

Dynamics of Evolving Social Groups

NOGA ALON, Tel-Aviv University and Microsoft Research

MICHAL FELDMAN, Tel-Aviv University and Microsoft Research

YISHAY MANSOUR, Tel-Aviv University and Microsoft Research

SIGAL OREN, Ben-Gurion University of the Negev

MOSHE TENNENHOLTZ, Technion - Israel Institute of Technology

Exclusive social groups are ones in which the group members decide whether or not to admit a candidate to the group. Examples of exclusive social groups include academic departments and fraternal organizations. In the present paper we introduce an analytic framework for studying the dynamics of exclusive social groups. In our model, every group member is characterized by his opinion, which is represented as a point on the real line. The group evolves in discrete time steps through a voting process carried out by the group's members. Due to homophily, each member votes for the candidate who is more similar to him (i.e., closer to him on the line). An admission rule is then applied to determine which candidate, if any, is admitted. We consider several natural admission rules including majority and consensus.

We ask: how do different admission rules affect the composition of the group in the long term? We study both growing groups (where new members join old ones) and fixed-size groups (where new members replace those who quit). Our analysis reveals intriguing phenomena and phase transitions, some of which are quite counterintuitive.

1. INTRODUCTION

Exclusive social groups (a.k.a. clubs) are those in which group members decide which new members to admit. Many of the social groups we are part of and are aware of are in fact exclusive groups. Examples that readily come to mind include academic departments and the National Academy of Sciences, where current members decide which members to accept. Additional examples from different areas of life are abundant, ranging from becoming a Freemason to getting the privilege to live in a condominium or a Kibbutz.

Some exclusive social groups, like academic departments, grow; others, like condominiums, have a fixed size. But many share a similar admission process: each member votes for a candidate he wants to admit to the group, and if the candidate receives sufficient votes he can join. Different groups require candidates to obtain different fractions of the votes to be admitted, from a simple majority to a consensus. Because members who join now affect those who will join in the future, different admission rules can lead to substantially different compositions as the group evolves. In particular, common wisdom suggests that requiring a greater fraction of group members to agree on a candidate increases the homogeneity of the group. But is this really so?

The broad question we address here is how different admission rules affect the composition of the types of members in the group, and how this composition evolves over time. The question comes in two flavors: (i) the growing group model, where the size of the group increases as new members join, and (ii) the fixed-size group model, where newly admitted members replace those who quit.

To answer this question we need to formalize several aspects of the group members and the way they vote. As common in the political and sociological literature (e.g., [DeGroot 1974; Friedkin and Johnsen 1990; Roemer 2001]), we assume that every group member and candidate has an “opinion,” which is a real number in the interval $[0, 1]$. For example, the opinion can be a political inclination on the spectrum between left and right or for the academic world how theoretical or applied one's research is.

The opinions of members and candidates form the basis for modeling the voting process. We assume that when making a choice between several candidates, each member chooses the one who is the most similar to himself. As opinions are real numbers, sim-

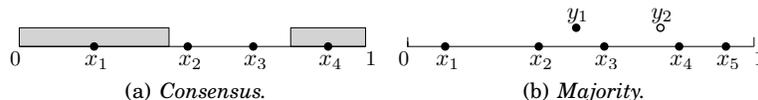


Fig. 1. Different admission rules: (a) for consensus, the intervals in which a new candidate can be accepted (marked in gray) are determined by the location of x_1 and x_4 . (b) for the majority rule, the candidate closer to the median, x_3 , will be accepted.

ilarity can be easily measured by the distance between opinions. This modeling choice is heavily grounded in the literature on homophily (e.g., [McPherson et al. 2001] and the references therein), stating that people prefer the company and tend to interact more with others who are more similar to them. In the words of Aristotle, people “love those who are like themselves.”

The use of homophily as the driving force behind the members votes ties the present paper to two large bodies of literature, one on opinion formation ([Ben-Naim et al. 2003; Deffuant et al. 2000; Hegselmann and Krause 2002]), the other on cultural dynamics ([Axelrod 1997; Flache and Macy 2011; Kempe et al. 2013]). Both bodies of work aim to understand the mechanisms by which individuals form their opinions (in the cultural dynamics literature it can be opinions on several issues) and how different mechanisms affect the distribution of opinions in society. In both bodies of work it is common to assume that similar individuals have a greater chance of influencing one another. Many of these models, however, are quite difficult to analyze, and therefore most of the literature has restricted its attention to models that operate on a fixed network that does not evolve over time. In a sense, we bypass this difficulty by de-emphasizing the network structure, and instead focus on the aggregate effect of group members choices. This modeling decision enables us to study the evolution of the social group over time as a function of the admission rule applied.

The present paper explores a variety of admission rules and their effects on the opinion distribution as the group evolves. We begin by presenting our results for growing groups, then proceed to fixed-size groups.

1.1. Growing Social Groups

Our models fit the following framework: each group member is characterized by his opinion $x_i \in [0, 1]$. A group of size k is denoted by $S(k) \in [0, 1]^k$. The admission process operates in discrete time steps. At each time step two candidates are considered for admission; their opinions, y_1, y_2 , are drawn from the uniform distribution $\mathcal{U}[0, 1]$. Each member, i , votes for the candidate that is more similar to him, that is, a candidate j , minimizing $|x_i - y_j|$. Finally, based on the members’ votes, an admission rule is used to determine which candidate is accepted to the group.

Consensus and majority. Two natural admission rules that come to mind are consensus (used, for example, by the Freemasons) and majority. In the first case, a candidate is accepted only if he receives the votes of all the members, so that in some steps no candidate is admitted. In the second case, the candidate preferred by the majority of members is admitted. As noted, it is natural to expect that requiring a greater fraction of the members to prefer one candidate over the other can only increase the homogeneity of the group. Our work suggests that this is not always the case. Indeed, if the consensus admission rule is applied, as the group grows only candidates who are more and more extreme can join. The reason for this is that all group members prefer the same candidate only when both candidates are close to one of the extremes. This is illustrated in Figure 1(a), where each member is positioned on the $[0, 1]$ interval according to his opinion. In the group depicted in the figure, if a candidate is admitted then it has to be the case that both candidates lie in one of the gray colored intervals.

Under the majority admission rule, as the group grows, the distribution of opinions of its members converges to a triangle distribution, with the median located at $1/2$. Convergence to this distribution happens regardless of the starting conditions, that is, the distribution of opinions in the initial group. Unlike in the case of consensus, when the majority rule is applied a new candidate joins the group at every step. To understand who this candidate is, we inspect the mechanics of the majority admission rule in greater detail. Consider two candidates, y_1 and y_2 , and assume that $y_1 < y_2$. Observe that because each member votes for the candidate who is more similar (closer) to himself, all the members to the left of $(y_1 + y_2)/2$ vote for y_1 , and members to the right of $(y_1 + y_2)/2$ vote for y_2 . It is easy to see that if $(y_1 + y_2)/2$ is located to the right of the median, the candidate who receives the majority of the votes is y_1 (i.e., the candidate on the left). In this case the candidate closer to the median is y_1 . Thus, the majority rule essentially prescribes that the candidate closer to the median is accepted into the group. This is illustrated in Figure 1(b), where $(y_1 + y_2)/2$ is located to the right of the median (x_3) and therefore y_1 , voted for by x_1 , x_2 and x_3 , is admitted into the group.

Note that the identity of the admitted candidate in each step is determined entirely by the location of the median. This means that in order to prove that the distribution of opinions in the group converges to the triangle distribution with the median at $1/2$ it is sufficient to show that the median converges to $1/2$. It is quite easy to show that if the median is located at $1/2 - \varepsilon$, the probability of accepting a candidate to the left of the median is less than half, therefore the median should move to the right. The main challenge lies in the fact that the analyzed process is discrete, which makes analyzing the magnitude of the move of the median technically more difficult.

Admission rules with special veto. The dichotomy between the composition of opinions in the group when using consensus as opposed to majority calls for understanding intermediate admission rules, where in order to admit a candidate, some given fraction, greater than half of the group, is required to prefer him. We study this question under a somewhat different model, assuming that the group originates with a founder, located at 1, who has a special veto power. Whenever two candidates apply, only the one closer to the founder (i.e., the right one) is considered for admission, and he is admitted if and only if an r -fraction of the group prefers him over the candidate on the left (otherwise, no candidate is admitted).

Our results here are quite intriguing: we show that this process exhibits a phase transition at $r = 1/2$. In particular, if $r < 1/2$, regardless of the initial conditions, the group converges to a continuous distribution. It is a truncated triangle distribution, characterized by the location of the $(1 - r)$ -quantile. At the same time, for $r > 1/2$ as the group grows, only candidates closer and closer to 0 are accepted into the group. These results resemble those we have presented above for the consensus admission rule, but are even stronger. Despite the excessive power granted to the founder of the group, who is located at 1, the group can entirely change its character and become one that admits only candidates that are close to 0.

Quantile-driven admission rules. Several of the admission rules mentioned above belong to a family we call “quantile-driven.” Under these rules the decision of which candidate to accept, if any, is determined solely by the location of the p -quantile for some value of p . The majority rule, for example, is a quantile-driven rule with $p = 1/2$. We show that for quantile-driven admission rules that have two rather natural properties the p -quantile always converges. This is a convenient tool for showing that in case of the majority rule, the median of the group converges to $1/2$. We also use it as part of the proof that for veto rules with $r < 1/2$ the $(1 - r)$ -quantile converges.

1.2. Fixed-Size Groups

Many groups have a fixed size and do not grow over time, as in the case of condominiums and committees. Often a committee member serves for a term, after which it is possible to extend his membership for an additional term. A natural way of doing so is to place the decision whether or not to extend his membership in the hands of the other committee members. This can be accomplished, for example, by comparing the candidate who finished his term with a potential replacement. The one who receives a p -fraction of the votes (for some given p) is the one who joins (or rejoins) the committee.

The fundamental questions that drive our analysis for fixed-size groups are similar to the ones we have analyzed for growing groups. Specifically, we are interested in understanding how the composition of a fixed-size committee can evolve over time, and how it is influenced by the admission rule.

It turns out that in the fixed-size group model our questions make sense even in an adversarial setting, that is, when both the member who is up for re-election and the potential replacement are chosen adversarially. The two aspects of fixed-size groups we are interested in are: (a) By how much can the opinion of committee members drift as the committee evolves? and (b) Can a committee member have immunity against replacement under some admission rules.

As in the case of growing groups, the answers to both questions depend on the value of p . For the majority rule (i.e., $p = 1/2$), for every initial configuration, the committee can move arbitrarily far from its initial location. In contrast, for any admission rule that requires even a single vote more than the standard majority, the drift becomes bounded and diminishes with p . In the extreme case, i.e., consensus, every admitted candidate is at distance of at most D from either boundary of the original configuration, where D denotes the diameter of the initial configuration.

Regarding the problem of immunity, our process exhibits an interesting phase transition. In particular, there exists a committee that grants immunity to one of its members (i.e., ensuring that this member can never be replaced) if and only if $p > 3/4$. Generally speaking, a committee in which a member has immunity has the following structure: there are two clusters of points located at the two extremes, and a single point (the median) located in the middle. We use the bound on the drift of the committee for any p greater than $1/2$ (as noted above) to show that neither of the two clusters can ever reach the median, so that the median of the committee effectively enjoys immunity.

