Robust Estimation of Tree Structured Gaussian Graphical Models

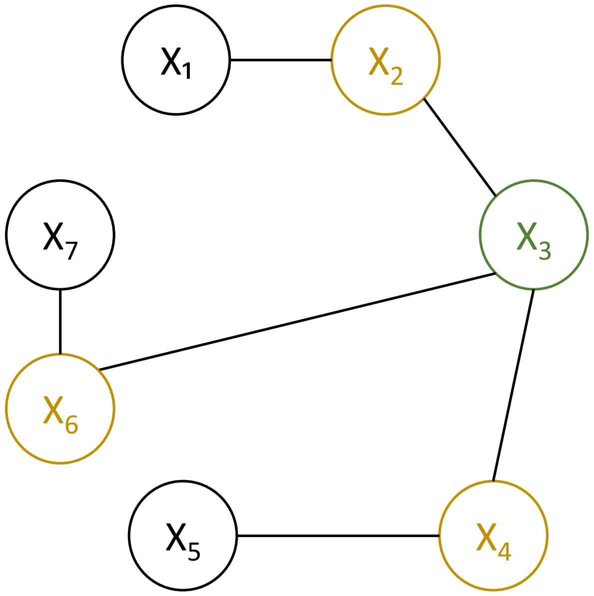
Ashish Katiyar, Jessica Hoffmann, Constantine Caramanis

The University of Texas at Austin

Undirected Probabilistic Graphical Models

- Represent conditional independence relations among random variables in the form of a graph.

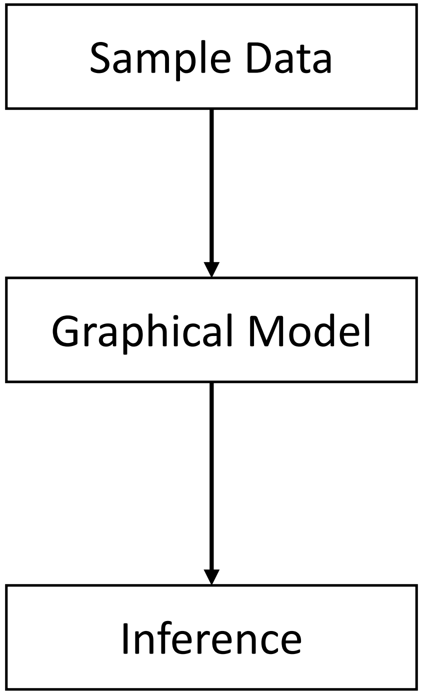
- Any random variable conditioned on the random variable it has an edge with, is independent of all the remaining random variables.



Why Graphical Models?

- Identify interactions among variables in large systems. (e.g. Gene Interaction Networks.)

- Makes inference in large scale systems tractable.

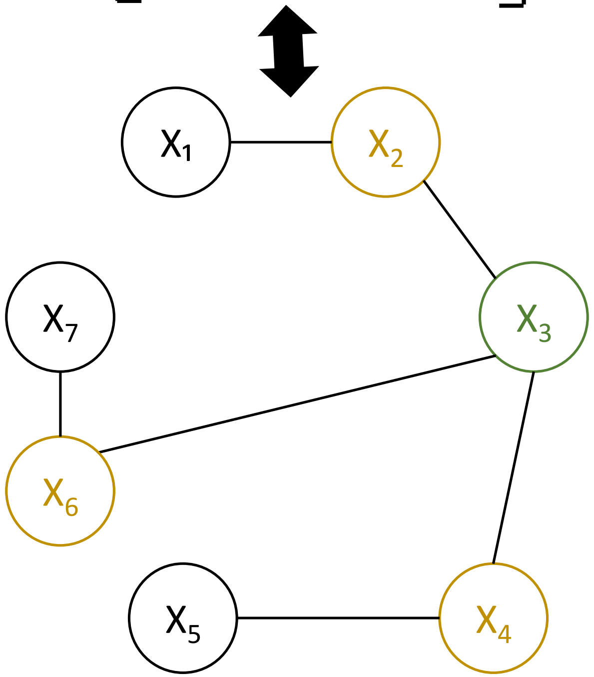


Gaussian Graphical Models

- $X = [X_1, X_2 \ldots, X_n]^T$- Jointly Gaussian random variables with covariance Σ^*.

