Decentralized Dynamics for Finite Opinion Games*

Diodato Ferraioli† Paul W. Goldberg‡ Carmine Ventre§

Abstract

Game theory studies situations in which strategic players can modify the state of a given system, due to the absence of a central authority. Solution concepts, such as Nash equilibrium, are defined to predict the outcome of such situations. In the spirit of the field, we then look at the computation of solution concepts by means of decentralized dynamics. These are algorithms in which players move in turns to improve their own utility and the hope is that the system reaches an "equilibrium" quickly.

We study these dynamics for the class of opinion games, recently introduced by [10]. These are games, important in economics and sociology, that model the formation of an opinion in a social network. We study best-response dynamics and show that the convergence to Nash equilibria is polynomial in the number of players. We also study a noisy version of best-response dynamics, called logit dynamics, and prove a host of results about its convergence rate as the noise in the system varies. To get these results, we use a variety of techniques developed to bound the mixing time of Markov chains, including coupling, spectral characterizations and bottleneck ratio.

*Work partially supported by EPSRC grant EP/G069239/1 "Efficient Decentralised Approaches in Algorithmic Game Theory" and by PRIN 2008 research project COGENT (COmputational and GamE-theoretic aspects of uncoordinated NeTworks), funded by the Italian Ministry of University and Research.

†Dipartimento di Informatica "Renato M. Capocelli", Università di Salerno, Italy. E-mail: `ferraioli@dia.unisa.it`.

‡Department of Computer Science, University of Liverpool, UK. E-mail: `P.W.Goldberg@liverpool.ac.uk`.

§School of Computing, Teesside University, UK. E-mail: `C.Ventre@tees.ac.uk`.

1 Introduction

Social networks are widespread in physical and digital worlds. The following scenario therefore becomes of interest. Consider a group of individuals, connected in a social network, who are members of a committee, and suppose that each individual has her own opinion on the matter at hand. How can this group of people reach *consensus*? This is a central question in economic theory, especially for processes in which people repeatedly average their own opinions. This line of work, see e.g. [1, 14, 19, 20], is based on a model defined by DeGroot [13]. In this model, each person i holds an opinion equal to a real number x_i, which might for example represent a position on a political spectrum. There is an undirected graph $G = (V, E)$ representing a social network, and node i is influenced by the opinions of her neighbors in G. In each time step, node i updates her opinion to be an average of her current opinion with the current opinions of her neighbors. A variation of this model of interest to our study is due to Friedkin and Johnsen [18]. In [18] it is additionally assumed that each node i maintains a persistent *internal belief* b_i, which remains constant even as node i updates her overall opinion x_i through averaging. (See Section 2 for the formal framework.)

However, as recently observed by Bindel et al. [10], consensus is hard to reach, the case of political opinions being a prominent example. The authors of [10] justify the absence of consensus by interpreting repeated averaging as a decentralized dynamics for selfish players. Consensus is not reached as players will not compromise further when this diminishes their *utility*. Therefore, these dynamics will converge to an equilibrium in which players might disagree; Bindel et al. study the cost of disagreement by bounding the price of anarchy in this setting.

In this paper, we continue the study of [10] and ask the question of how quickly equilibria are reached by decentralized dynamics in opinion games. We focus on the setting in which players have only a finite number of strategies available. This is motivated by the fact that in many cases although players have personal beliefs which may assume a continuum of values, they only have a limited number of strategies available. For example, in political elections, people have only a limited number of parties they can vote for and usually vote for the party which is *closer* to their own opinions. Motivated by several electoral systems around the world, we concentrate in this study on the case in which players only have two strategies available. This setting already encodes a number of interesting technical challenges as outlined below.

1.1 Our contribution

For the finite version of the opinion games considered in [10], we firstly note that this is a potential game [24, 22] thus implying that these games admit pure Nash equilibria. The set of pure Nash equilibria is then characterized (cf. Claim 2.2). We also notice the interesting fact that while the games in [10] have a price of anarchy of 9/8, our games have unbounded price of anarchy, thus implying that for finite games disagreeing has far more deep consequences on the social cost. These basic facts turn out to be useful in the study of decentralized dynamics for finite opinion games.

Given that the potential function is polynomial in the number of players, by proving that the potential decreases by a constant at each step of the best-response dynamics, we can prove that this dynamics quickly converges to pure Nash equilibria. This result is proved by "reducing" an opinion game to a version of it in which the internal beliefs can only take certain values. The reduced version is equivalent to the original one, as long as best-response dynamics is concerned. Note that the convergence rate for the version of the game considered in [10] is unknown.

In real life, however, there is some noise in the decision process of players. Arguably, people are not fully rational. On the other hand, even if they were, they might not exactly know what strategy represents the best response to a given strategy profile due to the incapacity to correctly determine their utility functions. To model this, we study *logit dynamics* [11] for opinion games. Logit dynamics features a *rationality level* $\beta \geq 0$ (equivalently, a noise level $1/\beta$) and each player is assumed to play a strategy with a probability which is proportional to the corresponding utility to the player and β. So the higher β is, the less noise there is and the more the dynamics is similar to best-response dynamics. Logit dynamics for potential games defines a Markov chain that has a nice structure. As in [5, 4] we exploit this structure to prove bounds on the convergence rate of logit dynamics to the so-called *logit equilibrium*. The logit equilibrium corresponds to the stationary distribution of the Markov chain. Intuitively, a logit equilibrium is a probability distribution over strategy profiles of the game; the distribution is concentrated around pure Nash equilibrium profiles[1]. It is observed in [4] how this notion enjoys a number of desiderata one would like solution concepts to have.

[1] It is worth to notice that the focus of best-response dynamics and logit dynamics is on two different solution concepts.

We prove a host of results on the convergence rate of logit dynamics that give a pretty much complete picture as β varies. We give an upper bound in terms of the cutwidth of the graph modeling the social network. The bound is exponential in β and the cutwidth of the graph, thus yielding an exponential guarantee for some topology of the social network. We complement this result by proving a polynomial upper bound when β takes a small value, namely, for β at most the inverse of the maximum degree of nodes of the graph. We complete the preceding upper bound in terms of the cutwidth with lower bounds. Firstly, we prove that in order to get an (essentially) matching lower bound it is necessary to evaluate the size of a certain subset of strategy profiles. When β is big enough relatively to this subset then we can prove that the upper bound is tight for any social network (specifically, we roughly need β bigger than $n \log n$ over the cutwidth of the graph). For smaller values of β, we are unable to prove a lower bound which holds for every graph. However, we prove that the lower bound holds in this case at both ends of the spectrum of possible social networks. In details, we look at two cases of graphs encoding social networks: cliques, which model monolithic, highly interconnected societies, and complete bipartite graphs, which model more sparse "antitransitive" societies. For these graphs, we firstly evaluate the cutwidth and then relate the latter to the size of the aforementioned set of states. This allows to prove a lower bound exponential in β and the cutwidth of the graph for (almost) any value of β. As far as we know, no previous result was known about the cutwidth of a complete bipartite graph; this might be of independent interest. The result on cliques is instead obtained by generalizing arguments in [21].

To prove the convergence rate of logit dynamics to logit equilibrium we adopt a variety of techniques developed to bound the mixing time of Markov chains. To prove the upper bounds we use some spectral properties of the transition matrix of the Markov chain defined by the logit dynamics, and coupling of Markov chains. To prove the lower bounds, we instead relay on the concept of bottleneck ratio and the relation between the latter and mixing time. (The interested reader might refer to [21] for a discussion of these concepts. Below, we give a quick overview of these techniques and state some useful facts.)

1.2 Related works

In addition to the papers mentioned above, our paper is related to the works on logit dynamics. This dynamics is introduced by Blume [11] and it is mainly adopted in the analysis of graphical coordination games [16, 26, 23], in which players are placed on vertices of a graph embedding social relations and each player wants to coordinate with neighbors: we highlight that an unique game is played on every edge, whereas, for opinion games, we need different games in order to encode beliefs (see below). Asadpour and Saberi [2] adopt the logit dynamics for analyzing a class of congestion games. However, none of these works evaluates the time the logit dynamics takes in order to reach the stationary distribution: this line of research is conducted in [5, 4].

A number of papers study the efficient computation of (approximate) pure Nash equilibria for 2-strategy games, such as, *party affiliation games* [17, 7] and *cut games* [9]. Similarly to these works, we focus on a class of 2-strategy games and study efficient computation of pure Nash equilibria; additionally we also study the convergence rate to logit equilibria.

Another related work is [15] by Dyer and Mohanaraj. They study graphical games, called *pairwise-interaction games*, and prove among other results, quick convergence of best-response dynamics for these games. However, our games do not fall in their class. The difference is that, in their case, there is a unique game being played on the edges of the graph; as noted above, we instead need a different game to encode the internal beliefs of the players.

2 The game

Let $G = (V, E)$ be an undirected connected graph[2] with $|V| = n$. Every vertex of the graph represents a player. Each player i has an *internal belief* $b_i \in [0, 1]$ and only two strategies or *opinions* are available, namely 0 and 1. Motivated by the model in [10], we define the utility of player i in a strategy profile $\mathbf{x} \in \{0, 1\}^n$ as

$$u_i(\mathbf{x}) = -\left((x_i - b_i)^2 + \sum_{j\,:\,(i,j)\in E} (x_i - x_j)^2\right) .$$

We call such a game an n-player opinion game on a graph G. Let $D_i(\mathbf{x}) = \{j : (i, j) \in E \wedge x_i \neq x_j\}$ be the set of neighbors of i that have an opinion different from i. Then $u_i(\mathbf{x}) = -(x_i - b_i)^2 - |D_i(\mathbf{x})|$.

[2] A number of papers, including [10], assume that the graph is weighted to model neighbors' different levels of influence. Here we focus on the case in which all neighbors exert the same kind of "political" weight.

2.1 Potential function

Let $D(\mathbf{x}) = \{(u, v) \in E \colon x_u \neq x_v\}$ be the set of *discording edges* in the strategy profile $\mathbf{x}$, that is the set of all edges in G whose endpoints have different opinions.

Claim 2.1. *The function*

$$\Phi(\mathbf{x}) = \sum_i (x_i - b_i)^2 + |D(\mathbf{x})| \tag{1}$$

is an exact potential function for the opinion game described above.

