The Astrophysical Journal, 967:172 (8pp), 2024 June 1 https://doi.org/10.3847/1538-4357/ad3afb

The Highest-redshift Balmer Breaks as a Test of ΛCDM

Charles L. Steinhardt[1,2], Albert Sneppen[1,2], Thorbjørn Clausen[1,2], Harley Katz[3], Martin P. Rey[3], and Jonas Stahlschmidt[1,2]
[1] Cosmic Dawn Center (DAWN), Denmark; steinhardt@nbi.ku.dk
[2] Niels Bohr Institute, University of Copenhagen, Jagtvej 120, DK-2100 Copenhagen Ø, Denmark
[3] Sub-department of Astrophysics, University of Oxford, DWB, Keble Road, Oxford OX1 3RH, UK
Received 2023 June 5; revised 2024 March 27; accepted 2024 March 28; published 2024 May 30

Abstract

Recent studies have reported tension between the presence of luminous, high-redshift galaxies and the halo mass functions predicted by standard cosmology. Here, an improved test is proposed using the presence of high-redshift Balmer breaks to probe the formation of early 10^4–$10^5\,M_\odot$ baryonic minihalos. Unlike previous tests, this does not depend upon the mass-to-light ratio and has only a slight dependence upon the metallicity, stellar initial mass function, and star formation history, which are all weakly constrained at high redshift. We show that the strongest Balmer breaks allowed at $z = 9$ using the simplest ΛCDM cosmological model would allow a D_{4000} as high as 1.26 under idealized circumstances and $D_{4000} \leqslant 1.14$ including realistic feedback models. Since current photometric template fitting to JWST sources infers the existence of stronger Balmer breaks out to $z \gtrsim 11$, upcoming spectroscopic follow-up will either demonstrate those templates are invalid at high redshift or imply new physics beyond "vanilla" ΛCDM.

Unified Astronomy Thesaurus concepts: Galaxy spectroscopy (2171); Galaxies (573); Quenched galaxies (2016)

1. Introduction

The standard cosmological paradigm and hierarchical merging predict the timing with which small, primordial density fluctuations under the influence of gravity produce massive, compact halos. Subsequently, baryons within those halos form stars and other structures recognizable as galaxies. Remarkably, recent studies have reported galaxies that are so massive, at such high redshifts, that they appear to have formed even before their halos could have had time to collapse (Steinhardt et al. 2016; Labbé et al. 2023) under the current standard model comprised of a cosmological constant dark energy and cold, collisionless dark matter (here termed "vanilla" ΛCDM). As a result, the presence of massive, early galaxies has been interpreted as evidence for new physics beyond vanilla ΛCDM (Steinhardt et al. 2016; Behroozi & Silk 2018; Boylan-Kolchin 2023).

However, this interpretation is challenging for three key reasons. First, observations directly find *luminous* galaxies at high redshift rather than *massive* ones. Stellar masses inferred from photometry are thus highly sensitive to assumptions on the mass-to-light ratio, which depend on the star formation history (SFH; Conroy 2013) and assumed stellar initial mass function (IMF; Steinhardt et al. 2023). Second, the possibility of a varying stellar baryon fraction (Finkelstein et al. 2015) will affect the inferred halo mass that is being compared to the measured stellar mass. Finally, once a halo has virialized, additional time is required for the baryons within that halo to form stars. Thus, the observed stellar masses should not be compared against halo formation at the observed redshift, but rather at some higher redshift. Determining which redshift is most relevant again depends upon the star formation history, which is typically one of the least-constrained properties in photometric template fitting, and even more so at high redshift (Laigle et al. 2016; Davidzon et al. 2017; Iyer & Gawiser 2017; Weaver et al. 2022). These significant systematic uncertainties makes it difficult to provide robust evidence for new physics beyond ΛCDM.

In this work, we propose an alternative, more robust test relying on the detection of the specific Balmer-break spectral feature at 3646 Å. The Balmer break comes from hydrogen atoms that have been excited into an $n = 2$ quantum state, and therefore is strongest at specific combinations of temperature and density such as those commonly found in the photospheres of A stars. Thus, the preconditions for producing a strong Balmer break are more complex than merely the creation of stellar mass. Rather, they reflect the characteristic timescale of stellar mass assembly, as younger stars have to evolve off their main sequence for A-type stars to dominate the galaxy's light. In fact, Balmer breaks are routinely used as indicators of old stellar populations in low-redshift surveys (e.g., Mobasher et al. 2005; Dunlop 2013; Shahidi et al. 2020), and traditionally interpreted to track a period of star formation and mass buildup and a subsequent quiescent phase lasting a few hundred Myr. Given that such timescales approach the Hubble time at $z \geqslant 8$, we argue here that their spectroscopic detection at high redshift would provide a significantly more robust test of structure formation within vanilla ΛCDM, whose interpretation does not require assumptions about the IMF and SFHs.

The quest for such high-redshift ($z \geqslant 8$) Balmer breaks is long-standing (e.g., Hashimoto et al. 2018; Laporte et al. 2021, 2023) but is rapidly expanding, with first photometric results from JWST implying $z = 9$–11 Balmer breaks (e.g., Adams et al. 2023; Atek et al. 2023; Furtak et al. 2023) that might hint toward new physics. So far, these candidates have lacked the spectroscopic confirmation of the Balmer break, which is key to distinguish them from a Lyman break (Hovis-Afflerbach et al. 2021) or from nebular emission producing

photometric excesses mimicking a break[4] (Roberts-Borsani et al. 2020; Carnall et al. 2023). Large-scale spectroscopic programs with JWST are nonetheless underway (e.g., JADES, A. Bunker et al. 2024, in preparation; CEERS, Finkelstein et al. 2023; GLASS, Treu et al. 2022), making a theoretical quantification of the earliest Balmer break expected from vanilla ΛCDM particularly timely.

Since a Balmer break requires a stellar population that is a few hundred Myr old, the timing can be broken up into three required steps, which we systematically explore in this paper. In Section 2, the formation time is calculated for the first virialized clumps capable of turning into stars. Once star formation begins, the minimum time necessary to turn off star formation, if only for a few hundred Myr, is estimated in Section 3. Finally, the time that must elapse between this turnoff and the appearance of a recognizable Balmer break is calculated in Section 4. In Section 5, these components are put together and it is shown that many current galactic spectral energy distribution models fit to existing photometry predict Balmer breaks, which, if verified spectroscopically, would already point toward new physics.

