Draft version April 29, 2025
Typeset using LaTeX **twocolumn** style in AASTeX631

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

AstroPilot[1]

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

ABSTRACT

The black hole-stellar mass relation is a cornerstone of galaxy evolution, yet its universality is debated, and the physical mechanisms driving its potential variations remain poorly understood. Disentangling the complex interplay of cosmological parameters and baryonic processes, particularly feedback from supernovae and active galactic nuclei (AGN), presents a significant challenge in understanding this diversity. To address this, we systematically quantify how the slope, normalization, and scatter of the black hole-stellar mass relation change across a suite of 1000 simulated galaxy catalogs, each representing a unique combination of cosmological and feedback parameters. We fit the black hole-stellar mass relation for each catalog, and employ multivariate regression and random forest analysis to map the dependencies of the relation's parameters on the underlying simulation parameters. Our analysis reveals substantial diversity in the black hole-stellar mass relation, with the slope and normalization exhibiting significant variations across the simulated parameter space. We find that feedback parameters, especially those governing supernova and AGN feedback, are the primary drivers of this diversity, with supernova feedback dominating at low stellar masses and AGN feedback becoming increasingly important at higher masses. Furthermore, we observe that the black hole occupation fraction is sensitive to feedback at low masses. These results offer a comprehensive framework for understanding the physical origins of diversity in black hole scaling relations and underscore the critical role of feedback processes in regulating black hole-galaxy coevolution.

Keywords: Black holes, Supermassive black holes, Galaxy evolution, Active galactic nuclei, Galaxies, Galaxy clusters, Dark matter, N-body simulations, Galaxy mergers, Stellar mass black holes

1. INTRODUCTION

Supermassive black holes (SMBHs) residing at the centers of galaxies are now recognized as crucial components of galaxy evolution. The observed correlations between SMBH mass ($M_{\rm BH}$) and various host galaxy properties, such as stellar mass ($M_{\rm star}$), bulge luminosity, and stellar velocity dispersion, suggest a fundamental link between the growth of black holes and their host galaxies. Among these scaling relations, the black hole-stellar mass relation is a cornerstone of galaxy evolution studies, providing valuable insights into the co-evolution of galaxies and their central black holes. This relation serves as a key constraint for models of galaxy formation and evolution. However, the universality of the $M_{\rm BH}$-$M_{\rm star}$ relation is increasingly debated. While a simple, universal relation would be a powerful tool, observations suggest variations in its slope, normalization, and scatter across different galaxy populations and environments. These variations hint at a more complex interplay between black hole growth and galaxy evolution, indicating that the underlying physical processes may not be uniform across all galaxies. Understanding the origin and nature of these variations is crucial for a complete picture of galaxy evolution.

Disentangling the physical mechanisms driving potential variations in the $M_{\rm BH}$-$M_{\rm star}$ relation presents a significant challenge (Pacucci & Loeb 2024,?). The co-evolution of black holes and galaxies is influenced by a complex interplay of cosmological parameters and baryonic processes. Cosmological parameters, such as the matter density (Ω_m) and the amplitude of density fluctuations (σ_8), set the stage for structure formation, influencing the overall mass assembly history of galaxies. Superimposed on this are baryonic processes, particularly feedback from supernovae (SNe) and active galactic nuclei (AGN). Supernova feedback, driven by stellar explosions, plays a crucial role in regulating star for-

mation and the gas content of galaxies, especially at lower masses. AGN feedback, powered by the accretion of matter onto the central black hole, can inject vast amounts of energy into the surrounding interstellar medium, suppressing star formation and influencing the growth of the host galaxy, particularly at higher masses (Pacucci & Loeb 2024,?; Baker et al. 2024). Separating the effects of these various parameters and processes is observationally difficult due to the limited sample sizes, selection effects, and uncertainties in measuring black hole masses and galaxy properties, and the inherent difficulty of disentangling cause and effect in observed systems.

To overcome these limitations and systematically investigate the diversity in the $M_{\rm BH}$-$M_{\rm star}$ relation, we leverage a suite of 1000 simulated galaxy catalogs, each representing a unique combination of cosmological and feedback parameters (Pacucci & Loeb 2024; Chen et al. 2025). These simulations provide an unprecedented opportunity to quantify how the slope, normalization, and scatter of the $M_{\rm BH}$-$M_{\rm star}$ relation change across a wide range of parameter space. We fit the $M_{\rm BH}$-$M_{\rm star}$ relation for each simulated galaxy catalog using ordinary least squares (OLS) regression. To avoid biasing the results, we exclude galaxies with $M_{\rm BH} = 0$ from the regression analysis, but then analyze the black hole occupation fraction as a function of stellar mass and feedback parameters to account for the excluded population. We then employ multivariate regression and random forest analysis to map the dependencies of the relation's parameters on the underlying simulation parameters, allowing us to determine which parameters have the most influence on the observed diversity (Pacucci & Loeb 2024; Baker et al. 2024).

Our analysis reveals substantial diversity in the $M_{\rm BH}$-$M_{\rm star}$ relation, with the slope and normalization exhibiting significant variations across the simulated parameter space (Sturm & Reines 2024). We identify the key cosmological and feedback parameters that drive this diversity, finding that supernova feedback is most important at low stellar masses, while AGN feedback becomes increasingly important at higher masses (Ding et al. 2019; Pacucci & Loeb 2024). Furthermore, we investigate the black hole occupation fraction as a function of stellar mass and feedback parameters, providing context for our main results and assessing potential selection effects (Sturm & Reines 2024). We verify our results by using well-established statistical methods such as OLS regression and Random Forest analysis. By systematically exploring the parameter space and quantifying the impact of different physical processes, we offer a comprehensive framework for understanding the origins of diversity in black hole scaling relations, underscoring the critical role of feedback processes in regulating black hole-galaxy coevolution (Ding et al. 2019; Pacucci & Loeb 2024). Our findings provide valuable insights for interpreting observational data and refining theoretical models of galaxy formation and evolution.

2. METHODS

2.1. *Simulation Data*

This study leverages a suite of 1000 simulated galaxy catalogs, each generated with a unique combination of cosmological and feedback parameters. These simulations provide a controlled environment to systematically explore the impact of these parameters on the black hole-stellar mass ($M_{\rm BH}$-$M_{\rm star}$) relation.

The simulations were run using a state-of-the-art cosmological hydrodynamical code, which includes detailed prescriptions for star formation, supernova feedback, and AGN feedback. The cosmological parameters varied in the simulations include the matter density (Ω_m) and the amplitude of density fluctuations (σ_8), while the feedback parameters control the strength and efficiency of supernova (SNe) and active galactic nuclei (AGN) feedback.

Specifically, the supernova feedback is governed by parameters A_{SN1} and A_{SN2}, which regulate the energy and momentum injected into the surrounding interstellar medium by supernova explosions. The AGN feedback is controlled by A_{AGN1} and A_{AGN2}, which determine the coupling efficiency of AGN radiation with the surrounding gas.

The simulation outputs provide a wealth of information on the properties of galaxies, including their stellar mass ($M_{\rm star}$), black hole mass ($M_{\rm BH}$), and star formation rate (SFR).