Proof. Given a strategy profile $\mathbf{x}$, player i experiences a negative utility or equivalently, a positive cost defined as $c_i(\mathbf{x}) = -u_i(\mathbf{x})$. As players try to minimize their cost, similarly, we show that the function Φ defined in (1) is a potential function of the opinion game to minimize. The difference in the cost of the player i when she switches from strategy x_i to strategy y_i is

$$c_i(\mathbf{x}) - c_i(\mathbf{x}_{-i}, y_i) = (x_i - b_i)^2 - (y_i - b_i)^2 + |D_i(\mathbf{x})| - |D_i(\mathbf{x}_{-i}, y_i)| \, .$$

The difference in the potential function between the two corresponding profiles is

$$\begin{aligned} \Phi(\mathbf{x}) - \Phi(\mathbf{x}_{-i}, y_i) &= \sum_j (x_j - b_j)^2 + |D(\mathbf{x})| + \sum_{j \neq i} (x_j - b_j)^2 + (y_i - b_i)^2 + |D(\mathbf{x}_{-i}, y_i)| \\ &= (x_i - b_i)^2 - (y_i - b_i)^2 + |D(\mathbf{x})| - |D(\mathbf{x}_{-i}, y_i)| \, . \end{aligned}$$

Discording edges not incident on i are not affected by the deviation of player i: let K be the number of such edges, then $|D(\mathbf{x})| = K + |D_i(\mathbf{x})|$ and $|D(\mathbf{x}_{-i}, y_i)| = K + |D_i(\mathbf{x}_{-i}, y_i)|$ and the claim follows. $\square$

A more convenient way to express the potential function above, useful in one of the proofs below, is the following: $\Phi(\mathbf{x}) = \sum_{e \in E} \Phi_e(\mathbf{x})$, where, for an edge $e = (i, j)$,

$$\Phi_e(\mathbf{x}) = \begin{cases} \alpha_e := \frac{b_i^2}{\Delta_i} + \frac{b_j^2}{\Delta_j} \, , & \text{if } x_i = x_j = 0 \, ; \\ \beta_e := \frac{b_i^2}{\Delta_i} + \frac{(1-b_j)^2}{\Delta_j} + 1 \, , & \text{if } x_i = 0 \text{ and } x_j = 1 \, ; \\ \gamma_e := \frac{(1-b_i)^2}{\Delta_i} + \frac{b_j^2}{\Delta_j} + 1 \, , & \text{if } x_i = 1 \text{ and } x_j = 0 \, ; \\ \delta_e := \frac{(1-b_i)^2}{\Delta_i} + \frac{(1-b_j)^2}{\Delta_j} \, , & \text{if } x_i = x_j = 1 \, . \end{cases} \tag{2}$$

and Δ_i represents the degree of i.

Interestingly, the potential function looks very similar to (but not the same as) the social cost

$$\mathsf{SC}(\mathbf{x}) = -\sum_{i=1}^{n} u_i(\mathbf{x}) = \sum_i (x_i - b_i)^2 + 2|D(\mathbf{x})| \, .$$

2.2 Nash equilibria and Price of Anarchy

Let B_i be the integer closer to the internal belief of the player i: that is, $B_i = 0$ if $b_i \leq 1/2$, $B_i = 1$ if $b_i > 1/2$. Moreover, let $N_i^s(\mathbf{x}) = |\{j \colon (i, j) \in E \text{ and } x_j = s\}|$ be the number of neighbors of i that play strategy s in the strategy profile $\mathbf{x}$.

The following claim shows that, for every player, it is preferable to select the opinion closer to his own belief if and only if at least half his neighborhood has selected this opinion. The only special cases occur when players have beliefs in $\{0, 1/2, 1\}$: if $b_i = 1/2$ player i will be additionally indifferent when exactly half (assuming that Δ_i is even) of his neighbors are playing the same strategy and the other half are playing the other strategy; if $b_i = 0$ or $b_i = 1$ player i will also be indifferent when Δ_i is odd and only $\lfloor \Delta_i/2 \rfloor$ neighbors are playing B_i.

Claim 2.2. *For every Nash equilibrium profile $\mathbf{x}$ it holds that for each player i*

$$x_i = \begin{cases} B_i \, , & \textit{if } N_i^{B_i}(\mathbf{x}) \geq \frac{\Delta_i}{2} \, ; \\ 1 - B_i \, , & \textit{otherwise.} \end{cases}$$

If $b_i = 1/2$, then $\mathbf{x}$ will be in equilibrium also if $x_i = 1 - B_i$, Δ_i is even and $N_i^{B_i}(\mathbf{x}) = \Delta_i/2$. If $b_i = 0$ or $b_i = 1$, then $\mathbf{x}$ will be in equilibrium also if $x_i = B_i$, Δ_i is odd and $N_i^{B_i}(\mathbf{x}) = \lfloor \Delta_i/2 \rfloor$.

Proof. We first prove that a profile for which the above conditions hold is a Nash equilibrium, and then we prove that every other profile is not in equilibrium. Let $\mathbf{x}$ be a profile for which the above conditions hold for every player and i be one such player. We consider first the case that $N_i^{B_i}(\mathbf{x}) \geq \Delta_i/2$: this means that we are considering $x_i = B_i$. But, for player i it is not convenient to play $1 - B_i$, since

$$u_i(\mathbf{x}_{-i}, 1 - B_i) = -\left[(1 - B_i - b_i)^2 + N_i^{B_i}(\mathbf{x})\right] \leq -\left[(B_i - b_i)^2 + \left(\Delta_i - N_i^{B_i}(\mathbf{x})\right)\right] = u_i(\mathbf{x})\,,$$

where the inequality easily follows for $B_i \in \{0, 1\}$. Similarly, if $N_i^{B_i}(\mathbf{x}) < \frac{\Delta_i}{2}$, we have $x_i = 1 - B_i$ and i has no incentive to switch to B_i, since

$$u_i(\mathbf{x}_{-i}, B_i) = -\left[(B_i - b_i)^2 + \left(\Delta_i - N_i^{B_i}(\mathbf{x})\right)\right] \leq -\left[(1 - B_i - b_i)^2 + N_i^{B_i}(\mathbf{x})\right] = u_i(\mathbf{x})\,.$$

Moreover if $b_i = 1/2$, Δ_i is even, $N_i^{B_i}(\mathbf{x}) = \Delta_i/2$ and $x_i = 1 - B_i$, we have $u_i(\mathbf{x}) = u_i(\mathbf{x}_{-i}, B_i)$; finally, if $b_i \in \{0, 1\}$, Δ_i is odd, $N_i^{B_i}(\mathbf{x}) = \lfloor \Delta_i/2 \rfloor$ and $x_i = B_i$, we have $u_i(\mathbf{x}) = u_i(\mathbf{x}_{-i}, 1 - B_i)$.

Now consider a profile $\mathbf{y}$ for which the conditions above do not hold for some player i: if $N_i^{B_i}(\mathbf{y}) \geq \Delta_i/2$, this means that $y_i = 1 - B_i$ and, in case Δ_i is even and $N_i^{B_i}(\mathbf{y}) = \Delta_i/2$ then $b_i \neq 1/2$; similarly, if $N_i^{B_i}(\mathbf{y}) < \frac{\Delta_i}{2}$, we have that $y_i = B_i$ and, in case Δ_i is odd and $N_i^{B_i}(\mathbf{y}) = \lfloor \Delta_i/2 \rfloor$ then $b_i \notin \{0, 1\}$. But, it is easy to check that in the former case $u_i(\mathbf{y}) < u_i(\mathbf{y}_{-i}, B_i)$ and in the latter case $u_i(\mathbf{y}) < u_i(\mathbf{y}_{-i}, 1 - B_i)$. ☐

Roughly speaking, Claim 2.2 shows that in a Nash equilibrium players tend to form large coalitions, by preferring to play what the majority plays to their own beliefs.

It is easy to check that this game has infinite Price of Anarchy. Consider the opinion game on a clique where each player has internal belief 0: the profile where each player has opinion 0 has social cost 0. By Claim 2.2, the profile where each player has opinion 1 is a Nash equilibrium and its social cost is $n > 0$. This is in sharp contrast with the bound $9/8$ proved in [10].

3 Best-response dynamics

Given two games we say they are *best-response equivalent* if each player has identical best responses to every combination of opponents' strategies. For the opinion games the following observation is straightforward.

Observation 3.1. *Let $\mathcal{G}$ be an opinion game where the player i has belief $b_i \in (0, 1/2)$: then $\mathcal{G}$ is best-response equivalent to the same game where the belief of i is set to $b_i = 1/4$. Similarly, if the player i has opinion $b_i \in (1/2, 1)$ the game is best-response equivalent to the same game where the belief of i is set to $b_i = 3/4$.*

The following theorem shows that, for this class of games, the best-response dynamics quickly converges to a Nash equilibrium.

Theorem 3.2. *The best-response dynamics for an n-player opinion game $\mathcal{G}$ converges to a Nash equilibrium after a polynomial number of steps.*

Proof. From Observation 3.1 we know that each opinion game is best-response equivalent to an opinion game where each player i has $b_i \in S = \left\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\right\}$. So, for a given opinion game $\mathcal{G}$ we construct a game $\mathcal{G}'$ with beliefs restricted to belong to S by "rounding" the beliefs of the original game and show that best-response dynamics converges quickly on $\mathcal{G}'$.

We begin by observing that for every profile $\mathbf{x}$, we have $0 \leq \Phi(\mathbf{x}) \leq n^2 + n$. Thus, the theorem follows by showing that at each time step the cost of a player decreases by at least a constant value. Fix $\mathbf{x}_{-i}$, the opinions of players other than i, and let x_i be the strategy currently played by player i and s be her best response. By definition of best response, we have $c_i(\mathbf{x}) > c_i(s, \mathbf{x}_{-i})$. We want to show

$$\Phi(\mathbf{x}) - \Phi(s, \mathbf{x}_{-i}) = c_i(x_i, \mathbf{x}_{-i}) - c_i(s, \mathbf{x}_{-i}) = (x_i - b_i)^2 - (s - b_i)^2 + |D_i(\mathbf{x})| - |D_i(s, \mathbf{x}_{-i})| \geq \frac{1}{2}\,. \quad (3)$$

(Recall that c_i denotes the cost of player i, defined as the opposite of the utility.) Since the difference between the squares is bounded from below by -1 then (3) follows in case $|D_i(\mathbf{x})| - |D_i(s, \mathbf{x}_{-i})| \geq 2$. For the sequel of the proof note that $c_i(x_i, \mathbf{x}_{-i}) - c_i(s, \mathbf{x}_{-i}) = -1 + 2(x_i + b_i) - 4x_ib_i + |D_i(\mathbf{x})| - |D_i(s, \mathbf{x}_{-i})|$.

