

The Linear Voting Model*

Colin Cooper¹ and Nicolas Rivera²

1 Department of Informatics, King's College London, London, UK
colin.cooper@kcl.ac.uk

2 Department of Informatics, King's College London, London, UK
nicolas.rivera@kcl.ac.uk

Abstract

We study voting models on graphs. In the beginning, the vertices of a given graph have some initial opinion. Over time, the opinions on the vertices change by interactions between graph neighbours. Under suitable conditions the system evolves to a state in which all vertices have the same opinion. In this work, we consider a new model of voting, called the Linear Voting Model. This model can be seen as a generalization of several models of voting, including among others, pull voting and push voting. One advantage of our model is that, even though it is very general, it has a rich structure making the analysis tractable. In particular we are able to solve the basic question about voting, the probability that certain opinion wins the poll, and furthermore, given appropriate conditions, we are able to bound the expected time until some opinion wins.

1998 ACM Subject Classification C.2.4 Distributed Systems, G.2 Discrete Mathematics, G.3 Probability and Statistics

Keywords and phrases Voter model, Interacting particles, Randomized algorithm, Probabilistic voting

Digital Object Identifier 10.4230/LIPIcs.ICALP.2016.144

1 Introduction

Graphs are very popular as a simple model of the complex environment in which individuals interact. In this paper we focus in voting models on finite graphs, in which vertices of a given graph have opinions and by interacting with their neighbours they change such opinions. Voting models can be used to mimic real-life situations such as the spread of opinions or infections in a society, the evolution of species or models of particle interaction in physics.

While many models has been proposed in the literature, we do not aim to propose a new particular model, but to unify some of the existing models in a tractable way. With this in mind, we propose a general model of voting, called the Linear Voting Model. This model subsumes several models proposed in the past, including, for example, the push model and the very popular pull model.

Even though the voter model has been widely studied in the case of infinite structures, one of the first rigorous studies on finite structures was made by Donnelly and Welsh [4]. In that work, the authors studied a continuous-time version of the pull voting model and, under the name of infection model, the push voting model. In the continuous time version, each vertex has an exponential clock and when it rings, the vertex selects a random neighbour and pulls its opinion (in the case of pull voting) or pushes its opinion on the neighbour (in the

* Research supported by EPSRC grant EP/M005038/1, “Randomized algorithms for computer networks”. N. Rivera was supported by funding from Becas CHILE.



© Colin Cooper and Nicolas Rivera;
licensed under Creative Commons License CC-BY

43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016).

Editors: Ioannis Chatzigiannakis, Michael Mitzenmacher, Yuval Rabani, and Davide Sangiorgi;
Article No. 144; pp. 144:1–144:12



Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany



case of push voting). On the other hand, Hassin and Peleg [6] and Nakata et al. [8] studied the discrete time version of pull voting, in which vertices do not have a clock but at each round each vertex synchronously pulls an opinion. Both papers considered the two-party model and studied its possible application to distributing computing, in particular to the agreement problem. The focus of [6] and [8] is on the probability that all vertices eventually adopt the opinion which was initially held by a given subset of vertices. They proved that the probability that opinion A wins is $d(A)/d(V)$, where $d(X)$ is the sum of the degrees of the vertices of $X \subseteq V$ and A is the set of vertices whose initial opinion was A .

The consensus time of G , i.e., the time needed for the vertices of a graph G to agree on an opinion during voting, has attracted a lot of attention, especially because a low consensus time implies a better distributed algorithm for the agreement problem. In the continuous-time setting, Oliveira [9] shows that the expected consensus time is $\mathcal{O}(H_{\max})$, where $H_{\max} = \max_{v,u \in V} H(v,u)$ and $H(v,u)$ is the hitting time of u of a random walk starting at vertex v . Furthermore, in a later work [10], Olivera proved that under certain conditions on the underlying graph G , the consensus time is concentrated around $2m(G)$, where $m(G)$ is the meeting time of two independent random walk starting in stationary distribution. It is however, not clear whether the continuous-time results apply to the discrete-time setting. Hassin and Peleg [6] using a dual process, the coalescing random walk, proved that the expected consensus time is $\mathcal{O}(m(G) \log(n))$, where $m(G)$ is the meeting time of independent discrete-time random walks, thus giving $\mathcal{O}(n^3 \log(n))$ in the worst case. By using the same approach, Cooper et al. [2] improved the previous result and proved that the consensus time is $\mathcal{O}(n/(\nu(1 - \lambda_2)))$, where n is the number of vertices of G and ν is a parameter that measures the regularity of the degree sequence, ranging from 1 for regular graphs to $\Theta(n)$ for the star graph. The result of Cooper et al. achieves an upper bound of $\mathcal{O}(n^3)$ in the worst case. Berenbrink et al. [1] used a more ad hoc approach and proved that the consensus time is $\mathcal{O}((d_{\text{ave}}/d_{\text{min}})(n/\Phi))$ where Φ is the conductance of the graph, and d_{ave} , d_{min} are the average and minimum degrees respectively.

The consensus time for the push model has not been so widely studied. Push voting is a particular class of the so-called Moran process. Díaz et al. [3] proved that the the consensus time is $\mathcal{O}(n^4 q)$ where q is the square of the sum of the inverses of the degree sequence of G , giving a consensus time of $\mathcal{O}(n^6)$ in the worst case.

1.1 Our model and results

Let $G = (V, E)$ be a graph with $|V| = n$. Define a configuration of opinions as a $n \times 1$ vector $\xi \in Q^V$, where $Q = \{0, 1\}$ for the two party model, or $Q = \{1, \dots, n\}$ if we want to allow more parties.

