Information Retention in Heterogeneous Majority Dynamics

Vincenzo Auletta[1], Ioannis Caragiannis[2], Diodato Ferraioli[1(✉)], Clemente Galdi[3], and Giuseppe Persiano[1]

[1] Università degli Studi di Salerno, Fisciano, Italy
{auletta,dferraioli}@unisa.it, giuper@gmail.com
[2] CTI "Diophantus", University of Patras, Patras, Greece
caragian@ceid.upatras.gr
[3] Università di Napoli "Federico II", Naples, Italy
clemente.galdi@unina.it

Abstract. A dynamics *retains* a specific information about the *starting* state of a networked multi-player system if this information can be computed from the state of the system also after several rounds of the dynamics. Information retention has been studied for the function that returns the majority of the states in systems in which players have states in $\{0, 1\}$ and the system evolves according to the majority dynamics: each player repeatedly updates its state to match the local majority among neighbors only. Positive and negative results have been given for probabilistic settings in which the initial states of the players are chosen at random and in worst-case settings in which the initial state is chosen non-deterministically.

In this paper, we study the (lack of) retention of information on the majority state (that is, which states appear in more players) for a generalization of the majority dynamics that we call *heterogeneous* majority dynamics. Here, each player x changes its state from the initial state $\mathbf{b}(x) \in \{0, 1\}$ to the opposite state $1 - \mathbf{b}(x)$ only if there is a surplus greater than a_x of neighbors that express that opinion. The non-negative player-dependent parameter a_x is called the *stubbornness* of x. We call *stubborn* the players which never change opinion when they are part of the majority. We give a complete characterization of the graphs that do not retain information about the starting majority; i.e., they admit a starting state for which the heterogeneous majority dynamics takes the system from a majority of 0's to a majority of 1's. We call this phenomenon "*minority becomes Majority*" (or mbM) and our main result shows that it occurs in all graphs provided that at least one player is non-stubborn. In other words, either no player in the majority will ever change its state (because they are all stubborn) or there is a starting configuration in which information regarding the majority is not retained and minority becomes Majority.

Our results are closely related to *discrete preference games*, a game-theoretic model of opinion formation in social networks: an interplay of internal belief (corresponding to the initial state of the player) and of

N. R. Devanur and P. Lu (Eds.): WINE 2017, LNCS 10674, pp. 30–43, 2017.
https://doi.org/10.1007/978-3-319-71924-5_3

social pressure (described by the heterogeneous majority dynamics). Our results show that, because of *local* strategic decisions, the *global* majority can be *subverted*.

1 Introduction

In this paper we study the *information retention* problem with respect to the *asynchronous heterogeneous majority dynamics*. In the *homogeneous majority dynamics* (or, simply, *majority dynamics*) players sit at the vertices of a *social graph*, each player starts with an opinion in $\{0, 1\}$ and repeatedly updates it to match the opinion of the majority of its neighbors. After some number of rounds, an election by majority takes place and we ask whether the *information* regarding the starting majority is *retained* in the election outcome, hence the name of *information retention*. To avoid ties it is assumed that the number of players is odd. The retention of information in majority dynamics has been studied in a probabilistic setting in which the initial opinions are independently conditioned on the majority and biased towards it. Positive and negative results have been given by [18] and, more recently, in [21] both for the *synchronous* model, in which all opinions are updated simultaneously, and the *asynchronous* model, in which at each round a single player updates her opinion. The retention of information in the majority dynamics in a worst-case setting has been first studied by Berger [9] that constructed a series of graphs in which the majority dynamics always results in the adoption of the opinions of the players in a small minority group. Actually, the phenomenon of a *minority becoming Majority* (or, the mbM phenomenon) is not restricted to some families of graphs but instead is a feature of the majority dynamics. Roughly speaking, every graph, except essentially for the complete graph and for the empty graph, admits an initial distribution of opinions which leads the minority opinion to become majority [3] (see also [6] for experimental results about this phenomenon).

Our contribution. In this paper we study the retention of information of majority in the worst-case (or the mbM phenomenon) for the asynchronous *heterogeneous* majority dynamics played on a graph G with vertices $\{1, \ldots, n\}$, each corresponding to a player with a binary state. In the *heterogeneous* majority dynamics, each player x is described by its *stubbornness* a_x that measures the willingness of the player to adopt (and keep) an opinion that differs from its initial opinion. More precisely, we distinguish between the *initial opinion* of a player x, called the *belief* $\mathbf{b}(x)$, and its *current opinion* $\mathbf{s}(x)$. The belief is internal to the player and is never explicitly revealed whereas the opinion is publicly known. A player x with belief $\mathbf{b}(x)$ can make a move from its current opinion $\mathbf{s}(x) = \mathbf{b}(x)$ to its revised opinion $\mathbf{s}(x) = 1 - \mathbf{b}(x)$ only if $d_{1-\mathbf{b}(x)}(x) - d_{\mathbf{b}(x)}(x) > a_x$ where, for $c = 0, 1$, d_c denotes the number of neighbors with opinion c. Similarly, x makes a move from $\mathbf{s}(x) = 1 - \mathbf{b}(x)$ to $\mathbf{s}(x) = \mathbf{b}(x)$ if $d_{1-\mathbf{b}(x)}(x) - d_{\mathbf{b}(x)}(x) \leq a_x$. A set of opinions is in *equilibrium* if no player can make a move.

We say that a pair $(G, (a_1, \ldots, a_n))$ consisting of a graph G (we assume that n is odd so that majority is well-defined) and stubbornness values for the players, is *subvertable* if there exist beliefs $(\mathbf{b}(1), \ldots, \mathbf{b}(n))$ with a majority of 0's and a sequence of moves that goes from the initial *truthful* state, in which $\mathbf{s}(x) = \mathbf{b}(x)$ for each vertex x, to an equilibrium state with a majority of 1's. We call such a belief assignment *subvertable*.

Our main contribution is a characterization of the subvertable pairs and shows that a pair $(G, (a_1, \ldots, a_n))$ is subvertable unless all players are *stubborn*. Roughly speaking, a stubborn player never changes its initial opinion if it happens to be in the starting majority. In order to formalize this definition, let us consider vertex x with $\mathbf{b}(x) = 0$ and $d_0(x)$ neighbors with opinion 0 and $d_1(x)$ neighbors with opinion 1, and suppose that the majority (that is at least $(n+1)/2$ vertices) has belief 0 (and, thus, $d_1(x) \leq (n-1)/2$). Clearly, if its degree $d(x)$ satisfies $d(x) \leq a_x$, then player x cannot make a move from $\mathbf{s}(x) = 0$ to $\mathbf{s}(x) = 1$. If $d(x) \geq n - a_x - 1$ then $d_1(x) - d_0(x) = 2d_1(x) - d(x) \leq n - 1 - d(x) \leq a_x$ and thus vertex x cannot make a move from $\mathbf{s}(x) = 0$ to $\mathbf{s}(x) = 1$. Instead, it is not hard to see that if $d(x) \in [a_x + 1, n - a_x - 2]$, there are values of $d_0(x)$ and $d_1(x)$ such that vertex x can move from $\mathbf{s}(x) = 0$ to $\mathbf{s}(x) = 1$ The same reasoning applies for vertices x with $\mathbf{b}(x) = 1$ in case majority is 1. We have thus the following definition.

