

First Light and Reionization Epoch Simulations (FLARES) – XIV. The Balmer/4000 Å breaks of distant galaxies

Stephen M. Wilkins^{1b, 1,2★}, Christopher C. Lovell^{1b, 1,3}, Dimitrios Irodotou^{1b, 4}, Aswin P. Vijayan^{1b, 5,6}, Anton Vikaeus^{1b, 7}, Erik Zackrisson^{1b, 7}, Joseph Caruana^{2,8}, Elizabeth R. Stanway^{1b, 9}, Christopher J. Conselice¹⁰, Louise T. C. Seeyave¹, William J. Roper^{1b, 1}, Katherine Chworowsky¹¹ and Steven L. Finkelstein¹¹

¹*Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QH, UK*

²*Institute of Space Sciences and Astronomy, University of Malta, Msida MSD 2080, Malta*

³*Institute of Cosmology and Gravitation, University of Portsmouth, Burnaby Road, Portsmouth PO1 3FX, UK*

⁴*Department of Physics, University of Helsinki, Gustaf Hällströmin katu 2, FI-00014 Helsinki, Finland*

⁵*Cosmic Dawn Center (DAWN), Rådmandsgade 62, 2200 København, Denmark*

⁶*DTU-Space, Technical University of Denmark, Elektrovej 327, DK-2800 Kgs. Lyngby, Denmark*

⁷*Observational Astrophysics, Department of Physics and Astronomy, Uppsala University, Box 516, SE-751 20 Uppsala, Sweden*

⁸*Department of Physics, Faculty of Science, University of Malta, Msida MSD 2080, Malta*

⁹*Department of Physics, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK*

¹⁰*Jodrell Bank Centre for Astrophysics, Department of Physics and Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, UK*

¹¹*Department of Astronomy, The University of Texas at Austin, Austin, Texas 78712-1205, USA*

Accepted 2023 November 14. Received 2023 November 14; in original form 2023 May 29

ABSTRACT

With the successful launch and commissioning of *JWST* we are now able to routinely spectroscopically probe the rest-frame optical emission of galaxies at $z > 6$ for the first time. Among the most useful spectral diagnostics used in the optical is the Balmer/4000 Å break; this is, in principle, a diagnostic of the mean ages of composite stellar populations. However, the Balmer break is also sensitive to the shape of the star formation history, the stellar (and gas) metallicity, the presence of nebular continuum emission, and dust attenuation. In this work, we explore the origin of the Balmer/4000 Å break using the SYNTHESIZER synthetic observations package. We then make predictions of the Balmer/4000 Å break using the First Light and Reionization Epoch Simulations at $5 < z < 10$. We find that the average break strength weakly correlates with stellar mass and rest-frame far-ultraviolet luminosity, but that this is predominantly driven by dust attenuation. We also find that break strength provides a weak diagnostic of the age but performs better as a means to constrain star formation and stellar mass, alongside the ultraviolet and optical luminosity, respectively.

Key words: methods: numerical – galaxies: evolution – galaxies: formation – galaxies: high-redshift – infrared: galaxies.

1 INTRODUCTION

A key goal in extragalactic astrophysics is to accurately measure various fundamental physical properties – including the star formation rate, stellar mass, metallicity, and dust attenuation – for the distant, high-redshift ($z > 4$), galaxy population. Doing so provides the essential constraints to galaxy formation models, ultimately providing insights into physical processes responsible for early galaxy formation and evolution.

Obtaining robust constraints on these properties requires deep near-infrared observations probing the rest-frame ultraviolet (UV) and optical emission. Early insights have come from the combination of *Hubble* and *Spitzer* (e.g. Hashimoto et al. 2018; De Barros et al. 2019; Endsley et al. 2021; Tacchella et al. 2022). However, these

insights were limited by the low sensitivity and resolution of *Spitzer*, and contamination from strong line emission (e.g. Stark et al. 2013; Wilkins et al. 2013, 2016, 2020). Through its sensitivity, wavelength-coverage, and spectroscopic capabilities, *JWST* is now overcoming these limitations, with the first spectroscopic measurements now emerging (e.g. Schaerer et al. 2022; Arrabal Haro et al. 2023; Boyett et al. 2023; Bunker et al. 2023a, b; Fujimoto et al. 2023; Matthee et al. 2023; Stiavelli et al. 2023; Trump et al. 2023; Trussler et al. 2023). Crucially, for this work, these spectroscopic studies now include detections of the rest-frame UV and optical continuum at sufficient signal to noise to enable the measurement of key continuum features (e.g. Bunker et al. 2023b; Cameron et al. 2023; Carnall et al. 2023; Hsiao et al. 2023; Looser et al. 2023; Vikaeus et al. 2023).

While nebular emission lines, which provide insights into the interstellar medium (ISM) and sources of ionizing photons, are a key target for spectroscopic observations, continuum features also contain invaluable information about the stellar populations, in

* E-mail: S.Wilkins@sussex.ac.uk

particular the older aspects of the stellar population. Arguably, the most prized continuum feature is the combined Balmer/4000 Å break feature (Bruzual 1983; Poggianti & Barbaro 1997). For a *simple* (single-aged) stellar population this feature is strongly sensitive to the age, making it a potentially powerful diagnostic of the star formation histories (SFHs) of galaxies (Worthey 1994; Poggianti & Barbaro 1997; Scoville et al. 2015).

Previous indications of notably strong Balmer breaks in high-redshift galaxies such as Hashimoto et al. (2018), Roberts-Borsani, Ellis & Laporte (2020), and Laporte et al. (2021) prompt further investigation into the underlying mechanisms responsible for producing such strong breaks. Importantly, we seek to discern whether current state-of-the-art hydrodynamic simulations can accommodate such detections – thus shedding new light on the debate of the onset of star formation in the early Universe.

In this work, we begin by exploring the physical properties influencing the strength of this feature. Specifically, we use the SYNTHESIZER¹ (Lovell et al., in preparation) synthetic observations code to explore how the choice of SFH, metallicity, nebular emission, dust, initial mass function (IMF), and choice of stellar population synthesis (SPS) models affect this observational feature. These are modelled using relatively simple parametrizations, in order to clarify their individual effects on this spectral break. Such features can be obscured in the more complex history and combination of physical properties described by full computational simulations for any given galaxy.

We then make detailed predictions of the Balmer break strength using the First Light and Reionization Epoch Simulations (FLARES; Lovell et al. 2021; Vijayan et al. 2021) for galaxies at high redshift ($5 \leq z \leq 10$). FLARES is a suite of hydrodynamical zoom simulations that, by virtue of its strategy, simulates a wide range of galaxies with ($M_{\star} = 10^8 - 10^{11} M_{\odot}$) at high redshift ($z > 5$). Unlike our simple modelling, which assumes a smooth parametrization of the SFH, FLARES galaxies have realistic, complex star formation and metal enrichment histories and star–dust geometries. *So far*, FLARES has been shown to produce an excellent match to observational constraints (Lovell et al. 2021; Vijayan et al. 2021; Roper et al. 2022; Tacchella et al. 2022; Wilkins et al. 2022, 2023a, b), including early constraints from *JWST* (Wilkins et al. 2023a, b; Adams et al. 2023). Using FLARES we explore how the predicted break strength correlates with stellar mass, far-UV luminosity, and key physical properties, including the specific star formation rate, age, dust attenuation, and optical mass-to-light ratio, in galaxies with realistically complex and stochastic SFHs.

