

Elementary

The chemical fingerprints of massive galaxy formation
over cosmic time

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr. S. de Rijcke,
volgens besluit van het college voor promoties
te verdedigen op donderdag 25 juni 2026
klokke 13:00 uur

door

Chloe Mai Tam Cheng

geboren te Toronto, Canada
in 1998

Promotores: Prof.dr. M. Kriek
Prof.dr. P. G. van Dokkum (Yale University)

Promotiecommissie: Prof.dr. H. Hoekstra
Prof.dr. I. A. G. Snellen
Dr. J. A. Hodge
Prof.dr. A. van der Wel (Universiteit Gent)
Dr. R. Bezanson (University of Pittsburgh)

ISBN: 978-94-6534-444-7

Cover:
Design and layout by Jessica Lin

The price of success is steep and I've never been able to distinguish it from the feeling of sacrifice. If I could hold success in my hand, it would be a beating heart.
– Weike Wang, *Joan is Okay* (2022)

*There is no problem in science that can be solved by a man that cannot be solved
by a woman.*
– Vera C. Rubin, *Bright Galaxies, Dark Matter* (1996)

CONTENTS

1	Introduction	1
1.1	Our working picture of massive galaxy formation	1
1.2	Stellar populations: tracers of galaxy formation	4
1.2.1	A stellar population cookbook	4
1.2.2	Using stellar populations to investigate chemical enrichment	7
1.2.3	Deducing stellar population properties with photometry	10
1.2.4	Stronger inferences with spectroscopy	12
1.3	Observational evidence for galaxy formation	13
1.3.1	Clues from the low-redshift Universe	15
1.3.2	Closing in on galaxy formation at higher redshifts	17
1.4	This thesis	20
1.5	Conclusions and outlook	21
2	Age and metal gradients in massive quiescent galaxies at $0.6 < z < 1.0$: implications for quenching and assembly histories	23
2.1	Introduction	24
2.2	Data and sample	26
2.2.1	LEGA-C spectra	26
2.2.2	Sample selection	27
2.3	Methods	29
2.3.1	Optimal extraction	29
2.3.2	Convolved R_e	31
2.3.3	Full spectrum fitting	33
2.4	Results	34
2.4.1	Observed gradients	34
2.4.2	Intrinsic metallicity gradients	39
2.5	Discussion	42
2.5.1	Colour gradients	42
2.5.2	Implications for galaxy evolution	45
2.5.2.1	Young quiescent galaxies	45
2.5.2.2	Growth of the quiescent galaxy population	46
2.5.2.3	Future work	47
2.5.3	Caveats	48
2.6	Summary and conclusion	50
	Appendices	53
2.A	Mock spectra tests	53

2.B Duplicate objects	56
3 Ages and metallicities of quiescent galaxies: confronting broad-band (UVJ) colours with stellar absorption lines	59
3.1 Introduction	60
3.2 Data and methods	62
3.2.1 LEGA-C	62
3.2.2 Spectral extraction and full-spectrum modelling	64
3.3 Results	65
3.3.1 Trends with age and metallicity	66
3.3.2 Comparison to stellar population models	67
3.4 Discussion	74
3.4.1 Comparison to previous work	74
3.4.2 Implications for galaxy evolution studies	75
3.4.3 Implications for SPS models	79
3.4.4 Caveats	81
3.5 Summary and conclusions	83
4 Clues to inside-out quenching in quiescent galaxies at $1.2 \lesssim z \lesssim 2.2$: Age, Fe-, and Mg-abundance gradients from $JWST$-SUSPENSE	87
4.1 Introduction	88
4.2 Data and sample	90
4.3 Methods	93
4.3.1 Correction of resampling noise	93
4.3.2 Spectral extraction	94
4.3.3 Determining distances and radii	95
4.3.4 Full spectrum fitting	97
4.4 Results	100
4.4.1 Spatially resolved ages and elemental abundances	100
4.4.2 Trends between gradients and galaxy parameters	104
4.5 Discussion	105
4.5.1 Implications for galaxy evolution	105
4.5.1.1 Quenching at $z \sim 2$	106
4.5.1.2 Galaxy evolution towards $z \lesssim 1$	107
4.5.2 Caveats	108
4.6 Summary and conclusions	110
Appendices	112
4.A Mock observations	112
4.B Wiggle correction algorithm	112
5 Hidden mass in early galaxies revealed by bottom-heavy initial mass functions	117
5.1 Article	118
5.2 Methods	127
5.2.1 Observing Strategy and Data Reduction	127
5.2.2 Full-spectrum fitting	128

5.2.3	Stellar velocity dispersions and virial masses	132
5.3	Extended Data	133
5.4	Supplementary Information	140
	Bibliography	140
	English Summary	163
	Nederlandse samenvatting	171
	Publications	179
	Curriculum Vitae	183
	Acknowledgments	187

1 | INTRODUCTION

1.1 Our working picture of massive galaxy formation

Massive galaxies are thought to form in two phases. In particular, their compact cores may have formed via in-situ star formation from cold gas flows, which quenched rapidly and efficiently at cosmic noon ($z \sim 2-3$), or even earlier. Subsequently, their outskirts may have been built up by merging with low-mass satellite galaxies, up until the present day (see Figure 1.1, Naab et al. 2009; Oser et al. 2010; Rodriguez-Gomez et al. 2016).

This two-phase model is in line with our cosmological framework. In the standard cold dark matter (Λ CDM) paradigm, structure forms hierarchically. All cosmic structures initially formed via the expansion of random quantum fluctuations following the Big Bang. As the Universe continued to expand and cool, gravitational instabilities caused these fluctuations to grow, condensing into galaxies and clusters of galaxies. Thus, in this model, small structures merge with larger structures to form the fabric of our Universe (White & Rees 1978; Spergel et al. 2007; Mo et al. 2010; Springel et al. 2018).

Observationally, the two-phase model is supported by findings that a large fraction ($\sim 40-60\%$) of massive galaxies were already quiescent by $z \sim 2-4$ (e.g., Smail et al. 2002; Daddi et al. 2003, 2005; Franx et al. 2003; Kriek et al. 2006, 2008a,b; Muzzin et al. 2013a; Spitler et al. 2014; Straatman et al. 2014, 2016). Moreover, these galaxies have been observed to grow in size by a factor of ~ 4 from $z \sim 2$ to $z \sim 0$, while their central densities (within ~ 1 kpc) remain largely unchanged. This size growth has largely been attributed to minor mergers (e.g., Daddi et al. 2005; Trujillo et al. 2006; Toft et al. 2007; Franx et al. 2008; van Dokkum et al. 2008; Bezanson et al. 2009; Damjanov et al. 2009, 2011; van der Wel et al. 2014). Furthermore, cosmological simulations (e.g. Illustris, FIRE, Millennium II, see Cooper et al. 2013; Somerville & Davé 2015; Anglés-Alcázar et al. 2017; Rodriguez-Gomez et al. 2016; Crain & van de Voort 2023) can reproduce these observations under the conditions of the two-phase model.

While this galaxy formation paradigm is strongly supported by observational and theoretical evidence, alternative formation mechanisms are also possible. For example, some studies propose that massive galaxies form through major, gas-rich

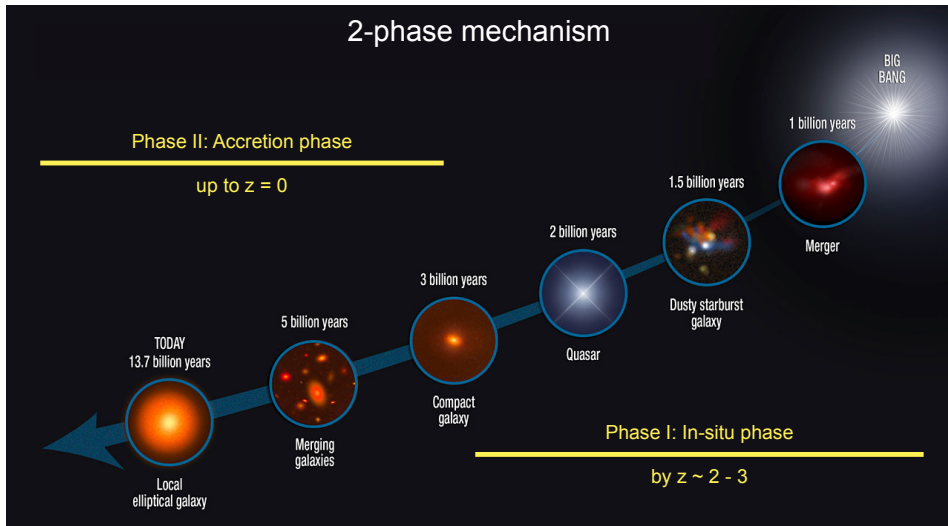


Figure 1.1: Illustration of the two-phase model of galaxy formation. The compact cores of today’s massive, early-type galaxies are thought to have formed via in-situ star formation, which quenched around $z \sim 2 - 3$. Their outskirts were then built up by the accretion of low-mass satellite galaxies up until $z \sim 0$ (Naab et al. 2009; Oser et al. 2010; Rodriguez-Gomez et al. 2016). Figure adapted from NASA/ESA/Sune Toft.

mergers (e.g., Khochfar & Silk 2006; Naab et al. 2007; Hopkins et al. 2009b). This scenario can explain the evolution in galaxy sizes, however models implementing major mergers imply excessive mass growth and low number densities compared to observations of present-day galaxies (Bezanson et al. 2009). Galaxy size growth could also be explained by the expansion of galaxies due to extreme mass loss from quasar feedback (Fan et al. 2008), however this mechanism under-predicts central densities of local elliptical galaxies (Bezanson et al. 2009). Furthermore, the evolution in galaxy metallicities cannot fully be explained by dry minor mergers, perhaps requiring contributions from major dry mergers (Beverage et al. 2024; Kriek et al. 2024). On the other hand, it may be that individual galaxies are not growing at all. Instead, larger galaxies may quench at later times as a result of the average density of the Universe increasing over cosmic time, a phenomenon known as progenitor bias (van Dokkum & Franx 2001; Carollo et al. 2013; Poggianti et al. 2013). Of course, individual galaxies will likely be affected by a combination of processes (van der Wel et al. 2008; Hopkins et al. 2010; Valentinuzzi et al. 2010; Oser et al. 2012; Newman et al. 2012; Nipoti et al. 2012; Barro et al. 2013), with, for example, different mechanisms affecting galaxies of different masses (Carollo et al. 1993; Rodriguez-Gomez et al. 2016) or acting at different redshifts (Belli et al. 2014, 2015, 2017; Wellons et al. 2015, 2016; Suess et al. 2021). Nevertheless, uncovering the dominant mechanisms responsible for galaxy assembly is crucial to our broader understanding of their evolution.

Equally important to our understanding of galaxy evolution are the possible

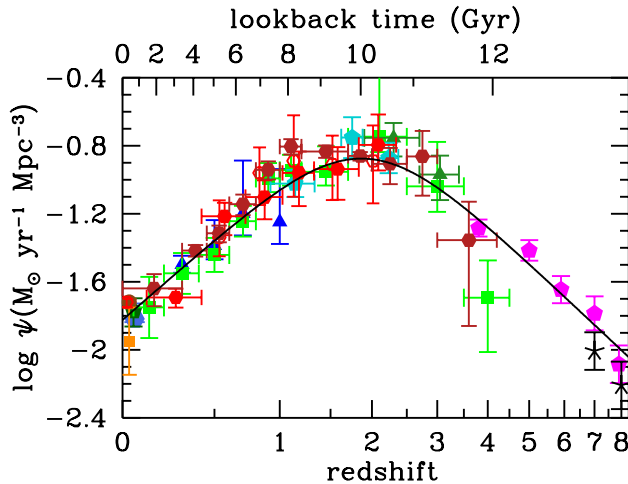


Figure 1.2: Observed star formation rate densities as a function of redshift. This Figure depicts the cosmic star formation history, suggesting that star-formation and quenching activity peaked at $z \sim 2$. Figure reproduced from Madau & Dickinson (2014).

mechanisms by which galaxies halt their star formation, or quench. In particular, the exact pathway towards quenching, and why it appears to peak at $z \sim 2 - 3$ in the two-phase picture, must be constrained (see Figure 1.2, Oser et al. 2010; Madau & Dickinson 2014; Rodriguez-Gomez et al. 2016; Cortese et al. 2021). There are a variety of proposed quenching mechanisms, both internal (e.g. morphological/gravitational quenching, Martig et al. 2009; Genzel et al. 2014; or feedback from star formation, active galactic nuclei, or accreting supermassive black holes, Di Matteo et al. 2005; Bower et al. 2006; Croton et al. 2006; Renzini 2009; Bluck et al. 2020) and environmental (e.g. halo-quenching, ram pressure stripping, tidal interactions, mergers, or harassment, Dekel & Birnboim 2006; van den Bosch et al. 2008; Weinmann et al. 2009; Wetzel et al. 2012). Regardless of the specific process, however, a galaxy’s cold gas reservoir must somehow be depleted, removed, or heated in order for star formation to cease (see, e.g., Cortese et al. 2021).

Disconcertingly, new observations from the state-of-the-art James Webb Space Telescope (*JWST*) have revealed extremely massive galaxies that were already quiescent in the very early Universe (Steinhardt et al. 2016; Carnall et al. 2023a,b, 2024; Valentino et al. 2023; Glazebrook et al. 2024; Gottumukkala et al. 2024; Xiao et al. 2024; de Graaff et al. 2025; Ito et al. 2025; Weibel et al. 2025). These findings challenge our theories of massive galaxy evolution, implying that galaxies formed and quenched faster and earlier than predicted by cosmological simulations (e.g. Baker et al. 2025; Lagos et al. 2025). Thus, these new observations are putting pressure not only on quenching and assembly theories, but also on cosmological theories of structure formation in the Universe.

1.2 Stellar populations: tracers of galaxy formation

In the era of *JWST*, it has become imperative for us to reveal the dominant modes of galaxy assembly and quenching, and the mechanisms by which subsequent evolution occurs. To achieve this goal, we can study the fossil records of galaxies over cosmic time, which are encoded in their detailed stellar population properties.

1.2.1 A stellar population cookbook

Stellar populations constitute the fundamental building blocks of galaxies (Baade 1944; Oort 1958) and represent groups of stars spanning a range of ages, masses, and chemical compositions. Interpreting the light from these stars thus requires a detailed understanding of the key, interconnected ingredients on which stellar populations are built, including the initial mass function, stellar evolution, and stellar spectra (Tinsley 1979; Walcher et al. 2011; Conroy 2013). These ingredients are described below and are summarized in the top row of Figure 1.3.

First, each star in a stellar population is born with a certain mass. The distribution of these birth stellar masses, known as the initial mass function (IMF, e.g. Salpeter 1955), is crucial for describing a stellar population. In particular, low-mass stars typically dominate the mass and number budget, while high-mass stars are much less numerous (see Conroy 2013). In a wide range of environments within our own Milky Way (MW), the IMF has been directly measured via resolved counts of individual stars. It has been found to be well-characterized by a single prescription (Salpeter 1955; Kroupa 2001; Chabrier 2003), typically parameterized as:

$$\Phi(M) = \frac{dN}{dM} = \begin{cases} k_1 M^{-1.3}, & 0.08 < M < 0.5 \\ k_2 M^{-2.3}, & 0.5 < M < 1.0 \\ k_3 M^{-2.35}, & M \geq 1.0 \end{cases} \quad (1.1)$$

where M is stellar mass and k_i represent normalization constants.

Each of the stars within a stellar population evolves according to a stellar evolutionary track. To describe the evolution of stars of all stellar masses, we can make use of isochrones in the Hertzsprung-Russell (HR) diagram. Isochrones trace the evolution in effective temperature (T_{eff}) and luminosity of a population with a certain age and metallicity, with both the timescale and possible evolutionary phases of the constituent stars being dependent on their initial masses. For example, dwarf stars ($< 0.8 M_{\odot}$) have remained on the main sequence since the Big Bang, while more massive stars evolve quickly and die in core-collapse supernovae (CCSNe), within a few Myr (e.g. Tinsley 1980; Dotter 2016, see the middle panel of the top row of Figure 1.3). These stellar evolution models must take into account complex processes including mass loss, rotation, binary interactions, and convection (see Conroy 2013).

Finally, each star within a stellar population emits light that we can observe as a spectrum, with stars of different masses and at different evolutionary stages having varying effects on the spectral shape, the presence and depths of different

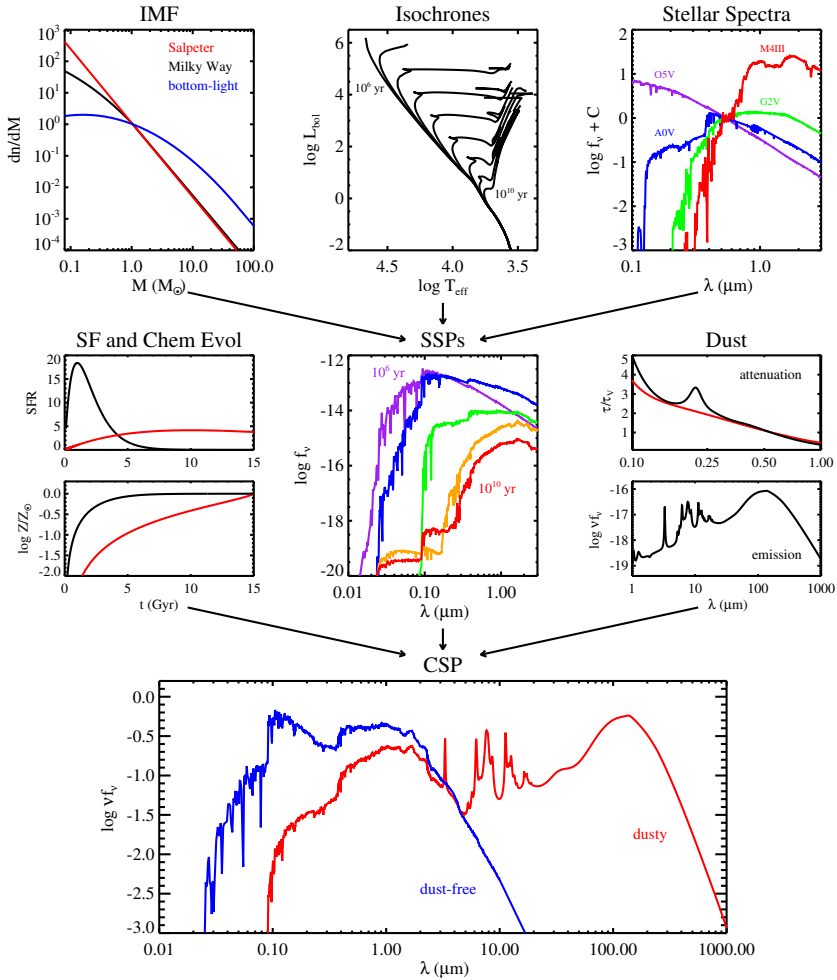


Figure 1.3: Summary of the procedure for building stellar population synthesis models. The key ingredients which make up stellar populations (see Section 1.2.1) and which are used to build simple stellar populations (see Section 1.2.3) are shown in the upper panels. The ingredients used to build composite stellar populations are shown in the middle panels. The bottom panel shows resulting composite stellar populations with and without an applied dust model. Figure reproduced from Conroy (2013).

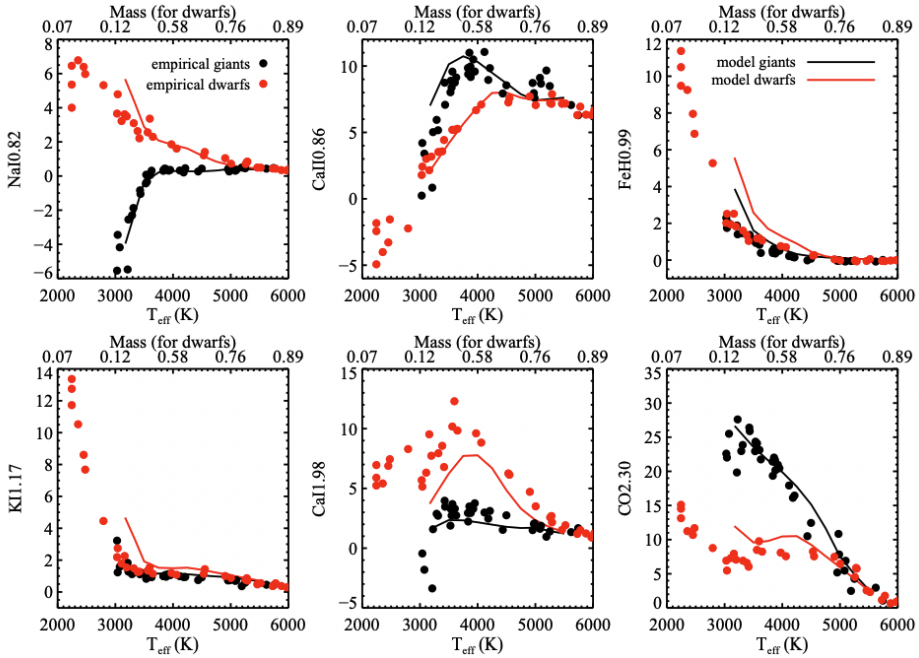


Figure 1.4: The dependence of surface-density-sensitive spectral indices on T_{eff} (converted to dwarf stellar masses via composite isochrones) for giant stars (black symbols) compared to dwarf stars (red symbols). Lines represent the strength of indices from model synthetic spectra generated in Conroy & van Dokkum (2012a). Points represent line strengths from empirical IRTF spectra (Cushing et al. 2005; Rayner et al. 2009). Figure reproduced from Conroy & van Dokkum (2012a).

spectral lines, and the total luminosity/scaling of the spectrum. In particular, main sequence O-type stars have high T_{eff} and bluer, luminous spectra that are dominated by He II lines. Strong He I lines and weak Balmer features (H I lines) begin to develop in B-type stars. These Balmer features are the strongest in A-type stars. Metal lines (e.g. Ca II, Fe I) also emerge in A-type stars and strengthen through F-type stars, becoming most prominent in solar- (G) and K-type stars. Molecular absorption bands (e.g. TiO, VO) begin to appear in the fainter, redder spectra of cooler M-type stars, with metal lines continuing to persist (e.g., Cannon & Pickering 1924; Morgan et al. 1943; Morgan & Keenan 1973; Gray & Corbally 2009; Carroll & Ostlie 2017).

In practice, we observe an integrated spectrum, which is the sum of the spectra of individual stars within a stellar population. Just as the properties of a single star affect its spectrum, the spectrum of a stellar population is influenced by the average properties of its constituent stars. In particular, while the stellar mass of a population is generally dominated by low-mass stars, as discussed above, the less-numerous high-mass stars monopolize the light that we see. This is due to the fact that high-mass stars have higher core temperatures and burn through their

fuel much faster than low-mass stars, resulting in a higher luminosity per unit mass. This difference in fuel consumption also results in shorter lifetimes. Thus, as a stellar population ages, its more massive, bluer stars evolve off of the main sequence and die (see the middle panel in the top row of Figure 1.3). At older ages, the low-mass stars therefore have an increased contribution, reddening the spectrum and lowering the luminosity.

While the dependencies of spectral shapes and line strengths on stellar population age seem relatively straightforward as described above, there are several other factors which complicate this picture. For example, evolved stellar stages, such as thermally-pulsating asymptotic giant branch (TP-AGB) stars, are prominent at young ages (< 2 Gyr), but their effects on the spectra are not well-constrained. These stars can have a significant effect on spectral shapes in the near-infrared, and on the strength of molecular bands such as CN, H₂O, and CO (Maraston 2005; Maraston et al. 2006; Kriek et al. 2010; Walcher et al. 2011; Conroy 2013).

Additionally, the metallicity of a stellar population can have a considerable effect on its spectrum. Metal lines, which are affected by age as discussed above, are of course also affected by metallicity. This results in a well-known degeneracy between age and metallicity, where an increase or decrease in stellar population age by a factor of three has nearly the same effect on the data as an increase or decrease in metallicity by a factor of two (Worthey 1994; Bruzual & Charlot 2003; Gallazzi et al. 2005). This issue is exacerbated by additional degeneracies between age and dust, and age and star-formation history (SFH, Bell & de Jong 2001; Papovich et al. 2001; Leja et al. 2017, 2019a,c).

Finally, the precise shape of the IMF can affect the strengths of the surface-gravity-sensitive features. This is due to the fact that the spectra of red dwarfs and red giants have different depths in features including Na I ($\sim 8180 - 8200 \text{ \AA}$), the calcium triplet (CaT, $\sim 8475 - 8725 \text{ \AA}$), and the Wing-Ford band ($\sim 9905 - 9945 \text{ \AA}$, Wing & Ford 1969), due to differences in surface gravity of these types of stars at the same T_{eff} (see Figure 1.4). For example, an excess of red dwarfs compared to what has been found in the MW (see Equation 1.1, also called a bottom-heavy IMF) would strengthen surface-gravity-sensitive features in integrated spectra, while a paucity of red dwarfs (i.e. a bottom-light IMF) may result in shallower features (see, e.g., Conroy & van Dokkum 2012b). This relationship between the IMF and surface-gravity-sensitive features is further complicated by the sensitivity of these features to individual abundances. Thus, strong degeneracies also exist between variations in the IMF and elemental abundances (e.g., Conroy & van Dokkum 2012a; Smith 2020; Lonoce et al. 2021).

1.2.2 Using stellar populations to investigate chemical enrichment

Detailed stellar population properties (i.e. elemental abundances and ages) provide a window into the processes that drove galaxy evolution. In particular, the total metal content of a galaxy reflects both its past star formation, as well as the interplay with the surrounding medium (i.e. inflows, outflows, and stripping). Additionally, elemental abundances can give us insight into the chemical enrichment

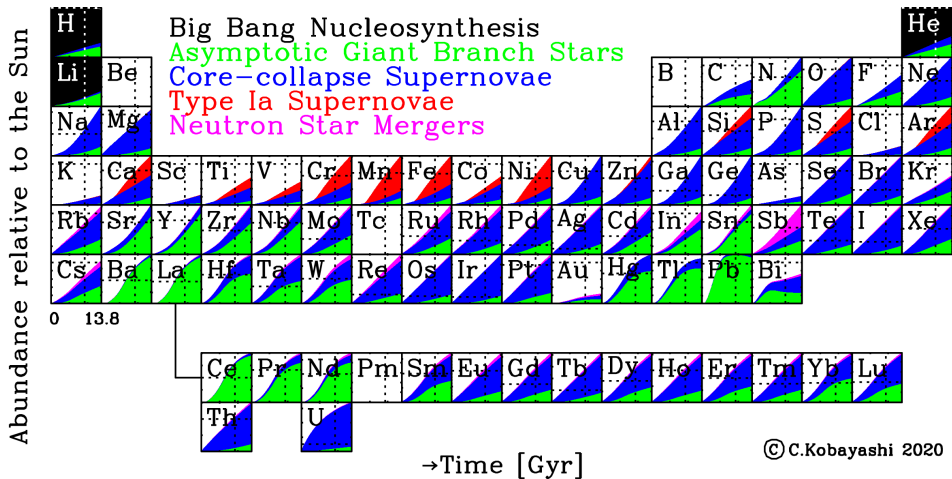


Figure 1.5: A visualization of the production pathways and timescales of elements in the periodic table. Dotted lines represent the values observed for a solar abundance pattern. Figure reproduced from Kobayashi et al. (2020), see also Nomoto et al. (2013).

and star formation histories of a galaxy, as various types of stars release different elements on varying timescales. We typically define the relative abundance of two arbitrary elements (X and Y) in terms of their relative number densities (N) relative to the solar value as

$$[X/Y] = \log \left(\frac{N_X}{N_Y} \right) - \log \left(\frac{N_X}{N_Y} \right)_{\odot} \quad (1.2)$$

(e.g., Maiolino & Mannucci 2019).

These relative abundances allow us to constrain star-formation and galaxy assembly histories, as different elements are enriched by distinct types of stars on varying timescales (e.g., Nomoto et al. 2013; Maiolino & Mannucci 2019; Kobayashi et al. 2020, see Figure 1.5 and the top panel of Figure 1.6). For example, the relative abundance of α elements (e.g. O, Ne, Mg) to iron-peak elements (e.g. Fe, Ni), expressed as $[\alpha/\text{Fe}]$, is an excellent tracer of star-formation timescales (Tinsley 1979; Maiolino & Mannucci 2019, although see Martín-Navarro 2016). This is a result of the fact that massive stars ($> 8 M_{\odot}$) undergoing CC SNe generally release α elements via shock heating and explosive nucleosynthesis of stellar material on relatively brief timescales, shortly after star formation begins.

In contrast, iron-peak elements (as well as elements such as Si, Ar, S, and Ca) are generally released via Type Ia supernovae (SNe) on longer timescales. These timescales, ranging from $\sim 40 - 50$ Myr to a few Gyr later than the α elements (depending on the assumed stellar IMF and SFH, see Figure 1.6), are due to the fact that Type Ia SNe occur after the formation of a white dwarf from a low-mass ($< 8 M_{\odot}$) star. This white dwarf formation can then be followed by mass exchange with a companion star, which can result in carbon-burning in the C-O core, triggering a Type Ia explosion. This process results in a delay between the

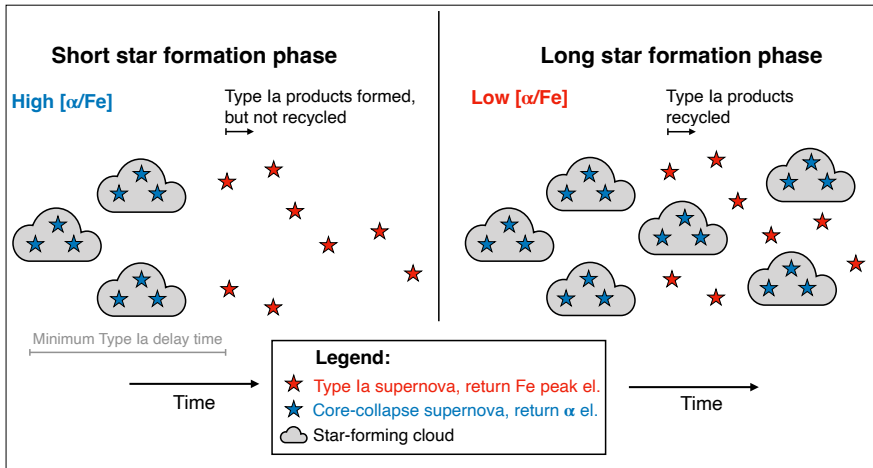
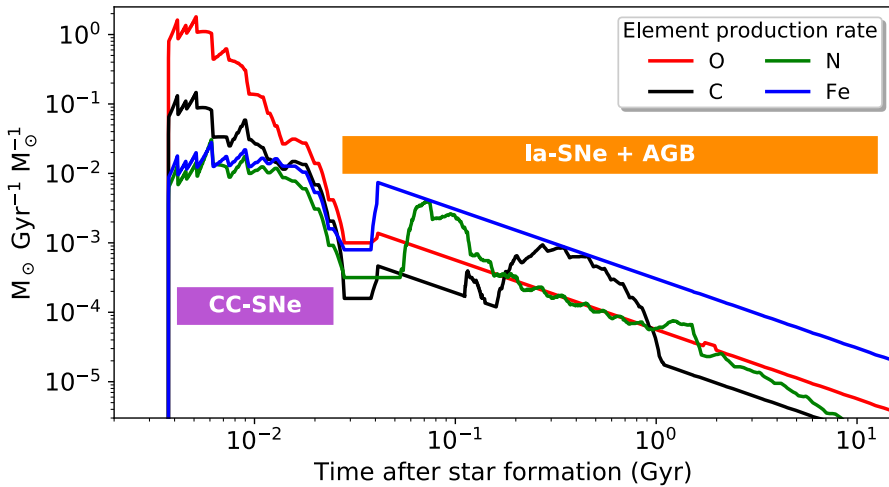


Figure 1.6: *Top:* Production rate of various elements from a solar-metallicity SSP model (from F. Vincenzo). The production rate (M_{\odot}/Gyr) of stars formed, normalized to $1 M_{\odot}$, is shown on the y -axis. The x -axis shows the amount of time after a single episode of SF. Oxygen (red line) has the shortest formation timescales, due to its primary production in core collapse supernovae. Iron (blue) is primarily produced by Type Ia supernovae. Both kinds of supernovae and asymptotic giant branch stars produce carbon (black). Nitrogen (green) is primarily produced by asymptotic giant branch stars. Figure adapted from Maiolino & Mannucci (2019). *Bottom:* Schematic diagram of how $[\alpha/\text{Fe}]$ can be used to trace star-formation timescales. Figure courtesy of Mariska Kriek and Aliza Beverage.

production of Type Ia SNe and the onset of star formation.

Finally, elements such as C and N are also enriched on intermediate timescales (i.e. longer than CC SNe but shorter than Type Ia SNe), ranging from $\sim 0.04 - 5$ Gyr. This delayed enrichment is due to the fact that these elements are primarily produced by intermediate mass stars ($2 M_{\odot} < M_{\text{star}} < 8 M_{\odot}$), including asymptotic giant branch (AGB) winds or Wolf Rayet stars.

By examining stellar metallicities and elemental abundance ratios, we can indirectly constrain galaxies' star formation and chemical enrichment histories. For example, a high $[\alpha/\text{Fe}]$ ratio is indicative of short star-formation timescales due to the quick progression of CC SNe. On the other hand, a low $[\alpha/\text{Fe}]$ ratio suggests long star-formation timescales. In this case, Type Ia SNe have had time to explode and enrich the interstellar medium, with their products being recycled into the next generation of stars, increasing the relative proportion of iron-peak elements (e.g., Maiolino & Mannucci 2019, see the bottom panel of Figure 1.6). However, while $[\alpha/\text{Fe}]$ is typically used as a proxy for star-formation timescales in this way, it is important to keep in mind that these abundance ratios are also affected by star-formation efficiency, late-time mergers, outflows, and the IMF (see, e.g., Tinsley 1979; Worthey et al. 1992; Zolotov et al. 2010; Martín-Navarro 2016; Andrews et al. 2017; Sybilka et al. 2018). In particular, a bottom-heavy IMF in conjunction with a high $[\text{Mg}/\text{Fe}]$ ratio implies unrealistically short and intense star formation episodes in nearby, massive, elliptical galaxies (Martín-Navarro 2016).

1.2.3 Deducing stellar population properties with photometry

The physical principles described above provide a framework for interpreting observations of stellar populations, which we can apply to real data. Photometric data, which are relatively inexpensive to obtain, are often the first step towards understanding stellar population properties for large samples of galaxies. These data are used to obtain broad-band spectral energy distribution (SED) shapes (either in an integrated or spatially resolved way), which are fit with stellar population synthesis (SPS) models in order to extract physical properties (Tinsley 1972; Searle et al. 1973; Larson & Tinsley 1978, see also Walcher et al. 2011; Conroy 2013).

SPS models are constructed by combining the key ingredients on which stellar populations are built (see Section 1.2.1). These ingredients are implemented in different ways depending on the model. First, these models depend on simple stellar populations (SSPs). Given a set of isochrones describing stellar evolution, a stellar spectral library (empirical, theoretical, or a combination), and an IMF, SSPs describe the evolution of the SED of a stellar population with a single age, metallicity, and abundance pattern. To generate a time- (t) and metallicity-dependent (Z) SSP spectrum (f_{SSP}), the above ingredients are generally combined as in Equation (1) in Conroy (2013) (although see Renzini & Buzzoni 1986; Maraston 1998; Bica

& Allain 1986 for alternative SSP constructions):

$$f_{\text{SSP}}(t, Z) = \int_{m_{\text{lo}}}^{m_{\text{up}}(t)} f_{\text{star}}[T_{\text{eff}}(M), \log g(M)|t, Z] \Phi(M) dM \quad (1.3)$$

where M is the initial (zero-age main sequence) stellar mass, $\Phi(M)$ is the IMF, f_{star} is a stellar spectrum, m_{lo} is typically taken to be the hydrogen burning limit (here we take this to be $0.08 M_{\odot}$), and $m_{\text{up}}(t)$ is dictated by stellar evolution. The isochrones determine the relation between T_{eff} and luminosity or surface gravity ($\log g$) for a given M , t , and Z . Combinations of SSPs, along with models for dust attenuation, star formation histories, and chemical evolution, can then be used to construct composite stellar populations (CSPs). This process is summarized in Figure 1.3.

