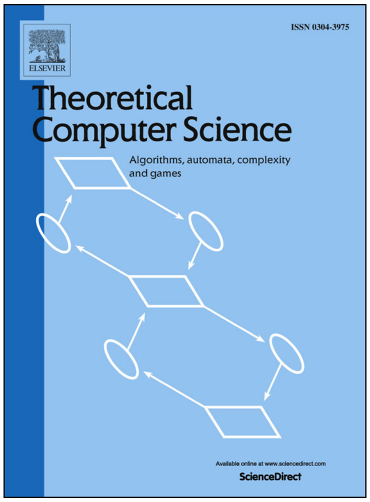
Social pressure in opinion dynamics

Diodato Ferraioli, Carmine Ventre



PII: S0304-3975(19)30452-9
DOI: https://doi.org/10.1016/j.tcs.2019.07.017
Reference: TCS 12102

To appear in: *Theoretical Computer Science*

Received date: 28 February 2018
Revised date: 18 March 2019
Accepted date: 15 July 2019

Please cite this article in press as: D. Ferraioli, C. Ventre, Social pressure in opinion dynamics, *Theoret. Comput. Sci.* (2019), https://doi.org/10.1016/j.tcs.2019.07.017

Social Pressure in Opinion Dynamics

Diodato Ferraioli
University of Salerno, Italy
dferraioli@unisa.it

Carmine Ventre
University of Essex, UK
c.ventre@essex.ac.uk

Abstract

Motivated by privacy and security concerns in online social networks, we study the role of social pressure in opinion dynamics. These are dynamics, introduced in economics and sociology literature, that model the formation of opinions in a social network. We enrich one of the most classical opinion dynamics, by introducing the pressure, increasing with time, to reach an agreement.

We prove that for clique social networks, the dynamics always converges to consensus (no matter the level of noise) if the social pressure is high enough. Moreover, we provide (tight) bounds on the speed of convergence; these bounds are polynomial in the number of nodes in the network provided that the pressure grows sufficiently fast. We finally look beyond cliques: we characterize the graphs for which consensus is guaranteed, and make some considerations on the computational complexity of checking whether a graph satisfies such a condition.

Keywords: Opinion Dynamics, Best Response Dynamics, Logit Dynamics

1 Introduction

Opinion dynamics focus on self-interested individuals, each with an *opinion* and all connected in some social network, in need of reaching a decision in a decentralized way (i.e., without a central authority dictating their actions). For example, they might be sitting on some hiring panel, members having their own favorite candidates and a job offer to be made, or they might be co-authors/reviewers deciding about the submission/notification of a paper.

This subject received large attention in economics and sociology literature. In particular, DeGroot [1974] defined the most prominent model for this setting, where the role of discussions in the decision making process is mathematically captured by individuals repeatedly averaging their own opinion with those of their neighbors. Friedkin and Johnsen [1990] considered a variant, in which each individual additionally maintains a persistent *internal belief*, which remains constant even as they update their opinions through averaging. These models have attracted much attention in recent literature, see e.g., [Bhalgat et al., 2010, Ferraioli et al., 2016, Chierichetti et al., 2018]. This line of work identifies the absence of consensus in many real-life situations and gives a game-theoretic explanation: individuals will not compromise any further when this increases their *cost*, defined as a measure of the distance between an individual's opinion and (i) her own belief; (ii) the opinions of her neighbors on the social network.

We here note, however, that in many cases – such as the examples mentioned above – there is a pressure to reach a consensus. Such a pressure augments as time (e.g., length of the meeting, the approaching deadline) goes on. Under which conditions does this pressure facilitate consensus? How long does it take to reach the consensus, if any?

As noted by Rajtmajer et al. [2016], these questions bear a certain degree of importance for security and privacy in Online Social Networks (OSNs) and, more generally, for distributed (multi-)access control policies. Access control to contents shared on OSNs is a central research topic in security and privacy. Typically, the uploader gets to decide the degree of access (e.g., "friends"; "friends of friends"; etc.) of a shared content (e.g., a picture of a group of people). However, the sensitivity to privacy issues of the uploader might differ from that of the others interested in the shared content (e.g., the others in the picture). A distributed protocol could involve all these individuals with the objective to reach a consensus on the accessibility of the content, in such a way to cater to the needs of everyone. The social pressure might be enforced, for example, by only publishing on the OSN upon consensus. Rajtmajer et al. [2016] introduce a number of important themes related to this applicative scenario, by defining a game-theoretic model reminiscent of the classical opinion dynamics by DeGroot [1974] and Friedkin and Johnsen [1990] (we refer to their paper for more details). However, as we discuss below, they fail to give (complete) answers to our questions of interest.

1.1 Our Contribution

In this work we formally analyze the process of opinion diffusion in environments with a time-increasing pressure to reach consensus, such as the ones described above, by extending and improving the contribution of Rajtmajer et al. [2016].

To this aim, we first introduce new dynamical models for the diffusion of opinion. Specifically, we follow Bhalgat et al. [2010] and Ferraioli et al. [2016], and we analyze the opinion dynamics of DeGroot [1974] and Friedkin and Johnsen [1990] within a game theoretic framework. However, we generalize the definitions of (noisy) best-response dynamics given by Bhalgat et al. [2010] and Ferraioli et al. [2016] along two main dimensions. Firstly, as detailed above, we introduce a non-decreasing pressure to coordinate opinions with the neighbors. Secondly, we do not restrict the way the individuals weigh disagreements. Related literature makes restrictive assumptions on the cost of disagreeing (either with a neighbor or with one's own belief). We here instead use any pair of functions, f and g, that measure the cost of having an opinion different from the belief and a neighbor's opinion, respectively.

In the case in which the underlying social network is a clique (as in many real world settings, e.g., the hiring committee example mentioned above), we prove bounds on the rate of convergence to consensus of these dynamics. We are able to closely describe the behavior of best-response dynamics: once individuals start to deviate from the initial opinion profile they move onto the opinion adopted by the majority; this way the initial majority keeps increasing until consensus. Incidentally, this also proves that each individual moves only once and, therefore, the rate of convergence – once deviations commence – is polynomial in the number of individuals. We determine the minimum level of social pressure needed to kick off deviations (as the ratio between the cost of having an opinion different from the belief and the cost of disagreeing with a neighbor) and give an instance showing that our bound on convergence is tight.

In many cases, it might not be realistic to assume that the individuals are always able to choose their best response, and they only have *bounded rationality.* Indeed, individuals with bounded rationality have been object of study of previous literature on opinion dynamics [Ferraioli et al., 2016], and they have been advocated also in the specific setting with external pressure analyzed in this work [Rajtmajer et al., 2016].

To model the bounded rational behavior of individuals, several *noisy* best-response update rules have been introduced. In this work we will focus on one of the most prominent, namely *logit update rule* [Blume, 1993], according to which the individual selected for update can adopt every opinion with a probability that is proportional to her advantage in adopting that opinion and to the rationality level $\beta > 0$. We prove that this dynamics always converges to consensus in at most n^3 steps, n being the number of individuals, as long as the social pressure is above a given threshold (that depends not only on the ratio between disagreements costs, but also on the rationality level β). This result is achieved by reducing the dynamics to a birth and death chain and evaluating the time the latter takes to reach the consensus. We stress that these are the first analytical results about the behavior of noisy dynamics in this setting.

For both dynamics, our results can be extended in several ways to accommodate more general settings (e.g., better rather than best responses; simultaneous rather than sequential moves; alternative noisy update rules; etc.).

We then move onto general social networks and characterize the stationary points (different from consensus) of best-response dynamics in terms of the existence of a certain cut of the graph. For a graph to admit an equilibrium in which individuals' opinions disagree there must exist a partition of vertices in (at least) two sets such that each vertex has more neighbors in its side of the partition than in the other(s). Can we recognize in polynomial-time whether a graph will guarantee consensus (i.e., will not admit such a cut)? We connect this question to the existence of certain locally-minimum cuts of the graph. While this has been proved to be an NP-complete problem in a recent follow-up paper [Auletta et al., 2018], we here show how to efficiently recognize (a subset of the) graphs for which consensus might not occur. Finally, we briefly focus on what happens beyond consensus: we analyze the "price of divergence" and bound how far from consensus the stationary points of best-response dynamics can be in these graphs; moreover, we show that our analysis can be immediately extended to characterize the topologies in which pressure can be ineffective not only to lead to consensus, but also to lead to a majority of individuals supporting a certain opinion.

We finally provide a case study based on recent negotiations between the EU and the UK about Brexit. We note how the observations developed in our theoretical analysis serve well to describe the current state of negotiations and suggest possible ways in which the current state can evolve.

1.2 Related Works

Understanding how opinions are formed and expressed in a social context has been object of extensive recent study, in AI and multiagent systems [Pryymak et al., 2012, Tsang and Larson, 2014, Grandi et al., 2015, Schwind et al., 2015], computer science at large [Acemoglu and Ozdaglar, 2011, Bindel et al., 2011, Mossel and Tamuz, 2017, Auletta et al., 2015], as well as, sociology, economics, physics, and epidemiology. In particular, the work of Friedkin and Johnsen [1990], which represents our starting point, has been largely studied recently and has then emerged as the principal model in the area. For example, Bindel et al. [2011] considered this model and proved that, under mild assumptions, whenever beliefs and opinions belong to $[0, 1]$, the repeated averaging process leads to a *unique* equilibrium, describing the opinion that each individual eventually expresses. Chierichetti et al. [2018] considered the case that opinions are discrete and bound the price of stability and price of anarchy of the games corresponding to this dynamics. Extensions of this model have been proposed by Bhawalkar et al. [2013], by Auletta et al. [2016] and by Bilò et al. [2016]. No of these works considers the effects on individuals of time-varying social pressure, hence their results are not comparable to ours.

A couple of works turn out to be closely related to this paper: Ferraioli et al. [2016] focus on the rate of convergence of opinion dynamics to equilibria under both best-response and logit dynamics. However, their model is simpler than ours (e.g., they consider binary opinion and fixed cost for disagreements) and does not consider time pressure: thus extending their results to our setting is not straightforward. A follow-up work by Auletta et al. [2019] provides conditions for convergence to consensus for a generalization of the model of Friedkin and Johnsen, that allows, among others, also to model time pressure. Nevertheless, due to the generality of the model, the conditions given therein are very abstract, and it is not trivial to translate them in the simple conditions for convergence to consensus given in our work. We also remark that Auletta et al. [2019] do not give any bound on the time the dynamics takes to converge to consensus. Moreover, they do not consider bounded rationality.

Logit dynamics have been introduced by Blume [1993] to model individuals with bounded rationality. The dynamics are founded on well-established measure theory tools for modeling limited cognitive capacities [McFadden, 1974], and has already been adopted for modeling the diffusion of information and opinions on social networks [Peyton Young, 2006, Montanari and Saberi, 2009, Ferraioli et al., 2016, Auletta et al., 2013b,a]. We stress that all these works consider a model of diffusion that is extremely simpler than ours, and, in particular, they do not consider the presence of increasing time pressure. Hence, their results do not naturally extend to our setting. Still, as we will see below, techniques defined in these works turned out to be useful for analyzing the behavior of the logit dynamics in our model.

It is worthy to compare our results with those given by Rajtmajer et al. [2016] as the models are very similar in spirit. There are a number of notable differences. Firstly, they consider simultaneous moves only, while we are able to deal as well with the arguably more realistic case of asynchronous moves, in which only some subset of individuals simultaneously update their opinions at each time step. Secondly, their analytical results only apply to cliques and a unique continuous set from which individuals choose beliefs and opinions; our results, instead, hold no matter which pair of sets individuals use for beliefs/opinions and the level of granularity of these sets (incidentally, discrete sets seem to fit better the application of opinion dynamics to OSNs). Thirdly, their convergence result requires an infinitely big level of social pressure while no guarantee on the speed of convergence is given; instead, we do give (tight) guarantees on convergence rate with “reasonable” values of pressure. Finally, they only provide experimental results for individuals with bounded rationality.

2 The Model

Let $G = (V, E)$ be a connected undirected graph with $|V| = n$. Every vertex of the graph represents an individual. Each individual i has a *belief* $b_i \in B$ (e.g., her preferred privacy setting in the OSN scenario) and can choose a (potentially) different *opinion* $x_i \in S$ (e.g., privacy setting in the OSN application). We do not make any assumption on the structure of beliefs and opinions. In particular, these sets can be either continuous or discrete. Moreover, it may be the case that $B \subseteq S$, or $S \subseteq B$, or that none of these relations hold. For every pair $x, y \in \bigcup_i (S \cup B)$, we will denote as $\operatorname{dist}(x, y) \in [0, 1]$ their *distance*. Multiple choices of distance are available: for example, if beliefs and opinions can be embedded in a metric space (e.g., they are values on a line), as it is the case in most of the previous literature on the topic, then $\operatorname{dist}(x, y)$ corresponds to the distance among these points in the metric space. Here, we instead consider a different concept of distance, known as *drastic distance* in literature about belief merging [Pigozzi, 2016]: for every $x, y \in \bigcup_i S$, we assume that $\operatorname{dist}(x, y) = 1$ if $x \neq y$, and $\operatorname{dist}(x, y) = 0$, otherwise. Note that we still allow the distance between the opinion and the belief of an individual to assume any value in $[0, 1]$. Note that our choice of distance fits particularly well our application scenarios. Indeed, our “binary” definition serves well our focus on consensus since it is not really relevant how “much” an individual is disagreeing with someone else but only that they disagree. Moreover, this definition allows us to consider even beliefs and opinions that cannot be embedded in a metric space (e.g., when beliefs and opinion are boolean formulas), but they are barely distinguishable. In this case, it would be impossible to evaluate the distance among opinions, but we can still assume that $\operatorname{dist}(x, y) = 1$ if $x \neq y$ and

```
Let $x_i^{(0)} = b_i$ for all $i$
Let $k = 0$
while $\exists\, (j, l) \in E$ s.t. $x_l^{(k)} \neq x_j^{(k)}$ do
    if $\exists i$ such that $x_i^{(k)} \notin BR_i(\mathbf{x}_{-i}^{(k)})$ then
        Let $i$ be such an individual
        Choose $x_i^{(k+1)} \in BR_i(\mathbf{x}_{-i}^{(k)})$
        Set $x_j^{(k+1)} = x_j^{(k)}$ for $j \neq i$
    else Set $x_i^{(k+1)} = x_i^{(k)}$ for all $i$
    Increment $k$ by 1
```

Algorithm 1: The dynamics

$\text{dist}(x, y) = 0$ otherwise. We leave the study of more nuanced definitions to future work.

