A Recursive Decomposition Framework for Causal Structure Learning in the Presence of Latent Variables—DiCoLa

DiCoLa: Divide-and-Conquer causal discovery in the presence of Latent variables

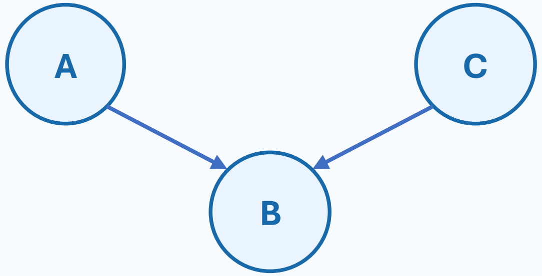
Speaker: Zheng Li

Authors: Zheng Li [1,2], Feng Xie [2], Shenglan Nie [2], Xichen Guo [2], Ruxin Wang [1], Hao Zhang [1]

1. Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences
2. Beijing Technology and Business University

Motivation: More Efficient Causal Discovery with Latent Variables

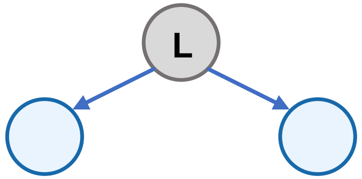
Why is Causal Discovery Important?



To understand how variables influence one another.

Challenge 1: Latent Variables

Latent confounders induce complex dependencies among observed variables.

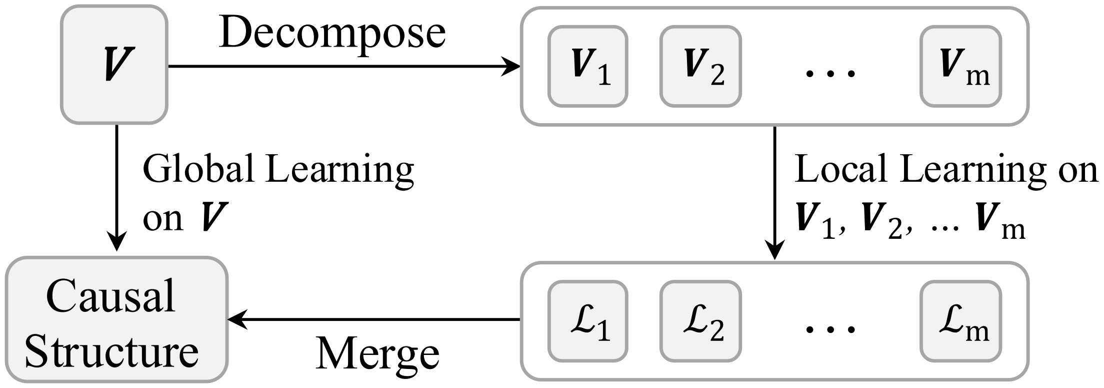


Challenge 2: Computational Complexity

Constraint-based methods may require exponentially many conditional independence tests as the dimensionality grows.

Goal: latent variables + more efficient + provable correctness

Divide-and-Conquer Methods & Their Limitations



A Divide-and-Conquer Strategy

✓ fewer variables per subproblem
✓ fewer CI tests

Main limitation

Causal Sufficiency Assumption

Most theoretically grounded frameworks rely on the critical assumption of **no latent variables**.

Severely Limited Applicability

Real-world data often contain unmeasured confounders.

With latent variables, existing divide-and-conquer methods lack correctness guarantees.

The Key Question

Can divide-and-conquer causal structure learning be generalized to settings with latent variables?

Why is this hard?

Incorrect decomposition may lose critical separating information required to recover the global structure.

