The Linear Voting Model: Consensus and Duality.*

Colin Cooper Nicolás Rivera

August 20, 2016

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.

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

Authors' address: Department of Informatics, King's College London, UK.

Email: {colin.cooper, nicolas.rivera}@kcl.ac.uk

1 Introduction.

Graphs are a simple model of the environment in which individual entities 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 [6]. 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 case of push voting). On the other hand, Hassin and Peleg [8] and Nakata et al.[10] 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 [8] and [10] 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 [11] 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 [12], 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 [8] 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. [3] 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. [2] used a more ad hoc approach and proved that the consensus time is $\mathcal{O}((d_{\text{ave}}/d_{\min})(n/\Phi))$ 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 instance of the Moran process. Díaz et al. [5] proved that the the consensus time is $\mathcal{O}(n^4q)$ 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 $Q \subseteq \{1, \ldots, n\}$ if we want to allow more choice of opinions.

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. It is a further assumption of the model that is $M_{ij} = 1$ then ij is an edge of G. 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 $\mathcal{M}$. The choice of l turns $\mathcal{M}$ into a probability space with measure l. Given an initial configuration ξ, we define the process $(\xi_t)_{t\geq 0}$, with t running over the non-negative integers, as

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

where M_t are i.i.d matrices sampled from l, and $M\xi$ is 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 $\xi' = M\xi$. 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 including the pull voting (asynchronous or synchronous) and the push voting model are given in Section 2.

We proceed to present our contribution. Theorem 1 of this paper gives the probability a particular opinion wins. This generalises the approach used in [8]. 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. Our approach is more similar to Levin et al. [9, Chapter 17] or Berenbrink [2].

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 2 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).

A configuration ξ is said to be in consensus if all the opinions in ξ are the same. Given $(\xi_t) \sim \mathcal{LVM}(l, \xi)$, we defined the consensus time τ_{cons} of (ξ_t) as

$$\tau_{\text{cons}} = \inf\{t \geq 0 : \xi_t \text{ is in consensus}\}.$$

Observe the consensus time is a stopping time with respect to $(\xi)_{t\geq 0}$ and that once the vertices reach consensus they never change their opinion again, thus the system reach a final state. 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*

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

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

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

where the minimum is over all $S \subseteq V$ except $S \neq \emptyset$ and $S \neq V$.

To make matters precise, let $\langle x, y \rangle = \sum x_i y_i$ denote the standard inner product. Given ξ and $\xi' = M\xi$, then $\mu(S) = \langle \mu, \xi \rangle$, and $\mu(S') = \langle \mu, \xi' \rangle = \langle \mu, M\xi \rangle$. For fixed S the quantity $Z(S) = \mu(S') - \mu(S)$ is a function of M. The expectation in (2) is with respect to the distribution l of M.

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*

$$\mathbb{E}(\tau_{\text{cons}}) \leq 64/\Psi. \tag{3}$$

Upper bounds on consensus time obtained from (3) are given in Section 4.1 for the models given in Section 2.

Finally, similarly to the voter model, the linear voter model has a dual process, which is also defined using the matrices M_t. Let $l \in \Pi$ be a distribution over matrices in $\mathcal{M}$. Consider a non-negative integer-valued vector f on $\mathbb{R}^V$. We define the discrete-time process $(f_t)_{t \geq 0}$ as

$$f_t = \begin{cases} f, & \text{if } t = 0, \\ M_{t-1}^\top f_{t-1}, & \text{if } t > 0, \end{cases} \tag{4}$$

where M_t are i.i.d. matrices sampled from l and $M^\top$ is the transpose of M. The process can be interpreted as follows. Vector f_t counts the number of particles at each of the vertices of the graph and $M_{t-1}^\top$ moves any particles on the vertices at time $t-1$. In particular if $M_{t-1}(v, u) = 1$ it means that all the particles on vertex v moves to vertex u. In this process once two or more particles come together they keep moving together and thus the process $(f_t)_{t \geq 0}$ is called the Coalescing process and we denote $(f_t)_{t \geq 0} \sim \mathcal{CP}(l, f)$. Observe that $\sum_{v \in V} f_t(v)$ is invariant over time. A standard choice of the initial vector f is the all 1 vector $\mathbf{1}$.

For a set of vertices $S \subseteq V$ define $f_t(S) = \sum_{v \in S} f_t(v)$. Consider a partition $\mathcal{S} = \{S_1, \ldots, S_m\}$ of the vertices V. We say the coalescing process is in agreement with respect to $\mathcal{S}$ at time t, if for some index $i \in \{1, \ldots, m\}$ we have $f_t(S_i) = n$, and consequently, $f(S_j) = 0$ for all $j \neq i$. More formally,

$$\{f_t \text{ is in agreement}\} \Leftrightarrow \{f_t(S_i) = n \text{ for some } i \in \{1, ..., m\}\} \tag{5}$$

Given a vector of opinion ξ we denote by $\mathcal{S}_\xi$ the natural partition given by the vertices with the same opinion, that is, if the set of opinions is Q, then $\mathcal{S}_\xi = \{S_a : a \in Q\}$ where $S_a = \{v : \xi(v) = a\}$. The next theorem states the relation between the linear voting process and the coalescing process.

