Draft version April 21, 2025
Typeset using LaTeX twocolumn style in AASTeX631

Evaluating the Stellar to Halo Mass Relation in Galaxies Generated by the SHARK Semi-Analytic Model

Jonah Walker[1]

[1] McMaster University
1280 Main St. W
Hamilton, ON L8S4L8, Canada

ABSTRACT

In the hierarchical model of structure formation, galaxies form and evolve within dark matter halos that grow through accretion and mergers. The stellar to halo mass relation (SHMR) characterizes the efficiency of galaxy formation by quantifying how baryonic structures (stellar mass) build up within these halos. This relation is a key constraint on galaxy formation models, reflecting the balance between star formation and feedback processes that quench star formation. This research evaluates how the SHMR varies across different galaxy environments within the SHARK semi-analytic model (SAM). SAMs work by applying physics-based models to precomputed dark matter halo merger trees. This approach is less computationally expensive when compared to hydrodynamic simulations, allowing us to test different physics models quickly and generate enormous samples of galaxies. SHARK is a novel SAM that is uniquely modular and open-source. We explore the SHMR as a function of galaxy properties and properties of their host environments, and from present day to when the Universe was half its current age (from redshift 0 to 1). Our results confirm that central galaxies, the most massive galaxy in a group residing in the centre of the halo, closely follow observed and simulated SHMR trends. In contrast, smaller galaxies orbiting the central (satellites) are quenched, exhibiting systematically lower SHMR, stellar mass, and cold gas mass—up to multiple orders of magnitude. Additionally, the SHMR shows remarkably little evolution over our redshift range, suggesting that the mechanisms that shape the SMHR are well in place when the Universe reached half its age.

Keywords: Galaxy Evolution — Dark Matter — Stellar-to-Halo Mass Relation — Semi-Analytic Modeling — SHARK

1. INTRODUCTION

Galaxies are enormous structures made up of gas, dust, and billions of stars, that exist inside an extended halo of dark matter (DM) particles. In the hierarchical model of structure formation, dark matter halos form in the early universe through gravitational collapse and merge with each other over time (White & Rees 1978). The gravity of the dense halos attracts the gas and dust which eventually forms a galaxy.

A galaxy can live in relative isolation in the field as the sole galaxy in an extended DM halo, or it can live in massive groups and clusters. The largest DM halos can contain thousands of small "satellite" galaxies in addition to the original and largest "central" galaxy. It is well understood the environment in which we find a galaxy has an influence on its evolution and observed properties. Early research into this includes Dressler (1980) for the effect of environment of morphology, as well as Kennicutt (1989) for the effect of environment of star formation rate.

This research evaluates this influence of galaxy environment on galaxy evolution using the stellar to halo mass relation (SHMR). Shown in Equation (1), the SHMR is a measure of the stellar mass of a galaxy relative to its host halo mass. It relates the amount of stars formed over a galaxy's lifetime to the mass of the host halo in which it lives.

$$SHMR = \frac{M_*}{M_{\mathrm{h,\,host}}} \tag{1}$$

Research measuring the SHMR has been done using both observational data and simulation data, with key breakthroughs occurring in the early 2000's. Marinoni & Hudson (2002) proposed the Halo Occupation Number, counting the number of galaxies in a virialized halo as a function of halo mass or total luminosity. Shortly after, the Conditional Luminosity Function, a function that yields the probability of finding a galaxy of a certain luminosity in a dark matter halo as a function of halo mass, was proposed (Yang et al. 2003; van den Bosch et al. 2003). Another key addition to this body of research is the mass luminosity relation derived from subhalo abundance matching in Vale & Ostriker (2004). By assuming there is a positive, monotonic relationship between halo mass and galaxy luminosity, they link an observed sample of galaxy luminosities to simulated dark-matter halo and subhalo masses. From this an empirical mass-luminosity relation is developed which is essentially the first measurement of the SHMR across a wide halo mass range.

Data from large-scale astronomical surveys broadened investigations into the SHMR. Many researchers, including Leauthaud et al. (2012) and Moster et al. (2012) use SDSS, Behroozi et al. (2013) combines SDSS and GALEX, Shuntov et al. (2022) uses COSMOS, and Behroozi et al. (2010) uses WMAP1 and 5. Using these sources of data, it was determined that the general structure of the SHMR is largely unchanging over time. Additionally, there is a characteristic halo mass in which a residing central can most efficiently convert baryons into stars. At $z = 0$, this mass is approximately $10^{12}\ M_\odot$, however this number decreases slightly with redshift. This peak halo mass effectively splits the SHMR into a low-mass and high-mass range. As such, stellar mass as a function of halo mass is often modelled as a broken power law, with a steep slope at the low-mass end and a shallower slope at the high-mass end. Below the efficiency peak, stellar mass increases with halo mass following $M_* \propto M_{\rm h}^{0.46}$ at $z \sim 0.2$–1 (Leauthaud et al. 2012) or alternatively $M_* \propto M_{\rm h}^{0.44}$ at low redshift, with little redshift evolution (Behroozi et al. 2013). The disjunction in the SHMR between high and low-mass halos may be the result of the high rate of in situ star formation in low mass galaxies, compared to high mass galaxies where ex situ (accreted) stars make up the majority of the stellar mass (Moster et al. 2012).

It is also known that the SHMR differs considerably between satellites and centrals (Zheng et al. 2005; Yang et al. 2009; Leauthaud et al. 2012; Shuntov et al. 2022). Most research defines the SHMR using centrals only, however when the populations are combined, central galaxies make up the bulk of the stellar mass in low mass halos, whereas satellites dominate the stellar mass in high mass halos (Shuntov et al. 2022).

The SHMR can be estimated observationally, but halo mass is challenging to measure directly. Recent efforts use simulations to measure this important relationship, such as Engler et al. (2020) using IllustrisTNG, Correa & Schaye (2020) using EAGLE, and Mitchell et al. (2015) using GALFORM. The research in this paper falls into this category, utilizing a novel semi-analytic model (SAM) of galaxy formation called SHARK.

SAMs are an established and efficient method of creating vast samples of galaxies and galaxy environments. They are used as an alternative to hydrodynamical simulations when researchers prioritize speed and a larger sample size. The two simulations differ in their handling of baryonic structures. In a hydrodynamic model, DM and baryonic matter directly interact to generate galaxies. As such, the model solves the gravity and fluid dynamics equations for a particle simultaneously. Conversely, a SAM simulates structure formation in a two-step process. Firstly, a DM-only N-body simulation generates a halo merger tree. Then, the SAM populates the DM universe with gas structures and galaxies. This method relieves considerable computational expense because DM-only N-body simulations are relatively easy for a computer, only needing to solve one equation (gravity) for a particle at a given snapshot.

SHARK (Lagos et al. 2018) is a uniquely open source[1] SAM that has been used to study a wide range of galaxy properties. Its code base is designed to be modular, written such that researchers can quickly and easily swap between different mathematical models for physics prescriptions governing galaxy evolution. SHARK has already proven to be an accurate SAM, as shown in papers such as Lagos et al. (2019) reproducing observed luminosity functions and Bravo et al. (2020) reproducing spectral energy distributions. An updated version, SHARK 2 (Lagos et al. 2024), incorporates revised prescriptions for key physical processes such as cooling, star formation, and feedback.

This work uses a version of SHARK built upon halo merger trees identified by HBT+HERONS (Ángel Chandro-Gómez et al. 2025), and is structured as follows. In Section 2 we introduce the SHARK dataset and describe the quantitative methods used to examine trends in SHMR scatter. Section 3 presents both recreated and novel visualizations generated from the dataset. Section 4 discusses our findings, and Section 5 summarizes the main conclusions of the paper.

2. GALAXY SAMPLE

There now exists a substantial body of research devoted to key metrics in SHARK galaxies. This study extends that effort by assessing whether trends in the SHMR are consistent with theoretical and observational

[1] Find the SHARK code base at: https://github.com/ICRAR/shark

expectations. Additionally, after verifying the accuracy of the SHMR in SHARK, we study the distribution of SHMR values across the dataset. To do this we employ a custom metric measuring residuals between galaxies and their expected SHMR value. We employ an exploratory data analysis framework to address three research objectives:

1. Assess the SHMR for central and satellite galaxies separately.
2. Examine SHMR trends through time.
3. Outline which galaxy properties influence the scatter in the SHMR.