Our Answer: YES

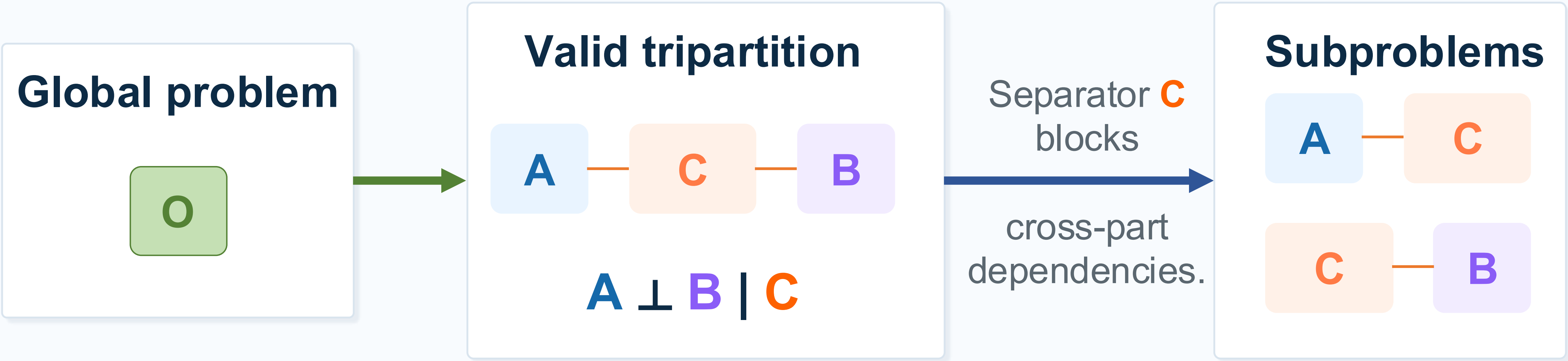
We propose **DiCoLa**, sound + complete if the base learner is.

The Fundamental Question

Under **what conditions** can global causal structure learning be decomposed into smaller subproblems while preserving separating information?

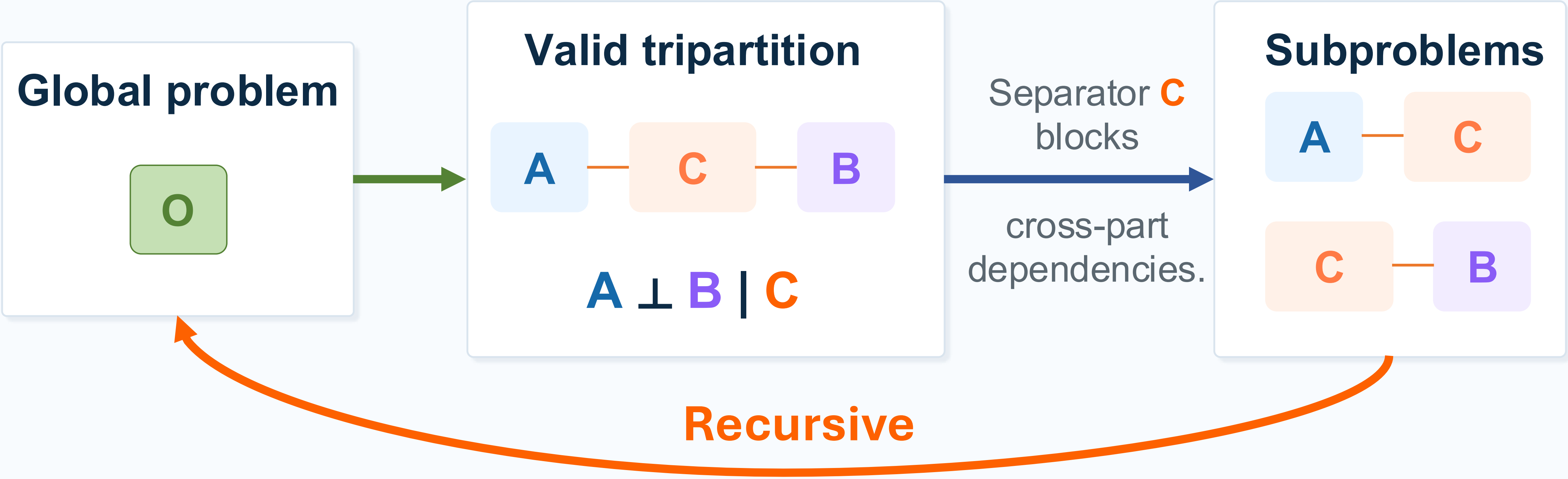
The Fundamental Question

Under **what conditions** can global causal structure learning be decomposed into smaller subproblems while preserving separating information?



The Fundamental Question

Under **what conditions** can global causal structure learning be decomposed into smaller subproblems while preserving separating information?