2. Forming Stars in the First Minihalos

The first stage in producing a Balmer break is to assemble a deep enough gravitational potential well to trigger radiative gas cooling and produce the compact clumps of baryons that will subsequently produce stars. Here, two approaches are taken and compared to estimate this formation time. First, an analytical estimate (as originally presented in Haiman & Loeb 1997) is updated to current cosmological parameters. There are major differences between the concordance cosmological parameters when the original calculation was performed in 1997 and at present, and the effects of each difference is evaluated. Second, the result is then compared to the formation time of the first Population III stars in numerical simulations.

There are significant differences between the assembly of baryons and dark matter. Although cold dark matter structures form via hierarchical merging, baryonic collapse is complicated by diffusion damping (Silk 1968) and gas pressure, which greatly reduce anisotropies on small scales. Relative velocities between the baryon and dark matter fluids can lead to an additional delay in baryon clumping (Tseliakhovich & Hirata 2010), which is neglected in this analysis. Including this effect would delay baryonic structure formation, star formation, and quenching even further, so the bound presented here is overly conservative.

As a result, the smallest baryonic fluctuations collapse slower than larger structures, and will not allow for accelerated construction via hierarchical merging. Instead, there is a characteristic mass scale corresponding to the first clumps of baryons to complete their collapse, likely between 10^4 and $10^5\,M_\odot$ ("minihalos"; Haiman & Loeb 1997). Given these fairly large gas masses, Population III stars likely form soon after the collapse of minihalos (see also, e.g., Nakamura & Umemura 2001; Bromm 2013 for reviews). The formation time of minihalos thus provides a lower bound for the age of the Universe at the birth of the first stars.

The Haiman & Loeb (1997) collapse time calculation assumed a cosmology with $(\Omega_b, \Omega_m, \Omega_\Lambda, h, \sigma_8) = (0.05, 1, 0, 0.50, 0.67)$. Thus, conversion to the Planck Collaboration et al. (2020) parameters of $(\Omega_b, \Omega_m, \Omega_\Lambda, h, \sigma_8) = (0.049, 0.315, 0.685, 0.674, 0.811)$ produces three significant adjustments:

1. Planck found an approximately 20% increase in σ_8, the linear amplitude of oscillation on scales $8\,h^{-1}$ Mpc, over the 1997 parameters. The collapse redshift is approximately linear in σ_8 (Haiman & Loeb 1997), so this will produce a collapse at 20% higher redshift than in the previous result.
2. The spherical collapse time is often described as proportional to $1/H$, where H is the Hubble parameter (Kolb & Turner 1990). However, this arises from dependence on the mass of the overdensity (or, equivalently, the matter density $\rho_m \propto \Omega_m H_0^2$), since the collapse time $t_c \propto M^{-1/2}$. Thus, although Planck finds $h = 0.674 \pm 0.005$,[5] larger than the 1997 value of $h = 0.5$, $\rho_m \propto \Omega_m h^2$ decreases since $\Omega_m = 0.315$ rather than 1 as was assumed prior to the discovery of dark energy. In total, $H(z \approx 1100)$ decreases by 27%, producing a corresponding increase in collapse time.
3. The same collapse time corresponds to a higher redshift when mapping age and redshift with the current concordance model, compared to the 1997 values.

Taking these effects together, the Haiman & Loeb (1997) calculations can be updated to modern cosmology. Since the first baryonic structures to collapse have masses of 10^4–$10^5\,M_\odot$, between $\sim 10^4$ and 10^6 such clumps comprise a typical protogalaxy. Thus, the most overdense of these will come from $\sim 4\sigma$–5σ initial fluctuations, and closer to 4σ for the earliest, and presumably smallest, protogalaxies (Miralda-Escudé 2003).[6] Therefore, early Population III star formation is estimated to start at $z \sim 32$ under vanilla ΛCDM structure formation (Figure 1). Slightly earlier than the $z \sim 27$ collapse time estimated from the 1997 cosmology.

This estimate of the earliest possible onset of star formation is in line with more realistic and modern cosmological simulations, with few reporting Population III star formation above $z \approx 30$–40 (see, e.g., Jaacks et al. 2018; Liu & Bromm 2020 for recent works, and Klessen & Glover 2023 for a review).

3. Suppressing Star Formation for Extended Periods

A Balmer break requires the light of O and B stars to disappear and be replaced by A-type stars. Thus it is necessary to not only initially form stars but also to avoid subsequently forming newer, younger stars that would outshine the older population (or for differential dust attenuation to obscure those newer, younger stars but not the older stellar population; e.g., Katz et al. 2019). The precise stellar-evolution timescales associated with this aging process are investigated in Section 4, but they are robustly $\geqslant 200$ Myr. Due to the lack of supermassive black holes at high redshift, the prime candidate to

[4] A handful, including Hashimoto et al. (2018), have spectroscopic redshift confirmation but none has rest-frame optical spectroscopy that confirms the Balmer break. An additional complication is that a photometric Balmer break is measured as a 4000 Å break, which can also be sensitive to dust and other physical processes that contribute to a measure such as D_{4000} but would be excluded by a more narrowly tailored metric that requires spectroscopy, as described in Section 4.

[5] The matter density would be slightly (but for this analysis, negligibly) larger than calculated here using h from local Cepheid and supernova observations (Riess et al. 2022).

[6] The first structure to collapse in a large observed field will naturally occur earlier, but the time for that galaxy to quench (Section 3) will be comparable.

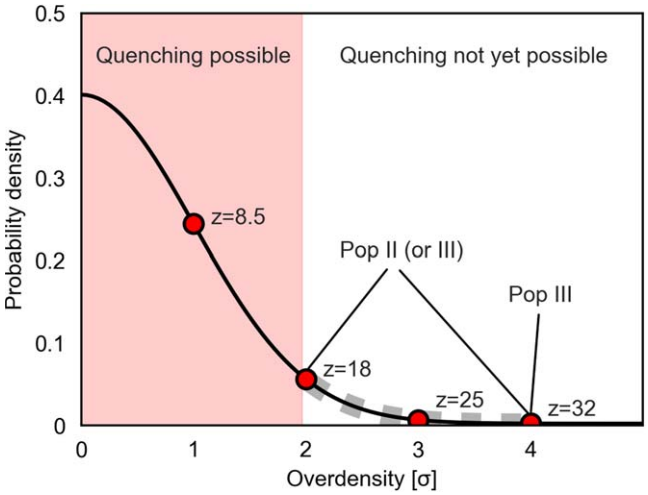

Figure 1. The distribution of overdensities that will form the first 10^4–10^5 $M_\odot$ baryonic minihalos is well approximated as Gaussian. The rare, most overdense regions (gray dashed) must form first Population III and later Population II stars. Subsequent minihalos (red) can be heated by previous star formation and might not necessarily form stars at collapse. Here, both semianalytical arguments (see accompanying text) and simulations find the boundary to lie at $\sim 2\sigma$ overdensities collapsing around $z \sim 18$. If a protogalaxy is discovered with a stellar population which quenched at $z > 18$, it would require new physics beyond "vanilla" ΛCDM.