This paper is structured as follows: We begin, in Section 2, to explore the impact of physical properties on the Balmer break using simple star formation and metal enrichment histories. Next, in Section 3, we introduce the FLARES project, including a detailed description of our spectral energy distribution (SED) modelling procedure. In Section 3.2, we make predictions for the break strength, including a comparison with current spectroscopic observations (Section 3.4), and an exploration of how the predicted break strength correlates with and is impacted by various physical properties (Section 3.3). We then summarize our findings in this work and present our conclusion in Section 4.

2 THEORETICAL BACKGROUND

In this work, we are interested in exploring predictions for the strength of the Balmer/4000 Å break feature. The Balmer break is a

sharp feature occurring at the Balmer limit (3645 Å) and is caused by bound-free absorption by electrons in the electron band $n = 2$. This feature is most prevalent in A stars ($1.4 - 2.1 M_{\odot}$) where effective temperatures of $\approx 10\,000$ K maximize the number of electrons in the $n = 2$ band. In hotter (OB) stars, an increasing fraction of hydrogen is ionized reducing the absolute number of electrons in $n = 2$, while in cooler stars the falling temperature reduces the number in $n = 2$. However, the spectra of cooler stars are increasingly affected by line blanketing. The combination of these features gives rise to a strong dependence slope around 4000 Å to the age of single-aged stellar population.

A popular metric of the break is D_{4000} defined by Bruzual (1983) using windows at [3750, 3950] Å and [4050, 4250] Å. However, the blue window of D_{4000} is redward of the Balmer limit (3645 Å) and thus not as useful for quantifying the break in younger galaxies. For this reason, we adopt the definition of Binggeli et al. (2019) and define the break strength using two windows at [3400, 3600] Å and [4150, 4250] Å instead. These windows are chosen to straddle the Balmer limit, and to also avoid strong nebular line emission while minimizing their separation to reduce the impact of dust attenuation. Throughout this work, we express the break strength simply as the logarithmic ratio of the red and blue window luminosities.

For young, star-forming galaxies, dominated by emission from O and B stars, and with little dust attenuation, this quantity is close to zero. For older stellar populations, this quantity becomes progressively more positive, with a 1 Gyr old simple stellar population reaching a strength of ≈ 0.5 .

To gain insight into the physical properties driving the break strength, we utilize the SYNTHESIZER (Lovell et al., in preparation) synthetic observations pipeline to make predictions assuming a simple star formation and metal enrichment history. SYNTHESIZER is designed to enable the creation of synthetic observations using both a particle- or parametric-based approach, and is designed to be more flexible, e.g. in terms of the choice of SPS model, IMF, or dust attenuation model. A (particle-only) precursor to SYNTHESIZER was used to generate predictions from FLARES as described in Section 3.1. Here, we use SYNTHESIZER to explore the sensitivity of the break to the star formation and metal enrichment history, reprocessing by dust and gas, as well as the choice of SPS model and IMF.

For this exploration, we employ smooth parametric SFHs, a single metallicity, and a simple screen model for reprocessing by dust and gas (i.e. stellar populations are all equally affected by dust). By default we assume the same SPS model and IMF as utilized by FLARES (Vijayan et al. 2021): v2.2.1 of the Binary Population And Spectral Synthesis (BPASS; Stanway & Eldridge 2018) and a Chabrier (2003) IMF extending to $300 M_{\odot}$. To model the contribution of nebular emission, we use v17.03 of the CLOUDY photoionization code (Ferland et al. 2017), assuming identical nebular and stellar metallicities, and assume a reference ionization parameter² of $U_{\text{ref}} = 0.01$. In Fig. 1, we use this model to visualize the break for four scenarios: a young 10 Myr burst of dust-free star formation (A), 100 Myr constant star formation with dust (C) and without dust (B), and a 1 Gyr old burst (D). The differences between these models are discussed in the proceeding sections.

²Here, the reference ionization parameter is the ionization parameter assumed for stellar populations with an age of 1 Myr and $Z = 0.01$. For stellar populations with different ages and metallicities the actual assumed ionization parameter is scaled $\propto Q^{1/3}$ – see Wilkins et al. (2020, 2023b) for more details.

¹<https://github.com/flaresimulations/synthesizer>

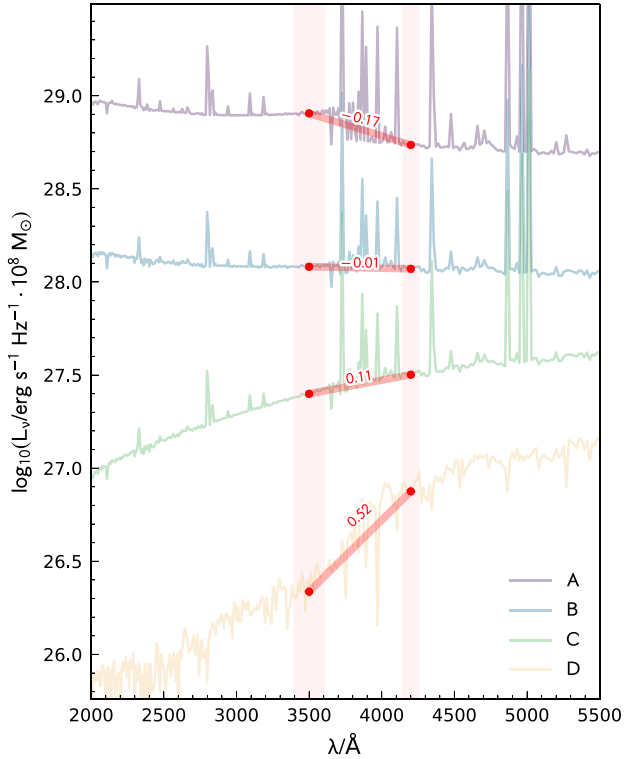


Figure 1. Visualization of the Balmer break in the context of four simple scenarios, each with the same stellar mass, $M_* = 10^8 M_\odot$. A: 10 Myr constant star formation and no dust. B: 100 Myr constant star formation and no dust. C: 100 Myr constant star formation and V-band optical depth $\tau_V = 1$. D: 1 Gyr old burst and no dust. Each scenario assumes $Z = 0.001$ and no escape of Lyman continuum photons (i.e. $f_{\text{esc}} = 0$). The two Balmer break windows are denoted by the shaded vertical bands.

2.1 Star formation history

We begin by exploring the impact of the SFH. In the top panel of Fig. 2, we show the break strength as a function of the *duration* of star formation assuming four different SFHs: an instantaneous burst, an exponentially declining ($\tau = -100$ Myr) SFH, a constant SFH, and an exponentially increasing ($\tau = 100$ Myr) SFH. In each case, we assume no escape of LyC photons (i.e. maximal nebular emission) and no dust attenuation. In the bottom panel, we instead express the *x*-axis in terms of the median (mass-weighted) age of the composite stellar population. The break strength broadly increases with increasing median age and star formation (SF) duration. However, the relationship between age and break strength is sensitive to the shape of the recent SFH. Models with recent star formation (i.e. constant or exponentially increasing) typically have bluer breaks for the same median age or duration. This immediately raises concerns about using the break strength as a diagnostic of the age for star-forming galaxies, at least in isolation.

