Structural Casual Models

Presented by Krish Doshi

What are Structural Casual Models?

1. A way of describing the relevant features of the world and how they interact with each other. Specifically, describing how nature assigns values to variables of interest.
2. Consists of two sets of variables U and V, and a set of functions f that assigns each variable in V (a value based on the values of other variablers in the model)

Types of Variables

1. Exogenous variables (in set U) are external to the model (we choose for whatever reason not to explain how they are caused) and are roots of the graphs

2. Endogenous variables (in set V) is a descendant of at least one exogenous variable

Causes

$$X \longrightarrow Y$$

1. X is a direct cause of variable Y if X appears in the function that assigns Y's value
2. X is a cause of Y if it is a direct cause of Y or of any ancestor of Y

$$X \longrightarrow Z \longrightarrow Y$$

*intransitive cases: if X->Z->Y and changes in X doesn't effect the value that Z takes

Product Decomposition

Any model with an acyclic graph, the joint distribution is given by the product of the conditional distributions

$$P(x_1, x_2, \ldots, x_n) = \prod_i P(x_i \mid pa_i)$$

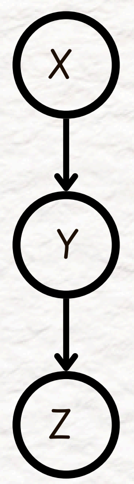
$$X \longrightarrow Y \longrightarrow Z$$

$$P(X=x, Y=y, Z=z) = P(X=x)P(Y=y|X=x)P(Z=z|Y=y)$$

Common Casual Graphs

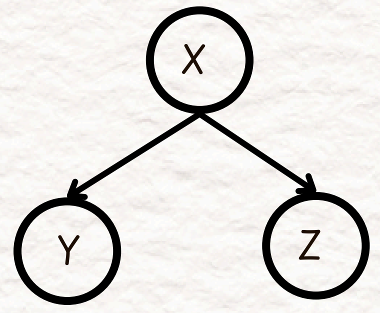
1. Chains

Three nodes and two edges with one edge directed into and one edge directed out of the middle variable. X and Z are independent by conditioning on Y if the error terms are independent.



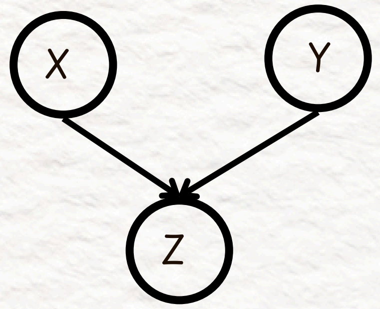
2. Forks

Three nodes with two arrows emanating from the middle variable (middle variable is the comon cause of the other two variables). Y and Z are independent by conditioning on X (conditioning on the common cause).



3. Colliders

When one node receives edges from two other nodes. X and Y are inconditionally independent, but are conditionally dependent conditioned on Z



Intervention

1. When we intervene on a variable in a model, we fix its value (denoted by “do”). Intervening on a variable essentially removes all the incoming arrows to that variable (need a randomized experiment to intervene on a variable).
2. Conditioning is where we narrow the focus to the subset of cases in which the variable takes the value we are interested in