Let $\mathcal{M}(V)$ be the set of all $n \times n$ matrices indexed by the elements of V , with exactly one 1 entry per row and all other elements 0. Also, define $\Pi(V)$ as the set of probability measures on $\mathcal{M}(V)$. If no confusion arises, we will just write \mathcal{M} instead of $\mathcal{M}(V)$ and Π instead $\Pi(V)$.

Let $l \in \Pi$ be a distribution over matrices in \mathcal{M} . Given an initial configuration ξ , we define the process $(\xi_t)_{t \geq 0}$, with t running over the non-negative integers, as

$$\xi_t = \begin{cases} \xi, & \text{if } t = 0, \\ M_{t-1} \xi_{t-1}, & \text{if } t > 0, \end{cases} \quad (1)$$

where M_t are i.i.d matrices sampled from l , and $M\xi$ is the standard matrix-vector multiplication. The above process is called a **linear voting model** with parameters (l, ξ) and it

is denoted by $(\xi_t) \sim \mathcal{LVM}(l, \xi)$. Clearly, $\xi_t(v)$ represents the opinion of vertex v at round t . Consider $M \in \mathcal{M}$ and $\xi' = M\xi$, then if all vertices have different opinions, we have that $\xi'(v) = \xi(w)$ if and only if $M(v, w) = 1$. Since M has only one 1 in each row, the voting is well-defined in the sense at every round each vertex adopts the opinion of only one vertex (including itself). Examples of linear voting models include the pull voting (asynchronous or synchronous) and the push voting model.

We proceed to present our contribution. Theorem 1 of this paper gives the probability a particular opinion wins. This generalises the approach used in [6]. Theorem 2 gives an upper bound to the expected consensus time. Our technique is qualitatively different from the approach of previous authors which depended on a detailed dualisation of the voting process, indeed, we follow an approach similar to Levin et al. [7, chapter 17] or Berenbrink [1].

Let $l \in \Pi$ and define the mean matrix H of l as

$$H = H(l) = \sum_{M \in \mathcal{M}} l(M)M.$$

From Lemma 4 we have that H is the transition Matrix of a Markov Chain with state space V . We denote by μ the stationary distribution of H (if any). Define the consensus time τ_{cons} as the first time all the opinions are the same, i.e, there exists c such that $\xi_{\tau_{\text{cons}}}(v) = c$ for all $v \in V$. Observe τ_{cons} is a stopping time and that c is the final opinion of the vertices. We have the following theorem about the winning probability.

► **Theorem 1.** *Let $(\xi_t) \sim \mathcal{LVM}(l, \xi)$ be a linear voting model with mean matrix H with $\xi \in \{0, 1\}^V$. Assume that H has a unique stationary distribution μ and that $\tau_{\text{cons}} < \infty$, then*

$$\mathbb{P}(\text{opinion 1 wins} | \xi_0 = \xi) = \sum_{v \in V} \mu(v)\xi(v).$$

The above theorem solves the winning probability problem under reasonable conditions, so we focus on the consensus time problem.

Consider the two party model and let S_t be the set of vertices with opinion 1 at the beginning of round t . Denote $\mu(S_t) = \sum_{v \in S_t} \mu(v)$, where μ is the stationary distribution of H , and $Z_t = \mu(S_{t+1}) - \mu(S_t)$. Let $\mu^* = \min_{v \in V} \mu(v)$. Define the quantity Ψ as

$$\Psi = \mu^* \min_{\substack{S \subseteq V \\ S \neq \emptyset, V}} \frac{\mathbb{E}(|Z_0| | S_0 = S)}{\min\{\mu(S), 1 - \mu(S)\}}, \quad (2)$$

where the minimum is over all $S \subseteq V$ except $S = \emptyset$ and $S = V$. Using the above definitions we prove the following theorem.

► **Theorem 2.** *Let $(\xi_t)_{t \geq 0} \sim \mathcal{LVM}(l, \xi)$ with $\xi \in \{0, 1\}^V$ be a voting model with $\Psi > 0$ then*

$$\mathbb{E}(\tau_{\text{cons}}) \leq 64/\Psi.$$

The structure of the paper is as follows. In Section 2 we introduce the model and give some examples to gain some intuition and demonstrate the flexibility of the model. In Section 3, we introduce the necessary notation to prove Theorem 1. In Section 4 we prove Theorem 2.

Notation. $G = (V, E)$ stands for a simple graph. We assume $|V| = n$. For $v \in V$ we denote by $N(v)$ the neighbourhood of v and define $d(v) = |N(v)|$. Moreover, given $X \subseteq V$, we define $d(X)$ as the sum of the degrees of the vertices in X . We use the notation $v \sim w$ to say that v and w are adjacent vertices. Q stands for the set of possible opinions, in general

$Q = \{0, 1\}$ or $Q = \{1, \dots, n\}$. We denote by \mathcal{M} the set of $n \times n$ matrices with exactly one 1 in each row and 0 in the other positions. Let Π be the set of probability distribution on \mathcal{M} , and $l \in \Pi$ be a given probability distribution over matrices in \mathcal{M} . M^\top denotes the transpose of the matrix M .

2 The linear voting model.

Recall the definition of a linear voting model. Given $l \in \Pi$ and $\xi \in Q^V$ we say $(\xi_t)_{t \geq 0} \sim \mathcal{LVM}(l, \xi)$ if $\xi_0 = \xi$ and $\xi_{t+1} = M_t \xi_t$, $t \geq 0$, where the M_t are i.i.d. samples from l . The following models are examples of linear voting.

- (a) **Synchronous pull model.** At each round each vertex samples a random neighbour and adopts the opinion of such neighbour.
- (b) **Asynchronous pull model.** At each round one vertex v is selected at random, then it samples a random neighbour and v adopts the opinion of this neighbour.
- (c) **Asynchronous push model.** At each round a vertex v is selected at random, then it samples a random neighbour and the neighbour adopts the opinion of v .
- (d) **Abusive push model.** At each round one vertex v is selected at random and the whole neighbourhood adopts the opinion of v .
- (e) **Pull-push model.** At each round one vertex v is selected at random, and two neighbours u_1, u_2 are selected randomly (with replacement). Then at the same time, u_1 adopts the opinion of v while v adopts the opinion of u_2 .

