Beyond Causal Discovery for Astronomy: Learning Meaningful Representations with Independent Component Analysis

Zehao Jin (金泽灏)[1,2,3,*], **Mario Pasquato**[3,4,5,6,7,†], **Benjamin L. Davis**[1,2,†], **Andrea Valerio Macciò**[1,2,8], **Yashar Hezaveh**[3,4,5,9]

[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, United Arab Emirates.
[3]Ciela Institute, Montréal, Canada.
[4]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]Max-Planck-Institut für Astronomie, Königstuhl 17, Heidelberg, 69117, Germany.
[9]Center for Computational Astrophysics, Flatiron Institute,
New York, NY, United States of America.

[*]Corresponding author. Email: zj448@nyu.edu
[†]These authors contributed equally to this work.

Abstract

We present the first steps toward applying causal representation learning to astronomy. Following up on previous work that introduced causal discovery to the field for the first time, here we solve a long standing conundrum by identifying the direction of the causal relation between supermassive black hole (SMBH) mass and their host galaxy properties. This leverages a score-based causal discovery approach with an exact posterior calculation. Causal relations between SMBHs and their host galaxies are further clarified by Independent Component Analysis (ICA). The astrophysical problem we focus on is one of the most important open issues in the field and one that has not seen a definitive resolution in decades. We consider the space of six physical properties of galaxies, subdivided by morphology: elliptical, lenticular, and spiral, plus SMBH mass. We calculate an exact posterior over the space of directed acyclic graphs for these variables based on a flat prior and the Bayesian Gaussian equivalent score. The nature of the causal relation between galaxy properties and SMBH mass is found to vary smoothly with morphology, with galaxy properties determining SMBH mass in ellipticals and vice versa in spirals. This settles a long-standing debate and is compatible with our theoretical understanding of galaxy evolution. ICA reveals a decreasing number of meaningful Independent Components (ICs) from ellipticals and lenticular to spiral. Moreover, we find that only one IC correlates with SMBH mass in spirals while multiple ones

38th Conference on Neural Information Processing Systems (NeurIPS 2024).

do in ellipticals, further confirming our finding that SMBH mass causes galaxy properties in spirals, but the reverse holds in ellipticals.

1 Introduction

Astronomy is an observational science and, as such, relies on empirical correlations to test theories. This gives rise to conundrums about the causal significance of measured quantities that cannot be resolved through experimental intervention. An example is the long-standing debate about the causal interpretation of the correlations between the mass of central SMBHs and the properties of their host galaxies [35, 61, 17, 19, 14, 60, 13, 56, 33, 18, 29, 62]. Recently, causal discovery techniques have been proposed as a way out of this impasse [44, 43, 32] and, more generally, there is increasing interest in treating astronomical problems using causality methods [see e.g., 45, 42].

However, this growing body of work typically takes for granted that the variables measured by astronomers are suitable for uncovering causal relations, without engaging with the problem of finding better representations for astronomical data in the spirit of causal representation learning. Our contribution is focused on filling this gap, albeit in a somewhat limited fashion, by applying ICA to the space of galaxy properties introduced by [44]. We show that the resulting ICs have a causal relation with the mass of the SMBH hosted by each galaxy, which depends on galaxy morphology. Namely, in elliptical galaxies the mass of the SMBH is an effect of multiple ICs representing combinations of observed galaxy properties, while in spiral galaxies the mass of the SMBH causes a single IC, which likely represents a coordinate corresponding to the physical coupling mechanism between black hole feedback and galaxy growth. This is in agreement with theoretical expectations from galaxy formation simulations [e.g., 66] and settles the debate on the causal interpretation of the correlations between the mass of central SMBHs and the properties of their host galaxies. Moreover, our contribution constitutes the first advancement toward applying causal representation learning to astronomy.

2 Sample

To explore the causal relationship between SMBHs and their host galaxies, we use the state-of-the-art dataset of a sample of 101 galaxies and their dynamically-measured SMBH masses. The dataset comprises seven variables of interest: dynamically-measured black hole mass ($M_\bullet$), central stellar velocity dispersion (σ_0), effective (half-light) radius of the spheroid[1] (R_e), the average projected density within R_e ($\langle\Sigma_\mathrm{e}\rangle$), total stellar mass ($M^*$), color ($W2 - W3$), and specific star formation rate (sSFR). Among these seven variables, σ_0, R_e, and $\langle\Sigma_\mathrm{e}\rangle$ cover the fundamental plane of elliptical galaxies [15]; while M^*, $W2 - W3$, and sSFR capture the star formation (see §A for additional information regarding the data).

3 Causal discovery methods

To represent the causal structure of the dataset, we use Directed Acyclic Graphs (DAGs). Each DAG encodes a set of conditional independencies, and DAGs that encode the same conditional independencies belong to the same Markov Equivalence Class (MEC). This choice assumes that no cyclical dependencies between variables exist. This is a reasonable assumption, given the clear differences in gas fractions and merger histories between the different morphological classes. To achieve a purely data-driven study, we adopt a uniform prior, giving equal prior probability, $P(G)$, to every one of the nearly 1.14×10^9 possible DAGs [41]. We calculate the exact posterior probabilities of every DAG given the data, $P(G \mid D)$, using the Bayesian Gaussian equivalent (BGe) score [20, 21, 34]. The BGe score gives the marginal likelihood by examining conditional independencies and ensures that DAGs belonging to the same MEC are scored equally.

[1] Throughout this article, we use the terms "bulge" and "spheroid" interchangeably to refer to the spheroid component of spiral and lenticular galaxies or the entirety of pure elliptical galaxies.

4 Causal discovery results

Among all possible causal structures, the most probable MEC and its corresponding DAGs for E, S0, and S galaxies are shown in Fig. 1. We find that in the most probable MEC for elliptical galaxies, the SMBH mass is a causal child, i.e., an effect of galaxy properties, while in the most probable MEC for spirals, the SMBH mass is a parent of galaxy properties (with lenticulars being in the middle).

The morphologically-dependent set trend holds not only in the most probable graphs, but is common over the entire posterior distribution. This can be quantified using edge and path marginals. Edge marginals are the posterior probability of a direct causal relation between two variables, marginalized over the causal structures of the other nodes. Similarly, path marginals provide the probability of a causal connection between two variables through a potentially indirect path (e.g., through intermediate nodes). These marginal causal structures can be represented in matrix form as shown in Fig. 2. The first row ($M_\bullet \to$ galaxy) and column (galaxy $\to M_\bullet$) of each matrix contain information pertaining to the inferred causal relationship between SMBH masses and their host galaxy properties.

Among all possible DAGs, the percentage of graphs exhibiting a direct edge from σ_0 to $M_\bullet$ is 78% in ellipticals, 72% in lenticulars, and only 22% in spirals. The path marginals in the bottom row of matrices support a similar picture, as by considering all possible paths relating these two nodes, we find that 79% of DAGs in ellipticals and 72% in lenticulars have σ_0 as an ancestor of $M_\bullet$, whereas this is the case in only 25% of DAGs in spirals.[2] A detailed physical interpretation of the causal structures found is presented in §B, along with discussions on unobserved confounders in §C and cyclicity in §D.

5 Independent component analysis

Astronomers characterize galaxies through properties that are selected according to subject-matter knowledge and, partly, convention. While these can be taken for granted as a starting point for causal discovery, one of the goals of our work is also to identify potentially better representations for the dataset, in terms of coordinates that may have a clearer causal connection with the outcome of interest, i.e., the mass of the SMBH hosted by any given galaxy. This is a causal representation learning task that in general would rely on detailed causal assumptions. As a first step, we consider ICA as a fast and effective causal representation learning technique under the assumption of an empty causal graph, given ICA's origins as a statistical technique to transform a set of observed signals into a set of statistically independent non-Gaussian components. This allows us to better cross-check and understand the causal structure we found in §4. In fact, ICA has been used to infer causal structures in linear non-Gaussian settings (e.g., ICA-LiNGAM [59]). In our case where the complex real world data is neither a perfect Gaussian or non-Gaussian, we exercise ICA with extra caution, keeping only non-Gaussian ICs.

We use the FastICA algorithm [30] implemented in `scikit-learn` to perform ICA on the space of six galaxy properties (σ_0, R_e, Σ_e, M^*, $W2-W3$, sSFR, without $M_\bullet$), with tolerance 0.0001 and function `logcosh`. As a pre-processing step, these are all log-transformed to ensure that differences in units of measurement do not affect our results. We run FastICA so that it outputs six ICs, but we are aware that not all of these are actually meaningful: in particular some end up having an approximately Gaussian distribution, as discussed below. To retain only meaningful ICs, we select non-Gaussian ones by means of a Shapiro-Wilk test [58], obtaining a p-value for each IC. We thus retain non-Gaussian ICs, defined for our purposes as having a p-value < 0.01, while we ignore the remaining ICs. We then calculate the correlation coefficient between $M_\bullet$ and the ICs. The normalized mixing matrix from ICA, Shapiro-Wilk test p-values, and correlation coefficients are shown in Fig. 3, with Gaussian ICs grayed out.

