
***Causa prima*: cosmology meets causal discovery for the first time**

Mario Pasquato

Département de Physique, Université de Montréal,
1375 Avenue Thérèse-Lavoie-Roux, Montreal, Canada
Mila - Quebec Artificial Intelligence Institute,
6666 Rue Saint-Urbain, Montréal, Canada
Ciela - Montreal Institute for Astrophysical Data Analysis
and Machine Learning, Montréal, Canada
Dipartimento di Fisica e Astronomia, Università di Padova,
Vicolo dell'Osservatorio 5, Padova, Italy
INFN-Padova, Via Marzolo 8, Padova, Italy
mario.pasquato@umontreal.ca

Zehao Jin

Center for Astrophysics and Space Science (CASS),
New York University Abu Dhabi, PO Box 129188, Abu Dhabi, UAE

Pablo Lemos

Département de Physique, Université de Montréal,
1375 Avenue Thérèse-Lavoie-Roux, Montreal, Canada
Mila - Quebec Artificial Intelligence Institute,
6666 Rue Saint-Urbain, Montréal, Canada
Ciela - Montreal Institute for Astrophysical Data Analysis
and Machine Learning, Montréal, Canada

Benjamin L. Davis

Center for Astrophysics and Space Science (CASS),
New York University Abu Dhabi, PO Box 129188, Abu Dhabi, UAE

Andrea Macciò

New York University Abu Dhabi, PO Box 129188, Abu Dhabi, United Arab Emirates
Center for Astrophysics and Space Science (CASS),
New York University Abu Dhabi, PO Box 129188, Abu Dhabi, UAE
Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany

Abstract

In astrophysics, experiments are impossible. We thus must rely exclusively on observational data. Other observational sciences increasingly leverage causal inference methods, but this is not yet the case in astrophysics. Here we attempt causal discovery for the first time to address an important open problem in astrophysics: the (co)evolution of supermassive black holes (SMBHs) and their host galaxies. We apply the Peter-Clark (PC) algorithm to a comprehensive catalog of galaxy properties to obtain a completed partially directed acyclic graph (CPDAG), representing a Markov equivalence class over directed acyclic graphs (DAGs). Central density and velocity dispersion are found to cause SMBH mass. We test the robustness of our analysis by random sub-sampling, recovering similar results. We also apply the Fast Causal Inference (FCI) algorithm to our dataset to relax the hypothesis of causal sufficiency, admitting unobserved confounds. Hierarchical SMBH assembly may provide a physical explanation for our findings.

1 Introduction

The puzzle of supermassive black hole (SMBH) and host galaxy coevolution is a long-standing open problem in astronomy [4, 9] More recently, the detection of high-redshift quasars reignited the debate on the formation mechanisms of SMBHs in the early Universe [10]. The main issue we face when trying to understand SMBH and host galaxy coevolution is the direction of causality: are galaxy properties driving SMBH assembly or is SMBH feedback shaping observed galaxy properties? While it is likely that a feedback loop exists, with SMBH mass at a given time affecting galaxy properties at a later time and galaxy properties in turn affecting SMBH mass at an even later time, we only have data about a given galaxy and its SMBH at one point in time.

Given this situation, most of the research on this topic in astronomy has been based either on identifying observational correlations [11, 5] or on developing models, based mainly on numerical simulations [16]. Here we break this pattern, introducing a method that to our knowledge has never before been applied in the field: we apply causal discovery to a state-of-the-art sample of galaxies.

2 Data

We consider a set of 83 early-type galaxies with directly-measured SMBH masses compiled by [6, 15, 3, 1, 12, 2, 13, 14]. The variables we consider are as follows: SMBH mass (M_{BH}), the stellar bulge (spheroid) mass (M_{sph}^*), central velocity dispersion ($\sigma_{0_{sph}}$), effective radius of the spheroid ($R_{e_{sph}}$), and the density at the SMBH sphere of influence ρ_{soi} measured by [14]. Fig. 1 summarizes the definitions underlying these variables. These variables are sufficient to characterize the SMBH and its environment, the central spheroid of the host galaxy. Masses are measured in Solar units, lengths in parsec, and velocities in km/s . Following standard practice, we consider the base-10 logarithms of these quantities. A summary table of the sample is reported in Tab. 1. A pair plot for these variables is shown in Fig. 2.

