

The Effects of Diversity in Aggregation Games

Petros Mol¹ Andrea Vattani¹ Panagiotis Voulgaris¹

¹University of California San Diego, 9500 Gilman Drive La Jolla, California 92093

pmol@cs.ucsd.edu avattani@cs.ucsd.edu pvoulgar@cs.ucsd.edu

Abstract: Aggregation of entities is a widely observed phenomenon in economics, sociology, biology and other fields. It is natural to ask how diverse and competitive entities can achieve high levels of aggregation. In order to answer this question we provide a game-theoretical model for aggregation. We consider natural classes of strategies for the individuals and show how this affects aggregation by studying the price of anarchy of the resulting game.

Our analysis highlights the advantages of populations with diverse strategies (*heterogeneous* populations) over populations where all individuals share the same strategy (*homogeneous* populations). In particular, we prove that a simple heterogeneous population composed of *leaders* (individuals that tend to invest) and *followers* (individuals that look for short-term rewards) achieves asymptotically lower price of anarchy compared to any homogeneous population, no matter how elaborate its strategy is.

This sets forth the question of how diversity affects the problem solving abilities of populations in general. We hope that our work will lead to further research in games with diverse populations and in a better understanding of aggregation games.

Keywords: aggregation, diversity, price of anarchy, price of stability.

1 Introduction

Aggregation of different entities manifests itself in several dynamic systems. Global population is one example: people are aggregated in a few dense urban areas rather than being distributed uniformly over the entire planet. Similar aggregation phenomena are observed in smaller scale as well: concentration of stores inside malls, abundance of restaurants around the center of a city, high density of students living near a university, and so on. Many more examples are encountered in the areas of economics, sociology, biology, and other fields.

Given the large body of evidence of aggregation phenomena, we wish to provide a theoretical model that explains it. The formalization we use in this paper models a population of t individuals inhabiting a world – represented as an undirected graph¹ with n nodes– and measures aggregation by the number of edges induced by the nodes occupied by the individuals. Since all the aforementioned examples are dynamic systems, evolving by means of choices taken by a large number of competitive entities, game theory provides an appropriate framework for analysis.

Note that optimizing aggregation in this form can be seen as an instance of the densest t -subgraph problem which is known to be NP-hard [13] and likely to be hard to approximate [8,13,14]. Given the complexity of the underlying problem it is natural to ask whether competitive entities are able to achieve high levels of aggregation. Specifically, we can ask: what strategies drive selfish behavior to form aggregated networks? In this work we consider a natural *class* of possible behaviors that players can follow, and we analyze the whole spectrum of games defined by this class. We identify behaviors that define games yielding high aggregation as well as subclasses of them that inherently incur low aggregation.

For specific players' behaviors, we measure the quality of aggregation by studying the Nash equilibria² of the corresponding game; i.e., placements of the population for which no individual has an incentive to move from its current position. Our main focus is the study of the *price of anarchy* in our games, which refers to the ratio of an optimum centralized solution to the *worst* Nash equilibrium [17,19]. In addition, we study the *price of stability* (also known as “optimistic price of anarchy”), which is defined as the ratio of the optimum to the *best* Nash equilibrium [3,4].

¹We assume that each node of the graph accommodates at most one individual.

²In this paper we are only concerned with pure Nash equilibria; i.e., equilibria based on deterministic strategies.

The price of stability is useful in applications where a central authority proposes a collective solution so that every player has no incentive to unilaterally deviate from it. On the other hand, the price of anarchy captures worst-case situations where no central coordination exists. A low price of anarchy implies good outcomes of the game even when players act exclusively in their own interest.

OUR RESULTS. In this work, we initiate a game-theoretic study of aggregation phenomena³ which can be considered as a competitive version of the k -induced subgraph problem. Our findings highlight the significance of *heterogeneity* in achieving high levels of aggregation. We show that *heterogeneous* populations (i.e., not all individuals follow the same strategy) composed of individuals following very simple, yet diverse, strategies, outperform *homogeneous* populations (i.e., all follow the same strategy) regardless of how sophisticated the strategy followed by the latter is. In particular, we show that a population composed of only two types of individuals, the *leaders* – who have a tendency to “invest” by moving to high degree nodes (in hope that other players will follow) – and the *followers* – who are more prudent and look for immediate rewards – achieves price of anarchy which is asymptotically lower than that achieved by *any* homogeneous population.

Our results suggest that the power of diversity manifests itself more significantly in large populations (i.e. when the number of players t is comparable with the number n of nodes), in which case the gap in the price of anarchy between heterogeneous and homogeneous populations can be as large as $\Omega(n)$. Interestingly, we show that homogeneous strategies cannot outperform the simple heterogeneous strategy even if they are provided with additional information about the parameters of the game.

For all the games we study the best-response dynamics and prove fast convergence. We also consider the price of stability of our games. In particular, for the population obtained as a mixture of followers and leaders the price of stability can be made arbitrarily close to 1 by tuning the mixing parameter (while preserving a low price of anarchy). We tighten this result by showing that no population (even heterogeneous)

³We emphasize that in this work we study *aggregation phenomena of populations* that occupy a network. Aggregation in our context has nothing to do with *information aggregation* which refers to economics mechanisms designed explicitly for the purpose of collecting and aggregating information.

can achieve optimal price of stability and low price of anarchy at the same time. Our results are summarized in Table 1.

DISCUSSION. In order to model heterogeneous populations one can take two possible views. In one, all players have the *same true payoff* (number of neighbors in the present context) but adopt *different strategies* towards optimizing their payoff. In this case, the global welfare is the sum of the players’ payoffs which corresponds to the number of induced edges. In the second view, there are two types of players with two *different payoff functions*, but the social welfare is *not* the sum of the players’ payoffs. The most natural view for this work is the first one: All players have the same ultimate goal, i.e. maximize their neighbors, and act strategically towards this goal.

RELATED WORK. The aggregation game we study in this work can be interpreted as a network formation game, where the subgraph induced by the individuals at equilibrium represents the created network. Several network creation games of different flavors have been considered in the literature and most of them are related to network design [3,4,10,12,23] and social networks [5,15]. One of the common settings [1,5,10,12] assumes that each player is associated to a particular node (during the entire course of the game) and can buy edges to any other node (i.e., the underlying graph is complete). The goal of each player is to minimize the distances to all other nodes paying as little as possible. Most of the work for this game aims to bound the price of anarchy [1,10,12]. Another line of work for network formation can be interpreted as a competitive version of the Steiner tree problem [3,4], and focus on bounding the price of stability of the game, since the price of anarchy can be $\Omega(n)$ in graphs of n nodes.

All the literature we mentioned associates costs with edges of the networks, so that good solutions try to avoid dense graphs. Instead, in the games we consider, edges are beneficial. A different line of work in social networks that takes into account this aspect is exchange theory. A vast body of empirical evidence in this field shows that high-degree nodes represent more powerful positions in networks [11,24]. In particular, the bargaining problem has received considerable attention and a study of it in general networks is provided in [16].

There has also been a lot of attention recently in circumventing high price of anarchy of certain games.

	Population	Price of Anarchy		Stability	Reference
		LB	UB		
Homogeneous	Followers	$\Theta(t)$ (even for $t = \Omega(n)$)		1	Obs. 1
	Leaders	∞		∞	Obs. 3
	Arbitrary	$\Theta(t)$ (even for $t = \Omega(n)$)			Thm. 4
Heterogeneous	Followers + Leaders	$\min\{\Theta(t), \Theta(n/t)\}$		$\Theta(1 + \epsilon)$	Thm. 5,6,8,9
	Informed homogeneous	$\Omega(n/t)$	$O(\min\{t, n/t\})$	$\Theta(1 + \epsilon)$	Thm. 14,13,15,9

Table 1: Summary of main results.