Theorem 3. *Let V a set of vertices and let $l \in \Pi$. Suppose that $(\xi_t) \sim \mathcal{LVM}(l, \xi)$ and $(f_t) \sim \mathcal{CP}(l, \mathbf{1})$. Then, for every $t \geq 0$,*

$$\mathbb{P}(\xi_t \textit{ is in consensus}) = \mathbb{P}(f_t \textit{ is in agreement with respect to } \mathcal{S}_\xi).$$

A particular case of Theorem 3 is when all vertices has different opinion in ξ. In such case agreement is reached at time t if and only if all particles are together at such time. The first time all particles are together is called the coalescing time and it is denoted by τ_{coalsc}. As corollary, when all vertices has different opinion we have that τ_{cons} has the same distribution as τ_{coalsc}.

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 1. *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 1. *The five models defined above are linear voting models.*

Proof.

We just prove it for models a) and b), 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 2. *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 $v, w \in V$,*

$$H(v, w) = \mathbb{P}(v \text{ adopts the opinion of } w \text{ at round } t) \tag{6}$$

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

$$\mathbb{E}(\xi_{t+1}|\xi_t) = \sum_{M\in\mathcal{M}} l(M)(M\xi_t) = \left(\sum_{M\in\mathcal{M}} l(M)M\right)\xi_t = H\xi_t. \tag{7}$$

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 (7)

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

□

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 4. *The mean matrix of the synchronous pull, asynchronous pull, push, abusive push, pull-push models are, respectively, $H_a = P$ [8] 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. $1 = \lambda_1 \geq \lambda_2 \ldots, \geq \lambda_n$. 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 3. *For $\lambda \neq 0$, the process $\left(\langle f, \xi_t \rangle / \lambda^t\right)_{t \geq 0}$ 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 (7) we have

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

□

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**.

Recall that the consensus time, τ_{cons}, is the first time in which all the vertices have the same opinion. Observe τ_{cons} is a stopping time with respect to $(\xi_t)_{t \geq 0}$.

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 [7] 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

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

□

Corollary 1. *Assume the same conditions of Theorem 1 but consider $Q = \{1 \ldots, , n\}$. 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 5. *Let G be a connected graph. Let A be the set of vertices whose initial opinion is 1. Given that the models reach consensus on G, let μ be the stationary distribution of H and let p be the probability that opinion 1 wins is*

(a) Synchronous pull model: $\mu(v) = d(v)/d(V)$, $p_a = d(A)/d(V)$

(b) Asynchronous pull model: $\mu(v) = d(v)/d(V)$, $p_b = d(A)/d(V)$

(c) Push model: $\mu(v) = C/d(v)$, where $C = 1/(\sum_{v \in V} d(v)^{-1})$; $p_c = C(\sum_{v \in A} d(v)^{-1})$

(d) Abusive pushing model: $\mu(v) = 1/n$, $p_d = |A|/n$

(e) Pull-push model : $\mu(v) = 1/n$, $p_e = |A|/n$.

Proof. We apply Theorem 1. For that we need to find the stationary distribution of the above models. The stationary distribution of P is $\mu(v) = d(v)/d(V)$, that gives us the result for synchronous pull. Observe that $(n-1)/nI + (1/n)P$ 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 $\mu'(v) = C/d(v)$, then as $F = F^\top$

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

proving that μ' 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 same 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 μ. Let $m_t = \langle \mu, \xi_t \rangle$, so that $(m_t)_{t\geq 0}$ is 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, then $\mu(S_t) = m_t$. 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 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

$$\eta_{t+1} \leq \eta_t + \rho_t Z_t, \tag{8}$$

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 study the process $\sqrt{\eta_t}$, in particular, $\mathbb{E}(\sqrt{\eta_t})$. Define $\Upsilon(S)$ by

$$\Upsilon(S) = \mathbb{E}\left(Z_t^2 \mathbb{1}_{\{\rho_t Z_t < 0\}} \middle| S_t = S\right) \tag{9}$$

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 4. *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. Let $S \subseteq V$ but $S \neq \emptyset$ and $S \neq V$. By conditioning on $S_t = S$, from equation (8) 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). It can be checked that $\rho_t Z_t / \eta_t \geq -1$. Indeed, from equation (8) we have $\rho_t Z_t \geq \eta_{t+1} - \eta_t \geq -\eta_t$. By taking expectations

$$\begin{aligned}
\mathbb{E}(\sqrt{\eta_{t+1}} | S_t = S) &\leq \sqrt{\eta(S)}\mathbb{E}\left(\sqrt{1 + \frac{\rho_t Z_t}{\eta_t}} \,\middle|\, S_t = S\right) \\
&= \sqrt{\eta(S)}\mathbb{E}\left(\left(\sqrt{1 + \frac{\rho_t Z_t}{\eta_t}}\right) \mathbb{1}_{\{\rho_t Z_t \geq 0\}} \,\middle|\, S_t = S\right) && (10) \\
&\quad + \sqrt{\eta(S)}\mathbb{E}\left(\left(\sqrt{1 + \frac{\rho_t Z_t}{\eta_t}}\right) \mathbb{1}_{\{\rho_t Z_t < 0\}} \,\middle|\, S_t = S\right). && (11)
\end{aligned}$$

