

The galaxy ultraviolet luminosity function from $z = 7$ to 20 in the COLIBRE simulations

Shengdong Lu¹,[★] Carlos S. Frenk¹, Cedric G. Lacey¹, Andrea Gehek², Joop Schaye³, Shaun Cole¹, Sownak Bose¹, Anna Durrant⁴, Nick Andreadis², Maarten Baes², Alejandro Benítez-Llambay⁵, Evgenii Chaikin^{1,3}, Camila Correa³, Robert A. Crain⁴, Filip Huško³, Robert J. McGibbon³, Sylvia Ploekinger⁶, Alexander J. Richings^{7,8}, Matthieu Schaller^{3,9} and James W. Trayford¹⁰

¹Institute for Computational Cosmology, Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK

²Department of Physics and Astronomy, Universiteit Gent, Proeftuinstraat 86 N3, B-9000 Ghent, Belgium

³Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands

⁴Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK

⁵Dipartimento di Fisica G. Occhialini, Università degli Studi di Milano Bicocca, Piazza della Scienza, 3 I-20126 Milano MI, Italy

⁶Department of Astrophysics, University of Vienna, Türkenschanzstrasse 17, A-1180 Vienna, Austria

⁷Centre for Data Science, Artificial Intelligence and Modelling, University of Hull, Cottingham Road, Hull HU6 7RX, UK

⁸E. A. Milne Centre for Astrophysics, University of Hull, Cottingham Road, Hull HU6 7RX, UK

⁹Lorentz Institute for Theoretical Physics, Leiden University, PO Box 9506, NL-2300 RA Leiden, the Netherlands

¹⁰Centre for Astrophysics Research, Department of Physics, Astronomy and Mathematics, University of Hertfordshire, College Lane, Hatfield, AL10 9AB, UK

Accepted 2026 September 18. Received 2026 August 26; in original form 2026 May 8

ABSTRACT

James Webb Space Telescope (JWST) has enabled the detection of galaxies in the earliest stages of cosmic history. We compare the ultraviolet luminosity functions (UVLFs) at redshifts $z = 7–20$ predicted by a set of new cosmological hydrodynamical simulations, COLIBRE (COLD ISM AND BETTER RESOLUTION), with observations, including those from *JWST*. The UV luminosities of COLIBRE galaxies are derived using the radiative transfer code SKIRT, which tracks stellar emission and its processing through the multiphase interstellar medium and dust distribution predicted by COLIBRE. We find that although COLIBRE is consistent with the observed evolution of the stellar mass function up to $z = 12$, its dust-attenuated UVLFs fall systematically below the observations at the bright end: at the number density of $10^{-6} \text{ Mpc}^{-3} \text{ mag}^{-1}$, the brightest galaxies are underluminous by $\approx 1 \text{ mag}$ at $z = 7$, increasing to $\approx 2.5 \text{ mag}$ at $z = 15$. Accounting for observational uncertainties brings the COLIBRE UVLFs closer to the observational data, but does not fully resolve the discrepancy. Ignoring dust attenuation allows COLIBRE to produce sufficiently bright galaxies at $7 \lesssim z \lesssim 12$, but at $z = 15$, COLIBRE still underpredicts the luminosities of the brightest galaxies, indicating the need for additional physical mechanisms to boost the UV luminosities at the earliest cosmic epochs, such as a ‘top-heavy’ stellar initial mass function. We fit the COLIBRE UVLFs with Schechter functions and calculate the evolution of the best-fitting parameters. We find that the galaxy number density decreases, the characteristic luminosity becomes fainter and the faint-end slope becomes steeper towards higher redshifts. The UV luminosity density decreases by a factor of several hundred from $z = 7$ to 15.

Key words: radiative transfer – dust, extinction – galaxies: evolution – galaxies: formation – galaxies: high-redshift – galaxies: luminosity function, mass function.

1 INTRODUCTION

Galaxies in the early Universe carry key information about the formation and evolution of cosmic structures. They reveal the transition of the Universe from an initially near-uniform state to the one we observe today filled with complex structures, reflecting physical processes, such as dark matter halo collapse, gas

cooling, star formation, and feedback loops. These early galaxies are thought to be the primary source of the ionizing photons responsible for cosmic reionization, which significantly influenced the transparency and subsequent evolution of the Universe (e.g. P. Madau & M. Dickinson 2014; B. E. Robertson et al. 2015).

Deep imaging with the *Hubble Space Telescope (HST)* and complementary *Spitzer*/IRAC observations had already extended galaxy detections to $z \approx 7–11$ (e.g. P. A. Oesch et al. 2016; G. W. Roberts-Borsani et al. 2016; G. W. Roberts-Borsani, R. S. Ellis & N. Laporte 2020; R. J. Bouwens et al. 2021), providing early

* E-mail: shengdong.lu@durham.ac.uk

constraints on the rest-frame optical properties and galaxy stellar masses in the first few hundred million years after the big bang. With the launch of the *James Webb Space Telescope* (*JWST*), our observational horizon expanded dramatically, reaching into the earliest epochs of cosmic history. For example, H. Atek et al. (2023) found two candidate galaxies at $z \approx 16$ in the first *JWST* deep field. More recently, *JWST* spectroscopy has pushed the frontier of spectroscopic confirmations to higher and higher redshifts, with galaxies now confirmed at $z \approx 14.3$ and ≈ 14.4 (S. Carniani et al. 2024; R. P. Naidu et al. 2026). F. Cullen et al. (2023) investigated the rest-frame ultraviolet (UV) continuum slope, β , of galaxies at $8 < z < 16$ using *JWST* and ground-based imaging, finding that high-redshift galaxies are generally blue, with a $\beta - M_{\text{UV}}$ (where M_{UV} is the absolute UV magnitude) relation and other properties consistent with those of young, metal-poor, moderately dust-reddened stellar populations.

Cosmological hydrodynamical simulations, such as EAGLE (R. A. Crain et al. 2015; J. Schaye et al. 2015), HORIZON-AGN (Y. Dubois et al. 2016; S. Kaviraj et al. 2017), FABLE (N. A. Henden et al. 2018), ILLUSTRISTNG (D. Nelson et al. 2019), SIMBA (R. Davé et al. 2019), FLAMINGO (J. Schaye et al. 2023), and MILLENNIUMTNG (R. Pakmor et al. 2023), have made significant progress in reproducing many key observed galaxy properties across cosmic time, including, for example, galaxy stellar mass functions (GSMFs), star formation rates (SFRs), sizes, morphologies, and chemical enrichment trends and have thus become crucial tools for interpreting astronomical observations and testing theoretical models of galaxy formation and evolution.

More recently, the COLIBRE (COLd ISM and Better RESolution) collaboration (J. Schaye et al. 2026; E. Chaikin et al. 2026a) introduced a new galaxy formation framework that explicitly models the multiphase interstellar medium (ISM), non-equilibrium H/He chemistry, and dust evolution. Spanning a wide range of volumes and resolutions, COLIBRE also reproduces key observed galaxy properties, such as the GSMF and its evolution (E. Chaikin et al. 2026b), the evolution of sizes (A. D. Ludlow et al. 2026) and resolved star formation laws (C. d. P. Lagos et al. 2026), making it a promising tool for studies of galaxy formation across cosmic time (J. Schaye et al. 2026; E. Chaikin et al. 2026b).

Most large-volume cosmological simulations (including COLIBRE) are calibrated to low-redshift observational constraints, especially the low- z GSMF, since stellar mass is a direct output of the simulations and can be estimated observationally from broadband spectral energy distribution (SED) fitting (e.g. R. A. Crain et al. 2015; D. Nelson et al. 2019). However, converting observed luminosities into stellar masses is a form of inverse modelling, which depends on a number of highly uncertain assumptions such as the star formation history, dust attenuation, metallicity, and the assumed stellar initial mass function (IMF; e.g. P. D. Mitchell et al. 2013; B. Lo Faro et al. 2017). These assumptions enter both into parametric SED-fitting approaches, which adopt specific functional forms for the star formation history (e.g. exponentially declining or delayed- τ models) and more flexible non-parametric methods, which avoid explicit functional forms but still rely on priors and finite-temporal resolution in reconstructing star formation histories (e.g. A. C. Carnall et al. 2019; J. Leja et al. 2019; C. Pacifici et al. 2023).

An alternative strategy is forward modelling, in which simulated galaxy properties (such as stellar mass and star formation history) are translated into directly observable quantities, such as luminosities, allowing a closer connection to the data

(S. McAlpine et al. 2016). While this approach is also model-dependent and relies, for example, on stellar population synthesis models with choices made in the simulations for the IMF and dust prescriptions, it is grounded in a single, internally self-consistent galaxy formation model. In particular, the same underlying physical framework simultaneously determines star formation, chemical enrichment, gas content in different phases and dust content, so tensions with observational luminosity functions can provide more direct diagnostics of the underlying physics. As a result, comparisons based on luminosity-related observables, which more naturally account for observational selection effects and survey characteristics, offer a powerful test of simulated galaxy populations, particularly at high redshift where stellar mass estimates remain highly uncertain.

The rest-frame UV luminosity function (UVLF) at high redshift quantifies the abundance and brightness of star-forming galaxies in the early Universe. Since rest-frame UV light traces recent star formation activity, the UVLF offers critical insights into the buildup of stellar mass, the efficiency of star formation and the ionizing photon budget during the epoch of reionization. Accurate measurements of the UVLF are therefore essential for constraining models of galaxy formation and the timeline of cosmic reionization. Recent deep observations from *JWST* have significantly improved the accuracy and depth of UVLF measurements at $z > 7$, providing crucial information on early galaxy populations (e.g. Y. Harikane et al. 2023, 2024; C. M. Casey et al. 2024; B. Robertson et al. 2024; R. P. Naidu et al. 2026).

Comparing the predictions for UVLFs at high redshift from cosmological simulations to the observations serves to assess the validity of current galaxy formation models and also to identify the key physical processes driving early galaxy evolution. Specific predictions for the UVLFs that *JWST* was expected to measure at high redshift in the Lambda-cold dark matter (Λ CDM) cosmology were presented, prior to the launch of *JWST*, by W. I. Cowley et al. (2018) using the Durham semi-analytic galaxy formation model, GALFORM (S. Cole et al. 2000; C. G. Lacey et al. 2016). A distinctive feature of this model is that the IMF in starburst episodes is assumed to be top-heavy (a modification introduced to explain the counts of faint submillimetre galaxies; C. M. Baugh et al. 2005).

Some of the first wave of *JWST* papers compared the data to a selection of models (not including that of W. I. Cowley et al. 2018) and concluded that bright galaxies at the highest redshifts are more abundant than predicted by the theoretical models they considered (e.g. S. L. Finkelstein et al. 2024). However, S. Lu et al. (2025) compared the pre-existing GALFORM predictions of UVLFs from W. I. Cowley et al. (2018) to the *JWST* and other observations and found remarkable agreement up to $z = 12$. An adjustment of the model to take into account the time-scale for the growth of dust grains extended this agreement out to the largest redshift for which there are data, $z \approx 14$.

Several other comparisons with theoretical predictions have been made since the publication of the *JWST* data. R. Kannan et al. (2023) calculated UVLFs for galaxies at $z > 10$ in the MILLENNIUMTNG hydrodynamical simulation. This model shows broad agreement with observations up to $z = 10$, but to come close to the data at $z = 10-12$ requires assuming that galaxies are dust free at these redshifts. Even in the absence of dust, the model underpredicts galaxy abundances by approximately an order of magnitude at $z = 15$. The authors conjectured that processes not included in the simulation, such as a variable stellar IMF or

contributions from Population III (Pop III) stars, could be necessary to resolve the discrepancy.

Similarly, L. Y. A. Yung et al. (2024), using the Santa Cruz semi-analytic model (R. S. Somerville & J. R. Primack 1999), presented results for the UVLF at high redshift, finding that the predicted galaxy number densities at $z > 11$ are at least an order of magnitude below the observed values even if galaxies are assumed to be dust free. They too suggested that a top-heavy IMF might eliminate the discrepancy. By contrast, R. S. Somerville et al. (2025) explored whether the observed abundance of UV-luminous galaxies at $z \gtrsim 10$ could be explained by invoking higher star formation efficiencies (SFEs) in dense environments. By introducing an *ad hoc* density-modulated SFE model into the Santa Cruz semi-analytic model, they showed that increasing the fraction of gas in dense clouds can reproduce the observed UVLFs over the high-redshift range considered, without requiring a top-heavy IMF. However, their analysis focused exclusively on high redshifts and did not present any evidence that the modified model remains consistent with observational constraints at lower redshift. Other recent studies have also compared model predictions with *JWST* UVLF measurements; some of these are discussed in Section 4.

In a series of three papers, we present a comprehensive study of the luminosity functions predicted by the COLIBRE simulations (J. Schaye et al. 2026; E. Chaikin et al. 2026a), across the full-wavelength range from the far-UV (FUV) to the submillimetre and over redshifts from $z = 0$ to ≈ 20 . We post-process the COLIBRE simulations with the 3D Monte Carlo radiative transfer code, SKIRT, to calculate the reprocessing of stellar emission by dust in the ISM (M. Baes et al. 2011; P. Camps & M. Baes 2015, 2020). We use for this the calibration-free COLIBRE-SKIRT pipeline developed by A. Gebek et al. (2026), who showed that when processed in this way, COLIBRE predicts a cosmic SED (CSED) at $z \approx 0$ that is consistent with local observations. In Paper I (S. Lu et al. 2026), using the same COLIBRE-SKIRT pipeline, we found remarkably good agreement at $z \approx 0$ between the predicted and observed LF from the FUV to the near-infrared (NIR) and from the far-infrared (FIR) to the submillimetre. This agreement suggests that COLIBRE successfully captures both the properties of stellar populations and the amount and spatial distribution of interstellar dust in present-day galaxies. In addition, E. Chaikin et al. (2026b) showed that COLIBRE roughly reproduces the observed evolution of the GSMF up to $z \approx 12$ and the number density of massive quenched galaxies up to $z \approx 7$.

In this second paper in the series, we adopt the same COLIBRE-SKIRT pipeline that has been shown to be successful at $z \approx 0$ and focus on the high-redshift UVLFs predicted by COLIBRE at $z = 7-20$, comparing them with observational data, particularly those from *JWST*. An important feature of this work is that it allows us to quantify the effects of dust attenuation using dust distributions computed self-consistently in COLIBRE with a live dust model (J. W. Trayford et al. 2026). Crucially, the COLIBRE model has been calculated down to $z = 0$ and shown to satisfy key observational constraints in the local Universe, as described above. Readers are referred to Paper I (S. Lu et al. 2026) for the study of the local ($z = 0$) LFs and to Paper III (Lu et al., in preparation) for the redshift evolution of broad-band LFs, again from the FUV to the submillimetre, up to intermediate redshifts ($z \approx 6$).

This paper is organized as follows: in Section 2, we introduce the COLIBRE simulations (Section 2.1), the galaxy sample selection method (Section 2.2), the radiative transfer code SKIRT and its setup (Section 2.3) and the compilation of observational data

for comparison (Section 2.4). The main results are presented in Section 3, where we study the correlation between stellar mass and galaxy UV magnitudes (Section 3.1) and compare the UVLFs of COLIBRE galaxies to observations (Section 3.2). We discuss our results in Section 4 and present a summary of our findings in Section 5. Throughout this paper, we adopt the cosmological parameters used in the COLIBRE simulations (taken from T. M. C. Abbott et al. 2022): present-day matter density, $\Omega_{m,0} = 0.306$; present-day baryon density, $\Omega_{b,0} = 0.0486$; linear theory amplitude of matter fluctuations on $8 \text{ Mpc } h^{-1}$ scales, $\sigma_8 = 0.807$; and dimensionless *Hubble* parameter, $h = 0.681$.

2 DATA AND METHOD

2.1 The COLIBRE simulations

COLIBRE¹ (J. Schaye et al. 2026; E. Chaikin et al. 2026a) is a new suite of cosmological hydrodynamical simulations of galaxy formation, performed with the open-source code SWIFT (M. Schaller et al. 2024). Below we briefly summarize the key aspects of the COLIBRE model. Readers are referred to J. Schaye et al. (2026) for a full description.

(i) *Hydrodynamics and gravity*. Gas dynamics are solved using the energy-density smoothed particle hydrodynamics scheme SPHENIX (J. Borrow et al. 2022). Gravitational forces are computed with a hybrid scheme combining a fast multipole method for short-range forces with a particle-mesh approach for long-range forces. Initial conditions are generated with MONOFONIC (O. Hahn et al. 2020; M. Michaux et al. 2021), starting from $z = 63$.

(ii) *Resolved multiphase ISM and non-equilibrium cooling*. Radiative cooling is followed down to temperatures of $\approx 10 \text{ K}$ (S. Ploekinger et al. 2025), enabling an explicit treatment of the cold, dense ISM phase. Cooling, heating, and the time evolution of chemical species are computed with the CHIMES network (A. J. Richings, J. Schaye & B. D. Oppenheimer 2014a, b). Hydrogen and helium are treated in non-equilibrium, while metal-line cooling is calculated element-by-element, accounting for depletion onto dust grains. The model includes molecular formation in the gas phase and on dust, shielding by gas and dust, photoelectric heating, cosmic rays, and radiation fields, which together regulate the ionization state, thermal balance, and molecular content of the cold ISM. For the metagalactic radiation field, COLIBRE adopts a modified redshift-dependent UV/X-ray background from C.-A. Faucher-Giguère (2020), with hydrogen and helium reionization completing at $z \approx 7$ and ≈ 3 , respectively. Reionization is therefore treated assuming a spatially uniform background, rather than through explicit radiative transfer, with local attenuation of photoionization, photodissociation and photoheating captured through an approximate treatment of gas and dust self-shielding.

(iii) *Dust evolution*. Dust grain formation, growth, and destruction are modelled self-consistently during the simulation using a multigrain dust evolution model (J. W. Trayford et al. 2026). The dust model is fully coupled to the cooling and chemistry calculations, such that the local dust abundance affects molecular formation rates, metal depletion and cooling efficiencies.

(iv) *Star formation and feedback*. Star formation follows a Schmidt law combined with a gravitational instability criterion

¹<https://www.colibre-simulations.org/>

Table 1. Summary of the COLIBRE simulations used in this work.

Simulation name	$L_{\text{box}}[\text{cMpc}]$	$m_{\text{gas}}[\text{M}_{\odot}]$	$m_{\text{DM}}[\text{M}_{\odot}]$	N_{gas}	N_{DM}	AGN feedback model	Minimum stellar mass [$\text{M}_{\odot}$]
L100m5	100	2.30×10^5	3.03×10^5	3008^3	4×3008^3	Thermal	5×10^5
L100m6	100	1.84×10^6	2.42×10^6	1504^3	4×1504^3	Thermal	5×10^6
L200m6	200	1.84×10^6	2.42×10^6	3008^3	4×3008^3	Thermal	5×10^6
L400m7	400	1.47×10^7	1.94×10^7	3008^3	4×3008^3	Thermal	5×10^7
L100m6h	100	1.84×10^6	2.42×10^6	1504^3	4×1504^3	Hybrid	5×10^6

Notes. Column (1): simulation name; (2): comoving side of the computational cube; (3): gas particle mass; (4): DM particle mass; (5): number of gas particles in the initial conditions; (6): number of DM particles in the initial conditions; (7): AGN feedback model; and (8): lower stellar mass limit of selected galaxies.

(F. S. J. Nobels et al. 2024), allowing an effective star formation relation to emerge from the resolved multiphase ISM. Feedback from massive stars includes early stellar feedback (stellar winds, radiation pressure, and H II regions) and core-collapse supernovae (SN, E. Chaikin et al. 2023; A. Benítez-Llambay et al. 2026). Type Ia SNe and chemical enrichment follow updated nucleosynthetic yields and delay-time distributions (C. A. Correa et al. 2026). Most colibre simulations use the purely thermal active galactic nucleus (AGN) feedback model of J. Schaye et al. (2026), while a subset employ instead the hybrid thermal-plus-kinetic-jet model of F. Huško et al. (2026), which combines thermally driven winds with kinetic jets.

(v) *Resolution and box sizes.* COLIBRE is available in three numerical resolutions, m5, m6, and m7, with corresponding gas particle masses of 2.30×10^5 , 1.84×10^6 , and $1.47 \times 10^7 \text{ M}_{\odot}$ and cube sizes ranging from 25 to 400 cMpc. Dark matter is super-sampled by a factor of 4 relative to baryonic particles, yielding similar dark matter and baryonic particle masses, in order to suppress spurious transfer of energy from dark matter to stars (A. D. Ludlow et al. 2019, 2023). A redshift-dependent gravitational softening length is adopted to ensure that the spatial resolution increases at high redshift, which is important for resolving the ISM structure of early galaxies.

(vi) *Calibration and galaxy identification.* The model is calibrated at the fiducial m7 resolution to reproduce a small set of key $z = 0$ galaxy scaling relations: the GSMF, the galaxy size–stellar mass relation (SSMR) and the black hole mass–stellar mass relation (BSMR, E. Chaikin et al. 2026a). The calibration to the $z = 0$ GSMF and SSMR is performed using Gaussian process emulators, while the calibration to the $z = 0$ BSMR is carried out manually and independently. After calibrating the m7 model, higher resolution simulations are not fully recalibrated. Instead, the best-fitting m7 parameters are adopted as a reference point and for the m6 and m5 resolutions, the SN and AGN feedback parameters are manually adjusted starting from this reference to recover a comparable level of agreement with the $z = 0$ GSMF, SSMR, and BSMR. Dark matter haloes are first identified with a friends-of-friends algorithm (M. Davis et al. 1985) and self-bound subhaloes are then found using HBT-HERONS (J. Han et al. 2018; V. J. Forouhar Moreno et al. 2025).

2.2 Galaxy sample selection

We follow the procedure outlined in Paper I to select galaxies for which to calculate FUV luminosities including the effects of dust attenuation. The key steps are summarized below:

(i) We select galaxies from three COLIBRE simulations of different resolutions and volumes (L100m5, L200m6, and L400m7), as

listed in Table 1, in order to derive UVLFs over a broad range of luminosity.²

(ii) We only analyse simulation outputs corresponding to COLIBRE ‘snapshots’ which contain more detailed information on the particles in galaxies, such as the dust mass fractions of each dust species, required for radiative transfer calculations. From the available snapshots, we select those at $z = 7, 9, 10, 12, 15$, and 20. In Fig. 1, we present the correlations between galaxy stellar mass and several global galactic properties for the three COLIBRE simulations at the selected redshifts, using galaxies above stellar mass limits of $5 \times 10^5 \text{ M}_{\odot}$ for m5, $5 \times 10^6 \text{ M}_{\odot}$ for m6, and $5 \times 10^7 \text{ M}_{\odot}$ for m7, which correspond to roughly 2–5 stellar particles.³ At $z = 20$, only three galaxies in L200m6 satisfy this stellar mass limit, while none do so in L400m7. We therefore consider only the results from L100m5 at this redshift.

