

The time-scales probed by star formation rate indicators for realistic, bursty star formation histories from the FIRE simulations

José A. Flores Velázquez,^{1,2} Alexander B. Gurvich^{1,2,★}, Claude-André Faucher-Giguère^{1,2}, James S. Bullock¹, Tjitske K. Starkenburg,² Jorge Moreno^{1,3}, Alexandres Lazar¹, Francisco J. Mercado¹, Jonathan Stern^{1,2}, Martin Sparre^{1,4,5}, Christopher C. Hayward^{1,6}, Andrew Wetzel^{1,7} and Kareem El-Badry^{1,8}

¹Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA

²Department of Physics and Astronomy and CIERA, Northwestern University, 1800 Sherman Ave, Evanston, IL 60201, USA

³Department of Physics and Astronomy, Pomona College, Claremont, CA 91711, USA

⁴Institut für Physik und Astronomie, Universität Potsdam, Karl-Liebknecht-Str 24/25, D-14476 Golm, Germany

⁵Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany

⁶Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA

⁷Department of Physics and Astronomy, University of California, Davis, CA 95616, USA

⁸Department of Astronomy and Theoretical Astrophysics Center, University of California Berkeley, Berkeley, CA 94720, USA

Accepted 2020 December 12. Received 2020 December 12; in original form 2020 August 19

ABSTRACT

Understanding the rate at which stars form is central to studies of galaxy formation. Observationally, the star formation rates (SFRs) of galaxies are measured using the luminosity in different frequency bands, often under the assumption of a time-steady SFR in the recent past. We use star formation histories (SFHs) extracted from cosmological simulations of star-forming galaxies from the FIRE project to analyse the time-scales to which the $H\alpha$ and far-ultraviolet (FUV) continuum SFR indicators are sensitive. In these simulations, the SFRs are highly time variable for all galaxies at high redshift, and continue to be bursty to $z = 0$ in dwarf galaxies. When FIRE SFHs are partitioned into their bursty and time-steady phases, the best-fitting FUV time-scale fluctuates from its ~ 10 Myr value when the SFR is time-steady to $\gtrsim 100$ Myr immediately following particularly extreme bursts of star formation during the bursty phase. On the other hand, the best-fitting averaging time-scale for $H\alpha$ is generally insensitive to the SFR variability in the FIRE simulations and remains ~ 5 Myr at all times. These time-scales are shorter than the 100 and 10 Myr time-scales sometimes assumed in the literature for FUV and $H\alpha$, respectively, because while the FUV emission persists for stellar populations older than 100 Myr, the time-dependent luminosities are strongly dominated by younger stars. Our results confirm that the ratio of SFRs inferred using $H\alpha$ versus FUV can be used to probe the burstiness of star formation in galaxies.

Key words: galaxies: high-redshift – galaxies: star formation – ultraviolet: galaxies.

1 INTRODUCTION

1.1 Measuring star formation rates observationally

In the regime where counting individual stellar objects and measuring their ages is impossible (e.g. the vast majority of extragalactic systems), star formation rates (SFRs) must be inferred from the observed integrated light from the current stellar population and surrounding ionized gas (e.g. Kennicutt & Evans 2012; Calzetti 2013). Two widely used indicators of star formation are the nebular $H\alpha$ recombination line (6563 Å) and the far-ultraviolet (FUV) continuum (~ 1350 – 1750 Å, for the GALEX bandpass). $H\alpha$ photons, primarily produced in the H II regions surrounding young massive stars, have long been used as an indicator of recent (~ 10 Myr) star formation due to the short lifetimes of the stars that produce them. On

the other hand, older (but still relatively young) stellar populations with masses above $\sim 5 M_{\odot}$ dominate the integrated spectrum in FUV and are thought to trace star formation on longer time-scales (up to ~ 100 Myr; Kennicutt & Evans 2012; Calzetti 2013). SFR indicators can thus be thought of as a convolution of the intrinsic ‘true’ star formation history (SFH) with a time-delayed response function. One can take advantage of this fact by combining observations of different SFR indicators to measure SFRs averaged over different times in a galaxy’s recent history (e.g. Weisz et al. 2012; Domínguez et al. 2015; Sparre et al. 2017).

There is a vast literature seeking to predict the properties of simple stellar populations (SSPs; notable examples include Leitherer et al. 1999; Bruzual & Charlot 2003; Conroy, Gunn & White 2009; da Silva, Fumagalli & Krumholz 2012; Eldridge et al. 2017). These SSPs model the evolution of populations of stars from a distribution of zero-age main-sequence (ZAMS) masses, known as an initial mass function, or IMF (see Kroupa et al. 1993 for one such example). SSPs make predictions for a number of important and useful properties

★ E-mail: agurvich@u.northwestern.edu

of the populations, including their spectral energy distributions (SEDs), supernova rates, distributions of stellar type, estimates of metal yields, distributions of surviving stellar masses, distributions of colours, and both ionizing photon rates and narrow-band FUV continuum emission as a function of age. The last two of these are the most relevant for the analysis presented here – because they are the basis for the SFR response functions mentioned above.

Once a luminosity response function is determined, the next step is to calibrate an SFR indicator by modelling a simple stellar population and the luminosity it emits. Thus, calibration requires making some assumption about the shape of the (recent) SFH. A common choice is a constant SFR, for which the calibration is the asymptotic luminosity of the population formed at that SFR (e.g. Buat et al. 2012; Kennicutt & Evans 2012). When suggested by observations (e.g. constrained by ratios of indicators that act on different time-scales) a ‘constant burst’ (i.e. a top hat of varying width) is sometimes used (e.g. Calzetti et al. 2007; Hao et al. 2011; Murphy et al. 2011). The appropriate calibration in this case (to recover the burst SFR), is the ratio of the luminosity of the stellar population formed during the burst at the inferred time delay of the observation (typically shortly following the burst) to the burst SFR.

A common definition for the time-scale that an indicator acts on is the time that it takes for the calibration population to reach ~ 90 per cent of its integrated luminosity. Following this definition, the FUV emission is thought to trace stars formed over the past ~ 100 Myr (Murphy et al. 2011), while $H\alpha$ emission trace stars formed in the past ~ 10 Myr (Hao et al. 2011). However, since even a stellar population formed in a single burst can emit significant UV light long after its FUV or $H\alpha$ emission peaks, the 90 per cent ‘persistence times’ can be substantially longer than the SFR time-scales to which FUV and $H\alpha$ light are sensitive. Moreover, the assumption of a constant SFH (or a single top-hat burst) is not necessarily accurate in real galaxies, which have can have complex and often ‘bursty’ SFHs. Thus, to properly interpret the SFRs implied by common observational indicators, we must understand the time-scales the indicators probe when applied to realistic SFHs, which may include highly time-variable periods.

1.2 The need to consider ‘bursty’ star formation

Deviations in the ‘indicated’ (or ‘measured’) SFR from the ‘true’ instantaneous (or ‘intrinsic’) SFR can be substantial when the SFR changes on time-scales shorter than the maximum lifetime of stars that contribute non-negligibly to the observed photons. So-called bursty star formation is a generic prediction of simulations of galaxy formation that include explicitly resolved star-forming regions and stellar feedback (Governato et al. 2012; Teyssier et al. 2013; Hopkins et al. 2014; Agertz & Kravtsov 2015; Sparre et al. 2017; Faucher-Giguère 2018). Recently, Iyer et al. (2020) characterized the SFR variability using a power spectral density (PSD) method in several different state-of-the-art galaxy formation models, including hydrodynamic simulations as well as semi-analytic and empirical models. They showed that, while different models predict different PSDs, the hydrodynamic simulations typically predict increasing short time-scale variability with decreasing galaxy mass. The SFR variability arises from a combination of effects, including the interplay between cosmological inflows, mergers, and recycling outflows on longer and intermediate time-scales (e.g. Anglés-Alcázar et al. 2017), and processes related to the lifecycle of molecular clouds in the interstellar medium (ISM) on shorter time-scales (e.g. Tacchella, Forbes & Caplar 2020).

There are also several observational indications that bursty SFRs are important in the real Universe (e.g. Weisz et al. (e.g. Weisz et al. 2012; Guo et al. 2016; Smit et al. 2016; Emami et al. 2019; Hirsten et al. 2019; Pelliccia et al. 2020)). Specifically, observations use the ratio of short and long time-scale tracers of SFR (e.g. the $H\alpha$ -to-UV ratio) as a signature of bursty star formation (e.g. Broussard et al. 2019). Since immediately after a burst of star formation the short-time-scale tracer decreases faster than the long-time-scale one, a burst leads to a ratio < 1 (unequal dust attenuation could also affect the $H\alpha$ -to-UV ratio and must also be taken into account, e.g. Koyama et al. 2019; Salim & Narayanan 2020). For example, Weisz et al. (2012) observed a decreasing $H\alpha$ -to-UV ratio with decreasing galaxy mass, suggesting that bursty star formation is more prevalent in dwarf galaxies (these results are echoed by Guo et al. 2016, who use $H\beta$ instead of $H\alpha$). There is also evidence of common SFR variability at high redshift, e.g. in the form of ‘extreme emission-line galaxies’, which have been interpreted as having recently experienced intense starbursts of duration ~ 15 Myr (van der Wel et al. 2011; Forrest et al. 2017). Understanding the prevalence and properties of bursty SFHs is of major importance for understand how galaxies form and evolve. The SFR–stellar mass (M_*) relation (also known as the ‘star formation main sequence’; Brinchmann et al. 2004; Elbaz et al. 2007; Noeske et al. 2007; Whitaker et al. 2012; Salim et al. 2014; Schreiber et al. 2015; Shivaei et al. 2015) suggests that galaxies form most of their stars in a steady process. However, as summarized above, simulations and observations both indicate that many galaxies undergo short starbursts, during which they lie well above this relation (see e.g. Willett et al. 2015; Sparre et al. 2017; Elbaz et al. 2018; Matthee & Schaye 2019). In detail, it is therefore possible that many galaxies frequently move above and below the mean SFR– M_* relation, rather than evolving smoothly along it (e.g. Orr et al. 2017). The scatter in the observed relation should then reflect, in part, the burstiness of star formation in galaxies (e.g. Caplar & Tacchella 2019). In particular, Sparre et al. (2017) compares the scatter in the observed SFR– M_* relation to simulations from the FIRE (Feedback in Realistic Environments) project¹ and find that, to first order, the time-variability of the SFHs in the simulations is consistent with the observed scatter in SFR– M_* (although some quantitative differences may remain to be explained).

Understanding the time-scales probed by SFR indicators is also relevant for observations outside the UV and for the interpretation of galaxies expected to deviate strongly from the main sequence, e.g. during or following major mergers. Using simulations of merging galaxies combined with detailed radiative transfer, Hayward et al. (2014) showed that the total infrared (IR) luminosity (when interpreted using standard assumptions)² overestimates the true SFR immediately after the merger-induced star formation burst. For major mergers in which a strong starburst is induced, the SFR inferred from the IR luminosity can overestimate the instantaneous SFR during the post-starburst phase by greater than two orders of magnitude.

1.3 This paper

In this paper, we ask the following questions: (1) What are the SFR time-scales probed by $H\alpha$ and FUV indicators when taking into account realistic SFHs; and (2) how do the results depend on SFR

¹FIRE project website: <http://fire.northwestern.edu>.

²Since IR light is produced by reprocessing of UV light by dust, it is expected to probe time-scales similar to FUV.

burstiness? To answer these questions, we use SFHs from the FIRE-2 simulations of galaxy formation. We also use the detailed data sets provided by the simulations to assess how the indicated SFR time-scales depend on redshift and galaxy mass. Since our analysis models the application of standard SFR indicators on the simulated SFHs, the results will inform the interpretation of both past and future observational SFR measurements.

The FIRE-2 simulations are the second-generation cosmological, zoom-in hydrodynamic simulations from the FIRE project (Hopkins et al. 2014, 2018). These high-resolution simulations include a detailed model for star formation and stellar feedback and produce galaxies that match a number of observed properties, which motivates their use as models for ‘realistic’ SFHs. The properties include the M_* –halomass (M_h) relation, mass–metallicity relations (Ma et al. 2016a), the Kennicutt–Schmidt (KS) relation (Hopkins et al. 2014; Orr et al. 2018), and constraints on galactic winds from observations of the circumgalactic medium (e.g. Faucher-Giguère et al. 2015, 2016; Hafen et al. 2017).