2.2 Metallicity

In Fig. 2, we also show predictions for three different metallicities: $Z = 0.0001$, 0.001 , and 0.01 . The break has a complex but relatively modest (<0.1 dex) metallicity dependence, which itself is also age dependent. This behaviour is driven, at least in part, by the contribution of nebular continuum emission. For young (<10 Myr) stellar populations, the break is stronger for lower metallicity populations

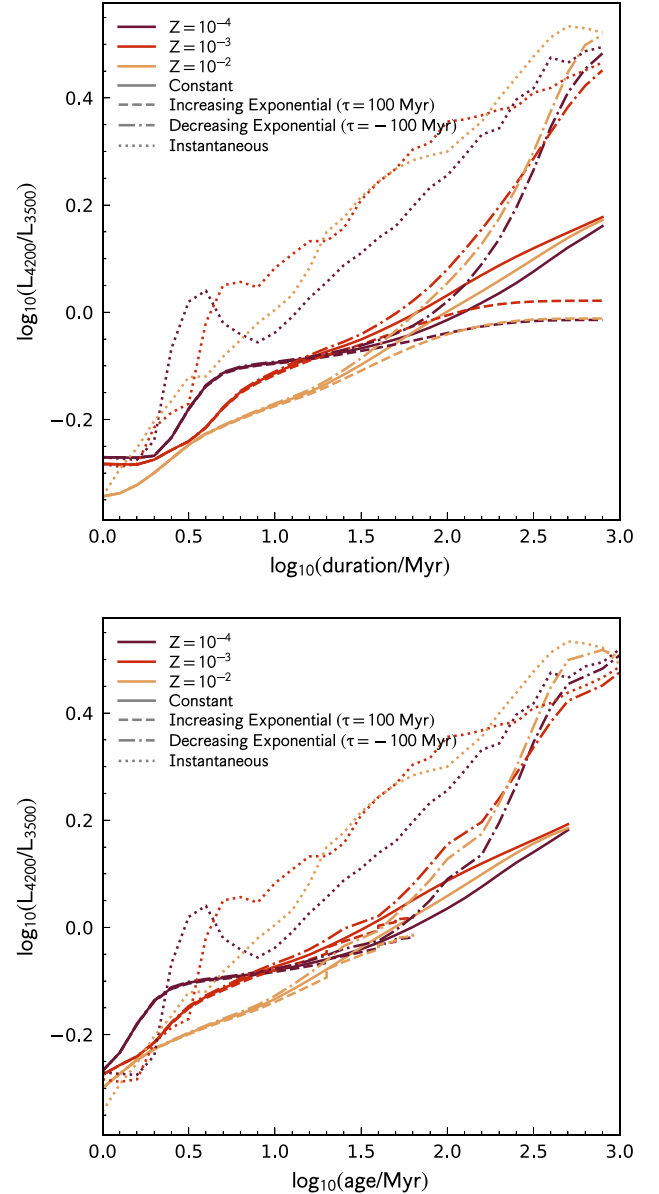


Figure 2. The sensitivity of the break strength to the SFH and metallicity. In the top panel, the *x*-axis is the duration of star formation while in the bottom we use median age instead – for an instantaneous burst the duration and age are equivalent while for a constant SFH the age is half the duration. In both panels, we show results assuming four different SFHs: an instantaneous burst (dotted line), exponentially decreasing star formation (dot-dashed line), constant star formation (solid line), and exponentially increasing star formation (dashed line), and three metallicities: $Z = 10^{-4}$, 10^{-3} , and 10^{-2} . The underlying SEDs assume no escape of Lyman-continuum photons (and hence include nebular emission) but have no dust attenuation/reddening.

but at older ages our intermediate metallicity ($Z = 0.001$) scenario gives rise to slightly stronger break.

2.3 Dust

For a monotonically declining attenuation curve, the break and UV continuum slope will be sensitive to dust. As dust attenuation is increased, the break strength will increase (redden). This is sensitive to the shape of the attenuation curve, but $\tau_V = 1$ dust

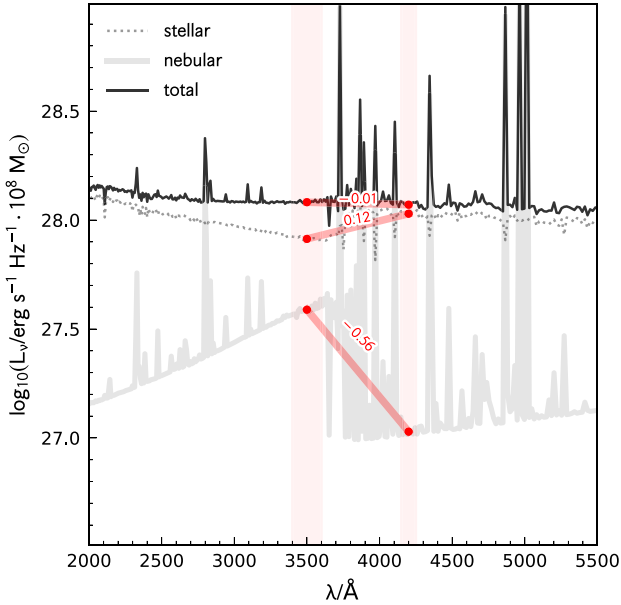


Figure 3. Stellar, nebular, and total SEDs for a dust-free star-forming galaxy (scenario B in Fig. 1) along with the strength of the Balmer break.

attenuation assuming a simple $\tau \propto \lambda^{-1}$ curve yields an increase in the break strength of ≈ 0.1 , comparable to significantly changing the metallicity or a 0.5 dex increase in age (for a instantaneous SFH). Since dust may preferentially affect galaxies with ongoing star formation, the impact would be to narrow the distribution of break strengths.

2.4 Lyman continuum escape fraction

While our definition of the break is chosen to minimize the impact of nebular line emission it is still susceptible to nebular continuum emission. Fig. 3 shows the pure stellar, nebular, and total (composite) spectrum assuming a 100 Myr constant SFH and no dust. While the intrinsic (stellar) spectrum implies a break strength of ≈ 0.1 , the addition of nebular continuum emission, which has a sharp negative (≈ -0.56) break, results in a shallower total break ≈ 0 . Since the contribution of nebular emission is sensitive to the SFH, and in particular the age, the impact of nebular emission is similarly sensitive. Fig. 4 shows the break strength assuming both $f_{\text{esc}} = 0$ and $f_{\text{esc}} = 1$. Nebular emission decreases the break strength by up to 0.2 dex for young ages, with the magnitude dropping for longer durations of star formation. Without a robust estimate of the contribution of nebular emission this suggests it would be extremely challenging to constrain accurate ages for galaxies with ages less than a few tens of Myr.