	ρ_{soi}	M_{BH}	$\sigma_{0_{sph}}$	$R_{e_{sph}}$	M_{sph}^*
mean	2.70	8.54	2.33	0.30	10.84
std	0.83	0.91	0.18	0.67	0.80
min	0.71	5.74	1.54	-1.24	8.03
max	4.67	10.30	2.59	1.49	12.26

Table 1: Summary of the early-type galaxy data set.

3 Methods

3.1 PC algorithm

The Peter and Clark (PC) algorithm [17] is one of the most well-known and time-tested causal discovery algorithms, used to infer causal relationships from observational data. The algorithm seeks to identify a skeletal structure of a causal graph by iteratively testing conditional independencies in the data. Initially, it assumes a fully-connected undirected graph among all variables. In the first phase, edges are removed based on conditional statistical independence tests, gradually increasing the conditioning set size. After obtaining the undirected skeleton, the algorithm enters the orientation phase, where it employs a set of rules to determine the direction of the remaining edges. However, it is not possible in general to orient all the edges, leading to a Partially Directed Acyclic Graph (PDAG), representing a Markov equivalence class of Directed Acyclic Graphs (DAGs). This graph captures the causal relationships that are identifiable from the observed data under certain assumptions, such as faithfulness, causal sufficiency, and acyclicity.

We use the PC algorithm in the implementation by [19] through the *causal-learn* library in Python, relying on the Kernel Conditional Independence test [KCI; 18]. The KCI test is a non-parametric method used to assess the conditional independence between two random variables given a set of conditioning variables. The principle underlying KCI is to exploit the properties of reproducing kernel Hilbert spaces (RKHS). The idea is to embed the distributions of the variables into high-dimensional

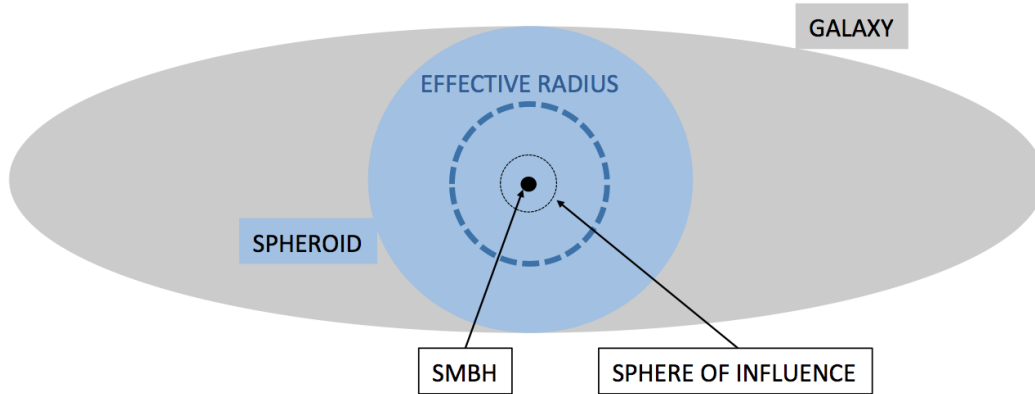


Figure 1: Early-type galaxy schematic, with a possible large-scale, rotationally-supported disk. The spheroid or bulge is the central, pressure (dispersion) supported part of the galaxy. The radius containing half its luminosity or effective radius gives a measure of its size. The SMBH is typically located in a central position. The sphere of influence of the SMBH corresponds to the region where stellar motions are dominated by the SMBH gravity.