Definition 1 (Stubborn vertex). *Vertex x with degree $d(x)$ and stubbornness a_x is* stubborn *if* $a_x \geq \min\{d(x), n - d(x) - 1\}$.

Clearly, if all vertices are stubborn then majority cannot be subverted as no vertex x in the majority will ever make a move from $\mathbf{s}(x) = \mathbf{b}(x)$ to $\mathbf{s}(x) = 1 - \mathbf{b}(x)$. The main result of this paper shows that:

if there is at least one non-stubborn vertex then there is a subvertable belief assignment.

We find that this sharp phase transition is highly surprising, since it implies that a minority could become majority (for some stubbornness levels) even in very dense graphs (e.g., clique minus a single edge) and in very sparse graphs (e.g., a graph consisting of a single edge plus isolated nodes). Hence, it highlights an interesting lack of robustness of social networks with respect to information retention. This weakness may be relevant to explain some phenomena arising on social media, such as the wide diffusion of misinformation.

To prove this result we design a polynomial-time algorithm that takes as input the social network G and players' stubbornness $a_1, \ldots, a_n$, such that there is at least one non-stubborn vertex, and returns a subvertable belief assignment for this instance. Actually, the algorithm considers the simplest belief assignment from which minority becomes Majority, namely one in which the minority consists of $\frac{n-1}{2}$ vertices. However, we remark that our characterization does not rule out that the subvertable belief assignment can have smaller minorities (even if there are instances on which only very large minorities can become majority, e.g. when there is a single non-stubborn vertex), and our algorithm can be adapted and optimized in order to find these minorities (see, e.g., [6]).

A possible interpretation of our result comes from a game theoretic model of how opinions are formed in societies (see the discussion on discrete preference games below). Within this context, the heterogeneous majority dynamics describes the social pressure on opinions expressed by the players in a social network. Our result shows that social networks are extremely vulnerable to social pressure since there always exists a subvertable majority unless all vertices are stubborn and never change their mind (in which case we do not have much of a social network). This is particularly negative as an external adversary might be able to orchestrate a sequence of steps of the underlying dynamics so as to reach the state in which majority is subverted. In principle, though, this could be very difficult since there could be different sequences of updates that lead to different equilibria with different majorities and the adversary has to be very careful in scheduling the best response moves.

Our characterization instead proves that, as long as we consider subvertable belief assignments with a minority of $\frac{n-1}{2}$ players, a stronger result is possible: there is always one single *swing* player whose best response in the initial state is to change its opinion and this leads to a state in which *any* sequence of moves leads to an equilibrium in which majority has been subverted. In other words, the adversary that wants to subvert the majority only has to influence the swing player and then the system will evolve without any further intervention towards an equilibrium in which majority is subverted. More precisely:

Definition 2. *A vertex u is said to be a* swing *vertex for subvertable belief assignment* $\mathbf{b}$ *with $\frac{n+1}{2}$ vertices with belief* 0 *if*

1. $\mathbf{b}(u) = 0$*;*
2. $d_1(u) - d_0(u) > a_u$*, that is, in the initial state, u can move from* $\mathbf{s}(u) = 0$ *to* $\mathbf{s}(u) = 1$*;*
3. *For every x with* $\mathbf{b}(x) = 1$*, it holds that $d'_x(0) - d'_x(1) \leq a_x$, where $d'_x(c)$ is the number of neighbors y of x with* $\mathbf{s}(y) = c$ *after u's move from* $\mathbf{s}(u) = 0$ *to* $\mathbf{s}(u) = 1$*. That is, after u's move no vertex with belief* 1 *can make a move from* 1 *to* 0*.*

Note that the definition above does not imply that the majority at equilibrium consists of only $\frac{n+1}{2}$ vertices with belief 1 (the initial $\frac{n-1}{2}$ plus the swing vertex). It may be indeed the case that other vertices with belief 0 will make moves from 0 to 1 after the move of the swing vertex u. Still, the third condition above implies that, after u's move, the number of vertices with opinion 1 is a majority and the size of this majority does not decrease.

Our main result then can be improved as follows:

if there exists at least one non-stubborn vertex, then there exists a subvertable belief assignment with a swing vertex.

It is natural to ask whether the characterization can be strengthened to take into account strong initial majorities (i.e., initial majorities of size at least $(1+\delta)\frac{n+1}{2}$ for some $0 < \delta < 1$). That is, to characterize the pairs (consisting of a social network and stubbornness levels) that admit at least a subvertable strong initial majority. We prove that no such characterization can be given by showing that there exists $\delta_{\max} \approx 0.85$ such that for all $0 < \delta < \delta_{\max}$ it is NP-hard to decide whether a given G and given stubbornness $a_1, \ldots, a_n$ admit a subvertable majority of size at least $(1+\delta)\frac{n+1}{2}$. That is, unless P = NP,

no polynomial-time algorithm exists that characterizes subvertable belief assignments for large initial majorities.

Related work. The majority dynamics and its generalizations are related to a line of research in social sciences that tries to model how opinions are formed and expressed in a social context. A simple classical model has been proposed by Friedkin and Johnsen [16] (see also [14]). Its main assumption is that each individual has a private initial belief and that the opinion she eventually expresses is the result of a repeated averaging between her initial belief and the opinions expressed by other individuals with whom she has social relations. The recent work of Bindel et al. [12] assumes that initial beliefs and opinions belong to $[0, 1]$ and considers the dynamics that repeatedly averages the opinions of the neighbors.

Ferraioli et al. [15] and Chirichetti et al. [13] considered a variant of this model, named *discrete preference games*, in which beliefs and opinions are discrete. These games are directly connected to the work in this paper. For this reason, below we give a more formal description of the games, highlight the conceptual link with the heterogeneous majority dynamics and briefly discuss the significance of our results in this context.

Previous results about these games focused on the rate of convergence of the game under different dynamics [15], and on the price of stability and price of anarchy [13]. Moreover, extensions of the model have been proposed along two main directions: some works assume that connections between nodes evolve over time so that players with similar opinions are more likely to be connected [10,11]; other works consider dynamics that try to capture more complex social relations (e.g., to allies and competitors or among more than two players) [1,4].

The problem of majority retention has been recently investigated even with respect to different dynamics: e.g., in [17], various negative results are proved with respect to a 3-state population protocol introduced in [2]. Similar problems have also been considered in the distributed computing literature, motivated by the need to control and restrict the influence of failures in distributed systems; e.g., see the survey by Peleg [19] and the references therein.

Discrete preference games. A *discrete preference game* consists of a n-vertex undirected graph G (the social network), *coefficients* $\alpha_1, \ldots, \alpha_n \in (0, 1)$ and *beliefs* $\mathbf{b}(1), \ldots, \mathbf{b}(n) \in \{0, 1\}$. Player i's strategy set consists of two possible *opinions* $\mathbf{s}(i) \in \{0, 1\}$ and the cost $c_i(\mathbf{s})$ of player i in state $\mathbf{s} = (\mathbf{s}(1), \ldots, \mathbf{s}(n)) \in \{0, 1\}^n$ is defined as $c_i(\mathbf{s}) = \alpha_i \cdot |\mathbf{s}(i) - \mathbf{b}(i)| + (1 - \alpha_i) \cdot \sum_{j \in N(i)} |\mathbf{s}(i) - \mathbf{s}(j)|$, where $N(i)$ is the set of neighbors of vertex i in G (i.e., friends in the social network). Note that the cost is the convex combination through α_i of two components that depend on whether the opinion coincides with the belief and on the strategies of the neighbors, respectively, and this models players that try to balance social acceptance (which would make the player pick the opinion that is the majority among its neighbors) and faithfulness to her own principles (which would make the player pick opinion equal to belief). Different values of α_i correspond to the different individual behaviors and reflect the heterogeneity of the society.