If $|D_i(\mathbf{x})| - |D_i(s, \mathbf{x}_{-i})| = 1$, we distinguish two sub-cases: if $x_i = 0$ then $c_i(x_i, \mathbf{x}_{-i}) > c_i(s, \mathbf{x}_{-i})$ implies $b_i > 0$, and by inspection for $b_i \in \{1/4, 1/2, 3/4, 1\}$ (3) holds; if $x_i = 1$ then $c_i(x_i, \mathbf{x}_{-i}) > c_i(s, \mathbf{x}_{-i})$ implies $b_i < 1$ and then it is easy to check that (3) is satisfied for $b_i \in \{0, 1/4, 1/2, 3/4\}$.

If $|D_i(\mathbf{x})| - |D_i(s, \mathbf{x}_{-i})| = 0$, we similarly distinguish two sub-cases: if $x_i = 0$, $c_i(x_i, \mathbf{x}_{-i}) > c_i(s, \mathbf{x}_{-i})$ implies $b_i > 1/2$, and by inspection for $b_i \in \{3/4, 1\}$ (3) is true; for $x_i = 1$, $c_i(x_i, \mathbf{x}_{-i}) > c_i(s, \mathbf{x}_{-i})$ implies $b_i < 1/2$ and then it is easy to check that (3) is satisfied for $b_i \in \{0, 1/4\}$.

Finally, we show that $|D_i(\mathbf{x})| \geq |D_i(s, \mathbf{x}_{-i})|$. Indeed, if by contradiction, $|D_i(\mathbf{x})| = |D_i(s, \mathbf{x}_{-i})| - k$, $k \geq 1$, then as $c_i(x_i, \mathbf{x}_{-i}) - c_i(s, \mathbf{x}_{-i}) = -1 + 2(x_i + b_i) - 4x_i b_i - k$ and since, for $x_i \in \{0, 1\}$, $-1 + 2(x_i + b_i) - 4x_i b_i \leq 1$ we would reach the contradiction $c_i(x_i, \mathbf{x}_{-i}) \leq c_i(s, \mathbf{x}_{-i})$. □

4 Logit Dynamics for Opinion Games

Let $\mathcal{G}$ be an opinion game as from the above; moreover, let $S = \{0, 1\}^n$ denote the set of all strategy profiles. For two vectors $\mathbf{x}, \mathbf{y} \in S$, we denote with $H(\mathbf{x}, \mathbf{y}) = |\{i \colon x_i \neq y_i\}|$ the Hamming distance between $\mathbf{x}$ and $\mathbf{y}$. The *Hamming graph* of the game $\mathcal{G}$ is defined as $\mathcal{H} = (S, \mathsf{E})$, where two profiles $\mathbf{x} = (x_1, \ldots, x_n), \mathbf{y} = (y_1, \ldots, y_n) \in S$ are adjacent in $\mathcal{H}$ if and only if $H(\mathbf{x}, \mathbf{y}) = 1$.

The *logit dynamics* for $\mathcal{G}$ runs as follows: at every time step (i) Select one player $i \in [n]$ uniformly at random; (ii) Update the strategy of player i according to the *Boltzmann distribution* with parameter β over the set $S_i = \{0, 1\}$ of her strategies. That is, a strategy $s_i \in S_i$ will be selected with probability

$$\sigma_i(s_i \mid \mathbf{x}_{-i}) = \frac{1}{Z_i(\mathbf{x}_{-i})} \, e^{\beta u_i(\mathbf{x}_{-i}, s_i)} \,, \tag{4}$$

where $\mathbf{x}_{-i} \in \{0, 1\}^{n-1}$ is the profile of strategies played at the current time step by players different from i, $Z_i(\mathbf{x}_{-i}) = \sum_{z_i \in S_i} e^{\beta u_i(\mathbf{x}_{-i}, z_i)}$ is the normalizing factor, and $\beta \geq 0$. As mentioned above, from (4), it is easy to see that for $\beta = 0$ player i selects her strategy uniformly at random, for $\beta > 0$ the probability is biased toward strategies promising higher payoffs, and for β that goes to ∞ player i chooses her best response strategy (if more than one best response is available, she chooses one of them uniformly at random).

The above dynamics defines a *Markov chain* $\{X_t\}_{t \in \mathbb{N}}$ with the set of strategy profiles as state space, and where the probability $P(\mathbf{x}, \mathbf{y})$ of a transition from profile $\mathbf{x} = (x_1, \ldots, x_n)$ to profile $\mathbf{y} = (y_1, \ldots, y_n)$ is zero if $H(\mathbf{x}, \mathbf{y}) \geq 2$ and it is $\frac{1}{n} \sigma_i(y_i \mid \mathbf{x}_{-i})$ if the two profiles differ exactly at player i. More formally, we can define the logit dynamics as follows.

Definition 4.1 (Logit dynamics [11])**.** *Let $\mathcal{G}$ be an opinion game as from the above and let $\beta \geq 0$. The* logit dynamics *for $\mathcal{G}$ is the Markov chain $\mathcal{M}_\beta = (\{X_t\}_{t \in \mathbb{N}}, S, P)$ where $S = \{0, 1\}^n$ and*

$$P(\mathbf{x}, \mathbf{y}) = \frac{1}{n} \cdot \begin{cases} \sigma_i(y_i \mid \mathbf{x}_{-i}), & \textit{if } \mathbf{y}_{-i} = \mathbf{x}_{-i} \textit{ and } y_i \neq x_i; \\ \sum_{i=1}^n \sigma_i(y_i \mid \mathbf{x}_{-i}), & \textit{if } \mathbf{y} = \mathbf{x}; \\ 0, & \textit{otherwise;} \end{cases} \tag{5}$$

where $\sigma_i(y_i \mid \mathbf{x}_{-i})$ is defined in (4)*.*

The Markov chain defined by (5) is ergodic. Hence, from every initial profile $\mathbf{x}$ the distribution $P^t(\mathbf{x}, \cdot)$ of chain X_t starting at $\mathbf{x}$ will eventually converge to a *stationary distribution* π as t tends to infinity.[3] As in [4], we call the stationary distribution π of the Markov chain defined by the logit dynamics on a game $\mathcal{G}$, the *logit equilibrium* of $\mathcal{G}$. In general, a Markov chain with transition matrix P and state space S is said to be *reversible* with respect to the distribution π if, for all $\mathbf{x}, \mathbf{y} \in S$, it holds that $\pi(\mathbf{x}) P(\mathbf{x}, \mathbf{y}) = \pi(\mathbf{y}) P(\mathbf{y}, \mathbf{x})$. If the chain is reversible with respect to π, then π is its stationary distribution. Therefore when this happens, to simplify our exposition we simply say that the matrix P is reversible. For the class of potential games the stationary distribution is the well known *Gibbs measure*.

Theorem 4.2 ([11])**.** *If $\mathcal{G} = ([n], \mathcal{S}, \mathcal{U})$ is a potential game with potential function Φ, then the Markov chain given by* (5) *is reversible with respect to the Gibbs measure $\pi(\mathbf{x}) = \frac{1}{Z} e^{-\beta \Phi(\mathbf{x})}$, where $Z = \sum_{\mathbf{y} \in S} e^{-\beta \Phi(\mathbf{y})}$ is the normalizing constant.*

[3] The notation $P^t(\mathbf{x}, \cdot)$, standard in Markov chains literature [21], denotes the probability distribution over states of S after the chain has taken t steps starting from $\mathbf{x}$.

Mixing time of Markov chains. One of the prominent measures of the rate of convergence of a Markov chain to its stationary distribution is the *mixing time*. For a Markov chain with transition matrix P and state space S, let us set

$$d(t) = \max_{\mathbf{x} \in S} \left\| P^t(\mathbf{x}, \cdot) - \pi \right\|_{\mathrm{TV}} ,$$

where the *total variation distance* $\|\mu - \nu\|_{\mathrm{TV}}$ between two probability distributions μ and ν on the same state space S is defined as

$$\|\mu - \nu\|_{\mathrm{TV}} = \max_{A \subset S} |\mu(A) - \nu(A)| = \frac{1}{2} \sum_{\mathbf{x} \in S} |\mu(\mathbf{x}) - \nu(\mathbf{x})| = \sum_{\substack{\mathbf{x} \in S: \\ \mu(\mathbf{x}) > \nu(\mathbf{x})}} (\mu(\mathbf{x}) - \nu(\mathbf{x})) .$$

For $0 < \varepsilon < 1/2$, the mixing time is defined as

$$t_{\text{mix}}(\varepsilon) = \min\{t \in \mathbb{N} \colon d(t) \leq \varepsilon\}.$$

It is usual to set $\varepsilon = 1/4$ or $\varepsilon = 1/2e$. If not explicitly specified, when we write t_{mix} we mean $t_{\text{mix}}(1/4)$. Observe that $t_{\text{mix}}(\varepsilon) \leq \lceil \log_2 \varepsilon^{-1} \rceil t_{\text{mix}}$.

4.1 Techniques

To derive our bounds, we employ several different techniques: *Markov chain coupling* and *spectral techniques* for the upper bound and *bottleneck ratio* for the lower bound. They are well-established techniques for bounding the mixing time; we next summarize them.

4.1.1 Markov chain coupling

A *coupling* of two probability distributions μ and ν on S is a pair of random variables (X, Y) defined on $S \times S$ such that the marginal distribution of X is μ and the marginal distribution of Y is ν. A *coupling of a Markov chain* $\mathcal{M}$ on S with transition matrix P is a process $(X_t, Y_t)_{t=0}^{\infty}$ with the property that X_t and Y_t are both Markov chains with transition matrix P. The importance of coupling for the mixing time of Markov chains is summarized in Appendix A.1. We here describe only the coupling that we use in the proof of Theorem 4.5 below. Our exposition follows the one in [5].