The galaxies used in this analysis are drawn from three simulation snapshots, which correspond to redshift values of $z = 0$, 0.5, and 1, which correspond to lookback times of 0, ~5.5, and 8 Gyr. Sample sizes in these snapshots are summarized in Table 1. A lower stellar mass cut of 10^8 $M_\odot$ was applied to ensure all galaxies are well-resolved. This cut is consistent with Lagos et al. (2018).

Table 1. Number of Galaxies by Type and Redshift

	$z = 0$	$z = 0.5$	$z = 1$
Total	802,774	818,144	817,000
Centrals	587,721	608,764	626,660
Satellites	191,400	175,443	154,499

Visualization of these datasets was done using Python. To expedite the research process and minimize processing time, only certain fields were imported from the simulation output. These are outlined in Table 2. A full description of all available fields for a SHARK galaxy can be found at *https://shark-sam.readthedocs.io/*.

2.1. *Mass Distribution Overview*

Figures 1 and 2 depict histograms and kernel density estimates of the two most important galaxy properties in our sample: halo mass and stellar mass. Figure 1 shows the halo mass distribution peaking near $\log_{10}(M_\mathrm{h}/M_\odot) \sim 11.5$ across redshifts. It also shows hierarchical growth, as the number of high mass halos increases as redshift approaches zero. In Figure 2, the stellar mass distribution begins at $\log_{10}(M_\star/M_\odot) = 8$, followed by a gradual decline toward higher masses. This behaviour is consistent across all redshifts and reflects the dominance of low-mass galaxies in the simulated population. The immediate rise at the low-stellar mass end is a result of the imposed stellar mass cut below 10^8 $M_\odot$.

2.2. *SHARK Physics Prescriptions*

This dataset is a result of the prescriptions implemented in SHARK to simulate physical processes in galaxy formation. These physical processes, as described in Bravo et al. (2020), are as follows:

1. collapse and merging of DM haloes;
2. phase changes of gas between H II, H I, and H_2;
3. accretion of gas on to haloes, which is modulated by the DM accretion rate;
4. shock heating and radiative cooling of gas inside DM haloes, leading to the formation of galactic discs via conservation of specific angular momentum of the cooling gas;
5. star formation in galaxy discs;
6. stellar feedback from the evolving stellar populations;
7. chemical enrichment of stars and gas;
8. growth of black holes via gas accretion and merging with other supermassive black holes;
9. heating by AGN;
10. photoionization of the intergalactic medium and intrahalo medium in low-mass haloes;
11. galaxy mergers driven by dynamical friction within common DM haloes, which can trigger starbursts and the formation and/or growth of spheroids;
12. collapse of globally unstable discs that also lead to starbursts and the formation and/or growth of bulges;
13. environmental processes affecting the gas content of satellite galaxies.

3. MEASUREMENTS

Figure 3 shows the star formation rate (SFR) as a function of stellar mass for SHARK galaxies across redshifts $z = 0$, 0.5, and 1. A prominent star-forming main sequence is evident at all redshifts, with a tight correlation between stellar mass and SFR. The pink curve represents the running medians of SFR at fixed stellar

Table 2. SHARK Galaxy Dataset Fields and Descriptions

Field Name	Description
`ID_galaxy`	Unique identifier for the galaxy.
`ID_subhalo`	Identifier of the subhalo hosting the galaxy.
`ID_halo`	Identifier of the main halo.
`type`	Galaxy type: 0 (central), 1 (satellite), 2 (orphan).
`Mhalo_host`	Mass of the host halo (in $M_\odot$).
`Mhalo`	Mass of the subhalo the galaxy resides in (in $M_\odot$).
`Mstar_bulge`	Stellar mass in the bulge component (in $M_\odot$).
`Mstar_disc`	Stellar mass in the disc component (in $M_\odot$).
`Mstar_all`	Total stellar mass (bulge + disc) (in $M_\odot$).
`Mstar_stripped`	Stellar mass stripped from the galaxy (in $M_\odot$).
`Mhotgas`	Mass of hot gas in the halo (in $M_\odot$).
`Mcoldgas_bulge`	Cold gas mass in the bulge (in $M_\odot$).
`Mcoldgas_disc`	Cold gas mass in the disc (in $M_\odot$).
`Mcoldgas_all`	Total cold gas mass (bulge + disc) (in $M_\odot$).
`SFR_bulge`	Star formation rate in the bulge (in $M_\odot$/yr).
`SFR_disc`	Star formation rate in the disc (in $M_\odot$/yr).
`SFR_all`	Total star formation rate (bulge + disc) (in $M_\odot$/yr).
`MBH`	Mass of the central black hole (in $M_\odot$).
`MBHacc_cold`	Accretion rate onto the black hole from cold gas (in $M_\odot$/yr).
`MBHacc_hot`	Accretion rate onto the black hole from hot gas (in $M_\odot$/yr).
`MBHacc_all`	Total black hole accretion rate (in $M_\odot$/yr).
`Mgas_lost_SF`	Gas mass lost due to star formation-driven outflows (in $M_\odot$).
`Mgas_lost_QSO`	Gas mass lost due to quasar-mode AGN feedback (in $M_\odot$).

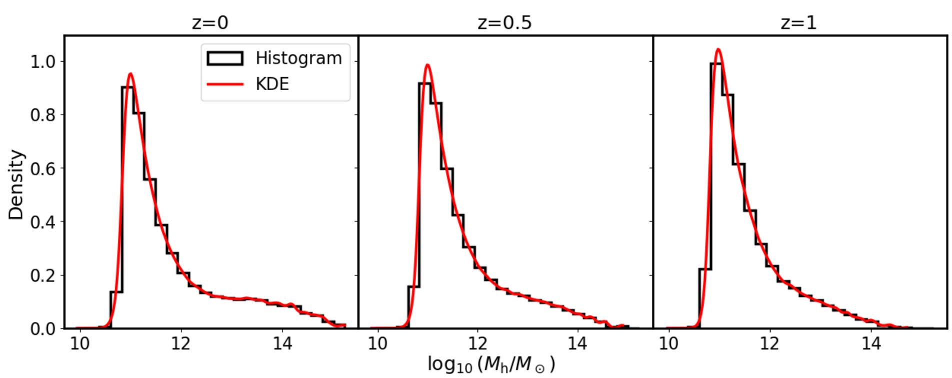

Figure 1. Halo mass distribution and KDE for three samples of SHARK galaxies across our redshift snapshots.

mass. The location and shape of this curve are consistent with observational results from Brinchmann et al. (2004) and simulated results from Lagos et al. (2024).

We define quenched galaxies as those with specific star formation rates (sSFRs) below 10^{-11} yr^{-1}, shown in each panel as a dashed orange line. As summarized in Table 3, a substantial fraction of SHARK satellite galaxies lie below this quenched threshold, particularly at lower redshifts. We find that centrals and satellites contribute differently to the quenched population, with satellites much more likely to be quenched across all redshifts.

Figure 4 is a common representation of the SHMR, depicting a galaxy's stellar mass as a function of halo mass. Overlaid are running median curves for this relation in previous studies, including Behroozi et al. (2013), Kravtsov et al. (2018), and Moster et al. (2012), all of which derive their relations using central galaxies exclusively. The SHMR computed in this work includes both central and satellite galaxies when evaluating the running median and interquartile range (IQR).

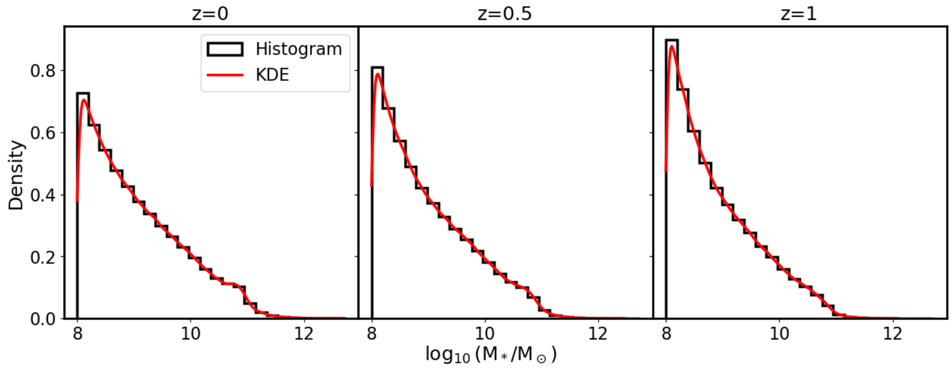

Figure 2. Stellar mass distribution and KDE for three samples of SHARK galaxies across our redshift snapshots. Note a stellar mass cut below $10^8\ M_\odot$ was implemented for this research.