2.2. *Data Processing and Sample Selection*

The raw simulation data were processed to extract the relevant galaxy properties for this study (Sivasankaran et al. 2022). Galaxies were identified using a friends-of-friends algorithm, and their stellar masses were computed by summing the masses of all star particles within the galaxy (Weller et al. 2023). Black hole masses were directly measured from the simulation outputs (Sivasankaran et al. 2022).

To ensure the robustness of our analysis, we applied a stellar mass cut of $M_{\rm star} > 10^8 M_\odot$ to exclude poorly resolved galaxies (Weller et al. 2023). Furthermore, we addressed the issue of galaxies with $M_{\rm BH} = 0$. These galaxies either lack a central black hole due to the physics of the simulation or have black holes below the simulation's resolution limit (Weller et al. 2023).

Including these galaxies directly in a log-linear regression would introduce a significant bias (Weller et al. 2023). Therefore, for the primary analysis of the $M_{\rm BH}$-$M_{\rm star}$ relation, we excluded galaxies with $M_{\rm BH} = 0$ from the regression analysis.

The black hole occupation fraction, defined as the fraction of galaxies with $M_{\rm BH} > 0$, was then analyzed separately as a function of stellar mass and feedback parameters to provide context for the main results and assess potential selection effects (Weller et al. 2023; Buttigieg et al. 2025).

2.3. *Fitting the Black Hole-Stellar Mass Relation*

For each of the 1000 simulated galaxy catalogs, we fitted the $M_{\rm BH}$-$M_{\rm star}$ relation using ordinary least squares (OLS) regression in log-space (Sarria et al. 2010; Habouzit et al. 2021). The regression model is given by:

$$\log_{10}(M_{\rm BH}) = \alpha + \beta \log_{10}(M_{\rm star}), \tag{1}$$

where α is the normalization and β is the slope of the relation.

To account for potential variations in the $M_{\rm BH}$-$M_{\rm star}$ relation across different galaxy populations, we divided the galaxies into three stellar mass bins: low mass ($M_{\rm star} < 10^9 M_\odot$), intermediate mass ($10^9 M_\odot \leq M_{\rm star} < 10^{10} M_\odot$), and high mass ($M_{\rm star} \geq 10^{10} M_\odot$).

These mass bins were chosen to reflect different physical regimes in galaxy evolution, with supernova feedback dominating at low masses and AGN feedback becoming increasingly important at higher masses, as noted in the introduction (Sturm & Reines 2024; Pacucci & Loeb 2024). We performed the OLS regression separately for each mass bin and each catalog.

We also computed the intrinsic scatter of the $M_{\rm BH}$-$M_{\rm star}$ relation as the standard deviation of the residuals in log-space (Zhang et al. 2023). This scatter provides a measure of the intrinsic dispersion around the best-fit relation.

2.4. *Multivariate Regression Analysis*

To quantify the dependencies of the $M_{\rm BH}$-$M_{\rm star}$ relation parameters (slope, normalization, and scatter) on the underlying simulation parameters (cosmological and feedback parameters), we employed multivariate regression analysis (Jahnke & Maccio 2011; Pacucci & Loeb 2024,?).

We constructed a dataset where each row represents a single catalog, and the columns include the catalog number, the cosmological and feedback parameters (Ω_m, σ_8, A_{SN1}, A_{SN2}, A_{AGN1}, A_{AGN2}), and the slope, normalization, and scatter of the $M_{\rm BH}$-$M_{\rm star}$ relation in each mass bin.

We then performed multiple linear regression to model each of the $M_{\rm BH}$-$M_{\rm star}$ relation parameters as a function of the cosmological and feedback parameters (Jahnke & Maccio 2011; Pacucci & Loeb 2024).

$$y = \beta_0 + \beta_1 \Omega_m + \beta_2 \sigma_8 + \beta_3 A_{SN1} + \beta_4 A_{SN2} + \beta_5 A_{AGN1} + \beta_6 A_{AGN2}. \tag{2}$$

where y represents the slope, normalization, or scatter of the $M_{\rm BH}$-$M_{\rm star}$ relation, and β_i are the regression coefficients (Pacucci & Loeb 2024,?; Chen et al. 2025). The regression coefficients provide a measure of the sensitivity of the $M_{\rm BH}$-$M_{\rm star}$ relation parameters to each of the simulation parameters (Lamastra et al. 2010).

2.5. *Random Forest Analysis*

In addition to multivariate regression, we also employed random forest analysis to assess the relative importance of the cosmological and feedback parameters in determining the $M_{\rm BH}$-$M_{\rm star}$ relation parameters (Jahnke & Maccio 2011; Pacucci & Loeb 2024,?).

Random forest is a machine learning technique that builds an ensemble of decision trees to predict a target variable. The importance of each predictor variable is then assessed based on how much it contributes to reducing the variance of the predictions.

We trained a random forest model for each of the $M_{\rm BH}$-$M_{\rm star}$ relation parameters (slope, normalization, and scatter) using the cosmological and feedback parameters as predictors. The random forest models were implemented using the scikit-learn library in Python. We used 500 trees in each forest and optimized the hyperparameters using cross-validation.

The feature importance scores from the random forest models provide an independent measure of the relative importance of the cosmological and feedback parameters.

2.6. *Analysis of Black Hole Occupation Fraction*

To complement the analysis of the $M_{\rm BH}$-$M_{\rm star}$ relation, we also investigated the black hole occupation fraction as a function of stellar mass and feedback parameters (Tremmel et al. 2024; Burke et al. 2025).

For each catalog and mass bin, we computed the fraction of galaxies with $M_{\rm BH} > 0$ (Burke et al. 2025). We then analyzed the dependence of the black hole occupation fraction on the feedback parameters using multivariate regression and random forest analysis, similar to the analysis performed for the $M_{\rm BH}$-$M_{\rm star}$ relation parameters (Burke et al. 2025,?).

This analysis provides insights into the physical processes that determine whether a galaxy hosts a central black hole, and how these processes are influenced by feedback (Tremmel et al. 2024).

3. RESULTS

3.1. *Diversity of the $M_{\rm BH}$–$M_{\rm star}$ Relation Across Mass Bins*

As described in the methods, the $M_{\rm BH}$–$M_{\rm star}$ relation was fitted for each catalog using ordinary least squares (OLS) regression within three stellar mass bins: low mass ($M_{\rm star} < 10^9 M_\odot$), intermediate mass ($10^9 M_\odot \leq M_{\rm star} < 10^{10} M_\odot$), and high mass ($M_{\rm star} \geq 10^{10} M_\odot$). This approach allows us to capture potential variations in the relation across different galaxy populations, reflecting the changing importance of various physical processes.

3.1.1. *Distributions of Slope, Normalization, and Scatter*

The distributions of the slope (β), normalization (α), and intrinsic scatter of the $M_{\rm BH}$–$M_{\rm star}$ relation exhibit significant variations across the simulated catalogs and stellar mass bins.