For every pair of strategy profiles $\mathbf{x} = (x_1, \ldots, x_n), \mathbf{y} = (y_1, \ldots, y_n) \in \{0,1\}^n$ we define a coupling (X_1, Y_1) of two copies of the Markov chain for which $X_0 = \mathbf{x}$ and $Y_0 = \mathbf{y}$. The coupling proceeds as follows: first, pick a player i uniformly at random; then, update the strategies x_i and y_i of player i in the two chains, by setting

$$(x_i, y_i) = \begin{cases} (0,0), & \text{with probability } \min\{\sigma_i(0 \mid \mathbf{x}), \sigma_i(0 \mid \mathbf{y})\}\,; \\ (1,1), & \text{with probability } \min\{\sigma_i(1 \mid \mathbf{x}), \sigma_i(1 \mid \mathbf{y})\}\,; \\ (0,1), & \text{with probability } \sigma_i(0 \mid \mathbf{x}) - \min\{\sigma_i(0 \mid \mathbf{x}), \sigma_i(0 \mid \mathbf{y})\}\,; \\ (1,0), & \text{with probability } \sigma_i(1 \mid \mathbf{x}) - \min\{\sigma_i(1 \mid \mathbf{x}), \sigma_i(1 \mid \mathbf{y})\}\,. \end{cases}$$

Three easy observations are in order: if $\sigma_i(0 \mid \mathbf{x}) = \sigma_i(0 \mid \mathbf{y})$ and player i is chosen, then, after the update, we have $x_i = y_i$; for every player i, at most one of the updates $(x_i, y_i) = (0,1)$ and $(x_i, y_i) = (1,0)$ has positive probability; if i is chosen for update, then the marginal distributions of x_i and y_i agree with $\sigma_i(\cdot \mid \mathbf{x})$ and $\sigma_i(\cdot \mid \mathbf{y})$ respectively, indeed, for $b \in \{0,1\}$, the probability that $x_i = b$ is

$$\min\{\sigma_i(b \mid \mathbf{x}), \sigma_i(b \mid \mathbf{y})\} + \sigma_i(b \mid \mathbf{x}) - \min\{\sigma_i(b \mid \mathbf{x}), \sigma_i(b \mid \mathbf{y})\} = \sigma_i(b \mid \mathbf{x})\,,$$

and the probability that $y_i = b$ is

$$\begin{aligned} &\min\{\sigma_i(b \mid \mathbf{x}), \sigma_i(b \mid \mathbf{y})\} + \sigma_i(1-b \mid \mathbf{x}) - \min\{\sigma_i(1-b \mid \mathbf{x}), \sigma_i(1-b \mid \mathbf{y})\} = \\ &= \min\{\sigma_i(b \mid \mathbf{x}), \sigma_i(b \mid \mathbf{y}) \ + (1 - \sigma_i(b \mid \mathbf{x})) - (1 - \max\{\sigma_i(b \mid \mathbf{x}), \sigma_i(b \mid \mathbf{y})\}) = \sigma_i(b \mid \mathbf{y})\,. \end{aligned}$$

For the path coupling technique (see Theorem A.2), the coupling described above is applied only to pairs of starting profiles which are adjacent in $\mathcal{H}$.

4.1.2 Relaxation time and spectral techniques

Another important measure related to the convergence of Markov chains is given by the *relaxation time*. Let P be the transition matrix of a Markov chain with finite state space S and let us label the eigenvalues of P in non-increasing order $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{|S|}$. It is well-known (see, for example, Lemma 12.1 in [21]) that $\lambda_1 = 1$ and, if P is irreducible and aperiodic, then $\lambda_2 < 1$ and $\lambda_{|S|} > -1$. We set $\lambda^\star$ as the largest eigenvalue in absolute value other than λ_1, that is, $\lambda^\star = \max_{i=2,\dots,|S|}\{|\lambda_i|\}$. The *relaxation time* t_{rel} of a Markov chain $\mathcal{M}$ is defined as

$$t_{\text{rel}} = \frac{1}{1-\lambda^\star}\,.$$

It turns out that mixing time and relaxation time are strictly related (see Appendix A.2). We here mention only a useful corollary of Lemma A.4.

In [3] it is proved that all the eigenvalues of the transition matrix of the logit dynamics for potential games are non-negative. We can then prove the following.

Corollary 4.3. *Let $\mathcal{G}$ be an n-player opinion game with profile space S and let P and π be the transition matrix and the stationary distribution of the logit dynamics for $\mathcal{G}$, respectively. For every pair of profiles $\mathbf{x}, \mathbf{y}$ we assign a path $\Gamma_{\mathbf{x},\mathbf{y}}$ on the Hamming graph $\mathcal{H}$. Then*

$$t_{\text{rel}} \leq 2n \max_{\substack{\mathbf{z},\mathbf{w}:\\ H(\mathbf{z},\mathbf{w})=1}} \frac{1}{\pi(\perp_{\mathbf{z},\mathbf{w}})} \sum_{\substack{\mathbf{x},\mathbf{y}:\\ (\mathbf{z},\mathbf{w})\in\Gamma_{\mathbf{x},\mathbf{y}}}} \pi(\mathbf{x})\pi(\mathbf{y})|\Gamma_{\mathbf{x},\mathbf{y}}|,$$

where for every pair of profiles $\mathbf{x}, \mathbf{y} \in S$ we set $\perp_{\mathbf{x},\mathbf{y}} = \arg\min\{\pi(\mathbf{x}), \pi(\mathbf{y})\}$ and $\top_{\mathbf{x},\mathbf{y}} = \arg\max\{\pi(\mathbf{x}), \pi(\mathbf{y})\}$.

Proof. Since all the eigenvalues of P are non-negative, it follows that $t_{\text{rel}} = \frac{1}{1-\lambda_2}$. Moreover, by reversibility of P, we have that for $\mathbf{z}, \mathbf{w} \in S$ such that $H(\mathbf{z}, \mathbf{w}) = 1$ it holds: $Q(\mathbf{z}, \mathbf{w}) = \pi(\perp_{\mathbf{z},\mathbf{w}})P(\perp_{\mathbf{z},\mathbf{w}}, \top_{\mathbf{z},\mathbf{w}}) \geq \frac{\pi(\perp_{\mathbf{z},\mathbf{w}})}{2n}$. Thus, the claim follows from Lemma A.4. □

4.1.3 Bottleneck ratio

Finally, an important concept to establish our lower bounds is represented by the *bottleneck ratio*. Consider an ergodic Markov chain with finite state space S, transition matrix P, and stationary distribution π. The probability distribution $Q(\mathbf{x}, \mathbf{y}) = \pi(\mathbf{x})P(\mathbf{x}, \mathbf{y})$ is of particular interest and is sometimes called the *edge stationary distribution*. Note that if the chain is reversible then $Q(\mathbf{x}, \mathbf{y}) = Q(\mathbf{y}, \mathbf{x})$. For any $L \subseteq S$, we let $Q(L, S \setminus L) = \sum_{\mathbf{x}\in L,\mathbf{y}\in S\setminus L} Q(\mathbf{x}, \mathbf{y})$. The bottleneck ratio of $L \subseteq S$, L non-empty, is

$$B(L) = \frac{Q(L, S \setminus R)}{\pi(L)}\,.$$

Useful facts about the bottleneck ratio, used in the sequel, are surveyed in Appendix A.3.

4.2 Upper bounds

4.2.1 For every β

Consider the bijective function $\sigma\colon V \to \{1, \dots, |V|\}$: it represents an ordering of vertices of G. Let $\mathcal{L}$ be the set of all orderings of vertices of G and set $V_i^\sigma = \{v \in V : \sigma(v) < i\}$. Then, the *cutwidth* of G is

$$\mathsf{CW}(G) = \min_{\sigma\in\mathcal{L}} \max_{1<i\leq|V|} |E(V_i^\sigma, V \setminus V_i^\sigma)|\,.$$

Theorem 4.4. *Let $\mathcal{G}$ be an n-player opinion game on a graph $G = (V, E)$. The mixing time of the logit dynamics for $\mathcal{G}$ is $t_{\text{mix}} \leq (1 + \beta) \cdot \mathsf{poly}\,(n) \cdot e^{\beta\Theta(\mathsf{CW}(G))}$.*

Proof. This proof is a generalization of a similar proof given by Berger et al. [8].

Consider the ordering of vertices of G that obtains the cutwidth. Fix $\mathbf{x}, \mathbf{y} \in S$ and let $v_1, v_2, \dots, v_d$ denote the indices (according to this ordering) of the vertices at which the profiles $\mathbf{x}$ and $\mathbf{y}$ differ; we consider the path $\Gamma_{\mathbf{x},\mathbf{y}} = (\mathbf{x}^0, \mathbf{x}^1, \dots, \mathbf{x}^d)$ on $\mathcal{H}$, where

$$\mathbf{x}^i = \left(y_1, \dots, y_{v_{i+1}-1}, x_{v_{i+1}}, \dots, x_n\right)\,.$$

(Above, we assume $v_{d+1} = n+1$). Notice that $\mathbf{x}^0 = \mathbf{x}$, $\mathbf{x}^d = \mathbf{y}$ and $|\Gamma_{\mathbf{x},\mathbf{y}}| \leq n$. For every edge $\xi = (\mathbf{x}^i, \mathbf{x}^{i+1})$ of $\mathcal{H}$, we consider the function Λ_ξ that assigns to every pair of profiles $\mathbf{x}, \mathbf{y}$ such that $\xi \in \Gamma_{\mathbf{x},\mathbf{y}}$, the following new profile

$$\Lambda_\xi(\mathbf{x}, \mathbf{y}) = \begin{cases} (x_1, \ldots, x_{v_{i+1}-1}, y_{v_{i+1}}, y_{v_{i+1}+1}, \ldots, y_n) & \text{if } \pi(\mathbf{x}^i) \leq \pi(\mathbf{x}^{i+1})\,; \\ (x_1, \ldots, x_{v_{i+1}-1}, x_{v_{i+1}}, y_{v_{i+1}+1}, \ldots, y_n) & \text{otherwise,} \end{cases}$$

where π denotes the stationary distribution (cf. Theorem 4.2). It is easy to see that Λ_ξ is an injective function: indeed, since ξ is known, if $\pi(\mathbf{x}^i) \leq \pi(\mathbf{x}^{i+1})$, then we can retrieve v_{i+1}, that is the first vertex where $\mathbf{x}^i$ and $\mathbf{x}^{i+1}$ differ and thus, selecting the first $v_{i+1} - 1$ vertices from $\Lambda_\xi(\mathbf{x}, \mathbf{y})$ and the remaining ones from $\mathbf{x}^i$ we are able to reconstruct $\mathbf{x}$ and, similarly, selecting the first $v_{i+1} - 1$ vertices from $\mathbf{x}^i$ and the remaining ones from $\Lambda_\xi(\mathbf{x}, \mathbf{y})$ we are able to reconstruct $\mathbf{y}$. Similarly, if $\pi(\mathbf{x}^i) > \pi(\mathbf{x}^{i+1})$, we can retrieve v_{i+2} and we can reconstruct $\mathbf{x}$ and $\mathbf{y}$ from $\Lambda_\xi(\mathbf{x}, \mathbf{y})$ and $\mathbf{x}^{i+1}$.