2.5 Choice of initial mass function

Any model SED is also sensitive to the choice of IMF. This is particularly relevant in the high-redshift Universe where the IMF is very uncertain (e.g. Bate 2023; Steinhardt et al. 2023). While FLARES assumes a constant IMF it is useful to explore the impact of the IMF in the context of our simple toy model. Fig. 5 shows the impact of changing the IMF to four alternatives: an IMF with a shallower ($\alpha_3 = 1$) high-mass slope than Salpeter ($\alpha_3 = 1.35$),

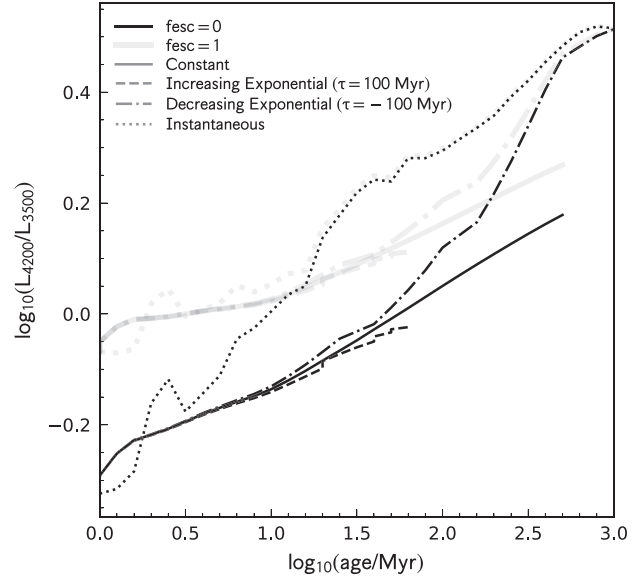


Figure 4. The sensitivity of the break strength to the SFH and Lyman continuum escape fraction assuming $Z = 0.005$.

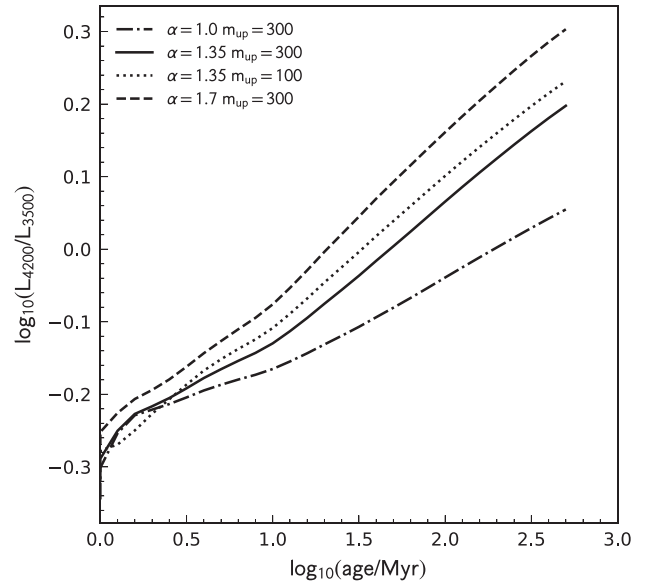


Figure 5. The sensitivity of the break strength to the age and choice of IMF assuming a constant SFH, $f_{\text{esc}} = 0$, $Z = 0.005$, and the BPASS SPS model. Four IMFs are shown: three with different values of the high-mass slope $\alpha_3 \in \{2, 2.35, 2.7\}$ and one with a $\alpha_3 = 2.35$ but lower upper-mass cut-off.

a Salpeter slope,³ a Salpeter slope but with a lower high-mass cut ($m_{\text{up}} = 100 M_{\odot}$), and a steeper slope ($\alpha_3 = 1.7$). In the first case, the result is more high-mass stars present producing a bluer break. For 100 Myr constant star formation duration, the slope is ≈ 0.1 dex smaller. Conversely assuming the steeper IMF results in a redder break, again by ≈ 0.1 dex assuming 100 Myr constant star formation.

³In our previous and following analysis, we assume the Chabrier (2003) IMF. This has a slightly shallower high-mass slope than Salpeter.

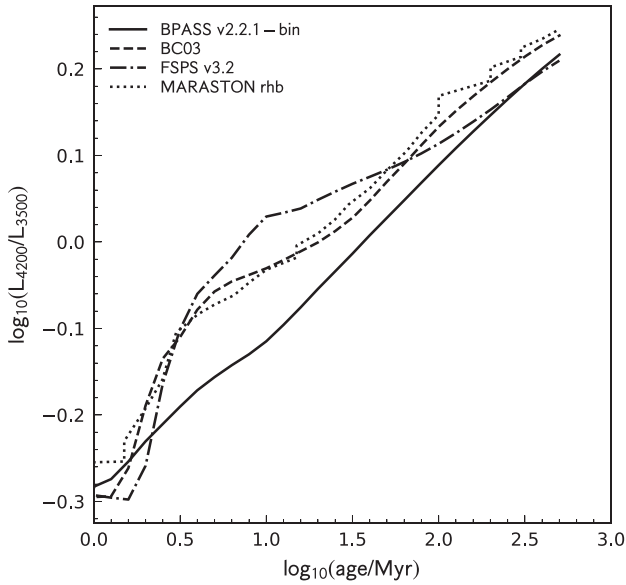


Figure 6. The sensitivity of the break strength to the duration of star formation and choice of SPS model assuming a constant SFH, $f_{\text{esc}} = 0$, and $Z = 0.005$. Four models are considered: BPASS v2.2.1 (our default), Chabrier (2003), the FSPS v3.2 model (Conroy & Gunn 2010), and the Maraston (2005) model.

Reducing the high-mass cut-off has a smaller effect, increasing the break strength by ≈ 0.025 .

2.6 Choice of stellar population synthesis model

The predicted SEDs of stellar populations are also sensitive to the choice of assumed SPS model. Since different SPS models follow stellar evolution and atmospheres in distinct ways, discrepancies between their predictions can occasionally occur (e.g. Byrne & Stanway 2023). To gauge the impact of this choice on our predictions in Fig. 6, we show how the choice of SPS model affects the break strength for three different models in addition to our default BPASS v2.2.1: Bruzual & Charlot (2003), the FLEXIBLE STELLAR POPULATION SYNTHESIS (FSPS; Conroy & Gunn 2010) v3.2 model, and the Maraston (2005) model. For ages 100–1000 Myr these models agree within 0.05 dex, however, for younger ages (~ 10 Myr) we find that BPASS yields significantly bluer (≈ -0.1) breaks.

3 PREDICTIONS FROM THE FIRST LIGHT AND REIONIZATION EPOCH SIMULATIONS

In this study, we make use of FLARES, introduced in Lovell et al. (2021) and Vijayan et al. (2021), and we defer the reader to that paper for full details. In short, FLARES is a suite of hydrodynamical resimulations utilizing the AGNdT9 variant of the EAGLE simulation project (Crain et al. 2015; Schaye et al. 2015) with identical resolution as the reference simulations. The core FLARES suite consists of 40 $14/h$ cMpc radius resimulations of regions selected from a large $(3.2 \text{ cGpc})^3$ dark matter only simulation (Barnes et al. 2017). The selected regions span a large range in overdensity (at $z \approx 4.7$): $\delta + 1 \approx -1 \rightarrow 1$ with overrepresentation of the extremes of the density distribution. This strategy yields galaxies stretching over a larger range in environment, mass, and luminosity than possible with a periodic volume using the same computational

resources. The simulated sample of galaxies, which stretched from $M_* = 10^8 - 10^{11} M_\odot$ (at $z = 5$), is well aligned with all but the most sensitive observations possible with *JWST*.

