Draft version April 23, 2025
Typeset using LaTeX **twocolumn** style in AASTeX631

Mapping the Diversity of the Black Hole-Stellar Mass Relation Across Cosmological and Feedback Parameter Space in Simulated Galaxies

AstroPilot[1]

[1] *Anthropic, Gemini & OpenAI servers. Planet Earth.*

ABSTRACT

The tight correlation between supermassive black hole mass and host galaxy stellar mass (the BH-M* relation) is a cornerstone of galaxy evolution, suggesting a fundamental co-evolutionary process. However, the physical mechanisms driving this relation and its potential diversity across different galaxy populations and cosmic epochs remain debated, making it crucial to understand its dependence on underlying physical processes for interpreting observations and refining theoretical models. Cosmological hydrodynamical simulations, while essential for studying galaxy and black hole co-evolution, are computationally expensive, and the simulated BH-M* relation is highly sensitive to various input physics, particularly feedback processes and cosmological parameters. Systematically exploring this complex parameter space to isolate the impact of individual or combined parameters on the relation's properties (slope, normalization, scatter) is therefore challenging due to the sheer number of required simulations. We address this challenge by leveraging an unprecedented dataset of cosmological hydrodynamical simulations that systematically vary key cosmological and feedback parameters. We quantify the slope, normalization, and scatter of the simulated BH-M* relation across this extensive parameter space, and by analyzing the resulting variations, we reveal the primary physical drivers responsible for the observed diversity and identify secondary dependencies, thereby providing crucial insights into the interplay between large-scale structure formation, baryonic physics, and black hole growth.

Keywords: Algorithms, Astronomical methods, Astronomical models, Astronomical simulations, Astronomy data analysis, Astronomy data modeling, Astronomy data reduction, Astronomy data visualization, Astronomy databases, Astronomy software, Astrostatistics, Compact objects, Computational astronomy, Computational methods, Cosmological parameters, Galaxies, Galaxy evolution, Linear regression, Markov chain Monte Carlo, Multivariate analysis, N-body simulations, Regression, Scaling relations, Stellar astronomy, Stellar physics, Stellar properties, Stellar remnants, Supernovae

1. INTRODUCTION

The observed tight correlation between the mass of supermassive black holes (SMBHs) residing at the centers of galaxies and the properties of their host galaxies, such as stellar mass ($M_{\rm star}$) or velocity dispersion (σ), stands as a cornerstone of modern astrophysics (e.g., **????**). These scaling relations, particularly the black hole-stellar mass relation (hereafter the BH-$M_{\rm star}$ relation), suggest a fundamental link and co-evolutionary process between the growth of SMBHs and the formation and evolution of their host galaxies. Understanding the origin and maintenance of these relations is crucial for a complete picture of cosmic structure formation.

The existence of such tight correlations across a wide range of galaxy masses and types implies that feedback processes, driven by energy and momentum released from accreting SMBHs (Active Galactic Nuclei, AGN) and stellar processes (supernovae, stellar winds), play a pivotal role in regulating both black hole growth and star formation within galaxies. These feedback mechanisms are thought to be essential for explaining observed galaxy properties, such as the quenching of star formation in massive galaxies and the regulation of black hole accretion itself.

The BH-$M_{\rm star}$ relation thus serves as a critical observational constraint for theoretical models of galaxy evolution (Reines & Volonteri 2015; Zhu et al. 2020). Reproducing the observed slope, normalization, and scatter

of this relation is a key test for any successful model of galaxy and black hole co-evolution.

While the average BH-$M_{\rm star}$ relation is well-established observationally, there is significant scatter around the mean relation, and evidence suggests potential secondary dependencies on other galaxy properties like morphology, star formation rate (SFR), environment, and redshift (Ding et al. 2019; Zhang et al. 2023; Pacucci & Loeb 2024). This observed diversity hints that the relation is not purely universal but can be modulated by the specific physical conditions and evolutionary paths of individual galaxies.

Quantifying this diversity and identifying its physical drivers is essential for moving beyond a simple average relation and understanding the complex interplay of processes that shape galaxy and black hole growth.

Disentangling the effects that contribute to the scatter and secondary dependencies requires theoretical tools capable of following the coupled evolution of gas, stars, dark matter, and black holes within a cosmological context (Mayer et al. 2023; Valentini & Dolag 2025).

Cosmological hydrodynamical simulations have become indispensable tools for studying the co-evolution of galaxies and their central black holes from early cosmic times to the present day (Ni et al. 2024; Valentini & Dolag 2025).

These simulations attempt to model the complex interplay of gravity, hydrodynamics, radiative cooling, star formation, stellar feedback, and black hole growth and feedback within a ΛCDM universe (Valentini & Dolag 2025).

By tracking the formation and evolution of galaxies within their cosmological environment, these simulations can naturally produce scaling relations between black holes and their hosts, offering a theoretical framework to interpret observations (Mayer et al. 2023; Ni et al. 2024).

However, the results of these simulations, particularly the properties of the simulated BH-$M_{\rm star}$ relation, are known to be highly sensitive to the specific implementations and parameterizations of sub-grid physics models, most notably those describing stellar and AGN feedback processes (Ding et al. 2019; Pacucci & Loeb 2024; Sturm & Reines 2024).

Different feedback models, or even different parameter choices within the same model, can lead to significant variations in the slope, normalization, and scatter of the simulated relation.

This sensitivity arises because feedback processes directly impact the supply of gas for star formation and black hole accretion, as well as the expulsion of gas from galaxies, thereby regulating the growth of both components.

The precise way feedback energy and momentum are coupled to the surrounding gas is often represented by efficiency parameters in simulation models.

Furthermore, the BH-$M_{\rm star}$ relation can also be influenced by cosmological parameters, such as the matter density ($\Omega_{\rm m}$) and the amplitude of initial density fluctuations (σ_8). These parameters set the large-scale structure of the universe and the growth history of dark matter halos, which in turn dictate the environment and merger history of galaxies (Sicilia et al. 2021).

Variations in cosmological parameters can affect the rate at which galaxies assemble mass and accrete gas, indirectly influencing both stellar and black hole growth (Sicilia et al. 2021; Pacucci & Loeb 2024).

Therefore, the simulated BH-$M_{\rm star}$ relation is a complex outcome of the interplay between large-scale cosmic structure formation and small-scale baryonic physics, both of which are governed by the input parameters of the simulation (Sicilia et al. 2021).

Systematically exploring the vast parameter space defined by these feedback and cosmological parameters to isolate their individual and combined impacts on the BH-$M_{\rm star}$ relation is computationally prohibitive using traditional methods (Maity et al. 2023).

These methods typically involve running a small number of simulations with fixed, distinct parameter sets, making it difficult to map the continuous dependence of galaxy properties on underlying physics (Vernardos & Fluke 2013).

Comparing results from different simulation projects (e.g., EAGLE, IllustrisTNG) is valuable but complicated by differences in numerical techniques and the full suite of sub-grid models employed.

A dedicated, systematic exploration within a single simulation framework is required to rigorously quantify parameter dependencies (Ciotti et al. 2022).

This work addresses this challenge by leveraging an unprecedented dataset of cosmological hydrodynamical simulations specifically designed to explore a wide range of cosmological and feedback parameter variations within a consistent framework (Schaye et al. 2023; Horowitz & Lukic 2025).

This dataset, generated through a systematic parameter study (Horowitz & Lukic 2025), provides a unique opportunity to move beyond comparing a few distinct simulation models and instead quantify the continuous dependence of galaxy and black hole properties on underlying physical parameters.

The extensive coverage of the parameter space allows for a detailed investigation into how variations in

key feedback efficiencies (e.g., for supernovae and AGN) (Valentini & Dolag 2025) and cosmological parameters (Ω_m, σ_8) (Horowitz & Lukic 2025) collectively shape the resulting BH-M_{star} relation.

By analyzing this rich dataset, we can overcome the limitations of traditional simulation comparisons and perform a comprehensive mapping of the relation's properties across the explored parameter space.

Our approach involves systematically analyzing the BH-M_{star} relation within each individual simulation run (referred to as a "catalog" in our dataset).

For each catalog, we quantify the properties of the relation, specifically its slope, normalization, and intrinsic scatter, using robust statistical methods (Sturm & Reines 2024).

We focus on galaxies with detected black holes ($M_{BH} > 0$), carefully considering the treatment of galaxies with zero black hole mass to avoid biasing our regression results, while also quantifying the black hole occupation fraction as a complementary diagnostic (Zhang et al. 2023; Sturm & Reines 2024).

To account for potential mass-dependent physics, where different feedback mechanisms might dominate in different mass regimes, we perform our analysis within distinct stellar mass bins (Zhang et al. 2023).

This per-catalog analysis provides a set of quantitative descriptors for the BH-M_{star} relation for each unique combination of cosmological and feedback parameters in our dataset (Ding et al. 2019).

Building upon the per-catalog analysis, we then treat the measured properties of the BH-M_{star} relation (slope, normalization, scatter) as dependent variables that are functions of the input cosmological and feedback parameters (Weller et al. 2023; Pacucci & Loeb 2024).

We employ multivariate statistical techniques, such as regression analysis, to quantify the sensitivity of these relation properties to variations in each parameter (Sicilia et al. 2021, 2022).

This allows us to build a quantitative model that describes how the BH-M_{star} relation changes as we vary the underlying physics parameters (Weller et al. 2023; Pacucci & Loeb 2024).