A line of work considers the noisy best-response dynamics, which reach high-quality states with high probability but only after exponentially many steps [9,18]. In [6,22], high price of anarchy is circumvented by centrally coordinating some of the players. In particular, [6] considers a model in which a globally optimal behavior (which brings to the optimum) is proposed and a fraction of the players follows this advise for a while but ultimately acts in a way that maximizes their utility. Finally, [7] considers a model where each player uses an experts learning algorithm to choose between an optimal (but untrusted) behavior and the best response dynamics. Observe that in our work we consider games with different classes of players, but (a) players in each class are not centrally controlled, and (b) none of the classes follows an optimal behavior (each class separately fails indeed to achieve low price of anarchy).

Related to our work is also a seminal study of *segregation* by economist Thomas Schelling [20,21]. The general formulation of the model proposed by Schelling assumes a population residing in the cells of a grid. Each cell has eight adjacent cells (including diagonal contact). Also, each individual of the population is either of type *A* or *B* (the type represents some characteristic such as race, ethnicity, etc.) and wants to have at least r adjacent individuals of its own type, where r is a satisfaction threshold common to all individuals. The system evolves in steps, and at each step an unsatisfied individual is selected and moved to a cell offering more neighbors of its own kind. Interestingly, experiments simulating this model display a high level of segregation of the two kind of individuals even with a mild threshold r (e.g., $r = 3$). Observe that the incentives of the individuals are in fact aggregation rules (as opposed to segregation rules), therefore Schelling’s model can be interpreted as a model of aggregation as well. Throughout the paper we will point out some relations of our games to the scenario proposed by Schelling.

ORGANIZATION. We discuss preliminaries in Section

2. Section 3 is dedicated to homogeneous populations and provides the theorem establishing their inherent high price of anarchy. In Section 4, we consider the population obtained by a mixture of followers and leaders and show that it yields both low price of anarchy and price of stability. In Section 5 we consider possible extensions. We conclude with future directions in Section 6.

2 Preliminaries

Consider an undirected graph $G = (V, E)$ and t players, where t possibly depends on $|V|$. For a placement $H \subseteq V$ (with $|H| = t$) of the players onto the graph, the *global welfare* is defined as the number of edges induced by H in G . We will say that a placement is *optimum* if it induces the maximum possible number of edges. For convenience we will often use the term optimum to also indicate the value of an optimum placement (that is, the number of induced edges). Note that it is NP-hard to find an optimum placement for general graphs because it is equivalent to solving the densest t -subgraph problem [13]. Also, this problem is likely to be hard to approximate since the best known centralized algorithm gives a $O(n^{1/4+\epsilon})$ approximation [8]. To circumvent this computational barrier, we are mainly interested in the case $t = \Theta(n)$ when the t -densest subgraph admits a constant factor approximation, but we show our results for general t as well.

Given a placement H of players onto a graph G , we let $\Gamma_H(u)$ denote the degree of the node u in H (that is, the number of adjacent individuals that a hypothetical player located in u would have under placement H). Similarly, we let $\bar{\Gamma}_H(u)$ be the degree of u in the graph obtained by G after removing the edges in H (that is, the number of *empty* adjacent positions that a hypothetical player located in u would have under H). Clearly for all placements $H \subseteq V$, $\deg_G(u) = \Gamma_H(u) + \bar{\Gamma}_H(u)$.

It is easy to see (Section 3.1) that if every player

plays the most natural strategy, that of greedily moving to a location with the highest number of neighbors, then the equilibria can be very poor compared to an optimum placement. Therefore, we will consider richer classes of games, where players might make decisions that take into account possible future benefits. In all our games, the way every player i decides where to move can be described in the following manner: player i “ranks” every (available) location in the graph through a *ranking function* $f_i(\cdot)$ that gives a score to each location (i.e. node) u of the graph with respect to the current configuration, and moves to the location with highest score. The functions $f_i(u)$ we consider are “local” to the location u , in the sense that $f_i(u)$ depends only on the current configuration of the neighborhood of u , i.e. on $\Gamma_H(u)$ and $\bar{\Gamma}_H(u)$.

Most of our proofs are obtained analyzing configurations reached by a *best-response dynamics*. Best-response dynamics studies the game in an evolving fashion. Specifically, the system evolves in steps: at each step a player is chosen and given the opportunity to move to a new better location with respect to its ranking function. The way players are chosen depends on some (possibly randomized) scheduling. We note that Schelling’s original work on segregation [20,21] also uses a best-response dynamics to model evolution.

NOTATION. We are interested in undirected and connected⁴ graphs. Given an undirected graph $G = (V, E)$ and any $S_1, S_2 \subseteq V$ we denote by E_{S_1, S_2} the set of edges with one endpoint in S_1 and the other in S_2 . When clear from the context, the same notation will be used for the cardinality of the edges from S_1 to S_2 . Abusing notation, we will use E_S instead of $E_{S, S}$. Since we only consider undirected graphs, $E_{S_1, S_2} = E_{S_2, S_1}$. Likewise we use \tilde{d}_{S_1, S_2} to denote the average degree of nodes in S_1 when considering edges only in E_{S_1, S_2} (notice that $\tilde{d}_{S_1, S_2} \neq \tilde{d}_{S_2, S_1}$). It is not hard to see that $\tilde{d}_{S, S} = \frac{2E_S}{|S|}$ while if $S \cap T = \emptyset$, $\tilde{d}_{S, T} = \frac{E_{S, T}}{|S|}$. Also if T_1, \dots, T_k form a partition of T then $E_{S, T} = \sum_{i=1}^k E_{S, T_i}$ and $\tilde{d}_{S, T} = \sum_{i=1}^k \tilde{d}_{S, T_i}$. Finally, we will use t and n to denote the *size of the population* and the *size of the graph* under consideration respectively.

⁴The connectivity requirement can be dropped for most of our results. However, there are cases for which disconnected graphs give rise to some rare pathologies (e.g., when the graph is an independent set) that need special care.

3 Homogeneous populations

In this section we analyze populations where all individuals have the same ranking function, i.e. $f_i = f_j$ for all players i, j . We call such populations *homogeneous*. We start by studying two very natural strategies and prove that both fail in achieving a low price of anarchy. We conclude the section by showing that a high price of anarchy is inherent in all homogeneous populations regardless of the ranking function they use.

3.1 A population of followers

We begin by looking at the most natural ranking function for the individuals which makes a player move to another (non-occupied) location if it offers more adjacent players than its current location. Formally, given a placement H of the players onto the graph, the ranking function of each player is defined by:

$$f(u) := \Gamma_H(u),$$

and the player is incentivized to move to another location v if $v \notin H$ and $\Gamma_{H'}(v) > \Gamma_H(u)$, where $H' = (H \setminus \{u\}) \cup \{v\}$. Note that the evolution of this game captures Schelling’s model provided that the satisfaction threshold r is large (see related work).

The price of anarchy and stability of this simple game are established in the following result which also demonstrates a fast convergence of the best-response dynamics. As such, an equilibrium is easy to find.

