. Reversible markov chains and random walks on graphs. https://www.stat.berkeley.edu/~aldous/RWG/book.pdf, 2002.

[2] Aris Anagnostopoulos, Luca Becchetti, Emilio Cruciani, Francesco Pasquale, and Sara Rizzo. Biased opinion dynamics: when the devil is in the details. *Information Sciences*, 593:49–63, 2022.

[3] Chalee Asavathiratham. *The Influence Model: A Tractable Representation for the Dynamics of Networked Markov Chains.* PhD thesis, Massachussets Institute of Technology, 2000.

[4] Michael Assaf and Mauro Mobilia. Metastability and anomalous fixation in evolutionary games on scale-free networks. *Physical review letters*, 109(18):188701, 2012.

[5] G. J. Baxter, R. A. Blythe, and A. J. McKane. Fixation and consensus times on a network: A unified approach. *Physical Review Letters*, 101(25):258701, December 2008.

[6] Petra Berenbrink, George Giakkoupis, Anne-Marie Kermarrec, and Frederik Mallmann-Trenn. Bounds on the voter model in dynamic networks. In Ioannis Chatzigiannakis, Michael Mitzenmacher, Yuval Rabani, and Davide Sangiorgi, editors, *43rd International Colloquium on Automata, Languages, and Programming, ICALP 2016, July 11-15, 2016, Rome, Italy*, volume 55 of *LIPIcs*, pages 146:1–146:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016.

[7] Colin Cooper, Martin Dyer, Alan Frieze, and Nicolás Rivera. Discordant voting processes on finite graphs. *SIAM Journal on Discrete Mathematics*, 32(4):2398–2420, 2018.

[8] Colin Cooper, Robert Elsässer, Hirotaka Ono, and Tomasz Radzik. Coalescing Random Walks and Voting on Connected Graphs. *SIAM Journal on Discrete Mathematics*, 27(4):1748–1758, January 2013. Publisher: Society for Industrial and Applied Mathematics.

[9] Emilio Cruciani, Hlafo Alfie Mimun, Matteo Quattropani, and Sara Rizzo. Phase transitions of the k-majority dynamics in a biased communication model. In *ICDCN '21: International Conference on Distributed Computing and Networking, Virtual Event, Nara, Japan, January 5-8, 2021*, pages 146–155. ACM, 2021.

[10] Josep Díaz, Leslie Ann Goldberg, David Richerby, and Maria Serna. Absorption time of the Moran process. *Random Structures & Algorithms*, 49(1):137–159, 2016.

[11] Peter Donnelly and Dominic Welsh. Finite particle systems and infection models. *Mathematical Proceedings of the Cambridge Philosophical Society*, 94(1):167–182, July 1983. Publisher: Cambridge University Press.

[12] Loke Durocher, Panagiotis Karras, Andreas Pavlogiannis, and Josef Tkadlec. Invasion dynamics in the biased voter process. In Luc De Raedt, editor, *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI 2022, Vienna, Austria, 23-29 July 2022*, pages 265–271. ijcai.org, 2022.

[13] Warren Ewens. *Mathematical Population Genetics I: Theoretical Introduction.* Springer, 2012.

[14] J. Glaz. Probabilities and moments for absorption in finite homogeneous birth-death processes. *Biometrics*, 35(4):813–816, 1979.

[15] Yehuda Hassin and David Peleg. Distributed probabilistic polling and applications to proportionate agreement. *Inf. Comput.*, 171(2):248–268, 2001.

[16] Laura Hindersin and Arne Traulsen. Counterintuitive properties of the fixation time in network-structured populations. *Journal of The Royal Society Interface*, 11(99):20140606, 2014.

[17] Varun Kanade, Frederik Mallmann-Trenn, and Thomas Sauerwald. On coalescence time in graphs: When is coalescing as fast as meeting?: Extended Abstract. In *Proceedings of the 2019 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, Proceedings, pages 956–965. Society for Industrial and Applied Mathematics, January 2019.

[18] Motoo Kimura. On the probability of fixation of mutant genes in a population. *Genetics*, 47(6):713–719, 1962.

[19] Johannes Lengler and Angelika Steger. Drift analysis and evolutionary algorithms revisited. *Combinatorics, Probability and Computing*, 27(4):643–666, July 2018.

[20] Hicham Lesfari, Frédéric Giroire, and Stéphane Pérennes. Biased majority opinion dynamics: Exploiting graph k-domination. In Luc De Raedt, editor, *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI 2022, Vienna, Austria, 23-29 July 2022*, pages 377–383. ijcai.org, 2022.

[21] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov Chains and Mixing Times.* American Mathematical Society, 2009.

[22] E. Lieberman, C. Hauert, and M.A. Nowak. Evolutionary dynamics on graphs. *Nature*, 433:312–316, 2005.

[23] Thomas M. Liggett. *Interacting Particle Systems*. Springer New York, 1985.

[24] Andrea Montanari and Amin Saberi. The spread of innovations in social networks. *Proceedings of the National Academy of Sciences*, 107(47):20196–20201, 2010.

[25] P. A. P. Moran. Random processes in genetics. *Mathematical Proceedings of the Cambridge Philosophical Society*, 54(1):60–71, January 1958.

[26] Martin A. Nowak. *Evolutionary Dynamics*. Harvard University Press, 2006.

[27] Roberto I. Oliveira and Yuval Peres. Random walks on graphs: new bounds on hitting, meeting, coalescing and returning. In *2019 Proceedings of the Sixteenth Workshop on Analytic Algorithmics and Combinatorics (ANALCO)*, pages 119–126. SIAM, 2019.

[28] Antonio F Peralta, Matteo Neri, János Kertész, and Gerardo Iñiguez. Effect of algorithmic bias and network structure on coexistence, consensus, and polarization of opinions. *Physical Review E*, 104(4):044312, 2021.

[29] Grant Schoenebeck and Fang-Yi Yu. Consensus of interacting particle systems on erdös-rényi graphs. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1945–1964. SIAM, 2018.

[30] Nobutaka Shimizu and Takeharu Shiraga. Quasi-majority functional voting on expander graphs. *arXiv preprint arXiv:2002.07411*, 2020.

[31] V. Sood, Tibor Antal, and S. Redner. Voter models on heterogeneous networks. *Physical Review E*, 77(4):041121, April 2008.