Backdoor Criterion

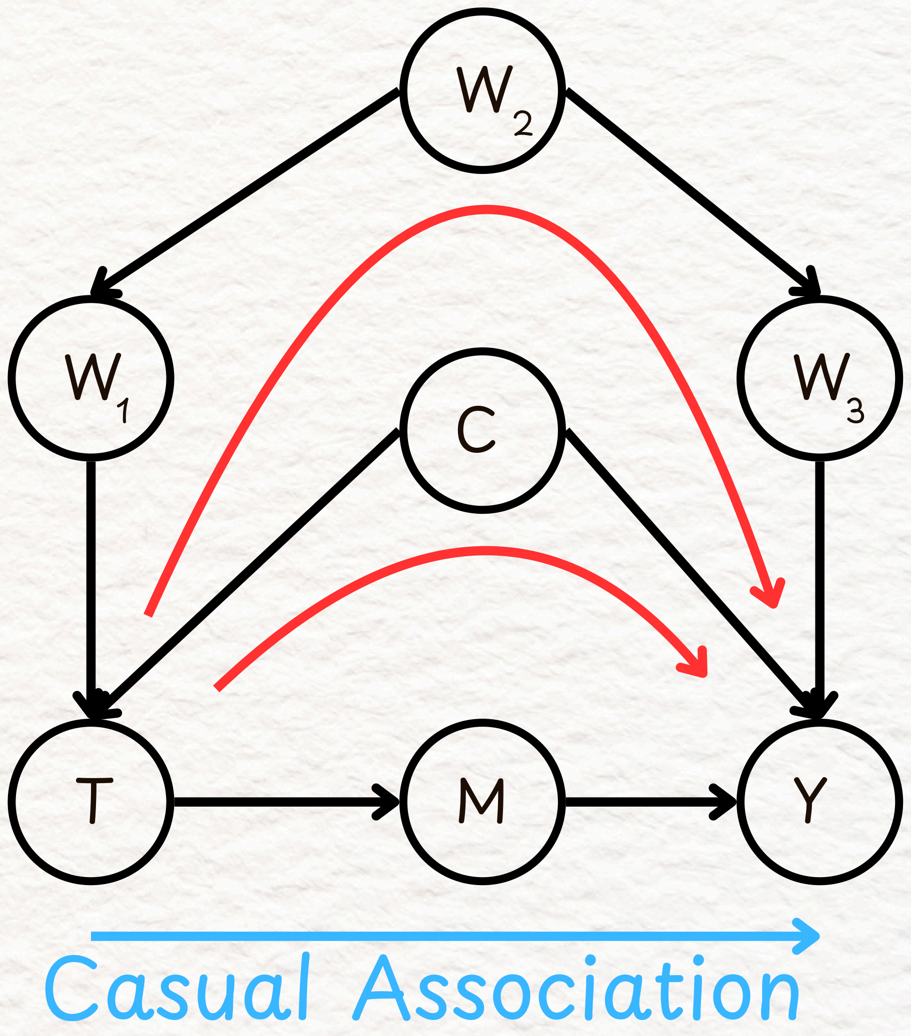
Want to be able to mimic an interventional distribution by just using observational data (blocking the non-casual association).

A set of variables W satisfies the backdoor criterion relative to T and Y if the following are true:

1. W blocks all backdoor paths from T to Y

2. W does not contain any descendants of T

$$P(y \mid do(t)) = P(y \mid t, W_2, C)$$

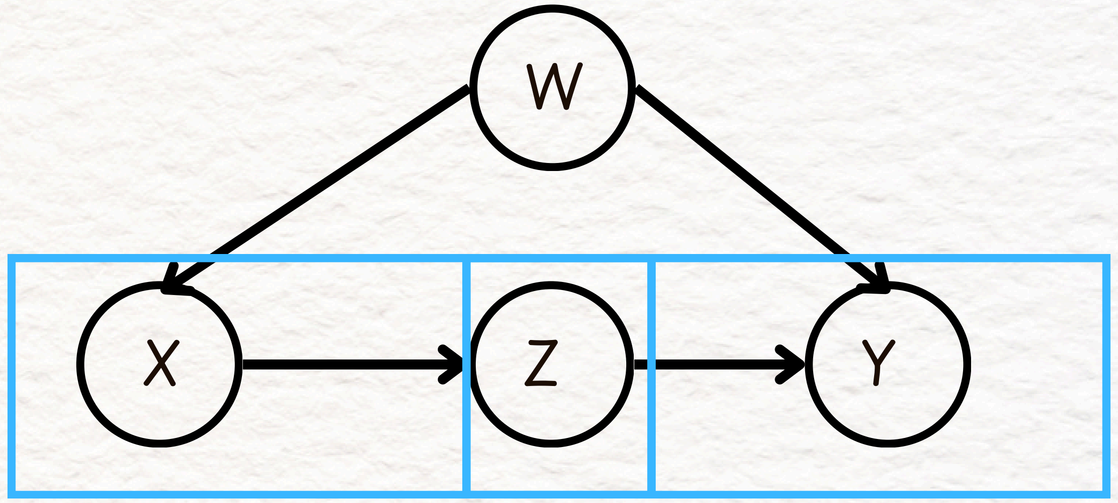


Frontdoor Criterion

Helps with finding the causal association when there is an unobserved confounding variable

A set of variables Z is said to satisfy the front-door criterion relative to an ordered pair of variables (X,Y) if

