Dynamics of Evolving Social Groups

NOGA ALON, Princeton University, USA and Tel-Aviv University, Israel
MICHAL FELDMAN, Tel-Aviv University and Microsoft Research, Israel
YISHAY MANSOUR, Tel-Aviv University, Israel
SIGAL OREN, Ben-Gurion University of the Negev, Israel
MOSHE TENNENHOLTZ, Technion—Israel Institute of Technology, Israel

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 this article, 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 run? 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.

14

CCS Concepts: • **Theory of computation → Algorithmic game theory**;

Additional Key Words and Phrases: Social networks, computational social choice, price of anarchy

ACM Reference format:
Noga Alon, Michal Feldman, Yishay Mansour, Sigal Oren, and Moshe Tennenholtz. 2019. Dynamics of Evolving Social Groups. *ACM Trans. Econ. Comput.* 7, 3, Article 14 (September 2019), 27 pages.
https://doi.org/10.1145/3355948

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.

The research of N. Alon was supported in part by ISF Grant No. 281/17, GIF Grant No. G-1347-304.6/2016, and the Simons Foundation. The work of M. Feldman was partially supported by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007–2013)/ERC Grant Agreement No. 337122 and by the Israel Science Foundation (Grant No. 317/17). The research of M. Tennenholtz was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant No. 740435).
Authors' addresses: N. Alon, Fine Hall, Princeton University, NJ, USA; email: nalon@math.princeton.edu; M. Feldman and Y. Mansour, Blavatnik School of Computer Science Schreiber Building, Tel Aviv University, Tel Aviv, Israel; emails: michal.feldman@cs.tau.ac.il, mansour@post.tau.ac.il; S. Oren, Alon Building, Department of Computer Science, Ben-Gurion University of the Negev, Beer Sheva, Israel; email: sigal3@gmail.com; M. Tennenholtz, Bloomfield Building, Technion—Israel Institute of Technology, Haifa, Israel; email: moshet@ie.technion.ac.il.

2167-8375/2019/09-ART14 $15.00
https://doi.org/10.1145/3355948

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 the case?

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., References [14, 19, 24]), 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, similarity can be easily measured by the distance between opinions. This modeling choice is heavily grounded in the literature on homophily (e.g., Reference [23] 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 article to two large bodies of literature, one on opinion formation [5, 13, 21], the other on cultural dynamics [3, 17, 22]. 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 article 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 we 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 $S(k) \in [0, 1]^k$. We denote by k_0 the initial size of the group. The admission process operates in discrete time steps. At each time step two candidates are

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.

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 $|x_i - y_j|$.[1] 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, then, 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 him, all the members to the left of $(y_1 + y_2)/2$ vote for y_1, while all the 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 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 shift of the median technically more difficult.

[1]We assume that all the members have the same voting power, which does not change throughout the process.

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 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, then 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 the p-quantile always converges for quantile-driven admission rules that have two rather natural properties. 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 opinions 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.

1.3 Related Work

Group formation and evolution is a central topic in Sociology [20]. Much of the research considers groups that evolve organically in the sense that there is no formal mechanism governing who can join the group. Research on this includes detecting social group and communities in networks (see Reference [18] for a comprehensive survey). Backstrom et al. [4] take a different approach and instead of detecting group attempt to understand how they evolve.

Closer in spirit to the current article is an experimental line of work originating from References [1, 2, 15]. This line of work studies endogenous group formation in the context of public good games. In particular, Ahn et al. study the majority admission rule and its affect on members contribution (in comparison to groups that anyone can join to). Cinyabuguma et al. [11] study a different process in which at each step the group votes on who should be removed from the group. Charness and Yang [9] study a more complex process in which a majority is required to remove someone from the group but a 60% majority is required to accept a member (or merge to another group). While the mechanisms in these experimental works are somewhat similar to the process of joining a social group in the current article, the focus of our article is different. In the experimental works the decision on whether to admit an individual depends on the expectations of the group of his contribution. While in the current article the emphasis is on the opinion of the members and the decisions on whether to admit a candidate depends on his opinion.

Also related is the literature on opinion formation in which, loosely speaking, agents form their opinion based on the opinions of their friends in a social network. A notable example is the Degroot model [14] in which at each time step individuals take an average of all their neighbors' opinions to form their opinion. Variants of it include process in which each node has some internal non-changing opinion [19] and bounded confidence in which nodes only take into account opinions of neighbors sufficiently similar to them [13, 21]. See Reference [8] for a thorough survey. References [6, 7, 10] take a utilitarian view towards opinion formation a node incurs some (cognitive) cost for having an opinion different than its neighbors and it aims to choose an opinion to minimize this cost. Using this approach, Reference [12] showed that opinion formation processes based on the DeGroot model are *not* polarizing.[2] Our work takes a different approach and instead of considering mechanisms in which an individual is forming his opinion, we consider mechanisms in which, in a sense, the whole group is forming its opinion.

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 counterintuitive, 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:

[2] Where polarization is defined as the sum of the nodes' costs.

