PROPORTIONAL HIERARCHICAL CLUSTERING

Anonymous authors
Paper under double-blind review

ABSTRACT

Hierarchical clustering represents data at multiple levels of granularity, but proportional fairness is traditionally only defined for a single partition. We therefore study how it can be extended to an entire hierarchy. Building on prior work on proportional clustering, we introduce two notions: the local core, which requires proportional fairness at each split, and the global core, which requires it at every partition in the hierarchy. We study both notions under the maximum loss and average loss objectives, where an agent's loss is, respectively, the maximum or average distance to other members of their cluster. We show that the two core notions are incomparable and that standard agglomerative clustering techniques fail to satisfy them. However, under the maximum loss objective, a hierarchical clustering in the local core always exists. On the way, we resolve an open question from Caragiannis et al. (2024) by showing that there always exists a non-centroid clustering in the core for $k = 2$, whereas for $k \geq 3$ the core can be empty (even for Euclidean distances). For the global core, we show that even arbitrary approximations of it need not exist. However, doubling the coalition size threshold allows us to compute an outcome in the global 16-core for maximum-loss and global $8n$-core for average-loss, where n is the number of data points.

1 INTRODUCTION

In clustering, one of the central tasks from data analysis and machine learning, we are given a set of n data points, typically in a metric space, and we want to group these points into clusters such that similar points are in the same cluster and dissimilar points are separated. Most studied in this regard is the partitioning into a fixed number k of pairwise disjoint clusters, with different objectives such as k-means. Another well-studied setting is *hierarchical clustering*, where the clusters form a hierarchy over the points. A hierarchical clustering is a sequence of clusterings, one for each possible number of clusters. Such a sequence starts with the clustering into n singleton clusters (i.e., each cluster contains a unique point), and iteratively creates the next clustering in the sequence by merging two existing clusters. The last clustering in the sequence then has one cluster containing all points. Hierarchical clusterings therefore provide information about the data at all levels of granularity. They have long been used in data mining and bioinformatics, and multi-level clusterings also appear in recent machine learning systems, such as hierarchical federated learning (Li et al., 2023; Yang et al., 2026; Shahid et al., 2025), trees of clusters over text embeddings for retrieval and memory in language-model pipelines (Sarthi et al., 2024; Rezazadeh et al., 2025), and multi-level hierarchies of clusters that summarize large collections of user interactions (Tamkin et al., 2024).

In settings in which the data points are related to individuals and the specific clustering may affect certain individuals, fairness considerations come into play. A classic example is facility location, where a city must decide where to place schools or hospitals. Each data point represents the location of a resident, and each resident wants to be as close as possible to one of the new facilities. For simplicity, we make the connection to individuals directly and refer to each data point as an agent. In particular, we follow the line of work on proportional fairness (Chen et al., 2019): a clustering is considered proportionally fair if no sufficiently large group of agents could "improve" by forming their own cluster. For classical clustering into k clusters, the canonical threshold for proportionality is $1/k$: the average cluster contains a $1/k$ fraction of the agents. Therefore, any group of agents of this size should be *entitled* to form its own cluster. For hierarchical clustering, however, the picture is less clear, since there are multiple clusterings into differently many clusters, and hence

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

multiple thresholds. In this work, we therefore study *how to measure the proportional fairness of a hierarchical clustering, and whether one can compute proportionally fair hierarchical clusterings.*

Work on proportional k-clustering can be distinguished by how agents evaluate the quality of a cluster. In centroid clustering, each cluster is represented by a center, and agents measure their loss by their distance to this center. In non-centroid clustering, agents instead evaluate a cluster by their distances to the other agents in it. Conceptually, hierarchical clustering is closer to the non-centroid setting: its goal is to describe how agents should be grouped at different levels of granularity, rather than to select representatives at each level. We therefore view a hierarchical clustering as a sequence of non-centroid clusterings. From this perspective, our starting point is the proportional fairness notion of Caragiannis et al. (2024) for non-centroid clustering. Their notion is based on the *core* from cooperative game theory: given a partition of n agents into k clusters no group of at least n/k agents should be able to improve (substantially) by forming their own cluster.

Our contributions. In this paper, we take two different perspectives of how to lift the notion of the core to hierarchical clustering, a *global* and a *local* perspective.

First, we study the *local core*. Intuitively, a hierarchical clustering can be modeled as a sequence of 2-clusterings. We start with the entire set of agents, then we partition this set into two subsets (i.e., we form a 2-clustering), continuing until each cluster is a singleton. The local core evaluates the proportionality of a hierarchy by the proportionality of these *local splits*: For each split in the hierarchical clustering the resulting 2-clustering should satisfy the core of Caragiannis et al. (2024).

Secondly, we study the *global core*. A hierarchical clustering can be seen as a sequence of non-centroid clusterings. The global core requires that every clustering in this sequence is in the core.

We start off by showing that standard agglomerative clustering methods satisfy neither the local nor the global core, nor any meaningful relaxation thereof (Section 3), including average linkage, which has provable guarantees for objective-based hierarchical clustering (Moseley & Wang, 2017; Cohen-Addad et al., 2019).

Focusing first on the local core (Section 4), we prove that, under the maximum loss objective, there always exists a hierarchical clustering in the local core. This is based on a simple algorithmic reduction to non-centroid 2-clustering: any α-approximation for the latter yields an α-approximation for the local core. On the way, we resolve an open question by Caragiannis et al. (2024) by showing that there always is a non-centroid 2-clustering in the core. Moreover, we show that for $k \geq 3$, there may be no k-clustering in the core, even if the points lie in a two-dimensional Euclidean space.

For the global core (Section 5), by contrast, no approximation is possible in general, and the global and local core do not imply each other in either direction. We therefore consider a relaxed notion in which the group-size threshold required for a coalition to block is doubled. For this relaxation, our algorithm ScoopNOrder returns a hierarchy in the global $(16, 2)$-core for the maximum loss and global $(8n, 2)$-core for the average loss.

Due to space constraints, proofs of results marked by ★ are deferred to the appendix.

Related work. For an overview of the vast topic of hierarchical clustering, we refer to two surveys (Murtagh & Contreras, 2012; Ran et al., 2022) and to some initial works on some of the applications (Sneath & Sokal, 1962; Jardine & Sibson, 1971; Castro et al., 2004). Recent work on fair hierarchical clustering (Ahmadian et al., 2020; Knittel et al., 2023a;b), like most work on fair clustering in machine learning (Chierichetti et al., 2017; Chhabra et al., 2023), focuses primarily on the notion of disparate impact, where one identifies some protected attributes (i.e., additional data next to the metric space) that should be treated fairly. Proportional fairness conceptually differs from disparate impact fairness in that it derives fairness only from the metric space itself and does not require predefined protected groups, a property it shares with individual fairness for clustering (Jung et al., 2020; Mahabadi & Vakilian, 2020; Chakrabarty & Negahbani, 2021).

Most central to our work is that of Caragiannis et al. (2024) on proportional non-centroid clustering. Their work takes inspiration from proportional centroid clustering, introduced by Chen et al. (2019), with follow-up work improving upper and lower bounds on approximation (Micha & Shah, 2020; Cookson et al., 2026a), studying core fairness in restricted metrics (Li et al., 2021), connecting it to other recently introduced concepts similar to proportionality (Aziz et al., 2024; Kellerhals & Peters,

2024), and providing algorithms to efficiently verify related axioms (He et al., 2026). Combining these two settings, Cookson et al. (2025) studied a joint model of non-centroid and centroid clustering under proportional fairness. The core has recently also been used as a fairness criterion in federated learning (Chaudhury et al., 2022). Recently, Kalayci & Micha (2026) considered proportionality in temporal panel selection in metric spaces, which similarly requires representation guarantees across a sequence of interdependent selections. Relatedly, Halpern et al. (2025) study federated assemblies, where representative assemblies are organized hierarchically and each parent assembly is selected from its children, giving rise to representation constraints across multiple levels of the hierarchy. Finally, Maity & Das (2026) consider fairness in hierarchical clustering from an individual fairness perspective, focusing on bounded distortion of local neighborhoods rather than proportional representation of groups.

Stating a simple objective for hierarchical clustering is not as trivial as it is for standard k-clustering. Our proportionality measures can also be seen as a way to specify the quality of a given hierarchical clustering. Dasgupta (2016) introduced an objective function for (similarity-based) hierarchical clustering. This objective was later generalized and extended to distance-based hierarchical clustering by Cohen-Addad et al. (2019); average linkage has constant-factor guarantees for the dual of this objective (Moseley & Wang, 2017) and for its distance-based variant (Cohen-Addad et al., 2019), and the objective is optimized directly by learned hierarchies (Chami et al., 2020; Zügner et al., 2022) and has been studied under differential privacy (Deng et al., 2025). Hierarchical clustering was also studied from an axiomatic approach: Carlsson & Mémoli (2010) specify three natural axioms inspired by the axioms by Kleinberg (2002) for k-clustering and show that the single linkage algorithm (defined in Section 3) is the unique method for distance-based hierarchical clustering satisfying these axioms.

Another approach for measuring the quality of a hierarchical clustering involves evaluating each clustering in the sequence using traditional objective functions for standard k-clustering (Plaxton, 2006; Dasgupta & Long, 2005; Lin et al., 2010). We similarly lift the core notion for non-centroid clustering to hierarchical clustering, in two different ways.

2 Model and Notation

Let $\mathcal{N}$ be a set of $n \geq 2$ *agents* in a (pseudo)-metric space $(\mathcal{N}, d)$ with a distance function $d\colon \mathcal{N} \times \mathcal{N} \to \mathbb{R}_{\geq 0}$ satisfying $d(x,x) = 0$, $d(x,y) = d(y,x)$, and $d(x,z) \leq d(x,y) + d(y,z)$ for all $x, y, z \in \mathcal{N}$. For a set of agents $A \subseteq \mathcal{N}$, we define $\operatorname{diam}(A) = \max_{x,y \in A} d(x,y)$, to be its *diameter* and we write $d|_A$ for the restriction of d to $A \times A$, i.e., $d|_A(x,y) \coloneqq d(x,y)$ for all $x, y \in A$. Thus $(A, d|_A)$ denotes the pseudo-metric space induced by $(\mathcal{N}, d)$ on A.

For an integer $1 \leq k \leq n$, a *k-clustering*, or simply a *clustering* is a partition $\mathcal{C} = \{C_1, \ldots, C_k\}$, i.e., $C_i \neq \emptyset$, $C_i \cap C_j = \emptyset$ for all $i \neq j$, and $\bigcup_{i=1}^{k} C_i = \mathcal{N}$. For $x \in \mathcal{N}$, we write $\mathcal{C}(x)$ for the unique cluster $C \in \mathcal{C}$ with $x \in C$.

A *hierarchical clustering* is a sequence $\mathcal{C}^1, \ldots, \mathcal{C}^n$, where $\mathcal{C}^k$ is a k-clustering, and for each $k \in \{1, \ldots, n-1\}$ there are two clusters $C, C' \in \mathcal{C}^{k+1}$ such that $\mathcal{C}^k = (\mathcal{C}^{k+1} \setminus \{C, C'\}) \cup \{C \cup C'\}$. We say that the clusters C, C' are *merged*, and denote the new cluster by $C_k \coloneqq C \cup C'$.

A hierarchical clustering $\mathcal{C}^1, \ldots, \mathcal{C}^n$ defines a rooted binary tree, denoted by $\mathcal{T}$, where each node corresponds to a subset of $\mathcal{N}$: There is a leaf for each singleton cluster $\{x\}$, and for each $k \in \{1, \ldots, n-1\}$, there is a node corresponding to C_k, whose two children are the nodes corresponding to C and C'. The root node corresponds to $C_1 = \mathcal{N}$. See Figure 1 for an example.

We measure clustering quality for each agent using loss functions. A loss function assigns to every nonempty cluster $C \subseteq \mathcal{N}$ and every agent $x \in C$ a value $\operatorname{loss}_x(C) \in \mathbb{R}_{\geq 0}$. The *average loss* of x in C is $\operatorname{loss}_x^{\text{avg}}(C) \coloneqq \frac{1}{|C|} \sum_{y \in C} d(x,y)$, and the *maximum loss* of x in C is $\operatorname{loss}_x^{\max}(C) \coloneqq \max_{y \in C} d(x,y)$. We focus on loss functions $\operatorname{loss}_{\cdot}(\cdot)$ that are bounded between average and maximum loss, i.e., for $C \subseteq \mathcal{N}$ and $x \in C$, $\operatorname{loss}_x^{\text{avg}}(C) \leq \operatorname{loss}_x(C) \leq \operatorname{loss}_x^{\max}(C)$. We call such loss functions *distance-bounded*. For a clustering $\mathcal{C}$ and an agent x, we write $\operatorname{loss}_x(\mathcal{C}) \coloneqq \operatorname{loss}_x(\mathcal{C}(x))$.