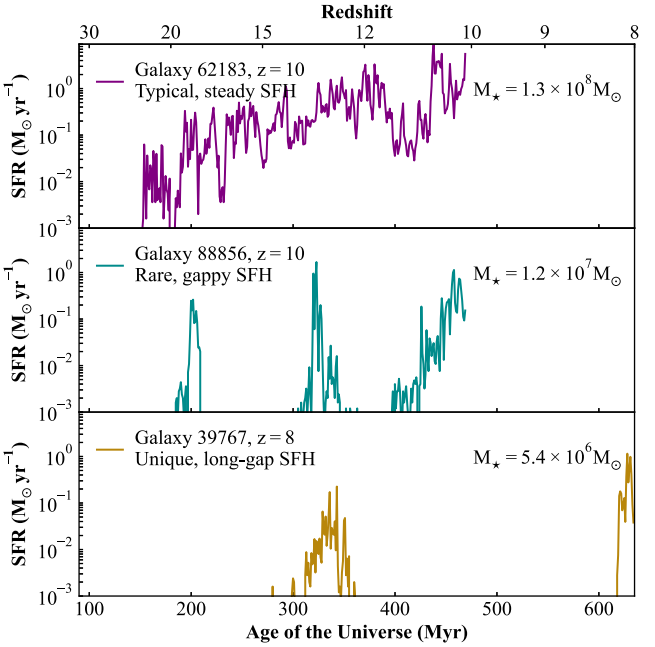

Figure 2. Sample star formation histories of high-redshift galaxies from data products of the SPHINX simulation (Rosdahl et al. 2018, 2022; Choustikov et al. 2024). Rising, continuous star formation histories are the most common occurrence (top) at $z = 8$, 9, and 10, but rare examples show extended gaps induced by stellar feedback (middle panels) in lower-mass objects. One, unique case of a very low-mass dwarf showcasing a $\geqslant$250 Myr gap is found (bottom), comparable to the stellar-evolution timescales necessary to produce strong Balmer breaks (Figure 3).

suppress star formation is energy input from stars, first through their radiation and subsequently via supernovae explosions.

Given the large uncertainties associated with modeling star formation and feedback at $z \geqslant 10$, here a back-of-the-envelope estimate is provided for generating enough stellar mass and feedback energy to quench star formation for several hundred million years. Then star formation histories are quantitatively examined with the state-of-the-art radiation hydrodynamical SPHINX simulation (Rosdahl et al. 2018, 2022).

In order to quench star formation, enough gas needs to be transformed into stars for them to significantly heat their surroundings and balance the very quick cooling times of the early Universe (e.g., White & Rees 1978; Dekel & Silk 1986). Most of the first minihalos form in isolation, in density peaks, and the short main-sequence lifetimes associated with massive Population III stars make them inefficient as a sustained heating source. Consequently, one can speculate that efficient quenching requires the larger overdensities of the Universe to collapse and provide the necessary stellar feedback (via a Population II stellar population). A 2σ overdensity collapses at $z = 18$ (Figure 1), which as below is a good estimate for when gaps can begin appearing in the star formation histories of high-redshift dwarf galaxies.

To estimate this more quantitatively, the star formation histories in Version 1 of the SPHINX Public Data Release (SPDR1) are used. SPHINX[20] models the cosmological formation of the first galaxies in an (20 comoving Mpc)^3 "average" volume (i.e., with particularly overdense regions favoring earlier collapse) with enough numerical resolution (10 pc spatial resolution and 2×10^5 $M_\odot$ dark matter particle mass) to start resolving the interstellar medium. The formation history is extracted from all stars within galaxies at $z = 10$, 9, and 8 that have SFR $\geqslant 0.3\, M_\odot\, \mathrm{yr}^{-1}$ to mimic a flux-limited survey (see Choustikov et al. 2024 for further details), and systematically searched for the earliest gaps in star formation greater than 100 and 250 Myr.

Selected examples of such star formation histories are shown in Figure 2. Nearly all star formation histories are continuous and rising (e.g., top panel), with several examples exhibiting star formation at $z \geqslant 20$ and validating the analytical timescales from Section 2. Four galaxies are found at $z = 10$ with $\geqslant$100 Myr gaps (middle panel), but none with longer. The bottom panel shows the only example of a low-mass, $z = 8$ galaxy exhibiting a $\geqslant$250 Myr gap, approaching the timescale necessary to create a strong Balmer break in the absence of dust (Section 4).

The lack of $\geqslant$250 Myr gaps at very high redshift ($z \geqslant 12$) in the SPHINX data suggests that of the star-forming populations that are likely observable with JWST, very few are expected to exhibit strong Balmer breaks (see also Binggeli et al. 2019). Nonetheless, a key learning from inspecting these star formation histories is that extended gaps due to stellar feedback can be generated as early as $z = 18$, even if the details of their lengths depend on the specific model and astrophysical conditions. In the remaining of this paper, it is thus assumed that quenching star formation for 300 Myr at $z \approx 18$ is within the realm of uncertainties of current models, and the following section quantifies the time needed for stars formed at that time to evolve off the main sequence and create strong Balmer breaks.