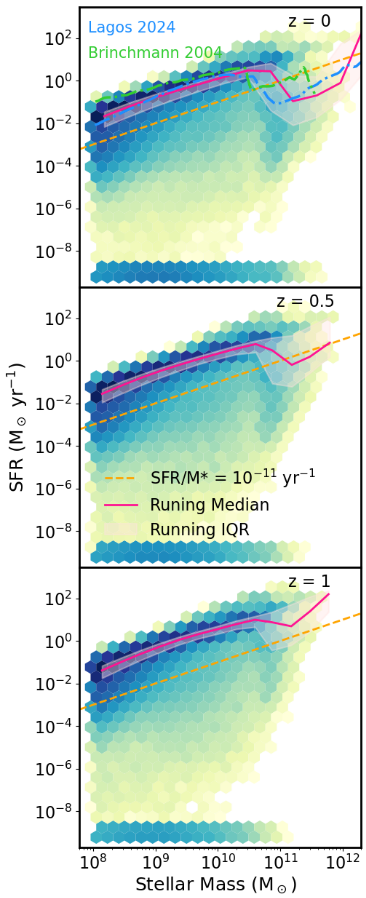

Figure 3. Star formation rate in units of solar mass per year as a function of stellar mass for a sample of SHARK galaxies at varying redshifts. The pink line forms the running median by connecting the median SFR values from equally spaced stellar mass bins. This line, along with the running medians from Brinchmann et al. (2004) (an observational study) and Lagos et al. (2024) (a SHARK study) trace the star forming main sequence.

Table 3. Proportion of galaxies below the sSFR quenched threshold (10^{-11} yr^{-1}).

Galaxy Category	z = 0	z = 0.5	z = 1
Centrals			
% Quenched	2.06	1.16	0.50
Satellites			
% Quenched	57.96	44.34	32.37

Despite this difference in sample, we find strong agreement with the central-galaxy-only trends at halo masses below $\sim 10^{12}\ M_\odot$. Above this value, our total-sample running median deviates, beginning with a sharp downward turn, eventually finding a plateau at constant stellar mass $M_* \approx 10^9\ M_\odot$.

This deviation reflects the dichotomy of properties between central and satellite galaxies. The environment of high-mass halos are dominated by small satellite galaxies compared to the relatively low amount of centrals with proportional stellar masses. The contribution of satellites depresses the total stellar mass relative to what would be expected from centrals alone.

Figure 5 depicts the considerable difference in SHMR running medians between central and satellite galaxies. For centrals, the stellar mass increases consistently with halo mass, mirroring trends reported among centrals in previous studies (i.e., Behroozi et al. (2013), Kravtsov et al. (2018), and Moster et al. (2012), plotted in Figure 4). In satellites, while our running median depicts a modest increasing trend, the wide IQR shows considerable scatter in the stellar mass distribution at fixed halo mass. We also observe a difference in density, where central galaxies are found in lower-mass halos, roughly in the range $10^{10}M_\odot \leq M_h/M_\odot \leq 10^{12}M_\odot$. Satellite galaxies are concentrated in high-mass halos of the approximate range $10^{12}M_\odot \leq M_h/M_\odot \leq 10^{15}M_\odot$. This separation in halo mass ranges further reinforces the

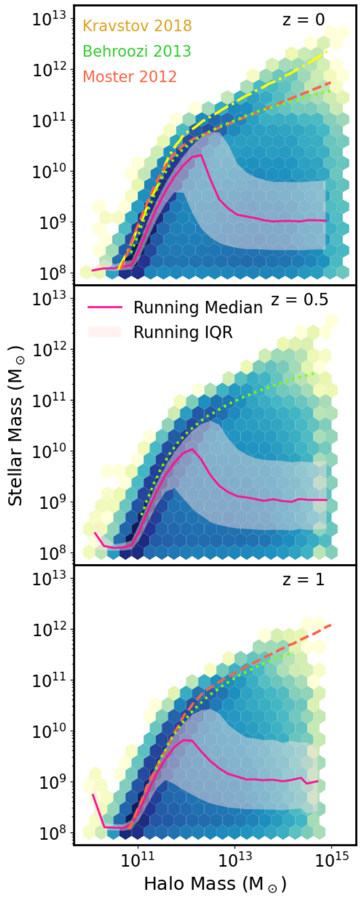

Figure 4. Stellar mass as a function of halo mass for a sample of SHARK galaxies across our redshift snapshots. The running median line for this dataset is shown in pink, and is calculated considering both central and satellite galaxies in our dataset. Running medians for this relation in Behroozi et al. (2013), Kravtsov et al. (2018), and Moster et al. (2012) consider central galaxies only.

distinction in SHMR behavior between the two galaxy types.

Figure 6 is another visualization of the SHMR, plotting the ratio of stellar and halo mass as a function of halo mass. Here, SHARK shows strong agreement with both observational data (Behroozi et al. 2013) and simulation results (Behroozi et al. 2019) for central galaxies at z=0. In central galaxies, the SHMR follows a parabolic curve with that peaks near $M_h = 10^{12} M_\odot$.

As discussed in our introduction, halo masses around $10^{12} M_\odot$ are widely considered optimal for efficient star formation in central galaxies, and this peak is expected to evolve with redshift (Kravtsov et al. 2018; Moster et al. 2012). Table 4 summarizes the halo mass at which the binned median SHMR is the greatest for each redshift snapshot in Figure 6. Our data indeed shows a tiny reduction in the halo mass peak.

Table 4. Halo masses corresponding to the peak SHMR values in Figure 6.

Redshift	Halo Mass at SHMR Peak ($M_\odot$)
$z = 0$	$10^{12.31}$
$z = 0.5$	$10^{12.30}$
$z = 1$	$10^{12.12}$

In central and satellite galaxies alike we observe the SHMR ratio is small, where $M_*/M_{\rm h,\,host} \leq 1$ for all M_h. It is understood that dark matter will always dominate the SHMR ratio, partially because star formation is not an efficient process; it is difficult to turn baryons into stars (Behroozi et al. 2013; Moster et al. 2012).

In the SHMR space, the stellar mass cutoff of $M_* < 10^8\, M_\odot$ translates into a diagonal boundary below which galaxies are omitted - most visible as a sharp lower edge in Figure 6. The boundary reflects the fact that low-mass galaxies below the cutoff are excluded from the analysis, and should not be interpreted as a physical lower bound of the SHMR distribution. However, the decline in SHMR with increasing halo mass among satellite galaxies is a genuine feature of the data. This is evidenced by the consistent downward slope observed in both the running upper quartile and median.

3.1. *Central SHMR Residual*

As can be seen in Figure 6, there is a lot of scatter in the SHMR for satellite galaxies at all redshifts. To explore this scatter, we calculate the *central SHMR residual* (Equation 2), which is the vertical distance between the SHMR of a galaxy to the SHMR/halo mass line for central galaxies (shown in the centrals column in Figure 6). This value is integral in determining how different galaxy properties correlate with scatter in the SHMR.

