








The evolution of the galaxy stellar mass function and star formation rates in the COLIBRE simulations from redshift 17 to 0

Evgenii Chaikin ^{1,2}★, Joop Schaye ¹, Matthieu Schaller ^{1,3}, Sylvia Ploekinger ⁴,
Alejandro Benítez-Llambay,⁵ Carlos S. Frenk,² Filip Huško,¹ Robert J. McGibbon ¹,
Alexander J. Richings ^{6,7} and James W. Trayford ⁸

¹Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands

²Institute for Computational Cosmology, Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK

³Lorentz Institute for Theoretical Physics, Leiden University, PO Box 9506, NL-2300 RA Leiden, the Netherlands

⁴Department of Astrophysics, University of Vienna, Türkenschanzstrasse 17, A-1180 Vienna, Austria

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

⁶Centre for Data Science, Artificial Intelligence and Modelling, University of Hull, Cottingham Road, Hull HU6 7RX, UK

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

⁸Institute of Cosmology and Gravitation, University of Portsmouth, Dennis Sciama Building, Burnaby Road, Portsmouth PO1 3FX, UK

Accepted 2026 April 8. Received 2026 February 25; in original form 2025 September 9

ABSTRACT

We investigate the evolution of the galaxy stellar mass function (GSMF) and star formation rates (SFRs) across cosmic time in the COLIBRE simulations of galaxy formation. COLIBRE includes a multiphase interstellar medium, radiative cooling rates coupled to a model for the evolution of dust grains, and employs prescriptions for stellar and active galactic nucleus (AGN) feedback calibrated to reproduce the $z = 0$ observed GSMF and stellar mass–size relation. We present the evolution of the GSMF from simulations at three resolutions: $m_{\text{gas}} \approx m_{\text{dm}} \sim 10^7$, 10^6 , and $10^5 M_{\odot}$, in cosmological volumes of up to 400^3 , 200^3 , and 100^3 cMpc^3 , respectively. We demonstrate that COLIBRE is consistent with the observed GSMF over the full redshift range for which there are observations to compare with ($0 < z < 12$), with maximum systematic deviations of ≈ 0.3 dex reached at $2 < z < 4$. We also examine the evolution of the star-forming main sequence, cosmic SFR density, stellar mass density, and galaxy quenched fraction, making predictions for both the fiducial COLIBRE model with thermally driven AGN feedback and its variant with hybrid (thermal + kinetic jet) AGN feedback, and finding good agreement with observations. Notably, we show that COLIBRE matches the number density of massive quiescent galaxies at high redshifts reported by *James Webb Space Telescope* (*JWST*), while predicting a stellar-to-halo mass relation that evolves little with redshift. We conclude that neither a redshift-dependent star formation efficiency, nor a variable stellar initial mass function, nor a deviation from the Λ cold dark matter model is necessary to reproduce the high-redshift *JWST* stellar masses and SFRs.

Key words: methods: numerical – galaxies: evolution – galaxies: formation – galaxies: general.

1 INTRODUCTION

The galaxy stellar mass function (GSMF) is a fundamental statistic for studying the evolution of galaxies over cosmic time (e.g. S. Cole et al. 2001; C. Conroy & R. H. Wechsler 2009; L. Pozzetti et al. 2010). Defined as the comoving number density of galaxies per logarithmic stellar mass bin, $dn(M_*)/d \log M_*$, the observationally inferred GSMF not only provides one of the most precise probes of the galaxy population in the observable Universe, but also serves as a critical constraint for state-of-the-art numerical simulations of galaxy formation (e.g. P. Torrey et al. 2014; R. A. Crain et al. 2015).

The shape and normalization of the GSMF are intrinsically linked to those of the halo mass function (HMF), which can be predicted with remarkable accuracy over a wide range of masses, redshifts, and cosmologies (e.g. A. Jenkins et al. 2001; J. Tinker et al. 2008; S. Bocquet et al. 2016; J. Wang et al. 2020). However, the slopes at the low- and high-mass ends of the GSMF deviate significantly from that of the HMF due to the halo mass-dependent efficiency with which haloes convert their share of inflowing baryonic matter into stars. At the high-mass end, the conversion efficiency is controlled by the long cooling times of hot gas and energy feedback from active galactic nuclei (AGNs), while in low-mass haloes the dominant controlling mechanism is energy feedback from stellar winds, radiation, and supernovae (SNe; e.g. A. J. Benson et al. 2003; R. G. Bower et al. 2006; R. S. Somerville et al. 2008; M. Vogelsberger et al. 2013; R. A. Crain et al. 2015; C. D. P. Lagos et al. 2018; P. Behroozi et al. 2019).

* E-mail: evgenii.chaikin@durham.ac.uk

Direct measurements, empirical modelling, and halo abundance matching and clustering (e.g. V. R. Eke et al. 2004; P. S. Behroozi, C. Conroy & R. H. Wechsler 2010; B. P. Moster et al. 2010) indicate that the ratio of stellar to halo mass peaks at around 3 per cent at the halo mass of $M_{200c} \sim 10^{12} M_{\odot}$ and drops steeply towards both higher and lower values of M_{200c} .

At low redshifts ($z \lesssim 1$), the GSMF has been tightly constrained by large spectroscopic and photometric surveys, such as the Galaxy And Mass Assembly Survey (GAMA; S. P. Driver et al. 2011; I. K. Baldry et al. 2012; J. Liske et al. 2015; S. P. Driver et al. 2022), the Cosmic Evolution Survey (COSMOS; N. Scoville et al. 2007; I. Davidzon et al. 2017; J. R. Weaver et al. 2022), the Deep Extragalactic Visible Legacy Survey (DEVILS; L. J. M. Davies et al. 2018; J. E. Thorne et al. 2021), and, very recently, the Cosmic Dawn Survey (Euclid Collaboration 2025b, c). These surveys, some of which also probe higher redshifts, have constructed galaxy samples at $z \lesssim 1$ that are complete down to stellar masses of $M_{*} \sim 10^7$ - $10^9 M_{\odot}$ (depending on redshift), encompassing up to $\sim 10^6$ objects over areas of sky of ~ 1 - 10^2 deg^2 . Multiwavelength photometry – spanning (far-)UV (ultraviolet) to (far-)IR (infrared) wavelengths – is typically available in ≈ 10 -30 bands per galaxy. These extensive data enable robust stellar mass determination via spectral energy distribution (SED) fitting, with systematic uncertainties generally within ≈ 0.15 dex at $z \approx 0$ (e.g. A. S. G. Robotham et al. 2020). Furthermore, very recently, K. Xu et al. (2025) probed the faint end of the $z \approx 0$ GSMF down to $\sim 10^6 M_{\odot}$ by applying the statistical Photometric Objects Around Cosmic-webs (PAC) method to the Dark Energy Spectroscopic Instrument (DESI; DESI Collaboration 2016) survey data. PAC combines a large-scale (but relatively shallow) spectroscopic sample with deep photometric data to statistically associate faint galaxies (lacking spectroscopy) with nearby spectroscopic tracers, enabling redshift determination and stellar mass estimates for galaxies too faint to be targeted spectroscopically.

At higher redshifts, inferring the GSMF becomes increasingly challenging due to the declining surface brightness of galaxies, which requires longer integration times to probe the low-mass end ($M_{*} \lesssim 10^9 M_{\odot}$). Additionally, reliable stellar mass estimates from SED fitting require coverage of the rest-frame optical and near-IR regions of the spectrum (e.g. J. Walcher et al. 2011). At $z \gtrsim 3$, these wavelengths are redshifted into spectral bands that are difficult to observe with ground-based telescopes due to atmospheric absorption, a limitation that can only be circumvented with space-based observatories.

Prior to *James Webb Space Telescope* (*JWST*), estimates of the high- z GSMF were primarily based on data collected by space-based (far-)IR instruments such as *Spitzer*/IRAC and *HST*/WFC3, often supported by additional observations from ground-based facilities. For instance, M. Kajisawa et al. (2009) and P. Santini et al. (2012) utilized deep *Spitzer*/IRAC and *HST*/WFC3 IR imaging to estimate the GSMF up to $z \approx 3$ and $z \approx 4.5$, respectively. Similarly, K. Duncan et al. (2014), M. Song et al. (2016), and M. Stefanon et al. (2021) combined data from *HST*/ACS, *HST*/WFC3, and *Spitzer*/IRAC to derive the GSMF for redshift ranges of $z \approx 4$ -7, $z \approx 4$ -8, and $z \approx 6$ -10, respectively. These studies relied on galaxy samples of $\sim 10^3$ - 10^4 objects, covering areas of less than 1 deg^2 , with most galaxies lacking spectroscopically confirmed redshifts.

The unprecedented IR sensitivity and resolution of *JWST* have prompted a re-evaluation of how stellar mass accumulated in the high-redshift Universe. In particular, *JWST* observations at $z \gtrsim 4$

reveal a more rapid growth of massive galaxies than previously anticipated (e.g. A. J. Bunker et al. 2023; A. C. Carnall et al. 2024; A. Weibel et al. 2024; T. Wang et al. 2025), suggesting that galaxy formation efficiencies $\gtrsim 30$ per cent may be necessary to reach the observed stellar masses ($M_{*} \gtrsim 10^{10} M_{\odot}$) within the age of the Universe at the observed redshift (e.g. Y. Harikane et al. 2023; C. M. Casey et al. 2024). The SEDs of these massive galaxies often indicate significant dust content and suggest that some of these galaxies are already quenched (i.e. no longer actively forming stars) at those high redshifts – a challenge that most theoretical models struggle to reproduce (e.g. F. Valentino et al. 2023; W. M. Baker et al. 2025a; A. de Graaff et al. 2025; C. D. P. Lagos et al. 2025; A. Weibel et al. 2025b). Some of these conclusions, however, may be affected by significant uncertainties in stellar mass estimates (within ≈ 0.5 dex, e.g. R. K. Cochrane et al. 2025, T. Wang et al. 2025), cosmic variance due to the limited surface areas probed by high- z surveys, and the reliance on photometric redshifts for most galaxies.

The pivotal role of the GSMF in tracing the build-up of stellar mass and the hierarchical assembly of galaxies in the Universe, combined with the wealth of observational data available at both low and high redshifts, has established the GSMF as one of the most widely used observables for calibrating and validating galaxy formation models. Numerous galaxy formation simulations, including ILLUSTRIS (M. Vogelsberger et al. 2014), EAGLE (J. Schaye et al. 2015), BAHAMAS (I. G. McCarthy et al. 2017), ILLUSTRISTNG (A. Pillepich et al. 2018), SIMBA (R. Davé et al. 2019), and FLAMINGO (J. Schaye et al. 2023), as well as some semi-analytic models (e.g. R. S. Somerville et al. 2008; Q. Guo et al. 2011; R. S. Somerville et al. 2012; B. M. B. Henriques et al. 2015; C. D. P. Lagos et al. 2018, 2024), have been calibrated (via parameter tuning) to reproduce the $z \approx 0$ GSMF or used it as a loose constraint on the model parameters. In contrast, high-redshift ($z \gtrsim 3$) GSMFs are typically predictions of these simulations and semi-analytic models, as they are rarely adopted as observational constraints during calibration. This makes high- z GSMFs a powerful test of the underlying physics in galaxy formation models, particularly when compared to recent observations from *JWST*.

We emphasize, however, that while galaxy stellar masses – and thus the GSMF – can be directly predicted in simulations, they are not directly measurable in observations. Instead, they must be inferred from photometric and/or spectroscopic data using stellar population synthesis (SPS) models, which depend on several assumptions, including the stellar initial mass function (IMF), dust attenuation, metallicity, and star formation history (SFH), each of which can introduce systematic uncertainties (e.g. J. Pforr, C. Maraston & C. Tonini 2012; B. Mobasher et al. 2015; S. Bellstedt & A. S. G. Robotham 2025; J. H. Lee et al. 2025).

In this work, we use the new COLIBRE cosmological hydrodynamical simulations of galaxy formation (E. Chaikin et al. 2026; F. Huško et al. 2026; J. Schaye et al. 2026) to investigate the build-up of stellar mass in the Universe across cosmic time. We compare simulation predictions with observational data from $z = 0$ up to the highest redshift with available observations ($z \approx 12$ for the GSMF) and make predictions out to $z = 17$. Unlike previous galaxy formation models applied to cosmological volumes of comparable size, COLIBRE includes a multiphase interstellar medium (ISM) with non-equilibrium chemistry for hydrogen and helium (S. Ploekinger et al. 2025), employs a detailed model for the formation and evolution of dust grains coupled to the

Table 1. COLIBRE simulations used in this work. Column (1): simulation name; column (2): size of the cosmological volume per dimension; column (3): the mean initial gas particle mass; column (4): the mean DM particle mass; column (5): number of gas particles in the initial conditions; column (6): number of DM particles in the initial conditions; column (7): Plummer-equivalent gravitational softening length used for baryons and DM; column (8): AGN feedback model used in the simulation; column (9): the redshift down to which the simulation has been run at the time of writing.

Name	L_{box} (cMpc)	m_{gas} (M_{\odot})	m_{dm} (M_{\odot})	N_{gas}	N_{dm}	ϵ_{soft}	AGN feedback model	z
m7 resolution								
L400m7	400	1.47×10^7	1.94×10^7	3008^3	4×3008^3	min(1.4 pkpc, 3.6 ckpc)	Thermal	0
L200m7	200	1.47×10^7	1.94×10^7	1504^3	4×1504^3	min(1.4 pkpc, 3.6 ckpc)	Thermal	0
L100m7	100	1.47×10^7	1.94×10^7	752^3	4×752^3	min(1.4 pkpc, 3.6 ckpc)	Thermal	0
L050m7	50	1.47×10^7	1.94×10^7	376^3	4×376^3	min(1.4 pkpc, 3.6 ckpc)	Thermal	0
L025m7	25	1.47×10^7	1.94×10^7	188^3	4×188^3	min(1.4 pkpc, 3.6 ckpc)	Thermal	0
L200m7h	200	1.47×10^7	1.94×10^7	1504^3	4×1504^3	min(1.4 pkpc, 3.6 ckpc)	Hybrid	0
m6 resolution								
L200m6	200	1.8×10^6	2.4×10^6	3008^3	4×3008^3	min(0.7 pkpc, 1.8 ckpc)	Thermal	0
L100m6	100	1.8×10^6	2.4×10^6	1504^3	4×1504^3	min(0.7 pkpc, 1.8 ckpc)	Thermal	0
L050m6	50	1.8×10^6	2.4×10^6	752^3	4×752^3	min(0.7 pkpc, 1.8 ckpc)	Thermal	0
L025m6	25	1.8×10^6	2.4×10^6	376^3	4×376^3	min(0.7 pkpc, 1.8 ckpc)	Thermal	0
m5 resolution								
L100m5	100	2.3×10^5	3.0×10^5	3008^3	4×3008^3	min(0.35 pkpc, 0.9 ckpc)	Thermal	4.6
L050m5	50	2.3×10^5	3.0×10^5	1504^3	4×1504^3	min(0.35 pkpc, 0.9 ckpc)	Thermal	0.9
L025m5	25	2.3×10^5	3.0×10^5	752^3	4×752^3	min(0.35 pkpc, 0.9 ckpc)	Thermal	0

chemistry (J. W. Trayford et al. 2026), and suppresses spurious transfer of energy from dynamically hot dark matter (DM) to dynamically-cooler stars by using four times as many DM particles as baryonic particles. In addition to the GSMF, we examine several other statistics closely tied to stellar mass evolution: the cosmic star formation rate (SFR) and stellar mass densities, the stellar-to-halo mass relation (SHMR), the star-forming main sequence (SFMS), the galaxy quenched fraction – stellar mass relation, and the number density of massive quiescent galaxies in the high-redshift Universe. The paper is structured as follows: Section 2 details the simulation methodology, Section 3 presents our findings, and Section 4 summarizes our key conclusions.

2 METHODS

2.1 The COLIBRE simulations

We use the COLIBRE simulations of galaxy formation, detailed in J. Schaye et al. (2026). All simulations were run using the astrophysical code SWIFT (M. Schaller et al. 2024), which utilizes task-based parallelism. Gas hydrodynamics is modelled with the density-energy smoothed particle hydrodynamics (SPH) scheme SPHENIX (J. Borrow et al. 2022). Gravity is solved using a Fourier-space particle-mesh method for long-range forces and the Fast Multipole Method for short-range forces, employing fixed softening lengths for baryons and DM.

The initial conditions for the COLIBRE simulations were generated at $z = 63$ by the MONOFONIC code (O. Hahn et al. 2020; M. Michaux et al. 2021) using second-order Lagrangian perturbation theory. The assumed cosmology for all simulations is the Λ CDM ‘3x2pt + all external constraints’ cosmology from T. M. C. Abbott et al. (2022): $\Omega_{\text{m},0} = 0.306$, $\Omega_{\text{b},0} = 0.0486$, $\sigma_8 = 0.807$, $h = 0.681$, $n_s = 0.967$, alongside a single massive neutrino species with a mass of 0.06 eV.

The COLIBRE simulations are available at three resolutions: m7 (gas and DM particle masses of $m_{\text{gas}} = 1.47 \times 10^7 M_{\odot}$ and $m_{\text{dm}} = 1.94 \times 10^7 M_{\odot}$, respectively), m6 ($m_{\text{gas}} = 1.8 \times 10^6 M_{\odot}$,

$m_{\text{dm}} = 2.4 \times 10^6 M_{\odot}$), and m5 ($m_{\text{gas}} = 2.3 \times 10^5 M_{\odot}$, $m_{\text{dm}} = 3.0 \times 10^5 M_{\odot}$). In contrast to previous cosmological simulations, which generally used similar numbers of DM particles and gas resolution elements in the initial conditions (with the DM particle mass of $m_{\text{dm}} = [(\Omega_{\text{m},0} - \Omega_{\text{b},0})/\Omega_{\text{b},0}] m_{\text{gas}} \approx 5 m_{\text{gas}}$), COLIBRE uses four times as many DM particles as baryonic particles to reduce spurious energy transfer from DM to stars, which can have adverse effects on galaxy stellar components (A. D. Ludlow, J. Schaye & R. Bower 2019; A. D. Ludlow et al. 2021, 2023; M. J. Wilkinson et al. 2023). At m7 resolution, the Plummer-equivalent gravitational softening length of baryons and DM is set to the minimum of 1.4 proper kpc (pkpc) and 3.6 comoving kpc (ckpc), while at m6 and m5 resolutions it is equal to min(0.7 pkpc, 1.8 ckpc) and min(0.35 pkpc, 0.9 ckpc), respectively.

In this work, we use the COLIBRE simulations at m7 resolution with comoving volumes of 400^3 and 200^3 cMpc³, at m6 resolution with a volume of 200^3 cMpc³, and at m5 resolution with 100^3 , 50^3 , and 25^3 cMpc³. Additionally, to assess the separate convergence with resolution and with cosmological volume (see Appendix A), we use smaller volumes at m7 and m6 resolutions: 100^3 cMpc³, 50^3 cMpc³, and 25^3 cMpc³. Table 1 provides a summary of all simulations used in this study. Unless specified otherwise, we use the largest available simulation at each resolution.¹

2.1.1 The subgrid model

The radiative cooling and heating rates for primordial elements and their free electrons are computed using the non-equilibrium thermochemistry solver CHIMES (A. J. Richings, J. Schaye & B.

¹At the time of writing, the m5 simulation in the 100^3 (50^3) cMpc³ volume is at $z = 4.6$ ($z = 0.9$). Therefore, unless stated otherwise, results at m5 resolution are shown from the largest available volume at a given redshift, while time-evolving quantities are interpolated around $z = 4.6$ and $z = 0.9$, switching from L100m5 to L050m5 and from L050m5 to L025m5, respectively.

D. Oppenheimer 2014a, b). The rates for nine metals – C, N, O, Ne, Mg, Si, S, Ca, and Fe – are provided by HYBRID-CHIMES (S. Ploeckinger et al. 2025), which are based on species fractions calculated by CHIMES under the assumption of ionization equilibrium and steady state chemistry, but rescaled to account for the difference between the non-equilibrium and equilibrium free electron number densities. The radiative cooling and heating rates account for the presence of a redshift-dependent homogeneous radiation background from distant galaxies and quasars and – in the ISM – an interstellar radiation field (ISRF), cosmic rays, dust, and shielding. Because COLIBRE does not include radiative transfer calculations, the strength of the ISRF, the cosmic ray rate, and the shielding column density are determined by a density- and temperature-dependent characteristic length-scale, based on the local Jeans length. The details of the combined non-equilibrium and quasi-equilibrium radiative cooling model used in COLIBRE can be found in S. Ploeckinger et al. (2025).

The gas chemical composition is tracked by modelling the abundances of 12 individual elements:² H, He, C, N, O, Ne, Mg, Si, Fe, Sr, Ba, and Eu. These abundances are diffused among SPH neighbours using a velocity shear-based subgrid model for turbulent mixing (C. A. Correa et al. 2026). Additionally, the COLIBRE simulations include a subgrid model for the evolution of dust grains, which is coupled to the chemistry and non-equilibrium cooling rates (J. W. Trayford et al. 2026). Graphite and silicate grains are produced by asymptotic giant branch (AGB) stars and SNe. Dust grains grow by accreting mass from the gas phase, are destroyed via thermal sputtering and shocks, and undergo size modifications through shattering and coagulation. Two grain sizes are explicitly modelled: 0.01 and 0.1 μm . The dust model is coupled to the CHIMES solver, where evolving dust abundances influence gas heating and cooling rates and facilitate the formation of molecular hydrogen.

Following F. S. J. Nobels et al. (2024), star-forming gas particles are identified using the gravitational instability criterion. To assign an SFR to these particles, we apply the M. Schmidt (1959) law, assuming a fixed star formation efficiency per free-fall time of $\varepsilon = 0.01$. Newly formed stellar particles represent simple stellar populations characterized by a G. Chabrier (2003) IMF. Stellar particles enrich their surrounding gas with metals through six chemical enrichment channels: AGB stars, type-Ia SNe, type-II SNe, neutron star mergers, common envelope jet SNe, and collapsars (see C. A. Correa et al. 2026, for further details).

Energy feedback from core-collapse supernovae (CC SNe) is implemented using a stochastic thermal-kinetic feedback prescription following E. Chaikin et al. (2023) with two modifications. First, the constant heating temperature of $\Delta T_{\text{SN}} = 10^{7.5}$ K is replaced by a density-dependent function, $\Delta T_{\text{SN}}(\rho_{\text{gas}}) \propto \rho_{\text{gas}}^{2/3}$, where ρ_{gas} is the density of the gas around the SN estimated at the time of feedback. Secondly, the constant energy per single CC SN (in units of 10^{51} erg), f_E , is replaced by a pressure-dependent function, $f_E(P_{\text{birth}})$, where P_{birth} represents the thermal pressure

of the parent gas particles measured at the time when they were converted into stellar particles. For details, see J. Schaye et al. (2026).

In addition to CC SN feedback, COLIBRE incorporates stochastic thermal energy feedback from type-Ia SNe, applying the same implementation used for CC SN thermal energy injections, but with $f_E = 1$. The number of individual type-Ia SNe per stellar particle is determined by an exponential delay time distribution function with a normalization of 1.54×10^{-3} individual type-Ia SNe per solar mass and an exponential time-scale of 2 Gyr (see Nobels et al., in preparation, for details).

COLIBRE also includes three early stellar feedback processes: H II regions, stellar winds, and radiation pressure (A. Benítez-Llambay et al. 2026). These processes are implemented stochastically, similarly to the energy feedback from SNe.

Supermassive black holes (SMBHs) are represented by black hole (BH) particles, seeded at high redshift by converting the densest gas particle in a halo into a BH particle. Seeding occurs when the halo mass, calculated on-the-fly using the Friends-of-Friends (FoF) algorithm, exceeds a threshold of $5 \times 10^{10} M_{\odot}$ at m7 resolution or $10^{10} M_{\odot}$ at m6 and m5 resolutions, provided the FoF halo does not already contain a BH particle. The mass accretion rate onto BHs is computed using the modified Bondi–Hoyle–Lyttleton formula, with turbulence and vorticity corrections from M. R. Krumholz, C. F. McKee & R. I. Klein (2006). Dynamical friction and BH-BH mergers are modelled following the prescriptions from Y. M. Bahé et al. (2022) and J. Schaye et al. (2026).

At each resolution, the COLIBRE simulations include models with purely thermal AGN feedback and hybrid AGN feedback (combining kinetic jets and thermal energy injections), with the largest COLIBRE volumes available only for the thermal models (see table 2 in J. Schaye et al. 2026). The primary focus of this work is on the COLIBRE simulations with purely thermal AGN feedback, though we provide a comparison between the hybrid and thermal models in Section 3.8.

Thermal AGN feedback is implemented deterministically, following C. M. Booth & J. Schaye (2009), but with BH particles heating their neighbouring gas by a temperature increment proportional to the BH (subgrid) mass, $\Delta T_{\text{AGN}} \propto m_{\text{BH}}$, rather than by a fixed temperature increment. The BH mass-dependent heating ensures that the frequency of AGN feedback events remains independent of BH mass at a fixed Eddington fraction, and that at the highest COLIBRE resolution, m5, the BH mass does not increase too steeply with galaxy stellar mass (see E. Chaikin et al. 2026 for details).

The hybrid AGN model detailed in F. Huško et al. (2026), additionally tracks the evolution of BH spin and distinguishes between three accretion states: the thick, thin and slim discs. The spin of each BH particle evolves over time due to gas accretion, BH–BH mergers, and jet-induced spin-down. The kinetic jet feedback mode is implemented deterministically: once a BH particle has accreted a mass whose energy equivalent exceeds that of a single jet event, two neighbouring gas particles are kicked in opposite directions, nearly aligned with the BH spin (within 7.5 degrees of the spin axis), with the jet kick velocity proportional to the square root of the BH (subgrid) mass, such that the injected energy scales linearly with BH mass. The thermal channel for energy injection is implemented identically to that in the purely thermal AGN feedback model, but with a spin-dependent

²This set of elements differs from those used in the prescription for gas radiative cooling. In particular, Sr, Ba, and Eu are not included in the cooling calculations, as their contributions to the total cooling rate are negligible. Conversely, Ca and S – elements not tracked in the COLIBRE chemistry but required for the radiative cooling – are assumed to have solar mass ratios relative to Si (M. Asplund et al. 2009).

radiative efficiency for the thin disc, while separate terms are used to represent the effects of winds in the thick and slim disc states. The amounts of energy available for injection through the thermal and kinetic jet channels are determined based on the BH's instantaneous accretion rate, expressed as a fraction of the Eddington rate, and its spin. In both the thermal and hybrid AGN feedback models, the BH accretion rate is capped at 100 times the Eddington rate, allowing for super-Eddington accretion events.

2.1.2 Calibration

As described in detail in E. Chaikin et al. (2026), the SN and AGN energy feedback in the COLIBRE simulations has been calibrated to reproduce the observed $z = 0$ GSMF from GAMA DR4 (S. P. Driver et al. 2022) and the observed $z \approx 0$ galaxy stellar mass–size relation reported by J. A. Hardwick et al. (2022). For the COLIBRE models with thermal AGN feedback, the calibration involved adjusting the strength and functional form of SN and AGN energy feedback. At m7 resolution, this was achieved by training Gaussian process emulators on ~ 200 simulations in a (50 Mpc)³ volume with varied SN and AGN feedback parameter values. The trained emulators were then fit to observational data in the stellar mass range $10^9 < M_*/M_\odot < 10^{11.3}$ to infer the best-fitting parameter values. At m6 and m5 resolutions, the calibration was performed through small manual adjustments of the subgrid parameters for SN and AGN feedback, relative to their best-fitting values determined using the emulators at m7 resolution. Additionally, and independently from the calibration of all other feedback parameters, the coupling efficiency of AGN feedback was tuned to match the $z = 0$ observed BH mass–stellar mass relation of massive galaxies for which dynamical measurements of BH masses are possible.

For the COLIBRE models incorporating hybrid AGN feedback, AGN feedback parameters were manually tuned at each resolution to match the same calibration data as for the thermal models, as well as the $z = 0.2$ AGN bolometric luminosity function (see F. Huško et al. 2026). Among the SN feedback parameters, only the normalization of the $\Delta T_{\text{SN}}(\rho_{\text{gas}}) \propto \rho_{\text{gas}}^{2/3}$ scaling was adjusted, while all other SN feedback parameters retained the same values as in the COLIBRE models with thermal AGN feedback at the corresponding resolution.

2.2 Halo finding and calculation of galaxy properties

We identify subhaloes using HBT-HERONS³ (V. J. Forouhar Moreno et al. 2025), which is an updated version of the Hierarchical Bound Tracing algorithm (HBT+) originally presented by J. Han et al. (2018). HBT-HERONS identifies structures by applying an iterative unbinding algorithm to particles within spatial FoF groups, starting at the highest redshift simulation output. Once self-bound objects are identified, their particles are tracked across subsequent simulation outputs to trace the evolution of these objects forward in time. This tracking improves the identification of satellite subhaloes, as particles remain associated with their subhaloes even after the subhalo is accreted into a larger host halo. Each candidate substructure undergoes additional checks for self-boundness and phase-space consistency to determine its

status – whether it remains resolved, has merged with another substructure, or has been disrupted.

We further process the output of HBT-HERONS using the Spherical Overdensity and Aperture Processor (SOAP; R. McGibbon et al. 2025) to obtain a comprehensive list of subhalo and galaxy properties in various apertures. This list includes, in particular, galaxy stellar masses and SFRs used in this work. For a given subhalo,

(i) We use the stellar mass computed by summing up the masses of individual stellar particles that are both gravitationally bound to the subhalo and located within three-dimensional (3D) spherical apertures of radius 50 pkpc. The aperture is centred on the position of the particle with the largest binding energy within the subhalo. We chose the 50 pkpc aperture because A. de Graaff et al. (2022) found it best reproduces masses inferred from virtual observations of the EAGLE simulations. We discuss the impact of the aperture choice on the GSMF in Appendix B, finding a noticeable effect only for stellar masses $M_* \gtrsim 10^{11} M_\odot$.

(ii) The SFR is calculated by summing the instantaneous SFRs of gas particles that are both bound to the subhalo and within a 50 pkpc 3D spherical aperture⁴. Additionally, in Appendix C, we explore the impact of using 10- and 100-Myr averaged SFRs instead of instantaneous SFRs, finding generally minor differences at all redshifts.

(iii) The specific SFR (sSFR) is calculated by dividing the subhalo's instantaneous SFR by its stellar mass, both measured at the same time and within the same apertures.

We use SOAP outputs at integer redshifts from $z = 20$ to $z = 12$; at integer and half-integer redshifts from $z = 12$ to $z = 10$; at integer, half-, and quarter-integer redshifts from $z = 10$ to $z = 1$; and at lower redshifts with a step of $\Delta z = 0.05$, except at $z = 0.15$, where we instead use $z = 0.14$ due to the structure of the COLIBRE output redshifts.

2.3 Assumptions in the analysis

Throughout this work, we assume a G. Chabrier (2003) IMF. All observational data based on E. E. Salpeter (1955) or P. Kroupa (2001) IMFs that we include for comparison with the simulations are converted to the G. Chabrier (2003) IMF using the conversion factors for SFRs and stellar masses provided by P. Madau & M. Dickinson (2014). Additionally, we account for differences in the assumed value of the dimensionless Hubble parameter h between cosmologies used in the comparison data, h_{obs} , and in COLIBRE, $h = 0.681$. This is done by applying the scaling factor $(h_{\text{obs}}/0.681)^\alpha$ to the reported observational volumes, stellar masses, and SFRs, where the exponent α is chosen according to each quantity's dependence on cosmological distances.

Random errors in observationally inferred stellar masses, arising from the application of SED-fitting algorithms to broad-band photometric data and uncertainties in SPS modelling, are estimated to be ≈ 0.1 - 0.2 dex at $z \lesssim 1$ (e.g. A. Gallazzi & E. F. Bell 2009; E. N. Taylor et al. 2011; J. S. Speagle et al. 2014; B. Mobasher et al. 2015; A. S. G. Robotham et al. 2020; G. T. Jones, E. R. Stanway & A. C. Carnall 2022; J. H. Lee et al. 2025), and ≈ 0.2 - 0.3

⁴We verified that the results of this work are largely insensitive to whether galaxy SFRs are computed using all star-forming gas particles within a 50 kpc aperture or only the gravitationally bound particles within that same distance.

³<https://hbt-herons.strw.leidenuniv.nl/>

dex or higher at $z \gtrsim 1$ (e.g. C. Conroy, J. E. Gunn & M. White 2009; C. Pacifici et al. 2023; P. Santini et al. 2023; R. K. Cochrane et al. 2025). These random errors affect the shape of the observed GSMF through the A. S. Eddington (1913) bias, particularly at the high-mass end. Due to the steep slope of the GSMF at high M_* , there are significantly more low-mass galaxies than high-mass ones, leading to a greater upward scatter of random errors in stellar mass of low-mass objects compared to the downward scatter of high-mass objects. As we will explore in Appendix D, this effect has an impact on any statistic related to galaxy stellar mass, including the GSMF, the SFMS, the galaxy quenched fraction, and the number density of massive quiescent galaxies.

