In Defense of Liquid Democracy

Daniel Halpern
Harvard University
dhalpern@g.harvard.edu

Joseph Y. Halpern*
Cornell University
halpern@cs.cornell.edu

Ali Jadbabaie†
MIT
jadbabai@mit.edu

Elchanan Mossel‡
MIT
elmos@mit.edu

Ariel D. Procaccia§
Harvard University
arielpro@seas.harvard.edu

Manon Revel¶
MIT
mrevel@mit.edu

Abstract

Liquid democracy is a voting paradigm that allows voters to choose between directly voting and transitively delegating their votes to other voters. While liquid democracy has been viewed as a system that can combine the best aspects of direct and representative democracy, it can also result in situations where few voters amass a large amount of influence. To analyze the impact of this shortcoming, we consider what has been called an *epistemic* setting, where voters decide on a binary issue for which there is a ground truth. Previous work has shown that under certain assumptions on the delegation model, the concentration of power is so severe that liquid democracy is less likely to identify the ground truth than direct voting. We examine different, arguably more realistic, classes of models, and prove they behave well by ensuring that (with high probability) there is a limit on concentration of power. Our proofs demonstrate that delegations can be treated as stochastic processes and that they can be compared to well-known processes from the literature — such as preferential attachment and multi-types branching process — that are sufficiently bounded for our purposes. Our results suggest that the concerns raised about liquid democracy can be overcome, thereby bolstering the case for this emerging paradigm.

*Supported in part by the NSF grants IIS-178108 and IIS-1703846, a grant from the Open Philanthropy Foundation, the ARO grant W911NF-17-1-0592, and the MURI grant W911NF-19-1-0217.

†Supported in part by the ARO MURI grant W911NF-19-1-0217; and the ONR 2016 Vannevar Bush Faculty Fellowship.

‡Supported in part by the 2020 ONR Vannevar Bush Faculty Fellowship, the NSF grant CCF-1918421, the ARO MURI grant W911NF-19-1-0217, and the Simons Investigator Award 622132.

§Supported in part by the NSF grants CCF-2007080, CCF-1733556 and IIS-2024287, and the ONR grant N00014-20-1-2488.

¶To whom correspondence may be addressed (mrevel@mit.edu). Supported in part by the Michael Hammer Fellowship and the ONR 2016 Vannevar Bush Faculty Fellowship.

1 Introduction

Liquid democracy is a voting paradigm that is conceptually situated between *direct democracy*, in which voters have direct influence over decisions, and *representative democracy*, where voters choose delegates who represent them for a period of time. Under liquid democracy, voters have a choice: they can either vote directly on an issue similar to direct democracy or delegate their vote to another voter, entrusting them to vote on their behalf. The defining feature of liquid democracy is that these delegations are *transitive*: if voter 1 delegates to voter 2 and voter 2 delegates to voter 3, then voter 3 votes (or delegates) on behalf of all three voters.

In recent years, liquid democracy has gained prominence around the world. The most impressive example is that of the German Pirate Party, which adopted the *LiquidFeedback* platform in 2010 [22]. Other political parties, such as the Net Party in Argentina and Flux in Australia, have run on the wily promise that once elected, their representatives would be essentially controlled by voters through a liquid democracy platform. Companies are also exploring the use of liquid democracy for corporate governance; Google, for example, has run a proof-of-concept experiment [17]. Practitioners, however, recognize that there is a potential flaw in liquid democracy, namely, the possibility of *concentration of power*, in the sense that certain voters amass an enormous number of delegations, giving them pivotal influence over the final decision. This scenario seems inherently undemocratic — and it is not a mere thought experiment. Indeed, in the LiquidFeedback platform of the German Pirate Party, a linguistics professor at the University of Bamberg received so many delegations that, as noted by Der Spiegel,[1] his "vote was like a decree."

Kahng et al. [21] examine liquid democracy's concentration-of-power phenomenon from a theoretical viewpoint and establish a troubling impossibility result in what has been called an *epistemic* setting, that is, one where there is a ground truth.[2] Informally, they demonstrate that, even under the strong assumption that voters only delegate to more "competent" voters, any "local mechanism" satisfying minimal conditions will, in certain instances, fall victim to a concentration of power, leading to relatively low accuracy. More specifically, Kahng et al. model the problem as a decision problem where voters decide on an issue with two outcomes, $\{0, 1\}$, where 1 is correct (the ground truth) and 0 is incorrect. Each of the voters $i \in \{1, \ldots, n\}$ is characterized by a *competence* $p_i \in [0, 1]$. The binary vote V_i of each voter i is drawn independently from a Bernoulli distribution, that is, each voter votes correctly with probability p_i. Under direct democracy, the outcome of the election is determined by a majority vote: the correct outcome is selected if and only if more than half vote for the correct outcome. Under liquid democracy, there exists a set of weights, weight_i for each $i \in [n]$, which represent the number of votes that voter i gathered transitively after delegation. If voter i delegates, then $\mathsf{weight}_i = 0$. The outcome of the election is then determined by a weighted majority; it is correct if and only if $\sum_{i=1}^{n} \mathsf{weight}_i V_i \geq n/2$. Kahng et al. also introduce the concept of a *delegation mechanism*, which determines whether voters delegate and, if so, to whom they delegate. They are especially interested in *local mechanisms*, where the delegation decision of a voter only depends on their local neighborhood according to an underlying social network. They assume that voters only delegate to those with strictly higher competence, which excludes the possibility of cyclic delegations. To evaluate liquid democracy, Kahng et al. [21] test the intuition that society makes more informed decisions under liquid democracy than under direct democracy

[1] http://www.spiegel.de/international/germany/liquid-democracy-web-platform-makes-professor-most-powerful-pirate-a-818683.html

[2] The use of the term "epistemic" in this context is well-established in the social choice literature [23, 28].

(especially given the foregoing assumption about upward delegation). To that end, they define the *gain* of a delegation mechanism to be the difference between the probability the correct outcome is selected under liquid democracy and the probability the correct outcome is selected under direct democracy. A delegation mechanism satisfies *positive gain* if its gain is strictly positive in some cases, and it satisfies *do no harm* if its loss (negative gain) is at most ε for *all* sufficiently large instances. The main result of Kahng et al. [21] is that local mechanisms can never satisfy these two requirements. Caragiannis and Micha [7] further strengthen this negative result by showing that there are instances where local mechanisms perform much worse than either direct democracy or dictatorship (the most extreme concentration of power).

These theoretical critiques undermine the case for liquid democracy: the benefits of delegation appear to be reversed by concentration of power. However, the negative conclusion relies heavily on modeling assumptions and has not been borne out by experiments [2]. In this paper, we provide a rebuttal by introducing an arguably more realistic model in which liquid democracy is able to avoid extreme concentration of power, thereby satisfying both do no harm and positive gain (for suitably defined extensions).

1.1 Our Contributions and Techniques

Our point of departure from the existing literature is the way we model delegation in liquid democracy. To emphasize these differences, instead of calling these delegation functions *mechanisms*, we instead call them delegation *models*, as they intend to capture independent voter behavior rather than prescribing to each voter to whom they must delegate. Our delegation models are defined by $M = (q, \varphi)$, where $q : [0, 1] \to [0, 1]$ is a function that maps a voter's competence to the probability they delegate and $\varphi : [0, 1]^2 \to \mathbb{R}_{\geq 0}$ maps a pair of competencies to a weight. In this model, each voter i votes directly with probability $1 - q(p_i)$ and, conditioned on delegating with probability $q(p_i)$, delegates to voter $j \neq i$ with probability proportional to $\varphi(p_i, p_j)$. Crucially, a voter does not need to "know" the competence of another voter to decide whether to delegate; rather, the delegation *probabilities* are merely influenced by competence in an abstract way captured by φ. Also, note that delegation cycles are possible, and we take a worst-case approach to deal with them: If the delegations form a cycle, then all voters in the cycle are assumed to be incorrect (vote 0).[3]

The most significant difference between our model of delegation and that of Kahng et al. [21] is that in our model, each voter has a chance of delegating to any other voter, whereas in their model, an underlying social network restricts delegation options. Our model captures a connected world where, in particular, voters may have heard of experts on various issues even if they do not know them personally. Although our model eschews an explicit social network, it can be seen as embedded into the delegation process, where the probability that i delegates to j takes into account the probability that i is familiar with j in the first place.

Another difference between our model and that of Kahng et al. [21] is that we model the competencies $p_1, \dots, p_n$ as being sampled independently from a distribution $\mathcal{D}$. While this assumption is made mainly for ease of exposition, it allows us to avoid edge cases and obtain robust results.

Our goal is to identify delegation models that satisfy (probabilistic versions of) positive gain and do no harm. Our first technical contribution, in Section 3, is the formulation of general conditions on the model and competence distribution that are sufficient for these properties to hold (Lemma 1). In particular, to achieve the more difficult do no harm property, we present conditions that guarantee

[3] In LiquidFeedback, delegation cycles are, in fact, ignored.

the maximum weight $\text{max-weight}(G_n)$ accumulated by any voter is sub-linear with high probability and that the expected increase in competence post-delegation is at least a positive constant times the population size. These conditions intuitively prevent extreme concentration of power and ensure that the representatives post-delegation are sufficiently better than the entire population to compensate for any concentration of power that does happen.

It then suffices to identify models and distribution classes that verify these conditions. A delegation model M and a competence distribution $\mathcal{D}$ induce a distribution over delegation instances that generates random graphs in ways that relate to well-known graph processes, which we leverage to analyze our models. Specifically, we introduce three models, all shown to satisfy do no harm and positive gain under *any* continuous distribution over competence levels. The first two models, *upward delegation* and *confidence-based* delegation, can be seen as interesting but somewhat restricted case studies, whereas the *general continuous delegation* model is, as the name suggests, quite general and arguably realistic. Despite the simplicity of the first two models, the three models, taken together, reveal the robustness of our approach.

Upward Delegation: In Section 4, we consider a model in which the probability of delegating p is exogenous and constant across competencies, and delegation only occur towards voters with strictly higher competence. That is, the probability that any voter i delegates is $q(p_i) = p$ and the weight that any voter i puts on another voter j is $\varphi(p_i, p_j) = \mathbb{I}_{\{p_j - p_i > 0\}}$. This model captures that there might be some reluctance to delegate regardless of the voter's competence but assumes that voters act in the interest of society only delegating to voters that are more competent then them.

To generate a random graph induced by such a model, one can add a single voter at a time in order of decreasing competence and allow the voter to either not delegate and create their own disconnected component, or delegate to the creator of any other component with probability proportional to p times the size of the component. This works because delegating to any voter in the previous components is possible (since they have strictly higher competence) and would result in the votes being concentrated in the originator of that component by transitivity. Such a process is exactly the one that generates a preferential attachment graph with a positive probability of not attaching to the existing components [30]. We can then show that, with high probability, no component grows too large, so long as $p < 1$. Further, there needs to be a constant improvement by continuity of the competence distribution, which ensures that a positive fraction of voters below a certain competence delegate to a positive fraction of voters with strictly higher competencies.

Confidence-Based Delegation: In Section 5, we consider a model in which voters delegate with probability decreasing in their competencies and choose someone at random when they delegate. That is, the probability $q(p_i)$ that any voter i delegates is decreasing in p_i and the weight that any voter i gives to any voter j is $\varphi(p_i, p_j) = 1$. In other words, in this model, competence does not affect the probability of receiving delegations, only the probability of delegating.

To generate a random graph induced by such a model, one can begin from a random vertex and study the delegation tree that starts at that vertex. A delegation tree is defined as a branching process, where a node i's "children" are the nodes that delegated to node i. In contrast to classical branching processes, the probability for a child to be born increases as the number of people who already received delegations decreases. Nevertheless, we prove that, with high probability, as long as a delegation tree is no larger than $O(\log n)$, our heterogeneous branching process is dominated by a sub-critical graph branching process [1]. We can then conclude that no component has size larger than $O(\log n)$ with high probability. Next, we show that the expected competence among the voters that do not delegate is strictly higher than the average one. Finally, given that no voter has weight larger than $O(\log n)$, we prove that a small number of voters end up in cycles with high probability. We can thus show that the conditions of Lemma 1 are satisfied.

General Continuous Delegation: Finally, we consider a general model in Section 6 where the likelihood of delegating is fixed and the weight assigned to each voter when delegating is increasing in their competence. That is, each voter i delegates with probability $q(p_i) = p$ and the weight that voter i places on voter j is $\varphi(p_i, p_j)$, where φ is continuous and increases in its second coordinate. Thus, in this model, the delegation distribution is slightly skewed towards more competent voters.

To generate a random graph induced by such a model, we again consider a branching process, but now voters j and k place different weights on i per φ. Therefore, voters have a *type* that governs their delegation behavior; this allows us to define a multi-type branching process with types that are continuous in $[0, 1]$. The major part of the analysis is a proof that, with high probability, as long as the delegation tree is no larger than $O(\log n)$, our heterogeneous branching process is dominated by a sub-critical Poisson multi-type branching process. To do so, we group the competencies into buckets that partition the segment $[0, 1]$ into small enough pieces. We define a new φ' that outputs, for any pair of competencies p_i, p_j, the maximum weight a voter from i's bucket could place on a voter from j's bucket. We can show that such a discrete multi-type branching process is sub-critical and conclude that no component has size larger than $O(\log n)$ with high probability. In a similar fashion to Confidence-Based Delegation, we also show that there is an expected increase in competence post-delegation.

1.2 Related work

Our work is most closely related to that of Kahng et al. [21], which was discussed in detail above. It is worth noting, though, that they complement their main negative result with a positive one: when the mechanism can restrict the maximum number of delegations (transitively) received by any voter to $o(\sqrt{\log n})$, do no harm and positive gain are satisfied. Imposing such a restriction would require a central planner that monitors and controls delegations. Gölz et al. [14] build on this idea: they study liquid democracy systems where voters may nominate multiple delegates and a central planner chooses a single delegate for each delegator in order to minimize the maximum weight of any voter.

Similarly, Brill and Talmon [6] propose allowing voters to specify ordinal preferences over delegation options and possibly restricting or modifying delegations in a centralized way. Caragiannis and Micha [7], and then Becker et al. [2] also consider central planners; they show that, for given competencies, the problem of choosing among delegation options to maximize the probability of a correct decision is hard to approximate. In any case, implementing these proposals would require a fundamental rethinking of the practice of liquid democracy. By contrast, our positive results show that *decentralized* delegation models are inherently self-regulatory, which supports the effectiveness of the current practice of liquid democracy.

More generally, there has been a significant amount of theoretical research on liquid democracy in recent years. To give a few examples: Green-Armytage [15] studies whether it is rational for voters to delegate their vote from a utilitarian viewpoint; Christoff and Grossi [8] examine a similar question but in the context of voting on logically interdependent propositions; Bloembergen et al. [3] and Zhang and Grossi [31] study liquid democracy from a game-theoretic viewpoint.

Further afield, liquid democracy is related to another paradigm called *proxy voting*, which dates back to the work of Miller [26]. Proxy voting allows voters to nominate representatives that have been previously declared. Cohensius et al. [10] study utilitarian voters that vote for the representative with the closest platform to theirs; they prove that the outcome of an election with proxy votes yields platforms closer to the median platform of the population than classical representative democracy. Their result provides a different viewpoint on the value of delegation.