Observation 1. *Consider the aggregation game with a population of followers. Then, for any connected graph, the price of stability is exactly 1, the price of anarchy is $O(t)$, and there are connected graphs of size n with price of anarchy as high as $\Omega(t)$ even for $t = \Theta(n)$. Finally, best-response dynamics converges in polynomial time.*

Proof. For the polynomial convergence of the best-response dynamics, define the simple (potential) function $R(H) = E_H$. Every time a player moves from u_1 to u_2 then the new configuration is $H' = H \setminus \{u_1\} \cup \{u_2\}$ with $R(H') - R(H) = \Gamma_{H \setminus \{u_1\}}(u_2) - \Gamma_{H \setminus \{u_1\}}(u_1) \geq 1$. Also $R(H) = O(t^2)$ for all placements H of size t and therefore the aggregation game with a population of followers reaches an equilibrium after at most $O(t^2) = O(n^2)$ iterations. For the price of stability, let $G_t = (V_t, E_t)$ be a t -densest subgraph of G . Then $H = V_t$ is an equilibrium (since G_t is a densest subgraph of G) and hence $\frac{E_t}{E_H} = 1$. For the

upper bound on the price of anarchy note that any subgraph of size t has at most $O(t^2)$ edges and any equilibrium has at least $\Omega(t)$ edges (since the graph is connected). Finally, for the lower bound consider the graph shown on the left in Figure 1 with $N = 0$ and $k = t$. The configuration shown is an equilibrium with $E_H = O(t)$ whereas the optimum subgraph of size t has $\Omega(t^2)$ edges. Therefore the price of anarchy can be as high as $\Omega(t) = \Omega(n)$ for $t = \Theta(n)$. \square

An optimal price of stability is appealing, however we show that in general, these optimal equilibria do not even satisfy some simple requirements. As explained in the related work section, the model provided by Schelling embeds a notion of satisfaction for the individuals⁵. Along the same lines, for any graph, we can classify Nash equilibria with respect to the minimum satisfaction among the individuals. Specifically, we say that a Nash equilibrium is r -stable if every individual has at least r adjacent players. We define the price of r -stability as the ratio of the optimum to the best r -stable equilibrium. As discussed previously, the price of stability can be interpreted as a reasonable solution proposed from a central authority to the players. In light of this, Observation 1 suggests that there are proposals of optimum value such that no individual has incentive to move. On the contrary, the following observation implies that if we were to look for proposals that guarantee even just a small amount of satisfaction for every individual (and assuming that satisfied individuals do not move), then the overall social welfare can be much worse than the optimum welfare⁶. The proof is in Appendix A.

Observation 2. *There are connected graphs of size n with price of 2-stability of $\Omega(n)$. In general, for connected graphs, the price of r -stability is $\Theta(t/r)$ for $r \geq 2$ and is exactly 1 for $r \in \{0, 1\}$.*

Finally, we note that deciding if there exists an r -stable equilibrium is NP-hard for any $r \geq 3$ [2].

3.2 A population of leaders

The previous game failed in providing a low price of anarchy due to the fact that the individuals were

⁵An individual is satisfied if it has at least r neighbors, for some threshold r common to all individuals. A satisfied individual has no incentive to move even if there exist available positions with more neighbors.

⁶A common measure for the social welfare of an equilibrium is the *egalitarian* objective function which is defined as the maximum player's utility. The argument above instead quantifies the quality of equilibria by their minimum player's utility.

short-sighted and did not look for long-term rewards. In particular, the followers' function failed to spot strategic positions in the graph. In this section we analyze a population of "leader" individuals that tries to overcome this issue. Specifically, individuals will move to high degree nodes even if they do not offer many adjacent individuals at the time of the move. In other words, individuals are investing in empty positions with the hope of gaining many adjacent players as the system evolves. Given their relation to common measures such as betweenness, high degree nodes play an important role in the study of power in social networks [11,24].

In order to account for high-degree nodes we define the following ranking function:

$$\ell(u) := \Gamma_H(u) + \bar{\Gamma}_H(u).$$

A player moves to a node v from u if $\ell(v) > \ell(u)$. Unfortunately, this population performs even worse than a population of "followers".

Observation 3. *The best-response dynamics of the game converges in polynomial time. However, there exist connected graphs for which all Nash equilibria have zero social welfare.*

Proof. For the convergence of best-response dynamics, consider the simple (potential) function $R(H) = \sum_{u \in H} \ell(u)$. Notice that every time a player moves $R(H)$ increases by at least 1. However $R(H)$ is bounded by $t(n-1)$ therefore the population achieves an equilibrium after at most $t(n-1)$ iterations. For the second part of the proof, consider the graph on the right of Figure 1 with $k = 5$. All its nodes have degree either 2 or 4. Hence, all individuals will eventually move to the nodes of degree 4. This placement induces no edges between individuals and hence the price of stability is infinite. \square

3.3 Lower bounds for homogeneous populations

In the previous sections we analyzed two simple kinds of populations: followers and leaders. For different reasons, both populations failed in ensuring a non-trivial price of anarchy. At this point, one could be tempted to think that more sophisticated ranking functions might result in lower prices of anarchy. In this section, we show that this is not the case. On the contrary, the seemingly naive strategy of followers gives (in asymptotic terms) the lowest possible price of anarchy among all homogeneous strategies. More

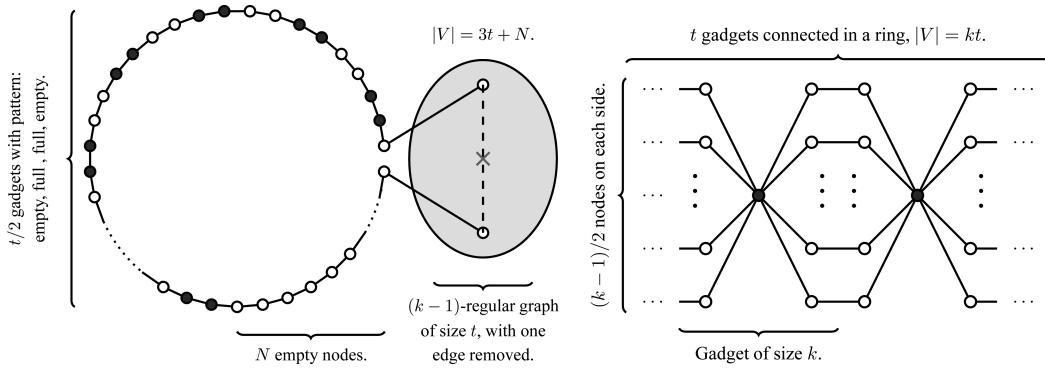


Figure 1: Graphs for lower bounds.

specifically, in Theorem 4, we show that *any* homogeneous population cannot yield low price of anarchy. Interestingly, the graphs used in Observations 1 and 3 entirely capture the hardness of achieving a low price of anarchy and play a central role in the proof of the aforementioned lower bound.

Theorem 4. *For any homogeneous population, there exist connected graphs for which the price of anarchy is $\Omega(t)$, even for $t = \Theta(n)$.*

Proof. In the case of homogeneous populations, the ranking function of the individuals can be represented as a table $s(i, j)$, where $s(i, j)$ denotes the value of the function at a node with i adjacent individuals and j adjacent empty positions. Below we show that for any relative ordering of the values⁷ $s(0, 2)$, $s(1, 1)$ and $s(0, n/3 - 1)$, there exist graphs and placements of individuals for which the price of anarchy is $\Omega(n)$. Suppose first that $s(0, 2) \geq s(1, 1)$. In that case, a simple ring with n nodes and t individuals with minimum distance at least 3 between any two individuals is an equilibrium with zero social welfare and therefore infinite price of anarchy. So, we may assume that $s(1, 1) > s(0, 2)$. If now $s(0, n/3 - 1) \geq s(1, 1) > s(0, 2)$, then, the placement shown on the right of Figure 1 (with $k = n/3$) is an equilibrium (note that every node in the graph has either degree $k - 1 = n/3 - 1$ or 2) which again yields infinite price of anarchy. The remaining case is $s(1, 1) > \max \{s(0, 2), s(0, n/3 - 1)\}$. In that case, the placement on the left of Figure 1 with $N = 0$ and $t = k = n/3$ is an equilibrium that induces $t/2$ edges, while the optimum is obtained by placing the t individuals in the $(k - 1)$ -regular graph of size t which yields $(t - 1)t/2 - 1$ induced edges. Therefore the price of anarchy in this case is $((t - 1)t - 2)/t = \Omega(t)$ which concludes the proof. \square

⁷Assume without loss of generality that n is an odd multiple of 3.

