

Calibrated, cosmological hydrodynamical simulations with variable IMFs I: method and effect on global galaxy scaling relations

Christopher Barber,¹★ Robert A. Crain² and Joop Schaye¹

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

²Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK

Accepted 2018 July 4. Received 2018 June 25; in original form 2018 April 24

ABSTRACT

The recently inferred variations in the stellar initial mass function (IMF) among local high-mass early-type galaxies may require a reinterpretation of observations of galaxy populations and may have important consequences for the predictions of models of galaxy formation and evolution. We present a new pair of cosmological, hydrodynamical simulations based on the EAGLE model that self-consistently adopts an IMF that, respectively, becomes bottom- or top-heavy in high-pressure environments for individual star-forming gas particles. In such models, the excess stellar mass-to-light (M/L) ratio with respect to a reference IMF is increased due to an overabundance of low-mass dwarf stars or stellar remnants, respectively. Crucially, both pressure-dependent IMFs have been calibrated to reproduce the observed trends of increasing excess M/L with central stellar velocity dispersion (σ_e) in early-type galaxies, while maintaining agreement with the observables used to calibrate the EAGLE model, namely the galaxy luminosity function, half-light radii of late-type galaxies, and black hole masses. We find that while the M/L excess is a good measure of the IMF for low-mass slope variations, it depends strongly on the age of the stellar population for high-mass slope variations. The normalization of the $[Mg/Fe]-\sigma_e$ relation is decreased (increased) for bottom- (top-)heavy IMF variations, while the slope is not strongly affected. Bottom-heavy variations have little impact on galaxy metallicities, half-light radii of early-type galaxies, or star formation rates, while top-heavy variations significantly increase these quantities for high-mass galaxies, leading to tension with observations.

Key words: methods: numerical – stars: luminosity function, mass function – galaxies: elliptical and lenticular, cD – galaxies: fundamental parameters – galaxies: star formation – galaxies: stellar content.

1 INTRODUCTION

The stellar initial mass function (IMF) is a crucial ingredient in the interpretation of galaxy observations as well as for predictions of models of galaxy formation. It defines the translation between physical quantities and observables, and is one of the largest sources of uncertainty in model predictions. In the Milky Way (MW), the IMF seems to be insensitive to environment, with a steep high-mass slope that flattens below $\sim 1 M_\odot$ (Kroupa 2001; Chabrier 2003; Bastian, Covey & Meyer 2010). Observational and theoretical studies alike nearly always adopt such a universal IMF in stellar evolution models, applying it to all galaxies, regardless of the conditions under which their stars were formed.

In the past decade, evidence for variations in the IMF has been steadily mounting, leading to a near-consensus that the IMF be-

comes ‘heavier’ in the regions of high global stellar velocity dispersion, σ , found in the centres of high-mass early-type galaxies (ETGs). In unresolved systems, the IMF is often parametrized by the excess stellar mass-to-light ratio (M/L) of the stars relative to the M/L one would derive spectroscopically assuming a standard IMF. The M/L -excess (hereafter MLE; also known as the ‘IMF mismatch parameter’)¹ is constrained observationally via several independent methods, including gravitational lensing (e.g. Auger et al. 2010; Treu et al. 2010; Spiniello et al. 2011; Barnabè et al. 2013; Posacki et al. 2015; Smith et al. 2015; Sonnenfeld et al. 2015; Collier, Smith & Lucey 2018), stellar population synthesis (SPS) modelling of IMF-sensitive spectral absorption features (e.g. Cenarro et al. 2003; Van Dokkum & Conroy 2010; Conroy & van Dokkum 2012;

¹We introduce the notation ‘MLE’ rather than the more-popular ‘ α ’ for the IMF mismatch parameter to avoid confusion in discussions involving abundances of α -elements.

* Email: cbar@strw.leidenuniv.nl

Spiniello et al. 2012; Ferreras et al. 2013; La Barbera et al. 2013; Spiniello et al. 2014; La Barbera, Ferreras & Vazdekis 2015; Rosani et al. 2018), or dynamical modelling of the stellar kinematics (e.g. Thomas et al. 2011; Dutton, Mendel & Simard 2012; Cappellari et al. 2013; Tortora, Romanowsky & Napolitano 2013; Li et al. 2017), with many of these studies employing a combination thereof. All three methods point to a strong trend of increasing MLE, and thus a ‘heavier’ IMF, with σ . Some studies find additional (and sometimes stronger) trends between the IMF and metallicity and/or α -enhancement, but there is still much debate on this issue (Conroy & van Dokkum 2012; La Barbera et al. 2013; McDermid et al. 2014; La Barbera et al. 2015; Martín-Navarro et al. 2015b).

Puzzlingly, constraints on the IMF seem to be inconsistent on a case-by-case basis. Smith (2014) has shown that for a sample of 34 ETGs, while both SPS and dynamical modelling imply heavier IMFs in high-mass ETGs, there seems to be no correlation between the MLE values derived using the two methods. Newman et al. (2017) compared the MLE derived using lensing, stellar dynamics, and SPS modelling for 3 lenses from the SINFONI Nearby Lens Locator Survey (SNELLS; Smith et al. 2015), also finding inconsistent results between the methods. Conversely, Lyubenova et al. (2016) finds consistent MLE values for SPS and dynamical modelling for a sample of 27 ETGs, arguing that inconsistencies found in other studies may be due to differences in aperture sizes, SPS models employed, or non-optimal dark matter (DM) halo corrections. These findings imply that the systematic errors involved in some of these analyses may not be well understood. Indeed, Clauwens, Schaye & Franx (2015) have shown that IMF trends inferred from stellar kinematics arise also in models assuming a universal IMF if the measurements and/or modelling errors have been underestimated. Furthermore, they found that the data shows an IMF dependence on distance from the Galaxy, suggesting the presence of systematic errors. These results imply that further study is required.

Although the majority of the evidence points towards a non-universal IMF, it is not clear *how* it varies. Dynamical modelling and gravitational lensing constrain only the dynamical M/L , and indicate that it is higher than expected assuming a stellar population with a fixed IMF. This generally implies that either the IMF is more bottom-heavy, leading to more low-mass dwarf stars that contribute significantly to the mass but not the total luminosity, or that the IMF is top-heavy, implying the extra mass comes from stellar remnants: black holes (BHs), neutron stars, and white dwarfs. Some spectroscopic IMF studies are thought to be able to constrain the shape of the low-mass end of the IMF, as a number of absorption features are sensitive to the surface gravity of stars and thus measure the ratio of dwarf-to-giant stars. The majority of these studies find that this ratio is higher in high- σ galaxies, but the means by which this is achieved is similarly unclear, since the increased ratio of dwarf-to-giant stars can be achieved either through a steepening of the IMF low-mass slope (e.g. Conroy & van Dokkum 2012; Conroy, van Dokkum & Villaume 2017), or steepening of the high-mass slope (e.g. La Barbera et al. 2013). On the other hand, $H\alpha$ and $g-r$ colours of local star-forming galaxies from the Galaxy And Mass Assembly (GAMA) survey imply that the high-mass end of the IMF becomes shallower in strongly star-bursting environments (Gunawardhana et al. 2011). The large variety of parametrizations of IMF variations makes comparison between different methods difficult, and has dramatic consequences for the uncertainty in the physical properties of galaxies inferred from observational surveys (Clauwens, Schaye & Franx 2016).

The consequences of a variable IMF on the predictions from galaxy formation models are unclear. While the IMF determines the present-day stellar M/L ratios of galaxies, it also governs the strength of stellar feedback and metal yields. For example a more top-heavy (bottom-heavy) IMF produces more (fewer) high-mass stars that end their lives as supernovae and return mass and energy to the interstellar medium (ISM), affecting the production and distribution of metals throughout the ISM. Metallicity affects the rate at which gas cools and forms future generations of stars, while stellar feedback governs the balance between the flow of gas into, and out of, galaxies, thus regulating star formation. The situation becomes even more complex with the inclusion of supermassive BHs, whose growth depends on the ability of supernova feedback to remove gas from the central regions of galaxies where such BHs reside (Bower et al. 2017). BH gas accretion generates active galactic nucleus (AGN) feedback, which is important for quenching star formation in high-mass galaxies. These processes are non-linear and deeply intertwined, rendering the question of how variations in the IMF would impact galaxies in such models non-trivial.

To address this question, recent studies have begun investigating the effect of IMF variations by post-processing cosmological simulations and semi-analytic models (SAMs), and by conducting self-consistent, small-scale, numerical simulations. In a post-processing analysis of the Illustris simulations, Blancato, Genel & Bryan (2017) study how variations in the IMF of individual star particles manifests as global IMF trends between galaxies, finding that the IMF of individual particles must vary much more strongly than the global trends imply in order to obtain the observed MLE- σ trends. Sonnenfeld, Nipoti & Treu (2017) use an evolutionary model based on dark matter-only numerical simulations to predict the evolution of the IMF in ETGs due to dry mergers from $z = 2$ to 0, finding that dry mergers tend to decrease the MLE of individual galaxies over time, while the correlation between the IMF and σ should remain time-invariant. Much can be learned from post-processing of such large-scale simulations, but such studies by construction neglect the effect that a variable IMF may have on galaxy properties during their formation and evolution due to the change in stellar feedback and metal yields.

IMF variations have also been investigated in SAMs of galaxy formation. Fontanot (2014) find that the variations at the high-mass end of the IMF have a much stronger effect on galaxy properties than variations at the low-mass end. By implementing the ‘integrated galactic IMF theory’ (Kroupa & Weidner 2003), which predicts that the IMF should become top-heavy in galaxies with high star formation rates (SFRs), into SAMs, Fontanot et al. (2017) and Gargiulo et al. (2015) both find that models with a variable IMF are better able to reproduce observed abundance scaling relations than those with a universal IMF. While such SAMs are useful as a computationally inexpensive method of exploring many types of IMF variations, they lack the ability to resolve the internal properties of galaxies, which may be important for IMF studies in light of recent evidence for significant radial gradients in the IMF in individual high-mass ETGs (Martín-Navarro et al. 2015a; La Barbera et al. 2016; van Dokkum et al. 2017; Oldham & Auger 2018; Sarzi et al. 2018; but see Alton, Smith & Lucey 2017; Davis & McDermid 2017).

Hydrodynamical simulations with the ability to resolve the internal structure of galaxies have been run with self-consistent IMF variations for a limited number of idealized galaxies. Bekki (2013) implement a density and metallicity-dependent IMF prescription to idealized chemodynamical simulations of dwarf-to-MW mass

galaxies. They find that a top-heavy IMF in the high-density environments of actively star-forming galaxies suppresses the formation of dense clumps and thus suppresses star formation, as well as increasing the overall metallicities and α -enhancement of such galaxies. Gutcke & Springel (2017) apply the local metallicity-dependent IMF of Martín-Navarro et al. (2015b) to numerical simulations of 6 MW-analogues using AREPO, finding a strong effect on the metallicity evolution in such systems. Guszejnov, Hopkins & Ma (2017) apply various prescriptions of IMF variations from giant molecular cloud (GMC) theory in a simulation of an individual MW analogue galaxy, finding that most prescriptions produce variations within the MW that are much stronger than observed. Such simulations are an excellent starting point to study the effect of IMF variations on galaxy formation and evolution, but are currently limited in statistics, especially for high-mass ETGs where the IMF is observed to vary the strongest.

In this paper, we present a pair of fully cosmological, hydrodynamical simulations, based on the EAGLE project (Crain et al. 2015; Schaye et al. 2015; hereafter referred to as C15 and S15, respectively), each of which includes a prescription for varying the IMF on a per-particle basis to become either bottom-heavy or top-heavy in high-pressure environments, while self-consistently modelling its consequences for feedback and heavy element synthesis. While a pressure-dependent IMF has been studied before using self-consistent, cosmological, and hydrodynamical simulations as part of the OverWhelmingly Large Simulations (OWLS) project (Schaye et al. 2010; Haas et al. 2013), the adopted IMF was in that case not motivated by the recent observations discussed above, and the OWLS models did not agree well with basic observables such as the galaxy luminosity function. In contrast, our prescription has been calibrated to broadly reproduce the observed relationship between the MLE and the central stellar velocity dispersion, and we verify that the simulations maintain good agreement with the observed luminosity function. It is the goal of this paper to investigate the effect that a variable IMF has on the properties of the galaxy population in the EAGLE model of galaxy formation, such as the galaxy stellar mass function (GSMF), luminosity function, SFRs, metallicities, α -enhancement, and sizes. In doing so, we may inform how the IMF should correlate with many galaxy observables, both across the galaxy population as well as within individual galaxies.

This paper is organized as follows. In Section 2 we describe the EAGLE simulations and the calibration of IMF variation prescriptions to match the local empirical MLE- σ correlations, and discuss how these prescriptions are self-consistently incorporated into the EAGLE model. Section 3.1 introduces the variable IMF simulations and details the resulting correlations between the galaxy-averaged IMF and central stellar velocity dispersion. In Section 3.2 we show that IMF variations have little effect on galaxy observables used to calibrate the reference EAGLE model, while Section 4 investigates the impact on predicted galaxy properties such as metallicity, α -enhancement, SFR, and sizes. Our conclusions are summarized in Section 5. Appendix A gives extra details regarding aperture effects and the IMF calibration, while Appendix B shows the effect of incrementally making individual physical processes in the simulations consistent with a variable IMF.

Paper II in this series will discuss trends between the MLE and global galaxy observables and determine which correlate most strongly with the MLE. In Paper III we will discuss the spatially resolved IMF trends within individual high-mass galaxies and the redshift-dependence of the MLE- σ relation. The simulation

data is publicly available at <http://icc.dur.ac.uk/Eagle/database.php> (McAlpine et al. 2016).

2 METHODS

In this section we describe the EAGLE model (Section 2.1) and our procedure of calibrating IMF variations in post-processing to match observed trends with galaxy velocity dispersion (Section 2.2), followed by a description of the modifications to the EAGLE model necessary to produce simulations that are self-consistent when including a variable IMF (Section 2.3).

2.1 The EAGLE simulations

In this study we use the EAGLE model (S15, C15) to study the effect of a variable IMF on predictions of galaxy properties. Here we briefly summarize the simulation model, but refer the reader to S15 for a full description.

EAGLE, short for ‘Evolution and Assembly of GaLaxies and their Environments’, is a suite of hydrodynamical, cosmological simulations aimed at studying the formation and evolution of galaxies from the early universe to $z = 0$. It was run with a modified version of the Tree-PM Smoothed Particle Hydrodynamics (SPH) code Gadget-3, last described by Springel (2005). The modifications to the SPH implementation, collectively known as Anarchy (Schaller et al. 2015, appendix A of S15), improve the treatment of artificial viscosity, time-stepping, and alleviate issues stemming from unphysical surface tension at contact discontinuities. Cubic volumes of up to (100 comoving Mpc)³ were simulated at various resolutions – in this paper we focus only on the ‘intermediate’ resolution simulations, with $m_{\text{gas}} = 1.6 \times 10^6 M_{\odot}$ and $m_{\text{DM}} = 9.7 \times 10^6 M_{\odot}$ for gas and dark matter particles, respectively. The gravitational softening is kept fixed at 2.66 co-moving kpc for $z > 2.8$ and at 0.70 proper kpc at lower redshifts. A Lambda cold dark matter cosmogony is assumed, with cosmological parameters chosen for consistency with Planck 2013: ($\Omega_b = 0.04825$, $\Omega_m = 0.307$, $\Omega_{\Lambda} = 0.693$, $h = 0.6777$; Planck Collaboration I 2014).

Physical processes acting on scales below the resolution limit of the simulation (termed ‘subgrid physics’) are modelled using analytic prescriptions whose inputs are quantities resolved by the simulation. The efficiency of feedback associated with the formation of stars and the growth of BHs was calibrated to match the observed $z = 0.1$ GSMF, galaxy sizes, and the $M_{\text{BH}}-M_{\star}$ relation.

Radiative cooling and photo-heating of gas are implemented element-by-element for the 11 elements most important for these processes, computing their heating and cooling rates via Cloudy assuming a Haardt & Madau (2001) UV and X-ray background (Wiersma, Schaye & Smith 2009a).

Star formation is implemented by converting gas particles into star particles stochastically with a probability proportional to their pressure, such that the simulations reproduce by construction of the empirical Kennicutt–Schmidt relation (Schaye & Dalla Vecchia 2008), renormalized for a Chabrier (2003, hereafter ‘Chabrier’) IMF. For a self-gravitating gaseous disc this star formation law is equivalent to the observed Kennicutt et al. (1998) surface density law. The density threshold for star formation increases with decreasing metallicity according to the model of Schaye (2004) to account for the metallicity dependence of the transition from the warm (i.e. $T \sim 10^4$ K) atomic to the cold ($T \ll 10^4$ K), molecular interstellar gas phase. Once stars are formed, their subsequent mass-loss is computed assuming a Chabrier IMF and the metallicity-dependent

stellar lifetimes of Portinari, Chiosi & Bressan (1998). Heavy element synthesis and mass-loss from stellar evolution are accounted for, including that due to winds from asymptotic giant branch stars and high-mass stars, as well as ejecta from core-collapse and type Ia supernovae (Wiersma et al. 2009b). Stellar feedback is implemented by stochastically injecting a fixed amount of thermal energy into some number of the surrounding gas particles (Dalla Vecchia & Schaye 2012), where the probability of heating depends on the local density and metallicity (S15).

Supermassive BHs are seeded in haloes that reach a Friends of Friends (FoF) mass of $10^{10} M_{\odot} h^{-1}$ by injecting a subgrid seed BH of mass $10^5 M_{\odot} h^{-1}$ into the most bound gas particle (Springel, Di Matteo & Hernquist 2005). BHs grow at the minimum of the Eddington rate and the Bondi & Hoyle (1944) rate for spherically symmetric accretion, taking into account angular momentum of infalling gas (Rosas-Guevara et al. 2015). BHs provide AGN feedback by building up an energy reservoir until they can heat at least one of their nearest neighbours by a minimum temperature, at which point they may stochastically heat their SPH neighbours (Booth & Schaye 2009). This procedure prevents gas from cooling too quickly after being heated, preventing overcooling.

We classify DM haloes using a FoF algorithm with a linking length of 0.2 times the mean interparticle spacing (Davis et al. 1985). Baryons are assigned to the halo (if any) of their nearest DM particle. Self-bound substructures within haloes, termed ‘subhaloes’, are then identified using the SUBFIND algorithm (Springel et al. 2001; Dolag et al. 2009). The ‘central’ subhalo within a halo is defined as the one containing the gas particle most tightly bound to the group, while all others are classified as ‘satellites’. We only consider subhaloes containing at least 100 star particles as resolved ‘galaxies’. For consistency with S15, we define stellar mass, M_* , as the mass of stars within a spherical aperture of radius 30 proper kpc around each galaxy. To compare with observations, we measured all other quantities, such as stellar velocity dispersion (σ_e), M/L , metallicity (Z), and α -enhancement ($[Mg/Fe]$), within a 2D circular aperture with the Sloan Digital Sky Survey (SDSS) r -band projected half-light radius, r_e , of each galaxy, observed along the z -axis of the simulation.

In the next section we discuss how we can use the Reference EAGLE simulations to calibrate a prescription that varies the IMF to match the observed trend between MLE and σ_e .

2.2 IMF calibration

The first goal of this paper is to implement a variable IMF into the EAGLE simulations that yields the observed trends of IMF with galaxy properties. While it is debated how the IMF varies as a function of metallicity or α -abundances, there is mounting evidence that the MLE in the centres of massive elliptical galaxies increases with stellar velocity dispersion (e.g. Treu et al. 2010; Cappellari et al. 2013; La Barbera et al. 2013; Spiniello et al. 2014; Li et al. 2017). This increase could be either due to a higher number of low-luminosity dwarf stars (‘bottom-heavy’ IMF) or stellar remnants (‘top-heavy’ IMF).

We follow Cappellari et al. (2013, hereafter C13) and define the MLE relative to the (M/L) one would obtain assuming a Salpeter IMF:

$$\text{MLE}_i = \log_{10}(M/L_i) - \log_{10}(M/L_i)_{\text{Salp}}, \quad (1)$$

where i denotes the observational filter in which the luminosity is measured. In the upper panel of Fig. 1, in green we plot the observed relation between SDSS r -band MLE_r and stellar velocity

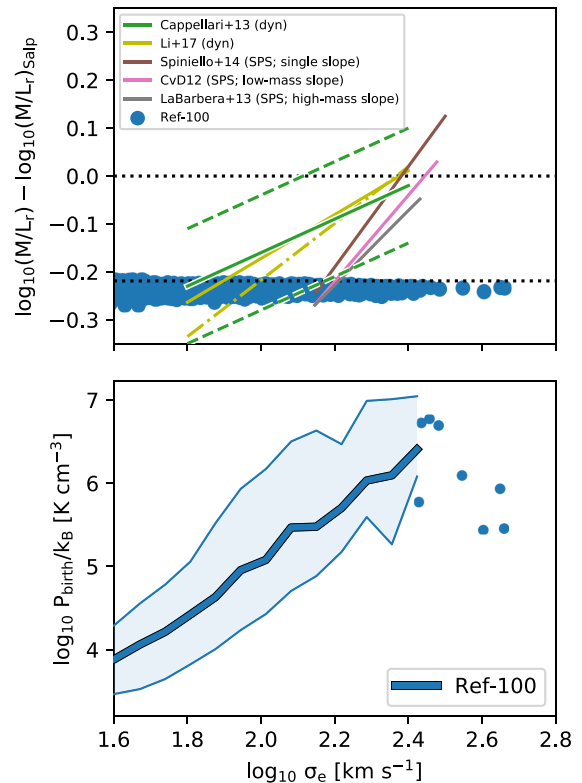


Figure 1. IMF-dependent properties of galaxies in the $(100 \text{ Mpc})^3$ EAGLE reference simulation (Ref-100) at $z = 0.1$. All quantities are measured within the half-light radius, r_e . Top panel: Stellar mass-to-light ratio excess (MLE) as a function of stellar velocity dispersion, σ_e . Horizontal dotted lines at MLE = 0 and -0.22 show the expected MLE for a Salpeter and Chabrier IMF, respectively. The observed trend from Cappellari et al. (2013) is shown as a green solid line, with the intrinsic scatter shown as dashed lines. Also shown are the observed trends from Li et al. (2017); yellow solid and dash-dotted for two different SPS models, respectively), Spiniello et al. (2014; brown solid), Conroy & van Dokkum (2012; pink solid), and La Barbera et al. (2013; grey solid). In brackets we indicate the method of IMF determination, either dynamical or spectroscopic, where for the latter case we also indicate the region of the IMF that is varied in the study. The reference model clearly does not reproduce the observed variation. Bottom panel: Stellar birth ISM pressure as a function of σ_e . The thick and thin lines show the median and 10–90th percentiles in σ_e bins. Where a bin has fewer than 10 galaxies, individual galaxies are shown. It is due to this correlation that we are able to vary the IMF for each individual star-forming gas particle as a function of its pressure in order to achieve a trend in integrated galactic IMF with σ_e , as observed.

dispersion, σ_e , both measured within r_e , obtained by C13 for high-mass elliptical galaxies in the ATLAS^{3D} survey. Also included are observed trends from Conroy & van Dokkum (2012), La Barbera et al. (2013), Spiniello et al. (2014), and Li et al. (2017). Note that for Li et al. (2017) we show the fits for elliptical and lenticular galaxies using two different SPS models to derive $(M/L)_{\text{Salp}}$. For comparison, we also show the same relation for galaxies in the $(100 \text{ Mpc})^3$ reference EAGLE simulation (hereafter Ref-100). As expected, the EAGLE galaxies lie along a line of constant $\text{MLE}_r \approx -0.22$, corresponding to the asymptotic value reached by a stellar population with constant SFR and a Chabrier IMF, and are clearly inconsistent with the observational trends. The goal of this paper is to implement a prescription for an IMF variation that reproduces

the observed C13 relation, and to investigate the effect it has on galaxy properties and observables.