Theoretical Foundation: M-Separation in MAGs

Valid tripartition

A — C — B

$A \perp B \mid C$

Theorem 1: Localization

For any $X \in A$ and $Y \in A \cup C$, X and Y are m-separated by a subset of $A \cup B \cup C$ **if and only if** they are m-separated by a subset of $A \cup C$

Theorem 2: Coverage

For any $X,Y \in C$, X and Y are m-separated by a subset of $A \cup B \cup C$ **if and only if** they are m-separated by a subset of $A \cup C$ or $B \cup C$

Theorem 3: Reconstructibility

For any X,Y, if they are m-connected in both subproblems: $A \cup C$, and $B \cup C$, then they must be m-connected in the full problem $A \cup B \cup C$.

How to Find Valid Decompositions?

Valid tripartition

A — C — B

$$A \perp B \mid C$$

How to Find Valid Decompositions?

UIG: undirected independence graph

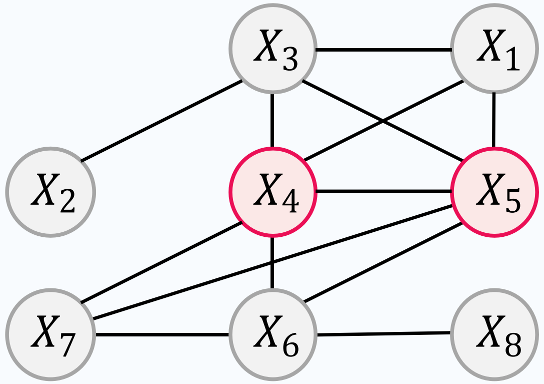
Valid tripartition

A — C — B

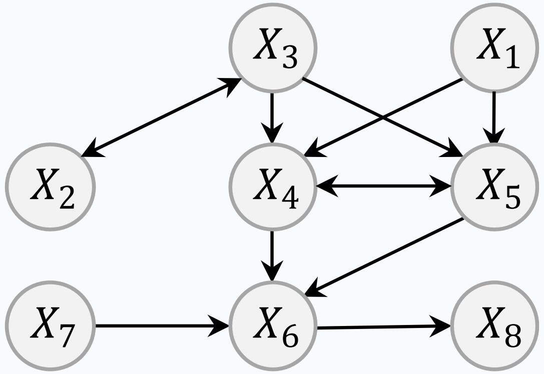
$A \perp B \mid C$

Key definition:

separator in UIG $\Rightarrow$ valid m-separation in the underlying MAG



UIG



MAG

How to Find Valid Decompositions?

UIG: undirected independence graph

Valid tripartition

A — C — B

$A \perp B \mid C$

Key definition:

separator in UIG $\Rightarrow$ valid m-separation in the underlying MAG

How to Learn UIGs?

How to Find Valid Decompositions?

UIG: undirected independence graph

Valid tripartition

A — C — B

$A \perp B \mid C$

Key definition:

separator in UIG $\Rightarrow$ valid m-separation in the underlying MAG

Key property:

Two node adjacent in the UIG if and only if they are present in each other's Markov blankets

How to Find Valid Decompositions?

How to Learn UIGs?

Markov Blanket Learning

Valid tripartition

$A \perp B \mid C$

How to Find Valid Decompositions?

Original Problem

O

Learn MBs

Markov blankets give local adjacency evidence.

Build UIG

Vertex separation in UIG implies m-separation in MAG.

Find separator C

Removing C disconnects the UIG.

