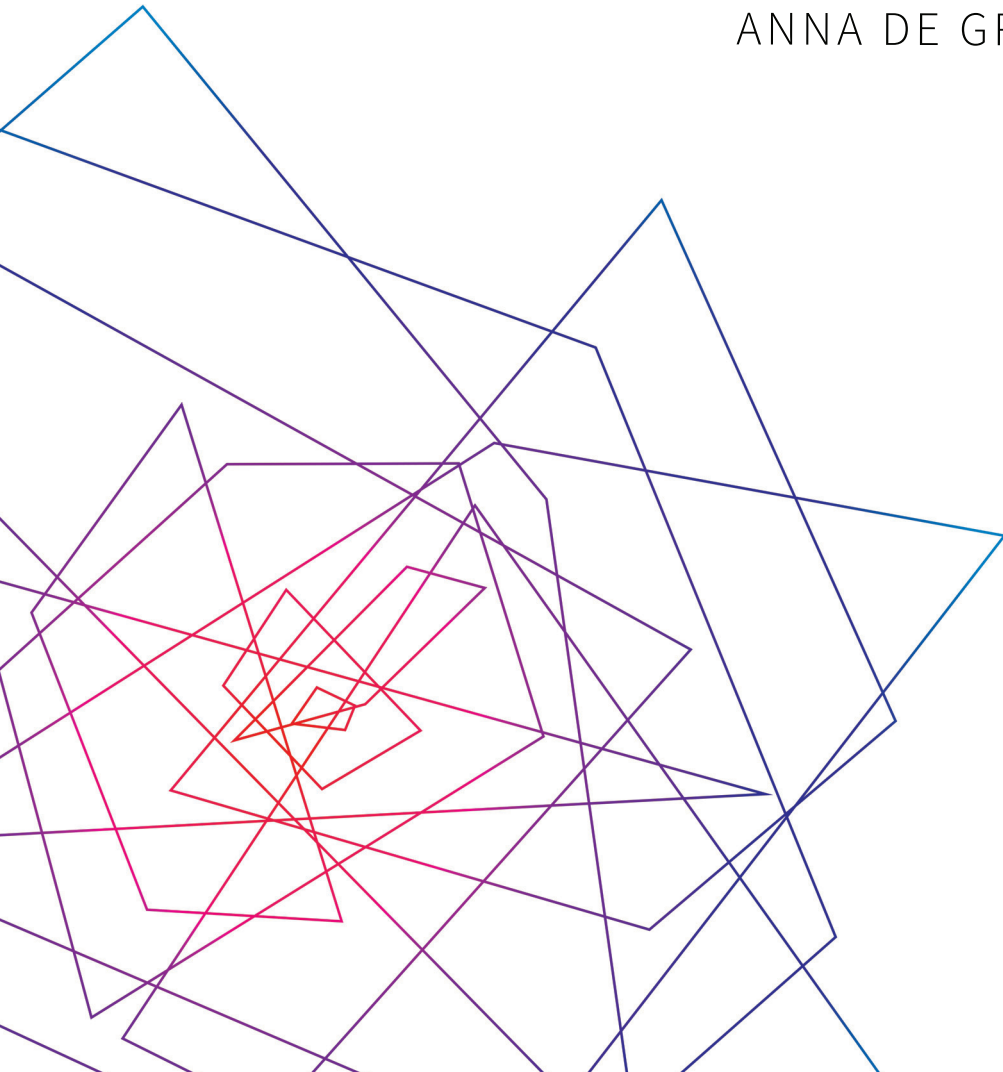# SHAPING MASSIVE GALAXIES

ANNA DE GRAAFF

Shaping Massive Galaxies

- The structural evolution of galaxies across $0 < z < 1$ -

Proefschrift

Anna Geraldine de Graaff

geboren te Aberdeen, Schotland
in 1995

Promotor:	Prof. dr. M. Franx
Co-promotor:	Dr. R. Bezanson

Promotiecommissie:	Prof. dr. H. J. A. Röttgering	
	Prof. dr. M. Kriek	
	Prof. dr. J. Schaye	
	Dr. M. Maseda	(University of Wisconsin-Madison)
	Prof. dr. A. van der Wel	(Universiteit Gent)

Ten thousand flowers in spring, the moon in autumn,
a cool breeze in summer, snow in winter.
If your mind isn't clouded by unnecessary things,
this is the best season of your life.

– Wumen Huikai

CONTENTS

1 | INTRODUCTION

1.1 The rich diversity of galaxies

A zoo of galaxies populates our Universe, of which the Milky Way provides just one, highly complex example. To understand the origins of our own galaxy and to be able to contextualise our place in the Universe, requires the study of nearby stars and distant galaxies.

The detailed mapping of the stars in the Milky Way has revealed that our galaxy alone already consists of multiple, intertwined structures. Whereas the central region comprises mainly old ($\sim 10\,\mathrm{Gyr}$) stars that together form a round bulge, the majority of stars (including the Sun) form a large disc, which is composed of prominent spiral arms and contains many gaseous star-forming regions. The disc can be decomposed into a thick and thin disc, with distinct differences in the ages and chemical compositions between the two groups of stars that make up these structures. Moreover, the bulge and disc are again encompassed by a stellar halo that is spherical in shape. These different observations all point toward a complex, multi-phase assembly history of the galaxy (for a review, see Bland-Hawthorn & Gerhard 2016), which still forms an active field of research.

Further from home, a new area of astronomy was opened up with the discovery of other galaxies in the nearby Universe. Over the past centuries, astronomers have catalogued thousands of objects in the night sky, perhaps one of the most famous of which being the Andromeda nebula. Although the nature of Andromeda and other nebulae had been subject to debate for longer, it was only one hundred years ago that Edwin Hubble discovered, based on the first accurate distance estimates, that Andromeda is in fact not a nebula, but an entirely different galaxy. Tremendous progress has since been made in the field of extragalactic astronomy. The few galaxies originally documented have been superseded by large imaging surveys with powerful telescopes that have by now catalogued millions of galaxies in the local Universe and much beyond.

Although the formation path of the Milky Way is unique, efforts to survey the night sky have revealed a large number of galaxies with similar properties to our own galaxy. Specifically, these galaxies are disc-like in shape and have ongoing star formation, and are therefore observed to have blue colours (Roberts & Haynes 1994). Despite their similarities, there is significant variation as well, as galaxies span a wide range in luminosity and size (Blanton et al. 2003). Moreover,

the morphologies of some can be described as pure discs, whereas others also have red central bulges that can comprise nearly half of the total stellar mass. Nevertheless, these galaxies are commonly grouped into a single class of 'spiral' or 'late-type' galaxies, a classification that was proposed by Hubble (1926, 1936) based on morphology alone, or 'star-forming' galaxies when selected by colour.

On the other hand, many of the more luminous galaxies contain no discs at all, and instead are spheroidal in shape, reminiscent of the central bulge of the Milky Way. In addition to their different shapes, these galaxies typically have very red colours, indicating old stellar populations in which no new stars are formed, and have light distributions that are highly centrally concentrated (e.g., Kauffmann et al. 2003). These objects are often categorised as 'elliptical' or 'early-type' galaxies on the Hubble sequence, a misnomer given that the red galaxies must represent an evolutionary phase that follows after the 'late-type' epoch of star formation, or are labelled 'quiescent' galaxies.

How we have arrived at the present-day Universe with its great diversity of galaxy shapes, morphologies, and colours is a challenging question. To grasp not only the differences, but also the similarities between different types of galaxies requires observations of local galaxies and their distant counterparts to piece together the average evolution histories. Remarkably, the foundations for the formation of galaxies were already established shortly after the Big Bang. Vast progress has been made in the past century in our understanding of the subsequent 13 billion years of evolution.

1.2   Hierarchical formation of structure

Shortly after the Big Bang all matter in the Universe was extremely dense, hot and ionised, and therefore opaque to radiation. As the Universe expanded and cooled with time, electrons and protons were able to combine and form neutral atoms. The corresponding decrease in free electrons and the greatly expanded space allowed for photons to stream freely for the first time (Peebles 1968), approximately $380\,000$ years after the Big Bang. Radiation from this epoch permeates the entire Universe and is still observable today, although strongly cooled down, and known as the cosmic microwave background (CMB; Penzias & Wilson 1965; Dicke et al. 1965).

The CMB is highly homogeneous and isotropic, with a spectral shape that is a near-perfect black body of temperature $T = 2.73\,\mathrm{K}$ (Fixsen et al. 1996; Bennett et al. 2013; Planck Collaboration et al. 2020). Deviations from this temperature across different regions of the sky are of order $\Delta T/T \sim 10^{-5}$, and are caused by very small fluctuations in the underlying matter density that were already present in the early Universe.

Encoded in the details of these temperature fluctuations is critical information on the nature of the matter in the Universe (e.g., White et al. 1994; Scott et al. 1995). The magnitude of the fluctuations and the associated spatial scales indicate that the Universe is composed of not only ordinary baryonic matter, but also non-baryonic 'dark' matter. The nature of dark matter is still unknown, but its inferred

properties are well-described by particles that are non-baryonic, weakly-interacting and had relatively low thermal velocities shortly after the Big Bang, and therefore referred to as cold dark matter (CDM).

The leading cosmological model for the last two decades has been the flat ΛCDM model, which describes the expansion history of the Universe and the time evolution of the energy densities of radiation, baryonic matter, cold dark matter, and the cosmological constant (Λ). This constant, a mysterious component that is also referred to as dark energy, is negligible in the early Universe, but comprises the majority of the energy density in the late-time Universe and governs the current accelerated expansion (Riess et al. 1998; Perlmutter et al. 1999).

Importantly, fits of the ΛCDM model to full-sky maps of the CMB have shown that there is approximately five times more dark matter than baryonic matter (Bennett et al. 2013; Planck Collaboration et al. 2020), which greatly affects the formation of structure. Whereas the baryons and photons were coupled in the early Universe, forming a fluid with complex oscillatory behaviour (Peebles & Yu 1970), the weakly-interacting dark matter was able to clump together by gravitational attraction. Very small primordial fluctuations in the matter density therefore grew into increasingly large density perturbations: this is the foundation for the hierarchical formation of structure.

Eventually, density perturbations became large enough that the local overdensities underwent gravitational collapse, forming dark matter haloes (e.g., Peacock 1999, Chapter 15, 17). These dark matter haloes formed the building blocks for galaxy formation, as the baryonic gas inside these haloes was able to cool and condense further. The first stars formed from this cooled gas, likely around ~ 100 million years after the Big Bang, giving rise to the very first galaxies.

At the present day, the initial density perturbations have grown into not only very large haloes, but also filamentary and sheet-like structures that span scales of $1 - 100\,\mathrm{Mpc}$. These structures are not distributed randomly, but form a cosmic web of dark and baryonic matter. Galaxies lie in the nodes of this web: rather than being islands of luminous stars, their evolution is strongly connected to the surrounding environment, as galaxies grow by accreting material from the cosmic web and through mergers with nearby haloes. Unlike dark matter, which can be described by the laws of gravity alone, the baryonic matter is subject to complex physical processes. How galaxies grew from small pockets of gas into the diverse stellar structures seen in the local Universe is a major question that is yet to be fully understood.

1.3   Assembly and evolution of massive galaxies

1.3.1   Probing galaxy evolution

Observations in the local Universe have provided great insight into the long evolutionary histories of galaxies. Different galaxy types have been found to correlate with not only morphology and colour, but also various spectral features, the galaxy dynamics and the local environment. From here a broad picture of the evolution

of galaxies has emerged, in which star-forming galaxies gradually grow in both size and mass, forming new stars as gas from the cosmic web accretes onto the disc. On the other hand, massive elliptical galaxies are thought to have gone through a more brief, intense period of star formation, followed by a quiescent phase in which mass growth occurred through mergers with neighbouring galaxies.

However, the archaeological research of nearby galaxies struggles to capture the details of these processes, as the characteristics of some physical mechanisms may be washed out over time or be degenerate with those of other mechanisms (e.g., the spectrum of an old stellar population of solar metallicity is extremely similar to that of a younger, but high-metallicity population of stars). Observations of more distant galaxies, i.e. at higher redshifts, are more difficult to obtain, but offer a necessary complement to the local studies, as they probe the physical state of galaxies at earlier epochs in the history of the Universe. The average evolution between galaxies at different redshifts can then be inferred statistically.

Furthermore, theoretical models allow to directly trace the evolution histories of individual galaxies. Hydrodynamical simulations are of particular interest, as these model the non-linear formation of structure and the formation of galaxies from dark and baryonic matter within large volumes (up to $\sim 10^6 \, \mathrm{Mpc}^3$), and thus have the ability to model multiple galaxies and their environments. These models help to assess our current theoretical understanding and provide critical insights into the intricate physical mechanisms that shape galaxies, assisting in the interpretation of our observations.

1.3.2   From light to mass

A major difficulty in the observations, regardless of redshift, stems from the fact that we are only able to measure the light emitted by galaxies. Measuring the mass distributions of galaxies from their light is critical to map the assembly histories of galaxies. The overall mass scale and its spatial distribution have to be inferred by exploiting measurements from imaging and spectroscopic data.

The sizes and shapes of galaxies are the easiest to estimate, as these are typically measured by fitting parametric models (Sérsic profiles; Sersic 1968) to single-band imaging at a rest-frame optical wavelength. These measurements, and their correlations with colour, reflect the stellar mass distributions within galaxies, and have been shown to quantify the dichotomy between the early- and late-type morphologies (e.g., Shen et al. 2003; Blanton et al. 2003; Kelvin et al. 2012): early-type galaxies typically have Sérsic indices of $n \approx 4$, whereas late-type galaxies have $n \approx 1$. In addition, at fixed luminosity, early-type galaxies are smaller in size than the late types. Moreover, the distributions of the projected axis ratios, quantifying the observed flattening, have been shown to correspond to different 3D shapes for the two populations. Early-type galaxies are typically rounder in shape, especially at high luminosity or stellar mass (e.g., Vincent & Ryden 2005; Padilla & Strauss 2008; van der Wel et al. 2009). In contrast, star-forming or late-type galaxies are well-described by a population of oblate (disc-like) shapes.

Adding in observations at other wavelengths, the colours of galaxies can be used to extract the stellar mass-to-light ratio (M_*/L). The total stellar mass of

a galaxy is then obtained by scaling this ratio with the total measured luminosity. After composing a spectral energy distribution (SED) from multi-wavelength imaging and, if available, rest-frame optical spectroscopy, the M_*/L can be modelled by fitting template spectra of different types of stellar populations. These templates are constructed from stellar population models that follow different star formation histories, and thus have different present-day star formation rates, ages and metallicities (e.g., Bruzual & Charlot 2003; Maraston 2005). Including the effects of dust, and under the assumption of a constant initial mass function (IMF), these models are typically described by at least 4, but ranging up to ≈ 20 free parameters (for a review, see Conroy 2013). For a sufficiently flexible model and wide parameter space, as well as high-quality data (i.e., a SED that spans a broad wavelength range at high signal-to-noise ratio; SNR), the M_*/L and stellar mass can be estimated with high precision, although the systematic uncertainties in the models can be as large as a factor of 2 (Conroy et al. 2009).

However, the majority of the mass in the Universe is not luminous, such as dark matter or gas, and is therefore not directly observable. Nevertheless, measurements of the internal dynamics of the stars offer a way to estimate the *total* mass within a galaxy. By means of the virial theorem, the mass enclosed within a radius r for galaxies that are in virial equilibrium is proportional to the mean square speed (σ) of the stars (Binney & Tremaine 1987, Chapter 4):

$$M(<r) \;\propto\; r\sigma^2 \,. \tag{1.1}$$

Typically, σ is approximated by the velocity dispersion (line width) of the stellar absorption lines in galaxy spectra, which are spatially integrated and projected along the line of sight. The mass obtained in this way is referred to as the dynamical mass, and is sensitive to all (dark and baryonic) mass within the radius r, as σ traces the scale and shape of the total gravitational potential.

To obtain a total mass estimate, requires a constant of proportionality that takes into account the 3D structure of the galaxy and the effect of the line-of-sight projection on σ. In practice, however, this comparison between dynamical and total masses is very challenging, as independent measurements of the total mass are available for a only a few dozen strongly lensing galaxies (Bolton et al. 2008; Koopmans et al. 2009). Instead, other studies have used calibrations based on dynamical Jeans models or empirical trends with Sérsic index (e.g., Cappellari et al. 2006; Taylor et al. 2010).

1.3.3   Measuring mass assembly

Large galaxy surveys, such as the Sloan Digital Sky Survey (SDSS; York et al. 2000) or the Cosmic Evolution Survey (COSMOS; Scoville et al. 2007), have obtained multi-wavelength photometry and spectroscopy for large numbers of galaxies. These data have provided stellar mass estimates, structural parameters, and velocity dispersions for $\sim 10^6$ galaxies, the majority of which are at $z \sim 0$, but with statistical samples ranging up to $z \sim 4$. Studies of the distributions of these galaxy properties and their evolution with redshift have attempted to reconstruct how galaxies built up their stellar mass.

The galaxy stellar mass function, describing the galaxy number density as a function of stellar mass, has been shown to evolve significantly between $0 < z < 4$, as the galaxy number density increases strongly toward lower redshift across the entire mass range (e.g., Marchesini et al. 2009; Ilbert et al. 2013; Muzzin et al. 2013a). However, when dividing the sample by colour, it becomes apparent that for star-forming galaxies only the less massive systems (i.e., below the 'knee' of the mass function, $\log(M_*/M_\odot) \lesssim 10.8$) grow significantly in mass with cosmic time. Quiescent galaxies, on the other hand, are extremely scarce at high redshifts, but their number density increases strongly with time, fully dominating the mass growth at the high mass end. At the present day, approximately 50% of all stellar mass is contained within quiescent galaxies (Bell et al. 2003). Crucially, these measurements demonstrate a correlation between the stellar mass and the galaxy type or star formation history.

Not only the mass functions, but also the distributions of the structural properties change with redshift. The number of galaxies with high Sérsic indices decreases strongly toward high redshift (Chevance et al. 2012), and high-redshift galaxies are therefore predominantly systems with exponential surface brightness profiles ($n \approx 1$). High-redshift galaxies, in particular the quiescent systems, are also significantly smaller in size (e.g., Trujillo et al. 2007; Franx et al. 2008). The shape distributions change as well: the high-redshift star-forming population contains a significant fraction of spheroidal and prolate shapes, in addition to a large fraction of oblate systems, particularly at lower masses (van der Wel et al. 2014b). Strikingly, the high-redshift population of quiescent galaxies contains a substantial fraction of oblate systems, which at the high mass end is a factor of three higher than observed in the local Universe (Chang et al. 2013).

To connect the multitude of observations into a coherent picture, requires a statistical framework. Galaxy scaling relations, and the redshift evolution thereof, can provide such a framework. A key relation is the strong correlation between the stellar masses and sizes of galaxies, which has been found to have a steeper slope and lower overall normalisation for quiescent galaxies, than the stellar mass-size relation of star-forming galaxies (e.g., Shen et al. 2003; Lange et al. 2015). The time evolution of this scaling relation also differs for the two populations, with the inferred average size growth of star-forming galaxies being broadly consistent with the gradual accretion of mass from the cosmic web, as described by the ΛCDM model. Quiescent galaxies, however, experience rapid evolution in their size at fixed mass, particularly at $z < 2$ (e.g., van der Wel et al. 2014a). This indicates a different growth mechanism for quiescent galaxies, likely caused by repeated merging with relatively small satellite galaxies (Bezanson et al. 2009; Naab et al. 2009), which may also explain the observed evolution in the Sérsic index and flattening.

Furthermore, stellar population properties have also been shown to correlate with stellar mass. More massive galaxies are typically older, have lower specific star formation rates, and higher stellar metallicities (e.g., Kauffmann et al. 2003; Brinchmann et al. 2004; Gallazzi et al. 2005). For some massive early-type galaxies, the central regions have been found to have relatively high α-element abundances (Thomas et al. 2005), indicative of a very rapid period of star formation. These

observations led to the idea of 'down-sizing' in the mass assembly of galaxies (e.g., Cowie et al. 1996; Thomas et al. 2005; Graves et al. 2007), in which more massive galaxies accumulated the majority of their stellar mass earlier in the history of the Universe. This may appear counter-intuitive, given that the hierarchical formation of structure implies a bottom-up scenario in which the smallest structures formed earliest. The two mechanisms are mutually compatible, due to the fact that the haloes in which galaxies form need to be sufficiently large to allow for gas to cool and star formation to initiate, and that massive galaxies ceased forming stars relatively early as the result of complex feedback processes (Neistein et al. 2006).

1.3.4   The Fundamental Plane

The measurements of the evolution of galaxies have shown that, as galaxies grow in stellar mass with time, the stellar population properties and structural properties change accordingly. The coupling between these processes, however, is still not well understood. Importantly, by only mapping the evolution in the stellar mass, our picture is incomplete.

Simulations of galaxy formation and evolution predict that the manner by which mass is assembled impacts the galaxy structure, as mergers and in-situ star formation (and the efficiency thereof) can leave different signatures. As a result, not only the stellar mass, but also the dark matter distributions within galaxies are expected to differ for different types of galaxies (e.g., Hopkins et al. 2008).

Valuable insight may therefore come from the Fundamental Plane (FP), which provides a framework that incorporates the dynamical masses of galaxies, and thereby connects stellar population properties and galaxy structure in a simplistic way. The FP was originally found as the empirical, planar scaling relation between the galaxy size, velocity dispersion, and surface brightness of local early-type galaxies (Djorgovski & Davis 1987; Dressler et al. 1987):

$$\log r_{\mathrm{e}} = a \log \sigma + b \log I_{\mathrm{e}} + c \,, \tag{1.2}$$

where r_{e} is the effective radius and I_{e} the mean surface brightness within r_{e}. The coefficients a and b describe the tilt of the plane, and c the zero point.

Although initially used as a distance indicator for nearby elliptical galaxies, the FP is also of astrophysical interest, as it is rooted in the scalar virial theorem, but differs slightly from this simple prediction: from the virial theorem we might expect $a = 2$ and $b = -1$, whereas observational studies local galaxies find $a \approx 1.4$ and $b \approx -0.8$ (e.g., Jorgensen et al. 1996; La Barbera et al. 2008; Hyde & Bernardi 2009). This deviation has strong physical implications for the properties and evolution of early-type galaxies.

Crucially, the three coefficients of the FP can be interpreted in terms of the dynamical mass-to-light ratio, M/L (Faber et al. 1987). The fact that the tilt of the observed FP differs from the expected virial plane suggests a dependence of M/L on mass for early-type galaxies. Several physical mechanisms have been proposed to explain this observed variation in M/L.

Systematic variations in the stellar populations of early-type galaxies, and thus in M_*/L, as a function of mass have been shown to account for only half of the

deviation of the observed FP from the virial plane (Hyde & Bernardi 2009; Graves & Faber 2010; La Barbera et al. 2010a). Therefore, the second half must be driven by variations in the galaxy structure as a function of mass: this can be either due to variation in the 3D stellar (kinematic) structure, also referred to as the non-homology of galaxies, or a systematic dependence of the dark matter mass fraction within galaxies on mass.

Non-homology is found by several studies to have a relatively small effect on the tilt (Cappellari et al. 2006; Bolton et al. 2007, 2008), although contrary results have also been reported (Bender et al. 1992; Graham & Colless 1997; Prugniel & Simien 1997; Trujillo et al. 2004; Desmond & Wechsler 2017). On the other hand, variations in the dark matter content are extremely difficult to measure observationally, but theoretical models have provided interesting insight. Simulations of galaxy mergers have shown that the mass dependence of M/L arises naturally among the descendants of major, dissipational galaxy mergers, as these types of mergers alter the baryonic and dark matter mass profiles (Boylan-Kolchin et al. 2006; Robertson et al. 2006; Hopkins et al. 2008). This possibly implies that all galaxies on the observed FP have a similar formation history. Within observational studies there is still debate regarding this interpretation, however, partially due to the large measurement uncertainty on the IMF, which is degenerate with variations in the dark matter fraction (e.g., Graves & Faber 2010; Bernardi et al. 2018).

Furthermore, not only the tilt, but also the zero point of the FP is directly related to M/L, as $c \propto -\log(M/L)$. The scatter about the zero point has been found to exceed the scatter expected due to measurement uncertainties, which implies that the FP is intrinsically not a true plane, but is better described by a dense cloud (Jorgensen et al. 1996). The offsets from the FP can therefore be interpreted as a variation in M/L, and have been shown to correlate with the stellar ages, metallicities, and α-element abundances of local early-type galaxies (Forbes et al. 1998; Gargiulo et al. 2009; Graves et al. 2009). Moreover, the scatter has been found to correlate with galaxy structure, specifically the Sérsic index and the dynamical-to-stellar mass ratio, which describes the amount of 'dark mass' within galaxies from either dark matter or missing stellar mass due to an incorrectly assumed IMF (Graves & Faber 2010; Bezanson et al. 2015).

By studying the details of the low-redshift FP, i.e. its tilt and scatter, we gain insight into the formation of local early-type galaxies. For instance, it has been shown that the combination of trends in the stellar population properties and structural properties throughout the FP can be explained by differences in the truncation times of the star formation in these galaxies, although dissipational mergers may likely also play a role (Gargiulo et al. 2009; Graves & Faber 2010).

Moving beyond the local Universe, the evolution of the FP provides an interesting metric of the time evolution of M/L for quiescent galaxies. If we assume that galaxies with old stellar populations evolve passively, which implies that they do not form or accrete new stars, M remains approximately constant. Measurements of the zero point of the FP at higher redshifts, and hence the change in M/L, then provide direct constraints on the formation epoch of the most massive quiescent galaxies (Franx 1993; van Dokkum & Franx 1996). Various results have

been reported that suggest quiescent galaxies formed very early, at $z \sim 3$, before the peak of star formation in the rest of the Universe (e.g., Treu et al. 2005; van der Wel et al. 2005; van Dokkum & van der Marel 2007; van de Sande et al. 2014).

In addition, different studies have claimed to measure a change in not only the zero point, but also the tilt of the high-redshift FP with respect to $z \approx 0$ (e.g., di Serego Alighieri et al. 2005; Jørgensen & Chiboucas 2013), although these results have been subject to debate (Holden et al. 2010; Saglia et al. 2010, 2016). These studies suggest that galaxies at higher redshift may follow a different relation between M/L and mass, which can be due to evolution in the stellar populations or the structures of high-redshift galaxies. The limited availability of measurements at $z > 0$ prevents strong conclusions to be drawn, but indicate that improved high-redshift measurements of the FP can add significantly to our understanding of the evolution of galaxies.

1.3.5   Towards a holistic view of the FP

The FP has been demonstrated to provide a powerful probe of galaxy evolution. However, different studies have shown that the measurements of the FP, and thus the physics inferred from detailed studies of the relation, depend strongly on the selection of the galaxy sample and the measurements used (e.g., the chosen photometric band; Hyde & Bernardi 2009; La Barbera et al. 2010a). For instance, galaxies in clusters populate the FP differently than galaxies in the field (La Barbera et al. 2010b; Saglia et al. 2010). Similarly, a sample of galaxies selected by luminosity results in a different FP than a (stellar) mass-selected sample (e.g., van der Wel et al. 2005).

At low redshift these effects can be measured and taken into account, owing to the wealth of data available from large surveys such as the SDSS. However, selection biases become increasingly problematic toward higher redshift. The bottleneck is the measurement of the stellar velocity dispersion, which requires high SNR spectroscopy at rest-frame optical wavelengths in order to measure robust absorption line profiles. Due to the surface brightness dimming of distant galaxies, and the shifting of prominent absorption lines toward longer wavelengths, it is difficult to obtain the necessary high-quality data. Moreover, to measure the sizes of distant galaxies requires high-resolution imaging, which can currently only be achieved with the *Hubble Space Telescope* (HST) that is not affected by the turbulent atmosphere of the Earth.

Therefore, to make the expensive observations more efficient, high-redshift measurements have typically focused on luminous quiescent galaxies in dense environments, or very bright objects in the field. In this way, data for a few hundred quiescent galaxies has been collected over the past two decades (e.g., van der Wel et al. 2005; Treu et al. 2005; Saglia et al. 2010; Holden et al. 2010; Jørgensen & Chiboucas 2013; van de Sande et al. 2014). Although this has provided strong constraints on the evolution of the most massive systems, particularly those in cluster environments, these are not representative of the broader population of massive galaxies. An open question is therefore how the average quiescent galaxy evolves ($\sim L^*$ galaxies with stellar masses similar to the Milky Way by $z \approx 0$), and

how this evolution compares with the measurements of the most massive galaxies.

Furthermore, so far nearly all studies of the FP have focused on early-type or quiescent galaxies alone. Samples are often specifically selected to include only galaxies that have high Sérsic indices, are round in shape, and red in colour with no significant Balmer emission lines. Star-forming galaxies are usually considered separately, and have been shown to obey their own dynamical scaling relation: there is a tight correlation between the rotation speed and luminosity or stellar mass of star-forming discs (Tully & Fisher 1977). Unlike the FP, this Tully-Fisher relation is a linear relation and independent of the galaxy size (Zwaan et al. 1995; Courteau & Rix 1999).

Few studies have shown, however, that both galaxy populations may be reconciled within one planar relation (Zaritsky et al. 2006, 2008). Two conditions need to be met in order to place star-forming and quiescent galaxies onto a single dynamical scaling relation: (i) in addition to the dispersion due to random motions of stars, the measure of σ has to explicitly include the line broadening due to rotation, and (ii) the differences in M_*/L need to be modelled and taken into account. These modifications result in the stellar mass FP, which differs from the luminosity FP by the replacement of the surface brightness with the stellar mass surface density (Hyde & Bernardi 2009). At low redshift, this has resulted in the finding of common stellar mass FP for all galaxies (Bezanson et al. 2015). There is also evidence that suggests this common relation may hold beyond $z \sim 0$, but a larger sample of star-forming and quiescent galaxies is needed to assess the statistical significance and physical implications of this finding.

These results are potentially very promising, as they allow to study the galaxy population as a whole. Instead of focusing solely on the differences between galaxy types, a unified scaling relation may also highlight their similarities, and thereby shed light on the coupling between different physical properties of galaxies. However, the interpretation of the stellar mass FP is challenging, due to the more complex range of structures and star formation histories that the star-forming galaxies introduce, and the difficulty in measuring the dark matter content in galaxies. Therefore, in addition to observations across a wide redshift range, theoretical models are needed to map the details of the FP and evaluate possible roles of different physical mechanisms.

1.4   Thesis summary

In this Thesis we explore the evolution in the M/L of massive galaxies in the context of the Fundamental Plane. By disentangling the effects of varying stellar population properties and structural properties, we evaluate what drives the observed variation in the M/L and assess how this evolves with time. This work is built on a combination of observational data and theoretical models, to form a holistic picture of the FP. Deep spectroscopic surveys are used to construct a representative sample of massive quiescent *and* star-forming galaxies across $0 < z < 1$. Cosmological simulations are used to assess the physical properties that may underlie the observed FP.

Chapter 2 presents a comprehensive analysis of the FP at $z \approx 0.8$, using a large sample of quiescent and star-forming galaxies. We base our sample on the LEGA-C Survey, a large spectroscopic survey of ~ 3000 $K_{\rm s}$-band selected galaxies at $0.6 < z < 1.0$ that provides very deep, rest-frame optical spectra. Further complemented by multi-wavelength photometry that covers ultra-violet to mid-infrared wavelengths, as well as high-resolution HST imaging, we obtain rest-frame colours and structural parameters. We perform SED modelling to estimate stellar masses and stellar population properties for the LEGA-C galaxies.

The constructed sample of 1419 galaxies is selected by the estimated stellar mass ($M_* > 10^{10.5}\,{\rm M_\odot}$), and covers a wide range in morphology, star formation activity and environment. We measure the rest-frame g-band FP spanned by this sample, and show that there is evidence for a slight evolution in the tilt of the FP with respect to $z \approx 0$, after accounting for selection effects. Examining the differences between the star-forming and quiescent population, we find that the two populations are distributed differently in the parameter space of the FP, as the star-forming population is offset from the quiescent population and shows much larger scatter. This reflects significant differences in the M/L, and we demonstrate that this is primarily due to differences in the stellar age and star formation activity, and to a lesser extent the dust attenuation.

However, by explicitly accounting for the differences in M_*/L among the sample, we find that both star-forming and quiescent galaxies follow the same stellar mass FP at $z \approx 0.8$. We show that the scatter about the stellar mass FP is approximately equal for both populations, forming an intrinsically tighter relation than the luminosity FP. Nevertheless, the scatter is still greater than the measurement uncertainties alone, which suggests there are systematic variations in M/M_* within the scatter. We examine whether the remaining scatter correlates with physical properties, but find no correlations with stellar population properties and the local overdensity, and only a weak dependence on the morphology, despite the strong non-homology of the sample. Overall, we show that, at fixed size and velocity dispersion, differences in M_*/L (for an assumed, universal IMF) across the sample account for approximately 54% of the variation in M/L, which implies that remainder must be caused by systematic fluctuations in the dark matter content or variations in the IMF between galaxies.

In **Chapter 3** we combine the LEGA-C data with a low-redshift sample from the SDSS to measure the redshift evolution in M/L and M/M_* across $0 < z < 1$. We find rapid evolution in the zero point of the g-band FP, and thus in M/L, which is stronger for quiescent galaxies than for star-forming galaxies. In comparison with previous studies, the evolution in M/L of quiescent galaxies is significantly stronger, which we show is largely due to the inclusion of less massive galaxies ($M_* < 10^{11}\,{\rm M_\odot}$) in the LEGA-C sample, with an additional weak effect from the large fraction of field galaxies rather than cluster galaxies.

On the other hand, the stellar mass FP is remarkably stable across cosmic time, as we measure no significant evolution in the zero point. At fixed size and velocity dispersion, the structural evolution in M/M_* is therefore negligible, and implies that the observed evolution in M/L must be caused by a combination of evolution in the stellar populations and the effects of progenitor bias. The fact that

star-forming and quiescent galaxies lie on the same, stable scaling relation across $0 < z < 1$, also suggests that any evolution in the size or velocity dispersion has to occur along the stellar mass FP: as galaxies grow in size or undergo a morphological transformation, there must be a corresponding change in the dynamical structure, such as to remain on the stellar mass FP.

In **Chapter 4** we evaluate how a fair comparison can be made between observations and cosmological simulations. As simulations model mass in the form of particles, which distinguish between gas, stars, dark matter and black holes, they cannot be directly compared with measurements based on the observed light from galaxies. Moreover, measurements extracted from simulations typically differ fundamentally from the methods used to analyse telescope data. We therefore use the stellar mass-size relation in the EAGLE cosmological simulations to examine the magnitude of these effects.

We create mock SDSS images of the projected stellar mass distributions and optical light distributions of $z = 0.1$ galaxies in the $100^3\,\mathrm{Mpc}^3$ simulation, and apply standard observational methods to measure the sizes, Sérsic indices and projected axis ratios. The use of a measurement method that is consistent with observational rather than other theoretical works, leads to a $0.06\,\mathrm{dex}$ difference in the inferred sizes of galaxies. Most importantly, we find a strong difference between the measured half-light radii and the half-mass radii, which is on average $0.1\,\mathrm{dex}$, but can be as large as $0.5\,\mathrm{dex}$, depending on the star formation activity, mean stellar age, and dust attenuation.

The stellar mass-size relation obtained with these mock observations is in significantly better agreement with the observed scaling relation than is the case for the size measurements that are conventional in theoretical studies. On the other hand, we find that the mock observations also reveal strong differences between simulated and real galaxies, as the distributions of the Sérsic indices and projected axis ratios differ substantially. This discrepancy is contrary to previous studies that used different measurement methods and found good agreement between the morphologies of simulated and real galaxies. Our work therefore highlights the importance of constructing mock observations and applying common observational analysis methods to fairly evaluate cosmological simulations.

Chapter 5 builds on these results, and investigates the physical drivers of a common stellar mass FP for quiescent and star-forming galaxies using the EAGLE simulations. First, we use the total mass enclosed within the effective radius and the stellar velocity dispersion to show that the simulated galaxies obey a tight total mass FP that is very close to the virial theorem. Despite a large diversity in the kinematic structures of the galaxy population, the effects of non-homology on the simulated FP are therefore small ($\lesssim 10\%$). This implies that the dynamical mass is a close approximation of the total mass within the effective radius.

Second, we demonstrate that when we use the stellar mass rather than the total mass, we obtain a stellar mass FP that deviates strongly from the virial plane. We show that this deviation is driven by the dark matter content of galaxies, as the dark matter fraction within the effective radius ($f_{\mathrm{DM}}(< r_{\mathrm{e}})$) is a smooth, power-law function of the galaxy size and stellar mass. Moreover, because it is this smooth variation in the dark matter fraction that sets the coefficients of the stellar mass

FP, both star-forming and quiescent galaxies lie on the same FP, with equally low scatter. We show that the variations in $f_{\mathrm{DM}}(< r_{\mathrm{e}})$ reflect more than just a dependence of M/M_* on mass, which may have been expected based on previous studies, but instead arise primarily from the large variation in r_{e} (and hence f_{DM}) at fixed stellar mass. In turn, the fluctuations in r_{e} are likely the product of multiple, complex physical processes.

Third, using luminosity-weighted mock observations of the size, stellar mass surface density, and the velocity dispersion, we demonstrate that observational biases have a significant influence on the measured coefficients of the stellar mass FP. Moreover, the luminosity weighting strongly increases the scatter in the relation, which we discuss is likely caused by the fact that the most luminous stars lie in dynamically-cold discs, and the resulting velocity dispersion is therefore strongly dependent on the inclination angle. Accounting for these effects, we show that the stellar mass FP in the EAGLE simulation broadly agrees with the observed scaling relation. However, when examined in detail, we find significant discrepancies. This likely reflects a fundamental difference between the 3D mass profiles of simulated galaxies and galaxies in the local Universe, which may arise from inaccuracies in the EAGLE subgrid model and its implementation in the simulation. Therefore, we suggest that the stellar mass FP may offer a new, straightforward measure of success for future cosmological simulations.

2 | THE FUNDAMENTAL PLANE IN THE LEGA-C SURVEY: UNRAVELLING THE M/L VARIATIONS OF MASSIVE STAR-FORMING AND QUIESCENT GALAXIES AT $z \sim 0.8$

ABSTRACT

We explore the connection between the kinematics, structures and stellar populations of massive galaxies at $0.6 < z < 1.0$ using the Fundamental Plane (FP). Combining stellar kinematic data from the Large Early Galaxy Astrophysics Census (LEGA-C) survey with structural parameters measured from deep *Hubble Space Telescope* imaging, we obtain a sample of 1419 massive ($\log(M_*/M_\odot) > 10.5$) galaxies that span a wide range in morphology, star formation activity and environment, and therefore is representative of the massive galaxy population at $z \sim 0.8$. We find that quiescent and star-forming galaxies occupy the parameter space of the g-band FP differently and thus have different distributions in the dynamical mass-to-light ratio ($M_{\rm dyn}/L_g$), largely owing to differences in the stellar age and recent star formation history, and, to a lesser extent, the effects of dust attenuation. In contrast, we show that both star-forming and quiescent galaxies lie on the same mass FP at $z \sim 0.8$, with a comparable level of intrinsic scatter about the plane. We examine the variation in $M_{\rm dyn}/M_*$ through the thickness of the mass FP, finding no significant residual correlations with stellar population properties, Sérsic index, or galaxy overdensity. Our results suggest that, at fixed size and velocity dispersion, the variations in $M_{\rm dyn}/L_g$ of massive galaxies reflect an approximately equal contribution of variations in M_*/L_g, and variations in the dark matter fraction or initial mass function.

Anna de Graaff, Rachel Bezanson, Marijn Franx, Arjen van der Wel,
Bradford Holden, Jesse van de Sande, Eric Bell, Francesco D'Eugenio,
Michael Maseda, Adam Muzzin, David Sobral, Caroline Straatman, Po-Feng Wu

2.1 Introduction

The stellar kinematics, sizes and luminosities of quiescent galaxies are strongly correlated, forming a tight scaling relation known as the Fundamental Plane (FP; e.g., Djorgovski & Davis 1987; Dressler et al. 1987; Jorgensen et al. 1996). Star-forming galaxies, on the other hand, have been shown to follow a linear scaling relation between the galaxy kinematics and luminosity (the Tully-Fisher relation; Tully & Fisher 1977). However, with few modifications to the FP, star-forming galaxies may be found to lie on the same planar scaling relation as the quiescent galaxy population, as was first demonstrated at $z \sim 0$ by Zaritsky et al. (2008). These observations raise the question of how galaxies settle onto the FP at higher redshift, and thus how the positions of galaxies within the FP, both at low and high redshifts, are related to different galaxy properties and their assembly histories.

In the local Universe, galaxies have bimodal distributions in their colours and structures. At high stellar mass, the majority of galaxies have low star formation rates (SFRs) and therefore red colours, in stark contrast with the blue, star-forming population that is dominant at lower stellar masses (Blanton et al. 2003). The colour bimodality becomes even more pronounced after correcting for reddening due to dust (Wyder et al. 2007; Taylor et al. 2015), and is tightly linked with the morphological type (Roberts & Haynes 1994; Kauffmann et al. 2003), as blue galaxies tend to form flattened disks with exponential surface brightness profiles. Red, quiescent galaxies, on the other hand, are rounder in shape and have more centrally concentrated light profiles. The morphological properties are also correlated with the dynamical structure: on average, quiescent galaxies have a lower (projected) angular momentum, with a subset being pressure-supported entirely, whereas the star-forming disks are dynamically cold and supported primarily by rotation (e.g., Romanowsky & Fall 2012; Cappellari 2016; van de Sande et al. 2018).

Moreover, at fixed stellar mass quiescent galaxies are systematically smaller than star-forming galaxies, a result which holds up to $z \sim 3$ (Franx et al. 2008; van der Wel et al. 2014a; Mowla et al. 2019; Suess et al. 2019). The rate of size growth also differs, pointing toward different growth mechanisms for disks (e.g, Mo et al. 1998; Somerville et al. 2008) and spheroids (e.g., Hopkins et al. 2009; Naab et al. 2009; Bezanson et al. 2009). On the other hand, differences in the colours and structures between the two populations begin to fade toward higher redshifts. The bimodality in colour extends at least to $z \sim 3$, but with bluer dust-corrected colours overall and with star-forming galaxies forming an increasingly larger fraction of the total population (Brammer et al. 2009; Whitaker et al. 2011; Muzzin et al. 2013a). Structurally, observations indicate that quiescent galaxies become more similar to the star-forming population at higher redshift, as they are more flattened and have less concentrated light profiles (Chevance et al. 2012; van der Wel et al. 2014b; Hill et al. 2019). Consistent with the observed flattened morphologies, Belli et al. (2017), Toft et al. (2017) and Newman et al. (2018) show that even very massive quiescent galaxies can have significant rotational support at $z \sim 2$, and Bezanson et al. (2018a) find a systematic increase in their rotational support at $z \sim 0.8$ with respect to $z \sim 0$.

Crucially, this leads to the question of how the evolution in colour is coupled to the observed growth in size and change in structure of galaxies. Scaling relations offer a statistical framework within which we can assess the properties of the bimodal galaxy population as well as possible evolutionary mechanisms. For quiescent galaxies, the most commonly studied relation is the FP, which connects the stellar velocity dispersion, effective radius and surface brightness with a remarkably low scatter (e.g., Djorgovski & Davis 1987; Dressler et al. 1987; Jorgensen et al. 1996). The zero point of the FP and the tilt with respect to the virial plane can be interpreted in terms of the dynamical mass-to-light ratio (M_{dyn}/L): the zero point is directly proportional to $\log(M_{\mathrm{dyn}}/L)$ (Faber et al. 1987), whereas the tilt of the FP reflects a dependence of M_{dyn}/L on mass, which can be due to systematic variations in the galaxy structure or the stellar population properties (e.g., Bender et al. 1992; Trujillo et al. 2004; Cappellari et al. 2006; Hyde & Bernardi 2009; Graves et al. 2009; Graves & Faber 2010; Cappellari et al. 2013b).

The low-redshift FP has been used extensively to study the properties and formation of the quiescent population. There is a correlation with stellar age and α-element abundance through the thickness of the FP of early-type galaxies at $z \sim 0$ (e.g., Forbes et al. 1998; Gargiulo et al. 2009; Graves et al. 2009), which Gargiulo et al. (2009) show is consistent with a dissipational merger formation scenario for early-type galaxies. By mapping galaxy properties throughout the FP, Graves & Faber (2010) found that the position perpendicular to the FP depends not only on the star formation history, but also on structural properties, and suggest that the link between these two is most readily explained by differences in the truncation time of star formation, although dissipational mergers may also play a role.

Studies of the FP at different redshifts provide additional constraints on the evolution of quiescent galaxies. The rapid change in the zero point of the FP, corresponding to a strong decrease in M_{dyn}/L toward higher redshift, has been used to estimate the formation epoch of massive quiescent galaxies (e.g., van Dokkum & Franx 1996; van der Wel et al. 2005; van Dokkum & van der Marel 2007; van de Sande et al. 2014). On the other hand, the redshift dependence of the tilt of the FP has been subject to debate, with several authors reporting a rotation in the FP at intermediate redshift with respect to the local FP (e.g. di Serego Alighieri et al. 2005; Jørgensen & Chiboucas 2013; Saracco et al. 2020). Others find no significant change in the tilt after taking into account selection effects (Holden et al. 2010), or only very weak evidence (Saglia et al. 2010, 2016), therefore leading to diverging conclusions on the mass dependence of the rate of change in M_{dyn}/L with redshift (e,g., di Serego Alighieri et al. 2005; Holden et al. 2010), as well as the slope of the stellar initial mass function (IMF; Renzini & Ciotti 1993).

The difficulty of measuring absorption line kinematics for faint sources has thus far restricted studies of the FP at higher redshifts to relatively small numbers of galaxies that are either very bright or reside in high-density environments (e.g., Holden et al. 2010; van de Sande et al. 2014; Beifiori et al. 2017; Prichard et al. 2017; Saracco et al. 2020). van de Sande et al. (2014) demonstrate that, as a result of their selection on luminosity, the colours of their sample are not representative of the main quiescent galaxy population, which steepens the inferred evolution in

M_{dyn}/L if left uncorrected. Moreover, the FP differs for galaxies in clusters and in the field at both low and intermediate redshifts (e.g., La Barbera et al. 2010b; Saglia et al. 2010; Joachimi et al. 2015), due to a systematic difference in age and possibly structure. These selection criteria, in addition to the effect of progenitor bias (van Dokkum & Franx 2001), can therefore lead to a significant bias in the inferred evolution of quiescent galaxies. The effects of selection biases are often difficult to model, however, particularly when the sample size is small.

Interestingly, Zaritsky et al. (2008), Bezanson et al. (2015), and more recently Aquino-Ortíz et al. (2020) have demonstrated that star-forming and quiescent galaxies may lie on the same planar scaling relation at low redshift, provided that both the stellar mass-to-light ratios (M_*/L) and rotation velocities are taken into account. The tilt and zero point of the mass FP, which is obtained by substituting the surface brightness in the luminosity FP with the stellar mass surface density, therefore appear to be insensitive to the significant variation in galaxy colour and structure. Bezanson et al. (2015) show that this result likely holds out to $z \sim 1$, although with a different zero point from the mass FP at $z \sim 0$. In apparent tension with observations of the Tully-Fisher relation of star-forming galaxies, which is independent of the galaxy size or surface brightness (e.g., Zwaan et al. 1995; Courteau & Rix 1999), these results suggest that previous FP analyses can be extended to the star-forming population, which would allow for the galaxy population to be studied as a whole and hence minimise the impact of selection effects and progenitor bias.

In this paper, we present the luminosity and mass FP of both star-forming and quiescent galaxies at $z \sim 0.8$ from the Large Early Galaxy Astrophysics Census (LEGA-C) survey (van der Wel et al. 2016; Straatman et al. 2018), which provides deep continuum spectroscopy for a large, K_{s}-band selected sample of galaxies at $0.6 < z < 1.0$. We explore systematic variations in the structural, environmental and stellar population properties within the scatter of the FP, to study the connection between the stellar populations and structures of massive galaxies at $z \sim 0.8$.

The paper is structured as follows. We describe the data sets used, the sample selection criteria and our spectral energy distribution (SED) modelling in Section 2.2. We examine the dependence of the scatter in the luminosity FP on variations in M_*/L in Section 2.3. We present the mass FP in Section 2.4 and discuss correlations with galaxy structure and environment. The implications of our findings are discussed in Section 2.5 and summarised in Section 2.6.

We assume a flat ΛCDM cosmology throughout, with $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and $\Omega_{\mathrm{m}} = 0.3$. All magnitudes are in the AB photometric system.

2.2 Data

2.2.1 The LEGA-C Survey

The LEGA-C survey (van der Wel et al. 2016; Straatman et al. 2018) is a deep spectroscopic survey conducted with the VIMOS spectrograph on the Very Large

Telescope, targeting massive galaxies at redshifts $0.6 < z < 1.0$ in the COSMOS field. The primary sample of the survey consists of ~ 3000 K_s-band magnitude selected objects, with a redshift-dependent limit $K_s = 20.7 - 7.5 \log[(1 + z)/1.8]$, corresponding to stellar masses of $\log(M_*/M_\odot) \gtrsim 10$. Each target was observed for a total of $\sim 20\,\mathrm{h}$ at a resolution of $R \sim 2500$ in the wavelength range $\sim 6300 - 8800\,\text{Å}$, resulting in spectra which reach a typical continuum signal-to-noise level of $S/N \approx 20\,\text{Å}^{-1}$. Here, we use the third data release of the LEGA-C survey, comprising 4209 spectra (including duplicate observations) which were reduced in a similar fashion to Straatman et al. (2018).

Integrated stellar velocity dispersions are measured from the absorption linewidths in the 1D optimally extracted spectra using the Penalized Pixel-Fitting code (pPXF; Cappellari & Emsellem 2004; Cappellari 2017). As described in full detail in Straatman et al. (2018) and Bezanson et al. (2018b), the continuum emission of each spectrum is modelled using a set of high-resolution synthetic stellar population templates, and the observed stellar velocity dispersion is measured as the Gaussian broadening of the best-fitting combination of templates. We note that this measurement differs from the *intrinsic* stellar velocity dispersion: absorption lines in the 1D, spatially-integrated spectrum can also be broadened by the (projected) rotational motions of a galaxy, and hence both the intrinsic velocity dispersion and rotational velocity contribute to the integrated velocity dispersion. The inclusion of rotational motion is important, as the resulting integrated velocity dispersion approximates the second velocity moment in the virial theorem (see Cappellari et al. 2006). These integrated velocity dispersions are, however, dependent on the inclination of galaxies with respect to the line of sight, especially for rotationally-supported systems. We explore the effect of inclination on our results in Section 2.4.5. We correct all measured dispersions to an aperture of one effective radius using the typical correction derived by van de Sande et al. (2013), $\sigma = 1.05 \times \sigma_{\mathrm{obs}}$. The same, constant correction is applied to all galaxies, which may be incorrect if there is a strong radial gradient in the profile of the velocity dispersion. However, as the aperture of the slit is on average only slightly larger than the typical effective radius of the LEGA-C galaxies, the choice of aperture correction does not have a large effect: our results and conclusions do not change significantly if we instead use the commonly adopted aperture correction by Cappellari et al. (2006), which takes into account the ratio of the slit aperture and the effective radius.

2.2.2   Ancillary data to LEGA-C

Morphological information in the rest-frame optical is available for nearly all LEGA-C galaxies from *HST* ACS F814W imaging in the COSMOS field (Scoville et al. 2007). Structural parameters are derived by fitting Sérsic profiles to the ACS imaging using GALFIT (Peng et al. 2010), following the procedures described in van der Wel et al. (2012, 2016). The Sérsic profile is parameterised by the Sérsic index n, the effective radius along the major axis a, and the ratio of the minor to major axis b/a. In the following, we consider only the circularised effective radius $r_e = \sqrt{ba}$, and correct all sizes to a rest-frame wavelength of $5000\,\text{Å}$, following van

der Wel et al. (2014a). We note that the circularised radius may not provide a
good estimate of the galaxy size for disk-like morphologies, as it is dependent on
the inclination angle of the system. However, as will be further discussed in Section 2.4.5, by using the circularised radius (as opposed to the major axis radius)
we are able to approximately counterbalance the dependence of the integrated velocity dispersion on the galaxy inclination, and thus mitigate the effects of galaxy
inclination on the FP. Lastly, we assume a nominal uncertainty of 10% on the measured sizes, and 5% on the integrated luminosity of the Sérsic profile (motivated
by van der Wel et al. 2012, Fig. 7).

The LEGA-C targets were selected from the K_s-selected UltraVISTA catalogue
constructed by Muzzin et al. (2013b), which consists of PSF-matched photometry
in 30 bands ranging from $0.15 - 24\,\mu m$. We measure rest-frame $U - V$ and $V - J$
colours from the multi-wavelength photometry using the EAZY template fitting
code (Brammer et al. 2008) with redshifts fixed to the spectroscopic redshifts, as
described in detail in Straatman et al. (2018).

We use the MAGPHYS code (da Cunha et al. 2008) to fit the photometric
SEDs and derive stellar population properties. MAGPHYS uses an energy balance recipe, which accounts for light absorbed by dust in the stellar birth clouds
being re-radiated in the infrared. To fit the SEDs, we use the infrared libraries
from da Cunha et al. (2008) and the Bruzual & Charlot (2003) stellar population templates, and assume a Chabrier IMF (Chabrier 2003), an exponentially
declining star formation history (SFH) with random bursts of star formation superimposed, and a two-component dust model (Charlot & Fall 2000). We fix the
redshift to the spectroscopic redshift and use only a subset of the UltraVISTA
photometry, consisting of all available broad bands ($uBgVrizYJHK_s$ as well as
the *Spitzer*/IRAC and *Spitzer*/MIPS photometry). For all SED-derived properties, we use the median of their posterior likelihood distribution and treat the
16th and 84th percentiles as 1σ uncertainties. We provide our catalogue of SED
properties used in this work in Appendix 2.A (Table 2.1), and also show a comparison between our stellar mass estimates and those presented in van der Wel et al.
(2016). Finally, we scale the stellar mass to a total stellar mass using the total
luminosity of the best-fit Sérsic profile (e.g., Taylor et al. 2010), a small correction
that typically increases the stellar mass by $\sim 2\%$.

2.2.3   Sample selection at $z \sim 0.8$

We select galaxies from the primary LEGA-C sample, using the flag $f_{\text{primary}} = 1$
and redshift restriction $0.6 \leq z \leq 1.0$ (2915 spectra, of which 294 are duplicate observations). Of this sample, we select all (2477) galaxies of stellar mass
$\log(M_*/M_\odot) \geq 10.5$. We exclude 51 spectra which do not meet the quality criteria described in Straatman et al. (2018) (e.g., flaws in the data reduction), as
well as those (365) with a $> 15\%$ uncertainty on the integrated stellar velocity
dispersion. Moreover, we require that the GALFIT fit has converged within the parameter constraints, leaving 1656 objects, of which 167 are duplicate observations.
We visually inspect the model and residual images and flag galaxies with significant
residual flux (f_{morph}), which show merger activity or for which a two-component
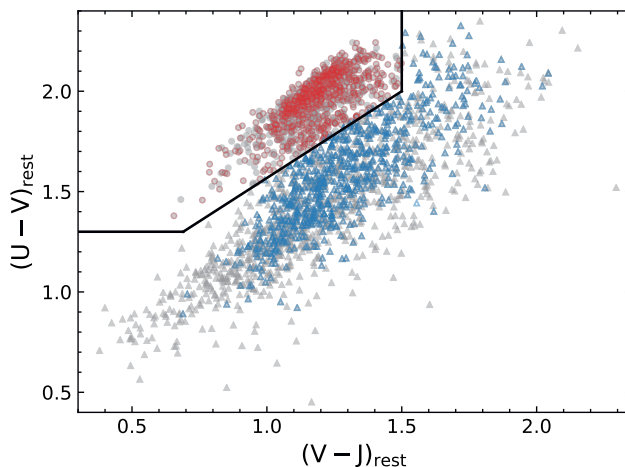
Figure 2.1: Rest-frame UVJ colours of galaxies in the primary sample of the LEGA-C survey at $0.6 < z < 1.0$. The selected sample of 1419 galaxies are highlighted in red (quiescent galaxies) and blue (star-forming galaxies), with solid lines showing the quiescent criteria from Muzzin et al. (2013a).

fit would be more appropriate (e.g., due to the presence of a point-source AGN, or star-forming clumps). Lastly, we flag objects that appear to be one system in the ground-based imaging, but are found to be close pairs of galaxies in the *HST* image. The resulting sample consists of 1489 unique objects, of which 66 are flagged as $f_{\mathrm{morph}} = 1$ and 28 are close pairs. We inspect the close galaxy pairs in this sample, and remove pairs (4) where the line broadening in the spectrum reflects their velocity offset, rather than the internal galaxy kinematics. Our final sample, for which $f_{\mathrm{morph}} = 0$, comprises 1419 galaxies. We note that including objects for which $f_{\mathrm{morph}} = 1$ introduces additional scatter, but does not change the results and conclusions in this paper.

We show the UVJ diagram of all (2621) primary LEGA-C galaxies at $0.6 < z < 1.0$ in Fig. 2.1, with the selected sample marked in red (quiescent) and blue (star-forming); we classify galaxies as quiescent and star-forming using the rest-frame $U - V$ and $V - J$ colours, following the Muzzin et al. (2013a) criteria:

$$U - V > 1.3 \tag{2.1}$$

$$V - J < 1.5 \tag{2.2}$$

$$U - V > 0.69 + 0.88\,(V - J)\,. \tag{2.3}$$

Our selected sample populates a large region in the colour-colour space, and is therefore representative of the massive galaxy population. It does not sample the bluest colours, which can be attributed to our selection on stellar mass: LEGA-C galaxies in the lower left corner of the UVJ diagram have a typical stellar mass of $\log(M_*/M_\odot) \approx 10.0$, and are therefore excluded. The S/N criterion imposed on the velocity dispersion does introduce some bias against (massive) galaxies with
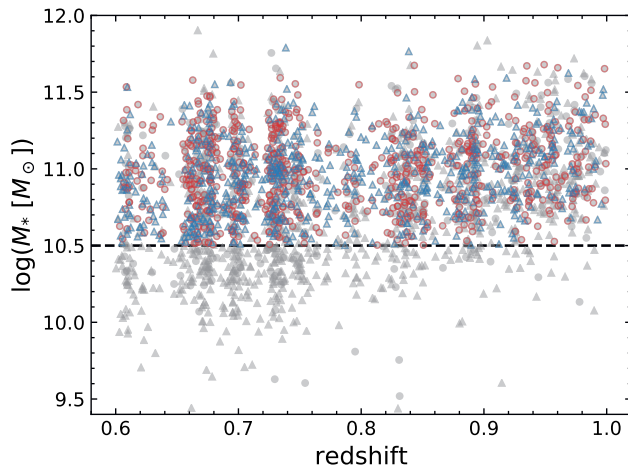
Figure 2.2: Stellar mass versus spectroscopic redshift of galaxies in the primary sample of the LEGA-C survey. The selected galaxies are marked in red and blue, indicating the UVJ quiescent and star-forming sample respectively. The dashed line shows the stellar mass criterion used to construct a representative sample of massive galaxies at $z \sim 0.8$. There are two discernible overdensities at $z \approx 0.67$ and $z \approx 0.73$, comprising $\sim 40\%$ of the sample.

very red rest-frame $V - J$ colours, typically corresponding to galaxies that are more strongly attenuated by dust and thus have a lower continuum S/N level in the spectra.

Fig. 2.2 shows the distribution of the selected sample as a function of the stellar mass and redshift. The dashed line marks a stellar mass of $\log(M_*/M_\odot) = 10.5$, above which we define our sample of LEGA-C galaxies (1419 objects) that is representative of galaxies of stellar mass $\log(M_*/M_\odot) \geq 10.5$ at $z \sim 0.8$.

2.2.4   Comparison sample at $z \sim 0$

We compile a reference sample of local galaxies by selecting galaxies in the redshift range $0.05 < z < 0.07$ from the 7th data release of the SDSS (DR7; Abazajian et al. 2009), for which $sciencePrimary = 1$, $reliable = 1$, $z\_warning = 0$, $sn\_median > 15$ and the uncertainty on the stellar velocity dispersion is $< 15\%$. To obtain stellar mass estimates that are comparable with the LEGA-C SED fits, we match the selected SDSS sample with the MAGPHYS-derived stellar mass catalogue by Chang et al. (2015). This has the advantage that (i) the same models and fitting method are used as in Section 2.2.1, and (ii) the photometry used spans a range in wavelength $(0.4 - 22\,\mu\text{m})$ that is similar to the UltraVISTA photometry, as Chang et al. (2015) cross-match the SDSS photometry with WISE. We use the structural parameters derived by Simard et al. (2011) from the 2D single Sérsic profile fits in the r-band. As there are multiple structural parameter catalogues

available for the SDSS, we examine the effect of our choice of the catalogue used in Appendix 2.B, finding no significant differences in the resulting FP. Following Section 2.2.1, we consider only the circularised effective radius, and correct the stellar mass estimates for missing flux using the total luminosity of the Sérsic profile. Selecting only galaxies of total stellar mass $\log(M_*/M_\odot) \geq 10.5$, our final sample contains 23,036 galaxies.

Moreover, we calculate rest-frame colours and luminosities using KCORRECT (Blanton & Roweis 2007), and distinguish between quiescent and star-forming galaxies using the rest-frame $u - r$ and $r - z$ colours and the colour cuts from Holden et al. (2012):

$$u - r > 2.26\,, \tag{2.4}$$

$$r - z < 0.75\,, \tag{2.5}$$

$$u - r > 0.76 + 2.5\,(r - z)\,. \tag{2.6}$$

Lastly, we consider the fact that the SDSS fibre spectra have an aperture diameter of $3''$, which covers only the central region of a galaxy at $z \approx 0.06$. We use publicly available data from the Mapping Nearby Galaxies at Apache Point Observatory survey (MaNGA; Bundy et al. 2015) of the SDSS DR15 (Blanton et al. 2017) to assess the effect of aperture size on the integrated stellar velocity dispersion, taking into account the dependence on the effective radius, Sérsic index, and axis ratio. As further detailed in Appendix 2.C, we hence derive a statistical aperture correction (typically $\sim 3\%$) to calculate the integrated stellar velocity dispersion within the effective radius from the fibre-derived SDSS DR7 velocity dispersions.

2.3   Luminosity Fundamental Plane

We begin by focusing on the fundamental plane in luminosity, specifically the luminosity measured in the rest-frame g-band. We measure the correlation between the residuals of the FP and various SED properties to explore the origin of the scatter in the FP and the differences between the star-forming and quiescent galaxy populations.

The FP describes the relation between the (integrated) stellar velocity dispersion (σ), surface brightness (I_e), and effective radius (R_e):

$$\log R_e = a \log \sigma + b \log I_e + c\,, \tag{2.7}$$

where the coefficients a and b describe the tilt of the plane, and c is the zero point. The parameters R_e and σ have units of kpc and $\mathrm{km\,s^{-1}}$ respectively, and $\log I_e \equiv -0.4\,\mu_e$, where μ_e is the mean surface brightness within the effective radius (see, e.g., Hyde & Bernardi 2009):

$$\mu_e = m + 2.5 \log\left(2\pi r_e^2\right) - 10 \log(1 + z)\,, \tag{2.8}$$

where m is the (rest-frame) apparent magnitude, and r_e is the effective radius in arcseconds.

2.3.1   Tilt of the FP

An accurate measurement of the tilt, such as in Hyde & Bernardi (2009), requires a detailed analysis of the sample completeness in both M_* and σ, as well as the uncertainties on all observed parameters. A full analysis of the tilt of the FP is beyond the scope of the current paper, and we will therefore assume minimal evolution in the tilt of the FP throughout, adopting the measurement of the rest-frame g-band plane ($I_\mathrm{e} = I_{\mathrm{e,\,g}}$) by Hyde & Bernardi (2009) for galaxies at $z \sim 0$, of $a = 1.404$ and $b = -0.761$.

However, as discussed in Section 2.1, there are several previous studies at variance with this assumption, as less massive galaxies of low M_dyn/L are likely to cause the FP to deviate more strongly from the virial plane toward higher redshift (see, e.g., Jørgensen & Chiboucas 2013). Therefore, we consider here the possibility of an evolution in the tilt and its effect on the results presented in the following sections.

Following an approach similar to Jorgensen et al. (1996) and Holden et al. (2010), we determine the best-fit values of a and b of the FP by minimising the sum of the absolute orthogonal deviations,

$$\Delta_\mathrm{LFP} = \frac{\left|\log R_\mathrm{e} - a \log \sigma - b \log I_{\mathrm{e,\,g}} - c\right|}{\sqrt{1 + a^2 + b^2}} \,. \tag{2.9}$$

We use the total completeness correction ('Tcor', see Straatman et al. 2018) as weights in the minimisation, such that less luminous galaxies receive a greater weight in the fitting procedure. This completeness correction accounts for the selection function of LEGA-C galaxies with respect to the full parent sample of K_s-band selected objects from the UltraVISTA catalogue, and includes a V_max correction. We note, however, that this completeness correction does not correct for the additional selection criteria imposed in Section 2.2.3, such as the maximum allowed uncertainty on the integrated velocity dispersion. To mitigate a bias against low-mass galaxies of high M_dyn/L, we impose a minimum velocity dispersion of $\log(\sigma/\mathrm{km\,s^{-1}}) > 2.1$: this limit corresponds to a completeness in $\log \sigma$ of $> 50\%$ up to $K_\mathrm{s} = 20.1$ (the magnitude limit comprising 90% of our sample).

As our data span a wide range in redshift and the zero point c changes significantly within $0.6 < z < 1.0$ (de Graaff et al. 2020), we restrict our fitting to a redshift range of $\Delta z = 0.10$. We measure the tilt in the range $0.65 < z < 0.75$, which encompasses the largest fraction of galaxies in our selected sample within the narrow window of $\Delta z = 0.10$ (602 objects; see Fig. 2.2). For comparison with previous studies, we use only the 325 quiescent galaxies within this redshift range. The best-fit parameters are $a = 1.29 \pm 0.18$ and $b = -0.62 \pm 0.04$ (where errors are estimated by bootstrapping the data).

This value of a is in good agreement with the value of $a = 1.40 \pm 0.05$ found by Hyde & Bernardi (2009), the measurement by Jorgensen et al. (1996) ($a = 1.24 \pm 0.07$), as well as the results by Holden et al. (2010), who found $a = 1.18 \pm 0.08$ and $a = 1.19 \pm 0.13$ at $z \sim 0$ and $z \sim 0.8$ respectively. The other parameter, b, appears to be in tension with these studies, including the assumed value of $b = -0.76 \pm 0.02$ by Hyde & Bernardi (2009) (a discrepancy of $\approx 4\sigma$).
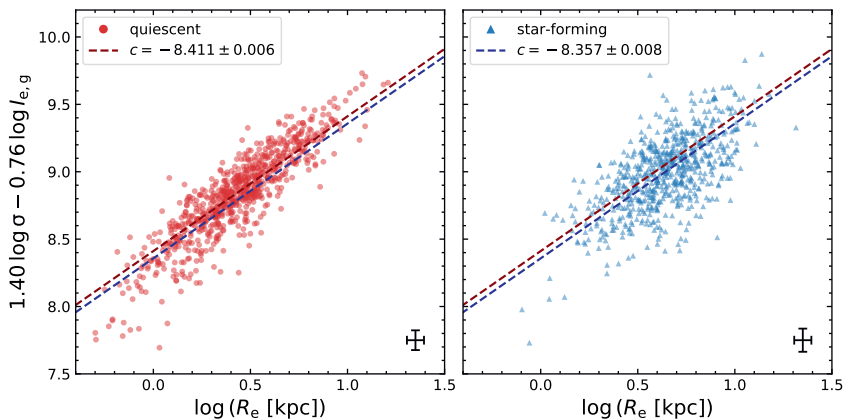
Figure 2.3: Edge-on view of the rest-frame g-band fundamental plane of quiescent (left) and star-forming (right) LEGA-C galaxies, assuming a fixed tilt from Hyde & Bernardi (2009). Star-forming and quiescent galaxies occupy different parts of the parameter space, as they differ in their best-fit zero points (dashed lines), effective radii, and scatter about the plane ($0.139 \pm 0.006\,\mathrm{dex}$ and $0.085 \pm 0.004\,\mathrm{dex}$ respectively). Observational uncertainties are similar for both populations, therefore indicating a significantly higher intrinsic scatter for star-forming galaxies.

To evaluate the dependence of the measured tilt of the FP on the fitting method used, we apply our method to the selected reference sample of $z \sim 0$ galaxies (Section 2.2.4). Imposing the same criterion of $\log(\sigma/\mathrm{km\,s^{-1}}) > 2.1$, we find $a = 1.296 \pm 0.015$ and $b = -0.732 \pm 0.004$. This is indeed slightly lower than the measurement by Hyde & Bernardi (2009), who used a more comprehensive fitting technique, and leaves a difference of $\approx 3\sigma$ in b with respect to the LEGA-C measurement.