Notably for our purposes, the FIRE simulations predict two main phases of star formation: bursty and time steady (e.g. Muratov et al. 2015; Faucher-Giguère 2018). At low redshift, \sim Milky Way-mass galaxies form long-lived stable discs with time-steady SFR (e.g. Gurvich et al. 2020). At high redshift, the progenitors of these galaxies have much more irregular morphologies and time-variable star formation histories (see e.g. Ma et al. 2017; Sparre et al. 2017; Stern et al. 2020). Fig. 1 illustrates the different morphologies predicted in the bursty and steady phases through the lens of the SFR indicators (FUV and $H\alpha$ light, uncorrected for dust obscuration). The left-hand panels ($z = 0.9$) feature a clumpy and amorphous distribution of star formation while the right-hand panels ($z \sim 0$) show coherent spiral patterns in the same galaxy but at a later time. In the FIRE simulations, star formation typically remains bursty until $z = 0$ in dwarf galaxies.

Although the present study is an important step in using detailed simulations to better understand the time-scales probed by different SFR indicators, we note a few limitations that should be borne in mind when using our results to interpret observations. First, as demonstrated quantitatively in the PSD comparison study by Iyer et al. (2020), different simulations predict different SFR variability statistics. Thus, in the bursty regimes, our results will apply specifically to the galaxy formation physics in the FIRE-2 simulations (including the stellar feedback models described in more detail below). Secondly, other indicators of SFR variability than the $H\alpha$ -to-FUV ratio that we model in this work can and have been used to empirically constrain SFR burstiness, including the 4000 Å break and $H\delta_A$ indices (e.g. Kauffmann 2014). It would be interesting to use simulations to also model these other indicators in the future. Thirdly, we must stress an important caveat regarding effects of dust, discussed in more detail in Section 4.3.1. Dust will in general attenuate FUV and $H\alpha$ light differently, as well as in ways that correlate with SFR variations. Thus, variations in observed FUV and $H\alpha$ light curves can follow substantially different patterns than the intrinsic star formation history. In this work, we do not model dust attenuation but instead assume that observational techniques are able accurately correct for dust and infer the non-attenuated FUV and $H\alpha$ fluxes.

The remainder of the paper is organized as follows. In Section 2, we describe the simulations in more detail, as well as how we define and analyse SFR indicators. In Section 3, we infer the best-fitting averaging time-scales as a function of cosmic time in galaxies of different masses. We discuss our results in Section 4 and in Section 5 we summarize and conclude.

2 METHODS

2.1 Star formation histories from FIRE-2 simulations

The FIRE-2 simulations are cosmological zoom-in simulations run in the meshless finite mass (MFM) mode of the GIZMO gravity + magnetohydrodynamic code.³ MFM is a Lagrangian, mesh-free, finite-mass method combining the advantages of traditional smooth particle hydrodynamics and grid-based methods (for numerical details and tests, see Hopkins 2015). The FIRE-2 physics model, which is described in full detail in Hopkins et al. (2018), includes radiative cooling for gas down to 10 K (including approximate treatments of fine-structure metal and molecular lines). Star particles representing simple stellar populations are formed in gas that is self-gravitating, dense ($n_H \geq 1000 \text{ cm}^{-3}$), and molecular (Hopkins, Narayanan & Murray 2013). Star particles return mass, metals, momentum, and energy in the ISM, at rates that are functions of the star particle’s age following the STARBURST99 population synthesis model (Leitherer et al. 1999). Stellar feedback is modelled on the scale of star-forming regions and includes supernovae (Types II and Ia), stellar winds from O, B, and AGB stars, photoelectric heating, photoionization, and radiation pressure.

We study analyse a range of galaxy masses, from dwarf galaxies up to $\sim L^*$. This allows us to study the different types of SFHs experienced by star-forming galaxies. We analyse eight simulations in total: four dwarfs with $z = 0$ halo mass in the range $M_h \sim 10^9 - 10^{11} M_\odot$ and four galaxies with halo mass $M_h \sim 10^{12}$, similar to the Milky Way. Table 1 catalogues some basic physical properties of these galaxies at redshift $z = 0$. The mass of baryonic resolution elements ranges from $m_b \approx 250 M_\odot$ (for dwarfs) to $\approx 7100 M_\odot$ (for the m12 galaxies). More details on resolution and other facets of the FIRE-2 simulations are given in Hopkins et al. (2018) and Garrison-Kimmel et al. (2017, 2019). All simulations include a sub-grid model for turbulent metal diffusion between nearby gas cells.

We use an ‘archaeological’ approach to reconstruct the SFHs of the galaxies in the FIRE simulations at high time resolution. Specifically, we first identify all of the stars in the galaxy’s main halo at redshift $z = 0$. To do so, we use the AMIGA halo finder (Gill, Knebe & Gibson 2004; Knollmann & Knebe 2009) to locate the main halo in our simulation volume and measure the virial radius (R_{vir}) and the halo centre of mass (measured using all of the dark, gas, and stellar components of the galaxy). Dark matter haloes are defined using the Bryan & Norman (1998) redshift-dependent overdensity criterion. We extract the star particles within a spherical volume of radius $5R_{*,\text{half}}$, where $R_{*,\text{half}}$ is the radius containing 50 per cent of the stars inside $0.15R_{\text{vir}}$. The time-dependent SFH is then reconstructed from the distribution of star particle formation times (SFTs) and star masses, self-consistently accounting for stellar mass-loss. This archaeological approach neglects the fact that some stars formed in galaxies other than the main progenitor (and then merged), but this is a good approximation for galaxies up to $\sim L^*$, which form most of their stars *in situ* (e.g. Anglés-Alcázar et al. 2017; Fitts et al. 2018).

From this high time-resolution SFH, we can construct SFHs with the SFR boxcar-averaged on an arbitrary time-scale t_{avg} , which we denote $\langle \text{SFR}_{\text{true}} \rangle_{t_{\text{avg}}}(t)$. Since, in observational applications, only the SFR before the time of observation is well defined, we define

³Information about GIZMO and a public version of the code are available at: <https://www.tapir.caltech.edu/~phopkins/Site/GIZMO>.

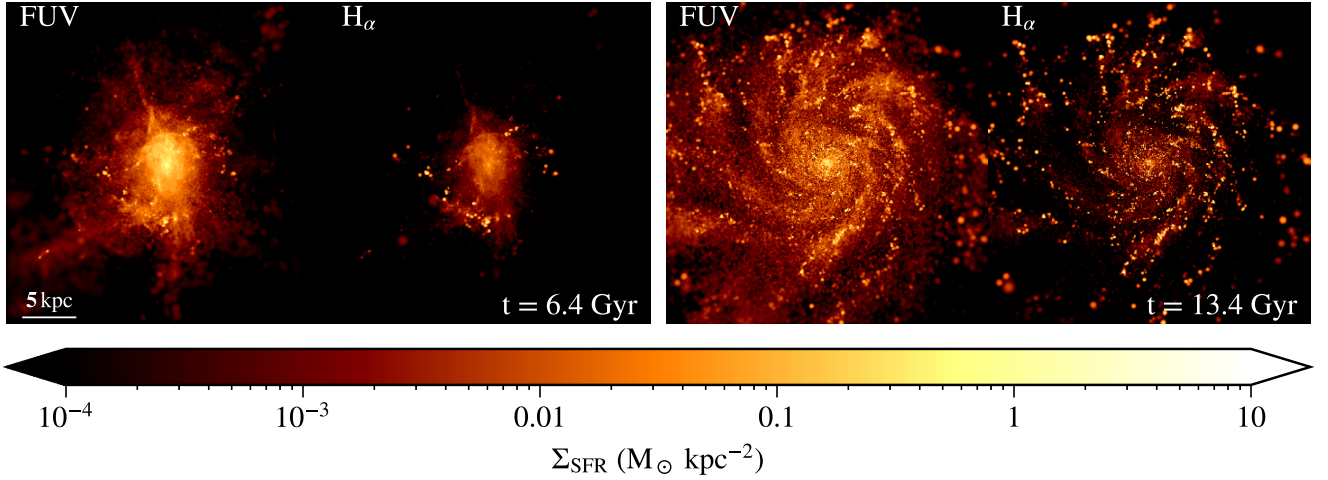


Figure 1. SFR surface density maps as traced by unattenuated FUV and $H\alpha$ emission for a simulated Milky Way-mass galaxy (m12i) at cosmic time $t = 6.4$ Gyr ($z = 0.9$, ≈ 50 Myr after a burst of star formation; left) and at $t = 13.4$ Gyr ($z = 0.03$, long after the SFR has settled to a steady state; right). The high-redshift panels on the left showcase the irregular morphology of the galaxy during the period of bursty star formation. The larger integrated FUV-indicated SFR also shows how the FUV-indicated SFR following a burst of star formation overestimates the more recent SFR indicated by $H\alpha$. In contrast, in the low-redshift panels on the right, which exhibit a well-ordered disc (viewed face-on) after the galaxy’s transition to time-steady star formation, have more similar total FUV- and $H\alpha$ -indicated SFRs. In these unattenuated maps, the $H\alpha$ light is more concentrated in young, bright star-forming regions, while there is more diffuse FUV light between spiral arms.

Table 1. Properties at $z = 0$ of the simulated galaxies analysed in this paper.

| Name | $M_h (M_\odot)^a$ | $M_* (M_\odot)^b$ | $R_{*,\text{half}} (\text{kpc})^c$ | f_g^d |
|------|----------------------|----------------------|------------------------------------|---------|
| m10q | 7.7×10^9 | 2.3×10^6 | 0.8 | 0.68 |
| m11i | 6.8×10^{10} | 1.0×10^9 | 3.7 | 0.61 |
| m11q | 1.4×10^{11} | 6.5×10^8 | 2.5 | 0.67 |
| m11d | 2.7×10^{11} | 4.5×10^9 | 6.9 | 0.57 |
| m12i | 9.4×10^{11} | 6.8×10^{10} | 2.9 | 0.20 |
| m12b | 1.1×10^{12} | 9.0×10^{10} | 2.8 | 0.16 |
| m12c | 1.1×10^{12} | 6.4×10^{10} | 3.4 | 0.22 |
| m12f | 1.3×10^{12} | 8.8×10^{10} | 4.0 | 0.24 |

Notes. ^aTotal dark matter and baryonic mass within R_{vir} , the virial radius defined as in Bryan & Norman (1998).

^bTotal stellar mass within $5R_{*,\text{half}}$.

^c3D radius containing half of the total stellar mass within 15 per cent of R_{vir} .

^dGas fraction $f_g = M_g/(M_* + M_g)$, where M_g is the gas mass within the galaxy.

these boxcar averages by averaging the *preceding* t_{avg} time interval preceding a given cosmic time t . Fig. 2 shows the SFH for one representative simulation in each mass bin.

For each galaxy, we identify the ‘end’ of bursty star formation by computing the root-mean-square relative error (RMSRE) in log-space using a 500 Myr-wide moving boxcar:

$$\text{RMSRE}_{\text{SFH}} = \sqrt{\frac{\langle (\log_{10}(\text{SFR})^2)_{500 \text{ Myr}} - (\log_{10}(\text{SFR}))_{500 \text{ Myr}}^2 \rangle}{\langle (\log_{10}(\text{SFR}))_{500 \text{ Myr}}^2 \rangle}}. \quad (1)$$

To account for periods when the SFR is identically zero, we add one-tenth of the minimum true SFR to the entire SFH (this choice does not affect any of our results). We consider the SFH to become ‘steady’ when $\text{RMSRE}_{\text{SFH}} \leq 0.3$. The bottom panel of Fig. 2 shows an example of this metric applied to the m12i simulation and how it accurately identifies the clear change in SFR behaviour in this case. Dwarf galaxies in FIRE are typically bursty all the way to $z = 0$, so no transition is identified for these galaxies.