4. From Quenching to Production of the Balmer Break

Finally, once a galaxy has quenched, enough time must elapse for the most massive stars to evolve off the main sequence so that the observed spectral energy distribution is dominated by stars with cool enough photospheres to produce a Balmer break. To gain a clean intuition of the timescales involved and how they relate to the characteristics of the

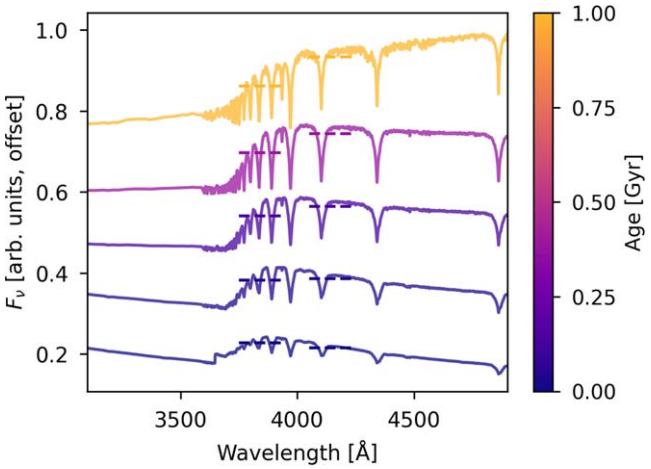

Figure 3. Spectral energy distributions around 4000 Å for simulated galaxies composed of a single stellar population of Population II stars with ages 25, 63, 158, 398, and 1000 Myr. The depth of the Balmer break grows rapidly with increased age to ∼300 Myr and more gradually after that. This strength of this feature is often described in terms of D_{4000} (Bruzual 1983; Poggianti & Barbaro 1997), the slope between the two regions indicated as dashed lines. D_{4000} is negative for very young populations and grows as the population ages toward 1 Gyr.

Balmer break, a grid of single stellar population (SSP) models was run using the Flexible Stellar Population Synthesis (FSPS; Conroy & Wechsler 2009) library with Planck Collaboration et al. (2020) parameters. Using an SSP is conservative as any more complex star formation history will take longer to produce a Balmer break. Here, a fiducial model with $[Z/H] = -1.7$, $A_V = 0.1$, and an IMF corresponding to a gas temperature of 60 K is chosen. As described in the remainder of this section, a comparison with previous work shows that the primary results here are not strongly sensitive either to the specific parameters or to the choice of the FSPS library, as a similar bound exists in studies using different methodology (Bruzual 1983; Poggianti & Barbaro 1997).

As the age of the stellar population increases, the spectral energy distribution begins to exhibit a visible Balmer break (Figure 3). For the remainder of this work,

$$D_{4000} = \frac{\int_{4050\mathring{A}}^{4250\mathring{A}} F_\nu \, d\lambda}{\int_{3750\mathring{A}}^{3950\mathring{A}} F_\nu \, d\lambda} \quad (1)$$

is used a standard measure of the depth of the break (Bruzual 1983; Poggianti & Barbaro 1997).

It should be emphasized that two other common definitions of the depth of the break exist. For the fiducial 300 Myr SSP used here, the Bruzual (1983) $F_{\nu,4150}/F_{\nu,3850}$ is 1.24. For comparison, an alternative definition of $D_{4000} = F_{\nu,4200}/F_{\nu,3500}$ (Binggeli et al. 2019) as used in BPASSv2.2.1 (Stanway & Eldridge 2018) for the same population yields $D_{4000} = 2.43$, and the narrowband $D_n(4000) = 1.11$ (Balogh et al. 1999), so it is critical to note which definition is used when comparing against observations.

The evolution of this measure as a function of age in is shown in Figure 4. For very young stellar populations, $D_{4000} < 1$ because of the blue continuum produced by the most massive stars. As the stellar population ages and young O and B stars evolve off the main sequence quickly, D_{4000} grows rapidly. This growth then becomes more gradual with time, as the last lower-mass B stars already exhibit weaker Balmer breaks and are slowly evolving off the main sequence. Galaxies in this latter phase are typically described as exhibiting Balmer breaks.

In order to model populations typical of those that could produce the earliest Balmer breaks, the following assumptions are tested:

1. As metallicity decreases, stellar photospheres increase in temperature at fixed mass. As a result, the Balmer break is produced by lower-mass stars than at higher metallicity, so the stellar population must be older for these lower-mass stars to dominate the spectrum and produce a Balmer break. Since the time required to produce a Balmer break is longer toward lower metallicity (Figure 4, left), here $[Z/H] = -1.7$ is chosen in an attempt to provide a conservative lower bound on the time required to produce a Balmer break. Although there is some justification for such a high Population II metallicity based on Galactic Population II stars (Baade 1944; Rix et al. 2022), this value is almost certainly an overestimate of the metallicity of very early Population II stars.
2. Although dust can contribute to an earlier production of a 4000 Å break, there is likely not enough time for dust to be produced in significant quantities by $z \sim 20$ (Gall et al. 2011; Ferrara et al. 2016). Here, a maximum extinction of $A_V = 0.1$ mag is taken as a conservative upper bound on high-redshift dust. Photometric template fitting finds that $A_V < 0.1$ is more likely even at $z = 10$ (Steinhardt et al. 2023). Regardless, the timescale for producing a Balmer break is relatively insensitive to A_V (Figure 4, middle) as long as the dust is evenly distributed.
3. Since the cosmic microwave background (CMB) temperature at $z = 20$ is slightly under 60 K, a bottom-lighter stellar IMF than in the Milky Way is very likely required (Low & Lynden-Bell 1976; Larson 1985; Sneppen et al. 2022; Steinhardt et al. 2023). A reduction in metallicity, as is likely at very high redshift, will act in the same direction. Here, the 60 K IMF prescription from Jermyn et al. (2018) is assumed for these first Population II stars. Nevertheless, the timing of Balmer-break formation is only very weakly dependent on the IMF, since the most massive stars must nearly all die regardless of their exact relative fractions of the stellar population (Figure 4, center).

From this systematic exploration, the largest Balmer breaks that can be produced at various redshifts can be calculated. For example, the largest Balmer break that can be produced by the time the Universe is approximately 550 Myr old (i.e., $z = 8.9$) has $D_{4000} = 1.26$. This includes both the timescales for initial collapse and subsequent quenching and an additional 340 Myr of quiescence, which is necessary to produce a Balmer break with $D_{4000} > 1.26$ after star formation. The maximum possible D_{4000} as a function of redshift is shown in Figure 5.

The ideal metric here would be $D_{n,4000}$, which is more narrowly designed to be sensitive specifically to the physics probed here. The D_{4000} measure described here enables comparison with previous work (Bruzual 1983; Poggianti & Barbaro 1997), but is also sensitive to dust and other physical processes beyond merely the age of the stellar population.