In agreement with previous measurements of the FP of quiescent galaxies (e.g., Jørgensen & Chiboucas 2013; Saracco et al. 2020) we thus find a slight change in the tilt toward higher redshift. We note that there may be small systematic effects contributing to this observed evolution, as the SDSS data and LEGA-C data differ systematically in their measurements of R_{e}, σ and L_g, as well as the galaxy selection function. We further investigate the redshift dependence of the tilt in Appendix 2.D, where we consider the full redshift range of LEGA-C as well as the effects of measurement uncertainties and selection bias.

Importantly, however, we have used the tilt measured in this section to verify that our assumption of no evolution does not affect our conclusions. If we adopt our measurement of the tilt, only the measurements of the zero points change significantly ($> 3\sigma$), although the relative difference between the zero points of the quiescent and star-forming populations remains. The observed correlations within the residuals from the FP in the following sections are also largely unchanged, as the correlation coefficients change only minimally in value.

2.3.2   Correlations between residuals from the FP and stellar population properties

We fit the zero point (c) of the plane for the quiescent and star-forming samples separately by minimising the mean absolute orthogonal deviation (Eq. 2.9) at fixed a and b. We calculate the scatter about the best-fit zero point as the normalised median absolute deviation (NMAD) in Δ_{LFP} (Eq. 2.9), and estimate uncertainties on both quantities using bootstrap resampling.

Fig. 2.3 shows an edge-on projection of the g-band FP, for both quiescent (red) and star-forming (blue) galaxies, with dashed lines indicating the respective best-fit zero point. Traditionally, studies of the FP have focused on quiescent galaxies only (e.g., Dressler et al. 1987; Jorgensen et al. 1996; van der Wel et al. 2004), as they form a tight sequence and can therefore be used as a distance indicator, or to study the evolution of the mass-to-light ratio (M_{dyn}/L). We confirm this result for the LEGA-C sample of quiescent galaxies, which has a scatter of 0.085 ± 0.004 dex. However, we also show that star-forming galaxies seem to follow the same tilt, albeit with a larger scatter, of 0.139 ± 0.006 dex. The star-forming galaxies occupy a different area of the parameter space: they are typically larger in size, consistent with the findings by van der Wel et al. (2014a), and their best-fit zero point ($c = -8.411 \pm 0.006$) is slightly lower than that of the quiescent population ($c = -8.357 \pm 0.008$, a difference of 5.4σ), which corresponds to a systematic offset of $\Delta \log I_{\mathrm{e,g}} = 0.071 \pm 0.013$ dex between the two populations.

We estimate the intrinsic scatter in the FP using Monte Carlo simulations: assuming a FP of zero intrinsic scatter, we self-consistently vary R_{e}, $I_{\mathrm{e,g}}$ and σ within the observational uncertainties (i.e., taking into account covariances between the different quantities), and calculate the resulting scatter in Δ_{LFP}. By doing so for 1000 simulations, we obtain a robust estimate of the scatter in Δ_{LFP} due to observational uncertainties alone. The remaining contribution to the observed scatter then is due to intrinsic variation about the plane. We find that the intrinsic scatter is slightly lower than the observed scatter, at 0.082 ± 0.005 dex and 0.134 ± 0.006 dex for the quiescent and star-forming samples respectively, indicating that the observed scatter is dominated by physical differences between galaxies. The value of M_{dyn}/L_g for the star-forming and quiescent populations therefore differs not only in the mean value, but also in the variance. This can reflect both (i) a difference in the structural properties, i.e. a systematically lower value of M_{dyn}/M_* for star-forming galaxies as well as an increased intrinsic scatter in $\log M_{\mathrm{dyn}}/M_*$, and (ii) a systematic difference in the *stellar* mass-to-light ratio (M_*/L_g) between the two populations. Considering the UVJ colour selection, a difference in M_*/L_g may be expected to contribute the systematic offset between the two zero points. Moreover, the wide range in colour spanned by the star-forming galaxies (Fig. 2.1) suggests that they are more strongly affected by dust attenuation, thus leading to a larger intrinsic scatter in the FP.

We demonstrate the dependence of the scatter on M_*/L_g in Figs. 2.4 & 2.5, where we show the residual from the FP in $\log I_{\mathrm{e,g}}$ as a function of the 4000 Å break ($D_{\mathrm{n}}4000$; Wu et al. 2018) and the Lick index $H\delta_A$, which are age indicators measured directly from the spectra, as well as the rest-frame $U - V$ and $V - J$

colours. Similar to the results from Graves et al. (2009) for quiescent galaxies at $z \sim 0$, we find a correlation with age (D_n4000, $H\delta_A$) through the thickness of the FP, which continues down toward younger, star-forming galaxies. Because $\Delta \log I_{e,g} \approx -\Delta \log M_{dyn}/L_g$, this correlation translates to a lower (higher) value of M_{dyn}/L_g for younger (older) galaxies. Our findings are also consistent with results by Jørgensen et al. (2019), who find increased Balmer line absorption ($H\zeta_A$) and lower M_{dyn}/L for quiescent galaxies in clusters at $z \sim 1$ with respect to early-type galaxies at $z \sim 0$, which they interpret as being due to a difference in age.

The residuals of the FP correlate even more strongly with the rest-frame $U-V$ and $V-J$ colours (Fig. 2.5), which in turn depend on a combination of dust attenuation, specific star formation rate (sSFR) and age (see, e.g., Leja et al. 2019b). Galaxies with positive values of $\Delta \log I_{e,g}$ are therefore not only younger on average, they may also have a higher sSFR or be less dust-obscured, or, a combination of both.

We explore these different contributions to the scatter using the results from our SED modelling (Section 2.2.2). Firstly, the upper left panel of Fig. 2.6 shows that the residual correlation with D_n4000 obtained from the best-fit SED models agrees well with that from the spectra (Fig. 2.4): in both cases there is a strong anti-correlation between $\Delta \log I_{e,g}$, and D_n4000, and the models are able to reproduce the observed bimodality, such that at fixed value of $\Delta \log I_{e,g}$ star-forming galaxies have a lower value of D_n4000. On the other hand, the models do not reproduce the observed, broad distribution in D_n4000, which may be due to incompleteness in the modelling itself or the result of degeneracy between the effects of age and dust on the observed SED.

The other panels of Fig. 2.6 show the residual correlations with the dust attenuation (A_V; measured from the best-fit SED model), the luminosity-weighted age (in the r-band) and the sSFR averaged over the last 100 Myr. For quiescent galaxies, the only significant correlation is with the stellar age. On the other hand, the scatter within the star-forming population correlates not only with age, but also weakly with the dust attenuation and, more strongly, with the sSFR.

The different intrinsic scatter for the star-forming and quiescent populations as well as the offset between the FP zero points (Fig. 2.3) are therefore, at least in part, due to significant differences in M_*/L_g between the two populations. We note that we also find the spread in all four observed properties (D_n4000, $H\delta_A$, $U-V$, $V-J$) to be slightly larger for the star-forming population than the quiescent population, which is consistent with their observed increased scatter in the FP. Interestingly, whereas the deviation between the best-fit zero points of the quiescent and star-forming samples is relatively small ($\Delta \log I_{e,g} \approx 0.07$ dex), we find that at a fixed value of D_n4000 or $(U-V)_{rest}$ the differences between the two populations can be up to three times greater ($\Delta \log I_{e,g} \sim 0.2$ dex), which may be due to variation in M_*/L_g, or differences in the structural properties. Thus far, we have neglected the effects of potential structural differences between the two populations, which we explore in full detail in the following section.
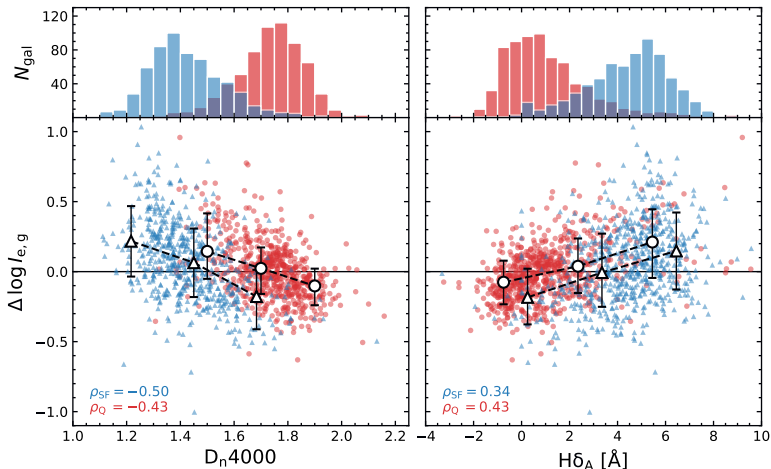
Figure 2.4: FP residual in $\log I_{e,g}$ versus the spectral age indices $D_n 4000$ (left) and Hδ_A (right). Red and blue markers indicate the quiescent and star-forming population respectively, with black open markers showing the running median and 16th and 84th percentiles. There is a strong correlation with $\Delta \log I_{e,g}$ in both panels (Spearman rank correlation coefficients, ρ, are denoted in each panel), albeit with large scatter, such that at fixed σ and R_e galaxies with higher surface brightness are younger. As the distributions in $D_n 4000$ and Hδ_A differ for the star-forming and quiescent galaxies, with the latter being older, this shows that stellar age is an important driver of the differences between the two populations in the g-band FP (Fig. 2.3).
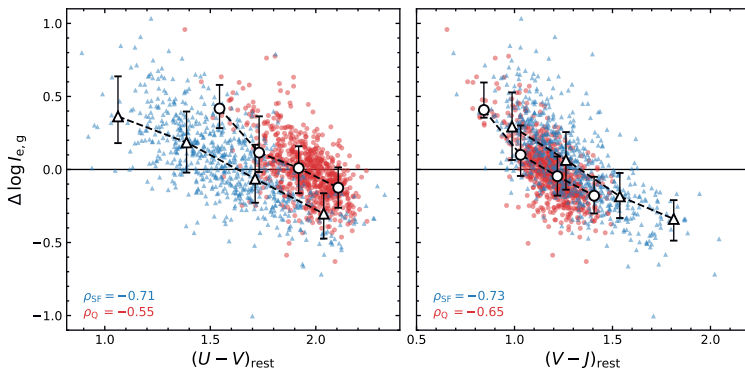
Figure 2.5: Correlation in the residuals from the FP with the rest-frame $U - V$ and $V - J$ colours. Symbols indicate the same as in Fig. 2.4. As $U - V$ and $V - J$ in turn correlate with properties of the stellar mass-to-light ratio (M_*/L_g), the strong correlations through the thickness of the FP suggest that variations in stellar age, dust attenuation, and star formation activity contribute significantly to the scatter in the FP, which we explore in Fig. 2.6.
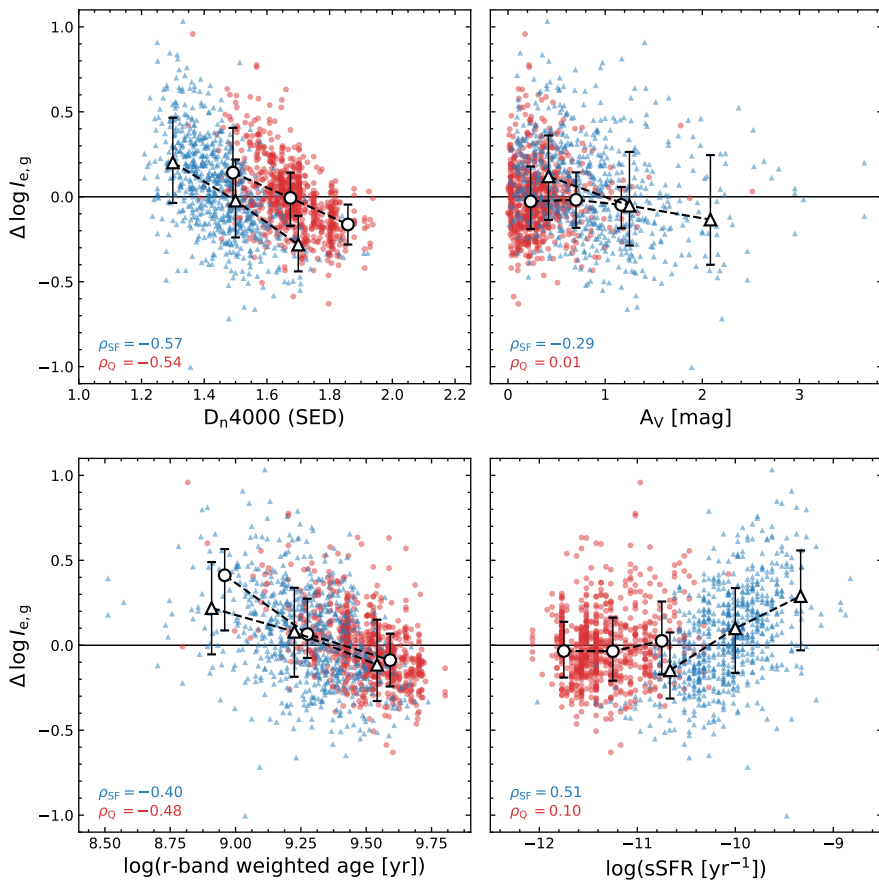
Figure 2.6: Correlation in the residuals from the FP with properties from the SED modelling (D_n4000 break, dust attenuation, luminosity-weighted stellar age, and specific star formation rate), which drive the stellar mass-to-light ratio. Symbols indicate the same as in Fig. 2.4. For the quiescent galaxies, only the variation in stellar age and D_n4000 (which, apart from the stellar age, is also dependent on the metallicity) contribute significantly to the intrinsic scatter of the g-band FP. Therefore, the increased intrinsic scatter for the star-forming population can, at least partially, be attributed to additional effects from variations in the sSFR and dust attenuation.

2.4 Mass Fundamental Plane

In this section we use the mass FP to explore the structural properties of galaxies within the parameter space of the FP, as well as the effect of environment. If we multiply the surface brightness of Eq. 2.7 by the M_*/L_g estimated from the SED modelling (Section 2.2.2), we obtain the stellar mass surface density (Σ_*), and hence the mass FP:

$$\log R_{\mathrm{e}} = \alpha \, \log \sigma + \beta \, \log \Sigma_* + \gamma, \qquad (2.10)$$

where α and β describe the tilt of the mass FP, and γ is the zero point.

2.4.1 Tilt of the mass FP

As in Section 2.3, we assume that the tilt of the FP does not vary significantly with redshift and adopt the results for the mass FP from Hyde & Bernardi (2009) of $\alpha = 1.629$ and $\beta = -0.84$, which was derived with an orthogonal fit to a large ($N \sim 50,000$) sample of early-type galaxies that takes into account both the measurement uncertainties and sample completeness. We again test the effect of this assumption using a more simple, orthogonal fit of the FP, and examine the possible redshift evolution of the tilt in more detail in Appendix 2.D.

We follow the same methodology as in Section 2.3.1, minimising the sum of the orthogonal deviations,

$$\Delta_{\mathrm{MFP}} = \frac{|\log R_{\mathrm{e}} - \alpha \, \log \sigma - \beta \, \log \Sigma_* - \gamma|}{\sqrt{1 + \alpha^2 + \beta^2}}, \qquad (2.11)$$

and using the total completeness corrections (Tcor) as weights. We include only quiescent galaxies in our fits for comparison with other FP studies, and exclude galaxies for which $\log(\sigma/\mathrm{km\,s^{-1}}) < 2.1$.

In the redshift range $0.65 < z < 0.75$ we measure a best-fit tilt of $\alpha = 1.56 \pm 0.12$ and $\beta = -0.68 \pm 0.03$ (where error bars are estimated through bootstrap resampling), which is significantly different from the assumed values by Hyde & Bernardi (2009). However, as in Section 2.3.1, we find that our measurement for the SDSS differs from the tilt found by Hyde & Bernardi (2009) due to differences in the methodology used. Both α and β measured from the LEGA-C data are consistent within $< 2\sigma$ with our best-fit parameters for the SDSS, of $\alpha = 1.432 \pm 0.012$ and $\beta = -0.736 \pm 0.003$. This remains the case even when we fit the entire LEGA-C sample combined (i.e., $0.6 < z < 1.0$), for which we find $\alpha = 1.49 \pm 0.10$ and $\beta = -0.70 \pm 0.02$, suggesting no significant rotation of the mass FP at $z \sim 0.8$ with respect to $z \sim 0$. Our results are in agreement with measurements by Zahid et al. (2016), who found no change in the tilt of the FP with respect to the SDSS for a sample of massive quiescent galaxies at $0.1 < z < 0.6$. Interestingly, these results seem to suggest that the measurement of the tilt of the mass FP, unlike the g-band FP, is not strongly dependent on the selection function, as was also recently shown by Bernardi et al. (2020) at low redshift.

However, as we have omitted the effect of measurement uncertainties in addition to a careful analysis of the selection function in our measurement of the tilt,
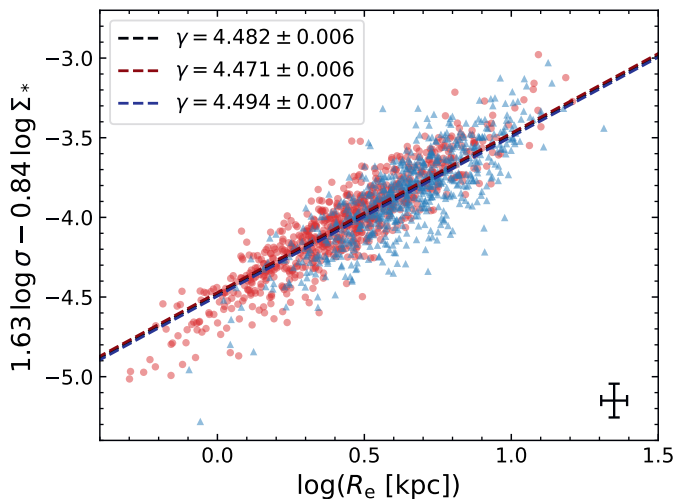
Figure 2.7: Edge-on view of the mass FP of quiescent (red) and star-forming (blue) LEGA-C galaxies. Dashed lines show the best-fit zero points for the star-forming, quiescent, and combined (black) samples, assuming a fixed tilt from Hyde & Bernardi (2009). The two populations lie on the same plane: the zero points differ by only ≈ 0.02 dex, and the intrinsic scatter is comparable for the quiescent and star-forming samples (0.107 ± 0.005 dex and 0.130 ± 0.009 dex in $\Delta \log R_{\mathrm{e}}$ respectively).

we choose to use the values by Hyde & Bernardi (2009) rather than our own measurement. We note that we do not use the more recent measurements by Bernardi et al. (2020), to adhere to the common convention of using circularised sizes in the FP, and to refrain from making assumptions on the effects of non-homology on the mass FP at higher redshifts. Although we do not use the tilt measured from the LEGA-C data in the rest of this paper, we have used this measurement to test the robustness of our results in the following sections against a different tilt, finding no qualitative differences.

2.4.2   Edge-on view of the mass FP

Analogous to Section 2.3.2, we fit the zero point (γ) by minimising the mean absolute orthogonal residuals at fixed α and β. Fig. 2.7 shows an edge-on projection of the mass FP, for both the star-forming (blue) and quiescent (red) galaxies. The best-fit zero points are indicated by dashed lines for the two populations separately (red, blue), as well as for the joint sample (black). Not only do both populations follow the same tilt, the star-forming and quiescent galaxies also have nearly equal zero points, with the two zero points deviating by 0.023 ± 0.009 dex (a systematic offset of $\Delta \log \Sigma_* = 0.027 \pm 0.011$ dex). This is consistent with results at low redshift by Zaritsky et al. (2008) and Bezanson et al. (2015), although Bezanson et al. (2015) find a slightly larger offset (≈ 0.05 dex) between the zero points of

the two populations at both $z \sim 0$ and $z \sim 0.7$. At $z \sim 0.7$, however, their offset is not statistically significant due to the sample size.

We find that the scatter in the mass FP is lower in comparison with the g-band FP, particularly so for the star-forming galaxies: the NMAD in Δ_{MFP} (Eq. 2.11) is 0.069 ± 0.003 dex and 0.085 ± 0.005 dex for the quiescent and star-forming samples respectively, and is consistent with the findings by Bezanson et al. (2015). Using Monte Carlo simulations, we estimate the intrinsic scatter for the quiescent and star-forming samples to be 0.058 ± 0.003 dex and 0.069 ± 0.005 dex respectively. Clearly, accounting for the M_*/L dramatically lowers both the total and intrinsic scatter of the star-forming population, although it is still slightly higher than the scatter within the quiescent population. Thus, unlike the g-band FP, all massive galaxies occupy the same region within the 3D parameter space of the effective radius, stellar mass surface density, and stellar velocity dispersion, regardless of their colour.

The remaining intrinsic scatter is low, but non-zero. In principle, a large number of galaxy properties may drive the intrinsic scatter in the mass FP: we discuss the effect of stellar populations on the FP in Section 2.4.3, the structural properties in Sections 2.4.4 & 2.4.5, and the effect of environment in Section 2.4.6.

2.4.3   Are the residuals from the mass FP correlated with stellar population properties?

In Fig. 2.8 we show the residual from the FP in $\log \Sigma_*$ as a function of the spectral properties $D_n 4000$ and $H\delta_A$ (left-hand panels), the rest-frame colours $U - V$ and $V - J$ (middle panels), and the SED-derived dust attenuation and stellar age (right-hand panels). Unlike the results of Figs. 2.4–2.6, we find no significant correlations with the different SED properties through the mass FP. There is only a very weak correlation with the spectral age indicators ($D_n 4000$ and $H\delta_A$), which may correspond to the very weak residual correlation between $\Delta \log \Sigma_*$ and A_V (upper right panel) or the sSFR (Spearman $\rho = 0.12$; shown in de Graaff et al. 2020, Fig. 3).

To first order, the lack of residual correlations within the scatter of the mass FP demonstrates the success of our SED modelling: if we neglect a potential correlation between structural and stellar population properties, and interpret the mass FP as arising from the virial theorem, then we would expect to find no correlation between the zero point γ and the stellar population properties of galaxies that are in virial equilibrium.

Thus far, we have simply used our SED models without questioning the underlying model assumptions, although we did show in Fig. 2.6 that the $D_n 4000$ index measured from the best-fit SEDs agree reasonably well with the measurements from the LEGA-C spectra. However, there are a large number of available SED fitting codes, with an even a larger parameter space of, e.g., possible star formation histories, dust laws and IMFs. For instance, in Appendix 2.A we compare our MAGPHYS masses to those derived with FAST (Kriek et al. 2009) and find significant, systematic differences between the two, casting doubt on the accuracy of the various stellar mass estimates.
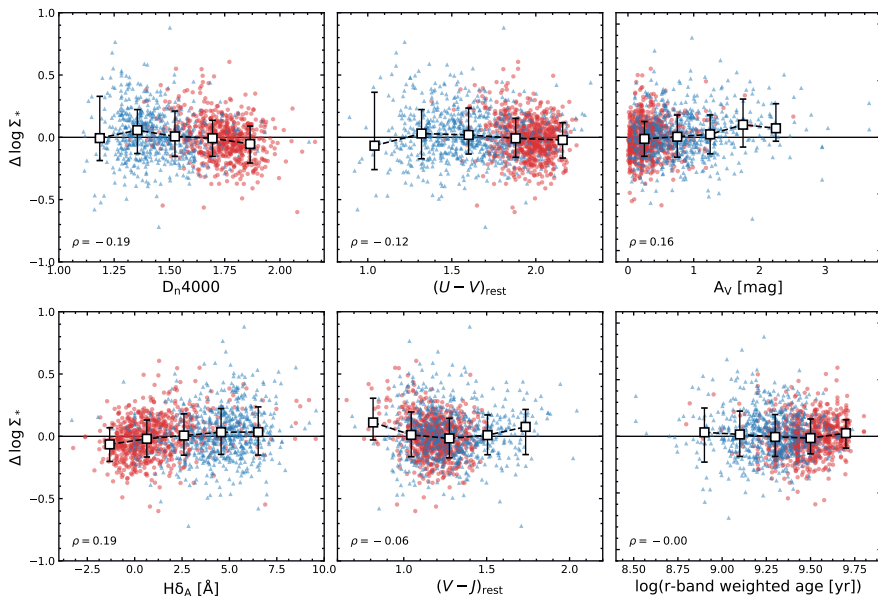
Figure 2.8: Residual in the mass fundamental plane in $\log \Sigma_*$ as function of the spectral age indices D_n4000 and $H\delta_A$ (top panels), and the rest-frame $U - V$ and $V - J$ colours. Red and blue markers indicate the quiescent and star-forming population respectively, with white squares showing the median and 16th and 84th percentiles of the total sample (with Spearman rank correlation coefficients, ρ, denoted in each panel). Contrary to the results in Fig. 2.4 for the g-band FP, we find no significant correlation with stellar population properties through the thickness of the mass FP.

Instead of using our modelled stellar masses to shed light on the FP, we can also ask whether the FP itself can provide information on the accuracy of the modelled M_*/L (see also van de Sande et al. 2015, who discuss the constraining power of M_{dyn}/L on stellar population properties). To do so, we calculate the M_*/L predicted by the mass FP:

$$\log \left(\frac{M_*}{L_g} \right)_{\text{FP}} = \log \Sigma_{*,\text{FP}} - \log I_{\text{e,g}} , \qquad (2.12)$$

where $I_{\text{e,g}}$ is the observed surface brightness and

$$\log \Sigma_{*,\text{FP}} = \left(\frac{1}{\beta} \right) \log R_{\text{e}} - \left(\frac{\alpha}{\beta} \right) \log \sigma - \left(\frac{\gamma}{\beta} \right) . \qquad (2.13)$$

The FP does not provide an absolute scaling of M_*/L_g, unless the value of γ is constrained otherwise (as done by Schechter et al. 2014). In Fig. 2.9 we therefore show $\log(M_*/L_g)_{\text{FP}} + (\gamma/\beta)$ versus the M_*/L_g estimated with MAGPHYS. The solid line has a unit slope, with the intercept set equal to the best-fit zero point of the mass FP (Fig. 2.7). The dashed line on the other hand shows the best fit
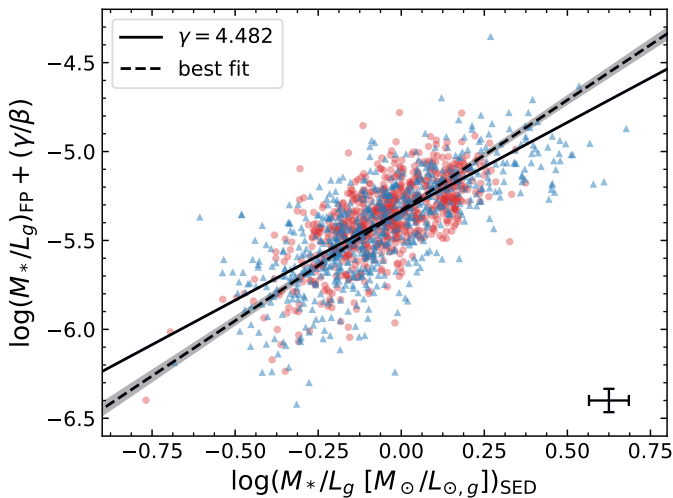
Figure 2.9: Comparison of the stellar mass-to-light ratio (M_*/L_g) predicted from the mass FP, and M_*/L_g estimated from multi-wavelength SED fitting with MAGPHYS, demonstrating that the SED modelling provides a reasonable approximation of M_*/L_g.

from an orthogonal distance regression that takes into account uncertainties in both variables, which gives a slope of $m = 1.24 \pm 0.03$. We note that the measured slope is only weakly dependent on the adopted tilt: if we instead use the measured tilt from Section 2.4.1, we find variations of order $\sim 1 - 2\sigma$ (e.g., $m = 1.27 \pm 0.04$ for the best-fit tilt at $0.65 < z < 0.75$).

Although the best-fit relation is statistically significantly different from a unit slope, the two different estimates of M_*/L_g agree remarkably well, considering that the only assumption made in calculating $(M_*/L_g)_{\mathrm{FP}}$ is that the mass FP has zero intrinsic scatter. Both estimates show a similar, large spread in M_*/L_g, and the scatter about the solid line is $\sigma_{\mathrm{NMAD}} = 0.117 \pm 0.004$ dex, partially driven by the uncertainties (of ~ 0.06 dex in either axis). Moreover, the systematic offset between the solid and dashed lines is < 0.05 dex for $\approx 75\%$ of the sample. Only toward extreme values of M_*/L_g do the systematic discrepancies become larger $(\sim 0.1$ dex), where the SED modelling also becomes more difficult (e.g., accurately predicting the effects of dust, or the recent star formation history) and the intrinsic scatter in the mass FP may become important.

2.4.4   Structural non-homology

The zero point of the mass FP is inversely proportional to ratio of the dynamical and stellar mass (i.e., $\gamma \propto \log(M_*/M_{\mathrm{dyn}})$), and therefore depends on the dark matter fraction within the effective radius, as well as the assumed IMF in the SED modelling. Considering structural properties only, one may expect a dependence of the zero point on the Sérsic index (n): n reflects the distribution of the stellar
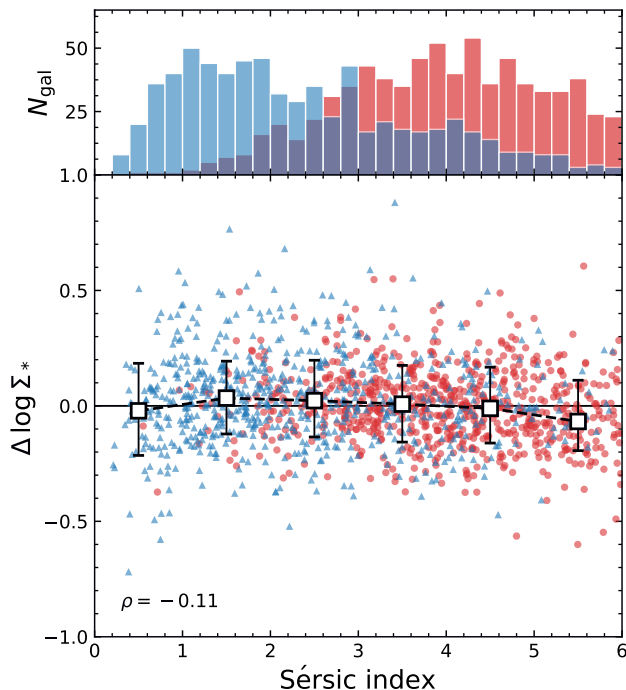
Figure 2.10: Residual in the mass FP in $\log \Sigma_*$ as a function of the Sérsic index. Symbols indicate the same as in Fig. 2.8. The star-forming and quiescent galaxies follow very different distributions in Sérsic index, yet, this has no significant effect on the scatter of the mass FP.

light, and hence the density profile of the stellar mass. Systematic differences in these density profiles may therefore lead to Sérsic-dependent variations in the velocity dispersion or the dark matter fraction within one $R_{\rm e}$. Bezanson et al. (2015) find a weak correlation between γ and n at $z \approx 0.06$; however, their sample at $z \sim 0.7$ contains too few objects to draw a conclusion on the non-homology of galaxies at higher redshift.

In Fig. 2.10, we show the residual from the FP in $\log \Sigma_*$ (for which $\Delta \log \Sigma_* \approx \Delta \log M_{\rm dyn}/M_*$) as a function of the best-fit Sérsic index for the significantly larger sample of LEGA-C galaxies. The median of the combined star-forming (blue) and quiescent (red) population, plotted as open squares, shows no dependence on the Sérsic index, except for the highest bin in Sérsic index. We confirm this result by performing a linear fit to the data, which indicates a very weak correlation of $\Delta \log \Sigma_* \propto (-0.020 \pm 0.004)\, n$ (Spearman rank correlation coefficient $\rho = -0.11$). The lack of an effect due to structural non-homology on the mass FP appears to be contradictory with previous measurements at $z \sim 0$, of both the FP (Bezanson et al. 2015) and direct measurements of $M_{\rm dyn}/M_*$ (e.g., Taylor et al. 2010). We discuss the implications of this result in Section 2.5.

2.4.5 Inclination effects

The third structural parameter of our Sérsic model is the observed ratio of the major and minor axes (b/a), which depends strongly on both the intrinsic morphology and the inclination angle of the system. For example, it provides an estimate of the inclination for systems that are intrinsically flat and axisymmetric.

Correlations between the projected axis ratio and M_{dyn}/L of quiescent galaxies have been predicted using the luminosity FP and Jeans modelling (e.g., Jorgensen et al. 1996; Cappellari et al. 2006), however, the effect on the observed FP is unclear. Bezanson et al. (2015) find a weak dependence of the integrated velocity dispersion on the projected axis ratio at $z \sim 0$, particularly for star-forming and low Sérsic index systems: flattened (low b/a) objects have an elevated integrated velocity dispersion, whereas the opposite is the case for round (high b/a) objects. This reflects the fact that for flattened, rotationally-supported systems, the integrated velocity dispersion is a combination of both the intrinsic velocity dispersion and the rotation along the line of sight, which is inclination-dependent.

As flattened galaxies become more common at higher redshift (van der Wel et al. 2014b; Hill et al. 2019), the effect of the inclination angle on the FP may become important. We evaluate this effect in Fig. 2.11, where we show the residual from the mass FP in $\log \sigma$ (rather than $\log \Sigma_*$) as a function of the projected axis ratio. Quiescent and star-forming galaxies are again indicated by red and blue symbols respectively, with the median of the full sample shown in black. For high values of b/a, the residual $\Delta \log \sigma$ is slightly negative: for round or face-on objects, the integrated velocity dispersion is lower than the velocity dispersion predicted from the mass FP in Eq. 2.10, as the contribution from rotational motion to the integrated velocity dispersion is minimised for systems at low inclination. Notably, this applies to both the quiescent and star-forming sample, suggesting that rotation is important for quiescent galaxies as well, and is further supported by the large number of highly flattened quiescent galaxies. The similarity between the projected axis ratio distributions of the star-forming and quiescent galaxies likely reflects a mixture of different intrinsic shapes within these galaxy populations, with both the star-forming and quiescent samples containing a significant fraction of disk-like morphologies as well as more spheroidal structures (see also Chang et al. 2013; van der Wel et al. 2014b). Additionally, the number of star-forming galaxies with low values of b/a may be slightly reduced by our selection on the SNR of the velocity dispersion (Section 2.2.3), as this results in a slight bias against highly reddened star-forming galaxies, which are more likely to be edge-on projections.

The anti-correlation between b/a and $\Delta \log \sigma$, however, does not continue toward low axis ratios, where we would expect the integrated velocity dispersion to be higher than the FP prediction due to an increased contribution from the rotational velocity. This can be attributed to our use of the circularised effective radius (Section 2.2.2), which is proportional to the square root of the axis ratio. For flattened objects, the smaller effective radius counteracts the increased velocity dispersion, resulting in a predicted velocity dispersion that is approximately equal to the observed value. The net effect of the random inclination angle on the FP therefore is to slightly enhance the scatter about the FP, contributing to the
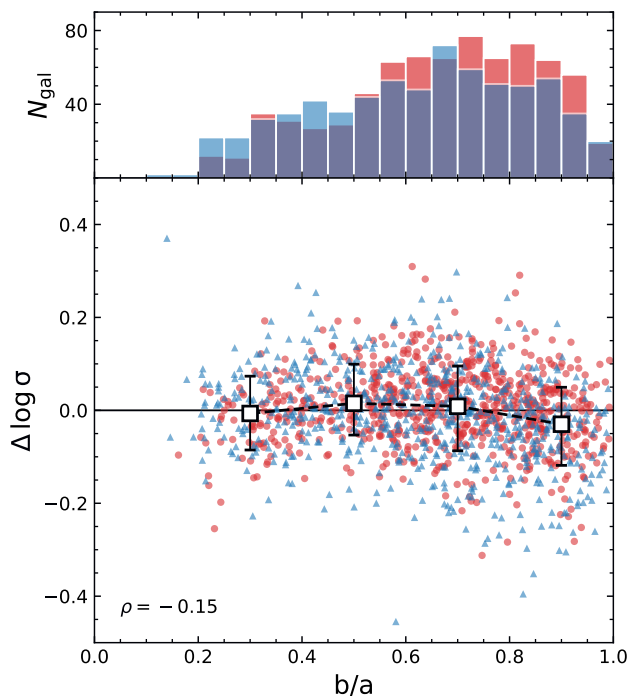
Figure 2.11: Residual from the mass FP in $\log \sigma$ as a function of the projected axis ratio (b/a). Symbols indicate the same as in Fig. 2.8. The integrated velocity dispersion is slightly lower than that predicted by the mass FP for rounder (higher b/a) systems, reflecting a minimal contribution of rotational motion to the integrated velocity dispersion for objects at low inclination angles. The effect of inclination is therefore a marginal increase in the intrinsic scatter in the FP.

intrinsic scatter derived in Section 2.4.2.

Indeed, Bernardi et al. (2020) show that the residuals of the FP correlate strongly with the axis ratio, if the major axis size is used rather than the circularised size. They hence demonstrate the importance of inclination effects on the FP, and show that the scatter in the FP can be further reduced by treating b/a as an additional variable in Eq. 2.7 or Eq. 2.10: by fitting a hyperplane to a sample of low-redshift elliptical and lenticular galaxies, they find that the tilt of the FP, i.e. the values of a and b, can differ by $\sim 2 - 3\sigma$ from the traditional (three parameter) FP, and that the scatter about the best-fit FP is decreased by up to 0.009 dex. Still, even after accounting for b/a as a separate variable, the effect of galaxy inclination remains apparent in the FP, as more highly inclined galaxies have a lower scatter about the plane than galaxies that are near face-on. These different effects are largest for S0 galaxies, and thus potentially even larger for star-forming disks.

2.4.6   Environment

Many previous studies of the luminosity FP have focused on clusters of galaxies (e.g., Jorgensen et al. 1996; van Dokkum & van der Marel 2007; Holden et al. 2010; Beifiori et al. 2017; Saracco et al. 2020), and explored differences in the properties of the FP between low and high density environments (e.g., van Dokkum et al. 2001; Cappellari et al. 2006; La Barbera et al. 2010b; Saglia et al. 2010; Joachimi et al. 2015). Burstein et al. (1990) first demonstrated that the effect of environment on the FP is expected to be small, as they found no dependence of the zero point on cluster richness. Using a large sample of early-type galaxies in the SDSS, La Barbera et al. (2010b) showed that the zero point of the luminosity FP indeed correlates weakly with the local galaxy density, regardless of the chosen passband. Joachimi et al. (2015) obtained similar results by considering the spatial correlation function of residuals in the r and i-band FP with the galaxy density field, and additionally find small systematic differences between central galaxies and satellites.

Interpreting the zero point of the plane as M_{dyn}/L, these results imply that galaxies in lower density environments have lower values of M_{dyn}/L than those in high density environments, and that central galaxies have higher M_{dyn}/L than satellites. A systematically lower luminosity-weighted age for field galaxies can explain their lower values in M_{dyn}/L as compared to cluster galaxies (van Dokkum & van der Marel 2007; La Barbera et al. 2010b), and is broadly consistent with the picture of hierarchical structure formation, from which we would expect galaxies to form earlier in highly dense environments. Joachimi et al. (2015) suggest that the lower value of M_{dyn}/L for satellite galaxies, which is not only lower than that of central galaxies, but also of field galaxies, can be attributed to the tidal stripping of dark matter and hot gas in the subhaloes as they fall into more massive haloes.

We explore the effect of environment on the FP by matching the LEGA-C sample with the Darvish et al. (2017) cosmic web catalogue (with a maximum matching radius of $1''$), which contains measurements of the projected density field of the COSMOS field out to $z = 1.2$, and categorises galaxies as 'central', 'satellite' or 'isolated'. This catalogue was constructed using the COSMOS2015 photometric redshift catalogue (Laigle et al. 2016) in the UltraVISTA-DR2 region (McCracken et al. 2012; Ilbert et al. 2013) following the adaptive weighted kernel smoothing method described in Darvish et al. (2015). In Fig. 2.12 we show in the top panel the residual from the g-band FP in $\log I_{\mathrm{e,g}}$ as a function of the projected overdensity, for both the quiescent (red) and star-forming (blue) sample. As the redshift distribution of the few galaxies at high overdensity is not representative of the full sample, we have corrected the values of $\Delta \log I_{\mathrm{e,g}}$ for the redshift evolution derived in de Graaff et al. (2020). The medians for galaxies classified as central, satellite or field ('isolated' in the catalogue by Darvish et al. 2017) are indicated by white symbols. There is a very weak anti-correlation between the residual in $\log I_{\mathrm{e,g}}$ and the overdensity, such that $\Delta \log I_{\mathrm{e,g}} \propto (-0.085 \pm 0.015) \log(1 + \delta)$. As this residual is inversely proportional to M_{dyn}/L, it is consistent with previous findings that galaxies in higher density environments have a higher value of M_{dyn}/L. When dividing our sample into satellites, centrals and field galaxies, we do not find any
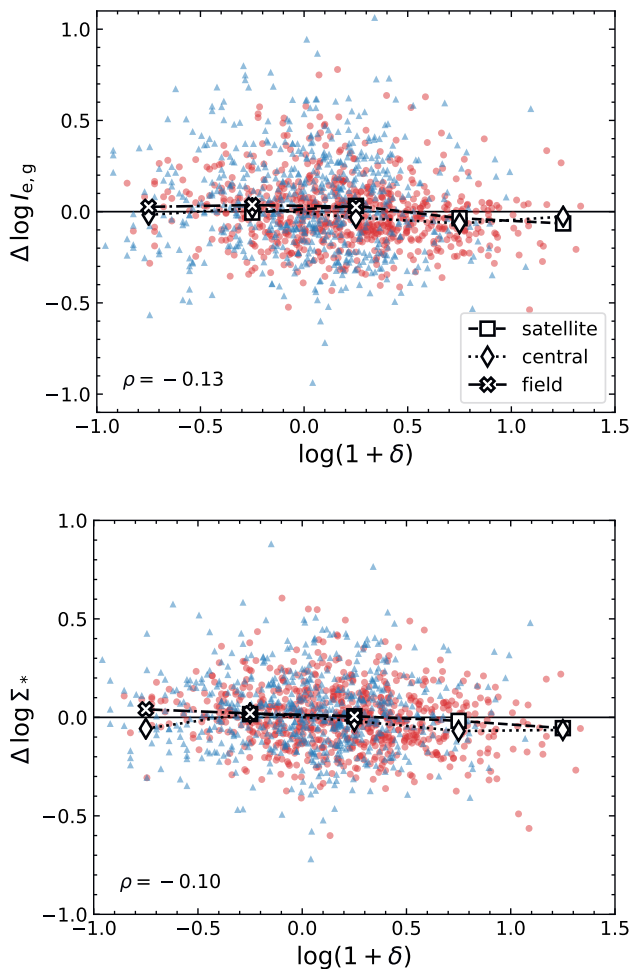
Figure 2.12: Residual in the *g*-band (top) and mass (bottom) FP as a function of the local overdensity (Darvish et al. 2017). Red and blue symbols indicate the quiescent and star-forming population respectively. White markers show the median of galaxies that are classified as central (diamonds), satellite (squares) or field (crosses) galaxies. We find no significant environmental dependence within the LEGA-C data for both the *g*-band and mass FP.

significant systematic differences between the subsamples, in contrast with the weak, but significant, effect found by Joachimi et al. (2015). However, our sample contains far fewer objects than these studies at low redshift, particularly so at high overdensity. Moreover, our measurements do not account for uncertainties in the density field estimation, which is particularly difficult to constrain precisely at low overdensities, and we therefore cannot draw any strong conclusions on the effect of environment on the FP.

Analogous to the top panel of Fig. 2.12, in the bottom panel we show the residual in $\log \Sigma_*$ of the mass FP as a function of the overdensity. We find an even weaker dependence of the zero point of the mass FP on environment, both in terms of overdensity, with $\Delta \log \Sigma_* \propto (-0.052 \pm 0.014) \log(1 + \delta)$, and galaxy type (satellite, central, field). Within the current galaxy sample and level of uncertainty, this suggests that at fixed R_e and σ the structural properties of galaxies in high density environments do not differ significantly from those in the field.

2.5 Discussion

2.5.1 Stellar populations

In agreement with many other studies (e.g., Jorgensen et al. 1996; Forbes et al. 1998; Wuyts et al. 2004; Gargiulo et al. 2009), we have shown that there is significant scatter in the luminosity FP, which cannot be attributed to measurement uncertainties alone. We find that the residuals from the FP correlate strongly with spectral features ($D_n 4000$, $H\delta_A$) as well as rest-frame colours ($U - V$, $V - J$). These correlations can be interpreted as systematic variations in M_*/L due to varying stellar ages, and in the case of the star-forming population, also different sSFRs and dust attenuation.

Previous results at low redshift, where significant residual correlations with stellar age are present in the FP (Forbes et al. 1998; Gargiulo et al. 2009; Graves et al. 2009), thus also hold at $z \sim 1$. Moreover, this correlation appears to be stronger in our sample as compared with both Gargiulo et al. (2009) and Graves et al. (2009). Graves & Faber (2010) show that variations in M_*/L contribute approximately 22% to the intrinsic thickness of the FP (i.e., $\mathrm{d} \log(M_*/L)/\mathrm{d}(\Delta \log I_e) \approx -0.22$), although depending on the stellar population modelling method used this value may be anywhere between $2\% - 53\%$.

However, these studies at low redshift focus on early-type galaxies alone, which are selected by morphology as well as insignificant $H\alpha$ or [O II]3727 line emissivity, whereas we here have extended the analysis to the full population of massive galaxies. The selected samples of early-types at $z \sim 0$ therefore likely consist of galaxies that span a narrower range in age and M_*/L. Moreover, at $z \sim 0$ the FP is often studied in the r-band, which may differ significantly from the rest-frame g-band considered here.

We evaluate the contribution of variations in M_*/L to the thickness of the FP in Fig. 2.13, using the different measurements of the tilt by Hyde & Bernardi (2009, Table 2) to obtain the FP in different rest-frame passbands. Firstly, we consider the observed (filled symbols) and intrinsic (open symbols) scatter in $\Delta \log I_{e,\lambda}$ at different wavelengths, for the quiescent (red), star-forming (blue), and combined (black) subsamples. The scatter about the mass FP (in $\Delta \log \Sigma_*$) is shown for reference. For the quiescent galaxies the observed scatter in the mass FP is approximately equal to that in the r, i and z-band FPs, whereas the intrinsic scatter in the mass FP is significantly lower than the luminosity FP, reflecting the relatively large uncertainty on the SED modelling in comparison with the observational error on the luminosity. More importantly, there is a significant decrease
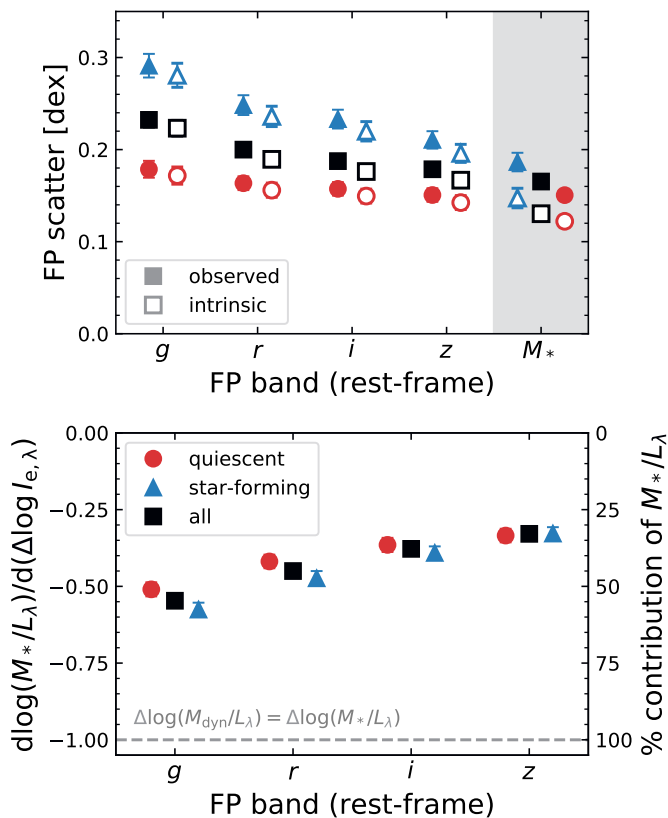
Figure 2.13: Effect of variation in the stellar mass-to-light ratio (M_*/L_λ) on the thickness of the FP. **Top:** Scatter in the FP in $\Delta \log I_{e,\lambda}$ for different rest-frame wavelengths, with solid and open symbols showing the observed and intrinsic scatter respectively. The scatter in the mass FP (in $\Delta \log \Sigma_*$) is shown for reference. **Bottom:** Contribution of M_*/L_λ to the residual from the FP in $\log I_{e,\lambda}$. The dashed line shows the maximum value, as $\Delta \log I_{e,\lambda} \approx \Delta \log(M_{\rm dyn}/L_\lambda)$. Both the observed and intrinsic scatter in the FP decrease toward longer wavelength, due to a decrease in the contribution from variations in M_*/L_λ.

in both the observed and intrinsic scatter toward longer wavelength, particularly so for the star-forming subsample. This reflects a lower contribution of M_*/L to the intrinsic scatter and suggests, unsurprisingly, that variations in the dust attenuation and recent star formation are most apparent at short wavelengths.

In the bottom panel we quantify the contribution of M_*/L variations using the SED-derived M_*/L estimates and a simple least-squares fit (to match the methods by Graves & Faber 2010). We note that we do not subtract the mean value of M_*/L along the (face-on) midplane, because the face-on FP is sparsely populated in comparison to the low-redshift studies, which together with the large uncertainties

on M_*/L makes a robust estimate of the mean M_*/L difficult. However, this mainly affects the uncertainty on the fit, and is unlikely to lead to a significant bias on the measured contribution of M_*/L.

We find that in the rest-frame g-band approximately 55% of the thickness of the FP is due to variations in M_*/L_g, with the contribution being slightly higher for star-forming galaxies ($\sim 58\%$, versus $\sim 51\%$ for quiescent galaxies). Stellar populations thus are the main driver of the intrinsic scatter in the g-band FP, exceeding the contributions of all other quantities examined in Section 2.4. On the other hand, Bernardi et al. (2020) recently showed that, for rotating systems, the use of the integrated velocity dispersion rather than the luminosity-weighted average of the second moment of the velocity (which is attainable from IFU data only; see Eq. 2.14) may also be a cause of substantial scatter in the FP. However, this additional scatter of approximately $\Delta \log \sigma \sim 0.03\,\mathrm{dex}$ (based on their Fig. A1) is still at least a factor ~ 3 lower than the contribution from stellar populations found here, and is further mitigated by the fact that this effect only becomes apparent in the case of very high S/N spectra.

Fig. 2.13 also shows that the dependence on M_*/L_λ is itself wavelength-dependent, such that the FP at longer wavelengths is less dominated by variations in M_*/L_λ. Interestingly, there is significant contribution from M_*/L_λ even at the longest wavelengths. Comparing with the results by Graves & Faber (2010) in the rest-frame r-band, we find that for our sample of quiescent galaxies the contribution from stellar populations is $\sim 42\%$. This is significantly higher than their measurement of 22% (for their preferred method of estimating M_*/L_r), but may be attributed to significant differences in the definition of quiescence: using D_n4000 as a proxy for age, if we select the 100 oldest (UVJ) quiescent galaxies in our sample, we find that variation in M_*/L_r contributes 23% to the thickness of the r-band FP.

Importantly, these measurements show that, under the assumption that the effects of dynamical non-homology are small (e.g., Bolton et al. 2008; Schechter et al. 2014), a significant fraction of the intrinsic scatter in the FP must arise variations in M_{dyn}/M_*, which may be due to variations in the IMF or the dark matter fraction. Our data currently lack a consistent measurement of the metallicity across the entire redshift range, as well as a measurement of the α-element abundance and other IMF-sensitive features (summarised in, e.g., van Dokkum & Conroy 2012), and we therefore cannot place constraints on the effect of IMF variations within the FP. On the other hand, we may expect the effect of IMF variations to be approximately as large as the uncertainties in the SED modelling (e.g., van de Sande et al. 2015), which would imply that the intrinsic scatter is dominated by fluctuations in the dark matter content.

For the quiescent LEGA-C galaxies, the significant correlations between the residuals from the luminosity FP and D_n4000 or $H\delta_A$, combined with the very weak correlations through the mass FP (Figs. 2.4 & 2.8), suggest that galaxies with younger luminosity-weighted ages, due to a later formation time or more extended star-formation history, have marginally higher values of $\Delta \log \Sigma_*$. If the effects of non-homology and IMF variations are small, this result implies that younger quiescent galaxies are slightly more baryon-dominated within $1\,R_e$. Although the

correlation between age and structure is very weak, in contrast with the strong correlation found by Graves & Faber (2010), this would be broadly consistent with the proposed scenario in which the truncation time of star formation determines the location of a galaxy within the parameter space of the FP.

However, the effect of galaxy merging, and how these trends apply to the star-forming population is still unclear. Recently, Ferrero et al. (2021) used cosmological hydrodynamical simulations to show that the tilt of the FP, of both star-forming and quiescent galaxies, can be explained entirely by variations in the dark matter fraction. A quantitative comparison with such simulations is challenging, as there are systematic mismatches between the observed and simulated sizes and velocity dispersions (van de Sande et al. 2019). However, hydrodynamical simulations of large volumes do qualitatively reproduce observed galaxy scaling relations, and therefore may also be able to shed light on the physical processes driving the intrinsic scatter in the FP, an analysis that we defer to a future work.

2.5.2 Structural non-homology

We have found that massive star-forming and quiescent galaxies lie on the same mass FP, with a comparable intrinsic scatter about the midplane (Fig. 2.7). Although the star-forming galaxies are typically slightly larger in size at fixed mass, their integrated velocity dispersion or stellar mass surface density tends to be lower, such that they fall on the same FP as the quiescent systems. The thickness of the mass FP is, unlike the g-band FP, largely uncorrelated with stellar population properties and can be interpreted as variation in $M_{\rm dyn}/M_*$. Under the assumption of a weakly varying IMF, the intrinsic scatter about the FP reflects a variation in the dark matter fraction within the effective radius.

Of particular interest then is the morphology, which we have modelled as a Sérsic profile. If the value of the Sérsic index reflects different underlying mass density profiles, we may expect it to correlate with the residuals in the mass FP. However, we find only a very weak correlation within the LEGA-C sample (Fig. 2.10). Interestingly, Bezanson et al. (2015) do find a weak dependence on Sérsic index within the mass FP at low redshift, for a sample of SDSS galaxies similar to the low-redshift sample considered here. In a different context, Cappellari et al. (2006) and Taylor et al. (2010) also demonstrate the importance of non-homology on the estimation of the dynamical mass of galaxies at $z \sim 0$. The lack of a correlation with Sérsic index in the mass FP in our sample is therefore surprising, as it seems to suggest that the dynamical masses of galaxies at $z \sim 1$ are independent of the observed Sérsic index. Any fluctuations in the dark matter fraction then simply reflect differences in the effective radii of galaxies, rather than the mass distribution itself.

This raises the question of how the difference in the structural dependence at $z \sim 0$ and $z \sim 1$ can be reconciled. One possibility is that the light profile evolves with redshift, while the underlying mass distribution does not change significantly, such that the mass FP is correlated with Sérsic index at $z \sim 0$, but not at $z \sim 1$. This scenario can be tested by measuring the colour gradients of galaxies to derive the Sérsic index and size of the stellar mass profile, instead of the rest-frame

5000 Å sizes used here. Suess et al. (2019) demonstrate that colour gradients are significantly steep especially at high stellar mass and are also dependent on redshift, and may therefore be important to take into account. Bernardi et al. (2019) show that, for a sample of very massive elliptical galaxies at $z \sim 0$, the accounting for stellar population gradients in galaxies can lead to a significant change in the inferred values of and variation in M_{dyn}/M_*. These gradients may then potentially act to wash out any significant dependence on Sérsic index through the thickness of the FP, although it is unclear how stellar population gradients affect the measurement of M_{dyn}/M_* for the population of late-type galaxies at $z \sim 0$, as well as galaxies at higher redshifts.

Secondly, if not a difference in the observed morphology, there may be differences in the derivation of the velocity dispersions between the various studies. For example, as opposed to the integrated velocity dispersion within $1\,R_{\mathrm{e}}$ used in this work, Taylor et al. (2010) use the central stellar velocity dispersion ($R_{\mathrm{e}}/8$); this difference in the aperture may lead to small systematic effects on the measured dispersions (see also Appendix 2.C). van Houdt et al. (2021) demonstrate using axisymmetric Jeans modelling that, at fixed mass, the dynamical masses of the LEGA-C galaxies do depend on Sérsic index, and do so in the same way as at $z \sim 0$. However, they also show that this dependence becomes apparent only when using the major axis size (rather than the circularised size) and after taking into account the effects of the slit aperture and the galaxy inclination (through the observed axis ratio) on the integrated velocity dispersion.

On the other hand, the lack of a residual correlation through the mass FP with Sérsic index does not imply that non-homology plays no role at all. The FP is tilted with respect to the virial plane, which may (in part) be due to a violation of the assumption of homology. Bezanson et al. (2013) compared the power-law relation between M_{dyn}/M_* and M_{dyn} for two different estimates of M_{dyn}, the first having a virial constant $K = 5$ (as in Section 2.D.1) and the second a Sérsic-dependent virial constant $K(n)$ (derived by Cappellari et al. 2006). At $z \sim 0$ the measured relation between M_{dyn}/M_* and M_{dyn} is slightly shallower for the Sérsic-dependent estimate of M_{dyn}, which indicates that non-homology contributes to the tilt of the FP, albeit a small effect (in agreement with findings by Cappellari et al. 2006). Moreover, Bernardi et al. (2020) show that by accounting for structural non-homology in their fits of the mass FP, as well as the galaxy inclination, they obtain a plane that is closer to the virial prediction.

By considering variations in the tilt of the mass FP, we observe a similar, weak effect. Taking the values of the tilt from Section 2.4.1, we find that the strongest tilt ($a = 1.432$ and $b = −0.736$) produces the weakest correlation with Sérsic index: $\Delta \log \Sigma_* \propto (−0.016 \pm 0.004)\,n$. Conversely, for the virial plane ($a = 2$ and $b = −1$) we find $\Delta \log \Sigma_* \propto (−0.026 \pm 0.005)\,n$. An evolution in the tilt, such that the mass FP becomes closer to the virial plane at higher redshift, may thus also bring the measurements at $z \sim 0$ and $z \sim 0.8$ into agreement.

2.6   Summary and conclusions

We have explored the connection between the structural and stellar kinematic properties of 1419 galaxies in the LEGA-C survey, which form a representative sample of the massive $(\log(M_*/M_\odot) > 10.5)$ galaxy population at $0.6 < z < 1$. In addition to the spectral and morphological properties obtained from the LEGA-C spectra and HST imaging respectively, we have performed SED modelling of multi-wavelength $(0.2 - 24\,\mu\mathrm{m})$ photometry to estimate stellar masses, as well as stellar population properties and the effect of dust attenuation. Separating our sample into star-forming and quiescent galaxies by the rest-frame UVJ colours, we have studied the effect of different structural, environmental and SED properties within the luminosity and mass FP. Our findings can be summarised as follows:

- There is significant scatter in the rest-frame g-band FP of quiescent galaxies, which exceeds the scatter due to measurement uncertainties. Star-forming galaxies also lie on the g-band FP, but with a different zero point and higher intrinsic scatter (Fig. 2.3). The residuals from the g-band FP correlate strongly with spectral age indicators ($\mathrm{D_n}4000$ and $\mathrm{H}\delta_\mathrm{A}$), as well as rest-frame colours ($U - V$, $V - J$). Using SED models, we interpret these correlations as being due to variation in the luminosity-weighted stellar age, and additionally for the star-forming sample, variation in the sSFR and dust attenuation.

- Both star-forming and quiescent galaxies lie on the same mass FP, with an approximately equal zero point and a comparable level of intrinsic scatter. In contrast with the g-band FP, we find no significant correlations in the residuals from the mass FP with different spectral and SED properties. Moreover, there is only a very weak correlation with Sérsic index and the observed axis ratio, corresponding to a minimal dependence on morphology for variations in M_dyn/M_* through the thickness of the FP.

- We evaluate the effect of environment on the FP, finding a very weak correlation between the residuals from the g-band FP and the projected galaxy overdensity, such that galaxies in high density environments have a marginally higher value of M_dyn/L, in line with previous studies that find galaxies at high overdensity to be slightly older. We find an even weaker correlation within the mass FP, suggesting that there is no significant structural difference between galaxies in low- and high-density environments at fixed size and velocity dispersion.

Overall, we find that variations in the M_*/L_g can account for $\sim 54\%$ of the thickness of the g-band FP. The other main contribution comes from variations in the dark matter content within $1\,R_\mathrm{e}$, or, variations in the IMF. Interestingly, the residuals in $\log \Sigma_*$ in the mass FP do not correlate strongly with morphology (Sérsic index), suggesting that the effect of structural non-homology is weak. Instead, variations in the galaxy size (at fixed mass) may play a more important role, as this leads to fluctuations in the dark matter fraction.

Future studies of IMF-sensitive spectral features or abundance measurements are required to quantify the role of IMF variations within the FP. On the other hand, the role of dark matter may well be explored with current cosmological hydrodynamical simulations, which are able to offer insight into the physical processes governing the properties of galaxies throughout the FP and the evolutionary processes that keep galaxies on the mass FP.

Acknowledgements

Table 2.1: Results of the MAGPHYS SED modelling.

ID	$\log(M_*/M_\odot)$	$\log(\mathrm{sSFR}/\mathrm{yr}^{-1})$	$\log(\mathrm{age/yr})$	A_V [mag]	$D_n 4000$
4792	$10.52^{+0.13}_{-0.00}$	$-10.32^{+0.30}_{-0.00}$	$9.26^{+0.21}_{-0.00}$	0.03	1.44
5786	$11.12^{+0.09}_{-0.05}$	$-10.72^{+0.10}_{-0.10}$	$9.41^{+0.16}_{-0.06}$	0.93	1.56
6859	$11.31^{+0.00}_{-0.00}$	$-11.02^{+0.00}_{-0.00}$	$9.20^{+0.00}_{-0.00}$	0.22	1.57
6890	$11.25^{+0.10}_{-0.09}$	$-11.02^{+0.35}_{-0.05}$	$9.24^{+0.04}_{-0.08}$	1.28	1.56
7002	$10.76^{+0.00}_{-0.09}$	$-10.87^{+0.00}_{-0.20}$	$9.36^{+0.00}_{-0.06}$	0.11	1.57

Notes. Values and formal error bars for the stellar mass, specific star formation rate, and luminosity-weighted age represent the 16th, 50th and 84th percentiles. The dust attenuation and Lick index $D_n 4000$ are measured from the best-fit SED. This table is available in its entirety in machine-readable format.

Appendix

2.A   Comparison of stellar mass estimates

In Section 2.2.2, we ran the MAGPHYS code (da Cunha et al. 2008) for broad-band photometry from the multi-wavelength catalogue by Muzzin et al. (2013b) to model the physical properties of the LEGA-C galaxies. We provide our catalogue of derived SED properties in Table 2.1. Our choice for MAGPHYS is motivated by our aim to minimise the systematic uncertainty in the measurement of the redshift evolution of the mass FP across $0 < z < 1$, and the public availability of the MAGPHYS modelling results for the SDSS by Chang et al. (2015). Our SED modelling differs from the results presented previously in van der Wel et al. (2016), who used the FAST code (Kriek et al. 2009) with different model assumptions and a different set of photometry.

In Fig. 2.14 we show a comparison between the best-fit stellar mass from van der Wel et al. (2016), and the median of the likelihood distribution of the stellar mass from MAGPHYS (as provided in Table 2.1). Red circles and blue triangles show quiescent and star-forming galaxies respectively, with the median shown as black open symbols. There is a clear offset between the two stellar mass estimates (blue and red dashed lines in the right-hand panel), with the masses inferred with MAGPHYS being systematically larger ($\sim 0.1-0.2\,\mathrm{dex}$). The offset is particularly significant for star-forming galaxies, and decreases slightly with increasing stellar mass.

One of the main differences between the modelling with FAST and MAGPHYS is the assumed form of the star formation history (SFH). The SFHs used for the FAST modelling are simply exponentially declining SFRs (τ models), whereas those for MAGPHYS additionally include random bursts of star formation. This can lead to significant changes in the inferred stellar ages and hence stellar masses, as the fits using the τ model SFHs can significantly underestimate the stellar mass (see, e.g., Pforr et al. 2012).
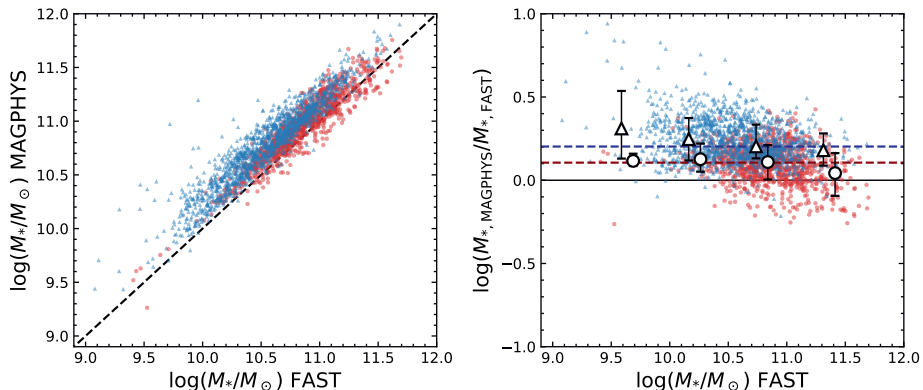
Figure 2.14: Comparison between the best-fit stellar mass from FAST (van der Wel et al. 2016) and the median of the stellar mass likelihood distribution from MAGPHYS for the primary sample of LEGA-C galaxies. Blue triangles and red circles indicate star-forming and quiescent galaxies respectively. In the right-hand panel, white markers show the median and 16th and 84th percentiles, with dashed lines indicating the median offset between the two mass estimates for the two populations. Stellar masses estimated with MAGPHYS are systematically larger than those from FAST, due to significant differences in the assumed star formation histories and dust attenuation model. The offset in stellar mass decreases slightly toward higher (FAST-derived) stellar mass and is larger for star-forming galaxies than quiescent galaxies, consistent with the findings by Leja et al. (2019a).

Moreover, the energy balance approach, combined with a different assumed dust model, may also change the inferred stellar mass. Whereas FAST applies a single dust screen, which in this case is the attenuation curve by Calzetti et al. (2000), MAGPHYS applies a two-component dust model (Charlot & Fall 2000) with different attenuation for stellar birth clouds and the diffuse interstellar medium, in better accord with observations of local galaxies (e.g., Calzetti et al. 2000).

Lastly, there are subtle differences in the photometry used. Although both works use the photometric catalogue by Muzzin et al. (2013b), van der Wel et al. (2016) use all available broad-band and medium-band filters, and exclude the *Spitzer*/MIPS 24 μm data. As the medium-band filters may suffer from large uncertainties in the zero points, and precise redshifts have already been measured from the LEGA-C spectra, we exclude these filters in our SED fitting. On the other hand, we do include the MIPS photometry, and make use of infrared libraries (da Cunha et al. 2008) and the energy balance recipe implemented in MAGPHYS to fit the mid-infrared data.

Our findings are broadly consistent with those by Leja et al. (2019a), who used a Bayesian approach to model the SEDs of galaxies at redshifts $0.5 < z < 2.5$ with a large number (14) of free parameters. They show that, in comparison with the results from FAST, the more complex model infers older stellar ages and therefore

systematically higher stellar masses, by $0.1 - 0.3$ dex. Moreover, similar to our result, the discrepancy between the two stellar mass estimates decreases slightly toward higher stellar mass. By using the stellar masses inferred with MAGPHYS, we therefore not only minimise systematic effects in our comparison of the mass FP at $z \sim 0$ and $z \sim 0.8$, but also adopt a stellar mass estimate that is likely to agree better with results from more sophisticated modelling.

2.B   Comparison of structural parameter estimates for the SDSS

In Sections 2.2.4, 2.3.1 and 2.4.1, we used the structural parameters measured in the r-band by Simard et al. (2011), which relies on imaging from the SDSS DR7, to measure the tilt of the FP of our low-redshift sample. However, by fitting Sérsic models on improved photometry from the SDSS DR9 for galaxies in the MaNGA survey (Bundy et al. 2015; Blanton et al. 2017), Fischer et al. (2019) show that the size estimates by Simard et al. (2011) may be biased. The resulting FP may therefore also change depending on the photometry and method of Sérsic modelling used.

Currently, there is no publicly available structural parameter catalogue that is based on the SDSS DR9 photometry for the larger spectroscopic sample of the SDSS. A direct assessment of the effect of this improved photometry on the FP is therefore not possible. Nevertheless, Fischer et al. (2017) demonstrate that the structural parameters measured by Meert et al. (2015) are largely unaffected by changes in the photometry, due to a different treatment of the sky background as compared with Simard et al. (2011).

By comparing the structural parameter catalogues by Simard et al. (2011) and Meert et al. (2015), we can therefore determine the extent to which the measured tilt of the FP depends on the catalogue used. Fig. 2.15 shows that the effective radii differ significantly between these two different catalogues, with the measurements by Simard et al. (2011) being systematically smaller toward larger radii (in agreement with findings by Fischer et al. 2019). This systematic discrepancy also affects the surface brightness and stellar mass surface density, which deviate most strongly toward low surface brightness or surface density.