An *equilibrium* state is defined to be a state $\mathbf{s} = (\mathbf{s}(1), \ldots, \mathbf{s}(n))$ for which there is no player i whose best response is to adopt strategy $1 - \mathbf{s}(i)$. More precisely, $\mathbf{s}$ is an *equilibrium* if for all i $c_i(\mathbf{s}) \leq c_i(1 - \mathbf{s}(i), \mathbf{s}_{-i})$, where we have used the standard game theoretic notation by which $(t, \mathbf{s}_{-i})$ denotes the vector $(\mathbf{s}(1), \ldots, \mathbf{s}(i-1), t, \mathbf{s}(i+1), \ldots, \mathbf{s}(n))$.

It is not difficult to see that the best-response dynamics of a discrete preference game with stubbornness coefficients $\alpha_1, \ldots, \alpha_n \in (0, 1)$ coincides with the heterogeneous majority dynamics with stubbornness $a_1, \ldots, a_n$ where $a_x = \left\lfloor \frac{\alpha_x}{1-\alpha_x} \right\rfloor$. The fact that the heterogeneous majority dynamics does not retain information about the majority state in the belief of the players translates to the possibility that the social network will express through opinions a majority that differs from the majority of the beliefs. It is thus possible that the *local* behavior of the players affects the *global* behavior of the network and that the social pressure felt by individual members of a social network has effects on the entire network.

2 Definitions and Overview

In this section we introduce the concepts of a bisection and of a good bisection and give an overview of the proof of our main result. Due to page limit most of the proofs are omitted. We refer the reader to the full version [5].

Good bisections yield subvertable belief assignments. A *bisection* $\mathcal{S} = (S, \overline{S})$ of a graph G with an odd number n of vertices is a partition of the vertices of G into two sets S and $\overline{S}$ of cardinality $\frac{n+1}{2}$ and $\frac{n-1}{2}$, respectively. We define the *advantage* $\mathsf{adv}_{\mathcal{S}}(x)$ of a vertex x with respect to bisection $\mathcal{S} = (S, \overline{S})$ as follows:

$$\mathsf{adv}_{\mathcal{S}}(x) = \begin{cases} W(x, S) - W(x, \overline{S}), & \text{if } x \in S; \\ W(x, \overline{S}) - W(x, S), & \text{if } x \in \overline{S}, \end{cases}$$

where $W(x, A)$ denotes the number of neighbors of x in the set A.

We say that a bisection $\mathcal{S} = (S, \overline{S})$ is *good* if

1. for every $x \in S$, $\mathsf{adv}_{\mathcal{S}}(x) \geq -a_x$;
2. there is $u \in S$ with $\mathsf{adv}_{\mathcal{S}}(u) \geq a_u + 1$.

Vertices $u \in S$ with $\mathsf{adv}_{\mathcal{S}}(u) \geq a_u + 1$ are called the *good vertices* of $\mathcal{S}$ and vertices $y \in S$ with $\mathsf{adv}_{\mathcal{S}}(y) < -a_y$ are called the *obstructions* of $\mathcal{S}$. The next lemma proves that if G has a good bisection then one can easily construct a subvertable belief assignment for G.

Lemma 1. *Let $\mathcal{S} = (S, \overline{S})$ be a good bisection for graph G and let u be a good vertex of $\mathcal{S}$. Then G admits a subvertable belief assignment* $\mathbf{b}$ *such that u is a swing vertex for* $\mathbf{b}$.

Minimal bisections. The technical core of our proof is the construction of a good bisection starting from a bisection $\mathcal{S}$ of minimal potential Φ. We define the *potential* Φ of a bisection $(S, \overline{S})$ as $\Phi(S, \overline{S}) = W(S, \overline{S}) + \frac{1}{2}\left(\sum_{x \in S} a_x - \sum_{y \in \overline{S}} a_y\right)$. We say that a bisection $\mathcal{S}$ has *k-minimal* potential if $\mathcal{S}$ minimizes the potential among all the bisections that can be obtained from $\mathcal{S}$ by swapping at most k vertices between S and $\overline{S}$. That is, $\mathcal{S}$ has *k-minimal* potential if, for all $A \subseteq S$ and for all $B \subseteq \overline{S}$, with $1 \leq |A| = |B| \leq k$, $\Phi(S, \overline{S}) \leq \Phi(S \backslash A \cup B, \overline{S} \backslash B \cup A)$. We will simply write that $\mathcal{S}$ has minimal potential whenever $\mathcal{S}$ has 1-minimal potential.

The next lemma proves some useful properties of minimal bisections.

Lemma 2. *Let $\mathcal{S} = (S, \overline{S})$ be a bisection of minimal potential. Then for all $x \in S$ and $y \in \overline{S}$, $\mathsf{adv}_{\mathcal{S}}(x) + \mathsf{adv}_{\mathcal{S}}(y) + 2W(x, y) \geq a_x - a_y$.*

Swapping vertices. To turn a minimal bisection $\mathcal{S}$ into a good bisection $\mathcal{T} = (T, \overline{T})$, we need at least one vertex in T with high advantage. One way to increase the advantage of a vertex $u \in S$ is to move vertices that are not adjacent to u away from S and to bring the same number of vertices that are adjacent to u into S. We define the *rank* of a vertex u with respect to bisection $\mathcal{S}$ as $\mathsf{rank}_{\mathcal{S}}(u) = \left\lceil \frac{a_u + 1 - \mathsf{adv}_{\mathcal{S}}(u)}{2} \right\rceil$. It is not hard to see that the rank is exactly the number of vertices that need to be moved. Note that a vertex u of $\mathsf{rank}_{\mathcal{S}}(u)$ has advantage $\mathsf{adv}_{\mathcal{S}}(u)$ such that $a_u - 2\mathsf{rank}_{\mathcal{S}}(u) + 1 \leq \mathsf{adv}_{\mathcal{S}}(u) \leq a_u - 2\mathsf{rank}_{\mathcal{S}}(u) + 2$. We next formalize the notion of swapping of vertices and prove that it is always possible to increase the advantage of a non-stubborn vertex x to $a_x + 1$.

Given a bisection $\mathcal{S} = (S, \overline{S})$ and a vertex u, a *u-pair* for $\mathcal{S}$ is a pair of sets (A_u, B_u) such that:

- if $u \in S$, then $A_u \subseteq S \cap \overline{N}(u)$ and $B_u \subseteq \overline{S} \cap N(u)$ with $|A_u| = |B_u| = \mathsf{rank}_{\mathcal{S}}(u)$;
- if $u \in \overline{S}$, then $A_u \subseteq S \cap N(u)$ and $B_u \subseteq \overline{S} \cap \overline{N}(u)$ with $|A_u| = \mathsf{rank}_{\mathcal{S}}(u)$ and $|B_u| = \mathsf{rank}_{\mathcal{S}}(u) - 1$.