► **Remark.** To be precise, the changes in the opinions happen at the end of a round t , prior to round $t + 1$. In particular if v adopts the opinion of w at round t , it means that at round $t + 1$, vertex v has the opinion of w at round t .

► **Lemma 3.** *The five models defined above are linear voting models.*

Proof Sketch. We just prove it for the first and second model. For the other models the proof is similar. Let ξ_t be the configuration of opinions at round t . In the synchronous pull voting at each round each vertex v samples a random neighbour $w(v)$ and then v adopts the opinion of $w(v)$. Call ξ_{t+1} the new configuration of opinion. We check that $\xi_{t+1} = M \xi_t$ where the (random) matrix M is given by $M(v, w(v)) = 1$ for all $v \in V$, and 0 for the others entries. It is straightforward to check that $M \xi_t(v) = M(v, w(v)) \xi(w(v)) = \xi_t(w(v)) = \xi_{t+1}(v)$ and also that M has only one 1 in each row and thus $M \in \mathcal{M}$.

For the asynchronous pull model, observe that only one vertex v is selected and then v adopts the opinion of a random vertex $w(v)$, while all other vertices keep their opinions unchanged. Call ξ_{t+1} the new configuration. Define M as $M(v, w(v)) = 1$, $M(u, u) = 1$ for all $u \neq v$ and 0 for all other entries (M is like the identity matrix, except in the column of v). It is not hard to check that the random matrix M mimics the asynchronous pull model, i.e. $\xi_{t+1} = M \xi_t$, and that $M \in \mathcal{M}$. ◀

Remember we define the mean matrix of $l \in \Pi$ as $H = H(l) = \sum_{M \in \mathcal{M}} l(M) M$. Since most of the models are described by rules rather than by giving the explicit distribution l , it might be hard to compute $H(l)$. Nevertheless, the following lemma helps us to compute H without exhibiting l explicitly.

► **Lemma 4.** *For any distribution l over matrices in \mathcal{M} , the matrix $H = H(l)$ is the transition matrix of a Markov chain. Moreover, for every $t \geq 0$, and $v, w \in V$,*

$$H(v, w) = \mathbb{P}(v \text{ adopts the opinion of } w \text{ at round } t). \quad (3)$$

Proof. Note that, as each element of M is a transition matrix (the rows sum up to 1), H is the convex combination of transition matrices and thus is a transition matrix. To prove the second part note that by conditioning on the configuration ξ_t we have that

$$\mathbb{E}(\xi_{t+1}|\xi_t) = \sum_{M \in \mathcal{M}} l(M)(M\xi_t) = \left(\sum_{M \in \mathcal{M}} l(M)M \right) \xi_t = H\xi_t. \quad (4)$$

Choose ξ_t such that the opinion of w is 1 and all other opinions are 0. Then the event $\{v$ adopts the opinion of w at round $t\}$ is equal to $\{\xi_{t+1}(v) = 1\}$. Thus, from equation (4)

$$\mathbb{P}(\xi_{t+1}(v) = 1|\xi_t) = \mathbb{E}(\xi_{t+1}(v)|\xi_t) = (H\xi_t)(v) = \sum_{w \in V} H(v, w)\xi_t(w) = H(v, w). \quad \blacktriangleleft$$

Let P be the transition matrix of a simple random walk on G , A the adjacency matrix of G and let I denote the identity matrix. Let $L = D - A$ be the combinatorial Laplacian where D is the diagonal matrix containing the degree sequence of G . Moreover, let F be the diagonal matrix defined by $F(v, v) = \sum_{w: w \sim v} 1/d(w)$. The next theorem gives the matrix H for the linear voting models used in our examples.

► **Theorem 5.** *The mean matrix of the synchronous pull, asynchronous pull, push, abusive push, pull-push models are, respectively, $H_a = P$ [6] and*

$$H_b = \frac{n-1}{n}I + \frac{1}{n}P, \quad H_c = I + \frac{1}{n}P^\top - \frac{1}{n}F, \quad H_d = I - \frac{1}{n}L, \quad H_e = \frac{1}{n}(P + P^\top) + \frac{n-1}{n}I - \frac{1}{n}F.$$

Proof Sketch. We compute H_a . Observe that $H_a(v, w)$ is the probability that v adopts the opinion of w . That happens only if the random neighbour selected for v is w . Then $H_a(v, w) = \frac{1}{d(v)}\mathbb{1}_{v \sim w}$, concluding that $H_a = P$. For H_b , remember that in asynchronous pull we select a random vertex v and then v adopts the opinion of a random neighbour $w(v)$. Observe that for a vertex u we have $H_b(u, u)$ is the probability that u adopts the opinion of u , i.e. the probability that u does not change the opinion. That happen with probability $(n-1)/n$. On the other hand if $w \sim v$ then we have $H_b(v, w) = 1/nd(v)$ because v has to be initially selected and then v has to select w from its neighbourhood. We conclude that $H_b = ((n-1)/n)I + (1/n)P$. The other cases are similar. ◀

3 Winning probability

The most basic question in any voting model is, ‘who wins?’. In order to answer this question we use some martingale arguments. Assume the two-party model, $Q = \{0, 1\}$. Since the mean matrix H of a linear voting model is a transition matrix, then all its eigenvalues lie in $[-1, 1]$. We order the eigenvalues in decreasing order, i.e. $1 = \lambda_1 \geq \lambda_2 \dots \geq \lambda_n$. Let λ be an eigenvalue of H^\top (H and H^\top have the same eigenvalues) with corresponding eigenvector f , that is $H^\top f = \lambda f$. Given $f, g \in \mathbb{R}^V$, we denote $\langle f, g \rangle = \sum_{v \in V} f(v)g(v)$ the standard inner product. Observe that $Q \subseteq \mathbb{R}$, so if $\xi \in Q^V$ and $f \in \mathbb{R}^V$, the inner product $\langle f, \xi \rangle = \sum_{v \in V} f(v)\xi(v)$ is well-defined.