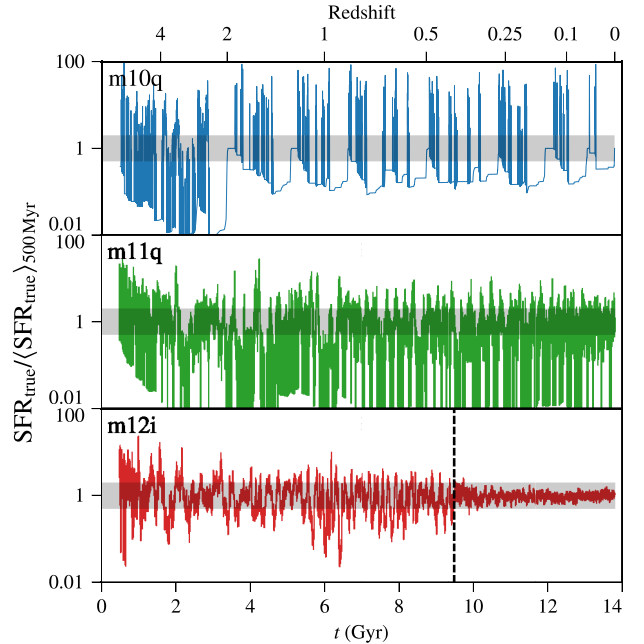


Figure 2. Star formation histories of three prototypical galaxies from our simulation sample, one from each halo mass bin considered (m10q, halo mass $M_h \sim 10^{10} M_\odot$ at $z = 0$; m11q, $M_h \sim 10^{11} M_\odot$; m12f, $M_h \sim 10^{12} M_\odot$). In each panel, the SFR is normalized by the 500 Myr running boxcar-averaged SFR. Dwarf galaxies (top two panels) have bursty SFHs at all redshifts. Milky Way-mass galaxies with $M_h(z = 0) \sim 10^{12} M_\odot$ (bottom panel) undergo a transition from a bursty SFR at high redshift to a time-steady SFR at low redshift (with some remaining, but much smaller fluctuations). The transition from bursty to steady SFR is indicated by the vertical black dashed line in the bottom panel. Shaded regions show a factor of 0.3 dex in either direction of unity.

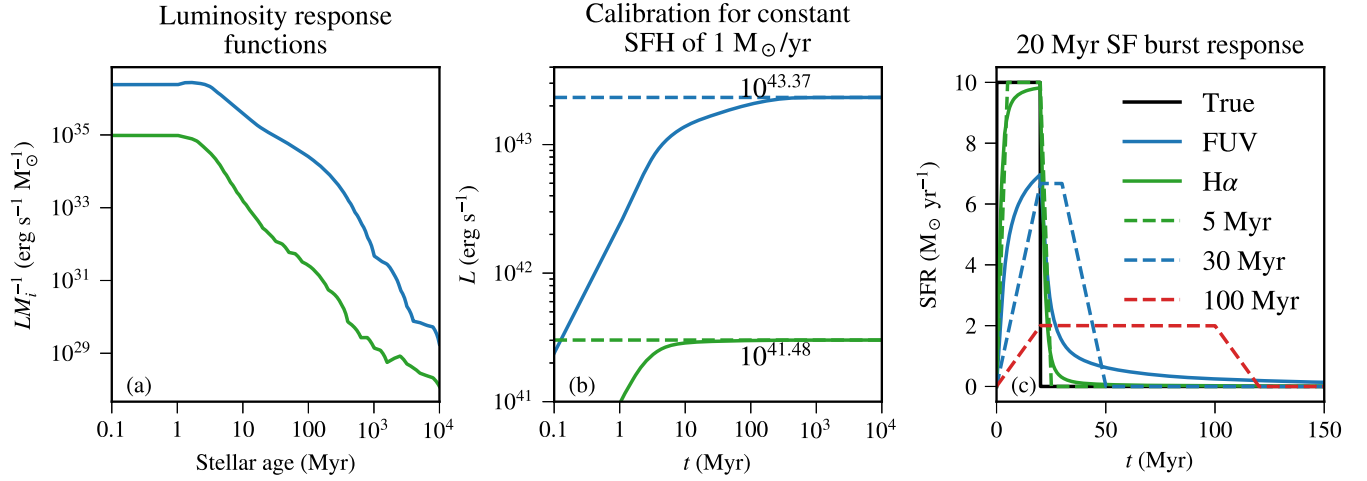


Figure 3. (a) Luminosity (per stellar mass formed) following a star formation burst in the FUV continuum and $H\alpha$ as a function of stellar age, based on a BPASS stellar population synthesis model. (b) Evolution of L_{FUV} and $L_{H\alpha}$ as a function of time assuming an SFH with a constant SFR of $1 \text{ M}_{\odot} \text{ yr}^{-1}$. SFR indicator calibration constants are defined as the asymptotic luminosity in each case, represented by dashed lines and annotated on each curve. (c) Indicated SFR response to a population of stars formed in a 20 Myr burst (solid black line). Coloured dashed lines show moving boxcar averages of the true SFR using different widths: 5 Myr (green), 30 Myr (blue), and 100 Myr (red).

2.2 Star formation rate indicators

We model stellar spectra as a function of age using version 2.2.1 of the Binary Population and Spectral Synthesis (BPASS) stellar population synthesis code (Eldridge et al. 2017). BPASS models the effects of binary evolution on stellar spectra and uses CLOUDY (Ferland et al. 1998) to compute nebular emissions, which is the main contributor to $H\alpha$ emission.⁴ We use a fixed BPASS model, which consists of a Kroupa IMF (Kroupa, Tout & Gilmore 1993) with a metallic mass fraction of $Z = 0.014$. To derive observationally indicated SFR values, SFR_{FUV} and $\text{SFR}_{H\alpha}$, we average the specific BPASS FUV flux over the wavelength interval 1556–1576 Å (near the centre of the GALEX FUV channel) and use the $H\alpha$ flux modelled by BPASS. This results in stellar age-dependent ‘response functions’, shown in Fig. 3(a).

In order to convert the modelled (‘observed’) luminosity into an indicated SFR, we assume a constant light-to-SFR ratio, i.e.

$$\text{SFR}_{\text{ind}} = C^{-1} L_{\text{ind}}, \quad (2)$$

where C is the SFR ‘calibration constant’. This constant can be determined by considering the asymptotic luminosity, $L_{\text{ind}}(t \rightarrow \infty)$, of a population of stars forming at a constant rate of SFR_{ind} . These calibration constants neglect obscuration by dust, which should be accounted for when attempting to measure the SFR of real galaxies. In principle, one can relax the assumption of time-constant SFR. In Section 3.2, we quantify the expected errors in indicated SFRs (relative to true SFRs), assuming that real SFHs are well modelled by the FIRE-2 simulations we analyse.

Fig. 3(b) shows L_{ind} as a function of time for $\text{SFR}_{\text{true}}(t) = 1 \text{ M}_{\odot} \text{ yr}^{-1}$. For the FUV continuum, $L_{\text{ind}} = \nu \langle L_{\nu} \rangle$ is evaluated at the middle of the FUV wavelength sample (1566 Å). Using equation (2), we find calibration constants of $C_{\text{FUV}} = 10^{43.33} \text{ erg s}^{-1} / \text{M}_{\odot} \text{ yr}^{-1}$ and

$C_{H\alpha} = 10^{41.42} \text{ erg s}^{-1} / \text{M}_{\odot} \text{ yr}^{-1}$. We note that each indicator reaches $L_{\text{ind}} \simeq 0.9 L_{\text{ind}}(t \rightarrow \infty)$ at $t \approx 100 \text{ Myr}$ and $t \approx 10 \text{ Myr}$, respectively – the often quoted time-scales for each of these indicators.

Fig. 3(c) shows the SFR response of these indicators applied to a population of stars formed by a 20 Myr burst of star formation at $10 \text{ M}_{\odot} \text{ yr}^{-1}$, beginning at $t = 0$. We compare the indicated SFRs to three moving ‘boxcar’ averages of width 5, 30, and 100 Myr. One important takeaway is that the 100 Myr running average is a poor match to the FUV indicated SFR at all times, which hints that the ‘effective’ FUV response time-scale is much shorter. On the other hand, the 5 Myr average agrees well with the $H\alpha$ indicated SFR.

2.3 Inferring the time-scales of SFR indicators

Next, we compare the true SFH to the ‘indicated’ SFHs using the calibrated indicators from Section 2.2. To measure the indicated $\text{SFR}_{\text{ind}}(t)$ (i.e. the SFR that would be inferred using a standard observational indicator, as defined in the previous section), we consider the total luminosity from the stars in a galaxy with $\text{SFT} < t$, i.e.

$$\text{SFR}_{\text{ind}}(t) = C_{\text{ind}}^{-1} \sum_{i, \text{SFT}_i < t} L_{\text{ind}}^i(t - \text{SFT}_i), \quad (3)$$

where $L_{\text{ind}}^i(t - \text{SFT}_i)$ corresponds to the response functions in Fig. 3(a), evaluated for star particle i with a stellar age $(t - \text{SFT}_i)$.

Our methodology for comparing the true and indicated SFRs is illustrated in Fig. 4. At any given time there is, in general, a non-zero deviation between the true and indicated SFRs. The goal of our analysis is to identify the *best-fitting boxcar averaging time-scale*, $t_{\text{avg}}^{\text{min}}$, that minimizes this difference between $\langle \text{SFR}_{\text{true}} \rangle_{t_{\text{avg}}}$ and a given SFR_{ind} . Physically, $t_{\text{avg}}^{\text{min}}$ is the width of the boxcar over which one would have to average the true SFR to match the indicated SFR at the time of an observation.

When the SFH fluctuates rapidly, many t_{avg} may produce the same $\langle \text{SFR}_{\text{true}} \rangle_{t_{\text{avg}}}$ that match the indicated SFR. To avoid this issue, we determine best-fitting $t_{\text{avg}}^{\text{min}}$ for *ensembles of observation times* in windows of size $T_w = 500 \text{ Myr}$ (longer than the typical short time-

⁴The STARBURST99 stellar population synthesis model used for the FIRE feedback model does not include binaries. As in previous work with BPASS (e.g. Ma et al. 2016b), we neglect possible dynamical differences (during the hydrodynamic simulations) between the different stellar evolution models but include the spectral differences that are most directly relevant to the present study.