Next, we evaluate the effect of these differences on the FP. We refit the FP using the catalogue by Meert et al. (2015) and following the methodology described in Sections 2.3.1 and 2.4.1. We note that we do not rederive the power-law coefficients of the corrections on the velocity dispersion (Appendix 2.C), as these corrections are very small and therefore are unlikely to have a significant effect on the measurement of the tilt. For the g-band FP, we find $a = 1.309 \pm 0.014$ and $b = -0.726 \pm 0.003$, which is in excellent agreement with the results found in Section 2.3.1, where we used the catalogue by Simard et al. (2011). Similarly, we find good agreement for the mass FP, with $\alpha = 1.437 \pm 0.012$ and $\beta = -0.730 \pm 0.003$. Given the large discrepancies found in Fig. 2.15, this may be surprising. However, in Fig. 2.16 we show that the change in the FP due to changes in the effective radii are relatively small, which can be explained by the fact that the uncertainties in
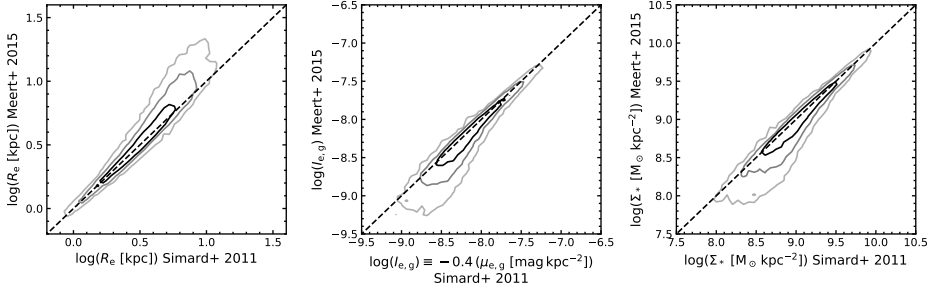
Figure 2.15: Comparison between the structural parameter catalogues by Simard et al. (2011) and Meert et al. (2015) for our SDSS sample at $z \approx 0.06$ (Section 2.2.4). Contours enclose 50%, 80% and 90% of the total sample, respectively. The two estimates of the effective radius (left) agree well for small galaxies, but become increasingly divergent at large radii. Correspondingly, the surface brightness (middle) and stellar mass surface density (right) are in strongest disagreement at low surface brightness and surface density. Despite these discrepancies, the tilt of the g-band and mass FP are unchanged when using the catalogue by Meert et al. (2015) rather than the Simard et al. (2011) catalogue, which can be attributed to the covariance between the galaxy size and surface brightness or stellar mass surface density (see Fig. 2.16).
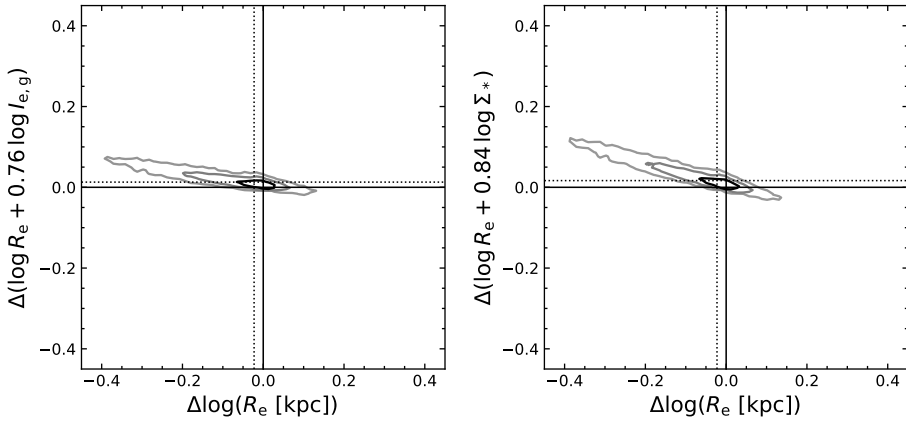
Figure 2.16: The change in the g-band FP (left) and mass FP (right) due to differences in the size estimates between the Simard et al. (2011) and Meert et al. (2015) catalogues, assuming a fixed tilt from Hyde & Bernardi (2009). Contours enclose 50%, 80% and 90% of the total sample, respectively, and dotted lines indicate the median values. Even a large change in the effective radius results in only a minor difference in the FP, which demonstrates that the uncertainties in $\log R_e$ and $\log I_{e,g}$ or $\log \Sigma_*$ are largely correlated along the FP. As a result, the tilt of the g-band and mass FP depend only very weakly on the choice of the structural parameter catalogue used.

$\log R_{\mathrm{e}}$ and $\log I_{\mathrm{e,g}}$ or $\log \Sigma_*$ correlate in a direction that is near-parallel to the FP itself. We therefore conclude that, although there are significant changes in the structural parameters between different catalogues, the FP itself is insensitive to these differences.

2.C   Velocity dispersion aperture corrections

As discussed in Section 2.2.1, the integrated velocity dispersion depends on the intrinsic velocity dispersion as well as the rotational velocity of a galaxy. The profiles of these quantities will vary with radius, and the integrated velocity dispersion will therefore depend on the aperture of the spectrum. The spectra of the SDSS galaxies (Section 2.2.4) were obtained with fibres that are $3''$ in diameter, whereas a typical galaxy in our sample at $z \approx 0.06$ has an effective radius of $r_{\mathrm{e}} \approx 5''$. The variation in galaxy sizes within the sample, and radial gradients in the integrated velocity dispersion may therefore lead to systematic uncertainties in the measured scaling relations. To derive a correction for the SDSS fibre velocity dispersions (σ_{fibre}) to the dispersion within a common physical aperture of $1\,r_{\mathrm{e}}$ (σ_{e}), we investigate the dependence of the integrated velocity dispersion on the aperture size and structural properties using integral field unit (IFU) spectroscopy.

We match the IFU data from the MaNGA survey of the SDSS DR15 (Bundy et al. 2015; Blanton et al. 2017) with our catalogue from the SDSS DR7, as well as the Sérsic profile fits by Simard et al. (2011). We select galaxies in the same way as in Section 2.2.4, but allow for a slightly wider redshift range of $0.04 < z < 0.08$ (median $z = 0.054$), and require that the flags `DAPQUAL=0` and `DRP3QUAL=0`, resulting in a selection of 702 galaxies. For each galaxy, we use the publicly available maps of the observed stellar velocity dispersion (corrected for the effect of instrumental resolution) and velocity field (Westfall et al. 2019) to calculate the flux-weighted second moment of the line-of-sight velocity:

$$\sigma_{\mathrm{aper}}^2 = \frac{\sum_i F_i \left(v_i^2 + \sigma_i^2\right)}{\sum_i F_i}, \tag{2.14}$$

where F_i is the g-band flux, v_i the velocity with respect to the galaxy centre and σ_i the observed velocity dispersion measured in the ith Voronoi bin. Bins are included only if at least 80% of their area lies within the specified aperture. We calculate σ_{aper} for two different apertures: circular apertures of $3''$ in diameter (σ_{3as}), and elliptical apertures defined by the effective radius (σ_{e}).

We use the results from the circular apertures to determine whether there are significant systematic effects between the velocity dispersions from the SDSS fibre spectra and MaNGA data, which can be due to differences in the observations themselves or in the reduction and analysis of the spectra. We find good agreement between the two measurements: there is only a small systematic offset, with a median of $(\sigma_{\mathrm{fibre}} - \sigma_{\mathrm{3as}}) = -3.0\,\mathrm{km\,s^{-1}}$, and scatter of 0.065 in the fractional difference ($\Delta\sigma = [\sigma_{\mathrm{fibre}} - \sigma_{\mathrm{3as}}]/\sigma_{\mathrm{3as}}$).

Next, we use the ratio of $\sigma_{\mathrm{fibre}}/\sigma_{\mathrm{e}}$ to examine the effect of aperture size. In Fig. 2.17, we show $\sigma_{\mathrm{fibre}}/\sigma_{\mathrm{e}}$ for all galaxies (grey symbols) as a function of their
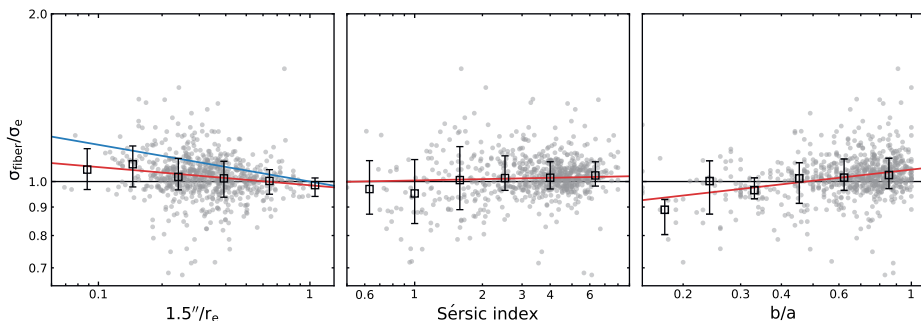
Figure 2.17: Ratio of the integrated velocity dispersion from the $3''$ SDSS fibre spectra and the MaNGA spectra within an aperture of one effective radius, calculated as the flux-weighted second velocity moment. Different panels show the dependence of this ratio on the effective radius, Sérsic index, and axis ratio in grey. Black squares and error bars represent the median and 16th and 84th percentiles. Red lines are the best-fit power laws for each parameter. For comparison, the left-hand panel also shows the result by Cappellari et al. (2006) in blue.

structural parameters (the circularised effective radius, Sérsic index, and axis ratio). Medians and percentiles (16th, 84th) are shown in black. There is a weak correlation between $\sigma_{\mathrm{fibre}}/\sigma_{\mathrm{e}}$ and the ratio of the aperture size, indicating a declining profile in the integrated velocity dispersion. On the other hand, for the few galaxies with low Sérsic index, σ_{fibre} appears to be systematically lower than σ_{e}, which may reflect a missing contribution from the rotational velocity. The third panel demonstrates this effect more clearly: for flattened systems, $\sigma_{\mathrm{fibre}}/\sigma_{\mathrm{e}}$ is significantly lower than for rounder objects.

Aperture corrections derived in previous studies usually take into account only the dependence on the ratio of the aperture and the effective radius ($r_{\mathrm{aper}}/r_{\mathrm{e}}$). E.g., Jorgensen et al. (1996) and Cappellari et al. (2006) derive a correction of the form:

$$\left(\frac{\sigma_{\mathrm{aper}}}{\sigma_{\mathrm{e}}}\right) = \left(\frac{r_{\mathrm{aper}}}{r_{\mathrm{e}}}\right)^{\alpha} . \tag{2.15}$$

Here, we use $\sigma_{\mathrm{aper}} = \sigma_{\mathrm{fibre}}$ and $r_{\mathrm{aper}} = 1.5''$, and also fit a power law relation to the Sérsic index and axis ratio:

$$\left(\frac{\sigma_{\mathrm{fibre}}}{\sigma_{\mathrm{e}}}\right) = \left(\frac{4}{n}\right)^{\beta} , \tag{2.16}$$

and

$$\left(\frac{\sigma_{\mathrm{fibre}}}{\sigma_{\mathrm{e}}}\right) = \left(\frac{0.6}{b/a}\right)^{\gamma} . \tag{2.17}$$

We fit each parameter separately, and take into account the small systematic offset between σ_{fibre} and σ_{3as}. The best-fit power law is shown in red in each panel in Fig. 2.17, which have exponents $\alpha = -0.033 \pm 0.003$, $\beta = -0.008 \pm 0.010$

and $\gamma = -0.067 \pm 0.012$. We also show the result by Cappellari et al. (2006) in blue, who used IFU spectroscopy for a sample of elliptical and lenticular galaxies and found a steeper relation of $\alpha = -0.066 \pm 0.035$. Importantly, however, our selection differs significantly from their sample, as we have not selected galaxies by morphology. Finally, we multiply the three correction factors and correct for the systematic offset between the SDSS fibre and MaNGA data, to calculate σ_{e} for each SDSS galaxy in our selection in Section 2.2.4. The correction to σ_{e} is typically small, with an average of 3%.

2.D   Tilt of the Fundamental Plane

Throughout this work we have assumed minimal evolution in the tilt of the FP and used a measurement of the tilt at low redshift, as accurate fitting of the FP is highly complex, and our results do not depend strongly on the assumed tilt. However, in Sections 2.3.1 & 2.4.1 we showed there is weak evidence for an evolution in the tilt of the FP, particularly so for the g-band FP. These measurements relied on a relatively simple planar fit to a subset of the data that is most complete in mass. Here, we further examine the redshift evolution of the tilt of the FP for the full sample of LEGA-C galaxies, and additionally take into account the measurement uncertainties and the effect of sample completeness.

We begin by writing the luminosity FP and mass FP as the power-law relations

$$R_{\mathrm{e}} \propto \sigma^a\, I_{\mathrm{e}}^b \quad \text{and} \quad R_{\mathrm{e}} \propto \sigma^\alpha\, \Sigma_*^\beta \,, \tag{2.18}$$

where $I_{\mathrm{e}} \propto L/R_{\mathrm{e}}^2$ and $\Sigma_* \propto M_*/R_{\mathrm{e}}^2$. Under the assumption of homology, i.e. $M_{\mathrm{dyn}} \propto R_{\mathrm{e}}\, \sigma^2$, the FP can be rewritten as a power-law relation between M_{dyn}, R_{e}, and M_{dyn}/L or M_{dyn}/M_* (for a full derivation, see Cappellari et al. 2006):

$$M_{\mathrm{dyn}}/L \propto M_{\mathrm{dyn}}^d\, R_{\mathrm{e}}^f \,, \tag{2.19}$$

or

$$M_{\mathrm{dyn}}/M_* \propto M_{\mathrm{dyn}}^\delta\, R_{\mathrm{e}}^\eta \,, \tag{2.20}$$

where the exponents d and f (or δ, η) depend on the tilt of the FP. If $f \ll d$ ($\eta \ll \delta$), then the tilt of the FP reflects the relation between M_{dyn}/L (M_{dyn}/M_*) and mass, as first proposed by Faber et al. (1987).

A measurement of the tilt depends strongly on the methodology used (e.g., a direct planar fit versus an orthogonal fit, see Hyde & Bernardi 2009), the sample completeness, and uncertainties on different parameters as well as their covariances (see also Magoulas et al. 2012). However, we can reduce some of these uncertainties by calibrating the relation between M_{dyn}/L and M_{dyn} directly (under the assumption that $f \ll d$, $\eta \ll \delta$), using an estimate of M_{dyn}:

$$M_{\mathrm{dyn}} = K\, \frac{R_{\mathrm{e}}\, \sigma^2}{G} \,, \tag{2.21}$$

where G is the gravitational constant and K is the virial coefficient, which in general depends on the structural properties of the galaxy. We set $K = 5$, which

was shown by Cappellari et al. (2006) to provide a good approximation for early-type galaxies. This may not provide a good estimate of K for late-type galaxies, however, the effect of the assumed virial constant, e.g. a Sérsic-dependent virial constant, is small (see also Fig 2.10 and Section 2.5).

2.D.1   Direct measurement of M_{dyn}/L_g vs. M_{dyn}

We show the relation between M_{dyn}/L_g and M_{dyn} in Fig. 2.18, in bins of $\Delta z = 0.10$. As previous measurements of the FP focused solely on early-type galaxies, we consider both the quiescent population alone (red circles), as well as the combined sample of quiescent and star-forming (blue) galaxies.

There is a strong correlation between M_{dyn}/L_g and M_{dyn}, in part due to the covariance between the two quantities. Moreover, the effect of sample incompleteness becomes apparent from this figure: our S/N selection on the velocity dispersion approximately translates to a selection on the rest-frame g-band magnitude, illustrated in Fig. 2.18 by shaded regions that cover $m_g > 22.5$. Toward lower M_{dyn} as well as higher redshift, this contributes to an apparent steepening of the observed power-law relation.

To estimate the exponent d, we therefore exclude the lowest-mass galaxies, requiring $\log(M_{\mathrm{dyn}}/M_\odot) > 10.6$. Fitting in logarithmic space, we use the orthogonal distance regression described by Hogg et al. (2010, Eq. 35), which takes into account the uncertainties in both axes and treats the (Gaussian) intrinsic variance orthogonal to the linear fit as a free parameter. We use the measurement uncertainties to estimate the covariance matrix for each galaxy with 1000 Monte Carlo simulations.

To account for the fact that galaxies of high M_{dyn}/L_K are less likely to be observed, we use the completeness correction Tcor (Section 2.3.1) to weight the covariance matrices, however, we now firstly renormalise these corrections in bins of M_{dyn}. The effect of this normalisation is (i) that at fixed M_{dyn} galaxies of high M_{dyn}/L receive a greater weight, and (ii) that galaxies of low M_{dyn}, where the completeness in M_{dyn}/L is lowest, do not introduce an extreme bias on the measured exponent.

For the quiescent sample, we find a weak evolution in the exponent between $0.6 \leq z < 0.9$ (Table 2.2), with measurements deviating by $\sim 1-2\sigma$ (where uncertainties on the fits are obtained from bootstrapping the data). The largest discrepancy is between the lowest ($0.6 \leq z < 0.7$) and highest redshift bins ($0.9 \leq z \leq 1.0$), which deviate by 2.2σ. However, in the highest redshift bin there are relatively few galaxies at low mass (see also Fig. 2.2), and the fit is therefore most affected by sample incompleteness.

Our measurements for the quiescent sample agree well with the results by Jørgensen & Chiboucas (2013), who measured $d = 0.44 \pm 0.09$ at $z = 0.54$ and $d = 0.55 \pm 0.08$ for quiescent cluster galaxies at $z \approx 0.85$ in the rest-frame B-band, as well as the work by Saracco et al. (2020), who found $d = 0.6 \pm 0.1$ at $z \approx 1.3$. Importantly, in both works the galaxy samples were selected in the I-band, which introduces a selection effect similar to the shaded regions in Fig. 2.18, and thus can steepen the inferred power-law.
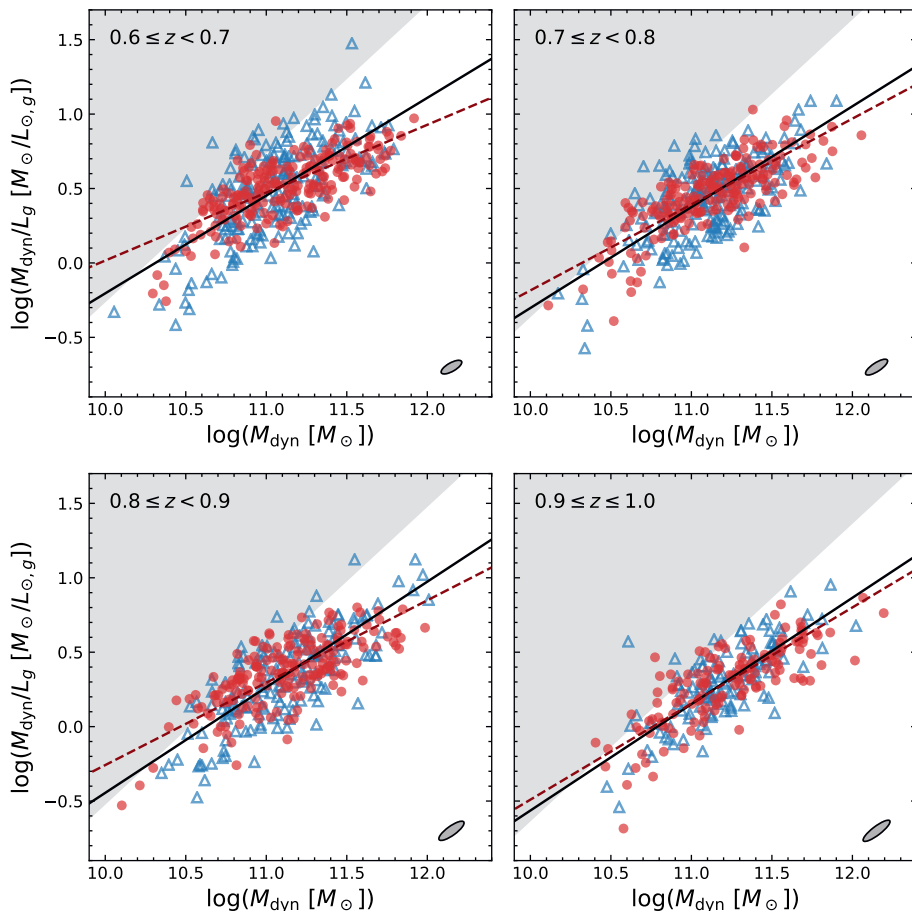
Figure 2.18: Relation between the dynamical mass-light ratio (M_{dyn}/L_g) and dynamical mass as a function of redshift. Dashed lines show linear fits to the quiescent galaxies (red circles) in each redshift bin; black lines show the result for the combined sample of quiescent and star-forming (blue triangles) galaxies (see Table 2.2). Ellipses show the typical measurement uncertainties. Shaded regions mark galaxies of $m_{\mathrm{g,rest}} > 22.5$ and illustrate the effect of sample selection in the I-band, common in previous studies of the FP, or a S/N limit for the velocity dispersion. The slope of the relation between M_{dyn}/L_g and M_{dyn} varies weakly with redshift, which can partially be attributed to incompleteness at high M_{dyn}/L_g toward higher redshift and lower M_{dyn}. As the slope is analogous to the tilt of the g-band FP, there is likely also a weak evolution in the tilt of the FP. Moreover, there is strong evolution in the intercept, as is expected from evolution in the stellar populations (see also de Graaff et al. 2020).

Table 2.2: Best-fit M_{dyn}/L_g vs. M_{dyn} exponents.

Redshift	Quiescent		Quiescent + star-forming	
	N_{gal}	d	N_{gal}	d
$0.05 < z < 0.07$	13,468	0.386 ± 0.004	20,508	0.514 ± 0.005
$0.6 \leq z < 0.7$	202	0.46 ± 0.04	411	0.64 ± 0.06
$0.7 \leq z < 0.8$	183	0.58 ± 0.06	393	0.67 ± 0.05
$0.8 \leq z < 0.9$	180	0.55 ± 0.05	349	0.66 ± 0.06
$0.9 \leq z \leq 1.0$	138	0.64 ± 0.07	266	0.70 ± 0.06

Notes. Only galaxies of $\log(M_{\mathrm{dyn}}/M_\odot) > 10.6$ are included in the fits.

Lastly, we compare our measured relations with the low-redshift SDSS sample, using Eq. 2.21 (with $K = 5$) to estimate M_{dyn} and applying the same fitting procedure to estimate d. The measurement relation is shallower than the LEGA-C measurements by $2 - 3\sigma$ ($0.6 < z < 0.9$), suggesting a weak evolution with redshift or a selection bias against galaxies that are faint in the rest-frame g-band, or, most likely, a combination of both.

We find a steepening of the slope when we include both quiescent and star-forming galaxies in the fit. Interestingly, there is no significant evolution within the LEGA-C sample ($0.6 < z < 1.0$) in this case, suggesting that the effect of sample selection is partially driving the observed evolution in d for the quiescent LEGA-C galaxies. On the other hand, there is weak evolution (at a level of $\approx 3\sigma$) in comparison with the SDSS fit when considering the full sample of star-forming and quiescent galaxies, pointing toward physical differences between the low and high-redshift samples.

In summary, we find evidence for an evolution in the power-law relation between M_{dyn}/L_g and M_{dyn} with redshift, and, by extension, the tilt of the FP. This evolution can be explained by an increasingly declining M_{dyn}/L for less massive galaxies toward higher redshift: this evolution is expected, as more galaxies are star-forming at higher redshifts, and this effect is strongest at lower masses ("downsizing", Cowie et al. 1996; Brinchmann et al. 2004). However, we find that the effects of sample selection and completeness also contribute significantly the observed redshift evolution. To determine the extent to which the evolution is of a physical origin, will require a more careful analysis of the selection function of both the SDSS and LEGA-C samples.

2.D.2   Direct measurement of M_{dyn}/M_* vs. M_{dyn}

Fig. 2.19 shows the relation between M_{dyn}/M_* and M_{dyn} at different redshifts, again with red and blue markers showing the quiescent and star-forming LEGA-C galaxies, respectively. The gray regions now illustrate the effect of our selection on stellar mass, which we used in Section 2.2.3 to homogenize our sample. The best-fit power laws are shown as dashed lines and solid lines (representing fits to the quiescent and full samples, respectively), the exponents (δ) of which are presented
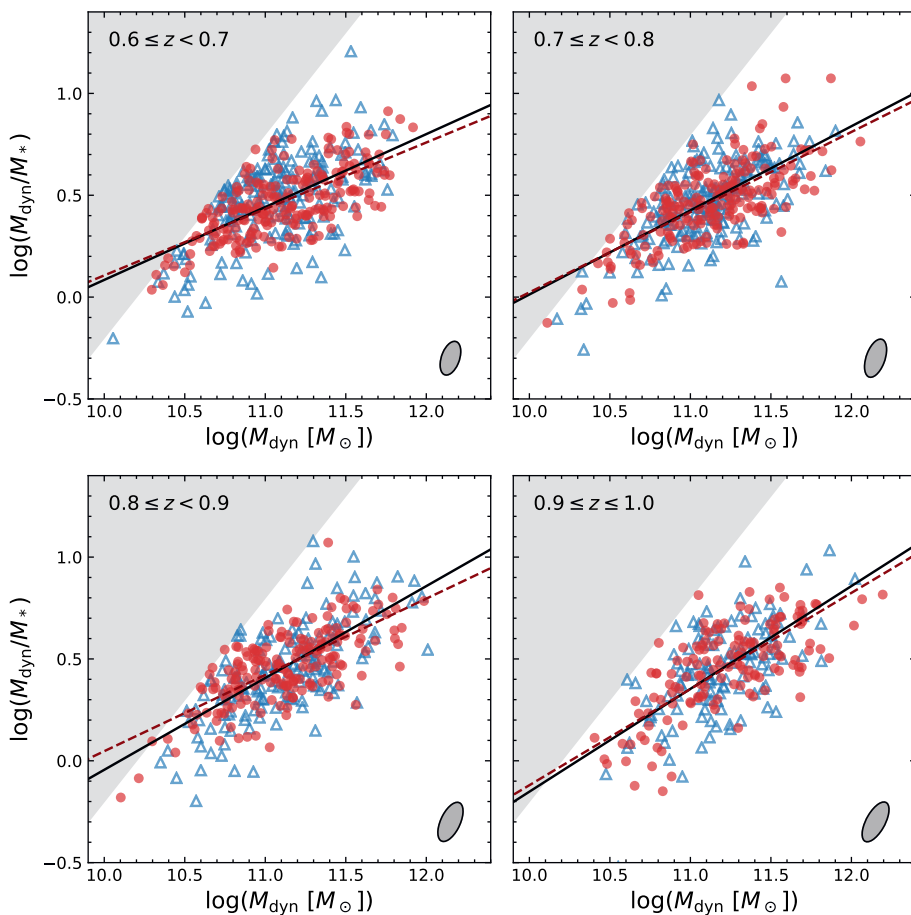
Figure 2.19: Relation between the dynamical-to-stellar mass ratio (M_{dyn}/M_*) and dynamical mass as a function of redshift. Symbols indicate the same as in Fig. 2.18. The effect of our selection on stellar mass $(\log(M_*/M_\odot) \geq 10.5)$ is indicated by the shaded regions. We find no significant variation with redshift in the slope of the relation between $\log(M_{\mathrm{dyn}}/M_*)$ and $\log(M_{\mathrm{dyn}})$ for the quiescent galaxies, and a weak evolution when considering the full sample (see Table 2.3).

in Table 2.3.

For the quiescent galaxies, we find no evolution in δ between $0.6 < z < 0.9$ and a reasonable agreement (a deviation $< 1.5\sigma$) with the measurement at low redshift. The highest redshift bin does diverge more strongly, but, as is apparent from Fig. 2.19, this measurement is likely strongly affected by an incompleteness in M_{dyn}/M_* at $\log(M_{\mathrm{dyn}}/M_\odot) \lesssim 10.9$. These results are consistent with our findings for the tilt of the FP in Section 2.4.1, as well as previous work by Bundy et al. (2007), who found no evolution in the relation between M_{dyn} and M_* between

Table 2.3: Best-fit M_{dyn}/M_* vs. M_{dyn} exponents.

Redshift	Quiescent		Quiescent + star-forming	
	N_{gal}	d	N_{gal}	d
Redshift	Quiescent		Quiescent + star-forming	
	N_{gal}	δ	N_{gal}	δ
$0.05 < z < 0.07$	13,468	0.313 ± 0.004	20,508	0.330 ± 0.003
$0.6 \leq z < 0.7$	202	0.33 ± 0.04	411	0.36 ± 0.03
$0.7 \leq z < 0.8$	183	0.39 ± 0.05	393	0.41 ± 0.04
$0.8 \leq z < 0.9$	180	0.37 ± 0.04	349	0.45 ± 0.04
$0.9 \leq z \leq 1.0$	138	0.50 ± 0.07	266	0.53 ± 0.05

**Notes.** Only galaxies of $\log(M_{\mathrm{dyn}}/M_\odot) > 10.6$ are included in the fits.

$z \sim 0$ and $z \sim 1$ for a sample of spheroidal galaxies.

When considering the combined sample of quiescent and star-forming galaxies, we do observe a weak evolution in δ, as we measure a slight steepening with redshift both within the LEGA-C sample itself and in comparison with the SDSS data. Moreover, these exponents are steeper than the fits to the quiescent galaxies for all redshift ranges. Interestingly, this is opposite to the result of an orthogonal fit to the FP (Section 2.4.1), where we find that the tilt of the FP is slightly closer to that of the virial plane at $z \sim 0.8$ than at $z \sim 0$. Additionally, the values of α and β are (marginally) closer to the virial prediction for the full sample than for the quiescent sample alone: $\alpha = 1.64 \pm 0.09$ and $\beta = -0.71 \pm 0.02$ (LEGA-C, $0.65 < z < 0.75$), whereas the fit to the quiescent sample gives $\alpha = 1.56 \pm 0.12$ and $\beta = -0.68 \pm 0.03$ ($\alpha = 1.467 \pm 0.014$ and $\beta = -0.730 \pm 0.004$ versus $\alpha = 1.432 \pm 0.012$ and $\beta = -0.736 \pm 0.003$, respectively, for the SDSS).

This apparent contradiction may reflect an increasing incompleteness in M_{dyn}/M_* toward higher redshift, with the difference in δ between the two samples at low redshift being due to the selection on M_* and the maximum allowed uncertainty on the velocity dispersion. We indeed find that the measurements for the SDSS data are in good agreement when we relax our stellar mass limit, with $\delta = 0.285 \pm 0.004$ and $\delta = 0.287 \pm 0.003$ for the quiescent and full SDSS samples, respectively. Alternatively, whereas the variations in M_{dyn}/L at fixed M_{dyn} are largely due to variations in M_*/L, the variations in M_{dyn}/M_* depend on variations in the IMF (and radial gradients therein; see Bernardi et al. 2019) as well as the dark matter content, which in turn depends on the galaxy size (as discussed in Section 2.5). Therefore, the discrepancy between the measurement of the tilt and the measurement of the relation between M_{dyn}/M_* and M_{dyn} may indicate that (i) the effects of stellar population gradients on the measurement of M_{dyn}/M_* cannot be neglected, or (ii) our assumption of minimal R_{e} dependence ($\eta \ll \delta$; Eq. 2.20) no longer holds, and that the measurement of δ alone therefore is insufficient to draw conclusions on the evolution of the tilt of mass FP.

3 | Tightly coupled morpho-kinematic evolution for massive star-forming and quiescent galaxies across 7 Gyr of cosmic time

ABSTRACT

We use the Fundamental Plane (FP) to measure the redshift evolution of the dynamical mass-to-light ratio (M_{dyn}/L) and the dynamical-to-stellar mass ratio (M_{dyn}/M_*). Although conventionally used to study the properties of early-type galaxies, we here obtain stellar kinematic measurements from the Large Early Galaxy Astrophysics Census (LEGA-C) Survey for a sample of ~ 1400 massive ($\log(M_*/M_\odot) > 10.5$) galaxies at $0.6 < z < 1.0$ that span a wide range in star formation activity. In line with previous studies, we find a strong evolution in M_{dyn}/L_g with redshift. In contrast, we find only a weak dependence of the mean value of M_{dyn}/M_* on the specific star formation rate, and a redshift evolution that likely is explained by systematics. Therefore, we demonstrate that star-forming and quiescent galaxies lie on the same, stable mass FP across $0 < z < 1$, and that the decrease in M_{dyn}/L_g toward high redshift can be attributed entirely to evolution of the stellar populations. Moreover, we show that the growth of galaxies in size and mass is constrained to occur within the mass FP. Our results imply either minimal structural evolution in massive galaxies since $z \sim 1$, or a tight coupling in the evolution of their morphological and dynamical properties, and establish the mass FP as a tool for studying galaxy evolution with low impact from progenitor bias.

Anna de Graaff, Rachel Bezanson, Marijn Franx, Arjen van der Wel,
Bradford Holden, Jesse van de Sande, Eric Bell, Francesco D'Eugenio,
Michael Maseda, Adam Muzzin, Camilla Pacifici, David Sobral,
Caroline Straatman, Po-Feng Wu

3.1   Introduction

Galaxies obey a tight scaling relation between size, velocity dispersion, and surface brightness or stellar mass surface density, known as the Fundamental Plane (FP) (e.g., Djorgovski & Davis 1987; Dressler et al. 1987; Jorgensen et al. 1996). The tilt and zero point of the luminosity FP are directly related to the dynamical mass-to-light ratio (M_{dyn}/L) (Faber et al. 1987), and the FP has therefore proven to be a valuable tool in studying the evolution in M_{dyn}/L of the quiescent galaxy population. The zero point in particular has been shown to evolve significantly with redshift, which places strong constraints on the formation epoch of massive quiescent galaxies (e.g., van Dokkum & Franx 1996; van der Wel et al. 2005).

However, Saglia et al. (2010, 2016) and Toft et al. (2012) suggest that evolution in the morphological or kinematic structure may be required to fully account for the observed evolution in the FP. Bezanson et al. (2013), on the other hand, demonstrate that when the surface brightness parameter in the FP is replaced by the stellar mass surface density, there is very little evolution in the resulting mass FP of massive quiescent galaxies to $z \sim 2$. These observations suggest that any redshift dependence of M_{dyn}/L is caused primarily by evolution in the stellar mass-to-light ratio (M_*/L), and that changes in the structure-dependent ratio of the total and stellar mass (M_{dyn}/M_*) are either minimal, or embedded in the FP.

High-redshift studies of the FP thus far, however, have been limited in sample size, with selections being biased toward either the densest environments or brightest objects (e.g., Holden et al. 2010; van de Sande et al. 2014; Beifiori et al. 2017; Prichard et al. 2017; Saracco et al. 2020), which populate the FP differently than typical galaxies in the field (see, e.g., Saglia et al. 2010; van de Sande et al. 2014). Extending these analyses to a more representative sample of the overall galaxy population is therefore crucial to understand the redshift evolution in M_{dyn}/L and M_{dyn}/M_*.

At low redshift, Zaritsky et al. (2008) and Bezanson et al. (2015) have shown that star-forming galaxies lie on the same surface as the quiescent galaxies, if both M_*/L and rotation velocities are taken into account. In de Graaff et al. (2021) we present the luminosity and mass FP of a large, K_{s}-band selected sample of galaxies drawn from the Large Early Galaxy Astrophysics Census (LEGA-C) Survey (van der Wel et al. 2016; Straatman et al. 2018), and find that star-forming and quiescent galaxies also lie on the same mass FP at $z \sim 0.8$.

In this Letter, we constrain the redshift evolution of the luminosity FP and mass FP between $0 < z < 1$, by using our representative sample of massive galaxies from the LEGA-C survey and a reference sample of local galaxies from the Sloan Digital Sky Survey (SDSS).

We assume a flat ΛCDM cosmology throughout, with $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and $\Omega_{\mathrm{m}} = 0.3$.

3.2   Data

3.2.1   LEGA-C Survey

Our sample is drawn from the third data release of the LEGA-C survey, a deep spectroscopic survey of ~ 3000 K_s-band selected galaxies at $0.6 < z < 1.0$ in the COSMOS field (van der Wel et al. 2016; Straatman et al. 2018), which provides accurate absorption line widths for a representative sample of the massive galaxy population at $z \sim 0.8$.

We describe the combined data set and our sample selection in detail in de Graaff et al. (2021). Briefly, we measure integrated stellar velocity dispersions, to which both the intrinsic velocity dispersion and projected rotational motions contribute, from the LEGA-C spectra (see Bezanson et al. 2018b; Straatman et al. 2018). We obtain structural parameters by fitting Sérsic profiles to ACS F814W imaging from the *Hubble Space Telescope* with GALFIT (Peng et al. 2010), and circularise the effective radii (i.e. $R_\mathrm{e} = \sqrt{ab}$). We derive stellar masses by fitting the galaxy spectral energy distributions (SEDs) with MAGPHYS (da Cunha et al. 2008) and measure rest-frame luminosities with EAZY (Brammer et al. 2008), using the multi-wavelength ($0.2 - 24\,\mu\mathrm{m}$) photometric catalogue by Muzzin et al. (2013b). We correct all masses and luminosities for missing flux using the total luminosity of the best-fit Sérsic profile (e.g., Taylor et al. 2010).

We select galaxies of stellar mass $\log(M_*/M_\odot) \geq 10.5$, and require a maximum uncertainty of 15% on the velocity dispersion. Moreover, we require that the GAL-FIT fit has converged, and remove galaxies which are significantly morphologically disturbed. Our final sample consists of 1419 galaxies. We use the rest-frame $U - V$ and $V - J$ colours and the selection criteria by Muzzin et al. (2013a) to define quiescent and star-forming subsamples.

3.2.2   SDSS

We obtain a reference sample of galaxies at $0.05 < z < 0.07$ from the 7th data release of the SDSS (Abazajian et al. 2009), matching the selection criteria and observables as closely as possible to the LEGA-C sample. Our selection and aperture corrections are detailed in de Graaff et al. (2021). Briefly, we require a maximum uncertainty on the stellar velocity dispersion of 15%, and correct the velocity dispersions to an aperture of $1\,R_\mathrm{e}$. We use stellar masses estimated from SED fitting with MAGPHYS (Chang et al. 2015), and structural parameters derived from single Sérsic models in the r-band (Simard et al. 2011). We circularise the effective radii, and correct all stellar masses using the total luminosity of the best-fit Sérsic profile. Our selection consists of 23,036 massive galaxies ($\log(M_*/M_\odot) \geq 10.5$).

Rest-frame colours and luminosities are calculated using KCORRECT (Blanton & Roweis 2007), and we differentiate between quiescent and star-forming galaxies using the rest-frame $u - r$ and $r - z$ colours and the criteria from Holden et al. (2012).

3.3 Evolution in M_{dyn}/L

The FP in luminosity, here taken as the rest-frame g-band luminosity, has the form:

$$\log R_{\mathrm{e}} = a \log \sigma + b \log I_{\mathrm{e,g}} + c, \tag{3.1}$$

where R_{e} is the effective radius, σ the integrated stellar velocity dispersion, and $I_{\mathrm{e,g}} = -0.4\,\mu_{\mathrm{e,g}}$, where $\mu_{\mathrm{e,g}}$ is the mean surface brightness within the effective radius, corrected for cosmological surface brightness dimming (see, e.g., Hyde & Bernardi 2009). The coefficients a and b describe the tilt of the plane, and c is the zero point.

We assume that the tilt of the FP does not evolve strongly with redshift (as shown in Holden et al. 2010, de Graaff et al. in prep.), and adopt the tilt derived by Hyde & Bernardi (2009), of $a = 1.404$ and $b = -0.761$, for both the SDSS and LEGA-C samples. We fit the zero point c of the FP for the SDSS sample by minimising the mean absolute orthogonal deviations from the FP,

$$\Delta_{\mathrm{FP}} = \frac{|\log R_{\mathrm{e}} - a \log \sigma - b \log I_{\mathrm{e,g}} - c|}{\sqrt{1 + a^2 + b^2}}. \tag{3.2}$$

Next, we determine for each LEGA-C galaxy the difference in $\log(M_{\mathrm{dyn}}/L_g)$ with respect to the SDSS sample, by firstly calculating the residual of the FP in $\log I_{\mathrm{e,g}}$:

$$\Delta \log I_{\mathrm{e,g}} = - \left(\Delta_{\mathrm{LFP}} - c_0 \right)/b, \tag{3.3}$$

where c_0 is the best-fit zero point to the SDSS data, and

$$\Delta_{\mathrm{LFP}} = \log R_{\mathrm{e}} - a \log \sigma - b \log I_{\mathrm{e,g}}. \tag{3.4}$$

We then make the common assumption that $\Delta \log I_{\mathrm{e,g}}$ is dominated by variations in M_{dyn}/L:

$$\Delta \log(M_{\mathrm{dyn}}/L_{\mathrm{g}}) \approx -\Delta \log I_{\mathrm{e,g}}. \tag{3.5}$$

We perform these calculations separately for the quiescent and star-forming populations, and show the observed redshift evolution of M_{dyn}/L_g in Fig. 3.1. Similar to many previous FP studies of quiescent galaxies (e.g., van der Wel et al. 2005; van Dokkum & van der Marel 2007), we find that M_{dyn}/L_g decreases with redshift, and show that this is also the case for the star-forming population. We determine the slope of the redshift evolution using a linear least squares fit, weighted by the observational errors, and estimate uncertainties on the fit via bootstrap resampling. The number of SDSS galaxies is significantly larger than the LEGA-C sample size, which effectively causes the fit to be forced through the best-fit zero point of the SDSS FP ($\Delta \log(M_{\mathrm{dyn}}/L_g) = 0$). Since this omits any potential systematic errors on the SDSS data, we fit to both the combined LEGA-C and SDSS data (solid lines) and the LEGA-C data only (dashed lines).

The results are presented in Table 3.1; the two different methods agree within 2σ and 1σ for the quiescent and star-forming samples respectively. Some small systematic discrepancies between the two different estimates for each subsample may be expected, considering that there are substantial differences in the measurements of the effective radii, velocity dispersions, and photometry between the SDSS and LEGA-C data.
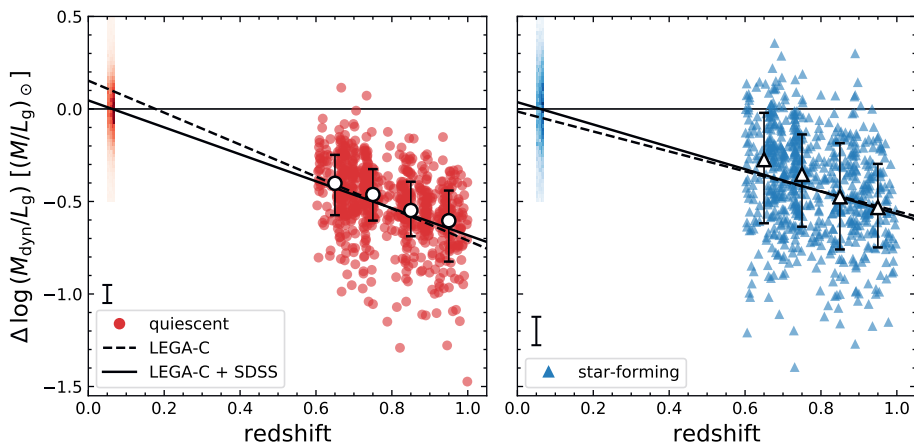
Figure 3.1: Redshift evolution of the dynamical mass-to-light ratio of quiescent (left) and star-forming (right) galaxies from the SDSS and LEGA-C samples. Linear fits to the LEGA-C data alone (dashed lines) and combined LEGA-C and SDSS sample (solid lines) show that there is a strong evolution in M_{dyn}/L_g with redshift, with the quiescent population evolving more rapidly than the star-forming population (Table 3.1).

3.3.1 Quiescent galaxies

We show a comparison with previous measurements of the redshift evolution in M_{dyn}/L of quiescent galaxies in Fig. 3.2, where coloured markers represent results obtained with the LEGA-C data and black symbols indicate different studies. Our result for the quiescent sample is consistent with the evolution of field galaxies in the rest-frame B-band measured by Treu et al. (2005) and Saglia et al. (2010, 2016), and slightly steeper than the bias-corrected measurement by van der Wel et al. (2005).

Other studies (e.g., van de Sande et al. 2014; Beifiori et al. 2017) deviate more significantly (typically $2-3\sigma$), which can largely be attributed to differences in the sample selection. Our selection generally differs from previous works in (i) the diversity of environment probed, with many studies focusing on galaxy clusters alone, or (ii) the mass range considered, as many studies have been limited to more massive galaxies.

van Dokkum & van der Marel (2007) and Saglia et al. (2010) have shown that the redshift evolution in M_{dyn}/L differs for cluster and field galaxies. If we restrict our fit to only those LEGA-C galaxies which are classified as being cluster members (Darvish et al. 2017), we also find a marginally shallower evolution of $\Delta \log(M_{\mathrm{dyn}}/L_g) \propto (-0.83 \pm 0.18)z$ as compared to the full LEGA-C sample.

Moreover, van der Wel et al. (2005) and others (e.g., Holden et al. 2010; Jørgensen & Chiboucas 2013) find evidence for a mass-dependent evolution of M_{dyn}/L, with low-mass galaxies evolving more rapidly than high-mass galaxies. We would therefore expect to find a steeper evolution for our sample (of
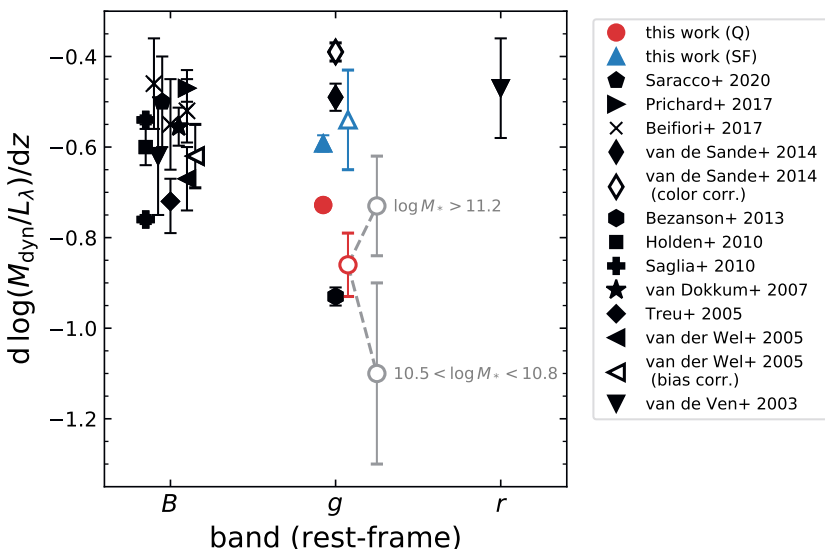
Figure 3.2: Comparison of the measured redshift evolution in M_{dyn}/L in different passbands. Red and blue markers show the results obtained in this paper for quiescent and star-forming galaxies respectively, for the LEGA-C sample (open) and combined LEGA-C and SDSS sample (solid). Black symbols show results from other studies of quiescent galaxies.

$\log(M_*/M_\odot) > 10.5)$ as compared with previous studies that typically select galaxies of $\log(M_*/M_\odot) \gtrsim 11$. We indeed find a mass dependence within our sample: if we fit only LEGA-C galaxies in the mass range $10.5 < \log(M_*/M_\odot) < 10.8$ or $\log(M_*/M_\odot) > 11.2$, we find $\Delta \log(M_{\mathrm{dyn}}/L_g) \propto (-1.1\pm0.2)z$ and $\Delta \log(M_{\mathrm{dyn}}/L_g) \propto (-0.73 \pm 0.11)z$ respectively.

Lastly, we note that the above measurements neglect the role of progenitor bias (van Dokkum et al. 2001): less massive galaxies tend to assemble and quench later than high-mass galaxies, such that galaxies of a fixed stellar mass at $z \sim 0$ will be younger than those at $z \sim 0.8$, and therefore also have a lower M_{dyn}/L_g. However, a full treatment of this effect on the FP is beyond the scope of this work.

3.3.2   Star-forming galaxies

The evolution of the star-forming population is significantly shallower than that of the quiescent population. Although the specific star formation rate (sSFR) decreases sharply toward $z \sim 0$ (Madau & Dickinson 2014), and M_{dyn}/L_g thus strongly increases, any low level of star formation will reduce the net increase in M_{dyn}/L_g. Moreover, progenitor bias plays a significant role: while young galaxies enter the massive star-forming population toward low redshift, many of the older galaxies become quiescent. The net effect is therefore a shallower observed evolution in M_{dyn}/L_g.

Table 3.1: Best-fit evolution in M_{dyn}/L_g and M_{dyn}/M_*.

Sample	$\mathrm{d}\log(M_{\mathrm{dyn}}/L_g)/\mathrm{d}z$	$\mathrm{d}\log(M_{\mathrm{dyn}}/M_*)/\mathrm{d}z$
$0.6 < z < 1.0$ Q	-0.86 ± 0.07	-0.05 ± 0.06
$0.6 < z < 1.0$ SF	-0.54 ± 0.11	-0.05 ± 0.08
$0.0 < z < 1.0$ Q	-0.728 ± 0.011	0.048 ± 0.009
$0.0 < z < 1.0$ SF	-0.604 ± 0.016	0.097 ± 0.011

Notes. Samples correspond to either the LEGA-C data ($0.6 < z < 1.0$) or combined SDSS and LEGA-C data ($0.0 < z < 1.0$) for the quiescent (Q) and star-forming (SF) populations.

3.4   Evolution in M_{dyn}/M_*

We obtain the mass FP by replacing the surface brightness ($I_{\mathrm{e,g}}$) by the stellar mass surface density ($\Sigma_* = M_*/(2\pi R_{\mathrm{e}}^2)$):

$$\log R_{\mathrm{e}} = \alpha \log \sigma + \beta \log \Sigma_* + \gamma, \tag{3.6}$$

where α and β describe the tilt, and γ is the zero point. Following the approach of Section 3.3, we adopt a fixed tilt of $\alpha = 1.629$ and $\beta = -0.84$ (Hyde & Bernardi 2009). We again fit the zero point of the SDSS sample (γ_0) for the star-forming and quiescent population separately, and calculate the residual of the FP in M_{dyn}/M_* for the LEGA-C galaxies:

$$\Delta \log(M_{\mathrm{dyn}}/M_*) \approx -\Delta \log \Sigma_* = (\Delta_{\mathrm{MFP}} - \gamma_0)/\beta, \tag{3.7}$$

where

$$\Delta_{\mathrm{MFP}} = \log R_{\mathrm{e}} - \alpha \log \sigma - \beta \log \Sigma_*. \tag{3.8}$$

In Fig. 3.3 we show $\Delta \log(M_{\mathrm{dyn}}/M_*)$ as a function of redshift for the star-forming (blue) and quiescent (red) LEGA-C and SDSS galaxies. As in Section 3.3, we perform a linear fit to the two populations separately, using the LEGA-C data only (dashed lines) and the combined LEGA-C and SDSS data (solid lines). The results are presented in Table 3.1.

For the quiescent galaxies, the two slopes are consistent within 1.6σ, and agree well with the lack of evolution found by Bezanson et al. (2013) for a sample of ~ 100 high-redshift quiescent galaxies. Our result demonstrates that the mass FP of the star-forming population also does not undergo a strong evolution.

Furthermore, we demonstrate that this result is not sensitive to the adopted definition of quiescence. The bottom panels of Fig. 3.3 show the dependence of $\Delta \log(M_{\mathrm{dyn}}/M_*)$ on the sSFR obtained from the SED fitting. There is only a weak correlation for both the SDSS and LEGA-C galaxies, as evidenced by linear fits to the data (black solid lines), with galaxies of high sSFR being on average slightly more baryon-dominated within $1\,R_{\mathrm{e}}$: $\mathrm{d}\log(M_{\mathrm{dyn}}/M_*)/\mathrm{d}\log(\mathrm{sSFR}) = -0.014 \pm 0.0005$ and $\mathrm{d}\log(M_{\mathrm{dyn}}/M_*)/\mathrm{d}\log(\mathrm{sSFR}) = -0.033 \pm 0.007$ for the SDSS and LEGA-C samples respectively.
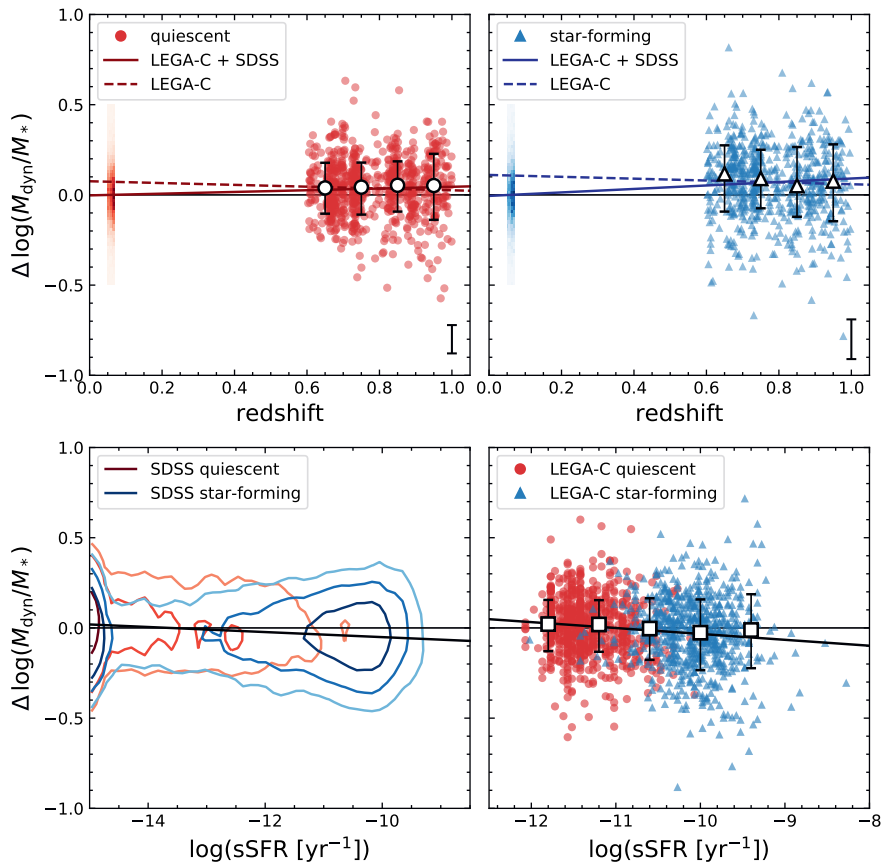
Figure 3.3: Evolution of the residuals from the mass FP of massive galaxies with redshift (top panels) and sSFR (bottom panels). Red and blue markers indicate the quiescent and star-forming LEGA-C galaxies respectively, with medians and 16th and 84th percentiles shown in black. The SDSS sample is represented by 2D histograms or contours that enclose 50%, 80% and 95% of each subsample (smoothed with a Gaussian filter of FWHM = 0.1 dex). Linear fits to the LEGA-C data (dashed lines) and combined LEGA-C and SDSS sample (solid lines) show that the redshift evolution in $\Delta \log(M_{\mathrm{dyn}}/M_*)$ is, at most, weak (see Table 3.1). Combined with the very weak correlation between $\Delta \log(M_{\mathrm{dyn}}/M_*)$ and the sSFR (solid lines; bottom panels), this implies that massive star-forming and quiescent galaxies lie on the same mass FP across $0 < z < 1$.

The LEGA-C data alone suggest that all galaxies lie on the same mass FP, irrespective of star formation activity and redshift. However, both Schechter et al. (2014) and Zahid et al. (2016) find a weak redshift evolution in the zero point of the mass FP of early-type galaxies, such that $\Delta \log(M_{\mathrm{dyn}}/M_*)$ increases slightly with redshift. We find a similar weak but significant evolution in M_{dyn}/M_* with

redshift for our combined LEGA-C and SDSS data, particularly so for the star-forming galaxies, raising the question of whether the observed evolution to $z \sim 0$ is due to structural evolution, or caused by systematic uncertainties.

In Fig. 3.1 we showed that the evolution of the luminosity FP is broadly consistent between the two data sets, suggesting that any systematic effects on the velocity dispersion, size, or luminosity are small. However, the stellar mass is an additional possible source of systematic error. Although we have mitigated potential biases between the SDSS and LEGA-C data by using the same models and software for the SED modeling of all galaxies, we caution that some differences remain, particularly in the photometry used. For instance, the aperture sizes differ systematically, the SED is sampled differently in wavelength space, and there may be systematic uncertainties in the calibration of the photometry. Overall this can lead to a systematic uncertainty of at least 0.05 dex between the SDSS and LEGA-C mass estimates: for example, we find lower stellar masses for our SDSS sample if we use the MPA-JHU catalog (Brinchmann et al. 2004), with a median offset of -0.05 dex compared to the masses from Chang et al. (2015). This would shift the SDSS data upward in Fig. 3.3, in closer agreement with the LEGA-C data. We therefore conclude that the observed weak evolution in the mean value of $\Delta \log(M_{\mathrm{dyn}}/M_*)$ is likely not significant, and caution against interpreting this as evidence for, e.g., evolution in the dark matter fraction or the initial mass function.

Systematics can also explain the discrepancy between our results and those by Bezanson et al. (2015), who found that the mass FP changes by $\Delta \log(M_{\mathrm{dyn}}/M_*) \sim 0.2 - 0.3$ dex between $0 \lesssim z \lesssim 0.7$. However, the SED modeling differs significantly for their low-redshift and high-redshift data, resulting in a systematic offset: when using the same methods, i.e. stellar masses from the MPA-JHU catalog for the SDSS and masses estimated with FAST (Kriek et al. 2009) for LEGA-C, we also find that $\mathrm{d} \log(M_{\mathrm{dyn}}/M_*)/\mathrm{d}z \approx 0.3$ dex.

Finally, we emphasize that although the residual from the FP in M_{dyn}/M_* is approximately constant across $0 < z < 1$ (Fig. 3.3), there is significant and systematic variation in M_{dyn}/M_* within the galaxy population itself. Fig. 3.4 shows a near face-on projection of the mass FP color-coded by the mean value of $\log(\widetilde{M}_{\mathrm{dyn}}/M_*)$ in bins of $\log R_{\mathrm{e}}$ and $\log \sigma$, where $\widetilde{M}_{\mathrm{dyn}}$ is calculated following Cappellari et al. (2006):

$$\widetilde{M}_{\mathrm{dyn}} = \frac{\beta(n) R_{\mathrm{e}} \sigma^2}{G}, \tag{3.9}$$

with $\beta(n) = 8.87 - 0.831n + 0.0241n^2$, where n is the Sérsic index and G the gravitational constant. While the zero point of the mass FP itself remains constant, individual galaxies may change in size and velocity dispersion with time, thus moving along the FP, and vary in M_{dyn}/M_*.

3.5   Discussion and Conclusions

In this Letter, we have measured the redshift evolution of the luminosity and mass FP of massive ($\log(M_*/M_\odot) \geq 10.5$) galaxies out to $z \sim 1$. Whereas previous studies suffered from significant selection bias, our sample of 1419 galaxies from the
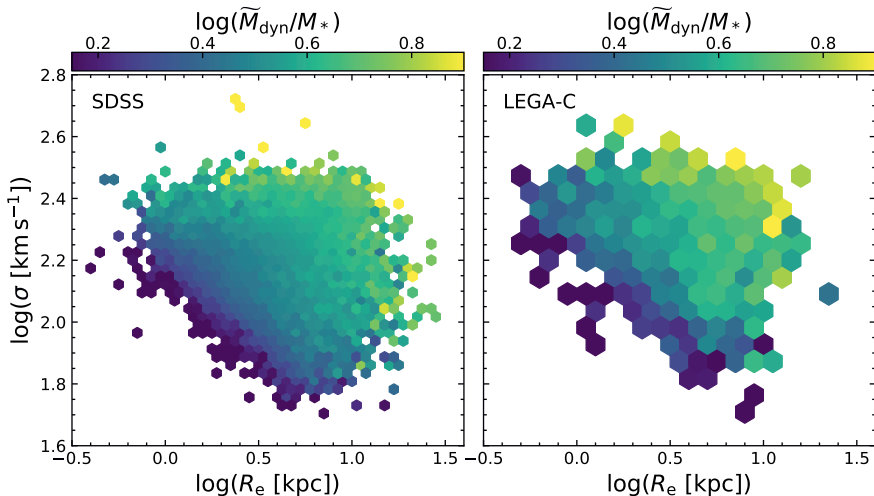
Figure 3.4: Variation in M_{dyn}/M_* along the mass FP. Panels show a near face-on projection of the mass FP, color-coded by the mean value of $\log(\widetilde{M}_{\mathrm{dyn}}/M_*)$ in each bin. Although Fig. 3.3 shows no evolution in the mean value of $\Delta\log(M_{\mathrm{dyn}}/M_*)$ with either redshift or sSFR, it is possible for individual galaxies grow with time, and thus undergo a change in M_{dyn}/M_*.

LEGA-C survey is highly homogeneous and representative of the massive galaxy population at $z \sim 0.8$ (van der Wel et al. 2016; de Graaff et al. 2021). We find that the star-forming and quiescent populations follow a steep evolution in M_{dyn}/L_g, yet, their evolution in M_{dyn}/M_* is remarkably weak: all massive galaxies lie on the same mass FP across $0 < z < 1$.

The stability of the mass FP implies that the evolution in the luminosity FP, and thus in M_{dyn}/L, is due to a combination of progenitor bias and evolution in the stellar populations alone: $\Delta\log(M_{\mathrm{dyn}}/L) = \Delta\log(M_*/L)$.

There is some room for evolution of the mass FP with redshift, however, if we assume that the weak evolution in Fig. 3.3 is physical, and not caused by systematic uncertainties. In this case, the weak dependence of the residuals from the FP on the sSFR and the different values of $\mathrm{d}\log(M_{\mathrm{dyn}}/M_*)/\mathrm{d}z$ for the star-forming and quiescent populations reflect structural differences.

In contrast, theoretical predictions (e.g., Hilz et al. 2013) and observations (e.g., van de Sande et al. 2013; Wuyts et al. 2016; Genzel et al. 2017) show that—within the effective radius—galaxies become more baryon-dominated at high redshift, whereas the best-fit evolution of our combined LEGA-C and SDSS data suggests the opposite. We emphasise that systematic observational uncertainties likely contribute to the observed offset between the SDSS and data at higher redshift. Moreover, we note that we have not accounted for baryonic mass in the form of gas, which may become increasingly important toward high redshift. We have also not included the effect of colour gradients, which may lead to an underestima-

tion of M_{dyn}/M_*, since mass-weighted sizes can be substantially smaller than the luminosity-weighted sizes used here (e.g., Szomoru et al. 2013; Chan et al. 2016).

The lack of evolution of the mass FP implies that the coupling of morphological and dynamical properties extends over a wide range in time, imposing strong constraints on the possible evolutionary pathways of galaxies. Quiescent galaxies for example, which have been shown to undergo significant size growth between $0 < z < 1$ (van der Wel et al. 2014a), must evolve dynamically such as to remain on the mass FP (Fig. 3.4).

Moreover, we find that the star-forming progenitors lie on the same scaling relation as their massive, quiescent descendants at low redshift. The mass FP therefore offers a tool to study the structural and kinematic evolution of galaxies with minimal impact from progenitor bias, by statistically tracking their trajectories along the plane.

Whether the mass FP can be used in a similar fashion at $z > 1$ or at lower mass, will require a larger number of stellar kinematic measurements at high redshift. Future studies will help to understand how galaxies settle onto the scaling relation, and whether galaxies become more baryon-dominated at high redshift.

Acknowledgements

4 | OBSERVED STRUCTURAL PARAMETERS OF EAGLE GALAXIES: RECONCILING THE MASS-SIZE RELATION IN SIMULATIONS WITH LOCAL OBSERVATIONS

ABSTRACT

We use mock images of $z = 0.1$ galaxies in the $100\,\mathrm{Mpc}$ EAGLE simulation to establish the differences between the sizes and morphologies inferred from the stellar mass distributions and the optical light distributions. The optical, r-band images used were constructed with a radiative transfer method to account for the effects of dust, and we measure galaxy structural parameters by fitting Sérsic models to the images with GALFIT. We find that the derived half-light radii differ systematically from the stellar half-mass radii, as the r-band sizes are typically $0.1\,\mathrm{dex}$ larger, and can deviate by as much as $\approx 0.5\,\mathrm{dex}$, depending on the dust attenuation and star formation activity, as well as the measurement method used. Consequently, we demonstrate that the r-band sizes significantly improve the agreement between the simulated and observed stellar mass-size relation: star-forming and quiescent galaxies in EAGLE are typically only slightly larger than observed (by $0.1\,\mathrm{dex}$), and the slope and scatter of the local relation are reproduced well for both populations. Finally, we compare the obtained morphologies with measurements from the GAMA survey, finding that too few EAGLE galaxies have bulge-like light profiles (Sérsic indices of $n \sim 4$). Despite the presence of a significant population of triaxial systems among the simulated galaxies, the surface brightness and stellar mass density profiles tend to be closer to exponential discs ($n \sim 1-2$). Our results highlight the need to measure the sizes and morphologies of simulated galaxies using common observational methods in order to perform a meaningful comparison with observations.

Anna de Graaff, James Trayford, Marijn Franx,
Matthieu Schaller, Joop Schaye, Arjen van der Wel
Monthly Notices of the Royal Astronomical Society, **511**, 2544, 2022

4.1 Introduction

The sizes and morphologies of galaxies are some of their most basic observable properties, and provide crucial insight into the formation of galaxies and the build-up of their stellar mass. Cosmological hydrodynamical simulations that aim to model a realistic universe are therefore expected to reproduce such fundamental characteristics. However, to determine the success of a given model requires a fair comparison between simulations and observations, as the latter can come with significant biases due to the systematic differences between the distribution of the light and the stellar mass.

Observationally, galaxy morphologies are highly diverse, but are usually grouped into two classes, of early-type (spheroidal or bulge-like) and late-type (more disc-like) systems. Importantly, these morphological types have been found to correlate with other properties: early-type galaxies are typically more massive than late-type galaxies, have significantly redder colours and lower star formation rates (e.g., Blanton et al. 2003; Kauffmann et al. 2003; Driver et al. 2006), and are often found to lie in denser environments (e.g., Dressler 1980; Gómez et al. 2003). Early-type galaxies thus appear to have followed very different evolutionary paths from late-type galaxies, although the precise mechanisms behind the quenching of star formation in galaxies and the possible link to a morphological transformation still represents an active area of research.

Furthermore, the stellar masses and sizes of both populations of galaxies have been shown to be correlated at low redshift (e.g., Shen et al. 2003; Lange et al. 2015), and this relation has been observed to exist at least up to $z \sim 3$ (e.g., Trujillo et al. 2006; van der Wel et al. 2014a; Mowla et al. 2019). The sizes of late-type galaxies can be linked back to the dependence of the halo angular momentum on halo mass (Mo et al. 1998). To zeroth order, the galaxy size reflects the size of the halo, but it further depends on the details of more complex processes, such as stellar feedback (e.g., Sales et al. 2010; Brook et al. 2011; DeFelippis et al. 2017), or the formation of a central bulge component through mergers or gravitational instabilities (e.g., Hernquist 1989; Dekel & Burkert 2014; Zolotov et al. 2015). For early-type galaxies, the mass-size relation is much steeper and evolves faster than is the case for the late-type population, suggesting a different formation history. Dry mergers are thought to play a significant role (Naab et al. 2009; Bezanson et al. 2009), and Shen et al. (2003) demonstrated that a simple model in which galaxies undergo repeated minor mergers, can describe both the slope and scatter of the observed mass-size relation of quiescent galaxies at $z \sim 0$ well.

As the stellar mass-size relation reflects fundamental processes in the formation and evolution of galaxies, it provides a key measure of success for theoretical models, and cosmological hydrodynamical simulations in particular. The latest generation of cosmological simulations all approximately reproduce the observed mass-size relations, e.g., the EAGLE simulations (Schaye et al. 2015; Furlong et al. 2017), Illustris-TNG (Genel et al. 2018), or SIMBA (Davé et al. 2019). Moreover, these simulations are able to form a diverse set of morphologies, as both star-forming discs and quiescent spheroids are formed (e.g., Snyder et al. 2015; Correa et al. 2017; Thob et al. 2019).

However, many of these studies are based on a comparison between the stellar mass distributions of simulated galaxies, and the optical light observed in photometric galaxy surveys. Additionally, there are often differences in the measurement techniques used: galaxy sizes in simulations are typically measured using a curve of growth method, whereas observational studies tend to fit parametric models to estimate galaxy sizes.