► **Lemma 6.** *The process $(\langle f, \xi_t \rangle / \lambda^t)_{t \geq 0}$ is a martingale with respect to $(\xi_t)_{t \geq 0}$*

Proof. Since $\langle f, \xi_t \rangle$ is bounded, we can check that $\mathbb{E}(\langle f, \xi_{t+1} \rangle | \xi_t) = \lambda \langle f, \xi_t \rangle$ and divide both sides by λ^{t+1} . By linearity of (conditional) expectation and equation (4) we have

$$\mathbb{E}(\langle f, \xi_{t+1} \rangle | \xi_t) = \langle f, H\xi_t \rangle = \langle H^\top f, \xi_t \rangle = \lambda \langle f, \xi_t \rangle. \quad \blacktriangleleft$$

Since H is a transition matrix, if the associated Markov chain is recurrent and aperiodic then the Markov chain has a unique stationary distribution. Denote this stationary distribution by μ . It is a classic result of the theory of finite Markov chains that μ , interpreted as a vector, is the unique eigenvector of H^\top with eigenvalue 1. We assume the vector μ is scaled so that $\sum_{v \in V} \mu(v) = 1$. Since, among all eigenvectors, μ is the most important we denote by $m_t = \langle \mu, \xi_t \rangle$ the martingale associated with the eigenvalue 1, and we call this martingale the **voting martingale**.

Proof of Theorem 1. Denote by $\mathbf{1}$ and $\mathbf{0}$ the vector where all components are 1 and 0 respectively. Since $(\xi_t)_{t \geq 0}$ always reaches consensus, it converges to $\mathbf{1}$ or $\mathbf{0}$ and thus $(m_t)_{t \geq 0}$ converges to 1 or 0. Moreover, $0 \leq m_t = \sum_{v \in V} \mu(v) \xi_t(v) \leq 1$ for every $\xi_t \in \{0, 1\}^V$, so $(m_t)_{t \geq 0}$ is a bounded martingale. These two properties of $(m_t)_{t \geq 0}$, together with the fact that τ_{cons} is a stopping time, allows us to apply the optional stopping theorem [5] to conclude $\mathbb{E}(m_0) = \mathbb{E}(m_{\tau_{\text{cons}}})$. Since $\xi_0 = \xi$ is a deterministic quantity then $\mathbb{E}(m_0) = m_0$. Moreover

$$\mathbb{E}(m_{\tau_{\text{cons}}}) = \langle \mu, \mathbf{1} \rangle \mathbb{P}(\xi_{\tau_{\text{cons}}} = \mathbf{1} | \xi_0 = \xi) + \langle \mu, \mathbf{0} \rangle \mathbb{P}(\xi_{\tau_{\text{cons}}} = \mathbf{0} | \xi_0 = \xi) = \mathbb{P}(\xi_{\tau_{\text{cons}}} = \mathbf{1} | \xi_0 = \xi).$$

Hence $\mathbb{P}(\xi_{\tau_{\text{cons}}} = \mathbf{1} | \xi_0 = \xi) = m_0 = \langle \mu, \xi \rangle$, therefore

$$\mathbb{P}(\text{opinion 1 wins} | \xi_0 = \xi) = \sum_{v \in V} \mu(v) \xi(v). \quad \blacktriangleleft$$

► **Corollary 7.** Assume the same conditions of Theorem 1 but consider $Q = \{1, \dots, n\}$. Suppose that $\xi \in Q^V$. Then the probability that $k \in Q$ wins is

$$\mathbb{P}(\xi_{\tau_{\text{cons}}} = k \mathbf{1} | \xi_0 = \xi) = \sum_{v \in V: \xi(v)=k} \mu(v).$$

Proof. Replace opinion k by opinion 1 and all other opinions by opinion 0, and then use Theorem 1 ◀

► **Theorem 8.** Let G be a connected graph. Let A be the set of vertices whose initial opinion is 1. Then, given that the models reach consensus, the probability p that opinion 1 wins is

- (a) *synchronous pull model*: $p_a = d(A)/d(V)$
- (b) *asynchronous pull model*: $p_b = d(A)/d(V)$
- (c) *push model*: $p_c = (\sum_{v \in A} d(v)^{-1}) / (\sum_{v \in V} d(v)^{-1})$
- (d) *abusive pushing model*: $p_d = |A|/n$
- (e) *pull-push model*: $p_e = |A|/n$.

Proof. We apply Theorem 1. For that we need to find the stationary distribution of the above models. The stationary distribution of P is $\pi(v) = d(v)/d(V)$, that gives us the result for synchronous pull. Observe that $(n-1)/n I + (1/n)P$ is a lazy version of the random walk of P , then it has the same stationary distribution, giving us the result for the asynchronous pull model. For the push model we just guess the stationary distribution and check it. Let $C = 1/(\sum_{v \in V} d(v)^{-1})$ and let $\pi'(v) = C/d(v)$, then as $F = F^\top$

$$\begin{aligned} (H_c^\top \pi)(v) &= ((I + \frac{1}{n}P - \frac{1}{n}F)\pi')(v) = \pi'(v) + \frac{1}{n} \sum_{w \in V} P(v, w) \pi'(w) - \frac{1}{n} F(v, v) \pi'(v) \\ &= \pi'(v) + \frac{1}{n} \sum_{w: w \sim v} \frac{1}{d(v)} \frac{C}{d(w)} - \frac{C}{d(v)n} \sum_{w: w \sim v} \frac{1}{d(w)} = \pi'(v) \end{aligned}$$

proving that π' is the stationary distribution of the mean matrix of the push model. For the abusive pushing model observe that as $I - (1/n)L$ is a symmetric matrix, its stationary distribution is uniform. H_e is also symmetric, giving the result for the push-pull model. ◀