To account for this Eddington bias, we add a log-normal scatter to galaxy stellar masses in the simulations, with a mean of zero and a redshift-dependent standard deviation of $\log_{10} M_*$ following the parametrization of P. Behroozi et al. (2019),

$$\sigma_{\text{random}}(z) = \min(\sigma_{\text{random},0} + \sigma_{\text{random},z} z, \sigma_{\text{random},\text{cap}}) \text{ dex}. \quad (1)$$

Based on the values of random errors reported in the literature, we choose $\sigma_{\text{random},0} = 0.1$, $\sigma_{\text{random},z} = 0.1$, and $\sigma_{\text{random},\text{cap}} = 0.3$. Unless stated otherwise, this scatter is included in all subsequent figures presenting statistics that depend on M_* . In particular, we note the added scatter also affects galaxy sSFR values, as M_* appears in the denominator of the sSFR definition.

In addition to random errors, SED-fitting codes introduce *systematic* uncertainties in the inferred stellar masses, as the assumptions underlying SED modelling – most notably the choice of SPS model, SFH parameterization, and dust attenuation – affect *all* galaxies within a given sample. By applying different SED-fitting pipelines with varying assumptions to the same data, systematic errors have been estimated to be ≈ 0.1 - 0.3 dex for $0 \lesssim z \lesssim 5$ (A. Muzzin et al. 2009; J. Leja et al. 2019; A. Katsianis et al. 2020; A. S. G. Robotham et al. 2020; G. T. Jones et al. 2022; J. H. Lee et al. 2025; M. Shuntov et al. 2025). At higher redshifts, where photometric data are more limited, these estimates become progressively more uncertain (e.g. C. Papovich et al. 2023; D. Narayanan et al. 2024; T. Harvey et al. 2025; T. Wang et al. 2025). However, R. K. Cochrane et al. (2025) showed that even when relying solely on *JWST*/NIRCam photometry, systematic errors remain within ≈ 0.4 dex up to $z \approx 10$; though these estimates neglect the effect of IMF choice, which can introduce additional systematic errors (e.g. A. Trinca et al. 2024).

Unlike the Eddington bias, which primarily affects the GSMF by *increasing* the number density of massive galaxies while having a negligible impact at low M_* , systematic errors can shift the GSMF in *either direction* – leftward or rightward. Therefore, we do not add a systematic error to the simulated galaxy stellar masses but instead show it for reference in the relevant plots (see below), assuming the parametrization,

$$\sigma_{\text{sys}}(z) = \min(\sigma_{\text{sys},0} + \sigma_{\text{sys},z} \sqrt{z}, \sigma_{\text{sys},\text{cap}}) \text{ dex}, \quad (2)$$

where $\sigma_{\text{sys},0} = 0.15$, $\sigma_{\text{sys},z} = 0.08$, and $\sigma_{\text{sys},\text{cap}} = 0.4$ based on the values reported in the literature. Finally, we note that systematic errors can vary with galaxy stellar mass by up to ≈ 0.15 dex relative to the average offset (see e.g. fig. 3 in J. Leja et al. (2019) and fig. F1 in M. Shuntov et al. 2025), an effect that we do not incorporate into equation (2).

3 RESULTS

3.1 Evolution of the galaxy stellar mass function

Fig. 1 shows the evolution of the GSMF in the COLIBRE simulations from $z = 0$ to $z = 4$, and Fig. 2 extends the evolution to higher redshifts, from $z = 5$ to $z = 17$.⁵ The GSMF is computed using fixed mass bins of 0.2 dex. The COLIBRE models with thermal AGN feedback at m7, m6, and m5 resolutions are depicted by the solid dark-red, orange, and light-blue curves, respectively. To highlight limitations due to resolution and volume, the curves transition from solid to dotted at stellar masses below 10 times the mean initial baryonic particle mass (i.e. where galaxies contain fewer than ~ 10 stellar particles) or where fewer than 10 galaxies are present per bin. We emphasize that the chosen resolution limit is shown only to guide the eye and does not necessarily correspond to the true resolution limit. The binned GSMF values for the L200m6 simulation at $0 < z < 17$ are provided in Appendix E, and those for other simulations are available on the COLIBRE website.⁶

As a comparison, we present the following observed GSMFs from the literature:

(i) The GSMF from GAMA DR4 at $z = 0$ (S. P. Driver et al. 2022), to which the COLIBRE simulations were calibrated. GAMA is a highly complete multiwavelength survey with over 200 000 spectroscopic redshifts. Stellar masses were inferred using UV-to-IR SED fitting with the Bayesian code PROSPECT (A. S. G. Robotham et al. 2020), assuming G. Bruzual & S. Charlot (2003) stellar population models. The analysis employed parametric SFHs modelled as a skewed-Normal distribution combined with a prescription for time-evolving metallicity (for more details, see S. Bellstedt et al. 2020). We use the GAMA GSMF derived from the data across four primary GAMA regions (in total covering ≈ 230 deg²), corrected to evolve to $z = 0$ and re-normalized to match the larger Sloan Digital Sky Survey (SDSS) area of ≈ 5012 deg².

(ii) The $0 < z < 4$ GSMF from the DEVILS survey (L. J. M. Davies et al. 2018), based on the D10-COSMOS field (≈ 1.5 deg²), as reported by J. E. Thorne et al. (2021). Stellar masses were derived by applying the PROSPECT code to ≈ 494 000 galaxies using the same data processing pipeline as for the GAMA survey.

(iii) The $0.2 < z < 7.5$ GSMF from the COSMOS2020 photometric redshift catalogue (J. R. Weaver et al. 2023b), based on $\sim 10^6$ objects observed over an effective area of ≈ 1.27 deg². Stellar masses were estimated via UV-to-IR SED fitting using the LEPHARE code (S. Arnouts & O. Ilbert 2011), assuming stellar population models of G. Bruzual & S. Charlot (2003) with fixed (i.e. non-evolving) metallicities and either exponentially declining or delayed SFHs. Photometric fluxes were measured across ~ 30 bands using the FARMER pipeline (J. R. Weaver et al. 2023a).

(iv) The $0.2 < z < 3$ GSMF from J. Leja et al. (2020) based on 58 461 galaxies in the 3D-*HST* survey (covering an effective area of ~ 900 arcmin²) and 48 443 galaxies in the COSMOS field (an effective area of ≈ 1.38 deg²). The authors employed the PROSPECT SED-fitting code (J. Leja et al. 2017) to model rest-frame UV-to-IR galaxy photometry. Stellar masses were estimated using

⁵The choice of $z = 17$ is motivated by the fact that this is the highest integer redshift for which our lowest resolution simulation (m7) contains a sufficient number of subhaloes with non-zero stellar masses (≈ 60), so that the GSMF can still be estimated.

⁶<https://colibre-simulations.org/>

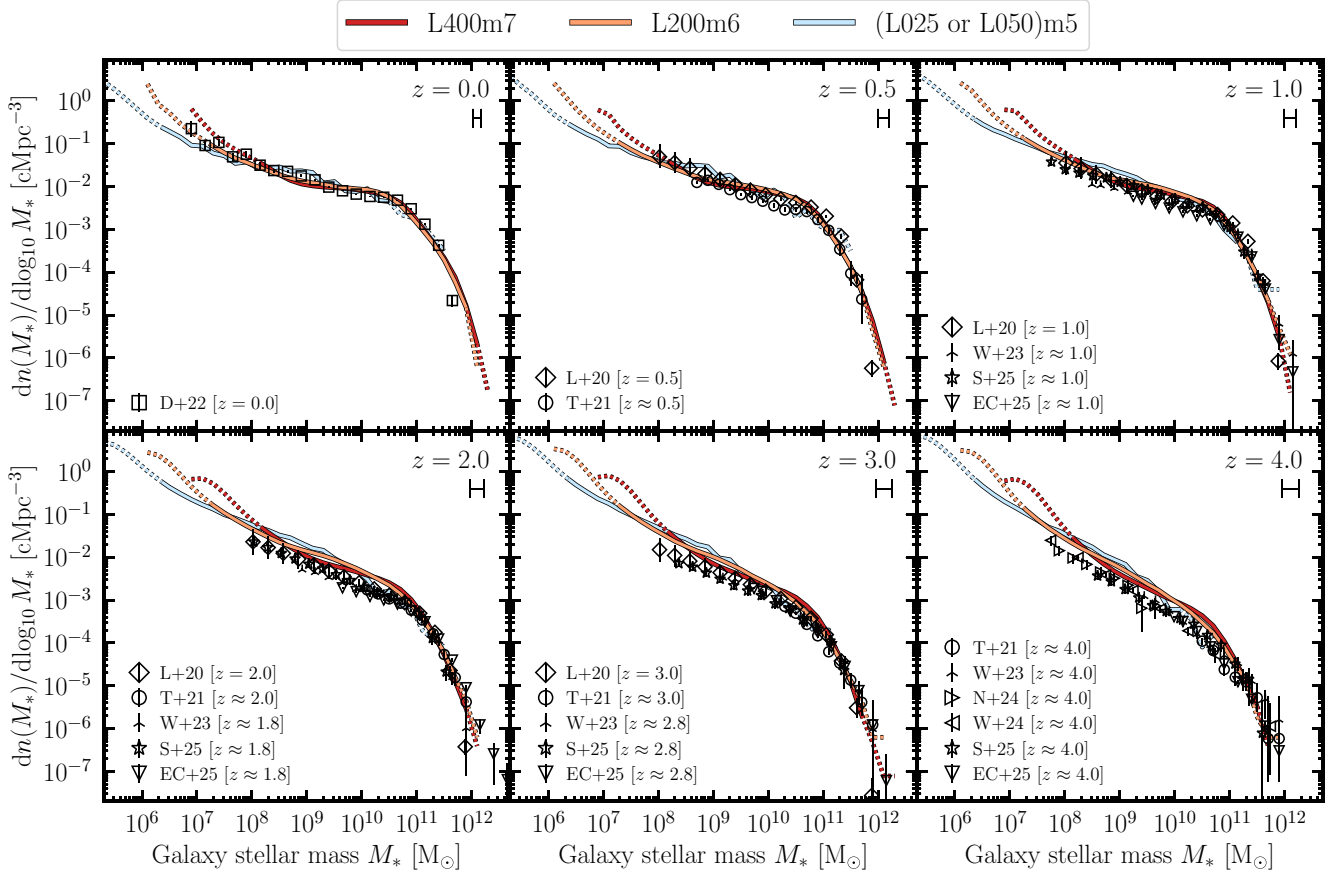


Figure 1. Evolution of the GSMF from $z = 0$ (top left) to $z = 4$ (bottom right) in the COLIBRE m7, m6, and m5 simulations using the largest cosmological volumes available (see Table 1), with solid dark-red, orange, and light-blue curves representing each resolution, respectively. To highlight resolution and volume limitations, the solid lines become dotted at stellar masses where galaxies are poorly resolved ($M_* < 10$ times the mean initial baryonic particle mass) or where the number of galaxies per 0.2 dex mass bin is less than 10. A log-normal scatter with a standard deviation given by equation (1) is added to galaxy stellar masses to account for Eddington bias. For reference, the (full extent of the) horizontal black error bar in each panel indicates the typical systematic uncertainty in stellar mass measurements, as given by equation (2). For comparison, we include a compilation of observational data from J. Leja et al. (2020), M. Stefanon et al. (2021), J. E. Thorne et al. (2021), S. P. Driver et al. (2022), J. R. Weaver et al. (2023b), A. Weibel et al. (2024), R. Navarro-Carrera et al. (2024), M. Shuntov et al. (2025), and Euclid Collaboration (2025b) (black symbols). COLIBRE is consistent with the selected observational data, given the systematic uncertainties on observationally inferred stellar masses, and achieves very good convergence with resolution at all explored redshifts.

the Flexible Stellar Population Synthesis model (C. Conroy et al. 2009), utilizing non-parametric SFHs and treating metallicity as a free parameter. All analysis was performed within a Bayesian framework.

(v) The $6 < z < 10$ GSMF from M. Stefanon et al. (2021) based on data from the GOODS, HUDF/XDF, and five CANDELS fields, covering a combined effective area of ≈ 731 arcmin². The authors constructed a sample of ~ 800 Lyman-break galaxies at $z \gtrsim 6$, including six galaxies at $z \sim 10$ and making use of deep IR data from *Spitzer*/IRAC at 3.6 and 4.5 μ m. Stellar masses were estimated using G. Bruzual & S. Charlot (2003) stellar population models, assuming a $0.2 Z_\odot$ metallicity and a constant SFH. SED fitting was performed using the FAST code (M. Kriek et al. 2009). Photometric redshifts were derived with EAZY (G. B. Brammer, P. G. van Dokkum & P. Coppi 2008).

(vi) The $4 < z < 9$ GSMF from A. Weibel et al. (2024) based on data from the *JWST*/NIRCam imaging programmes CEERS, PRIMER, and JADES, encompassing an effective area of ~ 500 arcmin². The authors compiled a sample of over 30 000 galaxies. Stellar masses and photometric redshifts were estimated using

SED fitting with the Bayesian framework BAGPIPES (A. C. Carnall et al. 2018). The authors adopted the BPASS-v2.2.1 stellar population models (E. R. Stanway & J. J. Eldridge 2018) with delayed- τ SFHs and treated metallicity as a free parameter.

(vii) The *JWST* GSMF from R. Navarro-Carrera et al. (2024) based on a sample of 3300 galaxies spanning $z \approx 3.5$ to 8.5. These galaxies were selected from *JWST* imaging in the HUDF (≈ 12 arcmin²) and the UKIDSS Ultra Deep Survey field (≈ 7 arcmin²), complemented by auxiliary *HST*/WFC3 and WFC3/ACS data in the same fields. Stellar masses were estimated using G. Bruzual & S. Charlot (2003) stellar population models with exponentially declining SFHs and metallicities of $Z = 0.2 Z_\odot$ and $1 Z_\odot$, as well as using C. Leitherer et al. (1999) models with constant SFHs and $Z = 0.05 Z_\odot$ and $0.4 Z_\odot$. Photometric redshifts were determined through SED fitting using LEPHARE.

(viii) The GSMF from the COSMOS-Web *JWST* photometric survey (M. Shuntov et al. 2025) based on imaging with four NIR-Cam filters. The data span a redshift range of $0.2 < z < 12$, containing 289 844 galaxies over an effective area of 0.53 deg², with ≈ 0.2 deg² also covered by one MIRI filter. Stellar masses were

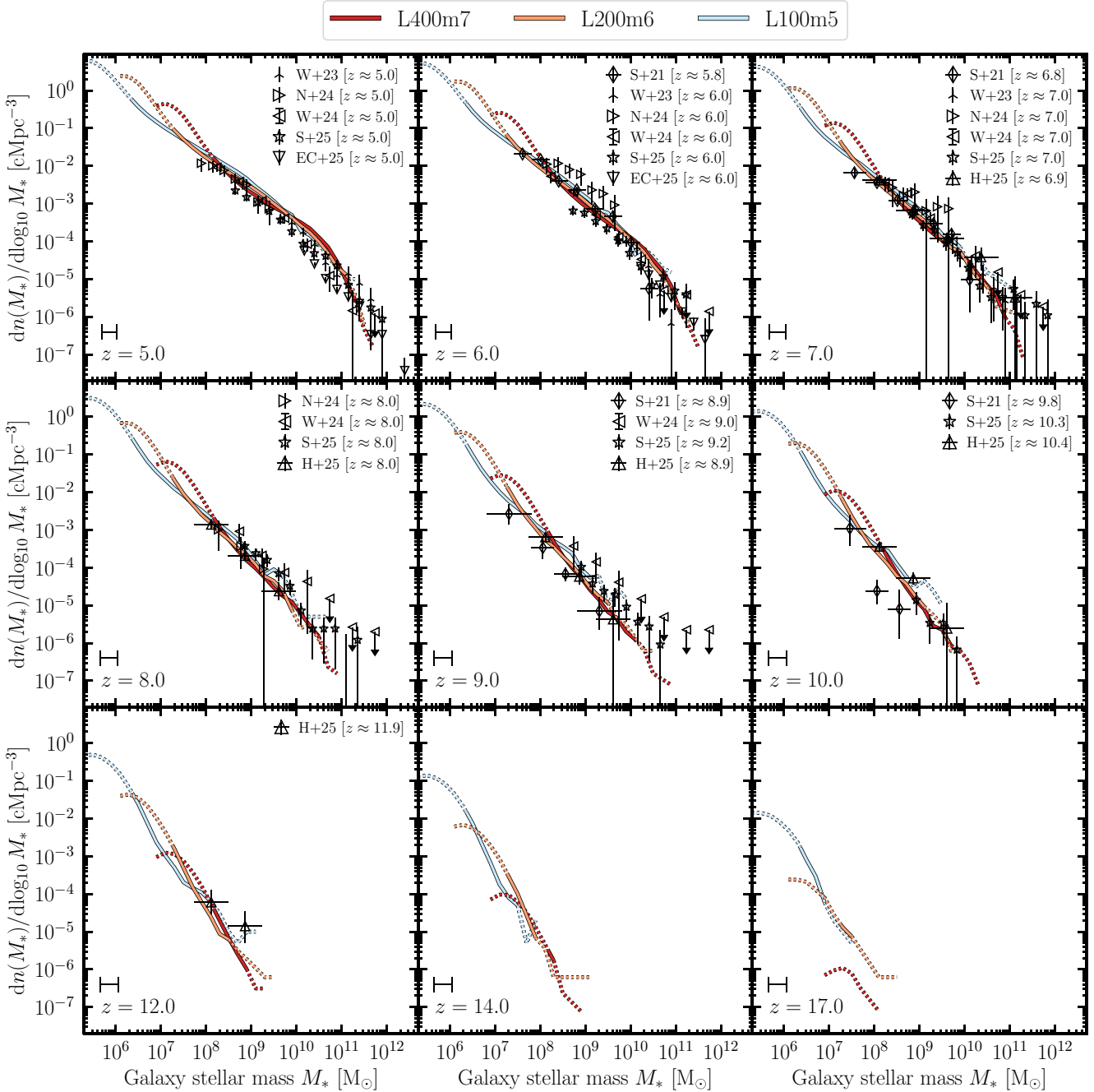


Figure 2. As Fig. 1, but for higher redshifts, ranging from $z = 5$ (top left) to $z = 17$ (bottom right). Black symbols represent observational data from M. Stefanon et al. (2021), J. R. Weaver et al. (2023b), A. Weibel et al. (2024), R. Navarro-Carrera et al. (2024), M. Shuntov et al. (2025), T. Harvey et al. (2025), and Euclid Collaboration (2025b). COLIBRE again shows very good convergence with resolution and is consistent with the observational data.

estimated using G. Bruzual & S. Charlot (2003) stellar population models with fixed metallicities, assuming delayed and exponential SFHs. SED fitting was performed using LEPHARE, applied to photometry in the UV-to-IR spectral range. The authors incorporated pre-*JWST* multiband imaging data in the COSMOS field alongside the *JWST* data.

(ix) The GSMF from T. Harvey et al. (2025) based on *JWST*/NIRCam imaging of 1120 galaxies at $6.5 < z < 13.5$ from multiple *JWST* fields, including CEERS, JADES, and PEARLS (covering a total area of ≈ 187 arcmin²). Stellar masses were

estimated using BAGPIPES, adopting the G. Bruzual & S. Charlot (2003) stellar population models with log-normal SFHs. The allowed metallicity values ranged from $0.005 Z_{\odot}$ to $5 Z_{\odot}$. Photometric redshifts were determined using EAZY.

(x) The GSMF from Euclid Collaboration (2025b) at $0.2 < z < 6.5$ derived using the Cosmic Dawn Survey pre-launch catalogues (Euclid Collaboration 2025c), with an effective area of ≈ 10.13 deg² and 2091 740 galaxies in the final sample after all selection criteria were applied. Multiwavelength data were provided by *Spitzer*/IRAC, the Hyper Suprime-Cam instrument

on the Subaru Telescope, and the MegaCam instrument on the Canada–France–Hawaii Telescope. Photometric redshifts were inferred using LEPHARE. Stellar masses were also derived with LEPHARE, using the same configuration and stellar templates as in COSMOS2020 (J. R. Weaver et al. 2023b).

We note that, except for the GSMF from J. Leja et al. (2020), none of the comparison data listed above have been corrected to remove the Eddington bias, which is included in our simulated GSMFs (as explained in Section 2.3).

It is evident from Figs 1 and 2 that the COLIBRE simulations are in good agreement with the observationally inferred GSMF at all redshifts where observational data are available. At $z = 0$, the excellent agreement of the COLIBRE m7 model is expected, as it was calibrated using emulators to match the GSMF from S. P. Driver et al. (2022). Although the m6 and m5 models were calibrated manually, they also show very good agreement, with deviations within ≈ 0.1 dex and no systematic offset from the observations.

For all COLIBRE resolutions, deviations from the data increase with redshift, peaking at $z \approx 2-4$, where the simulations predict a GSMF with number densities that are on average ≈ 0.3 dex higher than observed at $M_* \lesssim 10^{10.5} M_\odot$. The smallest discrepancies (around 0.2 dex) are found for the m7 and m6 models when compared to the measurements of J. Leja et al. (2020), whereas the largest deviations – up to ≈ 0.4 dex – occur in the m5 model relative to J. R. Weaver et al. (2023b) and M. Shuntov et al. (2025), as well as in the m7 and m6 models relative to Euclid Collaboration (2025b). While these minor discrepancies may point to limitations of the simulations, we also note that multiple studies have shown that, depending on the assumed SFH and SPS models, inferred stellar masses at $z \approx 2-4$ may be systematically offset by 0.1–0.3 dex relative to the true values (e.g. A. Katsianis et al. 2020; M. Shuntov et al. 2025). With this in mind, reducing the stellar masses in the simulations by 0.2 dex at $M_* \lesssim 10^{10.5} M_\odot$ would bring the simulations into much closer agreement with the data, within ≈ 0.2 dex.

Above $z = 4$, the differences between COLIBRE and observations become slightly smaller, on average falling within ≈ 0.25 dex for all resolutions and remaining at that level up to $z = 10$. Notably, from $z = 5$ to 10, COLIBRE is consistent with (albeit on average somewhat on the lower side of) the GSMF reported by M. Shuntov et al. (2025), which is based on the largest contiguous *JWST* imaging survey to date ($\approx 0.53 \text{ deg}^2$; C. M. Casey et al. 2023). Furthermore, COLIBRE reproduces the $7 < z < 10$ GSMF from T. Harvey et al. (2025), also based on *JWST* data, as well as the $6 < z < 9$ GSMF from M. Stefanon et al. (2021), which relies on pre-*JWST* observations. At the highest redshift probed by M. Stefanon et al. (2021), $z \approx 10$, COLIBRE overestimates their GSMF by up to ≈ 1 dex at $M_* \gtrsim 10^8 M_\odot$, supporting the scenario of a higher abundance of massive galaxies at early times, in line with *JWST* observations.

One notable exception to the good agreement between COLIBRE and observations is the discrepancy with the *JWST* measurements from R. Navarro-Carrera et al. (2024) at $z = 6$. At this redshift, the GSMFs from both M. Shuntov et al. (2025) and COLIBRE undershoot the GSMF reported by R. Navarro-Carrera et al. (2024) by ≈ 0.5 dex. M. Shuntov et al. (2025) suggest three possible explanations for the discrepancy in normalization: cosmic variance (the effective area of the *JWST* observations by R. Navarro-Carrera et al. 2024 is only $\approx 20 \text{ arcmin}^2$), photometric uncertainties, and/or systematic effects in the SED fitting.

Furthermore, at $7 < z < 9$, the COLIBRE GSMF falls more than 0.5 dex below the measurements of A. Weibel et al. (2024) for $M_* \gtrsim 10^{10} M_\odot$. However, A. Weibel et al. (2024) caution that the redshifts and stellar masses of the objects contributing to the GSMF at these high M_* and z are poorly constrained, and that the resulting GSMF values – nearly all of which are upper limits – may be significantly overestimated.

At the highest redshifts shown in Fig. 2, from $z = 12$ to 17, COLIBRE predicts a gradual evolution of the GSMF without sharp drop-offs, with the normalization changing by ≈ 2 dex over this redshift interval. At $z = 12$, we compare to the only available observational dataset, T. Harvey et al. (2025), who provide GSMF estimates based on *JWST*/NIRCam imaging at a median redshift of $z \approx 11.94$ and in two stellar mass bins: $M_* \sim 10^{8.125}$ and $10^{8.875} M_\odot$, each containing 5 objects (see their table 3). We find that COLIBRE is consistent with the lower mass bin within 1σ at all three resolutions but underpredicts the value in the higher mass bin by up to $\approx 2\sigma$, depending on the resolution. Note, however, that we are not accounting for potential systematic uncertainties. A shift in M_* by 0.3 dex, which is a relatively conservative estimate for the magnitude of the systematic error on stellar mass at $z = 12$, would bring the COLIBRE predictions at all three resolutions into agreement with the second bin within 1σ . Alternatively, increasing the assumed stellar mass uncertainty for simulated galaxies from our $z \geq 2$ default (and likely conservative for $z = 12$) value of 0.3 dex to 0.5 dex would also resolve the discrepancy (see Fig. D2 in Appendix D).

Finally, we note that at $z = 17$, the combination of different volumes and resolutions of COLIBRE enables us to predict a GSMF that spans more than two orders of magnitude in stellar mass: $10^{5.5} < M_*/M_\odot < 10^8$.

3.2 Stellar-to-halo mass relation

Having discussed the evolution of the COLIBRE GSMF, we turn to the SHMR, which is displayed in Fig. 3. Specifically, we show the evolution of SHMRs for central subhaloes at m7 (left panel), m6 (middle panel), and m5 (right panel) resolutions. The solid curves represent the median stellar-to-halo mass ratio for each 0.2-dex halo mass bin and the shaded regions mark the 16th to 84th percentile scatter. The colours indicate different redshifts, smoothly transitioning from yellow at $z = 11$ to dark blue at $z = 0.1$. The solid curves turn to dotted at the low- and high-mass ends, corresponding to haloes hosting galaxies with stellar masses $M_* < 10$ times the mean initial baryonic particle mass (i.e. containing fewer than ~ 10 stellar particles) and halo mass bins containing fewer than 10 objects, respectively.

For comparison, we take SHMRs from the semi-empirical models UNIVERSEMACHINE (P. Behroozi et al. 2019) and EMERGE (B. P. Moster et al. 2018) for central galaxies at $z = 0.1$. Both P. Behroozi et al. (2019) and B. P. Moster et al. (2018) used the peak halo mass defined within a sphere whose mean density is Δ times greater than the critical density of the Universe, with Δ being the redshift-dependent value from G. L. Bryan & M. L. Norman (1998) derived by fitting the results of V. R. Eke, S. Cole & C. S. Frenk (1996). In our analysis, we adopt the same G. L. Bryan & M. L. Norman (1998) halo mass definition, M_{halo} , but use the actual halo masses at each redshift instead of their peak values. As we study the SHMR of central subhaloes, this distinction should introduce negligible differences to the SHMR.

We find that at all redshifts, the COLIBRE SHMRs are well converged among the three resolutions for halo masses of $M_{\text{halo}} \gtrsim$

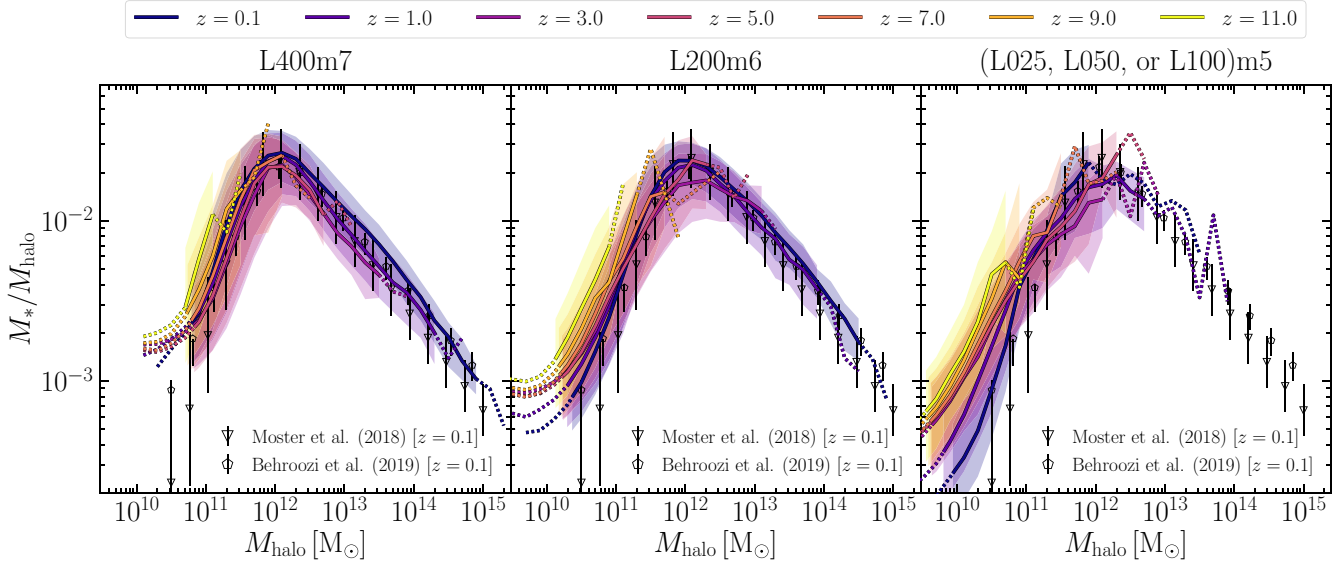


Figure 3. Evolution of the SHMR for central subhaloes in the COLIBRE simulations at m7 (left), m6 (middle), and m5 (right) resolutions. Solid curves represent the median stellar-to-halo mass ratios from the simulations, while the shaded regions show the 16th to 84th percentile scatter. Different colours indicate different redshifts, ranging from $z = 0.1$ (dark blue) to $z = 11$ (yellow). The solid lines switch to dotted at halo masses corresponding to galaxies with $M_* < 10$ times the mean initial baryonic particle mass, or where fewer than 10 objects fall within a given halo mass bin. For comparison, the semi-empirical SHMRs from B. P. Moster, T. Naab & S. D. M. White (2018) and P. Behroozi et al. (2019) at $z = 0.1$ are shown by black symbols. Here we do not add a log-normal scatter to galaxy stellar masses from the simulations because the Eddington bias is absent in the comparison data. The COLIBRE SHMR is consistent with the semi-empirical comparison data at $M_{\text{halo}} \gtrsim 10^{11} M_{\odot}$, where it shows excellent convergence with resolution, peaks at a halo mass $M_{\text{halo}} \sim 10^{12} M_{\odot}$, and exhibits only weak evolution with redshift.

$10^{11} M_{\odot}$. At lower M_{halo} , convergence degrades due to resolution limitations of the lower resolution simulations, with the SHMRs predicted by the m7 model flattening and approaching $\sim 10^{-3}$ from above as M_{halo} decreases to $\sim 10^{10} M_{\odot}$. This ‘floor’ of the median relation is expected, as the lowest non-zero stellar mass a subhalo can have at m7 resolution is comparable to the mass of a single stellar particle ($\sim 10^7 M_{\odot}$), which, at $M_{\text{halo}} \sim 10^{10} M_{\odot}$, results in a stellar-to-halo mass ratio of $\sim 10^{-3}$. At even lower halo masses ($M_{\text{halo}} \lesssim 10^9 M_{\odot}$), the median stellar-to-halo mass ratio drops to zero as the majority of these low-mass subhaloes have not formed any stellar particles (not shown).

At $z = 0.1$, COLIBRE agrees with the SHMRs reported by B. P. Moster et al. (2018) and P. Behroozi et al. (2019) for halo masses $M_{\text{halo}} \gtrsim 10^{11.5} M_{\odot}$, matching both the shape and normalization of the SHMR. For $M_{\text{halo}} \lesssim 10^{11.5} M_{\odot}$, the simulations remain in good agreement with P. Behroozi et al. (2019) but are only broadly consistent with B. P. Moster et al. (2018), which predict a slightly steeper decline of the SHMR at lower halo masses. We note, however, that both B. P. Moster et al. (2018) and P. Behroozi et al. (2019) derive their halo masses from DMO simulations, whereas COLIBRE includes baryonic effects throughout. For haloes in the mass range $10^{11.5} \lesssim M_{\text{halo}} \lesssim 10^{13} M_{\odot}$ ($10^{10} \lesssim M_{\text{halo}} \lesssim 10^{11.5} M_{\odot}$), halo masses may increase by ≈ 15 (30) per cent when measured from a DMO simulation (e.g. M. Schaller et al. 2015), thereby likely bringing COLIBRE into closer agreement with B. P. Moster et al. (2018) for $M_{\text{halo}} \lesssim 10^{11.5} M_{\odot}$. Another difference between COLIBRE and the two semi-empirical models lies in the choice of the halo finder. While COLIBRE employs HBT-HERONS, both B. P. Moster et al. (2018) and P. Behroozi et al. (2019) use ROCKSTAR (P. S. Behroozi, R. H. Wechsler & H.-Y. Wu 2013). V. J. Forouhar Moreno et al. (2025) found that differences in halo-finding

algorithms can introduce variations in the number density of central subhaloes of up to ≈ 10 per cent, which may, in turn, impact the SHMR.

