

A data-driven discovery of the causal connection between galaxy and black hole evolution

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}, Yashar Hezaveh^{3,4,5,11}

¹New York University Abu Dhabi,

P.O. Box 129188, Abu Dhabi, United Arab Emirates.

²Center for Astrophysics and Space Science (CASS), New York University Abu Dhabi,

P.O. Box 129188, Abu Dhabi, United Arab Emirates.

³Ciela Institute, Montréal, Canada.

⁴Mila - Quebec Artificial Intelligence Institute,

6666 Rue Saint-Urbain, Montréal, Canada.

⁵Département de Physique, Université de Montréal,

1375 Avenue Thérèse-Lavoie-Roux, Montréal, Canada.

⁶Dipartimento di Fisica e Astronomia, Università di Padova,

Vicolo dell’Osservatorio 5, Padova, Italy.

⁷Istituto di Astrofisica Spaziale e Fisica Cosmica (INAF IASF-MI),

Via Alfonso Corti 12, I-20133, Milan, Italy.

⁸Département d’Informatique et de Recherche Opérationnelle, Université de Montréal,

2920 Chemin de la Tour, Montréal, Canada.

⁹Department of Physics, School of Physics and Electronics, Hunan Normal University,

Changsha 410081, China.

¹⁰Purple Mountain Observatory, 10 Yuan Hua Road, Nanjing 210034, China.

¹¹Center for Computational Astrophysics, Flatiron Institute,

New York, NY, United States of America.

¹²Canadian Institute for Advanced Research Artificial Intelligence Chair.

¹³Canadian Institute for Advanced Research Senior Fellow.

¹⁴Institute for Astronomy, Zhejiang University, Hangzhou 310027, China.

¹⁵Max-Planck-Institut für Astronomie, Königstuhl 17, Heidelberg, 69117, Germany.

*Corresponding author. Email: zj448@nyu.edu

†These authors contributed equally to this work.

Correlations between galaxies and their supermassive black holes (SMBHs) have been observed, but the causal mechanisms remained unclear. The emerging field of causal inference now enables examining these relationships using observational data. This study, using advanced causal discovery techniques and a state-of-the-art dataset, reveals a causal link between galaxy properties and SMBH masses. In elliptical galaxies, bulge properties influence SMBH growth, while in spiral galaxies, SMBHs affect host galaxy properties, potentially through feedback in gas-rich environments. For spiral galaxies, SMBHs progressively quench star formation, whereas in elliptical galaxies, quenching is complete, and the causal connection has reversed. These findings support theoretical models of active galactic nuclei feedback regulating galaxy evolution and suggest further exploration of causal links in astrophysical and cosmological scaling relations.

Purely observational sciences have long relied on correlations between variables to assess the validity of theoretical models. However, observed correlations between two variables do not provide information about the direction of causality, making it impossible to discriminate between different causal mechanisms that predict the same correlational trends. While interventions (such as randomized controlled trials) are commonly used to identify causal factors in laboratory settings, this is rarely possible in observational fields such as astrophysics. Causal inference overcomes this limitation by exploiting the fact that different causal models produce distinct joint distributions of correlated variables with additional observables, allowing us to discriminate between these models and shed light on the direction of causality. With this, it becomes possible to investigate causal relationships in a purely data-driven manner.

Dynamically-measured SMBH host galaxies

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 spheroid¹ (R_e), the average projected density within R_e ($\langle \Sigma_e \rangle$), total stellar mass (M^*), color ($W2 - W3$), and specific star formation rate (sSFR). Among these seven variables, σ_0 , R_e , and $\langle \Sigma_e \rangle$ cover the fundamental plane of elliptical galaxies (1); while M^* , $W2 - W3$, and sSFR capture the star formation (see Table S1 and its pairplot in Fig. S1). M_\bullet values are curated from the literature on dynamical black hole mass measurements, and σ_0 values are obtained from the HyperLeda database (2). R_e and $\langle \Sigma_e \rangle$ measurements come from multi-component decompositions of surface brightness light profiles (primarily of 3.6 μm *Spitzer* Space Telescope imaging) from succeeding works (3–6).

¹Throughout 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.

M^* , $W2 - W3$, and sSFR are from the Wide-field Infrared Survey Explorer, WISE (7). The data has been used in a series of work related to black hole mass scaling relations (4, 5, 8–19), and more detailed information about the data can be found in §**Methods & Materials**. 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 $M_\bullet - \sigma_0$ relation (20, 21) in elliptical ($\epsilon = 0.31$ dex) vs. spiral galaxies ($\epsilon = 0.67$ dex) (12) and this choice is consistent with current understanding of quenching and hierarchical assembly (22).

A compendium of causal structures

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)². 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 (see §**Methods & Materials** for more details). 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 (23). We calculate the exact posterior probabilities of every DAG given the data, $P(G | D)$, using the Bayesian Gaussian equivalent (BGe) score (24–26). The BGe score gives the marginal likelihood by examining conditional independencies and ensures that DAGs belonging to the same MEC are scored equally.

Among all possible causal structures, the most probable MEC and its corresponding DAGs for E, S0, and S galaxies are shown in Fig. 1. More detailed information about the DAGs, MECs, and their exact posterior distributions can be found in §**Methods & Materials**. 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 ($M_\bullet \rightarrow$ galaxy) and column (galaxy $\rightarrow M_\bullet$) 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

²See §**Methods & Materials** for a brief introduction to causal inference.

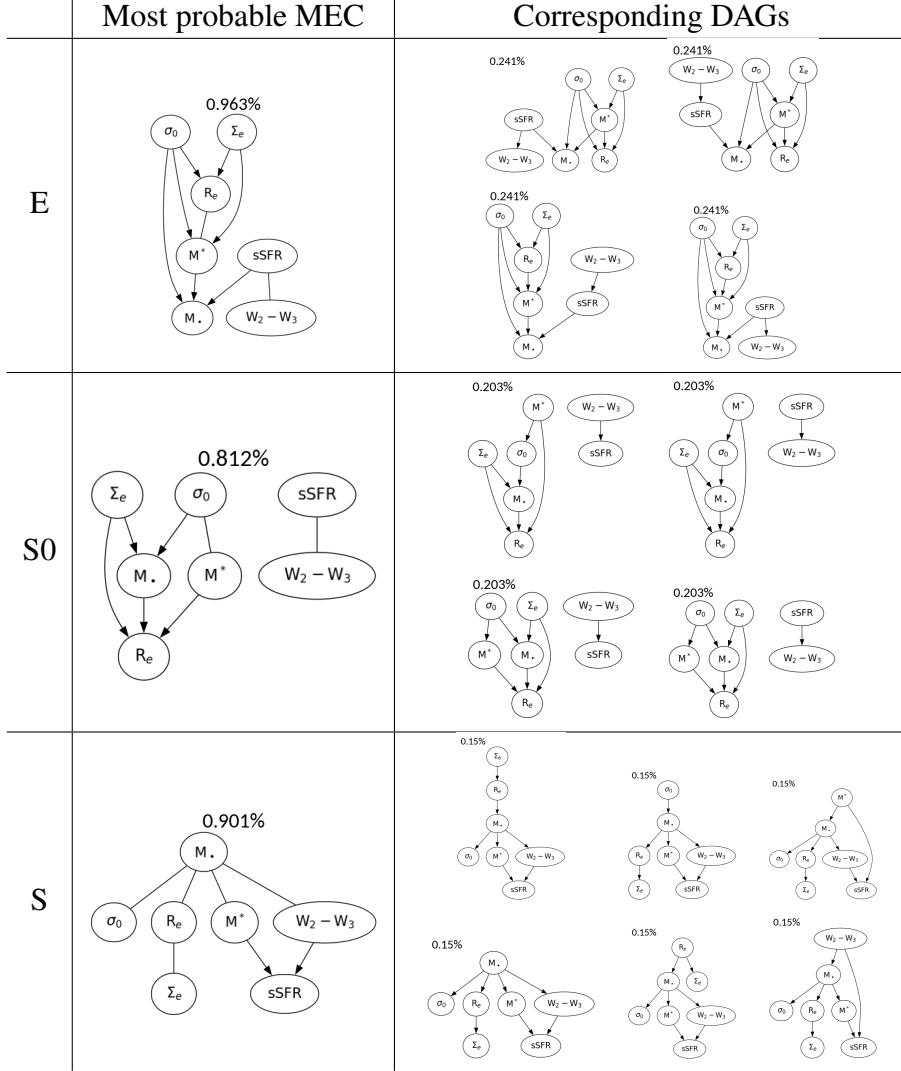


Figure 1: The most probable Markov Equivalence Class (MEC) for each morphology and their corresponding Directed Acyclic Graphs (DAGs). MECs are represented as Partially Directed Acyclic Graphs (PDAGs). Directed edges suggest the direction of causality. The undirected edge $A — B$ suggests both directions 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_* is strictly a child, while in spiral galaxies, M_* is *always* connected with four galaxy properties through four undirected edges, suggesting either M_* is the parent of all of the four galaxy properties, or M_* 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_* , 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}) (23). The MEC probabilities are the sum of their corresponding DAGs.

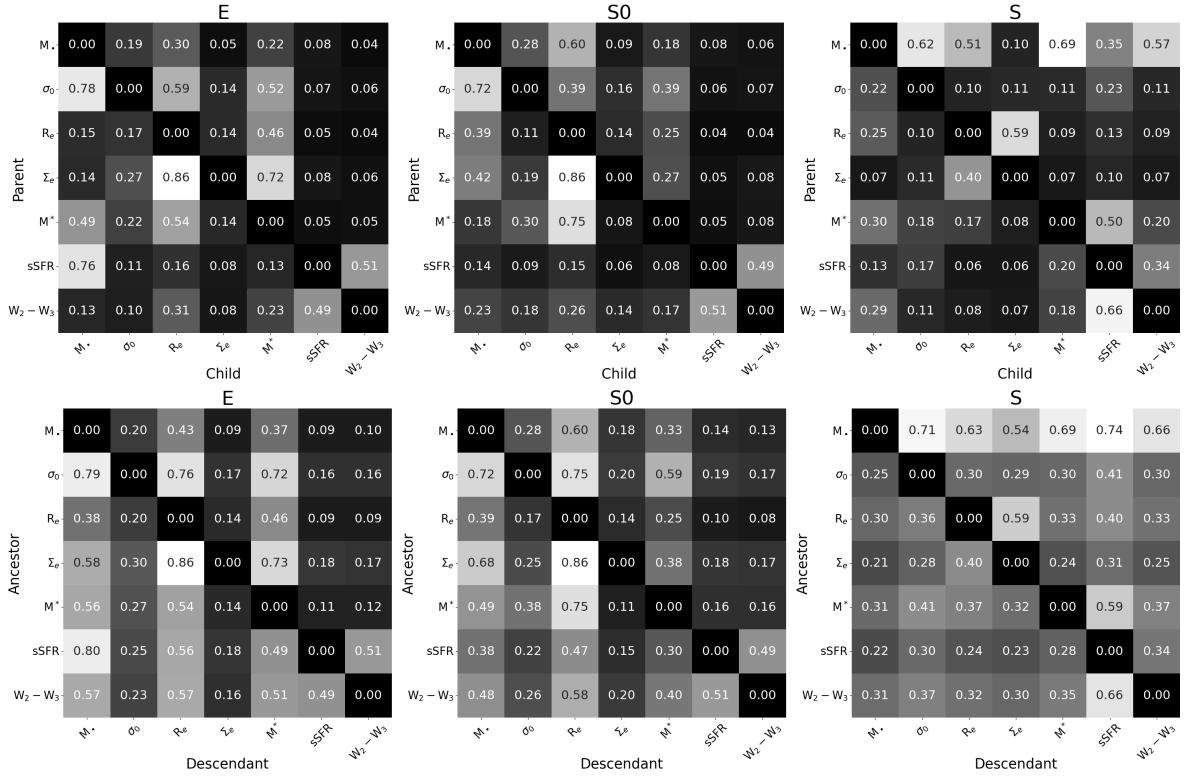


Figure 2: **Exact posterior edge marginals (top matrices) and path marginals (bottom matrices) for elliptical (left matrices), lenticular (middle matrices), and spiral (right matrices) galaxies.** Edge marginals give the probability of Parent \rightarrow Child through directed edges summed over all DAGs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

row 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. For comparison, the null results (i.e., the posterior from a uniform prior without any data) for the edge marginals are $P(\text{Parent}) = 29\%$, $P(\text{Child}) = 29\%$, and $P(\text{Disconnected}) = 42\%$; for the path marginals these probabilities are $P(\text{Ancestor}) = 42\%$, $P(\text{Descendant}) = 42\%$, and $P(\text{Disconnected}) = 16\%$ (see Fig. S2).

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 hydrodynamical cosmological simulations such as IllustrisTNG (27–31) show that elliptical galaxies with $\log(M^*/M_\odot) \geq 11$ 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% (32)³. 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 (35).

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 differently 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.⁴ 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% (32).

The six galaxy variables studied here can be split into the three parameters defining the fundamental plane (FP) of elliptical galaxies and three parameters related to star formation. The FP is a manifestation of dynamical equilibrium reached in the largely pressure-supported

³Here, Cannarozzo et al. (32) follow previous work (33, 34) and define a major merger as a stellar mass ratio greater than 1/4 between the two progenitors of a given galaxy.

⁴The coevolution of lenticular galaxies and their black holes is also strongly influenced by the presence of dust (36, 37).

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 our 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 (39).

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 $\langle \Sigma_e \rangle$ is the ancestor (86%) of R_e and that σ_0 is an ancestor (76%) of R_e . This implies $\langle \Sigma_e \rangle$ 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., $M \sim \sigma^2 R$).

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 (14, 35, 40).

We performed multiple verification tests of our results. We tested the same causal discovery approach on a set of semi-analytic models (SAMs) of galaxies with central black holes that have a clearly-defined (and customizable) causal direction for galaxy–SMBH coevolution (see more details in §**Methods & Materials**). In SAM galaxies, black hole feedback is actively affecting galaxies and is hard-coded to turn off as soon as a galaxy is quenched. Therefore, in SAM elliptical galaxies that become quenched, galaxy properties cause the black hole mass via the only remaining mechanism (i.e., mergers/accretion), and in SAM spiral galaxies with black hole feedback still on, black holes primarily cause galaxy properties. We conducted an additional check where the black hole feedback is manually turned off throughout the entire life of galaxies as “SAM no feedback” galaxies. The results we present in Fig. 3 indeed confirm the designed causal structure in the SAMs.

We also crosscheck our results with two alternative causal discovery methods, both *constraint*-

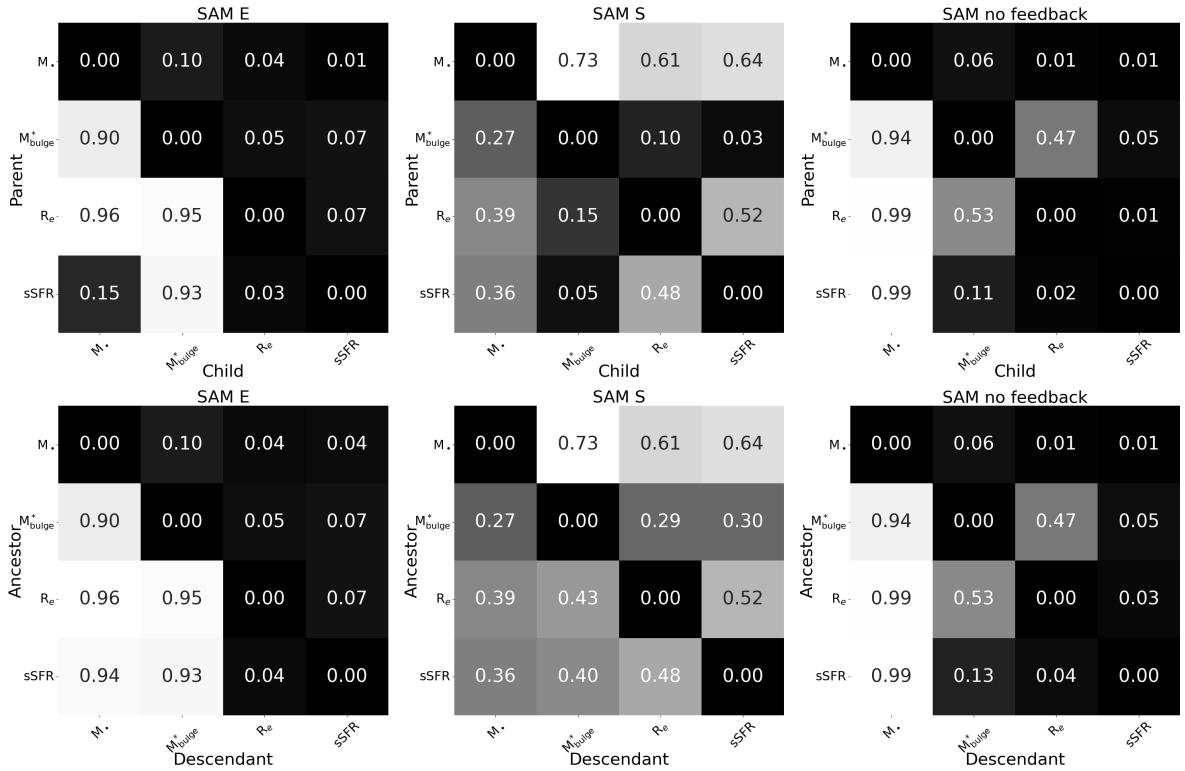


Figure 3: **Edge and path marginals for SAM galaxies.** These matrices are similar to those found in Fig. 2. Causal discovery is performed on ellipticals (SAM E), spirals (SAM S), and galaxies with black hole feedback intentionally turned off (SAM no feedback). Here, we are restricted to four parameters that are tracked in the SAMs.

based: the Peter-Clark, PC, (41) algorithm and the Fast Causal Inference, FCI, (42) 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. Details on these tests can be found in §**Methods & Materials**.