Let $x = \rho_t Z_t/\eta_t$. For $x \geq -1$ the following partial Taylor expansions are valid,

$$\sqrt{1+x} \leq 1 + \frac{x}{2}, \tag{12}$$

$$\sqrt{1+x} \leq 1 + \frac{x}{2} - \frac{x^2}{8} + \frac{x^3}{16}. \tag{13}$$

To upper bound (10) use (12), and for (11) use (13). Recall that, since $\mu(S_t)$ is a martingale, then $\mathbb{E}(Z_t|S_t = S) = 0$. After some rearrangement, we obtain

$$\begin{aligned}
\mathbb{E}(\sqrt{\eta_{t+1}}|S_t = S) &\leq \sqrt{\eta(S)} - \sqrt{\eta(S)}\mathbb{E}\left(\left(\frac{(\rho_t Z_t)^2}{8\eta_t^2} - \frac{(\rho_t Z_t)^3}{16\eta_t^3}\right)\mathbb{1}_{\{\rho_t Z_t<0\}}\middle|S_t = S\right) \\
&\leq \sqrt{\eta(S)} - \sqrt{\eta(S)}\mathbb{E}\left(\frac{Z_t^2}{8\eta_t^2}\mathbb{1}_{\{\rho_t Z_t<0\}}\middle|S_t = S\right) \\
&= \sqrt{\eta(S)} - \frac{\Upsilon(S)}{8\eta(S)^{3/2}} \leq \sqrt{\eta(S)} - \frac{\Upsilon}{8\eta(S)^{1/2}} \qquad (14)
\end{aligned}$$

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 (9) and $\Upsilon = \min(\Upsilon(S)/\eta(S))$. Remember that $\eta(\emptyset) = \eta(V) = 0$. If $S_t \neq \emptyset, V$ then the system is not in consensus and $\tau_{\text{cons}} > t$, thus

$$\begin{aligned}
\mathbb{E}(\sqrt{\eta_{t+1}}) &= \sum_{S\subseteq V}\mathbb{E}(\sqrt{\eta_{t+1}}|S_t = S)\mathbb{P}(S_t = S) = \sum_{S:S\neq\emptyset,V}\mathbb{E}(\sqrt{\eta_{t+1}}|S_t = S)\mathbb{P}(S_t = S) \\
&\leq \sum_{S:S\neq\emptyset,V}\left(\sqrt{\eta(S)} - \frac{\Upsilon}{8\eta(S)^{1/2}}\right)\mathbb{P}(S_t = S) \qquad (15) \\
&= \mathbb{E}(\sqrt{\eta_t}) - \sum_{S:S\neq\emptyset,V}\left(\frac{\Upsilon}{8\eta(S)^{1/2}}\right)\mathbb{P}(S_t = S|\tau_{\text{cons}} > t)\mathbb{P}(\tau_{\text{cons}} > t) \\
&= \mathbb{E}(\sqrt{\eta_t}) - \frac{\Upsilon}{8}\mathbb{E}\left(\frac{1}{\sqrt{\eta_t}}\middle|\tau_{\text{cons}} > t\right)\mathbb{P}(\tau_{\text{cons}} > t), \qquad (16)
\end{aligned}$$

where (15) follows using equation (14). As $1/x$ is convex for $x > 0$, apply Jensen's inequality to the random variable $x = \sqrt{\eta_t}$, to obtain

$$\mathbb{E}\left(\frac{1}{\sqrt{\eta_t}}\middle|\tau_{\text{cons}} > t\right) \geq \frac{1}{\mathbb{E}\left(\sqrt{\eta_t}|\tau_{\text{cons}} > t\right)} = \frac{\mathbb{P}(\tau_{\text{cons}} > t)}{\mathbb{E}\left(\sqrt{\eta_t}\right)}. \tag{17}$$

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 (17) into (16) we obtain

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

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

$$\Upsilon \sum_{t=0}^{T-1} \mathbb{P}(\tau_{\text{cons}} > t)^2 \leq 8 \frac{\mathbb{E}(\sqrt{\eta_0}) - \mathbb{E}(\sqrt{\eta_T})}{\sqrt{2}} \leq 4. \tag{18}$$

Let T be the defined as

$$T = \min\{t \geq 0 : \mathbb{P}(\tau_{\text{cons}} > t) < 1/2\}. \tag{19}$$

Therefore, from equation (18), 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

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

□

We need the following simple lemma.