To mitigate possible biases introduced in these comparison studies, much effort has gone into the post processing of simulations to produce realistic mock observations. At the core, these mock data all consist of optical images, which are created by modelling the spectral energy distributions (SEDs) of the stellar particles to estimate the total light emitted within a specified wavelength range. Further possible layers of complexity are the addition of a sky background and photon noise, and modelling of the effects of dust. Even without the inclusion of dust attenuation, these mock images have demonstrated the importance of colour gradients: sizes measured from simulated, optical images are generally larger than the corresponding stellar mass sizes (van de Sande et al. 2019), which is in line with observational findings (e.g., Szomoru et al. 2013; Mosleh et al. 2017; Suess et al. 2019). The mass-size relation is therefore also changed, and simulated galaxies are found to be larger than observed (Snyder et al. 2015; Bottrell et al. 2017b; van de Sande et al. 2019), although the galaxy populations in Illustris-TNG show relatively good correspondence with observations (Genel et al. 2018; Lin et al. 2021).

Most of the aforementioned studies, however, do not measure galaxy size in the same manner as observational studies, or do not model the effects of dust in their mock images. More progress on the latter front has been made in studies that measure galaxy morphologies from mock images created with radiative transfer codes, which model the dust absorption and scattering of light between the point of emission and an observer (e.g., SUNRISE, SKIRT, or POWDERDAY Jonsson 2006; Baes et al. 2011; Camps & Baes 2015; Narayanan et al. 2021). With these more realistic images, Rodriguez-Gomez et al. (2019) (Illustris-TNG) and Bignone et al. (2020) (EAGLE) found that galaxy morphologies at $z \sim 0$, as quantified by non-parametric methods (for a review, see Conselice 2014), agree well between simulations and observations.

To also make the measurement of galaxy sizes consistent with observations, requires fitting the mock surface brightness profiles with Sérsic models (Sersic 1968). These models are highly instructive, as they simultaneously measure the overall scale (size, luminosity) and morphology of a galaxy (quantified by the Sérsic index and the projected axis ratio). On the other hand, the modelling of Sérsic profiles is strongly dependent on the estimation and treatment of the sky background and noise within an image, and dedicated software for the robust extraction of structural parameters has therefore been developed (e.g., GALFIT, GIM2D; Peng et al. 2002; Simard et al. 2002). Using such software, Price et al. (2017) demonstrated the importance of the measurement method used on the inferred size, as the sizes of high-redshift galaxies in the MassiveFIRE simulations differ significantly between measurements with GALFIT and aperture-based methods. With a custom fitting method, Rodriguez-Gomez et al. (2019) found that the $z \sim 0$ mass-size

relation in Illustris-TNG depends only weakly on the method used to measure the half-light radius, but the Sérsic profile sizes of the simulated galaxies appear to be systematically larger than equivalent measurements from the Pan-STARRS 3π Steradian Survey.

Clearly, there are many factors at play when comparing simulations and observations: the physics implemented in the simulation (and the limited fidelity thereof), the level of 'realism' of the forward modelled mock data, and consistency in the analysis methods used. In this work, we aim to perform a consistent comparison between the structural properties of galaxies in the EAGLE simulation and galaxies from the Galaxy And Mass Assembly survey (GAMA; Driver et al. 2011; Liske et al. 2015; Baldry et al. 2018). We measure the structural parameters of the simulated galaxies using near-identical methods to large galaxy surveys, and do so for the projected stellar mass distributions, as well as optical images that include dust attenuation (from Trayford et al. 2017). This allows us to not only perform a robust comparison with observations from GAMA, but also to distinguish between the effects of colour gradients and differences in the measurement methods used.

We first describe the EAGLE simulations and the construction of the optical images used in Section 4.2. Section 4.3 discusses the subsequent creation of realistic mock images that include instrumental effects and noise, as well as the Sérsic profile modelling and associated quality control. We compare different measures of galaxy size in Section 4.4, and demonstrate how both the adopted measurement method and colour gradients (due to stellar population gradients and dust) within galaxies affect the overall mass-size relation. The morphological properties obtained with the Sérsic profile modelling are presented in Section 4.5 and compared with observations from GAMA. Finally, we discuss the implications of our findings in Section 4.6, and summarise our key results in Section 4.7.

4.2   Data

4.2.1   EAGLE simulations

The EAGLE simulations consist of a suite of smoothed particle hydrodynamics (SPH) simulations for a range of different volumes, resolutions, and subgrid models (Schaye et al. 2015; Crain et al. 2015). Here, we use the reference model run for the largest available comoving volume of $100^3\,\mathrm{Mpc}^3$ (L100N1504), which assumes a flat ΛCDM cosmology with cosmological parameters obtained from Planck Collaboration et al. (2014): $\Omega_\mathrm{m} = 0.307$, $\Omega_\mathrm{b} = 0.0482$ and $H_0 = 67.77\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. This simulation has a mass resolution of $9.7\times10^6\,\mathrm{M_\odot}$ for the dark matter particles, and $1.81 \times 10^6\,\mathrm{M_\odot}$ for the initial mass of the gas particles. As a result, galaxies of stellar mass $M_* \gtrsim 10^{10}\,\mathrm{M_\odot}$ are typically resolved by $\gtrsim 10^4$ stellar particles at $z \sim 0$. The Plummer-equivalent gravitational softening scale is $\epsilon = 0.70\,\mathrm{proper}$ kpc at $z < 2.8$, and the gravitational force starts to get softened on scales smaller than $2.8\epsilon \approx 2\,\mathrm{kpc}$. From hereon, we will use proper lengths for all quoted distances and sizes, unless stated otherwise.

Haloes are identified in EAGLE using the friends-of-friends algorithm, and self-bound substructures within haloes are identified using the SUBFIND algorithm

(Springel et al. 2001; Dolag et al. 2009). We follow the convention of Schaye et al. (2015) and define galaxies as the collection of particles that belong to a single substructure, with the galaxy stellar mass defined as the sum of the stellar particles enclosed within a spherical aperture with radius 30 kpc centred on the potential minimum. We focus our analysis on galaxies at $z = 0.1$ (snapshot 27), for which mock optical imaging created with SKIRT is available, as described in Section 4.2.2. Given the limited spatial resolution in the simulation, we impose a lower limit on the stellar mass of $M_* = 10^{10} \, \mathrm{M_\odot}$, as Ludlow et al. (2019) showed that galaxies below this stellar mass tend to have sizes smaller than the convergence radius of the dark matter, leading to the spurious transfer of energy from dark matter to stars via 2-body scattering. Selecting all galaxies of stellar mass $M_* \geq 10^{10} \, \mathrm{M_\odot}$ from the public EAGLE database (McAlpine et al. 2016), we obtain a sample of 3624 galaxies.

4.2.2   Galaxy images

Optical images, presented in Trayford et al. (2017), were generated by post-processing the EAGLE data with SKIRT (Baes et al. 2003, 2011; Camps & Baes 2015). The principle of the radiative transfer code SKIRT is to trace monochromatic 'photon packages' from a source to a specified detector using a Monte Carlo method. In this way, unlike with the commonly adopted method of applying a dust screen, representative 3D absorption and scattering of light due to dust are accounted for, thus creating a realistic image. We provide a brief summary of these data below, and refer the reader to Camps et al. (2016) and Trayford et al. (2017) for a detailed description of the procedures involved.

The stellar particles in the snapshot, provided they lie within a 30 kpc radius around the centre of the galaxy, form the source of the photon packages. As described in Trayford et al. (2015), each particle older than $> 100 \, \mathrm{Myr}$ is treated as a single stellar population, and assigned a SED with GALAXEV (Bruzual & Charlot 2003), using the initial mass, metallicity and stellar age from the simulation snapshot and assuming a Chabrier (2003) initial mass function. The spatial distribution of the light emitted by the particle is described by a truncated Gaussian distribution, with a smoothing length dependent on the distance to the 64th nearest neighbour.

For younger stars, the additional absorption by dust in the birth clouds needs to be taken into account. Given the limited mass resolution, however, this firstly requires a resampling of the recent star formation of the stellar particles with young ages ($< 100 \, \mathrm{Myr}$), which is done in a similar fashion to Trayford et al. (2015). Sub-particles older than 10 Myr are treated as described above, whereas younger populations are instead assigned SEDs using the MAPPINGS-III code (Groves et al. 2008), which models the emission and dust absorption within HII regions. The smoothing length for these young populations is taken to be dependent on their mass and the local gas density, although the net kernel (i.e., including the position) is equivalent to that of the other stellar particles. We note that the choice of the smoothing lengths sets the level of granularity in the final images, and Bignone et al. (2020) showed that this likely affects some of the non-parametric

morphological measurements. However, in Appendix 4.A we demonstrate that the smoothing has a negligible effect on the parametric morphologies measured in this work.

Dust in the diffuse interstellar medium (ISM) is modelled based on the properties and spatial distribution of the gas particles in the galaxy. Gas particles are smoothed using the SPH smoothing lengths, and the ISM is then discretised over an adaptive grid with a minimum grid cell size of 0.11 kpc. The dust mass within the grid cells is calculated from the star-forming as well as the cold ($T < 8000\,\mathrm{K}$) gas mass, by assuming a constant dust-to-metal mass ratio (Camps et al. 2016). Dust mass in the HII regions, already implemented through the MAPPINGS-III SEDs, is also accounted for. The composition of the dust grains is taken as the model by Zubko et al. (2004), a multi-component interstellar dust model that provides a good fit to the observed extinction curve of the Milky Way, as well as the diffuse infrared emission and abundance constraints.

With the source of emission and distribution of the dust defined, the SKIRT calculations are performed on a finely sampled wavelength grid (333 wavelengths in the range $0.28 - 2.5\mu$m), resulting in an integral field data cube. Broadband imaging is constructed by convolving the cube with an instrument response function and integrating along the wavelength direction. This is done for both the observed ($z = 0.1$) and rest frame for three different projections: face-on, edge-on, and random (the projection along the z-axis of the simulation box). Images have a field of view of $60 \times 60\,\mathrm{kpc}^2$ with a pixel scale of $0.234\,\mathrm{kpc\,pix}^{-1}$, which at $z = 0.1$ corresponds to an angular resolution of $0\rlap{.}{''}123\,\mathrm{pix}^{-1}$.

4.3 Methods

4.3.1 SDSS mock images

The images generated with SKIRT provide a realistic view of the optical emission of the simulated galaxies. However, unlike real observations of galaxies, these images do not include any instrumental effects or background noise. We therefore use the randomly oriented SKIRT images as a starting point to construct mock observations, specifically, to mimic typical image data from the Sloan Digital Sky Survey (SDSS). We choose to focus on only the r-band images (in the observed frame; Doi et al. 2010), as this is the wavelength range commonly used in observational studies.

In addition to creating mock SDSS images of the optical light, we construct mock 'images' of the stellar mass distributions directly from the simulation snapshot. These stellar mass maps are designed to have similar noise properties and resolution as the optical imaging, to allow for a robust comparison between the distributions of the optical emission and stellar mass.

4.3.1.1 Optical images

The initial images are 60 kpc on a side, which in most cases is significantly larger than the half-mass radius of the galaxy. However, more massive galaxies can have

large half-mass radii ($> 10\,\mathrm{kpc}$), or contain extended star-forming discs. Although a $30\,\mathrm{kpc}$ aperture may capture all the galaxy mass, at least for systems of $M_* < 10^{11}\,\mathrm{M_\odot}$ (Schaye et al. 2015), a large spatial extent can still be problematic in the Sérsic modelling, as the extended emission may get mistaken for background flux. We therefore add empty background pixels onto the sides of the images, such that they become $60'' \times 60''$ in area ($114^2\,\mathrm{kpc}^2$).

Next, we add a uniform background and convolve the image with a Gaussian point-spread function (PSF) to match the sky background and seeing in the SDSS imaging. We calculate the median value of the r-band 'sky' and 'psfWidth' from the photometric field catalogue of the ninth data release of the SDSS (DR9; Ahn et al. 2012), to set the sky background level ($\mu_{\mathrm{sky}} = 20.9\,\mathrm{mag\,arcsec}^{-2}$) and the full width at half maximum (FWHM $= 1''\!.39$) of the PSF, respectively. We note that the real PSF in the SDSS image data has a far more complex shape than the single Gaussian profile assumed here. However, in Appendix 4.B we show that a simple PSF model is sufficient for measuring parametric morphologies. After the convolution, we resample the image from a pixel scale of $0''\!.123\,\mathrm{pix}^{-1}$ to $0''\!.396\,\mathrm{pix}^{-1}$ to match the SDSS pixel resolution.

From the same SDSS DR9 catalogue we obtain typical values for the r-band detector gain ($G = 4.73\,e^-\,\mathrm{ADU}^{-1}$), the conversion factor from counts to fluxes (nMgyPerCount $= 0.0051\,\mathrm{nmgy\,ADU}^{-1}$), and the 'dark variance' (the combination of detector readout noise and the dark current; $\sigma_{\mathrm{dark}}^2 = 1.32\,\mathrm{ADU}^2$). The dark variance is added to the image to mimic detector effects, under the assumption that these electrons follow a Poisson distribution (i.e., $\mu = \sigma^2$), although this source of noise is insignificant in comparison to the sky background level (a factor ≈ 20 lower). Lastly, we convert the image to units of $e^-\,\mathrm{pix}^{-1}$.

The image now closely resembles the collection of photoelectrons by a detector, and these photoelectrons obey Poisson statistics. We can therefore create an image with a realistic noise level: for each pixel, we draw a random sample from the Poisson distribution with mean value equal to the number of electrons in that pixel ($\mu = N_{\mathrm{e,pix}}$). We also obtain a 'sigma image', an image with the same dimensions as the galaxy image that stores $G^{-1} \times \sqrt{N_{\mathrm{e,pix}}}$, which will be used as statistical weights in the two-dimensional Sérsic modelling (Section 4.3.2). We note that the image construction with SKIRT (Section 4.2.2) also introduces Poisson noise, however, this noise is well below the typical noise level in the SDSS (Trayford et al. 2017), therefore justifying the seemingly duplicate addition of photon noise.

As a final step, we divide the image by the gain and subtract the (previously added) sky background and dark variance from the noisy image, delivering the final mock SDSS image. Fig. 4.1 shows an example of an initial r-band image created with SKIRT, and the corresponding mock SDSS image (converted to physical flux units) that includes realistic noise and PSF smoothing.

Unlike the real SDSS data, these mock images do not contain any foreground or background sources, as only light from within a $30\,\mathrm{kpc}$ aperture is included. We have chosen to not implement this additional complexity, as Bottrell et al. (2017a) showed that the effect of crowding on the measurement of structural parameters is generally small, with the exception of very low surface brightness systems that are few in number.
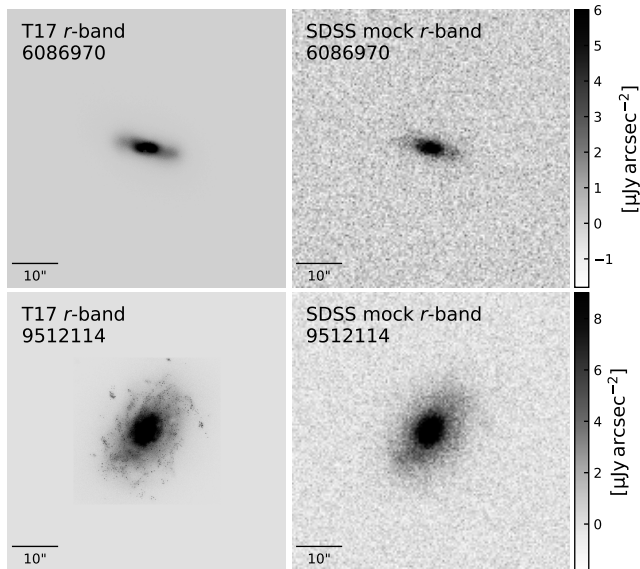
Figure 4.1: Examples of the r-band images constructed with SKIRT of galaxies at $z = 0.1$ (left; Trayford et al. 2017), and the corresponding mock SDSS images that include realistic instrumental and sky effects (right).

4.3.1.2   Stellar mass images

To construct images of the stellar mass distribution that match the noise and image resolution properties of the r-band images, we follow a similar methodology to the previous section, with few modifications. Rather than starting from the SKIRT data, we begin from the EAGLE particle data and select a box of size 114^3 kpc^3 centred around the potential minimum of the galaxy. Within this box, we select only the stellar particles that are identified as being part of the galaxy by the SUBFIND algorithm. In this way, analogous to the SKIRT images, neighbour galaxies are not included in the images. The current stellar mass of these particles is then projected in the $x - y$ plane of the simulation box to obtain an image of 512×512 pixels, which is the same orientation and of similar spatial resolution as the SKIRT data.

To be able to add realistic noise as described in Section 4.3.1.1, an effective mass-to-light ratio (Υ_{eff}) is required that describes the typical scaling between the r-band and stellar mass imaging. To obtain Υ_{eff}, we first compute the ratio (Υ) between the stellar mass of the galaxy (within the spherical aperture of radius 30 kpc), and the observed flux within a circular aperture of 30 kpc in the noise-free optical images. We then use the median of this distribution, $\Upsilon_{\text{eff}} = 10^{11}\,\text{M}_\odot\,\text{mJy}^{-1}$, to convert the stellar mass images to an effective flux and hence to a number of photoelectrons. We use a fixed value of Υ_{eff} for all galaxies, as the variation in Υ is relatively small: the standard deviation of 0.13 dex in Υ corresponds to variations in the image noise level of $\sim 15\%$, which we have found the Sérsic profile fitting
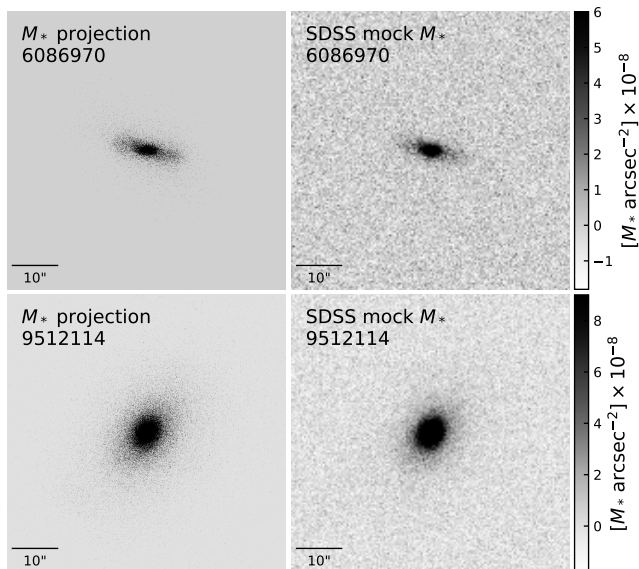
Figure 4.2: Examples of the stellar mass maps of galaxies at $z = 0.1$ created by projecting the stellar particles along the z-axis of the simulation box (left), and the corresponding mock SDSS 'stellar mass images' that include realistic instrumental and sky effects (right).

procedure (Section 4.3.2.2) to not be sensitive to.

As in Section 4.3.1.1, we add a uniform sky background level and smooth the image with a Gaussian PSF, using the same μ_{sky} and PSF FWHM as before. The image is then resampled to a pixel scale of $0\rlap{.}''396\,\mathrm{pix}^{-1}$, and the dark variance is added. We apply a Poisson noise model, and subtract the total (sky + dark variance) background to produce our final mock SDSS image of the stellar mass distribution. Fig. 4.2 shows an example of the initial $x - y$ projection of the stellar particles, and the mock SDSS image.

4.3.2   Sérsic modelling

We model the light and stellar mass profiles of the simulated galaxies by fitting a two-dimensional, parametric model to the mock imaging. This model, a single Sérsic profile (Sersic 1968), is described by five parameters: the total AB magnitude (m) or stellar mass ($M_{*,\mathrm{Sérsic}}$), the Sérsic index (n), the half-light or half-mass semi-major axis ($r_{\mathrm{e,maj}}$), the ratio of the semi-major and semi-minor axes (q), and the position angle (ϕ).

We describe our fitting procedure in detail in the following sections. In summary, we use a combination of SExtractor (Bertin & Arnouts 1996) and Galfit (Peng et al. 2010) to estimate the initial values of the Sérsic parameters and to find the best fitting parameter values, respectively. Both softwares are commonly used in observational studies that measure structural parameters of galaxies (e.g.,

Barden et al. 2012; Kelvin et al. 2012; van der Wel et al. 2012; Meert et al. 2015),
which enables us to perform a consistent comparison between simulated and ob-
servational results.

4.3.2.1   Initial parameter estimation

As the Sérsic model is described by five parameters, increased to seven free pa-
rameters by the addition of the galaxy centroid position, there is a vast parameter
space to be explored to find their optimal values. It is therefore crucial to provide
reasonable initial estimates of the Sérsic parameters to reduce the computational
cost, and avoid the fit to converge to a local, rather than global, minimum.

We use SExtractor (Bertin & Arnouts 1996) to detect the source(s) present
in each image and extract their photometric properties. Unlike real observations,
the mock images include only mass and light from the vicinity of the galaxy po-
tential minimum, and, in the majority of cases, there is thus only one source to
be found by SExtractor. However, merging systems or small satellites of larger
satellite galaxies may have been identified as a single galaxy by the subfind al-
gorithm, but show two (or more) spatially distinct components in the imaging.
We therefore run SExtractor with a setup akin to the 'cold mode' employed
by the GALAPAGOS code (for details, see Rix et al. 2004; Barden et al. 2012),
which was optimised to detect and deblend flux from bright sources. Specifically,
we use a relatively high detection threshold, requiring 3σ detections over 15 ad-
jacent pixels after smoothing with the default convolution kernel. To deblend the
detected object(s), we use a number of 64 subthresholds (the levels between the
detection threshold and maximum count value; DEBLEND_NTHRESH = 64)
and a minimum contrast of DEBLEND_MINCONT = 0.0001. For each image,
the corresponding sigma image (Section 4.3.1.1) is used to provide the algorithm
with the root mean square (RMS) noise level, and the background is set to a
fixed value of zero. We note that this procedure is vastly simplified in comparison
with observational data, due to the fact that our images contain just one or few
bright objects, and we have a perfect background subtraction and noise model.
Our SExtractor results are therefore only weakly sensitive to changes in the
parameters in the configuration file.

The output catalogue of SExtractor contains the centroid position ('X_IMAGE',
'Y_IMAGE'), total flux ('FLUX_AUTO'), half-light radius ('FLUX_RADIUS'),
ellipticity ($e \equiv 1-q$; 'ELONGATION') and position angle ($\theta \equiv \phi+90°$; 'THETA_IMAGE')
of each extracted source. We use these to set the initial values for the position,
m or $M_{*,\mathrm{Sersic}}$, q, and ϕ of the Sérsic model, respectively. For the initial value
of $r_{\mathrm{e,maj}}$, we follow the approach by Kelvin et al. (2012) and correct the circu-
larised radius from SExtractor to a major axis size, and account for the PSF
convolution:

$$r_{\mathrm{e,maj}} = \sqrt{\frac{r_{\mathrm{e,circ}}^2}{q} - 0.32\,\Gamma^2}\,, \qquad (4.1)$$

where Γ is the FWHM of the PSF. This leaves just one parameter, the Sérsic
index, which we set to an initial value of $n = 4$.
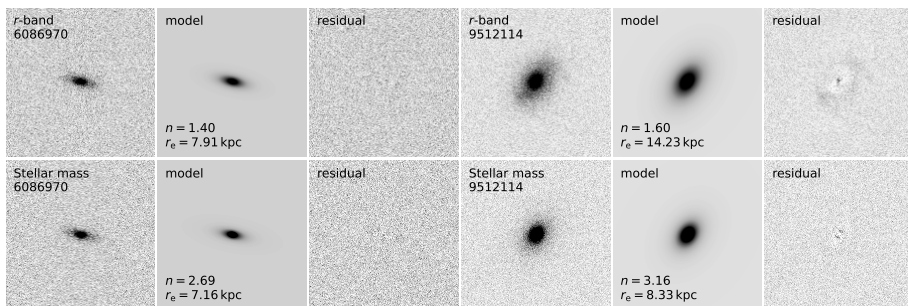
Figure 4.3: Results of the Sérsic profile modelling with GALFIT for the galaxies presented in Figs. 4.1 and 4.2. The top row shows the r-band image (left), best-fit Sérsic model (middle), and residual (right) of the two galaxies, respectively. The bottom row shows the corresponding results for the stellar mass images.

4.3.2.2   Sérsic profile fitting

We perform the Sérsic modelling with GALFIT (Peng et al. 2010), which uses the Levenberg-Marquardt algorithm to find the parameter values for which the total χ^2 value of the image is minimised. To do so, the mock image, sigma (RMS) image and PSF (from Sections 4.3.1.1 and 4.3.1.2) are provided as an input. We allow for multiple Sérsic profiles in the configuration file, such that satellite galaxies (if present) are fit simultaneously with the primary galaxy, and initial parameters for each profile are set as described in the previous section. The sky background is fixed to a value of zero, although we investigate the effect of allowing for a variable sky component in the section 4.3.2.5.

As a first pass, we do not place any constraints on the parameter values, to let the algorithm freely explore the parameter space. For most galaxies, this procedure leads to convergence with reasonable sizes and Sérsic indices. Occasionally however, the Sérsic index reaches implausible values (e.g., $n < 0.2$), and we therefore rerun the fits for these objects with an additional constraint of $0.2 < n < 8.0$ for the primary component only, which can lead to convergence within this range. In Fig. 4.3, we demonstrate the Sérsic modelling for the galaxies in shown in Figs. 4.1 and 4.2.

4.3.2.3   Flags

We assess the quality of the fits by three criteria, which translate into a single combined flag: any fit that has converged at the boundary of the allowed range in n is assigned a flag value of 1; a value of 2 is added to indicate images in which multiple components are fit simultaneously; a value of 4 is added to objects with bad fits, and is assigned on the basis of a visual inspection of the fits and residual images. This latter category consists of a mixture of objects, such as ongoing mergers that are simply not well described by Sérsic profiles, brightest cluster galaxies that have highly complex morphologies, or simply failed fits that are unrealistically large in size. In few cases (27), we find that the SEXTRACTOR-detected sources

Table 4.1: Best-fit r-band structural parameters and uncertainties from GALFIT. This table is available in its entirety online.

GalaxyID	m mag	r_{e} kpc	n	q	ϕ deg	flag
2	17.24 ± 0.00	2.77 ± 0.02	1.02 ± 0.02	0.44 ± 0.00	33.68 ± 0.35	0
13632	17.13 ± 0.01	3.02 ± 0.05	2.45 ± 0.07	0.74 ± 0.01	-39.91 ± 1.20	0
21794	17.69 ± 0.01	4.10 ± 0.06	1.26 ± 0.03	0.50 ± 0.01	10.58 ± 0.65	0
23302	17.47 ± 0.01	1.47 ± 0.02	3.50 ± 0.18	0.58 ± 0.01	55.00 ± 1.03	0
24478	18.19 ± 0.01	4.78 ± 0.10	1.01 ± 0.04	0.43 ± 0.01	62.11 ± 0.75	0

Table 4.2: Best-fit stellar mass structural parameters and uncertainties from GALFIT. This table is available in its entirety online.

GalaxyID	$\log(M_*/\mathrm{M}_\odot)$	r_{e} kpc	n	q	ϕ deg	flag
2	10.837 ± 0.000	2.61 ± 0.02	0.88 ± 0.02	0.42 ± 0.00	34.02 ± 0.24	0
13632	10.848 ± 0.004	2.63 ± 0.02	2.60 ± 0.06	0.77 ± 0.01	-40.93 ± 1.07	0
21794	10.592 ± 0.004	3.96 ± 0.05	1.22 ± 0.03	0.49 ± 0.00	9.91 ± 0.49	0
23302	10.647 ± 0.000	1.33 ± 0.02	2.72 ± 0.12	0.58 ± 0.01	52.99 ± 0.84	0
24478	10.383 ± 0.004	4.72 ± 0.07	0.92 ± 0.03	0.41 ± 0.01	63.50 ± 0.54	0

are overdeblended, due to strong dust lanes or star-forming clumps in the disc being detected as separate objects. This only affects the optical images, and for these few galaxies we redo the GALFIT fitting with a single component.

We provide the final catalogues of the best-fit Sérsic model parameters and flag values in Tables 4.1 and 4.2. For the results presented in the following sections, we filter out all galaxies that contain a flag value of 1 (18 r-band fits, 33 stellar mass fits), as these measurements of the Sérsic index and size are not robust. Galaxies with bad fits are also removed from the sample. Clearly, the definition of a 'bad' fit is subjective, however, $< 1\%$ of galaxies fall in this category (25 r-band fits, 10 stellar mass fits), and the population statistics are therefore likely unaffected (even at $M_* \gtrsim 10^{11.2}\,\mathrm{M}_\odot$, $< 20\%$ of galaxies are excluded). With these quality criteria applied, 3560 galaxies remain with good fits in both the r-band and stellar mass imaging.

4.3.2.4 Measurement uncertainties

Although GALFIT provides an estimated uncertainty on the measured structural parameters (limited to two decimal places), these tend to underestimate the true uncertainties (for discussion, see van der Wel et al. 2012). To obtain an estimate of the typical uncertainty on the different parameters, we create a second random noise realisation of the mock images, and repeat the Sérsic profile fitting for this set of images. By comparing the differences in the structural parameters between the two runs, we find that the scatter in $r_{\mathrm{e,maj}}$ corresponds to a typical uncertainty

of $\delta \log(r_{\mathrm{e,maj}}) = 0.03$ dex. Similarly, for the total magnitude and stellar mass $\delta \log(m_{\mathrm{Sérsic}}) = 0.05$ mag and $\delta \log(M_{*,\mathrm{Sérsic}}) = 0.02$ dex, respectively, and the Sérsic index is the hardest to constrain precisely, with $\delta \log(n) = 0.04$ dex. These values are broadly consistent with the uncertainties found by van der Wel et al. (2014a), given that the typical signal-to-noise ratio of our images S/N ≈ 100 (where the S/N is calculated using the pixels belonging to the galaxy as identified by SExtractor). Galaxies for which the Galfit estimates of the uncertainties are smaller than these values, are assigned the above typical values where needed.

4.3.2.5   Sky background estimation

As noted in Section 4.3.2.1, our images have perfect background subtraction by construction, which enables us to set the background to a fixed value of zero. However, obtaining an accurate background is often a challenge in observational studies, and commonly used tools such as SExtractor have been found to overestimate the sky background (Häussler et al. 2007).

The source extraction, and in particular the Sérsic modelling, is highly sensitive to the estimation of the background. To test whether the comparison we wish to make between our Sérsic fits and those from observational data is affected by sky background uncertainties, we rerun both SExtractor and Galfit on the mock imaging with a variable sky component. We note that this test does not capture all the complexities faced in observational studies, where the background usually varies spatially across the image, but serves as a test for any systematic effects from including a nuisance parameter in the Sérsic profile modelling.

From SExtractor, we obtain an initial estimate of the sky background in the image. We then create a sky component in the configuration file for Galfit, to be fitted simultaneously with the Sérsic profile(s). An accurate fit of the background by Galfit requires a sufficiently large area of background pixels in comparison with the area spanned by the galaxy itself. With an image size of $60''$ on a side, this is the case for the majority of the sample. For very large galaxies this area is insufficient, causing the background to be overestimated due to confusion between the sky background and low surface brightness emission from the object itself.

Fig. 4.4 shows the distribution of the sky background as determined by Galfit, for both the r-band and stellar mass fits. Both distributions peak at a value of zero and show small scatter: the median of 0.01 ADU (r-band) or 0.03 ADU (stellar mass) and standard deviation of 0.2 ADU are well below the typical galaxy flux per pixel of order $\sim 10^2$ ADU. For the stellar mass images, the asymmetric tail towards positive values of the sky background can be explained by the aforementioned effect of fitting the background in relatively small images. This effect is not present in the r-band images, as these images consist largely of empty background pixels (see Section 4.3.1.1). Most importantly, we find no systematic difference in the derived structural parameters between the fits with and without a variable sky component. The additional uncertainties on the structural parameters introduced by the variable sky component are also insignificant in comparison with the random uncertainties described in the previous section (4.3.2.4): $\delta \log(r_{\mathrm{e,maj}}) = 0.006$ dex, $\delta \log(m_{\mathrm{Sérsic}}) = 0.013$ mag or $\delta \log(M_{*,\mathrm{Sérsic}}) = 0.005$ dex, and $\delta \log(n) = 0.009$ dex.
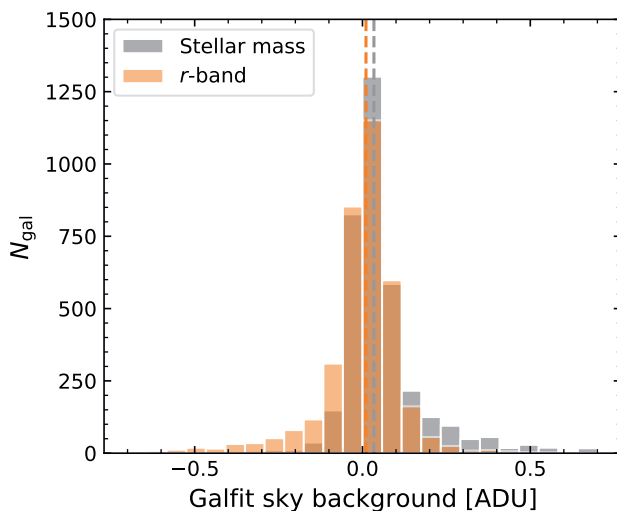
Figure 4.4: Distribution of the best-fit sky background level from GALFIT. The median of the distribution (dashed line) is close to zero for both the r-band (orange) and stellar mass (grey) images with little scatter ($0.2\,$ADU; where the unit ADU is related to the number of photoelectrons as described in Section 4.3.1.1), which is negligible in comparison to the typical galaxy flux per pixel of order $\sim 10 - 10^2\,$ADU. We therefore also find no systematic differences between the best-fit Sérsic models from the fits with a variable and fixed sky background.

4.4 Galaxy sizes

In this section we present the sizes measured with the Sérsic modelling, and evaluate how the estimated half-mass and half-light radii differ from commonly used measures of size from the public EAGLE catalogues (McAlpine et al. 2016). In addition to the different measurement methods, we examine the effects of gradients in the stellar population properties and dust attenuation on the observed size. We then assess the impact of different size (and stellar mass) estimates on the obtained stellar mass-size relation, and compare with the observed mass-size relation at $z \sim 0$.

4.4.1 Do simulated galaxies follow Sérsic profiles?

However, before we make these different comparisons, we begin by asking whether the Sérsic profile provides a good model for the surface brightness and density profiles of simulated galaxies. The simulation has finite resolution, set by both the mass of the particles and the gravitational softening scale of $\approx 2\,$kpc. Although Schaller et al. (2015) showed that the (3D) density profiles of the stellar and dark matter mass are on average well converged on scales $\gtrsim 2\,$kpc, for most galaxies the inner few kpc of the density profile will drive the fit of the Sérsic profile, as this is

where the majority of the high S/N flux is concentrated in the image.

To gauge whether the finite resolution leads to systematic deviations from the Sérsic model, we compare the Sérsic profiles with the azimuthally-averaged profiles from the mock images. We first extract the surface brightness (μ_r) and stellar mass surface density (Σ_*) profiles from the mock images and the best-fit models, by measuring the flux in elliptical apertures with the axis ratio and position angle from the best-fit Sérsic model. As we may expect the resolution to have a different effect on the profiles depending on the galaxy size itself, the sample is divided in three bins according to the half-light or half-mass radius.

In the upper panels of Fig. 4.5, we show the median, 16th and 84th percentiles of the observed μ_r profiles as a function of the half-light radius. For each size bin, the profiles are normalised to the median magnitude within the bin, and the scatter therefore represents a difference in the profile shape only. Underneath, we show the median, 16th and 84th percentiles of the residual profiles, which are calculated as the difference between the normalised profiles and best-fit models. For the largest size bin, the profiles are cut off at 30 kpc, because of the limited spatial extent of the SKIRT images (see Section 4.3.1.1). Similarly, the lower set of panels show the Σ_* profiles as a function of the half-mass radius, normalised to the median stellar mass in each bin, as well as the residual profiles.

The simulated galaxies are generally well described by the Sérsic models, as there are only some minor systematic features visible: in the left-hand panels (i.e., the smallest sizes), there is positive residual flux at $r \approx 2$ kpc ($r \approx 0.6\,r_{\rm e,maj}$), whereas the region around $r \approx 5$ kpc ($r \approx 1.7\,r_{\rm e,maj}$) is oversubtracted. At large radii, low surface brightness emission is also not fully captured by the single Sérsic profile. In the larger size bins, similar residual features appear around the same absolute radii (and thus at a smaller number of effective radii), suggesting that the limited resolution of the simulations has a small, systematic effect on the profiles.

Furthermore, we find that the Sérsic models perform well when comparing the integrated luminosity and stellar mass with the input data. Fig. 4.6 shows the difference between the total r-band magnitude ($m_{\rm Sérsic}$) and the magnitude measured within 30 kpc in the SKIRT image (i.e., not including noise or instrumental effects). Similarly, the right-hand panel shows the difference in the total stellar mass ($M_{*,\rm Sérsic}$) and the stellar mass within a 30 kpc aperture ($M_{*,30}$) as a function of $M_{*,30}$. Typically, 98% of the luminosity or stellar mass is recovered in the Sérsic fit. At high mass and high luminosity there is an increasingly stronger deviation, demonstrating that the 30 kpc aperture does not capture the full extent of the galaxy (as noted previously by Schaye et al. 2015). In Appendix 4.C, we also present the difference between $M_{*,\rm Sérsic}$ and the stellar mass obtained for other aperture sizes (50 kpc, 70 kpc, 100 kpc, and the entire subhalo mass), finding that $M_{*,\rm Sérsic}$ is approximately equivalent to the stellar mass enclosed within a spherical aperture of radius 70 kpc for very massive galaxies.

We therefore conclude that the Sérsic model provides a good description of both the stellar mass surface density profiles and the surface brightness profiles of EAGLE galaxies, and defer a further discussion of the minor systematic residuals to Section 4.6.2.
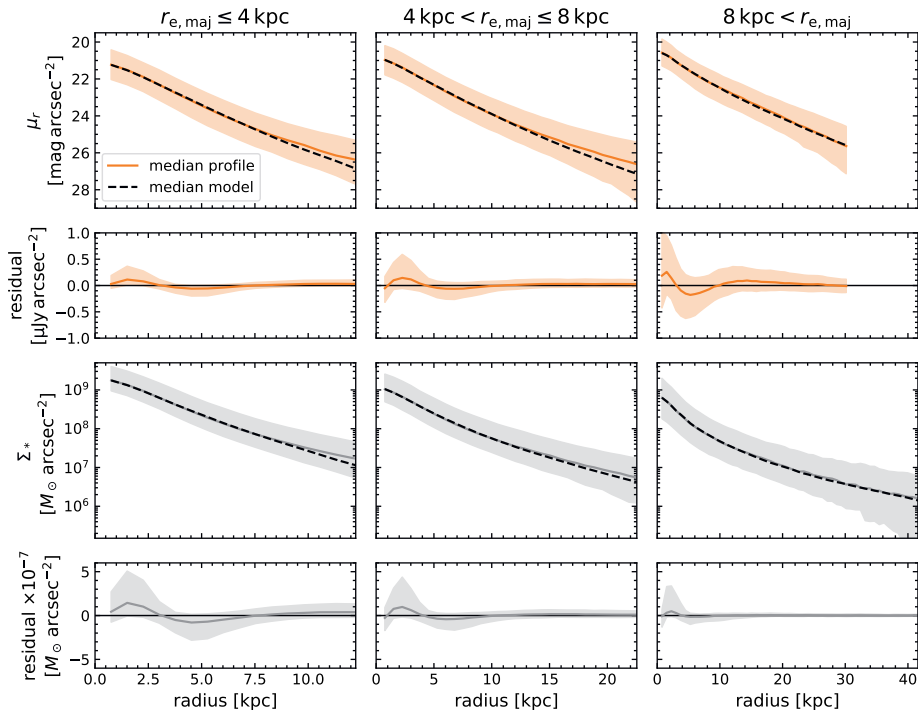
Figure 4.5: Surface brightness (orange) and stellar mass surface density (grey) profiles of EAGLE galaxies. The sample is divided into three bins of increasing half-light or half-mass radius, and the profiles are normalised to the median magnitude or stellar mass within each bin. Coloured lines and shaded areas in the first and third rows show the 16th, 50th, and 84th percentiles of the normalised profiles measured from the mock imaging, and dashed lines indicate the median of the normalised, best-fit Sérsic models in each panel. Surface brightness profiles are cut off at $r = 30$ kpc, corresponding to the size of the SKIRT images. The median, 16th and 84th percentiles of the residuals, calculated as the difference between the normalised profiles and models, are shown in the second and fourth rows (in linear scale, as opposed to the logarithmic scale used for the profiles). The surface brightness and density profiles closely follow Sérsic profiles, with only minor systematic features in the residuals.

4.4.2   Comparing different measures of size

Size estimates in the EAGLE data release are based on a growth-curve method and come in two variations (see also Furlong et al. 2017). The first method computes the total stellar mass belonging to a single subhalo within a spherical aperture of radius R centred around the minimum of the potential, after which spherical apertures of increasing radius are constructed to find the radius that encloses $M_*(< R)/2$. From hereon, we will refer to this half-mass radius as the 3D $r_{\mathrm{e,R}}$. The second method also uses the total stellar mass within a spherical aperture of radius R as starting point, however, the half-mass radius is now measured from a 2D
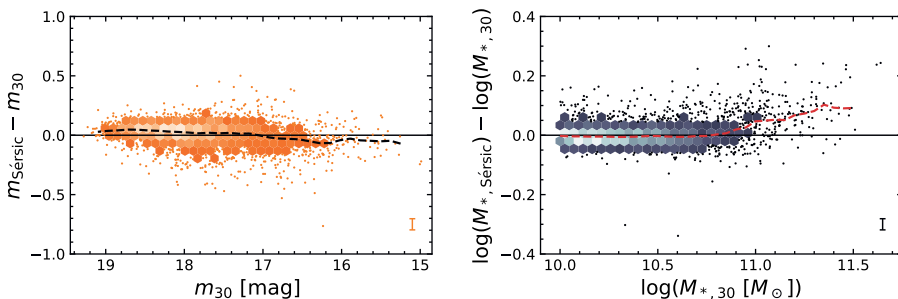
Figure 4.6: Recovered stellar mass and light with the Sérsic modelling. The left-hand panel shows the difference between the total magnitude of the Sérsic profile and the magnitude within a circular aperture of 30 kpc measured from the noise-free optical images, as a function of magnitude. The colour scale corresponds to the number density of data points, and individual galaxies are shown for sparsely populated regions (created using DENSITYPLOT; Krawczyk & Peters 2014); the dashed line shows the running median. Similarly, the right-hand panel shows the deviation between the total stellar mass of the Sérsic model and the conventional stellar mass of EAGLE galaxies (i.e., the total stellar particle mass within a spherical aperture of radius 30 kpc). On average, the models recover 98% of the mass and light within 30 kpc. This increases toward higher masses and luminosities, where the 30 kpc aperture does not capture the full extent of the galaxy.

projection of the stellar mass distribution: circular apertures of increasing radius are constructed to find the radius that encloses $M_*(< R)/2$. This computation is done for projections in three orthogonal planes, and the average of the three measurements then gives the 2D $R_{\mathrm{e,R}}$. With two different aperture sizes, $R = 30\,\mathrm{kpc}$ and $R = 100\,\mathrm{kpc}$, there are four different estimates of the half-mass radius in total.

The difference between the 3D and 2D sizes is significant, with the former being on average a factor of 4/3 larger, as is to be expected for spheroidal systems. More importantly, we find that this factor is not dependent on the galaxy mass or the sSFR. This is also apparent in Fig. 4.12 (discussed in Section 4.5.2), which shows that the median projected axis ratio is approximately constant across the six bins in stellar mass and sSFR. To compare with the Sérsic profile sizes, we can thus focus on just one of the two methods described above. In what follows, the results then translate to the other measure by a constant factor.

As the 2D $R_{\mathrm{e,R}}$ is by definition a circularised quantity, which differs systematically from the semi-major axis of the Sérsic profile by a factor $\sqrt{q}$, we choose to use the 3D $r_{\mathrm{e,R}}$ for our comparison. Fig. 4.7 shows the stellar half-mass radius from the GALFIT modelling as a function of the 3D $r_{\mathrm{e,30}}$ (left) and $r_{\mathrm{e,100}}$ (right). The bottom panels additionally show the difference between the two size estimates ($\log(r_{\mathrm{e,maj}}/r_{\mathrm{e,R}})$), together with the running median (dashed lines).

There are small, but significant, systematic differences between the 3D and Sérsic sizes. For small galaxies ($r_{\mathrm{e}} \lesssim 4\,\mathrm{kpc}$), the major axis sizes of the Sérsic fits are smaller by a constant factor of approximately $0.89\,r_{\mathrm{e,30}}$ and $0.86\,r_{\mathrm{e,100}}$ (or equivalently, a mean difference of $-0.053\,\mathrm{dex}$ and $-0.066\,\mathrm{dex}$, respectively).
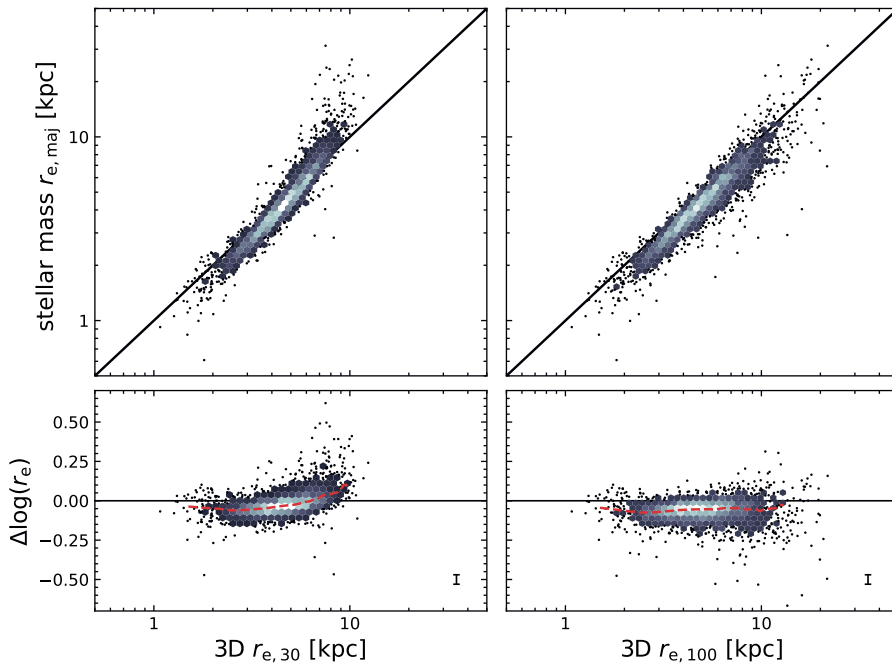
Figure 4.7: Comparison between different measures of the stellar half-mass radius. The upper panels show the semi-major axis of the best-fit Sérsic model to the stellar mass imaging versus the 3D half-mass radius within a spherical aperture of radius 30 kpc (left) and 100 kpc (right). The lower panels show the difference between the sizes as a function of radius, together with the running median (dashed lines). There is a systematic discrepancy between the different measures of size, which depends on the radius itself: at small radii, the Sérsic sizes are smaller by ≈ -0.05 dex (left) or ≈ -0.06 dex (right); this discrepancy decreases slightly toward larger radii in the right-hand panel, to a difference of ≈ -0.05 dex. The effect is significantly stronger in the left-hand panel, where the difference even changes sign, reflecting the fact that for massive galaxies the 30 kpc aperture underestimates the full extent of the galaxy.

However, there is a dependence on radius, particularly in the left-hand panel, where at large radii the Sérsic-derived half-mass radii are systematically larger. As also discussed in Section 4.4.1, for very massive galaxies the spherical aperture of 30 kpc is simply too small to encompass the full extent of the galaxy, and the 3D half-mass radii are therefore underestimated. For the larger aperture of 100 kpc this effect is greatly reduced, although there is still a slight increase in $\Delta \log(r_{\mathrm{e}})$ with increasing radius, with a mean difference of -0.053 dex for galaxies with $r_{\mathrm{e},100} > 4$ kpc. The size discrepancies found here appear to be slightly larger than the predictions by van de Ven & van der Wel (2021) (of $\Delta \log(r_{\mathrm{e}}) \approx 0.02$ dex), who derived an analytical prescription for the conversion from Sérsic profile sizes to 3D sizes. On the other hand, the two results are likely to be consistent when taking
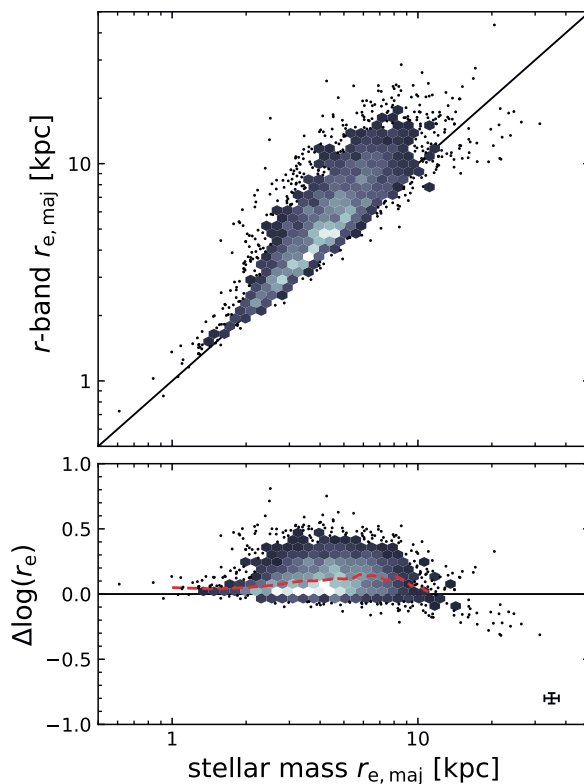
Figure 4.8: Comparison between the sizes obtained from the Sérsic profile fitting to the r-band and stellar mass image data. The dashed line shows the running median of the logarithmic difference between the two size estimates. Both quantities were estimated with the same methodology and using image data with similar noise and equal spatial resolution, and discrepancies can therefore be attributed entirely to radial variations in the mass-to-light ratio within galaxies. The half-light radii are systematically larger than the half-mass radii (typically, 25% larger), and this discrepancy increases slightly toward larger radii, albeit with large scatter.

into account the fact that the axis ratio distributions differ systematically between EAGLE and local observations (as oblate systems in EAGLE are not sufficiently flattened, see Section 4.5.2).

Thus far, however, we have only compared stellar half-mass radii, which give insight into the effect of different methodologies. This does not account for the effects of dust and stellar population gradients that affect the shapes of the light profiles, and hence also the inferred sizes. In Fig. 4.8, we show how the r-band half-light radii compare with the stellar half-mass radii. Here, the methodology is the same for both axes, and discrepancies are therefore entirely due to radial variations in the mass-to-light ratio (M_*/L_r).

The half-light radii are systematically larger than the half-mass radii, with an offset that depends weakly on radius. The two are comparable only for small galaxies ($r_{\mathrm{e,maj}} \lesssim 2\,\mathrm{kpc}$), which are mainly compact quiescent galaxies that may be expected to have only weak M_*/L_r gradients, although we caution that these galaxies are smaller than the PSF FWHM (2.6 kpc). The bottom panel shows the size difference as a function of the half-mass radius, together with the running median (dashed line): on average, the r-band radii are 40% larger (0.14 dex), with a median of 25% (0.10 dex). However, there is also significant scatter (of 0.13 dex), which is asymmetric with excesses of up to ~ 0.5 dex: gradients in M_*/L_r can thus have a great effect on the inferred size for individual galaxies.

Fig. 4.9 examines the origins of the size differences, by showing $\Delta \log(r_{\mathrm{e}})$ as a function of different galaxy properties. In addition to the stellar mass and specific star formation rate (sSFR) measured within a spherical aperture of radius 30 kpc, we extract the mean mass-weighted age and metallicity of the stellar particles within the same aperture. We note that, for visualisation purposes only, we have added a value of $0.01\,\mathrm{M_\odot\,yr^{-1}}$ to the instantaneous SFR before calculating the sSFR. Moreover, we estimate the dust attenuation in the r-band (A_{r}) by calculating the difference between the rest-frame absolute magnitudes with and without dust from Trayford et al. (2015, 2017), although we note that the attenuated magnitudes are only available for galaxies with a minimum of 250 dust particles (2590 galaxies).

The size discrepancy is independent of the stellar mass below $M_* \sim 10^{11}\,\mathrm{M_\odot}$, but at the high mass end the half-light radii become comparable to the half-mass radii. This may reflect minimal dust attenuation or colour gradients at these high masses, although this may also be partially due to the limited spatial extent of the original SKIRT images (60 kpc versus 114 kpc in the stellar mass images).

On the other hand, there is a strong correlation with the sSFR, and a similar trend is visible for the stellar age, with the youngest galaxies having much higher values of $\Delta \log(r_{\mathrm{e}})$ (by ≈ 0.15 dex) than the very oldest systems. Interestingly, we find no such correlation with the stellar metallicity.

Furthermore, edge-on galaxies (low projected axis ratios), which tend to have higher optical depths due to dust, have relatively large half-light radii. Observationally, this effect may be even stronger, as edge-on EAGLE galaxies are thicker than observed in the local Universe and thus also have significantly lower dust optical depths (see Trayford et al. 2017). The effects of dust are also visible in the lower right panel, which shows a weak, positive correlation between $\Delta \log(r_{\mathrm{e}})$ and the r-band dust attenuation. If we select only the highly star-forming galaxies ($\mathrm{sSFR} > 10^{-10.4}\,\mathrm{yr^{-1}}$), the effects of dust become even more pronounced: although the correlation with A_{r} becomes negligible, the anti-correlation with the axis ratio becomes slightly stronger (Spearman rank coefficient $\rho = -0.32$), which suggests that the dust geometry is an important factor. The increased spatial extent in the r-band imaging with respect to the stellar mass imaging can therefore be attributed to the presence of bright star-forming regions in the outskirts of galaxies and/or significant dust attenuation in the centre.
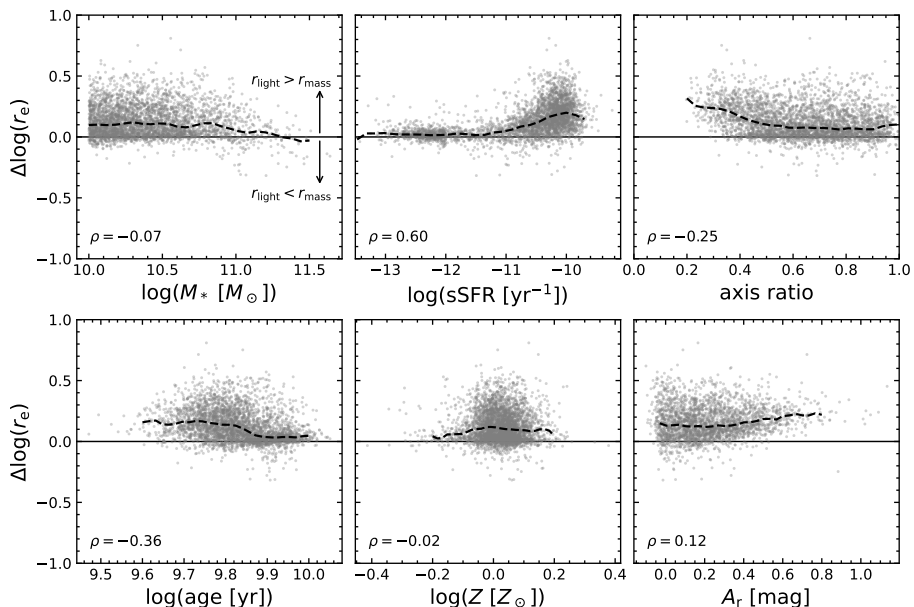
Figure 4.9: Difference between the r-band half-light radii and stellar half-mass radii ($\Delta \log(r_\mathrm{e})$) as a function of stellar population and dust properties. Dashed lines show the running median in each panel, and the value of the Spearman rank correlation coefficient (ρ) is indicated in each panel. The size difference is largely independent of the stellar mass and mass-weighted stellar metallicity, but depends strongly on the star formation activity, reflected by the positive correlation with the sSFR and negative correlation with the mass-weighted stellar age. Dust also has a significant effect, as edge-on galaxies show a stronger size discrepancy, and the r-band dust attenuation (A_r) correlates weakly with $\Delta \log(r_\mathrm{e})$. Star formation in the outskirts of galaxies, as well as dust attenuation in the central regions therefore likely drive the discrepancies between light and mass-weighted sizes.

4.4.3 Stellar mass-size relation

As described in Crain et al. (2015), for the EAGLE simulations the stellar mass-size relation of late-type galaxies at $z \sim 0$ was used in the calibration of the subgrid model parameters. Specifically, of the four subgrid models considered, three were rejected due to the simulations producing unrealistic size distributions for the massive galaxy population ($> 0.2\,\mathrm{dex}$ below the mass-size relation from Shen et al. 2003), and mass-size relations that decline with mass, rather than increase. Although the subgrid model was not fine-tuned to reproduce the observed mass-size relation, the low-redshift mass-size relation in EAGLE can also not be considered to be a true prediction of the simulation (Schaye et al. 2015). However, this calibration was done using stellar half-mass radii that were measured by fitting Sérsic profiles to the projected, azimuthally-averaged stellar mass density profiles, and used only the subset of EAGLE galaxies with Sérsic index $n < 2.5$.
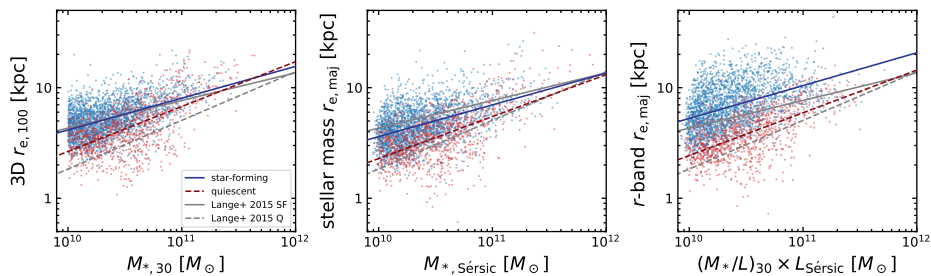
Figure 4.10: Stellar mass-size relation of quiescent (red) and star-forming (blue) EA-
GLE galaxies for three different measures of galaxy size and stellar mass. The left-hand
panel shows the stellar masses and half-mass radii from the public EAGLE catalogues,
which are measurements within spherical apertures of fixed radius. The middle panel
shows the half-mass radius and total mass of the best-fit Sérsic profiles to the stellar
mass imaging. On the right, the results from the r-band Sérsic fits are presented, with
stellar masses corrected to the total luminosity of the best-fit profile. Coloured lines
show the best-fit power law relations in each panel. Grey lines indicate the best-fit
r-band mass-size relations from Lange et al. (2015), for star-forming (solid) and quies-
cent (dashed) galaxies at $z \sim 0$ in the GAMA survey (where quiescence is defined by
the dust-corrected, rest-frame $u - r$ colour). Only in the right panel, with the fully
forward-modelled sizes, is there a clear separation between the star-forming and quies-
cent populations, and are both the slope and scatter about the relation comparable with
observations (see main text). The zero-point offsets are slightly higher than in GAMA,
however, indicating that both quiescent and star-forming galaxies in EAGLE are system-
atically larger (by $\approx 0.1\,\mathrm{dex}$) than observed in the local universe.

We showed previously that there are significant, systematic differences between dif-
ferent measures of size, which may therefore affect the inferred mass-size relation,
and potentially also the calibration of a subgrid model.

   We explore the effects of different size estimates on the mass-size relation in
Fig. 4.10. The left-hand panel shows results that are similar to the work by Furlong
et al. (2017), who presented the redshift evolution of the mass-size relation in the
EAGLE simulations. They defined the galaxy stellar mass as the mass enclosed
within a spherical aperture of radius 30 kpc, and the stellar half-mass radius as
the 3D $r_{\mathrm{e},100}$ (see also Section 4.4.2). Observations indicate different evolution for
late- and early-type galaxies, which holds true regardless of the method used to
define 'late' versus 'early' (by colour, morphology, or SFR; see, e.g., Shen et al.
2003). We therefore divide our sample by the instantaneous sSFR within the
30 kpc spherical aperture, and define the late-type or star-forming population as
having sSFR $> 10^{-11}\,\mathrm{yr}^{-1}$; the star-forming and quiescent galaxies in Fig. 4.10
are indicated in blue and red, respectively.

   Perhaps unsurprisingly, considering the subgrid model calibration, the star-
forming EAGLE galaxies closely follow the observed mass-size relation at $z \sim 0$
measured by Lange et al. (2015) using r-band Sérsic models of GAMA galaxies.
The grey solid and dashed lines show the single power-law fits ($r_{\mathrm{e}} = a\,(M_*/\mathrm{M_\odot})^b$)
to the star-forming and quiescent subsamples, respectively, where quiescence for

GAMA galaxies is defined using the rest-frame $u - r$ colour that is corrected for dust attenuation within the galaxy. If we fix the exponent to the local relation ($b = 0.25 \pm 0.02$), and perform a least-squares fit in logarithmic space to determine the normalisation a, we find excellent agreement between EAGLE ($\log(a) = -1.870 \pm 0.003$; where the error bar is obtained via bootstrap resampling) and GAMA ($\log(a) = -1.87 \pm 0.05$). If we instead fit both a and b simultaneously (coloured lines), we find an exponent $b = 0.287 \pm 0.009$ that is slightly steeper than observed, however, in good agreement with observations when considering the stellar mass limit imposed here ($M_* > 10^{10}\,M_\odot$ versus $M_* \gtrsim 10^9\,M_\odot$ for GAMA) and the fact that the mass-size relation has been shown to steepen toward high stellar mass (Shen et al. 2003).

The quiescent population, on the other hand, deviates strongly from the observed relation (dashed lines). Although the best-fit exponent of $b = 0.406 \pm 0.017$ is close to the observed value of $b = 0.44 \pm 0.02$, the normalisation is significantly higher: at fixed $b = 0.44$, the EAGLE galaxies are 0.14 ± 0.03 dex larger than observed ($\log(a) = -3.994 \pm 0.005$ versus $\log(a) = -4.14 \pm 0.03$).

However, the relations from Lange et al. (2015) are based on semi-major axis sizes from Sérsic models. The middle panel of Fig. 4.10 shows the result of using the half-mass radii obtained with the Sérsic profile fits for the stellar mass images. We emphasise that not only the size changes with respect to the left-hand panel, but also the stellar mass is replaced with the total mass of the best-fit Sérsic profile.

To take into account measurement uncertainties and the intrinsic scatter about the relation, we follow the maximum likelihood fitting method described by van der Wel et al. (2014a) to estimate the best-fit parameters of the power-law model. This method assumes there is intrinsic scatter (i.e., not due to measurement uncertainties) about the mass-size relation that follows a Gaussian distribution, and fits the intrinsic scatter as an additional variable to the zero point (a) and slope (b). Moreover, uncertainties in M_* are treated as an additional uncertainty in $\log(r_{\rm e})$.

We find that the slope of the relation for the Sérsic model sizes is changed minimally with respect to the aperture-based sizes, with $b = 0.287 \pm 0.010$ and $b = 0.379 \pm 0.016$ for the star-forming and quiescent samples, respectively. There is a significant change in the intercept, however, as this deviates by -0.060 ± 0.004 dex and -0.077 ± 0.008 dex respectively for the star-forming and quiescent populations. These values are in line with the systematic offsets found in Fig. 4.7, and slightly enhanced by the fact that the stellar masses are also marginally smaller than the aperture-derived masses for the majority of the sample (Fig. 4.6). The result of moving from 3D half-mass sizes to major axis sizes from the stellar mass Sérsic models is thus that the star-forming population appears systematically smaller than the observed mass-size relation by 0.06 dex. On the other hand, the agreement with observations is significantly improved for the quiescent population, although these galaxies are still systematically larger than observed by 0.06 dex.

Lastly, we take into account the effects of stellar population gradients and dust, by using the Sérsic fits to the r-band imaging rather than the stellar mass fits. Again, it is not only the parameter on the vertical axis that changes, but we also adjust the stellar mass: we correct the aperture-based mass ($M_{*,30}$) by multiplying

with the ratio of the total flux of the Sérsic profile and the flux measured within a circular aperture of 30 kpc.

The resulting mass-size relation in the right-hand panel differs from the other two panels by not just the zero points offsets, but also the scatter. The first effect is mainly in the relative difference between the star-forming and quiescent populations. Whereas the two populations overlap quite significantly when considering the half-mass radii, there is a larger separation when using the half-light radii. Interestingly, it is the star-forming population that changes with respect to the middle panel: the quiescent population is moved only slightly, as these sizes are larger than the observed relation by 0.10 ± 0.03 dex (at fixed $b = 0.44$, $\log(a) = -4.040 \pm 0.005$). If we also fit the exponent, we find $b = 0.386 \pm 0.015$, which is slightly shallower than the observed value (by 2.2σ), although this measurement may be affected by the limited spatial extent of the r-band images (as discussed in Section 4.4.2). On the other hand, the star-forming population moves towards much larger r_e at fixed stellar mass, and is 0.11 ± 0.05 dex larger than the observed relation ($\log(a) = -1.760 \pm 0.003$ at fixed $b = 0.25$). The best-fit exponent, $b = 0.297 \pm 0.012$, is slightly steeper than observed (by 1.9σ), but likely in good agreement with observations when taking into account the difference in the mass range used for the fitting.

Both populations are thus ≈ 0.1 dex larger at fixed stellar mass than observed, but the separation between the quiescent and star-forming populations matches that of the observed relations almost exactly. We can therefore conclude that colour gradients strongly affect the mass-size relation of the star-forming sample, and have only a moderate effect on the quiescent population.

The second difference with respect to the other panels is in the scatter in $\log(r_e)$ about the relation. For the star-forming population, the scatter in the half-light radii appears to be much closer to the observed scatter: using the normalised median absolute deviation (NMAD), we find $\sigma(\log r_e) = 0.19$ dex for the half-light radii, versus $\sigma(\log r_e) = 0.14$ dex and $\sigma(\log r_e) = 0.15$ dex for the 3D and Sérsic half-mass radii, respectively. On the other hand, the scatter for the quiescent population is approximately equal for all three measures of size (from left to right, $\sigma(\log r_e) = 0.16$ dex, $\sigma(\log r_e) = 0.15$ dex and $\sigma(\log r_e) = 0.15$ dex). Although these measurements are not provided explicitly by Lange et al. (2015), we obtain $\sigma(\log r_e) = 0.20$ (star-forming) and $\sigma(\log r_e) = 0.18$ (quiescent) for a $z \sim 0.1$ comparison sample selected from GAMA (sample selection described in Section 4.5.1). We note that the uncertainties on the size measurements in GAMA are expected to be larger than is the case for the EAGLE galaxies (e.g., due to additional uncertainties from the sky background). Therefore, whereas the observed scatter about the mass-size relation of star-forming galaxies agrees well between EAGLE and GAMA, the intrinsic scatter may be slightly too large for the EAGLE galaxies.

4.5   Galaxy morphologies

We now turn to the morphological properties of the EAGLE galaxies as quantified by the Sérsic index and projected axis ratio. We compare our results with a low-redshift sample of galaxies selected from the GAMA survey, which is approximately volume-limited above our stellar mass limit at $z \sim 0.1$. The optical imaging and derived data products of GAMA are largely based on SDSS imaging, and are therefore of similar quality to our constructed mock images and model fits.
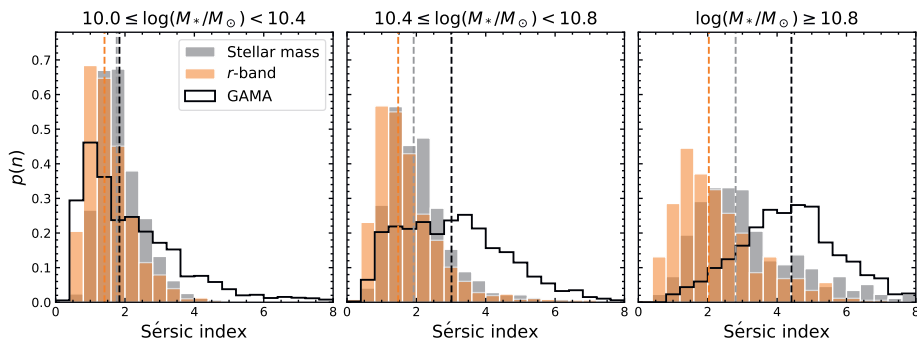
4.5.1   Sérsic indices

The Sérsic index characterises the shape of the surface brightness profile: a value of $n \sim 1$ describes an exponential profile, often found in late-type galaxies, whereas local early-type galaxies tend to be well approximated by profiles with a value of $n \sim 4$ (de Vaucouleurs profile). In Fig. 4.11, we show the probability distributions of the Sérsic indices measured from the r-band (orange) and stellar mass (grey) images. Observationally, the shape of the surface brightness profile has been shown to correlate with different physical properties, such as the mass (or luminosity), colour and spectral age indicators (e.g., Blanton et al. 2003; Kauffmann et al. 2003). The sample is therefore split into three bins in stellar mass ($M_{*,30}$), with dashed lines showing the median values in each panel.