According to well-established works in sociology and economics [DeGroot, 1974, Friedkin and Johnsen, 1990], the cost of individual i in an opinion profile $\mathbf{x} \in S^n$ depends only on its own belief and on the opinions currently adopted by individuals on the networks (and thus it does not depend on the private belief of other agents, or on the history of adopted opinions), as follows:

$$c_i(\mathbf{x}) = f_i(\text{dist}(x_i, b_i)) + \rho \cdot \sum_{(i,j)\in E} g_i(\text{dist}(x_i, x_j)),$$

where $\rho > 0$ represents the *social pressure* to reach consensus, $f_i \colon [0, 1] \to \mathbb{R}_{\geq 0}$ is a non-decreasing function, and $g_i \colon \{0, 1\} \to \mathbb{R}_{\geq 0}$ is a function such that $g_i(1) > g_i(0)$[1]. Here, f_i and g_i measure the effects to individual i of disagreeing with her own beliefs and a neighbor's opinion, respectively. Usually f_i and g_i are set to be weighted linear functions, i.e. $f_i(x) = w_i \cdot x$ for some $w_i > 0$, or weighted quadratic functions, i.e., $f_i(x) = w_i \cdot x^2$. Here, we keep them as general as possible; our results turn out to hold regardless of the specific choice of f_i and g_i. Nevertheless, for sake of presentation we assume that $f_i(0) = g_i(0) = 0$ for every i. We highlight that our results hold even if opinion/belief distances can be larger than 1 or $f_i(0), g_i(0) \neq 0$ by simply rescaling the functions f_i, g_i.

To simplify the exposition and present the main ideas, we will henceforth assume that $f_i = f$ and $g_i = g$ for every i. Moreover, we assume that B and S are finite; however, we stress that our results can be generalized to opinion/beliefs sets of infinite cardinality and to individual-specific functions.

Let $BR_i(\mathbf{x}_{-i}) = \arg\min_{x\in S} c_i(x, \mathbf{x}_{-i})$ for every $\mathbf{x}_{-i} \in S^{n-1}$. The *best-response consensus dynamics* is defined in Algorithm 1 under the assumption that the social pressure is non-decreasing in k, i.e., $0 < \rho^{(0)} \leq \rho^{(1)} \leq \rho^{(2)} \leq \ldots$, with $\lim_{k\to\infty} \rho^{(k)} \geq \rho^*$. Observe that when the algorithm terminates, then all individuals agree on the same opinion; when this occurs, we say that the corresponding opinion profile is a *consensus*. When the algorithm terminates, we will say that the algorithm (or the dynamics) *converges* (to consensus).

As discussed above, one novelty of our definition is in the role of the ρ's which, ultimately, resides on the value of ρ^*. In fact, ρ^* needs to be big enough to incentivize consensus, for otherwise, we either have a situation wherein no individual moves from her belief (when ρ^* is too small) or fall back to a generalized notion of opinion dynamics related to previous studies (when ρ^* is not big enough).

We are also interested in a noisy version of the dynamics of Algorithm 1 wherein individuals' responses are perturbed by some noise. Specifically, we look at *logit dynamics*, according to

[1] If $g_i(1) = g_i(0)$ then there would be no incentive to coordinate, i.e., no individual will move away from her b_i, no matter the social pressure to reach a consensus.

which individual i adopts opinion x_i as a response to $\mathbf{x}_{-i}$ with a probability proportional to $e^{-\beta c_i(x_i, \mathbf{x}_{-i})}$, with $\beta > 0$ (more details can be found in Sect. 3.2).

3 Clique Social Networks

3.1 Warm-up: Best Response Dynamics

Let us start by analyzing the behavior of the best-response dynamics on clique social networks. Whereas the results in this setting can appear quite straightforward, we stress that the bound we provide on the convergence time is tight. Moreover, the result is achieved by completely characterizing the behavior of the dynamics. This characterization can be very instructive, and may help in having a better understanding of the more complex cases studied in the paper, namely noisy dynamics and non-clique networks.

Theorem 1. *If G is a clique and $\rho^* > \frac{f(1)}{g(1)}$, then Algorithm 1 converges to consensus in time at most $n - 1 + k^*$, where*

$$k^* = \min \left\{ k \mid \rho^{(k)} > \frac{f(1)}{g(1)} \right\}.$$

Moreover, there is an instance in which the dynamics takes exactly $n - 1 + k^*$ steps to converge to consensus.

Proof. Given $\mathbf{x}^{(k)}$, the opinion profile at round k of the dynamics, we let $n_s^{(k)}$ denote the number of individuals adopting opinion $s \in S$ in $\mathbf{x}^{(k)}$.

Claim 1.1. *If $\rho^{(k)} > \frac{f(1)}{g(1)}$ and $\mathbf{x}^{(k)}$ is not a consensus then $\mathbf{x}^{(k+1)} \neq \mathbf{x}^{(k)}$.*

Proof. Since $\mathbf{x}^{(k)}$ is not a consensus, $n_s^{(k)} > 0$ for at least two opinions. Let x_+ be the opinion adopted by most individuals in $\mathbf{x}^{(k)}$ and x_- be the opinion adopted by least (non-zero) number of individuals in $\mathbf{x}^{(k)}$. For individual i adopting x_- we have

$$c_i(\mathbf{x}^{(k)}) - c_i(x_+, \mathbf{x}_{-i}^{(k)}) > -f(1) + \frac{f(1)}{g(1)} g(1) = 0,$$

where the inequality follows from $n_{x_+}^{(k)} \geq n_{x_-}^{(k)}$ and the lower bound to $\rho^{(k)}$. The test in Line 4 of the dynamics will be then true (as at least i will satisfy it) and thus $\mathbf{x}^{(k+1)} \neq \mathbf{x}^{(k)}$. □

By definition, k^* is the first round in which $\rho^{(k^*)} > \frac{f(1)}{g(1)}$. Clearly, if $\mathbf{x}^{(k^*)}$ is a consensus, we are done as the dynamics has converged in less that k^* steps. Thus we assume that $\mathbf{x}^{(k^*)}$ is not a consensus and, by Claim 1.1, conclude that $\mathbf{x}^{(k^*+1)} \neq \mathbf{x}^{(k^*)}$. Then, we set $s_{\max} = \arg\max_{s \in S} |\{i \in V : x_i^{(k^*+1)} = s\}|$, i.e., $s_{\max}$ denotes the[2] opinion adopted in $\mathbf{x}^{(k^*+1)}$ by the majority of individuals. Moreover, let $d(\mathbf{x}) = |\{i \in V \mid x_i \neq s_{\max}\}|$, i.e., $d(\mathbf{x})$ is the number of individuals with an opinion different from $s_{\max}$ in $\mathbf{x}$. Note that, in general, $d(\mathbf{x}) \leq n - 1$, while for a consensus profile $\mathbf{x}$, $d(\mathbf{x}) = 0$.

Claim 1.2. *For every $k \geq k^*$, if $\mathbf{x}^{(k+1)} \neq \mathbf{x}^{(k)}$ then $d(\mathbf{x}^{(k+1)}) = d(\mathbf{x}^{(k)}) - 1$.*

[2] In the proof of Claim 1.2, we prove that this opinion is unique.

Proof. For every $k \geq k^*$, if $\mathbf{x}^{(k+1)} \neq \mathbf{x}^{(k)}$, let i be the individual that deviates from $x_i = x_i^{(k)}$ to $y_i = x_i^{(k+1)}$. Clearly, $x_i \notin BR_i(\mathbf{x}_{-i}^{(k)})$ while $y_i \in BR_i(\mathbf{x}_{-i}^{(k)})$. Since i deviates in round $k+1$, we then have that

$$f(\mathrm{dist}(x_i, b_i)) + \rho^{(k)} g(1) \left(\sum_{s \neq x_i, y_i} n_s^{(k)} + n_{y_i}^{(k)} \right) = c_i(x_i, \mathbf{x}_{-i}^{(k)})$$

$$> c_i(y_i, \mathbf{x}_{-i}^{(k)}) = f(\mathrm{dist}(y_i, b_i)) + \rho^{(k)} g(1) \left(\sum_{s \neq x_i, y_i} n_s^{(k)} + (n_{x_i}^{(k)} - 1) \right).$$

Then

$$f(\mathrm{dist}(x_i, b_i)) - f(\mathrm{dist}(y_i, b_i)) + \rho^{(k)} g(1)(n_{y_i}^{(k)} - n_{x_i}^{(k)} + 1) \tag{1}$$

is positive.

We now prove a general result on the behavior of the dynamics, that implies the claim. Namely, we show that if $k \geq k^*$, then when individuals deviate, they move onto $s_{\max}$ and will never switch to a different opinion afterwards. To this aim, let K' be the set of rounds $k' \geq k^*$ in which a individual changes her opinion. We prove that there is a unique opinion adopted by the majority of individuals in $\mathbf{x}^{(k^*+1)}$ and, by induction on $k' \in K'$, that $n_{s_{\max}}^{k'}$ is increasing in k' and $n_s^{k'}$ is decreasing in k' for every $s \neq s_{\max}$.

The base case is $k' = k^*$. Let j denote the individual switching from $x_j = x_j^{(k^*)}$ to $y_j = x_j^{(k^*+1)}$ in round k^*. We want to prove that $y_j = s_{\max}$, i.e., y_j is the unique opinion adopted by more individuals in $\mathbf{x}^{(k^*+1)}$. Recall that y_j is the best response, and thus $y_j = \arg\min_{s \neq x_j \in S} c_i(s, \mathbf{x}_{-j}^{(k^*)})$. Since, for every $s \in S$ it holds that $\mathrm{dist}(s, b_j) \leq 1$ and $\rho^{(k^*)} g(1) > f(1)$, then we have $n_{y_j}^{(k^*)} = \max_{s \neq x_j \in S} n_s^{(k^*)}$. Moreover, since the difference between the f's in (1) is at most $f(1)$, $\rho^{(k^*)} g(1) > f(1)$, and $n_s^{(k^*)}$ is an integer for every $s \in S$, then (1) is satisfied if and only if then $n_{y_j}^{(k^*)} > n_{x_j}^{(k^*)} - 1$ which in turns implies, by the fact that $n_{y_j}^{(k^*)}$ and $n_{x_j}^{(k^*)}$ are integers, $n_{y_j}^{(k^*)} \geq n_{x_j}^{(k^*)}$. Hence, for every $s \neq y_j$

$$n_{y_j}^{(k^*+1)} > n_{y_j}^{(k^*)} \geq n_s^{(k^*)} \geq n_s^{(k^*+1)},$$

thus showing that y_j is the unique opinion for which we have that $n_{y_j}^{(k^*+1)} = \max_{s \in S} n_s^{(k^*+1)}$.

Assume now that the claim is true for $k'-1$; we prove it for k'. Let ℓ be the individual moving at round k'. By inductive hypothesis, $s_{\max}$ is the one opinion adopted by more individuals in $\mathbf{x}^{(k'-1)}$. So if $x_\ell^{(k'-1)} \neq s_{\max}$, by the same argument of the base case, ℓ will switch to $s_{\max}$. Moreover, no ℓ such that $x_\ell^{(k'-1)} = s_{\max}$ will deviate from $s_{\max}$. In fact, as noted above, this would require ℓ to move to a opinion adopted by at least as many individuals adopting $s_{\max}$ – this is impossible. □

It is not hard to see that the two claims above yield the upper bound. Observe that the result holds no matter which individual is selected at Line 5 and which best response is selected at Line 6 of the dynamics (if multiple choices are available). However, this choice influences $s_{\max}$, and, hence, the opinion on which the individuals converge.

For the lower bound, we need to prove the following claim.

Claim 1.3. *Suppose that for all individuals i and opinions $s \in S$ we have that $c_i(x_i^{(0)}, \mathbf{x}_{-i}^{(0)}) - c_i(s, \mathbf{x}_{-i}^{(0)}) \leq 0$ and $\mathbf{x}^{(0)}$ is not a consensus. Then $\mathbf{x}^{(k+1)} = \mathbf{x}^{(0)}$ for every k such that $\rho^{(k)} \leq \frac{f(1)}{\delta g(1)}$, where $\delta = \max_{s\colon n_s^{(0)} > 0} \{n_{s'}^{(0)} - n_s^{(0)} + 1 \mid s' \in S\}$.*

Proof. Suppose that $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)}$ and $\mathbf{x}^{(k)}$ is not a consensus. Then for all individuals i and $s \in S$, $c_i(x_i^{(k)}, \mathbf{x}_{-i}^{(k)}) - c_i(s, \mathbf{x}_{-i}^{(k)}) \le 0$. This yields

$$\rho^{(k)} g(1)(n_s^{(k)} - n_{x_i^{(k)}}^{(k)} + 1) \le f(\text{dist}(s, b_i)) - f(\text{dist}(x_i^{(k)}, b_i)).$$

Note that the RHS of this inequality cannot be $-f(1)$. Indeed, if this is the case, since $\rho^{(k)}$ is non-negative and $g(1) > 0$, the inequality would be violated for $n_s^{(k)} \ge n_{x_i^{(k)}}^{(k)}$ – this always occurs since $\mathbf{x}^{(k)}$ is not a consensus. Therefore, we can safely substitute the RHS of the inequality with $f(1)$. Then, we have that $\rho^{(k)} \le \frac{f(1)}{g(1)(n_s^{(k)} - n_{x_i^{(k)}}^{(k)} + 1)}$ for every i and every s.