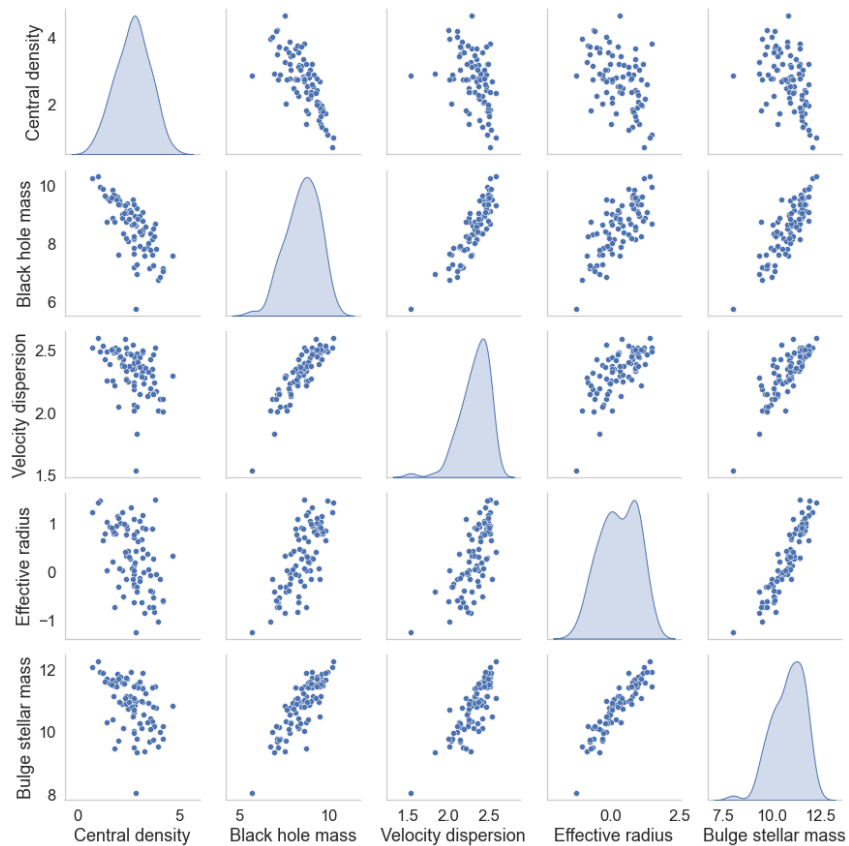


Figure 2: Pair plot of our data set of galaxy properties.

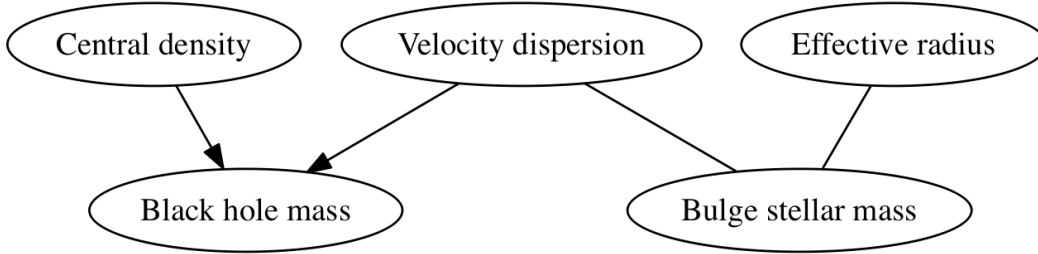


Figure 3: CPDAG learned by the PC algorithm on our data set. Directed edges represent causal dependencies, undirected edges represent lack of independence.

Hilbert spaces and then measure the independence in these spaces using a kernel-based distance metric. Essentially, if two variables are conditionally independent given a third variable, their joint distribution can be factorized into two separate distributions, and this factorization manifests as orthogonality in the RKHS. The KCI test statistically evaluates this orthogonality to decide whether conditional independence holds. By using kernel methods, the KCI test can capture nonlinear dependencies.

We set the threshold for rejecting the null hypothesis of conditional independence to a p-value of 0.1. This is a relatively generous threshold, erring towards rejection, meaning a graph with more edges. We also considered a value of 0.01, obtaining slightly different results, discussed in the following.

4 Results

Fig. 3 shows the Complete Partially Directed Acyclic Graph (CPDAG) which represents the Markov equivalence class learned by the PC algorithm on our data set. The main result is that SMBH mass is a direct effect of central density within the sphere of influence and of velocity dispersion.

4.1 Physical interpretation

Bondi–Hoyle–Lyttleton accretion [7] on a black hole predicts that mass growth is determined by the ambient density and the velocity dispersion of the medium. Similarly, repeated mergers of seed black holes, are likely governed by a similar dependence. with encounter cross-section being determined by relative velocity in the gravitational focusing limit. This is an indication that our causal discovery approach uncovered a genuine physical relation between bulge properties and SMBH mass, especially if we admit that currently observed quantities such as the stellar velocity dispersion are likely to trace primordial quantities.