3.1 Spectral energy distribution modelling

The SED modelling of galaxies in FLARES is described in depth in Vijayan et al. (2021). This builds on the work of Wilkins et al. (2016, 2018) and again we defer the reader to these papers for the details of our methodology. In short, we associate every star particle in the simulation with a pure stellar SED based on its mass, age, and metallicity using v2.2.1 of the BPASS (Stanway & Eldridge 2018) SPS library and assume a Chabrier (2003) IMF. Star particles are then associated with an ionization bound H II region calculated using version 17.01 of the CLOUDY photoionization code (Ferland et al. 2017). Specifically, we use the pure stellar spectrum as the incident radiation field, assuming a spherical geometry and that the total metallicity (but not detailed abundances) of the nebula is identical to the star particle. We also assume a covering fraction of 1 (corresponding to an LyC escape fraction of ≈ 0 for an ionization bound nebula), a hydrogen density of 10^2 cm^{-3} , and a metallicity and age dependent volume-averaged ionization parameter referenced at $t = 1$ Myr and $Z = 0.01$ of $\langle U \rangle = 0.01$. Since this work is focused on a continuum feature specifically chosen to avoid strong line emission, our results are not strongly sensitive to these choices other than LyC escape fraction. Following this, we associate young stellar populations (with age less than 10 Myr, following Charlot & Fall 2000, that birth clouds disperse along these time-scales) with a dusty birth cloud. For every star particle we then determine the line-of-sight surface density of metals Σ_Z in the x, y plane and use this to determine an ISM dust attenuation.

3.2 Trend with mass, luminosity, and redshift

We begin, in Fig. 7, by presenting the Balmer break as a function of the rest-frame far-UV luminosity and stellar mass, respectively, for $z \in \{5, 6, 7, 8, 9, 10\}$. First, this reveals a clear redshift evolution with the break strength increasing by ≈ 0.1 dex from $z = 9$ to $z = 5$. Secondly, the break strength also increases with luminosity, by ≈ 0.1 dex from $L = 10^{28} \rightarrow 10^{30} \text{ erg s}^{-1} \text{ Hz}^{-1}$ at $z = 5$.

To gain further insight into the origin of these trends, in Fig. 8 we present predictions at $z \in \{5, 7, 9\}$ both without dust attenuation alone (dashed line) and without both dust attenuation and nebular emission (i.e. pure stellar breaks, dotted line) compared to our full predictions. The predicted pure stellar breaks evolve with redshift but remain relatively flat with UV luminosity. This agrees with Wilkins et al. (2023a) who found that average ages only evolved subtly with stellar mass. While Wilkins et al. (2023a) showed that there was strong evolution in metallicities the modelling in Section 2.2 demonstrated that for typical SFHs in FLARES the impact of metallicity is limited.

We next explore the impact of nebular continuum emission. As noted in Section 2.4, nebular continuum emission is very blue across the break. Its inclusion in the FLARES modelling shifts predicted average break colours by -0.1 dex, comparable to that found using our simple toy model in Section 2.4. Since the impact of nebular emission is sensitive to the SFH, its impact is also redshift dependent, with galaxies at $z = 9$ more strongly affected than those at $z = 5$ due to the younger ages in the former.

Finally, we explore the impact of dust attenuation. In FLARES dust strongly correlates with intrinsic UV luminosity and stellar mass with a weaker correlation with attenuated UV luminosity (Vijayan et al.

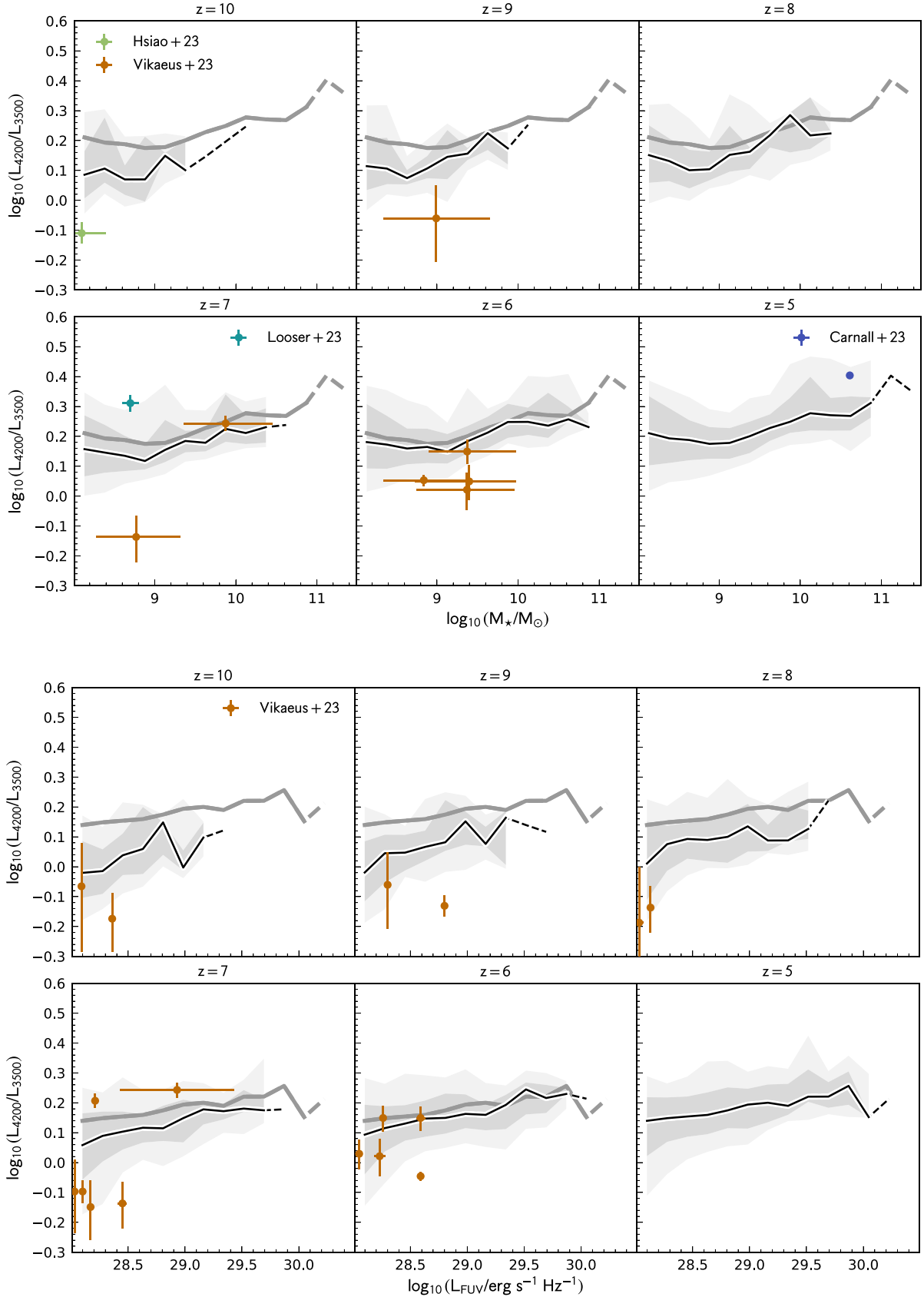


Figure 7. The predicted break strength as a function of the stellar mass (top) and attenuated rest-frame far-UV luminosity (bottom). The outlined dark line shows the median break strength while the two shaded regions show the 15.8–84.2th and 2.2–97.8th percentile ranges. The dashed line shows the median when the number of galaxies in the bin falls below 10. The thick line shows the $z = 5$ median relation to highlight the redshift evolution.