To measure galaxy properties, SPS models, using either SSPs or CSPs as their base, are fit to broadband SEDs to measure galaxy properties. These fits may be achieved using χ^2 minimization (e.g., Heavens et al. 2000; Tremonti et al. 2004; Ocvirk et al. 2006; Cid Fernandes et al. 2005; Walcher et al. 2006; Chilingarian et al. 2007; Tojeiro et al. 2007; Koleva et al. 2009), Markov chain Monte Carlo (MCMC, e.g., Conroy et al. 2009; Acquaviva et al. 2011; Conroy & van Dokkum 2012a; Conroy et al. 2018; Beverage et al. 2025), nested sampling (e.g., Chevallard & Charlot 2016; Carnall et al. 2018; Johnson et al. 2021; Beverage et al. 2025), or Principal Component Analysis (PCA) techniques (e.g., Connolly et al. 1995; Budavári et al. 2009), for example.

These model fits can reveal information about stellar mass-to-light (M/L) ratios and masses, SFHs, specific star-formation rates (sSFRs), and dust properties (e.g., Bell et al. 2003; Zibetti et al. 2009; Leja et al. 2019a; Abdurro'uf et al. 2021; Johnson et al. 2021; see also Walcher et al. 2011; Conroy 2013). They can also be used to infer galaxy ages and metallicities (e.g., Bell & de Jong 2000; MacArthur et al. 2004; Lee et al. 2007; Eminian et al. 2008). However, key uncertainties remain. For example, robust age and metallicity constraints are challenging to achieve due to the strong degeneracies between galaxy properties discussed in Section 1.2.1 (Worthey 1994; Bell & de Jong 2001; Papovich et al. 2001; Bruzual & Charlot 2003; Gallazzi et al. 2005; Leja et al. 2017, 2019b,c).

Moreover, star formation rates, stellar M/L ratios, and stellar masses are still uncertain, as they rely on the assumed shape of the IMF. The idea of a ‘universal’ IMF of the form in Equation 1.1 has arisen as a result of direct constraints in the MW and nearby regions (Sagar et al. 2001; Da Rio et al. 2012; Peña Ramírez et al. 2012; Andersen et al. 2017; Suárez et al. 2019; Damian et al. 2021), with most SPS models thus assuming this MW IMF. However, the number of stellar systems close enough for their stars to be resolved to directly measure the IMF is limited. As a result, these measurements are based on a biased sample and therefore possibly represent a biased IMF assumption. Theoretical studies also do not generally support the idea of a universal IMF (e.g., Schwarzschild & Spitzer 1953; Larson 1986; Kroupa 2001; Hennebelle & Chabrier 2008; Krumholz et al. 2011; Hopkins 2012, 2013; Chabrier et al. 2014). Thus, the IMF represents one of the most significant sources of systematic uncertainty in galaxy evolution studies.

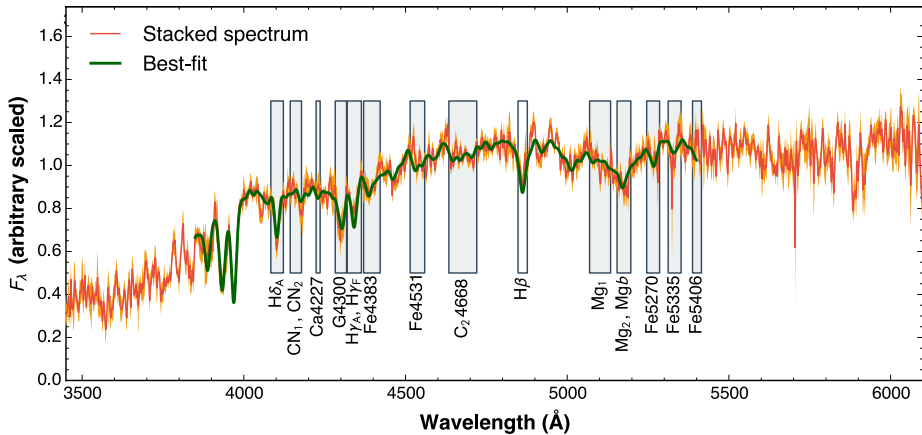


Figure 1.7: The stacked spectrum (red) of a sample of quiescent galaxies at $z \sim 1.6$, with 1σ uncertainties shown by the orange shaded region. Lick indices are indicated by grey rectangles. A best-fit model is shown by the green line (Onodera et al. 2015). Figure reproduced from Maiolino & Mannucci (2019).

The degeneracy between age and dust can be broken with longer wavelength photometry (i.e. UVJ colours; Burgarella et al. 2005; Labbé et al. 2005; Wuyts et al. 2007; Leja et al. 2017). However, while previous photometric studies have attempted to disentangle the degeneracy between age and metallicity (e.g., Bell & de Jong 2000; MacArthur et al. 2004; Lee et al. 2007; Eminian et al. 2008), this is not possible due to the nearly identical effects of age and metallicity on continuum shapes (Worthey 1994). Finally, the IMF cannot be constrained with photometry, as its subtle effect on surface-gravity-sensitive features cannot be detected using broad SED shapes. More robust constraints are therefore required to understand galaxy evolution in detail.

1.2.4 Stronger inferences with spectroscopy

To break the age-metallicity degeneracy, SPS models (as described in Section 1.2.3) are fit to spectroscopic data. These higher-spectral-resolution observations also give us crucial information about individual elemental abundances (i.e. separating metallicity into different elements) and the IMF.

There are two main methods by which we can extract detailed elemental abundance and IMF information from spectra. Historically, stellar population properties have been estimated using spectral indices (see Figure 1.7), such as the Lick/IDS set of indices (Burstein et al. 1984; Worthey et al. 1994, see also Fanelli et al. 1992; Alvarez et al. 2000; Cenarro et al. 2001; Serven et al. 2005). Indices are defined by a feature bandpass and a set of pseudocontinua, from which equivalent widths (EWs) can be measured and compared with SPS model predictions. Index fitting methods take advantage of the sensitivities of individual features to specific properties. For example, the strength of Balmer absorption lines in combination

with other age and metallicity-sensitive features (i.e. the depth of the 4000 Å break/ $D_n(4000)$ index, Hamilton 1985; Balogh et al. 1999), can allow us to constrain stellar population ages. In particular, Balmer lines are strongest in A-type stars (as discussed in Section 1.2.1), and are thus most sensitive to stellar populations around ~ 1 Gyr old (Trager & Somerville 2009; Conroy 2013; Maiolino & Mannucci 2019). Additionally, indices such as [MgFe]' and [Mg₂Fe] were defined to be sensitive to individual elemental abundances (Thomas et al. 2003; Bruzual & Charlot 2003). Finally, spectral features that are sensitive to the relative fraction of dwarf to giant stars (i.e. Na I, CaT, Wing-Ford band, see Section 1.2.1) can be used to constrain the IMF (e.g., Conroy & van Dokkum 2012a; Smith 2020; Lonoce et al. 2021).

In practice, however, stellar population properties contribute to the signal of the entire spectrum, and thus individual indices are also sensitive to secondary effects. Specifically, clean regions of the spectrum from which the continuum level can be estimated are rarely available. As a result, the strength of each index depends not only on the property of interest, but also on the properties to which the pseudocontinua are sensitive (Conroy 2013). Additionally, as discussed in Section 1.2.1, spectral lines are sensitive to multiple properties (e.g., Worthey 1994; Papovich et al. 2001; Bruzual & Charlot 2003; Gallazzi et al. 2005; McConnell et al. 2016; Zieleniewski et al. 2017; Lonoce et al. 2021).

To mitigate this issue and robustly break the degeneracies, we can fit entire spectra with full-spectrum SPS models, taking advantage of all available spectral information over a wide wavelength range (see Figure 1.7). Thus, under the assumption that SPS models are adequately detailed and accurate, it is possible to correctly measure ages, metallicities, and elemental abundances and, in some cases, separate them from IMF variations (Lonoce et al. 2021). This is demonstrated in Figure 1.8 for IMF and abundance variations specifically. Several different SPS models are shown, with each one having different variations in the two low-mass slopes and the low-mass cutoff of the IMF (see Equation 1.1), as well as in various elemental abundances. This Figure shows that, while an increase in the two low-mass slopes of the IMF can result in visually similar changes in the models near features such as NaD, TiO, Na I, and CaT, the changes near surface-gravity-sensitive features (Na I, CaT, Wing-Ford band) are slightly different with respect to each other. These features can thus be used to break the IMF and elemental abundance variation degeneracies. See Gu et al. (2022) for more details.

1.3 Observational evidence for galaxy formation

The tools described in Section 1.2 above have been applied extensively to photometric and spectroscopic observations of galaxies over cosmic time. These measurements have helped to build up our understanding of galaxy evolution. However, they have also raised additional questions and uncertainties as we have pushed to higher redshifts and more extreme galaxies. Here, we will summarize some of the main results and outstanding issues gleaned from observations at low (Section 1.3.1) and high (Section 1.3.2) redshifts.

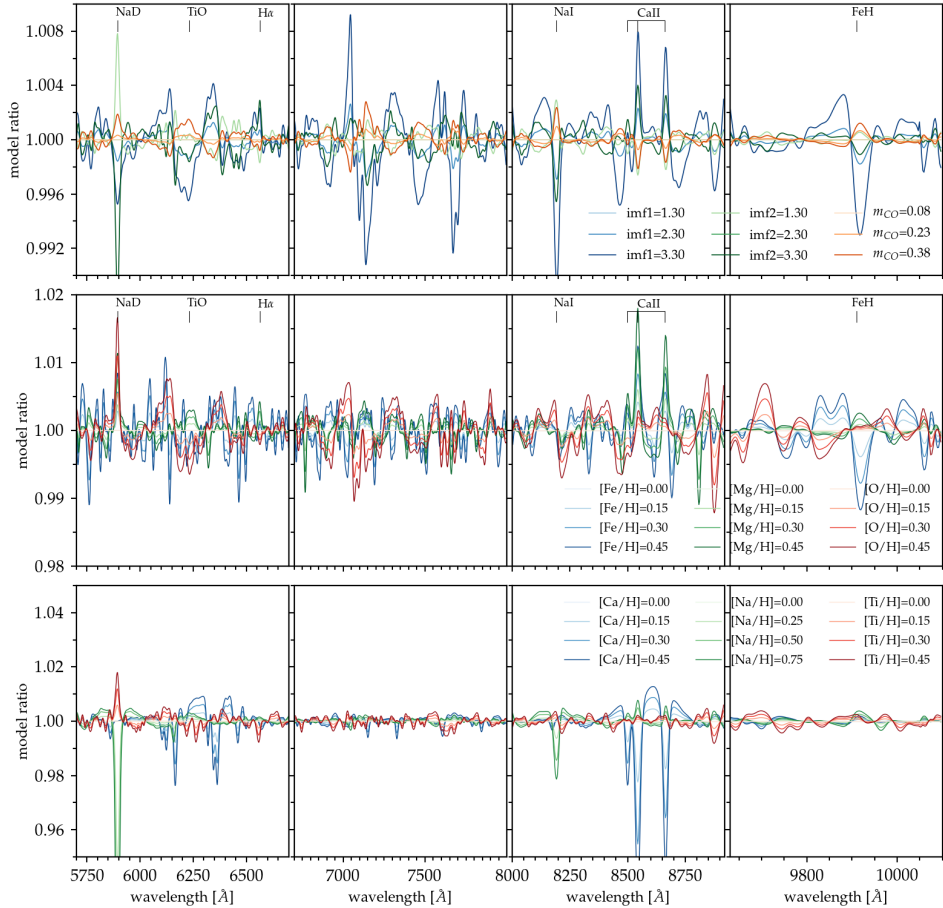


Figure 1.8: A demonstration of how the information from full-spectrum SPS models can be used to break degeneracies between stellar population properties. Several models are shown, with the parameters of each model varied relative to a reference model. The reference model has solar abundances and a Kroupa IMF. Each model is divided by this reference and normalized by a high-order polynomial. This model ratio is shown on the y -axis. Thus, this Figure shows SPS model sensitivity to changes in stellar population parameters as a function of wavelength. In the top panels, imf1 and imf2 represent variations in the low-mass slopes of the IMF and m_{CO} represents variations in the low-mass cutoff (see Equation 1.1). In the middle and bottom panels, regions that are sensitive to changes in elemental abundances (Fe, Mg, O, Ca, Na, and Ti) are shown. While increases in imf1 and imf2 appear to result in similar model ratio changes in regions near NaD, TiO, Na I, and CaT, the changes in the different IMF-sensitive features (Na I, CaT, and the Wing-Ford band) are slightly different. Thus, these three features can be used to break the degeneracy between IMF and elemental abundance variations. Figure reproduced from Gu et al. (2022).

1.3.1 Clues from the low-redshift Universe

Observations of massive, early-type galaxies (ETGs) in the low-redshift ($z \sim 0$) Universe have provided us with a great deal of insight into how massive galaxies have formed and evolved. These systems typically represent the latest stages of galaxy evolution. Their relatively high apparent magnitudes as a result of their close distances make them easier to observe compared to more distant galaxies. Additionally, they subtend a much larger solid angle on the sky, and thus we can study these objects at high spatial resolution.

By fitting SPS models to integrated photometric SEDs, we have learned that massive ETGs in the nearby Universe are composed of old stellar populations that quenched early and rapidly (e.g., Ferreras et al. 1999; Kaviraj et al. 2007; Pacifici et al. 2016; Carnall et al. 2018). Constraints based on integrated spectroscopy have reinforced these results. Studies using spectral indices have shown that the most massive quiescent galaxies at $z \sim 0$ also tend to be the oldest, most α -enhanced, and most metal-rich, consistent with the idea that massive galaxies formed their stars early and quickly (Trager et al. 2000; Gallazzi et al. 2005; Thomas et al. 2005, 2010).

More detailed measurements using full-spectrum fits have robustly confirmed that massive, compact galaxies are the oldest, most α -enhanced, and most metal-rich, reinforcing results from index studies, (e.g., Conroy et al. 2014; McDermid et al. 2015; Barone et al. 2018; Li et al. 2018). Crucially, full-spectrum fits to integrated spectra have revealed that nearby, massive ETGs have an excess of low-mass stars compared to the MW (i.e. a bottom-heavy IMF). This bottom-heaviness has been found to increase with increasing metallicity, $[\text{Mg}/\text{Fe}]$, and σ (e.g., Conroy & van Dokkum 2012b; Cappellari et al. 2012; Cappellari et al. 2013a; Spiniello et al. 2012; La Barbera et al. 2013; Ferreras et al. 2013; McDermid et al. 2014; Posacki et al. 2015; Cheng et al. 2023), and is in stark contrast with the idea of a universal IMF discussed in Section 1.2.3.

While integrated measurements have guided our understanding of the global properties of nearby galaxies, we can gain a detailed understanding of how these galaxies formed by studying the spatial distributions of stars and measuring stellar population gradients from high-spatial-resolution data. In particular, these gradients encode the build-up of stellar mass over cosmic time and allow us to differentiate quenching and assembly mechanisms. In the nearby Universe, spatially resolved photometry has revealed that low- z quiescent galaxies tend to have negative radial colour gradients, with redder centres and bluer outskirts (e.g., Peletier & Valentijn 1989; Franx & Illingworth 1990; Peletier et al. 1990b,a; Saglia et al. 2000; La Barbera et al. 2005; Suh et al. 2010; Tortora et al. 2010; Gonzalez-Perez et al. 2011; Liao & Cooper 2023). These colour gradients, which reflect the ages and/or metallicities of galaxies' constituent stellar populations (Sandage 1972), are consistent with the idea that galaxies were built up hierarchically (e.g., Larson 1974; Carlberg 1984; Kobayashi 2004; La Barbera et al. 2005; Tortora et al. 2010).

However, with resolved photometric data alone, it is difficult to determine the primary property that is driving these colour gradients, due to the aforementioned degeneracies between stellar properties. It is important to identify this driver,

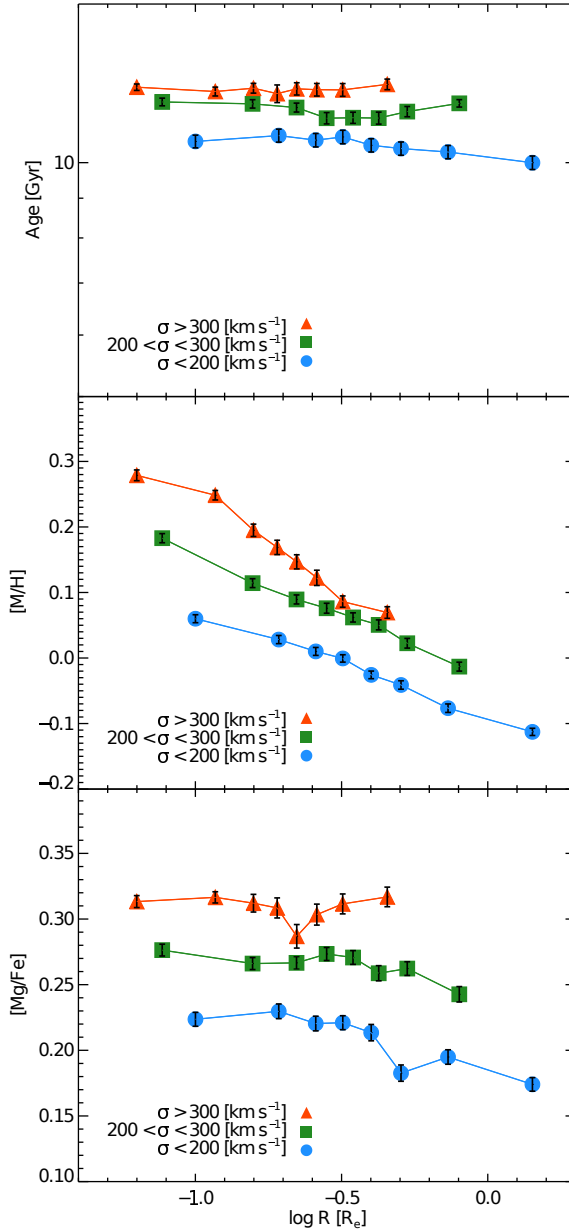


Figure 1.9: Average stellar population gradients in a sample of 45 massive elliptical galaxies at $z \sim 0$, from the CALIFA survey (Sánchez et al. 2012). Age gradients are shown in the top panel, metallicity gradients are shown in the middle, and [Mg/Fe] gradients are shown at the bottom, all as a function of radius. The spatially resolved stellar population parameters are constrained using spectral indices. The measurements are binned into 8 radius bins. In each panel, the average gradients are binned into velocity dispersion bins. Figure reproduced from Martín-Navarro et al. (2018).

as different properties are tied to different quenching and assembly mechanisms. For example, if age is primarily responsible for these colour gradients, this might indicate inside-out quenching, as this would result in older and therefore redder stellar populations being present in galaxy centres (e.g., Suess et al. 2019a). If metallicity is the primary culprit, this would suggest that compact, red galaxy cores accrete blue, low-mass, low-metallicity satellites, consistent with the two-phase formation model (e.g., Peletier & Valentijn 1989; Franx & Illingworth 1990; Peletier et al. 1990a; Davies et al. 1993; Vazdekis et al. 1997; Saglia et al. 2000; La Barbera et al. 2005; Suh et al. 2010; Tortora et al. 2010; Parikh et al. 2021; Liao & Cooper 2023).

As discussed in Section 1.2.4, these degeneracies can be more robustly disentangled using spectroscopic data. In particular, we have used high-spatial-resolution spectra to find that massive quiescent galaxies in the nearby Universe tend to have flat age and α -abundance gradients and negative metallicity gradients (e.g., Mehlert et al. 2003; Rose et al. 2005; Kuntschner et al. 2006; Sánchez-Blázquez et al. 2007; Koleva et al. 2011; Greene et al. 2013, 2015; Greene et al. 2019; Pastorello et al. 2014; González Delgado et al. 2015; Cook et al. 2016; Goddard et al. 2017; Martín-Navarro et al. 2018; San Roman et al. 2018; Ferreras et al. 2019; Oyarzún et al. 2019; Lacerna et al. 2020; Zheng et al. 2019; Santucci et al. 2020; Lee et al. 2023; Yoon et al. 2023; Parikh et al. 2024). Examples of these results are shown in Figure 1.9. These findings suggest that the negative colour gradients discussed above are driven by metallicity gradients. Additionally, there are indications that the IMF also varies with radius, from bottom-heavy in the cores of nearby, massive ETGs to MW-like in their outskirts (van Dokkum et al. 2017, though see Zieleniewski et al. 2017). Together, these results may support the two-phase model.

These archaeological studies of nearby galaxies have revealed key insights about massive quiescent galaxy formation. However, the abundances of stars in nearby galaxies can only give us information about the SFHs of all stars currently in the galaxy, including younger and metal-poor stars accreted during late-time mergers. These late-time mergers, as well as radial migration occurring over cosmic time, can also influence stellar population gradients. Additionally, inferring ages and SFHs becomes increasingly difficult for old stellar populations, as the variations in their spectral features become increasingly subtle. To truly understand the formation of massive quiescent galaxies, we must study their stellar population properties at earlier times.

1.3.2 Closing in on galaxy formation at higher redshifts

By studying massive quiescent galaxies at higher redshifts, we can examine more pristine stellar populations, closer to the epoch when they formed. This is especially important for galaxies at $z \sim 2 - 3$, which probe the epoch of peak galaxy growth and quenching in the two-phase model. (Oser et al. 2010; Whitaker et al. 2012; Madau & Dickinson 2014; Rodriguez-Gomez et al. 2016, see Section 1.1). Observing distant galaxies at the same level of detail as nearby galaxies is challenging, due in part to the fact that surface brightness dims cosmologically (as

$\sim (1+z)^4$). However, with cutting-edge technology, we have begun to apply the methods developed at low redshifts to high-redshift galaxies.

As we have moved to higher redshifts, integrated photometric data have been used to establish that massive quiescent galaxies were already in place by $z \sim 2-4$ (e.g., Smail et al. 2002; Daddi et al. 2003; Franx et al. 2003; Muzzin et al. 2013a; Spitler et al. 2014; Straatman et al. 2014, 2016). SED fits have furthermore indicated that distant massive quiescent galaxies tend to have old ages (Whitaker et al. 2010; Spitler et al. 2014; Straatman et al. 2014), large stellar masses (Muzzin et al. 2013a; Spitler et al. 2014; Straatman et al. 2014), and compact sizes (Trujillo et al. 2004, 2006; Daddi et al. 2005; van der Wel et al. 2008, 2014; van Dokkum et al. 2008). These findings align with the low- z results summarized in Section 1.3.1, implying that the progenitors of nearby galaxies formed their stars rapidly at early times and quenched quickly. However, similar to the low- z , photometric studies suffer from strong degeneracies, necessitating deep spectroscopy of these distant, massive systems.

Despite their importance in advancing our understanding of galaxy evolution, obtaining detailed spectroscopic data of massive quiescent galaxies at higher redshifts is extremely difficult. In particular, high-quality spectra of faint absorption lines are required. Observing these faint features is made all the more challenging by the fact that they are shifted into the near-infrared beyond $z \sim 1$, where instrument sensitivity typically decreases (especially from the ground, see, e.g., Vernet et al. 2011). Thus, incredibly deep spectra are needed to measure detailed stellar population parameters in distant massive quiescent galaxies.

With innovative technological advancements over the last ~ 15 years, spectroscopic surveys of galaxies in the early Universe have become possible. SPS model fits to deep spectra have revealed minimal evolution in elemental abundances since $z \sim 0.7$ (Choi et al. 2014; Beverage et al. 2021, 2023). However, beyond $z \sim 1$, distant galaxies have been found to have even more extreme properties than those in the nearby Universe, with low metallicities and α -enhanced stellar populations (e.g. Kriek et al. 2016; Kriek et al. 2019; Jafariyazani et al. 2020, 2025; Beverage et al. 2024; Carnall et al. 2022; Zhuang et al. 2023; Slob et al. 2024; Belli et al. 2025; Beverage et al. 2025). Additionally, these galaxies may have quenched rapidly, via neutral gas outflows (Park et al. 2024; Moretti et al. 2026). Moreover, these distant galaxies tend to be rotationally supported, compared to their dispersion-supported nearby counterparts (Slob et al. 2025). The precise mechanisms by which these early galaxies evolve into those we see at lower redshifts are still debated. For example, progenitor bias, late-time star formation, or major mergers may explain the evolution in abundances (Beverage et al. 2025), while minor mergers may be causing galaxies to lose rotation over cosmic time (Slob et al. 2025). Meanwhile, rapid quenching could be instigated by gas-rich major mergers, disc instabilities, or accretion of misaligned gas streams (Conselice et al. 2008; Dekel & Burkert 2014; Zolotov et al. 2015; Tacchella et al. 2016b; Park et al. 2024).

In reality, a combination of mechanisms are likely at play (e.g., Oser et al. 2012; Newman et al. 2012; Belli et al. 2014; Wellons et al. 2015; Rodriguez-Gomez et al. 2016; Beverage et al. 2025). Nonetheless, in order to understand the dominant processes influencing galaxy formation, we must examine spatially resolved stellar

populations. In contrast to the low-redshift Universe, however, spatially resolved measurements of distant massive quiescent galaxies have been, until recent years, *exclusively* limited to photometry. With these data, massive quiescent galaxies at $0.5 \lesssim z \lesssim 3$ have been found to have negative colour gradients, similar to results at low- z . Moreover, these colour gradients have been found to strengthen over cosmic time. These findings are consistent with two-phase formation, suggesting that the outskirts of compact, red cores are built up by the accretion of low-mass, bluer satellite galaxies (e.g., Wuyts et al. 2010; Guo et al. 2011; Gargiulo et al. 2012; Szomoru et al. 2013; Ciocca et al. 2017; Suess et al. 2019b, 2020, 2021, 2023; Miller et al. 2022, 2023, although see Pizzardo et al. 2026). However, alternative scenarios are possible, as these colour gradients could be signatures of inside-out quenching (e.g. Suess et al. 2019a). Spatially resolved spectroscopy is required to constrain the properties driving the observed colours in distant galaxies.

Initial spatially resolved spectroscopic studies have yet to converge on a coherent picture in the distant Universe, with existing work focusing on individual lensed and/or young galaxies, or relying on low-spectral-resolution data. For example, Jafariyazani et al. (2020) fit SPS models to spectroscopy of a single massive, lensed galaxy at $z \sim 2$ and found a flat age and [Mg/Fe] and negative metallicity gradient. While these are similar to findings in the nearby Universe, studies based on slitless grism spectroscopy have found diverse age and metallicity gradients in small samples of distant massive quiescent galaxies (Ditrani et al. 2022; Akhshik et al. 2023). Additionally, very young quiescent galaxies at $z \sim 3$ have been found to have varying age gradients, with indications of negative metallicity gradients (D’Eugenio et al. 2024; Pérez-González et al. 2025). Thus, our conclusions about galaxy formation from resolved spectroscopy at high redshifts have been limited by data quality, sample size, and sample selection.

The advent of *JWST* has furthermore allowed us to push to even higher redshifts, where the findings are more puzzling. In particular, as discussed in Section 1.1, *JWST* has facilitated the discovery of massive quiescent galaxies at much earlier times than predicted by current galaxy formation models (e.g., Carnall et al. 2023a,b, 2024; Valentino et al. 2023; Glazebrook et al. 2024; de Graaff et al. 2025; Ito et al. 2025, see also Baker et al. 2025; Lagos et al. 2025). The high stellar masses of $> 10^{11} M_{\odot}$ out to $z \sim 5$ can theoretically be produced, but only in the unlikely scenario that all available gas is transformed into stars at these early times (a star-formation efficiency of 100%, Carnall et al. 2024). These findings suggest that massive galaxies formed and quenched faster and earlier than can currently be explained.

The formation and evolution of massive quiescent galaxies thus remains an open question. In order to make progress in constraining assembly histories and quenching pathways, we require statistically significant measurements of spatially resolved stellar populations out to at least $z \sim 2$. Additionally, we must establish the robustness of inferred stellar masses by critically assessing the underlying SPS models and empirically constraining the IMF beyond the nearby Universe. These advancements are essential for achieving a holistic picture of massive galaxy formation over cosmic time.

1.4 This thesis

In this thesis, we provide some of the most detailed constraints to date on galaxy formation out to $z \sim 2$. Through flexible, full-spectrum modeling of deep integrated and spatially resolved spectra, we present novel measurements of stellar population gradients in distant galaxies, reveal discrepancies in SPS model predictions, and, for the first, time, robustly constrain the stellar IMF beyond the local Universe.

In **Chapter 2**, we present the first statistically significant measurements of stellar age and metal gradients in massive quiescent galaxies beyond the nearby Universe. We examine 456 massive ($10.3 \lesssim \log(M_*/M_\odot) \lesssim 11.8$) quiescent galaxies at $0.6 \lesssim z \lesssim 1.0$ from the LEGA-C survey and fit integrated and spatially resolved spectra using a flexible full-spectrum fitting code. We find that these galaxies tend to have flat age and [Mg/Fe] gradients and negative metallicity gradients. We forward model the [Fe/H] gradients in intrinsic space and find that they are significantly steeper than the observations.

We examine the observed gradients in three age bins and link high- z colour gradients to physical properties for the first time. Younger galaxies tend to have negative [Fe/H] and positive age gradients, suggesting that flat colour gradients in these galaxies are produced by the [Fe/H] and age effects cancelling each other. These results are consistent with a recent central starburst. Older galaxies have flat age and weaker, negative [Fe/H] gradients, suggesting that negative colour gradients in older quiescent galaxies are driven by metallicity gradients. The fact that age gradients flatten with time may be due to the fading of the central starburst, while the persistence of negative [Fe/H] gradients is consistent with galaxy outskirts being built up by minor mergers. Alternatively, the gradients could be inherited from these galaxies' star-forming progenitors.

In **Chapter 3**, we use detailed, spectroscopic ages and metallicities to challenge SPS model predictions from photometric data. We measure ages and metallicities by fitting a flexible, full-spectrum model to deep, continuum-normalized, absorption-line spectra of ~ 700 massive quiescent galaxies at $0.6 \lesssim z \lesssim 1.0$ from LEGA-C. We compare these measurements to independently measured, broadband UVJ colours.

We detect distinct sequences in age and metallicity in the rest-frame UVJ diagram, with age increasing along the quiescent sequence, and metallicity increasing nearly perpendicularly to the age trend. These findings indicate that age and metallicity behave differently in the UVJ diagram, suggesting that broadband UVJ colours may be used to predict relative ages and metallicities (though not precise values). However, while the age trend is consistent with predictions from SPS models, the metallicity trend conflicts with model expectations. Additionally, we investigate independent dynamical mass-to-light ratio trends in the UVJ diagram, finding that these trends also differ from model predictions. Thus, our results demonstrate that a reliance on model fits to photometry alone may lead to systematic biases in predicted galaxy properties. Our work illustrates that the underlying physical assumptions on which these models are built (e.g., non-solar abundance patterns, treatment of evolved stellar phases) must urgently be

re-assessed.

In **Chapter 4**, we measure spectroscopic stellar population gradients in a sample of massive quiescent galaxies at $z \sim 2$ for the first time. By modifying the method developed in Chapter 2, we fit deep, spatially resolved, *JWST*/NIRSpec-MSA spectra of 8 massive quiescent galaxies at $1.2 \lesssim z \lesssim 2.2$ from the *JWST*-SUSPENSE survey (Slob et al. 2024). We find that these galaxies tend to have negative age and flat [Fe/H] gradients, and may have positive [Mg/Fe] and [Mg/H] gradients. These results suggest that galaxy cores are older and possibly Mg-deficient, and thus formed faster and quenched more efficiently than their outskirts. However, our [Fe/H] and [Mg/Fe] gradients do not follow expectations from this picture.

These results are in contrast with the flat age and [Mg/Fe] and negative [Fe/H] gradients found at lower z , suggesting that gradients may evolve over cosmic time. This hypothesis is supported by our finding of a positive trend between age gradients and rotational support, and marginal trends between gradients and galaxy ages and velocity dispersions. Minor mergers can explain this evolution, by gradually destroying rotation, building up the metal poor outskirts, and diluting the initial age gradients. On the other hand, a progenitor bias scenario, where galaxies that quench at later times have negative [Fe/H] and flat age gradients, can also explain this evolution. In any case, these measurements have allowed us to gain a deeper understanding of galaxy assembly and quenching mechanisms, close to the epoch when these galaxies formed and quenched.

Finally, in **Chapter 5**, we present the first robust measurements of the stellar IMF in a sample of massive quiescent galaxies beyond the nearby Universe. We combine ultra-deep spectra of 9 massive quiescent galaxies at $z \sim 0.7$ from the *JWST*-IMPERNO program with deep, bluer spectra from LEGA-C. We fit the combined spectra with a flexible, full-spectrum model, allowing for variation in both detailed abundance patterns and the two low-mass slopes of the stellar IMF, enabling us to break strong degeneracies between IMF and abundance variations. We find that these distant galaxies have excess low-mass stars, indicating more bottom-heavy IMFs than the Milky Way. The oldest galaxy in our sample also has the most bottom-heavy IMF. As this galaxy may be a descendant of the extremely massive galaxies found at early times with *JWST*, our findings imply that these galaxies also had bottom-heavy IMFs at early times. The constraints that we find here suggest that their stellar masses are greater than originally reported, by a factor of ~ 4 . Our findings thus heighten the tension between observations of massive, early galaxies and current galaxy formation models.

1.5 Conclusions and outlook

With the aid of extremely deep ground- and space-based spectra and flexible stellar population models, our understanding of massive galaxy formation is rapidly evolving. This work has contributed substantially to clarifying our picture of massive galaxy formation from $z \sim 2$ to the present day. There are several pathways by which we can progress further. In particular, increased spatial resolution, larger

sample sizes, more sophisticated models, and cutting edge instruments will allow us to build upon these findings.

Increased spatial resolution with integral field spectroscopy can give us a more detailed view of assembly histories (e.g., Oyarzún et al. 2019). These insights can be gained by follow-up observations with state-of-the-art integral field unit (IFU) instruments including the *JWST*/NIRSpec-IFU and MUSE on the *VLT*. Looking forward, upcoming observations from near-infrared, high-spectral-resolution instruments such as HARMONI on the *Extremely Large Telescope (ELT)* and multi-object IFU instruments such as GIRMOS on *Gemini* will be crucial for obtaining statistically significant measurements of detailed stellar population gradients in the distant Universe.

While high-spatial-resolution measurements are the gold standard for understanding the details of galaxy formation, key insights can still be obtained from integrated spectra. *JWST* in particular has provided us with a wealth of these data. However most of them have low spectral resolution ($R \lesssim 1000$), while the spectroscopic insights discussed in this work can primarily only be obtained using at least medium resolution spectra ($R \gtrsim 1300$). By developing a sophisticated method to leverage low-resolution spectroscopy, we can increase our sample sizes to ~ 10 million galaxies over ~ 12 billion years of cosmic time.