4 Heterogeneous populations

In Section 3, we showed that neither a population of followers nor a population of leaders can achieve a low price of anarchy. Even worse, Theorem 4 suggests that we cannot hope for low price of anarchy when considering homogeneous populations. However, it leaves open the door for heterogeneous populations. It is natural to ask how many different “classes” of individuals are required in order to reduce the price of anarchy or even how complex the strategies of each class should be. In this section we settle both questions with a favorable answer that suggests an extreme separation between homogeneous and heterogeneous populations: while the naive populations of only leaders or only followers have high price of anarchy ($\Omega(t)$, even for $t = \Theta(n)$) when considered separately, we show that a simple heterogeneous population composed of a mixture of the two achieves a low price of anarchy, in particular a *constant* price of anarchy for $t = \Theta(n)$.

We also study the price of stability achieved by the heterogeneous population. Namely, in Theorem 8, we prove that the price of stability can be made arbitrarily close to 1 by tuning the mixing parameter (while maintaining a low price of anarchy). We conclude the section by proving that this is essentially the best price of stability one can achieve without increasing the price of anarchy. More specifically, we provide an impossibility theorem showing that no population can achieve an optimal price of stability and a low price of anarchy simultaneously (see Theorem 9).

4.1 Achieving low price of anarchy

Consider a λ -heterogeneous population (for some $0 < \lambda < 1$), with λt leaders (players with ranking function ℓ from Section 3.2) and $(1 - \lambda)t$ followers (players with ranking function $f(u)$ from Section 3.1).

The following theorem shows that the best-response dynamics of such a game converges in polynomial time and provides upper bounds for the price of anarchy. Interestingly, the price of anarchy of λ -heterogeneous populations with *constant* λ , is upper bounded by $O(\sqrt{n})$ and can be as low as constant when $t = \Theta(n)$.

Theorem 5. *Fix any $0 < \lambda < 1$ and any connected graph G of n nodes. Then the λ -heterogeneous population achieves a constant price of anarchy for $t = \Theta(n)$.*

In general, the price of anarchy is $O\left(\min\left\{\frac{1}{1-\lambda}t, \frac{1}{\lambda(1-\lambda)}\frac{n}{t}\right\}\right)$. In addition, best-response dynamics converges in polynomial time.

Proof. The proof of polynomial time convergence is a simple combination of Observations 1 and 3. First notice that the number of leaders' moves cannot exceed $O(\lambda tn)$ (every time a leader moves the potential function $\sum_{u:\text{leader}} \deg_G(u)$ increases by at least 1 and this sum cannot exceed the value λtn). Now conditioned on leaders not moving, followers' moves are also polynomially bounded (Observation 1). The two bounds together guarantee polynomial time convergence for the whole population.

Let now H be the set of nodes occupied by the population in any equilibrium. We will use $F, L \subseteq H$ to denote the set of nodes occupied (upon convergence) by followers and leaders respectively. We have $|L| = \lambda t, |F| = (1-\lambda)t$. Also let B the subset of nodes of any t -densest subgraph of the graph. We want to bound the price of anarchy, i.e. E_B/E_H . For the upper bound of $t/(1-\lambda)$, simply observe that $E_B < t^2/2$ while $E_H \geq (1-\lambda)t/2$ since the followers will have at least one neighbor because the graph is connected.

For the other bound, we define ℓ_0 (resp. f_0) to be the minimum value of the ranking function ℓ (resp. f) over the positions of the leaders (resp. followers) in H . That is, $\ell_0 = \min_{u \in L} \ell(u)$ and $f_0 = \min_{u \in F} f(u)$. We observe that every node in $B \setminus H$ can not have more than $f_0 + 1$ neighbors in H otherwise H would not be an equilibrium. Also there are at most t nodes in $B \setminus H$ and therefore $E_{B \setminus H, H} \leq (f_0 + 1) \cdot t$. On the other hand, $E_H \geq f_0 \cdot (1-\lambda)t/2$. By combining these two inequalities with the fact that $E_B \leq E_H + E_{B \setminus H, H} + E_{B \setminus H}$, we conclude that the price of anarchy is

$$\begin{aligned} \frac{E_B}{E_H} &\leq \frac{E_H}{E_H} + \frac{E_{B \setminus H, H}}{E_H} + \frac{E_{B \setminus H}}{E_H} \\ &\leq 1 + \frac{4}{1-\lambda} + \frac{E_{B \setminus H}}{E_H}. \end{aligned}$$

It remains to bound term $E_{B \setminus H}/E_H$. We start showing a lower bound for E_H . Observe that

$$\ell_0 \cdot \lambda t \leq \sum_{u \in L} \ell(u) \leq 2E_{L, H} + E_{L, V \setminus H}. \quad (1)$$

In addition, in any equilibrium H , the average degree $E_{V \setminus H, H}/|V \setminus H|$ from $V \setminus H$ to H is at most the average degree $E_{F, H}/|F|$ of F to H . Also we have $E_{V \setminus H, H} \geq E_{V \setminus H, L}$ from which $\frac{E_{F, H}}{|F|} \geq \frac{E_{V \setminus H, L}}{|V \setminus H|}$. This implies that $\frac{E_{F, H}}{(1-\lambda)t} \geq \frac{E_{V \setminus H, L}}{n}$ which entails $E_{V \setminus H, L} \leq \frac{n}{(1-\lambda)t} E_{F, H}$. Combining the last inequality with (1) we get

$$\begin{aligned} 2E_{L, H} + \frac{n}{(1-\lambda)t} E_{F, H} &\geq \ell_0 \lambda t \quad \Rightarrow \\ 2\frac{n}{(1-\lambda)t} (E_{L, H} + E_{F, H}) &\geq \ell_0 \lambda t \end{aligned}$$

Also we know that $2E_H \geq E_{L, H} + E_{F, H}$, therefore the above implies that $E_H \geq \ell_0 \lambda (1-\lambda) t^2 / (4n)$. We now bound $E_{B \setminus H}$. Note that the nodes in $B \setminus H$ cannot have degree more than ℓ_0 , otherwise H would not be an equilibrium, therefore $E_{B \setminus H} \leq \ell_0 \cdot t$. Combining the bounds on E_H and $E_{B \setminus H}$ we obtain $\frac{E_{B \setminus H}}{E_H} \leq 2 \frac{n}{\lambda(1-\lambda)t}$ from which the theorem follows. \square

The following theorem shows that the upper bound for the price of anarchy of the aforementioned heterogeneous strategy is asymptotically tight. We sketch the proof in Appendix B and defer the complete proof to the full version of the paper.