All mass bins show distributions that are skewed toward low values of n, indicating that the majority of galaxies are best described by profiles that closely resemble exponential discs. However, there is also an extended tail toward higher n, representing bulge-like profiles, which becomes more prominent at higher masses. This mass dependence is also apparent in the evolution of the median, as this increases from $n \approx 1.5$ ($n \approx 1.8$) in the lowest mass bin to $n \approx 2.0$ ($n \approx 2.8$) in the highest mass bin for the r-band (stellar mass) fits.

Interestingly, the r-band imaging shows systematically different Sérsic indices from the stellar mass imaging. The stellar mass profiles tend to be more concentrated in the centre, particularly at high stellar mass, with profiles that are closer to a classical de Vaucouleurs profile. In Appendix 4.A we demonstrate that this is not due to the smoothing lengths used to create the r-band images, as we find identical results for a smoothed version of the stellar mass images. Rather, colour gradients appear to have a strong effect on the shape of the light profile, as was also noted by Kelvin et al. (2012), who found systematic differences between their measurements of n in the r-band and at near-infrared wavelengths (e.g., the K-band). Younger stellar populations at larger radii have low M_*/L, particularly at shorter wavelengths, which may drive the Sérsic fit to lower observed values of n than expected from the underlying stellar mass profile. This is in line with the findings by Trayford et al. (2019), who showed that, based on the orbital properties of the stellar particles, younger stellar populations within EAGLE galaxies tend to reside in discs. Similarly, the effects of dust attenuation in the centre of the galaxy likely result in lower Sérsic indices at rest-frame optical wavelengths.

To compare more directly with observational data, we use the catalogue of single Sérsic profile fits to reprocessed SDSS r-band imaging from Kelvin et al.

Figure 4.11: Sérsic index distributions of EAGLE galaxies at $z = 0.1$, in bins of increasing stellar mass. Results of the fits to the stellar mass and r-band images are shown in grey and orange, respectively, with dashed lines indicating the medians in each panel. In black, we show the selected comparison sample from the GAMA survey with Sérsic profile fits in the r-band by Kelvin et al. (2012). All distributions are normalised such that their integral is equal to 1. The light and stellar mass profiles of the EAGLE galaxies are skewed toward low Sérsic indices, with only a slight increase in the median value of n toward higher stellar mass. In comparison with the GAMA data, EAGLE is deficient in bulge-like ($n \sim 4$) systems. The discrepancy between the distributions becomes stronger at higher stellar masses, and suggests a fundamental difference in the stellar mass density profiles of simulated and observed galaxies.

(2012). We select all galaxies within $0.06 < z < 0.12$ and match this morphological catalogue with the stellar masses from Driver et al. (2016), which were estimated using MAGPHYS (da Cunha et al. 2008). The stellar masses are then scaled to the total flux of the best-fit Sérsic profiles, and we select only galaxies with $\log(M_*/\mathrm{M_\odot}) \geq 10$. We exclude galaxies with Sérsic indices outside of $0.2 < n < 8$ for fair comparison with our own sample. Moreover, to filter out poor fits, we require that the reduced chi-squared value of the primary galaxy is within $0.5 < \chi^2_{\nu,\mathrm{pri}} < 2$. These criteria result in a final catalogue of 6554 GAMA galaxies with a median redshift of $z \approx 0.1$.

The GAMA survey is highly complete for the selected mass and redshift range, and can therefore readily be compared with the EAGLE sample, which is by construction volume-limited. The distributions of the Sérsic indices of the GAMA galaxies are shown in black in Fig. 4.11. For the lowest masses ($M_* \sim 10^{10.2}\,\mathrm{M_\odot}$), there is reasonable agreement between the observed and simulated r-band data, as both distributions peak around $n \sim 1.5$. However, the GAMA data show a less strongly peaked distribution at low n, and a more significant tail toward $n \sim 4$. This large number of bulge-like galaxy profiles is simply missing in EAGLE, and this discrepancy becomes even stronger at higher masses, where GAMA consists predominantly of high n systems. The fact that these discrepancies remain when comparing the stellar mass values of n with the GAMA data, suggests that it is the intrinsic mass distribution that differs from observations, rather than potential issues in the forward modelling (e.g., uncertainties in the dust properties and ge-

ometry). We further discuss the discrepant mass distribution of simulated galaxies in Section 4.6.2.
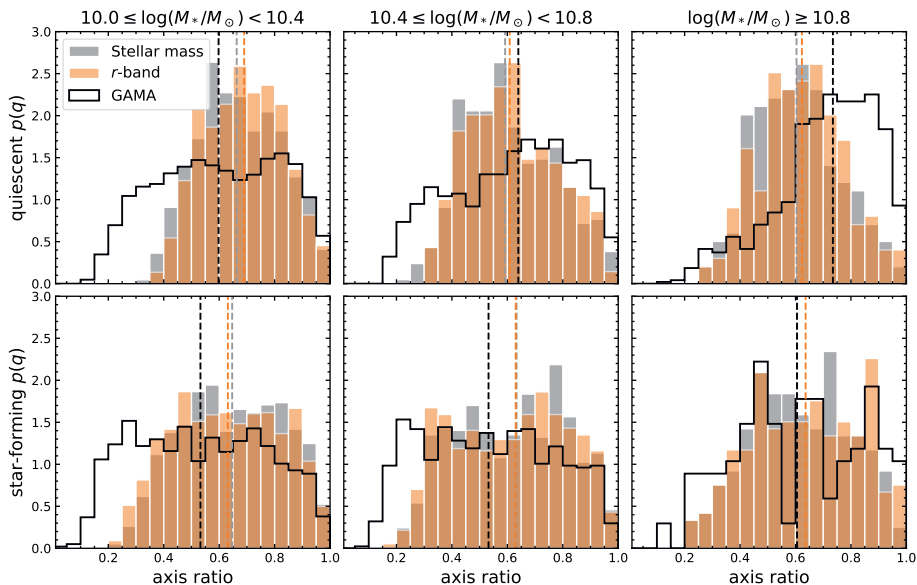
4.5.2   Axis ratios

The second morphological parameter is the ratio between the semi-major and semi-minor axes, which provides insight into the intrinsic shape of a system. However, due to projection effects, this cannot be done on an object-by-object basis. Rather, it is the distribution of axis ratios that is often used to infer the distribution of the intrinsic shapes for a sample of galaxies (see, e.g., Holden et al. 2012; Chang et al. 2013; van der Wel et al. 2014b).

As this is typically done separately for star-forming and quiescent galaxy populations, we divide our sample by the sSFR, as in Section 4.4.3. For GAMA we use the sSFR averaged over the last $100\,$Myr from the MAGPHYS SED modelling. The projected axis ratio distributions for these two populations are shown in Fig. 4.12, in bins of stellar mass. The orange and grey histograms show the r-band and stellar mass results for EAGLE, respectively, and black corresponds to the GAMA results; dashed lines indicate median values.

Using a Kolmogorov-Smirnov (KS) test, we find that both the r-band and stellar mass distributions agree very well. In contrast with the Sérsic index distributions (Fig. 4.11), which show that the stellar mass and light density profiles differ significantly, the stellar mass and light do trace the overall (3D) shapes of galaxies in the same manner. Moreover, we find that the star-forming and quiescent EAGLE galaxies indeed follow significantly ($> 3\sigma$) different distributions, apart from the highest mass bin, where the number of galaxies is also substantially smaller (~ 200 versus ~ 700 at lower masses). The distribution of quiescent EAGLE galaxies shows a peak around $q \approx 0.65$ in each panel, whereas the star-forming galaxies show a more uniform spread, which may be explained by a larger proportion of disc-like (oblate) systems within the star-forming population.

On the other hand, the median values of the distributions do not differ strongly between the two populations, nor show a dependence on the stellar mass, both of which are clear features in the GAMA data. The most significant difference between the observed and simulated data is at low axis ratios: a large number of GAMA galaxies are highly flattened (with $q \approx 0.2$), yet, these galaxies do not exist in the EAGLE simulations. As discussed by Trayford et al. (2017) and van de Sande et al. (2019), galaxies in EAGLE tend to be thicker than is observed, possibly as the result of the pressure floor that is imposed within the simulation. Moreover, the limited mass resolution of the dark matter particles in the simulation has been shown to lead to a heating of the baryonic particles via 2-body scattering (Ludlow et al. 2019, 2021).

Furthermore, there is a discrepancy between the GAMA and EAGLE data at high q for the quiescent galaxies. Whereas the EAGLE data show very little dependence on mass, the axis ratio distribution in GAMA is increasingly skewed toward high q at higher masses. This difference is most apparent in the highest mass bin, which peaks at $q \approx 0.8$ for the GAMA galaxies. Chang et al. (2013) showed that low-redshift galaxies in this mass range are mainly triaxial systems ($\approx 80\%$ of the

Figure 4.12: Projected axis ratio distributions of the EAGLE and GAMA comparison galaxies in bins of increasing stellar mass, further separated into quiescent (top panels) and star-forming (bottom panels) subsamples. Symbols indicate the same as in Fig. 4.11. Unlike in Fig. 4.11, the stellar mass and r-band fits show good agreement in all mass bins, and for both star-forming and quiescent galaxies. In line with observations, the quiescent subsamples show more strongly peaked distributions than the star-forming subsamples, which is consistent with a higher proportion of disc-like intrinsic shapes among the star-forming population. The main difference between the simulated and GAMA data is at low axis ratios, as there are no highly flattened systems in EAGLE, likely due to the imposed gas pressure floor and the limited resolution of the simulation. In the highest mass bin, the quiescent GAMA sample is skewed toward higher axis ratios, which implies that the GAMA galaxies are intrinsically rounder in shape than the EAGLE galaxies.

sample) with a mean intrinsic major-to-minor axis ratio of $C/A \approx 0.6$ and mean triaxiality parameter $T \approx 0.6$. For EAGLE, the highest mass bin may still contain a significant number of triaxial systems, but with more flattening along the intermediate or minor axes. Indeed, based on the 3D stellar mass distribution, Thob et al. (2019) showed that there is a significant population of triaxial systems and prolate systems in EAGLE, with $C/A \sim 0.4$ among galaxies in the red sequence. Contrary to the intermediate mass galaxies ($\log(M_*/\mathrm{M}_\odot) \lesssim 10.8$) that are not flattened enough, at the highest masses the simulation thus struggles to reproduce galaxies that are sufficiently round.

4.6 Discussion

4.6.1 The importance of apples-to-apples comparisons

The differences and similarities found between the structural properties of simulated and observed galaxies, only truly become apparent when using mock observations and the same measurement techniques as in large galaxy surveys. This can also lead to different conclusions: Correa et al. (2017) showed, using the stellar kinematics to identify spheroids and discs, that kinematic morphologies in EAGLE are tightly correlated with the $u - r$ colour, with central galaxies along the red sequence being dominated by spheroidal morphologies. Although we find some dependence of q and n on stellar mass and sSFR, the lack of $n \sim 4$ systems would lead to a different picture of the red sequence.

Perhaps more important, however, is the remarkable improvement in the stellar mass-size relation when M_*/L gradients are modelled, and the definition of galaxy size is made consistent with observations. The mass-size relation is often used as a key measure of success for cosmological simulations, and in the case of EAGLE also plays a role in the calibration of the subgrid model.

Yet, we have found that the 3D curve of growth methods commonly used to measure half-mass or half-light radii differ systematically from the semi-major axis sizes obtained with 2D Sérsic modelling. It is therefore difficult to directly compare the resulting mass-size relation with observations. Comparison with circularised sizes ($r_{e,circ} \equiv \sqrt{q}\, r_{e,maj}$), as done by, e.g., Genel et al. (2018) or van de Sande et al. (2019) using 2D growth-curve sizes, is possibly even more complex, as there is an additional dependence on the distribution of the projected axis ratios. Rather, the semi-major axis is the preferred measure of size here, as it is largely independent of inclination and intrinsic axis ratio for oblate systems, which is the most commonly found shape of $z \sim 0$ galaxies (Chang et al. 2013; van der Wel et al. 2014b, although the effects of dust complicate this slightly, as demonstrated in Fig. 4.9). When comparing with observations, this therefore allows to distinguish between a possible systematic offset in the sizes and a mismatch in the distribution of the intrinsic shapes.

By making a consistent comparison using the semi-major axis sizes, we have shown that the half-mass radii turn out systematically smaller than measurements in the r-band from GAMA (Lange et al. 2015), whereas there is significantly better agreement with observations for the quiescent population. The excellent agreement found previously between the 3D half-mass radii and observed r-band sizes of star-forming galaxies (Furlong et al. 2017) is therefore partially the result of the model calibration, and to some extent simply coincidence.

However, accounting for gradients in M_*/L with the use of the mock r-band imaging brings the star-forming population in good agreement again with the observed relation, and with a scatter that is closer to that observed. As also shown by van de Sande et al. (2019), the effect of using luminosity-weighted sizes rather than mass-weighted sizes is significant, and is further enhanced by the implementation of realistic dust attenuation in this work (see also Gadotti et al. 2010, for the effects of dust on the measurement of structural parameters). On

the other hand, the location of the quiescent population is changed only minimally within the mass-size plane, consistent with the expectation that these galaxies have less variation in M_*/L.

Remaining discrepancies, the sizes of both star-forming and quiescent EAGLE galaxies are approximately $0.1\,$dex larger at fixed mass, can be caused by a large number of factors within the simulation itself. Additionally, uncertainties in the radiative transfer modelling (e.g., the treatment of molecular clouds in the ISM) may introduce a systematic uncertainty on the M_*/L gradients, and hence the size measurements. The deviating shapes and morphologies of the simulated galaxies will also affect the simulated M_*/L gradients, as the results of the radiative transfer calculations are dependent on the geometry of both the stellar particles and dust (e.g., for the difference in the dust attenuation between thin and thick discs, see Trayford et al. 2017). Finally, it is also important to bear in mind that the stellar masses inferred with SED modelling carry large uncertainties ($\approx 0.3\,$dex at $z \sim 0$; Conroy et al. 2009), which can introduce a systematic uncertainty of similar magnitude in the observed mass-size relation (see Genel et al. 2018).

4.6.2   Mismatched density profiles and intrinsic shapes

We have demonstrated that the morphologies of EAGLE galaxies, as quantified by the Sérsic index, differ significantly from observations: at all stellar masses ($\log(M_*/\mathrm{M_\odot}) > 10$), there are too few galaxies with bulge-like ($n \sim 4$) light profiles. The fact that this discrepancy holds true also for the stellar mass surface density profiles, shows that the mass is distributed differently in simulated galaxies, and that observational effects (measuring light versus stellar mass, effects of dust attenuation) are of secondary importance.

We highlight this finding in Fig. 4.13, where we show the fraction of stellar mass enclosed within a fixed aperture of radius $2\,$kpc as a function of the total stellar mass of the best-fit model. Both the mass fractions and total masses are inferred from the best-fit Sérsic models, therefore demonstrating the physical difference between the Sérsic index distributions in EAGLE and GAMA: the stellar mass fractions in EAGLE are a factor ≈ 2 below the observed mass fractions in GAMA, irrespective of the stellar mass.

A deficiency in bulge-like systems was also found by Bottrell et al. (2017a,b), who constructed mock SDSS observations of galaxies in the Illustris simulations (Vogelsberger et al. 2014) and performed a two component (bulge + disc) Sérsic profile fitting to determine the bulge fractions. As also noted by Bottrell et al. (2017b), contrary to issues with early hydrodynamical simulations producing galaxies that were too bulge-like (e.g., Katz & Gunn 1991), it therefore appears that some of the more recent models struggle to form enough bulges. Rodriguez-Gomez et al. (2019) showed, using single Sérsic profile fits to mock Pan-STARRS imaging, that this is also the case for simulated galaxies in Illustris-TNG: although the galaxy sizes are in significantly better agreement with observations than was the case for Illustris, the morphologies are still too disc-like, as galaxies below a mass of $\log(M_*/\mathrm{M_\odot}) \approx 10.7$ follow profiles with $n \approx 1.5$. Only for the most massive galaxies ($\log(M_*/\mathrm{M_\odot}) \gtrsim 11$) is there good agreement between the simulated
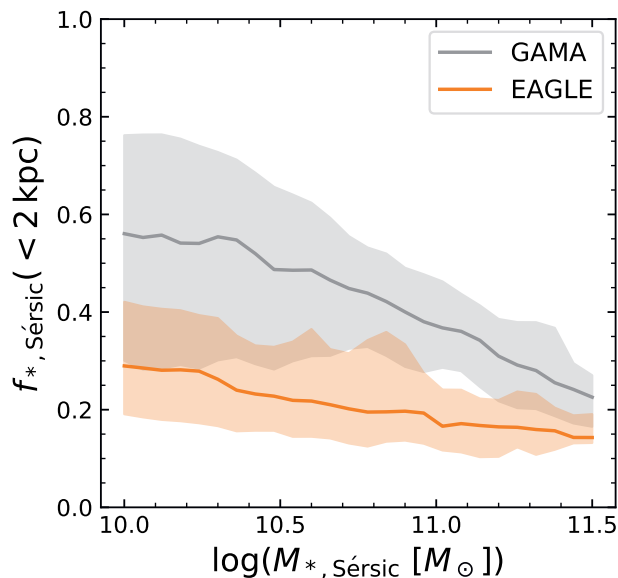
Figure 4.13: The stellar mass fraction within an aperture of radius 2 kpc measured from the best-fit Sérsic model, as a function of the total stellar mass inferred from the same model. Solid lines show the running median of the EAGLE (orange) and GAMA (grey) data, and shaded regions mark the 16th to 84th percentile range. In comparison with observations, the inferred stellar mass density profiles of EAGLE galaxies are deficient in mass at small scales.

and observed Sérsic indices.

Interestingly, non-parametric methods of quantifying morphology paint a slightly different picture. Based on the same optical imaging used in this work, but with slightly different noise and instrument resolution applied, Bignone et al. (2020) find that the distributions of nearly all commonly used non-parametric measures (e.g., Gini coefficient, Concentration parameter; Lotz et al. 2004) match well with observations from the GAMA survey. The non-parametric morphologies of galaxies in Illustris-TNG (Rodriguez-Gomez et al. 2019) are approximately equally successful in reproducing observations, demonstrating a substantial improvement with respect to previous measurements from the Illustris simulations (Snyder et al. 2015).

Based on the discrepant Sérsic indices alone, it may be tempting to conclude that there are improvements to be made in the physics implemented in the simulations, such as the feedback prescriptions. Yet, the non-parametric morphologies do not show a strong indication for this, particularly in EAGLE, where the non-parametric measures additionally correlate with the stellar mass and sSFR in the same way as in observations.

Reconciling the different outcomes of these two strategies of measuring galaxy

morphologies is not immediately obvious, however, the residual surface brightness and density profiles may provide some insight. As shown in Fig. 4.5, there are minor, systematic features in the residual flux (i.e., the difference between the mock image and best-fit Sérsic model). Although the excess low surface brightness emission at large radii ($\gtrsim 3\,r_e$) could simply reflect the fact that a two-component model (bulge+disc decomposition) would be a better description for the galaxy profiles, the features at smaller radii are not as easily 'fixed'.

The under-subtraction at $r < r_e$ ($\approx 2\,\text{kpc}$) followed by over-subtraction at $r \approx 1 - 2\,r_e$ ($\approx 5\,\text{kpc}$) suggests that the surface brightness profile declines more steeply than a $n \sim 2$ model describes. Similarly, the excess flux at $r \gtrsim 3\,r_e$ indicates that the profile is shallower than a $n \sim 2$ profile at large radii. A steep decline at small radii followed by a gradual decline at large radii is characteristic of a high Sérsic index profile ($n \gtrsim 4$). However, likely due to the high S/N in the central pixels, the fit is driven to low Sérsic indices. It therefore appears that the simulated galaxies are simply deficient in mass and light at the very centre in comparison to the rest of the galaxy, which may be the effect of the resolution limit in the simulation (see also Schaller et al. 2015) and the associated 2-body scattering of dark matter and baryonic particles (Ludlow et al. 2019, 2021). The pressure floor within the simulation may also play a role here, as the associated spatial scale of $\sim 1\,\text{kpc}$ likely affects the inner density mass density profiles, and therefore the measured Sérsic indices. These effects may also explain the similarities between the sizes and morphologies in the EAGLE and Illustris-TNG simulations (Genel et al. 2018; Rodriguez-Gomez et al. 2019), as, although the two simulations employ different physical models (e.g., the feedback prescriptions), both use a similar resolution and pressure floor.

A bulge+disc decomposition then also does not offer substantial improvement, because the profile shape in the centre deviates too strongly from a Sérsic profile. On the other hand, within the apertures used to calculate non-parametric morphological measures, these features in the light profiles may be washed out, and thus provide an explanation for their better consistency with observations. Restricting the Sérsic profile fitting to $r > 2\,\text{kpc}$ may then be an appropriate method to minimise the effects of the unrealistic profile shapes in the centres.

This breakdown in the density profiles at small scales would also help to reconcile the discrepant results found between the Sérsic index distributions and the projected axis ratios, as we may expect galaxies with $n \sim 1 - 2$ profiles to have disc-like (oblate) intrinsic shapes. Whereas this is likely the case for the star-forming population (Fig. 4.12), which only differ from observations by the lack of highly flattened galaxies, the quiescent galaxies show projected axis ratios that more plausibly reflect a large population of triaxial and prolate systems, as also shown to be present in EAGLE by Thob et al. (2019) and Trayford et al. (2019).

Although the quiescent galaxies are not as round as seen in observations, particularly at high stellar mass, the projected axis ratios show a picture that is closer to reality than would be concluded from the Sérsic indices alone, and is more consistent with the variety of bulges and discs found in studies that use kinematic morphologies as a proxy for the observed morphology (e.g., Correa et al. 2017; Clauwens et al. 2018). This then raises the question of whether Sérsic indices have
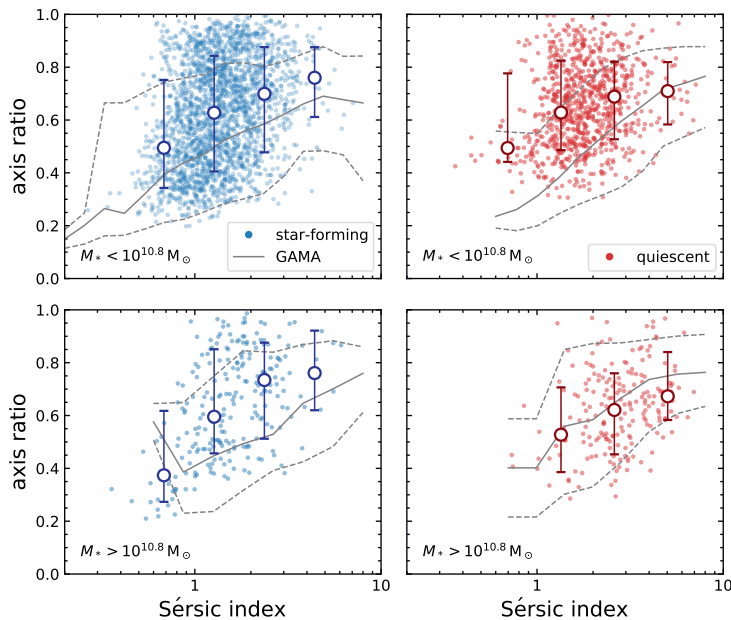
Figure 4.14: The r-band projected axis ratio versus Sérsic index, for star-forming (blue) and quiescent (red) EAGLE galaxies in two different stellar mass ranges (intermediate masses in top panels; high masses in bottom panels). Open circles show the median axis ratios of the EAGLE galaxies in logarithmic bins in Sérsic index, with error bars indicating the 16th and 84th percentiles. Grey lines show the 16th, 50th and 84th percentiles for the GAMA survey. The axis ratio generally increases toward higher Sérsic index, consistent with the picture of $n \sim 4$ galaxies being rounder in shape, although the scatter is large. Whereas Fig. 4.12 showed that massive quiescent galaxies in GAMA are intrinsically more round than is the case in EAGLE, this is partly due to the difference in the Sérsic index distributions: the shapes of the few bulge-like quiescent galaxies in EAGLE are in good agreement with observations.

any predictive power for the intrinsic shapes of simulated galaxies.

In Fig. 4.14 we show the projected axis ratios as a function of the Sérsic index for the r-band Sérsic models, with different panels separating the star-forming (left) and quiescent (right) populations, as well as massive and less-massive galaxies (top versus bottom panels, the boundary used being $M_* = 10^{10.8} \, \mathrm{M}_\odot$). In addition to showing the individual EAGLE galaxies, the open circles show the median axis ratios in bins of Sérsic index (with error bars showing the 16th and 84th percentiles). For comparison, we also show the running median for the comparison sample from GAMA (solid lines), with dashed lines indicating the 16th and 84th percentiles.

For the star-forming galaxies, the EAGLE data show a positive correlation between axis ratio and Sérsic index, albeit with large scatter. This is similar to the correlation and scatter present in the GAMA data, except for an offset toward

slightly higher axis ratio at fixed Sérsic index, which is likely explained by the lack of thin discs in EAGLE (as also discussed in Section 4.5.2).

The quiescent galaxies of $M_* < 10^{10.8}\,M_\odot$ (top right panel) do not show such a clear correlation, and thereby diverge from the trend seen in GAMA. On the other hand, the massive quiescent galaxies (bottom right panel) do show a slight increase in the median axis ratio toward higher Sérsic indices, and follow the observed correlation almost exactly, although the number of EAGLE galaxies (199) in this panel is relatively small.

The EAGLE simulations thus do appear to produce galaxies that resemble the classical picture of spheroidal galaxies with $n \sim 4$ light profiles, only not in sufficient number. As suggested previously, the resolution or the pressure floor may play a role at the small scales probed with these density profiles. Other, more physical effects could be in the implementation of the central black hole and stellar feedback in the simulations: the orbital structure of both the dark matter and stellar particles depend on the feedback mechanisms employed, with strong (black hole) feedback resulting in a higher fraction of box orbits and thus more strongly triaxial systems (Bryan et al. 2012). The observed axis ratio distribution may therefore offer an interesting constraint on the subgrid model, being an observable that is not as model-dependent as the light profile shape or definition of size.

4.7 Conclusions

Starting from the optical images of $z = 0.1$ EAGLE galaxies constructed with SKIRT by Trayford et al. (2017), we have created mock r-band images that have similar noise properties and resolution as photometric data from the SDSS. Following methods that are commonly used in observational studies, we have fitted Sérsic profiles to these mock observations using a combination of SEXTRACTOR and GALFIT, thus enabling an apples-to-apples comparison between the structural parameters of galaxies in EAGLE and local observations.

To be able to distinguish between the effects of different measurement techniques and the effects of variations in M_*/L (due to, e.g., recent star formation or dust attenuation), we have constructed a second set of images from the projection of the stellar mass particles. These stellar mass images are created such that the noise and resolution match the mock optical images.

Our findings can be summarised as follows:

- Galaxy sizes depend on the measure of size used, as there are systematic differences between the half-mass radii estimated with a curve-of-growth method (common in theoretical work) and the semi-major axes obtained with Sérsic profile modelling (common in observational studies). The magnitude of this discrepancy is on average $\approx 0.06\,\mathrm{dex}$, but is itself dependent on the galaxy size.

- Gradients in M_*/L due to radial variations in the star formation, stellar age and dust attenuation can have a large effect on the observed size: half-light radii are typically 25% larger than half-mass radii, but with large scatter

and outliers that deviate by as much as a factor ≈ 3. For quiescent galaxies, on the other hand, the light-weighted structural properties provide a good proxy of the mass-weighted properties.

- The measured stellar mass-size relation thus also depends strongly on the method used to determine the size (and corresponding stellar mass). Only for the r-band half-light radii estimated with the Sérsic modelling, is the mass-size relation in EAGLE in good agreement with observations for both star-forming and quiescent galaxies, albeit with a systematic offset of 0.1 dex.

- The Sérsic indices of EAGLE galaxies tend be lower than observed, due to a deficiency in bulge-like ($n \sim 4$) systems. A closer look at the surface brightness and mass density profiles shows that there is likely a deficiency in stellar mass (and hence light) at the very centres of the simulated galaxies.

- There is a lack of highly flattened objects among both the quiescent and star-forming population, likely due to the gas pressure floor and the limited resolution of the simulation. On the other hand, massive quiescent galaxies in EAGLE are not sufficiently round in shape, and appear to be more strongly triaxial than quiescent galaxies in GAMA.

Our work demonstrates that, for a fair comparison between the structural parameters of simulated and observed galaxies, it is crucial to account for the effects of M_*/L gradients within galaxies, as well as the systematic differences between various analysis techniques. This can be achieved by either deriving mass-weighted measurements from observations or, as shown here, by constructing realistic mock observations from simulations. Although computationally expensive, a realistic treatment of simulated data can truly provide a different picture of the simulated galaxy population.

Acknowledgements

(Australia), the AAO, and the participating institutions. The GAMA website is http://www.gama-survey.org/ .

Appendix

4.A   Impact of particle smoothing

As described in Section 4.2.2, the images created with SKIRT assume a truncated Gaussian profile for the spatial distribution of the stellar particles. The width of this distribution, the smoothing length, is set equal to the distance to the 64th nearest neighbour particle. On the other hand, to create the stellar mass images (Section 4.3.1.2), the stellar particles were treated as point sources.

If the effect of smoothing is large, we may expect to find less centrally concentrated light profiles, and thus lower Sérsic indices. As the size and total luminosity covary with the Sérsic index, this may also affect the obtained mass-size relation.

Given the computational cost of creating mock r-band images, we evaluate the effect of different smoothing lengths on the stellar mass images instead. We use the PY-SPHVIEWER code (Benitez-Llambay 2015) to construct stellar mass maps that include nearest neighbour smoothing. These images are created from the exact same particles used before, and thus have identical dimensions and orientation. The code then uses the 3D particle distribution to compute the distance to the 64th nearest neighbour for all particles, and provides a smoothed, projected image of 512×512 pixels. Finally, we process this image in the same fashion as described in Section 4.3.1.2 to apply PSF smoothing, degrade the pixel resolution, and add realistic noise.

We perform the Sérsic profile modelling for this set of smoothed images, and compare the resulting structural parameters to the fits without smoothing. Fig. 4.15 shows the distributions of the differences in the obtained stellar masses, half-mass radii, and Sérsic indices. All three distributions are centred around zero (median values indicated with dashed lines), and the scatter is consistent with the typical uncertainties discussed in Section 4.3.2.4. We can therefore conclude that the differences found in the structural parameters measured from the r-band and stellar mass images are not due to a difference in the applied smoothing.

4.B   Functional form of the PSF

To create realistic mock observations, we convolved the images with a PSF that takes the form of a circular Gaussian distribution (Section 4.3.1):

$$\mathrm{PSF}(\mathbf{r}) = \mathrm{PSF}(r) = \frac{1}{2\pi\sigma^2} \exp\left(\frac{-r^2}{2\sigma^2}\right), \tag{4.2}$$

where the variance σ^2 is related to the width of the distribution by FWHM $= 2\sqrt{2\ln 2} \times \sigma$.
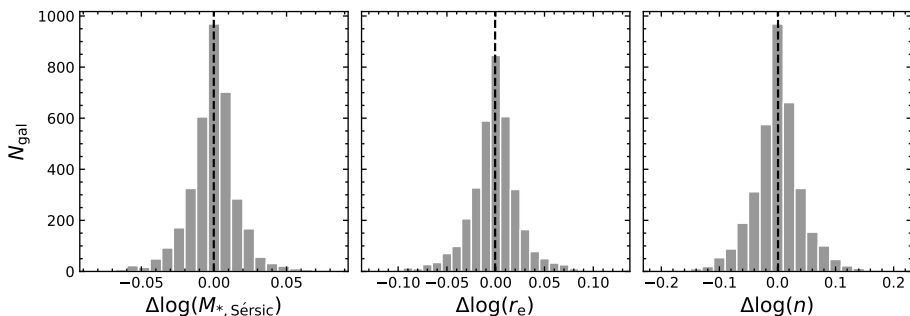
Figure 4.15: Differences between the structural parameters measured from stellar mass images with and without nearest neighbour smoothing. Median values are indicated with dashed lines. There is no systematic offset in the inferred stellar mass, half-mass radius and Sérsic index, and the scatter is consistent with the uncertainties discussed in Section 4.3.2.4. The differences between the r-band and stellar mass sizes and morphologies presented in this paper are therefore unlikely to be caused by a difference in the smoothing of the stellar particles.

However, the shape of the PSF is generally more complex than a single Gaussian distribution describes, as, in addition to a core component, there are typically extended wings. In the SDSS, the PSF has been modelled with various decompositions (see Stoughton et al. 2002), the simplest being a double Gaussian model, which was also used to perform surface brightness profile fitting to postage-stamp images of galaxies (to obtain 'model magnitudes'). This double Gaussian model is isotropic and takes three parameters:

$$\mathrm{PSF}(r) = \frac{(1-C)}{2\pi\sigma_1^2} \exp\left(\frac{-r^2}{2\sigma_1^2}\right) + \frac{C}{2\pi\sigma_2^2} \exp\left(\frac{-r^2}{2\sigma_2^2}\right), \qquad (4.3)$$

where σ_1^2 and σ_2^2 are the variances of the two components, and the constant C is the ratio of the overall amplitudes.

In Fig. 4.16, we show the radial profiles of these two different PSF shapes, using FWHM $= 1.39''$ for the single Gaussian PSF (solid black line). For the double Gaussian model (blue dashed line), we obtain typical parameters from the 'Field' catalogue described in Section 4.3.1.1: $[\sigma_1, \sigma_2, C] \approx [0.95\sigma, 2.05\sigma, 0.09]$. There is good agreement between the two PSFs at small radii, with the FWHM of the double Gaussian model being only slightly smaller (FWHM $= 1.33''$). For $r \gtrsim 3''$ the two models deviate increasingly strongly, by multiple orders of magnitude.

We quantify the difference between these two profile shapes by calculating the second moment of the distributions (see also Franx et al. 1989):

$$\langle r^2 \rangle = \frac{\iint_{\mathbb{R}^2} r^2 \, \mathrm{PSF}(\mathbf{r}) \, \mathrm{d}\mathbf{r}}{\iint_{\mathbb{R}^2} \mathrm{PSF}(\mathbf{r}) \, \mathrm{d}\mathbf{r}}, \qquad (4.4)$$
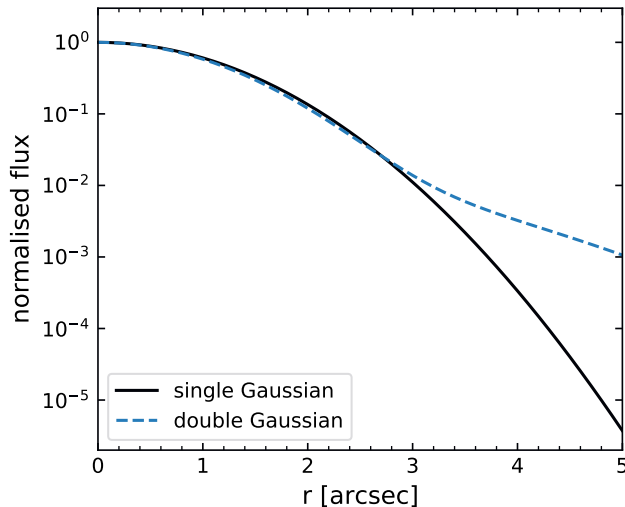
Figure 4.16: Radial profiles of the single Gaussian and double Gaussian models of the SDSS PSF. The two models are comparable at small scales, but deviate strongly in the outer wings.
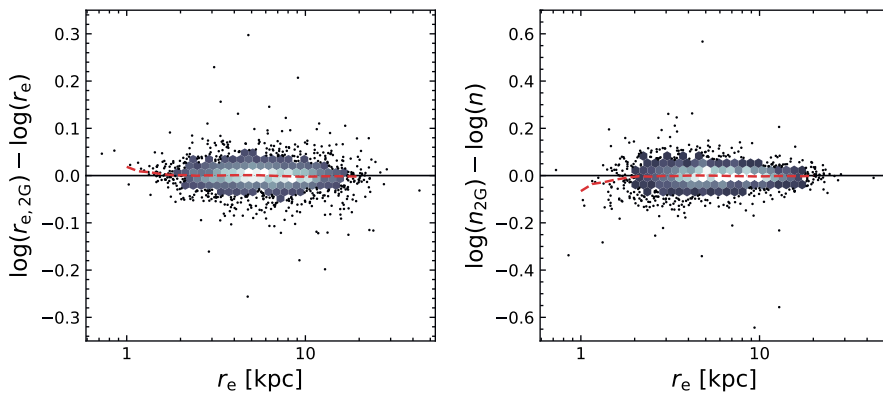
Figure 4.17: The difference in the inferred r-band size (left) and Sérsic index (right) between the images that are convolved with a single Gaussian PSF and a double Gaussian PSF. There is no systematic offset between the two sets of models (medians are indicated with a red, dashed line), and the scatter is consistent with the random uncertainties discussed in Section 4.3.2.4. Only for very small galaxies ($r_e \lesssim 2\,\text{kpc}$), where we may expect to see the largest impact of a difference in the PSF, does there appear to be a slight offset in both the size and Sérsic index, of up to $\Delta \log(r_e) \sim 0.05\,\text{dex}$ and $\Delta \log(n) \sim -0.1\,\text{dex}$. The choice of adopting a simple (single Gaussian) PSF shape thus has minimal impact on the results presented in this work.

which for the normalised, isotropic PSFs considered here reduces to

$$\langle r^2 \rangle = 2\pi \int_0^\infty r^2 \, \mathrm{PSF}(r) \, r \, \mathrm{d}r \, . \qquad (4.5)$$

For the single and double Gaussian models described previously, these moments are $\langle r^2 \rangle = 0.70 \, \mathrm{arcsec}^2$ and $\langle r^2 \rangle = 0.84 \, \mathrm{arcsec}^2$, respectively.

Whereas the difference in the FWHM is only 5%, the second moments differ by 20%, which may have a measurable effect on the constructed images and inferred structural parameters. This effect is expected to be largest for galaxies that have clumpy surface brightness profiles, as these would appear more smooth with the double Gaussian PSF, as well as for highly compact galaxies that would appear more extended.

We therefore use the optical images (Section 4.2.2) to evaluate whether the choice of the PSF model leads to systematic effects. A second set of mock r-band images is constructed in the exact same way as described in Section 4.3.1.1, except for the use of the double Gaussian model in the PSF smoothing instead of the single Gaussian model. We then run the Sérsic modelling pipeline on these images, and compute the difference in the obtained structural parameters.

We focus on the size and Sérsic index, as these are the parameters that are most likely to be affected by a change in the PSF. Fig. 4.17 shows the difference in the obtained half-light radius and Sérsic index, as a function of the half-light radius. Generally, the two sets of measurements agree very well, as there is no systematic offset and little scatter (consistent with the expected measurement uncertainties). Only at very small radii ($r_\mathrm{e} \lesssim 2 \, \mathrm{kpc}$) is there are slight difference between the two PSF models, as the use of the double Gaussian PSF leads to slightly larger sizes (up to $\Delta \log(r_\mathrm{e}) \sim 0.05 \, \mathrm{dex}$) and slightly lower Sérsic indices (up to $\Delta \log(n) \sim -0.1 \, \mathrm{dex}$).

Overall, we can therefore conclude that the shape of the PSF has a minimal influence on the inferred structural parameters, and that a simple PSF model is sufficient for measuring parametric morphologies.

4.C   Comparing Sérsic model stellar masses with aperture measurements

In Fig. 4.6, we found good agreement between the total stellar mass of the Sérsic models and the stellar mass measured within a spherical aperture of $30 \, \mathrm{kpc}$, except for the more massive ($M_* \gtrsim 10^{11} \, \mathrm{M}_\odot$) galaxies. For the comparison of the population statistics of simulated galaxies with observations, e.g. the mass-size relation or the stellar mass function, it may be of interest to evaluate which aperture best captures the stellar mass at the high mass end. Here, 'best' is defined as being as close to what is typically observed, which does not necessarily correspond to the true stellar mass of a galaxy.

To this end, we compare the total stellar mass of the best-fit Sérsic model (from fits to the mock stellar mass images) with different definitions of stellar mass available in the public EAGLE catalogues. Fig. 4.18 shows this comparison
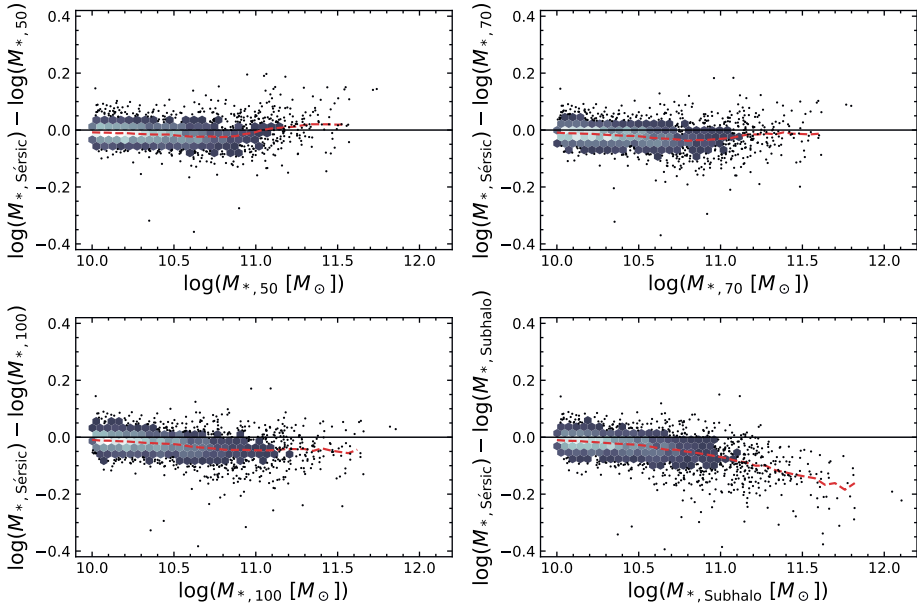
Figure 4.18: The difference between the total stellar mass of the best-fit Sérsic model and the mass of the stellar particles enclosed within a specified aperture, as a function of the aperture stellar mass. The top panels, as well as the bottom left panel, show the results for a spherical aperture of radius 50 kpc, 70 kpc, and 100 kpc. Red dashed lines indicate the running median. For comparison, the bottom right panel shows the total mass of all stellar particles belonging to the subhalo. If the aperture is too small (radius of 30 kpc, Fig. 4.6), the Sérsic model mass deviates strongly from the aperture mass at the at the high mass end. Conversely, the single Sérsic profile cannot capture all the stellar mass within the subhalo. An aperture of radius of ~ 70 kpc appears to give the best agreement between the two different measures of stellar mass, with an approximately constant offset of -0.02 dex across the entire range in stellar mass.

for stellar masses that are calculated as the sum of the stellar particle masses within spherical apertures of increasing radius (50 kpc, 70 kpc, 100 kpc). The bottom right panel compares the Sérsic model mass with the total stellar mass of the subhalo. Dashed lines show the running median in each panel.

Contrary to Fig. 4.6, where $M_{*,\mathrm{Sérsic}} > M_{*,30}$ toward high stellar mass, we find that if the aperture is too large (radius of 100 kpc, or the full subhalo), the Sérsic model significantly underestimates the total stellar mass at the high mass end. The $M_{*,70}$ mass is in best agreement with $M_{*,\mathrm{Sérsic}}$ at both lower and high stellar mass (with $M_{*,70}$ being 0.02 dex greater on average), suggesting that a spherical aperture of radius ~ 70 kpc will provide a measure that is most consistent with observations.

5 | A COMMON ORIGIN FOR THE FUNDAMENTAL PLANE OF QUIESCENT AND STAR-FORMING GALAXIES IN THE EAGLE SIMULATIONS

ABSTRACT

We use the EAGLE cosmological simulations to perform a comprehensive and systematic analysis of the $z = 0.1$ Fundamental Plane (FP), the tight relation between galaxy size, mass and velocity dispersion. We first measure the total mass and velocity dispersion (including both random and rotational motions) within the effective radius to show that simulated galaxies obey a total mass FP that is very close to the virial relation ($< 10\%$ deviation), indicating that the effects of non-homology are weak. When we instead use the stellar mass, we find a strong deviation from the virial plane, which is driven by variations in the dark matter content. The dark matter fraction is a smooth function of the size and stellar mass, and thereby sets the coefficients of the stellar mass FP without substantially increasing the scatter. Hence, both star-forming and quiescent galaxies obey the same FP, with equally low scatter (0.02 dex). We employ simulations with a variable stellar initial mass function (IMF) to show that IMF variations have a modest additional effect on this FP. Moreover, when we use luminosity-weighted mock observations of the size and spatially-integrated velocity dispersion, the inferred FP changes only slightly. However, the scatter increases significantly, due to the luminosity-weighting and line-of-sight projection of the velocity dispersions, and measurement uncertainties on the half-light radii. Importantly, we find significant differences between the simulated FP and observations, which likely reflects a systematic difference in the stellar mass distributions. Therefore, we suggest the stellar mass FP offers a simple test for cosmological simulations, requiring minimal post-processing of simulation data.

Anna de Graaff, Marijn Franx, Eric Bell, Rachel Bezanson,
Matthieu Schaller, Joop Schaye, Arjen van der Wel
Resubmitted to the Monthly Notices of the Royal Astronomical Society

5.1   Introduction

Quiescent galaxies have been found to obey a tight, planar scaling relation between the stellar velocity dispersion, size and surface brightness, which is known as the luminosity Fundamental Plane (FP; e.g., Djorgovski & Davis 1987; Dressler et al. 1987; Jorgensen et al. 1996). On the other hand, the star-forming population has been shown to follow a linear relation between the luminosity and kinematics, referred to as the Tully-Fisher (TF) relation (Tully & Fisher 1977). More recent work has demonstrated that the two galaxy populations can be reconciled within the framework of one scaling relation (e.g., Cortese et al. 2014; Bezanson et al. 2015; Aquino-Ortíz et al. 2020; de Graaff et al. 2021), by either modifying the TF relation (Cortese et al. 2014; Aquino-Ortíz et al. 2018) or the FP (e.g., Zaritsky et al. 2008; Hyde & Bernardi 2009; Aquino-Ortíz et al. 2020). However, these studies are largely empirically driven, aiming to construct a dynamical scaling relation with minimal scatter. A firmer theoretical footing is imperative to gain a better understanding of why different types of galaxies may lie on a single dynamical scaling relation.

5.1.1   Interpreting the luminosity FP

The properties of the FP of quiescent galaxies have typically been interpreted in terms of the dynamical mass-to-light ratio (M_{dyn}/L; as first suggested by Faber et al. 1987). The zero point of the FP is directly proportional to M_{dyn}/L, and the redshift evolution of the FP therefore directly traces the evolution in the M_{dyn}/L with cosmic time (e.g., van Dokkum & Franx 1996). Although at a fixed redshift the scatter about the FP is small ($\lesssim 0.1\,\mathrm{dex}$), it cannot be explained by measurement uncertainties alone (e.g., Jorgensen et al. 1996; Forbes et al. 1998). Rather, the scatter in the zero point can also be linked to variations in the M_{dyn}/L, and correlations between the offsets of galaxies from the FP and various stellar population and structural properties hence have provided insight into the formation histories of early-type galaxies (e.g., Gargiulo et al. 2009; Graves et al. 2009).

Furthermore, the tilt of the plane can also be interpreted by the M_{dyn}/L (e.g., Bender et al. 1992; Trujillo et al. 2004; Cappellari et al. 2006). Fundamentally, a tight scaling relation between the velocity dispersion (σ), effective radius (r_{e}) and surface brightness within r_{e} (I_{e}) is to be expected for systems that are in virial equilibrium. Interestingly, however, the FP is tilted with respect to this simple virial prediction:

$$r_{\mathrm{e}} \propto \sigma^a I_{\mathrm{e}}^b \tag{5.1}$$

where the coefficients a and b describe the tilt of the plane, which in the case of virial equilibrium would equal $a = 2$ and $b = -1$ for a homologous set of galaxies. In practice, values have been found to be in the range $a \approx [0.7, 1.5]$ and $b \approx [-0.9, -0.6]$, depending on the passband and fitting method used, as well as the redshift (e.g., Jorgensen et al. 1996; La Barbera et al. 2010a; Hyde & Bernardi 2009; Jørgensen & Chiboucas 2013). By rewriting Eq. 5.1 in terms of M_{dyn}/L

(using $M_{\mathrm{dyn}} \propto r_{\mathrm{e}}\sigma^2$ and $L = 2\pi r_{\mathrm{e}}^2 I_{\mathrm{e}}$; Cappellari et al. 2006),

$$\frac{M_{\mathrm{dyn}}}{L} \propto r_{\mathrm{e}}^{-2-(2+a)/(2b)} M_{\mathrm{dyn}}^{1+(a/2b)} \,, \tag{5.2}$$

which is constant if $a = 2$ and $b = -1$, it becomes apparent that the observed tilt of the FP reflects a correlation between M_{dyn}/L and M_{dyn} or r_{e}, the latter of which is often assumed to be subdominant.

Although these various studies of the FP have led to a consensus on the existence of a rotation of the FP with respect to the virial plane, the origins for this rotation and precise values of the tilt have been debated extensively for the past decades, without reaching a consensus. In addition to being highly sensitive to measurement choices and uncertainties, the tilt depends strongly on the chosen fitting method and sample selection biases (see, e.g., Hyde & Bernardi 2009; Magoulas et al. 2012), which leads to large uncertainties especially toward higher redshifts (e.g., Holden et al. 2010; Jørgensen & Chiboucas 2013; de Graaff et al. 2021). Nevertheless, multiple causes have been proposed to explain the observed deviation from the scalar virial theorem, which can be best understood by decomposing M_{dyn}/L and assessing how the different components scale with M_{dyn} (Hyde & Bernardi 2009):

$$\frac{M_{\mathrm{dyn}}}{L} = \frac{M_{\mathrm{dyn}}}{M_{\mathrm{tot}}} \frac{M_{\mathrm{tot}}}{M_*} \frac{M_*}{L} \,, \tag{5.3}$$

where M_* and M_{tot} are the stellar and total (dark matter and baryonic) mass, respectively.

First, the departure from the expected virial plane may reflect the fact that the assumption of homology is inaccurate, captured by the ratio $M_{\mathrm{dyn}}/M_{\mathrm{tot}} \neq 1$. Quantified using the Sérsic index (Sersic 1968), the effects of non-homology were shown by some studies to play a key role (Bender et al. 1992; Graham & Colless 1997; Prugniel & Simien 1997; Trujillo et al. 2004; Desmond & Wechsler 2017). However, others have found more modest or negligible contribution arising from variation in the galaxy structure, based on dynamical modelling or strong lensing results (e.g., Cappellari et al. 2006; Bolton et al. 2007, 2008; D'Eugenio et al. 2021).

Second, broader agreement has been reached on the magnitude of the contribution to the tilt from the mass dependence of M_{tot}/M_*, which depends on the 'dark' mass within galaxies. Crucially, this ratio is not simply the dark matter fraction (assuming negligible gas mass), but also includes missing mass due to uncertainties in the stellar initial mass function (IMF), as stellar masses that are estimated from spectral energy distribution (SED) modelling often rely on the assumption of a universal IMF and therefore carry a systematic uncertainty (for a review on SED modelling, see Conroy 2013). Although this dark component is expected to contribute significantly to the tilt of the FP ($\sim 50\%$; Renzini & Ciotti 1993; Hyde & Bernardi 2009; Graves & Faber 2010), distinguishing between the effects of variations in the IMF versus the dark matter fraction is challenging. Recent observational work based on simple dynamical models has suggested that IMF variations can fully explain the observed relation between M_{dyn}/M_* and M_{dyn}, by allowing for a non-universal IMF that can vary between galaxies as well as radially

within galaxies (Bernardi et al. 2018; Marsden et al. 2022). However, others have shown that this would be difficult to reconcile with the observed correlations of stellar population properties throughout the FP (Graves & Faber 2010), or have found evidence for variations in both the IMF and dark matter content in galaxies (e.g., Cappellari et al. 2013b,a).

The third component arises from variations in the stellar population properties across galaxies, i.e., variations in M_*/L. By evaluating the tilt of the FP in different passbands or by explicitly estimating M_*/L, the effects of M_*/L variations have been shown to be insufficient to fully explain the observed tilt, but may account for up to half of this tilt (e.g., La Barbera et al. 2008; Hyde & Bernardi 2009; Graves & Faber 2010; Bernardi et al. 2020; D'Eugenio et al. 2021).

5.1.2 The stellar mass FP

The effects of M_*/L variations across and along the FP can be addressed by explicitly estimating M_*/L independently, by fitting the spectral or photometric SEDs with stellar population models. We can then gain insight into M_{dyn}/M_* alone, a quantity that is of great interest, as it depends on the formation and structural evolution of galaxies, such as the effects of mergers (e.g., Hopkins et al. 2008). Zaritsky et al. (2006, 2008) first proposed the fundamental manifold, a 3D scaling relation within the 4D parameter space of the galaxy kinematics, size, surface brightness and M_*/L. Hyde & Bernardi (2009) showed that similar results can be achieved by modifying the FP, replacing the surface brightness by the stellar mass surface density (Σ_*):

$$r_{\mathrm{e}} \propto \sigma^{\alpha} \Sigma_*^{\beta} \,, \tag{5.4}$$

which is referred to as the stellar mass FP. Including M_*/L also results in a lower intrinsic scatter about the scaling relation (i.e., the scatter after accounting for measurement uncertainties) than the standard luminosity FP.

Importantly, Zaritsky et al. (2008) showed that this framework, which up to then had focused on dynamically-hot spheroids, can be extended to disc-like structures as well, if the dynamical measurement (σ) explicitly includes galaxy rotation in addition to the random motions of stars. Later work demonstrated that both star-forming and quiescent galaxies follow the same stellar mass FP, with nearly identical tilt, zero point and scatter (Bezanson et al. 2015; Aquino-Ortíz et al. 2020), and that this result holds out to $z \sim 1$ with minimal evolution in the FP (de Graaff et al. 2020, 2021).

These results appear to be at odds with the observation that star-forming galaxies obey the TF relation, which is explicitly independent of surface brightness or another third parameter (e.g., Zwaan et al. 1995; Courteau & Rix 1999; Meyer et al. 2008; Lelli et al. 2019). Furthermore, it casts doubt on earlier theoretical studies, which suggested that the dissipation of gas in galaxies plays a critical role in shaping the FP: using simulations of merging galaxies, dissipational mergers were shown to give rise to the observed tilt of the FP, with the tilt of the FP being preserved under further dissipationless mergers (Boylan-Kolchin et al. 2006; Robertson et al. 2006; Hopkins et al. 2008). As a result, Hopkins et al. (2008)

showed explicitly that discs and spheroids have a different dependence of M_{dyn}/M_* on mass.

5.1.3   The FP in cosmological simulations

Cosmological simulations may offer new insight into the origins of the FP, as, unlike the simulations of galaxy mergers, they do not a priori assume a formation channel for the FP. These large simulations have been shown to produce a wide diversity in galaxy morphologies and kinematic structures (e.g., Snyder et al. 2015; Correa et al. 2017; Thob et al. 2019), and also reproduce key observed relations such as the galaxy stellar mass function and stellar mass-size relation (e.g., Schaye et al. 2015; Genel et al. 2018; Davé et al. 2019), although simulations are typically calibrated to achieve these latter goals.

However, the galaxy structure and dynamics are not 'tuned' explicitly: the FP therefore poses both an interesting test of the realism of a simulation, and an opportunity to gain understanding of the drivers behind the relation itself. Focusing solely on early-type galaxies, different studies have shown that simulations such as Illustris, Illustris-TNG, Horizon-AGN and EAGLE form a FP that approximately resembles observations, but with significant variation in the measured tilt and scatter (Rosito et al. 2019b, 2021; D'Onofrio et al. 2020; Lu et al. 2020). Additionally, D'Onofrio et al. (2020) showed that galaxies follow a complex trajectory through the parameter space of the FP, and suggest that the low-redshift FP arises from a combination of galaxy mergers and the passive ageing of galaxies. On the other hand, Rosito et al. (2021) used the Horizon-AGN and Horizon-noAGN simulations to show that black hole feedback is a critical factor to reproduce the observed FP.

Taking a more holistic approach, Ferrero et al. (2021) evaluated the relation between the circular velocity, stellar mass and size for dispersion-dominated quiescent galaxies and rotation-dominated star-forming galaxies in the EAGLE and Illustris-TNG simulations. They suggest that, as a consequence of the stellar-halo mass relation, by which galaxies of fixed M_* occupy a narrow range in halo mass, galaxy size becomes the only differentiating parameter. Star-forming discs are larger than quiescent spheroids at fixed M_*, and therefore encompass relatively more dark matter within the effective radius. The TF and FP relations are therefore suggested to arise solely from variations in the dark matter fraction, with the TF relation being independent of surface brightness due to the independence of the circular velocity on size at large enough radii.

Although an intriguing result, it omits the fact that the observed structures of star-forming and quiescent differ not only in size, but also in morphology. As a result, the shape of the gravitational potential may be expected to vary as a function of galaxy type, leading to the aforementioned effects of non-homology on the FP. Furthermore, observational biases, due to M_*/L gradients in galaxies and differences in measurement methods, have been shown to have a significant effect on the obtained galaxy scaling relations and are important to take into account when comparing simulations and observations (e.g., Price et al. 2017; Bottrell et al. 2017a,b; van de Sande et al. 2019; de Graaff et al. 2022).

In this paper, we aim to assess the different effects of non-homology, the dark

matter content and observational uncertainties on the tilt and scatter of the stellar mass FP for both quiescent and star-forming galaxies. By using the EAGLE cosmological hydrodynamical simulations (Schaye et al. 2015), we systematically introduce one of these components at a time, and evaluate whether these results differ for quiescent and star-forming galaxies. We build on the mock observations and measurements presented in de Graaff et al. (2022) to arrive at a FP that is as close as possible to the observed FP, and show how selection biases affect the measurement and interpretation of the FP.

The simulations used and the different definitions of galaxy size, mass and velocity dispersion are described in Section 5.2. In Section 5.3 we present the simulated FP, and discuss the effects of non-homology and variations in the dark matter fractions. We introduce observational effects, measurement and selection biases in Section 5.4, where we demonstrate how the tilt of the FP is sensitive to these different effects. Moreover, we explore the possible additional complication of a non-universal IMF. We discuss these results in Section 5.5 and show how the FP and TF relation may be reconciled. Our main results are summarised in Section 5.6.

5.2   Data and methods

5.2.1   EAGLE simulations

The EAGLE simulations are a set of cosmological smoothed particle hydrodynamics (SPH) simulations (Schaye et al. 2015; Crain et al. 2015). These simulations all assume a flat ΛCDM cosmology with cosmological parameters from the Planck Collaboration et al. (2014) ($\Omega_{\rm m} = 0.307$, $\Omega_{\rm b} = 0.0482$ and $H_0 = 67.77\,{\rm km\,s^{-1}\,Mpc^{-1}}$), but vary in the volume, resolution and subgrid model used. In this work, we will focus mainly on the reference model with a volume of 100^3 comoving Mpc^3 (cMpc; L0100N1504), which has a mass resolution of $m_{\rm DM} = 9.7 \times 10^6\,{\rm M_\odot}$ and $m_{\rm b} = 1.81 \times 10^6\,{\rm M_\odot}$ for the dark matter particles and initial mass of the gas particles, respectively. With the Plummer-equivalent gravitational softening scale of $\epsilon = 0.70$ proper kpc at $z < 2.8$, this amounts to an effective spatial resolution of ≈ 2 proper kpc. The reference model assumes a Chabrier IMF (Chabrier 2003) for the star formation prescription, which together with the other subgrid prescriptions was calibrated to reproduce the $z = 0$ stellar mass function and stellar mass-size relation.

To assess the numerical convergence of our results, we use the smaller simulation of 25^3 cMpc^3 for the recalibrated model (L0025N0752), which has a resolution that is 8 times higher. Furthermore, to examine the effects of a non-universal IMF, we use the simulations by Barber et al. (2018). These simulations implement a variable IMF into the reference EAGLE model, by allowing either the low- or high-mass end of the Kroupa double power law IMF (i.e., above or below $0.5\,{\rm M_\odot}$; Kroupa 2001) to vary according to the pressure of the local interstellar medium. The models were calibrated to reproduce the scaling relation between the excess stellar mass-to-light ratio and stellar velocity dispersion that has been observed for

early-type galaxies in the local Universe (Cappellari et al. 2013a), while simultaneously matching other key observables such as the K-band luminosity function and the relation between the half-light radius and luminosity. The resulting bottom-heavy (LoML0050N0752) and top-heavy (HiML0050N0752) models were run in 50^3 cMpc^3 volumes, which can be readily compared with the reference run of the same volume.

We will focus on galaxies at $z = 0.1$, for which mock images of the light distributions that include realistic dust attenuation, noise and seeing are available from Trayford et al. (2017) and de Graaff et al. (2022). Throughout, galaxies are defined in the usual way, as the self-bound substructures that are identified within haloes by the SUBFIND algorithm (Springel et al. 2001; Dolag et al. 2009). This mechanism also allows for a separation of central and satellite galaxies, which is used in Section 5.3. Moreover, we distinguish between star-forming and quiescent galaxies based on the specific star formation rate (sSFR) measured within a spherical aperture of radius 30 proper kpc centred around the potential minimum (obtained from the online public database; McAlpine et al. 2016): quiescence is defined as sSFR $< 10^{-11}$ yr^{-1}. Lastly, in what follows all length units will be quoted as proper lengths unless explicitly noted otherwise.

5.2.2   Galaxy sizes and masses

As discussed in de Graaff et al. (2022), the sizes of galaxies depend strongly on whether these are measured from the stellar mass or optical light distributions. A secondary effect is the measurement technique used, i.e., whether quantities are measured with a growth curve method or by using parametric models. The mass that is enclosed within the effective radius then changes correspondingly.

Scaling relations, such as the FP, may be expected to be sensitive to these differences. To examine to what extent this makes a difference on the obtained FP, we will use multiple definitions of galaxy size and mass throughout the paper:

- $r_{\mathrm{e,3D}}$: the radius that encloses half of the stellar mass within a spherical aperture of radius 100 kpc centered around the potential minimum (see also Furlong et al. 2017). We consequently define the stellar mass within a spherical aperture of this radius as $M_*(< r_{\mathrm{e,3D}})$, and the total mass within the same aperture $M_{\mathrm{tot}}(< r_{\mathrm{e,3D}})$. The total mass is the sum of the dark matter, stellar, gas and black hole particle masses.

- $r_{\mathrm{e,*}}$: half-mass semi-major axis obtained from Sérsic profile fitting to projected images (along the z-axis of the simulation box) of the stellar mass distributions from de Graaff et al. (2022). $M_*(< r_{\mathrm{e,*}})$ is half of the stellar mass of the integrated, best-fit Sérsic model.

- $r_{\mathrm{e,r}}$: half-light semi-major axis obtained from Sérsic profile fitting to images of the optical light distributions in the r-band (again, using the random projection along the z-axis of the simulation box; Trayford et al. 2017; de Graaff et al. 2022). We obtain stellar masses $M_*(< r_{\mathrm{e,r}})$ by multiplying the mass-to-light ratio within a spherical aperture of radius 30 kpc (M_*/L_r) by half of the luminosity of the best-fit Sérsic profile.
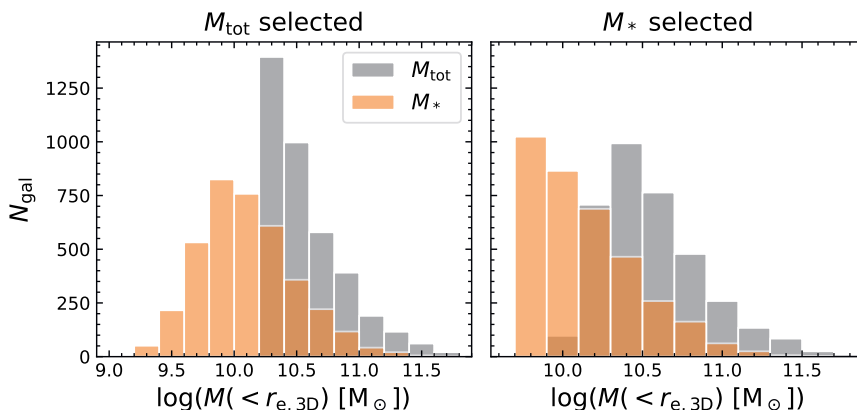
Figure 5.1: Distributions of the total mass and stellar mass enclosed within the 3D effective radius, for the samples selected by the total mass (left) and stellar mass (right). Total masses include all particle masses, i.e., dark matter, stellar, gas and black hole particles.

5.2.3   Sample selection

A key goal of this work is to quantify the effects of sample selection on the obtained FP. In large imaging surveys, the selection of galaxies is limited by total flux and/or surface brightness, depending on the apparent size and point spread function. For a chosen maximum distance, the flux-limited samples can then be used to construct sub-samples that form an accurate representation of the galaxy stellar mass function, down to a specified stellar mass limit.

Although cosmological simulations are by construction complete in mass, low-mass galaxies in the simulation are affected by the limited resolution of the simulation (resulting in, e.g., unreliable sizes; Ludlow et al. 2019, 2021). We therefore impose a selection on the galaxy mass, and construct two samples that are complete in (i) total mass and (ii) stellar mass. First, we calculate the total mass enclosed within $r_{e,3D}$, and select galaxies for which $M_{tot}(< r_{e,3D}) > 10^{10.2}\,M_{\odot}$ and that also contain $> 10^3$ stellar particles (96% of the sample contain $> 1 \times 10^4$ stellar particles). For the $100^3\,cMpc^3$ box (Section 5.2.1), this results in a sample of 3758 galaxies.

The second sample follows the selection of de Graaff et al. (2022): this sample is selected by requiring the aperture stellar mass $M_* > 10^{10}\,M_{\odot}$, and thus effectively selected by $M_*(< r_{e,3D}) > 10^{9.7}\,M_{\odot}$, with additional criteria imposed on the quality of the Sérsic profile fits. Namely, as discussed in detail in de Graaff et al. (2022), we require that the fit has converged within the parameter boundaries (removing 35 objects) and pass our visual inspection (by not showing strong residual features; removing 29 objects). We refer the reader to this previous work for further examples of the fitting procedure and discussion of the obtained sizes and morphologies. The stellar mass-selected sample consists of 3624 galaxies, of

which 3560 are flagged as having good Sérsic profile fits. The mass distributions for these two different selections are shown in Fig. 5.1. The two samples were constructed to contain an approximately equal number of objects, and to have a significant overlap, with 3183 galaxies appearing in both samples ($\approx 90\%$).

5.2.4   Velocity dispersion measurements

The third critical component that enters the FP is the velocity dispersion, which reflects the depth and shape of the gravitational potential. Observationally, this quantity is traced by the light emitted by stars, as their motion along the line of sight leads to a broadening of stellar absorption lines. Importantly, this motion can come from both the disordered motion and the ordered rotation of the stars (see also Section 5.1.2).

To systematically assess the impact of these different observational effects, we begin by measuring the velocity dispersion within spherical apertures of radius $r_{\mathrm{e,3D}}$. Following McAlpine et al. (2016), the kinetic energy of a collection of particles is calculated as

$$K = \frac{1}{2} \sum_i m_i (\mathbf{v} - \mathbf{v}_{\mathrm{pec}})^2 \,, \tag{5.5}$$

where m and $\mathbf{v}$ are the mass and velocity of the particle, respectively, and $\mathbf{v}_{\mathrm{pec}}$ is the peculiar velocity of the galaxy, which we calculate as the mass-weighted average velocity of the stellar particles within an aperture of $30\,\mathrm{kpc}$ centred around the potential minimum. We calculate the velocity dispersion within a radius r as the mean-square speed (thereby including both the random motion and rotation of the particles, see Binney & Tremaine 1987, Chapter 4.8.3), which depends on the kinetic energy and mass of the particles enclosed within the same radius:

$$\sigma(< r) = \sqrt{\frac{2K(< r)}{M(< r)}} \,. \tag{5.6}$$

We calculate two versions of this velocity dispersion: the stellar velocity dispersion $\sigma_*(< r_{\mathrm{e,3D}})$ that is based on the kinetic energy and mass of the stellar particles within $r_{\mathrm{e,3D}}$, and the total velocity dispersion $\sigma_{\mathrm{tot}}(< r_{\mathrm{e,3D}})$, which includes the dark matter, stellar and gas particles in Eq. 5.5 and 5.6. These two different velocity dispersions are compared in Fig. 5.2, which shows that the total velocity dispersion is systematically larger than the stellar velocity dispersion.