Proposition 2.1. *Consider a group $S(k_0)$, then, for any ε with probability 1 there exists $k_\varepsilon > k_0$ 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 $x_1(k), \ldots, x_k(k)$. The requirement for all the group members to agree 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 exists 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 $k_\varepsilon = \max\{k_\varepsilon^l, k_\varepsilon^r\}$. □

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.[3] 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 $h(\cdot)$ 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 $1/2$.[4] Hence, if the median converges to $1/2$, then 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 $(2\varepsilon)^2$. 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 $(2\varepsilon^2)k$. 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 applied to join the group, we cannot have more than roughly $4k\delta$ points in any interval of length δ. Thus, the

[3]In the case of a group of an even size, any consistent choice of the median will do. We note that the reformulation of the majority rule by using the median also serves as a tie breaking mechanism for the case that each of the candidates was voted for by exactly half of the members.

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

median will move to the right by at least a distance of about $2\varepsilon^2 k/(4k) = \varepsilon^2/2$. We have 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 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. As in this case the median might "get stuck" at some cluster of points. Instead of formalizing this intuition for the majority rule, 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:

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

While our main focus on this article is on the distribution of opinions that the group converge to, to provide a more complete picture in Appendix A, we present a brief discussion on the convergence rate of the majority and consensus rules.

3 GROWING GROUPS: QUANTILE-DRIVEN ADMISSION PROCESSES

In this section, we define and study a 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 $0 < p < 1$. As we will see this family captures natural admission rules (e.g., the majority rule). We focus on a subfamily of these admission rules and show that as the group evolves the location of the p-quantile converges.

We denote the p-quantile of a group $S(k)$ by $q_p(S(k))$ and 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 $|\{i | x_i < q_p(S(k))\}| \leq p \cdot k$.[5]

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 according to it 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. We denote by $f_p(q_p)$[6] 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

[5]For groups in which this definition admits more than a single choice for the p-quantile, any consistent choice will do.
[6]When p is clear from the context, we denote this function simply by $f(\cdot)$.

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 is using similar ideas to the ones we presented in our intuition for the convergence of the median 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 then show that the majority process is smooth:

Definition 3.3. An admission process in which at every step a candidate joins the group[7] 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 $q = 1/2$ 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. However, 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. □

3.1 Convergence of Smooth Admission Processes

Our main technical result states that the location of the p-quantile of a group that uses a smooth admission rule always converges to unique τ_p such that $f(\tau_p) = p$.[8] Formally, we show that:

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. While the intuition for the convergence of the p-quantile is quite simple, nailing down all of the proof details is quite complex. A key reason for this is that the acceptance probability depends on the location of the p-quantile that keeps changing.

We will need some notation for this proof. First, we let $\omega(k) = |\tau_p - q_p(S(k))|$. Then, we define the following two strictly increasing functions:

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

- $g_r(\omega) : [0, \tau_p] \rightarrow [0, p]$, $g_r(\omega) = p - f_p(\tau_p - \omega)$.
- $g_l(\omega) : [0, 1 - \tau_p] \rightarrow [0, 1 - p]$, $g_l(\omega) = f_p(\tau_p + \omega) - p$.

The function $g_r(\omega)$ is defined for cases in which the p-quantile is left of τ_p and given the distance of the p-quantile from τ_p (i.e., ω), it returns the probability of accepting a candidate in the interval $[\tau_p - \omega, \omega]$. To see why this is the case, observe that by definition $f(\tau_p) = p$. $g_l(\omega)$ is the symmetric function for cases in which the p-quantile is right of τ_p.

The formal proof of Theorem 3.5 has two main building blocks that are used iteratively to show that the p-quantile converges to τ_p. First, in Section 3.1.1, 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 3.1.2, we use Chernoff bounds to show that on the one hand the intervals are dense enough so that the p-quantile will remain long enough in the same interval. On the other hand, they are not too dense to prevent the p-quantile from moving a non-negligible distance towards τ_p. In Section 3.1.3, we carefully use these two building blocks on groups of growing size to show that indeed the p-quantile converges to τ_p.

3.1.1 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 3.6. *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) $g_r(\omega(k) - \sigma) > g_r(\omega(k))/2 > c_2 \cdot \sigma$.
(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 $S(k + t)$ will have at least $\frac{g_r(\omega(k))}{4} \cdot t$ members in the interval $[q_p(S(k)), q_p(S(k + t))]$. Furthermore, $\omega(S(k + t)) \leq \omega(S(k))$.

Proof. Informally, the second condition in the proposition implies that we can add t more members to the group and be sure that for all these additions the p-quantile was confined to the small interval $[q_p(S(k)) - \sigma, q_p(S(k)) + \sigma]$. As we will see the first condition can be used to get that while the quantile is in this interval the probability to accept a candidate in the interval $[0, q_p(S(k))]$ is at most $p - g_r(\omega(k))/2$. By applying Chernoff bounds, we get that in the t rounds at most $(p - g_r(\omega(k))/4)t$ members have joined the interval $[0, q_p(S(k))]$ and hence the p-quantile moved by at least $g_r(\omega(k))/4 \cdot t$ points closer to τ_p. The assumption that $\sigma < \omega(k)$ assures us that the addition of t new members will keep the p-quantile to the left of τ_p and hence $\omega(S(k + t)) \leq \omega(S(k))$.