2. GROWING GROUPS: CONSENSUS AND MAJORITY

2.1. Consensus

The first admission rule we analyze is consensus. Under this rule, a candidate is accepted only if all group members agree he is better than the other candidate. Even though it may initially seem counter intuitive, it is quite easy to see that when the consensus rule is applied as the group grows only members closer and closer to the two extremes will join the group:

PROPOSITION 2.1. *Consider a group $S(k_0)$, then, for any ε with probability 1 there exists k_ε such that for any $k > k_\varepsilon$ only candidates in $[0, \varepsilon]$ and $[1 - \varepsilon, 1]$ can be admitted to the group.*

PROOF. Denote the members of the group $S(k)$ ordered from left to right by $x_1(k), \dots, x_k(k)$. The requirement for all the group members to agree in order to admit a member implies that only members in the intervals $[0, 2x_1(k)]$ and $[2x_k(k) - 1, 1]$ can be admitted. The proof is completed by observing that if $x_1(k) > \varepsilon/2$ then the probability of accepting a candidate in $[0, \varepsilon/2]$ is at least $\varepsilon^2/4$ and hence with probability 1 there ex-

ists a step $k_\varepsilon^l > k_0$ such that $x_1(k_\varepsilon^l) < \varepsilon/2$. A symmetric argument shows the existence of k_ε^r such that $x_{k_\varepsilon^r}(k_\varepsilon^r) > 1 - \varepsilon/2$. Hence, the proposition holds with $k_\varepsilon = \max\{k_\varepsilon^l, k_\varepsilon^r\}$. \square

2.2. Majority

Under the majority rule a candidate that receives at least half of the votes is accepted to the group. As discussed in the introduction, the majority rule can be described as a function of the location of the median.¹ Denote the median of the group $S(k)$ by $m(S(k))$, the majority rule can be defined as follows:

Definition 2.2 (majority). Given two candidates y_1, y_2 , admit to the group $S(k)$ the candidate y_i minimizing $|m(S(k)) - y_i|$.

We show that for the majority decision rule, with high probability, the process converges to a distribution given by the triangle density function with a median located at $1/2$:

$$h(x) = \begin{cases} 4x & \text{for } 0 \leq x \leq 1/2 \\ 4 - 4x & \text{for } 1/2 < x \leq 1 \end{cases}$$

To prove that the distribution of opinions of the group members converges to the triangle distribution described above it is sufficient to show that with high probability the median converges to $1/2$. Indeed, if the median is located at $1/2$ then a candidate located at $x < 1/2$ will be admitted to the group with probability of $2x$. The reason for this is that this candidate is accepted to the group if and only if the other candidate is located at $[0, x]$ or $(1 - x, 1]$. Hence, the density function for $x \leq 1/2$ is $h(x) = 4x$. In a similar way we can compute the value for the density function for $x > 1/2$. Furthermore, if the median is in the interval $[1/2 - \varepsilon, 1/2 + \varepsilon]$, with probability $1 - O(\varepsilon)$ the same candidate will be chosen as in the case the median is exactly $1/2$.² Hence, if the median converges to $1/2$ the opinions distribution of the group converges to the triangle distribution $h(x)$.

We provide some informal intuition for the convergence of the median to $1/2$. Consider a group $S(k)$, such that $m(S(k)) = 1/2 - \varepsilon$ for $\varepsilon > 0$, and where the initial size of the group was $k_0 = 1$. (a symmetric argument holds for the case that $m(S(k)) = 1/2 + \varepsilon$). We observe that by symmetry the probability of accepting a candidate in $[0, m(S(k))]$ is the same as the probability of accepting a candidate in $[m(S(k)), 2m(S(k))]$. Also, the probability of accepting a candidate in $[2m(S(k)), 1]$ is $4\varepsilon^2$. Now, consider adding k more members to the group. By the previous probability computation the number of members added to the right-hand side of $m(S(k))$, in the k steps, exceeds that in its left side by roughly $4\varepsilon^2 k$. This means that the median should move by about $\frac{1}{2}(2\varepsilon)^2 k$ members to the right. Since altogether as the group grew from size 1 to $2k$, $4k$ candidates have attempted to get accepted to the group, we cannot have more than roughly $4k\delta$ points in any interval of length δ . Thus the median will move to the right by at least a distance of about $2\varepsilon^2 k / (4k) = \varepsilon^2/2$. We have thus shown that when we double the number of points the median increases from $1/2 - \varepsilon$ to roughly $1/2 - \varepsilon + \varepsilon^2/2$. In particular, this means that after roughly $1/\varepsilon$ doublings, the median will shift to about $1/2 - \varepsilon/2$.

The intuition above is lacking in two main aspects. First, we assumed that the probability of accepting a candidate in $[0, m(S(k))]$ remains fixed throughout all the k steps. However, this is not exactly true as in these k steps the median does not remain at the

¹In case of a group of an even size any consistent choice of the median will do.

²The formal reason for this is that with probability $1 - O(\varepsilon)$, $|y_1 - y_2| > 2\varepsilon$ and $|y_1 - (1 - y_2)| > 2\varepsilon$

same place. A second, more minor, issue is showing that, roughly speaking, it is always the case that each interval of size δ does not have too many members. Instead of formalizing this intuition we choose to take a more general approach: in the next section we define a family of admission rules that includes the majority rule and show convergence for each one of these rules. This gives us the following theorem for the majority rule:³

THEOREM 2.3. *Consider a group $S(k_0)$. For any $\varepsilon > 0$, with probability $1 - o(1)$, there exists k_ε , such that for any $k' > k_\varepsilon$, $|m(S(k')) - \frac{1}{2}| < \varepsilon$.*

3. GROWING GROUPS: QUANTILE-DRIVEN ADMISSION PROCESSES

In this section we define a quite broad family of admission rules with the common property that the choice of which candidate to accept (if at all) is determined by the location of the p -quantile for some value of $0 < p < 1$. We denote the p -quantile of a group $S(k)$ by $q_p(S(k))$ and formally define it as follows:

Definition 3.1. $q_p(S(k)) \in S(k)$ is a p -quantile of a group $S(k) \in [0, 1]^k$ if $|\{i|x_i \leq q_p(S(k))\}| \geq p \cdot k$ and $|\{i|x_i < q_p(S(k))\}| \leq p \cdot k$.⁴

Using this definition we define a *quantile-driven* admission rule:

Definition 3.2. An admission rule is *quantile-driven* if there exists a parameter $p \in [0, 1]$ such that for every $x \in [0, 1]$ the probability of accepting a member below x to the group $S(k)$ is only a function of x and the p -quantile of $S(k)$.

The majority rule is a quantile-driven rule as it prescribes that the candidate that is admitted to group is the one closer to the median. However, the consensus admission rule is not quantile-driven as the choice of which candidate (if at all) is admitted to the group is determined by both the 0-quantile and the 1-quantile.

We study quantile-driven admission processes which are admission processes in which candidates are admitted according to a quantile-driven admission rule. For these general processes we do not make any assumptions on the distribution that the candidates are drawn from or even on the number of the candidates. Even though, as earlier discussed for the specific processes we analyze in the present paper we assume that there are only 2 candidates that are drawn from the uniform distribution $\mathcal{U}[0, 1]$. We denote by $f_p(q_p)$ ⁵ the probability of accepting a candidate below the current location of the p -quantile q_p . Note, that $f_p(\cdot)$ is based both on the admission rule and the distribution that the candidates are drawn from. For example, for the admission process with the majority rule we have that:

$$f_{1/2}(q) = \begin{cases} 2q - 2q^2 & \text{for } q \leq 1/2 \\ 1 - 2q + 2q^2 & \text{for } q > 1/2 \end{cases}$$

An easy way for computing this function applies the same logic behind the intuition for showing that the median converges to 1/2. For example, if $q < 1/2$, then with probability $(1-2q)^2$ a candidate in the interval $[2q, 1]$ joins the committee. Furthermore, by symmetry the probability of a candidate to join $[0, q]$ is the same as the probability for joining $[q, 2q]$. Thus, we have that for $q < 1/2$, $f(q) = (1 - (1 - 2q)^2)/2 = 2q - 2q^2$.

We now define smooth quantile-driven admission processes and show that the majority process is smooth:

³In Section A.1 of the appendix we provide a short remark on the convergence rate.

⁴For groups in which this definition admits more than a single choice for the p -quantile, any consistent choice will do.

⁵When p is clear from the context we denote this function simply by $f(\cdot)$.

Definition 3.3. An admission process in which at every step a candidate joins the group⁶ is *smooth* if:

- (1) $f_p(\cdot)$ is a strictly increasing continuous function.
- (2) The probability of accepting a member in any interval of length δ is at least $c_1 \cdot \delta^2$ and at most $c_2 \cdot \delta$ for some constants c_1 and c_2 .

CLAIM 3.4. *The majority admission process is smooth.*

PROOF. Observe that both pieces of the function $f_{1/2}(\cdot)$ are continuous and strictly increasing for the appropriate range and hence the function is increasing and continuous (for continuity at $q = 1/2$ observe that for this value of q both pieces of the function attain the same value). Furthermore, observe that since in the majority rule a candidate is accepted at every step the probability of accepting a candidate in an interval of length δ is at least δ^2 . On the other hand, the probability of accepting a candidate in an interval of length δ is upper bounded by the probability that at least one of the candidates lies in the interval δ which is 2δ . \square

We show that the location of the p -quantile of a group that uses a smooth admission rule always converges to unique τ_p such that $f(\tau_p) = p$:⁷

THEOREM 3.5. *Consider a group $S(k_0)$ that uses a smooth admission process $f_p(\cdot)$. Let τ_p be the unique value satisfying $f_p(\tau_p) = p$. For any $\varepsilon > 0$, with probability $1 - o(1)$, there exists k'_ε such that for any $k' > k'_\varepsilon$, $|q_p(S(k')) - \tau_p| < \varepsilon$.*

We first provide some intuition on why smooth admission processes converge. Assume that $q_p(S(k)) < \tau_p$, the assumption that $f(\cdot)$ is strictly increasing implies that $f(q_p(S(k))) < p$ and hence the p -quantile has to move right. The upper bound on the probability of accepting a candidate in an interval of length δ allows us to show that the quantile will indeed keep moving right and will not “get stuck” at some cluster of points. The lower bound on the probability to accept a candidate provides us an assurance that the p -quantile cannot move too far when a small number of members is added.

We now give a taste of the way that the formal proof operates. The proof itself is provided in Appendix B. First, we let $\omega(k) = |\tau_p - q_p(S(k))|$. Then, we define the following two strictly increasing functions:

$$\begin{aligned} -g_r(\omega) &: [0, \tau_p] \rightarrow [0, p], g_r(\omega) = p - f_p(\tau_p - \omega). \\ -g_l(\omega) &: [0, 1 - \tau_p] \rightarrow [0, 1 - p], g_l(\omega) = f_p(\tau_p + \omega) - p. \end{aligned}$$

The crux of the proof is in the following proposition stated here for $q_p(S(k)) < \tau_p$:

PROPOSITION 3.6. *Consider adding t more members to a group $S(k)$, such that $q_p(S(k)) < \tau_p$. For any $\sigma < \tau_p - q_p(S(k))$ such that:*

- (1) $g_r(\omega(k) - \sigma) > g_r(\omega(k))/2 > c_2 \cdot \sigma$.
- (2) *Each of the intervals $[q_p(S(k)) - \sigma, q_p(S(k))]$ and $[q_p(S(k)), q_p(S(k)) + \sigma]$ contain at least t members.*

with probability at least $1 - e^{-\Theta(g_r(\omega(k))^2 \cdot t)}$, the group $S(k+t)$ includes at least $\frac{g_r(\omega(k))}{4} \cdot t$ members in the interval $[q_p(S(k)), q_p(S(k+t))]$. (this also implies that $\omega(S(k+t)) \leq \omega(S(k))$).

⁶For processes that do not exhibit this property we can restrict our attention to steps in which a candidate is accepted and normalize the function $f_p(\cdot)$ accordingly

⁷Such a τ_p always exists since $f(\cdot)$ is strictly increasing, $f(0) = 0$ and $f(1) = 1$.

Informally, for the conditions of the proposition to hold we need the quantile to be in a relatively small and dense interval, such that we are certain that for a long enough time the quantile will not leave this interval. When this is the case, we have that the quantile will move closer to τ_p by at least $\frac{g_r(\omega(k))}{4} \cdot t$ members. To “translate” the number of points into distance we first provide bounds on the density of each interval by using Chernoff bounds:

CLAIM 3.7. *let $\delta(k) = k^{-1/10}$. Let $k \geq \frac{k_0}{\delta(k)}$, where k_0 is the initial size of the group, for any interval I of length $|I| \geq \delta(k)$ the following holds with probability of at least $1 - \frac{2}{\delta(k)} \cdot e^{-\Theta(\delta(k)^2 \cdot k)}$:*

- (1) *The number of members of $S(k)$ in I is at least $c'_1 \cdot |I| \cdot \delta(k) \cdot k$ for $c'_1 < c_1$.*
- (2) *The number of members of $S(k)$ in I is at most $c'_2 \cdot |I| \cdot k$, for $c'_2 > c_2$.*

Finally, we combine these bound with careful repeated applications of Proposition 3.6.