Theorem 6. *For the λ -heterogeneous population, for any λ and t , there exists a connected graph of size n such that the price of anarchy is $\Omega\left(\min\left\{\frac{1}{1-\lambda}t, \frac{1}{\lambda(1-\lambda)}\frac{n}{t}\right\}\right)$.*

4.2 Price of stability and relation to price of anarchy

We now investigate the price of stability for the heterogeneous population. We need the following technical lemma to establish our main result.

Lemma 7. *Let $G = (V, E)$ be a graph and $G_t = (V_t, E_t)$ a densest t -size subgraph of G . Then $\forall k$ with $2 \leq k \leq t$ there exists a subgraph $G_k = (V_k, E_k)$ of size k such that $E_{V_k} \geq \frac{k(k-1)}{t(t-1)} E_{V_t}$.*

Proof. In fact we will prove that there exists a subgraph G_k of G_t which has at least a $\frac{k(k-1)}{t(t-1)}$ fraction of G_t 's edges. Let $S_k(G_t)$ be the set of all possible subgraphs of G_t that have size *exactly* k (clearly

$|S_k(G_t)| = \binom{t}{k}$) and $E_{V_k}^*$ be the number of edges of the optimum size- k subgraph of V_t . Each edge e of E_{V_t} belongs to exactly $\binom{t-2}{k-2}$ size- k subgraphs of G_t . Therefore, we can write $\sum_{H \in S_k(G_t)} E_H = \binom{t-2}{k-2} E_{V_t}$ which in turn implies $\binom{t}{k} E_{V_k}^* \geq \binom{t-2}{k-2} E_{V_t}$. The theorem follows by rearranging and simplifying the terms. \square

The following theorem shows that an almost optimal price of stability can be achieved while preserving a low price of anarchy.

Theorem 8. *For every constant $\epsilon > 0$, there exists a constant $\lambda = \lambda(\epsilon) > 0$ such that the mixed population with parameter λ achieves price of stability of at most $1 + \epsilon$.*

Proof. Set $\lambda = \lambda(\epsilon) = \frac{1}{2} \left(1 - \frac{1}{\sqrt{1+\epsilon}}\right) \in (0, 1)$. and let $\alpha = \alpha(\epsilon) = 1 - \lambda = \frac{1}{2} \left(1 + \frac{1}{\sqrt{1+\epsilon}}\right)$. Let H' be a αt -densest subgraph of G and D the subset of λt highest degree nodes in G . Consider a placement where the λt leaders are placed in D and the followers are placed in $H' \setminus D$. If $|H' \setminus D| < (1 - \lambda)t = \alpha t$ then the remaining followers are placed in arbitrary positions in the graph. Let H_0 be this initial placement. First notice that $H' \subseteq H_0$ since $\alpha t = (1 - \lambda)t = |H'| \geq |H' \setminus D|$. Allow the individuals move according to the best-response dynamics until they reach an equilibrium H . Note that, throughout the game, leaders won't move since they have initially been placed in the highest degree nodes. This means that only followers move and hence the total number of edges among individual can only increase ($E_H \geq E_{H_0} \geq E_{H'}$). Let now B be a t -densest subgraph of G . We then have

$$\begin{aligned} \frac{E_H}{E_B} &\geq \frac{E_{H'}}{E_B} \stackrel{\text{Lemma 7}}{\geq} \frac{\alpha t(\alpha t - 1)}{t(t-1)} \\ &= \frac{\alpha^2(t - 1/\alpha)}{t-1} = \alpha^2 \left(1 - \frac{1/\alpha - 1}{t-1}\right) \\ &\stackrel{t \geq 2}{\geq} \alpha^2(2 - 1/\alpha) \geq \frac{1}{1 + \epsilon} \end{aligned}$$

where in the last inequality we used the fact that by definition $\alpha \geq \frac{1}{\sqrt{1+\epsilon}}$ and $2\alpha - 1 = \frac{1}{\sqrt{1+\epsilon}}$. \square

Theorem 8 guarantees an arbitrarily good price of stability for the heterogeneous population. The following question arises: Can we achieve an optimum or quasi-optimum price of stability (that is, approaching 1) without increasing the price of anarchy? The following theorem provides a negative answer stating that any population (homogeneous or heterogeneous) that achieves a quasi-optimum price of stability *must*

have a high price of anarchy. We emphasize that the following theorem holds for *any* kind of heterogeneous population: specifically, it holds true even if every individual has a different strategy.

Theorem 9. *Consider any (even heterogeneous) population of t individuals. Suppose that for any connected graph the price of stability is at most $1 + 1/n$. Then, there exist connected graphs for which the price of anarchy is $\Omega(n)$.*

Proof. We consider the two graphs in Figure 2. We will show that quasi-optimum stability on the left graph implies $\Omega(n)$ price of anarchy on the right graph. Clearly the optimal arrangement H for the left graph is obtained by placing all the players on the ring which yields $E_H = t$. Also, *any* other arrangement gives at most $t - 1$ edges. Therefore if the price of stability is less than $1 + 1/n < t/(t-1)$, it should be the case that the players on the ring do not have incentive to move to the center of the star, or the bridge. Now consider the graph on the right where the star is replaced by a t -clique with one node removed (the t -densest subgraph has $t(t-1)/2 - 1$ edges). Notice that the nodes of the clique have the same value (degree) as the central node for the star⁸. So for any population with stability less than $1 + 1/n$, the ring is an equilibrium and therefore the price of anarchy can be as high as $(t-1)/2 - 1/t = \Omega(t) = \Omega(n)$. \square

5 Extensions

Under our model, we proved that any homogeneous strategy is bound to have high price of anarchy (i.e. $\Omega(t)$ even for $t = \Theta(n)$), while a mixture of leaders and followers achieves a low price of anarchy, in particular a constant price of anarchy for $t = \Theta(n)$. In this section we provide a less strict notion of leaders and show how it affects the price of anarchy (when mixed with followers). Moreover, we connect this new concept to a new, more powerful kind of homogeneous populations.

⁸Notice the existence of this extra node of degree 1 that is present in both graphs of Figure 2. Since we want the clique to connect with the ring, we need the bridge node but then have to remove one edge from the clique in order to maintain degree $t - 1$ for the nodes in the clique. However, removing this edge leaves the node shown on the bottom of the clique in the right graph with degree $t - 2$. The extra node added (which is adjacent to only this bottom node in the clique) is there to ensure degree $t - 1$ for the bottom node too.

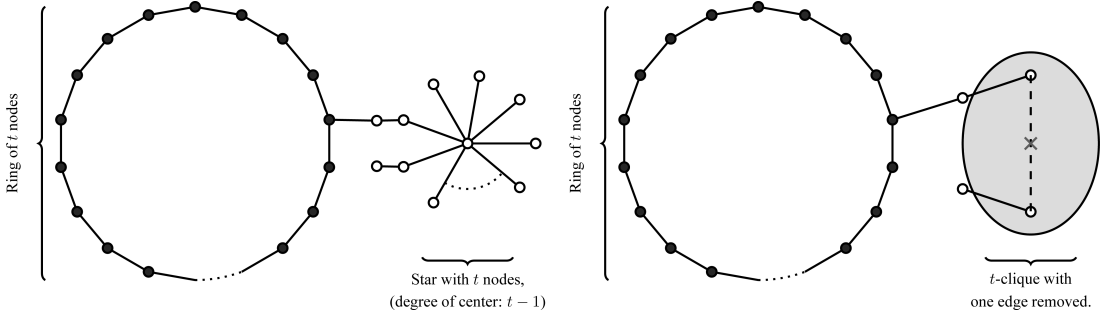


Figure 2: Graphs for Theorem 9.

