Federated Assemblies

Daniel Halpern
Harvard University

Ariel D. Procaccia
Harvard University

Ehud Shapiro
Weizmann Institute of Science

Nimrod Talmon
Ben-Gurion University

Abstract

A *citizens' assembly* is a group of people who are randomly selected to represent a larger population in a deliberation. While this approach has successfully strengthened democracy, it has certain limitations that suggest the need for assemblies to form and associate more organically. In response, we propose *federated assemblies*, where assemblies are interconnected, and each parent assembly is selected from members of its child assemblies. The main technical challenge is to develop random selection algorithms that meet new representation constraints inherent in this hierarchical structure. We design and analyze several algorithms that provide different representation guarantees under various assumptions on the structure of the underlying graph.

1 Introduction

Citizens' assemblies are a popular mechanism for democratic decision making [25, 26, 9, 16]. In the last decade, this paradigm has vastly grown in recognition and influence. In Europe, for example, governments have sponsored citizens' assemblies to inform national policy on constitutional questions (Ireland), climate change (France), and even nutrition (Germany). Technology companies like Meta [7] are also piloting (enormous) citizens' assemblies as a way of obtaining democratic inputs for AI governance and alignment.

While different assemblies may take somewhat different approaches, they all share two distinctive features. First, members of a citizen's assembly are *randomly selected* among volunteers. Second, members of the assembly engage in a long and substantial *deliberation* before reaching any conclusions.

The former feature is of great technical interest, as it is challenging to design a good random selection process. The goal is to achieve *descriptive representation*, in the sense that the assembly should reflect the composition of the population along multiple dimensions like gender, age, ethnicity and level of education; this is seen as a source of legitimacy for citizens' assemblies. However, since the pool of volunteers is typically skewed due to self-selection bias, uniform random selection will not yield descriptive representation and more sophisticated algorithmic solutions are required. Such algorithms, which are designed to achieve descriptive representation while optimizing fairness to volunteers, have been broadly deployed [19].

Our proposal: Federated assemblies. To our knowledge, the hundreds of citizens' assemblies convened around the world have all been independent of each other: in collaboration with practitioners, different countries, regions and municipalities have organized their own assemblies from the ground up.

By contrast, we propose a novel form of citizens' assemblies: *federated assemblies*. The most basic building block of a federated assembly is two assemblies (say, each representing the residents of a

city) that decide to federate, forming a new parent assembly (which represents the residents of both cities and discusses policy questions of mutual interest); crucially, *the members of the parent assembly are selected from the child assemblies*. More generally, a parent assembly can have more than two children, a child assembly can have more than one parent, and the parent assemblies themselves can federate. Overall, a federated assembly is represented by a directed acyclic graph, where nodes correspond to assemblies and an edge from x to y means that x federated with other assemblies to form the parent assembly y. The lowest level (non-federated) assemblies are *leaf assemblies*, which allow people to directly sign up as constituents. Even leaf assemblies need not represent only geographical entities, they can also correspond to issues (such as climate change) or identity groups (such as ethnic groups). Furthermore, people can sign up for multiple leaf assemblies — any that intersect their multi-faceted interests.

In our view, federated assemblies have several advantages over the current practice of citizens' assemblies. First, the process of forming a new assembly does not have to start from the very beginning, as its members are selected from child assemblies; therefore, the lengthy and costly process of recruiting volunteers can be avoided altogether and the bar for forming an assembly is significantly lowered. Second, in standard citizens' assemblies, the determination of which features to stratify over, and which values to assign to these features — which is made by the organizers — is sometimes controversial and gives rise to manipulation opportunities. By contrast, in federated assemblies, these "features" — which are induced by the structure of the graph — are self-determined. Third, in the spirit of *associative democracy* — "a model of democracy where power is highly decentralized and responsibility for civic well-being resides with like-minded civic associations" [15] — federated assemblies allow citizens to exercise power by forming organizations that are immediately integrated into a broader framework of governance. We believe that federated assemblies may be especially pertinent in the context of a *global citizens' assembly* — the holy grail of practitioners of deliberative democracy — as such an assembly could form organically as a federation of assemblies representing different countries, regions, and global issues.

Technical challenge and our results. Our proposal is undoubtedly radical and we acknowledge that the devil is in the details; we discuss some limitations in Section 6. Our goal in this paper is to address a key, technically challenging question that arises as we consider the implementation of federated assemblies: how should they be selected?

In the context of federated assemblies, we think of an assembly as satisfying descriptive representation if it reflects both its child assemblies and its constituents. Specifically, we wish to design a random selection process, where the members of each assembly are selected from its child assemblies, so that the following constraints are satisfied:

- *Individual Representation:* Let the assembly's *population* be the union of all (possibly overlapping) populations of its descendant leaf assemblies. Each member of this population should have an equal probability of being represented on the assembly. This constraint can be interpreted as realizing an *equality of power* ideal.
- *Ex ante representation of child assemblies:* The expected number of seats allocated to each child assembly should be proportional to the child assembly's population.[1]
- *Ex post representation of child assemblies:* The number of seats allocated to each child assembly should be proportional to the child assembly's population, rounded down, *ex post*. This ex post guarantee prevents situations where an unlucky draw leads to significant under-representation; it mirrors ex post quotas imposed on different features in the selection of standard citizens' assemblies.

Our results are primarily positive, showing that achieving various properties together is indeed tractable.

We begin by considering only the first two properties, individual representation and ex ante representation of child assemblies (Section 3). We design a simple algorithm (Algorithm 1) that is able to achieve both of these properties under some minor regularity conditions.

[1]When calculating a child's proportional share, if some populations overlap, we first split the weight of members in the intersection equally across their populations, e.g., if a member is in three child populations, they only contribute 1/3 to each.

Next, we throw ex post child representation into the mix (Section 4). This makes the problem much more challenging, so we focus on instances with additional structure. We begin with a very natural class, which we call *laminar instances*, where the assembly graph is a tree, and members are signed up for only one leaf assembly. This captures instances where assemblies represent hierarchical regions, e.g., city-level which feed into state-level which feed into national-level. For such instances, we give an algorithm (Algorithm 2) that achieves all of our desired properties. Next, we generalize laminar instances to a larger class, *semi-laminar*. These are rooted in a laminar instance, with multiple assemblies at each node that federate together, allowing constituents to, for example, organize both geographically and according to shared interests. It turns out that even this level of generality already adds quite a bit of complexity to the problem. We nonetheless devise a surprisingly intricate algorithm (Algorithm 3), that again, under mild regularity conditions, is able to achieve individual representation, ex ante representation of child assemblies, and approximate ex post representation of child assemblies up to an additive error of one.

Finally, in Section 5, we implement an algorithm based on column generation for convex programming, which, given any instance of our problem, computes a distribution satisfying the three properties. In addition to measuring the running time of the algorithm, our empirical results suggest that all of our properties can be achieved simultaneously in the general case, at least in practice.

Related work. There is a growing body of work on algorithms for randomly selecting citizens' assemblies, starting with the paper by Flanigan et al. [11]; there is significant interest in this topic in AI conferences, especially NeurIPS [10, 12, 8, 13]. The key challenge these papers address arises because of quotas imposed on multiple, overlapping features. By contrast, our community representation constraint amounts to stratified sampling with respect to a partition of the population, which is simple in the case of standard citizens' assemblies [3]. The difficulty of our problem stems from its graph structure and the need to also represent child assemblies, leading to technical questions that are quite different from prior work. We do note that the mathematical programming approach used by Flanigan et al. [11] to implement their (widely used) algorithm also works well in our case.

Conceptually, federated assemblies are somewhat related to *pyramidal democracy* [18]. In this scheme, citizens self-organize into small groups at the bottom of the pyramid, with each group nominating a delegate. In the next level, those delegates form groups, each again nominates a delegate, and so on. Our proposal differs in several significant ways, but most importantly, the selection of delegates is not random in pyramidal democracy — rather, it is up to each group to decide how to select its delegate — and there are no representation requirements; by contrast, the whole point of our work is to design a random selection process that satisfies representation constraints.

Our technical contribution fits more broadly within the literature on dependent rounding [14]. Arguably the closest in flavor is randomized rounding for flows [21], which has constraints similar to our inheritance ones. Furthermore, rounding the number of seats given to each child assembly (a step in several of our algorithms) is reminiscent of randomized apportionment [2, 20]. However, our inheritance and ex post child constraints do not fit neatly within existing frameworks and make it challenging to use off-the-shelf techniques directly. Instead, many of our results repeatedly invoke existing schemes, such as variations of the Birkhoff Von-Neumonn Theorem [4, 6], in nontrivial ways. We discuss these techniques in more detail when we use them.

2 Model

Let N be a finite set of people and G be a directed acyclic graph. Abusing notation slightly, we write $v \in G$ if v is a node in G.

For a node $v \in G$, let $\textsc{children}(v) = \{v' : v \to v' \text{ in } G\}$ be the set of nodes v' with a directed edge from v to v'. Let $\textsc{leaves}(G) = \{v : |\textsc{children}(v)| = 0\}$ be the leaf nodes of G. Each person $i \in N$ will be signed up for a nonempty set $L_i \subseteq \textsc{leaves}(G)$ of leaf nodes. For a set $L \subseteq \textsc{leaves}(G)$, we will also use $C^L = \{i : L_i = L\}$ for the set of people signed up for exactly the leaf nodes L. We refer to each C^L as an *equivalence class*, which collectively form a partition of N. For a leaf node $\ell \in G$, its *population* $N_\ell = \{i : \ell \in L_i\}$ is the set of all people signed up for it.

Let $\textsc{feds}(G) = \{v : |\textsc{children}(v)| > 0\}$ be the internal nodes of G, which we refer to as *federations*. For a federation $f \in \textsc{feds}(G)$, let $\textsc{descendants}(f) \subseteq \textsc{leaves}(G)$ be the set of all leaf nodes reachable from f in G. For a federation $f \in \textsc{feds}(G)$, its population is defined as

$N_f := \bigcup_{\ell \in \text{DESCENDANTS}(f)} N_\ell$. Note that this can equivalently be defined in terms of equivalence classes as $N_f = \bigcup_{L: L \cap \text{DESCENDANTS}(f) \neq \emptyset} C^L$.