4. GROWING GROUPS: SPECIAL VETO POWER

In this section we assume that the group has a founder with opinion 1. This founder has a special veto power in the sense that if a candidate is admitted to the group it will always be the candidate that the founder prefers (i.e., the right candidate). We term such rules *veto rules*. We study a family of veto rules characterized by a parameter r ($0 < r < 1$). Under each such rule if r -fraction of the group members agree that the right candidate is better than the left one then the right candidate joins the group. Else, no candidate is accepted in this step. Veto rules are also quantile-driven rules. To see why observe that, given two candidates located at y_1 and y_2 ($y_1 < y_2$) all the members to the left of $(y_1 + y_2)/2$ vote for y_1 while all the members to its right vote for y_2 . For veto rules the only candidate who has the potential to be admitted to the group is y_2 and he will be admitted if at least r -fraction of the group will vote for him. Putting this together we get that y_2 is admitted if at least a fraction r of the group is located to the right of $(y_1 + y_2)/2$. In particular this implies that y_2 will be accepted to the group $S(k)$ if $(y_1 + y_2)/2 \leq q_{1-r}(S(k))$. The reason for this is that a fraction greater than r of the group is located to the right of $(y_1 + y_2)/2$ and votes for y_2 . Hence, the family of veto rules can be described as follows:

Definition 4.1 (veto rules). Consider two candidates $y_1 < y_2$. y_2 will be admitted to the group $S(k)$ if and only if $(y_1 + y_2)/2 < q_{1-r}(S(k))$.

Recall that under veto rules, there are many steps in which none of the candidates joins the group. Since we want to track the changes in the group, we will only reason about the steps of the process in which a candidate is admitted. Hence, to compute the probability that the next candidate that is accepted to the group lies in some interval we will have to first compute the probability that any candidate is accepted when the $(1-r)$ -quantile is at $q_{(1-r)}(S(k))$. For simplicity throughout this section we denote $1-r$ by p :

CLAIM 4.2. *If $q_p \leq 1/2$, then the probability of accepting any candidate is $2q_p^2$. If $q_p > 1/2$, then the probability of accepting any candidate is $1 - 2(1 - q_p)^2$.*

PROOF. An easy method for computing the probability of accepting a candidate is using the geometric representation depicted in Figure 2. The diagonal line is $y_1 = 2q_p - y_2$ and the surface below it is the area such that the average of the two candidates y_1 and y_2 is below q_p . Thus, it includes all pairs of candidates (y_1, y_2) for which one of the candidates will be admitted to the group. For $q_p < 1/2$ this surface is a triangle with an area of $2q_p^2$. For $q_p > 1/2$ it is easier to compute the surface of the upper white

triangle and subtract this area from the unit square. Thus we have that the area of the pentagon that includes all pairs of candidates (y_1, y_2) such that one of the candidates is admitted to the group is $1 - 2(1 - q_p)^2$. \square

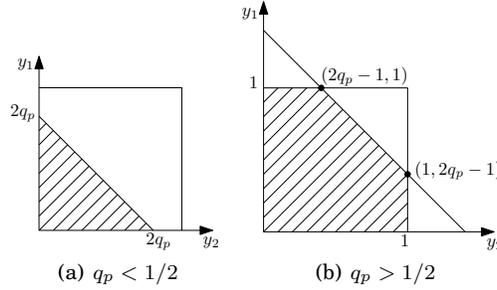


Fig. 2. The probability of accepting a candidate under veto rules: in both pictures the striped area includes all pairs of candidates (y_1, y_2) for which one of the candidates will be admitted to the group.

In the two subsections below we analyze the convergence of the $(1 - r)$ -quantile for different values of r . We establish the following phase transition: when $r > 1/2$ the $(1 - r)$ -quantile converges to 0 and when $r < 1/2$ the $(1 - r)$ -quantile converges to a specific value $1/2 < \tau_{1-r} < 1$ to be later determined. In both cases the distribution of opinions as the group grows is fully determined by the location of the $(1 - r)$ -quantile. Hence, when $r > 1/2$ we will see that as the group grows only candidates closer and closer to 0 will be accepted. For $r < 1/2$ the opinion distribution in the group will converge to a truncated triangle density distribution with a maximum located at τ_{1-r} as depicted in Figure 3.

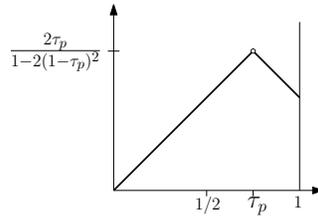


Fig. 3. A sketch of the density function that the group converges to for $p > 1/2 (r < 1/2)$.

4.1. $r > 1/2$: Convergence to 0

We show that for $p < 1/2$ (hence $r > 1/2$) as the group grows with high probability $q_p(S(k))$ is converging to 0. This implies that as the group grows only candidates closer and closer to 0 will be admitted. While the proof itself is somewhat technical the intuition behind it is rather simple: For any group $S(k)$ such that $q_p(S(k)) < 1/2$ the probability that the next accepted candidate lies in the interval $[0, q_p(S(k))]$ is exactly $1/2$. Recall that the right candidate is accepted if and only if $(y_1 + y_2)/2 < q_p(S(k))$. Thus a candidate located in the interval $[0, q_p(S(k))]$ will be chosen with probability $q_p(S(k))^2$. Also, note that in this case by Claim 4.2 the probability of accepting any candidate at all is $2q_p(S(k))^2$. Thus we have that the probability that the next accepted candidate lies in $[0, q_p(S(k))]$ is $1/2$. Recall that $q_p(S(k))$ is the location of the p -quantile

for $p < 1/2$. Roughly speaking the fact that the probability of accepting members to the left of $q_p(S(k))$ is greater than p implies that the p -quantile has to move left (towards 0). A similar argument for the case that $q_p(S(k)) > 1/2$ shows that in this case the probability to accept a candidate in $[0, q_p(S(k))]$ is greater than $1/2$ and hence the p -quantile should move left. Note that this is an example for an admission process which is not smooth ($f(\cdot)$ is not strictly increasing) but still converges.

The formal proof that the p -quantile indeed moves to the left gets more involved by the discrete nature of the process. This requires us to carefully track the changes in the location of the p -quantile to show that indeed as the group grows the p -quantile is moving to the left. As part of the proof, we actually prove a slightly stronger theorem which is that the $(p + \eta)$ -quantile (for $\eta = \frac{1-2p}{8}$) is converging to 0. The statement of the formal theorem we prove is the following:

THEOREM 4.3. *Consider a group $S(k_0)$. For any $\varepsilon > 0$, with probability $1 - o(1)$, there exists k_ε , such that for any $k' > k_\varepsilon$, $q_p(S(k')) < \varepsilon$.*

4.2. $r < 1/2$: Convergence to a Continuous Distribution

We show that for $p > 1/2$ ($r < 1/2$) as the group grows the p -quantile of the group is converging to the point $\tau_p = \frac{2p + \sqrt{2p^2 - p}}{1 + 2p} > 1/2$. If the p -quantile is at $q > 1/2$ the probability of a candidate $x < q$ to be the next accepted candidate is $\frac{x}{1 - 2(1 - q)^2}$. As with probability x a candidate below it appears and by Claim 4.2 for $q > 1/2$ the probability of any candidate to be accepted is $1 - 2(1 - q)^2$. Similarly we can compute the acceptance probability of a candidate $x > q$. By multiplying the probabilities by 2 we get the following density function (sketched in Figure 3 for $q = \tau_p$):

$$h(x, q) = \begin{cases} \frac{2x}{1 - 2(1 - q)^2} & \text{for } 0 \leq x \leq q \\ \frac{4q - 2x}{1 - 2(1 - q)^2} & \text{for } q < x \leq 1. \end{cases}$$

We observe that as q converges to τ_p the distribution of opinions in the group is converging to $h(x, \tau_p)$. This is because for values q close to τ_p the value of the function $h(x, q)$ is close to that of $h(x, \tau_p)$.

The proof that the p -quantile converges to τ_p acquires an additional level of complexity by the fact that the probability of the next accepted candidate to be in the interval $[0, q_p(S(k))]$ has a different expression for $q_p(S(k)) < 1/2$ and for $q_p(S(k)) > 1/2$. That is, in both cases with probability $q_p(S(k))^2$ both of the candidates will be in the interval $[0, q_p(S(k))]$ and hence a candidate in this interval will be accepted. However, since we condition on the event that a candidate is actually accepted we have to divide $q_p(S(k))^2$ by the probability that a candidate is accepted. Claim 4.2 shows that the probabilities of accepting a candidate for $q_p(S(k)) < 1/2$ and for $q_p(S(k)) > 1/2$ are different. In particular we have that for $q_p(S(k)) < 1/2$ the probability of the next accepted candidate to be in $[0, q_p(S(k))]$ is $1/2$ and for $q_p(S(k)) > 1/2$ this probability is $f(q_p(S(k))) = \frac{q_p(S(k))^2}{1 - 2(1 - q_p(S(k)))^2}$.

Luckily, the admission process for $q > 1/2$ (restricted to steps in which a candidate is admitted) is smooth and hence by Theorem 2.3 the p -quantile converges to τ_p . This implies that to show convergence it suffices to show that with high probability there exists some step $k_{1/2}$ such that from this step onwards the p -quantile remains above $q > 1/2$. This is done similarly to the proof showing that for $p > 1/2$ the p -quantile converges to 0. Here, when $q_p(S(k)) < 1/2$ the probability of the next accepted candidate to be below $q_p(S(k))$ is $1/2$. Since $q_p(S(k))$ denotes the location of the p -quantile

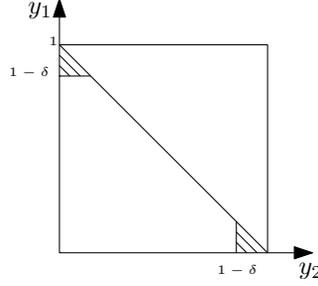


Fig. 4. The striped areas are the areas that if the point defined by the two candidates (y_1, y_2) is in one of them then a candidate in the interval $[1 - \delta/2, 1]$ will join the group.

for $p > 1/2$, $q_p(S(k))$ has to move right, at least until it passes $1/2$. Formally, we prove the following theorem:

THEOREM 4.4. *Consider a group $S(k_0)$. For any $\varepsilon > 0$, with probability $1 - o(1)$, there exists $k_\varepsilon > k_0$, such that for any $k' > k_\varepsilon$, $|q_p(S(k')) - \tau_p| < \varepsilon$.*

PROOF. We begin by observing that for $q > 1/2$ the admission rule is smooth. First, recall that for $q \geq 1/2$ we have that $f(q) = \frac{q^2}{1-2(1-q)^2}$. Note that due to the normalization it is indeed the case that each step a candidate is accepted to the group. Also, it is easy to verify that this is function is increasing and continuous for $q \in (1/2, 1]$. Next, observe that for $q \in (1/2, 1]$ the probability of accepting a candidate in every interval δ is at most $\frac{2\delta}{1-2(1-q)^2} \leq 4\delta$ and at least δ^2 . For the lower bound observe that the interval with the minimal acceptance probability is the last interval $[1 - \delta, 1]$. Now, by the geometric representation depicted in Figure 4 a candidate in $[1 - \delta, 1]$ will be accepted if the point defined by the two candidates is in one of the striped triangles in the figure. Hence the total probability of accepting a candidate in $[1 - \delta, 1]$ is at least δ^2 .