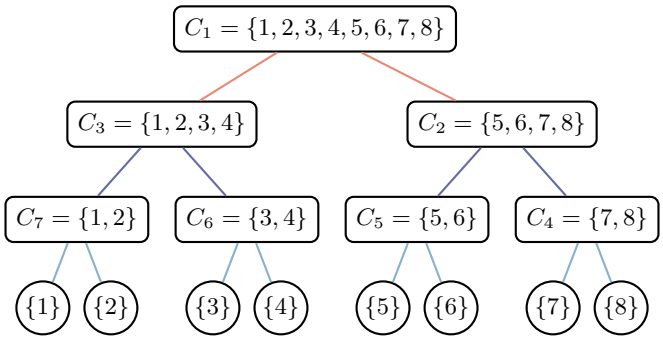

Figure 1: A hierarchical clustering represented as a tree. For instance, $\mathcal{C}^3 = \{\{1,2,3,4\},\{5,6\},\{7,8\}\}$.

Proportional fairness. Given $\alpha \geq 1$ and a loss function $\text{loss}_\cdot(\cdot)$ we say that a coalition $S \subseteq \mathcal{N}$ of agents α-*blocks* a clustering $\mathcal{C}$ if $\alpha \cdot \text{loss}_x(S) < \text{loss}_x(\mathcal{C})$ for all $x \in S$.

We recap two previous definitions of proportionality of Caragiannis et al. (2024). First, a k-clustering $\mathcal{C}$ is in the α-*core*, if there is no coalition of agents $S \subseteq \mathcal{N}$ with $|S| \geq n/k$ that α-blocks $\mathcal{C}$. They refer to the 1-core simply as the core. Similarly, a k-clustering $\mathcal{C}$ is in the (α, δ)-*core* if there is no coalition of agents $S \subseteq \mathcal{N}$ with $|S| \geq \delta \cdot n/k$ that α-blocks $\mathcal{C}$.

We extend these notions to the hierarchical setting. For our first notion, we evaluate the proportionality of the clusterings in the sequence constituing our hierarchical clustering. That is, we require that after $n - k$ merges, the resulting k-clustering is in the (approximate) core.

Definition 1 (global (α, δ)-core)**.** *Let $\alpha, \delta \geq 1$. We say that a hierarchical clustering $\mathcal{C}^1, \ldots, \mathcal{C}^n$ is in the* global (α, δ)-core *if for each $k \in \{1, \ldots, n\}$ the k-clustering $\mathcal{C}^k$ is in the (α, δ)-core.*

If $\delta = 1$, we say that $\mathcal{C}^1, \ldots, \mathcal{C}^n$ is in the *global α-core*. If additionally $\alpha = 1$, we say that the clustering is in the *global core*.

Our second notion imposes the proportionality condition locally at each internal node of the hierarchical clustering tree. That is, for each internal node C which is split into child clusters C_1, C_2 this split should be chosen proportionally (with regard to C).

Definition 2 (local α-core)**.** *Let $\alpha \geq 1$. A hierarchical clustering $\mathcal{C}^1, \ldots, \mathcal{C}^n$ with corresponding tree $\mathcal{T}$ is in the* local α-core *if for every internal (non-leaf) node C with children C_1 and C_2, the 2-clustering $\{C_1, C_2\}$ of C is in the α-core in the induced instance $(C, d\,|_C)$.*

The two hierarchical notions differ in the clusterings against which blocking coalitions are evaluated. For the global (α, δ)-core, coalitions may be arbitrary subsets of $\mathcal{N}$ and are compared to the global clustering $\mathcal{C}^k$ for all $k \in \{1, \ldots, n\}$. The local α-core instead applies the proportionality condition separately at each internal node. At a node C, only coalitions within C are considered, and they are compared only to the partition $\{C_1, C_2\}$. Thus, a coalition can block the partition at C only if it contains at least $|C|/2$ of the agents in C.

For example, consider the tree in Figure 1. The root cluster $\mathcal{N}$ is split into C_2 and C_3, and C_2 is further split into C_4 and C_5. At the merge where the clustering is $\{C_3, C_5, C_4\}$, a blocking coalition for the (α, δ)-core may contain agents from anywhere in $\mathcal{N}$ and must have size at least $\delta n/3$. In contrast, the local condition at node C_2 only considers coalitions $S \subseteq C_2$ of size at least $|C_2|/2$, and compares them to the two-cluster partition $\{C_5, C_4\}$.

As we will see, there are hierarchical clusterings that are in the local core but not in the global α-core, and hierarchical clusterings that are in the global core but not in the local α-core.

3 Limitations of Standard Agglomerative Clustering Methods

We briefly recall standard agglomerative hierarchical clustering methods. Each algorithm starts from the singleton partition $\mathcal{C}^n = \{\{i\} : i \in \mathcal{N}\}$ and iteratively merges clusters until only a single cluster remains, thereby inducing a hierarchical clustering tree $\mathcal{T}$.

Most standard agglomerative clustering methods proceed by repeatedly merging the pair of clusters with minimum inter-cluster distance. They differ only in the linkage criterion used to define this distance. We model such a criterion as a function $f\colon 2^{\mathcal{N}} \times 2^{\mathcal{N}} \to \mathbb{R}$, where $f(A, B)$ denotes the distance assigned to two nonempty clusters $A, B \subseteq \mathcal{N}$. We consider linkage functions standard if they satisfy $\min_{a\in A,\, b\in B} d(a, b) \leq f(A, B) \leq \max_{a\in A,\, b\in B} d(a, b)$, for all nonempty disjoint clusters $A, B \subseteq \mathcal{N}$. This condition captures a broad class of established linkage rules whose inter-cluster distance is determined by the pairwise distances between points in the two clusters. The extremes are known as the single-linkage and complete-linkage clustering methods. Further, average-linkage in its many variants typically lies between the two extremes.

For the broad class of hierarchical clustering methods using standard linkage functions we show that they cannot guarantee outputs in the global (α, δ)-core for any $\alpha \geq 1$ and a $\delta \in \Omega(n)$, or in the local α-core for any $\alpha \geq 1$, even with just four agents.

Theorem 1 (★). *For every $\alpha \geq 1$, every $n \geq 4$, every agglomerative hierarchical clustering algorithm with standard linkage function f and any distance-bounded loss, there is a metric space $(\mathcal{N}, d)$ with $|\mathcal{N}| = n$ for which any clustering $\mathcal{C}^1, \ldots, \mathcal{C}^n$ returned by the algorithm is not in the global (α, δ)-core for any $\delta \leq \lfloor n^2/4 \rfloor / n$, and not in the local α-core.*

4 Hierarchical Clusterings in the Local Core

Due to the impossibility of standard agglomerative hierarchical clustering algorithms achieving proportional fairness, we now turn to the design of new algorithms for our proportionality notions. We first consider the local core. This perspective is natural when hierarchical clustering represents nested decisions. Consider, for instance, the hierarchical federated learning setting in which in multiple steps clients are subdivided into increasingly homogeneous subgroups. We argue that the subdivision should be proportional at each split. If at least half of the clients would all incur lower loss by forming a subgroup together, then the split fails to represent them proportionally.

Since the local condition is imposed at each internal node separately, a natural approach is to build the hierarchical clustering top-down by repeatedly applying an algorithm for non-centroid clustering. Let $\mathcal{A}$ be an algorithm for non-centroid clustering that, given a metric space and an integer k, returns a k-clustering. Our algorithm, RecursiveSplit($\mathcal{A}$) (see Algorithm 2 in the appendix for a formal description), repreatedly takes an existing cluster (initially $\mathcal{N}$) and determines its children by computing a 2-clustering with $\mathcal{A}$. We say that algorithm $\mathcal{A}$ is an *α-core algorithm* for non-centroid k-clustering if, given any metric space $(\mathcal{N}, d)$, it returns a k-clustering in the α-core. We show that this guarantee transfers to the local core guarantee of RecursiveSplit($\mathcal{A}$).

Theorem 2 (★). *Let $\alpha \geq 1$. If $\mathcal{A}$ is an α-core algorithm for non-centroid 2-clustering, then* RecursiveSplit($\mathcal{A}$) *always returns a hierarchical clustering in the local α-core.*

In general, for non-centroid k-clustering under the maximum-loss objective, Caragiannis et al. (2024) showed that one can always obtain a k-clustering in the 2-core in polynomial time. They, however, left open whether one can obtain a clustering in the core. We confirm this for $k = 2$.

Theorem 3 (★). *There always exists a 2-clustering in the core under the maximum-loss objective.*

We prove Theorem 3 in two steps. First, we show that there is always a *fractional* clustering in the core, employing the theorem of Scarf (1967) on the non-emptyness of the core of balanced NTU games.[1] Then we show how to construct a "proper" 2-clustering from this fractional clustering.

To define *fractional* k-clusterings, let $\mathfrak{C} = \{C \subseteq \mathcal{N}\colon |C| \geq n/k\}$ be our set of possible blocking coalitions. For a single agent $x \in \mathcal{N}$, let $\mathfrak{C}_x = \{C \in \mathfrak{C}\colon x \in C\}$ be the set of coalitions containing x. Then a fractional clustering is an assignment of weights $\lambda\colon \mathfrak{C} \to [0, 1]$ such that $\sum_{C\in\mathfrak{C}_x} \lambda(C) \leq 1$ for every agent $x \in \mathcal{N}$; if equality holds for x, then we call x *covered*. The *loss* of x in λ is

$$\mathrm{loss}_x(\lambda) = \begin{cases} \max\limits_{\substack{C\in\mathfrak{C}_x \\ \lambda(C)>0}} \mathrm{loss}_x^{\max}(C) & \text{if } x \text{ is covered,} \\ \mathrm{diam}(\mathcal{N}) & \text{if } x \text{ is uncovered.} \end{cases}$$

[1] We note that a (different) approach based on Scarf's theorem was also recently used by Cookson et al. (2026b) to prove approximate core non-emptiness in centroid clustering.

A coalition $S \in \mathfrak{C}$ blocks a fractional clustering λ if $\text{loss}_x^{\max}(S) < \text{loss}_x(\lambda)$ for each $x \in S$. A fractional clustering λ is in the core if there is no coalition that blocks it. Proving that such a fractional clustering always exists comes down to proving that the associated NTU game is balanced.

Lemma 1 (★). *For every $k \geq 2$, there exists a fractional k-clustering in the core.*

Next, for $k = 2$, we show that any fractional 2-clustering can be translated to a "proper" 2-clustering in which each agent has at most the same loss that they received from the fractional clustering.

Lemma 2. *Given a fractional 2-clustering λ, there is a 2-clustering $\mathcal{C} = \{C_1, C_2\}$ with $\text{loss}_x^{\max}(\mathcal{C}) \leq \text{loss}_x(\lambda)$ for each $x \in \mathcal{N}$.*

Proof. For the fractional clustering λ, we define a conflict graph $G = (\mathcal{N}, E_\lambda)$ with $E_\lambda = \{\{x, y\} : x, y \in \mathcal{N}, d(x, y) > \min(\text{loss}_x(\lambda), \text{loss}_y(\lambda))\}$. An edge between two agents thus indicates that they cannot be in the same cluster in $\mathcal{C}$. It remains to show that G is bipartite as then we can take $\mathcal{C}$ to be the two sides of the bipartition.

We notice that $\sum_{C \in \mathfrak{C}} \lambda(C) = \sum_{x \in \mathcal{N}} \sum_{C \in \mathfrak{C}_x} \frac{1}{|C|} \lambda(C) \leq \sum_{x \in \mathcal{N}} \sum_{C \in \mathfrak{C}_x} \frac{2}{|\mathcal{N}|} \lambda(C) \leq 2$ with an equality in the last step occuring if and only if every agent $x \in \mathcal{N}$ has $\sum_{C \in \mathfrak{C}_x} \lambda(C) = 1$. If $E_\lambda = \emptyset$, any partition of $\mathcal{N}$ into two nonempty sets is a bipartition of G. Now suppose $E_\lambda \neq \emptyset$.

First, assume that every agent is covered. For any edge $\{x, y\} \in E_\lambda$, assume without loss of generality that $\text{loss}_x(\lambda) \leq \text{loss}_y(\lambda)$. No coalition C with $\lambda(C) > 0$ can contain both endpoints, as otherwise, $d(x, y) \leq \text{loss}_x^{\max}(C) \leq \text{loss}_x(\lambda)$, contradicting the definition of E_λ. Since both endpoints are covered, $\sum_{C \in \mathfrak{C}} \lambda(C) \geq \sum_{C \in \mathfrak{C}_x} \lambda(C) + \sum_{C \in \mathfrak{C}_y} \lambda(C) = 2$. Combining with the upper bound, this gives $\sum_{C \in \mathfrak{C}} \lambda(C) = 2$, implying that every coalition with positive weight contains either x or y. Hence, for any coalition C with $\lambda(C) > 0$ and any edge $\{x, y\}$, we have $x \in C$ if and only if $y \notin C$; thus $(C, \mathcal{N} \setminus C)$ is a bipartition.