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 (43), built on GFlowNets (44, 45). We sampled the posterior by training a DAG-GFN, giving results consistent with the exact-posterior approach (see §**Methods & Materials**).

Further insights can be gained by using time-series data and control variables in galaxy simulations (46) 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 supermassive black holes 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: (*i*) quenching starts in gas rich (i.e., spiral) galaxies, and hence there is a causal connection; and (*ii*) 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.

References

1. S. Djorgovski, M. Davis, **313**, 59 (1987).
2. D. Makarov, P. Prugniel, N. Terekhova, H. Courtois, I. Vauglin, **570**, A13 (2014).
3. G. A. D. Savorgnan, A. W. Graham, **222**, 10 (2016).
4. B. L. Davis, A. W. Graham, E. Cameron, **873**, 85 (2019).
5. N. Sahu, A. W. Graham, B. L. Davis, **876**, 155 (2019).

6. A. W. Graham, N. Sahu, **520**, 1975 (2023).
7. E. L. Wright, *et al.*, **140**, 1868 (2010).
8. A. W. Graham, N. Scott, **764**, 151 (2013).
9. N. Scott, A. W. Graham, J. Schombert, **768**, 76 (2013).
10. G. Savorgnan, *et al.*, **434**, 387 (2013).
11. G. A. D. Savorgnan, A. W. Graham, A. Marconi, E. Sani, **817**, 21 (2016).
12. N. Sahu, A. W. Graham, B. L. Davis, **887**, 10 (2019).
13. N. Sahu, A. W. Graham, B. L. Davis, **903**, 97 (2020).
14. A. W. Graham, N. Sahu, **518**, 2177 (2023).
15. B. L. Davis, A. W. Graham, M. S. Seigar, **471**, 2187 (2017).
16. B. L. Davis, A. W. Graham, E. Cameron, **869**, 113 (2018).
17. B. L. Davis, A. W. Graham, F. Combes, **877**, 64 (2019).
18. B. L. Davis, Z. Jin, **956**, L22 (2023).
19. B. Davis, Z. Jin, *American Astronomical Society Meeting Abstracts* (2024), vol. 56 of *American Astronomical Society Meeting Abstracts*, p. 152.06.
20. L. Ferrarese, D. Merritt, **539**, L9 (2000).
21. K. Gebhardt, *et al.*, **539**, L13 (2000).
22. V. Springel, *et al.*, **435**, 629 (2005).
23. 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>.
24. D. Geiger, D. Heckerman, *Proceedings of the Tenth International Conference on Uncertainty in Artificial Intelligence*, UAI'94 (Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1994), p. 235–243.
25. D. Geiger, D. Heckerman, *Ann. Statist.* **30**, 1412 (2002).
26. J. Kuipers, G. Moffa, D. Heckerman, *arXiv e-prints* p. arXiv:1402.6863 (2014).
27. F. Marinacci, *et al.*, **480**, 5113 (2018).

28. J. P. Naiman, *et al.*, **477**, 1206 (2018).
29. D. Nelson, *et al.*, **475**, 624 (2018).
30. A. Pillepich, *et al.*, **475**, 648 (2018).
31. V. Springel, *et al.*, **475**, 676 (2018).
32. C. Cannarozzo, *et al.*, **520**, 5651 (2023).
33. V. Rodriguez-Gomez, *et al.*, **449**, 49 (2015).
34. V. Rodriguez-Gomez, *et al.*, **458**, 2371 (2016).
35. K. Jahnke, A. V. Macciò, **734**, 92 (2011).
36. A. W. Graham, **521**, 1023 (2023).
37. A. W. Graham, **531**, 230 (2024).
38. J. Mould, *Frontiers in Astronomy and Space Sciences* **7**, 21 (2020).
39. J. Falcón-Barroso, R. F. Peletier, M. Balcells, **335**, 741 (2002).
40. A. W. Graham, **518**, 6293 (2023).
41. P. Spirtes, C. N. Glymour, R. Scheines, *Causation, prediction, and search* (MIT press, 2000).
42. P. Spirtes, *Proceedings of the Eighth International Workshop on Artificial Intelligence and Statistics*, T. S. Richardson, T. S. Jaakkola, eds. (PMLR, 2001), vol. R3 of *Proceedings of Machine Learning Research*, pp. 278–285. Reissued by PMLR on 31 March 2021.
43. T. Deleu, *et al.*, *Conference on Uncertainty in Artificial Intelligence (UAI)* (2022).
44. E. Bengio, M. Jain, M. Korablyov, D. Precup, Y. Bengio, *Advances in Neural Information Processing Systems (NeurIPS)* (2021).
45. Y. Bengio, *et al.*, *Journal of Machine Learning Research (JMLR)* (2023).
46. S. Waterval, *et al.*, *arXiv e-prints* p. arXiv:2403.03985 (2024).
47. J. Magorrian, *et al.*, **115**, 2285 (1998).
48. N. H. Soliman, A. V. Macciò, M. Blank, **525**, 12 (2023).
49. T. Di Matteo, V. Springel, L. Hernquist, **433**, 604 (2005).

50. J. Kormendy, L. C. Ho, **51**, 511 (2013).
51. T. M. Heckman, P. N. Best, **52**, 589 (2014).
52. J. Silk, M. J. Rees, **331**, L1 (1998).
53. D. Sijacki, V. Springel, T. Di Matteo, L. Hernquist, **380**, 877 (2007).
54. T. Di Matteo, J. Colberg, V. Springel, L. Hernquist, D. Sijacki, **676**, 33 (2008).
55. J. Schaye, *et al.*, **402**, 1536 (2010).
56. M. Gaspari, M. Ruszkowski, S. P. Oh, **432**, 3401 (2013).
57. R. L. Larson, *et al.*, **953**, L29 (2023).
58. M. Pasquato, N. Matsiuk, *Research Notes of the American Astronomical Society* **3**, 179 (2019).
59. G. W. Imbens, T. Lemieux, *Journal of econometrics* **142**, 615 (2008).
60. J. P. Ostriker, J. Spitzer, Lyman, R. A. Chevalier, **176**, L51 (1972).
61. X. Pang, *et al.*, **923**, 20 (2021).
62. S. L. Ellison, *et al.*, **487**, 2491 (2019).
63. E. Greenwood, *Experimental sociology: A study in method* (Columbia University Press, 1945).
64. M. Catelan, F. R. Ferraro, R. T. Rood, **560**, 970 (2001).
65. A. F. L. Bluck, *et al.*, **659**, A160 (2022).
66. T. D. Gebhard, M. J. Bonse, S. P. Quanz, B. Schölkopf, **666**, A9 (2022).
67. B. Schölkopf, *et al.*, *Proceedings of the National Academy of Science* **113**, 7391 (2016).
68. J. Runge, *et al.*, *Nature communications* **10**, 2553 (2019).
69. G. Di Capua, *et al.*, *Weather and Climate Dynamics* **1**, 519 (2020).
70. R. Spekkens, Causal inference lecture - 230306 (2023). PIRSA:23030069 see, <https://pirsa.org>:
71. J.-M. A. Allen, J. Barrett, D. C. Horsman, C. M. Lee, R. W. Spekkens, *Phys. Rev. X* **7**, 031021 (2017).

72. M. S. Leifer, R. W. Spekkens, **88**, 052130 (2013).
73. C. J. Wood, R. W. Spekkens, *New Journal of Physics* **17**, 033002 (2015).
74. D. Janzing, *arXiv e-prints* p. arXiv:0708.3411 (2007).
75. A. E. Allahverdyan, D. Janzing, *Journal of Statistical Mechanics: Theory and Experiment* **2008**, 04001 (2008).
76. M. Pasquato, Z. Jin, P. Lemos, B. L. Davis, A. V. Macciò, *arXiv e-prints* p. arXiv:2311.15160 (2023).
77. J. Pearl, *Causality* (Cambridge university press, 2009).
78. J. Pearl, M. Glymour, N. P. Jewell, *Causal inference in statistics : a primer* (Wiley, Chichester, West Sussex, 2016).
79. A. F. Cheng, *et al.*, **157**, 104 (2018).
80. J. M. Mooij, T. Claassen, *Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence (UAI)*, J. Peters, D. Sontag, eds. (PMLR, 2020), vol. 124 of *Proceedings of Machine Learning Research*, pp. 1159–1168.
81. D. Chickering, *Journal of Machine Learning Research* **3**, 507 (2002).
82. B. Huang, K. Zhang, Y. Lin, B. Schölkopf, C. Glymour, *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '18 (Association for Computing Machinery, New York, NY, USA, 2018), p. 1551–1560.
83. D. Heckerman, D. Geiger, D. Chickering, *Machine Learning* **20**, 197 (1995).
84. D. Madigan, J. York, D. Allard, *International Statistical Review / Revue Internationale de Statistique* **63**, 215 (1995).
85. J. Viinikka, A. Hyttinen, J. Pensar, M. Koivisto, *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, H. Lin, eds. (Curran Associates, Inc., 2020), vol. 33, pp. 6584–6594.
86. C. C. Emezue, A. Drouin, T. Deleu, S. Bauer, Y. Bengio, *ICML 2023 Workshop on Structured Probabilistic Inference & Generative Modeling* (2023).
87. M. J. Vowels, N. C. Camgoz, R. Bowden, *ACM Comput. Surv.* **55** (2022).
88. G. Menegozzo, D. Dall'Alba, P. Fiorini, *2022 IEEE 18th International Conference on Automation Science and Engineering (CASE)* (2022), pp. 2124–2131.

89. O. Ahmed, *et al.*, Causalworld: A robotic manipulation benchmark for causal structure and transfer learning (2020).
90. D. Kalainathan, O. Goudet, R. Dutta, *Journal of Machine Learning Research* **21**, 1 (2020).
91. M. Scutari (2014).
92. A. W. Graham, T. H. Jarrett, M. E. Cluver, **527**, 10059 (2024).
93. T. H. Jarrett, *et al.*, **946**, 95 (2023).
94. M. E. Cluver, *et al.*, **850**, 68 (2017).
95. J. H. Jeans, *Astronomy and cosmogony* (1928).
96. E. P. Hubble, *Realm of the Nebulae* (1936).
97. A. W. Graham, **487**, 4995 (2019).
98. K. Sachs, O. Perez, D. Pe'er, D. A. Lauffenburger, G. P. Nolan, *Science* **308**, 523 (2005).
99. P. Erdos, A. Renyi, *Publications of the Mathematical Institute of the Hungarian Academy of Sciences* **5**, 17 (1960).
100. J. Bradbury, *et al.*, JAX: composable transformations of Python+NumPy programs (2018).
101. S. Warshall, *J. ACM* **9**, 11–12 (1962).
102. Y. Zheng, *et al.*, *arXiv e-prints* p. arXiv:2307.16405 (2023).
103. A. A. Hagberg, D. A. Schult, P. J. Swart, *Proceedings of the 7th Python in Science Conference*, G. Varoquaux, T. Vaught, J. Millman, eds. (Pasadena, CA USA, 2008), pp. 11 – 15.
104. Y. Luo, X. Kang, G. Kauffmann, J. Fu, **458**, 366 (2016).
105. J. Fu, *et al.*, **434**, 1531 (2013).
106. Q. Guo, *et al.*, **413**, 101 (2011).
107. Q. Guo, *et al.*, **428**, 1351 (2013).
108. Planck Collaboration, *et al.*, **641**, A6 (2020).
109. B. M. B. Henriques, *et al.*, **451**, 2663 (2015).
110. A. W. Graham, C. C. Worley, **388**, 1708 (2008).

111. H. Bondi, **112**, 195 (1952).
112. L. Wang, *et al.*, **454**, 83 (2015).
113. M. Blank, A. V. Macciò, A. A. Dutton, A. Obreja, **487**, 5476 (2019).
114. OEIS Foundation Inc., Number of essential graphs with n nodes (in 1-1 correspondence with markov equivalence classes of acyclic digraphs). (2024). Published electronically at <http://oeis.org>.

Acknowledgments

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). 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. **Author contributions:** Z.J., B.L.D., M.P., and T.D. contributed to the writing of the article. B.L.D. supplied the observational data. Y.L. and X.K. provided semi-analytical models. C.C. provided analysis related to the NIHAO simulations. Z.J., T.D., M.P., and Y.H. developed the methodology and analyzed the data. Y.H., L.P.-L., Y.B., and P.L. critically reviewed the methodology. Z.J., B.L.D., A.V.M., and M.P. interpreted the results. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The data and code used for this work are available for download from the following GitHub repository: <https://github.com/ZehaoJin/causalbh>.

Supplementary materials

- Materials and Methods
- Figs. S1 to S17
- Table S1
- References (47–114)

Materials and Methods

Prior work

The existence of correlations between the mass of central SMBHs and the properties of their host galaxies has long been observationally established (20, 21, 47) and reproduced by specific prescriptions in numerical simulations (48). However, unveiling the causal structure underpinning these correlations has remained an open problem: does galaxy evolution influence the growth of SMBHs by regulating accretion, or do SMBHs shape their host galaxies' properties via feedback (49–56)? With the advent of the *James Webb Space Telescope*, this debate has been reinvigorated by the detection of massive high-redshift quasars (57).

The few attempts at identifying causal relations in the astrophysical literature focus on two variables at a time or on estimating causal coefficients given a causal structure (causal inference). Pasquato & Matsiuk (58) used a regression discontinuity approach (59) to measure the causal effect of galactic disk-shocking (60) on open star cluster properties. A similar approach was followed by Pang et al. (61) to measure the causal effect of a supernova explosion in the Vela OB2 complex on star formation.

Ellison et al. (62) used a matching strategy to measure the causal effect of galaxy mergers on active galactic nuclei (AGNs) activity. Matching is a popular way of controlling for confounds in quasi-experimental data, where assignment to treatment is not determined at random (63). A precursor to matching in the astrophysical literature is the study of “second-parameter pairs” in globular clusters (64): globular clusters were matched based on metallicity and other properties, looking for the reason one member of the pair displayed a hot horizontal branch and its match would not.

Bluck et al. (65) utilized a Random Forest classifier to extract causal insights from observations to find the most predictive parameters associated with the quenching of star formation. Gebhard et al. (66) applied a denoising technique based on causal principles, half-sibling regression (67), to exoplanet imaging. In physics, outside of the context of astronomy and cosmology, causal techniques have found direct application in geophysics (68) and climate science (69), and have functioned as a basis for theoretical development in quantum foundations (70–73) and thermodynamics (74, 75).

Our work builds upon a preliminary pilot study to explore causal connections in galaxy–SMBH systems (76).

Causal inference and discovery

The seminal book *Causality* (77) introduced operational definitions for the presence of several types of causal relations between different variables.⁵ While these build on empirically observ-

⁵We refer the interested reader to Pearl et al. (78) and two online courses found at <https://www.bradyneal.com/causal-inference-course> and <https://www.bilibili.com/video/BV1sJ41177sg> for further information about causal inference.

able statistical dependencies between pairs of variables, they leverage the presence of additional variables to break the symmetry inherent in such associations. For instance, if variables X and Y are dependent when conditioning on any set of other variables S (that is, they are persistently associated) and there exists a third variable Z such that (conditional on some S) X and Z are independent while Z and Y are dependent (there is something else, independent of X , with which Y is associated), then X is dubbed a *potential cause of Y* . In addition to *potential cause*, *Causality* (77) also defines the notions of *genuine cause* and *spurious association*.

Basic Causal structures

In the case of causally-related variables, there are three basic causal structures that can be represented as DAGs: *chains*, *forks*, and *colliders* (Fig. S3).

- In the case of a *chain*, X causes (\rightarrow) Z , and Z causes (\rightarrow) Y . In a chain model, X and Y are *not* independent ($X \not\perp\!\!\!\perp Y$) without conditioning on Z . For example, consider three standing dominoes in order X , Z , and Y . The falling of X will cause Z to fall, which in turn will cause Y to fall. However, when we condition on Z , the other two variables, X and Y , will be independent ($X \perp\!\!\!\perp Y | Z$). In other words, if we let domino Z fall, the subsequent domino Y will fall regardless of whether the prior domino X fell or not.
- In the case of a *fork*, a single variable Z , called a confounder, causally influences two other variables X and Y . For instance, consider the influence of rainy weather (Z) on both umbrella sales (X) and the number of people jogging outside (Y). On rainy days, more umbrellas are likely to be sold, and less people will go out for a jog. In a fork model, without conditioning on the confounder Z , the other variables, X and Y , will be dependent on each other ($X \not\perp\!\!\!\perp Y$). If one were to analyze umbrella sales and jogging activity without considering the weather, they will find them to be dependent. However, once we condition on the confounder Z and compare days with the same weather condition, umbrella sales and jogging activity should be independent of each other ($X \perp\!\!\!\perp Y | Z$).
- A *collider* refers to the case that two variables, X and Y , independently cause a third variable Z . Consider the tossing of two fair coins X and Y , and a bell Z that rings whenever both coins lands on heads (this example is still valid when Z is a bell that rings whenever at least one of the coins lands on heads, see Pearl et al. (78) for a detailed Bayesian proof). Without revealing if the bell rings or not, the head/tail states of two coins are independent to each other ($X \perp\!\!\!\perp Y$)—simply as how coin tosses naturally works. However, if we condition on the bell Z not ringing, knowing one of the coins landed on heads immediately informs us that the other coin landed on tails ($X \not\perp\!\!\!\perp Y | Z$), otherwise the bell would have rung.