The properties shown in Fig. 1 are: (1) the total mass (including baryons, dark matter, and black holes), denoted as M_{tot} ; (2) the mass of gas in the cool and dense phase, denoted as M_{cg} (defined as gas with $T < 10^{4.5} \text{ K}$ and $\rho_{\text{g}}/m_{\text{H}} > 10^{-1} \text{ cm}^{-3}$, where T and ρ_{g} are the gas temperature and density, respectively and m_{H} is the mass of a hydrogen atom); (3) the instantaneous SFR; and (4) the dust mass, denoted as M_{dust} . The total mass is calculated using all gravitationally bound particles, while the other global properties are measured within a spherical aperture of radius 50 kpc. All these quantities are taken from the Spherical Over-density and Aperture Processor (SOAP; R. McGibbon et al. 2025) catalogue, which contains a wide range of global properties measured within different apertures. As shown in Fig. 1, the three COLIBRE simulations with different resolutions produce very consistent correlations, which agree down to very low stellar masses (corresponding to about two stellar particles). This convergence test demonstrates the strong consistency of these relations across different COLIBRE resolutions.

(iii) For each COLIBRE simulation and at each redshift, we remove galaxies with stellar masses below the resolution limits adopted in Fig. 1. The remaining galaxies are then grouped in the stellar mass–SFR plane using bins of 0.5 dex in both stellar mass and SFR. From each bin, we randomly select 50 galaxies. If a bin contains fewer than 50 galaxies, all galaxies in that bin are included. For each bin we record the sampling fraction, defined as the ratio between the number of selected galaxies and the total number of galaxies in that bin, which is then used to upweight the contribution of the subsampled population to

²The COLIBRE galaxy data were accessed using the PYTHON packages SWIFTSIMIO (J. Borrow & A. Borrisov 2020; J. Borrow & A. J. Kelly 2021) and SWIFTGALAXY (K. A. Oman 2025).

³The approximate number of stellar particles in galaxies at the lower stellar-mass boundary varies slightly with resolution.

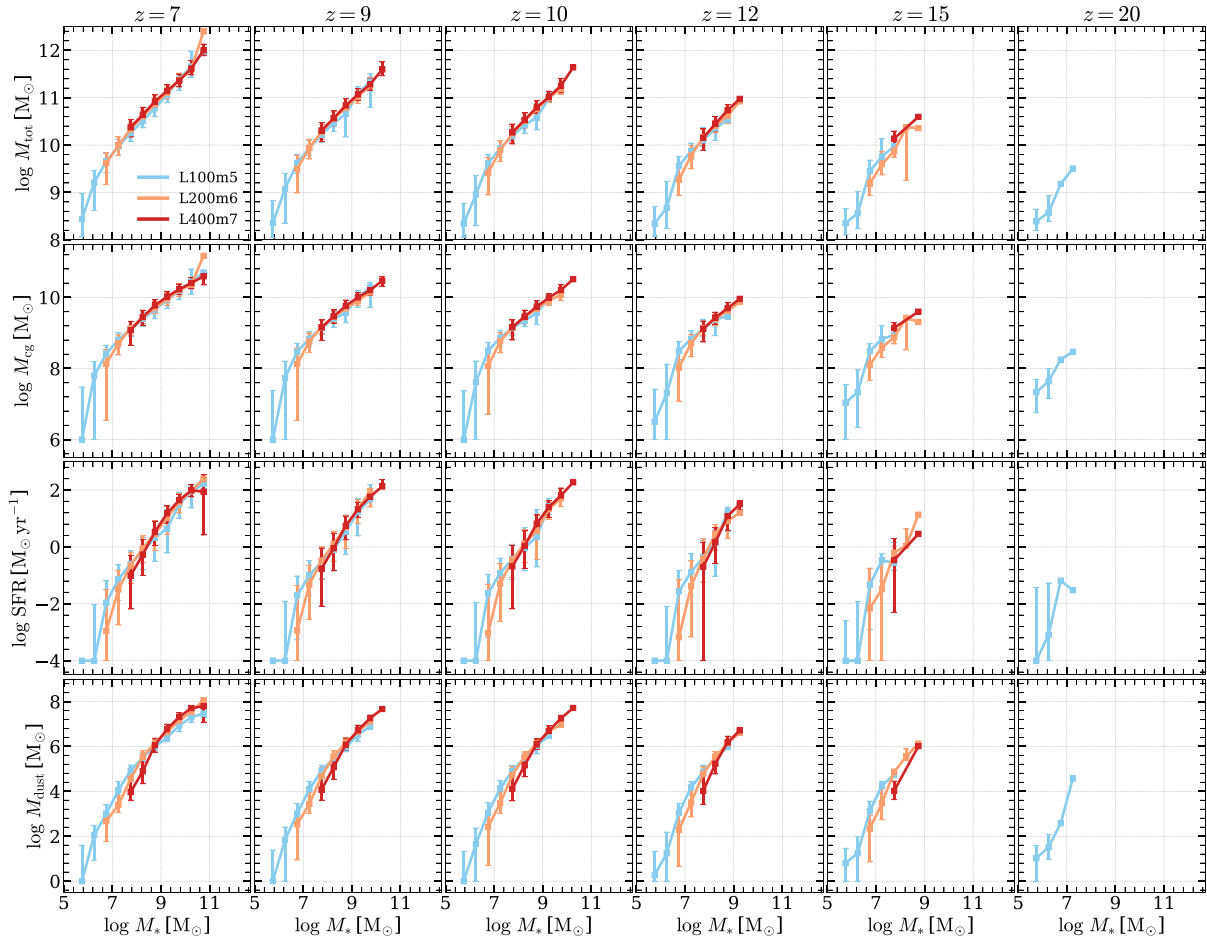


Figure 1. Correlations between galaxy stellar mass and other global properties: (1) the total mass (including baryons, dark matter, and black holes; M_{tot} ; top row); (2) the mass of gas in the cool, dense phase (i.e. $T < 10^{4.5}$ K and $\rho_{\text{g}}/m_{\text{H}} > 10^{-1}$ cm $^{-3}$), denoted as M_{cg} (second row); (3) the instantaneous SFR (third row), and (4) the total dust mass, denoted as M_{dust} (bottom row) of the galaxies (above certain stellar mass limits; see Table 1 for minimum stellar masses for COLIBRE simulations with different resolutions) from the three COLIBRE simulations at the selected redshifts. The correlations at different redshifts are shown in the different columns ($z = 7$ to 20 from left to right). The squares represent the median values of the corresponding quantity in the stellar mass bin, with the error bars indicating the 16th–84th percentile range (i.e. the $\pm 1\sigma$ region). Good convergence is seen for all relations among the three COLIBRE simulations.

the UVLF (see equation 2). The number of selected galaxies and their average fractions relative to the total galaxy population in the different COLIBRE simulations are listed in Table 2. Binning galaxies in the stellar mass–SFR plane, rather than by stellar mass alone, ensures that galaxies with different levels of star formation activity are represented within each mass range. This prevents the sample being dominated by either highly star-forming or quiescent systems and therefore better captures the diversity of UV emission that contributes to the UVLF. The target number of galaxies per bin is somewhat arbitrary and is chosen to balance computational cost with adequate sampling. We have verified that this sampling strategy provides sufficiently accurate estimates of the UVLFs across the whole redshift range considered here (see Appendix A).

In Fig. 2, we show the $\log M_* - \log \text{SFR}$ distribution of selected galaxy samples at $z = 7$ from the three COLIBRE simulations, colour-coded by their sampling fractions. Distributions of the selected samples at other redshifts ($z = 9, 10, 12, 15$, and 20) are shown in Appendix B (Fig. B1). As can be seen, rare galaxies with high stellar mass and high or low SFR have higher

sampling fractions, indicating a higher probability of being selected.

We adopt the purely thermal AGN feedback simulations as our default model. We have checked that the choice of AGN feedback model has little impact on the UVLFs at the redshifts of interest (see Appendix C1 for details).

2.3 Radiative transfer

We perform dust radiative transfer calculations for the samples of COLIBRE galaxies described above using the calibration-free COLIBRE-SKIRT pipeline of A. Gelek et al. (2026). Here, ‘calibration-free’ means that no SKIRT parameters were calibrated against luminosity-related observational constraints. Following A. Gelek et al. (2026), we use a rescaled B. T. Draine & A. Li (2007) dust model because it provides a better match to the local CSFD than the other dust models. However, as shown in Paper I, this choice has only a minor impact on the predicted LFs and is therefore not expected to significantly affect the UVLF results presented here, where dust is used only for attenuation. The COLIBRE-SKIRT pipeline is built upon the publicly available 3D

Table 2. The number of galaxies in the selected samples for calculation of the UVLFs using SKIRT at different redshifts in the three COLIBRE simulations. The percentages in parentheses represent the average sampling fraction (galaxies from different $\log M_*$ – $\log$ SFR grids have different sampling ratios) among galaxies above the stellar mass thresholds.

	L100m5	L200m6	L400m7
$z = 7$	3100 (0.44 per cent)	2415 (0.66 per cent)	1759 (0.82 per cent)
$z = 9$	2290 (0.56 per cent)	1781 (1.55 per cent)	1182 (3.30 per cent)
$z = 10$	2007 (0.71 per cent)	1468 (2.42 per cent)	949 (7.22 per cent)
$z = 12$	1428 (1.28 per cent)	971 (6.20 per cent)	521 (35.88 per cent)
$z = 15$	793 (4.67 per cent)	428 (33.59 per cent)	29 (100 per cent)
$z = 20$	158 (72.81 per cent)	–	–

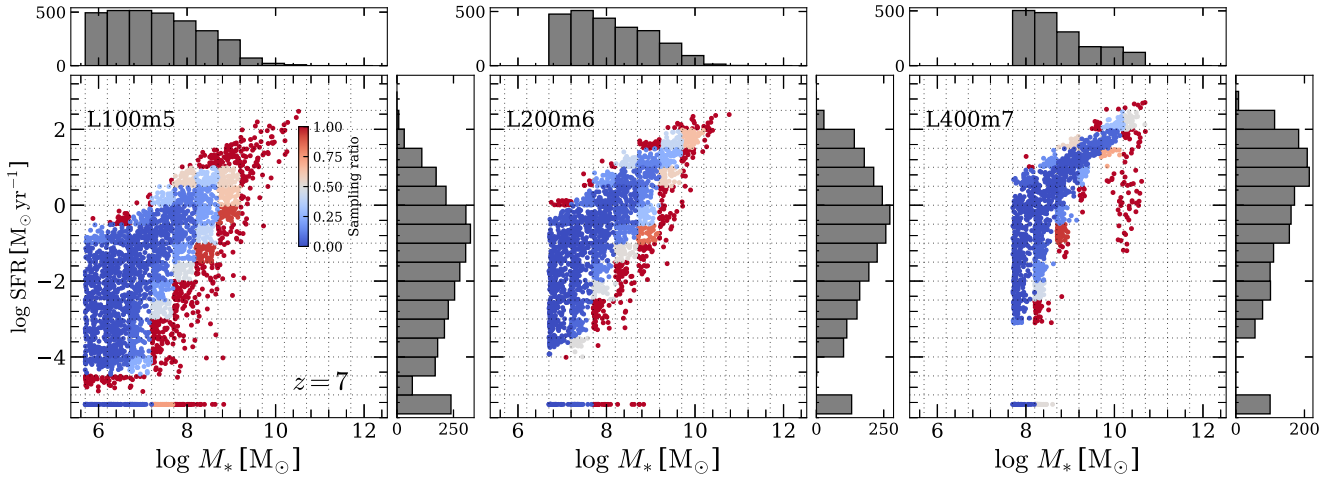


Figure 2. The distribution of galaxies selected for post-processing with SKIRT on the $\log M_* - \log$ SFR plane for L100m5, L200m6, and L400m7 (from left to right). For clarity, SFRs lower than $10^{-5.25} M_\odot \text{ yr}^{-1}$ are set to $10^{-5.25} M_\odot \text{ yr}^{-1}$. The grid delineated by dotted lines indicates the boundaries of the stellar mass and SFR bins used in the sample selection. The colours represent the sampling fraction of galaxies (i.e. the ratio between the number of galaxies selected from a bin and the total number of galaxies in that bin). The histograms show the distributions of stellar mass and SFR for the selected galaxies. Distributions at other redshifts are shown in Fig. B1.

Monte Carlo radiative transfer code SKIRT⁴ (M. Baes et al. 2011; P. Camps & M. Baes 2015, 2020). The code supports complex source and dust geometries derived directly from particle- or grid-based hydrodynamical simulations (M. Baes & P. Camps 2015) and employs adaptive spatial grids to efficiently resolve dust structures (P. Camps, M. Baes & W. Saftly 2013; W. Saftly et al. 2013; W. Saftly, M. Baes & P. Camps 2014).

A. Gebek et al. (2026) showed that the local CSED is broadly recovered by the calibration-free COLIBRE-SKIRT pipeline. In Paper I, we further showed that COLIBRE-SKIRT produces a galaxy population in remarkable agreement with measured local LFs from the FUV to the submillimetre, with the exception of the MIR, where emission of polycyclic aromatic hydrocarbons is important but not explicitly simulated by COLIBRE. Here, we broadly follow the SKIRT setup adopted in Paper I, which in turn follows the setup of A. Gebek et al. (2026), where the fiducial COLIBRE-SKIRT post-processing pipeline is described in full detail, with slight modifications to the SED wavelength range and the size of the dust discretization box, in order to reduce the computational cost. Below, we briefly summarize the setup.

(i) *Dust radiative transfer.* SKIRT self-consistently models dust absorption, scattering, heating and thermal re-emission (P.

Camps & M. Baes 2015), supporting various dust species, and complex geometries for both sources and dust distributions. Since we focus only on the UVLFs in this work and dust emission in the UV band is negligible, SKIRT is operated in `ExtinctionOnly` mode. The dust distribution is taken directly from COLIBRE, which tracks graphite and silicate grains of two representative sizes (large and small grains). Since COLIBRE provides only the masses of the small- and large-grain components for each dust particle, whereas SKIRT modelling requires a continuous grain size distribution, we reconstruct such a distribution by separately rescaling the small- and large-grain parts of an adopted dust model, divided at $0.03 \mu\text{m}$, such that their integrated masses match the corresponding dust masses predicted by the simulation. In this work, the continuous grain size distribution and the corresponding optical properties are adopted from B. T. Draine & A. Li (2007).

(ii) *Luminosity sources.* We decompose the luminosity sources into two components, evolved stars and star-forming regions. Stellar particles older than a threshold age of $t_R = 10 \text{ Myr}$ are treated as evolved stars and assigned simple stellar population templates from BPASS (v2.2.1; J. J. Eldridge et al. 2017; E. R. Stanway & J. J. Eldridge 2018), consistent with the early stellar feedback module adopted in COLIBRE (A. Benítez-Llambay et al. 2026) and assuming a G. Chabrier (2003) IMF, as in COLIBRE. Stellar populations younger than t_R are resampled using the star-forming gas (including the currently existing gas with

⁴Version 9; Available at <https://skirt.ugent.be>.

non-zero SFR averaged over time t_R and the parent gas of the young stellar particles) to better capture the emission from unresolved young stars. The emission from these star-forming regions is computed using the TODDLERS library (A. U. Kapoor et al. 2023, 2024), which models the time-dependent evolution of H II regions around young stars and includes the emission lines from the ionized gas in the H II regions. The TODDLERS library is also based on the BPASS (v2.2.1) library and is consistent with what we use for evolved stars. In this work, we adopt the dust-free version of TODDLERS, which currently does not include nebular continuum emission (see A. Gebek et al. 2026), such that attenuation by the dust distribution in COLIBRE is treated self-consistently in the subsequent SKIRT radiative transfer calculation. The luminosities of star-forming regions are determined using the SFR averaged over the preceding time interval t_R , assuming a constant star formation history over this period. As mentioned in Section 2.2, the lowest mass galaxies in our sample contain only 2–5 star particles. However, this limitation is partially removed in the SKIRT post-processing because these galaxies usually contain many more cool gas particles (the median ratio of the number of cool gas particles to star particles is ≈ 30). These cool gas particles enter both the resampling of very young (< 10 Myr) stellar populations and the modelling of the dust distribution for radiative transfer.

(iii) *Dust discretization box.* Given that galaxies at high redshift are much smaller than those at $z = 0$ and to optimize computational cost, we adopt a redshift-dependent discretization box for the dusty medium around each galaxy (unlike in Paper I, where a fixed-size discretization box of 100 kpc was used). This volume defines the spatial region within which the dust distribution is discretized onto the SKIRT grid. The dust discretization volume is centred at the location of the most bound particle and has a proper size given by

$$L_{\text{dust}} = 100 \text{ kpc} \times \frac{1.8}{0.7} \times \frac{1}{1+z}, \quad (1)$$

where z is the redshift. The factors 1.8 and 0.7 correspond to the comoving and maximum physical gravitational softening lengths, respectively, adopted in COLIBRE at m6 resolution. The factor $(1+z)^{-1}$ converts the comoving softening length into proper units. Although the comoving and maximum physical gravitational softening lengths vary with simulation resolution (see table 2 of J. Schaye et al. 2026), their ratio is fixed. We therefore adopt equation (1) for all simulations. We have verified that the resulting discretization box size is sufficiently large to encompass the dust distribution at high redshift in all simulations ($L_{\text{dust}} = 32.1$ kpc at $z = 7$ and 12.2 kpc at $z = 20$). In SKIRT, dust located outside this discretized medium volume is not included in the radiative transfer calculation. However, the radiation sources (stars and star-forming regions) are treated as gridless entities in SKIRT and are therefore not restricted to this volume, i.e. (gravitationally bound) stars and star-forming regions located outside the dust discretization box are still included.

(iv) The wavelength grid used here spans the range 0.09–0.2 μm , with 40 logarithmically spaced bins.

(v) We launch $10^{7.5}$ photon packets for each galaxy in the SKIRT computation. The photon packets are randomly sampled from the source emission spectra at their native spectral resolution (P. Camps & M. Baes 2020). We confirm that this number of photon packets is sufficient to resolve the SEDs of galaxies in the FUV band.

(vi) Galaxies are ‘observed’ with an idealized mock detector positioned along the z -axis of the simulation cube, corresponding to a random viewing orientation relative to the galaxy, at a distance of 10 Mpc. No point spread function (PSF) convolution or noise injection is applied. As a result, this procedure does not constitute a full end-to-end forward modelling of the observations, but instead yields idealized galaxy luminosities for direct comparison.

(vii) Luminosities are computed within a fixed projected circular aperture of proper radius⁵ 50 kpc. As shown in Fig. C3, the resulting UVLFs are virtually identical to those obtained when including all particles gravitationally bound to each galaxy, indicating that this aperture fully encompasses the galaxies at these redshifts. A detailed assessment of the impact of varying the aperture size on the UVLFs is presented in Appendix C3.

2.4 Observational data

The observational UVLFs used in this work for comparison include both new *JWST* measurements and earlier estimates from other facilities such as the *HST*. To compare our UVLFs with observational data at the same redshift, we adopt the following strategy: if an observational UVLF measurement provides a redshift range (i.e. lower and upper redshift limits) with a width no larger than 2, it is included when our target redshift (z_{target}) falls within this range, i.e. $z_{\text{obs,min}} \leq z_{\text{target}} \leq z_{\text{obs,max}}$ and $\Delta z \leq 2$, where $\Delta z \equiv z_{\text{obs,max}} - z_{\text{obs,min}}$ is the redshift bin width. If no redshift range is given and only a mean or median redshift (z_{obs}) is reported, the measurement is used only when $|z_{\text{obs}} - z_{\text{target}}| \leq 0.5$. Observational data with redshift bins wider than $\Delta z = 2$ are excluded to avoid contamination from lower redshift objects. As no observational measurements at $z = 20$ satisfy these selection criteria, we present the COLIBRE results at this redshift as genuine predictions to be tested against future observations. With this selection, we list in Table 3 the selected observational references at each redshift.

Differences in the *Hubble* constant assumed in the observations can introduce systematic biases in the calculation of galaxy distances, affecting both absolute magnitudes and number densities. To ensure a consistent comparison between our simulations and observational data, we rescale all absolute magnitudes in the observational data to the *Hubble* constant used in COLIBRE, using $M + 5 \log_{10}(h_{\text{sim}}/h_{\text{obs}})$, where h_{obs} is the *Hubble* constant stated in the observational study and h_{sim} is the *Hubble* constant in COLIBRE ($h_{\text{sim}} = 0.681$). In addition, all number densities are scaled to the *Hubble* constant in COLIBRE, through $\log_{10} \Phi + 3 \log_{10}(h_{\text{sim}}/h_{\text{obs}})$. This standard convention effectively removes the dependence on the *Hubble* constant in our comparison of simulations with observations and allows for a fair comparison across data sets adopting different cosmologies. We note that variations in other cosmological parameters, such as Ω_m , can also influence the inferred LFs by altering the luminosity distance and comoving volume estimates. However, these effects are generally small compared to those from the differences in h and they cannot be readily corrected using a simple analytic rescaling. For this reason, we do not apply an explicit correction for differences in Ω_m in this work.

Collecting observational data from multiple studies can be time-consuming and repetitive. To support reproducibility and make future comparisons easier, we provide all observational

⁵Note that only gravitationally bound particles are considered here.

Table 3. Observational data on the FUV luminosity functions used at each redshift (see Section 2.4 for our criteria for selecting the observational data). To make fair comparisons, all observational magnitudes and number densities are rescaled as described in the text.

Redshift	Pre- <i>JWST</i> references	<i>JWST</i> references
7	R. J. Bouwens et al. (2011, 2015, 2021, 2022); R. J. McLure et al. (2013); M. A. Schenker et al. (2013); R. A. A. Bowler et al. (2014); S. L. Finkelstein et al. (2015); M. Stefanon et al. (2019); and Y. Harikane et al. (2022)	R. Bouwens et al. (2023); R. A. Meyer et al. (2024); M. Franco et al. (2025); Y. Harikane et al. (2025); and S. Rojas-Ruiz et al. (2025)
9	R. J. McLure et al. (2013); R. J. Bouwens et al. (2015, 2021, 2022); T. Morishita et al. (2018); M. Stefanon et al. (2019); and R. A. A. Bowler et al. (2020)	R. Bouwens et al. (2023); R. J. Bouwens et al. (2023); C. T. Donnan et al. (2023); Y. Harikane et al. (2023, 2024, 2025); G. C. K. Leung et al. (2023); P. G. Pérez-González et al. (2023); N. J. Adams et al. (2024); C. T. Donnan et al. (2024); S. L. Finkelstein et al. (2024); R. A. Meyer et al. (2024); C. J. Willott et al. (2024); M. Franco et al. (2025); S. Rojas-Ruiz et al. (2025); and I. Chemerynska et al. (2026)
10	P. A. Oesch et al. (2014, 2018); R. J. Bouwens et al. (2015); T. Morishita et al. (2018); and M. Stefanon et al. (2019)	R. Bouwens et al. (2023); R. J. Bouwens et al. (2023); M. Castellano et al. (2023); C. T. Donnan et al. (2023, 2024); P. G. Pérez-González et al. (2023); N. J. Adams et al. (2024); Y. Harikane et al. (2024, 2025); C. J. Willott et al. (2024); M. Franco et al. (2025); S. Rojas-Ruiz et al. (2025); and I. Chemerynska et al. (2026)
12	–	R. Bouwens et al. (2023); R. J. Bouwens et al. (2023); Y. Harikane et al. (2023); P. G. Pérez-González et al. (2023); N. J. Adams et al. (2024); C. T. Donnan et al. (2024); B. Robertson et al. (2024); and C. J. Willott et al. (2024)
15	–	C. M. Casey et al. (2024); C. T. Donnan et al. (2024); S. L. Finkelstein et al. (2024); B. Robertson et al. (2024); Y. Harikane et al. (2025); and R. P. Naidu et al. (2026)
20	–	–

Notes. All observational UVLFs used in this work for comparison (uncorrected and together with their corresponding h_{obs} values) are available in machine-readable format on both the journal website and <https://icc.dur.ac.uk/data/>.

data sets used in this series of papers in machine-readable format on both the journal website and <https://icc.dur.ac.uk/data/>. For broader discoverability, these data sets are also listed on the DigitizeAstronomy website (<https://digastro.org/>), a community platform for sharing machine-readable data extracted from the astronomical literature. We hope this will be a useful resource for the community and encourage similar practices in future work.