In principle, in order to achieve this correlation, one could simply vary the IMF with the velocity dispersion of the galaxy in which it is born. However, this prescription lacks a physical basis, as there should not be any reason why a star born in a low-mass halo at high redshift should have direct ‘knowledge’ of the stellar velocity dispersion of its host galaxy. Indeed, it would have to know the future velocity dispersion of its host galaxy at $z \approx 0$ at the time it was born, which is infeasible to simulate. A more physical approach is to vary the IMF with respect to some gas property local to a star-forming gas particle at the time it is formed. This affords us the ability to seek connections between physical conditions and $z = 0$ observables, and to perform controlled experiments whereby the various consequences of a variable IMF are selectively enabled/disabled. Moreover, it is a philosophical choice of the EAGLE project to only allow subgrid routines to be ‘driven’ by physically meaningful properties, such as gas density, metallicity, or temperature.

Many physical models of the formation of the IMF on the scales of GMCs predict the IMF to depend on the temperature, density, and/or pressure of the GMC from which the stars form (e.g. Bate & Bonnell 2005; Jappsen et al. 2005; Bate 2009; Krumholz 2011; Hopkins 2012; Hennebelle & Chabrier 2013). One could in principle simply apply these models to star-forming gas particles in the EAGLE simulation using their individual densities and temperatures (as is done in Guszejnov et al. 2017), but it is not clear that such an approach is appropriate here given the much coarser resolution of EAGLE compared to current GMC-scale IMF simulations. Indeed, EAGLE does not resolve the cold phase of the ISM. An alternate approach is to vary the IMF with some parameter of the star-forming gas that is found to vary with stellar velocity dispersion, and attempt to calibrate this local dependence to obtain the observed global IMF–velocity dispersion relationship. One enticing possibility is to vary the IMF with the pressure at which gas particles are converted to star particles in the simulation (Schaye et al. 2010; Haas et al. 2013). Although the cold interstellar gas phase, which EAGLE does not attempt to model, will have very different densities and temperatures than the gas in EAGLE, pressure equilibrium implies that its pressure may be much more similar. However, note that the pressure in the simulation is smoothed on scales of $\sim 10^2 - 10^3$ pc, corresponding to L_{Jeans} of the warm ISM. Note as well that since the local SFR in EAGLE galaxies depends only on pressure, varying the IMF with pressure is equivalent to varying it with local SFR density.

In the lower panel of Fig. 1 we plot the mean r -band light-weighted ISM pressure at which stellar particles within (2D projected) r_e were formed, as a function of σ_e for galaxies in Ref-100 at $z = 0.1$. We see a strong correlation, where stars in galaxies with larger σ_e formed at higher pressures. Thus, by invoking an IMF that varies with birth ISM pressure, we can potentially match the observed $\text{MLE}_r - \sigma_e$ correlation.

To calibrate the IMF pressure-dependence to match the C13 trend, we post-processed the Ref-100 simulation using the flexible SPS (FSPS) software package (Conroy, Gunn & White 2009; Conroy & Gunn 2010). With FSPS, it is possible to generate tables of masses and luminosities in many common observational filters for simple stellar populations (SSPs) as a function of their age, metallicity, and IMF. Here we used the Basel spectral library (Lejeune, Cuisinier & Buser 1997, 1998; Westera et al. 2002) with Padova isochrones (Marigo & Girardi 2007; Marigo et al. 2008), but note that using the different available libraries would not affect our conclusions. Using FSPS in post-processing on Ref-100, star particles were re-

assigned masses and luminosities via interpolation of these tables, given their age, metallicity, initial mass, and birth ISM pressure. As a check, we verified that, for a Chabrier IMF, the SSP masses derived in post-processing using FSPS match the output masses of EAGLE stellar particles computed using the Wiersma et al. (2009b) models built into the simulation to within 2 per cent. However, the agreement between the models is not as good for IMFs with shallow high-mass slopes. Differences in how BH remnants from high-mass stars are treated between the two models result in small differences in mass for a Chabrier-like IMF, but when the high-mass IMF slope is shallow, BH masses begin to become important and these differences are amplified, resulting in ≈ 0.1 dex lower M_* from the Wiersma et al. (2009b) models than with FSPS for high-mass ($M_* > 10^{11} M_\odot$) galaxies with shallow high-mass IMF slopes (applicable to the HiM prescription, below). For consistency, we use stellar masses computed via FSPS for stellar M/L ratios as well as M_* throughout this paper. Note as well that we do not perform radiative transfer to estimate dust extinction. We do not expect dust to be very important here since we investigate mostly old, gas-poor galaxies and measure luminosities in the K or r -band, which are not as strongly affected by dust extinction as bluer wavelengths. However, we do neglect the luminosities of stellar particles with ages younger than 10 Myr, as such stars should still be embedded in their birth clouds, and thus are not expected to be observable (Charlot & Fall 2000).

We define the IMF piecewise as $dn/dM \propto M^x$, such that a Salpeter (1955) IMF has $x = -2.35$ for all M , and a Kroupa (2001, hereafter ‘Kroupa’) IMF has a slope of $x = -1.3$ and -2.3 for stellar masses below and above $0.5 M_\odot$, respectively. Consistent with the EAGLE reference model, we integrate the IMF from 0.1 to $100 M_\odot$.

We began the calibration with the Kroupa IMF since it is practically indistinguishable from the Chabrier IMF over this mass range, but is easier to work with due to its simpler double power-law shape. We tried different methods of varying the IMF, including varying the low-mass slope, high-mass slope, and the stellar mass at which the IMF transitions between these slopes. Varying only the transition mass to make the IMF more bottom-heavy in high-pressure environments (without changing the low-mass cut-off of $0.1 M_\odot$) did not yield a strong enough variation in the IMF to reproduce the observed trends. We briefly experimented with instead increasing the transition mass to make the IMF top-heavy in high-pressure environments (e.g. Fontanot et al. 2018), and found similar results to our ‘HiM’ prescription, outlined below.

We chose to vary the IMF with pressure according to two different prescriptions: one in which the low-mass slope is varied while the high-mass slope is kept fixed (hereafter referred to as LoM) and another where the high-mass slope is varied, keeping the low-mass slope fixed (HiM). These IMF prescriptions are depicted in the top row of Fig. 2. In both prescriptions, we vary the IMF slope between two fixed values x_{lowP} and x_{highP} that are asymptotically reached at low and high pressure, respectively, transitioning between them smoothly via a sigmoid function,

$$x = \frac{x_{\text{lowP}} - x_{\text{highP}}}{1 + \exp(2[\log_{10}(P/P_{\text{trans}})])} + x_{\text{highP}}. \quad (2)$$

Here P_{trans} defines the pressure (and thus the typical σ) at which the IMF transitions from light to heavy. We find that in both cases a value $\log_{10}(P_{\text{trans}}/k_B/[K \text{ cm}^{-3}]) = 5$ (corresponding to $\sigma_e \approx 80 \text{ km s}^{-1}$) works well for reproducing the C13 trend.

In the LoM case (top left-hand panel of Fig. 2), the slope from 0.5 to $100 M_\odot$ is kept fixed at $x = -2.3$ (as for a Kroupa IMF) but the low-mass slope (0.1 to $0.5 M_\odot$) is varied from $x_{\text{LoM, lowP}} = 0$

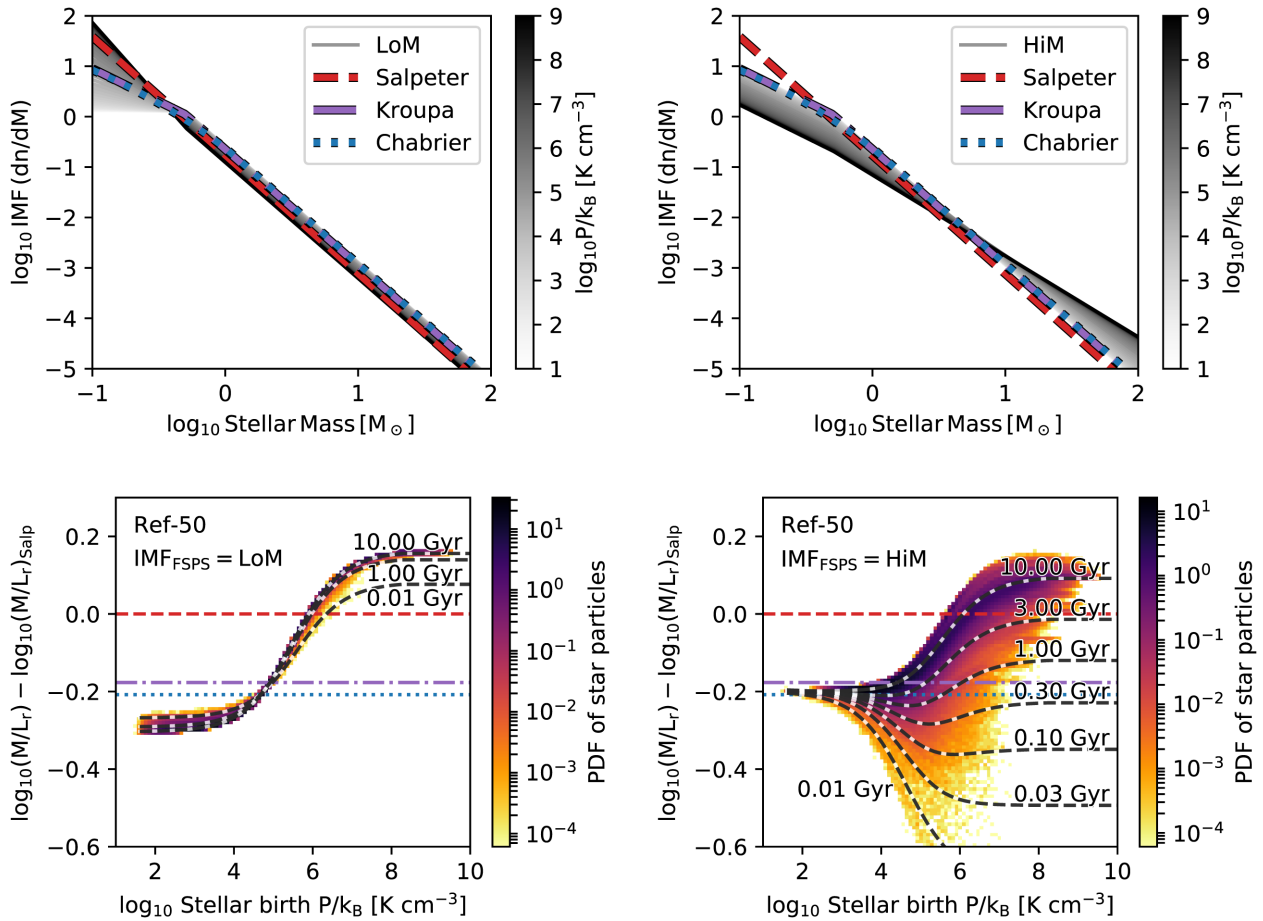


Figure 2. Top row: The two variable IMF prescriptions used in this study shown for a range in stellar birth ISM pressures (see grey-scale bar). Top left: Variable IMF in which the slope below $0.5 M_{\odot}$ is varied (hereafter called LoM) such that the IMF transitions from bottom-light at low P to bottom-heavy at high P . Top right: As in top left but instead varying the IMF slope *above* $0.5 M_{\odot}$ (hereafter HiM) such that it becomes top-heavy at high P . For all IMFs the integrated mass is normalized to $1 M_{\odot}$, causing the low-mass end of the HiM IMF to greatly decrease in normalization at high pressures. Bottom panels: 2D probability distribution functions of the r -band mass-to-light ratio excess (MLE_r) of individual star particles as a function of the pressure of the ISM out of which the star particles formed, for the Ref-50 simulation post-processed assuming LoM and HiM in the bottom left-hand and bottom right-hand panels, respectively. Black dashed lines show the MLE_r - P relation for SSPs at the indicated fixed ages. For reference, in all panels Salpeter, Kroupa, and Chabrier IMFs are shown in red dashed, purple dash-dotted, and blue dotted lines, respectively. Note the small scatter at fixed birth P for LoM, despite the wide range in the ages and metallicities of the stars. This shows that the MLE_r is a good proxy for the IMF when the high-mass slope is close to Salpeter. However, HiM yields a larger scatter in the MLE_r because in this case the MLE_r increases strongly with age at fixed P .

at low pressure to $x_{\text{LoM, highP}} = -3$ at high pressure. Note that this is by no means the only IMF variation prescription that reproduces the C13 trend, especially given the degeneracies between the slopes and the parameters of the sigmoid function, but we find that it is simple, intuitive, and works quite well at producing a clean trend between MLE_r and σ_e .

In the lower left-hand panel of Fig. 2 we plot the resulting MLE_r as a function of birth ISM pressure for individual star particles in Ref-50, post-processed with the LoM IMF prescription. With this IMF, stars born with $P/k_B \lesssim 10^4 \text{ K cm}^{-3}$ are bottom-light, while those with $P/k_B \gtrsim 10^6 \text{ K cm}^{-3}$ are bottom-heavy, with a smooth transition between these values. Such a prescription increases the fraction of dwarf stars in the stellar population at high pressure. This increases the mass and decreases the luminosity of ageing star particles, both leading to an increased MLE_r . Note the small amount of scatter at fixed birth P , despite the fact that stars of all ages and metallicities are plotted here. Thus, for low-mass slope variations, the MLE_r -parameter seems to be a good proxy for the IMF.

For our second variable IMF prescription, HiM (shown in the top right-hand panel of Fig. 2), we instead keep the IMF slope below $0.5 M_{\odot}$ fixed at $x = -1.3$ (the Kroupa value), while making the slope above this mass shallower at high pressures, again varying according to the sigmoid function of equation (2). Specifically, we have $x_{\text{HiM, lowP}} = -2.3$ and $x_{\text{HiM, highP}} = -1.6$, again with $\log_{10}(P_{\text{trans}}/k_B/[K \text{ cm}^{-3}]) = 5$. Similar ‘top-heavy’ forms of IMF variations have been proposed in the literature to explain the observed properties of strongly star-forming galaxies at both high and low redshifts (e.g. Baugh et al. 2005; Meurer et al. 2009; Haberman, Anderson & James 2010; Gunawardhana et al. 2011; Narayanan & Davé 2012, 2013; Zhang et al. 2018).

The lower right-hand panel of Fig. 2 shows MLE_r as a function of birth ISM pressure for individual star particles in Ref-50, this time post-processed assuming the HiM variable IMF. This prescription allows us to increase the MLE_r at high pressure by adding more stellar remnants such as BHs and neutron stars, while at the same time reducing the total luminosity of old stellar populations. Note that here the mass of ageing star particles is overall lower due to the

increased stellar mass-loss associated with the increased fraction of high-mass stars, but the stronger decrease in luminosity results in a net increase in the M/L ratio. Here we see much larger scatter than for LoM due to the fact that the MLE_r for a given star particle is no longer independent of age. The age-independence of the MLE for LoM was solely due to the fact that the high-mass slope is approximately the same as the reference (Salpeter) IMF. In that case, ageing the population removes roughly an equal fraction of mass and luminosity from the LoM IMF as it does from an SSP with a Salpeter IMF. For stars with a shallower high-mass IMF slope, the MLE_r is initially small, owing to the high luminosity, but increases over time as the luminosity decreases faster with age than for a Salpeter IMF. The resulting global correlations between the MLE_r and birth ISM pressure for individual galaxies in self-consistent simulations that include these IMF variations are shown in Appendix A.

These two IMF variation prescriptions were carefully calibrated by post-processing the Ref-100 simulation to reproduce the C13 trend between MLE_r and σ_e . Further details of this calibration procedure can be found in Appendix A. In the next sections we will confirm that this trend is still reproduced when the variable IMF is implemented self-consistently into a full cosmological hydrodynamical simulation.

As an aside, we also experimented with making the IMF become ‘top-light’, meaning that the high-mass slope becomes steeper, rather than shallower, at high pressure. This prescription was inspired by observational studies that infer IMF variations spectroscopically using the MILES SPS models (Vazdekis et al. 2010, 2012), which allow users to vary only the high-mass slope of the IMF, using the (perhaps confusingly nicknamed) ‘bimodal’ IMF of Vazdekis et al. (1996). Such studies find that the fraction of dwarf to giant stars increases with increasing σ in high-mass ETGs, which for this parametrization results in a steeper high-mass slope (or a top-light IMF; e.g. La Barbera et al. 2013, 2015). While we were able to obtain a match to the C13 trend with this bimodal parametrization in post-processing of the reference EAGLE simulations, we opted to use the LoM prescription instead due to the fact that the latter already increases the fraction of dwarf stars with less of an effect on feedback or metal production, making it more likely that the variable IMF model would match the galaxy observables used to originally calibrate the EAGLE reference model. Indeed, it has been shown by Martín-Navarro (2016) that the bimodal IMF prescription can have significant effects on the $[Mg/Fe]$ abundances in massive ETGs. Confirmation of the validity of such a top-light IMF prescription would require a fully self-consistent simulation, which we have not performed for this prescription. It would be interesting for future work to test how well these SPS models can fit IMF-sensitive absorption features using a ‘LoM’ IMF variation parametrization instead.

We also attempted to implement the local metallicity-dependent IMF prescription from Martín-Navarro et al. (2015b), where the high-mass slope of this bimodal IMF is shallower (steeper) than a Kroupa IMF at low (high) metallicities. This prescription was recently used by Clauwens et al. (2016) to reinterpret observational galaxy surveys and was implemented into hydrodynamical simulations of MW-analogues by Gutcke & Springel (2017). In post-processing of Ref-100, we found no clear trend between the MLE and σ_e when implementing this IMF variation prescription. We suspect that this may be partially due to the relatively flat mass–metallicity relation in high-mass galaxies in intermediate-resolution EAGLE (S15).

To provide an idea of the effect of these variable IMF prescriptions on the light output of galaxies, we generate images of galaxies using a modified version of the SKIRT radiative transfer code (Camps & Baes 2015). These modifications allow the user to generate images using spectral energy distribution templates from FSPS, for different variable IMF prescriptions. This new functionality in SKIRT is publicly available in a very general form at <http://www.skirt.ugent.be/>. In particular, it allows the user to specify for each star particle either the low-mass or high-mass slope of the IMF while keeping the other end fixed at the Kroupa value. In this way, one may vary the IMF according to many desired prescriptions, not only those presented in this paper.

We show in Fig. 3 RGB images of the SDSS gri central wavelengths of a massive elliptical galaxy from Ref-50, assuming a Kroupa (left-hand panel), LoM (middle), and HiM (right) IMF. For a Kroupa IMF we see clumps of blue, young star particles embedded in a white, diffuse, intragalactic stellar background. For LoM, the image looks almost identical to the Kroupa image since this IMF mostly adds very dim, low-mass stars to the stellar population, which do not strongly affect the light. On the other hand, in the HiM case the diffuse starlight is much dimmer than the young stars. This is to be expected because for the HiM IMF, older populations should be overall dimmer because a much higher proportion of their mass is invested into the high-mass stars that have since died off. Note, however, that since a top-heavy IMF produces in general more metals per stellar mass formed, the impact of dust on the HiM image is likely underestimated.

2.3 Preparations for self-consistent simulations with a variable IMF

The best way to test the full effect of a variable IMF on simulated galaxies is to run a new simulation that explicitly includes this IMF. This is because, for example the IMF affects the metals released into the ISM by stars, which then affect cooling rates, which further affect future star formation, and so on. Additionally, the IMF affects the available energy from supernovae to provide feedback and regulate star formation. Such effects cannot be accounted for in post-processing. In this section we describe modifications to the EAGLE code that were implemented to maintain self-consistency when adopting a variable IMF.

In EAGLE, the star formation law reproduces the empirical Kennicutt–Schmidt (KS) law (Kennicutt et al. 1998). This relation was originally derived by converting $H\alpha$ fluxes to SFRs assuming a Salpeter IMF. In the reference EAGLE model, S15 accounted for the lower (M/L) obtained from the assumed Chabrier IMF by dividing the normalization of the KS law by a factor of 1.65. This factor is the asymptotic ratio between the number of ionizing photons per solar mass formed after 100 Myr of evolution with a constant SFR as predicted by the Bruzual & Charlot (2003) model for a Chabrier and a Salpeter IMF. Because our IMF is not fixed, but varies with pressure, if we wish to maintain the same relationship between $H\alpha$ surface brightness and Σ_{gas} , we need to instead divide by a factor that is not constant but varies with pressure.