We now proceed to the formal 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 $[0, q_p(S(k))]$. We distinguish between two cases:

(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 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

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

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 bound the probability that the number of members that join the interval $[0, q_p(S(k))]$ in the next t steps is more than $(p - \frac{g_r(\omega(k))}{4}) \cdot t$. Observe that we can use here an upper bound on the probability of accepting a candidate in $[0, q_p(S(k))]$, since for a smaller probability it will be even less likely to accept more than $(p - \frac{g_r(\omega(k))}{4}) \cdot t$ members in the interval $[0, q_p(S(k))]$. Denote the number of candidates that joined the interval $[0, q_p(k))]$ by X, then

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

For the transition before the last, we use the assumption that $g_r(\omega(k)) < p$ for every $\omega(k) < \tau_p$. By the Chernoff bounds, we have that with high probability 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. In particular, we have that $q_p(S(k)) < q_p(S(k+t))$. By the assumption that $[q_p(S(k)), q_p(S(k)) + \sigma]$ contains at least t members for $\sigma < \omega(k)$, we have that $q_p(S(k+t)) < \tau_p$. Thus, we conclude that $\omega(S(k+t)) \leq \omega(S(k))$. □

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

Proposition 3.7. *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.

With probability at least $1 - e^{-\Theta(g_l(\omega(k))^2 \cdot t)}$, the group $S(k+t)$ will have at least $\frac{g_l(\omega(k))}{4} \cdot t$ members in the interval $[q_p(S(k+t)), q_p(S(k))]$. Furthermore, $\omega(S(k+t)) \leq \omega(S(k))$.

3.1.2 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 3.8. *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 $(1 - e^{-\frac{c_1 \cdot \delta(k)^2 t}{32}})$[9]:

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

(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 $(1 - e^{-\frac{c_2 \cdot \delta(k) t}{6}})$[10]:

$$Pr\left[X_J \geq (1+1) c_2 \cdot \frac{\delta(k)}{2}\right] \leq e^{-\frac{c_2 \cdot \frac{\delta(k)}{2} t}{3}} = e^{-\frac{c_2 \cdot \delta(k) t}{6}}.$$

Finally, we take a union bound over the bad events to show that all the segments have the right number of members with high probability. We get that the probability of a bad event is at most

$$\frac{2}{\delta(k)}\left(e^{-\frac{c_1 \cdot \delta(k)^2 t}{32}} + e^{-\frac{c_2 \cdot \delta(k) t}{6}}\right) = \frac{2}{\delta(k)} \cdot e^{-\Theta(\delta(k)^2 \cdot t)}. \qquad \square$$

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 3.9. *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 3.8, 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

[9]Observe that we can use here a lower bound on the probability of accepting a candidate in a segment of length $\frac{\delta(k)}{2}$, since for a higher probability it will be even less likely to accept less than $c_1 \cdot \frac{\delta(k)^2}{8} \cdot t$ members in interval J.
[10]Observe that we can use here an upper bound on the probability of accepting a candidate in a segment of length $\frac{\delta(k)}{2}$, since for a lower probability it will be even less likely to accept more than $c_2 \cdot \frac{\delta(k)}{2}$ members in interval J.

segments of the $\delta(k)/2$-partition. Thus, for the lower bound, we get that each interval I includes at least $2|I|/(3\delta(k))$ 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 $c_2' \cdot |I| \cdot k$. □

3.1.3 Putting It All Together. The essence of the proof of Theorem 3.5 is repeated application of Propositions 3.6 and 3.7 and Claim 3.9. To formalize this idea, we define a sequence of group sizes and then argue how the p-quantile changes between them. Recall that $\delta(k) = k^{-1/10}$ and Let $k_1 = \max\{k_\varepsilon, k_0^{10/9}\}$. Where, k_ε is chosen such that $g_l(\varepsilon/2) > 8 \cdot c_2' \cdot \delta(k_\varepsilon), g_r(\varepsilon/2) > 8 \cdot c_2' \cdot \delta(k_\varepsilon)$ 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 be clearer later on. 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 3.10. *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 3.9 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 Propositions 3.6 and 3.7 using $\sigma = \delta(k_i)$ and $t = c_1' \cdot \delta(k_i)^2 \cdot k_i$, and 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

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

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 3.10, we prove the following proposition:

Proposition 3.11. *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 $\omega(k_{a_{j+1}}) < \omega(k_{a_j}) - \frac{g_r(\frac{2}{3}\omega(k_{a_j}))}{8c_2'}$.
(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 3.9 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 3.10 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 3.10 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 $\omega(k_i) < \frac{2}{3}\omega(k_{a_j})$ is even simpler. In this case, by repeatedly applying Claim 3.10, 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 $\omega(k_{l+1}) < \omega(k_l)$.