By analyzing the coefficients or feature importance metrics from these models, we can identify which parameters exert the strongest influence on the slope, normalization, and scatter of the relation, and potentially uncover interactions between parameters (Sicilia et al. 2021, 2022).

Through this systematic mapping, we aim to reveal the primary physical drivers responsible for the diversity observed in the simulated BH-M_{star} relation across our parameter space (Jahnke & Maccio 2011).

We will quantify the relative importance of different feedback channels (e.g., thermal vs. kinetic AGN feedback, different supernova feedback implementations) and cosmological parameters in shaping the relation's characteristics (Pacucci & Loeb 2024,?).

Our analysis will provide crucial insights into how the efficiency and implementation of baryonic feedback processes, coupled with the large-scale cosmological environment, conspire to establish and maintain the observed scaling relations between black holes and galaxies (Jahnke & Maccio 2011).

We will also investigate potential secondary dependencies, such as the influence of star formation rate on the relation, within the context of varying feedback physics and across different mass regimes (Pacucci & Loeb 2024,?; Chen et al. 2025).

The results of this study will provide a quantitative link between the input physics of cosmological simulations and a key observable property of galaxies (Pacucci & Loeb 2024).

By identifying which parameters most strongly influence the BH-M_{star} relation, our findings can guide the development and calibration of future simulation models, helping to refine the sub-grid physics prescriptions necessary to reproduce observed galaxy populations (Zhang et al. 2023; Sturm & Reines 2024).

Furthermore, understanding the parameter dependence of the simulated relation's scatter and secondary dependencies can aid in interpreting the diversity observed in real galaxy populations and potentially constrain the underlying physical processes at play (Zhang et al. 2023; Sturm & Reines 2024).

This work represents a significant step towards building a predictive theoretical framework for galaxy and black hole co-evolution (Pacucci & Loeb 2024).

Looking ahead, the methodology developed here can be extended to explore the parameter dependence of other fundamental galaxy scaling relations, such as the stellar mass-halo mass relation or the mass-metallicity relation, within the same dataset (Giodini et al. 2013; D'Onofrio et al. 2021).

Future work could also involve applying more sophisticated statistical or machine learning techniques to model the complex, potentially non-linear dependencies of galaxy properties on simulation parameters (Bahar et al. 2022).

Exploring the redshift evolution of the parameter dependencies would also provide valuable insights into how the dominant physical processes shaping the BH-M_{star} relation change over cosmic time, offering a more complete picture of co-evolution across cosmic history (Giodini et al. 2013; Bahar et al. 2022).

2. METHODS

The primary objective of this work is to systematically quantify how the properties of the black hole–stellar mass ($M_{\rm BH}$–$M_{\rm star}$) relation in simulated galaxies vary across a multi-dimensional parameter space encompassing key cosmological and baryonic feedback parameters. To achieve this, we leverage a unique dataset of cosmological hydrodynamical simulations and employ a multi-step analysis methodology detailed below.

2.1. *Simulation Data and Parameter Space*

Our analysis is based on a suite of cosmological hydrodynamical simulations, each representing a distinct point in a predefined parameter space (Schaye et al. 2023). This suite includes variations in six key parameters: two cosmological parameters, the matter density Ω_m and the amplitude of matter fluctuations σ_8, and four parameters governing baryonic feedback processes, specifically two parameters for stellar feedback ($A_{\rm SN1}$, $A_{\rm SN2}$) and two for AGN feedback ($A_{\rm AGN1}$, $A_{\rm AGN2}$) (Schaye et al. 2023; Wright et al. 2024). Each simulation run, referred to as a 'catalog', provides a snapshot of the galaxy population at a specific redshift (implicitly assumed to be a fixed redshift for this analysis, likely $z = 0$ based on typical simulation outputs for scaling relations). The dataset comprises galaxy-level properties, including stellar mass ($M_{\rm star}$), black hole mass ($M_{\rm BH}$), and star formation rate (SFR), for a large sample of galaxies across all simulation catalogs (Schaye et al. 2023; Horowitz & Lukic 2025). The total dataset is stored in two main files: 'galaxies_full_optimal.parquet' containing galaxy properties and associated catalog parameters, and 'catalog_params_optimal.parquet' listing the parameters for each catalog.

2.2. *Data Preparation and Integration*

The initial step involves loading and preparing the simulation data for analysis. We load the galaxy-level data from 'galaxies_full_optimal.parquet' and the catalog-level parameters from 'catalog_params_optimal.parquet' (McGibbon & Khochfar 2023; Jung et al. 2024). To manage the potentially large size of the galaxy catalog, we employ efficient data loading techniques, such as chunked reading, to minimize memory footprint. While the galaxy DataFrame is structured to already include the parent catalog's parameters, we perform cross-checks to ensure data integrity and consistent association of each galaxy with its specific simulation parameters (Ω_m, σ_8, $A_{\rm SN1}$, $A_{\rm SN2}$, $A_{\rm AGN1}$, $A_{\rm AGN2}$) (McGibbon & Khochfar 2023; Jung et al. 2024). For the subsequent analysis, we select only the essential galaxy properties ($M_{\rm star}$, $M_{\rm BH}$, SFR) and the six simulation parameters, along with a unique catalog identifier ('catalog_number') (McGibbon & Khochfar 2023).

2.3. *Treatment of Galaxies with Zero Black Hole Mass*

A common characteristic of galaxy formation simulations is the presence of galaxies with zero assigned black hole mass ($M_{\rm BH} = 0$). This can represent galaxies that have not yet seeded a black hole, or cases where the black hole mass is below the simulation's resolution or seeding threshold (Bhowmick et al. 2023; Ni et al. 2024; Despali et al. 2025). Including these galaxies directly in a log-space regression of the $M_{\rm BH}$–$M_{\rm star}$ relation would introduce a strong bias towards lower $M_{\rm BH}$ values, particularly at low stellar masses (Habouzit et al. 2021).

Therefore, our primary analysis focuses on the scaling relation among galaxies with detected black holes (Ni et al. 2024).

2.3.1. *Exclusion from Regression*

For the core analysis of the $M_{\rm BH}$–$M_{\rm star}$ relation's slope, normalization, and scatter, we exclude all galaxies with $M_{\rm BH} = 0$ (Reines & Volonteri 2015; Li et al. 2023). This approach is justified by the focus on characterizing the relation where a black hole is present and resolved, and by the large sample sizes available in each stellar mass bin for galaxies with $M_{\rm BH} > 0$ (Reines & Volonteri 2015).

2.3.2. *Black Hole Occupation Fraction*

To provide context for the exclusion and to understand the prevalence of black holes across the parameter space, we perform a secondary analysis quantifying the black hole occupation fraction. This is defined as the fraction of galaxies in a given bin (e.g., stellar mass bin, within a specific catalog) that have $M_{\rm BH} > 0$ (Gallo et al. 2023; Tremmel et al. 2024). We analyze how this fraction varies with stellar mass and the simulation parameters (Tremmel et al. 2024; Burke et al. 2025).

2.3.3. *Robustness Check (Optional)*

As an optional robustness check, we may explore the use of censored regression techniques (e.g., a Tobit model) on a subset of catalogs (Martin & Mortlock 2024; Jing & Li 2024). This method can account for the lower limit imposed by non-detections ($M_{\rm BH} = 0$) and provide an estimate of the relation parameters that incorporates these galaxies (Martin & Mortlock 2024; Jing & Li 2024). However, given the computational cost and the primary focus on the resolved relation, this is considered a supplementary analysis.

2.4. *Galaxy Sample Stratification*

To investigate potential mass-dependent variations in the $M_{\rm BH}$–$M_{\rm star}$ relation and its dependence on simulation parameters, we stratify the galaxy sample based on stellar mass. Based on preliminary exploratory data analysis (EDA) and physical considerations (e.g., transition from SN-dominated to AGN-dominated feedback regimes) (Zhang et al. 2023; Sturm & Reines 2024; Pacucci & Loeb 2024), we define three stellar mass bins:

- Low Mass: $M_{\rm star} < 10^9\,M_\odot$
- Intermediate Mass: $10^9 \leq M_{\rm star} < 10^{10}\,M_\odot$
- High Mass: $M_{\rm star} \geq 10^{10}\,M_\odot$

These bins are chosen to be well-populated across the simulation suite and to potentially isolate different physical processes.

Additionally, we may perform further stratification by star formation rate (SFR) to explore secondary dependencies on galaxy activity. This could involve dividing galaxies within each mass bin into quiescent and star-forming populations based on a threshold in specific SFR (Vilella-Rojo et al. 2021; Murrell & Baldry 2025).