2 Model

There is a set of n voters, denoted $[n] = \{1, \ldots, n\}$. We assume voters are making a decision on a binary issue and there is a correct alternative and an incorrect alternative. Each voter i has a *competence level* $p_i \in [0, 1]$ which is the probability that i votes correctly. We denote the vector of competencies by $\vec{p}_n = (p_1, \ldots, p_n)$. When n is clear from the context, we sometimes drop it from the notation.

Delegation graphs: A *delegation graph* $G_n = ([n], E)$ on n voters is a directed graph with voters as vertices and a directed edge $(i, j) \in E$ denoting that i delegates their vote to j. Again, if n is clear from context, we occasionally drop it from the notation. The outdegree of a vertex in the delegation graph is at most 1 since each voter can delegate to at most one person. Voters that do not delegate have no outgoing edges. In a delegation graph G_n, the *delegations received* by a voter i, $\mathsf{dels}_i(G_n)$, is defined as the total number of people that (transitively) delegated to i in G_n, (i.e., the total number of ancestors of i in G_n). The *weight* of a voter i, $\mathsf{weight}_i(G_n)$, is $\mathsf{dels}_i(G_n)$ if i delegates, and 0 otherwise. We define $\mathsf{max\text{-}weight}(G_n) = \max_{i \in [n]} \mathsf{weight}_i(G_n)$ to be the largest weight of any voter and define $\mathsf{total\text{-}weight}(G_n) = \sum_{i=1}^{n} \mathsf{weight}_i(G_n)$. Since each vote is counted at most once, we have that $\mathsf{total\text{-}weight}(G_n) \leq n$. However, note that if delegation edges form a cycle, then the weight of the voters on the cycle and voters delegating into the cycle are all set to 0 and hence will not be counted. In particular, this means that $\mathsf{total\text{-}weight}(G_n)$ may be strictly less than n.[4]

Delegation instances: We call the tuple $(\vec{p}_n, G_n)$ a *delegation instance*, or simply an instance, on n voters. Let $V_i = 1$ if voter i would vote correctly if i did vote, and $V_i = 0$ otherwise. Fixed competencies $\vec{p}_n$ induce a probability measure $\mathbb{P}_{\vec{p}_n}$ over the n possible binary votes V_i, where $V_i \sim \text{Bern}(p_i)$. Given votes $V_1, \ldots, V_n$, we let X_n^D be the number of correct votes under direct democracy, that is, $X_n^D = \sum_{i=1}^{n} V_i$. We let $X_{G_n}^F$ be the number of correct votes under liquid democracy with delegation graph G_n, that is, $X_{G_n}^F = \sum_{i=1}^{n} \mathsf{weight}_i(G_n) \cdot V_i$. The probability that direct democracy and liquid democracy are correct are $\mathbb{P}_{\vec{p}_n}[X_n^D > n/2]$ and $\mathbb{P}_{\vec{p}_n}[X_{G_n}^F > n/2]$, respectively.

Gain of a delegation instance We define the *gain* of an instance as

$$\text{gain}(\vec{p}_n, G_n) = \mathbb{P}_{\vec{p}_n}[X_{G_n}^F > n/2] - \mathbb{P}_{\vec{p}_n}[X_n^D > n/2].$$

In words, it is the difference between the probability that liquid democracy is correct and the probability that majority is correct.

[4]This is a worst-case approach where cycles can only hurt the performance of liquid democracy, since this assumption is equivalent to assuming that all voters on the cycles vote incorrectly.

Randomization over delegation instances: In general, we assume that both competencies and delegations are chosen randomly. Each voter's competence p_i is sampled i.i.d. from a fixed distribution $\mathcal{D}$ with support contained in $[0,1]$. Delegations will be chosen according to a *model* M. A model $M = (q, \varphi)$ is composed of two parts. The first $q : [0,1] \to [0,1]$ is a function that maps competencies to the probability that the voter delegates. The second $\varphi : [0,1]^2 \to \mathbb{R}_{\geq 0}$ maps pairs of competencies to a weight. A voter i with competence p_i will choose how to delegate as follows:

- With probability $1 - q(p_i)$ they do not delegate.
- With probability $q(p_i)$, i delegates; i places weight $\varphi(p_i, p_j)$ on each voter $j \neq i$ and randomly sample another voter j to delegate to proportional to these weights. In the degenerate case where $\varphi(p_i, p_j) = 0$ for all $j \neq i$, we assume that i does not delegate.

A competence distribution $\mathcal{D}$, a model M, and a number n of voters induce a probability measure $\mathbb{P}_{\mathcal{D},M,n}$ over all instances $(\vec{p}_n, G_n)$ of size n.

We can now redefine the *do no harm (DNH)* and *positive gain (PG)* properties from Kahng et al. [21] in a probabilistic way.

Definition 1 (Probabilistic do no harm)**.** A model M satisfies *probabilistic do no harm* with respect to a class $\mathfrak{D}$ of distributions if, for all distributions $\mathcal{D} \in \mathfrak{D}$ and all $\varepsilon, \delta > 0$, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$,

$$\mathbb{P}_{\mathcal{D},M,n}[\text{gain}(\vec{p}_n, G_n) \geq -\varepsilon] > 1 - \delta.$$

Definition 2 (Probabilistic positive gain)**.** A model M satisfies *probabilistic positive gain* with respect to a class $\mathfrak{D}$ of distributions if there exists a distribution $\mathcal{D} \in \mathfrak{D}$ such that for all $\varepsilon, \delta > 0$, there exists $n_0 \in N$ such that for all $n \geq n_0$,

$$\mathbb{P}_{\mathcal{D},M,n}[\text{gain}(\vec{p}_n, G_n) \geq 1 - \varepsilon] > 1 - \delta.$$

3 Core Lemma

In this section, we prove the following key lemma, which provides sufficient conditions for a model M to satisfy probabilistic do no harm and probabilistic positive gain with respect to a class $\mathfrak{D}$ of distributions. This lemma will form the basis of all of our later results.

Lemma 1. *If M is a model, $\mathfrak{D}$ a class of distributions, and for all distributions $\mathcal{D} \in \mathfrak{D}$, there is an $\alpha \in (0,1)$ and $C : \mathbb{N} \to \mathbb{N}$ with $C(n) \in o(n)$ such that*

$$\mathbb{P}_{\mathcal{D},M,n}\left[\mathsf{max\text{-}weight}(G_n) \leq C(n)\right] = 1 - o(1) \tag{1}$$

$$\mathbb{P}_{\mathcal{D},M,n}\left[\sum_{i=1}^{n} \mathsf{weight}_i(G_n) \cdot p_i - \sum_{i=1}^{n} p_i \geq 2\alpha n\right] = 1 - o(1), \tag{2}$$

then M satisfies probabilistic do no harm. If in addition, there exists a distribution $\mathcal{D} \in \mathfrak{D}$ and an $\alpha \in (0,1)$ such that

$$\mathbb{P}_{\mathcal{D},M,n}\left[\sum_{i=1}^{n} p_i + \alpha n \leq n/2 \leq \sum_{i=1}^{n} \mathsf{weight}_i(G_n) \cdot p_i - \alpha n\right] = 1 - o(1), \tag{3}$$

then M satisfies probabilistic positive gain.

p This condition guarantees that representatives post-delegation are sufficiently more competent than the entire population to compensate for any concentration of power that does occur. Finally, condition (3) ensures that there exists a distribution for which, with high probability, the average competence pre-delegation is at most $1/2$ minus a constant, while the average competence post-delegation is at least $1/2$ plus a constant. This condition suffices to guarantee that the probability that liquid democracy is correct goes to 1 while direct democracy goes to 0.

Throughout many of the proofs, we will make use of the following well-known concentration inequality [18]:

Lemma 2 (Hoeffding's Inequality). *Let $Z_1, \cdots, Z_n$ be independent, bounded random variables with $Z_i \in [a, b]$ for all i, where $-\infty < a \leq b < \infty$. Then, for all $t \geq 0$:*

$$\mathbb{P}\left[\frac{1}{n}\sum_{i=1}^{n} Z_i - \mathbb{E}[Z_i] \geq t\right] \leq \exp\left(-\frac{2nt^2}{(b-a)^2}\right), \text{ and } \mathbb{P}\left[\frac{1}{n}\sum_{i=1}^{n} Z_i - \mathbb{E}[Z_i] \leq -t\right] \leq \exp\left(-\frac{2nt^2}{(b-a)^2}\right)$$

Proof of Lemma 1. We establish the two properties separately.

Probabilistic do no harm: We first show that a model M that satisfies conditions (1) and (2) satisfies probabilistic do no harm. Fix an arbitrary competence distribution $\mathcal{D} \in \mathfrak{D}$ and let α and C be such that (1) and (2) are satisfied. Without loss of generality, suppose that $C(n) \leq n$ for all n, as replacing any larger values of $C(n)$ with n will not affect (1) (since $\textsf{max-weight}(G_n) \leq n$ for all graphs G_n on n vertices). Fix $\varepsilon, \delta > 0$. We must identify some n_0 such that for all $n \geq n_0$, $\mathbb{P}_{\mathcal{D},M,n}[\text{gain}(\vec{p}_n, G_n) \geq -\varepsilon] > 1 - \delta$.

We will begin by showing there exists $n_1 \in \mathbb{N}$ such that for all instances $(\vec{p}_n, G_n)$ on $n \geq n_1$ voters, if both

$$\textsf{max-weight}(G_n) \leq C(n) \text{ and} \tag{4}$$

$$\sum_{i=1}^{n} \textsf{weight}_i(G_n) \cdot p_i - \sum_{i=1}^{n} p_i \geq 2\alpha n, \tag{5}$$

then

$$\text{gain}(\vec{p}_n, G_n) \geq -\varepsilon. \tag{6}$$

Since (4) and (5) each hold with probability $1 - o(1)$ by (1) and (2), for sufficiently large n, say $n \geq n_2$, they will each occur with probability at least $1 - \delta/2$. Hence, by a union bound, for all $n \geq n_2$, they both occur with probability at least $1 - \delta$. By taking $n_0 = \max(n_1, n_2)$, this implies that probabilistic do no harm is satisfied.

We now prove that, for sufficiently large n, (4) and (5) imply (6). First, we will show that

$$\text{gain}(\vec{p}_n, G_n) \geq -\mathbb{P}_{\vec{p}_n}[X_n^D > X_{G_n}^F]. \tag{7}$$

Indeed, we have that

$$\begin{aligned}\mathbb{P}_{\vec{p}_n}[X_n^D > n/2] &= \mathbb{P}_{\vec{p}_n}[X_n^D > n/2, X_{G_n}^F > n/2] + \mathbb{P}_{\vec{p}_n}[X_n^D > n/2, X_{G_n}^F \leq n/2] \\ &\leq \mathbb{P}_{\vec{p}_n}[X_{G_n}^F > n/2] + \mathbb{P}_{\vec{p}_n}[X_n^D > X_{G_n}^F]\end{aligned}$$

where the first transition holds by the law of total probability, and the second because the corresponding events are contained in each other. That is,

$$\left\{X_n^D > n/2, X_{G_n}^F > n/2\right\} \subseteq \left\{X_{G_n}^F > n/2\right\} \text{ and } \left\{X_n^D > n/2, X_{G_n}^F \leq n/2\right\} \subseteq \left\{X_n^D > X_{G_n}^F\right\}.$$

Re-arranging the terms above yields (7).

Hence, for our purpose, it suffices to show that (4) and (5) imply $\mathbb{P}_{\vec{p}_n}\left[X_n^D > X_{G_n}^F\right] \leq \varepsilon$. Intuitively, we will use (5) to show the expected value of X_n^D is well below the expected value of $X_{G_n}^F$. Then we will show both X_n^D and $X_{G_n}^F$ concentrate well around their means, where for the latter we will need (4). Together, these observations imply that $X_{G_n}^F > X_n^D$ with high probability.

Fix an instance $(\vec{p}_n, G_n)$ on n voters satisfying (4) and (5). We will show that for n large enough,

$$\mathbb{P}_{\vec{p}_n}\left[X_n^D < \sum_{i=1}^{n} p_i + \alpha n\right] > 1 - \varepsilon/2 \tag{8}$$

and

$$\mathbb{P}_{\vec{p}_n}\left[X_{G_n}^F > \sum_{i=1}^{n} \mathsf{weight}_i(G_n) \cdot p_i - \alpha n\right] > 1 - \varepsilon/2. \tag{9}$$

Note that since (5) holds for this instance, $\sum_{i=1}^n p_i + \alpha n \leq \sum_{i=1}^n \mathsf{weight}_i(G_n) \cdot p_i - \alpha n$. Therefore, when both events whose probability is considered in (8) and (9) hold, $X_n^D \leq X_n^F$. Hence,

$$\mathbb{P}_{\vec{p}_n}[X_n^D \leq X_{G_n}^F] \geq \mathbb{P}_{\vec{p}_n}\left[X_n^D < \sum_{i=1}^{n} p_i + \alpha n, X_{G_n}^F > \sum_{i=1}^{n} \mathsf{weight}_i(G_n) \cdot p_i - \alpha n\right] > 1 - \varepsilon$$

where the last inequality holds by a union bound. This implies that $\mathbb{P}_{\vec{p}_n}[X_n^D \leq X_{G_n}^F] < \varepsilon$, as needed.

It remains to be shown that (8) and (9) hold for sufficiently large n. For (8), this follows directly from Hoeffding's inequality (Lemma 2). To prove (9), first note that, as shown in Kahng et al. [21],

$$\mathrm{Var}_{\vec{p}_n}\left[X_{G_n}^F\right] = \sum_{i=1}^{n} \mathsf{weight}_i(G_n)^2 p_i(1-p_i) \leq \frac{1}{4}\sum_{i=1}^{n} \mathsf{weight}_i(G_n)^2 \leq \frac{1}{4}\sum_{i=1}^{\lceil n/C(n)\rceil} C(n)^2 < nC(n) \in o(n^2),$$

where the first inequality holds because $p(1-p)$ is upper bounded by $1/4$, the second because $\sum_{i=1}^n \mathsf{weight}_i(G_n) \leq n$ with each $\mathsf{weight}_i(G_n) \leq C(n)$ so the value is maximized by setting as many terms to $C(n)$ as possible, and the final inequality holds because $C(n) \leq n$.

Hence, by Chebyshev's inequality,

$$\mathbb{P}_{\vec{p}_n}\left[X_{G_n}^F \leq \mathbb{E}_{\vec{p}_n}\left[X_{G_n}^F\right] - \alpha n\right] \leq \frac{\mathrm{Var}_{\vec{p}_n}[X_{G_n}^F]}{(\alpha n)^2}.$$

This bound is $o(1)$ because the numerator is $o(n^2)$ and the denominators is $\Omega(n^2)$. This implies that for sufficiently large n, it will be strictly less than $\varepsilon/2$, so (9) holds.