3 RESULTS

3.1 Correlations between galaxy properties and UV magnitude

The SED (or spectrum) from SKIRT is the sum of emission from evolved stars and star-forming regions processed by the dust integrated over a given projected aperture (50 kpc radius in this work). Having the output SEDs from SKIRT, we follow the practice adopted in studies using *JWST* data (e.g. S. L. Finkelstein et al. 2024) and obtain the UV magnitude at $\lambda = 1500 \text{ \AA}$ by convolving the SED with a top-hat filter spanning 1450–1550 $\text{\AA}$. Throughout this work, we denote M_{UV} as the derived rest-frame AB absolute magnitude of galaxies.

In Fig. 3, we present the correlations between stellar mass and M_{UV} (both dust-attenuated and dust-free) for the selected galaxies over $z = 7\text{--}20$ in the three COLIBRE simulations listed in Table 1. At $z = 20$, only the L100m5 result is shown, as only three galaxies in L200m6 and none in L400m7 satisfy the stellar mass limit. As expected, at fixed redshift, both dust-attenuated and dust-free M_{UV} on average become brighter with increasing stellar mass. For both dust-attenuated and dust-free results, the three COLIBRE

simulations are in good agreement in terms of the $M_{\text{UV}} - M_*$ relation down to the lower stellar mass limit of the selected samples (see Table 1), indicating good numerical convergence between m5 and m6 and slightly poorer convergence for m7 (with the dust-free results showing even better consistency). Furthermore, the dust-attenuated and dust-free M_{UV} begin to diverge at $\log M_*/M_\odot \sim 7\text{--}8$ (except at $z = 20$, where they remain nearly identical), with the difference increasing toward higher stellar masses. This trend is expected because, on average, massive galaxies are more metal-enriched (C. A. Tremonti et al. 2004) and typically exhibit higher gas surface densities, which together imply larger dust column densities in their ISM and thus stronger attenuation. However, geometry, inclination and the relative star–dust distribution introduce significant scatter. An interesting feature is that the dust-attenuated $M_{\text{UV}} - M_*$ relation in L400m7 at $z = 7$ and 9 shows an ‘upturn’ at the highest stellar masses. This trend deserves further investigation in future work.

3.2 The UV luminosity functions

3.2.1 UVLFs for three COLIBRE simulations

In this work, COLIBRE UVLFs are computed in bins of width $\delta M = 0.5 \text{ mag}$, while the results are reported as galaxy number densities per magnitude. In each bin, the galaxy abundance is thus calculated as

$$\Phi_M = \frac{\sum_i 1/\eta_i}{\delta M \times L_{\text{box}}^3}, \quad (2)$$

where η_i is the sampling ratio (see Section 2.2) of the i th galaxy and L_{box} is the side length of the simulation cube. The

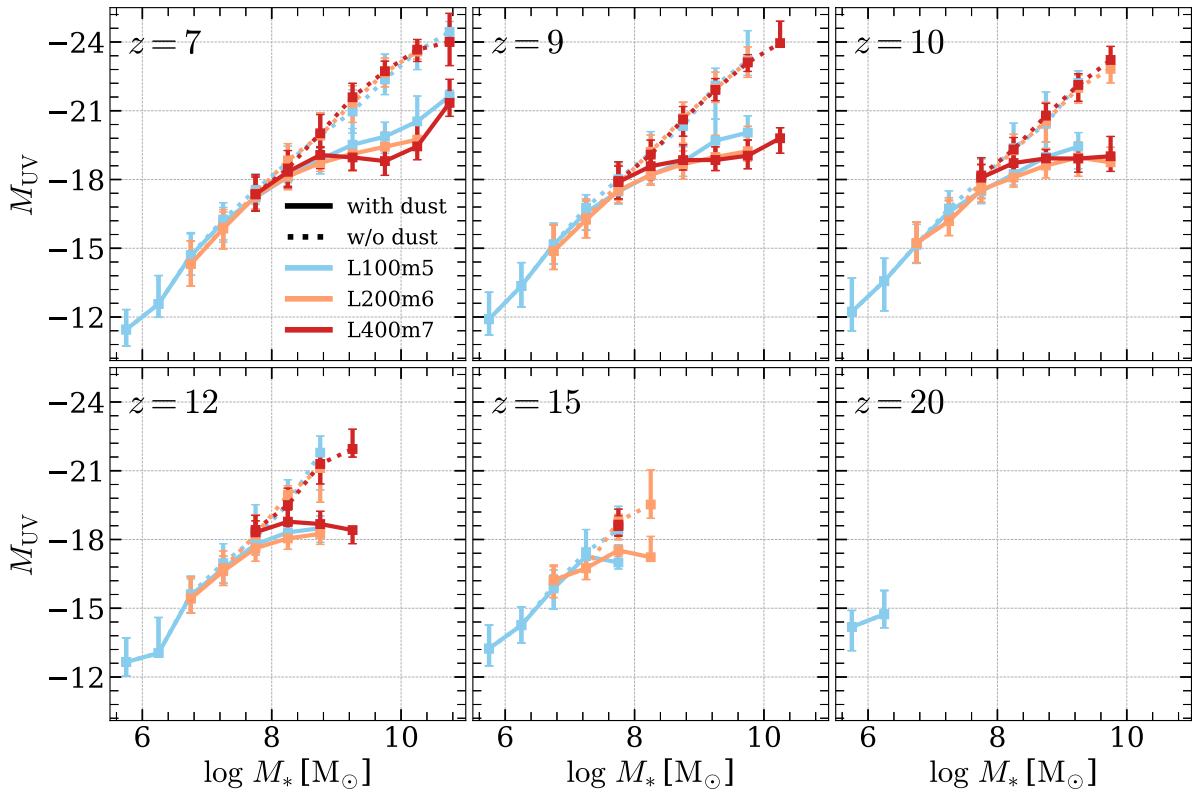


Figure 3. Correlations between the stellar mass of galaxies ($\log M_*$) and their UV magnitude, M_{UV} (solid: dust-attenuated; and dashed: dust free) at different redshifts (as indicated in each panel). Results for L100m5, L200m6, and L400m7 are indicated by different colours. The squares represent the median M_{UV} in the stellar mass bins, with the error bars indicating the interval from the 16th to 84th percentiles of M_{UV} in each stellar mass bin. COLIBRE simulations with different resolutions show consistent $M_{UV} - \log M_*$ relations across the redshift range $z = 7-15$, especially for dust-free M_{UV} .

summation is taken over all galaxies within the given M_{UV} bin. We confirm that the choice of magnitude bin width (e.g. 0.25, 0.5, 0.75, and 1) has only a minor influence on the UVLFs. In Fig. 4, we present the predicted UVLFs from three COLIBRE simulations (indicated with different colours) at $z = 7-20$. The predicted UVLFs in Fig. 4 are calculated from the dust-attenuated UV magnitudes from SKIRT. Observational data (see Section 2.4 for the selection criteria and Table 3 for the selected data sets) are overplotted for reference, but a detailed comparison is deferred to Section 3.2.3.

As shown in Fig. 4, the UVLFs from different COLIBRE simulations at varying resolutions exhibit good agreement over certain magnitude ranges. However, at magnitudes fainter than $M_{UV} \approx -18$, the UVLFs of L400m7 begin to deviate from those of the higher resolution simulations. Similarly, at magnitudes fainter than $M_{UV} \approx -16$, the UVLFs of L200m6 gradually diverge from those of L100m5. This is expected, as we limit our sample to galaxies with stellar mass above a certain threshold (roughly corresponding to 2–5 stellar particles; see Table 1), potentially leading to incompleteness at the faint end of each UVLF. When compared to Fig. 3, we find that the M_{UV} thresholds at which different COLIBRE simulations start to diverge roughly correspond to the median M_{UV} of galaxies near the stellar-mass lower limit. This implies that the UVLFs are already approximately converged even in galaxies containing only a few stellar particles. The good agreement between the COLIBRE simulations also suggests that the UVLFs are converged with respect to the simulation volume over the magnitude ranges considered.

3.2.2 Combining UVLFs from three simulations and estimating the errors

To construct a representative UVLF from COLIBRE across a wide magnitude range at each redshift, we combine the UVLFs from different COLIBRE simulations using the following procedure:

(i) Among the three COLIBRE simulations, we adopt the UVLFs from L200m6 as the default reference. Starting from $M_{UV} = -17.5$ and moving toward the fainter end, we compare the galaxy abundances from L200m6 and the higher resolution simulation, L100m5. If the difference between the two exceeds a threshold (0.25 dex) in two consecutive M_{UV} bins, we switch to using the UVLFs from L100m5 for all subsequent (fainter) bins. We note that we adopt $M_{UV} = -17.5$ as the starting point as it ensures sufficient reliability from both the numerical and statistical perspectives. According to Fig. 3, in L200m6, $M_{UV} = -17.5$ corresponds to a stellar mass of $\approx 10^{7.6} M_\odot$, which consists of ≈ 20 stellar particles. Additionally, as shown in Fig. 4, $M_{UV} = -17.5$ corresponds to a galaxy abundance of $\sim 10^{-4.4} \text{ Mpc}^{-3} \text{ mag}^{-1}$ at $z = 15$ (and even higher at lower redshifts) in L200m6, which translates to about 320 galaxies per magnitude in the L200 simulation cube. Practically, this choice also facilitates the construction of a smooth combined LF, which would be more difficult to obtain at other magnitudes. Therefore, this starting point guarantees both adequate resolution and robust statistical sampling.

(ii) Similarly, starting from $M_{UV} = -17.5$ and moving toward the brighter end, we compare the UVLFs from L200m6 and the simulation in the larger volume, L400m7. Again, if the difference

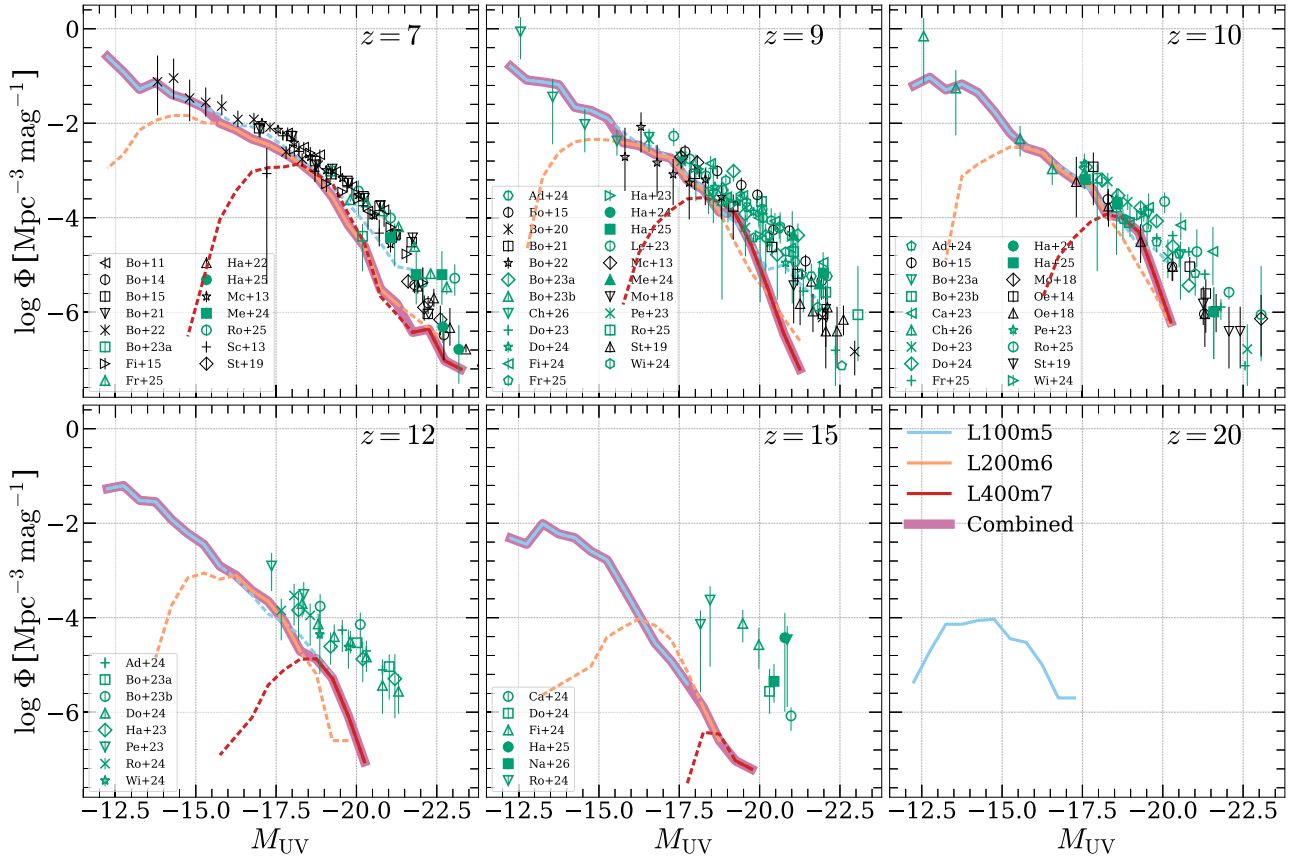


Figure 4. Comparison of (dust-attenuated) UVLFs between the predictions of COLIBRE (consisting of the UVLFs from three individual COLIBRE simulations and their combined UVLFs; see Section 3.2.2 for details) and observational data. We remind readers that the redshifts here correspond to the COLIBRE snapshots and are not the average redshifts of the observational data. Please see Section 2.4 for the selection criteria for the observational data and Table 3 for the list of selected observational data sets at each redshift. For the UVLFs from the three individual COLIBRE simulations, we use dashed lines to indicate those portions that are not included in the combined UVLFs. In each panel, we overplot observational data, grouped into three categories: (i) *JWST* data with spectroscopically confirmed redshifts (filled green symbols), (ii) *JWST* with photometric redshifts only (open green symbols), and (iii) photometric redshift-based measurements from earlier observations, such as *HST* (open black symbols). For clarity, we exclude data points labelled as upper or lower limits in the original paper.

exceeds 0.25 dex in two consecutive M_{UV} bins, we switch to using the UVLFs from L400m7 for all subsequent (brighter) bins.

The combined UVLFs are shown as the thick purple curves in Fig. 4, while measurements from individual COLIBRE simulations that are not used in the combination are shown as dashed curves.

The measurement uncertainties of the UVLFs for individual COLIBRE simulations, as well as for the combined UVLFs, are estimated using the bootstrap method. Since the UVLFs are computed from a subsample of galaxies (see Section 2.2), we assessed the statistical uncertainty arising from this subsampling by performing 1000 bootstrap resamplings of both the parent sample (i.e. the full set of galaxies in COLIBRE before any selection) and the selected subsample (for SKIRT modelling) for each COLIBRE simulation and each redshift. In each bootstrap iteration, we first randomly draw galaxies from the parent sample with replacement to create a ‘mock’ parent sample of the same size. Then, we randomly draw galaxies from the selected subsample (see Section 2.2) in each stellar mass–SFR bin, also with replacement, to create a ‘mock’ subsample (this procedure preserves the number of selected galaxies in each bin). For each

bootstrap iteration, we then recalculate the sampling ratio (i.e. the ratio between the number of selected galaxies and the total number of galaxies in the parent sample within each bin) for each (re)selected galaxy, which may appear multiple times in the bootstrap resampling, and then compute the corresponding ‘mock’ UVLF. Finally, the uncertainty in the UVLF is quantified as the standard deviation derived from the 1000 mock UVLFs. By bootstrapping both the parent sample and the selected subsample, we naturally capture both the fluctuations in the sampling ratio and the random variations in the detailed properties of the galaxies included in the selected subsample, reflecting statistical uncertainties arising from both the representativeness of the parent population and the stochasticity of the selection. These variations introduce additional fluctuations in the estimated UVLFs, which are used to quantify the measurement uncertainties.

The uncertainty in the combined UVLF is then obtained by applying the same combination strategy described in the previous paragraphs to each of the 1000 individual mock UVLFs. This yields 1000 mock combined UVLFs and the uncertainty in the combined UVLF is also quantified as the standard deviation of these 1000 realizations.

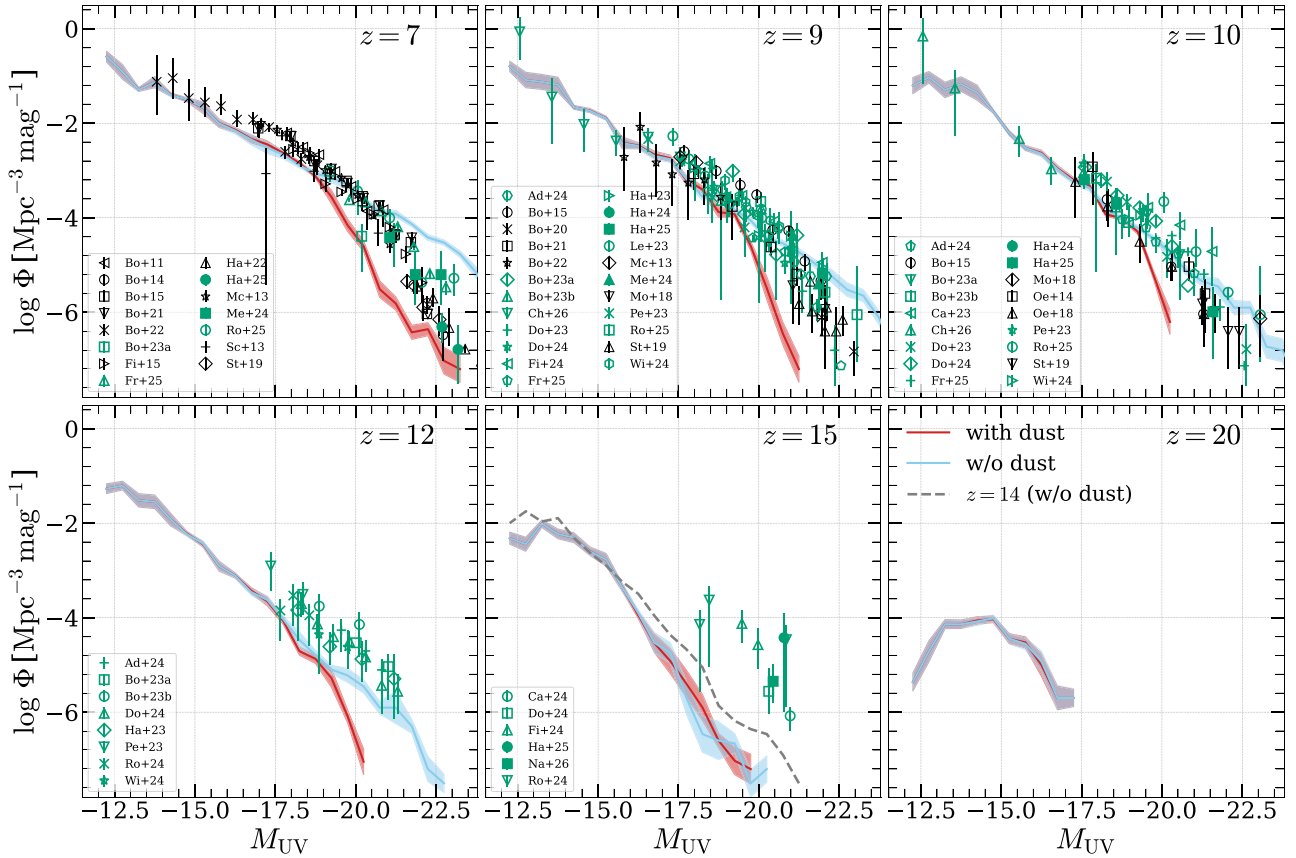


Figure 5. Comparison of the combined dust-attenuated and dust-free COLIBRE UVLFs with observational data. The shaded regions represent the $\pm 1\sigma$ ranges, calculated using the bootstrap method by perturbing both the parent sample (from which our random sample processed with SKIRT is drawn) and the selected sample 1000 times (see Section 3.2.2 for details). The grey dashed curve in the $z = 15$ panel shows the dust-free UVLF of COLIBRE at $z = 14$, providing a more precise comparison with observations that are based on samples with redshifts close to 14. The symbols for the observational data are as in Fig. 4.

3.2.3 Comparison with observations

The observational data used for comparison are selected as described in Section 2.4 and listed in Table 3. In Fig. 5, we compare the COLIBRE UVLFs with dust attenuation (red), including the measurement uncertainties estimated using the method described above, with the observational data. UVLFs constructed from earlier, pre-*JWST* data, which are all based on photometric redshifts, are indicated by open black symbols. UVLFs derived from *JWST* observations are shown as green symbols, with filled symbols indicating measurements based on fully spectroscopically confirmed high- z candidates, including those from Y. Harikane et al. (2024, 2025); R. A. Meyer et al. (2024); and R. P. Naidu et al. (2026)⁶ and open symbols indicating measurements based on photometric redshifts from *JWST*. To avoid confusion in the figure, we exclude data points reported as upper or lower limits. It can be seen that the measurements from *JWST* agree well with pre-*JWST* observations and that the results based on spec-

troscopically confirmed objects are consistent with those based on photometric redshifts.

As shown in Fig. 5, the galaxy abundance predicted by COLIBRE drops by a factor of $\sim 10^5$ from $M_{UV} \approx -15$ to ≈ -22 at $z = 7$ when dust attenuation is included (on average, $\Delta \log \Phi / \Delta M_{UV} \approx 0.7$ dex per magnitude). At higher redshift, e.g. $z = 15$, the decreasing trend becomes steeper with the galaxy abundance dropping by a factor of $\sim 10^4$ from $M_{UV} \approx -15$ to ≈ -19 ($\Delta \log \Phi / \Delta M_{UV} \approx 1$ dex per magnitude). This difference is most clearly seen in the upper left panel of Fig. 6, where the combined UVLFs at different redshifts are overplotted for comparison. At $z = 20$, the UVLF derived from L100m5 shows a decline similar to that at $z = 15$ for $M_{UV} < -15$, while at the faint end, it is affected by sample incompleteness.

When compared to the observed UVLFs, the COLIBRE dust-attenuated UVLFs show faint-end slopes at $z = 7-12$ that are broadly consistent with those inferred from observations, while the bright end is noticeably steeper. At $z = 7$, the COLIBRE dust-attenuated UVLF is systematically lower than the observations at the bright end. Equivalently, COLIBRE-SKIRT underpredicts the luminosities of the most UV-luminous galaxies, with an offset of ≈ 1 mag at a number density of $10^{-6} \text{ Mpc}^{-3} \text{ mag}^{-1}$. This number density corresponds to 8 galaxies per unit magnitude in the L200m6 simulation box and 64 galaxies per unit magnitude in the L400m7 simulation box. At $z = 9$ and 10, the COLIBRE UVLFs also agree fairly well with the observations at the faint end, while

⁶The UVLFs of R. A. Meyer et al. (2024) are based on First Reionization Epoch Spectroscopically Complete Observations (FRESCO) spectroscopically confirmed $H\beta + [O III]$ emitters. Since this $[O III]$ -selected sample is incomplete at the faint end owing to line-detection sensitivity and scatter in the $[O III]$ -to-UV ratio, we follow Y. Harikane et al. (2025) and show only the bright points with $M_{UV} \leq -21$.

