

Opinion Formation Games with Aggregation and Negative Influence^{*}

Markos Epitropou¹, Dimitris Fotakis², Martin Hoefer³, and
Stratis Skoulakis²

¹ Dept. of Electrical and Systems Engineering, University of Pennsylvania, USA
`mep@seas.upenn.edu`

² School of Electrical and Computer Engineering, NTU Athens, Greece
`fotakis@cs.ntua.gr`, `sskoul@corelab.ntua.gr`

³ Institut für Informatik, Goethe-Universität Frankfurt/Main, Germany
`mhoefer@cs.uni-frankfurt.de`

Abstract. We study continuous opinion formation games with aggregation aspects. In many domains, expressed opinions of people are not only affected by local interaction and personal beliefs, but also by influences that stem from global properties of the opinions in the society. To capture the interplay of such global and local effects, we propose a model of opinion formation games with aggregation, where we concentrate on the *average public opinion* as a natural way to represent a global trend in the society. While the average alone does not have good strategic properties as an aggregation rule, we show that with a reasonable influence of the average public opinion, the good properties of opinion formation models are preserved. More formally, we prove that a unique equilibrium exists in average-oriented opinion formation games. Simultaneous best-response dynamics converge to within distance ε of equilibrium in $O(n^2 \ln(n/\varepsilon))$ rounds, even in a model with *outdated information* on the average public opinion. For the Price of Anarchy, we show a small bound of $9/8 + o(1)$, almost matching the tight bound for games without aggregation. Moreover, some of the results apply to a general class of opinion formation games with negative influences, and we extend our results to the case where expressed opinions come from a restricted domain.

1 Introduction

The formation and dynamics of opinions are an important aspect in society and have been studied extensively for decades (see e.g., [16]). Opinion formation is based on information exchange, which is often *local* in the sense that socially connected people (e.g., family, friends, colleagues) interact more often and affect each other's opinion more strongly. Moreover, opinion formation is often *dynamic* in the sense that discussions and interactions lead to changes in the expressed opinions. With the advent of the internet and social media, local and

^{*} This work was supported by DFG Cluster of Excellence MMCI at Saarland University. Stratis Skoulakis is supported by a scholarship from the Onassis Foundation.

dynamic aspects of opinion formation have become ever more dominant. To capture opinion formation on a formal level, several models have been proposed (see e.g., [9, 12, 15, 6, 13, 4] for continuous opinions and [10, 20, 5] for discrete ones).

Motivation and Opinion Formation Model. We build on the influential model of Friedkin and Johnsen (FJ) [12] for continuous opinion formation, following the game-theoretic viewpoint of [6]. Each agent i holds an *intrinsic belief* $s_i \in [0, 1]$, which is private and invariant over time, and a *public opinion* $z_i \in [0, 1]$. Agent i selects her opinion so as to minimize the total (weighted) disagreement of z_i to her belief and to the opinions in her social neighborhood. In a dynamic setting, the agents start with their beliefs and in each round $t \geq 1$, update their opinion $z_i(t)$ to the minimizer of their disagreement cost, given the opinions of the others in round $t - 1$. The FJ model is extensively studied and has nice algorithmic properties. It admits a unique equilibrium [12, 6], which is approached quickly by the simultaneous best-response dynamics [13]. The Price of Anarchy (PoA) is $9/8$ for undirected social networks and $\Omega(n)$ for general directed networks [6]. Moreover, tight PoA bounds can be obtained by an elegant local smoothness argument both for undirected [4] and for directed [8] networks.

Despite these favorable properties, the FJ model disregards influences from global properties of the opinions, and also the nature of the dynamics of consensus formation. In many domains, public opinions are not only affected by local interaction and personal beliefs, as in e.g., [9, 12, 6, 13, 4, 7], but also by influences that stem from global properties of the opinions in the society. People are getting exposed to global trends, societal norms, results from voting and polling, etc., which are usually interpreted as the consensus view of the society and may affect opinion formation. Furthermore, groups of people (or networks of agents) often need to agree on a common action, even if their beliefs and/or their expressed opinions are totally different. This happens, e.g., when networked devices need to implement a common action, when people vote over a set of alternatives, or when a wisdom-of-the-crowd opinion is formed in a social network. In similar situations, an *aggregation rule* maps the public opinions to a *global* opinion that represents the consensus view on the issue at hand. E.g., in the FJ model, the global opinion might be the average or the median of the equilibrium opinions.

In presence of aggregation, an agent can also anticipate the impact of its chosen public opinion on the global one and might incorporate it in its choice. Hence, the disagreement cost should also reflect the distance of an agent's intrinsic belief to the global opinion. To address these issues, we consider a variant of the opinion formation game of [12, 6, 13] with opinion aggregation. Each agent i selects her opinion z_i so as to minimize:

$$C_i(\mathbf{z}) = w_i(s_i - z_i)^2 + \sum_{j \neq i} w_{ij}(z_j - z_i)^2 + \alpha_i(\text{aggr}(\mathbf{z}) - s_i)^2 . \quad (1)$$

In (1), $\mathbf{z} = (z_1, \dots, z_n) \in \mathbb{R}^n$ is the public opinion vector, $s_i \in [0, 1]$ is the belief of agent i , and $\text{aggr}(\mathbf{z})$ maps \mathbf{z} to a global opinion $\text{aggr}(\mathbf{z})$. The weights $w_{ij} \geq 0$ quantify how much the public opinion of agent j influences i , $w_i > 0$ quantifies i 's self-confidence, and $\alpha_i > 0$ quantifies the appeal of $\text{aggr}(\mathbf{z})$ to i .