Hence, given $\mathbf{x}^{(0)}$ and $\rho^{(k)}$ as from the hypothesis, the claim follows by a simple inductive argument. □

Now consider the following instance: $\text{dist}(s, b) \in \{0, 1\}$ for every opinion s and every belief b, $f(1) = g(1) = 1$ and every individual has a different belief. In this case $\mathbf{x}^{(0)}$ is not a consensus profile, $\delta = 1$ and if $\rho^{(0)} \le \frac{f(1)}{\delta g(1)}$, then $c_i(x_i^{(0)}, \mathbf{x}_{-i}^{(0)}) - c_i(s, \mathbf{x}_{-i}^{(0)}) \le 0$ for every i and s. The lower bound then follows from the three claims. □

Extensions. The arguments above can be easily extended in order to prove similar results even if f and g depend on i, and if B and S are continuous. In the first case, it is sufficient to redefine $k^* = \min \left\{ k \mid \rho^{(k)} > \frac{f_i(1)}{g_i(1)} \forall i \right\}$. In this last case, observe that only a discrete subset of these opinions are adopted at each step. Moreover, Claim 1.2 proves that if an opinion is not supported at step $k^* + 1$, then it will be never adopted. Hence, no changes are required to our proof in order to work with continuous opinion domains.

Moreover, similar behavior occurs even if: (i) individuals do not necessarily choose the best response, but they always choose a better response; (ii) more than one individual updates her opinion at each time step.

To prove convergence of better response dynamics, we need to generalize Claim 1.2 as follows. Let us define $\mathbf{d}(\mathbf{x})$ as an ordered vector $(d_1(\mathbf{x}), \ldots, d_m(\mathbf{x}))$, with $m = |S|$, where $d_i(\mathbf{x}) \le d_j(\mathbf{x})$ for every $1 \le i < j \le m$, and for every i there is a distinct $s_i \in S$ such that $d_i(\mathbf{x}) = |\{j \in V \mid x_j = s_i\}|$. In other words, $d_1(\mathbf{x})$ is the number of supporters of the less supported opinion, $d_2(\mathbf{x})$ is the number of supporters of the second less supported opinion, and so on. Given $\mathbf{x}$ and $\mathbf{y}$ we say that $\mathbf{d}(\mathbf{x}) \prec \mathbf{d}(\mathbf{y})$ if the first vector lexicografically precedes the latter. Observe that this implies that the vector $\mathbf{d}^* = (0, \ldots, 0, n)$ corresponding to consensus is minimum with respect to $\prec$.

Claim 1.4. *For every $k \ge k^*$, if $\mathbf{x}^{(k+1)} \ne \mathbf{x}^{(k)}$ then $\mathbf{d}(\mathbf{x}^{(k+1)}) \prec \mathbf{d}(\mathbf{x}^{(k)})$.*

Proof Sketch. As shown in the proof of Claim 1.2, every individual can decrease the cost only by changing her opinion to those opinions that are supported by more individuals than her current opinion. Hence, since $\mathbf{x}^{(k+1)}$ is achieved from $\mathbf{x}^{(k)}$ by having an individual to adopt a better response, that is an opinion that decreases her cost, the claim immediately follows. □

Hence, we can state the following result.

Proposition 2. *If G is a clique and $\rho^* > \frac{f(1)}{g(1)}$, then Algorithm 1 converges to consensus in time at most $k^* + \Theta(n \log m)$, where k^* is as in Theorem 1, even if at Line 6 $x_i^{(k+1)}$ is not set to a best response, but to any opinion that reduces the cost of i.*

Proof Sketch. From Claim 1.4, it follows that the number of supporters of the less supported opinion, say o, never increases, and when all the $d_1(\mathbf{x}^{(k^*)})$ individuals supporting this opinion have been given chance to update it, we achieve that this opinion is not supported by any individual. The same argument can then be repeated for the opinion that is less supported after o disappears, until the vector $\mathbf{d}^*$ (and thus the consensus) is reached. Observe that, if there are z available opinions, then the number of supporters of the less supported opinion cannot be larger than $\frac{n}{z}$. Hence, we have that the consensus is reached after at most

$$k^* + \sum_{i=1}^{m-1} \frac{n}{m+1-i} = k^* + n \sum_{i=2}^{m} \frac{1}{i} = k^* + \Theta(n \log m). \qquad \square$$

As for the case of multiple individuals selected at each time step, let us call *extremal instances* the opinion profiles in which there is no opinion supported by more individuals than any other opinion. For example, if n is even, then the profile $\tilde{\mathbf{x}}$ in which $n/2$ individuals have opinion 0 and the remaining half have opinion 1 is an extremal instance. Moreover, given an extremal instance $\mathbf{x}$, we define as *extremal schedule* for $\mathbf{x}$ any subset S of individuals such that (i) for every $i \in S$, i has one of the two most adopted opinions, and (ii) there are in S the same number of individuals with each of these two opinions. For example, an extremal schedule for the profile $\tilde{\mathbf{x}}$ described above, would be to select for update all individuals at the same time. Clearly, if the starting opinion profile is extreme, and at each time step the set of individuals that update their opinions is an extreme schedule, then the dynamics will never converge to consensus. However, we next show that the lack of convergence to consensus holds only for these extremal cases.

Proposition 3. *If G is a clique and $\rho^* > \frac{f(1)}{g(1)}$, then Algorithm 1 converges to consensus even if multiple individuals are selected for updating their opinion at each time step, as long as the starting profile is not an extremal instance, or there is a time step in which the set of individuals selected for update is not an extremal schedule.*

Proof Scketch. Consider first the case that the starting profile is not an extremal instance, i.e., there is an opinion $s_{\max}$ that is supported by more individuals than every other opinion. Then, according to Claim 1.2, every individual selected for update will adopt opinion $s_{\max}$, and thus a consensus on $s_{\max}$ will be eventually reached.

Suppose instead that there exists an opinion $s \neq s_{\max}$ that has the same number of supporters as $s_{\max}$, but there is a time step t in which we do not have an extremal schedule. Then, without loss of generality, we can suppose that, at this time step t, c supporters of $s_{\max}$, for some $c \geq 0$, and at least $c+1$ supporters of s are selected for update. According to Claim 1.2, the former individuals will adopt opinion s, and the latter ones will adopt opinion $s_{\max}$. Hence, at step $t+1$ there is a clear majority on $s_{\max}$, and thus a consensus is reached. $\square$

3.2 Noisy Dynamics

For sake of presentation, let us consider now $S = B = \{0, 1\}$ (we emphasize again that our arguments do generalize to different settings). Let also $\ell(\mathbf{x})$ be the number of 1's in the opinion profile $\mathbf{x}$. By using the notation of the proof of Theorem 1, we note that if $\mathbf{x} = \mathbf{x}^{(k)}$ for some $k \geq 0$, then $\ell(\mathbf{x}) = n_1^{(k)}$.

Now we assume that individuals only have bounded rationality and update their opinion according to a *logit update rule*. I.e., the rationality level is described by a parameter $\beta > 0$, and after k steps, given that the current opinion profile is $\mathbf{x} = \mathbf{x}^{(k)}$, individual i is selected at random for update and adopts opinion 1 with probability $P_i^k(\mathbf{x}) = \frac{e^{-\beta c_i(\mathbf{x}_{-i}, 1)}}{e^{-\beta c_i(\mathbf{x}_{-i}, 1)} + e^{-\beta c_i(\mathbf{x}_{-i}, 0)}}$. Note that for $\beta = 0$ (no rationality), individuals choose their actions uniformly at random; as

β increases the action that minimizes the cost has a larger probability to be selected; finally, as β tends to infinity, so the logit update tends to the best response update rule discussed above. By setting $\alpha(k) = e^{-\beta\rho^{(k)}g(1)}$ and $C = e^{-\beta f(1)}$, we have that

$$P_i^k(\mathbf{x}) = \begin{cases} \frac{\alpha(k)^{n-\ell(\mathbf{x})-1}}{\alpha(k)^{n-\ell(\mathbf{x})-1}+\alpha(k)^{\ell(\mathbf{x})}\cdot C^{-(1-b_i)}}, & \text{if } x_i = 0; \\ \frac{\alpha(k)^{n-\ell(\mathbf{x})}}{\alpha(k)^{n-\ell(\mathbf{x})}+\alpha(k)^{\ell(\mathbf{x})-1}\cdot C^{-(1-b_i)}}, & \text{if } x_i = 1. \end{cases}$$

Observe that if $k \geq k^*$ (as defined in Theorem 1) and $x_i = 0$, then $P_i^k(\mathbf{x}) > 1/2$ whenever $\ell \geq \frac{n}{2}$, and $P_i^k(\mathbf{x}) < 1/2$ whenever $\ell \leq \frac{n}{2} - 1$; if instead $\ell = \frac{n-1}{2}$, then $P_i^k(\mathbf{x}) > 1/2$ if $b_i = 1$ and $P_i^k(\mathbf{x}) \leq 1/2$, otherwise. Similarly, if $x_i = 1$, then $P_i^k(\mathbf{x}) > 1/2$ whenever $\ell \geq \frac{n+1}{2}$, and $P_i^k(\mathbf{x}) < 1/2$ whenever $\ell \leq \frac{n}{2}$; if instead $\ell = \frac{n+1}{2}$, then $P_i^k(\mathbf{x}) \geq 1/2$ if $b_i = 1$ and $P_i^k(\mathbf{x}) < 1/2$, otherwise.

We denote as τ the smallest integer for which $\mathbf{x}^{(\tau)}$ is such that $\ell(\mathbf{x}^{(\tau)}) \in \{0, n\}$. I.e., τ is the time the dynamics takes for reaching the consensus. Let $\mathbf{E}_{\mathbf{x}}[\tau] = \sum_{t\geq 0} t \cdot \Pr\left(\tau = t \mid \mathbf{x}^{(0)} = \mathbf{x}\right)$, i.e., the expectation of τ given that $\mathbf{x}^{(0)} = \mathbf{x}$. Fix $M = \max\left\{\frac{f(1)}{g(1)}, \frac{3}{\beta g(1)n}, \frac{2\log n}{\beta g(1)n}\right\}$. We prove the following theorem.

Theorem 4. *If G is a clique, $S = B = \{0,1\}$, and $\rho^* > M$, then, for every $\mathbf{x}$, $\mathbf{E}_{\mathbf{x}}[\tau] \leq n^3 + k^*$, where $k^* = \min\{k \mid \rho^{(k)} > M\}$.*

Before to go through the details of the proof of Theorem 4, let us briefly describe the main idea behind it. We first consider an *anonymous slowed-down* dynamics. That is, we build a new process with transition probability $Q_i^k(\mathbf{x})$ that enjoys two remarkable properties: (i) the probability that, after k steps, individual i selected for update adopts opinion o depends only on the number of individuals that actually have that opinion and not on their identities; (ii) this new dynamics is slower than the original, in the sense that the probability of getting closer to consensus is smaller than in the original dynamics. These properties of the new process are proved by carefully coupling the original dynamics with the new process.

The next step consists in translating this anonymous slowed-down process into a birth and death chain that counts the number of individuals with opinion 1. We show that the time that the birth and death chain takes to hit one of its extremes is, except for a delay of k^* steps, at least the same time that the process takes to reach the consensus. Again, this is proved through a coupling between the two processes.

We finally bound the time that the birth and death chain takes to hit one of its extremes. To this aim, we consider a reduced birth and death chain, and evaluate the hitting time of a single extreme of this chain, by using known techniques (see, e.g., [Auletta et al., 2012]). This bound requires that $\beta\rho^{(k^*)}g(1)$ is not too small, that in turn justifies our definition of M.

By choosing $\beta\rho^{(k^*)}g(1)$ to be sufficiently large ($1/\sqrt{n}$ suffices), one can prove, using the same arguments as above, that either a consensus is reached on opinion o or this opinion disappears in about n^3 steps.

The anonymous slowed-down dynamics. Consider an alternative process that works as follows: after k steps, given that the current opinion profile is $\mathbf{x} = \mathbf{x}^{(k)}$, an individual i is selected at random and adopts opinion 1 with probability $Q_i^k(\mathbf{x}) = P_i^k(\mathbf{x})$ if $k < k^*$, otherwise

$Q_i^k(\mathbf{x})$ is

$$\begin{cases} \frac{\alpha(k)^{n-\ell(\mathbf{x})-1}}{\alpha(k)^{n-\ell(\mathbf{x})-1}+\alpha(k)^{\ell(\mathbf{x})}\cdot C^{-1}}, & \text{if } x_i = 0 \text{ and } \ell(\mathbf{x}) > \frac{n-1}{2}; \\ \frac{\alpha(k)^{n-\ell(\mathbf{x})-1}}{\alpha(k)^{n-\ell(\mathbf{x})-1}+\alpha(k)^{\ell(\mathbf{x})}\cdot C}, & \text{if } x_i = 0 \text{ and } \ell(\mathbf{x}) \leq \frac{n-1}{2}; \\ \frac{\alpha(k)^{n-\ell(\mathbf{x})}}{\alpha(k)^{n-\ell(\mathbf{x})}+\alpha(k)^{\ell(\mathbf{x})-1}\cdot C^{-1}}, & \text{if } x_i = 1 \text{ and } \ell(\mathbf{x}) \geq \frac{n+1}{2}; \\ \frac{\alpha(k)^{n-\ell(\mathbf{x})}}{\alpha(k)^{n-\ell(\mathbf{x})}+\alpha(k)^{\ell(\mathbf{x})-1}\cdot C}, & \text{if } x_i = 1 \text{ and } \ell(\mathbf{x}) < \frac{n+1}{2}. \end{cases}$$

We would like to point out some important properties that this new dynamics enjoys. First, observe that for every two profiles $\mathbf{x}$ and $\mathbf{y}$ such that $\ell(\mathbf{y}) = n - \ell(\mathbf{x})$ and $x_i \neq y_i$, if $k \geq k^*$, then $1 - Q_i^k(\mathbf{x}) = Q_i^k(\mathbf{y})$.