The causal discovery method outlined in §3 & §4 reveals that galaxy properties cause $M_\bullet$ in E galaxies, while $M_\bullet$ causes galaxy properties in S galaxies. Interestingly, ICA results confirm this finding. Given that the ICs are independent by construction, they cannot have a common cause. Therefore, if $M_\bullet$ causes galaxy properties, it should be statistically associated (i.e., not independent) with one IC at most.

[2] For comparison, the null results (i.e., the posterior from a uniform prior without any data) for the edge marginals are $P(\mathrm{Parent}) = 29\%$, $P(\mathrm{Child}) = 29\%$, and $P(\mathrm{Disconnected}) = 42\%$; for the path marginals these probabilities are $P(\mathrm{Ancestor}) = 42\%$, $P(\mathrm{Descendant}) = 42\%$, and $P(\mathrm{Disconnected}) = 16\%$.

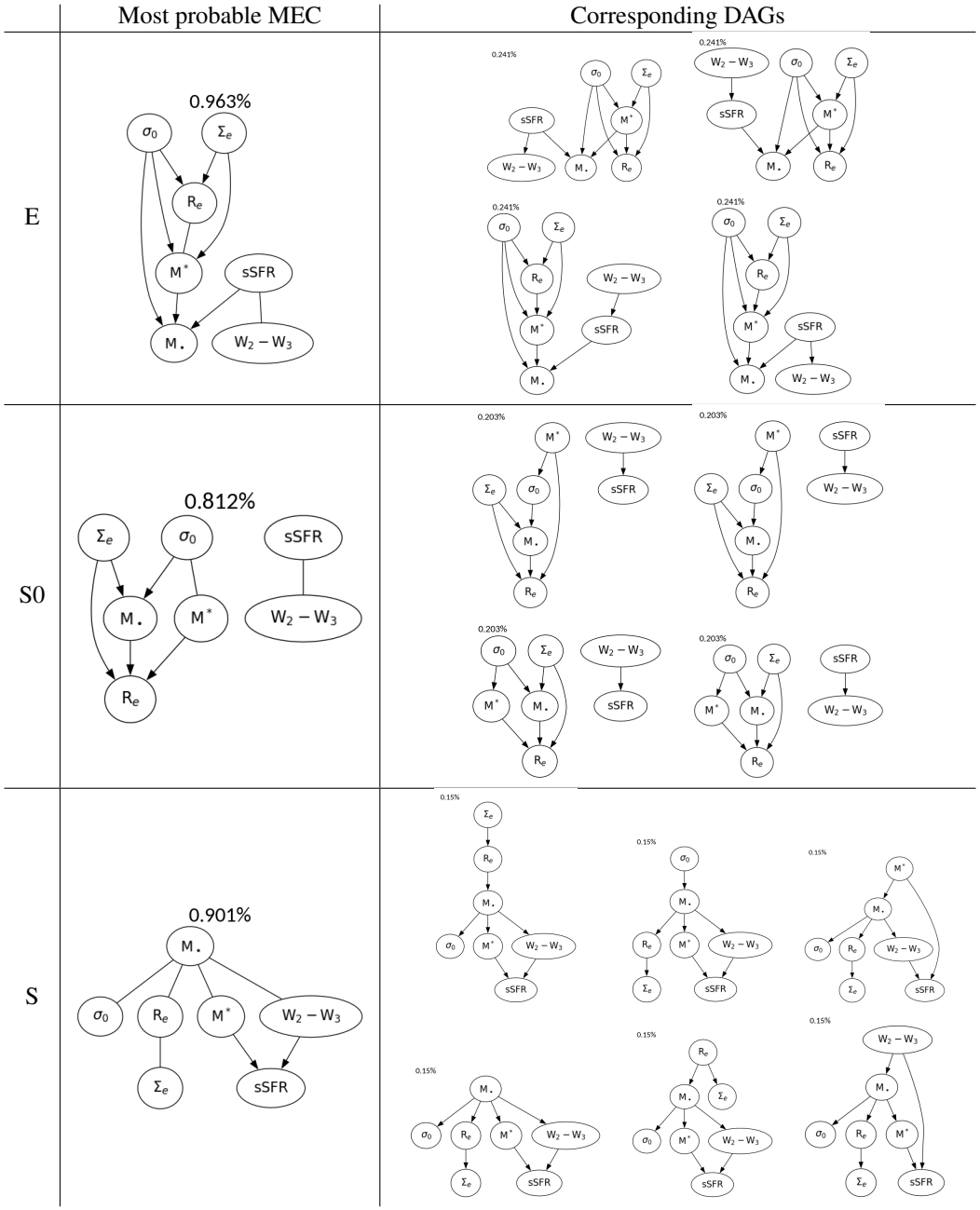

Figure 1: The most probable MEC for each morphology and their corresponding DAGs. MECs are represented as Partially Directed Acyclic Graphs (PDAGs). Directed edges suggest the direction of causality. The undirected edge A — B suggests both directions are possible (either $A \rightarrow B$ or $A \leftarrow B$), as long as no new MEC/conditional independencies are introduced by creating new colliders (i.e., two nodes both pointing towards a third node, $A \rightarrow C \leftarrow B$). In the ellipticals, $M_\bullet$ is strictly a child, while in spiral galaxies, $M_\bullet$ is *always* connected with four galaxy properties through four undirected edges, suggesting either $M_\bullet$ is the parent of all of the four galaxy properties, or $M_\bullet$ is the parent of three of the galaxy properties, and the child of the remaining one (as shown in the corresponding DAGs), ruling out more than one galaxy property pointing towards $M_\bullet$, since this creates a new collider and breaks the encoded conditional independencies. The percentage listed above each graph indicates the posterior probability of the graph, whereas the prior probability for each individual DAG is equal to the reciprocal of the total number of DAGs (approximately 8.78×10^{-10}) [41]. The MEC probabilities are the sum of their corresponding DAGs.

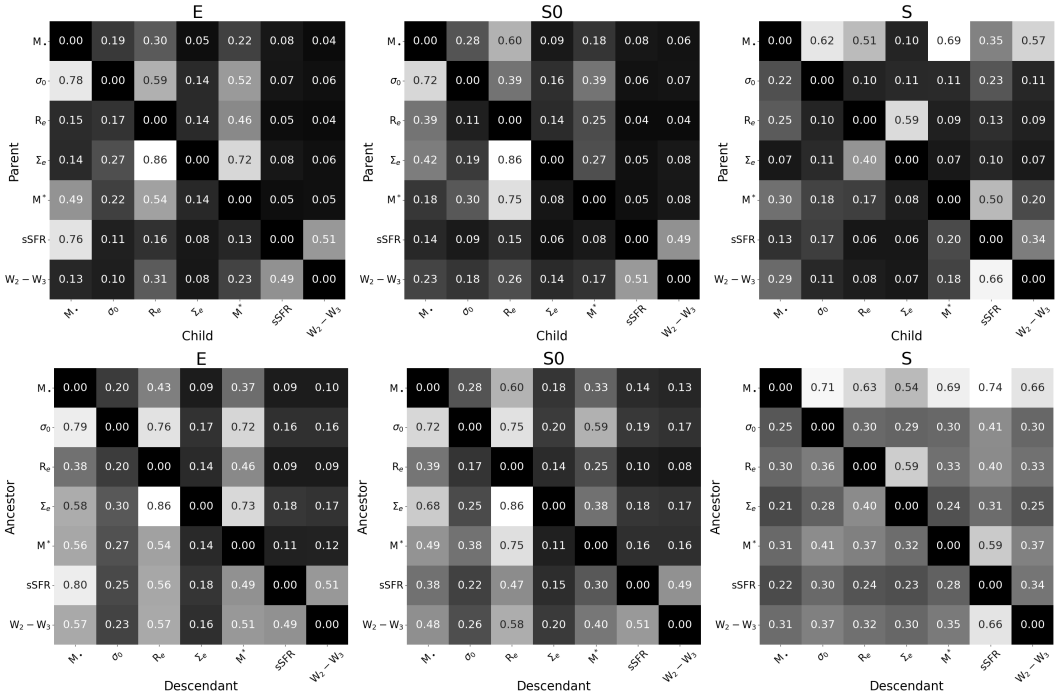

Figure 2: Exact posterior edge marginals (top matrices) and path marginals (bottom matrices) for elliptical (left matrices), lenticular (middle matrices), and spiral (right matrices) galaxies. Edge marginals give the probability of Parent → Child through directed edges summed over all DAGs and their probabilities, and path marginals give the probability of Ancestor → Descendant through both direct and indirect paths.