Now assume that there are uncovered agents and let $U \subseteq \mathcal{N}$ be the set of uncovered agents. Then $\sum_{C \in \mathfrak{C}} \lambda(C) < 2$. Since $\text{loss}_x(\lambda) = \text{diam}(\mathcal{N})$ for each $x \in U$, there cannot exist any edges between agents in U. Suppose next that there are two covered agents $y, z \in \mathcal{N} \setminus U$ with $\{y, z\} \in E_\lambda$. Then, as above, any C' with $y, z \in C'$ has $\lambda(C') = 0$. But this implies that the total weight of all coalitions would be exactly 2; a contradiction to the existence of uncovered agents. Thus, $(U, \mathcal{N} \setminus U)$ forms a bipartition. □

Proof of Theorem 3. Let λ be a fractional 2-clustering in the core (Lemma 1), i.e., for any coalition S of at least $n/2$ agents, there is an agent $x \in S$ with $\text{loss}_x(\lambda) \leq \text{loss}_x^{\max}(S)$. Thus, in the 2-clustering $\mathcal{C}$ obtained by Lemma 2, $\text{loss}_x^{\max}(\mathcal{C}) \leq \text{loss}_x(\lambda)$. So S does not block $\mathcal{C}$; thus $\mathcal{C}$ is in the core. □

As an immediate consequence of Theorems 2 and 3, we obtain the following.

Corollary 4. *For maximum loss, there always exists a hierarchical clustering in the local core.*

Remark 1. *The proof of Theorem 3 is existential, and it remains open whether a core-stable 2-clustering can be efficiently computed. The reduction by Caragiannis et al. (2024, Theorem 6) implies that checking core membership is coNP-hard for general k, although this does not imply hardness of finding a 2-clustering in the core. However, we can use* GREEDY CAPTURE *from Caragiannis et al. (2024) as $\mathcal{A}$ to obtain a hierarchical clustering in the local 2-core in polynomial time.*

We conclude the section with some further related results. First, while there always exists a fractional k-clustering in the core, there need not exist a "proper" k-clustering for any $k \geq 3$. This even holds true for Euclidean metric spaces.

Theorem 5 (★). *For every $k \geq 3$ there exists a metric space with $n \geq 9$ agents, where n is divisible by k, for which no k-clustering is in the α-core under the max-loss objective, for any $\alpha < 2^{1/5} < 1.149$. Moreover, for $k = 3$ there is a two-dimensional Euclidean space with $n = 9$ agents for which no k-clustering is in the α-core under the max-loss objective, for any $\alpha < 1.138$.*

Second, for the average-loss objective, we can slightly improve the $(2\lceil n/k \rceil - 3)$-core guarantee of Caragiannis et al. (2024) for non-centroid clustering by modifying their GREEDYCAPTURE algorithm. Rather than selecting a minimum-radius ball, we select a ball minimizing the average distance from its capturing agent. We refer to this variant as AVERAGEGREEDYCAPTURE.

Algorithm 1: SCOOPNORDER

Input : Normalized metric space $(\mathcal{N}, d)$, such that the smallest non-zero distance is 1
Output: Hierarchical clustering $\mathcal{C}^1, \ldots, \mathcal{C}^n$

Initialize a hierarchical clustering tree $\mathcal{T}$ by adding a singleton $\{x\}$ for each $x \in \mathcal{N}$.
Call a cluster C in $\mathcal{T}$ *active* if it does not yet have a parent (initially, all are active).
$\Delta \leftarrow \lceil \log_4(2\,\mathrm{diam}(\mathcal{N})) \rceil$
for $r = 0, 4^0, 4^1, \ldots, 4^\Delta$ **do**
  for $\ell = \lceil \log_2 n \rceil, \ldots, 2, 1$ **do**
    for $y \in \mathcal{N}$ **do**
      $\mathcal{Q} \leftarrow \{C \mid C \text{ is an active cluster and } C \subseteq B(y, r)\}$
      while *there is a set* $\mathcal{Q}' \subseteq \mathcal{Q}$ *with* $|\mathcal{Q}'| \geq 2$ *and* $|\bigcup \mathcal{Q}'| = 2^\ell$ **do**
        while $|\mathcal{Q}'| > 1$ **do**
          Take two minimum-size clusters $C, C' \in \mathcal{Q}'$ and merge them in $\mathcal{T}$.
          $\mathcal{Q}' \leftarrow (\mathcal{Q}' \setminus \{C, C'\}) \cup \{C \cup C'\}$
        $\mathcal{Q} \leftarrow \{C \mid C \text{ is an active cluster and } C \subseteq B(y, r)\}$
while *there is more than one active cluster in* $\mathcal{T}$ **do**
  Take two minimum-size active clusters C, C' and merge them in $\mathcal{T}$.
$\mathcal{C}^n \leftarrow \{\{x\} \mid x \in \mathcal{N}\}$
for $k = n - 1, \ldots, 1$ **do**
  Pick clusters $C, C' \in \mathcal{C}^{k+1}$ minimizing $|C \cup C'|$, such that $C \cup C'$ is a node in $\mathcal{T}$.
  $\mathcal{C}^k \leftarrow (\mathcal{C}^{k+1} \setminus \{C, C'\}) \cup \{C \cup C'\}$

Theorem 6 (★). *For the average-loss objective and $k < n$,* AVERAGEGREEDYCAPTURE *returns a clustering in the $\big(\lceil n/k \rceil - 1\big)$-core in polynomial time. Consequently,* RECURSIVESPLIT(AVERAGEGREEDYCAPTURE) *computes a hierarchical clustering in the local $(\lceil n/2 \rceil - 1)$-core in polynomial time.*

5 HIERARCHICAL CLUSTERINGS IN THE GLOBAL CORE

The local core guarantees obtained above do not imply a global core guarantee. For the local core, agents that are separated by an early split are never compared again. For the global core, however, these agents might still form a blocking coalition.

Firstly, we remark that, under the maximum-loss objective, for every $\alpha \geq 1$ there is a hierarchical clustering that is in the global core but not in the local α-core.

Theorem 7 (★). *For every $\alpha \geq 1$, there is a metric space with eight agents such that for the maximum loss objective, there is a hierarchical clustering that is in the global core but not in the local α-core.*

Second, we observe that the global α-core may be empty for every α. Since there always exists a clustering in the local core (Corollary 4), this implies a converse to Theorem 7. Therefore, the local and global core are incomparable.

Theorem 8 (★). *For every finite $\alpha \geq 1$ and for any distance-bounded loss, there exists an instance with six agents for which no hierarchical clustering is in the global α-core.*

To circumvent this impossibility, we consider a relaxation, the global (α, δ)-core, which increases the size bound at which coalitions may block a hierarchical clustering. For $\delta = 2$, we are able to give approximate guarantees of the global core.

Theorem 9. SCOOPNORDER *(Algorithm 1) produces a hierarchical clustering that is*

(1) in the $(16, 2)$-core for the maximum loss objective and

(2) in the $(8n, 2)$-core for the average loss objective.

SCOOPNORDER[2] proceeds in two phases. First, a hierarchical clustering tree $\mathcal{T}$ is built bottom-up: For increasing radii, we check for each agent whether a ball with the current radius around them (a "scoop") contains a set of clusters whose overall number of agents is a power of two. These clusters are then merged, again making sure to keep the power-of-two property. Once no further scoops can be found, we take the last remaining clusters and merge them in order of their size. The second phase then takes $\mathcal{T}$ and builds a hierarchical clustering $\mathcal{C}^1, \dots, \mathcal{C}^n$ out of them. For this, we need to fix an order of the merges; we simply order them by the size of the clusters.

We first prove that for any possible deviation set S, there is a large cluster whose diameter is bounded by the radius of S.

For $k \in \{1, \dots, n\}$, we write $\ell(k) \coloneqq \lceil \log_2(n/k) \rceil$.

Lemma 3. *Fix $k \in \{1, \dots, n\}$ and let $S \subseteq \mathcal{N}$ with $|S| \geq 2^n/k$. The tree $\mathcal{T}$ constructed by* SCOOPNORDER *contains a cluster C intersecting S with size at least $2^{\ell(k)}$ and*

(1) $\operatorname{diam}(C) \leq 16 \min_{x \in S} \operatorname{loss}_x^{\max}(S)$ *and*

(2) $\operatorname{diam}(C) \leq 8n \min_{x \in S} \operatorname{loss}_x^{\operatorname{avg}}(S)$.

Proof. Let $x \in S$ be the agent minimizing $\operatorname{loss}_x^{\max}(S)$, and let r be the first processed radius satisfying $r \geq 2 \operatorname{loss}_x^{\max}(S)$. Thus $r/4 < 2 \operatorname{loss}_x^{\max}(S)$ and we have $r < 8 \operatorname{loss}_x^{\max}(S)$.

If $\operatorname{loss}_x^{\max}(S) = 0$, then $r = 0$. Otherwise, $\operatorname{loss}_x^{\max}(S) \geq 1$, so the algorithm iterates over a radius r' with $r' < 8/4 \operatorname{loss}_x^{\max}(S)$. Thus, before processing r, every current cluster has diameter at most $2r' \leq r/2$. By the triangle inequality, each of the current clusters C intersecting S lies in $B(x, r)$, since $d(x, y) \leq \operatorname{loss}_x^{\max}(S) + \operatorname{diam}(C) \leq r$ for every $y \in C$. If one of these clusters has size at least $2^{\ell(k)}$, then we are done. Otherwise, their sizes are powers of two smaller than $2^{\ell(k)}$, and the sum of their sizes is at least $|S| \geq 2^{\ell(k)}$.

Among the clusters intersecting S, there is a subset whose cumulative size is exactly $2^{\ell(k)}$ since at this stage of the algorithm, all cluster sizes are powers of two. Call this subset $\mathcal{Q}'$. If in the for loop for radius r, there is an $\ell \geq \ell(k)$ and a $y \in \mathcal{N}$ such that $\bigcup \mathcal{Q}' \subseteq B(y, r)$ and $|\bigcup \mathcal{Q}'| = 2^\ell$, then either $\bigcup \mathcal{Q}'$ becomes a cluster of the required size, intersects S, and has diameter at most $2r$ or there is a different cluster created containing a member of S and thus also fulfilling the diameter bounds. In either case, $\mathcal{T}$ contains a cluster intersecting S of size at least $2^{\ell(k)}$ and of diameter at most $2r \leq 16 \operatorname{loss}_x^{\max}(S)$.

For average loss, choose $x \in S$ minimizing $\operatorname{loss}_x^{\operatorname{avg}}(S)$ and select the $2^{\ell(k)}$ closest members of S to x, including x. Since $|S| > 2^{\ell(k)}$, at least two members of S are at least as far from x as the farthest selected agent. The selected agents therefore lie in a ball around x of radius at most $\frac{1}{2} \sum_{y \in S} d(x, y) = \frac{|S|}{2} \operatorname{loss}_x^{\operatorname{avg}}(S)$.

The argument above applies to any set of at least $2^{\ell(k)}$ agents and any agent belonging to that set. Applied to the selected agents, it gives a cluster intersecting S, of size at least $2^{\ell(k)}$, and of diameter at most $16 \cdot \frac{|S|}{2} \operatorname{loss}_x^{\operatorname{avg}}(S) \leq 8n \min_{y \in S} \operatorname{loss}_y^{\operatorname{avg}}(S)$. □

The ordering step ensures that such a cluster C will be built in $\mathcal{C}^k$ or some $\mathcal{C}^t$ with $t < k$.

Lemma 4. *Let $\mathcal{C}^1, \dots, \mathcal{C}^n$ be a hierarchical clustering with clustering tree $\mathcal{T}$ returned by* SCOOPNORDER *and let C be a cluster of size at least $2^{\ell(k)}$ appearing in $\mathcal{T}$ for some $k \in \{1, \dots, n\}$. Then, for each $x \in C$, we have $\mathcal{C}^k(x) \subseteq C$.*

Proof. As illustrated in Figure 2, we distinguish between the clusters in $\mathcal{T}$ that were created in the main for loop (regular clusters) and those that were created in the subsequent while loop (odd clusters).

The size of every regular cluster is a power of two and thus consists of two clusters of equal size (also a power of two). Moreover, due to the last iteration of the for loop ($r = 4^\Delta$), we have that

[2]For easier readability, we state SCOOPNORDER for a metric space in which the smallest non-zero distance is 1. This can be obtained by dividing all distances by δ, where δ is the smallest non-zero distance.

