Draft version September 14, 2018
Typeset using LaTeX **twocolumn** style in AASTeX61

THE DIVERSITY AND VARIABILITY OF STAR FORMATION HISTORIES IN DIFFERENT SIMULATIONS

Kartheik Iyer, Sandro Tacchella, Lars Hernquist, Shy Genel, Chris Hayward, Neven Caplar, Phil Hopkins, Rachel Somerville, Romeel Dave, Ena Choi, and Viraj Pandya

ABSTRACT

As observational constraints on the star formation histories of galaxies improve due to higher S/N data and sophisticated analysis techniques, we have a better understanding of how galaxy scaling relations, such as the stellar mass star-formation rate and mass metallicity relations, evolve over the last 10 billion years. Despite these constraints, we have still a poor understanding of how individual galaxies evolve in these parameter spaces and therefore the physical processes that govern these scaling relations. We compile SFHs from an extensive set of simulations at $z = 0$ and $z = 1$, ranging from cosmological hydrodynamical simulations (Illustris, IllustrisTNG, MUFASA, SIMBA), zoom-in simulations (FIRE-2, VELA, Choi+17), semi-analytic models (Santa-Cruz SAM) and empirical models (UniverseMachine) to investigate the timescales on which star formation rates vary in different models.

We quantify the diversity in SFHs from different simulations using the Hurst index and use a random forest based analysis to quantify the factors that drive this diversity. We then use a power spectral density based analysis to quantify the SFH variability as a function of the different timescales on which different physical processes act. This allows us to assess the impact of different (stellar / black hole) feedback processes, which affect the star formation on short timescales, versus gas accretion physics, which affect the star formation on longer timescales. We find that variability on the longest timescales is not sufficient to explain the diversity in SFHs, with effects due to gas recycling, stellar feedback and winds contributing to the power on shorter timescales. Observational constraints in PSD space will help constrain the relative strengths of these processes.

[This is the report for my project during the Kavli Summer Program in Astrophysics (KSPA 2018).]

Keywords: galaxies: star formation — galaxies: evolution

Corresponding author: Kartheik Iyer
kgi1@physics.rutgers.edu

1. INTRODUCTION

A star formation history (SFH) is a stellar record of the various processes that shaped a galaxy since its birth to its current physical state. This includes both processes that enhance star formation, like the infall of pristine gas from cosmic filaments or wet mergers, and processes that quench star formation, such as tidal or ram pressure stripping of gas, or the heating of cold gas due to AGN or stellar feedback processes.

Recent years have seen a lot of improvements in simulating SFHs to greater physical accuracy in cosmological simulations by modeling processes such as feedback and black hole growth (Davé et al. 2016a; Vogelsberger et al. 2014a,b; Pillepich et al. 2017; Weinberger et al. 2016; Hopkins et al. 2014). In Semi-Analytic models, this corresponds to the addition to well-motivated subgrid recipes (Somerville et al. 2008, 2015; Yung et al. 2018; Brennan et al. 2016). In addition to these two methods, empirical models have developed techniques to estimate the SFHs of individual galaxies from global scaling relations and constraints (Behroozi et al. 2018; Tacchella et al. 2018; Moster et al. 2018). There have also been significant improvements in techniques for extracting information about the SFHs of individual galaxies from observations (Iyer & Gawiser 2017; Pacifici et al. 2016; Carnall et al. 2018; Leja et al. 2016; Heavens et al. 2000; Tojeiro et al. 2007), which give us an additional power to constrain galaxy evolution at different epochs, and observationally quantify the diversity of SFHs at a given stellar mass.

SFHs from both simulations and observations have been studied with the purpose of answering specific questions (Carnall et al. 2018; Tacchella et al. 2018; Torrey et al. 2017) - determining the rates at which galaxies grow and quench, and the timescales on which star formation rates are variable. It is harder to quantify the factors that set the diversity of SFHs within and across different simulations, as well as the effective timescales across which different physical processes such as galactic winds, stellar and AGN feedback leave imprints in the SFH. We propose a framework of analysing the SFHs of both individual galaxies as well as ensembles using their Power Spectral Densities (PSDs) to explicitly quantify the amount of energy in fluctuations of different timescales. Using this framework, we look at the SFHs from different simulations to quantify the timescales on which different physical processes affect the SFHs of galaxies.

2. DATASET:

For our analysis, we collect and analyze star formation histories from cosmological [Illustris, IllustrisTNG, MUFASA, SIMBA] and zoom-in [FIRE-2, Choi+17, VELA] hydrodynamical simulations, a s semi-analytic model [Santa-Cruz SAM] and an empirical model [UniverseMachine]. The SFHs from these sources have a wide range of diversity, both in the implementation of sub-grid physics like star formation and feedback prescriptions and the observations they are tuned to match. Analyzing them together using a common framework of tools allows us to understand both the differences between the different simulations. In addition to this, we adapt our framework to understand what sets the diversity and variability of star formation within individual simulations, which in turn allows us to better understand the processes that govern star formation and the timescales they act on. In the current work, we analyze these simulations at $z = 0$.

- Illustris (Vogelsberger et al. 2014b; Genel et al. 2014)

  The Illustris project is a large-scale hydrodynamical simulation of galaxy formation. The model includes recipes for primordial and metal-line cooling, stellar eveolution and feedback, gas recycling, chemical enrichment, supermassive black hole growth and AGN feedback.

- IllustrisTNG (Pillepich et al. 2017; Weinberger et al. 2016)

  A significantly updated version of the original Illustris project, IllustrisTNG carries over recipes for star formation and evolution, chemical enrichment, cooling, feedback with outflows, growth and multi-mode feedback from Illustris. In addition to this, it incorporates new black hole driven kinetic feedback at low accretion rates, magnetohydrodynamics and improvements to the numerical scheme. In addition to the regular TNG100 run, we use a number of $25\mathrm{Mpc}^3$ small boxes that vary the different physics models to quantify the effects of feedback and winds on the SFHs of galaxies. While the statistics of these SFHs may be subject to some cosmic variance in comparison with TNG100, the small-volume runs can be compared self-consistently with each other.

- MUFASA (Davé et al. 2016b)

  The MUFASA meshless hydrodynamical simulations include prescriptions for cooling and heating with Grackle, star formation and feedback from massive stars using scalings from FIRE (Hopkins et al. 2014) producing outflows that are most tightly correlated with mass, almost independent

of redshift. The model also includes feedback from long lived and AGB stars, and an implementation of quenching feedback that quenches galaxies by heating all the gas in massive halos (except gas that is self-shielded). At $z \sim 0$, their quenching mass is $\sim 10^{12} M_{\odot}$.

- SIMBA (Dave et al, in prep.)

  The SIMBA cosmological galaxy formation simulations are built on the MUFASA simulations including black hole growth and feedback. The simulation uses a torque limited BH accretion model, along with BH feedback that operates on two modes depending on the eddington rate (Anglés-Alcázar et al. 2017).

- Santa-Cruz SAM (Somerville et al. 2008, 2015; Yung et al. 2018; Brennan et al. 2016)

  The Santa-Cruz Semi-Analytic Model contains a number of well motivated sub-grid prescriptions that are used in conjunction to the Bolshoi-Planck dark matter N-body simulations' merger trees to construct populations of galaxies that are tuned to match observations at z=0. The model implements two modes of star formation in the model: a 'normal' mode following the Schmidt-Kennicutt relation along with exploding supernovae which drive outflows with recycling that occurs in isolated discs, and a 'starburst' mode that occurs as a result of a merger or internal disc instability.

- UniverseMachine (Behroozi et al. 2013, 2018)

  The UniverseMachine is an empirical model that determines the star formation rates of galaxies as a function of their host haloes' potential well depths, assembly histories, and redshifts. The model uses Bolshoi-Planck DM merger trees, and a variety of observational constraints including observed stellar mass functions, SFRs, quenched fractions, UV luminosity functions, UV-stellar mass relations, autocorrelation functions, and quenching dependence on environment to constrain its free parameters.

- FIRE-2 (Hopkins et al. 2014)

  The Feedback In Realistic Environments simulations consider a fully explicit treatment of multiphase ISM, stellar feedback. This is important since the M_*-M_{halo} relation is sensitive to feedback physics. Supernova feedback alone is not enough, radiative feedback (photo-heating and radiation pressure) is needed to destroy GMCs and enable efficient coupling of later SNe to gas. Feedback also produces reservoirs of gas for extended SF at late times. Gas particles in the simulation follows an ionized+atomic+molecular cooling curve, and star formation is only allowed in dense, molecular, self-gravitating regions above a density threshold. Stellar feedback includes contributions from radiation pressure, supernovae, stellar winds, photoionization and photo-electric heating.