The bisection $\mathcal{T}$ *associated* with the u-pair (A_u, B_u) for $\mathcal{S}$ is defined as

- if $u \in S$, $\mathcal{T} = (S \backslash A_u \cup B_u, \overline{S} \backslash B_u \cup A_u)$;
- if $u \in \overline{S}$, $\mathcal{T} = (\overline{S} \backslash B_u \cup A_u, S \backslash A_u \cup B_u)$.

Note that our choice for the size of A_u and B_u implies that in both cases $|T| = \frac{n+1}{2}$ as desired. The next lemma shows that u is a good vertex in the bisection associated with a u-pair.

Lemma 3. *For each bisection $\mathcal{S}$, let u be a vertex of the graph, (A_u, B_u) be a u-pair for $\mathcal{S}$, and $\mathcal{T}$ be the bisection associated to (A_u, B_u). Then $\mathsf{adv}_{\mathcal{T}}(u) \geq a_u + 1$.*

The problem now is to understand which vertex we have to choose for making it good. The next lemma says that stubborn vertices cannot be good vertices. But they are sort of neutral: indeed, they cannot be obstructions either.

Lemma 4. *For every bisection $\mathcal{S} = (S, \overline{S})$ and every stubborn vertex $x \in S$ it holds that $-a_x \leq \mathsf{adv}_{\mathcal{S}}(x) \leq a_x$.*

However, for every bisection $\mathcal{S}$ and every vertex u, a u-pair for $\mathcal{S}$ exists if and only if vertex u is non-stubborn, as showed by the next lemma.

Lemma 5. *For every bisection $\mathcal{S} = (S, \overline{S})$ and every vertex u, a u-pair for $\mathcal{S}$ exists if and only if u is non-stubborn.*

Hence, if there is a non-stubborn vertex u, there is a u-pair (A_u, B_u) for S, and u is certainly a good vertex for the bisection $\mathcal{T}$ associated to this u-pair. Therefore, if $\mathcal{T}$ is not good then it must be that there is a vertex y that is an obstruction for $\mathcal{T}$. In the last case, we will say that the vertex u, the u-pair (A_u, B_u) and the bisection $\mathcal{T}$ are *obstructed* by y. Most of the proof will be devoted to dealing with these obstructions.

3 Main Theorem

Our main result is the following.

Theorem 1. *Every graph G with an odd number of vertices and at least one non-stubborn vertex has a subvertable belief assignment* $\mathbf{b}$ *and a swing vertex u for* $\mathbf{b}$*. Moreover,* $\mathbf{b}$ *and u can be computed in polynomial time.*

We prove the theorem by exhibiting a polynomial-time algorithm (see Algorithm 1) that, given a graph G with an odd number of vertices, and at least one of which that is non-stubborn, returns a good bisection $\mathcal{S}$ and a good vertex u for $\mathcal{S}$. The theorem then follows from Lemma 1.

Input: A graph $G = (V, E)$ with $|V|$ odd and at least one non-stubborn vertex
Output: A pair $(\mathcal{S}, u)$ where $\mathcal{S}$ is a good bisection and u is its good vertex

$\mathcal{S} = (S, \overline{S})$ is a bisection of G of 3-minimal potential
$M=$ non-stubborn vertices of minimum rank in $\mathcal{S}$
if *there is* $u \in S$ *with* $\mathsf{adv}_{\mathcal{S}}(u) \leq -a_u - 1$ **then**
    Let $\mathcal{T} = (\overline{S} \cup \{u\}, S \setminus \{u\})$
    return $(\mathcal{T}, u)$
if *there is* $u \in S$ *with* $\mathsf{adv}_{\mathcal{S}}(u) \geq a_u + 1$ **then**
    return $(\mathcal{S}, u)$
if *there is* $u \in \overline{S}$ *with* $\mathsf{adv}_{\mathcal{S}}(u) \geq a_u + 1$ **then**
    Pick $w \in S$ and let $\mathcal{T} = (\overline{S} \cup \{w\}, S \setminus \{w\})$
    return $(\mathcal{T}, u)$
if *there is* $u \in S \cap M$ *with* $\mathsf{adv}_{\mathcal{S}}(u) < 0$ **then**
    Let $\mathcal{S}' = (\overline{S} \cup \{u\}, S \setminus \{u\})$
    Pick u-pair (A_u, B_u) for $\mathcal{S}'$
    Let $\mathcal{T}$ be the associated bisection
    return $(\mathcal{T}, u)$
if $M \cap \overline{S} \neq \emptyset$ **then return** `MinRankInNotS(`$\mathcal{S}$`)`
else return `MinRankInS(`$\mathcal{S}$`)`

Algorithm 1. Returns a good bisection and a good vertex

First, we note that the algorithm runs in time that is polynomial in the size of the input. Indeed, a bisection of 3-minimal potential at Line 1 can be efficiently computed through a local search algorithm [20], and all remaining steps only involve computationally easy tasks.

Next we prove that the algorithm is correct; that is, it outputs $(\mathcal{T}, u)$ where $\mathcal{T}$ is a good bisection and u is a good vertex for $\mathcal{T}$. Recall that, by Lemma 4, it is sufficient to check that $\mathsf{adv}_{\mathcal{T}}(u) \geq a_u + 1$ and that non-stubborn vertices $x \in S$ have $\mathsf{adv}_{\mathcal{T}}(x) \geq -a_x$.

The analysis of the algorithm can be divided in three parts: the warm-up cases, i.e., if Algorithm 1 stops before reaching Line 16; the case there is a non-stubborn vertex $u \in \overline{S}$ of minimum rank; and the case that every non-stubborn vertex u of minimum rank belongs to S. Due to the page limit, we only sketch the proof for the last and most interesting case, i.e. when the algorithm invokes procedure `MinRankInS` (described in the full version of the paper [5]).

Suppose then that the algorithm invokes procedure `MinRankInS`. In this case, all non-stubborn vertices of minimum rank belong to S. Moreover, all such vertices have non-negative advantage for otherwise the Algorithm would have stopped at Line 15.

Clearly, if `MinRankInS` stops at Line 5, Line 18, Line 26, Line 32, Line 39 or Line 49, then the bisection output is good and u is a good vertex for it.

Suppose now that `MinRankInS` stops at Line 9, Line 30, Line 36, Line 44, Line 47 or at Line 52. Since in all cases the algorithm returns a pair $(\mathcal{T}, v)$ where $\mathcal{T}$ is the bisection associated to a v-pair, then, by Lemma 3, $\mathsf{adv}_{\mathcal{T}}(v) \geq a_v + 1$. Thus, we only need to prove that $\mathsf{adv}_{\mathcal{T}}(x) \geq -a_x$ for every non-stubborn $x \in T$.

3.1 Properties of the Obstructions

Most of the work will be devoted to dealing with obstructions. Therefore, before proceeding, we give some useful properties of the obstructions, whose proof can be found in the full version.