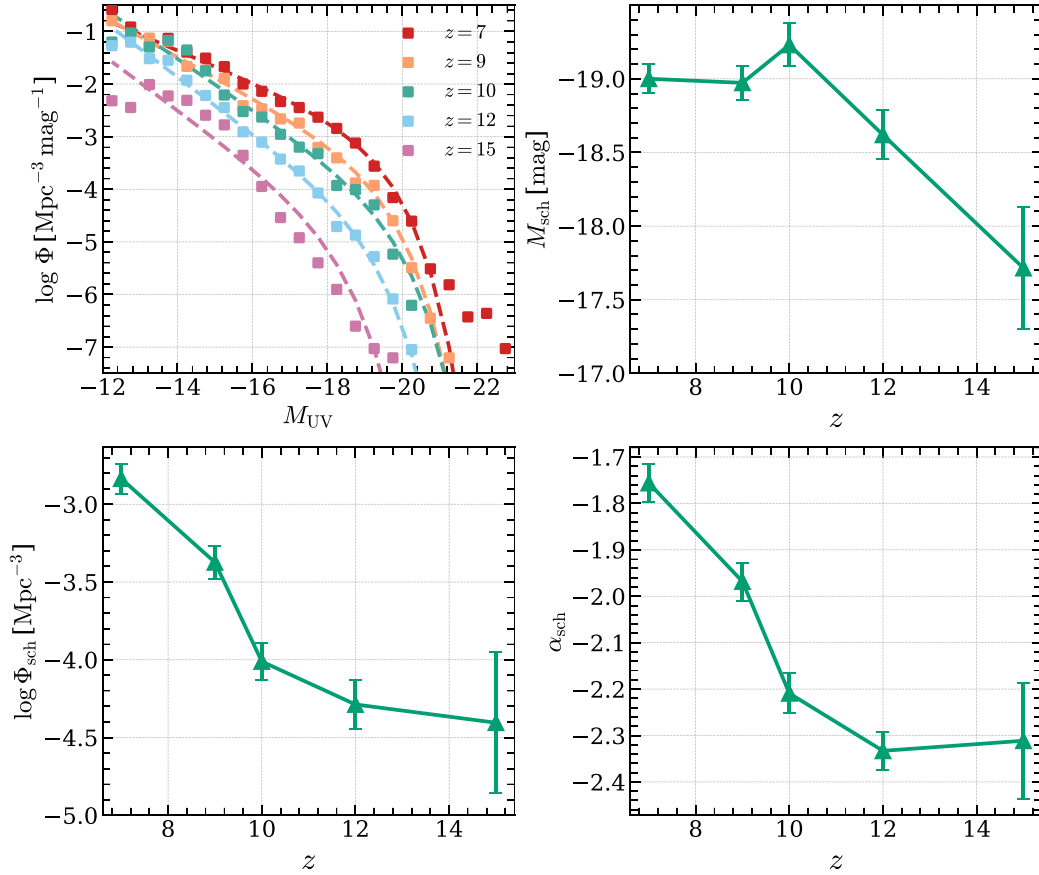


Figure 6. Upper left: P. Schechter (1976) function fits to the combined UVLFs in the COLIBRE simulations across redshifts $z = 7$ to 15 (see Section 3.2.2 for details on the combination method). Squares indicate galaxy number densities in each magnitude bin and dashed curves the corresponding Schechter fits. Other panels: redshift evolution of the Schechter parameters—characteristic absolute magnitude (M_{sch} ; upper right), characteristic number density ($\log \Phi_{\text{sch}}$; lower left), and faint-end slope (α_{sch} ; lower right). Error bars represent $\pm 1\sigma$ uncertainties, estimated with the bootstrap method (see Section 3.3 for details).

again underpredicting the luminosities of the most UV-luminous galaxies, with offsets of ≈ 1 mag at $z = 9$ and ≈ 1.2 mag at $z = 10$ at the same number density as above. The offset increases to ≈ 2.5 mag at $z = 15$.

The galaxy abundance at fixed M_{UV} decreases with redshift. For example, in COLIBRE, at $M_{\text{UV}} = -19.5$, it decreases from $\sim 10^{-4.5} \text{ Mpc}^{-3} \text{ mag}^{-1}$ at $z = 9$ to $\sim 10^{-6} \text{ Mpc}^{-3} \text{ mag}^{-1}$ at $z = 12$, indicating strong redshift evolution ($\Delta \log \Phi / \Delta z \approx 0.5$ at $M_{\text{UV}} = -19.5$; see also the upper left panel of Fig. 6), which is inconsistent with observations (e.g. Y. Harikane et al. 2023; S. L. Finkelstein et al. 2024) for which a milder redshift evolution of the UVLF at high- z is reported ($\Delta \log \Phi / \Delta z \approx 0.12$ at $M_{\text{UV}} \approx -19.5$ from $z = 9$ to 14 in S. L. Finkelstein et al. 2024; $\Delta \log \Phi / \Delta z \approx 0.05$ at $M_{\text{UV}} \approx -19.5$ from $z = 9$ to 12 in Y. Harikane et al. 2023⁷).

In Fig. 5, we also show the comparison between dust-free COLIBRE UVLFs (with measurement uncertainties) and the observations (which automatically include the effects of any dust attenuation present). We note that the dust-free results here are also from SKIRT modelling, including the emission from evolved stellar particles and the star-forming regions, but ignoring dust attenuation. We can see from Fig. 5 that the dust-free UVLFs dif-

fer from the dust-attenuated results only at the bright end, while the faint end remains nearly unchanged. This trend is consistent with the results shown in Fig. 3, where the difference between dust-attenuated and dust-free M_{UV} clearly increases with stellar mass.

At $z = 7$, the dust-free UVLFs are slightly lower than the observations at the faint end but substantially higher at the bright end. We note, however, that a completely dust-free scenario is inconsistent with observational results at this redshift, as Atacama Large Millimeter Array (ALMA) observations of $z \approx 7$ galaxies (e.g. the Reionization Era Bright Emission Line Survey programme; known as REBELS) already reveal significant rest-frame infrared emission, implying the presence of dust (L. Barrufet et al. 2023). At $z = 9$, the dust-free UVLF of COLIBRE agrees with the observations at the faint end, as does the dust-attenuated result, but is higher than the observations at the bright end. At $z = 10$, the dust-free COLIBRE UVLF roughly matches the observed UVLF across the whole luminosity range. At $z = 12$, the dust-free UVLF lies slightly below the observational data, while at $\Phi = 10^{-6} \text{ Mpc}^{-3} \text{ mag}^{-1}$, the dust-free UV luminosities of COLIBRE galaxies roughly match those inferred from the observations. At $z = 15$, even the dust-free UVLF remains lower than the observations over the entire luminosity range, meaning that COLIBRE-SKIRT underpredicts the luminosities of the most UV-luminous galaxies even when dust attenuation is ignored. The

⁷For Y. Harikane et al. (2023), we only provide a rough estimate of the galaxy abundance at $M_{\text{UV}} \approx -19.5$ for $z = 9$ and 12.

Table 4. The best-fitting Schechter parameters and the derived (dust-attenuated and dust-free) UV luminosity densities of the combined UVLFs of COLIBRE from $z = 7$ to 15. The fitting is performed over the range $-21 \leq M_{UV} \leq -12$. The measurement uncertainties of the Schechter parameters and UV luminosity density are evaluated using the bootstrap method (1000 iterations). The UV luminosity density is calculated by summing the UV luminosities of all eligible galaxies in the COLIBRE simulations, rather than by integrating the Schechter-function fits to the UVLFs (see Section 3.3 for more details).

z	M_{sch} [mag]	$\log \Phi_{sch}$ [Mpc $^{-3}$]	α_{sch}	$\log \rho_{UV}(\text{with dust})$ [erg s $^{-1}$ Hz $^{-1}$ Mpc $^{-3}$]	$\log \rho_{UV}(\text{w/o dust})$ [erg s $^{-1}$ Hz $^{-1}$ Mpc $^{-3}$]
7	-19.00 ± 0.10	-2.84 ± 0.10	-1.76 ± 0.04	25.46 ± 0.02	25.98 ± 0.02
9	-18.97 ± 0.11	-3.37 ± 0.10	-1.97 ± 0.04	24.90 ± 0.03	25.28 ± 0.02
10	-19.23 ± 0.15	-4.01 ± 0.12	-2.21 ± 0.04	24.57 ± 0.03	24.86 ± 0.02
12	-18.62 ± 0.16	-4.28 ± 0.16	-2.33 ± 0.04	23.89 ± 0.05	24.14 ± 0.04
15	-17.71 ± 0.41	-4.40 ± 0.45	-2.31 ± 0.12	22.60 ± 0.17	22.75 ± 0.14

luminosity offset is ≈ 2.5 mag at $\Phi = 10^{-6}$ Mpc $^{-3}$ mag $^{-1}$, similar to the dust-attenuated case, since the dust content at $z = 15$ is negligible.

We note that some of the observational data used for the comparison at $z = 15$ are actually based on galaxies at slightly lower redshifts. For example, the spectroscopic constraints from R. P. Naidu et al. (2026) rely on sources with $z_{spec} \approx 14$ (S. Carniani et al. 2024; R. P. Naidu et al. 2026). Motivated by this, we include the COLIBRE UVLF at $z = 14$, also calculated with SKIRT, in the $z = 15$ panel, shown as the grey dashed curve. The COLIBRE output at $z = 14$ is a ‘snapshot’ and does not contain the dust-species information required for SKIRT modelling. We therefore only present the dust-free UVLF at $z = 14$, noting that by $z \approx 15$ the dust-free UVLF is already very similar to the dust-attenuated one. As shown, while the COLIBRE UVLF at $z = 14$ is closer to the observations, the most UV-luminous galaxies are still fainter than observed, with an offset of ≈ 1.5 – 2 mag at $\Phi = 10^{-6}$ Mpc $^{-3}$ mag $^{-1}$.

More recently, K. C. Kreilgaard et al. (2026) and D. J. McLeod et al. (2026) reported lower UVLFs than many previous *JWST* photometric studies and hence results closer to the COLIBRE predictions. However, their redshift bins are not directly matched to those adopted here. This is particularly important since the UVLF evolves rapidly at these redshifts, especially beyond $z \approx 12$ (D. J. McLeod et al. 2026). In the highest redshift bin of D. J. McLeod et al. (2026), the detected candidates lie well below the upper edge of the adopted redshift interval; the resulting effective volume is therefore large, naturally leading to lower inferred number densities. This illustrates why such broad redshift bins cannot be used for a direct comparison with our target redshifts, even when the target redshift lies within the bin. We therefore do not include these measurements in our observational compilation, following the selection criteria stated in Section 2.4.

3.3 Schechter fitting and UV luminosity density

The P. Schechter (1976) function is widely used to parametrize the shape of observed LF at all wavelengths. The evolution of the LF is then often described in terms of the evolution of the Schechter function parameters. In this section, we fit Schechter functions to our combined (dust-attenuated) UVLFs at all the redshifts (except at $z = 20$, where the result is only from L100m5 and the bright end is not constrained) and study the redshift evolution of the Schechter parameters. The number density of galaxies within a luminosity range $[L, L + dL]$ according to the P. Schechter (1976)

function can be written as:

$$\Phi(L) dL = \Phi_{sch} \left(\frac{L}{L_{sch}} \right)^{\alpha_{sch}} \exp\left(-\frac{L}{L_{sch}}\right) \frac{dL}{L_{sch}}, \quad (3)$$

where $\Phi(L)$ denotes the number density of galaxies with luminosity L , and Φ_{sch} , L_{sch} , and α_{sch} represent the characteristic number density, characteristic luminosity, and faint-end slope of the UVLF, respectively. In this work, we adopt the magnitude-space formulation of the Schechter function, which can be written as:

$$\Phi(M) dM = 0.4 \ln(10) \Phi_{sch} 10^{-0.4(\alpha_{sch}+1)(M-M_{sch})} \times \exp[-10^{-0.4(M-M_{sch})}] dM, \quad (4)$$

where M_{sch} is the characteristic absolute magnitude.

We perform non-linear least-squares fitting over the M_{UV} range from -12 to -21 . During the fitting, each data point is assigned a weight inversely proportional to the square of its measurement error (in logarithmic space) described in Section 3.2.2, so data points with smaller errors have a larger influence on determining the best-fitting parameters. The derived Schechter parameters at different redshifts are listed in Table 4. In the upper left panel of Fig. 6, we show the combined UVLF predictions for different redshifts from COLIBRE with squares and the best-fitting Schechter functions with dashed curves. As can be seen, at $z = 7$, the extreme bright end of the COLIBRE UVLF cannot be well described by a Schechter function and there appears to be an excess of galaxies in these bins. This is consistent with the ‘upturn’ seen in the $M_{UV} - M_*$ relation of L400m7 at $z = 7$ (see Fig. 3).

In the other panels, we show the redshift evolution of the three Schechter parameters: the characteristic magnitude (M_{sch} ; upper right), the characteristic number density ($\log \Phi_{sch}$; lower left), and the faint-end slope (α_{sch} ; lower right). To obtain the measurement uncertainties of the Schechter parameters, we apply Schechter fitting to the 1000 mock combined UVLFs at each redshift (see Section 3.2.2 for details about obtaining the mock UVLFs) and calculate the standard deviation of the 1000 sets of Schechter parameters. As can be seen, all three parameters exhibit clear evolutionary trends with redshift, which are generally monotonic. Specifically, the characteristic absolute magnitude generally becomes fainter with increasing redshift, the faint-end slope tends to steepen and the characteristic number density generally declines.

We note that we did not compare with observational estimates of the Schechter parameters here, as many observational studies adopt different methods when applying Schechter fitting, for example, fixing the faint-end slope (e.g. Y. Harikane et al. 2023) or fixing both the faint-end slope and the characteristic magnitude

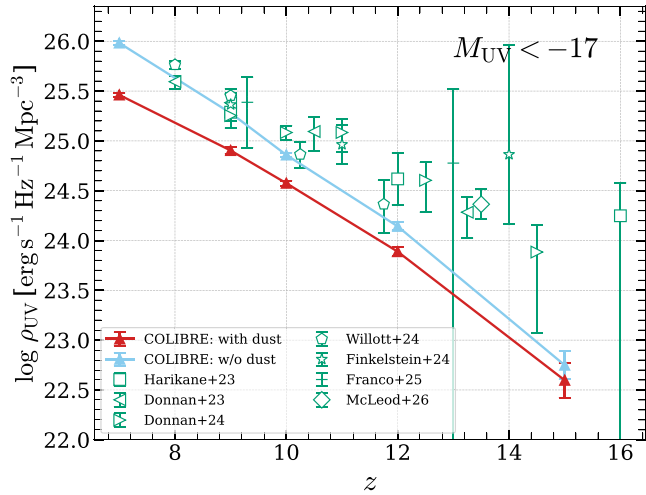


Figure 7. Redshift evolution of the UV luminosity density, ρ_{UV} , from COLIBRE (including all galaxies brighter than $M_{UV} = -17$) from $z = 7$ to 15, compared with *JWST* observations (dust-attenuated, including C. T. Donnan et al. 2023, 2024; Y. Harikane et al. 2023; S. L. Finkelstein et al. 2024; C. J. Willott et al. 2024; M. Franco et al. 2025; D. J. McLeod et al. 2026). The red line shows the dust-attenuated COLIBRE UV luminosity density and the blue line shows the dust-free results. The error bars represent the $\pm 1\sigma$ uncertainties, estimated with the bootstrap method by perturbing the parent sample (from which our random sample for SKIRT modelling is drawn) and the selected sample 1000 times (see Sections 3.2.2 and 3.3 for details).

(e.g. R. J. Bouwens et al. 2023). Comparing Schechter parameters is, in any case, not a straightforward way to compare predictions with observations. First, both the predicted and observed LFs may not be well described by a Schechter function. Secondly, observational Schechter-function fits may account for effects such as redshift evolution across the sample and photo- z uncertainties, whereas these effects are not included in our fits to the COLIBRE UVLFs. We therefore focus primarily on direct comparisons between the predicted and observed LFs.

For a further comparison with observations, we show, in Fig. 7, the redshift evolution of the UV luminosity density (ρ_{UV} ; both dust-attenuated and dust-free), together with recent observational results based on *JWST* observations (dust-attenuated; including C. T. Donnan et al. 2023, 2024; Y. Harikane et al. 2023; S. L. Finkelstein et al. 2024; C. J. Willott et al. 2024; M. Franco et al. 2025; D. J. McLeod et al. 2026). The observational results are selected using criteria similar to those described in Section 2.4: we only include measurements obtained over redshift intervals narrower than 2 and we ensure that all references use the same faint-end limit $M_{UV} < -17$ for calculating the UV luminosity density. To compute the UV luminosity density in COLIBRE, we calculate the contributions from galaxies of different luminosities separately in different simulation volumes (L100m5, L200m6, and L400m7). The boundaries that determine which volume is used for a given luminosity range are the same as those adopted in the construction of the combined UVLF (see Section 3.2.2). For each luminosity interval, we sum the UV luminosities of galaxies in the corresponding simulation, divide by the volume to obtain the luminosity-density contribution and then sum over all intervals brighter than the chosen threshold ($M_{UV} = -17$, following the practice in the above-mentioned references) to derive the global UV luminosity density.

As can be seen in Fig. 7, the UV luminosity density decreases towards higher redshifts, by a factor of ≈ 700 for the dust-attenuated case and by a factor of ≈ 1700 for the dust-free case from $z = 7$. Compared to observations, as expected, the dust-attenuated UV luminosity densities from COLIBRE are systematically lower at all redshifts, by 0.5–1 dex, consistent with the lower galaxy abundances in COLIBRE relative to observations (see Figs 4 and 5). The dust-free results, as expected, yield higher UV luminosity densities than the dust-attenuated ones. The dust-free COLIBRE UV luminosity density agrees reasonably well with the observational data up to $z = 10$, but the observational measurements plotted here have not been corrected for dust attenuation. At higher redshifts, most observationally inferred UV luminosity densities lie above the COLIBRE predictions. The main exception is the measurement from C. J. Willott et al. (2024) at $z \approx 12$, which is broadly consistent with the dust-free COLIBRE result. In addition, the measurements from M. Franco et al. (2025) at $z \approx 13$ (and also that from Y. Harikane et al. 2023 at $z \approx 16$) have very large uncertainties, such that the dust-free COLIBRE result, or its extrapolation in the case of the $z \approx 16$ point, remains within their 1σ ranges.

The difference between dust-free and dust-attenuated luminosity densities decreases toward higher redshifts, indicating weaker dust attenuation at these earlier times, consistent with Fig. 5. The derived UV luminosity densities and their uncertainties (estimated using 1000 bootstrap iterations) are listed in Table 4. We note, however, that the observational estimates of ρ_{UV} rely on fitting Schechter functions to the measured UVLFs and integrating over the fitted functions, which introduces additional uncertainty. For this reason, direct comparison of the UVLFs themselves provides a more direct discrimination between the simulations and the observations.

4 DISCUSSION

4.1 Caveats for COLIBRE model and SKIRT post-processing

As seen in Section 3, the dust-attenuated UVLFs in COLIBRE appear to be systematically lower than the observations, especially at the highest redshifts ($z = 12$ and 15), where even the dust-free UVLFs in COLIBRE remain lower than the observations (see Fig. 5). These differences may arise from physical processes that are not included in the current simulations as well as from numerical limitations. In the following, we discuss several caveats that may affect the predicted UVLFs.

First, COLIBRE does not include Population III (Pop III) stars. These (almost) metal-free stars are expected to form in the first galaxies and may contribute to the UV luminosity budget at very high redshift. Although Pop III stars are expected to form primarily in the earliest minihaloes (e.g. T. Abel, G. L. Bryan & M. L. Norman 2002; see V. Bromm & R. B. Larson 2004 for a review) and therefore directly affect only the faintest galaxies, their feedback can indirectly influence the evolution of more massive systems. In particular, Pop III stars can enrich the surrounding medium with metals and produce strong radiative and SN feedback, which affects the transition to Population II star formation and the early star formation histories of galaxies. As a result, their absence may influence the overall evolution of the galaxy population. For example, E. M. Ventura et al. (2024) made use of a semi-analytic model to study the impact of Pop III stars on UVLFs and reported that Pop III stars can have a significant effect on the total UVLF at $z = 12$ –16. In particular, they found that Pop III-dominated

systems can contribute significantly to the bright end of the UVLF at these redshifts, especially for a top-heavy Pop III IMF and may help explain the abundance of bright galaxies observed by *JWST* at $z \gtrsim 12$ without requiring an increased SFE.

Second, COLIBRE assumes a redshift-dependent, spatially uniform metagalactic ionizing and photodissociating background. This is likely to be a poor approximation during hydrogen reionization (which is completed at $z \approx 7$ in the UV-background model adopted by COLIBRE) when the radiation field is expected to be highly inhomogeneous. In particular, galaxies in overdense regions may experience a different ionizing background from those in more isolated environments, which could affect gas cooling, star formation histories and hence the predicted UV luminosities of high-redshift galaxies.

Third, COLIBRE assumes that the interstellar radiation field depends only on the local gas density (S. Ploekinger et al. 2025). This means that dense gas clouds that have not yet formed stars are nevertheless exposed to an interstellar radiation field, which may suppress star formation and delay the formation of the first stars. This approximation could therefore suppress the early build-up of the galaxy population and lead to a moderate underestimate of the predicted UVLFs at the highest redshifts considered here.

Fourth, the UV luminosities considered in this work do not include any contribution from AGN. While the UV emission of high-redshift galaxies is generally expected to be dominated by stellar light, AGN may contribute to the luminosities of the most massive systems (e.g. E. Giallongo et al. 2015; Y. Matsuoka et al. 2018; Y. Harikane et al. 2022), potentially affecting the brightest end of the UVLF. We note, however, that the highest redshift AGN currently known is at $z \approx 10.6$ (R. Maiolino et al. 2024a, b) and so AGN are unlikely to explain the difference between modelled and observed UVLFs at $z \gtrsim 12$.

In addition, nebular continuum emission is not included in the TODDLERS library used in SKIRT. Nebular continuum can modestly enhance the UV luminosities of very young stellar populations (e.g. D. Schaerer 2003; A. Raiter, R. A. E. Fosbury & H. Teimoorinia 2010) and may therefore slightly modify the UVLF, particularly at the bright end and at very high redshift where star formation episodes are dominated by young stellar populations. A test of the impact of nebular continuum on the UVLF can be found in A. Durrant et al. (2026), where nebular continuum is found to affect the UVLF at $z \gtrsim 9$, with the effect reaching up to ≈ 0.5 mag at the brightest end of the UVLF at $z = 15$. This is too small to reconcile the simulations with the observations in Fig. 5. The choices made in the COLIBRE-SKIRT pipeline may also introduce additional uncertainties in the predicted UVLFs. We refer the reader to A. Gebek et al. (2026) and Paper I of this series (S. Lu et al. 2026) for tests exploring the impact of different pipeline choices.