Let $E^\star = \{(j,k) \in E \colon j < v_{i+1} \text{ and } k \geq v_{i+1}\}$: observe that $|E^\star| \leq \mathsf{CW}(G)$. For any edge $e = (j,k) \in E^\star$, for every $\mathbf{x}, \mathbf{y} \in S$ and for every $\xi = (\mathbf{x}^i, \mathbf{x}^{i+1}) \in \Gamma_{\mathbf{x},\mathbf{y}}$, we distinguish two cases:

<u>If $x_j = y_j$ or $x_k = y_k$,</u> for all available values of x_j, y_j, x_k and y_k we show

$$\Phi_e(\mathbf{x}) + \Phi_e(\mathbf{y}) - \Phi_e(\perp_{\mathbf{x}^i,\mathbf{x}^{i+1}}) - \Phi_e(\Lambda_\xi(\mathbf{x}, \mathbf{y})) = 0\,,$$

where as above $\perp_{\mathbf{x}^i,\mathbf{x}^{i+1}} = \arg\min\{\pi(\mathbf{x}^i), \pi(\mathbf{x}^{i+1})\}$. Firstly, assume that $x_j = y_j$ and $\perp_{\mathbf{x}^i,\mathbf{x}^{i+1}} = \mathbf{x}^i$ which in turns implies that $\Lambda_\xi(\mathbf{x}, \mathbf{y}) = (x_1, \ldots, x_{v_{i+1}-1}, y_{v_{i+1}}, y_{v_{i+1}+1}, \ldots, y_n)$. We have:

$$\begin{aligned} \Phi_e(\mathbf{x}) + \Phi_e(\mathbf{y}) - \Phi_e(\perp_{\mathbf{x}^i,\mathbf{x}^{i+1}}) - \Phi_e(\Lambda_\xi(\mathbf{x}, \mathbf{y})) &= \Phi_e(x_j, x_k) + \Phi_e(y_j, y_k) - \Phi_e(y_j, x_k) - \Phi_e(x_j, y_k) \\ &= \Phi_e(x_j, x_k) + \Phi_e(x_j, y_k) - \Phi_e(x_j, x_k) - \Phi_e(x_j, y_k) = 0\,. \end{aligned}$$

It is not hard to check that the same is true for all the other possible cases arising.

<u>If $x_j \neq y_j$ and $x_k \neq y_k$,</u> similarly to the above, it is not hard to see that for all available values of x_j, y_j, x_k and y_k

$$\Phi_e(\mathbf{x}) + \Phi_e(\mathbf{y}) - \Phi_e(\perp_{\mathbf{x}^i,\mathbf{x}^{i+1}}) - \Phi_e(\Lambda_\xi(\mathbf{x}, \mathbf{y})) = \pm(\alpha_e + \delta_e - \beta_e - \gamma_e) = \pm 2\,,$$

where $\alpha_e, \beta_e, \gamma_e$ and δ_e are defined in (2). Moreover for $e = (j,k) \in E \setminus E^\star$ it holds:

$$\Phi_e(\mathbf{x}) + \Phi_e(\mathbf{y}) - \Phi_e(\perp_{\mathbf{x}^i,\mathbf{x}^{i+1}}) - \Phi_e(\Lambda_\xi(\mathbf{x}, \mathbf{y})) = 0$$

since, by construction, one of $\perp_{\mathbf{x}^i,\mathbf{x}^{i+1}}$ and $\Lambda_\xi(\mathbf{x}, \mathbf{y})$ has j-th and k-th entry of $\mathbf{x}$ and the other has j-th and k-th entry of $\mathbf{y}$. Thus, we have that for every $\mathbf{x}, \mathbf{y} \in S$ and for every $\xi = (\mathbf{x}^i, \mathbf{x}^{i+1}) \in \Gamma_{\mathbf{x},\mathbf{y}}$,

$$\begin{aligned} \Phi(\mathbf{x}) + \Phi(\mathbf{y}) - \Phi(\perp_{\mathbf{x}^i,\mathbf{x}^{i+1}}) - \Phi(\Lambda_\xi(\mathbf{x}, \mathbf{y})) &= \sum_{e \in E} \left(\Phi_e(\mathbf{x}) + \Phi_e(\mathbf{y}) - \Phi_e(\perp_{\mathbf{x}^i,\mathbf{x}^{i+1}}) - \Phi_e(\Lambda_\xi(\mathbf{x}, \mathbf{y}))\right) \\ &= \sum_{e \in E^\star} \left(\Phi_e(\mathbf{x}) + \Phi_e(\mathbf{y}) - \Phi_e(\perp_{\mathbf{x}^i,\mathbf{x}^{i+1}}) - \Phi_e(\Lambda_\xi(\mathbf{x}, \mathbf{y}))\right) \\ &\geq -2\mathsf{CW}(G)\,. \end{aligned} \tag{6}$$

Now let $\xi^\star = (\mathbf{z}, \mathbf{w})$ be the edge of $\mathcal{H}$ for which $\sum_{\substack{\mathbf{x},\mathbf{y}\colon \\ \xi^\star \in \Gamma_{\mathbf{x},\mathbf{y}}}} \frac{\pi(\mathbf{x})\pi(\mathbf{y})}{\pi(\perp_{\mathbf{z},\mathbf{w}})} |\Gamma_{\mathbf{x},\mathbf{y}}|$ is maximized. Applying Corollary 4.3, we obtain

$$\begin{aligned} t_{\text{rel}} &\leq 2n \sum_{\substack{\mathbf{x},\mathbf{y}\colon \\ \xi^\star \in \Gamma_{\mathbf{x},\mathbf{y}}}} \frac{\pi(\mathbf{x})\pi(\mathbf{y})}{\pi(\perp_{\mathbf{z},\mathbf{w}})} |\Gamma_{\mathbf{x},\mathbf{y}}| \leq 2n^2 \sum_{\substack{\mathbf{x},\mathbf{y}\colon \\ \xi^\star \in \Gamma_{\mathbf{x},\mathbf{y}}}} \frac{\pi(\mathbf{x})\pi(\mathbf{y})}{\pi(\perp_{\mathbf{z},\mathbf{w}})\pi(\Lambda_{\xi^\star}(\mathbf{x}, \mathbf{y}))} \pi(\Lambda_{\xi^\star}(\mathbf{x}, \mathbf{y})) \\ &\leq 2n^2 e^{2\beta \mathsf{CW}(G)} \sum_{\substack{\mathbf{x},\mathbf{y}\colon \\ \xi^\star \in \Gamma_{\mathbf{x},\mathbf{y}}}} \pi(\Lambda_{\xi^\star}(\mathbf{x}, \mathbf{y})) \leq 2n^2 e^{2\beta \mathsf{CW}(G)} \sum_{\mathbf{x}} \pi(\mathbf{x}) \leq 2n^2 e^{2\beta \mathsf{CW}(G)}\,, \end{aligned}$$

where the third inequality follows from Theorem 4.2 and (6), and the penultimate from the fact that Λ_ξ is injective.

The theorem follows from Theorem A.3 and by observing that, since $\Phi(\mathbf{x}) \geq 0$ for any strategy profile $\mathbf{x}$, Theorem 4.2 implies

$$\begin{aligned} \log\left((\pi_{\min}/4)^{-1}\right) &= \log\left(4 \sum_{\mathbf{x}} e^{-\beta(\Phi(\mathbf{x}) - \Phi_{\max})}\right) \\ &\leq \log\left(2^{n+2} \cdot e^{\beta \Phi_{\max}}\right) \leq \log\left(e^{n+2+\beta\Phi_{\max}}\right) = n + 2 + \beta \Phi_{\max}\,, \end{aligned} \tag{7}$$

where $\Phi_{\max} = \max_{\mathbf{x}} \Phi(\mathbf{x}) \leq n + |E| = \mathcal{O}(n^2)$. □

4.2.2 For small β

The following theorem shows that for small values of β the mixing time is polynomial. We remark that there are network topologies for which this theorem gives a bound higher than that guaranteed by Theorem 4.4 on the values of β for which the mixing time is polynomial.

Theorem 4.5. *Let $\mathcal{G}$ be an n-player opinion game on a connected graph G, with $n > 2$. Let $\Delta_{\max}$ be the maximum degree in the graph. If $\beta \leq 1/\Delta_{\max}$, then the mixing time of the logit dynamics for $\mathcal{G}$ is $\mathcal{O}(n \log n)$.*

Proof. Consider two profiles $\mathbf{x}$ and $\mathbf{y}$ that differ only in the strategy played by player j. W.l.o.g., we assume $x_j = 1$ and $y_j = 0$. We consider the coupling described in Section 4.1.1 for two chains X and Y starting respectively from $X_0 = \mathbf{x}$ and $Y_0 = \mathbf{y}$. We next compute the expected distance between X_1 and Y_1 after one step of the coupling.

Let N_i be the set of neighbors of i in the opinion game. Notice that for any player i, $\sigma_i(0 \mid \mathbf{x})$ only depends on x_k, for any $k \in N_i$, and $\sigma_i(0 \mid \mathbf{y})$ only on y_k, for any $k \in N_i$. Therefore, since $\mathbf{x}$ and $\mathbf{y}$ only differ at position j, $\sigma_i(0 \mid \mathbf{x}) = \sigma_i(0 \mid \mathbf{y})$ for $i \notin N_j$.