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

Last, observe that in the proof we basically applied Claim 3.10 $a_{j+1} - a_j$ times, hence by taking a union bound the assertion of the proposition holds with probability of at least $1 - (a_{j+1} - a_j)e^{-\Theta(k_{a_j}^{3/5})}$. □

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 3.11 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 until 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 3.11 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 3.10, 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 3.11, hence the theorem holds with probability $1 - \sum_{k=k_\varepsilon}^{\infty} e^{-\Theta(k^{3/5})}$.

4 GROWING GROUPS: SPECIAL VETO POWER

In this section, we assume that the group has a founder with opinion 1. Similar to other group members the founder prefers to admit members closer to him. However, differently from "regular" members the founder can always veto the left candidate. This means 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 at 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 $(y_1 + y_2)/2 \leq q_{1-r}(S(k))$. As in this case, a fraction of at least r 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 $(y_1 + y_2)/2 < q_{1-r}(S(k))$.

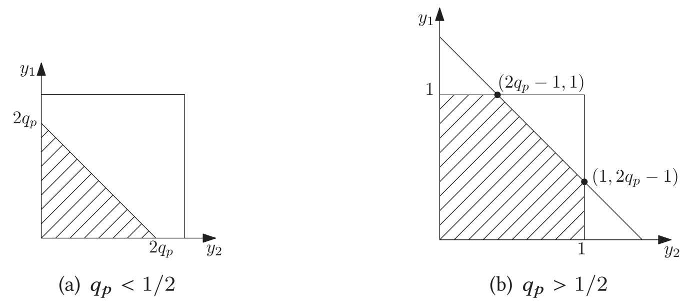

Fig. 2. The probability of accepting a candidate under veto rules: in both pictures the striped area includes all pairs of candidates (y_1, y_2) for which one of the candidates will be admitted to the group.

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 $1 - 2(1 - q_p)^2$. □

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 where $1 - r = p$.

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 $2q_p(S(k))^2$. This leads us to the following observation:

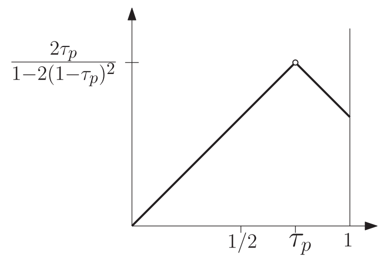

Fig. 3. A sketch of the density function that the group converges to for $r < 1/2$ ($p < 1/2$).

Observation 4.3. *For any group such that $q_p(S(k)) < 1/2$, 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 if $q_p(S(k)) < 1/2$ then 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 that 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 statement, which is that the $(p + \eta)$-quantile (for $\eta = \frac{1-2p}{8}$) is converging to 0. Formally, we prove the following theorem:

Theorem 4.4. *Consider a group $S(k_0)$. For any $\varepsilon > 0$, with probability $1 - e^{-\Omega(k_\varepsilon)}$, there exists k_ε, such that for any $k' > k_\varepsilon$, $q_p(S(k')) < \varepsilon$.*

Proof. Let $\eta = \frac{1-2p}{8}$ and let $\psi = \sqrt{\frac{1+2p+2\eta}{2(1-\eta)}}$, note that $\psi < 1$. The proof is composed of two main claims. First, in Claim 4.5 below, we consider increasing the group by adding ηk members and reason about the number of new members in the interval $[0, \psi \cdot q_{p+\eta}(S(k))]$. The main advantage of reasoning about additional ηk members is that we are guaranteed that in all those steps the p-quantile will always be left of $[0, q_{p+\eta}(S(k))]$. This makes reasoning about the acceptance probabilities considerably easier. Next, we clump together many of these ηk increments to increase the group's size by a factor of about $\frac{1}{2\eta}$. We show in Claim 4.6 that if we wait until the group's size is large enough then by multiple applications of Claim 4.5 with high probability the $(p + \eta)$-quantile (and hence the p-quantile) moves closer to 0 by a factor of ψ. □

Claim 4.5. *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 $[0, \psi \cdot q_{p+\eta}(S(k))]$ in the ηk steps is at least $(1 - \eta)\frac{\psi^2}{2} \cdot \eta k$.*

Proof. Note that by definition, we have that for every step k' of the ηk steps, $q_p(S(k')) \leq q_{p+\eta}(S(k))$.

- $q_p(S(k')) \leq 1/2$. In this case the probability to accept a candidate below $\psi \cdot q_{p+\eta}(S(k))$ is

$$\frac{(\psi \cdot q_{p+\eta}(S(k)))^2}{2(q_p(S(k')))^2} \geq \frac{(\psi \cdot q_{p+\eta}(S(k)))^2}{2(q_{p+\eta}(S(k)))^2} = \frac{\psi^2}{2}.$$

- $q_p(S(k')) > 1/2$. In this case the probability to accept a candidate below $\psi \cdot q_{p+\eta}(S(k))$ is

$$\frac{(\psi \cdot q_{p+\eta}(S(k)))^2}{1-2(1-(q_p(S(k')))^2} \geq \frac{(\psi \cdot q_{p+\eta}(S(k)))^2}{1-2(1-q_{p+\eta}(S(k)))^2} \geq \frac{\psi^2}{2}.$$