Now, by Theorem 2.3 we have that the p -quantile of a smooth admission rule converge to $\tau_p = \frac{2p + \sqrt{2p^2 - p}}{1 + 2p}$ with probability $1 - o(1)$. In Theorem 4.5 below we show that with high probability there exists some value of k starting which the admission rule is always smooth. Thus, we have that with high probability the p -quantile of the group converges to τ_p . \square

THEOREM 4.5. *Consider a group $S(k_0)$. There exists a time step $k_{1/2}$ such that with probability $1 - o(1)$ for any $k' > k_{1/2}$ the admission rule is smooth.*

PROOF. Pick $\eta = \frac{p-1/2}{4}$. We will show that with high probability there exists $k_{1/2} > k_0$ such that $q_{p-\eta}(S(k_{1/2})) > 1/2$ and for any $k' > k_{1/2}$, $q_p(S(k_{1/2})) > 1/2$. This implies that the admission rule is smooth for any group of size greater than $k_{1/2}$.

The proof follows a very similar structure to the proof of Theorem 4.3. We begin by showing that as we increase the group by ηk members at most $(p - 2\eta) \cdot \eta k$ members will join the interval $[0, q_{p-\eta}(S(k_i)) + \eta \cdot q_{p-\eta}(S(k_i))^2]$.

CLAIM 4.6. *For a group $S(k)$ such that $q_{p-\eta}(S(k)) < 1/2$, consider adding ηk members to the group $S(k)$, with probability $(1 - e^{-\Theta(\eta^3 k)})$. The number of members that joined the interval $[0, q_{p-\eta}(S(k)) + \eta \cdot q_{p-\eta}(S(k))^2]$ in the ηk steps at most $(p - 2\eta) \cdot \eta k$.*

PROOF. Note that by definition we have that for every step k' of the ηk steps, $q_p(S(k')) \geq q_{p-\eta}(S(k))$. We claim that this implies that the probability that the next

accepted candidate is below $q_{p-\eta}(S(k)) + \eta \cdot q_{p-\eta}(S(k))^2$ is at most $p - 3\eta$. To see why this is the case we first observe that with probability $q_{p-\eta}(S(k))^2$ both candidates are below $q_{p-\eta}(S(k))$ and hence a candidate below $q_{p-\eta}(S(k))$ is accepted. Also note that with probability at most $2\eta \cdot q_{p-\eta}(S(k))^2$ a member in the interval $[q_{p-\eta}(S(k)), q_{p-\eta}(S(k)) + \eta \cdot q_{p-\eta}(S(k))^2]$ joins the group, as this is an upper bound on the probability that a candidate in this interval shows up. Finally, as we only take into account steps in which a candidate was chosen we should divide the probabilities above by the probability of accepting a member. It is easy to see that the probability of accepting a member is minimized when $q_p(S(k')) = q_{p-\eta}(S(k))$ and hence the probability that the next accepted candidate is in the interval $[0, q_{p-\eta}(S(k)) + \eta \cdot q_{p-\eta}(S(k))^2]$ is at most

$$\frac{q_{p-\eta}(S(k))^2 + 2\eta \cdot q_{p-\eta}(S(k))^2}{2q_{p-\eta}(S(k))^2} = \frac{1}{2} + \eta = p - 3\eta.$$

By taking a Chernoff bound we get that with high probability the number of members accepted in the ηk steps in the interval $[0, q_{p-\eta}(S(k)) + \eta \cdot q_{p-\eta}(S(k))^2]$ is at most $(p - 2\eta) \cdot \eta k$. Denote by X the number of candidates accepted in the interval $[0, q_{p-\eta}(S(k)) + \eta \cdot q_{p-\eta}(S(k))^2]$, then

$$\Pr[X \geq (1 + \eta) \cdot (p - 3\eta) \cdot \eta k] \leq e^{-\frac{\eta^2(p-3\eta) \cdot \eta k}{3}} \leq e^{-\Theta(\eta^3 k)}$$

□

Let $k_1 \geq \frac{1}{\eta^6}$ (the larger k_1 is, the higher the probability the theorem holds is) and for any $i > 1$ let $k_{i+1} = (1 + \eta)k_i$. Also let $j = \lceil \log_{1+\eta} \frac{1}{\eta} \rceil$. We show that with high probability: $q_{p-\eta}(S(k_{i+j})) > q_{p-\eta}(S(k_i)) + \eta \cdot q_{p-\eta}(S(k_i))^2$.

CLAIM 4.7. *For $i > 1$, with probability $(1 - \sum_{l=1}^{j-1} e^{-\Theta(\eta^3(1+\eta)^l k_i)})$, either there exists $k_i < k' < k_{i+j}$ such that $q_{p-\eta}(S(k')) > 1/2$ or $q_{p-\eta}(S(k_{i+j})) > q_{p-\eta}(S(k_i)) + \eta \cdot q_{p-\eta}(S(k_i))^2$.*

PROOF. To prove the claim we apply Claim 4.6 j times. First we observe that with high probability for every i , $q_{p-\eta}(S(k_{i+1})) \geq q_{p-\eta}(S(k_i))$. The reason for this is that by Claim 4.6 we have that the number of members that joined in the ηk_i steps between k_i and k_{i+1} in the interval $[0, q_{p-\eta}(S(k_i)) + \eta \cdot q_{p-\eta}(S(k_i))^2]$ is at most $(p - 2\eta) \cdot \eta k$. Since the interval $[0, q_{p-\eta}(S(k_i))]$ is included in this interval we have that at most $(p - 2\eta) \cdot \eta k$ members were admitted to it. Hence, $q_{p-\eta}(S(k_{i+1})) \geq q_{p-\eta}(S(k_i))$.

Next, we consider all the members in the $k_{i+j} - k_i$ steps. The fact that $q_{p-\eta}(S(k_{i+1})) \geq q_{p-\eta}(S(k_i))$ implies that in the $k_{i+j} - k_i$ steps the number of candidates admitted to the interval $[0, q_{p-\eta}(S(k_i)) + \eta \cdot q_{p-\eta}(S(k_i))^2]$ is at most $(p - 2\eta)(k_{i+j} - k_i)$. We observe that in the worst case for the group $S(k_i)$ the interval $[0, q_{p-\eta}(S(k_i)) + \eta \cdot q_{p-\eta}(S(k_i))^2]$ contained k_i points. Thus, to prove the claim, we should show that $(p - 2\eta)(k_{i+j} - k_i) + k_i < (p - \eta)k_{i+j}$. Observe that

$$\begin{aligned} (p - 2\eta) \cdot (k_{i+j} - k_i) + k_i &= (p - \eta)k_{i+j} - \eta k_{i+j} + (1 - p + 2\eta)k_i \\ &< (p - \eta)k_{i+j} - k_i + (1 - p + 2\eta)k_i \\ &< (p - \eta)k_{i+j}. \end{aligned}$$

For the second transition we used the fact that $k_{i+j} = (1 + \eta)^{\lceil \log_{1+\eta} \frac{1}{\eta} \rceil} k_i > \frac{1}{\eta} k_i$.

Thus by taking a union bound over the bad events we get that the claim holds with probability at least

$$1 - \sum_{l=i}^{i+j-1} e^{-\Theta(\eta^3 k_l)}.$$

□

Lastly, we show that once we reached a step k_i such that the $(p - \eta)$ -quantile is above $1/2$, then with high probability the p -quantile of $S(k_{i+1})$ will also be above $1/2$.

CLAIM 4.8. *If $q_{p-\eta}(S(k_i)) > 1/2$, then with probability $(1 - e^{-\Theta(\eta^3 k_i)})$, $q_{p-\eta}(S(k_{i+1})) > 1/2$*

PROOF. Observe that since $q_p(S(k')) > 1/2$ for all k' of the ηk_i steps, the expected number of members accepted below $1/2$ in these ηk_i steps is at most $\frac{\eta k_i}{2}$. By taking a Chernoff bound, we have that with probability $1 - e^{-\frac{\eta^3 k_i}{6}}$ the number of candidates accepted in $[0, 1/2]$ is at most $(1/2 + \eta)\eta k_i < (p - \eta)k + i$:

$$\Pr[X \geq (1 + \eta) \cdot \frac{1}{2} \eta k_i] \leq e^{-\frac{\eta^2 \frac{1}{2} \eta k_i}{3}} = e^{-\frac{\eta^3 k_i}{6}}$$

□

The proof of Theorem is completed by observing that we can apply Claim 4.7 till we reach $k_{1/2}$ such that $q_{p-\eta}(k_{1/2}) > 1/2$. Once we reached $k_{1/2}$ we repeatedly apply Claim 4.8 to get that the $(p - \eta)$ -quantile stays above $1/2$ with high probability. By taking a (loose) union bound over the bad events we have that the probability of this is at least $1 - \sum_{i=1}^{\infty} e^{-\Theta(\eta^3 k_i)}$.

5. FIXED-SIZE GROUPS

We now turn our attention to groups of fixed size. As committees are a very good example for such groups throughout this section we will refer to the group as a committee. A committee x consisting of n members is represented by the location of its members' opinions on the real line: (x_1, \dots, x_n) with the convention that $x_i \leq x_{i+1}$ for every i . We consider an iterative process, where in each iteration one of the current committee members x_i can be replaced by a new candidate y . The member x_i is replaced by y if and only if at least $\lceil (n-1)/2 \rceil + \ell$ members weakly prefer y over x_i . This means that x_i is replaced if for at least $\lceil (n-1)/2 \rceil + \ell$ members x_j such that $j \neq i$, $|x_j - y| \leq |x_j - x_i|$. The case $\ell = 0$ corresponds to standard majority, and $\ell = \lfloor (n-1)/2 \rfloor$ corresponds to consensus.

We study two aspects of the evolution of fixed-size committees, namely the magnitude of drift of the committee (i.e., how far the committee can move from its initial configuration), and whether there exist committee members who are guaranteed immunity against replacement. We are able to answer both questions in the more demanding worst case framework. That is, we assume that both the members that might be replaced and the contender are chosen adversarially.

5.1. Magnitude of drift

It is easy to see that for usual majority ($\ell = 0$) the committee can move arbitrarily far when its initial configuration is an arithmetic progression $x_i = i$ (we simply keep replacing x_1 by $x_n + 1$.) More generally, in the next theorem we show that under the majority rule, any committee with distinct members can be transformed into an arithmetic progression and hence the drift from the initial configuration is unbounded.

THEOREM 5.1. *For every initial configuration in which all x_i 's are distinct, the committee can move arbitrarily far under the majority voting rule.*

PROOF. For the proof consider, for simplicity, the case of odd n (the case of even n is similar). Let the initial configuration be $x_1 < x_2 < \dots < x_{2k+1}$. Let $M = x_{k+1}$ be the median and let ϵ be a small positive real satisfying, say, $\epsilon k < M - x_k = x_{k+1} - x_k$ and $\epsilon k \leq x_{k+2} - M = x_{k+2} - x_{k+1}$. Now in step i ($1 \leq i \leq k$), replace x_i by $M - \epsilon i$, and in step $k+i$ ($1 \leq i \leq k$) replace x_{k+1+i} by $M + \epsilon i$. It is easy to verify that these replacements are legal (in fact, in each of them we have at least $k+1$ points that prefer the newcomer). Now we have an arithmetic progression and it can move arbitrarily far, by the observation above (in these steps we have only k points that prefer the new one). \square

In view of the above, it is interesting to show that even if ℓ is 1, the committee cannot move too far away. The following proposition establishes an upper bound on the distance the committee can move, as a function of ℓ and the diameter of the initial configuration $D = x_n - x_1$:

As before, we assume for simplicity that n is odd. The case of even n is similar (see remark C.1 in Appendix C).

THEOREM 5.2. *If $n = 2k + 1$, $1 \leq \ell \leq k$, and the initial configuration has diameter $D = x_n - x_1$, for any future configuration x' , it holds that $x'_{k-\ell+2} \leq x_n + \frac{Dk}{2\ell-1}$ and $x'_{k+\ell} \geq x_1 - \frac{Dk}{2\ell-1}$. The term $\frac{Dk}{2\ell-1}$ is tight up to a constant factor.*

The following lemma (which we prove in Appendix C facilitates the proof of Theorem 5.2.