An *instance* is a tuple $\mathcal{I} = \langle G, (N_v)_{v \in G} \rangle$ of the graph and the membership relationships. Given an instance $\mathcal{I}$ and a *target assembly size* n, our goal is to choose an *assembly* $A_v \subseteq N_v$ of size n for each node v. We will assume, in general, that each $|N_v| \geq n$ so that this is always possible. Furthermore, we require that each federation's assembly be drawn from its child assemblies, i.e., for $f \in \text{FEDS}(G)$, $A_f \subseteq \bigcup_{c \in \text{CHILDREN}(f)} A_c$. We call this the *inheritance* property. A vector $\boldsymbol{A} = (A_v)_{v \in G}$ satisfying these requirements an *assembly assignment* (or simply an assignment). Our algorithm for selecting assembly assignments will be random, and thus, their outputs will be distributions over assembly assignments, $\boldsymbol{\mathcal{A}}$. We will call such a distribution a *randomized assembly assignment* and use $\mathcal{A}_v$ to refer to the marginal distribution for the assembly at node v.

We would like our randomized assembly assignments to satisfy various properties. Some are ex post and should hold for all assignments in the support. Others will be ex ante and are simply properties of the distribution that hold in expectation.

Desired Properties. Arguably, the most important requirement is *individual representation*. A randomized assignment $\boldsymbol{\mathcal{A}}$ satisfies *individual representation* if for each node v and $i \in N_v$, $\Pr[i \in \mathcal{A}_v] = n/|N_f|$, that is, each person has an equal chance of being selected to the assembly.

The other flavor of requirements we have on solutions are with respect to child assemblies. For a federation $f \in \text{FEDS}(G)$ and child $c \in \text{CHILDREN}(f)$, we think of $|A_f \cap A_c|$ as the number of seats child c is allocated, and we would like this allocation to be at least as large as c "deserves." The question, however, is how to set these bounds. If the child populations N_c of a federation f are all pairwise disjoint, then a natural choice is that each c should be allocated an $|N_c|/|N_f|$ fraction of the n seats in A_f. For non-disjoint child populations, we generalize this by splitting a person's weight equally among all child populations they are a part of. More formally, for a federation f and member $i \in N_f$, define the *multiplicity of* i *at* f to be $m(i, f) = |\{c \in \text{CHILDREN}(f) : i \in N_c\}|$, the number of child nodes they are a member of. The *weighted population size* $w_{c,f}$ of a child federation $c \in \text{CHILDREN}(f)$ is defined by $w_{c,f} = \sum_{i \in N_c} \frac{1}{m(i,f)}$. Note that the definition implies that $|N_f| = \sum_{c \in \text{CHILDREN}(f)} w_{c,f}$. Hence, we would say that node c *deserves* a $w_{c,f}/|N_f|$ fraction of the seats in A_f. Define $q_{c,f} := w_{c,f}/|N_f|$ to be this fraction.

However, $n \cdot q_{c,f}$ need not be an integer. Hence, we consider two notions of fair child representation. First, we say that a randomized assignment $\boldsymbol{\mathcal{A}}$ satisfies *ex ante child representation* if, for each federation $f \in \text{FEDS}(G)$ and child $c \in \text{CHILDREN}(f)$, $\mathbb{E}[|\mathcal{A}_f \cap \mathcal{A}_c|] \geq n \cdot q_{c,f}$. Similarly, we will say that a randomized assignment satisfies *ex post child representation* if for each assignment $\boldsymbol{A}$ in the support and each federation $f \in \text{FEDS}(G)$ and child $c \in \text{CHILDREN}(f)$, $|A_f \cap A_c| \geq \lfloor n \cdot q_{c,f} \rfloor$. In other words, even if it is impossible to guarantee exact child representation, we can ensure that all children get at least the number of seats they deserve rounded down.

3 Ex Ante Child Representation

We begin our study with an algorithm that samples from a distribution satisfying inheritance, individual representation, and ex ante child representation under mild conditions, proving that these three properties are compatible. To gain intuition, we first informally present a simpler version of the algorithm tailored to the special case of assemblies of size $n = 1$. It works by selecting a single random order of all of N uniformly at random; the representative for node v is simply the highest-ranked member of N_v. Note that this allows for strikingly simple arguments for the various properties. Indeed:

- *Inheritance*: If $A_f = \{i\}$ for a federation f then $i \in N_c$ for some $c \in \text{CHILDREN}(f)$. Since $N_c \subseteq N_f$, i will be maximal among N_c, and hence, $A_c = \{i\}$.
- *Individual Representation*: Each $i \in N_v$ is equally likely to be the maximally ranked person among N_v, so they are a member of A_v with probability $1/|N_v|$.
- *Ex ante child representation*: A similar argument to inheritance implies that $A_c = A_f$ with probability $|N_c|/|N_f| \geq w_{c,f}$.

Algorithm 1 Selection algorithm with ex ante child representation guarantees

Input Graph G, populations $(N_v)_{v \in G}$, and assembly size n
Output Assembly assignment $(A_v)_{v \in G}$

1: Choose n linear orders over N, $\succ^1, \ldots, \succ^n$ independently uniformly at random
2: Select an additional linear order $\succ^{equiv}$ over N uniformly at random.
3: **for** $v \in G$ **do**
4: Let $i_{v,1}^{max}, \ldots, i_{v,n}^{max}$ be the maximally ranked people in $\succ^1, \ldots, \succ^n$ when restricted to N_v
5: **for** $L \subseteq \text{LEAVES}(G)$ **do**
6: $r_v^L \leftarrow |\{j : i_{j,v}^{max} \in C^L\}|$, i.e., the number of $i_{j,v}^{max}$ members in equivalence class C^L
7: Let B_v^L be the r_v^L highest ranked members of C^L in $\succ^{equiv}$
8: $A_v \leftarrow \bigcup_L B_v^L$
9: **return** $(A_v)_{v \in G}$

The challenge is to extend the foregoing algorithm to $n \geq 1$. One straightforward idea is to run the same algorithm n independent times, which would produce n singleton assemblies $A_v^1, \ldots, A_v^n$ for each node v, and subsequently set $A_v = \bigcup_j A_v^j$. At first glance, this seemingly maintains all of the properties of the $n = 1$ algorithm. However, this idea, unfortunately, does not quite compile; with some positive probability, we will have $j \neq j'$ with $A_v^j = A_v^{j'}$ so we do not end up with a size-n assembly. We can, nevertheless, remedy this by replacing selected members with distinct people from their equivalence class. This fix does impose an additional requirement that each nonempty equivalence class contains at least n people. However, this is a relatively mild condition since we view n as a reasonably small constant and populations as potentially very large. Later, we discuss how even this mild assumption can be relaxed while only slightly degrading the guarantees.

Theorem 1. *Assume that each nonempty equivalence class C^L satisfies $|C^L| \geq n$. Then, Algorithm 1 satisfies individual representation and ex ante child representation.*

Proof. Note that each $r_v^L \leq n$ and $r_v^L > 0$ only if C^L is nonempty. The requirement that each C^L be of size at least n ensures that line 7 is able to run, and we can always pick the b_v^L highest ranked members. Without this, line 7 may fail.

Since equivalence classes are disjoint and $\sum_L r_v^L = n$ (each $i_{v,j}^{max}$ will contribute to exactly one), we have that each $|A_v| = n$. Furthermore, if $i \in A_v$, there was another $i' \in N_v$ such that both i and i' are in the same equivalence class. Hence, $i \in N_v$ and the chosen assemblies are also valid.

We next show that each of the properties holds. For each $L \subseteq \text{LEAVES}(G)$, node v, and $j \leq n$, let $I_{v,j}^L = \mathbb{I}[i_{v,j}^{max} \in C^L]$ be the indicator variable that $i_{v,j}^{max}$ is signed up for the set of leaves L. As long as $i \in N_v$, we have that $\mathbb{E}[I_{v,j}^L] = |C^L|/|N_v|$ (over the randomness of the selected orders).

We begin with individual representation. Fix a node v and a person $i \in N_v$. Suppose we condition on a specific value of r_v^L, then (abusing notation slightly) over the randomness of $\succ^{equiv}$, we have that $\Pr[i \in \mathcal{A}_v : r_v^L] = r_v^L/|C^L|$ because each person $i \in C^L$ is equally likely to be in any of the $|C^L|$ positions. Next, note that $r_v^L = \sum_j I_{v,j}^L$. Hence, $\mathbb{E}[r_v^L] = n|C^L|/|N_v|$. Putting these together, we have that $\Pr[i \in \mathcal{A}_v] = n/|N_v|$, as needed.

Next, we show inheritance. Fix a federation $f \in \text{FEDS}(G)$ and an equivalence class C^L such that $C^L \subseteq N_f$. Note that there must be some child $c \in \text{CHILDREN}(f)$ such that $C^L \subseteq N_f$. We will show that for this choice of c, $B_f^L \subseteq B_c^L \subseteq A_c$, ex post. As this holds for every L, it follows that $A_f = \bigcup_L B_f^L \subseteq \bigcup_{c \in \text{CHILDREN}(f)} A_c$, ex post. To that end, it is sufficient to show that $r_f^L \leq r_c^L$, which implies that $B_f^L \subseteq B_c^L$. For this, we can simply show that for each j, $I_{v,f}^L \leq I_{v,c}^L$, ex post. Indeed, if the maximal selected member of $\succ_j$ when restricted to N_v is a member of C^L, then the same person will be maximal when restricted to N_c because $C^L \subseteq N_c$.