Simultaneously, in order to take full advantage of these state-of-the-art data, we must develop our models to be able to account for the diverse and extreme galaxies that we are finding in the distant Universe. To improve the physical assumptions on which current stellar population synthesis models are built, we must further our understanding of evolved stellar stages (e.g., Maraston 2005; Muzzin et al. 2009; Kriek et al. 2010), implement α -enhancements and variable abundance patterns (e.g., Coelho et al. 2007; Lee et al. 2009; Percival et al. 2009; Vazdekis et al. 2015; Knowles et al. 2021, 2023; Park et al. 2025; Byrne et al. 2025), and extend our models to redder wavelengths (e.g., Eftekhari et al. 2021, 2022) and younger ages.

Finally, progress towards robustly characterizing the early Universe will require both wide-area and deep, high-spectral-resolution surveys of more distant galaxies. Wide-field space-based instruments, such as the *Nancy Grace Roman Space Telescope*, will facilitate efficient observations of statistically significant samples of high-redshift galaxies across a range of environments. Simultaneously, deep spectroscopy with thirty-metre-class telescopes, such as the *ELT*, will provide the high-spatial- and spectral-resolution measurements needed to constrain early assembly histories on an individual galaxy basis. These next-generation facilities will allow us to paint a more complete picture of galaxy formation over cosmic time.

2 | AGE AND METAL GRADIENTS IN MASSIVE QUIESCENT GALAXIES AT $0.6 \lesssim z \lesssim 1.0$: IMPLICATIONS FOR QUENCHING AND ASSEMBLY HISTORIES

ABSTRACT

We present spatially resolved, simple stellar population equivalent ages, stellar metallicities, and abundance ratios for 456 massive ($10.3 \lesssim \log(M_*/M_\odot) \lesssim 11.8$) quiescent galaxies at $0.6 \lesssim z \lesssim 1.0$ from the Large Early Galaxy Astrophysics Census, derived using full-spectrum models. Typically, we find flat age and [Mg/Fe] gradients, and negative [Fe/H] gradients, implying iron-rich cores. We also estimate intrinsic [Fe/H] gradients via forward modelling. We examine the observed gradients in three age bins. Younger quiescent galaxies typically have negative [Fe/H] gradients and positive age gradients, possibly indicating a recent central starburst. Additionally, this finding suggests that photometrically measured flat colour gradients in young quiescent galaxies are the result of the positive age and negative metallicity gradients cancelling each other. For older quiescent galaxies, the age gradients become flat and [Fe/H] gradients weaken, though remain negative. Thus, negative colour gradients at older ages are likely driven by metallicity gradients. The diminishing age gradient may result from the starburst fading. Furthermore, the persistence of the [Fe/H] gradients may suggest that the outskirts are simultaneously built up by mergers with lower metallicity satellites. On the other hand, the gradients could be inherited from the star-forming phase, in which case mergers may not be needed to explain our findings. This work illustrates the need for resolved spectroscopy, instead of just photometry, to measure stellar population gradients. Extending these measurements to higher redshift is imperative for understanding how stellar populations in quiescent galaxies are assembled over cosmic time.

Chloe M. Cheng, Mariska Kriek, Aliza G. Beverage, Arjen van der Wel, Rachel Bezanson, Francesco D'Eugenio, Marijn Franx, Pavel E. Mancera Piña, Angelos Nersesian, Martje Slob, Katherine A. Suess, Pieter G. van Dokkum, Po-Feng Wu, Anna Gallazzi, and Stefano Zibetti
2024, *Monthly Notices of the Royal Astronomical Society*, **532**, 3604

2.1 Introduction

In the low-redshift ($z \sim 0$) Universe, spatially resolved stellar populations have given us many insights into how nearby galaxies formed. In particular, we have learned about the assembly histories of massive quiescent galaxies by studying their stellar populations out to large radii. These radial stellar population gradients encode the build-up of stellar mass and allow us to differentiate quenching mechanisms and assembly scenarios (see Maiolino & Mannucci 2019 for a review).

It has been found that low- z quiescent galaxies have negative radial colour gradients, with redder centres and bluer outskirts (i.e. Peletier & Valentijn 1989; Franx & Illingworth 1990; Peletier et al. 1990a,b; Saglia et al. 2000; La Barbera et al. 2005; Suh et al. 2010; Tortora et al. 2010; Gonzalez-Perez et al. 2011; Parikh et al. 2021; Liao & Cooper 2023). In addition, their α -element abundances are generally consistent with having no gradients, they have flat or mildly positive gradients in age (i.e. galaxy centres are the same age or slightly younger than their outskirts, though note that, e.g., Zibetti et al. 2020 found U-shaped age profiles), and they have mildly negative gradients in $[\text{Fe}/\text{H}]$ (i.e. galaxy centres are more metal-rich than their outskirts, Mehlert et al. 2003; Rose et al. 2005; Kuntschner et al. 2006; Sánchez-Blázquez et al. 2007; Koleva et al. 2011; Greene et al. 2013, 2015; Greene et al. 2019; Pastorello et al. 2014; González Delgado et al. 2015; Cook et al. 2016; Goddard et al. 2017; Martín-Navarro et al. 2018; San Roman et al. 2018; Ferreras et al. 2019; Oyarzún et al. 2019; Lacerna et al. 2020; Zheng et al. 2019; Santucci et al. 2020; Lee et al. 2023; Yoon et al. 2023; Parikh et al. 2024, etc.). Thus, colour gradients are thought to be primarily driven by metallicity gradients, with perhaps some contribution from age gradients (Peletier & Valentijn 1989; Franx & Illingworth 1990; Peletier et al. 1990b; Davies et al. 1993; Vazdekis et al. 1997; Saglia et al. 2000; La Barbera et al. 2005; Suh et al. 2010; Tortora et al. 2010; Parikh et al. 2021; Liao & Cooper 2023; Miller et al. 2023). These findings are indicative of inside-out growth, either because younger and/or lower metallicity stellar populations at the outskirts are accreted by minor mergers (e.g. Naab et al. 2009), or because late-time star formation occurs in the disc and not in the central bulge (i.e. Abramson et al. 2014; Hill et al. 2017).

While these findings have revealed a great deal about the assembly histories of nearby galaxies, low- z studies do not paint the entire picture. For example, if there has been a significant amount of merging in a galaxy's past, then the stars that we see today were born in a range of different progenitor galaxies with potentially very different formation histories. Additionally, radial migration can occur with time, so present-day stars may not be in the same places as when they were born (Maiolino & Mannucci 2019). Therefore, to study conditions in the main progenitors of today's elliptical galaxies and to constrain past merging activity, we need to observe quiescent galaxies closer to the epochs of formation and quenching, before the majority of this merger activity occurred.

At higher z , measurements of resolved stellar populations of quiescent galaxies have been almost exclusively limited to photometry. For example, Wuyts et al. (2010), Guo et al. (2011), Szomoru et al. (2012), Chan et al. (2016), Ciocca et al. (2017), Liu et al. (2017), Mosleh et al. (2017), Suess et al. (2019a,b, 2020, 2021),

and Miller et al. (2023) showed that quiescent galaxies between $0.5 \lesssim z \lesssim 2.5$ have negative colour gradients on average, similar to low- z results. However, it is not yet clear what physical property is driving the observed colour gradients at these redshifts. In particular, gradients in age, metallicity, and dust are possible contenders, but it is difficult to disentangle these effects beyond $z \sim 0$ (Suess et al. 2019a). For example, Gargiulo et al. (2012) found that the negative colour gradients in several galaxies at $z \sim 1.5$ could be explained either by a pure radial age *or* metallicity variation. Thus, complementary studies using high signal-to-noise (S/N) spectroscopy are needed to understand detailed star formation and assembly histories.

Unfortunately, spectroscopic studies of quiescent galaxies beyond the low- z universe are exceedingly challenging. Therefore, they have largely been based on the integrated light of distant quiescent galaxies. In general, these studies have found that quiescent galaxies tend to have extreme elemental abundances, with metal-poor and α -enhanced stellar populations compared to local galaxies (i.e. Kriek et al. 2016; Kriek et al. 2019; Jafariyazani et al. 2020, 2025; Beverage et al. 2021, 2023, 2024; Carnall et al. 2022; Zhuang et al. 2023). However, these findings from integrated light are biased toward the central regions of quiescent galaxies. In order to truly understand the evolution that took place, we require spatially resolved spectroscopic measurements of a statistically significant sample of distant quiescent galaxies.

Thus far, this has not been possible due to observational limitations, as we require high-quality spectra of faint absorption lines, which are shifted to the near-infrared beyond $z = 1$. As a result, spectroscopic measurements of spatially resolved stellar populations have only been achieved for a handful of individual distant systems. For example, Jafariyazani et al. (2020) examined a massive, lensed galaxy at $z \sim 2$ and found that, similar to galaxies at low- z , this galaxy has no age or [Mg/Fe] gradients and a marginally negative [Fe/H] gradient. Akhshik et al. (2023) examined eight lensed galaxies using *Hubble Space Telescope (HST)* grism spectroscopy and found diverse age gradients. They also measured metallicity gradients, but due to the low spectral resolution their uncertainties were extremely large. Finally, Pérez-González et al. (2025) used *James Webb Space Telescope (JWST)*/Near Infrared Spectrograph (NIRSpec) integral field unit (IFU) data to measure a strongly negative total metallicity gradient in a massive quiescent galaxy at $z \sim 3.7$.

To bridge this gap, we therefore need to resolve stellar populations in larger samples of distant quiescent galaxies. To achieve this goal, we require ultra-deep spectra with at least moderate spatial resolution, so that multiple resolution elements can be analysed. We also require a large, statistically significant sample of quiescent galaxies. The Large Early Galaxy Astrophysics Census (LEGA-C, van der Wel et al. 2016, 2021; Straatman et al. 2018), a high S/N, high spectral resolution spectroscopic survey of 3600 galaxies at $0.6 \lesssim z \lesssim 1$, has finally made this possible. In this work, we present a comprehensive study of spatially resolved stellar population parameters in a statistically significant sample of quiescent galaxies at $0.6 \lesssim z \lesssim 1$, using data from the third data release of LEGA-C.

This paper is organized as follows: in Section 2.2, we describe the LEGA-C

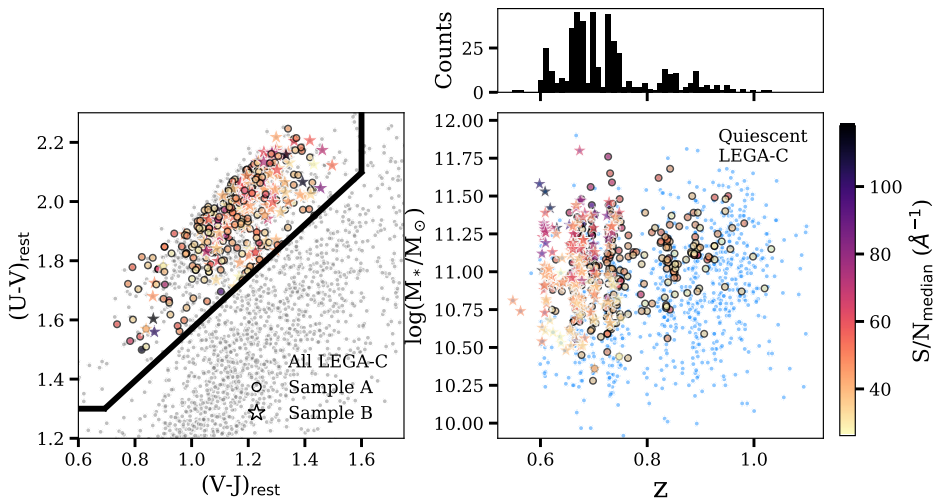


Figure 2.1: Quiescent galaxy sample studied in this work. In the left panel, we show the rest-frame UVJ diagram. The full LEGA-C sample is shown as grey points. In the right panel, we show stellar mass as a function of redshift. The full quiescent sample in LEGA-C is shown as blue points. In each panel, symbols are colour coded by the median S/N of the integrated spectrum. Circles represent galaxies for which we can only measure spatially resolved age and $[\text{Fe}/\text{H}]$ (Sample A). Stars represent galaxies for which we can measure spatially resolved age, $[\text{Fe}/\text{H}]$, and $[\text{Mg}/\text{Fe}]$ (Sample B). A histogram of the redshifts of the galaxies in Sample B is shown in the top panel.

data and our sample selection. In Section 2.3, we outline our spectral extraction, considerations for the blurred nature of the data due to observational seeing, and full-spectrum stellar population synthesis (SPS) modelling. We present our results and our predictions for gradients in intrinsic space in Section 2.4. We discuss the implications of our results for photometric measurements in the literature, galaxy assembly scenarios, and our assumptions in Section 2.5. Finally, our conclusions are presented in Section 2.6.

2.2 Data and sample

2.2.1 LEGA-C spectra

In this work we use spectroscopic data from the third data release of LEGA-C, a European Southern Observatory (ESO) Public Spectroscopic survey of 3600 galaxies between $0.6 \lesssim z \lesssim 1.0$. These galaxies are located in the Cosmic Evolution Survey (COSMOS) field (Scoville et al. 2007b) and were selected from the UltraVISTA (McCracken et al. 2012) K -band catalogue by Muzzin et al. (2013b). The data were collected over 128 nights using the VISIBLE MultiObject Spectrograph (VIMOS) on the ESO *Very Large Telescope* (VLT), providing deep (20-h integration), $R \sim 3500$ spectra with an average $\text{S/N} \sim 20 \text{ \AA}^{-1}$. The two-dimensional

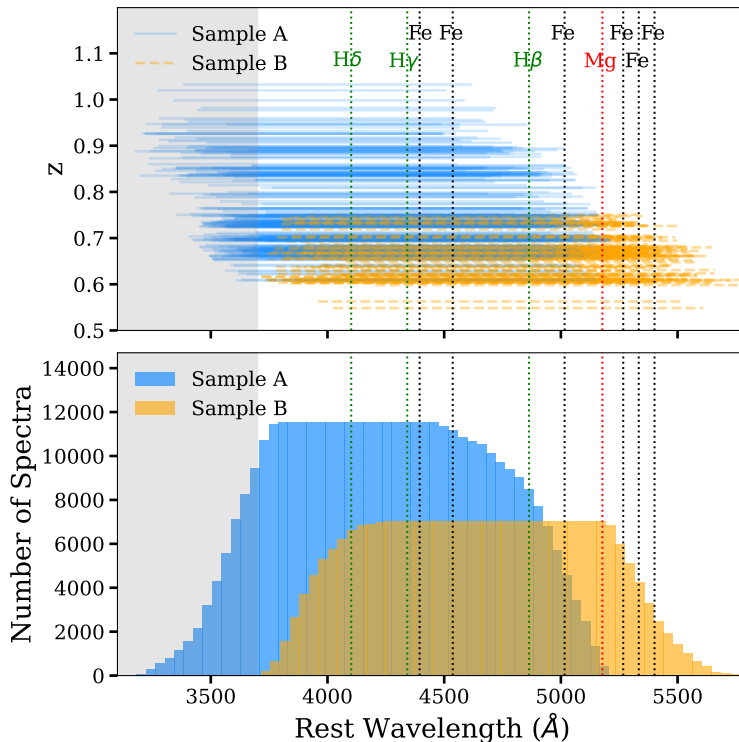


Figure 2.2: The wavelength coverage of our sample. In each panel, Sample A is indicated in blue and Sample B is indicated in orange. The shaded region indicates wavelengths that cannot be fit with ALF. We indicate key Balmer, Fe, and Mg features with vertical dotted lines. *Top:* the total wavelength ranges covered by all of the spectra in our sample. Rest wavelength is shown on the x -axis and redshift is shown on the y -axis. Galaxies in Sample A (for which we can measure age and metallicity gradients) are shown as solid lines and galaxies in Sample B (for which we can also measure Mg gradients) are shown as dashed lines. *Bottom:* histograms of the wavelength coverage of Samples A and B.

(2D) and reduced one-dimensional (1D) spectra are available on the ESO Science Archive Facility^{1,2}. See van der Wel et al. (2016, 2021), and Straatman et al. (2018) for details. In this study, we use the reduced 2D spectra.

2.2.2 Sample selection

We select quiescent galaxies from the LEGA-C catalogue via their rest-frame $U - V$ and $V - J$ colours (Wuyts et al. 2007). To make this selection, we first determine

¹http://archive.eso.org/eso/eso_archive_main.html .

²The reduced 1D spectra and catalogue have been released by ESO (<http://archive.eso.org/cms/eso-archive-news/Third-and-final-release-of-the-Large-Early-Galaxy-Census-LEGA-C-Spectroscopic-Public-Survey-published.html>), and are also available here: <https://users.ugent.be/~avdrwel/research.html#legac> .

the UltraVISTA colours (Muzzin et al. 2013b) for the LEGA-C galaxies with EAZY (Brammer et al. 2008) by fixing the spectroscopic z to those measured from the LEGA-C spectra (van der Wel et al. 2021). We use these colours to characterize galaxies as star forming or quiescent, employing the classification from Muzzin et al. (2013b), and select the quiescent galaxies. We show the rest-frame UVJ diagram in the left panel of Fig. 2.1, where galaxies in our sample are colour coded by their median rest-frame S/N per \AA . We remove galaxies with spectroscopic $z < 0.5$ from our sample. We also discard galaxies in LEGA-C’s mask 2, as we find that the noise spectrum is significantly underestimated. Furthermore, we require that each spectrum that we examine has a median rest-frame S/N $\gtrsim 20 \text{\AA}^{-1}$ and a maximum wavelength of at least 4450\AA for us to sufficiently recover ages and metallicities (these limits are determined using simulated galaxies, see Section 2.3.3 and Appendix 2.A). We perform a final sample cleaning during the fitting stage and discard individual galaxies whose age, $[Z/H]$, $[Fe/H]$, and $[Mg/H]$ posterior distributions are not well sampled. We also examine all ages and ensure that they are ≥ 1 Gyr as our fitting code is only valid for stellar populations older than 1 Gyr. See Section 2.3.3 for details about our fitting procedure.

We split our sample into two subsamples. For each galaxy, we make use of the full available wavelength range. For those galaxies with a maximum wavelength $\geq 5207 \text{\AA}$ (i.e. galaxies which include at least the bluest Mg feature, Mg b), we are able to measure age, metallicity, and Mg gradients (Sample B). However, for those galaxies whose spectra have wavelengths anywhere between $3700 \text{\AA} \leq \lambda < 5207 \text{\AA}$ (i.e. with no Mg features), we still expect to recover age and metallicity gradients (Sample A). In particular, spectra in Sample A have a wavelength range of at least $3700.0 \text{\AA} \lesssim \lambda \lesssim 4468.9 \text{\AA}$ which allows us to capture multiple Balmer and Fe lines. We show the wavelength ranges encompassed by each galaxy in our sample in the top panel of Fig. 2.2. We indicate galaxies in Sample A with solid lines and galaxies in Sample B with dashed lines. In the bottom panel, we show histograms of the total wavelength coverage of Sample A (blue) and Sample B (orange). In Fig. 2.1, we show the galaxies in Sample A as circles and the galaxies in Sample B as stars. In total, Sample A contains 285 galaxies while Sample B contains 171 galaxies (in other words, we can measure age and metallicity gradients for a total of 456 galaxies and Mg gradients for 171 galaxies).

Our final sample of 456 quiescent galaxies is shown in the right panel of Fig. 2.1, where spectroscopic redshift (from LEGA-C) is shown on the x -axis and stellar mass (computed by the LEGA-C team using MAGPHYS; da Cunha et al. 2008) is shown on the y -axis. The distribution of redshifts is also shown in the top panel. The rest of the quiescent LEGA-C sample is shown in blue. Our selection covers a wide range of redshifts ($0.5 \leq z \leq 1.1$) and stellar masses ($10.28 \leq \log(M_*/M_\odot) \leq 11.80$). We note that we include ~ 31 per cent of the total quiescent sample in LEGA-C. In the left panel of Fig. 2.1, we show that we sample the full distribution of quiescent galaxies in UVJ -space and our sample is thus representative of the UVJ quiescent sequence.

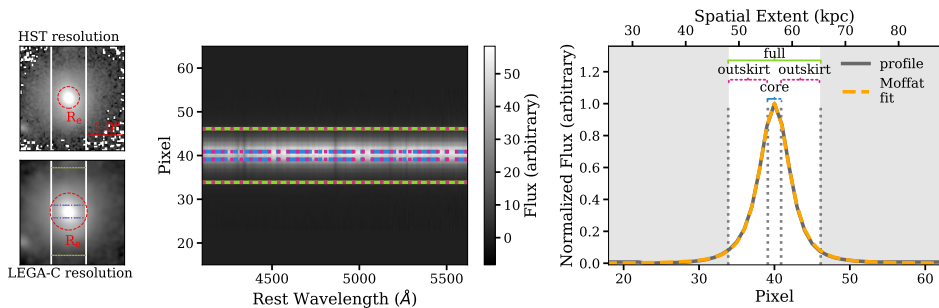


Figure 2.3: *Left:* COSMOS *HST* ACS images (Koekemoer et al. 2007; Scoville et al. 2007b) of an example galaxy. The top panel is the image at *HST*/F814W resolution, with the VIMOS slit overlaid. The bottom panel is the same image but convolved to the seeing of LEGA-C, which we compute in Section 2.3.2. We also outline the pixels that are included in each extraction, where solid lines show the full, integrated extraction region, dotted lines show the outskirts region, and dashed-dotted lines show the core region. The R_e is shown at the *HST* and LEGA-C resolutions in each panel and is indicated by the dashed circles. The remaining two panels show a demonstration of our optimal extraction procedure. In each panel, we outline the extraction regions similarly. *Middle:* an example 2D spectrum for the same galaxy as in the *HST* images. *Right:* the collapsed normalized flux profile of this 2D spectrum (solid curve), fit with a Moffat profile (dashed curve). The extraction regions are indicated with labels and dotted lines. Grey shaded regions show the pixels not included in our extracted spectra. On the top x -axis, we use the pixel scale of VIMOS (0.205 arcsec) to transform the pixels to physical units.

2.3 Methods

To determine spatially resolved stellar population parameters, we extract spatially resolved spectra, determine observed sizes in convolved space, and fit the spectra with full-spectrum SPS models. We describe our spectral extraction in Section 2.3.1, our size determination in Section 2.3.2, and our spectral fitting in Section 2.3.3.

2.3.1 Optimal extraction

To measure spatially resolved stellar population parameters in the LEGA-C galaxies, for each galaxy we obtain 1D spectra for two spatial bins and a 1D integrated spectrum using a custom optimal extraction routine applied to the 2D spectra.

For our optimal extraction routine, we first obtain the flux profile of each galaxy by collapsing the 2D spectrum over the wavelength axis. We then fit a Moffat (1969) profile to the flux profile, which is given by

$$M(\alpha, \gamma) = A \left(1 + \frac{(x - x_0)^2}{\gamma^2} \right)^{-\alpha}. \quad (2.1)$$

Here, we fit for A (a normalization factor), x_0 (the centre of the profile), and γ and α (the Moffat parameters). The full width at half-maximum (FWHM) of the

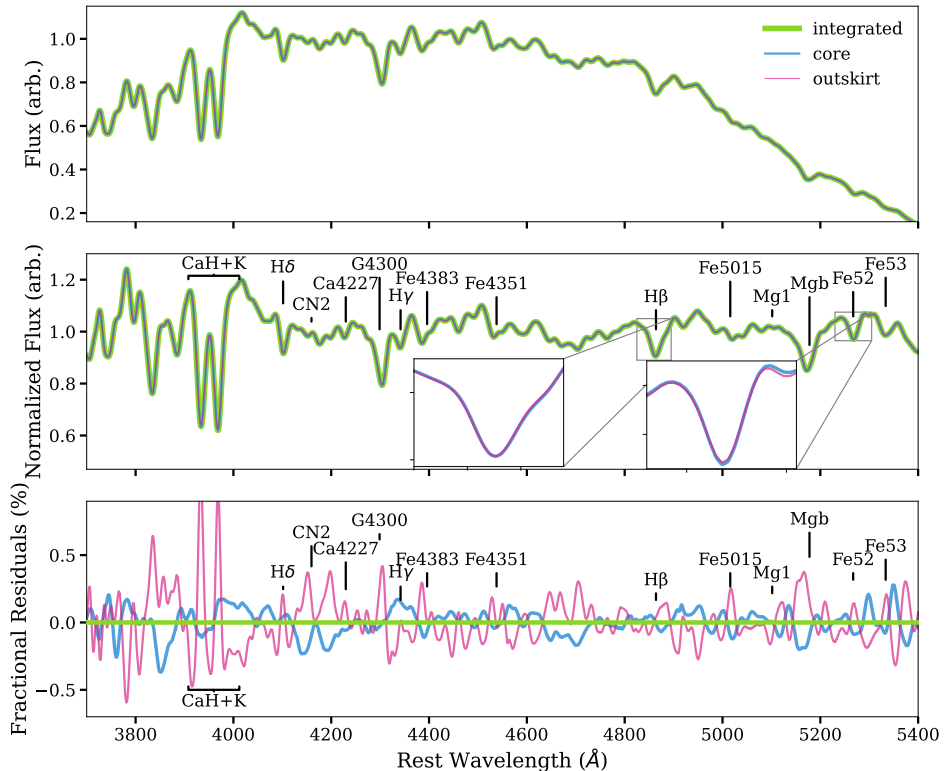


Figure 2.4: *Top:* extracted and stacked integrated, core, and outskirts spectra from our entire sample. We smooth each spectrum to a common velocity dispersion of $\sigma = 450$ km s $^{-1}$, then mean stack the spectra from each extraction. *Middle:* the same spectra but continuum normalized. We label key spectral absorption features. In the inset panel, we zoom-in on the core and outskirts spectra near the H β feature, which is sensitive to age, and near the Fe52 feature, which is sensitive to [Fe/H]. The two spectra are distinct in Fe52, but not in H β . *Bottom:* a representation of the fractional residuals of each stacked spectrum, where we divide each spectrum by the integrated spectrum.

Moffat profile is given by $FWHM = 2\gamma\sqrt{2^{1/\alpha} - 1}$. Moffat profiles were also used to extract the public LEGA-C spectra (van der Wel et al. 2021), and have been found to fit the wings of spectral profiles better than Gaussian fits (Moffat 1969; Trujillo et al. 2001).

We extract three 1D spectra for each galaxy: an integrated spectrum and two spatially binned spectra. To extract an integrated spectrum, we weight all (sub-)rows in the 2D spectrum by the Moffat profile and sum the weighted (sub-)rows with significant flux [i.e. (sub-)rows between the 3rd and 97th percentiles of the Moffat profile]. For the spatially binned spectra, we perform a similar weighted sum of the (sub-)rows in the middle 30 per cent of the Moffat profile only. These comprise the ‘core’ of the galaxy. To extract the ‘outskirts’, we perform a weighted sum of the remaining (sub-)rows (excluding the 3rd and 97th percentiles). A

demonstration of this procedure is shown in Fig. 2.3. We define these extraction regions to compromise between a large sample size and a minimum difference in S/N between each core and outskirt spectrum. In particular, the S/N of the core and outskirt spectra for each galaxy are often comparable. However, we have tested different sizes of extraction regions and find that the exact combination of rows that are included does not affect our conclusions. See Section 2.5.3 for more details.

For visualization purposes, we stack all of our rest-frame integrated, core, and outskirt spectra, respectively. Prior to stacking, we regrid the spectra to the same wavelength array using SPECTRES (Carnall 2017), continuum-normalize them, and smooth them all to a common velocity dispersion of $\sigma = 450 \text{ km s}^{-1}$ (the maximum σ of our sample). In Fig. 2.4, we show these spectra as well as their fractional residuals (where we divide each stacked spectrum by the stacked integrated spectrum). Visually, it can be seen that the three spectra are distinct. In particular, in the middle panel, we zoom in on two strong features: $\text{H}\beta$, which is sensitive to age, and an Fe feature near 5200 \AA , which is sensitive to $[\text{Fe}/\text{H}]$. The $\text{H}\beta$ feature is perhaps slightly weaker for the core spectrum. However, the Fe feature is distinctly stronger for the core spectrum compared to the outskirt spectrum. We discuss this further in Section 2.4.

2.3.2 Convolved R_e

A major consideration of our spectral extraction method is that our measurements will be significantly affected by the seeing of the telescope. In other words, our data and therefore our measurements are blurred by instrumental and atmospheric effects. It would therefore be a misrepresentation to display our results in units of the R_e measured in the *HST*/F814W images (Scoville et al. 2007b) via GALFIT (Peng et al. 2010) by the LEGA-C team. As such, we derive the *convolved* R_e for all galaxies using a method similar to Price et al. (2016), where the R_e is similarly blurred by the seeing of the observations.

We first determine the seeing by creating an idealized mock galaxy image in GALFIT (Peng et al. 2010) for each galaxy in our sample, using the structural parameters reported in LEGA-C. We generate the images with the same spatial resolution as LEGA-C, where we create square pixels of 0.205 arcsec by 0.205 arcsec , and with the same number of pixels as the LEGA-C spectra. We convolve this model image with a grid of different model point spread functions (PSFs). Each model PSF is a 2D Moffat kernel with different α and γ parameters, with α and γ ranging between $0.1 - 7.0$ (see equation 2.1). We then mask each model image with a rectangular aperture the same size as the VIMOS slit (1 arcsec wide) using the PHOTUTILS package (Bradley et al. 2024). We sum over the slit width to obtain a model flux profile (m). We centre m to the flux profile from the spectrum (f). We normalize m to f by multiplying m by a scaling factor defined as $\frac{\sum(m \times f)}{\sum m^2}$ (where we sum over all spatial elements) and compare it to f by calculating the reduced χ^2 (χ_{red}^2) value over the region that we consider in the optimal extraction (i.e. excluding the low-flux edges of the profiles). The seeing value which produces the minimum χ_{red}^2 corresponds to our best-fit seeing. Our median best-fit seeing

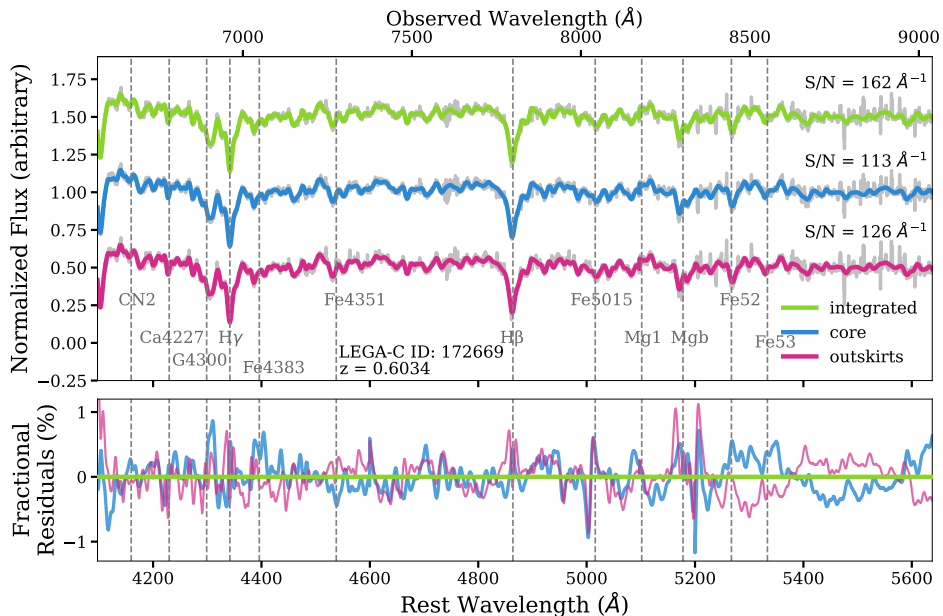


Figure 2.5: *Top:* example best-fitting ALF models to a $z \sim 0.6$ galaxy from our sample with high S/N. The spectra are shown in grey. The integrated fit is shown on top, the fit to the core spectrum is shown in the middle, and the fit to the outskirt spectrum is shown at the bottom. The median S/N of each bin is quoted above the respective fit. Key spectral features are labelled in grey. *Bottom:* similar to Fig. 2.4. The fractional residuals of each fit, where we divide each fit by the integrated fit.

is ~ 0.70 arcsec and, as expected, we find a similar seeing within each LEGA-C mask. We note that a similar method to determine the seeing in LEGA-C was used in van Houdt et al. (2021), although there *HST* Advanced Camera for Surveys (ACS)/F814W images were convolved with the best-fit seeing instead of GALFIT models. It is encouraging that our typical seeing is similar to that measured in van Houdt et al. (2021) (0.75 arcsec).

To calculate the convolved R_e for the LEGA-C spectra, we apply this best-fit seeing value to a similar model image of each galaxy, but with an increased spatial resolution (0.1 arcsec pixel^{-1}) and a larger number of pixels (50 times larger than the original model image). We place a series of circular apertures of increasing size on the enlarged image and perform aperture sums within each circle using PHOTUTILS. We then determine the radius which encloses 50 per cent of the light for this convolved model. For each galaxy, we use this value to determine the spatial extent of our core and outskirt spectra in units of this convolved R_e .

2.3.3 Full spectrum fitting

We fit each spectrum with the ABSORPTION LINE FITTER (ALF³), a full spectrum SPS model (Conroy & van Dokkum 2012a; Conroy et al. 2018), to derive our stellar population parameters. The ALF models are built on empirical simple stellar populations, created using the Mesa Isochrones and Stellar Tracks (MIST, Choi et al. 2016) and the Spectral Polynomial Interpolator (SPI, Villaume et al. 2017b).⁴ We use the Medium Resolution INT Library Of Empirical Spectra (MILES, Sánchez-Blázquez et al. 2006), the Extended Infrared Telescope Facility stellar library (E-IRTF, Villaume et al. 2017b), and a large sample of M-dwarf spectra (Mann et al. 2015) with SPI. In this way, ALF develops stellar spectra as a function of T_{eff} , surface gravity, and metallicity from a data-driven model.

The empirical parameter space spans $-2.0 \lesssim [\text{Fe}/\text{H}] \lesssim 0.5$ and $3.5 \lesssim \log(T_{\text{eff}}/\text{K}) \lesssim 3.9$ and is set by the combined E-IRTF and Mann et al. (2015) samples. We additionally use a theoretical stellar library (C3K, see Conroy & van Dokkum 2012a) to ensure the quality of interpolation at the boundaries of the empirical parameter space. The ALF models allow for variable abundance patterns by differentially including theoretical element response functions. To derive the $[\text{Mg}/\text{Fe}]$ versus $[\text{Fe}/\text{H}]$ relation, we use the Mg abundances for the MILES stellar library stars from Milone et al. (2011).

ALF first continuum normalizes the target spectrum by multiplying it by a high-order polynomial. It then samples the posteriors of 46 stellar parameters using a FORTRAN implementation of the Markov Chain Monte Carlo algorithm (MCMC) `emcee` (Foreman-Mackey et al. 2013), allowing for arbitrary variation in stellar age and detailed elemental abundance patterns. It fits for systematic parameters to characterize observed errors. Note that ALF can fit spectra between 3700 – 24000 Å and can be used for stellar populations that are older than 1 Gyr. For details, see Conroy & van Dokkum (2012a) and Conroy et al. (2018).

We fit each spectrum using 1024 walkers, 20000 burn-in steps, and a 1000 step production run. We examine 500 MCMC chains per fit in our analysis. In our implementation of ALF, we assume a Kroupa (2001) initial mass function and fit a single age. We initialize the age of each galaxy by drawing a random value from a uniform distribution centred at 3 Gyr to avoid the ALF walkers getting trapped at an unrealistically high age. We also set the upper limit of the age prior to be the age of the Universe at each galaxy’s redshift, plus 2 Gyr to allow for uncertainties. We do not fit the hot star component.

