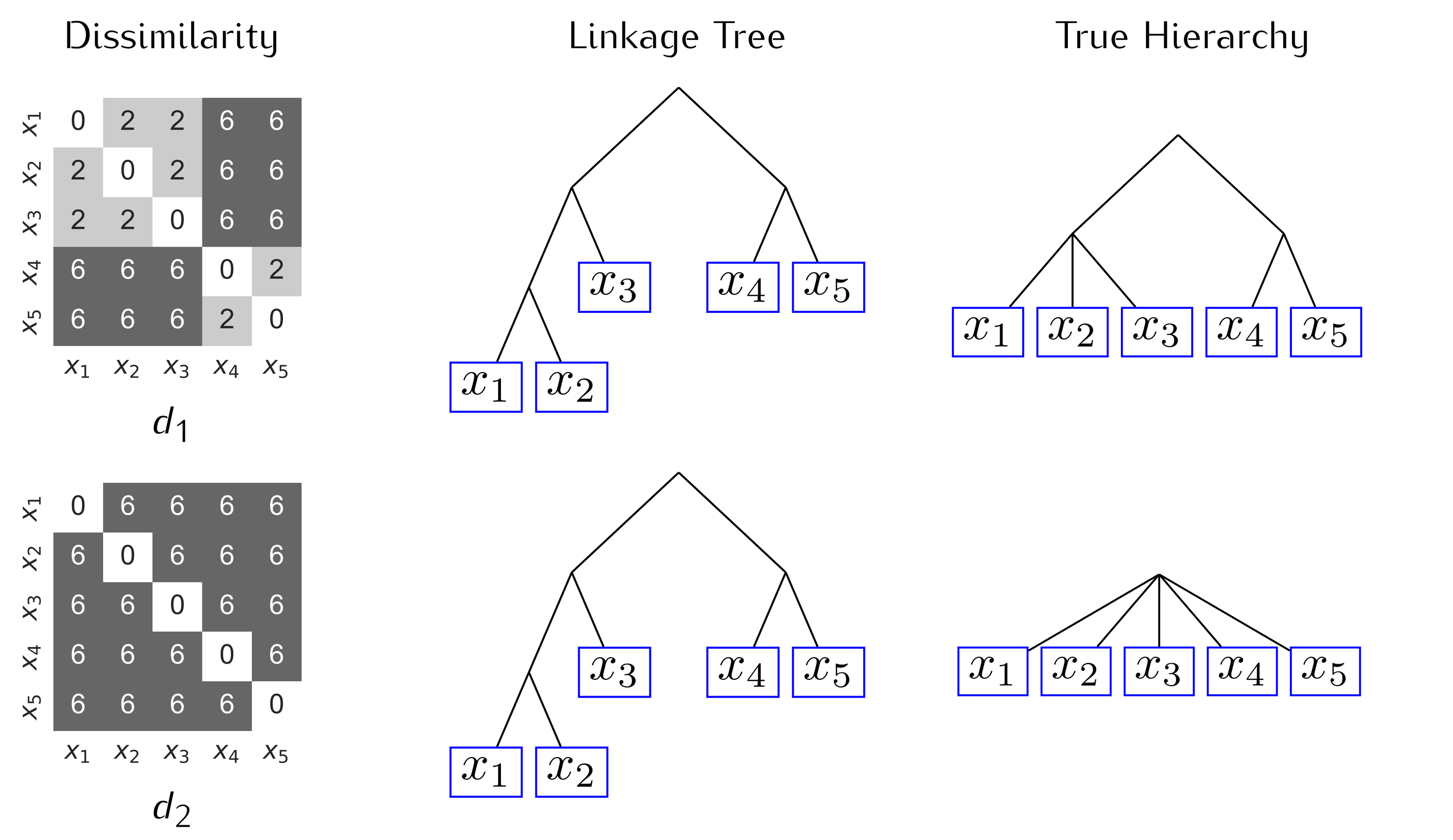
Beyond Binary Trees: Finding General Hierarchies