These three causal models each encode (conditional) independencies as discussed above and summarized in Fig. S3. Notice that chains and forks share the same (conditional) independencies, while colliders have a different set of (conditional) independencies. Chains and

forks are then considered as the same Markov Equivalence Class (MEC), while colliders belong to a different MEC. Note that these examples operate under the Markov assumption: $X \perp\!\!\!\perp_{\text{Graph}} Y | Z \Rightarrow X \perp\!\!\!\perp_{\text{Data}} Y | Z$, meaning that the (conditional) independencies encoded in a causal graph should appear in its data.⁶

Composite causal structures

In cases with more than three variables, such as in the right side of Fig. S3, variables can potentially be connected through multiple paths, with several chains, forks, or colliders. Following the (conditional) independencies encoded by chains, forks, and colliders, a path is *blocked* when conditioning on the middle variable of a chain or a fork and *unblocked* when not conditioning on the middle variable of a chain or a fork. Furthermore, a path is *blocked* when not conditioning on the middle variable of a collider and *unblocked* when conditioning on the middle variable of a collider. Two variables are defined to be *d-separated* if every path between them is blocked. Thus, *d-separated* variables are independent.⁷

In the Fig. S3, one will find $Z_2 \not\perp\!\!\!\perp X$ without any conditioning, since there is an unblocked chain path $Z_2 \rightarrow Z_3 \rightarrow X$. One should also find $Z_2 \perp\!\!\!\perp X | (Z_3, Z_1)$. Conditioning on Z_3 blocks the $Z_2 \rightarrow Z_3 \rightarrow X$ chain. Over the $Z_2 \rightarrow Z_3 \leftarrow Z_1 \rightarrow X$ path, although conditioning on Z_3 unblocks the $Z_2 \rightarrow Z_3 \leftarrow Z_2$ collider, the conditioning on Z_1 blocks the $Z_3 \leftarrow Z_1 \rightarrow X$ fork, making this path blocked. The remaining $Z_2 \rightarrow Y \leftarrow W \leftarrow X$ path is blocked by the collider $Z_2 \rightarrow Y \leftarrow W$ without conditioning on Y . Similarly, $Z_2 \not\perp\!\!\!\perp X | (Z_3, Z_1, Y)$.

Causal discovery from observational data

Causal discovery is most easily achieved through interventions. However, in observational fields such as astrophysics, interventions are rarely possible (79). In these cases, the field of causal discovery aims to reveal the causal relations between variables from purely observational data without interventions using alternative strategies.

Constraint-based methods. One of the most straightforward approaches to discovering causal structures from observational data is conducting conditional independence tests among variables since different MECs encode distinct conditional independencies.⁸ These approaches are generally referred to as *constraint-based methods*.

A commonly used constraint-based method is the Peter-Clark (41), PC, algorithm. The PC algorithm consists of three steps:

1. Start with a fully-connected, undirected graph among all variables, and remove edges based on conditional independence tests to arrive at a graph skeleton.

⁶Here, we are referring to the global Markov assumption, which is implied by the local Markov assumption. The local Markov assumption states that given its parents in a DAG, a node X is independent of all its non-descendants.

⁷With the exception of intransitive cases (78).

⁸Here, we adopt the faithfulness assumption, or the converse of the Markov assumption: $X \perp\!\!\!\perp_{\text{Data}} Y | Z \Rightarrow X \perp\!\!\!\perp_{\text{Graph}} Y | Z$.

2. Identify colliders with conditional independence tests and orient them.
3. Orient edges that are incident on colliders such that no new colliders will be constructed.

Note that in addition to the Markov assumption and faithfulness assumption, the PC algorithm further assumes causal sufficiency (i.e., no unobserved confounders) and acyclicity.

Another constraint-based method, the Fast Causal Inference (42), FCI, algorithm relaxes the assumption of causal sufficiency, allowing unobserved confounders. The FCI algorithm is based on the same independence testing procedure as the PC algorithm but differs at the stage of labeling and orienting edges.

It has also been proven that PC and FCI algorithms are sound and complete under cyclic settings (80). However, these methods only provide a point estimate for the true MEC without quantifying their uncertainties; this becomes particularly problematic when the number of data points is small and the reliability of conditional independence tests degrades.

Score-based methods. Instead of finding a single causal structure, we can adopt a Bayesian perspective and define a posterior over all possible DAGs, $P(G | D)$. To do this, *score-based methods* assign a numerical score to every DAG given the data. There are several possible ways to define such a score, such as the Bayesian Information Criterion (BIC) score (81), generalized score (82), and the Bayesian Gaussian equivalent (BGe) score (24–26).⁹

In an exact posterior approach, one evaluates the chosen score for every possible DAG, i.e., $P(G | D)$. However, the cost of an exact search grows super-exponentially as the number of variables (nodes) increases. For example, the number of possible DAGs for three variables is only 25 but exceeds 4×10^{18} for ten variables (see the online encyclopedia of integer sequences (23) for the number of DAGs corresponding to n nodes), making an exact posterior search quickly computationally intractable. For the current study, the total number of DAGs for seven variables is 1,138,779,265, which is at the limit of computational feasibility.

As a result, sampling algorithms have been developed to approximate the exact posterior distribution without going over all DAGs. Such approximation is often done with Markov Chain Monte Carlo (MCMC) methods, such as the MC3 algorithm (84) and Gadget (85). More recently, Deleu et al. (43) developed DAG–GFN as an alternative to MCMC, and showed that DAG–GFN compares favorably against other methods based on MCMC. Here, we also show that DAG–GFN does recover the exact posterior distribution fairly well under the SMBH–galaxy context in §Estimating exact posteriors with DAG–GFN. Some benchmark studies on causal discovery algorithms can be found in (68, 86–91).

Data

Our data is composed of a current state-of-the-art sample of galaxies with dynamically-measured SMBH masses. These include 145 nearby galaxies with a median luminosity distance of

⁹The scores listed above are for continuous data. The Bayesian Dirichlet equivalent (BDe) score (83) is one of the scores for discrete or categorical data.

21.5 Mpc that host SMBHs with directly resolved spheres of influence. From this parent sample, 101 out of 145 galaxies have all of the seven desired variables of interest for our study. We include the three parameters that define the fundamental plane of elliptical galaxies (*I*), i.e., central stellar velocity dispersion (σ_0), effective (half-light) radius of the spheroid¹⁰ (R_e), and the average projected density within R_e ($\langle \Sigma_e \rangle$). M_\bullet values are compiled by a series of progressive studies on black hole mass scaling relations (4, 5, 8–19). σ_0 values are collected from several works (12, 15, 17), which are obtained primarily from the HyperLeda database (2) and homogenized to produce an estimate of the mean velocity dispersion within an aperture of 595 pc. R_e and $\langle \Sigma_e \rangle$ measurements were produced via the multi-component decompositions of surface brightness light profiles (primarily of 3.6 μm *Spitzer* Space Telescope imaging) from succeeding works (3–6).¹¹

This choice of parameters also allows us to explore the well-known $M_\bullet - \sigma_0$ relation (20, 21). Indeed, one impetus for this study of SMBH–galaxy causality was the significant difference in intrinsic scatter (ϵ) observed in elliptical galaxies ($\epsilon = 0.31$ dex) vs. spiral galaxies ($\epsilon = 0.67$ dex) as determined by Sahu et al. (12). This implies that the $M_\bullet - \sigma_0$ relation is ≈ 2.3 times less accurate for predicting SMBH masses in spiral galaxies as opposed to elliptical galaxies. As shown in this work, this difference in the scatter of the relationship between morphological types foreshadows their inherent dichotomy in causal structures.

The remaining parameters we explore concern properties related to the star-formation rate (SFR) in galaxies. For this, we consider data from the Wide-field Infrared Survey Explorer (7), WISE, to provide the color, total stellar masses (M^*), and SFRs for our galaxies. These WISE values are all compiled from Graham et al. (92): M^* is derived from the prescriptions of Jarrett et al. (93) for W1 (3.4 μm) photometry and colors from WISE; and SFRs accounting for activity over the past 100 Myr is assessed via the WISE total integrated fluxes as per the calibrations of Cluver et al. (94). For WISE colors, we considered both W1–W2 (3.4 μm – 4.6 μm) and W2–W3 (4.6 μm – 12.1 μm) colors, but ultimately elected to conduct our analyses with only the latter color, which exhibits a greater range of diversity across morphological classes of galaxies. Rather than absolute SFR, we convert to specific star-formation rate (sSFR) by normalizing each SFR by the stellar mass of each galaxy (i.e., $\text{sSFR} \equiv \text{SFR}/M^*$).

We split the sample of galaxies into three morphological classes:

- highly-evolved, massive, gas-poor elliptical (E) galaxies, which have been exposed to the full range of feedback and merging processes throughout their long histories spanning large fractions of the age of the Universe,
- spiral (S) galaxies, at the opposite end of galaxy morphological classification schemes

¹⁰Here, 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.

¹¹ R_e (and $\langle \Sigma_e \rangle$) is calculated from the equivalent (i.e., geometric-mean) axis surface brightness profile of each galaxy. Radii computed along the equivalent axis of quasi-elliptical isophotes are equal to \sqrt{ab} , where a is the semi-major axis and b is the semi-minor axis of an isophote, and thus produce a circle with the *equivalent* area as the quasi-elliptical isophote.

(95–97), which are unlikely to have encountered any major mergers and still retain a large fraction of their gas, and

- lenticular (S0) galaxies, which represent a bridging population between E and S types.

Altogether, this gives us a sample of 35 elliptical, 38 lenticular, and 28 spiral galaxies for a total of 101 galaxies, each with six physical measurements of the host galaxy plus a dynamically-measured SMBH mass (see Table S1 and its pairplot in Fig. S1). All morphologies have been determined by the multi-component decompositions of surface brightness light profiles (3–6). Our general classification scheme defines elliptical galaxies as spheroids (with or without embedded disk components), lenticular galaxies as spheroids with extended disk components (without spiral structure), and spiral galaxies as disk galaxies (with classical bulges, pseudobulges, or no bulges) with spiral structure. For our purposes in this study, we have not considered barred morphologies as a distinct classification element.

As can be seen from the shape of the data in Fig. S1, the data is predominantly characterized by linear relations and appears normally distributed in their logarithmic form, which underpins the general assumption for calculating BGe scores¹².

Exact posterior methodology

We calculate the exact posteriors $P(G | D)$ of the BGe scores for all DAGs from seven variables, separately for elliptical, lenticular, and spiral galaxies. The posterior probability of a graph given the data $P(G | D)$ is proportional to the posterior probability of the data given a graph $P(D | G)$ under a uniform prior, through Bayes' rule $P(G | D) \propto P(D | G)P(G)$. Under the assumption of linear and Gaussian data, the BGe score gives the marginal likelihood that the distribution of the data sample $d = \{x_1, \dots, x_N\}$ of N variables is faithful (i.e., the data satisfies only and all the conditional independencies encoded by the DAG) to a hypothetical DAG model m^h as a product of local scores:

$$p(d |) = \prod_{i=1}^n \frac{p(d_i \cup \{X_i\} |)}{p(d_i |)}, \quad (1)$$

where i is the parent variables of the vertex i , and d^Y is the data restricted to the subset of data Y . The modularity (i.e., the full score is a product of local scores over all vertices i) of local scores ensures that all DAGs in the same MEC are scored equally, and simplifies the posterior calculation over a large amount of DAGs. The local scores are further characterized by

$$p(d |) = \left(\frac{N}{N+} \right)^{l/2} \frac{\Gamma_l((N + -n + l)/2)}{\pi^{lN/2} \Gamma_l((-n + l)/2)} \frac{|T|^{(-n+l)/2}}{|R|^{(N+-n+l)/2}}. \quad (2)$$

¹²The BGe score has also been shown to perform well in real-world scenarios where the data is not strictly linear and Gaussian, for example, in benchmark studies (43, 86) using real flow cytometry data (98), as well as simulated noisy data following an Erdős-Rényi model (99).

The detailed explanation and full derivation of Equations 1 and 2 can be found in Kuipers et al. (26). Empirically, many causal discovery methods based on the BGe score have been proven to successfully recover the ground truth causal structures in benchmark tests (43, 86).

The steps of calculating exact posteriors can be summarized in the following steps:

1. Generate all possible DAGs for N variables represented by $N \times N$ adjacency matrices A , with $A_{i,j} = 1$ if there is an arrow from node i to node j .
2. For every DAG, generate its transitive closure represented by an adjacency matrix.
3. Calculate the posterior probability for every DAG given the data with the BGe score following Equation 1 (the sum of the scores over all DAGs is equal to unity by construction).
4. Perform a weighted average on all DAG adjacency matrices according to their posterior probabilities to get the matrix of edge marginals.
5. Perform a weighted average on all transitive closure adjacency matrices according to their posterior probabilities to get the matrix of path marginals.

For a given value of N , steps 1 and 2 only need to be done once (i.e., the possible DAGs for N variables are unique), and only steps 3–5 need to be repeated for different datasets.

In this work, steps 2–5 are coded in a highly optimized and parallelized way on graphics processing units (GPUs), powered by a Python package JAX (100). The calculation of transitive closure adopts Warshall’s algorithm (101). The MECs for analysis are generated with a Python package causal-learn (102). The visualization of causal graphs is made possible through Python packages NetworkX (103) and PyGraphviz¹³.

Posterior distribution inspection

In addition to Fig. 1 (which shows the MEC with the highest posterior probability along with its corresponding DAGs) and Fig. 2 (which shows the edge and path marginals), here we take a deeper look at the posterior distribution. Fig. S4 shows the top four (in terms of posterior probabilities) MECs and Fig. S5 shows the top ten DAGs. The top graphs within each morphology class are similar to each other, and most of them convey the idea that elliptical galaxy properties → SMBH mass, in spirals SMBH mass → galaxy properties, and lenticulars occupy the middle ground. The paltry percentage of the total population for individual DAGs or MECs is not a rare and surprising phenomenon; due to the huge space of possible causal structures, the number of possible DAGs grows *super-exponentially* as the number of variables increases. The chance of drawing any DAG from a uniform distribution out of all possible DAGs is $8.781333053161975 \times 10^{-10}$, which is $\sim 10^6$ times smaller than the typical proportion around $\sim 10^{-3}$ we observed for the top DAGs (see Fig. S5).

¹³<https://pygraphviz.github.io/>

To better understand the relative posterior probability distribution and quantify the difference between graphs, in Fig. S6 we ordered DAGs and MECs by their posterior probabilities from highest to lowest. The posterior probability is shown as solid lines and labeled on the left y -axis, and the rapidly dropping curves show that a few leading DAGs and MECs are relatively much more probable than the DAGs and MECs in the long tail (note that the x -axis is in a logarithmic scale). The dashed lines and the right y -axis show the *structural Hamming distance* (SHD), a standard metric to evaluate the distance between graphs¹⁴, from each unique DAG or MEC to the most probable DAG or MEC. From the SHDs, the top few DAGs or MECs are more similar to each other with fewer edges away from each other, and less prominent DAGs or MECs are statistically more and more distinctive from the top ones. The SHD increases the slowest in spirals, making the posterior distribution of spirals the most unimodal, while the SHD grows fastest in lenticulars, reinforcing the picture that lenticulars, as the middle ground between ellipticals and spirals, have more sub-modes of causal structures and no clear dominance of one particular causal direction between black hole mass and galaxy properties. The probability distribution and SHD distribution together shows that despite the existence of some sub-modes, a general mode of causal structure is detected in ellipticals and spirals respectively, and this general mode can be visualized statistically via the edge marginals and path marginals in Fig. 2, as discussed in §Main.

Semi-analytical models (SAMs)

As an additional benchmark test for the methodology under the SMBH–galaxy context, we practiced the same causal discovery method on data generated by SAMs, where the ground truth causal direction is clearly defined, propagated through a series of coupled partial-differential equations, and is easily customizable. The result of this test has been presented in paragraphs related to Fig. 3 in §Main. Here we provide some details on the SAM dataset itself.

SAMs are powerful tools to model galaxy formation using dark matter halo merger trees from N -body simulations with some phenomenological descriptions of baryonic physical processes like cosmic reionization, hot gas cooling and cold gas infall, star formation and metal production, supernova feedback, gas stripping and tidal disruption of satellites, galaxy mergers, bulge formation, black hole growth, AGN feedback, etc. We adopt the model of Luo et al. (104), which is the resolution-independent version of the Munich galaxy formation model: L-Galaxies (mainly based on models of Fu et al. (105) and Guo et al. (106, 107)). The dark matter only N -body simulation is the JiuTian-1G simulation with 6144^3 dark matter particles in a $1 \text{ Gpc}/h$ cubic simulation box, based on Planck 2018 (108) cosmological parameters. In the model, there are two processes related to black hole growth and its feedback. The first is “quasar mode” where SMBHs can accrete cold gas directly during galaxy mergers. The other is “radio mode” where SMBHs can accrete hot gas continually from their host galaxies and inject energy into the hot atmosphere. The quasar mode is the main black hole growth channel, while

¹⁴ $\text{SHD}(G, G^*)$ counts the number of edges needed to add, delete, or revert, to move from one graph G to another graph G^* .

the radio mode is the main AGN feedback channel to suppress hot gas cooling. More details can be found in the supplementary material of Henriques et al. (109).

We use three groups of SAM galaxies: SAM E galaxies, SAM S galaxies, and SAM no-feedback galaxies. SAM E galaxies are galaxies with $B/T > 0.78$ (bulge-to-total ratio, $M_{\text{bulge}}^*/M^* > 0.78$ (110)), and SAM S galaxies are galaxies with $B/T \leq 0.78$. Additional stellar mass M^* cuts are applied such that the M^* distributions of SAM E and SAM S galaxies are similar to that of the real observational data used in this work for a fair comparison, as shown in Fig. S7. No cuts are applied to SAM no-feedback galaxies since they do not have any realistic counterparts and are generated solely for this test. This gives us 1189 SAM E galaxies, 1999 SAM S galaxies, and 2663 SAM no-feedback galaxies.