Other caveats include:

(i) *The dust model in COLIBRE*. Although the COLIBRE dust model has been shown to reproduce the $z = 0$ galaxy dust mass function and dust scaling relations well (see J. W. Trayford et al. 2026), it has not yet been tested over the redshift range studied in this work ($7 \leq z \leq 20$). Thus, the dust-attenuated results should be interpreted with caution. This is why we also present the dust-free results in Fig. 5. We discuss this point further in Section 4.3.1.

(ii) *Finite numerical resolution*. The finite mass and spatial resolution of the simulations may influence both the formation of low-mass galaxies and the internal structure of massive galaxies.

In particular, the star formation and feedback processes in galaxies may not be fully resolved, which could affect the predicted UV luminosities and therefore the UVLF. However, the convergence of COLIBRE simulations with resolution (see Fig. 4) suggests that this may not be a significant shortcoming.

(iii) *Limited simulation volume*. Since the simulations have a limited volume, the abundance of extremely rare and very luminous galaxies may not be fully sampled. This effect primarily impacts the brightest end of the UVLF, although the impact is expected to be modest given the relatively large volume of the simulations at m6 and m7 resolutions in which the brightest observed galaxies are well sampled.

4.2 Testing observational biases on the measured LFs

When comparing simulated LFs with observational data, several observational uncertainties and limitations should be kept in mind. First, observational LFs are typically calculated over finite-redshift bins, whereas we use simulation snapshots at a single redshift. If the LF evolves rapidly with redshift, this mismatch may lead to systematic differences, as the observed galaxy population is often biased toward the lower redshift end of the bin (M. J. Page & F. J. Carrera 2000).

Secondly, photometric-redshift errors can introduce systematic differences in the inferred high-redshift UVLFs. Since the bright end of the UVLF is steep and evolves rapidly with redshift, scatter in the estimated photo- z s can cause more galaxies to be scattered to higher rather than lower redshifts, thereby causing their inferred absolute magnitudes to appear brighter than their true values and boosting the inferred number density at the bright end (M. U. Subbarao et al. 1996; R. K. Sheth 2007). Such effects may be particularly relevant at high redshift, where photometric redshift uncertainties are significant (P. Arrabal Haro et al. 2023; R. Bouwens et al. 2023; S. Fujimoto et al. 2023; S. Serjeant & T. J. L. C. Bakx 2023) and most LFs rely entirely on photo- z estimates.

Thirdly, measurement uncertainties in luminosities can also introduce A. S. Eddington (1913) bias in which galaxies are scattered due to measurement errors across a steeply declining LF. In this case, the larger number of intrinsically faint galaxies scattered into brighter bins than vice versa by magnitude errors artificially enhances the inferred number density at the bright end.

Below, we investigate the impact of these observational effects on the UVLF by implementing them in the COLIBRE UVLFs. As these tests rely on the Schechter form of the COLIBRE UVLFs and aim to explain the difference between COLIBRE and observations, we only apply these tests to the COLIBRE UVLFs at $z \leq 15$.

4.2.1 The effect of redshift binning

In observations, UVLFs are typically estimated within finite-redshift bins, rather than at exact redshifts as in simulations. Consequently, the observed UVLFs at a given redshift represent an average over a redshift range. If the number of galaxies decreases rapidly with increasing redshift, this averaging tends to overestimate the galaxy abundance compared to the true value at the centre of the redshift bin (M. J. Page & F. J. Carrera 2000). The magnitude of this difference depends on the width of the redshift bin, to which we refer here as the ‘redshift binning effect’.

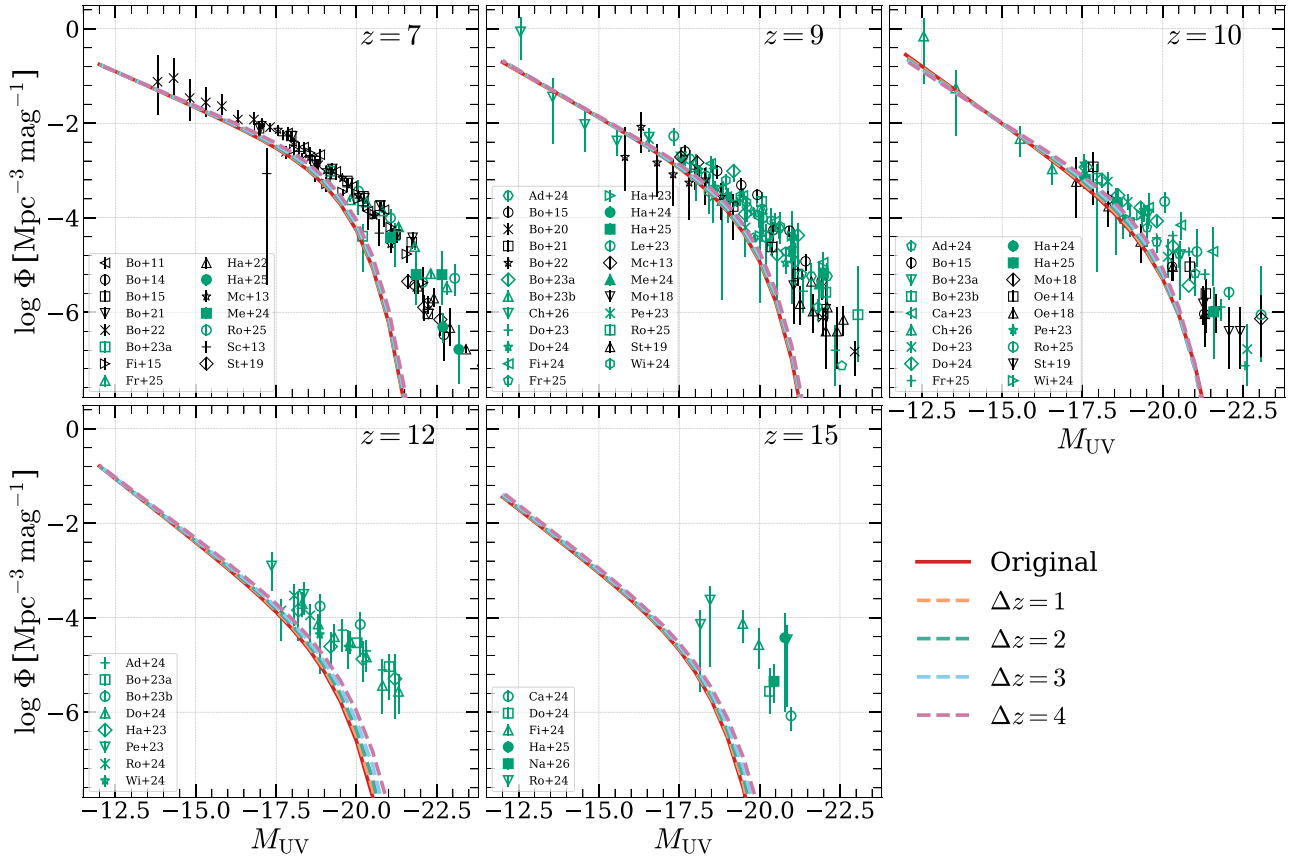


Figure 8. The effect of redshift binning on the predicted UVLFs. In each panel, the Schechter fit to the original (dust-attenuated) UVLF from COLIBRE (without the redshift binning effect) is shown with a red solid curve, while those including the binning effect are shown with dashed curves of different colour, corresponding to different values of Δz (i.e. the width of the redshift bin: $\Delta z = 1, 2, 3$, and 4 ; see Section 4.2.1 for details). The symbols for the observational data are the same as in Fig. 4.

To investigate this effect, we recalculate the UVLF at a given redshift, z , as the average value from $z - \Delta z/2$ to $z + \Delta z/2$, using:

$$\Phi(M, z) = \frac{\sum \Phi(M, z_i) \times \frac{dV}{dz}(z_i) \times \delta z}{\sum \frac{dV}{dz}(z_i) \times \delta z}, \quad (5)$$

where $\Phi(M, z_i)$ is the galaxy abundance at absolute magnitude M and redshift z_i ; δz is the thickness of the redshift slice (set to 0.1 here); and $\frac{dV}{dz}(z_i)$ is the comoving volume element per unit redshift at z_i , given by

$$\frac{dV}{dz}(z_i) = \frac{c}{H(z_i)} D_M^2(z_i) \Delta\Omega, \quad (6)$$

where c is the speed of light; $H(z_i)$ is the *Hubble* parameter at redshift z_i ; $D_M(z_i)$ is the transverse comoving distance at redshift z_i ; and $\Delta\Omega$ is the solid angle subtended by the survey area. The summation extends over the range $z - \Delta z/2$ to $z + \Delta z/2$. Since we compute the ‘real’ UVLFs of COLIBRE only at a subset of redshifts (i.e. $z = 7, 9, 10, 12$, and 15), we use the Schechter functions fitted to the UVLFs to interpolate or extrapolate them to other redshifts.

In Fig. 8, we show the comparison between the original COLIBRE UVLFs as represented by the best-fitting Schechter functions (dust-attenuated with a projected aperture size of 50 kpc) and those including the redshift binning effect (for different bin widths: $\Delta z = 1, 2, 3$, and 4 , indicated by dashed curves of different colour). As shown, larger Δz (i.e. wider redshift bins) only

results in slightly higher galaxy abundances and even with an extremely large redshift bin ($\Delta z = 4$), the COLIBRE UVLFs still fall short of the observations. As mentioned in Section 2.4, the largest observational redshift bin size allowed for comparison in this work is $\Delta z = 2$. Thus, accounting for the binning effect alone cannot explain away more than a very small part of the discrepancy between simulations and observations.

We note that some of the observational studies used for comparison have already tried to correct for the redshift binning effect. For example, C. T. Donnan et al. (2024) use the full redshift probability distribution for each galaxy, combined with a UV magnitude-based prior, to distribute each galaxy’s redshift probability across bins, thereby reducing the effect of redshift bin sizes in calculating the UVLF. Therefore, even after accounting for binning effects either in the simulations or in the corrected observational data, the discrepancy between the COLIBRE UVLFs and the observed abundances remains significant.

4.2.2 The effect of photo- z errors

For a galaxy with apparent UV magnitude, m_{UV} , its inferred absolute UV magnitude will be brighter than the true value if it is deemed to be at a higher redshift than its actual redshift. Because (1) low-redshift galaxies are intrinsically more numerous than high-redshift galaxies and (2) a fixed m_{UV} corresponds to a fainter M_{UV} at lower redshift (where the galaxy abundance is also higher), symmetric measurement uncertainties in the photo-

z scatter more low-redshift galaxies into higher redshift bins than the reverse. As a result, the number of galaxies in a given M_{UV} bin at high redshift is artificially increased, leading to an overestimation of the galaxy abundance relative to the true value.

Here, we investigate this effect using the following steps, where we again use the Schechter fit to the COLIBRE UVLF at each snapshot redshift:

(i) For a given redshift bin, $[z, z + \delta z]$, where δz is set to 0.1 here, we calculate the intrinsic number of galaxies within an absolute magnitude bin, $[M, M - \Delta M]$, as:

$$N_{\text{int}}(z) = \Phi(M, z) \times \frac{dV}{dz}(z) \times \delta z \times \Delta M. \quad (7)$$

(ii) Due to measurement uncertainties in photometric redshifts, the ‘observed’ number of galaxies in the given redshift and magnitude bin is lower than the intrinsic number (ignoring, at this stage, contributions from other redshift bins), due to photo- z errors scattering galaxies into other redshift bins. This can be written as:

$$N'_{\text{obs}} = N_{\text{int}} \times P_{\text{stay}}, \quad (8)$$

where P_{stay} is the probability that galaxies within $[z, z + \delta z]$ have observed photo- z values that remain in the same redshift bin.

(iii) Some galaxies in other redshift bins, $[z_i, z_i + \delta z]$, may be wrongly inferred to lie within $[z, z + \delta z]$ due to photo- z errors, with the number given by:

$$\Delta N_i = N_{\text{int}}(z_i) \times P_{\text{scatter},i}, \quad (9)$$

where $N_{\text{int}}(z_i)$ is the intrinsic number of galaxies in $[z_i, z_i + \delta z]$, calculated using equation (7). Note that here we refer to galaxies with the same apparent UV magnitude, rather than the same absolute UV magnitude, since apparent magnitudes are what we observe. $P_{\text{scatter},i}$ is the probability of galaxies in $[z_i, z_i + \delta z]$ being scattered into $[z, z + \delta z]$.

(iv) The final ‘observed’ number of galaxies in $[z, z + \delta z]$ within the absolute magnitude bin $[M, M - \Delta M]$ is:

$$N_{\text{obs}}^{\text{final}} = N'_{\text{obs}} + \sum_i \Delta N_i. \quad (10)$$

The summation extends over all redshift bins except the original bin, $[z, z + \delta z]$. The new ‘observed’ galaxy abundance can then be calculated accordingly.

(v) The intrinsic galaxy abundances used to compute N_{int} are inferred from our Schechter fits to the simulated UVLFs. We adopt the real (i.e. measured from COLIBRE galaxies) Schechter parameters of the UVLFs at $z = 7, 9, 10, 12$, and 15 , while those at other redshifts are obtained by interpolation (and extrapolation where necessary, for $z < 7$ and > 15).

(vi) The probabilities (i.e. the probability of galaxies staying in the same redshift bin, P_{stay} and the probabilities of galaxies being scattered in from other redshift bins, $P_{\text{scatter},i}$) are derived under the assumption that the measured photo- z distribution follows a Gaussian centred on the true redshift, with a fixed photo- z error, ($\sigma_z/(1+z)$), that does not vary with redshift.⁸ We ignore the effect of k -corrections in this simplified test.

Larger σ_z corresponds to a higher probability of galaxies being assigned an incorrect redshift, resulting in lower P_{stay} and higher $P_{\text{scatter},i}$ (for a given $[z_i, z_i + \delta z]$). In Fig. 9, we present the original dust-attenuated COLIBRE UVLFs (in Schechter form) and those including the photo- z bias (for different $\sigma_z/(1+z)$), indicated by dashed curves of different colour, together with observational results. As can be seen, larger $\sigma_z/(1+z)$ leads to higher abundances at all M_{UV} . At all redshifts, $z = 7-15$, $\sigma_z/(1+z) \approx 0.1 - 0.2$ is sufficient to explain the difference between the original COLIBRE predictions and the observations. However, from a compilation of observational studies of galaxies at $z > 6$ (R. Bouwens et al. 2023; R. J. Bouwens et al. 2023; C. T. Donnan et al. 2023; S. L. Finkelstein et al. 2023; Y. Harikane et al. 2023; D. J. McLeod et al. 2024), we find that the median reported value of $\sigma_z/(1+z)$ is only 0.05, with an upper limit given by the 84th percentile of about 0.1. This is smaller than the value required to account for the discrepancy found here. We note, however, that the quoted photo- z errors here are in fact uncertainties derived from the photometric-redshift estimation method, rather than true errors relative to the actual redshifts and may therefore be underestimated. In addition, catastrophic photo- z errors, which may scatter low-redshift interlopers into high-redshift samples or move galaxies between widely separated redshift bins (e.g. M. Bradač et al. 2026), are not taken into account.

B. Robertson et al. (2024, labelled Ro+24 in our figures and appearing only in the $z = 12$ and 15 panels) state that they correct for the effect of photo- z uncertainties. Their estimates are still higher than our original UVLFs (red solid curves, without accounting for photo- z errors). Furthermore, estimates based on spectroscopically confirmed galaxies (Y. Harikane et al. 2024 at $z = 9$ and 10 ; Y. Harikane et al. 2025 at $z = 7, 9, 10$, and 15 ; R. P. Naidu et al. 2026 at $z = 15$), which are not affected by photo- z uncertainties, are also higher than our original UVLFs. Together, these results indicate that photo- z errors alone cannot explain the discrepancy between the COLIBRE predictions and the observations.

We note that the test presented here is intended only as a simplified experiment. In particular, extrapolating the Schechter fits obtained at $z = 7-15$ to substantially lower redshifts is inherently uncertain and should not be interpreted as a physically motivated prediction of the UVLF evolution. The shape and normalization of the UVLFs are expected to evolve non-linearly with cosmic time and this evolution may flatten towards lower redshift. As such, our extrapolation likely overestimates the number density of low-redshift galaxies that could be scattered into high-redshift bins due to large photometric redshift uncertainties.

In addition, in this test we assume that the probability distribution of photometric redshift measurements is Gaussian and centred on the true redshift. We also ignore catastrophic photo- z errors. In practice, however, photometric redshift probability distributions are often non-Gaussian and can be significantly asymmetric, with the upper uncertainty typically larger than the lower one (e.g. R. Bouwens et al. 2023; R. J. Bouwens et al. 2023; C. T. Donnan et al. 2023; S. L. Finkelstein et al. 2023; Y. Harikane et al. 2023; D. J. McLeod et al. 2024) and real photo- z errors may be substantially larger than assumed here.

4.2.3 The effects of Eddington bias

Due to the decreasing abundance of galaxies from the faint to the bright end of the LF, errors in measurements of magnitudes

⁸This assumption is broadly consistent with the observations reported by R. Bouwens et al. (2023); R. J. Bouwens et al. (2023); C. T. Donnan et al. (2023); S. L. Finkelstein et al. (2023); Y. Harikane et al. (2023); and D. J. McLeod et al. (2024).

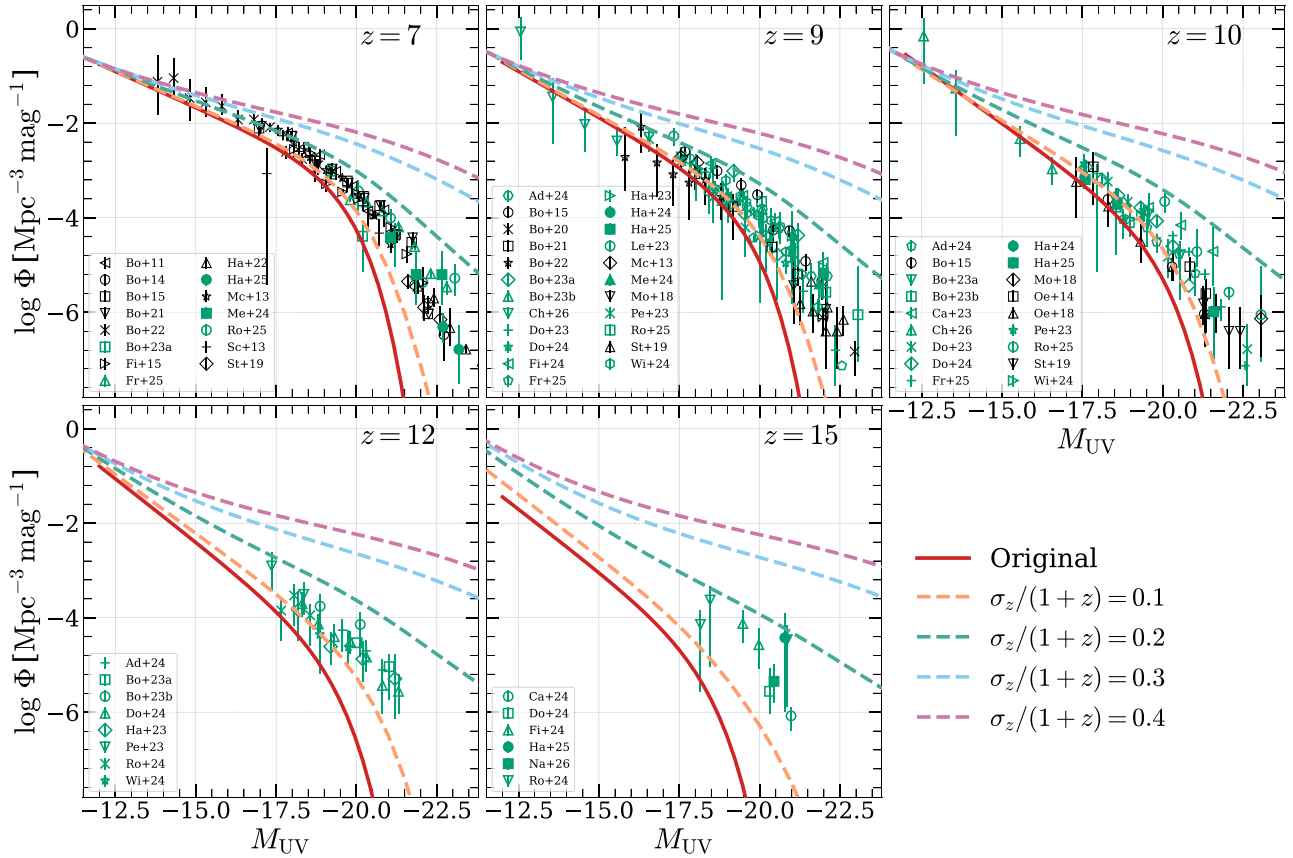


Figure 9. The effect of photo- z errors on the UVLFs. In each panel, the Schechter fit to the original (dust-attenuated) UVLF from COLIBRE (without the effect of photo- z errors) is shown with a red solid curve, while those including the photo- z error effect are shown with dashed curves of varying colour, corresponding to different values of σ_z , the measurement uncertainty of photo- z : $\sigma_z/(1+z) = 0.1, 0.2, 0.3$, and 0.4 (see Section 4.2.2 for details). We note that the larger values, $\sigma_z/(1+z) = 0.2, 0.3$, and 0.4 , are unlikely to represent realistic root-mean-square uncertainties and are included only to illustrate the potential effects of extreme photo- z errors. Our treatment assumes symmetric photo- z errors and does not include catastrophic failures. The symbols for the observational data are the same as in Fig. 4.

can cause the ‘observed’ number of galaxies in bright bins to be overestimated: more intrinsically faint galaxies are scattered into brighter bins than intrinsically bright galaxies into fainter bins. This effect is known as the Eddington bias (A. S. Eddington 1913).

To investigate the impact of Eddington bias, we assume that the measured M_{UV} follows a Gaussian distribution centred on the true value with a dispersion of $\sigma_{M_{UV}}$. The COLIBRE UVLFs (in the Schechter form) are then convolved with a Gaussian kernel of width $\sigma_{M_{UV}}$ to account for the observational uncertainty in M_{UV} .

In Fig. 10, we compare the original dust-attenuated COLIBRE UVLFs (in the Schechter form, without Eddington bias) to those including Eddington bias (with $\sigma_{M_{UV}} = 0.25, 0.5, 0.75$, and 1), alongside observational data. As can be seen, accounting for Eddington bias indeed brings the COLIBRE predictions closer to observations, but even $\sigma_{M_{UV}} = 1$ cannot fully explain the discrepancy at $z = 12$ and 15 . The estimated M_{UV} measurement uncertainties in observations are much smaller than this (with a median value of 0.2 mag and an 84th percentile value of 0.3 mag according to R. Bouwens et al. 2023; R. J. Bouwens et al. 2023; S. L. Finkelstein et al. 2023; Y. Harikane et al. 2023) and thus cannot reconcile the COLIBRE predictions with the data.