2.5. *Per-Catalog Regression Analysis*

The core of our analysis involves characterizing the $M_{\rm BH}$–$M_{\rm star}$ relation for each individual simulation catalog (Ding et al. 2019; Zhang et al. 2023; Pacucci & Loeb 2024). For each catalog and within each defined stellar mass bin, we perform the following steps:

1. Select galaxies belonging to the current catalog and mass bin.
2. Filter this selection to include only galaxies with $M_{\rm BH} > 0$.
3. Fit a linear model in log-space using Ordinary Least Squares (OLS) regression:
$$\log_{10}\left(\frac{M_{\rm BH}}{M_\odot}\right) = \alpha + \beta \log_{10}\left(\frac{M_{\rm star}}{M_\odot}\right)$$
Here, β represents the slope of the relation and α represents the normalization (the intercept at $\log_{10} M_{\rm star} = 0$, or more physically, $\log_{10} M_{\rm BH}$ at $M_{\rm star} = 1\,M_\odot$) (Reines & Volonteri 2015; Chen et al. 2025).
4. Record the best-fit values for the slope (β) and normalization (α).
5. Compute the intrinsic scatter of the relation within this bin. This is quantified as the standard deviation of the residuals ($\log_{10} M_{\rm BH,observed} - \log_{10} M_{\rm BH,predicted}$) from the OLS fit.
6. We flag results from bins with a small number of galaxies ($N < 20$) as potentially unreliable due to insufficient statistics.

This process yields a set of (β, α, scatter) values for each catalog and each stellar mass bin, representing the characteristics of the $M_{\rm BH}$–$M_{\rm star}$ relation under that specific set of simulation parameters (Zhu et al. 2020; Zhang et al. 2023; Sturm & Reines 2024).

2.6. *Mapping Relation Parameters Across Parameter Space*

Having characterized the $M_{\rm BH}$–$M_{\rm star}$ relation for each catalog, we then analyze how the derived parameters (β, α, scatter) depend on the six varied simulation parameters (Ω_m, σ_8, $A_{\rm SN1}$, $A_{\rm SN2}$, $A_{\rm AGN1}$, $A_{\rm AGN2}$) (Winkel et al. 2024; Chen et al. 2025).

2.6.1. *Parameter Dependence Analysis*

For each relation parameter (slope, normalization, scatter) and each stellar mass bin, we treat the parameter value as a dependent variable and the six simulation parameters as independent variables. We employ multivariate analysis techniques to quantify these dependencies (Recio-Blanco et al. 2005; Accurso et al. 2017).

- **Multiple Linear Regression:** We fit linear models of the form $P = c_0 + \sum_{i=1}^{6} c_i X_i$, where P is the relation parameter (β, α, or scatter), and X_i are the six simulation parameters. The coefficients c_i indicate the linear effect of each parameter on the relation property. Standardized coefficients can be used to compare the relative importance of different parameters.
- **Random Forest Regression:** This non-linear regression technique can capture more complex interactions and non-linear dependencies. It also provides a robust measure of feature importance, indicating which simulation parameters are most influential in determining the variation in the relation properties.
- **Partial Correlation Analysis:** We compute the partial correlation coefficient between each simulation parameter and each relation property, controlling for the effects of the other five simulation parameters. This helps isolate the unique contribution of each parameter.

2.6.2. *Visualization*

We visualize the dependencies using various plotting techniques, including scatter plots showing the relation parameter vs. individual simulation parameters (Kent

2017; Lan et al. 2021), heatmaps illustrating correlations (Kent 2017; Lan et al. 2021), and partial dependence plots from the Random Forest analysis to show the marginal effect of each parameter while averaging over others (Lan et al. 2021).

2.7. *Secondary Analyses*

In addition to the core analysis of the $M_{\rm BH}$–$M_{\rm star}$ relation parameters, we conduct secondary analyses to provide a more complete picture (Pacucci & Loeb 2024; Juodžbalis et al. 2024).

2.7.1. *Black Hole Occupation Fraction*

As mentioned in Section 2.3.2, we quantify the black hole occupation fraction ($f_{\rm occ}$) for each catalog and stellar mass bin (Gallo et al. 2023; Tremmel et al. 2024). We then analyze how $f_{\rm occ}$ depends on the simulation parameters, particularly the feedback parameters, using similar multivariate techniques as for the relation parameters (Tremmel et al. 2024). This helps understand how the seeding and initial growth of black holes are affected by the varied physics (Tremmel et al. 2024).

2.7.2. *SFR Dependence*

Within each stellar mass bin, we investigate whether there is a residual dependence of the $M_{\rm BH}$–$M_{\rm star}$ relation on galaxy SFR, after accounting for stellar mass. This can be done by examining the residuals of the $M_{\rm BH}$–$M_{\rm star}$ fit as a function of SFR, or by including SFR as an additional predictor in a multivariate regression for $M_{\rm BH}$ within each mass bin (Pacucci & Loeb 2024).

2.8. *Computational Implementation*

Given the large number of simulation catalogs and galaxies, computational efficiency is crucial. The per-catalog analysis (Section 2.5) is inherently parallelizable, as each catalog can be processed independently. We leverage this by distributing the analysis across multiple CPU cores (e.g., 8 cores) using parallel processing libraries such as 'multiprocessing' or 'joblib' (Singh et al. 2013; Martin et al. 2018). Catalogs are processed in batches to balance memory usage and CPU load. Efficient data access, including chunked reading and caching of intermediate selections, is employed to avoid memory bottlenecks (Mighell 2010).

We profile the runtime of the analysis for a single catalog and scale the approach to ensure the full suite can be processed within reasonable time constraints.

2.9. *Summary Workflow*

The overall analysis workflow can be summarized as follows:

1. Load and prepare the galaxy and catalog data, selecting relevant features.
2. For each simulation catalog:
3. For each stellar mass bin:
4. Select galaxies in the current catalog and mass bin with $M_{\rm BH} > 0$.
5. Fit the $\log_{10} M_{\rm BH}$ vs. $\log_{10} M_{\rm star}$ relation using OLS regression.
6. Record the slope, normalization, and scatter of the fit.
7. Compute the black hole occupation fraction for the bin.
8. Aggregate the relation parameters and occupation fractions for all catalogs and bins.
9. Analyze the dependence of the relation parameters and occupation fraction on the six simulation parameters using multivariate regression and correlation techniques.
10. Visualize the results and interpret the influence of cosmological and feedback parameters on the $M_{\rm BH}$–$M_{\rm star}$ relation and black hole occupation.

This systematic approach allows us to map the diversity of the $M_{\rm BH}$–$M_{\rm star}$ relation across the explored parameter space and identify the key physical drivers of its variations (Pacucci & Loeb 2024; Juodžbalis et al. 2024; Chen et al. 2025).

3. RESULTS

This section presents the results of our systematic investigation into the diversity of the black hole–stellar mass ($M_{\rm BH}$–$M_{\rm star}$) relation across a wide range of cosmological and feedback parameter space. We analyze a suite of 1,000 simulated galaxy catalogs, each representing a unique combination of cosmological parameters (Ω_m, σ_8) and feedback parameters (A_SN1, A_SN2, A_AGN1, A_AGN2). Our goal is to quantify how the slope, normalization, and scatter of the $M_{\rm BH}$–$M_{\rm star}$ relation, as well as the black hole occupation fraction, depend on these underlying physical parameters.

3.1. *overview and methodological recap*

For each simulated galaxy catalog, we fit the linear relation $\log_{10}(M_{\rm BH}/M_\odot) = \alpha + \beta \log_{10}(M_{\rm star}/M_\odot)$ within three distinct stellar mass bins: low mass ($M_{\rm star} < 10^9\,M_\odot$), intermediate mass ($10^9 \le M_{\rm star} < 10^{10}\,M_\odot$),

and high mass ($M_{\rm star} \geq 10^{10}\,M_\odot$). The fitting is performed only on galaxies hosting a black hole ($M_{\rm BH} > 0$). We determine the best-fit slope (β), normalization (α, evaluated at $\log_{10} M_{\rm star} = 0$), and intrinsic scatter (measured as the standard deviation of the residuals in $\log_{10} M_{\rm BH}$). We also quantify the black hole occupation fraction in each mass bin for each catalog.

To understand the dependence of these relation properties (slope, normalization, scatter, occupation fraction) on the six input catalog parameters (Ω_m, σ_8, A_SN1, A_SN2, A_AGN1, A_AGN2), we employ both multivariate linear regression and random forest analysis. Linear regression provides standardized coefficients indicating the direction and relative strength of linear dependencies, while random forest analysis provides a non-parametric measure of feature importance, capturing non-linear effects and interactions. Uncertainties on the fitted parameters are estimated via bootstrapping.

3.2. *diversity of the $M_{\rm BH}$–$M_{\rm star}$ relation across mass bins*

3.2.1. *distributions of slope, normalization, and scatter*

The $M_{\rm BH}$–$M_{\rm star}$ relation exhibits significant diversity across the 1,000 simulated catalogs, with the characteristics of the relation varying substantially depending on the stellar mass range considered.

Low Mass Bin ($M_{\rm star} < 10^9\,M_\odot$):—In the low mass regime, the distribution of fitted slopes is centered around a mean of 0.20 with a standard deviation of 0.09, spanning a range from −0.07 to 0.58. The normalization α has a mean of 4.26, a large standard deviation of 0.72, and ranges from 1.24 to 6.54. The intrinsic scatter in $\log_{10} M_{\rm BH}$ is relatively low, with a mean of 0.11 dex and a standard deviation of 0.05 dex. The black hole occupation fraction in this bin shows considerable variation across catalogs, with a mean of 0.72 and a standard deviation of 0.13.