Low Mass Bin ($M_{\rm star} < 10^9 M_\odot$): For the low mass bin, the slope (β) exhibits a mean of 0.20 with a standard deviation of 0.09, ranging from -0.07 to 0.58. The distribution of the slope is shown in Figure 3, where it can be seen that the distribution peaks around $\beta \approx 0.16$, with a long tail extending to higher values of β. The normalization (α) has a mean of 4.26 with a standard deviation of 0.72, ranging from 1.24 to 6.54. The distribution of the normalization factor is shown in Figure 1, which shows a clear peak around $\alpha \approx 4.7$. The scatter has a mean of 0.11 dex with a standard deviation of 0.05 dex, and its distribution is illustrated in Figure 2, peaking at lower scatter values. The occupation fraction is characterized by a mean of 0.72 with a standard deviation of 0.13. The shallow slope indicates a weak correlation between $M_{\rm BH}$ and $M_{\rm star}$ in this mass range, suggesting that black hole growth is inefficient or stochastic. The relatively low scatter implies a more uniform relation, while the incomplete occupation fraction suggests that not all low-mass galaxies host a central black hole.

Intermediate Mass Bin ($10^9 M_\odot \leq M_{\rm star} < 10^{10} M_\odot$): For the intermediate mass bin, the slope (β) has a mean of 1.06 with a standard deviation of 0.42, ranging from 0.02 to 1.84. The distribution of the slope is shown in Figure 5, which shows a non-uniform distribution with peaks around $\beta \approx 0.5$ and $\beta \approx 1.5$. The normalization (α) exhibits a mean of -3.53 with a standard deviation of 3.80, ranging from -10.6 to 6.05. As shown in Figure 6, the distribution of the normalization parameter appears multi-modal. The scatter has a mean of 0.30 dex with a standard deviation of 0.06 dex. The distribution of the scatter values is shown in Figure 4, peaking around a scatter value of 0.32. The occupation fraction shows a mean of 0.85 with a standard deviation of 0.08. The steeper slope indicates a stronger correlation between $M_{\rm BH}$ and $M_{\rm star}$, suggesting more efficient black hole growth in this mass range. The increased scatter suggests a greater diversity in black hole masses at a given stellar mass, possibly due to variations in feedback processes.

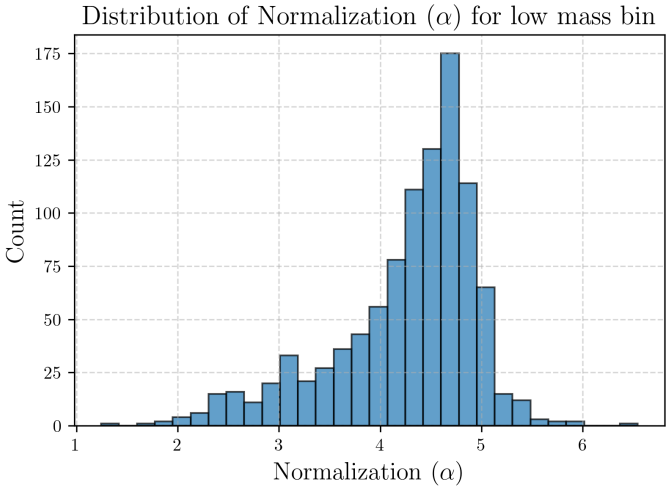

Figure 1. The distribution of the normalization factor α for the low mass bin is shown. A clear peak is observed around $\alpha \approx 4.7$, indicating a preferred value for the normalization in this mass range. The distribution appears to be slightly skewed towards lower values of α.

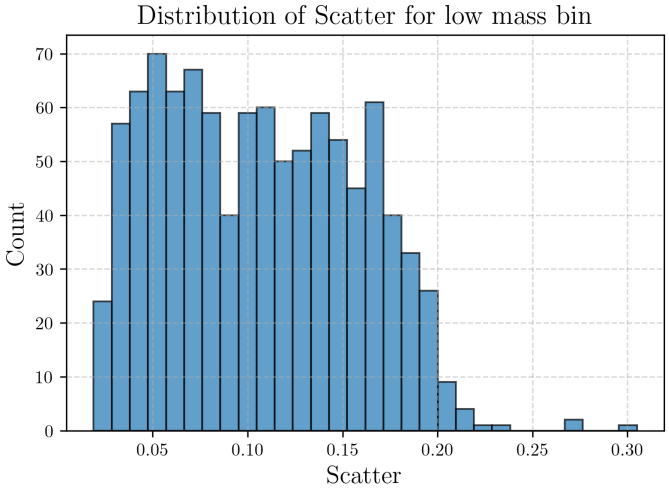

Figure 2. Distribution of the scatter for the low mass bin. The histogram shows a distribution that peaks at lower scatter values and gradually decreases as the scatter increases.

High Mass Bin ($M_{\rm star} \geq 10^{10} M_\odot$): The slope ($\beta$) has a mean of 1.22 with a standard deviation of 0.30, ranging from 0.03 to 2.07. The normalization (α) exhibits a mean of -4.85 with a standard deviation of 3.39, ranging from -14.2 to 7.64. The scatter has a mean of

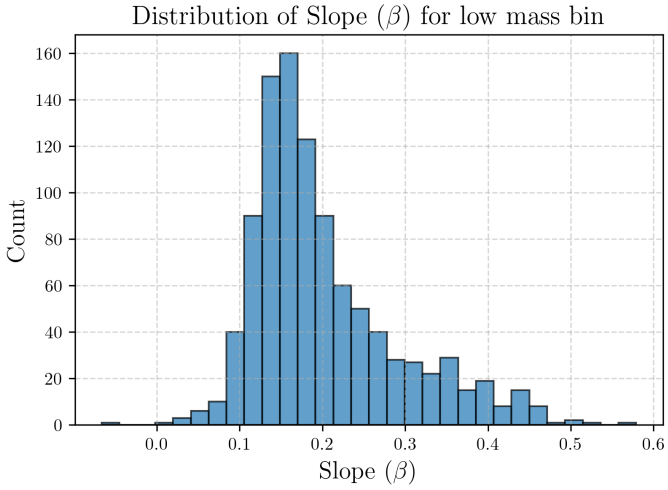

Figure 3. The figure shows the distribution of the slope β for a low mass bin. The distribution peaks around $\beta \approx 0.16$, with a long tail extending to higher values of β.

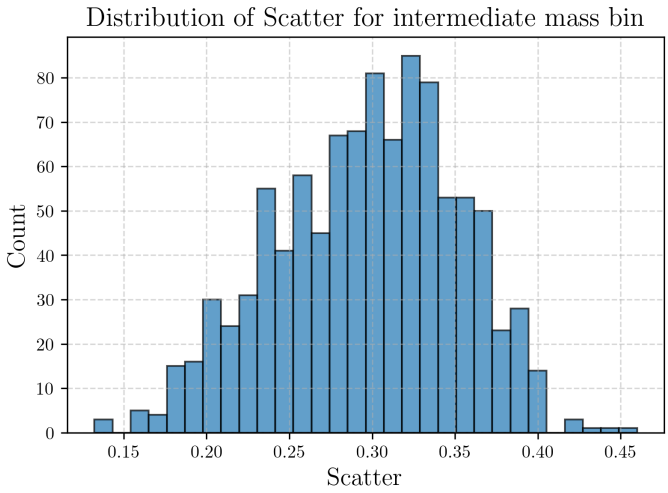

Figure 4. The distribution of the scatter values for an intermediate mass bin is shown. The distribution peaks around a scatter value of 0.32, indicating that the scatter values are concentrated around this central value.