Probabilistic positive gain: Fix a distribution $\mathcal{D} \in \mathfrak{D}$ and an $\alpha \in (0,1)$ such that (3) holds. We want to show that M satisfies probabilistic positive gain. Since $\mathcal{D} \in \mathfrak{D}$, it also satisfies (1) for some C. We show below that there exists an n_3 such that all instances $(\vec{p}_n, G_n)$ with $n \geq n_3$ voters satisfying (4) for which $\sum_{i=1}^n p_i + \alpha n \leq n/2 \leq \sum_{i=1}^n \mathsf{weight}_i(G_n) \cdot p_i - \alpha n$, we have that $\mathrm{gain}(\vec{p}_n, G_n) \geq 1 - \varepsilon$. As with the DNH part of the proof, since the events of (1) and (3) each hold with probability $1 - o(1)$, for sufficiently large n, say $n \geq n_4$, they each occur with probability at least $1 - \delta/2$. Hence, by a union bound, for all $n \geq n_4$, they both occur with probability $1 - \delta$. For $n_0 = \max(n_3, n_4)$, probabilistic positive gain is satisfied.

It remains to show that that if (1) and (3) hold for a specific instance $(\vec{p}_n, G_n)$, then $\text{gain}(\vec{p}_n, G_n) \geq 1-\varepsilon$ for sufficiently large n. Since $\mathcal{D} \in \mathfrak{D}$, (8) and (9) are both satisfied for sufficiently large n. When

$$\sum_{i=1}^{n} p_i + \alpha n \leq n/2 \leq \sum_{i=1}^{n} p_i - \mathsf{weight}_i(G_n) \cdot \alpha n$$

is satisfied as well, we get that $\mathbb{P}_{\vec{p}_n}\left[X_n^D > n/2\right] < \varepsilon/2$ and $\mathbb{P}_{\vec{p}_n}\left[X_{G_n}^L > n/2\right] > 1-\varepsilon/2$, so $\text{gain}(\vec{p}_n, G_n) > 1-\varepsilon$ is immediate. $\square$

In the following sections, we investigate natural delegation models and identify conditions such that the models satisfy probabilistic do no harm and probabilistic positive gain. In all instances, we will invoke Lemma 1 after showing that its sufficient conditions are satisfied.

4 Strictly Upward Delegation Model

We now turn to the analysis of a simple model that assumes that voters either do not delegate with fixed exogenous probability or delegate to voters that have a competence greater than their own.

Formally, for a fixed $p \in [0,1]$ we let $M_p^U = (q, \varphi)$ be the model consisting of $q(p_i) = p$ for all $p_i \in [0,1]$, and $\varphi(p_i, p_j) = \mathbb{I}_{\{p_j > p_i\}}$ for all $i, j \in [n]$. That is, voter i delegates with fixed probability p and puts equal weight on all the more competent voters. In other words, if voter i delegates, then i does so to a more competent voter chosen uniformly at random. Note that a voter with maximal competence will place 0 weight on all other voters, and hence is guaranteed not to delegate. We refer to M_p^U as the *Upward Delegation Model* parameterized by p.

Theorem 1 (Upward Delegation Model)**.** *For all $p \in (0,1)$, M_p^U satisfies probabilistic do no harm and probabilistic positive gain with respect to the class $\mathfrak{D}^C$ of all continuous distributions.*

Proof. To prove the theorem, we will prove that the Upward Delegation Model with respect to $\mathfrak{D}^C$ satisfies (1) and (2), which implies that the model satisfies probabilistic do no harm by Lemma 1. Later, we demonstrate a continuous distribution that satisfies (3), implying the model satisfies probabilistic positive gain.

Upward Delegation satisfies (1)*:* We show there exists $C(n) \in o(n)$ such that the maximum weight $\mathsf{max\text{-}weight}(G_n) \leq C(n)$ with high probability—that is, such that (1) holds. Fix some sampled competencies $\vec{p}_n$. Recall that each entry p_i in $\vec{p}_n$ is sampled i.i.d. from $\mathcal{D}$, a continuous distribution. Hence, almost surely, no two competencies are equal. From now on, we condition on this probability 1 event. Now consider sampling the delegation graph G_n. By the design of the model M_p^U, we can consider a random process for generating G_n that is isomorphic to sampling according to $\mathbb{P}_{D,M,n}$ as follows: first, order the competencies $p_{(1)} > p_{(2)} > \cdots > p_{(n)}$ (note that such strict order is possible by our assumption that all competencies are different) and rename the voters such that voter i has competence $p_{(i)}$; then construct G_n iteratively by adding the voters one at a time in decreasing order of competencies, voter 1 at time 1, voter 2 at time 2, and so on.

We start with the voter with the highest competence, voter 1. By the choice of φ, voter 1 places weight 0 on every other voter and hence by definition does not delegate. These voters form the first component in the graph G_n, which we call $C^{(1)}$. Then, we add voter 2 who either delegates to voter 1 joining component $C^{(1)}$ with probability p, or starts a new component $C^{(2)}$ with probability $1-p$. Next, we add voter 3. If $2 \in C^{(1)}$ (that is, if 2 delegated to 1), 3 either delegates to 1 (either directly or through 2 by transitivity) with probability p or she starts a new component $C^{(2)}$. If $2 \in C^{(2)}$, then 3 either delegates to 1 with probability $p/2$ and is added to $C^{(1)}$, or delegates to 2 with probability $p/2$ and is added to $C^{(2)}$, or starts a new component $C^{(3)}$. In general, at time t, if there are k existing components $C^{(1)}, \dots, C^{(k)}$, voter t either joins each component $C(j)$ with probability $\frac{p|C(j)|}{t-1}$ or starts a new component with probability $1-p$. To construct G_n, we run this process for n steps.

This is precisely the model introduced by Simon [30]. It has been studied under the name *infinite Polya urn process* [9] and is considered a generalization of the *preferential attachment model* (with a positive probability of not attaching to the existing graph).

Let $U_t^{(k)}$ be the size of the kth component, $C^{(k)}$, at time t. In general, our approach will be to show that each component $C^{(k)}$ remains below some $o(n)$ function by time n with high enough probability so that we can union bound over all possible $k \le n$ (there can never be more than n components in the graph). That is, we will show

$$\mathbb{P}_{\mathcal{D},M_p^U,n}[\max(U_n^{(1)}, \dots, U_n^{(n)}) > C(n)] \le \sum_{i=1}^{n} \mathbb{P}_{\mathcal{D},M_p^U,n}[U_n^{(k)} > C(n)] = o(1)$$

for some $C(n) \in o(n)$ to be chosen later. Hence, it will be useful to consider this process more formally from the perspective of the kth component, $C^{(k)}$. The kth component $C^{(k)}$ is "born" at some time $t \ge k$ when the kth person chooses to not delegate, at which point $U_t^{(k)} = 1$ (prior to this, $U_t^{(k)} = 0$). More specifically, the first component is guaranteed to be born at time $t = 1$ and for all other $k > 1$, it will be born at time $t \ge k$ with probability $\binom{t-1}{k-1}(1-p)^k p^{t-k}$, although these exact probabilities will be unimportant for our analysis. Once born, we have the following recurrence on $U_t^{(k)}$ describing the probability $C^{(k)}$ will be chosen at time t:

$$U_t^{(k)} = \begin{cases} U_{t-1}^{(k)} + 1 & \text{with probability } \frac{p \cdot U_{t-1}^{(k)}}{t-1} \\ U_{t-1}^{(k)} & \text{with probability } 1 - \frac{p \cdot U_{t-1}^{(k)}}{t-1}. \end{cases}$$

Let $W_t^{(k)}$ be the process for the size of component that is born at time k. That is, $W_k^{(k)} = 1$, and for $k > t$, $W_t^{(k)}$ follows the exact same recurrence as $U_t^{(k)}$. Note that since the kth component $C^{(k)}$ can only be born at time k or later, we have that $W_n^{(k)}$ stochastically dominates $U_n^{(k)}$ for all k. Hence, it suffices to show that

$$\sum_{i=1}^{n} \mathbb{P}_{\mathcal{D},M_p^U,n}[W_n^{(k)} > C(n)] = o(1). \tag{10}$$

Choose γ to be a constant such that $3/4 < \gamma < 1$ (say $\gamma = 7/8$); note that $p + (1-p)\gamma < p + (1-p) = 1$. Choose another constant δ such that $p + (1-p)\gamma < \delta < 1$. This additionally implies $3/4 < \gamma < \delta$. Finally, choose $C(n) = n^\delta$. We show that the probability that any component is of size greater than n^δ by time n (when the delegation process completes) approaches 0. That is, we show that $\mathbb{P}_{\mathcal{D},M_p^U,n}[\max(U_n^{(1)}, \dots, U_n^{(n)}) > n^\delta] = o(1)$ by showing that (10) holds for $C(n) = n^\delta$.

We split our analysis into two parts: the first consider the first n^γ components, while the second considers the last $n - n^\gamma$ components.

We first show that $\sum_{k=1}^{n^\gamma} \mathbb{P}[W_n^k > n^\delta] = o(1)$. Recall that $W_k^{(k)} = 1$ and we have the following recurrence for all $t > k$:

$$W_t^{(k)} = \begin{cases} W_{t-1}^{(k)} + 1 & \text{with probability } \frac{p \cdot W_{t-1}^{(k)}}{t-1} \\ W_{t-1}^{(k)} & \text{with probability } 1 - \frac{p \cdot W_{t-1}^{(k)}}{t-1}. \end{cases}$$

Our first goal is to show that the expectation of $W_n^{(k)}$ is upper bounded by

$$\mathbb{E}[W_n^{(k)}] \leq \frac{\Gamma(n+p)\Gamma(k)}{\Gamma(p+k)\Gamma(n)} \tag{11}$$

for all $k \leq n$, where Γ represents the Gamma function. By the tower property of expectation, for all $t \geq k+1$,

$$\begin{aligned} \mathbb{E}[W_t^{(k)}] &= \mathbb{E}[\mathbb{E}[W_t^{(k)} \mid W_{t-1}^{(k)}]] \\ &= \mathbb{E}[W_{t-1}^{(k)}(1 + \frac{p}{t-1})] \\ &= \mathbb{E}[W_{t-1}^{(k)}](1 + \frac{p}{t-1}). \end{aligned}$$

Thus,

$$\begin{aligned} \mathbb{E}[W_t^{(k)}] &= \mathbb{E}[W_{t-1}^{(k)}](1 + \frac{p}{t-1}) = \mathbb{E}[W_{t-2}^{(k)}](1 + \frac{p}{t-1})(1 + \frac{p}{t-2}) = \mathbb{E}[W_k^{(k)}] \prod_{i=k}^{t-1} (1 + \frac{p}{i}) \\ &= \prod_{i=k}^{t-1} (1 + \frac{p}{i}) = \prod_{i=k}^{t-1} (i+p)/i = \frac{(k-1)!}{(t-1)!} \frac{\prod_{i=0}^{t-1}(i+p)}{\prod_{i=0}^{k-1}(i+p)} = \frac{(k-1)!}{(t-1)!} \frac{\frac{\Gamma(p+t)}{\Gamma(p)}}{\frac{\Gamma(p+k)}{\Gamma(p)}} \\ &= \frac{\Gamma(t+p)\Gamma(k)}{\Gamma(p+k)\Gamma(t)}, \end{aligned}$$

where the first four equalities follow from the recursive formula for $\mathbb{E}[U_t^{(k)}]$, the fifth because $\mathbb{E}[W_k^{(k)}] = 1$, the sixth and seventh by rearranging terms, the eighth uses the fact that $\Gamma(x+1) = x\Gamma(x)$ for all $x \in \mathbb{R}$, and the last uses the fact that $\Gamma(n) = (n-1)!$ for all $n \in \mathbb{N}$. This proves (11).

We can now use Markov's inequality to show that for all k,

$$\mathbb{P}_{\mathcal{D}, M_p^U, n}\left[W_n^{(k)} > n^\delta\right] \leq \frac{\mathbb{E}[W_n^{(k)}]}{n^\delta} \leq \frac{1}{n^\delta} \cdot \frac{\Gamma(n+p)}{\Gamma(n)} \cdot \frac{\Gamma(k)}{\Gamma(k+p)}.$$

Hence,

$$\sum_{k=1}^{n^\gamma} \mathbb{P}_{\mathcal{D}, M_p^U, n}[W_n^{(k)} > n^\delta] \leq \sum_{k=1}^{n^\gamma} \frac{1}{n^\delta} \cdot \frac{\Gamma(n+p)}{\Gamma(n)} \cdot \frac{\Gamma(k)}{\Gamma(k+p)} = \frac{1}{n^\delta} \cdot \frac{\Gamma(n+p)}{\Gamma(n)} \cdot \sum_{k=1}^{n^\gamma} \frac{\Gamma(k)}{\Gamma(k+p)}.$$

What remains to be shown is that

$$\frac{1}{n^\delta} \cdot \frac{\Gamma(n+p)}{\Gamma(n)} \cdot \sum_{k=1}^{n^\gamma} \frac{\Gamma(k)}{\Gamma(k+p)} = o(1).$$

To do this, first note that $\frac{\Gamma(p+t)}{\Gamma(t)} = \Theta(t^p)$. Indeed, the fact that

$$(t+p-1)^p \leq \frac{\Gamma(p+t)}{\Gamma(t)} \leq (t+p)^p$$

follows from Gautschi's inequality [13], and both the upper and lower bounds are $\Theta(t^p)$. Because n^γ is an increasing function of n, we have that

$$\frac{1}{n^\delta} \cdot \frac{\Gamma(n+p)}{\Gamma(n)} \cdot \sum_{k=1}^{n^\gamma} \frac{\Gamma(k)}{\Gamma(k+p)} = \frac{1}{n^\delta} \cdot \Theta(n^p) \cdot \Theta\left(\sum_{k=1}^{n^\gamma} \frac{1}{k^p}\right). \tag{12}$$

Further,

$$\sum_{k=1}^{n^\gamma} \frac{1}{k^p} = \Theta\left(\int_1^{n^\gamma} \frac{1}{x^p}\, dx\right) = \Theta(n^{\gamma\cdot(1-p)}).$$

Hence, the left-hand side of (12) is $\Theta(n^{-\delta+p+\gamma\cdot(1-p)})$. By our choice of δ, $\delta > p + \gamma \cdot (1-p)$, so this implies that it is is $o(1)$, as desired.

Now consider the final $n - n^\gamma$ components. We will prove that $\mathbb{P}_{\mathcal{D},M_p^U,n}[W_n^{(n^\gamma+1)} > n^\delta] = o(1/n)$. Since $W_n^{(k)}$ stochastically dominates $W_n^{(k')}$ for all $k' \geq k$, this implies that $\mathbb{P}_{\mathcal{D},M_p^U,n}[W_n^{(k)} > n^\delta] = o(1/n)$ for all $k \geq n^\gamma +$. Hence,

$$\sum_{k=n^\gamma+1}^{n} \mathbb{P}_{\mathcal{D},M_p^U,n}\left[W_n^{(k)} > n^\delta\right] = o(1).$$

To do this, we compare the $W_t^{(n^\gamma+1)}$ process to another process, V_t. We define $V_0 = 1$, and for $t > 0$, take V_t to satisfy the following recurrence:

$$V_t = \begin{cases} V_{t-1} + 1 & \text{with probability } \frac{V_{t-1}}{t+n^\gamma} \\ V_{t-1} & \text{with probability } 1 - \frac{V_{t-1}}{t+n^\gamma}. \end{cases}$$

This is identical to the W recurrence with t shifted down by $n^\gamma + 1$ except without the p factor. Hence, $V_{n-n^\gamma+1}$ clearly stochastically dominates $W_n^{(n^\gamma+1)}$. For convenience in calculation, we will instead focus on bounding V_n which itself stochastically dominates $V_{n-n^\gamma+1}$.