Verification with PC and FCI

The PC and FCI algorithms, two *constraint-based methods* (in contrast with the *score-based method* adopted in this work), are also applied to the same observational data to cross-check our results. The details of these two time-tested algorithms are presented in §**Causal discovery from observational data**. We adopt the implementation of PC and FCI in the `Python` package `causal-learn` (102), and the results are reported in Fig. S8. The exact posterior result including edge/path marginals (Fig. 2) and the top MECs/DAGs (Fig. 1, Fig. S4, and Fig. S5) are generally consistent with the causal graphs found by PC and FCI.

In ellipticals, the PC algorithm finds σ_0 and sSFR cause M_\bullet . In our Bayesian approach, $\sigma_0 \rightarrow M_\bullet$ and sSFR $\rightarrow M_\bullet$ indeed have the highest and second highest edge/path marginals among the potential causes of M_\bullet in ellipticals. In spirals, the PC algorithm finds σ_0 , R_e , and M^* as effects of M_\bullet , and this is also consistent with the edge/path marginals reported in our Bayesian approach. The FCI algorithm produces results compatible with those of the PC algorithm, with the difference that, without the assumption of causal sufficiency, it leaves open the possibility that all the relations between SMBH mass and its causal parents are confounded by unobserved variables. Particularly, the double arrow between M_\bullet and sSFR in the lower left DAG of Fig. S8 may indicate an unobserved confounder, possibly the gas fraction, which in the future can be tested through hydrodynamical simulations where the gas fraction is more accessible.

Note that here we adopt a relatively high value of $\alpha = 0.15$, the significance cutoff for the p -value of conditional independence tests. Generally, a lower α value gives more false negative errors (i.e., fails to identify causal relations that exist), and a higher α results in more false positive errors (i.e., identifies causal relations that do not exist). Practically, the choice of α is often empirical and highly depends on the context. Here in our case, the conditional independence tests, which are the core of PC and FCI, suffer from the limited size of the dataset (35, 38, and 28 for E, S0, and S galaxies, respectively). We therefore selected a higher value of α to mitigate this. These limitations of PC and FCI are one of the main motivations for our adoption of a Bayesian approach by relatively comparing the posterior probabilities across all possible DAGs.

Possible unobserved confounders

Our posterior calculation approach implicitly adopts the assumption of causal sufficiency, i.e., assuming there are no unobserved confounders¹⁵. 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 our targets¹⁶. The edge and path marginals with distance included are presented in Fig. S9. 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.

Stability under observational errors

The variables used in this work are affected by observational errors and their marginal posterior probability distributions are given in Table S1 (assuming Gaussian posteriors). However, the causal structures explored so far have been calculated for the mean of these posteriors without considering their uncertainties. We now quantify the effect of this uncertainty on our inference of the causal structures. To do this, we draw samples from the posterior distribution of each variable to produce 100 mock datasets. The causal discovery method outlined in this article is repeated on each of these 100 randomly-sampled datasets to arrive at 100 pairs of different edge marginal and path marginal matrices for each of the three morphological types considered. The edge marginal and path marginal matrices are summarized in Fig. S10.

We find that, overall, the key findings of this study are robust against these uncertainties. For example, in ellipticals, the edge marginal between σ_0 and M_\bullet in both directions across random sampling realizations are 0.70 ± 0.09 and 0.26 ± 0.07 , giving a 3.84σ discrepancy (in other words, the probability that the inferred direction of causality is due to noise and the resulting uncertainties in the variables is about 10^{-4}). Fig. S11 shows the distribution of the edge marginals and path marginals for $\sigma_0 \rightarrow M_\bullet$. The difference between ellipticals and spirals is evident for all realizations.

¹⁵An unobserved confounder is a variable that is not included in the analysis but is a cause of two or more variables of interest.

¹⁶Our luminosity distances are adopted from Graham & Sahu (14). Indeed, this sample of dynamically-measured black hole masses comes from galaxies that are all in the local Universe (median $D_L = 19.3$ Mpc; $z = 0.00439$ (108)).

Stability under possible outliers

We also explored the possibility of individual outlier galaxies biasing the inferred causal relations. To do this, we performed leave-one-out cross-validation. For ellipticals, lenticulars, and spirals, respectively, one galaxy is taken out at a time, and causal discovery is performed repeatedly (e.g., for 35 elliptical galaxies this procedure will be repeated 35 times). The mean and standard deviation of the resulting marginals are shown in Fig. S12, and the marginals for $\sigma_0 \rightarrow M_\bullet$ are highlighted in Fig. S13. As can be seen, the fluctuations due to leave-one-out are much smaller than the uncertainties resulting from observational errors, suggesting that the results are not driven by any potential individual outlier galaxy.

Estimating posteriors with DAG–GFN

The general timescale to perform the exact posterior search for seven variables outlined in this work, including generating all possible DAGs, transitive closures, calculating posterior probabilities, and getting edge/path marginals, is approximately a few hours. However, as the number of possible DAGs grows by a factor of $\sim 10^2$ when the number of variables increases from seven to eight, and by a factor of $\sim 10^6$ when the number of variables increases from seven to nine, an exact search becomes impractical. Here, we explore DAG–GFN as a feasible way to approximate the posteriors as the number of variables increases.

The DAG–GFN method (43) uses the framework of Generative Flow Networks (44, 45), GFlowNets, to (approximately) sample from the posterior distribution. GFlowNets treat the problem of sampling from an unnormalized distribution over discrete and compositional objects as a sequential decision-making problem, where actions are taken by sampling from a learned policy at each step of generation. In the context of (Bayesian) causal discovery, DAGs are constructed one edge at a time, starting from the empty graph. The objective is to find a policy $\pi(G' | G)$ giving the probability of adding an edge to the DAG G to transform it into a new graph G' , such that sampling sequentially from it would yield samples from a distribution proportional to $R(G)$ (i.e., an unnormalised distribution). Deleu et al. (43) showed that such a policy satisfies

$$\frac{1}{|G| + 1} R(G')\pi(\text{stop} | G) = R(G)\pi(G' | G)\pi(\text{stop} | G'), \quad (3)$$

where $|G|$ is the number of edges in G , and $\pi(\text{stop} | G)$ is the probability of stopping the sequential process, effectively returning G as a sample of the posterior. To sample the posterior $P(G | D) \propto P(D | G)P(G)$ (by Bayes' rule), we can then use $R(G) = P(D | G)P(G)$.

DAG–GFN was trained on our data, and 10^5 DAGs were sampled from the trained network. The frequency of each sampled unique DAG gives the approximated posterior probability of that DAG. The marginals, as well as the top MECs and DAGs are presented in Extended Data Figs. 14–16, respectively. The approximated posteriors by DAG–GFN are highly consistent with the exact posteriors from our primary analysis. Visual inspection reveals that Fig. 2 and Fig. S14 present noticeable similarities.

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 → 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 (111). This latter quantity is found to be relatively constant in gas-rich galaxies, as confirmed by modern numerical simulations, like the NIHAO suite (112, 113) as shown in Fig. S17. This implies that accretion is fairly constant in all gas-rich galaxies, diminishing the causal relation galaxy → black hole.

Therefore, in spiral galaxies, the causal relation of galaxy → black hole is expected to be very weak compared to the black hole → 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 → galaxy direction is negligible compared to the galaxy → 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. This may be one of the reasons why we see many sub-modes in the posterior distributions of lenticulars as shown in Fig. S6. 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 humanity¹⁷, studies of samples of galaxies with different ages may provide observational clues about the presence or absence of cyclicity in future studies.

| Galaxy | $\log(M_\bullet)$ [M_\odot] | $\log(\sigma_0)$ [km s^{-1}] | $\log(R_e)$ [kpc] | $\log(\langle \Sigma_e \rangle)$ [$M_\odot \text{ pc}^{-2}$] | W2–W3 [mag] | $\log(M^*)$ [M_\odot] | $\log(\text{sSFR})$ [yr^{-1}] |
|------------------------|------------------------------------|--|----------------------|---|------------------|------------------------------|---|
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| 35 Elliptical Galaxies | | | | | | | |
| IC 1459 | 9.38 ± 0.18 | 2.47 ± 0.01 | 0.89 ± 0.09 | 2.89 ± 0.13 | 0.39 ± 0.06 | 11.24 ± 0.08 | -12.00 ± 0.10 |
| IC 4296 | 9.10 ± 0.09 | 2.52 ± 0.01 | 0.96 ± 0.31 | 2.87 ± 0.09 | 0.02 ± 0.08 | 11.47 ± 0.08 | -12.47 ± 0.11 |
| NGC 821 | 7.59 ± 0.19 | 2.30 ± 0.01 | 0.54 ± 0.01 | 2.91 ± 0.09 | 0.27 ± 0.13 | 10.64 ± 0.08 | -11.65 ± 0.11 |
| NGC 1275 | 8.90 ± 0.24 | 2.39 ± 0.02 | 1.24 ± 0.31 | 2.51 ± 0.13 | 3.00 ± 0.04 | 11.52 ± 0.09 | -9.81 ± 0.11 |
| NGC 1399 | 8.67 ± 0.06 | 2.52 ± 0.01 | 0.76 ± 0.09 | 3.01 ± 0.09 | 0.17 ± 0.07 | 11.23 ± 0.08 | -12.72 ± 0.13 |
| NGC 1407 | 9.65 ± 0.06 | 2.42 ± 0.01 | 0.80 ± 0.31 | 3.03 ± 0.12 | 0.07 ± 0.09 | 11.39 ± 0.08 | -15.69 ± 0.44 |
| NGC 1600 | 10.28 ± 0.04 | 2.52 ± 0.01 | 1.22 ± 0.09 | 2.66 ± 0.07 | -0.34 ± 0.10 | 11.71 ± 0.09 | -15.23 ± 0.44 |
| NGC 3091 | 9.61 ± 0.02 | 2.49 ± 0.01 | 1.15 ± 0.09 | 2.61 ± 0.17 | -0.21 ± 0.11 | 11.39 ± 0.09 | -15.09 ± 0.44 |
| NGC 3377 | 8.24 ± 0.23 | 2.13 ± 0.01 | 0.36 ± 0.01 | 2.77 ± 0.07 | -0.09 ± 0.08 | 10.13 ± 0.08 | -14.43 ± 0.44 |
| NGC 3379 | 8.63 ± 0.11 | 2.31 ± 0.00 | 0.43 ± 0.31 | 3.19 ± 0.14 | 0.12 ± 0.05 | 10.64 ± 0.08 | -14.94 ± 0.44 |

¹⁷Except in simulations, which we will investigate in future work.