Vice versa, if it is galaxy properties that cause $M_\bullet$, then it is possible for multiple ICs to be associated with $M_\bullet$. These causal structures we discovered at the previous stage thus indicate that one may find multiple ICs (of galaxy properties) correlated with $M_\bullet$ in E galaxies, while only one IC correlated with $M_\bullet$ in S galaxies. Fig. 3 shows that indeed in E galaxies, at least 2 ICs (IC_2 & IC_6) are correlated with $M_\bullet$, and in S galaxies only IC_1 is correlated with $M_\bullet$, as expected. This consistency further confirms the causal structures identified in this work.

It is worth noting that there are more meaningful ICs in E galaxies (three ICs, namely IC_5, IC_6, and IC_2)[3] than in S galaxies (2 ICs, namely IC_1 and IC_2). This trend suggests that E galaxies are intrinsically more *complicated* than S galaxies because they require more ICs to explain their structures. S galaxies are clearly defined, *rotationally-supported* systems that can be easily described by their more readily-measurable components related to their bulge plus large-scale disk structures. Whereas, E galaxies are *dispersion-supported* systems that consist only of a nearly featureless spheroidal component. In this sense, E galaxies are both more *complicated* (i.e., more meaningful ICs) and *complex* (i.e., chaotic and disordered systems) than S galaxies. This naturally follows the direction of increasing entropy as galaxies evolve from a more settled system (grand-design spirals) to a disorganized system (elliptical blobs of stars).

Interestingly, the ICs of elliptical and lenticular galaxies exhibit some clear similarities, while spiral galaxies behave entirely different. This is shown in Fig. 4 via the cosine similarity among ICs across morphologies. IC_5 of elliptical galaxies essentially coincides with IC_2 of lenticulars, whereas IC_6 of ellipticals corresponds to $-\mathrm{IC}_4$ of lenticulars. Notably, these are the most meaningful ICs (the least Gaussian) both for ellipticals and for lenticulars, and are related to color and star formation rate. No such correspondence emerges for spirals, where the physics of star formation is different due to the presence of large quantities of gas, onto which SMBH feedback impinges.

[3]Note that if instead of a fixed cutoff in p-value we used a visual approach such as the elbow method, the number of ICs could be raised to four, including IC_4 which also has a low p-value.

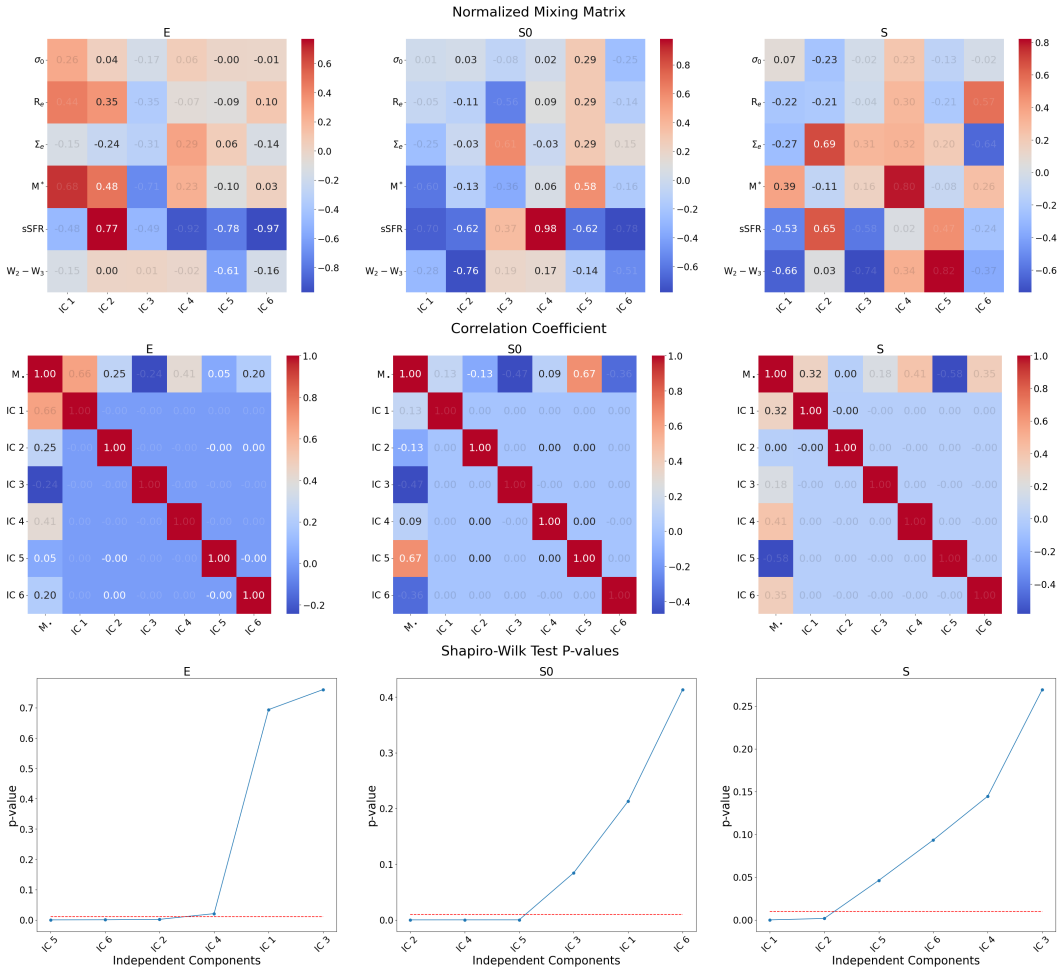

Figure 3: ICA, separately for E (left column), S0 (middle column), and S (right column) morphologies. Top row: Unit-norm normalized mixing matrix heat map for ICs with every variable except black hole mass. Middle row: correlation coefficients between every IC and black hole mass. Throughout the matrices in the top two rows, we display all of the cells with meaningless ICs with faded numbers. Bottom row: Shapiro-Wilk test of normality [58], showing the p-values of every IC. We adopt ICs with p-values lower than $\alpha = 0.01$ as meaningful ICs.

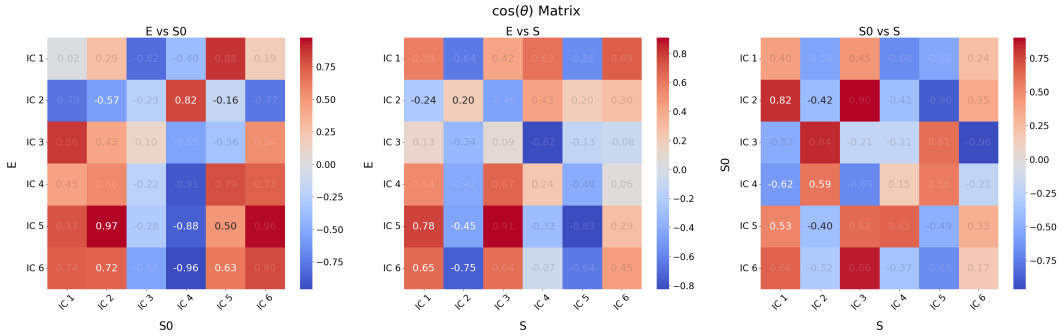

Figure 4: The cosine of angles between columns of mixing matrices among different morphologies. The cosine of the angle between two columns u and v is defined by $\cos\theta = \frac{u \cdot v}{|u| \cdot |v|}$. ICs across different morphologies are likely to carry similar interpretations when $\cos\theta$ is close to either 0 or 1. Similar to Fig. 3, we gray out cells without meaningful ICs.

6 Conclusions

We have presented the first development toward causal representation learning in the form of applying ICA to an astronomical issue involving the causal interpretation of empirically observed correlations, settling a long-standing debate about whether SMBHs determine the properties of their host galaxies or vice versa. Our main finding is that the causal direction runs from the black hole to the galaxy in the case of spiral galaxies, and in the opposite direction in the case of elliptical galaxies. This is consistent with the established theoretical understanding that SMBHs affect galaxy evolution in the gas-rich environment of spiral galaxies through feedback (e.g., by heating the gas through jets), while galaxy properties determine the growth of SMBH mass in ellipticals due to galaxy–galaxy merger rates.

Our finding relies on applying causal discovery in the form of a score-based method with an exact posterior calculation [32] and is confirmed by running ICA on the space of galaxy properties. In particular, ICA shows that in spirals only one IC correlates with SMBH mass, while multiple ICs do in ellipticals, which corroborates the results of causal discovery since multiple ICs can have a combined effect but not a common cause, given that they are independent by construction. Thus together, causal discovery and ICA enhance the interpretability of causal structures in complex datasets such as ours. Causal discovery and ICA can synergize effectively in data analysis, particularly in disentangling multivariate datasets. Here, we have utilized this synergy to provide a more robust interpretation of our first-of-its-kind causal discovery for astronomy by disentangling the causal structure of SMBH–galaxy coevolution.