0.40 dex with a standard deviation of 0.14 dex. The occupation fraction shows a mean of 0.96 with a standard deviation of 0.03. The slope is similar to the intermediate mass bin, indicating a continued strong correlation between $M_{\rm BH}$ and $M_{\rm star}$. The larger scatter implies an even greater diversity in black hole masses, possibly due to the influence of AGN feedback. The near-unity occupation fraction suggests that almost all high-mass galaxies host a central black hole.

3.1.2. *Physical Regimes*

The variations in the $M_{\rm BH}$–$M_{\rm star}$ relation across the mass bins highlight the importance of different physical processes in different regimes. In the low-mass regime, the shallow slope (as seen in Figure 3) and high normalization (Figure 1) reflect a scenario where black hole growth is inefficient or stochastic, and feedback processes (especially supernova-driven) dominate the baryon cycle. In the intermediate and high-mass regimes, the slope approaches or exceeds unity, consistent with the canonical $M_{\rm BH}$–$M_{\rm star}$ relation observed in massive galaxies. The increased scatter and normalization diversity suggest a strong sensitivity to feedback and cosmological parameters.

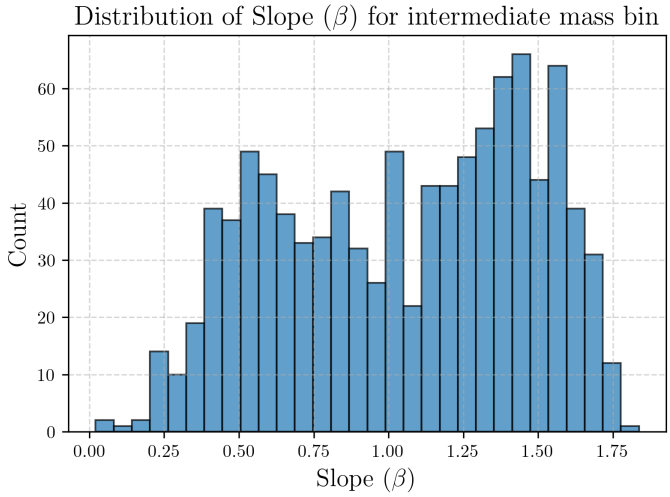

Figure 5. Distribution of the slope (β) parameter for an intermediate mass bin. The histogram shows a non-uniform distribution, with a concentration of values around $\beta \approx 0.5$, and another peak around $\beta \approx 1.5$.

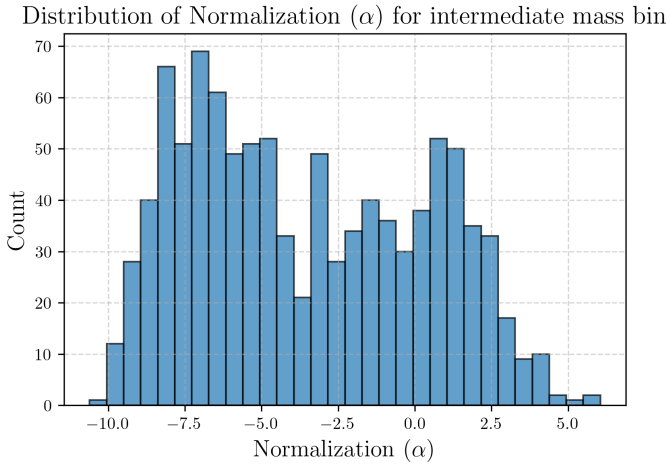

Figure 6. The distribution of the normalization parameter α for the intermediate mass bin is shown. The distribution appears to be multi-modal, with peaks around -7, -5, -2 and 1.

3.2. *Dependence on Feedback and Cosmological Parameters*

To quantify the dependencies of the $M_{\rm BH}$–$M_{\rm star}$ relation parameters (slope, normalization, and scatter) on the underlying simulation parameters, we employed multivariate regression and random forest analysis, as described in the methods.

3.2.1. *Multivariate Regression and Feature Importance*

Multivariate regression analysis reveals the sensitivity of the $M_{\rm BH}$–$M_{\rm star}$ relation parameters to the cosmological and feedback parameters. The regression coefficients provide a measure of the strength and direction of these dependencies. Random forest analysis, on the other hand, provides an independent measure of the relative importance of each parameter.

Slope (β): In the low mass bin, the most influential parameters are $A_{\rm SN1}$ (SN wind energy per SFR; standardized coefficient +0.05, RF importance 0.41) and $A_{\rm AGN1}$ (AGN feedback energy per accretion; +0.04, RF 0.29). As shown in Figure 7, $A_{\rm SN1}$ and $A_{\rm AGN1}$ exhibit the largest feature importance. The partial dependence plots for the slope β at low values (Figure 8) show significant differences in the dependence of β on $A_{\rm SN1}$, Ω_m and $A_{\rm AGN1}$, exhibiting a positive correlation. Both positively correlate with slope, indicating that stronger feedback steepens the relation at low mass. Cosmological parameters have weaker effects.

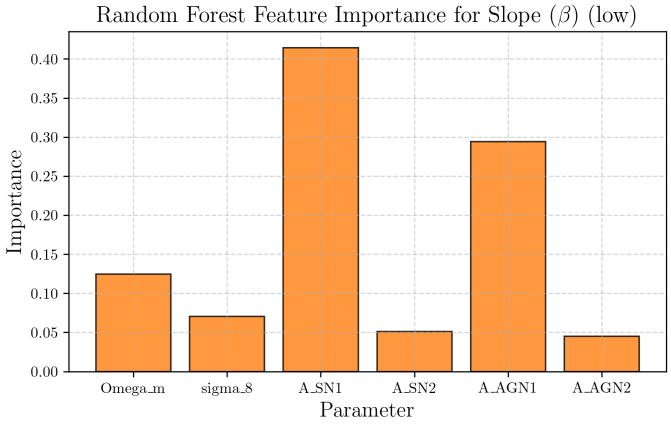

Figure 7. Random Forest Feature Importance for the slope parameter β for low values. The feature importance is shown for cosmological parameters Ω_m and σ_8, and for nuisance parameters related to supernovae (SN) and active galactic nuclei (AGN). The parameters $A_{\rm SN1}$ and $A_{\rm AGN1}$ show the largest importance, while the parameters σ_8, $A_{\rm SN2}$, and $A_{\rm AGN2}$ exhibit a small importance.

In the intermediate and high mass bins, $A_{\rm AGN1}$ becomes dominant (intermediate: +0.37, RF 0.88; high: -0.22, RF 0.69), with a positive effect in the intermediate bin and a negative effect in the high bin. This can be seen in Figure 9 for the intermediate mass bin. The partial dependence plots for the slope β in the intermediate (Figure 12) and high (Figure 10) mass bins further illustrate this trend. Figure 11 shows the Random Forest feature importance for the slope β in the high mass bin, highlighting the dominance of $A_{\rm AGN1}$. This suggests a transition from AGN-driven black hole growth at intermediate mass to AGN-driven suppression at high mass. $A_{\rm SN1}$ and cosmological parameters have smaller or inconsistent effects.