| 38 Lenticular Galaxies | | | | | | | |
|------------------------|--------------|-------------|--------------|-------------|--------------|--------------|---------------|
| NGC 3414 | 8.38 ± 0.05 | 2.38 ± 0.01 | 0.47 ± 0.09 | 3.07 ± 0.15 | 0.50 ± 0.07 | 10.63 ± 0.08 | -12.05 ± 0.10 |
| NGC 3585 | 8.49 ± 0.14 | 2.33 ± 0.01 | 0.90 ± 0.31 | 2.67 ± 0.10 | 0.06 ± 0.09 | 11.03 ± 0.09 | -13.16 ± 0.28 |
| NGC 3607 | 8.17 ± 0.17 | 2.35 ± 0.01 | 0.90 ± 0.31 | 2.74 ± 0.13 | 0.73 ± 0.06 | 11.13 ± 0.08 | -11.73 ± 0.09 |
| NGC 3608 | 8.63 ± 0.10 | 2.29 ± 0.01 | 0.66 ± 0.31 | 2.75 ± 0.09 | -0.13 ± 0.10 | 10.56 ± 0.09 | -14.86 ± 0.44 |
| NGC 3842 | 9.94 ± 0.12 | 2.49 ± 0.01 | 1.48 ± 0.09 | 2.03 ± 0.08 | -0.43 ± 0.07 | 11.45 ± 0.09 | -14.80 ± 0.44 |
| NGC 3923 | 9.47 ± 0.13 | 2.39 ± 0.01 | 0.92 ± 0.09 | 2.80 ± 0.13 | -0.03 ± 0.08 | 11.30 ± 0.08 | -15.60 ± 0.44 |
| NGC 4261 | 9.21 ± 0.08 | 2.47 ± 0.01 | 0.84 ± 0.31 | 2.89 ± 0.10 | 0.22 ± 0.08 | 11.17 ± 0.08 | -11.99 ± 0.10 |
| NGC 4291 | 8.97 ± 0.14 | 2.47 ± 0.01 | 0.27 ± 0.31 | 3.35 ± 0.14 | 0.00 ± 0.09 | 10.47 ± 0.08 | -14.77 ± 0.44 |
| NGC 4374 | 8.95 ± 0.04 | 2.44 ± 0.00 | 1.04 ± 0.31 | 2.59 ± 0.08 | -0.04 ± 0.07 | 11.14 ± 0.08 | -15.44 ± 0.44 |
| NGC 4472 | 9.36 ± 0.03 | 2.45 ± 0.00 | 1.01 ± 0.09 | 2.81 ± 0.08 | 0.28 ± 0.12 | 11.41 ± 0.08 | -12.08 ± 0.10 |
| NGC 4473 | 7.95 ± 0.22 | 2.25 ± 0.01 | 0.43 ± 0.31 | 2.96 ± 0.08 | 0.15 ± 0.08 | 10.53 ± 0.08 | -12.05 ± 0.11 |
| NGC 4486 | 9.85 ± 0.02 | 2.51 ± 0.01 | 0.85 ± 0.31 | 3.05 ± 0.08 | 0.33 ± 0.05 | 11.31 ± 0.08 | -12.02 ± 0.10 |
| NGC 4552 | 8.67 ± 0.04 | 2.40 ± 0.01 | 0.71 ± 0.31 | 2.68 ± 0.09 | 0.50 ± 0.10 | 10.77 ± 0.09 | -12.12 ± 0.11 |
| NGC 4621 | 8.59 ± 0.04 | 2.36 ± 0.01 | 0.88 ± 0.09 | 2.58 ± 0.10 | 0.42 ± 0.13 | 10.89 ± 0.08 | -11.82 ± 0.10 |
| NGC 4649 | 9.67 ± 0.10 | 2.52 ± 0.01 | 0.80 ± 0.09 | 3.04 ± 0.09 | 0.42 ± 0.10 | 11.24 ± 0.08 | -12.02 ± 0.09 |
| NGC 4697 | 8.10 ± 0.02 | 2.22 ± 0.00 | 1.09 ± 0.40 | 2.03 ± 0.08 | 0.09 ± 0.06 | 10.65 ± 0.08 | -12.11 ± 0.11 |
| NGC 4889 | 10.29 ± 0.33 | 2.59 ± 0.01 | 1.43 ± 0.09 | 2.38 ± 0.09 | -0.17 ± 0.09 | 11.72 ± 0.09 | -15.02 ± 0.44 |
| NGC 5077 | 8.85 ± 0.23 | 2.40 ± 0.01 | 0.64 ± 0.09 | 3.16 ± 0.17 | 0.22 ± 0.07 | 11.02 ± 0.08 | -11.94 ± 0.10 |
| NGC 5419 | 9.86 ± 0.14 | 2.54 ± 0.01 | 1.01 ± 0.01 | 2.87 ± 0.09 | 0.04 ± 0.12 | 11.64 ± 0.08 | -12.47 ± 0.12 |
| NGC 5576 | 8.20 ± 0.10 | 2.26 ± 0.01 | 0.76 ± 0.09 | 2.52 ± 0.09 | -0.23 ± 0.05 | 10.70 ± 0.08 | -15.00 ± 0.44 |
| NGC 5846 | 9.04 ± 0.04 | 2.38 ± 0.01 | 0.98 ± 0.31 | 2.64 ± 0.10 | -0.13 ± 0.08 | 11.18 ± 0.09 | -15.48 ± 0.44 |
| NGC 6251 | 8.77 ± 0.14 | 2.49 ± 0.03 | 1.16 ± 0.09 | 2.66 ± 0.09 | 1.05 ± 0.04 | 11.51 ± 0.08 | -11.49 ± 0.09 |
| NGC 7052 | 9.35 ± 0.02 | 2.45 ± 0.02 | 0.77 ± 0.09 | 3.04 ± 0.07 | 0.58 ± 0.05 | 11.22 ± 0.08 | -11.69 ± 0.10 |
| NGC 7619 | 9.35 ± 0.10 | 2.50 ± 0.01 | 1.11 ± 0.31 | 2.52 ± 0.07 | -0.01 ± 0.12 | 11.29 ± 0.08 | -15.11 ± 0.44 |
| NGC 7768 | 9.10 ± 0.15 | 2.46 ± 0.02 | 1.32 ± 0.31 | 2.36 ± 0.09 | -0.38 ± 0.05 | 11.44 ± 0.09 | -14.63 ± 0.44 |
| 28 Spiral Galaxies | | | | | | | |
| NGC 404 | 5.74 ± 0.10 | 1.54 ± 0.04 | -1.24 ± 0.31 | 3.64 ± 0.12 | 1.28 ± 0.05 | 8.85 ± 0.09 | -10.36 ± 0.16 |
| NGC 524 | 8.68 ± 0.10 | 2.37 ± 0.01 | 0.04 ± 0.31 | 3.83 ± 0.07 | 0.52 ± 0.06 | 11.10 ± 0.08 | -12.16 ± 0.10 |
| NGC 1023 | 7.62 ± 0.04 | 2.29 ± 0.01 | -0.41 ± 0.09 | 4.21 ± 0.09 | 0.18 ± 0.07 | 10.61 ± 0.08 | -12.35 ± 0.11 |
| NGC 1194 | 7.82 ± 0.04 | 2.17 ± 0.07 | -0.04 ± 0.40 | 3.96 ± 0.09 | 2.83 ± 0.04 | 10.46 ± 0.08 | -9.87 ± 0.09 |
| NGC 1316 | 8.16 ± 0.22 | 2.35 ± 0.01 | 0.14 ± 0.31 | 3.94 ± 0.30 | 0.65 ± 0.05 | 11.43 ± 0.08 | -11.96 ± 0.09 |
| NGC 1332 | 9.16 ± 0.06 | 2.47 ± 0.02 | 0.28 ± 0.40 | 3.68 ± 0.10 | 0.42 ± 0.05 | 10.88 ± 0.08 | -11.70 ± 0.11 |
| NGC 1374 | 8.76 ± 0.04 | 2.25 ± 0.01 | 0.03 ± 0.31 | 3.35 ± 0.08 | 0.12 ± 0.07 | 10.33 ± 0.08 | -14.63 ± 0.44 |
| NGC 2549 | 7.14 ± 0.23 | 2.15 ± 0.01 | -0.73 ± 0.09 | 4.25 ± 0.13 | 0.33 ± 0.06 | 9.97 ± 0.08 | -11.97 ± 0.15 |
| NGC 2778 | 7.14 ± 0.43 | 2.19 ± 0.01 | -0.63 ± 0.31 | 3.85 ± 0.14 | 0.12 ± 0.05 | 9.89 ± 0.08 | -11.70 ± 0.17 |
| NGC 2787 | 7.60 ± 0.05 | 2.28 ± 0.01 | -0.86 ± 0.31 | 4.16 ± 0.16 | 0.59 ± 0.04 | 9.80 ± 0.08 | -11.89 ± 0.14 |
| NGC 3115 | 8.94 ± 0.31 | 2.42 ± 0.01 | 0.19 ± 0.09 | 3.58 ± 0.08 | 0.14 ± 0.12 | 10.63 ± 0.08 | -12.60 ± 0.13 |
| NGC 3245 | 8.30 ± 0.11 | 2.32 ± 0.02 | -0.63 ± 0.09 | 4.50 ± 0.10 | 1.09 ± 0.04 | 10.45 ± 0.08 | -11.22 ± 0.09 |
| NGC 3384 | 7.02 ± 0.20 | 2.16 ± 0.01 | -0.52 ± 0.09 | 4.29 ± 0.08 | 0.24 ± 0.05 | 10.37 ± 0.08 | -11.85 ± 0.11 |
| NGC 3489 | 6.76 ± 0.06 | 2.02 ± 0.01 | -1.02 ± 0.31 | 4.74 ± 0.09 | 1.16 ± 0.04 | 10.14 ± 0.08 | -11.24 ± 0.09 |
| NGC 3665 | 8.76 ± 0.10 | 2.33 ± 0.02 | 0.33 ± 0.31 | 3.57 ± 0.09 | 1.33 ± 0.04 | 11.13 ± 0.08 | -11.27 ± 0.09 |
| NGC 3998 | 8.42 ± 0.18 | 2.42 ± 0.02 | -0.51 ± 0.40 | 4.20 ± 0.10 | 1.39 ± 0.05 | 10.30 ± 0.08 | -11.18 ± 0.09 |
| NGC 4026 | 8.26 ± 0.11 | 2.24 ± 0.01 | -0.83 ± 0.40 | 4.97 ± 0.13 | 0.52 ± 0.05 | 10.18 ± 0.08 | -13.23 ± 0.39 |
| NGC 4339 | 7.63 ± 0.12 | 2.05 ± 0.01 | -0.31 ± 0.31 | 3.48 ± 0.10 | 0.67 ± 0.14 | 10.02 ± 0.08 | -11.16 ± 0.12 |
| NGC 4342 | 8.65 ± 0.18 | 2.38 ± 0.01 | -0.29 ± 0.31 | 3.68 ± 0.07 | 0.31 ± 0.04 | 10.10 ± 0.08 | -14.40 ± 0.44 |
| NGC 4350 | 8.87 ± 0.14 | 2.26 ± 0.01 | 0.20 ± 0.31 | 3.08 ± 0.07 | 0.51 ± 0.06 | 10.35 ± 0.08 | -12.05 ± 0.10 |
| NGC 4371 | 6.83 ± 0.07 | 2.11 ± 0.01 | -0.15 ± 0.31 | 3.37 ± 0.19 | 0.64 ± 0.09 | 10.38 ± 0.08 | -12.02 ± 0.11 |
| NGC 4429 | 8.18 ± 0.03 | 2.24 ± 0.01 | -0.05 ± 0.31 | 3.76 ± 0.08 | 0.86 ± 0.05 | 10.75 ± 0.08 | -11.59 ± 0.09 |
| NGC 4434 | 7.84 ± 0.05 | 2.07 ± 0.01 | -0.25 ± 0.31 | 3.59 ± 0.07 | 0.00 ± 0.06 | 10.03 ± 0.08 | -14.33 ± 0.44 |
| NGC 4459 | 7.83 ± 0.08 | 2.24 ± 0.01 | -0.01 ± 0.31 | 3.69 ± 0.11 | 1.17 ± 0.10 | 10.56 ± 0.08 | -11.12 ± 0.09 |
| NGC 4526 | 8.66 ± 0.01 | 2.35 ± 0.02 | 0.06 ± 0.31 | 3.73 ± 0.10 | 1.14 ± 0.05 | 10.84 ± 0.08 | -11.38 ± 0.09 |
| NGC 4564 | 7.90 ± 0.12 | 2.19 ± 0.01 | -0.38 ± 0.09 | 3.96 ± 0.09 | 0.31 ± 0.05 | 10.12 ± 0.08 | -12.54 ± 0.13 |
| NGC 4578 | 7.28 ± 0.08 | 2.05 ± 0.02 | -0.31 ± 0.31 | 3.58 ± 0.07 | -0.24 ± 0.10 | 10.05 ± 0.08 | -14.35 ± 0.44 |
| NGC 4596 | 7.98 ± 0.20 | 2.15 ± 0.01 | -0.13 ± 0.09 | 3.64 ± 0.07 | 0.34 ± 0.08 | 10.54 ± 0.08 | -11.84 ± 0.09 |
| NGC 4742 | 7.13 ± 0.15 | 2.01 ± 0.01 | -0.61 ± 0.31 | 4.28 ± 0.09 | 0.40 ± 0.05 | 9.99 ± 0.09 | -11.79 ± 0.15 |
| NGC 4762 | 7.24 ± 0.05 | 2.15 ± 0.01 | -0.74 ± 0.31 | 4.38 ± 0.07 | 0.18 ± 0.07 | 10.56 ± 0.08 | -14.86 ± 0.44 |
| NGC 5018 | 8.00 ± 0.08 | 2.33 ± 0.01 | 0.05 ± 0.31 | 4.00 ± 0.09 | 0.89 ± 0.07 | 11.10 ± 0.08 | -11.40 ± 0.09 |
| NGC 5128 | 7.65 ± 0.13 | 2.01 ± 0.03 | 0.04 ± 0.40 | 3.75 ± 0.07 | 2.53 ± 0.03 | 10.86 ± 0.08 | -10.61 ± 0.09 |
| NGC 5252 | 9.03 ± 0.35 | 2.27 ± 0.06 | -0.15 ± 0.31 | 4.37 ± 0.09 | 2.27 ± 0.04 | 11.05 ± 0.08 | -10.30 ± 0.09 |
| NGC 5813 | 8.83 ± 0.04 | 2.37 ± 0.01 | 0.32 ± 0.31 | 3.44 ± 0.10 | 0.03 ± 0.12 | 11.10 ± 0.08 | -12.49 ± 0.12 |
| NGC 5845 | 8.41 ± 0.16 | 2.36 ± 0.01 | -0.20 ± 0.31 | 3.70 ± 0.11 | 0.53 ± 0.04 | 10.14 ± 0.08 | -11.75 ± 0.11 |
| NGC 6861 | 9.30 ± 0.22 | 2.59 ± 0.02 | 0.41 ± 0.31 | 3.29 ± 0.14 | 0.76 ± 0.05 | 10.84 ± 0.08 | -11.81 ± 0.09 |
| NGC 7332 | 7.06 ± 0.20 | 2.11 ± 0.01 | -0.59 ± 0.40 | 4.49 ± 0.10 | 0.46 ± 0.05 | 10.48 ± 0.08 | -11.78 ± 0.11 |
| NGC 7457 | 6.96 ± 0.26 | 1.83 ± 0.02 | -0.40 ± 0.31 | 3.32 ± 0.10 | 0.27 ± 0.11 | 9.92 ± 0.08 | -11.74 ± 0.14 |
| 28 Spiral Galaxies | | | | | | | |
| Circinus | 6.22 ± 0.08 | 2.17 ± 0.05 | -0.34 ± 0.03 | 3.86 ± 0.17 | 4.02 ± 0.03 | 10.04 ± 0.09 | -9.37 ± 0.13 |
| IC 2560 | 6.51 ± 0.09 | 2.14 ± 0.01 | -0.21 ± 0.03 | 3.29 ± 0.15 | 3.38 ± 0.04 | 10.43 ± 0.08 | -10.01 ± 0.09 |
| NGC 224 | 8.15 ± 0.19 | 2.19 ± 0.01 | -0.16 ± 0.00 | 3.67 ± 0.08 | 2.08 ± 0.04 | 10.71 ± 0.08 | -10.94 ± 0.09 |
| NGC 253 | 7.00 ± 0.30 | 1.98 ± 0.08 | -0.33 ± 0.01 | 3.61 ± 0.07 | 3.81 ± 0.04 | 10.43 ± 0.08 | -9.85 ± 0.09 |
| NGC 1097 | 8.38 ± 0.09 | 2.29 ± 0.01 | 0.13 ± 0.07 | 3.75 ± 0.07 | 3.41 ± 0.04 | 11.22 ± 0.08 | -10.04 ± 0.10 |
| NGC 1300 | 7.86 ± 0.31 | 2.34 ± 0.06 | -0.13 ± 0.10 | 3.19 ± 0.09 | 2.91 ± 0.04 | 10.56 ± 0.08 | -10.41 ± 0.09 |
| NGC 1320 | 6.77 ± 0.16 | 2.04 ± 0.04 | -0.70 ± 0.07 | 4.24 ± 0.09 | 3.34 ± 0.04 | 10.13 ± 0.09 | -9.68 ± 0.10 |
| NGC 1398 | 8.03 ± 0.08 | 2.29 ± 0.04 | 0.09 ± 0.04 | 3.58 ± 0.17 | 2.14 ± 0.04 | 11.17 ± 0.08 | -10.88 ± 0.09 |
| NGC 2273 | 6.95 ± 0.06 | 2.15 ± 0.03 | -0.55 ± 0.03 | 3.77 ± 0.15 | 3.14 ± 0.04 | 10.43 ± 0.08 | -10.02 ± 0.09 |
| NGC 2960 | 7.07 ± 0.04 | 2.22 ± 0.04 | -0.13 ± 0.05 | 3.89 ± 0.09 | 2.98 ± 0.04 | 10.72 ± 0.08 | -10.42 ± 0.09 |
| NGC 2974 | 8.23 ± 0.05 | 2.37 ± 0.01 | -0.17 ± 0.01 | 3.76 ± 0.12 | 1.36 ± 0.08 | 10.61 ± 0.08 | -11.28 ± 0.09 |
| NGC 3031 | 7.83 ± 0.09 | 2.18 ± 0.01 | -0.14 ± 0.01 | 3.65 ± 0.08 | 1.80 ± 0.03 | 10.57 ± 0.08 | -11.00 ± 0.09 |
| NGC 3079 | 6.28 ± 0.30 | 2.24 ± 0.03 | -0.46 ± 0.05 | 4.04 ± 0.17 | 3.64 ± 0.04 | 10.41 ± 0.08 | -9.84 ± 0.09 |
| NGC 3227 | 7.25 ± 0.25 | 2.10 ± 0.02 | 0.01 ± 0.03 | 3.38 ± 0.13 | 3.09 ± 0.04 | 10.65 ± 0.08 | -10.11 ± 0.09 |
| NGC 3368 | 6.89 ± 0.09 | 2.07 ± 0.01 | -0.60 ± 0.02 | 4.22 ± 0.14 | 2.13 ± 0.04 | 10.55 ± 0.08 | -10.81 ± 0.09 |
| NGC 3627 | 6.94 ± 0.05 | 2.10 ± 0.02 | -0.71 ± 0.07 | 4.33 ± 0.17 | 3.44 ± 0.04 | 10.63 ± 0.08 | -10.06 ± 0.09 |
| NGC 4151 | 7.29 ± 0.37 | 1.96 ± 0.05 | -0.25 ± 0.02 | 3.97 ± 0.13 | 2.82 ± 0.04 | 10.53 ± 0.08 | -9.89 ± 0.09 |
| NGC 4258 | 7.61 ± 0.01 | 2.12 ± 0.02 | -0.01 ± 0.06 | 3.27 ± 0.07 | 2.44 ± 0.04 | 10.56 ± 0.08 | -10.59 ± 0.09 |

| | | | | | | | |
|----------|-----------------|-----------------|------------------|-----------------|-----------------|------------------|-------------------|
| NGC 4303 | 6.78 ± 0.04 | 1.98 ± 0.04 | -0.70 ± 0.02 | 4.65 ± 0.07 | 3.87 ± 0.04 | 10.71 ± 0.09 | -9.86 ± 0.10 |
| NGC 4388 | 6.95 ± 0.09 | 2.00 ± 0.04 | 0.09 ± 0.02 | 3.07 ± 0.19 | 3.15 ± 0.04 | 10.12 ± 0.08 | -9.85 ± 0.09 |
| NGC 4501 | 7.31 ± 0.08 | 2.22 ± 0.02 | 0.22 ± 0.04 | 3.22 ± 0.07 | 3.05 ± 0.04 | 10.89 ± 0.08 | -10.35 ± 0.09 |
| NGC 4594 | 8.81 ± 0.03 | 2.35 ± 0.01 | 0.28 ± 0.02 | 3.48 ± 0.08 | 0.90 ± 0.05 | 11.06 ± 0.08 | -11.55 ± 0.09 |
| NGC 4699 | 8.28 ± 0.05 | 2.28 ± 0.02 | -0.64 ± 0.06 | 3.25 ± 0.18 | 2.20 ± 0.04 | 11.06 ± 0.08 | -10.86 ± 0.09 |
| NGC 4736 | 6.83 ± 0.10 | 2.03 ± 0.01 | -0.64 ± 0.01 | 4.46 ± 0.10 | 2.71 ± 0.04 | 10.38 ± 0.08 | -10.46 ± 0.09 |
| NGC 4826 | 6.20 ± 0.11 | 1.99 ± 0.02 | -0.38 ± 0.01 | 3.73 ± 0.11 | 2.21 ± 0.04 | 10.54 ± 0.08 | -10.68 ± 0.09 |
| NGC 4945 | 6.12 ± 0.30 | 2.07 ± 0.06 | -0.80 ± 0.14 | 3.78 ± 0.07 | 3.56 ± 0.03 | 10.23 ± 0.08 | -9.91 ± 0.09 |
| NGC 7582 | 7.74 ± 0.18 | 2.07 ± 0.02 | -0.32 ± 0.11 | 4.08 ± 0.17 | 3.29 ± 0.04 | 10.59 ± 0.08 | -9.64 ± 0.09 |
| UGC 3789 | 7.07 ± 0.04 | 2.03 ± 0.05 | -0.24 ± 0.01 | 3.59 ± 0.13 | 3.22 ± 0.04 | 10.51 ± 0.08 | -10.19 ± 0.09 |

Table 1: Sample of 101 galaxies with dynamical SMBH mass measurements. Column (1): galaxy name. Column (2): black hole mass. Column (3): central stellar velocity dispersion. Column (4): equivalent-axis, effective (half-light) radius of the spheroid component. Column (5): average projected density within R_e . Column (6): WISE W2–W3 color. Column (7): galaxy stellar mass. Column (8): specific star-formation rate, i.e., $\log(\text{sSFR}) \equiv \log(\text{SFR}) - \log(M^*)$.

