

This electronic thesis or dissertation has been downloaded from the King's Research Portal at <https://kclpure.kcl.ac.uk/portal/>



Voting Models on Graphs

Rivera Aburto, Nicolas Andres

Awarding institution:
King's College London

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

END USER LICENCE AGREEMENT



Unless another licence is stated on the immediately following page this work is licensed

under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International

licence. <https://creativecommons.org/licenses/by-nc-nd/4.0/>

You are free to copy, distribute and transmit the work

Under the following conditions:

- Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
- Non Commercial: You may not use this work for commercial purposes.
- No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Voting Models on Graphs

Nicolás Rivera

Thesis is submitted to King's College London, University of
London, in partial fulfilment of the requirements for the
degree of the Doctor of Philosophy



Department of Informatics

KING'S COLLEGE LONDON,
UNIVERSITY OF LONDON, U.K.

September 2017

To my family, and to Tami.

Acknowledgements

I would like to express my gratitude to my supervisor Colin Cooper for his continuous support and jokes. I will miss our meetings. My gratitude also goes to my co-supervisor Tomasz Radzik for his motivation and for endless afternoons of discussions. My thanks also to all my co-authors.

Also, I would like to thanks my friends in KCL, Andrew and Nan. I hope we can keep working together.

Many thanks to my friends Luke, Jessie, and Jova. Special thanks to my Chilean family in Oxford, Andrea, Fran, and Nico Raab.

I am grateful for the support, encouragement, and love of my parents and my brother throughout all these years. Thanks for taught me the value of hard work, and the importance of education.

Finally, I would like to acknowledge my moma, Tami. These four years in UK have been amazing. I am very glad to have you as my partner in crime, and I hope we can keep doing this forever. You are the best ♡.

This thesis was supported by Becas Chile.

Abstract

This thesis deals with voting processes. A Voting process models the exchange of opinions in a population of agents, commonly represented by vertices of a graph. Usually, an opinion represents the current state of the agent, and by interacting with neighbouring agents, such an opinion may change over time.

Depending on the context, the set of valid opinions can be, for instance, $\{0, 1\}$, $\{+, -\}$, $\{\text{agree}, \text{disagree}\}$, and $\{\text{healthy}, \text{infected}\}$ among others. The main research questions about a voting process are: i) Does the system reach consensus?, that is, if the system reaches a stable configuration where all vertices share the same opinion. ii) Subject to reaching consensus, what is the final opinion of the system? iii) How long does it take to reach consensus?

Throughout this thesis, we study three stochastic models of voting. Firstly, we introduce and study the Linear Voting model. The main motivation to introduce this model is to make a step toward unifying certain models of voting on graphs in a common framework. In this regard, our model is proven to be flexible enough to cover several other models as particular instances without compromising tractability. As a particular case, our model subsumes well-known models as the voter model (pull voting) and push voting. Moreover, due to its tractability, we are able to extend several of the well-known techniques used to study pull voting, to properties of this much richer model. Among the studied properties, we include consensus time, winning probabilities, and the construction of a dual process.

Secondly, we analyse the Coalescing and Branching random walks (COBRA) pro-

cess, which is a model of rumour spreading on a connected graph. Here, we establish a duality relation between the COBRA process and an infection process called BIPS. The BIPS process can be seen as a voting model with bias toward a fix opinion. The advantage of our approach, is that the BIPS process is much more tractable than the original COBRA process. By using this dual process, we obtain several results concerting the cover time of the Cobra process, which corresponds to the first time such that all vertices are informed.

Finally, we study three versions of discordant voting processes on graphs. In discordant voting, only vertices with different opinions are allowed to interact. In first place, we study discordant voting processes on several classes of graphs, showing that the expected consensus time can be polynomial in the number of vertices of the graph, or exponential, depending on the graph topology. Later, we define a general discordant process, parametrised in $\beta \in [0, 1]$, and study it on the complete graph. We compute the expected consensus time for all values of β , showing that several phase transitions occur as β moves from 0 to 1. Indeed, the expected consensus time is $\Theta(n \log n)$ when $\beta = 0$, $\Theta(n^2)$ when $\beta = 1/2$, and $\Theta(2^n)$ when $\beta = 1$.

Contents

List of Figures	6
1 Introduction	8
1.1 Dynamics on Networks	8
1.2 Previous Work	10
1.2.1 Other Models	17
1.3 Thesis Contribution and Organisation	19
1.4 Notation	21
1.5 About the Results.	22
2 The Linear Voting Model	23
2.1 Introduction	23
2.2 Definitions and Contributions.	25
2.3 The Model	29
2.3.1 Mean Matrix	32
2.4 Winning probability.	34
2.5 Consensus Time.	39
2.5.1 Introduction	39
2.5.2 Proof of Theorem 2	40
2.5.3 Consensus Time Examples	48
2.6 Coalescing Process and Duality	51
2.6.1 Coalescing Process	52

2.6.2	The Dual Relation	55
2.6.3	Coalescence Time and Meeting Times	58
2.7	Analysis on Standard Graph Families	61
2.7.1	Asynchronous Pull Model.	62
2.7.2	Push Model.	67
2.7.3	Abusive Pushing Model.	70
2.7.4	Oblivious Model.	70
3	Coalescing Branching Random Walks on Graphs	73
3.1	Introduction	73
3.2	Definitions and Contributions	75
3.3	Duality Between COBRA and BIPS Processes	80
3.4	Analysis of the BIPS process	82
3.4.1	Incremental Nature of BIPS	82
3.4.2	Sequential Analysis	86
3.5	The BIPS Process on Regular Graphs	90
3.5.1	Large Infection Size	91
3.5.2	Small Infections	97
4	Discordant Voting Model on General Graphs	104
4.1	Introduction	104
4.2	Birth-and-Death Chains	108
4.2.1	Push and Pull Chains	109
4.2.2	Push Chain: Bounds on Hitting Times	110
4.2.3	Biased Random Walk With Reflecting Barriers	113
4.2.4	Coupling	114
4.3	Voting on the Complete Graph K_n	115
4.3.1	Voting on the Cycle	116
4.4	Voting on the Star Graph S_n	127
4.5	Voting on the Double Star	136

4.6	Voting on the Barbell Graph	142
5	Discordant Voting Model on the Complete Graph	145
5.1	Birth-and-Death Chains	147
5.1.1	Push and Pull Chains	149
5.1.2	Coupling	150
5.2	Analysis of the Case $\beta \leq 1/2$	151
5.2.1	A General $n \log n$ Estimate	151
5.2.2	Case $\beta \in (0, 1/2)$: Proof of Theorem 69	153
5.2.3	Proof of Theorem 69.2 and 69.3	154
5.3	Analysis of the Case $\beta \in (1/2, 1)$	159
5.3.1	Estimation of $\mathbb{P}_N(T_0 < T_N^+)$	160
5.3.2	Estimation of $\pi(N)$	162
5.3.3	Proof of Theorem 70	164
A	Standard Probability Results	168

List of Figures

2.1	Left panel: Double star S_{24}^* . Right panel: Barbell K_{16}^*	62
2.2	Results of different models on different graphs. Each entry represent the expected consensus time of a model in a particular graph.	62
4.1	Comparison of expected time to consensus ($\mathbb{E}\tau_{\text{cons}}$) for discordant and ordinary asynchronous voting protocols on connected n -vertex graphs, starting from $R = B = n/2$	108
4.2	Examples of two configurations in C_{18}	117
4.3	Cases for discordant edge uv . ℓ'_x represents the length of the run containing discordant vertex x , for $x \in \{u, v, w, q\}$	119
4.4	Lower bound configuration in C_n	126
4.5	Pseudo-states for the push process on the Star	128
4.6	Pseudo-states for the pull process on the Star	134
4.7	Double star S_{24}^* with half of the vertices coloured blue, and half coloured red.	136
4.8	S_1 with r leaves coloured R	139

Chapter 1

Introduction

1.1 Dynamics on Networks

The understanding of the behaviour of a group of interacting agents is one of the central topics of many disciplines of computer science, mathematics, and physics among others. Some examples include the field of distributed computing, in which the objective is to coordinate a group of agents in order to solve a collaborative task. This is done by letting the agents to perform local computations and to share their (partial) knowledge with neighbouring agents. In physics, many particle systems attempt to model highly-dynamics environments by modelling the local dynamic of an agent, and local rules of interaction between them. A well-known instance is e.g., the Ising model, which is a model of ferromagnetism in statistical mechanics. In ecology, several models have been proposed to explain interactions between species in a community. A notable example is the Moran process, which is used to model the spread of a mutant gene in a community.

Generally, the interaction between agents creates high levels of complexity and thus, from a mathematical point of view, the construction of realistic and detailed systems leads to intractable models. As a consequence, it is hard to study them with the available technology. In this thesis, instead of studying complex and real-

istic models, we focus on simple models, expecting that the understanding of simple dynamics allows us to comprehend dynamics of more complex and real-life scenarios.

The main topic of study of this thesis are the so-called voting dynamics, or exchange of opinion dynamics. In these processes, agents have opinions, which can be modified as a reaction to the interaction with neighbouring agents, generally adopting the opinion of one of them.

Our starting point is to model agents as vertices in a graph, where the possible interactions are represented by the edges. The opinions will be represented as a function mapping vertices to some finite space which encodes their values (usually $\{0, 1\}$). The dynamics are introduced in a sort of algorithmic fashion: at every time step, we apply a rule that may change the current opinion of one or more vertices. Most of the rules we are going to consider use a source of randomness, i.e., randomised algorithms.

Following the idea above, in this thesis, we study three dynamics. In first place, we introduce and study the *Linear voting model*. This model subsumes well-known voting dynamics as the pull and push voting models. In second place, we study the so-called *coalescing and branching random walks*. Even though this process is not a voting dynamic, it can be studied by a voting/infection process via a duality relationship. Finally we study *discordant voting*, in which interactions between agents are restricted only to agents with different opinions.

A fundamental concept in the study of voting dynamics is the notion of consensus. We say that a voting dynamic reaches consensus if all the vertices reach a state where they share the same opinion, and such opinion never changes again. If a voting dynamic reaches consensus there are two important questions. What is the final opinion of the system? in other words, who wins the poll? How long does it take to reach consensus?

1.2 Previous Work

The study of voting models began with the independent works of Clifford and Sudbury [18], and Holley and Liggett [53], in which they introduced the so-called voter model. The voter model is a continuous-time process on a graph, in which each vertex has an initial opinion (0 or 1), and a (independent) Poisson clock that rings at rate 1. Then, when the clock of vertex v rings, v selects a random neighbour and *pulls* (adopts) its opinion. Since we believe the name ‘voter model’ is too general for a very specific model, we refer to this process as the (continuous-time) pull voting model.

This process has been extensively studied, especially in the case when the underlying graph is infinite. In particular in d -dimensional lattices and infinite d -regular trees. We refer to [59] for a general account of this setting, and for other similar continuous-time processes.

As usual, the behaviour of models on infinite and finite structures is quite different, and thus a proper study on finite structures is needed. The study of voting models on finite graphs started with the seminal paper of Donnelly and Welsh [33]. In such a work, the authors studied the continuous-time pull voting, and the continuous-time *push* voting model on finite graphs. The push model is quite similar to the pull model. The only difference is that when the clock of vertex v rings, vertex v selects a random neighbour w and *pushes* (forces) its opinion on w . In general, the behaviour of the pull and push models tends to be quite different, nevertheless, the two processes were proved to be equivalent on regular graphs [33].

The main tool used by Donnelly and Welsh to study the continuous-time pull and push voting models was a *dual* process. This dual process came from the idea of looking the voting process *backwards in time*. In the case of pull voting, the dual process is the so-called (continuous-time) coalescing random walk [53]. Here, we have independent continuous-time random walks starting from different vertices of the graph. Then, whenever two particles meet in a vertex, they coalesce (merge)

into one. For the push model, the dual process is a variation of the continuous-time coalescing random walk, but with more complicated transition rates. There is a deep relation between the voting process and its dual. Indeed, many properties of one process translate to a different property in the dual process. For example, one application is the computation of the probability that a certain opinion wins. Suppose we have a connected graph $G = (V, E)$ where each vertex has an initial opinion 0 or 1. For the pull model, it has been proved that the probability that a certain opinion, say 0, wins is $d(S)/d(V)$, where S is the set of vertices with initial opinion 0, and $d(X)$ is the sum of the degrees of the vertices of $X \subseteq V$. The connection with coalescing random walks comes from the fact that $\pi(S) = d(S)/d(V)$ is the stationary distribution of the set S of a random walk on the graph G . For the case of the push model, the probability that 0 wins is proportional to $\sum_{v \in S} d(v)^{-1}$, which corresponds to the stationary distribution of process similar to the coalescing random walks.

In the computer science community, the importance of voting models became apparent because of the application of these models to problems in distributed algorithms, in particular, to the *consensus* and *leader election* problems. These problems are described as follows. In the consensus problem, we have a group of indistinguishable agents with different opinions. Then, we need to provide the agents with an algorithm (the same algorithm for all of them) such that, after the execution of such an algorithm, the agents agree in a common opinion.

The leader election problem is similar to consensus problem, but instead of agreeing in an opinion, the agents have to choose a leader. It is known that those problems are impossible to solve if the agents use a deterministic algorithm because it is impossible to break the inner symmetry between the identical agents [61]. Due to this fact, it is fundamental to use a randomised algorithm to break the symmetry between the agents, and thus, to solve the problem.

The first application of voting models to distributed computing is found in the independent works of Nakata et al. [66], and Hassin and Peleg [52], where they use

the pull process to solve the consensus problem.

Due to the discrete nature of the interest of Nakata et al., and Hassin and Peleg, they studied a discrete version of the pull process. We call such a version the *discrete-time synchronous pull voting model*. Since this thesis is mainly about discrete time processes, we omit the reference to discrete time, and we just call it the synchronous pull model. The word *synchronous* means that all vertices act at the same time, and all vertices use the same (randomised) algorithm, nevertheless, their sources of randomness are independent. For example, in the case of the synchronous pull model, at each time-step and at the same time, all vertices (independently of each other) pull the opinion of a random neighbour.

In the aforementioned works, the authors provided new tools to analyse the pull model. The approach of [66] has a combinatorial nature, while the approach of [52] is more probabilistic. In particular, Hassin and Peleg proved the following. Suppose the agents are represented by vertices of a graph. If S_t is the set of vertices with opinion 0 at round t , then $(d(S_t))_{t \geq 0}$ is a martingale, where $d(S_t)$ is the sum of the degrees of the vertices in S_t . By using basic martingale properties, they proved that the probability that opinion 0 wins is $d(S_0)/d(V)$, recovering the result of Donnelly and Welsh for the discrete-time synchronous pull model. This martingale argument can be also applied to the continuous-time model.

Since voting processes can be seen as randomised algorithms to solve the consensus problem in distributed computing, it is very important to analyse the running time of such an algorithm. In this case, the algorithm ends when all vertices have the same opinion. The first time all vertices share a common opinion is called the *consensus time*. Consensus time is a very elusive random variable, and it requires a lot of work (and luck) to be able to compute simple functions of it, e.g., its expected value. Indeed, in the work of Hassin and Peleg, the authors choose to work with the dual process rather than the voting process in order to study the consensus time.

For the case of synchronous pull voting, its dual process is very similar to the continuous-time case. The dual process is described as follows. Consider independent

(discrete-time) random walks. At each time-step, all of them (independently) move to a random neighbour, and if two or more walks meet in a vertex, they coalesce (merge) into one. We call this process the (discrete-time) synchronous coalescing random walks. A very important quantity in this process is the coalescing time, which is defined as the first time all walks are together. Recall that the only difference with the continuous time version is that in the discrete-time process all the walks move at the same time-step, while in the continuous-time version a walk moves only when its own clock rings.

In either discrete or continuous time, the dual relationship establishes that the coalescing time of the coalescing random walks has the same distribution as the consensus time of the pull voting model given that all vertices have different initial opinions.

For the discrete-time coalescing random walks, Hassin and Peleg [52] proved that the expected coalescing time in a connected non-bipartite n -vertex graph G , is $\mathcal{O}(T_{\text{meet}} \log n)$, where T_{meet} is the expected meeting time of two independent random walks, starting from the worst initial configuration. The former result gives us that the expected consensus time is $\mathcal{O}(n^3 \log n)$ for any connected non-bipartite n -vertex graph. Cooper et al. [20] improved the previous result by proving that the consensus time is $\mathcal{O}(n/(\nu(1 - \lambda_2)))$, where n is the number of vertices of G , λ_2 is the second largest absolute eigenvalue, 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 $\mathcal{O}(n^3)$ for any graph. For random d -regular n -vertex graphs, Cooper et al. [23] proved that when $d \geq 3$, the expected coalescing time is asymptotically $2\theta_d = \frac{d-1}{d-2}n$. Very recently, Kanade et al. [56] proved that the expected coalescing time on any finite graph G is

$$\mathcal{O} \left(T_{\text{meet}} \left(1 + \sqrt{\frac{T_{\text{mix}}}{T_{\text{meet}}}} \log n \right) \right),$$

where T_{mix} is the mixing time of the random walk on G .

In the continuous-time coalescing random walks, Oliveira [67] shows that the

expected coalescing time is $\mathcal{O}(H_{\max})$, where $H_{\max} = \max_{v,u \in V} H(v,u)$, and $H(v,u)$ is the expected time that it takes to a random walk to reach vertex u starting from v . Furthermore, in a posterior work [68], Oliveira proved that under certain conditions on the graph G , the expected coalescing time is $2m(G)$, where $m(G)$ is the meeting time of two independent random walks starting from stationary distribution.

Even though the evident similarity between the continuous-time and discrete-time coalescing random walks, it is not clear whether the results in one setting can be translated into the other setting.

Recently, a more ad-hoc approach to bound the consensus time for synchronous pull voting was taken by Berenbrink et al [12]. Indeed, using the martingale defined by Hassin and Peleg [52], together with a potential function argument, they showed that the expected consensus time is $\mathcal{O}((d_{\text{ave}}/d_{\min})(n/\Phi))$, where Φ is the conductance of the graph, and d_{ave} and $d_{\min}$ are the average and minimum degree of the underlying graph, respectively.

Albeit active research has been done for the pull voting model, the push voting model has not been widely studied. One important exception to this, comes from the mathematical biology community, where a generalisation of a discrete-time version of push model has been studied, such a process is called the Moran process [63, 58]. The Moran process was introduced in ecology in order to understand the effect of mutant genes in a population. In the Moran process, vertices have one initial opinion (or gene), and opinions have a fitness, represented by a positive real number. At each round, a random vertex is chosen proportionally to the fitness of its current opinion, and then, it pushes its opinion to a random neighbour. When all opinions have the same fitness, we recover the (discrete-time) push model. In this setting, finding the probability that a particular opinion wins is computationally hard, and thus, tractable closed formulas are only available for very specific cases. Contrary to the previous, consensus time is bit more tractable and a few results have been obtained. In particular, Díaz et al. [31] proved that in the case with two opinions, one with fitness r and the other with fitness 1, the consensus time is $\mathcal{O}(n^6)$ if $r = 1$,

and $\mathcal{O}(n^4)$ and $\mathcal{O}(n^3)$ for the cases $r > 1$ and $r < 1$, respectively.

Since the incorporation of the pull model to the computer science literature, several variations of this model have been proposed in order to reach consensus faster. One example is the Best-of-two model. In this model, at each round, synchronously, each vertex v pulls the opinion of two random neighbours, and if those opinions are equal, v adopts such an opinion. This was first studied by Cooper et al. [21] on random regular graphs where vertices have opinions red or blue. There, they proved that if the imbalance between the number of red and blue opinions is large enough, then in $\mathcal{O}(\log n)$ steps the process reaches consensus, and also the final opinion is the majority opinion at the beginning of the process. The result was generalised in [22] to general graphs with good expansion properties, and recently, [26] extended the results to allow more than two opinions in the system. On random sparse graphs, [1] proved that a $\mathcal{O}(\log \log n)$ consensus time is achieved if instead of choosing two neighbours, each vertex chooses at least five (the vertex adopts the majority opinion of them), and the initial imbalance between randomly allocated red and blue opinions is large enough.

In computer science, there is a particular interest in voting protocol on the complete graph. In this setting, one of the most popular protocols is the best-of-three model. Here, in synchronous rounds, each vertex pulls the opinion of three random neighbours and adopts the opinion of the majority of those (ties are broken at random). This model was studied in [10] on the complete graph, allowing the vertices to have one of k different opinions. They proved that, if k is not too large, and if the imbalance between the majority and second majority opinion is large enough, then the process reaches consensus in $\mathcal{O}(k \log n)$ time-steps. For other protocols on the complete graph that reach consensus in a similar time, we refer the reader to [8], [9], and references therein. Very recently, a few works have developed protocols that finish in time $\mathcal{O}(\log k \log n)$ under appropriate (usually stronger) initial conditions, see [46], [11].

Beyond the algorithmic point of view, several other voting models have been stud-

ied in different settings. An interesting example is the noisy voter model, introduced by Granovsky and Madras [50]. The noisy voter model is similar to pull voting, as at each time-step, we sample (only one) random vertex v , and with probability p , we rerandomise its opinion, that is, set the opinion of v to a random opinion from $\{0, 1\}$. Otherwise, v pulls the opinion of a random neighbour as standard pull voting. The continuous-time version of the process is similar, but each vertex has a Poisson clock that rings at rate 1. This model is of particular significance as it is related to other statistical mechanics processes such as the Ising model [29]. While the noisy voter model does not reach consensus due to the rerandomisation probability, it reaches a stationary distribution. Its mixing time has been studied in a few works. In particular, for the case $p = 1$, the mixing time is $\mathcal{O}(n \log n)$, as the process corresponds to a random walk on the hypercube of 2^n vertices [57]. When the rerandomisation probability p is constant ($0 < p < 1$, independent of the size of the graph), Ramadas [71] proved that, in a continuous-time setting, the mixing time of the noisy voter model is $\mathcal{O}(\log n)$. Such results translate to a $\mathcal{O}(n \log n)$ mixing time for the discrete case. Recently, Cox, Peres and Steif [29] proved that, under appropriate conditions on the underlying graph, the mixing time of the noisy voter model is $n \log n / (2p)$ for the discrete-time case, and that it exhibits total variation cut-off, that is, convergence to stationary distribution occurs in a window of length $o(n \log n)$.

Another interesting family of voting processes are the so-called discordant voting models, where only vertices with different opinions are allowed to interact. Discordant voting originated in the complex networks community as a model of social evolution (see e.g. [51] or [72]). The general version of the model allows for *rewiring*. The interacting vertices can break edges and reconnect elsewhere. This serves as a model of social behaviour in which vertices either change their opinion, or their friends. Holme and Newman [54] investigated discordant voting as a model of a self-organising network which restructures itself based on the acceptance or rejection of differing opinions among social groups. The process is described as follows. At each time-step, a random discordant edge $e = (u, v)$ is selected (i.e., an edge whose two endpoints

have different opinions), and a random endpoint $x \in \{u, v\}$ is chosen. Then, with probability $1 - \alpha$, x pulls the opinion of y , otherwise, vertex x breaks the edge and rewires to a random vertex with the same opinion as itself. Simulations suggest the existence of a threshold behaviour in α . This was investigated further by Durrett et al. [36] for sparse random graphs of expected degree 4 (i.e. $G(n, 4/n)$). The paper studies two different rewiring strategies, rewire-to-random (reconnect to a random vertex), and rewire-to-same (reconnect to a vertex with the same opinion), and finds experimental evidence of a phase transition in both cases. Basu and Sly [7] made a formal analysis of rewiring for Erdos-Renyi graphs $G(n, 1/2)$ with $\alpha = 1 - \beta/n$, where $\beta > 0$ is constant. They found that for either strategy, if β is sufficiently small, then the network quickly disconnects maintaining the initial proportions of opinions. And, as β increases, the minority proportion starts decreasing. A subsequent paper by Durrett et al. [6] examines the phase transitions for the intermediate case of dense random graphs $G(n, 1/n^a)$ where $0 < a < 1$.

Another very popular model is majority voting. Majority voting is a deterministic discrete-time process in which at each time-step, synchronously, each vertex changes its opinion to the majority opinion among its neighbours. Goles and Olivos [49], using a potential function argument, proved that the processes converges in $\mathcal{O}(|E|)$ steps to a two-periodic dynamic, where either the opinions changes at every time-step between two configurations, or the opinions remain constant. This bound is proven tight for a class of dense graphs [44].

1.2.1 Other Models

Albeit we mainly focus on voting models, there are several related processes on graphs. Here, we introduce a few of them.

Rumour Spreading. An important family of processes are the so-called randomised gossip or rumour spreading protocols. Here, one vertex has a rumour, and at each round, each vertex performs an action to propagate the rumour in the graph. A

well-known example of these protocols are the *pull*, *push* and *push-pull* protocols, introduced by [30]. Here, we consider discrete-time steps. In the pull protocol, at each round, each non-informed vertex v chooses a random neighbour, and v becomes informed if such a neighbour is informed. In the push protocol each informed vertex chooses a random neighbour, and such a neighbour becomes informed. In the push-pull protocol, both actions are performed at the same time. A continuous-time version of the model can be defined. In this version, vertices have an independent Poisson clock, and they perform actions when their respective clocks ring. For more details about this process and results on several graph topologies, we refer the reader to [16], [32], [43], [47], [48], [69], [2], and references therein.

First Passage percolation This is a continuous-time process, which is very related to the continuous-time push protocol for rumour spreading. The main difference is that while in the push protocol, each vertex has a Poisson clock, in first passage percolation, each edge has a Poisson clock, and when such a clock rings, the information flows through the edge (so both endpoints are informed if at least one of them was informed). While on regular graphs, the continuous-time version of rumour spreading, and first passage percolation are essentially the same, they behave quite different on irregular graphs. In this setting, several results have been proved for different classes of graphs. We refer the reader to [5], [14], [42] and [55] for more details.

Contact process. The contact process is a continuous-time process on a graph. This process is very similar to rumour spreading, with the difference that informed vertices become uninformed at certain rate. As a consequence of this, it is more common to define the contact process in terms of an infection where vertices heal themselves at some rate. The behaviour of the process depends on the structure of the graph, and on the rate of healing. On one hand, on infinite graphs, the infection may fixate on the graph (last forever) or extinct. In graphs such as $\mathbb{Z}^d$ or regular trees, the main problem has been to understand what is the appropriate healing rate to have fixation or extinction of the infection, see e.g. [60], [13], [70].

In finite graphs, the picture is simpler as the infection always extinguishes. There-

fore, the main problem is to compute the extinction time, and to understand how the healing rate produces a transition between fast and slow extinction (namely, polynomially or exponentially slow on the number of vertices). Some results in this direction are presented in [45], [15] and [65].

1.3 Thesis Contribution and Organisation

This thesis proposes to study different models of voting, and other related processes, on finite graphs. To this end, each chapter of this thesis contributes to the development of the field by either defining a new process, generalising already known processes, or by studying properties of them. Since this thesis covers a wide range of topics within voting processes, each chapter is as self-contained as possible. In this way, it is possible to read chapters in any order, without the need to read more than a couple of definitions from other chapters. Almost all the results of this thesis can be understood with general knowledge of Probability Theory, which is assumed. Nonetheless, a small appendix section is included at the end of this thesis to remind the reader a few results that are frequently used.

We expect, throughout the chapters of this thesis, to give the reader a better understanding of the beautiful world of voting models, and, in general, of stochastic dynamics on finite graphs.

Chapter 2. In this chapter, we introduce the Linear Voting Model. This model is introduced as a step toward unifying several models of voting in a common framework. In particular, our model subsumes well-known models as the pull, and the push models. We extend some of the techniques used in the study of the well-known pull voting to our more general context. This chapter begins with the definition of the process and some examples. Then, we give an explicit computation of the probability that a particular opinion i wins the poll. Later, we study the consensus time of the linear voting model. We finalise, by describing a dual relationship between the linear voting model and a general system of coalescing random walks. The core of this

chapter was published in [27].

Chapter 3. In this chapter, we study the coalescing and branching random walks, or COBRA walks for short. This model is defined as follows. We have a connected graph G and a token on top of one of the vertices. At each round, each token in the graph generates $k \geq 1$ copies of itself, and each copy independently moves to a random neighbour. After the movement, if two or more tokens meet in a vertex, they coalesce into one. When $k = 1$, we recover a standard random walk on G . The main quantity of interest is the cover time, which is the first time such that all vertices have been visited by at least one token. Even though the COBRA process does not seem to be related to voting, we show a dual relationship between the COBRA walk and a process called *bias infection with persistent source* (BIPS), which is an infection model (which in turn is a voting model where the opinions are infected/healthy) with bias toward being infected. By studying the BIPS process, we are able to obtain bounds for the cover time of the COBRA walk for the case $k = 2$. In particular, we prove that for r -regular graphs with n vertices, the cover time is $\mathcal{O}((r/(1-\lambda) + r^2) \log n)$ and $\mathcal{O}(\log(n)/(1-\lambda)^2)$, where λ is the second eigenvalue of the transition matrix of the random walk on G . For general connected graphs with n vertices, m edges, and maximum degree $d_{\max}$, we prove the bounds $\mathcal{O}(m + d_{\max}^2 \log n)$ and $\mathcal{O}(m \log n)$. All the results hold with high probability and in expectation. The results of this chapter were published in [24] and [25], and they improve over all previous results about cover times, which are found in [39] and [62].

Chapters 4. This chapter deals with discordant voting dynamics. Here, only vertices with different opinions are allowed to interact. A vertex is discordant if at least one neighbour have a different opinion. We study three types of discordant voting models: pull, push, and oblivious. In discordant pull voting, at each round, a random discordant vertex pulls the opinion of a random neighbour with different opinion. Similarly, in discordant push voting, a random discordant vertex pushes (instead of pull) its opinion on a random neighbour with different opinion. In discordant oblivious voting, we choose a random discordant edge (both endpoints have

different opinion) and one of the endpoints pulls the opinion of the other endpoint. Here, we focus on the study of the consensus time. In particular, we study the three versions of discordant voting over different graph topologies, showing that the expected consensus time behaves differently than its non-discordant counterpart. Some of the results of chapter 4 were published in [19].

Chapter 5. This chapter continues the study of discordant voting. In particular, we study a more general model, which is parametrised in $\beta \in [0, 1]$. When β takes values 0, $1/2$, and 1, we recover the push, oblivious and pull discordant models, respectively. We carefully study the expected consensus time in the complete graph K_n , for all ranges of β , finding several phase transitions as β moves from 0 to 1.

1.4 Notation

The chapters of this thesis are mostly self-contained, and they introduce and recall the necessary notation. Nevertheless, some general notation is shared among all the chapters. Throughout all this thesis, $G = (V, E)$ stands for a simple graph, where V denotes the set of vertices and E the set of edges. All graphs are connected unless stated otherwise. We use the notation $v \sim w$ to say that v and w are adjacent vertices. For $v \in V$, we denote by $N(v) = \{w \in V : w \sim v\}$, the neighbourhood of v , and define the degree $d(v) = |N(v)|$.

In this work, we consider discrete-time random processes. We usually denote random processes by $(X_t)_{t \geq 0}$. Here, $\{t \geq 0\}$ refers to the set of non-negative integers $\{0, 1, 2, \dots\}$. Usually, we refer to these times as rounds, time-steps, or simply, times. So, the expressions, “at time t ”, “at round t ”, “at step t ”, and “at time-step t ”, are all equivalent.

We usually define voting processes as a combination of *pull* and *push* actions. We say that a vertex v *pulls* the opinion of vertex w , if v adopts or imitate the opinion of w . We say that v *pushes* its opinion on w , if w adopts the opinion of v . Additionally, we say that a process on a graph is synchronous if at each round, all

vertices perform a (random) action at the same time. Usually, each vertex performs the action independently of other vertices. The process is asynchronous, if at each round, only one vertex (usually chosen uniformly at random) performs an action.

Standard asymptotic notation is used throughout this thesis. Given two function $f, g : \{0, 1, \dots\} \rightarrow \mathbb{R}$ we say that $f = \mathcal{O}(g)$, if and only if there exists a constant $C > 0$, and n_0 such that $|f(n)| \leq C|g(n)|$, for all $n \geq n_0$. In particular, $f = \mathcal{O}(1)$ implies that f is bounded. Similarly, we say that $f = \Omega(g)$, if and only if there exists a constant $c > 0$, and n_0 such that $|f(n)| \geq c|g(n)|$, for all $n \geq n_0$. Finally, we say that $f = \Theta(g)$, if and only if $f = \mathcal{O}(g)$ and $f = \Omega(g)$.

We also write $f \ll g$ or $f = o(g)$ to denote that

$$\lim_{n \rightarrow \infty} \left| \frac{f(n)}{g(n)} \right| = 0.$$

Finally, we write $f \sim g$, if there exists a constant $c \in \mathbb{R} \setminus \{0\}$ such that

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = c.$$

1.5 About the Results.

All the results proved in this thesis are my own except the linear programming argument of Theorem 56, which is due to Colin Cooper and Martin Dyer.

Chapter 2

The Linear Voting Model

2.1 Introduction

In this chapter, we focus on voting models on finite graphs, in these processes vertices of a given graph have opinions, and by interacting with their neighbours, they change their 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, and interacting particles in physical systems.

To date, many models have been introduced and studied in the literature. Nevertheless, it is not always clear how such models relate to each other. In this chapter, we introduce the Linear Voting Model, a general voting model that aims to unify some of the existing approaches in a tractable way. This model subsumes several models proposed in the past, including, for example, the *push* model and the very popular *pull* model.

Our motivation in defining this model was to understand two techniques that were used to analyse synchronous pull voting on finite graphs, and to assess the possible applicability to other voting models. The aforementioned techniques are: i) A martingale argument that allows us to compute the probability that the final opinion of the system takes a particular value, ii) A time-reversal dual relation between pull vot-

ing and a system of coalescing random walks. These two techniques were introduced in the Computer Science community by Hassin and Peleg [52] in 2001, but their first appearance can be tracked down to the 1975 work of Holley and Liggett [53].

Let us give some intuition about the two techniques we mention above. Consider a graph $G = (V, E)$ with finite number of vertices. Moreover, consider synchronous pull voting, and assume that the vertices have initial opinion 0 or 1. In [52], the authors proved that the probability that opinion 0 wins is $d(A)/d(V)$, where A is the set of vertices with initial opinion 0, and $d(X)$ is the sum of the degrees of the vertices of $X \subseteq V$. This is proved by using the following proposition: Let A_t be the set of vertices with opinion 0 at time t , then $(d(A_t))_{t \geq 0}$ is a martingale. This fact, together with standard martingale methodology, gives a proof of the first statement.

Define the consensus time as the first time all vertices have the same opinion. Additionally, define a system of coalescing random walks on a graph G as follows. At each round, each particle performs one step of a random walk, and when two or more particles meet in one vertex, they coalesce (merge) into one, and they move together. The coalescing time is the first time all particles are together (and keep moving together). Alternatively, the coalescing time is the first time only one vertex is occupied by particles. The time-reversal duality between pull voting and coalescing random walks states that the distribution of the consensus time when all particles start with a different opinion, is equal to the distribution of the coalescing time. The name “time-reversal duality” comes from the fact that if we look at the voting process from time T to 0, we will see that its behaviour is similar to a set of coalescing random walks.

With these two concepts in mind, we design a general model that encodes different voting processes to which such ideas can be applied.

2.2 Definitions and Contributions.

Let V be a set of vertices with $|V| = n$. Define a configuration of opinions as a $n \times 1$ vector $\xi \in Q^V$, where $Q = \{0, 1\}$ for a two party model, or $Q \subseteq \{1, \dots, n\}$ 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. If no confusion arises, we will just write $\mathcal{M}$ instead of $\mathcal{M}(V)$.

Let l be a probability 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

$$\xi_t = \begin{cases} \xi, & \text{if } t = 0, \\ M_{t-1}\xi_{t-1}, & \text{if } t > 0, \end{cases} \quad (2.1)$$

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 $(\xi_t)_{t \geq 0} \sim \mathcal{LVM}(l, \xi)$. Note that $(\xi_t)_{t \geq 0}$ is a Markov chain on state space Q^V . The entry $\xi_t(v)$ represents the opinion of vertex v at round t . Consider $M \in \mathcal{M}$ and $\xi' = M\xi$. If all vertices have different opinions under ξ , 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 that at each round, each vertex adopts the opinion of only one vertex (including itself). This implies that vertices can only change their opinion to other opinions that currently exist in the system, and that they cannot create new ones. Examples of linear voting models, including pull voting (asynchronous or synchronous) and the push voting model, are given in Section 2.3.

We proceed to present our contributions. Theorem 1 of this chapter gives the probability that a particular opinion wins the poll. Our proof technique is based on a generalisation of the martingale argument used in [52]. We continue by giving the

necessary ingredients to enunciate our theorem. Let l be a probability distribution over $\mathcal{M}$, then define the mean matrix H of l as

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

From Lemma 11 below, it holds that H is the transition matrix of a Markov chain on state space V (not Q^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. Note that, if ξ_t is in consensus, so is ξ_{t+1} . Given $(\xi_t)_{t \geq 0} \sim \mathcal{LVM}(l, \xi)$, we define the consensus time τ_{cons} of $(\xi_t)_{t \geq 0}$ as

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

Observe that the consensus time is a stopping time with respect to $(\xi_t)_{t \geq 0}$, and that once the vertices reach consensus they never change their opinion again, thus the system reaches a final state. We have the following theorem about the winning probability.

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

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

The theorem above solves the winning probability problem under very reasonable conditions. Now, we focus on the consensus time problem. Theorem 2 below gives an upper bound to the expected consensus time. Our technique is qualitatively different from the duality approach used in the pull voting process, and is more similar to Levin et al. [57, Chapter 17] or Berenbrink et al. [12].

Consider the two party model, that is, $Q = \{0, 1\}$, and let S_t be the set of vertices with opinion 1 at the beginning of round t . Let $\mu(S) = \sum_{v \in S} \mu(v)$, where μ is the

¹a.s. stands for almost surely, that is, the event occurs with probability 1

stationary distribution of H and $S \subseteq V$. Additionally, consider the random variable $Z_t = \mu(S_{t+1}) - \mu(S_t)$ and $\rho_t = 2\mathbb{1}_{\{\mu(S_t) \leq \mu(S_t^c)\}} - 1$. Finally, define the quantity Υ as

$$\Upsilon = \min \left\{ \frac{\mathbb{E} \left(Z_t^2 \mathbb{1}_{\{\rho(S_t) Z(S_t) < 0\}} | S_t = S \right)}{\min\{\mu(S), \mu(S^c)\}} \right\}, \quad (2.2)$$

where the minimum is over all $S \subseteq V$ with $0 < \mu(S) < 1$. The definition of Υ is independent of t as the process is Markovian.

By using the above definitions, we prove the following theorem.

Theorem 2. *Let $(\xi_t)_{t \geq 0} \sim \mathcal{LVM}(l, \xi)$ be a linear voting model with an arbitrary number of initial opinions, and $\Upsilon > 0$. Then, there exists a constant $C > 0$ such that*

$$\mathbb{E}(\tau_{\text{cons}}) \leq C\Upsilon^{-1}, \quad (2.3)$$

for all initial configuration ξ .

Upper bounds on the consensus time obtained by Theorem 2 are given in Section 2.5.3 for the example models given in Section 2.3.

Finally, we extend the concept of duality used to analyse the pull voting process to our more general setting. For pull voting, this dual process corresponds to the so-called coalescing random walks (independent random walks that, on meeting, coalesce (merge) into one). It turns out that the dual process of a linear voting process is still a set of coalescing walks, but they move using a different set of rules.

Let l be a probability 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

$$f_t = \begin{cases} f, & \text{if } t = 0, \\ M_{t-1}^\top f_{t-1}, & \text{if } t > 0, \end{cases} \quad (2.4)$$

where M_t are independent and identically distributed (i.i.d.) matrices sampled from l , and $M^\top$ is the transpose of M . The process can be interpreted as follows. The vector $f_t \in \mathbb{R}^V$ counts the number of particles at each of the vertices of the graph at time t . The matrix $M_{t-1}^\top$ moves any particles on the vertices at the end of time $t-1$

(the result of such movement is observed at time t). In particular $M_{t-1}(v, u) = 1$ means that all the particles on vertex v at time $t - 1$ are in vertex u at time t . In this process, once two or more particles come together, they coalesce (merge), and they start moving together. Because of this, 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 constant over time. For practical purposes, the standard choice of the initial vector f is the all 1 vector $\mathbf{1}$, meaning that at the beginning of the process there is one particle on each vertex.

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

$$\{f_t \text{ is in agreement with respect to } \mathcal{S}\} \Leftrightarrow \{f_t(S_i) = n \text{ for some } i \in \{1, \dots, m\}\} \quad (2.5)$$

Given a vector of opinions 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 $S_a = \{v \in V : \xi(v) = a\}$. The next theorem states the relation between the linear voting process and the coalescing process.

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

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

A particular case of Theorem 3 is when each vertex has a different opinion in ξ . In such a case, agreement is reached at time t if and only if all particles are together at such a 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 have different opinions, we have that τ_{cons} has the same distribution as τ_{coalsc} .

Corollary 4. *Let V be a set of vertices and let l be a probability distribution over $\mathcal{M}$. Suppose that $(\xi_t)_{t \geq 0} \sim \mathcal{LVM}(l, \xi)$ where all the opinions of ξ are different, and $(f_t) \sim \mathcal{CP}(l, \mathbf{1})$. Then, for every $t \geq 0$,*

$$\mathbb{P}(\tau_{\text{cons}} \leq t) = \mathbb{P}(\tau_{\text{coalsc}} \leq t)$$

The rest of the chapter is structured as follows. In Section 2.3, we introduce the model and give some examples to aid intuition, and to demonstrate the flexibility of the model. In Section 2.4, we introduce the necessary notation and prove Theorem 1. In Section 2.5, we prove Theorem 2, and in Section 2.5.3, we use this theorem to find bounds for the expected consensus time of the example models of Section 2.3. Later, in Section 2.6, we state the dual relation between the linear voting model and the coalescing process. We finish the chapter computing the expected coalescing time of the example models on particular graph families.

Notation remainder. $G = (V, E)$ stands for a simple graph. We assume the cardinality of V is $|V| = n$, and the cardinality of E is m . For $v \in V$, we denote the neighbourhood of v by $N(v)$, and its degree by $d(v) = |N(v)|$. 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. The letter l usually denotes a probability distribution over matrices in $\mathcal{M}$. $M^\top$ denotes the transpose of the matrix M .

2.3 The Model

Recall the definition of a linear voting model. Given a probability distribution l over the matrices $\mathcal{M}$, and $\xi \in Q^V$, we say that $(\xi_t)_{t \geq 0} \sim \mathcal{LVM}(l, \xi)$, 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 .

Remark 5. *Since this is a discrete-time process, we usually refer to the time step as*

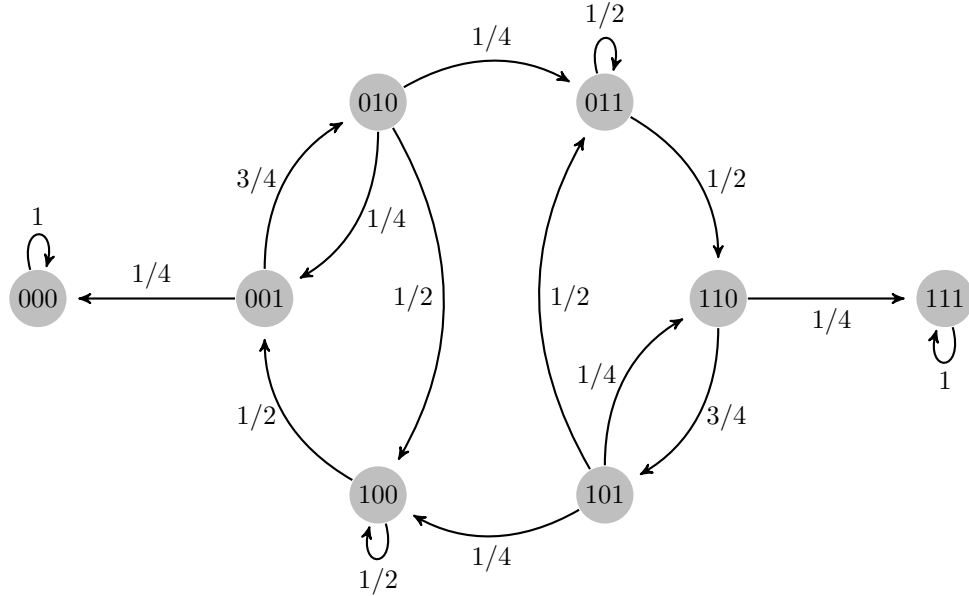
time-steps, steps or rounds. For example, if we talk about the configuration at round t , we mean ξ_t . The initial round is 0.

We illustrate the definition of the linear voting model with the following example.

Example 6. Suppose that we have vertices $V = \{1, 2, 3\}$, and that the opinions can be either 0 or 1. We consider a measure $l \in \Pi$ with probability mass on the following matrices

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix},$$

where $l(A_1) = l(A_2) = 1/4$ and $l(A_3) = 1/2$. Note the configurations are vectors of 0s and 1s, and that we have a total of eight configurations. Matrix A_1 exchanges the opinions of vertex 2 and 3, A_2 changes the opinion of vertex 3 to the opinion of 2, and A_3 rotates the opinions of all the vertices. The transition diagram of this Linear Voting process (which is a Markov chain on $\{0, 1\}^V$) can be plotted as follows



While the example above shows a toy example of a linear voting model, non-trivial examples can be constructed given a baseline structure where vertices interact. In particular, given a connected graph G , we define the following models.

- a) **Synchronous pull model.** At each round, each vertex samples a random neighbour and pulls (adopts) its opinion.
- b) **Asynchronous pull model.** At each round, one vertex v is selected uniformly at random. Then, this vertex pulls (adopts) the opinion of a random neighbour.
- c) **Asynchronous push model.** At each round, one vertex v is selected uniformly at random. Then, this vertex push (forces) its opinion on a random neighbour.
- d) **Abusive pushing model.** At each round, one vertex v is selected uniformly at random. Then, this vertex pushes (forces) its opinion to its whole neighbourhood.
- e) **Oblivious model.** At each round, one edge $e = (v, u)$ is selected uniformly at random. Then, v pulls the opinion of u with probability $1/2$, otherwise, u pulls the opinion of v .

Remark 7. *To be precise, the change in the opinions happens at the end of a round t , prior to round $t + 1$. In particular if v gets its opinion from w at round t , it means that at round $t + 1$, vertex v has the opinion of w at round t (either v pulls the opinion of w , or w pushes its opinion on v). Remember that the first round is $t = 0$.*

Remark 8. *There is no natural definition of a synchronous push model. This is because several vertices, with potentially different opinions, can push their opinion on a single vertex. On the other hand, it is possible to define a synchronous version of the oblivious model, in this case we can choose a random matching of the graph, instead of a single edge.*

Remark 9. *As noticed in [33], on regular graphs, the asynchronous pull, asynchronous push and oblivious models are all equivalent.*

Lemma 10. *The five models defined above are linear voting models.*

Proof. We give a proof for models a) and b), as for the other models the proof is similar. Let ξ_t be the configuration of opinions at round t . In synchronous pull voting, at each round, each vertex v pulls the opinion of a random neighbour $w(v)$. Let ξ_{t+1} be 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 other entries. It is straightforward to check that $\xi_{t+1}(v) = (M\xi_t)(v) = M(v, w(v))\xi_t(w(v)) = \xi_t(w(v))$. Also, by definition, 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 , which is randomly selected, pulls the opinion of a random neighbour $w(v)$, 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 row of v). It is not hard to check that the random matrix M mimics the asynchronous pull model, that is, $\xi_{t+1} = M\xi_t$, and that $M \in \mathcal{M}$. $\square$

2.3.1 Mean Matrix

Let l be probability distribution over $\mathcal{M}$. We define the mean matrix of l as

$$H = H(l) = \sum_{M \in \mathcal{M}} l(M)M. \quad (2.6)$$

Observe that in our examples, the models were described by rules (algorithms) rather than by an explicit distribution l . In practice, we expect that most of the models are described in this way. While this representation is useful as it gives more intuition than a long list of matrices with their respective probabilities, it is not clear how to compute the mean matrix $H(l)$, as we do not have an explicit form for the distribution l . The following lemma aims to provide a way of computing H without exhibiting l explicitly.