Finally, we show ex ante child representation. The key observation is that for a child $c \in \text{CHILDREN}(f)$, $|A_c \cap A_f| = \sum_{L : C^L \subseteq N_c} r_f^L$ because, as shown above, $r_f^L \leq r_c^L$, so the top r_f^L

Algorithm 2 Selection algorithm for laminar instances with ex post child representation guarantees

Input Graph G, populations $(N_v)_{v \in G}$, and assembly size n
Output Assembly assignment $(A_v)_{v \in G}$

1: **for** leaf $\ell \in \text{LEAVES}(G)$ **do**
2: Choose $A_\ell \subseteq N_\ell$ uniformly at random
3: **for** federation $f \in \text{FEDS}(G)$ in a topological sort **do**
4: Round $(n \cdot q_{c,f})_{c \in \text{CHILDREN}(f)}$ to an integral vector $(s_{c,f})_{c \in \text{CHILDREN}(f)}$ such that each $s_{c,f} \geq \lfloor w_{c,f} \rfloor$, $\mathbb{E}[s_c] = n \cdot w_{c,f}$, and $\sum_c s_{c,f} = n$
5: Select $B_{c,f} \subseteq A_c$ with $|B_{c,f}| = s_{c,f}$ uniformly at random.
6: Let $A_f \leftarrow \bigcup_c B_{c,f}$.
7: **return** $(A_v)_{v \in G}$

people from C^L are contained in both A_c and A_f. It follows that

$$\mathbb{E}[\mathcal{A}_c \cap \mathcal{A}_f] = \sum_{L : C^L \subseteq N_c} \mathbb{E}[r_f^L] = \sum_{L : C^L \subseteq N_c} \frac{|C^L| \cdot n}{|N_f|} = \frac{|N^c| \cdot n}{|N_f|} \geq q_{c,f} \cdot n,$$

as needed. □

We now discuss ways to implement a modification of Algorithm 1 that works even when $|C^L| < n$. The key idea is that Algorithm 1 will only fail to run if there is a node v such that $r_v^L > |C^L|$. In such cases, we can simply "reject" and restart the algorithm. Note that ex post guarantees will still be satisfied as long as the algorithm is able to terminate. For ex ante guarantees, as long as the probability of failure is at most some value p, the properties only degrade as a function of p. Specifically, each $i \in N_v$ will be selected in A_v with probability at least $n/|N_v| - p$, and for a child $c \in \text{CHILDREN}(f)$, their expected intersection will be at least $n(w_{c,f} - p)$. As long as $p \ll 1/|N_v|$, this will be a negligible loss. We show that, under mild conditions on the populations being reasonably large and equivalence classes being at least of size 3, this is indeed the case; see Appendix A for more details.

4 Ex Post Child Representation

In this section, we add ex post child representation to our list of requirements, albeit at a cost to the generality of our results.

4.1 Laminar Instances

We begin with a more restricted structure that captures many practical potential implementations of federated assemblies. The assumption is that G is a tree and that each $i \in N$ is signed up for exactly one leaf node, i.e., $|L_i| = 1$. This captures settings such as where the assemblies represent regions that form a hierarchy, i.e., city-level assemblies, which feed into state-level assemblies, which feed into national-level assemblies, and possibly beyond. Following set-theoretic terminology, we call instances satisfying these restrictions *laminar*.

Algorithm 2 works by going through the federations, allocating an integral number of seats to each child, and filling these seats directly from the child's assembly. The rounding (Line 4) can be done in a variety of ways using tools from the dependent rounding literature. Brewer and Hanif [5] provide a number of classical statistical methods for this; canonical examples from randomized algorithm design include *pipage rounding* [1] and variations of the *Birkhoff-von Neumann Theorem* [4, 6].

Theorem 2. *Algorithm 2 satisfies individual representation, ex ante child representation, and ex post child representation on laminar instances.*

While the proof is relegated to Appendix B, we give some intuition by discussing the naturalness of Algorithm 2, which allows for its relatively simple analysis. Specifically, we enforce ex ante and ex post child representation by first allocating the "correct" number of seats to each child. We then go through the tree iteratively, selecting members from their (already determined) child assemblies. We may hope that such ideas could be generalized to non-laminar instances, first allocating seats and

Algorithm 3 Algorithm for semi-laminar instances

Input Semi-laminar instance with graph G, populations $(N_v)_{v \in G}$, and assembly size n
Output Assembly assignment $(A_v)_{v \in G}$

for $(r, t) \in R \times \mathcal{T}$ **do**
  $s_{r,t} \leftarrow \text{ROUND}(n \cdot \frac{w_{r,t}}{|N_{r,t}|})$
for $r \in \text{LEAVES}(R)$ **do**
  $(B^{\text{SEL}}_{r,t}, B^{\text{UNS}}_{r,t})_{t \in T} \leftarrow \text{SAMPLELEAVES}((C^{\{r\} \times T})_{T \subseteq \mathcal{T}}, (s_{r,t})_{t \in \mathcal{T}}, n)$
for $r \in \text{FEDS}(R)$ in a topological sort **do**
  for $t \in \mathcal{T}$ **do**
    $B^{\text{SEL}}_{r,t}, B^{\text{SEL}}_{r,t} \leftarrow \text{SAMPLEFROMCHILDREN}(s_{r,t}, (B^{\text{SEL}}_{c,t}, B^{\text{UNS}}_{c,t}, w_{c,t}, |N_{c,t}|)_{c \in \text{CHILDREN}(r)})$
for $r \in R$ **do**
  for $t \in \mathcal{T}$ **do**
    $A_{(r,t)} \leftarrow B^{\text{SEL}}_{r,t} \cup B^{\text{UNS}}_{r,t}$
  $A_{(r,*)} \leftarrow \text{ROUNDANDSAMPLE}(n, (B^{\text{SEL}}_{r,t}, w_{r,t})_{t \in \mathcal{T}})$
return $(A_v)_{v \in G}$

then iteratively selecting members from the child assemblies, perhaps not uniformly at random, to account for people being in potentially different numbers of children. However, we give an example where this is not the case —relaxing either of the laminar assumptions (that G is a tree or that each $|L_i| = 1$) can lead to instances where satisfying all the properties is impossible using such algorithms. Instead, it is necessary to induce some other forms of correlation or relax the iterativeness, a challenge that leads to more complicated algorithms. We discuss this more formally in Appendix C.

4.2 Semi-Laminar Instances

We now turn to a generalization of laminar instances with a structure we view as quite practical for real-world implementation, as it allows people to organize both geographically and according to shared interests. Conceptually, there is an underlying laminar instance with graph R. In addition, there is a set $\mathcal{T}$ representing a set of *topics*. These may be various causes that people care about, say climate change or animal rights, or region-specific policy questions, such as budget allocation.

More specifically, the graph G has $|R| \cdot (|\mathcal{T}| + 1)$ nodes. $|R| \cdot |\mathcal{T}|$ of these are identified by members of $R \times \mathcal{T}$, i.e., there is a node (r, t) for each combination of region and topic. For each $t \in \mathcal{T}$, the set of nodes $R \times \{t\}$ is connected to form a copy of R. In addition, there is a set of nodes denoted $(r, *)$ for each $r \in R$. Each is a federation whose children are $\{r\} \times \mathcal{T}$. In other words, at each region, we have an assembly that represents all people with respect to all topics of that region.

In this graph, the leaves are nodes in $\text{LEAVES}(R) \times \mathcal{T}$. People can be signed up for any number of topics, but, as the underlying instance is laminar, we assume that each is a member of nodes in one region. Hence, we assume that each $L_i \subseteq \{r\} \times \mathcal{T}$ for some $r \in \text{LEAVES}(R)$.

Abusing notation slightly, we will write $N_r = \{i \mid L_i \subseteq \{r\} \times \mathcal{T}\}$ for all people in this region (technically $N_r = N_{(r,*)}$). Furthermore, a node (r, t) will have at most two parents in G: it will always have $(r, *)$ and possibly another node (r', t) if r has a parent (r') in R. Note that if (r', t) exists, the weighted population $w_{(r,t),(r',t)} = N_{(r,t)}$, the trivial weighting, because the children of (r', t) have disjoint populations. On the other hand, $w_{(r,t),(r,*)} = \sum_{i \in N_{(r,t)}} \frac{1}{|L_i|}$. For brevity, we will use $w_{r,t}$ to denote only the nontrivial weight of the node (r, t). Finally, note that, for $r \in \text{LEAVES}(R)$ and $T \subseteq \mathcal{T}$, we can write $C^{\{r\} \times T}$ for the equivalence class of people in leaf r signed up for topics T. Everybody must fall in exactly one of these equivalence classes.

We refer to an instance taking on the above structure as a *semi-laminar instance*. For such instances, we have the following algorithm shown in Algorithm 3, with additional helper functions formally defined in Algorithm 4 in Appendix D. The structure is essentially an extension of the algorithm for laminar instances. Ideally, for each topic, we could independently run Algorithm 2, and when we needed to select an assembly $(r, *)$, we could select the "correct" number of members from each $A_{(r,t)}$. However, this does not quite work because if people are signed up for more topics, this will

lead to them ending up in $A_{(r,*)}$ more frequently. To account for this, we essentially partition each $A_{(r,t)}$ into two pieces, $B^{\text{SEL}}_{r,t}$ of *(sel)ectable* people and $B^{\text{UNS}}_{(r,t)}$ of *(uns)electable* people, and run a separate laminar-like algorithm for each of these. The key idea is that when selecting members of $A_{(r,t)}$ for $A_{(r,*)}$, we will only select from members in $B^{\text{SEL}}_{r,t}$. While each $i \in N_{(r,t)}$ will have an equal chance of being in $A_{(r,t)}$, the more topics they are signed up for, the less likely they will be in $B^{\text{SEL}}_{r,t}$, so that, in aggregate, this will not increase their chances of being in $A_{(r,*)}$.

It should be said that the real complexity of this algorithm is hidden in the dependent rounding schemes of the helper functions. It requires careful balance and specific constraints to ensure the probabilities are exact and not subject to strange correlations that naïve implementations can impart. Furthermore, we need to avoid scenarios where the same person is selected twice for $A_{(r,*)}$ from two separate topics, which would not appear to have an easy solution. All of these schemes are implemented using variations of the Birkhoff-von Neumann algorithm. Budish et al. [6] give a class of constraints that are always possible to guarantee when rounding; we ensure that all of our rounding procedures take this form.