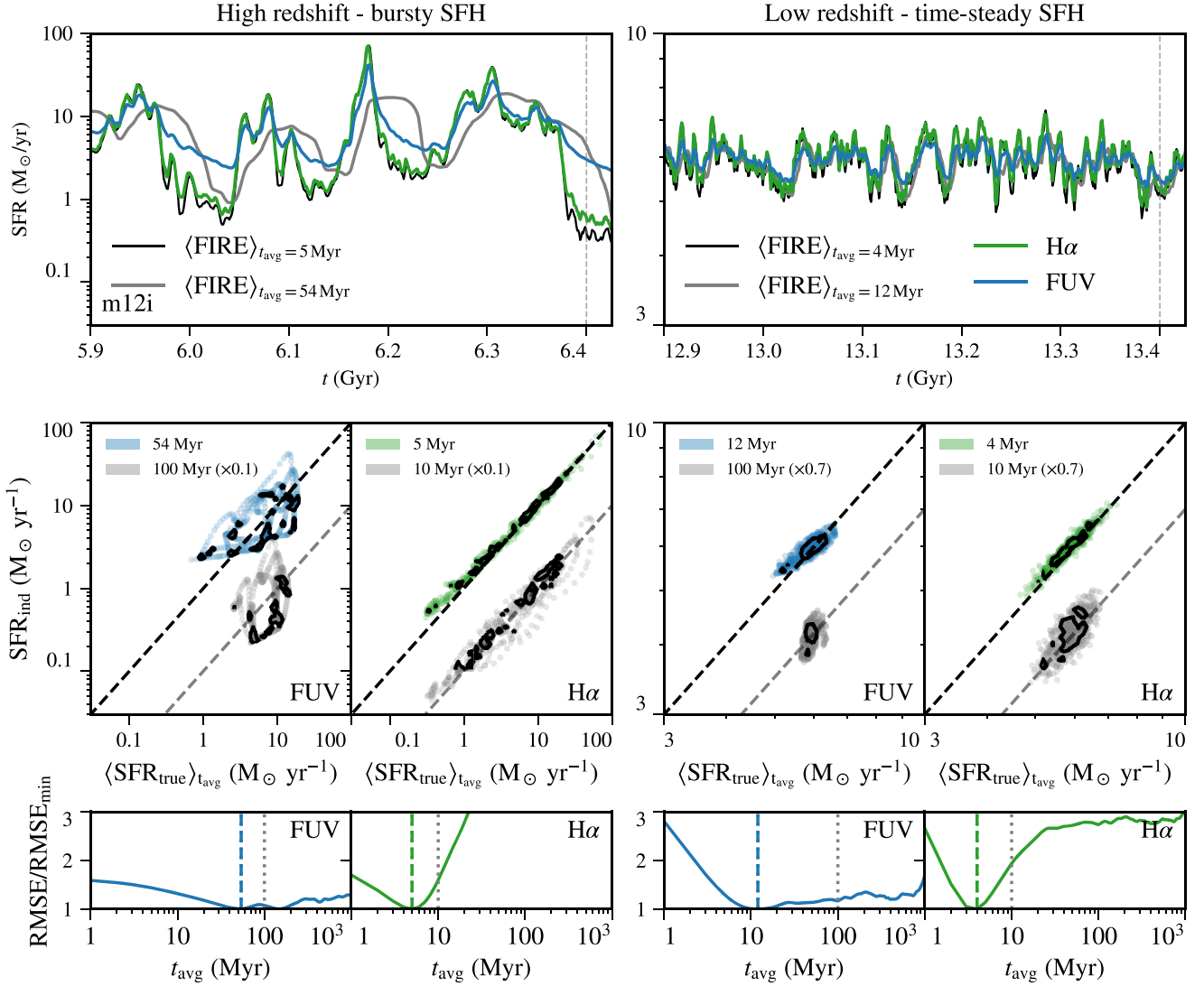


Figure 4. *Top:* ‘True’ and ‘indicated’ SFHs in a 500-Myr window at high (left) and low (right) redshift in m12i. True SFHs are plotted using best-fitting, boxcar-average equivalent time-scales for FUV (grey) and H α (black). Grey vertical lines are plotted to identify the times corresponding to the renderings in Fig. 1. At early times, the SFR frequently changes by a factor of >10 on time-scales ~ 30 – 50 Myr while at late times it varies only by $\lesssim 10$ per cent. *Middle:* Scatter plot of the indicated star formation rate, SFR_{ind} , versus the true star formation rate, boxcar-averaged using the best-fitting averaging time-scale (coloured). Contours containing 50 per cent of the points are plotted in black alongside a dashed line showing along the one-to-one relation. For comparison, scatter plots for SFR boxcar-averaged over time-scales of 100 and 10 Myr are also shown in grey in the FUV and H α panels, respectively (offset vertically by the factors noted on the figure for visual clarity). The smaller scatter in the coloured points versus the grey is visible by eye in all four panels. *Bottom:* RMSE between the indicated and time-averaged true SFHs as a function of boxcar averaging time-scale t_{avg} , normalized by the minimum RMSE. Vertical dashed lines (coloured) are plotted at the locations of the minima, while the dotted (grey) vertical lines indicate 100 and 10 Myr time-scales for reference. At both early and late times, H α has a well-defined minimum RMSE in almost all time windows. For FUV, on the other hand, the RMSE curve is typically much shallower, especially in the bursty high-redshift regime, implying that the equivalent boxcar average time-scale is not as well defined. Note that the value of $t_{\text{avg}}^{\text{min}}$ inferred in the high-redshift window shown here is not necessarily representative since the best-fitting time-scale fluctuates strongly following particularly extreme bursts of star formation like the one shown here (see Fig. 5).

scale variations during bursty phases). The increased accuracy in $t_{\text{avg}}^{\text{min}}$ resulting from introducing the time window comes at the cost of decreased ability to localize $t_{\text{avg}}^{\text{min}}$ in time within the window. We assign best-fitting values of $t_{\text{avg}}^{\text{min}}$ to the end of time windows (rather than the middle, for example). This ensures that $t_{\text{avg}}^{\text{min}}$ is computed using only SFRs that are causally connected to the observation time t . We quantify how our results depend the exact choice of T_w in Appendix A and find that they are stable for our fiducial choice of $T_w = 500$ Myr.

To find the best-fitting $t_{\text{avg}}^{\text{min}}$, we define another root-mean-square error (RMSE) that is a function of t_{avg} :

$$\text{RMSE}(t_{\text{avg}}) = \sqrt{\frac{\sum_i (\log_{10} \text{SFR}_{\text{ind}}^i - \log_{10} \langle \text{SFR}_{\text{true}} \rangle_{t_{\text{avg}}}^i)^2}{n}}, \quad (4)$$

where i is a sum over time samples spaced by 1 Myr covering the time window of total width T_w . This corresponds to the scatter with respect to the one-to-one line in log-space in the middle row of Fig. 4. Within a given time window, we repeat this process and

recompute SFHs of $\langle \text{SFR}_{\text{true}} \rangle_{t_{\text{avg}}}$ varying $1 \leq t_{\text{avg}} \leq 1000$ Myr in steps of 1 Myr (see the bottom row of Fig. 4). Lastly, we define the global minimum of this curve as $t_{\text{avg}}^{\text{min}}$. We repeat this procedure in $\sim 13\,000$ maximally overlapping windows (i.e. overlapping by 499 Myr) in order to construct histories of $t_{\text{avg}}^{\text{min}}$ over cosmic time. We omit the first 1 Gyr of the simulation to exclude transient effects related to the initial formation of the dark matter halo and galaxy.

3 RESULTS

3.1 Best-fitting time-scales

Fig. 4 summarizes our key results for two $T_w = 500$ Myr windows in m12i, a representative simulation of a Milky Way-mass halo. The left column focuses on high-redshift period, when the star formation is bursty, while the right column focuses on a low-redshift period when the SFR is relatively time-steady. The top row shows the SFHs during these periods. At early times (left-hand panel) the SFR changes by a factor of >10 on time-scales of ~ 30 –50 Myr whereas at late times (right-hand panel) the SFR changes only by $\lesssim 10$ per cent.

The middle-row panels show scatter plots of the true SFRs averaged over the best-fitting averaging time-scale, $t_{\text{avg}}^{\text{min}}$, versus the indicated SFRs for FUV and H α for the times corresponding to the right edge of the top-row panels. We also plot indicated SFHs for the often-quoted time-scales associated with each indicator (100 and 10 Myr), rather than $t_{\text{avg}}^{\text{min}}$, for comparison in grey. The points for the 100 and 10 Myr time-scales are offset vertically for clarity (by factors indicated on the panels). Contours containing 50 per cent of points in each case are shown in solid black. We see that the scatter for $t_{\text{avg}}^{\text{min}}$ is substantially smaller than the scatter for the reference 100 and 10 Myr time-scales.

The bottom row of Fig. 4 shows the RMSE as a function of t_{avg} , defined in equation (4), normalized by the RMSE at $t_{\text{avg}}^{\text{min}}$ (the RMSE minimum). We see little difference between the $t_{\text{avg}}^{\text{min}}$ for H α when comparing the high- and low-redshift results (5 Myr versus 4 Myr). This is because H α probes time-scales over which the SFR does not change drastically even in the bursty regime. Thus, H α tracks the true, instantaneous SFR fairly accurately in both the bursty and steady regimes.

FUV, on the other hand, shows more dependence on the burstiness of SFR. For the example shown in the right column of Fig. 4, when SFR is time-steady, we find a best-fitting $t_{\text{avg}}^{\text{min}} \approx 12$ Myr. This is in contrast to the much longer best-fitting $t_{\text{avg}}^{\text{min}} \approx 54$ Myr for the bursty period on the left experienced earlier by the same galaxy. As we discuss in more detail below, the best-fitting $t_{\text{avg}}^{\text{min}}$ can fluctuate strongly immediately following particularly extreme bursts of star formation, so the best fit for this example is not necessarily representative.

A concrete example of how FUV responds to a strongly changing SFR is shown in the top left panel of 4 (see the vertical dashed line at $t = 6.4$ Gyr). The left-hand panels of Fig. 1 show renderings of the FUV and H α light from the galaxy at this time. In this case, due to the strong drop in SFR shortly before the $t = 6.4$ Gyr ‘observation time’, the observed FUV light was emitted from stars that formed ~ 50 –100 Myr in the past. As a result, the FUV-indicated SFR is much larger than true instantaneous SFR, so that this indicated SFR is more representative of a time average extending to when the stars dominating the observed emission formed.

Next, we explore how $t_{\text{avg}}^{\text{min}}$ varies as a function of cosmic time and galaxy mass. Fig. 5 shows the evolution of $t_{\text{avg}}^{\text{min}}$ for FUV and H α over cosmic time for m10q, m11q, and all four m12 simulations listed in Table 1. For clarity, we show only results for a subset of the

$\sim 13\,000$ overlapping windows described in Section 2.3. Specifically, in each panel we show results for 56 equally spaced windows, each overlapping by 250 Myr with its neighbours.

We first focus on the two rightmost columns, corresponding to Milky Way-mass galaxies, which have dark matter haloes of mass $M_h \sim 10^{12} M_\odot$ at $z = 0$ (the m12 simulation plotted in red in the middle column is the same as in Fig. 4). These galaxies transition from bursty at high redshift to steady at later times. Vertical dashed lines indicate the times when bursty star formation ends for each simulation, using the definition in Section 2.1. We see that, when the SFR is time-steady (to the right of the dashed lines), the best-fitting averaging time-scales for H α and FUV are relatively stable (especially for H α) at 4–5 and 10–20 Myr, respectively. These ranges are indicated by the shaded regions in each row. As we found for the example in Fig. 4, the best-fitting $t_{\text{avg}}^{\text{min}}$ for H α is, in most cases, nearly the same even at high redshift when SFRs are highly variable. However, the best-fitting $t_{\text{avg}}^{\text{min}}$ for FUV fluctuates strongly in time, by factors up to >10 (i.e. up to values $t_{\text{avg}}^{\text{min}} > 100$ Myr). Thus, during bursty periods there is a large uncertainty in the actual time-scales probed by FUV-indicated SFRs.

The left column of Fig. 5 shows the results for dwarf galaxies in haloes of $z = 0$ mass $M_h \sim 10^{10}$ and $\sim 10^{11} M_\odot$ (m10q and m11q). The SFHs for these dwarf galaxies are bursty at all times. The more extreme burstiness for these dwarf galaxies (see the SFHs in Fig. 2) leads to more frequent excursions to very long values of $t_{\text{avg}}^{\text{min}}$. For m10q, the figure shows that values of $t_{\text{avg}}^{\text{min}} \gg 100$ Myr (up to ~ 1 Gyr) are found at some early times ($z \geq 2$). While it is beyond the scope of this paper to directly compare the simulated SFHs presented here to those observed in the real Universe, we note that the bursty SFHs in these dwarf galaxies may be consistent with the low quiescent fraction of observed dwarf galaxies at low redshift (e.g. Geha et al. 2012; Dickey et al. 2020).

The above results are illustrative of the main results of our analysis: (1) when the SFH is time-steady, $t_{\text{avg}}^{\text{min}}$ is relatively short for both H α and FUV (~ 5 and ~ 10 Myr, respectively); and (2) when the SFH is bursty, $t_{\text{avg}}^{\text{min}}$ can be much longer, especially for FUV due to its greater sensitivity to longer time-scales. As noted in Section 1.3, however, different simulations (with different models for star formation and stellar feedback) predict different SFR variability distributions, so the detailed quantitative results in this section (e.g. the range of $t_{\text{avg}}^{\text{min}}$ values in bursty periods) are expected to differ for different simulations. In steady periods, though, the best-fitting time-scales do not depend on the detail of the simulations.