The COLIBRE simulations predict only weak evolution.⁷ of the SHMR with redshift, with the normalization varying by only ≈ 0.5 dex between $z = 0.1$ and 11, while the overall shape remains largely unchanged. On average, the normalization increases weakly with redshift. At all resolutions and times, the peak of the SHMR never exceeds a value of 0.04. The most notable redshift-dependent variation is a monotonic shift of the high-mass end of the SHMR towards higher M_{halo} at lower redshifts, reflecting the fact that more massive haloes on average require more time to assemble. Importantly, the COLIBRE models at different resolutions show good convergence not only at $z = 0.1$ but also at all higher redshifts.

We note that the peak of the $z = 0.1$ SHMR in COLIBRE is $M_*/M_{\text{halo}} \approx 0.025$. Dividing this value by $\Omega_{b,0}/\Omega_{m,0}$ yields a baryon conversion efficiency into stars of ≈ 16 per cent. The logarithmic slope of the COLIBRE SHMR before (after) the peak is approximately 1 (−0.5). These values are comparable to those used by C. T. Donnan et al. (2025), who presented a simple theoretical model that reproduces both the high- z GSMF and high- z UV luminosity functions (UVLFs) measured by *JWST*. Their model combines an HMF with a redshift-independent galaxy formation efficiency (i.e. a non-evolving SHMR) and a fixed IMF, and assumes a mass-to-light ratio that evolves with redshift driven by

⁷We verified that the COLIBRE SHMR remains weakly dependent on redshift also when defining halo mass as M_{200c} – the mass enclosed within a sphere where the average density is 200 times the critical density of the Universe.

the increasingly younger stellar population at higher redshifts. The fact that the COLIBRE simulations also predict a (nearly) redshift-independent galaxy formation efficiency and are consistent with the high- z GSMF based on *JWST* data further supports the findings of the authors. At the same time, COLIBRE appears to be in tension with M. Shuntov et al. (2025), who argue that galaxy formation efficiencies as high as ≈ 0.5 may be necessary to match the observed GSMF in the $z \gtrsim 6$ Universe. As shown in Fig. 2, COLIBRE is broadly consistent with the GSMF from M. Shuntov et al. (2025), yet it predicts a nearly constant galaxy formation efficiency of ≈ 16 per cent at the peak of SHMR.

3.3 Cosmic stellar mass density

The top panel of Fig. 4 illustrates the evolution of the cosmic stellar mass density (CSMD) as a function of redshift (bottom x-axis) and cosmic time (top x-axis). The COLIBRE m7, m6, and m5 simulations are shown in dark-red, orange, and light-blue colours, respectively. The solid lines indicate the CSMD computed using the mass of all stellar particles that are present in the simulated volume. The dash-dotted lines represent the CSMD derived from only subhaloes with stellar masses equal to or exceeding $10^8 M_\odot$. This selection mimics a common observational method for estimating the CSMD by integrating a (double) P. Schechter (1976) function down to a stellar mass limit of $\sim 10^8 M_\odot$. The subhalo stellar masses are calculated by summing the masses of stellar particles that are gravitationally bound and located within 50 pkpc 3D apertures.

For clarity, the solid and dash-dotted lines for the m5 and m7 simulations are truncated below $z = 4$, where they overlap with the CSMD from the m6 simulation. The bottom panel shows the relative difference in CSMD between the m7 and m5 simulations compared to the m6 simulation. The relative difference for a simulation at resolution mX, where X represents 5, 6, or 7, is calculated as $\Delta \text{CSMD}_{\text{mX}} \equiv \text{CSMD}_{\text{mX}} / \text{CSMD}_{\text{m6}} - 1$, where CSMD_{mX} is the CSMD from the mX simulation. We compute these differences for two cases: using all stellar mass in the simulated volume (solid lines) and applying the $M_* \geq 10^8 M_\odot$ selection criterion (dash-dotted lines). Because the L100m5 and L50m5 simulations have not yet progressed to $z = 0$ at the time of writing (they are currently at $z = 4.6$ and $z = 0.9$, respectively; see Table 1), we show the CSMD at m5 resolution from the 100^3 cMpc^3 volume for $z > 4.6$, from 50^3 cMpc^3 between $z = 0.9$ and $z = 4.6$, and from the 25^3 cMpc^3 volume at lower redshifts. In Appendix A, we demonstrate that at $z < 4$, the CSMDs in these three volumes converge very well, while at higher redshifts the 100^3 cMpc^3 volume yields a slightly higher CSMD.

For comparison, we include CSMD estimates from the same ten studies shown in Figs 1 and 2, as well as additional measurements from S. P. Driver et al. (2018). Specifically, we show the $z = 0$ CSMD value from S. P. Driver et al. (2022), obtained by integrating the GAMA DR4 GSMF down to $M_* = 10^{6.75} M_\odot$. This estimate accounts for re-normalization of the GAMA fields to a larger region selected from the SDSS DR16, along with a correction for CSMD evolution between the GAMA sample's median observed redshift ($z \approx 0.079$) and $z = 0$. We also include CSMD estimates from J. Leja et al. (2020), M. Stefanon et al. (2021), J. E. Thorne et al. (2021), J. R. Weaver et al. (2023b), A. Weibel et al. (2024), M. Shuntov et al. (2025), R. Navarro-Carrera et al. (2024), T. Harvey et al. (2025), and Euclid Collaboration (2025b), all derived by integrating their respective best-fitting Schechter functions down to $M_* = 10^8 M_\odot$. Finally, we include the $0 < z < 5$ CSMD estimates

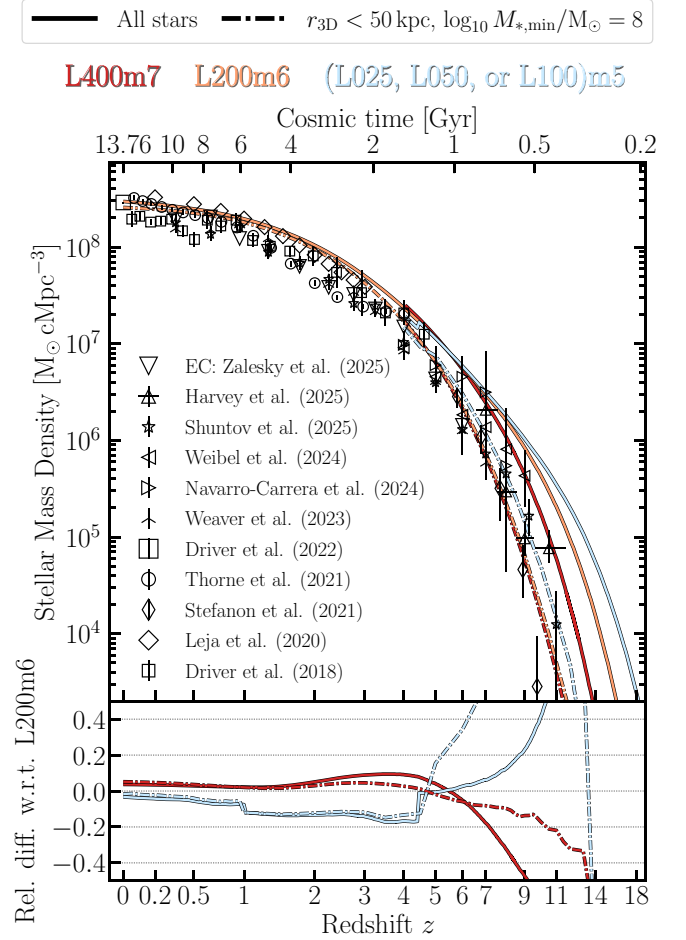


Figure 4. *Top panel:* Evolution of the cosmic stellar mass density (CSMD) in the COLIBRE m7, m6, and m5 simulations. Different resolutions are indicated by different colours. We show the total CSMD, including all stellar mass within the simulated volumes (solid lines), and the CSMD contributed only by galaxies with stellar masses $M_* \geq 10^8 M_\odot$ (dash-dotted lines), where the stellar mass is measured within 3D spherical apertures of radius 50 pkpc. While the total CSMD reflects the ‘raw’ simulation values, the CSMD shown by the dash-dotted lines incorporates common observational cuts and can be directly compared to observations. For clarity, the solid and dash-dotted lines for the m7 and m5 simulations are truncated at $z < 4$ due to their overlap with the m6 lines. For comparison, a compilation of recent observational measurements is shown as black symbols. No log-normal scatter is added to the simulated galaxy stellar masses, as the comparison data account for Eddington bias when estimating the CSMD. *Bottom panel:* Relative difference in the CSMD between simulations at different resolutions, shown for both versions of CSMD and using the m6 simulation as the baseline. The CSMD in COLIBRE agrees well with the observational data at all redshifts and shows very good convergence between resolutions for $z < 6$. Selecting only galaxies with $M_* \geq 10^8 M_\odot$ reduces the CSMD by more than 0.5 dex at $z \gtrsim 6$, while having only a minor impact below $z = 2$.

from S. P. Driver et al. (2018), based on a combined dataset of $\sim 10^6$ galaxies from GAMA DR2, G10-COSMOS, and the 3D-*HST* survey. In their analysis, the authors fit a spline to their GSMF measurements and integrated it down to $M_* \sim 10^7 M_\odot$, relying increasingly on extrapolation at higher redshifts.

We first compare the solid and dash-dotted lines of the same colour in the top panel, finding that applying a stellar mass threshold of $10^8 M_\odot$ to subhaloes contributing to the CSMD has

a noticeable impact at $z \gtrsim 2$. This threshold becomes increasingly significant at higher redshift, reducing the CSMD by about 0.5 and 1 dex at $z \sim 6$ and $z \sim 9$, respectively. These results are consistent with the findings of M. Furlong et al. (2015), who performed a similar analysis using the EAGLE simulations. Furthermore, while the $M_* = 10^8 M_\odot$ threshold accounts for most of the difference between the solid and dash-dotted curves, at $z \lesssim 2$ there is also a minor contribution from the aperture requirement⁸. At $z = 0$, COLIBRE predicts a total CSMD of $\approx 3 \times 10^8 M_\odot \text{ cMpc}^{-3}$, with ≈ 90 per cent bound to subhaloes of $M_* \geq 10^8 M_\odot$ and within 50 pkpc of their centres.

Next, by comparing the observational data with the dash-dotted lines, which best mimic the observational selection criteria, we find that the CSMD in COLIBRE is generally in good agreement with the observed CSMD across all redshifts. Specifically, the COLIBRE CSMD at $z = 0$ best matches the measurements from S. P. Driver et al. (2022); from $z = 0$ to $z = 1$, it is in good agreement with J. E. Thorne et al. (2021); between $z = 1$ and 4, COLIBRE remains consistent with the CSMD from S. P. Driver et al. (2018) and J. Leja et al. (2020); and at $4 < z < 12$, it agrees reasonably well with the CSMD reported by M. Shuntov et al. (2025). At the same time, COLIBRE slightly overestimates the CSMD reported by J. R. Weaver et al. (2023b), M. Shuntov et al. (2025), and Euclid Collaboration (2025b) at $1.5 < z < 4$, and underestimates the measurements from R. Navarro-Carrera et al. (2024) at $z > 6$ and from A. Weibel et al. (2024) at $z > 7$. These discrepancies are consistent with those seen in the GSMF (see Figs 1 and 2, and Section 3.1 for discussion). The COLIBRE CSMD is consistent with the observations from M. Stefanon et al. (2021) and T. Harvey et al. (2025) over the redshift range $6 < z < 9$, but at $z \approx 10$ it falls ≈ 0.5 dex below the measurement from T. Harvey et al. (2025), while exceeding that of M. Stefanon et al. (2021) by ≈ 1 dex.

Finally, in the lower panel, we observe that the relative differences in the CSMD between the m5 and m6 models, and between the m6 and m7 models, remain within ≈ 20 per cent at $0 < z < 5$, increase to within ≈ 40 per cent at $5 < z < 6$, and rise to beyond 40 per cent at higher redshifts. As we clarify in Appendix A, the sharp rise in the relative differences at $z > 7$ is caused by resolution effects, while the increase in the difference between the m6 and m5 resolutions at $5 < z < 7$ is due to the switch from the L050m5 simulation to L100m5 at $z = 5$, with the larger volume containing a massive protocluster at $z \sim 5-6$ that is atypical for a $(100 \text{ cMpc})^3$ volume and hence yields a higher CSMD compared to L200m6.

3.4 Cosmic star formation rate density

The top panel of Fig. 5 shows the evolution of the cosmic star formation rate density (CSFRD). The solid curves show the SFR from all star-forming gas particles in the simulation, while the dash-dotted curves indicate the SFR from gas particles that are bound to and located within 50 pkpc of subhaloes containing at least $10^7 M_\odot$ of stellar mass. This mass threshold is chosen to facilitate a more consistent comparison with *JWST* observations at high redshift. As in Fig. 4, we highlight the m6 simulation, while the CSFRD curves for the m7 and m5 simulations are truncated below $z = 4$ for visual clarity. The bottom panel

⁸The impact of this and other restrictions is discussed in detail in Appendix F.

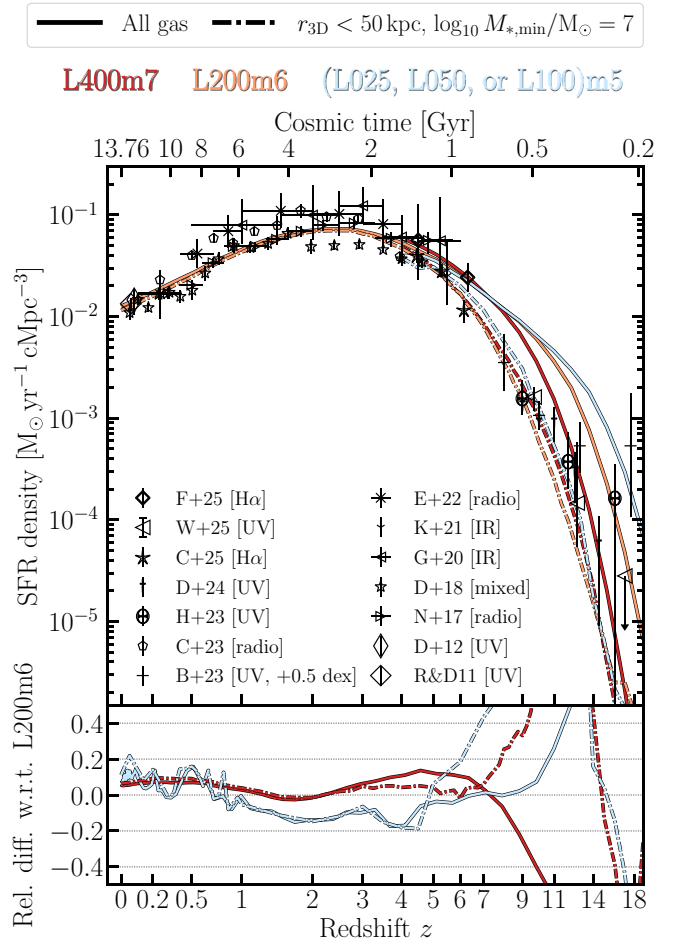


Figure 5. Evolution of the cosmic star formation rate density (CSFRD). The figure follows the same layout as Fig. 4. We show the total CSFRD, contributed by all star-forming gas particles within the simulated volumes (solid lines), and the CSFRD from galaxies with stellar masses $M_* \geq 10^7 M_\odot$ (dash-dotted lines). Black symbols represent a compilation of observational data from A. S. G. Robotham & S. P. Driver (2011), S. P. Driver et al. (2012), M. Novak et al. (2017), S. P. Driver et al. (2018), C. Gruppioni et al. (2020), Y. Khusanova et al. (2021), A. Enia et al. (2022), R. Bouwens et al. (2023), R. K. Cochrane et al. (2023), Y. Harikane et al. (2023), C. T. Donnan et al. (2024), A. Covelo-Paz et al. (2025), S. Fu et al. (2025), and A. Weibel et al. (2025a). The CSFRD in COLIBRE exhibits a plausible shape and normalization, showing good agreement with the data at all redshifts and very good numerical convergence at $z < 6$.

displays the differences in CSFRD relative to the m6 simulation. As in Fig. 4, we show the predictions of the COLIBRE m5 model from the 100^3 cMpc^3 volume for $z > 4.6$, from the 50^3 cMpc^3 volume for $0.9 < z < 4.6$, and from the 25^3 cMpc^3 volume at lower redshifts.

For comparison, we include a compilation of observational estimates of the CSFRD based on various SFR indicators. At $z \approx 0$, we show FUV-based estimates from S. P. Driver et al. (2012), derived from the GAMA DR1 data, and from A. S. G. Robotham & S. P. Driver (2011), which are based on a matched catalogue between the *GALEX* Medium Imaging Survey and SDSS DR7 data. At higher redshifts, we present SED-based measurements from S. P. Driver et al. (2018), derived from a combined dataset of the GAMA, G10-COSMOS, and 3D-*HST* surveys, yielding a total of 582 314 galaxies over $0 < z < 5$. We complement our

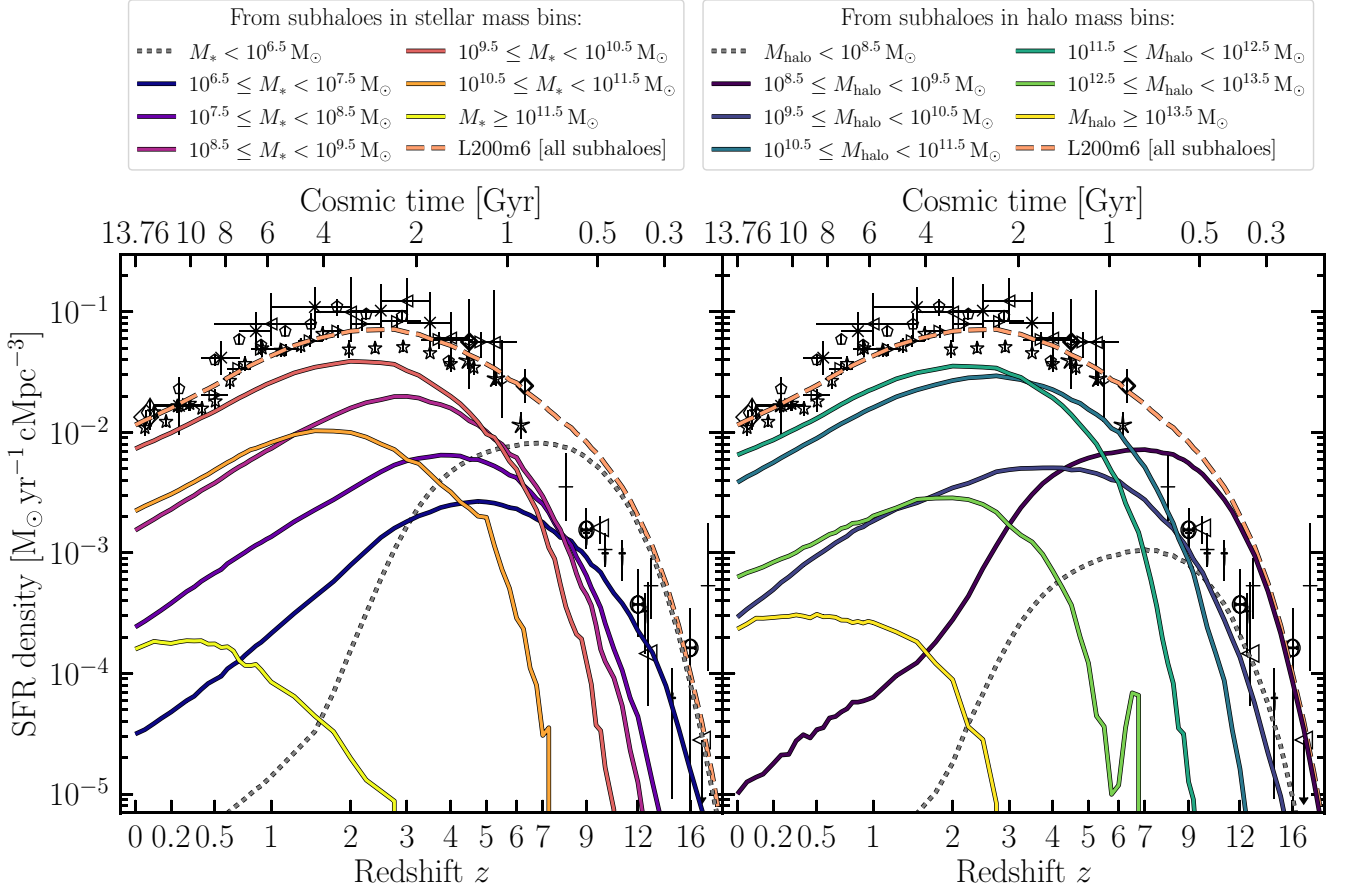


Figure 6. Cosmic star formation rate density (CSFRD) as a function of redshift in the COLIBRE L200m6 simulation, split into 1-dex stellar mass bins (*left panel*) and 1-dex halo mass bins (*right panel*). Differently coloured solid lines represent the contributions from each bin, except for the lowest mass bin, which is shown as a grey dotted line to highlight resolution limitations. The lower (upper) edge of the lowest (highest) stellar and halo mass bins is unbounded to include all remaining objects in the simulation at the low-mass (high-mass) end. The halo mass is the G. L. Bryan & M. L. Norman (1998) mass for central subhaloes and the total gravitationally bound mass for satellites. Subhalo SFRs are obtained by summing the instantaneous SFRs of gas particles that are bound to the subhalo and within a 50 kpc 3D spherical aperture. The thick dashed orange line represents the CSFRD calculated by summing the SFRs of all subhaloes within the simulation volume. To guide the eye, the black symbols show the same observational measurements as in Fig. 5. At $0 < z < 3$, subhaloes with masses of $M_{\text{halo}} \sim 10^{12} M_{\odot}$, which host galaxies with stellar masses of $\sim 10^{10} M_{\odot}$, dominate the CSFRD. At higher redshifts ($z > 3$), the primary contribution to the CSFRD shifts progressively towards subhaloes of lower mass, while the most massive objects ($M_{\text{halo}} \geq 10^{13.5} M_{\odot}$) contribute negligibly at all z .

comparison dataset with three radio continuum-based CSFRDs: measurements from the VLA COSMOS survey, derived from 3-GHz observations of 5 915 objects over $0.3 < z < 5$ (M. Novak et al. 2017); VLA observations of 554 radio-selected galaxies in the GOODS-N field, covering $0.1 < z < 3$ (A. Enia et al. 2022); and 150-MHz measurements from the LOFAR Two Metre Sky Survey, which include 55 581 radio sources spanning $0.1 < z \lesssim 4$ (R. K. Cochrane et al. 2023). At $z > 0.5$, we show CSFRD estimates from C. Gruppioni et al. (2020), based on 56 sources within the ALPINE multiwavelength survey (ECDFS and COSMOS fields) detected in sub-mm continuum by ALMA, as well as two high-redshift ($z \approx 4.5$ and ≈ 5.5) values of the CSFRD also measured by ALMA (Y. Khusanova et al. 2021). We further include H α -based measurements at $4 < z < 7$ from A. Covelo-Paz et al. (2025) and S. Fu et al. (2025). Both studies utilized *JWST*/NIRCam grism observations and derived CSFRD values by integrating their SFR functions down to $\approx 0.25 M_{\odot} \text{ yr}^{-1}$, corresponding to an absolute UV magnitude limit of $M_{\text{UV}} = -17$. Finally, at very high redshifts

($8 < z < 18$), we show the CSFRD based on *JWST*-measured UVLF reported by R. Bouwens et al. (2023), Y. Harikane et al. (2023), C. T. Donnan et al. (2024), and A. Weibel et al. (2025a). While the latter three studies estimated the CSFRD by integrating their measured UVLF down to -17 mag, R. Bouwens et al. (2023) adopted a fainter limit of -19 mag. To account for this difference, we shift the CSFRD values reported in R. Bouwens et al. (2023) by 0.5 dex upwards.

The COLIBRE simulations exhibit good numerical convergence for both particle selection criteria. At $0 < z < 6$, the relative differences between the simulations remain within ≈ 20 per cent. At $z > 6$, the convergence of the CSFRD from $M \geq 10^7 M_{\odot}$ galaxies becomes worse, driven by differences in the simulated volumes used at each resolution (see Appendix A). At even higher redshifts ($z \gtrsim 10$), the predicted total CSFRD begins to strongly diverge, with higher resolution simulations yielding systematically higher SFRs than their lower resolution counterparts. As clarified later in Fig. 6, this behaviour is driven by resolution effects. At $z \gtrsim$

10, the CSFRD is dominated by SFRs from low-mass subhaloes ($M_{\text{halo}} \lesssim 10^{9.5} M_{\odot}$). Higher resolution simulations resolve more of these low-mass objects and begin to track them at earlier times, which together lead to a higher CSFRD.

Turning to the comparison with observational data, at $z \approx 0$, COLIBRE predicts a CSFRD of $\approx 0.012\text{--}0.013 M_{\odot} \text{yr}^{-1} \text{cMpc}^{-3}$, which is consistent with the $z \sim 0$ measurements from S. P. Driver et al. (2012), A. S. G. Robotham & S. P. Driver (2011), and S. P. Driver et al. (2018). From $z = 0$ to 2, the CSFRD in the simulations increases by about 0.7 dex, closely following the measurements from M. Novak et al. (2017) and S. P. Driver et al. (2018). The simulated CSFRD peaks at $\approx 0.075 M_{\odot} \text{yr}^{-1} \text{cMpc}^{-3}$ between $z = 2$ and 3, in reasonable agreement with M. Novak et al. (2017) and S. P. Driver et al. (2018). However, it falls short by ≈ 0.3 dex compared to the measurements of C. Gruppioni et al. (2020), A. Enia et al. (2022), and R. K. Cochrane et al. (2023). The factor of ≈ 2 discrepancy at $z \approx 2$ between the CSFRD computed from intrinsic SFRs in simulations and that derived from direct SFR indicators in observations, with the former being lower, has been seen in nearly all modern galaxy formation simulations in representative volumes that predict a realistic $z = 0$ galaxy population, including EAGLE, ILLUSTRISTNG, and SIMBA (M. Furlong et al. 2015; A. Pillepich et al. 2018; R. Davé et al. 2019). Furthermore, the same discrepancy has long been seen in observational data themselves: the SFH derived from the observationally inferred CSMD is typically a factor of ≈ 2 lower around cosmic noon compared to the SFH constructed from observationally inferred SFRs (e.g. S. M. Wilkins, N. Trentham & A. M. Hopkins 2008; H. Yu & F. Y. Wang 2016; S. M. Wilkins, C. C. Lovell & E. R. Stanway 2019; M. Shuntov et al. 2025). The ubiquity of this discrepancy was clearly demonstrated by P. Behroozi et al. (2019), who calibrated the semi-empirical model UNIVERSEMACHINE by fitting it to the observed CSFRD and CSMD, among other constraints. Although by design the UNIVERSEMACHINE is highly flexible, the authors still had to introduce a redshift-dependent correction factor to the model's SFRs to satisfy the CSMD and CSFRD constraints simultaneously, which increased the 'observed' SFRs relative to the true SFRs predicted by the UNIVERSEMACHINE by ≈ 0.3 dex at $z = 2$. The source of this discrepancy remains unclear, but it is often attributed to systematic offsets in observational data, arising from uncertainties in the calibration of SFR indicators and assumptions regarding the dust attenuation curve, the stellar IMF, and/or gas metallicity.

Moving to even higher redshifts, we compare the COLIBRE CSFRD from subhaloes with $M_* \geq 10^7 M_{\odot}$ (dash-dotted curves) to H α - and UV-based *JWST* measurements at $z > 4$ from R. Bouwens et al. (2023), Y. Harikane et al. (2023), C. T. Donnan et al. (2024), A. Covelo-Paz et al. (2025), S. Fu et al. (2025), and A. Weibel et al. (2025a). At $4 < z < 8$, a stellar mass⁹ of $M_* \sim 10^7\text{--}10^8 M_{\odot}$ roughly corresponds to a rest-frame UV magnitude of $M_{\text{UV}} = -17$ (e.g. M. Song et al. 2016; M. Stefanon et al. 2021), to which all considered *JWST* data are normalized. At higher redshifts ($z > 8$), the $M_{\text{UV}} - M_*$ relation becomes increasingly uncertain; however, both theoretical models

⁹While in Fig. 5 we show the effect of applying the $M_* \geq 10^7 M_{\odot}$ cut, in Appendix F we additionally demonstrate the impact of a stricter $M_* \geq 10^8 M_{\odot}$ threshold. Compared to $M_* \geq 10^7 M_{\odot}$, adopting $M_* \geq 10^8 M_{\odot}$ further reduces the CSFRD predicted by the simulations by ≈ 0.03 dex at $z = 4$, increasing to about 0.15 dex at $z = 8$.

and observations suggest that galaxies with $M_* \sim 10^7 M_{\odot}$ should correspond to $-18 < M_{\text{UV}} < -16$ (e.g. P. Santini et al. 2023; R. Feldmann et al. 2025). COLIBRE agrees with the five said datasets within $\approx 1.5\sigma$ over $4 < z < 12$, while at $12 < z < 18$ it undershoots R. Bouwens et al. (2023) but remains roughly within the error bars – albeit on the lower side – of the measurements from C. T. Donnan et al. (2024), Y. Harikane et al. (2023), and A. Weibel et al. (2025a). We caution, however, that at these highest redshifts, the observationally inferred CSFRD relies heavily on extrapolation down to $M_{\text{UV}} = -17$, and that the CSFRD in the simulations is almost entirely dominated by subhaloes that have formed at most only a handful of stellar particles, indicating resolution limitations of the COLIBRE simulations.

Finally, we note that although the CSFRD from $M_* \geq 10^7 M_{\odot}$ subhaloes predicted by the m5 model becomes negligible above $z = 14$ due to the lack of $M_* \geq 10^7 M_{\odot}$ objects in the 100^3cMpc^3 volume at these very high redshifts, its total CSFRD (indicated by the solid light-blue curve) at $z = 14$ remains at $\sim 10^{-3} M_{\odot} \text{yr}^{-1} \text{cMpc}^{-3}$, decreasing by an order of magnitude by $z = 18$. At $8 \lesssim z \lesssim 14$, this total CSFRD exceeds *all JWST* measurements shown in Fig. 5 by ≈ 0.5 to ≈ 1.5 dex, which are based on integrating the UVLF down to $M_{\text{UV}} = -17$. This result implies that, due to the contribution from objects too faint to be currently observed – with $M_{\text{UV}} \gtrsim -17$ or, equivalently, $M_* \lesssim 10^7 M_{\odot}$ – the true CSFRD in the Universe at high redshifts may be significantly higher than current observational estimates suggest.

Fig. 6 decomposes the COLIBRE CSFRD into contributions from subhaloes in different mass bins (with an analogous decomposition of the CSMD presented in Appendix G). Here for clarity, we display only the m6 simulation. The CSFRD is calculated by summing the instantaneous SFRs of individual gas particles that are located within 50 pkpc of subhaloes and gravitationally bound.

In the left panel, the CSFRD is split into contributions from 1-dex stellar mass bins with centres (in log space) at $M_* = 10^6, 10^7, 10^8, 10^9, 10^{10}, 10^{11},$ and $10^{12} M_{\odot}$. As an exception, the lower edge of the lowest mass bin and the upper edge of the highest mass bin are left unbounded to include SFR contributions from subhaloes with $M_* < 10^{5.5} M_{\odot}$ (including those with zero stellar mass) and $M_* > 10^{12.5} M_{\odot}$, respectively. At each redshift, only subhaloes with stellar masses within the specified bin range contribute to the total SFR density of that bin. The corresponding CSFRDs from these subhaloes are shown as solid lines, with colours indicating different mass bins. To emphasize resolution limitations, the CSFRD from the lowest mass bin is represented by the dotted line in grey. The total CSFRD from all subhaloes in the simulated volume is depicted by the dashed orange line. The comparison data (black symbols) are the same as in Fig. 5. In the right panel of Fig. 6, we perform the same CSFRD decomposition as in the left panel, but based on halo mass instead of stellar mass. The (logarithmic) centres of the 1-dex halo mass bins are $M_{\text{halo}} = 10^8, 10^9, 10^{10}, 10^{11}, 10^{12}, 10^{13},$ and $10^{14} M_{\odot}$. As with the stellar mass decomposition, the lower edge of the lowest bin and the upper edge of the highest bin are unbounded. The halo mass is calculated using the definition from G. L. Bryan & M. L. Norman (1998) for central subhaloes and as the total gravitationally bound mass for satellites.

At $0 < z \lesssim 3$, the CSFRD is dominated by subhaloes located around the peak of the SHMR, with $M_{\text{halo}} \sim 10^{12} M_{\odot}$ and $M_* \sim 10^{10} M_{\odot}$. The second-largest contribution at these redshifts comes

from subhaloes in the $M_{\text{halo}} \sim 10^{11} M_{\odot}$ bin. Together, these haloes dominate the CSFRD up to $z \approx 6$, above which less massive subhaloes dominate.