Motivated by previous work on the wisdom of the crowd (see e.g., [16, Sec. 8.3], [14]), we concentrate on *average-oriented* opinion formation games, where the aggregation rule $\text{aggr}(\mathbf{z})$ maps \mathbf{z} to its average $\text{avg}(\mathbf{z}) = \sum_{j=1}^n z_j/n$. Then, the best response of each agent i to a public opinion vector \mathbf{z} is:

$$z_i = \frac{(w_i + \frac{\alpha_i}{n}) s_i + \sum_{j \neq i} (w_{ij} - \frac{\alpha_i}{n^2}) z_j}{w_i + \frac{\alpha_i}{n^2} + \sum_{j \neq i} w_{ij}} . \quad (2)$$

Contribution. The aggregation rule in (1) might significantly affect the dynamics and the equilibrium of opinion formation. This becomes evident in (2), where i 's influence from some opinions z_j can be negative. Negative influence models agent competition for dragging the average public opinion close to their intrinsic beliefs. An important side-effect is that the best-response (and equilibrium) opinions may become polarized and be pushed towards opposite directions, far away from the agent intrinsic beliefs. This is a significant departure from the FJ model, where the equilibrium opinions lie between the minimum and maximum intrinsic beliefs of the agents. Interestingly, we prove that the nice algorithmic properties of the FJ model are not affected – neither by negative influence nor by outdated information on the average opinion.

We show (Lemma 1) that average-oriented games admit a unique equilibrium, and simultaneous best-response dynamics converges to it within distance $\varepsilon > 0$ in $O(n^2 \ln(n/\varepsilon))$ rounds. For this result, all agents have access to the average public opinion in each round. Since the average is global information and thus difficult to monitor in large networks, we consider average-oriented opinion dynamics with outdated information. Here the average public opinion is announced to all the agents simultaneously every few rounds (e.g., a polling agency publishes this information every now and then). We prove (Theorem 1) that opinion dynamics with outdated information about the average converges to the unique equilibrium within distance $\varepsilon > 0$ after $O(n^2 \ln(n/\varepsilon))$ updates on the average. Both these results are proven for a more general setting with negative influence between the agents and with partially outdated information about the agent public opinions. The main point here is that negative influence and outdated information do not introduce undesirable oscillating phenomena to opinion dynamics.

In Section 4, we bound the PoA of average-oriented opinion formation games. We consider symmetric games, where $w_{ij} = w_{ji} \geq 0$ for all agent pairs $i \neq j$, all agents have the same self-confidence w and the same influence α from the average (for non-symmetric games the PoA is $\Omega(n)$, even without aggregation, see [6, Fig. 2]). We show (Theorem 2) that the PoA is at most $9/8 + O(\alpha/(wn^2))$. In general, this bound cannot be improved since for $\alpha = 0$, $9/8$ is a tight bound for the PoA under the FJ model [6]. While the proof builds on [4], local smoothness cannot be directly applied to symmetric average-oriented games, because the function $(\text{avg}(\mathbf{z}) - s_i)^2$ is not locally smooth. To overcome this difficulty, we combine local smoothness with the fact that the average opinion at equilibrium is equal to the average belief, a consequence of symmetry (Proposition 1).

A frequent assumption on continuous opinion formation is that agent beliefs and opinions take values in a finite interval of non-negative real numbers. By

scaling, one can then treat beliefs and opinions as numbers in $[0, 1]$. Here, we also assume that agent beliefs $s_i \in [0, 1]$. However, due to negative influence, the equilibrium opinions in our model may become polarized and end up far away from $[0, 1]$. We believe that such opinion polarization is natural and should be allowed under negative influence. Therefore, in Sections 3 and 4, we assume that public opinions can take arbitrary real values. Then, in Section 5, we also consider *restricted* average-oriented games with public opinions restricted to $[0, 1]$, and study how convergence properties and price of anarchy are affected.

Existence and uniqueness of equilibrium for restricted games follow from [18]. We prove (Theorem 3) that the convergence rate of opinion dynamics with negative influence and with outdated information is not affected by restriction of public opinions to $[0, 1]$. As for the PoA of restricted symmetric games, we consider the special case where $w_i = \alpha_i = 1$, for all agents i , and show that the PoA does not exceed $(\sqrt{2} + 2)^2/2 + O(\frac{1}{n})$ (Theorem 4). A technical challenge is that partial derivatives of the agent cost functions in the local smoothness inequality do not need to be 0 at equilibrium, due to the opinion restriction to $[0, 1]$. So, we first show that if $w_{ij} = 0$ for all $i \neq j$, the PoA is at most $1 + 1/n^2$. Then we combine the PoA of this simpler game with the local smoothness inequality of [4] and obtain an upper bound on the PoA of the restricted game. Due to lack of space, several proofs are omitted from this extended abstract.

Clearly, there are many alternative ways to model aggregation, which offer interesting directions for future research. For example, a possible aggregation is the *median* instead of the average. The median aggregation rule is prominent in Social Choice (see e.g., [17, 3]). However, it turns out that the FJ model with median aggregation has significantly less favorable properties. There are examples where median-oriented games lack exact equilibria (and, hence, convergence of best-response dynamics), but they can be shown to have approximate equilibria. A study of the median rule is beyond the scope of this paper.

Further Related Work. To the best of our knowledge, this is the first work to analyze the convergence of simultaneous best-response dynamics of the FJ model with negative influence and outdated information, or the price of anarchy of the FJ model with average opinion aggregation. However, there is some recent work on properties of opinion formation either with global information, or with negative influence, or where consensus is sought. We concentrate here on related previous work most relevant to ours. Discrete opinion formation is considered in [11] in the binary voter model, where each agent i has a certain probability of adopting the opinion of an agent outside i 's local neighborhood (this is conceptually equivalent to estimating the average opinion with random sampling). The authors analyze the convergence time and the probability that consensus is reached. Necessary and sufficient conditions under which local interaction in social networks with positive and negative influence reaches consensus are derived in [1]. Recently, a model of discrete opinion formation was introduced in [2] with generalized social relations, which include positive and negative influence. The authors show that generalized discrete opinion formation games admit a potential function, and thus, best-response dynamics converge to a Nash equilibrium.

2 Model and Preliminaries

Notation and Conventions. We define $[n] \equiv \{1, \dots, n\}$. For a vector \mathbf{z} , \mathbf{z}_{-i} is \mathbf{z} without its i -th coordinate and (z, \mathbf{z}_{-i}) is the vector obtained from \mathbf{z} if we replace z_i with z . Let $\mathbf{0} \equiv (0, \dots, 0)$ and \mathbb{I} be the $n \times n$ identity matrix. We use capital letters for matrices and lowercase letters for their elements such that, e.g., a_{ij} is the (i, j) element of a matrix A .