We recalibrate the star formation law by using the FSPS software to compute, for a given pressure, the ratio of the luminosity in the Galaxy Evolution Explorer (GALEX) far-ultraviolet (FUV)-band for a stellar population with a constant SFR, between the variable IMF and a population with a Salpeter IMF, i.e.

$$f_{\text{KS,mod}}(P) = \frac{L_{\text{FUV}}(\text{VarIMF}(P))}{L_{\text{FUV}}(\text{Salpeter})}. \quad (3)$$

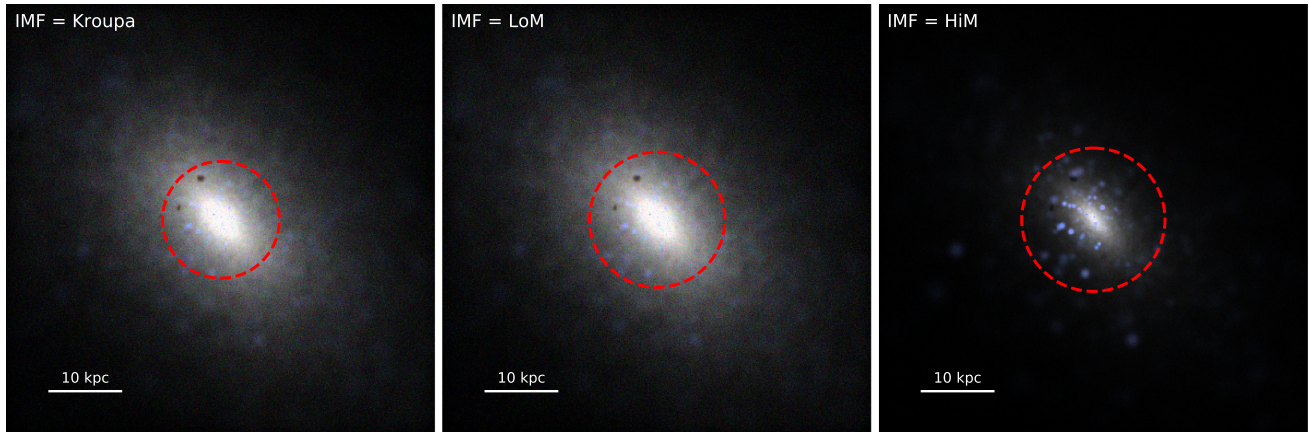


Figure 3. Images of a massive elliptical galaxy in the Ref-50 simulation, post-processed using SKIRT assuming three different variable IMF prescriptions. The images are 60 proper kpc per side and 60 proper kpc deep, centred on the galaxy. From left to right: Kroupa (universal), LoM, and HiM IMF prescriptions are implemented. The 2D projected r -band half-light radius is indicated in each panel as a dashed red circle. RGB colour channels correspond to SDSS g , r , and i peak wavelengths, respectively, normalized using the Lupton et al. (2004) scaling procedure. Assuming the LoM prescription, we produce a nearly identical image to that assuming a Chabrier IMF, while assuming the HiM prescription significantly reduces the luminosity of the diffuse stellar light due to a reduced fraction of low- and intermediate-mass stars, while increasing the fraction of very young stars.

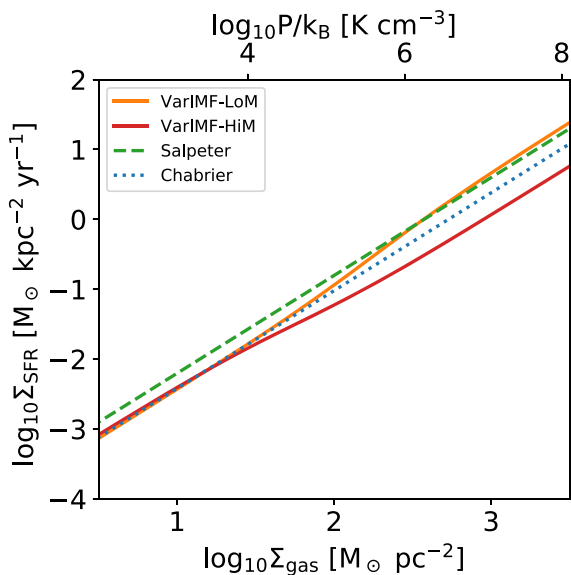


Figure 4. Star-formation law recalibration applied to gas particles in our simulations with a variable IMF. For reference, the laws calibrated for a Chabrier (used in Ref-50) and Salpeter IMF are shown in blue-dotted and green-dashed lines, respectively. The recalibrations that are used in our simulations with LoM and HiM are shown as orange and red solid lines, respectively. The pressure corresponding to the gas surface densities on the lower axis is shown along the top axis. To remain consistent with the observed KS law, at high pressure, SFRs are increased (decreased) by ≈ 0.3 dex relative to the reference simulations for simulations using the LoM (HiM) variable IMF prescription.

In Fig. 4 we plot the recalibrated star formation law as orange and red solid lines for LoM and HiM, respectively, and compare them to the original (Salpeter-derived) relation and EAGLE’s Chabrier IMF-corrected version. We used equation 8 of Schaye & Dalla Vecchia (2008) to convert gas pressure to gas surface density, assuming a gas mass fraction of unity and ratio of specific heats of $5/3$. In the low- P

regime, the normalization remains close to the reference EAGLE value, but at high pressures we multiply (divide) the normalization relative to reference EAGLE by a factor of ≈ 2 for LoM (HiM). Note that for a Chabrier IMF, by the above method we obtain $f_{\text{KS, mod}}(P) \approx 1.57$, not far from the factor 1.65 assumed in EAGLE. The difference here comes from the differences in the FSPS and BC03 models, and has no noticeable effect on our results.

We also make self-consistent the mass evolution of the stellar populations as well as the heavy element synthesis and mass ejected into the ISM from stellar winds and supernovae. This modification is straightforward since these processes already include an integration over the IMF in the EAGLE code.

Another consideration is that the IMF has a direct impact on the number of massive stars and thus the amount of stellar feedback energy that is returned to the ISM per unit stellar mass formed. We also make this self-consistent, which effectively results in a factor ≈ 2 less (more) feedback energy produced per stellar mass formed at high pressures for LoM (HiM). In the reference model, such a large change in the feedback efficiency can have significant effects on many galaxy properties (C15). However, in the case of our variable IMF simulations, the modified star formation law counteracts this effect, making the time-averaged feedback energy consistent with the reference model at fixed gas surface density. We refer the reader to Appendix B for further details regarding the individual impact of each of these effects on galaxy properties. As we will show in Section 3.2, performing variable-IMF simulations with these modifications yields excellent agreement with the observational diagnostics that were originally used to calibrate the subgrid feedback physics in the reference EAGLE model.

The SNIa rate per star particle in EAGLE depends only on the particle’s initial mass and an empirical delay time distribution function, calibrated to match the observed (IMF-independent) evolution of the SNIa rate density (S15). Because of the strong dependence of α -enhancement on SNIa rates, having these rates match observations directly is important. While an IMF-dependent SNIa rate model would be ideal from a theoretical point of view, it is precluded by the large uncertainties in parameters that would factor into such a model,

such as white dwarf binary fractions, binary separations, and merger rates. While the SNIa rates therefore do not depend directly on the IMF, they do depend on the star formation history of the simulation which can be affected by the IMF. We will show in Section 3.2.2 that the SNIa rates are not strongly affected in our variable IMF simulations.

In the next section we will present our simulations and discuss the resulting trend between the galaxy-averaged IMF and central stellar velocity dispersion. We discuss the impact of these variable IMF prescriptions on galaxy properties such as metal abundances and SFR in Section 4.

3 SELF-CONSISTENT SIMULATIONS WITH A VARIABLE IMF

We ran two new (50 Mpc)³ simulations with the same physics and resolution as the reference EAGLE model, except that we imposed two different IMF variation prescriptions. The IMF becomes either bottom-heavy (LoM) or top-heavy (HiM) when the pressure of the ISM out of which star particles are born is high. In Section 2.2 we described how in post-processing we calibrated the pressure dependencies to match the observed trend of excess M/L ratio with stellar velocity dispersion of Cappellari et al. (2013). Including the IMF variation prescriptions explicitly in these simulations allows the IMF variations to affect self-consistently the mass evolution, metal yields, and the stellar energetic feedback during the simulations. The simulations also include a recalibrated KS law normalization to account for the change in UV luminosity per stellar mass formed due to the variable IMF prescriptions (see Section 2.3 for details). Throughout we will refer to these simulations with bottom-heavy and top-heavy IMF prescriptions as LoM-50 and HiM-50, respectively, and the reference (50 Mpc)³ box (with a universal Chabrier IMF) as Ref-50. In Section 3.1 we present the resulting trends between the IMF, MLE_r , and σ_e in high-mass galaxies, while in Section 3.2 we show that both simulations agree well with the observational diagnostics used to calibrate the subgrid physics in the reference EAGLE model. Unless otherwise specified, all quantities are measured within a 2D aperture of radius r_e , the projected half-light radius in the r -band. This choice is motivated by the fact that many IMF studies (e.g. Cappellari et al. 2013) measure the IMF within such an aperture.

3.1 IMF versus stellar velocity dispersion

The trend between the IMF and central stellar velocity dispersion, σ_e , is the most prevalent correlation between the IMF and galaxy properties in the observational literature. In Fig. 5 we plot r -band light-weighted diagnostics related to the IMF as a function of σ_e for galaxies in the LoM-50 (left-hand column) and HiM-50 (right-hand column) simulations. As translucent coloured circles we show all galaxies with $\sigma_e > 10^{1.6}$ (≈ 40) km s⁻¹, coloured by the light-weighted mean pressure at which the stars within each galaxy's r_e were formed.

The upper row shows the r -band light-weighted mean IMF slope for individual galaxies, where in the upper left-hand and right-hand panels we plot the low-mass ($m < 0.5 M_\odot$) and high-mass ($m > 0.5 M_\odot$) slopes, respectively. As expected, for LoM-50 the IMF slope transitions from shallower than Kroupa to a steeper Salpeter-like slope with increasing σ_e , while for HiM-50, the high-mass slope becomes shallower than Salpeter with increasing σ_e , reaching values up to ≈ -1.8 , comparable to the shallowest slopes

inferred in local highly star-forming galaxies (Gunawardhana et al. 2011).

In the middle row we plot the resulting relation between MLE_r and σ_e for our variable IMF simulations. For both simulations, galaxies with $\sigma_e < 10^{1.8}$ (≈ 60) km s⁻¹ lie close to the Chabrier MLE_r value of -0.22 , with MLE_r increasing for higher-mass galaxies. To compare with C13, we select galaxies in a similar way to that study. The C13 sample consists of 260 early-type elliptical and lenticular galaxies selected morphologically based on whether they contain dust lanes or spiral arms, and is complete down to an absolute magnitude of $M_K = -21.5$ mag. We mimic their selection by first taking only galaxies with $M_K < -21.5$ mag (without any dust correction). This cut roughly corresponds to a stellar mass $\gtrsim 10^{10.5} M_\odot$ for all models (although the exact correspondence depends on the IMF assumed). Then, to select only ETGs, we make a cut in intrinsic $u^* - r^* > 2.0$, which roughly separates the blue cloud from the red sequence in EAGLE (Correa et al. 2017) and is similar to removing galaxies with specific SFR (sSFR) $\gtrsim 10^{-1.8}$ and $10^{-1.7}$ Gyr⁻¹ for LoM-50 and HiM-50, respectively. C13 additionally removes galaxies with very young stellar populations by excluding those with an H β stellar absorption line with equivalent width greater than 2.3 Å. McDermid et al. (2015) show that this cut corresponds roughly to an SSP age of 3.1 Gyr, which is already younger than any of our galaxies with $u^* - r^* > 2.0$. We refer to this selection as the ‘mock C13’ sample.

The mock C13 galaxies are highlighted as the opaque coloured circles in Fig. 5. When selecting galaxies in this manner, both simulations produce galaxies reasonably consistent with the C13 $MLE_r - \sigma_e$ trend, with the majority of galaxies lying within the intrinsic scatter.

For LoM-50, a least absolute deviation (LAD) fit to the $MLE_r - \sigma_e$ relation for these mock C13 galaxies yields a slope of 0.23 ± 0.07 , which agrees with the slope of 0.35 ± 0.06 reported by C13. However, our galaxies are offset by ≈ 0.05 dex above the C13 trend. This small discrepancy is partly due to the fact that the LoM prescription was initially calibrated using stars within an aperture larger than r_e , which we show in Appendix A can make a significant difference to the normalization of the $MLE_r - \sigma_e$ relation. Indeed, with a slightly larger choice of aperture, one can decrease the normalization of the LoM-50 $MLE_r - \sigma_e$ trend to match or even lie below the C13 trend. We caution that care with regards to aperture choices should be taken when comparing variable IMF claims between observational studies. Aperture choices vary between observational IMF studies (e.g. Conroy & van Dokkum 2012 and McDermid et al. 2014 measure M/L within $r_e/8$ and r_e , respectively) and even within them (McDermid et al. 2014 measure other properties like age and metallicity within $r_e/8$). Consistent apertures are crucial for making fair comparisons between such studies.

This positive offset is further increased (slightly) due to the fact that stars in LoM-50 tend to form from gas at slightly higher pressures than in Ref-50, which was used for the IMF calibrations. This can be seen in Fig. 6, where we show the distribution of gas birth ISM pressures and ages for stars within r_e of galaxies with $\sigma_e > 100$ km s⁻¹ for the two variable IMF simulations and Ref-50. This result is likely due to the weaker stellar feedback resulting from the more bottom-heavy IMF.

For HiM-50, while nearly all of the mock C13 galaxies lie within one standard deviation of the C13 $MLE_r - \sigma_e$ relation, an LAD fit is slightly shallower for the simulation, with a slope of 0.16 ± 0.09 (compared to the C13 slope of 0.35 ± 0.06). This shallower trend is the result of several factors. First, we are no longer sampling as

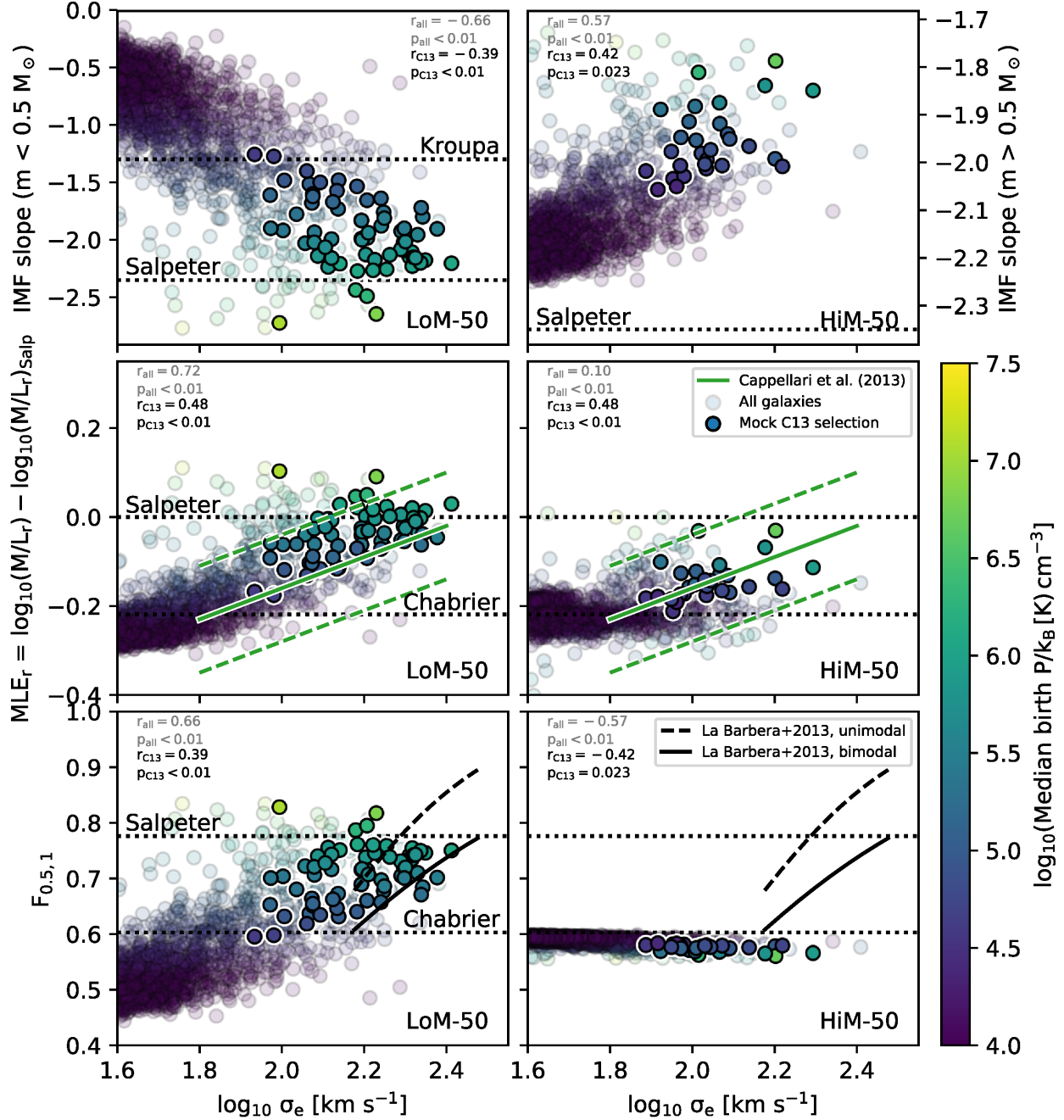


Figure 5. IMF diagnostics as a function of light-weighted stellar velocity dispersion, σ_e , all measured within the 2D, projected, r -band half-light radius, r_e , for all galaxies with $\sigma_e > 10^{1.6} \text{ km s}^{-1}$ in our LoM-50 (left-hand column) and HiM-50 (right-hand column) simulations at $z = 0.1$. Upper row: low-mass ($m < 0.5 M_\odot$) and high-mass ($m > 0.5 M_\odot$) r -band light weighted mean IMF slope for LoM-50 and HiM-50, respectively. Middle row: (M/L_r) -excess with respect to a Salpeter IMF. Lower row: Mass fraction of stars in the IMF with $m < 0.5 M_\odot$ relative to that for stars with $m < 1 M_\odot$, $F_{0.5,1}$ (equation 4). Expected values for fixed IMFs are indicated as dotted horizontal lines. To facilitate comparison with Cappellari et al. (2013, C13), we make a ‘mock C13’ selection of ETGs with $M_K < -21.5$ mag and intrinsic $u^* - r^* > 2$ (C13 cut; see text for details), indicated with opaque filled circles coloured by the light-weighted mean birth ISM pressure; points for galaxies outside this sample are translucent. The observed MLE- σ_e trend from C13 is shown as a green solid line, with the intrinsic scatter shown as dashed lines. Dwarf-to-giant mass fractions derived from the correlations between IMF slope and σ by La Barbera et al. (2013) are shown as black-solid and -dashed lines for bimodal and unimodal IMF parametrizations, respectively. The Pearson correlation coefficient, r , and its p -value are indicated in each panel for the full sample with $\sigma > 10^{1.6} \text{ km s}^{-1}$ and the C13 cut in grey and black, respectively.

many galaxies at high- σ_e ; Ref-50 had poor sampling to begin with for $\sigma_e > 10^{2.3} \text{ km s}^{-1}$ due to the limited simulation volume, which is now compounded by the fact that galaxies at fixed M_{DM} tend to have lower σ_e in HiM-50 than in Ref-50 due to the more efficient feedback.

A second factor is the impact that the stronger stellar feedback from HiM has on the times and gas pressures at which stars form in the simulation. In Fig. 6, we see that for HiM-50, stars in the centres of $\sigma_e > 100 \text{ km s}^{-1}$ galaxies typically form at lower pressures and later times than they did in Ref-50. This behaviour is due to the

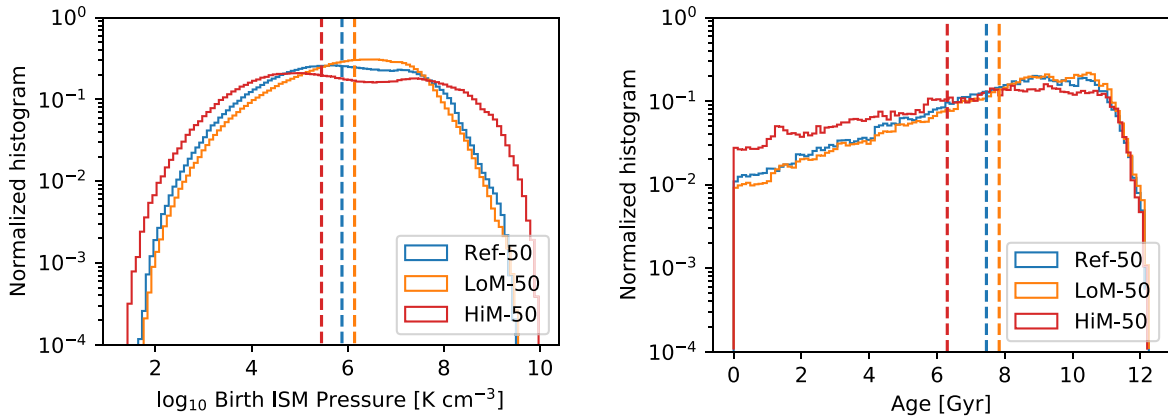


Figure 6. Properties of stars within the half-light radii of galaxies with $\sigma_e > 100 \text{ km s}^{-1}$ in our LoM-50 (orange) and HiM-50 (red) simulations compared with Ref-50 (blue). The left- and right-hand panels show birth ISM pressures and stellar ages, respectively. Vertical dashed lines denote medians. While LoM-50 matches Ref-50 reasonably well, HiM-50 produces stars with lower median birth ISM pressures and younger ages.

stronger stellar feedback delaying star formation to later times (and thus lower pressures). Consequently, galaxies in the simulation obtained IMFs with steeper high-mass slopes than expected, as well as less time to evolve, both of which lower the MLE relative to the post-processing analysis of Ref-50 (although this is not a strong effect for mock C13 galaxies; see Appendix A).

Despite the trend between MLE_r and σ_e being less clear for HiM-50 than LoM-50, the high- σ_e galaxies in HiM-50 are certainly not inconsistent with the C13 trend, and thus represent a conservative approach to studying top-heavy IMF variations in high-mass galaxies. Indeed, we will show in Paper II that this HiM IMF prescription causes the MLE to vary much more strongly with age than with σ_e .