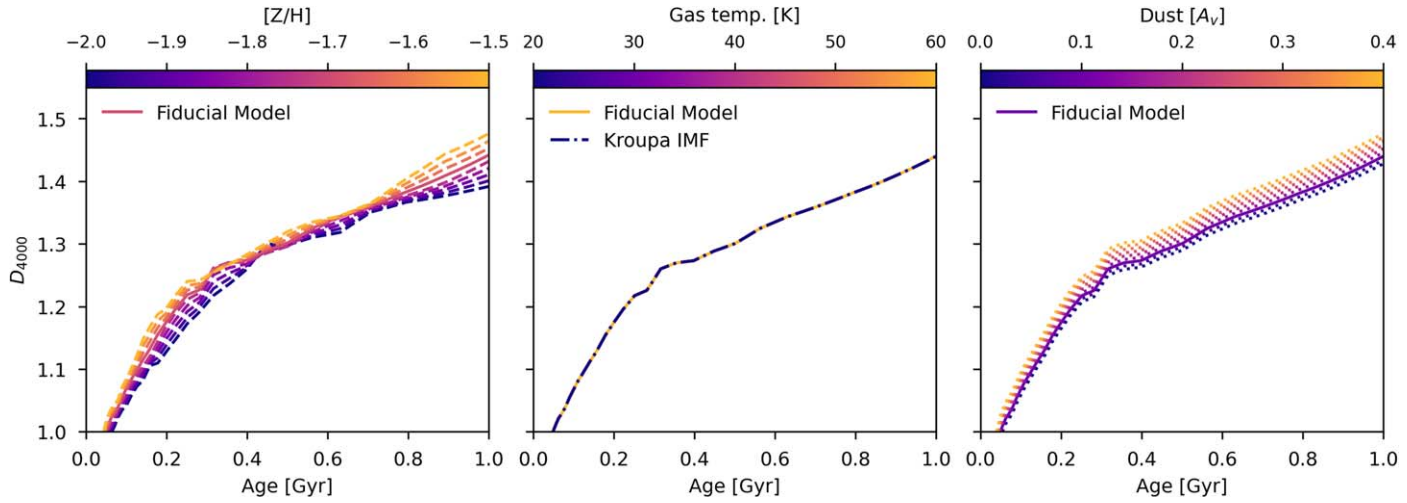

Figure 4. Growth of D_{4000} measuring the strength of the Balmer break as a function of stellar population age, varying metallicity (left), the IMF (middle) and dust extinction (right). For the fiducial model at $[Z/H] = -1.7$, $A_V = 0.1$, and an IMF corresponding to a gas temperature of 60 K, the Balmer break grows rapidly with age until $D_{4000} \approx 1.26$, which is followed by a more gradual growth. The timing of the earliest possible Balmer break is only weakly sensitive to assumptions of the IMF (middle) and extinction (right), but has a stronger dependence on metallicity (left).

It is tempting to argue that at high redshift, there has not yet been enough time for significant dust formation, and therefore the effects are negligible. However, it is also possible that this limited quantity of dust is highly concentrated near the sites of early star formation and can still produce significant attenuation. If so, there would be a meaningful difference between $D_{n,4000}$ and D_{4000}.

The fundamental problem, though, is that these predictions are being made only from photometric constraints, and $D_{n,4000}$ cannot be observationally determined from photometry. The argument being presented here is that the presumed D_{4000}, encompassing both Balmer and dust features, is physically implausible, requiring a fit at a lower redshift where the same features can instead be interpreted as emission lines. To that end, the constraints here are phrased in terms of D_{4000}, the more directly constrained quantity.

5. Discussion

The timing of the formation of the first stars and galaxies is intrinsically linked to structure formation within our cosmological model. Despite the difficulties of modeling such astrophysical processes, we have shown that strong Balmer breaks at high redshift can provide a test of early structure formation and potentially new physics beyond vanilla ΛCDM.

We find that a conservative upper limit for the highest-redshift Balmer break is $z \sim 9$, assuming a single burst of self-quenching very early star formation. Leveraging the state-of-the-art SPHINX simulation of early galaxy formation and reionization, we find that a more realistic estimate is closer to $z = 7.5$ given our current understanding of star formation and feedback in the first galaxies in an average-density environment of the Universe.

Interestingly, photometric observations of Balmer breaks in this regime have already been reported from photometric template fitting. This includes objects with $z_{\rm phot} \approx 7$–9 (Roberts-Borsani et al. 2020; Labbé et al. 2023; Laporte et al. 2023), $z_{\rm phot} \approx 9$–10 (Hashimoto et al. 2018; Laporte et al. 2021; Atek et al. 2023; Furtak et al. 2023), and even a couple of candidates at $z_{\rm phot} \approx 11$ (Adams et al. 2023). If taken at face value,[7] and uncertainties shown in Figure 5 are those reported by the original authors, even where the reported uncertainty is likely far smaller than truly attainable for photometric redshifts or than would be expected for D_{4000}. Best-fit template measurements of D_{4000} were provided by the original authors for Hashimoto et al. (2018), Laporte et al. (2021), Adams et al. (2023), Furtak et al. (2023), and Boyett et al. (2024), and estimated from figures for the remainder. A minimum D_{4000} uncertainty of $\sigma_m = 0.02$ is imposed in the interest of conservative uncertainty estimation. Several of these measurements are already hinting at new physics (see Figure 5). The boundaries shown here are for the fiducial model with $[Z/H] = -1.7$ and $A_V = 0.1$, but are only weakly sensitive to changes in the model (Figure 4) as the dominant uncertainty is in the measurement of D_{4000}.

However, there is an absolute necessity for spectroscopic confirmation of these breaks to draw robust conclusions. Indeed, both initial NIRSpec spectroscopy (Carnall et al. 2023) and template fits using physics tuned to ultra-high-redshift galaxies (Steinhardt et al. 2023) suggest these features may instead be due to strong nebular emission.

Conversely, a very weak Balmer break has already been inferred from spectroscopy at $z = 9.3127 \pm 0.0002$ (Boyett et al. 2024), although the spectrum only covers around half of the redder band needed to estimate D_{4000}. Similarly, the Hashimoto et al. (2018) object has spectroscopic confirmation of $z = 9.1096 \pm 0.0006$ from an 88 μm oxygen line, but the Balmer break is only detected photometrically.

One mechanism suggested for producing such an early Balmer break would be a combination of an older stellar population and dust-enshrouded younger population. This might allow a Balmer break to be produced even as a galaxy continues to form stars (Katz et al. 2019). Such a mechanism would mimic the appearance of quenching, and thus allow Balmer breaks in the blue region in Figure 5, where no new physics is required. Such a model was successfully used to explain MACS1149_JD1. However, most of the photometric Balmer breaks that have been reported would be deep enough

[7] Redshifts, D_{4000} values.