Next, note that the V_t process is isomorphic to the following classic Polya's urn process. We begin with two urns, one with a single ball and the other with n^γ balls. At each time, a new ball is added to one of the two bins with probability proportional to the bin size. The process V_t is isomorphic to the size of the one-ball urn after t steps. Classic results tell us that for fixed starting bin sizes a and b, as the number of steps grows large, the possible proportion of balls in the a-bin follows a $\mathrm{Beta}(a, b)$ distribution [25, 11, 29, 20, 24].

The mean and variance of such a Beta distribution would be sufficient to prove our necessary concentration bounds; however, for us, we need results after exactly $n - n^\gamma$ steps, not simply in the limit. Hence, we will be additionally concerned with the speed of convergence to this Beta distribution.

Let $X_n = \frac{V_n}{n}$ and $Z_n \sim \mathrm{Beta}(1, n^\gamma)$. From Janson [19], we know that the rate of convergence is such that, for any $p \geq 1$

$$\ell_p(X_n, Z_n) = \Theta(1/n) \tag{13}$$

where ℓ_p is the *minimal* L_p *metric*, defined as

$$\ell_p(X,Y) = \inf\left\{\mathbb{E}[|X' - Y'|^p]^{1/p} \mid X' \stackrel{d}{=} X, Y' \stackrel{d}{=} Y\right\},$$

which can be thought of as the minimal L_p norm over all possible couplings between X and Y. For our purposes, the only fact about the ℓ_p metric we will need is that $\ell_p(X,0) = \mathbb{E}[|X|^p]^{1/p}$ where 0 is the identically 0 random variable. Since ℓ_p is in fact a metric, the triangle inequality tells us that $\ell_p(0, X_n) \le \ell_p(0, Z_n) + \ell_p(Z_n, X_n)$, so, combining with (13), we have that

$$\mathbb{E}[|X_n|^p]^{1/p} \le \mathbb{E}[|Z_n|^p]^{1/p} + \Theta(1/n) \tag{14}$$

for all $p \ge 1$.

Note that since $Z_n \sim \text{Beta}(1, n^\gamma)$,

$$\mathbb{E}[Z_n] = \frac{1}{1+n^\gamma} = \Theta(n^{-\gamma})$$

and

$$\text{Var}[Z_n] = \frac{n^\gamma}{(2+n^\gamma)(1+n^\gamma)^2} = \Theta(n^{-2\gamma}).$$

Given these results, we are ready to prove that V_n is smaller than n^δ with probability $1 - o(1/n)$. Precisely, we want to show that $\mathbb{P}_{\mathcal{D},M_p^U,n}[X_n \ge n^{\delta-1}] = o(1)$. By Chebyshev's inequality,

$$\mathbb{P}_{\mathcal{D},M_p^U,n}[X_n \ge n^{\delta-1}] \le \frac{\text{Var}[X_n]}{(n^{\delta-1} - \mathbb{E}[X_n])^2}.$$

Inequality (14) with $p = 1$ along with the fact that X_n and Z_n are always nonnegative implies that $\mathbb{E}[X_n] \le \mathbb{E}[Z_n] + \Theta(1/n) = O(n^{-\gamma})$. Hence, $n^{\delta-1} - \mathbb{E}[X_n] = \Omega(n^{\delta-1})$ since $\delta - 1 > -1/2 > -\gamma$. We can therefore write:

$$\left(n^{\delta-1} - \mathbb{E}[X_n]\right)^2 = \Omega(n^{-2(\delta-1)}). \tag{15}$$

Inequality (14) with $p = 2$ implies that $\sqrt{\mathbb{E}[X_n^2]} \le \sqrt{\mathbb{E}[Z_n^2]} + \Theta(1/n)$. Hence,

$$\begin{aligned}
\mathbb{E}[X_n^2] &\le (\Theta(1/n) + \sqrt{\mathbb{E}[Z_n^2]})^2 \\
&\le (\Theta(1/n) + \sqrt{\mathbb{E}[Z_n]^2 + \text{Var}[Z_n]})^2 \\
&\le (\Theta(1/n) + \sqrt{\Theta(n^{-2\gamma})})^2 \\
&= (\Theta(1/n) + \Theta(n^{-\gamma}))^2 \\
&= \Theta(n^{-\gamma})^2 \\
&= \Theta(n^{-2\gamma}).
\end{aligned}$$

Next, note that $\text{Var}[X_n] \le \mathbb{E}[X_n^2]$, so

$$\text{Var}[X_n] = O(n^{-2\gamma}) \tag{16}$$

as well. Combining (15) and (16), we have that

$$\mathbb{P}_{\mathcal{D},M_p^U,n}[X_n \ge n^{\delta-1}] \le \frac{\text{Var}[X_n]}{(n^{\delta-1} - \mathbb{E}[X_n])^2} = O\left(n^{-2\gamma+2(1-\delta)}\right).$$

Since $-2\gamma + 2(1-\delta) < 1$, given our assumption that $3/4 < \gamma < \delta$, it follows that $\mathbb{P}_{\mathcal{D},M_p^U,n}[X_n \geq n^{\delta-1}] = o(1/n)$, which allows us to conclude that

$$\sum_{k=n^\gamma+1}^{n} \mathbb{P}_{\mathcal{D},M_p^U,n}[W_n^{(k)} > n^\delta] = o(1).$$

Since we showed earlier that $\sum_{k=1}^{n^\gamma} \mathbb{P}_{\mathcal{D},M_p^U,n}[W_n^{(k)} > n^\delta] = o(1)$, we have that

$$\sum_{k=1}^{n} \mathbb{P}_{\mathcal{D},M_p^U,n}[W_n^{(k)} > n^\delta] = o(1),$$

as needed. It follows that M_p^U satisfies (1).

Upward Delegation satisfies (2)*:* We will show there exists $\alpha \in (0,1)$ such that $\sum_{i=1}^n \mathsf{weight}_i(G_n) \cdot p_i - \sum_{i=1}^n p_i \geq 2\alpha n$ with high probability, so (2) is satisfied. Note that in the present scheme, cycles are impossible, so do need to worry about ignored voters.

Since $\mathcal{D}$ is a continuous distribution, there exists $a < b$ such that $\pi_a := \mathcal{D}[\{p : p < a\}] > 0$ and $\pi_b := \mathcal{D}[\{p : p > b\}] > 0$. Let $N_{a,n}(\vec{p}_n)$ be the number of voters in $\vec{p}_n$ with competence $p_i < a$ and $N_{b,n}(\vec{p}_n)$ be the number of voters with competence $p_i > b$. When we sample competencies, since each is chosen independently, $N_{a,n} \sim \text{Bin}(n, \pi_a)$ and $N_{b,n} \sim \text{Bin}(n, \pi_b)$. By Hoeffding's inequality (Lemma 2) and the union bound, with probability $1 - o(1)$, there will be at least $\pi_a/2 \cdot n$ voters with competence $p_i < a$ and $\pi_b/2 \cdot n$ voters with competence $p_i > b$. Indeed,

$$\begin{aligned} \mathcal{D}^n[N_{a,n} > \frac{n\pi_a}{2}, N_{b,n} > \frac{n\pi_b}{2}] &= 1 - \mathcal{D}^n[\{N_{a,n} \leq \frac{n\pi_a}{2}\} \cup \{N_{b,n} \leq \frac{n\pi_b}{2}\}] \\ &\geq 1 - (\mathcal{D}^n[N_{a,n} \leq \frac{n\pi_a}{2}] + \mathcal{D}^n[N_{b,n} \leq \frac{n\pi_b}{2}]) \\ &\geq 1 - \exp(-\frac{n\pi_a^2}{2}) - \exp(-\frac{n\pi_b^2}{2}), \end{aligned} \tag{17}$$

where the first line comes from De Morgan's law, the second from the union bound, and the last from Heoffding's inequality (Lemma 2).

Conditioned on this occurring, each voter with competence $p_i < a$ has probability at least $p\pi_b/2$ of delegating to a voter with competence at least b. As they each decide to do this independently, the number $N_{ab,n}$ of n voters deciding to do this stochastically dominates a random variable following the $\text{Bin}(\pi_a/2 \cdot n, p \cdot \pi_b/2)$ distribution. We can again apply Hoeffding's inequality to conclude that with probability $1 - o(1)$, at least $\pi_a \cdot \pi_b \cdot p/8 \cdot n$ voters do so. Indeed,

$$\begin{aligned} \mathcal{D}[N_{ab,n} > \frac{np\pi_a\pi_b}{8} \mid N_{a,n} > \frac{n\pi_a}{2}, N_{b,n} > \frac{n\pi_b}{2}] &\geq \mathcal{D}[\text{Bin}(\frac{n\pi_a}{2}, \frac{p\pi_b}{2}) > \frac{np\pi_a\pi_b}{8}] \\ &\geq 1 - \exp(-\frac{np^2\pi_a\pi_b^2}{4}), \end{aligned} \tag{18}$$

where the first inequality holds because $N_{ab,n}$ stochastically dominates the corresponding binomial random variable and the second holds by Hoeffding's inequality. Finally, using (17) and (18), we have

$$\begin{aligned} \mathcal{D}[N_{ab,n} > \frac{np\pi_a\pi_b}{8}] \geq \mathcal{D}[N_{ab,n} > \frac{np\pi_a\pi_b}{8} \mid N_{a,n} > \frac{n\pi_a}{2}, N_{b,n} > \frac{n\pi_b}{2}] \\ \cdot \mathcal{D}[N_{a,n} > \frac{n\pi_a}{2}, N_{b,n} > \frac{n\pi_b}{2}] \\ \geq 1 - o(1). \end{aligned}$$

Under these upward delegation models, delegations can only increase the total competence of all voters. Hence,

$$\sum_{i=1}^{n} \mathsf{dels}_i(G_n) \cdot p_i - \sum_{i=1}^{n} p_i \geq (b-a)N_{ab.n}.$$

Each of these $\pi_a \cdot \pi_b \cdot p/8 \cdot n$ voters results in a competence increase of at least $b-a$. Hence, under these high probability events, the total competence increase is at least $(b-a)\cdot\pi_a\cdot\pi_b\cdot p/8\cdot n$. Indeed, since $\mathcal{D}[N_{ab,n} > \frac{np\pi_a\pi_b}{8}] = 1-o(1)$, this implies $\mathcal{D}[\sum_{i=1}^{n} \mathsf{dels}_i(G_n)\cdot p_i - \sum_{i=1}^{n} p_i > \frac{np\pi_a\pi_b}{8}] = 1-o(1)$. By choosing $\alpha = \frac{p\pi_a\pi_b}{8}(b-a)$, we see that there is an $\alpha \cdot n$ increase in competence with high probability, as needed.

We have proved that for any continuous distribution $\mathcal{D}$, and for well-behaved realizations of $\vec{p}_n$ which occur with high probability, the graph generated from the random delegation process yields an increase in the expected sum of the votes of at least $\alpha \cdot n$. We can then conclude that M_p^U satisfies Equation (2) with respect to the class of continuous distributions.

Upward Delegation satisfies (3)*:* We now show that there exists a distribution $\mathcal{D}$ such that $\sum_{i=1}^{n} p_i + \alpha n \leq n/2 \leq \sum_{i=1}^{n} \mathsf{weight}_i(G_n)\cdot p_i - \alpha n$ with probability $1-o(1)$ for some $\alpha > 0$. This implies that the model satisfies probabilistic positive gain by Lemma 1, and will conclude the proof.

We take $\mathcal{D}$ to be $\mathcal{D}_\eta$, the uniform distribution $\mathcal{U}[0, 1-2\eta]$ for some small $0 < \eta < p/512$. Let $\alpha = \eta/2$. Clearly, $\mu_{\mathcal{D}_\eta}$, the mean of D_η, is $1/2-\eta$. Since each $p_i \overset{i.i.d.}{\sim} \mathcal{D}_\eta$, the p_is are bounded independent random variables with mean $1/2-\eta$, so Hoeffding's inequality directly implies that $\sum_{i=1}^{n} p_i \leq n/2 - n\eta/2 = n/2 - n\alpha$ with high probability.

Now consider $\mathcal{E}_F$, the event consisting of instances $(\vec{p}_n, G_n)$ such that $\sum_{i=1}^{n} \mathsf{weight}_i(G_n)\cdot p_i \geq n/2 + n\alpha$. We denote by $\mathcal{E}_D$ the event that $\sum_{i=1}^{n} p_i \geq n/2 - 3n\eta/2$. The same reasoning as before implies that $\Pr_{\mathcal{D}_\eta, M_p^U, n}(\mathcal{E}_D) = 1-o(1)$.

Let $a = 1/4 - \eta/2$ and $b = 1/2-\eta$, so we have that $\pi_a := \mathcal{D}_\eta[p_i < a] = 1/4$ and $\pi_b := \mathcal{D}_\eta[p_i > b] = 1/2$. We proved in the preceding derivation that $\sum_{i=1}^{n} \mathsf{weight}_i(G_n)\cdot p_i - \sum_{i=1}^{n} p_i > \frac{np\pi_1\pi_b}{8} = \frac{np}{64}(1-\eta)$ with high probability. Hence, if both this and $\mathcal{E}_D$ occur, which is the case with high probability, by the union bound, it follows that $\sum_{i=1}^{n} \mathsf{weight}_i(G_n)\cdot p_i > n/2 + n(\frac{p}{128}(1-\eta) - 3\eta/2)$ with high probability.

Since $\eta < \frac{p}{512} < 1/2$, we have that

$$\frac{p}{128}(1-\eta) - 3\eta/2 > \frac{p}{256} - 3\eta/2 > 2\eta - 3\eta/2 = \eta/2 = \alpha,$$

and we can conclude that $\mathcal{E}_F$ occurs with high probability. Hence, M_p^U satisfies Equation (3). $\square$

5 Confidence-Based Delegation Model

We now explore a model according to which voters delegate with probability that is strictly decreasing in their competence and when they do decide to delegate, they do so by picking a voter uniformly at random. This models the case where voters do not need to know anything about their peers' competencies, but do have some sense of their own competence, and delegate accordingly.

Formally, for any q, let $M_q^C = (q, \varphi^1)$ where $\varphi^1(p_i, p_j) = 1$ for all $i, j \in [n]$. Voter i puts equal weight on all the voters and hence samples one uniformly at random when they delegate. We refer to M_q^C as the Confidence-Based Delegation Model.

Theorem 2 (Confidence-Based Delegation Model)**.** *All models M_q^C with monotonically decreasing q satisfy probabilistic do no harm and probabilistic positive gain with respect to the class $\mathfrak{D}^C$ of all continuous distributions.*

Proof. We show that the Confidence-Based Model satisfies all of (1), (2), and (3).

Confidence-Based Delegation satisfies (1)*:* Fix some distribution $\mathcal{D} \in \mathfrak{D}^C$. We show there exists $C(n) \in O(\log n)$ such that (1) holds.