4 Consensus Time

In this section we assume the two-party model with opinions $Q = \{0, 1\}$. Let $(\xi_t)_{t \geq 0} \sim \mathcal{LVM}(l, \xi)$ be a linear voting model. Assume $H = H(l)$ has a unique stationary distribution and let $(m_t)_{t \geq 0}$ be the voting martingale defined in Section 3. We use the following convenient notation. Let S_t be the set of vertices with opinion 1 at the beginning of round t , let $\mu(S_t) = m_t = \langle \mu, \xi_t \rangle$, and let $Z_t = \mu(S_{t+1}) - \mu(S_t)$. Note that, since $\mu(S_t)$ is a martingale, $\mathbb{E}(Z_t | S_t = S) = 0$. The random variable Z_t gives us information about the change in the measure of the set S_t . A larger value of $|Z_t|$ implies voting finishes faster.

Let $\eta(S) = \min\{\mu(S), \mu(S^c)\}$, where $\mu(S^c) = 1 - \mu(S)$. Denote by η_t the process $\eta(S_t)$. Since $\mu(S_t) \in [0, 1]$ we have $\eta_t \in [0, 1/2]$. Recall that $\mu(V) = 1$ and $\mu(\emptyset) = 0$. Note that $\eta_{t+1} = \min\{\mu(S_t) + Z_t, \mu(S_t^c) - Z_t\}$. Noting that if $\eta_t = \mu(S_t)$, i.e. $\mu(S_t) \leq \mu(S_t^c)$, then

$$\eta_{t+1} \leq \mu(S_{t+1}) = \mu(S_t) + Z_t = \eta_t + Z_t,$$

and if $\eta_t = \mu(S_t^c)$, the same applies by noticing that $\mu(S_{t+1}^c) - \mu(S_t^c) = -Z_t$, i.e.

$$\eta_{t+1} \leq \mu(S_{t+1}^c) = \mu(S_t^c) - Z_t = \eta_t - Z_t,$$

then in both cases we get

$$\eta_{t+1} \leq \eta_t + \rho_t Z_t, \tag{5}$$

where $\rho_t = \rho(S_t) = 2\mathbb{1}_{\{\mu(S_t) \leq \mu(S_t^c)\}} - 1$. Observe $\rho_t \in \{-1, +1\}$. We are going to study the process $\sqrt{\eta_t}$, in particular, $\mathbb{E}(\sqrt{\eta_t})$. Define $\Upsilon(S)$ by

$$\Upsilon(S) = \mathbb{E}(Z_t^2 \mathbb{1}_{\{\rho_t Z_t < 0\}} | S_t = S) \tag{6}$$

and define $\Upsilon = \min \frac{\Upsilon(S)}{\eta(S)}$, where the minimum is over all $S \subseteq V$ but $S \neq \emptyset$ and $S \neq V$. With these ingredients we are ready to prove a technical lemma, which is fundamental for the proof of Theorem 2.

► **Lemma 9.** *Let $(\xi_t)_{t \geq 0} \sim \mathcal{LVM}(l, \xi)$ with $\xi \in \{0, 1\}^V$ be a voting model with $\Upsilon > 0$ then*

$$\mathbb{E}(\tau_{\text{cons}}) \leq 32/\Upsilon.$$

Proof. We borrow part of the argument from [1]. Let $S \subseteq V$ but $S \neq \emptyset$ and $S \neq V$. By conditioning on $S_t = S$, from equation (5) we have $\eta_{t+1} \leq \eta_t + \rho_t Z_t = \eta(S) + \rho_t Z_t$ (we replace η_t by $\eta(S)$ as $S_t = S$ is fixed). Then, by taking expectations

$$\begin{aligned} \mathbb{E}(\sqrt{\eta_{t+1}} | S_t = S) &\leq \sqrt{\eta(S)} \mathbb{E}\left(\sqrt{1 + \frac{\rho_t Z_t}{\eta_t}} \middle| S_t = S\right) \\ &= \sqrt{\eta(S)} \mathbb{E}\left(\left(\sqrt{1 + \frac{\rho_t Z_t}{\eta_t}}\right) \mathbb{1}_{\{\rho_t Z_t \geq 0\}} \middle| S_t = S\right) \end{aligned} \tag{7}$$

$$+ \sqrt{\eta(S)} \mathbb{E}\left(\left(\sqrt{1 + \frac{\rho_t Z_t}{\eta_t}}\right) \mathbb{1}_{\{\rho_t Z_t < 0\}} \middle| S_t = S\right). \tag{8}$$

Let $x = \rho_t Z_t / \eta_t$. It can be checked that $x \geq -1$. Indeed, from equation (5) we have $\rho_t Z_t \geq \eta_{t+1} - \eta_t \geq -\eta_t$, concluding $x \geq -1$.