$\|A\|$ and $\|\mathbf{z}\|$ denote the infinity norms of matrix A and vector \mathbf{z} , resp. We repeatedly use the standard properties of matrix norms without explicitly referring to them, i.e., (i) for any matrices A and B , $\|AB\| \leq \|A\| \|B\|$ and $\|A + B\| \leq \|A\| + \|B\|$; (ii) for any matrix A and any $\lambda \in \mathbb{R}$, $\|\lambda A\| \leq |\lambda| \|A\|$; and (iii) for any matrix A and any integer ℓ , $\|A^\ell\| \leq \|A\|^\ell$. Moreover, we use that for any $n \times n$ real matrix A with $\|A\| < 1$, $\sum_{\ell=0}^{\infty} A^\ell = (\mathbb{I} - A)^{-1}$.

Average-Oriented Opinion Formation. We consider average-oriented opinion formation games with n agents as introduced in Section 1. Wlog., we assume that agent beliefs $\mathbf{s} \in [0, 1]^n$. For the public opinions \mathbf{z} , we initially assume values in \mathbb{R} . In Section 5, we explain what changes if we restrict them to $[0, 1]$. An average-oriented game \mathcal{G} is *symmetric* if $w_{ij} = w_{ji}$ for all $i \neq j$, and $w_i = w$ and $\alpha_i = \alpha$ for all agents i . \mathcal{G} is *nonsymmetric* otherwise. If \mathcal{G} is symmetric, we let $w = 1$, by scaling other weights accordingly. Our convergence results hold for nonsymmetric games, our PoA bounds hold only for symmetric ones.

A vector \mathbf{z}^* is an *equilibrium* of an opinion formation game \mathcal{G} if for any agent i and any opinion z , $C_i(\mathbf{z}^*) \leq C_i(z, \mathbf{z}_{-i}^*)$, i.e., the agents cannot improve their individual cost at \mathbf{z}^* by unilaterally changing their opinions. The *social cost* $C(\mathbf{z})$ of \mathcal{G} is $C(\mathbf{z}) = \sum_{i \in N} C_i(\mathbf{z})$. An opinion vector \mathbf{o} is *optimal* if for any \mathbf{z} , $C(\mathbf{o}) \leq C(\mathbf{z})$. An optimal vector exists because the social cost function is proper. The *Price of Anarchy of \mathcal{G}* (PoA(\mathcal{G})) is $C(\mathbf{z}^*)/C(\mathbf{o})$, where \mathbf{z}^* is the unique equilibrium and \mathbf{o} an optimal vector.

It will be convenient to write (2) in matrix form. Let $S_i = w_i + \frac{\alpha_i}{n^2} + \sum_{j \neq i} w_{ij}$. We always assume that $S_i > 0$ so that $C_i(\mathbf{z})$ is strictly convex in z_i . We define two $n \times n$ matrices A and B . Matrix A has $a_{ii} = 0$, for all $i \in N$, and $a_{ij} = (w_{ij} - \frac{\alpha_i}{n^2})/S_i$, for all $j \neq i$. Matrix B is diagonal and has $b_{ii} = (w_i + \frac{\alpha_i}{n})/S_i$, for all $i \in N$, and $b_{ij} = 0$, for all $j \neq i$.

We assume that $\alpha_i \leq S_i \leq nw_i$, for all agents i (i.e., the agents neither are overwhelmed by the average opinion nor have extremely low self-confidence). This implies $\|A\| \leq 1 - \frac{2}{n^2}$, which is crucial for the convergence of best response dynamics. We term a matrix similar to A (i.e., with infinity norm less than 1 and 0s in its diagonal) *influence matrix*, and a matrix similar to B (i.e., diagonal one with positive elements) *self-confidence matrix*.

The simultaneous best-response dynamics of an average-oriented game \mathcal{G} starts with $\mathbf{z}(0) = \mathbf{s}$ and proceeds in rounds. In each round $t \geq 1$, the public opinion vector $\mathbf{z}(t)$ is:

$$\mathbf{z}(t) = A\mathbf{z}(t-1) + B\mathbf{s} \quad (3)$$

We refer to (3) and similar equations as *opinion formation processes*. An opinion formation process $\{\mathbf{z}(t)\}_{t \in \mathbb{N}}$ *converges* to a stable state \mathbf{z}^* if for all $\varepsilon > 0$, there

is a $t^*(\varepsilon)$, such that for all $t \geq t^*(\varepsilon)$, $\|\mathbf{z}(t) - \mathbf{z}^*\| \leq \varepsilon$. Iterating (3) over t (see also [13, Sec. 2]) implies that for all rounds $t \geq 1$,

$$\mathbf{z}(t) = A\mathbf{z}(t-1) + B\mathbf{s} = \dots = A^t\mathbf{s} + \sum_{\ell=0}^{t-1} A^\ell B\mathbf{s} . \quad (4)$$

Outdated Information of the Average Opinion. We study opinion formation when the agents have outdated information about the average public opinion. There is an infinite increasing sequence of rounds $0 = \tau_0 < \tau_1 < \tau_2 < \dots$ that describes an *update schedule* for the average opinion. At the end of round τ_p , the average $\text{avg}(\mathbf{z}(\tau_p))$ is announced to the agents. We refer to the rounds between two updates as an *epoch*, where rounds $\tau_p + 1, \dots, \tau_{p+1}$ comprise epoch p . The length of each epoch p , denoted by $k_p = \tau_{p+1} - \tau_p \geq 1$, is assumed to be finite. The update schedule is the same for all agents, but the agents might not be aware of it. They are only assumed to be aware of the most recent value of the average public opinion provided to them.

In this case, we need to distinguish in (2) and (3) between the influence from social neighbors, for which the most recent opinions $\mathbf{z}(t-1)$ are used, and the influence from the average public opinion, where possibly outdated information is used. As such, we now rely on three different $n \times n$ matrices D , E and B . Self-confidence matrix B is defined as before. Influence matrix D has $d_{ii} = 0$, for all $i \in [n]$, and $d_{ij} = w_{ij}/S_i$, for all $j \neq i$, and accounts for the influence from social neighbors. Influence matrix E has $e_{ii} = 0$, for all $i \in [n]$, and $e_{ij} = -\alpha_i/(n^2 S_i)$, for all $j \neq i$, and accounts for the influence from the average public opinion. By definition, $A = D + E$. Moreover, $\|D\| \leq 1 - 1/n$ and that $\|E\| \leq (n-1)/n^2$.