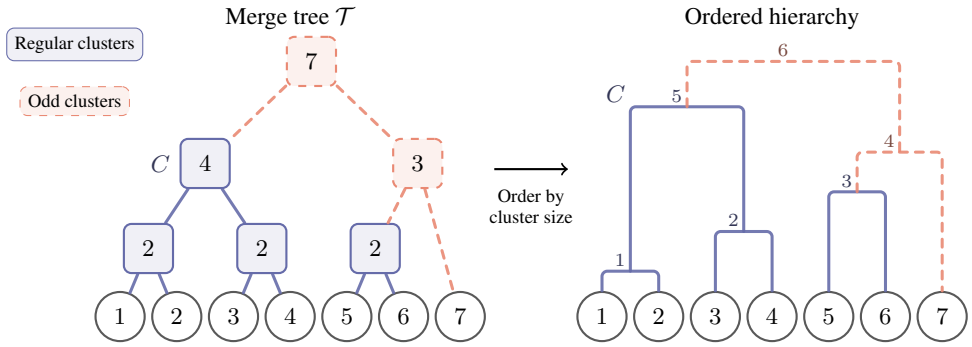

Figure 2: Ordering the merge tree $\mathcal{T}$ in SCOOPNORDER. Numbers at internal nodes of the tree give cluster sizes; numbers in the dendrogram give the merge order.

for each $i \in \{1, \dots, \lceil \log_2 n \rceil\}$, the (incomplete) hierarchical clustering tree $\mathcal{T}$ contains at most one cluster of size 2^i. This immediately implies that first, any odd cluster is created by merging two clusters of different size, and its size is not a power of two, and second, since the odd clusters are merged in the order of their size, each clustering $\mathcal{C}^i$ contains at most one odd cluster.

Let t be the smallest index such that each cluster in $\mathcal{C}^t$ has size at most $2^{\ell(k)}$. That is, $\mathcal{C}^t$ is obtained after completing all merges of size at most $2^{\ell(k)}$. Since $\mathcal{C}^t$ contains at most one odd cluster, and any two regular clusters of size at most $2^{\ell(k)-1}$ have already been merged (otherwise we chose t too large), all clusters but one in $\mathcal{C}^t$ have size $2^{\ell(k)}$. Thus, the number t of clusters in $\mathcal{C}^t$ is at most $t \leq \lceil n/2^{\ell(k)} \rceil \leq k$. Moreover, since $|C| \geq 2^{\ell(k)}$, we have $\mathcal{C}^t(x) \subseteq C$ for each $x \in C$. This is also true for every $t' \geq t$, and therefore also for $k \geq t$, which proves the claim. □

Proof of Theorem 9. Fix $k \in \{1, \dots, n\}$ and $S \subseteq \mathcal{N}$ with $|S| \geq 2n/k$. By Lemma 3, $\mathcal{T}$ contains a cluster C intersecting S such that $|C| \geq 2^{\ell(k)}$ and $\operatorname{diam}(C) \leq 16 \min_{x \in S} \operatorname{loss}_x^{\max}(S)$. By Lemma 4, we have $\mathcal{C}^k(x) \subseteq C$ for each $x \in C$. Therefore, for each $x \in C \cap S$, we have

$$\operatorname{loss}_x^{\max}(\mathcal{C}^k(x)) \leq \operatorname{diam}(C) \leq 16 \min_{y \in S} \operatorname{loss}_y^{\max}(S) \leq 16 \operatorname{loss}_x^{\max}(S).$$

Hence, S does not 16-block $\mathcal{C}^k$, and $\mathcal{C}^k$ is in the $(16, 2)$-core. Thus, $\mathcal{C}^1, \dots, \mathcal{C}^n$ is in the global $(16, 2)$-core for maximum loss. The argumentation for average loss is analogous. □

6 CONCLUSION

We introduced the concept of proportionality to hierarchical clustering and defined two variants of core stability: the global and the local core.

Our work leaves several open questions and possible future directions. The (non-)emptiness of the local core directly relates to the (non-)emptiness of the core for non-centroid clustering with $k = 2$. While we show that the core is always nonempty for $k = 2$ under the maximum loss, our proof is existential, and it remains open whether such a clustering can be computed in polynomial time. For the global (α, δ)-core, it would be interesting to see whether one can obtain guarantees for $1 < \delta < 2$.

Further, the auditing problem (given a hierarchical clustering, find the smallest α such that it is in the global (α, δ)-core or local α-core) has not yet been addressed, and is highly relevant for empirical analyses of hierarchical clustering algorithms.

Finally, we remark that there are also other ways to define hierarchical clusterings, for example by specifying at which "distance" two clusters are merged instead of specifying the order of merges. While the local core notion can be simply defined to work with the distance-based variant, this is not true for the global core, which heavily relies on the merge order. A distance-based (and order-invariant) version of the global core thus would be intriguing.

AI USE STATEMENT

In this work, we used generative AI tools for the development of Theorem 3 and Algorithm 1. The proof of Theorem 3 was developed with assistance from OpenAI's model GPT-5.6-Sol. We prompted the model to explore the $k = 2$ case using Scarf's balanced-core theorem. Through several rounds of discussion, the model suggested the conflict graph formulation and helped develop and present the resulting proof. We checked and revised the argument, corrected intermediate issues, and take full responsibility for the final proof and exposition. For Algorithm 1, the model suggested an initial algorithm after being prompted to improve upon a previous algorithm. We used this idea as a starting point, but rebuilt the algorithm and wrote its final presentation ourselves.

Additionally, we used generative AI tools to improve the language of the manuscript and to generate illustrative TikZ figures.

We have not used generative AI tools to help develop theoretical models or conceptual frameworks. The remaining tasks required to be disclosed (generating synthetic data sets, proposing or refining hypotheses, designing or provide feedback on research methodology or experiments, implementing methods, assisting with translation, cleaning and reformating datasets, supporting qualitative and thematic data analysis, interpreting results) do not apply to this work.

We have thoroughly reviewed all AI-assisted work and take responsibility for the final content of this work.

ETHICS STATEMENT

This work is theoretical and does not involve human participants or personal datasets. Our proportional fairness guarantees are defined relative to a given distance metric and do not imply fairness with respect to protected attributes.

REPRODUCIBILITY STATEMENT

The proofs of our results are provided in the main text and appendix, together with pseudocode for our algorithms. For the Euclidean space for Theorem 5, we also provide the code used to obtain the point coordinates in the supplementary material.

REFERENCES

Sara Ahmadian, Alessandro Epasto, Marina Knittel, Ravi Kumar, Mohammad Mahdian, Benjamin Moseley, Philip Pham, Sergei Vassilvitskii, and Yuyan Wang. Fair hierarchical clustering. In *Proceedings of the 33rd Conference on Neural Information Processing Systems (NeurIPS)*, 2020. URL https://proceedings.neurips.cc/paper/2020/hash/f10f2da9a238b746d2bac55759915f0d-Abstract.html.

Haris Aziz, Barton E. Lee, Sean Morota Chu, and Jeremy Vollen. Propotionally representative clustering. In *Proceedings of the 20th International Conference on Web and Internet Economics (WINE)*, 2024. URL http://www.cse.unsw.edu.au/~haziz/faircluster_report.pdf. Forthcoming.

Ioannis Caragiannis, Evi Micha, and Nisarg Shah. Proportional fairness in non-centroid clustering. In *Proceedings of the 37th Conference on Neural Information Processing Systems (NeurIPS)*, volume 37, pp. 19139–19166, 2024. URL https://papers.nips.cc/paper_files/paper/2024/hash/220cbc7435d6a56205c87d73d15d9eda-Abstract-Conference.html.

Gunnar Carlsson and Facundo Mémoli. Characterization, stability and convergence of hierarchical clustering methods. *Journal of Machine Learning Research*, 11(31):1425–1470, 2010. doi: 10.5555/1756006.1859898. URL https://www.jmlr.org/papers/v11/carlsson10a.html.

Rui M. Castro, Mark Coates, and Robert D. Nowak. Likelihood based hierarchical clustering. *IEEE Transactions on Signal Processing*, 52(8):2308–2321, 2004. doi: 10.1109/TSP.2004.831124. URL https://doi.org/10.1109/TSP.2004.831124.

Deeparnab Chakrabarty and Maryam Negahbani. Better algorithms for individually fair k-clustering. In *Proceedings of the 35th Conference on Neural Information Processing Systems (NeurIPS)*, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/6f221fcb5c504fe96789df252123770b-Abstract.html.

Ines Chami, Albert Gu, Vaggos Chatziafratis, and Christopher Ré. From trees to continuous embeddings and back: Hyperbolic hierarchical clustering. In *Proceedings of the 34th Conference on Neural Information Processing Systems (NeurIPS)*, 2020. URL https://proceedings.neurips.cc/paper/2020/hash/ac10ec1ace51b2d973cd87973a98d3ab-Abstract.html.

Bhaskar Ray Chaudhury, Linyi Li, Mintong Kang, Bo Li, and Ruta Mehta. Fairness in federated learning via core-stability. In *Proceedings of the 36th Conference on Neural Information Processing Systems (NeurIPS)*, 2022. URL https://proceedings.neurips.cc/paper_files/paper/2022/hash/25e92e33ac8c35fd49f394c37f21b6da-Abstract-Conference.html.

Xingyu Chen, Brandon Fain, Liang Lyu, and Kamesh Munagala. Proportionally fair clustering. In *Proceedings of the 36th International Conference on Machine Learning (ICML)*, volume 97, pp. 1032–1041, 09–15 Jun 2019. URL https://proceedings.mlr.press/v97/chen19d.html.

Anshuman Chhabra, Peizhao Li, Prasant Mohapatra, and Hongfu Liu. Robust fair clustering: A novel fairness attack and defense framework. In *Proceedings of the 11th International Conference on Learning Representations (ICLR)*, 2023. URL https://openreview.net/forum?id=4LMIZY7gt7h.

Flavio Chierichetti, Ravi Kumar, Silvio Lattanzi, and Sergei Vassilvitskii. Fair clustering through fairlets. In *Proceedings of the 31st Conference on Neural Information Processing Systems (NeurIPS)*, pp. 5029–5037, 2017. URL https://proceedings.neurips.cc/paper/2017/hash/978fce5bcc4eccc88ad48ce3914124a2-Abstract.html.

Vincent Cohen-Addad, Varun Kanade, Frederik Mallmann-Trenn, and Claire Mathieu. Hierarchical clustering: Objective functions and algorithms. *Journal of the ACM*, 66(4):26:1–26:42, 2019. doi: 10.1145/3321386. URL https://doi.org/10.1145/3321386.

Benjamin Cookson, Nisarg Shah, and Ziqi Yu. Unifying proportional fairness in centroid and non-centroid clustering. In *Proceedings of the 38th Conference on Neural Information Processing Systems (NeurIPS)*, pp. 35349–35386, 2025. URL https://proceedings.neurips.cc/paper_files/paper/2025/file/32b640528f5b67975562210f00c131ed-Paper-Conference.pdf.

Benjamin Cookson, Eva Deltl, and Yeeseok Oh. Improved lower bounds for proportionally fair clustering, 2026a. URL https://arxiv.org/abs/2606.07285.

Benjamin Cookson, Eva Deltl, and Yeeseok Oh. Optimally selecting representative agents from a metric space, 2026b. URL https://arxiv.org/abs/2608.29097.

Sanjoy Dasgupta. A cost function for similarity-based hierarchical clustering. In Daniel Wichs and Yishay Mansour (eds.), *Proceedings of the 48th Annual ACM Symposium on Theory of Computing (STOC)*, pp. 118–127. ACM, 2016. doi: 10.1145/2897518.2897527. URL https://doi.org/10.1145/2897518.2897527.

Sanjoy Dasgupta and Philip M. Long. Performance guarantees for hierarchical clustering. *Journal of Computer and System Sciences*, 70(4):555–569, 2005. ISSN 0022-0000. doi: https://doi.org/10.1016/j.jcss.2004.10.006. URL https://www.sciencedirect.com/science/article/pii/S0022000004001321.

Chengyuan Deng, Jie Gao, Jalaj Upadhyay, Chen Wang, and Samson Zhou. On the price of differential privacy for hierarchical clustering. In *Proceedings of the 13th International Conference on Learning Representations (ICLR)*, 2025. URL https://openreview.net/forum?id=yLhJYvkKA0.

Daniel Halpern, Ariel D. Procaccia, Ehud Shapiro, and Nimrod Talmon. Federated assemblies. In *Proceedings of the Thirty-Ninth AAAI Conference on Artificial Intelligence (AAAI'25)*. AAAI Press, 2025. ISBN 978-1-57735-897-8. doi: 10.1609/aaai.v39i13.33520. URL https://doi.org/10.1609/aaai.v39i13.33520.

Yu He, Jeremy Vollen, and Edith Elkind. Check, please: Verifiably fair clustering, 2026. URL https://arxiv.org/abs/2605.12317. arxiv:2605.12317.