- Choi+17 simulations (Choi et al. 2017, 2012)

  The Choi+17 zoom-in simulations run both with and without AGN feedback allow us to better quantify the effect of AGN on the SFHs of their host galaxies. The code uses the TreeSPH-code GADGET-3 (Springel 2005), with a BH-feedback model that includes thermal and momentum feedback, x-ray feedback. BH seeding occurs at $M_{vir} > 10^{11} M_{\odot}$.

- VELA (Ceverino et al. 2014; Zolotov et al. 2015)

  The VELA suite of hydrocosmological simulations considers a set of ~ 35 moderately massive galaxies, of which we consider 17 that have evolved to $z \sim 1$. The code implements gravity and hydrodynamics, gas and metal cooling, UV background photoionization, stochastic star formation, gas recycling and chemical enrichment, thermal feedback from supernovae (Ceverino et al. 2010, 2012), and a new implementation of feedback from radiation pressure (Ceverino et al. 2014), with no artificial shutdown of cooling. Stellar feedback from winds and SNe explosion causes heating at a constant rate over 40 Myr following star formation. This may not overcome the cooling rate, depending on gas conditions in the star-forming regions (Dekel & Silk 1986; Ceverino & Klypin 2009).

Figure 1 shows the distribution of stellar mass at $z = 0$ from all the different simulations. The different normalizations for the cosmological simulations corresponds to the difference in the volumes of the boxes they were run in. The zoom-in simulations have a much small number of galaxies, with FIRE-2 set up to produce massive dwarfs across a range of stellar masses, and the Choi+17 simulations focusing on massive galaxies impacted by AGN feedback.

3. THE DIVERSITY OF SFHS AT DIFFERENT MASSES AND REDSHIFTS

3.1. *Comparing star formation histories from different simulations*

Star formation histories in individual simulations show a wide range of trajectories that are influenced by a

multitude of factors like the dark matter accretion history and mergers, environment, inflows and outflows, and stellar and AGN feedback. In recent times, this has also been quantified by obervations, such as the SFHs of dwarfs in Weisz et al. (2011), ensemble studies like Pacifici et al. (2016), or the resolved SFHs from Hsieh et al. (2017).

For each simulation, we look at the SFHs for galaxies in a different bins of stellar mass, plotting them against normalized time. This allows us to qualitatively see how the SFHs formed their mass across cosmic time. We use a bin width of 0.2 dex for all simulations except the zoom-in simulations, for which we use a bin width of 0.5 dex centered on the masses in the figure. From Figure 2, we can see that Illustris galaxies as an ensemble have roughly constant SFRs for most of their lifetime, especially for the lower mass galaxies. Quenching due to improved wind and AGN feedback prescriptions in IllustrisTNG are also immediately apparent in the $M_* \sim 10^{10.5}, 10^{11} M_\odot$ bins as compared to Illustris, and with SIMBA compared to MUFASA. There is a considerable amount of difference in when galaxies form their stellar content in the lowest-stellar mass bin ranging from relatively early star formation for the UniverseMachine, MUFASA and SIMBA, with very different amounts of diversity, to constant and late-time star formation in the Illustris, TNG and SAM models. While MUFASA SFHs show the most diversity in SFHs at early times, the The UniverseMachine SFHs have a clear evolution with stellar mass, since they have been tuned to reproduce the SFR-M_* correlation and other scaling relations across a range of redshifts.

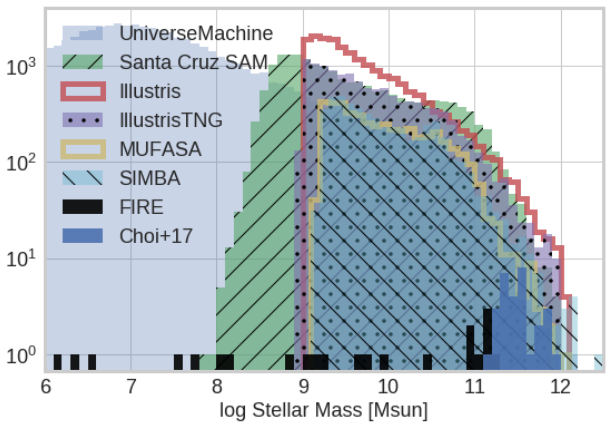

Figure 1. The distribution of galaxies in stellar mass at $z \sim 0$ for our samples from all the simulations and models we consider.

3.2. *The differences between the star formation history and mass accretion history:*

While many simulations track the stellar and dark matter accretion history ($M_*(t)$, $M_{DM}(t)$) of a galaxy over evolving snapshots, it is important to bear in mind the distinctions that these quantities possess from the star formation history, which we define as the star formation rate at each epoch in time (SFR(t)).

If all the star formation in a galaxy happened purely in-situ, ie. within the most massive progenitor, the star formation history and mass accretion history would be practically the same. However, every time a merger occurs, or a satellite gets accreted onto the main progenitor being tracked, this would affect the mass accretion and star formation histories in different ways. If the subhalo that gets accreted contains a significant amount of stellar content, as is the case for many major mergers, it results in a sudden spike in the mass accretion history. However, the star formation history would account for this by showing increased SFR at the times when those stars were formed. Thus, while the galaxy has the same mass at the time of observation, the distribution in time of when the galaxy obtained that mass can be quite different.

This difference is shown in Figure 3, where we look at the mass accretion history vs the star formation history for three different galaxies. For the least massive galaxy, the SFH and $M_*(t)$ are quite similar, since most of the mergers that occur do not bring in any additional stellar material, and gas that is accreted through inflows predominantly results in SFR within the central progenitor. On the other hand, the most massive galaxy grows mostly by mergers, which results in a mass accretion history that is very tightly coupled to the dark matter accretion history. Comparing this to the SFH shows that most of the stellar mass that is brought in through mergers is quite old, which results in a flattening of the star formation rate at late times. The middle panel shows an example for a galaxy that is a combination of the two effects, that probably experiences a few wet mergers but no major dry mergers, due to which all three quantities are rather closely related to each other.

It is important to keep in mind that the star formation histories are the more observationally accessible quantity, since it is possible to distinguish between stellar populations of different ages to make estimates of when they were formed. However, to make any claims about the mass accretion rate, it is extremely important to take the effects of mergers into account, especially at higher masses. While this can be estimated using empirical models (Behroozi et al. 2018; Moster et al. 2018), it is beyond the scope of this work.