Acknowledgments and Disclosure of Funding

This research was carried out on the high-performance computing resources at New York University Abu Dhabi. We acknowledge the usage of the HyperLeda database (`http://leda.univ-lyon1.fr`). Z.J. and M.P. wish to extend their heartfelt thanks to Jithendaraa Subramanian for providing in-depth support and clarifications regarding `DAG-GFN`, and to Michelle Liu for comments and discussion. Y.H. thanks Andrew Benson and Dhanya Sridhar for helpful discussions. Z.J. thanks Michael Blanton and Joseph Gelfand for useful suggestions. Z.J. genuinely thanks Mohamad Ali-Dib for his very timely help with HPC technical issues. **Funding:** This material is based upon work supported by Tamkeen under the NYU Abu Dhabi Research Institute grant CASS. This work is partially supported by Schmidt Futures, a philanthropic initiative founded by Eric and Wendy Schmidt as part of the Virtual Institute for Astrophysics (VIA).

References

[1] Emmanuel Bengio, Moksh Jain, Maksym Korablyov, Doina Precup, and Yoshua Bengio. Flow Network based Generative Models for Non-Iterative Diverse Candidate Generation. *Advances in Neural Information Processing Systems (NeurIPS)*, 2021.

[2] Yoshua Bengio, Tristan Deleu, Edward J Hu, Salem Lahlou, Mo Tiwari, and Emmanuel Bengio. GFlowNet Foundations. *Journal of Machine Learning Research (JMLR)*, 2023.

[3] Marvin Blank, Andrea V. Macciò, Aaron A. Dutton, and Aura Obreja. NIHAO - XXII. Introducing black hole formation, accretion, and feedback into the NIHAO simulation suite. *Monthly Notices of the Royal Astronomical Society*, 487(4):5476–5489, August 2019.

[4] H. Bondi. On spherically symmetrical accretion. *Monthly Notices of the Royal Astronomical Society*, 112:195, January 1952.

[5] Carlo Cannarozzo, Alexie Leauthaud, Grecco A. Oyarzún, Carlo Nipoti, Benedikt Diemer, Song Huang, Vicente Rodriguez-Gomez, Alessandro Sonnenfeld, and Kevin Bundy. The contribution of in situ and ex situ star formation in early-type galaxies: MaNGA versus IllustrisTNG. *Monthly Notices of the Royal Astronomical Society*, 520(4):5651–5670, April 2023.

[6] Benjamin Davis and Zehao Jin. Discovery of a Planar Black Hole Mass Scaling Relation for Spiral Galaxies. In *American Astronomical Society Meeting Abstracts*, volume 56 of *American Astronomical Society Meeting Abstracts*, page 152.06, February 2024.

[7] Benjamin L. Davis, Alister W. Graham, and Ewan Cameron. Black Hole Mass Scaling Relations for Spiral Galaxies. II. M $_{BH}$-M $_{*,tot}$ and M $_{BH}$-M $_{*,disk}$. *The Astrophysical Journal*, 869(2):113, December 2018.

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

[9] Benjamin L. Davis, Alister W. Graham, and Françoise Combes. A Consistent Set of Empirical Scaling Relations for Spiral Galaxies: The (v $_{max}$, M $_{oM}$)-(σ $_0$, M $_{BH}$, ϕ) Relations. *The Astrophysical Journal*, 877(1):64, May 2019.

[10] 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. *Monthly Notices of the Royal Astronomical Society*, 471(2):2187–2203, October 2017.

[11] Benjamin L. Davis and Zehao Jin. Discovery of a Planar Black Hole Mass Scaling Relation for Spiral Galaxies. *The Astrophysical Journal Letters*, 956(1):L22, October 2023.

[12] Tristan Deleu, António Góis, Chris Emezue, Mansi Rankawat, Simon Lacoste-Julien, Stefan Bauer, and Yoshua Bengio. Bayesian Structure Learning with Generative Flow Networks. *Conference on Uncertainty in Artificial Intelligence (UAI)*, 2022.

[13] Tiziana Di Matteo, Jörg Colberg, Volker Springel, Lars Hernquist, and Debora Sijacki. Direct Cosmological Simulations of the Growth of Black Holes and Galaxies. *The Astrophysical Journal*, 676(1):33–53, March 2008.

[14] 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.

[15] S. Djorgovski and Marc Davis. Fundamental Properties of Elliptical Galaxies. *The Astrophysical Journal*, 313:59, February 1987.

[16] Jesús Falcón-Barroso, Reynier F. Peletier, and Marc Balcells. Bulges on the Fundamental Plane of early-type galaxies. *Monthly Notices of the Royal Astronomical Society*, 335(3):741–752, September 2002.

[17] Laura Ferrarese and David Merritt. A Fundamental Relation between Supermassive Black Holes and Their Host Galaxies. *The Astrophysical Journal Letters*, 539(1):L9–L12, August 2000.

[18] M. Gaspari, M. Ruszkowski, and S. Peng Oh. Chaotic cold accretion on to black holes. *Monthly Notices of the Royal Astronomical Society*, 432(4):3401–3422, July 2013.

[19] 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. *The Astrophysical Journal Letters*, 539(1):L13–L16, August 2000.

[20] Dan Geiger and David Heckerman. Learning gaussian networks. In *Proceedings of the Tenth International Conference on Uncertainty in Artificial Intelligence*, UAI'94, page 235'Ă'Ş243, San Francisco, CA, USA, 1994. Morgan Kaufmann Publishers Inc.

[21] Dan Geiger and David Heckerman. Parameter priors for directed acyclic graphical models and the characterization of several probability distributions. *Ann. Statist.*, 30(5):1412–1440, 2002.

[22] Alister W. Graham. Invoking the virial theorem to understand the impact of (dry) mergers on the M_{bh}-σ relation. *Monthly Notices of the Royal Astronomical Society*, 518(4):6293–6304, February 2023.

[23] Alister W. Graham. Splitting the lentils: Clues to galaxy/black hole coevolution from the discovery of offset relations for non-dusty versus dusty (wet-merger-built) lenticular galaxies in the M_{bh}-$M_{*,spheroid}$ and M_{bh}-$M_{*,galaxy}$ diagrams. *Monthly Notices of the Royal Astronomical Society*, 521(1):1023–1044, May 2023.

[24] Alister W. Graham. Repainting the colour-mass diagrams by unearthing the green mountain: dust-rich S0 galaxies in the colour-(galaxy stellar mass) diagram, and the colour-(black hole mass) relations for dust-poor versus dust-rich galaxies. *Monthly Notices of the Royal Astronomical Society*, 531(1):230–250, June 2024.

[25] Alister W. Graham, T. H. Jarrett, and M. E. Cluver. Specific star formation rates in the M_{bh}-$M_{*,sph}$ diagram and the evolutionary pathways of galaxies across the sSFR-M_* diagram. *Monthly Notices of the Royal Astronomical Society*, 527(4):10059–10076, February 2024.

[26] Alister W. Graham and Nandini Sahu. Appreciating mergers for understanding the non-linear M_{bh}-$M_{*,spheroid}$ and M_{bh}-$M_{*,galaxy}$ relations, updated herein, and the implications for the (reduced) role of AGN feedback. *Monthly Notices of the Royal Astronomical Society*, 518(2):2177–2200, January 2023.

[27] Alister W. Graham and Nandini Sahu. Reading the tea leaves in the M_{bh}-$M_{*,sph}$ and M_{bh}-$R_{e,sph}$ diagrams: dry and gaseous mergers with remnant angular momentum. *Monthly Notices of the Royal Astronomical Society*, 520(2):1975–1996, April 2023.

[28] 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. *The Astrophysical Journal*, 764(2):151, February 2013.

[29] Timothy M. Heckman and Philip N. Best. The Coevolution of Galaxies and Supermassive Black Holes: Insights from Surveys of the Contemporary Universe. *Annual Review of Astronomy and Astrophysics*, 52:589–660, August 2014.

[30] A. Hyv'd'rinen and E. Oja. Independent component analysis: algorithms and applications. *Neural Networks*, 13(4):411–430, 2000.

[31] Knud Jahnke and Andrea V. Macciò. The Non-causal Origin of the Black-hole-galaxy Scaling Relations. *The Astrophysical Journal*, 734(2):92, June 2011.

[32] Zehao Jin, Mario Pasquato, Benjamin L. Davis, Tristan Deleu, Yu Luo, Changhyun Cho, Pablo Lemos, Laurence Perreault-Levasseur, Yoshua Bengio, Xi Kang, Andrea Valerio Macciò, and Yashar Hezaveh. A Data-driven Discovery of the Causal Connection between Galaxy and Black Hole Evolution. *The Astrophysical Journal*, submitted, 2024.