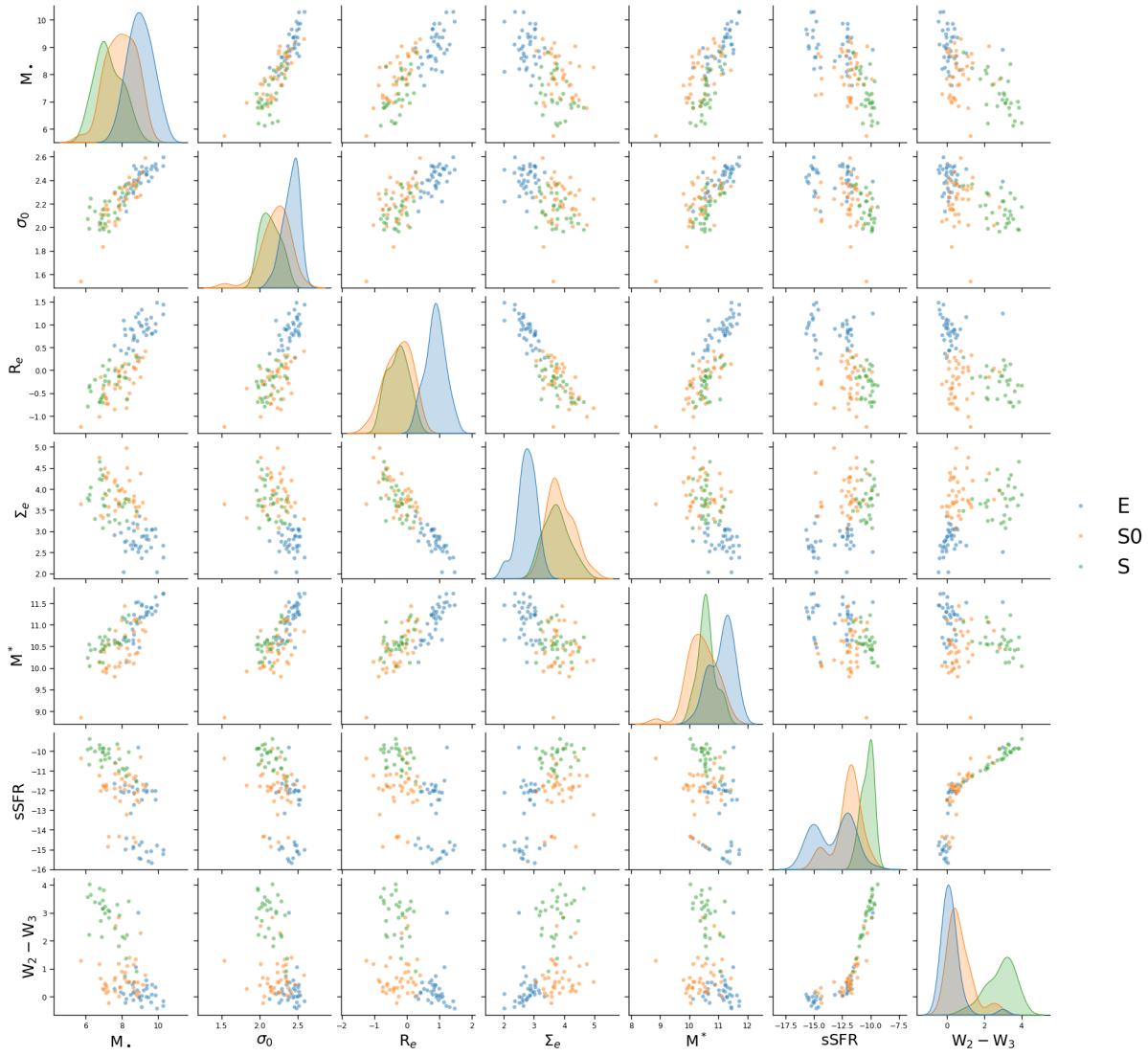


Figure 1: A pairplot of the data listed in Table S1.

| | | Edge Marginals of Uniform Prior | | | | | | | Path Marginals of Uniform Prior | | | | | | | |
|--------|-------|---------------------------------|------|------------|-------|------------|-------|--------|---------------------------------|-------------|------------|-------|------------|-------|--------|-------------|
| Parent | Child | M. | 0.00 | 0.29 | 0.29 | 0.29 | 0.29 | 0.29 | 0.29 | M. | 0.00 | 0.42 | 0.42 | 0.42 | 0.42 | 0.42 |
| | | σ_0 | 0.29 | 0.00 | 0.29 | 0.29 | 0.29 | 0.29 | 0.29 | σ_0 | 0.42 | 0.00 | 0.42 | 0.42 | 0.42 | 0.42 |
| | | R_e | 0.29 | 0.29 | 0.00 | 0.29 | 0.29 | 0.29 | 0.29 | R_e | 0.42 | 0.42 | 0.00 | 0.42 | 0.42 | 0.42 |
| | | Σ_e | 0.29 | 0.29 | 0.29 | 0.00 | 0.29 | 0.29 | 0.29 | Σ_e | 0.42 | 0.42 | 0.42 | 0.00 | 0.42 | 0.42 |
| | | M^* | 0.29 | 0.29 | 0.29 | 0.29 | 0.00 | 0.29 | 0.29 | M^* | 0.42 | 0.42 | 0.42 | 0.42 | 0.00 | 0.42 |
| | | $sSFR$ | 0.29 | 0.29 | 0.29 | 0.29 | 0.29 | 0.00 | 0.29 | $sSFR$ | 0.42 | 0.42 | 0.42 | 0.42 | 0.00 | 0.42 |
| | | $W_2 - W_3$ | 0.29 | 0.29 | 0.29 | 0.29 | 0.29 | 0.29 | 0.00 | $W_2 - W_3$ | 0.42 | 0.42 | 0.42 | 0.42 | 0.42 | 0.00 |
| | | M^* | - | σ_0 | R_e | Σ_e | M^* | $sSFR$ | M^* | - | σ_0 | R_e | Σ_e | M^* | $sSFR$ | $W_2 - W_3$ |

Figure 2: The edge marginals (left matrix) and path marginals (right matrix) for a uniform prior, i.e., all possible DAGs share the same probability.

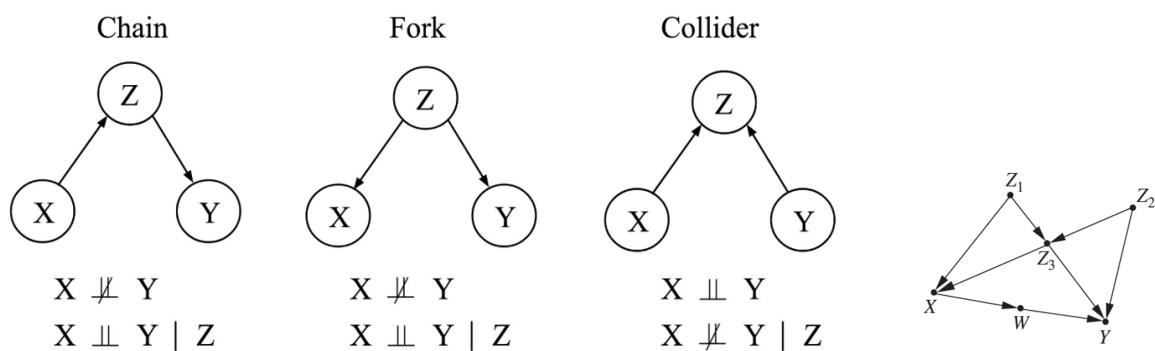


Figure 3: Three basic causal models, their (conditional) independencies, and a more complex graph on the right.

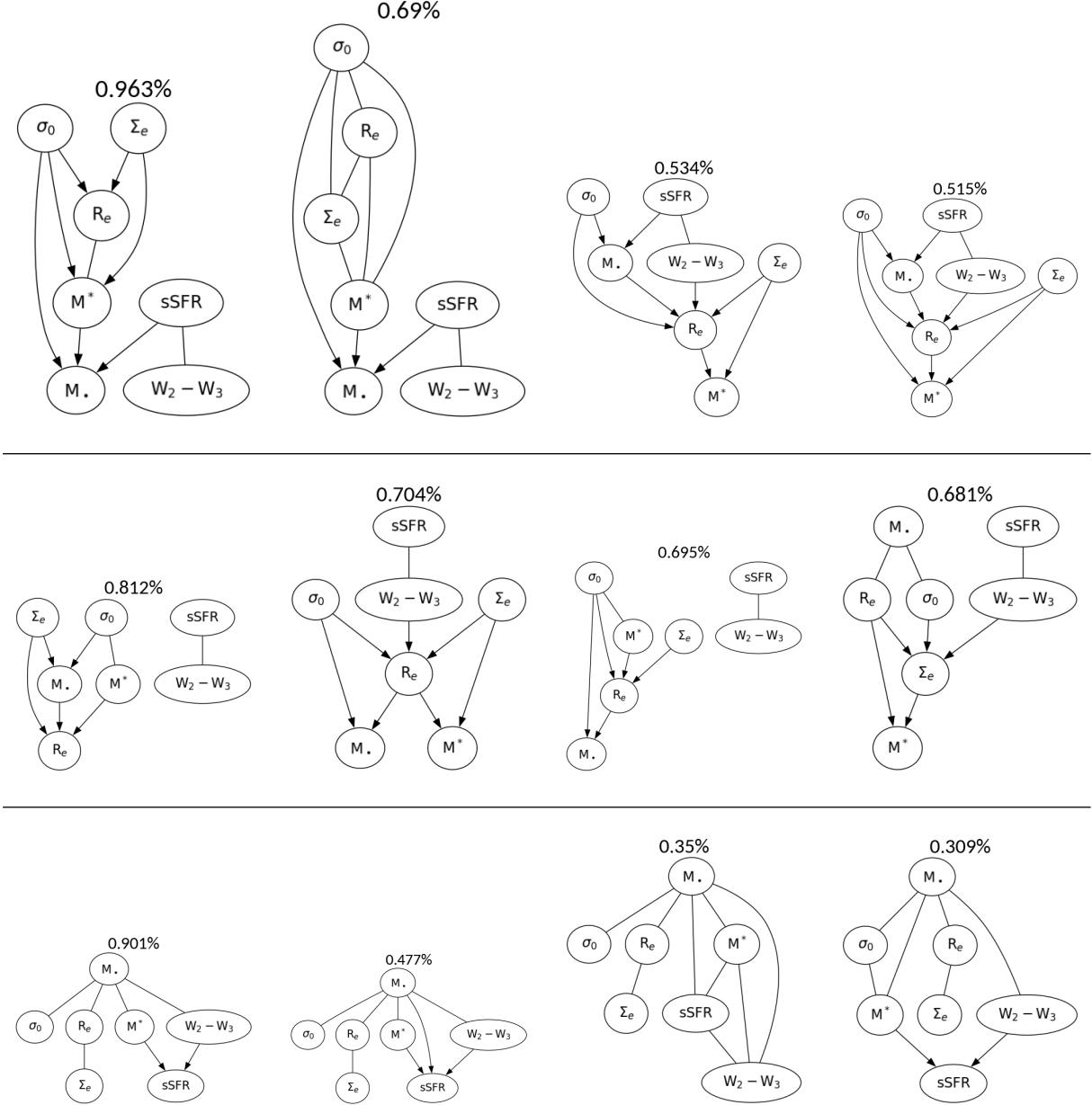


Figure 4: Exact posterior result for the top four Markov Equivalence Classes (MECs), represented as Partially Directed Acyclic Graphs (PDAGs) for elliptical (top panel), lenticular (middle panel), and spiral (bottom panel) galaxies. The posterior probability is labeled on top of each MEC, and is calculated by the sum of all DAG posterior probabilities within that MEC.

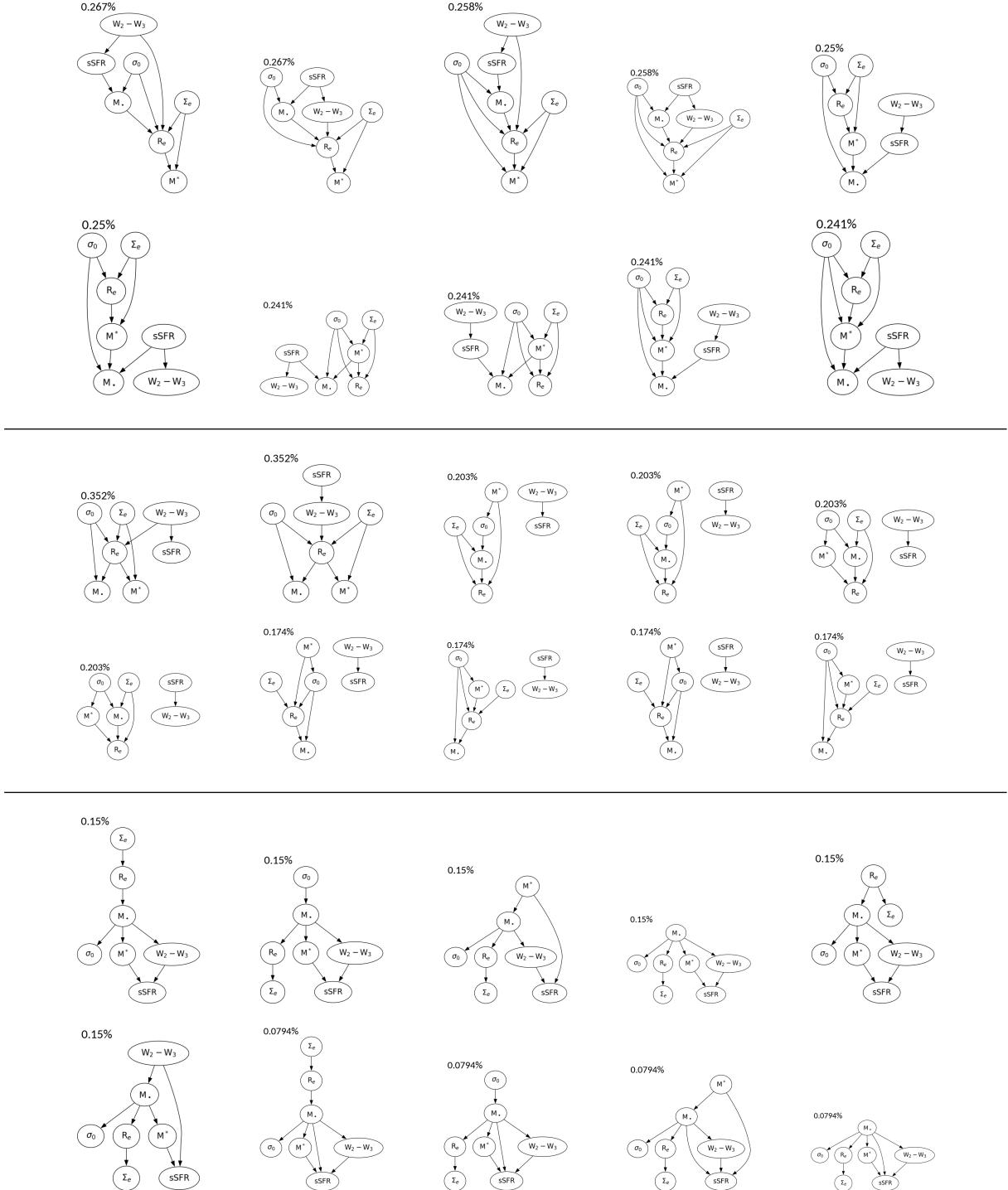


Figure 5: Exact posterior result for the top 10 most probable Directed Acyclic Graphs (DAGs) for elliptical (top panel), lenticular (middle panel), and spiral (bottom panel) galaxies. The percentage listed above each DAG indicates the posterior probability of the DAG, whereas the prior probability for every DAGs is equal to precisely $8.781333053161975 \times 10^{-10}$ (23).

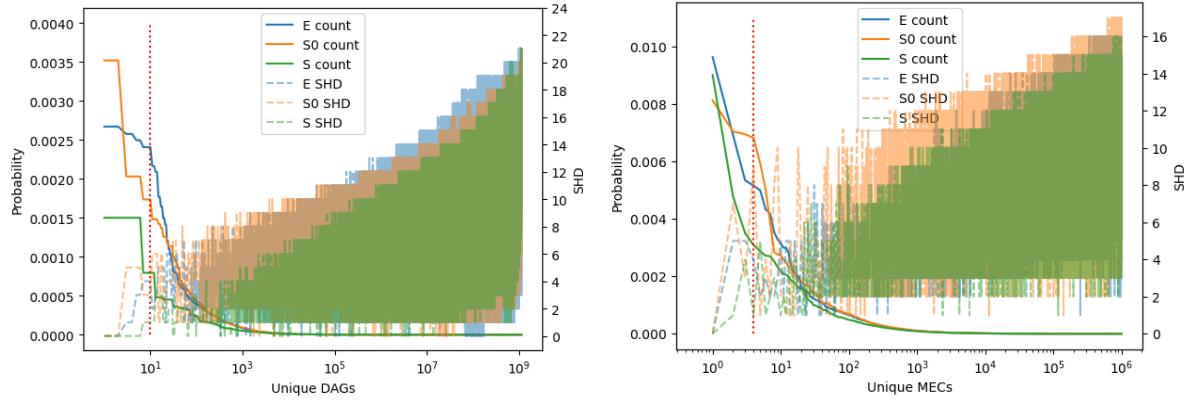


Figure 6: The exact posterior probability distribution and the structural Hamming distances (SHDs) to the most probable graph. There are in total 3.12510571×10^8 MECs in the case of seven nodes (114). Here, only the first 10^6 MECs are plotted for simplicity. The DAGs (left plot) and MECs (right plot) are ordered by their posterior probabilities from highest to lowest. The solid lines and the left y -axes show the posterior probability of the DAGs/MECs. The dashed lines and the right y -axes show the SHD, a measure of distance between graphs, from each DAG or MEC to the most probable DAG or MEC. The red dashed line marks the 10th DAG and the 4th MEC, which are shown in Extended Data Figs. 5 and 4, respectively.

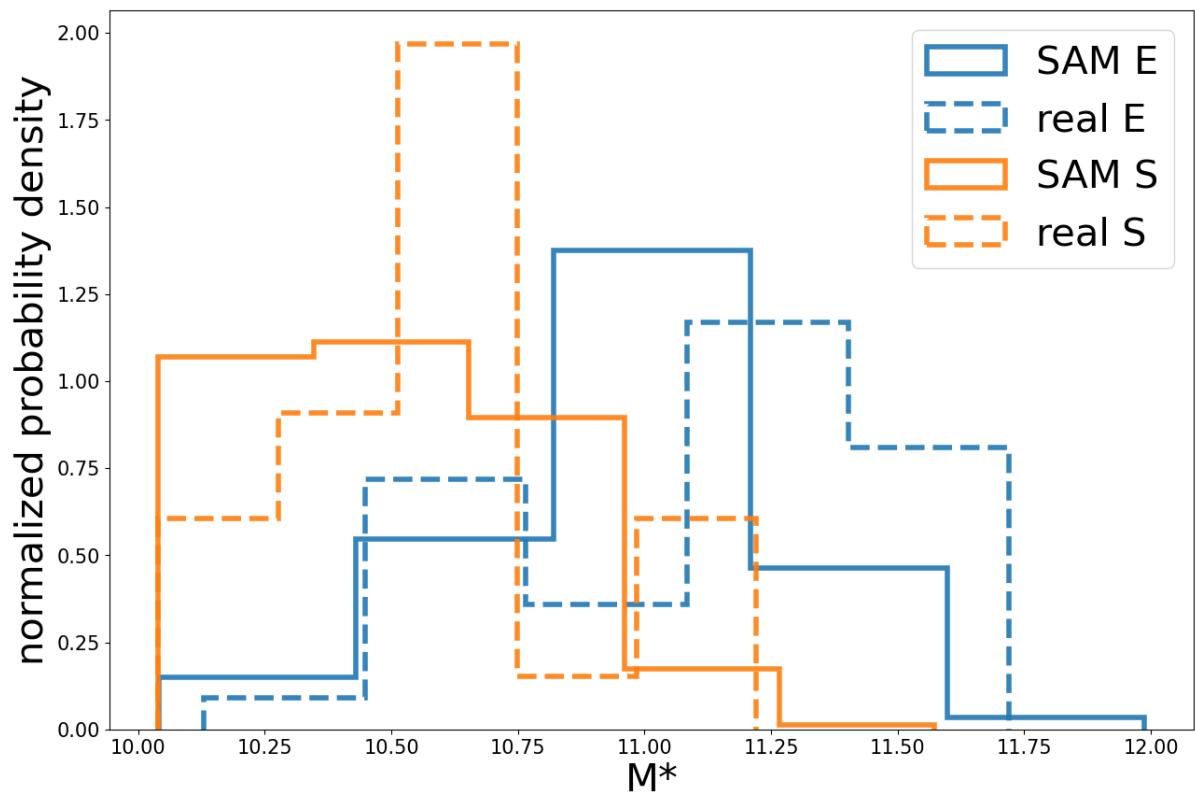


Figure 7: The M^* distribution for semi-analytical model (SAM) galaxies compared to the M^* distribution of real galaxies used in this work.