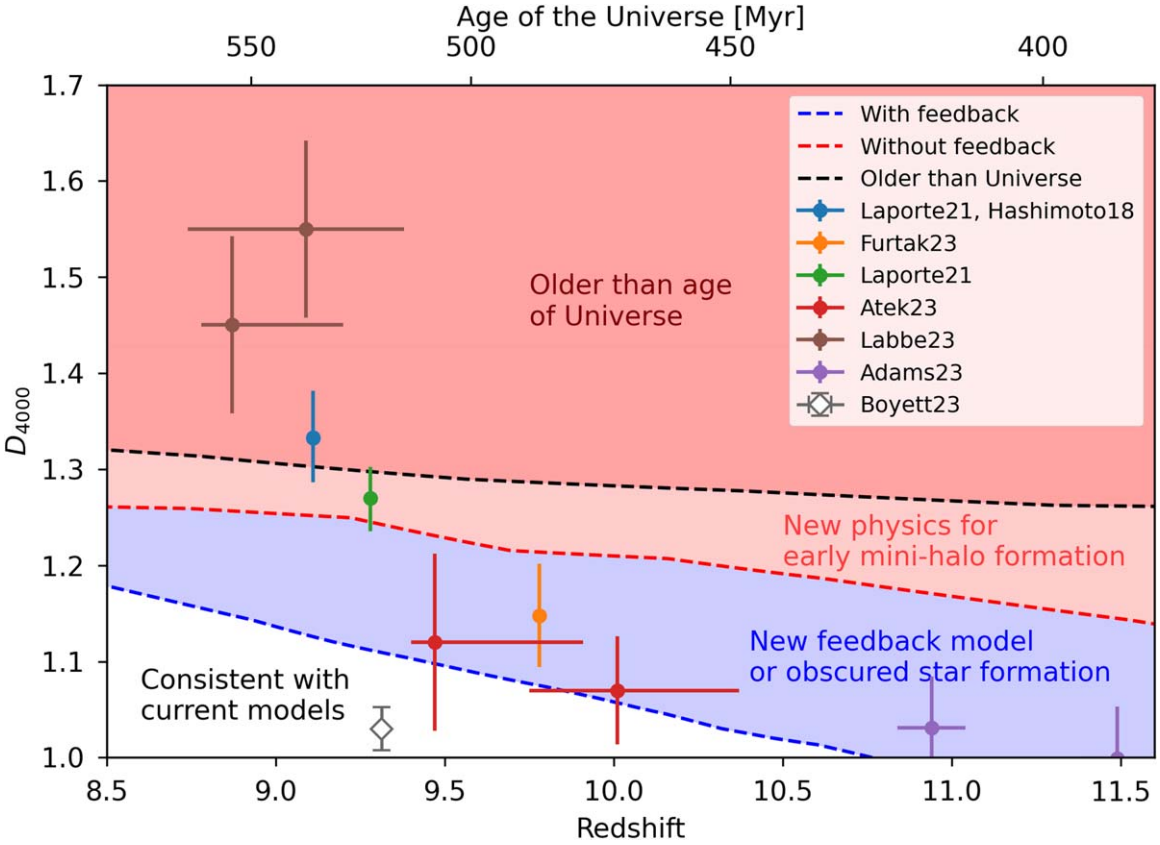

Figure 5. A selection of galaxies with photometrically selected Balmer breaks (Hashimoto et al. 2018; Laporte et al. 2021, 2023; Adams et al. 2023; Atek et al. 2023; Furtak et al. 2023; Labbé et al. 2023), as reported by the original authors (MACSJ1149-JD1 is first discussed in Hashimoto et al. (2018) with updated parameters taken from Laporte et al. 2021). The maximum possible $D_{4000}(z)$ as calculated here is shown for comparison. Taken at face value, the strongest, highest-redshift Balmer breaks predicted by current photometric studies are inconsistent with "vanilla" ΛCDM and require new physics. In fact, from our analysis of the Balmer break, the inferred stellar population would be substantially older than the age of the Universe for current cosmological parameters (Planck Collaboration et al. 2020) and require a far more exotic model. Weaker photometric breaks could be reconciled by alterations to feedback models, obscured star formation or new physics driving earlier minihalo collapse and structure formation. Spectroscopic confirmation is thus paramount to confirm the nature of these early Balmer breaks, with the current highest-redshift spectroscopically confirmed breaks (Boyett et al. 2024; Looser et al. 2024) being consistent with current models.

to require star formation earlier than expected for vanilla ΛCDM, which means if verified, they could not be explained by differential dust alone.

5.1. Are "Impossible" Photometric Balmer Breaks from JWST Genuine?

Previous studies have also found that the apparent stellar masses of these high-redshift Hubble and JWST galaxies are too high to have been produced by cold, collisionless dark matter (Steinhardt et al. 2016; Labbé et al. 2023). This tension becomes far stronger in combination with a deep presumed Balmer break, since producing the earliest possible Balmer breaks requires an older stellar population (Section 4) and likely a small stellar baryon fraction (Section 3), both of which greatly reduce the halo mass-to-light ratio. Models that constrain ultra-high-redshift objects to have physically plausible stellar IMFs and extinction instead produce far lower masses and younger stellar populations, consistent with vanilla ΛCDM (Steinhardt et al. 2023). These same templates also predict that the red photometric "bumps" are produced by strong nebular emission lines rather than a Balmer break, potentially resolving both problems via the same effect.

This explanation also predicts emission lines must be stronger than seen at lower redshifts, but comparable to those seen in early JWST spectroscopy at $z > 6$ (Cameron et al. 2023; Carnall et al. 2023; Sanders et al. 2023). In that way, upcoming spectroscopic follow-up of these early JWST candidates will not only determine which model is the better fit, but also simultaneously determine whether current cosmological models pass or fail both the stellar mass and Balmer-break tests.

The feature observed in the reddest bands (which is commonly interpreted as a Balmer break) has even been proposed as a selection mechanism in an attempt to confirm the ultrahigh redshifts of targets which would be difficult to pinpoint from a Lyman break alone (Labbé et al. 2023). However, as shown in this work, for any of these to truly be strong Balmer breaks would be inconsistent with vanilla ΛCDM. Regardless, such a selection mechanism will still work if the red photometric bumps are due to strong nebular emission lines, although it would imply the current inferred photometric redshifts are likely to be slightly overestimated. For example, stronger [OIII]+Hβ emission lines would produce a red bump in photometry at 15%–20% longer wavelengths than the Balmer-break feature, resulting in correspondingly lower redshifts. Finally, it is also possible that some of these photometric redshifts will turn out to have been catastrophic overestimates, as was already the case for some early JWST sources (e.g., Arrabal Haro et al. 2023). It should also be noted that many of these photometric redshifts were produced using methods that constrain the allowed templates based upon the redshift–age relation for vanilla ΛCDM, and this prior should

likely be removed for a proper cosmological test. However, spectroscopic follow-up will obviate the need for this prior as part of determining whether these galaxies are truly inconsistent with current cosmological models.