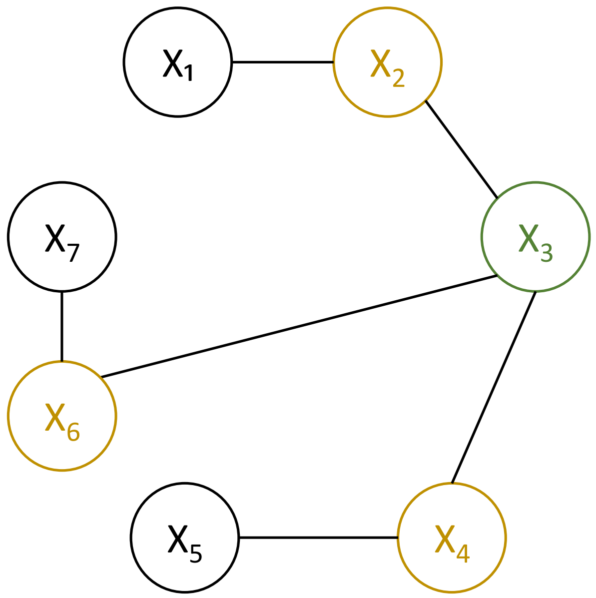
- Support of inverse covariance Ω^* gives the graphical model structure.

$$\Omega^* = \begin{bmatrix} X & X & 0 & 0 & 0 & 0 & 0 \\ X & X & X & 0 & 0 & 0 & 0 \\ 0 & X & X & X & 0 & X & 0 \\ 0 & 0 & X & X & X & 0 & 0 \\ 0 & 0 & 0 & X & X & 0 & 0 \\ 0 & 0 & X & 0 & 0 & X & X \\ 0 & 0 & 0 & 0 & 0 & X & X \end{bmatrix}$$



Effect Of Noise

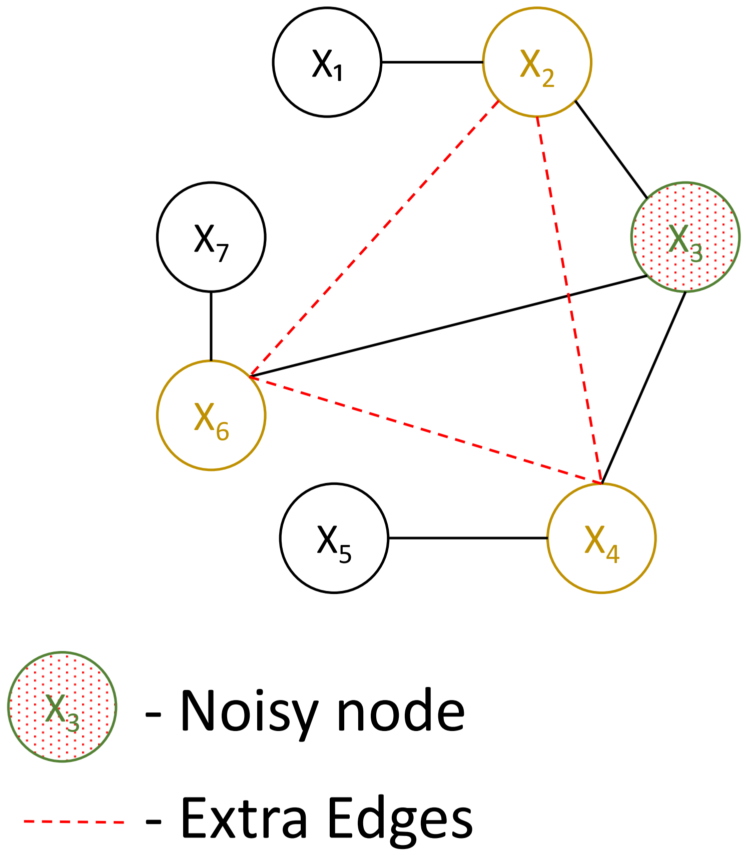
- Additive Gaussian noise in the random variables breaks down the conditional independence.
- Intuitively – Noisy samples do not convey the whole information.