Next, we apply a measurement that is in better agreement with observational methods. We use the size, axis ratio and position angle from the best-fit Sérsic profile to construct an elliptic cylindrical aperture, which is centred around the potential minimum and has a length along the z-axis of the simulation box of $\pm 50\,\mathrm{kpc}$. Selecting all particles within the aperture, we obtain the spatially-integrated line-of-sight velocity dispersion (σ_{los}; i.e., along the z-axis) by first calculating the weighted mean

$$\langle v_{\mathrm{los}} \rangle = \frac{\sum_i w_i v_{\mathrm{z},i}}{\sum_i w_i} \,, \tag{5.7}$$
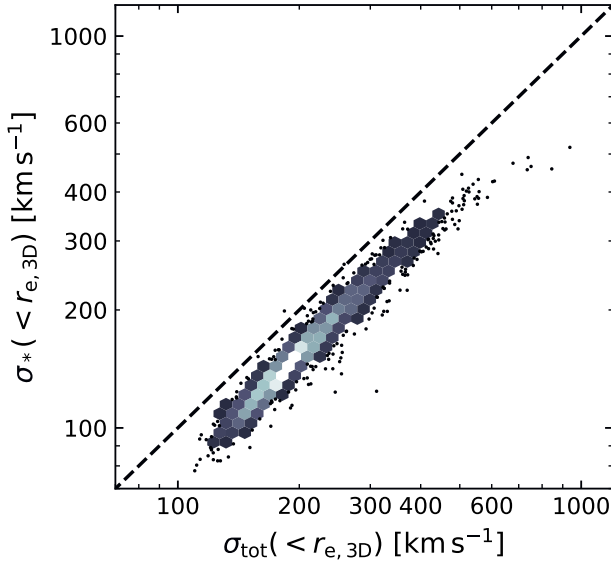
Figure 5.2: Comparison between the velocity dispersion of the stellar particles within $r_{e,3D}$ and the velocity dispersion of the stellar, dark matter, and gas particles within the same aperture (Eq. 5.5 and 5.6). Individual data points are shown for sparsely sampled areas of the figure (created using DENSITYPLOT; Krawczyk & Peters 2014). The total velocity dispersion is greater than the stellar velocity dispersion, suggesting that the dark matter particles are dynamically hotter.

followed by

$$\sigma_{los}^2 = \frac{\sum_i w_i (v_{z,i} - \langle v_{los} \rangle)^2}{\sum_i w_i} \,, \qquad (5.8)$$

where v_z is the velocity of the particle along the line of sight, and w_i is the weight. This spatially-integrated measurement of the velocity dispersion therefore also includes both the rotational and random motions of the particles along the line-of-sight direction. Using the Sérsic profile fits to the stellar mass images (see Section 5.2.2) for the apertures and the current mass of the stellar particles as weights, we obtain $\sigma_*(< r_{e,*})$. Similarly, using the Sérsic profile fits to the optical light and weighting by the luminosities of the particles, we obtain $\sigma_*(< r_{e,r})$. Here, we have chosen to use the r-band Sérsic profile fits to construct the apertures, but g-band luminosities for the weighting of the velocity dispersions, to mimic observations where the more prominent absorption lines are around ~ 5000 Å. We show in Appendix 5.C that the mismatch between the waveband chosen for these two different measurements has only a small effect.

We compare the three different stellar velocity dispersions in Fig. 5.3, which shows the luminosity-weighted and stellar mass-weighted velocity dispersion along the line of sight as a function of the 3D velocity dispersion calculated with Eq. 5.6.
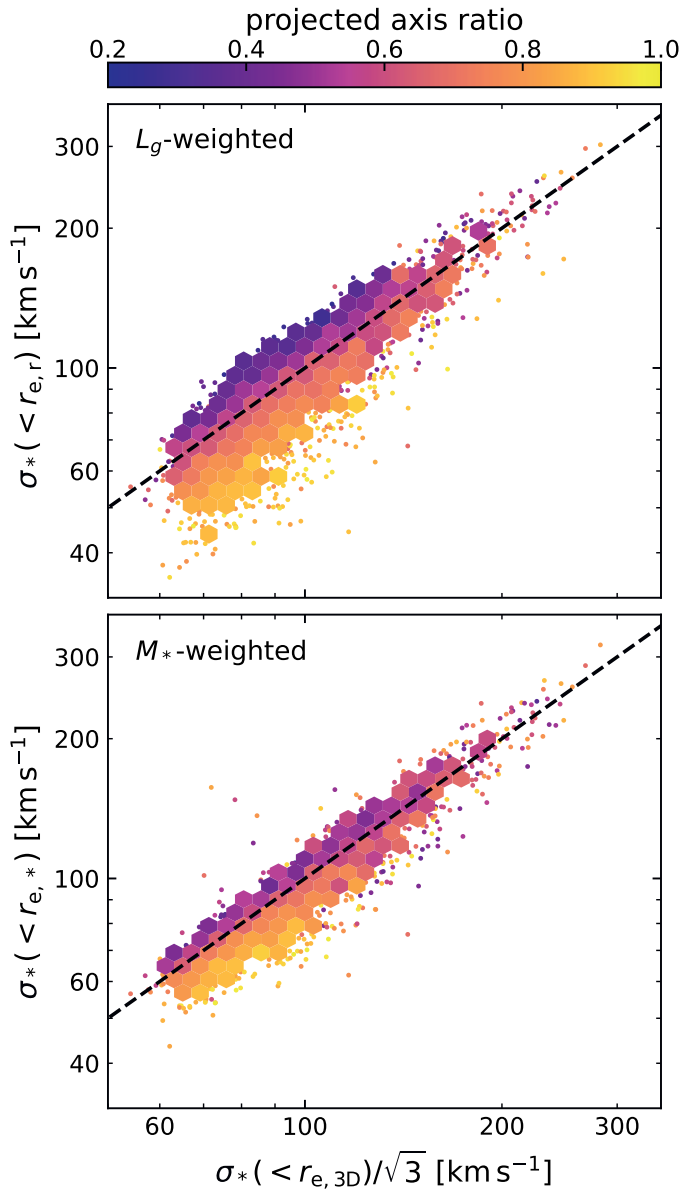
Figure 5.3: The spatially-integrated line-of-sight stellar velocity dispersions measured in elliptic cylindrical apertures versus the 3D stellar velocity dispersion, for the luminosity-weighted (top) and stellar mass-weighted (bottom) measures. The colour coding reflects the median projected axis ratio in each bin, demonstrating that the scatter about the unit slope can be attributed mainly to projection effects and the degree to which the rotational motion along the line of sight contributes to the measured velocity dispersion.

Although there is good agreement between the different measures, there is a large scatter that is particularly strong toward low σ_*. By colour coding the data with the projected axis ratios of the Sérsic models, it becomes apparent that this is due to projection effects: galaxies that are near face-on are observed to have a significantly lower velocity dispersion than edge-on galaxies, which is to be expected for oblate systems that are strongly rotating.

5.3   The simulated Fundamental Plane

As discussed in Section 5.1, the tilt of the FP reflects the deviation from the simple prediction of virial equilibrium for homologous systems: $M(r) \propto r\,\sigma^2$. In this Section, we present the theoretical perspective on the cause of these deviations, by evaluating the effects of structural non-homology and variations in the mass compositions of galaxies on the FP. Observational effects and selection biases will then be discussed in Section 5.4.

5.3.1   Dynamical tracers of the total galaxy mass

In Section 5.2.4 we defined two different tracers of the galaxy dynamics: a stellar velocity dispersion, and the total velocity dispersion. Clearly, σ_{tot} is a quantity that cannot be measured observationally, however, it may seem a natural choice when the aim is to recover M_{tot}. We begin by examining the planar relation

$$r_{\mathrm{e,3D}} \propto \sigma_{\mathrm{tot}}^{\alpha}(< r_{\mathrm{e,3D}})\, \Sigma_{\mathrm{tot}}^{\beta}\,, \tag{5.9}$$

where $\Sigma_{\mathrm{tot}} = M_{\mathrm{tot}}(< r_{\mathrm{e,3D}})/(\pi r_{\mathrm{e,3D}}^2)$, and the coefficients $\alpha = 2$ and $\beta = -1$ for homologous systems in virial equilibrium.

To aid in the visualisation of this 3D relation, we form narrow bins in $\log(M_{\mathrm{tot}}(< r_{\mathrm{e,3D}}))$, and show the relation between $\sigma_{\mathrm{tot}}(< r_{\mathrm{e,3D}})$ and $r_{\mathrm{e,3D}}$ for galaxies in the M_{tot}-selected sample in Fig. 5.4. Here, the sample is divided into star-forming (blue) and quiescent (red) galaxies, and shaded areas show the region in parameter space that is likely to be affected by the limited resolution of the simulation. We fit linear relations with a fixed slope ($m = -0.5$) to the data in each panel, which represent lines of constant M_{dyn} (dashed lines) and thus the tilt of the virial plane.

There is a tight sequence around these relations in all mass bins, except for the very highest mass bin that spans a broad range in mass (up to $10^{12.25}\,\mathrm{M_\odot}$). Moreover, toward higher total mass, the sequence itself shifts toward larger sizes and higher velocity dispersions. The assumption of virial equilibrium and homology, which would imply $M_{\mathrm{dyn}}/M_{\mathrm{tot}} = 1$, therefore seems to be reasonable. However, toward smaller sizes, there appears to be a systematic offset with respect to the dashed lines.

We fit the coefficients α and β by minimising the sum of the absolute orthogonal distances to the plane:

$$\Delta_{\mathrm{FP}} = \frac{|\log(r_{\mathrm{e,3D}}) - \alpha\,\log(\sigma_{\mathrm{tot}}(< r_{\mathrm{e,3D}})) - \beta\,\log(\Sigma_{\mathrm{tot}}) - \gamma|}{\sqrt{1 + \alpha^2 + \beta^2}}\,, \tag{5.10}$$
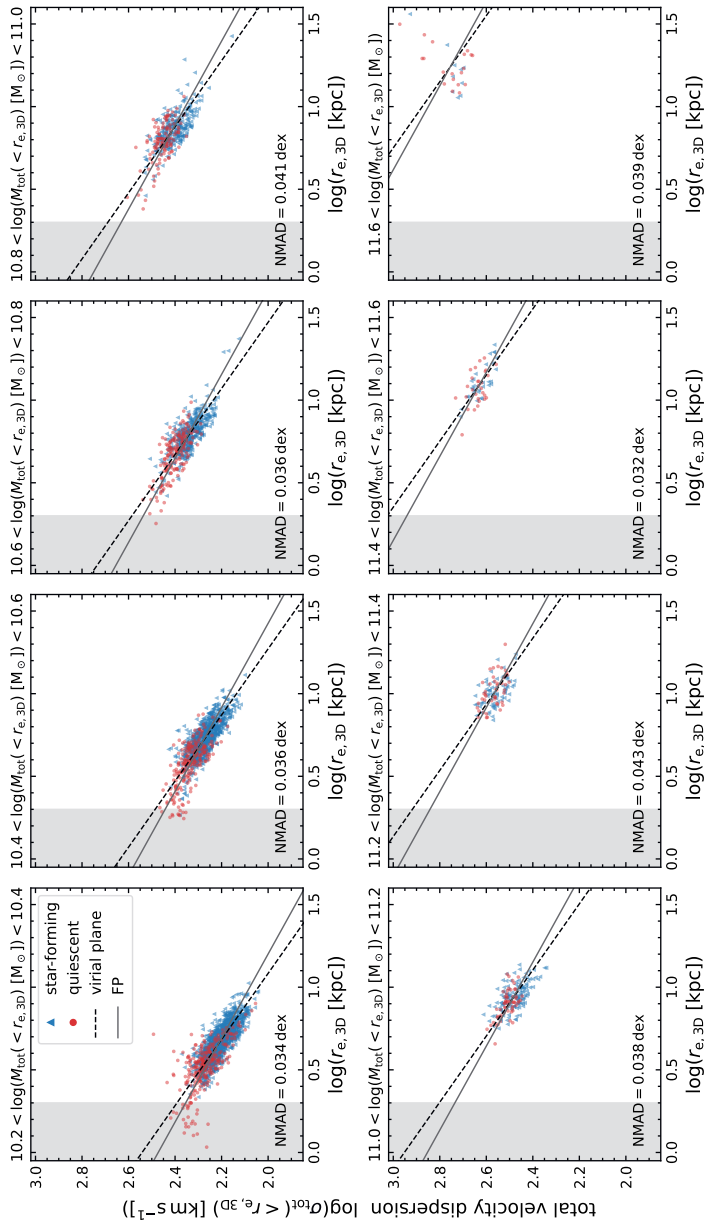
Figure 5.4: Velocity dispersion of all particles (dark matter, stellar, gas; see Section 5.2.4) as a function of the 3D half-mass radius, in bins of the total mass $M_{\rm tot}(<r_{\rm e,3D})$. The sample is divided into star-forming (blue) and quiescent (red) galaxies by the sSFR; the grey shaded area indicates $r_{\rm e,3D} < 2\,{\rm kpc}$, where measurements are likely to be affected by the resolution of the simulation. Dashed lines, with the intercept fit to the data in each panel, show the slope of the virial plane that galaxies are expected to follow if they are both in virial equilibrium and homologous. The best-fit total mass FP (solid lines; Table 5.1) deviates slightly from the virial plane due to the effects of non-homology; the orthogonal scatter about this line is printed in each panel.

Table 5.1: Best-fit coefficients of the total mass FP and stellar mass FP, and the orthogonal scatter about the planes, for different dynamical tracers and selection criteria. Galaxies for which $r_{e,3D} < 2\,$kpc are excluded from the fits.

Relation	Sample selection	α	β	γ	NMAD
$\log(r_{e,3D}) = \alpha \log(\sigma_{tot}(<r_{e,3D})) + \beta \log(\Sigma_{tot}) + \gamma$	M_{tot}	1.700 ± 0.010	-0.834 ± 0.004	4.40 ± 0.03	0.0176 ± 0.0003
	M_{tot} & quiescent	1.612 ± 0.018	-0.747 ± 0.011	3.79 ± 0.10	0.0184 ± 0.0009
	M_{tot} & star-forming	1.776 ± 0.015	-0.865 ± 0.006	4.49 ± 0.04	0.0158 ± 0.0005
$\log(r_{e,3D}) = \alpha \log(\sigma_*(<r_{e,3D})) + \beta \log(\Sigma_{tot}) + \gamma$	M_{tot}	1.799 ± 0.008	-0.940 ± 0.005	5.29 ± 0.03	0.0134 ± 0.0002
	M_{tot} & quiescent	1.769 ± 0.012	-0.885 ± 0.007	4.86 ± 0.06	0.0122 ± 0.0004
	M_{tot} & star-forming	1.848 ± 0.012	-0.949 ± 0.006	5.28 ± 0.03	0.0121 ± 0.0003
$\log(r_{e,3D}) = \alpha \log(\sigma_*(<r_{e,3D})) + \beta \log(\Sigma_{tot}) + \gamma$	M_*	1.829 ± 0.009	-0.953 ± 0.005	5.35 ± 0.04	0.0137 ± 0.0003
	M_* & quiescent	1.790 ± 0.011	-0.882 ± 0.006	4.80 ± 0.05	0.0124 ± 0.0004
	M_* & star-forming	1.871 ± 0.009	-0.963 ± 0.006	5.36 ± 0.04	0.0123 ± 0.0003
$\log(r_{e,3D}) = \alpha \log(\sigma_*(<r_{e,3D})) + \beta \log(\Sigma_*) + \gamma$	M_*	1.559 ± 0.008	-0.562 ± 0.003	2.26 ± 0.02	0.0186 ± 0.0004
	M_* & quiescent	1.514 ± 0.010	-0.564 ± 0.005	2.37 ± 0.04	0.0191 ± 0.0007
	M_* & star-forming	1.591 ± 0.011	-0.558 ± 0.004	2.17 ± 0.02	0.0180 ± 0.0005
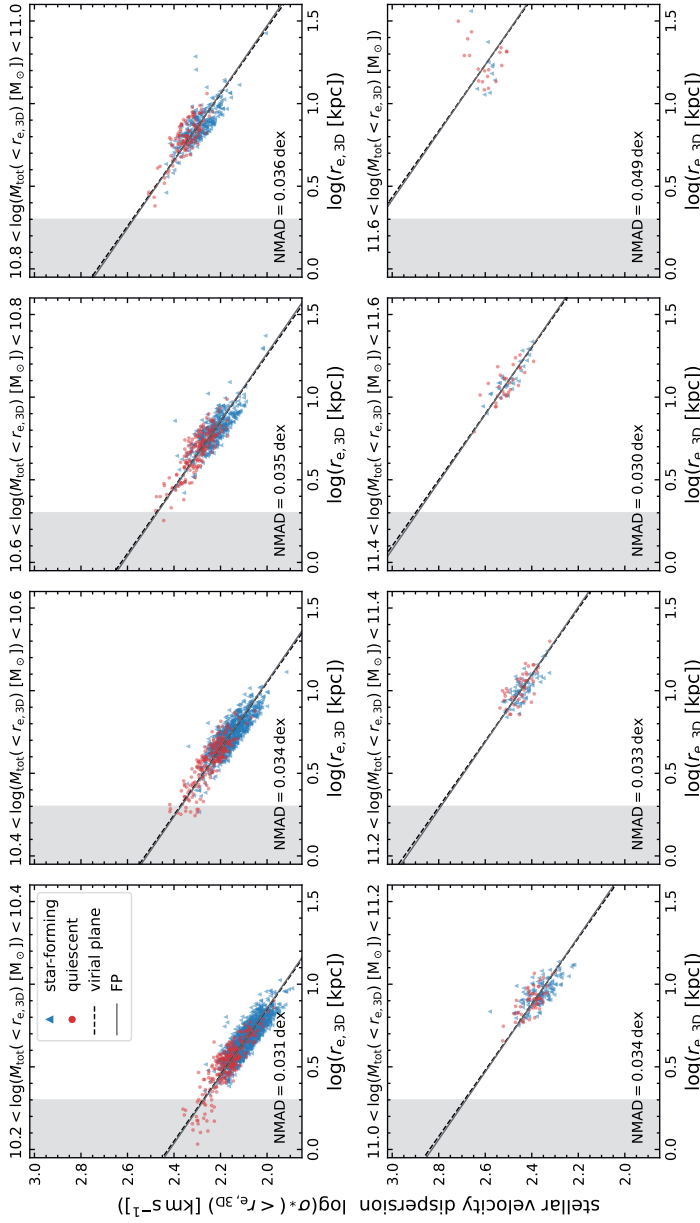
Figure 5.5: As Fig. 5.4, but showing the velocity dispersion of only the stellar particles. The scatter about the dashed lines is marginally lower, whereas the number of strong outliers (caused by recent mergers) is slightly increased. The tilt of the FP is closer to the virial plane when using σ_*, likely due to the decreased dependence on the galaxy structure and the fact that σ_{tot} is strongly influenced by the dynamically-hot dark matter particles (see main text).

where γ is the zero point of the plane. We obtain errors on the fit parameters by bootstrapping the data with 1000 subsamples. Excluding galaxies for which $r_{\rm e,3D} < 2\,\rm kpc$, we find that the coefficients deviate significantly from the virial plane (Table 5.1), with a stronger deviation for quiescent galaxies than for star-forming galaxies, and with very low scatter about the plane (0.018 dex). Solid lines in Fig. 5.4 show the best-fit total mass FP of the combined sample, and the scatter about these lines, calculated using the normalised median absolute deviation (NMAD), is indicated in each panel. We note that the scatter measured in Fig. 5.4 is larger than presented in Table 5.1 due to the finite bin widths used.

Next, in Fig. 5.5 we replace the total velocity dispersion by σ_*, which is expected to be a good tracer of the galaxy dynamics because of the collisionless nature of stellar orbits. The results are qualitatively similar to those of Fig. 5.4, except with slightly lower scatter. The overall scaling is also lower, as the stellar velocity dispersion is systematically lower than the total dispersion (Fig. 5.2). By inspecting the merger trees and images of the strong outliers that are visible in the figure, we find that these few systems are either currently merging with another galaxy or did so in their recent history, and have therefore likely not yet reached equilibrium.

Fitting the planar relation with $\sigma_*(< r_{\rm e,3D})$ instead of $\sigma_{\rm tot}(< r_{\rm e,3D})$ results in coefficients that are even closer, although still not equal, to the virial plane. Interestingly, the separate fits to the quiescent and star-forming subsamples are also in better agreement than before, and the scatter about these different planes is reduced even further (0.013 dex).

Whereas Figs. 5.4 and 5.5 focus on the qualitative differences between tilt of the FP and virial plane, we quantify the differences between the best-fit FP and the virial plane in Fig. 5.6. We compute the velocity dispersion predicted from the virial plane, which is equivalent to the circular velocity at the effective radius for a spherically symmetric mass distribution,

$$v_{\rm c}(r_{\rm e,3D}) = \sqrt{\frac{GM_{\rm tot}(< r_{\rm e,3D})}{r_{\rm e,3D}}} \,, \tag{5.11}$$

where G is the gravitational constant, and hence evaluate how the deviation between the measured velocity dispersion and this circular velocity depends on the half-mass radius. Fig. 5.6 shows that there is clearly a systematic offset between the zero-points of the total mass FP and virial plane, for both $\sigma_{\rm tot}$ (left) and σ_* (right), which we discuss in the following section (5.3.2). By measuring the Spearman rank correlation coefficients, we also find that there is a positive correlation between the deviation from the virial plane and the half-mass radius. This implies that the total mass FP is tilted with respect to the virial plane, and that this tilt is stronger for quiescent galaxies than for star-forming galaxies, and is consistent with the results of our planar fits (Table 5.1).

5.3.2   Effects of non-homology

The deviation of the different fits for the 'total mass' FP from the virial plane raises several questions. Most importantly, we may ask why the dynamical and
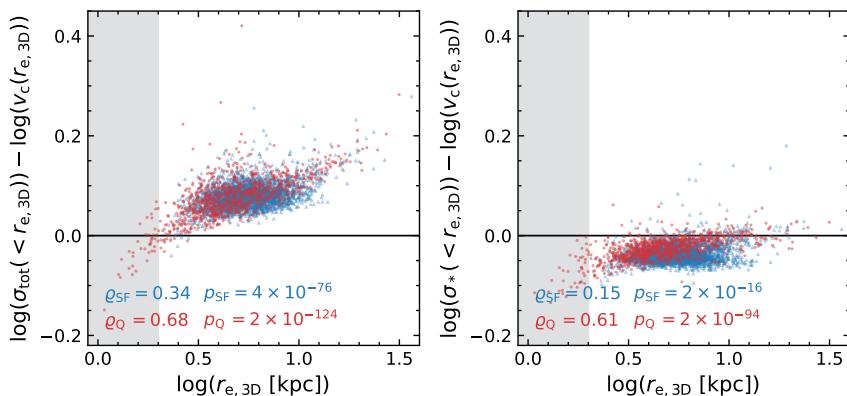
Figure 5.6: The offset between the FP and the virial plane, measured as the difference between the total (left) or stellar (right) velocity dispersion and the dispersion predicted by Eq. 5.11, as a function of the half-mass radius. There is not only a difference in the zero-point of the two planes, but also in the tilt: the positive Spearman rank correlation coefficients (and corresponding p-values) indicate that the FP is tilted with respect to the virial plane. This tilt is significant, and is stronger for quiescent galaxies than for star-forming galaxies.

total masses are different. Secondly, it is unclear why the use of the stellar velocity dispersion results in a FP that is closer to virial than is the case for the total velocity dispersion, given that the orbits of the cold dark matter are also collisionless and the gas fractions are small (see Section 5.3.3).

A difference between M_{dyn} and M_{tot} within the same spherical aperture of $r_{\mathrm{e,3D}}$ indicates that the measured velocity dispersion differs from the expected dispersion. Either the assumption of virial equilibrium does not hold, or the systems are not homologous. The first is unlikely, as the age of the galaxies at $z = 0.1$ is $\sim 10^{10}$ yr, and thus significantly larger than the crossing time ($\sim 10^8$ yr). Therefore, only for systems that have very recently merged with a significantly large neighbour, might we expect virial equilibrium to not have yet been established, which explains some of the apparent outliers in Figs. 5.4 and 5.5.

To examine the effects of non-homology, we again use the difference between the measured velocity dispersion and the velocity dispersion predicted from the virial plane with Eq. 5.11 ($\Delta \log \sigma$; equivalent to the offset between the measured FP and the virial plane). We then evaluate how this calculated deviation depends on different galaxy properties.

Fig. 5.7 shows $\Delta \log \sigma = \log(\sigma(r_{\mathrm{e,3D}})) - \log(v_c(r_{\mathrm{e,3D}}))$ for both the total and stellar velocity dispersion as a function of the instantaneous sSFR. This sSFR of course cannot be expected to drive the effects of structural non-homology, but may correlate with the galaxy structure and therefore lead to a correlation with $\Delta \log \sigma$. Indeed, Correa et al. (2017), Thob et al. (2019, Figs. 2 and 3) and de Graaff et al. (2022, Figs. 11, 12 and 14) show that the 3D shape and dynamical properties as well as the inferred projected structural parameters depend on the
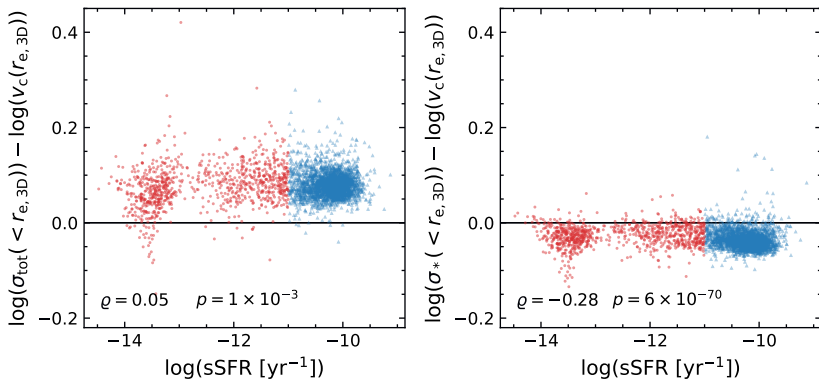
Figure 5.7: Difference between the measured total (left) or stellar (right) velocity dispersion and the dispersion predicted from the virial theorem (Eq. 5.11), as a function of the instantaneous sSFR. The Spearman rank correlation coefficients indicate that the sSFR has negligible impact on the deviation in σ_{tot}, and only a weak effect on σ_*.

colour, sSFR and stellar mass. We note that 419 galaxies have SFR $= 0\,\mathrm{M_\odot\,yr^{-1}}$, which for visualisation purposes only have been given an offset of $0.001\,\mathrm{M_\odot\,yr^{-1}}$ (corresponding to the cloud of points below sSFR $\lesssim -13\,\mathrm{yr^{-1}}$). The Spearman rank correlation coefficients (ϱ) indicate that the dependence on sSFR is, at most, weak. This is not unexpected, given that the individual fits to the quiescent and star-forming populations (Table 5.1) both deviate from the virial plane in an approximately equal way.

Next, we evaluate explicitly whether $\Delta \log \sigma$ correlates with differences in the galaxy structure. To do so, we use the Sérsic indices measured from the projected stellar mass distributions (Section 5.2.2), as well as the 3D structural parameters measured by Thob et al. (2019). The 3D stellar mass distributions were modelled with ellipsoids and quantified by the parameters $\epsilon_* = 1 - C/A$, which describes the flattening of the short axis (C) relative to the longest axis (A), and the triaxiality $T = (A^2 - B^2)/(A^2 - C^2)$, which also depends on the intermediate axis (B). A value of $T \approx 0$ thus corresponds to an oblate system, whereas $T \approx 1$ implies a prolate shape. The shape of the dark matter (within an approximately equal aperture as the stellar mass distribution) was measured in a similar way, and is quantified by the flattening parameter ϵ_{DM}. Furthermore, the kinematic structural parameters from Thob et al. (2019) provide information on the mean orbital properties of the stars, and are measured from cylindrical apertures that are aligned along the long axis with the total angular momentum of the stellar particles. The anisotropy in the velocity dispersion was then calculated as $\beta = 1 - (\sigma_z/\sigma_0)^2$, where σ_z is the stellar velocity dispersion along the long axis of the cylinder (the rotation axis of the galaxy) and σ_0 is the dispersion in the plane perpendicular to this axis, and thus reflects the degree of disordered motion along the radial or tangential direction. Finally, the quantity $\kappa_{\mathrm{co}} = K_{\mathrm{co}}^{\mathrm{rot}}/K$ measures the fraction of the total kinetic energy that is due to the co-rotation of stars along the axis defined by
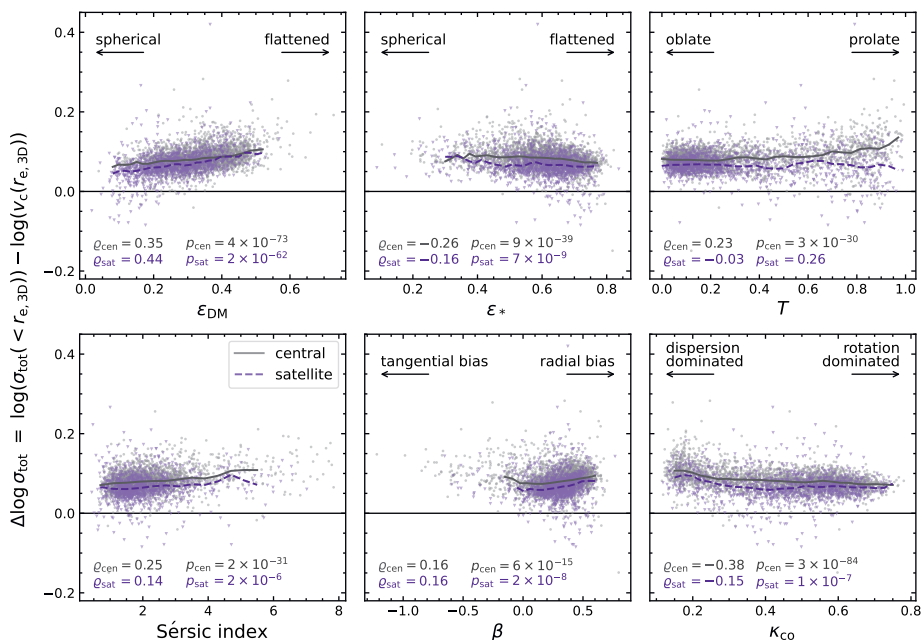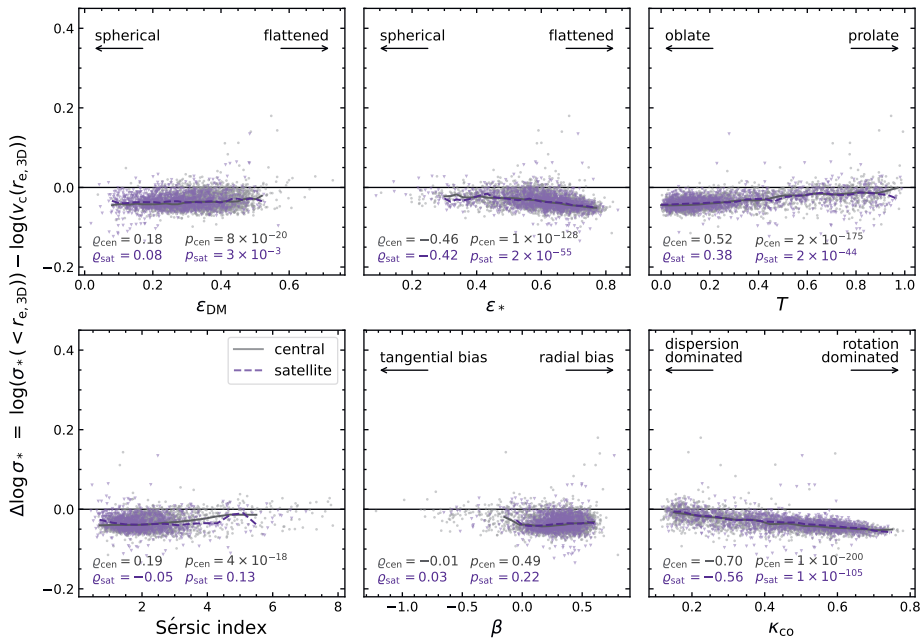
Figure 5.8: Deviation of the total velocity dispersion from the prediction of the virial plane (Eq. 5.11) as a function of various structural properties: the flattening of the 3D dark matter (ϵ_{DM}) and stellar mass (ϵ_*) distributions, triaxiality of the stellar mass (T), Sérsic index (n), anisotropy (β), and the co-rotational kinetic energy fraction (κ_{co}). Central (grey) and satellite (purple) galaxies are indicated separately, with solid and dashed lines showing the respective running medians in each panel. The correlations between $\Delta \log \sigma_{\mathrm{tot}}$ and the galaxy structure and dynamics demonstrate the effects of non-homology: galaxies have highly diverse morphologies and kinematic structures, which affects the measured value of σ_{tot} and hence causes a tilt in the FP with respect to the virial plane.

the total angular momentum. A value of $\kappa_{\mathrm{co}} = 1$ therefore corresponds to a dynamically-cold disc in which all stars follow circular orbits.

Figs. 5.8 and 5.9 show how $\Delta \log \sigma$ varies with these different structural properties, for both the total and stellar velocity dispersion, respectively. As the sSFR has minimal impact on the measured deviation, we omit the colour coding by sSFR in these figures. Instead, however, we distinguish between central (grey) and satellite (purple) galaxies, as the structural properties of satellite galaxies may be expected to be influenced by their local environment (e.g., through tidal stripping). The running median is plotted in each panel for the central galaxies (solid lines) and satellites (dashed lines).

Starting with σ_{tot}, we find correlations with all structural properties shown, particularly with the ϵ_{DM} and κ_{co}. These trends are generally stronger for central galaxies than the satellites, except for the ϵ_{DM}, and shows that the local environ-

Figure 5.9: Deviation of the stellar velocity dispersion from the prediction of the virial plane (Eq. 5.11) as a function of galaxy structure. Compared to Fig. 5.8, the stellar velocity dispersion is less dependent on the shape of the dark matter distribution ($\epsilon_{\rm DM}$), but instead depends more strongly on the structure of the stellar mass. For spherical, dispersion-supported systems σ_* approximates the virial plane prediction, but galaxies that are more flattened and strongly rotating diverge from this, due to the difference in the shape of the gravitational potential. Correlations are slightly weaker for satellite galaxies than centrals, which reflects the additional effect of the local environment on these systems.

ment has a small effect on the measured $\sigma_{\rm tot}$ and the other structural parameters. On the other hand, σ_* is less dependent on the structure of the dark matter, and instead depends strongly on the morphology and dynamics of the stellar mass.

The wide variety in shapes and structures among galaxies (discussed more extensively by Thob et al. 2019) clearly shows that the assumption of homology is incorrect. The measurement of $\sigma(< r_{\rm e,3D})$ reflects these variations in the density profiles: for instance, for oblate, rotating systems $\sigma_*(< r_{\rm e,3D})$ underestimates the total mass, whereas it is a good approximation of the total mass for more spherical systems with greater dispersion support. These differences in the structure therefore also lead to a deviation of the total mass FP from the virial plane. This is to be expected, as the virial plane (Eq. 5.11) assumes a spherically symmetric mass distribution, but the true circular velocities of galaxies depend on the shape of the mass distribution (see Binney & Tremaine 1987, Chapter 2). The fact that the FP with $\sigma_{\rm tot}$ deviates more strongly from the virial plane than σ_* can then be

attributed to the fact that σ_{tot} is sensitive to not only the stellar mass distribution, but also the dark matter.

However, as shown by Trayford et al. (2019), van de Sande et al. (2019) and de Graaff et al. (2022), the galaxy morphologies in the $100^3\,\mathrm{cMpc}^3$ EAGLE simulation are different from observed galaxies: the simulated galaxies tend to be thicker, and with significantly lower Sérsic indices. This is possibly the result of the pressure floor in the simulation, or due to the limited resolution, as Ludlow et al. (2019, 2021) showed that the 2-body scattering of the relatively massive dark matter particles with the baryonic particles in the simulation affects the resulting stellar mass density profiles. In Appendix 5.A we show results from the higher-resolution EAGLE simulations (described in Section 5.2.1), and demonstrate that the morphology (particularly the Sérsic index) is strongly dependent on the resolution, but that our conclusions on the effects of non-homology on the FP are robust to changes in the resolution.

Lastly, although we have explained the relative differences in $\Delta \log \sigma$, we have thus far neglected the fact that there is also a systematic offset in $\Delta \log \sigma_{\mathrm{tot}}$ visible in Fig. 5.7 and 5.8: σ_{tot} systematically overpredicts the total mass within $r_{\mathrm{e,3D}}$. This suggests that there must be a factor missing in Eq. 5.11, which can most plausibly be attributed to the assumptions made in obtaining the virial theorem. To arrive at the scalar virial theorem of $2K + W = 0$, where W is the gravitational potential energy, one has to assume that the mass density $\rho(r \rightarrow \infty) = 0$ (Binney & Tremaine 1987). Albeit a reasonable assumption for the stellar mass distribution, the distribution of the dark matter is more complex. The dark matter particles are more likely to be on highly eccentric orbits with semi-major axes that are significantly larger than the stellar half-mass radius of the galaxy. This can also be interpreted as a surface pressure term in the virial theorem, such that $2K + W + S_{\mathrm{p}} = 0$, with $S_{\mathrm{p}}/|W| < 0$ and hence $2K/|W| > 1$ (see also Shapiro et al. 2004). Therefore, we would expect to find $\sigma_{\mathrm{tot}} > v_{\mathrm{c}}$ (from Eq. 5.11), which is exactly what Fig. 5.7 and 5.8 show.

5.3.3   Variations in the dark matter fraction

Having quantified the $M_{\mathrm{dyn}}/M_{\mathrm{tot}}$ contribution to the tilt of the FP, we now add in the effects caused by the different mass compositions of galaxies, i.e. the contribution from stellar, gas and dark matter mass (we neglect the black hole mass, as this typically comprises $< 1\%$ of $M_{\mathrm{tot}}(< r_{\mathrm{e,3D}})$). This is also coupled with a change in the sample selection, as instead of using the M_{tot} complete sample, we from hereon focus on the M_*-selected sample and stellar velocity dispersions only.

First, we examine the effect of this change in the sample on the total mass FP. Because of the strong overlap between the two samples the coefficients are changed only weakly, although this is statistically significant. The total mass FP spanned by the M_*-selected sample is slightly closer to the virial plane than before, but the effects of non-homology discussed in the previous section still apply.

Second, we change from the total mass FP to the stellar mass FP (Eq. 5.4), and investigate the relation between $\sigma_*(< r_{\mathrm{e,3D}})$ and $r_{\mathrm{e,3D}}$ in narrow bins of $M_*(< r_{\mathrm{e,3D}})$. Fig. 5.10 differs from Fig. 5.5 only by the choice of the mass used
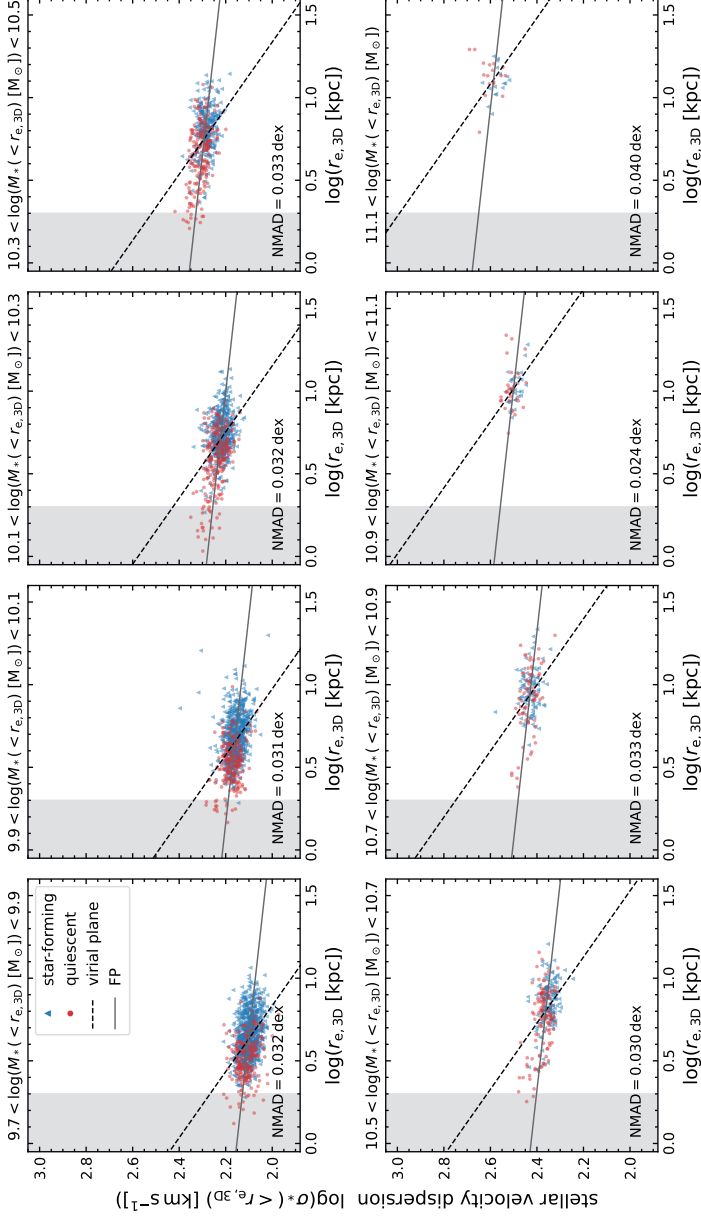
Figure 5.10: Stellar velocity dispersion (see Section 5.2.4) as a function of the half-mass radius, binned by the stellar mass $M_*(< r_{e,3D})$ rather than the total mass (Fig. 5.5). Symbols indicate the same as in Fig. 5.5. Galaxies clearly deviate strongly from the relations that they are expected to follow if $M_{dyn} \propto M_*$ (dashed lines show the slope corresponding to the virial plane). In addition to the effects of non-homology established in Section 5.3.2, variations in the dark matter and gas fractions within $r_{e,3D}$ must contribute to these deviations, and hence set the tilt of the stellar mass FP (solid lines).

to bin the data, with an additional small effect due to differences in the samples used. Dashed lines show the predicted slope in each panel for homologous galaxies in virial equilibrium, which is clearly a poor prediction. By fitting the tilt of the stellar mass FP, we find a much stronger deviation from the virial plane than before, particularly for the β coefficient, and with increased scatter (solid lines; Table 5.1).

We again define a predicted velocity dispersion, by replacing the total mass in Eq. 5.11 by the stellar mass:

$$\sigma_{\mathrm{pred}} = \sqrt{\frac{GM_*(< r)}{r}} \, , \tag{5.12}$$

using $r = r_{\mathrm{e,3D}}$ and evaluate how the dynamical and stellar mass differ from each other. The effects from non-homology discussed in the previous section (5.3.2) still hold here. However, there are now two new factors to consider: the dark matter (f_{DM}) and gas fraction (f_{gas}) within $r_{\mathrm{e,3D}}$.

In Fig. 5.11 we show the difference between the stellar velocity dispersion and the predicted dispersion ($\Delta \log \sigma_*$) as a function of $f_{\mathrm{DM}}(< r_{\mathrm{e,3D}})$ and $f_{\mathrm{gas}}(< r_{\mathrm{e,3D}})$. The upper panels are colour coded using the division into star-forming (blue) and quiescent (red) galaxies; the lower panels distinguish between central (grey) and satellite (purple) galaxies, for comparison with Fig. 5.9. There is a systematic offset in the obtained $\Delta \log \sigma_*$, as Eq. 5.12 misses a significant fraction of the total galaxy mass and therefore leads to a systematically lower value of σ_{pred}.

Most importantly, we find a very strong correlation between $\Delta \log \sigma_*$ and f_{DM}, for all four categories of galaxies. These trends are stronger than any of the correlations with galaxy structure found in Fig. 5.9, and therefore demonstrate that systematic variations in the dark matter fractions are the main driver of the tilt of the stellar mass FP. We find no correlation with the gas fraction for star-forming galaxies, which may be contrary to expectations, but the gas fractions are generally very low ($\lesssim 5\%$). On the other hand, there is a correlation among the satellite galaxies, which suggests an additional, weak effect of the local environment on the stellar mass FP.

Fig. 5.12 further examines the variation in f_{DM} across the stellar mass-size plane. This indicates that there is not simply a large variation in f_{DM}, but that the variation in f_{DM} is a smooth power-law function of both M_* and $r_{\mathrm{e,3D}}$. In turn, these variations result in the observed strong correlation between $\Delta \log \sigma_*$ and f_{DM} found in Fig. 5.11, and hence the tilt of the simulated stellar mass FP. The common FP for star-forming and quiescent galaxies can then be interpreted as the power-law relation $f_{\mathrm{DM}}(< r_{\mathrm{e,3D}}) \propto M_*^a r_{\mathrm{e,3D}}^b$ having similar coefficients a and b for both galaxy populations. We estimate the coefficients by minimising the sum of the offsets orthogonal to the planar relation:

$$\Delta_{\mathrm{DM}} = \frac{|\log(f_{\mathrm{DM}}(< r_{\mathrm{e,3D}})) - a \log(M_*) - b \log(r_{\mathrm{e,3D}}) - c|}{\sqrt{1 + a^2 + b^2}} \, , \tag{5.13}$$

where c is the zero-point of the relation and $M_* = 2M_*(< r_{\mathrm{e,3D}})$. We present the results of these fits in Table 5.2.
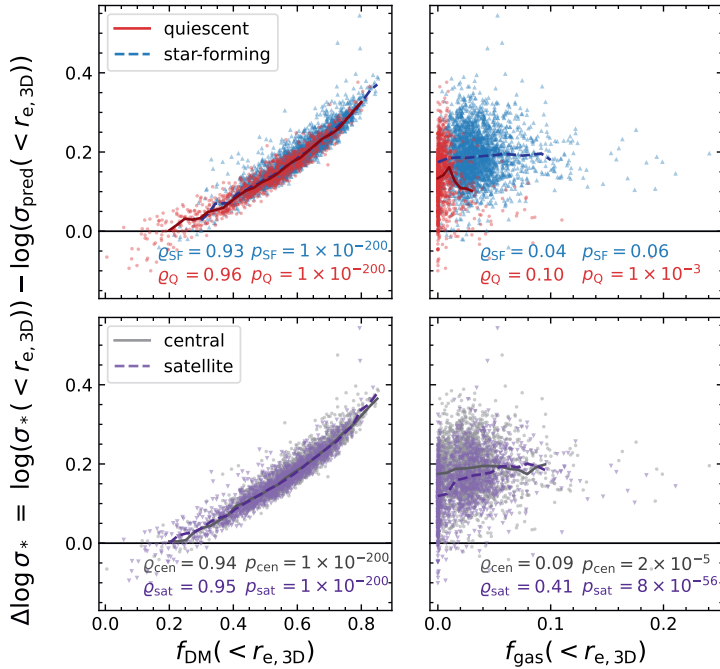
Figure 5.11: Deviation between the measured stellar velocity dispersion and the dispersion predicted from the stellar mass and stellar half-mass radius (Eq. 5.12; corresponding to the offset from the virial plane) versus the dark matter (left) and gas (right) fractions within the half-mass radius. Top panels distinguish between star-forming (blue) and quiescent (red) galaxies; bottom panels separate central (grey) and satellite (purple) galaxies. Solid and dashed lines show the running medians. The correlation with the dark matter fraction is stronger than for any other parameter (Fig. 5.9), indicating that it is the primary driver of the tilt of the stellar mass FP, for both star-forming and quiescent galaxies.

Table 5.2: Best-fit coefficients for the relation $\log(f_{\mathrm{DM}}(< r_{\mathrm{e,3D}})) = a \log(M_*) + b \log(r_{\mathrm{e,3D}}) + c$. Galaxies for which $r_{\mathrm{e,3D}} < 2\,\mathrm{kpc}$ are excluded from the fits.

Sample selection	a	b	c	NMAD
M_*	-0.214 ± 0.004	0.573 ± 0.009	1.58 ± 0.04	0.0312 ± 0.0006
M_* & quiescent	-0.317 ± 0.008	0.781 ± 0.015	2.54 ± 0.07	0.0339 ± 0.0014
M_* & star-forming	-0.194 ± 0.004	0.518 ± 0.008	1.41 ± 0.04	0.0280 ± 0.0007
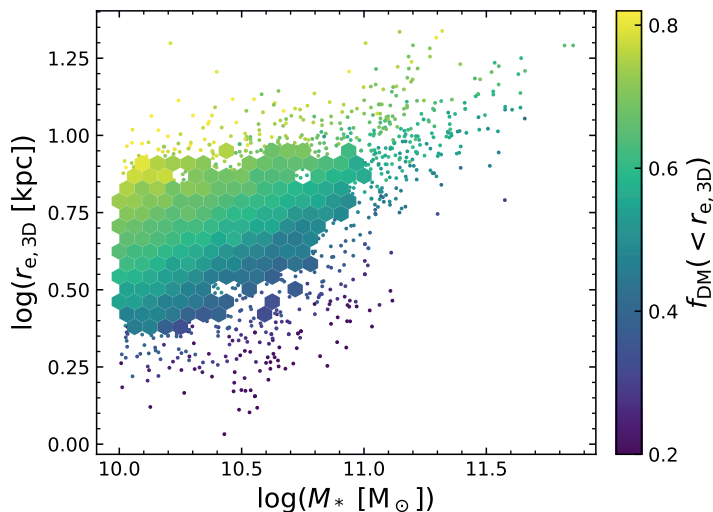
Figure 5.12: Total stellar mass versus the half-mass radius for all (quiescent and star-forming) galaxies, colour-coded by the dark matter fraction within the half-mass radius. The variation in f_{DM} is a smooth function of both size and stellar mass, which leads to the observed correlation between $\Delta \log \sigma_*$ and f_{DM} in Fig. 5.11. As a result, quiescent and star-forming galaxies lie on a common stellar mass FP that deviates strongly from the virial plane.

5.4   Observing the Fundamental Plane

Measurements of the FP in the previous section relied entirely on 3D measurements of the size, mass and velocity dispersion. To be able to compare with observations, we need to take into account the different observational effects that may bias the observed FP with respect to the intrinsic '3D FP'. Broadly, these are the effects of projection along a random viewing angle, differences in the measurement methods and associated measurement uncertainties, gradients in the M_*/L ratio and associated systematic uncertainties in the assumed IMF, and selection biases. In this section, all these effects are added in, to arrive at a realistic measurement of the FP.

5.4.1   Impact of projection effects and measurement biases

We use the mock observations described in Section 5.2, and begin with the measurements of the projected stellar mass distributions. As before, we create bins in stellar mass, which are now changed to the mass inferred from the best-fit Sérsic profile, and show the line-of-sight velocity dispersion as a function of the half-mass radius in Fig. 5.13. Star-forming and quiescent galaxies are again indicated separately, using blue and red symbols, respectively. The virial plane is shown as the dotted line in each panel for easy comparison with previous figures.

The main differences with respect to Fig. 5.10 are in the scatter: the offset in the velocity dispersion $\Delta \log \sigma_* = -0.007$ dex with a scatter of 0.05 dex, and the offset in the size $\Delta \log r_e = -0.06$ dex, with a scatter of 0.07 dex. The mass bins are also changed slightly, although this effect is small (offset of $\Delta \log M_* = -0.02$ dex with a scatter of 0.04 dex). These systematic offsets and scatter arise from projection effects, which particularly affect the velocity dispersions, and, for the sizes and masses, differences in the measurement methods and measurement uncertainties. We note that the projected velocity dispersions as measured in Section 5.2.4 are noise-free (nor include PSF smoothing), and the scatter therefore is purely from projection along the line of sight. However, measurement errors are expected to be subdominant, as we find that the typical uncertainty on the velocity dispersion for galaxies of $M* > 10^{10}\,\mathrm{M}_\odot$ at $z \sim 0$ in the Sloan Digital Sky Survey is ≈ 0.02 dex (SDSS; using the sample described in de Graaff et al. 2021).

To estimate the effect on the inferred stellar mass FP, we first calculate the stellar mass surface density within the elliptical aperture described by the $r_{e,*}$ and the axis ratio q_*: $\Sigma_* = M_*(< r_{e,*})/(\pi q_* r_{e,*}^2)$, with M_* being half of the stellar mass of the integrated Sérsic profile. We then measure the tilt of the stellar mass FP with the same orthogonal fitting used previously (dashed lines in Fig. 5.13; Table 5.3), and compare with the stellar mass FP measured from the 3D measurements (solid lines; Table 5.1). The scatter about this solid line is printed in each panel for comparison with Fig. 5.10.

The increased scatter affects mainly the α parameter of the tilt, likely due to the asymmetric scatter toward low σ_* from galaxies that are close to face-on. Moreover, the scatter about the FP itself is nearly doubled. The β parameter is largely unchanged, however, despite an offset and significant scatter in the size, which can be understood from the fact that changes in the size correlate in a direction that is near-parallel to the FP itself (see also Appendix B of de Graaff et al. 2021).

However, observational studies of the FP rarely use the effective radius as we have here, i.e., the semi-major axis size. Rather, the effective radius is often circularised, such that $r_{circ} = \sqrt{q}\,r_e$, as this may be a better approximation of the galaxy size for systems that are not oblate in shape (the fraction of low-redshift early-types that have prolate shapes or are triaxial). It also serves as a crude correction for the projection effects on σ_*, mitigating residual correlations with q throughout the FP. We therefore repeat our fits using the circularised size instead of the major axis size, and present the results in Table 5.3 and Appendix 5.B. Again, the changes to the β parameter are small, because of the covariance between r_{circ} and Σ_*. On the other hand, the α parameter depends strongly on the measure of size that is used, due to the corrective effect on the velocity dispersion. The FP resulting from the circularised sizes is closer to the intrinsic, 3D FP than is the case for the major axis sizes, despite the fact that the circularised size is a poor approximation of the galaxy size for oblate systems (i.e., most of the galaxy population). This also suggests that a properly-calibrated correction for the projected velocity dispersions, as derived empirically by van der Wel et al. (2022, submitted) using dynamical Jeans models, may provide even better results than the ad hoc correction from the circularised sizes.
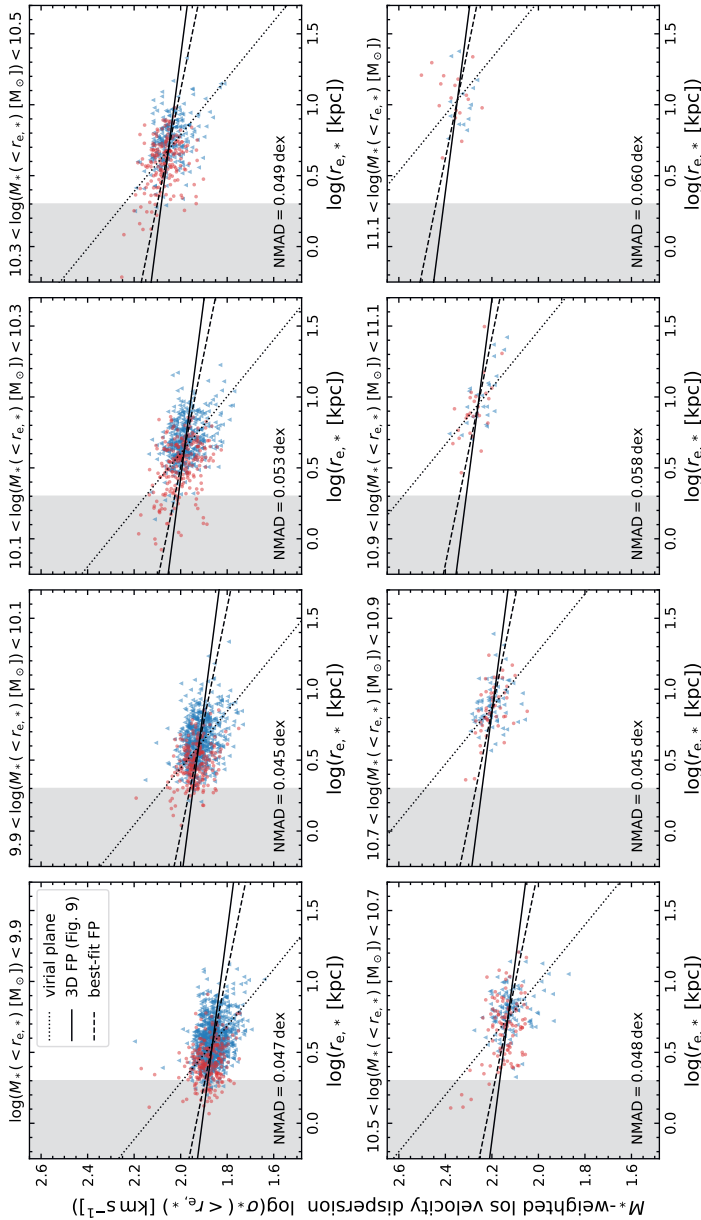
Figure 5.13: Stellar mass-weighted, line-of-sight velocity dispersion versus the half-mass radius from the best-fit Sérsic profile, in bins of the stellar mass $M_*(< r_{e,*})$. Star-forming and quiescent galaxies are indicated by blue and red symbols, respectively. The virial plane is shown as the dotted line, and solid lines show the stellar mass FP from fitting to the 3D aperture measurements (Table 5.1). The orthogonal scatter about this '3D FP' is printed in each panel for comparison with Fig. 5.10. Projection effects and differences in the applied measurement methods result in an increased scatter, and slightly alter the inferred stellar mass FP (dashed lines; Table 5.3).

Table 5.3: Best-fit coefficients of the stellar mass FP and orthogonal scatter, for different measures of galaxy size and velocity dispersion. Galaxies for which $r_{\rm e} < 2\,{\rm kpc}$ are excluded from the fits.

Relation	Sample selection	α	β	γ	NMAD
$\log(r_{\rm e,*}) = \alpha \log(\sigma_*(< r_{\rm e,*})) + \beta \log(\Sigma_*) + \gamma$	M_*	1.760 ± 0.018	-0.610 ± 0.006	2.40 ± 0.04	0.0353 ± 0.0008
	M_* & quiescent	1.62 ± 0.03	-0.550 ± 0.013	2.14 ± 0.09	0.043 ± 0.002
	M_* & star-forming	1.82 ± 0.02	-0.624 ± 0.006	2.42 ± 0.04	0.0326 ± 0.0009
$\log(r_{\rm circ,*}) = \alpha \log(\sigma_*(< r_{\rm e,*})) + \beta \log(\Sigma_*) + \gamma$	M_*	1.578 ± 0.015	-0.617 ± 0.005	2.71 ± 0.03	0.0311 ± 0.0006
	M_* & quiescent	1.49 ± 0.02	-0.584 ± 0.008	2.60 ± 0.07	0.0345 ± 0.0016
	M_* & star-forming	1.618 ± 0.016	-0.634 ± 0.006	2.77 ± 0.04	0.0296 ± 0.0007
$\log(r_{\rm e,r}) = \alpha \log(\sigma_*(< r_{\rm e,r})) + \beta \log(\Sigma_*) + \gamma$	M_*	1.407 ± 0.014	-0.578 ± 0.005	2.81 ± 0.04	0.0436 ± 0.0009
	M_* & quiescent	1.45 ± 0.03	-0.532 ± 0.015	2.33 ± 0.12	0.0478 ± 0.0017
	M_* & star-forming	1.38 ± 0.015	-0.576 ± 0.006	2.84 ± 0.05	0.0388 ± 0.0011
$\log(r_{\rm circ,r}) = \alpha \log(\sigma_*(< r_{\rm e,r})) + \beta \log(\Sigma_*) + \gamma$	M_*	1.266 ± 0.017	-0.545 ± 0.005	2.70 ± 0.04	0.0485 ± 0.0010
	M_* & quiescent	1.33 ± 0.02	-0.560 ± 0.010	2.69 ± 0.08	0.0378 ± 0.0014
	M_* & star-forming	1.19 ± 0.02	-0.546 ± 0.008	2.84 ± 0.06	0.0530 ± 0.0012

Lastly, we examine the effects of variations in M_*/L (for a universal IMF) within galaxies, which can lead to significantly larger sizes, depending on the star formation activity and dust attenuation. The velocity dispersions as calculated in Section 5.2.4 do not include the effects of dust, however, the measurements will be biased toward the location of the younger stellar populations. As a result, the scatter in the velocity dispersion (Fig. 5.3) is increased, as the younger stellar populations tend to lie in dynamically-cold discs (Trayford et al. 2019) and the projected velocity dispersions therefore will be more strongly dependent on the inclination angle.

Fig. 5.14 shows the equivalent of Fig. 5.13, but for the luminosity-weighted measurements. Except for the two highest mass bins, the scatter is significantly increased (by $\approx 50\%$), due to the greater scatter in the size ($0.14\,\mathrm{dex}$ in comparison with the 3D half-mass radii) and the velocity dispersion (scatter of $0.07\,\mathrm{dex}$ in comparison with the 3D dispersions). In Appendix 5.C, we investigate whether the strong increase in the scatter is caused by the inconsistency in the tracer used for the size and velocity dispersion, i.e., the use of half-light radii measured from r-band imaging that include the effects of dust, while the velocity dispersions are measured using the unattenuated g-band luminosities. We show that the strong increase in the scatter between the top panels of Figs. 5.13 and 5.14 is driven primarily by the change from a M_*-weighting to a L-weighting for the velocity dispersions, with the (in)consistency between the tracers being of lesser importance.

To measure the coefficients of the mock observed FP, we again calculate $\Sigma_* = M_*(< r_{\mathrm{e,r}})/(\pi q_{\mathrm{r}} r_{\mathrm{e,r}}^2)$, now using the stellar mass and axis ratio corresponding to the luminosity-weighted Sérsic model (see Section 5.2.2). As a result, the best-fit stellar mass FP has systematically different values for the α parameter (Table 5.3) than is the case for the mass-weighted measurements, and the scatter is further increased. We also perform the fits with the circularised half-light radii, and show the corresponding figure in Appendix 5.B. These again lead to a difference in α alone, as the result of the circularised size effectively compensating for the projection effects on σ_*.

5.4.2   Selection bias

All fits of the stellar mass FP thus far have been based on the stellar mass-selected sample. In observational studies of the FP, however, these galaxies would likely not all be selected: imaging and spectroscopic surveys have lower completeness at low luminosities, as well as toward low velocity dispersions (due to the limitation in the spectral resolution of the instrument, or a selection against velocity dispersions with large measurement uncertainties). As also shown by, e.g. Hyde & Bernardi (2009), these selection effects lead to a bias in the measured tilt of the FP.

Given the relatively high stellar mass (and therefore high luminosity) of our sample, we may expect all these galaxies to be identified in large imaging surveys of the $z \sim 0$ Universe, except for possibly very rare, very low surface brightness objects. Moreover, luminosity biases are relatively easily corrected for using standard Vmax corrections. We therefore only examine the effects of selection cuts
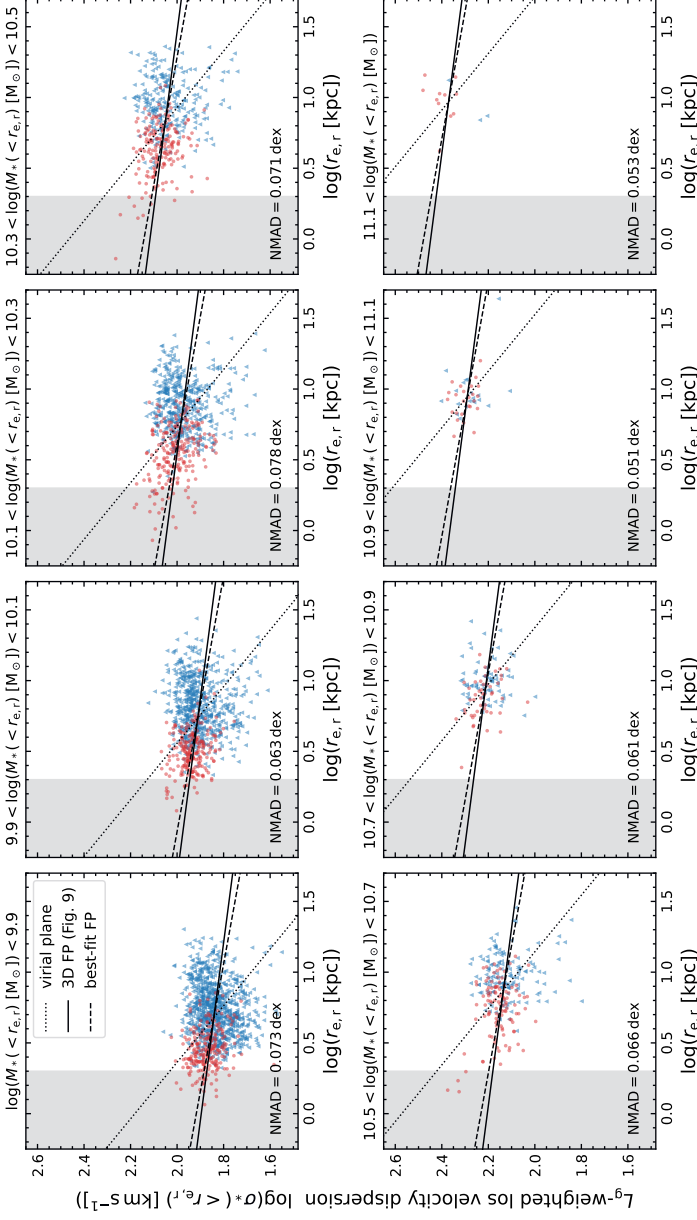
Figure 5.14: Projected, luminosity-weighted velocity dispersion versus the half-light radius from the best-fit Sérsic profile, in bins of the stellar mass $M_*(< r_{e,r})$ [$M_\odot$]. Symbols indicate the same as in Fig. 5.13. Gradients in the M_*/L ratio have a strong effect on the scatter, which is increased by $\approx 50\%$ in the lower mass bins, and make it more difficult to precisely retrieve the intrinsic (3D) stellar mass FP.
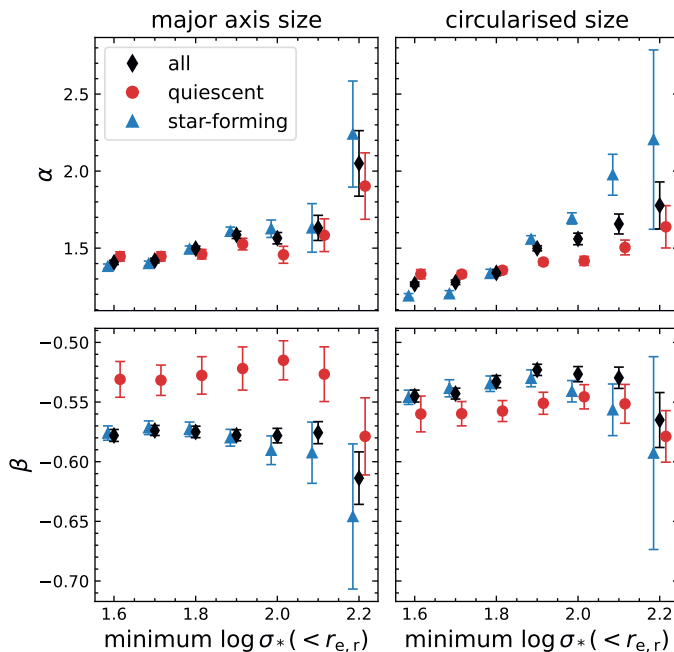
Figure 5.15: Effect of a selection bias in $\log \sigma_*$ on the measured tilt of the stellar mass FP for the luminosity-weighted measurements, using the major axis effective radius (left) and circularised effective radius (right) as the measure of size. Blue and red data points have been given a slight offset in $\log \sigma_*$ for visualisation purposes. Although the β parameter is quite stable (varying by $\lesssim 10\%$), the α parameter can vary strongly depending on the selection (differing by up to $\approx 80\%$ with respect to the complete sample).

in $\log \sigma_*$, as the dispersions do not scale trivially with luminosity or stellar mass and are susceptible to strong variation from the random projection on the sky.

We use the luminosity-weighted measurements, and measure the stellar mass FP after imposing different selections on $\log \sigma_*$ (i.e., horizontal cuts in Fig. 5.14). Fig. 5.15 shows the dependence of the parameters α and β on the different selections in $\log \sigma_*$, for the full sample (black), and the quiescent (red) and star-forming (blue) sub-samples. Both fits using the semi-major axis half-light radii (left) and circularised sizes (right) are shown.

We find results are qualitatively similar to the effects found by Hyde & Bernardi (2009): the β parameter varies only weakly with differences in the sample selection, particularly for the fits using the major axis sizes. On the other hand, the α parameter increases toward higher cuts in $\log \sigma_*$, which is a particularly strong effect for the fits using the circularised sizes. The star-forming population even reaches super-virial values, albeit with large uncertainties.

Although low-redshift studies of the FP will focus on samples of quiescent

galaxies with high completeness down to $\log(\sigma_*/\mathrm{km\,s}^{-1}) \approx 1.8$, these selection effects will likely become much more important toward higher redshift. Given the difficulty in measuring stellar absorption lines in high-redshift galaxies, the completeness boundary for the velocity dispersion shifts to $\log(\sigma_*/\mathrm{km\,s}^{-1}) \approx 2.1 - 2.2$ (e.g., Holden et al. 2010; de Graaff et al. 2021). At the same time, the structural properties of quiescent galaxies also change, becoming more disc-like and with stronger rotational support (Chang et al. 2013; Bezanson et al. 2018a), which will likely lead to selection effects that are more similar to that of the star-forming population in Fig. 5.15. To measure the evolution in the FP, then requires not just a fair comparison sample (i.e., a low-redshift sample for which the velocity dispersions are greater than the high-redshift completeness limit in $\log \sigma_*$), but also a correction factor to account for evolution in the (dynamical) structures of the galaxy population.