Appendix

A Unbiased asynchronous model: connections to lazy random walks

For every node $v \in V$, the expected state of v at time $t+1$, conditioned on $\mathbf{x}^{(t)} = \mathbf{x}$ is

$$\mathbf{E}\left[x_v^{(t+1)} \mid \mathbf{x}^{(t)} = \mathbf{x}\right] = \left(1 - \frac{1}{n}\right) x_v + \frac{1}{n}\left((1-\alpha)x_v + \frac{\alpha}{d_v}\sum_{u \in N(v)} x_u\right) = \left(1 - \frac{\alpha}{n}\right) x_v + \frac{\alpha}{n d_v}\sum_{u \in N(v)} x_u.$$

In vector form:

$$\mathbf{E}\left[\mathbf{x}^{(t)} \mid \mathbf{x}^{(t-1)} = \mathbf{x}\right] = \left(1 - \frac{\alpha}{n}\right)\mathbf{x} + \frac{\alpha}{n} P\mathbf{x},$$

where $P = D^{-1}A$ is the transition matrix of the simple random walk on G (with D the diagonal degree matrix and A the adjacency matrix of the graph) and I is the identity matrix. From this we obtain:

$$\mathbf{E}\left[\mathbf{x}^{(t)}\right] = \left(1 - \frac{\alpha}{n}\right)\mathbf{x}^{(t-1)} + \frac{\alpha}{n} P\mathbf{x}^{(t-1)} = \hat{P}\mathbf{x}^{(t-1)},$$

where $\hat{P} := \left(1 - \frac{\alpha}{n}\right) I + \frac{\alpha}{n} P$ is a row-stochastic matrix.[5] We finally obtain

$$\mathbf{E}\left[\mathbf{x}^{(t)}\right] = \hat{P}^t \mathbf{x}^{(0)}.$$

The matrix $\hat{P}$ corresponds to an ergodic Markov chain, whose left and right eigenvectors are the same of the transition matrix $P = D^{-1}A$ of the random walk on the underlying graph G. In particular, the main left eigenvector is the stationary distribution of the random walk on G, which corresponds to (2) in Section 3.

B Proof of Proposition 1

For completeness' sake in this section we give a full proof of Proposition 1. The proof makes use of some standard relations between a Markov chain and its corresponding *lazy* version.

Lazy Markov chains

Let P be a square row-stochastic matrix[6], let $\{X_t\}_t$ be a Markov chain with transition matrix P, let $\alpha \in (0,1)$, and let $P_1 = (1-\alpha)I + \alpha P$, where I is the identity matrix with the same dimension of P. Observe that P_1 is itself a stochastic matrix and that the Markov chain with transition matrix P_1 proceeds as follows: At each step, with probability $1-\alpha$ the chain stays where it is, and with probability α it does one step according to transition matrix P. We call the Markov chain $\{Y_t\}_t$ with transition Matrix P_1 the *lazy version* of Markov chain $\{X_t\}_t$ with parameter α.

Let $\{X_t\}_t$ and $\{Y_t\}_t$ be the Markov chains with transition matrices P and P_1, respectively. We can define the Markov chain $\{Y_t\}_t$ on the same probability space of $\{X_t\}_t$ using an extra random source as follows: Let $\{B_t\}_t$ be a sequence of independent and identically distributed Bernoulli random variables such that $B_t = 0$ with probability $1-\alpha$ and $B_t = 1$ with probability α, let $\sigma_t = \sum_{i=1}^t B_t$, for every $t = 1, 2, \ldots$, and let $\{Y_t\}$ be defined as follows

$$\begin{cases} Y_0 = X_0 & \\ Y_t = X_{\sigma_t} & \text{for } t = 1, 2, \ldots \end{cases} \tag{7}$$

[5] This trivially follows since the entries of each rows of P sum to 1.

[6] A matrix whose entries are in the interval $[0,1]$ and such that the entries of each row sum up to 1

Observe that, from the coupling in (7) it follows that, if chain $\{X_t\}$ visits the sequence of states $(x_0, x_1, x_2, \dots)$ then also the chain $\{Y_t\}_t$ visits the same states, in the same order, remaining on each state for $1/\alpha$ units of time in expectation. More formally, for a state x let $\tau^Y(x)$ be the first time the chain $\{Y_t\}_t$ hits state x. From the construction of the coupling it follows that, for each sequence of states $(x_0, x_1, \dots, x_t)$ such that $P(x_i, x_{i+1}) > 0$ for every $i = 0, \dots t-1$, it holds that

$$\mathbf{E}\left[\tau^Y(x_t) \,|\, (X_0, X_1, \dots, X_t) = (x_0, x_1, \dots, x_t)\right] = t/\alpha \tag{8}$$

Hence, if we have a theorem about the hitting time of some state (or set of states) for the original chain $\{X_t\}_t$, it directly applies also to the lazy chain, with a $1/\alpha$ multiplicative factor.

Proof of Proposition 1

Let $\Omega = \{0,1\}^n$ be the state space of all possible configurations for a set of n nodes where each node can be in state either 0 or 1. For an arbitrary graph G, let $\{X_t\}_t$ be the Markov chain with state space Ω describing the voter model on G and let $\{Y_t\}_t$ be the Markov chain with state space Ω describing our unbiased asynchronous variant. Notice that, according to the definition of our unbiased asynchronous variant, $\{Y_t\}_t$ is the lazy version of $\{X_t\}_t$ with parameter α. Using the coupling in (7) it thus can be defined on the same probability space of $\{X_t\}_t$ in a way that chain $\{X_t\}_t$ ends up with all nodes in state 1 if and only if chain $\{Y_t\}_t$ ends up with all nodes in state 1. Hence, the fixation probability in our unbiased asynchronous model is equal to the fixation probability in the voter model. Moreover, observe that the *consensus* time is the hitting time of the set of two states $\{(0,\dots,0),(1,\dots,1)\} \subseteq \{0,1\}^n$. Hence, according to (8), the expected consensus time of our unbiased asynchronous model equals the expected consensus time of the voter model multiplied by $1/\alpha$. □

C Computation of $T_k(n)$ in Theorem 1

Lemma C.1. *If $T(n) = B^{-1}\mathbf{1}$, then for each $k = 1, \dots, n-1$,*

$$T_k(n) = \frac{n-1}{\alpha}\left((n-k)(H_{n-1} - H_{n-k}) + k(H_{n-1} - H_{k-1})\right).$$