Effect Of Noise

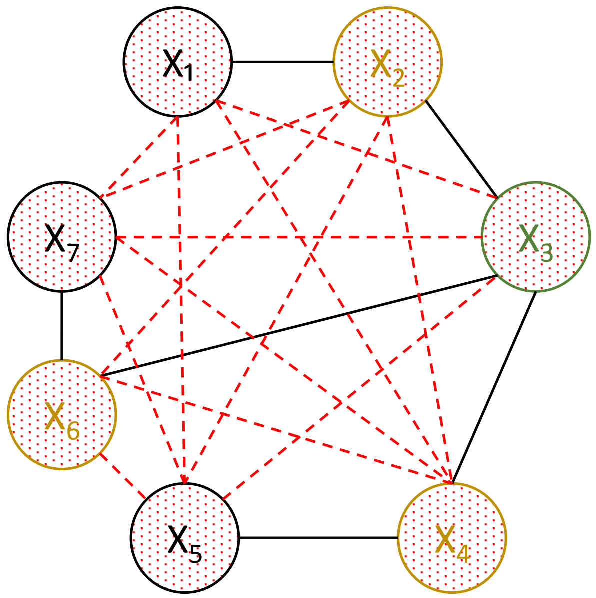
- Additive Gaussian noise in the random variables breaks down the conditional independence.

- Intuitively – Noisy samples do not convey the whole information.



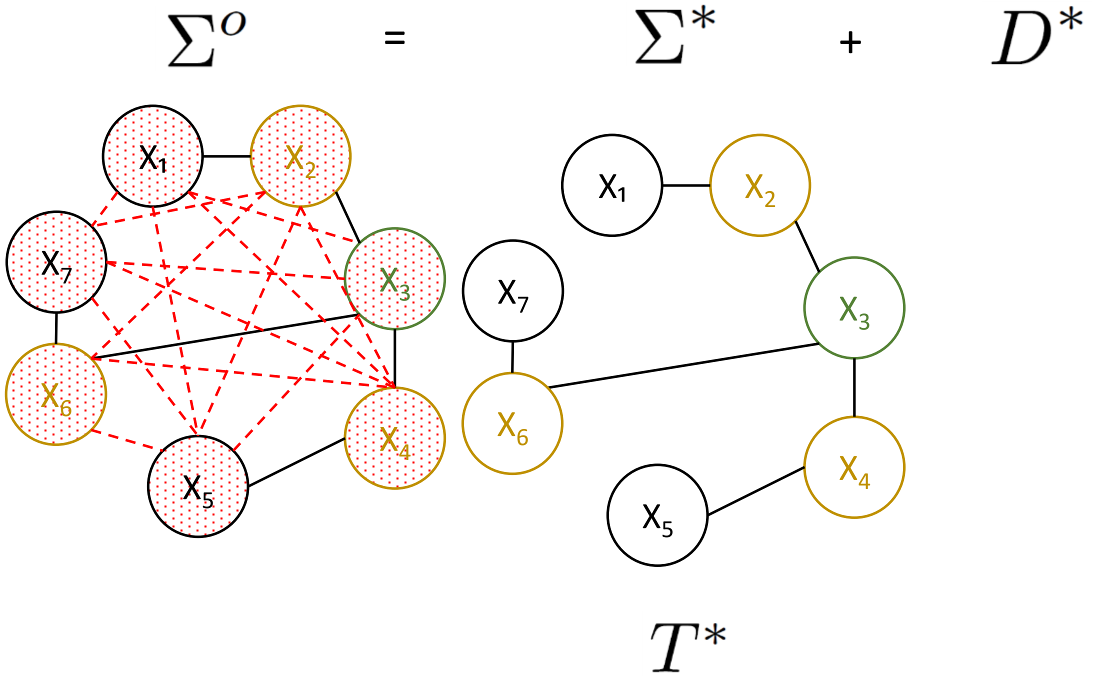
Effect Of Noise

- Additive Gaussian noise in the random variables breaks down the conditional independence.
- Example: if X – Y – Z is a Markov chain, then X and Z are no longer independent when conditioned on a *noisy* version of Y.



Problem Statement

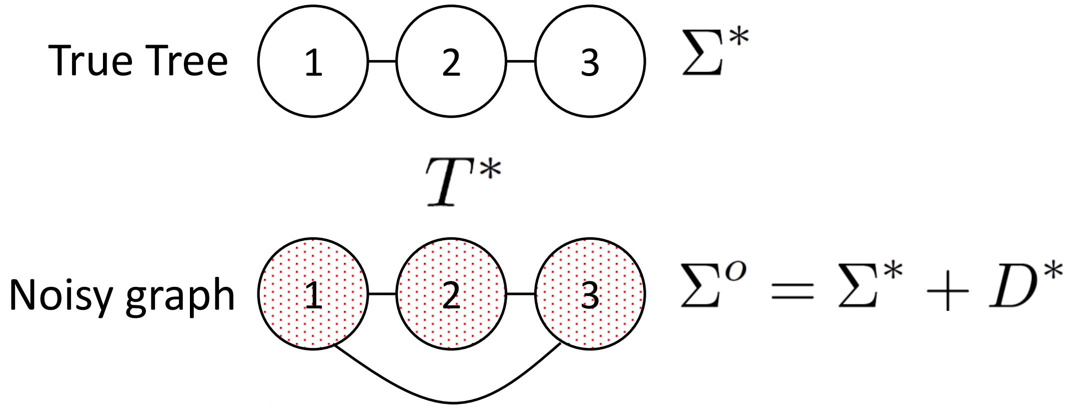
- Suppose the graphical model have a tree structure T^*.
- We observe $\Sigma^o = \Sigma^* + D^*$ where D^* is an unknown positive diagonal noise matrix.
- Goal: recover T^* given Σ^o.