Finally, note that all of these extra complexities mean that we impose some additional mild regularity conditions and achieve slightly degraded guarantees, at least for ex post child representation. The regularity conditions require that no weighted population is too close to zero or the entire population and that no child population is too dominant within its parent.

Theorem 3. *Fix a semi-laminar instance and assembly size n. Suppose there exist $\varepsilon, \delta > 0$ such that (i) for all $r, t \in R \times \mathcal{T}$, $\varepsilon \cdot |N_{r,t}| \leq w_{r,t} \leq (1-\varepsilon) \cdot |N_{r,t}|$, (ii) for all federations $f \in \text{FEDS}(G)$ and $c \in \text{CHILDREN}(f)$, $|N_c|/|N_f| \leq 1-\delta$, and (iii) $n \geq \frac{2}{\varepsilon \cdot \delta}$. Furthermore, suppose each $|N_v| \geq 4n$ and for all nonempty equivalence classes C^L, $|C^L| \geq 2$. Then, Algorithm 3 run on this instance satisfies individual representation, ex ante child representation, and approximate ex post child representation in that $|A_f \cap A_c| \geq \lfloor n \cdot q_{c,f} \rfloor - 1$ for all $f \in \text{FEDS}(G)$ and $c \in \text{CHILDREN}(f)$.*

The proof is relegated to Appendix E.

5 Experiments

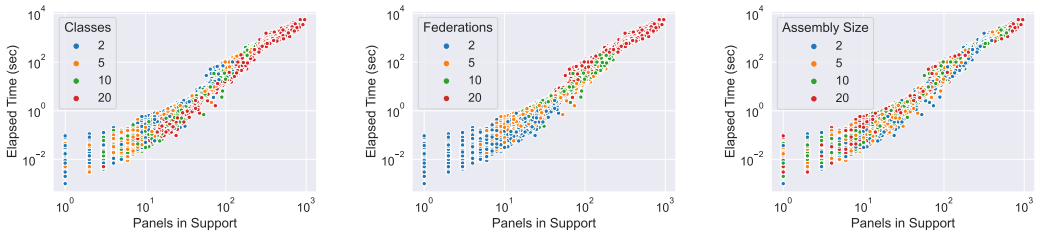
While we have given efficient algorithms for finding randomized solutions satisfying various properties in special cases, we focus now on how easy it is in practice to find randomized assignments in generality. To this end, we sample several thousand instances and use a brute force algorithm to try to find a randomized assignment satisfying all of our ex post and ex ante guarantees.

Algorithm. The computation of randomized assignments satisfying our guarantees is quite challenging. Our algorithm for this task uses convex optimization and integer linear programming (ILP) as subroutines. It is inspired by a related algorithm used for standalone citizens' assemblies, which is an extension of column generation to convex programming [11]. At a high level, we convert our problem to a smoother optimization problem by defining a convex loss over randomized assignments measuring how far a distribution is from satisfying all ex ante guarantees. The algorithm maintains a set of (deterministic) assembly assignments that all satisfy ex post guarantees. It alternates between two steps. First, it finds the distribution over just this set of assignments that minimizes the squared error. If this distribution achieves zero error (or, more accurately, some small additive error of .1% to avoid numerical issues), then every ex ante guarantee is satisfied; in that case, we have found a solution and subsequently return it. On the other hand, if it does not, the algorithm runs an ILP to find an assignment satisfying ex post guarantees that maximizes the loss gradient the current best possible point. We iterate these two steps until a solution is found.

Experimental Setup. We draw instances as follows. First, we fix a number of equivalence classes (2, 5, 10, or 20) and then sample their relative sizes from an exponential distribution. Next, we iterate through a number of federations (2, 5, 10, or 20). Each federation has a randomly selected set of already defined federations or equivalence classes as children.[2] This method allows for arbitrary DAGs to be sampled. We then test these instances with various assembly sizes n (2, 5, 10, or 20).

For each combination of parameters, we sampled and ran 100 instances. All optimizations were solved using Gurobi on an Amazon Web Services (AWS) instance with 128 vCPUs of a 3rd Gen

[2]This means that we allow federations to have direct members and child assemblies. However, this generality only makes the problem harder.



(a) Instances colored by the number of equivalence classes.

(b) Instances colored by the number of federations.

(c) Instances colored by the assembly size.

Figure 1: Scatter plots showing the time taken and number of panels in the support for each of the instances we ran on. Sub-plots show the same plot, colored by a parameter.

AMD EPYC running at 3.6GHz equipped with 1TB of RAM. We were able to run 64 instances in parallel, giving each thread two processors. Depending on the size of the instance, computation took anywhere from under a tenth of a second to multiple hours.

Results. Of the 6400 instances, on all but 15, the algorithm terminated, returning an optimal solution. These 15 had only the largest number of equivalence classes and federations (20 each). With better optimization, in practice, the problem of finding a randomized assignment satisfying all of our guarantees appears to be both feasible and tractable.

For the >99.7% of instances that did terminate, we use two metrics to measure the complexity of finding the distributions: the elapsed time and the number of panels added to the support. A scatter plot showing the relationship between these parameters is given in Figure 1. Each subfigure shows the same points but colors them depending on a different parameter so that we may see its effect. Overall, we see that both of these complexity measures are quite similar. Furthermore, increasing the number of equivalence classes and federations strongly correlates with increased complexity. Assembly size, on the other hand, appears to be relatively unimportant.

6 Discussion

The democratic governance of large-scale digital communities is an open problem. Key challenges include, first, the penetration of fake and duplicate digital identities (a.k.a. sybils), and second, the perils of large-scale online voting, which is considered to be untenable by some leading experts [17]. Federated assemblies can be viewed as a step in an effort to address these challenges. Our approach may address sybils by having the laminar core built from small local communities in which members know each other to be genuine and from communities that federate only if they trust each other to be genuine, for example by having sufficient intersection or actual relationships among them to base this trust on. Large-scale online voting is a nonissue as every federation, no matter how large, is governed by an assembly that engages in "small-scale" democracy.

Our approach also has some limitations. First and foremost, an obvious barrier to federated assemblies is the question of who would set up such assemblies, manage the infrastructure, and provide funding? One way to address this challenge is to use the conceptual framework and architecture of grassroots systems [22, 24] and to construct the application of grassroots federated assemblies as a grassroots platform [23], operated on peoples' smartphones without relying on any global resources other than the network itself.

Second, we modeled the problem as a static and single-shot: we simply needed to sample a single assignment fairly. In practice, however, these assemblies are dynamic and must be periodically updated. This is not an inherent limitation, however. Indeed, one solution is to resample fresh assemblies every fixed amount of time. However, this may get more challenging if the system is more malleable with members coming and going. We may hope for more "ex-post over time" properties to ensure no group is consistently receiving the short end of the stick. Furthermore, we may hope to allow for local changes to occur without completely refreshing all assemblies simultaneously, say rotating people in and out one at a time, and doing so with minimal changes to assemblies on the

opposite side of the graph. Modeling this well and defining useful "over-time" fairness properties seems to be challenging yet potentially impactful future work.

Finally, although we analytically solve well-motivated special cases, we leave open whether a randomized assignment satisfying all of our desiderata exists in the general case. In our extensive experiments, we have not found any infeasible instances, and we are therefore optimistic that existence can be guaranteed.

References

[1] A. A. Ageev and M. I. Sviridenko. Pipage rounding: A new method of constructing algorithms with proven performance guarantee. *Journal of Combinatorial Optimization*, 8:307–328, 2004.

[2] M. L. Balinski and H. P. Young. *Fair Representation: Meeting the Ideal of One Man, One Vote*. Rowman & Littlefield, 2010.

[3] G. Benadè, P. Gölz, and A. D. Procaccia. No stratification without representation. In *Proceedings of the 20th ACM Conference on Economics and Computation (EC)*, pages 281–314, 2019.

[4] G. Birkhoff. Three observations on linear algebra. *Universidad Nacional de Tucumán, Revista A*, 5:147–151, 1946.

[5] K. R. W. Brewer and M. Hanif. An introduction to sampling with unequal probabilities. In *Sampling With Unequal Probabilities*, pages 1–19. Springer, 1983.

[6] E. Budish, Y.-K. Che, F. Kojima, and P. Milgrom. Designing random allocation mechanisms: Theory and applications. *American Economic Review*, 103(2):585–623, 2013.

[7] N. Clegg. Bringing people together to inform decision-making on generative AI. Blog post, 2023. URL https://about.fb.com/news/2023/06/generative-ai-community-forum.

[8] S. Ebadian, G. Kehne, E. Micha, A. D. Procaccia, and N. Shah. Is sortition both representative and fair? In *Proceedings of the 36th Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2022.

[9] J. Fishkin. *Democracy When the People Are Thinking: Revitalizing Our Politics Through Public Deliberation*. Oxford University Press, 2018.

[10] B. Flanigan, P. Gölz, A. Gupta, and A. D. Procaccia. Neutralizing self-selection bias in sampling for sortition. In *Proceedings of the 34th Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2020.

[11] B. Flanigan, P. Gölz, A. Gupta, B. Hennig, and A. D. Procaccia. Fair algorithms for selecting citizens' assemblies. *Nature*, 596:548–552, 2021.

[12] B. Flanigan, G. Kehne, and A. D. Procaccia. Fair sortition made transparent. In *Proceedings of the 35th Annual Conference on Neural Information Processing Systems (NeurIPS)*, pages 25720–25731, 2021.

[13] B. Flanigan, J. Liang, A. D. Procaccia, and S. Wang. Manipulation-robust selection of citizens' assemblies. In *Proceedings of the 38th AAAI Conference on Artificial Intelligence (AAAI)*, pages 9696–9703, 2024.

[14] R. Gandhi, S. Khuller, S. Parthasarathy, and A. Srinivasan. Dependent rounding and its applications to approximation algorithms. *Journal of the ACM*, 53(3):324–360, 2006.