Proof. After inverting the matrix B as in [14, Eq. (5)],

$$T_k(n) = \frac{\sum_{s=1}^{n-1}\left(\sum_{\ell=1}^{m(s,k)} \prod_{i=1}^{\ell-1} q_i \prod_{j=\ell}^{s-1} p_j\right)\left(\sum_{\ell=1}^{M(s,k)} \prod_{i=s+1}^{n-\ell} q_i \prod_{j=n-\ell+1}^{n-1} p_j\right)}{\sum_{\ell=1}^{n} \prod_{i=1}^{n-\ell} q_i \prod_{j=n-\ell+1}^{n-1} p_j}, \tag{9}$$

where $m(s,k) := \min(s,k)$, and $M(s,k) := n - \max(s,k)$. We already observed that $p_i = rq_i$ for all $i = 1, \dots, n-1$, where $r = \alpha_{01}/\alpha_{10}$ (we do not yet substitute $r = 1$ because the following computation will also be useful in the biased setting). Hence,

$$\begin{aligned}
T_k(n) &= \frac{\sum_{s=1}^{n-1}\left(\sum_{\ell=1}^{m(s,k)} \prod_{i=1}^{\ell-1} q_i \prod_{j=\ell}^{s-1} rq_j\right)\left(\sum_{\ell=1}^{M(s,k)} \prod_{i=s+1}^{n-\ell} q_i \prod_{j=n-\ell+1}^{n-1} rq_j\right)}{\sum_{\ell=1}^{n} \prod_{i=1}^{n-\ell} q_i \prod_{j=n-\ell+1}^{n-1} rq_j} \\
&= \frac{\sum_{s=1}^{n-1}\left(\sum_{\ell=1}^{m(s,k)} r^{s-\ell} \prod_{i=1}^{s-1} q_i\right)\left(\sum_{\ell=1}^{M(s,k)} r^{\ell-1} \prod_{i=s+1}^{n-1} q_i\right)}{\sum_{\ell=1}^{n} r^{\ell-1} \prod_{i=1}^{n-1} q_i} \\
&= \frac{\sum_{s=1}^{n-1}\left(\prod_{i=1}^{s-1} q_i \cdot \prod_{i=s+1}^{n-1} q_i\right)\left(\sum_{\ell=1}^{m(s,k)} r^{s-\ell}\right)\left(\sum_{\ell=1}^{M(s,k)} r^{\ell-1}\right)}{\prod_{i=1}^{n-1} q_i \cdot \sum_{\ell=1}^{n} r^{\ell-1}} \\
&= \frac{\prod_{i=1}^{n-1} q_i \cdot \sum_{s=1}^{n-1} q_s^{-1}\left(\sum_{\ell=1}^{m(s,k)} r^{s-\ell}\right)\left(\sum_{\ell=1}^{M(s,k)} r^{\ell-1}\right)}{\prod_{i=1}^{n-1} q_i \cdot \sum_{\ell=1}^{n} r^{\ell-1}}
\end{aligned}$$

$$= \frac{\sum_{s=1}^{n-1} q_s^{-1} \left(\sum_{\ell=1}^{m(s,k)} r^{s-\ell}\right)\left(\sum_{\ell=1}^{M(s,k)} r^{\ell-1}\right)}{\sum_{\ell=1}^{n} r^{\ell-1}}. \tag{10}$$

Using now the assumption of the unbiased setting, $r = \alpha_{01}/\alpha_{10} = 1$ and the two inner summations in the numerator simplify to $m(s,k)$ and $M(s,k)$ respectively. After substituting $q_s = \alpha\frac{s(n-s)}{n(n-1)}$,

$$\begin{aligned}
T_k(n) &= \frac{\sum_{s=1}^{n-1}\left(\frac{n(n-1)}{\alpha s(n-s)} \cdot m(s,k) \cdot M(s,k)\right)}{n} \\
&= \frac{n-1}{\alpha}\left(\sum_{s=1}^{k-1} \frac{s(n-k)}{s(n-s)} + \sum_{s=k}^{n-1} \frac{k(n-s)}{s(n-s)}\right) \\
&= \frac{n-1}{\alpha}\left((n-k)\sum_{s=1}^{k-1} \frac{1}{n-s} + k\sum_{s=k}^{n-1} \frac{1}{s}\right) \\
&= \frac{n-1}{\alpha}\left((n-k)\sum_{s=n-k+1}^{n-1} \frac{1}{s} + k\sum_{s=k}^{n-1} \frac{1}{s}\right) \\
&= \frac{n-1}{\alpha}\left((n-k)(H_{n-1} - H_{n-k}) + k(H_{n-1} - H_{k-1})\right),
\end{aligned}$$

where $H_j = \sum_{i=1}^{j} i^{-1}$ is the j-th harmonic number.

□

D Equivalence of $\mathcal{M}_1$ and $\mathcal{M}_2$

Lemma D.1. *Assume $\mathcal{M}_1$ and $\mathcal{M}_2$ are initialized with the same distribution of opinions at time 0. Let $\mathbf{x} \in \{0,1\}^n$. Then, for every $t \geqslant 0$:*

$$\mathbf{P}_{\mathcal{M}_1}\left(\mathbf{x}^{(t)} = \mathbf{x}\right) = \mathbf{P}_{\mathcal{M}_2}\left(\mathbf{x}^{(t)} = \mathbf{x}\right).$$

Proof. Denote by $\mathbf{x}^{(t)}$ the state of the system at time t and assume $\mathbf{x}^{(t)} = \mathbf{x}$, where $\mathbf{x} \in \{0,1\}^n$. In particular, suppose that $x_u = a \in \{0,1\}$. Denote by d_a and d_{1-a} respectively the number of u's neighbours holding opinions a and $1-a$ at the end of iteration t. We begin by showing that the probabilities of u holding opinion a in iteration $t+1$ are the same in the two models. We first consider the case $x_u = a$ (i.e., we are interested in the probability that u does not change opinion between iterations t and $t+1$). For $\mathcal{M}_1$ we have:

$$\mathbf{P}_{\mathcal{M}_1}\left(x_u^{(t+1)} = a \mid \mathbf{x}^{(t)} = \mathbf{x}\right) = \frac{d_a}{d_u} + \frac{d_{1-a}}{d_u}(1-\alpha) = 1 - \alpha\frac{d_{1-a}}{d_u},$$

where we used $d_a + d_{1-a} = d_u$. For $\mathcal{M}_2$ we have:

$$\begin{aligned}
&\mathbf{P}_{\mathcal{M}_2}\left(x_u^{(t+1)} = a \mid \mathbf{x}^{(t)} = \mathbf{x}\right) = (1-\alpha) + \alpha\frac{d_a}{d_u} = 1 - \alpha\left(1 - \frac{d_a}{d_u}\right) \\
&= 1 - \alpha\frac{d_{1-a}}{d_u},
\end{aligned}$$

where the last equality follows since $d_u - d_a = d_{1-a}$.