4.2.4 Combined observational effects

As shown in the previous sections, all three observational effects bring the COLIBRE UVLF predictions closer to the observations. We now test the combined effect of these three observational biases. Since we consider only observational studies with a redshift range no larger than 2 (see Section 2.4), we adopt a redshift bin size of 2. For the photo- z and M_{UV} errors, we consider two sets of values, one corresponding to the median values reported in observational studies ($\sigma_z/(1+z) = 0.05$ and $\sigma_{M_{UV}} = 0.2$; R. Bouwens et al. 2023; R. J. Bouwens et al. 2023; C. T. Donnan et al. 2023; S. L. Finkelstein et al. 2023; Y. Harikane et al. 2023; D. J. McLeod et al. 2024) and another corresponding to their 84th percentiles ($\sigma_z/(1+z) = 0.1$ and $\sigma_{M_{UV}} = 0.3$).

In Fig. 11, we show (1) the original dust-attenuated COLIBRE UVLFs (in Schechter form), (2) the dust-attenuated COLIBRE UVLFs including a *single* observational effect (either the median values of the photo- z or M_{UV} errors, or a redshift bin size of 2), (3) the dust-attenuated COLIBRE UVLFs including the combination of all observational effects with the median photo- z and M_{UV} errors, (4) the dust-attenuated COLIBRE UVLFs including the combination of all observational effects with the 84th percentiles of the photo- z and M_{UV} errors, and (5) the dust-free COLIBRE UVLFs including the combination of all observational effects

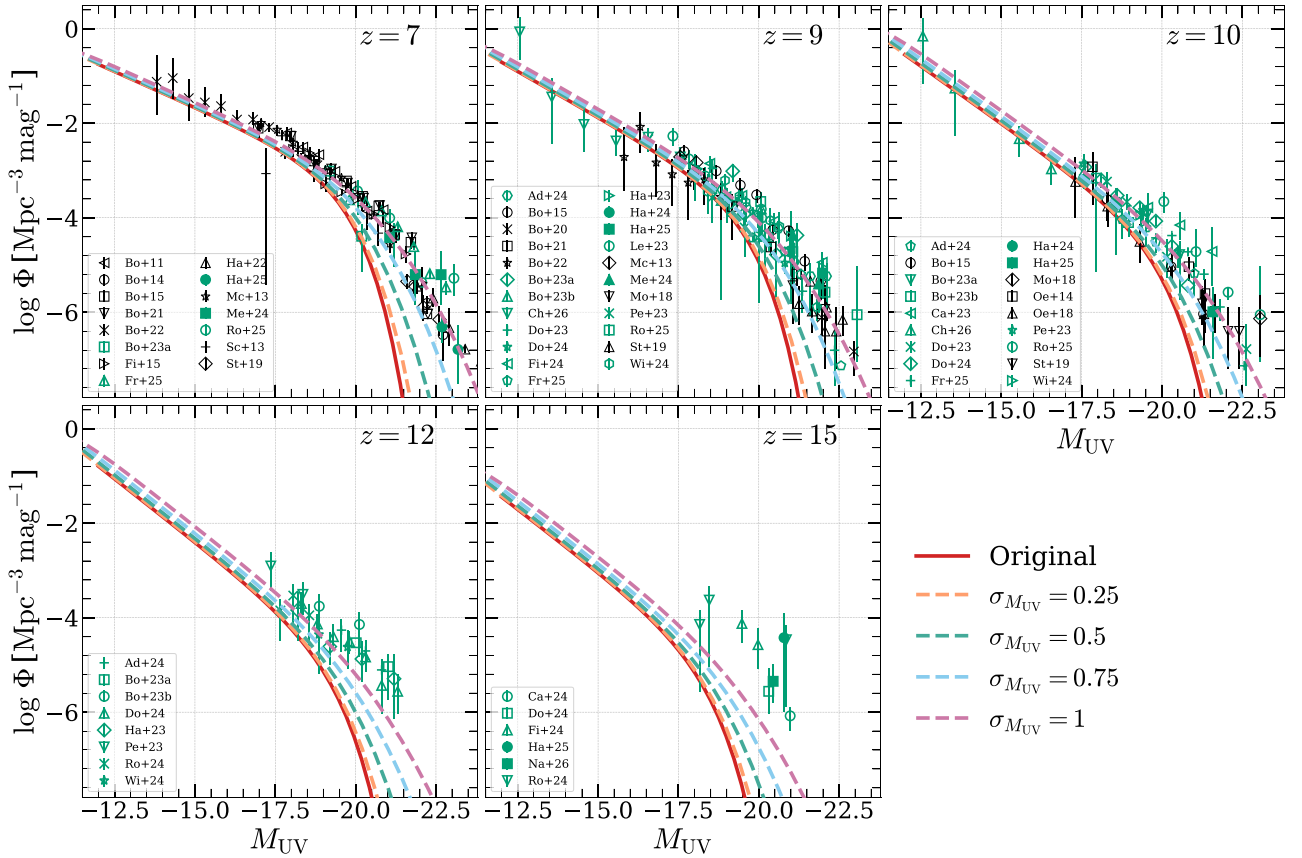


Figure 10. Effect of Eddington bias on the UVLFs (dust-attenuated). In each panel, the red solid curve shows the Schechter fit to the original dust-attenuated UVLF from COLIBRE without Eddington bias (i.e. without measurement uncertainty in M_{UV}), while the dashed curves show the results with Eddington bias (convolved with measurement uncertainties in M_{UV} of $\sigma_{M_{UV}} = 0.25, 0.5, 0.75$, and 1). The observational data points are the same as in Fig. 4.

with the 84th percentiles of the photo- z and M_{UV} errors (only for $z = 12$ and 15 panels). In cases (3)–(5), a redshift bin size of 2 is adopted.

As can be seen in the figure, combining all three uncertainties, assuming the median values of the photo- z and M_{UV} errors, shifts the dust-attenuated UVLF horizontally by $\lesssim 0.5$ mag at $\Phi = 10^{-6} \text{ Mpc}^{-3} \text{ mag}^{-1}$. However, the UV luminosities of galaxies at this number density remain fainter than observed. The combination assuming the 84th-percentile values of the photo- z and M_{UV} errors for the dust-attenuated COLIBRE UVLFs leads to rough agreement with observations at $z = 9$ and 10 , but still underestimates the galaxy luminosities at a given number density at $z = 12$ and 15 . Only when the effects of redshift binning and the 84th-percentile photo- z and M_{UV} errors are combined with the dust-free assumption at $z = 12$ and 15 (see the grey dotted curves in Fig. 11) do the COLIBRE UVLFs come close to the observations at $z = 12$ although they are still well below the data at $z = 15$. Considering that some studies, such as B. Robertson et al. (2024) and R. P. Naidu et al. (2026), may not be strongly affected by photo- z errors (which make the largest contribution to the shift of the UVLFs), as mentioned earlier in Section 4.2.2, additional physical mechanisms may still be required to boost the UV luminosities of high-redshift galaxies in order to explain the discrepancy between the COLIBRE UVLFs and the observations.

Throughout this section, we do not explicitly model the effects of cosmic variance. The COLIBRE UVLFs used at the bright end are measured from large simulation volumes, mainly L200m6 and L400m7, with volumes of 8×10^6 and $6.4 \times 10^7 \text{ cMpc}^3$, respectively. These volumes are several hundred to several thousand times larger than the effective volumes of current high-redshift *JWST* surveys, which are typically of order 10^3 – 10^5 cMpc^3 (e.g. M. Castellano et al. 2023; S. L. Finkelstein et al. 2024; I. Chemerynska et al. 2026). In addition, the COLIBRE simulations are periodic cosmological volumes whose mean density is fixed by the adopted cosmology, rather than randomly selected over or underdense survey fields (see sect. 2.2 of J. Schaye et al. 2026). We therefore regard the COLIBRE UVLFs as being close to the cosmic mean predictions.

On the observational side, the constraints compiled in Table 3 are drawn from multiple studies and from several independent or partially independent survey fields. The scatter among these measurements should therefore already reflect, at least partly, the effect of cosmic variance, together with differences in sample selection, photometric-redshift modelling, and completeness corrections. In addition, some published studies (e.g. C. T. Donnan et al. 2024; I. Chemerynska et al. 2026) attempt to account for the effects of cosmic variance in their error bars. We therefore do not further account for the effect of cosmic variance in our comparison.

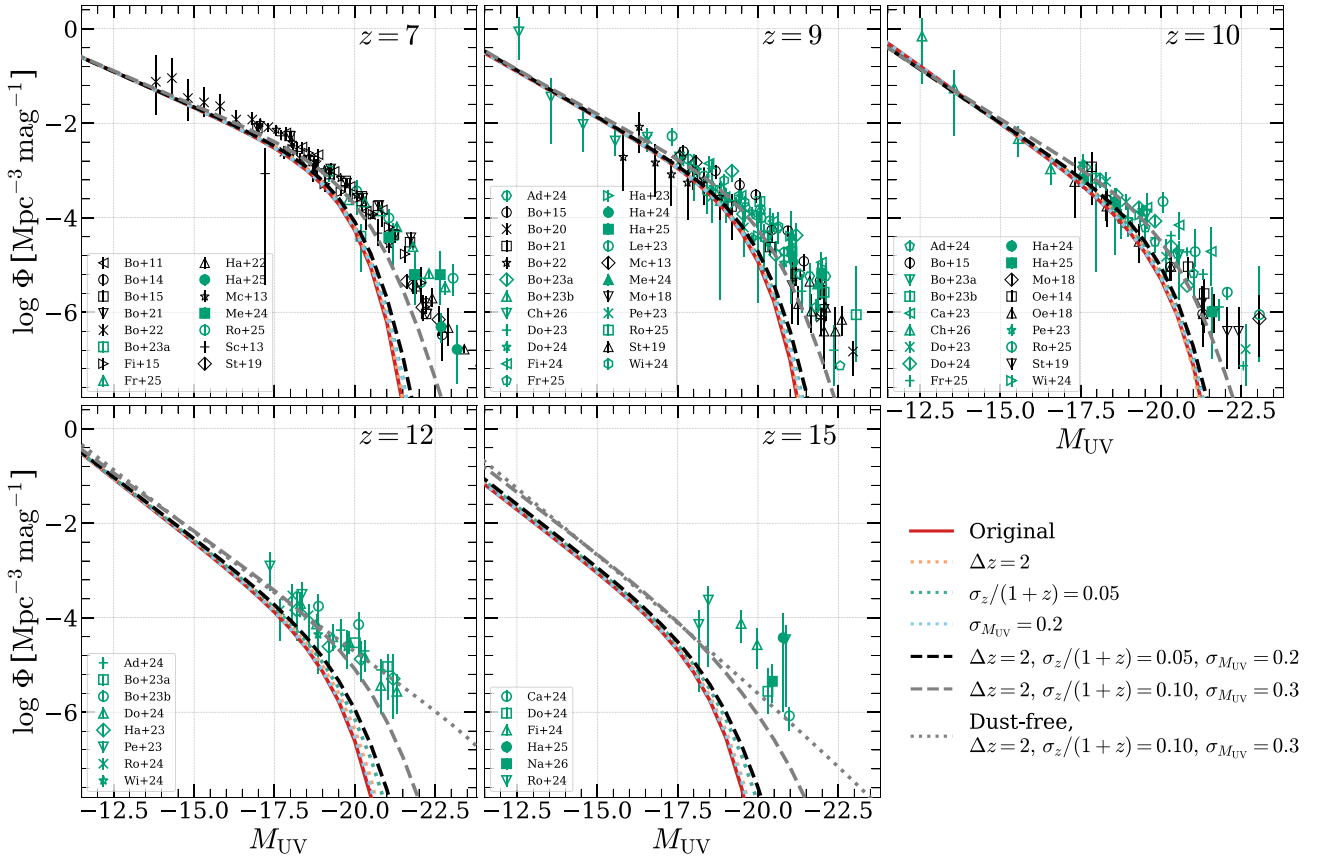


Figure 11. The combined effect of observational uncertainties on the dust-attenuated UVLFs (redshift binning, photometric redshift errors, and Eddington bias). In each panel, the original dust-attenuated UVLF from COLIBRE is shown by the red solid curve. The dotted curves in different colours show the UVLFs obtained when including a single observational effect (assuming median observational values or redshift bin size of 2). The black dashed curve shows the result including all three observational effects, assuming a redshift bin size of 2, a photometric redshift error of $\sigma_z/(1+z) = 0.05$ (the median value in observations), and an uncertainty in M_{UV} of 0.2 (also the median observational value). The grey dashed curve shows the result when adopting the 84th percentiles of the photometric redshift and M_{UV} uncertainties ($\sigma_z/(1+z) = 0.10$ and $\sigma_{M_{UV}} = 0.3$). The grey dotted curves in the $z = 12$ and 15 panels show the dust-free COLIBRE UVLFs combined with a redshift bin size of 2, a photo- z uncertainty of $\sigma_z/(1+z) = 0.10$, and an M_{UV} uncertainty of $\sigma_{M_{UV}} = 0.3$. The symbols for the observational data are the same as in Fig. 4.

4.3 Possible physical explanations

We have seen that observational uncertainties are not large enough to resolve the discrepancy between the COLIBRE UVLFs and observations. In this section, we discuss possible physical reasons that could explain the discrepancy.

4.3.1 Impact of dust content

Following the first wave of *JWST* papers claiming that the data were inconsistent with theoretical models within a Λ CDM framework (e.g. S. L. Finkelstein et al. 2023; I. Labbé et al. 2023), several studies have tried to explain the difference between predicted and observed UVLFs without challenging the entire Λ CDM paradigm. An early explanation was reduced dust attenuation at high redshift. For example, A. Ferrara, A. Pallottini & P. Dayal (2023), using an empirical model, argued that the measured UVLFs could be reproduced up to $z \approx 14$ if galaxies at $z \gtrsim 11$ contained negligible amounts of dust. More recently, F. Prada et al. (2026) also reported, using the empirical model UNIVERSEMACHINE, that the predicted dust-free UVLFs show good agreement with the observational data at $z = 12$. Moreover, A. P. Vijayan

et al. (2026) found that observational upper limits on dust masses at high redshift are only marginally consistent with the median dust mass–stellar mass relation predicted by COLIBRE, suggesting that dust grain growth in COLIBRE at $z \gtrsim 10$ may be more efficient than in the real Universe. Nevertheless, even in the limiting case of completely neglecting dust attenuation, our results in Fig. 5 show that the observed abundance of UV-bright galaxies at $z \gtrsim 10$ cannot be reproduced. Consistent with this finding, F. Iocco & L. Visinelli (2024) argued that, although dust attenuation may be minimal at the highest redshifts, an enhanced SFE is still required to reproduce the observed UVLFs and generate sufficiently luminous galaxies.

S. Lu et al. (2025) noted that the small age of the Universe at $z = 14$ – only 300 Myr – limits the amount of dust that can form from metals by grain growth in the ISM. This assumption is supported by recent observational constraints indicating very low dust content in galaxies at $z \gtrsim 12$ (I. Mitsuhashi et al. 2026). Using the Durham semi-analytic model, GALFORM, they showed that a reduced dust-to-metal ratio (relative to the Milky Way value) at $z = 12$ and 14 (due to the non-zero dust growth time-scale) brings the model predictions closer to observations. Additional factors beyond dust attenuation must be considered. S. Lu et al.

(2025) proposed that the missing factor is a top-heavy IMF in starbursts and they showed explicitly that the pre-*JWST*-launch W. I. Cowley et al. (2018) GALFORM predictions – which assume such an IMF – together with a realistic model for dust growth at very high redshifts, provide a good match to the *JWST* data over the entire redshift range (see below for further details).

4.3.2 Impact of the initial mass function

E. Chaikin et al. (2026b) compared the GSMFs in the COLIBRE simulations as a function of redshift with those inferred from observations. They found better agreement with the data than the level of agreement we find here between the predicted UVLFs and observations. This difference may reflect the fact that stellar masses are dominated by long-lived low-mass stars, whereas the UV luminosity is primarily produced by short-lived massive stars. A possible explanation of the difference is therefore that the simulations (with a universal G. Chabrier 2003 IMF) underestimate the relative abundance of massive stars that dominate the UV emission. Variations in the IMF that increase the fraction of massive stars would boost the UV luminosity without strongly altering the stellar mass function, making the top-heavy IMF advocated by W. I. Cowley et al. (2018) and others a plausible way to reconcile the predicted UVLFs with observations.

Observational studies have provided substantial evidence for the existence of a top-heavy IMF. For example, using C, N, and O isotopic abundance ratios in starburst galaxies, D. Romano et al. (2017) reported compelling evidence for a top-heavy IMF, suggesting that intensely star-forming environments may preferentially produce a larger fraction of massive stars. A. Sneppen et al. (2022) fitted a temperature-dependent IMF to galaxies in the COSMOS2015 catalogue (C. Laigle et al. 2016) and found that the IMF becomes increasingly top-heavy at higher redshifts. A. J. Cameron et al. (2024) reported two Lyman- α emitting galaxies at redshifts 5.9 and 7.9 that show signatures of exceptionally top-heavy IMFs. F.-H. Liang et al. (2021) argued that the IMF tends to be more top-heavy at lower metallicities (see also Z. Yan, T. Jerabkova & P. Kroupa 2017), further supporting the likelihood of a top-heavy IMF at high redshift.

A dual IMF – a solar neighbourhood-like form for quiescent/disc star formation and a top-heavy power law with slope $x = 1$ (compared with $x = 1.35$ for the high-mass end of a G. Chabrier 2003 IMF) in starbursts – has been adopted in the Durham semi-analytic model GALFORM (S. Cole et al. 2000; C. G. Lacey et al. 2016) since 2005 to match the observed number counts and redshift distribution of submillimetre galaxies (C. M. Baugh et al. 2005, who adopted an even more top-heavy IMF with $x = 0$). S. Lu et al. (2025) compared the GALFORM predictions of W. I. Cowley et al. (2018) for high- z galaxies, published before the launch of *JWST*, with the data and found that a combination of a top-heavy IMF and reduced dust content at high redshift yields good agreement with the observed UVLFs out to $z = 14$. However, the effect of a top-heavy IMF on stellar feedback, including the increased number of SNe expected for a top-heavy IMF, was not explicitly modelled and, although the effect of the IMF on metal enrichment was included, the model neglects other possible effects of the IMF on grain growth from metals. More recently, F. Fontanot et al. (2026) showed, using the semi-analytic model GAEA (G. De Lucia et al. 2024), that a variable IMF can improve agreement with observed UVLFs up to $z \approx 13$. However, although their increasingly top-heavy IMF affects the

stellar populations, returned fraction, SNe fraction, and chemical enrichment, it is not explicitly coupled to the stellar-feedback prescription. They also did not recalibrate the GAEA parameters after introducing the variable IMF.

E. R. Cueto et al. (2024), using the ASTRAEUS framework that couples a semi-analytic galaxy evolution model with a semi-numerical reionization scheme, argued that a top-heavy IMF alone is insufficient to account for the abundance of bright galaxies at high redshift. While such an IMF does enhance UV photon production per unit stellar mass, it also increases the strength of SN feedback, which suppresses further star formation and limits the net gain in UV luminosity. They argued that, as a result, the UVLFs remain largely consistent with those predicted under a Salpeter IMF (Salpeter 1955). In contrast, A. Hutter et al. (2025), using the same ASTRAEUS framework but adopting an IMF that becomes increasingly top-heavy in gas-rich, highly star-forming galaxies, found that such an evolving IMF can reproduce the observed UVLFs over $z = 5–15$. They showed that the luminosity boost is largest in massive galaxies, where it is less efficiently offset by SN feedback, thereby enhancing the abundance of UV-bright systems at $z > 10$.

Thus, assessing the effect of a top-heavy IMF on the high-redshift UVLF requires a self-consistent model that includes its impact on SN feedback, metal production, dust evolution, and stellar luminosities. This has recently been investigated by A. Durrant et al. (2026), who used COLIBRE to explore the effects of a variable IMF on the high-redshift UVLFs within a self-consistent forward-modelling framework. They find that a variable IMF that becomes increasingly top-heavy in high gas density regions goes a long way towards explaining the high-redshift UVLFs.

4.3.3 Other possibilities

R. S. Somerville et al. (2025) incorporated a density-modulated SFE into a semi-analytic galaxy formation model and showed that a higher effective SFE can reproduce or even exceed the observed number densities of UV-bright galaxies at $z \approx 6–17$. They further argued that this effect, together with decreasing dust attenuation and enhanced stochasticity in UV luminosities at earlier cosmic times, could help reproduce the observed evolution. However, the reduction in dust attenuation was implemented through a simplified post-processing prescription and the increased burstiness was likewise introduced in post-processing rather than being fully self-consistently predicted by the model. In addition, this updated model was not explicitly recalibrated to reproduce galaxy constraints at $z = 0$ so its validity is difficult to assess.

Another possibility is that bursty star formation, or more generally, increased scatter in UV luminosity at fixed halo mass, could boost the bright end of the high-redshift UVLF. C. A. Mason, M. Trenti & T. Treu (2023), using a simple empirical model based on halo formation time-scales derived from extended Press–Schechter theory, explored how variations in halo assembly histories, together with variations in SFE and dust attenuation, could affect the bright end of the high-redshift UVLF. They showed that scatter in halo formation time-scales can produce substantial scatter in the SFR at fixed halo mass, such that the youngest and most rapidly star-forming systems are scattered to brighter UV luminosities, thereby boosting the bright end of the UVLF.

G. Sun et al. (2023) used the high-redshift version of the FIRE-2 cosmological zoom-in simulations (X. Ma et al. 2018a, b) to study the bright end of the UVLFs at $8 \leq z \leq 12$. They showed

that, with a multichannel implementation of standard stellar feedback processes (including, e.g. SNe, stellar winds, and radiation from young stars), bursty star formation histories arise self-consistently in high-redshift galaxies and can reproduce the observed UVLFs over this redshift range without invoking a top-heavy IMF, non-standard cosmology, or strongly enhanced SFE. Similarly, using a minimal empirical framework, X. Shen et al. (2023) argued that variability in galaxy UV luminosities at fixed halo mass, potentially driven by bursty star formation and dust clearance during strong feedback cycles, can significantly increase the abundance of UV-bright galaxies and thereby help reconcile model predictions with *JWST* observations at high redshift. However, their analysis was based on a highly simplified empirical model, in which the UV-luminosity variability was introduced phenomenologically rather than being predicted self-consistently from an explicit model of galaxy formation.

While *ad hoc* models such as those discussed in this subsection, which introduce physical mechanisms explicitly designed to match the high-redshift observations, are useful guides to the sort of mechanisms that could be important, they cannot be regarded as fully successful models of galaxy formation, since there is no guarantee that they would produce a realistic galaxy population at later times including at the present day.

5 CONCLUSIONS

This paper is the second in a series of three in which we aim to study comprehensively the multiwavelength galaxy luminosity functions and their redshift evolution predicted by the new state-of-the-art hydrodynamical cosmological simulation set, COLIBRE (J. Schaye et al. 2026; E. Chaikin et al. 2026a), which resolves the cooling of the ISM down to 10 K, combined with the SKIRT radiative transfer code (M. Baes et al. 2011; P. Camps & M. Baes 2015, 2020). Specifically, in this paper, we study the UVLFs at high redshifts, $z = 7$ to 20 and compare them with observations, especially those from *JWST*. Our key conclusions are:

(i) COLIBRE simulations that span two orders of magnitude in numerical resolution show excellent convergence in the relations between stellar mass and other galaxy properties, such as total mass, mass of gas in the cool and dense phase, dust mass, SFRs, etc., down to very low stellar masses (corresponding to 2–5 stellar particles; Fig. 1).