We note that while the $\text{MLE}_r - \sigma_e$ trends for both simulations are consistent with dynamical IMF measurements, HiM-50 may not be consistent with spectroscopic IMF studies that are sensitive to the present-day fraction of dwarf to giant stars, which tend to find an increasing dwarf-to-giant ratio with increasing σ_e (e.g. La Barbera et al. 2013). To show this explicitly, following Clauwens et al. (2016), we compute for each galaxy the fraction of the mass in stars with $m < 0.5 M_\odot$ relative to the mass of stars with $m < 1 M_\odot$ given its IMF. Specifically, we compute

$$F_{0.5,1} = \frac{\int_{0.1}^{0.5} M \Phi(M) dM}{\int_{0.1}^1 M \Phi(M) dM}, \quad (4)$$

where $\Phi(M)$ is the IMF. This upper limit of $m < 1 M_\odot$ in the denominator roughly corresponds to the highest stellar mass expected in the old stellar populations of ETGs. These results are plotted in the lower row of Fig. 5, where, while $f_{0.5,1}$ increases strongly with σ_e for LoM-50 galaxies, it is relatively constant for HiM-50, remaining close to the value expected for a Chabrier IMF. This demonstrates that the increase in MLE_r for high- σ_e galaxies in HiM-50 is the result of excess mass in stellar remnants, rather than dwarf stars.

We compare with the results of La Barbera et al. (2013) by converting their IMF slopes for their 2SSP models with bimodal and unimodal IMF parametrizations to $F_{0.5,1}$. Note that we do not use their definition of $F_{0.5}$ which integrates the denominator in equation (4) to $100 M_\odot$, since $F_{0.5}$ is sensitive to the choice of IMF parametrization at fixed (present-day) $F_{0.5,1}$, which we show explicitly in Appendix C. In the LoM-50 case, $F_{0.5,1}$ agrees very well with the La Barbera et al. (2013) results of increasing mass

fraction of dwarf stars in high- σ_e galaxies. HiM-50, as expected, does not agree.

On the other hand, recall that the HiM prescription was motivated by the fact that highly star-forming galaxies have recently been found to have top-heavy IMFs. For example, Gunawardhana et al. (2011) have found that for local, bright ($M_r < -19.5$) star-forming galaxies, those with larger H α -inferred SFR surface density are redder than expected given their SFR and a standard IMF, implying that the high-mass IMF slope may be shallower in such systems. In Fig. 7, we compare with the results of Gunawardhana et al. (2011) by plotting the GALEX FUV-weighted high-mass slope of the IMF for star-forming (intrinsic $u^* - r^* < 2$) galaxies with $M_r < -19.5$ at $z = 0.1$ in HiM-50 as a function of the SFR surface density, defined as $\Sigma_{\text{SFR,Salp}} = \text{SFR}_{\text{Salp}} / (2\pi r_{e,\text{FUV}}^2)$, where SFR_{Salp} is the Salpeter-reinterpreted total SFR within a 3D aperture of radius 30 kpc and $r_{e,\text{FUV}}$ is the half-light radius in the FUV band. The Salpeter reinterpretation is performed by multiplying the true SFR by the ratio of the FUV flux relative to that expected for a Salpeter IMF, similar to that done by Clauwens et al. (2016). The result from Gunawardhana et al. (2011) is shown for two assumptions for the dust corrections. The positive trend for HiM-50 is qualitatively consistent with the observations, albeit slightly shallower. For reference, we include as a horizontal line the high-mass slope in all LoM-50 (as well as Ref-50) galaxies, corresponding to the Kroupa/Chabrier high-mass value. LoM-50 is, as expected, inconsistent with the observations since the high-mass slope is not varied in that model.

Another example comes from Meurer et al. (2009), who conclude that the increasing ratio of H α to FUV flux towards higher-pressure environments implies that the high-mass slope of the IMF may be becoming shallower in such environments. To compare with their data, we compute the flux of ionizing radiation, f_{ion} , by integrating the spectra output by FSPS up to 912 Å (as in Clauwens et al. 2016) and dividing by the flux in the FUV band, f_{FUV} . Since the ionizing flux is not identical to the H α flux, we normalize by the value of the ratio expected for a Salpeter IMF. In the lower panel of Fig. 7, we plot this ratio for our star-forming galaxies in Ref-50, LoM-50, and HiM-50 as a function of r -band luminosity surface density, Σ_r . We compare with the corresponding relation from Meurer et al. (2009) shown as a solid black line. Ref-50 and LoM-50 show a constant ‘Salpeter’-like $f_{\text{ion}}/f_{\text{FUV}}$ at all Σ_r . For $\Sigma_r > 8 L_{r,\odot} \text{ kpc}^{-2}$, HiM-50 galaxies increase in $f_{\text{ion}}/f_{\text{FUV}}$ with increasing Σ_r , in agreement with

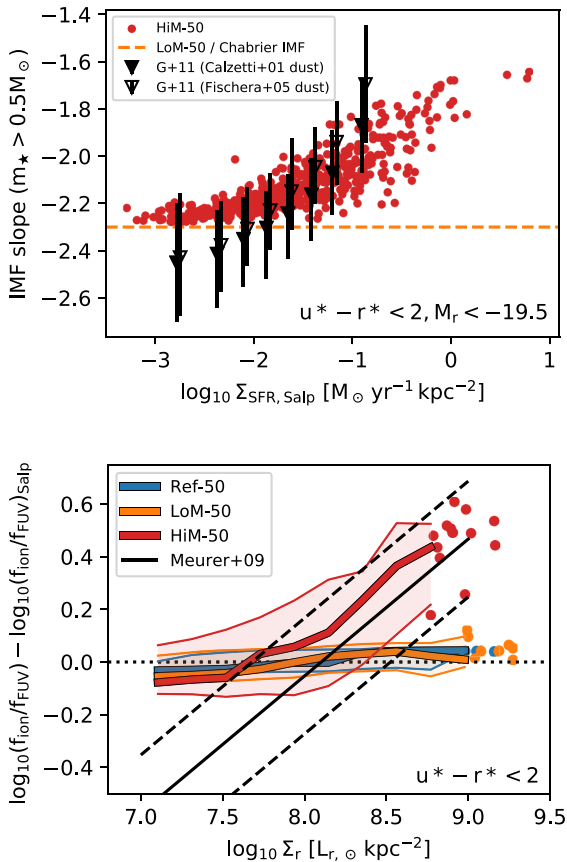


Figure 7. Upper panel: FUV-weighted high-mass ($m > 0.5 M_{\odot}$) IMF slope as a function of the Salpeter-reinterpreted SFR surface density, $\Sigma_{\text{SFR, Salp}}$, of star-forming (intrinsic $u^* - r^* < 2$) galaxies at $z = 0.1$. To compare with Gunawardhana et al. (2011), we only include galaxies also with $M_r < -19.5$. The horizontal orange dashed line shows the high-mass slope for all galaxies in LoM-50, corresponding to the Kroupa/Chabrier value. Open and filled triangles show results from the GAMA survey by Gunawardhana et al. (2011) assuming Calzetti (2001)/Cardelli, Clayton & Mathis (1989) and Fischera & Dopita (2005) dust corrections, respectively. Lower panel: Ratio of flux in ionizing photons to that in the GALEX FUV band, normalized to the value expected for a Salpeter IMF, as a function of r -band luminosity surface density for star-forming galaxies in Ref-50 (blue), LoM-50 (orange), and HiM-50 (red). Solid thick lines indicate medians, while filled regions denote 10–90th percentiles. The black solid line shows results from Meurer et al. (2009) for the flux in $H\alpha$ relative to FUV, normalized to the expected Salpeter value, for star-forming galaxies. Dashed black lines indicate rms residuals. The positive trends seen in observations are qualitatively reproduced for HiM-50 star-forming galaxies, while an IMF prescription that varies only the low-mass slope of the IMF by construction does not.

the observations. At lower Σ_r the relation flattens to the Salpeter value since in HiM we do not vary the IMF high-mass slope to values steeper than Salpeter. This result implies that the IMF high-mass slope may need to become steeper than Salpeter at low pressures to conform with these observations. However, it is not clear that the MLE– σ_e correlation would then still match the observed C13 trend, as this would increase the MLE at low pressure (or low σ_e), weakening the otherwise positive MLE– σ_e correlation.

Since strongly star-forming galaxies are the progenitors of present-day high-mass ETGs, it is not clear how to reconcile the observed top-heavy IMF in strongly star-forming galaxies with the observed bottom-heavy IMF seen in present-day ETGs as con-

strained by the dwarf-to-giant ratio (lower row of Fig. 5). These variable IMF simulations will thus be extremely useful in testing these different, possibly conflicting, IMF variation scenarios with galaxy formation models.

3.2 Subgrid calibration diagnostics

The success of the EAGLE model stems in part from calibrating the subgrid feedback parameters to match key observables (the GSMF, sizes, and BH masses) that are difficult to predict from first principles in cosmological simulations. Thus, a first check to see if the variable IMF simulations are reasonable is to verify that they also reproduce these calibration diagnostics.

3.2.1 Observable diagnostics

Since physical quantities like the GSMF and galaxy half-mass radii can only be derived from observables once an IMF is assumed, we must compare the ‘observable’ versions of the calibration diagnostics with the reference model and observations. The left-hand column of Fig. 8 shows the galaxy K -band luminosity function, 2D projected r -band half-light radii for late-type galaxies (LTGs), and BH masses as a function of luminosity for Ref-50 (blue), LoM-50 (orange), and HiM-50 (red).

We show the $z = 0.1$ K -band luminosity functions of our simulations in the top left-hand panel of Fig. 8, and compare them to observational data from the GAMA survey (Driver et al. 2012) down to $M_K - 5 \log_{10} h = -16$, corresponding to galaxies with ≈ 100 stellar particles, the resolution limit of the GSMF (S15). Both variable IMF runs agree well with Ref-50, with HiM-50 slightly underpredicting the luminosity function relative to Ref-50 by $\lesssim 0.1$ dex at all luminosities. This underprediction is likely caused by the stronger stellar feedback in HiM-50.

In the reference model, the density-dependence of the stellar feedback strength was calibrated to broadly match the observed sizes of LTGs. This calibration was necessary to prevent the formation of overly compact galaxies due to artificial thermal losses in high-density environments due to the limited resolution of the simulation. It is thus important to verify that our variable IMF simulations also reproduce the sizes of such galaxies. In the middle left-hand panel of Fig. 8, we plot the r -band 2D projected half-light radii as a function of M_r for all star-forming galaxies at $z = 0.1$ with at least 600 bound stellar particles, corresponding to $M_{\star} \approx 10^9 M_{\odot}$ or $M_r - 5 \log_{10} h < -17$. It was shown in S15 that galaxy sizes are converged down to this limit. To compare with observed sizes of LTGs, we plot only star-forming galaxies defined as those with intrinsic (dust-free) $u^* - r^* < 2$. The variable IMF simulations match Ref-50 well, with HiM-50 producing larger galaxies by ≈ 0.1 dex, which is less than the typical discrepancies between observational studies of galaxy sizes (e.g. Shen et al. 2003; Baldry et al. 2012). These slightly larger sizes may be caused by the fact that galaxies in HiM-50 typically form at later times, from higher angular momentum gas, giving them larger sizes for the same amount of stellar mass formed. This effect is compounded by the fact that stars in the central regions of these galaxies typically formed at higher pressures than those in the outskirts, and thus with more top-heavy IMFs, yielding dimmer present-day stellar populations and boosting the half-light radii to larger values.

For reference, we include the observed relation between Petrosian r -band half-light radius and r -band absolute magnitude for SDSS galaxies with Sersic indices $n < 2.5$ (Shen et al. 2003). Such indices correspond to surface density profiles typical of discy,

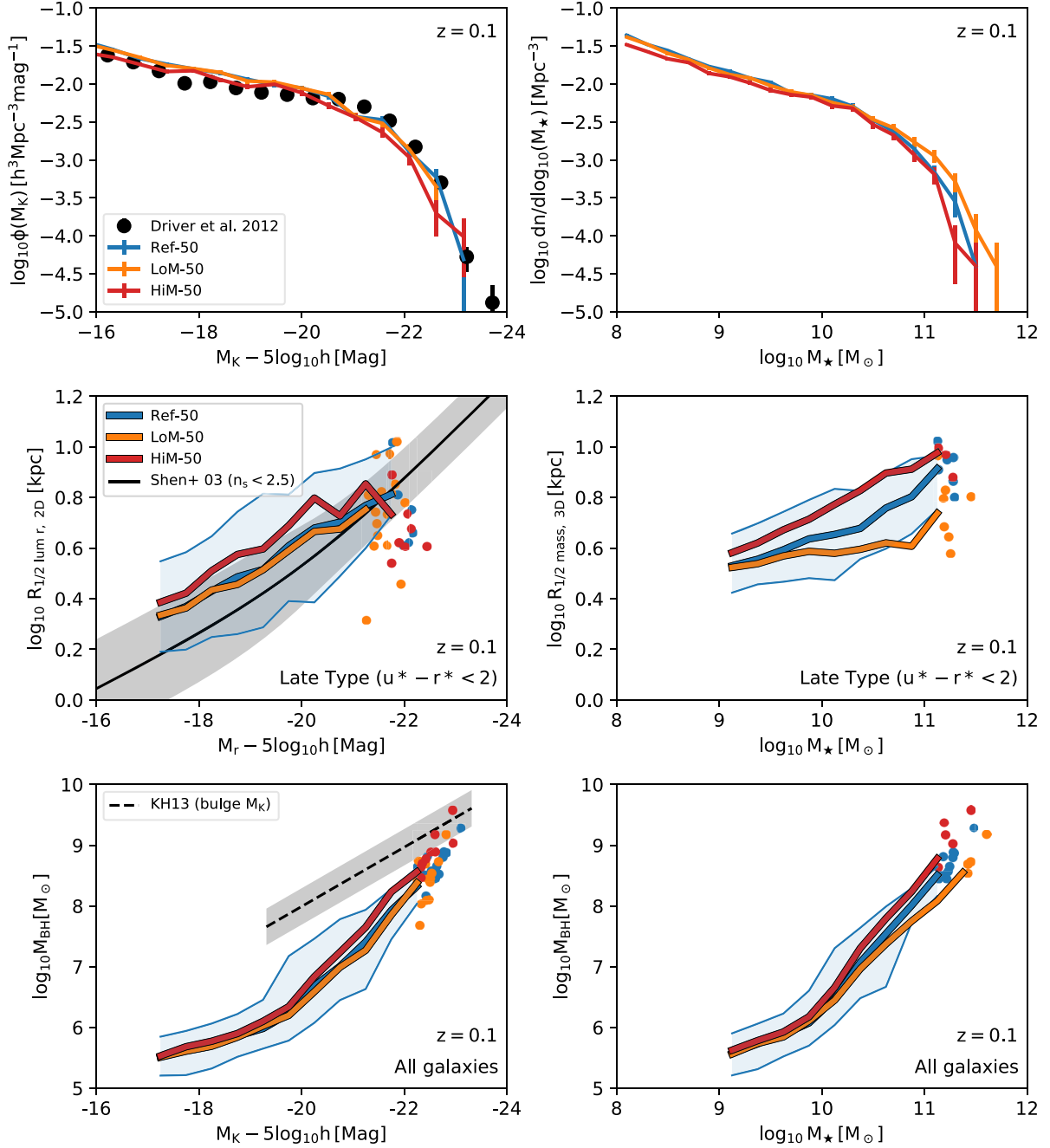


Figure 8. Left-hand column: Subgrid calibration diagnostics for the LoM-50 (orange), HiM-50 (red) simulations, compared to the reference model (Ref-50; blue) and observations for galaxies at $z = 0.1$. To remain consistent with S15, masses and luminosities (sizes) are measured within a 30 kpc 3D (2D) aperture. Top left-hand panel: K -band galaxy luminosity (M_K) function. Observational data from the GAMA survey are shown as black points with 1σ error bars (Driver et al. 2012). Middle left-hand panel: Projected K -band stellar half-light radius as a function of M_K for star-forming galaxies (intrinsic $u^* - r^* < 2.0$) with more than 600 star particles. Filled regions show the 10 to 90th percentile range for Ref-50; the other curves have similar scatter. Individual points are shown for bins containing fewer than 10 galaxies. Sersic r -band half-light radii from SDSS are shown for galaxies with Sersic indices $n_s < 2.5$ as a black solid line (Shen et al. 2003, the grey shaded region indicates 1σ scatter). Lower left-hand panel: As in the middle left-hand panel but showing the BH mass–galaxy M_K relation. For reference we show the observed relation with bulge luminosity (converted to AB magnitudes) from Kormendy & Ho (2013) as a dashed black line with intrinsic scatter indicated with the grey filled region. Right-hand column: Physical quantities corresponding to the subgrid calibration diagnostics shown in the left column. Top right-hand panel: GSMF. Middle right-hand and bottom right-hand panels: 3D half-mass radius and M_{BH} , respectively, as a function of M_{\star} . Galaxies in LoM-50 with $M_{\star} > 10^{10.5} M_{\odot}$ have higher masses at fixed number density and smaller physical sizes at fixed mass than in Ref-50 due to an excess (dark) mass of dwarf stars in the central regions. Both LoM-50 and HiM-50 produce galaxies with reasonably realistic luminosities, half-light radii, and BH masses.

star-forming galaxies. Note that a fairer comparison would be to also select simulated galaxies with $n < 2.5$. However, Shen et al. (2003) showed that the size–luminosity relation is not significantly affected for different reasonable choices of morphological discriminator. Indeed, S15 found that such an $n < 2.5$ cut selects 94 per cent of EAGLE galaxies with more than 600 star particles. We have checked that selecting all galaxies instead of only those with $u^* - r^* < 2$ makes little difference to this plot, so we expect that an $n < 2.5$ cut for the simulated galaxies would not change our result. While LoM-50 and Ref-50 agree relatively well with the data, HiM-50 systematically overpredicts galaxy sizes relative to SDSS by ≈ 0.2 dex, which differs from the observations at the 1σ level. Despite this good agreement for LTGs, we will show in Section 4.4 that the same is not true for ETGs in HiM-50.

The efficiency of AGN feedback was calibrated to match the normalization of the observed $M_{\text{BH}} - M_*$ relation by Booth & Schaye (2009) as part of the OWLS project (Schaye et al. 2010). Since it also gave good results for the much higher-resolution EAGLE simulations, this efficiency was adopted for the reference model. The lower left-hand panel of Fig. 8 shows M_{BH} as a function of M_K . Note that we use the actual M_{BH} values from the simulation rather than attempting to reinterpret them observationally. Both of the variable IMF simulations agree with Ref-50 extremely well for $M_K > -20$, while for more luminous objects there are slight variations of up to ≈ 0.1 and 0.3 dex above the Ref-50 relation for LoM-50 and HiM-50, respectively. These variations are much smaller than the scatter in the observed $M_{\text{BH}} - M_*$ relation and are an acceptable match to Ref-50. Note that the BH masses agree much more poorly if the KS law is recalibrated but feedback is not made self-consistent (see discussion in Section 2.3 and Appendix B).

The observed relation between M_{BH} and bulge K -band luminosity from Kormendy & Ho (2013) for classical bulges and elliptical galaxies is also shown in the lower left-hand panel of Fig. 8. Since we plot total rather than bulge K -band luminosity, we expect our results to fall to the right of the observed relation, which is indeed the case. At the brightest end, where most galaxies are elliptical, all of our simulations converge to the normalization of the observed relation.

We conclude that the luminosities, sizes, and BH masses of galaxies in the variable IMF simulations are reasonable, and match the Ref-50 run quite well. This is encouraging, since it means that we do not need to recalibrate the simulations to obtain an acceptable match to the observed luminosity function, galaxy sizes, and BH masses, even when including self-consistent feedback.

3.2.2 Physical diagnostics

While the observable calibration diagnostics are consistent with Ref-50, the same is not necessarily true for the associated physical quantities. In the right-hand column of Fig. 8, we plot the physical quantities corresponding to the diagnostics in the left-hand column, namely: the GSMF, 3D half-mass radii, and M_{BH} as a function of M_* for galaxies in our variable IMF simulations and Ref-50. We cannot compare with observations here, as these typically assume a universal IMF when deriving physical quantities. While it is possible in principle to reinterpret these physical quantities assuming a Chabrier IMF, thus allowing a comparison with observations, we choose to show the true physical quantities here to highlight the importance of IMF assumptions in the translation between observables and true physical quantities.

The GSMF is shown in the top right-hand panel. At low masses ($\lesssim 10^{10} M_\odot$), the variable IMF runs and Ref-50 match very well, owing to the fact that since these galaxies tend to have lower stellar birth ISM pressures, they form stars with an IMF similar to the Chabrier IMF used in the reference model. At higher masses ($M_* \gtrsim 10^{11} M_\odot$), HiM-50 is consistent with Ref-50 but LoM-50 predicts larger masses at fixed number density by ≈ 0.12 dex. Since the luminosity function in LoM-50 traces Ref-50 nearly perfectly, this difference in their GSMFs is purely due to the increased stellar mass required to produce a fixed luminosity for a bottom-heavy IMF, due to an excess mass of (dim) dwarf stars. Thus, this difference represents the error one incurs when converting K -band luminosity to stellar mass assuming a Chabrier IMF for a galaxy with an intrinsically bottom-heavy IMF. These results are consistent with the recent work of Clauwens et al. (2016) and Bernardi et al. (2018), who investigated the effect that bottom-heavy IMF variations in high-mass ETGs would have on the GSMF derived from SDSS galaxies.

Intrinsic sizes are also significantly affected by a variable IMF. In the middle right-hand panel of Fig. 8 we show the 3D half-mass radii of our galaxies as a function of M_* . As in the middle left-hand panel of Fig. 8, we show only LTGs. At fixed M_* , galaxies in LoM-50 and HiM-50 are smaller and larger, respectively, than Ref-50 by ≈ 0.2 dex at the highest masses. For LoM-50, the smaller physical sizes are due to the excess mass in dwarf stars in the central regions, while high-mass galaxies in HiM-50 are larger due to a deficit of low- and intermediate-mass stars in these regions which is not quite balanced by the excess mass in stellar remnants, as well as the stronger stellar feedback that tends to yield larger galaxies (e.g. Sales et al. 2010; Crain et al. 2015).