Split A / B

The Core Idea of DiCoLa

Divide

Find valid separators and split recursively

Learn

Run a base learner only on leaf subproblems

Merge

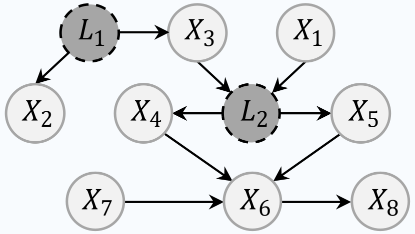
Combine local skeletons bottom-up

Orient

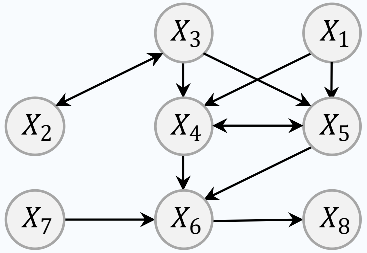
Apply PAG orientation rules globally

Framework, not a replacement: FCI / RFCI / FCI+ / L-MARVEL / ICD

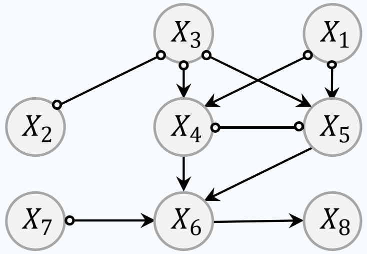