(ii) We carried out SKIRT radiative transfer calculations of both dust-free and dust-attenuated UV luminosities for a representative subsample of galaxies (see Fig. A1 for a test of the galaxy subsampling), in order to derive the UVLFs in three COLIBRE simulations of different volumes and numerical resolutions (L100m5, L200m6, and L400m7). The three COLIBRE simulations, when processed using SKIRT, also show excellent convergence in the $M_{UV} - M_*$ relation, both with and without dust attenuation (Fig. 3). The UVLFs derived from the different simulations show good convergence over their overlapping UV-luminosity ranges (Fig. 4). To obtain UVLFs over the widest accessible luminosity range, we combine the results from all three simulations (see Section 3.2.2).

(iii) The COLIBRE UVLFs show stronger redshift evolution than the observational data (e.g. Y. Harikane et al. 2023; S. L. Finkelstein et al. 2024), with the abundance at fixed M_{UV} decreasing more rapidly towards higher redshift. The individual, as well as the combined, dust-attenuated UVLFs in the three COLIBRE simulations lie systematically below the observational

determinations, particularly at the bright end and this discrepancy increases rapidly towards higher redshift (Fig. 4). At $z = 7$ and 9, the dust-free UVLFs in COLIBRE have a higher amplitude than the observed UVLFs at the bright end; at $z = 10$, they agree well with observations. At $z = 12$, although the dust-free UVLF in COLIBRE still lies slightly below the observational data, the UV magnitude of galaxies at $\Phi = 10^{-6} \text{ mag}^{-1} \text{ Mpc}^{-3}$ roughly matches the observed value. However, at $z = 15$, even the dust-free COLIBRE galaxies are ≈ 2.5 mag fainter than observed at this same number density. At $z = 20$, where only L100m5 contains a sufficient number of resolved galaxies, the dust-attenuated and dust-free UVLFs are nearly identical. In the absence of observational measurements satisfying our selection criteria, we present this result as a prediction for future observations (Fig. 5).

(iv) We fit Schechter functions to the combined, dust-attenuated UVLFs from COLIBRE and find that the Schechter parameters exhibit strong evolution with redshift: the characteristic magnitude, M_{sch} , becomes fainter; the faint-end slope, α_{sch} , becomes more negative (i.e. steeper); and the characteristic number density, $\log \Phi_{sch}$, decreases towards higher redshift (Fig. 6).

(v) The derived dust-attenuated and dust-free UV luminosity densities both decrease towards higher redshifts, by factors of ≈ 700 and ≈ 1700 , respectively. At $z > 10$, even the dust-free result is systematically lower than most observational estimates by ≈ 0.5 –1 dex, with larger discrepancies at higher redshifts, consistent with the lower galaxy abundance across all redshifts (Fig. 7).

(vi) Accounting for observational uncertainties in the measured UVLFs, including those resulting from finite-redshift bin sizes, measurement errors in photometric redshifts (assumed to be symmetric and ignoring catastrophic errors), and errors in M_{UV} (i.e. Eddington bias), brings the COLIBRE UVLFs closer to the observations. However, even the combination of all three effects does not fully resolve the discrepancy between the modelled and observed UVLFs when observationally reasonable values are assumed, indicating that additional physical mechanisms are required to boost the UV luminosities of high-redshift galaxies in the simulations (Figs 8–11).

There may be multiple mechanisms capable of boosting the UV luminosities of high-redshift galaxies, as discussed in Section 4.3. However, their overall impact on galaxy formation cannot be assessed in isolation, nor can it be robustly inferred from *ad hoc* modifications introduced solely to enhance the UV output, because such changes are also expected to affect other aspects of galaxy evolution. For example, a higher SFE may lead not only to higher UV luminosities, but also to stronger stellar feedback, more rapid gas consumption, enhanced metal enrichment, and potentially different dust production and attenuation histories. These coupled effects may, in turn, regulate subsequent star formation and partly offset the initial increase in UV emission (see also the discussion of the top-heavy IMF in Section 4.3.2). It is therefore necessary to test the effects of the proposed mechanisms within a self-consistent framework that simultaneously follows their impact on star formation, feedback, chemical enrichment, dust evolution, and the emergent luminosities of galaxies at all times.

ACKNOWLEDGEMENTS

We thank the referee for their constructive comments, which improved the paper. We thank Prof. Cheng Li for helpful comments on Eddington bias and Dr Fengwu Sun for comments

on *JWST* aperture sizes. We acknowledge support from STFC (ST/X001075/1). CSF acknowledges support from the European Research Council through Advanced Investigator grant DMIDAS (GA 786910). SB is supported by a UK Research and Innovation Future Leaders Fellowship (grant nos MR/V023381/1 and UKRI2044). AD is supported by an STFC doctoral studentship. NA acknowledges financial support by the Flemish Fund for Scientific Research (FWO – Vlaanderen) through the research grant G0C4723N. ABL acknowledges support by the Italian Ministry for Universities (MUR) program ‘Dipartimenti di Eccellenza 2023-2027’ within the Centro Bicocca di Cosmologia Quantitativa (BiCoQ) and support by UNIMIB’s Fondo Di Ateneo Quota Competitiva (project 2024-ATEQC-0050). EC acknowledges support from the Netherlands Organization for Scientific Research (NWO) through research programme Athena 184.034.002 and STFC consolidated grant ST/X001075/1. FH acknowledges funding from the Netherlands Organization for Scientific Research (NWO) through research programme Athena 184.034.002. SP acknowledges support by the Austrian Science Fund (FWF) through grant – DOI: 10.55776/V982. JT acknowledges support from an STFC Early Stage Research and Development grant (ST/X004651/1). 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 UK National e-Infrastructure.

DATA AVAILABILITY

The predicted UVLFs from COLIBRE, together with the compiled observational UVLFs used for comparison in this work, are included as Supplementary Material and can be obtained from both the journal website and <https://icc.dur.ac.uk/data/>. More data sets in a machine-readable format are available from the DigitizeAstronomy website (<https://digastro.org>).

REFERENCES

- Abbott T. M. C. et al., 2022, *Phys. Rev. D*, 105, 023520
 Abel T., Bryan G. L., Norman M. L., 2002, *Science*, 295, 93
 Adams N. J. et al., 2024, *ApJ*, 965, 169
 Arrabal Haro P. et al., 2023, *ApJ*, 951, L22
 Atek H. et al., 2023, *MNRAS*, 519, 1201
 Baes M., Camps P., 2015, *Astron. Comput.*, 12, 33
 Baes M., Verstappen J., De Looze I., Fritz J., Saftly W., Vidal Pérez E., Stalevski M., Valcke S., 2011, *ApJS*, 196, 22
 Barrufet L. et al., 2023, *MNRAS*, 522, 3926
 Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, *MNRAS*, 356, 1191
 Benítez-Llambay A. et al., 2026, *MNRAS*, 546, 268
 Borrow J., Borrisov A., 2020, *J. Open Source Softw.*, 5, 2430
 Borrow J., Kelly A. J., 2021, preprint ([arXiv:2106.05281](https://arxiv.org/abs/2106.05281))
 Borrow J., Schaller M., Bower R. G., Schaye J., 2022, *MNRAS*, 511, 2367
 Bouwens R., Illingworth G., Oesch P., Stefanon M., Naidu R., van Leeuwen I., Magee D., 2023, *MNRAS*, 523, 1009
 Bouwens R. J. et al., 2011, *ApJ*, 737, 90
 Bouwens R. J. et al., 2015, *ApJ*, 803, 34
 Bouwens R. J. et al., 2021, *AJ*, 162, 47
 Bouwens R. J., Illingworth G., Ellis R. S., Oesch P., Stefanon M., 2022, *ApJ*, 940, 55
 Bouwens R. J. et al., 2023, *MNRAS*, 523, 1036
 Bowler R. A. A. et al., 2014, *MNRAS*, 440, 2810
 Bowler R. A. A., Jarvis M. J., Dunlop J. S., McLure R. J., McLeod D. J., Adams N. J., Milvang-Jensen B., McCracken H. J., 2020, *MNRAS*, 493, 2059
 Bradač M. et al., 2026, preprint ([arXiv:2604.23668](https://arxiv.org/abs/2604.23668))
 Bromm V., Larson R. B., 2004, *ARA&A*, 42, 79
 Cameron A. J., Katz H., Witten C., Saxena A., Laporte N., Bunker A. J., 2024, *MNRAS*, 534, 523
 Camps P., Baes M., 2015, *Astron. Comput.*, 9, 20
 Camps P., Baes M., 2020, *Astron. Comput.*, 31, 100381
 Camps P., Baes M., Saftly W., 2013, *A&A*, 560, A35
 Carnall A. C., Leja J., Johnson B. D., McLure R. J., Dunlop J. S., Conroy C., 2019, *ApJ*, 873, 44
 Carniani S. et al., 2024, *Nature*, 633, 318
 Casey C. M. et al., 2024, *ApJ*, 965, 98
 Castellano M. et al., 2023, *ApJ*, 948, L14
 Chabrier G., 2003, *PASP*, 115, 763
 Chaikin E., Schaye J., Schaller M., Benítez-Llambay A., Nobels F. S. J., Ploekinger S., 2023, *MNRAS*, 523, 3709
 Chaikin E. et al., 2026a, *MNRAS*, 548, 300
 Chaikin E. et al., 2026b, *MNRAS*, 548, 740
 Chemerynska I. et al., 2026, *MNRAS*, 546, 2267
 Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, *MNRAS*, 319, 168
 Correa C. A. et al., 2026, *MNRAS*, 548, 645
 Cowley W. I., Baugh C. M., Cole S., Frenk C. S., Lacey C. G., 2018, *MNRAS*, 474, 2352
 Crain R. A. et al., 2015, *MNRAS*, 450, 1937
 Cueto E. R., Hutter A., Dayal P., Gottlöber S., Heintz K. E., Mason C., Trebitsch M., Yepes G., 2024, *A&A*, 686, A138
 Cullen F. et al., 2023, *MNRAS*, 520, 14
 Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieeantsoa M. H., Appleby S., 2019, *MNRAS*, 486, 2827
 Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, *ApJ*, 292, 371
 De Lucia G., Fontanot F., Xie L., Hirschmann M., 2024, *A&A*, 687, A68
 Donnan C. T. et al., 2023, *MNRAS*, 518, 6011
 Donnan C. T. et al., 2024, *MNRAS*, 533, 3222
 Draine B. T., Li A., 2007, *ApJ*, 657, 810
 Dubois Y., Peirani S., Pichon C., Devriendt J., Gavazzi R., Welker C., Volonteri M., 2016, *MNRAS*, 463, 3948
 Durrant A. et al., 2026, preprint ([arXiv:2607.16404](https://arxiv.org/abs/2607.16404))
 Eddington A. S., 1913, *MNRAS*, 73, 359
 Eldridge J. J., Stanway E. R., Xiao L., McClelland L. A. S., Taylor G., Ng M., Greis S. M. L., Bray J. C., 2017, *Publ. Astron. Soc. Aust.*, 34, e058
 Faucher-Giguère C.-A., 2020, *MNRAS*, 493, 1614
 Ferrara A., Pallottini A., Dayal P., 2023, *MNRAS*, 522, 3986
 Finkelstein S. L. et al., 2015, *ApJ*, 810, 71
 Finkelstein S. L. et al., 2023, *ApJ*, 946, L13
 Finkelstein S. L. et al., 2024, *ApJ*, 969, L2
 Fontanot F. et al., 2026, preprint ([arXiv:2603.22405](https://arxiv.org/abs/2603.22405))
 Forouhar Moreno V. J., Helly J., McGibbon R., Schaye J., Schaller M., Han J., Kugel R., Bahé Y. M., 2025, *MNRAS*, 543, 1339
 Franco M. et al., 2025, preprint ([arXiv:2508.04791](https://arxiv.org/abs/2508.04791))
 Fujimoto S. et al., 2023, *ApJ*, 949, L25
 Gebek A. et al., 2026, preprint ([arXiv:2607.14901](https://arxiv.org/abs/2607.14901))
 Giallongo E. et al., 2015, *A&A*, 578, A83
 Hahn O., Michaux M., Rampf C., Uhlemann C., Angulo R. E., 2020, Astrophysics Source Code Library, record ascl:2008
 Han J., Cole S., Frenk C. S., Benítez-Llambay A., Helly J., 2018, *MNRAS*, 474, 604
 Harikane Y. et al., 2022, *ApJS*, 259, 20
 Harikane Y. et al., 2023, *ApJS*, 265, 5
 Harikane Y., Nakajima K., Ouchi M., Umeda H., Isobe Y., Ono Y., Xu Y., Zhang Y., 2024, *ApJ*, 960, 56
 Harikane Y. et al., 2025, *ApJ*, 980, 138
 Henden N. A., Puchwein E., Shen S., Sijacki D., 2018, *MNRAS*, 479, 5385
 Hutter A., Cueto E. R., Dayal P., Gottlöber S., Trebitsch M., Yepes G., 2025, *A&A*, 694, A254
 Huško F. et al., 2026, *MNRAS*, 547, 324

- Iocco F., Visinelli L., 2024, *Phys. Dark Univ.*, 44, 101496
- Kannan R. et al., 2023, *MNRAS*, 524, 2594
- Kapoor A. U. et al., 2023, *MNRAS*, 526, 3871
- Kapoor A. U. et al., 2024, *A&A*, 692, A79
- Kaviraj S. et al., 2017, *MNRAS*, 467, 4739
- Kreilgaard K. C. et al., 2026, *ApJ*, 1009, 134
- Labbé I. et al., 2023, *Nature*, 616, 266
- Lacey C. G. et al., 2016, *MNRAS*, 462, 3854
- Lagos C. d. P. et al., 2026, *MNRAS*, 549, stag947
- Laigle C. et al., 2016, *ApJS*, 224, 24
- Leja J., Carnall A. C., Johnson B. D., Conroy C., Speagle J. S., 2019, *ApJ*, 876, 3
- Leung F. C. K. et al., 2023, *ApJ*, 954, L46
- Liang F.-H., Li C., Li N., Zhou S., Yan R., Mo H., Zhang W., 2021, *ApJ*, 923, 120
- Lo Faro B., Buat V., Roehly Y., Alvarez-Marquez J., Burgarella D., Silva L., Efstathiou A., 2017, *MNRAS*, 472, 1372
- Lu S., Frenk C. S., Bose S., Lacey C. G., Cole S., Baugh C. M., Helly J. C., 2025, *MNRAS*, 536, 1018
- Lu S. et al., 2026, *MNRAS*, preprint (arXiv:2605.02022)
- Ludlow A. D., Schaye J., Schaller M., Richings J., 2019, *MNRAS*, 488, L123
- Ludlow A. D., Fall S. M., Wilkinson M. J., Schaye J., Obreschkow D., 2023, *MNRAS*, 525, 5614
- Ludlow A. D. et al., 2026, *MNRAS*, 550, stag1319
- Ma X. et al., 2018a, *MNRAS*, 477, 219
- Ma X. et al., 2018b, *MNRAS*, 478, 1694
- Madau P., Dickinson M., 2014, *ARA&A*, 52, 415
- Maiolino R. et al., 2024a, *Nature*, 627, 59
- Maiolino R. et al., 2024b, *A&A*, 691, A145
- Mason C. A., Trenti M., Treu T., 2023, *MNRAS*, 521, 497
- Matsuoka Y. et al., 2018, *ApJ*, 869, 150
- McAlpine S. et al., 2016, *Astron. Comput.*, 15, 72
- McGibbon R., Helly J., Schaye J., Schaller M., Vandenbroucke B., 2025, *J. Open Source Softw.*, 10, 8252
- McLeod D. J. et al., 2024, *MNRAS*, 527, 5004
- McLeod D. J. et al., 2026, preprint (arXiv:2604.16666)
- McLure R. J. et al., 2013, *MNRAS*, 432, 2696
- Meyer R. A. et al., 2024, *MNRAS*, 535, 1067
- Michaux M., Hahn O., Rampf C., Angulo R. E., 2021, *MNRAS*, 500, 663
- Mitchell P. D., Lacey C. G., Baugh C. M., Cole S., 2013, *MNRAS*, 435, 87
- Mitsuhashi I. et al., 2026, *ApJ*, 1000, 159
- Morishita T. et al., 2018, *ApJ*, 867, 150
- Naidu R. P. et al., 2026, *Open J. Astrophys.*, 9, 56033
- Nelson D. et al., 2019, *Comput. Astrophys. Cosmol.*, 6, 2
- Nobels F. S. J., Schaye J., Schaller M., Ploekinger S., Chaikin E., Richings A. J., 2024, *MNRAS*, 532, 3299
- Oesch P. A. et al., 2014, *ApJ*, 786, 108
- Oesch P. A. et al., 2016, *ApJ*, 819, 129
- Oesch P. A., Bouwens R. J., Illingworth G. D., Labbé I., Stefanon M., 2018, *ApJ*, 855, 105
- Oman K. A., 2025, *J. Open Source Softw.*, 10, 9278
- Pacifici C. et al., 2023, *ApJ*, 944, 141
- Page M. J., Carrera F. J., 2000, *MNRAS*, 311, 433
- Pakmor R. et al., 2023, *MNRAS*, 524, 2539
- Pérez-González P. G. et al., 2023, *ApJ*, 951, L1
- Ploekinger S., Richings A. J., Schaye J., Trayford J. W., Schaller M., Chaikin E., 2025, *MNRAS*, 543, 891
- Prada F., Behroozi P., Ishiyama T., Pérez E., Klypin A., Wang X., Ruedas J., 2026, preprint (arXiv:2604.18683)
- Raiter A., Fosbury R. A. E., Teimoorinia H., 2010, *A&A*, 510, A109
- Richings A. J., Schaye J., Oppenheimer B. D., 2014a, *MNRAS*, 440, 3349
- Richings A. J., Schaye J., Oppenheimer B. D., 2014b, *MNRAS*, 442, 2780
- Roberts-Borsani G. W. et al., 2016, *ApJ*, 823, 143
- Roberts-Borsani G. W., Ellis R. S., Laporte N., 2020, *MNRAS*, 497, 3440
- Robertson B. et al., 2024, *ApJ*, 970, 31
- Robertson B. E., Ellis R. S., Furlanetto S. R., Dunlop J. S., 2015, *ApJ*, 802, L19
- Rojas-Ruiz S. et al., 2025, *ApJ*, 985, 80
- Romano D., Matteucci F., Zhang Z.-Y., Papadopoulos P. P., Ivison R. J., 2017, *MNRAS*, 470, 401
- Saftly W., Camps P., Baes M., Gordon K. D., Vandewoude S., Rahimi A., Stalevski M., 2013, *A&A*, 554, A10
- Saftly W., Baes M., Camps P., 2014, *A&A*, 561, A77
- Salpeter E. E., 1955, *ApJ*, 121, 161
- Schaerer D., 2003, *A&A*, 397, 527
- Schaller M. et al., 2024, *MNRAS*, 530, 2378
- Schaye J. et al., 2015, *MNRAS*, 446, 521
- Schaye J. et al., 2023, *MNRAS*, 526, 4978
- Schaye J. et al., 2026, *MNRAS*, 548, stag375
- Schechter P., 1976, *ApJ*, 203, 297
- Schenker M. A. et al., 2013, *ApJ*, 768, 196
- Serjeant S., Bakx T. J. L. C., 2023, *Nat. Astron.*, 7, 1143
- Shen X., Vogelsberger M., Boylan-Kolchin M., Tacchella S., Kannan R., 2023, *MNRAS*, 525, 3254
- Sheth R. K., 2007, *MNRAS*, 378, 709
- Sneppen A., Steinhardt C. L., Hensley H., Jermyn A. S., Mostafa B., Weaver J. R., 2022, *ApJ*, 931, 57
- Somerville R. S., Primack J. R., 1999, *MNRAS*, 310, 1087
- Somerville R. S., Yung L. Y. A., Lancaster L., Menon S., Sommovigo L., Finkelstein S. L., 2025, *MNRAS*, 544, 3774
- Stanway E. R., Eldridge J. J., 2018, *MNRAS*, 479, 75
- Stefanon M. et al., 2019, *ApJ*, 883, 99
- Subbarao M. U., Connolly A. J., Szalay A. S., Koo D. C., 1996, *AJ*, 112, 929
- Sun G., Faucher-Giguère C.-A., Hayward C. C., Shen X., Wetzel A., Cochrane R. K., 2023, *ApJ*, 955, L35
- Trayford J. W. et al., 2026, *MNRAS*, 545, staf2040
- Tremonti C. A. et al., 2004, *ApJ*, 613, 898
- Ventura E. M., Qin Y., Balu S., Wyithe J. S. B., 2024, *MNRAS*, 529, 628
- Vijayan A. P. et al., 2026, preprint (arXiv:2607.26058)
- Willott C. J. et al., 2024, *ApJ*, 966, 74
- Yan Z., Jerabkova T., Kroupa P., 2017, *A&A*, 607, A126
- Yung L. Y. A., Somerville R. S., Finkelstein S. L., Wilkins S. M., Gardner J. P., 2024, *MNRAS*, 527, 5929

SUPPLEMENTARY MATERIAL

Supplementary material is available at [MNRAS](#) online.

COLIBRE_Highz_UVLF.zip Obs_Highz_UVLF.zip

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

APPENDIX A: EFFECT OF RANDOM SAMPLING

As mentioned in Section 2.2, we do not perform radiative transfer on all galaxies in the COLIBRE simulations due to the computational cost, but instead model a representative subsample. To assess the impact of this random sampling strategy on UVLFs, we carried out the following test. We first construct a subset of the standard sample (i.e. the originally selected sample; see Section 2.2) using the following procedure: for galaxies in the original subsample in a given stellar mass–SFR bin, if the number of galaxies exceeds 25 (half of the original target number; see Section 2.2 for details), we randomly select 25 galaxies; if it is fewer than 25, we include all (originally selected) galaxies. The sampling ratios are reduced accordingly. This new, smaller sample is referred to as the reduced sample, relative to the standard sample listed in Table 2. In Fig. A1, we compare the UVLFs derived from

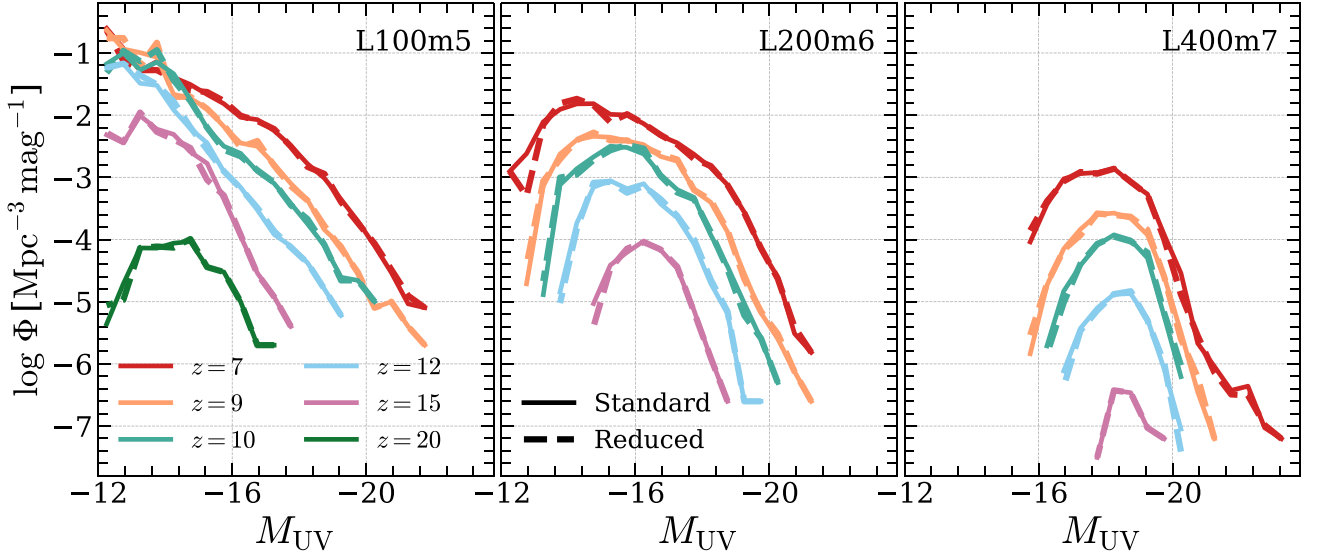


Figure A1. The effect of random sampling on the UVLFs (dust-attenuated). Results for different COLIBRE simulations are shown in separate panels (from left to right: L100m5, L200m6, and L400m7). In each panel, the UVLFs at different redshifts are indicated by different colours. The solid curves represent the UVLFs of the standard sample (i.e. the originally selected sample; see Section 2.2), while the dashed curves correspond to the UVLFs of the reduced sample (with smaller sample sizes; see Appendix A for details).

the standard and reduced samples across the three COLIBRE simulations with different resolutions. As shown, across all redshifts from $z = 7$ to 15 (and also $z = 20$ for L100m5), the UVLFs based on the reduced sample closely match those from the standard sample, indicating that even a smaller sample is sufficient to recover the UVLFs of the full galaxy population, highlighting the robustness and representativeness of our sampling method. This random sampling approach reduces the computational cost by approximately 99 per cent for L200m6 and L400m7 at $z = 7$, while preserving the accuracy of the resulting luminosity functions.