Moreover, for every two profiles $\mathbf{x}$ and $\mathbf{y}$ and for every two individuals i and j, we have $Q_i^k(\mathbf{x}) \geq Q_j^k(\mathbf{y})$ whenever one of the following two conditions are satisfied: either (i) $\ell(\mathbf{x}) > \ell(\mathbf{y})$, or (ii) $x_i = y_j$ and $\ell(\mathbf{x}) = \ell(\mathbf{y})$.

Finally, let us consider a profile $\mathbf{x}$. Observe that:

- $Q_i^k(\mathbf{x}) \leq P_i^k(\mathbf{x})$ if one of the following condition is satisfied: (i) $\ell(\mathbf{x}) > \frac{n-1}{2}$; (ii) $\ell(\mathbf{x}) = \frac{n}{2}$ and $x_i = 0$; (iii) $\ell(\mathbf{x}) = \frac{n+1}{2}$ and $x_i = 0$;

- $P_i^k(\mathbf{x}) \leq Q_i^k(\mathbf{x})$ if one of the following condition is satisfied: (i) $\ell(\mathbf{x}) < \frac{n+1}{2}$; (ii) $\ell(\mathbf{x}) = \frac{n}{2}$ and $x_i = 1$; (iii) $\ell(\mathbf{x}) = \frac{n-1}{2}$ and $x_i = 1$.

Roughly speaking, whenever the starting profile is sufficiently away from having half of individuals with opinion 1, this new dynamics decreases its distance from consensus more slowly than the original dynamics. We next formally state this property.

Lemma 5. *Let τ' be as τ but with respect to the anonymous slowed-down dynamics in place of the original one. Then, $\mathbf{E}_{\mathbf{x}}[\tau] \leq \mathbf{E}_{\mathbf{x}}[\tau']$.*

Proof. Let $\mathbf{x}^{(k)}$ be the opinion profile reached by the original dynamics after k steps, and $\mathbf{y}^{(k)}$ be the opinion profile reached by the new dynamics after the same number of steps. Given a profile $\mathbf{x}$, we sometimes consider the profile $\overline{\mathbf{x}}$ achieved by switching the opinion of every individual. Let us also define $\delta_{\mathbf{x}}(k) = \min\{\ell(\mathbf{x}^{(k)}), n - \ell(\mathbf{x}^{(k)})\}$, $\delta_{\mathbf{y}}(k) = \min\{\ell(\mathbf{y}^{(k)}), n - \ell(\mathbf{y}^{(k)})\}$, and $\Delta(k) = \delta_{\mathbf{y}}(k) - \delta_{\mathbf{x}}(k)$.

We will show that it is possible to couple the two dynamics so that, if $\Delta(0) \geq 0$, then $\Delta(k) \geq 0$ for every $k \geq 0$. The lemma then follows.

To this aim, let

$$\begin{aligned} &E_0^{(k)} \text{ be the event in which } \ell(\mathbf{x}^{(k)}) \geq \frac{n}{2} \text{ and } \ell(\mathbf{y}^{(k)}) \geq \frac{n}{2}, \\ &E_1^{(k)} \text{ be the event in which } \ell(\mathbf{x}^{(k)}) \geq \frac{n}{2} \text{ and } \ell(\mathbf{y}^{(k)}) < \frac{n}{2}, \\ &E_2^{(k)} \text{ be the event in which } \ell(\mathbf{x}^{(k)}) < \frac{n}{2} \text{ and } \ell(\mathbf{y}^{(k)}) \geq \frac{n}{2}, \\ &E_3^{(k)} \text{ be the event in which } \ell(\mathbf{x}^{(k)}) < \frac{n}{2} \text{ and } \ell(\mathbf{y}^{(k)}) < \frac{n}{2}. \end{aligned}$$

Let us consider permutations π_0, π_1, π_2, and π_3 of $[n]$ as follows. If $\Delta(k) < 0$, then they are arbitrary permutations. If instead $\Delta(k) \geq 0$, then if $E_0^{(k)}$ occurs, arbitrarily fix a set $A^{(k)} \subseteq \{i \colon x_i^{(k)} = 1 \text{ and } y_i^{(k)} = 0\}$ such that $|A^{(k)}| = \Delta(k)$, and consider π_0 such that $\pi_0(i) = i$ if $i \in A^{(k)}$, and $x_i^{(k)} = y_{\pi_0(i)}^{(k)}$ otherwise; if $E_1^{(k)}$ occurs, arbitrarily fix a set $A^{(k)} \subseteq \{i \colon x_i^{(k)} = 1 \text{ and } y_i^{(k)} = 1\}$ such that $|A^{(k)}| = \Delta(k)$, and consider π_1 such that $\pi_1(i) = i$ if $i \in A^{(k)}$, and

$x_i^{(k)} \neq y_{\pi_1(i)}^{(k)}$ otherwise; if $E_2^{(k)}$ occurs, arbitrarily fix a set $A^{(k)} \subseteq \{i \colon x_i^{(k)} = 0 \text{ and } y_i^{(k)} = 0\}$ such that $|A^{(k)}| = \Delta(k)$, and consider π_2 such that $\pi_2(i) = i$ if $i \in A^{(k)}$, and $x_i^{(k)} \neq y_{\pi_2(i)}^{(k)}$ otherwise; if $E_3^{(k)}$ occurs, arbitrarily fix a set $A^{(k)} \subseteq \{i \colon x_i^{(k)} = 0 \text{ and } y_i^{(k)} = 1\}$ such that $|A^{(k)}| = \Delta(k)$, and consider π_3 such that $\pi_3(i) = i$ if $i \in A^{(k)}$, and $x_i^{(k)} = y_{\pi_3(i)}^{(k)}$ otherwise. Roughly speaking, we are trying to match as many individuals as possible such that their opinions agree with the majority in the original dynamics with individuals whose opinions agree with the majority in the new dynamics.

The coupling picks i u.a.r., and, if $E_c^{(k)}$ occurs, with $c = 0, 3$, then it sets

$$\begin{aligned}
x_i^{(k+1)} &= y_{\pi_c(i)}^{(k+1)} = 1 && \text{with prob. } \min\{P_i^k(\mathbf{x}^{(k)}), Q_{\pi_c(i)}^k(\mathbf{y}^{(k)})\};\\
x_i^{(k+1)} &= y_{\pi_c(i)}^{(k+1)} = 0 && \text{with prob. } 1 - \max\{P_i^k(\mathbf{x}^{(k)}), Q_{\pi_c(i)}^k(\mathbf{y}^{(k)})\};\\
x_i^{(k+1)} &= 1, y_{\pi_c(i)}^{(k+1)} = 0 && \text{with prob. } \max\{0, P_i^k(\mathbf{x}^{(k)}) - Q_{\pi_c(i)}^k(\mathbf{y}^{(k)})\};\\
x_i^{(k+1)} &= 0, y_{\pi_c(i)}^{(k+1)} = 1 && \text{with prob. } \max\{0, Q_{\pi_c(i)}^k(\mathbf{y}^{(k)}) - P_i^k(\mathbf{x}^{(k)})\};
\end{aligned}$$

otherwise (i.e., if $E_c^{(k)}$ occurs with $c = 1, 2$) it sets

$$\begin{aligned}
x_i^{(k+1)} &= 1, y_{\pi_c(i)}^{(k+1)} = 0 && \text{with prob. } \min\{P_i^k(\mathbf{x}^{(k)}), 1 - Q_{\pi_c(i)}^k(\mathbf{y}^{(k)})\};\\
x_i^{(k+1)} &= 0, y_{\pi_c(i)}^{(k+1)} = 1 && \text{with prob. } \min\{1 - P_i^k(\mathbf{x}^{(k)}), Q_{\pi_c(i)}^k(\mathbf{y}^{(k)})\};\\
x_i^{(k+1)} &= y_{\pi_c(i}^{(k+1)} = 1 && \text{with prob. } \max\{0, P_i^k(\mathbf{x}^{(k)}) + Q_{\pi_c(i)}^k(\mathbf{y}^{(k)}) - 1\};\\
x_i^{(k+1)} &= y_{\pi_c(i)}^{(k+1)} = 0 && \text{with prob. } \max\{0, 1 - Q_{\pi_c(i)}^k(\mathbf{y}^{(k)}) - P_i^k(\mathbf{x}^{(k)})\}.
\end{aligned}$$

It is easy to check that above coupling correctly sets the update probabilities of both dynamics.

We now prove by induction that $\Delta(k) \geq 0$ for every $k \geq 0$. This is clearly true if $k = 0$, since we assume that $\mathbf{x}^{(0)} = \mathbf{y}^{(0)} = \mathbf{x}$.

Suppose now that the claim holds for k and let us prove it for $k+1$. Let i be the individual selected for update. First let us assume that $\ell(\mathbf{x}^{(k)}) \geq \frac{n}{2}$ and $x_i^{(k)} = 0$. We will distinguish several cases.

Suppose first that $\ell(\mathbf{y}^{(k)}) \geq \frac{n}{2}$ and $y_{\pi_0(i)}^{(k)} = 0$. Since $\ell(\mathbf{y}^{(k)}) \leq \ell(\mathbf{x}^{(k)})$ and $x_i^{(k)} = y_{\pi_0(i)}^{(k)}$, we have that $P_i^k(\mathbf{x}^{(k)}) \geq Q_i^k(\mathbf{x}^{(k)}) \geq Q_{\pi_0(i)}^k(\mathbf{y}^{(k)})$. Thus, with probability $Q_{\pi_0(i)}^k(\mathbf{y}^{(k)})$ we have that $\delta_{\mathbf{x}}(k+1) = \delta_{\mathbf{x}}(k) - 1$, $\delta_{\mathbf{y}}(k+1) = \delta_{\mathbf{y}}(k) - 1$, and thus $\Delta(k+1) = \Delta(k) \geq 0$; with probability $1 - P_i^k(\mathbf{x}^{(k)})$, we have that $\delta_{\mathbf{x}}(k+1) = \delta_{\mathbf{x}}(k)$, $\delta_{\mathbf{y}}(k+1) = \delta_{\mathbf{y}}(k)$, and thus $\Delta(k+1) = \Delta(k) \geq 0$; with remaining probability, we have that $\delta_{\mathbf{x}}(k+1) = \delta_{\mathbf{x}}(k) - 1$, $\delta_{\mathbf{y}}(k+1) = \delta_{\mathbf{y}}(k)$, and thus $\Delta(k+1) = \Delta(k) + 1 \geq 0$.

Suppose now that $\ell(\mathbf{y}^{(k)}) \geq \frac{n}{2}$ and $y_{\pi_0(i)}^{(k)} = 1$. Recall that our definition of π_0 implies that in this case $\Delta(k) \geq 1$. Then since $\ell(\mathbf{y}^{(k)}) < \ell(\mathbf{x}^{(k)})$, we have that $P_i^k(\mathbf{x}^{(k)}) \geq Q_i^k(\mathbf{x}^{(k)}) \geq Q_{\pi_0(i)}^k(\mathbf{y}^{(k)})$. Thus, with probability $Q_{\pi_0(i)}^k(\mathbf{y}^{(k)})$ we have that $\delta_{\mathbf{x}}(k+1) = \delta_{\mathbf{x}}(k) - 1$, $\delta_{\mathbf{y}}(k+1) = \delta_{\mathbf{y}}(k)$, and thus $\Delta(k+1) = \Delta(k) + 1 \geq 0$; with probability $1 - P_i^k(\mathbf{x}^{(k)})$, we have that $\delta_{\mathbf{x}}(k+1) = \delta_{\mathbf{x}}(k)$, $\delta_{\mathbf{y}}(k+1) \geq \delta_{\mathbf{y}}(k) - 1$ (if $\ell(\mathbf{y}^{(k)}) = n/2$, then $\delta_{\mathbf{y}}(k+1) = \delta_{\mathbf{y}}(k) - 1$, if $\ell(\mathbf{y}^{(k)}) = (n+1)/2$, then $\delta_{\mathbf{y}}(k+1) = \delta_{\mathbf{y}}(k)$, otherwise $\delta_{\mathbf{y}}(k+1) > \delta_{\mathbf{y}}(k)$), and thus $\Delta(k+1) \geq \Delta(k) - 1 \geq 0$; with remaining probability, we have that $\delta_{\mathbf{x}}(k+1) = \delta_{\mathbf{x}}(k) - 1$, $\delta_{\mathbf{y}}(k+1) \geq \delta_{\mathbf{y}}(k) - 1$, and thus $\Delta(k+1) \geq \Delta(k) \geq 0$.