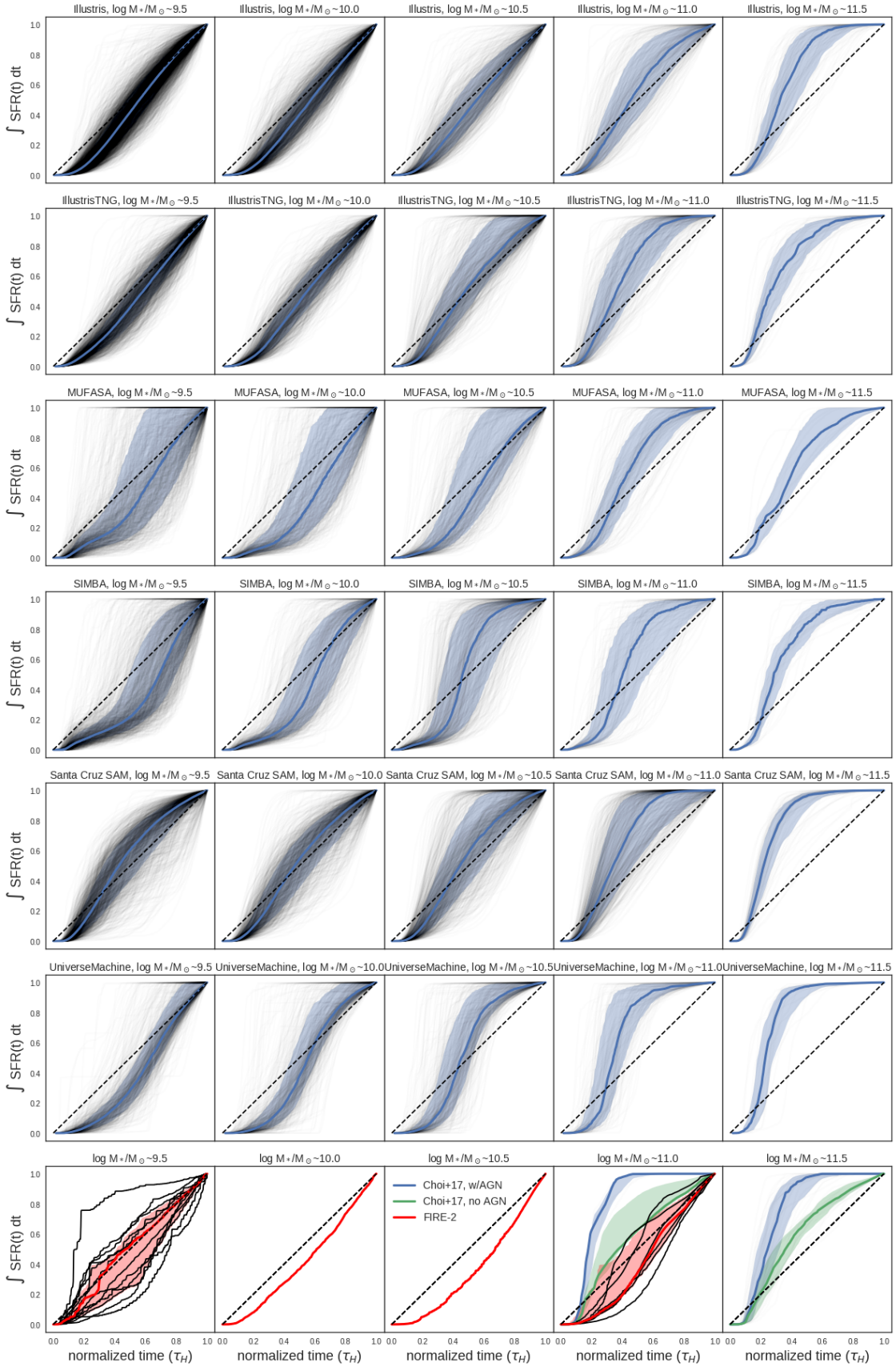

Figure 2. Plot illustrating the diversity of SFHs from the various cosmological simulations [Illustris, IllustrisTNG, MUFASA, SIMBA], zoom-in simulations [FIRE-2, Choi+17], Semi-Analytic [Santa-Cruz SAM] and empirical models [UniverseMachine] at $z \sim 0$ in different stellar mass ranges. The solid blue line and blue shaded region show the median and 16-84 percentile region for the SFHs respectively. The black dashed line would be the trajectory for a galaxy with constant SFR throughout its lifetime, with high SFRs given by slopes > 1 and low SFRs given by slopes < 1.

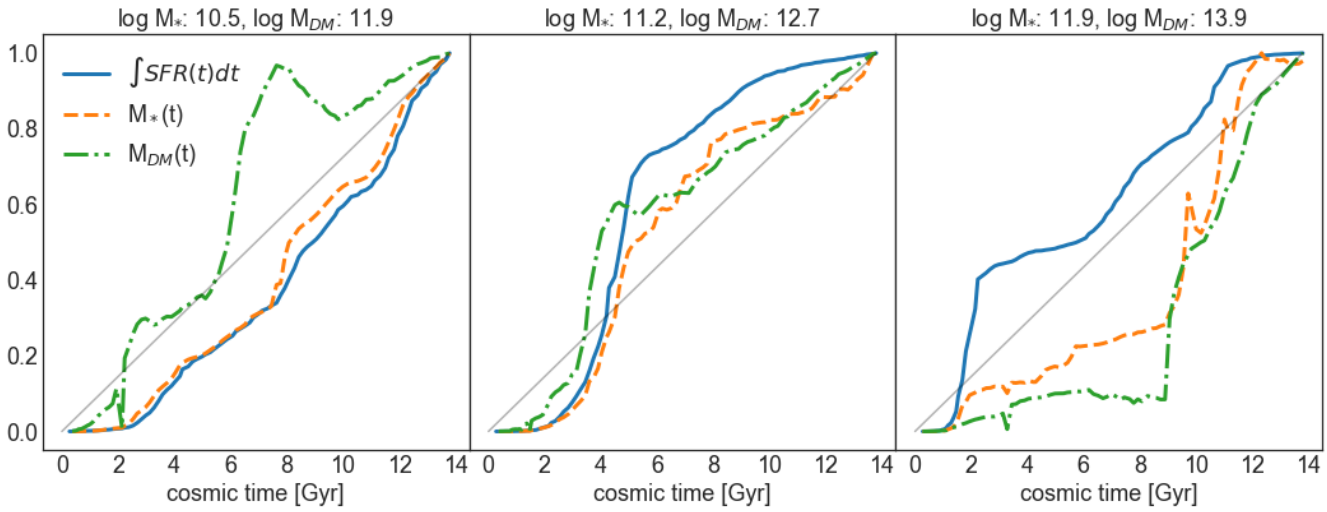

Figure 3. The difference between star formation histories and mass accretion histories, shown using the SFH, $M_*(t)$ and $M_{DM}(t)$ for three galaxies at different halo masses that have very different evolutionary pathways. All the curves have been normalized to have a maximum value of 1, in order to best compare their trajectories.

3.3. *Diversity as a function of the long-timescale variability of SFHs:*

To quantify the variability of star formation histories at the longest timescales, we use the Hurst parameter (Mandelbrot & Van Ness 1968; Kelson 2014), which is a measure of the long-timescale memory of a time-series and determines how smooth an SFH is. $H = 0.5$ is brownian motion, with $H \in [0.5, 1]$ encapsulating fractional Brownian motion, where the different timesteps aren't completely independent and are related by a covariance term. Calculating the Hurst index allows us to understand the long-timescale behaviour, and thus the diversity of SFHs. Kelson (2014) considers this and finds that the amount of diversity increases as H goes from 0.5 to 1. To determine the Hurst parameter for a given SFH, we use the `nolds.hurst_rs` module in python, which determines the hurst exponent using the rescaled range method. We find that the Hurst exponent correlates well with the diversity of SFHs in a given mass bin, with the SFHs closest to the 1:1 line having $H \sim 0.5$, and large excursions corresponding to $H \sim 1$.

We do this for all the different simulations that have a significant number of galaxies and summarize the results in Figure 4. The amount of deviation from the diagonal line in Figure. 2 is proportional to the average Hurst index at that stellar mass, with the diversity being proportional to the range in H that the SFHs span, as well as how close their values are to 1. Since the long-term memory of SFHs in our case also includes quenching due to AGN feedback, this is degenerate to some extent with the diversity quantified by the Hurst parameter. A set of SFHs with $H \sim 1$ might be more diverse than a set of SFHs with $H \sim 0.75$ since a larger H contains a bigger subset of possible SFHs than a smaller H, but it could also represent a larger overall deviation from the diagonal constant SFR line. For this reason we consider H values as well as their spread at a given stellar mass as the indicator of diversity. In a future analysis we will consider this effect with a modification that would make the SFHs quasi-stationary, since the H alone doesn't seem to be a good indicator of diversity at the highest stellar masses.

We see that Illustris and TNG are most diverse at $M_* \sim 10^{10.5} - 10^{11} M_\odot$ (largest spread in H). MUFASA and SIMBA have a large amount of diversity and become less diverse at highest masses. The SAMs have the largest amount of diversity at lower masses, with a smooth transition from diversity to quenching as we go to higher masses. UniverseMachine is less diverse than the SAM, MUFASA and SIMBA at all stellar masses, but more diverse than the Bolshoi-Planck dark matter merger trees that they are built off of. This likely comes from the perturbations to the SFRs from the baryonic physics that the merger trees do not account for.

3.4. *Explaining the diversity within a simulation using galaxy properties:*

We see that there is a variation of H with stellar mass, but there is also a large amount of scatter at a given mass. Since the SFH diversity depends on events like inflows, outflows and mergers in addition to feedback regulated star formation processes, it is important to analyze the full space of galaxy parameters that could correlate with the Hurst parameter - by looking at the variation