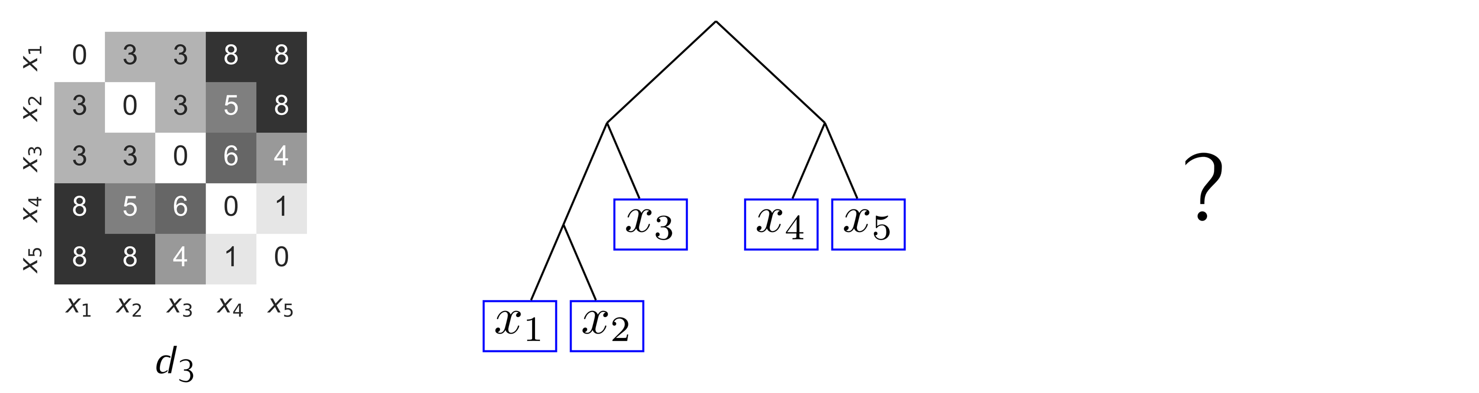
Maximilien Dreveton, Daichi Kuroda, Matthias Grossglauser, Patrick Thiran
INDY (Information and Network Dynamics), EPFL

Hierarchical Clustering

Hierarchical clustering: Representing data points $x \in \mathcal{X}$ by a tree that reflects the dissimilarity d.



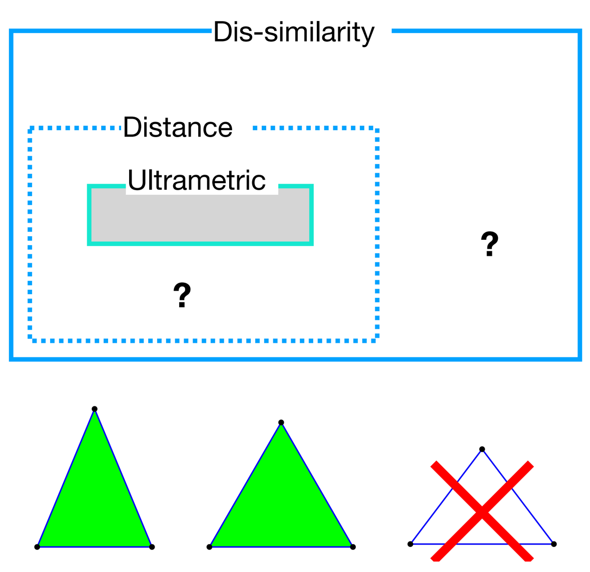
In the toy examples above, there is a natural hierarchy. But often it is a non-trivial task.