Suppose now that $\ell(\mathbf{y}^{(k)}) < \frac{n}{2}$ and $y_{\pi_1(i)}^{(k)} = 0$. Recall that our definition of π_1 implies that in this case $\Delta(k) \geq 1$. Then since $n - \ell(\mathbf{y}^{(k)}) = \ell(\overline{\mathbf{y}}^{(k)}) < \ell(\mathbf{x}^{(k)})$, we have that $P_i^k(\mathbf{x}^{(k)}) \geq Q_i^k(\mathbf{x}^{(k)}) \geq Q_{\pi_1(i)}^k(\overline{\mathbf{y}}^{(k)}) = 1 - Q_{\pi_1(i)}^k(\mathbf{y}^{(k)})$. Thus, with probability $1 - Q_{\pi_1(i)}^k(\mathbf{y}^{(k)})$ we have that

$\delta_{\mathbf{x}}(k+1) = \delta_{\mathbf{x}}(k) - 1$, $\delta_{\mathbf{y}}(k+1) = \delta_{\mathbf{y}}(k)$, and thus $\Delta(k+1) = \Delta(k) + 1 \geq 0$; with probability $1 - P_i^k(\mathbf{x}^{(k)})$, we have that $\delta_{\mathbf{x}}(k+1) = \delta_{\mathbf{x}}(k)$, $\delta_{\mathbf{y}}(k+1) \geq \delta_{\mathbf{y}}(k)$, (if $\ell(\mathbf{y}^{(k)}) = (n-1)/2$, then $\delta_{\mathbf{y}}(k+1) = \delta_{\mathbf{y}}(k)$, otherwise $\delta_{\mathbf{y}}(k+1) > \delta_{\mathbf{y}}(k)$), and thus $\Delta(k+1) \geq \Delta(k) \geq 0$; with remaining probability, we have that $\delta_{\mathbf{x}}(k+1) = \delta_{\mathbf{x}}(k) - 1$, $\delta_{\mathbf{y}}(k+1) \geq \delta_{\mathbf{y}}(k)$, (as above), and thus $\Delta(k+1) \geq \Delta(k) - 1 \geq 0$.

Suppose finally that $\ell(\mathbf{y}^{(k)}) < \frac{n}{2}$ and $y_{\pi_1(i)}^{(k)} = 1$. Since $n - \ell(\mathbf{y}^{(k)}) = \ell(\overline{\mathbf{y}}^{(k)}) \leq \ell(\mathbf{x}^{(k)})$ and $\overline{y}_{\pi_1(i)}^{(k)} = x_i^{(k)}$, we have $P_i^k(\mathbf{x}^{(k)}) \geq Q_i^k(\mathbf{x}^{(k)}) \geq Q_{\pi_1(i)}^k(\overline{\mathbf{y}}^{(k)}) = 1 - Q_{\pi_1(i)}^k(\mathbf{y}^{(k)})$. Thus, with probability $1 - Q_{\pi_1(i)}^k(\mathbf{y}^{(k)})$ we have that $\delta_{\mathbf{x}}(k+1) = \delta_{\mathbf{x}}(k) - 1$, $\delta_{\mathbf{y}}(k+1) = \delta_{\mathbf{y}}(k) - 1$, and thus $\Delta(k+1) = \Delta(k) \geq 0$; with probability $1 - P_i^k(\mathbf{x}^{(k)})$, we have that $\delta_{\mathbf{x}}(k+1) = \delta_{\mathbf{x}}(k)$, $\delta_{\mathbf{y}}(k+1) = \delta_{\mathbf{y}}(k)$, and thus $\Delta(k+1) = \Delta(k) \geq 0$; with remaining probability, we have that $\delta_{\mathbf{x}}(k+1) = \delta_{\mathbf{x}}(k) - 1$, $\delta_{\mathbf{y}}(k+1) = \delta_{\mathbf{y}}(k)$, and thus $\Delta(k+1) \geq \Delta(k) - 1 \geq 0$.

The remaining cases (i.e., when $\ell(\mathbf{x}^{(k)}) < \frac{n}{2}$ or $x_i^{(k)} = 1$) can be handled in a similar way. □

Translating the process into a birth and death chain. Consider now a birth and death chain on $\{0, \ldots, n\}$ such that the probability p_ℓ of going from state ℓ to state $\ell + 1$ is

$$p_\ell = \frac{n-\ell}{n} \cdot \begin{cases} \frac{\alpha(k^*)^{n-\ell-1}}{\alpha(k^*)^{n-\ell-1} + \alpha(k^*)^{\ell} \cdot C^{-1}}, & \text{if } \ell > \frac{n-1}{2}; \\ \frac{\alpha(k^*)^{n-\ell-1}}{\alpha(k^*)^{n-\ell-1} + \alpha(k^*)^{\ell} \cdot C}, & \text{otherwise.} \end{cases}$$

The probability q_ℓ of going from state ℓ to state $\ell - 1$ is

$$q_\ell = \frac{\ell}{n} \cdot \begin{cases} 1 - \frac{\alpha(k^*)^{n-\ell}}{\alpha(k^*)^{n-\ell} + \alpha(k^*)^{\ell-1} \cdot C^{-1}}, & \text{if } \ell \geq \frac{n+1}{2}; \\ 1 - \frac{\alpha(k^*)^{n-\ell}}{\alpha(k^*)^{n-\ell} + \alpha(k^*)^{\ell-1} \cdot C}, & \text{otherwise.} \end{cases}$$

With the remaining probability the chain remains in ℓ. We denote as $s(k)$ the position of this chain after k steps and as $\tau_{0,n}$ the smallest integer such that $s(\tau_{0,n}) \in \{0, n\}$. Let $\mathbf{E}_\ell[\tau_{0,n}]$ be the expectation of $\tau_{0,n}$ given that $s(0) = \ell$, i.e. $\mathbf{E}_\ell[\tau_{0,n}] = \sum_{t \geq 0} t \cdot \Pr(\tau_{0,n} = t \mid s(0) = \ell)$. We then have the following lemma.

Lemma 6. $\mathbf{E}_{\mathbf{x}}[\tau'] \leq k^* + \mathbf{E}_{\ell(\mathbf{x}^{(k^*)})}[\tau_{0,n}]$.

Proof. Let $\mathbf{x}^{(k)}$ be the opinion profile reached after k steps by the dynamics whose transition probabilities are described by Q. As above, we set $\delta(k) = \min\{\ell(\mathbf{x}^{(k)}), n - \ell(\mathbf{x}^{(k)})\}$, $\delta'(k) = \min\{s(k), n - s(k)\}$, and $\Delta(k) = \delta'(k) - \delta(k)$.

We will show that it is possible to couple this dynamics with the birth and death chain described above so that, if $\Delta(k^\star) \geq 0$, then $\Delta(k) \geq 0$ for every $k \geq k^\star$. The lemma then follows.

The coupling and the proof that it enjoys this desired property is very similar to the one described above. Anyway, for sake of completeness, we next describe it formally.

Given a profile $\mathbf{x}^{(k)}$ and the state $s(k)$ of the birth and death chain, if $\Delta(k) \geq 0$, then we define the profile $\mathbf{y}^{(k)}$ as follows:

- if $\ell(\mathbf{x}^{(k)}) \geq \frac{n}{2}$ and $s(k) \geq \frac{n}{2}$, arbitrarily fix a set $A^{(k)} \subseteq \{i \colon x_i^{(k)} = 1\}$ such that $|A^{(k)}| = \Delta(k)$, and set $y_i^{(k)} = 0$ if $i \in A^{(k)}$, and $y_i^{(k)} = x_i^{(k)}$ otherwise;
- if $\ell(\mathbf{x}^{(k)}) \geq \frac{n}{2}$ and $s(k) < \frac{n}{2}$, arbitrarily fix a set $A^{(k)} \subseteq \{i \colon x_i^{(k)} = 1\}$ such that $|A^{(k)}| = \Delta(k)$, and set $y_i^{(k)} = 1$ if $i \in A^{(k)}$, and $y_i^{(k)} = 1 - x_i^{(k)}$ otherwise;

- if $\ell(\mathbf{x}^{(k)}) < \frac{n}{2}$ and $s(k) \geq \frac{n}{2}$, arbitrarily fix a set $A^{(k)} \subseteq \{i\colon x_i^{(k)} = 0\}$ such that $|A^{(k)}| = \Delta(k)$, and set $y_i^{(k)} = 0$ if $i \in A^{(k)}$, and $y_i^{(k)} = 1 - x_i^{(k)}$ otherwise;

- if $\ell(\mathbf{x}^{(k)}) < \frac{n}{2}$ and $s(k) < \frac{n}{2}$, arbitrarily fix a set $A^{(k)} \subseteq \{i\colon x_i^{(k)} = 0\}$ such that $|A^{(k)}| = \Delta(k)$, and set $y_i^{(k)} = 1$ if $i \in A^{(k)}$, and $y_i^{(k)} = x_i^{(k)}$ otherwise;

(Observe that $\ell(\mathbf{y}^{(k)}) = s(k)$.) If $\Delta(k) < 0$, then $\mathbf{y}^{(k)}$ is an arbitrary profile such that $\ell(\mathbf{y}^{(k)}) = s(k)$.

The coupling then proceeds as follows: at each time step $k > k^\star$, we select an individual i uniformly at random and

- If either (i) $\ell(\mathbf{x}^{(k-1)}) \geq \frac{n}{2}$ and $s(k-1) \geq \frac{n}{2}$ or (ii) $\ell(\mathbf{x}^{(k-1)}) < \frac{n}{2}$ and $s(k-1) < \frac{n}{2}$, then with probability $\min\{Q_i^{k^\star}(\mathbf{y}^{(k-1)}), Q_i^{k-1}(\mathbf{x}^{(k-1)})\}$ set $x_i^{(k)} = 1$ and $s(k) = s(k-1) + 1 - y_i^{(k-1)}$, with probability $1 - \max\{Q_i^{k^\star}(\mathbf{y}^{(k-1)}), Q_i^{k-1}(\mathbf{x}^{(k-1)})\}$ set $x_i^{(k)} = 0$ and $s(k) = s(k-1) - y_i^{(k-1)}$, with probability $\max\{0, Q_i^{k-1}(\mathbf{x}^{(k-1)}) - Q_i^{k^\star}(\mathbf{y}^{(k-1)})\}$ set $x_i^{(k)} = 1$ and $s(k) = s(k-1) - y_i^{(k-1)}$, with probability $\max\{0, Q_i^{k^\star}(\mathbf{y}^{(k-1)}) - Q_i^{k-1}(\mathbf{x}^{(k-1)})\}$ set $x_i^{(k)} = 0$ and $s(k) = s(k-1) + 1 - y_i^{(k-1)}$.

- If either (i) $\ell(\mathbf{x}^{(k-1)}) \geq \frac{n}{2}$ and $s(k-1) < \frac{n}{2}$ or (ii) $\ell(\mathbf{x}^{(k-1)}) < \frac{n}{2}$ and $s(k-1) \geq \frac{n}{2}$, then with probability $\min\{1 - Q_i^{k^\star}(\mathbf{y}^{(k-1)}), Q_i^{k-1}(\mathbf{x}^{(k-1)})\}$ set $x_i^{(k)} = 1$ and $s(k) = s(k-1) - y_i^{(k-1)}$, with probability $\min\{Q_i^{k^\star}(\mathbf{y}^{(k-1)}), 1 - Q_i^{k-1}(\mathbf{x}^{(k-1)})\}$ set $x_i^{(k)} = 0$ and $s(k) = s(k-1) + 1 - y_i^{(k-1)}$, with probability $\max\{0, Q_i^{k-1}(\mathbf{x}^{(k-1)}) + Q_i^{k^\star}(\mathbf{y}^{(k-1)}) - 1\}$ set $x_i^{(k)} = 1$ and $s(k) = s(k-1) + 1 - y_i^{(k-1)}$, with probability $\max\{0, 1 - Q_i^{k^\star}(\mathbf{y}^{(k-1)}) - Q_i^{k-1}(\mathbf{x}^{(k-1)})\}$ set $x_i^{(k)} = 0$ and $s(k) = s(k-1) - y_i^{(k-1)}$.

It is easy to check that above coupling correctly sets the update probability both of the dynamics and of the birth and death chain.

We now prove by induction that if $\Delta(k^\star) \geq 0$, then $\Delta(k) \geq 0$ for every $k \geq k^\star$. Suppose indeed that the claim holds for $k-1$ and let us prove it holds also for k. Let i be the individual selected for update. First let us assume that $\ell(\mathbf{x}^{(k-1)}) \geq \frac{n}{2}$ and $x_i^{(k-1)} = 0$. Suppose first that $s(k-1) \geq \frac{n}{2}$ and $y_i^{(k-1)} = 0$. Since $\ell(\mathbf{y}^{(k-1)}) \leq \ell(\mathbf{x}^{(k-1)})$ and $x_i^{(k-1)} = y_i^{(k-1)}$, we have that $Q_i^{k-1}(\mathbf{x}^{(k-1)}) \geq Q_i^{k-1}(\mathbf{y}^{(k-1)}) \geq Q_i^{k^\star}(\mathbf{y}^{(k-1)})$. Thus, with probability $Q_i^{k^\star}(\mathbf{y}^{(k-1)})$ we have that $\delta(k) = \delta(k-1) - 1$, $\delta'(k) = \delta'(k-1) - 1$, and thus $\Delta(k) = \Delta(k-1) \geq 0$; with probability $1 - Q_i^{k-1}(\mathbf{x}^{(k-1)})$, we have that $\delta(k) = \delta(k-1)$, $\delta'(k) = \delta'(k-1)$, and thus $\Delta(k) = \Delta(k-1) \geq 0$; with remaining probability, we have that $\delta(k) = \delta(k-1) - 1$, $\delta'(k) = \delta'(k-1)$, and thus $\Delta(k) = \Delta(k-1) + 1 \geq 0$.