[33] 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.

[34] Jack Kuipers, Giusi Moffa, and David Heckerman. Addendum on the scoring of Gaussian directed acyclic graphical models. *arXiv e-prints*, page arXiv:1402.6863, February 2014.

[35] 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. *The Astronomical Journal*, 115(6):2285–2305, June 1998.

[36] D. Makarov, P. Prugniel, N. Terekhova, H. Courtois, and I. Vauglin. HyperLEDA. III. The catalogue of extragalactic distances. *Astronomy & Astrophysics*, 570:A13, October 2014.

[37] Federico Marinacci, Mark Vogelsberger, Rüdiger Pakmor, Paul Torrey, Volker Springel, Lars Hernquist, Dylan Nelson, Rainer Weinberger, Annalisa Pillepich, Jill Naiman, and Shy Genel. First results from the IllustrisTNG simulations: radio haloes and magnetic fields. *Monthly Notices of the Royal Astronomical Society*, 480(4):5113–5139, November 2018.

[38] Jeremy Mould. Understanding the Fundamental Plane and the Tully Fisher Relation. *Frontiers in Astronomy and Space Sciences*, 7:21, May 2020.

[39] Jill P. Naiman, Annalisa Pillepich, Volker Springel, Enrico Ramirez-Ruiz, Paul Torrey, Mark Vogelsberger, Rüdiger Pakmor, Dylan Nelson, Federico Marinacci, Lars Hernquist, Rainer Weinberger, and Shy Genel. First results from the IllustrisTNG simulations: a tale of two elements - chemical evolution of magnesium and europium. *Monthly Notices of the Royal Astronomical Society*, 477(1):1206–1224, June 2018.

[40] Dylan Nelson, Annalisa Pillepich, Volker Springel, Rainer Weinberger, Lars Hernquist, Rüdiger Pakmor, Shy Genel, Paul Torrey, Mark Vogelsberger, Guinevere Kauffmann, Federico Marinacci, and Jill Naiman. First results from the IllustrisTNG simulations: the galaxy colour bimodality. *Monthly Notices of the Royal Astronomical Society*, 475(1):624–647, March 2018.

[41] OEIS Foundation Inc. Number of acyclic digraphs (or dags) with n labeled nodes, entry a003024 in the on-line encyclopedia of integer sequences., 2024. Published electronically at `http://oeis.org`.

[42] Xiaoying Pang, Zeqiu Yu, Shih-Yun Tang, Jongsuk Hong, Zhen Yuan, Mario Pasquato, and M. B. N. Kouwenhoven. Disruption of Hierarchical Clustering in the Vela OB2 Complex and the Cluster Pair Collinder 135 and UBC 7 with Gaia EDR3: Evidence of Supernova Quenching. *The Astrophysical Journal*, 923(1):20, December 2021.

[43] Mario Pasquato. Bringing causality to astronomy, feb 2024. PIRSA:24020094 see, `https://pirsa.org`.

[44] Mario Pasquato, Zehao Jin, Pablo Lemos, Benjamin L. Davis, and Andrea V. Macciò. Causa prima: cosmology meets causal discovery for the first time. *arXiv e-prints*, page arXiv:2311.15160, November 2023.

[45] Mario Pasquato and Nadiia Matsiuk. Quasi-experimental Approach to Open Cluster Dynamics. *Research Notes of the American Astronomical Society*, 3(11):179, November 2019.

[46] Annalisa Pillepich, Dylan Nelson, Lars Hernquist, Volker Springel, Rüdiger Pakmor, Paul Torrey, Rainer Weinberger, Shy Genel, Jill P. Naiman, Federico Marinacci, and Mark Vogelsberger. First results from the IllustrisTNG simulations: the stellar mass content of groups and clusters of galaxies. *Monthly Notices of the Royal Astronomical Society*, 475(1):648–675, March 2018.

[47] Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, S. Basak, R. Battye, K. Benabed, J. P. Bernard, M. Bersanelli, P. Bielewicz, J. J. Bock, J. R. Bond, J. Borrill, F. R. Bouchet, F. Boulanger, M. Bucher, C. Burigana, R. C. Butler, E. Calabrese, J. F. Cardoso, J. Carron, A. Challinor, H. C. Chiang, J. Chluba, L. P. L. Colombo, C. Combet, D. Contreras, B. P. Crill, F. Cuttaia, P. de Bernardis, G. de Zotti, J. Delabrouille, J. M. Delouis, E. Di Valentino, J. M. Diego, O. Doré, M. Douspis, A. Ducout, X. Dupac, S. Dusini, G. Efstathiou, F. Elsner, T. A. Enßlin, H. K. Eriksen, Y. Fantaye, M. Farhang, J. Fergusson, R. Fernandez-Cobos, F. Finelli, F. Forastieri, M. Frailis, A. A. Fraisse, E. Franceschi, A. Frolov, S. Galeotta, S. Galli, K. Ganga,

R. T. Génova-Santos, M. Gerbino, T. Ghosh, J. González-Nuevo, K. M. Górski, S. Gratton, A. Gruppuso, J. E. Gudmundsson, J. Hamann, W. Handley, F. K. Hansen, D. Herranz, S. R. Hildebrandt, E. Hivon, Z. Huang, A. H. Jaffe, W. C. Jones, A. Karakci, E. Keihänen, R. Keskitalo, K. Kiiveri, J. Kim, T. S. Kisner, L. Knox, N. Krachmalnicoff, M. Kunz, H. Kurki-Suonio, G. Lagache, J. M. Lamarre, A. Lasenby, M. Lattanzi, C. R. Lawrence, M. Le Jeune, P. Lemos, J. Lesgourgues, F. Levrier, A. Lewis, M. Liguori, P. B. Lilje, M. Lilley, V. Lindholm, M. López-Caniego, P. M. Lubin, Y. Z. Ma, J. F. Macías-Pérez, G. Maggio, D. Maino, N. Mandolesi, A. Mangilli, A. Marcos-Caballero, M. Maris, P. G. Martin, M. Martinelli, E. Martínez-González, S. Matarrese, N. Mauri, J. D. McEwen, P. R. Meinhold, A. Melchiorri, A. Mennella, M. Migliaccio, M. Millea, S. Mitra, M. A. Miville-Deschênes, D. Molinari, L. Montier, G. Morgante, A. Moss, P. Natoli, H. U. Nørgaard-Nielsen, L. Pagano, D. Paoletti, B. Partridge, G. Patanchon, H. V. Peiris, F. Perrotta, V. Pettorino, F. Piacentini, L. Polastri, G. Polenta, J. L. Puget, J. P. Rachen, M. Reinecke, M. Remazeilles, A. Renzi, G. Rocha, C. Rosset, G. Roudier, J. A. Rubiño-Martín, B. Ruiz-Granados, L. Salvati, M. Sandri, M. Savelainen, D. Scott, E. P. S. Shellard, C. Sirignano, G. Sirri, L. D. Spencer, R. Sunyaev, A. S. Suur-Uski, J. A. Tauber, D. Tavagnacco, M. Tenti, L. Toffolatti, M. Tomasi, T. Trombetti, L. Valenziano, J. Valiviita, B. Van Tent, L. Vibert, P. Vielva, F. Villa, N. Vittorio, B. D. Wandelt, I. K. Wehus, M. White, S. D. M. White, A. Zacchei, and A. Zonca. Planck 2018 results. VI. Cosmological parameters. *Astronomy & Astrophysics*, 641:A6, September 2020.

[48] Vicente Rodriguez-Gomez, Shy Genel, Mark Vogelsberger, Debora Sijacki, Annalisa Pillepich, Laura V. Sales, Paul Torrey, Greg Snyder, Dylan Nelson, Volker Springel, Chung-Pei Ma, and Lars Hernquist. The merger rate of galaxies in the Illustris simulation: a comparison with observations and semi-empirical models. *Monthly Notices of the Royal Astronomical Society*, 449(1):49–64, May 2015.

[49] Vicente Rodriguez-Gomez, Annalisa Pillepich, Laura V. Sales, Shy Genel, Mark Vogelsberger, Qirong Zhu, Sarah Wellons, Dylan Nelson, Paul Torrey, Volker Springel, Chung-Pei Ma, and Lars Hernquist. The stellar mass assembly of galaxies in the Illustris simulation: growth by mergers and the spatial distribution of accreted stars. *Monthly Notices of the Royal Astronomical Society*, 458(3):2371–2390, May 2016.

[50] 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 $_{*,gal}$. *The Astrophysical Journal*, 876(2):155, May 2019.