5.1 Generalized β -leaders

In this section we consider a generalized definition for a population of leaders. Leaders as defined in Section 3.2 make decisions based solely on the total number of adjacent nodes regardless of how many of them are occupied. In other words leaders actions are somewhat indifferent towards aggregation⁹. We can obtain a more natural behavior considering a ranking function of the kind $\ell_\beta(u) := \Gamma_H(u) + \beta\bar{\Gamma}_H(u)$, where $0 \leq \beta \leq 1$. We call β -leaders individuals with this ranking function. The parameter β is the relative weight of an adjacent empty position to an adjacent individual. As such, it quantifies how much players are willing to invest in empty positions. Notice that for $\beta = 0$, $\ell_\beta(u)$ falls back to the ranking function of a *follower*, while for $\beta = 1$ we obtain a (pure) *leader*.

The convergence of the best-response dynamics in this game is not immediate. Define

$$\begin{aligned} R_\beta(H) &= \frac{1}{2} \sum_{u \in H} \Gamma(u) + \frac{\beta}{1+\beta} \sum_{u \in H} \bar{\Gamma}(u) \\ &= E_H + \frac{\beta}{1+\beta} E_{H, V \setminus H}, \end{aligned}$$

where H is some placement of the individuals onto the graph. By definition of R_β and ℓ_β ,

$$\sum_{u \in H} \ell_\beta(u) \geq R_\beta(H) \geq \frac{1}{2} \sum_{u \in H} \ell_\beta(u) \quad (2)$$

The following lemma establishes that R_β is a potential function, that is, all equilibria of the game are local optima of R_β and viceversa.

Lemma 10. *The function $R_\beta(H)$ strictly increases at each step of the best-response dynamics. Moreover, best-response dynamics converges in polynomial time.*

⁹To be more precise, leaders do seek aggregation even though they do so in a more indirect way (and not on a per step basis) by creating the conditions for better aggregation in the future. In fact this strategy turns out to be quite successful when leaders are mixed with followers.

Proof. Let H_1 be the subgraph induced by any placement of the individuals, $H_2 = H_1 \cup \{u_2\} \setminus \{u_1\}$ be the induced subgraph after one step of the best-response dynamics (where an individual moved from u_1 to u_2) and $H' = H_1 \cap H_2$. Then we have: $\Delta R = R_\beta(H_2) - R_\beta(H_1) = E_{H_2} - E_{H_1} + \frac{\beta}{1+\beta}(E_{H_2, V \setminus H_2} - E_{H_1, V \setminus H_1})$. We will show that this quantity is strictly positive. First we observe that $E_{H_1, V \setminus H_1} = E_{H', V \setminus H'} - E_{u_1, H'} + E_{u_1, V \setminus H'}$, which implies $E_{H_2, V \setminus H_2} - E_{H_1, V \setminus H_1} = -E_{u_2, H'} + E_{u_2, V \setminus H'} + E_{u_1, H'} - E_{u_1, V \setminus H'}$. Also we have $E_{H_2} - E_{H_1} = E_{u_2, H'} - E_{u_1, H'}$. Therefore,

$$\begin{aligned} \Delta R &= E_{u_2, H'} - E_{u_1, H'} \\ &\quad + \frac{1}{1+\beta}(-E_{u_2, H'} + E_{u_2, V \setminus H'} \\ &\quad + E_{u_1, H'} - E_{u_1, V \setminus H'}) \\ &= \frac{1}{1+\beta}[(E_{u_2, H'} + \beta E_{u_2, V \setminus H'}) \\ &\quad - (E_{u_1, H'} + \beta E_{u_1, V \setminus H'})] \\ &= \frac{1}{1+\beta}[\ell_\beta(u_2) - \ell_\beta(u_1)] > 0. \end{aligned}$$

□

Note that Theorem 4 implies that, for any constant $0 \leq \beta \leq 1$, a population composed exclusively of β -leaders has high price of anarchy. Analogously to the λ -heterogeneous population, we can consider a λ -mixture of β -leaders and followers. For this game we were not able to show convergence of the best-response dynamics. We leave as an open question if this game is a potential game and if it admits convergence for every $0 \leq \lambda, \beta \leq 1$. As for the quality of the equilibria the following theorem holds. The proof is similar in spirit to the proof of Theorem 5 and is omitted for presentation purposes.

Theorem 11. *Consider any $0 < \lambda < 1$, $0 \leq \beta \leq 1$, and any connected graph G of n nodes. Then for any equilibrium H , and t -densest subgraph B ,*

$$\frac{E_B}{E_H} =$$

$$\begin{cases} O(\frac{t}{1-\lambda}), & 0 \leq \beta < \frac{1}{t}, \\ O(\min\{\frac{t}{1-\lambda}, \frac{1}{\beta\lambda(1-\lambda)}\}), & \frac{1}{t} \leq \beta \leq \frac{1}{2(1-\lambda)} \frac{t}{n}, \\ O(\min\{\frac{t}{1-\lambda}, \frac{1}{\lambda(1-\lambda)} \frac{n}{t}\}), & \frac{1}{2(1-\lambda)} \frac{t}{n} < \beta \leq 1, \end{cases}$$

5.2 The effects of information

Theorem 4 implies that, for any fixed $0 \leq \beta \leq 1$, there exists a graph where a population of β -leaders behaves poorly. This raises a dual question: is it possible that for any graph there exists a β for which the β -leader population achieves high levels of aggregation? Note that this yields a new (somewhat less realistic¹⁰) model under which individuals are given more information about the game and they are allowed to adapt their strategies based on it. We show that if we let β depend on two additional quantities, the size n of the graph and the size t of the population, then there are (homogeneous) β -leader populations that achieve lower price of anarchy¹¹. However, we observe that the ranking functions need fine tuning and still cannot do better than the simple heterogeneous population we described. In particular, we can show (Theorem. 13, Appendix C) that an “informed” homogeneous population of $\beta(n, t)$ -leaders with $\beta(n, t) = \frac{\lambda n}{t}$, for $0 < \lambda < 1/4$, achieves price of anarchy $O(\min\{\frac{t}{1-4\lambda}, \frac{1}{\lambda(1-4\lambda)} \frac{n}{t}\})$. Moreover, this is (asymptotically) the best possible for any homogeneous population that is informed with t and n (Theorem 14, Appendix C). Finally, in Theorem 15 (Appendix C) we study the price of stability for this informed population.

6 Conclusions and future work

We have proved that in aggregation games, populations with diverse strategies achieve lower price of anarchy than homogeneous populations. Somewhat similarly, a number of recent results [6,7,22] circumvent high price of anarchy by considering mixtures of strategies, in the sense that some players might follow a “globally optimal” behavior. These results together with ours open new avenues for future research:

¹⁰In this section we consider ranking functions that can depend on information such as the graph size n and the size t of the population. The ranking functions we considered so far were not allowed to use this information. This is a valid constraint especially in dynamic situations where these quantities might change over time.

¹¹When the ranking function of an homogeneous population depends on only one of n and t , it is possible to extend the proof of Theorem 4 to show that the price of anarchy can be as high as $\Omega(n)$.

how does the number of different strategies affects the price of anarchy? Is there a connection between diverse strategies and the quality of equilibria for more general classes of games?

Our work also suggests that game theory can be a useful tool in analyzing a wider class of dynamic systems relevant to aggregation. Extending our analysis to segregation might lead to a better theoretical understanding of the seminal work of Schelling [20,21]. Finally, there are a number of possible extensions to aggregation games. For example one can consider aggregation of individuals of varying popularity or aggregation games over weighted graphs.