We next consider the case $x_u = 1-a$. We have:

$$\mathbf{P}_{\mathcal{M}_1}\left(x_u^{(t+1)} = a \mid \mathbf{x}^{(t)} = \mathbf{x}\right) = \frac{d_a}{d_u}\alpha,$$

$$\mathbf{P}_{\mathcal{M}_2}\left(x_u^{(t+1)} = a \mid \mathbf{x}^{(t)} = \mathbf{x}\right) = \alpha \frac{d_a}{d_u},$$

We have thus shown that

$$\mathbf{P}_{\mathcal{M}_1}\left(x_u^{(t)} = a \mid \mathbf{x}^{(t-1)} = \mathbf{x}\right) = \mathbf{P}_{\mathcal{M}_2}\left(x_u^{(t)} = a \mid \mathbf{x}^{(t-1)} = \mathbf{x}\right).$$

Moreover, since nodes' decisions are assumed *independently* by each node in every iteration both in $\mathcal{M}_1$ and $\mathcal{M}_2$, the above implies:

$$\mathbf{P}_{\mathcal{M}_1}\left(\mathbf{x}^{(t)} = \mathbf{x}' \mid \mathbf{x}^{(t-1)} = \mathbf{x}\right) = \prod_{u \in V} \mathbf{P}_{\mathcal{M}_1}\left(x_u^{(t)} = x'_u \mid \mathbf{x}^{(t-1)} = \mathbf{x}\right),$$

and the same of course holds for $\mathcal{M}_2$. As a consequence:

$$\mathbf{P}_{\mathcal{M}_1}\left(\mathbf{x}^{(t)} = \mathbf{x}' | \mathbf{x}^{(t-1)} = \mathbf{x}\right) = \mathbf{P}_{\mathcal{M}_2}\left(\mathbf{x}^{(t)} = \mathbf{x}' | \mathbf{x}^{(t-1)} = \mathbf{x}\right) \tag{11}$$

We complete the proof by inductively showing that

$$\begin{aligned} &\mathbf{P}_{\mathcal{M}_1}\left(\mathbf{x}^{(t-1)} = \mathbf{x}\right) = \mathbf{P}_{\mathcal{M}_2}\left(\mathbf{x}^{(t-1)} = \mathbf{x}\right) \\ &\Longrightarrow \mathbf{P}_{\mathcal{M}_1}\left(\mathbf{x}^{(t)} = \mathbf{x}\right) = \mathbf{P}_{\mathcal{M}_2}\left(\mathbf{x}^{(t)} = \mathbf{x}\right). \end{aligned}$$

To this purpose, note that our arguments above immediately imply that the claim is true for $t = 1$, if $\mathcal{M}_1$ and $\mathcal{M}_2$ are initialized in the same configuration $\mathbf{x}^{(0)}$. For the inductive step we have:

$$\begin{aligned} &\mathbf{P}_{\mathcal{M}_1}\left(\mathbf{x}^{(t)} = \mathbf{x}\right) \\ &= \sum_{\mathbf{z} \in \{0,1\}^V} \mathbf{P}_{\mathcal{M}_1}\left(\mathbf{x}^{(t)} = \mathbf{x} \mid \mathbf{x}^{(t-1)} = \mathbf{z}\right) \mathbf{P}_{\mathcal{M}_1}\left(\mathbf{x}^{(t-1)} = \mathbf{z}\right) \\ &= \sum_{\mathbf{z} \in \{0,1\}^V} \mathbf{P}_{\mathcal{M}_2}\left(\mathbf{x}^{(t)} = \mathbf{x} \mid \mathbf{x}^{(t-1)} = \mathbf{z}\right) \mathbf{P}_{\mathcal{M}_2}\left(\mathbf{x}^{(t-1)} = \mathbf{z}\right), \end{aligned}$$

where the last equality follows from (11) and from the inductive hypothesis. This concludes the proof. □

E Proof of Theorem 2

Proof of part (i). Note that given the state vector $\mathbf{x}^{(t)}$, the probability that $x_i^{(t+1)} = 1$ can be expressed as follows:

$$\mathbf{P}\left(x_i^{(t+1)} = 1 \,|\, \mathbf{x}^{(t)}\right) = (1-\alpha) x_i^{(t)} + \alpha \frac{\sum_{j \in V} a_{ij} x_j^{(t)}}{d_i},$$

since with probability $1-\alpha$ no opinion is copied to node i, and with probability α an opinion is copied to node i and the probability that it is opinion 1 is exactly $\sum_j a_{ij} x_j^{(t)} / d_i$.

Therefore, after defining $p_i(t) := \mathbf{P}\left(x_i^{(t)} = 1 \,|\, \mathbf{x}^{(t-1)}\right)$, we can write the model in the following form (called the *binary influence model* in [3]):

$$\begin{aligned} \mathbf{p}(t+1) &= P\mathbf{x}^{(t)} \\ \mathbf{x}(t+1) &= \mathcal{B}(\mathbf{p}(t+1)) \end{aligned}$$

where P is a row-stochastic matrix, namely

$$P_{ij} = (1-\alpha)\delta_{ij} + \alpha \frac{a_{ij}}{d_i}$$

where $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise, and $\mathcal{B}(\mathbf{p}(t))$ stands for a vector of n Bernoulli random variables, respectively with parameters $p_1(t), \dots, p_n(t)$. Observe that since each $x_i^{(t)}$ is binary, $\mathbf{p}(t) = \mathbf{E}\left[\mathbf{x}^{(t)} \,|\, \mathbf{x}^{(t-1)}\right]$. Thus, $\mathbf{E}\left[\mathbf{x}^{(t+1)} \,|\, \mathbf{x}^{(t)}\right] = \mathbf{p}^{(t+1)} = P\mathbf{x}^{(t)}$. Proceeding inductively we obtain

$$\begin{aligned}\mathbf{E}\left[\mathbf{x}^{(t)} \,|\, \mathbf{x}^{(0)}\right] &= \mathbf{E}\left[\mathbf{E}\left[\mathbf{x}^{(t)} \,|\, \mathbf{x}^{(t-1)}\right] \,|\, \mathbf{x}^{(0)}\right] \\ &= \mathbf{E}\left[P\mathbf{x}^{(t-1)} \,|\, \mathbf{x}^{(0)}\right] = P\mathbf{E}\left[\mathbf{x}^{(t-1)} \,|\, \mathbf{x}^{(0)}\right] = P^t\mathbf{x}^{(0)}.\end{aligned}$$