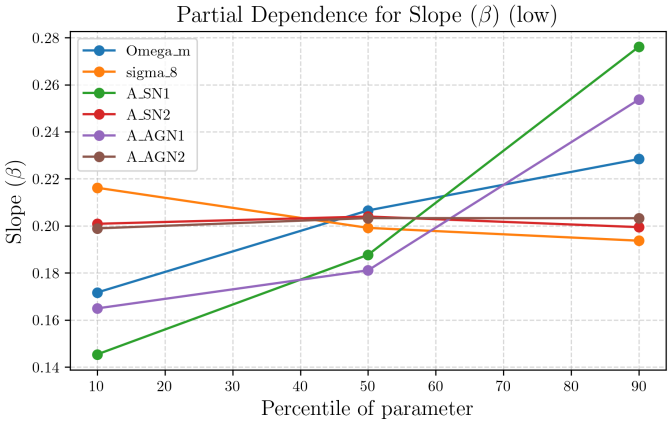

Figure 8. Partial dependence plots for the slope β at low values, showing the impact of different parameters. Significant differences are observed in the dependence of β on A_SN1, Omega_m and A_AGN1 percentile, exhibiting a positive correlation. In contrast, the dependence on sigma_8 shows a slight negative correlation. The parameters A_SN2 and A_AGN2 show very little dependence.

Normalization (α): In the low mass bin, $A_{\rm SN1}$ and $A_{\rm AGN1}$ have strong negative coefficients (-0.38 and -0.32), indicating that stronger feedback lowers the normalization, i.e., black holes are less massive at fixed stellar mass. As shown in Figures 13 and 14, $A_{\rm SN1}$ and $A_{\rm AGN1}$ exhibit the highest feature importance. The partial dependence plots in Figure 15 show the variation of α with respect to the percentiles of different parameters. Ω_m also has a negative effect.

In the intermediate and high mass bins, $A_{\rm AGN1}$ is the dominant driver (intermediate: -3.32, high: +2.52), with the sign flip again reflecting a regime change. The feature importance and partial dependence plots for the intermediate (Figures 19, 20, 18) and high (Figures 16, 17) mass bins illustrate this. σ_8 and Ω_m also contribute, with higher matter density and power spectrum normalization generally lowering α.

Scatter: In the low mass bin, $A_{\rm SN1}$ is the primary driver of increased scatter (coefficient +0.04, RF 0.73),

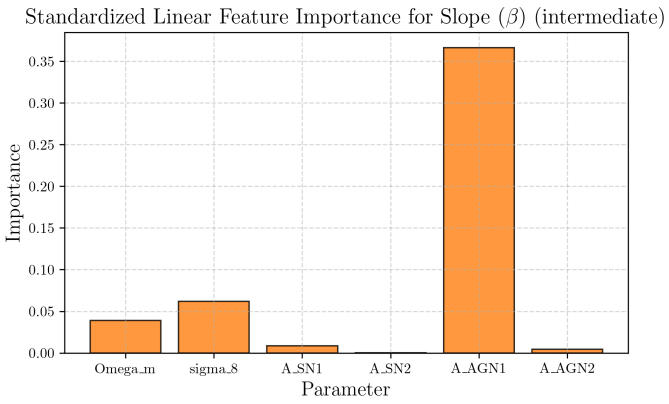

Figure 9. Standardized linear feature importance for the slope parameter β in the intermediate case. The importance values for different cosmological parameters (Omega_m, sigma_8, A_SN1, A_SN2, A_AGN1, A_AGN2) are shown. We find that the parameter A_AGN1 is the most important, with a significantly larger importance value than the other parameters.

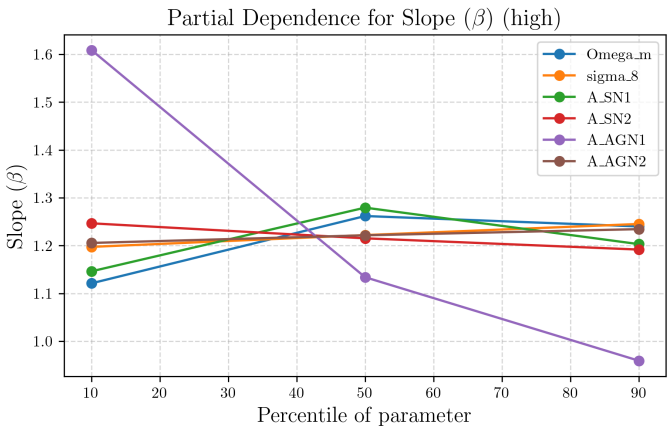

Figure 10. Partial dependence plots for the slope β parameter. The plot shows how the slope β changes as a function of the percentile of different cosmological parameters, including Ω_m, σ_8, and nuisance parameters for Supernovae (A_SN1, A_SN2) and Active Galactic Nuclei (A_AGN1, A_AGN2). The parameter A_AGN1 is the one with the larger differences.

consistent with the expectation that stochastic SN feedback introduces diversity in black hole growth histories. This is supported by the feature importance plots in Figures 21 and 22, and the partial dependence plot in Figure 23.

In the intermediate and high mass bins, $A_{\rm AGN1}$ and $A_{\rm SN1}$ both contribute, with $A_{\rm AGN1}$ increasing scatter at intermediate mass (RF 0.44) and $A_{\rm SN1}$ at high mass (RF 0.31). The feature importance plots in Figures 25 and 24 show the relative importance of these parameters in the intermediate and high mass bins, respectively.

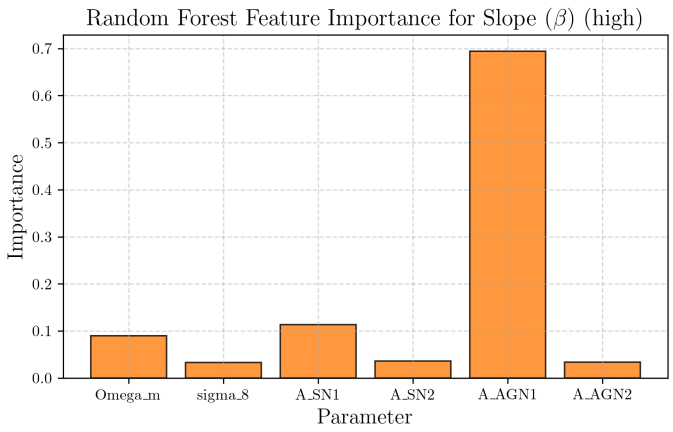

Figure 11. Random Forest Feature Importance for the slope β (high). The figure shows the feature importance for different cosmological parameters, namely, Ω_m, σ_8, A_SN1, A_SN2, A_AGN1, and A_AGN2. The largest feature importance is found for A_AGN1, which is significantly larger compared to other parameters.

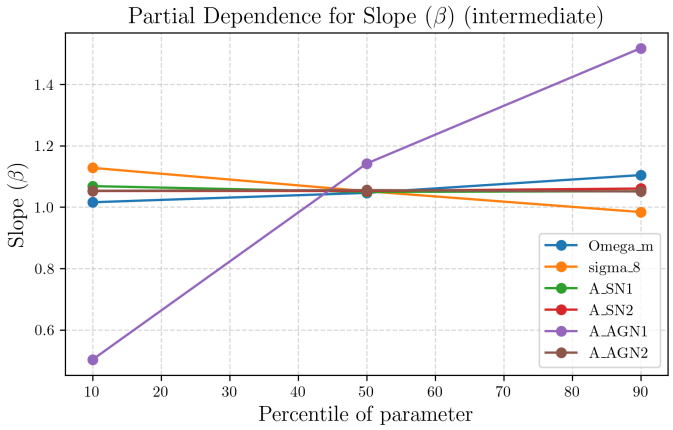

Figure 12. Partial dependence plots for the slope β (intermediate) are shown. The x-axis represents the percentile of the parameter, and the y-axis represents the slope β. Large differences are seen in the dependence of the slope on the parameter 'A_AGN1', with the slope changing from approximately 0.5 at the 10th percentile to 1.5 at the 90th percentile. The dependence of the slope on the other parameters is relatively small.