Figure 7 is a diagram depicting how the central SHMR residual is calculated. First, we isolate central galaxies and determine the full range of their host halo masses. Then, we split that range into equally spaced bins in log space, and compute the median SHMR in each bin. The central SHMR residual is then defined for each galaxy as the difference in log space between its SHMR and the median SHMR of its corresponding halo mass bin. We interpret this difference in units of dex, representing the offset in orders of magnitude.

$$\begin{aligned} &\textit{Central SHMR Residual} \\ &= \log(\text{SHMR of an individual galaxy}) \\ &- \log(\text{Median SHMR in that galaxy's bin}) \end{aligned} \tag{2}$$

Figure 8 is the central SHMR residual as a function of halo mass for both central and satellite galaxies across

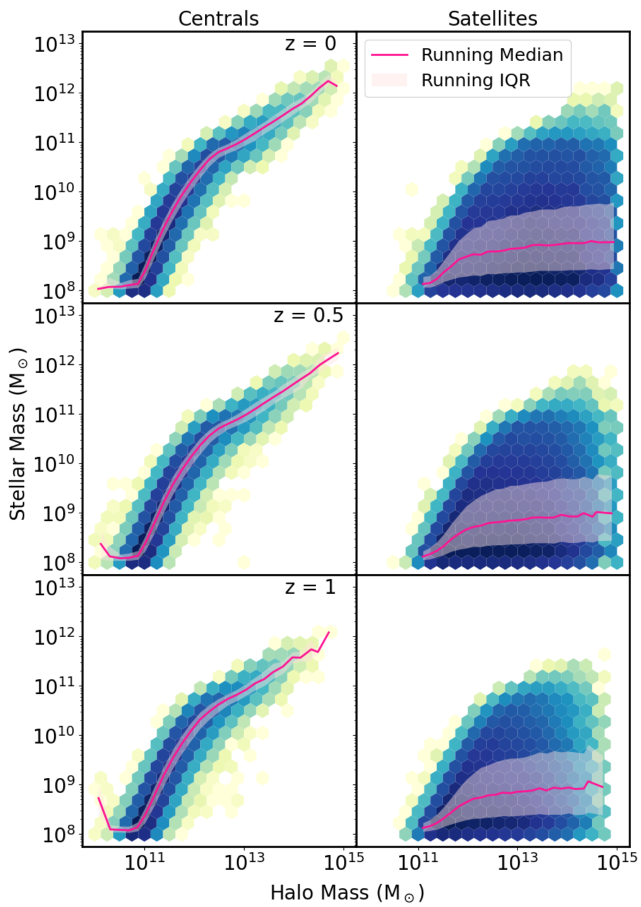

Figure 5. Stellar mass as a function of halo mass for a sample of SHARK galaxies across our redshift snapshots. The left column shows this relationship for central galaxies only. The right column shows this relationship for satellite galaxies only.

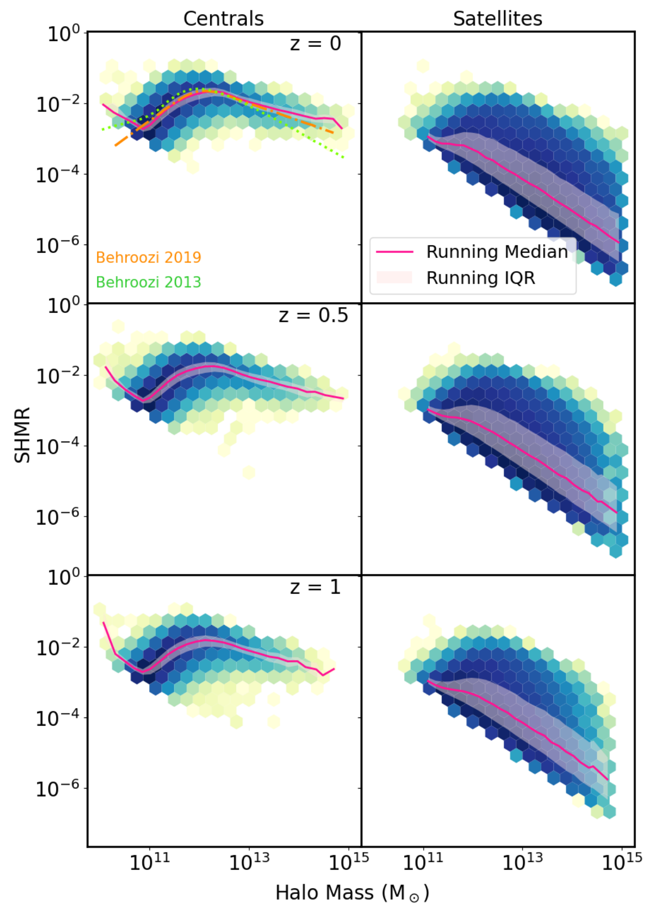

Figure 6. This figure depicts the SHMR as a ratio with respect to halo mass. The left column shows this relationship for central galaxies only. The right column shows this relationship for satellite galaxies only. The downward opening parabolic curve is a commonly observed pattern, shown in observational data (Behroozi et al. 2013) and simulation results (Behroozi et al. 2019). The peak of the parabola lies around $M_h = 10^{12} M_\odot$, decreasing to lower halo masses at higher redshifts in tiny steps, shown in Table 4.

three redshifts. To assess whether a statistically significant trend exists between halo mass and central SHMR residuals, we performed linear regressions on the full population of galaxies in each subsample. Due to the enormous sample size used in this analysis, even weak trends in SHMR residuals resulted in highly significant p-values. As a result, we focus our interpretation on slope and the coefficient of determination R^2, which more directly captures the explanatory power of each relationship.

For centrals, both the running median and the linear fit remain nearly flat across the full range of halo mass. A lack of correlation between central SHMR residuals and properties of central galaxies is an expected and reoccurring result. We specifically defined the central SHMR residual as the difference between a galaxy's SHMR and the median SHMR value for central galaxies in a similar halo mass bin. Therefore, our residuals for centrals should be symmetrically distributed around zero. The observed flatness of the running median and linear fit reinforces this expectation, suggesting that deviations from the SHMR for centrals are effectively random with respect to halo mass.

Residuals for satellite galaxies exhibit a clear decreasing trend with increasing halo mass, accompanied by moderate R^2 values across all redshifts. In higher mass halos, satellite galaxies are systematically beneath the central SHMR/halo mass line. Only satellites in low-mass halos can maintain high SHMR values comparable to those of centrals. Although this is partially a mathematical artifact of the halo mass and the SHMR being inversely proportional (see Equation 1), this relationship could indicate that processes that quench star formation in satellites are more prevalent or effective in

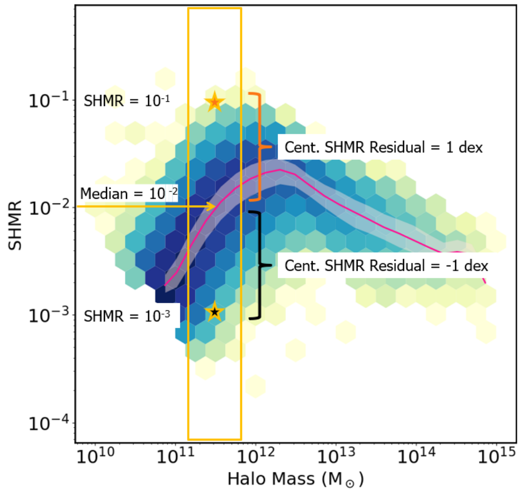

Figure 7. Diagram of how the Central SHMR Residual is calculated. This residual represents the difference in logspace between the SHMR of a galaxy (indicated by the black and yellow stars) and the median SHMR of its corresponding halo mass bin. In this example, a yellow bin spans $\sim 2 \times 10^{11} \leq M_h/M_{\odot} \leq\sim 8 \times 10^{11}$, and has a median SHMR of 10^{-2}. The residual is represented as an offset in log space, measured in units of dex (i.e., orders of magnitude) rather than as a raw difference.

higher mass halos. And because this trend is consistent across redshifts, it is likely that these processes remain uniformly active throughout cosmic time.

Figure 9 is the Central SHMR Residual as a function of stellar mass. Both centrals and satellites exhibit a positive trend as their residuals increase with stellar mass, although this relationship is much stronger in satellites.

Like in Figure 8, this is mainly the result of the SHMR being directly proportional to stellar mass. However, the stronger R^2 and the shape of the residual distribution suggests there could be a physical cause for the increasing trend. It is likely that these high mass satellites formed their stars before entering their host halo galaxy environment, using all their cold gas before it could be stripped by the environment. This is consistent with findings from hydrodynamical simulations showing that high-mass satellites typically form a substantial fraction of their stellar mass prior to infall and are more resilient to environmental disruption due to their deeper potential wells (Wang et al. 2017).