Lemma 11. *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 $v, w \in V$, it*

holds that

$$H(v, w) = \mathbb{P}(v \text{ gets its opinion from } w \text{ at round } t). \quad (2.7)$$

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

$$\mathbb{E}(\xi_{t+1}|\xi_t) = \sum_{M \in \mathcal{M}} l(M)(M\xi_t) = \left(\sum_{M \in \mathcal{M}} l(M)M \right) \xi_t = H\xi_t. \quad (2.8)$$

Choose $\xi_t = \xi$, such that the opinion of $\xi(w) = 1$, and $\xi(u) = 0$ for all $u \neq w$. Then, the event $\{v \text{ gets its opinion from } w \text{ at round } t\}$ is equal to $\{\xi_{t+1}(v) = 1\}$. Thus, from Equation (2.8)

$$\begin{aligned} \mathbb{P}(\xi_{t+1}(v) = 1 | \xi_t = \xi) &= \mathbb{E}(\xi_{t+1}(v) | \xi_t = \xi) = (H\xi)(v) = \sum_{z \in V} H(v, z)\xi(z) \\ &= H(v, w). \end{aligned}$$

□

We proceed to compute the mean matrix H for the five examples given before. Consider a connected graph G with n vertices and m edges, and let P be the transition matrix of a simple random walk on G , which is given by $P(v, w) = 1/d(v)$ if v and w are adjacent, 0 otherwise. Let A be the adjacency matrix of G , which is defined by $A(v, w) = 1$ if $v \sim w$, 0 otherwise. Let $L = D - A$ be the combinatorial Laplacian, where D is the diagonal matrix containing the degree sequence of G . Also, let I denote the $n \times n$ identity matrix. Moreover, let F be the diagonal matrix defined by $F(v, v) = \sum_{w: w \sim v} 1/d(w)$.

Theorem 12. *The mean matrix of the synchronous pull, asynchronous pull, push, abusive push, oblivious models are, respectively, $H_a = P$ and*

$$H_b = \frac{n-1}{n}I + \frac{1}{n}P, \quad H_c = I + \frac{1}{n}P^\top - \frac{1}{n}F, \quad H_d = I - \frac{1}{n}L, \quad H_e = I - \frac{1}{2m}L.$$

Proof Sketch. We compute H_a . From Equation (2.7), we have that $H_a(v, w)$ is the probability that v gets its opinion from w . In synchronous pull voting, this happens only if the random neighbour selected by 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 voting, a randomly selected vertex v pulls the opinion of a random neighbour w . Observe that for a vertex u we have $H_b(u, u)$ is the probability that u gets its opinion from u , i.e., the probability that u does not change the opinion. Such an event happens 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 to pull (probability $1/n$), and then v has to pull the opinion of w (with probability $1/d(v)$). We conclude that $H_b = ((n - 1)/n)I + (1/n)P$. The other cases are similar.

For H_c , remember that in asynchronous push model, a randomly selected vertex w pushes its opinion on a random neighbour v . Given two adjacent vertices $v, w \in V$, $H_c(v, w)$ is the probability that w pushes its opinion on v . Then

$$H_c(v, w) = \left(\frac{1}{n}\right) \left(\frac{1}{d(w)}\right) = \frac{1}{n} P^\top(v, w).$$

In the first equality, the first term corresponds to the probability that we choose w , and the second that w pushes on v . Then, as H is a transition matrix,

$$H_c(v, v) = 1 - \sum_{w:w \sim v} \frac{1}{d(w)n} = I(v, v) - \frac{1}{n} F(v, v).$$

We conclude

$$H_c = I + \frac{1}{n} P^\top - \frac{1}{n} F.$$

The matrices H_d and H_e are computed similarly.

□

2.4 Winning probability.

The most natural question in any voting model is: Who wins? In order to answer this question, we use a martingale argument. For now, assume the two-party model,

that is, $Q = \{0, 1\}$. Later, we will extend the result to the n -party model. Since the mean matrix H of a linear voting model is a transition matrix, then all its eigenvalues λ have absolute value less or equal than 1, and there is at least one eigenvalue equals to 1. Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of H , and assume that $\lambda_1 = 1$. 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 by $\langle f, g \rangle = \sum_{v \in V} f(v)g(v)$ the standard dot product. In this section, we interpret Q as subset of the real numbers, 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 13. *Suppose $\lambda \neq 0$ is a real eigenvalue of H with eigenvector f , then the process $(\langle f, \xi_t \rangle / \lambda^t)_{t \geq 0}$ is a martingale with respect to $(\xi_t)_{t \geq 0}$.*

Proof. As $\langle f, \xi_t \rangle$ is bounded for each $t \geq 0$, it suffices to show that $\mathbb{E}(\langle f, \xi_{t+1} \rangle | \xi_t) = \lambda \langle f, \xi_t \rangle$, and then to divide both sides by λ^{t+1} . By linearity of (conditional) expectation and by Equation (2.8), we have

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

□

We proceed to give the ingredients for the proof of Theorem 1. Assume that H , as a transition matrix, 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 $\sum_{v \in V} \mu(v) = 1$. Due to its importance in this work, we denote by $m_t = \langle \mu, \xi_t \rangle$ the martingale associated with the eigenvalue 1. We call this martingale the *voting martingale*.

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 $(\xi_t)_{t \geq 0}$ as

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

Observe that the consensus time is a stopping time with respect to the filtration given by $(\xi_t)_{t \geq 0}$. Note that once the vertices reach consensus, they never change their opinion again, and the system reaches a final state. We say that opinion q wins if q is the final opinion of the system. Here, we restate and prove Theorem 1.

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

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

Proof. Denote by $\mathbf{1}$ and $\mathbf{0}$ the vectors where all components are 1 and 0, respectively. Due to the fact that $\tau_{\text{cons}} < \infty$, $(\xi_t)_{t \geq 0}$ always reaches consensus, it converges to either $\mathbf{1}$ or $\mathbf{0}$, and hence $(m_t)_{t \geq 0}$ converges to 1 or 0 as t tends to infinity. Additionally, observe that, $0 \leq m_t = \sum_{v \in V} \mu(v) \xi_t(v) \leq 1$ 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, allow us to apply the optional stopping theorem (Theorem 86 in Appendix A) 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 $\mathbb{P}(\xi_{\tau_{\text{cons}}} = \mathbf{1} | \xi_0 = \xi) = m_0 = \langle \mu, \xi \rangle$, from which we conclude

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

□

Corollary 14. *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. Then apply Theorem 1. $\square$

We illustrate the use of Theorem 1 with the following example.

Example 15. Recall Example 6. In this case, we have $V = \{1, 2, 3\}$ and the opinion can be 0 or 1. We defined the probability measure l on the following matrices

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix},$$

where $l(A_1) = l(A_2) = 1/4$ and $l(A_3) = 1/2$. Therefore, the mean matrix H is

$$H = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 0 & 1/4 & 3/4 \\ 1/2 & 1/2 & 0 \end{pmatrix}.$$

In this case, H is the transition matrix of an ergodic Markov chain, whose stationary distribution is given by the vector $(3/10, 4/10, 3/10)$. Then, for the initial configuration of opinions $(0, 0, 1)$, the final opinion is 0 with probability $7/10$. For comparison purposes, in one hand, if we work with the Markov chain with all combinations of opinions as states (i.e., $2^{|V|}$ states), then we need to solve a system of $2^{|V|} \times 2^{|V|}$ linear equations to find the probability that a particular opinion wins. On the other hand, to compute the stationary distribution of the mean matrix H , we need to solve a $|V| \times |V|$ system of linear equations. Of course, we also need to be able to compute the mean matrix H .

Theorem 16. Let G be a connected graph, and let S 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. Then, for synchronous pull, asynchronous pull, push, abusive push and oblivious model, the associated μ and p are given by

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

(b) *Asynchronous pull model*: $\mu_b(v) = d(v)/d(V)$, $p_b = d(S)/d(V)$

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

(d) *Abusive pushing model*: $\mu_d(v) = 1/n$, $p_d = |S|/n$

(e) *Oblivious model*: $\mu_e(v) = 1/n$, $p_e = |S|/n$.

Proof. We apply Theorem 1. For that, we need to find the stationary distribution of the models above. Recall that for synchronous pull, the mean matrix is P , the transition matrix of a random walk on G . Then, it is well-known, that $\mu_a(v) = d(v)/d(V)$. For asynchronous pull, the mean matrix is given by $(n-1)/nI + (1/n)P$, which has the same eigenvectors as P , 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 $F = F^\top$, we have

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

thus μ' is the stationary distribution of the mean matrix of the push model. For the abusive pushing model, observe that $H_d = I - (1/n)L$ is a symmetric matrix, then its stationary distribution is uniform. H_e is also symmetric, then its stationary distribution is also uniform. $\square$

Remark 17. Under the assumption $\tau_{\text{cons}} < \infty$, Theorem 2 gives a satisfactory answer to the question of who wins. Nevertheless to check the condition $\tau_{\text{cons}} < \infty$ usually requires knowledge of the particular model in question. All the defined models reach consensus, except synchronous pull which requires the underlying graph to be non-bipartite [52].

2.5 Consensus Time.

2.5.1 Introduction

Let $(\xi_t)_{t \geq 0} \sim \mathcal{LVM}(l, \xi)$ be a linear voting model with $Q = \{0, 1\}$. 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 2.4.

In this section, we use the following notation. Let S_t be the set of vertices with opinion 1 at the beginning of round t , and denote by $\mu(S) = \sum_{v \in S} \mu(v)$ the stationary measure of the set $S \subseteq V$. Observe that $\mu(S_t) = m_t$.

Given $S \subseteq V$, define $\xi_S \in \{0, 1\}^V$ as the indicator vector of the set S . Additionally, define the random variable $Z(S)$ as

$$Z(S) = \langle \mu, M\xi_S \rangle - \langle \mu, \xi_S \rangle$$

where M is a random matrix with distribution l . The random variable $Z(S)$ represents the one step difference of the voting process. That is $Z(S) = \mu(S_{t+1}) - \mu(S_t)$ given that $S_t = S$, and since $(S_t)_{t \geq 0}$ is a Markov chain, the distribution of $Z(S)$ does not depend on t .

Additionally, define the set function $\rho : 2^V \rightarrow \mathbb{R}$ by

$$\rho(S) = 2\mathbb{1}_{\{\mu(S) \leq \mu(S^c)\}} - 1.$$

Note that ρ takes values -1 or 1 , and that $\rho(S) = -\rho(S^c)$. We also claim that $Z(S) = -Z(S^c)$. The proof of the claim is as follows. Denote by $\xi = \xi_S$ and $\xi^c = \xi_{S^c}$ the configurations of opinions represented by the sets S and S^c , respectively. Notice that $\xi + \xi^c = \mathbf{1}$, the vector where all components are 1. Let M be a random matrix with distribution l , then

$$\begin{aligned} Z(S) &= \langle \mu, M\xi \rangle - \langle \mu, \xi \rangle \\ &= \langle \mu, M(\mathbf{1} - \xi^c) \rangle - \langle \mu, (\mathbf{1} - \xi^c) \rangle \\ &= \langle \mu, \mathbf{1} \rangle - \langle \mu, M\xi^c \rangle - \langle \mu, \mathbf{1} \rangle + \langle \mu, \xi^c \rangle = -Z(S^c) \end{aligned}$$

For $S \subseteq V$ with $0 < \mu(S) < 1$, define the quantity $\Upsilon(S)$ by

$$\Upsilon(S) = \frac{\mathbb{E} \left(Z(S)^2 \mathbf{1}_{\{\rho(S)Z(S) < 0\}} \right)}{\min\{\mu(S), \mu(S^c)\}}, \quad (2.9)$$

and Υ by

$$\Upsilon = \min_S \Upsilon(S). \quad (2.10)$$

where the minimum is over all $S \subseteq V$ such that $0 < \mu(S) < 1$. Observe that, since $Z(S) = -Z(S^c)$ and $\rho(S) = -\rho(S^c)$, then $\Upsilon(S) = \Upsilon(S^c)$. Then, for $0 < \mu(S) \leq 1/2$,

$$\Upsilon(S) = \frac{\mathbb{E} \left(Z(S)^2 \mathbf{1}_{\{Z(S) < 0\}} \right)}{\mu(S)},$$

and by the symmetry $\Upsilon(S) = \Upsilon(S^c)$. Then, we have

$$\Upsilon = \min\{\Upsilon(S) : 0 < \mu(S) \leq 1/2\}, \quad (2.11)$$

which is useful for applications.

The main objective of this section is to prove Theorem 2, which is restated here.

Theorem 2. *Let $(\xi_t)_{t \geq 0} \sim \mathcal{LVM}(l, \xi)$ be a linear voting model with arbitrary number of initial opinions, and $\Upsilon > 0$. Then, there exists a constant $C > 0$ such that*

$$\mathbb{E}(\tau_{\text{cons}}) \leq C\Upsilon^{-1}, \quad (2.12)$$

for all initial configurations ξ .

In section 2.5.3, we provide upper bounds on the expected consensus time for the models defined in Section 2.3.

2.5.2 Proof of Theorem 2

Our proof is based on a potential function approach. Let $\eta(S) = \min\{\mu(S), \mu(S^c)\}$, where $\mu(S^c) = 1 - \mu(S)$. Denote by $(\eta_t)_{t \geq 0}$ the process $(\eta(S_t))_{t \geq 0}$, and let $(Z_t)_{t \geq 0} =$

$(Z(S_t))_{t \geq 0}$. 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 $\eta_{t+1} = \min\{\mu(S_t) + Z_t, \mu(S_t^c) - Z_t\}$. Also note that if $\eta_t = \mu(S_t)$, i.e., $\mu(S_t) \leq \mu(S_t^c)$, then

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

If $\eta_t = \mu(S_t^c)$, the same applies by observing that $\mu(S_{t+1}^c) - \mu(S_t^c) = -Z(S_t)$, and

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

In both cases we get

$$\eta_{t+1} \leq \eta_t + \rho_t Z_t, \quad (2.13)$$

where $\rho_t = \rho(S_t) = 2\mathbb{1}_{\{\mu(S_t) \leq \mu(S_t^c)\}} - 1$.

With these ingredients, we proceed to prove a technical lemma, which is fundamental for the proof of Theorem 2.

Lemma 18. *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*

$$\mathbb{P}(\tau_{\text{cons}} > T) \leq 1/2,$$

for all $T \geq \lfloor 32\eta_0/\Upsilon \rfloor$.

Proof.

Let $S \subseteq V$ but $S \neq \emptyset$ and $S \neq V$. By setting $S_t = S$ in Equation (2.13), 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 (2.13), we have $\rho_t Z_t \geq \eta_{t+1} - \eta_t \geq -\eta_t$. By taking expectations, and noting that $\eta_t = \eta(S) > 0$,

$$\begin{aligned} \mathbb{E}(\sqrt{\eta_{t+1}} | S_t = S) &\leq \sqrt{\eta(S)} \mathbb{E} \left(\sqrt{1 + \frac{\rho_t Z_t}{\eta_t}} \middle| S_t = S \right) \\ &= \sqrt{\eta(S)} \mathbb{E} \left(\left(\sqrt{1 + \frac{\rho_t Z_t}{\eta_t}} \right) \mathbb{1}_{\{\rho_t Z_t \geq 0\}} \middle| S_t = S \right) \end{aligned} \quad (2.14)$$

$$+ \sqrt{\eta(S)} \mathbb{E} \left(\left(\sqrt{1 + \frac{\rho_t Z_t}{\eta_t}} \right) \mathbb{1}_{\{\rho_t Z_t < 0\}} \middle| S_t = S \right). \quad (2.15)$$

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

$$\sqrt{1+x} \leq 1 + \frac{x}{2}, \quad (2.16)$$

and

$$\sqrt{1+x} \leq 1 + \frac{x}{2} - \frac{x^2}{8} + \frac{x^3}{16}. \quad (2.17)$$

To upper bound equation (2.14), we use Equation (2.16), and for equation (2.15), we use (2.17). Recall that, in view of the fact that $(\mu(S_t))_{t \geq 0}$ is a martingale, $\mathbb{E}(Z_t | S_t = S) = 0$. Then, after some rearrangement, we obtain

$$\begin{aligned} \mathbb{E}(\sqrt{\eta_{t+1}} | S_t = S) &\leq \sqrt{\eta(S)} - \sqrt{\eta(S)} \mathbb{E} \left(\left(\frac{(\rho_t Z_t)^2}{8\eta_t^2} - \frac{(\rho_t Z_t)^3}{16\eta_t^3} \right) \mathbb{1}_{\{\rho_t Z_t < 0\}} \middle| S_t = S \right) \\ &\leq \sqrt{\eta(S)} - \sqrt{\eta(S)} \mathbb{E} \left(\frac{Z_t^2}{8\eta_t^2} \mathbb{1}_{\{\rho_t Z_t < 0\}} \middle| S_t = S \right) \end{aligned} \quad (2.18)$$

$$\begin{aligned} &= \sqrt{\eta(S)} - \left(\frac{1}{8\eta(S)^{1/2}} \right) \left(\frac{\mathbb{E}(Z(S)^2 \mathbb{1}_{\{\rho(S)Z(S) < 0\}})}{\min(\mu(S), \mu(S^c))} \right) \\ &\leq \sqrt{\eta(S)} - \frac{\Upsilon}{8\eta(S)^{1/2}} \end{aligned} \quad (2.19)$$

In the second inequality, we used the fact that we are working in the event $\{\rho_t Z_t < 0\}$, and for the last inequality, we used the definition of Υ in Equation (2.10). If $S_t \notin \{\emptyset, V\}$, then the system is not in consensus and $\tau_{\text{cons}} > t$, thus

$$\begin{aligned} \mathbb{E}(\sqrt{\eta_{t+1}}) &= \sum_{S \subseteq V} \mathbb{E}(\sqrt{\eta_{t+1}} | S_t = S) \mathbb{P}(S_t = S) = \sum_{S: S \neq \emptyset, V} \mathbb{E}(\sqrt{\eta_{t+1}} | S_t = S) \mathbb{P}(S_t = S) \\ &\leq \sum_{S: S \neq \emptyset, V} \left(\sqrt{\eta(S)} - \frac{\Upsilon}{8\eta(S)^{1/2}} \right) \mathbb{P}(S_t = S) \end{aligned} \quad (2.20)$$

$$\begin{aligned} &= \mathbb{E}(\sqrt{\eta_t}) - \sum_{S: S \neq \emptyset, V} \left(\frac{\Upsilon}{8\eta(S)^{1/2}} \right) \mathbb{P}(S_t = S | \tau_{\text{cons}} > t) \mathbb{P}(\tau_{\text{cons}} > t) \\ &= \mathbb{E}(\sqrt{\eta_t}) - \frac{\Upsilon}{8} \mathbb{E} \left(\frac{1}{\sqrt{\eta_t}} \middle| \tau_{\text{cons}} > t \right) \mathbb{P}(\tau_{\text{cons}} > t), \end{aligned} \quad (2.21)$$

where (2.20) follows using Equation (2.19). As $1/x$ is convex for $x > 0$, we apply Jensen's inequality to the random variable $x = \sqrt{\eta_t}$, to obtain

$$\mathbb{E} \left(\frac{1}{\sqrt{\eta_t}} \middle| \tau_{\text{cons}} > t \right) \geq \frac{1}{\mathbb{E}(\sqrt{\eta_t} | \tau_{\text{cons}} > t)} = \frac{\mathbb{P}(\tau_{\text{cons}} > t)}{\mathbb{E}(\sqrt{\eta_t})}. \quad (2.22)$$

The last equality is due to the fact that the event $\{\tau_{\text{cons}} \leq t\}$ implies that the vertices reached consensus, and thus $S_t = \emptyset$ or $S_t = V$, and hence $\eta_t = 0$. 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 Equation (2.22) into (2.21), 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 $\eta_t \in [0, 1/2]$,

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

and by Equation (2.18), we have that $\mathbb{E}(\sqrt{\eta_t}) \leq \sqrt{\eta_0}$. Therefore

$$\frac{\Upsilon}{8} \mathbb{P}(\tau_{\text{cons}} > t)^2 \leq \sqrt{\eta_0}(\mathbb{E}(\sqrt{\eta_t}) - \mathbb{E}(\sqrt{\eta_{t+1}})).$$

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

$$\frac{\Upsilon}{8} \sum_{t=0}^{T-1} \mathbb{P}(\tau_{\text{cons}} > t)^2 \leq \sqrt{\eta_0}(\sqrt{\eta_0} - \mathbb{E}(\sqrt{\eta_T})) \leq \eta_0. \quad (2.23)$$

Observe that since $\mathbb{P}(\tau_{\text{cons}} > t)$ is non-increasing, then $\lim_{t \rightarrow \infty} \mathbb{P}(\tau_{\text{cons}} > t)$ exists. Suppose such a limit is not equal to 0, then by taking T large enough in both sides of Equation (2.23), we obtain a contradiction. Indeed, as $\Upsilon > 0$, the left-hand side would tend to infinity. We conclude that $\lim_{t \rightarrow \infty} \mathbb{P}(\tau_{\text{cons}} > t) = 0$.

Let T be the defined as

$$T = \min\{t \geq 0 : \mathbb{P}(\tau_{\text{cons}} > t) < 1/2\}. \quad (2.24)$$

From the fact that $\lim_{t \rightarrow \infty} \mathbb{P}(\tau_{\text{cons}} > t) = 0$, we deduce $T < \infty$. From the definition of T , the following upper bound follows from Equation (2.23),

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

hence $T \leq \lfloor 32\eta_0/\Upsilon \rfloor$. □

A straightforward application of Lemma 18 gives us the result of Theorem 2 for the two-party model, i.e., $Q = \{0, 1\}$. This is established in the following corollary.

Corollary 19. *Let $(\xi_t)_{t \geq 0} \sim \mathcal{LVM}(l, \xi)$ with $\xi \in Q^V$ where $Q = \{0, 1\}$, be a linear voting model with $\Upsilon > 0$, then*

$$\mathbb{E}(\tau_{\text{cons}}) \leq 64\Upsilon^{-1}. \quad (2.25)$$

Proof. From Lemma 18, we deduce that for time

$$T = \lfloor 32/\Upsilon \rfloor,$$

it holds $\mathbb{P}(\tau_{\text{cons}} \leq T) \geq 1/2$ independently of the initial opinion of the vertices. Hence, we assume the worst initial configuration of opinions. 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 (from the worst possible configuration), and look again after T steps until we reach consensus. As the probability that the process does not finish in T steps is at most $1/2$, we conclude that

$$\mathbb{E}(\tau_{\text{cons}}) \leq \sum_{k=1}^{\infty} kT \left(\frac{1}{2}\right)^k \leq 2T \leq 64\Upsilon^{-1}.$$

□

The full proof of Theorem 2, i.e., $Q = \{1, \dots, n\}$, is not as direct as Corollary 19. Indeed, it is much more complicated. The proof technique used in the proof of Theorem 2 below is motivated by Lemma 2.3 in [12]. In their proof, the authors show how to extend a bound for a two-party synchronous pull voting to several parties. We extend their argument to our more general setting.

Proof of Theorem 2.

Remember that we have n vertices taking opinions in the set $[n] = \{1, \dots, n\}$. We denote by $S_t^i \subseteq V$ the set of vertices with opinion i at time t , and by J_t the number of opinions in the system at time t , i.e., $J_t = |\{i \in [n] : |S_t^i| > 0\}|$. Clearly, $J_t \leq n$ since there are no more opinions than vertices. Moreover, $(J_t)_{t \geq 0}$ is non-increasing because when an opinion vanishes from the system, it does not return. For a configuration of opinions ξ , we define $J(\xi)$ as the number of different opinions in ξ .

Consider $q = \lfloor \log(n) / \log(6/5) \rfloor$, and define the following sets,

$$A_i = \{\xi \in [n]^V : (5/6)^{i+1}n < J(\xi) \leq (5/6)^i n\} \quad (2.26)$$

for $i = 0, \dots, q$. Observe that $(5/6)^{q+1} \leq 1/n$, and therefore

$$A_q = \{\xi \in [n]^V : 1 \leq J(\xi) \leq (5/6)^q n\}.$$

Clearly, $[n]^V = \bigcup_{i=0}^q A_i$ and the sets A_i are disjoint. Moreover, observe when ξ_t leaves A_i , it has to move to a configuration in A_j with $j > i$. This is because vanished opinions never appear again. Then, it is natural to consider the number of steps the process spends in states in A_i . Denote by $\tau_i = |\{t \geq 0 : \xi_t \in A_i\}|$ the total number of rounds the configurations of $(\xi_t)_{t \geq 0}$ belong to A_i . Then, after $\tau_q + \tau_{q-1} + \dots + \tau_i$ rounds, the configurations of opinions belong to some A_j with $j > i$. Observe that from the definition of q , it holds $6/5 \geq n(5/6)^q \geq 1$, thus

$$3.5831808 = \left(\frac{6}{5}\right)^7 \geq n \left(\frac{5}{6}\right)^{q-6} \geq \left(\frac{6}{5}\right)^6 = 2.985984 \geq \left(\frac{6}{5}\right)^5 > 2,$$

which allows us to conclude that $A_{q-6} = \{\xi \in [n]^n : J(\xi) = 3\}$, and that A_{q-5} contains configurations with at most two opinions. Therefore,

$$\sum_{i=0}^{q-5} \tau_i,$$

is the time needed to have one opinion in the system, i.e., to reach consensus. We claim that

$$\mathbb{E}(\tau_i) \leq 384(6/5)^i / (n\Upsilon) \quad (2.27)$$

for $i \leq q - 5$, and thus the expected time to reach consensus is

$$\sum_{i=0}^{q-5} \tau_i \leq \frac{384}{n\Upsilon} \sum_{i=0}^{q-5} \left(\frac{6}{5}\right)^i = \frac{384}{n\Upsilon} \left(5 \left(\frac{6}{5}\right)^{q-4} - 1\right) \leq \frac{C'}{\Upsilon}, \quad (2.28)$$

for a constant $C' > 0$.

In order to complete our proof, we need to check Equation (2.27). Fix $i \leq q - 5$, and suppose that at time T , the process is in configuration $\xi_T \in A_i$. We will prove later that for $T' = \lfloor 96(6/5)^i/(n\Upsilon) \rfloor$ the probability that $\xi_{T+T'} \in A_i$ is at most $3/4$. Indeed, we will prove that for any $\xi \in A_i$

$$\mathbb{P}(\xi_{T+T'} \in A_i | \xi_T = \xi) = \mathbb{P}(\xi_{T'} \in A_i | \xi_0 = \xi) \leq 3/4. \quad (2.29)$$

For now, assume that the inequality of Equation (2.29) holds true. The inequality above suggests to look at the process every T' steps. Denote by R the first value such that $\xi_{T+RT'} \notin A_i$. From Equation (2.29), we obtain that $\mathbb{P}(R > j) \leq (3/4)^j$, therefore $\mathbb{E}(R) \leq 4$, and thus the expected number of steps to leave A_i satisfies

$$\mathbb{E}(\tau_i) \leq 4T' \leq 384(6/5)^i/(n\Upsilon).$$

We proceed to prove Equation (2.29). The equality follows from the Markov property of the process $(\xi_t)_{t \geq 0}$ (strong Markov property if T is a stopping time). We just need to show that for any $\xi \in A_i$, $\mathbb{P}(\xi_{T'} \in A_i | \xi_0 = \xi) \leq 3/4$ holds true.

Let $i \leq q - 5$, let $\xi_0 = \xi \in A_i$, and let $k = (5/6)^i n$. Consider the following partitions of the opinions presented in ξ .

$$L = \{i \in [n] : 0 < \mu(S_0^i) \leq 3/k\}, \quad \text{and} \quad L^c = \{i \in [n] : \mu(S_0^i) > 3/k\}.$$

Denote by $J = |L| + |L^c|$, the total number of opinions in ξ . By definition of A_i , we have $5k/6 < J \leq k$. We claim that

$$|L| \geq J - k/3 \geq 5k/6 - k/3 = k/2 > 0.$$

To prove the claim, assume the opposite holds, i.e., $|L| < J - k/3$, then

$$1 \geq \sum_{i \in L^c} \mu(S_0^i) > |L^c|(3/k) = (J - |L|)(3/k) > (k/3)(3/k) = 1,$$

which is a contradiction, therefore, we conclude $|L| \geq J - k/3$.

Set $T' = \lfloor 96(6/5)^i/(n\Upsilon) \rfloor = \lfloor 96/(k\Upsilon) \rfloor$. Let Z_i be the indicator variable that takes value 1 if opinion i vanishes after T' steps (i.e., $|S_{T'}^i| = 0$), or if it wins the voting (i.e., $|S_{T'}^i| = n$), and value 0 otherwise.

We compute $\mathbb{P}(Z_i = 1 | \xi_0 = \xi)$ for $i \in L$. To that end, we replace all opinions different from i by 0, and we compute the probability that opinion i vanishes in T' extra steps. It is not hard to see that the later modification does not affect the probability of $\{Z_i = 1\}$, since we do not care about the fate of the other opinions. As $i \in L$, we have $\mu(S_0^i) \leq 3/k$. From Lemma 18, we have that by time $T' = \lfloor 32\mu(S_0^i)/\Upsilon \rfloor \leq \lfloor 96/(k\Upsilon) \rfloor$, consensus is reached in this two-party model with probability at least $1/2$, therefore $\mathbb{P}(Z_i = 1 | \xi_0 = \xi) \geq 1/2$. Hence

$$\mathbb{E} \left(\sum_{i \in L} Z_i \middle| \xi_0 = \xi \right) \geq |L|/2 \geq J/2 - k/6.$$

Therefore, by Markov's inequality,

$$\begin{aligned} \mathbb{P} \left(\sum_{i \in L} Z_i < L/3 \middle| \xi_0 = \xi \right) &= \mathbb{P} \left(L - \sum_{i \in L} Z_i \geq 2L/3 \middle| \xi_0 = \xi \right) \\ &\leq \frac{L - \mathbb{E}(\sum_{i \in L} Z_i | \xi_0 = \xi)}{(2L/3)} \\ &\leq \frac{L - L/2}{2L/3} \leq 3/4. \end{aligned} \tag{2.30}$$

Observe that if $\sum_{i \in L} Z_i > L/3$, then either one opinion wins the voting (hence there is one opinion in the system) or at least $L/3$ opinions vanish. Note that $L/3 \geq (1/3)((5/6)k - k/3) \geq (1/6)k$, thus $k - L/3 \leq (5/6)k$, as $L \geq k/2$.

If the event $\{\sum_{i \in L} Z_i > L/3\}$ holds then the number of opinions in the system at time T' , $J_{T'}$, satisfies

$$J_{T'} \leq \max \left\{ 1, J - \sum_{i \in L} Z_i \right\} \leq \max \left\{ 1, k - \frac{L}{3} \right\} \leq \max \left\{ 1, \frac{5k}{6} \right\} = 5k/6$$

From the above, and Equation (2.30), we conclude

$$\begin{aligned} \mathbb{P}(\xi_{T'} \in A_i | \xi_0 = \xi) &= \mathbb{P}(J_{T'} > (5/6)k | \xi_0 = \xi) \\ &\leq \mathbb{P} \left(\sum_{i \in L} Z_i \leq L/3 \middle| \xi_0 = \xi \right) \leq 3/4, \end{aligned} \tag{2.31}$$

thus proving Equation (2.29). □

2.5.3 Consensus Time Examples

We apply the previous theorems to our examples. We assume the two party model and use the same notation as in the last section. For our computations, we need to use some extra notation. Given $S \subseteq V$, denote by $E(S : S^c)$ the number of edges going from S to S^c , and by $d_S(v)$ the number of vertices of S adjacent to v . Observe that $E(S : S^c) = \sum_{v \in S} d_{S^c}(v) = \sum_{v \in S^c} d_S(v)$. For a given function $f : V \rightarrow \mathbb{R}^+$, we define $\mathcal{H}_f = \{S \subseteq V : 0 < f(S) \leq f(V)/2\}$, where $f(S) = \sum_{x \in S} f(x)$. For the function $\mathbf{1}(v) = 1$, we write $\mathcal{H}$ instead of $\mathcal{H}_1$. We denote the conductance of G by $\Phi(G) = \min_{S \in \mathcal{H}_d} \left\{ \frac{E(S : S^c)}{d(S)} \right\}$, where d is the degree function of the graph.

Recall the definition of Υ ,

$$\Upsilon = \min\{\Upsilon(S) : 0 < \mu(S) \leq 1/2\} = \min_{S \in \mathcal{H}_\mu} \Upsilon(S). \quad (2.32)$$

where for $0 < \mu(S) \leq 1/2$,

$$\Upsilon(S) = \frac{\mathbb{E}(Z(S)^2 \mathbf{1}_{\{Z(S) < 0\}})}{\mu(S)}.$$

Example 20. *Asynchronous pull model.*

For $S \subseteq V$, with $0 < \mu(S) \leq 1/2$, it holds

$$\mu(S)\Upsilon(S) = \mathbb{E}(Z(S)^2 \mathbf{1}_{\{Z(S) < 0\}}) = \sum_{v \in S} \frac{1}{n} \frac{d_{S^c}(v)}{d(v)} \left(\frac{d(v)}{m} \right)^2.$$

The result above follows because a negative change in Z is produced when a vertex x with opinion 1 (i.e $x \in S$) changes its opinion to 0. Then, with probability $1/n$, we select vertex x and in turn, x pulls the opinion of a neighbour y with probability $1/d(x)$. The stationary distribution of x is $\mu(x) = d(x)/m$. If y has opinion 1, then $Z_t = 0$, but if y has opinion 0, then $Z_t = -d(x)/d(V)$. Working out the expression above we get

$$\begin{aligned} \mathbb{E}(Z(S)^2 \mathbf{1}_{\{Z(S) < 0\}}) &= \sum_{v \in S} \frac{1}{n} \frac{d_{S^c}(v)}{d(v)} \left(\frac{d(v)}{m} \right)^2 \\ &= \sum_{v \in S} \frac{1}{n} \frac{d_{S^c}(v)d(v)}{m^2} \geq \frac{d_{\min}}{nm^2} E(S : S^c), \end{aligned} \quad (2.33)$$

therefore, from (2.10), it holds

$$\Upsilon \geq \frac{d_{\min}}{nm} \min_{S \in \mathcal{H}_d} \frac{E(S : S^c)}{d(S)} = \frac{d_{\min}}{nm} \Phi. \quad (2.34)$$

We conclude that $\mathbb{E}(\tau_{con}) = \mathcal{O}(nm/(d_{\min}\Phi))$. For r -regular graphs, we obtain $\mathbb{E}(\tau_{cons}) = \mathcal{O}(n^2/\Phi)$.

Example 21. *Asynchronous push model.*

Let $C = (\sum_{x \in V} d(x)^{-1})^{-1}$. Then,

$$\mathbb{E} (Z(S)^2 \mathbb{1}_{\{Z(S) < 0\}}) = \sum_{x \in S} \left(\frac{C}{d(x)} \right)^2 \sum_{y: y \sim x, y \in S^c} \frac{1}{n} \frac{1}{d(y)}.$$

This is explained as following. For this model, $Z(S) \leq 0$ only if an opinion is pushed on a vertex of S from a vertex in S^c . If an opinion is pushed on x , the difference is $Z(S) = -(C/d(x))$. The probability of changing x is given by the probability of selecting a neighbouring vertex $y \in S^c$ to push its opinion on x .

Then,

$$\mathbb{E} (Z(S)^2 \mathbb{1}_{\{Z(S) < 0\}}) \geq \frac{C^2}{d_{\max} n} \sum_{v \in S} \sum_{w \in S^c} \frac{\mathbb{1}_{v \sim w}}{d(v)d(w)}. \quad (2.35)$$

Define the function $J : V \rightarrow \mathbb{R}^+$ given by $J(v) = (d(v))^{-1}$. Then, by using that the stationary distribution is $\mu(v) = C/d(v)$, we have

$$\Upsilon \geq \frac{C}{nd_{\max}} \Psi = \frac{C}{nd_{\max}} \min_{S \in \mathcal{H}_J} \frac{\sum_{v \in S} \sum_{w \in S^c} \frac{\mathbb{1}_{v \sim w}}{d(v)d(w)}}{J(S)}. \quad (2.36)$$

In general, the parameter Ψ does not seem related to the classical graph parameters.

If the graph is regular of degree d , then

$$\Psi = \frac{1}{n^2} \Phi,$$

in which case $\mathbb{E}(\tau_{con}) = \mathcal{O}(n^2/\Phi)$, which agrees with the asynchronous pull model in Example 20. In general, it can be seen that $\Upsilon = \Omega(n^{-6})$, obtaining a universal upper bound for the consensus time of $\mathcal{O}(n^6)$.

Example 22. Abusive push model.

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

$$\mathbb{E} \left(Z(S)^2 \mathbf{1}_{\{Z(S) < 0\}} \right) = \sum_{x \in S^c} \frac{1}{n} \left(\frac{d_{S^c}(x)}{n} \right)^2.$$

Note that $Z(S) \leq 0$ if the vertex that pushes is in S^c . Then, with probability $1/n$, we sample a particular vertex x , which in turn pushes its opinion on all its neighbours. Since the stationary distribution for this model is $\mu(v) = 1/n$, then the change Z_t is $-d_{S^c}(x)/n$. Denote $\mathcal{H} = \{S \subseteq V : 0 < |S| \leq n/2\}$, then it holds that

$$\begin{aligned} \Upsilon &\geq \frac{1}{n^2} \min_{S \in \mathcal{H}} \left\{ \frac{\sum_{x \in S^c} d_{S^c}(x)^2}{|S|} \right\} \geq \frac{1}{n^2} \min_{S \in \mathcal{H}} \left\{ \frac{\sum_{x \in S^c} d_{S^c}(x)}{|S|} \right\} \\ &= \frac{1}{n^2} \min_{S \in \mathcal{H}} \frac{E(S : S^c)}{|S|}. \end{aligned} \quad (2.37)$$

The parameter $\min_{S \in \mathcal{H}} \frac{E(S : S^c)}{|S|}$ (minimum over all sets) is very similar to the graph conductance, indeed, for d -regular graphs $\min_{S \in \mathcal{H}} \frac{E(S : S^c)}{|S|} = d\Phi(G)$. In such a case, we have that

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

Example 23. Oblivious.

Our final example is the oblivious model. In this model, the stationary distribution is uniform. Then, the only way to produce a negative change in $Z(S)$ is by choosing a edge with endpoints in S and S^c , and making the vertex in S imitate the opinion of the one in S^c , which happens with probability $1/2$. Then

$$\mathbb{E} \left(Z(S)^2 \mathbf{1}_{\{Z(S) < 0\}} \right) = \frac{E(S : S^c)}{m} \frac{1}{2n^2},$$

and thus

$$\Upsilon(S) = \frac{E(S : S^c)}{2mn|S|}.$$

Recall we defined $\mathcal{H} = \{S \subseteq V : 0 < |S| \leq n/2\}$, hence

$$\Upsilon = \frac{1}{2mn} \min_{S \in \mathcal{H}} \frac{E(S : S^c)}{|S|}.$$

Observe that this parameter is similar to the one obtained in the abusive push model.

2.6 Coalescing Process and Duality

A remarkable property of the linear voting model is the existence of an associated time-reversed dual process. The existence of this process is important as it allows us to bypass the study of the original voting process by studying properties of the dual process.

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

$$f_t = \begin{cases} f, & \text{if } t = 0 \\ M_{t-1}^\top f_{t-1}, & \text{if } t > 0. \end{cases} \quad (2.38)$$

We call $(f_t)_{t \geq 0}$ the *coalescing process*, and we denote $(f_t)_{t \geq 0} \sim \mathcal{CP}(l, f)$. This process can be interpreted as follows. The vector f_t counts the number of particles at each of the vertices of the graph, and $M_{t-1}^\top$ moves the particles at time $t-1$. In particular, if $M_{t-1}(v, u) = 1$, it means that all the particles on vertex v at time $t-1$ are in vertex u at time t . In this process, once two or more particles come together, they coalesce (merge), and keep moving together. This is why the process $(f_t)_{t \geq 0}$ received the name of *coalescing process*. Observe that $\sum_{v \in V} f_t(v)$ is constant over time. For practical purposes, the standard choice of the initial vector f is the all 1 vector $\mathbf{1}$, indicating that we start the process with one particle at each vertex. For the rest of the chapter, we assume this is the case.

For a set of vertices $A \subseteq V$ define $f_t(A) = \sum_{v \in A} f_t(v)$. Consider a partition $\mathcal{A} = \{A_1, \dots, A_m\}$ of the vertices V . We say that the coalescing process is in agreement with respect to $\mathcal{A}$ at time t , if for some index $i \in \{1, \dots, m\}$, we have $f_t(A_i) = n$, and consequently, $f_t(A_j) = 0$ for all $j \neq i$. More formally,

$$\{f_t \text{ is in agreement with respect to } \mathcal{A}\} \Leftrightarrow \bigcup_{i=1}^m \{f_t(A_i) = n\}. \quad (2.39)$$

Given a vector of opinions ξ , we denote by $\mathcal{A}_\xi$, the natural partition given by the vertices with the same opinion, that is, if the set of opinions is Q , then $\mathcal{A}_\xi = \{A_a :$

$a \in Q\}$ where $A_a = \{v : \xi(v) = a\}$. The main result of this section is the proof Theorem 3, which is stated again here.

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

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

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

2.6.1 Coalescing Process

We begin by giving some intuition about the coalescing process $(f_t)_{t \geq 0} \sim \mathcal{CP}(l, \mathbf{1})$. 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 $g_M : V \rightarrow V$. Let χ_v be the characteristic vector for v . Then, we have that

$$g_M(v) = w \iff M(v, w) = 1 \iff M^\top \chi_v = \chi_w. \quad (2.40)$$

Let $(M_t)_{t \geq 0}$ be an i.i.d. sequence of matrices in $\mathcal{M}$ with distribution l , and let $(g_{M_t})_{t \geq 0}$ 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 vertex-valued process $(X_t)_{t \geq 0}$ where

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

As 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. The process $(X_t(v))_{t \geq 0}$ represents the trajectory of

particle that starts at v . Since we have one particle starting from each vertex, we have n processes simultaneously, in other words, we have $(X_t(v) : v \in V)_{t \geq 0}$ where $X_0(v) = v$ for each $v \in V$. The process $(X_t(v) : v \in V)_{t \geq 0}$ is the Markov chain for the trajectories of the whole configuration of particles in the coalescing process. Observe that when two particles meet, they keep moving together. The reason for this, is that 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., $f_t(w) = |\{v \in V : X_t(v) = w\}|$. Similarly to above,

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

Recall that the set $\mathcal{M}$ is 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. The set $\mathcal{M}$ is closed under multiplication, i.e., for $M_1, M_2 \in \mathcal{M}$, $M_1 M_2 \in \mathcal{M}$. 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

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

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 24. *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 $X_t(v)$. $\square$

In Section 2.3, we gave some examples of voting processes. The next step for us, is to find the dual processes associated to each one of them. Remember that in the voting process model, we sample a matrix M , and then we multiply the current vector of opinions by it. Note that the matrix M has the associated function g . By (2.40), 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 the opinion of w , i.e., $w = g(v)$, then in the coalescing process all particles in v are pushed to w (i.e., the particles, if any, move from v to w). In the same way, if v pushes its opinion on w , i.e., $g(w) = v$, then in the coalescing process, v pulls all the particles from w (i.e., the particles, if any, move from w to v).

As a general heuristic to find the dual of a linear voting model, replace the words pull by push, and vice versa, and replace the word opinion by particle(s).

Example 25. *Dual of synchronous pulling model.*

The model is described as follows. At each round, each vertex pulls the opinion of a random neighbour. Thus, in the dual process, each vertex v pushes its particles (if any) to a random neighbour of v . Note that this is equivalent to saying that each particle moves to an adjacent vertex selected uniformly 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 *Synchronous Coalescing Random Walks*. It has been extensively studied, see e.g., [3], [20], [28], [12], [56] and references therein.

Example 26. *Dual of asynchronous pulling model.*

Remember that at 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. At each round, select a random vertex v , and then v pushes its particles (if any) to a random neighbour w . This process is known as the *Asynchronous Coalescing Random Walks*, and they have been extensively studied in the continuous-time setting, while instead of having discrete time-steps, each vertex has an independent Poisson clock. We refer to [3], [67], and [68] for details.

Example 27. *Dual of asynchronous pushing model.*

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

Example 28. *Dual of abusive push model.*

In this model, at each round a random vertex v is selected to push its opinion on all its neighbours. In the dual, a random vertex v is selected, and then v pulls all the particles of its neighbours.

Example 29. *Dual of oblivious model.*

In this model, at each round, we choose a random edge, and a random endpoint to pull the opinion of the other endpoint. So, in the dual, at each round we choose a random edge, and a random endpoint, then if there are particles in such an endpoint, we move them to the other endpoint.

2.6.2 The Dual Relation

In this section, we present the relationship between the voting process and its coalescing particle dual process. This relation is in a distributional sense, that is, we do not have a coupling between the voting and coalescing processes but 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 $\mathcal{A}$ of the vertices V into the sets $A_1, \dots, A_m$. We say that the coalescing process is in agreement with respect to $\mathcal{A}$ at time t , if for some index $i \in \{1, \dots, m\}$, we have $f_t(A_i) = n$, and consequently, $f_t(A_j) = 0$ 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 such a case, the partition considers all the singletons of V , and thus, we have agreement in the coalescing process if and only if all particles are together, that is, all particles have coalesced into one. Here, we define the coalescing time τ_{coalsc} as the first time all particles are together.