Intermediate Mass Bin ($10^9 \leq M_{\rm star} < 10^{10}\,M_\odot$):—For intermediate mass galaxies, the slope distribution shifts towards steeper values, with a mean of 1.06 and a standard deviation of 0.42, ranging from 0.02 to 1.84. The normalization α shows even greater diversity than in the low mass bin, with a mean of −3.53 and a very large standard deviation of 3.80, spanning from −10.6 to 6.05. The distribution of α for this mass bin is illustrated in Figure 1, highlighting the wide range of values obtained across the parameter space. The intrinsic scatter increases significantly compared to the low mass bin, with a mean of 0.30 dex and a standard deviation of 0.06 dex. Figure 2 shows the distribution of scatter values in this bin, centered around ∼ 0.32 dex. The occupation fraction is higher and less variable, with a mean of 0.85 and a standard deviation of 0.08.

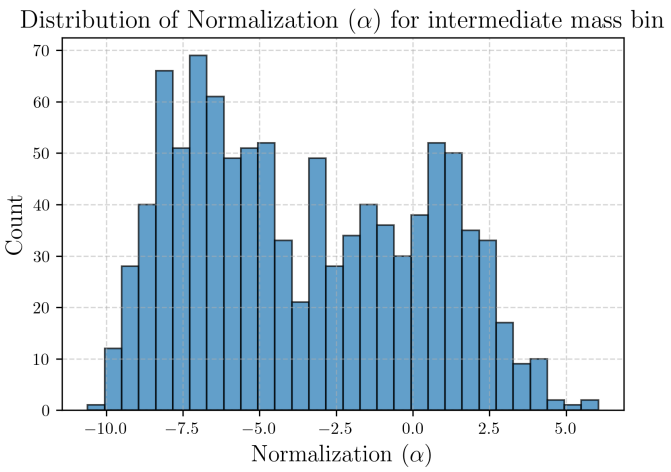

Figure 1. Histogram showing the distribution of the normalization parameter α for the intermediate mass bin. The distribution spans a wide range of values for α, indicating large differences in this parameter for objects within this mass range.

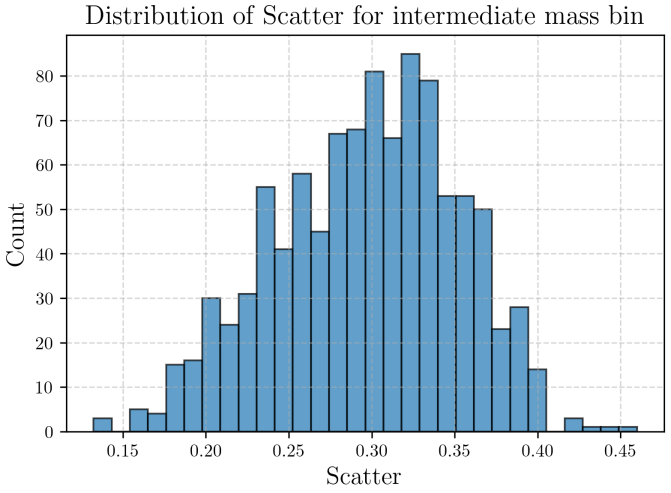

Figure 2. Histogram showing the distribution of Scatter for the intermediate mass bin. The distribution is centered around a value of approximately 0.32, with values ranging from about 0.14 to 0.46. The spread of the distribution indicates that a range of scatter values are found within this mass bin.

High Mass Bin ($M_{\rm star} \geq 10^{10}\,M_\odot$):—In the high mass bin, the mean slope is 1.22 with a standard deviation of 0.30, ranging from 0.03 to 2.07. The normalization α continues to exhibit large variations, with a mean of −4.85 and a standard deviation of 3.39, ranging from −14.2 to 7.64. The intrinsic scatter is highest in this

bin, with a mean of 0.40 dex and a standard deviation of 0.14 dex. The occupation fraction is close to unity and shows minimal variation, with a mean of 0.96 and a standard deviation of 0.03.

3.2.2. *physical regimes*

The observed trends in slope, normalization, scatter, and occupation fraction across mass bins suggest distinct physical regimes governing black hole and galaxy growth. The shallow mean slope (∼0.2) and relatively high normalization at low stellar masses indicate a regime where black hole growth is likely inefficient, stochastic, or decoupled from stellar mass assembly, possibly dominated by initial seeding conditions and strong supernova (SN) feedback. As stellar mass increases, the slope approaches or exceeds unity, consistent with the tightly coupled coevolution observed in more massive galaxies. The increasing scatter with stellar mass, coupled with the large diversity in normalization and slope across catalogs, points to a strong sensitivity to the specific feedback and cosmological parameters in the intermediate and high mass regimes. The rise in occupation fraction with mass reflects the increasing likelihood of a galaxy hosting a black hole as its potential well deepens and its formation history becomes more conducive to black hole seeding and retention.

3.3. *dependence on feedback and cosmological parameters*

We now detail the influence of the six input parameters on the properties of the $M_{\rm BH}$–$M_{\rm star}$ relation and the occupation fraction, as revealed by multivariate linear regression and random forest analysis.

3.3.1. *slope (β)*

Low Mass Bin:—In the low mass bin, the slope β is primarily influenced by SN feedback (A_SN1) and AGN feedback (A_AGN1). Linear regression shows positive standardized coefficients for both (+0.05 for A_SN1, +0.04 for A_AGN1), indicating that stronger feedback of either type tends to steepen the relation at low masses. Random forest analysis confirms their importance, with A_SN1 having the highest feature importance (0.41), followed by A_AGN1 (0.29). Cosmological parameters have weaker effects. The partial dependence of the slope on these parameters is shown in Figure 3, illustrating how varying A_SN1 and A_AGN1 across their range leads to larger changes in β compared to other parameters.

Intermediate Mass Bin:—The intermediate mass slope is strongly dominated by A_AGN1, which has a large positive standardized coefficient (+0.37) and the highest random forest importance (0.88). This is clearly shown in the standardized linear feature importance plot (Figure 4), where A_AGN1 stands out as significantly more important than other parameters. This suggests that stronger AGN feedback, particularly the mode parameterized by A_AGN1, drives a steeper $M_{\rm BH}$–$M_{\rm star}$ relation in this mass range. A_SN1 and cosmological parameters have significantly smaller effects.

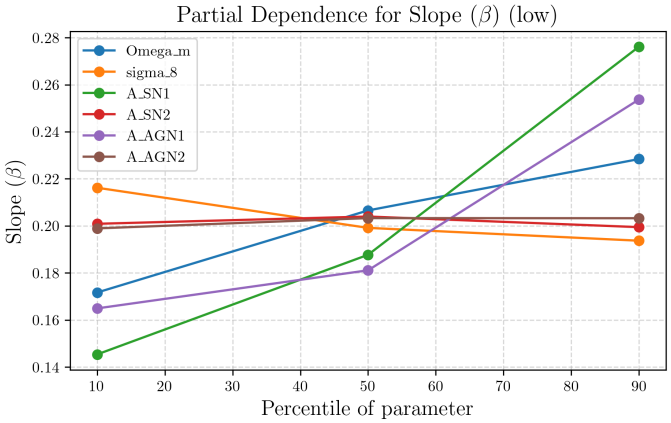

Figure 3. Partial dependence plot showing the variation of the slope β as a function of the percentile of different cosmological and astrophysical parameters. Each line represents the partial dependence for a specific parameter, as indicated in the legend, including cosmological parameters like $\Omega_{\rm m}$ and σ_8, and astrophysical parameters like A_SN1, A_SN2, A_AGN1, and A_AGN2. The plot reveals that some parameters, such as A_SN1, $\Omega_{\rm m}$, and A_AGN1, exhibit a strong dependence on their value, leading to large differences in the resulting slope β across their range. Other parameters, like A_SN2, A_AGN2, and σ_8, show a weaker dependence, resulting in smaller differences in β.

High Mass Bin:—In contrast to the intermediate bin, A_AGN1 has a dominant *negative* standardized coefficient (−0.22) and the highest random forest importance (0.69) for the slope in the high mass bin, as seen in Figure 5. The partial dependence plot in Figure 6 further illustrates this strong dependence on A_AGN1. This sign flip indicates a transition in the role of AGN feedback: while it promotes steeper slopes (likely by driving black hole growth relative to stellar growth) at intermediate masses, it leads to shallower slopes (possibly by suppressing black hole growth or enhancing stellar growth relative to black hole growth) at high masses. A_SN1 and cosmological parameters are less influential.

3.3.2. *normalization (α)*

Low Mass Bin:—The normalization α in the low mass bin is strongly negatively correlated with both A_SN1 (standardized coefficient −0.38) and A_AGN1 (−0.32). This is supported by the feature importance plots (Fig-

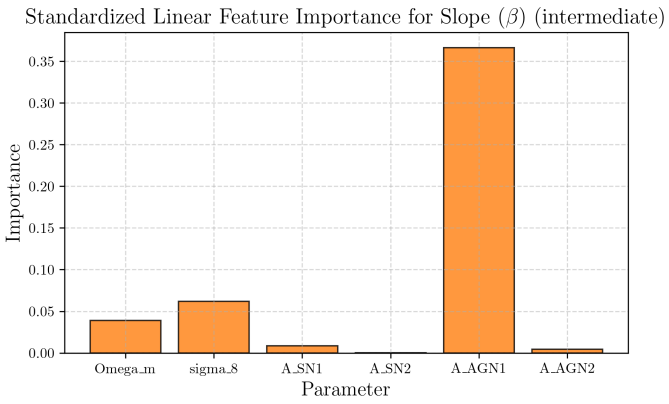

Figure 4. Standardized linear feature importance for the slope (β) in the intermediate case. The figure shows the relative importance of different parameters (Omega_m, σ_8, A_SN1, A_SN2, A_AGN1, A_AGN2) in determining the slope. A large difference in importance is observed, with the parameter A_AGN1 showing significantly higher importance compared to all other parameters.