Suppose now that $s(k-1) \geq \frac{n}{2}$ and $y_i^{(k-1)} = 1$. Recall that our definition of $\mathbf{y}^{(k-1)}$ implies that in this case $\Delta(k-1) \geq 1$. Then since $\ell(\mathbf{y}^{(k-1)}) < \ell(\mathbf{x}^{(k-1)})$, we have that $Q_i^{k-1}(\mathbf{x}^{(k-1)}) \geq Q_i^{k-1}(\mathbf{y}^{(k-1)}) \geq Q_i^{k^\star}(\mathbf{y}^{(k-1)})$. Thus, with probability $Q_i^{k^\star}(\mathbf{y}^{(k-1)})$ we have that $\delta(k) = \delta(k-1) - 1$, $\delta'(k) = \delta'(k-1)$, and thus $\Delta(k) = \Delta(k-1) + 1 \geq 0$; with probability $1 - Q_i^{k-1}(\mathbf{x}^{(k-1)})$, we have that $\delta(k) = \delta(k-1)$, $\delta'(k) \geq \delta'(k-1) - 1$ (if $\ell(\mathbf{y}^{(k-1)}) = n/2$, then $\delta'(k) = \delta'(k-1) - 1$, if $\ell(\mathbf{y}^{(k-1)}) = (n+1)/2$, then $\delta'(k) = \delta'(k-1)$, otherwise $\delta'(k) > \delta'(k-1)$), and thus $\Delta(k) \geq \Delta(k-1) - 1 \geq 0$; with remaining probability, we have that $\delta(k) = \delta(k-1) - 1$, $\delta'(k) \geq \delta'(k-1) - 1$, and thus $\Delta(k) \geq \Delta(k-1) \geq 0$.

Then, suppose that $s(k-1) < \frac{n}{2}$ and $y_i^{(k-1)} = 0$. Recall that our definition of $\mathbf{y}^{(k-1)}$ implies that in this case $\Delta(k-1) \geq 1$. Then since $n - \ell(\mathbf{y}^{(k-1)}) = \ell(\overline{\mathbf{y}}^{(k-1)}) < \ell(\mathbf{x}^{(k-1)})$, we have that $Q_i^{k-1}(\mathbf{x}^{(k-1)}) \geq Q_i^{k-1}(\overline{\mathbf{y}}^{(k-1)}) = 1 - Q_i^{k-1}(\mathbf{y}^{(k-1)}) \geq 1 - Q_i^{k^\star}(\mathbf{y}^{(k-1)})$. Thus, with probability $1 - Q_I^{k^\star}(\mathbf{y}^{(k-1)})$ we have that $\delta(k) = \delta(k-1) - 1$, $\delta'(k) = \delta'(k-1)$, and thus

$\Delta(k) = \Delta(k-1) + 1 \geq 0$; with probability $1 - Q_i^{k-1}(\mathbf{x}^{(k-1)})$, we have that $\delta(k) = \delta(k-1)$, $\delta'(k) \geq \delta'(k-1)$ (if $\ell(\mathbf{y}^{(k-1)}) = (n-1)/2$, then $\delta'(k) = \delta'(k-1)$, otherwise $\delta'(k) > \delta'(k-1)$), and thus $\Delta(k) \geq \Delta(k-1) \geq 0$; with remaining probability, we have that $\delta(k) = \delta(k-1) - 1$, $\delta'(k) \geq \delta'(k-1)$ (as above), and thus $\Delta(k) \geq \Delta(k-1) - 1 \geq 0$.

Finally, consider that $s(k-1) < \frac{n}{2}$ and $y_i^{(k-1)} = 1$. Then since $n - \ell(\mathbf{y}^{(k-1)}) = \ell(\overline{\mathbf{y}}^{(k-1)}) \leq \ell(\mathbf{x}^{(k-1)})$ and $\overline{y}_i^{(k-1)} = x_i^{(k-1)}$, we have that $Q_i^{k-1}(\mathbf{x}^{(k-1)}) \geq Q_i^{k-1}(\overline{\mathbf{y}}^{(k-1)}) = 1 - Q_i^{k-1}(\mathbf{y}^{(k-1)}) \geq 1 - Q_i^{k^\star}(\mathbf{y}^{(k-1)})$. Thus, with probability $1 - Q_i^{k^\star}(\mathbf{y}^{(k-1)})$ we have that $\delta(k) = \delta(k-1) - 1$, $\delta'(k) = \delta'(k-1) - 1$, and thus $\Delta(k) = \Delta(k-1) \geq 0$; with probability $1 - Q_i^{k-1}(\mathbf{x}^{(k-1)})$, we have that $\delta(k) = \delta(k-1)$, $\delta'(k) = \delta'(k-1)$, and thus $\Delta(k) = \Delta(k-1) \geq 0$; with remaining probability, we have that $\delta(k) = \delta(k-1) - 1$, $\delta'(k) = \delta'(k-1)$, and thus $\Delta(k) \geq \Delta(k-1) - 1 \geq 0$.

The remaining cases (i.e., when $\ell(\mathbf{x}^{(k-1)}) < \frac{n}{2}$ or $x_i^{(k-1)} = 1$) can be handled in a similar way. □

Bounding the hitting time of the extremes. In order to bound $\mathbf{E}_\ell\left[\tau_{0,n}\right]$, let us consider an alternative birth and death chain defined on the set of states $\{\frac{n}{2}, \dots, n\}$, such that the probability of a transition from state ℓ to state $\ell + 1$ is $p^*_\ell = p_\ell$ for every $\ell > \frac{n}{2}$, and it is $p^*_\ell = p_\ell + q_\ell$ if $\ell = \frac{n}{2}$, whereas the probability of a transition from state ℓ to state $\ell - 1$ is $q^*_\ell = q_\ell$ if $\ell > \frac{n}{2}$, and $q'_\ell = 0$ if $\ell = \frac{n}{2}$. (If n is odd, the set of states is $\{\frac{n+1}{2}, \dots, n\}$, the probability of a transition from state ℓ to state $\ell + 1$ is $p^*_\ell = p_\ell$ for every $\ell \geq \frac{n+1}{2}$, whereas the probability of a transition from state ℓ to state $\ell - 1$ is $q^*_\ell = q_\ell$ if $\ell > \frac{n+1}{2}$, and $q'_\ell = 0$ if $\ell = \frac{n+1}{2}$). It is not hard to see that, if $\ell \geq \frac{n}{2}$, then $\mathbf{E}_\ell\left[\tau_{0,n}\right] \leq \mathbf{E}_\ell\left[\tau^*_n\right]$, where τ^*_n is the first time step in which this new chain hits the state n (the case for $\ell \leq \frac{n}{2}$ is symmetric). Indeed, this immediately follows by coupling the two chains as follows: the new chain moves forward if and only if either the original chain is in a state $\ell \geq \frac{n}{2}$ and moves forward or it is in a state $\ell \leq \frac{n}{2}$ and moves backward; the new chain moves backward if and only if either the original chain is in a state $\ell > \frac{n+1}{2}$ and moves backward or it is in a state $\ell < \frac{n-1}{2}$ and moves forward; the new chain does not move if and only if n is odd, and either $\ell = \frac{n+1}{2}$ and the original chain moves backward or $\ell = \frac{n-1}{2}$ and the original chain moves forward.

We next show that $\mathbf{E}_\ell\left[\tau^*_n\right] \leq n^3$, whenever $\ell \geq \frac{n}{2}$. Therefore, we can conclude that for every starting profile $\mathbf{x}$, the dynamics converges to consensus in expected time $\mathbf{E}_\mathbf{x}\left[\tau\right] \leq \mathbf{E}_\mathbf{x}\left[\tau'\right] \leq k^* + \mathbf{E}_{\ell(\mathbf{x})}\left[\tau_{0,n}\right] \leq k^* + \mathbf{E}_\ell\left[\tau^*_n\right] \leq k^* + n^3$.

The proof of this last step follows standard arguments (see, e.g., [Auletta et al., 2012]). Indeed, it is well known that

$$\mathbf{E}_\ell\left[\tau^*_n\right] = \sum_{i=\ell+1}^{n} \sum_{j=\frac{n}{2}}^{i-1} \frac{1}{p^*_j} \prod_{m=j+1}^{i-1} \frac{q^*_m}{p^*_m}.$$

Now observe that

$$\begin{aligned}
\frac{q^*_m}{p^*_m} &= \frac{m\alpha(k^*)^{m-1}C^{-1}\left(\alpha(k^*)^{n-m-1} + \alpha(k^*)^m C^{-1}\right)}{(n-m)\alpha(k^*)^{n-m-1}\left(\alpha(k^*)^{n-m} + \alpha(k^*)^{m-1}C^{-1}\right)} \\
&= \frac{1 + \frac{h}{n}}{1 - \frac{h}{n}} \cdot \alpha(k^*)^{h-1}C^{-1} \cdot \frac{1+t}{1 + t\alpha(k^*)^2} \\
&\leq \frac{1 + \frac{h}{n}}{1 - \frac{h}{n}} \cdot \alpha(k^*)^{h+1}C^{-1}
\end{aligned}$$

for every $m \in \{\frac{n}{2} + 1, \dots, n-1\}$, where $h = 2m - n \geq 2$ and $t = \alpha(k^*)^{h+1}C^{-1}$.

Observe that

$$\alpha(k^*)^{h+1}C^{-1} = e^{-\beta[\rho^{(k^*)}g(1)(h-1)+f(1)]} \leq e^{-\beta\rho^{(k^*)}g(1)h},$$

where the last inequality follows because $\rho^{(k^*)} > \frac{f(1)}{g(1)}$. Moreover, if $m \leq \frac{3n}{4}$ (and, thus, $h \leq \frac{n}{2}$), then, since $e^{\frac{x}{1+x}} \leq 1 + x \leq e^x$,

$$\frac{1+\frac{h}{n}}{1-\frac{h}{n}} \leq e^{\frac{h}{n}} e^{\frac{h}{n-h}} \leq e^{\frac{3h}{n}}.$$

Thus, since $\rho^{(k^*)} > \frac{3}{\beta g(1) n}$,

$$\frac{q_m^*}{p_m^*} \leq e^{-h\left(\beta\rho^{(k^*)}g(1)-\frac{3}{n}\right)} \leq 1.$$

If $\frac{3n}{4} < m \leq n-1$ (and thus, $h > \frac{n}{2}$), then $\frac{m}{n-m} \leq n$. Then,

$$\frac{q_m^*}{p_m^*} \leq e^{-\left(\beta\rho^{(k^*)}g(1)h-\log n\right)} \leq e^{-\left(\frac{n}{2}\beta\rho^{(k^*)}g(1)-\log n\right)} \leq 1,$$

where in the last inequality we used that $\rho^{(k^*)} > 2\frac{\log n}{\beta g(1) n}$. Moreover, we have that

$$\frac{1}{p_j^*} = \frac{n}{n-j}\frac{\alpha(k^*)^{n-j-1} + \alpha(k^*)^j \cdot C^{-1}}{\alpha(k^*)^{n-j-1}} = \frac{n}{n-j}\left(1 + \alpha(k^*)^{2j-n+1} \cdot C^{-1}\right).$$

Since for every $j \in \left\{\frac{n}{2}, \ldots, n-1\right\}$ we have that $\frac{n}{n-j} \leq n$ and $\alpha(k^*)^{2j-n+1} \cdot C^{-1} \leq \alpha(k^*)^{2j-n} \leq 1$, then it follows that

$$\mathbf{E}_\ell\left[\tau_n^*\right] \leq \sum_{i=\ell+1}^{n}\sum_{j=\frac{n}{2}}^{i-1} n \leq n^3.$$

3.3 Extensions

Let us now briefly discuss how this proof can be extended to work also in different settings.

Multiple opinions. First of all, let us consider the case in which there are multiple available opinions, say d, and let o be the opinion that is adopted by the minority of individuals at the beginning. Let $P_i^k(\mathbf{x})$ be the probability according to the logit update rule that, after k steps, individual i when selected for update adopts opinion o given that the current opinion profile is $\mathbf{x}$. The same approach described above, i.e., define an anonymous slowed-down process, translate it into a birth and death chain, and bound the hitting time of the extremes of this chain, can be adopted to prove that, for every $M' \geq M$, if $\rho^* > M'$, then in at most $n^3 + k^*$ steps, with $k^* = \min\{k \mid \rho^{(k)} > M'\}$, either a consensus on o has been reached, or this opinion disappears. Suppose that this second event occurs. Observe that the behavior of the dynamics when conditioned on the event that o does not appear again in the next n^3 steps, is equivalent to the behavior of the dynamics when only $d-1$ opinions are available. Then, if o' is the opinion that is adopted by the minority of individuals when o disappears, we have that in the next n^3 steps, either a consensus on o' has been reached, or o' disappears. By repeating this argument, we have that conditioned on the event E that an opinion that disappears does not appear again within the next $(d-2)n^3$ steps, a consensus must be reached in $k^* + (d-1)n^3$ steps, i.e., $\mathbf{E}_{\mathbf{x}^{(k^*)}}\left[\tau \mid E\right] = (d-1)n^3$. However, by denoting with $\bar{E}$ the case that event E does not occur, we have

$$\max_{\mathbf{x}^{(k^*)}} \mathbf{E}_{\mathbf{x}^{(k^*)}}\left[\tau\right] \leq \max_{\mathbf{x}^{(k^*)}} \mathbf{E}_{\mathbf{x}^{(k^*)}}\left[\tau \mid E\right]\Pr(E) + \max_{\mathbf{x}^{(k^*)}} \mathbf{E}_{\mathbf{x}^{(k^*)}}\left[\tau \mid \bar{E}\right]\Pr(\bar{E})$$

$$\leq (d-1)n^3 + \left[(d-2)n^3 + \max_{\mathbf{x}^{(k^*)}} \mathbf{E}_{\mathbf{x}^{(k^*)}}\left[\tau\right]\right] \Pr(\bar{E}),$$

where the second inequality uses that $\mathbf{E}_{\mathbf{x}^{(k^*)}}\left[\tau \mid \bar{E}\right]$ is at most the time that the disappeared opinion takes to reappear plus the time to converge to consensus from the newly created profile. Then, by rearranging, we have that

$$\max_{\mathbf{x}^{(k^*)}} \mathbf{E}_{\mathbf{x}^{(k^*)}}\left[\tau\right] \leq (d-1)n^3 \cdot \frac{1+\Pr(\bar{E})}{1-\Pr(\bar{E})} = (d-1)n^3\left(1 + \frac{2\Pr(\bar{E})}{1-\Pr(\bar{E})}\right).$$

Now observe that the probability that an opinion that is not supported by any individual will be adopted in the next step is at most

$$\frac{1}{1+e^{\beta(\rho^{(k^*)}g(1)(n-1)-f(1))}} \leq e^{-\beta(\rho^{(k^*)}g(1)(n-1)-f(1))} \leq e^{-\beta\rho^{(k^*)}g(1)n},$$

where the last inequality uses that $\rho^{(k^*)} \geq \frac{f(1)}{g(1)}$. Hence, by union bound, $\Pr(\bar{E}) \leq \frac{(d-2)n^3}{e^{\beta\rho^{(k^*)}g(1)n}}$. By taking $\rho^{(k^*)} \geq \frac{\log 3(d-2)n^3}{\beta g(1)n}$, we then have that $\Pr(\bar{E}) \leq \frac{1}{3}$, and thus $\max_{\mathbf{x}^{(k^*)}} \mathbf{E}_{\mathbf{x}^{(k^*)}}\left[\tau\right] \leq 2(d-1)n^3$.