[15] N. Jones and H. Marsden. Associative democracy. In S. Toepler H. Anheier, editor, *International Encyclopedia of Civil Society*. Springer, 2010.

[16] H. Landemore. *Open Democracy: Reinventing Popular Rule for the Twenty-First Century*. Princeton University Press, 2020.

[17] S. Park, M. Specter, N. Narula, and R. L. Rivest. Going from bad to worse: From internet voting to blockchain voting. *Journal of Cybersecurity*, 7(1):tyaa025, 2021.

[18] M. Pivato. Pyramidal democracy. *Journal of Public Deliberation*, 5(1): article 8, 2009.

[19] A. D. Procaccia. A more perfect algorithm. *Scientific American*, 327(5):52–59, 2022.

[20] F. Pukelsheim. *Proportional Representation*. Springer, 2017.

[21] P. Raghavan and C. D. Tompson. Randomized rounding: A technique for provably good algorithms and algorithmic proofs. *Combinatorica*, 7(4):365–374, 1987.

[22] E. Shapiro. Grassroots distributed systems: Concept, examples, implementation and applications. *arXiv preprint arXiv:2301.04391*, 2023.

[23] E. Shapiro. A grassroots architecture to supplant global digital platforms by a global digital democracy. *arXiv preprint arXiv:2404.13468*, 2024.

[24] Ehud Shapiro. Grassroots social networking: Serverless, permissionless protocols for Twitter/LinkedIn/WhatsApp. In *Proceedings of the 3rd International Workshop on Open Challenges in Online Social Networks (OASIS)*, 2023.

[25] P. Stone. *The Luck of the Draw: The Role of Lotteries in Decision Making*. Oxford University Press, 2011.

[26] D. Van Reybrouck. *Against Elections: The Case for Democracy*. Random House, 2016.

A Extending Algorithm 1 to Smaller Equivalence Classes

We will analyze here the probability of Algorithm 1 needing more than $|C^L|$ people from each equivalence class C^L. It is sufficient to ensure this does not occur in the leaves, as in the proof of Theorem 1 internal nodes select fewer people from $|C^L|$ than their children.

Fix an equivalence class C^L and a leaf node $\ell \in L$. We can directly analyze the probability that more than $|C^L|$ people are selected from C^L for this assembly. Namely, for each of the n draws, the probability that it was a member from C^L is $|C^L|/|N_\ell|$. We are doing n draws of this. Hence, we wish to analyze the probability of $\Pr[\sum_{i=1}^{n} X_i > |C^L|]$, where each X_i is drawn from an independent Bernoulli with $\Pr[X_i = 1] = |C^L|/|N_\ell|$. By Chernoff bound says that for all X which is the sum of independent variables with $\mu = \mathbb{E}[X]$ and $\delta \geq 0$,

$$\Pr[X > (1+\delta)\mu] < \left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}}\right)^\mu \leq \left(\frac{e}{1+\delta}\right)^{(1+\delta)\mu}.$$

For our purposes, $\mu = n|C^L|/|N_\ell|$, and we wish to set δ such that $\mu(1+\delta) = |C^L|$, so $1+\delta = N^|ell|/n$. This implies that the probability of failure is at most

$$\left(\frac{e \cdot n}{|N_\ell|}\right)^{|C^L|}.$$

Suppose additionally that populations are not too small in the sense that $|N_\ell| \geq q|N|$ for some value q. Then, this is at most

$$\left(\frac{e \cdot n}{q \cdot |N|}\right)^{|C^L|}.$$

Finally, to ensure this happens for no leaves or equivalence classes, we need to union bound over the at most $|G|$ leaf nodes and $|N|$ equivalence classes, leading to a total probability of failure of at most

$$|N||G| \cdot \left(\frac{e \cdot n}{q|N|}\right)^{|C^\ell|}.$$

Now, suppose each $|C^\ell|$ is of size at most 3. Furthermore, assume that $|N| \gg n, 1/q, |G|$, which we would expect for reasonable instances. Thus, it makes sense to interpret this asymptotically as $|N|$ grows large compared to the rest of the terms. This leads to a bound of failure of $O(1/|N|^2)$, essentially negligible compared to all ex-ante guarantees (which are $\Theta(1/|N|)$).

B Proof of Theorem 2

Showing inheritance, ex ante child representation, and ex post child representation follow immediately from the definition of the algorithm.

It remains to establish individual representation. Fix a person i, and let $\ell \in \text{LEAVES}(G)$ be the leaf node they are signed up for. Note that A_ℓ is simply a random sample of n people from N_ℓ, so clearly $\Pr[i \in \mathcal{A}_\ell] = n/|N_\ell|$. Next, fix a federation $f^* \in \text{FEDS}(G)$ such that $i \in N_{f^*}$. Consider running the algorithm in a different order, sampling all vectors $(s_{c,f})_{c \in \text{CHILDREN}f}$ at the beginning before starting the algorithm, and then running according to these samples. Note that this leads to an equivalent process because each $(s_{c,f})_{c \in \text{CHILDREN}f}$ is sampled independently of everything else in the algorithm. Condition on a specific sample of these vectors. Let $\ell = v_0, v_1, \ldots, v_k = f$ be the path in G that leads from ℓ to f. Note that we can now directly compute the probability $i \in A_{f^*}$ because the only way to do so is if $i \in A_\ell$ and $i \in B_{v_{j-1},v_j}$ for each $j \geq 1$. Hence, this probability is exactly

$$\frac{n}{|N_\ell|} \cdot \prod_{j=1}^{k} \frac{s^{v_{j-1},v_j}}{n}.$$

To get the unconditional probability, we can simply take the expectation over all s values, i.e.,

$$\mathbb{E}\left[\frac{n}{|N_\ell|} \cdot \prod_{j=1}^{k} \frac{s_{v_{j-1}}^{v_j}}{n}\right].$$

Since each $s_{c,f}$ and $s_{c',f'}$ is sampled independently for $f \neq f'$, we can push the expectation in to get equality to

$$\frac{n}{|N_\ell|} \cdot \prod_{j=1}^{k} \frac{\mathbb{E}\left[s^{v_{j-1},v_j}\right]}{n} = \frac{n}{|N_\ell|} \cdot \prod_{j=1}^{k} \frac{n \cdot q_{v_{j-1},v_j}}{n} = \frac{n}{|N_\ell|} \prod_{j=1}^{k} \frac{|N_{v_{j-1}}|}{|N_{v_j}|} = \frac{n}{|N_{f^*}|},$$

as needed. □

C Impossibility for Iterative Algorithms

To formalize this impossibility result, we call an algorithm *topologically iterative* if it has the same structure as Algorithm 2, except line 2 and 5 are replaced with potentially different sampling schemes.

Proposition 1. *For all assembly sizes n, there exist instances both with G being a tree (but people signed up for multiple leaves) or people signed up for a single leaf (but G not being a tree) where no topologically iterative algorithm can simultaneously satisfy individual representation and ex ante child representation.*

Proof. We begin with an instance where G is a tree, but some people are signed up for multiple leaves. Fix an assembly size n. There will be a single federation f with $2n$ children $c_1, \ldots, c_{2n}$ which are all leaves. We will define populations by equivalence classes. There will be $2n$ sets of people $C^{\{c_j\}}$ of equal size that each are only signed up for c_j, i.e., $|C^{\{c_j\}}| = k$ for some arbitrary integer k. In addition, there will be a set $C^{\{c_1,\ldots,c_{2n}\}}$ of people signed up for all children assemblies. This will be $2n-1$ times as large as each individual group, so $|C^{\{c_1,\ldots,c_{2n}\}}| = (2n-1)k$. Note that, by symmetry, $q_{c_j,f} = 1/(2n)$ for each j.

Suppose our strategy now is to (1) sample $(A_{c_1}, \ldots, A_{c_{2n}})$ (from some distribution satisfying individual representation) (2) independently, sample an integral vector $(s_1, \ldots, s_{2n})$ such that $\sum_j s_j = n$ and $\mathbb{E}[s_j] = n/(2n) = 1/2$, and then (3) choose A_f by selecting s_j people from A_{c_j}. We now claim that we cannot select A_f such that it satisfies individual representation. Indeed fix an arbitrary $i \in C^{\{c_1\}}$. A necessary condition for $i \in A_f$ is that both $i \in A_{c_1}$ and $s_1 \geq 1$. Note that $\Pr[i \in \mathcal{A}_{c_1}] = \frac{n}{|C^{c_1}|+|C^{\{c_1,\ldots,c_{2n}\}}|} = \frac{1}{2k}$. Furthermore, $\Pr[s_1 \geq 1] \leq 1/2$ by Markov's inequality. Since these are selected independently, the probability they both occur is at most $\frac{1}{4k}$. However, individual representation ensures that $\Pr[i \in \mathcal{A}_f] = \frac{n}{|C^{\{c_1,\ldots,c_{2n}\}}|+\sum_j |C^{\{c_j\}}|} = \frac{n}{(4n-1)k} > \frac{1}{4k}$.

To convert this to an instance where G is not a tree, we can have an additional node for each equivalence class, and have these nodes point to the set of leaf assemblies that class was signed up for. Then, each person can sign up for this single corresponding leaf node rather than a larger set of nodes, and the same argument goes through. □

D Algorithm 3 Helper Functions

Here, we formalize the helper functions used in Algorithm 3, presented as Algorithm 4. All of the randomized rounding can be done using a variation of the Birkhoff Von-Neumonn Theorem. Namely, Budish et al. [6] give an algorithm to handle the following randomized rounding instances. Say we are given a set of values $p_1, \ldots, p_k$ and a set of *constraints* represented as a family of sets $\mathcal{I}$ where each $I \subseteq \mathcal{I}$ has $I \subseteq \{1, \ldots, k\}$. Our goal is to round these values to $x_1, \ldots, x_k$ such that $\mathbb{E}[x_i] = p_i$. Furthermore, we will do this such that for all $I \in \mathcal{I}$, $\lfloor \sum_{i \in I} p_i \rfloor \leq \sum_{i \in I} x_i \leq \lceil \sum_{i \in I} p_i \rceil$. Budish et al. [6] show that if $\mathcal{I}$ is a *bihierarchy*, then this is possible and can be done with a polynomial time algorithm. $\mathcal{I}$ is said to be a bihierarchy of there exists a partition $\mathcal{I} = \mathcal{I}_1 \cup \mathcal{I}_2$ such that for all pairs $I, I' \in \mathcal{I}_j$ of either partition, either $I \subseteq I'$, $I' \subseteq I$, or $I \cap I' = \emptyset$. One can check that all randomized roundings we do in the helper functions take this form.