We also observe modest redshift evolution in the satellite column of Figure 9. Both the slope and R^2 values decline slightly from $z = 0$ to $z = 1$, indicating a weakening relationship between stellar mass and SHMR residuals over time. The decreasing slope suggests that environmental processes that quench star formation accumulate gradually, and satellites at higher redshift have

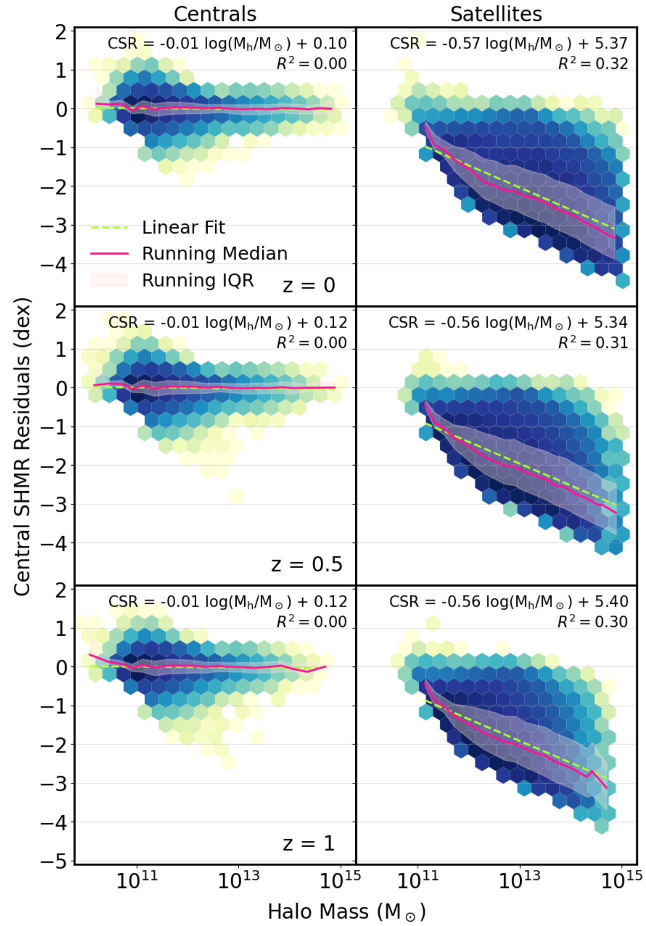

Figure 8. The relationship between the central SHMR residual and halo mass in central galaxies (left column) and satellite galaxies (right column) across our redshift snapshots. The pink line represents the running medians, and the green line is the output of a linear regression analysis on the galaxies of each subplot. The equation and coefficient of determination for each regression line is in the top right of each subplot. The associated p-values for every line was significant to a threshold of $< 10^{-10}$.

had less time to deviate from the central SHMR. Meanwhile, the mild weakening of the correlation implies that these environmental quenching effects do not impact all satellites equally, and could depend on factors like infall time, orbit, or host halo properties (Wang et al. 2017).

Figure 10 is the Central SHMR Residual as a function of the ratio of bulge mass to total mass (henceforth B/T) for the SHARK galaxies. We use B/T as a proxy for morphology, where a low B/T value is a disc dominated galaxy consistent with spirals, and a high value is bulge dominated and elliptical.

In satellites we observe a bimodal population of spirals and ellipticals, and it is clear from the linear fit that spiral galaxies deviate more from the SHMR than ellipti-

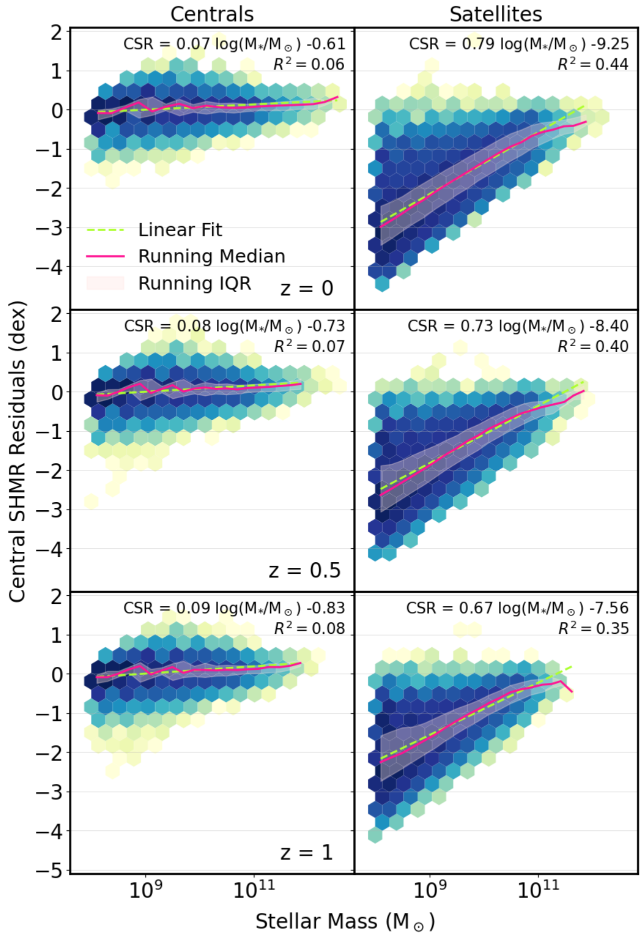

Figure 9. The relationship between the central SHMR residual and stellar mass in central galaxies (left column) and satellite galaxies (right column) across our redshift snapshots. The pink line represents the running medians, and the green line is the output of a linear regression analysis on the galaxies of each subplot. The equation and coefficient of determination for each regression line is in the top right of each subplot. The associated p-values for every line was significant to a threshold of $< 10^{-10}$.

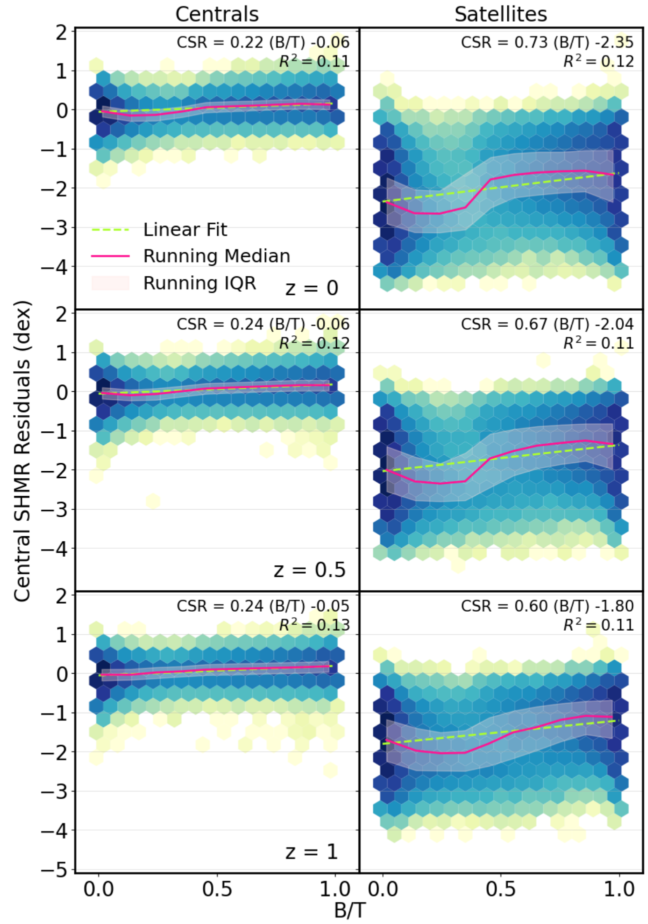

Figure 10. The relationship between the central SHMR residual and the bulge mass to total mass ratio in central galaxies (left column) and satellite galaxies (right column) across our redshift snapshots. The pink line represents the running medians, and the green line is the output of a linear regression analysis on the galaxies of each subplot. The equation and coefficient of determination for each regression line is in the top right of each subplot. The associated p-values for every line was significant to a threshold of $< 10^{-10}$.

cals. This is a slightly counterintuitive result, as spirals should be star forming thus increasing their SHMR, and ellipticals should have ceased star formation, theoretically falling behind in SHMR (Kennicutt 1998). What we are likely observing is the prevalence of gas stripping over the stripping of fully formed stars. Ellipticals have already formed all the stars they can, and in doing so reached a maximum SHMR, that can never be reduced because of the relative difficulty of stripping a fully formed star in SHARK. Conversely, upon infall, spirals have only converted a fraction of their cold gas to stars before being stripped. This leaves their stellar mass (thus SHMR) perpetually tiny.