Since the operator P is ergodic (being identical to the transition matrix of a lazy random walk on a connected graph),

$$\lim_{t\to\infty} \mathbf{E}\left[\mathbf{x}^{(t)} \,|\, \mathbf{x}^{(0)}\right] = \lim_{t\to\infty} P^t\mathbf{x}^{(0)} = \mathbf{1}\pi\mathbf{x}^{(0)},$$

where π is the left dominant eigenvector of P, normalized so that $\pi\mathbf{1} = 1$. Note that π represents the unique stationary distribution associated to P. In particular, since P is the transition matrix of a lazy random walk on the graph, $\pi_i = d_i / \sum_{j\in V} d_j$ and since by assumption $x_i^{(0)} = 1$ if $i \in W$ and 0 otherwise, we get

$$\lim_{t\to\infty} \mathbf{E}\left[\mathbf{x}^{(t)} \,|\, \mathbf{x}^{(0)}\right] = \mathbf{1}\pi\mathbf{x}^{(0)} = \mathbf{1}\frac{\sum_{u\in W} d_u}{\sum_{u\in V} d_u}.$$

Finally, recalling that $\mathbf{E}\left[x_u^{(t)} \,|\, \mathbf{x}^{(0)}\right] = \mathbf{P}\left(x_u^{(t)} = 1 \,|\, \mathbf{x}^{(0)}\right)$ since the $x_u^{(t)}$ are binary random variables, we get $\phi = (\sum_{u\in W} d_u)/(\sum_{u\in V} d_u)$. □

Proof of part (ii). We can achieve a bound on the expected consensus time by leveraging results that appear in a number of papers and are condensed in [1, Section 14.3.2] for the case of continuous-time random walks, while literature on the synchronous, discrete setting is sparser. For this reason, in the remainder we retrace the main points of the proof, adapting it to the discrete-time, synchronous case. To this purpose, we need some additional notation. Following [1] and [21], considered a node/state v of a (henceforth, ergodic and reversible) Markov chain, we denote by T_v^{hit} the *hitting time* of v, i.e., the number of steps till the chain reaches v for the first time. We denote by $\mathbf{E}_u\left[T_v^{\text{hit}}\right]$ the expected hitting time of v when the random walk starts at node u. Recall that the hitting time for a given graph is the maximum of $\mathbf{E}_u\left[T_v^{\text{hit}}\right]$ over all possible choices of u and v, i.e., $T^{\text{hit}} := \max_{u,v} \mathbf{E}_u\left[T_v^{\text{hit}}\right]$. Considered some probability distribution $\mathcal{D}$ over the states of the Markov chain, with a slight abuse of notation, we denote by $\mathbf{E}_{\mathcal{D}}\left[T_v^{\text{hit}}\right]$ the expected hitting time of v when the random walk starts at $u \sim \mathcal{D}$. Moreover, we denote by M_{uv} the *meeting time* of two independent copies of the Markov chain started at u and v respectively. Finally, the correspondence between our voter-like model and lazy random walks allows us to conclude that the consensus time T^{cons} follows the same distribution as the *coalescing time* of the corresponding, lazy random walk (for the reasons behind this fact, refer to [1, Section 14.3] or [15]). This time is defined with respect to a coalescing (possibly lazy) random walk over G: at time 0, we start n independent random walks, one per node/state in G. Assume there are x surviving walks at the end of step t; then, in step $t+1$, each of the x walks moves to a random neighbor, independently of the others; if two (or more) walks move to the same node in step $t+1$, they stick together thereafter, moving as a single one. Time T^{cons} is defined as the first step in which there is only one surviving walk. In light of the above considerations, in the remainder of this proof, we use both the terms "consensus time" and "coalescing time" to refer to T^{cons}.

Lemma E.1. *For a discrete, ergodic and reversible Markov chain, for every $u, v \in V$ we have:*

$$\max_{u,v} \mathbf{E}\left[M_{uv}\right] \leqslant \max_{u,v} \mathbf{E}_u \left[T_v^{\text{hit}}\right].$$

Proof. We follow the very same lines as the proof of [1, Proposition 14.5], which is given for continuous Markov chains. First of all, for $u, v \in V$, let X_t and Y_t be the chains at time t respectively started at u and v. For $x, y \in V$, we define the following function:

$$f(x,y) = \mathbf{E}_x\left[T_y\right] - \mathbf{E}_\pi\left[T_y\right].$$

Next, we define the following random variable:

$$S_t = \begin{cases} 2t + f(X_t, Y_t) & \text{if } 0 \leqslant t \leqslant M_{uv}, \\ S_{t-1} & \text{if } t > M_{uv}. \end{cases} \tag{12}$$

The S_t's form a martingale. It should be noted that here, both the length of the sequence (i.e., M_{uv}) and the values of the S_t's depend on the randomness of two walks started at u and v respectively. To prove this first result, we proceed in a way similar to the proof of [1, Proposition 3.3]. Clearly, we have $\mathbf{E}\left[S_t \mid S_1, \dots, S_{t-1}\right] = S_{t-1}$, whenever $t > M_{uv}$. Next, assume $(X_i, Y_i) = (x_i, y_i)$, for $i = 1, \dots, t-1$, with $x_i \neq y_i$ for every i. We first note that it is enough to prove that

$$\begin{aligned} \mathbf{E}\left[S_t \mid \bigcap_{i=1}^{t-1}(X_i, Y_i) = (x_i, y_i)\right] &= 2t + \mathbf{E}\left[f(X_t, Y_t) \mid \bigcap_{i=1}^{t-1}(X_i, Y_i) = (x_i, y_i)\right] \\ &= 2t + \mathbf{E}\left[f(X_t, Y_t) \mid (X_{t-1}, Y_{t-1}) = (x_{t-1}, y_{t-1})\right] \end{aligned}$$

where the second equality follows from the Markov property, which clearly also applies to $f(X_t, Y_t)$. We therefore have, if P is the transition matrix of a reversible Markov chain:

$$\begin{aligned} \mathbf{E}\left[S_t | S_1, \dots, S_{t-1}\right] &= 2t + \sum_{x,y} P(x_{t-1}, x) P(y_{t-1}, y) f(x,y) = 2t + \sum_x P(x_{t-1}, x) \sum_y P(y_{t-1}, y) f(x,y) \\ &\overset{(a)}{=} 2t + \sum_x P(x_{t-1}, x)(f(x, y_{t-1}) - 1) = 2t + \sum_x P(x_{t-1}, x) f(x, y_{t-1}) - 1 \\ &\overset{(a)}{=} 2t + f(x_{t-1}, y_{t-1}) - 2 = S_{t-1}, \end{aligned}$$

where (a) follows since, for $x \neq y$, by the one-step recurrence of $f(x,y)$ we get $f(x,y) = 1 + \sum_z P(x,z) f(z,y)$ (see [1, Proposition 3.3]).