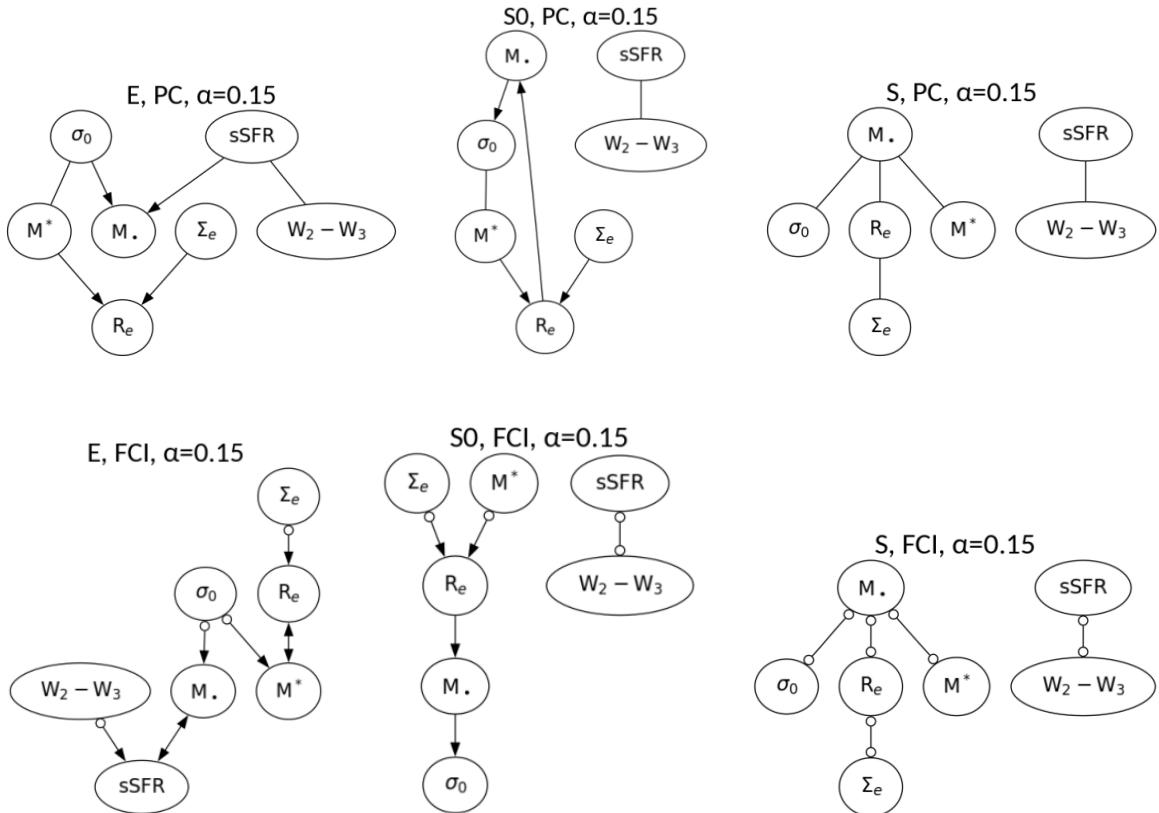


Figure 8: Graphs learned by the PC algorithm (upper row) and by the FCI algorithm (bottom row). The graphs in the top row use PDAGs to represent MECs of DAGs by leaving some edges undirected. The graphs in the bottom row are Partial Ancestral Graphs (PAGs) and introduce additional edge types: $A \longleftrightarrow B$ corresponding to a confounding relation (i.e., a third variable causes both A and B) and empty circles representing uncertainty regarding the ending symbol of the edge (i.e., $A \circ \rightarrow B$ may correspond to either $A \rightarrow B$ or to $A \longleftrightarrow B$, but rules out $B \rightarrow A$). The significance cutoff for conditional independence tests is set to $\alpha = 0.15$ in all graphs.

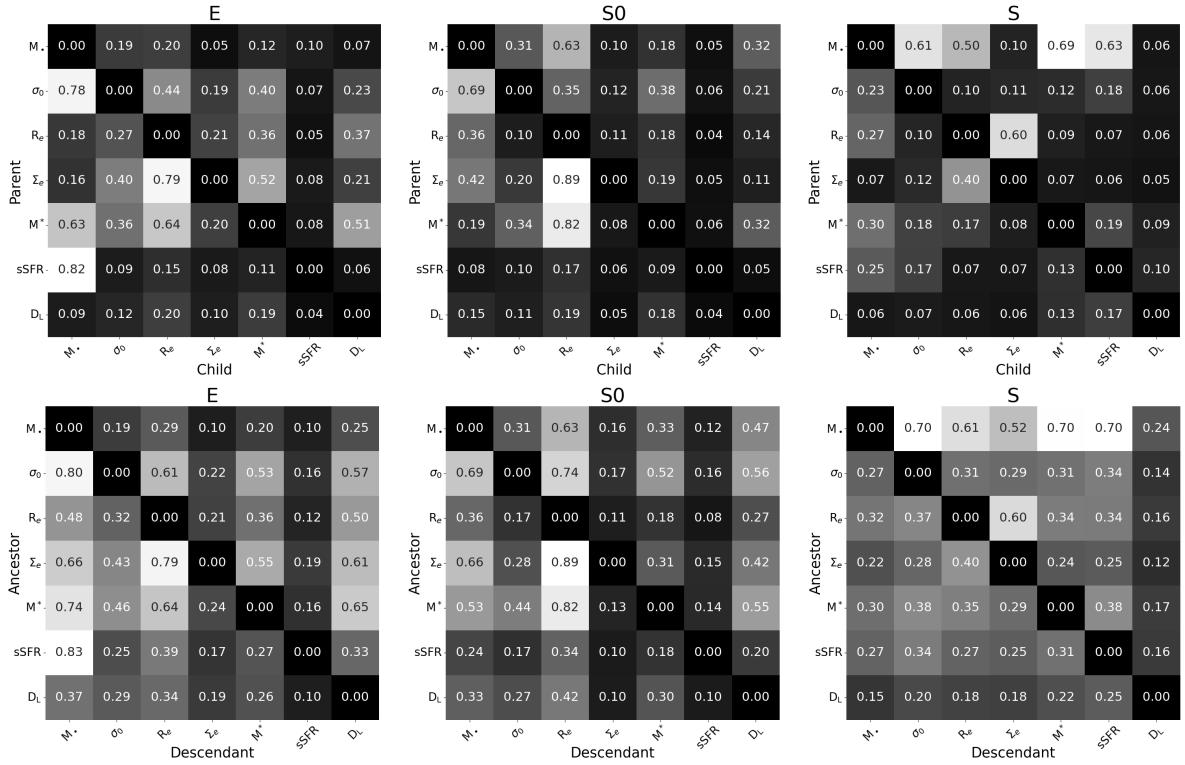


Figure 9: Edge marginals (top matrices) and path marginals (bottom matrices) with luminosity distance (D_L) as one of the variables.

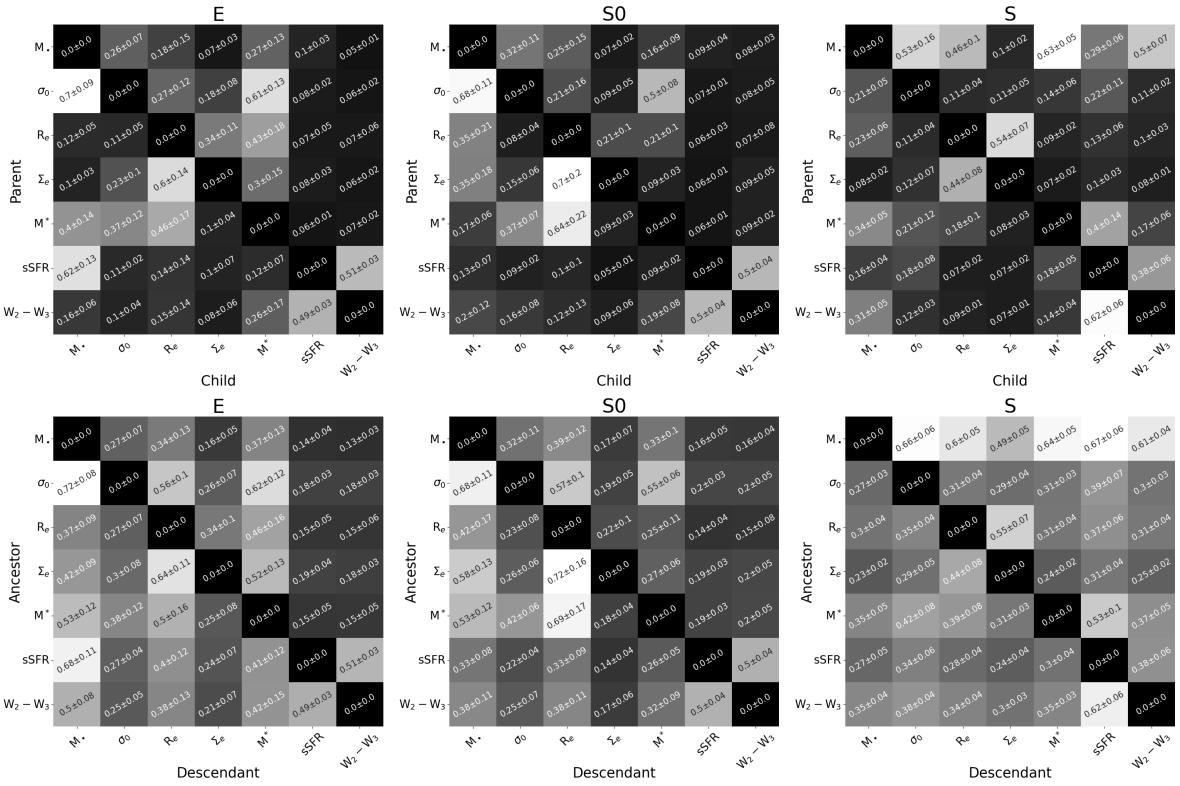


Figure 10: The mean and standard deviation of edge marginals (top matrices) and path marginals (bottom matrices) over 100 random sampling realizations for each morphological class.

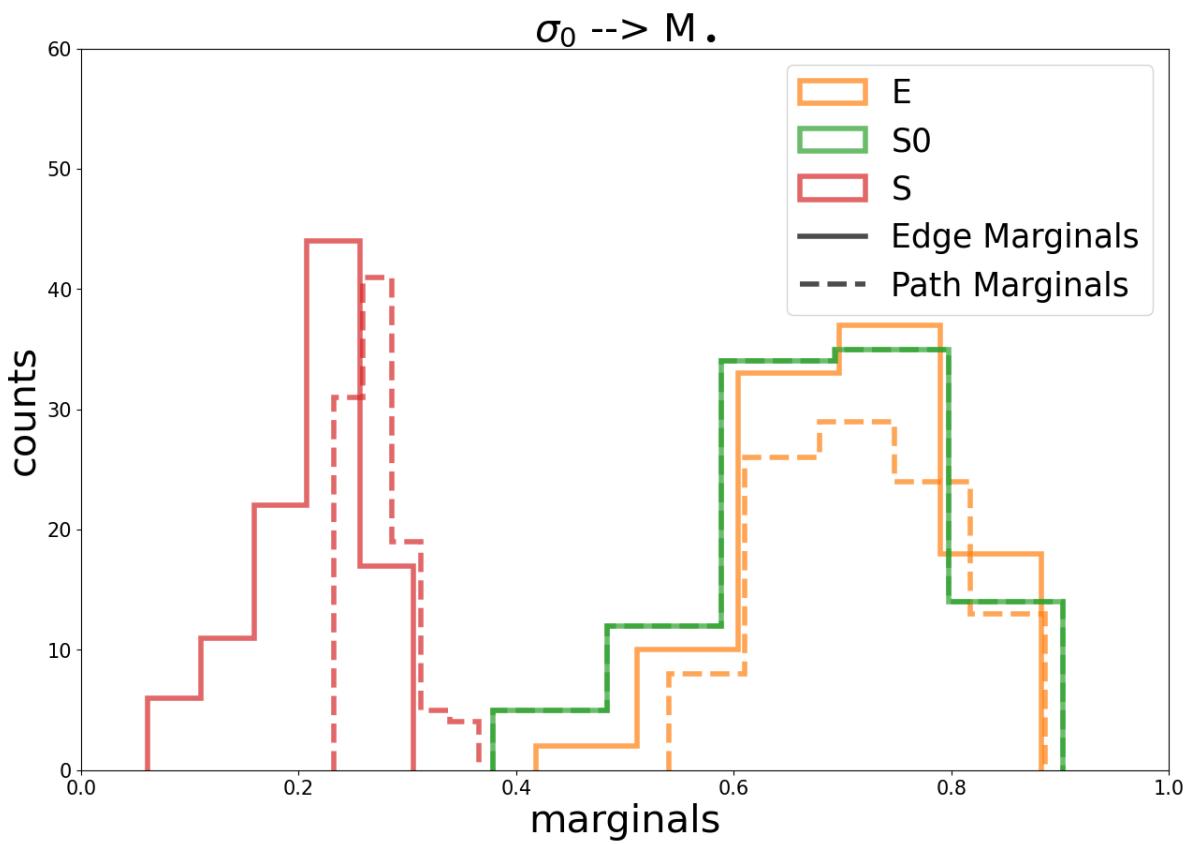


Figure 11: Edge marginal and path marginal distributions of $\sigma_0 \rightarrow M_\bullet$ over 100 random realizations for each morphological class.

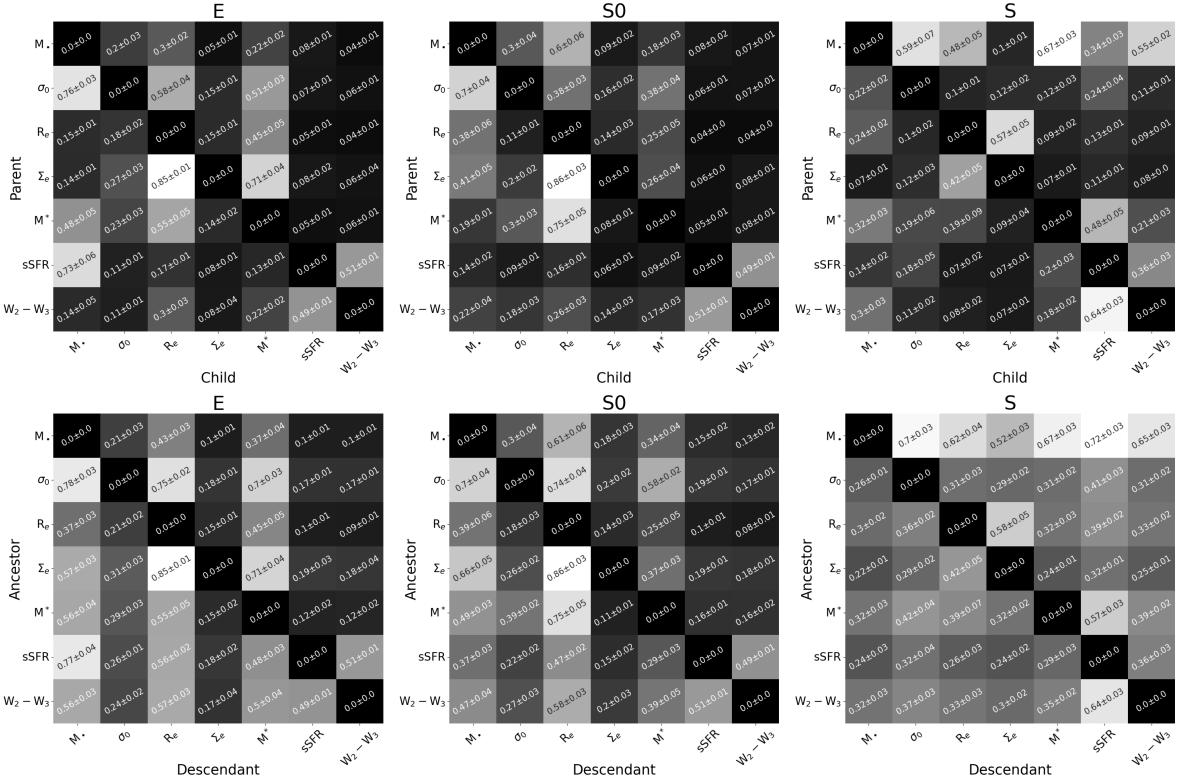


Figure 12: The mean and standard deviation of edge marginals (top matrices) and path marginals (bottom matrices) over all leave-one-out realizations.