At the beginning of the opinion formation process, $\mathbf{z}(0) = \mathbf{s}$. For each round t in epoch p , $\tau_p + 1 \leq t \leq \tau_{p+1}$, the agent opinions are updated according to:

$$\mathbf{z}(t) = D\mathbf{z}(t-1) + E\mathbf{z}(\tau_p) + B\mathbf{s} \quad (5)$$

Note that at the beginning of each epoch p , every agent i can subtract $z_i(\tau_p)$ from $n \text{avg}(\mathbf{z}(\tau_p))$ and compute the term $E\mathbf{z}(\tau_p)$ as $-\frac{\alpha_i}{n^2 S_i} (n \text{avg}(\mathbf{z}(\tau_p)) - z_i(\tau_p))$.

Opinion Formation with Negative Influence. An interesting aspect of average-oriented games is that the influence matrix A may contain negative elements. Motivated by this observation, we prove our convergence results for a general domain of opinion formation games that may have negative weights w_{ij} . Similarly to [6, 13], the individual cost function of each agent i is $C_i(\mathbf{z}) = w_i(z_i - s_i)^2 + \sum_{j \neq i} w_{ij}(z_i - z_j)^2$, and i 's best response to \mathbf{z}_{-i} is

$$z_i = \frac{w_i s_i + \sum_{j \neq i} w_{ij} z_j}{w_i + \sum_{j \neq i} w_{ij}} . \quad (6)$$

The important difference is that now some w_{ij} may be negative. We require that for each agent i , $w_i > 0$ and $S_i = w_i + \sum_{j \neq i} w_{ij} > 0$ (and thus, $C_i(\mathbf{z})$ is strictly convex in z_i). The matrices A and B are defined as before. Namely, $a_{ij} = w_{ij}/S_i$, for all $i \neq j$, and B has $b_{ii} = w_i/S_i$ for all i . We always require that $\|A\| < 1 - \beta$, for some $\beta > 0$ (β may depend on n). Simultaneous best-response dynamics is again defined by (3).

3 Convergence of Average-Oriented Opinion Formation

For any nonnegative influence matrix A with largest eigenvalue at most $1 - \beta$, [13, Lemma 5] shows that (3) converges to $\mathbf{z}^* = (\mathbb{I} - A)^{-1}B\mathbf{s}$ within distance ε in $O(\ln(\frac{\|B\|}{\varepsilon\beta})/\beta)$ rounds. We generalize [13, Lemma 5] to average-oriented games, where A may contain negative elements. Thus, we show the following lemma for generalized opinion formation games with negative influence between the agents.

Lemma 1. *Let A be any influence matrix, possibly with negative elements, with $\|A\| \leq 1 - \beta$, for some $\beta > 0$. Then, for any self-confidence matrix B , any $\mathbf{s} \in [0, 1]^n$ and any $\varepsilon > 0$, the opinion formation process $\mathbf{z}(t) = A\mathbf{z}(t-1) + B\mathbf{s}$ converges to $\mathbf{z}^* = (\mathbb{I} - A)^{-1}B\mathbf{s}$ within distance ε in $O(\ln(\frac{\|B\|}{\varepsilon\beta})/\beta)$ rounds.*

Since $\mathbb{I} - A$ is nonsingular, \mathbf{z}^* is the unique vector that satisfies $\mathbf{z}^* = A\mathbf{z}^* + B\mathbf{s}$. Thus, \mathbf{z}^* is the unique equilibrium of the corresponding opinion formation game. Moreover, since for average-oriented games $\|A\| \leq 1 - 2/n^2$, Lemma 1 implies that any average-oriented game admits a unique equilibrium $\mathbf{z}^* = (\mathbb{I} - A)^{-1}B\mathbf{s}$, and for any $\varepsilon > 0$, (3) converges to \mathbf{z}^* within distance ε in $O(n^2 \ln(n/\varepsilon))$ rounds.

We next extend Lemma 1 to the case where the agents use possibly outdated information about the average public opinion in each round. In fact, we establish convergence for a general domain with negative influence between the agents, which includes average-oriented opinion formation processes as a special case.

Theorem 1. *Let D and E be influence matrices, possibly with negative elements, such that $\|D\| \leq 1 - \beta_1$, $\|E\| \leq 1 - \beta_2$, for some $\beta_1, \beta_2 \in (0, 1)$ with $\beta_1 + \beta_2 > 1$. Then, for any self-confidence matrix B , any $\mathbf{s} \in [0, 1]^n$, any update schedule $0 = \tau_0 < \tau_1 < \tau_2 < \dots$ and any $\varepsilon > 0$, the opinion formation process (5) converges to $\mathbf{z}^* = (\mathbb{I} - (D + E))^{-1}B\mathbf{s}$ within distance ε in $O(\ln(\frac{\|B\|}{\varepsilon\beta})/\beta)$ epochs, where $\beta = \beta_1 + \beta_2 - 1 > 0$.*

Proof. We observe that $\mathbf{z}^* = (\mathbb{I} - (D + E))^{-1}B\mathbf{s}$ is the unique solution of $\mathbf{z}^* = D\mathbf{z}^* + E\mathbf{z}^* + B\mathbf{s}$ (as in Lemma 1, since $\|E + D\| \leq 1 - \beta$, with $\beta > 0$, the matrix $\mathbb{I} - (D + E)$ is non-singular). Hence, if (5) converges, it converges to \mathbf{z}^* . To show convergence, we bound the distance of $\mathbf{z}(t)$ to \mathbf{z}^* by a decreasing function of t and show an upper bound on $t^*(\varepsilon) = \min\{t : e(t) \leq \varepsilon\}$.