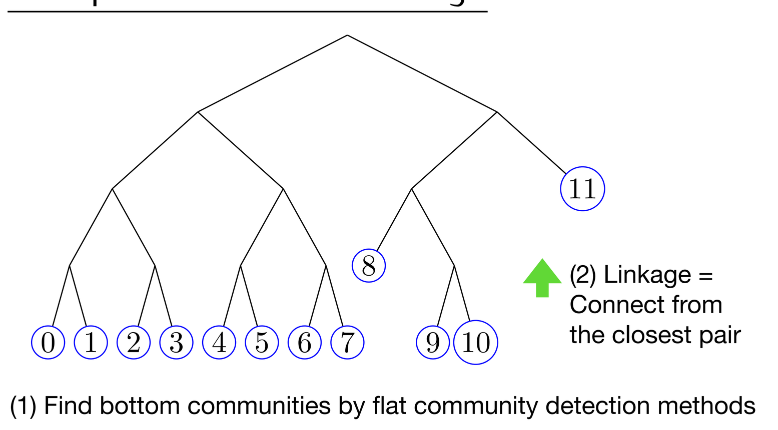
Actually, d needs to satisfy *ulatrametricity* to have a clear definition of hierarchy.

ultrametric distance: $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^+$

- $d(x, y) = d(y, x)$ (symmetry)
- $d(x, y) = 0$ iff $x = y$
- $d(x, z) < \max(d(x, y), d(y, z))$ (strong triangle inequality)



A Popular Method: Linkage



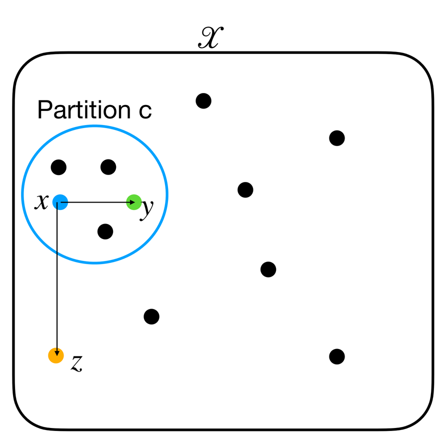
Limitations

- Overfit
  - Hallucinated levels
  - Infer hierarchy even when there is none
- Hierarchy is not always well-defined
  - Each hierarchical clustering outputs different trees and cannot decide which is the best

Nice Clusters

Definition: Nice Cluster $C \subseteq \mathcal{X}$ [6]
$C \subseteq \mathcal{X}$ is a nice cluster if for any $x, y \in C$ and $z \in \mathcal{X} \setminus C$,

$$d(x, z) - d(x, y) > 0. \qquad (1)$$



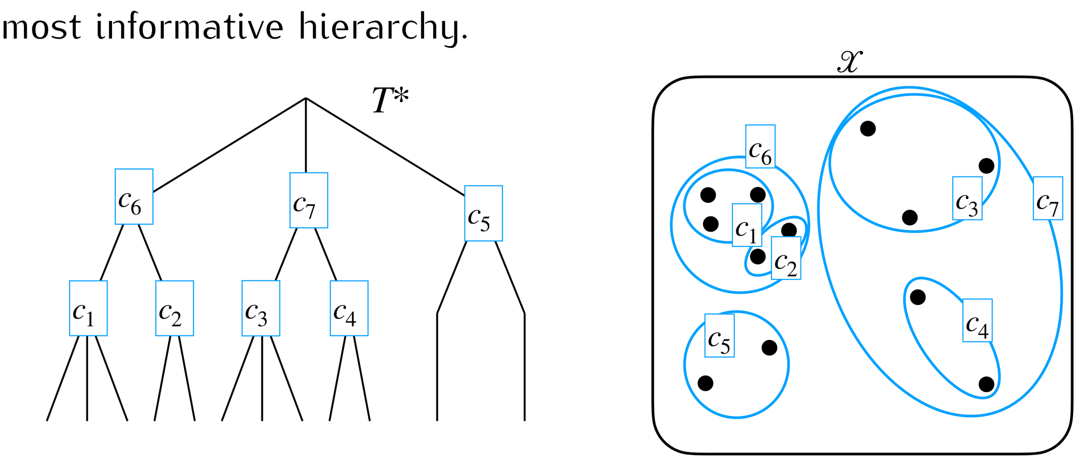
Most Informative Hierarchy

Theorem: The collection of all the nice clusters forms a hierarchy
Let be

$$T^* = \{C : C \subseteq \mathcal{X} \text{ that satisfies the nice cluster condition (1)}\}.$$

Then T^* can be represented as a rooted tree.