Nicholas Jardine and Robin Sibson. *Mathematical Taxonomy*. Wiley series in probability and mathematical statistiscs. Wiley, 1971.

Christopher Jung, Sampath Kannan, and Neil Lutz. Service in Your Neighborhood: Fairness in Center Location. In Aaron Roth (ed.), *1st Symposium on Foundations of Responsible Computing (FORC 2020)*, volume 156 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pp. 5:1–5:15, Dagstuhl, Germany, 2020. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. ISBN 978-3-95977-142-9. doi: 10.4230/LIPIcs.FORC.2020.5. URL https://drops.dagstuhl.de/entities/document/10.4230/LIPIcs.FORC.2020.5.

Yusuf Hakan Kalayci and Evi Micha. Temporal panel selection in ongoing citizens' assemblies. In *Proceedings of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2026. Forthcoming.

Leon Kellerhals and Jannik Peters. Proportional fairness in clustering: A social choice perspective. In *Proceedings of the 37th Conference on Neural Information Processing Systems (NeurIPS)*, pp. 111299–111317, 2024. URL http://papers.nips.cc/paper_files/paper/2024/hash/c981fd12b1d5703f19bd8289da9fc996-Abstract-Conference.html.

Jon M. Kleinberg. An impossibility theorem for clustering. In *Proceedings of the 15th Conference on Neural Information Processing Systems (NIPS)*, pp. 446–453. MIT Press, 2002. URL https://proceedings.neurips.cc/paper/2002/hash/43e4e6a6f341e00671e123714de019a8-Abstract.html.

Marina Knittel, Max Springer, John Dickerson, and Mohammad Hajiaghayi. Generalized reductions: making any hierarchical clustering fair and balanced with low cost. In *Proceedings of the 40th International Conference on Machine Learning (ICML)*. JMLR.org, 2023a.

Marina Knittel, Max Springer, John P. Dickerson, and MohammadTaghi Hajiaghayi. Fair, polylog-approximate low-cost hierarchical clustering. In *Proceedings of the 36th Conference on Neural Information Processing Systems (NeurIPS)*, 2023b. URL http://papers.nips.cc/paper_files/paper/2023/hash/8c533f53f76c4df9a7f08e7cb676d132-Abstract-Conference.html.

Bo Li, Lijun Li, Ankang Sun, Chenhao Wang, and Yingfan Wang. Approximate group fairness for clustering. In *Proceedings of the 38th International Conference on Machine Learning (ICML)*, pp. 6381–6391, 2021. URL https://proceedings.mlr.press/v139/li21j.html.

Youpeng Li, Xuyu Wang, and Lingling An. Hierarchical clustering-based personalized federated learning for robust and fair human activity recognition. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 7(1):20:1–20:38, 2023. doi: 10.1145/3580795. URL https://doi.org/10.1145/3580795.

Guolong Lin, Chandrashekhar Nagarajan, Rajmohan Rajaraman, and David P. Williamson. A general approach for incremental approximation and hierarchical clustering. *SIAM Journal on Computing*, 39(8):3633–3669, 2010. doi: 10.1137/070698257. URL https://doi.org/10.1137/070698257.

Sepideh Mahabadi and Ali Vakilian. Individual fairness for k-clustering. In *Proceedings of the 37th International Conference on Machine Learning (ICML)*, pp. 6586–6596, 2020. URL https://proceedings.mlr.press/v119/mahabadi20a.html.

Binita Maity and Shrutimoy Das. Individual fairness in hierarchical clustering, 2026. URL https://arxiv.org/abs/2608.25586.

Evi Micha and Nisarg Shah. Proportionally Fair Clustering Revisited. In *Proceedings of the 47th International Colloquium on Automata, Languages, and Programming (ICALP)*, volume 168, pp. 85:1–85:16, 2020. URL https://drops.dagstuhl.de/entities/document/10.4230/LIPIcs.ICALP.2020.85.

Benjamin Moseley and Joshua R. Wang. Approximation bounds for hierarchical clustering: Average linkage, bisecting k-means, and local search. In *Proceedings of the 31st Conference on Neural Information Processing Systems (NeurIPS)*, pp. 3094–3103, 2017. URL https://proceedings.neurips.cc/paper/2017/hash/d8d31bd778da8bdd536187c36e48892b-Abstract.html.

Fionn Murtagh and Pedro Contreras. Algorithms for hierarchical clustering: an overview. *WIREs Data Mining and Knowledge Discovery*, 2(1):86–97, 2012. doi: https://doi.org/10.1002/widm.53. URL https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/widm.53.

Greg Plaxton. Approximation algorithms for hierarchical location problems. *Journal of Computer and System Sciences*, 72(3):425–443, 2006. ISSN 0022-0000. doi: https://doi.org/10.1016/j.jcss.2005.09.004. URL https://www.sciencedirect.com/science/article/pii/S002200000500111X.

Xingcheng Ran, Yue Xi, Yonggang Lu, Xiangwen Wang, and Zhenyu Lu. Comprehensive survey on hierarchical clustering algorithms and the recent developments. *Artificial Intelligence Review*, 56(8):8219–8264, December 2022. ISSN 0269-2821. doi: 10.1007/s10462-022-10366-3. URL https://doi.org/10.1007/s10462-022-10366-3.

Alireza Rezazadeh, Zichao Li, Wei Wei, and Yujia Bao. From isolated conversations to hierarchical schemas: Dynamic tree memory representation for LLMs. In *Proceedings of the 13th International Conference on Learning Representations (ICLR)*, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/file/0382cb76309820f71c6eacd47b36ce71-Paper-Conference.pdf.

Parth Sarthi, Salman Abdullah, Aditi Tuli, Shubh Khanna, Anna Goldie, and Christopher D. Manning. RAPTOR: Recursive abstractive processing for tree-organized retrieval. In *Proceedings of the 12th International Conference on Learning Representations (ICLR)*, 2024. URL https://proceedings.iclr.cc/paper_files/paper/2024/hash/8a2acd174940dbca361a6398a4f9df91-Abstract-Conference.html.

Herbert E. Scarf. The core of an N person game. *Econometrica*, 35(1):50–69, 1967. doi: 10.2307/1909383.

Syeda Sanjana Shahid, Tara Salman, Mohamed Baza, and Marcio Andrey Teixeira. Hierarchical federated learning: Approaches, applications, and open challenges. *IEEE Access*, 13:190333–190353, 2025. doi: 10.1109/ACCESS.2025.3628455. URL https://doi.org/10.1109/ACCESS.2025.3628455.

P. H. A. Sneath and Robert R. Sokal. Numerical taxonomy. *Nature*, 193(4818):855–860, March 1962. doi: 10.1038/193855a0.

Alex Tamkin, Miles McCain, Kunal Handa, Esin Durmus, Liane Lovitt, Ankur Rathi, Saffron Huang, Alfred Mountfield, Jerry Hong, Stuart Ritchie, Michael Stern, Brian Clarke, Landon Goldberg, Theodore R. Sumers, Jared Mueller, William McEachen, Wes Mitchell, Shan Carter, Jack Clark, Jared Kaplan, and Deep Ganguli. Clio: Privacy-preserving insights into real-world AI use, 2024. URL https://arxiv.org/abs/2412.13678.

Xiaohong Yang, Tong Xie, Minghui Liwang, Chikai Shang, Yang Lu, Zhenzhen Jiao, Liqun Fu, and Seyyedali Hosseinalipour. Elsa: Efficient llm-centric split aggregation for privacy-aware hierarchical federated learning over the network edge, 2026. URL https://arxiv.org/abs/2601.13824. arXiv:2601.13824.

Daniel Zügner, Bertrand Charpentier, Morgane Ayle, Sascha Geringer, and Stephan Günnemann. End-to-end learning of probabilistic hierarchies on graphs. In *Proceedings of the 10th International Conference on Learning Representations (ICLR)*, 2022. URL https://openreview.net/forum?id=g2LCQwG7Of.

APPENDIX FOR ICLR-SUBMISSION #47062

A MISSING PROOFS FROM SECTION 3

A.1 PROOF OF THEOREM 1

Theorem 1 (★). *For every $\alpha \geq 1$, every $n \geq 4$, every agglomerative hierarchical clustering algorithm with standard linkage function f and any distance-bounded loss, there is a metric space $(\mathcal{N}, d)$ with $|\mathcal{N}| = n$ for which any clustering $\mathcal{C}^1, \ldots, \mathcal{C}^n$ returned by the algorithm is not in the global (α, δ)-core for any $\delta \leq \lfloor n^2/4 \rfloor / n$, and not in the local α-core.*

Proof. Let $B := \alpha \cdot n + 2$. Consider n agents placed on a line such that $d(1,2) = 1$ and $d(j, j+1) = (B-1) \cdot d(1,j)$ for every $j \in \{2, \ldots, n-1\}$ (as illustrated in Figure 3). Equivalently, $d(1,i) = B^{i-2}$ for every $i \in \{2, \ldots, n\}$.

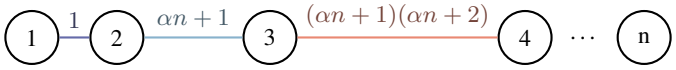

Figure 3: Counterexample for agglomerative methods on n agents.

We first show that every agglomerative algorithm with standard linkage function first merges $\{1\}$ and $\{2\}$, then merges the resulting cluster with $\{3\}$, and so on, until finally merging $\{1, \ldots, n-1\}$ with $\{n\}$. That is, after $n-1$ merges, it outputs the clustering sequence $\mathcal{C}^1, \ldots, \mathcal{C}^n$ with $\mathcal{C}^k = \big\{\{1, \ldots, (n-k)+1\}, \{(n-k)+2\}, \ldots, \{n\}\big\}$.

We prove this by induction on k. For $k = n$, this is true, since $\mathcal{C}^n$ is initialized as the singleton partition. Now suppose that after $n-k$ merges, the current clustering $\mathcal{C}^k$ is $\big\{\{1, \ldots, (n-k)+1\}, \{(n-k)+2\}, \ldots, \{n\}\big\}$. Let $C_k = \{1, \ldots, (n-k)+1\}$. Since f is a standard linkage function, $f(C_k, \{(n-k)+2\}) \leq d(1, (n-k)+2) = B^{n-k}$. On the other hand, for any two clusters $C, C' \in \mathcal{C}^k$ with $\{C, C'\} \neq \{C_k, \{(n-k)+2\}\}$, we have $f(C, C') \geq \min_{a \in C,\, b \in C'} d(a,b) \geq d((n-k)+2, (n-k)+3) > B^{n-k}$. Therefore, the next merge is between C_k and $\{(n-k)+2\}$, and so $\mathcal{C}^{k-1} = \big\{\{1, \ldots, (n-k)+2\}, \{(n-k)+3\}, \ldots, \{n\}\big\}$.

We now show that the returned hierarchical clustering is not in the global (α, δ)-core for any $\delta \leq \lfloor \frac{n^2}{4} \rfloor / n$.

Consider merge $k = \lceil 2\delta \rceil$. Then

$$\mathcal{C}^k = \big\{\{1, \ldots, (n-k)+1\}, \{(n-k)+2\}, \ldots, \{n\}\big\}.$$

Since $\delta \leq \lfloor \frac{n^2}{4} \rfloor / n$, we have $\delta \frac{n}{k} \leq (n-k)$. Thus a coalition of size $(n-k)$ is large enough to block at this level. Consider the coalition $S = \{1, \ldots, (n-k)\}$. Fix any $x \in S$. The cluster $\mathcal{C}^k(x)$ contains agent $(n-k)+1$. Since the loss is distance-bounded,

$$\begin{aligned} \mathrm{loss}_x(\mathcal{C}^k) &\geq \frac{1}{|\mathcal{C}^k(x)|} \sum_{y \in \mathcal{C}^k(x)} d(x,y) \\ &\geq \frac{1}{|\mathcal{C}^k(x)|} d(x, (n-k)+1) \geq \frac{1}{n} d((n-k), (n-k)+1). \end{aligned}$$

By construction,

$$d((n-k), (n-k)+1) = (B-1) \cdot d(1, (n-k)) = (B-1) \cdot \mathrm{diam}(S).$$

Therefore, $\mathrm{loss}_x(\mathcal{C}^k) \geq \frac{B-1}{n} \mathrm{diam}(S) > \alpha\, \mathrm{diam}(S)$. On the other hand, distance-boundedness also gives

$$\mathrm{loss}_x(S) \leq \max_{y \in S} d(x,y) \leq \mathrm{diam}(S).$$

Hence, $\mathrm{loss}_x(\mathcal{C}^k) > \alpha\, \mathrm{diam}(S) \geq \alpha\, \mathrm{loss}_x(S)$ for every $x \in S$. Thus S α-blocks $\mathcal{C}^k$. Therefore the returned hierarchical clustering is not in the global (α, δ)-core.