We first fit the integrated spectra, allowing for variation in all stellar parameters in ALF. For the spatially binned spectra (core and outskirts), we do not expect to be able to constrain all of the elemental abundances that ALF can fit at this S/N (see Appendix 2.A). Thus, to accurately constrain age, Fe, and Mg, we use the values of all other abundances from the integrated fits and fix these values in the fits to the spatially binned spectra. We inspect each of the three fits for each individual galaxy to ensure that the posterior distributions for the parameters of interest (age, Fe, and sometimes Mg) are well-sampled. For those galaxies

³<https://github.com/cconroy20/alf> .

⁴https://github.com/AlexaVillaume/SPI_Utils .

where the posteriors run up against the priors, we discard these fits as their ages, metallicities, or Mg abundances are not reliable. This is the case for < 10 per cent of our sample. This produces our final sample size of 456 galaxies.

There are several galaxies in LEGA-C that have been observed twice (111 galaxies in our sample). In our analysis, we treat these observations as different objects and fit them individually. However, we also use these objects to ensure the robustness of our fitting method, by comparing our results for the duplicate galaxies as well as to the results when we stack and fit the duplicate galaxies. In general, the results are consistent. For more details, see Appendix 2.B.

In the top panel of Fig. 2.5, we show example fits for one of our high S/N galaxies at $z \sim 0.6$. The best fit to the integrated spectrum is shown in green, the best fit to the core spectrum is shown in blue, and the best fit to the outskirts spectrum is shown in magenta. In the bottom panel, we divide each fit by the integrated fit, similar to the bottom panel of Fig. 2.4. The core and outskirts fits are distinct from each other.

We compare the results of our integrated fits to those of the overlapping galaxies fit in Beverage et al. (2023) (not shown) and find that our results are entirely consistent, despite the differences in the sample selection (see Beverage et al. 2023) and spectral extraction (see van der Wel et al. 2021). We will discuss the results of our integrated fits in more detail in a future study.

2.4 Results

In this section, we present the results of our ALF fits (Section 2.4.1). We also develop a model to explore what our observed gradients are expected to look like in intrinsic space (Section 2.4.2).

2.4.1 Observed gradients

Our main results are shown in Figs 2.6 and 2.7. In Fig. 2.6, we show histograms of the slope of each measured gradient (in convolved space). The median is indicated in blue with shaded regions indicating the bootstrapped uncertainties. The dashed black line indicates where the median would lie if there were no gradient. On average, we find that massive quiescent galaxies at $0.6 \lesssim z \lesssim 1.0$ have age gradients consistent with being flat (median $\Delta \log(\text{Age}(\text{Gyr})) / \Delta \log(R_{e,\text{convolved}}) = 0.007^{+0.002}_{-0.004}$), mildly negative metallicity gradients (median $\Delta[\text{Fe}/\text{H}] / \Delta \log(R_{e,\text{convolved}}) = -0.048^{+0.004}_{-0.009}$), and $[\text{Mg}/\text{Fe}]$ gradients consistent with being flat (median $\Delta[\text{Mg}/\text{Fe}] / \Delta \log(R_{e,\text{convolved}}) = -0.008 \pm 0.007$). Our results are qualitatively consistent with Fig. 2.4, where we can see that the $\text{H}\beta$ feature (sensitive to age) is visually not significantly different between the core and the outskirts. However, the Fe feature near 5200 Å is deeper in the core than in the outskirts.

In Fig. 2.7, we show the values of age, $[\text{Fe}/\text{H}]$, and $[\text{Mg}/\text{Fe}]$ as a function of radius in units of R_e (in convolved space, see Section 2.3.2). The values of each parameter have been split into equally sized stellar velocity dispersion (σ) bins for

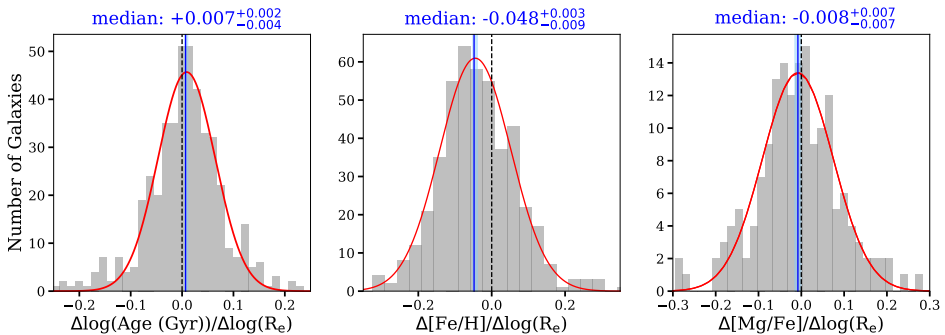


Figure 2.6: Histograms of our resolved stellar population gradients, normalized by R_e in convolved space. Age gradients are shown on the left, metallicity gradients are shown in the middle, and $[\text{Mg}/\text{Fe}]$ gradients are shown on the right. In each panel, we fit the histogram of the gradients from our entire sample (grey) with Gaussians (red curves) to emphasize the shape of the distribution. The median gradient is shown as a solid blue line, with the blue shaded region indicating the uncertainties on the median derived by bootstrap resampling. We also show where a flat median gradient would lie (black dashed line). Above each panel, we quote the value of the median gradient along with bootstrapped uncertainties.

the age and $[\text{Fe}/\text{H}]$ gradients (152 galaxies in each bin,⁵ and where σ is measured by the LEGA-C collaboration). These map approximately onto bins of stellar mass. The circles indicate the median in each bin with uncertainties derived by bootstrap resampling. These are plotted at the median R_e in convolved space in each bin. The data points are joined by lines to guide the eye. From this figure, we see that our results of flat age gradients, negative $[\text{Fe}/\text{H}]$ gradients, and flat $[\text{Mg}/\text{Fe}]$ gradients hold across a wide range of σ . We additionally recover the trends found in Beverage et al. (2023) for quiescent galaxies in LEGA-C. Specifically, they showed that age and $[\text{Fe}/\text{H}]$ increase with σ , while $[\text{Mg}/\text{Fe}]$ does not correlate with σ .

In Fig. 2.8, we also examine our gradients as a function of integrated galaxy stellar age, by splitting the sample into three equally sized age bins (152 galaxies in each bin⁶). We plot the running median in each age bin and the corresponding best linear fit to the running medians. We also show our individual measurements as grey contours. Splitting our sample into different age bins reveals intriguing trends. In particular, we find that the age gradients are slightly positive for the youngest galaxies and flatten with stellar population age. The metallicity gradients become weaker with increasing age, but are still negative at the oldest ages. We find no trend between the $[\text{Mg}/\text{Fe}]$ gradients and stellar population age (i.e. they are flat for all ages). We comment further on these trends in Section 2.5.

Our findings are consistent with the predictions of Wu et al. (2020), who used

⁵For the $[\text{Mg}/\text{Fe}]$ gradients there are 72, 57, and 42 galaxies in each increasing σ bin, respectively.

⁶For the $[\text{Mg}/\text{Fe}]$ gradients there are 38, 57, and 76 galaxies in each increasing age bin, respectively.

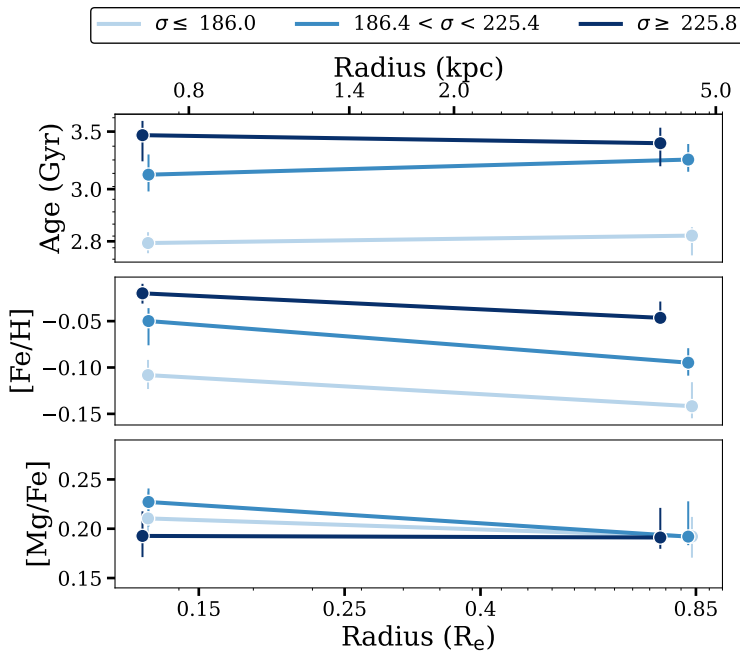


Figure 2.7: Stellar population parameters plotted as a function of radius (R_e and kpc) in convolved space. Ages are shown in the top panel, [Fe/H] is shown in the middle, and [Mg/Fe] is shown on the bottom. We split our sample into three velocity dispersion (σ) bins, with σ in units of km s^{-1} . These map approximately onto bins of stellar mass. We plot the median stellar population parameters with uncertainties derived by bootstrap resampling in each σ bin.

LEGA-C to propose that positive age gradients were required in younger quiescent galaxies to explain the increase in galaxy size with age. They are also consistent with the measurements of age gradients in young quiescent galaxies of D’Eugenio et al. (2020) at similar z and Pracy et al. (2013) and Wu (2021) at low z (and see also French 2021). In particular, D’Eugenio et al. (2020) measured Lick index gradients in the LEGA-C spectra for 17 post-starburst⁷ (PSB) galaxies as well as a control sample of typical quiescent galaxies. They found positive age gradients in the younger PSB galaxies. On the other hand, in their control sample of older quiescent galaxies, they found that the central stellar populations were older and more metal-rich than the outer stellar populations. However, their conclusions about the control sample were fundamentally limited by the use of only a few spectral indices; in contrast, our full-spectrum fits enable us to break the age–metallicity degeneracy, even for older quiescent galaxies (See Section 2.5.3 and Appendix 2.A). While we find that the older galaxies in our sample have flat age gradients, these results are otherwise qualitatively similar to what we find in this

⁷We emphasize that our galaxies are *not* necessarily PSB galaxies but are simply *younger* quiescent galaxies. Classifying galaxies in our sample as PSB is beyond the scope of this work.

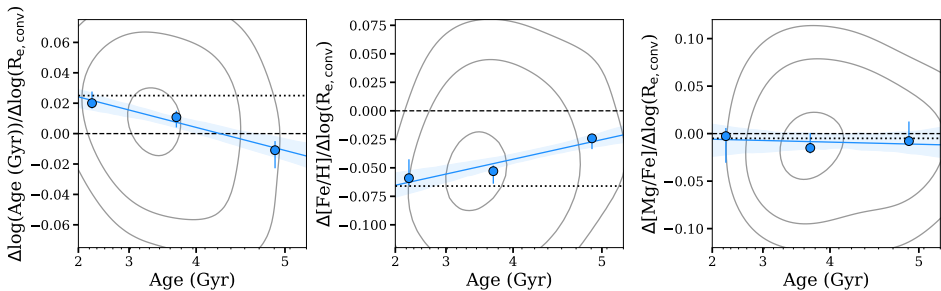


Figure 2.8: Spatially resolved stellar population gradients as a function of integrated age. In each panel, the 50th, 68th, and 95th percentiles of the whole sample are represented by the contours. The dashed line indicates where a flat gradient would lie. The points show the median gradients in three evenly sized bins of integrated age, with uncertainties derived by bootstrap resampling. The solid lines show linear fits to these running medians, and the shaded regions denote the 1σ uncertainties on these fits. The dotted line indicates where no trend lies.

work.

On the other hand, Setton et al. (2020) concluded that young quiescent galaxies at these redshifts have flat age gradients, based on their finding of a flat gradient of the $H\delta$ spectral index. None the less, our results do not necessarily disagree with those of Setton et al. (2020). We also examine the equivalent width of the $H\delta$ index in our spatially resolved spectra and find that the difference between $H\delta$ in the core and the outskirts is not significant. However, when we make use of the full spectral wavelength range instead of one spectral feature, we are able to recover a positive age gradient in our youngest age bin. This demonstrates that full spectral coverage is required to unveil gradients. Moreover, this may not be surprising as, in addition to being sensitive to age, $H\delta$ is also sensitive to star formation history (Worthey & Ottaviani 1997). Thus, it does not perfectly trace the age gradient.

We also examine our measured gradients in relation to several other physical parameters. We find no significant trends between our measured gradients and intrinsic R_e , spectroscopic redshift, stellar velocity dispersion, stellar mass, or Sérsic index. The fact that we do not find any significant trends with total stellar velocity dispersion or stellar mass contrasts with the results of Spolaor et al. (2010), González Delgado et al. (2015), Ferreras et al. (2019), Santucci et al. (2020), and Yoon et al. (2023) for low- z galaxies, where varying dependence between mass and different stellar population gradients has been found. For example, Spolaor et al. (2010) found a positive trend between metallicity gradient and mass but no strong dependence for age and $[Mg/Fe]$. Alternatively, Ferreras et al. (2019) found a weak dependence of metallicity gradient on velocity dispersion and a strong negative dependence of $[Mg/Fe]$ gradient on velocity dispersion. On the other hand, Sánchez-Blázquez et al. (2007), Pastorello et al. (2014), González Delgado et al. (2015), and Greene et al. (2015) found that metallicity gradients do not depend on mass at low z . Additionally, studies making use of cosmological simulations find, in particular, no dependence of metallicity gradient on M_* (Kobayashi 2004;

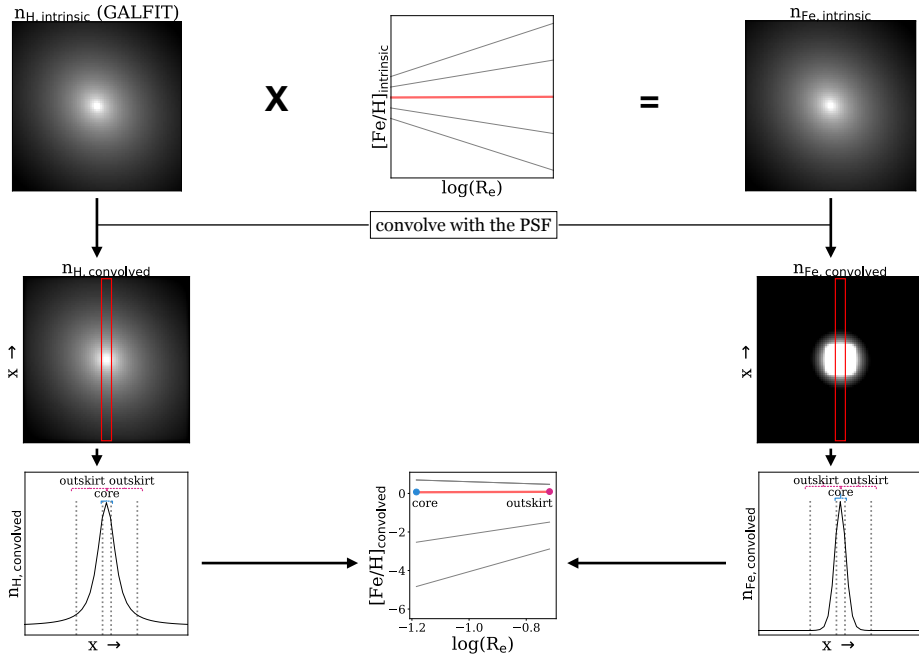


Figure 2.9: Graphical illustration of our intrinsic [Fe/H] gradient model. For a detailed description of the procedure, see the main body of the text. This process is repeated on a grid of 1000 intrinsic [Fe/H] gradient slopes. The slope that gives the minimum absolute residual between our measured slope and the model convolved slope is then our predicted intrinsic gradient.

Cook et al. 2016), except perhaps at very large radii (Cook et al. 2016, which are not reached by our measurements).

With this work, spatially resolved stellar population gradients have now been measured (with varying sample sizes) in massive quiescent galaxies out to $z \sim 1$. To obtain a detailed understanding of how these galaxies evolve from $z \sim 1$ to the present, we would need to quantitatively compare how the slopes of the stellar population gradients change over cosmic time. However, comparing results between different studies is challenging as there are often significant differences in the definition of gradients, data quality, and methods to derive ages and metallicities. In future work we will measure spatially resolved stellar population gradients using this same method out to $z \sim 3$ which will allow us to make a more homogeneous analysis of the evolution of gradients over cosmic time (see Section 2.5.2.3). However overall, our current findings of typically negative metallicity gradients and flat age and [Mg/Fe] gradients for older galaxies are qualitatively consistent with results from the literature, both at low z (e.g. Mehlert et al. 2003; Rose et al. 2005; Kuntschner et al. 2006; Sánchez-Blázquez et al. 2007; Greene et al. 2013, 2015; Greene et al. 2019; Pastorello et al. 2014; González Delgado et al. 2015; Martín-Navarro et al. 2018; San Roman et al. 2018; Ferreras et al. 2019; Oyarzún

et al. 2019; Zheng et al. 2019; Lacerna et al. 2020; Santucci et al. 2020; Yoon et al. 2023; Parikh et al. 2024) and at higher z (Jafariyazani et al. 2020). This suggests that these stellar population gradients were in place by at least $z \sim 1$ (with indications that this may even be true at $z \sim 2$, Jafariyazani et al. 2020) and are at least maintained until the current epoch.

2.4.2 Intrinsic metallicity gradients

While our detection of negative metallicity gradients is significant, it is important to note that our observations are affected by the seeing of the LEGA-C observations. To better understand the qualitative effect that the seeing has on our measurements, we develop a simple model to assess the metallicity gradients if we were to measure them in intrinsic space (i.e. in ideal conditions with no instrumental or atmospheric blurring).

The model that we create here is similar in principle to the forward model in Suess et al. (2019a) and is illustrated in Fig. 2.9. We first create an ‘intrinsic 2D hydrogen profile’. This is essentially the same GALFIT image that was generated in Section 2.3.2 as here we have assumed that the hydrogen column density profile (i.e. n_{H}) follows the light profile of the galaxy.

We use the hydrogen profile to create an intrinsic 2D [Fe/H] profile in linear space (i.e. $n_{\text{Fe}}/n_{\text{H}}$). To do this, we assume that the $n_{\text{Fe}}/n_{\text{H}}$ profile follows the expression

$$\log_{10} \left(\frac{n_{\text{Fe}}}{n_{\text{H}}} \right)_{\text{int}} = m_{\text{int}} \log(R_e) + b_{\text{int}} + \log_{10} \left(\frac{n_{\text{Fe}}}{n_{\text{H}}} \right)_{\odot}, \quad (2.2)$$

where m_{int} and b_{int} are the intrinsic slope and y -intercept of the [Fe/H] gradient, respectively, and we obtain $\log_{10}(n_{\text{Fe}}/n_{\text{H}})_{\odot} = -4.5$ from Asplund et al. (2009). We generate a grid of possible m_{int} (a grid of 1000 m_{int} with $-11 \lesssim m_{\text{int}} \lesssim 11$) which spans a much wider range than our measured $n_{\text{Fe}}/n_{\text{H}}$ gradient slopes (since we expect the intrinsic gradients to be stronger than the convolved ones, see also D’Eugenio et al. 2020). We set b_{int} ⁸ to the value of $n_{\text{Fe}}/n_{\text{H}}$ that we measure from the integrated fit. We generate a 2D radius profile in units of pixels. To avoid having a discontinuity in the centre of the simulated galaxy, we resample each pixel into ~ 500 subpixels, calculate the distance between each subpixel and the centre of the galaxy, and take the average of all of the subpixels in each pixel to be the distance between each pixel and the centre of the galaxy. We plug the grid of m_{int} ’s (the parameter for which we are trying to fit), b_{int} (which we fix to the integrated value), and 2D R_e profiles (which we fix to the values we calculate here) into equation 2.2 to obtain a grid of intrinsic 2D $n_{\text{Fe}}/n_{\text{H}}$ profiles.

We multiply the intrinsic 2D $n_{\text{Fe}}/n_{\text{H}}$ profiles by the intrinsic 2D n_{H} profile to obtain intrinsic 2D n_{Fe} profiles. We then convolve our intrinsic 2D n_{Fe} and n_{H} profiles with the best-fit seeing that we found in Section 2.3.2 and we collapse

⁸We note that, in this model, b_{int} is simply a scaling factor that indicates the metallicity at the centre of the galaxy. It does not affect that slopes of the gradients, the key information in which we are interested here.

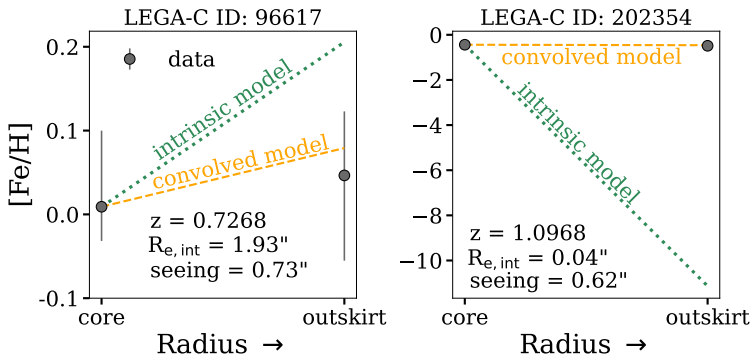


Figure 2.10: An illustration of how the intrinsic gradient model is applied to an example galaxy with a large R_e (left) and an example galaxy with a small R_e (right). We emphasize that this is a qualitative illustration of the intrinsic gradient slope that we obtain and the absolute $[\text{Fe}/\text{H}]$ values are not meant to be taken at face value. The grey circles indicate the spatially resolved $[\text{Fe}/\text{H}]$ values that we measure from the LEGA-C spectra, the dashed line is the recovered convolved model $[\text{Fe}/\text{H}]$ gradient, and the dotted line is the corresponding $[\text{Fe}/\text{H}]$ gradient that we predict in intrinsic space. In general, the intrinsic gradients are much steeper than those in convolved space.

these convolved images over the width of the VIMOS slit to obtain 1D profiles of n_{Fe} and n_{H} .

We weight the 1D profiles by the same weights that we use in our optimal extraction in Section 2.3.1 and sum each profile. We do this for each of our core (c) and outskirts (o) extractions to obtain a single value of each of $n_{\text{Fe},c}$, $n_{\text{H},c}$, $n_{\text{Fe},o}$, and $n_{\text{H},o}$ in convolved space for each grid point. We calculate the model convolved slopes:

$$m_{\text{conv}} = \frac{\left(\log \left(\frac{n_{\text{Fe},o}}{n_{\text{H},o}} \right) - \log \left(\frac{n_{\text{Fe},\odot}}{n_{\text{H},\odot}} \right) \right) - \left(\log \left(\frac{n_{\text{Fe},c}}{n_{\text{H},c}} \right) - \log \left(\frac{n_{\text{Fe},\odot}}{n_{\text{H},\odot}} \right) \right)}{\log(R_{e,o}) - \log(R_{e,c})} \quad (2.3)$$

and compare these to our measured slopes. We take the grid point with the smallest absolute residual between our model m_{conv} and our observed slope to be the intrinsic $[\text{Fe}/\text{H}]$ gradient of the galaxy.

We stress that this model serves to qualitatively demonstrate the effect that the seeing may be having on our observations, and that these intrinsic gradients should not be taken as the true gradients for the individual galaxies. As such, we do not model the uncertainties on our intrinsic gradient slopes. Our method works well for ~ 80 per cent of our sample (i.e. the model does not recover the observed $[\text{Fe}/\text{H}]$ gradient for 90 galaxies, due to a combination of their sizes being too small or their seeing being too large). We show a qualitative illustration of how the model is applied to two example galaxies (one with a large R_e and one with a small R_e in Fig. 2.10). We emphasize that this is a demonstration of how the best-fit convolved and intrinsic models look compared to the data. We do not make use of the results for individual galaxies here and consider only the average

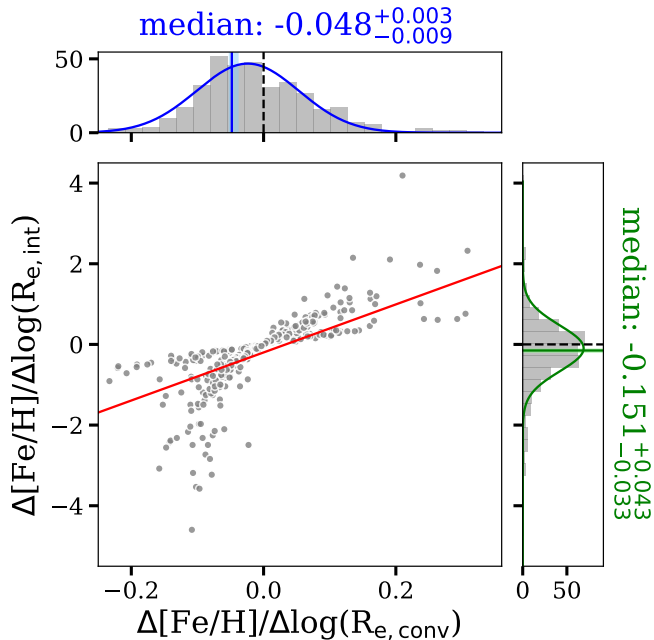


Figure 2.11: A comparison between our $[\text{Fe}/\text{H}]$ gradients measured in convolved space to those that we model in intrinsic space. The solid red line is the best fit from which we conclude that the intrinsic gradients are steeper than the observed ones by a factor of 5.96 ± 0.31 . A histogram of all intrinsic gradient predictions for our sample as well as a Gaussian fit to emphasize the shape of the distribution are shown on the right panel and the same for the measured gradients in convolved space is shown on top. We show the median gradients as solid lines with shaded regions indicating uncertainties derived by bootstrap resampling. The dashed lines show where flat gradients would lie. We quote the median gradient values and bootstrapped uncertainties above each panel. These are similar to the centre panel of Fig. 2.6.

trends.

In Fig. 2.11, we compare our observed $[\text{Fe}/\text{H}]$ gradients to our predicted $[\text{Fe}/\text{H}]$ intrinsic gradients. We show histograms for each type of gradient on the top and to the right, similar to Fig. 2.6. In the main panel, we fit a line to the comparison of the two types of gradients. The slope of this line is the average factor by which the intrinsic $[\text{Fe}/\text{H}]$ gradients are steeper than our observed gradients. In general, we find that the intrinsic $[\text{Fe}/\text{H}]$ gradients are steeper than our observed gradients, by a factor of 5.96 ± 0.31 . This indicates that the gradients we have measured here in massive quiescent galaxies at $0.6 \lesssim z \lesssim 1.0$ are significant.

We note that we only estimate the metallicity gradients here. Developing a similar model for the spatially resolved ages would be much more complex. However, we would expect any age gradients that we measure to be steeper in intrinsic space, with the effect being similar to what we find for the metallicity gradients. In particular, positive age gradients would be more positive in intrinsic

space and negative age gradients would be more negative. Thus, we would expect our discussion in Section 2.5 and our general conclusions to hold in intrinsic space.

The ‘intrinsic gradients’ that we measure here serve as a general demonstration of how much our measurements are affected by observational effects. None the less, it may still be interesting to compare our median intrinsic $[\text{Fe}/\text{H}]$ gradient of $\Delta[\text{Fe}/\text{H}] / \Delta \log(R_{\text{e,intrinsic}}) = -0.151^{+0.043}_{-0.033}$ to $[\text{Fe}/\text{H}]$ gradients in the literature. Again, we stress that it is difficult to compare results from different studies. If we take the measurements at face value, however, we see that our intrinsic $[\text{Fe}/\text{H}]$ gradients are generally shallower than $[\text{Fe}/\text{H}]$ gradients at lower z (e.g., Greene et al. 2015; Greene et al. 2019, etc.). For example, Greene et al. (2019) defined their $[\text{Fe}/\text{H}]$ gradients in a similar way and found a median of $\Delta[\text{Fe}/\text{H}] / \Delta \log(R) = -0.26$. We discuss this further in Section 2.5.2.

2.5 Discussion

2.5.1 Colour gradients

The spatially resolved spectroscopic measurements that we perform here are intimately linked with spatially resolved measurements from photometry. In cases where it has not been possible to obtain the high-quality spectra needed to resolve stellar populations, photometric data have been and are still used to examine spatially resolved stellar population properties.

Differences in stellar populations across the spatial distribution of a galaxy can be observed photometrically as colour gradients. In the local Universe, many studies have observed negative radial colour gradients in quiescent galaxies. These gradients have largely been attributed to negative metallicity gradients (i.e. Kormendy & Djorgovski 1989 and references therein, Peletier & Valentijn 1989; Franx & Illingworth 1990; Peletier et al. 1990b; Davies et al. 1993; Vazdekis et al. 1997; Saglia et al. 2000; La Barbera et al. 2005; Tortora et al. 2010, 2011), perhaps with a smaller contribution from age gradients (i.e. La Barbera & de Carvalho 2009; Tortora et al. 2010).

At higher redshifts, where even photometric observations have historically been more difficult to achieve, negative colour gradients have also been found on average in quiescent galaxies over a wide range in mass (i.e. Wuyts et al. 2010; Guo et al. 2011; Szomoru et al. 2012; Chan et al. 2016; Liu et al. 2017; Mosleh et al. 2017; Suess et al. 2019a; Miller et al. 2022, 2023; Setton et al. 2024). More specifically, it has been found that colour gradient strength increases with galaxy age, with younger quiescent galaxies having flatter colour gradients and older galaxies having more strongly negative colour gradients (Suess et al. 2020). This trend is further supported by the finding that PSB galaxies have flat colour gradients (Roche et al. 2010; Suess et al. 2021). However, using colours alone, it is impossible to break the age–metallicity degeneracy (Worthey 1994; Bruzual & Charlot 2003; Gallazzi et al. 2005). Thus, it is not necessarily clear what might be the origin of these colour gradients (or lack thereof) at higher redshifts. In particular, a redder colour might be driven by an older, more metal-rich, or dustier stellar population. For example, Miller et al. (2023) argued that metallicity gradients may be the primary

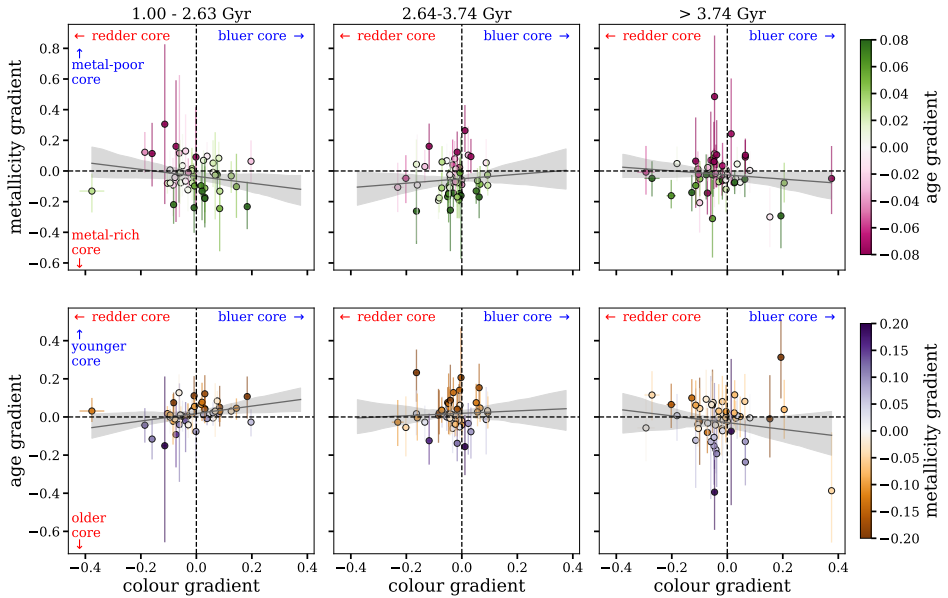


Figure 2.12: The relationship between our measured age and metallicity gradients and colour gradients. On the x -axis of each panel, we use the ratio between R_e measured in $HST/WFC3-F160W$ images (Cutler et al. 2022) to R_e measured in $HST/ACS-F814W$ images as a proxy for colour gradients ($\log(R_{e,F160W}/R_{e,F814W})$). In the top row, we show our $[\text{Fe}/\text{H}]$ gradients on the y -axis ($\Delta[\text{Fe}/\text{H}]/\Delta \log(R_e)$) and the symbols are colour coded by our age gradients ($\Delta \log(\text{Age}(\text{Gyr}))/\Delta \log(R_e)$). In the bottom row, we show our age gradients on the y -axis and the symbols are colour coded by metallicity gradients. We split this figure into three different age bins (the same as in Fig. 2.8). We show a linear least-squares fit to the points in each panel with the shaded region showing the uncertainties on the fit. We expect to see a positive trend if a particular gradient is the primary driver of the colour gradients. We see that there is a positive (negative) trend between age ($[\text{Fe}/\text{H}]$) and colour gradient at the youngest ages.

driver due the fact that they find negative colour gradients similar to low- z studies that have identified metallicity gradients as a primary driver, but this has yet to be shown definitively. Alternatively, Miller et al. (2022) and Setton et al. (2024) examined quiescent galaxies at $z \sim 2$ and ~ 4 , respectively with photometry and low-resolution spectroscopy, and found indications that the negative colour gradients are driven by dust. However, only a handful of quiescent galaxies were examined and metallicity was not considered. Furthermore, with low resolution spectroscopy, it is extremely challenging to break the age–metallicity degeneracy.

Our measurements can provide insight into the driver of observed colour gradients, at least at $0.6 \lesssim z \lesssim 1.0$. We examine the comparison between our spectroscopic measurements and photometric colour gradients in Fig. 2.12. In each panel, we show the ratio of R_e measured in $HST/WFC3-F160W$ images (Mowla et al. 2019) to R_e measured in $HST/ACS-F814W$ images (Scoville et al. 2007b) on the x -axis. This ratio is a proxy for colour gradient strength, as a positive

colour gradient (i.e. bluer centre), for example, would result in smaller sizes at bluer wavelengths (i.e. F814W). The F160W sizes are obtained from Cutler et al. (2022) and F814W sizes are measured by the LEGA-C collaboration, both using GALFIT (Peng et al. 2010). We exclude F160W sizes flagged as ‘bad’ in GALFIT. Unfortunately, the majority of the F160W sizes for our sample are not well-fitted by GALFIT so we are only able to calculate this ratio for a fraction of our sample. This is due to a combination of the imaging from Cutler et al. (2022) being relatively shallow and the fact that they do not have imaging for the entire Ultra-VISTA (McCracken et al. 2012) catalogue. In the top row, we show our measured $[\text{Fe}/\text{H}]$ gradients on the y -axis and the symbols are colour coded by our measured age gradients. In the bottom row, we show the reverse – measured age gradients are on the y -axis and the symbols are colour coded by measured $[\text{Fe}/\text{H}]$ gradients. We have split the sample into three age bins (the same ones as in Fig. 2.8). In each panel, we show a linear least-squares fit to the points, with the shaded region showing the uncertainties on the fit. If a particular gradient is the primary driver of the colour gradient, then we expect to see a positive trend between that gradient and the colour gradient strength.

For the younger galaxies ($1 \text{ Gyr} \leq \text{age} < 2.63 \text{ Gyr}$, left panel of Fig. 2.12), we find a positive trend between age and colour gradients and a negative trend between $[\text{Fe}/\text{H}]$ and colour gradients. This suggests that the flat colour gradients found in young quiescent galaxies by Suess et al. (2020) may actually be the result of the age and metal gradients compensating each other. In particular, a positive age gradient is expected to contribute to a bluer core and redder outskirts while a negative metallicity gradient results in the opposite. Thus, in colour space, the two gradients effectively cancel each other out. We note that it is beyond the scope of this paper to quantify how positive an age gradient would have to be to compensate a negative metallicity gradient, however our results are qualitatively indicative of this scenario.