3.2 Theoretical uncertainty due to unknown SFH

In the previous section, we quantified the best-fitting time-scales for FUV and H α indicators in a statistical sense. Even when the relevant time-scale is known, an important question is: how well does the indicated SFR match the true SFR boxcar-averaged over $t_{\text{avg}}^{\text{min}}$ when a measurement is made? In general, $(\text{SFR}_{\text{true}})_{t_{\text{avg}}^{\text{min}}} \neq \text{SFR}_{\text{ind}}$ because, depending on the SFH history, there may be no value of t_{avg} for which the indicated SFR exactly matches the boxcar average over the preceding t_{avg} . This introduces a minimum uncertainty (absent other information) in any measurement due to the unknown shape of the recent SFH, on top of the modelling uncertainties associated with the assumed stellar population synthesis model (including the IMF) and dust obscuration.

Fig. 6 quantifies this uncertainty by plotting distributions of the indicated-to-time-averaged-true-SFR ratio, $R_{\text{ind}}^{\text{ind}} \equiv \text{SFR}_{\text{ind}} / (\text{SFR}_{\text{true}})_{t_{\text{avg}}^{\text{min}}}$, for different averaging time-scales t_{avg} . To produce these distributions, we first partition the simulations listed

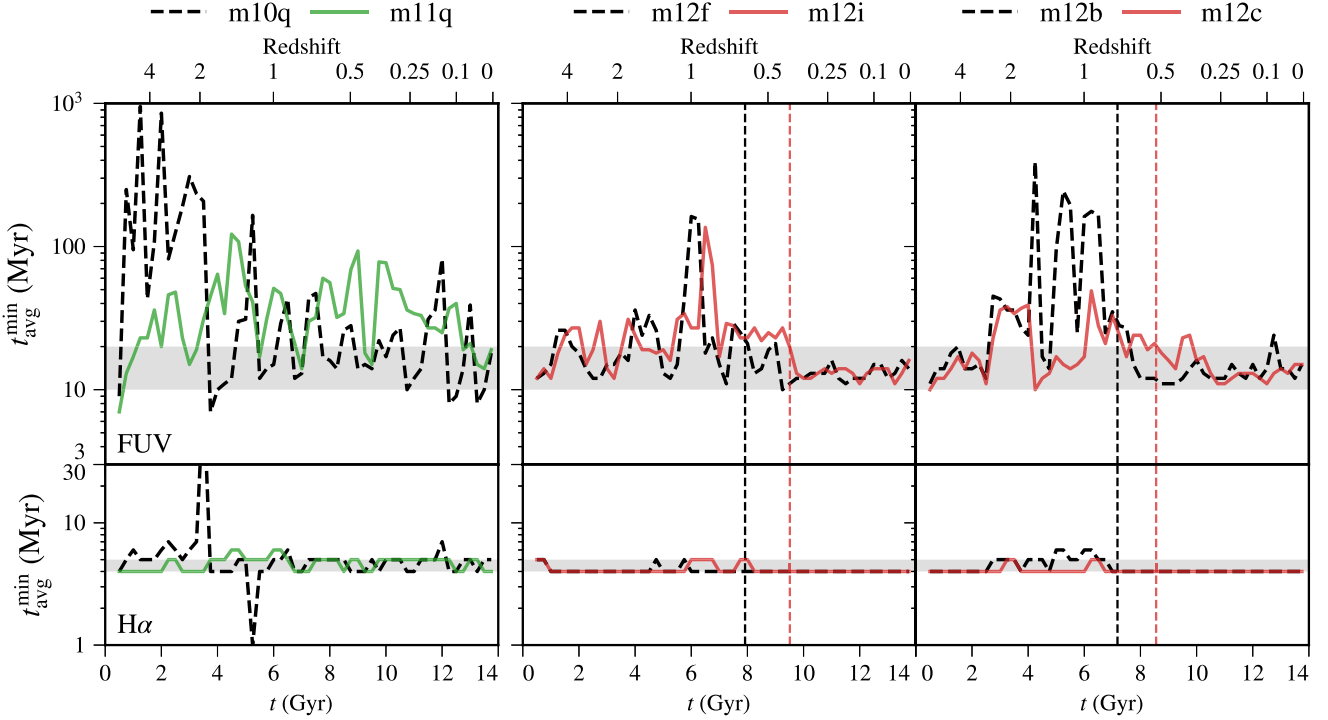


Figure 5. Best-fitting boxcar-average equivalent time-scale, $t_{\text{avg}}^{\text{min}}$, for both FUV (top) and H α (bottom) for different simulations as a function of cosmic time. Grey shaded regions are shown to help identify departures from the time-scales we identify in the time-steady limit for each indicator (10–20 Myr for FUV and 4–5 Myr for H α). *Left:* Simulations with halo mass $M_h \sim 10^{10} M_\odot$ and $M_h \sim 10^{11} M_\odot$ at $z = 0$. *Middle and right:* Simulations with $M_h \sim 10^{12} M_\odot$. The simulations of Milky Way-mass galaxies with $M_h \sim 10^{12} M_\odot$ experience a transition from bursty to time-steady SFR at the time indicated by the vertical dashed lines. For time periods when the SFR is bursty (regardless of halo mass) the best-fitting averaging time-scale for FUV fluctuates strongly and reaches much higher values than for time periods when the SFR is time-steady. For H α , the shorter best-fitting averaging time-scale 4–5 Myr is generally stable and insensitive to the shape of the SFH.

in Table 1 into their bursty and time-steady phases (the two rows in the figure). For each of these phases, we consider the FUV and H α indicators separately (the two columns). Then, for each panel, we evaluate the ratio $R_{\text{avg}}^{\text{ind}}$ for $t_{\text{avg}} = t_{\text{avg}}^{\text{min}}$, 100, 10, and 5 Myr at 1 Myr intervals (at each time, $t_{\text{avg}}^{\text{min}}$ is obtained by minimizing over a time window of total width $T_w = 500$ Myr, as described in Section 2.3). The distributions are shown for the results from all the simulations combined.

We see that, in almost all panels, $t_{\text{avg}}^{\text{min}}$ produces the smallest scatter around $R_{\text{avg}}^{\text{ind}} \approx 1$, as expected since $t_{\text{avg}}^{\text{min}}$ is defined as the best-fitting time-scale in each case. The $R_{\text{avg}}^{\text{ind}}$ distributions also confirm the results from the previous section that $t_{\text{avg}}^{\text{min}} \sim 10$ Myr for FUV and $t_{\text{avg}}^{\text{min}} \sim 5$ Myr for H α (see Fig. 5), since the distributions are similar when using either $t_{\text{avg}}^{\text{min}}$ or the corresponding constant averaging time-scale. On the other hand, the scatters for $t_{\text{avg}} = 100$ Myr and $t_{\text{avg}} = 10$ Myr are significantly larger in all panels, indicating again that these time-scales are typically much longer than those effectively probed by either FUV or H α .

To further quantify the $R_{\text{avg}}^{\text{ind}}$ distributions, we compute the standard deviations of $\log_{10} R_{\text{avg}}^{\text{ind}}$ for different cases and list the results in Table 2. For comparison, the typical observational uncertainty in SFRs measured through monochromatic indicators, combinations of monochromatic indicators, or SED modelling is ~ 0.1 – 0.3 dex (e.g. Hao et al. 2011; Koyama et al. 2015; Iyer et al. 2019). Thus, the 0.12–0.14 dex intrinsic scatter that we predict for FUV in the bursty regime for $t_{\text{avg}}^{\text{min}} \approx 10$ Myr, due solely to the ‘unknown’ preceding SFH, can be a significant additional source of uncertainty. On the hand, the intrinsic scatters ≈ 0.01 – 0.04 dex predicted for H α (in the

steady and bursty regimes) for $t_{\text{avg}}^{\text{min}} \approx 5$ Myr imply that this intrinsic uncertainty is usually subdominant.

4 DISCUSSION

4.1 Range of time-scales probed

So far, we have characterized the time-scales probed by SFR indicators by $t_{\text{avg}}^{\text{min}}$, which is the equivalent width of a boxcar-average response kernel. In reality, observational SFR indicators probe light emitted by stars with a range of different ages. This is especially true for FUV light, whose emission persists long after the H α emission powered by ionizing radiation becomes negligible (e.g. left-hand panel of Fig. 3). Another way to characterize the time-scales probed by an indicator is to partition the total luminosity observed at a time t by the age of the stars contributing to the light. At any given cosmic time t , we define $t_{50 \text{ per cent}}$ as the stellar age below which the combined luminosity in the band of interest is 50 per cent of the total luminosity, i.e. the light-weighted median stellar age. Similarly, we define t_x for $x = 65$ per cent, 90 per cent and 95 per cent. These time-scales reflect the extent to which older stars contribute to the observed luminosity compared to younger stars.

Fig. 7 illustrates how the total observed luminosity is contributed to by stars of different ages for two different cosmic times in m12i, one during a period in the bursty phase and one in the time-steady phase. Because the time-scale of H α is short compared to most SFR variations, it tracks the true SFR with relatively little time delay in either the bursty phase or the steady phase ($t_{50 \text{ per cent}} \sim 2$ – 4 Myr).

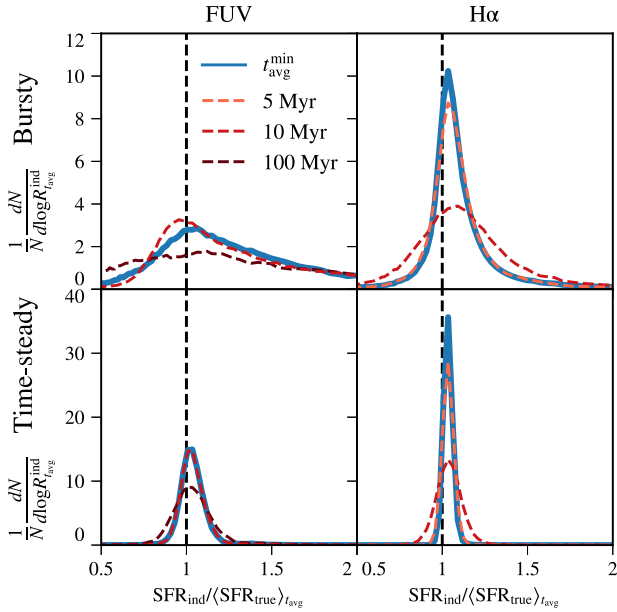


Figure 6. Distributions of the indicated-to-time-averaged-true-SFR ratio, $R_{t_{\text{avg}}}^{\text{ind}} \equiv \text{SFR}_{\text{ind}} / (\text{SFR}_{\text{true}})_{t_{\text{avg}}}$, for different averaging time-scales, t_{avg} , and SFR indicators (FUV, left column; H α , right column). SFHs from each simulation listed in Table 1 are partitioned into bursty (top row) and time-steady (bottom row) regimes and then combined to form each distribution plotted. Distributions using $t_{\text{avg}}^{\text{min}}$ are plotted in blue while different shades of red identify different fixed t_{avg} used to average the true SFR. Vertical dashed lines are plotted at a ratio of unity, while the standard deviations of these distributions in dex are compiled in Table 2. The scatter of the $\log_{10} R_{t_{\text{avg}}}^{\text{ind}}$ distribution, peaked near unity, is generally smallest for $t_{\text{avg}}^{\text{min}}$, though similar to the scatter for the fixed 10 and 5 Myr time-scales for FUV and H α , respectively.

Table 2. Intrinsic 1σ scatter in $\log_{10} (\text{SFR}_{\text{ind}} / (\text{SFR}_{\text{true}})_{t_{\text{avg}}})$ for different averaging time-scales in the bursty and time-steady star formation phases.

| SFR phase | $t_{\text{avg}}^{\text{min}}$ | FUV | | $t_{\text{avg}}^{\text{min}}$ | H α | |
|-------------|-------------------------------|--------------------|---------|-------------------------------|------------|--------|
| | | 10 Myr | 100 Myr | | 5 Myr | 10 Myr |
| Bursty | 0.143 ^b | 0.120 ^b | 0.242 | 0.035 | 0.041 | 0.094 |
| Time-steady | 0.025 | 0.026 | 0.042 | 0.011 | 0.014 | 0.030 |

Notes. ^aThis neglects uncertainties from observational effects, stellar population synthesis modelling, dust obscuration, etc.