We start by observing that if position j is chosen for update (this happens with probability $1/n$) then, by the observation above, both chains perform the same update. Since $\mathbf{x}$ and $\mathbf{y}$ differ only for player j, we have that the two chains are coupled (and thus at distance 0). Similarly, if player $i \neq j$ with $i \notin N_j$ is selected for update (which happens with probability $(n - \Delta_j - 1)/n$) we have that both chains perform the same update and thus remain at distance 1. Finally, let us consider the case in which $i \in N_j$ is selected for update. In this case, since $x_j = 1$ and $y_j = 0$, we have that $\sigma_i(0 \mid \mathbf{x}) \leq \sigma_i(0 \mid \mathbf{y})$. Therefore, with probability $\sigma_i(0 \mid \mathbf{x})$ both chains update position i to 0 and thus remain at distance 1; with probability $\sigma_i(1 \mid \mathbf{y}) = 1 - \sigma_i(0 \mid \mathbf{y})$ both chains update position i to 1 and thus remain at distance 1; and with probability $\sigma_i(0 \mid \mathbf{y}) - \sigma_i(0 \mid \mathbf{x})$ chain X updates position i to 1 and chain Y updates position i to 0 and thus the two chains go to distance 2. By summing up, we have that the expected distance $E[\rho(X_1, Y_1)]$ after one step of coupling of the two chains is

$$\begin{aligned}
E[\rho(X_1, Y_1)] &= \frac{n - \Delta_j - 1}{n} + \frac{1}{n} \sum_{i \in N_j} [\sigma_i(0 \mid \mathbf{x}) + 1 - \sigma_i(0 \mid \mathbf{y}) + 2 \cdot (\sigma_i(0 \mid \mathbf{y}) - \sigma_i(0 \mid \mathbf{x}))] \\
&= \frac{n - \Delta_j - 1}{n} + \frac{1}{n} \cdot \sum_{i \in N_j} (1 + \sigma_i(0 \mid \mathbf{y}) - \sigma_i(0 \mid \mathbf{x})) \\
&= \frac{n-1}{n} + \frac{1}{n} \cdot \sum_{i \in N_j} (\sigma_i(0 \mid \mathbf{y}) - \sigma_i(0 \mid \mathbf{x})) .
\end{aligned}$$

Let us now evaluate the difference $\sigma_i(0 \mid \mathbf{y}) - \sigma_i(0 \mid \mathbf{x})$ for some $i \in N_j$. Recall that $N_i^s(\mathbf{x})$ denotes the number of neighbors of i that have opinion s in the profile $\mathbf{x}$. Note that $N_i^0(\mathbf{y}) = N_i^0(\mathbf{x}) + 1$ and $N_i^1(\mathbf{x}) = N_i^1(\mathbf{y}) + 1 = \Delta_i - N_i^0(\mathbf{x})$. For sake of compactness we will denote with ℓ the quantity $e^{\beta(2b_i - 1 + 2N_i^1(\mathbf{x}) - \Delta_i)}$. By (4) we have

$$\sigma_i(0 \mid \mathbf{x}) = \frac{e^{-\beta(b_i^2 + N_i^1(\mathbf{x}))}}{e^{-\beta(b_i^2 + N_i^1(\mathbf{x}))} + e^{-\beta((1-b_i)^2 + \Delta_i - N_i^1(\mathbf{x}))}} = \frac{1}{1 + \ell} ,$$

and

$$\sigma_i(0 \mid \mathbf{y}) = \frac{e^{-\beta(b_i^2 + N_i^1(\mathbf{x}) - 1)}}{e^{-\beta(b_i^2 + N_i^1(\mathbf{x}) - 1)} + e^{-\beta((1-b_i)^2 + \Delta_i - N_i^1(\mathbf{x}) + 1)}} = \frac{1}{1 + \ell e^{-2\beta}} .$$

The function $\frac{1}{1+\ell e^{-2\beta}} - \frac{1}{1+\ell}$ is maximized for $\ell = e^{\beta}$. Thus

$$\sigma_i(0 \mid \mathbf{y}) - \sigma_i(0 \mid \mathbf{x}) \leq \frac{1}{1 + e^{-\beta}} - \frac{1}{1 + e^{\beta}} = \frac{2}{1 + e^{-\beta}} - 1 .$$

By using the well-known approximation $e^{-\beta} \geq 1 - \beta$ and since by hypothesis $\beta \leq 1/\Delta_{\max}$, we have

$$\sigma_i(0 \mid \mathbf{y}) - \sigma_i(0 \mid \mathbf{x}) \leq \beta \cdot \frac{1}{2 - \beta} \leq \frac{1}{\Delta_{\max}} \cdot \frac{\Delta_{\max}}{2\Delta_{\max} - 1} .$$

We can conclude that the expected distance after one step of the chain is

$$E[\rho(X_1, Y_1)] \leq \frac{n-1}{n} + \frac{1}{n} \cdot \frac{\Delta_j}{2\Delta_{\max} - 1} \leq \frac{n-1}{n} + \frac{2}{3n} = 1 - \frac{1}{3n} \leq e^{-\frac{1}{3n}} .$$

where the second inequality relies on the fact that $\Delta_{\max} \geq 2$, since the social graph is connected and $n > 2$. Since $\mathrm{diam}(\mathcal{H}) = n$, by applying Theorem A.2 with $\alpha = \frac{1}{3n}$, we obtain the theorem. □

4.3 Lower bound

Recall that $\mathcal{H}$ is the Hamming graph on the set of profiles of an opinion games on a graph G. The following observation easily follows from the definition of cutwidth.

Observation 4.6. *For every path on $\mathcal{H}$ between the profile $\mathbf{0} = (0, \dots, 0)$ and the profile $\mathbf{1} = (1, \dots, 1)$ there exists a profile for which there are at least $\mathsf{CW}(G)$ discording edges.*

From now on, let us write CW as a shorthand for $\mathsf{CW}(G)$, when the reference to the graph is clear from the context. For sake of compactness, we set $\mathbf{b}(\mathbf{x}) = \sum_i (x_i - b_i)^2$. We denote as $\mathbf{b}^\star$ the minimum of $\mathbf{b}(\mathbf{x})$ over all profiles with CW discording edges.

Let $R_\mathbf{0}$ ($R_\mathbf{1}$) be the set of profiles $\mathbf{x}$ for which a path from $\mathbf{0}$ (resp., $\mathbf{1}$) to $\mathbf{x}$ exists on $\mathcal{H}$ such that every profile along the path has potential value less than $\mathbf{b}^\star + \mathsf{CW}$. To establish the lower bound we use the technical result given by Theorem A.5 which requires to compute the bottleneck ratio of a subset of profiles that is weighted at most a half by the stationary distribution. Accordingly, we set $R = R_\mathbf{0}$ if $\pi(R_\mathbf{0}) \leq 1/2$ and $R = R_\mathbf{1}$ if $\pi(R_\mathbf{1}) \leq 1/2$. (If both sets have stationary distribution less than one half, the best lower bound is achieved by setting R to $R_\mathbf{0}$ if and only if $\Phi(\mathbf{0}) \leq \Phi(\mathbf{1})$.) W.l.o.g., in the remaining of this section we assume $R = R_\mathbf{0}$.

4.3.1 For large β

Let ∂R be the set of profiles in R that have at least a neighbor $\mathbf{y}$ in the Hamming graph $\mathcal{H}$ such that $\mathbf{y} \notin R$. Moreover let $\mathcal{E}(\partial R)$ the set of edges $(\mathbf{x}, \mathbf{y})$ in $\mathcal{H}$ such that $\mathbf{x} \in \partial R$ and $\mathbf{y} \notin R$: note that $|\mathcal{E}(\partial R)| \leq n|\partial R|$. The following lemma bounds the bottleneck ratio of R.

Lemma 4.7. *For the set of profiles R defined above, we have $B(R) \leq n \cdot |\partial R| \cdot e^{-\beta(\mathsf{CW} + \mathbf{b}^\star - \mathbf{b}(\mathbf{0}))}$.*

Proof. Since $\mathbf{0} \in R$, it holds $\pi(R) \geq \pi(\mathbf{0}) = \frac{e^{-\beta \mathbf{b}(\mathbf{0})}}{Z}$. Moreover, by (4) we have

$$\begin{aligned}
Q(R, \overline{R}) &= \sum_{\substack{(\mathbf{x},\mathbf{y}) \in \mathcal{E}(\partial R):\\ \mathbf{y} = (\mathbf{x}_{-i}, y_i)}} \frac{e^{-\beta\Phi(\mathbf{x})}}{Z} \frac{e^{\beta u_i(\mathbf{y})}}{e^{\beta u_i(\mathbf{x})} + e^{\beta u_i(\mathbf{y})}} \\
&= \sum_{\substack{(\mathbf{x},\mathbf{y}) \in \mathcal{E}(\partial R):\\ \mathbf{y} = (\mathbf{x}_{-i}, y_i)}} \frac{e^{-\beta\Phi(\mathbf{x})}}{Z} \frac{e^{-\beta\Phi(\mathbf{y})} e^{\beta(u_i(\mathbf{x}) + \Phi(\mathbf{x}))}}{e^{-\beta\Phi(\mathbf{x})} e^{\beta(u_i(\mathbf{x}) + \Phi(\mathbf{x}))} + e^{-\beta\Phi(\mathbf{y})} e^{\beta(u_i(\mathbf{x}) + \Phi(\mathbf{x}))}} \\
&= \frac{1}{Z} \sum_{(\mathbf{x},\mathbf{y}) \in \mathcal{E}(\partial R)} \frac{e^{-\beta\Phi(\mathbf{x})} e^{-\beta\Phi(\mathbf{y})}}{e^{-\beta\Phi(\mathbf{x})} + e^{-\beta\Phi(\mathbf{y})}} = \frac{1}{Z} \sum_{(\mathbf{x},\mathbf{y}) \in \mathcal{E}(\partial R)} \frac{e^{-\beta\Phi(\mathbf{y})}}{1 + e^{\beta(\Phi(\mathbf{x}) - \Phi(\mathbf{y}))}} \\
&\leq \frac{1}{Z} \sum_{(\mathbf{x},\mathbf{y}) \in \mathcal{E}(\partial R)} e^{-\beta\Phi(\mathbf{y})} \leq |\mathcal{E}(\partial R)| \cdot \frac{e^{-\beta(\mathbf{b}^\star + \mathsf{CW})}}{Z}.
\end{aligned}$$

The second equality follows from the definition of potential function which implies $\Phi(\mathbf{y}) - \Phi(\mathbf{x}) = -u_i(\mathbf{y}) + u_i(\mathbf{x})$ for $\mathbf{x}$ and $\mathbf{y}$ as above; last inequality holds because if by contradiction $\Phi(\mathbf{y}) < \mathbf{b}^\star + \mathsf{CW}$ then, by definition of R, it would be $\mathbf{y} \in R$, a contradiction. □

From Lemma 4.7 and Theorem A.5 we obtain a lower bound to the mixing time of the opinion games that holds for every value of β, every social network G and every vector $(b_1, \dots, b_n)$ of internal beliefs. However, it is not clear how close this bound is to the one given in Theorem 4.4. Nevertheless, by taking $b_i = 1/2$ for each player i and β high enough, we obtain the following theorem.

Theorem 4.8. *Let $\mathcal{G}$ be an n-player opinion game on a graph G. Then, there exist a vector of internal beliefs such that for $\beta = \Omega\left(\frac{n \log n}{\mathsf{CW}}\right)$ it holds $t_{\text{mix}} \geq e^{\beta\Theta(\mathsf{CW})}$.*

Proof. If $b_i = 1/2$ for every player i, from Lemma 4.7 and Theorem A.5, since $|\partial R| \leq 2^n$ then

$$t_{\text{mix}} \geq \frac{e^{\beta \mathsf{CW}}}{n 2^n} = e^{\beta \mathsf{CW} - n \log(2n)} = e^{\beta\Theta(\mathsf{CW})}. \qquad \square$$

4.3.2 For smaller β

Theorem 4.8 gives an almost tight lower bound for high values of β for each network topology. It would be interesting to prove a matching bound also for lower values of the rationality parameter: in this section we prove such a bound for specific classes of graphs: complete bipartite graphs and cliques.