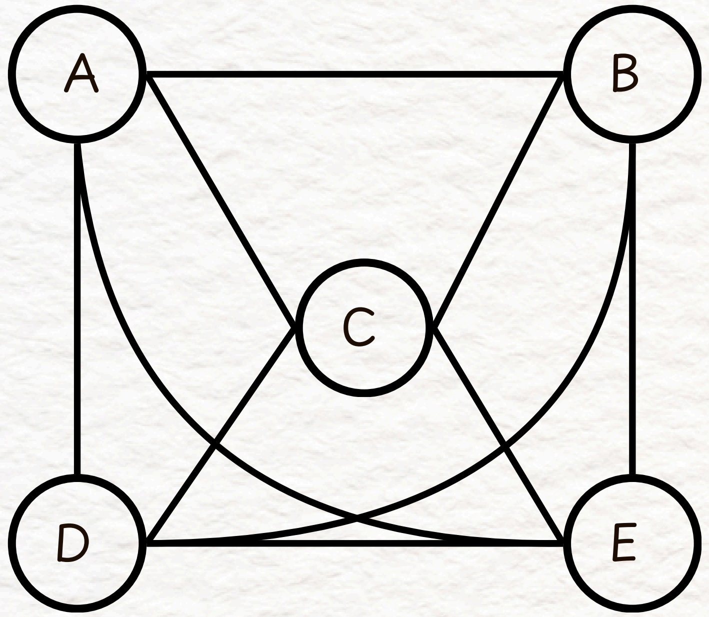
1. Z intercepts all directed paths from X to Y
2. There is no backdoor path from X to Z
3. All backdoor paths from Z to Y are blocked by X.



$$P(y \mid do(x)) = \sum_{w} P(z|x) \sum_{x'} P(y|x',z)P(x')$$

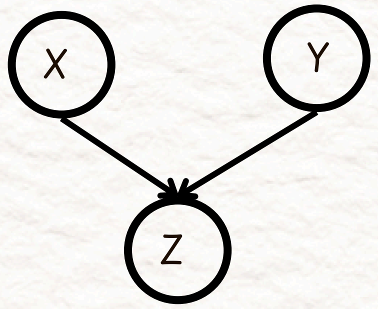
PC Algorithm

Start with Complete Graph



Identify the Skeleton

Start with complete undirected graph and remove edges X - Y where $X \perp\!\!\!\perp Y \mid Z$ for some (potentially empty) conditioning set Z, starting with the empty conditioning set and increasing the size

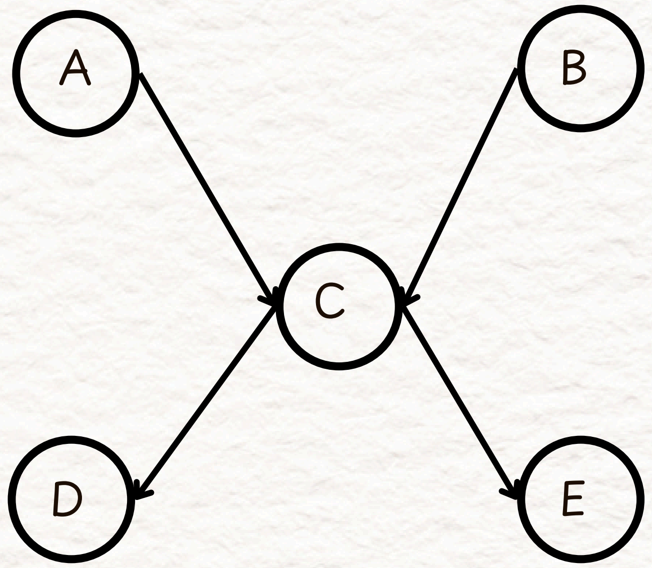


Identify Colliders and Orient Them

Now for any paths X - Z - Y in our working graph where the following are true:

1. We discovered that there is no edge between X and Y in our previous step
2. Z was not in the conditioning set that makes X and Y conditionally independent.