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

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

Proof. Let $(\xi_t)_{t \geq 0} \sim \mathcal{LVM}(\xi, l)$. If $t = 0$, then $\xi_t = \xi$ and the result holds. Note that if ξ is in consensus, the result also holds easily. Assume $t \geq 0$ and that ξ is not in consensus. Let Q be the set of possible opinions. For convenience of the proof, we assume that the opinions in Q are positive natural numbers. For $b \in Q$, let $A_b = \{v : \xi(v) = b\}$. Given ξ , define $\chi_b : V \rightarrow \{0, 1\}$ such that $\chi_b(v) = 1$ if and only if $\xi(v) = b$. Then, ξ can be written as

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

Let $\xi_t = M_{t-1} \dots M_0 \xi$, and note that $M = M_{t-1} \dots M_0$ is a matrix with exactly one 1 in each row, and all other entries 0. We say ξ_t is in consensus at a if $\xi_t = a \mathbf{1}$. Observe

that such a ξ_t can be written as

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

Observe that as χ_b indicates the set of vertices with opinion b in ξ (i.e. gives opinion b the value 1, and all other opinions the value 0), then $M \chi_b$ indicates the set of vertices with opinion b in $M \xi$. So, if $\xi_t = a \mathbf{1}$, it implies that $M_{t-1} \dots M_0 \chi_a = \mathbf{1}$, and $M_{t-1} \dots M_0 \chi_b = 0$ for $b \neq a$. Therefore, we have the following equalities between events

$$\begin{aligned} \{M_{t-1} \dots M_0 \xi = a \mathbf{1}\} &= \{M_{t-1} \dots M_0 \chi_a = \mathbf{1}\} \\ &= \{\langle M_{t-1} \dots M_0 \chi_a, \mathbf{1} \rangle = n\} = \{\langle \chi_a, M_0^\top \dots M_{t-1}^\top \mathbf{1} \rangle = n\}. \end{aligned} \quad (2.41)$$

Define $f'_t = M_0^\top \dots M_{t-1}^\top \mathbf{1}$. Observe that the event $\langle \chi_a, f'_t \rangle = n$ is equivalent to $f'_t(A_a) = \sum_{v \in A_a} f'_t(v) = n$. Consider the coalescing process $(f_t)_{t \geq 0}$ given by $f_t = M_{t-1}^\top \dots M_0^\top \mathbf{1}$ as defined in Equation (2.38). Note $(f_t)_{t \geq 0}$ and $(f'_t)_{t \geq 0}$ are different as $M_{t-1}^\top \dots M_0^\top \neq M_0^\top \dots M_{t-1}^\top$, but they have the same distribution (denoted $f'_t \stackrel{\mathcal{D}}{=} f_t$) as the matrices $M_0, \dots, M_{t-1}$ are i.i.d samples from a distribution l over matrices in $\mathcal{M}$, and thus the probability of obtaining one sequence of matrices or the opposite order is the same.

Since consensus can be attained by one and only one opinion, we have that

$$\begin{aligned} \mathbb{P}(\xi_t \text{ is in consensus}) &= \mathbb{P} \left(\bigcup_{a \in Q} \{M_{t-1} \dots M_0 \xi = a \mathbf{1}\} \right) \\ (\text{disjoint events}) &= \sum_{a \in Q} \mathbb{P}(\{M_{t-1} \dots M_0 \xi = a \mathbf{1}\}) \\ (\text{Equation (2.41)}) &= \sum_{a \in Q} \mathbb{P}(\langle \chi_a, f'_t \rangle = n) \\ &= \sum_{a \in Q} \mathbb{P}(f'_t(A_a) = n) \\ (f'_t \stackrel{\mathcal{D}}{=} f_t) &= \sum_{a \in Q} \mathbb{P}(f_t(A_a) = n) = \mathbb{P} \left(\bigcup_{a \in Q} \{f_t(A_a) = n\} \right). \end{aligned} \quad (2.42)$$

Finally, as in Equation (2.39), we have that the event

$$\{f_t \text{ is in agreement with respect to } \xi\}$$

can be written as $\bigcup_{a \in Q} \{f_t(A_a) = n\}$, finishing our proof. $\square$

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

$$\mathbb{P}(\tau_{\text{cons}} \leq t) = \mathbb{P}(\xi_t \text{ is in consensus}). \quad (2.43)$$

By Theorem 3, we have that $\mathbb{P}(\xi_t \text{ is in consensus})$ is equal to the probability that f_t is in agreement with respect to ξ at time t . On the other hand, suppose all the opinions of ξ are different, so $\mathcal{A}_\xi$ 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

$$\mathbb{P}(\tau_{\text{coalsc}} \leq t) = \mathbb{P}(f_t \text{ is in agreement with respect to } \xi). \quad (2.44)$$

As a corollary of Theorem 3 and equations (2.43) and (2.44), we have the following result.

Corollary 30. *Let $(\xi_t)_{t \geq 0} \sim \mathcal{LVM}(l, \xi)$ and $(f_t)_{t \geq 0} \sim \mathcal{CP}(l, \mathbf{1})$ and assume all vertices have different opinions in ξ . Then τ_{coalsc} and τ_{cons} have the same distribution.*

2.6.3 Coalescence Time and Meeting Times

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

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

As usual $\min(\emptyset) = \infty$. Define the worst-case expected meeting time, T_{meet} , as

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

We relate the meeting time and the coalescing time in the following variant of the Matthew's Method (see [3], Theorem 2.26).

Theorem 31. *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 $v_1, \dots, v_n$ be an arbitrary ordering of the states V . In this proof, the particle that starts at vertex v is simply called particle v . Define the cluster of particle v_1 as

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

Note that $C_0 = \{v_1\}$, and since after a meeting particles keep moving together, we have that $C_t \subseteq C_{t+1}$ for all $t \geq 0$.

Consider a uniformly random labelling $\pi_2, \dots, \pi_n$ of $V \setminus \{v_1\}$ (each vertex v is associated to a unique π_k), and for consistency, the label of v_1 is π_1 . For $2 \leq k \leq n$, define the hitting times $S_k = \min\{t \geq 0 : \{\pi_2, \dots, \pi_k\} \subseteq C_t\}$. Observe that $S_n = \min\{t \geq 0 : \{\pi_2, \dots, \pi_n\} \subseteq C_t\} = \tau_{\text{coalsc}}$. We compute the expected value of S_n . Note that

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

then, by taking expected value given the ordering π and the position of the particles at time S_i , we obtain that

$$\begin{aligned} \mathbb{E}((S_{i+1} - S_i) \mathbf{1}_{\{S_{i+1} > S_i\}} | \pi, X_{S_i}) &= \\ \mathbb{E}(\tau_{\text{meet}}(X_{S_i}(\pi_1), X_{S_i}(\pi_{i+1})) | \pi, X_{S_i}) \mathbf{1}_{\{S_{i+1} > S_i\}}. \end{aligned} \quad (2.45)$$

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, it holds $\mathbb{1}_{\{S_{i+1} > S_i\}} = \mathbb{1}_{\{X_{S_i}(\pi_1) \neq X_{S_i}(\pi_{i+1})\}}$ 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, by the strong Markov property (Theorem 87 in Appendix A),

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

By computing the expected value of Equation (2.45), 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 $\mathbb{P}(S_{i+1} > S_i) \leq 1/i$. Note the event $\{S_{i+1} > S_i\}$ is equivalent to the event $\{\pi_{i+1} \notin C_{S_i}\}$. Define $r : V \rightarrow [n]$ as a bijective function with $r(v_1) = 1$, and such that for every $v_i, v_j \in V$, we have that $r(v_i) < r(v_j)$ if vertex v_1 meets with v_i before v_1 meets v_j . In other words, r gives the order in which vertices meet with v_1 (if v_i and v_j meet v_1 at the same time, then $r(v_i) < r(v_j)$ if and only if $i < j$). Then

$$\{\pi_{i+1} \notin C_{S_i}\} \subseteq \bigcap_{k=2}^i \{r(\pi_{i+1}) > r(\pi_k)\}.$$

The above equality holds because if $\pi_{i+1} \notin C_{S_i}$, then π_{i+1} has to join the cluster of v_1 later than $\pi_1, \dots, \pi_i$. Finally, we claim that

$$\mathbb{P} \left(\bigcap_{k=2}^i \{r(\pi_{i+1}) > r(\pi_k)\} \middle| r \right) = 1/i.$$

The latter holds due to the fact that, given any fixed order r of the vertices, if we consider our random labelling π of the vertices, the probability that $r(\pi(i+1))$ is greater than all $r(\pi_k)$, $2 \leq k \leq i$, is $1/i$. $\square$

From the proof above, we have the following corollary.

Corollary 32. *Let $n = |V| \geq 2$, and assume that initially there are $m \leq n$ particles in the system. Then, independently of their starting position, the coalescing time of the m particles is*

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

2.7 Analysis on Standard Graph Families

In this section, we show results of the expected consensus time for four of our models on standard graph families. The models we consider are asynchronous pull, push, abusive pushing and oblivious model. Here, we assume that all vertices have different opinions at the beginning of the process. In order to avoid making this section extremely long, i) we avoid certain technicalities in our computations, and ii) we claim, without proof, that all our bounds are tight, that is, the upper bound matches the lower bound. In most of the cases, the lower bound can be obtained by computing T_{meet} (defined in Section 2.6.3), together with the fact that $\mathbb{E}(\tau_{\text{cons}}) = \mathbb{E}(\tau_{\text{coalsc}}) \geq T_{\text{meet}}$ (Theorem 3 and definition of T_{meet}).

The graphs that we analyse are the following. Complete graph, Cycle, Star graph, Double Star, and Barbell. We give a brief description of all of them.

Complete Graph K_n . The complete graph is a graph on n vertices where all pairs of different vertices are adjacent.

Cycle Graph C_n . The cycle graph is a graph on vertex set $\{0, \dots, n-1\}$, where vertices i and $i+1$ are adjacent (here, we understand that $0 = n$).

Star Graph S_n . The star graph is a graph on n vertices, $n-1$ edges, and a central vertex, which is adjacent with the remaining $n-1$ vertices.

Double Star Graph S_{2n}^* . This graph consist of two copies of a star S_n , such that the two central vertices are connected by an additional (central) edge. The total number of vertices is $2n$, and number of edges is $2(n-1) + 1$.

Barbell Graph K_{2n}^* . This graph consist of two copies of a complete graph K_n , with an additional edge joining two vertices in different cliques. In this graph, there are $2n$ vertices, and $2\binom{n}{2} + 1$ edges. Making an analogy with the double star, we call the extra edge the central edge, and its endpoints are the central vertices of each clique.

Figure 2.7 shows the Double Star and the Barbell graph. A summary of our results is given in Figure 2.2

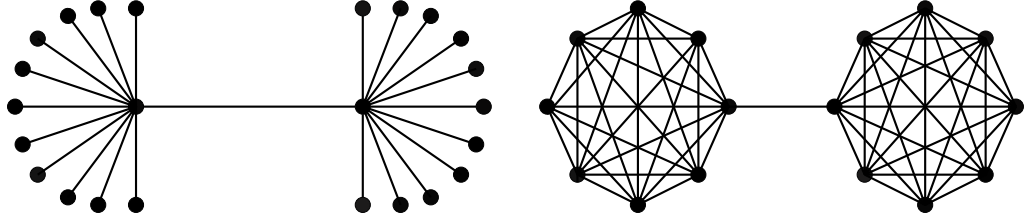


Figure 2.1: Left panel: Double star S_{24}^* . Right panel: Barbell K_{16}^*

	Async. Pull	Async. Push	Abusive Pushing	Oblivious
Complete graph K_n	$\mathcal{O}(n^2)$	$\mathcal{O}(n^2)$	1	$\mathcal{O}(n^2)$
Cycle C_n	$\mathcal{O}(n^3)$	$\mathcal{O}(n^3)$	$\mathcal{O}(n^3)$	$\mathcal{O}(n^3)$
Star graph S_n	$\mathcal{O}(n \log n)$	$\mathcal{O}(n^3)$	$\mathcal{O}(n)$	$\mathcal{O}(n^2)$
Double star S_n^*	$\mathcal{O}(n^2)$	$\mathcal{O}(n^4)$	$\mathcal{O}(n^2)$	$\mathcal{O}(n^2)$
Barbell graph	$\mathcal{O}(n^3)$	$\mathcal{O}(n^3)$	$\mathcal{O}(n^2)$	$\mathcal{O}(n^3)$

Figure 2.2: Results of different models on different graphs. Each entry represent the expected consensus time of a model in a particular graph.

2.7.1 Asynchronous Pull Model.

Complete Graph K_n .

Recall that from Theorem 3, it holds $\mathbb{E}(\tau_{\text{cons}}) = \mathbb{E}(\tau_{\text{coalsc}})$. From [3, Chapter 14], we have that $\mathbb{E}(\tau_{\text{coalsc}}) \sim n^2/2$.

Cycle Graph C_n .

From Theorem 2, Equation (21), it holds

$$\mathbb{E}(\tau_{\text{cons}}) = \mathcal{O}(n^2/\Phi),$$

and for a cycle $\Phi = \Theta(1/n)$, thus $\mathbb{E}(\tau_{\text{cons}}) = \mathcal{O}(n^3)$.

Star Graph S_n . From the dual relation of Theorem 3, and Theorem 31, we get $\mathbb{E}(\tau_{\text{cons}}) = \mathbb{E}(\tau_{\text{coalsc}}) = \mathcal{O}(T_{\text{meet}} \log n)$. We compute the meeting time of two particles. By ignoring the movement of all other particles, the two particles can be represented

by a 3-state Markov chain. Indeed, the states are

$$\begin{aligned}
(L, L) &= \{\text{both particles are in different leaves of } S_n.\} \\
(C, L) &= \{\text{One particle is on a leaf and the other in the central vertex of } S_n.\} \\
F &= \{\text{Both particles are together.}\}
\end{aligned} \tag{2.46}$$

The transition matrix is given as following.

$$\begin{pmatrix} (n-1)/(n+1) & 2/(n+1) & 0 \\ (n-1)/(n(n+1)) & (n-1)/(n+1) & 1/n \\ 0 & 0 & 1 \end{pmatrix},$$

where the first, second, and last columns represent states (L, L) , (C, L) and F , respectively. For a state A , let T_A be the expected time to reach state F . Let $T = (T_{(L,L)}, T_{(C,L)})$, and let M be the matrix containing the first two rows and columns of the matrix above. Then, T satisfies the following system of equations

$$T = MT + I\mathbf{1},$$

where $\mathbf{1} = (1, 1)$, and I is the 2x2 identity. This is equivalent to

$$\begin{cases} T_{(L,L)} = \frac{n-1}{n+1}T_{(L,L)} + \frac{2}{n+1}T_{(C,L)} + 1 \\ T_{(C,L)} = \frac{n-1}{n(n+1)}T_{(L,L)} + \frac{n-1}{n+1}T_{(C,L)} + 1 \end{cases}. \tag{2.47}$$

By solving $T = (I - M)^{-1}\mathbf{1}$, we obtain $T = (T_{(L,L)}, T_{(C,L)}) = (2n, 3n/2 - 1/2)$.

Therefore $\mathbb{E}(\tau_{con}) \leq 2n \log n$.

Double Star Graph $S_n - S_n$.

We consider the dual process. Let L_t be the number of leaves occupied by a particle at time t . Then

$$\mathbb{E}(L_{t+1}|L_t) \leq L_t - L_t/n + 2/n.$$

Observe that the above holds because if an occupied leaf is chosen, it moves to one of the central vertices, reducing L_t by 1. If one of the two central vertices is chosen,

then if a particle is there, it moves to a leave, possible increasing the number of leaves in 1. Then, for $k \geq 1$,

$$\mathbb{E}(L_t) \leq 2 + \mathbb{E}(L_0)(1 - 1/n)^t \leq 2 + ne^{-t/n}.$$

Thus, by choosing $T = 2\lceil \log n \rceil n$, we have

$$\mathbb{E}(L_T) \leq 2 + 1/n \leq 3.$$

Therefore

$$\mathbb{P}(L_n > 6) \leq \frac{\mathbb{E}(L_n)}{6} = 1/2.$$

We repeat the following process after we have at most 6 occupied leaves. Indeed, we wait T rounds, if after those T round the number of occupied leaves is more than 6, restart the process with all the particles, otherwise, we finish the process. One iteration of the above process finishes with probability at least $1/2$, then the above process has to be repeated at most 2 times in expectation. Then, the expected time to have at most 6 occupied leaves is at most $2T = 4\lceil \log n \rceil n$. Let, T_6 be the first time there are at most 6 occupied leaves. Then at such a time, there are at most 8 particles in the system (counting the central vertices). By the Strong Markov property (see Appendix A), we need to compute the expected coalescing time of 8 particles. By Corollary 32, this expected coalescing time of 8 particles is $\mathcal{O}(T_{\text{meet}})$ where T_{meet} is the worst-case expected meeting time of two particles.

Given two particles, their positions in the double star can be modelled by 6 states, they are:

$$\begin{aligned} (L, L) &= \{\text{Both particles on leaves attached of the same star}\} \\ (L, L') &= \{\text{Both particles on leaves of different star}\} \\ (C, L) &= \{\text{One particle on a central vertex, and one in a leaf, in the same star}\} \\ (C, L') &= \{\text{One particle on a central vertex, the other in a leaf, in different stars}\} \\ (C, C') &= \{\text{Both particles in different central vertices}\} \\ F &= \{\text{Both particles together}\} \end{aligned}$$

Remember that at each time we choose a random vertex, and move the particle in that vertex (if any) to a random neighbour. Then, the transition matrix of the two particle system is given by

$$\begin{pmatrix} \frac{2n}{2n+2} & 0 & \frac{2}{2n+2} & 0 & 0 & 0 \\ 0 & \frac{2n}{2n+2} & 0 & \frac{2}{2n+2} & 0 & 0 \\ \frac{(n-2)}{(2n+2)n} & 0 & \frac{2n}{2n+2} & \frac{1}{(2n+2)n} & 0 & \frac{n+1}{(2n+2)n} \\ 0 & \frac{n-1}{(2n+2)n} & \frac{1}{(2n+2)n} & \frac{2n}{2n+2} & \frac{1}{2n+2} & 0 \\ 0 & 0 & 0 & \frac{2n-2}{(2n+2)n} & \frac{2n}{2n+2} & \frac{2}{(2n+2)n} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Let M the submatrix containing the first 5 rows and columns. For any state A , let T_A be the expected hitting time of state F starting from state A . Let $T = (T_{(L,L)}, \dots, T_{(C,C')})$. Then, similarly as in the star, $T = (I - M)^{-1}\mathbf{1}$ (with I and $\mathbf{1}$ with the appropriate dimensions). Then,

$$T = (6n + o(n), 2n^2 + o(n^2), 5n + o(n), 2n^2 + o(n^2), 2n^2 + o(n^2)).$$

Hence, $T_{\text{meet}} = \Theta(n^2)$. We conclude that the expected coalescing time is $\mathcal{O}(n \log n) + \Theta(n^2)$.

Barbell Graph $K_n - K_n$. In the Barbell graph, we have two complete graphs K_n joined by an edge. Let us call those complete graphs, the left and right graph. We proceed to study the coalescing random walks process. Let R_t and L_t be the number of particles in the right and left graph, respectively. Let $Z_t = R_t + L_t$, then

$$\mathbb{E}(R_{t+1} + L_{t+1} | R_t, L_t) \leq R_t + L_t - L_t(L_t - 1)/n^2 - R_t(R_t - 1)/n^2.$$

The above holds because with probability R_t/n , we choose a particle in the right graph, and with probability at least $(R_t - 1)/n$, we reduce the number of particles by one (by moving the particle to an occupied vertex in the right graph). The same happens on the left graph. Since $Z_t = L_t + R_t$, for $Z_t \geq 4$, the quantity above is

maximised by considering $L_t = R_t = Z_t/2$. Then, we have that

$$\begin{aligned}
\mathbb{E}(Z_{t+1} \mathbb{1}_{\{Z_{t+1} \geq 4\}} | R_t, L_t) &\leq \mathbb{E}(Z_{t+1} \mathbb{1}_{\{Z_t \geq 4\}} | R_t, L_t) \\
&\leq (Z_t(1 + 1/n^2) - Z_t^2/(2n^2)) \mathbb{1}_{\{Z_t \geq 4\}} \\
&\leq Z_t \left(1 - \frac{1}{n^2}\right) \mathbb{1}_{\{Z_t \geq 4\}}.
\end{aligned} \tag{2.48}$$

So, $\mathbb{E}(Z_t \mathbb{1}_{\{Z_t \geq 4\}}) \leq (1 - n^{-2})^t \mathbb{E}(Z_0 \mathbb{1}_{\{Z_0 \geq 4\}}) = (1 - n^{-2})^t n \leq n e^{-t/n^2}$. Take $T = n^2 \lceil \log(2n) \rceil$, then by Markov's inequality we have

$$\mathbb{P}(Z_t \geq 4) \leq \mathbb{P}(Z_T \mathbb{1}_{\{Z_T \geq 4\}} > 0) \leq 1/2.$$

Denote by T_4 the time there are 4 particles in the system. Then, by the same geometric argument used in the double star, it holds $\mathbb{E}(T_4) \leq \mathcal{O}(n^2 \log n)$. Similarly as in the double star, we need to find T_{meet} . We claim that $T_{\text{meet}} = \Theta(n^3)$, so $\mathbb{E}(\tau_{\text{cons}}) = \Theta(n^3)$. We represent the meeting process of two particles with 6 states. A central vertex in the barbell is one of the two vertices joined by the special edge that connects the cliques. A vertex is usual, if it is not central. Similarly as in the double star, we define the following states.

$$\begin{aligned}
(U, U) &= \{\text{Both particles on usual vertices in the same clique}\} \\
(U, U') &= \{\text{Both particles on usual vertices in different cliques}\} \\
(C, U) &= \{\text{One particle on a central, the other on a usual vertex, same clique}\} \\
(C, U') &= \{\text{One particle on a central, the other on a usual vertex, diff. cliques}\} \\
(C, C') &= \{\text{Both particles in different central vertices}\} \\
F &= \{\text{Both particles together}\}
\end{aligned}$$

It is implicit that in state (U, U) the particles are in different vertices. Using the states above, and by the definition of the dual of asynchronous pull voting process, the transition matrix is given by

$$\begin{pmatrix} \frac{n(n-1)-2}{n(n-1)} & \frac{1}{n(n-1)} & 0 & 0 & 0 & \frac{1}{n(n-1)} \\ 0 & \frac{n(n-1)-1}{n(n-1)} & 0 & \frac{1}{n(n-1)} & 0 & 0 \\ \frac{(n-2)}{2n^2} & 0 & \frac{2n-3}{2n-2} & \frac{1}{2n^2} & 0 & \frac{2n-1}{2n^2(n-1)} \\ 0 & \frac{n-1}{2n^2} & \frac{1}{2n^2} & \frac{2n-3}{2n-2} & \frac{1}{2n(n-1)} & 0 \\ 0 & 0 & 0 & \frac{n-1}{n^2} & \frac{n-1}{n} & \frac{1}{n^2} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Using the same notation as in the previous cases, we solve $T = (I - M)^{-1}\mathbf{1}$ to obtain that, the dominant terms of the vector T are

$$(n^3, 2n^2, n^3, 2n^3, 2n^3),$$

thus $T_{\text{meet}} = \Theta(n^3)$, and therefore $\mathbb{E}(\tau_{\text{coalsc}}) = \mathcal{O}(n^2 \log n) + \Theta(n^3)$.

2.7.2 Push Model.

Complete Graph K_n . This process is equivalent to Asynchronous Pull Voting on K_n , thus $\mathbb{E}(\tau_{\text{cons}}) \sim n^2/2$.

Cycle Graph C_n . This process is equivalent to Asynchronous Pull Voting on a Cycle, thus $\mathbb{E}(\tau_{\text{cons}}) = \mathcal{O}(n^3)$.

Star Graph S_n . We work with the dual process. Here, we choose a random vertex v to pull all the particles of a random neighbour. Suppose there are k particles in the system. We compute an upper bound for the expected time until the first meeting, i.e., the first time we have $k - 1$ particles. Suppose that the k particles are in the leaves (which clearly takes more time for the first meeting than having one particle in the central vertex). Then, if the central vertex pulls one of the particles of those k leaves, and then one of the $k - 1$ occupied leaves is chosen to pull, then one particle is removed. Such an event happen with probability

$$\left(\frac{1}{n} \frac{k}{n-1}\right) \left(\frac{k-1}{n}\right) = \frac{n^2(n-1)}{k(k-1)}.$$

Observe that such an event uses two rounds, so the expected time for the first meeting of k particles is less than $2n^2(n-1)/(k(k-1))$. Also, observe that it takes just one step to go from n particles to $n-1$. Then, the expected time to have only one particle is

$$\mathbb{E}(\tau_{\text{coalsc}}) \leq 1 + 2n^2(n-1) \sum_{k=2}^{n-1} \frac{1}{k(k-1)} = \mathcal{O}(n^3) \quad (2.49)$$

and, from the dual relation $\mathbb{E}(\tau_{\text{cons}}) = \mathbb{E}(\tau_{\text{coalsc}})$.

Double Star Graph $S_n - S_n$.

Suppose there are k particles in one star, l in the other star, and that $k+l \geq 4$. Then, similarly as in the star, we assume such particles are in the leaves of their respective stars. Then, in two rounds, and with probability at least $\frac{k(k-1)}{n^2(n-1)}$, one of the k particles meet with other in the same star, and with probability $\frac{l(l-1)}{n^2(n-1)}$ the same event happens in the other star. So, if in total there are $m = k+l$ particles, the probability that in two steps, two particles meet is at least

$$\min_{k+l=m; k, l \geq 0} \left\{ \frac{l(l-1)}{n^2(n-1)} + \frac{k(k-1)}{n^2(n-1)} \right\} \geq \frac{m^2/2 - m}{n^2(n-1)} \geq \frac{m^2}{4n^2(n-1)}.$$

The last inequality holds because $m \geq 4$. Let T_3 be the first time-step there are 3 particles in the system. Then

$$\mathbb{E}(T_3) = \mathcal{O} \left(n^3 \sum_{m=4}^{\infty} m^{-2} \right) = \mathcal{O}(n^3).$$

By the strong Markov property, we just need to find an upper bound for the coalescing time for three particles starting from any possible configuration. By Corollary (32), the expecting coalescing time of three particles is $\Theta(T_{\text{meet}})$, where T_{meet} is the worst-case expected meeting time of two particles.

We use the same representation used in the pull case. By using states (L, L) , (L, L') , (C, L) , (C, L') , (C, C') and F , the transition matrix is given by

$$\begin{pmatrix} \frac{n^2-1}{n^2} & 0 & \frac{1}{n^2} & 0 & 0 & 0 \\ 0 & \frac{n^2-1}{n^2} & 0 & \frac{1}{n} & 0 & 0 \\ \frac{(n-2)}{2n} & 0 & \frac{n^2+n-2}{2n^2} & \frac{1}{2n^2} & 0 & \frac{n+1}{2n^2} \\ 0 & \frac{n-1}{2n} & \frac{1}{2n^2} & \frac{n^2+n-2}{2n^2} & \frac{1}{2n^2} & 0 \\ 0 & 0 & 0 & \frac{n-1}{n} & \frac{n-1}{n^2} & \frac{1}{n^2} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

With the same notation as before, we solve $T = (I - M)^{-1}\mathbf{1}$, and obtain that the dominant terms in T are

$$(2n^3, n^4, 2n^3, n^4, n^4)$$

so $T_{\text{meet}} = \Theta(n^4)$, and $\mathbb{E}(\tau_{\text{coalsc}}) = \mathcal{O}(n^3) + \Theta(n^4)$.

Barbell Graph $K_n - K_n$. Note this graph is almost regular, so we expect a similar behaviour to the asynchronous pull voting. We work with the coalescing particles. The analysis is similar to the double star, but the central vertices in the Barbel graph are the two vertices joined by the edge between the two cliques. When the number of particles is greater than or equal to 4, the exact same analysis done for the double star holds here, so the expected number of steps to have 3 particles is $\mathcal{O}(n^3)$. For the rest, we need to upper bound the expected meeting time of three particles, which is $\Theta(T_{\text{meet}})$. We claim that $T_{\text{meet}} = \Theta(n^3)$. We represent the meeting process using the same states used for the barbell graph in the asynchronous pull model, i.e., (U, U) , (U, U') , (C, U) , (C, U') , (C, C') and F . The transition matrix is given by

$$\begin{pmatrix} \frac{n^2-n-1}{n^2-n} & 0 & \frac{1}{n^2} & 0 & 0 & \frac{1}{n(n-1)} \\ 0 & \frac{n^2-1}{n^2} & 0 & \frac{1}{n^2} & 0 & 0 \\ \frac{n-2}{2n(n-1)} & 0 & \frac{2n^2-n-2}{2n^2} & \frac{1}{2n^2} & 0 & \frac{2n-1}{2n^2(n-1)} \\ 0 & \frac{1}{2n} & \frac{1}{2n^2} & \frac{2n^2-n-2}{2n^2} & \frac{1}{2n^2} & 0 \\ 0 & 0 & 0 & \frac{1}{n} & \frac{n^2-n-1}{n^2} & \frac{1}{n^2} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Using the same notation as the previous case, we solve $T = (I - M)^{-1}\mathbf{1}$. Then, the leading terms in T are

$$\left(\frac{2}{3}n^3, \frac{5}{3}n^3, \frac{2}{3}n^3, \frac{5}{3}n^3, \frac{5}{3}n^3\right).$$

Hence, $T_{\text{meet}} = \Theta(n^3)$, and so, $\mathbb{E}(\tau_{\text{coalsc}}) = \Theta(n^3)$.

2.7.3 Abusive Pushing Model.

Complete Graph K_n . The process finishes in one step since all vertices are connected.

Cycle Graph C_n . From Example 22, we have $\mathbb{E}(\tau_{\text{cons}}) = \mathcal{O}(n^2/(d\Phi))$, but for a cycle, $\Phi = \Theta(1/n)$. Therefore

$$\mathbb{E}(\tau_{\text{cons}}) = \mathcal{O}(n^3). \quad (2.50)$$

Star Graph S_n . The process finishes if and only if the central vertex is chosen to push. Then $\mathbb{E}(\tau_{\text{cons}}) = n$.

Double Star Graph $S_n - S_n$. The process finishes if we choose both central vertices in two consecutive steps. Then, the expected consensus time is $\mathbb{E}(\tau_{\text{cons}}) = \mathcal{O}(n^2)$.

Barbel Graph $K_n - K_n$. The same argument for the double-star works here, where the role of the central vertices is played by the two adjacent vertices that belongs to different cliques. Then $\mathbb{E}(\tau_{\text{cons}}) = \mathcal{O}(n^2)$.

2.7.4 Oblivious Model.

Complete Graph K_n . This process is equivalent to Asynchronous Pull Voting on K_n , thus $\mathbb{E}(\tau_{\text{cons}}) \sim n^2/2$.

Cycle Graph C_n . This process is equivalent to Asynchronous Pull Voting on a Cycle, thus $\mathbb{E}(\tau_{\text{cons}}) = \mathcal{O}(n^3)$.

Star Graph S_n . We recall the dual process. At each round, we choose a edge and we move the particles (if any) of one random endpoint to the other. Let $N = n - 1$

be the number of edges of the star. We work with the dual process. We mimic the case of asynchronous push voting on S_n . Suppose there are $k \leq n - 1$ particles in the star (It takes just one step to go from n particles to $n - 1$). Let us bound the expected time to reduce the number of particles by one. Clearly, the worst case is when the particles are on the leaves. In two steps, the following event happens with probability $k(k - 1)/(2N^2)$: in one step, we chose an edge adjacent to one occupied leaf (probability $k/(n - 1)$), and with probability $1/2$ such a particle moves to the central vertex. In the second step, we choose an edge adjacent to another occupied leaf (with probability $(k - 1)/(n - 1)$), then two particles meet. Therefore, we need to wait at most $4N^2/(k(k - 1))$ steps in expectation to reduce the number of particles by one. Then

$$\mathbb{E}(\tau_{\text{coalsc}}) \leq 4N^2 \sum_{k=2}^{\infty} \frac{1}{k(k - 1)} = \mathcal{O}(N^2) = \mathcal{O}(n^2). \quad (2.51)$$

Double Star Graph $S_n - S_n$. Let $N = 2(n - 1) + 1$ be the total number of edges. Similarly to the push model, suppose there are k particles in one star, l in the other star, and that $k + l \geq 4$. We proceed to bound the expected number of steps until the first meeting occurs. We assume that all particles are on the leaves (worst-case scenario). Then, in two rounds, with probability at least $(k(k - 1))/(2N^2)$, two particles in one star meet, and with probability $l(l - 1)/(2N^2)$, two particles in the other star meet. This holds because with probability k/N , we choose an edge adjacent to an occupied leave in the first start, and with probability $1/2$, the particle moves to the central vertex. Then, with probability $(k - 1)/N$, we choose another edge adjacent to one of the $k - 1$ occupied leaves, and then the two particles meet. Hence, if there are m particles in the system, by optimising over $k + l = m$, we have that the probability of a meeting in two steps is less than

$$\frac{m^2/2 - m}{8(n - 1)^2} \geq \frac{m^2}{4N^2}.$$

The last inequality holds because $m = k + l \geq 4$. So the expected time to have 3 particles is at most $\mathcal{O}(N^2) = \mathcal{O}(n^2)$ since $\sum_{m=4}^{\infty} m^{-2} = \mathcal{O}(1)$.

As the previous cases, we need to compute T_{meet} . We use the same 6 states used in the pull and push model. The transition matrix is

$$\begin{pmatrix} \frac{N-1}{N} & 0 & \frac{1}{N} & 0 & 0 & 0 \\ 0 & \frac{N-1}{N} & 0 & \frac{1}{N} & 0 & 0 \\ \frac{(n-2)}{2N} & 0 & \frac{2N-n-1}{2N} & \frac{1}{2N} & 0 & \frac{1}{N} \\ 0 & \frac{n-1}{2N} & \frac{1}{2N} & \frac{2N-n-1}{2N} & \frac{1}{2N} & 0 \\ 0 & 0 & 0 & \frac{N-1}{2N} & \frac{N-1}{2N} & \frac{1}{N} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Solving $T = (I - M)^{-1}\mathbf{1}$, we obtain that the leading terms of T are given by $(2n^2, 4n^2, 2n^2, 4n^2, 4n^2)$, and thus $T_{\text{meet}} = \Theta(n^2)$, concluding that $\mathbb{E}(\tau_{\text{cons}}) = \Theta(n^2)$.

Barbell Graph $K_n - K_n$. Let $N = 2\binom{n}{2} + 1$ be the total number of edges. Suppose there are k particles in one clique, l particles in the other, and $k + l \geq 4$. Note that to have one meeting in the first clique we just need to choose an edge such that both endpoints are occupied by a particle. That happens with probability $\binom{k}{2}/N$. This is proportional to the probability as in the double star process (also here, we just need one step, instead of two). Then, following the same argument as in the double star, the expected time to have three particles is $\mathcal{O}(N)$. As usual, we need to compute T_{meet} . We use the 6 states representation. The transition matrix is

$$\begin{pmatrix} \frac{N-2}{N} & 0 & \frac{1}{N} & 0 & 0 & \frac{1}{N} \\ 0 & \frac{N-1}{N} & 0 & \frac{1}{N} & 0 & 0 \\ \frac{(n-2)}{2N} & 0 & \frac{2N-n-1}{2N} & \frac{1}{2N} & 0 & \frac{1}{N} \\ 0 & \frac{n-1}{2N} & \frac{1}{2N} & \frac{2N-2n+1}{2N} & \frac{n-1}{2N} & 0 \\ 0 & 0 & 0 & \frac{N-1}{2N} & \frac{N-1}{2N} & \frac{1}{N} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

By solving $T = (I - M)^{-1}\mathbf{1}$, we obtain that the dominant terms in T are $(2n^2, n^3, 3n^2, n^3, n^3)$, and thus $T_{\text{meet}} = \Theta(n^3)$. Therefore, $\mathbb{E}(\tau_{\text{cons}}) = \mathcal{O}(N) + \Theta(n^3) = \Theta(n^3)$.

Chapter 3

Coalescing Branching Random Walks on Graphs

3.1 Introduction

In this chapter, we study the *coalescing-branching* random walk process (COBRA walk or COBRA process, for short) for propagating information through a connected n -vertex graph. The COBRA process, which was introduced in [38], can be viewed as spreading a single item of information throughout an undirected graph in synchronised rounds. That is, at each round, each vertex which has received the information in the previous step (possibly, simultaneously from more than one neighbour and not necessarily for first time), *pushes* (sends) the information to k randomly selected neighbours, but it does not keep the information (unless it receive the information from another vertex). The COBRA process is typically studied for integer branching factors $k \geq 2$ (with the case $k = 1$ corresponding to a random walk).

The main quantity of interest in information propagation processes (and of this chapter) is the time taken to inform (or visit) all vertices. By analogy with a random walk, this is referred to as the *cover time*.

In the literature, there already exists some w.h.p.¹ cover time results for the COBRA process. For example, the work of Dutta et al. [39] includes the following results for the case $k = 2$.

- i. For the complete graph K_n all vertices are visited in $\Theta(\log n)$ rounds.
- ii. For regular constant degree expanders, the cover time is $\mathcal{O}(\log^2 n)$.
- iii. For the D -dimensional grid on n vertices, the cover time is $\mathcal{O}(n^{1/D}(\log n)^j)$, for some $j \geq 1$.
- iv. For n -vertex trees, the cover time is $\mathcal{O}(n \log n)$, which is tight for the star graph.

The following improved bounds were shown in [62]. Indeed, they showed that

- i. For r -regular n -vertex graphs, the cover time is $\mathcal{O}((r^4/\phi^2) \log^2 n)$, where ϕ denotes the conductance of G , defined by

$$\phi = \min_{S \subseteq V} \left\{ \frac{E(S : S^c)}{r \min\{|S|, |S^c|\}} \right\},$$

where $E(S : S^c)$ is the number of edges with one endpoint in S , and one in S^c , and we use the convention $0/0 = \infty$.

- ii. For D -dimensional grid on n vertices, the cover time is $\mathcal{O}(D^2 n^{1/D})$, and this result is tight for D constant.
- iii. For general n -vertex graphs, the cover time is $\mathcal{O}(n^{11/4} \log n)$.

The main contribution of this chapter is to show improvements on the previous bounds on the cover time of the COBRA process.

For arbitrary connected graphs, we improve the $\mathcal{O}(n^{11/4} \log n)$ bound to $\mathcal{O}(m + (d_{\max})^2 \log n) = \mathcal{O}(n^2 \log n)$, where $d_{\max}$ is the maximum vertex degree, and m is

¹With high probability (w.h.p.), with probability at least $1 - n^{-c}$, for some constant $c > 0$.

the number of edges of the graph. Also, we improve the $\mathcal{O}(n^{11/4} \log n)$ bound to $\mathcal{O}(m \log n)$.

For an r -regular connected graph G , we show two bounds: $\mathcal{O}((r^2 + r/(1 - \lambda)) \log n)$ and $\mathcal{O}(1/(1 - \lambda)^2 \log n)$, where λ is the second largest eigenvalue of the transition matrix of the random walk on G . The former bound improves the latter for the case when $1 - \lambda = o(1/\sqrt{r})$. Since $1 - \lambda \geq \phi^2/2$, the former bound for regular graphs improves also the $\mathcal{O}((r^4/\phi^2) \log^2 n)$ bound given in [62].

Our main tool in analysing the COBRA process is a duality relation between this process and a particular discrete voting or epidemic process, which we call the *biased infection with persistent source* (BIPS) process. This process is interesting by itself, and the establishment of the duality between COBRA and BIPS can be considered as one of the main contributions of this chapter. The BIPS process is described as follows. A fixed vertex v is the source of an infection and remains permanently infected. At each time-step, each vertex u , other than v , selects k neighbours, independently and uniformly, then u is infected the next time-step if and only if at least one of the selected neighbours is infected. This can be seen as a voting model that gives preferences to the opinion “infected” over the opinion “healthy”.

3.2 Definitions and Contributions

Consider a connected graph $G = (V, E)$, an integer $k \geq 1$, and a subset of vertices $W \subseteq V$. The COBRA process, with starting set W and branching factor k , is the set-process $(W_t)_{t \geq 0}$ with $W_0 = W$, and the set W_{t+1} generated as follows. Each vertex $v \in W_t$ independently chooses k neighbours uniformly at random with replacement. Denote such a set of neighbours by $Y(v)$, then, W_{t+1} is defined by $W_{t+1} = \bigcup_{v \in W_t} Y(v)$. Note that a vertex in W_t does not necessarily belong to W_{t+1} . We can think of W_t as a set of vertices carrying a piece of information, and then each of them passes the information to k random neighbours to generate the new set of informed vertices W_{t+1} . For $W_0 = \{u\}$, let $\mathbf{cov}(u) = \min\{T : \bigcup_{t=0}^T W_t = V\}$ be the first round such

that each vertex has been informed at least once starting from $W_0 = \{u\}$.

We proceed to present the main results of this chapter.

Theorem 33. *Let G be a connected graph with n vertices, m edges and maximum vertex degree $d_{\max}$. For the COBRA process with branching factor $k = 2$, and for each $u \in V$, $\mathbf{cov}(u)$ is*

$$\mathcal{O}(m + (d_{\max})^2 \log n), \quad (3.1)$$

and

$$\mathcal{O}(m \log n), \quad (3.2)$$

w.h.p. and in expectation, where the hidden constants in the $\mathcal{O}$ notation do not depend on vertex u .

For a connected r -regular graph G with adjacency matrix $A(G)$, let $P = A(G)/r$ denote the transition matrix of the random-walk on G . Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigenvalues of matrix P ordered in a non-increasing sequence. Thus $\lambda_1 = 1 > \lambda_2 \geq \dots, \geq \lambda_n \geq -1$. Let $\lambda = \lambda(G) = \max_{i=2, \dots, n} |\lambda_i|$ be the second largest eigenvalue (in absolute value). Our second result gives a bound on the cover time of COBRA for regular graphs in terms of the *eigenvalue gap* $1 - \lambda$, and the degree r . We note that, for a connected graph, $1 - \lambda > 0$ if and only if the graph is not bipartite. Then the same bound can be derived for bipartite graphs if we consider the *lazy* COBRA process, which allows each vertex to also select itself with probability $1/2$. For the lazy process, we have to consider the transition matrix of a lazy random walk, whose eigenvalues are all greater than or equal to 0. In any case, the following theorem holds.

Theorem 34. *Let G be a connected regular n -vertex graph with eigenvalue gap $1 - \lambda \gg \sqrt{(\log n)/n}$. Then, for the COBRA process with branching factor $k = 2$, and for each $u \in V$, it holds that*

$$\mathbf{cov}(u) = \left(\left(\frac{r}{1 - \lambda} + r^2 \right) \cdot \log n \right), \quad (3.3)$$

and

$$\mathbf{cov}(u) = \mathcal{O}\left(\frac{\log n}{(1-\lambda)^2}\right) \quad (3.4)$$

w.h.p. and in expectation, where the hidden constants in the $\mathcal{O}$ notation do not depend on vertex u .

Remark 35. *All our results are given for $k = 2$. We decide to fix $k = 2$ because for greater values of k , the cover time is smaller (as the process inform more vertices per round), so $k = 2$ provides an interesting upper bound for the cover time for all $k \geq 2$. The case $k = 1$ corresponds to a random walk on the graph, and we have nothing new to say about it.*

The COBRA process can be seen as a type of multiple random walks processes, so it is tempting to try to analyse it using techniques developed for such processes. Previous work on multiple random walks include [4, 17, 23, 40], where cover times were analysed for various classes of graphs. The analyses of the COBRA process given in Dutta et al. [39] and Mitzenmacher et al. [62] use a number of tools from multiple random walks, but the applicability of those tools turns out to be limited because the random walks in COBRA are highly dependent. In this regard, we propose an alternative approach. Instead of directly analysing the COBRA walks, we analyse a related epidemic process, called BIPS. We show that the BIPS and the COBRA process are dual under time reversal, and thus properties of one process can be obtained by studying related properties in the other process.

Biased Infection with Persistent Source (BIPS): Consider a connected graph $G = (V, E)$ and an integer $k \geq 1$. Also, consider a special vertex v , which is the source of an infection. We consider the set-process $(A_t)_{t \geq 0}$ defined as follows. Let $A_0 = \{v\}$. Given A_t , each vertex $u \in V$ other than v , selects independently and uniformly with replacement k neighbours, and becomes a member of A_{t+1} if and only if at least one of the k selected neighbours is in A_t . Additionally, $v \in A_t$ for all $t \geq 0$. We call A_t the infected set at time t . Observe that the source v is always infected.

Finally, for ease of notation, stating that $A_0 = \{v\}$ is equivalent to stating that v is the source of the infection process.

The BIPS process is a discrete epidemic process of the SIS (Susceptible-Infected-Susceptible) type, in which vertices (other than the source v) refresh their infected state at each step by contacting k randomly chosen neighbours. The presence of a persistent (or corrupted) source means that, almost surely, all vertices of the underlying graph eventually become infected. The BIPS process is of independent interest since in the context of epidemics, certain viruses exhibit the property that a particular host can remain persistently infected.

Our main results for the COBRA process follow from the duality relationship between COBRA and BIPS. To avoid confusion between the BIPS process $(A_t)_{t \geq 0}$ and the COBRA process $(W_t)_{t \geq 0}$, we use the notation $\mathbb{P}(\cdot)$ for probabilities in the BIPS process, and $\widehat{\mathbb{P}}(\cdot)$ in the COBRA process. Additionally, for the COBRA process, let $\text{Hit}(v) = \min\{t \geq 0 : v \in W_t\}$ be the first time v is visited.

Theorem 36. *Let G be a connected graph. Consider a COBRA process $(W_t)_{t \geq 0}$ and a BIPS process $(A_t)_{t \geq 0}$, both with parameter $k \geq 1$. Then, for each $v \in V$, $W \subseteq V$, and $t \geq 0$, we have*

$$\widehat{\mathbb{P}}(\text{Hit}(v) > t | W_0 = W) = \mathbb{P}(W \cap A_t = \emptyset | A_0 = \{v\}).$$

For the BIPS process, we define $\mathbf{infec}(v)$ as the first time all vertices are infected starting from a source v . We prove the following two theorems about the BIPS process.

Theorem 37. *Let G be a connected graph with n vertices, m edges and maximum vertex degree $d_{\max}$. For every $v \in V$, the infection time $\mathbf{infec}(v)$ of the BIPS process with $k = 2$ is*

$$\mathcal{O}(m + (d_{\max})^2 \log n), \tag{3.5}$$

and also

$$\mathcal{O}(m \log n), \tag{3.6}$$

with probability at least $1 - \mathcal{O}(1/n^3)$, where the hidden constants in the $\mathcal{O}$ notation do not depend on vertex v .

Theorem 38. *Let G be a connected r -regular n -vertex graph with $1 - \lambda \gg \sqrt{\log n/n^2}$. Then for every $v \in V$, the infection time $\mathbf{infec}(v)$ of the BIPS process with $k = 2$, is*

$$\mathcal{O}\left(\left(\frac{r}{1-\lambda} + r^2\right) \log n\right) \quad (3.7)$$

and

$$\mathcal{O}\left(\frac{\log n}{(1-\lambda)^2}\right) \quad (3.8)$$

with probability at least $1 - \mathcal{O}(1/n^3)$, where the hidden constants in the $\mathcal{O}$ notation do not depend on vertex v .

Theorems 33 and 34 follow from Theorems 37 and 38, respectively, and from Theorem 36. Then for any two vertices $u, v \in V$, and any $T \geq 0$, applying Theorem 36 with $W_0 = \{u\}$ gives

$$\begin{aligned} \widehat{\mathbb{P}}(\text{Hit}(v) > T | W_0 = u) &= \mathbb{P}(u \notin A_T | A_0 = v) \\ &\leq \mathbb{P}(A_T \neq V | A_0 = v) \\ &= \mathbb{P}(\mathbf{infec}(v) > T). \end{aligned} \quad (3.9)$$

Also, Theorem 37 states that there is a constant $c > 0$ such that for $T = c(m + (d_{\max})^2 \log n)$ we have $\mathbb{P}(\mathbf{infec}(v) > T) = \mathcal{O}(1/n^3)$, where the hidden constant in the $\mathcal{O}$ notation does not depend on v . Hence, by the union bound

$$\begin{aligned} \widehat{\mathbb{P}}(\mathbf{cov}(u) > T) &\leq \sum_{v \in V} \widehat{\mathbb{P}}(\text{Hit}(v) > T | W_0 = u) \\ &\leq \sum_{v \in V} \mathbb{P}(\mathbf{infec}(v) > T) = n\mathcal{O}(n^{-3}) = \mathcal{O}(n^{-2}). \end{aligned} \quad (3.10)$$

This proves the first part of Theorem 33, i.e., $\mathbf{cov}(u) \leq T$ with probability $1 - \mathcal{O}(1/n^2)$. The same argument can be used to prove the second part. To see that the

²This means that $\lim_{n \rightarrow \infty} \frac{\sqrt{\log n/n}}{1-\lambda} \rightarrow 0$

expected value of $\mathbf{cov}(u)$ is $\mathcal{O}(T)$, consider restarting the COBRA process after T steps from any vertex in the current set W_T , and stop the process if all vertices have been covered. If the graph has not yet been covered, look again in T steps. Repeat until the graph is covered. A simple analysis shows that the expected cover time is $\mathcal{O}(T)$. We obtain Theorem 34 from the corresponding Theorem 38 in an analogous way.