4.2 Robustness

The PC algorithm relies on (conditional) independence tests between variables. With finite data, the test may fail to find a dependence that is actually present, or vice versa deem that two variables are dependent when they in fact are not. To test this, we change the value of the threshold for conditional independence used by the KCI test from 0.1 to 0.01. The resulting CPDAG is almost identical to the one found initially (shown in Fig. 3) with the only difference being that the edge between velocity dispersion and bulge stellar mass is missing. The main result regarding SMBH mass being caused by central density and velocity dispersion stands unaffected.

We also test the reliability of the results we obtained by re-running our analysis on a subset of the data. We repeat 10 times a procedure where we randomly exclude 10% of our data set and run the PC algorithm on the remaining 90%. This results in five unique different CPDAGs. Three are essentially minor variations on the original computed on the whole data set and are presented in Fig. 5. These account for 60% of the CPDAGs we computed in this fashion. The other two are shown in Fig. 6 (accounting for 30% of the runs) and in Fig. 7 (accounting for 10% of the runs). The last two causal structures differ from the one we found using the whole data set in that black hole mass affects bulge stellar mass, but the result that velocity dispersion and central density cause SMBH mass, still holds.

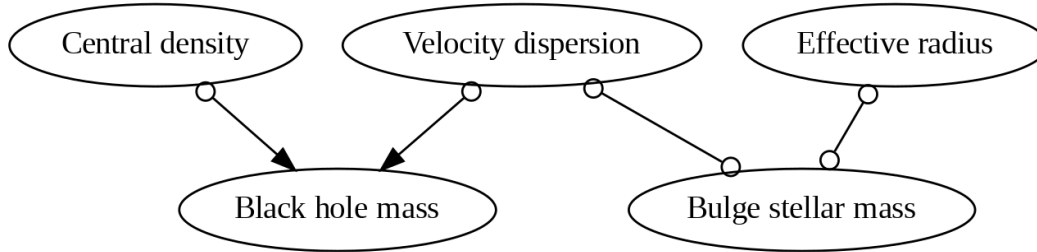


Figure 4: PAG learned by the FCI algorithm on our data set. Directed edges represent causal dependencies, undirected edges represent lack of independence, like in Fig. 3. However here empty circles denote a relation that may be due to an unobserved confound.

4.3 Limitations

Early galactic evolution is rife with quantities we are not able to measure directly. The causal sufficiency assumption underlying the PC algorithm, that is the absence of unobserved confounds, is, therefore, a strong assumption to make. There is some precedent in the literature suggesting that, indeed, observed correlations may be the result of a common mechanism of accretion, rather than a direct causal link [8]. An additional source of confounds is the dynamical evolution of the system itself: past values of any variable may in principle causally affect both the current values of that variable and the current values of other variables. In this regard it should also be noted that we did not consider the possibility of cyclic causal relations. We are considering testing our approach in the future on simulation time-series data, where a ground truth is available and the dynamical nature of the data generating process is under our control.

To address the issue of unobserved confounds we applied the Fast Causal Discovery [FCI; 17] algorithm to our data set. FCI learns a Partial Ancestral Graph (PAG), shown for our data in Fig. 4. PAGs contain additional edge types with respect to CPDAGs. In particular the edges marked with an empty circle represent a situation where there is an association between variables (the null hypothesis of statistical independence has been rejected) but the nature of the association may not be causal. If this is the case, the association is due to one or more unobserved confounds. The FCI algorithm on our dataset indeed cannot decide whether central density and velocity dispersion cause SMBH mass (as found by the PC algorithm) or whether the observed association is due to unobserved confounds. It does however rule out SMBH mass causing central density and velocity dispersion.

5 Conclusions

We applied causal discovery via the PC and FCI algorithms for the first time to a state-of-the-art data set of dynamically measured SMBHs and their host early-type galaxy properties. The causal structure we learned from this data suggests that SMBH mass is the effect of central velocity dispersion and density at the SMBH sphere of influence. The result is robust to changes in hyperparameters such as the threshold for statistical significance in the KCI independence test used by the PC algorithm and to random subsampling of our data set. From the physical point of view, this suggests that the bulge properties determine the SMBH mass by controlling SMBH assembly, which likely takes place through repeated mergers of seed black holes.