DiCoLa Illustration



(a) DAG



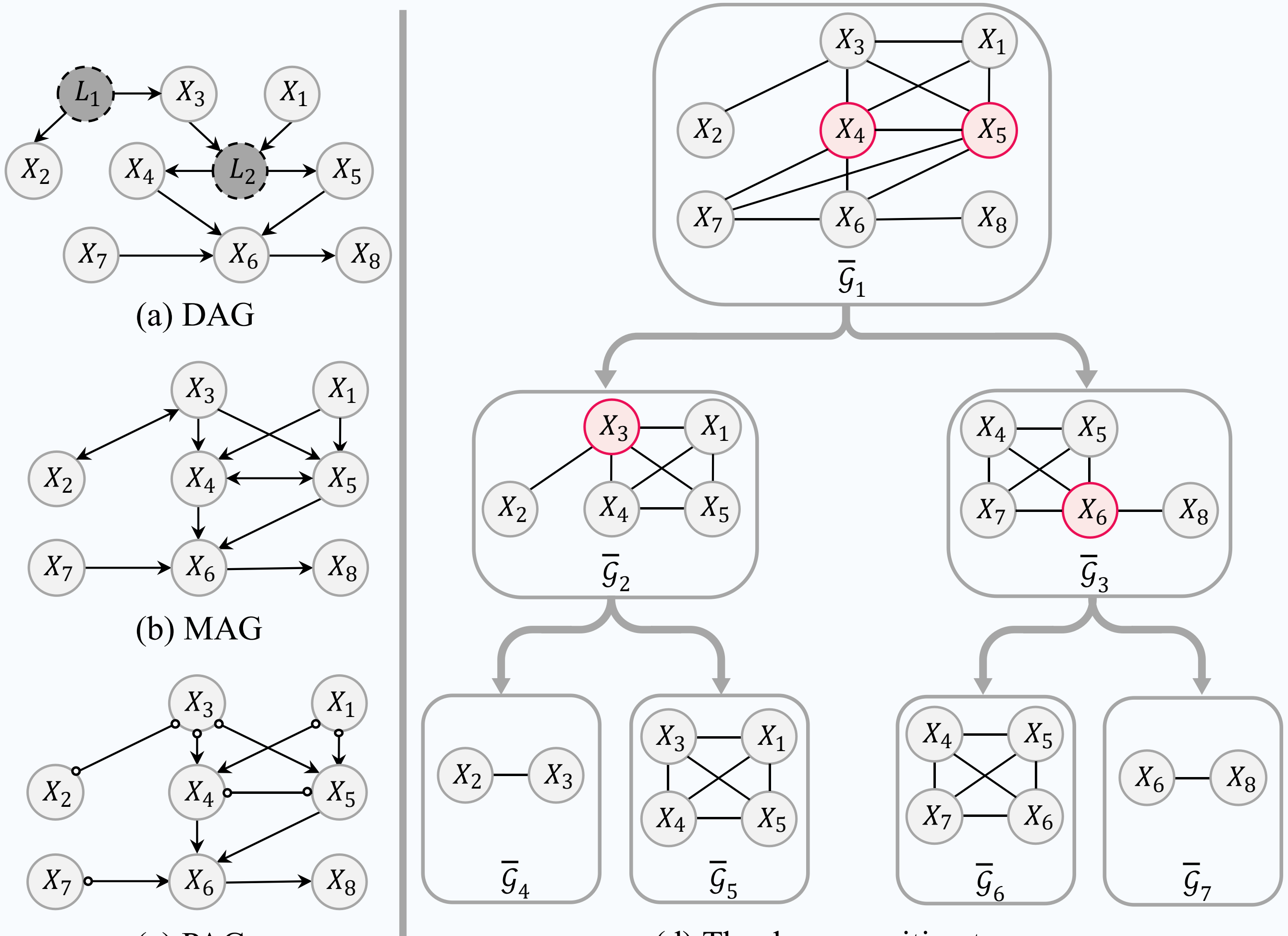
(b) MAG



(c) PAG

01 **Start from observed variables**

DiCoLa Illustration



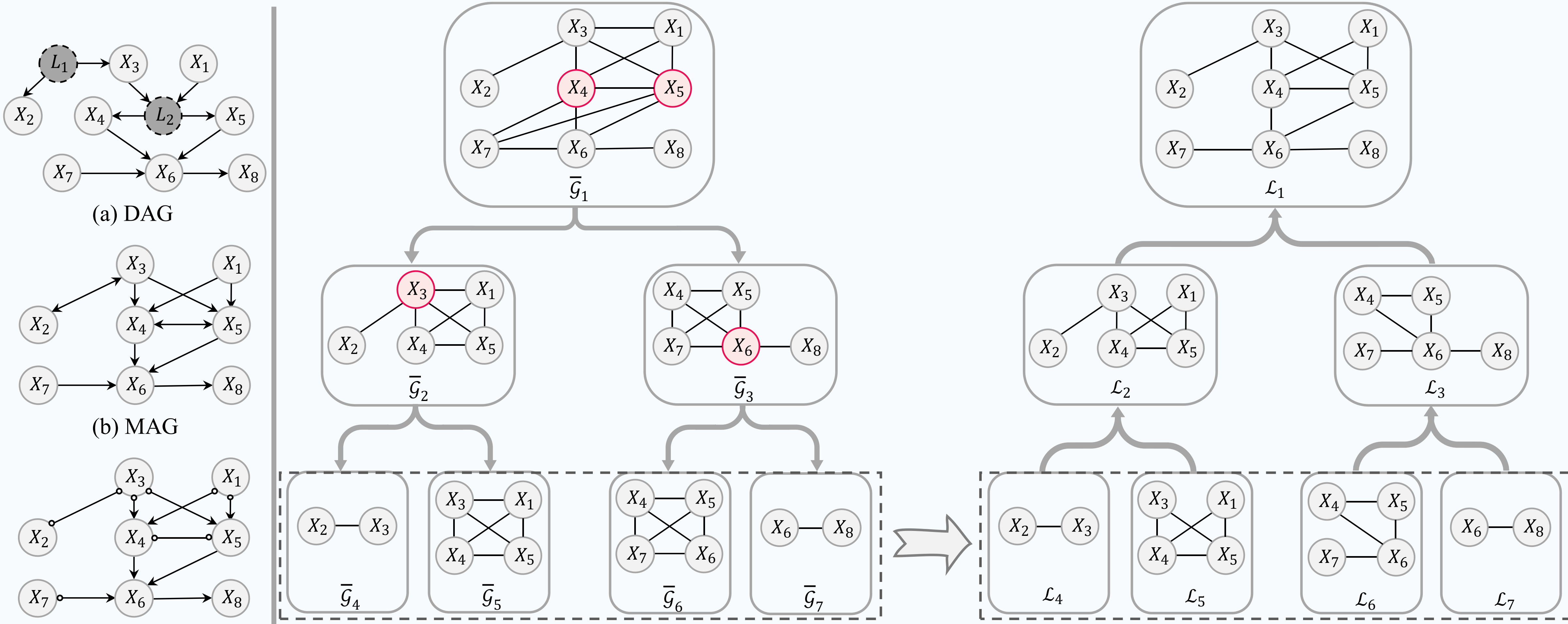
(c) PAG

(d) The decomposition tree

01 **Start from observed variables**

02 **Split by valid separators**

DiCoLa Illustration



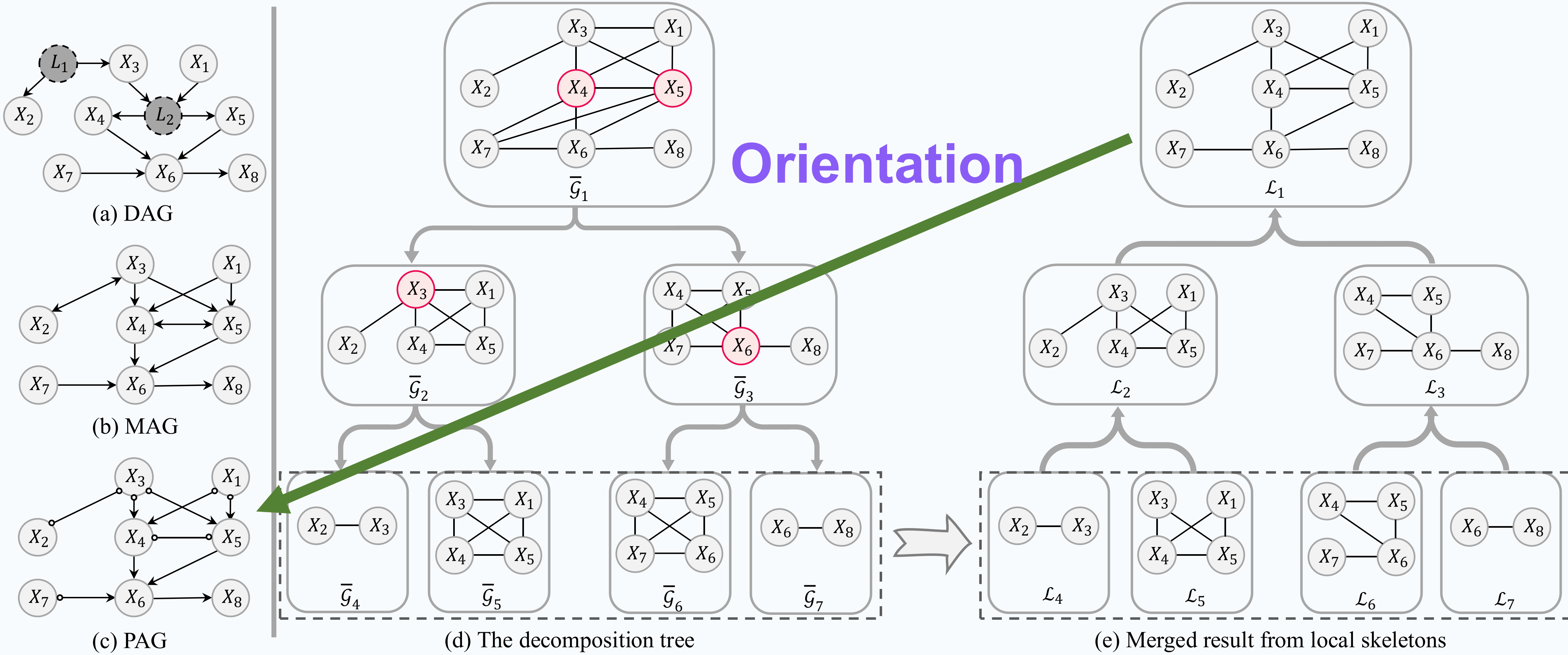
(a) DAG (b) MAG (c) PAG (d) The decomposition tree (e) Merged result from local skeletons

01 Start from observed variables

02 Split by valid separators

03 Learn locally and merge globally

DiCoLa Illustration



01 **Start from observed variables**

02 **Split by valid separators**

03 **Learn locally and merge globally**

Theoretical Guarantee & Complexity

Soundness & completeness

If the base learner is sound and complete, DiCoLa preserves these guarantees.