For the older galaxies ($> 2.63 \text{ Gyr}$, middle and right panels of Fig. 2.12), we see a flat relationship between colour and age gradients. The relationship with $[\text{Fe}/\text{H}]$ gradient is also mostly flat. We cannot conclude much from this figure, which may not be surprising as the observed age and metal gradients are weaker for these galaxies and the uncertainties on the colour gradients are large. However, while no clear trend is seen with $[\text{Fe}/\text{H}]$ gradients, the lack of trend between age and colour gradients may suggest that colour gradients at older ages are most likely driven primarily by $[\text{Fe}/\text{H}]$ gradients. Of course, though our observed stellar population gradients can, qualitatively, explain the observed colour gradients, we cannot rule out that dust is contributing to the colour gradients as well (see Miller et al. 2022; Setton et al. 2024) as we do not consider the effects of dust in our analysis. In particular, if the cores of these galaxies are dustier than their outskirts, then their observed colour gradients may be driven primarily by dust instead of age or metallicity. More work is needed to determine the primary driver of colour gradients at $z > 0$.

2.5.2 Implications for galaxy evolution

In our analysis, we find that there are distinct differences in the stellar population gradients of galaxies at different ages, with younger galaxies having positive age gradients and stronger metallicity gradients, and older galaxies having flat age gradients and weaker metallicity gradients. This suggests that the cores of the younger galaxies are younger and more metal-rich than their outskirts, and the cores of older galaxies, while still being more metal-rich, are the same age as their outskirts. Furthermore, we find flat $[\text{Mg}/\text{Fe}]$ gradients at all ages, which implies that the core and outer regions of the galaxies in our sample have similar star formation time-scales (see e.g. Maiolino & Mannucci 2019). However, we caution that it is difficult to interpret our $[\text{Mg}/\text{Fe}]$ gradients as we measure them for a much smaller sample compared to our age and $[\text{Fe}/\text{H}]$ gradients. We note that there is a significant amount of scatter in our gradient measurements for individual galaxies, however these average trends still have interesting implications for galaxy assembly histories. Here we will discuss the possible scenarios suggested by our results, first for the younger galaxies and then for the older galaxies.

2.5.2.1 Young quiescent galaxies

We find that massive quiescent galaxies in the youngest age (i.e. $1 \text{ Gyr} \leq \text{age} < 2.63 \text{ Gyr}$ old) bin of our sample tend to have positive age gradients and negative metallicity gradients. In other words, their cores tend to be younger and more metal-rich than their outskirts. Thus, this core stellar population may be indicative of a central starburst that occurred just before the galaxy stopped forming stars. These observed age gradients were predicted observationally by Wu et al. (2020) for young, compact quiescent galaxies, to explain their observed increase in half-light radii with age. As the central starburst fades, the older, more extended population will become more dominant, resulting in an increase of the half-light radius (see also Whitaker et al. 2012; Yano et al. 2016; Almaini et al. 2017; Maltby et al. 2018). Pathak et al. (2021) also showed that the majority of their IllustrisTNG-simulated young quiescent galaxies have positive age gradients as a result of recent central starbursts (although a significant portion of their young quiescent galaxy sample was also consistent with having flat age profiles). In this context, it is furthermore interesting to note that positive age gradients may also be present in the compact star-forming population, as their star formation is more centrally concentrated than the stellar population that is already present (e.g., Barro et al. 2016b, 2017; Tadaki et al. 2017; Popping et al. 2017). Therefore, compact star-forming galaxies which were rapidly quenched may indeed be the progenitors of young quiescent galaxies (Barro et al. 2013, 2014a,b, 2016a, 2017; D'Eugenio et al. 2020).

Our results provide insights into the quenching mechanisms of young quiescent galaxies. In particular, we can discriminate between two popular suggested mechanisms in the literature at these redshifts: wet compaction events or wet, gas-rich major mergers. In the case of wet compaction, clumps of gas migrate towards the centre, compacting the galaxy by decreasing the R_e . This occurs preferentially at high redshifts. Galaxies may then experience extended star formation in the outskirts after the gas in the core is depleted. In particular, simulations have shown

that following the compaction phase, compact galaxies may be encircled by an extended disc of young stars and low-density gas. Quenching post-compaction is then expected to occur inside–out (Bell 2008; Bell et al. 2012; Fang et al. 2013; Li et al. 2015; Tacchella et al. 2015a, 2016a; Zolotov et al. 2015; González Delgado et al. 2016; Lin et al. 2017; Ellison et al. 2018; Sánchez et al. 2018; Spilker et al. 2019; Breda et al. 2020; Nelson et al. 2021; Abdurro’uf et al. 2023; Avila-Reese et al. 2023; Lapiner et al. 2023). In this way, a compaction event would be expected to produce negative age gradients (Dekel & Burkert 2014; Zolotov et al. 2015; Tacchella et al. 2016a). On the other hand, mergers can trigger a strong starburst by driving gas to the centre of the galaxy, producing a positive age gradient (Hopkins et al. 2008; Snyder et al. 2011; Wellons et al. 2015; Pathak et al. 2021). Thus, our results support a merger-triggered central starburst event, or another quenching mechanism that results in a central starburst just prior to quenching. In this regard, our conclusions are consistent with those of Setton et al. (2020).

2.5.2.2 Growth of the quiescent galaxy population

In contrast to the mildly positive age gradients that we find in younger galaxies, older quiescent galaxies in our sample tend to have flat age and [Mg/Fe] gradients and weaker (but still negative) [Fe/H] gradients. In this section, we will discuss what these observations may imply for the evolution of individual quiescent galaxies, as well as for the growth of the quiescent galaxy population over time.

Our finding that the outskirts of massive quiescent galaxies have lower metallicities compared to their cores is supportive of inside–out formation, which has been suggested by many previous studies (e.g., van der Wel et al. 2008; Bezanson et al. 2009; Hopkins et al. 2009a; Naab et al. 2009; Oser et al. 2010; van Dokkum et al. 2010; Moster et al. 2013; van de Sande et al. 2013; Abramson et al. 2014; Somerville & Davé 2015; Nelson et al. 2016; Hill et al. 2017; Tacchella et al. 2019; Suess et al. 2019a, 2020; Santucci et al. 2020; Pulsoni et al. 2021; Conselice et al. 2022; Beverage et al. 2023). For example, the measurement of small sizes of distant quiescent galaxies points to a scenario in which quiescent galaxies build up their outskirts by the accretion of lower-metallicity, younger galaxies over cosmic time (e.g., Trujillo et al. 2004, 2006; Franx et al. 2008; van der Wel et al. 2008, 2014; Bezanson et al. 2009; Naab et al. 2009; van Dokkum et al. 2010). In this scenario, the lack of an age gradient in older quiescent galaxies may be explained by central rejuvenation, as discussed in Section 2.5.2.1. This may lower the average age (and [Mg/Fe]) in galaxy cores to a similar age (and [Mg/Fe]) to that of the accreted galaxies. Thus, our observed gradients may support inside–out growth driven by *minor mergers* (e.g., Bezanson et al. 2009; Naab et al. 2009; Oser et al. 2010, etc.) associated with central rejuvenation. The minor merger mechanism is also supported by our finding of shallower intrinsic [Fe/H] gradients compared to low- z studies such as Greene et al. (2019) (Section 2.4.2), as minor mergers would cause [Fe/H] gradients to become steeper towards lower z (Naab et al. 2009; Oser et al. 2010). On the other hand, our results disfavour inside–out growth driven by *late-time star formation* (i.e., Abramson et al. 2014; Nelson et al. 2016). In this case, the outskirts of quiescent galaxies would be built up by this late-time star

formation, resulting in younger stellar populations in the outskirts (i.e. a negative age gradient).

Another popular scenario that can explain the size growth of the quiescent galaxy population over cosmic time is that of progenitor bias, where galaxies that quench at later times are larger (van Dokkum & Franx 2001; Carollo et al. 2013; Poggianti et al. 2013). Age and metallicity gradients are generally expected in this case (Larson 1974; White 1980; Kobayashi 2004), and the models of Pipino et al. (2008) and Pipino et al. (2010) have shown that it is possible to produce negative $[\text{Fe}/\text{H}]$ gradients and flat $[\alpha/\text{Fe}]$ gradients. Thus, in this scenario, individual galaxies evolve passively and the gradients are directly inherited from the star-forming phase. Additionally, galaxies in our older age bin quenched at earlier times ($z_{\text{form}} \gtrsim 2$), while galaxies in our younger age bins quenched at later times ($0.9 \lesssim z_{\text{form}} \lesssim 1.4$). High- z star-forming galaxies have indeed been found to have negative metallicity gradients (e.g., Jones et al. 2010, 2013, 2015) and negative or flat age gradients (Tripodi et al. 2024; Shen et al. 2024), consistent with our findings. Furthermore, colour gradients of star-forming galaxies have been found to become stronger at lower z (Mosleh et al. 2017; Suess et al. 2019a,b). Thus, our findings are qualitatively consistent with the progenitor bias scenario as well.

In summary, our observed gradients are consistent with both the minor merger and progenitor bias scenarios. In fact, it is likely that both are playing a role, as suggested by, for example, van der Wel et al. (2008), Hopkins et al. (2010), Valentinuzzi et al. (2010), Oser et al. (2012), Newman et al. (2012), Nipoti et al. (2012), and Barro et al. (2013). For instance, it may be that mergers are only important for the most massive galaxies ($\log(M_*/M_\odot) > 11$, Carollo et al. 1993; Rodriguez-Gomez et al. 2016). It is also possible that different mechanisms can contribute to the assembly of specific kinds of galaxies at specific redshifts (Belli et al. 2014, 2015, 2017; Wellons et al. 2015, 2016). For example, Suess et al. (2021) found progenitor bias to be primarily important in slow-quenching green valley galaxies at low redshifts. None the less, Beverage et al. (2024) and Kriek et al. (2024) also argue for some contribution from major mergers to explain the metal evolution and dynamical masses of distant quiescent galaxies.

The scatter in our gradient measurements could be used as an additional probe to understand the importance of the different processes. However, the uncertainties in our individual measurements are quite large, and the significant amount of scatter in our observed gradients can be attributed primarily to these measurement uncertainties. Thus, further observations are required to conclude which process may be dominant.

2.5.2.3 Future work

Our findings contribute to a vast body of work that aims to understand the build-up of the quiescent galaxy population. However, much still remains to be done to understand which of the scenarios that we have outlined in the previous sections may describe the primary mechanism of galaxy assembly, and to learn how the negative $[\text{Fe}/\text{H}]$ gradients originate. There are two main ways in which we can expand upon this work: going to higher redshifts and increasing our spatial

resolution.

Examining spatially resolved stellar population gradients at higher redshifts will allow us to understand how these gradients have evolved over time. For example, Suess et al. (2019b) found that colour gradients at $z \lesssim 1.0$ (i.e. the regime explored in this study) remain relatively constant, consistent with the lack of trend that we see between our observed gradients and z (Section 2.4.1). However, at higher redshifts ($1.0 \lesssim z \lesssim 2.5$), they found that colour gradients evolve rapidly. It is important to investigate whether similar evolution happens in spectroscopic gradients. Understanding how gradients change over time will in turn allow us to better constrain the primary mechanism of galaxy assembly. In particular, in the scenario of hierarchical formation, it is predicted that the compact central regions of galaxies are formed at $z \sim 2 - 3$ and are built up by accretion at later times (Oser et al. 2010; Rodriguez-Gomez et al. 2016). In this case, we would expect minor mergers to strengthen [Fe/H] gradients toward $z \sim 0$. However, for example, Jafariyazani et al. (2020) found a steeper [Fe/H] gradient in an individual lensed galaxy at $z \sim 2$ than measurements at low z and our predicted intrinsic [Fe/H] gradients (Section 2.4.2). Probing closer to the epoch of formation of these compact cores with a larger sample may allow us to point to a dominant assembly mechanism. In an upcoming study, we will extend the redshift range of this analysis up to $z \sim 3$ with ultra-deep, medium resolution spectra from the Spectroscopic Ultradeep Survey Probing Extragalactic Near-infrared Stellar Emission (SUSPENSE), a Cycle 1 *JWST*/NIRSpec survey of 20 distant quiescent galaxies at $1 < z < 3$ (Slob et al. 2024).

Increased spatial resolution is also important for the accurate interpretation of stellar population gradients. In particular, Oyarzún et al. (2019) found that average radial metallicity profiles of low- z galaxies from the Mapping Nearby Galaxies at APO (MaNGA) survey are not linear and thus a single gradient value does not fully characterize metallicity profiles. Specifically, they found that metallicity profiles flatten in galaxy outskirts. This points to stellar accretion from merging galaxies, showing that higher resolution data can also provide key insights about galaxy assembly mechanisms. In this study we do not recover detailed stellar population profiles (i.e. with more than two resolution elements). This is primarily because we do not have the required S/N in smaller spatial bins to reliably recover ages and metallicities. Higher S/N, higher spatial resolution spectra can make this possible. We hope to achieve this in future work with *JWST*-SUSPENSE as well. Additionally, IFU data from the Multi Unit Spectroscopic Explorer (MUSE) or the Enhanced Resolution Imager and Spectrograph (ERIS) on the *VLT* may provide more detailed measurements.

2.5.3 Caveats

The depth and resolution of the LEGA-C spectra in combination with the ALF full spectrum models have allowed us to measure statistically significant spectroscopic age and metal gradients beyond the low- z Universe for the first time. However, there are several caveats that we address here and that should be kept in mind when interpreting our results.

The most impactful consideration is the fact that our data are highly blurred by atmospheric seeing and instrumental effects, combined with the fact that the VIMOS slit and our 1D spectral extraction regions are relatively narrow compared to the size of the galaxy. Thus, contamination between the core and outskirt regions is a concern for our gradient measurements. However, we would expect any gradient that we are able to detect in the convolved space of the LEGA-C data to be even stronger in intrinsic space. Therefore, the fact that we are able to detect differences between the central and outer regions of the galaxies at all is extremely encouraging and indicates that the gradients are significant and likely stronger than what we report here. This fact and the effect of any potential contamination between the core and the outskirts is also demonstrated by our intrinsic model in Section 2.4.2. Future work with data from the space-based *JWST* will partially mitigate this issue.

Additionally, we argue that our optimal extraction procedure in Section 2.3.1 does not affect our conclusions. Recall from Section 2.3.1 that we do not include the outermost wings of the spectra (the outermost 3 per cent of the spectral rows on each side of the profile, see shaded regions in the right panel of Fig. 2.3), as we find that they introduce a lot of noise. Our detailed extraction routine thus maximizes our sample size while also maintaining a similar S/N between the core and the outskirts for most objects. We perform several tests and find that modifications to the exact extraction do not impact our conclusions. For example, we test including all of the spectral rows, using the profile itself to weight the extraction, and testing different proportions of rows in the core and the outskirts (i.e. while discarding the outer 6 per cent of the wings, we have tested a 47 per cent/47 per cent split and a 44 per cent/50 per cent split). In each test, we are able to recover the same trends as presented in Section 2.4, with the median values for each gradient in each test within uncertainties of each other. Thus, the wings of the spectra do not contribute additional information to the gradients in these galaxies with the S/N that we have and our precise extraction routine does not impact our conclusions.

Another concern with this method of spectral binning is contamination of the core spectrum with outskirt information and vice versa, due to inclination effects. For galaxies which we observe face-on, this is not an issue as the core and outskirts are distinct in the 2D spectra. However, for galaxies which we observe edge-on or with some inclination, contamination may be an issue. We find that the distribution of gradients in age, metallicity, and Mg are approximately the same in bins of increasing axis ratio, so we do not expect this to be a significant issue.

We also perform tests to understand whether oversubtraction of the sky background could result in the false detection of gradients. Oversubtraction would preferentially affect the outskirts and artificially weaken the absorption lines compared to the core, potentially creating a gradient where none exists. However, this would affect all features across the spectrum equally, and we see in the inset panels of Fig. 2.4 that this is not the case. Specifically, the $H\beta$ feature is not significantly different between the core and the outskirts, whereas the comparably-deep Fe52 feature is distinct between the core and outskirt regions. Moreover, we test whether we would be able to recover a gradient if the sky background was

oversubtracted. For a subsample of galaxies, we create a simulated oversubtracted outskirts spectrum by subtracting a constant value of 1 per cent of the maximum of the core spectrum from the core spectrum and fit this with ALF. In this test, we typically find flat age, [Fe/H], and [Mg/Fe] gradients. Thus, our detection of gradients is not the result of oversubtraction of the sky background.

Finally, the age–metallicity degeneracy is a well-known issue in stellar population studies, whereby the effects of age and metallicity on integrated light can be confused (Worthey 1994; Bruzual & Charlot 2003; Gallazzi et al. 2005). This can severely bias measurements of ages and metallicities. We confirm that the depth, resolution, and spectral coverage of the LEGA-C spectra allow us to disentangle the effects of age and metallicity on our spectra. In particular, our gradient measurements are not impacted by the age–metallicity degeneracy. We show this in Appendix 2.A, where we simulate different kinds of age and metallicity gradients in mock spectra, and find that we are able to recover the input gradients.

2.6 Summary and conclusion

In this paper, we presented spatially resolved stellar population gradients for 456 massive quiescent galaxies at $0.6 \lesssim z \lesssim 1.0$ from the LEGA-C survey. We extracted 1D integrated spectra as well as 1D spectra comprising the core and outskirts of each galaxy from the LEGA-C 2D spectra. We measured spatially resolved ages, [Fe/H], and [Mg/Fe] by fitting the spectra with a flexible, full-spectrum SPS model, ALF. We also forward-modelled what we would expect our [Fe/H] gradients to look like in unblurred, intrinsic space. Our main conclusions are summarized here:

- On average, massive quiescent galaxies at these redshifts have flat age (median $\Delta \log(\text{Age}(\text{Gyr}))/\Delta \log(R_{e,\text{convolved}}) = 0.007^{+0.002}_{-0.004}$) and [Mg/Fe] (median $\Delta[\text{Mg}/\text{Fe}] / \Delta \log(R_{e,\text{convolved}}) = -0.008 \pm 0.007$) gradients and negative [Fe/H] gradients (median $\Delta[\text{Fe}/\text{H}] / \Delta \log(R_{e,\text{convolved}}) = -0.048^{+0.004}_{-0.009}$), corresponding to more iron-rich galaxy cores. These results are consistent with what has been found for quiescent galaxies in the literature, both in the local Universe (e.g., Greene et al. 2015; Martín-Navarro et al. 2018) and at $z \sim 2$ (Jafariyazani et al. 2020).
- In intrinsic space, we find that our [Fe/H] gradients are, on average, expected to be a factor of 6.0 ± 0.3 steeper than in observed space. Thus, our intrinsic gradients are much steeper than our measured gradients, with a median of $\Delta[\text{Fe}/\text{H}] / \Delta \log(R_{e,\text{convolved}}) = -0.15^{+0.04}_{-0.03}$.
- When we split our galaxy sample into three equal age bins, we find that the youngest quiescent galaxies have positive age gradients, while the older quiescent galaxies have flat age gradients. The [Fe/H] gradients weaken with age, but remain negative at all ages. The [Mg/Fe] gradients remain flat for all age bins, implying that the outer and inner regions formed over similar timescales (see e.g. Maiolino & Mannucci 2019).

- These findings suggest that photometrically measured flat colour gradients in young quiescent galaxies (and PSB galaxies) at high redshifts (Suess et al. 2020, 2021) may be the result of the positive age and negative metallicity gradients compensating each other. Meanwhile, negative colour gradients found in older quiescent galaxies (Wuyts et al. 2010; Guo et al. 2011; Szomoru et al. 2012; Chan et al. 2016; Liu et al. 2017; Mosleh et al. 2017; Suess et al. 2019a, 2020; Miller et al. 2022, 2023; Setton et al. 2024) are likely driven primarily by metallicity gradients.
- Furthermore, our results indicate that the gradients we observe in young quiescent galaxies may be due to a recent central starburst. These findings support a merger-triggered central starburst event as the primary quenching mechanism for young quiescent galaxies out to $z \sim 1$ (also predicted to produce/maintain positive age gradients).
- Our finding of negative metallicity gradients is supportive of a picture of inside-out growth via minor mergers of less massive, lower metallicity satellite galaxies. On the other hand, our results are also consistent with the progenitor bias scenario, with gradients being inherited directly from the star-forming phase. In reality, it is likely that both minor mergers and progenitor bias are playing a role in the assembly of massive quiescent galaxies.

This work represents the first study of detailed spectroscopic stellar population gradients in a statistically significant quiescent galaxy sample beyond the low-redshift Universe. These observations show that simply using colour gradients does not capture the full breadth of stellar population gradients in distant galaxies and have given us deeper insights into the possible mechanisms of galaxy formation and evolution. However, we are still limited by observational effects due to ground-based observations and moderate redshifts, as the rest-frame optical shifts into the near-infrared beyond $z = 1$. With its unprecedented sensitivity at near-infrared wavelengths and its high spatial resolution, *JWST* will allow us to significantly improve upon the current work and push it to higher redshifts in the near future.

Acknowledgments

We thank the anonymous referee for a useful report that improved this manuscript. We thank the LEGA-C team for making their dataset public. We also thank Piyush Sharda, Jesse van de Sande, Ivana van Leeuwen, and Colin Yip for useful conversations. This work was performed using the compute resources from the Academic Leiden Interdisciplinary Cluster Environment (ALICE) provided by Leiden University. This work also used the Dutch national e-infrastructure with the support of the Samenwerkende Universitaire Rekenfaciliteiten (SURF) Cooperative using grant no. EINF-6344 which is financed by the Dutch Research Council (NWO). MK acknowledges funding from the NWO through the award of the Vici grant VI.C.222.047 and National Science Foundation (NSF) Astronomy and Astrophysics Research Grant (AAG) AST-1909942. RB acknowledges support from the Research Corporation for Scientific Advancement (RCSA) Cottrell Scholar Award

ID No: 27587 and from the National Science Foundation NSF-CAREER grant # 2144314. FDE acknowledges support by the Science and Technology Facilities Council (STFC), by the European Research Council (ERC) through Advanced Grant 695671 ‘QUENCH’, and by the UK Research and Innovation (UKRI) Frontier Research grant RISEandFALL. PEMP acknowledges support from the NWO through the Veni grant VI.Veni.222.364. P.F.W. acknowledges funding through the National Science and Technology Council grant 111-2112- M-002-048-MY2. A.G. acknowledges support from Italian National Institute for Astrophysics (INAF)-Minigrant-2022 ‘LEGA-C’ 1.05.12.04.01.

Data Availability

This study makes use of data from the LEGA-C survey. The two-dimensional and reduced one-dimensional spectra can be obtained from the ESO Science Archive Facility (http://archive.eso.org/eso/eso_archive_main.html). The reduced spectra and catalogue have been released by ESO (<http://archive.eso.org/cms/eso-archive-news/Third-and-final-release-of-the-Large-Early-Galaxy-Census-LEGA-C-Spectroscopic-Public-Survey-published.html>) and are also available here: <https://users.ugent.be/~avdrwel/research.html#legac>. For more details, please see van der Wel et al. (2016), van der Wel et al. (2021), and Straatman et al. (2018).

Individual stellar population ages and elemental abundances will be published in an upcoming work. Other data products generated in the course of this work (i.e. individual gradients and intrinsic model gradients) will be made available upon reasonable request.

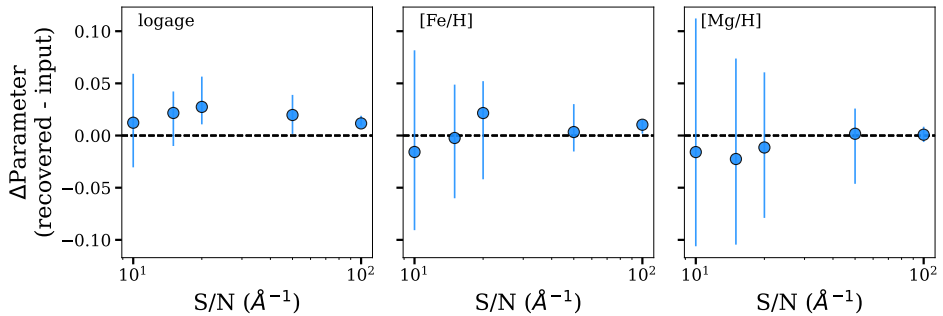


Figure 2.13: Mock spectra tests to determine the minimum S/N per Å that we need in order to recover age (left panels), [Fe/H] (middle panels), and Mg abundances (right panels) when we fix the other elemental abundances to their input values.

Appendix

2.A Mock spectra tests

We perform several tests with simulated mock LEGA-C spectra as discussed in Section 2.2.2. In general, we make use of the `write_a_model` framework in ALF to generate mock spectra. We give the mock spectra specific values of σ , age, and elemental abundances, as well as an S/N of 1000 \AA^{-1} . We then scale the mocks to different S/N. To do this, we take the noise spectra from real LEGA-C galaxies and scale the noise to the desired S/N. We also randomly sample the noise from a Gaussian distribution and add this random noise to our mock flux to achieve a more realistic simulated spectrum.

First, we test the S/N that we need to recover ages, [Fe/H], and Mg abundances in spatially resolved bins by generating several sets of mock spectra with typical properties (solar abundances, $\sigma = 150 \text{ km s}^{-1}$, and an age of 3 Gyr) at varying levels of S/N. For each S/N bin, we generate 40 mock spectra and fit each with ALF, leaving age, [Fe/H], and [Mg/H] free and fixing all other abundances to their input values (similar to our core and outskirt fits for our real data). We take the differences between the recovered and input parameters from each fit. We compute the medians of all of these differences and plot the median and 1σ errors of the distribution of medians in Fig. 2.13. From this test, we determine that we need $S/N \gtrsim 20 \text{ \AA}^{-1}$ to reliably recover the parameters of interest. Choi et al. (2014) performed similar tests and came to a similar conclusion for quiescent galaxies at low- z .

We also use these mocks to test which elemental abundances we are able to recover by fixing different combinations of parameters, including different elements and velocity dispersion. We find that our results are robust when we fix all elemental abundances except for Fe and sometimes Mg (depending on the wavelength range, see Section 2.2.2), and leave age and velocity dispersion free.

We note that there is a very small offset between the input and recovered ages

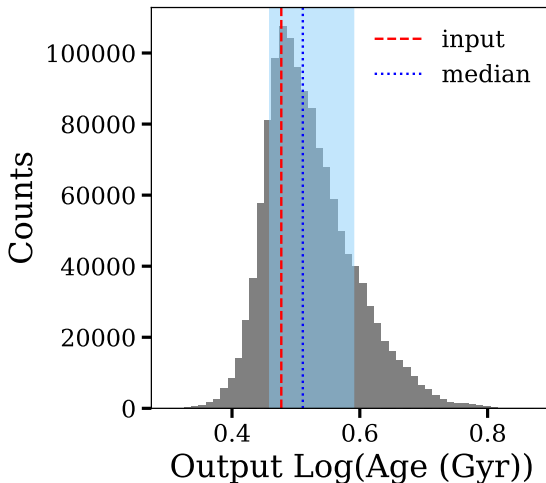


Figure 2.14: The distribution of the stacked MCMC chains for the recovered age from 40 mock spectra with $S/N = 20 \text{ \AA}^{-1}$. The median of this distribution is indicated by the dotted line, with the shaded regions indicating the uncertainties. The dashed line indicates the input age value.

in the leftmost panel of Fig. 2.13. We examine this by stacking all 40 of the MCMC age chains for the $S/N = 20 \text{ \AA}^{-1}$ test and plotting the distribution in Fig. 2.14. The distribution has a long tail towards old ages, which skews the median toward a slightly older age. However, the peak of the distribution is precisely at the input age, which shows that ALF is indeed recovering the input age. We still use the median age in our analysis as the offset is very small. Additionally, we note that because the ages are systematically offset to slightly older ages, this does not affect our conclusions as we examine the difference between the age in the core and the outskirts, and not the absolute value of the age.

Since the LEGA-C galaxies are at varying redshifts, they have variable rest-frame wavelength ranges, so we perform a second test to determine the shortest wavelength range for which we can still recover $[\text{Fe}/\text{H}]$. We similarly generate mock spectra as for the S/N tests, but with varying maximum wavelengths. Specifically, we generate mock spectra with wavelengths ranging from $3700 \text{ \AA} \leq \lambda \leq (4200 - 5000) \text{ \AA}$ to supplement the real data. We give the mock spectra $\sigma = 150 \text{ km s}^{-1}$, an age of 3 Gyr, and a solar abundance pattern. We generate 20 mock spectra for each wavelength range. In this test, we treat our simulated spectra like our real integrated spectra and leave all parameters free in the ALF fits. We determine the differences between our recovered and input $[\text{Fe}/\text{H}]$ and plot the medians in Fig. 2.15 (similar to Fig. 2.13). We are able to reliably recover $[\text{Fe}/\text{H}]$ down to a maximum wavelength of 4450 \AA . We note that the accuracy of the age recovery is less dependent on the upper wavelength range, as there are many Balmer absorption lines redward of 3700 \AA .

Finally, we note that there is an anticorrelation between our measured gra-

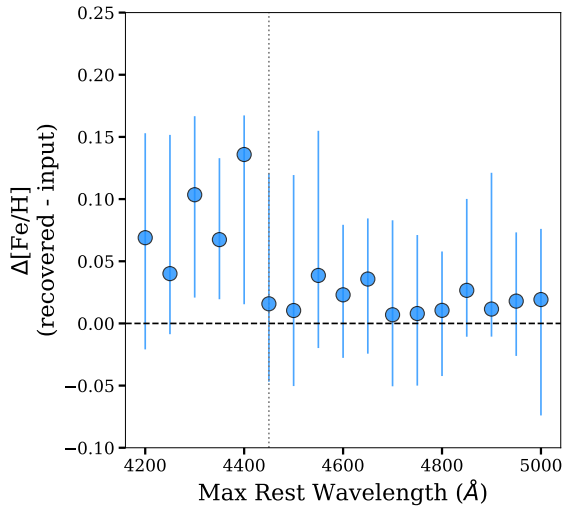


Figure 2.15: Mock spectra tests to determine the minimum maximum wavelength that can be used to reliably measure metallicity gradients. This minimum maximum wavelength cut-off (4450 Å) is indicated by the vertical dotted line (i.e. spectra need a maximum wavelength of at least 4450 Å in order for us to recover metallicity gradients). On the x -axis, we show the maximum simulated wavelength. On the y -axis, we show the difference between the recovered and input $[\text{Fe}/\text{H}]$. Mock spectra are given a median S/N of 20 \AA^{-1} (the S/N limit of our study).

dients in age and metallicity, so we assess whether this may be caused by the well-known age–metallicity degeneracy (Worthey 1994; Bruzual & Charlot 2003; Gallazzi et al. 2005). We find that this is not the case. In particular, even if the individual measured values of age and metallicity are affected by the degeneracy, the measured *gradients* in these quantities should not change. To test this, we generate several sets of 47 simulated LEGA-C galaxies with different combinations of age and metallicity gradients (i.e. no gradients, an age gradient but no metallicity gradient, a metallicity gradient but no age gradient, etc.). We adopt the S/N of the corresponding real LEGA-C galaxies for each simulated set of spectra (ranging from a median value of between $\sim 20 - 113 \text{ \AA}^{-1}$ in the simulated core spectra and $\sim 20 - 126 \text{ \AA}^{-1}$ in the simulated outskirt spectra). We find that we are always able to recover the input gradient within uncertainties. We show this in Fig. 2.16, where the comparison between our median recovered and input age gradients are shown in the top panel and the same for metallicity is shown in the bottom panel.

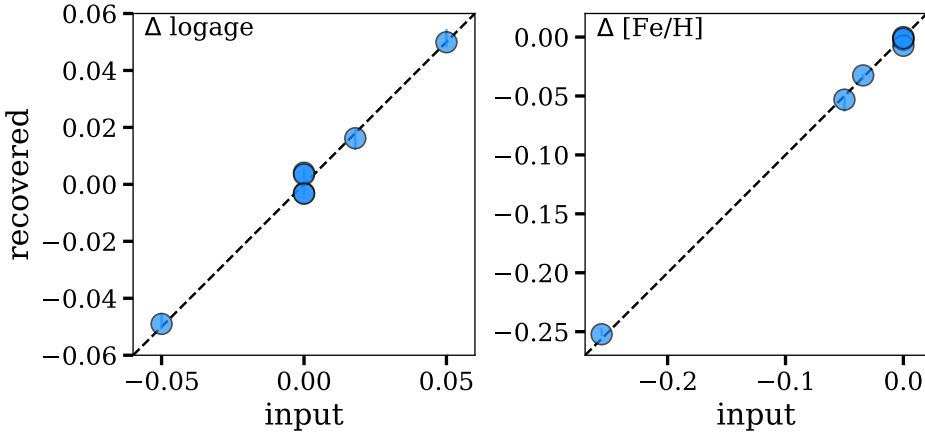


Figure 2.16: Mock spectra tests to demonstrate that we are able to recover age and metallicity gradients in our data. The S/N of the simulated spectra is the same as that of the corresponding real LEGA-C spectra. On the x -axes we show the input gradient and on the y -axes we show the recovered gradient. The median gradients for $\log(\text{age})$ are shown in the left panel and the median gradients for $[\text{Fe}/\text{H}]$ are shown in the right panel. Figure adapted for this thesis.

2.B Duplicate objects

There are several galaxies in LEGA-C that have been observed twice, in two different masks. In the main body of the paper, we have completed our analysis treating each duplicate observation as an individual galaxy. Here, we compare the fits to the galaxies in one mask to those in the other mask. Here we call them Mask ‘A’ and Mask ‘B’, which should not be confused for LEGA-C’s actual masks (numbered from 1 to 102). We also compare our results to results from a fit to the stack of both duplicate observations. We create these stacks for the integrated, core, and outskirts regions. We do not smooth the spectra to the same resolution, as they are spectra of the same object. We regrid the spectra to the same wavelength range (the different LEGA-C masks have slightly different wavelength coverage) using SPECTRES (Carnall 2017) and mean stack the spectra. In the stacked core and outskirts, we fix the abundances other than Fe and Mg to the values from the fits to the stacked integrated spectra. In Fig. 2.17, we plot the difference in ages between the outskirts and core regions in the top panel and the difference in $[\text{Fe}/\text{H}]$ between the outskirts and core regions in the bottom panel. We compare the results from Mask A and Mask B and colour code by the stacked results. In general, the results between all three fits are consistent within error bars. This is an independent check of the robustness of the results.

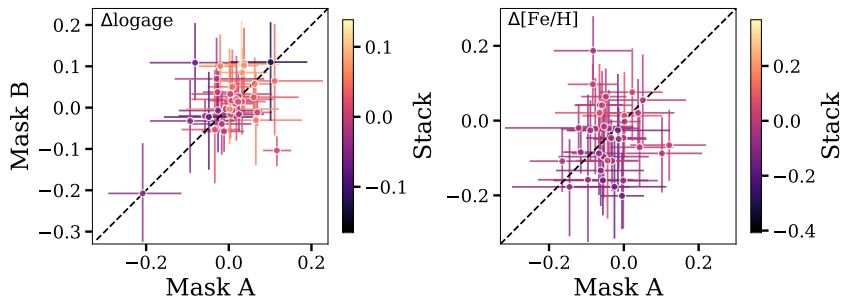


Figure 2.17: A comparison of the duplicate objects that are observed twice (in different masks) in LEGA-C. The differences between the outskirts and core ages are shown in the left panel and the differences between the outskirts and core $[\text{Fe}/\text{H}]$ values are shown in the right panel. The results from one mask (Mask ‘A’) are shown on the x -axis and the results from the other mask (Mask ‘B’) are shown on the y -axis. The symbols in each panel are colour coded by the stacked results. Figure adapted for this thesis.