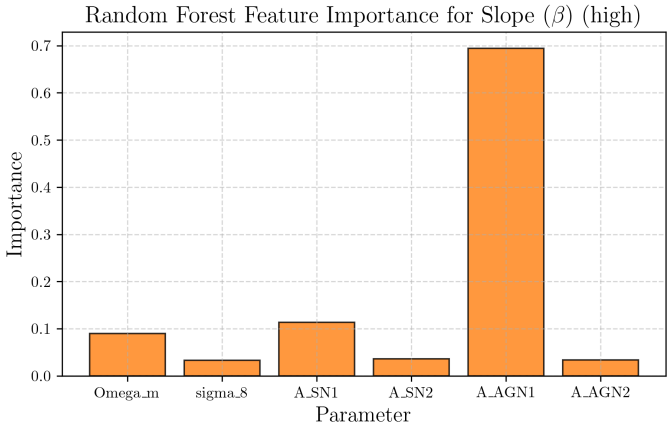

Figure 5. Figure showing the Random Forest feature importance for the high slope (β). The importance of various cosmological (Ω_m, σ_8) and astrophysical ($A_{\rm SN1}$, $A_{\rm SN2}$, $A_{\rm AGN1}$, $A_{\rm AGN2}$) parameters is displayed. Large differences in importance are evident, with $A_{\rm AGN1}$ being the dominant feature, while others exhibit significantly lower importance.

ure 7 and Figure 8), which show A_SN1 and A_AGN1 as the most important parameters for α in this mass range. This means that stronger feedback, both SN and AGN, tends to lower the normalization, resulting in less massive black holes at a fixed stellar mass of 1 $M_\odot$. Ω_m also shows a negative correlation.

Intermediate Mass Bin:—A_AGN1 is the dominant driver of normalization diversity in the intermediate mass bin, with a large negative standardized coefficient (-3.32). Its overwhelming importance is evident in the random forest feature importance plot (Figure 9)

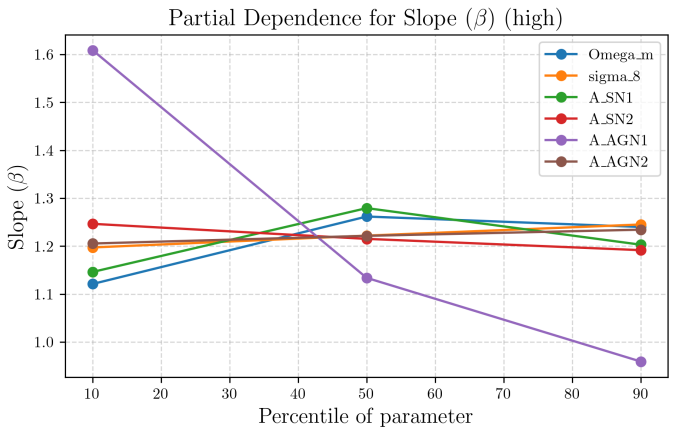

Figure 6. Partial dependence of the slope β on the percentiles of different parameters. The lines show how the value of β changes when a specific parameter is varied across its 10th, 50th, and 90th percentiles, while other parameters are held at their median values. Large differences in the slope β are observed for some parameters, such as A_AGN1, while others show smaller variations.

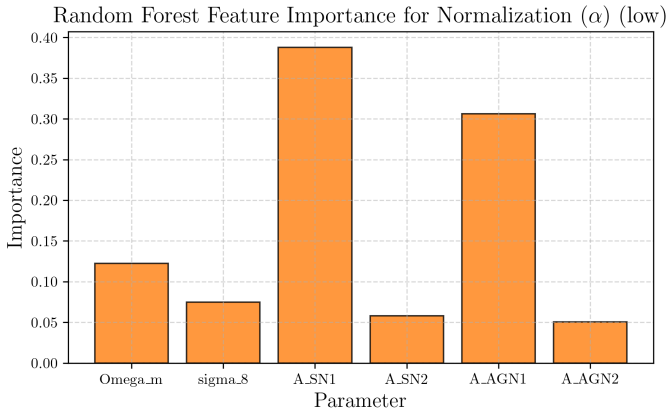

Figure 7. Random Forest feature importance for the normalization parameter α in the low regime. The bar chart shows the relative importance of different cosmological and astrophysical parameters (Omega_m, sigma_8, A_SN1, A_SN2, A_AGN1, A_AGN2) as determined by a Random Forest model. Large differences in importance are seen, with parameters A_SN1 and A_AGN1 showing significantly higher importance compared to others like sigma_8, A_SN2, and A_AGN2.

and the partial dependence plot (Figure 10). Stronger A_AGN1 leads to significantly lower normalizations. σ_8 and Ω_m also contribute, with higher values generally lowering α.

High Mass Bin:—Similar to the slope, the effect of A_AGN1 on normalization flips sign in the high mass bin, showing a large positive standardized coefficient ($+2.52$). Its dominance is again highlighted in the ran-

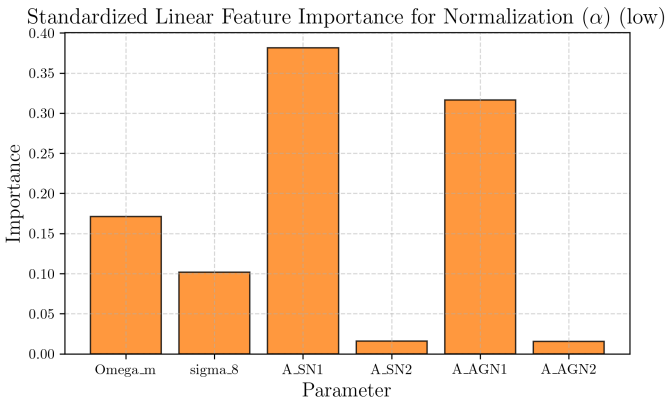

Figure 8. Standardized linear feature importance for normalization (α) in the low-α regime. Significant differences in importance are evident across the parameters, with A_SN1 and A_AGN1 being the most important, Omega_m and sigma_8 moderately important, and A_SN2 and A_AGN2 having negligible importance.

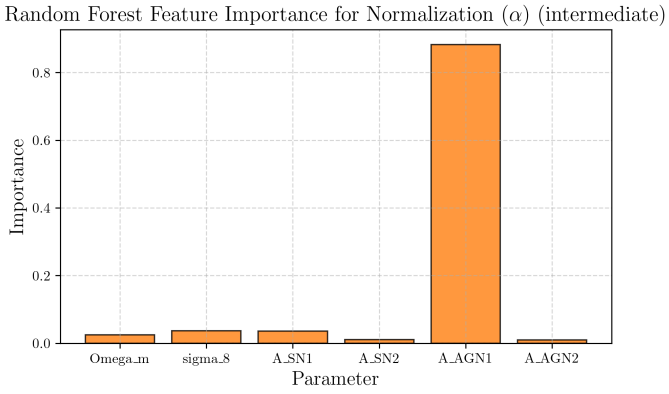

Figure 9. Random Forest feature importance for the normalization parameter α. The figure displays the relative importance of various cosmological and astrophysical parameters (Omega_m, sigma_8, A_SN1, A_SN2, A_AGN1, A_AGN2) in predicting the value of α using a Random Forest model. Large differences in importance are observed, with the parameter A_AGN1 showing significantly higher importance compared to all other parameters, which exhibit very low importance.

dom forest feature importance plot (Figure 11). This implies that stronger A_AGN1 leads to *higher* normalizations at high masses, again suggesting a complex, mass-dependent role for this feedback mode. σ_8 and Ω_m also have notable effects, with higher σ_8 tending to increase α and higher Ω_m tending to decrease it.

3.3.3. *scatter*

Low Mass Bin:—The intrinsic scatter in the low mass bin is primarily driven by A_SN1, which has a positive standardized coefficient (+0.04) and the highest random

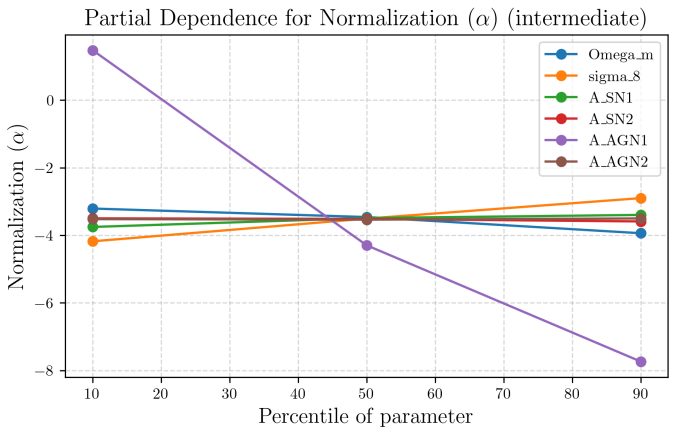

Figure 10. Partial dependence of the normalization parameter α. The plot shows how α changes as each individual parameter (Ω_m, σ_8, A_SN1, A_SN2, A_AGN1, and A_AGN2) is varied across its 10th, 50th, and 90th percentiles, while other parameters are held at their median. A large dependence is observed for the parameter A_AGN1, where α varies significantly. In contrast, the parameters Ω_m, σ_8, A_SN1, A_SN2, A_AGN1, and A_AGN2 show relatively small dependence on α.