The rest of this chapter is as follows. First, we prove the dual relationship. Later, we study the incremental nature of the BIPS process in order to give a proof of Theorem 37. We finalise by studying the BIPS process on regular graphs, and then proving Theorem 38. This is done in two steps. We first analyse infections with a large number of infected vertices, and then infections with a small number of infected vertices.

Notation remainder. $G = (V, E)$ stands for a graph. All graphs are connected, and we assume $|V| = n$. For $v \in V$, we denote by $N(v)$ the neighbourhood of v and, in general, for $A \subseteq V$, we define $N(A) = \bigcup_{v \in A} N(v)$. For $v \in V$, we define the degree of v , $d(v)$, as $|N(v)|$, and we denote the maximum degree by $d_{\max}$. Moreover, given $X \subseteq V$, we define $d(X)$ as the sum of the degrees of the vertices in X . Given a set X , we define $d_X(u)$ as the number of neighbour of u in X , i.e. $d_X(u) = |N(u) \cap X|$. We usually denote by $(W_t)_{t \geq 0}$ the COBRA process and by $(A_t)_{t \geq 0}$ the BIPS process. Finally, $\mathbf{infec}(v)$ is the random variable representing the time the BIPS process, starting from source v , needs to infect the whole graph.

3.3 Duality Between COBRA and BIPS Processes

We proceed to prove Theorem 36. Recall that $(W_t)_{t \geq 0}$ and $(A_t)_{t \geq 0}$ denote the COBRA and BIPS processes, respectively, and that we use the notation $\mathbb{P}(\cdot)$ for probabilities in the BIPS process, and $\widehat{\mathbb{P}}(\cdot)$ for probabilities in the COBRA process. Moreover, to simplify notation, we will write “ $A_0 = v$ ” for the frequently appearing condition “ $A_0 = \{v\}$.”

Theorem 36. *Let G be a connected graph. Consider a COBRA process $(W_t)_{t \geq 0}$ and a BIPS process $(A_t)_{t \geq 0}$, both with parameter $k \geq 1$. For the COBRA process, let $\text{Hit}(v) = \min\{t \geq 0 : v \in W_t\}$. Then, for each $v \in V$, $W \subseteq V$, and $t \geq 0$, we have*

$$\widehat{\mathbb{P}}(\text{Hit}(v) > t | W_0 = W) = \mathbb{P}(W \cap A_t = \emptyset | A_0 = \{v\}).$$

Proof. Observe that the claim is trivial if $v \in W$, since both probabilities are 0. We assume that $v \notin W$, and proceed by induction on t . For $t = 0$, the claim is true because both probabilities are 1. Assume the claim is true for a fixed $t \geq 0$, we will prove it for $t + 1$.

Consider the BIPS process at step $t + 1$. Denote by B_x the random k -set of neighbours chosen by vertex x . Then $x \in A_{t+1}$ if and only if $A_t \cap B_x \neq \emptyset$. For convenience, we set $B_v = \{v\}$. For a subset $S \subseteq V$, define $X(S) = \bigcup_{x \in S} B_x$. It is an assumption of the model that, at step $t + 1$, for any fixed set $B \subseteq V$, the event $\{X(S) = B\}$ is independent of A_t , and thus of the event $\{B \cap A_t = \emptyset\}$. Also $X(S)$ is independent of A_0 . Then

$$\begin{aligned} \mathbb{P}(B \cap A_t = \emptyset, X(W) = B | A_0 = v) \\ = \mathbb{P}(B \cap A_t = \emptyset | A_0 = v) \mathbb{P}(X(W) = B). \end{aligned}$$

Note that for any $S \subseteq V$, the following event equalities hold

$$\begin{aligned} \{S \cap A_{t+1} = \emptyset\} &= \left\{ \bigcap_{x \in S} \{B_x \cap A_t = \emptyset\} \right\} = \left\{ \left(\bigcup_{x \in S} B_x \right) \cap A_t = \emptyset \right\} \\ &= \{X(S) \cap A_t = \emptyset\}. \end{aligned}$$

Therefore,

$$\begin{aligned} \mathbb{P}(W \cap A_{t+1} = \emptyset | A_0 = v) &= \mathbb{P}(X(W) \cap A_t = \emptyset) \\ &= \sum_{B: B \subseteq V} \mathbb{P}(B = A_t \neq \emptyset, X(W) = B | A_0 = v) \\ &= \sum_{B: B \subseteq V} \mathbb{P}(B \cap A_t = \emptyset | A_0 = v) \mathbb{P}(X(W) = B). \end{aligned} \tag{3.11}$$

For any set B , the induction hypothesis gives

$$\mathbb{P}(B \cap A_t = \emptyset | A_0 = v) = \widehat{\mathbb{P}}(\text{Hit}(v) > t | W_0 = B).$$

By substituting the above into Equation (3.11), we get

$$\mathbb{P}(W \cap A_{t+1} = \emptyset | A_0 = v) = \sum_{B: B \subseteq V} \widehat{\mathbb{P}}(\text{Hit}(v) > t | W_0 = B) \mathbb{P}(X(W) = B) \quad (3.12)$$

Consider round 0 in the COBRA process, and recall that $W_0 = W$. For any vertex $u \in W$, let $Y(u)$ be the random set chosen (to push to) by u in the first round of the COBRA process, and define $Y(W) = \bigcup_{u \in W} Y(u)$. Observe that $W_1 = Y(W)$. As $v \notin W$, for any $u \in W$, it holds that for any $B \subseteq V$,

$$\mathbb{P}(X(u) = B) = \widehat{\mathbb{P}}(Y(u) = B).$$

Note that the variables $X(u)$ are independent for $u \in W$, and that the same holds for the variables $Y(u)$. Hence, we have that for any $B \subseteq V$,

$$\mathbb{P}(X(W) = B) = \widehat{\mathbb{P}}(Y(W) = B).$$

By substituting the last equality in equation (3.12), we get

$$\begin{aligned} \mathbb{P}(W \cap A_{t+1} = \emptyset | A_0 = v) &= \sum_{B: B \subseteq V} \widehat{\mathbb{P}}(\text{Hit}(v) > t | W_0 = B) \mathbb{P}(X(W) = B) \\ &= \sum_{B: B \subseteq V} \widehat{\mathbb{P}}(\text{Hit}(v) > t | W_0 = B) \widehat{\mathbb{P}}(Y(W) = B) \\ &= \widehat{\mathbb{P}}(\text{Hit}(v) > t + 1 | W_0 = W). \end{aligned}$$

□

3.4 Analysis of the BIPS process

3.4.1 Incremental Nature of BIPS

In this section, we consider the BIPS process with source v on any connected graph, not necessarily regular. To study the BIPS process, instead of tracking the

infected set $(A_t)_{t \geq 0}$, we will track its degree, $d(A_t)$. Given an infected set $A \subseteq V$, we define the subset of vertices $B = B(A)$ and $C = C(A)$ as

$$\begin{aligned} B &= \{u \in V : N(u) \subseteq A\}, \\ C &= (N(A) \cup \{v\}) \setminus B, \end{aligned} \tag{3.13}$$

The set B contains, possible with the exception of the source v , all vertices that will be surely infected in the next round (because all their neighbours are infected). On the other hand, the set C represents the vertices that might be infected, depending on their (random) choice. Notice that if a vertex does not belong to $B \cup C$, then it is surely not infected in the next round.

For an infected set A_t , we denote by B_t and C_t its associated sets $B(A_t)$ and $C(A_t)$, respectively. We claim that if $A_t \neq V$ then $C_t \neq \emptyset$. The reason for this is the following. If v does not belong to B_t , then $v \in C_t$. Suppose this is not the case. Let $A_t(v)$ be the sets of all vertices connected to v via a path whose vertices are in A_t . Clearly $A_t(v) \subseteq A_t \neq V$. By definition, it exists an edge (u, w) with $u \in A_t(v)$ and $w \in A_t^c$. Then $u \notin B_t$ but $u \in N(A_t)$, then $u \in C_t$. We state the claim above as the next lemma.

Lemma 39. *Let $(A_t)_{t \geq 0}$ be a BIPS process, then the event $\{A_t \neq V\}$ implies the event $\{C_t \neq \emptyset\}$*

With the definitions of B_t and C_t , we analyse the difference $d(A_{t+1}) - d(A_t)$. Observe the following relation between A_t , B_t , and C_t ,

$$d(B_t) + \sum_{u \in C_t} d_{A_t}(u) = \sum_{u \in V} d_{A_t}(u) = \sum_{u \in A_t} d(u) = d(A_t). \tag{3.14}$$

The first equality holds because only vertices in B_t and C_t have neighbours in A_t , and also because all the neighbours of vertices in B_t belong to A_t .

Let $X_{t,u}$ be the indicator variable taking value 1 if and only if $u \in A_t$, then

$$\begin{aligned}
d(A_{t+1}) &= \sum_{u \in V} d(u)X_{t+1,u} \\
&= \sum_{u \in B_t} d(u) + \sum_{u \in C_t} d(u)X_{t+1,u} \\
&= d(B_t) + \sum_{u \in C_t} d(u)X_{t+1,u} \\
&= d(A_t) - \sum_{u \in C_t} d_{A_t}(u) + \sum_{u \in C_t} d(u)X_{t+1,u} \tag{3.15} \\
&= d(A_t) + \sum_{u \in C_t} (d(u)X_{t+1,u} - d_{A_t}(u)). \tag{3.16}
\end{aligned}$$

Equation (3.15) follows from our observation in Equation (3.14).

Lemma 40. *Let $A \subset V$, such that $A \neq V$, yet $v \in A$. Then*

$$\mathbb{E}(d(A_{t+1}) - d(A_t) | A_t = A) = \sum_{u \in C_t} \left(d_A(u) - \frac{d_A(u)^2}{d(u)} \right) \geq \frac{|C_t|}{2}.$$

Proof. Given $A_t = A$, we have that the random variables $X_{t+1,u}$ are independent of each other. Moreover, for $u \neq v$, it holds

$$\mathbb{P}(X_{t+1,u} = 0 | A_t = A) = \left(1 - \frac{d_A(u)}{d(u)} \right)^2. \tag{3.17}$$

The above holds because $X_{t+1,u} = 0$, if and only if u chooses two neighbours outside A_t . If $u \in C_t \setminus \{v\}$, we have that

$$\begin{aligned}
\mathbb{E}(d(u)X_{t+1,u} - d_A(u) | A_t = A) &= d(u) \left(1 - \left(1 - \frac{d_A(u)}{d(u)} \right)^2 \right) - d_A(u) \\
&= d(u) \left(2 \frac{d_A(u)}{d(u)} - \left(\frac{d_A(u)}{d(u)} \right)^2 \right) - d_A(u) \\
&= d_A(u) \left(1 - \frac{d_A(u)}{d(u)} \right) \geq \left(1 - \frac{1}{d(u)} \right) \geq \frac{1}{2}. \tag{3.18}
\end{aligned}$$

In the last two inequalities, we use that for $u \in C_t$ it holds $0 < d_A(u) < d(u)$, thus $d(u) \geq 2$, and that $d_A(u) \left(1 - \frac{d_A(u)}{d(u)} \right)$ is minimised when $d_A(u)$ is equal to 1 or

$d(u) - 1$. If $v \in C_t$ then $d(v)X_{t+1,v} - d_{A_t}(v) = d(v) - d_{A_t}(v) \geq 1$, because $d(v) > d_{A_t}(v)$, otherwise v belongs to B_t instead of C_t .

Summing up over all $u \in C_t$, and by Equation (3.18), the following holds

$$\mathbb{E}(d(A_{t+1}) - d(A_t) | A_t = A) = \mathbb{E} \left(\sum_{u \in C_t} (d(u)X_{t+1,u} - d_{A_t}(u)) \middle| A_t = A \right) \geq \frac{|C_t|}{2}. \quad (3.19)$$

□

Recall that $\mathbf{infec}(v)$ is the time it takes the BIPS process with source v to infect the whole graph G .

Theorem 41. *Let G be a connected graph with n vertices and m edges. Then for every $v \in V$, it holds that $\mathbb{E}(\mathbf{infec}(v)) = \mathcal{O}(m)$.*

Proof. Consider any $T > 0$. Then by Lemma 40, it holds that

$$2m \geq \mathbb{E}(d(A_T)) \geq d(A_0) + \sum_{k=0}^{T-1} \mathbb{E} \left(\frac{|C_k|}{2} \right). \quad (3.20)$$

We compute a lower bound for $\mathbb{E}(|C_k|)$. By Lemma 39, we have that if the event $\{\mathbf{infec}(v) > k\}$ holds then $\{|C_k| > 0\}$ holds as well. Hence,

$$\mathbb{E}(|C_k|) = \mathbb{E}(|C_k| | |C_k| \geq 1) \mathbb{P}(|C_k| \geq 1) + 0 \geq \mathbb{P}(|C_k| \geq 1) \geq \mathbb{P}(\mathbf{infec}(v) > k).$$

We conclude that for all $T \geq 1$,

$$4m \geq \sum_{k=0}^{T-1} \mathbb{P}(\mathbf{infec}(v) > k),$$

hence, by taking T tending to infinity, we deduce

$$4m \geq \sum_{k=0}^{\infty} \mathbb{P}(\mathbf{infec}(v) > k) = \mathbb{E}(\mathbf{infec}(v)).$$

□

The next corollary proves Equation (3.6) of Theorem 37 and, via duality (Theorem 36), Equation (3.2) of Theorem 33.

Corollary 42. *There exists a constant $c > 0$ such that*

$$\mathbb{P}(\mathbf{infec}(v) > cm \log n) \leq n^{-3},$$

for all large enough n .

Proof. By Markov's inequality

$$\mathbb{P}(\mathbf{infec}(v) \geq 2\mathbb{E}(\mathbf{infec}(v))) \leq 1/2.$$

Consider the following algorithm. At time $T = 2\mathbb{E}(\mathbf{infec}(v))$, we check if the process infected the whole graph. If so, we stop, otherwise we drop all the infection and restart the algorithm until it stops. Due to the fact that we infect the graph with probability at least $1/2$ by time $2\mathbb{E}(\mathbf{infec}(v))$, the probability that we iterate the algorithm more than $3\lceil \log_2 n \rceil$ times is at most n^{-3} . The conclusion follows from the fact that $\mathbb{E}(\mathbf{infec}(v)) \leq 4m$, as claimed in Theorem 41. $\square$

3.4.2 Sequential Analysis

The proof of Equation (3.5) of Theorem 37 requires a more subtle argument. One of the difficulties of studying the BIPS process is that the one-step difference, i.e., $d(A_{t+1}) - d(A_t)$, has a huge range. For instance, in the complete graph, the whole graph may become infected or healthy (by exception of the source) in one step. This discourages us from attempting to use a raw concentration inequality to prove that $\mathbf{infec}(v)$ is concentrated around its mean, and thus, by Theorem 41, $\mathbf{infec}(v) = \mathcal{O}(m)$. To face the problem above, one option is to study the actual distribution of $d(A_{t+1}) - d(A_t)$, nevertheless, this is very hard to do as a consequence of the non-monotonic behaviour of A_t . Instead we follow a different path, we use the fact that $d(A_{t+1}) - d(A_t)$, given A_t , can be written as the sum of independent random variables whose range is an interval of length at most $d_{\max}$. If the maximum degree $d_{\max}$, i.e., the range of these random variables is not too large, then we can prove that $\mathbf{infec}(v) = \mathcal{O}(m)$.

We proceed to formalise the idea above. Recall Equation (3.16) given by

$$d(A_{t+1}) = d(A_t) + \sum_{u \in C_t} (d(u)X_{t+1,u} - d_{A_t}(u)). \quad (3.21)$$

where $X_{t,u}$ indicates where $u \in A_t$, or not. By recursively applying Equation (3.21), for any $t \leq \mathbf{infec}(v)$, we have

$$d(A_t) = d(v) + \sum_{\tau=0}^{t-1} \sum_{u \in C_\tau} (d(u)X_{\tau+1,u} - d_{A_\tau}(u)). \quad (3.22)$$

From Lemma 39, it holds that $|C_\tau| > 0$ for all $0 \leq \tau < \mathbf{infec}(v)$, hence, for all $t \leq \mathbf{infec}(v)$, we write Equation (3.22) as

$$d(A_t) = d(v) + \sum_{l=1}^{\nu(t)} Y_l, \quad (3.23)$$

where, $\nu(0) = 0$, $\nu(t) \equiv \sum_{\tau=0}^{t-1} |C_\tau|$ for $t \geq 1$, and $Y_{\nu(\tau)+i} \equiv d(u)X_{\tau+1,u} - d_{A_\tau}(u)$, for $1 \leq i \leq |C_\tau|$, where u is the i -th smallest vertex of C_τ in some arbitrary but fixed ordering of the vertices V . Since $\nu(0) = 0$, and $1 \leq |C_\tau| \leq n$, we have

$$t \leq \nu(t) < \nu(t-1) + n. \quad (3.24)$$

We say that round t , with $t < \mathbf{infec}(v)$, consists of $|C_t|$ steps, with the random variable $Y_{\nu(t)+i}$ corresponding to the i -th step of this round. Thus, we can view the BIPS process as a sequence of single steps which are grouped into rounds.

Even though the BIPS process finishes at round $\mathbf{infec}(v)$, the sequence $(A_t)_{t \geq 0}$ is defined in the natural way for all $t \geq 1$. For $t \geq \mathbf{infec}(v)$, $A_t = V$ and thus $d(A_t) = 2m$. The sequence (Y_l) is defined for $1 \leq l < \nu(\mathbf{infec}(v))$, that is, until the completion of the BIPS process. For technical convenience, we set $Y_l = 1$ for all $l \geq \nu(\mathbf{infec}(v))$, so the process (Y_l) can be defined for all $l \geq 0$. The choice of the value 1 will become clear later.

Observe that the random variables Y_l are not independent. Indeed, the distribution of Y_l depends on the values of the variables Y_i , for $1 \leq i \leq l-1$. In this way, for any fixed $l \geq 1$, and an arbitrary sequence of numbers $y_1, y_2, \dots, y_{l-1}$, we have two

possibilities. Either the given sequence of numbers is not a feasible realisation of the sequence of random variables $Y_1, Y_2, \dots, Y_{l-1}$, or it is feasible realisation and shows in full the evolution of the BIPS process until step $l-1$, determining the distribution of the variable Y_l .

In particular, if $Y_1 = y_1, Y_2 = y_2, \dots, Y_{l-1} = y_{l-1}$, then by starting from the known initial sets A_0 and C_0 , and using the fact that the vertices of C_t are considered according to a fixed ordering of all vertices of V , we can keep track of the values of $Y_1, Y_2, \dots$ to identify the vertices in A_1 (which also gives the set C_1). This can also be done for the vertices in A_2, A_3 , and so on. Finally, either the process has completed before step l , so $Y_l \equiv 1$, or we identify the round t which includes step l . In the latter case, we would be able to recover the set $A_t \subsetneq V$ of vertices infected at round t , and the vertex u considered in step l . In both cases, we get the distribution of the random variable Y_l .

Equation (3.23) implies that instead of analysing the sequence $(d(A_t))_{t \geq 0}$, we can analyse the sequence of sums $R_q = \sum_{l=1}^q Y_l$, $q \geq 0$. There is a technical complication here because only for those $q = \nu(t)$, the value of R_q corresponds to the value of $d(A_t)$. This means that a large value of some R_q does not immediately guarantee a large value of $d(A_t)$. However, an appropriately long sequence $R_q, R_{q+1}, \dots, R_{q'}$ with large values would imply a large value of some $d(A_t)$ as there must be an index in $\{q+1, \dots, q'\}$ which corresponds to a value $\nu(t)$ for some t .

More precisely, we have the following relationship between the sequences $(d(A_t))_{t \geq 1}$ and $(R_q)_{q \geq 1}$. For each $1 \leq k \leq 2m - d(v)$, and each $t \geq 1$, the following relation holds

$$\{d(A_t) < d(v) + k\} \subseteq \{\exists q : t \leq q < tn \wedge R_q < k\}. \quad (3.25)$$

The above holds for the following reason. Consider an execution of the BIPS process such that $d(A_t) < d(v) + k$. From (3.23), $R_{\nu(t)} = \sum_{l=1}^{\nu(t)} Y_l < k$, and from (3.24), $t \leq \nu(t) < tn$. Thus $R_q < k$, for some $t \leq q < tn$.

We proceed to derive a lower bound on the conditional expectation of Y_l given

the values of the variables $Y_1, Y_2, \dots, Y_{l-1}$. If these values show that the BIPS process has already infected the whole graph (that is, $l > \nu(T)$), then $Y_l \equiv 1$ and $\mathbb{E}(Y_l|Y_1, Y_2, \dots, Y_{l-1}) = 1$. Otherwise, let u denote the vertex corresponding to Y_l , let t denote the index of the current round (that is, the round which includes step l), and let $A = A_t$ and $C = C_t$. As mentioned above, u , t and A_t (and thus C_t) are fully defined by the values of variables $Y_1, Y_2, \dots, Y_{l-1}$. If u is the source v , then $v \in C = C(A)$, and so $d_A(v) \leq d(v) - 1$, therefore $Y_l \geq 1$. If $u \neq v$, then

$$\begin{aligned} \mathbb{E}(Y_l|Y_1, Y_2, \dots, Y_{l-1}) &= d(u) \left(1 - \left(1 - \frac{d_A(u)}{d(u)} \right)^2 \right) - d_A(u) \\ &\geq 1 - \frac{1}{d(u)} \geq \frac{1}{2} \end{aligned} \quad (3.26)$$

The inequalities (explained in equation (3.18)) hold because $u \in C$. Therefore, in all cases, we have

$$\mathbb{E}(Y_l|Y_1, Y_2, \dots, Y_{l-1}) \geq \frac{1}{2}. \quad (3.27)$$

Lemma 43. *For any $q \geq 1$, it holds that*

$$\mathbb{P} \left(\sum_{l=1}^q Y_l \leq \frac{q}{2} - t \right) \leq \exp \left(-\frac{t^2}{8qd_{\max}^2} \right).$$

Proof. For $l \geq 1$ consider $W_l = Y_l - \mathbb{E}(Y_l|Y_1, \dots, Y_{l-1})$. From the definition of the random variables Y_l we have that $|W_l| \leq 2d_{\max}$. Also, it holds that

$$\mathbb{E}(W_l|Y_1, \dots, Y_{l-1}) = 0. \quad (3.28)$$

Therefore, $S_q = \sum_{l=1}^q W_l$ is a martingale with respect to the filtration induced by the variables (Y_l) . Also for $q \geq 1$, it holds that $|S_q - S_{q-1}| = |W_q| \leq 2d_{\max}$. From Azuma's inequality (Theorem 85 in Appendix A), it holds that

$$\mathbb{P}(S_q \leq -t) \leq \exp \left(\frac{-t^2}{8qd_{\max}^2} \right). \quad (3.29)$$

Thus with probability at least $1 - \exp(-t^2/(8qd_{\max}^2))$, we have

$$\sum_{l=1}^q Y_l > \sum_{l=1}^q \mathbb{E}(Y_l|Y_1, \dots, Y_{l-1}) - t \geq \frac{q}{2} - t. \quad (3.30)$$

The last inequality of (3.30) comes from (3.27). $\square$

The proof of Equation (3.5) of Theorem 37 follows from Lemma 44 (proved below) by choosing $k = 2m - d(v)$.

Lemma 44. *Consider the BIPS process on a connected graph with n vertices, m edges and the maximum vertex degree $d_{\max}$. For any constant $C > 0$, define $C' = 64(C+3)$. Then, for any $1 \leq k \leq 2m - d(v)$, and $t(k) = 4k + C'(d_{\max})^2 \log n$,*

$$\mathbb{P}(\exists t \geq t(k) : d(A_t) < d(v) + k) = \mathcal{O}(n^{-C}).$$

Proof. From Equation (3.25) and Lemma (43), it holds

$$\begin{aligned} \mathbb{P}(\exists t \geq t(k) : d(A_t) < d(v) + k) &= \mathbb{P}\left(\bigcup_{t \geq t(k)} \{\exists q : t \leq q \leq tn \wedge R_q < k\}\right) \\ &\leq \mathbb{P}\left(\bigcup_{q \geq t(k)} \{R_q < k\}\right) \\ &\leq \sum_{q=t(k)}^{\infty} \exp\left(-\frac{(q/2 - k)^2}{8qd_{\max}^2}\right) \\ &\leq \sum_{q=t(k)}^{\infty} \exp\left(-\frac{(q/2 - k)}{32d_{\max}^2}\right) \tag{3.31} \\ &\leq \sum_{j=0}^{\infty} \exp\left(-\frac{(C'/2)d_{\max}^2 \log(n) + j}{32d_{\max}^2}\right) \\ &\leq \mathcal{O}(d_{\max}^2) \exp\left(-\frac{C'd_{\max}^2 \log n}{64d_{\max}^2}\right) \\ &= \mathcal{O}(d_{\max}^2)n^{-(C+3)} = \mathcal{O}(n^{-C}). \tag{3.32} \end{aligned}$$

The inequality in (3.31) holds because $(q/2 - k)/q \geq 1/4$ for all $q \geq t(k)$. $\square$

3.5 The BIPS Process on Regular Graphs

In the analysis of the BIPS process on regular graphs, we track the size of the current infection set rather than the degree of this set. This analysis is done in

two phases. The first phase deals with small infection sizes, while the second phase considers large infections.

We begin with analysing the second phase since it is easier to deal with.

3.5.1 Large Infection Size

We begin our analysis by giving a lower bound of the size of $|A_{t+1}|$ given A_t .

Lemma 45. *Let G be a connected r -regular graph on n vertices, with $\lambda < 1$ where λ is the second absolute eigenvalue of the random-walk transition matrix. Let A_t be the size of the infected set after step t of the BIPS process with $k = 2$, then*

$$\mathbb{E}(|A_{t+1}| \mid A_t = A) \geq |A|(1 + (1 - \lambda^2)(1 - |A|/n)). \quad (3.33)$$

Proof. A direct computation gives us

$$\begin{aligned} \mathbb{E}(|A_{t+1}| \mid A_t = A) &= 1 + \sum_{u \in V \setminus \{v\}} \mathbb{P}(X_{t+1,u} = 1 \mid A_t = A) \\ &= 1 + \sum_{u \in V \setminus \{v\}} (1 - (1 - d_A(u)/r)^2) \\ &\geq \sum_{u \in V} (1 - (1 - d_A(u)/r)^2) \\ &= \sum_{u \in V} \left(2 \frac{d_A(u)}{r} - \frac{d_A(u)^2}{r^2} \right) \\ &= 2A - \sum_{u \in V} \frac{d_A(u)^2}{r^2}. \end{aligned} \quad (3.34)$$

In the last step, we use the fact that $\sum_{u \in V} d_A(u) = \sum_{u \in A} d(u) = r|A|$. Let $P = P(G)$ be the transition matrix of a simple random walk on G . Let $P(x, A) = \sum_{y \in A} P(x, y) = d_A(x)/r$. From (3.34), we have

$$\mathbb{E}(|A_{t+1}| \mid A_t = A) \geq 2A - \sum_{x \in V} P(x, A)^2. \quad (3.35)$$

Observe that $\sum_{x \in V} P(x, A)^2 = \langle P\mathbb{1}_A, P\mathbb{1}_A \rangle = \|P\mathbb{1}_A\|_2^2$, where $\mathbb{1}_A$ is the characteristic (indicator) vector of A . $P\mathbb{1}_A$ is the standard matrix-vector product. As P is

symmetric, it has an orthonormal basis of right eigenvectors $f_1, \dots, f_n$, i.e., $\|f_i\|_2 = 1$, $\langle f_i, f_j \rangle = 0$ for $i \neq j$. For any vector g , $g = \sum_{i=1}^n \langle g, f_i \rangle f_i$ and $\|g\|_2^2 = \sum_{i=1}^n \langle g, f_i \rangle^2$. Here $f_1 = (1/\sqrt{n})$ is the unique eigenvector with eigenvalue 1, and $\langle \mathbb{1}_A, f_1 \rangle = A/\sqrt{n}$. Thus,

$$\begin{aligned}
\|P\mathbb{1}_A\|_2^2 &= \|P \sum_{i=1}^n \langle \mathbb{1}_A, f_i \rangle f_i\|_2^2 = \|\sum_{i=1}^n \langle \mathbb{1}_A, f_i \rangle P f_i\|_2^2 \\
&= \|\sum_{i=1}^n \langle \mathbb{1}_A, f_i \rangle \lambda_i f_i\|_2^2 = \sum_{i=1}^n \langle \mathbb{1}_A, f_i \rangle^2 \lambda_i^2 \|f_i\|_2^2 \\
&\leq (1 - \lambda^2) \langle \mathbb{1}_A, f_1 \rangle^2 + \lambda^2 \sum_{i=1}^n \langle \mathbb{1}_A, f_i \rangle^2 \\
&= (1 - \lambda^2) \frac{|A|^2}{n} + \lambda^2 \|\mathbb{1}_A\|_2^2 \\
&= (1 - \lambda^2) \frac{|A|^2}{n} + \lambda^2 |A|. \tag{3.36}
\end{aligned}$$

Hence, equations (3.35) and (3.36) imply

$$\mathbb{E}(|A_{t+1}| \mid |A_t| = A) \geq 2|A| - \lambda^2 |A| - (1 - \lambda^2) \frac{|A|^2}{n},$$

which is equivalent to (3.33). $\square$

A direct application of the lemma above allows us to analyse the second phase of the process which begins when the number of infected vertices is greater than $K \log n / (1 - \lambda)^2$, for a large enough constant $K > 0$. Indeed, the first lemma below considers the case when the number of the infected vertices is between $K \log(n) / (1 - \lambda)^2$ and $9n/10$, and the second lemma, the case when the number of infected vertices is at least $9n/10$.

Lemma 46. *Let G be a connected r -regular n -vertex graph, and consider the BIPS process on G from some step $t > 0$. There exist constants $c > 0$ and $K > 0$ such that, if $1 - \lambda \geq c\sqrt{\log n/n}$ and $|A_t| \geq K \log n / (1 - \lambda)^2$ then $9/10$ of the whole graph is infected within $\mathcal{O}(\log n / (1 - \lambda))$ additional steps with probability at least $1 - n^{-4}$.*

Proof. Assume A_t has size less or equal than $9n/10$ but greater than $K \log n / (1 -$

$\lambda)^2$, then from Lemma 45,

$$\begin{aligned}\mathbb{E}(A_{t+1}|A_t) &\geq A_t(1 + (1 - \lambda^2)(1 - 9/10)) \\ &\geq A_t \left(1 + \frac{1 - \lambda}{10}\right).\end{aligned}$$

Let $\varepsilon = \sqrt{10 \log n / A_t}$. Remember that, given A_t , the size of A_{t+1} is the sum of independent Bernoulli random variables (with potentially different parameters). By using the Chernoff bound (Theorem 84, in appendix A) for the lower tail of the sum of Bernoulli random variables, we get

$$\begin{aligned}\mathbb{P}(A_{t+1} < (1 - \varepsilon)\mathbb{E}(A_{t+1}|A_t)|A_t) &\leq e^{-\varepsilon^2 \mathbb{E}(A_{t+1}|A_t)/2} \\ &= \exp\left(-\frac{5 \log n}{A_t} \mathbb{E}(A_{t+1}|A_t)\right) \\ &\leq \exp\left(-5 \log n \left(1 + \frac{1 - \lambda}{10}\right)\right) \leq \frac{1}{n^5}.\end{aligned}\tag{3.37}$$

Choose $K = 4000$. By hypothesis $A_t \geq K \log n / (1 - \lambda)^2$, so $\varepsilon \leq (1 - \lambda)/20$. Therefore, with probability at least $1 - n^{-5}$ we have

$$\begin{aligned}A_{t+1} &\geq (1 - \varepsilon)\mathbb{E}(A_{t+1}|A_t) \\ &\geq A_t \left(1 + \frac{1 - \lambda}{10}\right) \left(1 - \frac{1 - \lambda}{20}\right) \\ &\geq A_t \left(1 + \frac{1 - \lambda}{23}\right).\end{aligned}$$

Finally, we have that after $23/(1 - \lambda)$ rounds, the size of infection has at least doubled. Hence, with probability at least $1 - 23(\log n)n^{-5}/(1 - \lambda) \geq 1 - n^{-4}$, after $\mathcal{O}(\log n/(1 - \lambda))$ rounds, the infection covers at least $9n/10$ vertices. $\square$

Lemma 47. *Let G be a connected n -vertex r -regular graph with $1 - \lambda \geq c\sqrt{\log n/n}$, for a suitably large constant c . If the BIPS process starts with at least $(9/10)n$ infected vertices, then with probability at least $1 - n^{-5}$ the whole graph is infected within $T \leq 8 \log n / (1 - \lambda)$ rounds.*

Proof. For convenience, let A_0 and B_0 be the size of the infected and non-infected sets at the beginning of this phase, and let us set $q = 9/10$. Hence $A_0 \geq qn$. Let A_t and B_t be their respective sizes after t rounds. From (3.33), we get

$$\mathbb{E}(A_{t+1}|A_t = A) \geq A + (n - A)(1 - \lambda^2)A/n. \quad (3.38)$$

The corresponding inequality for B_{t+1} is

$$\begin{aligned} \mathbb{E}(B_{t+1}|B_t) &\leq B_t - B_t(1 - \lambda^2)A_t/n \\ &= B_t(1 - (1 - \lambda^2)A_t/n). \end{aligned} \quad (3.39)$$

Let $|A_t| = k$. By applying the law of total probability and Equation (3.39), we get

$$\begin{aligned} \mathbb{E}(B_{t+1}) &= \sum_{k=qn}^n \mathbb{E}(B_{t+1}|B_t = n - k)\mathbb{P}(B_t = n - k) + \mathbb{E}(B_{t+1}|A_t < qn)\mathbb{P}(A_t < qn) \\ &\leq \sum_{k=qn}^n (n - k)(1 - (1 - \lambda^2)k/n)\mathbb{P}(B_t = n - k) + n\mathbb{P}(A_t < qn) \\ &\leq \sum_{k=qn}^n (n - k)(1 - (1 - \lambda^2)q)\mathbb{P}(B_t = n - k) + n\mathbb{P}(A_t < qn) \\ &\leq (1 - (1 - \lambda^2)q)\mathbb{E}(B_t) + n\mathbb{P}(A_t < qn). \end{aligned} \quad (3.40)$$

We next prove that

$$\mathbb{P}(A_t < qn) \leq tn^{-8}. \quad (3.41)$$

To check that the last inequality holds, consider the event $E_t = \{A_i \geq qn, i = 0, \dots, t\}$. We are going to prove that E_t has high probability. Indeed

$$\begin{aligned} \mathbb{P}(E_t) &= \mathbb{P}(E_t|E_{t-1})\mathbb{P}(E_{t-1}) + \mathbb{P}(E_t|E_{t-1}^c)\mathbb{P}(E_{t-1}^c) \\ &\geq \mathbb{P}(E_t|E_{t-1})\mathbb{P}(E_{t-1}). \end{aligned}$$

Observe that A_t depends only on A_{t-1} , since it is a Markov chain, then

$$\mathbb{P}(E_t|E_{t-1}) = \mathbb{P}(A_t \geq qn|A_{t-1} \geq qn),$$

and by a standard coupling argument

$$\mathbb{P}(A_t \geq qn | A_{t-1} \geq qn) \geq \mathbb{P}(A_t \geq qn | A_{t-1} = \lceil qn \rceil).$$

Choose $\varepsilon = \sqrt{16 \log n / qn}$, then by the Chernoff bound

$$\begin{aligned} \mathbb{P}(A_{t+1} < (1 - \varepsilon)\mathbb{E}(A_{t+1} | A_t = \lceil qn \rceil) | A_t = \lceil qn \rceil) &\leq e^{-\varepsilon^2 \mathbb{E}(A_{t+1} | A_t = \lceil qn \rceil) / 2} \\ &= e^{-8 \log n} = \frac{1}{n^8}. \end{aligned}$$

Since we assume that $1 - \lambda \geq c\sqrt{\log n / n}$ for a suitably large constant c , we have $A_t = qn \geq K \log n / (1 - \lambda)^2$, with $K = 4000$, so $\varepsilon \leq (1 - \lambda) / 15$. Thus, with probability at least $1 - n^{-8}$, we have

$$\begin{aligned} A_{t+1} &\geq (1 - \varepsilon)\mathbb{E}(A_{t+1} | A_t = \lceil qn \rceil) \\ &\geq qn \left(1 + \frac{1 - \lambda}{10}\right) \left(1 - \frac{1 - \lambda}{15}\right) \geq qn. \end{aligned}$$

Therefore, $\mathbb{P}(A_t \geq qn | A_{t-1} = \lceil qn \rceil) \geq 1 - n^{-8}$. We conclude that

$$\mathbb{P}(E_t) \geq (1 - n^{-8})^t \geq 1 - tn^{-8}.$$

Observe that $\mathbb{P}(A_t \geq qn) \geq \mathbb{P}(E_t) \geq 1 - tn^{-8}$, so Equation (3.41) holds.

Let us return to Equation (3.40). By using equation (3.41), we have

$$\mathbb{E}(B_{t+1}) \leq (1 - (1 - \lambda^2)q)\mathbb{E}(B_t) + tn^{-7}. \quad (3.42)$$

Denote $\theta = 1 - (1 - \lambda^2)q$, then by iterating Equation (3.42), and by using $B_0 \leq (1 - q)n$, we get

$$\mathbb{E}(B_t) \leq \theta^t(1 - q)n + O(t^2 n^{-7}) \leq n\theta^t + O(t^2 n^{-7}).$$

Choosing $T = 6 \log n / \log(1/\theta)$, and applying Markov's inequality, we get

$$\begin{aligned} \mathbb{P}(B_T \geq 1) &\leq \mathbb{E}(B_T) \leq n\theta^T + O(T^2 n^{-7}) \\ &= n^{-5} + O(T^2 n^{-7}). \end{aligned} \quad (3.43)$$

We just need to prove that $T^2 n^{-7} = \mathcal{O}(n^{-5})$. For that we check that $T = \mathcal{O}(n)$. Observe that for $0 < \theta < 1$, we have $1 - \theta \leq \log(1/\theta)$, and thus

$$\begin{aligned} T &= 6 \log n / (\log(1/\theta)) \leq 6 \log n / (1 - \theta) \\ &\leq 6 \log n / (q(1 - \lambda^2)) \leq 6 \log n / (q(1 - \lambda)) \\ &\leq \mathcal{O}(\log n / (1 - \lambda)) = \mathcal{O}(n), \end{aligned} \tag{3.44}$$

where the last bound follows from the assumption $1 - \lambda \geq c\sqrt{(\log n)/n}$. We conclude that $\mathbb{P}(B_T \geq 1) = \mathcal{O}(n^{-5})$ from equations (3.43) and (3.44). $\square$

In order to apply the two lemmas above together, we do the following. Start a BIPS process with $|A_0| \geq K \log n / (1 - \lambda)$ where K is the constant of Lemma 46. Let T be the first time such that $|A_T| \geq 9n/10$. From Lemma 46, we know that there exists a large constant C , such that $T \leq C \log n / (1 - \lambda)$ with probability at least $1 - n^{-3}$.

Let $\mathcal{H} = \{A \subseteq V : |A| \geq 9n/10\}$ and let $t \geq 8 \log n / (1 - \lambda)$, then

$$\begin{aligned} \mathbb{P}(A_{T+t} = V | A_T) &= \sum_{A \in \mathcal{H}} \mathbb{P}(A_{T+t} = V | A_T = A) \mathbb{P}(A_T = A) \\ &= \sum_{A \in \mathcal{H}} \mathbb{P}(A_t = V | A_0 = A) \mathbb{P}(A_T = A) \end{aligned} \tag{3.45}$$

$$\geq \sum_{A \in \mathcal{H}} (1 - n^{-5}) \mathbb{P}(A_T = A) = 1 - n^{-5}. \tag{3.46}$$

Equation (3.45) holds true as a consequence of the strong Markov property (Theorem 87 in Appendix A). Equation (3.46) follows from Lemma 47. We conclude that it takes $T + t = \mathcal{O}(\log n / (1 - \lambda))$ steps to finish the process with probability at least $1 - 2n^{-3}$.

Corollary 48. *Suppose that we start a BIPS process with infection size $|A_0| \geq K \log n / (1 - \lambda)^2$ where K is large enough, then the process infects the whole graph in $\mathcal{O}(\log n / (1 - \lambda))$ rounds with probability at least $1 - 2n^{-3}$.*

3.5.2 Small Infections

As seen in Corollary 48, when the size of the infection is fairly large, it is not so difficult to prove that the whole graph is infected in $\mathcal{O}(\log n/(1-\lambda))$ rounds. In this section, we prove that in $\mathcal{O}(\log n/(1-\lambda)^2)$ and $\mathcal{O}((r/(1-\lambda) + r^2) \log n)$ rounds the infection reaches the necessary size of Corollary 48, which leads us to the proof of Theorem 34.

The main difference between small and large infections is that, due to the concentration behaviour of large infections, it is fairly easy to track the size of A_t , and to prove that it increases in each round by a substantial amount. On the other hand, it is very hard to actually have track of A_t for small size. Indeed, when the size of the infection is rather small, in one step the infection can either grow or shrink depending on several factors, including randomness, graph structure, location of the infected vertices, etc.

In this section, we provide the analysis of the early stages of the BIPS process which will lead us to the proof of Theorem 38. The main part of the analysis is stated in Lemma 49, which gives us the proof of Equation (3.6) of Theorem 38. After that, we combine Lemmas 49 and 44 to provide a proof of Equation (3.5).

We begin our analysis by introducing the necessary notation. Define the quantity $\Delta = K \log n/(1-\lambda)^2$ where K is the constant of Lemma 46. Note that once the infection has size $|A_t| \geq \Delta$, it has enough size to apply the results of the previous section. Let $T \geq t \geq 0$ be two integers and $\alpha > 0$. Define the event $\mathcal{A}_{t,\alpha}$ by

$$\mathcal{A}_{t,\alpha} = \{|A_t| \geq \alpha\} \cup \bigcup_{i=0}^{t-1} \{|A_i| \geq \Delta\}, \quad (3.47)$$

$\mathcal{E}_{t,\alpha}$ by

$$\mathcal{E}_{t,\alpha} = \bigcap_{i=t}^{\infty} \mathcal{A}_{i,\alpha}, \quad (3.48)$$

$\mathcal{E}_{t,\alpha}^T$ by

$$\mathcal{E}_{t,\alpha}^T = \bigcap_{i=t}^T \mathcal{A}_{i,\alpha}. \quad (3.49)$$

The event $\mathcal{A}_{t,\alpha}$ says that the infection size at round t is at least α , unless it has already hit the final target of Δ at some earlier round. The event $\mathcal{E}_{t,\alpha}$ says that the infection size is at least α (the intermediate target) at round t and will not drop below α before reaching the final target Δ , or it has already reached Δ before round t . Observe that due to the source vertex, it holds that $|A_t| \geq 1$ for all t , therefore $\mathbb{P}(\mathcal{E}_{0,1}) = 1$.

In Lemma 49 below we have two intermediate targets α and β for the infection size, where $1 \leq \alpha < \beta \leq \Delta$. The lemma says (roughly) that for some appropriately large specified $T = T(\alpha, \beta, n, \lambda)$, if the infection size is at least α at some round t , then w.h.p. it will reach the second threshold of β within the subsequent T rounds.

Lemma 49. *Let G be a connected n -vertex regular graph with $1 - \lambda \gg \sqrt{\log n/n}$. Let $\beta > \alpha \geq 1$, and let $t \geq 0$ be a non-negative integer. Choose $T = \frac{6}{\alpha} \left(\frac{\beta}{1-\lambda} + \frac{2C \log n}{(1-\lambda)^2} \right)$ where $C \geq 1$. For large enough n , it holds that*

$$\mathbb{P}(\mathcal{E}_{t+T,\beta}) \geq \mathbb{P}(\mathcal{E}_{t,\alpha}) - 2n^{-C}. \quad (3.50)$$

Observe that due to the source vertex, it holds that $|A_t| \geq 1$ for all t , therefore $\mathbb{P}(\mathcal{E}_{0,1}) = 1$.