We start by considering the class of complete bipartite graphs $K_{m,m}$.

Theorem 4.9. *Let $\mathcal{G}$ be an n-player opinion game on $K_{m,m}$. Then, there exist a vector of internal beliefs such that, for every $\beta = \Omega\left(\frac{1}{m}\right)$, we have $t_{\text{mix}} \geq \frac{e^{\beta\Theta(\mathsf{CW})}}{n}$.*

To prove the theorem above, we start by evaluating the cutwidth of $K_{m,m}$: in particular, we characterize the best ordering from which the cutwidth is obtained. We will denote with A and B the two sides of the bipartite graph.

Claim 4.10. *The ordering that obtains the cutwidth in $K_{m,m}$ is the one that selects alternatively a vertex from A and a vertex from B. Moreover, the cutwidth of $K_{m,m}$ is $\lceil m^2/2 \rceil$.*

Proof. Let $(T, V \setminus T)$ be a cut of the graph, we denote with t the size of T: we also denote t_A as the number of vertices of A in T and t_B as the number of vertices of B in T. Obviously, $t = t_A + t_B$. Given t_A and t_B, the size of the cut $(T, V \setminus T)$ will be $t_A(m - t_B) + t_B(m - t_A) = mt - 2t_A(t - t_A)$. It is immediate to check that for every fixed t the cut is minimized when $\lceil t/2 \rceil$ vertices of T are taken from A and the remaining ones from B. Therefore, the cutwidth is achieved by an ordering which selects alternatively vertices from the two sides of the graph and is then given by the maximum over t of

$$\left\lceil \frac{t}{2} \right\rceil \left(m - \left\lfloor \frac{t}{2} \right\rfloor\right) + \left\lfloor \frac{t}{2} \right\rfloor \left(m - \left\lceil \frac{t}{2} \right\rceil\right).$$

The above function is equal to $mt - \frac{t^2-1}{2}$ for t odd and $m - t^2/2$ for t even. Both these fuctions are maximized for $t = m$. However, this may be impossible to achieve when for example t is odd and m is even. Nevertheless, a simple case analysis on the parity of m and t shows that the maximum is achieved for $t = m-1, m, m+1$ when m is even and for $t = m$ for m odd, resulting in a cutwidth of $\lceil m^2/2 \rceil$. □

The following lemma gives a bound to the size of ∂R for this graph.

Lemma 4.11. *For the opinion game on the graph $K_{m,m}$ with $b_i = 1/2$ for every player i, there exists a constant c_1 such that $|\partial R| \leq e^{c_1\sqrt{\mathsf{CW}}}$.*

Proof. Since $b_i = 1/2$ for every player i, we have that $\mathbf{b}(\mathbf{x}) = n/4$ for every profile $\mathbf{x}$. Therefore, by definition of R, all profiles in R (and therefore ∂R) have less then CW discording edges. Indeed, for $\mathbf{x} \in R$ we have $\mathbf{b}(\mathbf{x}) + |D(\mathbf{x})| = \Phi(\mathbf{x}) < \mathbf{b}^\star + \mathsf{CW}$. Moreover, if a profile $\mathbf{y}$ has less then $\mathsf{CW} - m$ discording edges, then $\mathbf{y}$ is not in ∂R as a state neighbor of $\mathbf{y}$ has at most $m-1$ additional discording edges.

Consequently, to bound the size of ∂R, we need to count the number of profiles in R that have potential between $\mathbf{b}^\star + \mathsf{CW} - m$ and $\mathbf{b}^\star + \mathsf{CW} - 1$ (i.e., the number of profiles with at least $\mathsf{CW} - m$ and at most $\mathsf{CW} - 1$ discording edges). To count that, we consider two sets L_0 and L_1: we start by setting $L_0 = V$ and $L_1 = \emptyset$. We take vertices from L_0 and sequentially move them to L_1. We can think of L_0 as the set of vertices with opinion 0 and L_1 as the set of vertices with opinion 1: this way we can model a path from $\mathbf{0}$ to $\mathbf{1}$ in the Hamming graph. The number $M(t)$ of edges between L_0 and L_1 after t moves is the number of discording edges in the social graph when vertices in L_0 have opinion 0 and vertices in L_1 have opinion 1. We have

$$\left\lceil \frac{t}{2} \right\rceil \left(m - \left\lfloor \frac{t}{2} \right\rfloor\right) + \left\lfloor \frac{t}{2} \right\rfloor \left(m - \left\lceil \frac{t}{2} \right\rceil\right) \leq M(t) \leq mt\,,$$

where the lower bound follows from the structural proof of minimum cuts contained in Claim 4.10.

Let t_1 be the largest integer such that for all possible ways to choose $t_1 - 1$ vertices in L_0 and move them in L_1, the number of edges between L_0 and L_1 is less than $\mathsf{CW} - m$, i.e.

$$(t_1 - 1)m < \mathsf{CW} - m \Rightarrow t_1 = \left\lfloor \frac{\mathsf{CW}}{m} \right\rfloor = \left\lfloor \frac{m}{2} \right\rfloor = \left\lfloor \frac{n}{4} \right\rfloor.$$

Let t_2 be the smallest integer such that for all possible ways to move t_2+1 vertices from L_0 to L_1, the number of edges between L_0 and L_1 is at least CW, i.e.

$$\left\lceil \frac{t_2+1}{2} \right\rceil \left(m - \left\lfloor \frac{t_2+1}{2} \right\rfloor \right) + \left\lfloor \frac{t_2+1}{2} \right\rfloor \left(m - \left\lceil \frac{t_2+1}{2} \right\rceil \right) \geq \mathsf{CW}\,,$$

that, as showed in Claim 4.10, means $t_2 = m-2$ for m even and $t_2 = m-1$ for m odd. Then, we can conclude $t_2 \leq m-1$.

By the definition of t_1, all profiles with at most $t_1 - 1$ players with opinion 1 are not in ∂R and, by definition of t_2, all profiles with at least t_2+1 players with opinion 1 are not in R. Thus, we have

$$|\partial R| \leq \sum_{i=t_1}^{t_2} \binom{n}{i} \leq \sum_{i=t_1}^{t_2} \left(\frac{n \cdot e}{i} \right)^i \leq \sum_{i=t_1}^{t_2} (5e)^i = \frac{(5e)^{t_2+1} - (5e)^{t_1}}{5e-1} \leq (5e)^{t_2+1} \leq (5e)^m \leq\ e^{3m}\,, \tag{8}$$

where in the third inequality we used the fact that $i \geq t_1 > n/5$, in the penultimate the fact that $t_2 < m$ and lastly the fact that $5^m \leq e^{2m}$ for $m \geq 0$. The lemma follows since $m \leq \sqrt{2}\sqrt{\mathsf{CW}}$. □

Proof of Theorem 4.9. If $b_i = 1/2$ for every player i, from Lemmata 4.7 and 4.11, we have

$$B(R) \leq n \cdot e^{c_1\sqrt{\mathsf{CW}}} \cdot e^{-\beta \mathsf{CW}} \leq n \cdot e^{-\beta \mathsf{CW}(1-c_2)},$$

where $c_2 = \frac{c_1\sqrt{\mathsf{CW}}}{\beta \mathsf{CW}} < 1$ since by hypothesis $\beta > \frac{c_1}{\sqrt{\mathsf{CW}}} = \Omega(1/m)$; we also notice that c_2 goes to 0 as β increases. The theorem follows from Theorem A.5. □

We remark that it is possible to prove a result similar to Theorem 4.9 also for the clique K_n: the proof follows from a simple generalization of Theorem 15.3 in [21] and by observing that the cutwidth of a clique is $\lfloor n^2/4 \rfloor$.

5 Conclusions and open problems

In this work we analyze two decentralized dynamics for binary opinion games: the best-response dynamics and the logit dynamics. As for the best-response dynamics we show that it takes time polynomial in the number of players to reach a Nash equilibrium, the latter being characterized by the existence of clusters in which players have a common opinion. On the other hand, for the logit dynamics we show polynomial convergence when the level of noise is high enough and that it increases as β grows.

It is important to highlight, as noted above, that the convergence time of the two dynamics are computed with respect to two different equilibrium concepts, namely Nash equilibrium for the best-response dynamics and logit equilibrium for the logit dynamics. This explains why the convergence times of these two dynamics asymptotically diverge even though the logit dynamics becomes similar to the best response dynamics as β goes to infinity.

Theorem 4.4 and 4.8 which prove bounds to the convergence of logit dynamics can also be read in a positive fashion. Indeed, for social networks that have a bounded cutwidth, the convergence rate of the dynamics depends only on the value of β. (We highlight that checking if a graph has bounded cutwidth can be done in polynomial time [25].) In general, we have the following picture: as long as β is less than the maximum of (roughly) $\frac{\log n}{\mathsf{CW}}$ and $\frac{1}{\Delta}$ the convergence time to the logit equilibrium is polynomial. Moreover, Theorem 4.8 shows that for β lower bounded by (roughly) $\frac{n \log n}{\mathsf{CW}}$ the convergence time to the logit equilibrium is super-polynomial. Then for some network topology, there is a gap in our knowledge which is naturally interesting to close.

In [6] the concept of metastable distributions has been introduced in order to predict the outcome of games for which the logit dynamics takes too much time to reach the stationary distribution for some value of β. It would be interesting to investigate existence and structure of such distributions for our opinion games.

We also note that our proofs for logit dynamics can be extended to the case in which the social graph is weighted. In such a setting, however, we obtain non-matching bounds: it would be interesting to develop more sophisticated techniques in order to get tight bounds.

References

[1] Daron Acemoglu and Asuman Ozdaglar. Opinion dynamics and learning in social networks. *Dynamic Games and Applications*, 1:3–49, 2011.

[2] Arash Asadpour and Amin Saberi. On the inefficiency ratio of stable equilibria in congestion games. In Stefano Leonardi, editor, *Internet and Network Economics*, volume 5929 of *Lecture Notes in Computer Science*, pages 545–552. Springer Berlin / Heidelberg, 2009.

[3] V. Auletta, D. Ferraioli, F. Pasquale, P. Penna, and G. Persiano. Convergence to equilibrium of logit dynamics for strategic games. Full version of [4], currently in preparation.

[4] Vincenzo Auletta, Diodato Ferraioli, Francesco Pasquale, Paolo Penna, and Giuseppe Persiano. Convergence to equilibrium of logit dynamics for strategic games. In *Proceedings of the 23rd ACM symposium on Parallelism in algorithms and architectures*, SPAA '11, pages 197–206. ACM, 2011.