So, summarizing, we obtain the following result.

Theorem 7. *If G is a clique, $S = B = \{0,\ldots,d-1\}$, and $\rho^* > M' = \max\left\{\frac{f(1)}{g(1)}, \frac{\log 3(d-2)n^3}{\beta g(1)n}\right\}$, then, for every $\mathbf{x}$, $\mathbf{E}_{\mathbf{x}}\left[\tau\right] \leq 2(d-1)n^3 + k^*$, where $k^* = \min\{k \mid \rho^{(k)} > M'\}$.*

It is immediate to see that, by opportunely redefining M', all above arguments can be made to work even if $B \neq S$ and f and g depend on i.

Different noisy update rules. Our approach can be adapted to work also with noisy update rules different from logit. As an example of this adaption, we consider the *mistake model*, that assumes that at each time step one individual is selected uniformly at random, and the selected individual with probability $1-\varepsilon \geq \frac{1}{2}$ adopts the best response, (if more than one best responses exist, then the current opinion will be selected if it is a best response, while an arbitrary best response will be chosen otherwise), and with remaining probability she will adopt a randomly chosen opinion. Hence, supposing, for simplicity, that $B = S = \{0,1\}$ we can describe the process, by setting the probability $P_i^k(\mathbf{x})$ that the individual i selected for update after k steps adopts opinion 1 given that the current opinion profile is $\mathbf{x}$, as follows:

$$P_i^k(\mathbf{x}) = \begin{cases} 1-\varepsilon, & \text{if } f(1-b_i) - f(0-b_i) + \rho^{(k)}g(1)(n-2\ell(\mathbf{x})+1) \geq 0 \text{ and } x_i = 1; \\ 1-\varepsilon, & \text{if } f(1-b_i) - f(0-b_i) + \rho^{(k)}g(1)(n-2\ell(\mathbf{x})-1) > 0 \text{ and } x_i = 0; \\ \varepsilon, & \text{otherwise.} \end{cases}$$

It is then possible to define the anonymous slowed-down process Q_i^k as follows: $Q_i^k(\mathbf{x}) = P_i^k(\mathbf{x})$ if (i) $k < k^\star$ where $k^* = \min\{k \mid \rho^{(k)} > \frac{f(1)}{g(1)}\}$, or (ii) $x_i = 1$ and $\ell(x) \neq \frac{n+1}{2}$, or (iii) $x_i = 0$ and $\ell(x) \neq \frac{n-1}{2}$; otherwise, we have $Q_i^k(\mathbf{x}) = \varepsilon$.

The anonymous process can then be converted in the birth and death chain such that $p_\ell = \frac{n-\ell}{n}(1-\varepsilon)$ if $\ell > \frac{n-1}{2}$ and $p_\ell = \frac{n-\ell}{n}\varepsilon$, otherwise, while $q_\ell = \frac{\ell}{n}\varepsilon$ if $\ell \leq \frac{n+1}{2}$ and $p_\ell = \frac{\ell}{n}(1-\varepsilon)$, otherwise.

It is, finally, immediate to bound the time that this birth and death chain takes to hit one of its extremes, just as done above. Hence, we can conclude with the following theorem.

Theorem 8. *If G is a clique, $S = B = \{0,1\}$, and $\rho^* > \frac{f(1)}{g(1)}$, then, for every $\mathbf{x}$, $\mathbf{E}_{\mathbf{x}}\left[\tau\right] \leq n^3 + k^*$, where $k^* = \min\{k \mid \rho^{(k)} > \frac{f(1)}{g(1)}\}$.*

Different selection rules. Next we consider the case that multiple individuals are selected at each time step for updating their opinions. It is not hard to see that the approach introduced in this work can be useful also in this setting. In particular, the reduction from the original process to the anonymous slowed-down process does not depend on the specific selection rule. Hence, we still can implement such a reduction. However, it is impossible to translate this new process into a birth and death chain. Still, since the process is anonymous, i.e., the update rules do not depend on the identities of individuals, it is still possible to use (possibly more complex) techniques to bound the time that the new process takes to reach the consensus. Note that both these bounds and the condition on ρ^* needed for achieving them, may turn out to be quite different from the ones in Theorem 4.

We highlight that if there is a sufficiently large probability of making a wrong choice when there is not a clear majority, then with noisy dynamics it is possible to reach the consensus quickly even for these selection rules for which convergence was impossible without noise.

4 Other Social Networks

We begin by characterizing the *stationary points* of our best-response dynamics, i.e., profiles in which no individual will have an incentive to change their opinion in any of the next steps of the dynamics. Given graph G, individual i and profile $\mathbf{x}$ we denote with $N_x^i(\mathbf{x})$, for $x \in S$, the number of neighbors of i in G adopting opinion x in $\mathbf{x}$.

Lemma 9. *If $\rho^* > \frac{f(1)}{g(1)}$, then the profile $\mathbf{x}$ can be a stationary point for the dynamics on $G = (V, E)$ if and only if for all individuals i and every opinion $y_i \neq x_i$ such that $N_{y_i}^i(\mathbf{x}) > 0$, it holds that*

$$N_{x_i}^i(\mathbf{x}) - N_{y_i}^i(\mathbf{x}) \geq \begin{cases} 1 & \textit{if } f(\operatorname{dist}(y_i, b_i)) < f(\operatorname{dist}(x_i, b_i)); \\ 0 & \textit{otherwise.} \end{cases}$$

Proof. Suppose that $\mathbf{x}^{(k)} = \mathbf{x}$ for some k such that $\rho^{(k)} > \frac{f(1)}{g(1)}$. We have that $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)}$ if and only if for all $i \in V$, and for every $y_i \neq x_i$, $c_i(\mathbf{x}^{(k)}) \leq c_i(y_i, \mathbf{x}_{-i}^{(k)})$. This is equivalent to

$$N_{x_i}^i(\mathbf{x}) - N_{y_i}^i(\mathbf{x}) \geq \frac{f(\operatorname{dist}(x_i, b_i)) - f(\operatorname{dist}(y_i, b_i))}{\rho^{(k)} g(1)}. \tag{2}$$

We first observe that this implies that $N_{x_i}^i(\mathbf{x}) > 0$. Suppose, instead, that $N_{x_i}^i(\mathbf{x}) = 0$. Then, since G is connected, it must exist y_i such that $N_{x_i}^i(\mathbf{x}) \geq 1$. Then, for this pair of opinions, it turns out that the LHS of (2) is at most -1. Hence, the condition cannot be satisfied, since the RHS is at least $-\frac{f(1)}{\rho^{(k)} g(1)} > -1$, where the inequality follows from $\rho^{(k)} > \frac{f(1)}{g(1)}$.

However, $N_{x_i}^i(\mathbf{x}) > 0$, and thus $N_{x_i}^i(\mathbf{x}) \geq 1$, in turn implies that (2) is trivially satisfied when $N_{y_i}^i(\mathbf{x}) = 0$, since the RHS is at most $\frac{f(1)}{\rho^{(k)} g(1)} < 1$, where the inequality follows again from $\rho^{(k)} > \frac{f(1)}{g(1)}$.

Consider then $y_i \neq x_i$ such that $N_{y_i}^i(\mathbf{x}) > 0$. If $f(\operatorname{dist}(y_i, b_i)) \geq f(\operatorname{dist}(x_i, b_i))$, then the RHS of (2) is in $\left[-\frac{f(1)}{\rho^{(k)} g(1)}, 0\right] \subseteq (-1, 0)$, since $\rho^{(k)} > \frac{f(1)}{g(1)}$. However, since $N_x^i(\mathbf{x}^{(k)})$ and $N_y^i(\mathbf{x}^{(k)})$ are integers, then in this case (2) is equivalent to the desired condition.

If $f(\operatorname{dist}(y_i, b_i)) < f(\operatorname{dist}(x_i, b_i))$, then the RHS of (2) is in $\left(0, \frac{f(1)}{\rho^{(k)} g(1)}\right] \subseteq (0, 1)$, and thus even in this case (2) is equivalent to the desired condition.

Moreover, since $\rho^{(k+1)} \geq \rho^{(k)}$, the inequality will hold also at round $k+1$; then, by induction, $\mathbf{x}$ is stationary. □

Note that a consensus profile $\mathbf{x}$ is always a stationary point for the dynamics, no matter the graph G, as $N_y^i(\mathbf{x}) = 0$ for all i and all $y \neq x_i$. However, for some graphs G, there might be additional stationary points that are different from consensus – in such cases, we say that the dynamics on G *diverges*.

Given $G = (V, E)$, for $v \in V$ and $A \subset V$, we let $N_v(A) = |\{j \in A \mid (v, j) \in E\}|$. We say that G is *well-partitioned* if V can be partitioned in sets $V_0, V_1, \ldots, V_{m-1}$, with $m > 1$, such that, for all $b, c \in \{0, 1, \ldots, m-1\}$ with $b \neq c$ and $v \in V_b$, $N_v(V_b) \geq N_v(V_c)$.

Theorem 10. *If $\rho^* > \frac{f(1)}{g(1)}$, the dynamics on G diverges if and only if G is well-partitioned.*

Proof. Let us start by proving that whenever G is well-partitioned then there is an instance on which the dynamics on G diverges. Indeed, by hypothesis, there exists a partition of V in $V_0, V_1, \ldots, V_{m-1}$, with $m > 1$ such that for all $v \in V_b$, $N_v(V_b) \geq N_v(V_c)$, for every $b, c \in \{0, 1, \ldots, m-1\}$ with $b \neq c$. We can then naturally define $B = S = \{0, 1, \ldots, m-1\}$ and the belief of $i \in V_b$ to be b; moreover we can also set $\rho^{(0)} = \rho^*$ and $f(\alpha) = f(1)$ for every $\alpha > 0$. Therefore, by Lemma 9 we can conclude that $\mathbf{x}^{(0)}$ is a stationary point different from consensus.

For the other direction, let $\mathbf{x}$ be a stationary point of the dynamics on G different from consensus. Since $\rho^* > \frac{f(1)}{g(1)}$, Lemma 9 yields a partition of the vertices of the graph as requested, i.e., $i \in V_b$ iff $x_i = s_b$ where we rewrite w.l.o.g. the opinion set as $S = \{s_0, \ldots, s_{m-1}\}$. □

Basically, Theorem 10 proves that individuals are content of having reached consensus within their own cluster/community, and they do not care for general consensus.

Note that, as shown in Claim 1.1, $\rho^* > \frac{f(1)}{g(1)}$ is necessary in order to incentivize consensus.

The Price of Divergence. If the dynamics on a graph diverges, then it is possible that the opinions of the individuals fail to converge to consensus. However, one can still ask whether pressure is useful in these cases to get as close as possible to this goal. Indeed, if pressure is absent or too low, then it is possible that a stationary point of the dynamics consists of every individual having a different opinion (for example, when $f(1) = g(1) = 1$ and each individual has a different belief). Then, it is natural to ask whether pressure enables the system to converge to a stationary point in which only few different opinions are adopted.

Unfortunately, we show that this is not the case. Indeed, consider a cycle with nodes $0, 1, \ldots, n-1$, with even n. For $i = 0, \ldots, n/2 - 1$, assume that individuals $2i$ and $2i+1$ have belief b_i. Hence, in $\mathbf{x}^{(0)}$ individuals adopt $n/2$ different opinions and none of them is adopted by a majority. Moreover, if $\rho^{(0)} > \frac{f(1)}{g(1)}$, then the profile $\mathbf{x}^{(0)}$ is a stationary point of the dynamics, since it satisfies the conditions of Lemma 9. Thus, even with a strong pressure it is still possible to have $O(n)$ different opinions at equilibrium.

On the other hand, it is not hard to see that if there are no isolated nodes, the example above cannot be pushed further (and thus it is not possible to achieve n different opinions at equilibrium as in the case without high enough pressure). Indeed, at equilibrium every non-isolated node must have at least one neighbor with the same opinion, for otherwise Lemma 9 would not be satisfied. Hence, the total number of different opinions that can be adopted in a stationary point cannot be larger than $n/2$.

4.1 Understanding Well-Partitioned Graphs

The result above provides a characterization of the social networks on which the dynamics may diverge in terms of well-partitioned graphs. However, it seems hard to give a more explicit and topological characterization of these graphs. The related literature on "clustering" of graphs [Gharan and Trevisan, 2014, Peng et al., 2015, Kolev and Mehlhorn, 2016], for example, focuses only on algorithmic characterizations (wherein, naturally, algorithms are intended to run in polynomial-time). We now follow a similar avenue.