5.4.3 Contribution of IMF variations

Although we have measured in detail the stellar mass FP and its dependence on different observational effects, one potentially significant systematic uncertainty remains due to the assumption of a universal IMF in the simulations. We use the simulations with a pressure-dependent bottom-heavy ('LoM') and top-heavy ('HiM') IMF, which varies both between and within galaxies (further described in Section 5.2.1), to assess the magnitude of this uncertainty on the FP within the EAGLE simulations.

We compare the variable IMF runs with the $50^3 \, \mathrm{cMpc}^3$ simulation that uses the reference model ('Ref'), which assumes a fixed, Chabrier IMF. Because observational studies typically assume a universal IMF, we can no longer use a sample selection based on the summed stellar particle masses for a fair comparison between the different simulations. Instead, we therefore select by the r-band luminosity: luminosities for the stellar particles were computed using the FSPS software, and based on the age, metallicity and IMF of each particle (for details, see Barber et al. 2018). By comparing the total luminosity within a spherical aperture of radius 30 kpc to the stellar mass within the same aperture for galaxies in the Ref simulation, we find that a minimum (rest-frame) r-band luminosity of $10^{9.85} \, L_{\odot,r}$ ($M_r \approx -20.0$) results in a selection completeness of $\gtrsim 50\%$ down to a stellar mass of $10^{10} \, \mathrm{M}_\odot$. Applying this limit provides a sample of 527, 528, and 415 galaxies in the Ref, LoM and HiM simulations, respectively. As before, we divide these samples into star-forming and quiescent subsamples using the boundary of $\mathrm{sSFR} = 10^{-11} \, \mathrm{yr}^{-1}$. We note that the SFR and stellar mass are both dependent on the IMF, but the effect on the sSFR and thus the definition of quiescence is negligible due to the approximately equal change in the SFR and the stellar mass (see also Clauwens et al. 2016).

We subsequently extract 3D aperture measurements as used in Section 5.3, measuring the particle masses and stellar velocity dispersions within the 3D half-mass radii. To estimate the effect of an incorrect assumption for the IMF requires luminosity-based measurements. We use measurements from the online catalogues (McAlpine et al. 2016; Barber et al. 2018), as individual particle luminosities are
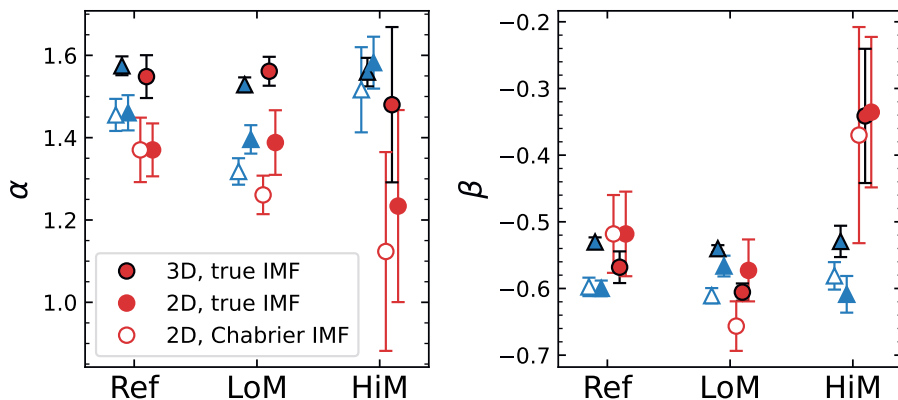
Figure 5.16: Measured tilt of the stellar mass FP for galaxies in the 50^3 cMpc^3 EAGLE simulations for the reference, bottom-heavy IMF (LoM) and top-heavy IMF (HiM) models. Circles and triangles show results for the quiescent and star-forming populations, respectively. Black outlined symbols indicate the fits to the 3D aperture measurements (as in Table 5.1); coloured symbols show the fits to the projected measurements using the true stellar mass (filled symbols) and Chabrier-reinterpreted stellar mass (open symbols), respectively. The effects of a variable IMF are small for the intrinsic (3D) stellar mass FP. However, by assuming an incorrect IMF, the inferred tilt of the stellar mass FP can be changed by $\approx 10\%$, and up to $\approx 20\%$ for quiescent galaxies.

not available. The size used is the circular half-light radius in the r-band, $r_{\rm e,2D}$, based on the total luminosity within a $30\,{\rm kpc}$ spherical aperture[1]; the velocity dispersion ($\sigma_*(< r_{\rm e,2D})$) is measured within a circular aperture of radius $r_{\rm e,2D}$ in projection along the z-axis of the simulation box, and weighted by the r-band luminosities of the particles. For the stellar mass, we use (i) the true stellar mass within the circular aperture (i.e., based on the varying IMF; $M_*(< r_{\rm e,2D})$) and (ii) the stellar mass within the same aperture that is reinterpreted under the assumption of a Chabrier IMF ($M_{*,{\rm Chab}}(< r_{\rm e,2D})$; Chabrier 2003). The latter quantity is calculated by multiplying the r-band luminosity with the M_*/L_r ratio that is obtained for the particles when these are evolved with a Chabrier IMF (using FSPS), and therefore allows for a comparison with observations (Barber et al. 2018).

The tilt of the stellar mass FP is measured in the same way as before, with the results shown in Fig. 5.16 for the three simulations. The 3D aperture measurements are shown as the black circles (triangles) for the quiescent (star-forming) populations. The results for the Ref simulation are consistent with the measurements in

---

[1]We note that these sizes are free from measurement uncertainties and do not include the effects of dust, therefore leading to considerably lower scatter than seen in Section 5.4.1. Moreover, these circular half-light radii are smaller than the 3D half-mass radii by $\approx 25\%$ for all galaxies, which differs strongly from the Sérsic model half-light radii that are larger by $\approx 25\%$ and smaller by $\approx 10\%$ in comparison with the 3D half-mass radii for star-forming and quiescent galaxies, respectively.

Table 5.1. For the LoM model, the fits to these 3D aperture measurements are very similar to those of the Ref model, with only a marginally lower value of β. The star-forming population in the HiM simulation is also consistent with the other models, and only the quiescent galaxies diverge, particularly in the measurement of β, although with a large measurement uncertainty.

Because of the differences in the projected measurements with respect to the previous sections of this paper, measurements of the tilt based on the 2D quantities cannot easily be compared to the results of Table 5.3. However, comparison of the 2D and 3D measurements using the true stellar masses gives insight into the effects of measurement biases: the filled coloured symbols show that there is indeed a small difference in the measured tilt due to a combination of effects from the projection, aperture definition and M_*/L gradients.

Comparing the three simulations, we find that the measurements of the tilt for the LoM model are close to those of the Ref model ($< 10\%$ difference), whereas the HiM model deviates more strongly. We calculate the difference in the measured (projected) stellar velocity dispersion and the predicted velocity dispersion (using the true stellar mass; Eq. 5.12) and examine the drivers of these FPs in Fig. 5.17. Instead of the dark matter fraction, we compute the 'dark mass' fraction as $1 - f_*(<r_{\rm e,2D})$, where $f_* \equiv M_*/M_{\rm tot}$ and is calculated using 3D apertures of radius $r_{\rm e,2D}$. If the gas fraction is negligible, f_* simply measures the dark matter fraction. The LoM model leads to slightly lower dark fractions than the Ref model, particularly for the quiescent galaxies. On the other hand, galaxies in the HiM simulation are strongly dark matter-dominated within $r_{\rm e,2D}$. Despite this difference, there is a strong correlation between $\Delta \log \sigma_*$ and $1 - f_*$ for both models. We also show the correlations with the luminosity-weighted average IMF slope within the circular aperture of $r_{\rm e,2D}$. Although these correlations are strong, the fluctuations in $1 - f_*$ still dominate, indicating that variations in the dark matter fractions are still the primary driver of the simulated stellar mass FP. This likely also explains why we find little variation in the tilt between the different simulations.

Finally, we can quantify the effect of IMF variations on the observed tilt of the stellar mass FP, by using the 2D measurements and comparing the fits obtained for the true stellar masses and the reinterpreted stellar masses, shown as the open symbols in Fig. 5.16. As expected for the Ref model, the Chabrier and Chabrier-reinterpreted IMF measurements result in identical fits. For the LoM model both α and β are slightly lower in value for the Chabrier-reinterpreted measurements, and lower than is measured for the Ref model (by $\approx 10 - 25\%$), an effect that is stronger for the quiescent galaxy population. Star-forming galaxies in the HiM model are not affected significantly by a change in the assumed IMF, and agree well with the fit to the Ref model. On the other hand, the quiescent population does show a very different tilt from the Ref model. Although the formal statistical uncertainty on the fit is large (due to a small sample size and likely a small number of outliers affecting the fitting), the result itself is of significance, as the comparison is between three simulations with equal initial conditions.

In Fig. 5.18 we show the corresponding change to the results of Fig. 5.17, obtained by calculating the velocity dispersion predicted from the Chabrier-reinterpreted stellar mass, and similarly the reinterpreted stellar mass fraction $(1 - f_{*,\rm Chab})$. For
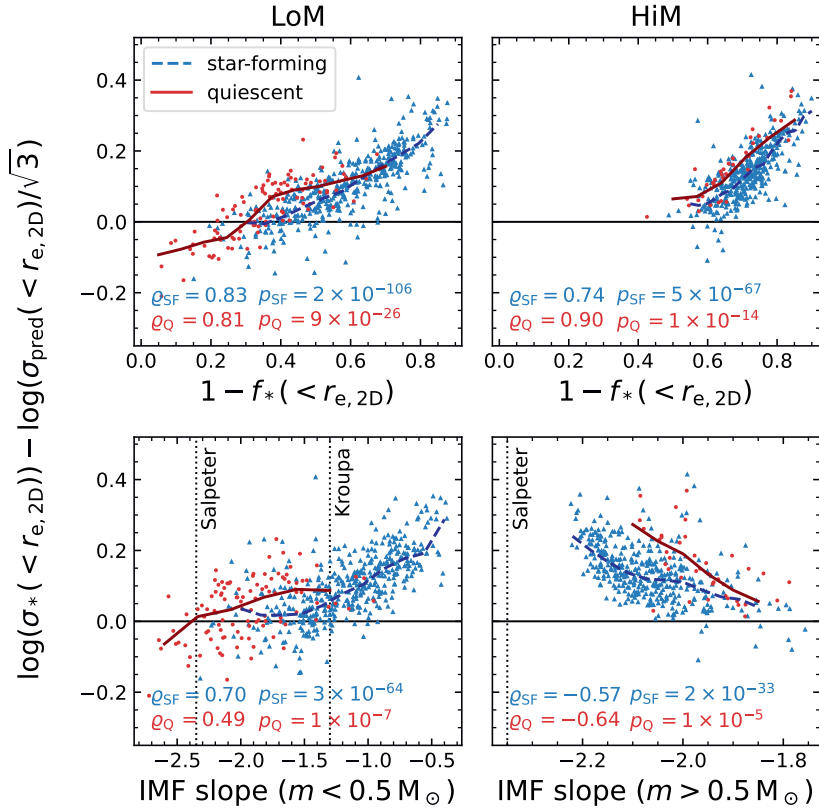
Figure 5.17: Deviation between the measured stellar velocity dispersion and the predicted velocity dispersion for galaxies in the LoM and HiM variable IMF simulations, as a function of the dark mass fraction and luminosity-weighted IMF slope within the circular aperture of $r_{e,2D}$. Galaxies in the LoM simulation are slightly less dark matter-dominated than in the reference model (in comparison with Fig. 5.11), whereas the HiM model produces strongly dark matter-dominated galaxies. In both cases, there is a somewhat stronger correlation between $\Delta \log \sigma_*$ and $1-f_*$ than with the IMF slope, indicating that fluctuations in the dark matter content are the primary driver of the stellar mass FP, and likely explains the weak variation in the inferred tilt between the different simulations (Fig. 5.16).

the LoM model, the effect of an incorrectly assumed IMF largely removes the correlation with the IMF slope for the quiescent galaxies, and strongly reduces the effect for the star-forming galaxies. Instead, the missing stellar mass is interpreted as extra dark matter, resulting in high values of and a strong correlation with $1 - f_{*,\mathrm{Chab}}$. The difference in the tilt between the 'true IMF' and Chabrier IMF measurements then must stem from the mismatch between the density profile that is traced by the measured velocity dispersion and the (incorrectly) estimated stellar mass surface density. Interestingly, the correlations for galaxies in the HiM
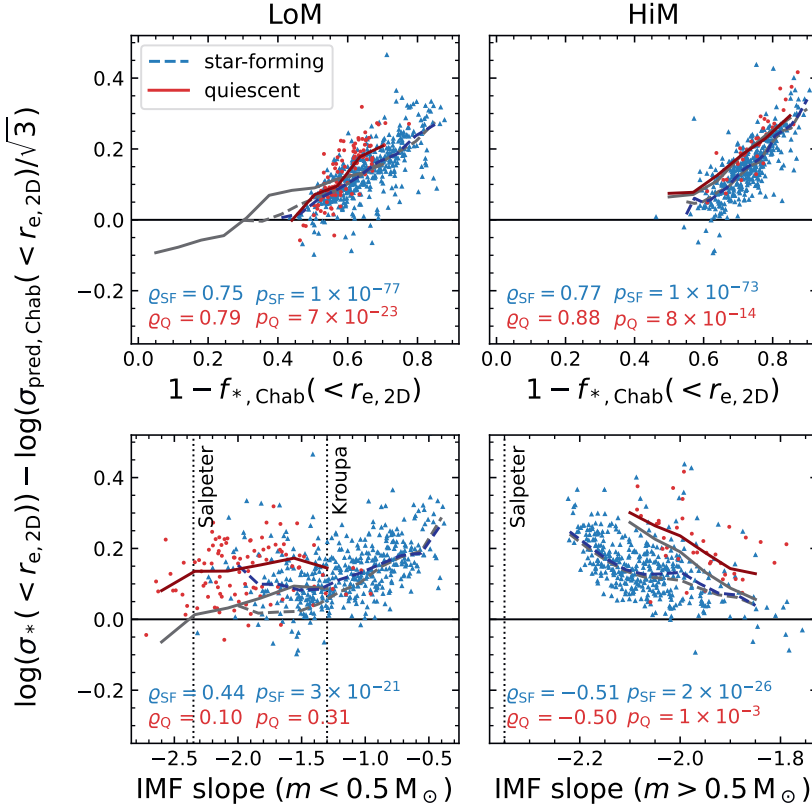
Figure 5.18: Deviation between the measured stellar velocity dispersion and the velocity dispersion predicted from the Chabrier-reinterpreted stellar mass for galaxies in the LoM and HiM variable IMF simulations, as a function of the dark mass fraction and luminosity-weighted IMF slope within the circular aperture of $r_{\mathrm{e,2D}}$. The Chabrier IMF underestimates the true stellar mass of galaxies in the LoM simulation, leading to increased dark mass fractions compared with Fig. 5.17 (grey lines show the medians from Fig. 5.17), and weakened correlations with the IMF slope. Galaxies in the HiM simulation are largely unaffected by a difference in the assumed IMF, which may reflect strong structural differences between galaxies in the HiM simulation and the other models.

simulation are largely unchanged with respect to Fig. 5.17, which is consistent with the very small changes found in the tilt between the 'true IMF' and Chabrier IMF fits. Possibly, the top-heavy IMF strongly affects the structural properties of galaxies in the HiM model, such that the profile of the dark matter and corresponding variations in the dark matter content dominate the stellar mass FP (see also Section 5.5.2), and a reinterpretation of the stellar mass with a different IMF therefore has a comparatively small impact.

Overall, in comparison with the Ref model, the parameters of the observed

stellar mass FPs using the reinterpreted stellar masses differ from each other by at most $\approx 9\%$ for star-forming galaxies and at most $\approx 29\%$ for quiescent galaxies. For star-forming galaxies, this effect is thus of similar magnitude to the effects of non-homology (Section 5.3.2); for quiescent galaxies, the uncertainty on the IMF has a larger effect, although this is still subdominant to the effects of variations in the dark matter content.

5.5   Discussion

5.5.1   Interpreting the tilt and scatter of the FP

We set out to measure the low-redshift stellar mass FP, as its tilt and scatter in the zero point trace $M_{\rm dyn}/M_*$, and therefore reflect the structural properties and assembly of galaxies. By firstly constructing the total mass FP, we have found that although there is significant variation in the structural properties among the simulated galaxy population, the non-homology of this population has only a weak effect on the tilt of the FP. The resulting deviation from the tilt expected under the assumption of homology is strongest for the population of quiescent galaxies, but still $< 10\%$. In other words, within the effective radius the stellar velocity dispersion, measured by taking into account both the random and rotational motions of stars (see Section 5.2.4), provides a good proxy of the circular velocity (although with an offset of $\approx 10\%$). Hence, $M_{\rm dyn} \propto M_{\rm tot}$, which broadly agrees with observational findings based on dynamical modelling or strongly lensing systems (e.g., Cappellari et al. 2006; Bolton et al. 2007, 2008; Li et al. 2018).

Instead, as also suggested by earlier theoretical work (e.g., Boylan-Kolchin et al. 2006; Robertson et al. 2006; Hopkins et al. 2008) and more recently specifically for the EAGLE and IllustrisTNG simulations (Ferrero et al. 2021), we have shown that systematic variations in the dark matter fraction as a smooth function of the size and stellar mass drive the strong, observed deviation of the stellar mass FP from the virial plane. Variations in the IMF can affect the inferred dark matter mass fraction, and thereby contribute to the tilt as well, but have a smaller effect. Interestingly, this is true for both the stellar mass FP of the quiescent and star-forming galaxy populations, and leads to a near identical tilt of and scatter about the two simulated FPs: within the parameter space of the mass, velocity dispersion and size, we can therefore regard these two populations as forming a single distribution.

Observationally, it is difficult to constrain $f_{\rm DM}$, particularly in a way that is independent from the measurement of the FP itself (i.e., through strong lensing), to allow for a robust measurement of potential gradients in $f_{\rm DM}$ along the FP. Mock observations from the simulations therefore are important to estimate the effects of observational uncertainties and test the role of $f_{\rm DM}$ in the interpretation of the observed FP. We have found that simply the effects from the random projection of galaxies on the sky, leading to differences in the measured sizes and the velocity dispersion, alter the measured FP by $> 5\sigma$, although this can be remedied somewhat by adopting a circularised size rather than a major axis size (or, likely, an alternative correction factor that directly corrects σ_* for the random inclination
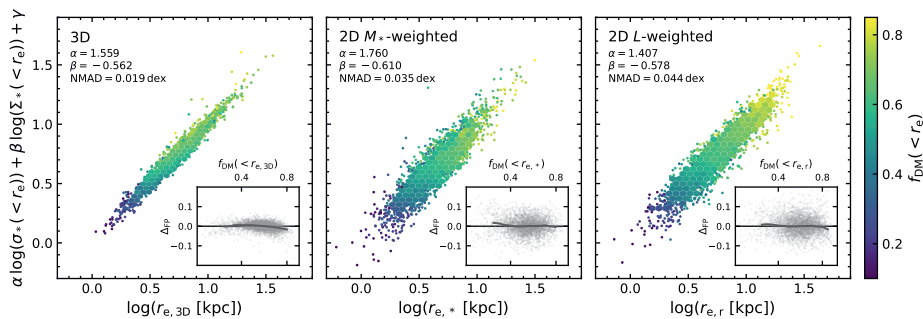
Figure 5.19: Edge-on view of the stellar mass FP for the intrinsic FP (relying on 3D aperture measurements; left), and the mock observed FPs using measurements weighted by the stellar mass (middle) and the observed light (right). The colour scale shows the dark matter fraction within the effective radius, and insets show the offset from the FP as a function of the same dark matter fraction. Systematic variations in $f_{\rm DM}$ across the galaxy population set the tilt of the stellar mass FP. The remaining scatter about the FPs anti-correlates only weakly with $f_{\rm DM}$ (Spearman $\rho = -0.23$), and is therefore driven by random scatter and measurement uncertainties.

angle). Moreover, gradients in M_*/L also have a significant effect on the retrieved FP, bringing very good consistency in the β parameter (associated with Σ_*) with the 3D measurements. On the other hand, the α parameter (associated with σ_*) is biased low and strongly dependent on the sample selection.

The edge-on projections of the intrinsic and mock observed stellar mass FPs are shown in Fig. 5.19, where the colour scale illustrates the correlation with $f_{\rm DM}$ along the FP. In comparison with observational results, we find that the mock observed EAGLE FP broadly agrees with measurements of the stellar mass FP of quiescent galaxies (Hyde & Bernardi 2009; Bernardi et al. 2020) as well as the K-band luminosity FP (La Barbera et al. 2010a; Magoulas et al. 2012) when taking into account differences in the sample selection, as these studies have found $\alpha \approx 1.5 - 1.6$ (and $\alpha \approx 1.9$ for the most massive early-type galaxies, of $M_* > 10^{11}\,{\rm M_\odot}$). We emphasise that the simulation was not explicitly tuned to reproduce the FP – only the stellar mass-size relation was used to reject unrealistic subgrid models (Schaye et al. 2015) – and the inferred FP and the drivers of the relation can therefore be considered to be predictive. We note that, unfortunately, the FP of star-forming galaxies has as of yet not been measured explicitly, and we therefore cannot compare those results with observations.

In detail, however, there are some significant differences between the simulated and observed $z \sim 0$ stellar mass FP of quiescent galaxies. The values for β differ more strongly, with $\beta \approx -0.8$ in observations, whereas the lowest value of β measured in the EAGLE simulations is $\beta = -0.66 \pm 0.04$ for the bottom-heavy IMF model. The top-heavy IMF model on the other hand (with $\beta = -0.37 \pm 0.16$) is disfavoured, although not ruled out given the large uncertainties on our measurements. The physical interpretation of the discrepancy in β is further discussed in Section 5.5.2.2. The measured scatter also differs by a factor ≈ 2,

with the simulated FP having both a lower observed scatter and intrinsic scatter.

The difference in the measured scatter can easily be attributed to the fact that the observed stellar mass FP includes additional scatter due to the measurement uncertainties on the stellar masses. Yet, the nonzero (intrinsic) scatter in the observed stellar mass FP has been suggested to arise from physical effects, namely variations in $f_{\rm DM}$ or the IMF through the thickness of the FP (Graves & Faber 2010). The insets in Fig. 5.19 show the residuals from the simulated FPs as a function of the dark matter fraction: we find only a weak anti-correlation between $\Delta_{\rm FP}$ and $f_{\rm DM}$ (Spearman $\rho = -0.23$) for the 3D FP (left-hand panel), which indicates that galaxies with higher Σ_* than predicted from the FP have slightly lower $f_{\rm DM}$ (since $\Delta \log \Sigma_* \propto \Delta_{\rm FP}$, see Eq. 5.10). Interestingly, we also find a correlation with the co-rotating fraction $\kappa_{\rm co}$ (Section 5.3.2) within the scatter, of $\rho = 0.35$. We do not find any correlations with other bulk galaxy properties, e.g., with mass, velocity dispersion or SFR. This suggests that the scatter in the 3D FP may not be entirely random, but still be partially due to physical effects (e.g., recent mergers). On the other hand, for the mock observed FPs the scatter is predominantly driven by measurement uncertainties. Therefore, the difference found between the intrinsic scatter of the simulated and observed FP may most easily be explained by an underestimation of the uncertainties in M_*/L from the SED modelling. This includes the systematic uncertainty due to the assumed IMF, which we find to lead to an increase of $\approx 10\%$ in the measured scatter in the simulated FPs.

Despite the differences between simulated and observed FP, it is interesting to explore the physical origins of the simulated relation. Of course, although we have demonstrated that the variation in $f_{\rm DM}$ can largely explain the tilt of the FP, with observational uncertainties and biases muddying the picture, the quantity $f_{\rm DM}$ in itself is merely a consequence of other factors. From Fig. 5.10 we see that, at fixed M_*, the scatter in σ_* is relatively small, but there is strong variation in $r_{\rm e}$. The variation in $r_{\rm e}$ correlates with $f_{\rm DM}$ (Fig. 5.19), and is in line with the suggestion by Ferrero et al. (2021) that galaxy size is the main differentiating parameter between galaxies. The question of which physical mechanisms drive the FP can therefore be recast as: what causes the scatter in $r_{\rm e}$ at fixed M_*?

This effectively reduces the FP to a 2D scaling relation, namely the stellar mass-size relation, and the question of which galaxy properties show correlations along this plane. Observational studies have shown that galaxy size correlates with colour, age, metallicity and α-element enhancement at fixed mass (Franx et al. 2008; Scott et al. 2017; Barone et al. 2020, 2022), and suggested an additional dependence on the structural properties (Sérsic index, level of rotational support) and environment. For instance, at fixed stellar mass, more compact galaxies have been found to be older and to have higher metallicities, as well as greater α-element enhancements. However, these measurements are difficult to interpret, as the stellar population properties are typically luminosity-weighted, global quantities, and therefore are difficult to relate to the overall star formation histories and merger histories of galaxies.

Cosmological simulations may offer valuable insight here, as they allow to trace individual particles within the formation history of a galaxy, and therefore to dis-

tinguish between the in-situ and ex-situ growth of galaxies. Furlong et al. (2017) demonstrated that there is a dependence on the sSFR across the stellar mass-size plane for star-forming galaxies in the EAGLE 100 Mpc simulation, with larger galaxies having higher sSFR at fixed M_*, and a correlation with the mass assembly timescale for simulated quiescent galaxies, such that more extended quiescent galaxies assembled later in cosmic time. Rosito et al. (2019a) showed that this also translates to observable measurements, finding trends with not just age and metallicity, but also radial gradients therein, across the dynamical mass-size plane in the EAGLE simulations, and suggest that this in turn correlates with the stellar spin parameter. Therefore, combining these different ideas, investigating the in-situ and ex-situ growth and associated timescales across the stellar mass-size plane and linking these to observable measures may deliver powerful insight into the assembly of galaxies and the physical origins of the tilt of the FP, for simulated, and likely also for observed, populations of galaxies.

5.5.2 Reconciling the FP with the TF relation

We have found that, at least intrinsically, the simulated star-forming galaxies lie on a total mass FP and stellar mass FP that are approximately the same as the FPs spanned by the quiescent population, and with equally low scatter. On the other hand, star-forming galaxies have been shown to obey the TF relation in both observations and the EAGLE simulations (e.g., Tully & Fisher 1977; Schaye et al. 2015; Ferrero et al. 2017). These two findings may appear to be contradictory, as the FP is explicitly dependent on surface brightness, yet, extensive literature has shown that the TF relation does not correlate with a third parameter (e.g., Zwaan et al. 1995; Courteau & Rix 1999; Meyer et al. 2008; Lelli et al. 2019).

5.5.2.1 Star-forming galaxies can simultaneously obey the FP and TF relation

There is a slight difference between the two relations in the measure of the kinematics that is used: whereas the FP takes the spatially-integrated velocity dispersion, the TF relation uses the inclination-corrected rotational velocity. Although different in nature, both serve as a proxy for the circular velocity, because the integrated velocity dispersion accounts for both the disordered motion and rotation of the stars (see Section 5.2.4). Furthermore, there is a difference in the aperture that is considered: rotational velocities are often measured in the outskirts of the disk, whereas the FP probes the effective radius or even smaller radii. The stellar mass or baryonic mass TF relation, described by $M \propto v_c^\mu$, has a slope of $\mu \approx 3-4$ depending on the aperture chosen (e.g., see Lelli et al. 2019).

If we rewrite the FP in a form that is closer to that of the TF relation, we obtain

$$M \propto \sigma^{-\alpha/\beta} r_e^{(1+2\beta)/\beta}.$$ (5.14)

Focusing on the 3D measurements of the size, stellar velocity dispersion and the total mass, this results in

$$M_{\rm tot} \propto \sigma_*^{1.94} r_{\rm e,3D}^{0.96},$$ (5.15)

for the M_*-selected star-forming population (Table 5.1). As expected from Section 5.3.2, this is very close to the viral relation, which indicates that the star-forming galaxies form a near-homologous sample, and corresponds roughly to the group of galaxies with $T \approx 0$, $\epsilon_* \approx 0.7$ and $\kappa_{\rm co} \approx 0.6$ in Fig. 5.9. Although σ_* has a small systematic offset from $v_{\rm c}$ in this parameter range, because of the small impact of non-homology for this sample we can conclude that $\sigma_* \propto v_{\rm c}(r_{\rm e})$, provided that σ_* does not suffer from projection effects. Hence,

$$M \propto \sigma_*^{-\alpha/\beta} r_{\rm e}^{(1+2\beta)/\beta} \propto v_{\rm c}^{\mu}(r_{\rm e}) r_{\rm e}^{\nu} . \tag{5.16}$$

If we now set $M = M_*$, and continue with the 3D measurements to avoid projection effects on σ_*, we obtain (using the bottom rows of Table 5.1)

$$M_*(< r_{\rm e,3D}) \propto v_{\rm c}^{2.85}(r_{\rm e,3D}) r_{\rm e,3D}^{0.21} , \tag{5.17}$$

which is effectively the TF relation, as $\nu \ll \mu$. Moreover, this value of μ corresponds very well to observationally measured values: e.g., using integral field unit (IFU) data from the SAMI Survey, Bloom et al. (2017) found $\mu^{-1} = 0.31 \pm 0.09$ ($\mu \approx 3.2$), whereas Lelli et al. (2019) found $\mu = 3.06 \pm 0.08$ for the baryonic TF relation of galaxies in the SPARC dataset using their smallest aperture of $1.3r_{\rm e}$. We stress that the choice of aperture is critical here, as Lelli et al. (2019) show, using the exact same sample, that the velocity of the flat part of the rotation curve results in a slope of $\mu = 3.85 \pm 0.09$.

Star-forming galaxies in EAGLE are thus simultaneously compatible with the stellar mass FP and the stellar mass TF relation. We indeed find that the scatter is lower for the FP than the TF relation, finding an orthogonal scatter of 0.0221 ± 0.0005 dex for the TF relation obtained from an orthogonal linear fit to the stellar masses and circular velocities calculated in Section 5.3.2 (with a best-fit slope $\mu = 3.24 \pm 0.03$). This is a marginal, although statistically significant, difference of 0.0035 ± 0.0006 dex with respect to the scatter in the stellar mass FP, and may help to explain why observational studies might not find a correlation with size within the TF relation.

It is unclear, however, to which extent this result can be translated to the observed FP and TF relation. The tilt of the simulated FP differs from observations, and indicates that there may be fundamental discrepancies between the (dynamical) structures of simulated and observed galaxies. As is discussed more extensively in Section 5.5.2.2, we expect this to be of particular importance for quiescent galaxies, but the star-forming galaxies may likely also be affected.

Nevertheless, we can examine why the TF arises from the FP in the simulations, by dividing Eq. 5.17 by Eq. 5.15:

$$\frac{M_*}{M_{\rm tot}} \propto \sigma_*^{0.91} r_{\rm e,3D}^{-0.75} , \tag{5.18}$$

which under the assumption of homology, $\sigma_* \approx v_{\rm c}$ and $M_{\rm tot} \approx M_{\rm dyn}$, reduces to

$$\frac{M_*}{M_{\rm tot}} \propto M_{\rm tot}^{0.46} r_{\rm e,3D}^{-1.21} . \tag{5.19}$$
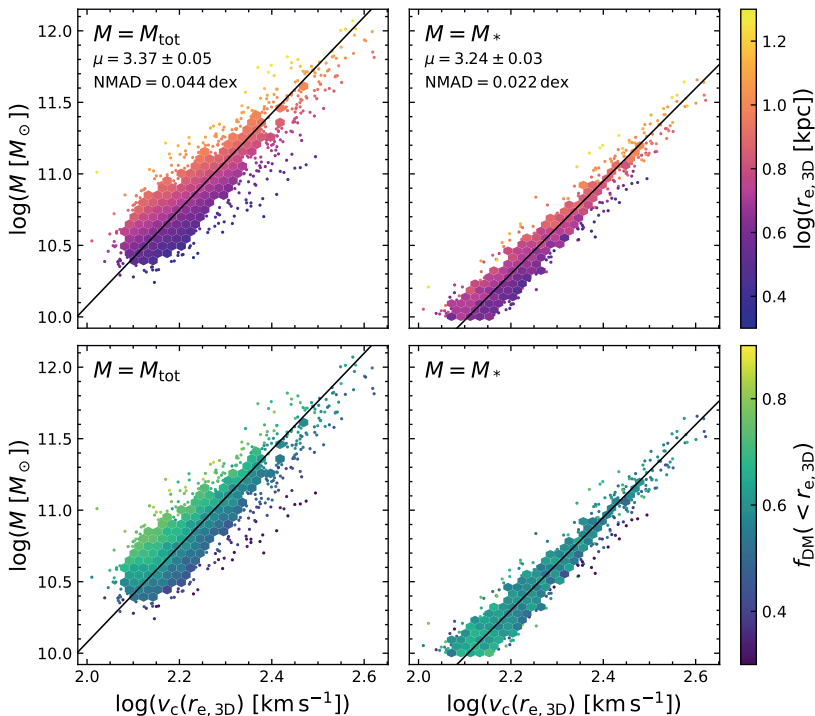
Figure 5.20: Tully-Fisher relation of star-forming galaxies, using the total mass (left) and stellar mass (right) and circular velocities calculated with Eq. 5.11. Solid lines show the relations obtained with orthogonal distance regression, and the orthogonal scatter is printed in each panel. The offset from the total mass TF relation correlates with the half-mass radius (top), and in turn also with the dark matter fraction (bottom). Because of the correlation between $r_{e,3D}$ and f_{DM}, the stellar mass TF relation is nearly as tight as the stellar mass FP.

This relation can be interpreted in terms of the stellar and dark matter density profiles (provided that the gas fractions are low). At fixed M_{tot} the stellar-to-total mass ratio decreases strongly with radius, which implies that the dark matter fraction rises rapidly. Ferrero et al. (2021) showed that the half-mass radii of star-forming EAGLE galaxies lie in the dark matter-dominated regions of galaxies, i.e., $r_{e,3D} > r_c$ with the 'critical radius' (r_c) defined as $M_{DM}(< r_c) = M_*(< r_c)$. As a result, M_*/M_{tot} depends mainly on the mass profile of the dark matter, $M_{DM}(< r)$, which increases monotonically with radius for a NFW profile. The small scatter in the stellar-halo mass relation ($\approx 0.15\,\text{dex}$ for galaxies in the mass range considered here; Matthee et al. 2017) implies that galaxies of fixed M_* have similar M_{DM} profiles. Variations in $r_{e,3D}$ therefore are largely responsible for the slope of the TF relation: Fig. 5.20 shows that although there is no tight TF relation between M_{tot} and v_c, there is a tight relation between M_* and v_c due to the near-perfect correlation between $r_{e,3D}$ and f_{DM} at fixed v_c.

Moreover, there is a weaker dependence on M_{tot}, such that at fixed size, more massive galaxies are relatively more baryon-dominated within $r_{\mathrm{e,3D}}$. This overall scaling with mass thus reflects the compactness of the stellar mass distribution, and depends on the mass assembly history of galaxies, e.g., through variations in the star formation efficiency within $r_{\mathrm{e,3D}}$ or the merger history. For star-forming galaxies, this possibly reflects the build-up of central bulges in more massive galaxies.

5.5.2.2   A TF relation for quiescent galaxies?

The same reasoning of the previous section can be applied to the quiescent galaxy population. Although the effects of non-homology are stronger for this population, we can obtain a relation similar to Eq. 5.17 for the quiescent population: $M_*(< r_{\mathrm{e,3D}}) \propto \sigma_*^{2.68} r_{\mathrm{e,3D}}^{0.23}$. This form is close to that of the Faber-Jackson relation (FJ; Faber & Jackson 1976), which is the linear scaling relation between the mass (or luminosity) and velocity dispersion for early-type galaxies.

However, whereas the r_{e} dependence found here is just as weak as for the star-forming galaxies, observational studies have shown that the scatter in the FJ relation correlates significantly with galaxy size, therefore motivating the use of the FP (e.g., Djorgovski & Davis 1987; Dressler et al. 1987). This mismatch between the observed and simulated FJ relation suggests that either projection effects on the observed σ_* depend on galaxy size, or, more plausibly, that there is a discrepancy between the simulated and observed FP.

To achieve the weak dependence of M_* on r_{e}, requires that the tilt of the FP $\beta \approx -0.5$ with no strong restriction on α (see Eq. 5.14). Throughout, we have found $\beta \approx -0.56$ with minimal variation, despite significant effects from random inclination angles and M_*/L gradients, and different sample selection effects, and therefore describes a FP that can be easily reconciled with the TF relation.

Yet, observational studies of early-type galaxies at $z \sim 0$ have measured a different tilt, with $\beta = -0.84 \pm 0.02$ or $\beta = -0.776 \pm 0.019$ for the stellar mass FP (Hyde & Bernardi 2009; Bernardi et al. 2020). Similarly, for the luminosity FP, which in principle may differ slightly in the tilt due to M_*/L variations, measurements have consistently resulted in $\beta \approx -0.8$ for early-type galaxies (e.g., Jorgensen et al. 1996; La Barbera et al. 2010a; Cappellari et al. 2013b). These measurements all point to a much weaker dependence of M_*/M_{tot} on size, i.e., we would expect Fig. 5.12 (showing f_{DM} in the stellar mass-size plane) to look very different for observed galaxies. This suggests that there are either differences in the dark matter density profiles, or differences in the stellar mass density profiles with respect to the simulated galaxies.

Other theoretical studies of the FP using cosmological simulations have noted a similar systematic discrepancy in the tilt: although these use different sample selections (i.e., a selection of early-type galaxies by the Sérsic index or kinematic structure), measurements, and fitting methods, Lu et al. (2020) found that the luminosity FP in the IllustrisTNG-100 simulation has a tilt of $\beta = -0.63$ and Rosito et al. (2021) reported $\beta = -0.54$ for the stellar mass FP in the Horizon-AGN simulation. Moreover, Ferrero et al. (2021) showed that the FJ relation

of early-type galaxies is systematically offset from the observed relation in both IllustrisTNG and EAGLE.

As also proposed by the aforementioned studies, the difference between the observed and simulated quiescent galaxy population is likely related to the fact that the sizes of simulated galaxies are systematically too large in comparison with observations, and, correspondingly, the velocity dispersions too low (e.g., Genel et al. 2018; Rodriguez-Gomez et al. 2019; van de Sande et al. 2019; de Graaff et al. 2022). In addition, the morphological properties also differ from observations, with star-forming galaxies not being flattened enough, and quiescent galaxies not being sufficiently round and having Sérsic indices that are too low. This implies that the stellar mass distributions diverge from real galaxies, which can be caused by several effects, such as the limited resolution, the gas pressure floor imposed in the simulation, or the details of the star formation and feedback prescriptions in the subgrid models, including the choice of the adopted IMF.

Interestingly, the three cosmological simulations employ different subgrid models, yet all result in a similarly divergent mass-size relation and FP. The simulations do have similar resolutions, and EAGLE and Illustris-TNG both use a pressure floor with an associated spatial scale of $\approx 1\,\mathrm{kpc}$, which likely affects the stellar mass density profiles. An increased resolution for the EAGLE simulations leads to improved Sérsic indices (i.e., more realistic 1D profiles) and smaller half-mass radii, but similar 3D shapes and velocity dispersions (Appendix 5.A; Thob et al. 2019). Despite the improvements, the obtained stellar mass FP is similar to that in the simulations at standard resolution. Therefore, either due to the pressure floor or inaccuracies in the subgrid model (e.g., the implementation of the feedback processes, or the IMF), simulated quiescent galaxies do not obtain the correct shapes and dynamical properties.

We suggest that, as a result, the 3D stellar mass distributions are too 'puffy', with sizes that are larger than observed or ellipticities that are lower than observed, and thereby containing relatively more dark matter within the effective radius. Even for the EAGLE model that assumes a bottom-heavy IMF, evidence for which has been found in low-redshift early-type galaxies (e.g., van Dokkum & Conroy 2010; Auger et al. 2010), the inferred tilt deviates by $> 3\sigma$ from observations and is likely too dark matter-dominated. In the terminology of the previous section, this would mean that the effective radii of quiescent galaxies are not small enough in comparison with their critical radii.

On the other hand, Mukherjee et al. (2022) performed a strong lensing analysis to compare the inferred dark matter fractions of massive early-type galaxies in simulations and observations, and found good agreement between the projected measurement of $f_{\mathrm{DM}}(< r_{\mathrm{e}})$ from EAGLE and observed lenses, and only a slight discrepancy for the smaller aperture of $f_{\mathrm{DM}}(< r_{\mathrm{e}}/2)$. However, their study focuses on the most massive galaxies in EAGLE ($M_* \gtrsim 10^{11}\,\mathrm{M_\odot}$), thereby probing a different regime than considered here. Moreover, as we showed, even though the 1D mass profiles may appear realistic, the 3D structures of the dark matter and stellar mass can still differ and lead to systematic differences in σ_* between observations and simulations. Although relying on projected measurements, the FP is sensitive to the 3D structure, and therefore the stellar mass distribution relative to the dark
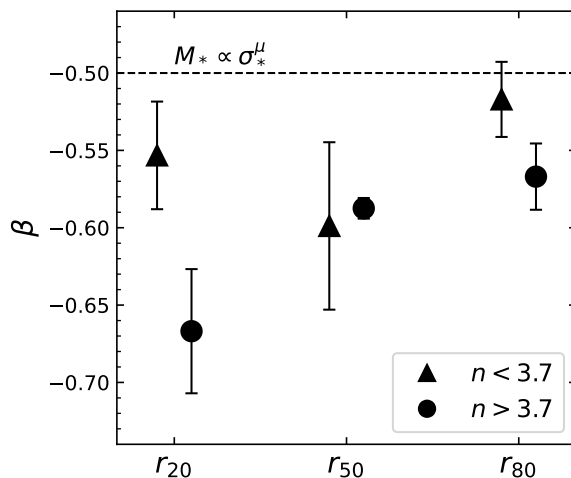
Figure 5.21: Tilt parameter β of the stellar mass FP as a function of the aperture size, for galaxies with low and high Sérsic indices in the high-resolution EAGLE simulation. The line of $\beta = -0.5$ implies no dependence of the stellar mass on size, i.e., the existence of a perfect TF or FJ relation. There is a weak trend toward this value of β for larger aperture sizes, indicating that the dark matter content plays an increasingly important role in the tilt of the stellar mass FP.

matter mass distribution.

If our assertion, that the greater relative importance of the dark matter in the simulated galaxies affects the inferred FP, is correct, then we would expect to find a dependence of the FP tilt on the chosen aperture. So far, we set this aperture to be half of the enclosed mass or light. Miller et al. (2019) proposed the use of the radii enclosing 20% (r_{20}) and 80% (r_{80}) of the light (or stellar mass) distribution instead of the half-light (half-mass) radius, as these sizes are suggested to be more closely linked to the star formation history and halo mass, respectively.

We use the high-resolution, $25^3\,\mathrm{cMpc^3}$ EAGLE simulation to measure the 3D stellar mass r_{20} and r_{80} radii (where the percentiles are calculated using the stellar mass enclosed within a spherical aperture of radius $100\,\mathrm{kpc}$), and measure the stellar velocity dispersions within the same apertures. Given the very small number of galaxies with sSFR $< 10^{-11}\,\mathrm{yr^{-1}}$, we instead divide the sample in two equal-sized subsamples of high and low Sérsic indices (split at $n = 3.7$). Fig. 5.21 shows the measured value of the β parameter of the stellar mass FP as a function of the aperture size for the two subsamples. Although a strong conclusion is not possible due to the small sample size and the limited resolution affecting the measurements for r_{20}, there seems to be a trend in the expected direction: larger apertures are more dark matter-dominated, and result in a higher measured value for β. Moreover, high Sérsic index galaxies have systematically lower (more negative) β, in line with the suggestion that galaxy structure is correlated with the dark matter fraction and hence the tilt of the stellar mass FP.

We can therefore conclude that the stellar mass FP offers an important measure of success for the realism of early-type galaxies in cosmological simulations. Lu et al. (2020) similarly proposed the use of the scaling relation between M_{dyn}/L and M_{dyn}, however, to obtain a realistic estimate of L requires significant effort in the post-processing of a simulation (e.g., Trayford et al. 2017). Instead, the stellar mass FP is easily measured, and provides an equally valuable assessment of the 3D stellar mass profile.

Lastly, it is interesting to note that the problem of an inconsistency in the measured tilt occurs mainly at low redshift, as Lu et al. (2020) and Rosito et al. (2021) report no evolution and weak evolution in β with redshift, respectively. At the same time, observational work does show evidence for evolution in the tilt of the FP, with values reported in the range $\beta \approx [-0.7, -0.5]$ at $z \approx 1$ (e.g., Saglia et al. 2010, 2016; Jørgensen & Chiboucas 2013; Jørgensen et al. 2019; Saracco et al. 2020; de Graaff et al. 2021). The change in the tilt may be correlated with the observed structural evolution across the same redshift range, as quiescent galaxies become smaller and more disc-like in shape toward higher redshift (Chang et al. 2013) and with greater rotational support (Bezanson et al. 2018a). This also appears to be supported by the fact that Bernardi et al. (2020) report a systematically higher value of β for low-redshift S0 galaxies than elliptical galaxies.

If the tilt of the FP indeed depends on the average structural properties of the selected galaxy population (e.g., discs versus spheroids), then we would expect star-forming and quiescent galaxies to lie on a single stellar mass FP at higher redshifts (as has been observed; de Graaff et al. 2021), but to span increasingly divergent FPs toward $z \sim 0$. The fact that Bezanson et al. (2015) find star-forming galaxies to be consistent with the stellar mass FP of quiescent galaxies at $z \approx 0.05$, may then be due to the larger scatter from measurement errors and projection effects for the star-forming population (which affect both the measured velocity dispersion and size), or be caused by the small apertures in which the velocity dispersions were measured, probing only the bulge-like centres of star-forming galaxies at low redshifts. Recently completed large IFU surveys of low-redshift galaxies can shed light on whether the stellar mass FP of star-forming galaxies is truly the same as for quiescent galaxies, and simultaneously offers a direct comparison with the TF relation within the same aperture.

5.6 Conclusions

We have used the EAGLE cosmological simulations to measure the tilt and scatter of the stellar mass FP ($r_{\mathrm{e}} \propto \sigma^{\alpha} \Sigma_{*}^{\beta}$) for a mass-selected sample of galaxies at $z = 0.1$ ($M_{*} \gtrsim 10^{10}\,\mathrm{M}_{\odot}$). From measurements of the total and stellar masses and velocity dispersions within 3D spherical apertures defined by the half-mass radii, we have evaluated the different drivers of the simulated FP. By comparing with measurements of the masses, sizes and stellar velocity dispersions obtained from realistic mock observations, we have quantified the effects of observational uncertainties and the sample selection on the inferred scaling relation.

Our results can be summarised as follows:

- We use the measured total masses and 3D stellar velocity dispersions to show that, within the effective radius, the simulated galaxies obey a total mass FP that is very close to the virial relation. The stellar velocity dispersions, which take into account both the random and streaming motions of the stars, thus provide a good approximation for the circular velocities (deviating by $\approx 10\%$), with only a weak effect from the galaxy environment (i.e., the classification into central/satellite systems). Therefore, despite significant variation in the structural properties among the simulated galaxy population, the effects of this non-homology on the simulated FP are weak. The velocity dispersion of all (dark matter, stellar and gas) particles deviates more strongly from the circular velocity, due to the dynamically-hot dark matter particles.

- Replacing the total mass by the stellar mass, we find that star-forming and quiescent galaxies span a nearly identical stellar mass FP within the EAGLE simulations, with equally low scatter (0.019 dex). The stellar mass FP deviates strongly from the virial relation, which is driven by variations in the dark matter fraction within the effective radius ($f_{\rm DM}$), with negligible impact from variations in the gas content. We show that $f_{\rm DM}$ is a smooth function of the size and stellar mass, and therefore sets the tilt of the stellar mass FP. We find that the remaining scatter in this FP anti-correlates only very weakly with $f_{\rm DM}$, and correlates weakly with the degree of rotational support.

- For the star-forming galaxies in the simulations, we demonstrate that they are simultaneously compatible with the stellar mass FP and the linear Tully-Fisher relation, provided that both relations are evaluated within the same aperture of the effective radius. The scatter about the TF relation is only slightly higher (0.022 dex) than the stellar mass FP.

- We create mock observations to show that the projection of galaxies at a random inclination angle along the line of sight affects both the measured sizes and (spatially-integrated) velocity dispersions. These effects can change the inferred tilt of the simulated stellar mass FP by $\approx 10\%$, and increase the scatter by a factor of ≈ 2. When we use luminosity-weighted measurements instead of M_*-weighted measurements, the tilt of the stellar mass FP is changed by a similar amount, but in the opposite direction. The α parameter (associated with σ_*) in particular is highly sensitive to these changes, and also depends strongly on the sample selection. The scatter about the mock FP is further increased by $\approx 30\%$, which we show is caused by the luminosity weighting of the velocity dispersions.

- Systematic uncertainties in the assumed IMF can have a significant effect on the inferred tilt of the stellar mass FP: the parameters of the tilt change by up to $\approx 30\%$ for simulations that employ an observationally-motivated, variable IMF with respect to the stellar mass FP measured for the standard EAGLE model that assumes a universal Chabrier IMF. Nevertheless, we find

that regardless of the adopted IMF, variations in f_{DM}, which are themselves correlated with the IMF, are the main driver of the FP.

- However, although the tilt and scatter of the measured mock FPs broadly agree with observational results, we find significant differences as well. Regardless of the adopted IMF or the resolution of the simulation, the β parameter (associated with Σ_*) differs by $> 5\sigma$ from local observations. The imposed pressure floor in the simulation, or inaccuracies in the subgrid model likely lead to substantial differences in the 3D stellar mass distributions of the simulated galaxies with respect to local observations. In addition, the standard resolution used in the EAGLE simulations leads to Sérsic indices that are too low. The effects of non-homology are therefore possibly also weaker for simulated galaxies than for the real Universe.

Our work indicates that f_{DM} is the dominant factor that sets the properties of the FP in the simulations, which in turn is most likely caused by the large variation in r_{e} at fixed M_*. We have found that the correlations between the variations in f_{DM} and r_{e} at fixed M_* naturally give rise to both a FP and TF relation for star-forming galaxies within the aperture of $1\,r_{\mathrm{e}}$, although it is unclear to what extent this is the result of the systematic discrepancies in the mass distributions of EAGLE galaxies with respect to observations. For the galaxy population as a whole, it further raises the question of which physical mechanisms may be responsible for the variation in r_{e} at fixed M_*, as this may provide valuable insight into the physical origins of dynamical scaling relations such as the FP, as well as the TF relation.

Acknowledgements

Data Availability

The data underlying this article are available through `http://icc.dur.ac.uk/Eagle`. Data products (velocity dispersions, masses) created for this article are available upon reasonable request to the corresponding author.

Appendix

5.A   High-resolution simulation results

In Section 5.3 we investigated the tilt of the FP and showed that there are correlations with the morphological and dynamical properties of galaxies, as well as

their mass compositions (particularly f_{DM}). However, the structural properties of galaxies in the EAGLE simulations have been found to deviate from observed galaxies (e.g., Trayford et al. 2017; van de Sande et al. 2019; de Graaff et al. 2022). Ludlow et al. (2019, 2021) showed that this is at least in part due to the limited resolution of the simulations, as the 2-body scattering of dark matter and baryonic particles causes a dynamical heating of the baryons, which affects the galaxy size and likely also other structural properties.

These effects can be alleviated by increasing the resolution: the high-resolution $25^3\,\mathrm{Mpc}^3$ EAGLE simulation (RecalL0025N0752; Section 5.2.1) therefore provides an important test for the robustness of the conclusions drawn from simulations at the standard resolution. Using this simulation, Thob et al. (2019) showed that their measured structural properties are not affected by a change in the resolution.

We perform a similar test for the Sérsic profile modelling, selecting all (78) galaxies of $M_* > 10^{10}\,\mathrm{M}_\odot$ (within a spherical aperture of radius $30\,\mathrm{kpc}$) in the high-resolution simulation. We also select 71 galaxies from the reference model simulation that was run at standard resolution, and has the same volume and initial conditions as the high-resolution simulation (RefL0025N0376). We follow the methodology described in de Graaff et al. (2022) to fit Sérsic profiles to the projected stellar mass distributions, and show in Fig. 5.22 how the Sérsic indices and projected axis ratios differ between the RefL0025N0376 and RecalL0025N0752 simulations. The increased resolution has a strong effect on the measured Sérsic indices, and largely resolves the previously found discrepancy between the RefL0100N1504 simulation and observations in the local Universe.

On the other hand, the projected axis ratios are largely unchanged, and galaxies in the high-resolution simulation are only slightly rounder. This suggests that the 3D shapes do not depend on the resolution, as also found by Thob et al. (2019). We therefore also assess the effect of resolution on the half-mass radius and the stellar velocity dispersion within this spherical aperture. The distributions are shown in Fig. 5.23: in line with the higher Sérsic indices (indicating more centrally concentrated mass distributions), we find that the half-mass radii are systematically smaller in the high-resolution simulation. However, the velocity dispersions show no dependence on the resolution, and are still smaller than observed (see van de Sande et al. 2019).

Next, we assess the convergence of the results obtained in Section 5.3. We apply the exact same methodology as before, by constructing a sample that is complete in M_{tot} and measuring the different mass components within $r_{\mathrm{e,3D}}$, as well as $\sigma_{\mathrm{tot}}(< r_{\mathrm{e,3D}})$, $\sigma_*(< r_{\mathrm{e,3D}})$ and the differences ($\Delta \log \sigma$) with the predicted velocity dispersion, Eq. 5.11.

Fig. 5.24 presents the equivalent of Figs. 5.8 and 5.9 for the high-resolution simulation, showing the correlation between $\Delta \log \sigma$ and different structural parameters. We show $\Delta \log \sigma_{\mathrm{tot}}$ (filled circles) and $\Delta \log \sigma_*$ (open circles) within the same figure, with solid lines indicating the running medians. Given the small number of objects (79), we omit the separation into central and satellite galaxies. Compared with Fig. 5.8, we find that the correlations for $\Delta \log \sigma_{\mathrm{tot}}$ are weaker, suggesting that the results found before are partially driven by the effects of the limited resolution (likely particularly that of the dark matter particles). On the
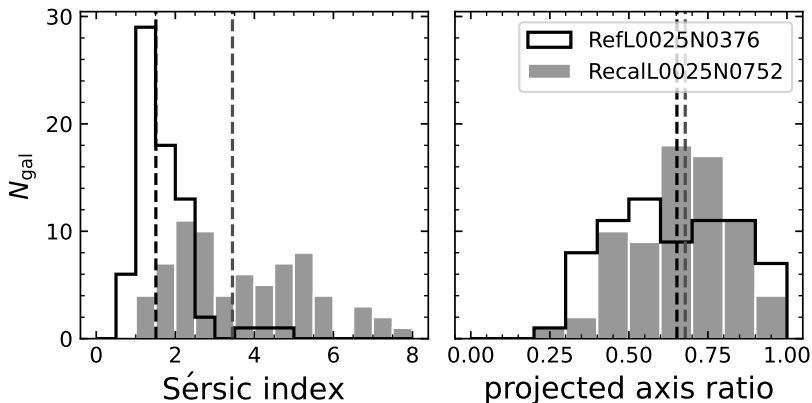
Figure 5.22: Probability density distribution of the Sérsic index (left) and axis ratio (right) of the projected stellar mass distribution for EAGLE galaxies of $M_* > 10^{10}\,\mathrm{M_\odot}$. The black histogram shows the results for the $25^3\,\mathrm{Mpc}^3$ simulation run at standard resolution; the grey histogram shows the distribution for galaxies in the high-resolution simulation. The Sérsic index is strongly dependent on the resolution: the standard resolution does not produce a sufficient number of bulge-like ($n \approx 4$) systems in comparison with observations, which is largely solved by the increased resolution. On the other hand, there is little change in the projected axis ratios, with galaxies in the high-resolution simulation being only slightly rounder.
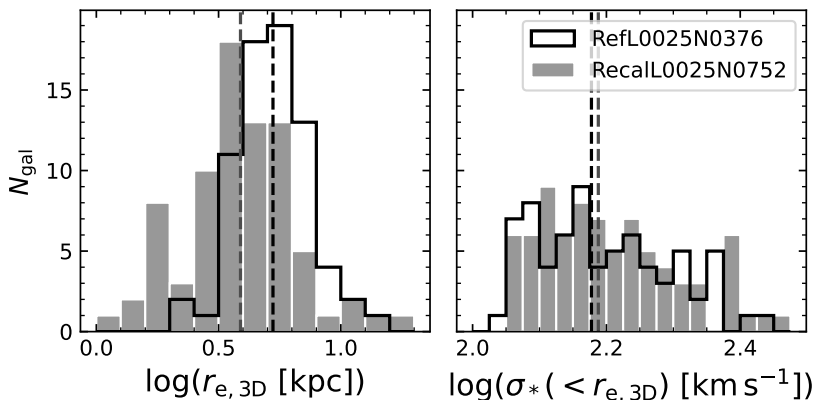
Figure 5.23: Probability density distribution of the 3D stellar half-mass radius (left) and the stellar velocity dispersion within this radius (right) for EAGLE galaxies of $M_* > 10^{10}\,\mathrm{M_\odot}$. The black histogram shows the results for the $25^3\,\mathrm{Mpc}^3$ simulation run at standard resolution; the grey histogram shows the distribution for galaxies in the high-resolution simulation. The increased resolution leads to more compact galaxies, with higher Sérsic indices (Fig. 5.22). However, the velocity dispersions remain unchanged, and are therefore still lower than in local observations.
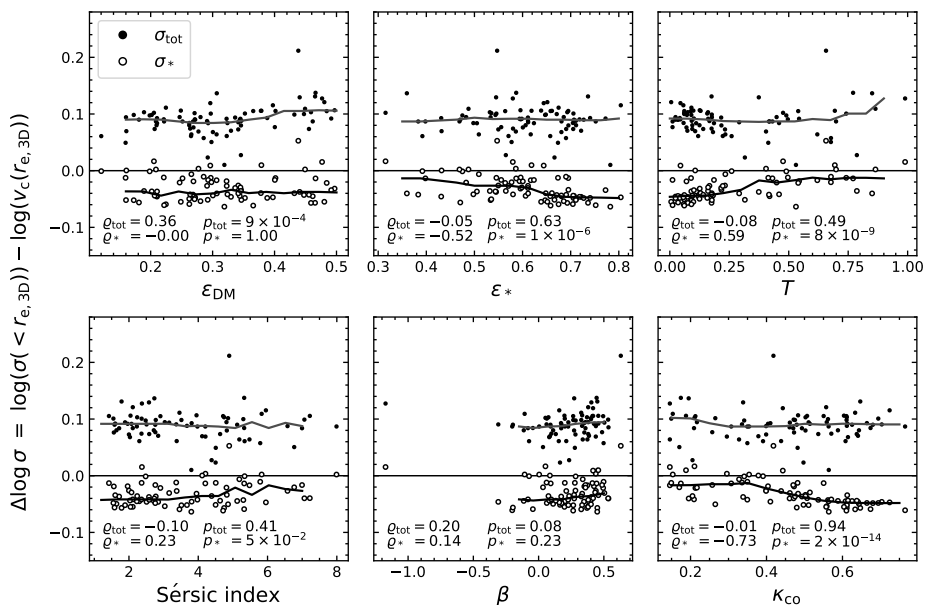
Figure 5.24: Analogous to Fig. 5.8 and 5.9, the deviation in the total (filled symbols) and stellar (open symbols) velocity dispersion from the predicted velocity dispersion of Eq. 5.11 as a function of different structural properties, for galaxies in the high-resolution RecalL0025N0752 simulation. Lines show the running medians in each panel. Although the number of massive galaxies galaxies in the high-resolution simulation is limited, the measured correlations are similar to those found for the standard resolution. Despite a dependence of the galaxy morphology on the simulation resolution (Fig. 5.22), the effects of non-homology shown in Section 5.3.2 are not affected significantly.

other hand, the results for $\Delta \log \sigma_*$ are nearly identical to those found in Fig. 5.9. Our conclusions on the effects of non-homology on the total mass FP therefore are robust to changes in the resolution. Moreover, the systematic offset between σ_{tot} and σ_* remains, which indicates that this is not an effect of the resolution, and gives credence to the interpretation discussed in Section 5.3.2.

Lastly, we use the M_*-selected sample from above to assess the stellar mass FP in the high-resolution simulation. Following Section 5.3.3, we measure $\Delta \log \sigma_*$ using Eq. 5.12 and examine the relation with the dark matter and gas fraction. Fig. 5.25 shows star-forming (blue) and quiescent (red) galaxies separately: we find that the main conclusion, that f_{DM} is the main driver of the stellar mass FP, is unchanged. Unlike the results of Fig. 5.11, we find a weak correlation with the gas fraction for star-forming galaxies, although this is at low statistical significance (p-value of 0.041).

In conclusion, we find good convergence between the results obtained with the simulation at standard and high resolution, despite differences in some of the morphological properties between the two different sets of simulations.
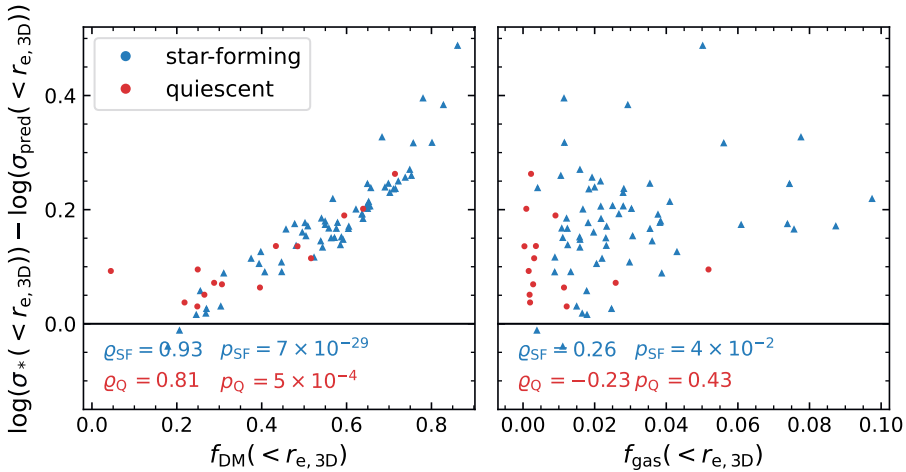
Figure 5.25: Deviation between the measured stellar velocity dispersion and the dispersion predicted from the stellar mass and half-mass radius (Eq. 5.12) versus the dark matter (left) and gas (right) fractions within the half-mass radius, for galaxies in the high-resolution simulation. Unlike Fig. 5.11, there is a weak correlation with the gas fraction for star-forming galaxies (p-value = 0.041). However, the conclusion that the variation in f_{DM} is the primary driver of the stellar mass FP does not depend on the resolution of the simulation.

5.B   Stellar mass FP with circularised sizes

In Section 5.4.1 we presented the stellar mass FP, and showed the relation between the projected major axis size and line-of-sight velocity dispersion for different stellar mass bins (Figs. 5.13 and 5.14). However, observational studies often use circularised sizes rather than major axis sizes, which we showed to result in a FP that is in better agreement with the intrinsic stellar mass FP, as the circularised sizes provide an ad hoc correction for the random inclination angles of galaxies.

In Figs. 5.26 and 5.27 we show the circularised size as a function of the line-of-sight velocity dispersion, binned by the stellar mass, for the stellar mass-weighted and luminosity-weighted measurements, respectively. These differ from Figs. 5.13 and 5.14 by only the measure of size used. The velocity dispersions are unchanged, as these are spatially-integrated measurements within elliptical apertures (see Section 5.2.4).

As is to be expected, the circularised sizes are smaller than the major axis sizes, with an average offset of $-0.1\,\mathrm{dex}$. Most importantly, however, the scatter in $\log r_{\mathrm{circ}}$ changes as well: there is a wide spread in the distribution of the projected axis ratios (de Graaff et al. 2022), and the circularised size can differ from the major axis size by $\approx -0.35\,\mathrm{dex}$ for a galaxy that is projected edge-on ($q \approx 0.2$). At the same time, from the top panel of Fig. 5.3 we can see that, at fixed intrinsic dispersion $\sigma_*(< r_{\mathrm{e,3D}})$, the observed line-of-sight velocity dispersion is a factor
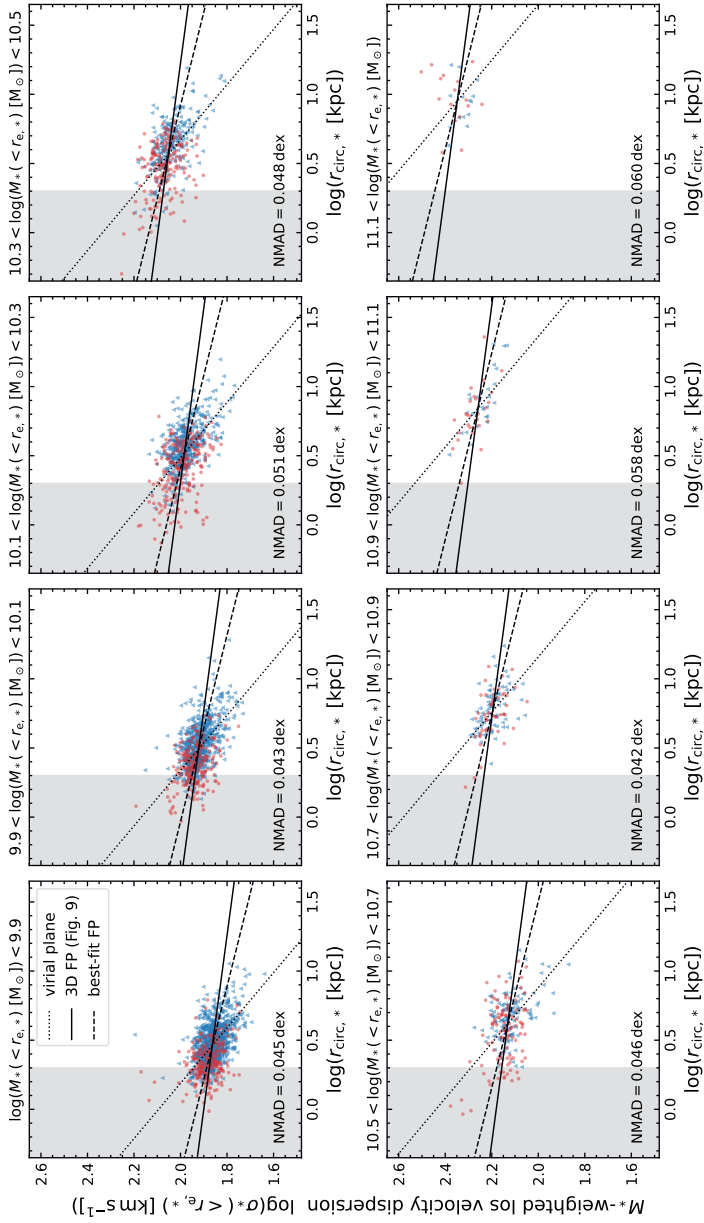
Figure 5.26: As Fig. 5.13, but showing the circularised effective radius rather than the major axis size (obtained from the best-fit Sérsic profile to the projected stellar mass distributions).
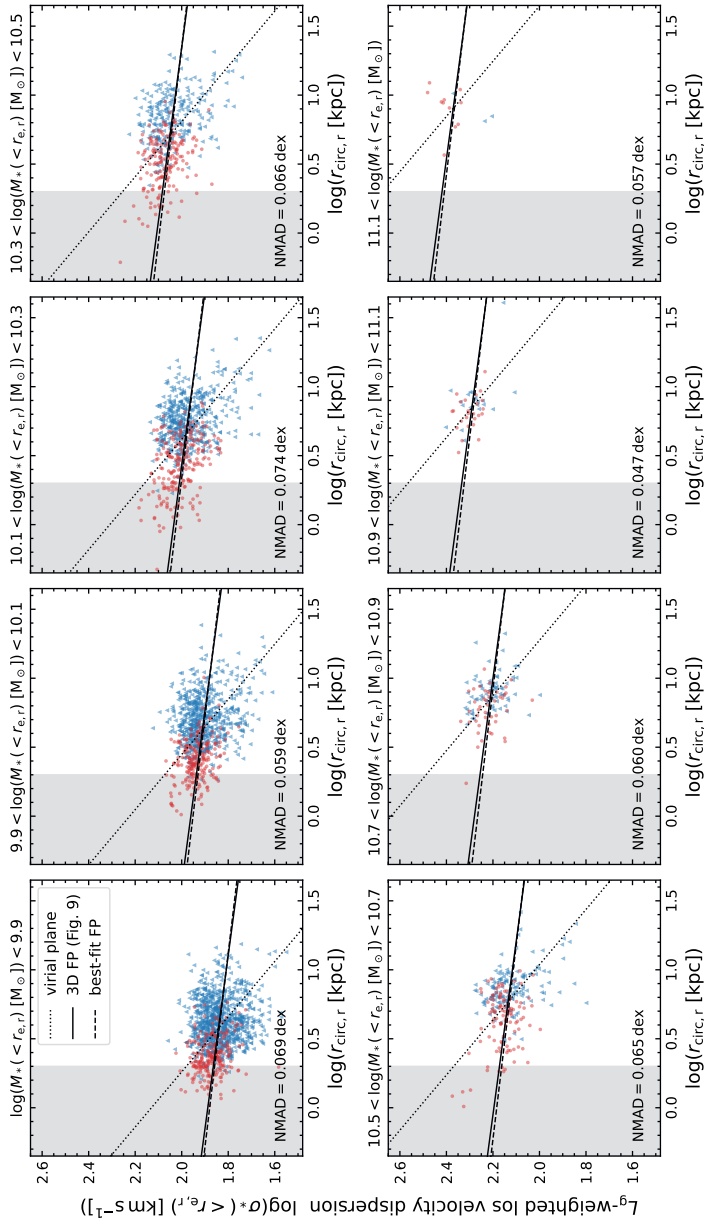
Figure 5.27: As Fig. 5.14, but showing the circularised effective radius rather than the major axis size (obtained from the best-fit Sérsic profile to the mock r-band images).