We call T^* most informative hierarchy.



Properties

1. T^* coincides with the natural hierarchy when d is ultrametric
2. $T^* \subseteq T_{linkage}$ always holds even though $T_{linkage}$ may not be valid; therefore T^* can be reconstructed by trimming vertices of $T_{linkage}$ that violate (1)

Handling Noise

$T_\varepsilon \in \mathcal{H}(\mathcal{X}, d, \varepsilon)$ satisfies for any $t \in T_\varepsilon$:

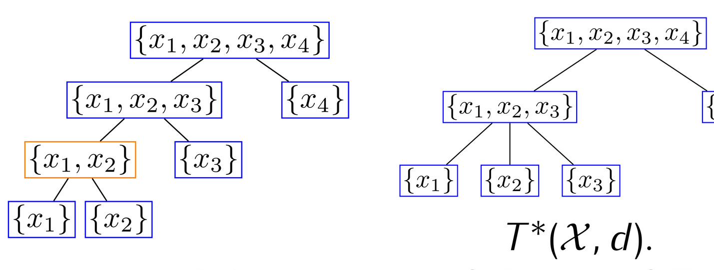
$$\min_{x,y \in t, z \in \mathcal{X} \setminus t} d(x, z) - d(x, y) > \varepsilon. \qquad (2)$$

$$\begin{pmatrix} 0 & 2 & 2 & 3 \\ 2 & 0 & 2 & 3 \\ 2 & 2 & 0 & 3 \\ 3 & 3 & 3 & 0 \end{pmatrix}$$

Dissimilarity d

$$\begin{pmatrix} 0 & 1.8 & 2.2 & 2.9 \\ 1.8 & 0 & 2.1 & 3.1 \\ 2.2 & 2.1 & 0 & 2.95 \\ 2.9 & 3.1 & 2.95 & 0 \end{pmatrix}$$

Dissimilarity $\hat{d}$.



$\varepsilon < 0.4$

$T^*(\mathcal{X}, d)$.
$0.4 \leq \varepsilon < 0.7$

Application: Hierarchical Community Detection

Based on [1], we propose the following algorithm:

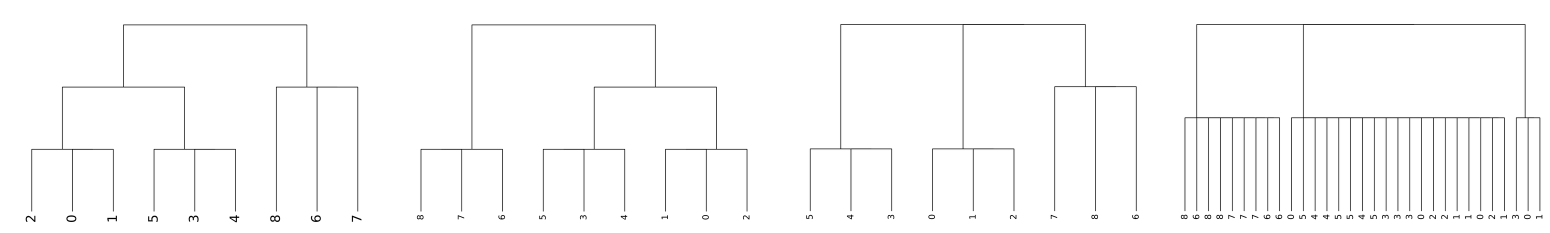
1. Find bottom communities $\hat{\mathcal{X}}$ by a flat community detection algorithm
2. Define $\hat{s} : \hat{\mathcal{X}} \times \hat{\mathcal{X}} \to [0, 1]$ as edge densities between the pair of communities $\hat{\mathcal{X}} \times \hat{\mathcal{X}}$
3. Apply average linkage to obtain $\hat{T}_{\text{linkage}}$
4. Trim every vertex of $\hat{T}_{\text{linkage}}$ that violates (2) to obtain $\hat{T}$

Table 1: Performance of HCD algorithms on 20 ABCD graphs (absence of hierarchies).