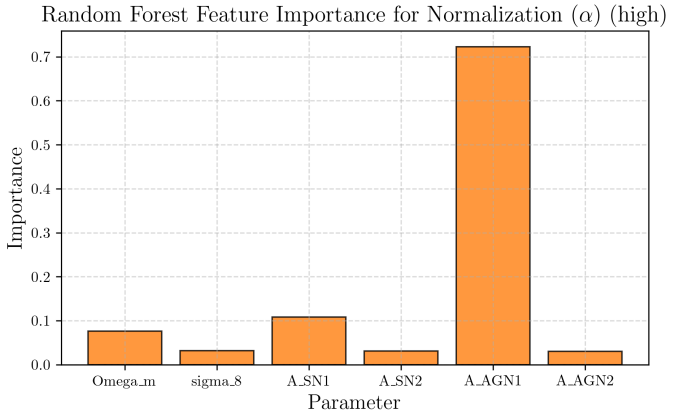

Figure 11. Random Forest feature importance for predicting the normalization parameter α. The figure shows the relative importance of different cosmological and astrophysical parameters (Omega_m, sigma_8, A_SN1, A_SN2, A_AGN1, A_AGN2) as determined by a Random Forest model. A large difference in importance is observed, with the parameter A_AGN1 showing significantly higher importance compared to all other parameters.

forest importance (0.73), as shown in Figure 12 and Figure 13. The partial dependence plot (Figure 14) confirms the strong influence of A_SN1 and Ω_m on scatter in this regime. This is consistent with the expectation that stochastic SN feedback introduces significant diversity in the early growth phases of black holes in shallow potential wells.

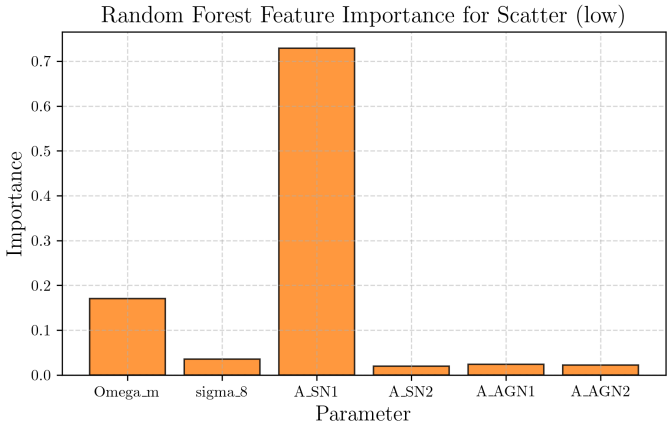

Figure 12. Feature importance of various parameters for predicting the scatter (low) using a Random Forest model. Large differences in importance are observed. The parameter A_SN1 shows the highest importance, followed by Ω_m. The other parameters (σ_8, A_SN2, A_AGN1, and A_AGN2) exhibit significantly lower importance.

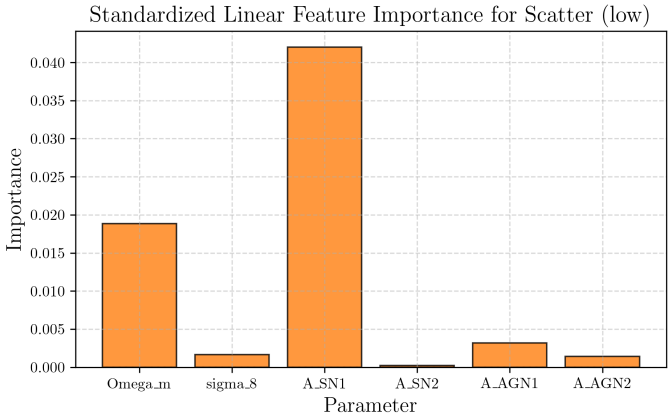

Figure 13. Standardized linear feature importance for various cosmological and astrophysical parameters, specifically for the low scatter case. The parameters shown are $\Omega_{\rm m}$, σ_8, and amplitudes for supernova (A_SN1, A_SN2) and AGN (A_AGN1, A_AGN2) feedback models. Large differences in importance are observed, with A_SN1 showing the highest importance, followed by $\Omega_{\rm m}$. The other parameters exhibit significantly lower importance.

Intermediate Mass Bin:—Both A_AGN1 and A_SN1 contribute to the scatter in the intermediate mass bin. A_AGN1 shows the highest random forest importance (0.44), suggesting it is a key driver of scatter in this regime, potentially due to variations in the efficiency or timing of AGN feedback relative to galaxy assembly.

High Mass Bin:—In the high mass bin, A_SN1 and A_AGN1 remain important drivers of scatter, with A_SN1 having a slightly higher random forest importance (0.31) than A_AGN1 (0.28). The standardized linear feature importance (Figure 15) also shows A_AGN1 and A_SN1 as the most important parameters for scatter in this bin. This indicates that both SN and AGN feedback contribute to the diversity around the mean relation even at high masses.

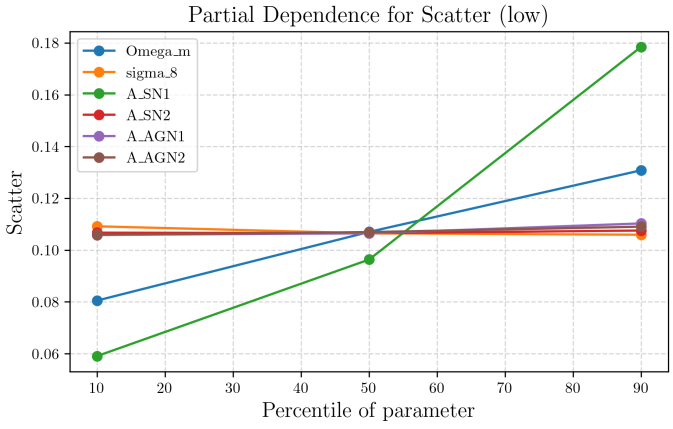

Figure 14. Partial dependence of the Scatter quantity on the percentile of various cosmological and nuisance parameters. The plot shows how the Scatter changes as each parameter is varied across its distribution (represented by percentiles), while other parameters are held at their median values. Large differences in Scatter are observed for the cosmological parameter $\Omega_{\rm m}$ and the supernova amplitude parameter $A_{\rm SN1}$, indicating a strong dependence. In contrast, the Scatter shows only small differences across the percentiles of σ_8, $A_{\rm SN2}$, $A_{\rm AGN1}$, and $A_{\rm AGN2}$, suggesting a weaker dependence on these parameters.

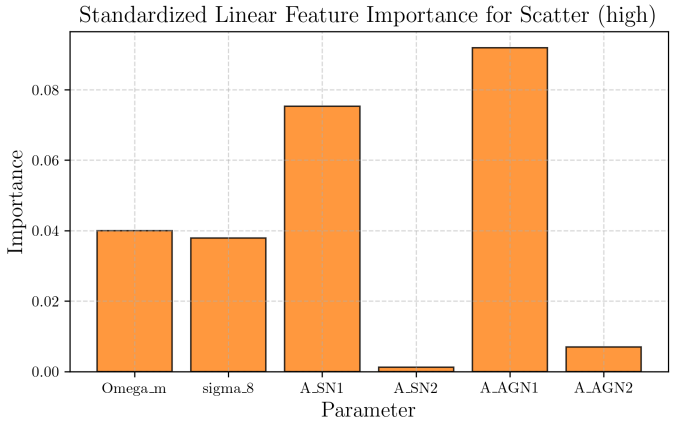

Figure 15. Standardized linear feature importance for scatter (high). The importance of different parameters (Ω_m, σ_8, $A_{\rm SN1}$, $A_{\rm SN2}$, $A_{\rm AGN1}$, $A_{\rm AGN2}$) is displayed. Large differences in importance are observed, with $A_{\rm AGN1}$ and $A_{\rm SN1}$ showing significantly higher importance compared to $A_{\rm SN2}$ and $A_{\rm AGN2}$. Ω_m and σ_8 have moderate importance.

3.3.4. *partial dependence and 2d interactions*

Partial dependence plots, such as those shown for slope and normalization in Figures 3, 6, and 10, reveal non-linear relationships between feedback parameters and the properties of the $M_{\rm BH}$–$M_{\rm star}$ relation. For instance, increasing A_AGN1 can initially steepen the slope and increase scatter, but these effects may saturate or even reverse at very high feedback strengths, depending on the mass bin. Two-dimensional heatmaps (not shown here) highlight interactions between parameters. For example, the joint effect of A_AGN1 and A_AGN2 on the high mass slope and normalization is significant, with the steepest slopes and highest normalizations often occurring when A_AGN1 is high and A_AGN2 is low, suggesting that the relative balance between different AGN feedback modes is crucial at high masses.

3.3.5. *cosmological parameters*

While feedback parameters are the dominant drivers of diversity, cosmological parameters Ω_m and σ_8 have secondary but non-negligible effects. As seen in Figures 3 and 14, higher Ω_m generally leads to lower normalizations and increased scatter across mass bins. Higher σ_8 can either increase or decrease normalization depending on the mass bin and interacts with feedback parameters in complex ways, particularly influencing the normalization and scatter in the intermediate and high mass regimes.