Proof. We compute the probability of $\mathcal{E}_{t+T,\beta}^c$. Write

$$\begin{aligned} \mathbb{P}(\mathcal{E}_{t+T,\beta}^c) &\leq \mathbb{P}(\mathcal{E}_{t+T,\beta}^c \cap \mathcal{E}_{t,\alpha}) + \mathbb{P}(\mathcal{E}_{t,\alpha}^c) \\ &\leq \sum_{s=t+T}^{\infty} \mathbb{P}(\mathcal{A}_{s,\beta}^c \cap \mathcal{E}_{t,\alpha}) + \mathbb{P}(\mathcal{E}_{t,\alpha}^c) \end{aligned} \quad (3.51)$$

We focus on $\mathbb{P}(\mathcal{A}_{s,\beta}^c \cap \mathcal{E}_{t,\alpha})$ for $s \geq t + T$. Recall that

$$\mathcal{A}_{s,\beta}^c = \{|A_s| < \beta\} \cap \bigcap_{i=0}^{s-1} \{|A_i| \leq \Delta\}.$$

For $s > t$, define the event $\mathcal{B}_{s,t,\alpha} = \bigcap_{i=t}^s \{\Delta \geq |A_i| \geq \alpha\}$. Then

$$\begin{aligned}
\bigcap_{i=0}^{s-1} \{|A_i| \leq \Delta\} \cap \mathcal{E}_{t,\alpha} &\subseteq \bigcap_{i=0}^{s-1} \{|A_i| \leq \Delta\} \cap \mathcal{E}_{t,\alpha}^T \\
&= \left(\bigcap_{i=0}^{s-1} \{|A_i| \leq \Delta\} \cap \bigcap_{i=t}^T \left(\{|A_i| \geq \alpha\} \cup \bigcup_{j=0}^{i-1} \{|A_j| > \Delta\} \right) \right) \\
&= \bigcap_{i=t}^{s-1} \{\Delta \geq |A_i| \geq \alpha\} = \mathcal{B}_{s-1,t,\alpha}
\end{aligned} \tag{3.52}$$

Therefore $\mathcal{A}_{s,\beta}^c \cap \mathcal{E}_{t,\alpha} \subseteq \{|A_s| < \beta\} \cap \mathcal{B}_{s-1,t,\alpha}$. Let $\phi = \log \delta > 0$ where $\delta = (1 + (1 - \lambda^2)/2)$, thus

$$\begin{aligned}
\mathbb{P}(\mathcal{A}_{s,\beta}^c \cap \mathcal{E}_{t,\alpha}) &\leq \mathbb{P}(\{|A_s| < \beta\} \cap \mathcal{B}_{s-1,t,\alpha}) \\
&= \mathbb{P}(e^{-\phi|A_s|} \mathbf{1}_{\mathcal{B}_{s-1,t,\alpha}} > e^{-\phi\beta}) \\
&\leq \mathbb{E}(e^{-\phi|A_s|} \mathbf{1}_{\mathcal{B}_{s-1,t,\alpha}}) e^{\phi\beta}.
\end{aligned} \tag{3.53}$$

To ease notation, we write $\mathcal{B}_s$ instead of $\mathcal{B}_{s,t,\alpha}$. We focus on getting a good estimate of $\mathbb{E}(e^{-\phi|A_s|} \mathbf{1}_{\mathcal{B}_{s-1}})$. Define $G(s)$ by $G(s) = \mathbb{E}(e^{-\phi|A_s|} \mathbf{1}_{\mathcal{B}_{s-1}})$. Let $\mathcal{F}_s = \sigma(A_0, \dots, A_s)$, then for $s > t$, it holds

$$\begin{aligned}
G(s) &= \mathbb{E}(\mathbb{E}(e^{-\phi|A_s|} \mathbf{1}_{\mathcal{B}_{s-1}} | \mathcal{F}_{s-1})) \\
&\leq \mathbb{E}(\mathbf{1}_{\mathcal{B}_{s-1}} \mathbb{E}(e^{-\phi|A_s|} | \mathcal{F}_{s-1})).
\end{aligned} \tag{3.54}$$

Remember that given A_{s-1} , the event that $u \in V$ belongs to A_s is independent of other vertices, thus

$$\begin{aligned}
\mathbb{E}(e^{-\phi|A_s|} | \mathcal{F}_{s-1}) &= \prod_{u \in V} (e^{-\phi} \mathbb{P}(u \in A_s | A_{s-1}) + \mathbb{P}(u \notin A_s | A_{s-1})) \\
&= \prod_{u \in V} (1 - (1 - e^{-\phi}) \mathbb{P}(u \in A_s | A_{s-1})) \\
&\leq \prod_{u \in V} \exp(-(1 - e^{-\phi}) \mathbb{P}(u \in A_s | A_{s-1})) \\
&= \exp(-(1 - e^{-\phi}) \mathbb{E}(|A_s| | A_{s-1})) \\
&\leq \exp(-(1 - e^{-\phi}) |A_{s-1}| (1 + (1 - \lambda^2)(1 - |A_s|/n))).
\end{aligned} \tag{3.55}$$

The last inequality follows from Lemma 45. By substituting the last expression into Equation (3.54), we obtain $G(s) \leq \mathbb{E}(\mathbb{1}_{\mathcal{B}_{s-2}} h(A_{s-1}))$, where

$$\begin{aligned} h(A) &= \mathbb{1}_{\{\Delta \geq |A| \geq \alpha\}} \exp(-|A|(1 - e^{-\phi})(1 + (1 - \lambda^2)(1 - |A|/n)) \\ &\leq \mathbb{1}_{\{\Delta \geq |A| \geq \alpha\}} \exp(-|A|(1 - e^{-\phi})(1 + (1 - \lambda^2)(1 - \Delta/n)) \\ &\leq \mathbb{1}_{\{\Delta \geq |A| \geq \alpha\}} \exp(-|A|(1 - e^{-\phi})(1 + (1 - \lambda^2)/2)). \end{aligned} \quad (3.56)$$

The last step holds for large n , indeed, by our assumption over $1 - \lambda$ we have $|A| \leq \Delta \ll n$. Hence, $1 - \Delta/n \geq 1/2$ for large enough n . Write $\delta = (1 + (1 - \lambda^2))/2$, then the following holds.

$$\begin{aligned} h(A) &\leq \mathbb{1}_{\{\Delta \geq |A| \geq \alpha\}} \exp(-|A|(1 - e^{-\phi})\delta) \\ &= \mathbb{1}_{\{\Delta \geq |A| \geq \alpha\}} \exp(-|A|((1 - e^{-\phi})\delta - \phi)) \exp(-\phi|A|) \\ &\leq \mathbb{1}_{\{\Delta \geq |A| \geq \alpha\}} \exp(-\alpha((1 - e^{-\phi})\delta - \phi)) \exp(-\phi|A|). \end{aligned} \quad (3.57)$$

The last inequality is due to the fact that $x + e^{-x} - 1 \geq 0$ for all x . Therefore

$$\begin{aligned} G(s) &\leq \mathbb{E}(\mathbb{1}_{\mathcal{B}_{s-2}} h(A_{s-1})) \\ &\leq \exp(-\alpha((1 - e^{-\phi})\delta - \phi)) \mathbb{E}(e^{-\phi|A_{s-1}|} \mathbb{1}_{\mathcal{B}_{s-2}} \mathbb{1}_{\{\Delta \geq |A| \geq \alpha\}}) \end{aligned} \quad (3.58)$$

$$\leq \exp(-\alpha((1 - e^{-\phi})\delta - \phi)) G(s-1). \quad (3.59)$$

Fix $s = t + t'$, applying equation (3.59) recursively, and using the fact that $G(t) \leq 1$, we obtain

$$G(s) \leq G(t) \exp(-t'\alpha((1 - e^{-\phi})\delta - \phi)) \leq \exp(-t'\alpha((1 - e^{-\phi})\delta - \phi)). \quad (3.60)$$

Then, for $s = t + t'$, it holds that

$$\begin{aligned} \mathbb{P}(\mathcal{A}_{s,\beta}^c \cap \mathcal{E}_{t,\alpha}) &\leq \exp(-t'\alpha((1 - e^{-\phi})\delta - \phi) + \phi\beta) \\ &= \exp(-(s - t)\alpha((1 - e^{-\phi})\delta - \phi) + \phi\beta). \end{aligned} \quad (3.61)$$

Denote $\gamma = (1 - e^{-\phi})\delta - \phi$. Returning to equation 3.51, by summing up from $s = t + T$

to infinity, we have

$$\begin{aligned}
\sum_{s=t+T}^{\infty} \mathbb{P}(\mathcal{A}_{s,\beta}^c \cap \mathcal{E}_{t,\alpha}) &\leq \sum_{i=0}^{\infty} \exp(-(i+T)\alpha\gamma + \phi\beta) \\
&= e^{-\alpha\gamma T + \phi\beta} \sum_{i=0}^{\infty} e^{-\alpha\gamma i} \\
&= \frac{e^{-\alpha\gamma T + \phi\beta}}{1 - e^{-\alpha\gamma T}}.
\end{aligned} \tag{3.62}$$

We proceed to find an upper bound of $\phi\beta - \alpha\gamma T$. By using the definition of γ, ϕ and δ in terms of $x = (1 - \lambda^2)/2$, we get

$$\begin{aligned}
-\alpha\gamma T + \phi\beta &= -\alpha T((1 - \delta^{-1})\delta - \log(\delta)) + \log(\delta)\beta \\
&= -\alpha T(\delta - 1 - \log(\delta)) + \log(\delta)\beta \\
&= -\alpha T(x - \log(1+x)) + \log(1+x)\beta \\
&= -\alpha T x + (\alpha T + \beta) \log(1+x).
\end{aligned} \tag{3.63}$$

By using that $\log(1+x) \leq x - x^2/2 + x^3/3$ for $x \leq 1$, we get

$$\begin{aligned}
-\alpha\gamma T + \phi\beta &\leq -\alpha T x + (\alpha T + \beta) \log(1+x) \\
&\leq -\alpha T x + (\alpha T + \beta)(x - x^2/2 + x^3/3)
\end{aligned} \tag{3.64}$$

$$\begin{aligned}
&= x \left(\beta \left(1 - \frac{x}{2} + \frac{x^2}{3} \right) - \frac{\alpha T x}{2} \left(1 - \frac{2x}{3} \right) \right) \\
&\leq x \left(\beta - \frac{\alpha T x}{3} \right).
\end{aligned} \tag{3.65}$$

The last equality comes from the fact that $1 - x/2 + x^2/3$ is decreasing for $x \in [0, 1/2]$, and $(1 - \lambda)/2 \leq 1/2$ (for large n). Choose $T = \frac{3\beta}{\alpha x} + \frac{3C \log n}{\alpha x^2}$ for $C \geq 1$, and conclude that

$$-\alpha\gamma T + \phi\beta \leq -C \log n.$$

Substituting the inequality above into Equation (3.62), we obtain

$$\begin{aligned}
\sum_{s=t+T}^{\infty} \mathbb{P}(\mathcal{A}_{s,\beta}^c \cap \mathcal{E}_{t,\alpha}) &\leq \frac{e^{-\alpha\gamma T + \phi\beta}}{1 - e^{-\alpha\gamma T}} \\
&\leq \frac{e^{-C \log n}}{1 - e^{-C \log n}} \leq 2n^{-C},
\end{aligned} \tag{3.66}$$

for large n . Finally, observe that

$$\begin{aligned} T &= \frac{3\beta}{\alpha x} + \frac{3C \log n}{\alpha x^2} \\ &\leq \frac{6}{\alpha} \left(\frac{\beta}{1-\lambda} + \frac{2C \log n}{(1-\lambda)^2} \right). \end{aligned} \quad (3.67)$$

From equations 3.51 and 3.66, we conclude that

$$\mathbb{P}(\mathcal{E}_{t+T,\beta}^c) \leq 2n^{-C} + \mathbb{P}(\mathcal{E}_{t,\alpha}^c).$$

□

By using the lemma above, we can prove that in $\mathcal{O}(\log n/(1-\lambda)^2)$ rounds, the number of infected vertices is large enough to apply Corollary 48.

Corollary 50. *Let G be a connected n -vertex r -regular graph with $1-\lambda \gg \sqrt{\log n/n}$. Consider the BIPS process with $|A_0| = 1$ and let T be the first time such that $|A_T| \geq K \log n/(1-\lambda)^2$ where K is the constant of Lemma 46. Then, there exists a large enough constant L such that $T \leq L \log n/(1-\lambda)^2$, with probability at least $1 - 4n^{-3}$*

Proof. Fix $C = 3$. We first apply Lemma 49 with t given by

$$t = \frac{12(C+1) \log n}{(1-\lambda)^2},$$

and $\beta = 2 \log n/(1-\lambda)$, and by using the fact that $\mathbb{P}(\mathcal{E}_{0,1}) = 1$, we get

$$\mathbb{P}(\mathcal{E}_{t,\beta}) \geq 1 - 2n^{-C}. \quad (3.68)$$

In this point, recall that $\Delta = K \log n/(1-\lambda)^2$, where K is a large enough constant (as large as needed in Lemma 46). Choose $t' = 12(C+K)/(1-\lambda)$, and $\beta' = \Delta$, and apply Lemma 43 to obtain

$$\mathbb{P}(\mathcal{E}_{t+t',\Delta}) \geq \mathbb{P}(\mathcal{E}_{t,\beta}) - 2n^{-C} \geq 1 - 4n^{-C}. \quad (3.69)$$

Observe that the event $\mathcal{E}_{t+t',\Delta}$ implies that there exists a time $s \leq t + t'$ such that $|A_s| \geq \Delta$, then we conclude that $T \leq t + t' \leq L \log n/(1-\lambda)^2$ with probability at least $1 - 4n^{-3}$. □

In a similar fashion as Equation (3.46), we can use Corollary 50 and Corollary 48 to conclude the result of Equation (3.8) of Theorem 38.

Finally, by using the previous results, the proof of Equation (3.7) of Theorem 38 is fairly easy. Indeed, by following the same arguments above, we only need to show that the number of infected vertices is $\Omega((\log n)/(1 - \lambda))$ in $\mathcal{O}((r/(1 - \lambda) + r^2) \log n)$ rounds with high probability. This is given by Lemma 44, which we restate for the case of regular graph.

Corollary 51. *Consider the BIPS process on a connected r -regular graph with n vertices. For any constant $C > 0$ define $C' = 64(C + 3)$, then for any $1 \leq k \leq n$ and $t(k) = 4rk + C'r^2 \log n$,*

$$\mathbb{P}(\exists t \geq t(k) : |A_t| < k) = \mathcal{O}(n^{-C}). \quad (3.70)$$

By using the Corollary 51 with $t = \mathcal{O}((r/(1 - \lambda) + r^2) \log n)$ and $\alpha = \Omega(\log n/(1 - \lambda))$, it holds that

$$\mathbb{P}(\mathcal{E}_{t,\alpha}) \geq \mathbb{P}(\forall s \geq t : |A_s| \geq \alpha) = 1 - \mathcal{O}(n^{-3}),$$

which is the same as in equation (3.68) in the proof of Corollary 50. After that, we continue with exactly the same proof to obtain that in

$$\mathcal{O}((r/(1 - \lambda) + r^2) \log n + 1/(1 - \lambda)) = \mathcal{O}((r/(1 - \lambda) + r^2) \log n)$$

rounds, the number of infected vertices is at least $K \log n/(1 - \lambda)^2$, thus proving Equation (3.7) of Theorem 38.

Chapter 4

Discordant Voting Model on General Graphs

4.1 Introduction

This chapter is dedicated to the study of *discordant voting*. In this model, we consider a connected graph whose vertices hold one of two different opinions, red (R) or blue (B), i.e., this is a two-party model. In discordant voting, interactions occur between two neighbouring vertices with different opinions. After an interaction happens, the two involved vertices share the same colour. The quantity of interest is the time to reach consensus, that is, the first round such that all vertices have the same colour. We denote this time by τ_{cons} .

An edge whose endpoint have different colours (i.e., one vertex is coloured red and the other one blue) is said to be discordant. A vertex is discordant if it is incident with a discordant edge.

Discordant voting is an asynchronous process, that is, only a pair of vertices interact at each time-step. There are several ways to update the colours of the interacting vertices. Here, we consider three of them.

Push: Pick a random discordant vertex and push its colour on a random discordant

neighbour.

Pull: Pick a random discordant vertex and pull the colour of a random discordant neighbour.

Oblivious: Pick a random endpoint of a random discordant edge and push the colour on the other endpoint.

Remark 52. *Observe that once the process reaches consensus, there are no more discordant vertices and the process stops. For consistency, we extend the process beyond the consensus time by keeping the final configuration constant over time.*

We are interested in computing $\mathbb{E}(\tau_{\text{cons}})$, the expected consensus time, for different graph topologies. Intuitively, we expect these protocols to reach consensus faster than the standard asynchronous pull and push models introduced in Chapter 2, because discordant models only consider interaction between vertices with different opinion. While the above intuitive statement is true for some topologies, it is not true in general, indeed, we will show that the expected consensus time can be exponentially large.

Perhaps, even more surprising is the fact that for discordant voting using the oblivious protocol, the expected consensus time in any connected graph depends only on the initial number of vertices of each colour (red or blue), and it is independent of any other possible structure of the graph.

Whichever discordant edge is chosen, the number of blue (resp. red) vertices in the graph increases (resp. decreases) by one with probability $1/2$ at each step. This is equivalent to an unbiased random walk on the path with vertices $\{0, 1, \dots, n\}$ with absorbing barriers (see Feller [41, XIV.3]).

Remark 53. Oblivious protocol. *Let τ_{cons} be the consensus time in the discordant voting process starting from any initial colouring with r red vertices and $n - r$ blue vertices. Then for any connected n vertex graph, the consensus time for the discordant oblivious model is $\mathbb{E}\tau_{\text{cons}} = r(n - r)$.*

Starting with an equal number of red and blue vertices, the oblivious protocol takes $\mathbb{E}\tau_{\text{cons}} \sim n^2/4$ steps for any connected graph. For ordinary oblivious voting (cf. Chapter 2), the performance of the oblivious protocol can also depend on the number of edges m . In the worst case, the expected wait time to hit the last red-blue edge is m , so the ordinary case takes $\mathbb{E}\tau_{\text{cons}} = \mathcal{O}(mn^2)$ steps.

While in oblivious voting, the graph topology does not affect the speed of the process, this is not true for the discordant pull and push models. We begin by analysing the complete graph K_n .

Theorem 54. *Let τ_{cons} be the time to consensus in discordant voting. Then, for the complete graph K_n , it holds that $\mathbb{E}\tau_{\text{cons}}(\text{Push}) = \Theta(n \log n)$, and $\mathbb{E}\tau_{\text{cons}}(\text{Pull}) = \Theta(2^n)$.*

Remark 55. *We say that $\mathbb{E}\tau_{\text{cons}} = \mathcal{O}(f(n))$, if there exists a constant C such that for all initial colourings, it holds $\mathbb{E}\tau_{\text{cons}} \leq Cf(n)$ for all large n . We say that $\mathbb{E}\tau_{\text{cons}} = \Omega(f(n))$, if there exists an initial colouring and a constant $c > 0$ such that $\mathbb{E}\tau_{\text{cons}} \geq cf(n)$, for all large n . Finally, $\mathbb{E}(\tau_{\text{cons}}) = \Theta(f(n))$, if $\mathbb{E}(\tau_{\text{cons}}) = \mathcal{O}(f(n))$ and $\mathbb{E}(\tau_{\text{cons}}) = \Omega(f(n))$.*

From the previous result, we conclude that in the complete graph K_n , different protocols give very different expected consensus times, which vary from $\Theta(n \log n)$ for push, to $\Theta(n^2)$ for oblivious, to $\Theta(2^n)$ for pull. On the basis of this evidence, our initial view was that there should be a meta-theorem of the ‘push is faster than oblivious, oblivious is faster than pull’ type. Intuitively, this is supported by the following argument. Suppose red (R) is the larger colour class. Then, choosing a discordant vertex uniformly at random, favours the election of the larger class. In the push process, red vertices push their opinion more often, which tends to increase the size of R . Conversely, the pull process tends to re-balance the set sizes. If R is larger, it is recoloured more often, making the pull process slower.

For the cycle C_n , we prove that all three protocols have similar expected time to consensus; a result which is consistent with our possible meta-theorem above.

Theorem 56. *Let τ_{cons} be the time to consensus in discordant voting. Then, for the cycle C_n , the discordant Push, Pull and Oblivious protocols have $\mathbb{E}\tau_{\text{cons}} = \Theta(n^2)$.*

At this point, based on our two previous results, we are left with a choice. We either produce evidence for a relationship of the form $\mathbb{E}\tau_{\text{cons}}(\text{Push}) = \mathcal{O}(\mathbb{E}\tau_{\text{cons}}(\text{Pull}))$ for general graphs, or refute it. Mossel and Roch [64] found slow convergence of the iterated prisoners dilemma problem (IPD) on caterpillar trees. Intuitively push voting is aggressive, whereas pull voting is altruistic, and thus, similar to cooperation in the IPD. Motivated by this, we found simple counter-examples, namely the star graph S_n and the double star S_n^* . The star is the graph on n vertices such that a central vertex is connected to the remaining $n - 1$ vertices, and no other pair of vertices is adjacent. The double star S_n^* is given by two disjoint stars $S_{n/2}$ which are connected by a special edge via their central vertices.

Theorem 57. *Let τ_{cons} be the time to consensus in discordant voting. Then, for the star graph S_n , $\mathbb{E}\tau_{\text{cons}}(\text{Push}) = \Theta(n^2 \log n)$, and $\mathbb{E}\tau_{\text{cons}}(\text{Pull}) = \mathcal{O}(n^2)$.*

For the double star S_n^* , $\mathbb{E}\tau_{\text{cons}}(\text{Push}) = \Omega(2^{n/5})$, and $\mathbb{E}\tau_{\text{cons}}(\text{Pull}) = \mathcal{O}(n^3)$.

Thereupon, little remains of the possibility of a meta-theorem, except for the vague hope that at least one of the push and pull protocols always has expected polynomial time to consensus. However, this is disproved by the example of the *barbell graph*, which consists of two cliques of size $n/2$ joined by a single edge.

Theorem 58. *Let τ_{cons} be the time to consensus in discordant voting, then for the barbell graph on n vertices, $\mathbb{E}\tau_{\text{cons}}(\text{Push}) = \Omega(2^{n/5})$, and $\mathbb{E}\tau_{\text{cons}}(\text{Pull}) = \Theta((\sqrt{2})^n)$.*

A summary of the results of Theorems 54, 56, 57, and 58 is given in Figure 4.1. The column for ordinary asynchronous pull, push and oblivious voting follows from the results obtained in Section 2.7 of Chapter 2.

A major obstacle in the analysis of discordant voting, is that the effect of re-colouring a vertex is not always monotone, and the analysis of the protocols depends a lot on the graph topology. Indeed, for each of the graphs studied, the way to bound

	Discordant voting			Ordinary voting		
	Push	Pull	Obliv.	Push	Pull	Obliv.
Complete graph K_n	$\Theta(n \log n)$	$\Theta(2^n)$	$n^2/4$	$\mathcal{O}(n^2)$	$\mathcal{O}(n^2)$	$\mathcal{O}(n^2)$
Cycle C_n	$\Theta(n^2)$	$\Theta(n^2)$		$\mathcal{O}(n^3)$	$\mathcal{O}(n^3)$	$\mathcal{O}(n^3)$
Star graph S_n	$\Theta(n^2 \log n)$	$\mathcal{O}(n^2)$		$\mathcal{O}(n^3)$	$\mathcal{O}(n \log n)$	$\mathcal{O}(n^2)$
Double star S_n^*	$\Omega(2^{n/5})$	$\mathcal{O}(n^3)$		$\mathcal{O}(n^4)$	$\mathcal{O}(n^2)$	$\mathcal{O}(n^2)$
Barbell graph	$\Omega(2^{n/5})$	$\Theta(2^{n/2})$		$\mathcal{O}(n^3)$	$\mathcal{O}(n^3)$	$\mathcal{O}(n^3)$

Figure 4.1: Comparison of expected time to consensus ($\mathbb{E}\tau_{\text{cons}}$) for discordant and ordinary asynchronous voting protocols on connected n -vertex graphs, starting from $R = B = n/2$.

$\mathbb{E}\tau_{\text{cons}}$ differs. The proof of the pull voting result for the cycle C_n in particular, is somewhat delicate, and requires an analysis of the optimum of a linear program based on a potential function.

While our proofs vary from graph to graph, the general proof methodology is to map the process to a Birth-and-Death chain on state space $\{0, \dots, n\}$. In section 4.2, we state the general results about Birth-and-Death chains that are used for the proofs. After section 4.2, we then prove Theorems 54, 56, 57 and 58 in that order.

4.2 Birth-and-Death Chains

A Markov chain $(X_t)_{t \geq 0}$ is said to be a Birth-and-Death chain on state space $S = \{0, \dots, N\}$, if given $X_t = i$, the possible values of X_{t+1} are $i+1, i$ or $i-1$ with probability p_i, r_i and q_i , respectively. Note that $q_0 = p_N = 0$. In this chapter, unless stated otherwise, we consider no self-loops, i.e., we assume $r_i = 0$. We also consider reflecting barriers, i.e., $p_0 = q_N = 1$, and that all states are reachable from any state, i.e., $p_i > 0$ for $i \in \{0, \dots, N-1\}$, and $q_i > 0$ for $i \in \{1, \dots, N\}$. Given a random variable Y , we denote by $\mathbb{E}_i Y$, the expected value of random variable Y given that

the chain starts from i (i.e., $X_0 = i$). Finally, we define the hitting time of state i as $T_i = \min\{t \geq 0 : X_t = i\}$.

We summarise the results we require on Birth-and-Death chains. Most of them are standard in the Markov chain literature, and they can be found in, e.g., Chapter 2 of [57].

We say that a probability distribution π satisfies the detailed balance equations, if

$$\pi(i)P(i, j) = \pi(j)P(j, i), \text{ for all } i, j \in S. \quad (4.1)$$

All Birth-and-Death chains with $p_i = P(i, i+1)$ and $q_i = P(i, i-1)$ such that all states can be reached, can be shown to satisfy the detailed balance equations. From this, it can be proved that

$$\mathbb{E}_{i-1}T_i = \frac{1}{q_i\pi(i)} \sum_{k=0}^{i-1} \pi(k) \quad (4.2)$$

An equivalent formulation is $\mathbb{E}_0T_1 = 1/p_0 = 1$, and in general,

$$\mathbb{E}_{i-1}T_i = \sum_{k=0}^{i-1} \frac{1}{p_k} \frac{q_{k+1} \cdots q_{i-1}}{p_{k+1} \cdots p_{i-1}}, \quad \text{for } i \in \{1, \dots, N\}. \quad (4.3)$$

In writing this expression, we follow the convention that if $k = i-1$ then $\frac{q_{k+1} \cdots q_{i-1}}{p_{k+1} \cdots p_{i-1}} = 1$, so that the last term in the sum is $1/p_{i-1}$. Also, note that the final index k in the sum is $N-1$, i.e., we never divide by $p_N = 0$.

Let T_M be the number of transitions needed to reach state M for the first time, i.e., the hitting time of M . For any $M \leq N$, we have that $\mathbb{E}_0T_M = \sum_{i=1}^M \mathbb{E}_{i-1}T_i$. For example, $\mathbb{E}_0T_1 = \frac{1}{p_0} = 1$ and $\mathbb{E}_0T_2 = 1 + \frac{1}{p_1} + \frac{q_1}{p_0p_1}$ etc. Thus, for $M \geq 1$,

$$\mathbb{E}_0T_M = \sum_{i=1}^M \mathbb{E}_{i-1}T_i = \sum_{i=1}^M \sum_{k=0}^{i-1} \frac{1}{p_k} \prod_{j=k+1}^{i-1} \frac{q_j}{p_j}. \quad (4.4)$$

4.2.1 Push and Pull Chains

We proceed to describe two important chains that feature in our analysis: the push and pull chains. For our purposes, these chains have state space in $\{0, 1, \dots, N\}$, where $N = n/2$ and n is an even positive integer.

Push Chain. The push chain $(Z_t)_{t \geq 0}$ is a Birth-and-Death chain whose state space is $\{0, \dots, N\}$, and the transition probabilities $p_i = P(i, i+1)$ are given by

$$p_i = \begin{cases} 1, & \text{if } i = 0 \\ 1/2 + i/n + \delta/n, & \text{if } i \in \{1, \dots, n/2 - 1\} \\ 0, & \text{if } i = n/2. \end{cases} \quad (4.5)$$

where $q_i = 1 - p_i$, and $\delta \in \{-1, 0, +1\}$.

Pull Chain. The pull chain $(\bar{Z}_t)_{t \geq 0}$ is a Birth-and-Death chain whose state space is $\{0, \dots, N\}$, and the transition probabilities $\bar{p}_i = \bar{P}(i, i+1)$ are given by

$$\bar{p}_i = \begin{cases} 1, & \text{if } i = 0 \\ 1/2 - i/n - \delta/n, & \text{if } i \in \{1, \dots, n/2 - 1\} \\ 0, & \text{if } i = n/2 \end{cases}. \quad (4.6)$$

Again, $\bar{q}_i = 1 - \bar{p}_i$ and $\delta \in \{-1, 0, +1\}$ is fixed.

For $1 \leq i \leq N-1$, the pull chain is the push chain with the probabilities reversed, i.e., $\bar{p}_i = q_i$.

4.2.2 Push Chain: Bounds on Hitting Times

Push Chain: Upper bound on hitting time.

Lemma 59. *For any $M \leq N$, let $\mathbb{E}_0 T_M$ be the expected hitting time of M in the push chain $(Z_t)_{t \geq 0}$ starting from state 0. Then, for all $\delta \in \{-1, 0, 1\}$,*

$$\mathbb{E}_0 T_M \leq 2N \log M + O(1).$$

Proof. By using (4.4), and by recalling the notational convention given below (4.3), we can change the order of summation to obtain

$$E_0 T_M = \sum_{k=0}^{M-1} \sum_{i=k+1}^M \frac{1}{p_k} \frac{q_{k+1} \cdots q_{i-1}}{p_{k+1} \cdots p_{i-1}} = \frac{1}{p_{M-1}} + \sum_{k=0}^{M-2} \sum_{i=k+1}^{M-1} \frac{1}{p_k} \frac{q_{k+1} \cdots q_{i-1}}{p_{k+1} \cdots p_{i-1}}. \quad (4.7)$$

By using (4.5), we observe that for $1 \leq k \leq N-2$, it holds $q_k/p_k \geq q_{k+1}/p_{k+1}$, $q_1/p_1 \leq 1$, and for $2 \leq k \leq N-1$, that $q_k/p_k < 1$. As $p_0 = 1$, we upper bound $E_0 T_M$ by

$$E_0 T_M \leq M + \frac{1}{p_{M-1}} + \sum_{k=1}^{M-2} \frac{1}{p_k} \sum_{i=k+1}^{M-1} \left(\frac{q_{k+1}}{p_{k+1}} \right)^{i-k-1}, \quad (4.8)$$

and

$$\sum_{k=1}^{M-2} \frac{1}{p_k} \sum_{\ell=0}^{\infty} \left(\frac{q_{k+1}}{p_{k+1}} \right)^{\ell} = \sum_{k=1}^{M-2} \frac{1}{p_k} \frac{1}{1 - \frac{q_{k+1}}{p_{k+1}}} = \sum_{k=1}^{M-2} \frac{p_{k+1}}{p_k} \frac{1}{p_{k+1} - q_{k+1}}. \quad (4.9)$$

As $q_k = 1 - p_k$, $p_k - q_k = 2p_k - 1 > 0$ for all $k \in \{2, \dots, N-1\}$, then $\frac{1}{p_k - q_k} = \frac{N}{k+\delta}$. For all $k \in \{1, \dots, N-2\}$, we have $\frac{p_{k+1}}{p_k} \leq 2$. By using Equation (4.8) with the upper bounds given in Equation (4.9), we obtain the required conclusion. $\square$

Push Chain: Lower bound on hitting time.

Lemma 60. *Let $\delta = 0$. Let $\mathbb{E}_0 T_M$ be the expected hitting time of M in the push chain $(Z_t)_{t \geq 0}$ starting from state 0. There exists a constant C such that for any N such that $\sqrt{N} \leq M = o(N^{3/4})$,*

$$\mathbb{E}_0 T_M \geq CN \log(M/\sqrt{N}).$$

Proof. For $0 < x < 1$,

$$\frac{1-x}{1+x} = \exp \left\{ -2 \left(x + \frac{x^3}{3} + \dots + \frac{x^{2\ell+1}}{2\ell+1} + \dots \right) \right\}.$$

Thus with $N = n/2$

$$\prod_{j=k+1}^{i-1} \frac{q_j}{p_j} = \prod_{j=k+1}^{i-1} \frac{1-j/N}{1+j/N} \quad (4.10)$$

$$\begin{aligned} &= \exp \left\{ -2 \left(\sum_{j=k+1}^{i-1} \frac{j}{N} + \sum_{j=k+1}^{i-1} \frac{(j/N)^3}{3} + \dots + \sum_{j=k+1}^{i-1} \frac{(j/N)^{2\ell+1}}{2\ell+1} + \dots \right) \right\} \\ &= \exp \{-2\mathcal{F}\}. \end{aligned} \quad (4.11)$$

If $f(s)$ is non-negative and monotone increasing, then $\sum_{s=k+1}^{i-1} f(s) \leq \int_k^i f(s) ds$.

Therefore, the sum of terms in $(j/N)^3$ and above in $\mathcal{F}$ can be bounded above by

$$\sum_{\ell \geq 1} \sum_{j=k+1}^{i-1} \frac{(j/N)^{2\ell+1}}{2\ell+1} \leq \sum_{\ell \geq 1} \frac{1}{(2\ell+1)N^{2\ell+1}} \int_k^i x^{2\ell} dx \quad (4.12)$$

$$\begin{aligned} &\leq \sum_{\ell \geq 1} \frac{1}{(2\ell+1)N^{2\ell+1}} \cdot \frac{i^{2\ell+2}}{2\ell+2} \\ &= \mathcal{O}\left(\frac{i^4}{N^3}\right) \sum_{\ell \geq 1} \frac{1}{(2\ell+1)(2\ell+2)} = \mathcal{O}\left(\frac{i^4}{N^3}\right). \end{aligned} \quad (4.13)$$

Thus, for all $i \leq M = o(N^{3/4})$, it holds

$$\mathcal{F} = \frac{i(i-1)}{2N} - \frac{k(k+1)}{2N} + \mathcal{O}\left(\frac{i^4}{N^3}\right) = \frac{i^2}{2N} - \frac{k^2}{2N} - \frac{i+k}{2N} - o(1).$$

Replacing $\mathcal{F}$ with the upper bound given above, we obtain a lower bound on the term of Equation (4.11). In particular,

$$\prod_{j=k+1}^{i-1} \frac{q_j}{p_j} \geq \exp\left\{-\frac{i^2}{N} + \frac{k^2}{N} + o(1)\right\}. \quad (4.14)$$

Observe the hidden constant in the $o(1)$ depends on N but it is uniform on i . Then, by using such a lower bound in Equation (4.7), we obtain

$$\begin{aligned} \mathbb{E}_0 T_M &\geq (1 - o(1)) \sum_{i=0}^M \sum_{k=0}^{i-1} \frac{1}{p_k} \exp\left(-\frac{i^2}{N}\right) \exp\left(\frac{k^2}{N}\right) \\ &\geq 2(1 - o(1)) \sum_{i=0}^M \exp\left(-\frac{i^2}{N}\right) \sum_{k=0}^{i-1} \exp\left(\frac{k^2}{N}\right). \end{aligned} \quad (4.15)$$

The last line follows because in the regime $k \leq M$, we have $p_k \sim 1/2$. We proceed to find a lower bound for the last term on the right-hand side of equation (4.15). Let

$$\sigma(i) = \sum_{k=0}^{i-1} \exp\left(\frac{k^2}{N}\right), \quad (4.16)$$

and let $\beta = (1/2) \log 2 \approx 0.34$. We claim that if $i \geq \sqrt{N}$, then

$$\sigma(i) \geq (1 - o(1)) \frac{\beta N}{2i} e^{i^2/N}. \quad (4.17)$$

Indeed, let $A = \lfloor \beta N/i \rfloor$. Then

$$\sigma(i) \geq \sum_{k=i-A}^{i-1} \exp(k^2/N) \geq A \exp\left(\frac{(i-A)^2}{N}\right) \geq A \exp\left(\frac{i^2 - 2iA}{N}\right). \quad (4.18)$$

Using the definition A , we have that $Ai/N \leq \beta$. Then

$$\sigma(i) \geq A \exp\left(\frac{i^2 - 2iA}{N}\right) = \lfloor \beta N/i \rfloor \exp\left(-2\beta + \frac{i^2}{N}\right) = \frac{\lfloor \beta N/i \rfloor}{2} \exp\left(\frac{i^2}{N}\right). \quad (4.19)$$

Finally, by replacing Equation (4.17) in Equation (4.15) we get

$$\begin{aligned} \mathbb{E}_0 T_M &\geq 2(1 - o(1)) \sum_{i=\lfloor \sqrt{N} \rfloor}^M \exp\left(-\frac{i^2}{N}\right) \sigma(i) \geq \beta N(1 - o(1)) \sum_{i=\lfloor \sqrt{N} \rfloor}^M \frac{1}{i} \\ &\geq CN \log(M/\sqrt{N}), \end{aligned} \quad (4.20)$$

where $C > 0$ is an appropriate constant. $\square$

4.2.3 Biased Random Walk With Reflecting Barriers

Given $p \in (0, 1)$, a *biased random walk* with parameter p , $(Y_t)_{t \geq 0}$, is a Birth-and-Death chain on state space $S = \{0, 1, \dots, N\}$ whose transition probabilities $p_i = p(i, i+1)$ are given by

$$p(i, i+1) = \begin{cases} 1, & \text{if } i = 0 \\ p, & \text{if } i \in \{1, \dots, N-1\} \\ 0, & \text{if } i = N \end{cases} \quad (4.21)$$

and $q_i = 1 - p_i$.

We are interested in obtaining a lower bound for the hitting time of state $M \leq N$, i.e. T_M , starting from state 1. Starting from state 1, we say that a *run* is a sequence of steps that finishes in state 0 or state M . The run is a failure if it finishes in state

0. After a failure, due to the reflecting barrier, we return to state 1, and a new run starts. Let K be the number of runs needed to reach state M , and let X_i be the length of the i -th run. Then

$$T_M = (K - 1) + \sum_{i=1}^K X_i. \quad (4.22)$$

In the equation above, the term $(K - 1)$ counts the extra step the chain performs from state 0 to 1 after a failure. Then, as each run has at least one step, it holds that $T_M \geq 2K - 1 \geq K$.

Recall that a run is a failure if it finishes in state 0. Since the runs start from state 1, from [41, Chapter XIV], a run is a failure with probability

$$\frac{(q/p)^M - q/p}{(q/p)^M - 1}.$$

Therefore, as the runs are independent,

$$\mathbb{P}(T_M > k) \geq \mathbb{P}(K > k) = \left(\frac{(q/p)^M - q/p}{(q/p)^M - 1} \right)^k. \quad (4.23)$$

4.2.4 Coupling

Another simple yet useful tool to study Birth-and-Death processes is coupling. Coupling allows us to compare chains with difficult and less tractable transition probabilities with simpler chains. In particular, for the interest of this chapter, we will compare with the push and pull chains, and the biased random walk. Here, we state two coupling results that feature in our analysis.

Lemma 61. *Consider states $\{0, 1, \dots, N\}$, and two Birth-and-Death chains $(X_t)_{t \geq 0}$ and $(\tilde{X}_t)_{t \geq 0}$ with parameters $(p_i, q_i)_{i=1}^N$ and $(\tilde{p}_i, \tilde{q}_i)_{i=1}^N$, respectively ($r_i = \tilde{r}_i = 0$) and suppose that $p_i \leq \tilde{p}_i$ for all i . Hence, if $X_0 = \tilde{X}_0$, it is possible to couple $(X_t)_{t \geq 0}$ and $(\tilde{X}_t)_{t \geq 0}$ such that $X_t \leq \tilde{X}_t$ for all $t \geq 0$. Moreover, the expected hitting time of state $M \in \{0, 1, \dots, N\}$ in chain $(\tilde{X}_t)_{t \geq 0}$ is smaller than in the chain $(X_t)_{t \geq 0}$.*

Proof. Consider the following construction. If $X_t \neq \tilde{X}_t$, then the two chains choose their next state independently with their corresponding probabilities. If $X_t = \tilde{X}_t = i$, we sample a uniform (in $[0, 1]$) random variable U (independent of everything), and set X_{t+1} and $\tilde{X}_{t+1}$ as follows.

$$X_{t+1} = \begin{cases} i + 1 & U \leq p_i, \\ i - 1 & \text{otherwise,} \end{cases} \quad \tilde{X}_{t+1} = \begin{cases} i + 1 & U \leq \tilde{p}_i, \\ i - 1 & \text{otherwise.} \end{cases}$$

Since $p_i \leq \tilde{p}_i$, it is clear that if $X_t = \tilde{X}_t = i$, then $X_{t+1} \leq \tilde{X}_{t+1}$. This prevents the chain $(X_t)_{t \geq 0}$ from overtaking the chain $(\tilde{X}_t)_{t \geq 0}$, when they are in the same position. Note that as the chains only increases or decreases by 1, a necessary condition for one chain to overtake the other, is that they are in the same position. Finally, as both start from the same position, it holds that $X_t \leq \tilde{X}_t$ for all $t \geq 0$. From here, we deduce that the expected hitting time of state M is smaller in $(\tilde{X}_t)_{t \geq 0}$ than in $(X_t)_{t \geq 0}$. $\square$

Following the proof of Lemma 61, we can prove the following.

Lemma 62. *Consider states $\{0, 1, \dots, N\}$ and a Birth-and-Death chain $(X_t)_{t \geq 0}$ with parameters $(p_i, q_i, r_i)_{i=1}^N$. Let $(Y_t)_{t \geq 0}$ be a biased random walk with reflecting barriers with parameter p . Suppose that $p_i + r_i \leq p$ for all $i \geq 1$. Then, it is possible to couple $(X_t)_{t \geq 0}$ and $(Y_t)_{t \geq 0}$ such that $X_t \leq Y_t$, for all $t \geq 0$, given that they start from the same state. Thus, the hitting time of state $M \in \{0, 1, \dots, N\}$ in chain $(Y_t)_{t \geq 0}$ is smaller than in chain $(X_t)_{t \geq 0}$.*

4.3 Voting on the Complete Graph K_n .

Consider the complete graph K_n . Let B_t and R_t be the number of blue vertices and red vertices at time t , respectively. Then, the probability B increases at a given step is B_t/n for the push model, whereas in the pull process it is $R_t/n = 1 - B_t/n$.

Suppose n is even (the case n odd can be handled similarly). The chain defined by $Y_t = \max\{R_t, B_t\} - n/2$ is a Birth-and-Death chain. We proceed to study the

time that the process $(Y_t)_{t \geq 0}$ needs, to reach $N = n/2$ starting from 0, i.e., when the initial state is half red vertices and half blue vertices.

Proof of Theorem 54: Push process. For the push model, the process $(Y_t)_{t \geq 0}$ is identical to the push chain $(Z_t)_{t \geq 0}$ with transitions given by Equation (4.5) with $\delta = 0$. This was analysed in Section 4.2, where we proved that the expected time to reach state $N = n/2$ is $\Theta(n \log n)$.

Proof of Theorem 54: Pull process. For the pull model, the process $(Y_t)_{t \geq 0}$ is identical to the pull chain $(\bar{Z}_t)_{t \geq 0}$ with transitions given by Equation (4.21) with $\delta = 0$. Then, to compute the expected consensus time, we just need to compute the hitting time of N . We use Equation (4.4) for such a task.

To begin with, observe that $w_k = \binom{n}{N+k}$ with $k = 0, 1, \dots, N$, satisfies the detailed balance equations (4.1). Hence, we have $\pi(k) = w_k/W$, where $W = w_0 + w_1 + \dots + w_N$.

It follows from Equation (4.2) that

$$\mathbb{E}_{i-1} T_i = \frac{2n}{n+2i} \cdot \frac{1}{\binom{n}{N+i}} \cdot \sum_{k=0}^{i-1} \binom{n}{N+k}.$$

By choosing $i = N$, we have

$$\mathbb{E}_{N-1} T_N = \sum_{k=0}^{N-1} \binom{n}{N+k} = \frac{1}{2} \left(2^n - 2 + \binom{n}{N} \right) = \Omega(2^n). \quad (4.24)$$

This gives us a lower bound on the expected time to reach state N . On the other hand, the following upper bound

$$\sum_{i=1}^N \mathbb{E}_{i-1} T_i \leq 2 \cdot 2^n \cdot \sum_{i=1}^N \frac{1}{\binom{n}{N+i}} = O(2^n),$$

follows from a result of Sury [73] that states

$$\sum_{i=1}^N \frac{1}{\binom{n}{N+i}} = \frac{n+1}{2^n} \sum_{i=0}^n \frac{2^i}{i+1} = O(1).$$

4.3.1 Voting on the Cycle

An n -cycle G , with $V = [n]$, has $E = \{(i, i+1) : i \in [n]\}$, where we identify vertex $n+i$ with vertex i . See Fig. 4.2.

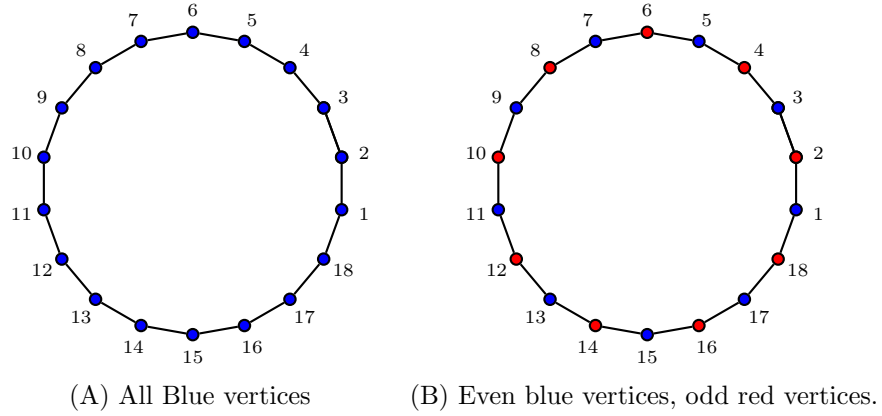


Figure 4.2: Examples of two configurations in C_{18}

Let $\xi = \xi_t$ denote the configuration of colours of the discordant voting process at time t , with $\xi(i)$ being the colour of vertex i . Let $K(\xi)$ denote the set of discordant edges of ξ , and let $k(\xi) = |K(\xi)|$. Let $D = D(\xi)$ denote the set of discordant vertices in ξ .

We say that $i+1, i+2, \dots, i+j$ is a *run* of length j if $\xi(i) \neq \xi(i+1) = \xi(i+2) = \dots = \xi(i+j) \neq \xi(i+j+1)$. A *singleton* is a run of length 1, that is, a single vertex. Those single vertices require special treatment, as they lie in two discordant edges. Note that the number of runs in ξ is equal to the number of discordant edges k . Also, observe that k is even, because red and blue runs must alternate, so we write $r(\xi) = \frac{1}{2}k(\xi)$, and $k_0 = 2r_0 = k(X_0)$. Thus, $r(\xi)$ is the number of runs of a given colour. With the above notation, we have that the consensus time τ_{cons} is the first $t \geq 0$ such that $k(\xi_t) = r(\xi_t) = 0$.

Suppose the k runs in ξ have lengths $\ell_1, \ell_2, \dots, \ell_k$, and let $s(\xi)$ denote the number of singletons. Clearly $\sum_{i=1}^k \ell_i = n$, and the number of discordant vertices is $|D| = 2k - s$, so $k \leq |D| \leq 2k$.

We wish to determine the expected consensus time τ_{cons} , starting from an arbitrary configuration X_0 , of the discordant push or pull process. In these processes, the run lengths behave rather like symmetric random walks on the line. However, an

analysis using classical random walk techniques [41] seems problematic. There are two main difficulties. Firstly, the k “walks” (run lengths) are correlated. If a run is long, the adjacent runs are likely to be shorter, and vice versa. Secondly, when the vertex changing opinion is a singleton, then the lengths of the two adjacent runs are combined, so the total number of runs suddenly decreases in two.

In this setting, we will use the random walk view only to give a lower bound on the convergence time. For the upper bound, we use a different approach based on a potential function.

Upper Bound

We proceed to find an upper bound for the expected consensus time on the cycle. We begin by defining the potential function

$$\psi(\xi) = \sum_{i=1}^k \sqrt{\ell_i},$$

where ℓ_i are the lengths of the runs in the cycle. Observe that $\psi(\xi) = 0$ if and only if $k(\xi) = 0$. An important feature of ψ is that it is the sum of a strictly concave function of the run lengths ℓ_i ($i \in [k]$). Almost any other potential function with these properties would give similar results.

Lemma 63. *For any configuration ξ on the n -cycle with k runs, $\psi(\xi) \leq \sqrt{kn}$.*

Proof. If $k = 0$, this is clearly true. Otherwise, if $k \geq 2$, by concavity, we have $\psi(\xi)/k = \frac{1}{k} \sum_{i=1}^k \sqrt{\ell_i} \leq \sqrt{\frac{1}{k} \sum_{i=1}^k \ell_i} = \sqrt{n/k}$, so $\psi(\xi) \leq \sqrt{kn}$. $\square$

Observe that $k(\xi_{t+1}) = k(\xi_t)$ at step t of either the push or pull process, unless vertex changing its colour is a singleton, in which case we might have $k(\xi_{t+1}) = k(\xi_t) - 2$.