APPENDIX B: SAMPLE DISTRIBUTION AT VARIOUS REDSHIFTS

In Fig. B1, we present the distribution of selected samples of galaxies from the COLIBRE simulations on the $\log M_*$ – $\log \text{SFR}$ plane (as in Fig. 2) for redshifts $z = 9$ to 20. As shown, similarly to the trend observed for $z = 7$ in Fig. 2, rare objects with high stellar mass and high or low SFR have a higher selection probability.

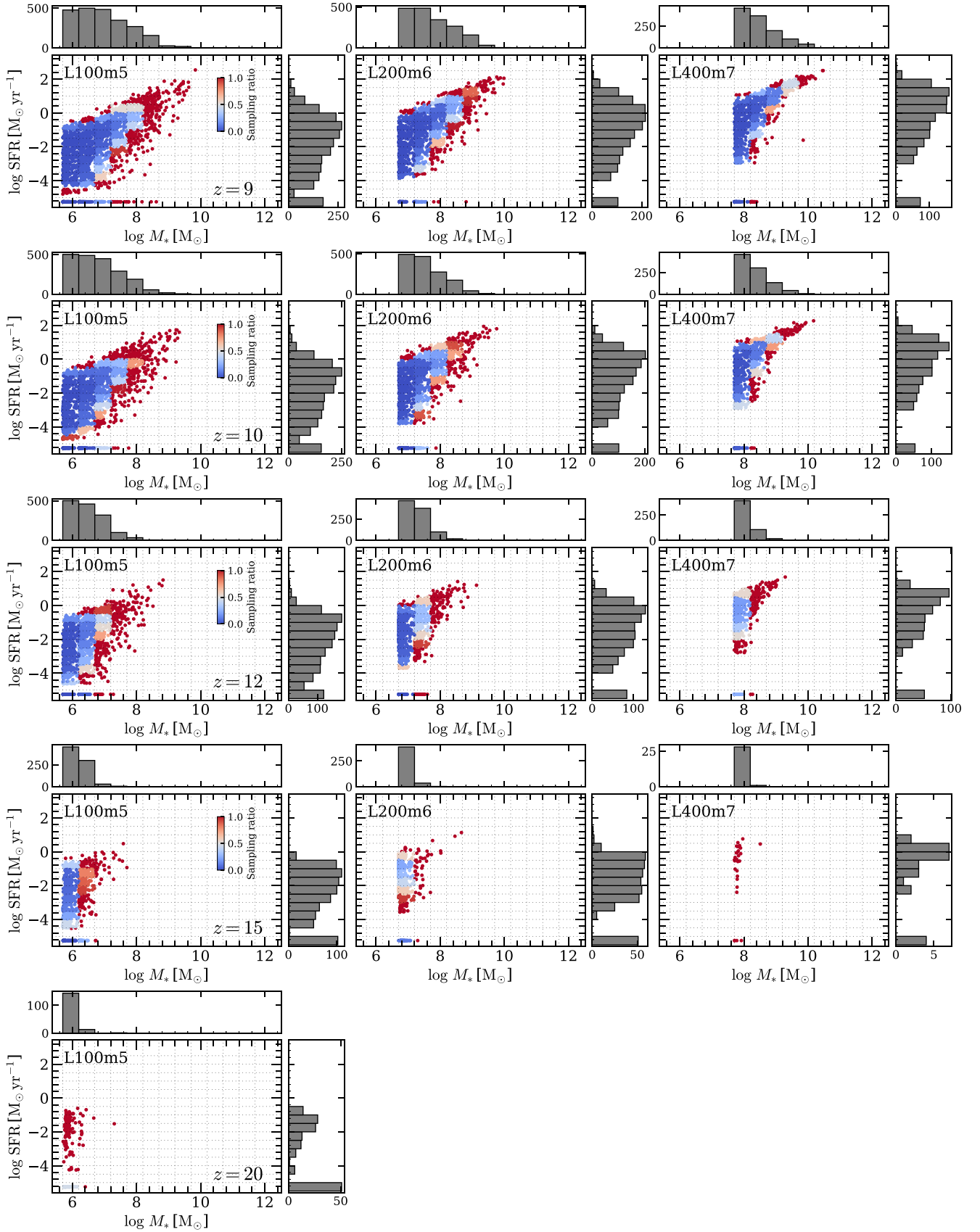


Figure B1. The distribution of selected galaxies on the $\log M_*$ – $\log \text{SFR}$ plane from $z = 9$ to 20 (from top to bottom). The symbols are as in Fig. 2.

APPENDIX C: EXPLORING THE IMPACT OF COLIBRE AND SKIRT SETUP CHOICES

C1 The effect of the feedback model

In Fig. C1, we investigate the effect of the AGN feedback model on the UVLFs. We present results from two COLIBRE L100m6 simulations that employ different AGN feedback models: the purely thermal model (solid; default in this work) and the hybrid (thermal + kinetic jet) model (dashed; see F. Huško et al. 2026 for details). The UVLFs, both dust-attenuated and dust-free, produced by the two models show only minor differences, with the slight discrepancy at the bright end likely due to the small number of galaxies in the simulation volume.

C2 The effect of star-forming region resampling

As mentioned in Section 2.3, we follow the fiducial SKIRT setup for COLIBRE described in A. Gebek et al. (2026) and resample the young stellar particles using the star-forming gas (both the currently existing star-forming gas and the parent gas of the young stellar particles) to mitigate the limited resolution of the simulations. Including the parent gas particles of recently formed stars ensures that the resampling conserves the mass of young

stars and captures the full star formation activity over the past 10 Myr. This approach enforces a constant star formation history over the past 10 Myr, such that the resulting stellar population includes stars with ages from nearly 0 to 10 Myr. This is important because the youngest stars make a disproportionate contribution to the ionizing UV luminosity. The impact of this resampling on the UV luminosities decreases with increasing numerical resolution, as the young stellar populations are then better directly resolved. Thus, one of the primary motivations for adopting this treatment is also to improve convergence across different resolutions. Another motivation is to interface with the TODDLERS model for star-forming regions. However, a limitation of this star-forming region resampling scheme is that it does not maintain full self-consistency with the dust modelling. In particular, the inferred UV flux is computed from resampled stellar populations that are, by construction, too young to have contributed dust through core-collapse SNe, while the dust content is not adjusted accordingly. In practice, however, this inconsistency is unlikely to be important for most galaxies, since their dust content is expected to be dominated by dust growth in the ISM enriched by earlier generations of stars. It may nevertheless be more relevant for the youngest systems and should therefore be borne in mind when interpreting the UV luminosities derived with this method.

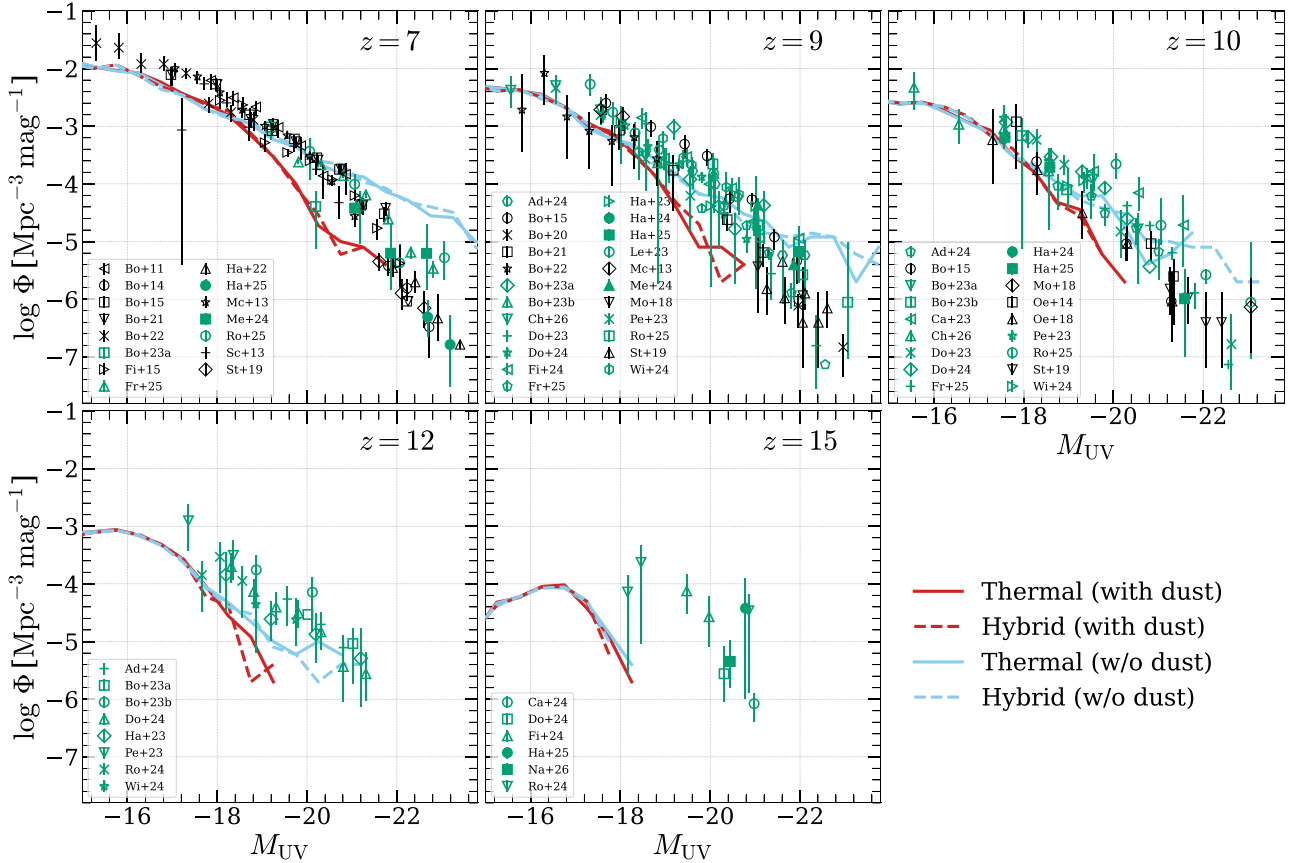


Figure C1. Effect of the AGN feedback model on the UVLFs. In each panel, the UVLFs with the thermal AGN feedback model (the default in this paper) and the hybrid (thermal + jet) feedback model (see F. Huško et al. 2026 for details) are shown with solid and dashed curves, respectively. Dust-attenuated and dust-free results are shown in red and blue, respectively. The COLIBRE L100m6 and L100m6h simulations are used for this test. The symbols for the observational data are the same as in Fig. 4.

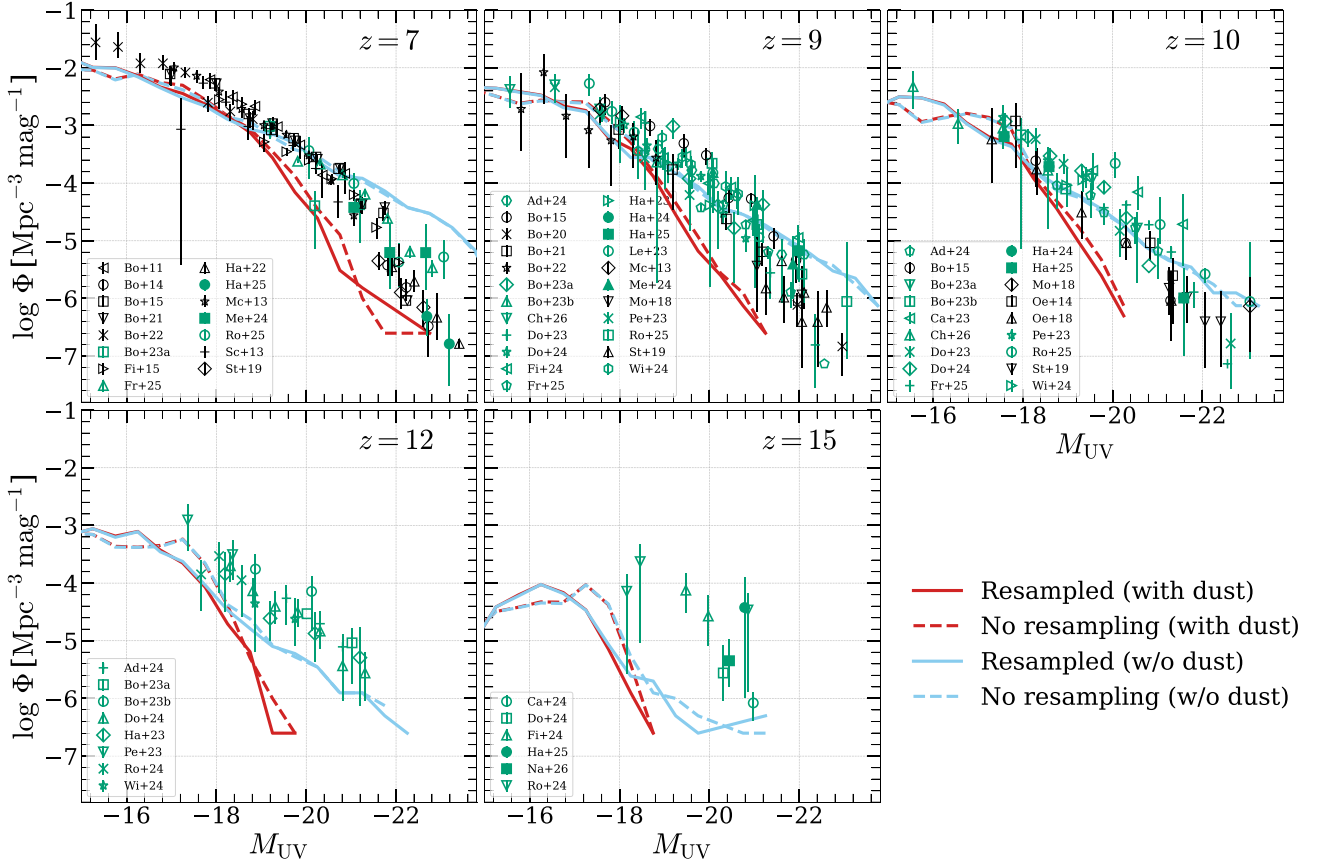


Figure C2. Effect of star-forming region resampling on the UVLFs. In each panel, the solid curve represents results with resampling (default in this work), while dashed curves represent results without resampling. Dust-attenuated and dust-free results are shown in red and blue, respectively. The L200m6 COLIBRE simulation is used for this test. Observational data points are the same as in Fig. 4.

In Fig. C2, we investigate the effect of this SFH resampling by comparing the UVLFs with resampling (the default in this paper) and those without resampling (i.e. including stellar particles of all ages and excluding the star-forming gas). As can be seen, at $z \leq 10$, the resampling procedure slightly reduces the dust-attenuated UV luminosities of galaxies at a given number density, except at the brightest end at $z = 7$, while the dust-free UV luminosities are nearly unchanged. At higher redshifts ($z \geq 12$), the effect of SFH resampling becomes larger for the dust-attenuated results and the dust-free results also show measurable differences. These differences in the UV luminosities are expected to propagate into the UV luminosity density, with SFH resampling lowering ρ_{UV} , particularly at $z \geq 12$. Nevertheless, the differences introduced by SFH resampling are not sufficiently large to account for the discrepancy between the predicted and observed UVLFs. Moreover, the increasingly close agreement between the dust-attenuated and dust-free results towards higher redshift indicates that the low dust content of these galaxies limits the impact of dust-related choices in the SKIRT radiative transfer modelling.

C3 The effect of the projected aperture size

In this section, we investigate the influence of the (projected) aperture size on the UVLFs. As mentioned in Section 2.3, we

adopt a projected circular aperture with a radius of 50 kpc for calculating the galaxy luminosity, to be consistent with other studies from the COLIBRE project (although a spherical aperture is typically used in some other studies). In Fig. C3, we present the combined, dust-attenuated UVLFs using different projected aperture sizes ($1 R_{*,0.5}$, $3 R_{*,0.5}$, $5 R_{*,0.5}$, 50 kpc, and an aperture including all gravitationally bound particles; here, $R_{*,0.5}$ is the half-stellar-mass radius provided by the SOAP catalogue).

As expected, UVLFs measured within larger apertures exhibit higher amplitudes, since larger apertures contain more particles (including both stellar particles and star-forming gas). The differences in UVLFs across aperture sizes can be as large as ~ 1 mag in the horizontal direction, highlighting the importance of aperture choice when comparing UVLFs. The UVLFs obtained with a 50 kpc aperture are nearly identical to those including all gravitationally bound particles, indicating that at $z > 7$, a 50 kpc aperture is sufficiently large to enclose essentially all particles associated with a given galaxy.

JWST photometric measurements of high-redshift galaxies do not rely on a single fixed physical aperture. Instead, they typically use small apertures to optimize the signal-to-noise ratio and then apply aperture corrections to estimate total fluxes. For example, Y. Harikane et al. (2023) measured colours in 0.3 arcsec-diameter circular apertures on PSF-matched images and estimated total magnitudes by applying an aperture correction defined relative

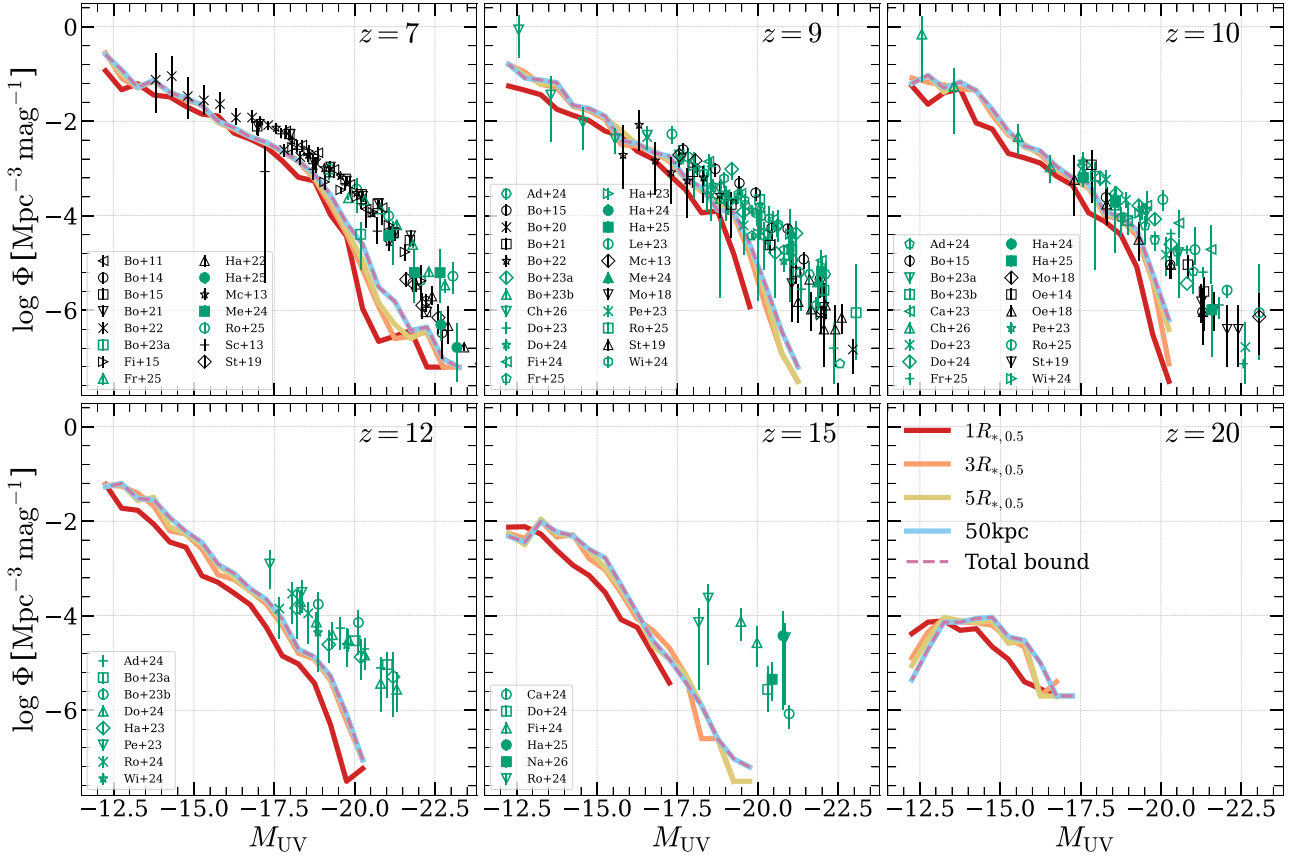


Figure C3. Comparisons of COLIBRE UVLFs (the dust-attenuated combined UVLFs for $z = 7\text{--}15$ and UVLF from L100m5 for $z = 20$) with different projected aperture sizes. In this work, we adopt a proper projected aperture of 50 kpc (blue) as the default. The symbols for observational data are the same as in Fig. 4.

to MAG_AUTO. S. L. Finkelstein et al. (2023), by contrast, adopted small elliptical Kron apertures (Kron factor 1.1; minimum radius 1.6) and derived total fluxes using a two-step, simulation-based aperture-correction procedure. Thus, although the observational implementations are different, both studies ultimately aim to recover total magnitudes rather than the raw fluxes measured within the initial apertures. In this context, our use of a 50 kpc

aperture is intended to encompass the full UV-emitting extent of each simulated galaxy and thereby provide an effectively total magnitude, making the comparison with these *JWST*-based measurements reasonable.

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