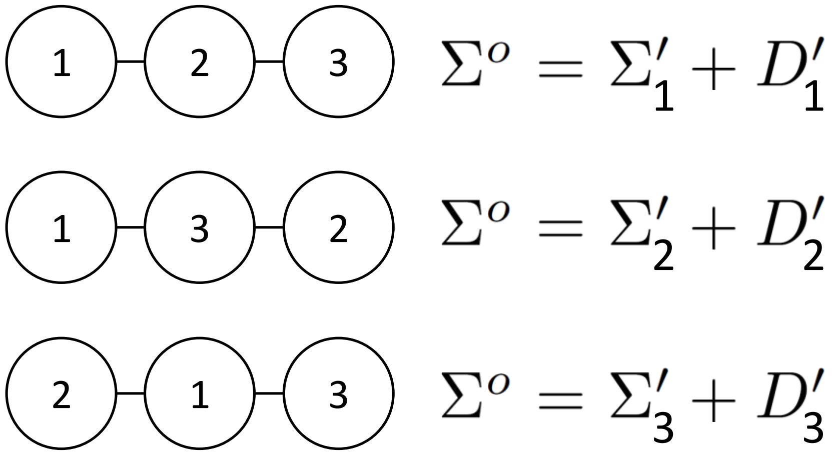
Bad News! **Unidentifiability**

- Even for *arbitrarily small* noise the problem is unidentifiable!

- There are covariance matrices that differ only on diagonal entries, but their inverses have different sparsity pattern.



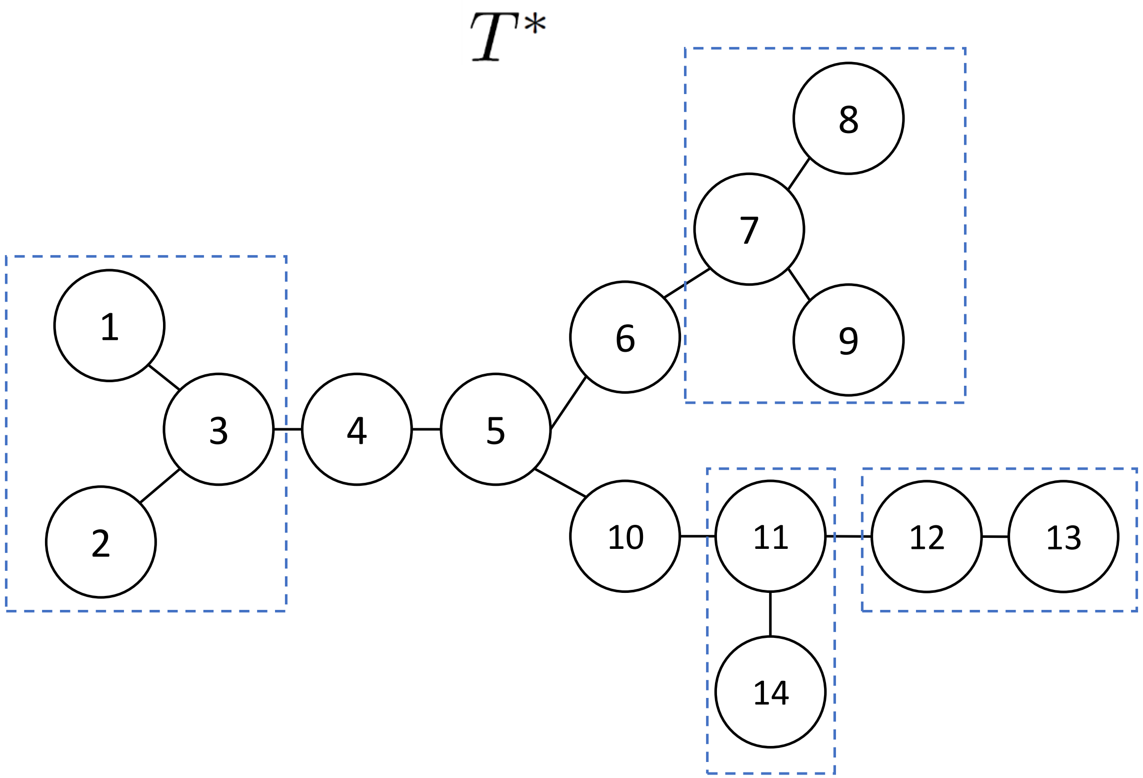
(a) Underlying tree T^* and the noisy graphical model.