Hence, it is always the case that the probability to accept a candidate below $\psi \cdot q_{p+\eta}(S(k))$ is at least $\frac{\psi^2}{2}$. Thus, in expectation in the ηk steps at least $\frac{\psi^2}{2} \cdot \eta k$ candidates in the interval $[0, \psi \cdot q_{p+\eta}(S(k))]$ join the group. Let X be the number of candidates accepted below $\psi \cdot q_{p+\eta}(S(k))$. By taking a Chernoff bound, we get that with high probability X is at least $(1-\eta)\frac{\psi^2}{2} \cdot \eta k$:

$$Pr\left[X \leq (1-\eta) \cdot \frac{\psi^2}{2} \cdot \eta k\right] \leq e^{\frac{-\eta^3 \cdot \psi^2 \cdot k}{4}}. \qquad \square$$

We now reason about groups of growing sizes. 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 $j = \lceil \log_{1+\eta} \frac{1}{2\eta} \rceil$. We show that with high probability: $q_{p+\eta}(S(k_{i+j})) < \psi \cdot q_{p+\eta}(S(k_i))$.

Claim 4.6. *For $i > 1$, with probability* $(1 - \sum_{l=1}^{j-1} e^{-\Theta(\eta^3(1+\eta)^l k_i)})$, $q_{p+\eta}(S(k_{i+j})) < \psi \cdot q_{p+\eta}(S(k_i))$.

Proof. To prove the claim, we apply Claim 4.5 j times. First, we observe that with high probability for every i, $q_{p+\eta}(S(k_{i+1})) \leq q_{p+\eta}(S(k_i))$. The reason for this is that by Claim 4.5, we have that the number of members that joined in the ηk_i steps between k_i and k_{i+1} in the interval $[0, \psi \cdot q_{p+\eta}(S(k))]$ is at least

$$(1-\eta) \cdot \frac{\psi^2}{2} \cdot \eta k_i \geq (1-\eta) \cdot \frac{1+2p+2\eta}{4(1-\eta)} \cdot \eta k_i =^{(*)} \left(p + \frac{5}{2}\eta\right) \cdot \eta k_i > (p+\eta) \cdot \eta k_i.$$

The transition marked by $(*)$ uses that $1 = 8\eta + 2p$ by the definition of η. Now, since $\psi < 1$, we get that the number of members that joined $[0, q_{p+\eta}(S(k_i))]$ is also at least $(p+\eta) \cdot \eta k_i$, and hence we have that $q_{p+\eta}(S(k_{i+1})) \leq q_{p+\eta}(S(k_i))$. Therefore, the number of members that joined the interval $[0, \psi \cdot q_{p+\eta}(S(k_i))]$ between steps k_i and k_{i+j} is at least $(p + \frac{5}{2}\eta)(k_{i+j} - k_i)$. To complete the proof, we show that this number is greater than $(p+\eta)k_{i+j}$. To this end, observe that j was chosen such that $k_{i+j} = \frac{1}{2\eta}k_i + c$ for some $c \geq 0$. Thus, we have that

$$\left(p + \frac{5}{2}\eta\right)(k_{i+j} - k_i) = \left(p + \frac{5}{2}\eta\right)\left(\frac{1}{2\eta} - 1\right)k_i + \left(p + \frac{5}{2}\eta\right)c.$$

We now separately bound the coefficient of k_i:

$$\begin{aligned}\left(p + \frac{5}{2}\eta\right)\left(\frac{1}{2\eta} - 1\right) &= (p+\eta)\frac{1}{2\eta} + \frac{3}{2}\eta\left(\frac{1}{2\eta}\right) - \left(p + \frac{5}{2}\eta\right) \\ &= (p+\eta)\frac{1}{2\eta} + \frac{3}{4} - \left(p + \frac{5}{2}\eta\right) > (p+\eta)\frac{1}{2\eta}.\end{aligned}$$

Hence, we have that $q_{p+\eta}(S(k_{i+j})) < \psi \cdot q_{p+\eta}(S(k_i))$ as required. To compute the probability that the claim assertion holds, we can take a union bound on the bad events in Claim 4.5 and get that the claim holds with probability of at least $1 - \sum_{l=i}^{i+j-1} e^{-\Theta(\eta^3 k_l)} = 1 - \sum_{l=1}^{j-1} e^{-\Theta(\eta^3(1+\eta)^l k_i)}$. $\square$

To complete Theorem 4.4 proof, we can simply apply Claim 4.6 repeatedly and get that each time we increase the group by a factor of $(1+\eta)^j$ the distance of the $(p+\eta)$-quantile from 0 is decreasing by an extra factor of ψ. Hence, for any ε there exists k_ε such that $q_{p+\eta}(S(k_\varepsilon)) < \varepsilon$ and for any $k > k_\varepsilon$ it holds that $q_{p+\eta}(S(k_\varepsilon)) < \varepsilon$.