Lemma 5. *Let X be an integrable random variable with mean 0 then*

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

Proof. Let $X^+ = X \mathbb{1}_{\{X>0\}}$ and $X^- = |X| \mathbb{1}_{\{X<0\}}$. 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 4 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 $\rho_t Z_t < 0$ then $|Z_t| > 0$, which means that at least one vertex changes its opinion, and $|Z_t| \geq \mu^* = \min_{v \in V} \mu(v)$. From there

$$\begin{aligned} \Upsilon(S) &= \mathbb{E}\left(Z_t^2 \mathbb{1}_{\{\rho_t Z_t < 0\}} \middle| S_t = S\right) = \mathbb{E}\left((\rho_t Z_t \mathbb{1}_{\{\rho_t Z_t < 0\}})^2 \middle| S_t = S\right) \\ &\geq \mu^* \mathbb{E}\left(|\rho_t Z_t \mathbb{1}_{\{\rho_t Z_t < 0\}}| \middle| S_t = S\right). \end{aligned} \tag{20}$$

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 5 in equation (20), 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

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

□

4.1 Consensus time examples

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 $0/0 = \infty$.

Example 1. ***Asynchronous pulling model.***

For this model

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

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

$$\mathbb{E}(|Z_t||S_t = S) \;=\; \frac{1}{nd(V)} \left(\sum_{v\in S} d_{S^c}(v) + \sum_{v\in S^c} d_S(v) \right) = \frac{2E(S:S^c)}{nd(V)} \tag{21}$$

therefore from (2)

$$\Psi = \frac{d_{\min}}{d(V)} \frac{2}{n} \min_S \frac{E(S:S^c)}{\min\{d(S),d(S^c)\}} \tag{22}$$

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 regular expanders, which is optimal up to a constant. For the cycle, $\mathcal{O}(n^3)$ optimal as well.

Example 2. ***Asynchronous push model.***

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

$$\mathbb{E}(|Z_t||S_t = S) \;=\; 2\frac{C}{n} \sum_{v\in S} \sum_{w\in S^c} \frac{\mathbb{1}_{v\sim w}}{d(v)d(w)}. \tag{23}$$

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

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

In general, the parameter Ψ does not seem related to the classical graph parameters. If the graph is regular of degree d, then

$$\Psi = \frac{2}{n^2}\Phi,$$

in which case $\mathbb{E}(\tau_{con}) = \mathcal{O}(n^2/\Phi)$, which agrees with the asynchronous pull model in Example 2.

Example 3. ***Abusive push model.***

We continue with the abusive pushing model on a graph G.

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

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

$$\Psi = \frac{2}{n^2} \min_S \frac{E(S : S^c)}{\min\{|S|, |S^c|\}} \tag{24}$$

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 4. ***Pull-push model.***

Our final example is the pull-push model. In this model the stationary distribution is uniform. 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 the push neighbour in S and the pull one in S^c, or vice versa. In either case, the change in $|Z_t|$ will be of $1/n$, then

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

Then

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

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

5 Coalescing process and duality

One of the most interesting properties of the linear voting model is the associated dual process. The dual process is important because we can study properties in the dual process and obtain an associated result in the original process.

Coalescing Process.

We recall the definition of the dual process from (4). Consider a finite set of states V. Let $(M_t)_{t\geq 0}$ be an i.i.d. collection of random matrices sampled with distribution $l \in \Pi$ and let $f \in \mathbb{R}^V$ be a non-negative integer vector. Define the process $(f_t)_{t\geq 0}$ as follows.

$$f_t = \begin{cases} f, & \text{if } t = 0 \\ M_{t-1}^\top f_{t-1}, & \text{if } t \geq 0. \end{cases} \tag{25}$$

For the above process we denote $(f_t) \sim \mathcal{CP}(l, f)$. We usually set $f = \mathbf{1}$ indicating that we start the process with one particle in each vertex.

Let us explain the dual process intuitively. For a given M and $v \in V$, let $w \in V$ be the unique element such that $M(v, w) = 1$ and define $g_M(v) = w$. This induces a function $g_M : V \to V$. It is also defines a process X on vertices, where $X(v) = w$ iff M is such that $g_M(v) = w$. Let χ_v be the characteristic vector for v. We have that

$$X(v) = w \iff g_M(v) = w \iff M(v, w) = 1 \iff M^\top \chi_v = \chi_w. \tag{26}$$

Let $f' = M^\top f$, then

$$\begin{aligned} f'(w) &= \sum_{v\in V} M^\top(w, v) = \sum_{v\in V} M(v, w) \\ &= \sum_{v\in V} \mathbb{1}(g_M(v) = w) = |\{v : g_M(v) = w\}| = |\{v : X(v) = w\}|. \end{aligned}$$

Let $(M_t)_{t\geq 0}$ be an i.i.d. sequence of matrices in $\mathcal{M}$ with distribution l and $(g_{M_t})_{t\geq 0}$ be the corresponding g-functions as defined above. Define $X_{t+1}(v)$ by

$$X_{t+1}(v) = g_{M_t}(X_t(v)),$$

where $X_0(v) = v$. This defines a vector valued process X_t where

$$X_{t+1} = g_{M_t}(X_t).$$

Since the matrices M_t are random samples, we have that $(X_t(v))_{t\geq 0}$ is a random process, indeed, it is a Markov chain. Observe that if initially we have one particle on each vertex, we have n Markov chains, $(X_t(v) : v \in V)_{t\geq 0}$ where $X_0(v) = v$ for each v. The process $(X_t(v) : v \in V)_{t\geq 0}$ is the Markov chain for the trajectories of the whole configuration of particles. Observe that when two particles meet, they keep moving together. Why? because if $X_t(v) = X_t(w)$, then, $X_{t+1}(v) = g_{M_t}(X_t(v)) = g_{M_t}(X_t(w)) = X_{t+1}(w)$.