Lemma 6. *Let $u \in \overline{S}$ be a vertex of minimum rank for the bisection $\mathcal{S}$ and let y be an obstruction for u. Then $y \in \overline{S}$. Similarly, let $u \in S$ be a vertex of minimum rank for the bisection $\mathcal{S}$ and assume there is no vertex of minimum rank in $\overline{S}$. If y is an obstruction for u, then $y \in S$.*

Lemma 7. *Let $\mathcal{S}$ be a bisection and let u be a vertex of minimum rank in $\overline{S}$. Let $\mathcal{T}$ be the bisection associated with a u-pair (A_u, B_u) for $\mathcal{S}$. If vertex y is an obstruction for $\mathcal{T}$, then $\mathsf{adv}_{\mathcal{S}}(y) \leq -a_y + 2\mathsf{rank}_{\mathcal{S}}(u) - 3$. Moreover, for every non-stubborn $v \in S$ if $\mathsf{adv}_{\mathcal{S}}(v) + \mathsf{adv}_{\mathcal{S}}(y) + 2W(v, y) \geq a_v - a_y$, then v is adjacent to y, v has minimum rank and $\mathsf{adv}_{\mathcal{S}}(y) \geq -a_y + 2\mathsf{rank}_{\mathcal{S}}(u) - 4$.*

Lemma 8. *Let $\mathcal{S}$ be a bisection and suppose that there is no vertex in $\overline{S}$ with minimum rank. Let u be a vertex of minimum rank in S. Let $\mathcal{T}$ be the bisection associated with a u-pair (A_u, B_u) for $\mathcal{S}$. Suppose there is an obstruction y for $\mathcal{T}$ with $\mathsf{adv}_{\mathcal{S}}(y) < 0$ and $\mathsf{rank}_{\mathcal{S}}(y) > \mathsf{rank}_{\mathcal{S}}(u)$. Let $\mathcal{S}' = (\overline{S} \cup \{y\}, S \backslash \{y\})$. Then $\mathsf{rank}_{\mathcal{S}'}(y) \leq \mathsf{rank}_{\mathcal{S}}(u)$.*

Lemma 9. *Let $\mathcal{S}$ be a bisection and let u be a vertex of minimum rank in S. Let $\mathcal{T}$ be the bisection associated with a u-pair (A_u, B_u) for $\mathcal{S}$. Suppose there is an obstruction y for $\mathcal{T}$ with $\mathsf{adv}_{\mathcal{S}}(y) \geq 0$. Then y has minimum rank $\ell = \left\lceil \frac{a_y+1}{2} \right\rceil$ and $\mathsf{adv}_{\mathcal{S}}(y) = 0$.*

3.2 `MinRankInS` stops at Line 9

In this case, we have that u is a vertex of S with minimum rank ℓ. Vertex y is an obstruction of bisection $\mathcal{T}$ associated with u-pair (A_u, B_u), and $\mathsf{adv}_{\mathcal{S}}(y) < 0$. By Lemma 6, $y \in S$. Observe that $\mathsf{rank}_{\mathcal{S}}(y) > \ell$, for otherwise Algorithm 1 would have stopped at Line 15. From Lemma 8, we obtain that $\mathsf{rank}_{\mathcal{S}_0}(y) \leq \ell$. We remind the reader that $\mathcal{S}_0 = (\overline{S} \cup \{y\}, S \backslash \{y\})$ and $\mathcal{T}_0 = (\overline{S} \cup \{y\} \backslash A_y \cup B_y, S \backslash \{y\} \cup A_y \backslash B_y)$.

For every non-stubborn $x \in T_0 \backslash \{y\}$, $\mathsf{adv}_{\mathcal{T}}(x)$ can be written as: $W(x, \overline{S}) - W(x, S) + 2W(x, y) - 2W(x, A_y) + 2W(x, B_y)$.

If $x \in \overline{S} \backslash A_y$, then

$$\begin{aligned}\mathsf{adv}_{\mathcal{T}}(x) &= \mathsf{adv}_{\mathcal{S}}(x) + 2W(x,y) - 2W(x, A_y) + 2W(x, B_y) \\ &\geq \mathsf{adv}_{\mathcal{S}}(x) + 2W(x,y) - 2|A_y| \\ &= \mathsf{adv}_{\mathcal{S}}(x) + 2W(x,y) - 2\mathsf{rank}_{\mathcal{S}'}(y) \geq \mathsf{adv}_{\mathcal{S}}(x) + 2W(x,y) - 2\ell.\end{aligned}$$

Since $\mathsf{rank}_{\mathcal{S}}(y) > \ell$, then $\mathsf{adv}_{\mathcal{S}}(y) \leq a_y - 2\ell$. By applying Lemma 2 to $y \in S$ and $x \in \overline{S}$ we obtain that $\mathsf{adv}_{\mathcal{S}}(x) + 2W(x,y) \geq -\mathsf{adv}_{\mathcal{S}}(y) + a_y - a_x \geq -a_x + 2\ell$. Hence $\mathsf{adv}_{\mathcal{T}}(x) \geq -a_x$.

Finally, if $x \in B_y$, then $x \in S$ and, by definition of y-pair, $W(x, y) = 1$. Therefore we have

$$\begin{aligned}\mathsf{adv}_{\mathcal{T}}(x) &= -\mathsf{adv}_{\mathcal{S}}(x) + 2W(x,y) - 2W(x, A_y) + 2W(x, B_y) \\ &\geq -\mathsf{adv}_{\mathcal{S}}(x) - 2W(x, A_y) + 2 \\ &\geq -\mathsf{adv}_{\mathcal{S}}(x) - 2\mathsf{rank}_{\mathcal{S}'}(y) + 2 \geq -\mathsf{adv}_{\mathcal{S}}(x) - 2\ell + 2\end{aligned}$$

Since ℓ is the minimum rank, it must be the case that $\mathsf{rank}_{\mathcal{S}}(x) \geq \ell$ which implies that $\mathsf{adv}_{\mathcal{S}}(x) \leq a_x + 2 - 2\ell$. Therefore, $\mathsf{adv}_{\mathcal{T}}(x) \geq -a_x$.

3.3 `MinRankInS` reaches Line 19

We remind the reader that in this case $u \in S$ is a non-stubborn vertex of minimum rank ℓ and y is an obstruction to bisection $\mathcal{T}$ associated with u-pair (A_u, B_u) for $\mathcal{S}$. By Lemma 6, $y \in S$. Note also that $\mathsf{adv}_{\mathcal{S}}(y) \geq 0$, for otherwise `MinRankInS` would have stopped at Line 9. From Lemma 9, it then follows that $\mathsf{adv}_{\mathcal{S}}(y) = 0$ and $\mathsf{rank}_{\mathcal{S}}(y) = \ell = \left\lceil \frac{a_y+1}{2} \right\rceil$. Moreover, given a y-pair (A_y, B_y) of $\mathcal{S}_1 = (\overline{S} \cup \{y\}, S \backslash \{y\})$, y_1 is either a vertex of $\left(\overline{S} \cup \{y\} \backslash A_y\right) \cap \overline{N}(y)$ with $\mathsf{adv}_{\mathcal{S}}(w) = a_y - a_{y_1}$ or it is an obstruction to bisection $\mathcal{T}_1$ associated with this pair. Note that, by Lemma 6, even in this last case $y_1 \in \overline{S} \cup \{y\} \backslash A_y$.

Properties of y and y_1. We need to state some properties of y and y_1 before proving that the bisections returned by `MinRankInS` after Line 19 are good.