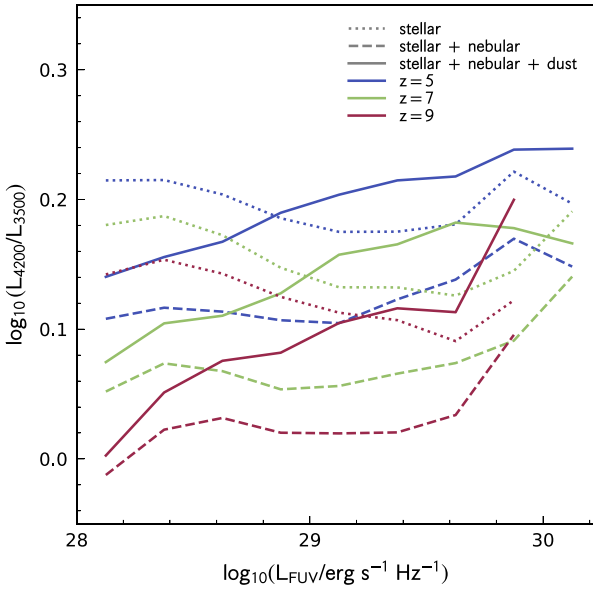


Figure 8. The impact of reprocessing by dust and gas on the median break strength at $z \in \{5, 7, 9\}$. The three sets of lines show predictions assuming no-reprocessing (pure stellar, dotted line), including nebular emission (dashed line), and including both dust and nebular emission (solid lines).

2021, 2022). Since dust attenuation will redden/increase the break, this results in larger predicted breaks in the most luminous galaxies.

In summary, pure stellar break strengths show no correlation with stellar mass and only mild redshift evolution (≈ 0.06 dex, $z = 9 \rightarrow 5$). The inclusion of dust and nebular emission introduces a mass dependence, and strengthens the redshift evolution to ≈ 0.1 dex ($z = 9 \rightarrow 5$).

Our results can be compared to those of Binggeli et al. (2019), who derived Balmer break statistics for $M_* \geq 10^8 M_\odot$ galaxies using the Shimizu et al. (2016), FIRE-2 (Ma et al. 2019), and FirstLight (Ceverino, Klessen & Glover 2018) simulations, albeit for a more limited redshift range ($z = 7-9$) than presented here. For galaxies in the stellar mass range $10^8-10^9 M_\odot$ where Binggeli et al. (2019) have sufficiently many simulated objects at all redshifts, the median Balmer break strength agrees with our predictions to within 0.05 dex at $z = 7, 8$, and 9. However, the FIRE-2 simulations appear to produce a wider distribution of Balmer breaks at each redshift, with more outliers at both $\log_{10}(L_{4200}/L_{3500}) < 0$ and > 0.3 .

3.3 Correlation with physical properties

We now explore how the predicted break strength correlates with key physical properties, including the specific star formation rate, age, dust attenuation, and the rest-frame optical mass-to-light ratio. These relationships are shown in Fig. 9.

First, we find a clear strong ($r = -0.8$) inverse relationship between the break strength and specific star formation rate. Here, the star formation rate is measured on a time-scale of 10 Myr but this assumption does not make a significant difference to the correlation. The correlation with the mass-weighted median age is significantly weaker ($r = 0.37$). For the bulk of galaxies, which have break strengths 0.0–0.2 there is no correlation with age. However, galaxies with the largest breaks tend to be older and vice versa. This lack of correlation reflects two issues: first, the break is also sensitive to the shape of the SFH as well as other properties (nebular

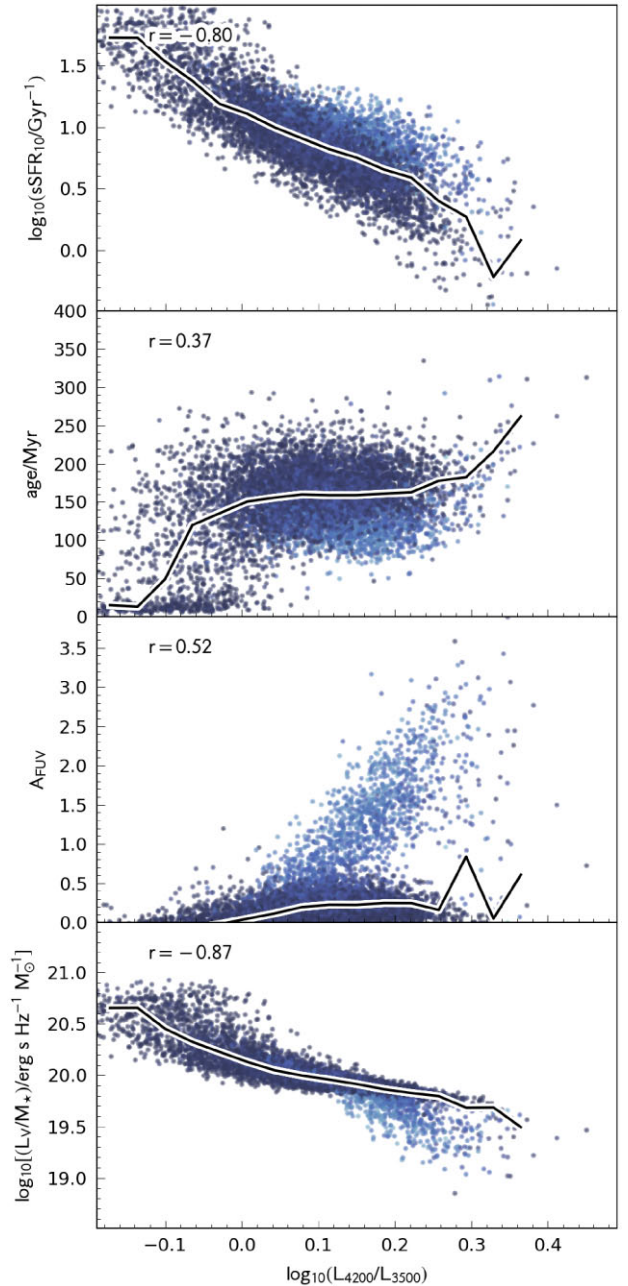


Figure 9. The correlation between the predicted break strength and the specific star formation rate, age, dust attenuation, and V-band mass-to-light ratio for galaxies at $z = 5$. Individual objects are colour-coded by the rest-frame far-UV luminosity. The solid line shows the median value. For each quantity and diagnostic the Pearson correlation coefficient r is also presented.

emission and dust) but also that galaxies in FLARES exhibit a relatively tight range of ages (Wilkins et al. 2023a). The correlation with dust attenuation is more complex. Most galaxies in FLARES (with $L_{\text{FUV}} > 10^{28} \text{ erg s}^{-1} \text{ Hz}^{-1}$) have relatively low attenuation (i.e. $A_{\text{FUV}} < 0.5$, see Vijayan et al. 2021). However, the most intrinsically luminous galaxies in FLARES have increasing attenuation. Since increasing dust attenuation results in larger breaks there is a branch of increasingly dusty galaxies with increasing break strengths. Thus while the dustiest galaxies all have strong breaks, the break itself is not an unambiguous tracer of dust attenuation.