The $M_{\text{BH}} - M_*$ relation is plotted in the bottom right-hand panel of Fig. 8. The relation for LoM-50 is shallower than for Ref-50 owing to the increased M_* of the most-massive galaxies. The discrepancy for HiM-50 is similar to that for the $M_{\text{BH}} - M_K$ relation, and is reasonably consistent with Ref-50.

As a check, we reinterpreted M_* and the half-mass radii of these galaxies assuming a Chabrier IMF by multiplying their K -band luminosities by the M/L ratio they would have had if they had evolved with a Chabrier IMF. We refer to this reinterpreted mass as $M_{*, \text{Chab}}$. Doing so puts the masses, sizes, and BH masses of the LoM-50 galaxies into excellent agreement with Ref-50 by decreasing the inferred masses of high-mass galaxies by a factor ≈ 2 , but has little effect for HiM-50 (not shown).

Finally, we investigate the SNIa rate evolution in the simulations. As mentioned in Section 2.3, the predicted rates do not explicitly depend on the IMF, but they are affected by the star formation history of the simulations. Fig. 9 compares the evolution of the SNIa rate density with observations compiled by Graur et al. (2014). The rates are not strongly affected in either variable IMF simulation, where the agreement with observations is about as good as it is for Ref-50. This is encouraging, since a strong deviation from the observed rates would require a recalibration of the empirical time-delay function. All differences here are due to the effect that the IMF has on the star formation history, which is slightly delayed in HiM-50 relative to Ref-50.

4 EFFECT OF VARIABLE IMFS ON GALAXY PROPERTIES

In this section we investigate the effect that our variable IMF prescriptions have on the predicted properties of galaxies in the $(50 \text{ Mpc})^3$ self-consistent variable IMF simulations relative to Ref-50

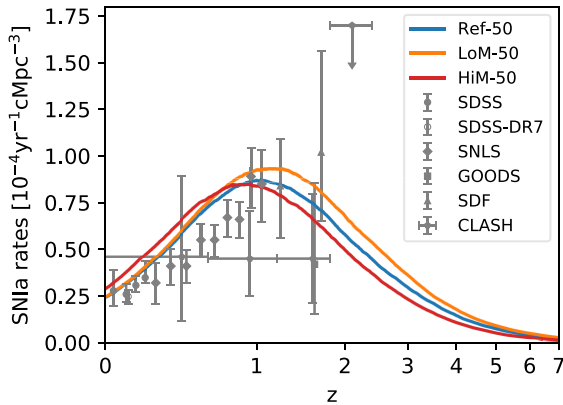


Figure 9. Cosmic supernova Ia rate density as a function of redshift in the LoM-50 (orange), HiM-50 (red), and Ref-50 (blue) simulations. Grey data points show observations compiled by and classified as the ‘most accurate and precise measurements’ by Graur et al. (2014): SDSS Stripe 82 (Dilday et al. 2010), SDSS-DR7 (Graur & Maoz 2013), SNLS (Perrett et al. 2012), GOODS (Dahlen et al. 2008), SDF (Graur et al. 2014), and CLASH (Graur et al. 2014). The 1σ systematic and statistical uncertainties are indicated with error bars. The SNIa rates are not strongly affected in the variable IMF simulations and match observations about as well as the reference model.

and observations. Specifically, Sections 4.1, 4.2, 4.3, and 4.4 discuss the effect on α -enhancement, metal abundances, SFRs, and ETG sizes, respectively. All results are shown for galaxies at $z = 0.1$; properties at higher redshift will be discussed in a future work.

4.1 α -enhancement

One of the strongest effects that the variable IMF has in the simulations is on the abundances of α -elements in high- σ_e galaxies. In the upper panel of Fig. 10 we show $[\text{Mg}/\text{Fe}]$ as a function of σ_e for ETGs (defined as those with $u^* - r^* > 2.0$) in our variable IMF simulations.

Segers et al. (2016) have already shown that the trend of α -enhancement with stellar velocity dispersion in the reference simulations agrees well with observations of quiescent ETGs although the normalization of the relation is too low by about a factor of 2. We note that nucleosynthetic yields are uncertain at the factor of 2 level (e.g. Wiersma et al. 2009b), while uncertainties in stellar population modelling lead to systematic errors in the observed values of $[\alpha/\text{Fe}]$ of ~ 0.1 dex. Thus, the normalization of the trend is not nearly as constraining as the slope. Least absolute deviation fits to these galaxies above $\sigma_e > 10^{1.8} \text{ km s}^{-1}$ yield slopes of 0.33 ± 0.07 , 0.26 ± 0.04 , and 0.40 ± 0.12 for Ref-50, LoM-50, and HiM-50, respectively. While the slopes for Ref-50 and HiM-50 agree with the observed values of 0.33 ± 0.01 (Thomas et al. 2010) and 0.33 ± 0.03 (Conroy, Graves & van Dokkum 2014), that for LoM-50 is somewhat shallower.

We have investigated the cause of the difference in the normalization and slopes of these trends, and find that the culprit is the difference in the metal yields in the different simulations. In high-mass galaxies, LoM-50 produces fewer massive stars per unit stellar mass formed and thus less Mg is synthesized for future generations of stars, leading to reduced $[\text{Mg}/\text{Fe}]$. On the other hand, HiM-50 produces many more massive stars per unit stellar mass formed, increasing the Mg yields. By independently switching on the different effects of the variable IMF in smaller $(25 \text{ Mpc})^3$ boxes, we confirmed that indeed the yields, rather than the feedback or the KS law renormalization, drive these differences. While the differences

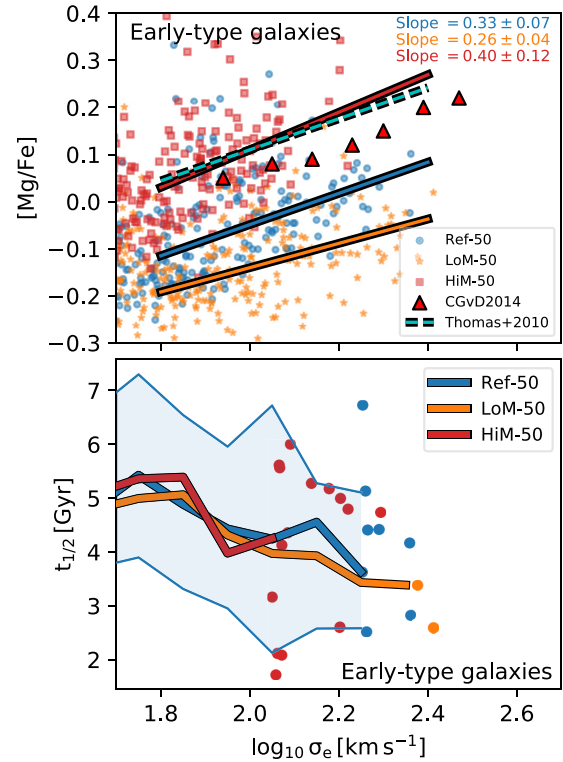


Figure 10. Top panel: Stellar $[\text{Mg}/\text{Fe}]$ as a function of stellar velocity dispersion for galaxies at $z = 0.1$ in LoM-50 (orange stars) and HiM-50 (red squares) compared to Ref-50 (blue circles). All quantities are SDSS r -band light-weighted and measured within the projected half-light radius, r_e . To facilitate a fairer comparison with observations, included are only ETGs ($u^* - r^* > 2.0$). The observed trend for ETGs from Thomas et al. (2010) is shown as a dashed cyan line, while the observed trend for quiescent galaxies from Conroy et al. (2014) is shown as red triangles. Least absolute deviation fits for ETGs with $\sigma_e > 10^{1.8} \text{ km s}^{-1}$ in each simulation are shown as solid coloured lines, with slopes labelled. The $[\text{Mg}/\text{Fe}]$ abundances in LoM-50 (HiM-50) are normalized lower (higher) relative to Ref-50 for high-mass galaxies. While the slopes of the $[\text{Mg}/\text{Fe}]$ – σ_e relation for both Ref-50 and HiM-50 are consistent with observations, that for LoM-50 is somewhat shallower. Bottom panel: Median formation time of stars within r_e for the same galaxy samples as the upper panel. The positive $[\text{Mg}/\text{Fe}]$ – σ_e correlation reflects the star formation histories of ETGs.

in the normalization are not large, it is concerning that the $[\text{Mg}/\text{Fe}]$ values in LoM-50 fall so far below those from observations, especially since, following Portinari et al. (1998), the Mg yields have already been doubled in the reference model with respect to the standard yields.

It is interesting that our IMF variations have such little effect on the slope of the $[\text{Mg}/\text{Fe}]$ – σ_e relation. Recent studies of IMF variations in SAMs have concluded that a top-heavy IMF in rapidly star-forming environments (which occurs in the high-redshift progenitors of present-day high-mass ETGs) may be necessary to produce a positive correlation in this relation, a result of higher Mg abundances due to a larger number of SNI (Gargiulo et al. 2015). In the EAGLE model, a variable IMF is not necessary to reproduce this slope, as the positive trend comes from the quenching of star formation via AGN feedback, preventing much of the Fe from type Ia SNe from being incorporated into future stellar populations (Segers et al. 2016). It is thus encouraging that the slope does not become even steeper with a top-heavy IMF in EAGLE. Although the difference in Mg abundance between HiM-50 and Ref-50 in-

creases with σ_e , the same is true for the Fe abundances, maintaining the slope of the $[\text{Mg}/\text{Fe}] - \sigma_e$ relation.

To help support this quenching scenario, in the lower panel of Fig. 10 we plot the median formation time, $t_{1/2}$, of stars within the 2D projected half-light radius, r_e , of the same ETGs shown in the upper panel. Here we see a trend of decreasing $t_{1/2}$ with increasing σ_e for all three simulations, in qualitative agreement with recent results from the ATLAS^{3D} survey (McDermid et al. 2015). These results support the idea that short star formation histories leads to higher $[\text{Mg}/\text{Fe}]$ ratios in ETGs.

This quenching scenario is also supported by semi-analytic modelling performed by De Lucia, Fontanot & Hirschmann (2017), although they find that abrupt quenching of high-mass galaxies prevents them from reaching high enough metallicities at $z = 0$, requiring a variable IMF to match both the mass–metallicity relation and the α -enhancement of high-mass galaxies simultaneously (see also Arrigoni et al. 2010). We will show in the next section that our HiM-50 simulation does indeed match the slope of the observed mass–metallicity relation better than the reference model.

4.2 Metallicities

We also investigate the effect of the variable IMFs on the stellar mass–metallicity relation where, in the top panel of Fig. 11, we plot the stellar metallicity, Z , measured within r_e , as a function of Chabrier-reinterpreted stellar mass, $M_{\star, \text{Chab}}$, for all galaxies with $M_{\star, \text{Chab}} > 10^8 M_{\odot}$. Interestingly, the total metallicities are largely unchanged in the LoM-50 run relative to Ref-50, with only a slight (< 0.1 dex) decrease in Z for $M_{\star, \text{Chab}} > 10^{11} M_{\odot}$. However, for HiM-50, the mass–metallicity relation is much steeper, with $Z/Z_{\odot} \approx 0.4$ dex higher in high-mass ($M_{\star, \text{Chab}} \sim 10^{11} M_{\odot}$) galaxies. This difference comes from the much higher production of metals per unit stellar mass formed from an IMF with a shallow high-mass slope.

We compare the simulated trends with the observed relation for SDSS galaxies from Gallazzi et al. (2005). For $M_{\star, \text{Chab}} > 10^{10.5} M_{\odot}$, the mass–metallicity relation is observed to flatten off, as is also seen in Ref-50 and LoM-50, albeit at lower mass than observed. HiM-50, on the other hand, shows no sign of a saturating metallicity at high mass, which may be in tension with observations. Zahid et al. (2014) argue that this saturation occurs when the gas-phase abundances are high enough that during star-formation, the metal mass that is removed from the ISM and locked up into low-mass stars is comparable to that produced and released back into the ISM by high-mass stars. HiM-50 keeps increasing in metallicity because for a top-heavy IMF, many more supernovae are produced per dwarf star formed, producing more metals than are being locked up, delaying saturation to higher metallicities. Note, however, that De Rossi et al. (2017) have shown that the mass–metallicity relation in EAGLE saturates in part due to AGN feedback quenching star formation and ejecting metal-rich gas out of high-mass galaxies. This implies that the Zahid et al. (2014) explanation is at least incomplete. Indeed, we will show in Section 4.3 that galaxies with $M_{\star} \gtrsim 10^{11} M_{\odot}$ have typically higher SFRs in HiM-50 than in Ref-50, possibly contributing to their higher metallicities.

Given the large uncertainties in the simulated nucleosynthetic yields (Wiersma et al. 2009b) and the calibration of the metallicity indicators applied to observations (e.g. Kewley & Ellison 2008), the slope of the mass–metallicity relation is more constraining than its normalization. Because our (50 Mpc)³ simulations only contain galaxies up to $M_{\star, \text{Chab}} \sim 10^{11} M_{\odot}$, it is not clear if the lack of flattening at the high-mass end is actually inconsistent with the ob-

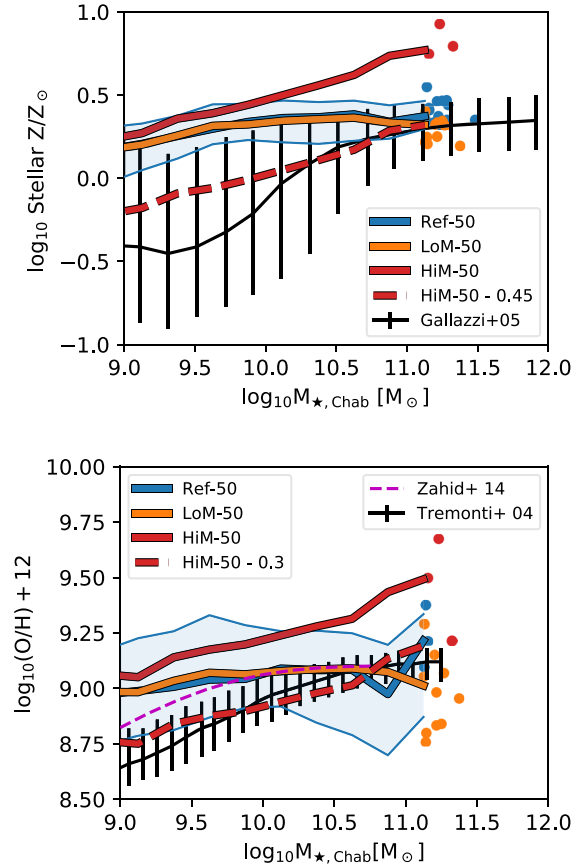


Figure 11. Metal abundances in LoM-50 (orange) and HiM-50 (red) compared to Ref-50 (blue) as a function of $M_{\star, \text{Chab}}$ for all galaxies at $z = 0.1$. Medians are indicated by solid lines and filled regions show the 10th to 90th percentile ranges. Abundances are measured in projection within one effective radius, r_e , while $M_{\star, \text{Chab}}$ is measured within a 30 kpc 3D aperture. Top: Stellar metallicity. Observations reported by Gallazzi et al. (2005) are shown as the black solid line with error bars indicating the scatter. Bottom: Gas-phase oxygen abundance for star-forming gas. For each galaxy, abundances are measured for all bound star-forming gas particles. The fit to observations from Zahid et al. (2014) is plotted as a magenta dashed line. Observations from Tremonti et al. (2004) are shown as black error bars. To ease comparison of the slopes with observations, the thick, red dashed curve lowers the trend for HiM-50 in both panels to match the normalization of the observations at $M_{\star, \text{Chab}} \approx 10^{11} M_{\odot}$. Stellar and gas-phase metallicities in high-mass galaxies in LoM-50 are consistent with Ref-50 but rise ≈ 0.4 dex higher for HiM-50 at the highest masses, in tension with the flattening in the observed relation above $10^{10} M_{\odot}$.

servations. To facilitate visual comparison of the slopes, we lower the normalization of the HiM-50 relation until the high-mass median value matches the value from Gallazzi et al. (2005), which we show as a dashed-red line in Fig. 11. Indeed, the reduced HiM-50 trend agrees with Gallazzi et al. (2005) better than is the case for Ref-50 or LoM-50, which begin to flatten at lower stellar masses than observed galaxies. This result is also in qualitative agreement with some SAMs that find that the slope of the mass–metallicity relation and the α -enhancement can be simultaneously reproduced with similar top-heavy IMF variations (De Lucia et al. 2017). The better agreement may also be related to the stellar feedback, as S15 found that the higher efficiency of star formation required to match the GSMF in the high-resolution simulation Recal-L025N0752 resulted in stronger outflows that decreased the metallicity of the ISM

enough to better match the observed mass–metallicity relations. Nevertheless, this rescaling is inconsistent with the radiative cooling rates in the simulation, so yields in HiM-50 would need to be rescaled in the simulation input parameters to verify that the better agreement of HiM-50 persists.

It is unclear how stellar metallicities derived from spectroscopic observations depend on the assumed IMF. Because the observationally inferred gas-phase metallicity may be less sensitive to the assumed IMF, we plot in the lower panel of Fig. 11 the mass-weighted gas-phase oxygen abundances for star-forming gas in our variable IMF simulations as a function of $M_{\star, \text{Chab}}$ (although gas-phase abundances may still be affected by IMF assumptions in observational studies, see e.g. Paalvast & Brinchmann 2017). We compare our simulations with the observations of Tremonti et al. (2004) and the fit to observations of Zahid et al. (2014) for $z = 0.1$. Note that we again only focus on the slope, rather than the absolute values, as a function of mass. Our trends here are qualitatively similar to those found for the stellar metallicities in the top panel of Fig. 11, with perhaps more scatter. HiM-50 continues increasing towards high masses while metallicities in Ref-50 and LoM-50 tend to flatten off above $10^{10} M_{\odot}$. Lowering the HiM-50 trend by 0.3 dex brings it into reasonable agreement with Tremonti et al. (2004), and yields a better match than Ref-50 or LoM-50. Again, self-consistent rescaling of the yields would be required to confirm this better agreement. We conclude that none of the models are ruled out by the observed trends between α -enhancement or metallicity and mass. For HiM-50 the absence of any flattening in the mass–metallicity relation is in tension with observations but larger volume simulations are required to judge the severity of the problem.

4.3 Star formation

As mentioned in earlier sections, SFRs can be strongly affected by the IMF, both the SFRs inferred from (virtual) observations as well as the true SFRs in the simulations due to the IMF’s effect on metallicity and stellar feedback. Fig. 12 details the effect of these variable IMF prescriptions on star formation in the simulations as a function of stellar mass. We show both the true quantities (dashed lines) and the corresponding values after interpreting the M_{\star} and SFR from the K -band and FUV light, respectively, under the assumption of a Chabrier IMF (solid lines). The reinterpreted SFR, SFR_{Chab} , is computed by multiplying the actual SFR by the ratio between the GALEX FUV-band luminosity of the stars and that they would have had if evolved with a Chabrier IMF, as also done by Clauwens et al. (2016). We note that this modifies the SFR by only ~ 10 per cent for most galaxies, leading to nearly identical trends of true and Chabrier-reinterpreted SFR with M_{\star} . Note as well that only the Chabrier-interpreted values (solid lines) should be compared with the observations.

The top left-hand panel of Fig. 12 shows the ‘galaxy formation efficiency’ of galaxies for our variable IMF simulations, defined as the ratio $M_{\star, \text{Chab}}/M_{\text{DM}}$ normalized to the cosmic baryon fraction. Both variable IMF simulations agree with Ref-50 (solid thick lines), as well as with abundance matching results from Behroozi et al. (2013) and Moster et al. (2013) shown as thin solid lines. For LoM-50, the true efficiency is shifted towards higher values for $M_{\star, \text{Chab}} > 10^{11} M_{\odot}$ (dashed orange line), where the excess stellar mass originates from dwarf stars. Interestingly, the true efficiencies for HiM-50 galaxies are quite close to their Chabrier-interpreted values. This implies that any correction to the observed galaxy formation efficiencies to account for IMF variations depends sensitively on how such variations are parametrized.

The specific SFR (sSFR), measured within 3D spherical apertures of radius 30 kpc, are shown in the upper right-hand panel of Fig. 12 for star-forming galaxies, defined as those with $u^* - r^* < 2$. In all simulations, the sSFRs are consistent for $M_{\star, \text{Chab}} < 10^{10.5} M_{\odot}$. At higher masses, sSFRs turn upwards for HiM-50, while they decrease for LoM-50 and Ref-50. The decrease in sSFR is consistent with the observed relations for local star-forming galaxies from the GAMA survey (Davies et al. 2016) as well as SDSS combined with Wide-field Infrared Survey Explorer (WISE) data (Chang et al. 2015),² while the upturn at high mass in HiM-50 is in tension with the observations. This upturn is not as strong for the true sSFR (red-dashed line), implying that the discrepancy for HiM-50 is partially due to reinterpreting the SFRs assuming a Chabrier IMF.

These higher sSFRs in HiM-50 are due to a lack of quenching at high masses. This can be seen in the middle left-hand panel of Fig. 12, where we show total SFR_{Chab} as a function of $M_{\star, \text{Chab}}$ for all galaxies at $z = 0.1$. Here it can be seen that in both Ref-50 and LoM-50, the positive trend between SFR_{Chab} and $M_{\star, \text{Chab}}$ flattens at $M_{\star, \text{Chab}} > 10^{10.5} M_{\odot}$ with a large amount of scatter towards low SFR_{Chab} as galaxies are quenched. However, in HiM-50, the trend becomes even steeper with less scatter at high mass. This behaviour leads to a lower passive fraction at high mass for HiM-50 galaxies, as shown in the middle right-hand panel of Fig. 12, where we define the passive fraction as the fraction of galaxies with intrinsic $u^* - r^* > 2$ in bins of M_{\star} . These low passive fractions in HiM-50 contrast with the increasing passive fraction with increasing M_{\star} seen in observations of SDSS galaxies (Moustakas et al. 2013), as well as Ref-50 and LoM-50.