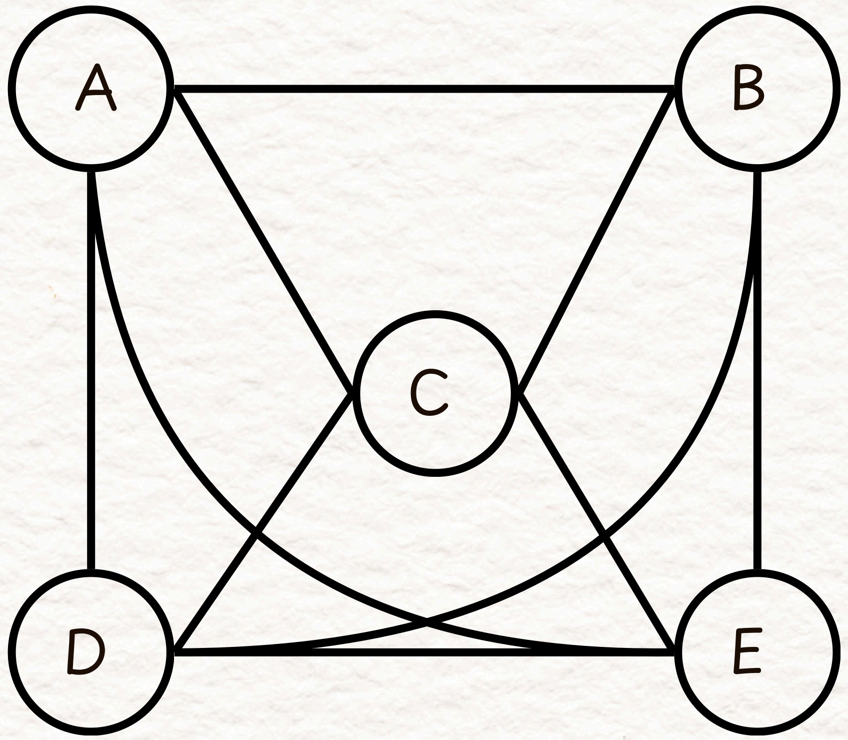
Then we know that X - Z - Y forms an immorality



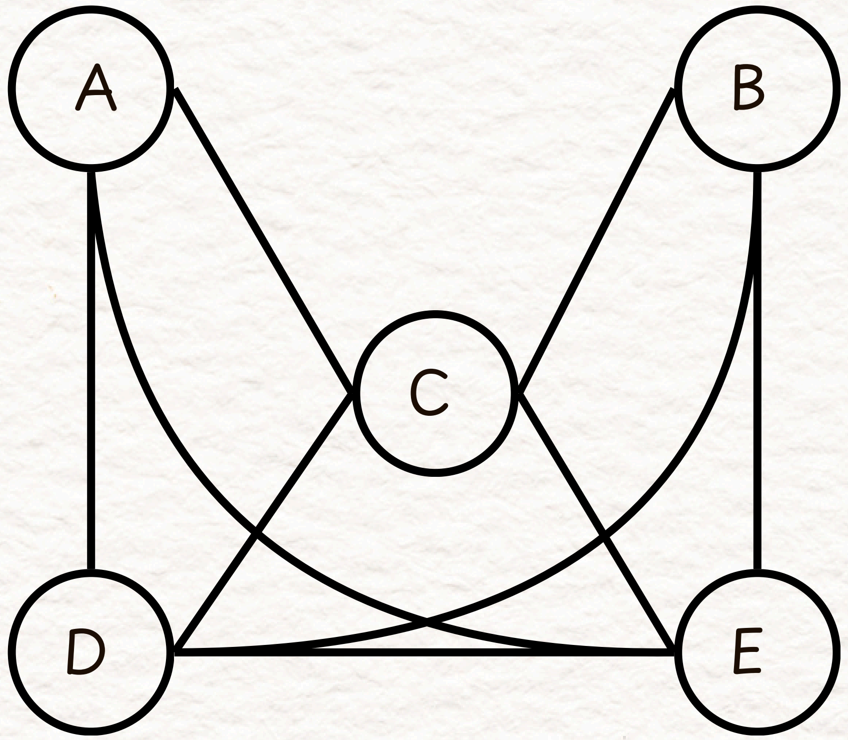
Orient Qualifying edges that are incident on colliders

Use fact that we discovered all colliders to orient more edges. Any edge Z-Y part of a partially directed path of the form X->Z-Y, where there is no edge connecting X and Y can be oriented as Z->Y