Lemma 10. *$W(y, y_1) = 0$.*

Lemma 11. *If $\mathsf{adv}_{\mathcal{S}}(y_1) \neq a_y - a_{y_1}$, then a_y is even and $\mathsf{adv}_{\mathcal{S}}(y_1) = a_y - a_{y_1} + 1$.*

Lemma 12. *If $\mathsf{adv}_{\mathcal{S}}(y_1) \neq a_y - a_{y_1}$, then $\mathsf{adv}_{\mathcal{S}}(w) \geq a_y - a_w + 1 - W(w, y)$ for every $w \in \overline{S}$.*

Lemma 13. *$\mathsf{rank}_{\mathcal{S}_2}(y) = \mathsf{rank}_{\mathcal{S}_2}(y_1) = \ell$.*

Lemma 14. *For every $u \in S_2$ and every $v \in \overline{S}_2 \backslash \{y\}$, we have that if $\mathsf{adv}_{\mathcal{S}}(y_1) = a_y - a_{y_1}$ or $u = y_1$, then $\mathsf{adv}_{\mathcal{S}_2}(u) + \mathsf{adv}_{\mathcal{S}_2}(v) + 2W(u, v) \geq a_u - a_v$, else $\mathsf{adv}_{\mathcal{S}_2}(u) + \mathsf{adv}_{\mathcal{S}_2}(v) + 2W(u, v) = a_u - a_v + c + 2W(y, v) + 2W(u, y_1) - 2W(u, y) - 2W(v, y_1)$, for $c \geq \max\{0, 2W(u, v) - 2W(y_1, u) - W(y, v)\}$.*

Lemma 15. *For every $u \in S \backslash \{y_1\}$ and $v \in \overline{S} \backslash \{y\}$, if $W(u, y) = 1$ and $W(u, y_1) = 0$, then $\mathsf{adv}_{\mathcal{S}_2}(u) + \mathsf{adv}_{\mathcal{S}_2}(v) + 2W(u, v) \geq a_u - a_v - 1$.*

`MinRankInS` stops at Line 30. Therefore there is $v \in S_2$, whose rank in $\mathcal{S}_2$ is less than $\mathsf{rank}_{\mathcal{S}_2}(y) = \ell$. Note that, since $\mathsf{adv}_{\mathcal{S}_2}(v) = \mathsf{adv}_{\mathcal{S}}(v) + 2W(v, y_1) - 2W(v, y)$,

$$\begin{aligned} \mathsf{rank}_{\mathcal{S}_2}(v) &= \left\lceil \frac{a_v + 1 - \mathsf{adv}_{\mathcal{S}}(v) - 2W(v, y_1) + 2W(v, y)}{2} \right\rceil \\ &= \left\lceil \frac{a_v + 1 - \mathsf{adv}_{\mathcal{S}}(v)}{2} \right\rceil - W(v, y_1) + W(v, y) \\ &= \mathsf{rank}_{\mathcal{S}}(v) - W(v, y_1) + W(v, y). \end{aligned}$$

Hence, $\mathsf{rank}_{\mathcal{S}_2}(v) < \ell$ if and only if $\mathsf{rank}_{\mathcal{S}}(v) = \ell$ (that is, v has minimum rank in $\mathcal{S}$), $W(v, y_1) = 1$ and $W(v, y) = 0$. From this we obtain that for every vertex v with $\mathsf{rank}_{\mathcal{S}_2}(v) < \ell$, it holds that $\mathsf{adv}_{\mathcal{S}}(v) \geq 0$ (since v has minimum rank in S and no vertex of minimum rank in S with negative advantage can exist, otherwise a good bisection was returned at Line 15 of Algorithm 1), and, $\mathsf{adv}_{\mathcal{S}_2}(v) \geq 2$. We also observe that every vertex $x \in \overline{S}_2 = \overline{S} \cup \{y\} \backslash \{y_1\}$ has $\mathsf{rank}_{\mathcal{S}_2}(x) \geq \ell$. If $x = y$, then this follows from Lemma 13. If $x \neq y$, then it follows since $\mathsf{rank}_{\mathcal{S}}(x) \geq \ell + 1$, and the rank decreases of at most one when two vertices are swapped.

Moreover, if `MinRankInS` stops at Line 30, then the bisection $\mathcal{T}_2$ associated to v-pair (A_v, B_v) for $\mathcal{S}_2$ has an obstruction y_2. By Lemma 6, $y_2 \in S_2 \backslash A_v$. Suppose that $\mathsf{adv}_{\mathcal{S}_2}(y_2) \geq 0$, then, from Lemma 9, it follows that $\mathsf{adv}_{\mathcal{S}_2}(y_2) = 0$ and has minimum rank, i.e., $\mathsf{rank}_{\mathcal{S}_2}(y_2) = \ell - 1$. However, this is a contradiction, since we showed that if $\mathsf{rank}_{\mathcal{S}_2}(y_2) = \ell - 1$, then $\mathsf{adv}_{\mathcal{S}_2}(y_2) \geq 2$.

It must be then the case that $\mathsf{adv}_{\mathcal{S}_2}(y_2) < 0$ and $\mathsf{rank}_{\mathcal{S}_2}(y_2) \geq \ell$. Then, by Lemma 8, $\mathsf{rank}_{\mathcal{S}_3}(y_2) \leq \ell - 1$, where $\mathcal{S}_3 = (\overline{S}_2 \cup \{y_2\}, S_2 \cup \{y_2\})$. It must be also the case that either $\mathsf{rank}_{\mathcal{S}}(y_2) \geq \ell + 1$ or $\mathsf{rank}_{\mathcal{S}}(y_2) = \ell$, $W(y_2, y) = 1$ and $W(y_2, y_1) = 0$. Indeed, $\mathsf{rank}_{\mathcal{S}}(y_2) \geq \ell$, since ℓ is the minimum rank in $\mathcal{S}$. If $\mathsf{rank}_{\mathcal{S}}(y_2) = \ell$, then $\mathsf{adv}_{\mathcal{S}}(y_2) \geq 0$. Thus, if $W(y_2, y) = 0$ and $W(y_2, y_1) = 1$, then $\mathsf{rank}_{\mathcal{S}_2}(y_2) = \ell - 1$, a contradiction. If $W(y_2, y) = W(y_2, y_1)$, then $\mathsf{adv}_{\mathcal{S}_2}(y_2) = \mathsf{adv}_{\mathcal{S}}(y_2) \geq 0$, still a contradiction.

We now can prove that the bisection $\mathcal{T}_3$ returned at Line 30 is good. Recall that $\mathcal{T}_3$ is associated to y_2-pair (A_{y_2}, B_{y_2}) for $\mathcal{S}_3$, i.e., $\mathcal{T}_3 = (\overline{S}_2 \cup \{y_2\} \backslash A_{y_2} \cup B_{y_2}, S_2 \backslash \{y_2\} \cup A_{y_2} \backslash B_{y_2})$, where $|A_{y_2}| = |B_{y_2}| = \mathsf{rank}_{\mathcal{S}_3}(y_2) \leq \ell - 1$.