3 | AGES AND METALLICITIES OF QUIESCENT GALAXIES: CONFRONTING BROAD-BAND (UVJ) COLOURS WITH STELLAR ABSORPTION LINES

ABSTRACT

For decades, studying quiescent galaxies beyond $z \sim 1$ has been challenging due to the reliance on photometric spectral energy distributions, which are highly susceptible to degeneracies between age, metallicity, dust, and star-formation history. Only recently has deep, rest-frame, optical spectroscopy made robust metallicity and age measurements possible, allowing us to empirically assess their effects on continuum shapes. To this end, we measure ages and metallicities of ~ 700 massive ($10.2 \lesssim \log(M_*/M_\odot) \lesssim 11.8$) quiescent galaxies at $0.6 \lesssim z \lesssim 1.0$ from the Large Early Galaxy Astrophysics Census (LEGA-C) via continuum-normalized, absorption-line spectra, and compare with independent rest-frame $U - V$ and $V - J$ colours. Age increases along the quiescent sequence as both colours redden, consistent with stellar population synthesis (SPS) model predictions. Metallicity increases perpendicularly to the age trend, with higher metallicities at redder $U - V$ and bluer $V - J$ colours. Thus, age and metallicity behave differently in the UVJ diagram. Moreover, this trend conflicts with SPS model expectations of increasing metallicity approximately along the quiescent sequence. Independent dynamical mass-to-light ratio trends also differ dramatically from SPS model predictions. These results demonstrate that relying on model fits to continuum shapes alone may lead to systematic biases in ages, metallicities, and stellar masses. The cause of these data-model disparities may stem from non-solar abundance patterns in quiescent galaxies or the treatment of evolved stellar phases in the models. Resolving these discrepancies is crucial, as photometric data remain central even with *JWST*.

Chloe M. Cheng, Mariska Kriek, Aliza G. Beverage, Martje Slob, Rachel Bezanson, Marijn Franx, Joel Leja, Pavel E. Mancera Piña, Katherine A. Suess, Arjen van der Wel, Jesse van de Sande, and Pieter G. van Dokkum

2025, *Monthly Notices of the Royal Astronomical Society*, **540**, 1527

3.1 Introduction

Stellar populations of massive quiescent galaxies provide insight into the physical processes driving galaxy formation and evolution. In particular, tracing these populations over cosmic time is crucial for distinguishing different galaxy assembly and quenching mechanisms (e.g., Matteucci 1994; Trager et al. 2000; Conroy et al. 2014; Choi et al. 2014; Peng et al. 2015; Kriek et al. 2016; Kriek et al. 2019; Spitoni et al. 2017; Maiolino & Mannucci 2019; Trussler et al. 2020; Beverage et al. 2021, 2024, 2025; Cheng et al. 2024). These investigations subsequently provide constraints on cosmological simulations and theoretical models for galaxy formation and evolution (Maraston et al. 2006; Maiolino & Mannucci 2019).

A common approach to understanding quiescent stellar populations is to fit broad-band spectral energy distribution (SED) shapes with stellar population synthesis (SPS) models (Tinsley 1972; Searle et al. 1973; Larson & Tinsley 1978). Broad-band photometric data are relatively inexpensive to obtain and primarily reveal information about stellar mass-to-light (M/L) ratios, specific star-formation rates, and dust attenuation (e.g., Sawicki & Yee 1998; Bell & de Jong 2001; Papovich et al. 2001; Bell et al. 2003; Gallazzi & Bell 2009; Wuyts et al. 2009; Zibetti et al. 2009; Reddy et al. 2010, 2012; Leja et al. 2019b; Abdurro'uf et al. 2021; Johnson et al. 2021; see Conroy 2013 as well). They can also indicate ages and metallicities (e.g., Bell & de Jong 2000; MacArthur et al. 2004; Whitaker et al. 2010; Pacifici et al. 2016; Nersesian et al. 2024, 2025); however, these measurements are challenging due to well-known degeneracies between various stellar properties, including age, metallicity, star-formation history (SFH), and dust (Worthey 1994; Bell & de Jong 2001; Papovich et al. 2001; Bruzual & Charlot 2003; Gallazzi et al. 2005; Leja et al. 2017, 2019a,b).

While the degeneracy between age and dust can be broken with longer wavelength data (i.e. UVJ colours; Burgarella et al. 2005; Labbé et al. 2005; Wuyts et al. 2007; Leja et al. 2017), the age–metallicity degeneracy cannot reliably be broken with photometric data alone (Worthey 1994; Lee et al. 2007; Eminian et al. 2008; Conroy 2013; Nersesian et al. 2024, 2025). In this case, spectroscopic data can be used to disentangle and infer these quantities robustly. By fitting Lick indices (Burstein et al. 1984; Worthey et al. 1994) or the full spectrum (Conroy & van Dokkum 2012a; Conroy et al. 2018) with SPS models, one can directly quantify ages, metallicities, and elemental abundances. However, high-quality spectra are required to make these measurements. Thus, while this has been done extensively in local quiescent galaxies (e.g., Gonzalez et al. 1993; Caldwell & Rose 1998; Trager et al. 2000; Vazdekis 2001; Bruzual & Charlot 2003; Gallazzi et al. 2005; Thomas et al. 2005; Graves & Schiavon 2008; Choi et al. 2014; Conroy et al. 2014; Gu et al. 2018; Bernardi et al. 2023), these spectra were historically difficult to obtain beyond $z \sim 1$ as faint absorption features needed to measure ages and metallicities are shifted into the near-infrared. With developments allowing for ultra-deep ground- and space-based spectra, these analyses are now also being extended to high redshifts (up to $z \sim 3$), where massive quiescent galaxies are found to be metal-poor and α -enhanced compared to their low- z counterparts (e.g., Kriek et al. 2016; Kriek et al. 2019; Jafariyazani et al. 2020, 2025; Beverage

et al. 2021, 2023, 2024, 2025; Carnall et al. 2022; Zhuang et al. 2023).

Ideally, to maximize the information being considered when modelling stellar populations, one can fit photometric and spectroscopic data simultaneously (see e.g. Kriek et al. 2008b; Belli et al. 2019; Tacchella et al. 2022; Akhshik et al. 2023; Cappellari 2023; Kaushal et al. 2024; Slob et al. 2024; Nersesian et al. 2025). While powerful constraints on SFHs and, therefore, quenching mechanisms can be obtained by leveraging the complementary strengths of spectroscopy and photometry to break the degeneracies discussed above (Tacchella et al. 2022), models often struggle to reproduce both types of data simultaneously (e.g., Fumagalli et al. 2016; Carnall et al. 2024; Kriek et al. 2024; Slob et al. 2024). Additionally, different available SPS models can lead to different and inconsistent stellar population properties (e.g., Lee et al. 2007; Eminian et al. 2008; Muzzin et al. 2009; Conroy & Gunn 2010; Walcher et al. 2011; Pacifici et al. 2023; Whitler et al. 2023), in particular when including redder wavelength data for quiescent galaxies (van de Sande et al. 2015). These discrepancies may be due to the fact that there are a variety of uncertainties associated with SPS models, as different models make varying assumptions about stellar evolution and implement different spectral libraries, as well as to the choices that the user can make for the models and priors within the available parameter spaces (Walcher et al. 2011; Conroy 2013, see also Coelho et al. 2020; Byrne & Stanway 2023; Nersesian et al. 2024; Slob et al. 2024; Bellstedt & Robotham 2025). Thus, resulting galaxy properties may be uncertain and systematically biased. In combination with the modelling degeneracies discussed above, possible template mismatches are therefore concerning.

To overcome some of these issues and better understand the inconsistencies between both continuum shapes and absorption lines and different SPS models, we require a large sample of distant quiescent galaxies with high-quality (i) broadband photometry and (ii) age and metallicity measurements from deep, absorption-line spectra that are independent of the shape of the continuum. This is now possible with the Large Early Galaxy Astrophysics Census (LEGA-C; van der Wel et al. 2016, 2021; Straatman et al. 2018), a deep spectroscopic survey of galaxies at $0.6 \lesssim z \lesssim 1.0$ with UltraVISTA photometry (McCracken et al. 2012). In this work, we model the spectra of ~ 700 quiescent galaxies from LEGA-C with the ABSORPTION LINE FITTER (ALF, which continuum-normalizes spectra prior to fitting, Conroy & van Dokkum 2012a; Conroy et al. 2018). We combine rest-frame UVJ colours with independent, SSP-equivalent ages and metallicities to understand the relationships between these quantities as well as how our data compare to popular SPS models.

This paper is organized as follows: in Section 3.2 we summarize the LEGA-C data set and detail the full-spectrum fitting employed to derive galaxy ages and metallicities. In Section 3.3 we present our results. In particular, we discuss relationships between age and colour and metallicity and colour in Section 3.3.1. We also compare our results to predictions from two popular SPS models, the Flexible Stellar Population Synthesis models (FSPS; Conroy et al. 2009; Conroy & Gunn 2010) and the Bruzual & Charlot (2003) models, in Section 3.3.2. We compare our results with previous work, discuss the implications for galaxy evolution studies and SPS models, and outline some caveats of our work in Section 3.4. Finally, we

summarize and conclude in Section 3.5.

Throughout this work, we assume a flat Lambda cold dark matter (Λ CDM) cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$. All magnitudes are given in the AB-magnitude system (Oke & Gunn 1983). A Kroupa (2001) initial mass function is assumed throughout.

3.2 Data and methods

In this work, we use data from LEGA-C (van der Wel et al. 2021). We provide a brief summary of the survey and sample selection in Section 3.2.1 and direct the reader to van der Wel et al. (2021) and Cheng et al. (2024) (C24 hereafter) for more details. In Section 3.2.2, we summarize the methods used to derive stellar population ages and elemental abundances for each data set.

3.2.1 LEGA-C

We use spectroscopic data from the third data release of LEGA-C, a European Southern Observatory (ESO) Public Spectroscopic survey of 3600 galaxies at $0.6 \lesssim z \lesssim 1.0$. Galaxies were selected from the UltraVISTA (McCracken et al. 2012) K -band catalogue by Muzzin et al. (2013b) and are located in the Cosmic Evolution Survey (COSMOS) field (Scoville et al. 2007b). The data were collected using the VISible MultiObject Spectrograph (VIMOS) on the ESO *Very Large Telescope* (VLT) over 128 nights, providing deep (20-h integration), $R \sim 3500$ spectra with signal-to-noise (S/N) $\sim 20 \text{ \AA}^{-1}$ on average. The raw 2D and reduced 1D spectra are available on the ESO Science Archive Facility^{1,2}. See van der Wel et al. (2016, 2021), and Straatman et al. (2018) for details.

We use the LEGA-C sample selected in C24. We summarize the sample selection and stellar population parameter derivation here, but refer the reader to C24 for details. In particular, C24 determined rest-frame UVJ colours using EAZY (Brammer et al. 2008), via the Muzzin et al. (2013b) K -band catalogue³ with the spectroscopic redshifts fixed to those measured from the LEGA-C spectra (van der Wel et al. 2021), and selected quiescent galaxies by employing the classification from Muzzin et al. (2013a). C24 also removed galaxies with spectroscopic redshifts < 0.5 . Additionally, galaxies in LEGA-C's mask 2 were discarded as their noise spectra are significantly underestimated. A median rest-frame S/N $\gtrsim 20 \text{ \AA}^{-1}$ was required, as well as a maximum wavelength of at least 4450 \AA in order

¹http://archive.eso.org/eso/eso_archive_main.html .

²The reduced 1D spectra and catalogue have been released by ESO (<http://archive.eso.org/cms/eso-archive-news/Third-and-final-release-of-the-Large-Early-Galaxy-Census-LEGA-C-Spectroscopic-Public-Survey-published.html>). They are also available here: <https://users.ugent.be/~avdrwel/research.html#legac> .

³There is a known systematic offset in the Muzzin et al. (2013b) photometry due to zero-point corrections (van der Wel et al. 2021). This offset is corrected in the UltraVISTA DR3 catalogue which does not cover the entire LEGA-C field, and thus we use the original Muzzin et al. (2013b) catalogue. We test our analysis for the common galaxies (~ 290) in the two catalogues. The redetermined UVJ colours are consistent between both sets of photometry, and the use of the DR3 photometry does not change our conclusions.

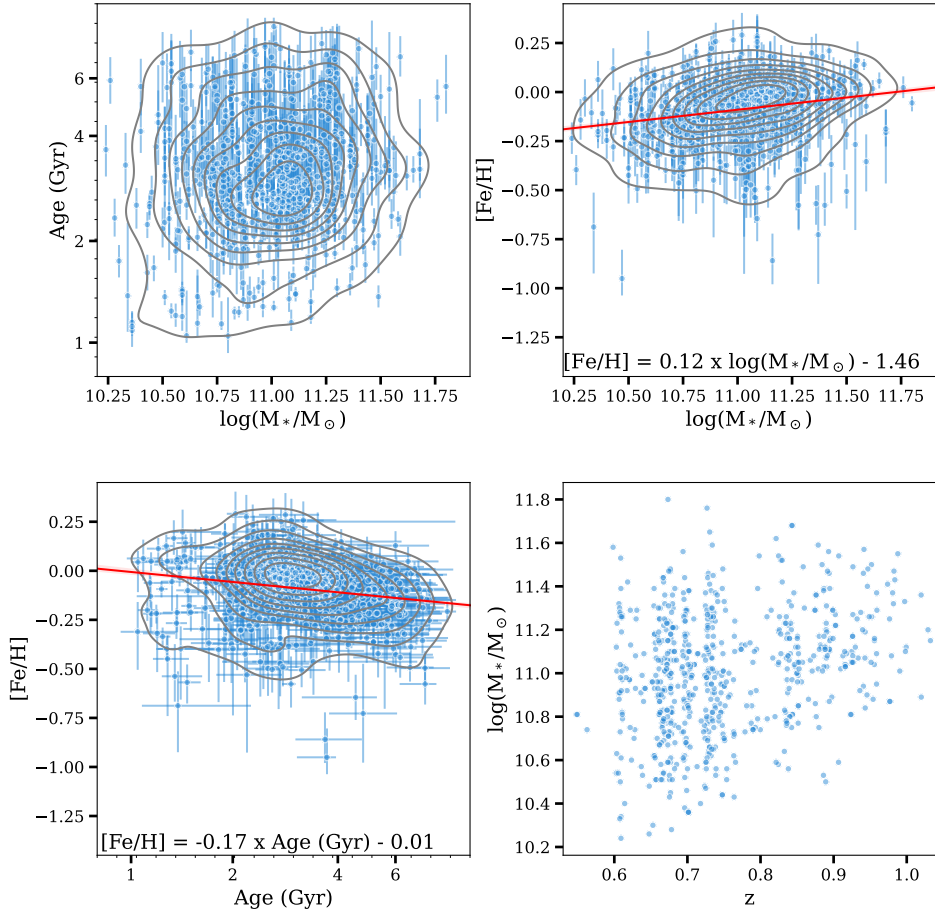


Figure 3.1: Properties of our final sample of massive quiescent galaxies from LEGA-C. In the top left panel, we show SSP-equivalent age as a function of stellar mass. In the top right panel, we show $[\text{Fe}/\text{H}]$ as a function of stellar mass. In the bottom left panel, we show $[\text{Fe}/\text{H}]$ as a function of age. In the bottom right panel, we show stellar mass as a function of redshift. We fit linear relations to the points in the top right and bottom left panels, respectively (solid lines with shaded uncertainties) and we state the corresponding best-fitting relations at the bottom of these panels. Stellar masses are measured by the LEGA-C team using MAGPHYS (da Cunha et al. 2008) and SSP-equivalent ages and $[\text{Fe}/\text{H}]$ abundances are derived from our ALF fits.

to sufficiently recover ages and metallicities. Note that the sample size in C24 was limited by the S/N of the spatially resolved spectra. However, this study is concerned with the full, integrated spectra. As such, we are able to add 244 more quiescent galaxies with $S/N \gtrsim 20 \text{ \AA}^{-1}$.

3.2.2 Spectral extraction and full-spectrum modelling

In this paper, we use the 1D spectra from C24, which were extracted from the 2D LEGA-C spectra as part of a spatially resolved analysis. In summary, an optimal extraction procedure was performed to obtain spatially resolved and integrated spectra, by fitting the spectral flux profiles with Moffat (1969) profiles. To extract the integrated spectra, these fits were used to perform weighted sums of all rows with significant flux (i.e. excluding (sub-)rows outside of the 3rd and 97th percentiles of the Moffat profiles to reduce noise, with the sums weighted by the Moffat profiles). For consistency, we perform the same extraction of the integrated spectra for the additional 244 galaxies in our sample. See C24 for details.

C24 derived stellar population parameters from the LEGA-C spectra using the ABSORPTION LINE FITTER (ALF⁴; Conroy & van Dokkum 2012a; Conroy et al. 2018). Empirical simple stellar populations (SSPs) constructed using the Mesa Isochrones and Stellar Tracks (MIST; Choi et al. 2016) and the Spectral Polynomial Interpolator (SPI; Villaume et al. 2017b)⁵ provide the foundation for the ALF models. ALF employs the Medium Resolution INT Library Of Empirical Spectra (MILES; Sánchez-Blázquez et al. 2006), the Extended Infrared Telescope Facility stellar library (E-IRTF; Villaume et al. 2017b), and a large sample of M-dwarf spectra (Mann et al. 2015) with SPI to generate stellar spectra as a function of effective temperature (T_{eff}), surface gravity, and metallicity from a data-driven model.

The empirical parameter space is set by the combination of the E-IRTF and Mann et al. (2015) samples. It spans $-2.0 \lesssim [\text{Fe}/\text{H}] \lesssim 0.5$ and $3.5 \lesssim \log(T_{\text{eff}}/\text{K}) \lesssim 3.9$. To ensure high-quality interpolation at the boundaries of the empirical parameter space, ALF also makes use of a theoretical stellar library (C3K; see Conroy & van Dokkum 2012a). By differentially including theoretical element response functions, ALF allows for variable abundance patterns. The elemental abundances of the base stellar population are estimated in Milone et al. (2011) and used in conjunction with the differential response functions to estimate the absolute abundances.

During the fitting procedure, ALF removes the continuum from the observations by fitting a high-order Chebyshev polynomial to the ratio of the data to the model. It then implements a Fortran version of the Markov chain Monte Carlo (MCMC) algorithm `emcee` (Foreman-Mackey et al. 2013) to sample the posteriors of 46 stellar population parameters. In particular, it allows for arbitrary variation in stellar population age and detailed elemental abundance patterns. To characterize observed errors, ALF also fits for systematic parameters. Note that ALF fits spectra between 3700 and 24000 Å and stellar populations older than 1 Gyr. For details, see Conroy & van Dokkum (2012a) and Conroy et al. (2018).

C24 fit each spectrum with 1024 walkers, 20000 burn-in steps, and a 1000-step production run. For each fit, 500 MCMC chains were examined. A Kroupa (2001) initial mass function (IMF) was assumed and a single stellar population age was fit. To initialize the age of each galaxy and to avoid the ALF walkers getting

⁴<https://github.com/cconroy20/alf> .

⁵https://github.com/AlexaVillaume/SPI_Utils .

trapped at an unrealistically high initial age, a random value was drawn from a uniform distribution centred at 3 Gyr. Additionally, the upper limit of the age prior was set to be the age of the Universe at the redshift of each galaxy, plus 2 Gyr to allow for uncertainties. The hot star component was not fit. Variation in all stellar parameters in ALF was allowed. Galaxies for which ALF fit an age < 1 Gyr were discarded (5 galaxies), as the abundances are extrapolated for ages younger than 1 Gyr and are unreliable (Conroy et al. 2018). Furthermore, the fits for galaxies where the posteriors ran up against the age, $[Z/H]$, and $[Fe/H]$ priors were discarded as these fits would result in unreliable ages and metallicities (~ 6 per cent of the sample). There are a variety of reasons why the parameter space may not be fully explored for these objects, for example low S/N, low levels of residual star formation, or non-stellar absorption. We fit the 244 galaxies that we add to this sample in exactly the same way as C24. Finally, we visually inspect all of the spectra and remove any objects near the edge of the UVJ box with strong $[O_{III}]$ emission (10 galaxies), which could potentially host active galactic nuclei or have low levels of star formation. The final LEGA-C sample examined in this study includes 683 quiescent galaxies.

We show the properties of our final sample in Fig. 3.1. In the top panels, we show, respectively, SSP-equivalent ages and $[Fe/H]$ abundances versus stellar masses (computed using MAGPHYS, da Cunha et al. 2008, and obtained via private communication with the LEGA-C team). In the bottom-left panel, we show our SSP-equivalent ages versus SSP-equivalent $[Fe/H]$ abundances. In the bottom right panel, we show stellar masses versus spectroscopic redshifts. We remove two galaxies with MAGPHYS masses $< 10^{10} M_{\odot}$ as the flux levels of their spectra are abnormally low. Overall, the galaxies in our sample are at $0.6 \lesssim z \lesssim 1.0$ with $10.2 \lesssim \log(M_*/M_{\odot}) \lesssim 11.8$ and $-1.3 \lesssim [Fe/H] \lesssim 0.3$, and are $1.0 - 8.6$ Gyr old.

A subset of 135 quiescent LEGA-C galaxies were fit using ALF in Beverage et al. (2023). They fit the 1D extracted spectra from the LEGA-C public release. We find that the stellar population parameters for the galaxies that we have in common (110 galaxies) are fully consistent between the two studies, despite subtle differences in the spectral extraction procedures (not shown). We also qualitatively recover the negative relation between $[Fe/H]$ and age (Beverage et al. 2021) and the slightly positive relation between $[Fe/H]$ and mass (Beverage et al. 2023), as shown by the red lines in Fig. 3.1. We state the linear fits in the corresponding panels in Fig. 3.1.

3.3 Results

In this paper we compare detailed stellar population properties derived from absorption lines to colours derived from broad SED shapes. In Section 3.3.1, we explore how age and metallicity vary on the UVJ diagram. In Section 3.3.2, we compare our results with predictions from popular SPS models.

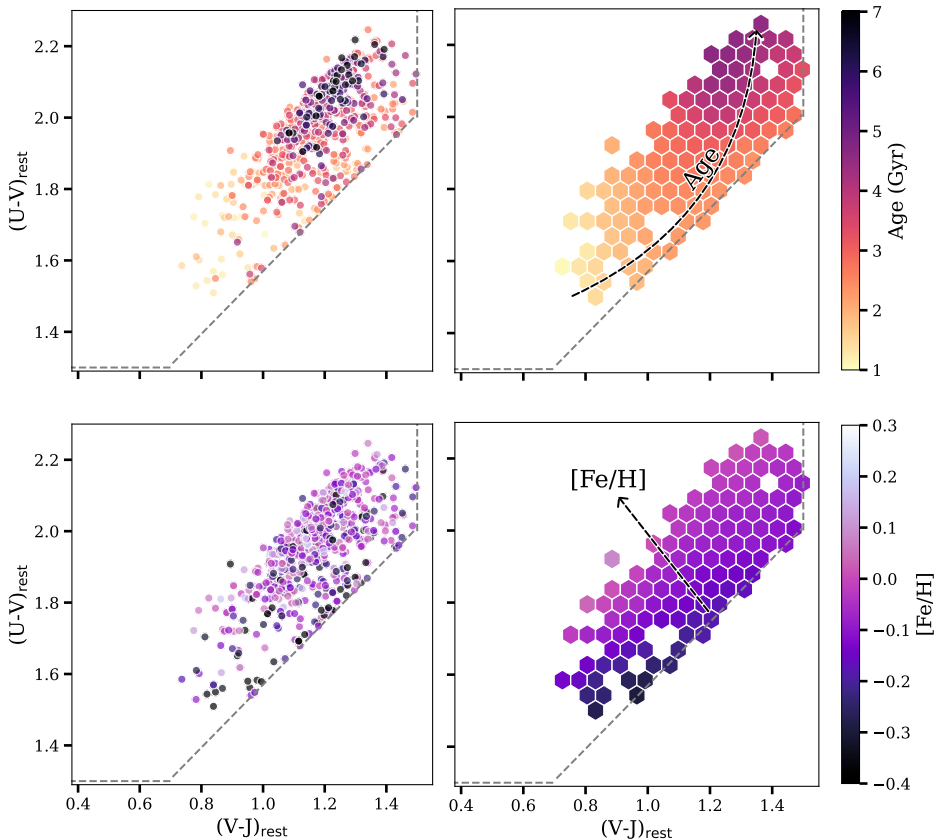


Figure 3.2: Rest-frame UVJ diagrams of quiescent galaxies at $0.6 \lesssim z \lesssim 1.0$ from the LEGA-C survey. *Left:* In these panels, we colour code the symbols by their measured ages (top) and their measured $[\text{Fe}/\text{H}]$ values (bottom), as derived by fitting absorption lines in continuum-normalized spectra with ALF. *Right:* In these panels, we LOESS-smooth the ages and metallicities from the left panels (Cappellari et al. 2013a; Cleveland & Devlin 1988). In each panel, we plot an arrow showing the general, qualitative direction of the dependence between age and colour (top) and metallicity and colour (bottom).

3.3.1 Trends with age and metallicity

In Fig. 3.2, we show several rest-frame UVJ diagrams colour coded by age and metallicity. In the left panels, we present the results for the individual galaxies in our LEGA-C sample. The symbols are colour coded by their measured SSP-equivalent ages (top) and $[\text{Fe}/\text{H}]$ values (bottom), both measured via absorption lines in continuum-normalized spectra (Section 3.2.2). In the right panels, we smooth the data using the Locally Weighted Regression (LOESS) 2D algorithm (Cleveland & Devlin 1988) via the LOESS Python package (Cappellari et al. 2013a).

There is a strong sequence in age in the top panels. In particular, age increases approximately parallel to the quiescent sequence, from ~ 1 Gyr in the bluest

regions of the UVJ diagram to $\sim 3 - 5$ Gyr in the reddest regions. There is some curvature to this trend, with the oldest galaxies populating the reddest end of the $U - V$ axis, but located slightly blueward on the $V - J$ axis. Interestingly, $[\text{Fe}/\text{H}]$ behaves differently, increasing approximately perpendicular to the quiescent sequence, from subsolar ($-0.3 \lesssim [\text{Fe}/\text{H}] \lesssim -0.2$) near the edge of the UVJ box to supersolar ($[\text{Fe}/\text{H}] \sim 0.08$) as one moves across the quiescent sequence. We indicate the approximate, qualitative trends that we describe here with arrows in the right panels.

This observed age trend has been suggested in previous work (e.g., Whitaker et al. 2012, 2013; Mendel et al. 2015; Belli et al. 2019; Díaz-García et al. 2019). However, in these studies, galaxy ages were derived either via the broad-band photometric SEDs alone, or by including the SEDs in simultaneous fits to spectra and photometry. In contrast, our ages are completely independent of continuum shape as they are derived directly from the continuum-normalized LEGA-C spectra.

Moreover, for the first time, we observe a trend in metallicity as derived from absorption lines. While ages and metallicities are known to be degenerate (Worthey 1994; Bruzual & Charlot 2003; Gallazzi et al. 2005), previous studies that examined the relationship between ages and broad-band colours (e.g., Whitaker et al. 2012, 2013; Mendel et al. 2015; Belli et al. 2019; Díaz-García et al. 2019) did not take metallicities into account⁶ as they were poorly constrained for distant quiescent galaxies. However, it is now possible to measure robust metallicities beyond the local Universe and disentangle them from ages using a combination of deep LEGA-C spectra and ALF (Conroy & van Dokkum 2012b; Choi et al. 2014; Beverage et al. 2021, 2023; Gu et al. 2022; Cheng et al. 2024). As our measured age and $[\text{Fe}/\text{H}]$ trends are distinct, this demonstrates that observed reddening along the quiescent sequence is indeed primarily due to an increase in age, as previously suggested (e.g., Whitaker et al. 2012, 2013; Mendel et al. 2015; Belli et al. 2019; Díaz-García et al. 2019). Furthermore, the differing trends may indicate that age and metallicity are not strongly degenerate in this UVJ space. However, this does not imply that we can simply derive metallicities and ages directly from UVJ colours as there are many other fitting degeneracies that must be considered. Thus, we do not attempt to quantify the strength of the age and metallicity trends here, but simply state that qualitative trends exist. We discuss this further in Section 3.4.1.

3.3.2 Comparison to stellar population models

Our current understanding of stellar populations in unresolved systems from both broad-band photometry and high-quality spectroscopy primarily relies on SPS models (Tinsley 1972; Searle et al. 1973; Larson & Tinsley 1978). Building these models involves combining isochrones, spectral libraries, and the IMF to construct simple stellar populations (SSPs). Every model implements these components in slightly different ways. Additionally, there are several aspects of population synthesis that still contain uncertainties (Renzini 2006; Walcher et al. 2011; Conroy

⁶Note that Díaz-García et al. (2019) examined trends between UVJ colours and metallicity, however, they derived total metallicity from photometry alone and used a different SPS model.

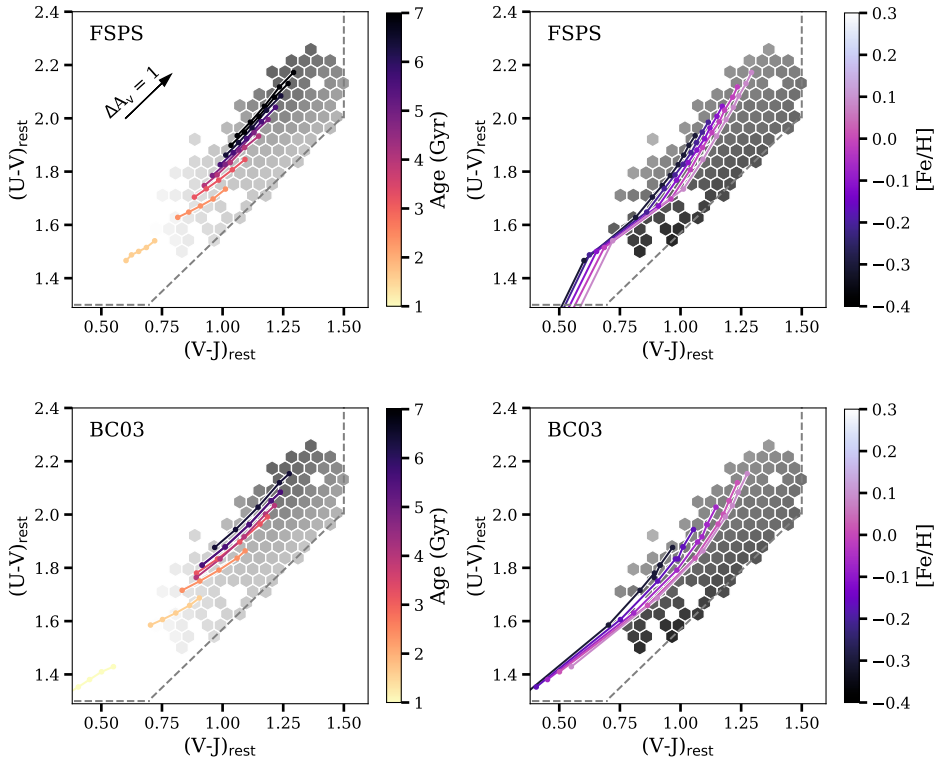


Figure 3.3: The same LOESS-smoothed rest-frame UVJ diagrams as in the right panels of Fig. 3.2 (grey) but with stellar evolution tracks overplotted. Tracks produced using PYTHON-FSPS (Johnson et al. 2023, see also Conroy et al. 2009; Conroy & Gunn 2010) are shown in the top panels and tracks produced using BAGPIPES (Carnall et al. 2018, which uses the Bruzual & Charlot 2003, or BC03, SPS models) are shown in the bottom panels. For each model, we produce tracks spanning the age ($\sim 1 - 8$ Gyr) and $[\text{Fe}/\text{H}]$ ($-0.31 < [\text{Fe}/\text{H}]/\text{dex} < 0.08$) ranges of our sample. The tracks in the left and the right panels are the same tracks for each model. On the left, we emphasize the age trend and the ages are indicated by colour-coded points, connected by lines to guide the eye. On the right, we emphasize the $[\text{Fe}/\text{H}]$ trend, with metallicities indicated by colour-coded points connected by lines. In general, we see the same kind of sequence in age in the stellar evolution tracks from both models that we see in our data (left panels). On the other hand, the metallicity sequence of the tracks seems to be perpendicular to that observed in the LEGA-C data (right panels).

2013). For example, all models must make approximations for processes such as convection, stellar rotation, mass loss, binary stars, and advanced stages of stellar evolution (e.g. thermally pulsating asymptotic giant branch, or TP-AGB, stars). Many codes treat these processes differently (compare the models of Bruzual & Charlot 2003 with those of Maraston 2005, for instance, and see Muzzin et al. 2009). Additionally, spectral libraries are often incomplete (Coelho et al. 2020;

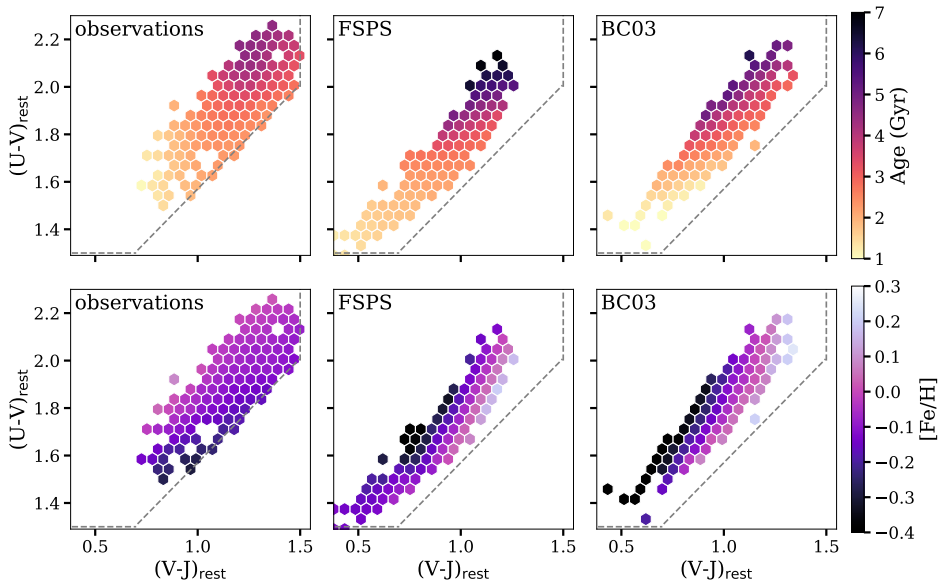


Figure 3.4: LOESS-smoothed rest-frame UVJ diagrams for the observations and the models. For the models, we input the ages and metallicities of each galaxy in our LEGA-C sample and output the predicted rest-frame UVJ colours, with some random scatter added to simulate observational uncertainties. The diagrams are colour coded by age in the top panels and by $[\text{Fe}/\text{H}]$ in the bottom panels. The same diagrams as in the right panels of Fig. 3.2 are shown for the data in the left panels, the diagrams for the FSPS models are shown in the middle panels, and the diagrams for the BC03 models are shown in the right panels. Age follows the quiescent sequence in all three panels. However, the trend in $[\text{Fe}/\text{H}]$ is inconsistent between the data and the models.

Byrne & Stanway 2023; Bellstedt & Robotham 2025). Thus, different SPS models can produce inconsistent results (see, e.g., Lee et al. 2007; Eminian et al. 2008; Muzzin et al. 2009; Conroy & Gunn 2010; Walcher et al. 2011; van de Sande et al. 2015; Coelho et al. 2020; Byrne & Stanway 2023). Moreover, they are often not able to reliably reproduce observations, either from photometry, spectra, or combined photometric and spectroscopic data, and in cases where models can reproduce the observations, the physics implied by each model can be conflicting (e.g., Lee et al. 2007; Fumagalli et al. 2016; Byrne & Stanway 2023; Cappellari 2023; Pacifici et al. 2023; Kaushal et al. 2024; Nersesian et al. 2024; Slob et al. 2024; Tortorelli et al. 2024; Bevacqua et al. 2025; Kim et al. 2025).