Note that when sampling an instance $(\vec{p}_n, G_n)$, the probability an arbitrary voter i chooses to delegate is precisely $p := \mathbb{E}_{\mathcal{D}}[q]$. To see this, consider how a voter i chooses whether to delegate: they first sample a competence $p_i \sim \mathcal{D}$ and then sample whether or not to delegate from $\mathrm{Bern}(q(p_i))$. Treating this as a single process, it is clear that the overall probability of choosing to delegate is exactly $\mathbb{E}_{\mathcal{D}}[q]$ by integrating out the competence.

Further, since $\mathcal{D}$ is continuous and q is monotonically decreasing, $p \in (0, 1)$. When a voter does decide to delegate, they do so by picking another voter uniformly at random. Hence, we can consider the marginal distribution of delegation graphs directly (ignoring the competencies). We will show that when sampling a delegation graph, for any specific voter i, with probability $1 - o(1/n)$, $\mathsf{dels}_i(G_n) \leq C(n)$, which implies $\mathsf{weight}_i(G_n) \leq C(n)$. A union bound over all n voters implies $\mathsf{max\text{-}weight}(G_n) \leq C(n)$ with probability $1 - o(1)$.

To that end, we will describe a branching process similar to the well-known *graph branching process* [1], which has the property that the distribution of its size exactly matches the distribution of $\mathsf{dels}_i(G_n)$ for an arbitrary voter i. We will compare this process to a known graph branching process that has size at most $O(\log n)$ with high probability. We will show our process is sufficiently dominated such that it too has size at most $O(\log n)$ with high probability. The branching process works as follows. Fix our voter i. We sample which other voters end up in i's "delegation tree" (i.e., its ancestors in G_n) dynamically over a sequence of time steps. As is standard for these processes, all voters V will be one of three types, live, dead, or neutral. Dead voters are those whose "children" (i.e., voters who delegate to them) we have already sampled. Live voters are voters who have decided to delegate, but whose children have not yet been sampled. Neutral voters are still in the "pool" and have yet to commit to a delegation. At time zero, i is a live voter, there are no dead voters, and all other voters $V \setminus \{i\}$ are neutral. At each time step, we take some live voter j, sample which of the neutral voters choose to delegate to j, add these voters as live vertices, and update j as dead. The procedure ends when there are no more live vertices, at which point the number of delegations received by i is simply the total number of dead vertices.

Let us now describe this more formally. Following the notation of Alon and Spencer [1], let Z_t denote the number of voters we sample to delegate at time t. Let Y_t be the number of live vertices at time t; we have that $Y_0 = 1$. At time t, we remove one live vertex and add Z_t more, so we have the recursion $Y_t = Y_{t-1} - 1 + Z_t$. We let N_t be the number of neutral vertices at time t. We have that $N_0 = n - 1$, and $N_t = N_{t-1} - Z_t$. Note that after t time steps, there are t dead vertices and Y_t live ones, so this is equivalent to $N_t = n - 1 - t - Y_t$. To sample Z_t, we fix some live voter j and ask how many of the neutral voters chose to delegate to j, conditioned on them not delegating to any of the dead voters. Note that when sampling at this step, there are $t - 1$ dead voters and conditioned on the neutral voters not delegating to the dead ones, the probability they delegate to any of the other $n - t$ individuals (not including themselves) is exactly $\frac{p}{n-t}$, equally split between them for a total delegation probability of p. Hence $Z_t \sim \mathrm{Bin}(N_{t-1}, \frac{p}{n-t}) \sim \mathrm{Bin}(n - t - Y_{t-1}, \frac{p}{n-t})$. We denote by $\mathfrak{X}_{n,p}^D$ the random variable that counts the size of this branching process, i.e., the number of time steps until there are no more live vertices. Note that the number of delegations received by any voter has the same distribution as $\mathfrak{X}^D_{n,p}$.

Choose some constant p' such that $p < p' < 1$. We will be comparing the $\mathfrak{X}^D_{n,p}$ to a graph branching process $\mathfrak{X}^G_{n,p'}$. The graph branching process is nearly identical, except the probability each of the neutral vertex joins our component is independent of the number of dead vertices and is simply $\frac{p'}{n}$. In other words, $Z_t \sim \mathrm{Bin}(N_{t-1}, \frac{p'}{n})$. A key result about this branching process is the probability of seeing a component of a certain size ℓ decreases exponentially with ℓ. In other words, there is some constant c such that

$$\mathbb{P}_{\mathcal{D},M_q^C,n}[\mathfrak{X}^G_{n,p'} \leq c\log(n)] = 1 - o(1/n).$$

Take $C(n) = c\log(n)$. Note that as long as t is such that $\frac{p}{n-t} \leq \frac{p'}{n}$, the sampling in the delegation branching process is dominated by the sampling in this graph branching process. Hence, as long as $\frac{p}{n-C(n)} \leq \frac{p'}{n}$, $\mathbb{P}[\mathfrak{X}^D_{n,p} \leq c\log(n)] \geq \mathbb{P}[\mathfrak{X}^G_{n,p'} \leq c\log(n)]$. Since $C(n) \in O(\log n)$, this is true for sufficiently large n, so for such n, $\mathbb{P}[\mathfrak{X}^D_{n,p} \leq c\log(n)] = 1 - o(1/n)$. By a union bound over all n voters, this implies the desired result.

Confidence Based Delegation satisfies (2)*:* Let $\bar{q}$ be such that $\bar{q}(x) = 1 - q(x)$, so $\bar{q}$ represents the probability someone with competence x does not delegate and let $q^+(x) = \bar{q}(x)x$. Let $\mu_{\mathcal{D}}$ the mean of the competence distribution $\mathcal{D}$. We first show that

$$\frac{\mathbb{E}_{\mathcal{D}}[q^+]}{\mathbb{E}_{\mathcal{D}}[\bar{q}]} > \mu_{\mathcal{D}}.$$

Indeed, since both x and $\bar{q}(x)$ are strictly increasing functions of x, the Fortuin–Kasteleyn–Ginibre (FKG) inequality [12] tells us that $\mathbb{E}_{\mathcal{D}}[q^+] > \mathbb{E}_{\mathcal{D}}[\bar{q}] \cdot \mathbb{E}_{\mathcal{D}}[x] = \mathbb{E}_{\mathcal{D}}[\bar{q}] \cdot \mu_{\mathcal{D}}$. This implies that the expected competence conditioned on not delegating is strictly higher than the overall expected competence.

Let $\mu^* = \frac{\mathbb{E}_{\mathcal{D}}[q^+]}{\mathbb{E}_{\mathcal{D}}[\bar{q}]}$ be this expected competence. We will show that for any constant $\gamma > 0$, with high probability, both $\sum_{i=1}^n p_i \leq (\mu + \gamma)n$ and $\sum_{i=1}^n \mathsf{weight}_i(G)p_i \geq (\mu^* - \gamma)n$. If we choose $\gamma = (\mu^* - \mu)/3$ and $\alpha = \gamma/2$, it follows that, with high probability,

$$\sum_{i=1}^n \mathsf{weight}_i(G)p_i - \sum_{i=1}^n p_i \geq 2\alpha n,$$

implying that (2) is satisfied.

Since the p_is are bounded independent variables, it follows directly from Heoffding's inequality that $\sum_{i=1}^n p_i \leq n(\mu + \gamma)$ with high probability, so we now focus on showing $\sum_{i=1}^n \mathsf{weight}_i(G) \cdot p_i \geq (\mu^* - \gamma)n$ with high probability. To do this, we will first show that, with high probability, the delegation graph G satisfies $\mathsf{dels}_i(G) \leq C(n)$ for all i and $\mathsf{total\text{-}weight}(G) \geq n - C(n)\log^2 n$.

We showed in the earlier part of this proof that $\mathsf{dels}_i(G) \leq C(n)$ with high probability. We will now prove that $\mathbb{P}_{\mathcal{D},M_q^C,n}[\mathsf{total\text{-}weight}(G) \geq n - C(n)\log^2 n \mid \mathsf{dels}_i(G) \leq C(n)] = 1 - o(1)$. To do this, we will first bound the number of voters that, with high probability, end up in cycles. Fix a voter i and sample i's delegation tree. Voter i will only end up in a cycle if i chooses to delegate to someone in this delegation tree. Since we are conditioning on $\mathsf{dels}_i(G) \leq C(n)$, the maximum size of this tree is $C(n)$. Hence, the total φ weight that voter i places on someone in the tree is at most $C(n)$, while the total weight they place on all voters is $n - 1$. Hence, the probability that i

delegates to someone in their tree can be at most $p \cdot C(n)/(n-1)$. Since this is true for each voter i, the expected number of voters in cycles is at most $np\frac{C(n)}{(n-1)} \in O(\log n)$. By Markov's inequality, the probability that more than $\log^2 n$ voters are in cycles is at most $np\frac{C(n)}{(n-1)\log^2 n} = O(1/\log n) = o(1)$.

Next, since we have conditioned on $\mathsf{dels}_i(G) \leq C(n)$, no single voter, and in particular no single voter in a cycle, can receive more than $C(n)$ delegations. So conditioned on the high probability event that there are at most $\log^2 n$ voters in cycles, there are at most $C(n)\log^2 n$ voters that delegate to those in cycles. This implies that $\mathsf{total\text{-}weight}(G) \geq n - C(n)\log^2 n + \log^2 n$ with high probability.

We now show that, conditioned on the graph satisfying these properties, the instance $(\vec{p}, G)$ satisfies $\sum_{i=1}^n \mathsf{weight}_i(G) \cdot p_i \geq n(\mu^* - \gamma)$ with high probability. Note that the competencies satisfy that those that don't delegate are drawn i.i.d. from the distribution of competencies conditioned on not delegating, which has mean μ^*. Fix an arbitrary graph G satisfying the properties. Suppose M is the set of voters that do not delegate. Note that for each $i \in M$, $\mathsf{weight}_i(G) \leq C(n)$, by assumption. Further $\sum_{i\in M} \mathsf{weight}_i(G) \geq n - C(n)\log^2(n)$. Hence, when we sample the non-delegator p_is, $\mathbb{E}[\sum_{i\in M} \mathsf{weight}_i(G) \cdot p_i] \geq (n - C(n)\log^2(n)) \cdot \mu^*$. Moreover,

$$\mathrm{Var}[\sum_{i\in M} \mathsf{weight}_i(G) \cdot p_i] \leq \sum_{i\in M} \mathsf{weight}_i(G)^2 \leq C(n) \cdot n.$$

This follows from the fact that $\mathrm{Var}[p_i] \leq 1$ and that we have fixed the graph G and hence $\mathsf{weight}_i(G)$ for each i, so these terms can all be viewed as constants. In addition, we know that, for each voter i, $\mathsf{weight}_i(G) \leq C(n)$, and $\sum_{i=1}^n \mathsf{weight}_i(G) \leq n$. Hence, we can directly apply Chebyshev's inequality:

$$\begin{aligned}
\mathbb{P}_{\mathcal{D}, M_q^C, n}\left[\sum_{i\in M} \mathsf{weight}_i(G) p_i < n(\mu^* - \gamma)\right] &< \frac{\mathrm{Var}[\sum_{i\in M} \mathsf{weight}_i(G) p_i]}{(\mathbb{E}[\sum_{i\in M} \mathsf{weight}_i(G) p_i] - n(\mu^* - \gamma))^2} \\
&\leq \frac{nC(n)}{(\gamma n - C(n)\log^2(n)\mu^*)^2} \\
&\in o(1),
\end{aligned}$$

where the final step holds because the numerator is $o(n^2)$ and the denominator is $\Omega(n^2)$. Hence, $\sum_{i\in M} \mathsf{weight}_i(G) p_i \geq n(\mu^* - \gamma)$ with high probability, as needed.

To summarize, we have proved that, conditioned on $\mathsf{dels}_i(G) \leq C(n)$ for all i and $\mathsf{total\text{-}weight}(G) \geq n - C(n)\log^2 n$, $\sum_{i=1}^n \mathsf{weight}_i(G) \cdot p_i \geq n(\mu^* - \gamma/3)$ occurs with high probability. Given that, conditioned on $\mathsf{dels}_i(G) \leq C(n), \mathsf{total\text{-}weight}(G) \geq n - C(n)\log^2 n$ occurs with high probability and that $\mathsf{dels}_i(G) \leq C(n)$ occurs with high probability, we can conclude by the chain rule that the intersection of these events hold with high probability. Given that the probability of any of this event is greater than the probability of the intersection, we can conclude that $\sum_{i=1}^n \mathsf{weight}_i(G) \cdot p_i \geq n(\mu^* - \gamma/3)$ occurs with probability $1 - o(1)$, as desired.

Confidence-Based Delegation satisfies (3)*:* We finally show there exists a distribution $\mathcal{D}$ such that $\sum_{i=1}^n p_i + \alpha n \leq n/2 \leq \sum_{i=1}^n \mathsf{weight}_i(G_n) \cdot p_i - \alpha n$ with probability $1 - o(1)$. This implies that the model M_q^C satisfies probabilistic positive gain by Lemma 1.

Using the notation of the analogous proof in Section 4, let $\mathcal{D}_\eta = \mathcal{U}[0, 1 - 2\eta]$ for $\eta \in [0, 1/2)$. As a function of η, $\frac{\mathbb{E}_{\mathcal{D}_\eta}[q^+]}{\mathbb{E}_{\mathcal{D}_\eta}[\bar{q}]}$, the expected competence conditioned on not delegating, is continuous. Moreover, if $\eta = 0$, $\frac{\mathbb{E}_{\mathcal{D}_0}[q^+]}{\mathbb{E}_{\mathcal{D}_0}[\bar{q}]} > \mu_{\mathcal{D}_0} = 1/2$. Hence, for small enough $\eta > 0$, $\frac{\mathbb{E}_{\mathcal{D}_\eta}[q^+]}{\mathbb{E}_{\mathcal{D}_\eta}[\bar{q}]} > 1/2 > \mu_{\mathcal{D}_\eta}$.

We choose $\mathcal{D}_\eta$ to be our distribution for this choice of η. As in the previous section, let $\mu^*_{\mathcal{D}_\eta} = \frac{\mathbb{E}_{\mathcal{D}_\eta}[q^+]}{\mathbb{E}[\bar{q}]}$. Note that $\mu_{\mathcal{D}_\eta} = 1/2 - \eta$. Let $\gamma = \min(\frac{1/2-\mu_{\mathcal{D}_\eta}}{2}, \frac{\mu^*_{\mathcal{D}_\eta}-1/2}{2})$ and $\alpha = \gamma$. By the earlier argument for (2), we have that with high probability $\sum_{i=1}^n p_i \leq n(\mu+\gamma) \leq n/2 - \alpha n$, and $\sum_{i=1}^n \mathsf{weight}_i(G) p_i \geq n(\mu^* - \gamma) \geq n/2 + \alpha n$.

By the union bound, we have that both occur simultaneously with high probability, so (3) holds. $\square$

6 Continuous General Delegation Model

Finally, we study a model in which voters delegate with fixed probability, and they do so by picking a voter according to a continuous increasing delegation function. This is a general model in which delegations can either go to more or less competent neighbors but where more competent voters are more likely to be chosen over less competent ones.

Formally, let $M^S_{p,\varphi} = (q^p, \varphi)$ where q^p is a constant function equal to p, that is, $q^p(x) = p$ for all $x \in [0,1]$, and $\varphi(x,y)$ is non-zero, continuous, and increasing in y. We then have the following.