Acknowledgments and Disclosure of Funding

M. P. acknowledges financial support from the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 896248. This material is based upon work supported by Tamkeen under the NYU Abu Dhabi Research Institute grant CASS.

References

- [1] Benjamin L. Davis, Alister W. Graham, and Ewan Cameron. Black Hole Mass Scaling Relations for Spiral Galaxies. I. $M_{BH} - M_{*,sph}$. *ApJ*, 873(1):85, March 2019.

- [2] Benjamin L. Davis, Alister W. Graham, and Françoise Combes. A Consistent Set of Empirical Scaling Relations for Spiral Galaxies: The $(v_{max}, M_{oM})-(\sigma_0, M_{BH}, \phi)$ Relations. *ApJ*, 877(1):64, May 2019.
- [3] Benjamin L. Davis, Alister W. Graham, and Marc S. Seigar. Updating the (supermassive black hole mass)-(spiral arm pitch angle) relation: a strong correlation for galaxies with pseudobulges. *MNRAS*, 471(2):2187–2203, October 2017.
- [4] Tiziana Di Matteo, Volker Springel, and Lars Hernquist. Energy input from quasars regulates the growth and activity of black holes and their host galaxies. *Nature*, 433(7026):604–607, February 2005.
- [5] Karl Gebhardt, Ralf Bender, Gary Bower, Alan Dressler, S. M. Faber, Alexei V. Filippenko, Richard Green, Carl Grillmair, Luis C. Ho, John Kormendy, Tod R. Lauer, John Magorrian, Jason Pinkney, Douglas Richstone, and Scott Tremaine. A Relationship between Nuclear Black Hole Mass and Galaxy Velocity Dispersion. *Astrophysical Journal Letters*, 539(1):L13–L16, August 2000.
- [6] Alister W. Graham and Nicholas Scott. The $M_{BH}-L_{spheroid}$ Relation at High and Low Masses, the Quadratic Growth of Black Holes, and Intermediate-mass Black Hole Candidates. *ApJ*, 764(2):151, February 2013.
- [7] Fred Hoyle and RA Lyttleton. On the accretion theory of stellar evolution. *Monthly Notices of the Royal Astronomical Society, Vol. 101, p. 227*, 101:227, 1941.
- [8] Knud Jahnke and Andrea V. Macciò. The Non-causal Origin of the Black-hole-galaxy Scaling Relations. *ApJ*, 734(2):92, June 2011.
- [9] John Kormendy and Luis C. Ho. Coevolution (Or Not) of Supermassive Black Holes and Host Galaxies. *Annual Review of Astronomy and Astrophysics*, 51(1):511–653, August 2013.
- [10] Rebecca L. Larson, Steven L. Finkelstein, Dale D. Kocevski, Taylor A. Hutchison, Jonathan R. Trump, Pablo Arrabal Haro, Volker Bromm, Nikko J. Cleri, Mark Dickinson, Seiji Fujimoto, Jeyhan S. Kartaltepe, Anton M. Koekemoer, Casey Papovich, Nor Pirzkal, Sandro Tacchella, Jorge A. Zavala, Micaela Bagley, Peter Behroozi, Jaclyn B. Champagne, Justin W. Cole, Intae Jung, Alexa M. Morales, Guang Yang, Haowen Zhang, Adi Zitrin, Ricardo O. Amorín, Denis Burgarella, Caitlin M. Casey, Óscar A. Chávez Ortiz, Isabella G. Cox, Katherine Chworowsky, Adriano Fontana, Eric Gawiser, Andrea Grazian, Norman A. Grogan, Santosh Harish, Nimish P. Hathi, Michaela Hirschmann, Benne W. Holwerda, Stéphanie Juneau, Gene C. K. Leung, Ray A. Lucas, Elizabeth J. McGrath, Pablo G. Pérez-González, Jane R. Rigby, Lise-Marie Seillé, Raymond C. Simons, Benjamin J. Weiner, Stephen M. Wilkins, L. Y. Aaron Yung, and The CEERS Team. A CEERS Discovery of an Accreting Supermassive Black Hole 570 Myr after the Big Bang: Identifying a Progenitor of Massive $z > 6$ Quasars. *arXiv e-prints*, page arXiv:2303.08918, March 2023.
- [11] John Magorrian, Scott Tremaine, Douglas Richstone, Ralf Bender, Gary Bower, Alan Dressler, S. M. Faber, Karl Gebhardt, Richard Green, Carl Grillmair, John Kormendy, and Tod Lauer. The Demography of Massive Dark Objects in Galaxy Centers. *Astronomical Journal*, 115(6):2285–2305, June 1998.
- [12] Nandini Sahu, Alister W. Graham, and Benjamin L. Davis. Black Hole Mass Scaling Relations for Early-type Galaxies. I. $M_{BH}-M_{*,sph}$ and $M_{BH}-M_{*,gal}$. *ApJ*, 876(2):155, May 2019.
- [13] Nandini Sahu, Alister W. Graham, and Benjamin L. Davis. Revealing Hidden Substructures in the $M_{BH}-\sigma$ Diagram, and Refining the Bend in the $L-\sigma$ Relation. *ApJ*, 887(1):10, December 2019.
- [14] Nandini Sahu, Alister W. Graham, and Benjamin L. Davis. The (Black Hole Mass)-(Spheroid Stellar Density) Relations: $M_{BH}-\mu$ (and $M_{BH}-\Sigma$) and $M_{BH}-\rho$. *ApJ*, 927(1):67, March 2022.
- [15] Giulia A. D. Savorgnan, Alister W. Graham, Alessandro Marconi, and Eleonora Sani. Supermassive Black Holes and Their Host Spheroids. II. The Red and Blue Sequence in the $M_{BH}-M_{*,sph}$ Diagram. *ApJ*, 817(1):21, January 2016.