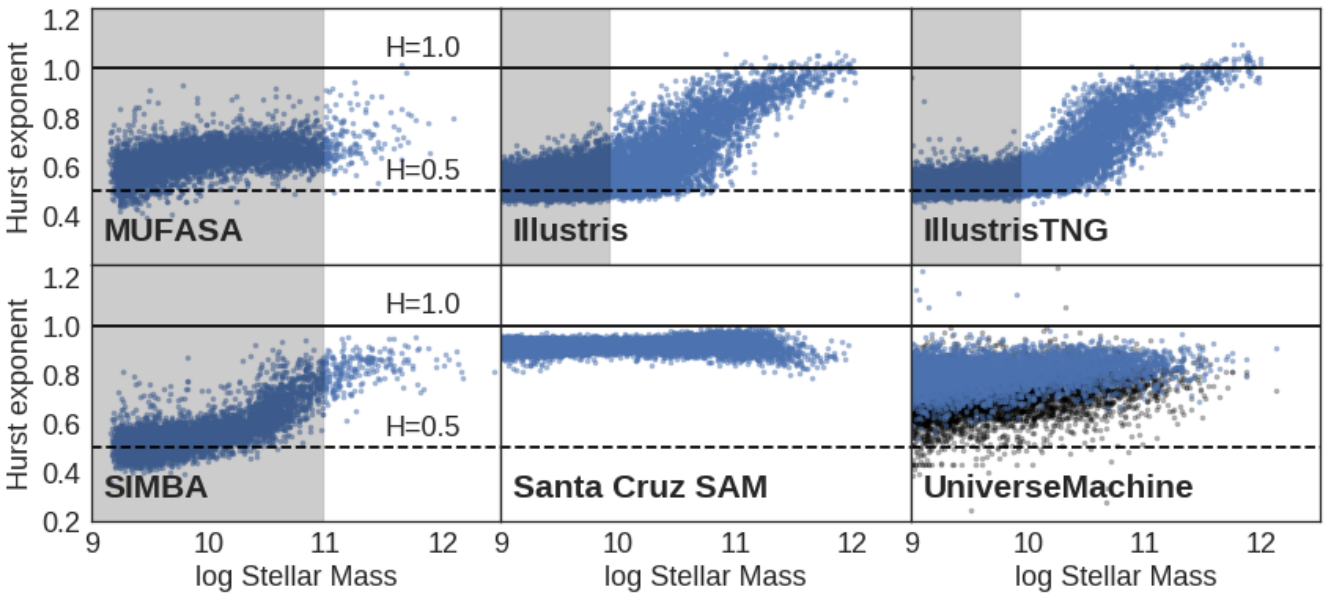

Figure 4. The hurst exponent calculated for all the simulations. The shaded region shows where contributions to shot noise make the SFH unreliable on timescales of $\sim 1Gyr$. Black points in the bottom-right panel show the Hurst exponents for the Bolshoi-Planck dark matter only merger trees that were used in computing the UniverseMachine SFHs.

with respect to quantities like halo mass, gas mass, black hole mass and accretion rate, gas phase metallicity, age and size of the stellar disk/bulge, etc. However, this is hard to visualize since the parameter space is heavily multidimensional.

However, machine learning algorithms can be used for feature selection in cases like this, allowing us to find the number of significant parameters that explain the variance in the Hurst parameter. We use Extremely Randomized Trees, implemented via `scipy.ensemble.ExtraTreesRegressor` to break down the variance of the Hurst parameter as a function of these labels. The results of performing this analysis on the $z \sim 0$ IllustrisTNG sample are shown in Figure 5. We find that majority of the variance in the Hurst parameter is explained by the four dimensional parameter space consisting of stellar mass, black hole mass, halo mass and the stellar size of the galaxy. The correlation of the Hurst parameter with stellar mass is interesting, but ultimately a product of the stochastic nature of star formation via the production and destruction of clouds. It is interesting, however, to see that the variance in the Hurst parameter is not explained simply by the stellar mass, but by four properties. While the first three are not completely independent and correlate almost equally well with the long timescale correlation (given by the Hurst index) at high masses, they are significantly uncorrelated at lower halo masses, which the random forest picks up on. For example, at a roughly constant stellar mass of $10^{10.4} < M_* < 10^{10.6}$, the scatter in the Hurst index in Figure 4 correlates strongly with the Black Hole mass, since the galaxies that have larger black holes are driven to stronger feedback, and thus have H closer to 1. The tertiary effect of Halo mass enters through the difference between smooth gas accretion and merger-driven growth, as discussed in Sec. 3.2. However, to understand this effect fully, it is important to look at not just the correlation on the longest timescales, but at the full power spectra of galaxy SFHs, in order to better understand which factors influence the variability on different timescales.

Since two of the quantities (stellar and halo mass) are integrated quantities that depend on the overall SFH and dark matter accretion history, respectively, it tells us that about half of the diversity comes from how the galaxy builds up its mass - i.e. from feedback regulated star formation on different timescales, and about 40% comes from its merger history. The rest of the variance is due to spatial effects that influence the distribution of star forming clumps in the galaxy, and how old/stable the disk is. Thus, to fully understand the diversity of SFHs, we now need to look at the variability of SFHs on different timescales. We look at this in more detail in the next section.

4. THE VARIABILITY OF SFHS AT DIFFERENT TIMESCALES

To explain the diversity of SFHs at a given epoch, we saw that it is not enough to look at the variability at the longest timescales, since physical processes that

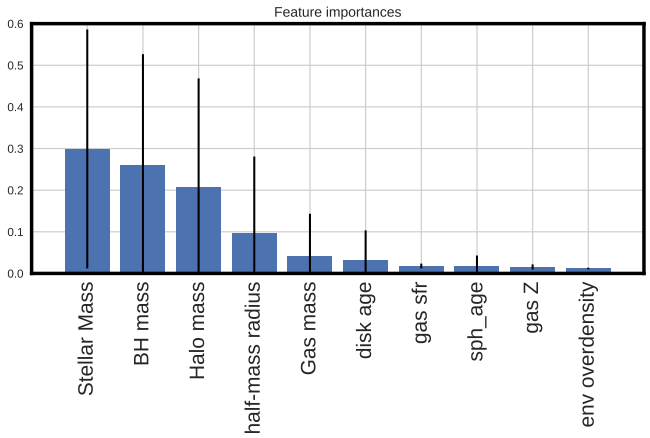

Figure 5. Results of a Random Forest analysis to explain the variance of the Hurst parameter as a function of the physical properties of a galaxy.

impact the SFH across a range of timescales influence the overall diversity of SFHs.

4.1. *Power Spectral Density*

For a continuous time series $\psi(t)$, the Power Spectral Density (PSD) is defined in terms of the Fourier transform $f(k) = \int dt e^{ikt}\psi(t)$ as $PSD(k) = |f(k)|^2$. While a SFH is a continuous time series in theory, practically we extract a discrete series from the simulations with some time step Δt. In this scenario, we define the PSD as a summation

$$PSD(f) = \frac{(\Delta t)^2}{\tau_H}\left|\sum_{n=1}^{N} SFR(t_n)e^{-ifn\Delta t}\right|^2 \quad (1)$$

where the SFH is sampled at discrete times t_n and τ_H is the Hubble time. We perform this computation using Welch's method (Welch 1967), implemented in the `scipy.signal.welch` module.

We wish to perform our PSD analysis on SFHs in both SFR and log SFR spaces. This effectively amounts to a weighting scheme for the SFHs, and in future work it might be advantageous to devise suitable weighting schemes for tackling specialized problems. In our case, going to log SFRs allows us to study parts of the SFH with very low SFRs, that are important regimes for phenomena like quenching or temporary quiescence. However, we find that the threshold of sensitivity to shot noise for log SFRs is much higher than the linear counterpart, allowing us to perform this analysis on only the highest-mass galaxies in our sample. On the other hand, linear SFR allows us to look at peaks of star formation that lie well above the threshold shot-noise limit, to identify features due to gas infall and mergers.

In our simulations, we often have time bins where the star formation rate is 0, due to the discreteness of the star particles. While this doesn't cause a significant effect in the linear analysis, except for the lowest-mass galaxies, it stops us from a straightforward extension of the analysis in log SFR space. While it is possible to look at the gas SFR at times when the simulation snapshots are available, this is not available at all time steps and does not give us a complete picture of the galaxy's SFH. In all the simulations that we consider, star formation rates in the gas at each time step act as a probability for when a star particle is formed. To account for this artificial discreteness, we re-bin the star formation history into time-bins of variable duration such that each bin has a minimum of one star particle in it. We then re-bin this corrected SFH into time-bins with the same original width as before, ensuring non-zero (if extremely low) SFRs in each bin. This better corresponds with what we see in nature, as well as being a more physical approximation of the processes that form stars in galaxies.

4.2. *Finding the lowest timescales we can probe*

Both cosmological (Illustris, TNG, MUFASA, SIMBA) and zoom-in simulations (VELA, FIRE-2, Choi+17) have limits on the lowest SFR possible in any given time bin that is set by the size of the star particles they use. All the simulations listed above turn gas into a star particle probabilistically depending on whether certain temperature and density conditions are met. In practice, this introduces portions in the SFH where the $SFR = 0$, punctuated by small spikes which contain $\mathcal{O}(1)$ star particles. The effect of this on the power spectrum is to introduce white noise on the timescales where the SFR is probabilistically populated by discrete star particles. Looking at the PSD of individual galaxies, we see the effects of this effectively Poisson-distributed 'shot noise' as a flattening as we approach short timescales. This depends on the amount of time the SFH spends in the vicinity of the minimum SFR threshold, set by

$$\langle SFR_{min}\rangle = \langle M_{*,sp}\rangle / t_{PSD} \quad (2)$$

where $M_{*,sp}$ is the average stellar mass of the star particles in the simulation, and t_{PSD} is the timescale being probed. From this relation, we see that the effects of shot noise on the PSD are greater on short timescales, as well as for simulations that have more massive star particles. However, finding the amount of time SFHs at a given stellar mass spend below SFR_{min} is a nontrivial task, depending on the shape of the SFH itself, and the number of the fluctuations around the median shape that could take it below SFR_{min}.