[51] Nandini Sahu, Alister W. Graham, and Benjamin L. Davis. Revealing Hidden Substructures in the M $_{BH}$-σ Diagram, and Refining the Bend in the L-σ Relation. *The Astrophysical Journal*, 887(1):10, December 2019.

[52] Nandini Sahu, Alister W. Graham, and Benjamin L. Davis. Defining the (Black Hole)-Spheroid Connection with the Discovery of Morphology-dependent Substructure in the M_{BH}-n_{sph} and M_{BH}-$R_{e,sph}$ Diagrams: New Tests for Advanced Theories and Realistic Simulations. *The Astrophysical Journal*, 903(2):97, November 2020.

[53] G. Savorgnan, A. W. Graham, A. Marconi, E. Sani, L. K. Hunt, M. Vika, and S. P. Driver. The supermassive black hole mass-Sérsic index relations for bulges and elliptical galaxies. *Monthly Notices of the Royal Astronomical Society*, 434(1):387–397, September 2013.

[54] G. A. D. Savorgnan and A. W. Graham. Supermassive Black Holes and Their Host Spheroids. I. Disassembling Galaxies. *The Astrophysical Journals*, 222(1):10, January 2016.

[55] 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. *The Astrophysical Journal*, 817(1):21, January 2016.

[56] Joop Schaye, Claudio Dalla Vecchia, C. M. Booth, Robert P. C. Wiersma, Tom Theuns, Marcel R. Haas, Serena Bertone, Alan R. Duffy, I. G. McCarthy, and Freeke van de Voort. The physics driving the cosmic star formation history. *Monthly Notices of the Royal Astronomical Society*, 402(3):1536–1560, March 2010.

[57] Nicholas Scott, Alister W. Graham, and James Schombert. The Supermassive Black Hole Mass-Spheroid Stellar Mass Relation for Sérsic and Core-Sérsic Galaxies. *The Astrophysical Journal*, 768(1):76, May 2013.

[58] S. S. SHAPIRO and M. B. WILK. An analysis of variance test for normality (complete samples). *Biometrika*, 52(3-4):591–611, 12 1965.

[59] Shohei Shimizu, Patrik O. Hoyer, Aapo Hyv228;rinen, and Antti Kerminen. A linear non-gaussian acyclic model for causal discovery. *Journal of Machine Learning Research*, 7(72):2003–2030, 2006.

[60] Debora Sijacki, Volker Springel, Tiziana Di Matteo, and Lars Hernquist. A unified model for AGN feedback in cosmological simulations of structure formation. *Monthly Notices of the Royal Astronomical Society*, 380(3):877–900, September 2007.

[61] Joseph Silk and Martin J. Rees. Quasars and galaxy formation. *Astronomy & Astrophysics*, 331:L1–L4, March 1998.

[62] 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. *Monthly Notices of the Royal Astronomical Society*, 525(1):12–23, October 2023.

[63] Peter Spirtes. An anytime algorithm for causal inference. In Thomas S. Richardson and Tommi S. Jaakkola, editors, *Proceedings of the Eighth International Workshop on Artificial Intelligence and Statistics*, volume R3 of *Proceedings of Machine Learning Research*, pages 278–285. PMLR, 04–07 Jan 2001. Reissued by PMLR on 31 March 2021.

[64] Peter Spirtes, Clark N Glymour, and Richard Scheines. *Causation, prediction, and search*. MIT press, 2000.

[65] Volker Springel, Rüdiger Pakmor, Annalisa Pillepich, Rainer Weinberger, Dylan Nelson, Lars Hernquist, Mark Vogelsberger, Shy Genel, Paul Torrey, Federico Marinacci, and Jill Naiman. First results from the IllustrisTNG simulations: matter and galaxy clustering. *Monthly Notices of the Royal Astronomical Society*, 475(1):676–698, March 2018.

[66] Volker Springel, Simon D. M. White, Adrian Jenkins, Carlos S. Frenk, Naoki Yoshida, Liang Gao, Julio Navarro, Robert Thacker, Darren Croton, John Helly, John A. Peacock, Shaun Cole, Peter Thomas, Hugh Couchman, August Evrard, Jörg Colberg, and Frazer Pearce. Simulations of the formation, evolution and clustering of galaxies and quasars. *Nature*, 435(7042):629–636, June 2005.

[67] Liang Wang, Aaron A. Dutton, Gregory S. Stinson, Andrea V. Macciò, Camilla Penzo, Xi Kang, Ben W. Keller, and James Wadsley. NIHAO project - I. Reproducing the inefficiency of galaxy formation across cosmic time with a large sample of cosmological hydrodynamical simulations. *Monthly Notices of the Royal Astronomical Society*, 454(1):83–94, November 2015.

[68] Stefan Waterval, Andrea V. Macciò, Tobias Buck, Aura Obreja, Changhyun Cho, Zehao Jin, Benjamin L. Davis, Keri L. Dixon, and Xi Kang. HELLO project: high-z evolution of large and luminous objects. *Monthly Notices of the Royal Astronomical Society*, 533(2):1463–1484, September 2024.

[69] Edward L. Wright, Peter R. M. Eisenhardt, Amy K. Mainzer, Michael E. Ressler, Roc M. Cutri, Thomas Jarrett, J. Davy Kirkpatrick, Deborah Padgett, Robert S. McMillan, Michael Skrutskie, S. A. Stanford, Martin Cohen, Russell G. Walker, John C. Mather, David Leisawitz, III Gautier, Thomas N., Ian McLean, Dominic Benford, Carol J. Lonsdale, Andrew Blain, Bryan Mendez, William R. Irace, Valerie Duval, Fengchuan Liu, Don Royer, Ingolf Heinrichsen, Joan Howard, Mark Shannon, Martha Kendall, Amy L. Walsh, Mark Larsen, Joel G. Cardon, Scott Schick, Mark Schwalm, Mohamed Abid, Beth Fabinsky, Larry Naes, and Chao-Wei Tsai. The Wide-field Infrared Survey Explorer (WISE): Mission Description and Initial On-orbit Performance. *The Astronomical Journal*, 140(6):1868–1881, December 2010.

A Data

$M_\bullet$ values are curated from the literature on dynamical black hole mass measurements, and σ_0 values are obtained from the HyperLeda database [36]. R_e and $\langle\Sigma_\mathrm{e}\rangle$ measurements come from multi-component decompositions of surface brightness light profiles (primarily of $3.6\,\mu\mathrm{m}$ *Spitzer* Space Telescope imaging) from succeeding works [54, 8, 50, 27]. M^*, $W2-W3$, and sSFR are from the Wide-field Infrared Survey Explorer, WISE [69, 25]. The data has been used in a series of works related to black hole mass scaling relations [28, 57, 53, 55, 50, 51, 52, 26, 10, 7, 8, 9, 11, 6]. To investigate the effect of galaxy morphologies on the underlying causal structure, we further split our sample into 35 elliptical (E), 38 lenticular (S0), and 28 spiral (S) galaxies. This choice is motivated by the observed difference in intrinsic scatter (ϵ) in the $M_\bullet$–σ_0 relation [17, 19] in elliptical ($\epsilon = 0.31$ dex) vs. spiral galaxies ($\epsilon = 0.67$ dex) [51], and this choice is consistent with the current understanding of quenching and hierarchical assembly [e.g., 66]. See Fig. 5 for a visualization of the data.