Finally, we also explore the correlation of the break strength with the optical rest-frame V -band mass-to-light ratio. This reveals a strong positive ($r = 0.87$) correlation. Galaxies with the strongest breaks thus tend to have larger stellar masses for the same V -band luminosity. This arises due to sensitivity of both the break and mass-to-light ratio to dust attenuation and the SFH. This reinforces the utility of the break as a diagnostic of stellar masses.

3.4 Comparison with observational direct constraints

At the time of writing, the first spectroscopic measurements of galaxies at $z > 5$ made by *JWST* are emerging, including the detection of the continuum at sufficiently high signal to noise to enable the measurement of the break (e.g. Bunker et al. 2023b; Cameron et al. 2023; Carnall et al. 2023; Hsiao et al. 2023; Looser et al. 2023; Vikaeus et al. 2023). Here, we compare our predictions with those studies that have published (Hsiao et al. 2023; Vikaeus et al. 2023) or made available (via private communication Carnall et al. 2023; Looser et al. 2023) observational constraints on the Balmer break using our metric.

Carnall et al. (2023) present NIRSspec observations of a massive quiescent galaxy (GS-9209) at $z \approx 4.7$ detecting the rest-frame 3000–9000 Å continuum at high signal to noise. This object has a measured break of $\log_{10}(L_{4200}/L_{3500}) \approx 0.4$. This places it at the extreme end, but within the range of our predictions. Since GS-9209 was initially selected as an extremely red, and thus potentially quiescent source (Caputi et al. 2004), its location at the extreme end of our predicted distribution is unsurprising.

Looser et al. (2023) report the discovery of JADES-GS-z7-01-QU, a potentially quiescent galaxy at $z = 7.3$ from *JWST* Advanced Deep Extragalactic Survey (JADES; Bunker et al. 2023a) observations. The break strength of this source is ≈ 0.3 , though this value may be overestimated as the red window coincides with noisy features in spectrum. Nevertheless, like GS-9209, JADES-GS-z7-01-QU exhibits a strong break, lying at the extreme end of the predicted distribution. However, GS-9209 and JADES-GS-z7-01-QU are unlikely to be representative of the wider population of high-redshift galaxies and it is therefore premature to make any inferences about the model.

Hsiao et al. (2023) present NIRSspec observations of MACS0647-JD, the triply lensed source found to be at $z \approx 10.2$. Unlike GS-9209 and JADES-GS-z7-01-QU, MACS0647-JD is not selected to be quiescent and consistent with being typical star-forming galaxies. For MACS0647-JD, Hsiao et al. (2023) report $M_* = 10^{8.1} M_{\odot}$ and $\log_{10}(L_{4200}/L_{3500}) \approx -0.11$. In contrast to GS-9209 and JADES-GS-z7-01-QU, MACS0647-JD lies *below* our predictions and in fact just outside the central 95 per cent range. However, since this is a single object, and only marginally inconsistent, it is again impossible to draw firm inferences about FLARES.

Vikaeus et al. (2023) present an analysis of NIRSspec observations of 23 spectroscopically confirmed galaxies at $6 < z < 12$ using public *JWST*/NIRSspec data from the cycle 1 GO 1433 and GO 2282 programs in addition to public spectroscopic data from JADES. Within the uncertainties, this sample broadly brackets the FLARES predictions.

These and other recent observations demonstrate *JWST*'s extraordinary potential to probe the rest-frame SEDs of galaxies in the distant Universe. With many more spectroscopic observations planned over the coming cycles *JWST* will soon place strong constraints on the break providing a new tool to test galaxy formation models like FLARES.

4 CONCLUSIONS

In this work, we have explored predictions for the strength of the Balmer/4000 Å break predicted by the FLARES project. Prior to this we explored how various parameters (star formation and metal enrichment history, dust attenuation, and LyC escape fraction) and assumptions (SPS model and IMF) affect the break using a simple toy model.

(i) Our toy modelling reveals that while the strength of the break is sensitive to the age (or duration) it is also sensitive to the shape of the SFH, metallicity, dust attenuation, and LyC (ionizing photon) escape fraction. The impact of both the shape of the SFH and LyC escape fraction can shift the break strength by ~ 0.2 dex. Metallicity is less important but can drive differences of up to ≈ 0.1 dex.

(ii) This modelling also reveals the sensitivity of the break strength to the choice of SPS model and IMF. The choice of the SPS model can change the break strength by > 0.1 dex, with the largest impact at star formation durations of ~ 10 Myr. The impact of the IMF grows with increasing duration of star formation, increasing to up to 0.3 dex for $\alpha_3 = 1.0 \rightarrow 1.7$ after a few hundred Myr of star formation.

(iii) FLARES predicts a flat relationship between attenuated far-UV luminosity or mass and the *intrinsic* break strength. Including nebular emission reduces the break strength by ≈ 0.1 dex while including dust increases the break strength and produces a slightly positive relationship between attenuated far-UV luminosity or mass and the break strength. The average break strength also varies with redshift, increasing by around 0.1 dex from $z = 9 \rightarrow 5$.

(iv) The break strength inversely correlates with both specific star formation rate and optical mass-to-light ratio but only weakly correlates with the average age, and then only at the extremes.

(v) At the time of writing there have only been small number of robust spectroscopic measurements of the break strength, and these are biased samples. However, this will soon change, ultimately providing a new constraint on galaxy formation models in the distant Universe.