(b) Graphical models for 3 different decompositions of Σ^o.

Good News! **Limited** Unidentifiability

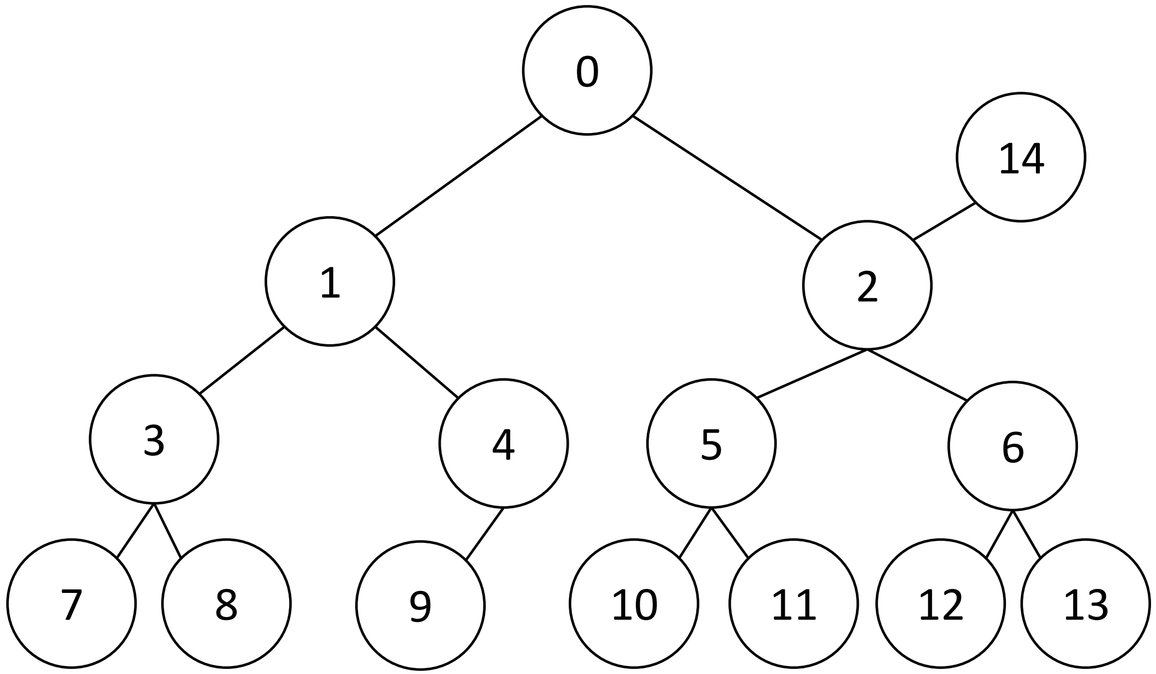
- The ambiguity in tree structure is highly limited.
- The only ambiguity is between a leaf node and its immediate neighbor.



Trees formed by permuting nodes within the dotted regions form an equivalence class $\mathcal{T}_{T^*}$.

Proof - Key Idea

- Off-diagonal covariance entries have information about the tree structure.
- They can be used to categorize any set of 4 nodes as a **star** or **non-star**.
- **Non-star** – Exactly 2 nodes lie in one subtree.