In the lower left-hand panel of Fig. 12, we show that these higher SFRs in HiM-50 galaxies are a consequence of a larger cold gas fraction. We compute cold gas masses within apertures of 70 pkpc following Crain et al. (2017), where for each gas particle, we compute the mass of neutral hydrogen following the prescription of Rahmati et al. (2013), accounting for self-shielding and assuming a Haardt & Madau (2001) ionizing UV background.³ While the ratio between star-forming gas mass and $M_{\star, \text{Chab}}$ decreases steeply with increasing $M_{\star, \text{Chab}}$ in Ref-50 and LoM-50, this fraction drops less steeply in HiM-50, lying ≈ 0.5 dex above Ref-50 at $M_{\star, \text{Chab}} \gtrsim 10^{11} M_{\odot}$. The higher cold gas fractions in HiM-50 galaxies at $z = 0.1$ are likely due to burstier stellar feedback ejecting more gas out of galaxies at high- z . This effect causes a delay in the peak of star formation – we will investigate the time-dependent properties of galaxies in these variable IMF simulations in a future paper.

We compare our results in the lower left-hand panel of Fig. 12 with the median cold gas fractions of galaxies from the mass-selected extended GALEX Arcibo SDSS (xGASS) and extended CO Legacy Database (xCOLD) GASS surveys (Saintonge et al. 2017, Catinella et al. 2018). These observations match all of our simulations well for $M_{\star, \text{Chab}} > 10^{9.5} M_{\odot}$ but are too high in the lowest-mass bin. This low-mass tension is likely due to the fact that atomic hydrogen masses are not converged in the ‘intermediate’ resolution EAGLE model, especially in the range $M_{\star, \text{Chab}} \sim 10^{9-10} M_{\odot}$ where they are lower by nearly an order of magnitude relative to higher-resolution models (see Crain et al. 2017). Thus, the use of higher-resolution simulations may resolve this tension with observations for low-mass galaxies.

²We show results for galaxies from their online catalogue with $\text{FLAG}=1$, $\text{sSFR} > 10^{-2} \text{Gyr}^{-1}$, and M_{\star} above their mass completeness limit.

³As in Catinella et al. (2018), to account for helium we multiply the neutral hydrogen mass by 1.3 to obtain total cold gas mass.

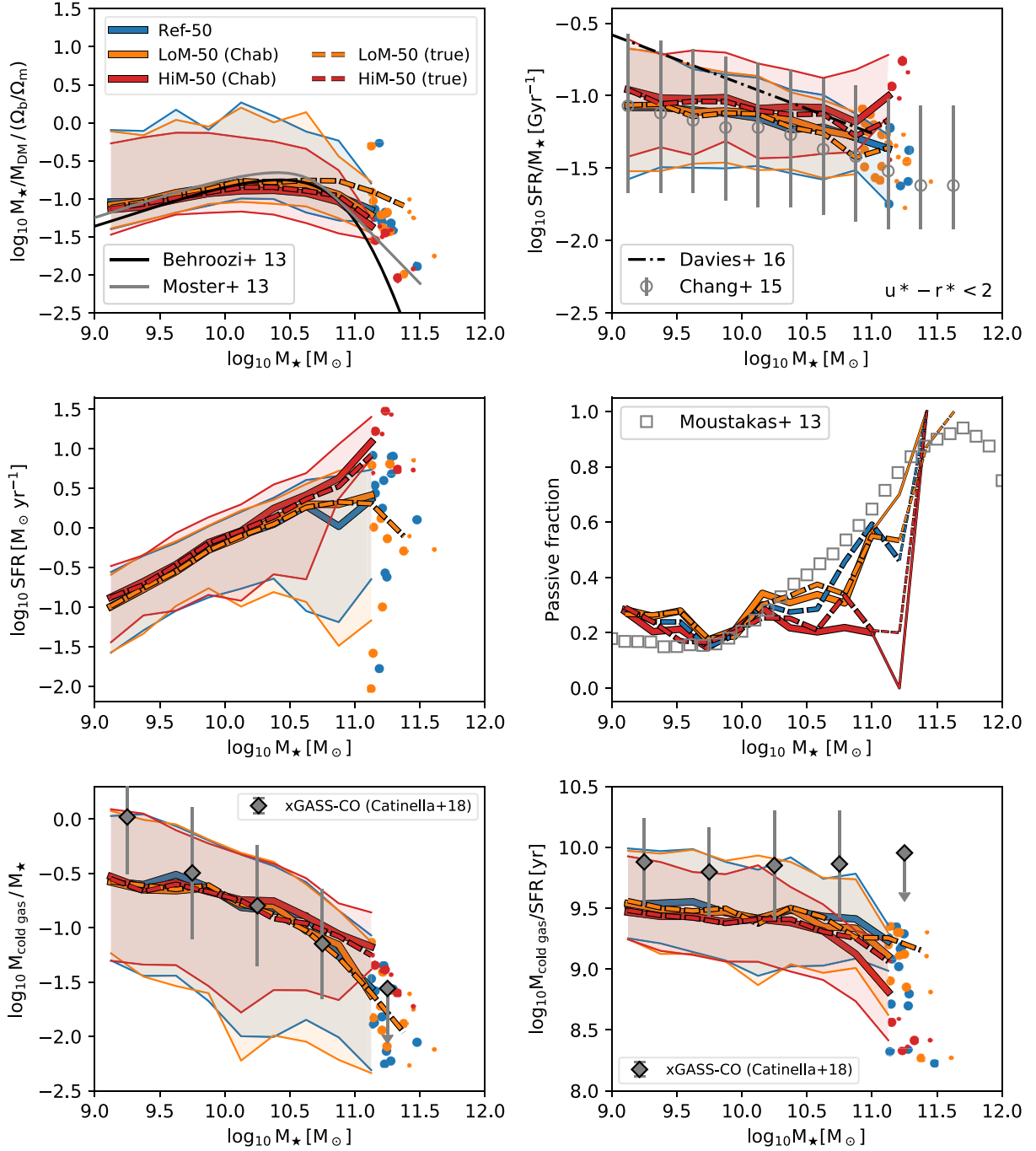


Figure 12. Star formation properties of all galaxies with $M_{\star, \text{Chab}} > 10^9 M_{\odot}$ at $z = 0.1$ as a function of stellar mass in our LoM-50 (orange) and HiM-50 (red) simulations, compared with Ref-50 (blue). Solid lines show results when reinterpreting the M_{\star} and SFR assuming a Chabrier IMF, while dashed lines show true values. In reading order from top left to bottom right: Galaxy formation efficiency, sSFR for star-forming galaxies, SFR for all galaxies, passive fractions, neutral hydrogen fractions, and gas consumption time-scale. Star-forming galaxies are defined as those with intrinsic $u^* - r^* < 2$. Medians are indicated as thick lines and shaded regions mark the 10–90th percentiles for the Chabrier-inferred lines. For bins with fewer than 10 galaxies, we show individual galaxies, where large and small dots refer to Chabrier-inferred and true values, respectively; for passive fractions these bins are indicated with thinner lines. Thin solid black and grey lines show abundance matching results from Behroozi, Wechsler & Conroy (2013) and Moster, Naab & White (2013), respectively. Grey circles show the median sSFR for SDSS+WISE galaxies from $z = 0$ to 0.2 with $\text{sSFR} > 10^{-2} \text{ Gyr}^{-1}$, with error bars showing the 10–90th percentiles (Chang et al. 2015). The black dash-dotted line in the upper right-hand panel shows a fit to the observed sSFRs for star-forming galaxies from $z = 0$ to 0.1 from the GAMA survey (Davies et al. 2016). Passive fractions for SDSS galaxies are shown as grey squares (Moustakas et al. 2013). Median neutral hydrogen fractions and gas consumption time-scales from the xGASS and xCOLDGASS surveys are shown as grey diamonds (Catinella et al. 2018); error bars denote 1σ scatter while arrows indicate upper limits on the median. The results for LoM-50 are consistent with Ref-50, but for HiM-50 the SFRs are too high for $M_{\star, \text{Chab}} > 10^{11} M_{\odot}$, resulting in lower passive fractions at high mass. These higher SFRs in HiM-50 are due to a higher gas fraction in high-mass galaxies.

The SFR change is not due to the renormalization of the star formation law (see Section 2.3). To show this, we plot the gas consumption time-scale, parametrized by the cold gas mass divided by the SFR, in the lower right-hand panel of Fig. 12 for all galaxies in each of our simulations. In all of our simulations the true value of this time-scale (dashed lines) is quite constant and is barely affected by the variable IMF (by at most 0.1 dex). We thus conclude that a higher cold gas mass fraction, rather than the renormalization of the star formation law, is responsible for the higher SFRs in high-mass HiM-50 galaxies. We also compare these results with the observed xGASS-CO sample, who also find a nearly constant relation. There is, however, a systematic ≈ 0.4 dex offset towards lower gas consumption time-scales relative to the observations. The reason for this offset is unclear, but is consistent with the expected systematic uncertainties associated with different SFR calibrators (compare results from Davies et al. 2016 and Chang et al. 2015 in the upper right-hand panel of Fig. 12).

4.4 ETG galaxy sizes

We showed in Section 3.2.1 that the observed sizes of LTGs are reproduced in our variable IMF simulations. We now investigate if this agreement persists for ETGs as well. In the left-hand panel of Fig. 13, we plot the half-light radius, r_e , as a function of $M_{*,\text{Chab}}$ for non-star-forming galaxies, defined as those with $u^* - r^* > 2$. While the relation for LoM-50 agrees well with Ref-50, HiM-50 ETGs are strongly offset to larger sizes at fixed $M_{*,\text{Chab}}$.

The reasons for this offset in HiM-50 are the same as those responsible for the smaller ≈ 0.1 dex offset seen for LTGs in Section 3.2.1. Stronger stellar feedback, later formation times, and the fact that the luminosities of the central, top-heavy regions of the galaxies are dimmer relative to a Chabrier IMF than the less top-heavy outskirts, all inflate the sizes at fixed $M_{*,\text{Chab}}$ relative to Ref-50. These effects are exacerbated in ETGs due to the fact that they are older, since old, top-heavy stellar populations are much dimmer than those with a Chabrier IMF, causing ETGs to shift towards lower $M_{*,\text{Chab}}$ at fixed r_e .

For comparison, we plot the observed relations from SDSS (Shen et al. 2003) and the GAMA survey (Lange et al. 2015) for $z < 0.1$ galaxies with Sersic index $n_s > 2.5$. We multiply the Shen et al. (2003) sizes by a factor 1.075 to convert from z - to r -band half-light radii (see Kelvin et al. 2012; Lange et al. 2015). Both LoM-50 and Ref-50 match the observed relations well for $M_{*,\text{Chab}} > 10^{10} M_\odot$, but HiM-50 ETGs are too large by $\approx 0.2 - 0.3$ dex, in tension with the observations.

In the right-hand panel of Fig. 13 we plot the true 3D half-mass radius as a function of true M_* for the same galaxies. As was the case for star-forming galaxies in Fig. 8, the physical sizes of high-mass ETGs in LoM-50 are smaller than in Ref-50 by $\approx 0.2 - 0.3$ dex at fixed M_* due to higher mass fractions of dwarf stars in their central regions. The positive offset in HiM-50 galaxies seen in the observable diagnostics is also present in the physical ones since, as with the light, the masses of the central, old stellar populations with top-heavy IMFs are lower than those with a Chabrier IMF, pushing these galaxies to lower M_* and larger half-mass radii.

5 SUMMARY AND CONCLUSIONS

We have modified the reference EAGLE cosmological, hydrodynamical simulations to self-consistently include a prescription for a stellar IMF that varies per-star particle as a function of the ISM pressure at which it was formed. Two prescriptions are explored: in

both cases we begin with the Kroupa double-power law IMF and vary the slope either below or above $0.5 M_\odot$ (hereafter referred to as LoM and HiM, respectively; see Fig. 2). For each prescription the dependence of the slope on birth ISM pressure was calibrated such that the observed Cappellari et al. (2013, hereafter C13) trend of increasing excess stellar mass-to-light (M/L) ratio with central stellar velocity dispersion, σ_e , is roughly recovered. For LoM this recovery is accomplished by an increase in the fraction of low-mass stars, while for HiM the increasing mass fraction in stellar remnants and decreased luminosity are responsible.

The calibration of the pressure-dependence of the IMF was performed by post-processing the $(100 \text{ Mpc})^3$ reference EAGLE simulation. From this post-processing procedure, we have found that

(i) In order to reproduce the observed trends between the stellar M/L ratio excess (MLE) relative to that expected for a Salpeter IMF (equation 1), and central stellar velocity dispersion (σ_e), LoM and HiM must, respectively, become more bottom- and top-heavy in higher-pressure environments (although this may also be possible with a ‘top-light’ HiM prescription, see Section 2.2). Since ISM pressures typically decrease with the age of the universe (e.g. Crain et al. 2015), these IMF prescriptions are implicitly time-dependent.

(ii) The MLE is only an excellent proxy for the IMF (i.e. independent of age and metallicity) when the high-mass slope is close to the ‘reference’ IMF (which here is Salpeter). For IMFs with shallower high-mass slopes, the MLE becomes strongly age-dependent for stars less than a few Gyr old (Fig. 2).

We ran two new $(50 \text{ Mpc})^3$ simulations with the same physics and resolution as the reference EAGLE model, but now each including one of our variable IMF prescriptions, which we refer to as LoM-50 and HiM-50, respectively. These simulations use variable nucleosynthetic yields, a star formation law, and stellar feedback consistent with their locally varying IMFs. Our conclusions are as follows:

(i) Both variable IMF simulations are broadly consistent with the observed trend between MLE and σ_e of C13 (Fig. 5). However, the trend in HiM-50 is less clear than for LoM-50 due to a lack of high- σ_e galaxies in the former, likely caused by burstier stellar feedback.

(ii) Galaxies in LoM-50 are consistent with the increasing fraction of dwarf stars towards higher σ_e in ETGs inferred by spectroscopic IMF studies, while those in HiM-50 do not show such a trend (Fig. 5). On the other hand, star-forming galaxies in HiM-50 show increasing ratios of ionizing flux to FUV flux with increasing r -band surface brightness, in agreement with recent observations, while LoM-50 galaxies show no trend (Fig. 7). It is unclear how to reconcile these apparently conflicting observations.

(iii) Relative to Ref-50, stellar ages and birth ISM pressures are largely unchanged in LoM-50, while HiM-50 produces younger stars at lower birth ISM pressures on average. This change may be due to the stronger stellar feedback from a top-heavy IMF (Fig. 6).

(iv) Observational proxies for the EAGLE subgrid calibration diagnostics [galaxy K -band luminosity function, r -band half-light radii (r_e) of LTGs, and BH masses] are consistent with the reference model for LoM-50. The same is true for HiM-50, except that the r_e of LTGs in HiM-50 are larger, but only by ≈ 0.1 dex, at fixed r -band luminosity (left column of Fig. 8).

(v) Stellar masses and LTG half-mass radii are larger and smaller, respectively, by $\approx 0.1 - 0.2$ dex in LoM-50 relative to Ref-50 (right-hand column of Fig. 8). This difference is due to an excess of (dim)

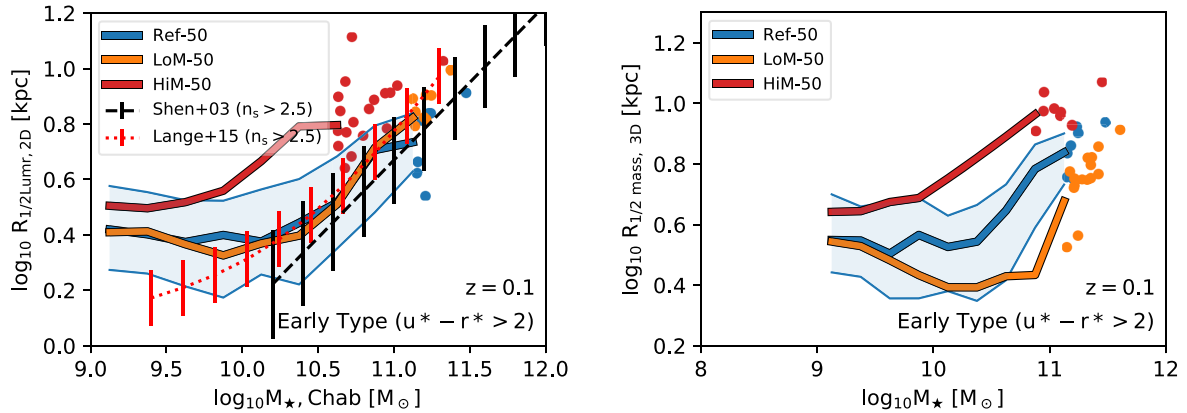


Figure 13. Size– M_{\star} relation for ETGs, defined as those with intrinsic $u^* - r^* > 2$, in Ref-50 (blue), LoM-50 (orange), and HiM-50 (red) at $z = 0.1$. Left-hand panel: 2D projected r -band half-light radius as a function of Chabrier-reinterpreted stellar mass, $M_{\star, \text{Chab}}$. The black-dashed and red-dotted lines with error bars show the observed trend from SDSS and GAMA, respectively, for galaxies with Sérsic index $n_s > 2.5$ (Shen et al. 2003; Lange et al. 2015). For the GAMA data we assume errors of 0.1 dex based on visual inspection of their 25–75th percentiles, while for the SDSS data we assume the errors are the same as those for their same relation for LTGs. Right-hand panel: Physical 3D half-mass radius as a function of M_{\star} . While the half-light radii of ETGs in Ref-50 and LoM-50 are consistent with observations, ETGs in HiM-50 are larger at fixed $M_{\star, \text{Chab}}$ by $\approx 0.2 - 0.4$ dex. As is the case for LTGs (see middle row of Fig. 8), LoM-50 ETGs are physically smaller by $\approx 0.2 - 0.3$ dex at fixed M_{\star} relative to Ref-50.

dwarf stars that increase the mass (but not the light) in the central regions of these galaxies as a result of the bottom-heavy IMF.

We also investigated the effect that the IMF has on predicted global galaxy scaling relations. Our results are as follows:

(i) While the slopes of the $[\text{Mg}/\text{Fe}] - \sigma_e$ relation in the Ref-50 and HiM-50 simulations for high-mass ($\sigma_e > 60 \text{ km s}^{-1}$), early-type ($u - r > 2$) galaxies (0.33 ± 0.07 and 0.40 ± 0.12 , respectively) are consistent with the observed relation, LoM-50 produces a relation with slightly (but significantly) shallower slope of 0.26 ± 0.04 , compared to observed values 0.33 ± 0.01 (Thomas et al. 2010) and 0.33 ± 0.03 (Conroy et al. 2014) (Fig. 10). The normalization of the $[\text{Mg}/\text{Fe}] - \sigma_e$ relation in HiM-50 is ≈ 0.15 dex higher than in Ref-50, bringing it into agreement with observations, while it is $\approx 0.05 - 0.1$ dex lower in LoM-50, ≈ 0.2 dex below the observed relation. However, given the large systematic uncertainties in the normalization for both observations and simulations, these differences in normalization are not significant.

(ii) Stellar and gas-phase metallicities in LoM-50 are consistent with Ref-50, but in HiM-50 both quantities increase steeply with $M_{\star, \text{Chab}}$, the stellar mass inferred under the assumption of a Chabrier IMF, above $10^{10} M_{\odot}$ with no sign of flattening at higher mass. This is contrary to the flattening seen in Ref-50 and LoM-50, and is possibly inconsistent with similar flattening seen in observations.

(iii) The relations between $M_{\star, \text{Chab}}/M_{\text{DM}}$ and $M_{\star, \text{Chab}}$ for galaxies in LoM-50 and HiM-50 are consistent with Ref-50. Adopting true M_{\star} in place of $M_{\star, \text{Chab}}$ increases $M_{\star, \text{Chab}}/M_{\text{DM}}$ by ≈ 0.3 dex for LoM-50 galaxies with $M_{\star} > 10^{10.5} M_{\odot}$, but has little effect for HiM-50 galaxies. Specific SFRs of galaxies with $M_{\star, \text{Chab}} \gtrsim 10^{10.5} M_{\odot}$ are higher in HiM-50 than in Ref-50, resulting in a lower passive fraction that does not rise with stellar mass up to at least $M_{\star, \text{Chab}} = 10^{11} M_{\odot}$, in tension with observations (Fig. 12). This higher SFR is a result of a higher star-forming gas fraction in high-mass galaxies, likely due to the burstier feedback being more efficient at ejecting gas from galaxies at early times, delaying star-formation to lower z .

(iv) While the half-light radii of ETGs in LoM-50 are consistent with Ref-50 and observations, those in HiM-50 with $M_{\star, \text{Chab}} > 10^{10} M_{\odot}$ are about a factor of 2 larger at fixed $M_{\star, \text{Chab}}$, inconsistent with observations. The larger half-light radii are likely due

to stronger stellar feedback and the stronger dimming of old stellar populations with a top-heavy IMF relative to a Chabrier IMF, increasing r_e and decreasing $M_{\star, \text{Chab}}$, respectively.

The results of this project are intended to aid in the interpretation of evidence for IMF variations in real galaxies, especially in terms of *how* the IMF varies, either at the high-mass or the low-mass end. While a high-mass slope variation cannot be definitively ruled out by the present analysis, the model in which the IMF varies at the low-mass end (as well as the reference model with a Chabrier IMF) produces galaxies that match observations much more closely than one in which the high-mass slope is varied.

It is intriguing that observations that are sensitive to the low-mass and high-mass slopes of the IMF prefer LoM and HiM, respectively, with neither model matching all observations simultaneously (Figs 5 and 7). This result could indicate that IMF variations are more complex than those explored in this paper, or that systematic uncertainties in models used to constrain the IMF observationally are underestimated.

This paper lays the groundwork for further analysis of the predicted trends between the IMF (parametrized by the MLE) and galaxy properties, which will be explored in a series of upcoming papers (Papers II and III). Paper II will investigate the trends between the MLE and global properties across the galaxy population, uncovering the observable properties that are predicted to correlate with the MLE most strongly. In Paper III we delve into spatially resolved properties of individual galaxies, exploring how IMF variations affect radial gradients in M/L , metal abundances, and MLE to further expose the differences in the predictions due to the non-universality of the IMF. Paper III will also investigate the time dependence of the IMF and its effect on the evolution of galaxies in our simulations.