The vector process $(f_t)_{t\geq 0}$ counts the number of particles on each vertex at time $t \geq 0$, i.e. $f_t(w) = |\{v : X_t(v) = w\}|$. Similarly to above,

$$X_{t+1}(v) = w \iff g_{M_t}(g_{M_{t-1}} \cdots g_{M_0}(v)) = w \iff M_t^\top \cdots M_0^\top \chi_v = \chi_w.$$

The set $\mathcal{M}$ is closed under multiplication, i.e. if $M_1, M_2 \in \mathcal{M}$, then $M_1M_2 \in \mathcal{M}$. Let $U = M_0 \cdots M_t$, then $U \in \mathcal{M}$ and $U^\top \chi_v = \chi_w$ if and only if $U(v,w) = 1$. As $f = \sum_{v \in V} \chi_v$, we have

$$f_{t+1} \;=\; M_t^\top \cdots M_0^\top f \;=\; U^\top f \;=\; \sum_{v \in V} U^\top \chi_v,$$

and

$$f_{t+1}(w) = |\{v : U^\top \chi_v = \chi_w\}| = |\{v : X_{t+1}(v) = w\}|.$$

Finally we remark that the fact that $U = M_0 \cdots M_t$ (i.e. $U\xi = M_0 \cdots M_t \xi$) explains the notion that a coalescing process is the time reversal of the corresponding linear voting process. At this point we might guess that $(X_t(v))_{t \geq 0}$ is a Markov chain with transition matrix $H(l)$, i.e. the mean matrix of l.

Proposition 1. *The Markov chain $X_t(v)$ has transition matrix $H(l)$.*

Proof. Suppose that $X_t(v) = w$, then

$$X_{t+1}(v) = g_{M_t}(X_t(v)) = g_{M_t}(w).$$

Then

$$\mathbb{P}(X_{t+1}(v) = u | X_t(v) = w) = \mathbb{P}(g_{M_t}(w) = u) = \mathbb{P}(g_M(w) = u),$$

where we use the fact that M_t has the same distribution as M because they are i.i.d samples. Finally,

$$\mathbb{P}(g_M(w) = u) = \mathbb{P}(M(w,u) = 1) = \sum_{M \in \mathcal{M}} l(M) M(w,u) = H(w,u).$$

Therefore, $H(l)$ is the transition matrix of $X_t(v)$. □

In Section 2 we gave some examples of voting process, let us show the dual processes of these examples. Remember that in the voting process we sample a matrix M and then we multiply by it. Note that the matrix M has the associated function g. By (26), we have that all particles in v move to $g(v)$ in the coalescing process. In words, if in the linear voting model, v pulls from w, i.e. $w = g(v)$, then in the coalescing process all particles in v moves to w. If in the linear voting model v pushes on w, i.e. $g(w) = v$, then in the coalescing process all the particles from w move to v.

Example 5. ***Dual of synchronous pulling model.***

The model is described as follows. In each round each vertex pulls the opinion of a random neighbour. Thus, in the dual each vertex v pushes the particle in v (if any) to a random neighbour of v. Note this is equivalent to saying that each particle moves to a adjacent vertex selected uniform at random. Thus each particle moves as a random walk on G until two or more of them meet, and after meeting the particles move together. This model is known simply as a *Coalescing Random Walk*. It has been extensively studied, see e.g. [1], [3], [4], [11] and [12] for further references.

Example 6. ***Dual of asynchronous pulling model.***

Remember that in each round we select a random vertex v and v pulls the value of a random neighbour, say w. Then the dual is described as follows. In each round select a random vertex v and then push the particles of v (if any) to a random neighbour w. We call this dual the pushing particles model.

Example 7. ***Dual of asynchronous pushing model.***

This model is similar to the asynchronous pulling model. In each round we select a random vertex v, and v pushes its value on a random neighbour, say w. In the dual at each round we select a random vertex v and then v pulls the particles of a random neighbour w. This model is called the pulling particles model.

Example 8. ***Dual of abusing push model.***

In this model in each round a random vertex v is selected and then v pushes its opinion on all its neighbours. In the dual a random vertex v is selected and then v pulls all the particles in its neighbours. We call this the abusing thief model.

Example 9. ***Dual of pull-push model.***

The model is self-dual.