We first prove that for every $x \in \overline{S}_2 \cup \{y_2\} \backslash A_{y_2}$, we have that $\mathsf{adv}_{\mathcal{T}_3}(x) \geq -a_x$. If $x \neq y$, we distinguish two cases. If $\mathsf{rank}_{\mathcal{S}}(y_2) \geq \ell+1$, then by applying Lemma 2 to $y_2 \in S$ and $x \in \overline{S}$, we have that $\mathsf{adv}_{\mathcal{S}}(x) + 2W(x, y_2) \geq -\mathsf{adv}_{\mathcal{S}}(y_2) + a_{y_2} - a_x \geq -a_x + 2\ell$, where we used that $\mathsf{rank}_{\mathcal{S}}(y_2) \geq \ell + 1$ and thus $\mathsf{adv}_{\mathcal{S}}(y_2) \leq a_{y_2} - 2\ell$. Then, $\mathsf{adv}_{\mathcal{S}_2}(x) = \mathsf{adv}_{\mathcal{S}}(x) + 2W(x, y) - 2W(x, y_1) \geq -a_x + 2\ell - 2$. If $\mathsf{rank}_{\mathcal{S}}(y_2) < \ell + 1$, then, as stated above, it must be the case that $\mathsf{rank}_{\mathcal{S}}(y) = \ell$, $W(y_2, y) = 1$ and $W(y_2, y_1) = 0$. Then, from Lemma 15, it holds that $\mathsf{adv}_{\mathcal{S}_2}(x) + 2W(x, y_2) \geq -\mathsf{adv}_{\mathcal{S}_2}(y_2) + a_{y_2} - a_x - 1 \geq -a_x + 2\ell - 1$, where we used that $\mathsf{rank}_{\mathcal{S}_2}(y_2) = \mathsf{rank}_{\mathcal{S}}(y_2) + 1 = \ell + 1$ and thus $\mathsf{adv}_{\mathcal{S}}(y_2) \leq a_{y_2} - 2\ell$.

Hence, in both cases, we have $\mathsf{adv}_{\mathcal{T}_3}(x) \geq \mathsf{adv}_{\mathcal{S}_3}(x) - 2W(x, A_{y_2}) \geq \mathsf{adv}_{\mathcal{S}_2}(x) + 2W(x, y_2) - 2(\ell - 1) \geq -a_x + 2\ell - 1 - 2(\ell - 1) \geq -a_x + 1$. If $x = y$, then, by using that $\mathsf{adv}_{\mathcal{S}_2}(y) = -\mathsf{adv}_{\mathcal{S}}(y)$ since $W(y, y_1) = 0$, we have $\mathsf{adv}_{\mathcal{T}_3}(y) \geq \mathsf{adv}_{\mathcal{S}_3}(y) - 2(\ell - 1) = \mathsf{adv}_{\mathcal{S}_2}(y) + 2W(y, y_2) - 2(\ell - 1) = -\mathsf{adv}_{\mathcal{S}}(y) + 2W(y, y_2) - 2(\ell - 1)$. We showed above that $\mathsf{adv}_{\mathcal{S}}(y) = 0$ and $\ell = \left\lceil \frac{a_y+1}{2} \right\rceil \leq \frac{a_y+2}{2}$. Hence, $\mathsf{adv}_{\mathcal{T}_3}(y) \geq -a_y + 2W(y, y_2) \geq -a_y$.

Finally, we prove that for all $x \in B_{y_2}$, $\mathsf{adv}_{\mathcal{T}_3}(x) \geq -a_x$. Recall that $B_{y_2} \subseteq S_2 \backslash \{y_2\}$ and $W(x, y_2) = 1$ for all $x \in B_{y_2}$. We have two cases. If $x \neq y_1$, then $\mathsf{adv}_{\mathcal{T}_3}(x) \geq -\mathsf{adv}_{\mathcal{S}_3}(x) - 2(\ell - 1) = -\mathsf{adv}_{\mathcal{S}_2}(x) + 2W(x, y_2) - 2(\ell - 1) = -\mathsf{adv}_{\mathcal{S}}(x) + 2W(x, y) - 2W(x, y_1) - 2(\ell - 2) \geq -a_x + 2\ell - 2 + 2W(x, y) - 2W(x, y_1) + 2 - 2(\ell - 1) \geq -a_x$, where we used that $\mathsf{rank}_{\mathcal{S}}(x) \geq \ell$ and thus $\mathsf{adv}_{\mathcal{S}}(x) \leq a_x - 2\ell + 2$. If $x = y_1$, then $\mathsf{adv}_{\mathcal{T}_3}(y_1) \geq -\mathsf{adv}_{\mathcal{S}_3}(y_1) - 2(\ell - 1) = -\mathsf{adv}_{\mathcal{S}_2}(y_1) + 2W(y_1, y_2) - 2(\ell - 1) = \mathsf{adv}_{\mathcal{S}}(y_1) + 2 - 2(\ell - 1)$, where we used that $\mathsf{adv}_{\mathcal{S}_2}(y_1) = -\mathsf{adv}_{\mathcal{S}}(y_1)$ since $W(y, y_1) = 0$ and $W(y_1, y_2) = 1$ because $y_1 \in B_{y_2}$. Since $\mathsf{adv}_{\mathcal{S}}(y_1) \geq a_y - a_{y_1} \geq 2(\ell - 1) - a_{y_1}$, then $\mathsf{adv}_{\mathcal{T}_3}(y_1) \geq -a_{y_1} + 2 \geq -a_{y_1}$.

`MinRankInS` stops at Line 36. In this case y and y_1 have minimum rank in $\mathcal{S}_2$ and there is a vertex $w \in S \backslash \{y\} \cup \{y_1\}$ of minimum rank ℓ and negative advantage in $\mathcal{S}_2$. Note that if $\mathsf{rank}_{\mathcal{S}_2}(w) \leq \mathsf{rank}_{\mathcal{S}}(w)$, then $\mathsf{adv}_{\mathcal{S}_2}(w) \geq \mathsf{adv}_{\mathcal{S}}(w)$. Thus, since in $\mathcal{S}$ all vertices of minimum rank have non-negative advantage, it must be the case that $\mathsf{rank}_{\mathcal{S}}(w) = \ell + 1$, and $W(w, y) = 0$ and $W(w, y_1) = 1$. Thus the w-pair defined at Line 35 can be constructed.

Consider now the bisection $\mathcal{S}_4$ defined at Line 34. Observe that $\mathsf{adv}_{\mathcal{S}_4}(w) = -\mathsf{adv}_{\mathcal{S}_2}(w)$ and therefore

$$\begin{aligned}\mathsf{rank}_{\mathcal{S}_4}(w) &= \left\lceil \frac{a_w + 1 - \mathsf{adv}_{\mathcal{S}_4}(w)}{2} \right\rceil = \left\lceil \frac{a_w + 1 + \mathsf{adv}_{\mathcal{S}_4}(w)}{2} \right\rceil - \mathsf{adv}_{\mathcal{S}_4}(w) \\ &= \left\lceil \frac{a_w + 1 - \mathsf{adv}_{\mathcal{S}_2}(w)}{2} \right\rceil + \mathsf{adv}_{\mathcal{S}_2}(w) = \mathsf{rank}_{\mathcal{S}_2}(w) + \mathsf{adv}_{\mathcal{S}_2}(w),\end{aligned}$$

that is at most $\ell - 1$ since w has rank ℓ and negative advantage in $\mathcal{S}_2$.