In the simplest case, given an SFH that is simply a constant $SFR_{const} = \psi_{mean}$ + stochastic fluctuations $SFR_{fluct} = (N(0, \psi_\sigma))$, the distribution of SFR(t) at any given time is simply given by a gaussian $N(\psi_{mean}, \psi_\sigma)$. $\psi_{mean} = M_*/\tau_H$ is set by the stellar mass of the galaxy, where τ_H is the age of the universe at the epoch of interest. The amount of time any SFH at a given stellar mass spends below a threshold SFR is then given by

$$
\begin{aligned}
&t(SFR < SFR_{min}|M_*, z) \\
&= \tau_H \int_{-\infty}^{SFR_{min}} exp\left(-\frac{(SFR - \psi_{mean})^2}{(\psi_\sigma)^2}\right) dSFR \\
&= \frac{\tau_H \psi_\sigma \sqrt{\pi}}{2} \left(1 + erf(\frac{M_{*,sp}/t_{PSD} - M_*/\tau_H}{\psi_\sigma})\right)
\end{aligned}
$$

Using this, we can set a threshold on the amount of shot noise, and limit our analysis to timescales above that. We then consider two more complicated cases,

- Where the SFH has a power spectrum with a non-trivial power law ($PSD(f) \propto f^{-2}$) as seen in the simulations at long timescales, and postulated in Kelson (2014), and
- Where the median SFH is not stationary, i.e. Evolves with time following the SFR-M* correlation (Behroozi et al. 2018; Moster et al. 2018).

Since deriving the average time an SFH spends at the shot noise limit analytically is involved for these cases, we perform a series of numerical experiments by generating mock SFHs that satisfy these criteria and modeling the effects of discrete star particles in the same way as the simulations. To do this, we discretize the mock SFH by rounding the SFR in each time bin to its nearest number of star particles, and consider the excess as the gas probability that a star particle will be formed in that time bin.

We then compute the power spectra for these SFHs before and after the discretization procedure, and quantify the timescale at which the divergence from the original PSD exceeds a certain threshold (here 0.2 dex). We also tried fitting the PSD corresponding to the discretized SFH with a broken power-law to quantify the timescale at which the transition from $\alpha = 2$ to white noise ($\alpha = 0$) happens, and find that our results do not significantly change. Based on these numerical experiments, figure 7 shows the thresholds for the case where we consider the case with the a constant SFR with $\alpha = 2$, and figure 6 for the case where we consider an ensemble of SFHs that yield the Schreiber et al. (2015) SFR-M* correlation at different epochs. Results of repeating the analysis using log SFR as opposed to linear SFR are shown in Appendix A. For all cases, the figures can be read in two ways:

- Read horizontally, the figures give the minimum timescale to which we can study the PSDs for galaxies in a given stellar mass bin at a particular epoch.
- Read vertically, the figures give the minimum SFR (and therefore the minimum stellar mass) needed to probe a certain timescale or regime of the PSDs of galaxies.

Below these stellar masses (and timescales) the effects of discretization of the star particles begins to dominate the SFRs, and thus the PSDs. While the solid black lines are the results of the experiment for SFHs that produce an SFR-M* correlation with ~ 0.3 dex scatter, the figures also contain two additional curves for each star particle mass, showing the sensitivity curves for the case with ~ 0.1 dex and 0.5 dex scatter. We find that the sensitivity improves with an increase in the scatter, since it corresponds to more power in the PSD at all timescales - this results in a larger portion of the PSD above the shot noise threshold. Since the analysis with log SFR effectively re-weights the SFHs such that lower SFRs (and thus the shot noise due to discrete star particles) is on a more comparable footing to the regions of the SFH with high SFR, the sensitivity to short timescales is much worse for this case, as seen in Appendix A.

4.3. *What sets the variability within a simulation?*

We use the IllustrisTNG 25 Mpc^3 small volume runs to quantify the effects of changing the strengths of feedback and winds, as well as the way they are implemented. This allows us to see the effects on the PSDs in different stellar mass bins, and better understand the timescales on which the feedback and wind prescriptions influence the SFHs of individual galaxies.

From Figure.8, we see that TNG with the slower Illustris winds (as compared to TNG), cause a mass-dependent deficit in the PSD on timescales of 2-4 Gyr. This can also be explicitly seen for the case of slower winds in TNG itself. Corresponding to the slower winds case, stronger winds result in a mass-dependent increase in the PSD on $\sim 1Gyr$ timescales. In comparison to other effects like AGN, the effects of winds manifest across a wide range of effects in the PSD. For the highest mass bin, faster winds seem to reduce the variability at longer $\sim 3-4Gyr$ and shorter $< 100Myr$ timescales and redistribute it on intermediate $\sim 250Myr$ timescales.

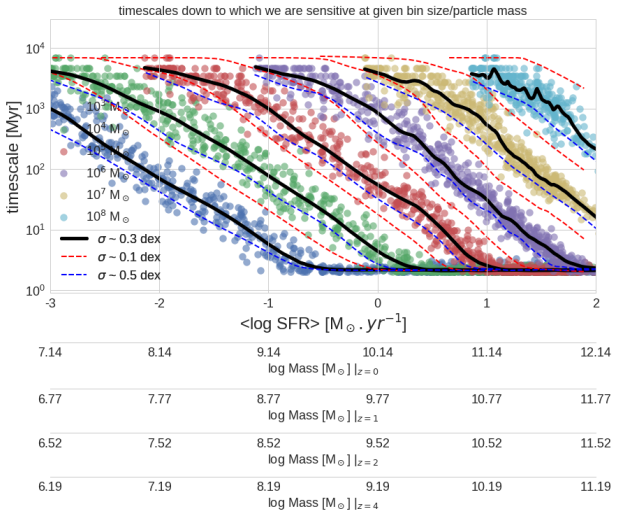

Figure 6. Quantifying the lowest timescales we can probe at different stellar masses for galaxies that follow the SFR-M$_*$ correlation from Schreiber et al. (2015) with perturbations that add different amounts of scatter to the relation (between 0.1 to 0.5 dex) for different masses for the star particles following the procedure described in sec. 4.2.

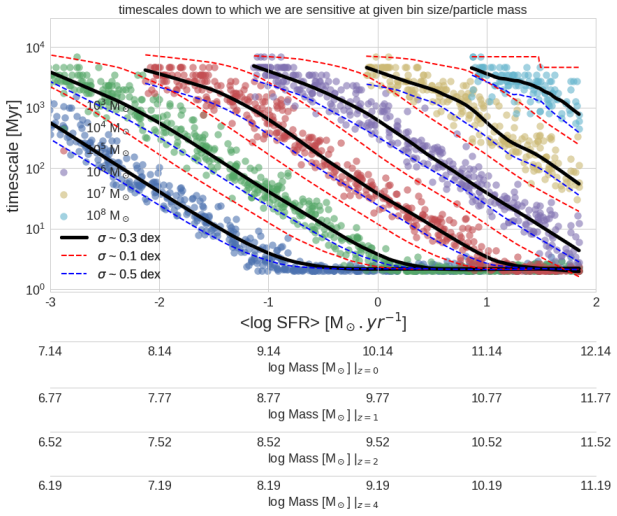

Figure 7. Quantifying the lowest timescales we can probe at different stellar masses for galaxies with a constant SFH + perturbations that add different amounts of scatter to the relation (between 0.1 to 0.5 dex) for different masses for the star particles following the procedure described in sec. 4.2.

The lack of AGN kinetic feedback (no BH kin fdbk) seems to lower the PSD on the longest timescales (> 10 Gyr), which might be due to less efficient quenching since the lack of kinetic feedback implies that not as much gas is being blown out of the galaxy by AGN driven winds.

4.4. *The Power Spectral Density of galaxies at different masses and redshifts*

In Figure 9 we present the power spectral densities for all galaxies in a given stellar mass bin, for all the different simulations that we consider. The PSDs are generated using all the galaxies in a mass bin for a given simulation, truncated to 1000 randomly chosen galaxies if there are more than that many at any given mass, using the same bins as in Figure 2. We truncate the PSD below the thresholds shown in Figure 7 in each stellar mass bin corresponding to the mean stellar masses of the star particles in these simulations. Since the SAM and UniverseMachine do not have star particles, they do not need to undergo this truncation procedure. For the FIRE-2 simulation, the resolution is usually high enough that the limit is never reached in these plots. The shaded regions show the 16th to 84th percentiles for the PSDs of each simulation.