Let $v_t = v \in D(\xi_t)$ be the active vertex at time $t < \tau_{\text{cons}}$, i.e., the vertex selected to push in the push rule, or to pull in the pull rule. Let δ_v be the expected change in ψ given the current configuration ξ_t and the active vertex v_t , i.e.

$$\delta_v = \mathbb{E}[\psi(\xi_{t+1}) - \psi(\xi_t) | v_t = v, \xi_t].$$

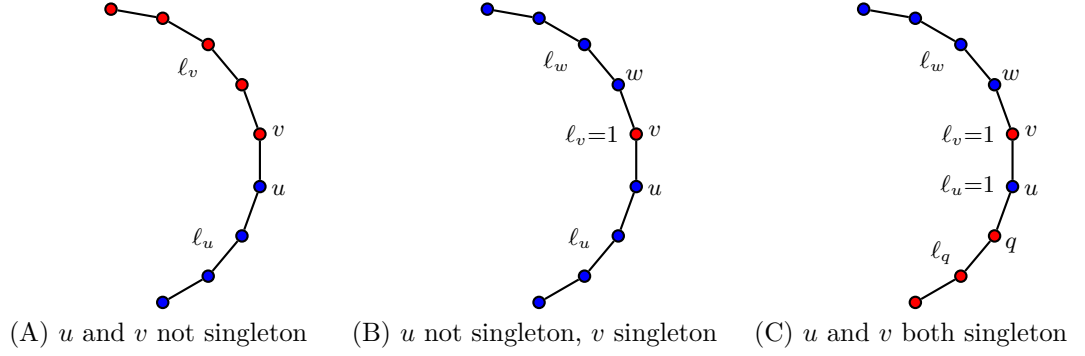


Figure 4.3: Cases for discordant edge uv . ℓ'_x represents the length of the run containing discordant vertex x , for $x \in \{u, v, w, q\}$.

If there are $|D|$ discordant vertices, the total expected change δ in ψ is

$$\Delta = \mathbb{E}[\psi(\xi_{t+1}) - \psi(\xi_t) | \xi_t] = \left(\frac{1}{|D|} \sum_{v \in D} \delta_v \right) \mathbf{1}_{\{t < \tau_{\text{cons}}\}}. \quad (4.25)$$

We will show that Δ is negative, so $\psi(\xi_t)$ is, in expectation, monotonically decreasing with respect to t . Unfortunately, we cannot simply bound δ_v for each $v \in D$, since it is possible to have $\delta_v > 0$. Thus, we consider discordant *edges*. We partition the set K of discordant edges uv into three subsets:

1. $A = \{uv : u \text{ and } v \text{ not singleton}\},$
2. $B = \{uv : u \text{ not singleton, } v \text{ singleton}\},$
3. $C = \{uv : u \text{ and } v \text{ both singleton}\}.$

The partition above is explained in Figure 4.3.

Note that the total number of runs k may change only if $uv \in B \cup C$. For each $uv \in K$ define δ_{uv} as

$$\delta_{uv} = \begin{cases} \delta_u + \delta_v, & uv \in A; \\ \delta_u + \frac{1}{2}\delta_v, & uv \in B; \\ \frac{1}{2}\delta_u + \frac{1}{2}\delta_v, & uv \in C. \end{cases}$$

Observe that each singleton is in two discordant edges, and all other discordant vertices in only one. Moreover, note that each run is bounded by two discordant vertices. Therefore

$$\delta = \frac{1}{|D|} \sum_{v \in D} \delta_v = \frac{1}{|D|} \sum_{uv \in K} \delta_{uv}. \quad (4.26)$$

We proceed to show that $\delta_{uv} < 0$, for all $uv \in K$. We consider the sets A , B and C separately. So far, the analysis has been identical for pull and push voting. Now, we must distinguish them. First we consider the discordant push process.

Push Voting

We use the following useful notation. For a discordant vertex $v \in D$, we write ℓ'_v for the size of the run of v (i.e., the run containing v). Now, we analyse the change δ_{uv} for each edge uv in the sets A , B , and C .

Set A . Suppose edge $uv \in A$. If u pushes, then the run of u increases its size, while the run of v decreases it (and similarly if v pushes), then

$$\delta_v = \sqrt{\ell'_v + 1} - \sqrt{\ell'_v} + \sqrt{\ell'_u - 1} - \sqrt{\ell'_u},$$

and

$$\delta_u = \sqrt{\ell'_v - 1} - \sqrt{\ell'_v} + \sqrt{\ell'_u + 1} - \sqrt{\ell'_u}.$$

Hence

$$\begin{aligned} \delta_{uv} &= (\sqrt{\ell'_v + 1} + \sqrt{\ell'_v - 1} - 2\sqrt{\ell'_v}) + (\sqrt{\ell'_u + 1} + \sqrt{\ell'_u - 1} - 2\sqrt{\ell'_u}) \\ &\leq -\frac{1}{4}((\ell'_v)^{-3/2} + (\ell'_u)^{-3/2}). \end{aligned} \quad (4.27)$$

The last inequality is given by Lemma 64 below.

Lemma 64. *For all $\ell \geq 1$, $\sqrt{\ell + 1} + \sqrt{\ell - 1} \leq 2\sqrt{\ell} - \frac{1}{4}\ell^{-3/2}$.*

Proof. First, we prove the inequality $\sqrt{1+x} + \sqrt{1-x} \leq 2 - \frac{1}{4}x^2$, for all $x \leq 1$. By squaring both sides, the inequality is true if $2 + 2\sqrt{1-x^2} \leq 4 - x^2 + \frac{1}{16}x^4$. This is true

if $\sqrt{1-y} \leq 1 - \frac{1}{2}y$, with $y = x^2$. Squaring both sides, this is $1 - y^2 \leq 1 - y^2 + \frac{1}{4}y^4$, which is clearly true. Now, letting $x = 1/\ell$, $\sqrt{\ell+1} + \sqrt{\ell-1} \leq 2\sqrt{\ell} - \frac{1}{4}\ell^{-3/2}$ is equivalent to $\sqrt{1+x} + \sqrt{1-x} \leq 2 - \frac{1}{4}x^2$ with $x \leq 1$. $\square$

Set B. Suppose edge $uv \in B$ where v is a singleton. Let w be the other discordant neighbour of v (See Figure 4.3.B). In one hand, if vertex v is chosen, it pushes its opinion on u or w with equal probability, then

$$\begin{aligned} \delta_v &\leq \frac{1}{2}(\sqrt{\ell'_u - 1} - \sqrt{\ell'_u} + \sqrt{2} - 1 + \sqrt{\ell'_w - 1} - \sqrt{\ell'_w} + \sqrt{2} - 1) \\ &\leq \sqrt{2} - 1. \end{aligned} \quad (4.28)$$

Since $\sqrt{\ell-1} \leq \sqrt{\ell}$, we have $\delta_v \leq \sqrt{2} - 1$. On the other hand, if u pushes its opinion, then the runs of vertices u , v and w merge. Here, we distinguish two cases. If the runs of u and w are different (i.e., there are more than 2 runs in the configuration), then the runs of u, v and w merge

$$\delta_u = \sqrt{\ell'_w + \ell'_u + 1} - \sqrt{\ell'_w} - \sqrt{\ell'_u} - 1 \leq \sqrt{3} - 3. \quad (4.29)$$

The last inequality is from Lemma 65 below. If the runs of u and v are the same, then $\ell'_u = n - 1$ and therefore, when u pushes, the process reaches consensus (and the number of runs is 0), thus

$$\delta_u = -\sqrt{\ell'_u} - 1 = -\sqrt{n-1} - 1 \leq \sqrt{3} - 3. \quad (4.30)$$

Thus

$$\delta_{uv} \leq \frac{1}{2}(\sqrt{2} - 1) + \sqrt{3} - 3 < -1 \leq -\frac{1}{2}((\ell'_v)^{-3/2} + (\ell'_u)^{-3/2}). \quad (4.31)$$

Lemma 65. For all $\ell_1, \ell_2 \geq 1$, $\sqrt{\ell_1} + \sqrt{\ell_2} + 1 \geq \sqrt{\ell_1 + \ell_2 + 1} + (3 - \sqrt{3})$.

Proof. Consider $f(\ell_1, \ell_2) = \sqrt{\ell_1} + \sqrt{\ell_2} + 1 - \sqrt{\ell_1 + \ell_2 + 1} + (\sqrt{3} - 3)$. Then, for all $\ell_1, \ell_2 > 0$,

$$\frac{\partial f}{\partial \ell_i} = \frac{1}{2\sqrt{\ell_i}} - \frac{1}{2\sqrt{\ell_1 + \ell_2 + 1}} > 0 \quad (i = 1, 2).$$

Hence $f(\ell_1, \ell_2) \geq f(1, 1) = 0$ for all $\ell_1, \ell_2 \geq 1$. $\square$

Set C . Let $uv \in C$. Let w be the other discordant neighbour of v , and q the other discordant neighbours of u (See Figure 4.3.C). If v is pushing, then with half probability, it pushes to w , making the run of w shorter and its own run larger. With the other half probability, it pushes to u , thus the runs of v , u and q merge. Therefore

$$\delta_v \leq \frac{1}{2}(\sqrt{\ell'_w - 1} - \sqrt{\ell'_w} + \sqrt{2} - 1 + \sqrt{\ell'_q + 2} - \sqrt{\ell'_q - 2}).$$

From Lemma 65 with $\ell_1 = 1$, it holds that $\sqrt{\ell + 2} - \sqrt{\ell} - 2 \leq \sqrt{3} - 3$. Thus, using that $\sqrt{\ell - 1} \leq \sqrt{\ell}$, we obtain

$$\delta_v \leq \frac{1}{2}(\sqrt{2} - 1 + \sqrt{3} - 3) < -0.425.$$

By symmetry, the same analysis holds for δ_u , and thus $\delta_u < -0.425$. Using the definition of δ_{uv} for $uv \in C$ we get

$$\delta_{uv} < -0.425 \leq -\frac{1}{5}((\ell'_v)^{-3/2} + (\ell'_u)^{-3/2}).$$

From the above three cases, we conclude that for all $uv \in K$, it holds that

$$\delta_{uv} < -\frac{1}{5}((\ell'_v)^{-3/2} + (\ell'_u)^{-3/2}).$$

By using the result for set A , B and C , and Equation (4.26), we obtain

$$\delta = \frac{1}{|D|} \sum_{v \in D} \delta_v = \frac{1}{|D|} \sum_{uv \in K} \delta_{uv} \leq -\frac{1}{5|D|} \sum_{uv \in K} (\ell_v^{-3/2} + \ell_u^{-3/2}). \quad (4.32)$$

Observe that in the sum $\sum_{uv \in K} ((\ell'_v)^{-3/2} + (\ell'_u)^{-3/2})$, each run is represented two times (one time of each edge adjacent to the run), so in terms of the length of the runs $\ell_1, \dots, \ell_k$, we have that

$$\frac{1}{5|D|} \sum_{uv \in K} (\ell_v^{-3/2} + \ell_u^{-3/2}) = \frac{2}{5|D|} \sum_{i=1}^k \ell_i^{-3/2} \geq \frac{1}{5k} \sum_{i=1}^k \ell_i^{-3/2},$$

where in the last inequality we use that $|D| \leq 2k$. For Equation (4.25) and (4.26) we obtain

$$\mathbb{E}(\psi(\xi_{t+1})|\xi_t) \leq \psi(\xi_t) - \mathbb{1}_{\{t < \tau_{\text{cons}}\}} \left(\frac{1}{5\kappa} \sum_{i=1}^k \ell_i^{-3/2} \right). \quad (4.33)$$

Since $f(x) = x^{-3}$ is a convex function, then for any random variable X , it holds by Jensen's inequality that $\mathbb{E}(f(X)) \geq f(\mathbb{E}(X))$, so

$$\frac{1}{k} \sum_{i=1}^k \ell_i^{-3/2} \geq \left(\frac{1}{k} \sum_{i=1}^k \sqrt{\ell_i} \right)^{-3} = \left(\frac{k}{\psi(\xi_t)} \right)^3. \quad (4.34)$$

Therefore,

$$\mathbb{E}(\psi(\xi_{t+1})|\xi_t) \leq \psi(\xi_t) - \mathbb{1}_{\{t < \tau_{\text{cons}}\}} \left(\frac{1}{5} \left(\frac{k}{\psi(\xi_t)} \right)^3 \right) = \psi(\xi_t) - \mathbb{1}_{\{t < \tau_{\text{cons}}\}} \left(\frac{k^3}{5\psi(\xi_t)^3} \right). \quad (4.35)$$

Hence, by using Lemma 63,

$$\mathbb{E}(\psi(\xi_{t+1})|\xi_t) - \psi(\xi_t) \leq -\mathbb{1}_{\{t < \tau_{\text{cons}}\}} \left(\frac{1}{5} k^3 / (kn)^{3/2} \right) = -\mathbb{1}_{\{t < \tau_{\text{cons}}\}} \left(\frac{1}{5} (k/n)^{3/2} \right). \quad (4.36)$$

We proceed to use Equation (4.36) to bound the expected consensus time. To this end, we need to divide the process into different phases. Suppose that, initially, the process starts with $2r_0$ runs. For $r \leq r_0$, define T_r as

$$T_r = \min\{t \geq 0 : k(\xi_t) = 2r\}. \quad (4.37)$$

Clearly, T_r is a stopping time, and $T_r < T_{r-1}$. Moreover, for all $t \in [T_r, T_{r-1})$, we have that $k(\xi_t)$ is constant and equal to $2r$. We call such an interval, phase r . Observe that the time to reach consensus is equal to T_0 , and can be written as $T_0 = \sum_{i=1}^{r_0} T_i - T_{i-1}$. We proceed to find a bound for $\mathbb{E}(T_{r-1} - T_r)$ for $1 \leq r \leq r_0$.

Let $m_r = \mathbb{E}(T_{r-1} - T_r)$. Clearly, $m_r \geq 1$. Additionally, define $\gamma_r = \frac{1}{5} m_r (2r/n)^{3/2}$. Consider the process $Q_t = \psi(\xi_{T_r+t+1}) + (t+1)\gamma_r$ which stops at time $t = T_{r-1} - T_r$, i.e., $Q_t = Q_{t \wedge (T_{r-1} - T_r)}$.

Observe that Equation (4.36), implies that $(Q_t)_{t \geq 0}$ is a supermartingale with respect to $(\xi_{t+T_r})_{t \geq 0}$. Indeed, let ξ be a configuration with $2r$ runs. Then,

$$\begin{aligned} \mathbb{E}(Q_{t+1}|\xi_{t+T_r} = \xi, \{t < T_{r-1} - T_r\}) &= \mathbb{E}(\psi(\xi_{T_r+t+1})|\xi_{t+T_r} = \xi) + (t+1)\gamma_r \\ &= \mathbb{E}(\psi(\xi_1)|\xi_0 = \xi) + (t+1)\gamma_r \end{aligned} \quad (4.38)$$

$$\begin{aligned} &\leq \psi(\xi) + (t+1)\gamma_r - \frac{1}{5} (k/n)^{3/2} \\ &\leq \psi(\xi) + t\gamma_r = Q_t \end{aligned} \quad (4.39)$$

Equation (4.38) follows from the Strong Markov Property (See Theorem 87 in Appendix A). Then, since $Q_{t+1} = Q_t$ given the event $\{t \geq T_{r-1} - T_r\}$, we have that

$$\mathbb{E}(Q_{t+1} | \xi_{t+T_r}) \leq \mathbb{1}_{\{t < T_{r-1} - T_r\}} (\psi(\xi_{t_r+t}) + t\gamma_r) + \mathbb{1}_{\{t \geq T_{r-1} - T_r\}} Q_t = Q_t.$$

Since $(Q_t)_{t \geq 0}$ is a bounded supermartingale, and $T_{r-1} - T_r$ is a stopping time, by the optional stopping theorem (Theorem 86 in Appendix A), we have

$$\mathbb{E}(Q_0) \leq \mathbb{E}(Q_{T_{r-1}-T_r}) = \mathbb{E}(\psi(\xi_{T_{r-1}})) - m_r \gamma_r.$$

Observe that $\mathbb{E}(Q_0) = \mathbb{E}(\psi(\xi_{T_r}))$. To ease notation, define $\varphi_r = \mathbb{E}[\psi(\xi_{T_r})]$, then the equation above can be written as

$$\varphi_{r-1} - \varphi_r \geq \gamma_r m_r = \frac{1}{5} m_r (2r/n)^{3/2}, \quad (4.40)$$

and this equation holds for any $r \in 1, \dots, r_0$.

With the ingredients above, we proceed to find a bound for $\mathbb{E}(\tau_{\text{cons}})$. From Lemma 63, $\varphi_r \leq \sqrt{2rn}$. Then, from (4.40), we have $m_r \leq 5\sqrt{2rn}(2r/n)^{-3/2} \leq (5/2)n^2/r$. Thus

$$\mathbb{E}(\tau_{\text{cons}}) = \sum_{j=1}^{r_0} m_j \leq (5/2)n^2 \sum_{j=1}^{r_0} 1/j < (5/2)n^2(\ln r_0 + 1).$$

Since $r_0 \leq n/2$, this gives an absolute bound of $(5/2)n^2 \ln(en/2) = O(n^2 \log n)$. However, we can improve this with a more careful analysis.

Let $x_r = \varphi_r - \varphi_{r-1} \geq 0$, for $r \in [r_0]$, so $\varphi_r = \sum_{i=1}^r x_i \leq \sqrt{2rn}$. Also, from (4.40), we have $m_r \leq 5x_r(n/2r)^{3/2} = (5/2)\sqrt{2}n^{3/2}x_r/r^{3/2}$, so

$$\mathbb{E}\tau_{\text{cons}} = \sum_{j=1}^{r_0} m_j < 10\sqrt{2}n^{3/2} \sum_{j=1}^{r_0} x_j/r^{3/2}.$$

Thus, $\mathbb{E}\tau_{\text{cons}}$ is bounded above by T^* , the optimal value of the following linear program.

$$\begin{aligned} T^* &= \max (5/2)\sqrt{2}n^2 \sum_{r=1}^{r_0} x_r/r^{3/2} \\ \text{such that } \sum_{j=1}^r x_j &\leq \sqrt{2rn} & (r \in [r_0]) \\ x_j &\geq 0 & (j \in [r_0]). \end{aligned} \quad (4.41)$$

This linear program can be solved easily by a greedy procedure. In fact, it is a polymatroidal linear program [35], but we give a self-contained proof for this simple case, using linear programming duality.

Lemma 66. *Let $0 < b_1 < b_2 < \dots < b_\nu$ and $c_1 > c_2 > \dots > c_\nu > 0$. Then the linear program $\max \sum_{j=1}^\nu c_j x_j$ subject to $\sum_{j=1}^r x_j \leq b_r$, $x_r \geq 0$ ($r \in [\nu]$) has optimal solution $x_1 = b_1$, $x_j = b_j - b_{j-1}$ ($j = 2, 3, \dots, \nu$).*

Proof. This solution has objective function value $c_1 b_1 + c_2(b_2 - b_1) + \dots + c_\nu(b_\nu - b_{\nu-1})$. The dual linear program is $\min \sum_{i=1}^\nu b_i y_i$ subject to $\sum_{i=j}^\nu y_i \geq c_j$, $y_j \geq 0$ ($j \in [\nu]$), and has feasible solution $y_\nu = c_\nu$, $y_j = c_j - c_{j+1}$ ($j \in [\nu - 1]$). Then the dual objective function has value $b_\nu c_\nu + b_{\nu-1}(c_{\nu-1} - c_\nu) + \dots + b_1(c_1 - c_2)$. However,

$$c_1 b_1 + c_2(b_2 - b_1) + \dots + c_\nu(b_\nu - b_{\nu-1}) = b_\nu c_\nu + b_{\nu-1}(c_{\nu-1} - c_\nu) + \dots + b_1(c_1 - c_2).$$

Since the objective function values are equal, it follows that the two solutions are optimal in the primal and dual respectively. $\square$

Thus, the optimal solution to (4.41) is $x_r = \sqrt{2r} - \sqrt{2(r-1)} = \sqrt{2r}(1 - \sqrt{1 - 1/r}) \leq \sqrt{2/r}$, for $r \in [r_0]$, since $1 - y \leq \sqrt{1 - y}$ for $0 \leq y \leq 1$. Thus

$$\begin{aligned} T^* &\leq (5/2)\sqrt{2}n^2 \sum_{j=1}^{r_0} x_r/r^{3/2} \\ &\leq (5/2)\sqrt{2}n^2 \sum_{j=1}^{r_0} \sqrt{2}/(\sqrt{r} r^{3/2}) \\ &= 5n^2 \sum_{r=1}^{r_0} 1/r^2 \leq \frac{5\pi^2}{6}n^2, \end{aligned} \tag{4.42}$$

since $\sum_{r=1}^\infty 1/r^2 = \pi^2/6$. We conclude that $\mathbb{E}(\tau_{\text{cons}}) \leq T^* = O(n^2)$.

Pull Voting

The case of pull voting is similar, but the calculations for the sets A, B and C are changed as follows.

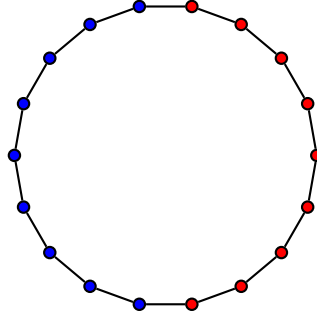


Figure 4.4: Lower bound configuration in C_n

Set A. The analysis for this case is identical the set A in the push case, except that δ_u and δ_v are interchanged. Hence $\delta_{uv} \leq -\frac{1}{4}(\ell_v^{-3/2} + \ell_u^{-3/2})$, as before.

Set B. $\delta_v = \sqrt{\ell_u + \ell_w + 1} - \sqrt{\ell_u} - \sqrt{\ell_w} - 1 \leq \sqrt{3} - 3$, using Lemma 65. Also $\delta_u = \sqrt{2} + \sqrt{\ell_u - 1} - \sqrt{\ell_u} - 1 \leq \sqrt{2} - 1$. Thus $\delta_{uv} \leq \sqrt{2} - 1 + \frac{1}{2}(\sqrt{3} - 3) < -0.22 \leq -\frac{1}{10}(\ell_v^{-3/2} + \ell_u^{-3/2})$.

Set C. $\delta_v = \sqrt{\ell_w + 2} - \sqrt{\ell_w} - 2 < \sqrt{3} - 3$, from Lemma 65 with $\ell = 1$. Similarly $\delta_u < \sqrt{3} - 3$, so $\delta_{uv} \leq \sqrt{3} - 3 < -1.25 < -\frac{1}{2}(\ell_v^{-3/2} + \ell_u^{-3/2})$.

Hence, we have $\delta_{uv} < -\frac{1}{10}(\ell_v^{-3/2} + \ell_u^{-3/2})$ for all $uv \in K$, whereas we had $\delta_{uv} < -\frac{1}{5}(\ell_v^{-3/2} + \ell_u^{-3/2})$ for push voting. Thus, the estimated rate of convergence is only half than the one for push voting. The rest of the analysis follows the same lines as before, except that the convergence time estimates are doubled. However, we still conclude that $\mathbb{E}(\tau_{\text{cons}}) = O(n^2)$.

Lower Bound

Consider an n -cycle, with n even (odd case is similar). To prove $\mathbb{E}\tau_{\text{cons}} = \Omega(n^2)$, we consider the configurations with two runs of equal length, one blue and one red. That is, we have $k = 2$ and $\ell_1 = \ell_2 = n/2$. See Figure 4.4.

Let R_t be the number of red vertices at time step t , i.e., R_t represents the length of the red run at time t if the process have not finished. Given that $1 < R_t < n - 1$ then the number of discordant vertices is four, and two of them are red. Then, the

number of red vertices increases in 1 with probability $1/2$, and decreases in 1 with the same probability. Therefore, R_t behaves like a random walk when $R_t \in [1, n-1]$. Observe that to finish the process, it is necessary that at some point the number of red vertices is either 1 or $n-1$. Since R_t behaves like a random walk, the time to reach 1 or $n-1$ starting from $R_0 = n/2$ is equal to $(n/2 - 1)^2/4 = \Omega(n^2)$, obtaining the claimed lower bound.

4.4 Voting on the Star Graph S_n

We proceed to prove Theorem 57 for the case of the star graph S_n on n vertices. In view of the fact that all the leaves of the star behave in the same way, each colouring of the star can be represented by a triple (r, b, X) , where r is the number of red leaves, b the number of blue leaves, and X represents the colour of the central vertex, i.e., $X \in \{R, B\}$. Observe that $r + b = n - 1$. Note that interactions only occurs between the central vertex and the set of leaves with different colour. We proceed to write the transition probabilities for the Markov chain $(Z_t)_{t \geq 0}$ moving on state space $\{(r, b, X)\}$. This transition are different for discordant push and discordant pull voting.

Push Voting on the Star

We begin by writing the transition probabilities for discordant push voting. Note that there are two possible actions, either the central vertex pushes its opinion on a leave, or a leaf pushes its opinion on the central vertex. Then, the transitions from state (r, b, R) to state $(r + 1, b - 1, R)$ happen with probability $1/(b + 1)$, and to state (r, b, B) with probability $b/(b + 1)$. Similarly, the transition probability from state (r, b, B) to state $(r - 1, b + 1, B)$ is $1/(r + 1)$, and to state (r, b, R) is $r/(r + 1)$. For purposes of the discussion, we group the states (r, b, R) and (r, b, B) into a single pseudo-state $S(r)$. $S(r)$ represents the two states with r red leaves, so r moves from 0

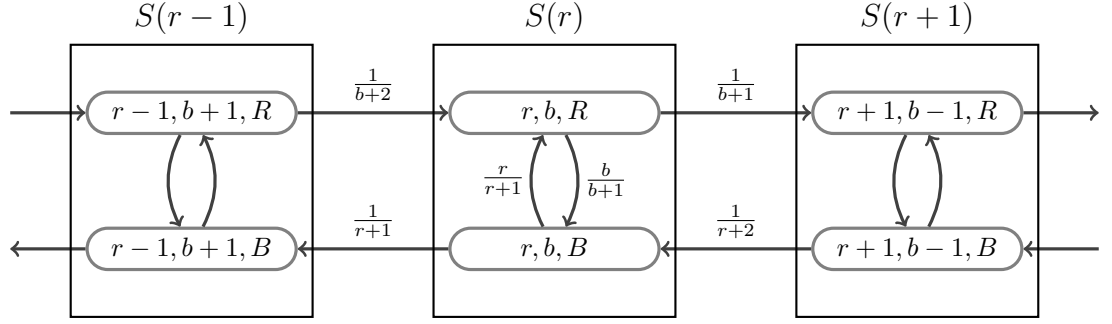


Figure 4.5: Pseudo-states for the push process on the Star

to $n-1$. Observe that when the Markov chain reaches pseudo states $S(0)$ or $S(n-1)$, it reaches them through state $(n-1, 0, R)$ or $(0, n-1, B)$, respectively. Those state represent colouring without discordant vertices and thus are the final state of the system. Thus, unless, we start from states $(0, n-1, R)$ or $(n-1, 0, B)$, from pseudo states $S(0)$ and $S(n-1)$, are essentially absorbing. The transition probabilities of this Markov chain are illustrated in figure 4.5.

We proceed to describe the transitions between pseudo-states $S(r)$. Let $X \in \{R, B\}$. Denote by $P_X(r, r+1)$ the probability that the next visited pseudo-state is $S(r+1)$ starting from pseudo-state $S(r)$ with central vertex X , and similarly, define $P_X(r, r-1)$ (Here, the definition does not consider transitions from $S(r)$ to $S(r)$). By analysing the transition of Figure 4.5, we have that

$$\begin{aligned}
 P_R(r, r+1) &= \frac{1}{b+1} \left(1 + \frac{b}{b+1} \frac{r}{r+1} + \cdots + \left(\frac{b}{b+1} \frac{r}{r+1} \right)^k + \cdots \right) \\
 &= \frac{1}{b+1} \frac{1}{1 - \frac{rb}{(r+1)(b+1)}} = \frac{r+1}{r+b+1} = \frac{r+1}{n}.
 \end{aligned} \tag{4.43}$$

The equalities above hold because we can only enter $S(r+1)$ via state $(r+1, b-1, R)$. If we start from (r, b, R) , the chain either jump in to state $(r+1, b-1, R)$ in one step, or we move to (r, b, B) and go back to (r, b, R) , and so on. In the second step, we use that $b+r$ is the total number of leaves, i.e., $n-1$.

Similarly,

$$P_B(r, r+1) = \frac{r}{r+1} P_R(r, r+1) = \frac{r}{n} \quad (4.44)$$

holds because in one step we jump to (r, b, R) and then we apply the previous computation of $P_R(r, r+1)$.

In any case, we have that from $S(r)$, the next visited pseudo-state is $S(r+1)$ with probability at most $(r+1)/n$ and at least r/n . That is, let $(Z_t)_{t \geq 0}$ be a Markov chain moving in states $\{(r, b, X)\}$, starting from pseudo state $S(r)$. Let T' be the first time $(Z_t)_{t \geq 0}$ moves to $S(r+1)$ or $S(r-1)$. Then,

$$\frac{r}{n} \leq \mathbb{P}(Z_{T'} \in S(r+1) | Z_0 \in S(r)) \leq \frac{r+1}{n}. \quad (4.45)$$

We need one last ingredient before computing lower and upper bounds for the expected consensus time. Observe that states (r, b, R) and (b, r, B) are equivalent in the sense that one is exactly the other after flipping the colours. Thus, the consensus time starting from (r, b, R) or (b, r, B) is the same. Therefore, we can glue states (r, b, R) with (b, r, B) into one, and thus we glue pseudo-states $S(0)$ and $S(n-1)$, $S(1)$ and $S(n-2)$, and in general $S(i)$ with $S(n-1-i)$. Note that if n is even then, $S(n/2-1)$ is glued with $S(n/2)$, while if n is odd, then $S((n-1)/2)$ is not glued with another state. For the rest of the discussion we assume n is odd. The case for n even can be handled similarly.

For $i = \{0, \dots, (n-1)/2\}$, define a new set of pseudo-states $Q(i)$ which contains the equivalent states $S((n-1)/2+i)$ and $S((n-1)/2-i)$. Without considering the self-transition probabilities, the pseudo-state $Q(0)$ is reflecting because from $S((n-1)/2)$ the chain jumps to either pseudo-state $S((n-1)/2-1)$ or to $S((n-1)/2+1)$ (both are members of $Q(1)$). Additionally, $Q((n-1)/2)$ is absorbing as it contains the two absorbing pseudo-states $S(0)$ and $S(n-1)$. In terms of the random walk $(Z_t)_{t \geq 0}$ moving on the space state $\{(r, b, X)\}$, denote by Z_t^Q the current pseudo-state $Q(i)$ to which Z_t belongs to, and define the stopping times $T_i = \min\{t > T_{i-1} : Z_t^Q \neq Z_{T_{i-1}}^Q\}$, and $T_0 = 0$. Then, define the thinned version of the original walk $(Z_t)_{t \geq 0}$, namely

$(\tilde{Z}_k)_{k \geq 0}$, by $\tilde{Z}_k = Z_{T_k}$. The process $(\tilde{Z}_k)_{k \geq 0}$ on state space $\{(r, b, X)\}$ is just the original Markov chain $(Z_t)_{t \geq 0}$ observed only when it transitioned to a new pseudo-state $Q(i)$. Due to the strong Markov property, the process $(\tilde{Z}_k)_{k \geq 0}$ is a Markov chain.

We are ready to compute upper and lower bounds for the expected consensus time. Fix a pseudo-state $Q(M)$. In order to estimate the expected hitting time of $Q(M)$, denoted by $T(M)$, we first count the transition between pseudo-states $Q(i)$, while ignoring the self-transitions. Then, we count the number of self transitions performed by the walk. In other words, we first estimate the value $K = K(M)$ such that $\tilde{Z}_K = Z_{T_K} \in Q(M)$. Then

$$\begin{aligned} T(M) &= (T_K - T_{K-1}) + (T_{K-1} - T_{K-2}) + \dots + (T_1 - T_0) \\ &= \sum_{k=1}^K (T_k - T_{k-1}) \\ &= \sum_{k=1}^{\infty} (T_k - T_{k-1}) \mathbf{1}_{\{K(M) \geq k\}}, \end{aligned}$$

and observe that $T_k - T_{k-1}$ is the number of self-jumps in the k -th pseudo-state Q visited by the Markov chain $(Z_t)_{t \geq 0}$. Note that $T_k - T_{k-1}$ is independent of $K(M)$, as the number of self-jump in the k -th pseudo-state is independent of the movement of the chain between pseudo-states $Q(i)$. As $T(M)$ is finite almost surely (finite state space), we have

$$\mathbb{E}(T(M)) = \sum_{k=1}^{\infty} \mathbb{E}((T_k - T_{k-1}) \mathbf{1}_{\{K(M) \geq k\}}). \quad (4.46)$$

Observe that

$$\mathbb{E}((T_k - T_{k-1}) \mathbf{1}_{\{K(M) \geq k\}}) = \mathbb{E}((T_k - T_{k-1}) | K(M) \geq k) \mathbb{P}(K(M) \geq k) \quad (4.47)$$

Let $A = \bigcup_{i=0}^{M-1} Q(i)$, i.e., A is the union of all states (r, b, X) such that they belong to one of the pseudo-states $Q(i)$ for $i \leq M$. Then the event $\{K(M) \geq k\}$ is equivalent to

$$\{K(M) \geq k\} = \bigcup_{s \in A^k} \{Z_{T_i} = s_i, \forall i \in \{0, \dots, k-1\}\},$$

that is, all possible path that do not use states in $Q(M)$. By the strong Markov property, it holds that

$$\begin{aligned}
& \mathbb{E}((T_k - T_{k-1}) \mid K(M) \geq k) = \\
& \mathbb{E} \left((T_k - T_{k-1}) \middle| \bigcup_{s \in A^k} \{Z_{T_i} = s_i, \forall i \in \{0, \dots, k-1\}\} \right) = \\
& \sum_{s \in A^{k+1}} \mathbb{E}((T_k - T_{k-1}) \mid Z_{T_{k-1}} = s_{k-1}) \frac{\mathbb{P}(Z_{T_i} = s_i, \forall i \in \{0, \dots, k\})}{\mathbb{P}(\bigcup_{s \in A^{k+1}} \{Z_{T_i} = s_i, \forall i \in \{0, \dots, k\}\})} = \\
& \sum_{s \in A^{k+1}} \mathbb{E}(T_1 \mid Z_0 = s_{k-1}) \frac{\mathbb{P}(Z_{T_i} = s_i, \forall i \in \{0, \dots, k\})}{\mathbb{P}(\bigcup_{s \in A^{k+1}} \{Z_{T_i} = s_i, \forall i \in \{0, \dots, k\}\})}
\end{aligned} \tag{4.48}$$

Observe that $\mathbb{E}(T_1 \mid Z_0 = s)$ denotes the expected number of time-steps needed to exit the pseudo-state of $s = (r, b, X)$, say $Q(i)$. We proceed to compute such an expectation. Recall that pseudo-state $Q(i)$ contains the pseudo-states $S((n-1)/2+i)$ and $S((n-1)/2-i)$, and those states are equivalent, so the time to leave them are the same. Without loss of generality, suppose $Z_0 = s \in S((n-1)/2+i)$. Let $r = (n-1)/2+i$ and $b = (n-1)/2-i$. The two states of $S(r)$ are (r, b, R) and (b, r, B) . For simplicity (and making an abuse of notation), we call those states by R and B . Then, transition from R to B happen with probability $b/(b+1)$ and from B to R with probability $r/(r+1)$, all other transitions move the chain outside $S(r)$ (see Figure (4.5)).

Let C_R and C_B be the number of steps before leaving pseudo-state $S(r)$ starting from state R and B , respectively, and let $\lambda = (br)/((b+1)(r+1))$. Then, for $k \geq 0$, we have

$$\mathbb{P}(C_R \geq k+2) = \frac{b}{b+1} \mathbb{P}(C_B \geq k+1) = \lambda \mathbb{P}(C_R \geq k),$$

and then

$$\begin{aligned}
\mathbb{P}(C_R \geq 2k) &= \mathbb{P}(C_B \geq 2k) = \lambda^k, \\
\mathbb{P}(C_R \geq 2k+1) &= \frac{b}{b+1} \mathbb{P}(C_B \geq 2k), \\
\mathbb{P}(C_B \geq 2k+1) &= \frac{r}{r+1} \mathbb{P}(C_R \geq 2k).
\end{aligned} \tag{4.49}$$

Thus,

$$\begin{aligned}
\mathbb{E}(C_R) &= \sum_{k=1}^{\infty} \mathbb{P}(C_R \geq k) = \sum_{k=1}^{\infty} \mathbb{P}(C_R \geq 2k) + \sum_{k=0}^{\infty} \mathbb{P}(C_R \geq 2k+1) \\
&= \left(\sum_{k=0}^{\infty} \lambda^k \right) - 1 + \frac{b}{b+1} \sum_{k=0}^{\infty} \lambda^k \\
&= \frac{1}{1-\lambda} - 1 + \frac{b}{b+1} \left(\frac{1}{1-\lambda} \right) \leq 2/(1-\lambda) \\
&= \frac{2(b+1)(r+1)}{b+r+1} = \frac{2(b+1)(r+1)}{n}.
\end{aligned} \tag{4.50}$$

By setting $r = n - 1 - b$ and optimising over b , we obtain that $\mathbb{E}(C_R) \leq cn$, for some universal constant c independent of r, b , and n . By symmetry, we also have that $\mathbb{E}(C_B) \leq cn$. Therefore, if $s = (r, b, R)$, then $\mathbb{E}(T_1 | Z_0 = s) = \mathbb{E}(C_R)$, otherwise $s = (r, b, B)$, and then $\mathbb{E}(T_1 | Z_0 = s) = \mathbb{E}(C_B)$. In any case, we have that

$$\mathbb{E}(T_1 | Z_0 = s) \leq cn.$$

By replacing the above in Equation (4.48), we obtain

$$\begin{aligned}
&\sum_{s \in A^{k+1}} \mathbb{E}(T_1 \mid Z_0 = s_{k-1}) \frac{\mathbb{P}(Z_{T_i} = s_i, \forall i \in \{0, \dots, k\})}{\mathbb{P}(\bigcup_{s \in A^{k+1}} \{Z_{T_i} = s_i, \forall i \in \{0, \dots, k\}\})} \\
&\leq \sum_{s \in A^{k+1}} cn \frac{\mathbb{P}(Z_{T_i} = s_i, \forall i \in \{0, \dots, k\})}{\mathbb{P}(\bigcup_{s \in A^{k+1}} \{Z_{T_i} = s_i, \forall i \in \{0, \dots, k\}\})} = cn.
\end{aligned} \tag{4.51}$$

Therefore, from Equation (4.47) it holds that

$$\mathbb{E}((T_k - T_{k-1}) \mathbb{1}_{\{K(M) \geq k\}}) \leq cn \mathbb{P}(K(M) \geq k).$$

Then, from Equation (4.46),

$$\mathbb{E}(T(M)) \leq cn \sum_{k \geq 1} \mathbb{P}(K(M) \geq k) = cn \mathbb{E}(K(M)). \tag{4.52}$$

To compute $K(M)$, we analyse the process $(\tilde{Z}_k)_{k \geq 0}$. Recall that $K(M) = \min\{k \geq 0 : \tilde{Z}_k \in Q(M)\}$. Additionally, remember that the process $(\tilde{Z}_k)_{k \geq 0}$ is the process $(Z_t)_{t \geq 0}$ viewed only at times T , such that Z_T moves to a new pseudo-state $Q(i)$. Using the bound of equations (4.45), we have that

$$\begin{aligned} & \mathbb{P}(\tilde{Z}_{k+1} \in Q(r+1) | \tilde{Z}_k \in Q(r)) \\ &= \mathbb{P}(\tilde{Z}_{k+1} \in S((n-1)/2 + r + 1) | \tilde{Z}_k \in S((n-1)/2 + r)) \\ &\geq \frac{1}{2} - \frac{1}{2n} + \frac{r}{n} \geq \frac{1}{2} + \frac{r}{n} - \frac{1}{n}, \end{aligned} \tag{4.53}$$

and

$$\mathbb{P}(\tilde{Z}_k \in Q(1) | \tilde{Z}_{k-1} \in Q(0)) = 1,$$

thus we can couple the process $(\tilde{Z}_k)_{k \geq 0}$ with a birth-and-death chain $(W_k)_{k \geq 0}$ on space state $\{Q(0), \dots, Q((n-1)/2)\}$ with $Q(0)$ being a reflecting barrier, $Q((n-1)/2)$ absorbing, and $p_i = 1/2 + i/n - 1/n$ (This is a Push chain with $\delta = -1$, see Section 4.2.1). It is clear that this new chain reaches state M slower than the process $(\tilde{Z}_k)_{k \geq 0}$, because $(W_k)_{k \geq 0}$ has less probability to go to greater states (see Lemma 61 for the formal statement). Using the result of Lemma 59, we obtain that

$$\mathbb{E}_0(K(M)) = \mathcal{O}(n \log M),$$

for any $M \in \{2, \dots, (n-1)/2\}$. Therefore, from Equation (4.52) we obtain

$$\mathbb{E}_0(T(M)) = Cn\mathbb{E}(K(M)) = \mathcal{O}(n^2 \log M). \tag{4.54}$$

Replace M with $(n-1)/2$ to obtain the desired upper bound.

For the lower bound, fix $M = \lfloor n^{3/5} \rfloor$. It is not hard to see that for $i \leq M$, it holds that $((n-1)/2 + i) \sim ((n-1)/2 - i) \sim n/2$, asymptotically. Then, since $Q(i) = S((n-1)/2 + i) \cup S((n-1)/2 - i)$, it holds from Equation (4.50) that $\mathbb{E}(C_R) \geq c'n$ for a small constant $c' > 0$, and by symmetry the same holds for $\mathbb{E}(C_B)$. Therefore, from Equation (4.46) we get

$$\mathbb{E}(T(M)) \geq cn\mathbb{E}(K(M)).$$

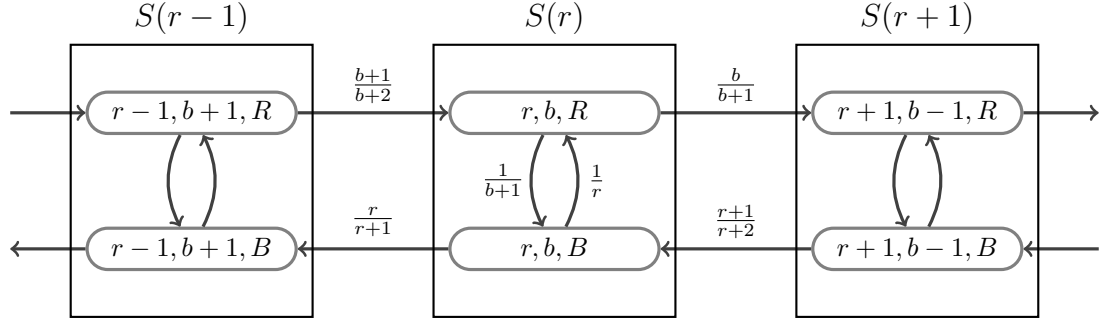


Figure 4.6: Pseudo-states for the pull process on the Star

We compute a lower bound for $\mathbb{E}(K(M))$. Similarly as we did for the upper bound, we couple the process $(\tilde{Z}_k)_{k \geq 0}$ with a birth-and-death chain on space state $\{Q(0), \dots, Q((n-1)/2)\}$ with $Q(0)$ being a reflecting barrier, $Q((n-1)/2)$ absorbing and $p_i = 1/2 + i/n + 1/n$, which is an upper bound for the probability of $(\tilde{Z}_k)_{k \geq 0}$ transitioning from $Q(i)$ to $Q(i+1)$. Then, the new birth-and-death chain reaches state M faster. From Lemma (60), we obtain that

$$\mathbb{E}_0(K(M)) = \Omega(n \log n),$$

therefore, we conclude that

$$\mathbb{E}_0(T(M)) = \Omega(n^2 \log n).$$

Pull Voting on the Star

As for discordant push voting on the star, we group the states (r, b, R) and $(r-1, b+1, B)$ into a single pseudo-state $S(r)$. The transition probabilities within or between $S(r+1)$ or $S(r-1)$ are shown in Figure 4.6, and are obtained by calculations similar to the push case. In the final pseudo-state $S(n-1)$ on the right, the state $(n, 0, R)$ is absorbing. Hence, the state $(n, 0, B)$ cannot be reached, unless we start from it, which is not important for finding an upper bound, since in one time-step, we move away for such a state. By symmetry, a similar situation happens with $S(0)$. Our objective is to find an upper and lower bound to the hitting time of $S(n-1)$ or

$S(0)$.

Suppose the chain is in state (r, b, R) of $S(r)$. The probability of a direct transition from (r, b, R) to $(r + 1, b - 1, R)$ is $b/(b + 1)$. This occurs when a blue leaf vertex is chosen and pulls the colour of the red central vertex. We say a *run* is a sequence of transitions which keeps the colour of the central vertex unchanged. Let $\rho(r, x, R)$ be the run given by the sequence of transitions

$$(r, b, R) \rightarrow (r + 1, b - 1, R) \rightarrow \cdots \rightarrow (x - 1, n - x + 1, R) \rightarrow (x, n - x, R).$$

Then

$$\mathbb{P}(\rho(r, x, R)) = \frac{n - r}{n - r + 1} \frac{n - r - 1}{n - r} \cdots \frac{n - x + 1}{n - x + 2} = \frac{n - x + 1}{n - r + 1}.$$

The probability a run starting from $(r, n - r, R)$ finishes by absorption at $(n, 0, R)$ is

$$\mathbb{P}(\rho(r, n, R)) = \frac{1}{n - r + 1} \geq \frac{1}{n}.$$

Each run is terminated by absorption, or by a change of colour of the central vertex, say from R to B . In the latter case, this marks the start of a new run (possibly of length zero) in the opposite direction. Starting from $(r, n - r, R)$, let X be the number of changes of colour of the central vertex from R to B , or vice versa, before absorption at $(n, 0, R)$ or $(0, n, B)$. Let L_i the length of the i -th run, then the time for consensus τ_{cons} , of the pull process is given by

$$\tau_{\text{cons}} = (X - 1) + \sum_{i=1}^X L_i.$$

Here, the $(X - 1)$ term represents the number of time-steps where the central vertex change colour. Clearly, $L_i \leq n - 1$, as there are at most $n - 1$ discordant vertices. Thus,

$$\tau_{\text{cons}} \leq X - 1 + (n - 1)X \leq nX.$$

Note that, given the initial position of the run, each run is independent of all previous runs. Moreover, each run finishes in absorption with probability at least $1/n$. Let Y be the first occurrence of a success of independent trials with success probability $1/p$, i.e., Y is a geometric random variable with success probability $1/p$, then it is not hard to see that $\mathbb{E}X \leq \mathbb{E}Y = n$, so $\mathbb{E}(\tau_{\text{cons}}) = O(n^2)$.

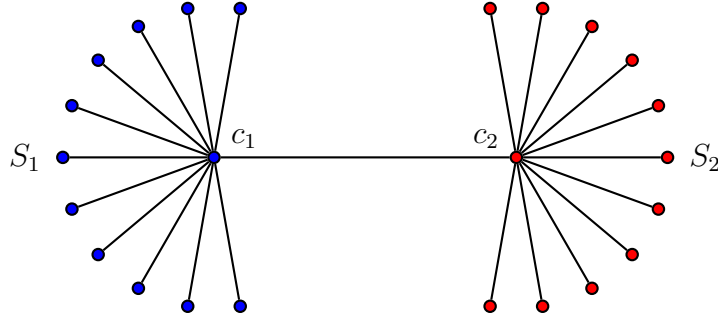


Figure 4.7: Double star S_{24}^* with half of the vertices coloured blue, and half coloured red.