5.2. Cosmological Models for z > 9 Balmer Breaks

At present, it would seem that the most likely outcome is that spectroscopic follow-up will not discover Balmer breaks in contradiction with vanilla ΛCDM, but rather that the spectral energy distributions predicted from lower-redshift analogs are not representative of the first galaxies. However, given the possibility that some of these galaxies will indeed exhibit ultra-high-redshift Balmer breaks, it is worth considering what modifications to vanilla ΛCDM would be needed to produce them.

It is worth noting that the discovery of a very old stellar population would be in disagreement not directly with ΛCDM, but rather with the the ΛCDM parameters measured by current experiments and observations. A version of ΛCDM with very different parameters could allow the age of the Universe to be far higher at $z \sim 10$ than with the Planck Collaboration et al. (2020) cosmological parameters. However, the difference between Planck measurements and supernova-based parameters (Riess et al. 2022) is relatively small (around 30–40 Myr at $z \sim 10$). The Planck measurements used here would provide the oldest Universe at fixed redshift, so they are the most conservative current bound.

Two possible solutions include modifying the IMF so that massive stars never form and the Balmer break can form earlier in the evolution of the SSP, or modifying the matter power spectrum (e.g., via a modification to dark matter or other process) to allow for earlier structure formation. The former is a poor physical solution as due to increased heating from the CMB and other sources, the high-redshift IMF is expected to be bottom-light rather than bottom-heavy (Jermyn et al. 2018).

Modifying dark matter interactions will also impact structure formation. Warm dark matter models typically inhibit structure formation (Murray et al. 2013; Pacucci et al. 2013; Schneider et al. 2013), so a cold dark matter (CDM) scenario is still most likely. Self-interacting CDM models (Spergel & Steinhardt 2000; Tulin & Yu 2018) may be a more promising approach for earlier structure formation. Baryon–dark matter interactions have also been proposed (Barkana 2018) to hasten gas cooling and star formation in an attempt to explain possible detection of the global 21 cm absorption signal (Bowman et al. 2018; but see also Barkana et al. 2018).

Primordial magnetic fields (Turner & Widrow 1988; Ratra 1992) are known to enhance power at high k-modes of the matter power spectrum due to the Lorentz force (e.g., Wasserman 1978; Kim et al. 1996). This can lead to structure formation earlier than vanilla ΛCDM, although measurements/upper limits of the high-redshift electron optical depth (e.g., Planck Collaboration et al. 2020; Heinrich & Hu 2021) constrain the amount of possible early structure formation (Katz et al. 2021) due to its impact on HI reionization.

Another possibility would be a model that produces primordial structure. Primordial black holes have been proposed via several mechanisms (Carr & Kühnel 2020; Villanueva-Domingo et al. 2021). These would then seed more rapid growth of the first baryonic minihalos, and thus the first stars.

The most extreme photometric Balmer breaks, if verified, would require a stellar population which is significantly older than the age of the Universe using the Planck Collaboration et al. (2020) cosmology (Figure 5). Such a population could not be produced even by primordial structure formation or dark matter modifications, and would instead be entirely incompatible with ΛCDM models. Given the overwhelming observational evidence for the ΛCDM cosmological paradigm (see Bull et al. 2016), this would seem to be the least likely outcome.

Fortunately, the answer to this question is likely to be provided in the very near future. JWST/NIRSpec follow-up of several photometrically selected $z > 9$ candidates with red bumps has already been performed, with the results currently in preparation. If even a single one of these sources indeed exhibits a clear Balmer break with sufficiently high D_{4000}, as predicted by many photometric template fits, then vanilla ΛCDM will be ruled out. However, if they all instead caused by strong nebular emission, modifications in templates for high redshift along the lines of those proposed for bottom-light IMFs (Sneppen et al. 2022; Steinhardt et al. 2023) will be required instead.

Acknowledgments

The authors would like to thank Sebastien Aagaard, Kit Boyett, Lukas Furtak, Zoltan Haiman, Thomas Harvey, Troels Haugbølle, Anne Hutter, Nicolas Laporte, Tobias Looser, Bahram Mobasher, Vadim Rusakov, Luka Vujeva, and Darach Watson for helpful comments and discussions. The Cosmic Dawn Center (DAWN) is funded by the Danish National Research Foundation under grant No. 140. M.R. is supported by the Beecroft Fellowship funded by Adrian Beecroft.

ORCID iDs

Charles L. Steinhardt https://orcid.org/0000-0003-3780-6801
Albert Sneppen https://orcid.org/0000-0002-5460-6126
Thorbjørn Clausen https://orcid.org/0009-0006-7165-3828
Harley Katz https://orcid.org/0000-0003-1561-3814
Martin P. Rey https://orcid.org/0000-0002-1515-995X
Jonas Stahlschmidt https://orcid.org/0009-0007-4720-7438

References

Adams, N. J., Conselice, C. J., Ferreira, L., et al. 2023, MNRAS, 518, 4755
Arrabal Haro, P., Dickinson, M., Finkelstein, S. L., et al. 2023, Natur, 622, 707
Atek, H., Chemerynska, I., Wang, B., et al. 2023, MNRAS, 524, 5486
Baade, W. 1944, ApJ, 100, 137
Balogh, M. L., Morris, S. L., Yee, H. K. C., Carlberg, R. G., & Ellingson, E. 1999, ApJ, 527, 54
Barkana, R. 2018, Natur, 555, 71
Barkana, R., Outmezguine, N. J., Redigol, D., & Volansky, T. 2018, PhRvD, 98, 103005
Behroozi, P., & Silk, J. 2018, MNRAS, 477, 5382
Binggeli, C., Zackrisson, E., Ma, X., et al. 2019, MNRAS, 489, 3827
Bowman, J. D., Rogers, A. E. E., Monsalve, R. A., Mozdzen, T. J., & Mahesh, N. 2018, Natur, 555, 67
Boyett, K., Trenti, M., Leethochawalit, N., et al. 2024, NatAs, 8, 657
Boylan-Kolchin, M. 2023, NatAs, 7, 731
Bromm, V. 2013, RPPh, 76, 112901
Bruzual, A. G. 1983, ApJ, 273, 105
Bull, P., Akrami, Y., Adamek, J., et al. 2016, PDU, 12, 56
Cameron, A. J., Saxena, A., Bunker, A. J., et al. 2023, A&A, 677, A115
Carnall, A. C., Begley, R., McLeod, D. J., et al. 2023, MNRAS, 518, L45
Carr, B., & Kühnel, F. 2020, ARNPS, 70, 355
Choustikov, N., Katz, H., Saxena, A., et al. 2024, MNRAS, 529, 3751