In terms of stellar mass, the CSFRD at $z \gtrsim 6$ is dominated by the lowest stellar mass bin, reflecting the resolution limit of the m6 simulation. The galaxies in this bin either contain only a few stellar particles (with a mass of $m_* \sim 10^6 M_{\odot}$) or have not yet formed any stars. Focussing again on the right panel, we find that the majority of these galaxies must reside in subhaloes with masses of $10^{8.5} \lesssim M_{\text{halo}}/M_{\odot} \lesssim 10^{9.5}$, as these subhaloes dominate the $z \gtrsim 6$ CSFRD¹⁰

The most massive objects ($M_{\text{halo}} \gtrsim 10^{13.5} M_{\odot}$, $M_* \gtrsim 10^{11.5} M_{\odot}$) contribute negligibly to the CSFRD at all redshifts due to their rarity. At $z \approx 0$, their contribution to the total CSFRD is only a few per cent, which gets even smaller at higher redshifts. We emphasize, however, that the mass used to split the CSFRD into bins is the actual mass at each redshift. This implies that the lower mass progenitors of the $z \approx 0$ objects with $M_{\text{halo}} \sim 10^{14} M_{\odot}$ contribute to progressively lower mass bins at higher redshifts.

3.5 Evolution of the star-forming main sequence

An alternative but equally important diagnostic of stellar mass evolution, alongside the GSMF, is the relation between galaxy SFR and its stellar mass. The SFR of star-forming galaxies is known to follow a remarkably tight sequence in the stellar mass – SFR plane, known as the SFMS, with a scatter of 0.2–0.3 dex (e.g. J. Brinchmann et al. 2004; K. E. Whitaker et al. 2012). This strong correlation between the SFR and stellar mass has been observed to hold from $z = 0$ up to $z \approx 6$, with the SFR at a fixed stellar mass declining at lower redshifts as $\text{SFR} \propto (1+z)^{\alpha}$, where $2 < \alpha < 4$ (e.g. J. S. Speagle et al. 2014; P. Popesso et al. 2023).

Modern galaxy formation models, including EAGLE, ILLUSTRIS, and SIMBA, have successfully reproduced the qualitative features of the observed SFMS over $0 < z < 5$. However, at $1 < z < 4$, the simulated SFMS typically falls short in normalization by 0.2–0.5 dex compared to the observed relation (e.g. R. S. Somerville & R. Davé 2015; R. Davé et al. 2019; M. Donnari et al. 2019; A. Katsianis et al. 2020). Growing evidence suggests that the longstanding offset between simulated and observed SFMS arises from methodological inconsistencies in how SFRs are measured in simulations versus observations, as already briefly discussed in Section 3.4 in the context of the CSFRD. While simulations predict SFRs directly, observational SFR estimates rely on a series of assumptions in stellar population modelling – including the assumed dust attenuation curve, gas metallicity, and SFHs – all of which may introduce significant systematic errors (e.g. S. Bellstedt et al. 2020; C. Pacifici et al. 2023).

Bearing these caveats in mind, we now turn to the SFMS in COLIBRE. Because there is no unique way to define the SFMS (e.g. J. Leja et al. 2022), we begin by comparing several definitions of the SFMS in the COLIBRE m6 simulation over a limited redshift range ($0.1 < z < 6$; Fig. 7). After selecting a fiducial definition, we use it to compute the SFMS in the COLIBRE simulations at

all resolutions and compare these predictions with a range of observational data within an extended redshift range ($0.1 < z < 13$; Fig. 8).

Fig. 7 shows the evolution of the SFMS in the L200m6 simulation at redshifts $z = 0.1, 0.5, 1, 2, 4$, and 6 (different panels). The colour scale of the pixels indicates the fraction of galaxies per 0.1 dex bin with a given stellar mass and SFR, normalized by the total number of galaxies across all SFR bins at the same M_* . We show the COLIBRE SFMS defined in three different ways:

- (i) *solid orange lines*: the median SFR of active galaxies, where ‘active’ refers to galaxies whose SFR lies within 1 dex of (or higher than) the SFMS, which is determined iteratively;
- (ii) *short-dashed light-green lines*: the mode of the SFR distribution of all galaxies (both star-forming and quiescent) at fixed M_* ;
- (iii) *dash-dotted red lines*: the mean of the SFR distribution of all galaxies (both star-forming and quiescent) at fixed M_* ;

where in all three cases, the SFMS is computed in 0.2-dex (as opposed to 0.1-dex) stellar mass bins. For reference, the diagonal thin dotted line in each panel corresponds to $\text{sSFR} = 0.2/t_{\text{H}}(z)$, providing an alternative criterion for selecting active galaxies: $\text{sSFR} \geq 0.2/t_{\text{H}}(z)$, where $t_{\text{H}}(z)$ is the Hubble time at redshift z . The SFMSs are displayed only for mass bins that contain at least 30 galaxies and for which the fraction of quenched galaxies (defined¹¹ as $\text{sSFR} < 0.2/t_{\text{H}}(z)$, i.e. independently of the SFMS) is less than 50 per cent, ensuring that the first two SFMS definitions always correspond to the star-forming galaxy population.

In Fig. 7, we see that COLIBRE predicts a realistic SFMS, with the SFR increasing roughly linearly with galaxy stellar mass up to a critical value of $M_* \sim 10^{10} - 10^{11} M_{\odot}$, above which the SFMS begins to flatten. This critical mass increases monotonically with redshift. A second important observation from Fig. 7 is that, at all redshifts ($0 < z < 6$), the first and second SFMS definitions produce nearly identical SFMSs, while the third definition yields an SFMS with a marginally higher normalization at $z \geq 2$, but is otherwise consistent with the first two definitions. Repeating the same test using the m5 simulation (not shown), we found similar trends. However, we noticed that due to the much smaller number of galaxies per bin available in the m5 volume compared to m6, the second definition – based on the peak of the galaxy number density at fixed M_* – yields a slightly noisier SFMS, whereas the third definition – based on the mean SFR – becomes more sensitive to outliers, biasing the SFMS towards higher SFR values. Based on these findings, we adopt the first definition (the median SFR of galaxies with $\text{SFR} > 0.1 \times$ the SFR at the SFMS) as our fiducial choice and use it for comparisons with observational data in Fig. 8, though we note that using either of the other two definitions would have little to no effect on our conclusions.

Fig. 7 also shows the $0.2 < z < 3$ measurements from the Euclid Q1 data release (Euclid Collaboration 2025a), which is based on the Euclid Deep Fields (an effective area of $\approx 63 \text{ deg}^2$) and includes over 4×10^6 objects above the 95 per cent stellar mass

¹⁰The spike at $6.5 < z < 6.75$ in the CSFRD from the $10^{12.5} \leq M_{\text{halo}}/M_{\odot} < 10^{13.5}$ bin is caused by a single starburst galaxy with $M_{\text{halo}} \approx 10^{12.52} M_{\odot}$ ($\approx 10^{12.56} M_{\odot}$) at $z = 6.75$ ($z = 6.5$) whose SFR drops from $\approx 5 \times 10^2 M_{\odot} \text{ yr}^{-1}$ at $6.5 < z < 6.75$ to $\approx 10^2 M_{\odot} \text{ yr}^{-1}$ at $z = 6$.

¹¹We do not define the quenched fraction based on a distance cut from the SFMS at fixed M_* because the SFMS begins to bend downward at $M_* \gtrsim 10^{10} M_{\odot}$, which would bias the quenched fractions at the high-mass end. However, we verified that defining the quenched fraction as the fraction of galaxies with $\text{SFR} < 0.1 \times$ the SFR at the linearly extrapolated SFMS (fit to the simulation data within $10^9 < M_*/M_{\odot} < 10^{10}$) yields results very similar to the definition based on $\text{sSFR} < 0.2/t_{\text{H}}(z)$.

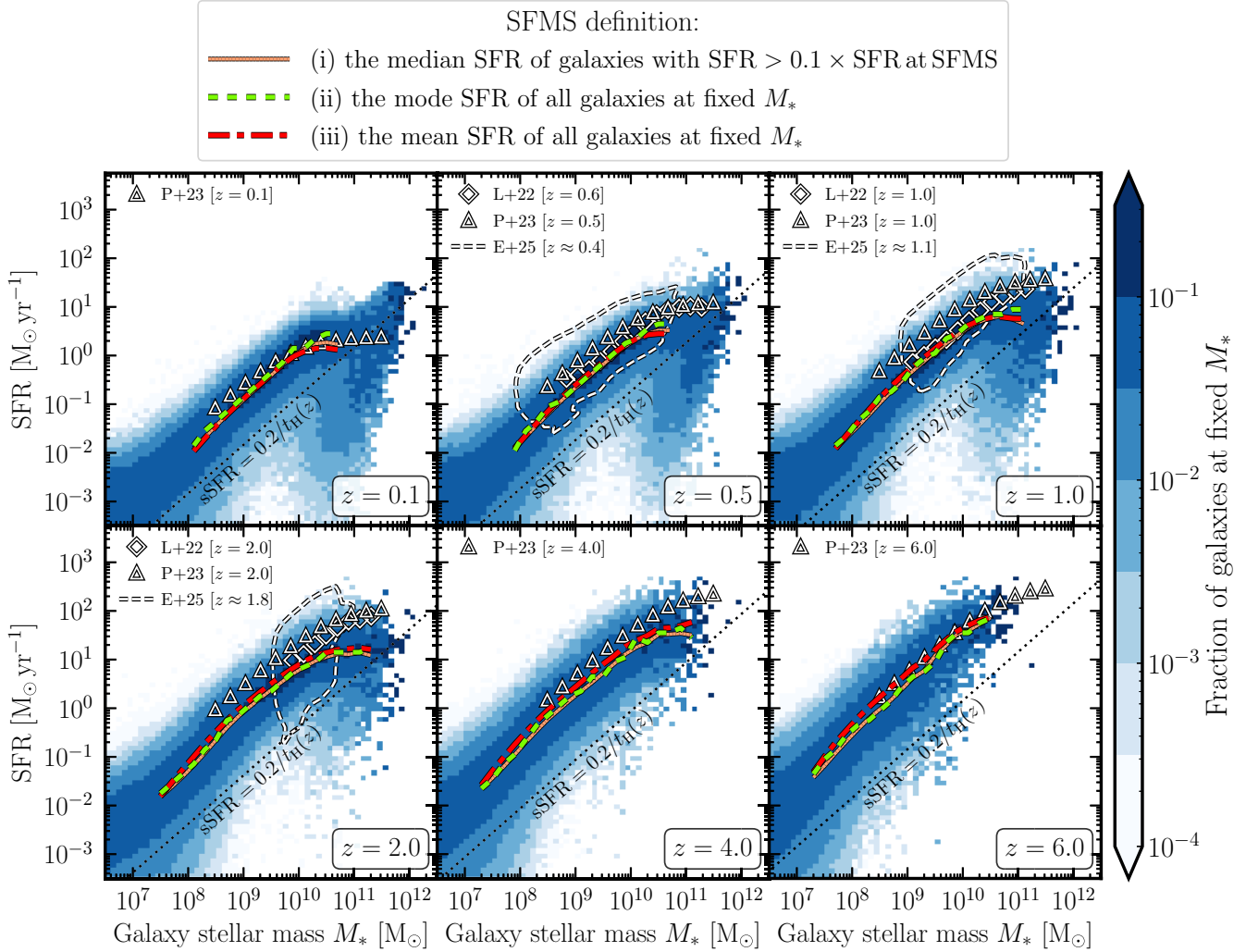


Figure 7. SFMS in the L200m6 simulation at $z = 0.1, 0.5, 1, 2, 4,$ and 6 (different panels). The colour scale shows the fraction of galaxies per pixel with a given M_* and SFR, normalized by the total number of galaxies across all pixels at the same M_* . The solid orange lines represent the COLIBRE SFMS, defined as the median SFR of star-forming galaxies ($\text{SFR} > 0.1 \times$ the SFR at the SFMS determined iteratively). The short-dashed light-green and red dash-dotted curves show alternative SFMS definitions based on the mode and mean of the SFR distribution at fixed M_* , respectively. For reference, the diagonal thin dotted line shows $\text{sSFR} = 0.2/t_H(z)$, which is commonly used as a sSFR threshold for defining star-forming galaxies. Furthermore, also for reference, white diamonds and triangles indicate the observed SFMSs from J. Leja et al. (2022) and P. Popesso et al. (2023), respectively, while the white dashed contours mark the regions enclosing 95 per cent of the sample of star-forming galaxies from Euclid Collaboration (2025a). All three definitions of the SFMS in the simulation yield very similar results at all redshifts.

completeness limit up to $z = 3$. Specifically, from this sample, we show the distribution of star-forming galaxies in the SFR – M_* plane in three redshift bins (median values of $z = 0.37, 1.15,$ and 1.83), indicated by white dashed contours enclosing 95 per cent of the star-forming galaxies (i.e. 2σ of the star-forming distribution). The distribution of galaxies in the L200m6 simulation (i.e. pixels coloured in different shades of blue) largely covers the 2σ contours from Euclid Collaboration (2025a) at all three redshifts. Finally, Fig. 7 includes the observed SFMSs from J. Leja et al. (2022) and P. Popesso et al. (2023), indicated by white diamonds and triangles, respectively. These two SFMSs are shown for reference only, and we defer discussion of how well COLIBRE reproduces these two (and other) observed relations to the next figure.

Fig. 8 shows the evolution of the COLIBRE SFMS from $z = 0.1$ to $z = 13$. The thick solid lines in dark-red, orange, and light-blue denote the SFMS in the COLIBRE m7, m6, and m5 simulations,

respectively. The shaded regions in the corresponding colours indicate the 16th to 84th percentiles of the SFR distribution of galaxies whose SFR is within 1 dex of the SFMS or higher.

For comparison, we include the SFMSs from several recent observational studies. We show the $0 < z < 9$ SFMS best-fitting relation derived from the DEVILS survey (J. E. Thorne et al. 2021); the $0.2 < z < 3$ SFMS from J. Leja et al. (2022), defined as the ‘ridge line’ in the SFR – M_* plane and based on the 3D-*HST* and COSMOS catalogues, with photometry and SED fitting methods identical to those in J. Leja et al. (2020); the SFMS best-fitting relation from P. Popesso et al. (2023, their equation 15) based on the authors’ comprehensive compilation of 27 recent studies spanning $0 < z < 6$, which we further extrapolated to $z = 9$; and the best-fitting formula for the SFMS from M. P. Koprowski et al. (2024), derived through the stacking of FIR imaging of $\sim 100\,000$ K-band-selected galaxies observed with *Herschel* and James Clerk

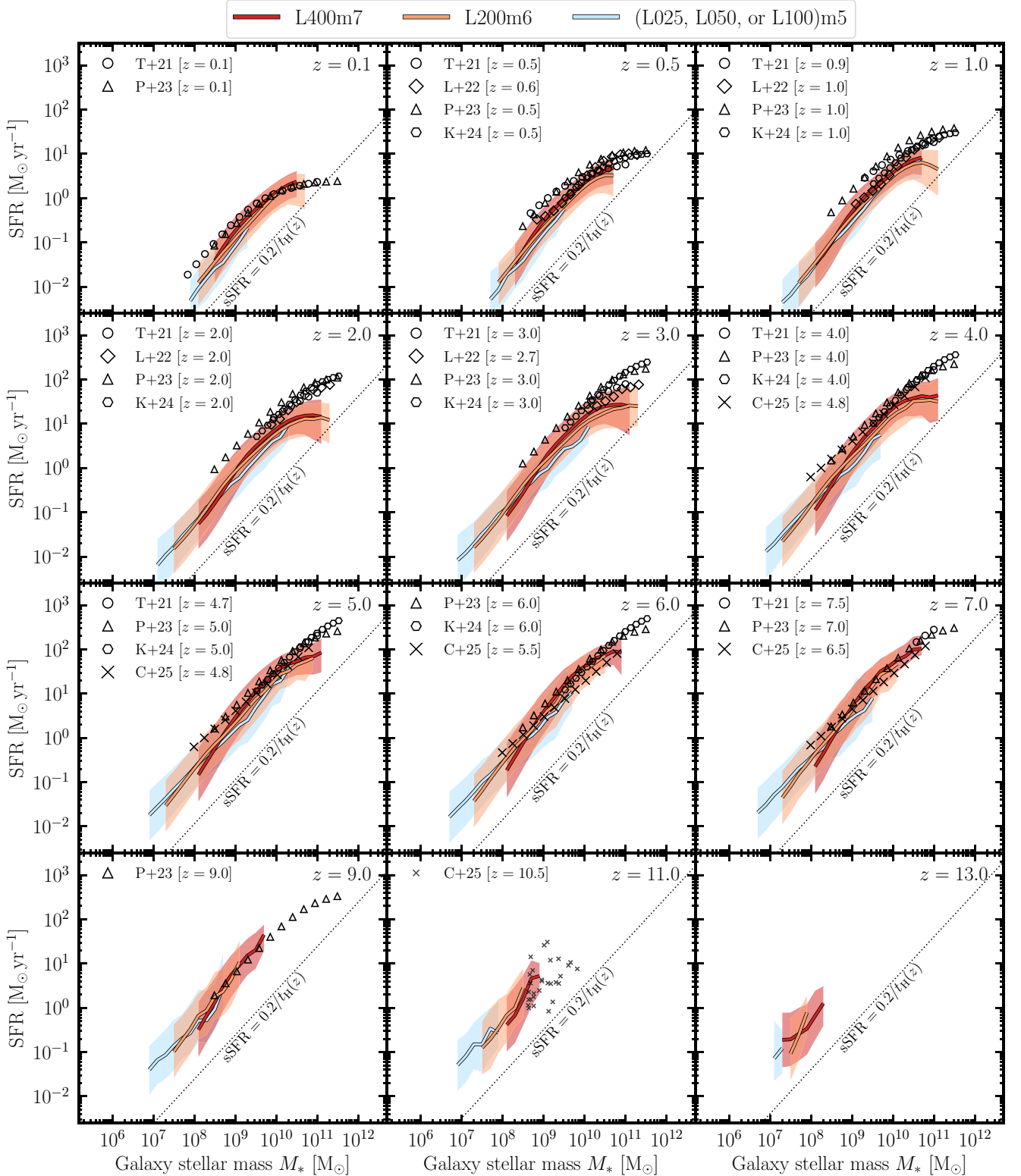


Figure 8. Evolution of the SFMS from $z = 0.1$ (top left) to $z = 13$ (bottom right) in the m7, m6, and m5 COLIBRE simulations, shown in dark-red, orange, and light-blue, respectively. The thick solid lines represent the COLIBRE SFMS, defined as the median SFR of star-forming galaxies, identified iteratively by $\text{SFR} > 0.1 \times$ the SFR at the SFMS. The shaded regions indicate the 16th to 84th percentiles of the SFR of star-forming galaxies ($\text{SFR} > 0.1 \times$ the SFR at the SFMS) in each mass bin. A compilation of observed SFMS data from J. E. Thorne et al. (2021), J. Leja et al. (2022), P. Popesso et al. (2023), M. P. Koprowski et al. (2024), and J. W. Cole et al. (2025) are shown as black symbols. For reference, the diagonal thin dotted line indicates another commonly used criterion for defining star-forming galaxies: $\text{sSFR} \geq 0.2/t_{\text{H}}(z)$. COLIBRE shows good agreement with observational data at low ($z \lesssim 0.5$) and high ($z \gtrsim 5$) redshifts. At intermediate redshifts ($0.5 < z < 4$), the simulations underestimate the observed SFMS in all datasets at $M_* \gtrsim 10^{10.5} M_{\odot}$, but remain broadly consistent with J. Leja et al. (2022) and M. P. Koprowski et al. (2024) at lower stellar masses.

Maxwell Telescope in the UKIDSS Ultra Deep Survey and COSMOS fields over $0.5 < z < 6$. Finally, we show the $4.5 < z < 12$ SFMS from J. W. Cole et al. (2025), based on CEERS data obtained with *JWST*, supplemented by *HST/ACS* and *HST/WFC3* data. Their sample includes 1 863 galaxies, with stellar masses and SFRs derived through SED fitting using BAGPIPES and BPASS stellar evolution models. At $z = 11$, we plot their individual object measurements due to the relatively small number of galaxies.

The SFMS in the m5, m6, and m7 COLIBRE models qualitatively matches the observational data at all examined redshifts. At $z \geq 5$, COLIBRE reproduces both the normalization and slope of the SFMS reported by J. E. Thorne et al. (2021), P. Popesso et al. (2023), and J. W. Cole et al. (2025). For $1 < z < 4$, the COLIBRE SFMS is ≈ 0.2 - 0.6 dex below all observed SFMSs at the high-mass end ($M_* \gtrsim 10^{10.5} M_\odot$), while still broadly matching the SFMS of J. Leja et al. (2022) and, to a somewhat lesser degree, M. P. Koprowski et al. (2024) at $M_* \lesssim 10^{10.5} M_\odot$. Below $z = 1$, the COLIBRE SFMS once again becomes broadly consistent with J. E. Thorne et al. (2021) and P. Popesso et al. (2023). While COLIBRE shows good convergence with resolution at $z \gtrsim 5$, at lower redshifts the higher resolution models tend to predict systematically lower SFRs at fixed stellar mass, leading to larger deviations from the observations. This lower normalization of the SFMS in the higher resolution simulations appears to correlate with a slightly higher normalization of their GSMFs at similar or lower M_* and corresponding redshifts (see Fig. 1), which may point to small but systematic differences in stellar mass growth between different resolution levels.

The fact that COLIBRE falls short of the SFMS from J. E. Thorne et al. (2021) and P. Popesso et al. (2023) by 0.2 - 0.6 dex near cosmic noon warrants further discussion. As previously mentioned, a discrepancy of this magnitude between theoretical models and observations at $1 < z < 4$ is common in galaxy formation studies, and is believed to stem largely from methodological inconsistencies between theoretical predictions and observations. One way to mitigate these discrepancies is therefore to mock-observe simulated galaxies rather than relying on the internal SFRs and stellar masses predicted by the simulations. This approach involves generating realistic SEDs based on the distribution of stellar particle ages and metallicities, which are then processed through the same observational pipelines used for real data to infer SFRs and stellar masses, ensuring a more consistent comparison. By applying such a method to the EAGLE simulations, A. Katsianis et al. (2020) showed that the common tension between observed and simulated SFMS relations is largely alleviated.

Alternatively, the tension can be reduced by improving observational techniques to minimize inevitable systematic errors in the inferred stellar masses and SFRs, which would yield values closer to the true ones, thereby enabling a more consistent comparison with the intrinsic values from simulations (see e.g. Fig. 1 in E. J. Nelson et al. 2021). Among the most promising improvements are the use of Bayesian SED-fitting codes that model galaxy SFHs in a non-parametric way. Unlike traditional parametric models, which assume fixed functional forms for SFHs (e.g. a constant or exponentially declining SFR), non-parametric methods offer greater flexibility, enabling a more accurate inference of the complex and diverse shapes of the SFHs found in real galaxies. A notable example of the non-parametric approach is the work of J. Leja et al. (2022, shown as black diamonds in Fig. 8), who analysed data from the COSMOS and 3D-*HST* surveys over the redshift range $0.2 < z < 3$ using the Bayesian SED-fitting code PROSPECTOR with non-parametric SFHs. They found

systematically higher stellar masses and lower SFRs compared to studies relying on parametric SFHs combined with SFR indicators such as IR, IR + UV, H α , or radio. The authors attributed the SFR discrepancy to dust heating by older stellar populations, as inferred by PROSPECTOR, while arguing that studies reporting higher SFRs have (mis)attributed dust heating to younger stars, leading to overestimated SFRs. In contrast, the increase in stellar masses is ascribed to the contribution of an additional component of old stars captured by non-parametric SFHs (see the discussion in J. Leja et al. 2020). Overall, J. Leja et al. (2022) showed that their observationally inferred SFMS aligns closely with predictions from galaxy formation simulations, including EAGLE (J. Schaye et al. 2015), ILLUSTRISTNG (A. Pillepich et al. 2018), and SIMBA (R. Davé et al. 2019), thereby resolving (or at least strongly mitigating) the longstanding discrepancies between observed and simulated SFRs.

To sum up, in Fig. 8 we find that COLIBRE reproduces the observed SFMS reasonably well at $0 < z < 1$ and $z > 4$, whereas at intermediate redshifts the discrepancy between COLIBRE and observations is ≈ 0.2 - 0.6 dex. While we cannot rule out the possibility that this discrepancy is due to intrinsic limitations of the simulations, and that accounting for systematic errors in observations might even exacerbate it, the growing evidence in the literature suggests that the discrepancy between simulations and observations around cosmic noon is dominated by systematic errors on the observational side, which can likely be alleviated by moving simulated galaxies into observational space and/or by improving observational techniques to reduce these systematic errors.

3.6 Evolution of the galaxy quenched fraction

Another valuable statistic derived from SFRs is the fraction of quenched galaxies at a given stellar mass. These are galaxies that are forming stars at such a low rate that their stellar mass remains nearly unchanged over a Hubble time. Accurately reproducing quenching is important, as quenched galaxies account for more than ≈ 40 per cent of the observed CSMD at $z \approx 0$ (e.g. J. R. Weaver et al. 2023b). Observations indicate that the fraction of quenched galaxies increases with cosmic time at fixed stellar mass, and with stellar mass at fixed redshift (see G. De Lucia et al. 2024, for a recent review on cosmic quenching). The primary mechanisms driving galaxy quenching include (i) AGN feedback in massive galaxies ($M_* \gtrsim 10^{10} M_\odot$), (ii) environmental effects such as tidal and ram pressure stripping in satellite galaxies of massive haloes, and (iii) stellar feedback in low-mass galaxies.

We show the fraction of quenched galaxies as a function of stellar mass in Fig. 9. Each panel corresponds to a different redshift, ranging from $z = 0.1$ (top left) to $z = 7$ (bottom right). The quenched fraction in the simulations (shown as solid lines) is defined as the fraction of galaxies within a given stellar mass bin with sSFRs below $0.2/t_H(z)$ (black dotted lines in Fig. 7). Such thresholds – which compare galaxy sSFRs to a fraction of either the Hubble time or the age of the Universe at a given redshift – are commonly used in the literature to distinguish between quiescent and star-forming galaxy populations (e.g. M. Franx et al. 2008; A. Gallazzi et al. 2014; A. C. Carnall et al. 2024; T. A. Russell et al. 2025; W. M. Baker et al. 2025a). At low stellar masses ($M_* \lesssim 10^9 M_\odot$), the solid lines become dotted to emphasize that the rise in quenched fractions may be driven by limited numerical resolution (see below). The solid lines also become dotted at the high-mass end, indicating that there are fewer than 10 objects per

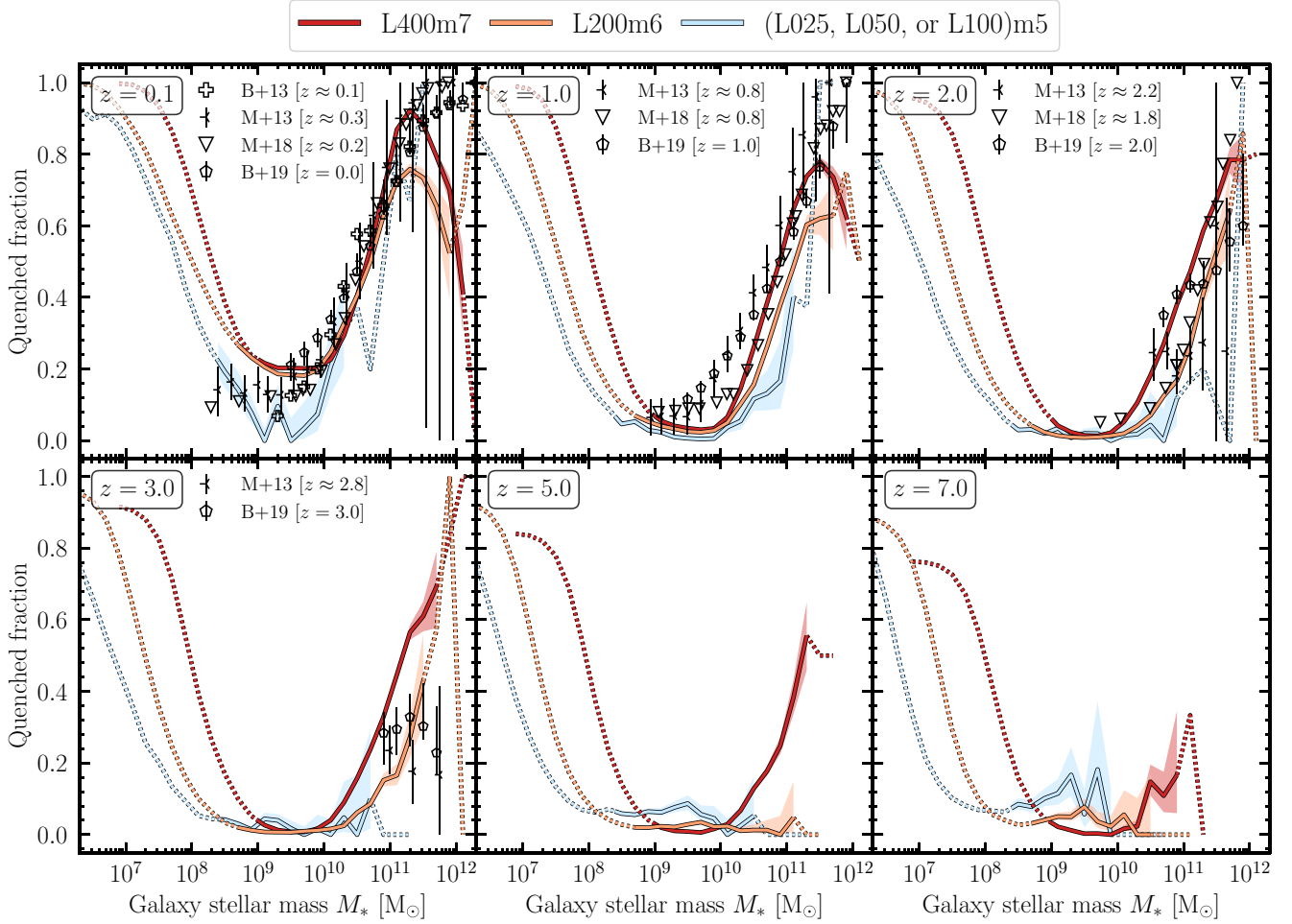


Figure 9. Evolution of the galaxy quenched fraction from $z = 0.1$ to $z = 7$ (different panels). Thick solid lines show the quenched fraction in the simulations, where galaxies are considered quenched if their $\text{sSFR} < 0.2/t_{\text{H}}(z)$. Shaded regions indicate the uncertainty in the predicted quenched fractions, calculated as the Clopper–Pearson interval at the 68 per cent confidence level. At low stellar masses ($M_* \lesssim 10^9 M_\odot$), the solid lines become dotted to indicate that the increase in quenched fraction may be an artefact of limited numerical resolution. The lines also become dotted at high M_* where the number of objects per bin drops below 10. Black symbols represent a compilation of observational data from A. Muzzin et al. (2013), A. E. Bauer et al. (2013), B. P. Moster et al. (2018), and P. Behroozi et al. (2019). Overall, COLIBRE shows good agreement with the data at all redshifts. However, at $z \sim 0$, the most massive galaxies ($M_* \gtrsim 10^{11.5} M_\odot$) appear more star-forming than observed.

stellar mass bin. The shaded regions indicate the uncertainty in the quenched fractions, computed as the Clopper–Pearson interval at the 68 per cent confidence level.

For comparison, we include the $0 < z < 4$ quenched fractions inferred from the semi-empirical model EMERGE (B. P. Moster et al. 2018), where quiescent galaxies are defined using $\text{sSFR} < 0.3/t_{\text{H}}(z)$, and the $0 < z < 5$ quenched fractions inferred from the semi-empirical model UNIVERSEMACHINE (P. Behroozi et al. 2019), where quiescent galaxies satisfy $\text{sSFR} < 10^{-11} \text{ yr}^{-1}$. Additionally, we include the $z \approx 0.1$ quenched fractions from GAMA reported by A. E. Bauer et al. (2013), which were re-calculated for the $\text{sSFR} < 10^{-11} \text{ yr}^{-1}$ threshold by P. Behroozi et al. (2019). Finally, we incorporate the $0.2 < z < 4$ quenched fractions from the COSMOS/UltraVISTA Survey (A. Muzzin et al. 2013), where a quenched galaxy population is defined using UVJ colour selection.

COLIBRE matches the $z \approx 0$ observational data reasonably well, with the exception of the dip at $M_* \gtrsim 10^{11.5} M_\odot$ and the increase at $M_* \lesssim 10^9 M_\odot$. The COLIBRE quenched fraction for

galaxies with $M_* \sim 10^9\text{--}10^{10} M_\odot$ stays at ≈ 20 per cent in the m6 and m7 models, and at ≈ 10 per cent in the m5 model. These values increase steeply for higher stellar masses, reaching $\approx 75\text{--}90$ per cent by $M_* \sim 10^{11.5} M_\odot$, in agreement with observations. For even higher stellar masses ($M_* \gtrsim 10^{11.5} M_\odot$), which correspond to brightest cluster galaxies (BCGs) and are only sampled in the volumes of the m7 and m6 simulations, the quenched fraction drops from $\approx 75\text{--}90$ per cent to $\approx 40\text{--}60$ per cent. This drop is consistent with A. Muzzin et al. (2013) but is not seen in the other datasets we compare to.