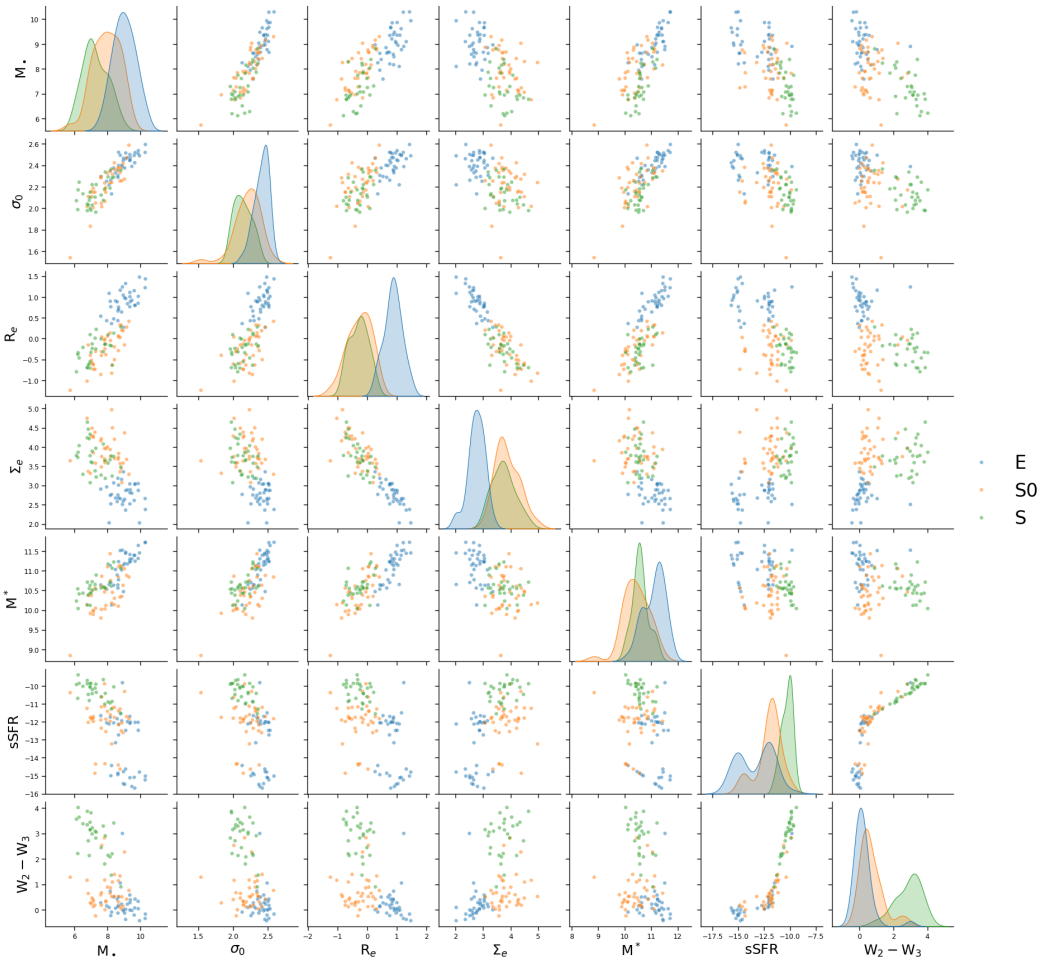

Figure 5: A pairplot of the data for 101 galaxies, separated into elliptical (E), lenticular (S0), and spiral (S) morphologies.

B Discussion on black hole and galaxy evolution

B.1 Causal connections for galaxy evolution

We find that these results are consistent with theoretical models of galactic evolution. Ellipticals are highly-evolved galaxies, being the result of a large number of galactic mergers. Modern hydrodynam-

ical cosmological simulations such as IllustrisTNG [37, 39, 40, 46, 65] show that elliptical galaxies with $\log(M^*/\mathrm{M}_\odot) \geq 11$ are generally the end result of two or more major merger events, such that the typical present-day fraction of stars with *ex situ* origins is greater than 50% [5].[4] In even more general terms, the process of successive mergers will act to erase the preexisting causal connection from the SMBH to its host galaxy and establish new correlations via the central limit theorem [31].

During a merger, the SMBHs at the center of each merging galaxy play no role in the large-scale dynamics; it is the galaxy properties (chiefly size and mass) that shape the galaxy mergers and their outcomes. Central SMBHs are passively driven to the bottom of the post-galaxy-merger potential well by dynamical friction, eventually merging together. So it stands to reason to expect that in ellipticals, the distribution of SMBH masses is determined by that of galaxy properties and *not* vice versa.

For spiral galaxies, this is not the case, since they experience at most a few relatively minor mergers. Unlike elliptical galaxies, spirals are predominantly composed of *in situ* stellar populations. Causal relations between SMBH mass and galaxy properties may thus be set primordially in a secular coevolution phase, and they are not erased by mergers. As a result, spiral galaxies behave markedly different compared to ellipticals. Interestingly, lenticulars appear to lie in-between, as expected, based on the fact that lenticulars have undergone enough mergers to erase spiral structure while still maintaining an extended disk structure, but are not yet comparable to ellipticals in terms of mass and pressure support.[5] Moreover, by extension of Cannarozzo et al.'s results to all early-type (i.e., lenticular and elliptical) galaxies, all but the most massive lenticular galaxies should still maintain *in situ* stellar fractions greater than 50% [5].

The six galaxy variables studied here can be split into the three parameters defining the fundamental plane (FP) of elliptical galaxies [15] and three parameters related to star formation. The FP is a manifestation of dynamical equilibrium reached in the largely pressure-supported stellar dynamics of massive elliptical galaxies [38]. Moreover, it is a consequence of the merger formation of these galaxies via dissipation and feedback that ultimately places them on the FP. Although only 35/101 of the galaxies are ellipticals, the classical bulges of lenticular and spiral galaxies are also governed by the FP. Indeed, it has been found that the bulges of type S0–Sbc galaxies tightly follow the same FP relation as ellipticals [16].

The matrices in Fig. 2 also provide information about the causal nature of the observed FP relationship. By looking at the path marginals for elliptical galaxies (bottom left), we find that $\langle\Sigma_\mathrm{e}\rangle$ is the ancestor (86%) of R_e and that σ_0 is an ancestor (76%) of R_e. This implies $\langle\Sigma_\mathrm{e}\rangle$ and σ_0 are both upstream of R_e, confirming that the density and dynamics of stellar populations in an elliptical galaxy govern its size. Furthermore, we find that there is nearly no chance that M^* is disconnected from R_e (i.e., $54\% + 46\% = 100\%$, they are *never* d-separated, thus *always* correlated), indicating the existence of a size–mass relation due to the virial theorem (i.e., $M \sim \sigma^2 R$).

B.2 Causal active galactic nuclei feedback

From Fig. 2, we find that, in spirals, $M_\bullet$ is the ancestor (74%) of sSFR, in lenticulars, there is no dominant causal direction between the two parameters (38% and 14%), while in ellipticals, $M_\bullet$ becomes the descendant (80%) of the galaxy's sSFR. This can be interpreted as a direct consequence of the presence or absence of gas through active galactic nuclei (AGN) feedback. If there is a substantial gas reservoir (as in spirals), the SMBH is the ancestor since its feedback is responsible for shutting down star formation and hence stopping the growth of stellar mass. With a dearth of gas, as in ellipticals, even large AGN bursts will not affect the stellar mass, and thus the SMBH cannot be an ancestor of galaxy properties. This is further supported by the fact that we find that $M_\bullet$ is the parent (69%) of M^* in spirals, but becomes the descendant (56%) or child (49%) of M^* in elliptical galaxies. However, it is true that in the absence of gas, mergers are the main pathway for SMBH growth, and this will also cause the SMBH to become a descendant or child in hierarchical assembly [31, 26, 22].

[4] Here, Cannarozzo et al. [5] follow previous work [48, 49] and define a major merger as a stellar mass ratio greater than 1/4 between the two progenitors of a given galaxy.

[5] The coevolution of lenticular galaxies and their black holes is also strongly influenced by the presence of dust [23, 24].

We also crosscheck our results with two alternative causal discovery methods, both *constraint-based*: the Peter-Clark, PC, [64] algorithm and the Fast Causal Inference, FCI, [63] algorithm, which both yield consistent results with the exact posterior approach. Additionally, we test the inclusion of the distance to galaxies as a substitute variable, exploring the possibility of it being a hidden confounder. We find that the causal relations identified are not altered. Furthermore, we find that the inferred causal relations are robust to observational errors using random sampling and to possible outlier galaxies using leave-one-out cross-validation.

The exact posterior methodology employed here for causal discovery is a powerful tool for ascertaining causal structures in a purely data-driven manner. However, for problems with more variables, this exact approach becomes computationally intractable due to the combinatorial increase in the number of possible DAGs. In these cases, it remains possible to quantify the posterior over DAGs through posterior samples generated with samplers such as `DAG-GFN` [12], built on GFlowNets [1, 2]. We sampled the posterior by training a `DAG-GFN`, giving results consistent with the exact-posterior approach.

Further insights can be gained by using time-series data and control variables in galaxy simulations [e.g., 68] to test the causal findings and explanations presented here. With knowledge of the underlying causal structures and mechanisms behind galaxy–SMBH coevolution, it should ultimately be possible to create physically-motivated black hole mass scaling relations.

We present the first data-driven evidence on the direction of the causal relationship between SMBHs and their host galaxies. Our findings reveal that in elliptical galaxies, bulge properties influence SMBH growth, whereas in spiral galaxies, SMBHs shape galaxy characteristics. The process of quenching can be causally explained as follows:

1. quenching starts in gas rich (i.e., spiral) galaxies, and hence there is a causal connection
2. the quenching is over in elliptical galaxies, where we only see the end product of such quenching, and the causal connection is now reversed.

These findings support theoretical models of galactic evolution driven by feedback processes and mergers. The successful application of causal discovery to this astrophysical dataset paves the way for a deeper understanding of the fundamental physical processes driving galaxy evolution and establishes causal discovery as a powerful tool for data-driven breakthroughs across various scientific disciplines.