Conroy, C. 2013, ARA&A, 51, 393
Conroy, C., & Wechsler, R. H. 2009, ApJ, 696, 620
Davidzon, I., Ilbert, O., Laigle, C., et al. 2017, A&A, 605, A70
Dekel, A., & Silk, J. 1986, ApJ, 303, 39
Dunlop, J. S. 2013, in The First Galaxies, ed. T. Wiklind, B. Mobasher, & V. Bromm (Berlin: Springer), 223
Ferrara, A., Viti, S., & Ceccarelli, C. 2016, MNRAS, 463, L112
Finkelstein, S. L., Bagley, M. B., Ferguson, H. C., et al. 2023, ApJL, 946, L13
Finkelstein, S. L., Song, M., Behroozi, P., et al. 2015, ApJ, 814, 95
Furtak, L. J., Shuntov, M., Atek, H., et al. 2023, MNRAS, 519, 3064
Gall, C., Hjorth, J., & Andersen, A. C. 2011, A&ARv, 19, 43
Haiman, Z., & Loeb, A. 1997, ApJ, 483, 21
Hashimoto, T., Laporte, N., Mawatari, K., et al. 2018, Natur, 557, 392
Heinrich, C., & Hu, W. 2021, PhRvD, 104, 063505
Hovis-Afflerbach, B., Steinhardt, C. L., Masters, D., & Salvato, M. 2021, ApJ, 908, 148
Iyer, K., & Gawiser, E. 2017, ApJ, 838, 127
Jaacks, J., Thompson, R., Finkelstein, S. L., & Bromm, V. 2018, MNRAS, 475, 4396
Jermyn, A. S., Steinhardt, C. L., & Tout, C. A. 2018, MNRAS, 480, 4265
Katz, H., Laporte, N., Ellis, R. S., Devriendt, J., & Slyz, A. 2019, MNRAS, 484, 4054
Katz, H., Martin-Alvarez, S., Rosdahl, J., et al. 2021, MNRAS, 507, 1254
Kim, E.-J., Olinto, A. V., & Rosner, R. 1996, ApJ, 468, 28
Klessen, R. S., & Glover, S. C. O. 2023, ARA&A, 61, 65
Kolb, E. W., & Turner, M. S. 1990, The Early Universe, Vol. 69 (Boston, MA: Addison-Wesley)
Labbé, I., van Dokkum, P., Nelson, E., et al. 2023, Natur, 616, 266
Laigle, C., McCracken, H. J., Ilbert, O., et al. 2016, ApJS, 224, 24
Laporte, N., Ellis, R. S., Witten, C. E. C., & Roberts-Borsani, G. 2023, MNRAS, 523, 3018
Laporte, N., Meyer, R. A., Ellis, R. S., et al. 2021, MNRAS, 505, 3336
Larson, R. B. 1985, MNRAS, 214, 379
Liu, B., & Bromm, V. 2020, MNRAS, 497, 2839
Looser, T. J., D'Eugenio, F., Maiolino, R., et al. 2024, Natur, 629, 53
Low, C., & Lynden-Bell, D. 1976, MNRAS, 176, 367
Miralda-Escudé, J. 2003, Sci, 300, 1904
Mobasher, B., Dickinson, M., Ferguson, H. C., et al. 2005, ApJ, 635, 832
Murray, S. G., Power, C., & Robotham, A. S. G. 2013, A&C, 3, 23
Nakamura, F., & Umemura, M. 2001, ApJ, 548, 19
Pacucci, F., Mesinger, A., & Haiman, Z. 2013, MNRAS, 435, L53
Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A6
Poggianti, B. M., & Barbaro, G. 1997, A&A, 325, 1025
Ratra, B. 1992, ApJL, 391, L1
Riess, A. G., Yuan, W., Macri, L. M., et al. 2022, ApJL, 934, L7
Rix, H.-W., Chandra, V., Andrae, R., et al. 2022, ApJ, 941, 45
Roberts-Borsani, G. W., Ellis, R. S., & Laporte, N. 2020, MNRAS, 497, 3440
Rosdahl, J., Blaizot, J., Katz, H., et al. 2022, MNRAS, 515, 2386
Rosdahl, J., Katz, H., Blaizot, J., et al. 2018, MNRAS, 479, 994
Sanders, R. L., Shapley, A. E., Topping, M. W., Reddy, N. A., & Brammer, G. B. 2023, ApJ, 955, 54
Schneider, A., Smith, R. E., & Reed, D. 2013, MNRAS, 433, 1573
Shahidi, A., Mobasher, B., Nayyeri, H., et al. 2020, ApJ, 897, 44
Silk, J. 1968, ApJ, 151, 459
Sneppen, A., Steinhardt, C. L., Hensley, H., et al. 2022, ApJ, 931, 57
Spergel, D. N., & Steinhardt, P. J. 2000, PhRvL, 84, 3760
Stanway, E. R., & Eldridge, J. J. 2018, MNRAS, 479, 75
Steinhardt, C. L., Capak, P., Masters, D., & Speagle, J. S. 2016, ApJ, 824, 21
Steinhardt, C. L., Kokorev, V., Rusakov, V., Garcia, E., & Sneppen, A. 2023, ApJ, 951, 40
Treu, T., Roberts-Borsani, G., Bradac, M., et al. 2022, ApJ, 935, 110
Tseliakhovich, D., & Hirata, C. 2010, PhRvD, 82, 083520
Tulin, S., & Yu, H.-B. 2018, PhR, 730, 1
Turner, M. S., & Widrow, L. M. 1988, PhRvD, 37, 2743
Villanueva-Domingo, P., Mena, O., & Palomares-Ruiz, S. 2021, FrASS, 8, 87
Wasserman, I. 1978, ApJ, 224, 337
Weaver, J. R., Kauffmann, O. B., Ilbert, O., et al. 2022, ApJS, 258, 11
White, S. D. M., & Rees, M. J. 1978, MNRAS, 183, 341