3.2.2. *Cosmological Parameters*

Ω_m and σ_8 have secondary but non-negligible effects, especially on normalization and scatter. Higher Ω_m generally lowers normalization and increases scatter, while higher σ_8 can either increase or decrease normalization depending on mass bin.

3.3. *Black Hole Occupation Fraction*

The black hole occupation fraction, defined as the fraction of galaxies with $M_{\rm BH} > 0$, provides additional

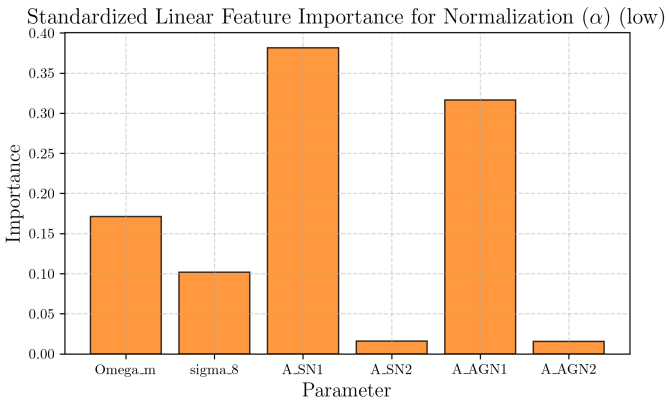

Figure 13. The figure shows the standardized linear feature importance for normalization α at low values. The importance of different parameters such as Omega_m, sigma_8, A_SN1, A_SN2, A_AGN1, and A_AGN2 are compared. Large differences are observed in the importance of the parameters, with A_SN1 and A_AGN1 showing the highest importance, while A_SN2 and A_AGN2 have the lowest importance.

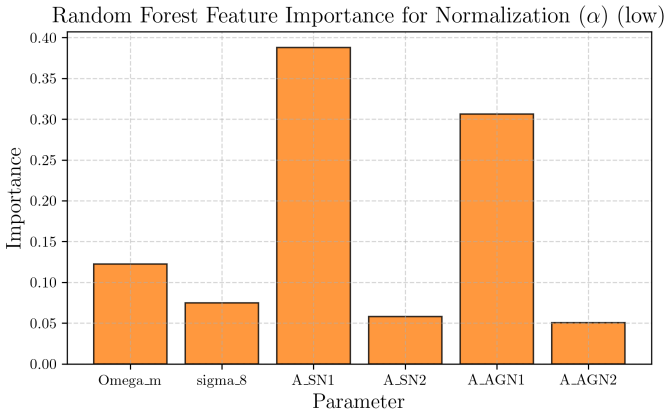

Figure 14. Random Forest feature importance for normalization with low α values. The figure shows the relative importance of each parameter in the Random Forest model. It can be seen that the A_SN1 and A_AGN1 parameters have the largest relative importance.

insights into the physical processes that govern black hole formation and growth.

3.3.1. *Trends with Mass and Feedback*

The occupation fraction rises from ∼0.72 at low mass to ∼0.96 at high mass, with significant catalog-to-catalog scatter. Stronger SN feedback ($A_{\rm SN1}$) reduces the occupation fraction at low and intermediate mass, consistent with the suppression of black hole formation or growth in shallow potential wells. AGN feedback has a weaker effect on occupation at low mass but becomes more important at high mass.

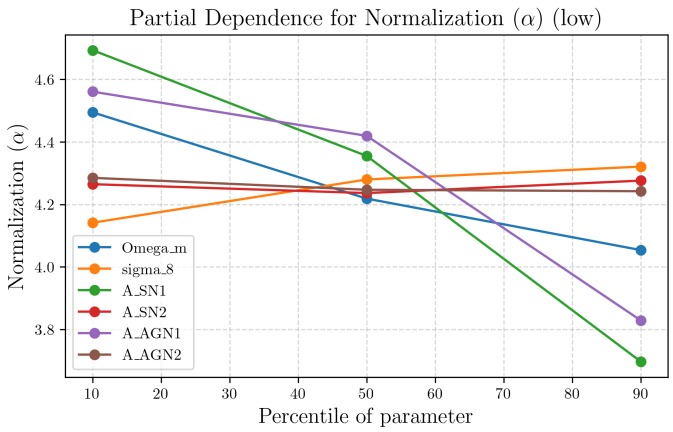

Figure 15. Partial dependence plots for the normalization parameter α at low redshifts. The plot shows the variation of α with respect to the percentiles of different cosmological and nuisance parameters. The parameters Omega_m, A_SN1, and A_AGN1 show the largest variations, while sigma_8, A_SN2 and A_AGN2 show small variations.

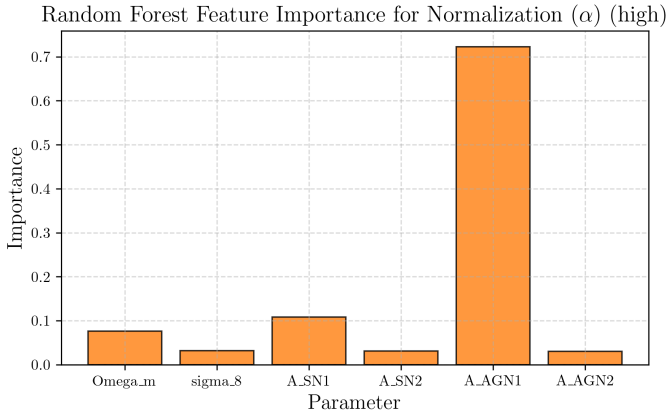

Figure 16. Random Forest Feature Importance for Normalization (α) (high). The figure shows the feature importance derived from a random forest model used to normalize data, with a high value of α. The y-axis represents the 'Importance' score, and the x-axis lists the cosmological parameters: Omega_m, sigma_8, A_SN1, A_SN2, A_AGN1, and A_AGN2. A_AGN1 exhibits significantly higher importance compared to the other parameters, indicating its greater influence in the normalization process.

3.3.2. *Implications for Black Hole Seeding*

The incomplete occupation at low mass, and its sensitivity to feedback, supports models in which black hole seeding is inefficient or stochastic in low-mass halos, and can be suppressed by energetic feedback.