Now, for every $x \in \overline{S}_2 \backslash A_w$, we have $\mathsf{adv}_{\mathcal{T}_5}(x) \geq \mathsf{adv}_{\mathcal{S}_2}(x) + 2W(x, w) - 2\mathsf{rank}_{\mathcal{S}_4}(w) \geq -\mathsf{adv}_{\mathcal{S}_2}(w) + a_w - a_x - 2(\ell - 1)$, where we used that, by Lemma 14, $\mathsf{adv}_{\mathcal{S}_2}(x) + 2W(x, w) \geq -\mathsf{adv}_{\mathcal{S}_2}(w) + a_w - a_x + 2W(y, x) + 2W(w, y_1) - 2W(w, y) -$

$2W(x, y_1) \geq -\mathsf{adv}_{\mathcal{S}_2}(w) + a_w - a_x$. Since $\mathsf{rank}_{\mathcal{S}_2}(w) = \ell$, then $\mathsf{adv}_{\mathcal{S}_2}(w) \leq a_w - 2\ell + 2$, from which we achieve that $\mathsf{adv}_{\mathcal{T}_5}(x) \geq -a_x$.

Finally, take $x \in B_w \subseteq S_2$. We have $\mathsf{adv}_{\mathcal{T}_5}(x) \geq -\mathsf{adv}_{\mathcal{S}_2}(x) - 2\mathsf{rank}_{\mathcal{S}_4}(w) \geq -\mathsf{adv}_{\mathcal{S}_2}(x) - 2(\ell - 1)$. However, by hypothesis, w has minimum rank among the non-stubborn vertices and thus it must be the case that $\mathsf{rank}_{\mathcal{S}_2}(x) \geq \mathsf{rank}_{\mathcal{S}_2}(w) = \ell$ which implies that $\mathsf{adv}_{\mathcal{S}_2}(x) \leq a_x - 2\ell + 2$. Therefore, $\mathsf{adv}_{\mathcal{T}}(x) \geq -a_x$.

There is still a missing case, for which we refer the reader to the full version.

4 Lower Bound

We next show that deciding if it is possible to subvert the majority when starting from a weaker minority is a computationally hard problem, even if we start with a minority of size very close to $\frac{n-1}{2}$. The main result of this section is given by the following theorem.

Theorem 2. *For every constant $0 < \varepsilon < \frac{133}{155}$, it is NP-hard to decide whether in a graph G with n vertices there exists a subvertable belief assignment with at most $\frac{n-1}{2}(1 - \varepsilon)$ vertices in the initial minority.*

The proof of Theorem 2 uses essentially the same gadgets as a similar proof in [3], but tuned for the current setting.

5 Open Problems

While this work proves information retention for the heterogeneous majority dynamics in unweighted social network when only one player is allowed to update her state at each time step, it would be interesting to understand what happens if one considers weighted graphs or concurrent updates. Preliminary experimental results along this direction have been given in [6]. It would be also interesting to investigate the extent at which the mbM phenomenon occurs if one considers noisy variants of the heterogeneous majority dynamics, see, e.g., [7,8]. Finally, one can be interested in understanding how probable the minority becomes majority phenomenon is, and how is this frequency related to the topological properties of the network.

Acknowledgments. This work was partially supported by a Caratheodory research grant E.114 from the University of Patras and by the "GNCS – INdAM". Part of this work done while G. Persiano was visiting Google, New York.

References

1. Acar, E., Greco, G., Manna, M.: Group reasoning in social environments. In: Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems, AAMAS 2017, 8–12 May 2017, São Paulo, Brazil, pp. 1296–1304 (2017)
2. Angluin, D., Aspnes, J., Eisenstat, D.: A simple population protocol for fast robust approximate majority. Distrib. Comput. **21**(2), 87–102 (2008)

3. Auletta, V., Caragiannis, I., Ferraioli, D., Galdi, C., Persiano, G.: Minority becomes majority in social networks. In: Markakis, E., Schäfer, G. (eds.) WINE 2015. LNCS, vol. 9470, pp. 74–88. Springer, Heidelberg (2015). https://doi.org/10.1007/978-3-662-48995-6_6
4. Auletta, V., Caragiannis, I., Ferraioli, D., Galdi, C., Persiano, G.: Generalized discrete preference games. In: Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI 2016, New York, NY, USA, 9–15 July 2016, pp. 53–59 (2016)
5. Auletta, V., Caragiannis, I., Ferraioli, D., Galdi, C., Persiano, G.: Information retention in heterogeneous majority dynamics. arXiv preprint arXiv:1603.02971 (2016)
6. Auletta, V., Caragiannis, I., Ferraioli, D., Galdi, C., Persiano, G.: Robustness in discrete preference games. In: Proceedings of the 2017 International Conference on Autonomous Agents & Multiagent Systems, 8–12 May 2017, Sao Paulo (2017)
7. Auletta, V., Ferraioli, D., Pasquale, F., Penna, P., Persiano, G.: Logit dynamics with concurrent updates for local interaction games. In: Bodlaender, H.L., Italiano, G.F. (eds.) ESA 2013. LNCS, vol. 8125, pp. 73–84. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40450-4_7
8. Auletta, V., Ferraioli, D., Pasquale, F., Persiano, G.: Mixing time and stationary expected social welfare of logit dynamics. Theor. Comput. Syst. **53**(1), 3–40 (2013)
9. Berger, E.: Dynamic monopolies of constant size. J. Comb. Theor. Ser. B **83**(2), 191–200 (2001)
10. Bhawalkar, K., Gollapudi, S., Munagala, K.: Coevolutionary opinion formation games. In: STOC, pp. 41–50. ACM (2013)
11. Bilò, V., Fanelli, A., Moscardelli, L.: Opinion formation games with dynamic social influences. In: Cai, Y., Vetta, A. (eds.) WINE 2016. LNCS, vol. 10123, pp. 444–458. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-54110-4_31
12. Bindel, D., Kleinberg, J.M., Oren, S.: How bad is forming your own opinion? In: Proceedings of the 52nd Annual IEEE Symposium on Foundations of Computer Science (FOCS), pp. 55–66 (2011)
13. Chierichetti, F., Kleinberg, J.M., Oren, S.: On discrete preferences and coordination. In: Proceedings of the 14th ACM Conference on Electronic Commerce (EC), pp. 233–250 (2013)
14. DeGroot, M.H.: Reaching a consensus. J. Am. Stat. Assoc. **69**(345), 118–121 (1974)
15. Ferraioli, D., Goldberg, P.W., Ventre, C.: Decentralized dynamics for finite opinion games. Theor. Comput. Sci. **648**, 96–115 (2016)
16. Friedkin, N.E., Johnsen, E.C.: Social influence and opinions. J. Math. Sociol. **15**(3–4), 193–205 (1990)
17. Mertzios, G.B., Nikoletseas, S.E., Raptopoulos, C.L., Spirakis, P.G.: Determining majority in networks with local interactions and very small local memory. Distrib. Comput. **30**(1), 1–16 (2017)
18. Mossel, E., Neeman, J., Tamuz, O.: Majority dynamics and aggregation of information in social networks. Auton. Agent. Multi-agent Syst. **28**(3), 408–429 (2014)
19. Peleg, D.: Local majorities, coalitions and monopolies in graphs: a review. Theor. Comput. Sci. **282**, 231–257 (2002)
20. Schäffer, A.A., Yannakakis, M.: Simple local search problems that are hard to solve. SIAM J. Comput. **20**(1), 56–87 (1991)
21. Tamuz, O., Tessler, R.J.: Majority dynamics and the retention of information. Isr. J. Math. **206**(1), 483–507 (2015)