Theorem 3 (Continuous General Delegation Model)**.** *All models $M^S_{p,\varphi}$ with $p \in (0,1)$ and φ that is non-zero, continuous, and increasing in its second coordinate satisfy probabilistic do no harm and probabilistic positive gain with respect to the class $\mathfrak{D}^C$ of all continuous distributions.*

The proof of this theorem is relegated to Appendix A. It is similar in structure to the proof of Theorem 2, but requires much more intricate analysis to account for different "types" of voters resulting in distinct delegating behavior depending on competence.

7 Discussion

This paper is, to the best of our knowledge, the first to study (decentralized) liquid democracy models in which concentration of power is unlikely to occur. While we focused on a setup without an underlying social network (i.e., there is no restriction on whom a voter may delegate to), we can extend all of our results to a model where a directed social network is first sampled, and then a (q, φ)-model is followed. The social network must be sampled such that each voter's neighbors are chosen uniformly at random, although the number of such neighbors could follow any small-tailed distribution. Intuitively, delegation proportional to weighting i's neighbors (rather than the entire population) can be shown to be "equivalent" to a possibly different weighting over the entire population. (This extension does not carry over well to undirected networks, since if voters have a small number of neighbors, we would expect many 2-cycles to form after delegation, which, under the worst-case cycle approach, may not be canceled out by the overall increase in competence.)

Our paper further relies on a set of assumptions and modeling choices that are worth discussing.

- First, the assumption that there exists a ground-truth alternative seems ill-suited for some voting scenarios. We think of this assumption as viable in the legislative process, where one option can be objectively superior to another with respect to a concrete metric, even if that metric is not always apparent beforehand. For example, an economic policy can be evaluated with respect to the metric of maximizing gross domestic product in five years. However, some decisions are inherently subjective.

- Second, like previous papers [21, 7, 2], we assume that voters vote independently. Admittedly, this is not a realistic assumption; relaxing it, as it was relaxed for the classic Condorcet Jury Theorem [16, 27], is a natural direction for future work.

- Third, again, like much of the previous work on liquid democracy, our models do not take strategic considerations into account. It would be interesting to bridge our models and those that do capture game-theoretic issues in liquid democracy [3, 31].

More generally, our work aims to provide a better understanding of a prominent shortcoming of liquid democracy, namely the concentration of power. But there are others. For example, any voter can see the complete delegation graph under current liquid democracy systems — a feature that helps voters make informed delegation decisions (because one's vote can be transitively delegated). This may lead to voter coercion, however, and the tradeoff between transparency and security is poorly understood. Nevertheless, there are many reasons to be excited about the potential of liquid democracy [4]. We believe that our results provide another such reason and hope that our techniques will be useful in continuing to build the theoretical understanding of this compelling paradigm.

References

[1] Noga Alon and Joel H. Spencer. 2016. *The Probablistic Method*. John Wiley & Sons.

[2] Ruben Becker, Gianlorenzo D'Angelo, Esmaeil Delfaraz, and Hugo Gilbert. 2021. When Can Liquid Democracy Unveil the Truth? arXiv:2104.01828.

[3] Daan Bloembergen, Davide Grossi, and Martin Lackner. 2019. On rational delegations in liquid democracy. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI)*. 1796–1803.

[4] Christian Blum and Christina Isabel Zuber. 2016. Liquid democracy: Potentials, problems, and perspectives. *Journal of Political Philosophy* 24, 2 (2016), 162–182.

[5] Béla Bollobás, Svante Janson, and Oliver Riordan. 2007. The phase transition in inhomogeneous random graphs. *Random Structures & Algorithms* 31, 1 (2007), 3–122.

[6] Markus Brill and Nimrod Talmon. 2018. Pairwise Liquid Democracy.. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI)*. 137–143.

[7] Ioannis Caragiannis and Evi Micha. 2019. A Contribution to the Critique of Liquid Democracy. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI)*. 116–122.

[8] Zoé Christoff and Davide Grossi. 2017. Binary Voting with Delegable Proxy: An Analysis of Liquid Democracy. In *Proceedings of the 16th Conference on Theoretical Aspects of Rationality and Knowledge (TARK)*. 134–150.

[9] Fan Chung, Shirin Handjani, and Doug Jungreis. 2003. Generalizations of Polya's urn problem. *Annals of Combinatorics* 7, 2 (2003), 141–153.

[10] Gal Cohensius, Shie Mannor, Reshef Meir, Eli Meirom, and Ariel Orda. 2017. Proxy Voting for Better Outcomes. In *Proceedings of the 16th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*. 858–866.

[11] Florian Eggenberger and George Pólya. 1923. Über die statistik verketteter vorgänge. *Zeitschrift für Angewandte Mathematik und Mechanik* 3, 4 (1923), 279–289.

[12] Cees M. Fortuin, Pieter W. Kasteleyn, and Jean Ginibre. 1971. Correlation inequalities on some partially ordered sets. *Communications in Mathematical Physics* 22, 2 (1971), 89–103.

[13] Walter Gautschi. 1959. Some elementary inequalities relating to the gamma and incomplete gamma function. *Journal of Mathematics and Physics* 38, 1 (1959), 77–81.

[14] Paul Gölz, Anson Kahng, Simon Mackenzie, and Ariel D. Procaccia. 2018. The fluid mechanics of liquid democracy. In *Proceedings of the 14th Conference on Web and Internet Economics (WINE)*. 188–202.

[15] James Green-Armytage. 2015. Direct voting and proxy voting. *Constitutional Political Economy* 26, 2 (2015), 190–220.

[16] Olle Häggström, Gil Kalai, and Elchanan Mossel. 2006. A law of large numbers for weighted majority. *Advances in Applied Mathematics* 37, 1 (2006), 112–123.

[17] Steve Hardt and Lia C. R. Lopes. 2015. Google votes: A liquid democracy experiment on a corporate social network. *Technical Disclosure Commons* (2015).

[18] Wassily Hoeffding. 1963. Probability Inequalities for Sums of Bounded Random Variables. *J. Amer. Statist. Assoc.* 58, 301 (1963), 13–30.

[19] Svante Janson. 2020. Rate of convergence for traditional Pólya urns. *Journal of Applied Probability* 57, 4 (2020), 1029–1044.

[20] Norman L Johnson and Samuel Kotz. 1978. *Urn Models and Their Application: An Approach to Modern Discrete Probability Theory.* Wiley.

[21] Anson Kahng, Simon Mackenzie, and Ariel D. Procaccia. 2021. Liquid democracy: An algorithmic perspective. *Journal of Artificial Intelligence Research* 70 (2021), 1223–1252.

[22] Christoph Carl Kling, Jérôme Kunegis, Heinrich Hartmann, Markus Strohmaier, and Steffen Staab. 2015. Voting behaviour and power in online democracy: A study of LiquidFeedback in Germany's Pirate Party. In *Proceedings of the 9th International AAAI Conference on Web and Social Media (ICWSM).*

[23] Christian List and Robert E. Goodin. 2001. Epistemic Democracy: Generalizing the Condorcet Jury Theorem. *Journal of Political Philosophy* 9, 3 (2001), 277–306.

[24] Hossam Mahmoud. 2009. *Pólya Urn Models.* CRC Press.

[25] Andrei A. Markov. 1917. Sur quelques formules limites du calcul des probabilités. *Bulletin de l'Académie des Sciences* 11, 3 (1917), 177–186.

[26] James C. Miller. 1969. A program for direct and proxy voting in the legislative process. *Public choice* 7, 1 (1969), 107–113.

[27] Shmuel Nitzan and Jacob Paroush. 2017. Collective decision making and jury theorems. *The Oxford Handbook of Law and Economics* 1 (2017).

[28] Marcus Pivato. 2012. A statistical approach to epistemic democracy. *Episteme* 9, 2 (2012), 115–137.

[29] George Pólya. 1930. Sur quelques points de la théorie des probabilités. *Annales de l'institut Henri Poincaré* 1, 2 (1930), 117–161.

[30] Herbert A. Simon. 1955. On a class of skew distribution functions. *Biometrika* 42, 3/4 (1955), 425–440.

[31] Yuzhe Zhang and Davide Grossi. 2021. Power in liquid democracy. In *Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI).* 5822–5830.

A Proof of Theorem 3

Fix $M^S_{p,\varphi}$ and and $\mathcal{D} \in \mathfrak{D}^C$. Note that since φ is continuous and always positive on the compact set $[0,1]^2$, φ is in fact uniformly continuous and there are bounds $L, U \in \mathbb{R}^+$ such that φ is bounded in the interval $[L, U]$. Additionally, we can assume without loss of generality that for all $x \in [0,1]$, $\mathbb{E}_{\mathcal{D}}[\varphi(x,\cdot)] = 1$. Indeed, $\mathbb{E}_{\mathcal{D}}[\varphi(x,\cdot)]$ is a positive, continuous function of x, so replacing φ by $\varphi'(x,y) = \varphi(x,y)/\mathbb{E}_{\mathcal{D}}[\varphi(x,\cdot)]$ induces the same model and satisfies the desired property.

The Continuous General Delegation Model satisfies (1)*:* Our goal is to show there is some $C(n) \in O(\log n)$ such that, with high probability, no voter receives more than $C(n)$ delegations. To do this, just as in the proof of Theorem 2, we consider a branching process of the delegations received beginning with some voter i. We will show that under minimal conditions on the sampled competencies (which all occur with high probability), this branching process will be dominated by a well-known *subcritical multi-type Poisson branching process* [5], which has size $O(\log n)$ with high probability.

For a fixed competence vector $\vec{p}_n$, the branching process for the number of delegations received by a voter i works as follows. We keep track of three sets of voters: those that are live at time t (L_t), those dead at time t (D_t), and those neutral at time t (N_t). Unlike in the proof of Theorem 2, where it was sufficient to keep track of the number of voters in each category, here we must keep track of the voter identities as well, as they do not all delegate with the same probability. At time zero, the only live voter is voter i and the rest are neutral, so $L_0 = \{i\}$, $D_0 = \emptyset$, and $N_0 = [n] \setminus \{i\}$. As long as there are still live voters, we sample the next set of delegating voters Z_t in time t by choosing some live voter $j \in R_{t-1}$ and sampling its children. Once j's children are sampled, j becomes dead, and j's children become live. All voters that did not delegate and were not delegated to remain neutral. The children are sampled independently; the probability they are included is the probability they delegate to j conditioned on them not delegating to the dead voters in D_{t-1}. For each voter $k \in N_{t-1}$, k will be included with probability

$$p \cdot \frac{\varphi(p_k, p_j)}{\sum_{k' \in [n] \setminus (D_{t-1} \cup \{k\})} \varphi(p_k, p_{k'})}.$$

This is precisely the probability k delegates to j conditioned on them not delegating to any voter in D_{t-1}. We continue this process until there are no more live voters, at which point the number of delegations is simply the number of dead voters, or equivalently, the total number of time steps. We denote by $\mathfrak{X}^D_{\vec{p}_n,i}$ the size of the branching process parameterized by competencies $\vec{p}_n$ and a voter $i \in [n]$.

Our goal will be to compare $\mathfrak{X}^D_{\vec{p}_n,i}$ to the outcome of a well-known multi-type Poisson branching process. In this branching process, there are a fixed finite number k of types of voters.[5] The process itself is parameterized by a $k \times k$ matrix M, where $M_{\tau\tau'}$ is the expected number of children of type τ' a voter of type τ will have. The process is additionally parameterized by the type $\tau \in [k]$ of the starting voter. The random variable Y_t keeps track of the number of live voters of each type; it is a vector of length k, where the τth entry is the number of live voters of type τ. Hence, $Y_0 = e_\tau$, the (basis) vector with a 1 in entry τ and an entry 0 for all other types. We sample children by taking an arbitrary live voter of type τ' (the τ' component in Y_{t-1} must be positive, indicating that there is such a voter), and sampling its children Z_t, which is also a vector of length k, each entry indicating

[5] In the literature, these are often called *particles*, but to be consistent with our other branching processes, we call them voters here.

the number of children of that type. The vector Z_t is sampled such that the τ'' entry is from the $\text{Pois}(M_{\tau'\tau''})$ distribution. That is, children of different types are sampled independently from a Poisson distribution, with the given expected value. We have the recursion $Y_t = Y_{t-1} + Z_t - e_{\tau'}$.

Note that this means that there is no "pool" of voters to choose from; in fact, it is possible for this process to grow unboundedly large (see [1, Section 11.6] for the classical description of the single-type Poisson branching process). Nonetheless, this process will still converge often enough to remain useful. We denote by $\mathfrak{X}^P_{M,\tau}$ the random variable that gives the size of this branching process, parameterized by expected-children matrix M and starting voter type $\tau \in [k]$. Such a branching process is considered *sub-critical* if the largest eigenvalue of M is strictly less than 1 [5]. In such a case, if we begin with voter of any type $\tau \in [k]$, the probability of the branching process surviving ℓ steps decreases exponentially in ℓ. Hence, there is some c such that for all $\tau \in [k]$,

$$\mathbb{P}[\mathfrak{X}^P_{M,\tau} \le c\log(n)] = 1 - o(1/n).$$

To compare these branching processes, we make a sequence of adjustments to the original branching process that at each step creates a dominating branching process slightly closer in flavor to the multi-type Poisson. In the end, we will be left with a sub-critical multi-type Poisson process that we can bound.

Fix some $\varepsilon > 0$, which is a parameter in all of our steps. Later, we will choose ε to be sufficiently small (specifically, such that $p\frac{(1+\varepsilon)^3}{1-2\varepsilon} < 1$) to ensure that the Poisson branching process is sub-critical. To convert from our delegation branching process to the Poisson branching process, we take a voter's type to be their competence (which completely characterizes their delegation behavior). However, to compare to the Poisson process, there must be a finite number of types. Hence, we partition the interval $[0,1]$ into B buckets, each of size $1/B$, such that voters in the same bucket delegate and are delegated to "similarly". We choose B large enough such that all points in $[0,1]^2$ within a distance of $\sqrt{2}/B$ of each other differ in φ by at most $L \cdot \varepsilon$. (Recall that the range of φ is in the interval $[L, U]$.) This is possible since φ is uniformly continuous. Further, this implies any points $(x,y), (x',y')$ within a square with side length $1/B$ have the property that $\varphi(x,y) \le \varphi(x',y') + L \cdot \varepsilon \le (1+\varepsilon) \cdot \varphi(x',y')$. Note that B depends only on φ and ε, and hence is a constant with respect to the number of voters n.

We say a voter i is of type τ if $\frac{\tau-1}{B} < p_i \le \frac{\tau}{B}$ for $1 \le \tau \le B$ (with a non-strict inequality for $\tau = 1$, so 0 is of type 1). Let $S_\tau = (\frac{\tau-1}{B}, \frac{\tau}{B}]$ be the set of competencies of type τ (except that, in the case that $\tau = 1$, we take S_1 to be the closed interval $[0, \frac{1}{B}]$). Let $\pi_\tau = \mathcal{D}[S_\tau]$ be the probability that a voter has type τ. Since the types form a partition of $[0,1]$, we have that $\sum_{\tau=1}^{B} \pi_\tau = 1$.