ACKNOWLEDGEMENTS

We would like to thank Adam Carnall and Tobias Looser for providing measurements of the break using our chosen diagnostic. We thank the EAGLE team for their efforts in developing the EAGLE simulation code. This work used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The equipment was funded by BEIS capital funding via STFC capital grants ST/K00042X/1, ST/P002293/1, ST/R002371/1, and ST/S002502/1, Durham University, and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure. SMW and WJR thank STFC for support through ST/X001040/1. CCL acknowledges support from a Dennis Sciama fellowship funded by the University of Portsmouth for the Institute of Cosmology and Gravitation. DI acknowledges support by the European Research Council via ERC Consolidator Grant KETJU (no. 818930) and the CSC – IT Center for Science, Finland. APV acknowledges support from the Carlsberg Foundation (grant no. CF20-0534). The Cosmic Dawn Center (DAWN) is funded by the Danish National Research Foundation under grant no. 140. AV and EZ acknowledge funding from the Swedish National Space Agency. EZ also acknowledges grant 2022–03804 from the Swedish Research Council. LTCS is supported by an STFC studentship. We also wish to acknowledge the following open source software packages used in the analysis: SCIPY (Virtanen et al. 2020), ASTROPY (Robitaille et al. 2013),

MATPLOTLIB (Hunter 2007), and WEBPLOTDIGITIZER (Rohatgi 2020). We list here the roles and contributions of the authors according to the Contributor Roles Taxonomy (CRediT)⁴: **Stephen M. Wilkins**: Conceptualization, Data curation, Methodology, Investigation, Formal Analysis, Visualization, Writing – original draft. **Christopher C. Lovell**, **Aswin P. Vijayan**: Data curation, Methodology, Writing – review & editing. **Joseph Caruana**, **Katherine Chworowsky**, **Chris Conscience**, **Steven Finkelstein**, **Dimitrios Irodotou**, **William Roper**, **Louise T. C. Seeyave**, **Elizabeth Stanway**, **Anton Vikaeus**, **Erik Zackrisson**: Writing – review & editing.

DATA AVAILABILITY

The data associated with the paper will be made publicly available at <https://flaresimulations.github.io> on the acceptance of the manuscript.

REFERENCES

- Adams N. J. et al., 2023, *ApJ*, preprint (arXiv:2304.13721)
- Arrabal Haro P. et al., 2023, *ApJ*, 951, L22
- Barnes D. J. et al., 2017, *MNRAS*, 471, 1088
- Bate M. R., 2023, *MNRAS*, 519, 688
- Binggeli C. et al., 2019, *MNRAS*, 489, 3827
- Boyett K. et al., 2023, *ApJL*, 940, L52
- Bruzual A. G., 1983, *ApJ*, 273, 105
- Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
- Bunker A. J. et al., 2023a, *A&A*, preprint (arXiv:2306.02467)
- Bunker A. J. et al., 2023b, *A&A*, 677, A88
- Byrne C. M., Stanway E. R., 2023, *MNRAS*, 521, 4995
- Cameron A. J. et al., 2023, *A&A*, 677, A115
- Caputi K. I., Dunlop J. S., McLure R. J., Roche N. D., 2004, *MNRAS*, 353, 30
- Carnall A. C. et al., 2023, *Nature*, 619, 716
- Ceverino D., Klessen R. S., Glover S. C. O., 2018, *MNRAS*, 480, 4842
- Chabrier G., 2003, *PASP*, 115, 763
- Charlot S., Fall S. M., 2000, *ApJ*, 539, 718
- Conroy C., Gunn J. E., 2010, *ApJ*, 712, 833
- Crain R. A. et al., 2015, *MNRAS*, 450, 1937
- De Barros S., Oesch P. A., Labbé I., Stefanon M., González V., Smit R., Bouwens R. J., Illingworth G. D., 2019, *MNRAS*, 489, 2355
- Endsley R., Stark D. P., Chevallard J., Charlot S., 2021, *MNRAS*, 500, 5229
- Ferland G. J. et al., 2017, *Rev. Mex. Astron. Astrofis.*, 53, 385
- Fujimoto S. et al., 2023, *ApJ*, 949, L25
- Hashimoto T. et al., 2018, *Nature*, 557, 392
- Hsiao T. Y.-Y. et al., 2023, *ApJ*, preprint (arXiv:2305.03042)
- Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90
- Laporte N., Meyer R. A., Ellis R. S., Robertson B. E., Chisholm J., Roberts-Borsani G. W., 2021, *MNRAS*, 505, 3336
- Looser T. J. et al., 2023, preprint (arXiv:2302.14155)
- Lovell C. C., Vijayan A. P., Thomas P. A., Wilkins S. M., Barnes D. J., Irodotou D., Roper W., 2021, *MNRAS*, 500, 2127
- Ma X. et al., 2019, *MNRAS*, 487, 1844
- Maraston C., 2005, *MNRAS*, 362, 799
- Matthee J., Mackenzie R., Simcoe R. A., Kashino D., Lilly S. J., Bordoloi R., Eilers A.-C., 2023, *ApJ*, 950, 67
- Poggianti B. M., Barbaro G., 1997, *A&A*, 325, 1025
- Roberts-Borsani G. W., Ellis R. S., Laporte N., 2020, *MNRAS*, 497, 3440
- Robitaille T. P. et al., 2013, *A&A*, 558, A33
- Rohatgi A., 2020, Webplotdigitizer, Version 4.4, <https://automeris.io/WebPlotDigitizer>
- Roper W. J., Lovell C. C., Vijayan A. P., Marshall M. A., Irodotou D., Kuusisto J. K., Thomas P. A., Wilkins S. M., 2022, *MNRAS*, 514, 1921
- Schaerer D., Marques-Chaves R., Barrufet L., Oesch P., Izotov Y. I., Naidu R., Guseva N. G., Brammer G., 2022, *A&A*, 665, L4
- Schaye J. et al., 2015, *MNRAS*, 446, 521
- Scoville N., Faisst A., Capak P., Kakazu Y., Li G., Steinhardt C., 2015, *ApJ*, 800, 108
- Shimizu I., Inoue A. K., Okamoto T., Yoshida N., 2016, *MNRAS*, 461, 3563
- Stanway E. R., Eldridge J. J., 2018, *MNRAS*, 479, 75
- Stark D. P., Schenker M. A., Ellis R., Robertson B., McLure R., Dunlop J., 2013, *ApJ*, 763, 129
- Steinhardt C. L., Kokorev V., Rusakov V., Garcia E., Sneppen A., 2023, *ApJ*, 951, L40
- Stiavelli M. et al., 2023, *ApJ*, 957, L18
- Tacchella S. et al., 2022, *ApJ*, 927, 170
- Trump J. R. et al., 2023, *ApJ*, 945, 35
- Trussler J. A. A. et al., 2023, *MNRAS*, 523, 3423
- Vijayan A. P., Lovell C. C., Wilkins S. M., Thomas P. A., Barnes D. J., Irodotou D., Kuusisto J., Roper W. J., 2021, *MNRAS*, 501, 3289
- Vijayan A. P. et al., 2022, *MNRAS*, 511, 4999
- Vikaeus A. et al., 2023, preprint (arXiv:2309.02504)
- Virtanen P. et al., 2020, *Nat. Methods*, 17, 261
- Wilkins S. M. et al., 2013, *MNRAS*, 435, 2885
- Wilkins S. M., Feng Y., Di-Matteo T., Croft R., Stanway E. R., Bunker A., Waters D., Lovell C., 2016, *MNRAS*, 460, 3170
- Wilkins S. M., Feng Y., Di-Matteo T., Croft R., Lovell C. C., Thomas P., 2018, *MNRAS*, 473, 5363
- Wilkins S. M. et al., 2020, *MNRAS*, 493, 6079
- Wilkins S. M. et al., 2022, *MNRAS*, 517, 3227
- Wilkins S. M. et al., 2023a, *MNRAS*, 518, 3935
- Wilkins S. M. et al., 2023b, *MNRAS*, 522, 4014
- Worthey G., 1994, *ApJS*, 95, 107

⁴<https://credit.niso.org/>

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.