Start with Complete Graph

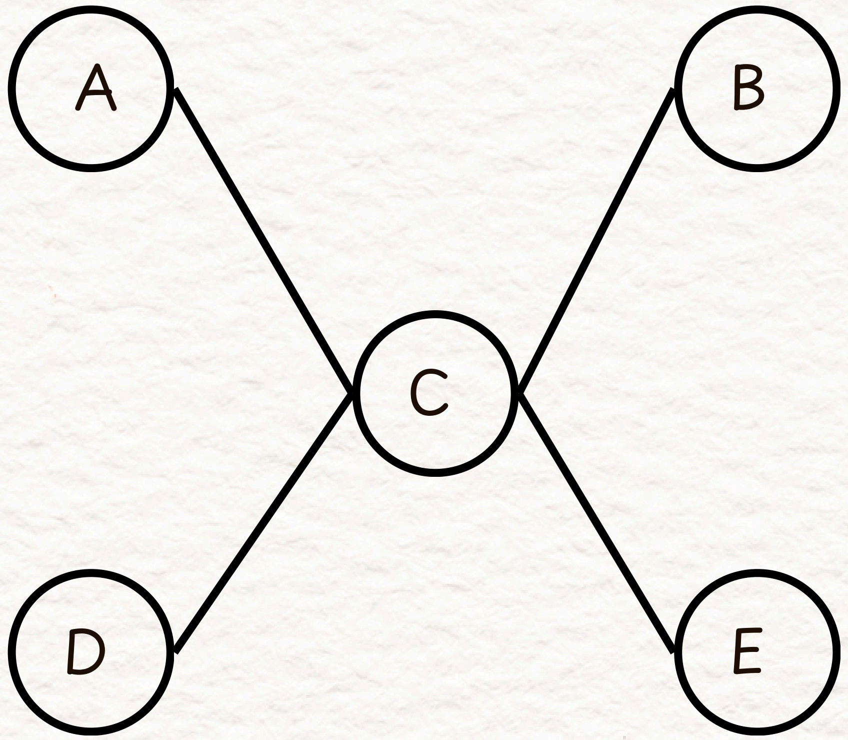


Identifying the Skeleton



$A \perp\!\!\!\perp B \mid \{\}$

Identifying the Skeleton

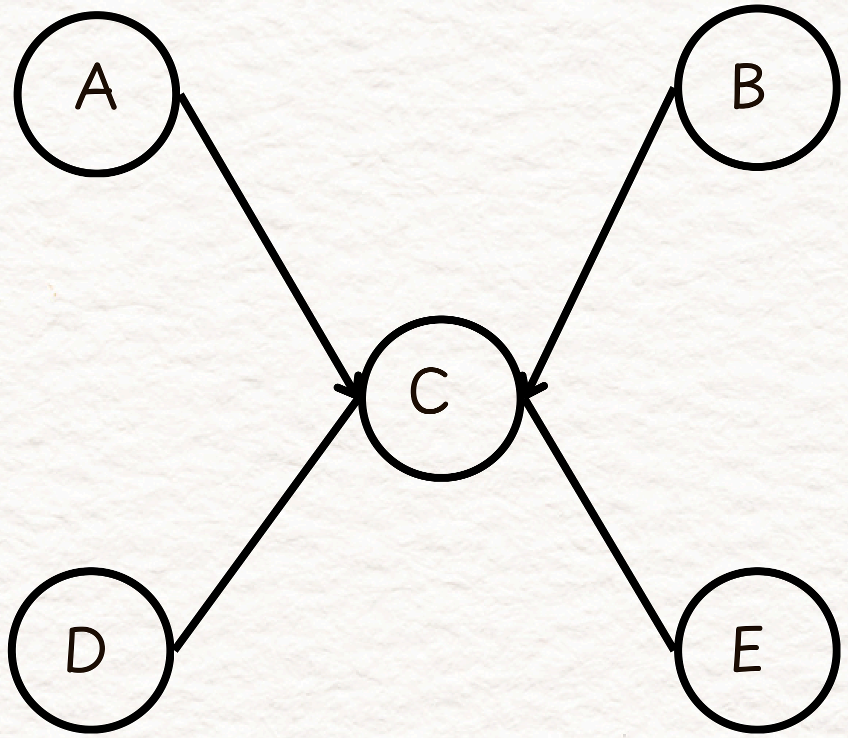


$A \perp\!\!\!\perp B \mid \{\}$

$\forall$ other pairs (X,Y), $X \perp\!\!\!\perp Y \mid \{C\}$

*Keeping in mind that this doesn't work for A and B conditioned on C since it is a collider