For any two types τ, τ', we define

$$\varphi'(\tau,\tau') = \sup_{(x,y)\in S_\tau \times S_{\tau'}} \varphi(x,y).^6$$

We abuse notation by extending φ' to operate directly on competencies in $[0,1]$ by first converting competencies to types and then applying φ'. Then, φ' has the property that for any $p_i, p_j \in [0,1]$,

$$\varphi(p_i,p_j) \le \varphi'(p_i,p_j) \le (1+\varepsilon)\varphi(p_i,p_j).$$

We have that for all τ, if $x \in S_\tau$, then

$$\sum_{\tau'=1}^{B} \varphi'(\tau,\tau')\pi_{\tau'} = \mathbb{E}_{\mathcal{D}}[\varphi'(x,\cdot)] \le (1+\varepsilon) \cdot \mathbb{E}_{\mathcal{D}}[\varphi(x,\cdot)] = (1+\varepsilon).$$

[6] Note that, because φ is increasing in its second coordinate, one can actually write $\tilde{\varphi}(\tau,\tau') = \sup_{x\in S_\tau} \varphi(x, \frac{\tau'}{B})$.

Hence, we define

$$\tilde{\varphi}(\tau, \tau') = \varphi'(\tau, \tau') \cdot \frac{(1+\varepsilon)}{\sum_{\tau''=1}^{B} \varphi'(\tau, \tau'')\pi_{\tau''}}.$$

We again abuse notation to allow $\tilde{\varphi}$ to operate directly on competencies. We have that $\tilde{\varphi}(x, y) \geq \varphi'(x, y) \geq \varphi(x, y)$ for all competencies $x, y \in [0, 1]$ and further, for all τ, $\sum_{\tau'=1}^{B} \tilde{\varphi}(\tau, \tau')\pi_{\tau'} = 1 + \varepsilon$.

The Poisson branching process we will eventually compare to is one with B types parameterized by the expected-children matrix M, where

$$M_{\tau\tau'} = p\frac{(1+\varepsilon)^2}{1-2\varepsilon}\tilde{\varphi}(\tau, \tau').$$

First, we show that M has largest eigenvalue strictly less than 1 (for our choice of ε), so that the branching process will be subcritical. Indeed, M has only positive entries, so we need only exhibit an eigenvector with all nonnegative entries such that the associated eigenvalue is strictly less than 1. The Perron-Frobenius theorem tells us this eigenvalue must be maximal.

The eigenvector we consider is $\vec{\pi} = (\pi_1, \ldots, \pi_B)$ (which has nonnegative entries, as each π_τ is a probability). We show it has eigenvalue $p\frac{(1+\varepsilon)^3}{1-2\varepsilon}$, strictly less than 1 due to our choice of ε. Indeed, we have that

$$(M\vec{\pi})_\tau = \sum_{\tau'=1}^{B} \pi_\tau \tilde{\varphi}(\tau, \tau')\pi_{\tau'} = \pi_\tau p \frac{(1+\varepsilon)^3}{1-2\varepsilon}$$

by the definition of $\tilde{\varphi}$. Hence, $\vec{\pi}$ is our desired eigenvector.

Since $\mathfrak{X}^P_{M,\tau}$ is sub-critical for all τ, we have that there is some c such that for all $\tau \in [B]$, $\mathbb{P}[\mathfrak{X}^P_{M,\tau} \leq c\log(n)] = 1 - o(1/n)$. We take $C(n) = c\log(n)$.

Now we consider our branching process, $\mathfrak{X}^D_{\vec{p},i}$. To make the comparison, we will need some minimal concentration properties. We first show that the sampled competencies $\vec{p}$ satisfy these properties with high probability, and then show that, conditioned on these properties, the branching process $\mathfrak{X}^D_{\vec{p},i}$ is easily comparable to a Poisson process. The properties are the following:

1. For each voter $i \in [n]$, $\sum_{j \neq i} \varphi(p_i, p_j) \geq (1-\varepsilon) \cdot n$.
2. For each type $\tau \in [B]$, the number of voters of type τ, $|\{i \mid p_i \in S_\tau\}| \leq (1+\varepsilon)\pi_\tau n$.

For the first property, fix the competence p_i of a single voter i. Then when sampling the p_js, $\sum_{j \neq i} \varphi(p_i, p_j)$ is the sum $n-1$ independent variables, all in the interval $[L, U]$, with mean 1. Hence, by Hoeffding's inequality, for all competencies c, $\mathcal{D}^n[\sum_{j\neq i} \varphi(p_i, p_j) \geq (1-\varepsilon)n \mid p_i = c] = 1 - o(1/n)$, where the $o(1/n)$ term is independent of c. By the law of total probability, this implies that even when p_i is sampled as well, the $1 - o(1/n)$ bound continues to hold. By a union bound over all n voters, this holds for everybody with probability $1 - o(1)$.

For the second property, note that the number of voters of type τ follows a $\mathrm{Bin}(n, \pi_\tau)$ distribution. A simple application of Hoeffding's inequality implies that for this τ, $|\{i \mid p_i \in S_\tau\}| \leq (1+\varepsilon)\pi_\tau n$ (note that this holds even in the extreme cases where $\pi_\tau = 0$ or $\pi_\tau = 1$). As the number B of types is fixed and independent of n, a union bound over all B types implies this holds for all τ with probability $1 - o(1)$.

Now fix some voter competencies $\vec{p}$ such that both properties hold. We will first upper bound the probability a voter of type τ delegates to a voter of type τ'. Hence, we can compare our branching process to one with these larger probabilities, and this will only dominate our original process.

To that end, since $|D_{t-1}| = t-1 \leq t$ (recall that D_{t-1} consists of the dead voters at time $t-1$), using the first property, we have that for all $i \in [n]$,

$$\sum_{j \in [n] \setminus (D_{t-1} \cup \{i\})} \varphi(p_i, p_j) \geq (1-\varepsilon)n - U \cdot t.$$

Hence, as long as $t \leq \varepsilon n/U$, $\sum_{j \in [n] \setminus (D_{t-1} \cup \{i\})} \varphi(p_i, p_j) \geq (1-2\varepsilon)n$.

Including the fact that $\varphi(p_i, p_j) \leq \tilde{\varphi}(p_i, p_j)$ for all p_i and p_j, we have that for all time steps $t \leq \varepsilon n/U$,

$$p \cdot \frac{\varphi(p_i, p_j)}{\sum_{k' \in [n] \setminus (D_{t-1} \cup \{k\})} \varphi(p_i, p_{k'})} \leq \frac{p}{n} \cdot \frac{\tilde{\varphi}(p_i, p_j)}{(1-2\varepsilon)}.$$

Note that for sufficiently large n, $C(n) \leq \varepsilon n/U$, so from now on we restrict ourselves to such n.

Further, note that by the second property, there will never be more than $(1+\varepsilon)\pi_\tau n$ neutral voters of type τ. Hence, if we take a voter of type τ' at time step $t \leq C(n)$, the number of children it will have of type τ will be stochastically dominated by a $\mathrm{Bin}((1+\varepsilon)\pi_\tau n, \frac{p}{n} \cdot \frac{\tilde{\varphi}(p_i,p_j)}{(1-2\varepsilon)})$, and this is independent for each τ. As n grows large, this distribution approaches a $\mathrm{Pois}(p\frac{(1+\varepsilon)}{1-2\varepsilon}\tilde{\varphi}(\tau, \tau'))$. In particular, this means that for sufficiently large n, it will be stochastically dominated by a $\mathrm{Pois}(p\frac{(1+\varepsilon)^2}{1-2\varepsilon}\tilde{\varphi}(\tau, \tau'))$ distribution (note the extra $(1+\varepsilon)$ factor). Hence, if voter i is of type τ, up to time $t \leq C(n)$, $\mathfrak{X}^D_{\vec{p},i}$ is dominated by $\mathfrak{X}^P_{M,\tau}$, so

$$\mathbb{P}_{\mathcal{D}, M^S_{p,\varphi} \cdot n}[\mathfrak{X}^D_{\vec{p},i} \geq C(n)] \geq \mathbb{P}_{\mathcal{D}, M^S_{p,\varphi} \cdot n}[\mathfrak{X}^P_{M,\tau} \geq C(n)] = 1 - o(1/n).$$

A union bound over all n voters tells us this is true for all voters simultaneously with probability $1 - o(1)$, as needed.

The Continuous General Delegation Model satisfies (2)*:* To show (2) holds, we first show the following.

Let $\mu_\mathcal{D}$ be the mean of the competence distribution $\mathcal{D}$. For a fixed x, let $\varphi^+_x(y)$ be the function $\varphi(x, y) \cdot y$. We show that there is some $c > 0$ such that for all $x \in [0, 1]$,

$$\mathbb{E}_\mathcal{D}[\varphi^+_x] \geq \mu_\mathcal{D} + c. \tag{19}$$

Indeed, if we view $\mathbb{E}_\mathcal{D}[\varphi^+_x]$ as a function of x for $x \in [0, 1]$, first note that it is a continuous function on a compact set, and hence it attains its minimum. Further, for all $x \in [0, 1]$, since $\varphi(x, y)$ and y are both increasing functions of y, by the FKG inequality [12],

$$\mathbb{E}_\mathcal{D}[\varphi^+] > \mathbb{E}_\mathcal{D}[\varphi(x, \cdot)] \cdot \mu_\mathcal{D} = \mu_D,$$

since, by assumption, $\mathbb{E}_\mathcal{D}[\varphi(x, \cdot)] = 1$. Hence, this attained minimum must be strictly larger than μ, implying (19).

Since $\varphi(x, y)$ is normalized so that $\mathbb{E}_{y \sim \mathcal{D}}[\varphi(x, y)] = 1$, $\mathbb{E}_{y \sim \mathcal{D}}[\varphi(x, y) \cdot y]$ is the expected competence of the voter to whom someone of competence x delegates to (prior to other competencies being drawn). Hence, (19) tells us that "on average", all voters (regardless of competence) tend to delegate to those with competence strictly above the mean. Ideally, we would choose $\alpha \approx c/2$ and hope that some concentration result tells us that the weighted competencies post-delegation will be strictly above $\mu + c/2$ (the mean of all competencies will be close to μ by standard concentration results). However, proving this concentration result is surprisingly subtle, as there are many dependencies between different voter delegations. Indeed, if one voter with high competence and many delegations chooses to delegate "downwards" (that is, to someone with very low competence), this can cancel out all of the "expected" progress we had made thus far. Hence, the rest of this proof involves proving concentration *does* in fact hold. We prove this by breaking up the process of sampling instances into much more manageable pieces, where, in each, as long as nothing goes "too" wrong, concentration will hold.

In particular, we will prove that for all $\gamma > 0$, with high probability,

$$\sum_{i=1}^{n} \mathsf{weight}_i(G) \cdot p_i - \sum_{i=1}^{n} p_i \geq (c(1-p) - \gamma)n. \tag{20}$$

Fix such a γ. As in the previous part, fix $\varepsilon > 0$ which will paramaterize our steps. We will later choose ε sufficiently small to get our desired result (precisely ε such that $6\varepsilon + \varepsilon^2 < \gamma$). By choosing $\gamma < c(1-p)$, this value is positive, so we can choose $\alpha = \frac{c(1-p)-\gamma}{2}$ which proves Equation (2)

To that end, we define a sequence of six sampling steps that together are equivalent to the standard sampling process with respect to $\mathcal{D}$ and $M^S_{p,\varphi}$. In each step, we will show that with high probability, nothing "goes wrong", and conditioned on nothing going wrong in all these steps, we will get the α improvement that we desire. The six steps are as follows:

1. Sample a set $M \subseteq [n]$ of voters that choose not to delegate. Each voter is included independently with probability p.

2. Sample competencies p_i for $i \in [n] \setminus M$. Each p_i is sampled i.i.d. from $\mathcal{D}$.

3. Sample competencies p_j for $j \in M$. Each p_j is sampled i.i.d. from $\mathcal{D}$.

4. Sample a set $R \subseteq [n] \setminus M$ of delegators that delegate to those in M. Each voter $i \in [n] \setminus M$ is included independently with probability $\frac{\sum_{j \in M} \varphi(p_i, p_j)}{\sum_{j \in [n] \setminus \{i\}} \varphi(p_i, p_j)}$, that is, the total φ weight they put on voters in M divided by the total φ weight they put on all voters.

5. Sample delegations of voters in $[n] \setminus (M \cup R)$. At this point, we are conditioning on such voters delegating, and when they do delegate, they do so to voters in $[n] \setminus M$. Hence, for each $i \in [n] \setminus (M \cup R)$, they delegate to $j \in [n] \setminus (M \cup \{i\})$ with probability $\frac{\varphi(p_i, p_j)}{\sum_{j' \in [n] \setminus (M \cup \{i\})} \varphi(p_i, p_{j'})}$.

6. Sample delegations of voters in R. At this point, we are conditioning on such voters delegating to those in M. Hence, for each $i \in R$, they choose to delegate to $j \in M$ with probability $\frac{\varphi(p_i, p_j)}{\sum_{j' \in M} \varphi(p_i, p_{j'})}$.

We now analyze each step, describing what could "go wrong". Let $\mathcal{E}_1, \ldots, \mathcal{E}_6$ be the events that nothing goes wrong in each of the corresponding steps. We define these events formally below. Our goal is to show that $\mathbb{P}_{\mathcal{D}, M^S_{p,\varphi}, n}[\mathcal{E}_1 \cap \cdots \cap \mathcal{E}_6] = 1 - o(1)$.

• Let $\mathcal{E}_1$ be the event that $(p-\varepsilon) \cdot n \leq |M| \leq (p+\varepsilon) \cdot n$. Note that M is the sum of n independent Bernouilli random variables with success probability p. It follows directly from a union bound over both variants of Hoeffding's inequality that

$$\mathbb{P}_{\mathcal{D}, M^S_{p,\varphi}, n}[(p-\varepsilon) \cdot n \leq |M| \leq (p+\varepsilon) \cdot n] = 1 - o(1).$$

• Let $\mathcal{E}_2$ be the event that $\sum_{i\in[n]\setminus M} p_i \le n(\mu+\varepsilon)(1-p+\varepsilon)$. Note that $\sum_{i\in[n]\setminus M} p_i$ is the sum of $n-|M|$ i.i.d. random variables with mean μ. Conditioning on event $\mathcal{E}_1$, $|M|$ is lower bounded by $n(p-\varepsilon)$, implying that $n-|M| \le n(1-p+\varepsilon)$ as well. It follows from Lemma 2 that

$$\mathbb{P}_{\mathcal{D},M^S_{p,\varphi},n}\left[\sum_{i\in[n]\setminus M} p_i \le n(\mu+\varepsilon)(1-p+\varepsilon) \;\middle|\; \mathcal{E}_1\right] = 1-o(1)$$

which, combined with $\mathbb{P}_{\mathcal{D},M^S_{p,\varphi},n}[\mathcal{E}_1] = 1-o(1)$, proves that $\mathcal{E}_1 \cap \mathcal{E}_2$ occurs with probability $1-o(1)$.