Because both centrals and satellites exhibit a bimodal distribution of galaxies, a linear relationship between B/T and the central SHMR residual is not a good assumption. Although outside the scope of this research, I believe a piecewise linear function or sigmoid function would be a more accurate fit for this relationship.

Figure 11 is the central SHMR residual as a function of cold gas fraction, calculated in Equation (3).

$$f_{\text{gas, cold}} = \frac{M_{\text{cold}}}{M_{\text{cold}} + M_{\text{hot}} + M_{*}} \tag{3}$$

The distribution of galaxies in this plot supports the idea of widespread gas stripping in SHARK satellites. Here, the satellites with the smallest amounts of cold gas have the greatest scatter below the SHMR line. The

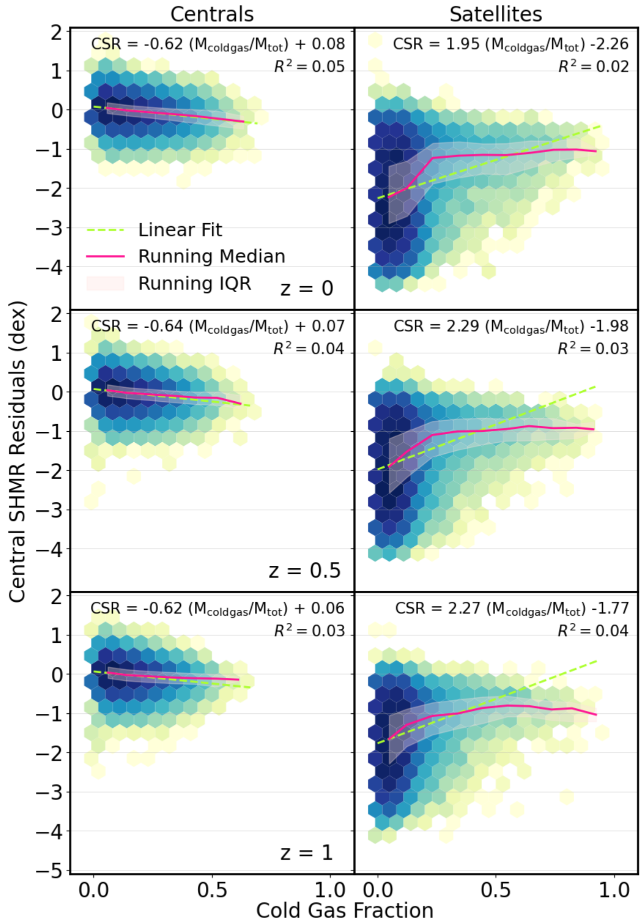

Figure 11. The relationship between the central SHMR residual and cold gas fraction in central galaxies (left column) and satellite galaxies (right column) across our redshift snapshots. The pink line represents the running medians, and the green line is the output of a linear regression analysis on the galaxies of each subplot. The equation and coefficient of determination for each regression line is in the top right of each subplot. The associated p-values for every line was significant to a threshold of $< 10^{-10}$.

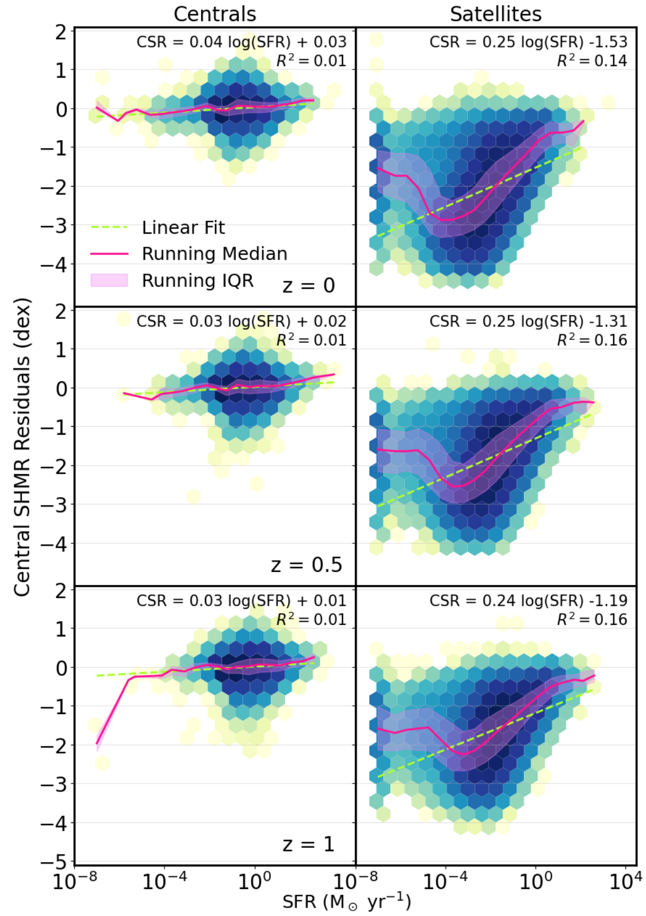

Figure 12. The relationship between the central SHMR residual and star formation rate in central galaxies (left column) and satellite galaxies (right column) across our redshift snapshots. The pink line represents the running medians, and the green line is the output of a linear regression analysis on the galaxies of each subplot. The equation and coefficient of determination for each regression line is in the top right of each subplot. The associated p-values for every line was significant to a threshold of $< 10^{-10}$.

steep relationship confirms that cold gas content has a tight, positive relationship to central SHMR residuals. This supports a scenario where satellites lose cold gas through stripping processes, preventing further star formation and leading to lower-than-expected stellar masses at fixed halo mass.

In Figure 12, we can see the effect of star formation rate on the residuals. While centrals cluster near $SFR \approx 1M_\odot/yr$, satellites occupy a broader distribution at lower star formation rate with a positive slope. The strong relation in the satellite population indicates that galaxies retaining higher SFRs remain closer to the central SHMR line. In contrast, those with suppressed star formation - likely due to cold gas stripping - show the most negative residuals.

Additional plots are included in the Appendix to further support the results presented. Appendix A contains extended central SHMR residual plots, while Appendix B provides a detailed breakdown of gas-related trends.

4. DISCUSSION

Our approach to study the galaxy properties influencing scatter in the SHMR in SHARK leads us to the conclusion that environmental quenching drives satellite galaxies towards systematically lower SHMR values compared to centrals.

Figures 1 to 6 show that SHARK generated a sample of ∼800,000 galaxies whose properties closely match previous observational and simulation studies. These galaxies display key astronomical properties as they clearly undergo hierarchical growth in Figure 1 and form a star forming main sequence in Figure 3. SHARK galaxies are extremely accurate when it comes to SHMR values, with our data in Figures 4, 5, 6 closely agreeing with previous studies on the statistic. Additionally, in all plots not using the central SHMR relation, the galaxies showed no redshift dependence, as predicted in the literature.

The central SHMR residual is calculated in such a way that it displays trends in the scatter of the SHMR when plotted. Figure 10 plots the relationship between the residuals and the bulge mass to total mass ratio as a proxy for morphology. The trendline in this figure indicates that spiral satellite galaxies are more scattered than their elliptical counterparts. The figure also indicates the SHMR for all satellites are globally suppressed, as their residuals are mostly negative. This is an interesting result as spiral galaxies are typically the more star forming morphology. The fact that these SHMR values are low means that these galaxies are not forming stars as they should be, which strongly indicates that spirals are quenched whenever they are in a host halo.

Figure 11 adds to this picture by indicating that satellites are being quenched through gas stripping. At higher redshifts the galaxies with the lowest cold gas fraction did not deviate from the SHMR as much as they do now. This is evidenced by the regression intercept becoming more negative with increasing redshift. An increased prevalence of galaxies with extremely low cold gas fraction would stack values in the bottom left corner of the plot and decrease the y intercept. Table 5 shows the prevalence of satellite galaxies whose cold gas fraction is depleted at our redshift snapshots. Because the number of cold gas depleted galaxies is increasing over time, we know there must be considerable gas stripping occurring.