3.3.6. *summary of feature importance*

In summary, the analysis clearly delineates the primary drivers of $M_{\rm BH}$–$M_{\rm star}$ relation diversity across mass scales:

- **Low Mass:** SN feedback (A_SN1) is the dominant factor influencing slope, normalization, and scatter (Figures 3, 7, 8, 12, 13, 14).
- **Intermediate Mass:** AGN feedback (A_AGN1) becomes the primary driver of slope and normalization (Figures 4, 9, 10), with SN feedback still contributing to scatter.
- **High Mass:** AGN feedback (A_AGN1) remains dominant (Figures 5, 6, 11), but its effect on slope and normalization reverses sign, indicating a transition from promoting black hole growth relative to stellar mass to potentially regulating or suppressing it. Both A_SN1 and A_AGN1 contribute significantly to scatter (Figure 15).

Cosmological parameters play a modulating role, particularly affecting normalization and scatter at higher masses.

3.4. *black hole occupation fraction*

3.4.1. *trends with mass and feedback*

The black hole occupation fraction, defined as the fraction of galaxies hosting a black hole with $M_{\rm BH} > 0$, shows a strong dependence on stellar mass and feedback parameters. It rises from a mean of $\sim$0.72 in the low mass bin to $\sim$0.85 in the intermediate bin and nearly unity ($\sim$0.96) in the high mass bin.

Crucially, the occupation fraction, particularly at low and intermediate masses, is sensitive to feedback. Stronger SN feedback (A_SN1) significantly reduces the occupation fraction in the low and intermediate mass bins. This is consistent with scenarios where energetic SN-driven winds can expel gas necessary for black hole formation or early growth, or even eject nascent black holes from shallow potential wells. AGN feedback has a weaker effect on occupation at low mass but becomes more relevant at higher masses, although its impact on occupation is less pronounced than its impact on the scaling relation properties for black hole hosts.

3.4.2. *implications for black hole seeding*

The incomplete occupation fraction at low stellar masses and its strong dependence on feedback parameters provide insights into black hole seeding mechanisms. The results support models where seeding is not universally efficient in all halos, and where early feedback processes, particularly from supernovae, can prevent or suppress the formation or retention of black holes in low-mass galaxies. The significant catalog-to-catalog variation in occupation fraction at low mass highlights that the fraction of galaxies hosting a black hole is a sensitive probe of the interplay between seeding physics and early feedback.

3.5. *secondary dependencies and sfr*

While our primary analysis focused on the $M_{\rm BH}$–$M_{\rm star}$ relation and its dependence on the six input parameters, we note that galaxy properties like star formation rate (SFR) are also influenced by feedback. In our simulations, SFR shows a weak correlation with $M_{\rm star}$ ($r = 0.10$) and is also affected by feedback parameters. Although we did not explicitly stratify the $M_{\rm BH}$–$M_{\rm star}$ relation by SFR, the dominant influence of feedback parameters on both $M_{\rm BH}$ and $M_{\rm star}$ suggests that the primary drivers of the $M_{\rm BH}$–$M_{\rm star}$ relation diversity are the direct effects of feedback on black hole and stellar growth, rather than indirect effects mediated solely through SFR. Future work could explore the $M_{\rm BH}$–$M_{\rm star}$–SFR fundamental plane across this parameter space.

3.6. *comparison with observational constraints*

Comparing our simulated results to observational constraints provides a crucial validation step. The mean slopes ($\beta \sim 1.1$–1.2) and scatters ($\sim$0.3–0.4 dex) in the intermediate and high mass bins are broadly consistent with the canonical $M_{\rm BH}$–$M_{\rm star}$ relation observed in local massive galaxies (e.g., Kormendy & Ho 2013; McConnell & Ma 2013). However, the simulated catalogs exhibit a significantly wider range of slopes, normalizations, and scatters than typically reported from single observational samples, particularly at low masses. This suggests that the observed scatter and potential secondary dependencies in the real universe might be partly attributable to variations in feedback efficiency and cosmological environment. The simulated occupation fractions at high mass are consistent with observational inferences of near-unity occupation in massive galaxies. At low masses, the simulated occupation fractions are lower and highly variable, aligning with recent observational efforts suggesting incomplete occupation in dwarf galaxies (e.g., Greene et al. 2020).

3.7. *synthesis and physical interpretation*

Our results provide a comprehensive picture of how feedback and cosmology shape the $M_{\rm BH}$–$M_{\rm star}$ relation. The diversity observed across the parameter space is primarily driven by feedback, with a clear transition in the dominant feedback mode and its effect from low to high stellar masses. SN feedback is crucial in the low mass regime, influencing both the scaling relation properties and the likelihood of a galaxy hosting a black hole. AGN feedback becomes increasingly important at higher masses, driving the slope and normalization, but its role transitions from promoting relative black hole growth at intermediate masses to potentially regulating or suppressing it. Cosmological parameters modulate these relations, highlighting the interconnectedness of baryonic and dark matter physics in galaxy evolution. The sensitivity of the low-mass occupation fraction to feedback provides a direct link to black hole seeding physics, suggesting that early feedback can be a critical factor in determining whether a galaxy hosts a black hole. The broad range of relation properties across the simulated catalogs suggests that the observed scatter and potential secondary dependencies in the $M_{\rm BH}$–$M_{\rm star}$ relation are natural outcomes of variations in the underlying physical processes governing galaxy and black hole coevolution.

4. CONCLUSIONS

This work presents the first systematic exploration of the diversity of the black hole–stellar mass ($M_{\rm BH}$–$M_{\rm star}$) relation across a multi-dimensional parameter space encompassing key cosmological and baryonic feedback parameters in a large suite of hydrodynamical simulations. By quantifying the slope, normalization, and scatter of the relation, as well as the black hole occupation fraction, for 1,000 distinct simulation catalogs, we have mapped how these properties depend on variations in Ω_m, σ_8, and parameters governing stellar and AGN feedback ($A_{\rm SN1}$, $A_{\rm SN2}$, $A_{\rm AGN1}$, $A_{\rm AGN2}$).

Our analysis, performed by fitting the $M_{\rm BH}$–$M_{\rm star}$ relation in stellar mass bins ($< 10^9\,M_\odot$, 10^9–$10^{10}\,M_\odot$, $> 10^{10}\,M_\odot$) for galaxies with resolved black holes, reveals substantial diversity in the relation's properties across the explored parameter space. The mean slope transitions from shallow (~ 0.2) at low stellar masses to steeper (~ 1.1–1.2) at intermediate and high masses, consistent with a shift in the dominant physical processes governing black hole growth and galaxy evolution. The intrinsic scatter in the relation increases with stellar mass, from ~ 0.1 dex at low mass to ~ 0.4 dex at high mass. The black hole occupation fraction rises from $\sim 72\%$ at low mass to nearly 96% at high mass, exhibiting significant catalog-to-catalog variation, particularly at lower masses.

Using multivariate regression and feature importance analysis, we identified the primary drivers of this diversity. Baryonic feedback parameters, especially the efficiency of supernova feedback ($A_{\rm SN1}$) and AGN feedback ($A_{\rm AGN1}$), are the dominant factors shaping the $M_{\rm BH}$–$M_{\rm star}$ relation and the occupation fraction. The influence of feedback is strongly mass-dependent: SN feedback is most impactful at low stellar masses, primarily affecting the normalization, scatter, and occupation fraction. AGN feedback becomes increasingly important at intermediate and high masses, driving variations in slope, normalization, and scatter. Notably, the sign of the dependence on $A_{\rm AGN1}$ flips between the intermediate and high mass bins, suggesting a transition from AGN-driven black hole growth and coevolution to AGN-driven quenching and regulation of both black hole and galaxy growth at the highest masses. Cosmological parameters (Ω_m, σ_8) play a secondary but non-negligible role, primarily modulating the normalization and scatter of the relation.

The sensitivity of the black hole occupation fraction to feedback parameters, particularly the suppression by strong SN feedback at low masses, provides crucial insights into black hole seeding and early growth mechanisms. Our results support scenarios where energetic feedback can prevent or delay black hole formation or growth in shallow potential wells, leading to the incomplete and variable occupation fractions observed in the

simulations and increasingly inferred for low-mass galaxies in the real universe.

Overall, this study demonstrates that the observed diversity and intrinsic scatter in the $M_{\rm BH}$–$M_{\rm star}$ relation are natural consequences of variations in the efficiency and implementation of baryonic feedback processes within the cosmological context. The broad range of scaling relation properties predicted by the simulations highlights the importance of accurately modeling feedback physics to understand black hole–galaxy coevolution and provides a framework for interpreting observational constraints across different galaxy populations and cosmic epochs. This work underscores the power of large-scale parameter studies in simulations to disentangle the complex interplay of physical processes that shape fundamental galaxy scaling relations.