To compute the probability that the assertion of the theorem holds, we take a union bound over all the bad events and get that the assertion holds with probability at least

$$1 - \sum_{i=1}^{\infty} e^{-\Theta(\eta^3 k_i)} = 1 - \sum_{i=1}^{\infty} e^{-\Theta(\eta^3 (1+\eta)^{i-1} k_1)}.$$

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 $q > 1/2$, then 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 any candidate is accepted at all we have to divide $q_p(S(k))^2$ by the probability that a candidate is accepted, which is different for $q_p(S(k)) < 1/2$ and for $q_p(S(k)) > 1/2$. 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 $f(q_p(S(k)) = \frac{q_p(S(k)^2}{1-2(1-q_p(S(k))^2}$.

Fortunately, the admission process when $q_p(S(k)) > 1/2$ (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 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 4.7. *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 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

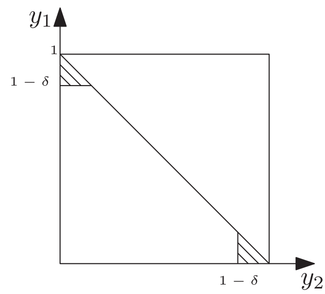

Fig. 4. The striped areas are the areas that if the point defined by the two candidates (y_1, y_2) is in one of them then a candidate in the interval $[1-\delta, 1]$ will join the group.

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 converges to $\tau_p = \frac{2p+\sqrt{2p^2-p}}{1+2p}$ with probability $1-o(1)$. In Proposition 4.8 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. □

Proposition 4.8. *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 4.4. We begin by showing that as we add ηk members to the group at most $(p-2\eta)\cdot\eta k$ members will join the interval $[0, q_{p-\eta}(S(k)) + \eta\cdot q_{p-\eta}(S(k))^2]$. □

Claim 4.9. *Consider adding ηk members to a group $S(k)$ such that $q_{p-\eta}(S(k)) < 1/2$. With probability $(1-e^{-\Theta(\eta^3 k)})$ the number of members that will join the interval $[0, q_{p-\eta}(S(k)) + \eta\cdot q_{p-\eta}(S(k))^2]$ in the ηk steps is 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 it 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 $[0, q_{p-\eta}(S(k)) + \eta\cdot q_{p-\eta}(S(k_i))^2]$ is at most $(p-2\eta)\cdot\eta k$. Denote by X the number of candidates accepted in the interval $[0, q_{p-\eta}(S(k)) + \eta \cdot q_{p-\eta}(S(k_i))^2]$, 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 4.10. *For $i > 1$, with probability $(1 - \sum_{l=1}^{j-1} e^{-\Theta(\eta^3 (1+\eta)^l k_i)})$, either $q_{p-\eta}(S(k_{i+j})) > q_{p-\eta}(S(k_i)) + \eta \cdot q_{p-\eta}(S(k_i))^2$ or there exists $k_i < k' < k_{i+j}$ such that $q_{p-\eta}(S(k')) > 1/2$.*

Proof. To prove the claim, we apply Claim 4.9 j times. Note that we can apply Claim 4.9 as long as $S(k_l) \leq 1/2$ for $i < l \leq i+j$. If for some l we have that $S(k_l) > 1/2$, then the claim holds. 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 4.9, 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 use 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 $1 - \sum_{l=i}^{i+j-1} e^{-\Theta(\eta^3 k_l)}$. □

Last, 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 4.11. *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 any step 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 $(1/2 + \eta)\eta k_i < (p-\eta)k + i$:

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

The proof of the proposition is completed by observing that we can apply Claim 4.10 until 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 4.11 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: $(x_1, \ldots, x_n)$ 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 $|x_j - y| \leq |x_j - x_i|$ for at least $\lceil (n-1)/2 \rceil + \ell$ members x_j such that $j \neq i$. 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: (1) the magnitude of drift of the committee (i.e., how far the committee can move from its initial configuration), and (2) 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 member 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.

Proposition 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 $x_1 < x_2 < \cdots < x_{2k+1}$. Let $M = x_{k+1}$ be the median and let ε be a small positive real satisfying, say, $\varepsilon k < x_{k+1} - x_k$ and $\varepsilon k \leq x_{k+2} - x_{k+1}$. Now, in step i ($1 \leq i \leq k$), replace x_i by $M - \varepsilon i$, and in step $k + i$ ($1 \leq i \leq k$), replace x_{k+1+i} by $M + \varepsilon i$. 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 as described in the observation above by repeatedly replacing x_1 by $x_n + 1$ this arithmetic progression can move arbitrarily far. □

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

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', then 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.*

Proof. We show that $x'_{k-\ell+2} \leq x_n + \frac{Dk}{2\ell-1}$. By symmetry the same argument implies that $x'_{k+\ell} \geq x_1 - \frac{Dk}{2\ell-1}$. The proof relies on the following Lemma: □

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

Before providing the proof of Lemma 5.3, we will use it to prove Theorem 5.2. 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 $j \in \{k-\ell+2, \ldots, k+1\}$, 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 $2\sum_{j=k-\ell+2}^{k} d(x_j, x'_j) + d(x_{k+1}, x'_{k+1}) \leq kD$.[11] Now, substitute $d(x_j, x'_j) \geq Dt$ for every $j \in \{k-\ell+2, \ldots, k+1\}$ 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 $x_1, \ldots, x_{2k+1}$ 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 $\delta = \Theta(\ell/k^2)$. 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 $\sum_{i\geq 0}(1-\delta)^i = 1/\delta$ 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.