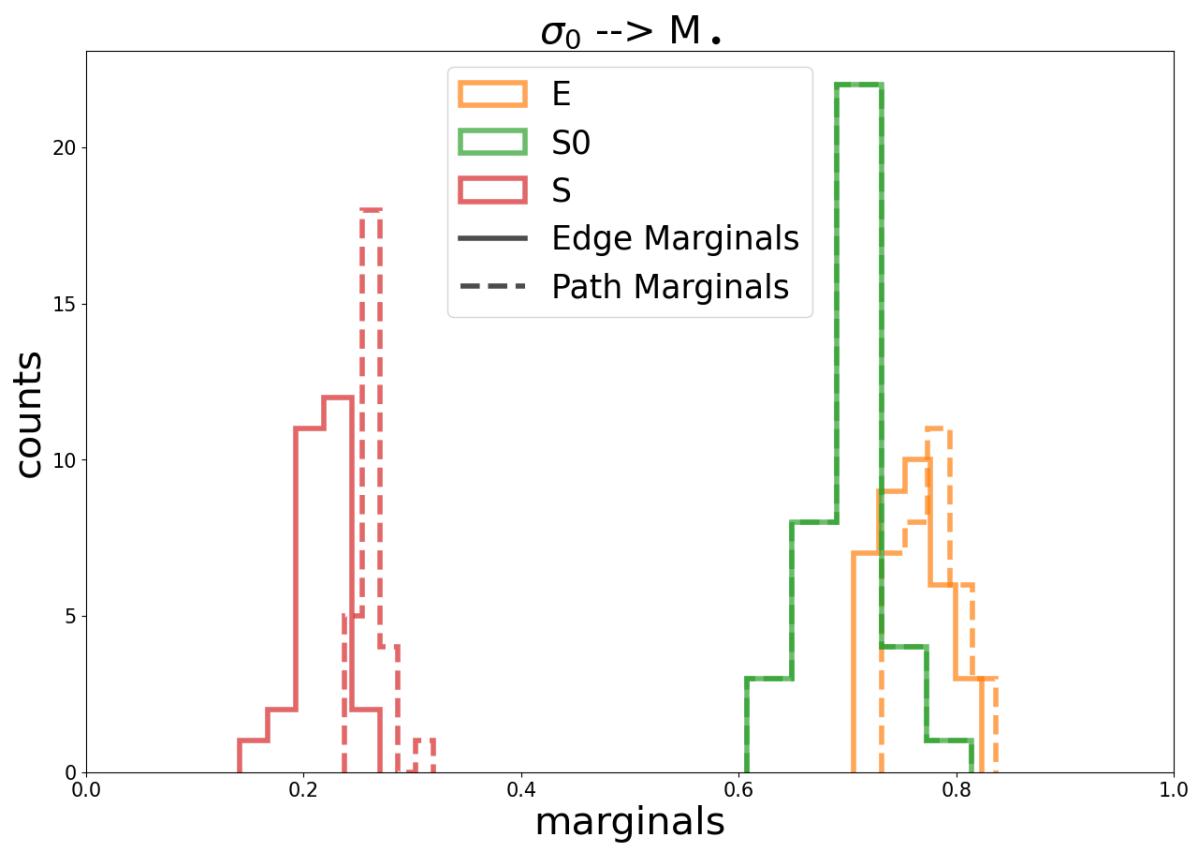


Figure 13: **Edge marginal and path marginal distributions of $\sigma_0 \rightarrow M_\bullet$ over all leave-one-out realizations.**

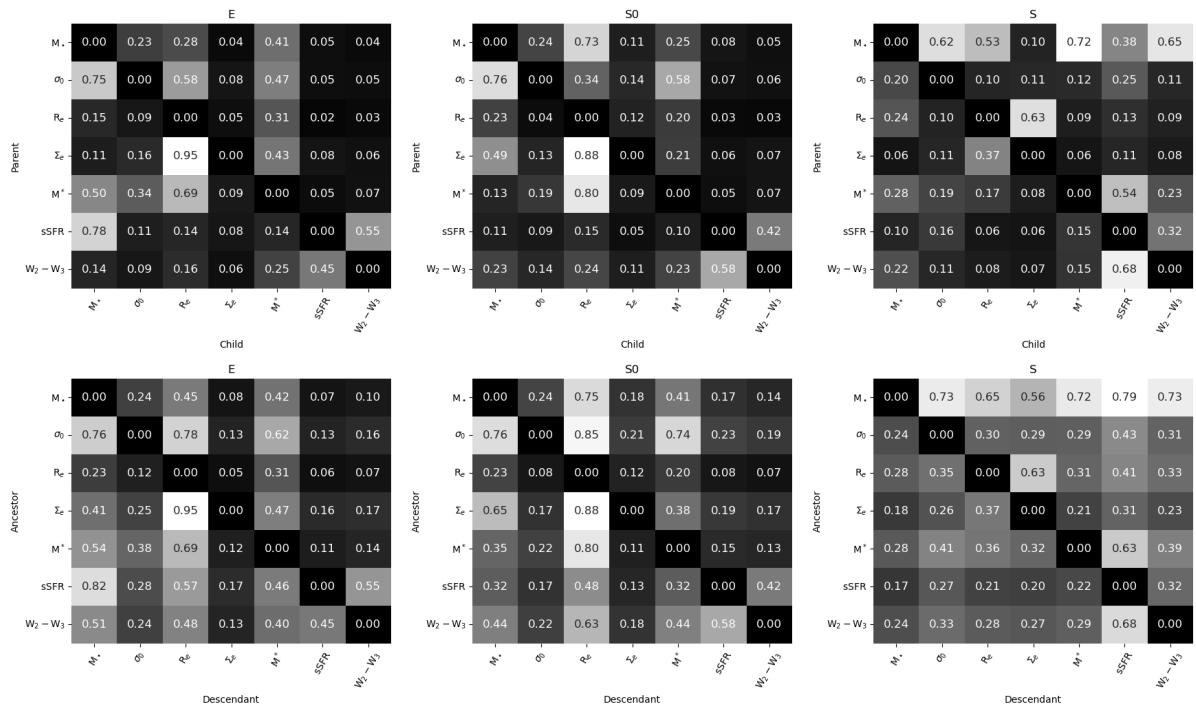


Figure 14: Edge marginals (top matrices) and path marginals (bottom matrices) approximated by DAG-GFN.

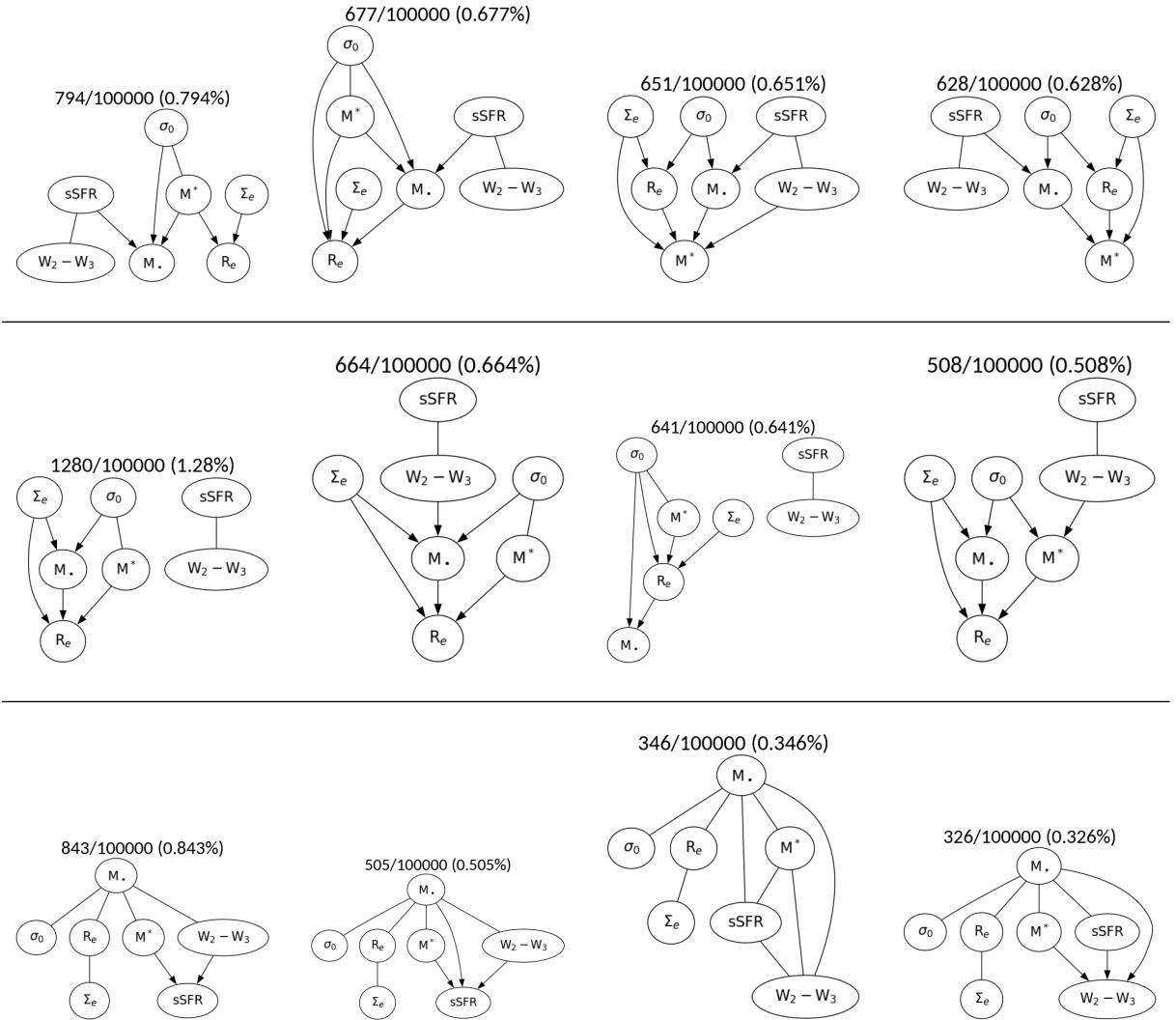


Figure 15: **Top four MECs sampled by DAG-GFN for elliptical (top panel), lenticular (middle panel), and spiral (bottom panel) galaxies.**

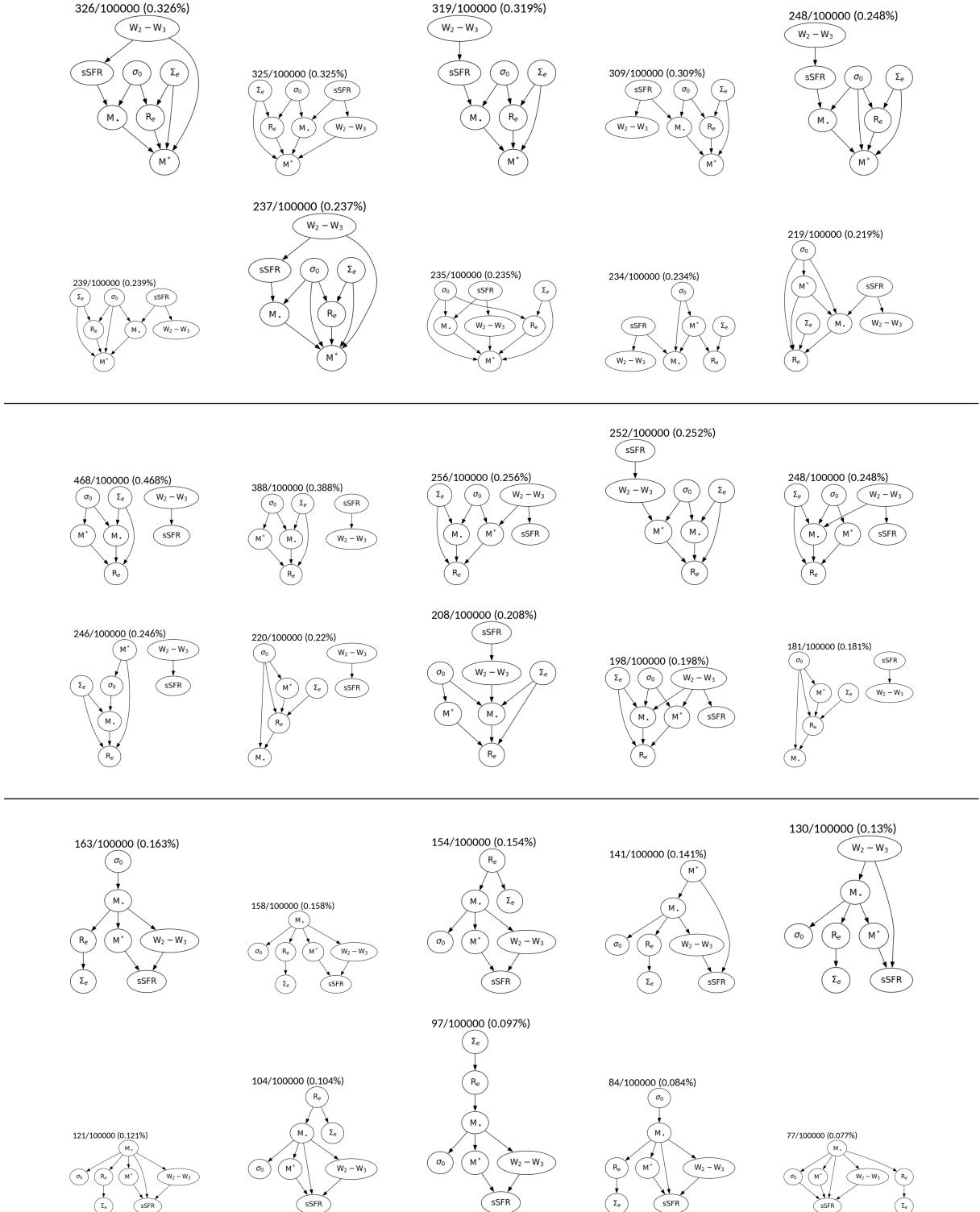


Figure 16: Top 10 DAGs sampled by DAG-GFN for elliptical (top panel), lenticular (middle panel), and spiral (bottom panel) galaxies.

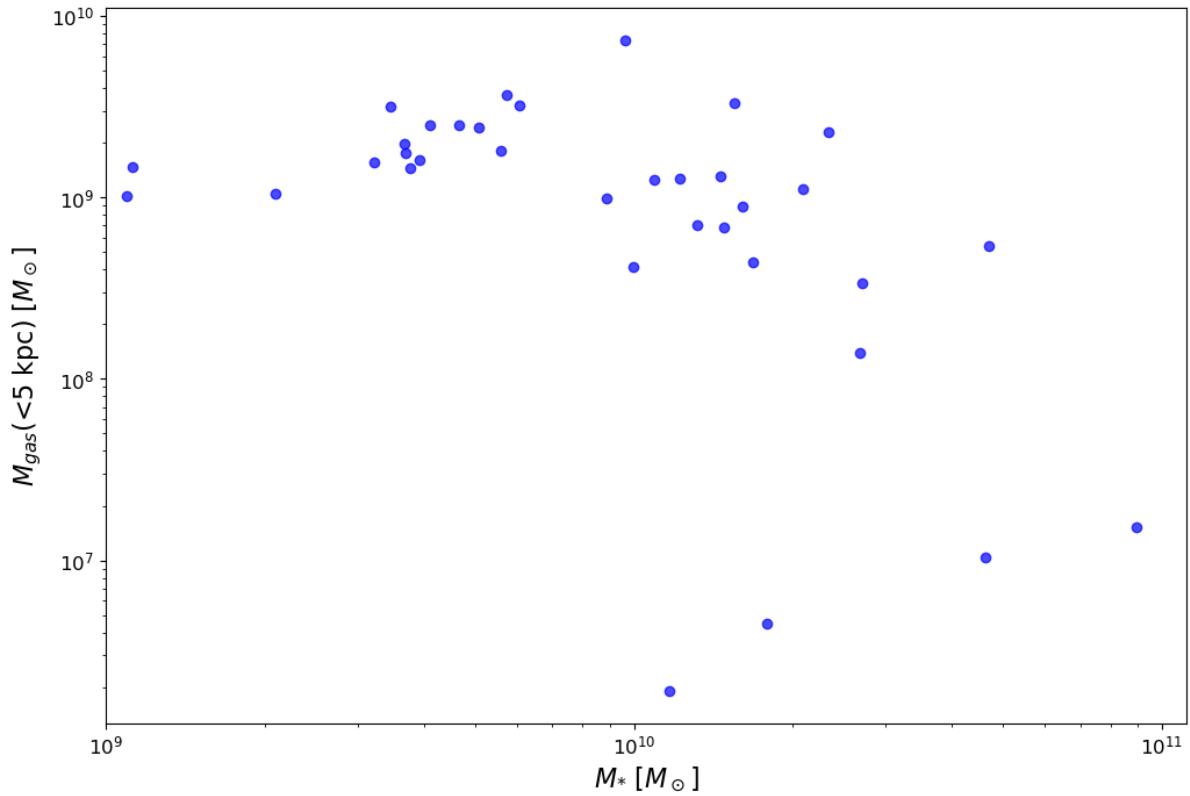


Figure 17: Gas mass within 5 kpc versus total stellar mass in NIHAO simulated galaxies (II2). The central gas mass is fairly constant in gas rich galaxies, implying that gas accretion onto the black hole, which is mainly regulated by the black hole mass and the local gas density (III), is also quite uniform across galaxies, weakening the galaxy → SMBH causal relation in spirals.