5.1 The dual relation

We present the relation between the voting process and its coalescing particle dual. This relation is in a distributional sense, that is, we do not have a coupling between voting and coalescing processes but do we have distributional equalities, in particular we have a relation between the probability of being in consensus in the voting process and the probability of agreement in the coalescing process.

We recall the definition of the agreement time. Consider a partition $\mathcal{A} = \{A_1, \ldots, A_m\}$ of the vertices V. We say the coalescing process is in agreement at time t with respect to $\mathcal{A}$, if for some index $i \in \{1, \ldots, m\}$ we have $f(A_i) = n$, and consequently, $f(A_j) = 0$ for all $j \neq i$. An important partition is the one given by a configuration of opinions of the voting process. Let ξ be a vector of opinions, then define $\mathcal{A}_\xi$ as the partition of vertices of V into sets with the same opinion in ξ. A particular case is when all vertices have different opinions in ξ. In that case the partition the vertices into singletons and thus we have agreement in the coalescing particles if and only if all particles are together, that is, all particles coalesce into one. For this case, we define the coalescing time τ_{coalsc} as the first time all particles are together.

For convenience we restate Theorem 3

Theorem 6. *Let V a set of vertices and let $l \in \Pi$. Suppose that $(\xi_t) \sim \mathcal{LVM}(l, \xi)$ and $(f_t) \sim \mathcal{CP}(l, \mathbf{1})$. Then, for every $t \geq 0$,*

$$\mathbb{P}(\xi_t \textit{ is in consensus}) = \mathbb{P}(f_t \textit{ is in agreement with respect to } \mathcal{S}_\xi).$$

Proof of Theorem 3. Let $(\xi_t) \sim \mathcal{LVM}(\xi, l)$. If $t = 0$ and $\xi_t = \xi$, the result is trivial. Assume that ξ is not in consensus. Let Q be the set of possible opinions. For $b \in Q$ let $S_b = \{v : \xi(v) = b\}$. Let χ_b be the characteristic vector for S_b. Then ξ can be written as

$$\xi = \sum_{b \in Q} b\chi_b.$$

Let $\xi_t = M_{t-1}...M_0\xi$. We say ξ_t is in consensus at a if $\xi_t = a\mathbf{1}$. If so, as

$$M_{t-1}...M_0\xi = M_{t-1}...M_0(\sum_{b\in Q} b\chi_b) = \sum_{b\in Q} bM_{t-1}...M_0\chi_b = a\mathbf{1},$$

which (assuming $a \neq 0$) is equivalent to $M_{t-1}...M_0\chi_a = \mathbf{1}$, and thus $M_{t-1}...M_0\chi_b = 0$ for $b \neq a$. Therefore

$$M_{t-1}...M_0\xi = a\mathbf{1} \Leftrightarrow M_{t-1}...M_0\chi_a = \mathbf{1} \Leftrightarrow \langle M_{t-1}...M_0\chi_a, \mathbf{1}\rangle = n \Leftrightarrow \langle \chi_a, M_0^\top...M_{t-1}^\top\mathbf{1}\rangle = n. \quad (27)$$

Define $f_t' = M_0^\top...M_{t-1}^\top\mathbf{1}$. Observe the event $\langle\chi_a, f_t'\rangle = n$ is equivalent to $f_t'(S_a) = \sum_{v\in S_a} f_t'(v) = n$. Consider the coalescing process at time t, $f_t = M_{t-1}^\top...M_0^\top\mathbf{1}$ as defined in equation (25). It is clear that f_t and f_t' has the same distribution (denoted $f_t' \overset{\mathcal{D}}{=} f_t$) since the matrices are iid. Since consensus can be attained by one and only one opinion, we have that

$$\begin{aligned}
\mathbb{P}(\xi_t \text{ is in consensus }) &= \mathbb{P}(\bigcup_{a\in Q}\{M_{t-1}\cdots M_0\xi = a\mathbf{1}\}) \\
\text{(disjoint events)} &= \sum_{a\in Q}\mathbb{P}(\{M_{t-1}\cdots M_0\xi = a\mathbf{1}\}) \\
\text{(equation (27))} &= \sum_{a\in Q}\mathbb{P}(\langle\chi_a, f_t'\rangle = n\}) \\
&= \sum_{a\in Q}\mathbb{P}(f_t'(S_a) = n) \\
(f_t' \overset{\mathcal{D}}{=} f_t) &= \sum_{a\in Q}\mathbb{P}(f_t(S_a) = n) = \mathbb{P}\left(f_t(S_a) = n \text{ for some } a \in Q\right) \qquad (28)
\end{aligned}$$

Finally, as in equation (5), we have that the event $\{f_t \text{ is in agreement with respect to } \xi\}$ can be written as $\{\exists_{a\in Q} f_t(S_a) = n\}$, finishing our proof. □

Remember that if ξ_t is in consensus then it is in consensus for all successive times. Then

$$\mathbb{P}(\tau_{\text{cons}} \leq t) = \mathbb{P}(\xi_t \text{ is in consensus}). \quad (29)$$