We note that several studies have fit the LEGA-C sample with different models (e.g., Chauke et al. 2018; Beverage et al. 2021, 2023; Barone et al. 2022; Cappellari 2023; Kaushal et al. 2024; Nersesian et al. 2024). However, in this work, we have access to both photometric colours from broad-band SEDs and completely independent measurements of ages and metallicities from continuum-normalized, absorption-line spectra. Thus, we are in a position to compare our results to predictions from popular SPS models to better understand the uncertainties and

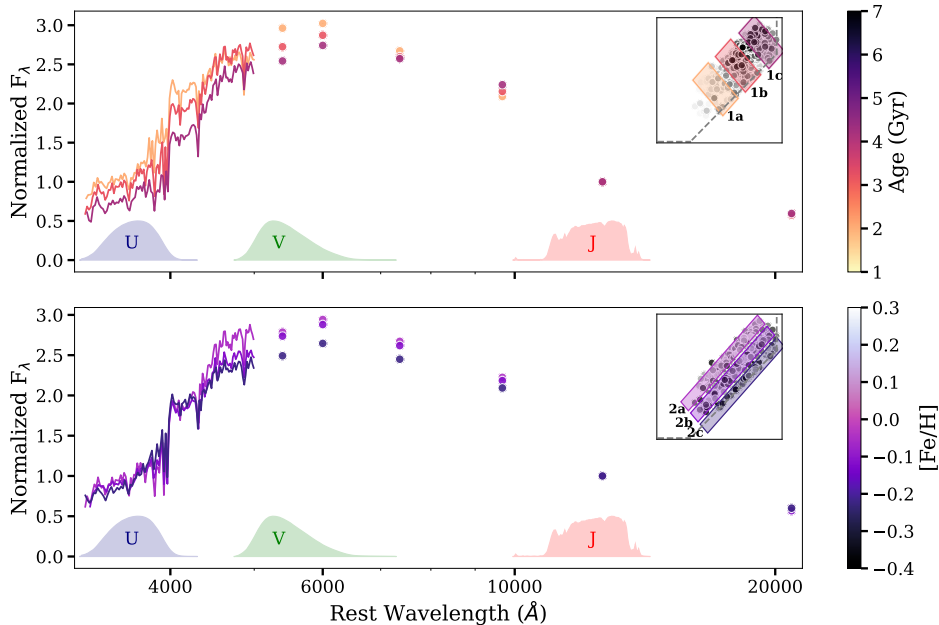


Figure 3.5: A comparison of stacked LEGA-C spectra. In the top panel, we select three bins along the quiescent sequence and approximately along our observed age sequence. In the bottom panel, we select three bins perpendicular to the quiescent sequence and approximately along our observed metallicity sequence. We normalize the spectra over 4400 – 4500 Å and mean-stack the spectra in each bin. We renormalize each stack by the photometric rest-frame J band. We also normalize and mean-stack the UltraVISTA photometry beyond 5000 Å for the same galaxies in each bin in the same way. We colour code the stacked spectra by the median age or metallicity in each bin and median-bin the spectra in bins of 50 pixels for visibility. We state each curve’s median age, $[\text{Fe}/\text{H}]$, sSFR averaged over the last 0.01 Gyr (measured via MAGPHYS by the LEGA-C team), and the number of galaxies in each stack in Table 3.1. The UVJ filter response curves are indicated in each panel. In the inset panels, we show the same rest-frame UVJ diagrams as in the left panels of Fig. 3.2 but greyscaled, with the bins indicated by the shaded regions. This figure serves to illustrate our results, as in the top panel the difference in ages in the three stacked spectra (along the observed age sequence) is greater than in the bottom panel (along the observed $[\text{Fe}/\text{H}]$ sequence), while the opposite is true for the difference in $[\text{Fe}/\text{H}]$.

discrepancies we discussed above. To facilitate this comparison, in Fig. 3.3 we overplot stellar evolution tracks generated using SPS models on the same LOESS-smoothed UVJ -diagrams as in Fig. 3.2 but with the hexbins greyscaled. In the top panels, we show tracks computed using PYTHON-FSPS (Johnson et al. 2023, and see also Conroy et al. 2009; Conroy & Gunn 2010), implemented with the MIST isochrones (Choi et al. 2016; Dotter 2016) and MILES spectral library (Sánchez-Bláquez et al. 2006; Falcón-Barroso et al. 2011). In the bottom panels, we similarly show tracks computed using BAGPIPES (Carnall et al. 2018), which makes use

Table 3.1: Median age, [Fe/H], sSFR over 0.01 Gyr, and number of galaxies in each stack for the stacked spectra in Fig. 3.5.

Box	Age (Gyr)	[Fe/H]	sSFR (0.01 Gyr) (M_{\odot}/yr)	Number of galaxies
1a	2.01	-0.14	-10.99	48
1b	3.19	-0.09	-11.37	297
1c	4.25	-0.07	-11.46	131
2a	3.66	-0.05	-11.58	297
2b	3.22	-0.12	-11.21	185
2c	2.67	-0.22	-10.73	62

of the 2016 version of the Bruzual & Charlot (2003) SPS models (BC03 hereafter), implemented with the Padova isochrones (Alongi et al. 1993; Bressan et al. 1993; Fagotto et al. 1994a,b; Girardi et al. 1996) and MILES spectral library. In both cases, we model a stellar population with a τ SFH truncated after 1 Gyr and no dust (i.e., $A_V = 0$, see Section 3.4.4 for a discussion of the impacts of our assumed SFH and A_V). We indicate with an arrow the direction in which dust would move the tracks according to the Calzetti et al. (2000) attenuation law. From the models, we compute UVJ colours for a grid of ages and metallicities, spanning $\sim 1 - 8$ Gyr (the age of the Universe at $z \sim 0.6$) and the metallicity range of our sample ($-0.31 < [\text{Fe}/\text{H}]/\text{dex} < 0.08$). We colour code the tracks by stellar population age in the left panels, where we connect points of the same age by lines to guide the eye. We colour code the tracks by metallicity in the right panels, where we assume $[\text{Fe}/\text{H}] \approx \log(Z/Z_{\odot})$ (see Beverage et al. 2025 and Section 3.4.3).

In the left panels of Fig. 3.3, we see that both models have age trends that are qualitatively similar to those seen in our observations. In particular, ages generally increase along the quiescent sequence, with the trend steepening at the reddest tip of the quiescent sequence. This comparison indicates that SPS models predict similar UVJ colours as a function of age to what we observe for the LEGA-C galaxies.

However, the same cannot be said when we examine the metallicity panels (right panels of Fig. 3.3). Here, both models do not follow the trends that we see in our observations, and perhaps even show a sequence in the perpendicular direction (i.e. while our metallicities increase approximately perpendicular to the quiescent sequence, the metallicities of the models seem to increase more along the quiescent sequence). This finding suggests that models like FSPS and BC03 may not accurately predict broad-band colours for varying metallicities.

We note that the model tracks in Fig. 3.3 only span a small range in the UVJ quiescent box and that the real galaxies (grey hexagons) show a much broader distribution. This is likely due to photometric uncertainties, differences between the real SFH and the assumed τ -model SFH, differences between the real dust properties and the assumed dust extinction ($A_V = 0$), or template mismatches. To assess the significance of our results and evaluate whether we expect to be able to observe the metallicity trend that we see in the UVJ diagram, we generate

the predicted UVJ colours for each galaxy in our sample assuming the FSPS and BC03 models. In particular, we input the measured SSP-equivalent ages and metallicities of our sample and output rest-frame UVJ colours for each galaxy. Uncertainties on the rest-frame fluxes are estimated using the median S/N of the observed photometric data points (between 1 and 7 bands, depending on the filter) which are close in rest-frame wavelength. For the U , V , and J bands, we find typical uncertainties of ~ 3.5 , ~ 1.0 , and ~ 1.0 per cent, respectively. We then add random scatter to the model fluxes and use these to calculate realistic U , V , and J colours by randomly sampling from a normal distribution multiplied by each of these typical observed relative errors.

We plot the LOESS-smoothed UVJ diagrams colour coded by age and metallicity for the observations, the FSPS models, and the BC03 models in Fig. 3.4. In the top panels, the age trend is qualitatively consistent between the observations and the models. In particular, ages always increase along the quiescent sequence, with some curvature near the top right corner of the UVJ box. In the bottom panels, however, the trend in $[\text{Fe}/\text{H}]$ is not consistent between the observations and the models, with the trend being nearly perpendicular between the two, consistent with what we see in Fig. 3.3. This test indicates that we can realistically expect to detect the trend with metallicity in our data. Thus, our result is significant. Additionally, even if we are underestimating the typical relative observed errors on the UVJ colours as described above, we find that we are still able to recover the trends with typical uncertainties of up to 5 per cent for each of the U , V , and J bands (this percentage allows the model panels to approximately display the same spread in UVJ colours as the observations). We discuss possible additional sources of scatter in Section 3.4.4.

Finally, in Fig. 3.5, we visualize how the observed SEDs and spectra of galaxies change along and perpendicular to the quiescent sequence, in roughly similar directions to those in which age and metallicity increase, respectively. In the top panel, we select three bins along the quiescent sequence (i.e. approximately along the age sequence). In the bottom panel, we select three bins perpendicular to the quiescent sequence (i.e. approximately along the metallicity sequence). These bins are shown by the shaded rectangles on the rest-frame UVJ diagrams in the inset panels. We normalize the spectra over the region from 4400 to 4500 Å and mean-stack the spectra in each bin. We then normalize each stack by the UltraVISTA photometry in the rest-frame J band. We also normalize and mean-stack the UltraVISTA photometry beyond 5100 Å for the same galaxies in each bin in the same way. In particular, as the galaxies in our sample are at similar redshifts, we mean-stack the photometry in the observed frame and de-redshift the median wavelength of each band when plotting the points in Fig. 3.5. In the top panel, we colour code the stacked spectra and photometry by the median age in each bin and in the bottom panel we colour code by the median metallicity in each bin. We median-bin the spectra in bins of 50 pixels for ease of visualization. The UVJ filter response curves are also indicated in each panel. We state the median age, $[\text{Fe}/\text{H}]$, specific star-formation rate (sSFR) over 0.01 Gyr (computed using MAGPHYS, da Cunha et al. 2008, and obtained via private communication with the LEGA-C team), and the number of galaxies in each stack in Table 3.1.

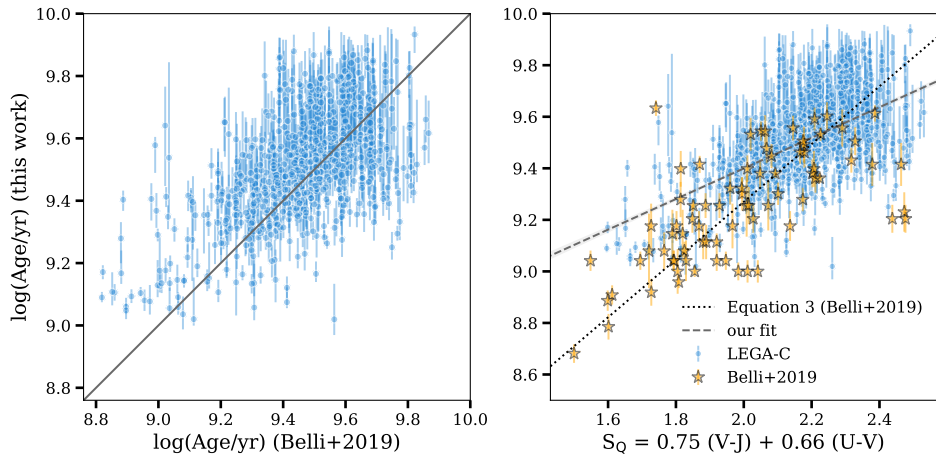


Figure 3.6: *Left:* A comparison between the ages that we measure by using ALF to fit absorption lines in continuum-normalized spectra (y -axis) to those derived using equation (3) in B19 (i.e. using only the UVJ colours, x -axis). The one-to-one relation is represented by the solid line. In general, we find older ages compared to those predicted from the B19 age–colour relation. However, overall, the ages of the LEGA-C galaxies are relatively consistent with the predicted ages from B19. *Right:* SSP-equivalent ages plotted as a function of the S_Q axis defined in equation (2) in B19. We also reproduce the points from fig. (8) in B19 (orange stars). The dotted line is the age–colour relation from B19. The dashed line shows our illustrative linear fit to the points in this plot, with shaded regions indicating the 1σ uncertainties. Our fit differs qualitatively from the age–colour relation derived in B19. Furthermore, there is a lot of scatter that is not captured by a linear relation on these axes.

This figure serves as an illustration of our results. In particular, the difference in ages in the three stacked spectra in the top panel (~ 1 Gyr on average) is greater than in the bottom panel (~ 0.6 Gyr on average). On the other hand, the difference in $[\text{Fe}/\text{H}]$ values in the top panel (~ 0.05 dex on average) is smaller than in the bottom panel (~ 0.1 dex on average). Thus, the selected bins in the top panel are tracing the age sequence that we observe with very little dependence on $[\text{Fe}/\text{H}]$. When we examine the combined spectra and SEDs, we see that age is primarily shifting the spectra to redder wavelengths and making the 4000 \AA break more pronounced. Meanwhile, the selected bins in the bottom panel are tracing the observed $[\text{Fe}/\text{H}]$ sequence with a smaller dependence on age compared to the top panel. Metallicity is not shifting the spectra in the same way as age in the top panel (i.e. the spectra all have approximately the same peak wavelength in different metallicity bins), but it acts to make the galaxies redder in $U - V$. However, as a caveat we note that this is not a perfect comparison as metallicity and age are changing in both panels, since the observed age and metallicity trends are not exactly along and perpendicular to the quiescent sequence, respectively. The observed shifts in the spectra here may also be affected by other degenerate effects, such as dust, SFH, or sSFR. See Section 3.4.4 for details.

3.4 Discussion

In Section 3.3, we examined trends between spectroscopic ages and metallicities and independent UVJ colours. We found that observed reddening along the quiescent sequence in the UVJ diagram is primarily due to age, while metallicity increases approximately perpendicular to the quiescent sequence (Section 3.3.1). We also investigated whether these trends can be reproduced by popular SPS models. While these models predict similar UVJ colours as a function of age compared to what we see in the data, this is not the case for metallicity (Section 3.3.2).

Here, we discuss the implications of our results. In Section 3.4.1, we compare our results to previous studies, in particular Belli et al. (2019). We discuss the implications for galaxy evolution studies in general, and for derived stellar M/L ratios and masses specifically, in Section 3.4.2. We examine the implications of our results for SPS models in Section 3.4.3. Finally, in Section 3.4.4, we discuss caveats of our work that should be taken into account.

3.4.1 Comparison to previous work

In Section 3.3.1, we used independent measurements of age and UVJ colours to demonstrate that reddening along the quiescent sequence is primarily due to age, as previously suggested by observations (e.g., Whitaker et al. 2012, 2013; Mendel et al. 2015; Belli et al. 2019; Díaz-García et al. 2019) and cosmological simulations (e.g., Akins et al. 2022). In Belli et al. (2019) (B19 hereafter), this trend was used to derive a relationship between ages in quiescent galaxies at $1.0 < z < 2.5$ and UVJ colours (see their equation (3))⁷. This relation has since been used to predict quiescent galaxy ages using only UVJ colours (e.g., Suess et al. 2019b; Clausen et al. 2025). However, in B19 and other previous studies, the broad-band photometric SEDs were included when deriving galaxy ages. In contrast, our ages are completely independent of continuum shape as they are derived directly from the continuum-normalized LEGA-C spectra, by only fitting absorption lines using ALF. Thus, this study provides an independent check of the age trend along the quiescent sequence in the UVJ diagram.

In the left panel of Fig. 3.6, we compare our ages measured from the LEGA-C spectra using ALF to those derived using equation (3) in B19 (i.e. predicted from UVJ colours). In general, we find older ages compared to those predicted from the B19 age–colour relation. This difference is somewhat surprising as we measure SSP-equivalent ages while B19 considers mass-weighted ages, and SSP-equivalent ages are typically found to be younger than mass-weighted ages (Trager & Somerville 2009). However, the difference is small, and the overall ages that we find for the LEGA-C galaxies are relatively consistent with the B19 predicted ages, even though we examine a different redshift range. Thus, the B19 relation is able to predict approximate quiescent galaxy ages down to $z \sim 0.6$.

⁷Díaz-García et al. (2019) derived similar relationships between stellar population parameters (including ages and metallicities) and UVJ colours. However, they examined photometric data only, in bands offset from U , V , and J . Thus, we examine only the relationship from Belli et al. (2019) here.

While we confirm that the B19 relation can be used in a general sense, we find that it may be too simplistic to derive precise ages for *individual* galaxies. In the right panel of Fig. 3.6, we plot our SSP-equivalent ages as a function of the $S_Q = 0.75(V - J) + 0.66(U - V)$ axis defined in B19 (see their equation (2)). We overplot the age–colour relation derived by B19 (dotted black line, see their equation (3)) as well as their individual data points (orange stars). As an illustrative comparison, we fit a similar linear relation to our galaxies on these axes, shown by the dashed grey line, with shaded regions indicating the 1σ uncertainty on our fit. Our fit differs from the age–colour relation in B19. In particular, we find a qualitatively shallower slope. Moreover, there is a significant amount of scatter that is not captured well by either fit, likely due in part to the fact that B19 have younger galaxies (ALF cannot fit galaxies younger than 1 Gyr, see Section 3.2.2), and because the relationship between UVJ colours and galaxy ages is not actually linear (see the curvature in the age trend in the top right panel of Fig. 3.2). Thus, we do not give the functional form of our linear fit as we do not believe that it accurately represents a quantitative relationship between UVJ colours and galaxy ages. We also apply multiple linear and polynomial regression algorithms to attempt to derive quantitative relationships between UVJ colours and our metallicities and ages (not shown). Specifically, we attempt to find the best linear or polynomial combination of $U - V$ and $V - J$ colours that minimizes the residuals in age and metallicity, similar to the S_Q axis defined in Belli et al. (2019). However, we find that no simple functional form fits the data well and captures the relationships between UVJ colours and ages and metallicities.

Accordingly, we suggest that the UVJ diagram can, as found previously, be used to identify whether a quiescent galaxy is relatively older or younger, but it may not be possible to specify a precise value of age based on an individual galaxy’s UVJ colours. Additionally, we show that the UVJ diagram can also be used to estimate whether a quiescent galaxy is relatively more metal-rich or metal-poor, though the scatter in metallicity is larger than in age (and again, it is difficult to specify an exact metallicity based on UVJ colours alone, see also Leja et al. 2019c). Other fitting degeneracies discussed in Section 3.1 could be complicating these relationships, for example degeneracies between age, dust, and SFH (see Section 3.4.3 for further discussion). Looking forward, more advanced fitting techniques, such as random forest algorithms or neural networks, in addition to combining photometry and spectra and developing SPS models further, may be able to provide a quantitative translation between UVJ colours and ages and metallicities for individual galaxies.

3.4.2 Implications for galaxy evolution studies

In Section 3.3.2, we examined how trends between our spectroscopically measured ages and metallicities and independent photometric UVJ colours compare to predictions from popular SPS models. We found that, while the age trend is generally reproduced well by both FSPS and BC03, these models do not recover the metallicity trend. Here, we discuss the implications of our results for galaxy evolution studies in general.

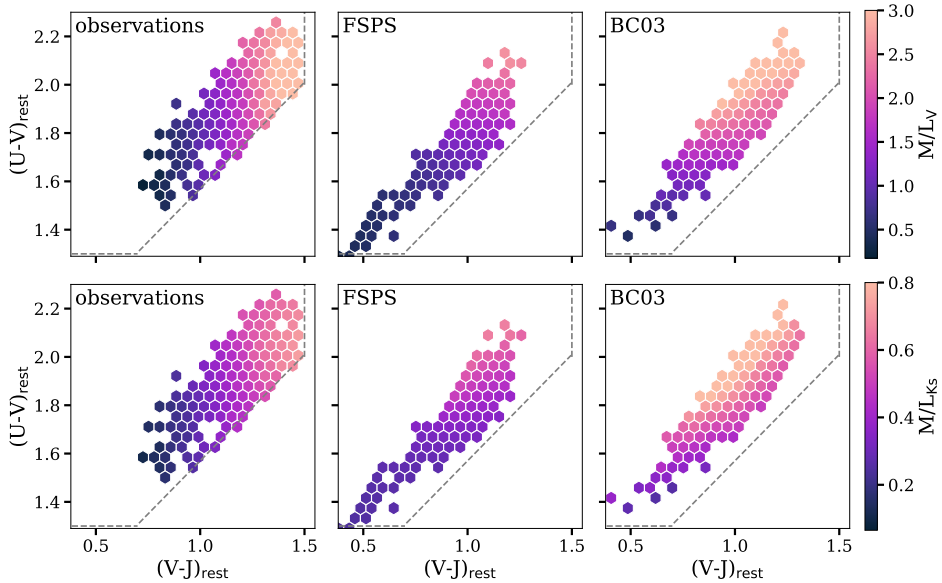


Figure 3.7: LOESS-smoothed rest-frame UVJ diagrams for the data and the models, where the diagrams are colour coded by the M/L ratio. The LEGA-C UVJ diagrams are shown in the left column, the FSPS diagrams are shown in the middle column, and the BC03 diagrams are shown in the right column. This is similar to Fig. 3.4, where we input the ages and metallicities of each galaxy in our LEGA-C sample into the models, and output the predicted rest-frame UVJ colours (from which we compute luminosities) and stellar masses. For the LEGA-C sample, we compute M/L ratios using independently measured, dynamically scaled virial masses from van der Wel et al. (2022). The top row is colour coded by M/L ratios in the V band (M/L_V) and the bottom row is colour coded by M/L ratios in the K_s band (M/L_{K_s}). While the M/L ratio increases approximately along the quiescent sequence in most cases, the trend curves in opposite directions for the data and the models at the reddest tip of the quiescent sequence. Additionally, the trends in the M/L ratios predicted by the two models are not consistent with each other.

SPS models have been used for decades to fit photometric and spectroscopic data of quiescent galaxies and interpret their formation and evolution. However, if these models cannot fully reproduce observations as we suggest here, we may need to re-evaluate certain aspects of our understanding of galaxy evolution. For example, while the age trends that we find in this work are qualitatively consistent with those predicted by the models, individual age measurements may not be reliable due to degeneracies with other fitting parameters (Worthey 1994; Bell & de Jong 2001; Papovich et al. 2001; Bruzual & Charlot 2003; Gallazzi et al. 2005; Leja et al. 2017, 2019a,b). In particular, an incorrect metallicity will be compensated by an incorrect estimation of age, dust, or SFH. Accordingly, derived metallicities, SFHs, and dust components may be systematically biased (Leja et al. 2017; Carnall et al. 2019; Leja et al. 2019a,b; Pacifici et al. 2023; Bellstedt & Robotham 2025). Thus, our work may have significant implications for our understanding of the

SFHs of massive galaxies in the early Universe as derived from SED modelling (e.g., Glazebrook et al. 2017, 2024; Carnall et al. 2023b, 2024; de Graaff et al. 2025; Bevacqua et al. 2025; Nanayakkara et al. 2025), or for why massive quiescent galaxies are still quite dusty (e.g., Gobat et al. 2018; Lee et al. 2024; Setton et al. 2024; Siegel et al. 2025).

Another set of key quantities that could be affected is the M/L ratios and therefore stellar masses. To explore this issue, we compare trends between M/L ratios in our data and those predicted by the models. It has been found previously that M/L ratios, similar to ages, have a tight relationship with rest-frame colours (e.g., Bell & de Jong 2001; van de Sande et al. 2015). Thus, in Fig. 3.7 we show a similar comparison between our observations and SPS models as in Fig. 3.4, but now colour coding the rest-frame UVJ diagrams with M/L ratios.

In the top row we show the V -band M/L ratios (M/L_V) and in the bottom row we show the K_s -band M/L ratios (M/L_{K_s}). We show the LEGA-C galaxies in the left panels, where we compute M/L ratios using dynamical-model-calibrated virial masses derived in van der Wel et al. (2022). We derive luminosities from the EAZY V - and K_s -band colours (redetermined as in Section 3.2, using the LEGA-C spectroscopic redshifts). The virial masses were scaled to dynamical masses obtained via Jeans modelling in van Houdt et al. (2021), and are defined as

$$M_{\text{vir}} = K(n)K(q) \frac{\sigma_{*,\text{int}}'^2 R_{\text{sma}}}{G} \quad (3.1)$$

(equation (9) in van der Wel et al. 2022), where $K(n) = 8.87 - 0.831n + 0.0241n^2$ and where n is the Sérsic index (Cappellari et al. 2013b), $K(q) = (0.87 + 0.38e^{-3.78(1-q)})^2$ and where q is the projected axis ratio $q = b/a$, $\sigma_{*,\text{int}}'^2$ is the spatially integrated stellar velocity second moment from LEGA-C, and R_{sma} is the semimajor axis of the ellipse that contains 50 per cent of the projected Sérsic light model from the best-fitting Jeans model.

We use dynamical masses instead of stellar masses as they are independent of SPS models (see also van de Sande et al. 2015), similar to our age and metallicity measurements in the left panels of Fig. 3.4. As a caveat, we note that they include both baryonic and dark matter contributions⁸. The dark matter component may systematically bias the M/L ratios to higher values; however, the qualitative trends along the UVJ diagram would be the same assuming a similar dark matter fraction within one R_e .

In the middle and right panels of Fig. 3.7, we show the stellar M/L ratios for the FSPS and BC03 models, respectively, derived for each galaxy using the assumed age and metallicity (with added random scatter for the UVJ colours as in Section 3.3.2) from which we compute luminosities. As an additional caveat, we note that our treatment of dust could affect the observed luminosities while the assumed IMF will affect the model M/L ratios. We discuss this further in Section 3.4.4.

Similar to the metallicity panels in Fig. 3.4, the qualitative trends between M/L ratios in the data and in the models are not consistent. In particular, while the

⁸As these are quiescent galaxies, we neglect the gas mass.

M/L ratio increases roughly along the quiescent sequence in most panels, the data and models curve in opposite directions at the reddest tip of the quiescent sequence. Moreover, in the bottom-right panel, the M/L_{K_s} ratio increases approximately perpendicular to the quiescent sequence in the BC03 models, while this is not the case for the data. Thus, the data-model inconsistencies that we find in Section 3.3.2 also apply to stellar mass measurements.

Our results are consistent with those of van de Sande et al. (2015), who compared M/L -colour relations to those predicted by SPS models for massive quiescent galaxies out to $z \sim 2$. They found that the BC03, FSPS, and Maraston & Strömbäck (2011) SPS models were not able to simultaneously reproduce the relationships that they found between dynamical g -band M/L ratios and $(g - z)$ rest-frame colours and dynamical K -band M/L ratios and $(g - K)$ rest-frame colour. Furthermore, Bellstedt & Robotham (2025) explored the impacts of different stellar libraries and model assumptions on SPS-model-derived stellar masses and found that the primary impact on stellar masses was in assuming a fixed, solar metallicity, followed by stellar library differences (see also Jones et al. 2022) and SFH parameterizations (see also Carnall et al. 2019; Leja et al. 2019a; Lower et al. 2020; Suess et al. 2022; Haskell et al. 2024).

Additionally, the middle and right panels in Fig. 3.7 demonstrate that there are significant differences in the M/L ratio trends between the models themselves. This was also indicated by van de Sande et al. (2015), who found that introducing a variable IMF improved the match between observed and predicted M/L -colour relations for certain models, but not all (see also Bellstedt & Robotham 2025). Similarly, Muzzin et al. (2009) found that different SPS models result in best-fitting stellar masses that differ by a factor of up to 1.5. Finally, Lee et al. (2025) found systematic differences in M/L ratios measured using four different SPS models (Vazdekis et al. 2016, BC03, Plat et al. 2019, and FSPS).

Given these findings, it would be informative to understand whether ages, metallicities, or other quantities are primarily driving these M/L ratio differences. To this end, we examine relationships between our measured ages and metallicities and M/L ratios (not shown), and find that there may be a correlation between M/L and age in our observations, albeit with large scatter. There is no clear relation between M/L and metallicity in our observed LEGA-C galaxies. Additionally, we examine the ratio between the observed and modelled M/L ratios (not shown). In this case, there is also no clear relation with age or metallicity, indicating that we do not observe any specific biases for specific properties. While this is unexpected given the inconsistencies between the models and the data in UVJ space, it is not surprising that this analysis does not yield any further insights as the observed M/L ratios derived from dynamical masses rely on many assumptions, have significant uncertainties, and include dark matter (van Houdt et al. 2021; van der Wel et al. 2022). Thus, understanding the driver of M/L ratio differences between the data and the models will benefit from a more in-depth investigation.

3.4.3 Implications for SPS models

In Section 3.3.2, we showed that popular SPS models such as FSPS and BC03 can broadly reproduce the age trend in the UVJ diagram, but cannot reproduce the $[\text{Fe}/\text{H}]$ trend. We also showed that these models fail to predict the observed M/L ratio trends in Section 3.4.2. In this section, we discuss possible explanations for these data-model inconsistencies.

First, the observed discrepancies could be attributed to differences in the models used to measure stellar population parameters from absorption lines (i.e., ALF) and the models used to produce the UVJ colours (i.e., FSPS, BC03). While both FSPS and ALF use the same MIST isochrones (Choi et al. 2016) and MILES stellar libraries (Sánchez-Blázquez et al. 2006), the major difference between the SPS models considered here and ALF is that ALF allows for a variable abundance pattern while FSPS and BC03 assume solar-scaled abundance patterns. If models with α -enhancements or variable abundance patterns predict different continuum shapes compared to models with solar-scaled abundance patterns, this could explain the differences between the observations and model predictions in Figs 3.4 and 3.7. For example, Choi et al. (2019) found that the variable abundance pattern implemented in ALF was necessary to simultaneously reproduce spectra and broad-band, rest-frame, optical colours (ugr). Vazdekis et al. (2015) found that α -enhancement and a bottom-heavy IMF are required to reproduce observed photometric colours. Several other groups have also introduced α -enhanced models (e.g., Coelho et al. 2007; Lee et al. 2009; Percival et al. 2009; Knowles et al. 2021, 2023; Byrne et al. 2025; Park et al. 2025). Thus, implementing more widespread use of non-solar abundance patterns, especially when studying α -enhanced stellar populations such as those in distant massive quiescent galaxies, may be a promising way forward in explaining the discrepancies in UVJ space. In fact, Park et al. (2025) recently⁹ presented α -MC, a new α -enhanced implementation of the FSPS models. In the future, it will be interesting to generate the models in Section 3.3.2 with α -MC to examine how α -enhancement affects the observed trends in stellar properties.

In this context, it is interesting to note that these different approaches (solar-scaled versus variable abundance patterns) lead to different fitting results. For example, Nersesian et al. (2024) computed ages and metallicities for the LEGA-C galaxies using PROSPECTOR (Johnson et al. 2021, which makes use of the FSPS models, Conroy et al. 2009; Conroy & Gunn 2010). We find that there is good agreement between the PROSPECTOR and ALF ages (albeit with a large amount of scatter), but poor agreement between the metallicities (not shown, A. Nersesian, private communication). This was also found in Beverage et al. (2025), who fit the spectra of massive quiescent galaxies at $1 < z < 3$ with a PYTHON version of ALF (ALF α , Beverage 2024; Beverage et al. 2025) and compared their ages and metallicities to PROSPECTOR results from Slob et al. (2024). Thus, implementing non-solar abundance patterns in standard fitting codes may improve the consistency in age and metallicity measurements between the different codes as well.

⁹At the time of writing, the α -enhanced FSPS models from Park et al. (2025) were not available as part of PYTHON-FSPS for us to test here.

Another potential reason for the discrepancies between the models and the data (from a modelling perspective) is the contribution of evolved stellar stages. Specifically, the TP-AGB phase is prominent at low metallicities for young galaxies (< 2 Gyr) and can contribute significantly to the M/L ratio (Walcher et al. 2011). The TP-AGB phase is implemented differently in every SPS model, and its exact contribution is debated (compare Maraston 2005, Maraston et al. 2006, and Lu et al. 2025, who found that a prominent TP-AGB phase reproduced observations the best, with Kriek et al. 2010 who found that a prominent TP-AGB phase overpredicted near-infrared luminosities, and see Bevacqua et al. 2025 as well). These differences can have varying impacts on SED fits. In particular, Muzzin et al. (2009) fit the near-infrared SEDs of young galaxies, where the effects of the TP-AGB should be most evident, with different SPS models, and found that the best-fitting stellar population parameters differed by a factor of up to 3. Thus, agreement on the implementation of TP-AGB stars and other evolved stellar stages is needed.

Other aspects of SPS models could help alleviate the discrepancy with observations. For instance, it will also be important to reconcile different spectral libraries, as they have been found to produce different results. In particular, Byrne & Stanway (2023) compared multiple spectral libraries using the same SPS model and found that the discrepancies with spectroscopic quantities when using different libraries were worse for metallicities than for ages, as the 4000 Å break was relatively insensitive to library differences. Similar results were found by Bellstedt & Robotham (2025). We note that ALF makes use of the C3K stellar library, which Byrne & Stanway (2023) found to be the most robust out of their tested libraries. Finally, different SPS models do not always agree on the definition of metallicity. For example, total metallicity is often computed using a combination of $[Mg/Fe]$ and $[Fe/H]$ (e.g. Thomas et al. 2003) when Mg and Fe are both measured individually. However, we note that the opacities are primarily driven by iron. Moreover, as discussed above, both FSPS and BC03 assume solar-scaled abundances, while massive quiescent galaxies are α -enhanced. In particular, Beverage et al. (2025) showed that the best-fitting FSPS metallicities follow $[Fe/H]$, and not $[Z/H]$, for galaxies that are α -enhanced (like the ones in this sample). See also Choi et al. (2014) and Leja et al. (2019a).

In summary, our results demonstrate that standard SPS models do not predict the correct UVJ colours for quiescent galaxies when given stellar metallicities or ages (even when taking into account dust and extended SFHs, see Section 3.4.4). Thus, we emphasize the need for caution when interpreting or assuming metallicities or stellar masses predicted from photometry, or derived via solar-abundance-scaled codes that rely on the shape of the continuum (see also Nersesian et al. 2024; Bellstedt & Robotham 2025; Beverage et al. 2025). More work is needed to implement non-solar abundance patterns (see e.g. Park et al. 2025), understand the contribution of continuum shapes and advanced stellar stages, and reconcile the ingredients in different SPS models.

3.4.4 Caveats

Here, we discuss the main caveats that should be taken into consideration when interpreting our results, including whether our observed trends may be driven by other physical properties or assumptions that we have made, and our use of ALF.

Throughout this work, we assume that there is no dust in the LEGA-C galaxies and we do not implement a dust law in our generation of FSPS and BC03 models. However, it is likely that at least some of the LEGA-C galaxies are affected by dust. In particular, the location of our data in the UVJ diagram in Figs 3.3, 3.4, and 3.7 is slightly redder than expected from the models, especially when we compare with FSPS, and dust could be contributing to this offset. The presence of dust would also introduce additional scatter in the SPS model panels (middle and right) in Figs 3.4 and 3.7. We test these possibilities by generating the same models as shown in Figs 3.4 and 3.7, but assuming a Calzetti et al. (2000) dust law with (i) FAST (Kriek et al. 2009, 2018) A_V values, measured by and obtained via private communication with the LEGA-C team for each galaxy (although these values will be heavily affected by the degeneracies and model inaccuracies that we discuss throughout the paper), and (ii) A_V values that scale with age (as dust approximately follows the age trend in the UVJ diagram, see the arrow in Fig. 3.3), such that $A_V = 0.5$ at 1 Gyr, decreasing down to an age of 3 Gyr, beyond which $A_V = 0$. We correct the model FSPS and BC03 luminosities for the respective dust extinction in each case (not shown). In all cases, dust moves the distribution along the arrow shown in Fig. 3.3, but does not provide a better match between the observations and the models. However, we note that implementing the FAST A_V values for each galaxy broadens the distribution of the models in UVJ space in the middle and right panels of Figs 3.4 and 3.7 to better match the observations. Dust also acts to increase the M/L ratios; however, it does not change their distribution in UVJ space, preserving the discrepancy between the observations and the models. While we find that dust does not drive the disagreement between the observations and the models, we note that correlations between galaxy properties could alter the distribution of galaxies in UVJ space. Thus, it is still possible that better A_V constraints could bring the observations and models into agreement. Further exploration of methods to disentangle dust from other physical properties is required.