E Proof of Theorem 3

There are a variety of conditions we must check, namely, that Algorithm 3 can successfully run to completion, it returns a valid assembly assignment, and that the returned assignment satisfies our ex ante and ex post guranteess

Algorithm 4 Helper functions

function ROUND(x)
  return $\lceil x \rceil$ with probability $x - \lfloor x \rfloor$ and $\lfloor x \rfloor$ otherwise.
function SAMPLELEAVES($(C^T)_{T \subseteq \mathcal{T}}$, $(s_t)_{t \in \mathcal{T}}$, n)
  $p_t^T \leftarrow \frac{|C^T|/|T|}{\sum_{T':t \in T'} |C^{T'}|/|T'|} \cdot s_t$ for all $T \subseteq \mathcal{T}$ with nonempty C^T and $t \in T$.
  Round all $(p_t^T)_{t,T}$ to $(\alpha_t^T)_{t,T}$ such that $\mathbb{E}[\alpha_t^T] = p_t^T$, $\alpha_t^T \leq \lceil p_t^T \rceil$,$\forall t, \sum_{T:t \in T} \alpha_t^T = s_t$, and $\forall T, \sum_{t \in T} \alpha_t^T \leq \lceil \sum_{t \in T} p_t^T \rceil$.
  for $T \subseteq \mathcal{T}$ **do**
    Choose $(D_t^T)_{t \in T}$ of sizes $(\alpha_t^T)_{t \in T}$ randomly from C^T such they are each disjoint
  for $t \in \mathcal{T}$ **do**
    $B_t^{\text{SEL}} \leftarrow \bigcup_{T:t \in T} D_t^T$
  for $t \in \mathcal{T}$ **do**
    $q_t^T \leftarrow \frac{|C^T|(1-1/|T|)}{\sum_{T':t \in T'} |C^{T'}|(1-1/|T'|)} \cdot (n - s_t)$ for all $T \subseteq \mathcal{T}$ with nonempty C^T and $t \in T$
    Round $(q_t^T)_T$ to (β_t^T) such that each $\beta_t^T \leq \lceil q_t^T \rceil$, $\sum_T \beta_t^T = n - s_t$, and $\mathbb{E}[\beta_t^T] = q_t^T$.
    for $T : t \in T$ **do**
      Choose E_t^T be a random sample of β_t^T members of $C^T \setminus \{D_t^T\}$.
    $B_t^{\text{UNS}} \leftarrow \bigcup_{T:t \in T} E_t^T$.
  return $(B_t^{\text{SEL}}, B_t^{\text{UNS}})_{t \in \mathcal{T}}$
function SAMPLEFROMCHILDREN(s, n, $(B_c^{\text{SEL}}, B_c^{\text{UNS}}, w_c, |N_c|)_c$)
  Let $x_c^{\text{SEL}} = s \cdot \frac{w_c}{\sum_{c'} w_{c'}}$ and $x_c^{\text{UNS}} = (n - s)\frac{|N_c| - w_c}{\sum_{c'} |N_{c'}| - w_{c'}}$ for all c.
  Round each x_c^j for $j \in \{\text{SEL}, \text{UNS}\}$ and child c to γ_c^j such that $\mathbb{E}[\gamma_c^{\text{SEL}}] = x_c^{\text{SEL}}$, $\gamma_c^{\text{SEL}} \geq \lfloor x_c^{\text{SEL}} \rfloor$, for $\sum_c \gamma_c^{\text{SEL}} = s$, $\sum_c \gamma_c^{\text{UNS}} = n - s$, and for all c, $\gamma_c^{\text{SEL}} + \gamma_c^{\text{UNS}} \geq \lfloor x_c^{\text{SEL}} + x_c^{\text{UNS}} \rfloor$.
  for all c **do**
    Let D_c^{SEL} be a random sample of size γ_c^{SEL} from B_c^{SEL}
    Let D_c^{UNS} be a random sample of size γ_c^{UNS} from B_c^{UNS}.
  $B_c^{\text{SEL}} \leftarrow \bigcup_c D_c^{\text{SEL}}$
  $B_c^{\text{UNS}} \leftarrow \bigcup_c D_c^{\text{UNS}}$
  return $B_c^{\text{SEL}}, B_c^{\text{UNS}}$
function ROUNDANDSAMPLE(SIZE, $((S_1, x_1), \ldots (S_k), (x_k))$
  Let $(y_1, \ldots, y_k) = \text{SIZE} \cdot (\frac{x_1}{\sum_j x_j}, \ldots, \frac{x_k}{\sum_j x_j})$
  Round $(x_1, \ldots, x_k)$ to $(\gamma_1, \ldots, \gamma_k)$ such that $\mathbb{E}[\gamma_i] = x_i$ and $\lfloor x_i \rfloor \leq \gamma_i \leq \lceil x_i \rceil$ and $\sum_i \gamma_i = \text{SIZE}$.
  for $i = 1, \ldots, k$ **do**
    Let $D_i \subseteq B_i$ be a random sample of size γ_i
  return $\bigcup_i D_i$

Fix an arbitrary rounding $s_{r,t}$ for $r \in R$ and $t \in \mathcal{T}$ from line 2. We will take for now that these are fixed constants satisfying $\lfloor n \cdot \frac{w_{r,t}}{|N_{r,t}|} \rfloor \leq s_{r,t} \leq \lceil n \cdot \frac{w_{r,t}}{|N_{r,t}|} \rceil$, only considering randomness from other sampling. In the end, we will deal with the randomness over these $s_{r,t}$ values as well to prove some ex ante guarantees.

We will show for all $(r, t) \in R \times \mathcal{T}$, the following properties hold:

1. Each $B_{r,t}^{\text{SEL}}, B_{r,t}^{\text{UNS}} \subseteq N_{(r,t)}$.
2. Each $|B_{r,t}^{\text{SEL}}| = s_{r,t}$, $|B_{r,t}^{\text{UNS}}| = n - s_{r,t}$.
3. Each pair $B_{r,t}^{\text{SEL}}$ and $B_{r,t}^{\text{UNS}}$ are disjoint.
4. All of $(B_{r,t}^{\text{SEL}})_{t \in \mathcal{T}}$ are pairwise disjoint.
5. For $i \in N_{(r,t)}$, $\Pr[i \in B_{r,t}^{\text{SEL}}] = \frac{s_{r,t}}{|L_i| \cdot w_{r,t}}$ and $\Pr[i \in B_{r,t}^{\text{UNS}}] = \frac{(n - s_{r,t})(1 - 1/|L_i|)}{|N_{r,t}| - w_{r,t}}$.

Fix a topic $t \in T$. We will show these properties for all r by structural induction on the graph R. We begin by proving it is the case for leaf nodes $r \in \text{LEAVES}(R)$. Fix such an r.

For such an r, we simply need to analyze the SAMPLELEAVES function, showing that it can run successfully and produces an output satisfying the desired properties.

Properties 1– 4 are immediate from the algorithm assuming it can run to completion. However, we must show that the random choices made on lines 7 and 14 are feasible, in the sense that the set we are selecting from has enough people. To that end, fix a set $T \subseteq \mathcal{T}$, and consider the execution of line 7. Note that this is successful as long as $\sum_t \alpha_t^T \le |C^T|$. Indeed, we have $\sum_t \alpha_t^T \le \lceil \sum_{t \in T} p_t^T \rceil$ by assumption of the rounding. Expanding this, for each t,

$$p_t^T = \frac{|C^T|/|T|}{w_{r,t}} \cdot s_t \le \frac{|C^T|/|T|}{w_{r,t}} \cdot \left\lceil \frac{n \cdot w_{r,t}}{|N_{r,t}|} \right\rceil.$$

Let $\lambda = w_{r,t}/|N_{r,t}|$. We have that $|N_{r,t}| \ge 4n$, so this is at most

$$\frac{|C^T|/|T|}{4\lambda n} \cdot \lceil \lambda n \rceil.$$

Now, since $\lambda \ge \varepsilon$, $\lambda n \ge \frac{2\lambda}{\varepsilon\delta} \ge 1$. Thus $\frac{\lceil \lambda n \rceil}{2\lambda n} \le 1$, and we have $p_t^T \le |C^T|/|T|$. Summing up over t, we get that $\sum_{t \in T} p_t^T \le |C^T|$. Thus, $|C^T| \ge \lceil |C^T| \rceil = \lceil \sum_{t \in T} p_t^T \rceil$ since $|C^T|$ is an integer.