As in the proof of Lemma 1, for each round $t \geq 1$, we define $e(t) = \|\mathbf{z}(t) - \mathbf{z}^*\|$ as the distance of the opinions at time t to \mathbf{z}^* . For convenience, we also define

$$f(\beta_1, \beta_2, k) = (1 - \beta_1)^k + (1 - \beta_2) \frac{1 - (1 - \beta_1)^k}{\beta_1}.$$

For any fixed value of $\beta_1, \beta_2 \in (0, 1)$ with $\beta_1 + \beta_2 > 1$, $f(\beta_1, \beta_2, k)$ is a decreasing function of k . Indeed, the derivative of f with respect to k is equal to $\ln(1 - \beta_1)(1 - \beta_1)^k(1 - \frac{1 - \beta_2}{\beta_1})$, which is negative, because $1 > (1 - \beta_2)/\beta_1$, since $\beta_1 + \beta_2 > 1$.

We next show that (i) for any epoch $p \geq 0$ and any round k , $0 \leq k \leq k_p$, in epoch p , $e(\tau_p + k) \leq f(\beta_1, \beta_2, k)e(\tau_p)$; and (ii) that in the last round $\tau_{p+1} = \tau_p + k_p$ of each epoch $p \geq 0$, $e(\tau_{p+1}) \leq (1 - \beta)e(\tau_p)$. The first claim shows that

the distance to equilibrium decreases from each round to the next within each epoch, while the second claim shows that the distance to equilibrium decreases geometrically from the last round of each epoch to the last round of the next epoch. Combining the two claims, we obtain that for any epoch $p \geq 0$ and any round k , $0 \leq k \leq k_p$, in epoch p , $e(\tau_p + k) \leq f(\beta_1, \beta_2, k)(1 - \beta)^p e(0)$. Therefore, for any update schedule $\tau_0 < \tau_1 < \tau_2 < \dots$, the opinion formation process (5) converges to $(\mathbb{I} - (D + E))^{-1} B \mathbf{s}$ in $O(\ln(e(0)/\varepsilon)/\beta)$ epochs.

To prove (i), we fix any epoch $p \geq 0$ and apply induction on k . The basis, where $k = 0$, holds because $f(\beta_1, \beta_2, 0) = 1$. For any round k , with $1 \leq k \leq k_p$, in p , we have that:

$$\begin{aligned} e(\tau_p + k) &= \|D\mathbf{z}(\tau_p + k - 1) + E\mathbf{z}(\tau_p) + B\mathbf{s} - (D\mathbf{z}^* + E\mathbf{z}^* + B\mathbf{s})\| \\ &\leq \|D\| \|\mathbf{z}(\tau_p + k - 1) - \mathbf{z}^*\| + \|E\| \|\mathbf{z}(\tau_p) - \mathbf{z}^*\| \\ &\leq (1 - \beta_1)e(\tau_p + k - 1) + (1 - \beta_2)e(\tau_p) \\ &\leq (1 - \beta_1)f(\beta_1, \beta_2, k - 1)e(\tau_p) + (1 - \beta_2)e(\tau_p) = f(\beta_1, \beta_2, k)e(\tau_p). \end{aligned}$$

The first inequality follows from the properties of matrix norms. The second inequality holds because $\|D\| \leq 1 - \beta_1$ and $\|E\| \leq 1 - \beta_2$. The third inequality follows from the induction hypothesis. Finally, we use that for any $k \geq 1$, $(1 - \beta_1)f(\beta_1, \beta_2, k - 1) + 1 - \beta_2 = f(\beta_1, \beta_2, k)$.

To prove (ii), we fix any epoch $p \geq 0$ and apply claim (i) to the last round $\tau_{p+1} = \tau_p + k_p$, with $k_p \geq 1$, of epoch p . Hence, $e(\tau_{p+1}) = \|\mathbf{z}(\tau_p + k_p) - \mathbf{z}^*\| \leq f(\beta_1, \beta_2, k_p)e(\tau_p)$.

We next show that $f(\beta_1, \beta_2, k_p) \leq 2 - (\beta_1 + \beta_2) = 1 - \beta$, which concludes the proof of the claim. The inequality holds because for any integer $k \geq 1$, $f(\beta_1, \beta_2, k)$ is a convex function of β_1 . For a formal proof, we fix any $k \geq 1$ and any $\beta_2 \in (0, 1)$, and consider the functions $g(x) = (1 - x)^k + \frac{1 - (1 - x)^k}{x}(1 - \beta_2)$ and $h(x) = 2 - \beta_2 - x$, where $x \in [1 - \beta_2, 1]$ (since we assume that $\beta_1 \in (0, 1)$ and that $\beta_1 > 1 - \beta_2$). For any fixed value of $\beta_2 \in (0, 1)$, $h(x)$ is a linear function of x with $h(1 - \beta_2) = 1$ and $h(1) = 1 - \beta_2$. For any fixed value of $k \geq 1$ and $\beta_2 \in (0, 1)$, $g(x)$ is a convex function of x with $g(1 - \beta_2) = 1 = h(1 - \beta_2)$ and $g(1) = 1 - \beta_2 = h(1)$. Therefore, for any $\beta_1 \in [1 - \beta_2, 1]$, $g(\beta_1) \leq h(\beta_1) = 2 - (\beta_1 + \beta_2)$.

To obtain an upper bound on $e(0) = \|\mathbf{s} - \mathbf{z}^*\|$, we work as in the proof of Lemma 1, using the fact that $\|D + E\| \leq 1 - \beta$, and show first that $\|(\mathbb{I} - (D + E))^{-1}\| \leq 1/\beta$ and then that $\|\mathbf{z}^*\| \leq \|B\|/\beta$. Since $\mathbf{z}(0) = \mathbf{s}$, we have that $e(0) = \|\mathbf{s} - \mathbf{z}^*\| \leq 1 + \|B\|/\beta$. Using the fact that for each epoch $p \geq 0$ and for every round k , $0 \leq k \leq k_p$, in p , $e(\tau_p + k) \leq f(\beta_1, \beta_2, k)(1 - \beta)^p e(0)$, we obtain that $t^*(\varepsilon) = O(\ln(\frac{\|B\|}{\varepsilon\beta})/\beta)$ epochs. \square

For average-oriented games, $D + E = A$, $\|D\| \leq 1 - 1/n$ and $\|E\| \leq (n - 1)/n^2$. Hence, applying Theorem 1 with $\beta \geq 1/n^2$, we conclude that for any $\varepsilon > 0$, the opinion formation process (5) with outdated information about $\text{avg}(\mathbf{z}(t))$ converges to $\mathbf{z}^* = (\mathbb{I} - A)^{-1} B \mathbf{s}$ within distance ε in $O(n^2 \ln(n/\varepsilon))$ epochs.