[5] Vincenzo Auletta, Diodato Ferraioli, Francesco Pasquale, and Giuseppe Persiano. Mixing time and stationary expected social welfare of logit dynamics. In Spyros Kontogiannis, Elias Koutsoupias, and PaulG. Spirakis, editors, *Algorithmic Game Theory*, volume 6386 of *Lecture Notes in Computer Science*, pages 54–65. Springer Berlin Heidelberg, 2010.

[6] Vincenzo Auletta, Diodato Ferraioli, Francesco Pasquale, and Giuseppe Persiano. Metastability of logit dynamics for coordination games. In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '12, pages 1006–1024. SIAM, 2012.

[7] Maria-Florina Balcan, Avrim Blum, and Yishay Mansour. Improved equilibria via public service advertising. In *Proceedings of the twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '09, pages 728–737. Society for Industrial and Applied Mathematics, 2009.

[8] Noam Berger, Claire Kenyon, Elchanan Mossel, and Yuval Peres. Glauber dynamics on trees and hyperbolic graphs. *Probability Theory and Related Fields*, 131:311–340, 2005.

[9] Anand Bhalgat, Tanmoy Chakraborty, and Sanjeev Khanna. Approximating pure nash equilibrium in cut, party affiliation, and satisfiability games. In *Proceedings of the 11th ACM conference on Electronic commerce*, EC '10, pages 73–82. ACM, 2010.

[10] D. Bindel, J. Kleinberg, and S. Oren. How bad is forming your own opinion? In *Foundations of Computer Science (FOCS), 2011 IEEE 52nd Annual Symposium on*, pages 57 –66, oct. 2011.

[11] Lawrence E. Blume. The statistical mechanics of strategic interaction. *Games and Economic Behavior*, 5(3):387 – 424, 1993.

[12] R. Bubley and M. Dyer. Path coupling: A technique for proving rapid mixing in markov chains. volume 0, page 223, Los Alamitos, CA, USA, 1997. IEEE Computer Society.

[13] Morris H. DeGroot. Reaching a consensus. *Journal of the American Statistical Association*, 69(345):118–121, 1974.

[14] Peter M. DeMarzo, Dimitri Vayanos, and Jeffrey Zwiebel. Persuasion bias, social influence, and unidimensional opinions. *The Quarterly Journal of Economics*, 118(3):909–968, 2003.

[15] Martin Dyer and Velumailum Mohanaraj. Pairwise-interaction games. In Luca Aceto, Monika Henzinger, and Jirí Sgall, editors, *Automata, Languages and Programming*, volume 6755 of *Lecture Notes in Computer Science*, pages 159–170. Springer Berlin / Heidelberg, 2011.

[16] Glenn Ellison. Learning, local interaction, and coordination. *Econometrica*, 61(5):1047–71, September 1993.

[17] Alex Fabrikant, Christos Papadimitriou, and Kunal Talwar. The complexity of pure nash equilibria. In *Proceedings of the thirty-sixth annual ACM symposium on Theory of computing*, STOC '04, pages 604–612, New York, NY, USA, 2004. ACM.

[18] Noah E. Friedkin and Eugene C. Johnsen. Social influence and opinions. *The Journal of Mathematical Sociology*, 15(3-4):193–206, 1990.

[19] Benjamin Golub and Matthew O. Jackson. Naïve learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, 2(1):112–49, February 2010.

[20] Matthew O. Jackson. *Social and Economic Networks*. Princeton University Press, Princeton, NJ, USA, 2008.

[21] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, 2006.

[22] Dov Monderer and Lloyd S. Shapley. Potential games. *Games and Economic Behavior*, 14(1):124 – 143, 1996.

[23] A. Montanari and A. Saberi. Convergence to equilibrium in local interaction games. In *Foundations of Computer Science, 2009. FOCS '09. 50th Annual IEEE Symposium on*, pages 303 –312, oct. 2009.

[24] RobertW. Rosenthal. A class of games possessing pure-strategy nash equilibria. *International Journal of Game Theory*, 2:65–67, 1973.

[25] Dimitrios Thilikos, Maria Serna, and Hans Bodlaender. Constructive linear time algorithms for small cutwidth and carving-width. In Gerhard Goos, Juris Hartmanis, Jan van Leeuwen, D. Lee, and Shang-Hua Teng, editors, *Algorithms and Computation*, volume 1969 of *Lecture Notes in Computer Science*, pages 192–203. Springer Berlin / Heidelberg, 2000.

[26] H. Peyton Young. The diffusion of innovations in social networks. In By L. Blume and Steven N. Durlauf, editors, *Economy as an evolving complex system*, volume 3 of *Proceedings volume in the Santa Fe Institute studies in the sciences of complexity Santa Fe Institute Studies on the Sciences of Complexity*, pages 267–282. Oxford University Press US, 2006.

A Review of mixing time techniques

A.1 Markov chain coupling

When the two coupled chains $(X_t, Y_t)_{t=0}^{\infty}$ start at $(X_0, Y_0) = (\mathbf{x}, \mathbf{y})$, we write $\mathbf{P}_{\mathbf{x},\mathbf{y}}(\cdot)$ and $\mathbf{E}_{\mathbf{x},\mathbf{y}}[\cdot]$ for the probability and the expectation on the space $S \times S$ where the two coupled chains are both defined. We denote by τ_{couple} the first time the two chains meet; that is,

$$\tau_{\text{couple}} = \min\{t \colon X_t = Y_t\} .$$

Often couplings of Markov chains with the property that for $s \geq \tau_{\text{couple}}$, $X_s = Y_s$ are considered. The following theorem, which follows from Proposition 4.7 and Theorem 5.2 in [21] establishes the importance of this tool.

Theorem A.1 (Coupling)**.** *Let $\mathcal{M}$ be a Markov chain with finite state space S and transition matrix P. For each pair of states $\mathbf{x}, \mathbf{y} \in S$ consider a coupling (X_t, Y_t) of $\mathcal{M}$ with starting states $X_0 = \mathbf{x}$ and $Y_0 = \mathbf{y}$. Then*

$$\left\| P^t(\mathbf{x}, \cdot) - P^t(\mathbf{y}, \cdot) \right\|_{\mathrm{TV}} \leq \mathbf{P}_{\mathbf{x},\mathbf{y}}\left(X_t \neq Y_t\right) = \mathbf{P}_{\mathbf{x},\mathbf{y}}\left(\tau_{\text{couple}} > t\right) .$$

Sometimes it is difficult to specify a coupling and to analyze the coupling time τ_{couple} for each pair of starting states $\mathbf{x}$ and $\mathbf{y}$. The *path coupling* theorem says that it is sufficient to define a coupling only for pairs of Markov chains starting from *adjacent* states and an upper bound on the mixing time can be obtained if each of these couplings contracts their distance on average. More precisely, consider a Markov chain $\mathcal{M}$ with state space S and transition matrix P; recall that $\mathcal{H} = (S, \mathsf{E})$ is the Hamming graph and let $w : \mathsf{E} \to \mathbb{R}$ be a function assigning weights to the edges such that $w(e) \geq 1$ for every edge $e \in \mathsf{E}$; for $\mathbf{x}, \mathbf{y} \in S$, we denote by $\rho(\mathbf{x}, \mathbf{y})$ the weight of the shortest path in $\mathcal{H}$ between $\mathbf{x}$ and $\mathbf{y}$. The following theorem holds.

Theorem A.2 (Path Coupling [12])**.** *Suppose that for every edge $(\mathbf{x}, \mathbf{y}) \in \mathsf{E}$ a coupling (X_t, Y_t) of $\mathcal{M}$ with $X_0 = \mathbf{x}$ and $Y_0 = \mathbf{y}$ exists such that $\mathbf{E}_{\mathbf{x},\mathbf{y}}\left[\rho(X_1, Y_1)\right] \leq e^{-\alpha} \cdot w(\mathbf{x}, \mathbf{y})$ for some $\alpha > 0$. Then*

$$t_{\text{mix}}(\varepsilon) \leq \frac{\log(\text{diam}(\mathcal{H})) + \log(1/\varepsilon)}{\alpha}$$

where $\text{diam}(\mathcal{H})$ *is the (weighted) diameter of $\mathcal{H}$.*

A.2 Relaxation time and spectral techniques

The relaxation time is related to the mixing time by the following theorem (see, for example, Theorems 12.3 and 12.4 in [21]).

Theorem A.3 (Relaxation time)**.** *Let P be the transition matrix of a reversible, irreducible, and aperiodic Markov chain with state space S and stationary distribution π. Then*

$$(t_{\text{rel}} - 1)\log 2 \leq t_{\text{mix}} \leq \log\left(\frac{4}{\pi_{\min}}\right) t_{\text{rel}} ,$$

where $\pi_{\min} = \min_{\mathbf{x} \in S} \pi(\mathbf{x})$.

Bounds on relaxation time can be obtained by using the following lemma (see Corollary 13.24 in [21]).

Lemma A.4. *Let P the transition matrix of an irreducible, aperiodic and reversible Markov chain with state space S and stationary distribution π. Consider the graph $G = (S, E)$, where $E = \{(\mathbf{x}, \mathbf{y}) \colon P(\mathbf{x}, \mathbf{y}) > 0\}$, and to every pair of states $\mathbf{x}, \mathbf{y} \in S$ we assign a path $\Gamma_{\mathbf{x},\mathbf{y}}$ from $\mathbf{x}$ to $\mathbf{y}$ in G. We define*

$$\rho = \max_{e = (\mathbf{z}, \mathbf{w}) \in E} \frac{1}{Q(e)} \sum_{\substack{\mathbf{x}, \mathbf{y} \colon \\ e \in \Gamma_{\mathbf{x},\mathbf{y}}}} \pi(\mathbf{x})\pi(\mathbf{y})|\Gamma_{\mathbf{x},\mathbf{y}}| .$$

Then $\frac{1}{1-\lambda_2} \leq \rho$.

A.3 Bottleneck ratio bounds

We use the following theorem to derive lower bounds to the mixing time (see, for example, Theorem 7.3 in [21]).

Theorem A.5 (Bottleneck ratio)**.** *Let $\mathcal{M} = \{X_t \colon t \in \mathbb{N}\}$ be an irreducible and aperiodic Markov chain with finite state space S, transition matrix P, and stationary distribution π. Then the mixing time is*

$$t_{\text{mix}} \geq \max_{L \colon \pi(L) \leq 1/2} \frac{1}{4B(L)} \,.$$