Algorithm 2: RECURSIVESPLIT($\mathcal{A}$)

Input : Metric space $(\mathcal{N}, d)$
Output: Hierarchical clustering $\mathcal{C}^1, \dots, \mathcal{C}^n$

$\mathcal{C}^1 \leftarrow \{\mathcal{N}\}$;
Initialize a queue $\mathcal{Q} \leftarrow (\mathcal{N})$;
for $k = 1, \dots, n-1$ **do**
  Choose and remove a non-singleton cluster C from $\mathcal{Q}$;
  $(C_1, C_2) \leftarrow \mathcal{A}(C, d|_C, 2)$;
  $\mathcal{C}^{k+1} \leftarrow \big(\mathcal{C}^k \setminus \{C\}\big) \cup \{C_1, C_2\}$;
  if $|C_1| > 1$ **then**
    Add C_1 to $\mathcal{Q}$;
  if $|C_2| > 1$ **then**
    Add C_2 to $\mathcal{Q}$;
return $\mathcal{C}^1, \dots, \mathcal{C}^n$

It remains to show that the returned hierarchical clustering is not in the local α-core. Let $\mathcal{T}$ be the tree corresponding to the clustering. Then the two children of the root node are $C_1 = \{1, \dots, n-1\}$ and $C_2 = \{n\}$. Consider the coalition $S = \{1, \dots, n-2\} \subseteq C_1$. Since $n \geq 4$, we have $|S| = n - 2 \geq n/2$. Thus S is large enough to block the local 2-clustering at the root. Fix any $x \in S$. The cluster C_1 contains agent $n-1$. By distance-boundedness, $\mathrm{loss}_x(C_1) \geq \frac{1}{|C_1|}\sum_{y \in C_1} d(x,y) \geq \frac{d(x,n-1)}{|C_1|} \geq \frac{d(n-2,n-1)}{n-1}$. By construction, $d(n-2, n-1) = (B-1)d(1, n-2) = (B-1)\,\mathrm{diam}(S)$. Thus, $\mathrm{loss}_x(C_1) \geq \frac{B-1}{n-1}\,\mathrm{diam}(S) > \alpha\,\mathrm{diam}(S)$.

Therefore, for every $x \in S$, $\mathrm{loss}_x(C_1) > \alpha\,\mathrm{loss}_x(S)$. Thus S α-blocks $\{C_1, C_2\}$, and the hierarchical clustering is not in the local α-core. □

B MISSING PROOFS FROM SECTION 4

B.1 PROOF OF THEOREM 2

Theorem 2 (★). *Let $\alpha \geq 1$. If $\mathcal{A}$ is an α-core algorithm for non-centroid 2-clustering, then* RECURSIVESPLIT($\mathcal{A}$) *always returns a hierarchical clustering in the local α-core.*

Proof. Suppose $\mathcal{A}$ is an α-core algorithm for non-centroid 2-clustering. Let $\mathcal{C}^1, \dots, \mathcal{C}^n$ be the hierarchical clustering returned by RECURSIVESPLIT($\mathcal{A}$) (Algorithm 2) and let $\mathcal{T}$ be its corresponding tree.

For every internal node C of $\mathcal{T}$, RECURSIVESPLIT($\mathcal{A}$) obtains C_1 and C_2 by running $\mathcal{A}$ on the induced metric space $(C, d|_C)$. Since $\mathcal{A}$ returns a non-centroid 2-clustering $\{C_1, C_2\}$ for C in the α-core for every internal node C, the hierarchical clustering returned by RECURSIVESPLIT($\mathcal{A}$) is in the local α-core. □

B.2 PROOF OF LEMMA 1

A cooperative game with non-transferable utility (NTU game) consists of a set of agents $\mathcal{N}$ and, for each nonemty coalition $\emptyset \neq S \subseteq \mathcal{N}$, a closed set $\emptyset \subsetneq V(S) \subsetneq \mathbb{R}^S$ of payoff vectors attainable by S with the following properties:

- $V(S)$ is *comprehensive*, i.e., for all $\mathbf{a} \in V(S)$ and all $\mathbf{b} \in \mathbb{R}^S$ with $\mathbf{b} \leq \mathbf{a}$ (where the inequality holds for every coordinate), we have $\mathbf{b} \in V(S)$.
- $V(S)$ is *bounded from above*, i.e., there is a constant M such that $a_i \leq M$ for all $\mathbf{a} \in V(S)$ and $i \in S$.

The set $V(S)$ represents the payoff or utility vectors that the members of S can obtain on their own, wherein for $\mathbf{a} \in V(S)$ and $i \in S$, a_i is the payoff obtainable by agent i.

For a coalition $S \subseteq \mathcal{N}$, define $\widehat{V}(S) \coloneqq V(S) \times \mathbb{R}^{\mathcal{N} \setminus S}$. Thus $\widehat{V}(S) \subseteq \mathbb{R}^{\mathcal{N}}$ consists of all the payoff vectors that are feasible for the agents in S, with no restrictions on the payoffs of agents outside S.

A family of nonnegative weights $(\lambda_S)_{\emptyset \neq S \subseteq \mathcal{N}}$ is *balanced* if $\sum_{S: i \in S} \lambda_S = 1$ for every $i \in \mathcal{N}$.

A coalition S *strictly blocks* a payoff vector $\mathbf{a} \in \mathbb{R}^{\mathcal{N}}$ if there exists $\mathbf{b} \in V(S)$ such that $b_i > a_i$ for every $i \in S$.

We use the following classical result of Scarf (1967).

Theorem 10 (Scarf's Theorem (Scarf, 1967))**.** *Consider an NTU game consisting of a set $\mathcal{N}$ of agents and payoff vectors $V(S)$ for each $\emptyset \neq S \subseteq \mathcal{N}$. If, for every balanced family $(\lambda_S)_{\emptyset \neq S \subseteq \mathcal{N}}$,*

$$\bigcap_{S: \lambda_S > 0} \widehat{V}(S) \subseteq V(\mathcal{N}),$$

then there exists a payoff vector $\mathbf{a}^\star \in V(\mathcal{N})$ that is not strictly blocked by any coalition $\emptyset \neq S \subseteq \mathcal{N}$.

Intuitively, Scarf's Theorem on balanced NTU games guarantees the existence of a feasible payoff vector for all agents that no coalition can strictly improve upon for all of its members.

Lemma 1 (★). *For every $k \geq 2$, there exists a fractional k-clustering in the core.*

Proof. Let $S \subseteq \mathcal{N}$. We define

$$V(S) \coloneqq \begin{cases} \{\mathbf{a} \in \mathbb{R}^{\mathcal{N}} : \text{ there is a fractional clustering } \lambda \text{ with } a_i \leq -\operatorname{loss}_i(\lambda)\} & \text{if } S = \mathcal{N}, \\ \{\mathbf{a} \in \mathbb{R}^S : a_i \leq -\operatorname{loss}_i(S) \text{ for every } i \in S\} & \text{if } n/k \leq |S| < n, \\ \{\mathbf{a} \in \mathbb{R}^S : a_i \leq -\operatorname{diam}(\mathcal{N})\} & \text{if } S < n/k. \end{cases}$$

We note that the corresponding NTU game is both comprehensive and bounded from above by 0. Further, any balanced family $(\lambda_S)_{\emptyset \neq S \subseteq \mathcal{N}}$ is equivalent to a fractional clustering in which we ignore the sets of size less than n/k. Lastly, by definition, we have $a_i \leq -\operatorname{loss}_i(\lambda)$ for every $\mathbf{a} \in \bigcap_{S: \lambda_S > 0} \widehat{V}(S)$ and hence also $\mathbf{a} \in V(\mathcal{N})$. Thus, there exists an unblocked $\mathbf{a} \in V(\mathcal{N})$ and therefore also a fractional clustering in the core. □

B.3 Proof of Theorem 5

Theorem 5 (★). *For every $k \geq 3$ there exists a metric space with $n \geq 9$ agents, where n is divisible by k, for which no k-clustering is in the α-core under the max-loss objective, for any $\alpha < 2^{1/5} < 1.149$. Moreover, for $k = 3$ there is a two-dimensional Euclidean space with $n = 9$ agents for which no k-clustering is in the α-core under the max-loss objective, for any $\alpha < 1.138$.*

Proof. We first focus on the result for general metric spaces before turning to the points for Euclidean spaces.

General metric spaces. Fix $k = 3$ and let $n \geq 9$ divisible by 3. We create three sets of agents, G_1, G_2, and G_3, which we will henceforth call *groups*, with

$$|G_1| = |G_3| = \frac{n}{k} - 1, \ |G_2| = \frac{n}{k} - 2,$$

and $d(i, j) = 0$ for each $i, j \in G_t$ and each $t \in \{1, 2, 3\}$. The remaining four agents will be called a, b, c, and w. We have the following distances for all agent–group pairs (as illustrated in Figure 4).

	G_1	G_2	G_3
a	$2^{\frac{1}{5}}$	1	$(1 + 2^{\frac{4}{5}})$
b	$(1 + 2^{\frac{1}{5}})$	2	$2^{\frac{4}{5}}$
c	$2^{\frac{2}{5}}$	$(1 + 2^{\frac{3}{5}})$	$2^{\frac{3}{5}}$

Moreover, we have $d(a, b) = 1$, and agent w is very far away from all other agents, that is, $d(w, i) = M$ for all $i \neq w$ and a sufficiently large $M \gg 1$. The remaining undefined distances can be chosen as the lengths of shortest paths on the weighted graph induced by the specified edges.

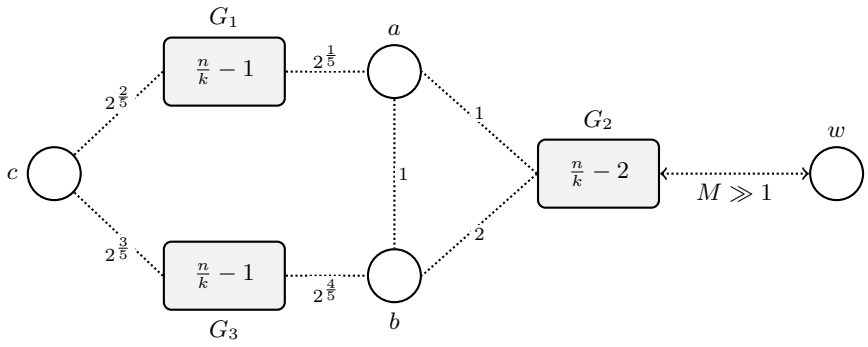

Figure 4: Counterexample for non-centroid max-loss: edge labels specify their length; the distance between any two points is the length of the shortest path.

In our proof, we focus on the five following coalitions of size $\frac{n}{k}$:

$$S_1 \coloneqq G_1 \cup \{a\}, \quad S_2 \coloneqq G_1 \cup \{c\}, \quad S_3 \coloneqq G_2 \cup \{a, b\},$$
$$S_4 \coloneqq G_3 \cup \{b\}, \quad S_5 \coloneqq G_3 \cup \{c\}.$$

For the sets S_1, S_2, S_4 and S_5, and each agent i within that set, we have

$$\mathrm{loss}_i(S_1) = 2^{\frac{1}{5}},\ \mathrm{loss}_i(S_2) = 2^{\frac{2}{5}},\ \mathrm{loss}_i(S_4) = 2^{\frac{4}{5}},\ \text{ and }\ \mathrm{loss}_i(S_5) = 2^{\frac{3}{5}}.$$

For S_3 and each $i \in S_3 \setminus \{a\}$ we have $\mathrm{loss}_i(S_3) = 2$ while the loss of agent a is $\mathrm{loss}_a(S_3) = 1$.

Let $\mathcal{C} = (C_1, C_2, C_3)$ be any 3-clustering of the gadget. One cluster must contain w, and any such cluster has a loss of M for all its members, unless w is placed alone as a singleton. Observe that the cluster containing w contains at most $\frac{n}{k}$ members, otherwise the cluster's agents without w form a blocking coalition for an α-core, with $\alpha \geq M/2$.

We distinguish between two cases, namely whether or not $\mathcal{C}$ contains a *mixed cluster*. Here, we call a cluster in $\mathcal{C}$ mixed if it contains agents from more than one of the three groups G_1, G_2, G_3.

CASE 1: THERE IS NO MIXED CLUSTER. We claim that, if $\mathcal{C}$ has a separate cluster for each of the groups G_1, G_2, G_3, then one of $S_1, \ldots, S_5$ is an α-blocking coalition.

If each group lies in its own cluster, then one of the three must also contain w. As mentioned above, we assume without loss of generality the cluster containing w has size at most n/k.