For $x \geq -1$ the following partial Taylor expansions are valid,

$$\sqrt{1+x} \leq 1 + \frac{x}{2}, \quad (9)$$

$$\sqrt{1+x} \leq 1 + \frac{x}{2} - \frac{x^2}{8} + \frac{x^3}{16}. \quad (10)$$

To upper bound (7) use (9), and for (8) use (10). Recall that, since $\mu(S_t)$ is a martingale, then $\mathbb{E}(Z_t|S_t = S) = 0$. After some rearrangement, we obtain

$$\begin{aligned} \mathbb{E}(\sqrt{\eta_{t+1}}|S_t = S) &\leq \sqrt{\eta(S)} - \sqrt{\eta(S)} \mathbb{E} \left(\left(\frac{(\rho_t Z_t)^2}{8\eta_t^2} - \frac{(\rho_t Z_t)^3}{16\eta_t^3} \right) \mathbb{1}_{\{\rho_t Z_t < 0\}} \middle| S_t = S \right) \\ &\leq \sqrt{\eta(S)} - \sqrt{\eta(S)} \mathbb{E} \left(\frac{Z_t^2}{8\eta_t^2} \mathbb{1}_{\{\rho_t Z_t < 0\}} \middle| S_t = S \right) \\ &= \sqrt{\eta(S)} - \frac{\Upsilon(S)}{8\eta(S)^{3/2}} \leq \sqrt{\eta(S)} - \frac{\Upsilon}{8\eta(S)^{1/2}} \end{aligned} \quad (11)$$

In the second inequality we used the fact that we are working in $\{\rho_t Z_t < 0\}$ and after that we used the definition of $\Upsilon(S)$ from (6) and $\Upsilon = \min(\Upsilon(S)/\eta(S))$. Remember that $\eta(\emptyset) = \eta(V) = 0$, then

$$\begin{aligned} \mathbb{E}(\sqrt{\eta_{t+1}}) &= \sum_{S \subseteq V} \mathbb{E}(\sqrt{\eta_{t+1}}|S_t = S) \mathbb{P}(S_t = S) = \sum_{S: S \neq \emptyset, V} \mathbb{E}(\sqrt{\eta_{t+1}}|S_t = S) \mathbb{P}(S_t = S) \\ &\leq \sum_{S: S \neq \emptyset, V} \left(\sqrt{\eta(S)} - \frac{\Upsilon}{8\eta(S)^{1/2}} \right) \mathbb{P}(S_t = S) \end{aligned} \quad (12)$$

$$\begin{aligned} &= \mathbb{E}(\sqrt{\eta_t}) - \sum_{S: S \neq \emptyset, V} \left(\frac{\Upsilon}{8\eta(S)^{1/2}} \right) \mathbb{P}(S_t = S | \tau_{\text{cons}} > t) \mathbb{P}(\tau_{\text{cons}} > t) \\ &= \mathbb{E}(\sqrt{\eta_t}) - \frac{\Upsilon}{8} \mathbb{E} \left(\frac{1}{\sqrt{\eta_t}} \middle| \tau_{\text{cons}} > t \right) \mathbb{P}(\tau_{\text{cons}} > t), \end{aligned} \quad (13)$$

where (12) follows using equation (11). As $1/x$ is convex for $x > 0$, apply Jensen's inequality to the random variable $x = \sqrt{\eta_t}$, to obtain

$$\mathbb{E} \left(\frac{1}{\sqrt{\eta_t}} \middle| \tau_{\text{cons}} > t \right) \geq \frac{1}{\mathbb{E}(\sqrt{\eta_t} | \tau_{\text{cons}} > t)} = \frac{\mathbb{P}(\tau_{\text{cons}} > t)}{\mathbb{E}(\sqrt{\eta_t})}. \quad (14)$$

The last equality holds because the event $\{\tau_{\text{cons}} \leq t\}$ implies that the vertices reached consensus, then $S_t = \emptyset$ or $S_t = V$, hence $\eta_t = 0$, and then

$$\begin{aligned} \mathbb{E}(\sqrt{\eta_t}) &= \mathbb{E}(\sqrt{\eta_t} | \tau_{\text{cons}} > t) \mathbb{P}(\tau_{\text{cons}} > t) + \mathbb{E}(\sqrt{\eta_t} | \tau_{\text{cons}} \leq t) \mathbb{P}(\tau_{\text{cons}} \leq t) \\ &= \mathbb{E}(\sqrt{\eta_t} | \tau_{\text{cons}} > t) \mathbb{P}(\tau_{\text{cons}} > t). \end{aligned}$$

By substituting (14) into (13) we obtain

$$\mathbb{E}(\sqrt{\eta_{t+1}}) \leq \mathbb{E}(\sqrt{\eta_t}) - \frac{\Upsilon}{8} \frac{\mathbb{P}(\tau_{\text{cons}} > t)^2}{\mathbb{E}(\sqrt{\eta_t})},$$

then as $\eta_t \in [0, 1/2]$

$$\frac{\Upsilon}{8} \mathbb{P}(\tau_{\text{cons}} > t)^2 \leq \mathbb{E}(\sqrt{\eta_t}) (\mathbb{E}(\sqrt{\eta_t}) - \mathbb{E}(\sqrt{\eta_{t+1}})) \leq \frac{\mathbb{E}(\sqrt{\eta_t}) - \mathbb{E}(\sqrt{\eta_{t+1}})}{\sqrt{2}}.$$

Summing from $t = 0$ up to a time $T - 1$ we have

$$\Upsilon \sum_{t=0}^{T-1} \mathbb{P}(\tau_{\text{cons}} > t)^2 \leq 8 \frac{\mathbb{E}(\sqrt{\eta_0}) - \mathbb{E}(\sqrt{\eta_T})}{\sqrt{2}} \leq 4. \tag{15}$$

Let T be the minimum time t such that $\mathbb{P}(\tau_{\text{cons}} > t) < 1/2$, then for every $t < T$ we have $\mathbb{P}(\tau_{\text{cons}} > t) \geq 1/2$. Therefore, from equation (15), it holds that

$$T \leq 16/\Upsilon.$$

Note that our bound for T is independent of the initial position, so we assume the worst case. We compute $\mathbb{E}(\tau_{\text{cons}})$ by looking at the process every T steps. If at round T the process finished then $\tau_{\text{cons}} \leq T$, otherwise, we restart the process and look again after T steps until we reach consensus. As the probability the process does not finish in T steps is at most $1/2$, we conclude that

$$\mathbb{E}(\tau_{\text{cons}}) \leq \sum_{k=1}^{\infty} kT \left(\frac{1}{2}\right)^k \leq 2T \leq \frac{32}{\Upsilon}. \quad \blacktriangleleft$$

We need the following simple lemma.