4 The PoA of Symmetric Average-Oriented Games

We proceed to bound the PoA of average-oriented opinion formation games. We now concentrate on the most interesting case of symmetric games, since nonsymmetric opinion formation games can have a PoA of $\Omega(n)$, even if $\alpha = 0$ (see e.g., [6, Fig. 2]). We recall that for symmetric games, $w_{ij} = w_{ji}$ for all agent pairs i, j , and $w_i = 1$ and $\alpha_i = \alpha$, for all agents i .

Our analysis generalizes a local smoothness argument put forward in [4, Sec. 3.1]. A function $C(\mathbf{z})$ is (λ, μ) -locally smooth [19] if there exist $\lambda > 0$ and $\mu \in (0, 1)$, such that for all $\mathbf{z}, \mathbf{x} \in \mathbb{R}^n$,

$$C(\mathbf{z}) + (\mathbf{x} - \mathbf{z})^T C'(\mathbf{z}) \leq \lambda C(\mathbf{x}) + \mu C(\mathbf{z}), \quad (7)$$

where $C'(\mathbf{z}) = (\frac{\partial C_1(\mathbf{z})}{\partial z_1}, \frac{\partial C_2(\mathbf{z})}{\partial z_2}, \dots, \frac{\partial C_n(\mathbf{z})}{\partial z_n})$ is the vector with the partial derivative of $C_i(\mathbf{z})$ with respect to z_i , for each agent i . At the equilibrium \mathbf{z}^* , $C'(\mathbf{z}^*) = \mathbf{0}$. Hence, applying (7) for the equilibrium \mathbf{z}^* and for the optimal solution \mathbf{o} , we obtain that $\text{PoA} \leq \lambda/(1 - \mu)$. For symmetric games without aggregation, [4, Sec. 3.1] shows that for any $\mathbf{s} \in [0, 1]^n$, the cost function $\sum_{i=1}^n (z_i - s_i)^2 + \sum_{i \in N} \sum_{j \neq i} w_{ij} (z_i - z_j)^2$ is (λ, μ) -locally smooth for any $\lambda \geq \max\{1/(4\mu), 1/(\mu + 1)\}$. Using $\lambda = 3/4$ and $\mu = 1/3$, we obtain that the PoA of symmetric opinion formation games without aggregation is at most $9/8$ [4], which is tight [6, Fig. 1].

This elegant approach cannot be directly generalized to symmetric average-oriented games, because the function $\sum_{i \in N} (\text{avg}(\mathbf{z}) - s_i)^2$ is not (λ, μ) -locally smooth for any $\mu < 1$. So, instead of trying to find λ, μ so that (7) holds for all $\mathbf{z} \in \mathbb{R}^n$, we identify values of λ, μ such that (7) holds for all opinion vectors \mathbf{z} with $\text{avg}(\mathbf{z}) = \text{avg}(\mathbf{s})$. This suffices for bounding the PoA, since we need to apply (7) only for the optimal opinion vector \mathbf{o} and the equilibrium opinion vector \mathbf{z}^* . Moreover, the following proposition shows that for the equilibrium vector \mathbf{z}^* , we have that $\text{avg}(\mathbf{z}^*) = \text{avg}(\mathbf{s})$.

Proposition 1. *Let \mathbf{z}^* be the equilibrium and \mathbf{s} the agent belief vector of any symmetric average-oriented opinion formation game. Then, $\text{avg}(\mathbf{z}^*) = \text{avg}(\mathbf{s})$.*

Based on Proposition 1, we show that the PoA of symmetric average-oriented games tends to $9/8$, which is the PoA of symmetric opinion formation games without aggregation.

Theorem 2. *Let \mathcal{G} be any symmetric average-oriented opinion formation game with n agents and influence $\alpha \geq 0$ from the average public opinion. Then, $\text{PoA}(\mathcal{G}) \leq \frac{9}{8} + O(\frac{\alpha}{n^2})$.*

Proof. We find appropriate parameters $\lambda > 0$ and $\mu \in (0, 1)$ such that (7) holds for any $\mathbf{x} \in \mathbb{R}^n$ and any $\mathbf{z} \in \mathbb{R}^n$ with $\text{avg}(\mathbf{z}) = \text{avg}(\mathbf{s})$. Since the equilibrium \mathbf{z}^* of any symmetric average-oriented game \mathcal{G} has $\text{avg}(\mathbf{z}^*) = \text{avg}(\mathbf{s})$, by Proposition 1, $\text{PoA}(\mathcal{G}) \leq \lambda/(1 - \mu)$.

We divide agent's i personal cost $C_i(\mathbf{z})$ into three parts $C_i(\mathbf{z}) = F_i(\mathbf{z}) + I_i(\mathbf{z}) + A_i(\mathbf{z})$, where $F_i(\mathbf{z}) = \sum_{j \neq i} w_{ij} (z_i - z_j)^2$, $I_i(\mathbf{z}) = (x_i - s_i)^2$ and $A_i(\mathbf{z}) = (\text{avg}(\mathbf{z}) - s_i)^2$. Following the previous notation:

$$\begin{aligned}
F(\mathbf{z}) &= \sum_{i \in N} F_i(\mathbf{z}) = \sum_{i \in N} \sum_{j \neq i} w_{ij} (z_i - z_j)^2 = 2 \sum_{i,j:i \neq j} w_{ij} (z_i - z_j)^2 \\
I(\mathbf{z}) &= \sum_{i \in N} I_i(\mathbf{z}) = \sum_{i \in N} (z_i - s_i)^2 = (\mathbf{z} - \mathbf{s})^T (\mathbf{z} - \mathbf{s}) \\
A(\mathbf{z}) &= \sum_{i \in N} A_i(\mathbf{z}) = \alpha \sum_{i \in N} (\text{avg}(\mathbf{z}) - s_i)^2 = \alpha (\text{avg}(\mathbf{z}) - \mathbf{s})^T (\text{avg}(\mathbf{z}) - \mathbf{s}) .
\end{aligned}$$