Next, fix r, $t \in \mathcal{T}$, and T with $t \in T$, and consider the run of line 14 with these parameters. Note that this sample will be successful as long as $|C^T| \ge \alpha_t^T + \beta_t^T$. We additionally have $\alpha^T \le \lceil p_t^T \rceil$ and $\beta_t^T \le \lceil q_t^T \rceil$. Therefore, we have

$$\begin{aligned}
&\alpha_t^T + \beta_t^T \\
&\le \lceil p_t^T + q_t^T \rceil + 1 \\
&\le \left\lceil \frac{|C^T|/|T|}{w_{r,t}} \left\lceil \frac{n \cdot w_{r,t}}{|N_{r,t}|} \right\rceil + \frac{|C^T|(1 - 1/|T|)}{|N_{r,t}| - w_{r,t}} \left\lceil \frac{n \cdot (|N_{r,t}| - w_{r,t})}{|N_{r,t}|} \right\rceil \right\rceil + 1 \\
&= \left\lceil |C^T| \left((1/|T|) \cdot \left(\frac{1}{w_{r,t}} \cdot \left\lceil \frac{n \cdot w_{r,t}}{|N_{r,t}|} \right\rceil \right) \right. \right. \\
&\qquad \left. \left. + (1 - 1/|T|) \left(\frac{1}{|N_{r,t}| - w_{r,t}} \right) \left\lceil \frac{n \cdot (|N_{r,t}| - w_{r,t})}{|N_{r,t}|} \right\rceil \right) \right\rceil + 1.
\end{aligned}$$

Now, since $(1/|T|) + (1 - 1/|T|) = 1$, this is a convex combination between two terms. Thus, we may upper bound it by the maxmium of the two as

$$\le \left\lceil |C^T| \max \left(\frac{1}{w_{r,t}} \cdot \left\lceil \frac{n \cdot w_{r,t}}{|N_{r,t}|} \right\rceil, \left(\frac{1}{|N_{r,t}| - w_{r,t}} \right) \left\lceil \frac{n \cdot (|N_{r,t}| - w_{r,t})}{|N_{r,t}|} \right\rceil \right) \right\rceil + 1.$$

Again, let $\lambda = w_{r,t}/|N_{r,t}| \le w_{r,t}/4n$, we can expand

$$\frac{1}{w_{r,t}} \cdot \left\lceil \frac{n \cdot w_{r,t}}{|N_{r,t}|} \right\rceil \le \frac{\lceil \lambda n \rceil}{4n\lambda}.$$

Since $n\lambda \ge \varepsilon \cdot \frac{2}{\varepsilon\delta} \ge 2$, we have that $\frac{\lceil \lambda n \rceil}{4n\lambda} \le 1/2$. The same argument applies for the second term swapping λ with $1 - \lambda$, since we also know $1 - \lambda \ge \varepsilon$. Thus, since we have also assumed $|C^T| \ge 2$, this simplifies to $\lceil |C^T|/2 \rceil + 1 \le |C^T|$, as needed.

Finally, we handle property 5. Fix $i \in N_{(r,t)}$ with $L_i = \{r\} \times T$, and note that $t \in T$. Now, by symmetry, each member of C^T is equally likely to be in each of $B_{r,t}^{\text{SEL}}$ and $B_{r,t}^{\text{UNS}}$. Thus, the probability is simply the expected number of seats going to C^T divided by $|C^T|$. This is precisely $p_t^T/|C^T|$ and $q_t^T/|C^T|$, which expand to our desired values.

Next, we proceed to the inductive step. Fix an internal node $r \in \text{FEDS}(R)$ and topic $t \in \mathcal{T}$. Suppose all of the properties hold $B_{c,t}^{\text{SEL}}$ and $B_{c,t}^{\text{UNS}}$ for $c \in \text{CHILDREN}(r)$. To show they hold for this r and t, we must consider the SAMPLEFROMCHILDREN step. We will show it successfully runs, producing $B_{r,t}^{\text{SEL}}$ and $B_{r,t}^{\text{UNS}}$ satisfying all of the properties.

The two places SAMPLEFROMCHILDREN may fail are on lines 21 and 22 if we wish to sample more people than are available. Line 21 can run successfully as long as $\gamma_c^{\text{SEL}} \leq |B_{c,t}^{\text{SEL}}| = s_{c,t}$ for each $c \in \text{CHILDREN}(r)$. Fix such a child c. Note that by the rounding, we have that $\gamma_c^{\text{SEL}} \leq \lceil s \cdot \frac{w_{c,t}}{\sum_{c'} w_{c',t}} \rceil$ Note that the denominator $\sum_{c'} w_{c',t} = w_{r,t}$ since children are disjoint. Furthermore, we have that $s_{r,t} \leq \lceil n \cdot \frac{w_{r,t}}{|N_{r,t}|} \rceil$ and $s_{r,t} \geq \lfloor n \cdot \frac{w_{c,t}}{|N_{c,t}|} \rfloor$. Therefore, it is sufficient to show

$$\left\lceil \left\lceil n \cdot \frac{w_{r,t}}{|N_{r,t}|} \right\rceil \cdot \frac{w_{c,t}}{w_{r,t}} \right\rceil \leq \left\lfloor n \cdot \frac{w_{c,t}}{|N_{c,t}|} \right\rfloor.$$

This is implied by

$$\left\lceil n \cdot \frac{w_{r,t}}{|N_{r,t}|} \right\rceil \cdot \frac{w_{c,t}}{w_{r,t}} + 1 \leq n \cdot \frac{w_{c,t}}{|N_{c,t}|},$$

which itself is implied by

$$(n \cdot \frac{w_{r,t}}{|N_{r,t}|} + 1) \cdot \frac{w_{c,t}}{w_{r,t}} + 1 \leq n \cdot \frac{w_{c,t}}{|N_{c,t}|}.$$

Finally, note that

$$\left(n \cdot \frac{w_{r,t}}{|N_{r,t}|} + 1\right) \cdot \frac{w_{c,t}}{w_{r,t}} + 1 \leq n \cdot \frac{w_{r,t}}{|N_{r,t}|} + 2,$$

so this entire inequality follows from

$$n \cdot \frac{w_{r,t}}{|N_{c,t}|} - n \cdot \frac{w_{r,t}}{|N_{r,t}|} \geq 2.$$

Now, we have that

$$n \cdot \frac{w_{r,t}}{|N_{c,t}|} - n \cdot \frac{w_{r,t}}{|N_{r,t}|} = n \cdot \frac{w_{r,t}}{|N_{c,t}|}(1 - \frac{|N_{c,t}|}{|N_{r,t}|}) \geq n \cdot \varepsilon \cdot \delta \geq 2,$$

by assumption on n.

A similar argument works for line 22. Fix $c \in \text{CHILDREN}(r)$. We need to show that

$$\left\lceil (n - s_{r,t}) \cdot \frac{|N_{c,t}| - w_{c,t}}{\sum_{c'} |N_{c',t}| - w_{c',t}} \right\rceil \leq n - s_{c,t}$$

for all $c \in \text{CHILDREN}(r)$. Expanding definitions, this is implied by

$$\left\lceil \left\lceil n \left(\frac{|N_{r,t}| - w_{r,t}}{|N_{r,t}|} \right) \right\rceil \frac{|N_{c,t}| - w_{c,t}}{|N_{r,t}| - w_{r,t}} \right\rceil \leq \left\lfloor n \cdot \frac{|N_{c,t}| - w_{c,t}}{|N_{c,t}|} \right\rfloor.$$

Doing the same expansion on the ceilings, this inequality is implied by

$$n \cdot \frac{|N_{c,t}| - w_{c,t}}{|N_{r,t}|} + 2 \leq n \cdot \frac{|N_{c,t}| - w_{c,t}}{|N_{c,t}|}.$$

Finally, we have that

$$n \cdot \frac{|N_{c,t}| - w_{c,t}}{|N_{c,t}|} - n \cdot \frac{|N_{c,t}| - w_{c,t}}{|N_{r,t}|} = n \cdot \frac{|N_{c,t}| - w_{c,t}}{|N_{c,t}|} \cdot \left(1 - \frac{|N_{c,t}|}{|N_{r,t}}\right) \geq n \cdot \varepsilon \cdot \delta \geq 2.$$

Now that we have proved the function runs successfully, we prove the properties. We have that $B_{r,t}^{\text{SEL}} \subseteq \bigcup_{c \in \text{CHILDREN}(r)} B_{c,t}^{\text{SEL}} \subseteq \bigcup_{c \in \text{CHILDREN}(r)} N_{(c,t)} = N_{(r,t)}$, and a symmetric argument holds for $B_{r,t}^{\text{SEL}}$. The size holds because the sets $B_{c,t}^{\text{SEL}}$ are pairwise disjoint, so we never select the same person twice, and end up with $s_{r,t}$ distinct people. Again, a symmetric argument works for $B_{r,t}^{\text{UNS}}$. The disjointness holds trivially because it held for the children sets, along with distinct child regions having disjoint populations. Finally, fix $i \in N_{(r,t)}$, and suppose $c \in \text{CHILDREN}(r)$ is the unique child such that $i \in N_c$. Then the only way i can be in $B_{(r,t)}^{\text{SEL}}$ is if i was selected to $B_{(c,t)}^{\text{SEL}}$, and is then subsequently in the subset selected to join $B_{(r,t)}^{\text{SEL}}$. The probability of being in $B_{(c,t)}^{\text{SEL}}$ is $\frac{s_{c,t}}{|L_i| \cdot w_{c,t}}$ by induction. Furthermore, we sample γ_c^{SEL} with $\mathbb{E}[\gamma_c^{\text{SEL}}] = \frac{s_{r,t}}{w_{c,t}} w_{r,t}$ and selected γ_c^{SEL} out of the $s_{c,t}$

people uniformly at random from $B^{\text{SEL}}_{(c,t)}$. Thus conditioned on $i \in B^{\text{SEL}}_{(c,t)}$, the probability i is in $B^{\text{SEL}}_{(s,t)}$ is $\frac{s_{r,t}}{w_{c,t}} w_{r,t}$. Hence, the overall probability $i \in B^{\text{SEL}}_{(s,t)}$ is

$$\frac{s_{c,t}}{|L_i| \cdot w_{c,t}} \cdot \frac{\frac{s_{r,t}}{w_{c,t}} w_{r,t}}{s_{c,t}} = \frac{s_{r,t}}{|L_i| \cdot w_{r,t}},$$

as needed. A symmetric argument holds for $B^{\text{UNS}}_{(r,t)}$.