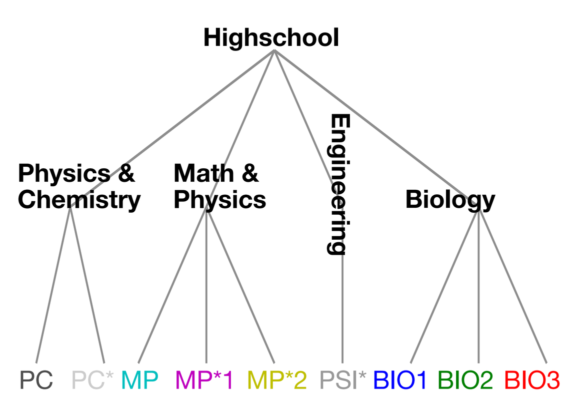
	$\hat{k}$	$\hat{k} = k$	$\rho(z^*, \hat{z})$	$\rho(T^*, \hat{T})$	$\hat{T} = T^*$
Our Algorithm	10 ± 0	20/20	0.97 ± 0.0038	0.97 ± 0.0038	20/20
Nested sEEP [4]	10 ± 0	20/20	0.97 ± 0.0038	0.95 ± 0.051	16/20
Nested DCBM [5]	10 ± 0.77	18/20	0.99 ± 0.0047	0.97 ± 0.077	18/20

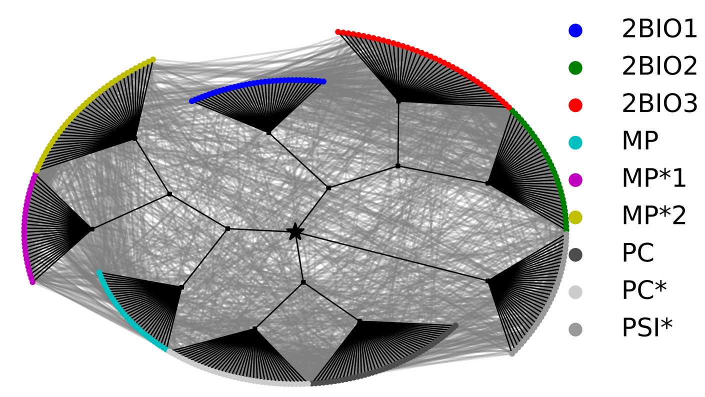
Table 2: Performance of HCD algorithms on HDCBMs (all possible hierarchies with 9 leaves).

	$\hat{k}$	$\hat{k} = k$	$\rho(z^*, \hat{z})$	$\rho(T^*, \hat{T})$	$\hat{T} \simeq T^*$
Our Algorithm	8.4 ± 0.84	62%	0.93 ± 0.10	0.98 ± 0.044	61.6%
Nested sEEP [4]	8.4 ± 0.84	62%	0.93 ± 0.10	0.94 ± 0.0812	36.8 %
Nested DCBM [5]	20 ± 10.3	0.01%	0.81 ± 0.098	0.64 ± 0.154	0.0 %



True hierarchy | Our Algorithm | Nested sEEP [4] | Nested DCBM [5]





References

1. M. Dreveton et. al. When does bottom-up beat top-down in hierarchical community detection? *arXiv*, 2023.
2. S. Dasgupta. A cost function for similarity-based hierarchical clustering. *STOC '16*, 2016.
3. T. Li et. al. Hierarchical community detection by recursive partitioning. *JASA*, 2022.
4. M. T. Schaub, J. Li, & L. Peel. Hierarchical community structure in networks. *Physical Review E*, 2023.
5. T. P. Peixoto. Hierarchical block structures and high-resolution model selection in large networks. *Physical Review X*, 2014.
6. M. Ackerman, & S. Dasgupta. . Incremental clustering: The case for extra clusters. *NeurIPS*, 2014.