Expectation of meeting time. Consider any two states u and v. We have from the definition of S_t:

$$S_{M_{uv}} = 2M_{uv} + f(X_{M_{uv}}, Y_{M_{uv}}),$$

hence:

$$\mathbf{E}\left[S_{M_{uv}}\right] = 2\mathbf{E}\left[M_{uv}\right] + \mathbf{E}_{(X_t,Y_t)}\left[\mathbf{E}_{X_{M_{uv}}}\left[T_{Y_{M_{uv}}}\right] - \mathbf{E}_\pi\left[T_{Y_{M_{uv}}}\right]\right] = 2\mathbf{E}\left[M_{uv}\right] - \mathbf{E}_{(X_t,Y_t)}\left[\mathbf{E}_\pi\left[T_{Y_{M_{uv}}}\right]\right],$$

where the second equality follows since, for every realization of (X_t, Y_t), we have $X_{M_{uv}} \equiv Y_{M_{uv}}$ by definition of M_{uv}. Next,

$$\mathbf{E}_{(X_t,Y_t)}\left[\mathbf{E}_\pi\left[T_{Y_{M_{uv}}}\right]\right] = \sum_y \mathbf{E}_\pi\left[T_y\right] \mathbf{P}\left(X_t \text{ and } Y_t \text{ meet at } y\right) \leqslant \max_{ij} \mathbf{E}_i\left[T_j\right],$$

which implies:

$$\mathbf{E}\left[S_{M_{uv}}\right] \geqslant 2\mathbf{E}\left[M_{uv}\right] - \max_{ij} \mathbf{E}_i\left[T_j\right]. \tag{13}$$

On the other hand, since S_t is a martingale, we can apply the optional stopping theorem, whence:

$$\mathbf{E}\left[S_{M_{uv}}\right] = \mathbf{E}\left[S_0\right] = \mathbf{E}\left[f(u,v)\right] = \mathbf{E}_u\left[T_v\right] - \mathbf{E}_\pi\left[T_v\right] \tag{14}$$

Putting together (13) and (14) and noting that $\mathbf{E}_u\left[T_v\right] \leqslant \max_{ij} \mathbf{E}_i\left[T_j\right]$, we conclude that:

$$2\mathbf{E}\left[M_{uv}\right] \leqslant \mathbf{E}\left[S_{M_{uv}}\right] + \max_{ij} \mathbf{E}_i\left[T_j\right] = \mathbf{E}_u\left[T_v\right] - \mathbf{E}_\pi\left[T_v\right] + \max_{ij} \mathbf{E}_i\left[T_j\right] \leqslant 2\max_{ij} \mathbf{E}_i\left[T_j\right],$$

whence the thesis of Lemma E.1.

□

Using the aforementioned, well-known equivalence between (lazy) coalescing random walks and the voter model (see [15] or [1, Section 14.3]), we next provide an upper bound on $\mathbf{E}\left[T^{\text{cons}}\right]$, which also provides a bound on the convergence time of the voter model. Note that this result is only provided for the sake of completeness and holds for any reversible Markov chain.

First of all, we fix an arbitrary ordering of the nodes of G, let it be $u_1, \dots, u_n$, and we label the random walks accordingly. We then consider the following, equivalent formulation of the coalescing random walk process: whenever two or more random walks meet at step t, they coalesce and follow thereafter the future path of the lower-labeled random walk. Note that we deterministically have $T^{\text{cons}} \geqslant \max_v M_{u_1 v}$, with the inequality possibly strict.

Next, let $m = \max_{u,v} \mathbf{E}\left[M_{u,v}\right]$ and note that $m \leqslant T^{\text{hit}}$ from Lemma E.1. We have:

$$\begin{aligned}
\mathbf{E}\left[T^{\text{cons}}\right] &= \sum_{t=0}^{\infty} \mathbf{P}\left(T^{\text{cons}} \geqslant t\right) \leqslant \sum_{t=0}^{\infty} \mathbf{P}\left(\max_v M_{u_1 v} \geqslant t\right) \\
&\leqslant \sum_{t=0}^{\infty} \min\left\{1, \sum_{v \neq u_1} \mathbf{P}\left(M_{u_1 v} \geqslant t\right)\right\}.
\end{aligned}$$

On the other hand, by proceeding as in [1, Section 2.4.3]) to derive (2.20) we obtain, for every $v \in V$: $\mathbf{P}\left(M_{u_1 v} \geqslant t\right) \leqslant e^{-\lfloor \frac{t}{m} \rfloor}$, whence:

$$\begin{aligned}
\mathbf{E}\left[T^{\text{cons}}\right] &\leqslant \sum_{t=0}^{\infty} \min\left\{1, ne^{-\lfloor \frac{t}{m} \rfloor}\right\} \leqslant \sum_{t=0}^{\infty} \min\left\{1, ne^{-\frac{t-1}{m}}\right\} \\
&< 1 + \int_0^{\infty} \min\left\{1, ne^{-\frac{t-1}{m}}\right\} dt \\
&\overset{(a)}{=} 1 + \int_0^{1+m\ln n} 1 dt + n \int_{1+m\ln n}^{\infty} e^{-\frac{t-1}{m}} dt \\
&= 2 + m\ln n + m \leqslant (\ln n + 3) T^{\text{hit}},
\end{aligned}$$

where (a) follows from the fact that $ne^{-\frac{t-1}{m}} < 1$ for $t > 1 + m\ln n$.