We see that:

- At longer timescales, ($1Gyr+$) the shapes and slopes of the PSDs are different based on what models they use for winds, AGN feedback, and other quenching mechanisms - both at low and high stellar masses.
- We see that SIMBA generally has a higher value of PSD than MUFASA, with the difference growing as we go to higher masses. This can be linked to the inclusion of AGN with torque-limited accretion, which indirectly channels power from disk instabilities into the regulation of SFR over long timescales. The difference on the longest timescales decreases for the highest mass bins, where the quenching prescription for the MUFASA galaxies produces similar SFHs.
- We also see that IllustrisTNG diverges from Illustris at $\sim$ 1 Gyr timescales, due to an improved metallicity dependent prescription for winds blowing gas and stars out of the galaxy, and see a mass dependent divergence at $\sim 10Gyr$ due to an improved AGN feedback model.
- On the longest timescales, we see that the FIRE-2 dwarfs have less power since they do not currently contain an implementation for AGN feedback and thus don't experience episodes of periodic quenching due to it.
- The PSDs generally have a slope of ~ 2 in the 100 Myr to 1 Gyr timescales at all masses. This might be tied to the stellar mass - dark matter cross power spectrum, which also has a slope of $\sim$ 2

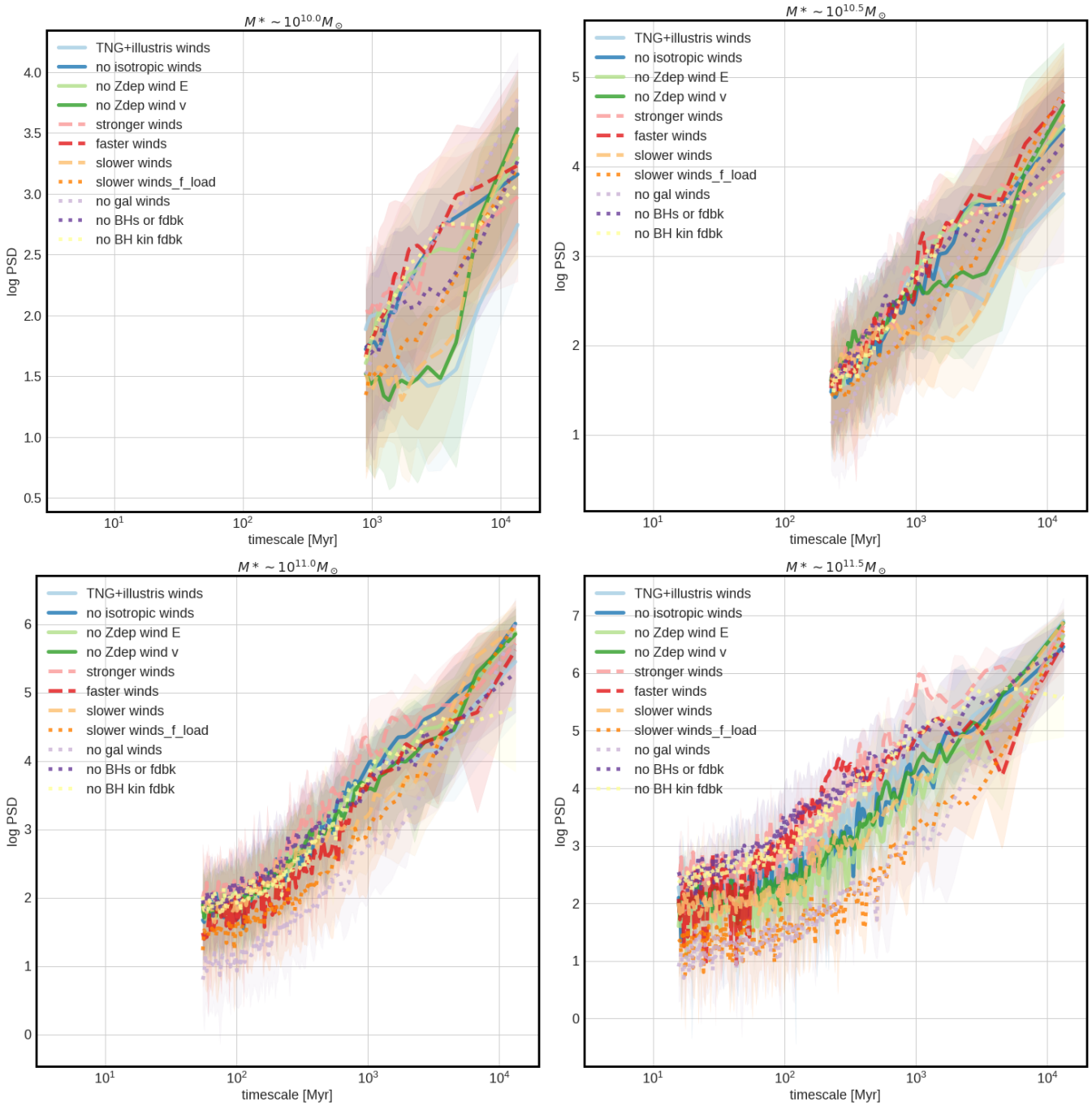

Figure 8. The power spectral densities of galaxies in a series of small-box IllustrisTNG runs with different conditions. The lines and shaded regions show the median and 68 percentile distributions at each point in time. The PSDs are truncated at the point below which shot noise contamination begins to flatten them out. The different models allow us to observe the pasts of the PSD that are affected by different physical conditions, allowing us to understand the contributions at different timescales.

throughout, indicating that it includes the contributions from the stochastic mergers of subhalos and galaxies leading to their growth. It is presumbable that star formation on these timescales is effectively approximated by brownian motion - similar to Kelson (2014), and other studies that look at the scatter around the SFR-M_* correlation.

- The FIRE-2 dwarfs have significantly higher power than the other simulations on the 100 Myr to 1 Gyr timescales, indicating that feedback does play