To better understand the shape of the quenched fraction– M_* relation at the high-mass end and the importance of AGN feedback in driving it, we re-ran the COLIBRE m7 simulation in a $(50 \text{ cMpc})^3$ cosmological volume without SMBHs. This test confirmed that the rise in the quenched fraction from $M_* \sim 10^{10}$ to $\sim 10^{11.5} M_\odot$ is driven by AGN feedback in massive haloes, while the decrease at $M_* \gtrsim 10^{11.5} M_\odot$ is likely an indication of cooling flows in the BCGs, which the COLIBRE AGN feedback model cannot fully counteract, resulting in overly high SFRs in these

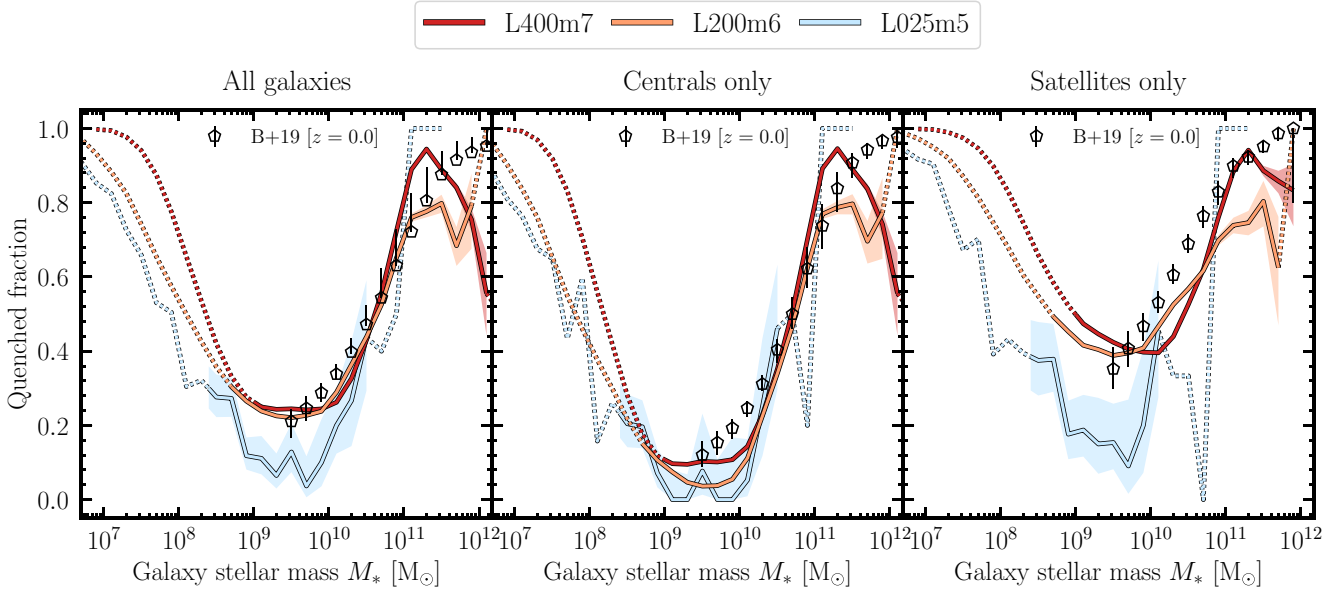


Figure 10. Quenched fractions at $z = 0$ computed for all galaxies (left panel), central galaxies (middle panel), and satellite galaxies (right panel). Black symbols show data from P. Behroozi et al. (2019). COLIBRE predicts that, at fixed stellar mass, satellite galaxies are on average more quenched than centrals, in agreement with P. Behroozi et al. (2019). The m5 model predicts systematically lower satellite quenched fractions compared to m6 and m7, due to its relatively small cosmological volume, which lacks massive host haloes (see Appendix A for details).

objects. However, the dip at $M_* \gtrsim 10^{11.5} M_\odot$ disappears when we increase the assumed log-normal scatter on the stellar mass from ≈ 0.1 dex (as given by equation (1) at $z \approx 0$) to $\gtrsim 0.3$ dex (see Fig. D1 in Appendix D). Similarly, the dip also disappears if the SFR is computed within an aperture of $\lesssim 10$ pkpc (comparable to the physical size at $z = 0.1$ of a fibre with a diameter of a few arcseconds in surveys such as SDSS or GAMA), rather than the default 50 pkpc aperture used for all quantities in this work (not shown).

Below $M_* \sim 10^9 M_\odot$, the increase in galaxy quenched fraction with decreasing stellar mass is likely driven (or at least exacerbated) by numerical effects due to insufficient sampling of the galaxy SFR by star-forming gas particles. As stellar mass decreases, the (star-forming) gas mass also declines, which worsens SFR sampling. Eventually (around a stellar mass of $M_* \sim 10^8 - 10^9 M_\odot$ depending on resolution and redshift), the sampling becomes so sparse that galaxies frequently lack any star-forming gas particles, which artificially boosts the quenched fraction at lower stellar masses. To indicate this limitation of the simulations, in Fig. 9, we switch the solid lines to dotted at stellar masses of $1 \times 10^9 M_\odot$, $0.5 \times 10^9 M_\odot$, and $0.25 \times 10^9 M_\odot$ for the m7, m6, and m5 resolutions, respectively (i.e. $\propto m_{\text{gas}}^{1/3}$), ignoring for simplicity the dependence on redshift. We observe that the increase in quenched fraction at low M_* becomes less severe at higher resolutions but remains prominent even in the m5 simulation. A similar trend in quenched fraction was found in the EAGLE simulations by M. Furlong et al. (2015), although the differences between resolutions were somewhat larger.

At higher redshifts ($z \gtrsim 1$), COLIBRE remains in reasonably good agreement with the observational data for $M_* \gtrsim 10^9 M_\odot$, with the quenched fraction steadily decreasing with increasing redshift at all stellar masses. However, caution should be taken when drawing more detailed conclusions, as differences in the criteria for defining quiescent galaxy populations between

the observations and COLIBRE may influence the results. For galaxies with $M_* \sim 10^9 - 10^{10} M_\odot$, COLIBRE predicts a nearly zero quenched fraction by $z = 1$, which remains negligible also at higher redshifts. For $M_* \sim 10^{11} M_\odot$, the quenched fraction decreases from $\approx 70-80$ per cent at $z = 0.1$ to 20-40 per cent by $z = 2$, and further drops to 5-30 per cent by $z = 5$.

Although the $z = 0.1$ quenched fractions show good convergence with resolution for $10^{10} \lesssim M_*/M_\odot \lesssim 10^{11}$, at $z \geq 1$ the higher resolution simulations predict systematically lower quenched fractions (i.e. weaker quenching) than m7 in this mass range. The explanation for this discrepancy is likely twofold. First, the gas surrounding SMBHs in the higher resolution simulations is, on average, denser than in the m7 simulation, which may increase radiative cooling losses, thereby reducing the effectiveness of AGN feedback on the host galaxy. Secondly, according to the COLIBRE prescription for thermal AGN feedback, BHs inject energy into neighbouring gas particles such that the gas temperature increases by ΔT_{AGN} upon injection. Because the gas particle mass is a factor of 8 (64) lower in the m6 (m5) simulation compared to m7, the energy required to heat a gas particle by the same ΔT_{AGN} is also 8 (64) times lower in m6 (m5) than in m7. In other words, less energy is injected per AGN event in higher resolution simulations, making the AGN feedback less bursty and potentially reducing its efficiency in expelling gas and quenching the galaxy. This may also explain why, at $z = 0.1$, the dip in the galaxy quenched fraction at $M_* \gtrsim 10^{11.5} M_\odot$ is slightly more pronounced at m6 resolution than at m7.

Lastly, at $z = 7$, the COLIBRE m7 model predicts a quenched fraction of about 20 per cent at $M_* \sim 10^{11} M_\odot$, while the cosmological volumes of the m6 and m5 simulations are not large enough to make robust predictions at such high stellar mass and redshift.

Fig. 10 shows the $z = 0$ quenched fractions for all galaxies (left panel), central galaxies (middle panel), and satellite galaxies (right panel). For comparison, we include observational data

from P. Behroozi et al. (2019), which also provides quenched fractions separately for centrals, satellites, and the overall galaxy population.

Satellites in COLIBRE exhibit systematically higher quenched fractions than centrals at fixed stellar mass, a trend consistent at all resolutions. For galaxies with $M_* \sim 10^9\text{-}10^{10} M_\odot$, the quenched fraction of centrals ranges between 0 and 10 per cent, whereas satellites show values of $\approx 15\text{-}40$ per cent, reflecting the important role of environmental effects. The impact of environmental effects on the quenched fraction diminishes for more massive satellites ($M_* \gtrsim 10^{10} M_\odot$). This can be explained by the deeper gravitational potential wells of more massive satellites, which allow them to better retain their gas reservoirs within the host environment (as seen in some numerical simulations and observations; e.g. P. H. Goubert et al. 2024, though see also A. R. Wetzel, J. L. Tinker & C. Conroy 2012; E. Kukstas et al. 2023), and by the fact that they are, on average, accreted later and therefore spend less time within their host haloes. Instead, quenching in these massive satellites should primarily be driven by internal processes (in particular, AGN feedback), in a manner similar to central galaxies. As a result, the quenched fractions of massive satellites become increasingly similar to those of central galaxies at comparable stellar masses. A detailed study of environmental effects on the galaxy quenched fraction and other galaxy properties will be presented in Bahé et al. (in preparation).

For central galaxies, the quenched fraction shows good convergence in the mass range $10^{10} \lesssim M_*/M_\odot \lesssim 10^{11}$ and reasonable convergence at the low ($M_* \lesssim 10^{10} M_\odot$) and high ($M_* \gtrsim 10^{11} M_\odot$) mass ends, where the higher resolution simulations tend to predict somewhat lower values. For satellite galaxies, the convergence is worse: in the $10^9 \lesssim M_*/M_\odot \lesssim 10^{10}$ range, the quenched fraction in the m5 model is a factor of ≈ 2 lower than in the m6 and m7 counterparts. The cause for this discrepancy is the relatively small cosmological volume used in the m5 simulation, which biases the predicted satellite quenched fraction towards lower values due to the lack of massive haloes ($M_{\text{halo}} \gtrsim 10^{13.5} M_\odot$). In Fig. A3 of Appendix A, we show that the bias is significantly reduced when all three resolutions are compared within the same cosmological volume (25^3 cMpc^3).

3.7 The abundance of massive quiescent galaxies at high redshift

We next quantify the abundance of quenched galaxies as a function of redshift and assess whether COLIBRE predicts a number density of massive quiescent objects at high redshifts consistent with recent *JWST* constraints.

Fig. 11 shows the redshift evolution of the number density of quiescent galaxies at $1.75 < z < 9.25$. Unlike in Figs 9 and 10, here we define quenched galaxies as those with $\text{sSFR} < 10^{-10} \text{ yr}^{-1}$, to facilitate a fair comparison with the observational data (discussed later). We present the number densities of quiescent galaxies in COLIBRE for several stellar mass bins commonly used in observational studies: $M_* > 10^{10} M_\odot$ (top panel), $M_* > 10^{10.5} M_\odot$ (middle panel), and $M_* > 10^{11} M_\odot$ (bottom panel). The shaded regions indicate Poisson uncertainties in the simulation predictions. We show the predictions from the largest COLIBRE simulations at each resolution: L400m7, L200m6, and L100m5 (colours). Unlike in previous figures, we do not switch from the L100m5 simulation to L050m5 below $z = 4.6$, as the 50^3 cMpc^3 volume is too small for a meaningful comparison (see Fig. A2 in Appendix A).

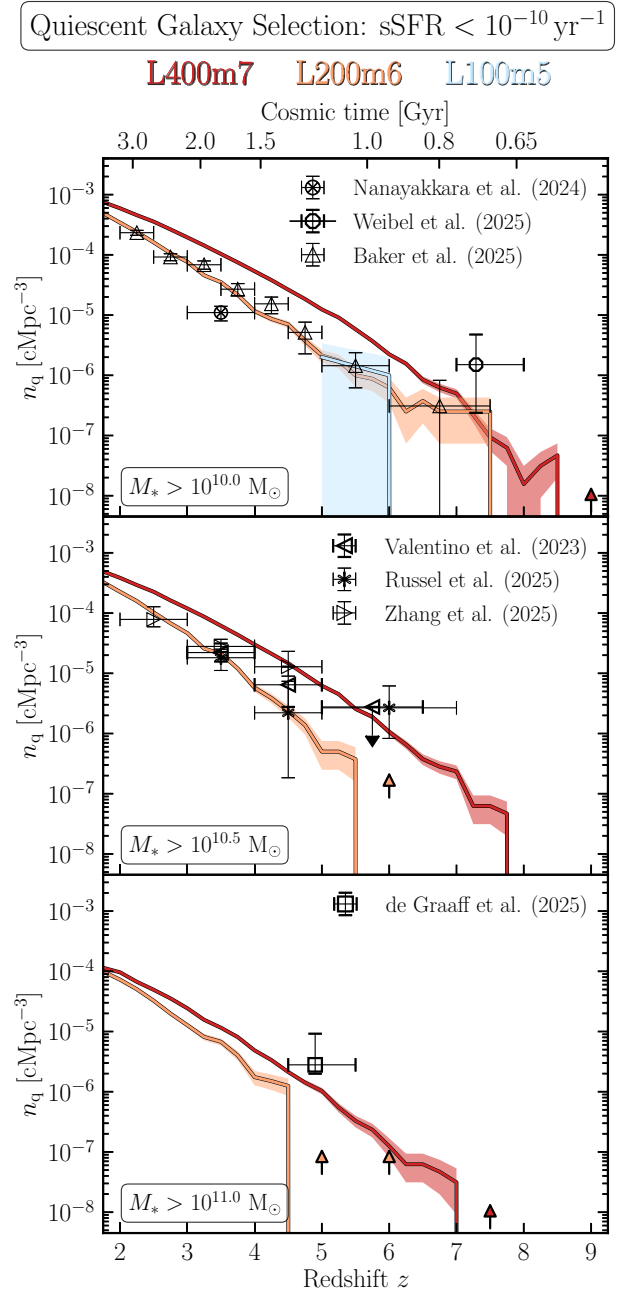


Figure 11. Evolution of the comoving number density of massive quiescent galaxies in the COLIBRE L400m7, L200m6, and L100m5 simulations (coloured curves). Quiescent galaxies are defined as those with $\text{sSFR} < 10^{-10} \text{ yr}^{-1}$, mimicking observational selection criteria at high redshift. We show the number densities of quiescent galaxies with stellar masses $M_* > 10^{10} M_\odot$ (top panel), $M_* > 10^{10.5} M_\odot$ (middle panel), and $M_* > 10^{11} M_\odot$ (bottom panel). The shaded regions indicate the Poisson uncertainties on the simulation predictions. The upward-pointing arrows denote non-zero number density values in the regime of small-number statistics at high redshift and are coloured consistently with the corresponding curves (see the text for details). For comparison, we include *JWST* data with similar M_* and sSFR thresholds from F. Valentino et al. (2023), T. Nanayakkara et al. (2024), T. A. Russell et al. (2025), A. Weibel et al. (2025b), A. de Graaff et al. (2025), W. M. Baker et al. (2025b), and Y. Zhang et al. (2025). All *JWST* data up to $z \approx 7$ are consistent with at least one of the COLIBRE simulations, although the m6 model systematically predicts a lower number density of quiescent galaxies than m7 at fixed redshift.

Following Section 2.3, the number densities predicted by the simulations are computed using galaxy stellar masses with a log-normal scatter (equation (1)) to account for Eddington bias. The effect of Eddington bias on the number density of quiescent galaxies is discussed in Appendix D, where we show that adding the scatter improves the agreement with observations (see Fig. D3). At very high redshift, the number of quiescent objects may fluctuate between zero and a few because of small-number statistics. In such cases, we use upward-pointing arrows (instead of plotting a continuous curve) to indicate redshift bins where the number density of quiescent galaxies is greater than zero but both adjacent bins are zero. The arrows are coloured consistently with the curves corresponding to the different resolutions, with their tips marking the corresponding non-zero number density values.

For comparison, in Fig. 11 we include *JWST* data from F. Valentino et al. (2023), T. Nanayakkara et al. (2024), T. A. Russell et al. (2025), A. Weibel et al. (2025b), A. de Graaff et al. (2025), W. M. Baker et al. (2025b), and Y. Zhang et al. (2025). Briefly,

(i) T. Nanayakkara et al. (2024) reported 12 quiescent galaxy candidates with spectroscopically confirmed redshifts in the redshift range $z = 3-4$. All galaxies have stellar masses $M_* > 10^{10} M_\odot$ and were classified as quiescent using the UVJ colour selection criterion. Additionally, the authors showed that all but one galaxy have $\text{sSFR} < 10^{-10} \text{ yr}^{-1}$ over the past 10 Myr.

(ii) A. Weibel et al. (2025b) reported the discovery of a massive quiescent galaxy at redshift $z = 7.29$ within the $\sim 300 \text{ arcmin}^2$ survey area covered by PRIMER and CEERS. Using *JWST*/NIRCam and MIRI photometry, along with *JWST*/NIRSpec PRISM spectroscopy from the RUBIES programme, they estimated the galaxy’s stellar mass to be $M_* \approx 10^{10.2} M_\odot$, with an sSFR of $< 10^{-10} \text{ yr}^{-1}$ over the past 50 Myr.

(iii) F. Valentino et al. (2023) found ≈ 80 quiescent galaxy candidates with photometric redshifts of $3 < z < 5$ based on NIRCam imaging in 11 *JWST* fields, covering a total effective area of $\approx 145 \text{ arcmin}^2$. From this study, we plot the ‘strict’ galaxy subsample with stellar masses $M_* > 10^{10.6} M_\odot$ that were classified as quiescent based on rest-frame UVJ colours.

(iv) T. A. Russell et al. (2025) identified ~ 100 quiescent galaxies at redshifts $3 < z < 7$ in the CEERS, NEP, and JADES fields (covering a total area of $\approx 144.45 \text{ arcmin}^2$), using *JWST*/NIRCam imaging data. The galaxies were classified as quiescent if their sSFR satisfied $\text{sSFR} < 0.2/t_{\text{age}}(z)$, where $t_{\text{age}}(z)$ is the age of the Universe at redshift z . At $z = 5$, this threshold corresponds to an $\text{sSFR} \approx 1.7 \times 10^{-10} \text{ yr}^{-1}$. In Fig. 11, we plot the ‘robust’ subsample of galaxies from this study with stellar masses $M_* > 10^{10.6} M_\odot$ (12 galaxies in total).

(v) A. de Graaff et al. (2025) identified one quiescent galaxy with $M_* \sim 10^{11} M_\odot$ at $z = 4.9$ within the RUBIES survey, observed using both *JWST*/NIRCam and *JWST*/NIRSpec. An effective surface area used in the study was $\sim 100 \text{ arcmin}^2$. The authors defined the galaxy as quiescent based on its spectrum and determined that the galaxy has an sSFR over the past 100 Myr of just $\sim 10^{-11} \text{ yr}^{-1}$. Since the authors do not report an uncertainty on their measurement, we estimate it assuming a Poisson distribution with an unknown mean λ and an observed count of $k = 1$. Adopting a flat prior on λ and using Bayesian inference, the posterior is a gamma distribution with shape parameter $\alpha = 2$. We take the 16th and 84th percentiles of this distribution as an estimate of the uncertainty.

(vi) W. M. Baker et al. (2025b) analysed a sample of 745 massive quiescent galaxies over the redshift range $2 < z < 7$, using *JWST*/NIRCam imaging from a compilation of public *JWST* fields covering a combined area of over 800 arcmin^2 . Photometric redshifts were derived using EAZY, and quiescent galaxies were identified based on an sSFR threshold of $\text{sSFR} < 0.2/t_{\text{age}}(z)$. For our comparison, we consider only the subset of their sample with stellar masses $M_* > 10^{10} M_\odot$.

(vii) Y. Zhang et al. (2025) presented a sample of 20 massive ($M_* > 10^{10.3} M_\odot$) quiescent galaxies at $2 < z < 5$ with *JWST*/NIRSpec PRISM spectra from the RUBIES survey, with a total surface area of $\approx 150 \text{ arcmin}^2$. We plot their subsample of galaxies satisfying $\text{sSFR} < 10^{-10} \text{ yr}^{-1}$ and $M_* > 10^{10.5} M_\odot$.

We find that the number density of massive quiescent galaxies, n_q , predicted by the COLIBRE simulations increases steeply with decreasing redshift, rising by more than four orders of magnitude between $z = 9$ and $z = 2$ in all three stellar mass bins. Furthermore, comparing the different panels at fixed redshift, we see that n_q decreases as the stellar mass threshold increases from $M_* = 10^{10} M_\odot$ to $10^{11} M_\odot$. Thanks to the relatively large volumes of the m7 and m6 simulations, quiescent galaxies with $M_* > 10^{10} M_\odot$ are sampled out to $z = 9$ and $z = 7.5$ in the m7 and m6 simulations, respectively. For the most massive quiescent galaxies ($M_* > 10^{11} M_\odot$), the m7 (m6) simulation predicts a non-zero number density at $z \leq 7.5$ ($z \leq 6$). The predictions from the m5 simulation are present only over a very narrow redshift interval, $5 < z < 6$, and only for the lowest stellar mass bin ($M_* > 10^{10} M_\odot$). This is expected: at $z \lesssim 5$ the simulation is not yet available, while at higher redshifts the $(100 \text{ cMpc})^3$ volume of the m5 simulation is too small to yield robust predictions, particularly in the higher stellar mass bins.

When comparing the m6 and m7 simulations at fixed redshift and for the same stellar mass bin, we find that the predicted values of n_q are systematically higher in the m7 simulation, with differences ranging from ≈ 0.2 to $\approx 1 \text{ dex}$ ¹². As shown in Appendix A, these discrepancies cannot be due to the eight times larger volume of the m7 simulation. Instead, they arise from weaker, more gradual quenching at high redshifts in the higher resolution COLIBRE simulations – a trend already observed in the evolution of the galaxy quenched fraction (see Fig. 9). As discussed in Section 3.6, a likely explanation for this behaviour is stronger radiative cooling losses at higher resolution, combined with less bursty AGN feedback, leading to overall weaker AGN feedback and consequently weaker quenching of massive galaxies.

All *JWST* measurements are reproduced by at least one of the COLIBRE simulations. In the $M_* > 10^{10} M_\odot$ mass bin, the m6 and m7 simulations both predict $n_q \sim 10^{-6.5} \text{ cMpc}^{-3}$ at $z \approx 7$, which lies within $\approx 1\sigma$ of the measurements by A. Weibel et al. (2025b) and W. M. Baker et al. (2025b). The m6 simulation is also consistent with W. M. Baker et al. (2025b) at all lower redshifts probed by the observations ($2 < z < 6$), as well as with the measurement of T. Nanayakkara et al. (2024) at $z \approx 3.5$. The

¹²Unlike the m6 and m7 models, the m5 and m6 simulations formally converge for the lowest stellar mass bin in the narrow redshift range where the predictions of the L100m5 simulation are available ($5 < z < 6$). However, given the tendency of the 100^3 cMpc^3 volume to produce higher n_q than larger volumes at these redshifts (see Fig. A2), the predictions of the m5 model in a larger volume would likely lie slightly below those of L200m6.

m5 model, where available, is likewise consistent with W. M. Baker et al. (2025b). In contrast, the m7 model overpredicts the observationally inferred number densities from W. M. Baker et al. (2025b) and T. Nanayakkara et al. (2024) by ≈ 0.25 dex to ≈ 1 dex at these redshifts. In the $M_* > 10^{10.5} M_\odot$ bin, the m7 and m6 simulations predict quiescent galaxy number densities of $n_q \sim 10^{-6} \text{ cMpc}^{-3}$ at $z \approx 6$ and $z \approx 5$, respectively, rising in both cases to $\sim 10^{-3.5} \text{ cMpc}^{-3}$ by $z \approx 2$. The high-redshift measurements reported by F. Valentino et al. (2023) and T. A. Russell et al. (2025) at $z \approx 6$, as well as by Y. Zhang et al. (2025) at $z \approx 4.5$, are reproduced by the m7 simulation, while the m6 simulation shows good agreement with the lower redshift bins ($z \lesssim 4$) reported in these studies. For the highest mass bin, the m7 (m6) simulation predicts a number density of $\approx 1.5 \times 10^{-6} \text{ cMpc}^{-3}$ ($\approx 10^{-7} \text{ cMpc}^{-3}$) at $z \approx 4.9$, placing it within $\approx 1.5\sigma$ ($\approx 3\sigma$) of the single galaxy-based estimate by A. de Graaff et al. (2025).

We note that earlier galaxy formation simulations and semi-analytic models that predict reasonably realistic low-redshift galaxy populations typically strongly underpredict the number density of massive quiescent galaxies observed by *JWST* at $z \gtrsim 5$ (see e.g. fig. 6 in A. Weibel et al. 2025b and the section ‘Theoretical predictions’ in A. de Graaff et al. 2025). The improved performance of COLIBRE in this metric may be attributed to several factors. First, unlike most earlier galaxy simulations, COLIBRE does not implement an effective equation of state (EoS) to model the ISM. The absence of an EoS results in denser and cooler gas within the ISM, which can enable more rapid BH accretion. Consequently, the more efficiently growing SMBHs may trigger stronger AGN feedback, facilitating the quenching of massive galaxies at earlier times. Secondly, COLIBRE does not impose a cap on SMBH accretion at the Eddington rate. This approach also allows faster SMBH growth and potentially stronger quenching, particularly at high redshifts where super-Eddington accretion events are expected to be more common (e.g. F. Huško et al. 2025). A detailed investigation of the relevance of these factors will be conducted in future work.

Furthermore, as explained in Section 2.3, galaxy stellar masses in this work include a log-normal scatter with a standard deviation given by equation (1) to account for Eddington bias, which is present in observational data due to non-negligible uncertainties in the inferred stellar masses. As we show in Fig. D3 in Appendix D, the number density of massive quiescent galaxies is particularly sensitive to this scatter, which on average boosts the value of n_q at fixed redshift and stellar mass bin. This effect may not have been fully accounted for in some previous works comparing simulations with observations, potentially biasing simulation-based estimates of n_q towards lower values.

Lastly, we note that many earlier galaxy formation simulations were conducted in cosmological volumes of $\sim 100^3 \text{ cMpc}^3$, which strongly limits the non-zero number density of quenched galaxies that can be predicted at high redshift. For example, C. C. Lovell et al. (2023) showed that the highest redshift at which a non-zero abundance of massive quenched systems is found in the EAGLE simulations increases from $z \approx 5$ in the original EAGLE volume of 100^3 cMpc^3 to $z \approx 8$ in the FLARES simulations, which are zoom-in simulations selected from a parent volume of 3.2^3 Gpc^3 and employing the same EAGLE galaxy formation model.

We emphasize that the number density of quiescent galaxies presented in this work is computed using an $\text{sSFR} < 10^{-10} \text{ yr}^{-1}$ threshold to define quiescent galaxies. In all but one of the observational studies that we show for comparison in Fig. 11, galaxies

were selected using (or satisfy) a similar sSFR threshold (see precise definitions above) to maximize consistency in the comparison and minimize potential bias in our conclusions. Furthermore, very recently, our results were confirmed by Á. Chandro-Gómez et al. (2025), who conducted a detailed study of the properties of massive quenched galaxies in COLIBRE at high redshift, finding generally good agreement with *JWST* data while adopting $\text{sSFR} < 0.2/t_{\text{age}}(z)$ to define quiescent systems.

3.8 Comparison of models with thermal and hybrid AGN feedback

Thus far, we have discussed exclusively the COLIBRE simulations with thermal AGN feedback. However, the COLIBRE suite also includes models with hybrid (thermal + kinetic jet) AGN feedback. In this section, we make use of these additional simulations to investigate how differences in AGN feedback prescriptions impact the key galaxy properties examined in this study: stellar masses, SFRs, and quenched fractions. We emphasize that both the thermal and hybrid models were calibrated to reproduce the $z = 0$ GSMF from S. P. Driver et al. (2022) and the $z \approx 0$ stellar mass–size relation from J. A. Hardwick et al. (2022) in the stellar mass range $10^9 < M_*/M_\odot < 10^{11.3}$, as well as the $z \approx 0$ masses of SMBHs in massive galaxies. This implies that the properties of galaxies other than their $z = 0$ stellar masses, half-mass radii, and SMBH masses should not necessarily be expected to match, especially at higher redshifts and outside the calibrated stellar mass range, where no observational constraints were applied to tune the simulations. Additionally, although the primary difference between the thermal and hybrid models lies in the AGN feedback prescription, we note that one of the SN feedback parameters also differs slightly. Namely, at all resolutions, the normalization of the gas density-dependent heating temperature $\Delta T_{\text{SN}}(\rho_{\text{gas}}) \propto \rho_{\text{gas}}^{2/3}$ is increased in the hybrid model, resulting in lower heating temperatures at fixed gas density and therefore weaker SN feedback. This adjustment was introduced to counterbalance the stronger AGN feedback in the hybrid model in intermediate-mass galaxies (see F. Huško et al. 2026 for details).

For clarity, we present a comparison of the COLIBRE thermal and hybrid models at a single resolution level: m7. We selected this resolution¹³ to maximize galaxy number statistics at the high-mass end, as the largest COLIBRE volumes are available for m7 resolution. Specifically, the largest m7 volume in which the hybrid model was run is $(200 \text{ cMpc})^3$, whereas the thermal model is available in $(200 \text{ cMpc})^3$ and $(400 \text{ cMpc})^3$ volumes. Therefore, to ensure a consistent comparison, we use the $(200 \text{ cMpc})^3$ volume for both models.

The top, middle, and bottom row of Fig. 12 show, respectively, the GSMF, the SFMS, and the galaxy quenched fraction, all in the m7 COLIBRE simulations using thermal (dark-red) and hybrid (yellow) AGN feedback. The three columns, from left to right, correspond to redshifts $z = 0.1$, $z = 2$, and $z = 5$. The solid lines and shaded regions in the second and third rows are defined in the same way as in Figs 8 and 9, respectively. The redshifts $z = 0.1$, 2, and 5 have been chosen to highlight various evolutionary phases of the galaxy populations:

¹³We have verified that the general trends discussed below for m7 resolution also hold at m6 resolution.

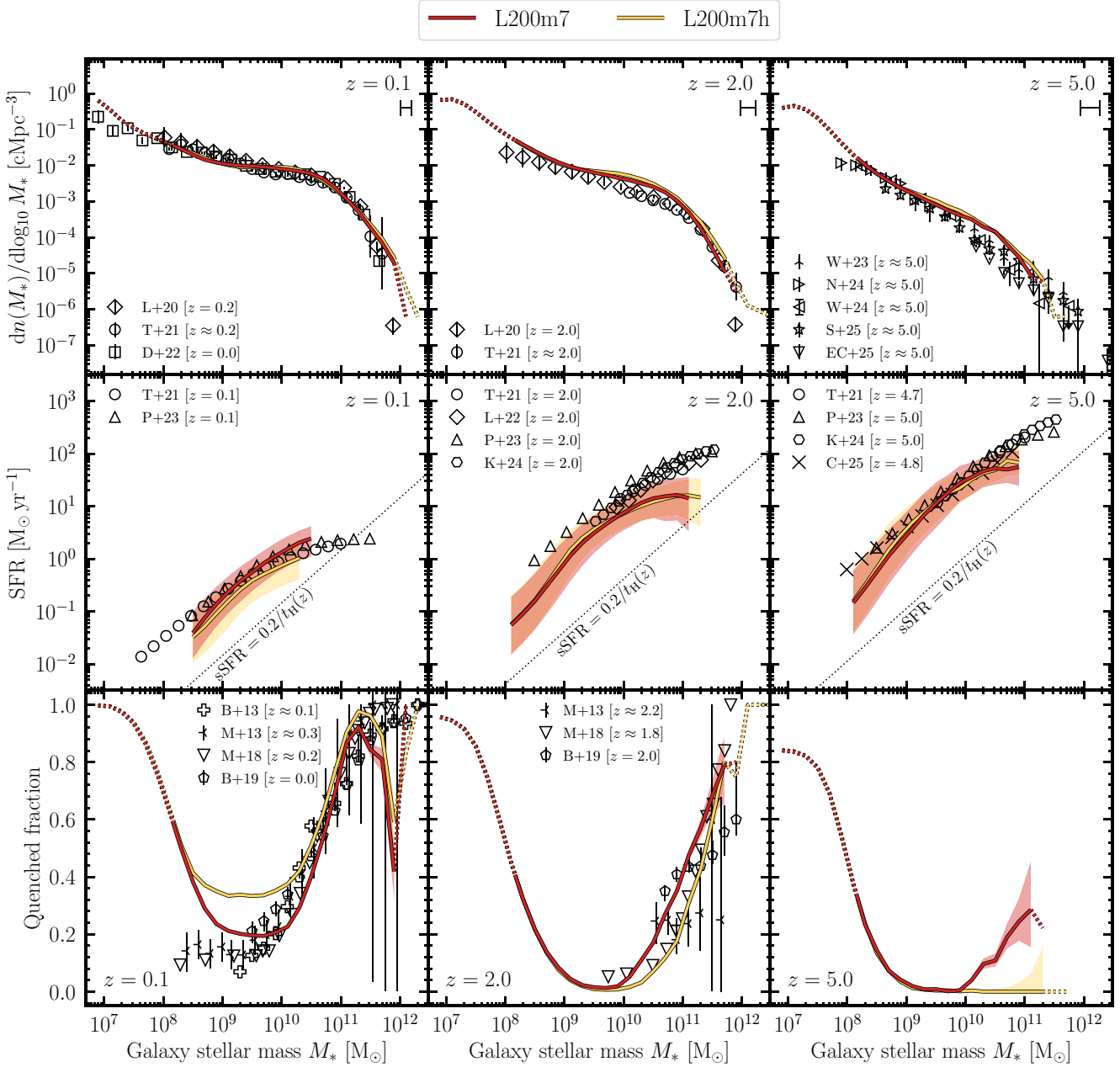


Figure 12. Comparison of the COLIBRE m7 simulations in a $(200 \text{ cMpc})^3$ volume with thermal (dark-red) and hybrid (yellow) AGN feedback at $z = 0.1$ (left), $z = 2$ (middle), and $z = 5$ (right). *Top row:* GSMF. *Middle row:* SFMS, defined as the median SFR of galaxies with $\text{SFR} > 0.1 \times$ the SFR at the SFMS. *Bottom row:* galaxy quenched fraction, where quenched galaxies are defined as $\text{sSFR} < 0.2/t_{\text{H}}(z)$. The shaded regions in the middle row indicate the 16th to 84th percentiles of the SFR distribution of galaxies with $\text{SFR} > 0.1 \times$ the SFR at the SFMS, while those in the bottom row show the uncertainty in the quenched fractions, calculated using the Clopper–Pearson interval at the 68 per cent confidence level. For reference, the horizontal black error bar in the top panels marks the systematic uncertainty in stellar mass measurements given by equation (2), while the diagonal thin dotted lines in the middle panels indicate $\text{sSFR} = 0.2/t_{\text{H}}(z)$. The observational data indicated by black symbols are the same as in Figs 1, 2, 8, and 9. While the hybrid and thermal models yield nearly identical GSMFs and SFMSs, the hybrid model shows weaker quenching at $z = 2$ and $z = 5$, but stronger quenching at $z = 0.1$ compared to the thermal model.