Hence, the social cost is $C(\mathbf{z}) = F(\mathbf{z}) + I(\mathbf{z}) + A(\mathbf{z})$. We let $F'(\mathbf{z}) = (\frac{\partial F_1(\mathbf{z})}{\partial z_1}, \dots, \frac{\partial F_n(\mathbf{z})}{\partial z_n})$, $I'(\mathbf{z}) = (\frac{\partial I_1(\mathbf{z})}{\partial z_1}, \dots, \frac{\partial I_n(\mathbf{z})}{\partial z_n})$ and $A'(\mathbf{z}) = (\frac{\partial A_1(\mathbf{z})}{\partial z_1}, \dots, \frac{\partial A_n(\mathbf{z})}{\partial z_n})$ be the vectors with the partial derivatives of $F_i(\mathbf{z})$, $I_i(\mathbf{z})$ and $A_i(\mathbf{z})$, respectively, with respect to z_i , for each agent i . Note that $A'(\mathbf{z}) = (2\alpha/n)(\text{avg}(\mathbf{z}) - \mathbf{s})$. The following two propositions are proven in [4, Sec. 3.1].

Proposition 2 ([4]). *For any symmetric matrix $W = (w_{ij})$, any $\mathbf{z}, \mathbf{x} \in \mathbb{R}^n$, and any $\lambda > 0$ and $\mu \in (0, 1)$ with $\lambda \geq 1/(4\mu)$,*

$$F(\mathbf{z}) + (\mathbf{x} - \mathbf{z})^T F'(\mathbf{z}) \leq \lambda F(\mathbf{x}) + \mu F(\mathbf{z}) .$$

Proposition 3 ([4]). *For any $\mathbf{z}, \mathbf{x}, \mathbf{s} \in \mathbb{R}^n$, $\lambda > 0$ and $\mu \in (0, 1)$ with $\lambda \geq 1/(\mu + 1)$, it holds that $I(\mathbf{z}) + (\mathbf{x} - \mathbf{z})^T I'(\mathbf{z}) \leq \lambda I(\mathbf{x}) + \mu I(\mathbf{z})$.*

Using Proposition 1 and increasing the right-hand side by a small fraction of $I(\mathbf{x})$ and $I(\mathbf{z})$, we can prove an upper bound on $A(\mathbf{z}) + (\mathbf{x} - \mathbf{z})^T A'(\mathbf{z})$.

Proposition 4. *For any $\alpha > 0$, any $\mathbf{z}, \mathbf{x}, \mathbf{s} \in \mathbb{R}^n$ with $\text{avg}(\mathbf{z}) = \text{avg}(\mathbf{s})$, any $\delta \geq 0$, and any $\lambda > 0$ and $\mu \in (0, 1)$ such that $\lambda\mu \geq \alpha/n^2$,*

$$A(\mathbf{z}) + (\mathbf{x} - \mathbf{z})^T A'(\mathbf{z}) \leq \delta A(\mathbf{x}) + \mu I(\mathbf{x}) + (1 - \delta + 2\lambda)A(\mathbf{z}) + \mu I(\mathbf{z}) . \quad (8)$$

Applying Propositions 2 and 3 with $\lambda = 3/4$ and $\mu = 1/3$, and Proposition 4, and summing up the corresponding inequalities, we obtain that for any $\delta \geq 0$, and any $\lambda > 0$ and $\mu \in (0, 1)$ with $\lambda\mu \geq \alpha/n^2$,

$$\text{PoA}(\mathcal{G}) \leq \frac{\max\{3/4, \delta\} + \mu}{1 - \max\{1/3, 1 - \delta + 2\lambda\} - \mu} \quad (9)$$

If α/n^2 is small enough, e.g., if $\alpha/n^2 \leq 1/2400$, we use $\delta = 3/4$, $\lambda = 1/24$ and $\mu = 24\alpha/n^2$ in (9) and obtain that $\text{PoA}(\mathcal{G}) \leq 9/8 + O(\frac{\alpha}{n^2})$. Otherwise, we use $\mu = 1/3$, $\lambda = 3\alpha/n^2$ and $\delta = 6\alpha/n^2 + 2/3$, and obtain that $\text{PoA}(\mathcal{G}) = O(\frac{\alpha}{n^2})$. \square

5 Average-Oriented Games with Restricted Opinions

A frequent assumption in the literature on opinion formation is that agent beliefs come from a finite interval of nonnegative real numbers. Then, by scaling we can assume beliefs $s_i \in [0, 1]$. If the influence matrix A is nonnegative, then since

$b_{ii} + \sum_{j=1}^n a_{ij} = 1$ for all $i \in [n]$, we have that the equilibrium opinions are $\mathbf{z}^* = (\mathbb{I} - A)^{-1} B \mathbf{s} \in [0, 1]^n$. In contrast, for the more general domain we treat here, an important side-effect of negative influence is that the best-response (and equilibrium) opinions may not belong to $[0, 1]$. Motivated by this observation, we consider a *restricted* variant of opinion formation games, where the (best-response and equilibrium) opinions are restricted to $[0, 1]$.

To distinguish restricted opinion formation processes from their unrestricted counterparts, we use $\mathbf{y}(t)$ to denote the opinion vectors restricted to $[0, 1]^n$. For restricted average-oriented games and restricted games with negative influence, the best-response opinion y_i of each agent i to \mathbf{y}_{-i} is computed by (2) and (6), respectively. But now, if the resulting value is $y_i < 0$, we increase it to $y_i = 0$, while if $y_i > 1$, we decrease it to $y_i = 1$. Since the individual cost $C_i(\mathbf{y})$ is a strictly convex function of y_i , the restriction of y_i to $[0, 1]$ results in a minimizer $y_i^* \in [0, 1]$ of $C_i(y, \mathbf{y}_{-i})$. So, the restricted opinion formation process is

$$\mathbf{y}(t) = [A\mathbf{y}(t-1) + B\mathbf{s}]_{[0,1]}, \quad (10)$$

where $[\cdot]_{[0,1]}$ is the restriction of opinions $\mathbf{y}(t)$ to $[0, 1]^n$. The influence matrix A (and the influence matrices D and E for processes with outdated information) and the self-confidence matrix B are defined as in unrestricted opinion formation.