^bNote that in this case the 10 Myr average produces a scatter smaller than (but similar to) the $t_{\text{avg}}^{\text{min}}$ average. This is possible because the minimization procedure used to obtain $t_{\text{avg}}^{\text{min}}$ is different from minimizing the width of the $R_{t_{\text{avg}}}^{\text{ind}}$ distribution (see Section 2.3). As a result, the time-scale that minimizes the RMSE in a 500 Myr window is not guaranteed to also minimize the 1σ scatter of the corresponding log-ratio distribution (Fig. 6). Moreover, stochasticity in the SFH, especially in the bursty regime, can introduce noise in the numerical estimates of scatter.

On the other hand, FUV exhibits a much larger difference between the two phases of star formation, with $t_{50 \text{ per cent}}$ fluctuating from ~ 5 to ~ 50 Myr in the bursty phase, before stabilizing in the time-steady phase.

Fig. 8 demonstrates in more detail how the stellar age distribution of observed FUV photons fluctuates by comparing $t_{\text{avg}}^{\text{min}}$ to different t_x for m12i as a function of cosmic time. Both $t_{\text{avg}}^{\text{min}}$ and t_x fluctuate much more strongly during the bursty phase (left of the vertical dashed black line) than in the more steady SFR phase (right of the

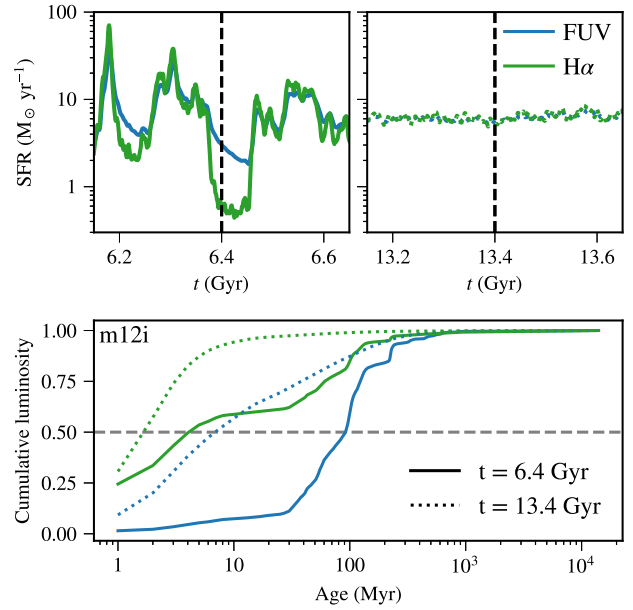


Figure 7. *Top:* SFHs near cosmic times $t = 6.4$ Gyr ($z = 0.9$ and ≈ 50 Myr after a burst of star formation; left) and $t = 13.4$ Gyr ($z = 0.03$ and long after the SFR has settled to a steady state; right) as indicated by FUV (blue) and H α (green) for m12i. *Bottom:* Cumulative luminosity in FUV and H α as a function of stellar age, at $t = 6.4$ Gyr (solid curves) and $t = 13.4$ Gyr (dotted curves). A horizontal dashed line is plotted at 50 per cent to identify the intersection with each cumulative luminosity curve, which we define as $t_{50 \text{ per cent}}$ (see Section 4.1). For FUV, $t_{50 \text{ per cent}}$ is strongly dependent on the recent SFH and is larger when the SFR is bursty compared to when it is time-steady. For H α , $t_{50 \text{ per cent}}$ also varies depending on the recent SFH but by a smaller factor ($\sim 2\times$ in the example shown).

vertical line). The fluctuations can cause the different t_x curves to reach values well above the quasi-steady values predicted at late times. This is especially so for the smaller percentiles, which are more sensitive to short time-scale SFR fluctuations. As expected from previous studies, the asymptotic value of $t_{90 \text{ per cent}} \sim 100$ Myr is consistent with the time-scale sometimes associated with FUV, defined such that FUV reach 90 per cent of its asymptotic luminosity for a constant SFH (see Section 1.1). In general, the distribution and time evolution of $t_x \text{ per cent}$ values depend on the details of the SFH, and so can depend on galaxy mass and redshift. However, for all simulations that achieve time-steady SFRs at late-times, the $t_x \text{ per cent}$ that best matches the late-time $t_{\text{avg}}^{\text{min}}$ is $t_{65 \text{ per cent}}$.

4.2 H α -to-FUV ratio as a gauge of SFR burstiness and its relation to $t_{\text{avg}}^{\text{min}}$ and $t_{50 \text{ per cent}}$

Since FUV and H α probe the SFR averaged over different time-scales, the ratio of FUV to H α can be used as a probe the SFR variability (e.g. Fumagalli, da Silva & Krumholz 2011; Weisz et al. 2012; Broussard et al. 2019). As the best-fitting time-scales $t_{\text{avg}}^{\text{min}}$ and $t_{50 \text{ per cent}}$ are also sensitive to changes in the SFR, we investigate whether the observed H α -to-FUV ratio can be used to infer more accurate (i.e. more applicable to the specific time corresponding to the observation) time-scales than the ones listed above that are implied statistically by the FIRE-2 simulated SFHs.

Fig. 9 shows a scatter plot of the H α -to-FUV ratio versus $t_{\text{avg}}^{\text{min}}$ (top panel) and $t_{50 \text{ per cent}}$ (bottom panel), for m12i and cosmic times $1 \text{ Gyr} < t < 13.8 \text{ Gyr}$ spaced by 1 Myr (this includes both times when the SFR is bursty and when it is steady). This figure focuses on

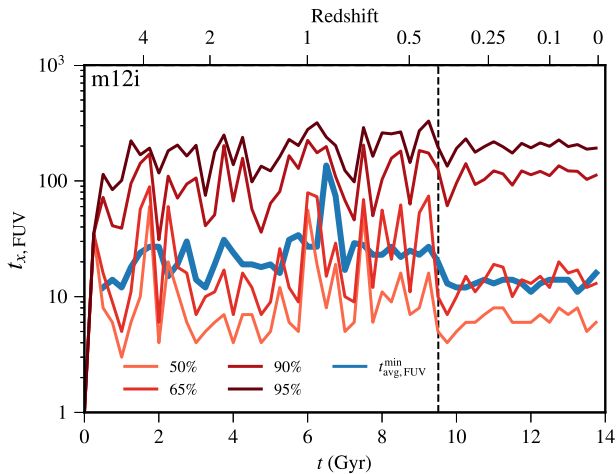


Figure 8. FUV light-weighted percentile stellar ages, t_x , for different percentile values ($x = 50$ per cent, 65 per cent, 90 per cent, and 95 per cent) as a function of cosmic time in m12i. The best-fitting boxcar-average equivalent time-scale, $t_{\text{avg}}^{\text{min}}$, is also shown for comparison (solid blue), along with the end of bursty star formation (vertical dashed line). Both $t_{\text{avg}}^{\text{min}}$ and t_x fluctuate much more strongly during the bursty phase. The fluctuations can cause the different t_x curves to reach values well above the quasi-steady values predicted at late times, especially, for the smaller percentiles, which are more sensitive to short time-scale SFR fluctuations.

FUV time-scales because, owing to their longer values, they are more sensitive to variations in the SFH than $\text{H}\alpha$. For $t_{\text{avg}}^{\text{min}}$, we see almost no correlation with the SFR ratio. This is because, as we defined it, $t_{\text{avg}}^{\text{min}}$ is computed by minimizing over a long time window $T_w = 500$ Myr. As a result, since the SFR (and consequently the $\text{H}\alpha$ -to-FUV ratio) fluctuates on time-scales smaller than T_w , during bursty phases $t_{\text{avg}}^{\text{min}}$ is a poor measure of the very recent (or ‘local’) SFH, which is the portion of the SFH probed by the instantaneous $\text{H}\alpha$ -to-FUV ratio. Thus, within a single window while $t_{\text{avg}}^{\text{min}}$ remains effectively constant the $\text{H}\alpha$ -to-FUV ratio takes many values (both less than and larger than unity) tracing the horizontal lines observed in the top panel of Fig. 9, and washing out any correlation between the two.

On the other hand, like the $\text{H}\alpha$ -to-FUV ratio, $t_{50 \text{ per cent}}$ is directly sensitive to the SFH at times immediately preceding the time of observation. This results in a much tighter anticorrelation between $t_{50 \text{ per cent}}$ and the $\text{H}\alpha$ -to-FUV ratio, visible in the bottom panel of the figure. The slope of relationship is negative because a smaller $\text{H}\alpha$ -to-FUV ratio implies that (on average) the SFR has decreased from the earlier times probed by FUV light to the more recent times probed by $\text{H}\alpha$ light. This implies that the stellar age of the stars below which 50 per cent of the observed FUV luminosity was produced, i.e. $t_{50 \text{ per cent, FUV}}$, is larger (i.e. formed during a recent burst of star formation). Quantitatively, the best-fitting slope of the power-law relation plotted in the bottom panel of the figure is -2.42 for m12i, though it can vary with galaxy mass and redshift as the statistics of SFR variations evolve. In addition to splitting the SFH into two distinct bursty and time-steady phases, and quantifying the best-fitting time-scales in each regime as we have done above, a Fourier analysis of the SFH can quantify the variability on different time-scales (Tacchella et al. 2016; Caplar & Tacchella 2019; Iyer et al. 2020; Wang & Lilly 2020). Caplar & Tacchella (2019) point out that, in their formalism, both the slope of the PSD on short time-scales and the ‘decorrelation time-scale’, τ_{break} (the time-scale after which the PSD resembles white noise) are accessible from observations of the $\text{H}\alpha$ -to-FUV ratio. Iyer et al.

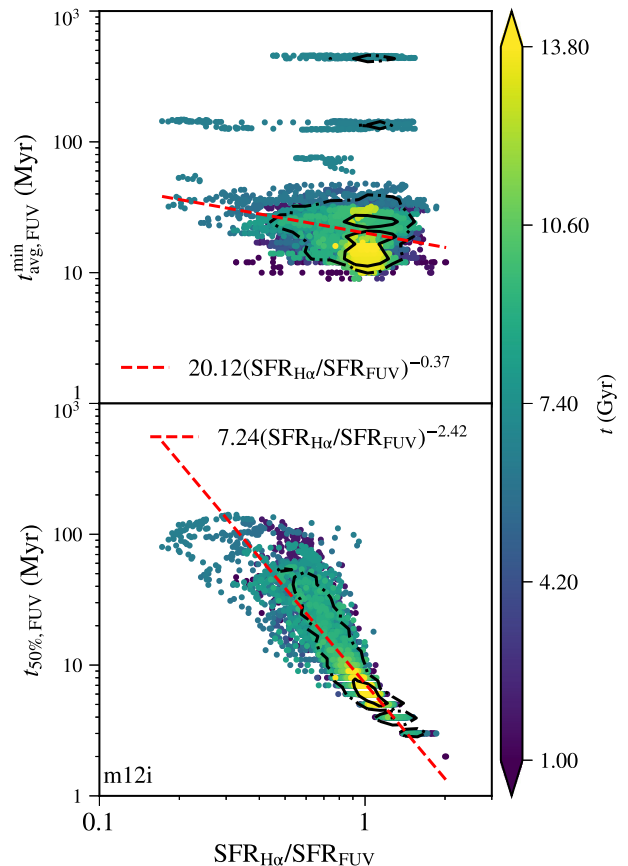


Figure 9. Scatter plots of the best-fitting boxcar-average equivalent time-scale, $t_{\text{avg}}^{\text{min}}$ (top) and the light-weighted median stellar age, $t_{50 \text{ per cent}}$, for FUV versus the instantaneous $\text{H}\alpha$ -to-FUV ratio in m12i. Each dot represents the value at one point of the full SFH, sampled at 1 Myr intervals (in each case, the first 1 Gyr is omitted to avoid edge effects near $t = 0$). 50th and 90th percentile contours are plotted in solid and dot-dashed black, respectively. A least-squares best-fitting line is plotted in dashed red, with parameters indicated in each panel. For $t_{\text{avg}}^{\text{min}}$, there is almost no correlation with the SFR ratio because $t_{\text{avg}}^{\text{min}}$ is computed by minimizing over a long time window $T_w = 500$ Myr (see Section 2.3). As a result, $t_{\text{avg}}^{\text{min}}$ is a poor measure of the recent SFH probed by the instantaneous $\text{H}\alpha$ -to-FUV ratio in bursty phases. On the other hand, like the $\text{H}\alpha$ -to-FUV ratio, $t_{50 \text{ per cent}}$ is directly sensitive to the recent SFH and is thus much more tightly anticorrelated with $\text{SFR}_{\text{H}\alpha}/\text{SFR}_{\text{FUV}}$.