ACKNOWLEDGEMENTS

We are grateful to the anonymous referee for constructive feedback which increased the overall quality of the paper. We thank Fabio Fontanot for useful comments. CB thanks Bart Clauwens, Madusha Gunawardhana, Padraig Alton, Richard Bower, Tom The-

uns, Marijn Franx, and Matthieu Schaller for insightful discussions regarding this project. This work used the DiRAC Data Centric system at Durham University, operated by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). This equipment was funded by BIS National E-infrastructure capital grant ST/K00042X/1, STFC capital grants ST/H008519/1 and ST/K00087X/1, STFC DiRAC Operations grant ST/K003267/1 and Durham University. DiRAC is part of the National E-Infrastructure. RAC is a Royal Society University Research Fellows. We also gratefully acknowledge PRACE for awarding us access to the resource Curie based in France at Très Grand Centre de Calcul. This work was sponsored by the Dutch National Computing Facilities Foundation (NCF) for the use of supercomputer facilities, with financial support from the Netherlands Organization for Scientific Research (NWO). This research made use of `ASTROPY`, a community-developed core `PYTHON` package for Astronomy (`Astropy Collaboration 2013`).

REFERENCES

- Alton P. D., Smith R. J., Lucey J. R., 2017, *MNRAS*, 468, 1594
- Arrighi M., Trager S. C., Somerville R. S., Gibson B. K., 2010, *MNRAS*, 402, 173
- Astropy Collaboration, 2013, *A&A*, 558, A33
- Auger M. W., Treu T., Gavazzi R., Bolton A. S., Koopmans L. V., Marshall P. J., 2010, *ApJ*, 721, L163
- Baldry I. K. et al., 2012, *MNRAS*, 421, 621
- Barnabè M., Spiniello C., Koopmans L. V., Trager S. C., Czoske O., Treu T., 2013, *MNRAS*, 436, 253
- Bastian N., Covey K. R., Meyer M. R., 2010, *ARA&A*, 48, 339
- Bate M. R., 2009, *MNRAS*, 392, 1363
- Bate M. R., Bonnell I. A., 2005, *MNRAS*, 356, 1201
- Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, *MNRAS*, 356, 1191
- Behroozi P. S., Wechsler R. H., Conroy C., 2013, *ApJ*, 770, 57
- Bekki K., 2013, *MNRAS*, 436, 2254
- Bernardi M. et al., 2018, *MNRAS*, 475, 757
- Blancato K., Genel S., Bryan G., 2017, *ApJ*, 845, 21
- Bondi H., Hoyle F., 1944, *MNRAS*, 104, 273
- Booth C. M., Schaye J., 2009, *MNRAS*, 398, 53
- Bower R. G., Schaye J., Frenk C. S., Theuns T., Schaller M., Crain R. A., McAlpine S., 2017, *MNRAS*, 465, 32
- Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
- Calzetti D., 2001, *PASP*, 113, 1449
- Camps P., Baes M., 2015, *Astron. Comput.*, 9, 20
- Cappellari M. et al., 2013, *MNRAS*, 432, 1862 (C13)
- Cardelli J., Clayton G., Mathis J., 1989, in Allamandola L., Tielens A., eds, Proc. IAU Symp. 135, Interstellar Dust, Santa Clara, California, p. 5, <http://adsabs.harvard.edu/abs/1989IAUS..135P...5C>
- Catinella B. et al., 2018, *MNRAS*, 476, 875
- Canarro A. J., Gorgas J., Vazdekis A., Cardiel N., Peletier R. F., 2003, *MNRAS*, 339, L12
- Chabrier G., 2003, *ApJ*, 586, 1
- Chang Y. Y., Wel A. V. D., Cunha E. D., Rix H. W., 2015, *ApJS*, 219, 9
- Charlot S., Fall S. M., 2000, *ApJ*, 539, 718
- Clauwens B., Schaye J., Franx M., 2015, *MNRAS*, 449, 4091
- Clauwens B., Schaye J., Franx M., 2016, *MNRAS*, 462, 2832
- Collier W. P., Smith R. J., Lucey J. R., 2018, *MNRAS*, 473, 1103
- Conroy C., Graves G. J., van Dokkum P. G., 2014, *ApJ*, 780, 33
- Conroy C., Gunn J. E., 2010, *ApJ*, 712, 833
- Conroy C., Gunn J. E., White M., 2009, *ApJ*, 699, 486
- Conroy C., van Dokkum P., Villaume A., 2017, *ApJ*, 837, 8
- Conroy C., van Dokkum P. G., 2012, *ApJ*, 760, 71
- Correa C. A., Schaye J., Clauwens B., Bower R. G., Crain R. A., Schaller M., Theuns T., Thob A. C., 2017, *MNRAS*, 472, L45
- Crain R. A. et al., 2015, *MNRAS*, 450, 1937 (C15)
- Crain R. A. et al., 2017, *MNRAS*, 464, 4204
- Dahlen T., al. E., 2008, *ApJ*, 681, 462
- Dalla Vecchia C., Schaye J., 2012, *MNRAS*, 426, 140
- Davies L. J. et al., 2016, *MNRAS*, 461, 458
- Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, *ApJ*, 292, 371
- Davis T. A., McDermid R. M., 2017, *MNRAS*, 464, 453
- De Lucia G., Fontanot F., Hirschmann M., 2017, *MNRAS*, 466, L88
- De Rossi M. E., Bower R. G., Font A. S., Schaye J., Theuns T., 2017, *MNRAS*, 472, 3354
- Dilday B. et al., 2010, *ApJ*, 713, 1026
- Dolag K., Borgani S., Murante G., Springel V., 2009, *MNRAS*, 399, 497
- Driver S. P. et al., 2012, *MNRAS*, 427, 3244
- Dutton A. A., Mendel J. T., Simard L., 2012, *MNRAS*, 422, L33
- Ferreras I., La Barbera F., de la Rosa I. G., Alexandre V., de Carvalho R. R., Falcón-Barroso J., Ricciardelli E., 2013, *MNRAS*, 429, L15
- Fischer J., Dopita M., 2005, *ApJ*, 619, 340
- Fontanot F., 2014, *MNRAS*, 442, 3138
- Fontanot F., De Lucia G., Hirschmann M., Bruzual G., Charlot S., Zibetti S., 2017, *MNRAS*, 464, 3812
- Fontanot F., De Lucia G., Xie L., Hirschmann M., Bruzual G., Charlot S., 2018, *MNRAS*, 475, 2467
- Gallazzi A., Charlot S., Brinchmann J., White S. D., Tremonti C. A., 2005, *MNRAS*, 362, 41
- Gargiulo I. D. et al., 2015, *MNRAS*, 446, 3820
- Graur O., Maoz D., 2013, *MNRAS*, 430, 1746
- Graur O. et al., 2014, *ApJ*, 783, 28
- Gunawardhana M. L. P. et al., 2011, *MNRAS*, 415, 1647
- Guszejnov D., Hopkins P. F., Ma X., 2017, *MNRAS*, 472, 2107
- Gutcke T. A., Springel V., 2017, preprint ([arXiv:1710.04222](https://arxiv.org/abs/1710.04222))
- Haardt F., Madau P., 2001, in Neumann D. M., Tran J. T. V., eds, Clusters of Galaxies and the High Redshift Universe Observed in X-rays. Savoie, France, p. 64, <http://adsabs.harvard.edu/abs/2001cgch.confE..64H>
- Haas M. R., Schaye J., Booth C. M., Vecchia C. D., Springel V., Theuns T., Wiersma R. P. C., 2013, *MNRAS*, 435, 2931
- Habergham S. M., Anderson J. P., James P. A., 2010, *ApJ*, 717, 342
- Hennebelle P., Chabrier G., 2013, *ApJ*, 770, 150
- Hopkins P. F., 2012, *MNRAS*, 423, 2037
- Jappsen A.-K., Klessen R. S., Larson R. B., Li Y., Mac Low M.-M., 2005, *A&A*, 435, 611
- Kelvin L. S. et al., 2012, *MNRAS*, 421, 1007
- Kennicutt R. C., Jr. 1998, *ApJ*, 498, 541
- Kewley L. J., Ellison S. L., 2008, *ApJ*, 681, 1183
- Kormendy J., Ho L. C., 2013, *ARA&A*, 51, 511
- Kroupa P., 2001, *MNRAS*, 322, 231
- Kroupa P., Weidner C., 2003, *ApJ*, 598, 1076
- Krumholz M. R., 2011, *ApJ*, 743, 110
- La Barbera F., Ferreras I., Vazdekis A., 2015, *MNRAS*, 449, L137
- La Barbera F., Ferreras I., Vazdekis A., de la Rosa I. G., de Carvalho R. R., Trevisan M., Falcón-Barroso J., Ricciardelli E., 2013, *MNRAS*, 433, 3017
- La Barbera F., Vazdekis A., Ferreras I., Pasquali A., Cappellari M., Martín-Navarro I., Schönebeck F., Falcón-Barroso J., 2016, *MNRAS*, 457, 1468
- Lange R. et al., 2015, *MNRAS*, 447, 2603
- Lejeune T., Cuisinier F., Buser R., 1997, *A&AS*, 125, 229
- Lejeune T., Cuisinier F., Buser R., 1998, *A&AS*, 130, 65
- Li H. et al., 2017, *ApJ*, 838, 77
- Lupton R. H., Blanton M. R., Fekete G., Hogg D. W., O'Mullane W., Szalay A. S., Wherry N., 2004, *PASP*, 116, 133
- Lyubanova M. et al., 2016, *MNRAS*, 463, 3220
- Marigo P., Girardi L., 2007, *A&A*, 469, 239
- Marigo P., Girardi L., Bressan A., Groenewegen M. A. T., Silva L., Granato G. L., 2008, *A&A*, 482, 883
- Martín-Navarro I., 2016, *MNRAS*, 456, L104
- Martín-Navarro I., La Barbera F., Vazdekis A., Falcón-Barroso J., Ferreras I., 2015a, *MNRAS*, 447, 1033
- Martín-Navarro I. et al., 2015b, *ApJ*, 806, L31
- McAlpine S., Helly J. C., Schaller M., Trayford J. W., Qu Y., Furlong M., Bower R. G., Crain R. A., 2016, *Astron. Comput.*, 15, 72

- McDermid R. M. et al., 2014, *ApJ*, 792, 5
- McDermid R. M. et al., 2015, *MNRAS*, 448, 3484
- Meurer G. R. et al., 2009, *ApJ*, 695, 765
- Moster B. P., Naab T., White S. D., 2013, *MNRAS*, 428, 3121
- Moustakas J. et al., 2013, *ApJ*, 767, 50
- Narayanan D., Davé R., 2012, *MNRAS*, 423, 3601
- Narayanan D., Davé R., 2013, *MNRAS*, 436, 2892
- Newman A. B., Smith R. J., Conroy C., Villaume A., van Dokkum P., 2017, *ApJ*, 845, 18
- Oldham L., Auger M., 2018, *MNRAS*, 474, 4169
- Paalvast M., Brinchmann J., 2017, *MNRAS*, 470, 1612
- Perrett K. et al., 2012, *AJ*, 144, 59
- Planck Collaboration I, 2014, *A&A*, 571, A1
- Portinari L., Chiosi C., Bressan A., 1998, *A&A*, 334, 38
- Posacki S., Cappellari M., Treu T., Pellegrini S., Ciotti L., 2015, *MNRAS*, 446, 493
- Rahmati A., Pawlik A. H., Raičević M., Schaye J., 2013, *MNRAS*, 430, 2427
- Rosani G., Pasquali A., La Barbera F., Ferreras I., Vazdekis A., 2018, *MNRAS*, 476, 5233
- Rosas-Guevara Y. M. et al., 2015, *MNRAS*, 454, 1038
- Saintonge A., Catinella B., Tacconi L. J., Kauffmann G., Genzel R., Cortese L., Davé R., Fletcher T. J., 2017, *ApJS*, 233, 22
- Sales L. V., Navarro J. F., Schaye J., Vecchia C. D., Springel V., Booth C. M., 2010, *MNRAS*, 409, 1541
- Salpeter E. E., 1955, *ApJ*, 121, 161
- Sarzi M., Spiniello C., La Barbera F., Krajnović D., van den Bosch R., 2018, *MNRAS*, 478, 4084
- Schaller M., Dalla Vecchia C., Schaye J., Bower R. G., Theuns T., Crain R. A., Furlong M., McCarthy I. G., 2015, *MNRAS*, 454, 2277
- Schaye J., 2004, *ApJ*, 609, 667
- Schaye J., Dalla Vecchia C., 2008, *MNRAS*, 383, 1210
- Schaye J. et al., 2010, *MNRAS*, 402, 1536
- Schaye J. et al., 2015, *MNRAS*, 446, 521 (S15)
- Segers M. C., Schaye J., Bower R. G., Crain R. A., Schaller M., Theuns T., 2016, *MNRAS*, 461, L102
- Shen S., Mo H. J., White S. D. M., Blanton M. R., Kauffmann G., Voges W., Brinkmann J., Csabai I., 2003, *MNRAS*, 343, 978
- Smith R. J., 2014, *MNRAS*, 443, L69
- Smith R. J., Alton P., Lucey J. R., Conroy C., Carter D., 2015, *MNRAS*, 454, L71
- Sonnenfeld A., Nipoti C., Treu T., 2017, *MNRAS*, 465, 2397
- Sonnenfeld A., Treu T., Marshall P. J., Suyu S. H., Gavazzi R., Auger M. W., Nipoti C., 2015, *ApJ*, 800, 94
- Spiniello C., Koopmans L. V., Trager S. C., Czoske O., Treu T., 2011, *MNRAS*, 417, 3000
- Spiniello C., Trager S., Koopmans L. V., Conroy C., 2014, *MNRAS*, 438, 1483
- Spiniello C., Trager S. C., Koopmans L. V., Chen Y. P., 2012, *ApJ*, 753, 32
- Springel V., 2005, *MNRAS*, 364, 1105
- Springel V., Di Matteo T., Hernquist L., 2005, *MNRAS*, 361, 776
- Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, *MNRAS*, 328, 726
- Thomas D., Maraston C., Schawinski K., Sarzi M., Silk J., 2010, *MNRAS*, 404, 1775
- Thomas J. et al., 2011, *MNRAS*, 415, 545
- Tortora C., Romanowsky A. J., Napolitano N. R., 2013, *ApJ*, 765, 8
- Tremonti C. A. et al., 2004, *ApJ*, 613, 898
- Treu T., Auger M. W., Koopmans L. V. E., Gavazzi R., Marshall P. J., Bolton A. S., 2010, *ApJ*, 709, 1195
- van Dokkum P., Conroy C., Villaume A., Brodie J., Romanowsky A., 2017, *ApJ*, 841, 23
- Van Dokkum P. G., Conroy C., 2010, *Nature*, 468, 940
- Vazdekis A., Casuso E., Peletier R. F., Beckman J. E., 1996, *ApJS*, 106, 307
- Vazdekis A., Ricciardelli E., Cenarro A. J., Rivero-González J. G., Díaz-García L. A., Falcón-Barroso J., 2012, *MNRAS*, 424, 157
- Vazdekis A., Sánchez-Blázquez P., Falcón-Barroso J., Cenarro A. J., Beasley M. A., Cardiel N., Gorgas J., Peletier R. F., 2010, *MNRAS*, 404, 1639
- Westera P., Lejeune T., Buser R., Cuisinier F. A. G. B., 2002, *A&A*, 381, 524
- Wiersma R. P. C., Schaye J., Smith B. D., 2009a, *MNRAS*, 393, 99
- Wiersma R. P. C. R., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009b, *MNRAS*, 399, 574
- Zahid H. J., Dima G. I., Kudritzki R.-P., Kewley L. J., Geller M. J., Hwang H. S., Silverman J. D., Kashino D., 2014, *ApJ*, 791, 130
- Zhang Z.-Y., Romano D., Ivison R. J., Papadopoulos P. P., Matteucci F., 2018, *Nature*, 558, 260

APPENDIX A: APERTURE EFFECTS AND IMF CALIBRATION DETAILS

In Fig. A1, we show the effect of aperture choice on the MLE- σ_e relation for mock C13 galaxies in LoM-50 and HiM-50. Comparing the left- and right-hand columns, we see that the global IMF, measured over all stellar particles in each galaxy, underestimates the MLE values when measured within r_e , by ≈ 0.1 dex. Choosing an even smaller aperture of $r_e/2$ increases the difference further to ≈ 0.1 dex. We did not measure the IMF within $r_e/8$, as for many galaxies this is below the resolution limit of the simulations.

Fig. A2 outlines the process of calibrating the variable IMF prescriptions LoM and HiM for mock C13 galaxies in the left- and right-hand columns, respectively, by showing the MLE- σ_e relation for different steps in the process. The top row shows the Ref-100 simulation [with the same resolution as Ref-50 but with a volume (100 Mpc^3), where the $z = 0.1$ masses and luminosities of the stars were recomputed in post-processing assuming they evolved according to either the LoM or HiM variable IMF prescription (upper left- and right-hand panels, respectively)]. The positive trend up to high σ_e is clear in both cases. While the best-fitting relation in LoM-50 is offset from the C13 relation, the slopes are consistent. As mentioned in Section 3.1, this offset is due to the fact that this IMF variation prescription was originally calibrated using apertures larger than r_e . The smaller aperture size in this plot excludes stars with Chabrier-like IMFs, increasing the MLE for the entire relation. For HiM-50, the normalization of the MLE- σ_e relation is much closer to the C13 relation, but the slope is slightly shallower. Since these calibrations were done by eye, a perfect match to the C13 slope is not expected. Indeed, the agreement with C13 is still very good, since most of our points lie within the 1σ scatter of their relation.

The middle row shows the same IMF prescriptions applied to the Ref-50 simulation. Here we are missing the high- σ_e galaxies due to the smaller box size, but the positive trend is still significant in both cases. Finally, the bottom row shows the results for LoM-50 and HiM-50, which were run with the LoM and HiM IMF prescriptions self-consistently included. The trend for LoM is preserved in LoM-50, with a slightly higher normalization. As mentioned in Section 3.1, this is due to the typically larger birth ISM densities at which stars are born in LoM-50 relative to Ref-50. On the other hand, the trend in HiM-50 is slightly weaker due to smaller values of σ_e as well as younger ages. Better statistics at high- σ_e may be required to determine if galaxies in HiM-50 are inconsistent with the C13 trend.

Fig. A3 shows the MLE are a function of the median birth pressure of stars within r_e of galaxies at $z = 0.1$ in LoM-50 (left) and HiM-50 (right). MLE correlates extremely well with pressure in LoM-50, but more weakly in HiM-50. The larger scatter in HiM-50 is due to the age-dependence of MLE for a shallow high-mass IMF slope.

APPENDIX B: SELF-CONSISTENCY TESTS

Making the simulations completely self-consistent with the variable IMF while simultaneously ensuring that the subgrid calibration diagnostics remained consistent with the reference model was a challenging and painstaking process. Due to the non-linear process of galaxy formation and its connection with the IMF, it can be unclear which effect of the variable IMF is responsible for changes in the galaxy properties. Thus, we investigated the effect of a variable IMF in smaller, $(25 \text{ Mpc})^3$ boxes, adding new effects of the variable IMFs one at a time until they became self-consistent. Fig. B1 shows this process for LoM-50, and its effects on the subgrid physics calibration diagnostics: the K -band luminosity (M_K) function, the $r_e - M_K$ relation, and the $M_{\text{BH}} - M_K$ relation. To emphasize discrepancies in M_{BH} , we plot medians relative to the Ref-25 simulation in the right-hand panel.

First we allowed the IMF to affect only the yields in the simulation, while keeping the feedback and star formation law consistent with that used for the reference (Chabrier IMF) model. This model we refer to as ‘LoM-25 yields’, represented by the orange curves in Fig. B1. For all calibration diagnostics, this model agrees very well with Ref-25, which is not surprising since metallicities are not strongly affected in the LoM-50 runs.

We next added the effect of modifying the physical star-formation law such that the observed star formation law, i.e. intrinsic UV surface brightness as a function of gas surface density, was preserved, shown as the green curve. In this case the simulation produced a larger number of high-mass galaxies, while reducing M_{BH} by over

0.5 dex for these galaxies, without strongly affecting the sizes. This result can be explained by self-regulation of stellar and BH growth. Since the feedback per stellar mass formed was still set as ‘Chabrier’, the increased normalization of the SF law (i.e. the smaller gas consumption time-scale) would result in stronger stellar feedback at fixed gas surface density. Thus, galaxies naturally decrease the gas density (and thus the SFR) until the feedback returns to the value appropriate for self-regulation (i.e. outflows roughly balance inflows). This lower gas density reduces the ability of gas to accrete onto BHs, reducing their accretion rates and thus their final $z = 0.1$ masses. With lower BH masses, AGN feedback is suppressed, causing stellar feedback to compensate, resulting in higher M_* and thus brighter M_K .

To alleviate this issue, we made the feedback self-consistent, shown by the red curves in Fig. B1. With decreased stellar feedback per unit stellar mass, the effect of the modified star-formation law is cancelled out such that the amount of feedback at fixed gas surface density is more consistent with the reference model. This model is much more consistent with Ref-25 and with the calibration data, and is our fiducial model.

Fig. B2 shows the same results but for the HiM IMF and corresponding HiM-25 simulations. Modifying the yields (orange curve) increases the sizes slightly and BH masses more strongly. Since the HiM IMF affects metallicities much more strongly than the LoM IMF, this can be seen as the effect of increasing the cooling rate due to higher metal yields, increasing the ability of gas to accrete onto BH particles and the importance of AGN feedback.