≈ 2 greater for edge-on systems in comparison with face-on systems. As $\log \sigma_*$ is unchanged in the fits of the stellar mass FP and in Figs. 5.26 and 5.27, it is this change in the scatter in $\log r_{\rm circ}$ that counteracts the projection effects on $\log \sigma_*$ and hence alters the inferred FP.

5.C   Effects of luminosity-weighting on the FP

In Section 5.4.1 we found that the mock observations of the sizes and velocity dispersions introduce significant scatter in the FP. The M_*-weighted mock sizes and velocity dispersions shown in Fig. 5.13 indicate that this is likely due to the random projection of galaxies along the line of sight, as well as the uncertainties on the half-mass radii, as these were measured from mock images with realistic noise and PSF smoothing.

However, in Fig. 5.14 we found that the use of luminosity-weighted measurements further increases the scatter by $\approx 50\%$ for the less massive galaxies, despite the fact that these measurements were extracted using the exact same methodology. Whereas the measurements in Fig. 5.13 are both weighted by M_*, the measurements in Fig. 5.14 use slightly different tracers: the sizes were measured from r-band images that include the effects of dust attenuation (Section 5.2.2), but the velocity dispersions were measured using the unattenuated g-band luminosities of the stellar particles, which are spatially-integrated measurements within elliptical apertures defined by the r-band Sérsic profiles.

Therefore, we explore whether the inconsistency in the tracer used causes the strong increase in the scatter. We create images of the unattenuated rest-frame g-band light, and follow the methodology described in de Graaff et al. (2022) to construct mock images and fit Sérsic profiles. Next, we use these g-band Sérsic profiles to construct elliptical apertures, and hence obtain consistent, spatially-integrated velocity dispersions that are weighted by the g-band luminosities of the particles.

We show the resulting relation between the g-band half-light radii (major axis sizes) and velocity dispersions in Fig. 5.28. The scatter is approximately equal to the scatter found in Fig. 5.14 that used the inconsistent, luminosity-weighted measurements. We therefore conclude that it is the luminosity-weighting itself that leads to an increase in the scatter, rather than the difference between the tracers. This is likely caused by the fact that the younger stellar particles (which have low M_*/L) are dynamically colder, and the line-of-sight velocity dispersion is therefore more strongly dependent on the inclination angle of the galaxy.
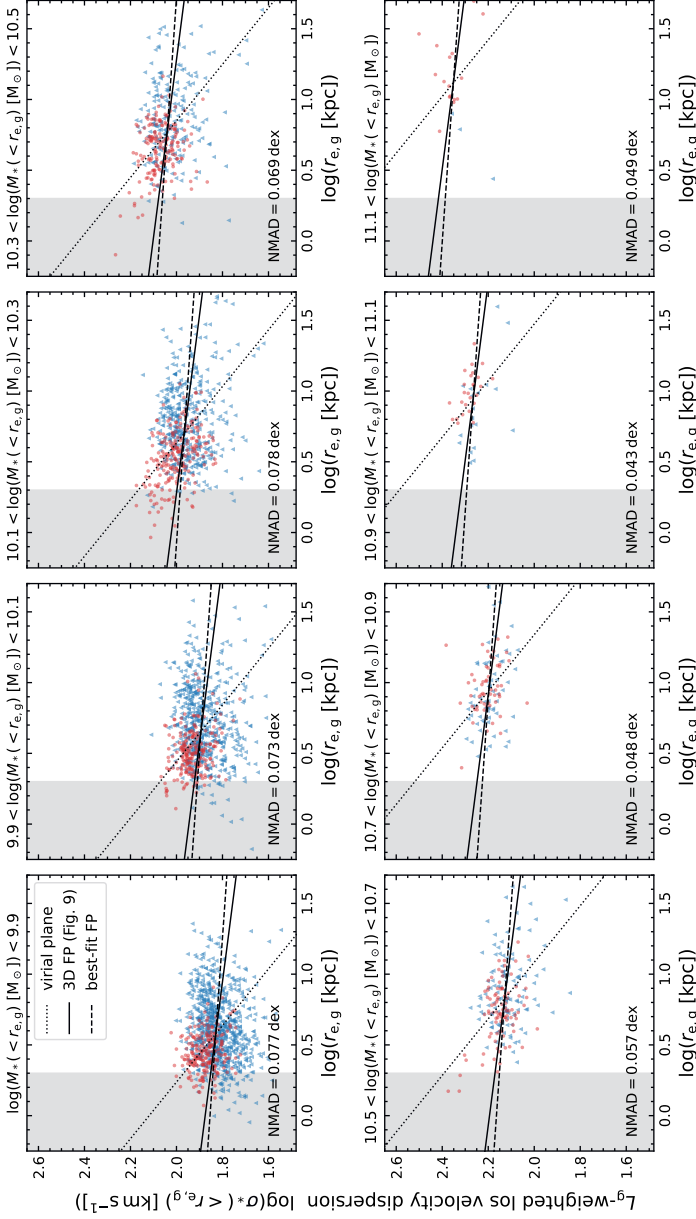
Figure 5.28: As Fig. 5.14, but using a consistent tracer (the rest-frame g-band luminosity) for both the sizes and velocity dispersions: the effective radii are measured from g-band images that do not include the effects of dust attenuation using the methodology described in de Graaff et al. (2022), and the spatially-integrated stellar velocity dispersions are measured within elliptical apertures defined by these unattenuated g-band half-light radii. The strong increase in the scatter between Figs. 5.13 and 5.14 is therefore not due to an inconsistency in the tracers within Fig. 5.14. Rather, it is the luminosity-weighting itself that causes a significant increase in the scatter.

Bibliography

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, ApJS, 182, 543

Ahn, C. P., Alexandroff, R., Allende Prieto, C., et al. 2012, ApJS, 203, 21

Aquino-Ortíz, E., Sánchez, S. F., Valenzuela, O., et al. 2020, ApJ, 900, 109

Aquino-Ortíz, E., Valenzuela, O., Sánchez, S. F., et al. 2018, MNRAS, 479, 2133

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33

Auger, M. W., Treu, T., Gavazzi, R., et al. 2010, ApJ, 721, L163

Baes, M., Davies, J. I., Dejonghe, H., et al. 2003, MNRAS, 343, 1081

Baes, M., Verstappen, J., De Looze, I., et al. 2011, ApJS, 196, 22

Baldry, I. K., Liske, J., Brown, M. J. I., et al. 2018, MNRAS, 474, 3875

Barber, C., Crain, R. A., & Schaye, J. 2018, MNRAS, 479, 5448

Barden, M., Häußler, B., Peng, C. Y., McIntosh, D. H., & Guo, Y. 2012, MNRAS, 422, 449

Barone, T. M., D'Eugenio, F., Colless, M., & Scott, N. 2020, ApJ, 898, 62

Barone, T. M., D'Eugenio, F., Scott, N., et al. 2022, MNRAS, 512, 3828

Beifiori, A., Mendel, J. T., Chan, J. C. C., et al. 2017, ApJ, 846, 120

Bell, E. F., McIntosh, D. H., Katz, N., & Weinberg, M. D. 2003, ApJS, 149, 289

Belli, S., Newman, A. B., & Ellis, R. S. 2017, ApJ, 834, 18

Bender, R., Burstein, D., & Faber, S. M. 1992, ApJ, 399, 462

Benitez-Llambay, A. 2015, py-sphviewer: Py-SPHViewer v1.0.0

Bennett, C. L., Larson, D., Weiland, J. L., et al. 2013, ApJS, 208, 20

Bernardi, M., Domínguez Sánchez, H., Brownstein, J. R., Drory, N., & Sheth, R. K. 2019, MNRAS, 489, 5633

Bernardi, M., Domínguez Sánchez, H., Margalef-Bentabol, B., Nikakhtar, F., & Sheth, R. K. 2020, MNRAS, 494, 5148

Bernardi, M., Sheth, R. K., Dominguez-Sanchez, H., et al. 2018, MNRAS, 477, 2560

Bertin, E. & Arnouts, S. 1996, A&AS, 117, 393

Bezanson, R., Franx, M., & van Dokkum, P. G. 2015, ApJ, 799, 148

Bezanson, R., van der Wel, A., Pacifici, C., et al. 2018a, ApJ, 858, 60

Bezanson, R., van der Wel, A., Straatman, C., et al. 2018b, ApJ, 868, L36

Bezanson, R., van Dokkum, P. G., Tal, T., et al. 2009, ApJ, 697, 1290

Bezanson, R., van Dokkum, P. G., van de Sande, J., et al. 2013, ApJ, 779, L21

Bignone, L. A., Pedrosa, S. E., Trayford, J. W., Tissera, P. B., & Pellizza, L. J. 2020, MNRAS, 491, 3624

Binney, J. & Tremaine, S. 1987, Galactic dynamics

Bland-Hawthorn, J. & Gerhard, O. 2016, ARA&A, 54, 529

Blanton, M. R., Bershady, M. A., Abolfathi, B., et al. 2017, AJ, 154, 28

Blanton, M. R., Hogg, D. W., Bahcall, N. A., et al. 2003, ApJ, 594, 186

Blanton, M. R. & Roweis, S. 2007, AJ, 133, 734

Bloom, J. V., Croom, S. M., Bryant, J. J., et al. 2017, MNRAS, 472, 1809

Bolton, A. S., Burles, S., Treu, T., Koopmans, L. V. E., & Moustakas, L. A. 2007, ApJ, 665, L105

Bolton, A. S., Treu, T., Koopmans, L. V. E., et al. 2008, ApJ, 684, 248

Bottrell, C., Torrey, P., Simard, L., & Ellison, S. L. 2017a, MNRAS, 467, 1033

Bottrell, C., Torrey, P., Simard, L., & Ellison, S. L. 2017b, MNRAS, 467, 2879

Boylan-Kolchin, M., Ma, C.-P., & Quataert, E. 2006, MNRAS, 369, 1081

Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503

Brammer, G. B., Whitaker, K. E., van Dokkum, P. G., et al. 2009, ApJ, 706, L173

Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, MNRAS, 351, 1151

Brook, C. B., Governato, F., Roškar, R., et al. 2011, MNRAS, 415, 1051

Bruzual, G. & Charlot, S. 2003, MNRAS, 344, 1000

Bryan, S. E., Mao, S., Kay, S. T., et al. 2012, MNRAS, 422, 1863

Bundy, K., Bershady, M. A., Law, D. R., et al. 2015, ApJ, 798, 7

Bundy, K., Treu, T., & Ellis, R. S. 2007, ApJ, 665, L5

Burstein, D., Faber, S. M., & Dressler, A. 1990, ApJ, 354, 18

Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682

Camps, P. & Baes, M. 2015, Astronomy and Computing, 9, 20

Camps, P., Trayford, J. W., Baes, M., et al. 2016, MNRAS, 462, 1057

Cappellari, M. 2016, ARA&A, 54, 597

Cappellari, M. 2017, MNRAS, 466, 798

Cappellari, M., Bacon, R., Bureau, M., et al. 2006, MNRAS, 366, 1126

Cappellari, M. & Emsellem, E. 2004, PASP, 116, 138

Cappellari, M., McDermid, R. M., Alatalo, K., et al. 2013a, MNRAS, 432, 1862

Cappellari, M., Scott, N., Alatalo, K., et al. 2013b, MNRAS, 432, 1709

Chabrier, G. 2003, PASP, 115, 763

Chan, J. C. C., Beifiori, A., Mendel, J. T., et al. 2016, MNRAS, 458, 3181

Chang, Y.-Y., van der Wel, A., da Cunha, E., & Rix, H.-W. 2015, ApJS, 219, 8

Chang, Y.-Y., van der Wel, A., Rix, H.-W., et al. 2013, ApJ, 773, 149

Charlot, S. & Fall, S. M. 2000, ApJ, 539, 718

Chevance, M., Weijmans, A.-M., Damjanov, I., et al. 2012, ApJ, 754, L24

Clauwens, B., Schaye, J., & Franx, M. 2016, MNRAS, 462, 2832

Clauwens, B., Schaye, J., Franx, M., & Bower, R. G. 2018, MNRAS, 478, 3994

Conroy, C. 2013, ARA&A, 51, 393

Conroy, C., Gunn, J. E., & White, M. 2009, ApJ, 699, 486

Conselice, C. J. 2014, ARA&A, 52, 291

Correa, C. A., Schaye, J., Clauwens, B., et al. 2017, MNRAS, 472, L45

Cortese, L., Fogarty, L. M. R., Ho, I. T., et al. 2014, ApJ, 795, L37

Courteau, S. & Rix, H.-W. 1999, ApJ, 513, 561

Cowie, L. L., Songaila, A., Hu, E. M., & Cohen, J. G. 1996, AJ, 112, 839

Crain, R. A., Schaye, J., Bower, R. G., et al. 2015, MNRAS, 450, 1937

da Cunha, E., Charlot, S., & Elbaz, D. 2008, MNRAS, 388, 1595

Darvish, B., Mobasher, B., Martin, D. C., et al. 2017, ApJ, 837, 16

Darvish, B., Mobasher, B., Sobral, D., Scoville, N., & Aragon-Calvo, M. 2015, ApJ, 805, 121

Davé, R., Anglés-Alcázar, D., Narayanan, D., et al. 2019, MNRAS, 486, 2827

de Graaff, A., Bezanson, R., Franx, M., et al. 2020, ApJ, 903, L30

de Graaff, A., Bezanson, R., Franx, M., et al. 2021, ApJ, 913, 103

de Graaff, A., Trayford, J., Franx, M., et al. 2022, MNRAS, 511, 2544

DeFelippis, D., Genel, S., Bryan, G. L., & Fall, S. M. 2017, ApJ, 841, 16

Dekel, A. & Burkert, A. 2014, MNRAS, 438, 1870

Desmond, H. & Wechsler, R. H. 2017, MNRAS, 465, 820

D'Eugenio, F., Colless, M., Scott, N., et al. 2021, MNRAS, 504, 5098

di Serego Alighieri, S., Vernet, J., Cimatti, A., et al. 2005, A&A, 442, 125

Dicke, R. H., Peebles, P. J. E., Roll, P. G., & Wilkinson, D. T. 1965, ApJ, 142, 414

Djorgovski, S. & Davis, M. 1987, ApJ, 313, 59

Doi, M., Tanaka, M., Fukugita, M., et al. 2010, AJ, 139, 1628

Dolag, K., Borgani, S., Murante, G., & Springel, V. 2009, MNRAS, 399, 497

D'Onofrio, M., Chiosi, C., Sciarratta, M., & Marziani, P. 2020, A&A, 641, A94

Dressler, A. 1980, ApJ, 236, 351

Dressler, A., Lynden-Bell, D., Burstein, D., et al. 1987, ApJ, 313, 42

Driver, S. P., Allen, P. D., Graham, A. W., et al. 2006, MNRAS, 368, 414

Driver, S. P., Hill, D. T., Kelvin, L. S., et al. 2011, MNRAS, 413, 971

Driver, S. P., Wright, A. H., Andrews, S. K., et al. 2016, MNRAS, 455, 3911

Faber, S. M., Dressler, A., Davies, R. L., et al. 1987, in Nearly Normal Galaxies. From the Planck Time to the Present, ed. S. M. Faber, 175

Faber, S. M. & Jackson, R. E. 1976, ApJ, 204, 668

Ferrero, I., Navarro, J. F., Abadi, M. G., Benavides, J. A., & Mast, D. 2021, A&A, 648, A124

Ferrero, I., Navarro, J. F., Abadi, M. G., et al. 2017, MNRAS, 464, 4736

Fischer, J. L., Bernardi, M., & Meert, A. 2017, MNRAS, 467, 490

Fischer, J. L., Domínguez Sánchez, H., & Bernardi, M. 2019, MNRAS, 483, 2057

Fixsen, D. J., Cheng, E. S., Gales, J. M., et al. 1996, ApJ, 473, 576

Forbes, D. A., Ponman, T. J., & Brown, R. J. N. 1998, ApJ, 508, L43

Franx, M. 1993, PASP, 105, 1058

Franx, M., Illingworth, G., & Heckman, T. 1989, AJ, 98, 538

Franx, M., van Dokkum, P. G., Förster Schreiber, N. M., et al. 2008, ApJ, 688, 770

Furlong, M., Bower, R. G., Crain, R. A., et al. 2017, MNRAS, 465, 722

Gadotti, D. A., Baes, M., & Falony, S. 2010, MNRAS, 403, 2053

Gallazzi, A., Charlot, S., Brinchmann, J., White, S. D. M., & Tremonti, C. A. 2005, MNRAS, 362, 41

Gargiulo, A., Haines, C. P., Merluzzi, P., et al. 2009, MNRAS, 397, 75

Genel, S., Nelson, D., Pillepich, A., et al. 2018, MNRAS, 474, 3976

Genzel, R., Schreiber, N. M. F., Übler, H., et al. 2017, Nature, 543, 397

Gómez, P. L., Nichol, R. C., Miller, C. J., et al. 2003, ApJ, 584, 210

Graham, A. & Colless, M. 1997, MNRAS, 287, 221

Graves, G. J. & Faber, S. M. 2010, ApJ, 717, 803

Graves, G. J., Faber, S. M., & Schiavon, R. P. 2009, ApJ, 698, 1590

Graves, G. J., Faber, S. M., Schiavon, R. P., & Yan, R. 2007, ApJ, 671, 243

Groves, B., Dopita, M. A., Sutherland, R. S., et al. 2008, ApJS, 176, 438

Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357

Häussler, B., McIntosh, D. H., Barden, M., et al. 2007, ApJS, 172, 615

Hernquist, L. 1989, Nature, 340, 687

Hill, A. R., van der Wel, A., Franx, M., et al. 2019, ApJ, 871, 76

Hilz, M., Naab, T., & Ostriker, J. P. 2013, MNRAS, 429, 2924

Hogg, D. W., Bovy, J., & Lang, D. 2010, arXiv e-prints, arXiv:1008.4686

Holden, B. P., van der Wel, A., Kelson, D. D., Franx, M., & Illingworth, G. D. 2010, ApJ, 724, 714

Holden, B. P., van der Wel, A., Rix, H.-W., & Franx, M. 2012, ApJ, 749, 96

Hopkins, P. F., Bundy, K., Murray, N., et al. 2009, MNRAS, 398, 898

Hopkins, P. F., Cox, T. J., & Hernquist, L. 2008, ApJ, 689, 17

Hubble, E. P. 1926, ApJ, 64, 321

Hubble, E. P. 1936, Realm of the Nebulae

Hunter, J. D. 2007, Computing in Science Engineering, 9, 90

Hyde, J. B. & Bernardi, M. 2009, MNRAS, 396, 1171

Ilbert, O., McCracken, H. J., Le Fèvre, O., et al. 2013, A&A, 556, A55

Joachimi, B., Singh, S., & Mandelbaum, R. 2015, MNRAS, 454, 478

Jonsson, P. 2006, MNRAS, 372, 2

Jørgensen, I. & Chiboucas, K. 2013, AJ, 145, 77

Jorgensen, I., Franx, M., & Kjaergaard, P. 1996, MNRAS, 280, 167

Jørgensen, I., Hunter, L. C., O'Neill, C. R., et al. 2019, ApJ, 881, 42

Katz, N. & Gunn, J. E. 1991, ApJ, 377, 365

Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003, MNRAS, 341, 54

Kelvin, L. S., Driver, S. P., Robotham, A. S. G., et al. 2012, MNRAS, 421, 1007

Koopmans, L. V. E., Bolton, A., Treu, T., et al. 2009, ApJ, 703, L51

Krawczyk, C. & Peters, T. 2014, densityplot: v1.0.0

Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2009, ApJ, 700, 221

Kroupa, P. 2001, MNRAS, 322, 231

La Barbera, F., Busarello, G., Merluzzi, P., et al. 2008, ApJ, 689, 913

La Barbera, F., de Carvalho, R. R., de La Rosa, I. G., & Lopes, P. A. A. 2010a,
    MNRAS, 408, 1335

La Barbera, F., Lopes, P. A. A., de Carvalho, R. R., de La Rosa, I. G., & Berlind,
    A. A. 2010b, MNRAS, 408, 1361

Laigle, C., McCracken, H. J., Ilbert, O., et al. 2016, ApJS, 224, 24

Lange, R., Driver, S. P., Robotham, A. S. G., et al. 2015, MNRAS, 447, 2603

Leja, J., Johnson, B. D., Conroy, C., et al. 2019a, ApJ, 877, 140

Leja, J., Tacchella, S., & Conroy, C. 2019b, ApJ, 880, L9

Lelli, F., McGaugh, S. S., Schombert, J. M., Desmond, H., & Katz, H. 2019, MNRAS, 484, 3267

Li, H., Mao, S., Cappellari, M., et al. 2018, MNRAS, 476, 1765

Lin, T., Weipeng, L., Yang, W., & Napolitano, N. R. 2021, arXiv e-prints, arXiv:2109.08838

Liske, J., Baldry, I. K., Driver, S. P., et al. 2015, MNRAS, 452, 2087

Lotz, J. M., Primack, J., & Madau, P. 2004, AJ, 128, 163

Lu, S., Xu, D., Wang, Y., et al. 2020, MNRAS, 492, 5930

Ludlow, A. D., Fall, S. M., Schaye, J., & Obreschkow, D. 2021, arXiv e-prints, arXiv:2105.03561

Ludlow, A. D., Schaye, J., Schaller, M., & Richings, J. 2019, MNRAS, 488, L123

Madau, P. & Dickinson, M. 2014, ARA&A, 52, 415

Magoulas, C., Springob, C. M., Colless, M., et al. 2012, MNRAS, 427, 245

Maraston, C. 2005, MNRAS, 362, 799

Marchesini, D., van Dokkum, P. G., Förster Schreiber, N. M., et al. 2009, ApJ, 701, 1765

Marsden, C., Shankar, F., Bernardi, M., et al. 2022, MNRAS, 510, 5639

Matthee, J., Schaye, J., Crain, R. A., et al. 2017, MNRAS, 465, 2381

McAlpine, S., Helly, J. C., Schaller, M., et al. 2016, Astronomy and Computing, 15, 72

McCracken, H. J., Milvang-Jensen, B., Dunlop, J., et al. 2012, A&A, 544, A156

Meert, A., Vikram, V., & Bernardi, M. 2015, MNRAS, 446, 3943

Meyer, M. J., Zwaan, M. A., Webster, R. L., Schneider, S., & Staveley-Smith, L. 2008, MNRAS, 391, 1712

Miller, T. B., van Dokkum, P., Mowla, L., & van der Wel, A. 2019, ApJ, 872, L14

Mo, H. J., Mao, S., & White, S. D. M. 1998, MNRAS, 295, 319

Mosleh, M., Tacchella, S., Renzini, A., et al. 2017, ApJ, 837, 2

Mowla, L. A., van Dokkum, P., Brammer, G. B., et al. 2019, ApJ, 880, 57

Mukherjee, S., Koopmans, L. V. E., Tortora, C., et al. 2022, MNRAS, 509, 1245

Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013a, ApJ, 777, 18

Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013b, ApJS, 206, 8

Naab, T., Johansson, P. H., & Ostriker, J. P. 2009, ApJ, 699, L178

Narayanan, D., Turk, M. J., Robitaille, T., et al. 2021, ApJS, 252, 12

Neistein, E., van den Bosch, F. C., & Dekel, A. 2006, MNRAS, 372, 933

Newman, A. B., Belli, S., Ellis, R. S., & Patel, S. G. 2018, ApJ, 862, 126

Padilla, N. D. & Strauss, M. A. 2008, MNRAS, 388, 1321

Peacock, J. A. 1999, Cosmological Physics

Peebles, P. J. E. 1968, ApJ, 153, 1

Peebles, P. J. E. & Yu, J. T. 1970, ApJ, 162, 815

Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, AJ, 124, 266

Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2010, AJ, 139, 2097

Penzias, A. A. & Wilson, R. W. 1965, ApJ, 142, 419

Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, ApJ, 517, 565

Pforr, J., Maraston, C., & Tonini, C. 2012, MNRAS, 422, 3285

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014, A&A, 571, A16

Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A6

Price, S. H., Kriek, M., Feldmann, R., et al. 2017, ApJ, 844, L6

Prichard, L. J., Davies, R. L., Beifiori, A., et al. 2017, ApJ, 850, 203

Prugniel, P. & Simien, F. 1997, A&A, 321, 111

Renzini, A. & Ciotti, L. 1993, ApJ, 416, L49

Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, AJ, 116, 1009

Rix, H.-W., Barden, M., Beckwith, S. V. W., et al. 2004, ApJS, 152, 163

Roberts, M. S. & Haynes, M. P. 1994, ARA&A, 32, 115

Robertson, B., Cox, T. J., Hernquist, L., et al. 2006, ApJ, 641, 21

Rodriguez-Gomez, V., Snyder, G. F., Lotz, J. M., et al. 2019, MNRAS, 483, 4140

Romanowsky, A. J. & Fall, S. M. 2012, ApJS, 203, 17

Rosito, M. S., Pedrosa, S. E., Tissera, P. B., et al. 2021, A&A, 652, A44

Rosito, M. S., Tissera, P. B., Pedrosa, S. E., & Lagos, C. D. P. 2019a, A&A, 629, L3

Rosito, M. S., Tissera, P. B., Pedrosa, S. E., & Rosas-Guevara, Y. 2019b, A&A, 629, A37

Saglia, R. P., Sánchez-Blázquez, P., Bender, R., et al. 2010, A&A, 524, A6

Saglia, R. P., Sánchez-Blázquez, P., Bender, R., et al. 2016, A&A, 596, C1

Sales, L. V., Navarro, J. F., Schaye, J., et al. 2010, MNRAS, 409, 1541

Saracco, P., Gargiulo, A., La Barbera, F., Annunziatella, M., & Marchesini, D. 2020, MNRAS, 491, 1777

Schaller, M., Frenk, C. S., Bower, R. G., et al. 2015, MNRAS, 451, 1247

Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, MNRAS, 446, 521

Schechter, P. L., Pooley, D., Blackburne, J. A., & Wambsganss, J. 2014, ApJ, 793, 96

Scott, D., Silk, J., & White, M. 1995, Science, 268, 829

Scott, N., Brough, S., Croom, S. M., et al. 2017, MNRAS, 472, 2833

Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1

Sersic, J. L. 1968, Atlas de Galaxias Australes (Observatorio Astronomico, Universidad Nacional de Cordoba)

Shapiro, P. R., Iliev, I. T., Martel, H., Ahn, K., & Alvarez, M. A. 2004, arXiv e-prints, astro

Shen, S., Mo, H. J., White, S. D. M., et al. 2003, MNRAS, 343, 978

Simard, L., Mendel, J. T., Patton, D. R., Ellison, S. L., & McConnachie, A. W. 2011, ApJS, 196, 11

Simard, L., Willmer, C. N. A., Vogt, N. P., et al. 2002, ApJS, 142, 1

Snyder, G. F., Torrey, P., Lotz, J. M., et al. 2015, MNRAS, 454, 1886

Somerville, R. S., Barden, M., Rix, H.-W., et al. 2008, ApJ, 672, 776

Springel, V., White, S. D. M., Tormen, G., & Kauffmann, G. 2001, MNRAS, 328, 726

Stoughton, C., Lupton, R. H., Bernardi, M., et al. 2002, AJ, 123, 485

Straatman, C. M. S., van der Wel, A., Bezanson, R., et al. 2018, ApJS, 239, 27

Suess, K. A., Kriek, M., Price, S. H., & Barro, G. 2019, ApJ, 877, 103

Szomoru, D., Franx, M., van Dokkum, P. G., et al. 2013, ApJ, 763, 73

Taylor, E. N., Franx, M., Brinchmann, J., van der Wel, A., & van Dokkum, P. G. 2010, ApJ, 722, 1

Taylor, E. N., Hopkins, A. M., Baldry, I. K., et al. 2015, MNRAS, 446, 2144

Thob, A. C. R., Crain, R. A., McCarthy, I. G., et al. 2019, MNRAS, 485, 972

Thomas, D., Maraston, C., Bender, R., & Mendes de Oliveira, C. 2005, ApJ, 621, 673

Toft, S., Gallazzi, A., Zirm, A., et al. 2012, ApJ, 754, 3

Toft, S., Zabl, J., Richard, J., et al. 2017, Nature, 546, 510

Trayford, J. W., Camps, P., Theuns, T., et al. 2017, MNRAS, 470, 771

Trayford, J. W., Frenk, C. S., Theuns, T., Schaye, J., & Correa, C. 2019, MNRAS, 483, 744

Trayford, J. W., Theuns, T., Bower, R. G., et al. 2015, MNRAS, 452, 2879

Treu, T., Ellis, R. S., Liao, T. X., et al. 2005, ApJ, 633, 174

Trujillo, I., Burkert, A., & Bell, E. F. 2004, ApJ, 600, L39

Trujillo, I., Conselice, C. J., Bundy, K., et al. 2007, MNRAS, 382, 109

Trujillo, I., Förster Schreiber, N. M., Rudnick, G., et al. 2006, ApJ, 650, 18

Tully, R. B. & Fisher, J. R. 1977, A&A, 500, 105

van de Sande, J., Kriek, M., Franx, M., Bezanson, R., & van Dokkum, P. G. 2014, ApJ, 793, L31

van de Sande, J., Kriek, M., Franx, M., Bezanson, R., & van Dokkum, P. G. 2015, ApJ, 799, 125

van de Sande, J., Kriek, M., Franx, M., et al. 2013, ApJ, 771, 85

van de Sande, J., Lagos, C. D. P., Welker, C., et al. 2019, MNRAS, 484, 869

van de Sande, J., Scott, N., Bland-Hawthorn, J., et al. 2018, Nature Astronomy, 2, 483

van de Ven, G. & van der Wel, A. 2021, ApJ, 914, 45

van der Wel, A., Bell, E. F., Häussler, B., et al. 2012, ApJS, 203, 24

van der Wel, A., Chang, Y.-Y., Bell, E. F., et al. 2014b, ApJ, 792, L6

van der Wel, A., Franx, M., van Dokkum, P. G., & Rix, H. W. 2004, ApJ, 601, L5

van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2005, ApJ, 631, 145

van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014a, ApJ, 788, 28

van der Wel, A., Noeske, K., Bezanson, R., et al. 2016, ApJS, 223, 29

van der Wel, A., Rix, H.-W., Holden, B. P., Bell, E. F., & Robaina, A. R. 2009, ApJ, 706, L120

van Dokkum, P. G. & Conroy, C. 2010, Nature, 468, 940

van Dokkum, P. G. & Conroy, C. 2012, ApJ, 760, 70

van Dokkum, P. G. & Franx, M. 1996, MNRAS, 281, 985

van Dokkum, P. G. & Franx, M. 2001, ApJ, 553, 90

van Dokkum, P. G., Franx, M., Kelson, D. D., & Illingworth, G. D. 2001, ApJ, 553, L39

van Dokkum, P. G. & van der Marel, R. P. 2007, ApJ, 655, 30

van Houdt, J., van der Wel, A., Bezanson, R., et al. 2021, ApJ, 923, 11

Vincent, R. A. & Ryden, B. S. 2005, ApJ, 623, 137

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261

Vogelsberger, M., Genel, S., Springel, V., et al. 2014, MNRAS, 444, 1518

Westfall, K. B., Cappellari, M., Bershady, M. A., et al. 2019, AJ, 158, 231

Whitaker, K. E., Labbé, I., van Dokkum, P. G., et al. 2011, ApJ, 735, 86

White, M., Scott, D., & Silk, J. 1994, ARA&A, 32, 319

Wu, P.-F., van der Wel, A., Gallazzi, A., et al. 2018, ApJ, 855, 85

Wuyts, S., Förster Schreiber, N. M., Wisnioski, E., et al. 2016, ApJ, 831, 149

Wuyts, S., van Dokkum, P. G., Kelson, D. D., Franx, M., & Illingworth, G. D. 2004, ApJ, 605, 677

Wyder, T. K., Martin, D. C., Schiminovich, D., et al. 2007, ApJS, 173, 293

York, D. G., Adelman, J., Anderson, John E., J., et al. 2000, AJ, 120, 1579

Zahid, H. J., Damjanov, I., Geller, M. J., Hwang, H. S., & Fabricant, D. G. 2016, ApJ, 821, 101

Zaritsky, D., Gonzalez, A. H., & Zabludoff, A. I. 2006, ApJ, 638, 725

Zaritsky, D., Zabludoff, A. I., & Gonzalez, A. H. 2008, ApJ, 682, 68

Zolotov, A., Dekel, A., Mandelker, N., et al. 2015, MNRAS, 450, 2327

Zubko, V., Dwek, E., & Arendt, R. G. 2004, ApJS, 152, 211

Zwaan, M. A., van der Hulst, J. M., de Blok, W. J. G., & McGaugh, S. S. 1995, MNRAS, 273, L35

Nederlandse Samenvatting

De Melkweg

Onze Melkweg is op een donkere, heldere nacht te zien als een lichte band van sterren aan de hemel. Achter dit zwakke licht schuilt een ingewikkelde structuur. Inmiddels weten we dat ons sterrenstelsel bestaat uit ruim 100 miljard sterren en grote hoeveelheden gas en stof, die samen een platte schijf vormen met daarbinnen grote spiraalarmen.

Echter, de Melkweg is, in al haar complexiteit, slechts een van de vele sterrenstelsels in het heelal. Deze ontdekking is relatief nieuw: ongeveer honderd jaar geleden werd voor het eerst aangetoond dat enkele waargenomen 'gasnevels' zó ver weg staan (zoals Andromeda), dat het op zichzelf staande sterrenstelsels moeten zijn. Sindsdien is het vakgebied van de extragalactische sterrenkunde exponentieel gegroeid. Dankzij moderne telescopen, die zeer efficiënt sterrenstelsels over de gehele hemel en tot op grote afstanden in kaart hebben gebracht, weten we nu van het bestaan en de eigenschappen van miljoenen sterrenstelsels.

De grote vraag die rijst: hoe is de Melkweg ontstaan, en hoe verhoudt deze geschiedenis zich tot de rest van het universum? Hoewel er veel vooruitgang is geboekt in de afgelopen decennia, is de zoektocht naar het antwoord hierop nog in volle gang.

Een extragalactische tweedeling

Onder de miljoenen sterrenstelsels die zijn ontdekt heerst grote diversiteit. Desalniettemin kan de verscheidenheid aan sterrenstelsels worden samengevat in een beperkt aantal soorten. Deze classificatie, ontworpen door Edwin Hubble, bevat twee belangrijke types: de spiraalvormige stelsels, en de elliptische stelsels (zie Figuur 1). Een derde categorie is die van de onregelmatige stelsels, dat wil zeggen, zonder kenmerkende structuur (ook wel 'trainwrecks' genoemd).

Het meest voorkomend in aantal zijn de spiraalvormige, stervormende stelsels, zo ook onze Melkweg. Deze stelsels hebben in gemeen dat ze ongeveer dezelfde structuur hebben, bestaande uit een platte schijf met spiraalarmen. Ze hebben veelal ook dezelfde blauwe kleur, veroorzaakt door de aanwezigheid van veel jonge sterren. Daar tegenover staan de elliptische stelsels, ook wel passieve stelsels genoemd: deze sterrenstelsels hebben een rondere vorm en een egale structuur. Daarnaast bevatten deze stelsels voornamelijk oude sterren en vrijwel geen gas waarvan nieuwe sterren kunnen worden gevormd, en zijn zij daarom roder van kleur ('rood en dood').
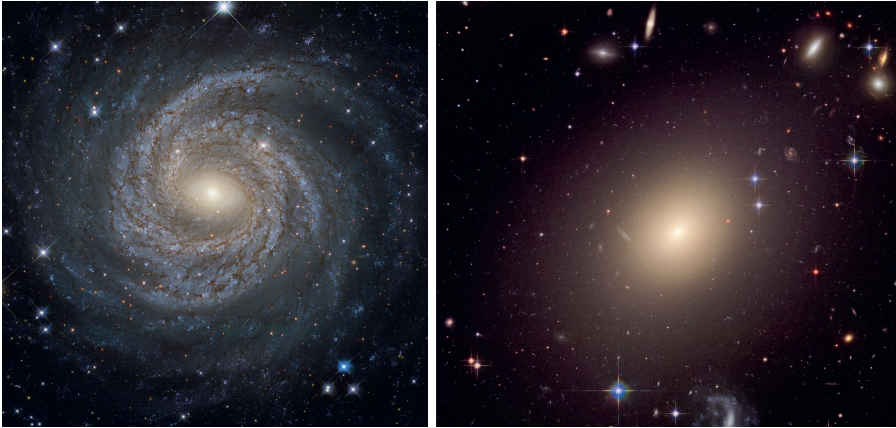
Figuur 1: Voorbeeld van een spiraalvormig sterrenstelsel (NGC 6814; links) en een elliptisch sterrenstelsel (Abell S0740; rechts) in het nabije heelal. *Bron: NASA, ESA, en The Hubble Heritage Team*

Het feit dat sommige sterrenstelsels spiraalvormig zijn en andere elliptisch, suggereert dat er grote verschillen zijn in de ontstaansgeschiedenis van deze stelsels. Er wordt daarom actief onderzoek gedaan om te begrijpen wat de verschillen, en overeenkomsten, zijn tussen sterrenstelsels.

Hier wordt een combinatie van waarnemingen voor gebruikt: gezien de evolutie van sterrenstelsels traag verloopt (over miljarden jaren), kunnen we niet de ontwikkeling van individuele stelsels volgen. Wel kunnen we gebruik maken van het feit dat de snelheid van het licht constant is, en we daardoor van verre sterrenstelsels het licht zien dat miljarden jaren geleden was uitgezonden. Door waarnemingen van nabije en verre sterrenstelsels te combineren, kunnen we een reconstructie maken van de tijdlijn van de geschiedenis van het heelal en de sterrenstelsels daarbinnen.

Ontstaan en evolutie van sterrenstelsels

De oorsprong van sterrenstelsels ligt in de allervroegste periode van het heelal. Vlak na de oerknal (13,8 miljard jaar geleden) was alle materie extreem heet en nagenoeg gelijk verdeeld: het heelal zag er in alle richtingen hetzelfde uit (isotroop) en met een gelijke samenstelling (homogeen). De nadruk hier ligt op het woord *nagenoeg*: piepkleine fluctuaties in de dichtheid van de materie werden met de tijd groter onder de invloed van de zwaartekracht, terwijl de materie langzaamaan afkoelde. Tegenwoordig zijn deze kiemen uitgegroeid tot gigantische structuren van enkele miljarden lichtjaren groot (zo'n 10^{20} meter).

Een belangrijke ontdekking is dat deze structuren bestaan uit meer dan alleen de materie zoals wij deze op Aarde kennen. Volgens huidige theorieën, die een zeer goede beschrijving van onze waarnemingen vormen, bestaat er vijf keer meer materie dan wat wij kunnen zien. Dit mysterieuze component wordt ook wel donkere materie genoemd, en heeft speciale eigenschappen. Hoewel gewone materie

allerlei natuurkundige en scheikundige processen kan doorlopen, waarbij bijvoorbeeld zware atomen en complexe moleculen worden gevormd, hebben de deeltjes van donkere materie nauwelijks interactie met elkaar dan wel met gewone materie.

Toch speelt donkere materie een belangrijke rol in de geschiedenis van het heelal en de sterrenstelsels. In het vroege heelal klontert de donkere materie namelijk sneller samen, en vormt zo de basis van de structuren op grote schaal. Deze structuren hebben een vorm vergelijkbaar met een spinnenweb ('kosmisch web'): lange draden (filamenten) van donkere materie verbinden de knopen van het web. Deze knopen van donkere materie worden ook wel halo's genoemd, en zijn de geboorteplaats van sterrenstelsels.

De gewone materie, in de vorm van heet gas, klontert verder samen binnenin de halo's, en koelt af via straling. Wanneer de dichtheid van het gas hoog genoeg is en de temperatuur laag genoeg, vormen zich de eerste sterren. Met de tijd wordt steeds meer gas aangevoerd vanuit het kosmische web en neemt het aantal sterren toe, wat samen een jong sterrenstelsel vormt.

Hoewel de details van deze processen onduidelijk zijn, gezien de eerste generatie sterren nog nooit is waargenomen, weten we wel met zekerheid dat er 500 miljoen jaar na de oerknal al sterrenstelsels bestonden. Wat later, ongeveer 2-3 miljard jaar na de oerknal, begint de tweedeling tussen sterrenstelsels al langzaam zichtbaar te worden. Echter, zijn er nog grote verschillen tussen deze jonge sterrenstelsels en de spiraalvormige en elliptische stelsels in het huidige heelal.

Ten eerste, zoals redelijkerwijs te verwachten, zijn de jonge stelsels kleiner en minder massief. De verschillen tussen de verre sterrenstelsels en die van vandaag de dag gaan echter verder dan alleen hun formaat. Ten tweede zijn de sterren in de vroege stelsels gemiddeld een stuk jonger, en bevatten de sterren en het gas een andere chemische samenstelling, met minder 'metalen' (een verwarrende term onder sterrenkundigen voor elementen met hogere massa dan waterstof). Ten derde is er een groot verschil in de structuur van de sterrenstelsels, gezien waarnemingen van verre sterrenstelsels aantonen dat de sterren en het gas binnen deze stelsels dikke schijven vormen. Bovendien, in tegenstelling tot de sterren in een spiraalstelsel die in een vrijwel circulaire baan rond het centrum van het stelsel draaien, zijn de bewegingen van de sterren en het gas binnen deze jonge stelsels een stuk turbulenter.

Onder de structuur valt ook de rol van de donkere materie op de galactische schaal. De halo is grofweg een factor 10-100 maal groter dan het sterrenstelsel zelf, en 100-1000 keer zwaarder dan de totale massa van alle sterren binnen het stelsel. Toch domineert de gewone materie op kleine schaal, binnenin het sterrenstelsel, gezien binnen een straal van een paar lichtjaren slechts een klein deel van de massa "donker" is. Hier schijnt ook evolutie in te bestaan, omdat voor jonge sterrenstelsels de fractie aan donkere materie op kleine schaal lager lijkt te zijn dan voor oudere stelsels, hoewel dit berust op slechts een klein aantal, lastig te verkrijgen metingen.

Het waarnemen van sterrenstelsels

Uniek aan de wetenschap van de sterrenkunde is dat het heelal ons laboratorium vormt, en het ontwerpen van 'experimenten' daardoor een ander karakter aan-

neemt. Het bepalen van de meest basale eigenschappen, zoals een gewicht of een afstand, is niet triviaal, niet in de laatste plaats omdat onze metingen puur berusten op het licht dat wordt uitgezonden door de verre sterrenstelsels. Telescopen verrichten metingen van dit licht, en door middel van de juiste waarneemstrategie kan men hieruit de eigenschappen en ontwikkeling van sterrenstelsels afleiden.

Op basis van een enkel beeld ('image') krijgen we inzicht in de helderheid en de distributie van het licht binnen een sterrenstelsel, en daarmee de geprojecteerde (2-dimensionale) vorm en de relatieve grootte (de hoekdiameter). Om de absolute grootte te bepalen is meer informatie nodig, namelijk de afstand tot het stelsel. Dit is mogelijk door het meten van kleuren: de verhouding tussen de helderheid van het licht op verschillende golflengtes.

Het meten van verschillende kleuren is in meerdere opzichten cruciaal. Zoals eerder benoemd, correleert de kleur met het type sterrenstelsel (de Melkweg is bijvoorbeeld relatief blauw), en biedt daardoor inzicht in de eigenschappen van de stellaire populaties. Een belangrijke observatie is echter dat verre sterrenstelsels allemaal rood blijken te zijn, wat tegen onze verwachting in lijkt te gaan. Dit wordt veroorzaakt door de uitdijing van het heelal: jonge sterrenstelsels, die ver van ons staan, zijn intrinsiek blauw, maar hun licht wordt tijdens de lange reis door het heelal langzaam uitgerekt tot langere (en dus rodere) golflengtes. Dit effect heet roodverschuiving, en de mate van roodverschuiving is direct afhankelijk van de afstand tussen de Aarde en het waargenomen object. Door meerdere kleuren te meten kan de roodverschuiving van een sterrenstelsel worden bepaald, en daarmee de afstand tot en grootte van het stelsel.

Wanneer de afstand eenmaal bekend is, biedt de combinatie van de totale helderheid en de verschillende kleuren ook inzicht in het gewicht en andere eigenschappen van de sterpopulatie van een sterrenstelsel. Door deze metingen te vergelijken met modellen van sterpopulaties, krijgen we een schatting van de gemiddelde leeftijd en totale massa van de sterren, evenals de mate van stervorming.

Hoewel deze massaschattingen relatief eenvoudig te verkrijgen zijn, berusten ze op meerdere onzekere aannames, wat leidt tot een onzekerheid van een factor 2-3 in de massa. Bovendien betreft deze schatting alleen de *stellaire* massa, en gaat dus voorbij aan de eventuele extra massa van het gas en de donkere materie binnen het sterrenstelsel.

Een methode om de *totale* massa van een sterrenstelsel te meten is gebaseerd op de bewegingen van sterren binnen het stelsel. Zo is de snelheid van een ster draaiend in een circulaire baan rond een sterrenstelsel enkel afhankelijk van de straal van de baan en de totale massa binnen deze straal. Het meten van individuele sterren is alleen mogelijk voor zeer nabije sterrenstelsels, maar het meten van een grotere populatie van sterren biedt ook inzicht in de typische snelheid van de sterren, en daarmee een schatting van de totale massa. Voor deze bepaling is wel een spectrum nodig: voor sterren die zich van ons af bewegen, zullen de spectraallijnen een grotere roodverschuiving hebben dan voor sterren die zich iets naar ons toe bewegen. De spectraallijnen in het spectrum van een sterrenstelsel zijn daardoor verbreed, en deze verbreding kan eenvoudig worden omgerekend naar een snelheid. Dit wordt ook wel de snelheidsdispersie genoemd, en beschrijft de snelheid waarmee sterren ten opzichte van elkaar bewegen.

In tegenstelling tot de waarnemingen in de vorm van beelden zijn spectra erg lastig te verkrijgen, gezien al gauw een hele nacht nodig is om genoeg data te verzamelen voor een object met lage helderheid. Hoewel er ongeveer een miljoen spectra zijn van relatief nabije, heldere sterrenstelsels, zijn er tot nog toe slechts een paar honderd spectra genomen van jongere sterrenstelsels met een hogere rood-verschuiving, die minder helder zijn. Een groot deel van de conclusies over verre sterrenstelsels is dus gebaseerd op een kleine hoeveelheid data. Bovendien, gezien de moeilijkheid van het verkrijgen van de spectra, zijn deze vooral genomen voor de helderste verre sterrenstelsels, die zeer waarschijnlijk niet representatief zijn voor de algehele populatie van sterrenstelsels. Om de evolutie van sterrenstelsels beter te begrijpen en bestaande theorieën te testen, is het daarom cruciaal om meer spectra te verzamelen en analyseren van verre sterrenstelsels.

Dit Proefschrift

Centraal in dit proefschrift staat de vraag wat de eigenschappen van sterrenstelsels op hogere roodverschuiving zijn, en hoe deze zich verhouden tot de sterrenstelsels van nu. In het bijzonder ligt de focus op de structuur van sterrenstelsels: de massa en de verdeling daarvan binnen het stelsel.

Voor de beantwoording maken we gebruik van een groot aantal waarnemingen met de Very Large Telescope in Chili. Deze waarnemingen maken deel uit van de LEGA-C Survey, en bestaan uit diepe spectra van ongeveer 3000 massieve sterrenstelsels met een roodverschuiving van 0.6 tot 1.0, toen het heelal ongeveer de helft van de huidige leeftijd was (6-8 miljard jaar oud). Van deze spectra meten we de snelheidsdispersies en andere eigenschappen van de sterrenstelsels, zoals een schatting van de leeftijd. Verder gebruiken we bestaande waarnemingen van de ruimtetelescoop *Hubble* om de grootte van en de lichtdistributie binnen de stelsels te meten. Ook meten we, met behulp van een verzameling aan data van verschillende telescopen uit de literatuur, de kleuren van dezelfde stelsels en maken zo een schatting van de stellaire massa's.

Uniek aan deze uitgebreide dataset is niet alleen de omvang, maar ook de selectie: in tegenstelling tot eerdere studies zijn de sterrenstelsels in LEGA-C representatief voor de populatie van massieve sterrenstelsels op de hogere rood-verschuiving van ongeveer $z \approx 0.8$. Daarom kunnen we, door vergelijking met data op lage roodverschuiving in het nabije heelal, de evolutie in de eigenschappen van sterrenstelsels onderzoeken.

Daarnaast maken we gebruik van grote computersimulaties, die het ontstaan en de evolutie van sterrenstelsels simuleren op basis van de meest recente theo-rieën. Aan de ene kant bieden deze simulaties een handvat voor de interpretatie van de waargenomen data, en aan de andere kant kunnen we op deze manier onze huidige theorieën over de ontwikkeling van sterrenstelsels testen.

In **Hoofdstuk 2** presenteren we ten eerste de verschillende waarnemingen en ma-ken we een zorgvuldige selectie van sterrenstelsels uit de LEGA-C Survey, op basis van strikte kwaliteitseisen. Vervolgens tonen we aan dat er voor deze sterrenstelsels een sterke correlatie is tussen de grootte, de snelheidsdispersie, en de helderheid.

Deze relatie staat beter bekend als de 'fundamental plane': sterrenstelsels liggen op een vlak binnen deze 3-dimensionale ruimte. Dit is eerder ook aangetoond op zowel lagere als hogere roodverschuiving voor elliptische en passieve sterrenstelsels.

Een nieuwe bevinding is dat de spiraalvormige en stervormende sterrenstelsels bijna dezelfde relatie lijken te volgen, maar met een grotere spreiding. We laten zien dat deze spreiding enkel wordt veroorzaakt door grote verschillen in de eigenschappen van de sterpopulaties van sterrenstelsels, zoals de leeftijd en mate van stervorming. Wanneer we in plaats van de helderheid kijken naar de relatie tussen de grootte, snelheidsdispersie en de stellaire massa, vinden we één 'stellar mass fundamental plane': zowel passieve als stervormende sterrenstelsels volgen dezelfde 2-dimensionale relatie met een erg lage spreiding, ondanks de enorme aanwezige verschillen in de eigenschappen van de sterrenstelsels.

We vergelijken deze resultaten met waarnemingen van het nabije heelal in **Hoofdstuk 3**. Er is sterke evolutie in de fundamental plane, wat aangeeft dat er over de periode van 8 miljard jaar significante veranderingen hebben plaatsgevonden in de stellaire populaties of in de structuren van sterrenstelsels. Verrassend genoeg vinden we echter geen evolutie in de stellar mass fundamental plane. Dit suggereert dat, voor een stelsel van een bepaalde grootte en snelheidsdispersie, het verschil tussen een hoge en lage roodverschuiving voornamelijk wordt veroorzaakt door een verandering in de stellaire populatie. Overigens betekent dit niet dat individuele sterrenstelsels niet van structuur veranderen. Integendeel, sterrenstelsels veranderen met de tijd van structuur, maar deze evolutie is sterk gereguleerd, in de zin dat bijvoorbeeld een verandering in de grootte van het stelsel direct is gekoppeld aan een verandering in de snelheidsdispersie.

In **Hoofdstuk 4** betrekken we voor het eerst de computersimulaties erbij. We maken gebruik van de EAGLE simulaties en onderzoeken hoe deze simulaties ter ondersteuning van de waarnemingen kunnen dienen. Aangezien de simulaties de 3-dimensionale massadistributies van sterrenstelsels modelleren, wat substantieel verschilt van hoe we sterrenstelsels waarnemen, creëren we eerst 'namaak'-beelden van de ruim 3000 gesimuleerde sterrenstelsels. Deze beelden zijn gemaakt door het licht dat de gesimuleerde sterrenstelsels uitstralen te modelleren en te projecteren, en bevatten, net als echte waarnemingen, ruis en de effecten van atmosferische turbulentie.

Door de gebruikelijke waarneemtechnieken toe te passen voor de analyse van deze beelden, vinden we dat er grote verschillen zijn tussen de structuren gemeten van de beelden en die van de 3-dimensionale massadistributies uit de simulatie. We laten zien dat dit een significant effect heeft op de gemeten relatie tussen de grootte en de stellaire massa van sterrenstelsels, een relatie die vaak als succescriterium wordt gebruikt voor simulaties. De EAGLE simulaties reproduceren de waargenomen relatie zeer goed, inclusief de spreiding binnen de relatie. Echter, vinden we ook enkele belangrijke verschillen: de structuren van de gesimuleerde sterrenstelsels zijn minder compact in vergelijking met de echte waarnemingen. Daarnaast zijn de spiraalvormige stelsels niet plat genoeg, en de elliptische stelsels niet rond genoeg.

We zetten deze analyse voort in **Hoofdstuk 5**, waar we de oorsprong van de fundamental plane relatie doorgronden met behulp van dezelfde simulaties. Eerst

laten we zien dat er voor alle sterrenstelsels een zeer nauwe relatie bestaat tussen de totale massa, de grootte, en de snelheidsdispersie. Deze relatie ligt ook dichtbij de theoretische verwachting, met kleine afwijkingen die kunnen worden verklaard door verschillen in de structuren tussen de gesimuleerde sterrenstelsels.

Vervolgens vervangen we de totale massa door de stellaire massa, en tonen aan dat de gesimuleerde sterrenstelsels ook op één stellar mass fundamental plane liggen, net als de waargenomen sterrenstelsels. In tegenstelling tot de waarnemingen, bieden de simulaties inzicht in de oorsprong van deze relatie. We vinden dat de hoeveelheid donkere materie binnen een sterrenstelsel sterk afhankelijk is van de grootte en de stellaire massa. Deze afhankelijkheid is vervolgens verantwoordelijk voor de precieze verhoudingen in de relatie tussen de stellaire massa, de grootte en de snelheidsdispersie van de stellar mass fundamental plane. Dit verklaart ook waarom we, ondanks de tweedeling binnen de populatie van sterrenstelsels, slechts één relatie vinden: de variatie in de donkere materie massa als functie van de grootte en de stellaire massa is nagenoeg hetzelfde voor zowel passieve als stervormende sterrenstelsels.

Ten slotte gebruiken we de eerder gemaakte beelden van de simulaties, en maken we ook metingen van de snelheidsdispersies van de gesimuleerde stelsels op een manier lijkend op echte waarnemingen. We tonen aan dat waarneemeffecten een significante invloed hebben op de gemeten structuur en fundamental plane van sterrenstelsels, en dat dit met name de spreiding in de relatie beïnvloedt. Door zo een eerlijke vergelijking te maken tussen de simulaties en waarnemingen, vinden we dat de massadistributies van de gesimuleerde sterrenstelsels systematisch verschillen van de waargenomen stelsels. Het is daardoor lastig te zeggen of de interpretatie van de oorsprong van de fundamental plane relatie zoals hierboven beschreven ook direct toepasbaar is op de waargenomen stelsels. Aan de andere kant toont dit werk aan dat de fundamental plane relatie, wat op zichzelf een zeer eenvoudig te meten relatie is, een nieuwe, simpele test biedt om na te gaan hoe goed een computersimulatie het heelal nabootst.

Toekomstige simulaties, waarvan de sterrenstelsels nog dichter bij de werkelijkheid zullen liggen, zullen moeten uitwijzen wat de rol van donkere en gewone materie in de massadistributie op de galactische schaal precies is. Het voordeel van de gesimuleerde sterrenstelsels ten opzichte van de waargenomen stelsels, waarvan we slechts een momentopname krijgen, is dat de ontwikkeling vanaf het ontstaan van de eerste sterren volledig kan worden gevolgd. Er valt dus nog een hoop te leren over de tijdsevolutie van sterrenstelsels op basis van deze simulaties.

Tegelijkertijd zijn nieuwe waarnemingen uiteraard ook van belang, die bijvoorbeeld kunnen helpen bij het verbeteren en aanscherpen van de theorieën aan de hand waarvan de simulaties worden ontwikkeld. Bovendien bieden nieuwe telescopen en instrumenten, zoals de Atacama Large Millimeter/submillimeter Array of de *James Webb Space Telescope*, een zeer hoge resolutie waarmee zelfs de verste sterrenstelsels scherp kunnen worden waargenomen en dus de structuren van sterrenstelsels kunnen worden ontrafeld. In de komende jaren kunnen we verwachten grote vooruitgang te boeken in onze kennis over de ontwikkeling van sterrenstelsels.

Publications

Refereed publications

1. *The Mass Scale of High-Redshift Galaxies: Virial Mass Estimates Calibrated with Stellar Dynamical Models from LEGA-C*
van der Wel, A., van Houdt, J., Bezanson, R., Franx, M., D'Eugenio, F., Straatman, C., Bell, E.F., Muzzin, A., Sobral, D., Maseda, M., **de Graaff, A.**, Holden, B., ApJ in press

2. *The LEGA-C and SAMI Galaxy Surveys: Quiescent Stellar Populations and the Mass-Size Plane across 6 Gyr*
Barone, T., D'Eugenio, F., Scott, N., Colless, M., Vaughan, S.P., van der Wel, A., Fraser-McKelvie, A., **de Graaff, A.**, van de Sande, J., Wu, P-F., Bezanson, R., Brough, S., Bell, E.F., Croom, S.M., Cortese, L., Driver, S., Gallazzi, A.R., Muzzin, A., Sobral, D., Bland-Hawthorn, J., Bryant, J.J., Goodwin, M., Lawrence, J.S., Lorente, N.P.F., Owers, M.S., 2022, MNRAS, 512, 3828

3. *Observed structural parameters of EAGLE galaxies: reconciling the mass-size relation in simulations with local observations*
de Graaff, A., Trayford, J., Franx, M., Schaller, M., Schaye, J., van der Wel, A., 2022, MNRAS, 511, 2544

4. *LEGA-C: Analysis of Dynamical Masses from Ionized Gas and Stellar Kinematics at $z \sim 0.8$*
Straatman, C., van der Wel, A., van Houdt, J., Bezanson, R., Bell, E.F., van Dokkum, P., D'Eugenio, F., Franx, M., Gallazzi, A., **de Graaff, A.**, Maseda, M.V., Meidt, S., Muzzin, A., Sobral, D., Wu, P-F, ApJ, 928, 126

5. *Ubiquitous OII Emission in Quiescent Galaxies at $z \sim 0.85$ from the LEGA-C Survey*
Maseda, M., Van der Wel, A., Franx, M., Bell, E.F., Bezanson, R., Muzzin, A., Sobral, D., D'Eugenio, F., Gallazzi, A., **de Graaff, A.**, Leja, J., Straatman, C., Whitaker, K.E., Williams, C.C., Wu, P-F., 2021, ApJ, 923, 18

6. *Stellar Dynamical Models for 797 $z \sim 0.8$ Galaxies from LEGA-C*
van Houdt, J., van der Wel, A., Bezanson, R., Franx, M., D'Eugenio, F., Barisic, I., Bell, E.F., Gallazzi, A., **de Graaff, A.**, Maseda, M.V., Pacifici,

C., van de Sande, J., Sobral, D., Straatman, C., Wu, P-F., 2021, ApJ, 923, 11

7. *The Large Early Galaxy Astrophysics Census (LEGA-C) Data Release 3: 3000 High-Quality Spectra of Ks-selected galaxies at z > 0.6*
Van der Wel, A., Bezanson, R., D'Eugenio, F., Straatman, C., Franx, M., van Houdt, J., Maseda, M.V., Gallazzi, A., Wu, P-F., Pacifici, C., Barisic, I., Brammer, G.B., Munoz-Mateos, J.C., Vervalcke, S., Zibetti, S., Sobral, D., **de Graaff, A.**, Calhau, J., Kaushal, Y., Muzzin, A., Bell, E.F., van Dokkum, P.G., 2021, ApJS, 256, 44

8. *The Fundamental Plane in the LEGA-C Survey: Unraveling the M/L Ratio Variations of Massive Star-forming and Quiescent Galaxies at z ∼ 0.8*
de Graaff, A., Bezanson, R., Franx, M., van der Wel, A., Holden, B., van de Sande, J., Bell, E.F., D'Eugenio, F., Maseda, M.V., Muzzin, A., Sobral, D., Straatman, C.M.S.; Wu, P-F., 2021, ApJ, 913, 103

9. *Tightly Coupled Morpho-kinematic Evolution for Massive Star-forming and Quiescent Galaxies across 7 Gyr of Cosmic Time*
de Graaff, A., Bezanson, R., Franx, M., van der Wel, A., Bell, E.F., D'Eugenio, F., Holden, B., Maseda, M.V., Muzzin, A., Pacifici, C., van de Sande, J., Sobral, D., Straatman, C.M.S., Wu, P-F., 2020, ApJL, 903, L30

10. *Probing the missing baryons with the Sunyaev-Zel'dovich effect from filaments*
de Graaff, A., Cai, Y-C., Heymans, C., Peacock, J.A., 2019, A&A, 624, A48.

Submitted publications

1. *A common origin for the Fundamental Plane of quiescent and star-forming galaxies in the EAGLE simulations*
de Graaff, A., Franx, M., Bell, E.F., Bezanson, R., Schaller, M., Schaye, J., van der Wel, A., resubmitted to MNRAS (arXiv e-prints, arXiv:2207.13491)

CURRICULUM VITAE

I was born in Aberdeen, Scotland, but spent the majority of my childhood in The Hague. In the last two years at my secondary school, the Christelijk Gymnasium Sorghvliet, I also attended the Pre-University College in Leiden, where I participated in a wide variety of lectures and projects. Astronomy caught my interest, and my final-year project supervised by Dr. Ivo Labbé confirmed my choice.

Opting for extra adventure, I pursued an integrated master's degree in physics and astronomy at the University of Edinburgh. There, I skipped plenty of lectures to instead build a radio telescope with my friends. I was also an active (committee) member of several student organisations: the local running club, Dutch Society and the Young Scientific Researchers Association. I graduated in May 2018 with First Class Honours, and was awarded for being the top student of my cohort.

My first encounter with scientific research was in 2015, during an exchange year at the University of California, Berkeley. With Dr. Ryan Trainor, I studied the sizes of Lyman-α emitters in *Hubble* imaging. The next summer I took part in the LEAPS programme organised by Leiden Observatory and ESA, where I analysed the extended UV emission in nearby galaxies with Dr. Meiert Grootes. Back in Edinburgh, I completed my Senior Honours project with Dr. Yan-Chuan Cai and Prof. Catherine Heymans, which resulted in the detection of some of the elusive warm-hot baryons in the cosmic web that had been labelled as "missing" for decades. The publication, written with the additional help of Prof. John Peacock, sparked a tumultuous response from the press and scientific community. For my MPhys project, I characterised high-redshift Lyman-α emitters and Lyman Break galaxies in large, ground-based imaging surveys with Prof. Ross McLure. Importantly, these mentors taught me the fundamentals of being a scientist: to ask the right questions, trust my own judgement, and communicate with others.

In September 2018 I started my doctoral research as a de Sitter fellow in Leiden, supervised by Prof. Marijn Franx. Given the freedom to pursue my own interests, I studied the structures of galaxies with state-of-the-art observations and simulations. I have worked with a large number and variety of collaborators, most notably the LEGA-C team. Although the pandemic cancelled many of the trips that I had planned and hoped for, I was fortunate to be able to visit Dr. Rachel Bezanson and Prof. Pieter van Dokkum in early 2020. I have also presented my work at many different conferences and team meetings, both in person and virtually.

In October 2022 I will join the Max Planck Institute for Astronomy in Heidelberg as a postdoctoral fellow, and continue my work at the interface of observations and theory. In particular, as a member of the NIRSpec GTO team since 2018, I look forward to exploring the new data from the *James Webb Space Telescope*.

Acknowledgements

The journey towards the completion of this thesis has been quite the challenge. Three chapters of this thesis were written while in self-isolation due to the pandemic, making for a rather lonely PhD experience. Here, I would like to thank all those that have provided support and, most importantly, made these couple of years more fun.

To start, I want to express my gratitude to the support staff at Leiden Observatory, who were always more than happy to provide help with any practicalities needed. A special thank you to Monica and Marjan, who helped organise many of the Friday borrels and other social events. Ruud, thank you for fixing my bricked laptop, this thesis would otherwise likely have been delayed by months.

When it comes to the science, my collaborators have formed a crucial component to this work and my development as an astronomer. Rachel, Arjen, Eric, Francesco and all others of the LEGA-C team: you provided a welcoming and encouraging environment where I could learn and discuss not just the nitty-gritty details of data processing and analysis, but also different philosophies behind the reading and writing of papers. Michael, Bernd, Hans-Walter, thank you for helping me navigate the world of NIRSpec, the complexity of which reaches far beyond the science. Joop, and other simulators in Leiden, you offered a nice and much-needed counterpart to the world of observers. The galaxies journal club in Leiden provided a fun scientific playground, thanks to Leindert, Themiya, Ken, Sarah, and many others. Lastly, Marijn, I might have a learnt a thing or two from you these past years, it's been all types of fun.

The department in Leiden is truly something special. It was an honour to serve the almighty Big Dipper together with Martijn (and Mantas). Mantas, you were the worst, best and only officemate – I hope you're proud of your medal. Hiddo, you were always up for anything fun, and we still need to complete that hike to the other side of the Grand Canyon. Kirsty, you're my favourite Australian. Thank you for organising the best Halloween parties. Fraser, thank you for always sharing in my dark humour. Marta, I like you (even) more than I do Nala, I promise. Pooneh, our quarantine bubble during the first wave of covid would not have been the same without you. Last but not least, a special thank you to Anniek for enthusiastically showing all of us how to dance the floss.

To the generation of PhD's before me: Jeroen, Dilovan, Andrej, Bas, Patrick, Frits, Gabriella, Dario, Turgay, Lukasz, Christos and Alex, you showed us young 'uns the ropes. Thanks for all the good times and the advice on how to PhD. Jit, thank you for all your support during some of the less fun times. Lydia, Lammim,

Marina, Folkert, I very much enjoyed all our discussions about science, life and beer. To the new/younger folks, thank you for breathing new life into the borrel and the KL puzzles. Josh, you're too nice. Roi, Joey, Nicco, Thijs, Erik, Roland, Margot, Christian, Christiaan, Alfred, Elia, Ivana, Evgenii, Elina, Sil, Jelle, Laura: I probably owe all of you a beer. To Amy, all I can say is: I told you so.

My postdoc/adult friends, Matus and Matthieu, thank you for being my unofficial mentors and teaching me the basics of academia. Matus, I feel very fortunate to be able to call you my friend, but let's agree to never cycle 236 km in torrential rain again. Matthieu, thanks for all the snark. Andrés, let's ride!

To counterbalance some of these astronomers, I'm happy to have people in my life that give me a reality check every now and then. Jamie, Lou, I'm very proud of our longstanding friendship, and our ability to share, understand and celebrate both the small things and major milestones in life. Ryan, do you realise it has been 9 years? I can't wait for more trips to Belgium, to chat about life, food and beer. Julie, Ana, despite having taken very different paths in life, somehow we manage to pick up the conversation right where we left off whenever we talk. Thank you for hosting me for several holidays. Friends from Edinburgh, Khush, Ori, Sarah, Ellen, Luciana and Alex, I'm very happy to still be in touch, and get updates on all your various pets.

Aan mijn familie, ik ben erg blij en dankbaar dat ik de laatste jaren zo dicht bij huis heb kunnen zijn, zeker tijdens de meerdere turbulente periodes. Ik zal de spontane fietsritjes naar Den Haag zeker gaan missen. Boris, Gabriëlle, Tijmen (en Hera) het is altijd een gezellige boel bij jullie thuis, met vrije plaats voor zowel de leuke als de meer serieuze gesprekken. Mam, jouw levenslessen over eigenwijs en zelfstandig zijn, zijn zeker goed van pas gekomen de afgelopen jaren. Pap, volgens mij ben jij mijn grootste fan, voor zover sterrenkundigen groupies kunnen hebben. Toen je ruim 10 jaar terug een "offhand" opmerking maakte dat ik wel met James Webb zou kunnen werken tijdens mijn studie, was mijn eerste gedachte dat ik daar vast niet goed genoeg voor zou zijn. Maar, ouders blijken wel vaker gelijk te hebben.

Lastly, to my grandma, who studied chemistry in Leiden in the 1940s, but was informed that there was no space for women in the lab. Sadly, you did not get to see me pursue my PhD. Sexism is unfortunately still rife in academia. Nevertheless, may this thesis serve as a testament to the fact that there is a place for women in the sciences, including at the University of Leiden.