Table 5. Count of satellite galaxies with a depleted cold gas fraction over redshift. We define depleted as cold gas fraction $< 10\%$.

Redshift	**Count of Depleted Satellites**
$z = 0$	$147,794$
$z = 0.5$	$139,027$
$z = 1$	$121,622$

Figure 12 shows the result of that gas stripping: a drop in star formation. Satellite galaxies with the lowest SFRs also show the most negative SHMR residuals, while those still forming stars remain closer to the central SHMR. The strong trend across redshifts reinforces the link between quenching and environmental suppression of gas in satellites.

Taken together, the results presented in Figures 10 through 12 provide compelling evidence that environmental regulation of gas and star formation is the dominant mechanism shaping the SHMR residuals in satellite galaxies. Future work could build on this analysis in several ways. First, it would be valuable to test whether the trends observed in SHARK also appear in other galaxy formation models, such as GALFORM or IllustrisTNG. This would clarify whether the strong link between SHMR residuals and gas/SFR is a robust prediction or a model-dependent feature. Second, the results could be compared directly to observational data - for example, by matching the SHMR residual trends seen here to those from SDSS or GAMA surveys, using proxy measures for cold gas and star formation. This would allow us to test whether the physical mechanisms identified in the simulation are also shaping real galaxies. Third, a future study could track individual satellite galaxies through time to understand the role of infall, orbital history, and stripping events in shaping SHMR residuals. This would connect the static trends found here to the dynamic processes that drive them.

5. CONCLUSION

In this work, we present a detailed analysis of the stellar-to-halo mass relation (SHMR) in galaxies produced by the SHARK semi-analytic model. We began by confirming that SHARK produces realistic galaxy populations in terms of stellar mass, halo mass, SFR, and overall SHMR shape. From there, we investigated which physical properties of galaxies contribute to scatter around the SHMR, with a particular focus on satellite galaxies.

Addressing our first research objective to evaluate central and satellite SHMR values separately, our analysis showed a clear dichotomy between the two. The SHMR for central galaxies is tight, with residuals exhibiting little to no dependence on galaxy morphology, gas content, or star formation activity. In contrast, satellite galaxies showed strong, consistent trends: residuals became increasingly negative in systems with low cold gas fractions, low star formation rates, and high bulge-to-total ratios.

Addressing our second research objective to evaluate SHMR trends over time, we observed little to no dependence of redshift trends on the SHMR. The plots of stellar mass versus halo mass and SHMR ratio versus halo mass were uniform across redshift except for a slight decline in peak SHMR consistent with observational results. Some time dependent trends were observed over time when plotting central, which we correlate to external factors quenching and stripping satellites.

Addressing our final research objective to outline which galaxy properties influence the scatter in the SHMR, we showed that cold gas fraction and star formation rate are the strongest predictors of SHMR residuals in satellites. Structural features like B/T, stellar mass, and halo mass play a secondary role. These results suggest that many satellites experience quenching as a result of gas stripping. This widespread quenching causes satellites to fall to low SHMR values. Overall, the satellite SHMR appears to be shaped by external processes, whereas central galaxies evolve in a more internally regulated manner.

These results highlighted the importance of environment in SHARK, playing an integral role shaping the stellar mass content of satellite galaxies. By visualizing SHMR trends this study provides support that the physics prescriptions and halo finder in SHARK produces an accurate and useful galaxy sample across redshifts. Additionally, by linking SHMR residuals to specific physical properties, this study provided a clearer picture of the processes that drive galaxy evolution in dense environments and established a framework for future work using both simulations and observations.

ACKNOWLEDGMENTS

I would like to thank Dr. Laura Parker and Dr. Matías Bravo for their guidance, feedback, and support throughout the project. I would also like to thank Megan Oxland for her help collecting the SHARK dataset used in this paper.

Software: numpy (Harris et al. 2020), scipy (Virtanen et al. 2020), matplotlib (Hunter 2007), pandas (McKinney 2010), h5py (Collette 2013), Jupyter (Kluyver et al. 2016)

REFERENCES

Behroozi, P., Wechsler, R. H., Hearin, A. P., & Conroy, C. 2019, MNRAS, 488, 3143, doi: 10.1093/mnras/stz1182

Behroozi, P. S., Conroy, C., & Wechsler, R. H. 2010, ApJ, 717, 379, doi: 10.1088/0004-637X/717/1/379

Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, ApJ, 770, 57, doi: 10.1088/0004-637X/770/1/57

Bravo, M., Lagos, C. d. P., Robotham, A. S. G., Bellstedt, S., & Obreschkow, D. 2020, Monthly Notices of the Royal Astronomical Society, 497, 3026–3046, doi: 10.1093/mnras/staa2027

Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, MNRAS, 351, 1151, doi: 10.1111/j.1365-2966.2004.07881.x

Collette, A. 2013, Python and HDF5 (O'Reilly Media, Inc.)

Correa, C. A., & Schaye, J. 2020, Monthly Notices of the Royal Astronomical Society, 499, 3578, doi: 10.1093/mnras/staa3053

Dressler, A. 1980, ApJ, 236, 351, doi: 10.1086/157753

Engler, C., Pillepich, A., Joshi, G. D., et al. 2020, Monthly Notices of the Royal Astronomical Society, 500, 3957, doi: 10.1093/mnras/staa3505

Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357

Hunter, J. D. 2007, Computing in Science & Engineering, 9, 90

Kennicutt, R. C. 1998, Annual Review of Astronomy and Astrophysics, 36, 189, doi: 10.1146/annurev.astro.36.1.189

Kennicutt, Jr., R. C. 1989, ApJ, 344, 685, doi: 10.1086/167834

Kluyver, T., et al. 2016, in Positioning and Power in Academic Publishing: Players, Agents and Agendas (IOS Press), 87–90

Kravtsov, A. V., Vikhlinin, A. A., & Meshcheryakov, A. V. 2018, Astronomy Letters, 44, 8, doi: 10.1134/S1063773717120015

Lagos, C. d. P., Tobar, R. J., Robotham, A. S. G., et al. 2018, MNRAS, 481, 3573, doi: 10.1093/mnras/sty2440

Lagos, C. d. P., Robotham, A. S. G., Trayford, J. W., et al. 2019, MNRAS, 489, 4196, doi: 10.1093/mnras/stz2427

Lagos, C. d. P., Bravo, M., Tobar, R., et al. 2024, Monthly Notices of the Royal Astronomical Society, 531, 3551, doi: 10.1093/mnras/stae1024

Leauthaud, A., Tinker, J., Bundy, K., et al. 2012, ApJ, 744, 159, doi: 10.1088/0004-637X/744/2/159

Marinoni, C., & Hudson, M. J. 2002, ApJ, 569, 101, doi: 10.1086/339319

McKinney, W. 2010, in Proceedings of the 9th Python in Science Conference, ed. S. van der Walt & J. Millman, 51–56

Mitchell, P. D., Lacey, C. G., Baugh, C. M., & Cole, S. 2015, Monthly Notices of the Royal Astronomical Society, 456, 1459, doi: 10.1093/mnras/stv2741

Moster, B. P., Naab, T., & White, S. D. M. 2012, Monthly Notices of the Royal Astronomical Society, 428, 3121, doi: 10.1093/mnras/sts261

Shuntov, M., McCracken, H. J., Gavazzi, R., et al. 2022, A&A, 664, A61, doi: 10.1051/0004-6361/202243136

Vale, A., & Ostriker, J. P. 2004, MNRAS, 353, 189, doi: 10.1111/j.1365-2966.2004.08059.x van den Bosch, F. C., Yang, X., & Mo, H. J. 2003, MNRAS, 340, 771, doi: 10.1046/j.1365-8711.2003.06335.x

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261, doi: 10.1038/s41592-019-0686-2

Wang, M. Y., Fattahi, A., Cooper, A. P., et al. 2017, MNRAS, 468, 4887, doi: 10.1093/mnras/stx742

White, S. D. M., & Rees, M. J. 1978, MNRAS, 183, 341, doi: 10.1093/mnras/183.3.341

Yang, X., Mo, H. J., & van den Bosch, F. C. 2003, MNRAS, 339, 1057, doi: 10.1046/j.1365-8711.2003.06254.x