Acknowledgements

We would like to thank the anonymous reviewers for valuable comments regarding the content of the paper and the presentation of the results.

The first author was supported by NSF grants CNS-0831536 and CNS-0716790. The second author was supported by NSF grant CCF-0905645. The last author was supported by NSF grant CNS-0831536.

References

- [1] A. Albers, S. Eilts, E. Even-Dar, Y. Mansour, and L. Roditty. On Nash equilibria for a network creation game. In *Proceedings of ACM-SIAM Symposium on Discrete Algorithms*, 2006, pp. 89–98.
- [2] O. Amini, D. Peleg, S. Pérennes, I. Sau, and S. Saurabh. Degree-Constrained Subgraph Problems: Hardness and Approximation Results. In *Approximation and Online Algorithms*, 2008, pp. 29–42.
- [3] E. Anshelevich, A. Dasgupta, J. Kleinberg, Tardos, T. Wexler, and T. Roughgarden. The Price of Stability for Network Design with Fair Cost Allocation. In *Proceedings of IEEE Symposium on Foundations of Computer Science*, 2004, pp. 295–304.
- [4] E. Anshelevich, A. Dasgupta, E. Tardos, and T. Wexler. Near-Optimal Network Design with Selfish Agents. In *Proceedings of ACM Symposium on Theory of Computing*, 2003, pp. 511–520.
- [5] V. Bala and S. Goyal. A non-cooperative model of network formation. *Econometrica*, 68, 2000, pp. 1181–1229.

- [6] M.-F. Balcan, A. Blum, and Y. Mansour. Improved equilibria via public service advertising. In *Proceedings of ACM-SIAM Symposium on Discrete Algorithms*, 2009.
- [7] M.-F. Balcan, A. Blum, and Y. Mansour. Circumventing the price of anarchy: leading dynamics to good behavior. In *ICS*, 2010.
- [8] A. Bhaskara, M. Charikar, E. Chlamtac, U. Feige, and A. Vijayaraghavan. Detecting High Log-Densities – an $O(n^{1/4})$ Approximation for Densest k -Subgraph. To appear in *Proceedings of ACM Symposium on Theory of Computing*, 2010.
- [9] L. Blume. How noise matters. In *Games and Economic Behavior*, 44:251-271,2003.
- [10] J. Corbo and D. Parkes. The price of selfish behavior in bilateral network formation. In *Proceedings of ACM Symposium on Principles of Distributed Computing*, 2005, pp 99–107.
- [11] R. M. Emerson. Power-dependence relations. *American Sociological Review*, 27:31–40, 1962.
- [12] A. Fabrikant, A. Luthra, E. Maneva, C. Papadimitriou, and C. Shenker, On a Network Creation Game. In *Proceedings of ACM Symposium on Principles of Distributed Systems*, pp. 247–351, 2003.
- [13] U. Feige, G. Kortsarz, and D. Peleg. The Dense k -Subgraph Problem. In *Algorithmica*, 29(3):410–421, 1999.
- [14] D. Goldstein, M. Langberg. Title: The Dense k Subgraph problem. In arXiv:0912.5327, 2009.
- [15] M. Jackson. A survey of models of network formation: Stability and efficiency. In *Group Formation Economics: Networks, Clubs and Coalitions*. Cambridge University Press, Cambridge, UK, 2005.
- [16] J. Kleinberg, E. Tardos. Balanced Outcomes in Social Exchange Networks. In *Proceedings of ACM Symposium on Theory of Computing*, 2008.
- [17] E. Koutsoupias and C.H. Papadimitriou. Worst-case equilibria. In *Lecture Notes in Computer Science*, 1563:404–413, 1999.
- [18] J.R. Marden and J.S. Shamma. Revisiting log-linear learning: asynchrony, completeness, and payoff-based implementation. Submitted, 2008.
- [19] C. Papadimitriou. Algorithms, Games, and the Internet. In *Proceedings of ACM Symposium on Theory of Computing*, 749–753, 2001.
- [20] T. Schelling. Dynamic Models of Segregation. In *Journal of Mathematical Sociology*, 1:143–186, 1972.
- [21] T. Schelling. *Micromotives and Macrobehavior*. Norton, 1978.
- [22] Y. Sharma and D.P. Williamson. Stackelberg thresholds in network routing games or the value of altruism. In *ACM Conference on Electronic Commerce*, 2007.
- [23] A. Vetta. Nash Equilibria in Competitive Societies with Applications to Facility Location, Traffic Routing and Auctions. In *Proceedings of IEEE Symposium on Foundations of Computer Science*, pp. 416–425, 2002.
- [24] D. Willer (editor). *Network Exchange Theory*. Praeger, 1999.

A Proof of Observation 2

For $r = 0$ or $r = 1$ the statement is obvious since for connected graphs, all players in any optimum placement H are 1-stable (they have at least one adjacent player). For $r = 2$ consider the graph shown in Figure 3. Clearly the optimal size- t subgraph has $O(t^2)$ edges. However, the only 2-stable placement is the one where all t players are placed on the ring of size t (notice that the clique can accommodate at most $t-1$ players). Hence, the best 2-stable equilibrium has $O(t)$ edges which implies that the price of 2-stability is $\Omega(t) = \Omega(n)$. For $r > 2$ the proof is identical (we need only replace the ring with an r -regular graph).

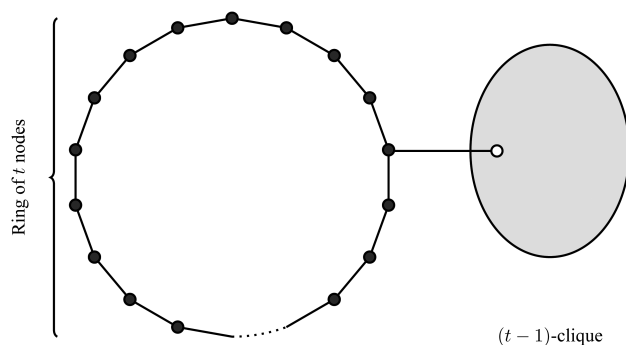


Figure 3: 2-stable placement with high price of anarchy.

B Proof sketch of Theorem 6

We present a construction of a disconnected graph for simplicity. It is not hard to extend it to a connected graph with similar price of anarchy. Consider a graph

G that contains a t -clique (let V_c be the set of the corresponding nodes) and a bipartite graph (V_1, V_2) with $|V_1| = \lambda t$ and $|V_2| = n_0 = n - \lambda t - t$. Each node of V_1 has $t-1$ edges with nodes from V_2 so that the degree of the nodes in V_2 is equally distributed and is at most $\lceil \frac{\lambda t(t-1)}{n_0} \rceil \leq 1 + \frac{\lambda t(t-1)}{n_0}$. The densest t -subgraph is the t -clique which gives $t(t-1)/2$ edges. Now notice that if all the leaders are placed on the nodes of V_1 and all the followers on the highest degree nodes of V_2 we get an equilibrium (call this placement H). Notice that each follower will have at most $1 + \frac{\lambda t(t-1)}{n_0}$ edges, so the total number of edges is $(1 - \lambda)t(1 + \frac{\lambda t(t-1)}{n_0})$. Now if $1 < \frac{\lambda t(t-1)}{n_0}$, then $E_{V_c}/E_H > \frac{n_0}{2\lambda t(1-\lambda)}$, while if $1 > \frac{\lambda t(t-1)}{n_0}$, $E_{V_c}/E_H > \frac{t-1}{2(1-\lambda)}$.