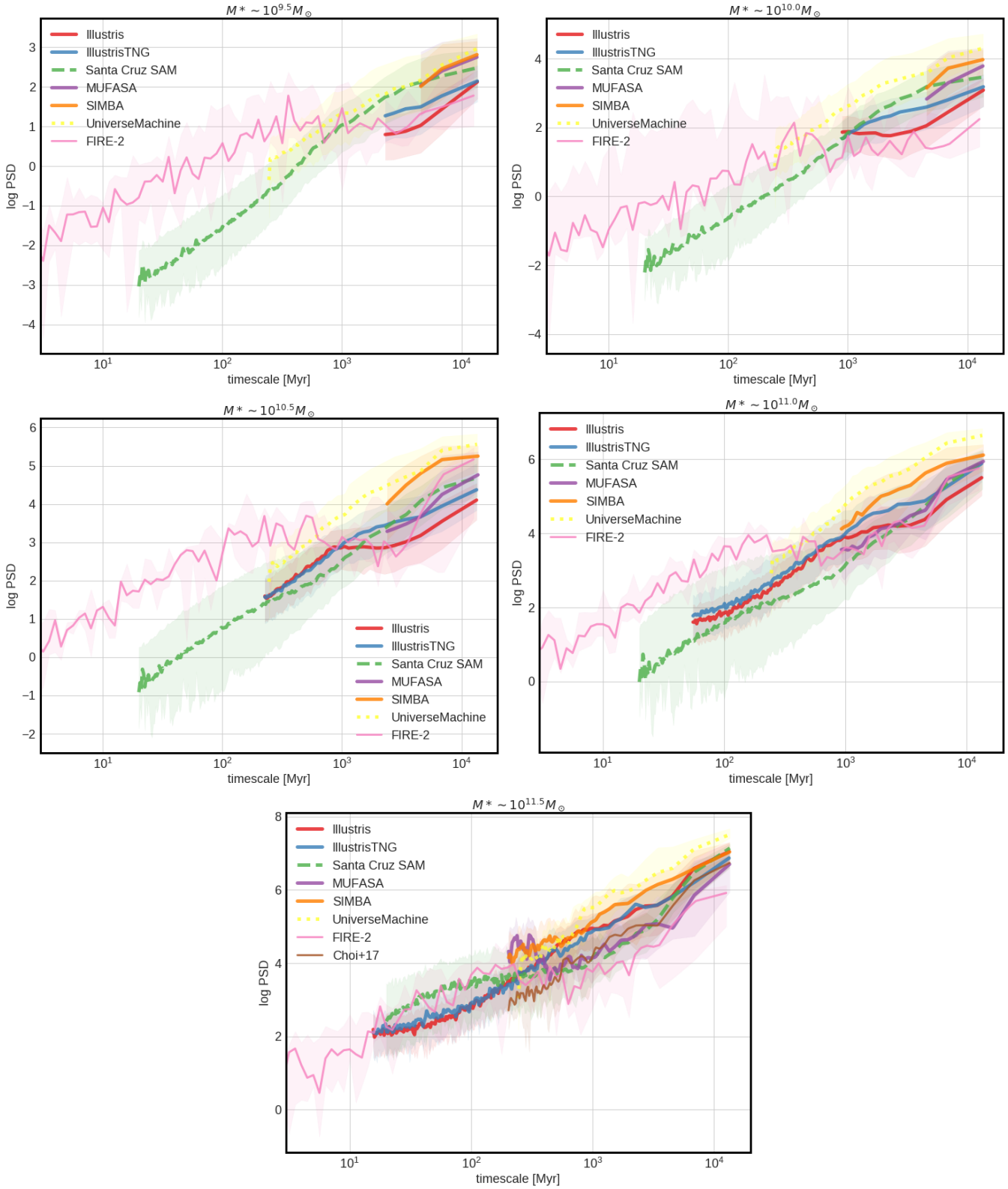

Figure 9. The power spectral densities of galaxies across our collection of simulations in different mass bins. The lines and shaded regions show the median and 68 percentile distributions at each point in time. The PSDs are truncated at the point below which shot noise contamination begins to flatten them out.

a very important role in regulating the SFRs of these galaxies.

- For the highest mass bin ($M_* \sim 10^{11.5} M_\odot$), almost all the PSDs show very similar shapes, with slopes

~ 2 that corresponds to a Hurst parameter of 1. This is in line with what we see in Figure 4.

- For the highest mass bin, the MUFASA PSD shows a broad peak at $\sim 200-400$ Myr timescales. This contribution to the signal comes from the wind decoupling time, which is $\sim$ 250 Myr at $z \sim 0$. Gas particles that get converted into wind particles get ejected into the ISM but don't escape. At the end of the recoupling timescale, these particles are converted back into gas particles to form stars again. This isn't seen in Illustris because the winds are stronger, but can be seen in a small-box run of IllustrisTNG (0502) that has slower winds or (0504) which has slower winds with a fixed loading factor.

5. DISCUSSION AND FUTURE WORK

5.1. *Model comparisons using observations*

We see that the PSDs of galaxies at a given stellar mass can vary considerably depending on the choice of simulation and the timescale of interest. While some of this variance might be due to random-number effects in the simulations Genel et al. (2018), the majority arises from the different implementations of the physical recipes that govern the heating, cooling, inflows and outflows of gas within galaxies. Therefore, if we can observationally constrain the PSD at a given timescale or the PSD slope across a range of timescales, it is possible to discriminate between these physical models and determine which effects are dominant at a given timescale as well as the relative strengths of the different effects.

Observational studies have shown that it is possible to probe a range of timescales from the integrated light of galaxies, with spectra that have sufficiently high S/N and resolution (Tojeiro et al. 2007; Pacifici et al. 2016; Leja et al. 2016). It is therefore possible to construct spectral probes of different timescales to perform this kind of PSD-based model comparison. For example, using $SFR_{H\alpha}/SFR_{UV}$ yields a probe of $\sim 5-10/\sim 70-100$ Myr timescales. While this is theoretically possible, it is important to take into account the degeneracies present due to observational factors like dust attenuation and metallicity, which will contribute to the statistical uncertainties of estimating the PSD for a given range of timescales.

5.2. *Routes into Chaos*

Galaxies are massively turbulent systems with large numbers of degrees of freedom. Writing down a set of coupled differential equations to model the star formation rate as a function of parameters of the system might allow us to understand the physics of how star formation is regulated in galaxies at different stellar masses, but it is hard to estimate the number of degrees of freedom we might need, and the exact forms through which these are coupled.

5.3. *Future work*

- Repeating the PSD analysis in log SFR space. While this constitutes an effective re-weighting of the SFH, it allows us to better examine the signatures of processes like quenching in the PSDs. Given that star formation rates are distributed roughly lognormally in observations, this might also be a more physically motivated basis to study the PSDs.
- Alternative time-series analysis tools: Wavelets and Structure Functions. Applications of these methods might give us information that is complementary to the Power Spectral Densities.
- Making SFHs quasi-stationary, and the effects on longest timescales. The Hurst analysis accurately quantifying the diversity in this regime.
- Further analyzing the IllustrisTNG small boxes - using cross power spectra to explicitly quantify the effects of changing physics
- Running the SAM with different strengths for feedback and winds. Independently confirming the effects of wind and AGN strengths on PSDs.
- Correlating spatial turbulence, disk stability and morphological factors to temporal PSD slopes.
- Obtaining constraints on physical phenomena by comparing the SFHs from different simulations to observationally reconstructed SFHs on longer ($1-10$ Gyr) timescales, and using burstiness indicators on short ($10-100$ Myr) timescales.

6. CONCLUSIONS

The distributions of galaxy star formation histories look very different across simulations. Analyzing the long-timescale variability of these SFHs alone is not enough to explain the diversity of SFHs within and across simulations. Using power spectral density analysis allows us to create a uniform framework to compare the variability of SFHs across different timescales. Looking at runs of the IllustrisTNG simulation with different physics models allows us to understand the imprints that different physical processes leave on the star formation histories of galaxies, and understand the differences between the PSDs of different simulations as a function

of the differences in their implementations of physical phenomena like stellar winds and AGN feedback.

ACKNOWLEDGEMENTS

This work was initiated as a project for the Kavli Summer Program in Astrophysics held at the Center for Computational Astrophysics of the Flatiron Institute in 2018. The program was co-funded by the Kavli Foundation and the Simons Foundation. We would like to thank Peter Behroozi for making the UniverseMachine SFHs publicly available. We thank them for their generous support. KI gratefully acknowledges support from Rutgers University.

REFERENCES

Anglés-Alcázar, D., Faucher-Giguère, C.-A., Quataert, E., et al. 2017, Monthly Notices of the Royal Astronomical Society: Letters, 472, L109

Behroozi, P., Wechsler, R., Hearin, A., & Conroy, C. 2018, arXiv preprint arXiv:1806.07893

Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, The Astrophysical Journal, 770, 57

Brennan, R., Pandya, V., Somerville, R. S., et al. 2016, Monthly Notices of the Royal Astronomical Society, stw2690

Caplar, N., Lilly, S. J., & Trakhtenbrot, B. 2017, The Astrophysical Journal, 834, 111

Carnall, A., McLure, R., Dunlop, J., & Davé, R. 2018, Monthly Notices of the Royal Astronomical Society, 480, 4379

Ceverino, D., Dekel, A., & Bournaud, F. 2010, Monthly Notices of the Royal Astronomical Society, 404, 2151

Ceverino, D., Dekel, A., Mandelker, N., et al. 2012, Monthly Notices of the Royal Astronomical Society, 420, 3490

Ceverino, D., & Klypin, A. 2009, The Astrophysical Journal, 695, 292

Ceverino, D., Klypin, A., Klimek, E. S., et al. 2014, Monthly Notices of the Royal Astronomical Society, 442, 1545

Choi, E., Ostriker, J. P., Naab, T., & Johansson, P. H. 2012, The Astrophysical Journal, 754, 125

Choi, E., Ostriker, J. P., Naab, T., et al. 2017, The Astrophysical Journal, 844, 31

Davé, R., Rafieferantsoa, M. H., Thompson, R. J., & Hopkins, P. F. 2016a, arXiv preprint arXiv:1610.01626

Davé, R., Thompson, R., & Hopkins, P. F. 2016b, Monthly Notices of the Royal Astronomical Society, 462, 3265

Dekel, A., & Silk, J. 1986, The Astrophysical Journal, 303, 39

Genel, S., Vogelsberger, M., Springel, V., et al. 2014, Monthly Notices of the Royal Astronomical Society, 445, 175

Genel, S., Bryan, G. L., Springel, V., et al. 2018, arXiv preprint arXiv:1807.07084

Heavens, A. F., Jimenez, R., & Lahav, O. 2000, Monthly Notices of the Royal Astronomical Society, 317, 965

Hopkins, P. F., Kereš, D., Oñorbe, J., et al. 2014, Monthly Notices of the Royal Astronomical Society, 445, 581

Hsieh, B., Lin, L., Lin, J., et al. 2017, The Astrophysical Journal Letters, 851, L24