By Theorem 3 we have that $\mathbb{P}(\xi_t \text{ is in consensus}) = \mathbb{P}(f_t \text{ is in agreement with respect to } \xi)$. On the other hand, suppose all the opinion of ξ are different, then S_ξ is a partition of V into singletons, then f_t is in agreement if and only if all particles are together, but if at time t all particles are together, then at all later times this property holds, and

$$\mathbb{P}(\tau_{\text{cons}} \leq t) = \mathbb{P}(f_t \text{ is in agreement with respect to } \xi). \quad (30)$$

As a corollary of Theorem 3 and equations (29) and (30), we have the following result

Corollary 2. *Let $(\xi_t) \sim \mathcal{LVM}(l, \xi)$ and $(f_t) \sim \mathcal{CP}(l, \mathbf{1})$ and assume all vertices have different opinions in ξ. Then*

$$\tau_{\text{coalsc}} \overset{\mathcal{D}}{=} \tau_{\text{cons}}. \quad (31)$$

5.2 Coalescence time

Let $(f_t) \sim CP(l, \mathbf{1})$ be a coalescing process and let $(X_t(v) : v \in V)_{t\geq 0}$ be the trajectories of the particles in such process. Remember that $X_t(v)$ gives the position at time t of the particle starting at v. Define $\tau_{\text{meet}}(v,w)$, the meeting time of $X_t(v)$ and $X_t(w)$, by

$$\tau_{\text{meet}}(v,w) \equiv \min\{t \geq 0 : X_t(v) = X_t(w)\}.$$

As usual $\min(\emptyset) = \infty$. Define T_{meet} as

$$T_{\text{meet}} \equiv \max_{v,w\in V} \mathbb{E}(\tau_{\text{meet}}(v,w)).$$

We relate the meeting time and the coalescing time in the following variant of the Matthew's Method (see [1], Theorem 2.26).

Theorem 7. *Let $n = |V| \geq 2$ and assume that the expected meeting time of every pair of particles is finite then*

$$\mathbb{E}(\tau_{\text{coalsc}}) \leq h_{n-1} T_{\text{meet}},$$

where $h_n = \sum_{k=1}^{n} \frac{1}{k}$.

Proof. Let $v_1, \ldots, v_n$ an arbitrary ordering of the states V. In this proof particle v is the particle that starts at vertex v. Define the cluster of particle v_1 as

$$C_t = \{v \in V : X_t(v) = X_t(v_1)\}.$$

Note that $v_1 \in C_0$ and that $C_t \subseteq C_{t+1}$ for all $t \geq 0$. Define the hitting time $T_0 = 0$, and for $k > 0$

$$T_k = \min\{t > T_{k-1} : C_t \neq C_{t-1}\}.$$

As usual, $\min\{\emptyset\} = \infty$. The hitting time T_k is exactly the k-th time in which particle v_1 meets with some other particles. Let $K = \max\{k \geq 0 : T_K < \infty\}$, i.e. K is the number of times the cluster increases in size until it has size n and T_K is the time at which the cluster has size n, and thus $\tau_{\text{coalsc}} = T_K$.

Consider a random ordering $\pi_2, \ldots, \pi_n$ of $V\setminus\{v_1\}$, and for consistency define $\pi_1 = v_1$. For $2 \leq k \leq n$ define the hitting times $S_k = \min\{t \geq 0 : \{\pi_2, \ldots, \pi_k\} \subseteq C_t\}$. Observe that $S_n = \min\{t \geq 0 : \{\pi_2, \ldots, \pi_n\} \subseteq C_t\} = \tau_{\text{coalsc}} = T_K$. We compute the expected value of S_n.

$$S_n = S_2 + \sum_{i=2}^{n-1}(S_{i+1} - S_i) = S_2 + \sum_{i=2}^{n-1}(S_{i+1} - S_i)\mathbb{1}_{\{S_{i+1} > S_i\}}.$$

Taking expected values given the ordering π and the position of the particles up to time S_i, we obtain that

$$\mathbb{E}((S_{i+1} - S_i)\mathbb{1}_{\{S_{i+1}>S_i\}}|\pi, \{X_t, t \leq S_i\}) =$$
$$\mathbb{E}(\tau_{\text{meet}}(X_{S_i}(\pi_1), X_{S_i}(\pi_{i+1}))|\pi, \{X_t, t \leq S_i\})\mathbb{1}_{\{S_{i+1}>S_i\}}. \tag{32}$$

The last equality holds because $(S_{i+1} - S_i)$ represents the time when two particles starting from positions $X_{S_i}(\pi_1)$ and $X_{S_i}(\pi_{i+1})$ meet, i.e. $(S_{i+1} - S_i) = \tau_{\text{meet}}(X_{S_i}(\pi_1), X_{S_i}(\pi_{i+1}))$. Moreover