No correctness loss from recursive decomposition

Computational gain

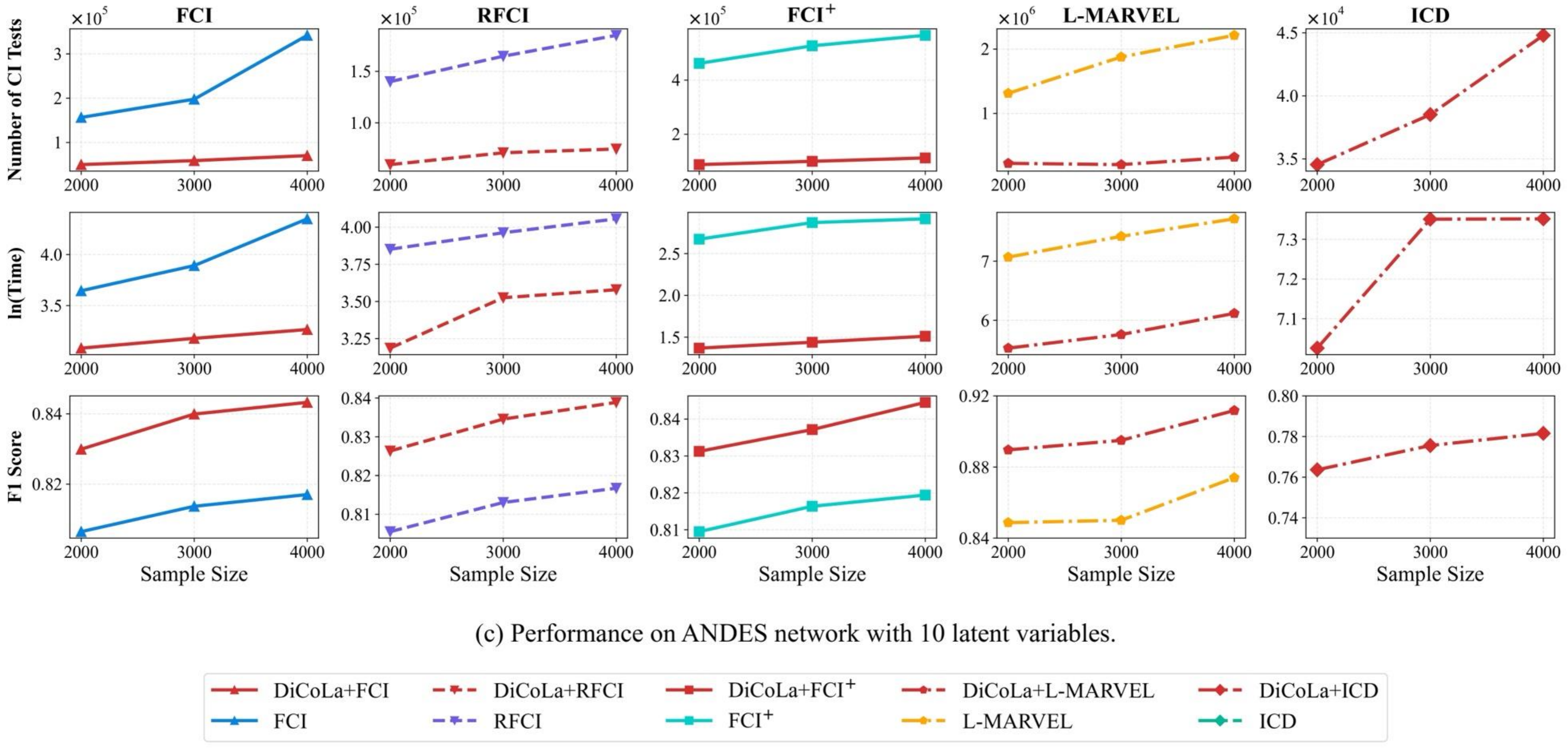
Global learner	$O(n^2 \cdot 2^n)$
DiCoLa + Base	$O(m \cdot {k_{max}}^2 \cdot 2^{k_{max}})$

k_{max} = largest leaf subproblem size

m = number of leaf problems

When $k_{max} \ll$ n, the exponential bottleneck shrinks.

Simulations



(c) Performance on ANDES network with 10 latent variables.

Takeaway: DiCoLa makes causal discovery with latent variables much cheaper while keeping accuracy competitive.

Fewer CI tests & lower runtime

The DiCoLa+Base curves are consistently lower on efficiency metrics.

F1 is maintained or improved

The computational saving does not come at the cost of structure quality.

Conclusions

Theory

First divide-and-conquer foundation for latent-variable causal structure learning.

Algorithm

Recursive decomposition framework that wraps standard base learners.

Empirics

Fewer CI tests and lower runtime, while maintaining graph quality.

Divide-and-conquer causal discovery is possible beyond settings without latent variables

Dense Structures

The efficiency gain of DiCoLa is most pronounced on sparse or well-decomposable graphs.

Thank You

GitHub: zhengli0060/DiCoLa-ICML2026

Email: zhengli0060@gmail.com