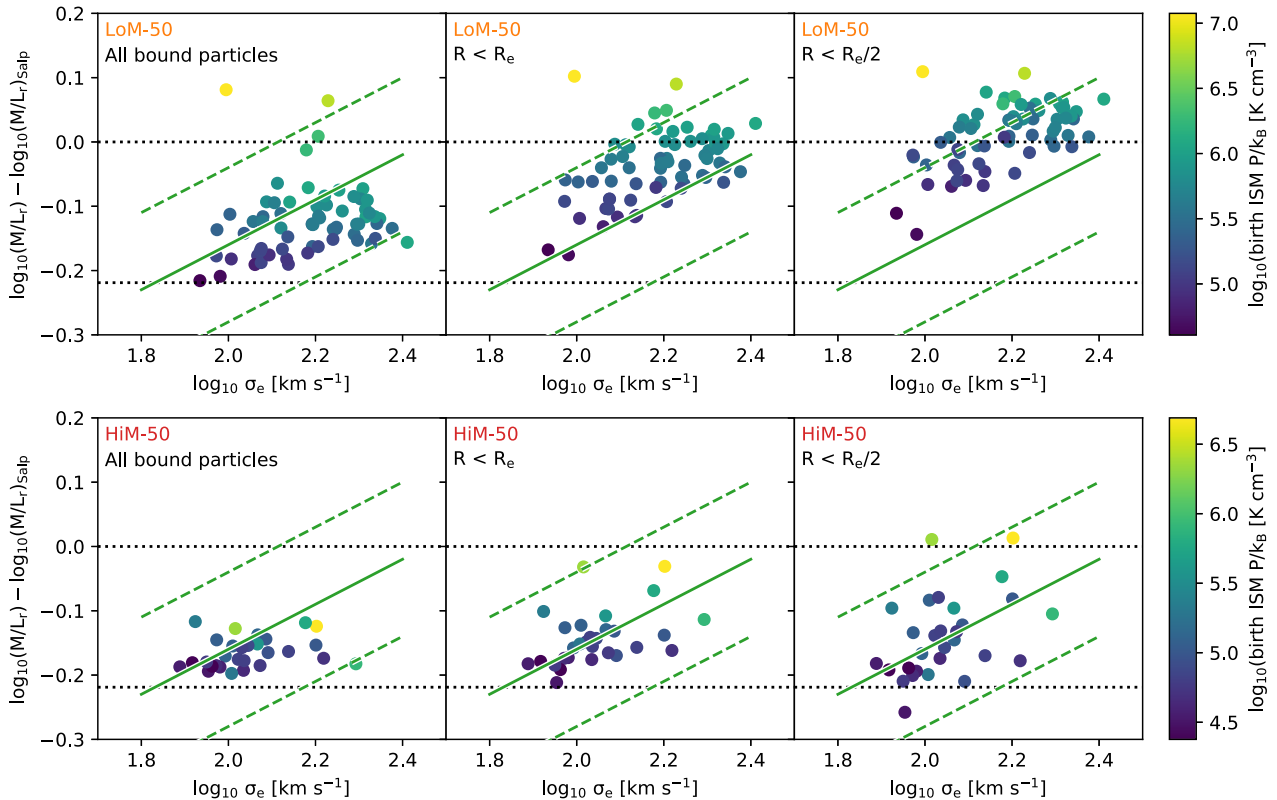


Figure A1. Effect of aperture choice on IMF measurement. Mass-to-light excess (MLE) over Salpeter as a function of stellar velocity dispersion for mock C13 galaxies in our LoM-50 (top row) and HiM-50 (bottom row) simulations at $z = 0.1$. From left to right, the MLE is measured for all bound stars, stars within $1 r_e$, and stars within $0.5 r_e$, respectively. For reference, in all panels we plot the $\text{MLE}_r - \sigma_e$ relation of Cappellari et al. (2013) with 1σ scatter as green-solid and green-dashed lines, respectively. Horizontal dashed lines indicate MLE_r for Salpeter and Chabrier IMFs from top to bottom, respectively. The value of MLE is sensitive to the precise choice of aperture.

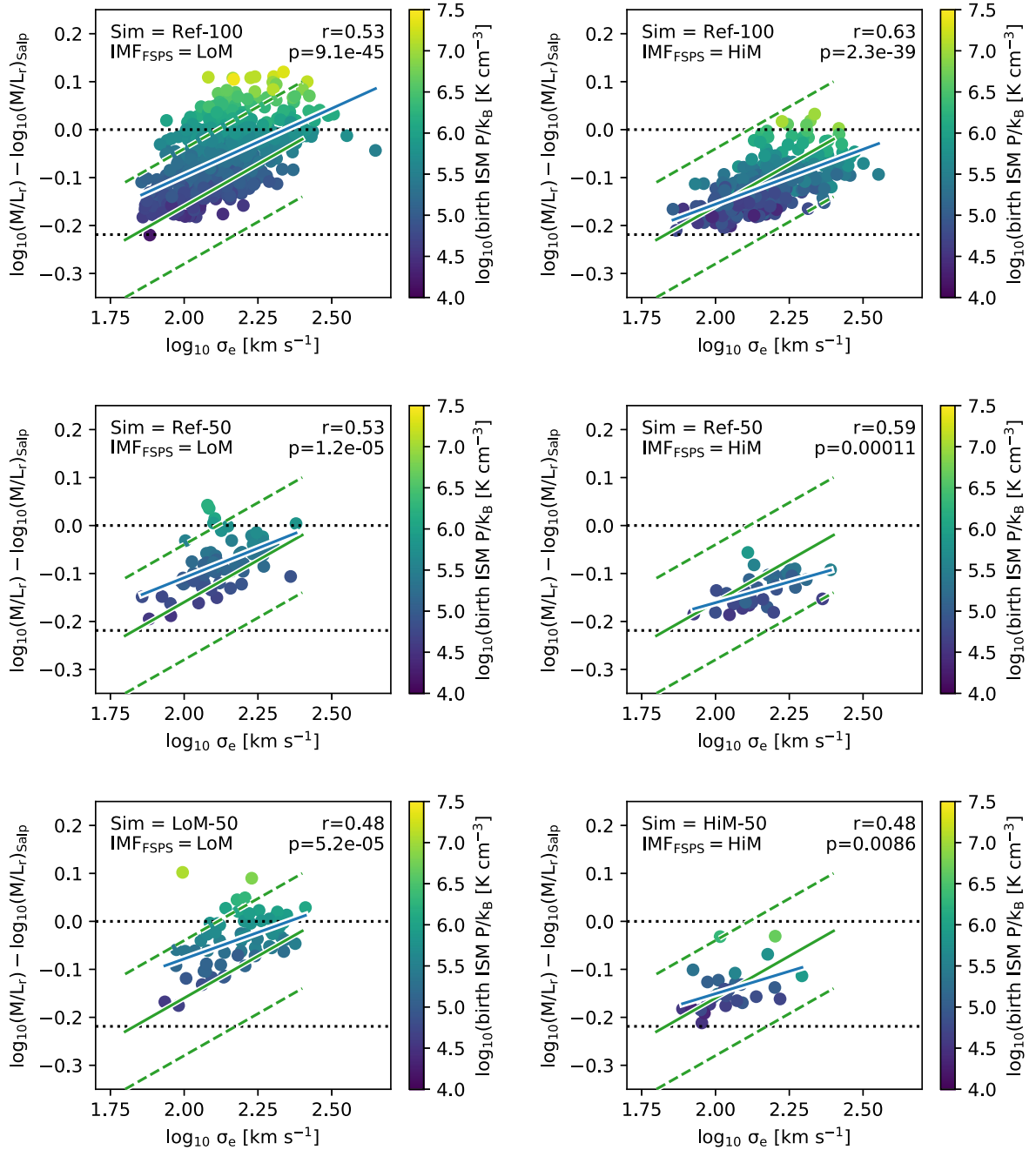


Figure A2. Variable IMF prescription calibration. MLE_r as a function of σ_e for galaxies consistent with a C13 selection. All quantities are measured within r_e . Simulations in the left- and right-hand columns assume the LoM and HiM IMF variation prescription, respectively. Top row: Post-processed Ref-100 simulation. Middle row: Post-processed Ref-50 simulation. Bottom row: Self-consistent LoM-50 (left) and HiM-50 (right) simulations. Least-absolute deviation fits are shown as blue solid lines, and the Spearman r coefficient and its associated p -value are indicated in the upper right-hand corner of each panel. The calibrated correlation between the MLE_r and σ_e is preserved in the LoM-50 simulation, but is reduced in HiM-50.

Reducing the physical star-formation law at high pressure so as to maintain the same observed law makes the situation worse, further increasing BH masses and sizes, while also strongly suppressing the bright end of the M_K function. The situation is essentially reversed relative to the LoM-25 case: the gas density is increased in order to obtain strong enough stellar feedback for self-regulation, which increases the BH masses and AGN feedback, which lowers stellar masses at the high-mass end of the GSMF. Again, including self-consistent stellar feedback (red curve) removes the need to

change the gas densities, correcting the stellar and BH masses. The sizes are still slightly larger, but this is not a large difference from Ref-25.

APPENDIX C: THE DWARF-TO-GIANT RATIO

Spectroscopic IMF studies are sensitive to the ratio of dwarf-to-giant stars in the present-day stellar populations of the galaxies for which they constrain the IMF. La Barbera et al. (2013) have concluded that as long as models that differ only in IMF parametrization fit the

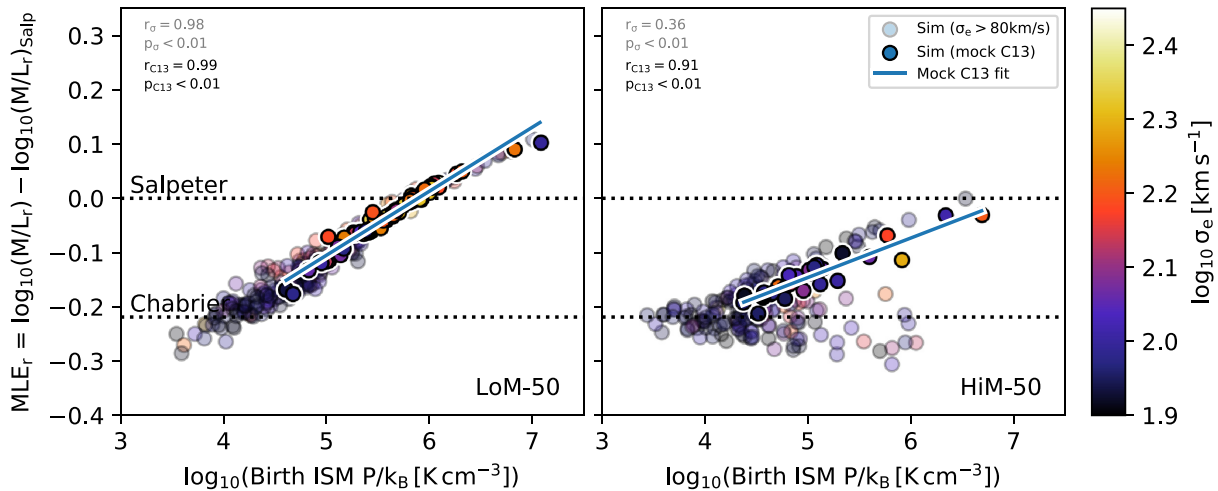


Figure A3. MLE_r as a function of birth ISM pressure for LoM-50 and HiM-50 (left- and right-hand panels, respectively), coloured by central stellar velocity dispersion, σ_e . All quantities are r -band light weighted, measured within 2D projected r_e . Galaxies with $\sigma_e > 80 \text{ km s}^{-1}$ are shown as translucent circles, while those that fall under our ‘mock C13’ selection criteria are shown as opaque circles.

IMF-sensitive stellar absorption features equally well, they yield the same dwarf-to-giant ratio. In that study, this ratio is defined as (their eq. 4)

$$F_{0.5} = \frac{\int_{0.1}^{0.5} M \Phi(M) dM}{\int_{0.1}^{100} M \Phi(M) dM}, \quad (\text{C1})$$

where $\Phi(M)$ is the IMF. However, as noted by Clauwens et al. (2016), spectroscopic IMF studies are not able to constrain the denominator since in most ETGs, the stellar populations are so old that the highest stellar mass remaining is $\approx 1 M_\odot$. Thus, while this definition of $F_{0.5}$ is unique for a given IMF, it is not clear that it is unique for a given mass fraction of dwarf stars at the present day. To investigate this, we employ the definition used by Clauwens et al. (2016), where the ratio is instead defined with respect to stars with $m < 1 M_\odot$, $F_{0.5,1}$ (equation 4). $F_{0.5,1}$ is a more accurate representation of the present-day dwarf-to-giant mass ratio, to which spectroscopic IMF studies are sensitive. If it is true that as long as the present-day IMF-sensitive spectral features are well-fit, the choice of IMF parametrization does not affect $F_{0.5}$, it must also be true that it does not affect $F_{0.5,1}$.

In Fig. C1, we plot $F_{0.5}$ as a function of $F_{0.5,1}$ for a variety of IMF parametrizations. Note that both of these quantities are unique for a given IMF. As orange and red solid lines we show results for the LoM and HiM IMF parametrizations, respectively, where the range of values plotted corresponds to the ranges over which the IMFs are allowed to vary in our model (see Section 2.2). As expected, for LoM the dwarf-to-giant ratio spans a wide range under either definition, with both increasing with the steepening of the low-mass slope. HiM also decreases for both dwarf-to-giant ratio definitions as the high-mass slope becomes shallower, but follows a much steeper trend than in the LoM case. This shows that $F_{0.5}$ is much more sensitive to changes in the high-mass slope than is $F_{0.5,1}$. This result indicates that, at a fixed present-day dwarf-to-giant ratio, $F_{0.5}$ is sensitive to the parametrization of the IMF.

To further illustrate this point, we also include in Fig. C1 results for the bimodal IMF parametrization of Vazdekis et al. (1996), where we plot results for the range of high-mass slopes recovered for high-mass ETGs found by La Barbera et al. (2013), corresponding to $\Gamma_b = 1$ to 3 (or $x_{\text{Bimodal}} = -2$ to -4 according to the IMF slope convention used in our paper). Unsurprisingly, the bimodal IMF line follows the HiM trend and extends it to higher dwarf-to-giant ratios as the high-mass slope steepens. We also include the same trend for a single power-law (or ‘unimodal’) IMF in brown. Here we also show results for slopes recovered for high-mass ETGs by La Barbera et al. (2013), with $\Gamma = 0.8$ to 2 (or $x = -1.8$ to -3). The trend between $F_{0.5}$ and $F_{0.5,1}$ is also monotonic for this prescription, but again follows a separate track.

It is clear that, for a given IMF parametrization, $F_{0.5}$ is a good tracer of the present-day dwarf-to-giant ratio. However, for a fixed present-day dwarf-to-giant ratio, $F_{0.5,1}$, the corresponding zero-age value, $F_{0.5}$, is extremely sensitive to the choice of IMF parametrization. This is particularly true at $F_{0.5,1} \approx 0.7$, where the range of $F_{0.5}$ values ranges from sub-Chabrier to super-Salpeter, depending on the IMF parametrization employed.

It is thus interesting that La Barbera et al. (2013) find that models with different IMF parametrization are consistent in $F_{0.5}$. However, at fixed $F_{0.5}$, the difference in $F_{0.5,1}$ values between bimodal and unimodal IMFs is not large, and is comparable to the typical differences in $F_{0.5}$ seen between these two IMF parametrizations found by La Barbera et al. (2013). This is because IMF prescriptions that vary the high-mass slope are more sensitive to $F_{0.5}$ than to $F_{0.5,1}$, as can be seen by the steep slopes of the bimodal and unimodal lines in Fig. C1. It would thus be interesting to see if spectroscopic analyses would still yield consistent $F_{0.5}$ values under the assumption of a LoM-like IMF parametrization, as the differences could be greater.

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

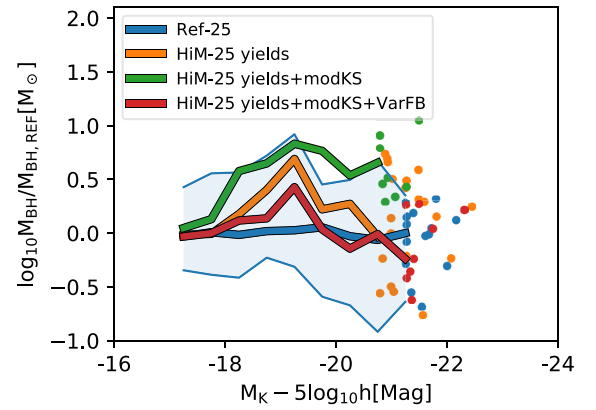
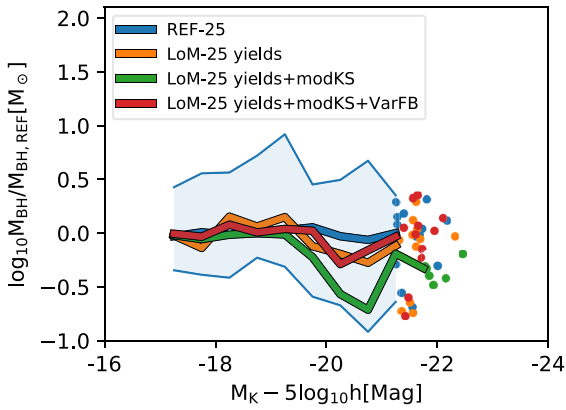
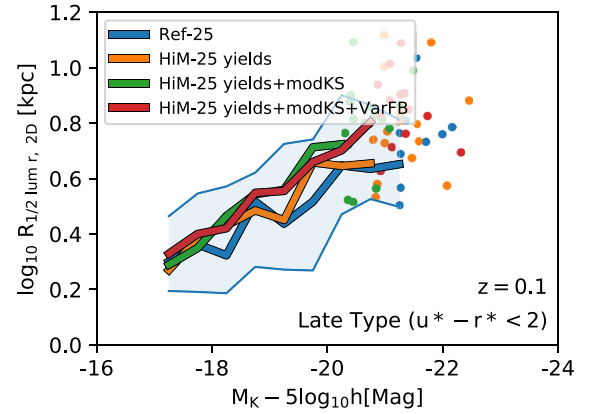
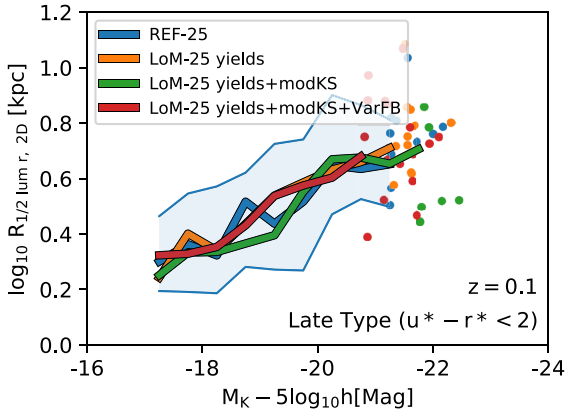
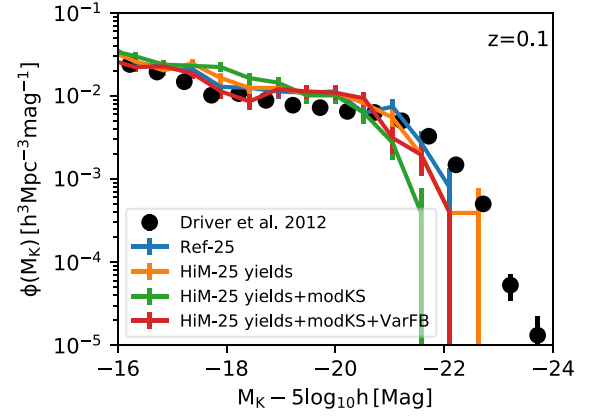
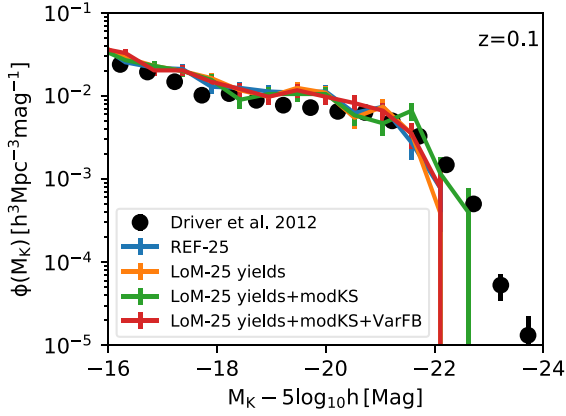


Figure B1. As in the left column of Fig. 8 but for simulations with the same model as LoM-50 but for a $(25 \text{ Mpc})^3$ volume [rather than $(50 \text{ Mpc})^3$]. In blue we show the reference 25 Mpc box (Ref-25), while other colours show simulations with LoM, adding new effects of the variable IMF one at a time. In orange we show the effect of only changing the stellar yields while keeping everything else fixed at the reference (Chabrier) prescription. In green we self-consistently modify the star-formation law. In red we further make the stellar feedback self-consistent with the variable IMF, this being our fiducial LoM model. The left-hand panel shows the K -band luminosity function. The middle panel shows half-light radius as a function of K -band absolute magnitude. To clarify the deviations of M_{BH} from the Ref-25 model, the right-hand panel shows the median M_{BH} relative to the median for Ref-25, rather than absolute M_{BH} as in Fig. 8. The ‘yields’ model and our fiducial, self-consistent model match the calibrated values of the Ref model very closely, while the model ‘yields+modKS’ deviates with higher luminosities and lower BH masses.

Figure B2. As in Fig. B1 but for the HiM prescription. Out of these three models, our fiducial, self-consistent model does the best job at matching the reference model and hence the calibration diagnostics, especially for BH masses.

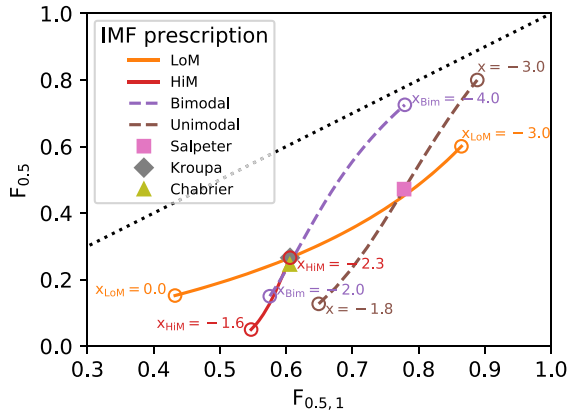


Figure C1. Effect of IMF parametrization on the inferred dwarf-to-giant ratio. We show the mass fraction of stars with $m < 0.5 M_{\odot}$ relative to the total mass of the IMF, $F_{0.5}$ (equation C1), as a function of the same fraction but relative to the mass $< 1 M_{\odot}$, $F_{0.5,1}$ (equation 4). The former is the dwarf-to-giant ratio defined by La Barbera et al. (2013, their equation 4), while the latter is that defined in Clauwens et al. (2016). A dotted black line shows the 1:1 relation. As orange and red solid lines we show the two variable IMF prescriptions used in this work, LoM and HiM, respectively. The dashed purple line shows the bimodal IMF prescription of Vazdekis et al. (1996), while the dashed brown line shows the relation for a single power law IMF slope. The values for Salpeter, Kroupa, and Chabrier IMFs are indicated with filled symbols (see legend). Open circles mark the values of the IMF slope for the range over which the IMF is varied in each parametrization. At fixed $F_{0.5,1}$, the $F_{0.5}$ depends sensitively on the parametrization assumed.