Identifying the Colliders

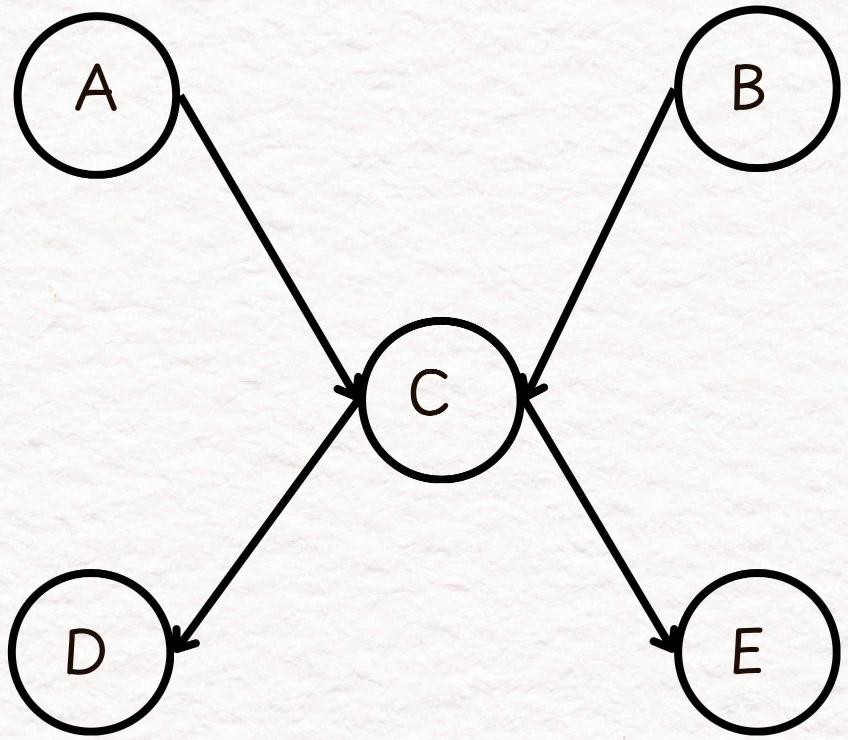


There is no edge between A and B

$A \perp\!\!\!\perp B \mid \{\}$

C is not in the conditioning set that makes A and B conditionally independent

Orienting Other Edges



Cannot orient the other edges towards C or else we would have formed more colliders, which would have been detected in previous step

Other Algorithms

PC algorithm is one of the most commonly used algorithms, but there are also other algorithms/variants which work better in some other situations

- PC-stable tackles the problem of order dependence (order in which variables are analyzed in skeleton phase providing different skeletons) problem
- FCI algorithm works with no assumed casual sufficiency (no unmeasured common causes of any pairs of variables aka no latent confounders)
- CCD algorithm works with no assumed acyclicity

SCMs in Real Life

Draft version January 15, 2025
Typeset using LaTeX **twocolumn** style in AASTeX631

Causal Discovery in Astrophysics: Unraveling Supermassive Black Hole and Galaxy Coevolution

Zehao Jin (金泽灏) [1,2,3], Mario Pasquato [3,4,5,6,7,*] Benjamin L. Davis [1,2,*] Tristan Deleu [4,8]
Yu Luo (罗煜) [9,10], Changhyun Cho [1,2] Pablo Lemos [3,4,5] Laurence Perreault-Levasseur [3,4,5,11]
Yoshua Bengio [4,8,12,13] Xi Kang (康熙) [14,10] Andrea Valerio Macciò [1,2,15] and Yashar Hezaveh [3,4,5,11]

[1] New York University Abu Dhabi, P.O. Box 129188, Abu Dhabi, United Arab Emirates
[2] Center for Astrophysics and Space Science (CASS), New York University Abu Dhabi, P.O. Box 129188, Abu Dhabi, UAE
[3] Montréal Institute for Astrophysical Data Analysis and Machine Learning (Ciela), Montréal, Canada
[4] Montréal Institute for Learning Algorithms (Mila), Quebec Artificial Intelligence Institute, 6666 Rue Saint-Urbain, Montréal, Canada
[5] Département de Physique, Université de Montréal, 1375 Avenue Thérèse-Lavoie-Roux, Montréal, Canada
[6] Dipartimento di Fisica e Astronomia, Università di Padova, Vicolo dell'Osservatorio 5, Padova, Italy
[7] Istituto di Astrofisica Spaziale e Fisica Cosmica (INAF IASF-MI), Via Alfonso Corti 12, I-20133, Milan, Italy
[8] Département d'Informatique et de Recherche Opérationnelle, Université de Montréal, 2920 Chemin de la Tour, Montréal, Canada
[9] Department of Physics, School of Physics and Electronics, Hunan Normal University, Changsha 410081, China
[10] Purple Mountain Observatory, 10 Yuan Hua Road, Nanjing 210034, China
[11] Center for Computational Astrophysics, Flatiron Institute, New York, NY, United States of America
[12] Canadian Institute for Advanced Research Artificial Intelligence Chair
[13] Canadian Institute for Advanced Research Senior Fellow
[14] Institute for Astronomy, Zhejiang University, Hangzhou 310027, China
[15] Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany

(Received September 16, 2024; Revised December 10, 2024; Accepted December 10, 2024)

Submitted to ApJ

Thank You