Remark for the case $\alpha \leqslant 1/2$. Note that whenever $\alpha \leqslant 1/2$, it is possible to get a stronger bound on the expected consensus time. In the remainder, we write the transition matrix of the process in compact form as: $P = (1-\alpha)I + \alpha Q$, where Q is the transition matrix associated to a random walk on G, namely, $Q_{ij} = \frac{a_{ij}}{d_i}$. We denote by $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$ the eigenvalues of Q. Being G undirected, Q and P define reversible Markov chains. Moreover, if $\mathbf{w}$ is a right eigenvector of Q with eigenvalue λ:

$$P\mathbf{w} = (1-\alpha)\mathbf{w} + \alpha Q\mathbf{w} = (1-\alpha+\alpha\lambda)\mathbf{w},$$

i.e., $\mathbf{w}$ is also a right eigenvector of P with eigenvalue $\lambda' = 1 - \alpha + \alpha\lambda$. In particular, we have $\lambda' \geqslant 0$, whenever $\alpha \leqslant 1/(1-\lambda)$. Since Q is stochastic, we have $\lambda_i \in [-1, 1]$ for every i. Therefore, whenever $\alpha \leqslant \frac{1}{2}$, we also have $\lambda'_i \geqslant 0$, for every i, namely P is positive semidefinite (PSD). In this case, we can apply a number of recent results for lazy random walks [8, 17, 27]. In particular, [27, Theorem 6] immediately implies that the expected time of convergence to consensus, which corresponds to the coalescing time of n coalescing random walks on G satisfies:

$$\mathbf{E}\left[T^{\text{cons}}\right] \leqslant c\, T^{\text{hit}},$$

where c is a universal constant and T^{hit} is the largest expected hitting time on G. □

F Biased fixation probability on regular graphs

The analysis from the subsection 4 can be extended from the clique to any Δ-regular graph as follows [10]. Let S be the set of nodes with opinion 1 at any given step. The transition probabilities of the birth-and-death chain on $\{0, 1, \ldots, n\}$ will now generally depend on S, and not just on the size $k = |S|$. For example, if p_S is the probability of transitioning from the set S of nodes with opinion 1 to a set of size of $|S| + 1$,

$$\begin{aligned} p_S &= \sum_{u \notin S} \mathbf{P}\left(u\right) \mathbf{P}\left(v \in S \,|\, u\right) \alpha_{01} \\ &= \sum_{u \notin S} \frac{1}{n} \frac{|N(u) \cap S|}{\Delta} \alpha_{01} \\ &= \frac{\alpha_{01}}{n\Delta} |\partial S|, \end{aligned}$$

where $N(u)$ is the set of nodes adjacent to u and ∂S is the cut through S.

Similarly, if q_S is the probability of transitioning from the set S of nodes with opinion 1 to a set of size $|S| - 1$,

$$q_S = \frac{\alpha_{10}}{n\Delta} |\partial \bar{S}| = \frac{\alpha_{10}}{n\Delta} |\partial S|.$$

The key observation is that the ratio $p_S/q_S = \alpha_{01}/\alpha_{10} = r$ equals the fitness and does not depend on the set S at all. Therefore, the process can still be cast as a birth-and-death process if we ignore all steps in which the number of nodes with opinion 1 does not change: the presence of such steps is irrelevant for the fixation probability. Therefore, by the same analysis of the previous section, we obtain a fixation probability of $(1 - r^{-k})/(1 - r^{-n})$ when starting from a configuration with k nodes holding opinion 1.

G Proof of Theorem 4

Define the matrix B as in the proof of Theorem 1; note that in this setting $p_i = r q_i$ for all $i = 1, \ldots, n-1$, where $r = \alpha_{01}/\alpha_{10}$. Recall the formulation for the expected consensus time starting from k nodes with opinion 1, namely the k-th entry of the vector $T(n) = B^{-1}\mathbf{1}$, which by proceeding as in the proof of Lemma C.1 can be computed as

$$T_k(n) = \frac{\sum_{s=1}^{n-1} q_s^{-1} \left(\sum_{\ell=1}^{m(s,k)} r^{s-\ell}\right) \left(\sum_{\ell=1}^{M(s,k)} r^{\ell-1}\right)}{\sum_{\ell=1}^{n} r^{\ell-1}},$$

with $m(s,k) := \min(s,k)$, and $M(s,k) := n - \max(s,k)$.

The above expression for $T_k(n)$, although exact, is very hard to understand asymptotically, therefore we proceed to bound it asymptotically with simpler expressions. For the remainder of the proof we assume $r < 1$ (i.e. $\alpha_{01} < \alpha_{10}$); this is without loss of generality, up to a relabeling of the opinions.

Using a difference of geometric sums, we can write

$$\sum_{\ell=1}^{\min(s,k)} r^{s-\ell} = \sum_{\ell=0}^{s-1} r^{\ell} - \sum_{\ell=0}^{\max(-1,s-k-1)} r^{\ell} = \frac{1-r^s}{1-r} - \frac{1-r^{\max(0,s-k)}}{1-r} = \frac{r^{\max(0,s-k)} - r^s}{1-r}.$$

Therefore, substituting $q_s = \alpha_{10}\frac{s(n-s)}{n(n-1)}$ and computing the other geometric sums,

$$\begin{aligned} T_k(n) &= \frac{\sum_{s=1}^{n-1} \frac{n(n-1)}{\alpha_{10}s(n-s)} \left(\frac{r^{\max(0,s-k)}-r^s}{1-r}\right)\left(\frac{1-r^{M(s,k)}}{1-r}\right)}{\frac{1-r^n}{1-r}} \\ &= \frac{n(n-1)}{\alpha_{10}(1-r)(1-r^n)} \sum_{s=1}^{n-1} \frac{(r^{\max(0,s-k)} - r^s)(1-r^{M(s,k)})}{s(n-s)}. \end{aligned}$$

Note that the $\max(\cdot)$ in the exponents (also hidden in $M(s,k) = n - \max(s,k)$) can be removed by splitting the sum into two parts at $s = k$. In particular, we also have

$$T_k(n) = \frac{n(n-1)}{\alpha_{10}(1-r)(1-r^n)} \left(\sum_{s=1}^{k-1} \frac{(1-r^s)(1-r^{n-k})}{s(n-s)} + \sum_{s=k}^{n-1} \frac{(r^{s-k}-r^s)(1-r^{n-s})}{s(n-s)} \right).$$

Note that $T_k(n) = \frac{n(n-1)}{\alpha_{10}(1-r)(1-r^n)}(S_1 - S_2)$, where

$$\begin{aligned} S_1 &= (1-r^{n-k}) \left[\sum_{s=1}^{k-1} \frac{1}{s(n-s)} - \sum_{s=1}^{k-1} \frac{r^s}{s(n-s)} \right] \\ S_2 &= (r^{-k}-1) \left[r^n \sum_{s=k}^{n-1} \frac{1}{s(n-s)} - \sum_{s=k}^{n-1} \frac{r^s}{s(n-s)} \right] \end{aligned}$$