(i) At $z = 5$, the GSMFs in the thermal and hybrid models are nearly indistinguishable, with the hybrid GSMF being marginally higher and both GSMFs reproducing the observational data. The SFMSs are also nearly identical. However, the quenched fractions differ: the thermal model predicts a monotonic increase in the quenched fraction with stellar mass, from ≈ 0 per cent at $M_* \sim 10^{10} M_{\odot}$ to ≈ 30 per cent at $M_* \sim 10^{11} M_{\odot}$, whereas the hybrid

model predicts a quenched fraction of zero in the same stellar mass range.

(ii) At $z = 2$, the SFMSs in the hybrid and thermal models remain indistinguishable. Quenching continues to be stronger in the thermal model, although the differences are no longer as pronounced: in the hybrid model, the quenched fraction increases with stellar mass from ≈ 0 per cent at $M_* \sim 10^{10} M_{\odot}$ to

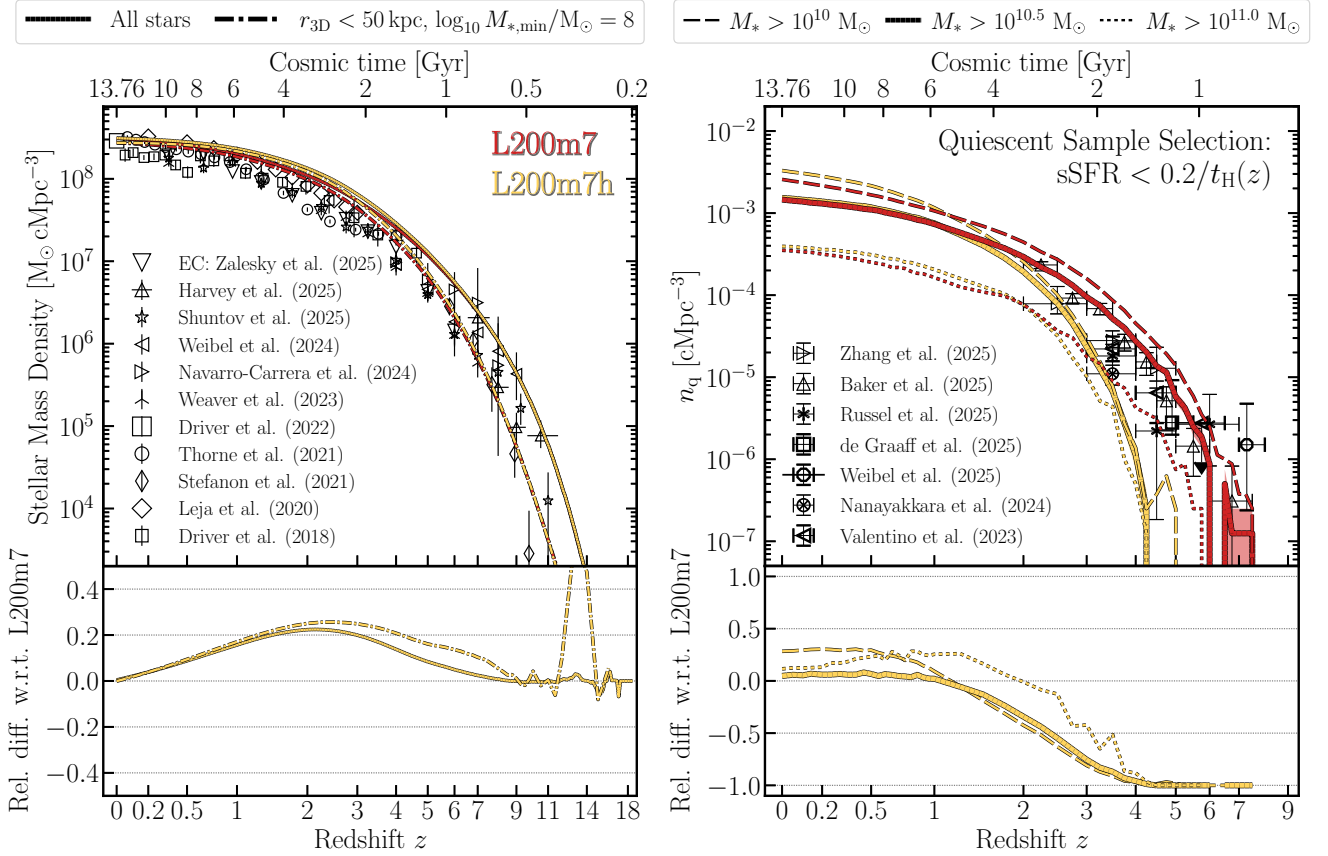


Figure 13. Evolution of the cosmic stellar mass density (CSMD, *left*) and the number density of massive quiescent galaxies, n_q , (*right*) in the COLIBRE L200m7 simulations with thermal (dark-red) and hybrid (yellow) AGN feedback. The top panels show simulation predictions (curves) compared to observational data (black symbols), while the bottom panels display the relative differences between the two simulations, defined as in Fig. 4. Different line styles indicate the different selection criteria used to compute the CSMD and n_q in the simulations (see the main text for details). At $z \gtrsim 1$, the hybrid model predicts a higher CSMD than the thermal model, owing predominantly to weaker quenching at early cosmic times. However, by $z = 0$, the CSMDs of the two models converge, as quenching becomes stronger in the hybrid model than in the thermal model at $z \lesssim 1$.

≈ 30 per cent at $M_* \sim 10^{11} M_\odot$, whereas in the thermal model, it rises from ≈ 0 to ≈ 40 per cent over the same mass range. At even higher stellar masses ($M_* \gtrsim 10^{11.5} M_\odot$), the quenched fractions in the two models converge, while continuing to rise, reaching ≈ 80 per cent by $M_* \approx 5 \times 10^{11} M_\odot$. At $z = 2$, non-negligible differences also emerge in the GSMF, with the hybrid model predicting values that are $\approx 0.05 - 0.1$ dex higher than those of the thermal model at $M_* \gtrsim 10^{9.5} M_\odot$.

(iii) At $z = 0.1$, the GSMFs of the thermal and hybrid models again become indistinguishable, while the normalization of the SFMS is about 0.1 dex lower in the hybrid model. Interestingly, the trend in quenched fractions reverses: the hybrid model now exhibits a higher quenched fraction, reaching ≈ 35 per cent at $10^9 \lesssim M_*/M_\odot \lesssim 10^{10}$, compared to ≈ 20 per cent in the thermal model. This is a consequence of the stronger AGN feedback produced by the hybrid model, which extends into the dwarf galaxy regime (see F. Huško et al. 2026) and will be studied in detail in future work. At higher stellar masses ($M_* \gtrsim 10^{11} M_\odot$), the quenched fractions in both models converge, with the hybrid model predicting marginally higher values. At even greater masses ($M_* \gtrsim 10^{11.5} M_\odot$), the quenched fractions in both models decline to around 50 per cent, suggesting that neither thermal energy injection nor kinetic jets can fully offset the strong cooling flows in the most massive galaxies. We note, however, that the quenched fractions at the high-mass end

are particularly sensitive to uncertainties in stellar mass due to Eddington bias, which may artificially elevate their values (see Appendix D).

The hybrid model exhibits significantly lower quenched fractions at high redshifts due to weaker AGN feedback, resulting mainly from slower BH growth at high z , as well as from the fact that jets in the hybrid model require longer time-scales to quench the host galaxy (see fig. G1 in Á. Chandro-Gómez et al. 2025). Although both the thermal and hybrid AGN feedback models use the modified Bondi–Hoyle–Lyttleton formula (with turbulence and vorticity corrections from M. R. Krumholz et al. 2006) to compute BH mass accretion rates, the latter further modifies the accretion rate by multiplying it by the accretion efficiency, a dimensionless parameter $0 \leq \varepsilon_{\text{accr}} \leq 1$. Physically, this parameter reduces the accretion rate into the BH event horizon relative to that at the scales of the BH accretion disc to account for mass loss due to accretion disc winds (see F. Huško et al. 2026, for details). At very high accretion rates (Eddington fractions greater than 1), which are readily reached in the COLIBRE simulations at high redshift, $\varepsilon_{\text{accr}}$ is set to 1 per cent, which slows down the BH mass growth in the hybrid model relative to the thermal model (see Huško et al., in preparation, for further details).

Over time, the lower quenched fractions in the hybrid model, combined with slightly weaker SN feedback, allow more stellar

mass to form – resulting in a marginally elevated GSMF in the hybrid model at $M_* \gtrsim 10^{9.5} M_\odot$ by $z = 2$. By $z = 0.1$, the GSMFs in both models converge, as they were calibrated to match the same observational constraints. Mathematically, the low- z quenching in the hybrid model is expected to be stronger than in the thermal model, to ensure that the same amount of stellar mass is formed by $z = 0$. Physically, this stronger quenching arises from the hybrid model’s more efficient AGN feedback, which at low redshifts is dominated by the kinetic jet mode, which experiences less radiative losses compared to thermal AGN feedback.

To better understand these differences between the two models, Fig. 13 shows the time evolution of the CSMD (left panel) and the number density of massive quiescent galaxies (right panel). The CSMD is plotted using two different apertures and galaxy stellar mass selection criteria, as in Fig. 4, while the number density of massive quiescent galaxies is shown for three different stellar mass thresholds, as in Fig. 11. In Fig. 13, we use a criterion for quiescent galaxies of $\text{sSFR} < 0.2/t_H(z)$, which is roughly equivalent to a fixed threshold of $\text{sSFR} < 10^{-10} \text{ yr}^{-1}$ at $3 < z < 7$, but approaches $\text{sSFR} < 10^{-11} \text{ yr}^{-1}$ at $z = 0$. The top row shows the predictions from the simulations using the thermal and hybrid models, along with their comparison to observational data. The bottom row presents the relative differences between the models, using the thermal model as the baseline.

Both models form similar amounts of stellar mass at $z \gtrsim 7$. Between $z = 7$ and $z = 2$, the CSMDs in the two models diverge, with the differences peaking at $z \approx 2$, where the hybrid model has formed ≈ 25 per cent more stellar mass. From $z = 2$ to 0, the CSMDs converge, with the differences approaching zero by $z \approx 0$, as both models were calibrated to reproduce the same observed $z = 0$ GSMF.

The behaviour of the number density of quiescent galaxies is in line with the stronger quenching predicted by the thermal model at high redshifts, as shown in Fig. 12. The thermal model predicts a significantly higher number density of quiescent galaxies at $z \gtrsim 2$ across all three stellar mass bins, allowing it to better match the *JWST* data compared to the hybrid model. By contrast, quenching in the hybrid model becomes as strong as that in the thermal model by $z \approx 1$ and exceeds it by about 30 per cent by $z = 0$ for galaxies with $M_* > 10^{10} M_\odot$. The number density of galaxies with $M_* > 10^{10.5} M_\odot$ remains similar between the two models at $0 < z < 1$, while the abundance of the most massive objects ($M_* > 10^{11} M_\odot$) is slightly higher in the hybrid model at $z < 2$, but converges to that of the thermal model by $z = 0$.

4 CONCLUSIONS

In this work, we have studied the evolution of stellar masses and SFRs in new COLIBRE cosmological hydrodynamical simulations of galaxy formation (E. Chaikin et al. 2026; J. Schaye et al. 2026). We analysed a suite of fiducial COLIBRE simulations covering a range of cosmological volumes and resolutions: $(400 \text{ cMpc})^3$ and $(200 \text{ cMpc})^3$ at m7 resolution with a gas particle mass of $m_{\text{gas}} = 1.47 \times 10^7 M_\odot$; $(200 \text{ cMpc})^3$ at m6 resolution with $m_{\text{gas}} = 1.8 \times 10^6 M_\odot$; and $(100 \text{ cMpc})^3$, $(50 \text{ cMpc})^3$, and $(25 \text{ cMpc})^3$ at m5 resolution with $m_{\text{gas}} = 2.3 \times 10^5 M_\odot$. Our main focus was the evolution of the GSMF, which we predicted over $0 < z < 17$. Additionally, we investigated the evolution of the SHMR, SFMS, galaxy quenched fraction, CSFRD, CSMD, and the number density of massive quiescent galaxies at high redshift. To facilitate a fair comparison with observations, where inferred stellar masses

have non-negligible errors, the GSMF, SFMS, galaxy quenched fraction, and number density of massive quiescent galaxies were computed using stellar masses of simulated galaxies that include a log-normal scatter with a standard deviation given by equation (1). Finally, we studied the differences between two AGN feedback implementations that are available in COLIBRE: the fiducial model with thermal energy injections and a hybrid model combining thermal and kinetic jet feedback. Our key findings are as follows:

(i) The COLIBRE GSMF is consistent with the observationally inferred data over the entire redshift range where the observations are available ($0 < z < 12$), while demonstrating very good convergence with resolution over a factor of 64 in gas particle mass (Figs 1 and 2). The largest systematic deviations from the observations, ≈ 0.3 dex, occur at $z \approx 2-4$. At much higher redshifts ($12 < z < 17$), which lack observational measurements, COLIBRE predicts a smoothly evolving GSMF, with the normalization changing by ≈ 2 dex between $z = 12$ and $z = 17$ (Fig. 2). At the highest redshift probed by the simulations, $z = 17$, the combined GSMF predicted by the COLIBRE m5 and m6 resolutions still spans more than 2 dex in stellar mass.

(ii) COLIBRE predicts that the SHMR for central subhaloes peaks at $M_{\text{halo}} \sim 10^{12} M_\odot$ with a value of $M_*/M_{\text{halo}} \approx 0.025$ and logarithmic slopes to the left and to the right of the peak of, respectively, ≈ 1.0 and -0.5 (Fig. 3). At all resolutions, the shape and normalization of the SHMR only mildly change with redshift and at $z = 0.1$ are consistent with semi-empirical models from B. P. Moster et al. (2018) and P. Behroozi et al. (2019).

(iii) The CSMD is consistent with observationally inferred values at $0 < z < 12$ (Fig. 4). At $z = 0$, COLIBRE predicts a total CSMD of $\approx 3 \times 10^8 M_\odot \text{ cMpc}^{-3}$, with ≈ 90 per cent residing in subhaloes of $M_* \geq 10^8 M_\odot$ within 50 pkpc of their centres.

(iv) The CSFRD reproduces observationally inferred values reasonably well for $0 < z < 18$ (Fig. 5). In particular, at $8 < z < 18$, COLIBRE predictions for the CSFRD from subhaloes with $M_* \geq 10^7 M_\odot$ are consistent with *JWST* UV measurements reported by Y. Harikane et al. (2023), C. T. Donnan et al. (2024), and A. Weibel et al. (2025a). At $z \approx 2$, COLIBRE captures some observational trends while underestimating others by 0.2-0.5 dex, reflecting the well-known tension between intrinsic and some observed SFRs during this epoch. By $z = 0$, the predictions are again consistent with observations, yielding a total CSFRD of $\approx 0.012-0.013 M_\odot \text{ yr}^{-1} \text{ cMpc}^{-3}$.

(v) Subhaloes near the peak of the SHMR, characterized by $M_{\text{halo}} \sim 10^{12} M_\odot$ and $M_* \sim 10^{10} M_\odot$, dominate the CSFRD from $z = 0$ to $z \approx 3$. At $z \gtrsim 6$, lower mass subhaloes ($M_{\text{halo}} \lesssim 10^{9.5} M_\odot$) become the primary contributors to the CSFRD. Meanwhile, the contribution from the most massive objects ($M_{\text{halo}} \gtrsim 10^{13.5} M_\odot$) remains negligible at all redshifts (Fig. 6).

(vi) Similarly to the CSFRD, COLIBRE reproduces the observed SFMS reasonably well at $0 < z < 1$ and $z > 4$, but underpredicts the data at intermediate redshifts (Fig. 8). Specifically, between $1 < z < 4$, COLIBRE predicts an SFMS that is lower by $\approx 0.2-0.6$ dex compared to the values reported by J. E. Thorne et al. (2021) and P. Popesso et al. (2023), although it agrees much better with the measurements by J. Leja et al. (2022). At $z < 5$, higher resolution simulations yield an SFMS with a slightly yet systematically lower normalization, leading to somewhat worse agreement with observations, whereas at higher redshifts, the convergence with resolution is generally good.

(vii) COLIBRE predicts a gradual increase in the fraction of quenched galaxies with decreasing redshift, consistent with observations (Fig. 9). At $z \approx 0$, galaxies with $M_* \sim 10^9\text{-}10^{10} M_\odot$ exhibit the lowest quenched fraction ($\approx 10\text{-}20$ per cent), while those with $M_* \sim 10^{11.5} M_\odot$ have the highest ($\approx 75\text{-}90$ per cent). However, for even more massive galaxies ($M_* \gtrsim 10^{11.5} M_\odot$), the quenched fraction decreases to $\approx 40\text{-}60$ per cent, suggesting that AGN feedback does not completely suppress cooling flows in these massive haloes. At fixed stellar mass, COLIBRE predicts higher quenched fractions for satellites compared to centrals due to environmental effects, consistent with comparison data (Fig. 10).

(viii) COLIBRE is consistent with recent measurements by *JWST* of the number density of massive quiescent galaxies at $2 < z < 7$ (Fig. 11). At $z \approx 7$, COLIBRE predicts a comoving number density of galaxies with $M_* > 10^{10} M_\odot$ and $s\text{SFR} < 10^{-10} \text{ yr}^{-1}$ of $\sim 10^{-6.5} \text{ cMpc}^{-3}$, which rises to $\sim 10^{-3.5} \text{ cMpc}^{-3}$ by $z \approx 2$. Moreover, in the largest COLIBRE volume (400^3 cMpc^3), quiescent galaxies with $M_* > 10^{11} M_\odot$ are found out to $z = 7.5$, with a number density of $\sim 10^{-8} \text{ cMpc}^{-3}$ at $z \approx 7.5$ and $\sim 10^{-6} \text{ cMpc}^{-3}$ at $z \approx 5$. Predictions from the largest available m5 simulation (100^3 cMpc^3) are available only over a very narrow redshift range ($5 < z < 6$), owing to the relatively small cosmological volume and because the simulation has thus far progressed only down to $z = 4.6$. At a fixed redshift and stellar mass bin, the m6 model predicts a somewhat lower number density than the m7 model, with differences ranging between ≈ 0.2 and 1 dex.

(ix) The COLIBRE model with hybrid AGN feedback predicts significantly weaker quenching by AGN at $z \gtrsim 2$ compared to the thermal model, resulting in a $\approx 1\text{-}2$ dex lower number density of massive quiescent galaxies at these redshifts (Fig. 13). This weaker AGN quenching, combined with the slightly weaker SN feedback in the hybrid model, leads to an increase in the CSMD at $0.5 < z < 7$ and an elevated GSMF around $z = 2$ relative to the thermal model (Figs 12 and 13). By contrast, at $z = 0$, the hybrid model predicts stronger quenching, with the CSMD and GSMF converging to those of the thermal model.

In closing, we highlight that the COLIBRE simulations are consistent with both the low- and high- z observationally inferred GSMF, including the most recent measurements from *JWST*. The simulations achieve this agreement using a standard universal stellar IMF (G. Chabrier 2003) and a standard Λ CDM cosmology, predicting a nearly redshift-independent baryon conversion efficiency in DM haloes, with a peak value of ≈ 16 per cent. This result challenges recent claims based on *JWST* data suggesting that efficiencies much greater than ~ 10 per cent are necessary to explain the observations of high-mass galaxies at redshifts $z \gtrsim 7$ (e.g. A. Weibel et al. 2024; M. Shuntov et al. 2025). However, we caution that the stellar masses for *JWST* galaxies were inferred using a similar IMF to the one assumed in COLIBRE and that rest-frame UVLF comparisons between COLIBRE and observations have not yet been made.

We also note that, given the likely presence of systematic uncertainties in the observational analyses, we cannot exclude the possibility that some of the agreement between COLIBRE and the data found in this work is fortuitous. In future work, we will perform forward modelling to compare COLIBRE predictions with data in observer space, providing an important consistency check. Further follow-up work will investigate early quenching mechanisms of massive galaxies in COLIBRE and compare low- and high-redshift luminosity functions predicted by the simulations

with observational data, to provide further theoretical insights into the evolution of galaxy stellar mass and SFRs in the observable Universe.

ACKNOWLEDGEMENTS

We thank Yannick M. Bahé, Simon Driver, Victor J. Forouhar Moreno, Anna de Graaff, Claudia del P. Lagos, Christopher Lovell, Kai Wang, John Weaver, and the anonymous referee for their useful comments. This work used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The equipment was funded by BEIS capital funding via STFC capital grants ST/K00042X/1, ST/P002293/1, ST/R002371/1 and ST/S002502/1, Durham University, and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure. This project has received funding from the Netherlands Organization for Scientific Research (NWO) through research programme Athena 184.034.002. ABL acknowledges support by the Italian Ministry for Universities (MUR) programme ‘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 STFC consolidated grant ST/X001075/1. CSF acknowledge support from European Research Council (ERC) Advanced Grant DMIDAS (GA 786910). SP acknowledges support from the Austrian Science Fund (FWF) through project V 982-N. Data analysis was carried out using SWIFTSIMIO (J. Borrow & A. Borrisov 2020; J. Borrow & A. J. Kelly 2021).

DATA AVAILABILITY

All GSMF data presented in this work are available on <https://colibre-simulations.org/>. The remaining data underlying this article will be shared on reasonable request to the corresponding author. The public version of the SWIFT simulation code can be found on www.swiftsim.com. The SWIFT modules related to the COLIBRE galaxy formation model will be integrated into the public version after the public release of COLIBRE.

REFERENCES

- Abbott T. M. C. et al., 2022, *Phys. Rev. D*, 105, 023520
 Arnouts S., Ilbert O., 2011, Astrophysics Source Code Library, record ascl:1108.009
 Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, *ARA&A*, 47, 481
 Bahé Y. M. et al., 2022, *MNRAS*, 516, 167
 Baker W. M. et al., 2025a, *MNRAS*, 539, 557
 Baker W. M. et al., 2025b, *A&A*, 702, A270
 Baldry I. K. et al., 2012, *MNRAS*, 421, 621
 Bauer A. E. et al., 2013, *MNRAS*, 434, 209
 Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, *MNRAS*, 488, 3143
 Behroozi P. S., Conroy C., Wechsler R. H., 2010, *ApJ*, 717, 379
 Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, *ApJ*, 762, 109
 Bellstedt S., Robotham A. S. G., 2025, *MNRAS*, 540, 2703
 Bellstedt S. et al., 2020, *MNRAS*, 498, 5581
 Benítez-Llambay A. et al., 2026, *MNRAS*, 546, stag268
 Benson A. J., Bower R. G., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2003, *ApJ*, 599, 38
 Bocquet S., Saro A., Dolag K., Mohr J. J., 2016, *MNRAS*, 456, 2361
 Booth C. M., Schaye J., 2009, *MNRAS*, 398, 53

- Borrow J., Borrisov A., 2020, *Journal of Open Source Software*, 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
- Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, *MNRAS*, 370, 645
- Brammer G. B., van Dokkum P. G., Coppi P., 2008, *ApJ*, 686, 1503
- Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, *MNRAS*, 351, 1151
- Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
- Bryan G. L., Norman M. L., 1998, *ApJ*, 495, 80
- Bunker A. J. et al., 2023, *A&A*, 677, A88
- Carnall A. C., McLure R. J., Dunlop J. S., Davé R., 2018, *MNRAS*, 480, 4379
- Carnall A. C. et al., 2024, *MNRAS*, 534, 325
- Casey C. M. et al., 2023, *ApJ*, 954, 31
- Casey C. M. et al., 2024, *ApJ*, 965, 98
- 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., 2026, *MNRAS*, 548, stag300
- Chandro-Gómez Á. et al., 2025, preprint ([arXiv:2512.16208](https://arxiv.org/abs/2512.16208))
- Cochrane R. K. et al., 2023, *MNRAS*, 523, 6082
- Cochrane R. K., Katz H., Begley R., Hayward C. C., Best P. N., 2025, *ApJ*, 978, L42
- Cole J. W. et al., 2025, *ApJ*, 979, 193
- Cole S. et al., 2001, *MNRAS*, 326, 255
- Conroy C., Wechsler R. H., 2009, *ApJ*, 696, 620
- Conroy C., Gunn J. E., White M., 2009, *ApJ*, 699, 486
- Correa C. A. et al., 2026, *MNRAS*, 548, stag645
- Covelo-Paz A. et al., 2025, *A&A*, 694, A178
- Crain R. A. et al., 2015, *MNRAS*, 450, 1937
- Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, *MNRAS*, 486, 2827
- Davidzon I. et al., 2017, *A&A*, 605, A70
- Davies L. J. M. et al., 2018, *MNRAS*, 480, 768
- de Graaff A., Trayford J., Franx M., Schaller M., Schaye J., van der Wel A., 2022, *MNRAS*, 511, 2544
- de Graaff A. et al., 2025, *Nat. Astron.*, 9, 280
- De Lucia G., Fontanot F., Xie L., Hirschmann M., 2024, *A&A*, 687, A68
- DESI Collaboration, 2016, preprint ([arXiv:1611.00036](https://arxiv.org/abs/1611.00036))
- Donnan C. T. et al., 2024, *MNRAS*, 533, 3222
- Donnan C. T., Dunlop J. S., McLure R. J., McLeod D. J., Cullen F., 2025, *MNRAS*, 539, 2409
- Donnari M. et al., 2019, *MNRAS*, 485, 4817
- Driver S. P. et al., 2011, *MNRAS*, 413, 971
- Driver S. P. et al., 2012, *MNRAS*, 427, 3244
- Driver S. P. et al., 2018, *MNRAS*, 475, 2891
- Driver S. P. et al., 2022, *MNRAS*, 513, 439
- Duncan K. et al., 2014, *MNRAS*, 444, 2960
- Eddington A. S., 1913, *MNRAS*, 73, 359
- Eke V. R., Cole S., Frenk C. S., 1996, *MNRAS*, 282, 263
- Eke V. R. et al., 2004, *MNRAS*, 355, 769
- Enia A. et al., 2022, *ApJ*, 927, 204
- Euclid Collaboration, 2025a, preprint ([arXiv:2503.15314](https://arxiv.org/abs/2503.15314))
- Euclid Collaboration, 2025b, preprint ([arXiv:2504.17867](https://arxiv.org/abs/2504.17867))
- Euclid Collaboration, 2025c, *A&A*, 695, A229
- Feldmann R. et al., 2025, *MNRAS*, 536, 988
- Forouhar Moreno V. J., Helly J., McGibbon R., Schaye J., Schaller M., Han J., Kugel R., Bahé Y. M., 2025, *MNRAS*, 543, 1339
- Franx M., van Dokkum P. G., Förster Schreiber N. M., Wuyts S., Labbé I., Toft S., 2008, *ApJ*, 688, 770
- Fu S. et al., 2025, *ApJ*, 987, 186
- Furlong M. et al., 2015, *MNRAS*, 450, 4486
- Gallazzi A., Bell E. F., 2009, *ApJS*, 185, 253
- Gallazzi A., Bell E. F., Zibetti S., Brinchmann J., Kelson D. D., 2014, *ApJ*, 788, 72
- Goubert P. H., Bluck A. F. L., Piotrowska J. M., Maiolino R., 2024, *MNRAS*, 528, 4891
- Gruppioni C. et al., 2020, *A&A*, 643, A8
- Guo Q. et al., 2011, *MNRAS*, 413, 101
- Hahn O., Michaux M., Rampf C., Uhlemann C., Angulo R. E., 2020, *Astrophysics Source Code Library*, record ascl:2008.024
- Han J., Cole S., Frenk C. S., Benítez-Llambay A., Helly J., 2018, *MNRAS*, 474, 604
- Hardwick J. A., Cortese L., Obreschkow D., Catinella B., Cook R. H. W., 2022, *MNRAS*, 509, 3751
- Harikane Y. et al., 2023, *ApJS*, 265, 5
- Harvey T. et al., 2025, *ApJ*, 978, 89
- Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, *MNRAS*, 451, 2663
- Huško F., Lacey C. G., Roper W. J., Schaye J., Briggs J. M., Schaller M., 2025, *MNRAS*, 537, 2559
- Huško F. et al., 2026, *MNRAS*, 547, stag324
- Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, *MNRAS*, 321, 372
- Jones G. T., Stanway E. R., Carnall A. C., 2022, *MNRAS*, 514, 5706
- Kajisawa M. et al., 2009, *ApJ*, 702, 1393
- Katsianis A. et al., 2020, *MNRAS*, 492, 5592
- Khusanova Y. et al., 2021, *A&A*, 649, A152
- Koprowski M. P., Wijesekera J. V., Dunlop J. S., McLeod D. J., Michałowski M. J., Lisiecki K., McLure R. J., 2024, *A&A*, 691, A164
- Kriek M., van Dokkum P. G., Labbé I., Franx M., Illingworth G. D., Marchesini D., Quadri R. F., 2009, *ApJ*, 700, 221
- Kroupa P., 2001, *MNRAS*, 322, 231
- Krumholz M. R., McKee C. F., Klein R. I., 2006, *ApJ*, 638, 369
- Kukstas E. et al., 2023, *MNRAS*, 518, 4782
- Lagos C. D. P., Tobar R. J., Robotham A. S. G., Obreschkow D., Mitchell P. D., Power C., Elahi P. J., 2018, *MNRAS*, 481, 3573
- Lagos C. D. P. et al., 2024, *MNRAS*, 531, 3551
- Lagos C. D. P. et al., 2025, *MNRAS*, 536, 2324
- Lee J. H. et al., 2025, *AJ*, 169, 185
- Leitherer C. et al., 1999, *ApJS*, 123, 3
- Leja J., Johnson B. D., Conroy C., van Dokkum P. G., Byler N., 2017, *ApJ*, 837, 170
- Leja J. et al., 2019, *ApJ*, 877, 140
- Leja J., Speagle J. S., Johnson B. D., Conroy C., van Dokkum P., Franx M., 2020, *ApJ*, 893, 111
- Leja J. et al., 2022, *ApJ*, 936, 165
- Liske J. et al., 2015, *MNRAS*, 452, 2087
- Lovell C. C. et al., 2023, *MNRAS*, 525, 5520
- Ludlow A. D., Schaye J., Bower R., 2019, *MNRAS*, 488, 3663
- Ludlow A. D., Fall S. M., Schaye J., Obreschkow D., 2021, *MNRAS*, 508, 5114
- Ludlow A. D., Fall S. M., Wilkinson M. J., Schaye J., Obreschkow D., 2023, *MNRAS*, 525, 5614
- Madau P., Dickinson M., 2014, *ARA&A*, 52, 415
- McCarthy I. G., Schaye J., Bird S., Le Brun A. M. C., 2017, *MNRAS*, 465, 2936
- McGibbon R., Helly J. C., Schaye J., Schaller M., Vandenbroucke B., 2025, *J. Open Source Softw.*, 10, 8252
- Michaux M., Hahn O., Rampf C., Angulo R. E., 2021, *MNRAS*, 500, 663
- Mobasher B. et al., 2015, *ApJ*, 808, 101
- Moster B. P., Somerville R. S., Maulbetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, *ApJ*, 710, 903
- Moster B. P., Naab T., White S. D. M., 2018, *MNRAS*, 477, 1822
- Muzzin A., Marchesini D., van Dokkum P. G., Labbé I., Kriek M., Franx M., 2009, *ApJ*, 701, 1839
- Muzzin A. et al., 2013, *ApJ*, 777, 18
- Nanayakkara T. et al., 2024, *Sci. Rep.*, 14, 3724
- Narayanan D. et al., 2024, *ApJ*, 961, 73
- Navarro-Carrera R., Rinaldi P., Caputi K. I., Iani E., Kokorev V., van Mierlo S. E., 2024, *ApJ*, 961, 207
- Nelson E. J. et al., 2021, *MNRAS*, 508, 219