Finally, there is one more place where the algorithm can potentially fail, which is in the RoundAndSample call on line 30. To ensure that this line can run, we need to make sure that for each region r and topic t, $|B^{\text{SEL}}_{r,t}| = s_{r,t} \leq \lceil n \cdot \frac{w_{r,t}}{|N_{(r,t)}|} \rceil$ is sufficiently large to be able to sample from. Note that we will take at most $\lceil n \cdot \frac{w_{r,t}}{\sum_{t'} w_{r,t'}} \rceil$ people. Furthermore, this denominator is $|N_r|$. So, we simply need to show

$$\left\lceil n \cdot \frac{w_{r,t}}{|N_r|} \right\rceil \leq \left\lfloor n \cdot \frac{w_{r,t}}{|N_{r,t}|} \right\rfloor.$$

Note that since $\left\lfloor n \cdot \frac{w_{r,t}}{|N_{r,t}|} \right\rfloor$ is an integer, this is implied by

$$n \cdot \frac{w_{r,t}}{|N_r|} \leq \left\lfloor n \cdot \frac{w_{r,t}}{|N_{r,t}|} \right\rfloor.$$

Again, let $\lambda = \frac{w_{r,t}}{|N_{r,t}|}$. We then have

$$\begin{aligned} n \cdot \frac{w_{r,t}}{|N_r|} &= n \cdot \lambda \cdot \frac{|N_{r,t}|}{|N_r|} \\ &\leq n \cdot \lambda \cdot (1 - \delta) \\ &= n\lambda - n\lambda\delta. \end{aligned}$$

Now, since $\lambda \geq \varepsilon$, $n\lambda\delta \geq 2$. Therefore, this is at most

$$n\lambda - 2 \leq \lfloor n\lambda \rfloor = \left\lfloor n \cdot \frac{w_{r,t}}{|N_{r,t}|} \right\rfloor,$$

as needed.

We have now shown that Algorithm 3 can execute to completion. In what we remains we show all of the properties that it satisfies. First, note that each assembly in the final assignment is of size n because it is composed of a disjoint union of sets whose sizes add up to n.

Next, we show that approximate ex post child representation holds. Fix a region r. We first consider $(r, *)$. Note that for each t, we select at least $\lfloor n \cdot \frac{w_{r,t}}{|N_r|} \rfloor$ members from $B^{\text{SEL}}_{(r,t)} \subseteq A_{(r,t)}$. Thus, $A_{(r,t)} \cap A_{(r,*)} \geq \lfloor n \cdot \frac{w_{r,t}}{|N_r|} \rfloor$, the exact child representation guarantees. Now fix an internal region $r \in \text{FEDS}(R)$, a child $c \in \text{CHILDREN}(R)$ and a topic $t \in T$. We show that $|A_{(c,t)} \cap A_{(r,t)}| \geq \lfloor n \cdot \frac{N_{(c,t)}}{N_{(r,t)}} \rfloor - 1$. Indeed, $|A_{(c,t)} \cap A_{(r,t)}| = |B^{\text{SEL}}_{c,t} \cap B^{\text{SEL}}_{(r,t)}| + |B^{\text{UNS}}_{c,t} \cap B^{\text{UNS}}_{(r,t)}|$. These two sizes are randomly selected on line 19 of SelectFromChildren, and their sum will be $\gamma^{\text{SEL}}_c + \gamma^{\text{UNS}}_c$. There is a constraint that

$$\gamma^{\text{SEL}}_c + \gamma^{\text{UNS}}_c \geq \lfloor x^{\text{SEL}}_c + x^{\text{UNS}}_c \rfloor = \left\lfloor s_{r,t} \cdot \frac{w_{c,t}}{w_{r,t}} + (n - s_{r,t}) \cdot \frac{|N_{(c,t)}| - w_{c,t}}{|N_{(r,t)}| - w_{r,t}} \right\rfloor.$$

We will show that

$$s_{r,t} \cdot \frac{w_{c,t}}{w_{r,t}} + (n - s_{r,t}) \cdot \frac{|N_{(c,t)}| - w_{c,t}}{|N_{(r,t)}| - w_{r,t}} \leq \frac{|N_{(c,t)}|}{|N_{(r,t)}|} - 1. \tag{1}$$

which implies the ex post guarantees. Note that if $s_{(r,t)}$ were equal to its expectation $n \cdot \frac{w_{r,t}}{|N_{(r,t)}|}$, then the sum simplifies to exactly $\frac{|N_{(c,t)}|}{|N_{(r,t)}|}$. Note that by rounding $s_{(r,t)}$ to a neighboring integer, one of $s_{(r,t)}$ and $(n - s_{r,t})$ is larger than its expectation, and the other is smaller. However, this different is at most one. Furthermore, they are multipled by either $\frac{w_{c,t}}{w_{r,t}}$ or $\frac{|N_{(c,t)}| - w_{c,t}}{|N_{(r,t)}| - w_{r,t}}$, terms which are at most

1. Hence, even after this rounding, the sum inside can differ from the expectation by at most one, and we get the lower bound of $\frac{|N_{(c,t)}|}{|N_{(r,t)}|} - 1$, as needed.

Next, we consider ex ante child representation. For a node $(r, *)$, this is straightforward, as we directly round to get in expectation $n \cdot \frac{w_{(r,t)}}{|N_r|}$ selected from $B^{\text{SEL}}_{r,t} \subseteq A_{(r,t)}$. For an internal $r \in \text{FEDS}(R)$, child $c \in \text{CHILDREN}(r)$, and topic $t \in \mathcal{T}$, the overlap of $|A_{(r,t)} \cap A_{(c,t)}|$ is precisely the number of people selected from $B^{\text{SEL}}_{c,t}$ for $B^{\text{SEL}}_{r,t}$ plus the number of people selected from $B^{\text{UNS}}_{c,t}$ for $B^{\text{UNS}}_{r,t}$, as these sets are disjoint. The expected number of each is $s_{(r,t)} \cdot \frac{w_{c,t}}{w_{r,t}}$ and $(n - s_{(r,t)}) \cdot \frac{|N_{(c,t)}| - w_{c,t}}{|N_{(r,t)}| - w_{r,t}}$. If we take expectation over $s_{(r,t)}$ as well, by linearity, we have that the expected overlap is

$$n \cdot \frac{w_{r,t}}{|N_{(r,t)}|} \cdot \frac{w_{c,t}}{w_{r,t}} + \left(n - n \cdot \frac{w_{r,t}}{|N_{(r,t)}|}\right) \cdot \frac{|N_{(c,t)}| - w_{c,t}}{|N_{(r,t)}| - w_{r,t}} = n \cdot \frac{|N_{(c,t)}|}{|N_{(r,t)}|},$$

as needed.

Finally, we show individual representation. First, fix a node $(r, *)$ and a person $i \in N_r$ signed up for leaf nodes $\{r'\} \times T$. For each $t \in T$, we have that $i \in B^{\text{SEL}}_{(r,t)}$ with probability $\frac{s_{(r,t)}}{|T| \cdot w_{(r,t)}}$. In expectation, $n \cdot \frac{w_{(r,t)}}{|N_r|}$ will be selected from $B^{\text{SEL}}_{(r,t)}$ to be in $A_{(r,*)}$. hence, conditioned on $i \in B^{\text{SEL}}_{(r,t)}$, i will be selected with probability $\frac{n \cdot \frac{w_{(r,t)}}{|N_r|}}{s_{r,t}}$. This means the total probability of i being in both $B^{\text{SEL}}_{(r,t)}$ and $A_{(r,*)}$ is $\frac{n}{|T| \cdot |N_r|}$. Note that i being in each of $B^{\text{SEL}}_{(r,t)}$ cannot happen simultaneously, because these sets are disjoint. Therefore, the total probability i is in $A_{(r,*)}$ is the sum of the probabilities of all of these events, and therefore $\frac{n}{|N_r|}$.

Finally, consider a node (r, t), and a member $i \in N_{(r,t)}$. We have that the probability of $i \in B^{\text{SEL}}_{(r,t)}$ is $\frac{s_{r,t}}{|L_i| \cdot w_{r,t}}$ and the probability of $i \in B^{\text{UNS}}_{(r,t)}$ is $\frac{(n - s_{r,t})(1 - 1/|L_i|)}{|N_{(r,t)}| - w_{r,t}}$. Note that these are mutually exclusive events because the sets are disjoint. Hence, the total probability of $i \in A_{(r,t)}$ is the sum

$$\frac{s_{r,t}}{|L_i| \cdot w_{r,t}} + \frac{(n - s_{r,t})(1 - 1/|L_i|)}{|N_{(r,t)}| - w_{r,t}}.$$

By lienarity of expectation over the sampling of $s_{r,t}$, we have that the complete probability is

$$\frac{n \cdot \frac{w_{r,t}}{|N_{(r,t)}|}}{|L_i| \cdot w_{r,t}} + \frac{(n - n \cdot \frac{w_{r,t}}{|N_{(r,t)}|})(1 - 1/|L_i|)}{|N_{(r,t)}| - w_{r,t}} = \frac{n}{|L_i|}|N_{(r,t)}| + \frac{n \cdot (1 - 1/|L_i|)}{|N_{(r,t)}|} = \frac{n}{|N_{(r,t)}|},$$

showing individual representation. □

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification:

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification:

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory Assumptions and Proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

Justification:

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental Result Reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification:

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [Yes]

Justification:

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (https://nips.cc/public/guides/CodeSubmissionPolicy) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so "No" is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (https://nips.cc/public/guides/CodeSubmissionPolicy) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental Setting/Details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

   Answer: [Yes]

   Justification:

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment Statistical Significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

   Justification: We show scatter plots of all samples.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.

- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments Compute Resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification:

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code Of Ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics https://neurips.cc/public/EthicsGuidelines?

   Answer: [Yes]

   Justification:

   Guidelines:

   - The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader Impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Justification: Our work does not have plausible negative societal impacts, as far as we can see.

    Guidelines:

    - The answer NA means that there is no societal impact of the work performed.
    - If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.

- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

    Answer: [NA]

    Justification:

    Guidelines:

    - The answer NA means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [NA]

    Justification:

    Guidelines:

    - The answer NA means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.

- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New Assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [NA]

    Justification:

    Guidelines:

    - The answer NA means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and Research with Human Subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [NA]

    Justification:

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional Review Board (IRB) Approvals or Equivalent for Research with Human Subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [NA]

    Justification:

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.