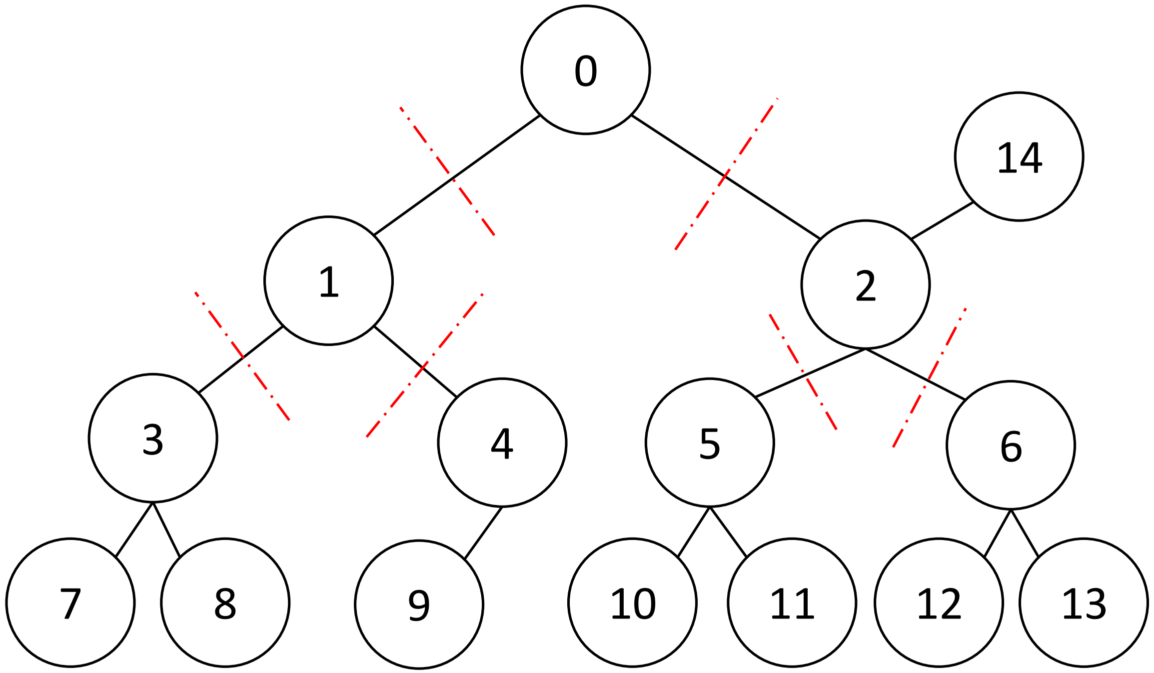


Node Set	Classification
{0, 1, 3, 7}	Non-Star
{7, 8, 9, 10}	Non-Star
{14, 2, 10, 11}	Non-Star
{7, 9, 1, 6}	Star
{1, 14, 10, 6}	Star

Proof – Key Idea

- Categorization of any set of 4 nodes as star/non-star defines all possible partitions in 2 subtrees with minimum 2 nodes.

- Thus the off diagonal elements define the tree upto the equivalence class $\mathcal{T}_{T^*}$.