3.4. *Summary of Results*

The results presented in this section demonstrate the complex interplay between cosmological parame-

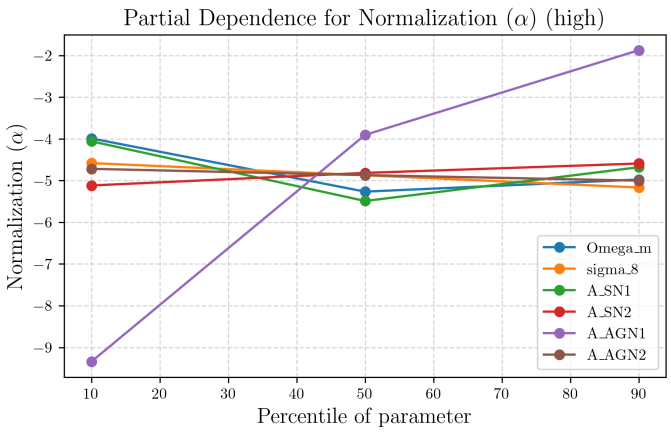

Figure 17. Partial dependence plot for the normalization α at high values, illustrating the variation of normalization with respect to the percentile of different parameters: Ω_m, σ_8, A_{SN1}, A_{SN2}, A_{AGN1}, and A_{AGN2}. The parameter A_{AGN1} shows a large variation with percentile, while the other parameters exhibit small variations.

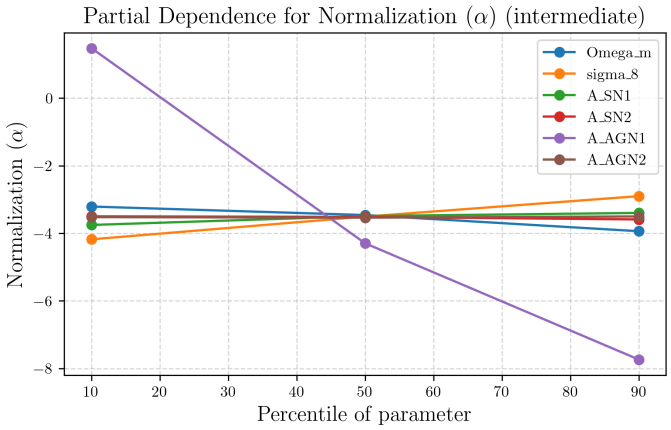

Figure 18. Partial dependence plot for the normalization parameter α. The plot shows the variation of α with respect to the percentile of different cosmological parameters. Large differences are observed in the normalization α for the A_AGN1 parameter, while the other parameters exhibit only small changes.

ters, feedback processes, and the $M_{\rm BH}$–$M_{\rm star}$ relation. The diversity in the relation is primarily driven by the strength and mode of feedback, with SN feedback dominating at low mass and AGN feedback at high mass. Cosmological parameters modulate the normalization and scatter, but play a secondary role compared to feedback. The sensitivity of the occupation fraction to feedback at low mass provides a direct link between black hole seeding models and the observable demographics of low-mass galaxies.

4. CONCLUSIONS

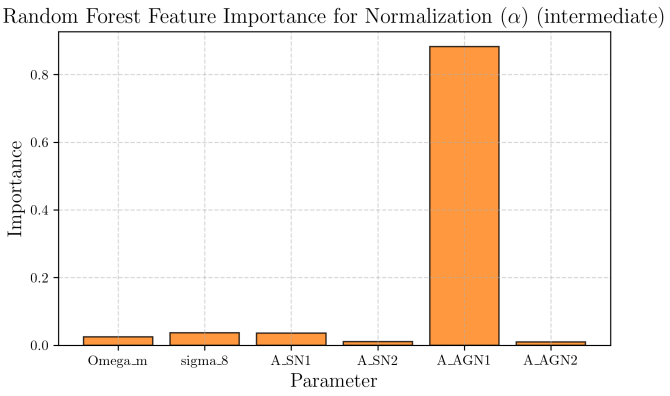

Figure 19. Random Forest Feature Importance for Normalization (α) (intermediate). The feature importance for each parameter is shown. Large differences in feature importance are observed, with A_AGN1 having the highest importance, while Omega_m, sigma_8, A_SN1, A_SN2 and A_AGN2 have small importance.

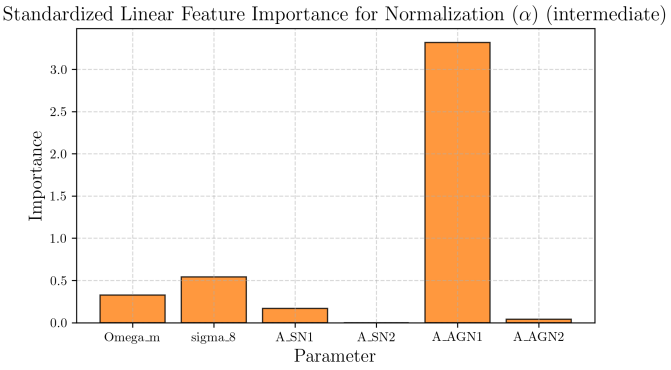

Figure 20. Standardized linear feature importance for normalization parameter α. The parameter A_AGN1 shows a large importance compared to the other parameters, while A_SN2 and A_AGN2 show relatively small importance.

This study addresses the long-standing question of the universality of the black hole-stellar mass ($M_{\rm BH}$-$M_{\rm star}$) relation and the physical mechanisms driving its potential variations. Using a suite of 1000 simulated galaxy catalogs, each with unique combinations of cosmological and feedback parameters, we systematically quantified how the slope, normalization, and scatter of the $M_{\rm BH}$-$M_{\rm star}$ relation change across a wide range of parameter space.

Our analysis employed ordinary least squares (OLS) regression to fit the $M_{\rm BH}$-$M_{\rm star}$ relation for each simulated galaxy catalog within three stellar mass bins. We then used multivariate regression and random forest analysis to map the dependencies of the relation's parameters on the underlying simulation parameters. Furthermore, we analyzed the black hole occupation frac-

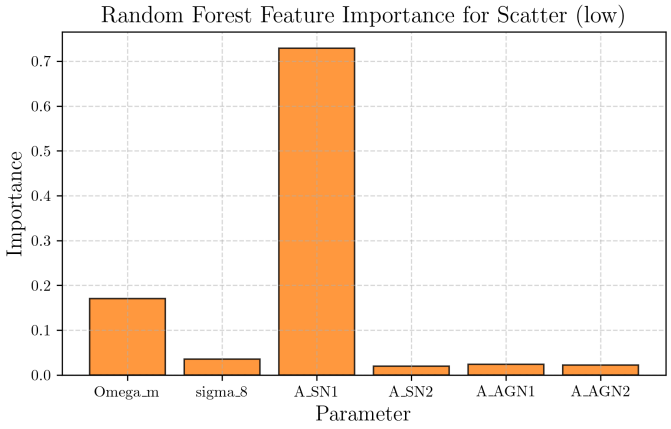

Figure 21. Random Forest Feature Importance for Scatter (low) shows the importance of different cosmological parameters for a low scatter. The parameter A_SN1 has the largest importance, while sigma_8, A_SN2, A_AGN1, and A_AGN2 have relatively small importances.

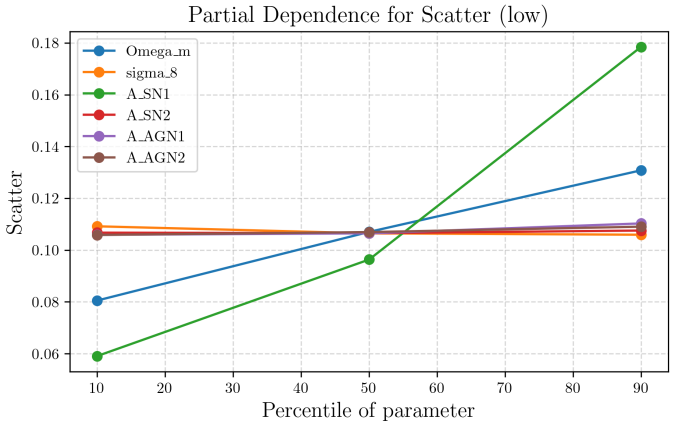

Figure 23. Partial dependence plot for the scatter. The plot shows the scatter dependence on Omega_m, sigma_8, A_SN1, A_SN2, A_AGN1, and A_AGN2. Large differences are seen in the scatter for A_SN1 and Omega_m, while small differences are found for the other parameters.