- Nobels F. S. J., Schaye J., Schaller M., Ploekinger S., Chaikin E., Richings A. J., 2024, *MNRAS*, 532, 3299
- Novak M. et al., 2017, *A&A*, 602, A5
- Obreschkow D., Murray S. G., Robotham A. S. G., Westmeier T., 2018, *MNRAS*, 474, 5500
- Pacifici C. et al., 2023, *ApJ*, 944, 141
- Papovich C. et al., 2023, *ApJ*, 949, L18
- Pförr J., Maraston C., Tonini C., 2012, *MNRAS*, 422, 3285
- Pillepich A. et al., 2018, *MNRAS*, 473, 4077
- Ploekinger S., Richings A. J., Schaye J., Trayford J. W., Schaller M., Chaikin E., 2025, *MNRAS*, 543, 891
- Popesso P. et al., 2023, *MNRAS*, 519, 1526
- Pozzetti L. et al., 2010, *A&A*, 523, A13
- Richings A. J., Schaye J., Oppenheimer B. D., 2014a, *MNRAS*, 440, 3349
- Richings A. J., Schaye J., Oppenheimer B. D., 2014b, *MNRAS*, 442, 2780
- Robotham A. S. G., Driver S. P., 2011, *MNRAS*, 413, 2570
- Robotham A. S. G., Bellstedt S., Lagos C. d. P., Thorne J. E., Davies L. J., Driver S. P., Bravo M., 2020, *MNRAS*, 495, 905
- Russell T. A. et al., 2025, *MNRAS*, 544, 4482
- Salpeter E. E., 1955, *ApJ*, 121, 161
- Santini P. et al., 2012, *A&A*, 538, A33
- Santini P. et al., 2023, *ApJ*, 942, L27
- Schaller M. et al., 2015, *MNRAS*, 451, 1247
- 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
- Schmidt M., 1959, *ApJ*, 129, 243
- Scoville N. et al., 2007, *ApJS*, 172, 1
- Shuntov M. et al., 2025, *A&A*, 695, A20
- Somerville R. S., Davé R., 2015, *ARA&A*, 53, 51
- Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008, *MNRAS*, 391, 481
- Somerville R. S., Gilmore R. C., Primack J. R., Domínguez A., 2012, *MNRAS*, 423, 1992
- Song M. et al., 2016, *ApJ*, 825, 5
- Speagle J. S., Steinhardt C. L., Capak P. L., Silverman J. D., 2014, *ApJS*, 214, 15
- Stanway E. R., Eldridge J. J., 2018, *MNRAS*, 479, 75
- Stefanon M., Bouwens R. J., Labbé I., Illingworth G. D., Gonzalez V., Oesch P. A., 2021, *ApJ*, 922, 29
- Taylor E. N. et al., 2011, *MNRAS*, 418, 1587
- Thorne J. E. et al., 2021, *MNRAS*, 505, 540
- Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, *ApJ*, 688, 709
- Torrey P., Vogelsberger M., Genel S., Sijacki D., Springel V., Hernquist L., 2014, *MNRAS*, 438, 1985
- Trayford J. W. et al., 2026, *MNRAS*, 545, staf2040
- Trinca A., Schneider R., Valiante R., Graziani L., Ferrotti A., Omukai K., Chon S., 2024, *MNRAS*, 529, 3563
- Valentino F. et al., 2023, *ApJ*, 947, 20
- Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, *MNRAS*, 436, 3031
- Vogelsberger M. et al., 2014, *MNRAS*, 444, 1518
- Walcher J., Groves B., Budavári T., Dale D., 2011, *Ap&SS*, 331, 1
- Wang J., Bose S., Frenk C. S., Gao L., Jenkins A., Springel V., White S. D. M., 2020, *Nature*, 585, 39
- Wang T. et al., 2025, *ApJ*, 988, L35
- Weaver J. R. et al., 2022, *ApJS*, 258, 11
- Weaver J. R. et al., 2023a, *ApJS*, 269, 20
- Weaver J. R. et al., 2023b, *A&A*, 677, A184
- Weibel A. et al., 2024, *MNRAS*, 533, 1808
- Weibel A. et al., 2025a, preprint (arXiv:2507.06292)
- Weibel A. et al., 2025b, *ApJ*, 983, 11
- Wetzel A. R., Tinker J. L., Conroy C., 2012, *MNRAS*, 424, 232
- Whitaker K. E., van Dokkum P. G., Brammer G., Franx M., 2012, *ApJ*, 754, L29
- Wilkins S. M., Trentham N., Hopkins A. M., 2008, *MNRAS*, 385, 687
- Wilkins S. M., Lovell C. C., Stanway E. R., 2019, *MNRAS*, 490, 5359
- Wilkinson M. J., Ludlow A. D., Lagos C. d. P., Fall S. M., Schaye J., Obreschkow D., 2023, *MNRAS*, 519, 5942
- Xu K. et al., 2025, *MNRAS*, 540, 1635
- Yu H., Wang F. Y., 2016, *ApJ*, 820, 114
- Zhang Y. et al., 2025, preprint (arXiv:2508.08577)

APPENDIX A: CONVERGENCE WITH RESOLUTION AND VOLUME

In this section, we test the convergence of the COLIBRE simulations with numerical resolution and with cosmological volume. Because some subgrid parameter values in the COLIBRE models differ slightly between resolutions (see table 1 in J. Schaye et al. 2026, for a summary of these differences), convergence with resolution should be understood as ‘weak convergence’, following the terminology of J. Schaye et al. (2015).

We emphasize that in COLIBRE, the strength of SN and AGN feedback was optimized to reproduce the observed $z = 0$ GSMF and the size–stellar mass relation in the stellar mass range $10^9 < M_*/M_\odot < 10^{11.3}$ (see E. Chaikin et al. 2026, for details). Furthermore, the coupling efficiency of AGN feedback was tuned to match the observed $z = 0$ BH mass–stellar mass relation of massive galaxies. The calibration was performed independently at each resolution, and no constraints at higher redshift were applied. Consequently, the COLIBRE predictions at higher redshifts or outside the mass range used for calibration are not guaranteed to converge between different resolution levels, particularly for statistics other than those used in the calibration.

A1 Convergence of the cosmic stellar mass density

Fig. A1 shows the evolution of the CSMD where the left-hand panels compare the COLIBRE m5, m6, and m7 simulations, all in a $(25 \text{ cMpc})^3$ volume, while the right-hand panels compare $(25 \text{ cMpc})^3$, $(50 \text{ cMpc})^3$, $(100 \text{ cMpc})^3$, and $(200 \text{ cMpc})^3$ volume simulations, all at m6 resolution. The layout of the panels is the same as in Fig. 4.

The CSMD exhibits very good convergence with resolution at $z \lesssim 5$, with relative differences between adjacent resolutions within ≈ 20 per cent. This convergence holds both for the total CSMD, computed using all stellar particles in the simulation (solid curves), and for the CSMD contributed only by subhaloes with stellar masses $M_* \geq 10^8 M_\odot$ (dash-dotted curves). At higher redshifts ($z \gtrsim 6$), the total CSMDs begin to diverge, with higher resolution simulations predicting systematically higher values of the CSMD. This trend is expected, as the total CSMD at these high redshifts is dominated by SFRs of the least massive subhaloes (see Fig. 6), whose abundance is sensitive to numerical resolution.

The CSMD also shows very good convergence with cosmological volume at fixed resolution for $z \lesssim 4$, which holds for both the total CSMD and that contributed only by $M_* \geq 10^8 M_\odot$ subhaloes. At higher redshifts, convergence remains good for the total CSMD up to $z \approx 14$, with relative differences between simulations in different volumes remaining within ≈ 30 per cent. This behaviour is expected, as at high redshifts the total CSMD is dominated by the least massive objects ($M_* \lesssim 10^{6.5} M_\odot$; see Fig. G1), which are abundant in all four volumes. In contrast, the CSMD from $M_* \geq 10^8 M_\odot$ subhaloes depends sensitively on cosmological volume above $z \approx 4$. On average, larger volumes yield higher CSMDs at high redshifts, as they are more likely to host $M_* \geq 10^8 M_\odot$

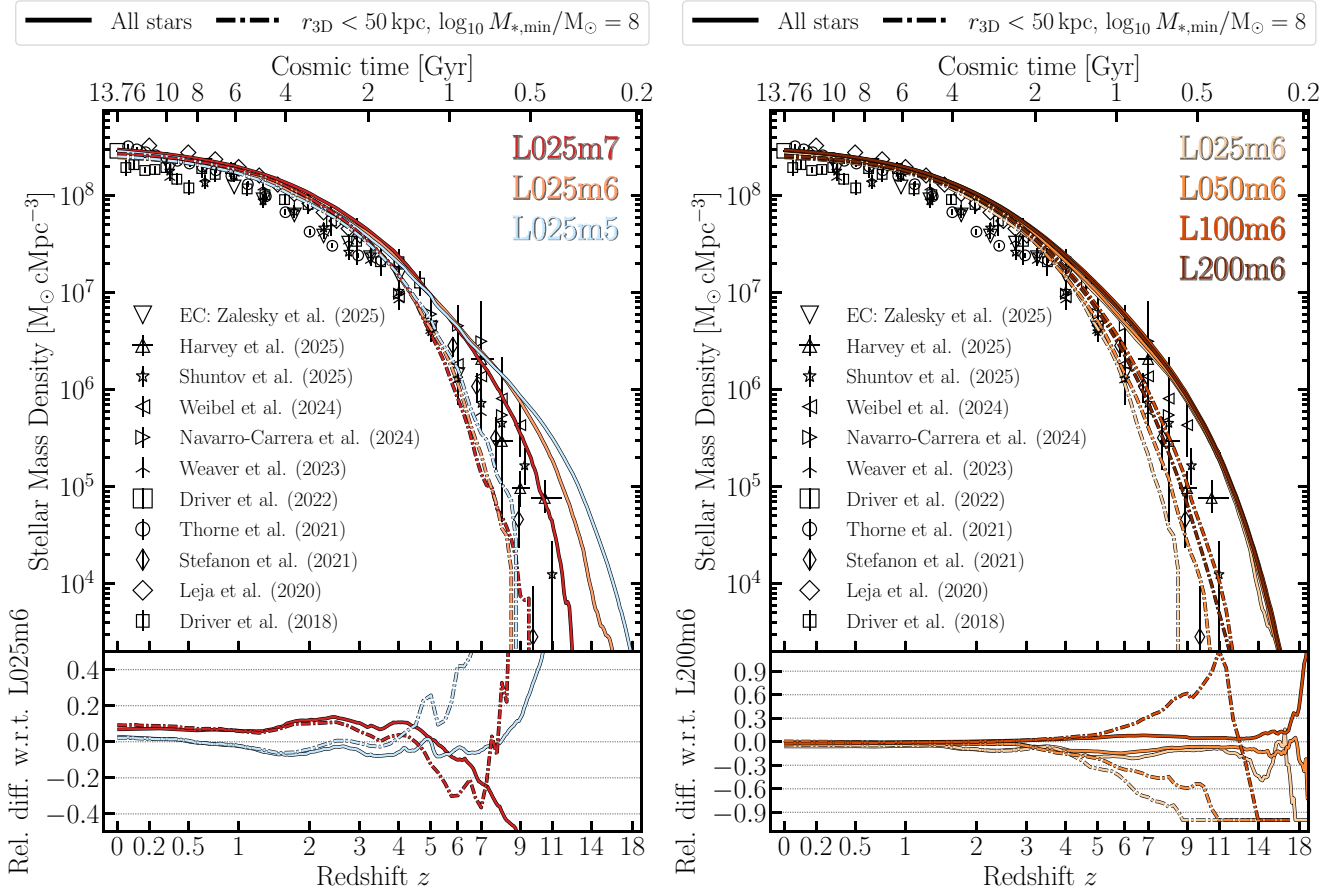


Figure A1. The impact of numerical resolution (*left*) and cosmological volume (*right*) on the evolution of the CSMD. The left-hand panels compare the m5, m6, and m7 resolution simulations in the $(25 \text{ cMpc})^3$ volume, while the right-hand panels compare $(25 \text{ cMpc})^3$, $(50 \text{ cMpc})^3$, $(100 \text{ cMpc})^3$, and $(200 \text{ cMpc})^3$ volumes at m6 resolution. The layout of the panels is the same as in Fig. 4. The CSMD exhibits very good convergence with resolution (cosmological volume) at $z < 5$ ($z < 4$), with relative differences remaining below ≈ 20 per cent.

galaxies. One exception to this trend is the $(100 \text{ cMpc})^3$ volume, which predicts a slightly higher CSMD than that in the larger $(200 \text{ cMpc})^3$ volume due to a rare overdensity in its ICs that forms a massive protocluster by $z \sim 5$, atypical for a $(100 \text{ cMpc})^3$ volume.

A2 Convergence of the number density of massive quiescent galaxies

Fig. A2 shows the evolution of the comoving number density of massive quiescent galaxies predicted by the COLIBRE m7 simulations in different cosmological volumes: $(50 \text{ cMpc})^3$, $(100 \text{ cMpc})^3$, $(200 \text{ cMpc})^3$, and $(400 \text{ cMpc})^3$, coloured in progressively darker shades of red for larger volumes. The comparison data (black symbols) are the same as in Fig. 11, and the number densities are computed for the same three stellar

mass bins: $M_* > 10^{10} M_\odot$ (top panel), $M_* > 10^{10.5} M_\odot$ (middle panel), and $M_* > 10^{11} M_\odot$ (bottom panel). For reference, thin horizontal dotted lines indicate the minimum non-zero number density of quiescent galaxies resolvable in each cosmological volume.

While predictions from volumes of $(100 \text{ cMpc})^3$ or larger converge at $z \lesssim 5$, the $(50 \text{ cMpc})^3$ volume is too small to yield meaningful predictions above $z \approx 3.5$. In particular, its minimum resolvable number density of $\sim 10^{-5} \text{ cMpc}^{-3}$ makes it unsuitable for comparison with high-redshift ($z \gtrsim 4$) observations, which typically report number densities comparable to or below this level. The $(100 \text{ cMpc})^3$ volume extends the predictions of non-zero number densities out to $z \approx 6.5$ for the adopted sSFR and M_* cuts, with a minimum resolved number density of $10^{-6} \text{ cMpc}^{-3}$; however, its predictions at $z \gtrsim 4$ are subject to small-number statistics.

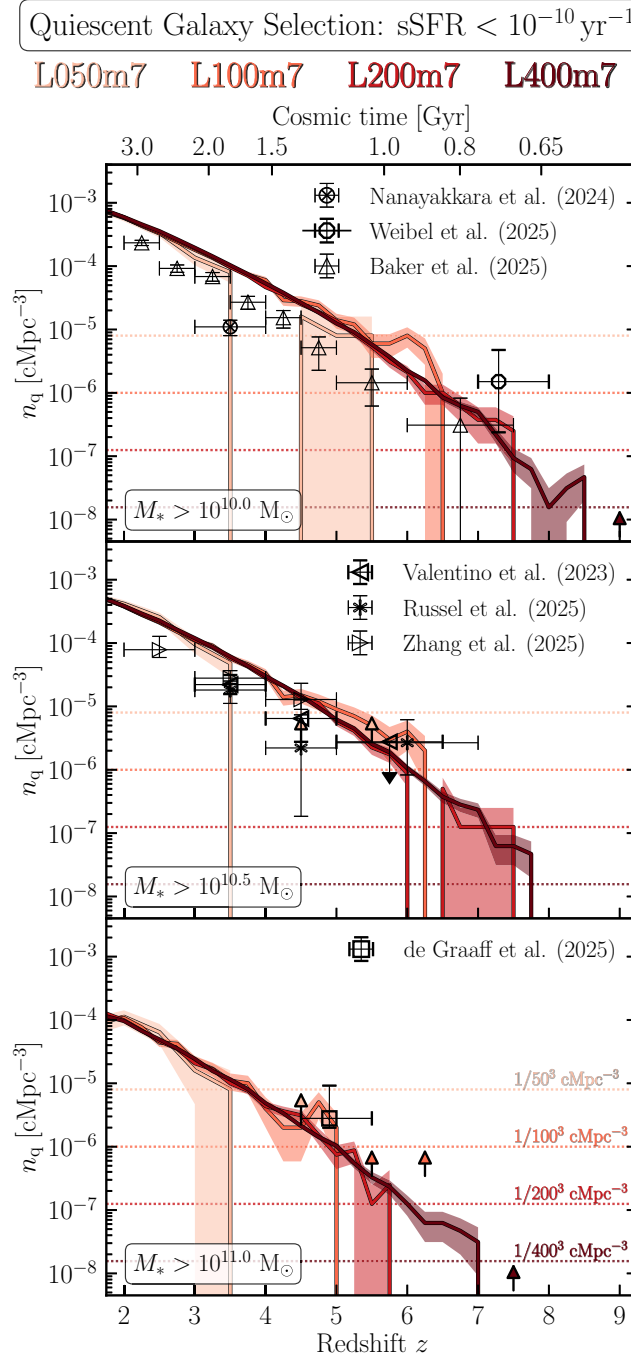


Figure A2. As Fig. 11, but shows the comoving number density of massive quiescent galaxies predicted by the COLIBRE m7 simulations in different cosmological volumes: $(50 \text{ cMpc})^3$, $(100 \text{ cMpc})^3$, $(200 \text{ cMpc})^3$, and $(400 \text{ cMpc})^3$, indicated by different shades of red. For reference, the thin horizontal dotted lines mark the minimum non-zero number density of quiescent galaxies that each volume can resolve. While predictions from volumes of $(100 \text{ cMpc})^3$ and larger converge at $z \lesssim 5$, the $(50 \text{ cMpc})^3$ volume is too small to provide meaningful predictions at $z \gtrsim 3.5$.

A3 Convergence of the galaxy quenched fraction

Fig. A3 shows the $z = 0$ quenched fraction for all, central, and satellite galaxies predicted by the simulations at m7, m6, and m5 resolutions, all using the $(25 \text{ cMpc})^3$ cosmological volume.

We find that the convergence of the satellite quenched fraction between m5 and the two lower resolutions improves significantly when the same cosmological volume is used, compared to Fig. 10, which shows the quenched fraction in the simulations at different resolutions in different volumes.

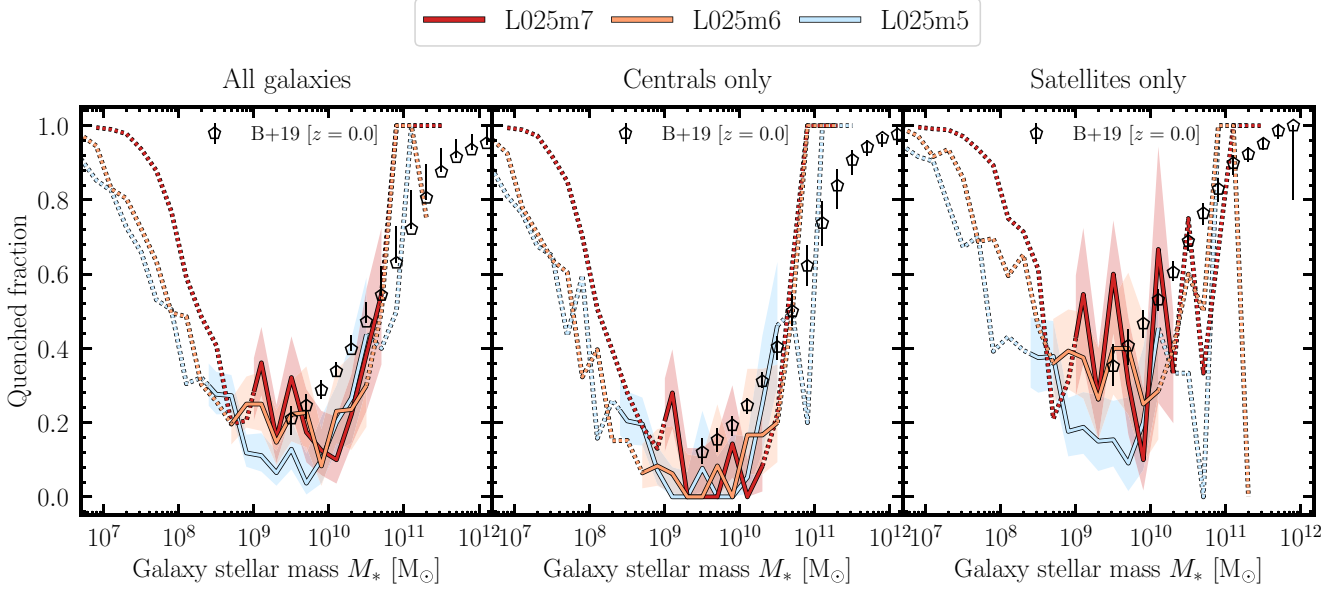


Figure A3. As in Fig. 10, but with all simulations using the same cosmological volume of $(25 \text{ cMpc})^3$. Adopting the same cosmological volume at all COLIBRE resolutions improves the convergence of the quenched fraction for satellite galaxies, compared to Fig. 10, where the m5 simulation uses a much smaller volume than the m6 and m7 counterparts.

APPENDIX B: THE EFFECT OF THE CHOICE OF APERTURE

In this section, we explore the impact of the aperture choice on the COLIBRE GSMF. As described in Section 2.2, subhalo stellar masses in this work are computed by summing the masses of individual stellar particles that (i) are gravitationally bound to their respective subhaloes and (ii) are located within 50 pkpc 3D apertures. The aperture is centred on the particle (stars, gas, DM, or BHs) with the largest (i.e. most negative) gravitational binding energy within each subhalo.

Fig. B1 shows the $z = 0$ GSMF in the COLIBRE L200m6 simulation for different aperture choices. The dark-blue, violet, and orange curves correspond to stellar masses computed using bound stellar particles within 30, 50, or 100 pkpc 3D apertures, respectively, with 50 pkpc being our fiducial choice. The yellow curve represents the case where stellar masses are calculated by summing all bound stellar particles within subhaloes, without imposing an aperture constraint. Unlike in Figs 1 and 2, here we do not add log-normal scatter to galaxy stellar masses to account for Eddington bias, as this would introduce additional scatter in the GSMF between different aperture cases, making their comparison less straightforward.

We find that the COLIBRE GSMF is independent of the aperture choice for $M_* \lesssim 10^{11} M_\odot$. However, stellar masses of the most massive objects in the volume ($M_* \gtrsim 10^{11.5} M_\odot$) can vary by up to ≈ 0.5 dex depending on the aperture, with the larger apertures resulting in higher M_* . These results are consistent with a similar analysis done for the EAGLE simulations (see fig. 6 in J. Schaye et al. 2015).

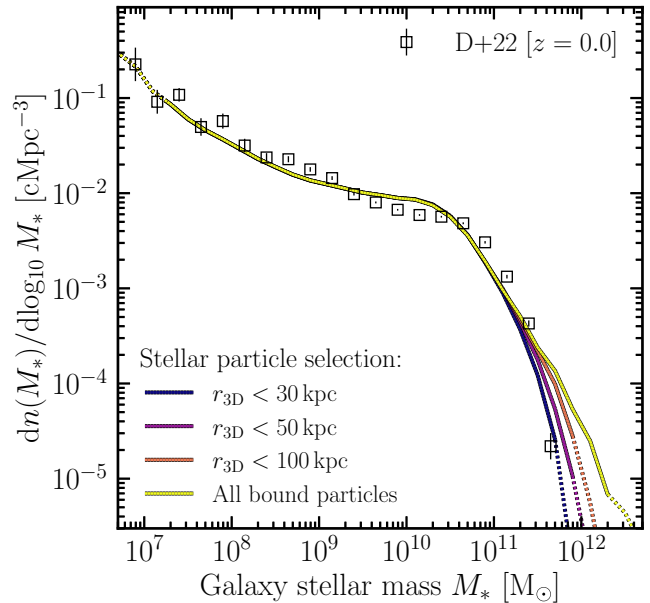


Figure B1. The $z = 0$ GSMF in the COLIBRE L200m6 simulation for four different aperture choices used to calculate subhalo stellar masses. The dark-blue, violet, and orange curves correspond to 3D spherical apertures with radii of 30, 50, and 100 pkpc, respectively, while the yellow curve represents the case where no aperture cut is applied. In all cases, stellar masses are computed using only stellar particles gravitationally bound to subhaloes. Unlike in previous figures, here we do not add log-normal scatter to galaxy stellar masses to account for Eddington bias. For reference, the observed $z = 0$ GSMF from S. P. Driver et al. (2022) is shown as black squares. The choice of aperture has a negligible impact on the GSMF for stellar masses $M_* \lesssim 10^{11} M_\odot$, but becomes increasingly important at higher M_* .

APPENDIX C: AVERAGE STAR FORMATION RATES

In this work, gas SFRs are calculated as instantaneous SFRs. By contrast, observational measurements typically trace SFRs averaged over longer time-scales: ~ 10 Myr using recombination lines such as $H\alpha$, and ~ 100 Myr from UV continuum emission (e.g. P. Madau & M. Dickinson 2014). In this section, we explore the impact of time-averaging on the predicted SFRs and assess the validity of using instantaneous SFRs in comparisons with observational data.

The left (right) column of Fig. C1 shows the correlation between the SFR averaged over the past 10 (100) Myr of lookback time and the instantaneous SFR. Results are presented for the COLIBRE L200m6 simulation at redshifts $z = 0.1$ (top row), $z = 4$ (middle row), and $z = 9$ (bottom row). All SFRs are computed using gas particles bound to subhaloes and located within 50 kpc of the subhaloes' centres. Solid lines represent the median of the averaged SFRs computed in 0.2-dex bins of instantaneous SFR,

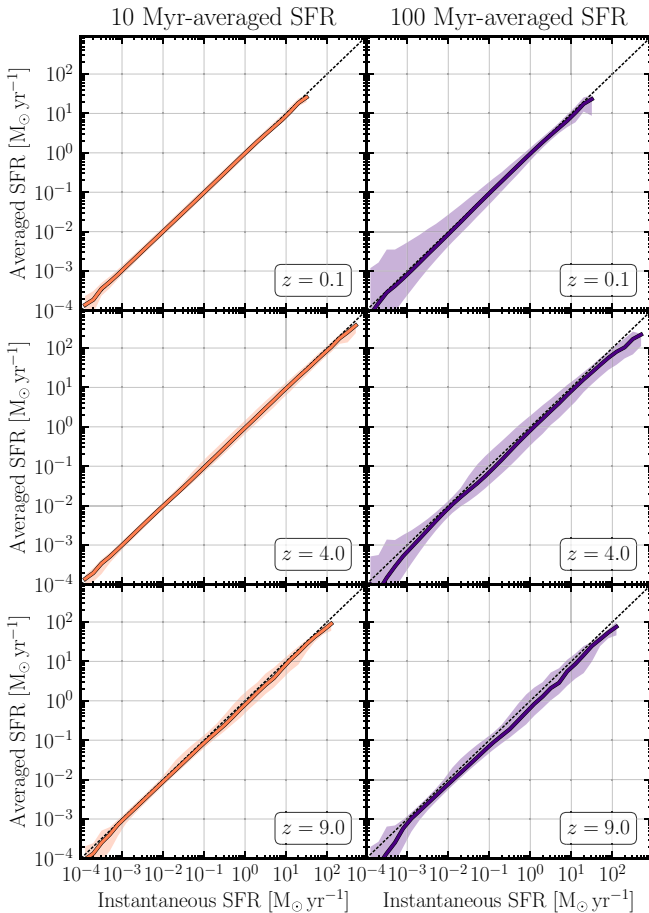


Figure C1. The galaxy SFR averaged over 10 Myr (left column) and 100 Myr (right column) of lookback time, plotted against the instantaneous galaxy SFR in the COLIBRE L200m6 simulation at redshifts $z = 0.1$ (top row), $z = 4$ (middle row), and $z = 9$ (bottom row). Solid lines represent the median relations, while shaded regions denote the 5th to 95th percentile scatter. For reference, the black diagonal line shows the $y = x$ relation. On average, we find little to no systematic deviation between the instantaneous and averaged SFRs, while the scatter increases with the length of the averaging interval.

while shaded regions denote the 5th to 95th percentile scatter. For reference, the black diagonal line indicates the $y = x$ relation.

We find no systematic deviations between the instantaneous and averaged SFRs, except in the $z = 9$ bin, where the 100-Myr averaged SFR is marginally lower than the instantaneous value, by up to ≈ 0.1 dex. As expected, the scatter is larger for the 100-Myr time-scale than for the 10-Myr one, since the longer average captures star formation events further in the past. The close agreement among all three SFR definitions justifies our use of instantaneous SFRs throughout the main part of this work.

APPENDIX D: THE EFFECT OF EDDINGTON BIAS

As outlined in Section 2.3, random errors in observationally inferred stellar masses, combined with the exponential decline in the number of galaxies at the high-mass end of the GSMF, lead to Eddington bias, i.e. the dominance of upscatter over downscatter between adjacent galaxy mass bins. Because observed stellar masses inherently carry non-negligible uncertainties, this bias affects all galaxy statistics based on stellar mass – including the GSMF, the SFMS, and the galaxy quenched fraction – unless explicitly corrected for (e.g. D. Obreschkow et al. 2018; J. Leja et al. 2020). In contrast, this bias is absent in statistics derived from raw simulation data, where galaxy stellar masses are precisely known.

To ensure more consistent comparisons between COLIBRE and observations, throughout this work we assumed a log-normal scatter in galaxy stellar masses of $\sigma_{\text{random}}(z) = \min(0.1 + 0.1z, 0.3)$ dex to represent random errors on stellar masses (see equation 1). In this section, we examine in detail how incorporating this scatter affects the GSMF, the SFMS, the galaxy quenched fraction, and the number density of massive quiescent galaxies predicted by COLIBRE.

D1 The effect of Eddington bias on the GSMF, SFMS, and quenched fraction at redshift 0

Fig. D1 presents the $z = 0$ GSMF (top panel), SFMS (middle panel), and quenched fraction (bottom panel) from the COLIBRE m6 simulation in the $(200 \text{ cMpc})^3$ volume. The coloured lines depict scenarios where a log-normal scatter with zero mean and a standard deviation ranging from $\sigma_{\text{random}} = 0.1$ dex (dark blue) to 0.5 dex (yellow) is added to galaxy stellar masses. The grey lines represent the case without added scatter. The black symbols indicate either observational measurements or predictions from semi-empirical models constrained by observations. As in previous figures showing the SFMS and quenched fractions, the SFMS is defined as the median SFR of galaxies with $\text{SFR} > 0.1 \times$ the SFR at the SFMS, while the quenched fraction is defined as the fraction of galaxies with $\text{sSFR} < 0.2/t_{\text{H}}(z)$.

We find that adding log-normal scatter to galaxy stellar masses significantly affects the GSMF and quenched fraction, while the impact on the SFMS is relatively minor. The most pronounced differences occur at $M_* \gtrsim 10^{11} M_{\odot}$ and for $\sigma_{\text{random}} \geq 0.3$ dex. For example, a log-normal scatter with a standard deviation of 0.4 dex increases the quenched fraction at $M_* \sim 3 \times 10^{11} M_{\odot}$ from ≈ 70 to ≈ 90 per cent and raises the GSMF by ≈ 0.4 dex. In contrast, adopting a scatter with $\sigma_{\text{random}} = 0.1$ dex – the value used in COLIBRE at $z = 0$ – yields negligible differences in both the GSMF and quenched fraction.