$\mathbb{1}_{\{S_{i+1}>S_i\}} = \mathbb{1}_{\{X_{S_i}(\pi_1)\neq X_{S_i}(\pi_{i+1})\}}$ because if $X_{S_i}(\pi_1) = X_{S_i}(\pi_{i+1})$ then $(S_{i+1}-S_i) = 0$. As $\mathbb{1}_{\{X_{S_i}(\pi_1)\neq X_{S_i}(\pi_{i+1})\}}$ is a function of the random order π and of the position of the particles at time S_i, we can take $\mathbb{1}_{\{S_{i+1}>S_i\}}$ out of the conditional expectation. Also

$$\mathbb{E}(\tau_{\text{meet}}(X_{S_i}(\pi_1), X_{S_i}(\pi_{i+1}))|\pi, \{X_t, t\leq S_i\}) \leq T_{\text{meet}}.$$

Taking the expected value of equation (32) we get

$$\mathbb{E}(S_{i+1}-S_i) \leq T_{\text{meet}}\mathbb{P}(S_{i+1}>S_i).$$

We need to prove that $\mathbb{P}(S_{i+1}>S_i)\leq 1/i$. For such task let us define the following sets, for $1\leq i\leq K$

$$A_i = C_{T_i}\setminus C_{T_{i-1}}.$$

Clearly the sets $\{A_i\}$ are disjoint and their union is $V\setminus\{v_1\}$. Also, note that for every $i\leq n$ there exists $l\leq K$ such that $S_i = T_l$ because by time S_i the cluster of π_1 increases its size and thus that time corresponds to one of the T_k values. If so, $S_{i+1}>S_i=T_l$ means that $\pi_{i+1}\notin B_l = \cup_{k=1}^{l} A_k$. On the other hand we have that for every $k\leq i$ there exists $l_k\leq l$ such that $\pi_k\in A_{l_k}\subseteq B_l$. Let $f:\{2,\dots,n\}\to\mathbb{N}$ be such that $f(k)=j$ if and only if $\pi_k\in A_j$. Note that from the definition of f we have $\pi_k\in A_{f(k)}$ and that $l=\max\{f(k):k=1,\dots,i\}$. Then

$$\begin{aligned}
\{S_{i+1}>S_i\} &= \{\pi_{i+1}\notin B_l\}\\
&= \bigcap\{\pi_{i+1}\notin \bigcup_{j=1}^{l} A_j\}\\
&= \bigcap\{\pi_{i+1}\notin \bigcup_{k=2}^{i} A_{f(k)}\}\\
&= \bigcap_{k=2}^{i}\{f(i+1)>f(k)\} \qquad (33)
\end{aligned}$$

Right now the problem is reduced to prove $\mathbb{P}(\bigcap_{k=2}^{i}\{f(i+1)>f(k)\})\leq 1/i$. Define $r:V\to[n]$ as the bijective function with $r(v_1)=1$ and such that for every $v,w\in V$ we have that $r(v)<r(w)$ if vertex v_1 meets with v before v_1 meets w; and $r(v)<r(w)$ if v_1 meets v and w at the same time but $v<w$. Basically, r gives the order in which vertices meet with v_1. Observe that $\{f(i+1)>f(k)\}\subseteq\{r(\pi_{i+1})>r(\pi_k)\}$, then

$$\mathbb{P}\left(\bigcap_{k=2}^{i}\{f(i+1)>f(k)\}\right)\leq \mathbb{P}\left(\bigcap_{k=2}^{i}\{r(\pi_{i+1})>r(\pi_k)\}\right),$$

but $\pi_2,\dots,\pi_n$ is a random ordering of the vertices and thus all order are equally likely, then

$$\mathbb{P}(r(\pi_{i+1})>r(\pi_2),\dots,r(\pi_{i+1})>r(\pi_i)\}) = 1/i.$$

thus

$$\mathbb{P}(S_{i+1}>S_i)\leq 1/i,$$

therefore

$$\mathbb{E}(S_n)\leq T_{\text{meet}}\left(1+\frac{1}{2}+\cdots+\frac{1}{n-1}\right),$$

finishing the proof. □

References

[1] David Aldous and James Allen Fill. Reversible markov chains and random walks on graphs, 2002. Unfinished monograph, recompiled 2014, available at `http://www.stat.berkeley.edu/$\sim$aldous/RWG/book.html`.

[2] 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.

[3] 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.

[4] J Theodore Cox. Coalescing random walks and voter model consensus times on the torus in zd. *The Annals of Probability*, pages 1333–1366, 1989.

[5] 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.

[6] Peter Donnelly and Dominic Welsh. Finite particle systems and infection models. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 94, pages 167–182. Cambridge Univ Press, 1983.

[7] Rick Durrett. *Probability: theory and examples.* Cambridge university press, 2010.

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

[9] David Asher Levin, Yuval Peres, and Elizabeth Lee Wilmer. *Markov chains and mixing times.* AMS Bookstore, 2009.

[10] 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.

[11] Roberto Oliveira. On the coalescence time of reversible random walks. *Transactions of the American Mathematical Society*, 364(4):2109–2128, 2012.

[12] Roberto Oliveira. Mean field conditions for coalescing random walks. *The Annals of Probability*, 41(5):3420–3461, 2013.