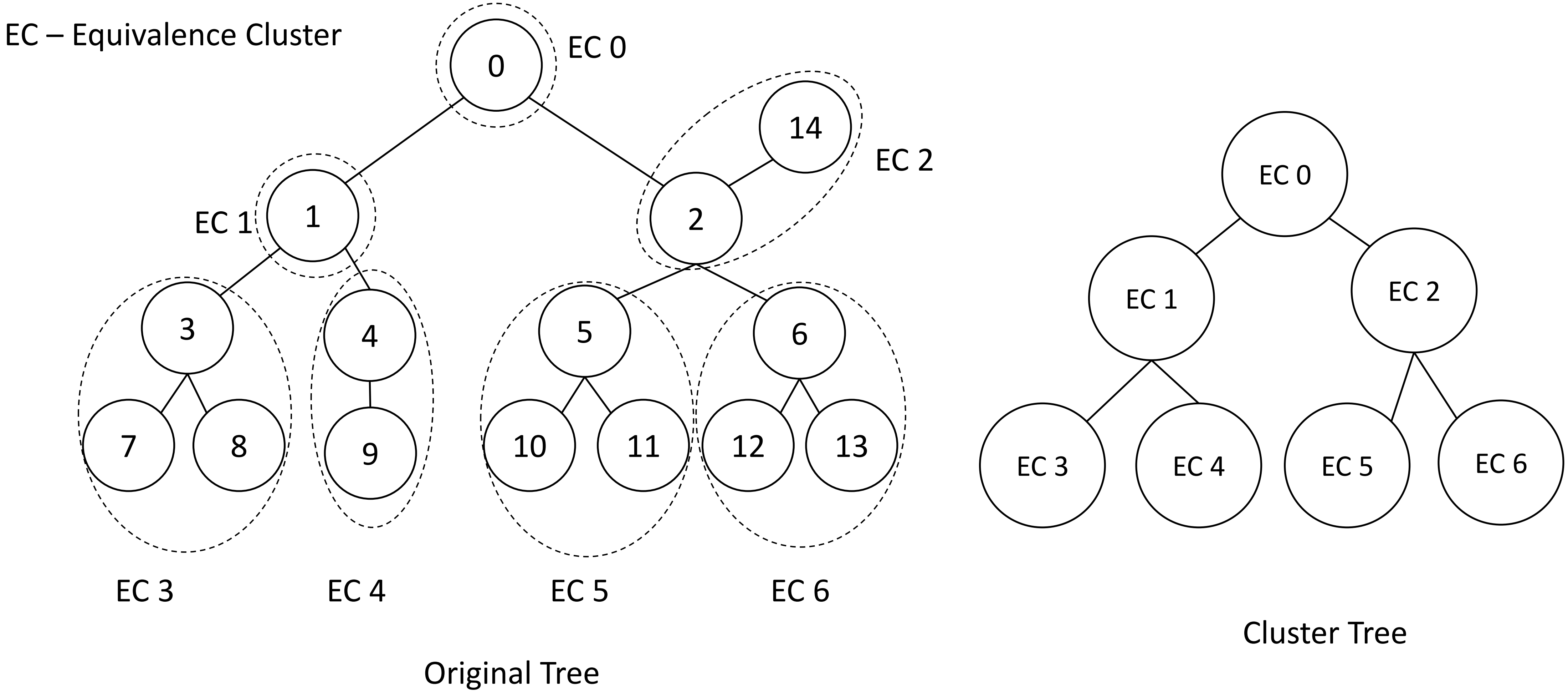


Identifiability with Side Information

- **Diagonal Majorization Condition –** If the precision matrix is known to have diagonal entries greater than the absolute off diagonal entries.

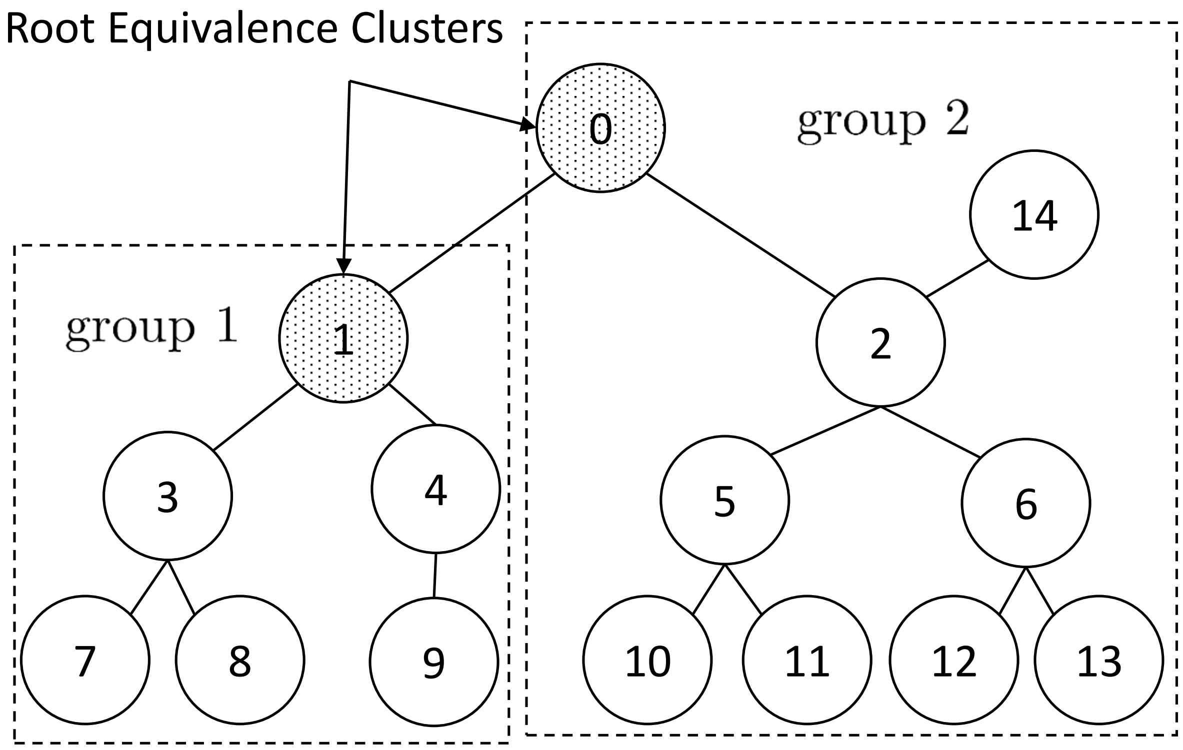
- **Minimum Eigenvalue Condition –** If a lower bound on the minimum eigenvalue of the covariance is known and the noise variance is smaller than this lower bound.