For the first term S_1 above:

$$(1-r^{n-k}) \left[\sum_{s=1}^{k-1} \frac{1}{s(n-s)} - \sum_{s=1}^{k-1} \frac{r^s}{s(n-s)} \right] = \frac{1-r^{n-k}}{n} \left[H_{n-1} + (H_{k-1} - H_{n-k}) - n \sum_{s=1}^{k-1} \frac{r^s}{s(n-s)} \right],$$

where we used

$$\begin{aligned} \sum_{s=1}^{k-1} \frac{1}{s(n-s)} &= \frac{1}{n} \sum_{s=1}^{k-1} \left(\frac{1}{s} + \frac{1}{n-s} \right) = \frac{1}{n} \left(\sum_{s=1}^{k-1} \frac{1}{s} + \sum_{s=n-k+1}^{n-1} \frac{1}{s} \right) \\ &= \frac{1}{n}[H_{n-1} + (H_{k-1} - H_{n-k})]. \end{aligned}$$

For the second term S_2 above:

$$(r^{-k}-1) \left[r^n \sum_{s=k}^{n-1} \frac{1}{s(n-s)} - \sum_{s=k}^{n-1} \frac{r^s}{s(n-s)} \right] = \frac{r^{-k}-1}{n} \left[r^n[H_{n-1} - (H_{k-1} - H_{n-k})] - n \sum_{s=k}^{n-1} \frac{r^s}{s(n-s)} \right],$$

where we used

$$\begin{aligned} \sum_{s=k}^{n-1} \frac{1}{s(n-s)} &= \frac{1}{n} \sum_{s=k}^{n-1} \left(\frac{1}{s} + \frac{1}{n-s} \right) = \frac{1}{n} \left[\sum_{s=1}^{n-1} \left(\frac{1}{s} + \frac{1}{n-s} \right) - \sum_{s=1}^{k-1} \left(\frac{1}{s} + \frac{1}{n-s} \right) \right] \\ &= \frac{1}{n}[2H_{n-1} - H_{n-1} - (H_{k-1} - H_{n-k})] \end{aligned}$$

$$= \frac{1}{n}[H_{n-1} - (H_{k-1} - H_{n-k})]$$

Putting the two equalities for S_1 and S_2 together,

$$\begin{aligned}
n(S_1 - S_2) = (1 - r^{n-k}) &\left[H_{n-1} + (H_{k-1} - H_{n-k}) - n \sum_{s=1}^{k-1} \frac{r^s}{s(n-s)} \right] + \\
&- (r^{-k} - 1) \left[r^n (H_{n-1} - (H_{k-1} - H_{n-k})) - n \sum_{s=k}^{n-1} \frac{r^s}{s(n-s)} \right] \\
&= (1 - 2r^{n-k} + r^n) H_{n-1} + (1 - r^n)(H_{k-1} - H_{n-k}) + \\
&- (1 - r^{n-k}) n \sum_{s=1}^{k-1} \frac{r^s}{s(n-s)} + (r^{-k} - 1) n \sum_{s=k}^{n-1} \frac{r^s}{s(n-s)} \\
&= (1 - 2r^{n-k} + r^n) H_{n-1} + (1 - r^n)(H_{k-1} - H_{n-k}) + \\
&- (1 - r^{n-k}) \sum_{s=1}^{k-1} \left[\frac{r^s}{s} + \frac{r^s}{n-s} \right] + (r^{-k} - 1) \sum_{s=k}^{n-1} \left[\frac{r^s}{s} + \frac{r^s}{n-s} \right] \\
&= (1 - 2r^{n-k} + r^n) H_{n-1} + (1 - r^n)(H_{k-1} - H_{n-k}) + \\
&- 2 \sum_{s=1}^{n-1} \frac{r^s}{s} + r^{n-k} \sum_{s=1}^{k-1} \left[\frac{r^s}{s} + \frac{r^s}{n-s} \right] + r^{-k} \sum_{s=k}^{n-1} \left[\frac{r^s}{s} + \frac{r^s}{n-s} \right] \\
&= (1 - 2r^{n-k} + r^n) H_{n-1} + (1 - r^n)(H_{k-1} - H_{n-k}) + \\
&- 2(1 - r^{n-k}) \sum_{s=1}^{n-1} \frac{r^s}{s} + r^{-k}(1 - r^n) \sum_{s=k}^{n-1} \left[\frac{r^s}{s} + \frac{r^s}{n-s} \right].
\end{aligned}$$

Note that, using the assumption $r < 1$,

- $(1 - 2r^{n-k} + r^n) H_{n-1} + (1 - r^n)(H_{k-1} - H_{n-k}) < 2H_{n-1} < 2 \log n + 2$
- $-2(1 - r^{n-k}) \sum_{s=1}^{n-1} \frac{r^s}{s} < -2(1 - r) r$
- $r^{-k}(1 - r^n) \sum_{s=k}^{n-1} \left[\frac{r^s}{s} + \frac{r^s}{n-s} \right] < 2r^{-k}(1 - r^n) \sum_{s=k}^{n-1} r^s = 2r^{-k}(1 - r^n) \frac{r^k - r^n}{1-r} < \frac{2(1-r^n)}{1-r}$

Hence it follows that $n(S_1 - S_2) = O(\log n)$. Therefore, when r and α_{10} are constant with respect to n and for any value of k it holds:

$$T_k(n) = \frac{n(n-1)}{\alpha_{10}(1-r)(1-r^n)} (S_1 - S_2) = O(n \log n).$$

Note that such a bound is asymptotically tight for some values of k. In fact, let us consider $k = \frac{n+1}{2}$ (for odd values of n). Note that:

- $(1 - 2r^{n-k} + r^n) H_{n-1} + (1 - r^n)(H_{k-1} - H_{n-k}) = (1 - 2r^{n-k} + r^n) H_{n-1} > (1 - 2r^{\frac{n-1}{2}}) H_{n-1}$
- $-2(1 - r^{n-k}) \sum_{s=1}^{n-1} \frac{r^s}{s} > -2(1 - r^{n-k}) \sum_{s=1}^{\infty} \frac{r^s}{s} = 2(1 - r^n) \ln(1 - r)$
- $r^{-k}(1 - r^n) \sum_{s=k}^{n-1} \left[\frac{r^s}{s} + \frac{r^s}{n-s} \right] > 2r^{-k}(1 - r^n) \frac{r^k}{k} > \frac{2(1-r^n)}{n}$

Hence it follows that there exists k (namely $k = \frac{n+1}{2}$) such that $n(S_1 - S_2) = \Omega(\log n)$. Therefore, when r and α_{10} are constant with respect to n, for $k = \frac{n+1}{2}$ it holds:

$$T_k(n) = \frac{n(n-1)}{\alpha_{10}(1-r)(1-r^n)} (S_1 - S_2) = \Omega(n \log n).$$