LEMMA 5.3. *Let the configuration before a step be $x = (x_1, x_2, \dots, x_{2k+1})$, and the configuration after a step in which y has been added and x_i been dropped be $x' = (x'_1, x'_2, \dots, x'_{2k+1})$. If the median moved to the right, then the sum of distances from the median has decreased by at least $2 \sum_{j=k-\ell+2}^k d(x_j, x'_j) + d(x_{k+1}, x'_{k+1})$.*

We now prove Theorem 5.2.

PROOF. We show that $x'_{k-\ell+2} \leq x_n + \frac{Dk}{2\ell-1}$. By symmetry the same argument implies that $x'_{k+\ell} \geq x_1 - \frac{Dk}{2\ell-1}$. Consider any configuration x' during the process. We show that if $x'_{k-\ell+2} \geq x_n + Dt$, then $t \leq \frac{k}{2\ell-1}$. If $x'_{k-\ell+2} \geq x_n + Dt$, then for every $j \in \{k-\ell+2, \dots, k+1\}$, the point x_j has moved at least Dt to the right. It is easy to see that the sum of distances from the median is always bounded by $D\lfloor n/2 \rfloor$. Therefore in the original configuration the sum of distances is at most kD . By Lemma 5.3, it must hold that $2 \sum_{j=k-\ell+2}^k d(x_j, x'_j) + d(x_{k+1}, x'_{k+1}) \leq kD$.⁸ Now, substitute $d(x_j, x'_j) \geq Dt$ for every $j \in \{k-\ell+2, \dots, k+1\}$ to get $2(\ell-1)Dt + Dt \leq kD$, or equivalently $t \leq \frac{k}{2\ell-1}$, as desired.

To see that this is asymptotically tight, consider a profile x_1, \dots, x_{2k+1} with $x_{i+1} - x_i = (1 - \delta)^{i-1}$, where δ is chosen to make point $x_{k-\ell+2}$ equally distanced from points x_1 and x_{2k+2} , where x_{2k+2} is defined by the same geometric progression (i.e., $x_{2k+2} = x_{2k+1} + (1 - \delta)^{2k}$). Here, it can be shown that $\delta = \Theta(k/\ell^2)$. The process continues iteratively by always considering the next point in the geometric progression versus the current smallest point in the profile. The new candidate continues to be chosen over the smallest point at all iterations. The process converges to a point at distance

⁸Note that Lemma 5.3 applies to a single change, and we are discussing a sequence of changes. However, the sum of distances of the j -th point is lower bounded by the distance from its initial to final location.

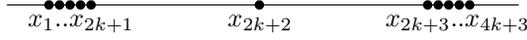


Fig. 5. An example of the configuration that the median has immunity if at least $3k + 3$ votes are required to remove a member.

$\sum_{i \geq 0} (1 - \delta)^i = 1/\delta$ from x_1 . The distance that point $x_{k-\ell+2}$ moved is roughly $1/\delta = \Theta(k^2/\ell)$, whereas the diameter of the initial configurations is at most $2k$. Thus, the distance that $x_{k-\ell+2}$ moved is $\Theta(Dk/\ell)$, as claimed. \square

For the case of consensus (i.e., $\ell = k$), we establish a stronger bound on the shift of the committee. In the Appendix C we prove that:

PROPOSITION 5.4. *For the case of consensus (i.e., $\ell = k$), if $n \geq 3$ and the initial configuration has diameter $D = x_n - x_1$, then every new element that will be added to the committee during the process is at least $x_1 - D$ and at most $x_n + D$.*

5.2. Immunity

Given n , the majority needed to replace an existing member, and an initial configuration, we say that a committee member has *immunity* if it can never be replaced by the process above. We show that a phase transition occurs at a majority of $\frac{3}{4}n$. For simplicity of presentation we assume that $n = 4k + 3$.

THEOREM 5.5. *Let $n = 4k + 3$. There exists an initial configuration in which a member has immunity if and only if a majority of at least $3k + 3$ is required.*

In Figure 5 we give an example of such a configuration in which the median has immunity. It consists of two clusters, each of size $2k + 1$, and an additional point which is the median. Each cluster is located at a different side of the median and sufficiently far from it. We now sketch the proof showing that the median of this committee has immunity. Observe that in order to remove a member in the left cluster at least $k + 1$ members of the left cluster have to prefer the contender over the existing member. This means that informally we can consider the left cluster as an independent committee requiring a majority of at least $\lceil (n - 1)/2 \rceil + 1$. Thus, as long as the left cluster is sufficiently far from the median, we can apply Theorem 5.2 to show that the drift of the left cluster is bounded. As the same argument holds for the right cluster we have that the median will stay a median. The previous argument relied on the fact that the majority required to remove a candidate is large enough such that the number of votes required separately from each cluster is greater than half its size. We now provide a formal proof of this argument:

PROPOSITION 5.6. *If $n = 4k + 3$ and a majority of at least $3k + 3$ is required, then there exists a configuration in which the median has immunity.*

PROOF. Suppose a majority of $3k + 2 + \ell$ is required, where $1 \leq \ell \leq k$. Consider the configuration where $x_{2k+1} - x_1 \leq d$, $x_{4k+3} - x_{2k+3} \leq D$ and the median $M = x_{2k+2}$ is of distance greater than $\frac{Dk}{2\ell-1}$ from x_{2k+1} and from x_{2k+3} . We claim that the median has immunity.

Consider the $2k + 1$ members to the left of M . Any element x_i of them can be replaced by a new candidate y only if at least $k + \ell$ elements out of the $2k + 1$ elements prefer y to x_i (otherwise, y has a majority of at most $3k + \ell + 1$, which is not sufficient). By applying Theorem 5.2 to this set of $2k + 1$ members, we get that in any future configuration $x'_{k-\ell+2} \leq x_{2k+1} + \frac{kd}{2\ell-1}$. But since the median is of distance greater than $\frac{Dk}{2\ell-1}$ from x_{2k+1} , there are at least $k - \ell + 2$ elements to the left of the median throughout the whole process. Analogously, it can be shown that there are always at least $k - \ell + 2$

members to the right of the median (by applying the assertion that $x'_{k+\ell} \geq x_1 - \frac{Dk}{2\ell-1}$ from Theorem 5.2 to the $2k+1$ members to the right of the median). Now observe that as long as there are at least $k-\ell+2$ elements in each side of the median, it cannot be replaced. Indeed, for every new candidate y , there are at most $3k+\ell$ members who prefer y to the median, while the required number is at least $3k+\ell+2$. \square

It is interesting to note that, one can show that the median can guarantee an even stronger property than immunity, namely to always remain the median. This can be done by slightly modifying the previous example, so that the distance between the median and each of the two points x_{2k+1} and x_{2k+3} is greater than, say, kd . Since no element from the left set can ever be above $x_{2k+1} + kd$ and no element from the right set can ever be below $x_{2k+3} - kd$ (see remark following the proof of Theorem 5.2), the original median remains the median forever.

In Appendix C we show that the other direction also holds, that is:

PROPOSITION 5.7. *If $n = 4k+3$ and a majority of at most $3k+2$ is required to replace an existing member, then for any initial configuration no element has immunity.*

6. CONCLUDING REMARKS

In this paper we initiate the study of evolving social groups, and the effects of different admission rules on their long-run compositions. In our models, each group member is represented by a point in $[0, 1]$ representing his opinion. Every group member prefers candidates located closer to him to candidates that are further away because of homophily. We consider stochastic models where in each step two random candidates appear and voted for by the current group members. In the case of a fixed-size group, our analysis holds even in an adversarial model.

The framework we present extends itself to several exciting directions. First there are more families of admission rules that are worth studying. One such family is the p -majority which we only studied for fixed-size groups. Recall that for growing groups we have analyzed a variant of it that gave special veto power to the founder located at 1. We suspect that for growing groups the family of p -majority admission rules also exhibits a phase transition: for $p > 3/4$ as the group grows only candidates close to the extremes will join it; for $p < 3/4$ the distribution of opinions in the group will converge to some continuous distribution. Additional interesting extensions include considering candidates that arrive according to a not necessarily uniform distribution on $[0, 1]$, and analyzing a process in which at every step more than 2 candidates apply. While these extensions lead to interesting questions we believe that the models we have considered in this paper already shed light on real life processes involving the dynamics of evolving groups; the present paper provides a framework and tools for further exploration of this direction.

REFERENCES

- Robert Axelrod. 1997. The Dissemination of Culture. *Journal of Conflict Resolution* 41, 2 (April 1997), 203–226.
- Eli Ben-Naim, Paul Krapivsky, and Sidney Redner. 2003. Bifurcations and Patterns in Compromise Processes. *Physica D* 183 (2003).
- Guillaume Deffuant, David Neau, Frederic Amblard, and Gérard Weisbuch. 2000. Mixing beliefs among interacting agents. *Advances in Complex Systems* 3 (2000), 87–98.
- Morris H. DeGroot. 1974. Reaching a Consensus. *J. Amer. Statist. Assoc.* 69 (1974), 118–121.

- Andreas Flache and Michael W Macy. 2011. Local Convergence and Global Diversity From Interpersonal to Social Influence. *Journal of Conflict Resolution* 55, 6 (2011), 970–995.
- Noah E. Friedkin and Eugene C. Johnsen. 1990. Social Influence and Opinions. *Journal of Mathematical Sociology* 15, 3-4 (1990), 193–205.
- Rainer Hegselmann and Ulrich Krause. 2002. Opinion dynamics and bounded confidence: Models, analysis and simulation. *Journal of Artificial Societies and Social Simulation* 5, 3 (2002).
- David Kempe, Jon Kleinberg, Sigal Oren, and Aleksandrs Slivkins. 2013. Selection and Influence in Cultural Dynamics. In *Proc. 14th ACM Conference on Electronic Commerce*.
- Miller McPherson, Lynn Smith-Lovin, and James M. Cook. 2001. Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology* 27 (2001), 415–444.
- John E. Roemer. 2001. *Political Competition: Theory and Applications*. Harvard University Press.
- Nicholas C Wormald. 1995. Differential equations for random processes and random graphs. *The annals of applied probability* (1995), 1217–1235.

A. PROOFS FROM SECTION 2

A.1. A remark about the convergence rate of the majority and consensus rules

It is interesting to note that the convergence of the majority process is very slow. Indeed, suppose that when there are t points selected already, the median is $1/2 - g(t)$. Then, by the reasoning above, in the next step the probability that the chosen point is to the right of the median exceeds the probability it is on its left by $(2g(t))^2$. This means that the median, on the average, steps by $\frac{1}{2}(2g(t))^2$ units to the right in each step. For small values of $g(t)$ the density of points in the relevant range is about 2δ points in an interval of length δ . This means that on the average the median increases by about $2g(t)^2/(2t) = g(t)^2/t$ in a step. We thus get that $g(t) - g(t+1)$ is essentially $g(t)^2/t$ implying that $g = g(t)$ satisfies the following differential equation: $g' = -g^2/t$. Solving we get $1/g = \ln t + c$ or equivalently $t = Ce^g$. C can be solved from the initial conditions. Thus, for example, if we start with $t_0 = 500$ (which is close to $10e^4$) and the median for that t is $1/4 = 1/2 - 1/4$, we get that it will take close to $t = 10e^{1/\varepsilon}$ steps to get to a median $1/2 - \varepsilon$. More information about how to prove that discrete random processes converge with high probability to the solution of a differential equation can be found in [Wormald 1995].

In contrast, in the consensus model convergence is fast: it is easy to see that for any initial configuration, after t steps, with high probability every newly elected member lies in $[0, O(1/\sqrt{t})] \cup [1 - O(1/\sqrt{t}), 1]$.

B. PROOFS FROM SECTION 3

In the following section we prove Theorem 2.3. Recall that the theorem we wish to prove is the following:

Consider a group $S(k_0)$ that uses a smooth admission process $f_p(\cdot)$. Let τ_p be the unique value satisfying $f_p(\tau_p) = p$. For any $\varepsilon > 0$, with probability $1 - o(1)$, there exists k'_ε , such that for any $k' > k'_\varepsilon$, $|q_p(S(k')) - \tau_p| < \varepsilon$.