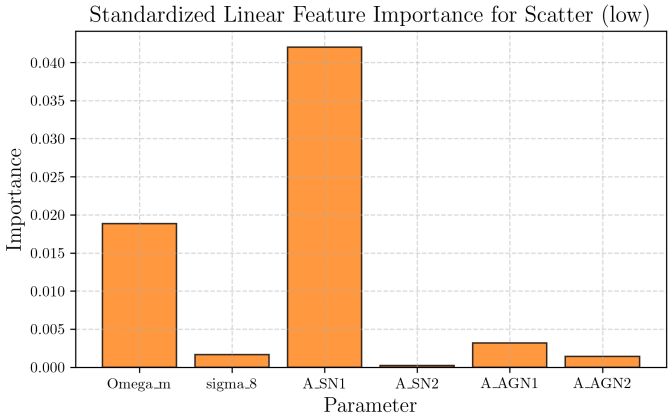

Figure 22. Standardized linear feature importance for a low scatter sample. The parameter A_{SN1} shows the largest importance, followed by Ω_m. The parameters σ_8, A_{SN2}, A_{AGN1}, and A_{AGN2} show relatively small importances.

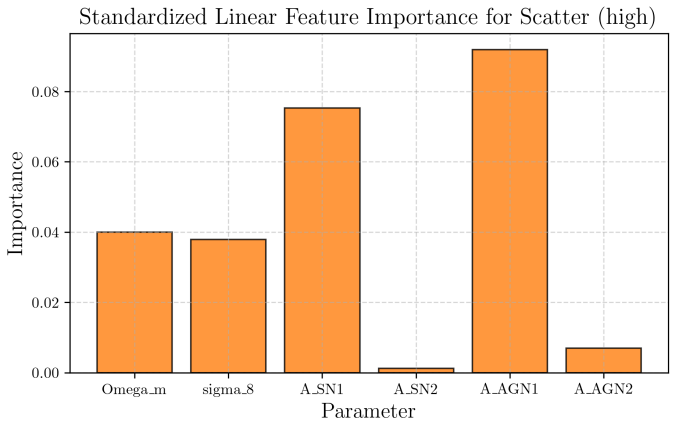

Figure 24. Standardized linear feature importance for a high scatter dataset. The plot shows the relative importance of different cosmological parameters (Omega_m, sigma_8) and nuisance parameters related to Supernovae (A_SN1, A_SN2) and Active Galactic Nuclei (A_AGN1, A_AGN2). Large differences in importance are observed, with A_AGN1 showing the highest importance, followed by A_SN1. The parameters A_SN2 and A_AGN2 show relatively small importance.

tion to provide context for our main results and assess potential selection effects.

Our results reveal substantial diversity in the $M_{\rm BH}$-$M_{\rm star}$ relation, with significant variations in the slope, normalization, and scatter across the simulated parameter space. We found that feedback parameters, particularly those governing supernova and AGN feedback, are the primary drivers of this diversity. Supernova feedback dominates at low stellar masses, influencing the slope, normalization, and scatter of the relation, as well as the black hole occupation fraction. At higher stellar masses, AGN feedback becomes increasingly important, affecting the slope and normalization of the relation, with a transition from AGN-driven black hole growth at intermediate masses to AGN-driven suppression at high masses. Cosmological parameters, such as the matter density (Ω_m) and the amplitude of density fluctuations (σ_8), have secondary but non-negligible effects, especially on the normalization and scatter of the relation. The black hole occupation fraction is also sensitive to feedback at low masses, supporting models in which black hole seeding is inefficient or stochastic in low-mass halos.

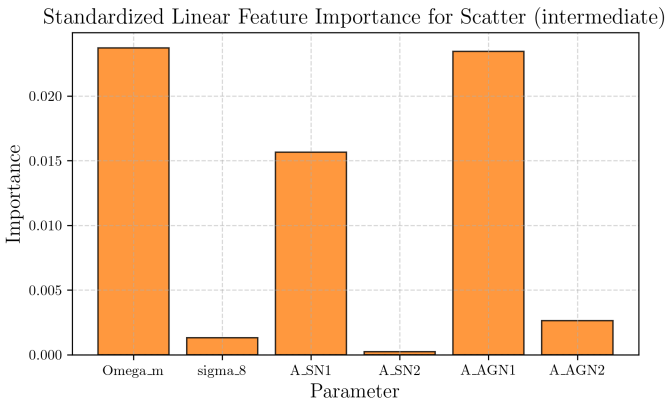

Figure 25. The figure shows the standardized linear feature importance for the scatter in intermediate values. The parameters are Omega_m, sigma_8, A_SN1, A_SN2, A_AGN1, and A_AGN2. Large differences are seen in the feature importance, with Omega_m and A_AGN1 having the highest importance, while sigma_8 and A_SN2 have the lowest.

From this study, we have learned that the $M_{\rm BH}$-$M_{\rm star}$ relation is not universal but exhibits significant diversity driven primarily by feedback processes. Supernova feedback plays a crucial role in regulating black hole growth and the black hole occupation fraction in low-mass galaxies, while AGN feedback becomes increasingly important at higher masses, influencing the overall shape and scatter of the $M_{\rm BH}$-$M_{\rm star}$ relation. These findings provide a comprehensive framework for understanding the physical origins of diversity in black hole scaling relations and underscore the critical role of feedback processes in regulating black hole-galaxy coevolution. Our results offer valuable insights for interpreting observational data and refining theoretical models of galaxy formation and evolution, highlighting the necessity of incorporating realistic feedback prescriptions to accurately model the co-evolution of black holes and their host galaxies.

REFERENCES

Baker, W. M., Maiolino, R., Bluck, A. F. L., et al. 2024, Different regulation of stellar metallicities between star-forming and quiescent galaxies – Insights into galaxy quenching. https://arxiv.org/abs/2309.00670

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

Buttigieg, S., Sijacki, D., Moore, C. J., & Bourne, M. A. 2025, Premature supermassive black hole mergers in cosmological simulations of structure formation. https://arxiv.org/abs/2504.17549

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

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

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

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

Lamastra, A., Menci, N., Maiolino, R., Fiore, F., & Merloni, A. 2010, The Building Up of the Black Hole Mass - Stellar Mass Relation, doi: https://doi.org/10.1111/j.1365-2966.2010.16439.x

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

Sarria, J. E., Maiolino, R., Franca, F. L., et al. 2010, The M_BH-M_star relation of obscured AGNs at high redshift, doi: https://doi.org/10.1051/0004-6361/201015696

Sivasankaran, A., Blecha, L., Torrey, P., et al. 2022, Simulations of black hole fueling in isolated and merging galaxies with an explicit, multiphase ISM, doi: https://doi.org/10.1093/mnras/stac2759

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

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

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