(2020) directly quantifies the PSD of the ‘true’ SFHs in the FIRE-2 simulations.

In future work, it would be interesting to expand our analysis to connect more directly to PSD statistics, as well as to more systematically quantify the time-scales probed by SFR indicators as a function of the observed $\text{H}\alpha$ -to-FUV ratio.

4.3 Caveats and possible extensions

4.3.1 Dust attenuation

In this work, we have neglected attenuation of UV light by dust (e.g. Calzetti et al. 2000; Reddy et al. 2016). This assumption is reasonable for dwarf galaxies that contain little dust (e.g. Broussard et al. 2019), but attenuation by dust is in general critical to model to obtain correctly normalized SFR measurements. Since attenuation by dust depends on wavelength, it can also affect the ratio of different

SFR indicators, such as the $H\alpha$ -to-FUV ratio discussed above. Dust could modify the time-scales probed by an SFR indicator if the attenuation that it produces is time-dependent and correlated with the light sources (Koyama et al. 2019; Salim & Narayanan 2020). For example, if dust initially obscures star-forming regions but is cleared by stellar feedback after some time, this could modify the effective response function of the indicator. Properly modelling this kind of effect requires accounting for the exact spatial distribution and time-dependence of both the dust and light sources in galaxies, along with a detailed treatment of radiative transfer.

Zoom-in simulations such as the FIRE-2 simulations analysed in this paper can be combined with dust radiative transfer codes (e.g. Jonsson 2006; Camps & Baes 2015; Narayanan et al. 2020) to model these effects. This would also make it possible to study SFR indicators resulting from the processing of direct stellar radiation by dust, such as indicators based on IR radiation. A preliminary analysis of our simulations including dust radiative transfer suggests that, especially during bursty phases, there are substantial time-varying effects from dust attenuation. These effects are generally larger on FUV than on $H\alpha$, and we find that the magnitude of dust attenuation is strongly correlated with the true SFR, such that the emergent FUV flux can be lower when SFR is higher.

In this paper, we have assumed that observational techniques can correct for dust attenuation sufficiently accurately to recover intrinsic FUV and $H\alpha$ fluxes that can be compared to our idealized analysis, but in the future it will be important to more explicitly forward-model the radiative transfer to assess when this can be done robustly.

4.3.2 IMF sampling and variations

Another possible extension of this work would be to include the effects of stochastic sampling of the IMF, which could have interesting effects especially in dwarf galaxies (e.g. da Silva, Fumagalli & Krumholz 2014). Stochastic sampling of the IMF can have two different kinds of effects related to SFR burstiness. The first is an apparent signature of SFR variability in observed light. For example, even if a galaxy has a perfectly steady SFH, at low SFR its $H\alpha$ emission will fluctuate due to the fact that massive OB stars, which dominate the production of ionizing photons, form only rarely. In the limit of very low SFRs, these massive stars can be so rare that their number fluctuates significantly from time to time. The consequent, unavoidable light-curve fluctuations introduce biases in SFRs inferred using indicators calibrated under the assumption of a steady luminosity (see Section 2.2). The light-curve variability will also cause fluctuations in the $H\alpha$ -to-FUV ratio and could give the incorrect impression that the underlying SFH is fluctuating more than it really is.

The other effect of stochastic IMF sampling is dynamical and arises through the stellar feedback. In the FIRE-2 simulations analysed in this paper, the feedback processes are IMF averaged, meaning e.g. that each star particle has a radiative luminosity and supernova rate proportional to the rates for a stellar population that fully samples the IMF. For dwarf galaxies for which the mass of individual star particles can be $\approx 250 M_\odot$ (our m10q run) or lower, some star particles should in reality contain no O star, which would significantly modify the feedback from those star individual particles. Although such dynamical effects may be significant for some of the lowest-mass dwarfs and/or for certain sub-grid models (e.g. Applebaum et al. 2020), explicit comparisons with re-simulations that do include the effects of stochastic sampling on feedback suggest that stochastic IMF sampling does not substantially modify the SFH

statistics for the FIRE-2 simulations analysed in this paper (Su et al. 2018; Wheeler et al. 2019).

Additionally, if the IMF varies with environment or galaxy properties (e.g. van Dokkum & Conroy 2010), the luminosity response functions for the SFR indicators would also vary, potentially altering the time-scales of SFR indicators. Broussard et al. (2019) showed that the ratio of the near-ultraviolet (NUV, $\sim 2750 \text{ \AA}$)-to- $H\alpha$, η , is an additional indicator of bursty star formation on time-scales similar those identified for FUV here. They find that when the slope of the high-mass end of the IMF is steepened, the effective the time-scale probed by the η ratio decreases. This is because emission in both $H\alpha$ and the NUV becomes even more dominated by short-lived O and B stars. IMF variations would also introduce dynamical feedback effects which could affect the SFHs predicted by simulations.

5 CONCLUSIONS

In this paper, we quantify the time-scales probed by the FUV and $H\alpha$ indicators commonly used to measure SFRs by using star formation histories from FIRE-2 cosmological zoom-in simulations (Hopkins et al. 2018). A key goal of our study is to investigate the effects of realistic SFHs, including secular trends as well as short-term variability (‘burstiness’), on the time-scales probed by standard indicators calibrated to constant SFHs (e.g. Kennicutt & Evans 2012). The SFHs for Milky Way-mass galaxies in the FIRE simulations exhibit a transition from a highly time-variable, or bursty, phase at early times to a time-steady phase at late times. Moreover, dwarf galaxies are bursty at all epochs (Fig. 2). Our analysis is on based mock, unattenuated FUV and $H\alpha$ light curves produced by combining SFHs from FIRE-2 simulations with the BPASS stellar population synthesis model, which includes the effects of binary stars (Eldridge et al. 2017).

We define $t_{\text{avg}}^{\text{min}}$, the best-fitting boxcar-average equivalent time-scale, which is a statistical measure computed in 500 Myr-wide moving windows that cover representative samples of short-time-scale SFR fluctuations (see Section 2.3). We find that $t_{\text{avg}}^{\text{min}} \sim 5 \text{ Myr}$ for $H\alpha$, independent of whether the SFH is bursty or time-steady. On the other hand, $t_{\text{avg}}^{\text{min}}$ fluctuates more during the bursty phase of star formation, from $t_{\text{avg}}^{\text{min}} \sim 10 \text{ Myr}$ to $> 100 \text{ Myr}$ in cases following particularly extreme bursts of star formation. However, when the SFR becomes time-steady, the FUV $t_{\text{avg}}^{\text{min}}$ approaches $\sim 10 \text{ Myr}$ (Fig. 5). These best-fitting boxcar-average equivalent time-scales are notable because theoretical models often assume that FUV and $H\alpha$ observations probe SFRs boxcar-averaged over the much longer time-scales of 100 and 10 Myr, respectively (e.g. Orr et al. 2018; Hani et al. 2020). This difference arises because, for example, while stellar populations emit significant FUV for ages exceeding 100 Myr, the integrated UV emission is strongly weighted towards younger, massive stars.

In order to further characterize the distribution of stellar ages contributing to each SFR indicator, we also define $t_{50 \text{ per cent}}$ as the light-weighted median stellar age, and analogous quantities for different age percentiles. Unlike $t_{\text{avg}}^{\text{min}}$, $t_{50 \text{ per cent}}$ is a local measure of the recent SFH, in the sense that it is computed directly based on the SFH immediately preceding the time of interest (or observation) rather than being a statistical best fit over a large sampling window. We focused our analysis of the stellar age distribution on FUV (Fig. 8), which probes a wider age range. For FUV, $t_{50 \text{ per cent}}$ fluctuates strongly, from ~ 5 to $\sim 50 \text{ Myr}$ in the bursty phase for the example shown in the figure. As expected from previous studies, the asymptotic value of $t_{90 \text{ per cent}} \sim 100 \text{ Myr}$ is consistent with the time-scale sometimes associated with FUV, defined such that FUV

reach 90 per cent of its asymptotic luminosity for a constant SFH. The distribution and time evolution of $t_{x \text{ per cent}}$ values in general depend on the details of the SFH. However, we find that for all simulations that achieve time-steady SFRs at late-times, the $t_{x \text{ per cent}}$ that best matches the late-time $t_{\text{avg}}^{\text{min}}$ is $t_{65 \text{ per cent}}$.

We also investigated how the observed $H\alpha$ -to-FUV ratio can tell us more about the time-scale probed by FUV light. Since $H\alpha$ and FUV are sensitive to light from stars of different ages, the observed ratio is sensitive to the shape of the recent SFH. In turn, the shape of the recent SFH affects $t_{50 \text{ per cent, FUV}}$. Indeed, the simulations predict a strong anticorrelation between $t_{50 \text{ per cent, FUV}}$ and $\text{SFR}_{H\alpha}/\text{SFR}_{\text{FUV}}$ (Fig. 9). This anticorrelation arises because, in periods following a burst of star formation, the FUV luminosity is dominated by older stars formed during the burst (corresponding to large $t_{50 \text{ per cent, FUV}}$) while the $H\alpha$ luminosity is dominated by younger stars. Thus, following a burst after which the true SFR decays $\text{SFR}_{H\alpha}/\text{SFR}_{\text{FUV}}$ decreases to <1 . More generally, our results support the use of the $H\alpha$ -to-FUV ratio as an observational probe of SFR variability (e.g. Weisz et al. 2012).

Finally, we note that there are several avenues for extending the present study, which could lead to further important insights. These include the effects of dust attenuation and reprocessing of stellar light into IR indicators, stochastic sampling and possible environmental variations of the IMF, as well as the use of more detailed panchromatic data. It is also important to reiterate that the detailed quantitative results we have presented apply to the set of FIRE-2 simulations we have analysed. Different galaxy formation simulations predict different SFR variability statistics (Iyer et al. 2020), due to e.g. different models for star formation and stellar feedback. For this reason, comparisons with observational statistics that quantitatively constrain the SFR variability in observed galaxy populations (e.g. Sparre et al. 2017; Broussard et al. 2019; Caplar & Tacchella 2019; Wang & Lilly 2020), will play a key role, in combination with the methodology presented in this paper, in more precisely quantifying the time-scales probed by different indicators.

ACKNOWLEDGEMENTS

This paper represents the effort of José Antonio Flores Velázquez, whose life was tragically taken on 2019 August 14 before he could see this project finished. José was a beloved son, brother, friend, and colleague, with a warm smile and a big heart. Those who knew him were inspired by his unwavering optimism and passion for his work. On top of being a great person, he was also a remarkably talented astronomer, with great potential as exemplified by his drive to learn things by himself while also recognizing when to reach out for help. His refusal to sacrifice his identity also made, and continues to make, him a role model to folks who struggle to embody their culture in academic settings. The field of Astronomy lost a brilliant scientist and our goal is to share his work with the community.

The authors thank the anonymous referee for their careful review of this manuscript. JAFV was supported by a National Science Foundation Graduate Research Fellowship Program under grant DGE-1839285. ABG was also supported by an NSF-GRFP under grant DGE-1842165 and was additionally supported by NSF grants DGE-0948017 and DGE-145000. CAFG was supported by NSF through grants AST-1412836, AST-1517491, AST-1715216, and CAREER award AST1652522, by NASA through grants NNX15AB22G and 17-ATP17-0067, by STScI through grants HST-GO-14681.011, HST-GO-14268.022-A, and HST-AR-14293.001-A, and by a Cottrell Scholar Award from the Research Corporation for Science Advancement. JSB, AL, and FJM were supported by NSF grants AST-1910346 and AST-1518291. Support for JM is provided by the

NSF (AST Award Number 1516374), and by the Harvard Institute for Theory and Computation, through their Visiting Scholars Program. JS is supported as a CIERA Fellow by the CIERA Postdoctoral Fellowship Program (Center for Interdisciplinary Exploration and Research in Astrophysics, Northwestern University). AW received support from NASA through ATP grant 80NSSC18K1097 and HST grants GO-14734, AR-15057, AR-15809, and GO-15902 from STScI; the Heising-Simons Foundation; and a Hellman Fellowship. KE was supported by an NSF-GRFP and a Hellman fellowship from UC Berkeley. The Flatiron Institute is supported by the Simons Foundation. Numerical calculations were run on the Quest computing cluster at Northwestern University, the Wheeler computing cluster at Caltech, XSEDE allocations TG-AST130039, TG-AST120025, and TG-AST140023, Blue Waters PRAC allocation NSF.1713353, and NASA HEC allocations SMD16-7592, SMD-16-7561 and SMD-17-1204. We honour the invaluable labour of the maintenance and clerical staff at our institutions, whose contributions make our scientific discoveries a reality. This research was conducted on Potawatomi Indigenous land.