► **Lemma 10.** *Let X be an integrable random variable with mean 0 then*

$$\mathbb{E}(|X| \mathbf{1}_{\{X < 0\}}) = \mathbb{E}(|X|)/2.$$

Proof. Let $X^+ = X \mathbf{1}_{\{X > 0\}}$ and $X^- = |X| \mathbf{1}_{\{X < 0\}}$. Clearly $X = X^+ - X^-$ and $|X| = X^+ + X^-$. Then we have the system of equations

$$\begin{aligned} \mathbb{E}(X^+) - \mathbb{E}(X^-) &= \mathbb{E}(X) = 0, \\ \mathbb{E}(X^+) + \mathbb{E}(X^-) &= \mathbb{E}(|X|). \end{aligned}$$

Then $\mathbb{E}(X^-) = \mathbb{E}(X^+) = \mathbb{E}(|X|)/2$. ◀

We proceed with the proof of Theorem 2

Proof of Theorem 2. From Lemma 9 we have

$$\mathbb{E}(\tau_{\text{cons}}) \leq 32/\Upsilon,$$

where $\Upsilon = \min \frac{\Upsilon(S)}{\eta(S)}$ and the minimum is over all $S \subseteq V$ other than $S = \emptyset$ and $S = V$. Observe that if $|Z_t| > 0$, it means that at least one vertex changes its opinion, thus $|Z_t| \geq \mu^* = \min_{v \in V} \mu(v)$. From there

$$\begin{aligned} \Upsilon(S) &= \mathbb{E}(Z_t^2 \mathbf{1}_{\{\rho_t Z_t < 0\}} | S_t = S) = \mathbb{E}((\rho_t Z_t \mathbf{1}_{\{\rho_t Z_t < 0\}})^2 | S_t = S) \\ &\geq \mu^* \mathbb{E}(|\rho_t Z_t \mathbf{1}_{\{\rho_t Z_t < 0\}}| | S_t = S) \end{aligned} \tag{16}$$

Note that $\mathbb{E}(\rho_t Z_t | S_t = S) = \rho(S) \mathbb{E}(Z_t | S_t) = 0$ because $\mu(S_t)$ is a martingale. Using Lemma 10 in equation (16), gives $\Upsilon(S) \geq \mu^* \mathbb{E}(|Z_t| | S_t)/2$. Hence $\Upsilon \geq \mu^* \min \frac{1}{2} \frac{\mathbb{E}(|Z_t| | S_t)}{\eta(S)}$. Recalling the definition of Ψ in equation 2, we conclude $\Upsilon \geq \Psi/2$ and therefore

$$\mathbb{E}(\tau_{\text{cons}}) \leq \frac{64}{\Psi}. \quad \blacktriangleleft$$

We apply the above theorems to our examples. We use the following notation. Given $S \subseteq V$, denote by $E(S : S^c)$ the number of edges going from S to S^c . Denote by $d_S(v)$ the number of vertices of S adjacent to v . Observe that $E(S : S^c) = \sum_{v \in S} d_{S^c}(v) = \sum_{v \in S^c} d_S(v)$. We denote the graph conductance by $\Phi(G) = \min_{S \subseteq V} \frac{E(S : S^c)}{\min\{d(S), d(S^c)\}}$ where $0/0 = \infty$.

144:10 The Linear Voting Model

► **Example 11.** Consider the asynchronous pulling model on a graph G .

$$\mathbb{E}(|Z_t| | S_t = S) = \sum_{v \in S} \frac{d(v)}{d(V)} \frac{1}{n} \frac{d_{S^c}(v)}{d(v)} + \sum_{v \in S^c} \frac{d(v)}{d(V)} \frac{1}{n} \frac{d_S(v)}{d(v)}.$$

Why? With probability $1/n$ we select vertex v and this vertex selects a random neighbour w with probability $1/d(v)$, and adopts its opinion. The stationary distribution of v is $\mu(v) = d(v)/d(V)$. If w has the same opinion as v , then $Z_t = 0$, but if w has the opposite opinion then $|Z_t| = d(v)/d(V)$. Then

$$\mathbb{E}(|Z_t| | S_t = S) = \frac{1}{nd(V)} \left(\sum_{v \in S} d_{S^c}(v) + \sum_{v \in S^c} d_S(v) \right) = \frac{2E(S : S^c)}{nd(V)} \quad (17)$$

therefore from (2)

$$\Psi = \frac{d_{\min}}{d(V)} \frac{2}{n} \min_S \frac{E(S : S^c)}{\min\{d(S), d(S^c)\}} \quad (18)$$

Hence we conclude that $\mathbb{E}(\tau_{con}) = \mathcal{O}(nd(V)/d_{\min}\Phi)$. This gives a consensus time of $\mathcal{O}(n^2)$ for expanders, which is optimal up to a constant. For the cycle, $\mathcal{O}(n^3)$ optimal as well.