• Let $\mathcal{E}_3$ be the event consisting of all instances $(\vec{p}, G)$ such that

$$\frac{\sum_{j\in M}\varphi(p_i,p_j)\cdot p_j}{\sum_{j\in M}\varphi(p_i,p_j)} \ge \frac{(1-\varepsilon)}{(1+\varepsilon)}(\mu+c)$$

for all $i \in [n]\setminus M$.

We show $\mathcal{E}_3$ occurs with high probability conditional on $\mathcal{E}_1$ and $\mathcal{E}_1$ (conditioning on $\mathcal{E}_2$ is unnecessary. but makes the final statement easier). Fix a set of voters M and p_i for $i\in[n]\setminus M$ satisfying $\mathcal{E}_1$ and $\mathcal{E}_2$. For each $i\in[n]\setminus M$, we will show that with probability $1-o(1/n)$, when we sample the p_js for $j\in M$, they satisfy

$$\sum_{j\in M}\varphi(p_i,p_j) \le |M|(1+\varepsilon) \tag{21}$$

and

$$\sum_{j\in M}\varphi(p_i,p_j)\cdot p_j \ge |M|(1-\varepsilon)(\mu+c). \tag{22}$$

(21) follows from the fact that $\sum_{j\in M}\varphi(p_i,p_j)$ is the sum of $|M|$ bounded independent random variables

with mean $\mathbb{E}_{y\sim\mathcal{D}}[\varphi(p_i,y)] = 1$. By Hoeffding's inequlaity, since $|M|$ is linear in n, $\sum_{j\in M}\varphi(p_i,p_j)$ is at most $|M|(1+\varepsilon)$ with probability $1-o(1/n)$.

(22) follows from the fact that $\sum_{j\in M}\varphi(p_i,p_j)\cdot p_j$ is also the sum of $|M|$ bounded independent random variables with mean $\mathbb{E}_{\mathcal{D}}[\varphi^+_{p_i}]$. Again, since we have conditioned on $\mathcal{E}_1$, $|M|$ is lower bounded by $(p-\varepsilon)n$, which by Hoeffding's inequality implies that $\sum_{j\in M}\varphi(p_i,p_j)p_j$ is at least $|M|(1-\varepsilon)\mathbb{E}_{\mathcal{D}}[\varphi_{p_i}]$ with probability $1-o(1/n)$.

Finally, we can conclude via a union bound that $\frac{\sum_{j\in M}\varphi(p_i,p_j)\cdot p_j}{\sum_{j\in M}\varphi(p_i,p_j)} \ge \frac{(1-\varepsilon)}{(1+\varepsilon)}(\mu+c)$ with probability $1-o(1/n)$ for any $i\in[n]\setminus M$. Hence, by another union bound over the at most n voters $i\in[n]\setminus M$, $\frac{\sum_{j\in M}\varphi(p_i,p_j)\cdot p_j}{\sum_{j\in M}\varphi(p_i,p_j)} \ge \frac{(1-\varepsilon)}{(1+\varepsilon)}(\mu+c)$ for all $i\in[n]\setminus M$ with high probability.

By the law of total probability, $\mathcal{E}_3$ conditioned on $\mathcal{E}_1$ and $\mathcal{E}_2$ occurs with probability $1-o(1)$, which proves that $\mathcal{E}_1\cap\mathcal{E}_2\cap\mathcal{E}_3$ occurs with probability $1-o(1)$ by the chain rule.

• Let $\mathcal{E}_4$ be the entire sample space. Nothing can "go wrong" during this sampling step. So trivially, $\mathcal{E}_1\cap\mathcal{E}_2\cap\mathcal{E}_3\cap\mathcal{E}_4$ occurs with probability $1-o(1)$.

• Let $\mathcal{E}_5$ be the event that $\mathsf{dels}_i(G) \le C(n)$ for all $i\in[n]\setminus M$ and $\mathsf{total\text{-}weight}(G) \ge n - C(n)^2\log(n)$ in the subgraph G sampled (i.e., with delegations only from voters not in R or M). We will show $\mathcal{E}_5$ occurs with high probability even when we sample a full delegation graph (that is, samples delegations for all voters), which implies it continues to hold even when we sample only some delegations (recall that at this step we have only sampled delegations from voters in $[n] \setminus (M \cup R)$).

The proof of this is very similar to the one in Theorem 2, with one extra step to allow for different φ weights.

It was proved in the previous part of this proof that, for all voters i, we have that $\mathsf{dels}_i(G) \leq C(n)$ with probability $1 - o(1)$ (not conditioned on anything) when we sample entire delegation graphs, so we can safely condition on this fact. We now prove that $\mathbb{P}_{\mathcal{D},M^S_{p,\varphi},n}[\mathsf{total\text{-}weight}(G_n) \geq n - O(\log^3 n) \mid \mathsf{dels}_i(G) \leq C(n)] = 1 - o(1)$.

We begin by bounding the number of voters that end up in cycles. Fix some voter i, and let us begin by sampling their delegation tree.

Since we are conditioning on the tree having size at most $C(n)$, the most weight that voter i can place on all of the voters in i's delegation tree is $U \cdot C(n)$. The minimum weight that i can place on all voters is $L(n-1)$. Hence, the probability that i delegates to someone in i's tree conditional the delegation tree having size at most $C(n)$ is at most $p \cdot \frac{U \cdot C(n)}{L \cdot (n-1)}$. Since i was arbitrary, this implies that the expected number of voters in cycles can be at most $n \cdot p \cdot \frac{U \cdot C(n)}{L \cdot (n-1)} \in O(\log n)$.

Applying Markov's inequality just as in the analogous proof in the previous section, the probability that more than $\log^2 n$ voters are in cycles is at most $np\frac{UC(n)}{L(n-1)\log^2 n} = O(1/\log n) = o(1)$. Further, the total number of people that could delegate to voters in cycles is at most $C(n)$ times the number of voters in cycles. Hence, with probability $1 - o(1)$, there are at most $C(n) \cdot \log^2 n$ voters delegating to those in cycles. This implies the desired bound. Hence, we have proved that $\mathbb{P}_{\mathcal{D},M^S_{p,\varphi},n}[\mathcal{E}_5] = 1 - o(1)$. Since we have already shown that $\mathbb{P}_{\mathcal{D},M^S_{p,\varphi},n}[\mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}_3 \cap \mathcal{E}_4] = 1 - o(1)$, a union bound implies that $\mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}_3 \cap \mathcal{E}_4 \cap \mathcal{E}_5$ occurs with probability $1 - o(1)$ as well.

• We now consider the sixth step. To define $\mathcal{E}_6$, we need some new notation. Fix competencies $\vec{p}$ and a partial delegation graph G such that $(\vec{p}, G)$ is in the first five events. We define Q_i for $i \in R$ to be the random variable representing the competence of the voter to whom i delegates. Since we know i delegates to a voter in M, note that

$$Q_i(G) = p_j \text{ with probability } \frac{\varphi(p_i, p_j)}{\sum_{j' \in M} \varphi(p_i, p_{j'})} \text{ for all } j \in M.$$

Let $\mathcal{E}_6$ be the event consisting of all instance $(\vec{p}, G)$ such that that $\sum_{i \in R} \mathsf{dels}_i(G) \cdot Q_i(G) \geq \frac{(1-\varepsilon)^2}{1+\varepsilon}(\mu+c)(1-p-2\varepsilon)\cdot n$. We show that $\mathbb{P}_{\mathcal{D},M^S_{p,\varphi},n}[\mathcal{E}_6 \mid \mathcal{E}_1 \cap \cdots \cap \mathcal{E}_5] = 1 - o(1)$. This, combined with the the fact $\mathbb{P}_{\mathcal{D},M^S_{p,\varphi},n}[\mathcal{E}_1 \cap \cdots \cap \mathcal{E}_5] = 1 - o(1)$ (shown earlier), implies that $\mathbb{P}_{\mathcal{D},M^S_{p,\varphi},n}[\mathcal{E}_1 \cap \cdots \cap \mathcal{E}_6] = 1 - o(1)$. It follows from the definition of Q_i that

$$\mathbb{E}[Q_i] = \sum_{j \in M} \frac{\varphi(p_i, p_j)}{\sum_{j' \in M} \varphi(p_i, p_{j'})} \cdot p_j = \frac{\sum_{j \in M} \varphi(p_i, p_j) \cdot p_j}{\sum_{j \in M} \varphi(p_i, p_j)}.$$

By conditioning on $\mathcal{E}_3$, we have that $\mathbb{E}[Q_i] \geq \frac{(1-\varepsilon)}{(1+\varepsilon)}(\mu+c)$ for each $i \in R$. Hence, $\mathbb{E}[\sum_{i \in R} \mathsf{dels}_i(G) \cdot Q_i] \geq (n - |M| - C(n)^2 \log(n)) \cdot \frac{1+\varepsilon}{1-\varepsilon} \cdot (\mu + c)$, since we are conditioning on $\mathcal{E}_3$ and $\mathcal{E}_5$. Further, for sufficiently large n, $C(n)^2 \log(n) \leq \varepsilon n$; since we are conditioning on $\mathcal{E}_1$, $|M| \leq (p + \varepsilon)n$, so we have that for sufficiently large n,

$$\mathbb{E}[\sum_{i \in R} \mathsf{dels}_i(G) \cdot Q_i] \geq (1 - p - 2\varepsilon) \cdot \frac{1+\varepsilon}{1-\varepsilon} \cdot (\mu + c) \cdot n \in \Omega(n).$$

Next, consider $\mathrm{Var}[\sum_{i\in R} \mathsf{dels}_i(G)\cdot Q_i]$. Since each Q_i takes on values in $[0,1]$, $\mathrm{Var}[Q_i] \le 1$. Further, each summand is independent, as each Q_i is independent and we have fixed G, so we can can view $\mathsf{dels}_i(G)$ as a constant. Hence, $\mathrm{Var}[\sum_{i\in R} \mathsf{dels}_i(G)\cdot Q_i] \le \sum_{i\in R} \mathsf{dels}_i(G)^2 \in o(n^2)$ since, for all i, $\mathsf{dels}_i(G) \le C(n) \in O(\log n)$ and $\sum_i \mathsf{dels}_i(G) \le n$. Hence,

$$\begin{aligned}
&\mathbb{P}_{\mathcal{D},M^S_{p,\varphi},n}[\sum_{i\in R} \mathsf{dels}_i(G)\cdot Q_i < \frac{(1-\varepsilon)^2}{1+\varepsilon}(\mu+c)(1-p-2\varepsilon)\cdot n]\\
&\le \mathbb{P}_{\mathcal{D},M^S_{p,\varphi},n}[\sum_{i\in R} \mathsf{dels}_i(G)\cdot Q_i < (1-\varepsilon)\mathbb{E}[\sum_{i\in R} \mathsf{dels}_i(G)\cdot Q_i]]\\
&\le \frac{\mathrm{Var}[\sum_{i\in R} \mathsf{dels}_i(G)\cdot Q_i]}{\varepsilon^2\cdot \mathbb{E}[\sum_{i\in R} \mathsf{dels}_i(G)\cdot Q_i]^2} \in o(1)
\end{aligned}$$

where the second inequality is due to Chebyshev's inequality, which is $o(1)$ because the numerator is $o(n^2)$ and the denominator is $\Omega(n^2)$. This implies the desired result.

Finally, we show that for all instance $(\vec{p}, G) \in \mathcal{E}_1 \cap \cdots \cap \mathcal{E}_6$, (2) holds, and hence so does (20). We have that $\sum_{i=1}^n w_i(G)\cdot p_i = \sum_{i\in R} \mathsf{dels}_i(G)\cdot Q_i(G) + \sum_{j\in M} p_j$, because in G each voter $i \in R$ delegates all of their $\mathsf{dels}_i(G)$ votes to the voter in M with competence $Q_i(G)$. Hence, $\sum_{i=1}^n w_i(G)\cdot p_i - \sum_{i=1}^n p_i = \sum_{i\in R} \mathsf{dels}_i(G)\cdot Q_i(G) - \sum_{i\in[n]\setminus(M\cup R)} p_i$. Since $(\vec{p}, G) \in \mathcal{E}_2$, we have that $\sum_{i\in[n]\setminus M} p_i \le n(\mu+\varepsilon)(1-p+\varepsilon)$. Since $(\vec{p}, G) \in \mathcal{E}_6$, we have that $\sum_{i\in R} \mathsf{dels}_i(G)\cdot Q_i(G) \ge \frac{(1-\varepsilon)^2}{1+\varepsilon}(\mu+c)(1-p-2\varepsilon)\cdot n$. Hence, this difference is at least

$$\begin{aligned}
((\mu+c)(1-p-2\varepsilon) - (\mu+\varepsilon)(1-p+\varepsilon))n &\ge (c(1-p) - 3\varepsilon\mu - 2\varepsilon c - (1-p)\varepsilon - \varepsilon^2)n\\
&\ge (c(1-p) - 6\varepsilon - \varepsilon^2)n
\end{aligned}$$

where the second inequality holds because, $c, (1-p), \mu \le 1$. By choosing ε such that $6\varepsilon + \varepsilon^2 \le \gamma$ ($\varepsilon = \min(\gamma/7, 1)$ will do), (20) follows.

The Continuous General Delegation Model Satisfies (3)*:* We now show that there exists a distribution $\mathcal{D}$ and $\alpha > 0$ such that $\sum_{i=1}^n p_i + \alpha n \le n/2 \le \sum_{i=1}^n \mathsf{weight}_i(G_n)\cdot p_i - \alpha n$ with probability $1 - o(1)$. This implies that the model $M^S_{p,\varphi}, n$ satisfies probabilistic positive gain by Lemma 1.

As in earlier arguments, let $\mathcal{D}_\eta = \mathcal{U}[0, 1-2\eta]$ for $\eta \in [0, 1/2)$. Note that

$$f(\eta) = \inf_{x\in[0,1]} \left\{\mathbb{E}_{D_\eta}[\varphi^+_x]\right\}\cdot(1-p) - 3\eta/2$$

is a continuous function of η.

Moreover, $f(0) > 0$. Hence, for sufficiently small $\eta > 0$, $f(\eta) > 0$.

Consider $\mathcal{D}_\eta$ for some $\eta > 0$ such that $f(\eta) > 0$. Let $\alpha = \min(\eta/2, f(\eta)/2)$. Since $\mu_{\mathcal{D}_\eta} = 1/2 - \eta$, by Hoeffding's inequality, $\sum_{i=1}^n p_i \le (1/2 - \eta/2)n \le n/2 - \alpha n$ with high probability.

Next, note that we can choose $c = \inf_{x\in[0,1]} \left\{\mathbb{E}_{D_\eta}[\varphi^+_x]\right\}$ in order to satisfy (19). Hence, by choosing $\gamma = f(\eta)/2$, it follows from (20) that

$$\sum_{i=1}^n \mathsf{weight}_i(G)\cdot p_i - \sum_{i=1}^n p_i \ge (c(1-p) - f(\eta)/2)n = (3\eta/2 + f(\eta)/2)n \ge (3\eta/2 + \alpha)n$$

with high probability. Further, by Hoeffding's iequality, $\sum_{i=1}^{n} p_i \geq (1/2 - 3\eta/2)n$ with high probability, so by the union bound applied to these inequalities,

$$\sum_{i=1}^{n} \mathsf{weight}_i(G) \cdot p_i \geq n/2 + \alpha n$$

with high probability, as needed. $\square$