Proof of Lemma 5.3. 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. In particular, this means that $x_{k+1} < y < x_{k+2}$. Let S' denote the sum of distances from the (new) median in x'. Since the distance between x_{k+1} and 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. □

We note that Theorem 5.2 also holds for committees of even size. If $n = 2k$, 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

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

of the right median. In both cases this distance is counted k times). We leverage this observation to show that Lemma 5.3 also holds for even committees. In a similar way to the odd case, if the left median moves right, then the sum of distances from (either one of) the medians decreases by at least $2\sum_{j=k-\ell+1}^{k} d(x'_j, x_j)$.

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

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 only members in $[x_1 - D, x_n + D]$ will ever join the committee.*

Proof. It suffices to show we never admit a member with opinion greater than $x_n + D$, as by symmetry the same argument implies we do not admit 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 $x_1 < x_2 \cdots < x_n$ and the configuration after a step in which y has been added be $x'_1 < x'_2 < \cdots < x'_n$. 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 $y < x_n$ and $y < x_2$, then $x'_1 = y, x'_2 = x_2, x'_n = x_n$, and the claim easily follows. 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 $x'_n \leq x_n$ $x'_1 = x_1$ and $x'_2 \leq x_2$ implying the desired result.
- If $y < x_1$, then $x'_1 = y \geq 2x_{n-1} - x_n$, $x'_2 = x_1$ 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, for any future configuration $(x''_1, x''_2, \ldots x''_n)$, we have that $x''_n \leq x''_n + x''_2 - x''_1 \leq x_n + D$, and the Proposition holds. □

We note that 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.

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.*

Proof. 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 first sketch

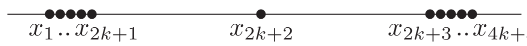

Fig. 5. An example of the configuration that the median has immunity if at least $3k + 3$ votes are required to remove a member.

the proof showing that the median of this committee has immunity and then present the formal proof. Observe that 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. Next, we formalize this intuition:

Suppose a majority of $3k + 2 + \ell$ is required, where $1 \leq \ell \leq k$. Pick some $d > 0$ and 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 $\frac{kd}{2\ell-1}$ 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 $x'_{k+\ell} \geq x_1 - \frac{kd}{2\ell-1}$ 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 $3k + \ell + 2$. □

It is interesting to note 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 instance, 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 $x_{2k+3} - kd$ the original median remains the median forever.

Finally, we show that the other direction also holds, that is:

Proposition 5.6. *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.*

Proof. 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. Let δ be a small real number satisfying $\delta < \frac{\min_j (x_j - x_{j-1})}{n}$ and δ divides $x_j - x_{j-1}$ for every j. We describe the process in stages. For every $j = 1, \ldots, 2k + 2$, stage j begins in a configuration where the smallest (i.e., left-most) j points form an arithmetic progression, distanced δ from each another, and ends in a configuration where x_j has been removed and the smallest $j + 1$ points form an arithmetic progression distanced δ from each another. This is done as follows.

Let $x_1, \ldots, x_j$ denote the smallest j points in stage j, and suppose these points form an arithmetic progression. The arithmetic progression progresses toward x_{j+1} as follows. Replace x_1 by $z_1 = x_j + \delta$, then x_2 by $z_2 = z_1 + \delta$, and so on, until the left-most $j + 1$ points form an arithmetic progression, distanced δ from one another. (Note that this process may repeat more than j steps.) In every replacement of x_j by z_j there are at least $3k + 2$ members who prefer z_j over x_j — these are the $2k + 1$ members to the right of the median (who are obviously closer to z_j than to x_j), the median itself, and at least k out of the $2k + 1$ members to the left of the median (by the fact that the smallest j points form an arithmetic progression). By the choice of δ all points including x_j have been replaced by the time the arithmetic progression reaches x_{j+1}. In stage $2k + 2$, we follow this process with $\delta' = \delta - \epsilon$ for some $\epsilon < \delta$ (to avoid tie breaking). In the end of stage $2k + 2$ of this process, the median has been replaced as well, as claimed.

Note that this process can be simplified in the stages that are strictly to the left of the median, and only as we get closer to the median, we need to be more careful. But for ease of presentation, we describe a uniform process.

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

6 CONCLUDING REMARKS

In this article, 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. In fact, after the preliminary version of our article was published, Feldheim and Feldheim [16] showed that when the two candidates in each step are independent random points in $[0, \infty)$ chosen according to any probability distribution, when applying the special veto rule when the founder is located at 0,[12] for any $r < 1$ the r-quantile will never drift away to infinity. This is in contrast to the case of distributions supported on a finite interval where our results show that, for example, with uniform distribution on $[0, 1]$ for $r > 1/2$ the quantile drifts to 1.

While all these extensions lead to interesting questions, we believe that the models we have considered in this article already shed light on real-life processes involving the dynamics of evolving groups; the present article provides a framework and tools for further exploration of this direction.