REFERENCES

Accurso, G., Saintonge, A., Catinella, B., et al. 2017, Deriving a multivariate CO-to-H_2 conversion function using the [CII]/CO(1-0) ratio and its application to molecular gas scaling relations, doi: https://doi.org/10.1093/mnras/stx1556

Bahar, Y. E., Bulbul, E., Clerc, N., et al. 2022, The eROSITA Final Equatorial-Depth Survey (eFEDS): X-ray Properties and Scaling Relations of Galaxy Clusters and Groups, doi: https://doi.org/10.1051/0004-6361/202142462

Bhowmick, A. K., Blecha, L., Torrey, P., et al. 2023, Representing low mass black hole seeds in cosmological simulations: A new sub-grid stochastic seed model. https://arxiv.org/abs/2309.15341

Burke, C. J., Natarajan, P., Baldassare, V. F., & Geha, M. 2025, Multi-wavelength constraints on the local black hole occupation fraction. https://arxiv.org/abs/2410.11177

Chen, Y., Gu, Q., Fan, J., et al. 2025, The relation between black hole spin, star formation rate, and black hole mass for supermassive black holes. https://arxiv.org/abs/2503.03223

Ciotti, L., Ostriker, J. P., Gan, Z., et al. 2022, A Parameter Space Exploration of High Resolution Numerically Evolved Early Type Galaxies Including AGN Feedback and Accurate Dynamical Treatment of Stellar Orbits, doi: https://doi.org/10.3847/1538-4357/ac70c7

Despali, G., Moscardini, L., Nelson, D., et al. 2025, Introducing the AIDA-TNG project: galaxy formation in alternative dark matter models. https://arxiv.org/abs/2501.12439

Ding, X., Silverman, J., Treu, T., et al. 2019, The mass relations between supermassive black holes and their host galaxies at 1<z<2 with HST-WFC3, doi: https://doi.org/10.3847/1538-4357/ab5b90

D'Onofrio, M., Marziani, P., & Chiosi, C. 2021, Past, present and Future of the Scaling Relations of Galaxies and Active Galactic Nuclei. https://arxiv.org/abs/2109.06301

Gallo, E., Hodges-Kluck, E., Treu, T., et al. 2023, The black hole occupation fraction of local dwarf galaxies with AXIS. https://arxiv.org/abs/2311.09161

Giodini, S., Lovisari, L., Pointecouteau, E., et al. 2013, Scaling relations for galaxy clusters: properties and evolution, doi: https://doi.org/10.1007/s11214-013-9994-5

Habouzit, M., Li, Y., Somerville, R. S., et al. 2021, Supermassive black holes in cosmological simulations I: M_BH-M_star relation and black hole mass function, doi: https://doi.org/10.1093/mnras/stab496

Horowitz, B., & Lukic, Z. 2025, Differentiable Cosmological Hydrodynamics for Field-Level Inference and High Dimensional Parameter Constraints. https://arxiv.org/abs/2502.02294

Jahnke, K., & Maccio, A. 2011, The non-causal origin of the black hole-galaxy scaling relations, doi: https://doi.org/10.1088/0004-637X/734/2/92

Jing, T., & Li, C. 2024, Regression for Astronomical Data with Realistic Distributions, Errors and Non-linearity. https://arxiv.org/abs/2411.08747

Jung, M., Roca-Fàbrega, S., hoon Kim, J., et al. 2024, The AGORA High-resolution Galaxy Simulations Comparison Project. V: Satellite Galaxy Populations In A Cosmological Zoom-in Simulation of A Milky Way-mass Halo, doi: https://doi.org/10.3847/1538-4357/ad245b

Juodžbalis, I., Maiolino, R., Baker, W. M., et al. 2024, A dormant, overmassive black hole in the early Universe, doi: https://doi.org/10.1038/s41586-024-08210-5

Kent, B. R. 2017, Editorial: Techniques and Methods for Astrophysical Data Visualization, doi: https://doi.org/10.1088/1538-3873/aa5fa6

Lan, F., Young, M., Anderson, L., et al. 2021, Visualization in Astrophysics: Developing New Methods, Discovering Our Universe, and Educating the Earth. https://arxiv.org/abs/2106.00152

Li, J. I.-H., Shen, Y., Ho, L. C., et al. 2023, The Sloan Digital Sky Survey Reverberation Mapping Project: The Black Hole Mass−Stellar Mass Relations at $0.2 \lesssim z \lesssim 0.8$. https://arxiv.org/abs/2301.04177

Maity, B., Paranjape, A., & Choudhury, T. R. 2023, A fast method of reionization parameter space exploration using GPR trained SCRIPT, doi: https://doi.org/10.1093/mnras/stad2984

Martin, D., Jose, J., & Longland, R. 2018, On the parallelization of stellar evolution codes, doi: https://doi.org/10.1186/s40668-018-0025-5

Martin, W., & Mortlock, D. J. 2024, Robust Bayesian regression in astronomy. https://arxiv.org/abs/2411.02380

Mayer, L., Capelo, P. R., Zwick, L., & Matteo, T. D. 2023, Direct formation of massive black holes via dynamical collapse in metal-enriched merging galaxies at $z \sim 10$: fully cosmological simulations. https://arxiv.org/abs/2304.02066

McGibbon, R., & Khochfar, S. 2023, Multi-Epoch Machine Learning 2: Identifying physical drivers of galaxy properties in simulations, doi: https://doi.org/10.1093/mnras/stad1811

Mighell, K. J. 2010, CRBLASTER: A Parallel-Processing Computational Framework for Embarrassingly-Parallel Image-Analysis Algorithms, doi: https://doi.org/10.1086/656566

Murrell, G. G., & Baldry, I. K. 2025, Star formation rate density as a function of galaxy mass at z < 0.2 with MUSE and GAMA surveys. https://arxiv.org/abs/2410.08036

Ni, Y., Chen, N., Zhou, Y., et al. 2024, The Astrid Simulation: Evolution of black holes and galaxies to z=0.5 and different evolution pathways for galaxy quenching. https://arxiv.org/abs/2409.10666

Pacucci, F., & Loeb, A. 2024, The Redshift Evolution of the $M_\bullet - M_\star$ Relation for JWST's Supermassive Black Holes at $z > 4$, doi: https://doi.org/10.3847/1538-4357/ad3044

Recio-Blanco, A., Aparicio, A., Piotto, G., Angeli, F. D., & Djorgovski, S. G. 2005, Multivariate Analysis of Globular Cluster's Horizontal Branch Morphology: searching for the second parameter, doi: https://doi.org/10.1051/0004-6361:20053006

Reines, A. E., & Volonteri, M. 2015, Relations Between Central Black Hole Mass and Total Galaxy Stellar Mass in the Local Universe, doi: https://doi.org/10.1088/0004-637X/813/2/82

Schaye, J., Kugel, R., Schaller, M., et al. 2023, The FLAMINGO project: cosmological hydrodynamical simulations for large-scale structure and galaxy cluster surveys, doi: https://doi.org/10.1093/mnras/stad2419

Sicilia, A., Lapi, A., Boco, L., et al. 2021, The Black Hole Mass Function Across Cosmic Times I. Stellar Black Holes and Light Seed Distribution, doi: https://doi.org/10.3847/1538-4357/ac34fb

—. 2022, The Black Hole Mass Function Across Cosmic Times II. Heavy Seeds and (Super)Massive Black Holes, doi: https://doi.org/10.3847/1538-4357/ac7873

Singh, N., Browne, L.-M., & Butler, R. 2013, Parallel Astronomical Data Processing with Python: Recipes for multicore machines, doi: https://doi.org/10.1016/j.ascom.2013.04.002

Sturm, M. R., & Reines, A. E. 2024, A Breakdown of the Black Hole - Bulge Mass Relation in Local Active Galaxies. https://arxiv.org/abs/2406.06675

Tremmel, M., Ricarte, A., Natarajan, P., et al. 2024, An Enhanced Massive Black Hole Occupation Fraction Predicted in Cluster Dwarf Galaxies, doi: https://doi.org/10.33232/001c.116617

Valentini, M., & Dolag, K. 2025, Hydrodynamic methods and sub-resolution models for cosmological simulations. https://arxiv.org/abs/2502.06954

Vernardos, G., & Fluke, C. J. 2013, A new parameter space study of cosmological microlensing, doi: https://doi.org/10.1093/mnras/stt1076

Vilella-Rojo, G., Logroño-García, R., López-Sanjuan, C., et al. 2021, J-PLUS: The star formation main sequence and rate density at d < 75 Mpc, doi: https://doi.org/10.1051/0004-6361/202039156

Weller, E. J., Pacucci, F., Natarajan, P., & Matteo, T. D. 2023, Over-massive Central Black Holes in the Cosmological Simulations ASTRID and Illustris TNG50, doi: https://doi.org/10.1093/mnras/stad1362

Winkel, N., Bennert, V. N., Remigio, R. P., et al. 2024, Combining Direct Black Hole Mass Measurements and Spatially Resolved Stellar Kinematics to Calibrate the $M_{\rm BH}$-$\sigma_\star$ Relation of Active Galaxies. https://arxiv.org/abs/2411.02488

Wright, R. J., Somerville, R. S., del P. Lagos, C., et al. 2024, The baryon cycle in modern cosmological hydrodynamical simulations. https://arxiv.org/abs/2402.08408

Zhang, Y., Ouchi, M., Gebhardt, K., et al. 2023, The Stellar Mass - Black Hole Mass Relation at $z \sim 2$ Down to $\mathcal{M}_{\rm BH} \sim 10^7 M_\odot$ Determined by HETDEX, doi: https://doi.org/10.3847/1538-4357/acc2c2

Zhu, P., Ho, L. C., & Gao, H. 2020, The Correlation between Black Hole Mass and Stellar Mass for Classical Bulges and the Cores of Ellipticals, doi: https://doi.org/10.3847/1538-4357/abcaa1