Iyer, K., & Gawiser, E. 2017, The Astrophysical Journal, 838, 127

Kelson, D. D. 2014, arXiv preprint arXiv:1406.5191

Leja, J., Johnson, B. D., Conroy, C., van Dokkum, P. G., & Byler, N. 2016, arXiv preprint arXiv:1609.09073

Mandelbrot, B. B., & Van Ness, J. W. 1968, SIAM review, 10, 422

Moster, B. P., Naab, T., & White, S. D. 2018, Monthly Notices of the Royal Astronomical Society, 477, 1822

Pacifici, C., Oh, S., Oh, K., Lee, J., & Yi, S. K. 2016, arXiv preprint arXiv:1604.02460

Pillepich, A., Springel, V., Nelson, D., et al. 2017, Monthly Notices of the Royal Astronomical Society, 473, 4077

Schreiber, C., Pannella, M., Elbaz, D., et al. 2015, Astronomy & Astrophysics, 575, A74

Somerville, R. S., Hopkins, P. F., Cox, T. J., Robertson, B. E., & Hernquist, L. 2008, Monthly Notices of the Royal Astronomical Society, 391, 481

Somerville, R. S., Popping, G., & Trager, S. C. 2015, Monthly Notices of the Royal Astronomical Society, 453, 4337

Tacchella, S., Bose, S., Conroy, C., Eisenstein, D. J., & Johnson, B. D. 2018, arXiv preprint arXiv:1806.03299

Tojeiro, R., Heavens, A. F., Jimenez, R., & Panter, B. 2007, Monthly Notices of the Royal Astronomical Society, 381, 1252

Torrey, P., Vogelsberger, M., Hernquist, L., et al. 2017, arXiv preprint arXiv:1711.11039

Vogelsberger, M., Genel, S., Springel, V., et al. 2014a, Monthly Notices of the Royal Astronomical Society, 444, 1518

—. 2014b, Nature, 509, 177

Weinberger, R., Springel, V., Hernquist, L., et al. 2016, Monthly Notices of the Royal Astronomical Society, 465, 3291

Weisz, D. R., Dalcanton, J. J., Williams, B. F., et al. 2011, The Astrophysical Journal, 739, 5

Welch, P. 1967, IEEE Transactions on audio and electroacoustics, 15, 70

Yung, L., Somerville, R. S., Finkelstein, S. L., Popping, G., & Davé, R. 2018, arXiv preprint arXiv:1803.09761

Zolotov, A., Dekel, A., Mandelker, N., et al. 2015, Monthly Notices of the Royal Astronomical Society, 450, 2327

APPENDIX

A. VARIATIONS ON TESTS PROBING THE LOWEST TIMESCALES TO WHICH WE CAN PROBE THE PSD

We repeat the analysis described in Sec. 4.2 using log SFRs instead of linearly distributed SFRs. While this has some advantages such as being able to observe the change in the SFh due to quenching better, it is significantly worse for analysing short timescales at all masses. This happens due to two reasons: firstly, going to log SFR reduces decreases the magnitude of the fluctuations in the SFH, leading to less power at all timescales. Secondly, it decreases the S/N of the power spectrum, due to the increased contribution to the fluctuations from discrete star particles. The sensitivity curves for this analysis are shown in Figure 10, where we see that while it is possible to probe the long-timescale modes (at $\geq 1Gyr$) of the SFH across a wide range of stellar masses, probing the shorter timescales is limited to only the highest mass ($M_* > 10^{11} M_\odot$) galaxies.

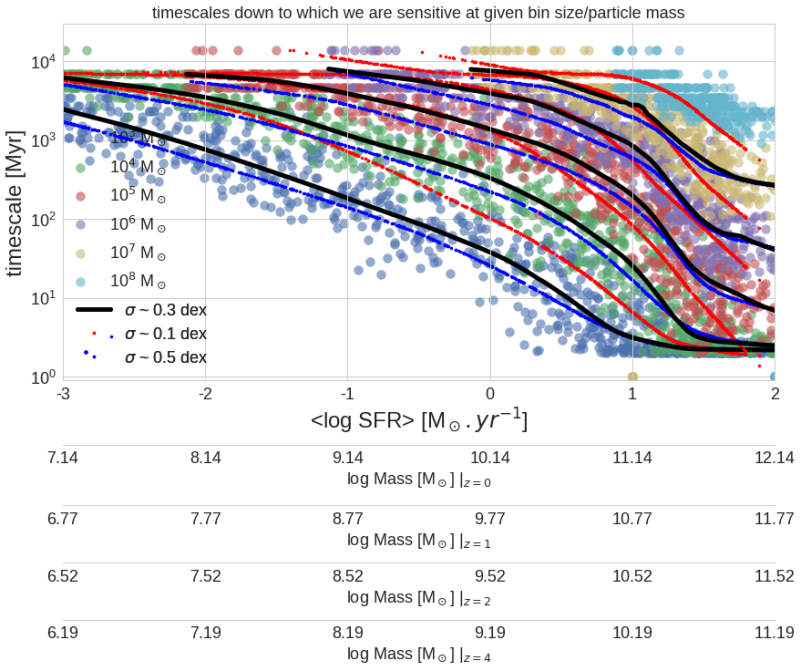

Figure 10. Quantifying the lowest timescales we can probe at different stellar masses for galaxies that follow the SFR-M_* correlation from Schreiber et al. (2015) with perturbations that add different amounts of scatter to the relation (between 0.1 to 0.5 dex) for different masses for the star particles following the procedure described in sec. 4.2. This figure shows the thresholds corresponding to an analysis in log SFR space.

B. STRUCTURE FUNCTIONS

As an alternate method of analysis, we also compute structure functions (Caplar et al. 2017), which quantify the variance in SFR as a function of their separation. The structure function at a timescale (τ) is defined as,

$$SF(\tau)^2 = \frac{1}{N(\tau)} \sum_{i<j} [SFR(t_i) - SFR(t_j)]^2 \qquad \tau = t_i - t_j \tag{B1}$$

The advantage of the structure function formalism is that it can be naturally extended to analyzing an ensemble of SFHs using the same criterion as that for analyzing a single SFH. As a consistency check, we find that the median SF^2 for the sample is not very different from the structure function of the entire ensemble of SFHs in any given mass/redshift bin.

Performing the ensemble analysis could let us probe lower timescales than currently possible with the PSDs. However, the SF makes an assumption about stationarity, which may not hold true for galaxies with a hurst index $H \neq 0.5$.

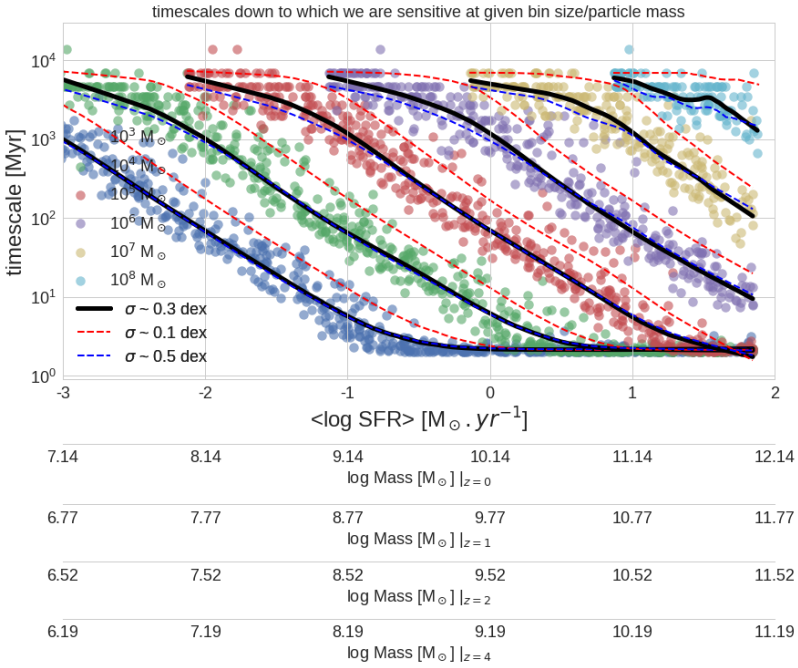

Figure 11. Quantifying the lowest timescales we can probe at different stellar masses for galaxies with a constant SFH + perturbations that add different amounts of scatter to the relation (between 0.1 to 0.5 dex) for different masses for the star particles following the procedure described in sec. 4.2. This figure shows the thresholds corresponding to an analysis in log SFR space.

To make sure this is the case, we can subtract the long timescale fluctuations from the SFH, leaving essentially the shorter timescale variability that we analyze using the structure function. This allows us to probe a complementary range of timescales to extend the current analysis with.