G_1 with w. Then coalition $S_2 = G_1 \cup \{c\}$ α-blocks $\mathcal{C}$: Clearly, the loss of each agent $i \in G_1$ in S_2 is smaller than in $\mathcal{C}$ by a factor greater than α. Since $|\mathcal{C}(w)| \leq n/k$, agent c must be clustered with agents from G_2 or G_3; thus its loss is

$$\mathrm{loss}_c(\mathcal{C}) \geq 2^{\frac{3}{5}} > \alpha 2^{\frac{2}{5}} = \alpha\, \mathrm{loss}_c(S_2).$$

G_3 with w. Suppose first that b is clustered with G_1. Then due to our above assumption of having no mixed cluster, the third cluster contains G_2. For c, this implies that it is clustered either with b or with G_2, thus

$$\mathrm{loss}_c(\mathcal{C}) \geq 2^{\frac{3}{5}} > \alpha 2^{\frac{2}{5}} = \alpha\, \mathrm{loss}_c(S_2).$$

And for all $i \in G_1$

$$\mathrm{loss}_i(\mathcal{C}) \geq d(b, i) > 2^{\frac{3}{5}} > \alpha 2^{\frac{2}{5}} = \alpha\, \mathrm{loss}_i(S_2).$$

So $S_2 = G_1 \cup \{c\}$ α-blocks $\mathcal{C}$.

If b is clustered with G_2, then we claim that $S_4 = G_3 \cup \{b\}$ α-blocks $\mathcal{C}$. Again this is obvious for the members of G_3, and for b, we have

$$\mathrm{loss}_b(\mathcal{C}) \geq 2 > \alpha 2^{\frac{4}{5}} = \alpha\, \mathrm{loss}_b(S_4).$$

G_2 with w.

- If b is clustered with G_2 and w, then $S_3 = G_2 \cup \{a, b\}$ α-blocks $\mathcal{C}$. The improvement is clear for all agents clustered with w. For agent a we have $\text{loss}_a(\mathcal{C}) \geq 2^{\frac{1}{5}} > \alpha = \alpha \, \text{loss}_a(S_3)$.
- If b is clustered with G_1, leading to a loss of at least $(1 + 2^{\frac{1}{5}})$ for all its members, then $S_2 = G_1 \cup \{c\}$ α-blocks, as $\text{loss}_c(\mathcal{C}) \geq 2^{\frac{3}{5}} > \alpha 2^{\frac{2}{5}} = \alpha \, \text{loss}_c(S_2)$.
- Let b be clustered with G_3. If a and c are clustered with G_1, then $S_1 = G_1 \cup \{a\}$ α-blocks as $\text{loss}_i(\mathcal{C}) \geq 2^{\frac{2}{5}} > \alpha \cdot 2^{\frac{1}{5}} = \text{loss}_i(S_1)$ for all $i \in S_1$. Otherwise, one of a and c is not clustered with G_1. If c is with $G_3 \cup \{b\}$ or with $G_2 \cup \{w\}$, then $S_5 = G_3 \cup \{c\}$ α-blocks $\mathcal{C}$ as $\text{loss}_i(\mathcal{C}) \geq 2^{\frac{4}{5}} > \alpha \cdot 2^{\frac{3}{5}} = \alpha \, \text{loss}_i(S_5)$ for all $i \in S_5$. If instead a is clustered with $G_3 \cup \{b\}$ or with $G_2 \cup \{w\}$, then $S_1 = G_1 \cup \{a\}$ α-blocks $\mathcal{C}$ as $\text{loss}_i(\mathcal{C}) \geq 2^{\frac{2}{5}} > \alpha \cdot 2^{\frac{1}{5}} = \alpha \, \text{loss}_i(S_1)$ for all $i \in S_1$.

In all cases, we find an α-blocking coalition for some $\alpha \geq 2^{\frac{1}{5}}$.

CASE 2: THERE IS A MIXED CLUSTER. We claim that one of $S_1, \ldots, S_4$ α-blocks $\mathcal{C}$.

Suppose that there is only one mixed cluster and call it C_1. As the other two clusters are not mixed, C_1 must contain all members of one of the groups. Assume the mixed cluster contains w. Since the cluster containing w cannot have more than $\frac{n}{k}$ agents, C_1 must contain the agents of G_2, and one other agent x from G_1 or G_3. Now consider the clustering $\mathcal{C}'$ where x is moved to the cluster containing the remaining members of its group. Then $\mathcal{C}'$ does not contain a mixed cluster and therefore, by Case 1, it is not in the α-core. Now observe that for every agent $i \in N$ we have $\text{loss}_i(\mathcal{C}') \leq \text{loss}_i(\mathcal{C})$; thus if a coalition S α-blocks $\mathcal{C}'$, then it also α-blocks $\mathcal{C}$. Consequently, $\mathcal{C}$ is also not in the α-core.

Hence, we may assume our clustering $\mathcal{C}$ contains a mixed cluster that does not contain w. In any such clustering $\mathcal{C}$, every agent in a mixed cluster has loss at least $1 + 2^{\frac{1}{5}}$, and every agent in the cluster containing w has loss at least $M \gg (1 + 2^{\frac{1}{5}})$. If there is an agent with a smaller loss, then it lies in the third cluster, which then is not mixed. It follows that there are at least two groups whose members all incur a loss of at least $1 + 2^{\frac{1}{5}}$ in $\mathcal{C}$: One group in the cluster containing w and at least one group in the mixed cluster. We call such groups *deviating*.

Suppose that G_1 is deviating. If a and c are part of the same cluster, then $S_1 = G_1 \cup \{a\}$ is α-blocking: For each $i \in G_1$, we have

$$\alpha \cdot \text{loss}_i(S_1) = \alpha \cdot 2^{\frac{1}{5}} < 2^{\frac{2}{5}} < 1 + 2^{\frac{1}{5}} \leq \text{loss}_i(\mathcal{C}),$$

and for a we have

$$\alpha \cdot \text{loss}_a(S_1) = \alpha \cdot 2^{\frac{1}{5}} < 2^{\frac{2}{5}} < 2^{\frac{1}{5}} + 2^{\frac{2}{5}} = d(a, c) \leq \text{loss}_a(\mathcal{C}).$$

Similarly, if b and c are part of the same cluster, then $S_2 = G_1 \cup \{c\}$ is α-blocking as for each $i \in G_1$,

$$\alpha \cdot \text{loss}_i(S_2) = \alpha \cdot 2^{\frac{2}{5}} < 2^{\frac{1}{5}} \cdot 2^{\frac{2}{5}} \leq 1 + 2^{\frac{1}{5}} = \text{loss}_i(\mathcal{C}),$$

and for c we have

$$\alpha \cdot \text{loss}_c(S_2) = \alpha \cdot 2^{\frac{2}{5}} < 2^{\frac{3}{5}} + 2^{\frac{4}{5}} = d(b, c).$$

Finally, if a or c are clustered with w or with G_3, then again S_1 or S_2 α-blocks, because both a and c face losses exceeding $2^{\frac{1}{5}}$ or $2^{\frac{2}{5}}$, respectively, i.e., by a factor strictly larger than $2^{\frac{1}{5}}$.

If G_1 is not among the deviating groups, then b cannot be clustered with G_1 (otherwise G_1 would deviate, as $d(b, G_1) = 1 + 2^{\frac{1}{5}}$). Hence b is either clustered with w or in a mixed cluster, and in both cases its loss exceeds 2. Thus $S_4 = G_3 \cup \{b\}$ is α-blocking: For each $i \in S_4$, we have

$$\alpha \cdot \text{loss}_i(S_1) = \alpha \cdot 2^{\frac{4}{5}} < 2 \leq \text{loss}_i(\mathcal{C}).$$

Thus in every clustering where at least one cluster is mixed, one of the coalitions $S_1, \ldots, S_4$ strictly blocks by a factor of at least $2^{\frac{1}{5}}$, contradicting α-core stability.

Finally, to extend to arbitrary $k \geq 3$ and n divisible by k, add $(k - 3)$ dummy clusters of size $\frac{n}{k}$ each, with intra-distance 0 and distance M to every outside node. In any feasible k-clustering, these dummy clusters must stand alone, so the effective choice reduces to clustering the gadget $\{a, b, c, w, G_1, G_2, G_3\}$.

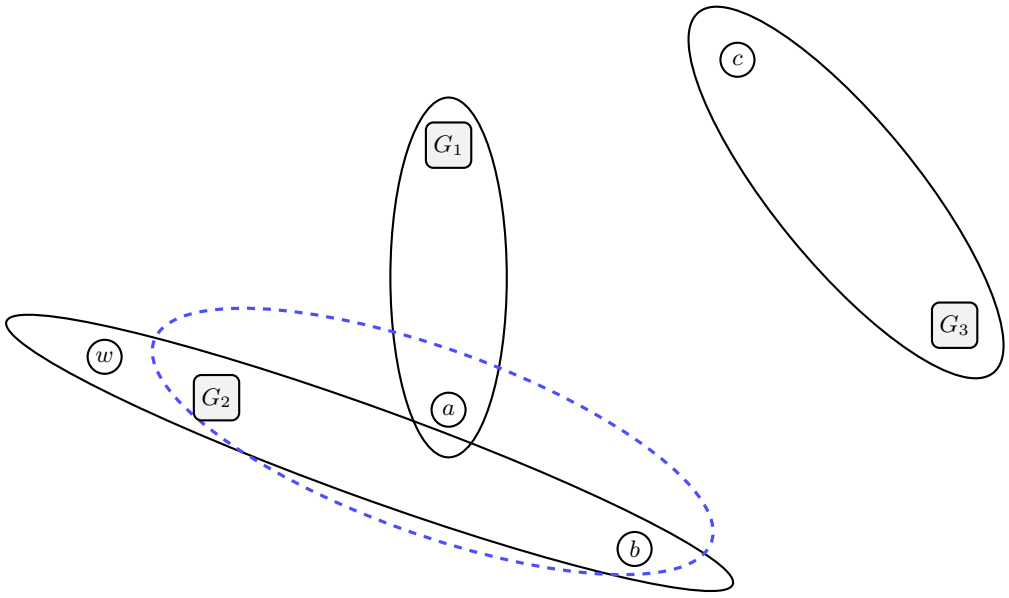

Figure 5: The points for the counterexample in two-dimensional Euclidean space (rotated by 90 degrees). w was placed closer to the remaining points for an easier illustration. The black ellipses depict the clustering that is in the 1.139-core and the blue (dashed) ellipse is the 1.138-blocking coalition.

Algorithm 3: CORECHECK

Input : Metric space $(\mathcal{N}, d)$, clustering $\mathcal{C}$, and loss function loss
Output: Smallest $\alpha \geq 1$ such that $\mathcal{C}$ is in the α-core, or ∞ if no finite factor exists

$\alpha \leftarrow 1$; $n \leftarrow |\mathcal{N}|$; $k \leftarrow |\mathcal{C}|$;
for *each* $S \subseteq \mathcal{N}$ *with* $|S| \geq n/k$ **do**
  for *each* $i \in S$ **do**
    $\rho_i \leftarrow \text{loss}_i(\mathcal{C}) / \text{loss}_i(S)$;
  $\alpha \leftarrow \max\{\alpha, \min_{i \in S} \rho_i\}$;
return α

Points for the Euclidean space. To prove the lower bound for the Euclidean space for $n = 9$ and $k = 3$ uses the same set of points as in Theorem 5, with each (group of) points having roughly the same "function"; see Figure 5 for an illustration.

The points were obtained by optimizing a system of inequalities and are as follows.

$$
\begin{aligned}
a &: (0.00000000,\ 0.00000000),\\
b &: (-0.52633831,\ -0.70311454),\\
c &: (1.32270616,\ -1.09187983),\\
G_1 &: (1.00000000,\ 0.00000000),\\
G_2 &: (0.04545127,\ 0.87711815),\\
G_3 &: (0.31938656,\ -1.91276308),\\
w &: (10.00000000,\ 10.00000000).
\end{aligned}
$$

To find an α-core clustering with the smallest α for the set of points, we applied Algorithm 3 to each possible clustering $\mathcal{C}$. □

B.4 Proof of Theorem 6

Theorem 6 (★). *For the average-loss objective and $k < n$,* AverageGreedyCapture *returns a clustering in the $(\lceil n/k \rceil - 1)$-core in polynomial time. Consequently,* RecursiveSplit(AverageGreedyCapture) *computes a hierarchical clustering in the local $(\lceil n/2 \rceil - 1)$-core in polynomial time.*

Proof. We construct the clustering iteratively. Let $\mathcal{N}'$ denote the set of remaining agents. If $|\mathcal{N}'| < n/k$, we add $\mathcal{N}'$ to the clustering and stop. Otherwise, for every $i \in \mathcal{N}'$, let C_i consist of i and its $\lceil n/k \rceil - 1$ closest agents in $\mathcal{N}'$. Choose an agent $i^\star \in \mathcal{N}'$ minimizing $\mathrm{loss}_i^{\mathrm{avg}}(C_i)$, add $C_{i^\star}$ to the clustering, and remove $C_{i^\star}$ from $\mathcal{N}'$. If fewer than k clusters are created, repeatedly select an arbitrary agent from a non-singleton cluster and make it a singleton cluster until there are exactly k clusters. We show that the resulting clustering $\mathcal{C}$ is in the $(\lceil n/k \rceil - 1)$-core.