4.5 Voting on the Double Star

Push Voting on the Double Star

A *double star* S_{2n+2}^* comprises two stars S_1, S_2 , each with n leaves, and their central vertices c_1, c_2 joined by an edge. Let $\xi_t : V \rightarrow \{R, B\}$ be the configuration of colours of the vertices $v \in V$ at time t . We will show that the expected consensus time for the push process on S_{2n+2}^* can be exponential in n .

Theorem 67. *The push process on the double star with $2n+2$ vertices has worst-case convergence time $\Omega(2^{2n/5})$.*

Proof. We assume that the initial configuration ξ_0 satisfies $\xi_0(v) = B$ for $v \in S_1$, and $\xi_0(v) = R$ for $v \in S_2$ (See Figure 4.7). Then, in order to achieve consensus, we must have that either the vertices of S_1 become R or the vertices of S_2 become B . Let T be the first time S_1 becomes R and let T' be the first time S_2 becomes B . Observe that they can be potentially infinite, but at least one of them must be finite. Then, we have that the consensus time τ_{cons} satisfies

$$\tau_{\text{cons}} \geq \min(T, T').$$

For now, we restrict our attention to S_1 in order to analyse T . Let $r(t)$ be defined as the number of red leaves in S_1 at time t . Define the sequence of stopping times

T_i as follows. $T_0 = 0$ and for $i \geq 1$, define $T_{i+1} = \min\{t > T_i : r_t \neq r_{T_i}\}$, i.e., T_i is the i -th time such that the central vertex c_1 pushes its opinion on a leave of S_1 . Define the process $(\tilde{r}_k)_{k \geq 0} = (r_{T_k})_{k \geq 0}$. To avoid confusion, we denote by t the times referring to the actual discordant process, and by k the times referring the process $(\tilde{r}_k)_{k \geq 0}$. For $0 < M \leq n$, define $K(M)$ as $\min\{k : \tilde{r}_k = M\}$. Clearly $K(M) \leq T$. This is because the $K(M) \leq K(n) \leq \tau_{\text{cons}}$.

With the assumptions above, the process $(\tilde{r}_k)_{k \geq 0}$ is not Markovian. This is because we do not have information about the colour of the central vertex c_1 . Nevertheless, we can condition on the colour of c_1 at time T_k . Let X_t the colour of c_1 and Y_t the colour of c_2 , at time t of the process $(r_t)_{t \geq 0}$. Then, it holds

$$\begin{aligned}
\mathbb{P}(\tilde{r}_{k+1} = r + 1 | \tilde{r}_k = r) &= \mathbb{P}(\tilde{r}_{k+1} = r + 1 | \tilde{r}_k = r, X_{T_k} = R) \mathbb{P}(X_{T_k} = R | \tilde{r}_k = r) \\
&\quad + \mathbb{P}(\tilde{r}_{k+1} = r + 1 | \tilde{r}_k = r, X_{T_k} = B) \mathbb{P}(X_{T_k} = B | \tilde{r}_k = r) \\
&= \mathbb{P}(\tilde{r}_{k+1} = r + 1 | \tilde{r}_k = r, X_{T_k} = R) \mathbb{P}(X_{T_k} = R | \tilde{r}_k = r).
\end{aligned} \tag{4.55}$$

The last equality holds because if the central edge is B , the number of red leaves cannot increase. We proceed to compute $\mathbb{P}(\tilde{r}_{k+1} = r + 1 | \tilde{r}_k = r, X_{T_k} = R)$. By the Strong Markov property, such a probability is equivalent to $\mathbb{P}(r_{T_1} = r + 1 | r_0 = r, X_0 = R)$. We compute an upper bound for it. We assume that the colour of c_2 (the whole S_2) is always R . The latter assumption clearly increases the probability above as the colour of X_0 has more chances to be R , and then, the next time that c_1 pushes its opinion, it is more likely it will be pushing a red opinion. At this point, in the neighbourhood of c_1 , there are $r + 1$ vertices of colour R (the red leaves and c_2) and b vertices of colour B (just the leaves) with $r + b = n$. Ignoring the actions of the leaves of S_2 , and assuming c_2 is always red, there are two possible events that do not decrease the number of red vertices. i) the central vertex pushes its opinion, increasing the number of red leaves, or ii) a blue neighbour of c_1 pushes its opinion, changing the central vertex to blue. Then, it is necessary that a red neighbour of the central vertex pushes its opinion to change back the central colour to red, returning

to the original state of the process. The first event happen with probability $(b+1)^{-1}$, the second with probability $(b/(b+1))((r+1)/(r+2))$. Then, we get

$$\begin{aligned}
\mathbb{P}(r_{T_1} = r+1 | r_0 = r, X_0 = R) &\leq \mathbb{P}(r_{T_1} = r+1 | r_0 = r, X_0 = R, Y_0 = R) \\
&\leq \sum_{j=0}^{\infty} \frac{1}{b+1} \left(\frac{b}{b+1} \frac{r+1}{r+2} \right)^j \\
&= \frac{r+2}{r+b+2} = \frac{r+2}{n+2}.
\end{aligned} \tag{4.56}$$

Notice that for $r \leq (n-8)/5$, the above probability is less or equal than $1/5$. From now, we are interested in the hitting time of state $\lfloor (n-8)/5 \rfloor$ of the process $(\tilde{r}_k)_{k \geq 0}$. We couple the process $(\tilde{r}_k)_{k \geq 0}$ with a biased random walk $(Y_k)_{k \geq 0}$ on states $\{0, \dots, \lfloor (n-8)/5 \rfloor\}$ with parameter $p = 1/5$ and reflecting barrier at 0. Since, for $i \leq \lfloor (n-8)/5 \rfloor$, the probability $p = 1/5$ is greater than the probability that $\tilde{r}_k$ increases in one step given $r_{k-1} = i$, then we can couple $(\tilde{r}_k)$ and the biased random walk $(Y_k)_{k \geq 0}$ so that $\tilde{r}_k \leq Y_k$ for all $k \geq 0$ (See Lemma 62). Then, $(Y_t)_{t \geq 0}$ has smaller hitting times for all states $M \leq \lfloor (n-8)/5 \rfloor$, i.e., smaller than $K(M)$ which in turn is less than T . From Equation (4.23), we have that

$$\begin{aligned}
\mathbb{P}(T > k) \geq \mathbb{P}(K(M) > k) &= \left(\frac{4^M - 4}{4^M - 1} \right)^k \\
&= \left(1 - \frac{3}{4^M - 1} \right)^k \\
&\leq 1 - \frac{3k}{4^M - 1} \leq 1 - \frac{k}{4^M}.
\end{aligned} \tag{4.57}$$

With the result above, we are ready to finish the argument including the rest of the graph, i.e., S_2 . Note that by symmetry, the T and T' have the same distribution (even though they are highly dependent). In particular, $\mathbb{P}(T \leq k) = \mathbb{P}(T' \leq k)$.

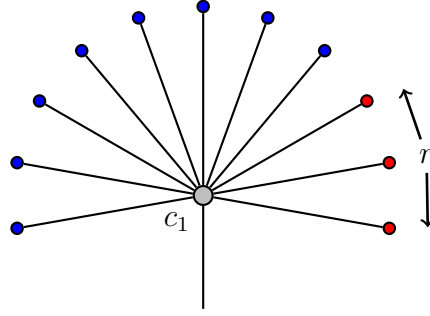


Figure 4.8: S_1 with r leaves coloured R

Then, by Markov's inequality, $(k+1)\mathbb{P}(\tau_{\text{cons}} \geq k+1) \leq \mathbb{E}(\tau_{\text{cons}})$, therefore

$$\begin{aligned}
\mathbb{P}(\tau_{\text{cons}}) &\geq k\mathbb{P}(\tau_{\text{cons}} > k) \\
&\geq k\mathbb{P}(\min(T, T') > k) \\
&= k(1 - \mathbb{P}(\{T \leq k\} \cup \{T' \leq k\})) \\
(\text{Union bound}) &\geq k(1 - 2\mathbb{P}(T \leq k)) \\
(\text{Equation (4.57)}) &= k(1 - 2(k4^{-M})) = k - 2k^24^{-M}. \tag{4.58}
\end{aligned}$$

Now choose $M = \lfloor (n-8)/5 \rfloor$ and $k = (1/4)4^M$, then $2k^24^{-M} = (1/8)4^M$ and thus

$$k - 2k^24^{-M} = (1/8)4^M - (1/16)4^M = \Theta(4^{n/5}) = \Theta(2^{2n/5}).$$

□

Pull Voting on the Double Star

Lemma 68. *Let τ_{cons} be the expected time to complete discordant pull voting on the double star of $2n+2$ vertices. Then for any starting configuration $\mathbb{E}\tau_{\text{cons}} = \mathcal{O}(n^3)$.*

Proof. Our proof mimics the proof for discordant pull voting on the star graph. If the centres c_1, c_2 are the same colour (say red) we call the central edge monochromatic. If the central vertices are both red (e.g.), a run is a sequence of steps in which a blue leaf vertex is chosen at each step and pulls the red colour from one of the central

vertices. Let ξ_t the configurations of colours at time t . Consider the stopping times $(M_i)_{i \geq 1}$ and $(T_i)_{i \geq 0}$ defined as following.

$$T_0 = 0,$$

and,

$$M_i = \mathbf{infec}\{t \geq T_{i-1} : \text{the central edge is monochromatic in } \xi_t\}.$$

and also,

$$T_i = \mathbf{infec}\{t \geq M_i : \text{the central edge is non-monochromatic in } \xi_t \text{ or } \xi_t \text{ is in consensus}\}.$$

With the above definition, we have $0 \leq M_1 \leq T_1 \leq M_2 \leq T_2 \dots$. Note that T_k denotes the end of a run. Such a run ends with a non-monochromatic edge or with consensus. Let $K = \min\{k \geq 0 : T_K = \tau_{\text{cons}}\}$ where τ_{cons} is the consensus time. Then, we have that τ_{cons} can be written as

$$\tau_{\text{cons}} = \sum_{k=1}^{\infty} ((T_k - M_k) + (M_k - T_{k-1})) \mathbf{1}_{\{K \geq k\}} \quad (4.59)$$

Here, $L_i = T_i - M_i$ represents the length of the i -th run, and $J_i = M_i - T_{i-1}$ represents the number of time-steps needed to have a monochromatic central edge after the end of the previous run. Since we are working in a finite graph, τ_{cons} is finite almost surely, and then

$$\begin{aligned} \mathbb{E}(\tau_{\text{cons}}) &= \sum_{k=1}^{\infty} \mathbb{E}((L_k + J_k) \mathbf{1}_{\{K \geq k\}}) \\ &= \sum_{k=1}^{\infty} \mathbb{E}((L_k + J_k) | \{K \geq k\}) \mathbb{P}(K \geq k). \end{aligned} \quad (4.60)$$

Observe that $L_k \leq 2n + 1$, since there are at most $2n + 1$ discordant edges in the graph as the central edge is monochromatic.

In order to compute $\mathbb{E}(J_k | \{K \geq k\})$, define $\mathcal{N}$ as the set of all configurations of colours ξ , such that they do not have a monochromatic central edge. Then, observe that $\{K \geq k\} = \{\xi_{T_{k-1}} \in \mathcal{N}\}$. Therefore

$$\mathbb{E}(J_k | \{K \geq k\}) = \sum_{\xi \in \mathcal{N}} \mathbb{E}(J_k | \xi_{T_{k-1}} = \xi) \frac{\mathbb{P}(\xi_{T_{k-1}} = \xi)}{\mathbb{P}(\xi_{T_{k-1}} \in \mathcal{N})}. \quad (4.61)$$

By the strong Markov property $\mathbb{E}(J_k | \xi_{T_{k-1}} = \xi) = \mathbb{E}(J_1 | \xi_0 = \xi)$, which is just the expected number of steps needed to obtain a monochromatic central edge starting from ξ . We proceed to compute its expected value.

Consider a configuration ξ whose central edge is non-monochromatic. Let r_1, b_1 be the number of R and B leaves in S_1 (resp. r_2, b_2 in S_2), and let $b_1 + b_2 = b$. Moreover, denote the probability of becoming monochromatic in one step by $\phi(r_1, b_1, r_2, b_2)$. Then

$$\phi(r_1, b_1, r_2, b_2) \geq \min \left\{ \frac{2}{b_1 + r_2 + 2}, \frac{2}{r_1 + b_2 + 2} \right\} \geq \frac{2}{2n + 2} = \frac{1}{n + 1}.$$

Here, the minimum is taken depending on whether c_1 is R or B . In both cases, we have that for a configuration in $\mathcal{N}$, the chances of becoming monochromatic is at least $1/(n + 1)$. Then, in expectation, we need at most $n + 1$ steps, and thus $\mathbb{E}(J_1 | \xi_0 = \xi) \leq n + 1$. Using this result and the fact that $L_k \leq 2n + 1$, from Equation (4.60) we obtain

$$\mathbb{E}(\tau_{\text{cons}}) \leq (3n + 2) \sum_{k=1}^{\infty} \mathbb{P}(K \geq k) = (3n + 2)\mathbb{E}(K). \quad (4.62)$$

To compute $\mathbb{E}(K)$, we use a similar argument as for $\mathbb{E}(J_1 | \xi_0)$. In particular, note that runs are independent given the initial configuration. Then, independently of the initial configuration, we compute an absolute lower bound p for the probability that a given run finishes, and then we couple K with a geometric random variable G with success probability p , such that $K \leq G$ and thus $\mathbb{E}(K) \leq p^{-1}$.

We proceed to find a universal lower bound for the probability that a run finishes in consensus. Without loss of generality, suppose the central edge is R and let b the number of B leaves in the double star. Suppose $b \geq 2$, then the probability that a leaf with colour B pulls the opinion is $b/(b + 2)$ if both c_1 and c_2 are discordant, and $b/(b + 1)$ in case only one is discordant. In any case, to obtain a lower bound of this probability we assume both are discordant. If $b = 1$ then, the probability that the only leaf with colour B pulls is $1/2$, because only one of c_1 and c_2 is discordant.

Then, the probability that the run finishes in consensus when there are initial b leaves with colour B is

$$\left(\frac{b}{b+2} \frac{b-1}{b+1} \frac{b-2}{b} \cdots \frac{4}{6} \frac{3}{5} \frac{2}{4} \right) \frac{1}{2} = \frac{3}{5(b+2)(b+1)} \geq \frac{3}{5(n+2)(n+1)}. \quad (4.63)$$

From our previous analysis, we get $\mathbb{E}(K) \leq 5(n+2)(n+1)/3$. Finally, from Equation (4.62) we deduce that the expected consensus time is

$$\mathbb{E}(\tau_{\text{cons}}) = \mathcal{O}(n^3).$$

□

4.6 Voting on the Barbell Graph

The barbell or dumbbell graph of $2n$ vertices, B_{2n} , is given by two disjoint cliques S_1 and S_2 of size n joined by a single edge e . Let $c_1 \in S_1$ and $c_2 \in S_2$ be the two vertices joined by the special edge. The barbell graph has $N = 2n$ vertices and $2\binom{n}{2} + 1$ edges.

Push Voting on the Barbell

In order to prove the exponential lower bound, we start with the following configuration: all vertices in S_1 are B , and all vertices in S_2 are R . We repeat the same argument used for discordant push voting on the double star, that is, to reach consensus either S_1 has to change its colour to all R , or S_2 to all B .

We use the same formalism as in the push case in the double star. Let T be the first time all S_1 is B and T' the first time all S_2 is R . Clearly, the consensus time satisfies $\tau_{\text{cons}} \geq \min(T, T')$, and note that T and T' can be potentially infinity, but at least one is finite.

We concentrate on T . Instead of computing the expected value of T , we count the number of actions concerning vertices of S_1 , say G . Also, in order to decrease the value of T and G , we assume that the colour of all vertices in S_2 is R , and

that they remain unchanged over time. This only helps to decrease G , as sometimes vertex $c_2 \in S_2$ will push an R colour on a vertex of S_1 via the edge joining both cliques. Let $(r_t)_{t \geq 0}$ be the number of R vertices in S_1 after the t -th action concerting vertices of S_1 , and let $(M_t)_{t \geq 0}$ be the number of discordant vertices. To compute $\mathbb{P}(r_{t+1} = r_t + 1 | r_t, M_t)$ we need to be careful. In particular, we need to consider the cases that the colour of c_1 is R or B . Let X be the colour of c_1 , and remember we fix the whole S_2 to be red, then

$$\mathbb{P}(r_{t+1} = r_t + 1 | r_t, M_t, X = B) = \frac{r_t + 1}{M_t} = \frac{r_t + 1}{n + 1}.$$

The above probability holds because there are r_t blue vertices in S_1 plus the extra red vertex c_2 . Also, note that all vertices of S_1 are discordant as well as vertex c_2 . Moreover, note that with probability $1/(n + 1)$ we sample c_1 to push, and with probability $1/n$ it pushes its opinion on c_2 , but since we set the opinion of c_2 to be fixed, this does not change b_t at all, then

$$\mathbb{P}(r_{t+1} = r_t | r_t, M_t, X = B) = \frac{1}{n(n + 1)}.$$

Finally,

$$\mathbb{P}(r_{t+1} = r_t - 1 | r_t, M_t, X = B) = 1 - \frac{r_t + 1}{n + 1} - \frac{1}{n(n + 1)}.$$

For the case $X = R$, a slightly simpler argument, gives us,

$$\mathbb{P}(r_{t+1} = r_t + 1 | r_t, M_t, X = R) = \frac{r_t}{n},$$

and

$$\mathbb{P}(r_{t+1} = r_t - 1 | r_t, M_t, X = R) = 1 - \frac{r_t}{n}.$$

Also, notice that $\mathbb{P}(r_t = 1 | r_t = 0) = 1$. In any case, in the regime $1 \leq r_t \leq M = \lfloor (n - 10)/5 \rfloor$, we have that

$$\max_{1 \leq r \leq M} \left\{ \max \left\{ \frac{r + 1}{n + 1} + \frac{1}{n(n + 1)}, \frac{r}{n} \right\} \right\} \leq 1/5,$$

and thus, from Lemma 62, the process $(r_t)_{t \geq 0}$ can be coupled with a biased random walk $(Y_t)_{t \geq 0}$ on state space $\{0, \dots, M\}$ with reflecting barriers and parameter $p = 1/5$.

Then, we have $r_t \leq Y_t$ for all $t \geq 0$, when they start from the same state. From here, rest of the argument is exactly the same as the argument for discordant push in the double star (see proof after Equation (4.56)). We conclude the expected consensus time is $\Omega(2^{2n/5})$ for the Barbell graph on $2n$ vertices.

Pull Voting on the Barbell

We suppose we have reached a configuration in which there is only one B vertex, which is symmetric to the case of having only one R vertex. Suppose the unique B vertex is in S_1 . We modify our process so that the system reaches consensus faster. To do that, at each round we only select vertices in S_1 , and assume the final colour will be red. If the final colour is blue, then we must also recolour all S_2 and we will eventually reach a configuration where all vertices are B except for one red vertex. Note that even if the vertex c_1 of the bridge edge $e = (c_1, c_2)$ is blue, the interaction between S_1 and S_2 does not affect the outcome. Indeed, it only slows the process because there are more red vertices that can potentially pull the blue vertex. So, in order to speed up the process, and thus obtain a lower bound on the expected consensus time, we erase the edge connecting both cliques, so S_1 became a clique, disconnected from S_2 .

Here, we have a clique with $N/2$ vertices. We use a result from the proof of Theorem 54 for K_n as given in Section 4.3. Inequality (4.24) shows that the expected time for pull voting to reach consensus in K_n , when all but one vertex is red is $\Omega(2^n)$. Hence, the expected time to finish in our modified process is $\Omega(2^n) = \Omega((\sqrt{2})^{2n})$.

Chapter 5

Discordant Voting Model on the Complete Graph

In this chapter, we introduce the discordant β -Push-Pull model, which corresponds to a generalisation of the discordant voting model on the complete graph K_n . Consider $\beta \in [0, 1]$, then the β -Push-Pull model is defined as follows. At each step, with probability β , we execute one step of discordant pull, otherwise, we execute one step of discordant push. Recall that discordant voting is asynchronous, i.e., in each round exactly one vertex changes its opinion. We remark that on the complete graph, push, oblivious, and pull models correspond to particular cases of the β -Push-Pull model, with $\beta = 0, 1/2$, and 1 , respectively.

In the previous chapter, it was proved that, in a complete graph of n vertices, the expected time to finish discordant push is $\Theta(n \log n)$, while for pull and oblivious models is $\Theta(2^n)$ and $\Theta(n^2)$, respectively. Therefore, as these models correspond to a particular instantiation of the β -Push-Pull model ($\beta = 0$ for push, $\beta = 1/2$ for oblivious, and $\beta = 1$ for pull), it is natural to assess how the expected consensus time $\mathbb{E}\tau_{\text{cons}}$ varies between the complexity orders, as β varies from zero to one. In particular, how exactly does the β -Push-Pull process make the transition in $\mathbb{E}\tau_{\text{cons}}$ from order $n \log n$ (push: $\beta = 0$), to order n^2 (oblivious: $\beta = 1/2$), and finally to

order 2^n (pull: $\beta = 1$). The following theorems give the answer.

Theorem 69. *Let $0 \leq \beta \leq 1/2$. Let $\mathbb{E}\tau_{\text{cons}}$ be the expected time to reach consensus in the β -Push-Pull process on the complete graph K_n , starting from $R_0, B_0 = n/2$.*

1. *If $\beta \in [0, 1/2)$ is independent of n , then $\mathbb{E}\tau_{\text{cons}} = \Theta\left(\frac{n \log n}{(1-2\beta)}\right)$.*
2. *If $\beta = 1/2 - \varepsilon$ where $\varepsilon \gg \frac{1}{n}$, then $\mathbb{E}\tau_{\text{cons}} = \Theta\left(\frac{n}{\varepsilon} \log(n\varepsilon)\right)$.*
3. *If $\beta = 1/2 - \varepsilon$ where $\varepsilon = \mathcal{O}\left(\frac{1}{n}\right)$, then $\mathbb{E}\tau_{\text{cons}} = \Theta(n^2)$.*

Theorem 70. *Let $1/2 \leq \beta \leq 1$. Let $\mathbb{E}\tau_{\text{cons}}$ be the expected time to reach consensus in the β -Push-Pull process on the complete graph K_n , starting from $R_0, B_0 = n/2$.*

1. *If $\beta = 1/2 + \varepsilon$ where $\varepsilon = \mathcal{O}\left(\frac{1}{n}\right)$, then $\mathbb{E}\tau_{\text{cons}} = \Theta(n^2)$.*
2. *If $\beta = 1/2 + \varepsilon$ where $\varepsilon \geq \frac{1}{n}$ but $\varepsilon = o(1)$, then*

$$\begin{aligned} \mathbb{E}\tau_{\text{cons}} &= \Theta(\sqrt{n}/\varepsilon^{3/2}) \exp\left(\frac{n}{2\varepsilon} \sum_{j \geq 1} \frac{1}{(2j-1)(2j)} (2\varepsilon)^{2j}\right) \\ &= \Theta(\sqrt{n}/\varepsilon^{3/2}) \exp(n\varepsilon(1 + \mathcal{O}(\varepsilon^2))). \end{aligned} \quad (5.1)$$

3. *If $\beta = 1/2 + \varepsilon = 1 - \delta$ where ε, δ are constants independent of n , then*

$$\mathbb{E}\tau_{\text{cons}} = \Theta(\sqrt{n}) \left(2(1-\delta)^{1-\delta} \delta^\delta\right)^{n/(1-2\delta)}. \quad (5.2)$$

4. *If $\beta = 1 - \delta$ where $\delta = o(1)$ but $\delta = \Omega(n^{-1})$, then*

$$\mathbb{E}\tau_{\text{cons}} = \Theta(\sqrt{n\delta}) 2^n (4\delta/e)^{n\delta}. \quad (5.3)$$

5. *If $\beta = 1 - \delta$ where $\delta \ll \frac{1}{n}$, then*

$$\mathbb{E}\tau_{\text{cons}} = \Theta(1) 2^n \exp(-\delta n \log n). \quad (5.4)$$

In particular,

6. *If $\beta = 1 - \delta$ where $\delta = \mathcal{O}\left(\frac{1}{n \log n}\right)$, then $\mathbb{E}\tau_{\text{cons}} = \Theta(2^n)$.*

From the previous result, we observe that the expected time to reach consensus transitions from $\Theta(n \log n)$ to $\Theta(n^2)$, and beyond, in a window of width $o(1)$ around $\beta = 1/2$. Therefore given a complexity order between $n \log n$ to 2^n , by using Theorems 69 and 70, we can always find a parameter $\beta \in [0, 1]$ such that the expected consensus time of the β -Push-Pull model attains such an order. As an example of this, for any $a > 1$, the value of $\beta \sim 1/2 + o(1)$ that gives us $\mathbb{E}\tau_{\text{cons}} = \Theta(n^a)$, is given by the following corollary.

Corollary 71. *To obtain an expected completion time of $\Theta(n^a)$, $a > 1$ constant, choose β as follows. Let $c > 0$ constant, then*

1. *If $1 < a < 2$ then put $\beta = 1/2 - \varepsilon$ where $\varepsilon = c(\log n)/n^{a-1}$.*
2. *If $a = 2$ put $\beta = 1/2 + \varepsilon$ where $|\varepsilon| = c/n$.*
3. *If $a > 2$ then put $\beta = 1/2 + \varepsilon$ where $\varepsilon = c((a - 2) \log n + (3/2) \log \log n) / n$.*

5.1 Birth-and-Death Chains

Remark 72. *This section contains various results presented in Section 4.2 of Chapter 4. In order to have more independent chapters, we repeat some of the results here.*

A Markov chain $(X_t)_{t \geq 0}$ is said to be a Birth-and-Death chain on state space $S = \{0, \dots, N\}$ if given $X_t = i$ then the possible values of X_{t+1} are $i + 1, i$, or $i - 1$, with probability p_i, r_i , and q_i , respectively. Note that $q_0 = p_N = 0$.

In this section, we assume that $r_i = 0, p_0 = 1, q_N = 1, p_i > 0$ for $i \in \{0, \dots, N-1\}$ and $q_i > 0$ for $i \in \{1, \dots, N\}$. Given a random variable Y , we denote by $\mathbb{E}_i Y$, the expected value of random variable Y when the chain starts from i (i.e., $X_0 = i$). Finally, we define the hitting time of state i as $T_i = \min\{t \geq 0 : X_t = i\}$.

We proceed to summarise the results that we require on Birth-and-Death chains (see Levin, Peres, and Wilmer [57, chapter 2.5]).

We say that a probability distribution π satisfies the detailed balance equations, if

$$\pi(i)P(i, j) = \pi(j)P(j, i), \text{ for all } i, j \in S. \quad (5.5)$$

Birth-and-Death chains with $p_i = P(i, i+1)$, $q_i = P(i, i-1)$ can be shown to satisfy the detailed balance equations. Thus, we can write $\pi(i)$ in terms of $\pi(0)$,

$$\pi(1) = \frac{\pi(0)p_0}{q_1} \quad \text{and} \quad \pi(i) = \pi(0) \frac{p_0 \cdots p_{i-1}}{q_1 \cdots q_i} \text{ for } i \geq 2. \quad (5.6)$$

By normalising π , we obtain a stationary distribution. The normalisation constant is

$$\pi(0) = \left(1 + \frac{p_0}{q_1} + \sum_{j=2}^n \frac{p_0 \cdots p_{j-1}}{q_1 \cdots q_j} \right)^{-1}. \quad (5.7)$$

As the Birth-and-Death process is recurrent, the stationary distribution satisfies

$$\pi(i) = \frac{1}{\mathbb{E}_i(T_i^+)},$$

where $\mathbb{E}_i(T_i^+)$ is the expected first return time to state i (starting from state i). It follows from this, (see e.g. [57]) that

$$\mathbb{E}_{i-1}T_i = \frac{1}{q_i\pi(i)} \sum_{k=0}^{i-1} \pi(k). \quad (5.8)$$

By replacing the value of π given in equation 5.6, we have that $\mathbb{E}_0T_1 = 1/p_0 = 1$, and in general,

$$\mathbb{E}_{i-1}T_i = \sum_{k=0}^{i-1} \frac{1}{p_k} \frac{q_{k+1} \cdots q_{i-1}}{p_{k+1} \cdots p_{i-1}}, \quad \text{for } i \in \{1, \dots, N\}. \quad (5.9)$$

For the sum above, we adopt the following convention. If $k = i-1$ then $\frac{q_{k+1} \cdots q_{i-1}}{p_{k+1} \cdots p_{i-1}} = 1$, so that the last term in the sum is $1/p_{i-1}$. Note also that the final index k on p_k is $k = N-1$, i.e., we never divide by $p_N = 0$.

Starting from state 0, let T_M be the number of transitions needed to reach state M for the first time. For any $M \leq N$, we have that $\mathbb{E}_0T_M = \sum_{i=1}^M \mathbb{E}_{i-1}T_i$. For example, $\mathbb{E}_0T_1 = \frac{1}{p_0} = 1$ and $\mathbb{E}_0T_2 = 1 + \frac{1}{p_1} + \frac{q_1}{p_0p_1}$ etc. Thus, for $M \geq 1$

$$\mathbb{E}_0T_M = \sum_{i=1}^M \mathbb{E}_{i-1}T_i = \sum_{i=1}^M \sum_{k=0}^{i-1} \frac{1}{p_k} \prod_{j=k+1}^{i-1} \frac{q_j}{p_j}. \quad (5.10)$$

5.1.1 Push and Pull Chains

From now, we assume n is even, and denote $N = n/2$. In order to find the time to reach consensus in the β -Push-Pull process, we can consider the chain $Z_t = \max\{R_t, B_t\} - n/2$, where R_t denotes the number of red vertices at time t , $B_t = n - R_t$ represents the number of blue vertices. Note that $(Z_t)_{t \geq 0}$ is a Birth-and-Death chain with state space $\{0, 1, \dots, N\}$. We call $(Z_t)_{t \geq 0}$ the β -Push-Pull chain. Observe as well that the consensus time is equivalent to the first time that $Z_t = N$, i.e., the hitting time T_N . We proceed to define two Birth-and-Death chains which underlie our analysis. Those chains have states $\{0, 1, \dots, N\}$. The transition probabilities from state i are given by $P(i, i+1)$, and $Q(i, i+1) = 1 - P(i, i+1)$.

Push Chain. Let Y_t be the state occupied by the push chain at step $t \geq 0$. The transition probabilities $P_i = P(i, i+1)$ from $Y_t = i$ are given by

$$P_i = \begin{cases} 1, & \text{if } i = 0 \\ 1/2 + i/n, & \text{if } i \in \{1, \dots, n/2 - 1\} \\ 0, & \text{if } i = n/2 \end{cases} \quad (5.11)$$

Pull Chain. Let $\bar{Y}_t$ be the state occupied by the pull chain at step $t \geq 0$. Given that $\bar{Y}_t = i$, the transition probabilities $\bar{P}_i = \bar{P}(i, i+1)$ are given by

$$\bar{P}_i = \begin{cases} 1, & \text{if } i = 0 \\ 1/2 - i/n, & \text{if } i \in \{1, \dots, n/2 - 1\} \\ 0, & \text{if } i = n/2 \end{cases} \quad (5.12)$$

For $1 \leq i \leq N - 1$, the pull chain is the push chain with reversed probabilities, i.e., $\bar{P}_i = Q_i$.

The β -Push-Pull chain $(Z_t)_{t \geq 0}$ which represents the β -Push-Pull process is a mixture of the push and pull chains with transition probabilities $p_i = (1 - \beta)P_i + \beta\bar{P}_i$. The transition probabilities of $Z_t = \max\{R_t, B_t\} - n/2$ for the β -Push-Pull chain are

as follows. Let $p_i = \mathbb{P}(Z_{t+1} = Z_t + 1 \mid Z_t = i)$, then

$$p_i = \begin{cases} 1, & \text{if } i = 0 \\ 1/2 + (1 - 2\beta)i/n, & \text{if } i \in \{1, \dots, n/2\} \\ 0, & \text{if } i = n/2 \end{cases} \quad (5.13)$$

If $n = 2N + 1$ is odd then the only difference is that the states i of the chain are $i \in \{1, \dots, N + 1\}$. We can analyse this case by comparing with the processes with $n = 2N$ and $n = 2N + 2$. This makes no difference to our analysis.

5.1.2 Coupling

Another simple yet useful tool to study Birth-and-Death processes is coupling. Coupling allows us to compare chains with difficult and less tractable transition probabilities with simpler ones. In particular, for the interest of this chapter, we will compare to the push and pull chains, and the biased random walk. Here we restate two coupling results, introduced in the previous chapter, that feature in our analysis.

Lemma 73. *Consider states $\{0, 1, \dots, N\}$, and two Birth-and-Death chains $(X_t)_{t \geq 0}$ and $(\tilde{X}_t)_{t \geq 0}$ with parameters $(p_i, q_i)_{i=1}^N$ and $(\tilde{p}_i, \tilde{q}_i)_{i=1}^N$, respectively ($r_i = \tilde{r}_i = 0$). For $X_0 = \tilde{X}_0$ and, $p_i \leq \tilde{p}_i$ for all i , then it is possible to couple $(X_t)_{t \geq 0}$ and $(\tilde{X}_t)_{t \geq 0}$ such that $X_t \leq \tilde{X}_t$ for all $t \geq 0$. Moreover, the expected hitting time of state $m \in \{0, 1, \dots, N\}$ in chain $(\tilde{X}_t)_{t \geq 0}$ is smaller than in the chain $(X_t)_{t \geq 0}$.*

Proof. Consider the following construction. If $X_t \neq \tilde{X}_t$, then the two chains choose their next state independently with their corresponding probabilities. If $X_t = \tilde{X}_t = i$, we sample a uniform (in $[0, 1]$) random variable U (independent of everything), and set X_{t+1} and $\tilde{X}_{t+1}$ as follows.

$$X_{t+1} = \begin{cases} i + 1 & U \leq p_i, \\ i - 1 & \text{otherwise,} \end{cases} \quad \tilde{X}_{t+1} = \begin{cases} i + 1 & U \leq \tilde{p}_i, \\ i - 1 & \text{otherwise.} \end{cases}$$

Since $p_i \leq \tilde{p}_i$, it is clear that if $X_t = \tilde{X}_t = i$, then $X_{t+1} \leq \tilde{X}_{t+1}$. This prevents the chain $(X_t)_{t \geq 0}$ from overtaking the chain $(\tilde{X}_t)_{t \geq 0}$, when they are in the same position. Note that as the chains only increases or decreases by 1, a necessary condition for one chain to overtake the other, is that they are in the same position. Finally, as both start from the same position, it holds that $X_t \leq \tilde{X}_t$ for all $t \geq 0$. From here, we deduce that the expected hitting time of state m is smaller in $(\tilde{X}_t)_{t \geq 0}$ than in $(X_t)_{t \geq 0}$. $\square$

Let $M \in \{0, 1, \dots, n/2\}$ be the starting position of the β -Push-Pull chain $(Z_t)_{t \geq 0}$. Moreover, let $T_N = T_N(\beta)$ be the time taken to reach position N when the parameter of the chain is β , and let $\mathbb{E}_M T_N$ be the expectation of T_N starting from state M . An immediate corollary of Lemma 73 is the following.

Lemma 74. *Let $0 \leq \beta' \leq \beta \leq 1$, and $N \geq M$. Then $\mathbb{E}_M T_N(\beta') \leq \mathbb{E}_M T_N(\beta)$.*

Proof. Note that a chain with parameter β has transition probabilities

$$p_i(\beta) = 1/2 + (1 - 2\beta)i/n,$$

which is smaller than

$$p_i(\beta') = 1/2 + (1 - 2\beta')i/n.$$

From Lemma 73, the result holds. $\square$

5.2 Analysis of the Case $\beta \leq 1/2$

5.2.1 A General $n \log n$ Estimate

We develop a general recipe to obtain a $\Theta(n \log n)$ estimate of the time to reach consensus. First, we show the upper bound.

Theorem 75. *Consider a Birth-and-Death process over $\{0, \dots, n\}$ with $p_0 = q_n = 1$. If the following conditions hold:*

1. *For all $k \in \{1, \dots, n-2\}$, $\frac{p_k}{q_k} \leq \frac{p_{k+1}}{q_{k+1}}$.*

2. There exists a constant $C_1 > 0$ such that for all $k \in \{1, \dots, n-1\}$:

$$\frac{1}{p_k - q_k} \leq C_1 \frac{n}{k}.$$

3. For all $k \in \{1, \dots, n-2\}$ there exists a constant $C_2 > 0$ such that

$$\frac{p_{k+1}}{p_k} \leq C_2.$$

Then,

$$\mathbb{E}_0(T_n) \leq C_1 C_2 n \log n + \mathcal{O}(n).$$

Proof. To begin with, by changing order of summation in (5.10) it follows that

$$E_0 T_n = \sum_{i=1}^n \frac{1}{p_k} \sum_{k=0}^{i-1} \prod_{j=k+1}^{i-1} \frac{q_j}{p_j} = \sum_{k=0}^{n-1} \frac{1}{p_k} \sum_{i=k+1}^n \frac{q_{k+1} \cdots q_{i-1}}{p_{k+1} \cdots p_{i-1}}. \quad (5.14)$$

By the first condition, the ratios q_j/p_j are decreasing (as function of j), and the second condition implies $\frac{q_{k+1}}{p_{k+1}} < 1$. Thus,

$$\mathbb{E}_0 T_n \leq \sum_{k=0}^{n-1} \frac{1}{p_k} \sum_{m=0}^{\infty} \left(\frac{q_{k+1}}{p_{k+1}} \right)^m.$$

Working on the sum above,

$$\sum_{k=0}^{n-1} \frac{1}{p_k} \sum_{m=0}^{\infty} \left(\frac{q_{k+1}}{p_{k+1}} \right)^m = \sum_{k=0}^{n-1} \frac{1}{p_k} \frac{1}{1 - \frac{q_{k+1}}{p_{k+1}}} = \sum_{k=0}^{n-1} \frac{p_{k+1}}{p_k} \frac{1}{p_{k+1} - q_{k+1}}.$$

Finally, by using the upper bounds given in the second and third conditions we obtain the claimed result. $\square$

Theorem 76. Consider a birth-and-death process over $\{0, \dots, n\}$. If the following conditions hold:

1. For all $k \geq 1$, $p_k \leq p_{k+1}$ and $q_k \geq q_{k+1}$.

2. There exists a constant C_1 such that for all $k \in \{1, \dots, n\}$,

$$\frac{1}{p_k - q_k} \geq C_1 \frac{n}{k}.$$

3. There exists a constant C_2 such that for all $k \in \{1, \dots, n\}$,

$$\frac{q_k}{p_k} \leq \left(1 - \frac{C_2 k}{n}\right),$$

then,

$$\mathbb{E}_0(T_n) \geq \left(\frac{C_1}{2}\right) n \log \left(\frac{n}{2}\right) + \mathcal{O}(n).$$

Proof. Using (5.14) we have

$$E_0 T_n = \sum_{i=1}^n \frac{1}{p_k} \sum_{k=0}^{i-1} \prod_{j=k+1}^{i-1} \frac{q_j}{p_j} \geq \sum_{i=1}^n \frac{1}{p_i} \sum_{k=0}^{i-1} \left(\frac{q_i}{p_i}\right)^{i-k-1} = \sum_{i=1}^n \frac{1}{p_i - q_i} \left(1 - \left(\frac{q_i}{p_i}\right)^i\right).$$

Thus

$$\mathbb{E}_0 T_n \geq \sum_{i=\lfloor \sqrt{n} \rfloor}^n C_1 \frac{n}{i} \left(1 - \left(1 - \frac{C_2 i}{n}\right)^i\right) \geq \sum_{i=\lfloor \sqrt{n} \rfloor}^n C_1 \frac{n}{i} - \sum_{i=\lfloor \sqrt{n} \rfloor}^n C_1 \frac{n}{i} \exp \left\{-\frac{C_2 i^2}{n}\right\}. \quad (5.15)$$

In the above, the first term is give by $C_1 n(H(n) - H(\lfloor \sqrt{n} \rfloor - 1))$ with $H(m)$ the m -th harmonic number. For the second term, we use the following bound.

$$\begin{aligned} \sum_{i=\lfloor \sqrt{n} \rfloor}^n C_1 \frac{n}{i} \exp \left\{-\frac{C_2 i^2}{n}\right\} &\leq C_1 \sqrt{n} \sum_{i=\lfloor \sqrt{n} \rfloor}^n \exp \left\{-\frac{C_2 i^2}{n}\right\} \\ &\leq C_1 \sqrt{n} \int_0^n \exp \left\{-\frac{C_2 x^2}{n}\right\} dx \leq n \frac{C_1 \sqrt{\pi}}{2\sqrt{C_2}}. \end{aligned}$$

Our final estimate becomes

$$\mathbb{E}_0 T_n \geq C_1 n(H(n) - H(\lfloor \sqrt{n} \rfloor) - 1) - n \frac{C_1 \sqrt{\pi}}{2\sqrt{C_2}} \sim \frac{C_1}{2} n \log n + \mathcal{O}(n).$$

□

5.2.2 Case $\beta \in (0, 1/2)$: Proof of Theorem 69

Proof of Theorem 69.1

The proof of the Item 1 of Theorem 69 is just an application of the result obtained in Section 5.2.1. We apply the theorems above to the Markov chain $(Z_t)_{t \geq 0}$ whose transition probabilities are given by Equation (5.11).

In particular, we apply Theorems 75 and 76. For the application of those theorems, we have to be careful with the endpoints of the chains involved. The Markov chain $(Z_t)_{t \geq 0}$ moves from 0 to $n/2$ while the chains of Theorems 75 and 76 move from 0 to n .

Consider the β -Push-Pull model with fixed $\beta \in [0, \frac{1}{2})$. Note that when $\beta = 0$, we recover the Push model, and hence, it is straightforward to verify the conditions of Theorem 75. For a general $\beta \in [0, 1/2)$, observe that p_k/q_k is an increasing function on k for $k \in \{1, \dots, n/2 - 2\}$, which is the first condition. Also note that

$$\frac{1}{p_k - q_k} = 2(1 - 2\beta) \frac{n}{k},$$

which enables us to take $C_1 = 2(1 - 2\beta)$ for the second condition. Finally $p_{k+1}/p_k = 1 + \mathcal{O}(1/n)$ for $k \in \{1, \dots, n - 2\}$. Those three conditions give us that β -Push-Pull model satisfies

$$\begin{aligned} \mathbb{E}_0 T_n / 2 &\leq \frac{1}{2(1 - 2\beta)} (1 + \mathcal{O}(1/n)) \frac{n}{2} \log(n/2) + \mathcal{O}(n) \\ &= \frac{1}{4(1 - 2\beta)} (1 + \mathcal{O}(1/n)) n \log(n/2) + \mathcal{O}(n). \end{aligned}$$

As for the lower bound of $\mathbb{E}_0(T_n)/2$, we use Theorem 76. The first condition is true, for the second condition we use $C_1 = 2(1 - 2\beta)$. The last condition can be checked with $C_2 = 4(1 - 2\beta)$. Then, we get

$$\mathbb{E}_0 T_n / 2 \geq \frac{1}{8(1 - 2\beta)} n \log(n/2) + \mathcal{O}(n).$$

Note the two inequalities above give us very good estimates. Indeed, the lower and the upper bound are equal up to a multiplicative constant of 2.

5.2.3 Proof of Theorem 69.2 and 69.3

We start by saying that the case $\beta = 1/2$ makes the Birth-and-Death chain $(Z_t)_{t \geq 0}$ (with transition probabilities given by Equation (5.11)) a simple random walk in a line, thus the time to reach consensus is $\Theta(n^2)$.

Consider $\varepsilon = \varepsilon_n \rightarrow 0$, $\varepsilon > 0$ and choose $\beta = 1/2 - \varepsilon$. We can assume that $\beta > 0$ for every n . Define $\delta = \frac{1}{\varepsilon}$ and assume that $\delta < n/2$. To simplify notation, we define $N = n/2$.

Theorem 77. *Let $\varepsilon = \varepsilon_n$, $\varepsilon > 0$ and $\varepsilon \rightarrow 0$, and $\delta = \delta_n$ be two constants such that $\varepsilon = \frac{1}{\delta}$ and $\delta < N/2$. Then, for large n , β -Push-Pull model with $\beta = \frac{1}{2} - \varepsilon$ we have*

$$\mathbb{E}_0 T_{n/2} \leq 2 \exp(8\varepsilon) \frac{n}{\varepsilon} (\log(n\varepsilon)) + \mathcal{O}(n\delta),$$

and there exists a constant $K = K_\varepsilon = \Theta(1)$ such that

$$\mathbb{E}_0 T_{n/2} \geq K \frac{n}{\varepsilon} \log\left(\frac{n}{2}\varepsilon\right).$$

The proof of Theorem 69.2 and 69.3 is a direct consequence of Theorem 77. Indeed, for $\mathcal{O}(1/n) = \varepsilon > 2/N = 4/n$, we obtain a $\Theta(n^2)$ estimate directly from Theorem 77. For $4/n \geq \varepsilon \geq 0$, we can couple the process with parameter $\beta = 1/2 + \varepsilon$ between two processes, one process with parameter $\beta = 1/2 + 4/n$, and other with $\beta = 1/2$. In both cases the process take $\Theta(n^2)$ steps, in expectation, to reach state N .

We continue by proving Theorem 77.

Proof of Theorem 77.