C Informed populations

Lemma 12. *Consider the aggregation game with the above $\beta(n, t)$ -leaders population with parameter λ . For any connected graph $G = (V, E)$ of n nodes and any equilibrium H , it holds that $E_H \geq (1 - 4\lambda)R_{n,t}(H)$.*

Proof. Let ℓ_0 be the minimum value of the ranking function over the individuals in H (i.e. $\ell_0 = \min_{u \in H} \ell_\beta(u)$), and let u_0 be a node in H that has $\ell_\beta(u_0) = \ell_0$. We note that for connected graphs $\ell_0 \geq 1$. Since we are at equilibrium, the individual placed in u_0 does not have any incentive to move, so the rankings of all the nodes in $V \setminus H$ after the player is removed must be at most ℓ_0 . This shows that every node in $V \setminus H$ has at most ℓ_0 adjacent individuals in $H \setminus \{u_0\}$, and at most $\ell_0 + 1$ neighbors in H . We conclude that $E_{H, V \setminus H} \leq (\ell_0 + 1)(n - t) \leq 2\ell_0(n - t)$. Also, inequality (2) implies that $R_\beta(H) \geq \ell_0 t/2$. Combining the above inequalities we get $E_{H, V \setminus H} \leq 4(n - t)R_\beta(H)/t$. Now write

$$\begin{aligned} E_H &= R_\beta(H) - \lambda \frac{t}{n + \lambda t} E_{H, V \setminus H} \\ &\geq R_\beta(H) - \lambda \frac{4(n - t)}{n + \lambda t} R_\beta(H) \\ &\geq (1 - 4\lambda)R_\beta(H), \end{aligned}$$

which concludes the proof. \square

The following theorem provides the desired bound on the price of anarchy.

Theorem 13. *Consider the aggregation game with the informed population above with parameter λ . For any connected graph $G = (V, E)$ of n nodes, the price*

of anarchy is at most

$$\min \left\{ \frac{1}{1 - 4\lambda} t, 1 + \frac{2}{\lambda(1 - 4\lambda)} \frac{n}{t} \right\}.$$

Proof. Let H be the set of nodes occupied by the population in any equilibrium and ℓ_0 be the minimum value of ranking function ℓ_β achieved over all individuals in H . Recall that for a connected graph $\ell_0 \geq 1$, and inequality (2) implies that $R_\beta(H) \geq \ell_0 t/2$. First we show that E_H cannot be too small. By Lemma 12 we get

$$E_H \geq (1 - 4\lambda)R_\beta(H) \geq (1 - 4\lambda)\ell_0 t/2. \quad (3)$$

The first part of the bound on the price of anarchy now is immediate since every optimal solution has at most $t^2/2$ edges. For the second part of the bound, let B be a densest subgraph of size t of G . All the nodes in $B \setminus H$ have degree at most $\ell_0 n/(\lambda t)$, or else H is not an equilibrium. As a result $E_{B \setminus H, V} \leq |B \setminus H| \ell_0 n/(\lambda t) \leq \ell_0 n/\lambda$. Therefore the price of anarchy is given by:

$$\begin{aligned} \frac{E_B}{E_H} &\leq \frac{E_H + E_{B \setminus H, V}}{E_H} \leq 1 + \frac{\ell_0 n}{\lambda E_H} \\ &\leq 1 + \frac{2n}{\lambda(1 - 4\lambda)t}. \end{aligned}$$

\square

Theorem 13 shows that the informed population under consideration achieves a price of anarchy of $O(n/t)$ for any $t = \Omega(\sqrt{n})$. The following theorem tightens this result by showing that this is the best possible for homogeneous informed populations.

Theorem 14. *For any homogeneous informed population, there exist connected graphs of n nodes for which the price of anarchy is $\Omega(n/t)$ for any t such that $\sqrt{n} \leq t \leq n/4$ and $n \bmod t = 0$.*

Proof. Consider a homogeneous informed population of t individuals with some informed ranking function. Without loss of generality we can assume that for each n and t this function is represented as a table $s_{n,t}(i, j)$, where the value $s_{n,t}(i, j)$ is the value of the ranking function in a location with i adjacent individuals and j adjacent empty positions. We will show that for any assignment of the values $s_{n,t}(0, 2)$, $s_{n,t}(1, 1)$ and $s_{n,t}(0, (n/t) - 1)$, there are graphs of n nodes for which the price of anarchy is $\Omega(n/t)$.

We proceed by cases. First suppose that $s_{n,t}(0, 2) \geq s_{n,t}(1, 1)$. Then, placing the individuals on a ring of n nodes such that they are at distance at least 2 each

other yields an equilibrium of zero social welfare (i.e., the price of anarchy is infinite).

Therefore we can assume that $s_{n,t}(1, 1) > s_{n,t}(0, 2)$. Suppose that $s_{n,t}(0, (n/t) - 1) \geq s_{n,t}(1, 1) > s_{n,t}(0, 2)$. Then, consider the graph on the right of Fig. 1 with $k = n/t$. Note that every node in the graph has either degree $(n/t) - 1$ or 2. Thus, placing the individuals onto the nodes of degree $(n/t) - 1$ is an equilibrium with infinite price of anarchy.

The remaining case is $s_{n,t}(1, 1) > \max(s_{n,t}(0, 2), s_{n,t}(0, (n/t) - 1))$. Consider the graph on the left of Fig. 1 with¹² $N = n - 3t$ and $k = n/t$. Now if we place the t individuals on the ring in groups of two at distance two each other (as shown in the figure) we obtain an equilibrium. The value of this placement is $t/2$, while the optimum is obtained by placing the individuals in the $(n/t) - 1$ regular graph which yields a value of $(n - t)/2 - 1$. Therefore the price of anarchy in this case is $(n - t - 2)/t = \Omega(n/t)$. \square

We close this section by analyzing the price of stability for the informed population of parameter λ . For any $0 < \lambda < 1/4$, the price of stability is a constant. Moreover, by tuning the parameter λ , we can make the price arbitrarily close to one.

Theorem 15. *For every constant $\epsilon > 0$, there exists a constant $\lambda = \lambda(\epsilon) > 0$ such that, for any connected graph $G = (V, E)$, the homogeneous informed population with parameter λ achieves price of stability of at most $1 + \epsilon$.*

Proof. Set $\lambda = \frac{\epsilon}{4(1+\epsilon)} < \frac{1}{4}$ and consider any connected graph G of n nodes. Let B be a densest t -size subgraph of G . Place the individuals on the nodes in B and let them move according to the best-response dynamics until they achieve an equilibrium. Let $H \subseteq V$ be the set of nodes occupied by the population in this equilibrium. Lemma 10 shows that the ranking function R_β strictly increases after each step of the best response dynamics, therefore $R_\beta(H) \geq R_\beta(B)$. Now we have:

$$\begin{aligned} E_H &\geq (1 - 4\lambda)R_\beta(H) \geq (1 - 4\lambda)R_\beta(B) \\ &\geq (1 - 4\lambda)E_B. \end{aligned}$$

As such, the price of stability is $E_B/E_H \leq 1/(1 - 4\lambda) = 1 + \epsilon$. \square

The proof of Theorem 15 shows that $\lambda = \Theta(\epsilon)$. This

¹²For $t \geq \sqrt{n}$ it is always possible to construct a $(k - 1)$ -regular graph with size t and degree $n/t - 1$.

observation and Theorem 13 imply that, for any $\epsilon = \epsilon_n \omega(1/n)$, the informed population achieves price of stability at most $1 + \epsilon_n$ and price of anarchy $o(n)$. This result is tightly complemented by Theorem 9 in Section 4.2.