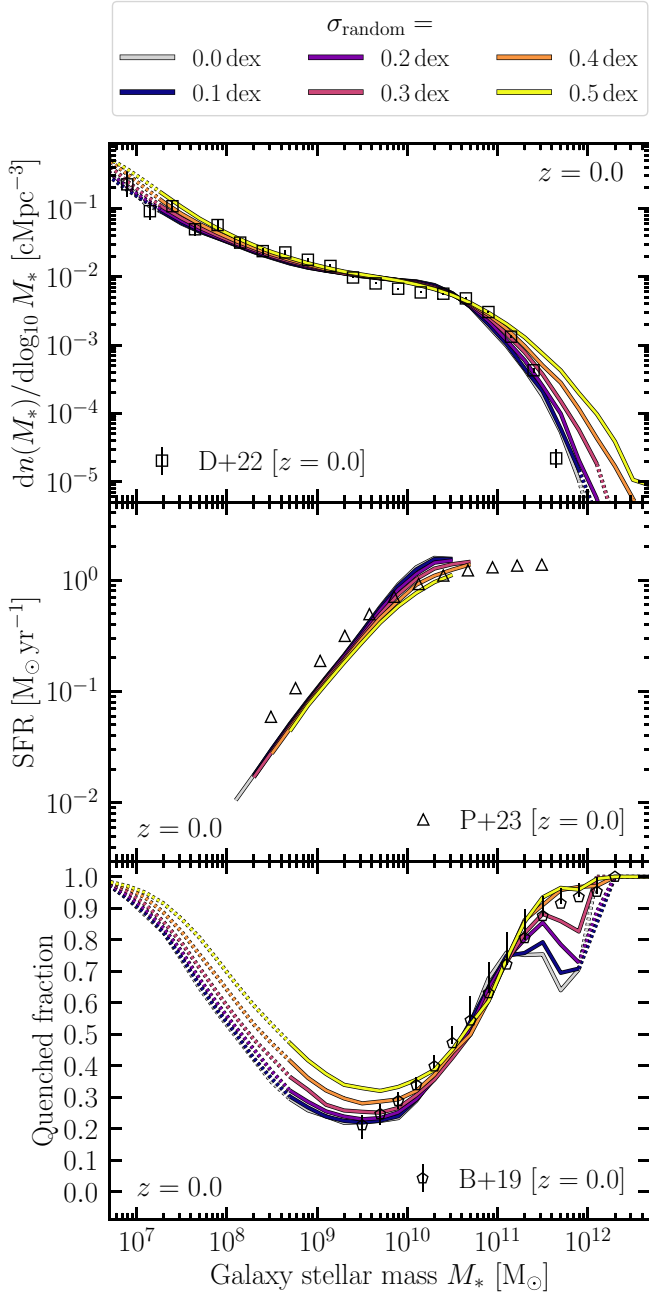


Figure D1. The GSMF (*top*), the SFMS (*middle*), and the galaxy quenched fraction (*bottom*) in the COLIBRE L200m6 simulation at $z = 0$ after adding different amounts of log-normal scatter to galaxy stellar masses to account for Eddington bias (colours). The black symbols in the top, middle, and bottom panels indicate the data from, respectively, S. P. Driver et al. (2022), P. Popesso et al. (2023), and P. Behroozi et al. (2019). The grey-coloured lines correspond to the case with no scatter (i.e. $\sigma_{\text{random}} = 0.0$ dex). Due to the Eddington bias, the scatter affects all three relations, particularly at $M_* \gtrsim 10^{11} M_{\odot}$, where the effect becomes significant for $\sigma_{\text{random}} \geq 0.3$ dex.

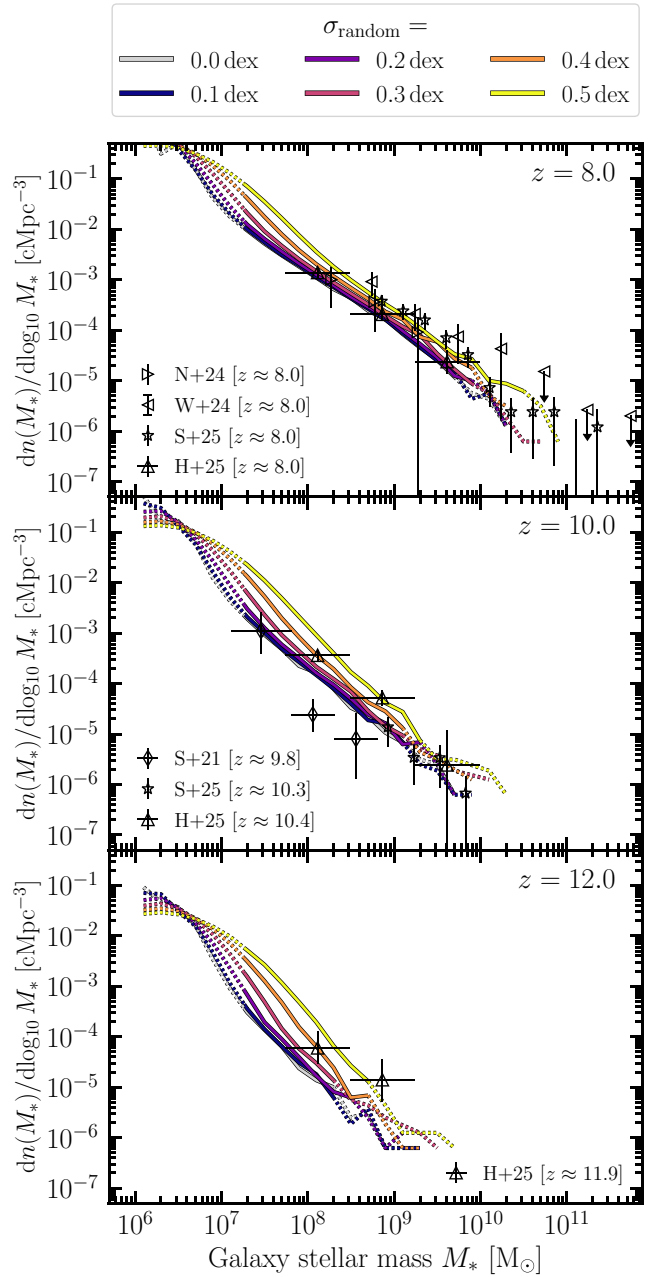


Figure D2. As Fig. D1, but shows only the GSMF, with each panel corresponding to a different redshift: $z = 8$ (*top*), $z = 10$ (*middle*), and $z = 12$ (*bottom*). The observational data (black symbols) are the same as in Fig. 2 at the corresponding redshifts. The Eddington bias is important in bringing the simulation predictions (solid curves) into better agreement with the observational data at the high-mass end, particularly at $z = 12$.

D2 The effect of Eddington bias on the GSMF at high redshifts

Fig. D2 extends the upper panel of Fig. D1 by showing the effect of Eddington bias on the GSMF at three additional redshifts: $z = 8$ (*top*), $z = 10$ (*middle*), and $z = 12$ (*bottom*). Unlike at $z = 0$, the Eddington bias significantly affects the GSMF already above $M_* \sim 10^7 M_{\odot}$, as the GSMF is steeper in this stellar mass range at high redshifts compared to $z = 0$. Adding a log-normal scatter improves the agreement between the simulation

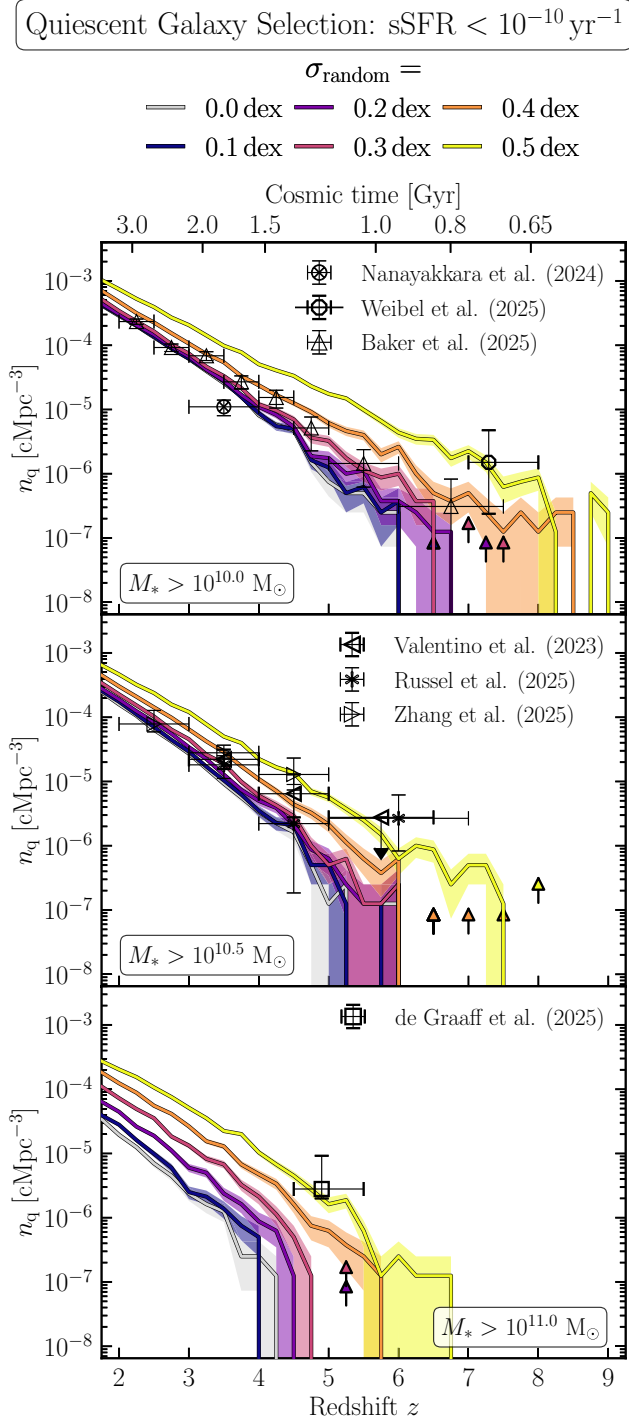


Figure D3. As Fig. D1, but shows the effect of Eddington bias on the comoving number density of massive quiescent galaxies. The layout of the figure is the same as Fig. 11, with different panels corresponding to different stellar mass bins. The Eddington bias plays an important role in bringing the simulation predictions (solid curves) into better agreement with the observational data (black symbols), especially at higher redshifts and in higher stellar mass bins.

and observations at the high-mass end. In particular, adopting $\sigma_{\text{random}} = 0.5$ dex at $z = 12$ brings the COLIBRE GSMF to within 1σ of the value in the higher mass bin reported by T. Harvey et al. (2025).

D3 The effect of Eddington bias on the number density of massive quiescent galaxies

Fig. D3 shows the effect of Eddington bias on the evolution of the comoving number density of massive quiescent galaxies (n_q) in the L200m6 simulation. The figure layout matches Fig. 11, including the observational data (black symbols). As in Fig. 11, different panels correspond to different stellar mass bins for quiescent galaxies contributing to n_q : $M_* > 10^{10} M_\odot$ (top), $M_* > 10^{10.5} M_\odot$ (middle), and $M_* > 10^{11} M_\odot$ (bottom). In the simulation, quiescent galaxies satisfy $s\text{SFR} < 10^{-10} \text{ yr}^{-1}$. Note that while curves in Fig. D3 correspond to fixed σ_{random} values (as indicated in the legend), Fig. 11 uses σ_{random} from equation (1), which depends on redshift but saturates at 0.3 dex for $z \geq 2$. Adding log-normal scatter affects both galaxy stellar masses and their sSFRs, as stellar mass appears in the definition of sSFR.

Fig. D3 demonstrates that the number density of massive quiescent galaxies is highly sensitive to the added log-normal scatter, with larger standard deviations of the scatter generally yielding higher n_q at fixed redshift and stellar mass bin. The effect strengthens at higher redshifts and/or for higher stellar mass bins. For instance, adopting $\sigma_{\text{random}} = 0.5$ dex increases n_q in the $M_* > 10^{10} M_\odot$ and $M_* > 10^{11} M_\odot$ bins by more than an order of magnitude at $z > 4.5$ and $z > 2$, respectively. All observational data shown in Fig. D3 can be matched within $\approx 1\sigma$ for at least one value of σ_{random} between 0 and 0.5 dex.

APPENDIX E: BINNED GALAXY STELLAR MASS FUNCTION FROM COLIBRE

To facilitate comparisons between COLIBRE and other studies, Tables E1 and E2 present the binned GSMF values from the COLIBRE L200m6 simulation with thermal AGN feedback (the orange lines in Figs 1 and 2). The left-most column lists the central values of 0.2 dex stellar mass bins, while the remaining columns provide GSMF values at different redshifts: from $z = 0$ to $z = 7$ in Table E1 and from $z = 8$ to $z = 17$ in Table E2. These values account for Eddington bias as the GSMF is calculated based on galaxy masses including random errors, as described by equation (1).

Table E1. The binned GSMF from the COLIBRE simulation in the $(200 \text{ cMpc})^3$ volume at m6 resolution with thermal AGN feedback. The left-most column shows the central values of 0.2 dex stellar mass bins, while the remaining columns present GSMF values at redshifts from $z = 0$ to $z = 7$. These values account for Eddington bias by the inclusion of log-normal random errors in galaxy stellar masses, as described by equation (1).

$\log_{10}(M_*/M_\odot)$	$\log_{10}(\Phi/[\text{cMpc}^{-3}\text{dex}^{-1}])$									
	$z = 0.0$	$z = 0.1$	$z = 0.5$	$z = 1$	$z = 2$	$z = 3$	$z = 4$	$z = 5$	$z = 6$	$z = 7$
6.5	-0.28	-0.26	-0.18	-0.05	0.19	0.29	0.28	0.19	0.06	-0.09
6.7	-0.53	-0.51	-0.43	-0.33	-0.08	0.01	-0.01	-0.09	-0.21	-0.36
6.9	-0.73	-0.71	-0.64	-0.55	-0.35	-0.30	-0.34	-0.44	-0.56	-0.71
7.1	-0.90	-0.88	-0.82	-0.74	-0.59	-0.58	-0.65	-0.77	-0.92	-1.09
7.3	-1.06	-1.04	-0.98	-0.91	-0.79	-0.81	-0.90	-1.06	-1.24	-1.45
7.5	-1.21	-1.19	-1.13	-1.06	-0.97	-1.00	-1.12	-1.30	-1.52	-1.77
7.7	-1.33	-1.32	-1.26	-1.20	-1.12	-1.18	-1.31	-1.51	-1.75	-2.03
7.9	-1.43	-1.42	-1.37	-1.32	-1.27	-1.34	-1.48	-1.70	-1.96	-2.26
8.1	-1.53	-1.52	-1.48	-1.44	-1.40	-1.48	-1.65	-1.87	-2.14	-2.46
8.3	-1.63	-1.62	-1.58	-1.55	-1.53	-1.63	-1.79	-2.03	-2.33	-2.67
8.5	-1.72	-1.71	-1.68	-1.66	-1.65	-1.74	-1.92	-2.17	-2.48	-2.86
8.7	-1.80	-1.79	-1.77	-1.75	-1.74	-1.85	-2.05	-2.32	-2.67	-3.03
8.9	-1.86	-1.85	-1.83	-1.82	-1.82	-1.94	-2.17	-2.47	-2.84	-3.25
9.1	-1.91	-1.90	-1.88	-1.86	-1.89	-2.04	-2.31	-2.64	-3.04	-3.44
9.3	-1.94	-1.95	-1.91	-1.91	-1.96	-2.16	-2.46	-2.81	-3.22	-3.67
9.5	-1.99	-1.98	-1.96	-1.96	-2.05	-2.28	-2.62	-3.00	-3.41	-3.92
9.7	-2.02	-2.02	-2.00	-2.01	-2.15	-2.43	-2.76	-3.18	-3.61	-4.17
9.9	-2.05	-2.05	-2.05	-2.09	-2.26	-2.56	-2.94	-3.37	-3.85	-4.26
10.1	-2.07	-2.08	-2.12	-2.17	-2.38	-2.70	-3.10	-3.58	-4.04	-4.65
10.3	-2.13	-2.13	-2.19	-2.26	-2.51	-2.87	-3.30	-3.76	-4.28	-4.97
10.5	-2.24	-2.25	-2.29	-2.38	-2.65	-3.05	-3.53	-4.01	-4.62	-5.20
10.7	-2.43	-2.44	-2.46	-2.54	-2.85	-3.29	-3.77	-4.31	-4.90	-5.73
10.9	-2.70	-2.70	-2.72	-2.81	-3.09	-3.55	-4.10	-4.67	-5.36	-
11.1	-3.00	-2.99	-3.05	-3.16	-3.43	-3.93	-4.51	-4.88	-	-5.90
11.3	-3.32	-3.35	-3.44	-3.58	-3.81	-4.38	-4.88	-5.43	-6.20	-
11.5	-3.72	-3.73	-3.87	-4.07	-4.22	-4.88	-5.73	-6.20	-	-
11.7	-4.23	-4.26	-4.45	-4.66	-4.81	-5.36	-6.20	-	-	-
11.9	-4.82	-4.88	-5.36	-5.30	-5.36	-6.20	-6.20	-	-	-
12.1	-6.20	-5.73	-6.20	-5.90	-6.20	-6.20	-	-	-	-

Table E2. As Table E1, but at redshifts ranging from $z = 8$ to $z = 17$.

$\log_{10}(M_*/M_\odot)$	$\log_{10}(\Phi/[\text{cMpc}^{-3}\text{dex}^{-1}])$									
	$z = 8$	$z = 9$	$z = 10$	$z = 11$	$z = 12$	$z = 13$	$z = 14$	$z = 15$	$z = 16$	$z = 17$
6.5	-0.30	-0.53	-0.80	-1.11	-1.45	-1.83	-2.24	-2.69	-3.18	-3.68
6.7	-0.55	-0.77	-1.04	-1.33	-1.67	-2.03	-2.44	-2.90	-3.36	-3.89
6.9	-0.89	-1.11	-1.37	-1.65	-1.98	-2.32	-2.72	-3.15	-3.66	-4.10
7.1	-1.28	-1.51	-1.76	-2.04	-2.36	-2.68	-3.08	-3.52	-3.98	-4.55
7.3	-1.68	-1.93	-2.21	-2.49	-2.80	-3.13	-3.53	-3.95	-4.42	-4.88
7.5	-2.04	-2.33	-2.64	-2.98	-3.31	-3.64	-3.99	-4.52	-5.09	-5.12
7.7	-2.32	-2.66	-3.03	-3.39	-3.81	-4.10	-4.57	-5.16	-5.30	-
7.9	-2.59	-2.95	-3.39	-3.77	-4.21	-4.62	-5.16	-5.36	-5.90	-
8.1	-2.81	-3.25	-3.64	-4.05	-4.59	-4.97	-5.43	-6.20	-5.90	-6.20
8.3	-3.06	-3.45	-3.91	-4.43	-5.06	-5.73	-6.20	-6.20	-	-
8.5	-3.25	-3.70	-4.22	-4.65	-5.20	-5.90	-6.20	-6.20	-	-6.20
8.7	-3.48	-3.93	-4.42	-4.90	-5.51	-5.73	-	-	-	-
8.9	-3.66	-4.21	-4.74	-5.30	-5.73	-	-	-	-	-
9.1	-3.92	-4.52	-5.03	-5.90	-	-	-6.20	-	-	-
9.3	-4.23	-4.76	-5.25	-5.90	-6.20	-	-	-	-	-
9.5	-4.35	-4.93	-5.51	-6.20	-6.20	-	-	-	-	-
9.7	-4.61	-5.36	-6.20	-	-	-	-	-	-	-
9.9	-5.12	-5.51	-6.20	-6.20	-	-	-	-	-	-
10.1	-5.60	-	-	-	-	-	-	-	-	-
10.3	-5.51	-6.20	-	-	-	-	-	-	-	-
10.5	-5.90	-6.20	-	-	-	-	-	-	-	-
10.7	-6.20	-	-	-	-	-	-	-	-	-

APPENDIX F: THE IMPACT OF PARTICLE SELECTION ON THE CSMD AND CSFRD

Here we examine how different particle selection criteria affect the CSFRD and CSMD predicted by the simulations. Fig. F1 presents the evolution of CSFRD and CSMD in the COLIBRE L200m6 simulation. The left- and right-hand panels display, respectively, the CSFRD and CSMD computed using the following criteria:

- (i) CSFRD and CSMD based on all star-forming gas particles and all stellar particles in the simulated volume, respectively (solid curves);
- (ii) CSFRD and CSMD based on only those particles that are bound to subhaloes (short-dashed curves);
- (iii) As (ii) but the particles (gas for CSFRD and stars for CSMD) are additionally restricted to lie within 50 pkpc 3D apertures (long-dashed curves);
- (iv) As (iii) but only subhaloes with stellar mass $M_* \geq 10^7 M_\odot$ contribute to the CSFRD and CSMD (dotted curves);

- (v) As (iii) but only subhaloes with stellar mass of $M_* \geq 10^8 M_\odot$ contribute to the CSFRD and CSMD (dash-dotted curves).

We note that of these, criteria (i) and (iv) were used in Fig. 5, and criteria (i) and (v) in Fig. 4.

Fig. F1 shows that different particle selection criteria influence the CSFRD and CSMD in various ways. First, the difference between the CSFRD (CSMD) computed using all gas (stellar) particles and that based only on bound gas (stellar) particles remains negligible across all redshifts. Secondly, restricting gas (stellar) particles to lie within 50 pkpc apertures reduces the CSFRD (CSMD) by ≈ 10 percent at $z = 0$, compared to using all bound particles. This difference diminishes with increasing redshift and becomes negligible at $z > 3$ for both the CSFRD and CSMD. Finally, restricting the CSFRD and CSMD to subhaloes with $M_* \geq 10^7 M_\odot$ and $M_* \geq 10^8 M_\odot$, which is relevant for comparisons with high-redshift observations, has little impact at $0 < z < 3$, but becomes increasingly significant at higher redshifts.

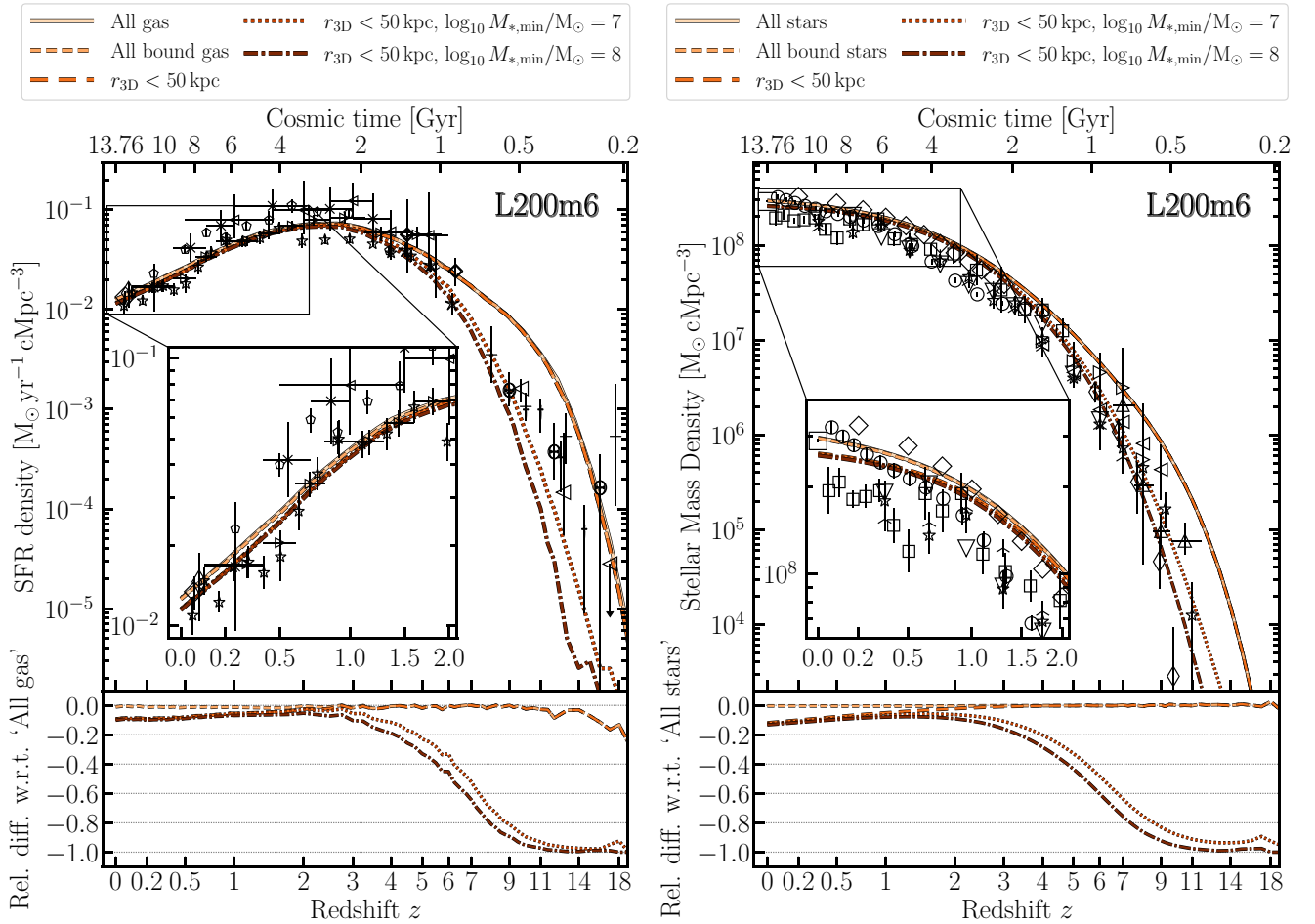


Figure F1. Evolution of the cosmic star formation rate density (CSFRD; *left*) and CSMD (*right*) in the COLIBRE L200m6 simulation, calculated using different particle selection criteria (lines with varying styles and colours). The CSFRD (CSMD) based on all gas (stellar) particles in the simulated volume is shown by the solid lines, while the short-dashed lines indicate the selection of all bound gas (stellar) particles. The long-dashed, dotted, and dash-dotted lines correspond to selections that restrict the calculation of the CSFRD and CSMD to particles that are both bound and within 50 pkpc 3D apertures, with the dotted and dash-dotted lines further limiting the calculations to subhaloes with $M_* \geq 10^7 M_\odot$ and $M_* \geq 10^8 M_\odot$, respectively. The layouts of the left- and right-hand panels match those in Figs 5 and 4, respectively, including the same set of comparison data (black symbols). Additionally, an inset in the bottom left corner shows a zoom-in on the evolution of CSFRD and CSMD at $0 < z < 2$. The restrictions to subhaloes with $M_* \geq 10^7 M_\odot$ and $M_* \geq 10^8 M_\odot$ result in significant, though similar, differences at $z > 4$, while the 50 pkpc aperture constraint has a small but non-negligible impact at $z < 2$.

The choice of minimum stellar mass threshold – $M_* \geq 10^7 M_\odot$ versus $M_* \geq 10^8 M_\odot$ – has a negligible effect below $z = 4$, but becomes progressively more important at higher z : with the differences in the CSFRD (CSMD) increasing from ≈ 0.06 (≈ 0.07) dex at $z = 5$ to ≈ 0.5 (≈ 0.6) dex by $z = 12$.

APPENDIX G: STELLAR MASS DENSITY FROM SUBHALOES IN DIFFERENT MASS BINS

Fig. G1 shows the CSMD in the L200m6 simulation, decomposed into contributions from subhaloes in different mass bins. It follows the same structure as Fig. 6, but displays the CSMD instead of the CSFRD. The left (right) panel shows the decomposition by stellar (halo) mass. The CSMD in each mass bin is computed by summing the masses of individual stellar particles that are gravitationally bound to subhaloes and located within 50 pkpc.

The $0 < z < 3$ CSMD is dominated by $M_* \sim 10^{10}$ - $10^{11} M_\odot$ galaxies. From $z = 3$ to 5, $M_* \sim 10^9 M_\odot$ galaxies also make a significant contribution to the total CSMD, and at $z > 5$, the CSMD is dominated by the least massive objects ($M_* \lesssim 10^{6.5} M_\odot$, comparable to the mass of a few stellar particles at m6 resolution).

By contrast, the galaxies with $M_* \gtrsim 10^{11.5} M_\odot$ never dominate the CSMD.

In terms of halo mass, subhaloes near the peak of the SHMR, with $M_{\text{halo}} \sim 10^{12} M_\odot$, dominate the CSMD at $0 < z < 3$. Between $z = 3$ and 5, subhaloes with $M_{\text{halo}} \sim 10^{11} M_\odot$ become similarly dominant. At $z > 5$, the CSMD is primarily contributed by $M_{\text{halo}} \sim 10^9 M_\odot$ subhaloes. Objects with $M_{\text{halo}} \lesssim 10^{8.5} M_\odot$ and those with $M_{\text{halo}} \gtrsim 10^{13.5} M_\odot$ never dominate the CSMD.

At $z = 0$, about 50 per cent of the CSMD is found in $10^{10.5} < M_*/M_\odot < 10^{11.5}$ galaxies, ≈ 37 per cent in $10^{9.5} < M_*/M_\odot < 10^{10.5}$ galaxies, $10^{8.5} < M_*/M_\odot < 10^{9.5}$ and $M_* > 10^{11.5} M_\odot$ galaxies contribute ≈ 6 per cent and ≈ 5 per cent, respectively, and the remaining ≈ 2 per cent of the CSMD is in $M_* < 10^{8.5} M_\odot$ objects. In terms of halo mass, ≈ 45 per cent of the $z = 0$ CSMD resides in subhaloes with $10^{11.5} < M_{\text{halo}}/M_\odot < 10^{12.5}$, while those with $10^{12.5} < M_{\text{halo}}/M_\odot < 10^{13.5}$ and $10^{10.5} < M_{\text{halo}}/M_\odot < 10^{11.5}$ each contribute about 22 per cent. The most massive subhaloes, in the mass bin $M_{\text{halo}} > 10^{13.5} M_\odot$, contribute only ≈ 7 per cent to the $z = 0$ CSMD. Finally, subhaloes with $10^{9.5} < M_{\text{halo}}/M_\odot < 10^{10.5}$ account for ≈ 4 per cent, while those with $M_{\text{halo}} < 10^{9.5} M_\odot$ contribute the remaining fraction (less than ≈ 1 per cent).

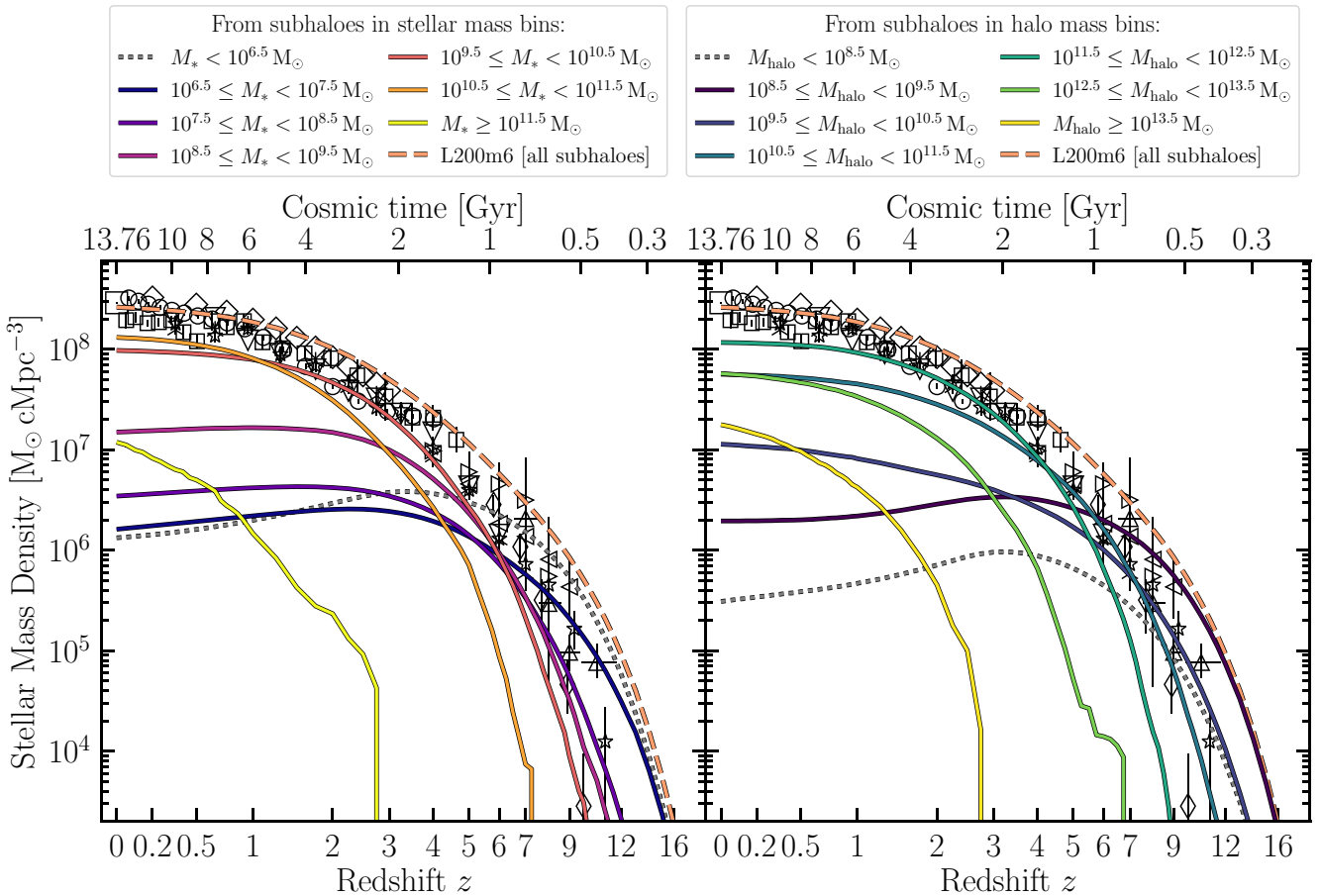


Figure G1. As Fig. 6, but showing the CSMD instead of the cosmic star formation rate density. The observational data (black symbols) are as in Fig. 4. Between $0 < z < 3$, the CSMD is dominated by galaxies with stellar masses of $M_* \sim 10^{10}$ - $10^{11} M_\odot$. At higher redshifts ($3 < z < 5$), galaxies with $M_* \sim 10^9 M_\odot$ begin to contribute significantly, and by $z > 5$, the largest contribution comes from the lowest mass systems ($M_* \sim 10^6$ - $10^7 M_\odot$).

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