- [16] Nadine H. Soliman, Andrea V. Macciò, and Marvin Blank. Co-Evolution vs. Co-existence: The effect of accretion modelling on the evolution of black holes and host galaxies. *MNRAS*, 525(1):12–23, October 2023.
- [17] Peter Spirtes, Clark N Glymour, and Richard Scheines. *Causation, prediction, and search*. MIT press, 2000.
- [18] Kun Zhang, Jonas Peters, Dominik Janzing, and Bernhard Schölkopf. Kernel-based Conditional Independence Test and Application in Causal Discovery. *arXiv e-prints*, page arXiv:1202.3775, February 2012.
- [19] Yujia Zheng, Biwei Huang, Wei Chen, Joseph Ramsey, Mingming Gong, Ruichu Cai, Shohei Shimizu, Peter Spirtes, and Kun Zhang. Causal-learn: Causal Discovery in Python. *arXiv e-prints*, page arXiv:2307.16405, July 2023.

A Additional figures

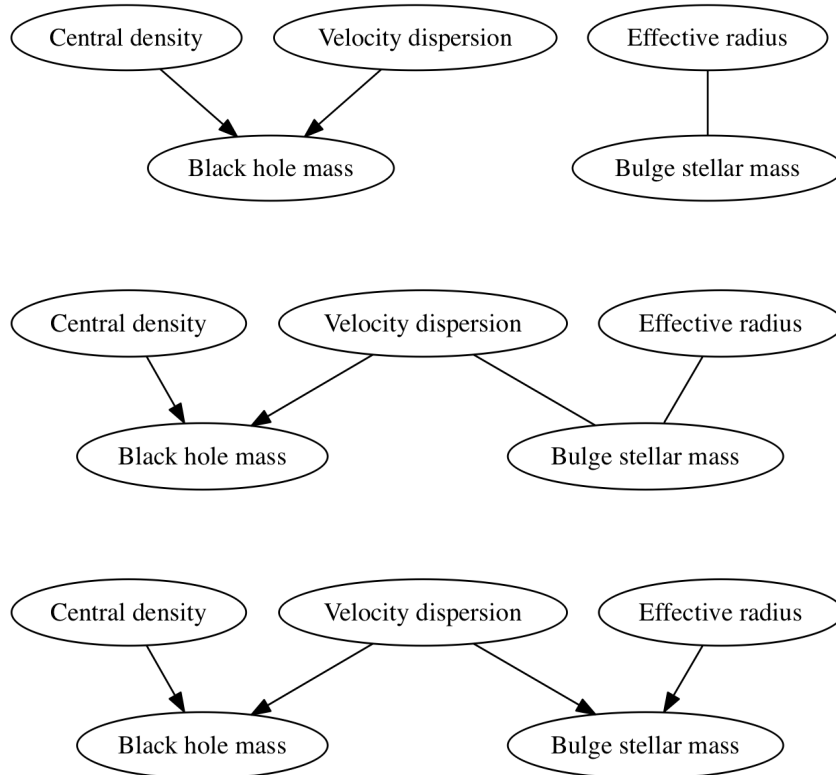


Figure 5: Three CPDAGs learned by the PC algorithm on a random subsample containing 90% of the original data set.

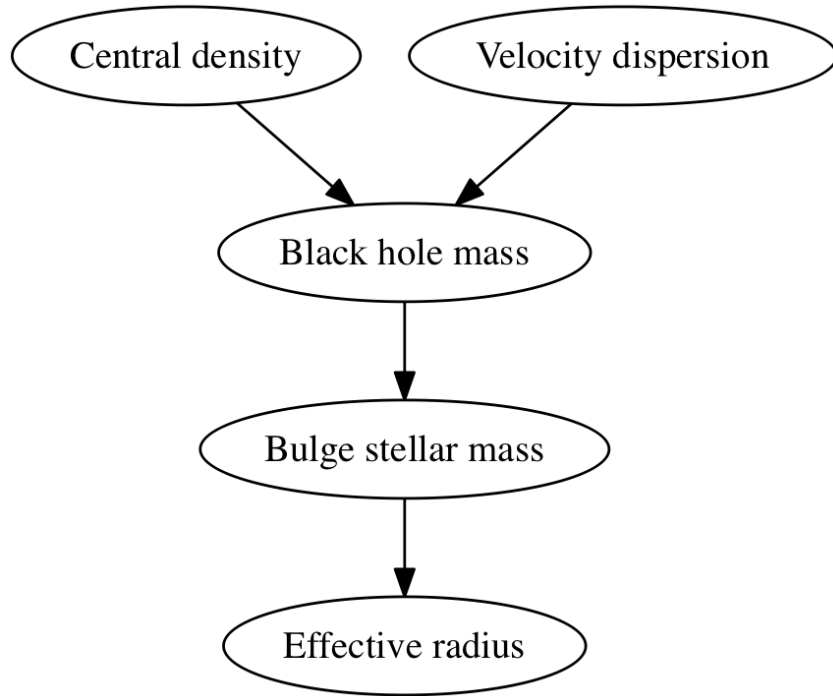


Figure 6: CPDAG learned by the PC algorithm on a random subsample containing 90% of the original data set.

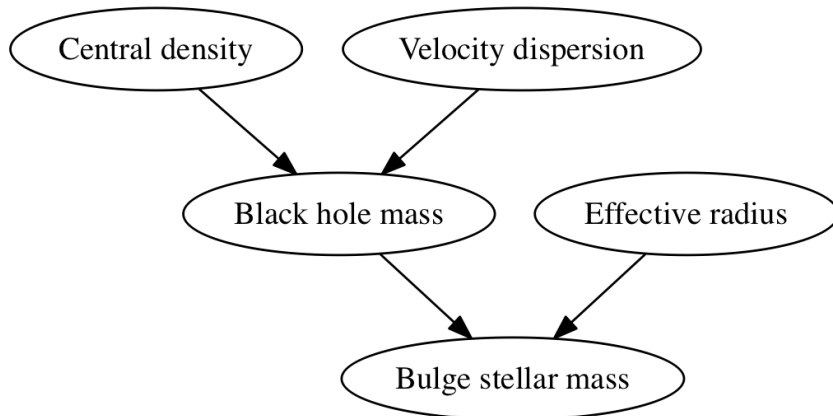


Figure 7: CPDAG learned by the PC algorithm on a random subsample containing 90% of the original data set.