Let $S \subseteq \mathcal{N}$ be any subset of agents with $|S| \geq n/k$. Let $C = C_{i^\star}$ be the first cluster constructed by the algorithm that intersects S, and choose $j \in C \cap S$. At the beginning of the iteration in which C is constructed, no agent of S has yet been removed. Hence $S \subseteq \mathcal{N}'$.

With this, we now get that $\sum_{x \in C} d(j,x) \leq \sum_{x \in C \setminus \{j\}} (d(j,i^\star) + d(i^\star,x)) = (\lceil n/k \rceil - 2) d(i^\star, j) + \sum_{x \in C} d(i^\star, x) \leq (\lceil n/k \rceil - 1) \sum_{x \in C} d(i^\star, x)$. Thus, by the choice of $i^\star$, we have $\mathrm{loss}_j^{\mathrm{avg}}(C) \leq (\lceil n/k \rceil - 1)\, \mathrm{loss}_{i^\star}^{\mathrm{avg}}(C) \leq (\lceil n/k \rceil - 1)\, \mathrm{loss}_j^{\mathrm{avg}}(S)$. The same holds after splitting off singletons: if $C' \subseteq C$ is the final cluster containing j, the triangle inequality gives $\mathrm{loss}_j^{\mathrm{avg}}(C') \leq \frac{|C'|-1}{|C'|} \sum_{x \in C} d(i^\star, x) \leq (|C| - 1)\, \mathrm{loss}_{i^\star}^{\mathrm{avg}}(C)$. Hence, the clustering is in the $(\lceil n/k \rceil - 1)$-core.

We now show that the analysis is tight for $k = 2$ and $n \geq 6$ using the instance depicted in Figure 6. In this instance, one agent is located at each of $-(n/2)^2, -1, \varepsilon, (n/2) - 2$ agents are located at 0, and $(n/2) - 1$ agents are located at 1, where $0 < \varepsilon < 1/4$.

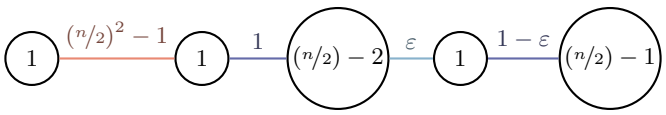

Figure 6: Example to show lower bounds for AverageGreedyCapture in Theorem 6.

The algorithm first selects the agents at 1 together with the agent at ε, since a center at 1 has average loss $\frac{1-\varepsilon}{(n/2)}$, which is strictly smaller than at any other location. The remaining agents form the second cluster.

Now consider the coalition S of the $(n/2)$ agents at $-1, 0, \varepsilon$. The agent at ε improves by a factor $\frac{((n/2)-1)(1-\varepsilon)}{1+((n/2)-1)\varepsilon}$ which tends to $(n/2) - 1$ as ε approaches 0. The agents at 0 and -1 improve by factors $\frac{((n/2)^2+1)}{(1+\varepsilon)}$ and $\frac{((n/2)^2+(n/2)-3)}{((n/2)-1+\varepsilon)}$, respectively, both greater than $(n/2) - 1$. Thus, for every $\alpha < (n/2) - 1$, choosing ε sufficiently small makes S α-block the returned clustering.

By Theorem 2 we can thus use AverageGreedyCapture as $\mathcal{A}$ in Algorithm 2 to show that RecursiveSplit(AverageGreedyCapture) returns a hierarchical clustering in the local $(\lceil n/2 \rceil - 1)$-core in polynomial time. □

C Missing proofs from Section 5

C.1 Proof of Theorem 7

Theorem 7 (★). *For every $\alpha \geq 1$, there is a metric space with eight agents such that for the maximum loss objective, there is a hierarchical clustering that is in the global core but not in the local α-core.*

Proof. Let $D := \alpha + 1$. Consider the metric space on agents $\{1, \ldots, 8\}$ induced by the shortest-path metric of the following weighted graph, illustrated in Figure 7.

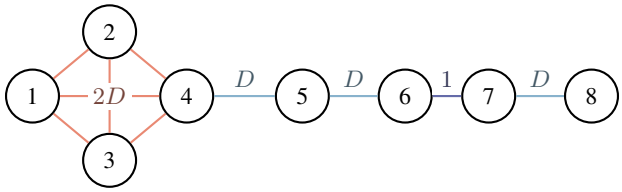

Figure 7: Metric on agents $\{1, \ldots, 8\}$ inducing the shortest-path distances. Every edge inside the clique $\{1, 2, 3, 4\}$ has length $2D$.

The agents $\{1, 2, 3, 4\}$ form a clique with edge length $2D$. Moreover, $d(4, 5) = d(5, 6) = d(7, 8) = D$ and $d(6, 7) = 1$. All remaining distances are shortest-path distances.

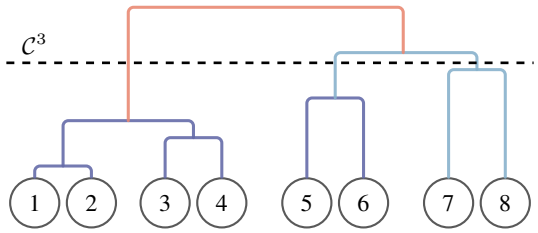

Figure 8: Dendrogram representation of a hierarchical clustering that is in the global core, but not in the local α-core.

We first show that the hierarchical clustering $\mathcal{C}^1, \ldots, \mathcal{C}^8$ depicted in Figure 8 is in the global core. For $t \in \{0, 1, 2, 3\}$ agents $5, 6, 7, 8$ are singletons and hence have loss zero. Thus none of them can belong to a blocking coalition. Among the agents $1, 2, 3, 4$, all pairwise distances are equal to $2(\alpha + 1)$, so no subset of these agents can make all of its members strictly better off. Hence no coalition blocks at these levels.

In clustering $\mathcal{C}^4$, the newly formed cluster is $C_4 = \{5, 6\}$. Agent 5 already has minimum possible loss among coalitions of size two. Agent 6 could improve by pairing with agent 7, but agent 7 is still a singleton and therefore would not strictly improve. Thus there is no blocking coalition.

Now consider $k = 3$ and the clustering $\mathcal{C}^3 = \big\{\{1, 2, 3, 4\}, \{5, 6\}, \{7, 8\}\big\}$. At this cut there are three clusters, hence any blocking coalition must have size at least $\frac{n}{3} = \frac{8}{3}$. Thus only coalitions of size at least 3 may block. The current loss of every agent in $\{5, 6\}$ and $\{7, 8\}$ is $\alpha + 1$. For such an agent to strictly improve, every other member of a blocking coalition would have to be at distance strictly less than $\alpha+1$. But the only distance between the two pairs that is smaller than $\alpha+1$ is $d(6, 7) = 1$. Thus the only possible improving coalition involving these agents is the pair $\{6, 7\}$, which has size 2 and is therefore too small to block at this global level. Any coalition of size at least 3 containing one of the agents $5, 6, 7, 8$ gives at least one of its members loss at least $\alpha + 1$, so it cannot block.

Lastly, at $k = 2$, any blocking coalition must have size at least $\frac{n}{2} = 4$. The pair $\{7, 8\}$ cannot participate in a blocking coalition of size at least 4, because any such coalition containing one of these agents also contains an agent at distance at least $2\alpha+1$, so not all its members strictly improve. Similarly for $\{5, 6\}$, since any such coalition also contains two additional agents, one of which is at distance at least $3\alpha + 1$. It remains to consider coalitions contained in $\{1, 2, 3, 4\}$. Any coalition of size at least 4 contained in $\{1, 2, 3, 4\}$ must be exactly $\{1, 2, 3, 4\}$, and therefore contains agent 4. This coalition cannot strictly improve agent 4, whose current loss is at most $2(\alpha + 1)$. Hence no coalition of size at least 4 blocks.

Therefore $\mathcal{C}^1, \ldots, \mathcal{C}^8$ is in the global core. However, $\mathcal{C}^1, \ldots, \mathcal{C}^8$ is not in the local α-core: Consider the node C_2 in $\mathcal{T}$ with children $C_4 = \{5, 6\}$ and $C_3 = \{7, 8\}$. The coalition $S = \{6, 7\}$ of size at least $\frac{4}{2}$ is α-blocking for C_2, since $\mathrm{loss}_x(\{C_5, C_4\}) = \alpha + 1 > \alpha = \alpha \cdot \mathrm{loss}_x(S)$. □

C.2 Proof of Theorem 8

Theorem 8 (★). *For every finite $\alpha \geq 1$ and for any distance-bounded loss, there exists an instance with six agents for which no hierarchical clustering is in the global α-core.*

Proof. Fix $\alpha \geq 1$. We construct an instance with six agents $\mathcal{N} = \{1, \ldots, 6\}$ placed on the real line. Let $D := 3\alpha + 1$ and $M := 6\alpha(D+1) + 1$. The consecutive distances are $d(1,2) = 1$, $d(2,3) = D$, $d(3,4) = M$, $d(4,5) = D$, $d(5,6) = 1$ (as illustrated in Figure 9). Let $L := \{1,2,3\}$ and $R := \{4,5,6\}$ denote the left-most and right-most agents respectively.



Figure 9: Counterexample instance on six agents.

We show that every hierarchical clustering contains a clustering that is α-blocked by some coalition.

Consider an arbitrary hierarchical clustering. At the merge with two clusters, there are two cases.

First, suppose the two clusters are exactly L and R. Then, immediately before this merge, the hierarchical clustering had three clusters. Hence, one of L or R must already exist at the previous merge, i.e., when there were three clusters. For $n = 6$ and three clusters, any coalition of size at least $\frac{6}{3} = 2$ can block. Observe that the two close pairs α-block their respective triples. If the cluster L appears at some merge at which coalitions of size 2 may block, then $S = \{1,2\}$ α-blocks. For both agents $x \in S$, we have $\mathrm{loss}_x(S) \leq \mathrm{loss}_x^{\max}(S) = 1$. On the other hand, since loss is distance-bounded, $\mathrm{loss}_x(L) \geq \mathrm{loss}_x^{\mathrm{avg}}(L)$. For $x = 1$, $\mathrm{loss}_1^{\mathrm{avg}}(L) = \frac{D+2}{3}$, while for $x = 2$, $\mathrm{loss}_2^{\mathrm{avg}}(L) = \frac{D+1}{3}$. Since $D = 3\alpha + 1$, we obtain$\mathrm{loss}_x(L) \geq \mathrm{loss}_x^{\mathrm{avg}}(L) \geq \frac{D+1}{3} > \alpha \geq \alpha \cdot \mathrm{loss}_x(S)$. Thus $S = \{1,2\}$ α-blocks L. Symmetrically, if the cluster R appears at a level at which coalitions of size 2 may block, then $S = \{5,6\}$ α-blocks. Thus the newly created triple is α-blocked by either $\{1,2\}$ or $\{5,6\}$, a contradiction.

Second, suppose the two clusters are not exactly L and R. Then at least one of the two clusters contains agents from both sides of the middle gap, i.e., it contains some agent from L and some agent from R. Since there are only two clusters, it follows that either every agent in L belongs to a cluster that crosses the middle gap, or every agent in R belongs to a cluster that crosses the middle gap. Assume without loss of generality that this holds for L.

Consider the coalition $S = L$. Since the cluster containing each $x \in L$ crosses the middle gap, it contains some agent $y \in R$. Hence $d(x,y) \geq M$. By distance-boundedness,

$$\mathrm{loss}_x(\mathcal{C}) \geq \frac{1}{|\mathcal{C}(x)|} \sum_{z \in \mathcal{C}(x)} d(x,z) \geq \frac{1}{|\mathcal{C}(x)|} d(x,y) \geq \frac{M}{6}.$$

On the other hand, for every $x \in S = L$, we have $\mathrm{loss}_x(S) \leq \max_{z \in S} d(x,z) \leq \mathrm{diam}(S) \leq D+1$. Therefore, $\mathrm{loss}_x(\mathcal{C}) \geq \frac{M}{6} > \alpha(D+1) \geq \alpha \cdot \mathrm{loss}_x(S)$ for every $x \in S$. Thus this coalition α-blocks the two-cluster partition. Since at the two-cluster level coalitions of size at least $\frac{6}{2} = 3$ may block and $|S| = 3$, this is a feasible blocking coalition. The case in which every agent in R belongs to a cluster crossing the middle gap is symmetric.

In all, this shows that no hierarchical clustering for this metric space is in the global α-core. □