The proof is quite complicated hence we first provide a brief outline of the proof. We repeat propositions and claim that were already presented in the main body to make the proof easier to follow.

The proof has two main building blocks that are used iteratively to show that the p -quantile converges to τ_p . First, in Section B.1, we show that if the p -quantile is in a relatively small and dense interval then it will move closer to τ_p by a certain number of points which is a function of the density of the interval it is in. In the second building block, in Section B.2 we use Chernoff bounds to show that on one hand the intervals are dense enough so that the p -quantile will remain long enough in the same interval. But, on the other hand, they are not too dense to prevent from the p -quantile to move a non-negligible distance towards τ_p . In the rest of the proof, Section B.3, we carefully use these two building blocks on groups of growing size to show that indeed the p -quantile converges to τ_p .

B.1. If the p -quantile is confined to a small interval then it moves closer to τ_p

We now formally show that if the p -quantile is in a relatively small and dense interval then it will move closer to τ_p by a certain number of points. We first state and prove the proposition for the case that $q_p(S(k)) < \tau_p$ and then provide the statement for the symmetric case.

PROPOSITION B.1. *Consider adding t more members to a group $S(k)$, such that $q_p(S(k)) < \tau_p$. For any $\sigma < \omega(k)$ such that:*

$$(1) \quad g_r(\omega(k) - \sigma) > g_r(\omega(k))/2 > c_2 \cdot \sigma.$$

(2) Each of the intervals $[q_p(S(k)) - \sigma, q_p(S(k))]$ and $[q_p(S(k)), q_p(S(k)) + \sigma]$ contain at least t members.

The following hold with probability at least $1 - e^{-\Theta(g_r(\omega(k))^2 \cdot t)}$.

(1) $\omega(S(k+t)) \leq \omega(S(k))$.

(2) The group $S(k+t)$ contains at least $\frac{g_r(\omega(k))}{4} \cdot t$ members in the interval $[q_p(S(k)), q_p(S(k+t))]$. (i.e., the p -quantile moved by at least $\frac{g_r(\omega(k))}{4} \cdot t$ members as the group size was increased by t .)

PROOF. Note that since each of the intervals $[q_p(S(k)) - \sigma, q_p(S(k))]$ and $[q_p(S(k)), q_p(S(k)) + \sigma]$ include at least t points, then for each step k' in the next t steps we have that $q_p(S(k')) \in [q_p(S(k)) - \sigma, q_p(S(k)) + \sigma]$. Using this we compute an upper-bound on the probability of accepting a candidate in $[0, q_p(S(k))]$:

(1) For every step k' such that $q_p(S(k')) < q_p(S(k))$ we have that the probability of accepting a candidate in $[0, q_p(S(k))]$ is at most the probability of accepting a candidate in $[0, q_p(S(k'))]$, which is $f(q_p(S(k')))$, plus the probability of accepting a candidate in $[q_p(S(k')), q_p(S(k))]$ which is at most $c_2 \cdot (q_p(S(k)) - q_p(S(k'))) < c_2 \cdot \sigma$. Since $f(\cdot)$ is an increasing function, $q_p(S(k')) < q_p(S(k))$, and by using our assumptions on σ , we have that:

$$f(q_p(S(k'))) + c_2 \cdot \sigma \leq f(q_p(S(k))) + c_2 \cdot \sigma = p - g_r(\omega(k)) + c_2 \cdot \sigma \leq p - g_r(\omega(k))/2.$$

(2) For every step k' such that $q_p(S(k')) \geq q_p(S(k))$ the probability of accepting a candidate in $[0, q_p(S(k))]$ is at most the probability of accepting a candidate in $[0, q_p(S(k'))]$ which is:

$$f(q_p(S(k'))) \leq f(q_p(S(k)) + \sigma) = p - g_r(\omega(k) - \sigma) \leq p - g_r(\omega(k))/2,$$

where in the last transition we used our assumptions on σ .

Hence the probability of accepting a candidate in $[0, q_p(S(k))]$ is at most $p - g_r(\omega(k))/2$. We can now use Chernoff bounds to compute the probability that the number of members that join the interval $[0, q_p(S(k))]$ in the next t steps is more than $(p - \frac{g_r(\omega(k))}{4}) \cdot t$. Denote the number of candidates that joined the interval $[0, q_p(k))]$ by X , then:

$$\begin{aligned} Pr[X \geq (1 + \frac{g_r(\omega(k))}{4p}) \cdot (p - \frac{g_r(\omega(k))}{2}) \cdot t] &\leq e^{-\frac{(\frac{g_r(\omega(k))}{4p})^2 \cdot (p - \frac{g_r(\omega(k))}{2}) \cdot t}{3}} \\ &\leq e^{-\frac{(\frac{g_r(\omega(k))}{4p})^2 \cdot (g_r(\omega(k)) - \frac{g_r(\omega(k))}{2}) \cdot t}{3}} \\ &= e^{-\frac{g_r(\omega(k))^2}{96p} \cdot t} \end{aligned}$$

For the transition before the last we use the assumption that $g_r(\omega(k)) < p$ for every $\omega(k) < \tau_p$. This implies that $\omega(S(k+t)) \leq \omega(S(k))$ (since the number of member that joined $[0, q_p(S(k))]$ is less than $p \cdot t$ and in particular that the number of members in the interval $[q_p(S(k)), q_p(S(k+t))]$ is at least $\frac{g_r(\omega(k))}{4} \cdot t$). The last implies that in the group $S(k+t)$ the number of points separating $q_p(S(k))$ and $q_p(S(k+t))$ is at least $\frac{g_r(\omega(k))}{4} \cdot t$, as required. \square

The proof for the symmetric case is very much similar hence we only state the corresponding proposition without repeating the proof:

PROPOSITION B.2. Consider adding t more members to a group $S(k)$, such that $q_p(S(k)) > \tau_p$. For any $\sigma < \omega(k)$ such that:

(1) $g_l(\omega(k) - \sigma) > g_l(\omega(k))/2 > c_2 \cdot \sigma$.

(2) Each of the intervals $[q_p(S(k)) - \sigma, q_p(S(k))]$ and $[q_p(S(k)), q_p(S(k)) + \sigma]$ contain at least t members.

The following hold with probability at least $1 - e^{-\Theta(g_l(\omega(k))^2 \cdot t)}$:

- (1) $\omega(S(k+t)) \leq \omega(S(k))$.
- (2) The group $S(k+t)$ contains at least $\frac{g_l(\omega(k))}{4} \cdot t$ members in the interval $[q_p(S(k)), q_p(S(k+t))]$.

B.2. Density Bounds

We now provide bounds on the density of the group in every interval and every step. In particular, consider a group of size k and let $\delta(k) = k^{-1/10}$, we will show that for large enough k , each interval I of length $|I| \geq \delta(k)$ contains at least $c'_1 \cdot |I| \cdot \delta(k) \cdot k$ members and at most $c'_2 \cdot |I| \cdot k$ members. We begin by partitioning the $[0, 1]$ interval into equal segments of length $\delta(k)/2$:

LEMMA B.3. Consider adding t new members to the group $S(k)$. With high probability $(1 - \frac{2}{\delta(k)} \cdot e^{-\Theta(\delta(k)^2 \cdot t)})$ for every segment J in the $\delta(k)/2$ -partition :

- (1) The number of members accepted in the t steps to J is at least $c_1 \cdot \frac{\delta(k)^2}{8} \cdot t$.
- (2) The number of members accepted in the t steps to J is at most $c_2 \cdot \delta(k) \cdot t$.

PROOF. We first compute the probability that a specific interval J has the right number of candidates and then apply a union bound to show that the lemma holds for all intervals simultaneously. Throughout this proof we denote the number of accepted candidates that are located in an interval J by X_J .

- (1) **Lower bound** - By the assumption that the admission process is smooth we have that the probability of accepting a candidate in a segment of length $\frac{\delta(k)}{2}$ is at least $c_1 \cdot \frac{\delta(k)^2}{4}$. Thus, by taking a Chernoff bound we get that the number of candidates accepted to interval J is at least $c_1 \cdot \frac{\delta(k)^2}{8} \cdot t$ with probability $(1 - e^{-\frac{c_1 \cdot \delta(k)^2 t}{32}})$:

$$Pr[X_J \leq (1 - 0.5) \cdot c_1 \cdot \frac{\delta(k)^2}{4} t] \leq e^{-\frac{\frac{1}{4} \cdot c_1 \cdot \frac{\delta(k)^2}{4} t}{2}} = e^{-\frac{c_1 \cdot \delta(k)^2 t}{32}}.$$

- (2) **Upper bound** - By the assumption that the admission process is smooth we have that the probability of accepting a candidate in a segment of length $\frac{\delta(k)}{2}$ is at most $c_2 \cdot \frac{\delta(k)}{2}$. Thus, by taking a Chernoff bound we get that the number of candidates accepted to interval J is at most $c_2 \cdot \delta(k) \cdot t$ with probability $(1 - e^{-\frac{c_2 \cdot \delta(k) t}{24}})$:

$$Pr[X_J \geq (1 + 0.5)c_2 \cdot \frac{\delta(k)}{2} t] \leq e^{-\frac{\frac{1}{4} \cdot c_2 \cdot \frac{\delta(k)}{2} t}{3}} = e^{-\frac{c_2 \cdot \delta(k) t}{24}}.$$

Finally we take a union bound to show that all the segments have the right number of members with high probability:

$$\frac{2}{\delta(k)} (e^{-\frac{c_1 \cdot \delta(k)^2 t}{32}} + e^{-\frac{c_2 \cdot \delta(k) t}{24}}) \leq \frac{2}{\delta(k)} \cdot e^{-\Theta(\delta(k)^2 \cdot t)}.$$

□

Next, we use the bounds on the smaller consecutive segments to show that any interval of length greater than $\delta(k)$ contains the “right” number of members:

CLAIM B.4. Let $k \geq \frac{k_0}{\delta(k)}$, where k_0 is the initial size of the group, for any interval I of length $|I| \geq \delta(k)$ the following holds with probability of at least $1 - \frac{2}{\delta(k)} \cdot e^{-\Theta(\delta(k)^2 \cdot k)}$:

- (1) The number of members of $S(k)$ in I is at least $c'_1 \cdot |I| \cdot \delta(k) \cdot k$ for $c'_1 < c_1$.
- (2) The number of members of $S(k)$ in I is at most $c'_2 \cdot |I| \cdot k$, for $c'_2 > c_2$.

PROOF. In Lemma B.3 we partitioned the interval $[0, 1]$ to disjoint segments of length $\frac{\delta(k)}{2}$ and proved bounds on the number of members in each small segment. To show that similar density bounds hold for *any* interval of length at least $\delta(k)$, we observe that any interval of size $|I| \geq \delta(k)$ is contained in an interval I^l of length at most $2|I|$ consisting of consecutive segments of our $\delta(k)/2$ -partition and contains an interval I^s of length at least $\max\{|I| - \delta(k), \frac{\delta(k)}{2}\} \geq |I|/3$ of consecutive segments of the $\delta(k)/2$ -partition. Thus, for the lower bound we get that each interval I includes at least $|I|/(6\delta(k))$ segments of the $\delta(k)/2$ -partition. Since each of these segments includes at least $c_1 \cdot \frac{\delta(k)^2}{8} \cdot (k - k_0)$ members and $k \geq \frac{k_0}{\delta(k)}$, we have that there exists a constant c'_1 such that the number of members in I is at least $c'_1 \cdot |I| \cdot \delta(k)$.