APPENDIX

A 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

[12]Recall that in the special veto rule, we studied the founder was located at1. When the founder is located at 0, choose the smaller candidate if at least r fraction of the existing ones prefer her, else choose nobody.

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 $1/4 = 1/2 - 1/4$, then 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 Reference [25].

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 $[0, O(1/\sqrt{t})] \cup [1 - O(1/\sqrt{t}), 1]$.

REFERENCES

[1] Toh-Kyeong Ahn, R. Isaac, and Timothy C. Salmon. 2008. Endogenous group formation. *J. Public Econ. Theory* 10, 2 (2008), 171–194.

[2] Toh-Kyeong Ahn, R. Mark Isaac, and Timothy C. Salmon. 2009. Coming and going: Experiments on endogenous group sizes for excludable public goods. *J. Public Econ.* 93, 1–2 (2009), 336–351.

[3] Robert Axelrod. 1997. The dissemination of culture. *J. Conflict Resolut.* 41, 2 (Apr. 1997), 203–226.

[4] Lars Backstrom, Dan Huttenlocher, Jon Kleinberg, and Xiangyang Lan. 2006. Group formation in large social networks: Membership, growth, and evolution. In *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 44–54.

[5] Eli Ben-Naim, Paul Krapivsky, and Sidney Redner. 2003. Bifurcations and patterns in compromise processes. *Physica D* 183 (2003).

[6] Kshipra Bhawalkar, Sreenivas Gollapudi, and Kamesh Munagala. 2013. Coevolutionary opinion formation games. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*. ACM, 41–50.

[7] David Bindel, Jon Kleinberg, and Sigal Oren. 2011. How bad is forming your own opinion? In *Proceedings of the 52nd IEEE Symposium on Foundations of Computer Science.*

[8] Claudio Castellano, Santo Fortunato, and Vittorio Loreto. 2009. Statistical physics of social dynamics. *Rev. Modern Phys.* 81 (2009), 591–646.

[9] Gary Charness and Chun-Lei Yang. 2014. Starting small toward voluntary formation of efficient large groups in public goods provision. *J. Econ. Behav. Organiz.* 102 (2014), 119–132.

[10] Flavio Chierichetti, Jon Kleinberg, and Sigal Oren. 2013. On discrete preferences and coordination. In *Proceedings of the 14th ACM Conference on Electronic Commerce.*

[11] Matthias Cinyabuguma, Talbot Page, and Louis Putterman. 2005. Cooperation under the threat of expulsion in a public goods experiment. *J. Public Econ.* 89, 8 (2005), 1421–1435.

[12] Pranav Dandekar, Ashish Goel, and David T. Lee. 2013. Biased assimilation, homophily, and the dynamics of polarization. *Proc. Natl. Acad. Sci. U.S.A.* 110, 15 (2013).

[13] Guillaume Deffuant, David Neau, Frederic Amblard, and Gérard Weisbuch. 2000. Mixing beliefs among interacting agents. *Adv. Complex Syst.* 3 (2000), 87–98.

[14] Morris H. DeGroot. 1974. Reaching a consensus. *J. Amer. Statist. Assoc.* 69 (1974), 118–121.

[15] Karl-Martin Ehrhart and Claudia Keser. 1999. Mobility and cooperation: On the run. CIRANO Working Papers, CIRANO. https://EconPapers.repec.org/RePEc:cir:cirwor:99s-24.

[16] Naomi Dvora Feldheim and Ohad Noy Feldheim. 2016. Mean and minimum of independent random variables. *arXiv preprint arXiv:1609.03004* (2016).

[17] Andreas Flache and Michael W. Macy. 2011. Local convergence and global diversity from interpersonal to social influence. *J. Conflict Resolut.* 55, 6 (2011), 970–995.

[18] Santo Fortunato. 2010. Community detection in graphs. *Phys. Rep.* 486, 3–5 (2010), 75–174.

[19] Noah E. Friedkin and Eugene C. Johnsen. 1990. Social influence and opinions. *J. Math. Sociol.* 15, 3–4 (1990), 193–205.

[20] Frederick C. Gamst. 1991. Foundations of social theory. *Anthropol. Work Rev.* 12, 3 (1991), 19–25.

[21] Rainer Hegselmann and Ulrich Krause. 2002. Opinion dynamics and bounded confidence: Models, analysis and simulation. *J. Artific. Societ. Soc. Simul.* 5, 3 (2002).

[22] David Kempe, Jon Kleinberg, Sigal Oren, and Aleksandrs Slivkins. 2013. Selection and influence in cultural dynamics. In *Proceedings of the 14th ACM Conference on Electronic Commerce*.

[23] Miller McPherson, Lynn Smith-Lovin, and James M. Cook. 2001. Birds of a feather: Homophily in social networks. *Annu. Rev. Sociol.* 27 (2001), 415–444.

[24] John E. Roemer. 2001. *Political Competition: Theory and Applications*. Harvard University Press.

[25] Nicholas C. Wormald. 1995. Differential equations for random processes and random graphs. *Ann. Appl. Probabil.* (1995), 1217–1235.

Received March 2017; revised March 2018; accepted October 2018