C Discussion on unobserved confounders

Our posterior calculation approach implicitly adopts the assumption of causal sufficiency, i.e., assuming there are no unobserved confounders[6]. With the presence of an unobserved confounder, non-existing causal relations might be falsely identified. Some potential unobserved confounders, such as the reserve of gas or merger history, are practically difficult to observe but are already integrated into our interpretation. However, the distance from us to galaxies does not directly play any role in galaxy formation theory nor in our interpretation, but might influence multiple variables we examined, since our ability to measure all these seven variables decreases as distance increases and thus bias our sample towards nearby and more massive BHs/galaxies. Therefore, we examined the impact of distance by performing causal discovery with distance as one of the seven variables.

Since $W2 - W3$ and sSFR are highly degenerate with each other, we replaced $W2 - W3$ with D_L, the luminosity distance to the targets[7]. The edge and path marginals with distance included are presented in Fig. 6. Comparing against the original marginals without distance (Fig. 2), the presence of distance barely changes any previously identified causal relations, since the edge and path marginals between galaxy properties and SMBH masses remains unchanged with or without the inclusion of distance.

[6] An unobserved confounder is a variable that is not included in the analysis, but is a cause of two or more variables of interest.

[7] The luminosity distances are adopted from [26]. Indeed, this sample of dynamically-measured black hole masses comes from galaxies that are all in the local Universe (median $D_L = 19.3\,\mathrm{Mpc}$; $z = 0.00439$ according to [47]).

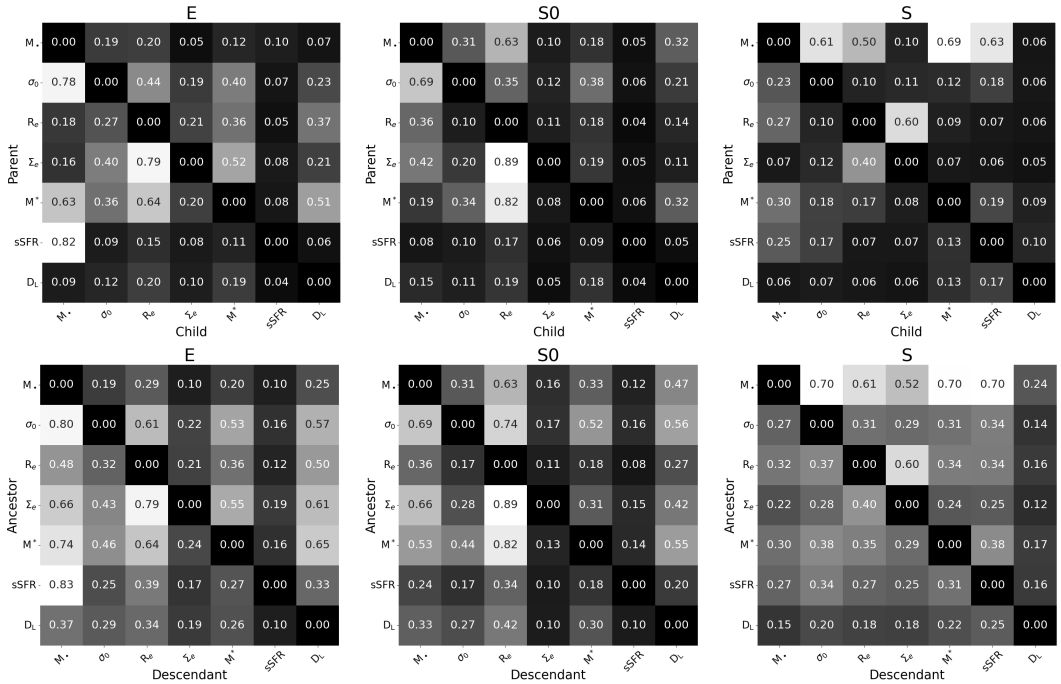

Figure 6: Edge marginals (top matrices) and path marginals (bottom matrices) with luminosity distance (D_L) as one of the variables.

D Discussion on cyclicity

By calculating the posterior probabilities of all possible DAGs, we implicitly assumed acyclicity, i.e., no loops in a graph. In fact, the existence of feedback loops between black hole mass and galaxy properties (i.e., having black hole mass causing the galaxy properties, and then galaxy properties also causing black hole mass at the same time) is trivial in ellipticals and spirals according to galaxy formation theory. Black holes affect their host galaxies through black hole feedback, a process that heats the gas and pushes gas out to starve star formation, while galaxies also affect the central black hole through mergers and accretion. In an ideal spiral galaxy, there have been (at most) only minor mergers, thus killing off the merger path of galaxy $\rightarrow$ black hole.

The accretion onto the black hole is mainly regulated by the black hole mass itself and the gas density in the central region [4]. This latter quantity is found to be relatively constant in gas-rich galaxies, as confirmed by modern numerical simulations, like the NIHAO suite [67, 3] as shown in Figure 7. This implies that accretion is fairly constant in all gas-rich galaxies, diminishing the causal relation galaxy $\rightarrow$ black hole.

Therefore, in spiral galaxies, the causal relation of galaxy $\rightarrow$ black hole is expected to be very weak compared to the black hole $\rightarrow$ galaxy direction. On the other hand, ellipticals are in short supply of gas, therefore the central SMBH lacks the media in which to project its energy to regulate star formation. As a result, in ellipticals, the black hole $\rightarrow$ galaxy direction is negligible compared to the galaxy $\rightarrow$ black hole path enabled by major mergers.

In all (in spirals and ellipticals), one of the causal directions between SMBHs and galaxies is expected to considerably overwhelm the other, making the causal structure acyclic. The lenticulars, however, might have both major mergers and black hole feedback simultaneously, thus being more cyclic in their causal structure. To fully identify cyclic causal structures, time-series data is usually required. While in our case of SMBH–galaxy coevolution, which happens on a timescale of billions of years, obtaining time-series data is impossible within the lifetime of humanity[8], studies of samples of galaxies with different ages may provide observational clues about the presence or absence of cyclicity in future studies.

[8]Except in simulations, which we will investigate in future work.

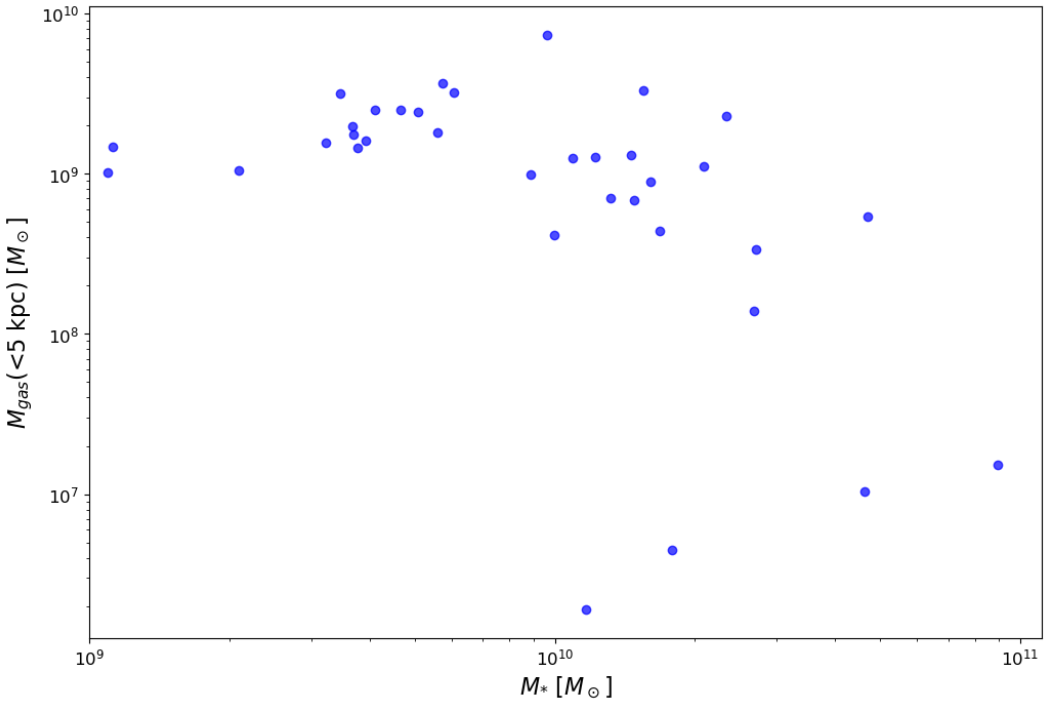

Figure 7: Gas mass within 5 kpc versus total stellar mass in NIHAO simulated galaxies [67]. The central gas mass is fairly constant in gas-rich galaxies, implying that gas accretion onto the black hole, which is mainly regulated by the black hole mass and the local gas density [4], is also quite uniform across galaxies, weakening the galaxy $\rightarrow$ SMBH causal relation in spirals.