We show a general result for restricted opinion formation processes that is equivalent to Theorem 1. As in Section 3, we prove our result for the more general setting of negative influence. Using Theorem 3, we can immediately bound the convergence time for restricted average-oriented processes. The proof of the following is similar to the proof of Theorem 1.

Theorem 3. *Let D and E be influence matrices, possibly with negative elements, such that $\|D\| \leq 1 - \beta_1$, $\|E\| \leq 1 - \beta_2$, for some $\beta_1, \beta_2 \in (0, 1)$ with $\beta_1 + \beta_2 > 1$. Then, for any self-confidence matrix B , any $\mathbf{s} \in [0, 1]^n$, any update schedule $0 = \tau_0 < \tau_1 < \tau_2 < \dots$, the restricted opinion formation process $\mathbf{y}(t) = [D\mathbf{y}(t-1) + E\mathbf{y}(\tau_p) + B\mathbf{s}]_{[0,1]}$ converges to the unique equilibrium point \mathbf{y}^* of $\mathbf{y}'(t) = [(D + E)\mathbf{y}'(t-1) + B\mathbf{s}]_{[0,1]}$. For any $\varepsilon > 0$, $\mathbf{y}(t)$ is within distance ε to \mathbf{y}^* after $O(\ln(\frac{1}{\varepsilon})/\beta)$ epochs, where $\beta = \beta_1 + \beta_2 - 1$.*

We also bound the PoA of restricted symmetric average-oriented games. Due to opinion restriction to $[0, 1]$, the average opinion at equilibrium may be far from $\text{avg}(\mathbf{s})$. Therefore, we cannot rely on Proposition 4 anymore. Moreover, the PoA of restricted games increases fast with α (e.g., if $\mathbf{s} = (0, \dots, 0, 1/n)$, $w_{ij} = 0$ for all $i \neq j$, and $\alpha = n^2$, $\text{PoA} = \Omega(n)$). Therefore, we restrict our attention to the case where $\alpha = w = 1$ and show that the PoA of restricted symmetric average-oriented games remains constant. An interesting intermediate result of our analysis is that if all agents only value the distance of their opinion to their belief and to the average, i.e., if $w_{ij} = 0$ for all $i \neq j$, the PoA of such games is at most $1 + 1/n^2$.

Theorem 4. *Let \mathcal{G} be any symmetric average-oriented opinion formation game with $n \geq 2$ agents, $w = \alpha = 1$, and opinions restricted to $[0, 1]$. Then, $\text{PoA}(\mathcal{G}) \leq (\sqrt{2} + 2)^2/2 + O(\frac{1}{n})$, where $(2 + \sqrt{2})^2/2 < 5.8285$.*

References

1. C. Altafini. Consensus Problems on Networks With Antagonistic Interactions. *IEEE Transactions on Automatic Control*, 58(4):935-946, 2013.
2. V. Auletta, I. Caragiannis, D. Ferraioli, C. Galdi, and G. Persiano. Generalized Discrete Preference Games. In *Proc. of the 25th International Joint Conference on Artificial Intelligence (IJCAI '16)*, pp. 53-59, 2016.
3. S. Barberà. An introduction to strategyproof social choice functions. *Social Choice and Welfare*, 18:619-653, 2001.
4. K. Bhawalkar, S. Gollapudi, and K. Munagala. Coevolutionary opinion formation games. In *Proc. of the 45th ACM Symposium on Theory of Computing (STOC '13)*, pp. 41-50, 2013.
5. V. Bilò, A. Fanelli, and L. Moscardelli. Opinion formation games with dynamic social influences. In *Proc. of the 12th Conference on Internet and Network Economics (WINE '16)*, LNCS 10123, pp. 444-458, 2016.
6. D. Bindel, J.M. Kleinberg, and S. Oren. How bad is forming your own opinion? In *Proc. of the 52nd IEEE Symposium on Foundations of Computer Science (FOCS '11)*, pp. 57-66, 2011.
7. B. Chazelle and C. Wang. Inertial Hegselmann-Krause systems. *IEEE Transactions on Automatic Control*, to appear, 2017.
8. Po-An Chen, Yi-Le Chen, and Chi-Jen Lu. Bounds on the Price of Anarchy for a more General Class of directed Graphs in Opinion Formation Games. *Operations Research Letters*, 44(6):808-811, 2016.
9. M.H. DeGroot. Reaching a consensus. *Journal of the American Statistical Association*, 69:118-121, 1974.
10. D. Ferraioli, P.W. Goldberg, and C.Ventre. Decentralized dynamics for finite opinion games. In *Proc. of the 5th Symposium on Algorithmic Game Theory (SAGT '12)*, LNCS 7615, pp. 144-155, 2012.
11. B. Fotouhi and M.G. Rabbat. The Effect of Exogenous Inputs and Defiant Agents on Opinion Dynamics with Local and Global Interactions. *IEEE Journal of Selected Topics in Signal Processing*, 7(2):347-357, 2013.
12. N.E. Friedkin and E.C. Johnsen. Social influence and opinions. *Journal of Mathematical Sociology*, 15(3-4):193-205, 1990.
13. J. Ghaderi and R. Srikant. Opinion dynamics in social networks with stubborn agents: Equilibrium and convergence rate. *Automatica*, 2014.
14. B. Golub and M.O. Jackson. Naïve Learning in Social Networks and the Wisdom of Crowds. *American Economic Journal: Microeconomics*, 2(1):112-149, 2010.
15. R. Hegselmann and U. Krause. Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal Artificial Societies and Social Simulation*, 5, 2002.
16. M.O. Jackson. *Social and Economic Networks*. Princeton University Press, 2008.
17. H. Moulin. On strategy-proofness and single-peakedness. *Public Choice*, 35:437-455, 1980.
18. J.B. Rosen. Existence and uniqueness of equilibrium points in concave n -person games. *Econometrica*, 33:520-534, 1965.
19. T. Roughgarden and F. Schoppmann. Local smoothness and the price of anarchy in splittable congestion games. *J. Economic Theory*, 156:317-342, 2015.
20. E. Yildiz, A. Ozdaglar, D. Acemoglu, A. Saberi, and A. Scaglione. Binary opinion dynamics with stubborn agents. *ACM Transactions on Economics and Computation*, 1(4):19:1-19:30, 2013.