Similarly, for the upper bound this implies that each segment of length $\frac{\delta(k)}{2}$ contains at most $c_2 \cdot \delta(k)(k - k_0)$ members who joined the group in the admission process. Thus the interval I^l includes at most $2c_2 \cdot |I| \cdot (k - k_0)$ such members. Note that in the worst case all the k_0 initial members were located in the interval I . By the assumption that $k \geq \frac{k_0}{\delta(k)}$, we have that the number of members in this interval is at most $2c_2 \cdot |I| \cdot (k - k_0) + k_0 \leq 2c_2 \cdot |I| \cdot k + k \cdot \delta(k)$. Since $|I| \geq \delta(k)$ we have that there exists a constant c'_2 such that the number of members in interval I is at most $c'_2 \cdot |I| \cdot k$. \square

B.3. Putting it all Together

The essence of Theorem 2.3 is repeated application of Propositions B.1 and B.2 and Claim B.4. To formalize this idea we define a sequence of group sizes and then argue how the p -quantile changes between them. Let $k_1 = \max\{k_\varepsilon, k_0^{10/9}\}$. Where, k_ε is chosen be such that $g_l(\varepsilon/2), g_r(\varepsilon/2) > 8 \cdot c'_2 \cdot \delta(k_\varepsilon)$ and for every $\omega > \varepsilon/2$, $g_r(\omega - \delta(k_\varepsilon)) > g_r(\omega)/2$ and $g_l(\omega - \delta(k_\varepsilon)) > g_l(\omega)/2$. Such k_ε exists by continuity. The exact reasoning behind this choice of k_ε will become clearer later. Next, for every $i \geq 1$ let $k_{i+1} = (1 + c'_1 \cdot \delta(k_i)^2) \cdot k_i$, we show that the p -quantile cannot get too far from τ_p and under some conditions it gets closer to τ_p :

CLAIM B.5. Suppose that each of the intervals $[q_p(S(k_i)) - \delta(k_i), q_p(S(k_i))]$ and $[q_p(S(k_i)), q_p(S(k_i)) + \delta(k_i)]$ contains at least $c'_1 \cdot \delta(k_i)^2 \cdot k_i$ members. Then, the following holds with probability at least $1 - e^{-\Theta(k_i^{3/5})}$:

- (1) For every $k_i < k' \leq k_{i+1}$, $\omega(S(k')) < \omega(S(k_i)) + \delta(k_i)$.
- (2) Each of the intervals $[q_p(S(k_{i+1})) - \delta(k_{i+1}), q_p(S(k_{i+1}))]$ and $[q_p(S(k_{i+1})), q_p(S(k_{i+1})) + \delta(k_{i+1})]$ contains at least $c'_1 \cdot \delta(k_{i+1})^2 \cdot k_{i+1}$ members. More generally, every interval I of length $|I| \geq \delta(k_{i+1})$ includes at least $c'_1 \cdot |I| \cdot \delta(k_{i+1}) \cdot k_{i+1}$ members and at most $c'_2 \cdot |I| \cdot k_{i+1}$ members.
- (3) If $q_p(S(k_i)) < \tau_p$ and $g_r(\omega(k_i) - \delta(k_i)) > g_r(\omega(k_i))/2 > c_2 \cdot \delta(k_i)$, then, the number of members in the interval $[q_p(S(k_i)), q_p(S(k_{i+1}))]$ in the group $S(k_{i+1})$ is at least $\frac{g_r(\omega(k_i))}{4} \cdot c'_1 \cdot \delta(k_i)^2 \cdot k_i$ (in particular $q_p(S(k_i)) \leq q_p(S(k_{i+1})) < \tau_p$).
- (4) If $q_p(S(k_i)) > \tau_p$ and $g_l(\omega(k_i) - \delta(k_i)) > g_l(\omega(k_i))/2 > c_2 \cdot \delta(k_i)$, then, the number of members in the interval $[q_p(S(k_{i+1})), q_p(S(k_i))]$ in the group $S(k_{i+1})$ is at least $\frac{g_l(\omega(k_i))}{4} \cdot c'_1 \cdot \delta(k_i)^2 \cdot k_i$ (in particular $q_p(S(k_i)) \geq q_p(S(k_{i+1})) > \tau_p$).

PROOF. Observe that statement (1) holds simply by the assumption that each of the intervals $[q_p(S(k_i)) - \delta(k_i), q_p(S(k_i))]$ and $[q_p(S(k_i)), q_p(S(k_i)) + \delta(k_i)]$ contains at least $c'_1 \cdot \delta(k_i)^2 \cdot k_i = k_{i+1} - k_i$ members. Thus, when increasing the group by $c'_1 \cdot \delta(k_i)^2 \cdot k_i$ members the p -quantile cannot move a distance greater than $\delta(k_i)$.

Next, recall that $k_1 \geq \frac{k_0}{\delta(k_1)}$. Thus, we can apply Claim B.4 and get that statement (2) holds with probability at least $1 - \frac{2}{\delta(k_{i+1})} \cdot e^{-\Theta(\delta(k_{i+1})^2 \cdot k_{i+1})}$.

For the last two statements we apply Proposition B.1 and Proposition B.2 using $\sigma = \delta(k_i)$ and $t = c'_1 \cdot \delta(k_i)^2 \cdot k_i$ we have that the two statements hold with probabilities at least $1 - e^{-\Theta(g_r(\omega(k_i))^2 \cdot \delta(k_i)^2 \cdot k_i)}$ and $1 - e^{-\Theta(g_l(\omega(k_i))^2 \cdot \delta(k_i)^2 \cdot k_i)}$ respectively. By the assumption that $\delta(k_i) < g_r(\omega(k_i))/(2 \cdot c_2)$ for statement (3) and that $\delta(k_i) < g_l(\omega(k_i))/(2 \cdot c_2)$ for statement (4) we have that we can bound each of these probabilities by $1 - e^{-\Theta(\delta(k_i)^4 \cdot k_i)}$.

Thus, by taking a union bound we have that the claim holds with probability at least $1 - \frac{2}{\delta(k_{i+1})} \cdot e^{-\Theta(\delta(k_{i+1})^2 \cdot k_{i+1})} - e^{-\Theta(\delta(k_i)^4 \cdot k_i)}$. By using the fact that $\delta(k_i) = k_i^{-1/10}$ and $k_{i+1} > k_i$ we can bound this probability by

$$1 - 2k_{i+1}^{1/10} \cdot e^{-\Theta(k_{i+1}^{4/5})} - e^{-\Theta(k_i^{3/5})} \geq 1 - e^{-\Theta(k_i^{3/5})}.$$

□

We are now ready to show that the p -quantile indeed gets closer to τ_p . To this end, larger increments of the group's size are required. Thus, we define the following series a_j such that $2k_{a_j} \leq k_{a_{j+1}}$ and for every $i < a_{j+1}$, $k_i < 2k_{a_j}$, and consider the changes in the p -quantile from $S(k_{a_j})$ to $S(k_{a_{j+1}})$. Based on Claim B.5 we prove the following proposition:

PROPOSITION B.6. *For every j , with probability of at least $1 - (a_{j+1} - a_j)e^{-\Theta(k_{a_j}^{3/5})}$:*

- (1) If $\omega(k_{a_j}) > \frac{3}{4}\varepsilon$, then $\omega(k_{a_{j+1}}) < \omega(k_{a_j}) - \frac{g_r(\frac{3}{4}\omega(k_{a_j}))}{8c'_2}$.
- (2) Else, $\omega(k_{a_{j+1}}) < \frac{3}{4}\varepsilon + \delta(k_{a_j})$.

PROOF. We present the proof for $q_p(S(k_{a_j})) < \tau_p$, as the proof for the symmetric case is identical. We begin by considering the case that for all steps k_i such that $k_{a_j} \leq k_i < k_{a_{j+1}}$ we have that $\omega(k_i) \geq \frac{2}{3}\omega(k_{a_j})$. Recall that $k_1 > k_\varepsilon$ and k_ε was chosen such that:

- For every $\omega > \varepsilon/2$, $g_r(\omega - \delta(k_\varepsilon)) > g_r(\omega)/2$. This implies that for every $a_j < i < a_{j+1}$, $g_r(\omega(k_i) - \delta(k_i)) > g_r(\omega(k_i))/2$, since $g_r(\cdot)$ is a strictly increasing function and $\delta(k_i) < \delta(k_\varepsilon)$.
- $g_r(\varepsilon/2) > 8 \cdot c'_2 \cdot \delta(k_\varepsilon)$. This implies that for every $a_j < i < a_{j+1}$, $g_r(\omega(k_i))/2 > 8 \cdot c'_2 \cdot \delta(k_i) > c_2 \cdot \delta(k_i)$ since $c'_2 > c_2$ and $\delta(k_i) < \delta(k_\varepsilon)$.

Next, we apply Claim B.4 and get that with probability at least $1 - \frac{2}{\delta(k_{a_j})} \cdot e^{-\Theta(\delta(k_{a_j})^2 \cdot k_{a_j})}$ both intervals $[q_p(S(k_{a_j})), q_p(S(k_{a_j})) + \delta(k_{a_j})]$ and $[q_p(S(k_{a_j})) - \delta(k_{a_j}), q_p(S(k_{a_j}))]$ contain at least $c'_1 \cdot \delta(k_{a_j})^2 \cdot k_{a_j}$ members. Since we established that $g_r(\omega(k_i) - \delta(k_i)) > g_r(\omega(k_i))/2 > c_2 \cdot \delta(k_i)$, we can now apply Claim B.5 repeatedly for $a_j \leq i < a_{j+1}$. We get that the number of members in the interval $[q_p(S(k_i)), q_p(S(k_{i+1}))]$ in the group $S(k_{i+1})$ is at least $\frac{g_r(\omega(k_i))}{4} \cdot c'_1 \cdot \delta(k_i)^2 \cdot k_i$ and in particular $q_p(S(k_i)) \leq q_p(S(k_{i+1})) < \tau_p$.

Finally, we sum over all the indices i , $a_j \leq i < a_{j+1}$ to get that the number of members that joined the group in the interval $[q_p(S(k_{a_j})), q_p(S(k_{a_{j+1}}))]$ is at least:

$$\sum_{i=a_j}^{a_{j+1}-1} \frac{g_r(\omega(k_i))}{4} \cdot c'_1 \cdot \delta(k_i)^2 \cdot k_i \geq \frac{g_r(\frac{2}{3}\omega(k_{a_j}))}{4} \cdot \sum_{i=a_j}^{a_{j+1}-1} c'_1 \cdot \delta(k_i)^2 \cdot k_i = \frac{g_r(\frac{2}{3}\omega(k_{a_j}))}{4} \cdot (k_{a_{j+1}} - k_{a_j})$$

Note that by our construction of the series a_j we have that $k_{a_j} \leq k_{a_{j+1}}/2$, hence, the number of members that joined the interval $[q_p(S(k_{a_j})), q_p(S(k_{a_{j+1}}))]$ is at least $\frac{g_r(\frac{2}{3}\omega(k_{a_j}))}{8} \cdot k_{a_{j+1}}$. Finally observe that by applying Claim B.5 over $S(k_{a_{j+1}})$ we have that in the group $S(k_{a_{j+1}})$ every interval $|I|$ of length $|I| \geq \delta(k_{a_{j+1}})$ contains at most $c'_2 \cdot |I| \cdot k_{a_{j+1}}$ members. This means that if $\frac{g_r(\frac{2}{3}\omega(k_{a_j}))}{8c'_2} > \delta(k_{a_{j+1}})$ (as we assumed), then, the length of the interval $[q_p(S(k_{a_j})), q_p(S(k_{a_{j+1}}))]$ is at least $\frac{g_r(\frac{2}{3}\omega(k_{a_j}))}{8c'_2}$ as required.

The case in which there exists a step $k_{a_j} < k_i < k_{a_{j+1}}$ such that that $\omega(k_i) < \frac{2}{3}\omega(k_{a_j})$ is even simpler. In this case, by repeatedly applying Claim B.5 we have that:

- If $\omega(k_l) < \frac{2}{3}\omega(k_{a_j})$, then $\omega(k_{l+1}) < \frac{2}{3}\omega(k_{a_j}) + \delta(k_l)$.
- If $\omega(k_l) \geq \frac{2}{3}\omega(k_{a_j})$, then $\omega(k_{l+1}) < \omega(k_l)$.