Cluster Tree

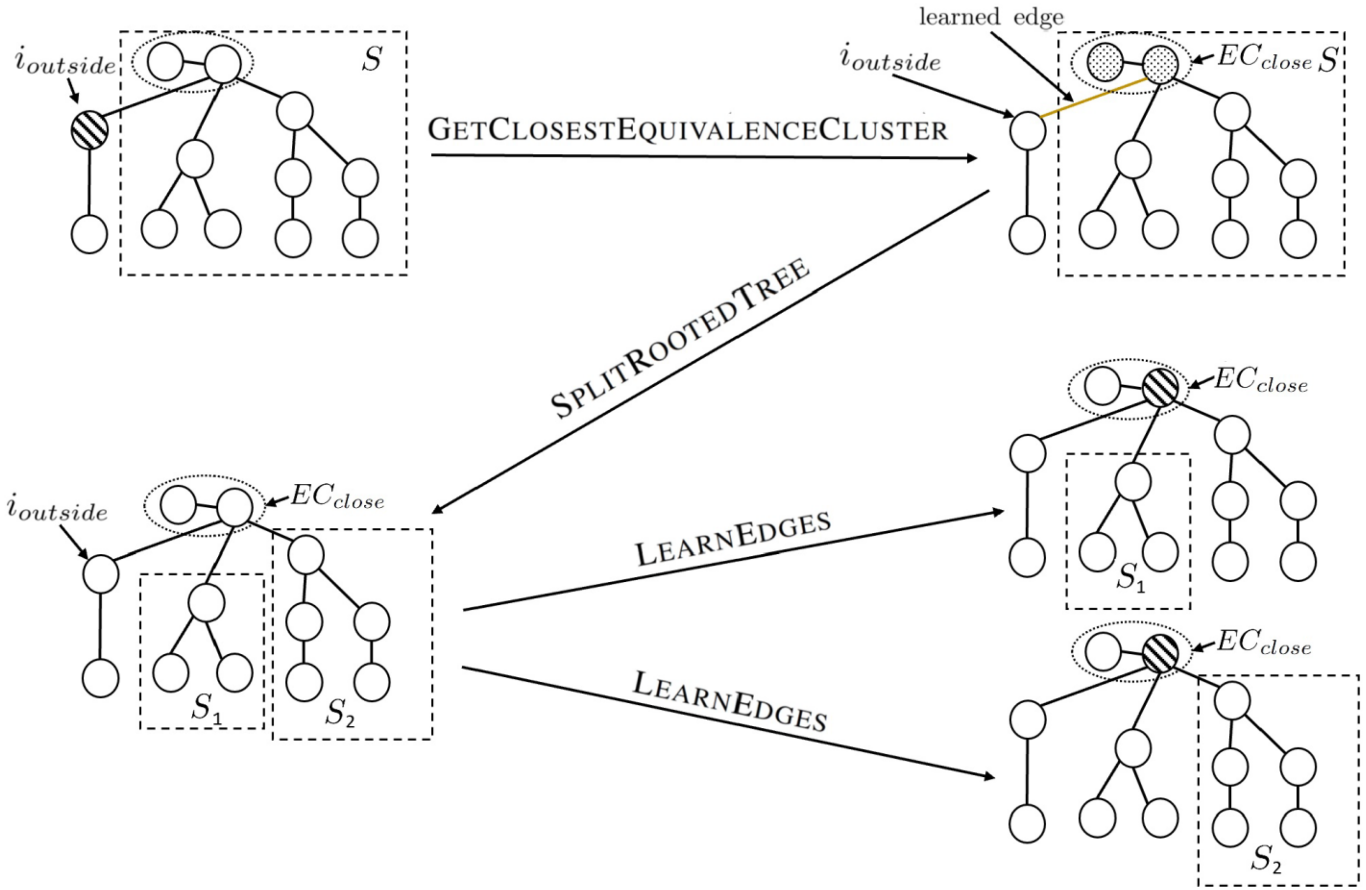


Algorithm - Initialization

- Split the tree in 2 subtrees.
- Find the root equivalence equivalence cluster of each subtree.



Algorithm – Recursion step



Conclusion

- Unidentifiability of learning tree structured Gaussian Graphical Models in presence of noise.
- Identifiability conditions with side information.
- Algorithm to find the tree structure.