► **Example 12.** Consider the push model on a graph G . Let $C = (\sum_{v \in V} d(v)^{-1})^{-1}$.

$$\mathbb{E}(|Z_t| | S_t = S) = \sum_{v \in S} \frac{C}{d(v)} \sum_{w: w \sim v, w \in S^c} \frac{1}{nd(w)} + \sum_{v \in S^c} \frac{C}{d(v)} \sum_{w: w \sim v, w \in S} \frac{1}{nd(w)}.$$

The above equation holds because to change the opinion of a vertex $v \in S$, the push model needs to select a vertex $w \in S^c$ adjacent to v and then w needs to push its opinion on v . That happens with probability $1/(nd(w))$. In such case, the change in $|Z_t|$ is $\mu(v) = C/d(v)$. The same applies if $v \in S^c$. Then

$$\mathbb{E}(|Z_t| | S_t = S) = \frac{2C}{n} \sum_{v \in S} \sum_{w \in S^c} \frac{\mathbb{1}_{v \sim w}}{d(v)d(w)}. \quad (19)$$

By using the notation $J(S) = \sum_{v \in S} d(v)^{-1}$ and that the stationary distribution is $\mu(v) = C/d(v)$ we have

$$\Psi = \frac{2C}{nd_{\max}} \min_S \frac{\sum_{v \in S} \sum_{w \in S^c} \frac{\mathbb{1}_{v \sim w}}{d(v)d(w)}}{\min\{J(S), J(S^c)\}}.$$

The parameter Ψ does not seem related to the classical graph parameters.

► **Example 13.** We continue with the abusive push model on a graph G .

$$\mathbb{E}(|Z_t| | S_t = S) = \sum_{v \in S} \frac{1}{n} \frac{d_{S^c}(v)}{n} + \sum_{v \in S^c} \frac{1}{n} \frac{d_S(v)}{n}.$$

The above equation holds because with probability $1/n$ we sample a vertex v . Then v pushes its opinion on all its neighbours. Since the stationary distribution for this model is $\mu(v) = 1/n$, then the change in $|Z_t|$ is $d_{S^c}(v)/n$ if $v \in S$ and $d_S(v)/n$ if $v \in S^c$. Then $\mathbb{E}(|Z_t| | S_t) = \frac{2}{n^2} E(S : S^c)$. Then it holds that

$$\Psi = \frac{2}{n^2} \min_S \frac{E(S : S^c)}{\min\{|S|, |S^c|\}} \quad (20)$$

The parameter $\min_S \frac{E(S:S^c)}{\min\{|S|, |S^c|\}}$ is very similar to the graph conductance, indeed, for d -regular graphs $\min_S \frac{E(S:S^c)}{\min\{|S|, |S^c|\}} = d\Phi(G)$. In such case we have that

$$\mathbb{E}(\tau_{\text{cons}}) = \mathcal{O}\left(\frac{n^2}{d\Phi}\right).$$

That gives us a $\mathcal{O}(n^2/d)$ time for regular expanders, which is optimal when the degree is constant. For the complete graph it gives us $\mathcal{O}(n)$, which is far from optimal, since the abusive push model finishes in just one round on the complete graph. For a cycle it gives us a $\mathcal{O}(n^3)$ time which is optimal.

► **Example 14.** Our final example is for the pull-push model. In this model the stationary distribution is uniform. Then the only way to produce a positive change in $|Z_t|$ is that when the random vertex v is chosen to pull and push, it selects one neighbour in S and the other in S^c . In that case, the change in $|Z_t|$ will be of $1/n$, then

$$\mathbb{E}(|Z_t| | S_t = S) = \sum_{v \in V} \frac{1}{n^2} \frac{d_{S^c}(v)d_S(v)}{d(v)^2}.$$

Then

$$\Psi = \frac{1}{n^2} \min_S \sum_{v \in V} \frac{d_{S^c}(v)d_S(v)}{d(v)^2} / \min\{|S|, |S^c|\}.$$

Once again Ψ does not seem related to the classical graph parameters.

5 Discussion

In this paper we introduced and studied the linear voting model. The model can be seen as a generalisation of many models of voting. Despite its generality, the process is tractable and we can compute the probability that a given opinion wins. Moreover, by using a suitable potential function we were able to provide a bound for the expected consensus time. Furthermore, applying this bound in specific cases led to classical graph parameters, such as conductance, as well to other less familiar, or even new, parameters.

Future work includes the study of particular models on interesting graph families, like expanders, transitive graphs or random graphs, as well as the development of new techniques to analyse the model.

References

- 1 Petra Berenbrink, George Giakkoupis, Anne-Marie Kermarrec, and Frederik Mallmann-Trenn. Bounds on the voter model in dynamic networks. *arXiv preprint arXiv:1603.01895*, 2016.
- 2 Colin Cooper, Robert Elsasser, Hirotaka Ono, and Tomasz Radzik. Coalescing random walks and voting on connected graphs. *SIAM Journal on Discrete Mathematics*, 27(4):1748–1758, 2013.
- 3 Josep Díaz, Leslie Ann Goldberg, George B Mertzios, David Richerby, Maria Serna, and Paul G Spirakis. Approximating fixation probabilities in the generalized moran process. *Algorithmica*, 69(1):78–91, 2014.
- 4 Peter Donnelly and Dominic Welsh. Finite particle systems and infection models. *Mathematical Proceedings of the Cambridge Philosophical Society*, 94(01):167–182, 1983.
- 5 Rick Durrett. *Probability: theory and examples*. Cambridge university press, 2010.

144:12 The Linear Voting Model

- 6 Yehuda Hassin and David Peleg. Distributed probabilistic polling and applications to proportionate agreement. *Information and Computation*, 171(2):248–268, 2001.
- 7 David Asher Levin, Yuval Peres, and Elizabeth Lee Wilmer. *Markov chains and mixing times*. AMS Bookstore, 2009.
- 8 Toshio Nakata, Hiroshi Imahayashi, and Masafumi Yamashita. Probabilistic local majority voting for the agreement problem on finite graphs. In *Computing and Combinatorics*, pages 330–338. Springer, 1999.
- 9 Roberto Oliveira. On the coalescence time of reversible random walks. *Transactions of the American Mathematical Society*, 364(4):2109–2128, 2012.
- 10 Roberto Oliveira. Mean field conditions for coalescing random walks. *The Annals of Probability*, 41(5):3420–3461, 2013.