Thus, by induction we have that for any $l > i$, $\omega(k_l) < \frac{2}{3}\omega(k_{a_j}) + \delta(k_{a_j})$ and hence $\omega(k_{a_{j+1}}) < \frac{2}{3}\omega(k_{a_j}) + \delta(k_{a_j})$. Note, that for this case it is possible that for some l , $q_p(S(k_l)) > \tau_p$ however by our choice of k_ε , it would still be the case that $\omega(k_l) < \frac{2}{3} \cdot \omega(k_{a_j}) + \delta(k_{a_j})$.

The proof of the second statement is identical to the second case of the first statement, for any $l > i$:

- If $\omega(k_l) < \frac{3}{4}\varepsilon$, then $\omega(k_{l+1}) < \frac{3}{4}\varepsilon + \delta(k_l)$.
- If $\omega(k_l) \geq \frac{3}{4}\varepsilon$, then $\omega(k_{l+1}) < \omega(k_l)$.

Thus, by induction we have that $\omega(k_{a_{j+1}}) < \frac{3}{4}\varepsilon + \delta(k_{a_j})$.

Lastly, observe that in the proof we basically applied Claim B.5 $a_{j+1} - a_j$ times, hence by taking a union bound the assertion of the proposition holds with probability of at least $1 - (a_{j+1} - a_j)e^{-\Theta(k_{a_j}^{3/5})}$. \square

Finally we are ready to complete the proof of Theorem 2.3. To this end, we do an induction over the series k_{a_j} . As long as $\omega(k_{a_j}) > \frac{3}{4}\varepsilon$ we can apply Proposition B.6 repeatedly and get that $\omega(k_{a_{j+1}}) < \omega(k_{a_j}) - \frac{g_r(\omega(k_{a_j})/2)}{8c'_2}$. In particular, as long as $\omega(k_{a_j}) > \frac{3}{4}\omega(k_{a_j})$, the distance to τ_p is reduced by at least $\frac{g_r(\omega(k_{a_j})/2)}{8c'_2}$ hence after at most $\frac{8c'_2}{g_r(\omega(k_{a_j})/2)}$ iterations the p -quantile is closer to τ_p by a factor of at least $3/4$. We can continue doing so till we reach a distance of $\frac{3}{4}\varepsilon$. Let $k'_\varepsilon > k_1 \geq k_\varepsilon$ be such that $\omega(k'_\varepsilon) \leq \frac{3}{4}\varepsilon$. For any $k_{a_{j+1}} > k'_\varepsilon$, the second statement in Proposition B.6 tells us that $\omega(k_{a_{j+1}}) < \frac{3}{4}\varepsilon + \delta(k_\varepsilon) < \varepsilon$. The proof is then completed by noticing that Claim B.5 which we used in the induction actually guarantees that for any $k' > k_{a_{j+1}}$ we have that $\omega(k') < \frac{3}{4}\varepsilon + 2\delta(k_\varepsilon) < \varepsilon$. To prove the theorem we repeatedly applied Proposition B.6, hence the theorem holds with probability $1 - \sum_{k=k_\varepsilon}^{\infty} e^{-\Theta(k^{3/5})}$.

C. PROOFS FROM SECTION 5

Proof of Proposition 5.4. Recall that we want to show that for the consensus admission rule (i.e., $\ell = k$), if $n \geq 3$ and the initial configuration has diameter $D = x_n - x_1$, then every new element that will be added to the committee during the process is at least $x_1 - D$ and at most $x_n + D$.

It suffices to show we never get an element bigger than $x_n + D$, as by symmetry the same argument implies we do not get one below $x_1 - D$.

We claim that during the process the quantity $x_n + x_2 - x_1$ does not increase. To prove it let the configuration before a step be $x_1 < x_2 \dots < x_n$ and the configuration after a step in which y has been added be $x'_1 < x'_2 < \dots < x'_n$. We have to show that $x'_n + x'_2 - x'_1 \leq x_n + x_2 - x_1$.

Since y cannot replace x_i for $2 \leq i \leq n-1$, There are only two possible cases.

Case 1: The new element y replaced x_1 . In this case $x_1 \leq y \leq 2x_2 - x_1$. If $y < x_n$ then if $y < x_2$ the claim is trivial and otherwise $x'_n = x_n$, $x'_1 = x_2$ and $x'_2 \leq y \leq 2x_2 - x_1$, implying that

$$x'_n + x'_2 - x'_1 \leq x_n + (2x_2 - x_1) - x_2 = x_n + x_2 - x_1,$$

as needed.

If $y > x_n$ then $x'_1 = x_2$, $x'_2 = x_3 \leq x_n$ and $x'_n = y \leq 2x_2 - x_1$ and hence

$$x'_n + x'_2 - x'_1 \leq (2x_2 - x_1) + x_n - x_2 = x_n + x_2 - x_1.$$

Case 2: The new element y replaced x_n . In this case $2x_{n-1} - x_n \leq y \leq x_n$. If $y > x_1$ then $x'_n \leq x_n$ and $x'_2 - x'_1 \leq x_2 - x_1$ implying the desired result. Otherwise $x'_1 = y \geq 2x_{n-1} - x_n$, $x'_2 = x_1$ and $x'_n = x_{n-1}$. Hence

$$x'_n + x'_2 - x'_1 \leq x_{n-1} + x_1 - (2x_{n-1} - x_n) = x_n - (x_{n-1} - x_1) \leq x_n \leq x_n + x_2 - x_1.$$

This completes the proof of the claim.

Note, now, that in the beginning $x_n + x_2 - x_1 \leq x_n + D$. Therefore, if at some point during the process we have a configuration $(x''_1, x''_2, \dots, x''_n)$, then $x''_n \leq x''_n + x''_2 - x''_1 \leq x_n + D$, by the claim. This completes the proof.

Remark: It is easy to see that if $n = 2$ this is not true and that the above is tight, namely we can add elements as close as we wish to $x_n + D$ or to $x_1 - D$ with appropriate initial configurations. ■

Proof of Lemma 5.3. Let the configuration before a step be $x = (x_1, x_2, \dots, x_{2k+1})$, and the configuration after a step in which y has been added and x_i been dropped be $x' = (x'_1, x'_2, \dots, x'_{2k+1})$. Recall that we want to show that if the median moved to the right, then the sum of distances from the median has decreased by at least $2 \sum_{j=k-\ell+2}^k d(x_j, x'_j) + d(x_{k+1}, x'_{k+1})$.

For the conditions of the lemma to hold, it must be that $y > x_{k+1}$, $x_i < x_{k-\ell+2}$ and $x_{k-\ell+2}$ (weakly) prefers y to x_i (i.e., $d(x_i, x_{k-\ell+2}) \geq d(x_{k-\ell+2}, y)$). Let $S = \sum_{j=1}^{2k+1} d(x_j, x_{k+1})$ be the sum of distances from the median in configuration x . We distinguish between two cases.

Case 1: y is the new median. Let S' denote the sum of distances from the (new) median in x' . Since the distance between x_{k+1} and y is added to k elements and subtracted from k elements, we have $S' = S - d(x_i, x_{k+1})$. It holds that $2 \sum_{j=k-\ell+2}^k d(x'_j, x_j) + d(x'_{k+1}, x_{k+1}) = 2d(x_{k-\ell+2}, x_{k+1}) + d(x_{k+1}, y)$. Therefore, to establish the assertion of the lemma we need to show that $d(x_i, x_{k+1}) \geq 2d(x_{k-\ell+2}, x_{k+1}) + d(x_{k+1}, y)$, or equivalently (by substituting $d(x_i, x_{k+1}) = d(x_i, x_{k-\ell+2}) + d(x_{k-\ell+2}, x_{k+1})$) that $d(x_i, x_{k-\ell+2}) \geq$

$d(x_{k-\ell+2}, x_{k+1}) + d(x_{k+1}, y)$. But the right hand side is exactly $d(x_{k-\ell+2}, y)$, which is at most $d(x_i, x_{k-\ell+2})$ by the fact that $x_{k-\ell+2}$ has chosen y over x_i , as desired.

Case 2: x_{k+2} is the new median (here, $y > x_{k+2}$). Let S' denote the sum of distances from the (new) median in x' . Since the distance between x_{k+1} and x_{k+2} is added to k elements and subtracted from k elements, we have $S' = S - d(x_i, x_{k+1}) + d(y, x_{k+2})$. It holds that $2 \sum_{j=k-\ell+2}^k d(x'_j, x_j) + d(x'_{k+1}, x_{k+1}) = 2d(x_{k-\ell+2}, x_{k+1}) + d(x_{k+1}, x_{k+2})$. Therefore, to establish the assertion of the lemma we need to show that $d(x_i, x_{k+1}) - d(y, x_{k+2}) \geq 2d(x_{k-\ell+2}, x_{k+1}) + d(x_{k+1}, x_{k+2})$, or equivalently (by substituting $d(x_i, x_{k+1}) = d(x_i, x_{k-\ell+2}) + d(x_{k-\ell+2}, x_{k+1})$ and rearranging) that $d(x_i, x_{k-\ell+2}) \geq d(x_{k-\ell+2}, x_{k+1}) + d(x_{k+1}, x_{k+2}) + d(x_{k+2}, y)$. But the right hand side is exactly $d(x_{k-\ell+2}, y)$, which is at most $d(x_i, x_{k-\ell+2})$ by the fact that $x_{k-\ell+2}$ has chosen y over x_i , as desired.

Remark C.1. If $n = 2k$ (i.e., n is even), then there are two medians. It is easy to verify that the sum of distances of the points from the left median equals the sum of their distances from the right median (as the difference is that in the distance from the left median, the distance between the right and left medians is counted for all the points to the right of the left median, and in the distance from the right median, it is counted for all the points to the left of the right median. In both cases this distance is counted k times). We leverage this observation to show, in a similar way to the odd case, that if the left median moves right, then the sum of distances from (either one of) the medians decreases by at least $2 \sum_{j=k-\ell+1}^k d(x'_j, x_j)$.

■

Proof of Proposition 5.7. Recall that we want to show for $n = 4k + 3$ and a majority of at most $3k + 2$ is required to replace an existing member, then for any initial configuration no element has immunity.

We describe the process that removes the $2k + 2$ points up to, and including, the median. An analogous process can be applied to the points to the right of the median, showing that all points can be removed. We describe the process in stages, each stage $j = 1, \dots, 2k + 2$ handles points x_1, \dots, x_j . For ease of presentation, when clear in the context we denote the i th point of a given configuration (not necessarily the initial one) by x_i .

For every $j = 1, \dots, 2k + 2$, stage j begins in a configuration where the smallest j points form an arithmetic progression, and ends in a configuration where x_j has been removed and the smallest $j + 1$ points form an arithmetic progression. This is done as follows. Suppose the smallest j points form an arithmetic progression. We replace x_1 with $z_1 = x_j - \delta_j$, where δ_j divides the difference between x_{j+1} and x_j and is smaller than $\frac{x_{j+1} - x_j}{j+1}$. We then replace x_2 by $z_2 = z_1 - \delta_j$, and so on, until the smallest j points form an arithmetic progression with difference δ_j . In every such replacement there are at least $3k + 2$ members who prefer the new point to the old one — these are the $2k + 1$ members to the right of the median, the median itself, and at least k out of the $2k + 1$ members to the left of the median. For simplicity, let x_1, \dots, x_j denote the new locations, i.e., $x_i = z_{j-i}$. Now we would like the j point to progress to x_{j+1} . Now, replace x_1 by $z_1 = x_j + \delta_j$ (again, following a similar argument to the one above, at least $3k + 2$ points prefer x'_1 to x_1). We again rename x_1, \dots, x_j to denote the new locations. We continue the process until we reach x_{j+1} and the $j + 1$ left-most points form an arithmetic progression. By the choice of δ_j all points including x_j have been replaced, and are still all smaller than x_{j+1} . Repeat this process until we reach the median and replace it. Then we follow the same process from the right side.

Note that in the beginning we can make the process simpler, and only as we get close to the median we have to be careful, but we prefer to keep the description uniform, for ease of presentation.

Remark: Note that the process above establishes a stronger property. That is, not only can we ensure to omit the element x_i for every fixed i , we can omit *all* elements of the committee together in one process. ■