DATA AVAILABILITY

The data supporting the plots within this article are available on reasonable request to the corresponding author. A public version of the GIZMO code is available at <http://www.tapir.caltech.edu/~phopkin/Site/GIZMO.html>. Additional data including simulation snapshots, initial conditions, and derived data products are available at <http://fire.northwestern.edu/data/>.

REFERENCES

- Agertz O., Kravtsov A. V., 2015, *ApJ*, 804, 18
 Anglés-Alcázar D., Faucher-Giguère C.-A., Kereš D., Hopkins P. F., Quataert E., Murray N., 2017, *MNRAS*, 470, 4698
 Applebaum E., Brooks A. M., Quinn T. R., Christensen C. R., 2020, *MNRAS*, 492, 8
 Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, *MNRAS*, 351, 1151
 Broussard A. et al., 2019, *ApJ*, 873, 74
 Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
 Bryan G. L., Norman M. L., 1998, *ApJ*, 495, 80
 Buat V. et al., 2012, *Highlights of Astronomy*, 16, 495
 Calzetti D., 2013, in Falcón-Barroso J., Knapen J. H., eds, *Star Formation Rate Indicators*. Cambridge Univ. Press, Cambridge, p. 419
 Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, *ApJ*, 533, 682
 Calzetti D. et al., 2007, *ApJ*, 666, 870
 Camps P., Baes M., 2015, *Astron. Comput.*, 9, 20
 Caplar N., Tacchella S., 2019, *MNRAS*, 487, 3845
 Conroy C., Gunn J. E., White M., 2009, *ApJ*, 699, 486
 da Silva R. L., Fumagalli M., Krumholz M., 2012, *ApJ*, 745, 145
 da Silva R. L., Fumagalli M., Krumholz M. R., 2014, *MNRAS*, 444, 3275
 Dickey C. M. et al., 2020, *ApJ*, preprint ([arXiv:2010.01132](https://arxiv.org/abs/2010.01132))
 Domínguez A., Siana B., Brooks A. M., Christensen C. R., Bruzual G., Stark D. P., Alavi A., 2015, *MNRAS*, 451, 839
 Elbaz D. et al., 2007, *A&A*, 468, 33
 Elbaz D. et al., 2018, *A&A*, 616, A110
 Eldridge J. J., Stanway E. R., Xiao L., McClelland L. A. S., Taylor G., Ng M., Greis S. M. L., Bray J. C., 2017, *PASA*, 34, e058
 Emami N., Siana B., Weisz D. R., Johnson B. D., Ma X., El-Badry K., 2019, *ApJ*, 881, 71
 Faucher-Giguère C.-A., 2018, *MNRAS*, 473, 3717
 Faucher-Giguère C.-A., Hopkins P. F., Kereš D., Muratov A. L., Quataert E., Murray N., 2015, *MNRAS*, 449, 987

Faucher-Giguère C.-A., Feldmann R., Quataert E., Kereš D., Hopkins P. F., Murray N., 2016, *MNRAS*, 461, L32

Ferland G. J., Korista K. T., Verner D. A., Ferguson J. W., Kingdon J. B., Verner E. M., 1998, *PASP*, 110, 761

Fitts A. et al., 2018, *MNRAS*, 479, 319

Forrest B. et al., 2017, *ApJ*, 838, L12

Fumagalli M., da Silva R. L., Krumholz M. R., 2011, *ApJ*, 741, L26

Garrison-Kimmel S. et al., 2017, *MNRAS*, 471, 1709

Garrison-Kimmel S. et al., 2019, *MNRAS*, 487, 1380

Geha M., Blanton M. R., Yan R., Tinker J. L., 2012, *ApJ*, 757, 85

Gill S. P. D., Knebe A., Gibson B. K., 2004, *MNRAS*, 351, 399

Governato F. et al., 2012, *MNRAS*, 422, 1231

Guo Y. et al., 2016, *ApJ*, 833, 37

Gurvich A. B. et al., 2020, *MNRAS*, 498, 3664

Hafen Z. et al., 2017, *MNRAS*, 469, 2292

Hani M. H., Hayward C. C., Orr M. E., Ellison S. L., Torrey P., Murray N., Wetzel A., Faucher-Giguère C.-A., 2020, *MNRAS*, 493, L87

Hao C.-N., Kennicutt R. C., Johnson B. D., Calzetti D., Dale D. A., Moustakas J., 2011, *ApJ*, 741, 124

Hayward C. C. et al., 2014, *MNRAS*, 445, 1598

Hirtenstein J. et al., 2019, *ApJ*, 880, 54

Hopkins P. F., 2015, *MNRAS*, 450, 53

Hopkins P. F., Narayanan D., Murray N., 2013, *MNRAS*, 432, 2647

Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, *MNRAS*, 445, 581

Hopkins P. F. et al., 2018, *MNRAS*, 480, 800

Iyer K. G., Gawiser E., Faber S. M., Ferguson H. C., Kartaltepe J., Koekemoer A. M., Pacifici C., Somerville R. S., 2019, *ApJ*, 879, 116

Iyer K. G. et al., 2020, *MNRAS*, 498, 430

Jonsson P., 2006, *MNRAS*, 372, 2

Kauffmann G., 2014, *MNRAS*, 441, 2717

Kennicutt R. C. Jr, Evans N. J., 2012, *ARA&A*, 50, 531

Knollmann S. R., Knebe A., 2009, *ApJ*, 182, 608

Koyama Y. et al., 2015, *MNRAS*, 453, 879

Koyama Y., Shimakawa R., Yamamura I., Kodama T., Hayashi M., 2019, *PASJ*, 71, 8

Kroupa P., Tout C. A., Gilmore G., 1993, *MNRAS*, 262, 545

Leitherer C. et al., 1999, *ApJ*, 123, 3

Ma X., Hopkins P. F., Faucher-Giguère C.-A., Zolman N., Muratov A. L., Kereš D., Quataert E., 2016a, *MNRAS*, 456, 2140

Ma X., Hopkins P. F., Kasen D., Quataert E., Faucher-Giguère C.-A., Kereš D., Murray N., Strom A., 2016b, *MNRAS*, 459, 3614

Ma X., Hopkins P. F., Feldmann R., Torrey P., Faucher-Giguère C.-A., Kereš D., 2017, *MNRAS*, 466, 4780

Matthee J., Schaye J., 2019, *MNRAS*, 484, 915

Muratov A. L., Kereš D., Faucher-Giguère C.-A., Hopkins P. F., Quataert E., Murray N., 2015, *MNRAS*, 454, 2691

Murphy E. J. et al., 2011, *ApJ*, 737, 67

Narayanan D. et al., 2020, *ApJ*, preprint (arXiv:2006.10757)

Noeske K. G. et al., 2007, *ApJ*, 660, L43

Orr M. E. et al., 2017, *ApJ*, 849, L2

Orr M. E. et al., 2018, *MNRAS*, 478, 3653

Pelliccia D. et al., 2020, *ApJ*, 896, L26

Reddy N. A., Steidel C. C., Pettini M., Bogosavljević M., 2016, *ApJ*, 828, 107

Salim S., Narayanan D., 2020, *ARA&A*, 58, 529

Salim S., Lee J. C., Ly C., Brinchmann J., Davé R., Dickinson M., Salzer J. J., Charlot S., 2014, *ApJ*, 797, 126

Schreiber C. et al., 2015, *A&A*, 575, A74

Shivaei I. et al., 2015, *ApJ*, 815, 98

Smit R., Bouwens R. J., Labbé I., Franx M., Wilkins S. M., Oesch P. A., 2016, *ApJ*, 833, 254

Sparre M., Hayward C. C., Feldmann R., Faucher-Giguère C.-A., Muratov A. L., Kereš D., Hopkins P. F., 2017, *MNRAS*, 466, 88

Stern J. et al., 2020, *ApJ*, preprint (arXiv:2006.13976)

Su K.-Y. et al., 2018, *MNRAS*, 480, 1666

Tacchella S., Dekel A., Carollo C. M., Ceverino D., DeGraf C., Lapiner S., Mand elker N., Primack Joel R., 2016, *MNRAS*, 457, 2790

Tacchella S., Forbes J. C., Caplar N., 2020, *MNRAS*, 497, 698

Teyssier R., Pontzen A., Dubois Y., Read J. I., 2013, *MNRAS*, 429, 3068

van der Wel A. et al., 2011, *ApJ*, 742, 111

van Dokkum P. G., Conroy C., 2010, *Nature*, 468, 940

Wang E., Lilly S. J., 2020, *ApJ*, 895, 25

Weisz D. R. et al., 2012, *ApJ*, 744, 44

Wheeler C. et al., 2019, *MNRAS*, 490, 4447

Whitaker K. E., van Dokkum P. G., Brammer G., Franx M., 2012, *ApJ*, 754, L29

Willett K. W. et al., 2015, *MNRAS*, 449, 820

APPENDIX: DEPENDENCE OF OUR RESULTS ON WIDTH T_w OF THE TIME WINDOW

In Section 2.3, we define $t_{\text{avg}}^{\text{min}}$, the best-fitting boxcar-average equivalent time-scale for ensembles of points in SFHs grouped in moving windows of fiducial width $T_w = 500$ Myr. In Fig. A1, we show that the choice of T_w , for sufficiently large T_w , does not affect the inferred average value of $t_{\text{avg}}^{\text{min}}$. For H α (bottom panel), $t_{\text{avg}}^{\text{min}}$ is largely insensitive to the window size so long as it is larger than $T_w = 50$ Myr (save for the first 1 Gyr, which is affected by edge effects near $t = 0$). On the other hand, FUV (top panel) shows stronger variations for $T_w < 250$ Myr, but above this value $t_{\text{avg}}^{\text{min}}$ is well converged.

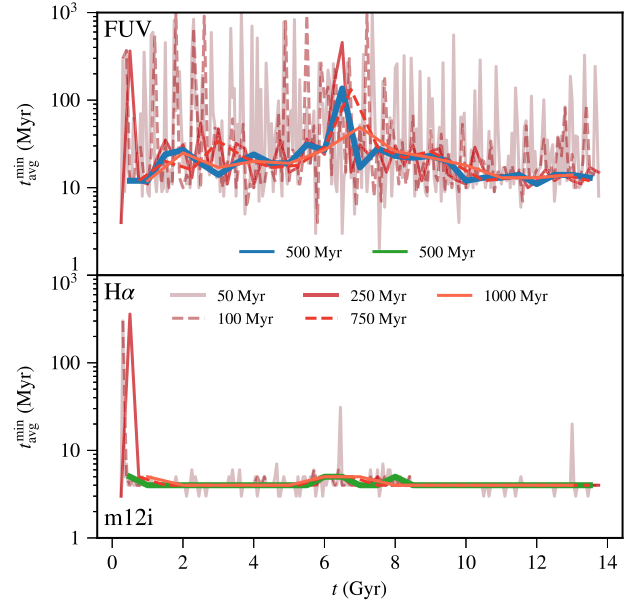


Figure A1. History of $t_{\text{avg}}^{\text{min}}$ over cosmic time for FUV (top) and H α (bottom) in non-overlapping windows for different choices of window width T_w (indicated by colour) in m12i. For sufficiently large T_w ($\gtrsim 250$ Myr for FUV and $\gtrsim 50$ Myr for H α), $t_{\text{avg}}^{\text{min}}$ is relatively insensitive to the choice of T_w .

This paper has been typeset from a \LaTeX file prepared by the author.