—. 2009, ApJ, 695, 900, doi: 10.1088/0004-637X/695/2/900

Zheng, Z., Berlind, A. A., Weinberg, D. H., et al. 2005, ApJ, 633, 791, doi: 10.1086/466510

Ángel Chandro-Gómez, del P. Lagos, C., Power, C., et al. 2025, On the accuracy of dark matter halo merger trees and the consequences for semi-analytic models of galaxy formation. https://arxiv.org/abs/2501.07677

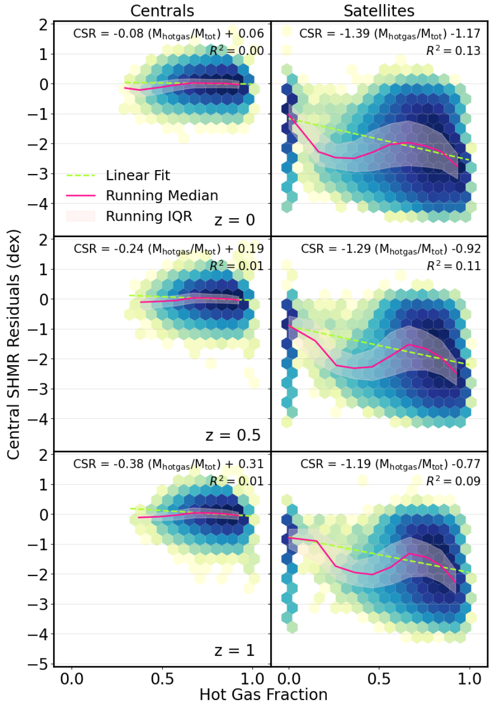

Figure A1. The relationship between the central SHMR residual and hot gas fraction in central galaxies and satellite galaxies across our redshift snapshots. The pink line represents the running medians, and the green line is the output of a linear regression analysis on the galaxies of each subplot.

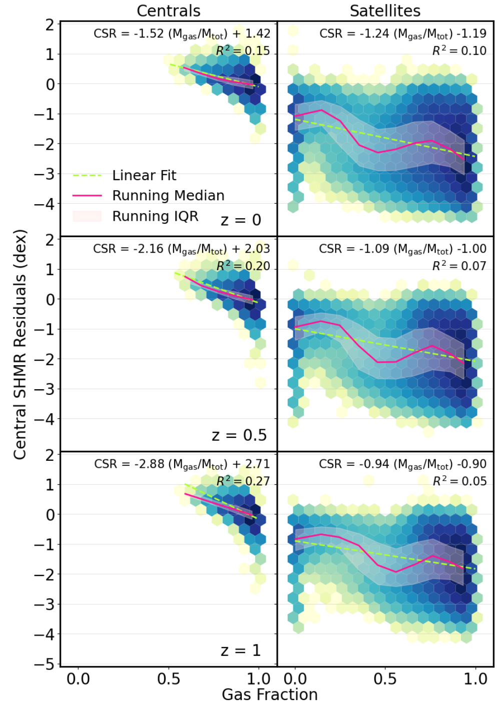

Figure A2. The relationship between the central SHMR residual and gas fraction in central galaxies and satellite galaxies across our redshift snapshots. The pink line represents the running medians, and the green line is the output of a linear regression analysis on the galaxies of each subplot..

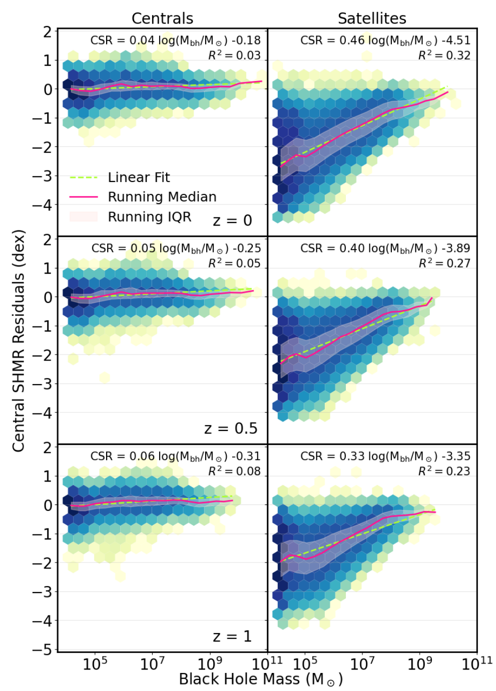

Figure A3. The relationship between the central SHMR residual and black hole mass in central galaxies and satellite galaxies across our redshift snapshots. The pink line represents the running medians, and the green line is the output of a linear regression analysis on the galaxies of each subplot.

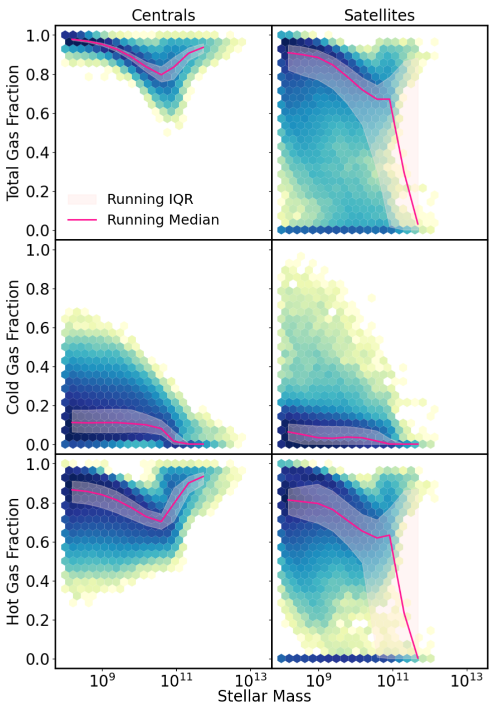

Figure B1. Breakdown of gas fractions with respect to stellar mass at redshift 0.

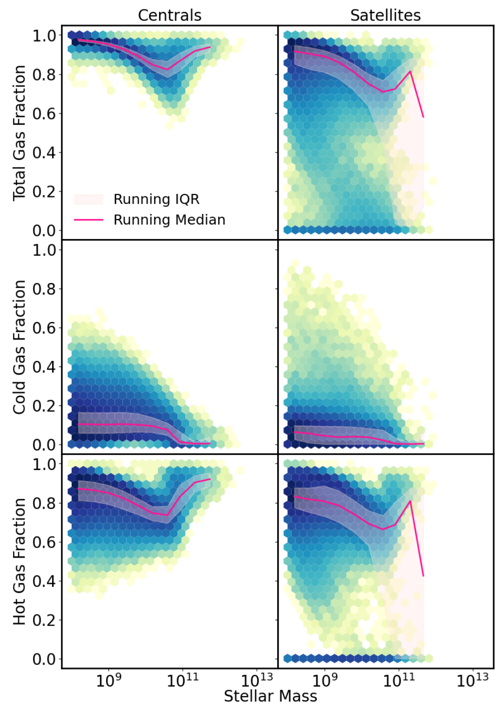

Figure B2. Breakdown of gas fractions with respect to stellar mass at redshift 0.5.

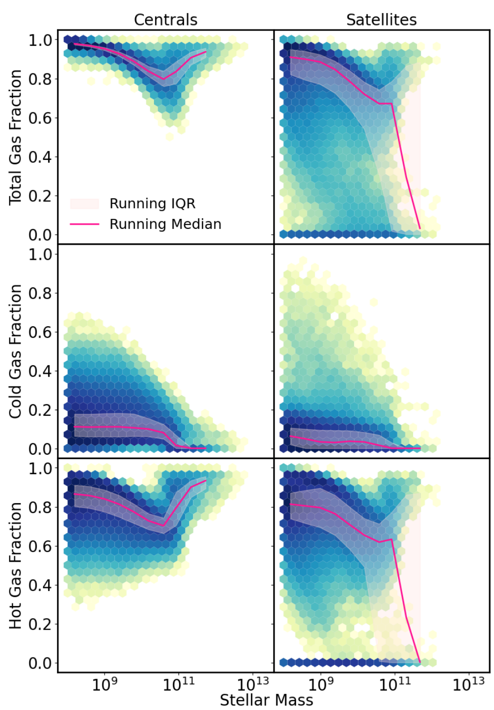

Figure B3. Breakdown of gas fractions with respect to stellar mass at redshift 1.