Determining whether a graph $G = (V, E)$ is well-partitioned is, in fact, equivalent to the graph having a *non-singleton locally-minimal cut*, i.e., a partition (L, R) of vertices V such that $|L|, |R| > 1$ and there is no vertex $v \in L$ ($v \in R$, respectively) such that $E(L \setminus \{v\}, R \cup \{v\}) < E(L, R)$ ($E(L \cup \{v\}, R \setminus \{v\}) < E(L, R)$, resp.), where $E(A, B) = |\{(u, v) \in E\colon u \in A, v \in B\}|$. Roughly speaking, a non-singleton locally-minimal cut is a cut such that each side contains at least two elements, and moving one vertex to the another set of the partition does not diminish the weight of the cut.

Theorem 11. *A connected graph G is well-partitioned if and only if G has a* non-singleton locally minimal cut *of a graph $G = (V, E)$.*

Proof. Consider first the "if" direction. Let $(A, V \setminus A)$ be a non-singleton locally minimal cut of G. We can then set $m = 2$ and use $V_0 = A$ and $V_1 = V \setminus A$ as the partition needed by the definition of well-partitioned graphs.

Assume, indeed, that this partition does not witness that G is well-partitioned. Then there exists $b \in \{0, 1\}$ and $v \in V_b$ such that $N_v(V_b) < N_v(V_{1-b})$. But then moving v from V_b to V_{1-b} would give rise to a new cut of smaller size; a contradiction with the fact that $(A, V \setminus A)$ is locally minimal.

As for the "only if" direction, consider a graph G that does not have a non-singleton locally minimal cut (i.e., all locally minimal cuts have one side being a singleton). Suppose that a partition $(V_0, \ldots, V_{m-1})$ of G exists that witnesses that G is well-partitioned. Consider, then, the cut $(A, V \setminus A)$ such that $A = V_0$. Since G is well-partitioned, then for each $v \in A$, it turns out that $N_v(A) \geq N_v(V \setminus A)$, and thus the size of the cut cannot decrease if v is moved from A to $V \setminus A$. Similarly, for every $v \in V_c \subseteq V \setminus A$, with $c = 1, \ldots, m$, we have that $N_v(V \setminus A) \geq N_v(V_c) \geq N_v(A)$, and thus the size of the cut cannot decrease if v is moved from $V \setminus A$ to A. Then, we conclude that $(A, V \setminus A)$ is a locally minimal cut, and it must be a non singleton cut, since every partition of a well-partitioned connected graph G must contain at least two elements. However, this is a contradiction. □

Unfortunately, deciding if a graph contains a non-singleton locally minimal cut has been recently proven to be an NP-complete problem [Auletta et al., 2018].

In order to bypass this hardness result, one can focus on *non-singleton minimum cut*. Indeed, there is a polynomial-time algorithm to establish whether a graph has a non-singleton minimum cut: we can use the algorithm of Karger [1993] to enumerate (with high probability) all of the (roughly $|V|^2/2$) min-cuts (of an undirected, unweighted graph) in polynomial-time; then we can simply check the size of both ends of the cut.

Clearly, a non-singleton minimum cut is also locally minimal, thus its existence is sufficient to conclude that G is well-partitioned. However, such a condition is not necessary already for $m = 2$, as we are going to discuss next. For an even number n, let G be the clique K_n with a perfect matching removed, i.e., $E = \{(i, j) : j \neq i\} \setminus \{(2i, 2i - 1) : n/2 \geq i \geq 1\}$. Consider the partition in which even vertices are in V_0 and odd vertices in V_1. Each vertex has exactly half of its neighbors in V_b and exactly half in V_{1-b} for $b = 0, 1$. By Theorem 10, the dynamics diverges on G. However, G does not have a non-singleton min-cut, since all the min-cuts of G are of the kind $(\{v\}, V \setminus \{v\})$, $v \in V$.

4.2 Convergence to Majority

Whenever convergence to consensus cannot be achieved, one may ask whether the dynamics still converges to some weaker form of agreement, as, for example, an opinion being adopted by a strict majority of the population. Unfortunately, time pressure can be ineffective also to achieve this goal. Consider, indeed, the cycle example discussed above. It is immediate to see that at equilibrium, each of the $n/2$ opinions is supported by only 2 individuals, regardless the pressure.

It is not hard to see that we can characterize the social networks on which convergence to majority is allowed in a way similar to what we did above for consensus. Specifically, given $G = (V, E)$, we say that G is *majority-partitioned* if it is well-partitioned and, moreover, there are $i, j \in \{0, \dots, m-1\}$ such that $|V_i| = |V_j| \geq |V_k|$ for every $k \neq i, j$. Similarly, G is *absolutely-partitioned* if it is well-partitioned and, moreover, $|V_i| \leq n/2$ for every $i \in \{0, \dots, m-1\}$. Then it is immediate to see that we can extend Theorem 10 to prove the following characterization of social networks for which pressure to consensus always leads one opinion to be adopted by a strict majority (absolute majority, respectively) of individuals.

Theorem 12. *If $\rho^* > \frac{f(1)}{g(1)}$, there is an instance such that the dynamics on G converge to a stationary point in which no opinion is supported by strictly more individuals than any other opinion if and only if G is majority-partitioned.*

Similarly, if $\rho^* > \frac{f(1)}{g(1)}$, there is an instance such that the dynamics on G converge to a stationary point in which no opinionis supported by the absolute majority of individuals if and only if G is absolutely-partitioned.

Unfortunately, it is not hard to see that the reduction given in Auletta et al. [2018] for proving that it is hard to decide whether a graph is well-partitioned can be opportunely padded with isolated vertices keeping the opportune opinion, to prove that it is NP-hard also to decide whether a graph is majority-partitioned or absolutely-partitioned.

5 A case study: Brexit

A particularly interesting case study comes from Brexit, and more specifically the negotiations under Article 50 of the Treaty on European Union (EU) for the withdrawal of the UK from the EU. The negotiations have to end by March 29, 2019 (two years from UK's notification) with an agreement; in absence of any, the so-called "no deal" outcome would arguably cause significant disruption to the UK and the EU. Our findings can be applied to Brexit and can explain the reasons behind an absence of an agreement to date. Since the two-year deadline is a device to exert pressure to converge, we can use our model and results to understand and motivate the evolution of the negotiation and the (current absence of) an outcome.

Linking back to our model, the set of beliefs B could contain all options ranging from EU membership to leaving on WTO terms, whilst S, the set of opinions, can be a yes/no answer to a confidence vote in the Prime Minister. The distance between a belief and a different opinion can be assumed to be 1, especially in the current polarized environment of the politics (in the UK). As discussed above, we leave open the problem to study less strict notions of distance, e.g., the distance between WTO and EU membership should be higher than the distance between EU and EEA membership; such a study would help explain whether more accommodating positions between people with different opinions would have led to an agreement more quickly.

The EU side has quickly reached a consensus on the goals that needed to be satisfied by the withdrawal agreement, within the international legal framework. Many political commentators have used the word 'unity' to describe the EU in the process. We can translate this into our framework and say that the social network representing the EU member states is a clique, and, in turn, consensus has been reached quickly as theorized by our results.

At the time of writing, the UK is however still negotiating internally despite the no deal deadline being few weeks away, with the consequent pressure to find a solution. There are different factions within the parties in the House of Commons, each holding a different and strong belief about what the future relationship with the EU should be; this has led to a deadlock which is preventing from reaching consensus (or even a simple majority). To map this state of affairs to our findings, we can note how the social network underlying the UK parliament is far from being a clique and is actually clustered in a number of groups. This explains why despite the extremely high pressure, there is still no consensus. Actually, some of the recent political moves (splinters from both the main parties uniting to form a new group) aim at modifying the social network in order to break the deadlock. Our results explain that this strategy will be successful as long as the new graph will have no locally minimum cuts.

6 Conclusions

In this work we initiated the study of opinion dynamics with a social pressure towards consensus. For clique social networks, we have been able to give a complete picture of what happens both with fully rational individuals and bounded rational ones. Much more is left to understand for different social network topologies: Can we bound the time that the dynamics takes in order to converge to consensus or to a generic stationary point? Does the convergence to consensus in non-clique graphs take a path similar to the one described by Claim 1.2, with the majority opinion becoming more and more supported as time goes on? Simulations may be an useful tool for answering this question.

In this work we focused on unweighted graphs. Naturally, it would be interesting to see to which extent our results hold on weighted networks. Note that, for example, it is not hard to see that for specific weight assignment of edges, the best response dynamics does not converge to consensus even if the underlying graph is a clique.

It would be interesting also to evaluate how bounded rationality influences the evolution of opinions in non-clique network topologies. Do noisy dynamics, such as logit dynamics, converge to the consensus whenever best response does?

Acknowledgements

A preliminary version of this paper appeared as [Ferraioli and Ventre, 2017]. Diodato Ferraioli was supported by "GNCS–INdAM". Carmine Ventre was supported by the EPSRC grant EP/M018113/1.

References

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

Vincenzo Auletta, Diodato Ferraioli, Francesco Pasquale, and Giuseppe Persiano. Metastability of logit dynamics for coordination games. In *SODA '12*, pages 1006–1024, 2012.

Vincenzo Auletta, Diodato Ferraioli, Francesco Pasquale, Paolo Penna, and Giuseppe Persiano. Logit dynamics with concurrent updates for local interaction games. In *ESA '13*, pages 73–84, 2013a.

Vincenzo Auletta, Diodato Ferraioli, Francesco Pasquale, and Giuseppe Persiano. Mixing time and stationary expected social welfare of logit dynamics. *Theory of Computing Systems*, 53 (1):3–40, 2013b.

Vincenzo Auletta, Ioannis Caragiannis, Diodato Ferraioli, Clemente Galdi, and Giuseppe Persiano. Minority becomes majority in social networks. In *WINE '15*, pages 74–88, 2015.

Vincenzo Auletta, Ioannis Caragiannis, Diodato Ferraioli, Clemente Galdi, and Giuseppe Persiano. Generalized discrete preference games. In *IJCAI '16*, pages 53–59, 2016.

Vincenzo Auletta, Diodato Ferraioli, and Gianluigi Greco. Reasoning about consensus when opinions diffuse through majority dynamics. In *IJCAI*, pages 49–55, 2018.

Vincenzo Auletta, Angelo Fanelli, and Diodato Ferraioli. Consensus in opinion formation processes in fully evolving environments. In *AAAI'19*, 2019.

Anandd Bhalgat, Tanmoy Chakraborty, and Sanjeev Khanna. Approximating pure Nash equilibrium in cut, party affiliation, and satisfiability games. In *EC '10*, pages 73–82, 2010.

Kshipra Bhawalkar, Sreenivas Gollapudi, and Kamesh Munagala. Coevolutionary opinion formation games. In *STOC '13*, pages 41–50, 2013.

Vittorio Bilò, Angelo Fanelli, and Luca Moscardelli. Opinion formation games with dynamic social influences. In *WINE '16*, pages 444–458, 2016.

David Bindel, Jon Kleinberg, and Sigal Oren. How bad is forming your own opinion? In *FOCS '11*, pages 55–66, 2011.

Lawrence E. Blume. The statistical mechanics of strategic interaction. *Games and Economic Behavior*, 5:387–424, 1993.

Flavio Chierichetti, Jon Kleinberg, and Sigal Oren. On discrete preferences and coordination. *Journal of Computer and System Sciences*, 93:11–29, 2018.

Morris H. DeGroot. Reaching a consensus. *J. American Statistical Association*, 69:118–121, 1974.

Diodato Ferraioli and Carmine Ventre. Social pressure in opinion games. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, pages 3661–3667. AAAI Press, 2017.

Diodato Ferraioli, Paul W. Goldberg, and Carmine Ventre. Decentralized dynamics for finite opinion games. *Theoretical Computer Science*, 648:96 – 115, 2016.

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

Shayan Oveis Gharan and Luca Trevisan. Partitioning into expanders. In *SODA '14*, pages 1256–1266, 2014.

Umberto Grandi, Emiliano Lorini, and Laurent Perrussel. Propositional opinion diffusion. In *AAMAS '15*, pages 989–997, 2015.

David R. Karger. Global min-cuts in rnc, and other ramifications of a simple min-cut algorithm. In *SODA '93*, pages 21–30, 1993.

Pavel Kolev and Kurt Mehlhorn. A note on spectral clustering. In *ESA '16*, pages 57:1–57:14, 2016.

Daniel McFadden. *Conditional logit analysis of qualitative choice behavior*, pages 105–142. Number 2. 1974.

Andrea Montanari and Amin Saberi. Convergence to equilibrium in local interaction games. In *FOCS '09*, pages 303–312, 2009.

Elchanan Mossel and Omer Tamuz. Opinion exchange dynamics. *Probability Surveys*, 14:155–204, 2017.

Richard Peng, He Sun, and Luca Zanetti. Partitioning well-clustered graphs: Spectral clustering works! In *COLT '15*, pages 1423–1455, 2015.

Hobart Peyton Young. The diffusion of innovations in social networks. *The economy as an evolving complex system III: Current perspectives and future directions*, 267, 2006.

Gabriella Pigozzi. Belief merging and judgment aggregation. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2016 edition, 2016.

Oleksandr Pryymak, Alex Rogers, and Nicholas R. Jennings. Efficient opinion sharing in large decentralised teams. In *AAMAS '12*, pages 543–550, 2012.

Sarah Rajtmajer, Anna Squicciarini, Christopher Griffin, Sushama Karumanchi, and Alpana Tyagi. Constrained social energy minimization for multi-party sharing in online social networks. In *AAMAS '16*, pages 680–688, 2016.

Nicolas Schwind, Katsumi Inoue, Gauvain Bourgne, Sebastien Konieczny, and Pierre Marquis. Belief revision games. In *AAAI'15*, pages 1590–1596, 2015.

Alan Tsang and Kate Larson. Opinion dynamics of skeptical agents. In *AAMAS '14*, pages 277–284, 2014.