For this proof, we define $\theta = 4\varepsilon/n$. We proceed to find estimates for Equation (5.14). Note that $p_i = 1/2 + 2\varepsilon i/n$, and $q_i = 1 - p_i$, hence

$$\frac{q_i}{p_i} = \exp\left(\log\left(1 - \frac{4\varepsilon i}{n}\right) - \log\left(1 + \frac{4\varepsilon i}{n}\right)\right). \quad (5.16)$$

Consider $|x| < 1$. Then it holds that $\log(1+x) = \sum_{k \geq 1} (-1)^{k+1} \frac{x^k}{k}$ for $|x| < 1$. Since i is at most $\frac{n}{2}$ and $\varepsilon < 1/2$ we can apply the logarithm expansion on Equation (5.16), hence

$$\frac{q_i}{p_i} = \exp\left(-2 \sum_{k \geq 1} \frac{(\theta i)^{2k-1}}{2k-1}\right), \quad (5.17)$$

where $\theta = \frac{4\varepsilon}{n}$. Note that $\theta i \leq 2\varepsilon$. We proceed to estimate the sum in Equation (5.17). Write

$$\begin{aligned} -2 \sum_{k \geq 1} \frac{(\theta i)^{2k-1}}{2k-1} &= -2\theta i \left(1 + \frac{(\theta i)^2}{3} + \frac{(\theta i)^4}{5} + \frac{(\theta i)^6}{7} + \dots \right) \\ &\geq -2\theta i \left(1 + \frac{(2\varepsilon)^2}{3} + \frac{(2\varepsilon)^4}{5} + \frac{(2\varepsilon)^6}{7} + \dots \right) \\ &\geq -\frac{2\theta i}{1-4\varepsilon^2}. \end{aligned}$$

Noticing that the above sum is clearly upper bounded by $-2\theta i$, we conclude that for every $\varepsilon < 1/2$,

$$\frac{q_i}{p_i} \in \left[\exp\left(-\frac{2\theta i}{1-4\varepsilon^2}\right), \exp(-2\theta i) \right]. \quad (5.18)$$

Upper Bound. We begin by finding a good upper bound for Equation (5.14). We shall start by substituting Equation (5.18) into Equation (5.14). This, along with the fact that $\frac{1}{p_i} \leq 2$, gives us

$$\begin{aligned} E_0 T_N &\leq 2 \sum_{k=0}^{N-1} \sum_{i=k+1}^N \exp\left(-2\theta \sum_{j=k+1}^{i-1} j\right) \\ &= 2 \sum_{k=0}^{N-1} \sum_{i=k+1}^N \exp(-\theta(i^2 - k^2 - k - i)) \\ &= 2 \sum_{k=0}^{N-1} \sum_{i=k+1}^N \exp(-\theta(i^2 - k^2)) \exp(\theta(i + k)) \\ &\leq 2 \exp(4\varepsilon) \sum_{k=0}^{N-1} \sum_{i=k+1}^N \exp(-\theta(i^2 - k^2)). \end{aligned} \quad (5.19)$$

As $2 \exp(4\varepsilon) = \Theta(1)$, we just need to estimate the double sum of equation (5.19). To deal with the sum of Equation (5.19), we make the substitution $l = i - k$ to obtain

$$\begin{aligned} \sum_{k=0}^{N-1} \sum_{i=k+1}^N \exp(-\theta(i^2 - k^2)) &= \sum_{k=0}^{N-1} \sum_{l=1}^{N-k} \exp(-\theta(l^2 + 2lk)) \\ &= \sum_{l=1}^N \exp(-\theta l^2) \sum_{k=0}^{N-l} \exp(-2\theta lk). \end{aligned} \quad (5.20)$$

Using an integral approximation, we obtain

$$\begin{aligned}
\sum_{l=1}^N \exp(-\theta l^2) \sum_{k=0}^{N-l} \exp(-2\theta l k) &= \sum_{l=1}^N \exp(-\theta l^2) \int_0^{N-l} \exp(-2\theta l k) dk + \mathcal{O}(N) \\
&= \mathcal{O}(n) + \frac{1}{2\theta} \sum_{l=1}^N \exp(-\theta l^2) \frac{1}{l} (1 - \exp(-2\theta l(N-l))) \\
&\leq \mathcal{O}(n) + \frac{1}{2\theta} \sum_{l=1}^N \exp(-\theta l^2) \frac{1}{l} (1 - \exp(-4\varepsilon l)).
\end{aligned}$$

We separate the sum of Equation (5.20) in three parts; $1 \leq l \leq \lfloor \delta \rfloor$, $\lfloor \delta \rfloor < l \leq \lfloor \sqrt{N\delta} \rfloor$, and $\lfloor \sqrt{N\delta} \rfloor < l \leq N$. The function $f(x) = (1 - e^{-4\varepsilon x})/x$ is monotone decreasing for $x \geq 0$. By a series expansion, we have $\lim_{x \rightarrow 0} f(x) = 4\varepsilon$. Hence.

$$\sum_{l=1}^{\lfloor \delta \rfloor} f(l) \leq 4\varepsilon \delta = 4, \quad \sum_{l=\delta+1}^{\lfloor \sqrt{N\delta} \rfloor} f(l) \leq \sum_{l=\delta+1}^{\lfloor \sqrt{N\delta} \rfloor} \frac{1}{l} \leq \log \sqrt{N/\delta} + O(1).$$

Therefore, when $1 \leq l \leq \sqrt{N\delta}$, as $N = n/2$, $\delta = 1/\varepsilon$ and $\theta = 4\varepsilon/n$, Equation (5.20) becomes

$$\sum_{l=1}^{\sqrt{N\delta}} \exp(-\theta l^2) \sum_{k=0}^{N-l} \exp(-2\theta l k) \leq \mathcal{O}(n) + \mathcal{O}\left(\frac{4 + \log \sqrt{n\varepsilon}}{\theta}\right) = \mathcal{O}\left(\frac{n}{\varepsilon} \log(n\varepsilon)\right). \quad (5.21)$$

The last part is to sum from $\lfloor \sqrt{N\delta} \rfloor + 1$ to N . As $\varepsilon \geq 4/n$, then $l \geq \lceil \sqrt{N\delta} \rceil \geq 2$, and $\theta l^2 \geq 2$

$$\begin{aligned}
\sum_{l=\lceil \sqrt{N\delta} \rceil}^N \exp(-\theta l^2) \sum_{k=0}^{N-l} \exp(-2\theta l k) &= \mathcal{O}(n) + \frac{1}{2\theta} \sum_{l=\lceil \sqrt{N\delta} \rceil}^N \frac{\exp(-\theta l^2)}{l} \\
&= \mathcal{O}(n) + \mathcal{O}\left(\frac{1}{\theta}\right) \int_2^\infty \frac{e^{-k^2/2}}{k} dk \\
&= \mathcal{O}\left(\frac{1}{\theta}\right) = \mathcal{O}(n\delta).
\end{aligned} \quad (5.22)$$

By combining the sums of equations (5.21) and (5.22), we obtain the desired upper bound.

Lower Bound. We proceed to compute a lower bound for Equation (5.14). Let $\theta' = \theta/(1 - 4\varepsilon^2)$. Similarly, as in the upper bound, substitute Equation (5.18) into Equation (5.14). Followed by using $\frac{1}{p_i} \geq 1$, we perform the substitution $l = i - k$ to obtain

$$\begin{aligned}
E_0 T_N &\geq \sum_{k=0}^{N-1} \sum_{i=k+1}^N \exp \left(-2\theta' \sum_{j=k+1}^{i-1} j \right) \\
&\geq \sum_{k=0}^{N-1} \sum_{i=k+1}^N \exp \left(-\theta' (i^2 - k^2) \right) \\
&= \sum_{l=1}^N \exp \left(-\theta' l^2 \right) \sum_{k=0}^{N-l} \exp \left(-2\theta' lk \right). \tag{5.23}
\end{aligned}$$

Consider the sum of Equation (5.23) when l goes from $\lceil \delta \rceil$ to $\lfloor \sqrt{n\delta}/2 \rfloor$. For l in such a regime, note that $\theta' l^2 \leq \frac{\varepsilon \delta n}{n(1-4\varepsilon)} = \frac{1}{1-4\varepsilon} \leq 5$, if we assume $\varepsilon \leq 1/5$ (which is true for sufficiently large n). Thus, a lower bound is given by

$$\begin{aligned}
\sum_{l=1}^N \exp \left(-\theta' l^2 \right) \sum_{k=0}^{N-l} \exp \left(-2\theta' lk \right) &\geq \sum_{l=\lceil \delta \rceil}^{\lfloor \sqrt{n\delta} \rfloor} \exp \left(-\theta' l^2 \right) \sum_{k=0}^{N-l} \exp \left(-2\theta' lk \right) \\
&\geq \sum_{l=\lceil \delta \rceil}^{\lfloor \sqrt{n\delta} \rfloor} e^{-5} \int_{k=0}^{N-l} \exp \left(-2\theta' lk \right) \\
&\geq \frac{e^{-5}}{2\theta'} \sum_{l=\lceil \delta \rceil}^{\lfloor \sqrt{n\delta} \rfloor} \frac{1}{l} (1 - \exp \left(-2\theta' l(N-l) \right)) \\
&\geq \frac{e^{-5}}{2\theta'} (1 - e^{-2}) \sum_{l=\lceil \delta \rceil}^{\lfloor \sqrt{n\delta} \rfloor} \frac{1}{l} \tag{5.24} \\
&= \frac{e^{-5}}{2\theta'} (1 - e^{-2}) (\log(\sqrt{n/\delta}) + \Theta(1)).
\end{aligned}$$

Then, there exists a bounded (away from zero) value $K = K_\varepsilon > 0$, such that

$$\frac{e^{-5}}{2\theta'} (1 - e^{-2}) (\log(\sqrt{n/\delta}) + \Theta(1)) \geq K \log(\sqrt{n/\delta}).$$

Therefore, by using that $\theta = 4\varepsilon/n$, from the last equation we deduce that

$$\mathbb{E}_0 T_N \geq \frac{K}{\theta} \log \sqrt{N/\delta} = \Theta \left(\frac{n}{\varepsilon} \log(n\varepsilon) \right).$$

The explanation of Equation (5.24) is due to the following argument. Recall that $\delta \leq N/2$ and $N = n/2$, then we have that

$$\left\lfloor \sqrt{n\delta}/2 \right\rfloor \leq \sqrt{2N\delta}/2 \leq \sqrt{N^2}/2 \leq \lceil N/2 \rceil. \quad (5.25)$$

Note that the function $l(N-l)$ is non-decreasing for $l \in \{1, \dots, \lceil N/2 \rceil\}$. Then, from the fact that $\theta' = \theta/(1-4\varepsilon) \geq \theta = 4\varepsilon/n$, for all $l \in \left\{ \lceil \delta \rceil, \dots, \left\lfloor \sqrt{n\delta}/2 \right\rfloor \right\}$, it holds that

$$2\theta' l(N-l) \geq 2\theta' \lceil \delta \rceil (N - \lceil \delta \rceil) \geq 8 \frac{\varepsilon}{n} \delta (N - \delta) \geq 8 \frac{(N - N/2)}{n} \geq 2.$$

We conclude that $\exp(-2\theta' l(N-l)) \leq \exp(-2)$, which explains Equation (5.24). □

5.3 Analysis of the Case $\beta \in (1/2, 1)$

We start by providing the necessary tools. In this section, we assume that n is divisible by 4 (other cases can be bounded by this). Let $N = n/2$, and define the commute time $C[0, N] = \mathbb{E}_0 T_N + \mathbb{E}_N T_0$. Since $\beta \in (1/2, 1)$, we have that the bias of the random walk is toward 0, then $\mathbb{E}_N T_0 \leq \mathbb{E}_0 T_N$. Therefore, we get

$$\mathbb{E}_0 T_N \leq C[0, N] \leq 2\mathbb{E}_0 T_N. \quad (5.26)$$

Hence, $C[0, N]$ is a good estimate for $\mathbb{E}_0 T_N$. To compute $C[0, N]$, we use the following lemma, whose proof is standard and can be found in [3].

Lemma 78. *Let a, b be two states of an ergodic Markov chain. Let T_a be the hitting time of state a , and let T_b^+ be the first return time to state b , then*

$$C[a, b] = \frac{1}{\pi_b \mathbb{P}_b(T_a < T_b^+)}.$$

We obtain a way to compute $C[0, N]$ by applying Lemma 78. We need to estimate $\pi(N)$ and $\mathbb{P}_N(T_0 < T_N^+)$.

5.3.1 Estimation of $\mathbb{P}_N(T_0 < T_N^+)$

Define $(X_t)_{t \geq 0}$ as a biased random walk with parameter p ($p \in (0, 1)$). Recall that a biased random walk with parameter p is a Birth-and-Death process on $\{0, \dots, N\}$ with reflecting barriers at 0 and N , and such that for any $i \in \{1, \dots, N-1\}$, we have $\mathbb{P}(X_{t+1} = i+1 | X_t = i) = p$ and $\mathbb{P}(X_{t+1} = i-1 | X_t = i) = q = 1-p$. We define $g_N(p)$ the probability that starting from N , the Markov chain $(X_t)_{t \geq 0}$ reaches state 0 before returning to N . From [41, Chapter XIV], we get the following result.

Lemma 79. *For $p \in (0, 1)$ and $p \neq 1/2$, we have $g_N(p) = \frac{1-(p/q)}{1-(p/q)^N} \geq q - p$.*

By using the Lemma above and a coupling argument we have the following result.

Lemma 80. *For $\beta \in (1/2, 1)$ we have*

$$g_N(1-\beta) \geq \mathbb{P}_N(T_0 < T_N^+) \geq \frac{1}{2} g_{N/2}(3/4 - \beta/2).$$

Proof. For the upper bound, consider a biased random walk $(X_t)_{t \geq 0}$ with parameter $p = 1-\beta$ on state space $\{0, \dots, N\}$. Let $\beta = 1/2 + \varepsilon$. Then $p_i = 1/2 + (1-2\beta)i/n = 1/2 - \varepsilon(2i/n)$. As $1-\beta = 1/2 - \varepsilon$, we have $p_i > 1-\beta$ for all $i \in \{1, \dots, N-1\}$. Thus, we can couple $(X_t)_{t \geq 0}$ with $(Z_t)_{t \geq 0}$ such that $\mathbb{P}(X_t \leq Z_t, \forall t \geq 0 | X_0 = Z_0 = N) = 1$ (see Lemma 73). In particular, we get that $g_N(1-\beta) \geq \mathbb{P}_N(T_0 < T_N^+)$.

For the lower bound we choose $M = N/2$. Consider the stopping time $T = \min(T_M, T_N^+)$, then by the Strong Markov property (Theorem 87 in Appendix A), we have that

$$\begin{aligned} \mathbb{P}_N(T_0 < T_N^+) &= \mathbb{E}(\mathbb{P}_N(T_0 < T_N^+ | X_T)) \\ &= \mathbb{P}(T_0 < T_N^+ | X_T = M) \mathbb{P}_N(X_T = M) \\ &= \mathbb{P}_M(T_0 < T_N^+) \mathbb{P}_N(T_M < T_N^+). \end{aligned} \tag{5.27}$$

To estimate $\mathbb{P}_N(T_M < T_N^+)$, we restrict our chains to state space $\{M, \dots, N\}$, with a new reflecting barrier at M . In the restricted chain, we have $p_i \in [1/2 - \varepsilon, 1/2 - \varepsilon/2]$, for all states $i \in \{M+1, \dots, N-1\}$. Thus $p_i \leq 1/2 - \varepsilon/2 = 3/4 - \beta/2$. Applying the same coupling argument as before, we get $g_{N/2}(3/4 - \beta/2) \leq \mathbb{P}_N(T_M < T_N^+)$.

Then, the lower bound follows from a coupling argument between $(Z_t)_{t \geq 0}$ on $\{M, \dots, N\}$ and a biased random walk with $p = 3/4 - \beta/2$ on state space $\{M, \dots, N\}$, along with the observation that, since the bias of the chain $(Z_t)_{t \geq 0}$ is toward 0, we have $\mathbb{P}_M(T_0 < T_N) \geq 1/2$. $\square$

By using Lemma 79 together with Lemma 80 we obtain the following Lemma.

Lemma 81. *For any $\beta \in (1/2, 1)$ that may depend on n , we have*

$$\frac{1}{1 - ((1 - \beta)/\beta)^N} \frac{2\beta - 1}{\beta} \geq \mathbb{P}_N(T_0 < T_N^+) \geq \frac{2\beta - 1}{4}. \quad (5.28)$$

In particular, if $\varepsilon \geq c/n$ then

$$\mathbb{P}_N(T_0 < T_N^+) = \Theta(\varepsilon). \quad (5.29)$$

Proof. The bounds in (5.28) are from Lemma 79. The upper bound of (5.29) comes from

$$\left(\frac{1 - \beta}{\beta}\right)^N = \left(\frac{1 - 2\varepsilon}{1 + 2\varepsilon}\right)^N \leq \exp\left(-\frac{4\varepsilon N}{1 + 2\varepsilon}\right) \leq \exp(-2\varepsilon N) \leq e^{-c}.$$

It follows that the first term on the left hand side of (5.28) is at most $1/(1 - e^{-c})$, which is a bounded quantity for any constant $c > 0$ or for c tending to infinity. $\square$

Lemma 82. *For $\varepsilon n = \Omega(1)$ and $\varepsilon \leq 1/2$, it holds*

$$\pi(0) = \Theta\left(\sqrt{\frac{\varepsilon}{n}}\right). \quad (5.30)$$

Proof. Consider that $p_i = 1/2 - 2\varepsilon i/n$ and $q_i = 1 - p_i$. Note that for i constant, $p_i, q_i \sim 1/2$. Thus, Equation (5.7) combined with $p_0 = q_N = 1$, tells us that

$$\pi(0)^{-1} = 1 + \frac{p_0}{q_1} + \sum_{j=2}^N \frac{p_0 \cdots p_{j-1}}{q_1 \cdots q_j} = \Theta\left(\sum_{j=1}^{N-1} \frac{p_1 \cdots p_j}{q_1 \cdots q_j}\right). \quad (5.31)$$

From $p_i = 1/2 - 2\varepsilon i/n$ and $q_i = 1/2 + 2\varepsilon i/n$, we get $p_i/q_i = 1 - \frac{8\varepsilon i/n}{1 + 4\varepsilon i/n}$. Thus, as $i \leq N = n/2$, and $0 < \varepsilon \leq 1/2$, we obtain

$$1 - \frac{8\varepsilon i}{n} \leq \frac{p_i}{q_i} = 1 - \frac{8\varepsilon i}{n + 4\varepsilon i} \leq 1 - \frac{4\varepsilon i}{n}.$$

For $|x| < 1$, $\exp(-x/(1-x)) \leq 1-x \leq \exp(-x)$, so

$$\exp\left(-\frac{8\varepsilon i}{n-8\varepsilon i}\right) \leq \frac{p_i}{q_i} \leq \exp\left(-\frac{4\varepsilon i}{n}\right).$$

Provided $i \leq N/4$, we have $p_i/q_i \geq \exp(-16\varepsilon i/n)$, and thus

$$\sum_{j=1}^{N/4} \prod_{i=1}^j \exp\left(-\frac{16\varepsilon i}{n}\right) \leq \pi(0)^{-1} \leq 2 \sum_{j=1}^{N/2} \prod_{i=1}^j \exp\left(-\frac{4\varepsilon i}{n}\right), \quad (5.32)$$

or equivalently

$$\pi(0)^{-1} = \Theta(1) \sum_{j=1}^{N/2} \exp\left(-\Theta(\varepsilon) \frac{j^2}{n}\right). \quad (5.33)$$

To finish the proof, we need to check that the sum above is $\Theta(1)\sqrt{n/\Theta(\varepsilon)}$

$$\begin{aligned} \sum_{j=1}^{N/2} \exp\left(-\Theta(\varepsilon) \frac{j^2}{n}\right) &= \Theta(1) \int_0^{N/2} \exp\left(-\Theta(\varepsilon) \frac{x^2}{n}\right) dx \\ &= \Theta(1) \sqrt{\frac{n}{\Theta(\varepsilon)}} \int_0^{\Theta(\sqrt{\varepsilon n})} \exp(-x^2) dx \\ &= \Theta(1) \sqrt{\frac{n}{\Theta(\varepsilon)}} \Phi(\Theta(\sqrt{\varepsilon n})). \end{aligned}$$

Observe that $\Phi(x)$ goes to $1/2$ as x approaches infinity, and it goes to 0 when x approaches 0. Hence, if $\sqrt{\varepsilon n} = \Omega(1)$ then $\Phi(\Theta(\sqrt{\varepsilon n})) = \Theta(1)$.

□

5.3.2 Estimation of $\pi(N)$

We express $\pi(N)$ in terms of $\pi(0)$. By Equation (5.6) we have that

$$\pi(N) = \pi(0) \frac{p_0 \cdots p_{N-1}}{q_1 \cdots q_{N-1}} = \pi(0) R.$$

Note that $p_0 = 1$ and using $N = n/2$, then $\beta = 1 - \delta$

$$p_i = \frac{1}{2} + (1 - 2\beta) \frac{i}{n} = \frac{1}{2N} (N - (1 - 2\delta)i).$$

In this way we can write R as

$$R = \prod_{i=1}^{N-1} \frac{N - (1 - 2\delta)i}{N + (1 - 2\delta)i}. \quad (5.34)$$

Lemma 83.

i. Let $c > 1$ be a positive constant. If $1/(cn) \leq \delta < 1/2$, then

$$R = \Theta(1) \frac{1}{\sqrt{\delta}} \left(\frac{1}{2(1-\delta)^{1-\delta}\delta^\delta} \right)^{n/(1-2\delta)}. \quad (5.35)$$

ii. Let c be a positive constant. If $\delta \leq 1/(cn)$, then

$$R = \Theta(1) \frac{\sqrt{n}}{2^n} \exp(\delta n \log n). \quad (5.36)$$

Proof. We proof the two equations above separately.

Proof of Equation (5.35). Let $\alpha = 1/(1-2\delta)$, then from equation (5.34) we can write R as

$$\begin{aligned} R &= \prod_{k=1}^{N-1} \frac{\alpha N - k}{\alpha N + k} = \frac{(\alpha N - 1) \cdots (\alpha N - (N - 1))}{(\alpha N + 1) \cdots (\alpha N + (N - 1))} \\ &= \frac{\alpha}{\alpha - 1} \frac{\Gamma(\alpha N) \Gamma(\alpha N)}{\Gamma((\alpha + 1)N) \Gamma((\alpha - 1)N)}, \end{aligned}$$

where $\Gamma(z) = (z-1)\Gamma(z-1)$. Provided z is at least a small positive constant, we can perform an asymptotic expansion of the Gamma function [74] as given by

$$\Gamma(z) = \sqrt{2\pi} z^{z-1/2} e^{-z} \left(1 + \mathcal{O}\left(\frac{1}{z}\right) \right).$$

Thus, simplifying extensively gives

$$\begin{aligned} R &= \Theta(1) \sqrt{\frac{\alpha+1}{\alpha-1}} \left(\frac{\alpha^{2\alpha}}{(\alpha+1)^{\alpha+1}(\alpha-1)^{\alpha-1}} \right)^N \\ &= \Theta(1) \sqrt{\frac{\alpha+1}{\alpha-1}} \left(\left(\frac{\alpha}{\alpha+1} \right)^{\alpha+1} \left(\frac{\alpha}{\alpha-1} \right)^{\alpha-1} \right)^N \\ &= \Theta(1) \sqrt{\frac{1-\delta}{\delta}} \left(\left(\frac{1}{2-2\delta} \right)^{1-\delta} \left(\frac{1}{2\delta} \right)^\delta \right)^{2N/(1-2\delta)}. \end{aligned}$$

Equation (5.35) follows directly from this.

Proof of Equation (5.36). We use the expression for R given in (5.34). This is equivalent to

$$R = \prod_{k=1}^{N-1} \frac{N-k}{N+k} \prod_{k=1}^{N-1} \frac{1+2\delta k/(N-k)}{1-2\delta k/(N+k)}. \quad (5.37)$$

The first product can be written as

$$\prod_{k=1}^{N-1} \frac{N-k}{N+k} = 2 \frac{N!N!}{(2N)!} = \Theta(1) \frac{\sqrt{n}}{2^n}.$$

If $\delta < 1/cn$, then the denominator in the second product of (5.37) is $1 - \mathcal{O}(\delta)$. For the numerator, we have $2\delta k/(N-k) < 1$, so

$$1 + 2\delta k/(N-k) = \exp \left(2\delta \frac{k}{N-k} - O \left(\frac{\delta k}{N-k} \right)^2 \right).$$

Here

$$\sum_{k=1}^{N-1} \frac{k}{N-k} = \mathcal{O}(N) + N \log N,$$

and

$$\delta^2 \sum_{k=1}^{N-1} \left(\frac{k}{N-k} \right)^2 = \mathcal{O}(\delta^2 N^2) = \mathcal{O}(1).$$

Thus

$$\prod_{k=1}^{N-1} \frac{1 + 2\delta k/(N-k)}{1 - 2\delta k/(N+k)} = \Theta(1 + \delta) \exp(\delta n \log n),$$

completing the proof of (5.36). □

5.3.3 Proof of Theorem 70

We apply the various results we have accumulated so far to

$$\mathbb{E}_0 T_N = \Theta(1) \frac{1}{\pi(N) \mathbb{P}_N(T_0 < T_N^+)} \quad (5.38)$$

and

$$\pi(N) = \pi(0)R. \quad (5.39)$$

Here, $\mathbb{P}_N(T_0 < T_N^+)$ is given by Lemma 80, $\pi(0)$ by Lemma 82, and R by Lemma 83.

From Equation (5.35) with $\beta = 1/2 + \varepsilon = 1 - \delta$ (i.e., $\delta = 1/2 - \varepsilon$, $1 - 2\delta = 2\varepsilon$), we obtain

$$R = \Theta(1) \left((1 + 2\varepsilon)^{1+2\varepsilon} (1 - 2\varepsilon)^{1-2\varepsilon} \right)^{-n/4\varepsilon}. \quad (5.40)$$

For $0 < x \leq 1/2$,

$$(1+x)^{1+x}(1-x)^{1-x} = \exp \left(\sum_{\substack{j \geq 2 \\ j \text{ even}}} \frac{2}{(j-1)j} x^j \right). \quad (5.41)$$

Substitute $x = 2\varepsilon$ in Equation (5.41), and apply it to Equation (5.40) to obtain

$$R = \Theta(1) \exp \left(-\frac{n}{4\varepsilon} \sum_{\substack{j \geq 2 \\ j \text{ even}}} \frac{2}{(j-1)j} (2\varepsilon)^j \right). \quad (5.42)$$

From Equation (5.28), we get

$$\Theta(\varepsilon) \frac{1}{1 - e^{-\Theta(n\varepsilon)}} \geq \mathbb{P}_N(T_0 < T_N^+) \geq \Theta(\varepsilon). \quad (5.43)$$

We proceed to prove all the different cases of Theorem 70.

Theorem 70: Part 1. Case $0 \leq \varepsilon \leq c/n$.

For the lower bound, we use a coupling argument (Lemma 74). Indeed, we can couple the our chain $(Z_t)_{t \geq 0}$ of parameter $\beta = 1/2 + \varepsilon$ with a faster chain of parameter $\beta' = 1/2 < \beta$. Thus, from Theorem 69.3, $\mathbb{E}_0 T_N = \Omega(n^2)$ holds.

We next prove that $\mathbb{E}_0 T_N = \mathcal{O}(n^2)$. Suppose that $\varepsilon = c/n$ where $c > 0$ is constant. From Equation (5.42),

$$\pi(N) = \pi(0)R = \pi(0)\Theta(e^{-c}),$$

In consequence, $\pi(N) = \Theta(\pi(0))$. For $\beta \geq 1/2$, it follows from $\pi(j+1) = \pi(j)p_j/q_{j+1}$ that $\pi(j) \geq \pi(j+1)$. Thus $\pi(N) = \Theta(1/n)$ because

$$1 = \sum_{j=0}^N p_j \leq (N+1)p_N \leq C(N+1)p_0 \leq C \sum_{j=0}^N p_j = C.$$

Then, from Equation (5.38) and the right hand side of Equation (5.43)

$$\mathbb{E}_0 T_N(1/2 + c/n) = \frac{1}{\pi(N)\mathbb{P}_N(T_0 < T_N^+)} = \mathcal{O}\left(\frac{n}{\varepsilon}\right) = \mathcal{O}\left(\frac{n^2}{c}\right).$$

Hence, for any constant c , and all $\varepsilon \leq c/n$, from Lemma 74, we have that $\mathbb{E}_0 T_N(1/2 + \varepsilon) \leq \mathbb{E}_0 T_N(1/2 + c/n) = \mathcal{O}(n^2/c)$. Combining this with the result (from above) that $\mathbb{E}_0 T_N = \Omega(n^2)$ gives us that $\mathbb{E}_0 T_N = \Theta(n^2)$.

Theorem 70: Part 2. Case $c/n \leq \varepsilon = o(1)$.

We use Equation (5.38) together with (5.39), i.e.,

$$\mathbb{E}_0 T_N = \Theta(1) \frac{1}{\pi(0) R \mathbb{P}_N(T_0 < T_N^+)}.$$

From Lemma 82, we have

$$\pi(0) = \Theta(\sqrt{\varepsilon/n}).$$

Equation (5.43) gives $\mathbb{P}_N(T_0 < T_N^+) = \Theta(\varepsilon)$. Combining this with Equation (5.42) gives

$$\mathbb{E}_0 T_N = \Theta(1) \frac{\sqrt{n}}{\varepsilon^{3/2}} \exp \left(-\frac{n}{4\varepsilon} \sum_{\substack{j \geq 2 \\ j \text{ even}}} \frac{2}{(j-1)j} (2\varepsilon)^j \right) = \Theta(1) \frac{\sqrt{n}}{\varepsilon^{3/2}} \exp(n\varepsilon(1 + \mathcal{O}(\varepsilon^2))).$$

For the second equality, we used that $\varepsilon \leq 1/2$ and $\sum_{J \geq 1} 1/(J(J+1)) = 1$.

Theorem 70: Parts 3 and 4. In this regime we have that $\beta = 1/2 + \varepsilon = 1 - \delta$ with δ constant or $\delta = o(1)$ but $\delta \gg 1/n$. We use Equation (5.38) together with (5.39), i.e.,

$$\mathbb{E}_0 T_N = \Theta(1) \frac{1}{\pi(0) R \mathbb{P}_N(T_0 < T_N^+)} . \quad (5.44)$$

We just need to replace the values in the formula above. From Lemma 82, we have $\pi(0) = \Theta(n^{-1/2})$. From Equation (5.35), we get

$$R = \Theta(1) \frac{1}{\sqrt{\delta}} \left(\frac{1}{2(1-\delta)^{1-\delta} \delta^\delta} \right)^{n/(1-2\delta)},$$

and from Equation (5.43), $\mathbb{P}_N(T_0 < T_N^+) = \Theta(\varepsilon) = \Theta(1)$. Therefore

$$\mathbb{E}_0 T_N = \Theta(\sqrt{n}) \left(2(1-\delta)^{1-\delta} \delta^\delta \right)^{n/(1-2\delta)}.$$

Theorem 70: Parts 5 and 6.

In this regime we have that $\beta = 1/2 + \varepsilon = 1 - \delta$ with $\delta \ll 1/n$. Again, we use Equation (5.38) together with (5.39), i.e.,

$$\mathbb{E}_0 T_N = \Theta(1) \frac{1}{\pi(0) R \mathbb{P}_N(T_0 < T_N^+)} \quad (5.45)$$

As for parts 3 and 4, we just need to replace the values in the formula above. From Lemma 82, we get $\pi(0) = \Theta(n^{-1/2})$. From Equation (5.36),

$$R = \Theta(1) \frac{\sqrt{n}}{2^n} \exp(\delta n \log n).$$

Finally, from Equation (5.43), $\mathbb{P}_N(T_0 < T_N^+) = \Theta(\varepsilon) = \Theta(1)$. From there, we get

$$\mathbb{E}_0 T_N = \Theta(1) 2^n \exp(-\delta n \log n).$$

Note that for $\delta = \mathcal{O}(n \log n)^{-1}$, the exponential term becomes a constant, giving us the $\Theta(2^n)$ estimate.

Appendix A

Standard Probability Results

Theorem 84. Chernoff-Hoeffding bound [34, Section 1.6]

Let $X = \sum_{i=1}^n X_i$, where $(X_i)_{i=1}^n$ are independently distributed random variables in $[0, 1]$. Then, for all $\varepsilon > 0$,

$$\mathbb{P}(X > (1 + \varepsilon)\mathbb{E}(X)) \leq \exp\left(-\frac{\varepsilon^2}{3}\mathbb{E}(X)\right), \quad (\text{A.1})$$

and,

$$\mathbb{P}(X < (1 - \varepsilon)\mathbb{E}(X)) \leq \exp\left(-\frac{\varepsilon^2}{2}\mathbb{E}(X)\right). \quad (\text{A.2})$$

Theorem 85. Azuma's Inequality [34, Section 5.2]

Let $(M_k)_{k \geq 0}$ be a discrete-time martingale with respect to a filtration $(\mathcal{F}_k)_{k \geq 0}$. Suppose that for each $k > 0$ there exists a_k, b_k such that

$$a_k \leq M_k - M_{k-1} \leq b_k.$$

Then, for all $t > 0$,

$$\mathbb{P}(M_n > X_0 + t), \mathbb{P}(M_n < X_0 - t) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right) \quad (\text{A.3})$$

Theorem 86. Doob's Optional-Stopping Theorem [75, Section 10.10]

Let $(M_k)_{k \geq 0}$ be a discrete-time supermartingale with respect to a filtration $(\mathcal{F}_k)_{k \geq 0}$,

and let T be a stopping time with respect to such a filtration. Then M_T is integrable and

$$\mathbb{E}(M_T) \leq \mathbb{E}(M_0), \quad (\text{A.4})$$

if one of the following situations hold:

- T is bounded, that is, there exists a (deterministic) $n \in \mathbb{N}$, such that $T \leq n$.
- (M_k) is bounded, i.e., there exists a (deterministic) $x \in \mathbb{R}$, such that $|M_k| < x$ for all $k \geq 0$.
- $\mathbb{E}(T) < \infty$, and, for some (deterministic) $x \in \mathbb{R}$, it holds,

$$|M_{k+1} - M_k| \leq x,$$

for all $k \geq 0$.

Moreover, inequality of Equation (A.4) becomes a equality if $(M_k)_{k \geq 0}$ is a martingale.

Theorem 87. Strong Markov Property [37, Section 5.2]

Let $(X_k)_{k \geq 0}$ be a discrete-time Markov chain on finite state space S and transition matrix P . Let T be a stopping time of $(X_k)_{k \geq 0}$, such that $T < \infty$. Then, conditional on $X_T = i$, the process $(X_{T+t})_{t \geq 0}$ is independent of $X_0, \dots, X_{T-1}$, and moreover, it has the same distribution as $(X_t)_{t \geq 0}$ conditional on $X_0 = i$. In particular, for all $j \in S$, we have that

$$\mathbb{P}(X_{T+1} \in j | X_0, \dots, X_T = i) = \mathbb{P}(X_{T+1} \in j | X_T = i) = P(i, j).$$

Bibliography

- [1] Mohammed Amin Abdullah and Moez Draief. Global majority consensus by local majority polling on graphs of a given degree sequence. *Discrete Appl. Math.*, 180:1–10, 2015.
- [2] Hüseyin Acan, Andrea Collevecchio, Abbas Mehrabian, and Nick Wormald. On the push&pull protocol for rumor spreading. *SIAM J. Discrete Math.*, 31(2):647–668, 2017.
- [3] 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/~aldous/RWG/book.html>.
- [4] Noga Alon, Chen Avin, Michal Koucký, Gady Kozma, Zvi Lotker, and Mark R. Tuttle. Many random walks are faster than one. *Combin. Probab. Comput.*, 20(4):481–502, 2011.
- [5] Hamed Amini, Moez Draief, and Marc Lelarge. Flooding in weighted sparse random graphs. *SIAM J. Discrete Math.*, 27(1):1–26, 2013.
- [6] Anirban Basak, Rick Durrett, and Yuan Zhang. The evolving voter model on thick graphs. *arXiv preprint arXiv:1512.07871*, 2015.
- [7] Riddhipratim Basu and Allan Sly. Evolving voter model on dense random graphs. *Ann. Appl. Probab.*, 27(2):1235–1288, 2017.

- [8] L. Becchetti, A. Clementi, E. Natale, F. Pasquale, and R. Silvestri. Plurality consensus in the gossip model. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 371–390. SIAM, Philadelphia, PA, 2015.
- [9] L. Becchetti, A. Clementi, E. Natale, F. Pasquale, and L. Trevisan. Stabilizing consensus with many opinions. In *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 620–635. ACM, New York, 2016.
- [10] Luca Becchetti, Andrea Clementi, Emanuele Natale, Francesco Pasquale, Riccardo Silvestri, and Luca Trevisan. Simple dynamics for plurality consensus. volume 30, pages 293–306, 2017.
- [11] Petra Berenbrink, Tom Friedetzky, George Giakkoupis, and Peter Kling. Efficient plurality consensus, or: the benefits of cleaning up from time to time. In *43rd International Colloquium on Automata, Languages, and Programming*, volume 55 of *LIPICs. Leibniz Int. Proc. Inform.*, pages Art. No. 136, 14. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2016.
- [12] Petra Berenbrink, George Giakkoupis, Anne-Marie Kermarrec, and Frederik Mallmann-Trenn. Bounds on the voter model in dynamic networks. 55:Art. No. 146, 15, 2016.
- [13] Carol Bezuidenhout and Geoffrey Grimmett. The critical contact process dies out. *Ann. Probab.*, 18(4):1462–1482, 1990.
- [14] B. Bollobás and Y. Kohayakawa. On Richardson’s model on the hypercube. pages 129–137, 1997.
- [15] Christian Borgs, Jennifer Chayes, Ayalvadi Ganesh, and Amin Saberi. How to distribute antidote to control epidemics. *Random Structures Algorithms*, 37(2):204–222, 2010.

- [16] Stephen Boyd, Arpita Ghosh, Balaji Prabhakar, and Devavrat Shah. Randomized gossip algorithms. *IEEE Trans. Inform. Theory*, 52(6):2508–2530, 2006.
- [17] Andrei Z. Broder, Anna R. Karlin, Prabhakar Raghavan, and Eli Upfal. Trading space for time in undirected s - t connectivity. *SIAM J. Comput.*, 23(2):324–334, 1994.
- [18] Peter Clifford and Aidan Sudbury. A model for spatial conflict. *Biometrika*, 60:581–588, 1973.
- [19] Colin Cooper, Martin Dyer, Alan Frieze, and Nicolás Rivera. Discordant voting processes on finite graphs. In *43rd International Colloquium on Automata, Languages, and Programming*, volume 55 of *LIPICs. Leibniz Int. Proc. Inform.*, pages Art. No. 145, 13. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2016.
- [20] Colin Cooper, Robert Elsässer, Hirotaka Ono, and Tomasz Radzik. Coalescing random walks and voting on connected graphs. *SIAM J. Discrete Math.*, 27(4):1748–1758, 2013.
- [21] Colin Cooper, Robert Elsässer, and Tomasz Radzik. The power of two choices in distributed voting. In *Automata, languages, and programming. Part II*, volume 8573 of *Lecture Notes in Comput. Sci.*, pages 435–446. Springer, Heidelberg, 2014.
- [22] Colin Cooper, Robert Elsässer, Tomasz Radzik, Nicolás Rivera, and Takeharu Shiraga. Fast consensus for voting on general expander graphs. In *Distributed computing*, volume 9363 of *Lecture Notes in Comput. Sci.*, pages 248–262. Springer, Heidelberg, 2015.
- [23] Colin Cooper, Alan Frieze, and Tomasz Radzik. Multiple random walks in random regular graphs. *SIAM J. Discrete Math.*, 23(4):1738–1761, 2009/10.
- [24] Colin Cooper, Tomasz Radzik, and Nicolas Rivera. The coalescing-branching random walk on expanders and the dual epidemic process. In *Proceedings of*

- the 2016 ACM Symposium on Principles of Distributed Computing, PODC '16*, pages 461–467, New York, NY, USA, 2016. ACM.
- [25] Colin Cooper, Tomasz Radzik, and Nicolás Rivera. Improved cover time bounds for the coalescing-branching random walk on graphs. In *Proceedings of the 29th ACM Symposium on Parallelism in Algorithms and Architectures, SPAA '17*, pages 305–312, New York, NY, USA, 2017. ACM.
 - [26] Colin Cooper, Tomasz Radzik, Nicolás Rivera, and Takeharu Shiraga. Fast plurality consensus in regular expanders. In *Distributed computing*, volume 91 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 13:1–13:16, Dagstuhl, Germany, 2017. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
 - [27] Colin Cooper and Nicolás Rivera. The Linear Voting Model. In *43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016)*, volume 55 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 144:1–144:12, Dagstuhl, Germany, 2016. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
 - [28] J. T. Cox. Coalescing random walks and voter model consensus times on the torus in $\mathbf{Z}^d$. *Ann. Probab.*, 17(4):1333–1366, 1989.
 - [29] J. Theodore Cox, Yuval Peres, and Jeffrey E. Steif. Cutoff for the noisy voter model. *Ann. Appl. Probab.*, 26(2):917–932, 2016.
 - [30] Alan Demers, Dan Greene, Carl Hauser, Wes Irish, John Larson, Scott Shenker, Howard Sturgis, Dan Swinehart, and Doug Terry. Epidemic algorithms for replicated database maintenance. In *Proceedings of the Sixth Annual ACM Symposium on Principles of Distributed Computing, PODC '87*, pages 1–12, New York, NY, USA, 1987. ACM.

- [31] 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.
- [32] Benjamin Doerr, Mahmoud Fouz, and Tobias Friedrich. Social networks spread rumors in sublogarithmic time. In *STOC’11—Proceedings of the 43rd ACM Symposium on Theory of Computing*, pages 21–30. ACM, New York, 2011.
- [33] Peter Donnelly and Dominic Welsh. Finite particle systems and infection models. volume 94, pages 167–182, 1983.
- [34] Devdatt P. Dubhashi and Alessandro Panconesi. *Concentration of measure for the analysis of randomized algorithms*. Cambridge University Press, Cambridge, 2009.
- [35] F. D. J. Dunstan and D. J. A. Welsh. A greedy algorithm for solving a certain class of linear programmes. *Math. Programming*, 5:338–353, 1973.
- [36] Richard Durrett, James P. Gleeson, Alun L. Lloyd, Peter J. Mucha, Feng Shi, David Sivakoff, Joshua E. S. Socolar, and Chris Varghese. Graph fission in an evolving voter model. *Proc. Natl. Acad. Sci. USA*, 109(10):3682–3687, 2012.
- [37] Rick Durrett. *Probability: theory and examples*, volume 31 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, fourth edition, 2010.
- [38] Chinmoy Dutta, Gopal Pandurangan, Rajmohan Rajaraman, and Scott Roche. Coalescing-branching random walks on graphs. In *Proceedings of the Twenty-fifth Annual ACM Symposium on Parallelism in Algorithms and Architectures*, SPAA ’13, pages 176–185, New York, NY, USA, 2013. ACM.
- [39] Chinmoy Dutta, Gopal Pandurangan, Rajmohan Rajaraman, and Scott Roche. Coalescing-branching random walks on graphs. *ACM Trans. Parallel Comput.*, 2(3):20:1–20:29, November 2015.

- [40] Robert Elsässer and Thomas Sauerwald. Tight bounds for the cover time of multiple random walks. *Theoret. Comput. Sci.*, 412(24):2623–2641, 2011.
- [41] William Feller. *An introduction to probability theory and its applications: volume I*, volume 3. John Wiley & Sons New York, 1968.
- [42] James Allen Fill and Robin Pemantle. Percolation, first-passage percolation and covering times for Richardson’s model on the n -cube. *Ann. Appl. Probab.*, 3(2):593–629, 1993.
- [43] Nikolaos Fountoulakis, Konstantinos Panagiotou, and Thomas Sauerwald. Ultra-fast rumor spreading in social networks. In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1642–1660. ACM, New York, 2012.
- [44] Silvio Frischknecht, Barbara Keller, and Roger Wattenhofer. Convergence in (social) influence networks. In *Distributed computing*, volume 8205 of *Lecture Notes in Comput. Sci.*, pages 433–446. Springer, Heidelberg, 2013.
- [45] Ayalvadi Ganesh, Laurent Massoulié, and Don Towsley. The effect of network topology on the spread of epidemics. In *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, volume 2, pages 1455–1466. IEEE, 2005.
- [46] Mohsen Ghaffari and Merav Parter. A polylogarithmic gossip algorithm for plurality consensus. In *Proceedings of the 2016 ACM Symposium on Principles of Distributed Computing*, pages 117–126. ACM, 2016.
- [47] George Giakkoupis. Tight bounds for rumor spreading in graphs of a given conductance. In *28th International Symposium on Theoretical Aspects of Computer Science*, volume 9 of *LIPICs. Leibniz Int. Proc. Inform.*, pages 57–68. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2011.

- [48] George Giakkoupis. Tight bounds for rumor spreading with vertex expansion. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 801–815. ACM, New York, 2014.
- [49] E. Goles and J. Olivos. Periodic behaviour of generalized threshold functions. *Discrete Math.*, 30(2):187–189, 1980.
- [50] Boris L. Granovsky and Neal Madras. The noisy voter model. *Stochastic Process. Appl.*, 55(1):23–43, 1995.
- [51] T Gross and H Sayama. Adaptive networks: Theory, models, and data, 2009.
- [52] Yehuda Hassin and David Peleg. Distributed probabilistic polling and applications to proportionate agreement. *Inform. and Comput.*, 171(2):248–268, 2001.
- [53] Richard A. Holley and Thomas M. Liggett. Ergodic theorems for weakly interacting infinite systems and the voter model. *Ann. Probability*, 3(4):643–663, 1975.
- [54] Petter Holme and Mark EJ Newman. Nonequilibrium phase transition in the coevolution of networks and opinions. *Physical Review E*, 74(5):056108, 2006.
- [55] Svante Janson. One, two and three times $\log n/n$ for paths in a complete graph with random weights. *Combin. Probab. Comput.*, 8(4):347–361, 1999. Random graphs and combinatorial structures (Oberwolfach, 1997).
- [56] Varun Kanade, Frederik Mallmann-Trenn, and Thomas Sauerwald. On coalescence time in graphs—when is coalescing as fast as meeting? *arXiv preprint arXiv:1611.02460*, 2016.
- [57] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson.

- [58] Erez Lieberman, Christoph Hauert, and Martin A Nowak. Evolutionary dynamics on graphs. *Nature*, 433(7023):312, 2005.
- [59] Thomas Liggett. *Interacting particle systems*, volume 276. Springer Science & Business Media, 2012.
- [60] Thomas M. Liggett. *Stochastic interacting systems: contact, voter and exclusion processes*, volume 324 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999.
- [61] Nancy A. Lynch. *Distributed algorithms*. The Morgan Kaufmann Series in Data Management Systems. Morgan Kaufmann, San Francisco, CA, 1996.
- [62] Michael Mitzenmacher, Rajmohan Rajaraman, and Scott Roche. Better bounds for coalescing-branching random walks. In *Proceedings of the 28th ACM Symposium on Parallelism in Algorithms and Architectures*, pages 313–323. ACM, 2016.
- [63] P. A. P. Moran. Random processes in genetics. volume 54, pages 60–71, 1958.
- [64] Elchanan Mossel and Sébastien Roch. Slow emergence of cooperation for win-stay lose-shift on trees. *Machine Learning*, 67(1):7–22, 2007.
- [65] Thomas Mountford, Jean-Christophe Mourrat, Daniel Valesin, and Qiang Yao. Exponential extinction time of the contact process on finite graphs. *Stochastic Process. Appl.*, 126(7):1974–2013, 2016.
- [66] Toshio Nakata, Hiroshi Imahayashi, and Masafumi Yamashita. Probabilistic local majority voting for the agreement problem on finite graphs. In *Computing and combinatorics (Tokyo, 1999)*, volume 1627 of *Lecture Notes in Comput. Sci.*, pages 330–338. Springer, Berlin, 1999.
- [67] Roberto Imbuzeiro Oliveira. On the coalescence time of reversible random walks. *Trans. Amer. Math. Soc.*, 364(4):2109–2128, 2012.

- [68] Roberto Imbuzeiro Oliveira. Mean field conditions for coalescing random walks. *Ann. Probab.*, 41(5):3420–3461, 2013.
- [69] K. Panagiotou and L. Speidel. Asynchronous rumor spreading on random graphs. *Algorithmica*, 78(3):968–989, 2017.
- [70] Robin Pemantle. The contact process on trees. *Ann. Probab.*, 20(4):2089–2116, 1992.
- [71] Harishchandra Ramadas. Mixing of the noisy voter model. *Electron. Commun. Probab.*, 19:no. 17, 9, 2014.
- [72] Hiroki Sayama, Irene Pestov, Jeffrey Schmidt, Benjamin James Bush, Chun Wong, Junichi Yamanoi, and Thilo Gross. Modeling complex systems with adaptive networks. *Comput. Math. Appl.*, 65(10):1645–1664, 2013.
- [73] B. Sury. Sum of the reciprocals of the binomial coefficients. *European J. Combin.*, 14(4):351–353, 1993.
- [74] Eric W. Weisstein. Stirling’s series. From MathWorld—A Wolfram Web Resource. Last visited on 26/08/2017. URL: <http://mathworld.wolfram.com/StirlingsSeries.html>.
- [75] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.