Additionally, in our FSPS and BC03 model generation, we assume a relatively simple SFH with a short, truncated burst. We also assume a single age in our ALF fits. However, if multiple stellar populations are present (see e.g. Carnall et al. 2024; Nanayakkara et al. 2025), this could also play a role in the offset that we see between the data and the models. We test the effects of several different SFHs, by extending the burst, varying the τ value, and implementing a delayed- τ model. We find that altering the assumed SFH may help to explain some of the scatter that we see in the right-most panels of Figs 3.4 and 3.7, but does not account for all of it. Crucially, implementing different SFHs does not generate the trends that we observe in age, metallicity, or M/L ratios, with these parameters behaving the same way in these test models as with our assumed simple SFH.

There may be other effects creating the trends in stellar population parameters

that we see that are not inherent to the SPS models. For example, the age and metallicity measurements from which we derive our main results are based solely on our fits using ALF, which may not be as accurate for galaxies near the edge of the UVJ box; these galaxies could potentially be dusty post-starburst galaxies (Zick et al. 2018) or have low levels of star formation (Akins et al. 2022). Since our conclusion of a metallicity trend depends on this region of the UVJ diagram, we test shifting the UVJ box by 0.1 magnitudes toward redder $U - V$ colours to examine the effect of excluding potential contaminant galaxies near the edge of the UVJ box. We find that both the age and metallicity trends are still present and follow the same qualitative direction as in Fig. 3.2, although the age trend is slightly tighter. Thus, the inclusion of the galaxies towards the edge of the UVJ box does not affect our conclusions.

To quantitatively examine the potential biases of using ALF, we additionally test whether ALF could be predicting lower $[\text{Fe}/\text{H}]$ abundances for a quiescent population which had a recent minor starburst. In this case, the equivalent widths of absorption lines would be decreased by an additional, mostly featureless, continuum contribution. To test this scenario, we select 33 quiescent galaxies from LEGA-C that lie along the quiescent sequence, but away from the edge of the UVJ box. For each selected LEGA-C galaxy, we generate a young (0.1 Gyr) SED with FSPS that is 6 per cent of the luminosity of the corresponding LEGA-C galaxy in the V band, and add the young SED to the LEGA-C spectrum to make 33 composite spectra that still look visually quiescent. We determine this luminosity percentage by estimating the UVJ colours of the composite spectra such that the composite galaxies lie near the edge of the UVJ box. Note that we give the young SEDs the same metallicity as derived from the fiducial fit to the corresponding LEGA-C spectrum. We also test giving the young SEDs higher and lower metallicities and find that this does not affect our conclusions. We fit the composite spectra with ALF in the same way as the corresponding LEGA-C spectra. We find that including a young population on top of an older quiescent one results in $[\text{Fe}/\text{H}]$ values that are consistent with the $[\text{Fe}/\text{H}]$ values that we derive from the fiducial fits within uncertainties. However, we note that in many cases the fits are driven to slightly lower median $[\text{Fe}/\text{H}]$ values compared to the fiducial case. In particular, we find that an average decrease of ~ 0.24 in $U - V$ colour and ~ 0.04 in $V - J$ colour results in an average decrease in $[\text{Fe}/\text{H}]$ of ~ 0.05 dex. When we compare this to our expectations from a similar decrease in $U - V$ and $V - J$ in our observations in the lower right panel of Fig. 3.2, we see that the average observed decrease in $[\text{Fe}/\text{H}]$ is ~ 0.11 dex. Therefore, low levels of star formation may contribute to the observed metallicity trend, and in general we caution that composite stellar populations near the edge of the UVJ box may lead to lower metallicities. However, based on the test presented here, these low levels of star formation cannot fully explain the observed metallicity trend, and we note that such a recent burst would generate emission lines, which we do not observe in our spectra.

In the future, it will still be important to confirm our ALF fitting results via other means. Encouragingly, we note that our observed metallicity trend is consistent with the results of Nersesian et al. (2025), who simultaneously fit photometry

and spectroscopy in a large sample of star-forming and quiescent LEGA-C galaxies. Another possibility is to estimate ages and metallicities via traditional Lick indices (Worthey et al. 1994), which could potentially provide us with insight into whether our results are robust. However, these measurements would not have the same independence as our ALF measurements, as SPS models like FSPS or BC03 are required to interpret Lick indices. Additionally, this method will likely not be able to provide age and metallicity estimates to the same level of accuracy as ALF as it only makes use of a few individual features while ALF considers the full spectrum (see e.g. Kriek et al. 2016). Other methods of constraining ages and metallicities are beyond the scope of this work, but should be considered in the future.

Finally, we test whether the metallicity trend could be caused by distinct types of galaxies populating different parts of the UVJ diagram. For example, if a specific population with low metallicity is preferentially found towards the edge of the UVJ box, this could drive the observed trend. For quiescent galaxies, properties that are typically correlated with lower metallicity include lower velocity dispersion, lower mass, or higher redshift (e.g., Smith et al. 2009; Leethochawalit et al. 2019; Beverage et al. 2023, 2024, 2025). We do not find any preference of lower velocity dispersion or lower mass populations towards the edge of the UVJ box. There is a slight preference of lower metallicities for higher redshift galaxies, which are indeed closer to the edge of the UVJ box. It is still unclear why these galaxies fall in this region of the UVJ diagram; however, we note that our sample also contains fewer galaxies at these redshifts. Larger samples of high redshift galaxies will be required going forward to examine whether there is a true dependence on redshift. The only additional property that is preferentially found towards the edge of the box is slightly higher sSFR. However, at the same mass, we do not expect quiescent galaxies with higher sSFR to have lower metallicity. In fact, we would expect the opposite trend, as galaxies that quench at later times and have longer star-formation time-scales will have higher $[\text{Fe}/\text{H}]$ values (e.g., Beverage et al. 2021, 2024, 2025).

3.5 Summary and conclusions

The age–metallicity degeneracy (Worthey 1994; Bruzual & Charlot 2003; Gallazzi et al. 2005) poses a major challenge to interpreting photometric data of quiescent galaxies. Measuring metallicities has historically been challenging, especially beyond the local Universe, where low S/N data make it difficult to constrain weak metallicity-sensitive features. As a result, metallicity has typically been left unconstrained or fixed to solar values, leading to (systematic) uncertainties in stellar population studies. Consequently, the red colours of quiescent galaxies, for example in the well-known UVJ diagram, were fully attributed to age (e.g., Whitaker et al. 2012, 2013; Mendel et al. 2015; Belli et al. 2019; Díaz-García et al. 2019), and our knowledge of how metallicity affected the shapes of galaxy continua was completely dependent on solar-scaled models, which become increasingly uncertain at earlier times.

In this work we compared, for the first time, spectroscopically derived ages and metallicities (from continuum-normalized, absorption-line spectra) to independently measured broad-band colours of quiescent galaxies beyond the low-redshift Universe. In particular, we examined a sample of ~ 700 massive quiescent galaxies at $0.6 \lesssim z \lesssim 1.0$ from the LEGA-C survey (van der Wel et al. 2021). We fit the continuum-normalized, absorption-line spectra using ALF (Conroy & van Dokkum 2012a; Conroy et al. 2018) and recovered SSP-equivalent ages and $[\text{Fe}/\text{H}]$ values. We compared these measurements to UltraVISTA (McCracken et al. 2012) UVJ colours (Muzzin et al. 2013b) by colour coding the UVJ diagrams of these galaxies by age and metallicity. We also compared our observed trends between spectroscopic ages and metallicities and broad-band UVJ colours to predictions from popular SPS models including FSPS (Conroy et al. 2009; Conroy & Gunn 2010) and BC03 (Bruzual & Charlot 2003). Our findings can be summarized as follows:

- Galaxy age generally increases along the quiescent sequence in the UVJ diagram. This was implied in Whitaker et al. (2012, 2013), Mendel et al. (2015), Belli et al. (2019), and Díaz-García et al. (2019); however, for the first time, we detected this trend using age measurements that are independent of UVJ colours, while also accounting for stellar metallicity. We also found that there is a slight curvature to the trend, with the oldest galaxies populating the reddest end of the $U - V$ axis, but located slightly blueward on the $V - J$ axis. Thus, we concluded that a simple linear relation between UVJ colours and ages does not represent the age trend very well at these redshifts, in contrast to the findings of Belli et al. (2019) at $1.0 < z < 2.5$.
- Stellar metallicity increases approximately perpendicular to the quiescent sequence. *Thus, the sequence in metallicity does not follow the sequence in age*, in contrast to predictions from stellar population models (Bruzual & Charlot 2003; Conroy et al. 2009; Conroy & Gunn 2010). This finding indicates that the interpretation of the quiescent sequence as an age sequence (Whitaker et al. 2012, 2013; Mendel et al. 2015; Belli et al. 2019; Díaz-García et al. 2019) is robust.
- In general, SPS models are able to qualitatively reproduce our observed age trend, though the observed galaxies are on average redder than model predictions. This difference could be explained by some dust attenuation.
- However, the models predict that metallicity increases nearly perpendicular to the trend that we observe in the data. In other words, SPS models predict very different UVJ colours for a given metallicity than the colours that we observe.
- These data-model inconsistencies likely impact our understanding of stellar population properties of quiescent galaxies derived from photometric data, as an incorrect metallicity will be compensated by age, dust, or SFH, which may subsequently bias M/L ratios. To explore this further, we examined the trends between M/L ratios and the UVJ diagram measured in our observed

LEGA-C galaxies and predicted by FSPS and BC03. For our observations, we made use of independent M/L ratios derived from dynamically scaled virial masses. Similar to the metallicity trend, we found that the M/L ratio trends predicted by SPS models behave very differently from those which we observe. This difference indicates that stellar masses inferred from M/L ratios measured via SPS models may also suffer from model inaccuracies.

Our work suggests that, while state-of-the-art SPS models are powerful tools in deriving stellar population properties, further development is required to accurately recover the stellar metallicities, ages, and M/L ratios of quiescent galaxies. In particular, the implementation of non-solar abundance patterns and a better understanding of the contribution from evolved stellar stages may improve the mismatch between observations and SPS model predictions, as well as the discrepancies between different SPS models. The need to improve model predictions will only become more urgent as we continue to push the frontier of quiescent galaxy studies to higher redshifts with *JWST*.

Acknowledgments

We thank the anonymous referee for taking the time to give us useful feedback that improved this manuscript. We thank the LEGA-C team for making their data set public. We also thank Brian Lorenz, Angelos Nersesian, and Colin Yip for useful conversations. This work was performed using the compute resources from the Academic Leiden Interdisciplinary Cluster Environment (ALICE) provided by Leiden University. This work also used the Dutch national e-infrastructure with the support of the Samenwerkende Universitaire Rekenfaciliteiten (SURF) Cooperative using grant no. EINF-6344 and grant no. EINF-10017 which are financed by the Dutch Research Council (NWO). MK acknowledges funding from the NWO through the award of the Vici grant VI.C.222.047 and NSF AAG grant AST-1909942. PEMP acknowledges the support from the Dutch Research Council (NWO) through the Veni grant VI.Veni.222.364.

Data Availability

This study makes use of data from the LEGA-C survey. The 2D and reduced 1D spectra can be obtained from the ESO Science Archive Facility (http://archive.eso.org/eso/eso_archive_main.html). The reduced spectra and catalogue have been released by ESO (<http://archive.eso.org/cms/eso-archive-news/Third-and-final-release-of-the-Large-Early-Galaxy-Census-LEGA-C-Spectroscopic-Public-Survey-published.html>) and are also available here: <https://users.ugent.be/~avdrwel/research.html#legac>. For details, see van der Wel et al. (2016, 2021) and Straatman et al. (2018).

The re-determined UVJ colours as well as ages, metallicities, and $[Mg/Fe]$ values are available in an online catalogue. Other data products generated in the course of this work will be made available upon reasonable request.

4

CLUES TO INSIDE-OUT QUENCHING IN QUIESCENT GALAXIES AT $1.2 \lesssim z \lesssim 2.2$: AGE, FE-, AND MG-ABUNDANCE GRADIENTS FROM *JWST*-SUSPENSE

ABSTRACT

Spatially resolved stellar populations of massive quiescent galaxies at cosmic noon provide powerful insights into star-formation quenching and stellar mass assembly mechanisms. Previous photometric studies have revealed that the cores of these galaxies are redder than their outskirts. However, spectroscopy is needed to break the age-metallicity degeneracy and uncover the driver of these colour gradients. In this work, we derive the age and elemental abundance gradients for eight distant ($1.2 \lesssim z \lesssim 2.2$), massive ($10.3 \lesssim \log(M_*/M_\odot) \lesssim 11.1$) quiescent galaxies by fitting full-spectrum models to ultra-deep NIRSpec-MSA spectroscopy from the *JWST*-SUSPENSE survey. We find that these galaxies have negative age and flat [Fe/H] gradients as well as tentative indications of positive [Mg/H] and [Mg/Fe] gradients. These results suggest that galaxy cores are older and perhaps also Mg deficient compared to galaxy outskirts. The age gradients may indicate inside-out quenching, while Mg-deficient cores could suggest rapid gas expulsion as the central quenching mechanism. Thus, galaxy cores may have formed faster and quenched more efficiently than their outskirts. In this scenario, however, our [Fe/H] and [Mg/Fe] gradients are still puzzling, and further investigation is required to understand the nature of [Mg/H] gradients in massive quiescent galaxies at these redshifts. Our results contrast with those of lower-redshift studies, which find flat age and [Mg/Fe] gradients and negative metallicity gradients. Additionally, we find a positive trend between age gradients and rotational support and marginal trends between [Fe/H] gradients and galaxy velocity dispersions and ages. We discuss our findings in the context of galaxy growth scenarios, including minor mergers and progenitor bias. With this work, we present the first stellar population gradients from NIRSpec-MSA spectroscopy in the current largest sample of distant quiescent galaxies.

Chloe M. Cheng, Martje Slob, Mariska Kriek, Aliza G. Beverage, Guillermo Barro, Rachel Bezanson, Anna de Graaff, Natascha M. Förster Schreiber, Brian Lorenz, Danilo Marchesini, Ignacio Martín-Navarro, Adam Muzzin, Andrew B. Newman, Sedona H. Price, Katherine A. Suss, Arjen van der Wel, Jesse van de Sande, Pieter G. van Dokkum, and Daniel R. Weisz

2026, *Astronomy & Astrophysics*, in press

4.1 Introduction

Spatially resolved measurements of stellar populations are critical for our understanding of galaxy assembly histories and star-formation quenching mechanisms. In the local Universe, key insights into the formation and assembly of massive early-type galaxies have been gained by measuring stellar population gradients out to large radii (e.g. Peletier & Valentijn 1989; Franx & Illingworth 1990; Kuntschner et al. 2006; Goddard et al. 2017; Li et al. 2018; Santucci et al. 2020; Parikh et al. 2024). In particular, findings of flat α -element abundance gradients, flat or mildly positive age gradients, and mildly negative metallicity gradients (implying iron-rich cores) suggest that these nearby galaxies have experienced inside-out growth (e.g. Mehlert et al. 2003; La Barbera et al. 2005; Greene et al. 2013, 2015; Greene et al. 2019; Martín-Navarro et al. 2018; Zibetti et al. 2020). In this scenario, the compact central cores of massive galaxies are formed at cosmic noon ($z \sim 2-3$) or earlier, with their outer wings building up via minor mergers towards $z \sim 0$ (e.g. Bezanson et al. 2009; Hopkins et al. 2009a; Naab et al. 2009; Oser et al. 2010; van de Sande et al. 2013; van Dokkum et al. 2014; Rodríguez-Gomez et al. 2016).

Advancements in observations and modelling have allowed these measurements to be extended to higher redshifts ($z \sim 1$). In particular, negative colour gradients, (e.g. Gargiulo et al. 2012; Suess et al. 2019a,b, 2020, 2021; Miller et al. 2023; van der Wel et al. 2024), flat age and α -element abundance gradients, and negative metallicity gradients contribute to the inside-out growth picture (Cheng et al. 2024; see also D’Eugenio et al. 2020), with galaxy outskirts being built up by the accretion of low-metallicity (i.e. bluer) satellite galaxies. In this context, however, the flat age and [Mg/Fe] gradients are more difficult to explain, as they require that the accreted satellites have similar ages and star-formation histories as the central galaxy. On the other hand, we note that this may be expected for satellite galaxies (see e.g. Pasquali et al. 2010; Gallazzi et al. 2021; Oyarzún et al. 2023).

The resolved stellar populations of galaxies at even earlier times are crucial to assessing the inside-out growth picture. High-resolution photometric studies beyond $z \sim 1$ may in fact support this scenario, as colour gradients have been observed to strengthen between $z \sim 2$ and $z \sim 0$, with galaxies’ blue outskirts building up over time (Suess et al. 2019b, although note that this scenario is still under debate due to new evidence from JWST; see e.g. Suess et al. 2022; Martorano et al. 2026; McGrath et al. 2026). These colour gradients suggest that the compact cores of galaxies accrete bluer low-mass satellites (e.g. Wuyts et al. 2010; Guo et al. 2011; Gargiulo et al. 2012; Szomoru et al. 2013; Ciocca et al. 2017; Suess et al. 2019b, 2020, 2021, 2023; Miller et al. 2022, 2023; Setton et al. 2024).

On the other hand, the strengthening of colour gradients (as well as the increase in galaxy size) over cosmic time could be explained by larger galaxies with stronger gradients quenching at later times (i.e. progenitor bias; van Dokkum & Franx 2001; Carollo et al. 2013; Poggianti et al. 2013; Keating et al. 2015; Damjanov et al. 2019, 2023). In this case, gradients in galaxies at $z \gtrsim 1$ would be inherited from their star-forming progenitors. This would suggest that both age and metallicity contribute to the colour gradients, as distant star-forming galaxies have been found to have negative metallicity (e.g. Jones et al. 2015) and age gradients (Tripodi

et al. 2024; Shen et al. 2024). Colour gradients at $z \gtrsim 1$ could also be signatures of inside-out quenching (e.g. Suess et al. 2019a), as older central stellar populations result in redder cores. Furthermore, dust could also play a role (e.g. Miller et al. 2023; Setton et al. 2024), as age, metallicity, and dust are strongly degenerate in broadband spectral energy distributions (Worthey 1994; Bell & de Jong 2001; Bruzual & Charlot 2003; Gallazzi et al. 2005; Leja et al. 2019a).

To be able to discriminate between the scenarios described above, the physical property that may be driving the observed colour gradients needs to be determined. We must therefore measure robust stellar population gradients in massive quiescent galaxies out to at least $z \sim 2$, the peak of the quenching and assembly era (Oser et al. 2010; Whitaker et al. 2012; Rodriguez-Gomez et al. 2016). To achieve this goal, deep, spatially resolved rest-frame optical spectra of a sample of massive quiescent galaxies are required to obtain detailed measurements of absorption lines. However, these measurements are extremely challenging beyond the local Universe. Thus, spectroscopic gradient measurements exist for only a small number of relatively young quiescent galaxies beyond $z \sim 1$. For example, Jafariyazani et al. (2020) found a flat age and [Mg/Fe] gradient and a slightly negative [Fe/H] gradient in a massive lensed galaxy at $z \sim 2$ using deep MOSFIRE spectra. Ditrani et al. (2022) used *Hubble Space Telescope* (*HST*) grism spectra and found negative metallicity and diverse age gradients in four distant quiescent galaxies. Akhshik et al. (2023) examined similar data and found diverse age and metallicity gradients in eight galaxies. However, due to line blending, metallicities obtained from low-resolution spectroscopy ($R \sim 100$) may be highly uncertain ($\sim 0.3 - 0.5$ dex, Akhshik et al. 2023).

Detailed stellar population maps have also been achieved with *James Webb Space Telescope* (*JWST*) integral field unit (IFU) spectroscopy, but only for very young individual quiescent galaxies. For example, Pérez-González et al. (2025) examined a massive young (0.6 Myr old) galaxy at $z \sim 3.7$ and found a strongly negative total metallicity gradient, a slightly negative age gradient, and a slightly positive dust gradient. Additionally, D'Eugenio et al. (2024) examined a young (0.5 Gyr old) galaxy at $z \sim 3$ and found a flat age gradient. To come to a consensus about the behaviour of spatially resolved stellar populations in distant quiescent galaxies, larger, more representative samples that also include older, fainter galaxies are needed. While *JWST*-IFU observations such as those in Pérez-González et al. (2025) and D'Eugenio et al. (2024) are ideal, it is prohibitively expensive to obtain these data for larger samples.

To avoid the expense of IFU observations, we can take advantage of the micro shutter assembly (MSA), *JWST*'s multi-object spectrometer on the Near Infrared Spectrograph (NIRSpec). This goal has been achieved with the *JWST*-Spectroscopic Ultradeep Survey Probing Extragalactic Near-infrared Stellar Emission (SUSPENSE), an ultra-deep spectroscopic survey of 20 quiescent galaxies at $1 \lesssim z \lesssim 3$ (Slob et al. 2024). SUSPENSE leverages the NIRSpec-MSA to achieve both medium spectral and moderate spatial resolutions for many galaxies in a single pointing. In this work, we took advantage of the unique capabilities of this instrument to present robust spatially resolved measurements of age, [Fe/H], [Mg/Fe], and [Mg/H] in a sample of eight massive quiescent galaxies at

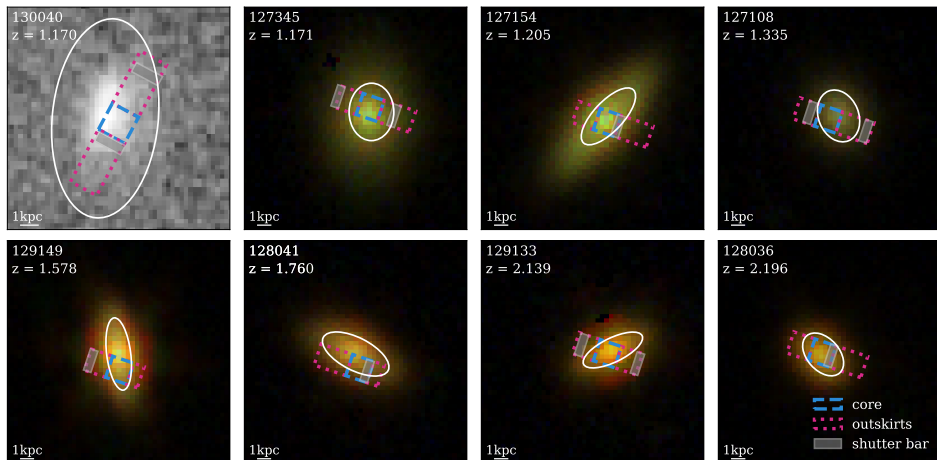


Figure 4.1: Colour image cut-outs of the eight distant quiescent galaxies in our sample. Seven galaxies have NIRCcam imaging, which we combine into RGB images here. Galaxy 130040 does not have NIRCcam imaging, so we show a cut-out of the COSMOS *HST*/ACS F814W image (Koekemoer et al. 2007; Scoville et al. 2007b; note that this image is rotated by 90 degrees compared to the NIRCcam images). We highlight the regions that we combined for the core and outskirts spectra in blue and pink, respectively (see Section 4.3.2). We also indicate where shutter bars lie within these core and outskirts regions with shaded rectangles. We indicate $1 R_e$ in convolved space for each galaxy with an ellipse (see Section 4.3.3).

$1.2 \lesssim z \lesssim 2.2$ from SUSPENSE.

This paper is organised as follows: In Section 4.2 we describe the spectroscopic data from SUSPENSE and our sample selection. In Section 4.3 we outline our methods to correct spectral resampling artefacts, extract spatially resolved spectra, determine de-projected distances, and measure spatially resolved stellar population parameters. In Section 4.4 we present our resulting gradients. We discuss our results and their implications for galaxy quenching mechanisms and assembly histories in Section 4.5 and present our conclusions in Section 4.6. Throughout this work we assume a flat Λ cold dark matter cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and a Kroupa (2001) initial mass function (IMF). All magnitudes are given in the AB-magnitude system (Oke & Gunn 1983).

4.2 Data and sample

The spectroscopic data in this work were drawn from *JWST*-SUSPENSE¹ (ID: 2110², PIs: Kriek and Beverage), an ultra-deep (16.4 hours on-source), NIRSpect-MSA/G140M-F100LP ($R \sim 1300$) survey of 20 massive quiescent galaxies at $1 \lesssim z \lesssim 3$ (Slob et al. 2024, see also Beverage et al. 2025). Targets were identified

¹<https://suspense.strw.leidenuniv.nl/> .

²[doi:10.17909/y6rb-fn24](https://doi.org/10.17909/y6rb-fn24).

Table 4.1: Sample of massive quiescent galaxies from *JWST*-SUSPENSE examined in this work.

ID	$z_{\text{spec}}^{\text{a}}$	$\log(M/M_{\odot})^{\text{a}}$	Kinematic Properties ^b			R_{e}^{c}		Axis ratio (b/a) ^b		Radius ^c		$S/N_{4520-4820\text{\AA}}^{\text{d}}$	
			V_{re}/σ_0	σ_0 (km/s)	intrinsic ^b (kpc)	convolved (kpc)	core ($R_{\text{e,conv}}$)	outskirt ($R_{\text{e,conv}}$)	core ($R_{\text{e,conv}}$)	outskirt ($R_{\text{e,conv}}$)	core (\AA^{-1})	outskirt (\AA^{-1})	
130040	1.170	11.1	$0.63^{+0.06}_{-0.07}$	285^{+5}_{-5}	5.12^{e}	5.57	5.57	0.53^{e}	0.17	0.45	21.2	19.9	
127345	1.171	10.7	-	-	1.52 ± 0.66	2.13	2.13	0.78	0.40	0.95	36.8	31.1	
127154	1.205	10.8	-	242^{+4}_{-6}	2.12 ± 0.41	2.72	2.72	0.41	0.72	1.24	54.2	30.7	
127108	1.335	10.3	$0.54^{+0.12}_{-0.20}$	210^{+5}_{-10}	1.43 ± 0.68	2.02	2.02	0.75	0.51	0.97	29.2	21.7	
129149	1.578	11.0	$0.49^{+0.07}_{-0.12}$	387^{+4}_{-34}	1.00 ± 0.08	2.77	2.77	0.33	0.73	1.56	91.7	52.7	
128041	1.760	10.7	$1.48^{+0.07}_{-0.06}$	226^{+1}_{-1}	1.60 ± 0.42	2.77	2.77	0.45	0.75	1.01	58.2	43.7	
129133	2.139	11.1	$1.19^{+0.07}_{-0.07}$	257^{+2}_{-3}	1.58 ± 0.83	2.43	2.43	0.37	0.53	1.33	59.1	29.1	
128036	2.196	11.0	$1.18^{+0.09}_{-0.08}$	210^{+3}_{-6}	1.05 ± 0.25	1.85	1.85	0.65	0.55	1.09	47.3	33.2	

^a Presented in Slob et al. (2024).

^b Presented in Slob et al. (2025).

^c See Section 4.3.3.

^d See Section 4.2. Note that these S/N values are quoted per rest-frame \AA .

^e Presented in Griffith et al. (2012).

from the UltraVISTA (McCracken et al. 2012) K_s -band DR3 catalogue (Muzzin et al. 2013b), and were selected to be quiescent using the Muzzin et al. (2013a) UVJ classification. The SUSPENSE galaxies were observed at two dither configurations, by offsetting two nearly identical MSA configurations by eight shutters in the dispersion direction, trying to maintain the same relative aperture pattern for each target. However, due to micro shutter defects, this was not possible for all targets. Thus, a small number of galaxies were observed in different positions in each configuration (see Section 4.3.2). A two-point nod pattern was also implemented for each galaxy, with a two-micro-shutter cross-dispersion offset to avoid self-subtraction of the extended targets during the data reduction. After identifying the optimal configuration, the standard three-shutter slitlets were extended by hand where possible, ranging from 3 to 7 micro shutters in length. Filler star-forming targets were also added to the configuration. The ‘unconstrained’ (midbar) option was implemented for the source centring constraint, and thus some galaxies are centred behind MSA bars (see Fig. 4.1).

The data were reduced via a customised version of v1.12.5 of the *JWST* Science Calibration Pipeline (Bushouse et al. 2023) and version 1183 of the Calibration Reference Data System (CRDS). As we only make use of the reduced 2D spectra in this work (see Section 4.3.2), we summarise the relevant data reduction steps here, which were performed for each dither separately. In particular, the master bias frame and dark current were subtracted and detector artefacts and jumps due to cosmic rays were removed. Count-rate frames were obtained by fitting the slope of each pixel and $1/f$ correlated vertical read out noise was then removed using the GRIZLI correction algorithm (Brammer 2023). The data were background subtracted using the average of all frames observed in the same visit and dither but in the opposite nod. Flat field, barshadow, and pathloss corrections were applied to the 2D spectra for each galaxy, which were then flux-calibrated. The calibrated 2D spectra were rectified and resampled to a common reference frame. A custom outlier detection algorithm was implemented, and the final 2D spectrum was constructed by performing an inverse-read-noise-weighted combination of the 2D frames. See Slob et al. (2024) for details.

We selected a subset of the quiescent galaxies from SUSPENSE. In particular, we required each spectrum (integrated, core, and outskirts, see Section 4.3.2) to have a rest-frame $S/N \gtrsim 20 \text{ \AA}^{-1}$ between $4520\text{--}4820 \text{ \AA}$ (the region that all spectra have in common) to derive robust stellar population parameters. We determined these limits by performing tests with simulated observations, similar to Cheng et al. (2024) (see Appendix 4.A). Additionally, we discarded four galaxies, due to insufficient wavelength coverage (IDs 129982 and 130208), or due to the presence of strong emission lines that may be associated with active galactic nuclei (AGN, IDs 128452 and 130647, see also Beverage et al. 2025). Our selection does not lead to a bias in mass, size, or integrated age compared to the total SUSPENSE quiescent sample.

Image cut-outs of our final sample of eight quiescent galaxies (40% of the total quiescent SUSPENSE sample) are shown in Fig. 4.1. Seven galaxies have *JWST*/Near Infrared Camera (NIRCam) imaging from the Cosmological Evolution Survey (COSMOS)-Web survey (Casey et al. 2023). We obtained these images

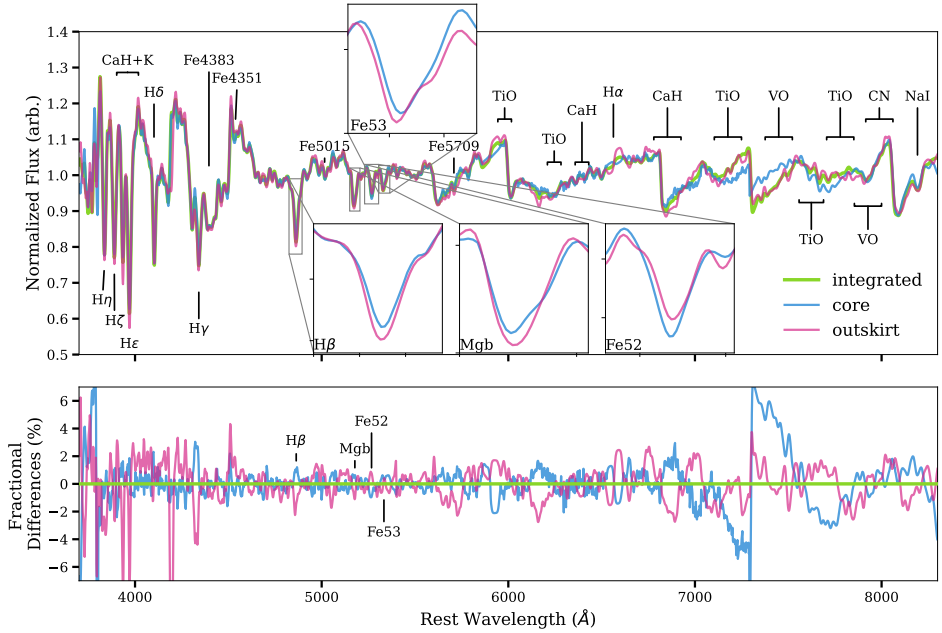


Figure 4.2: *Top:* Stack of continuum-normalised spectra of all galaxies. The integrated stack is shown in green, the core stack is shown in blue, and the outskirt stack is shown in magenta. The key spectral features are labelled. The inset panels show a zoom-in on the $H\beta$ ($\sim 4849\text{--}4878\text{ \AA}$, sensitive to age), Mgb ($\sim 5162\text{--}5194\text{ \AA}$, sensitive to $[Mg/H]$), $Fe52$ ($\sim 5247\text{--}5287\text{ \AA}$, sensitive to age and $[Fe/H]$), and $Fe53$ ($\sim 5314\text{--}5454\text{ \AA}$, sensitive to age and $[Fe/H]$) features in the core and outskirt spectra. *Bottom:* Fractional differences of each stacked spectrum. We divided each spectrum by the stacked integrated spectrum.

from the Cosmic Dawn Centre (DAWN) *JWST* Archive (DJA³, see Brammer 2023 and Valentino et al. 2023) and combined the F115W, F277W, and F444W data into RGB images using the `ASTROPY make_lupton_rgb` function (this applies an inverse hyperbolic sine scaling, Lupton et al. 2004; Astropy Collaboration et al. 2022). Galaxy 130040 does not have NIRCcam imaging, so we show the COSMOS *HST*/Advanced Camera for Surveys (ACS) F814W image (Koekemoer et al. 2007; Scoville et al. 2007b). We report properties of our sample in Table 4.1.

4.3 Methods

4.3.1 Correction of resampling noise

Spatially resolved observations obtained with *JWST* are affected by undersampling of the point spread function (PSF). This undersampling produces an artefact called ‘resampling noise’, which presents itself as periodic flux variations as a function

³<https://dawn-cph.github.io/dja/index.html> .

of wavelength (Smith et al. 2007; Law et al. 2023; Perna et al. 2023; Dumont et al. 2025; Newman et al. 2025). These variations are colloquially referred to as ‘wiggles’. This behaviour has been observed in the previous generation of space-based telescopes (see e.g. Dressel et al. 2007; Smith et al. 2007; Anderson 2016), and in *JWST* IFU instruments (e.g. Law et al. 2023; Perna et al. 2023; Dumont et al. 2025; Newman et al. 2025). We also observed these wiggles in the individual spectral rows of the SUSPENSE NIRSpec-MSA 2D spectra. The wiggles do not strongly affect integrated observations, as they are averaged out when the flux is integrated over large apertures (see Appendix 4.B and Law et al. 2023; Perna et al. 2023; Newman et al. 2025). However, the wiggles are a significant concern for spatially resolved studies as they can affect the shapes of continua and broad spectral features, potentially biasing spectral modelling (Perna et al. 2023). We corrected the wiggles by implementing an algorithm based on those presented by Perna et al. (2023)⁴ and Dumont et al. (2025)⁵, customising these routines for our specific case. See Appendix 4.B for details.

4.3.2 Spectral extraction

For each galaxy, we applied a custom extraction routine to the wiggle-corrected 2D spectrum to obtain three 1D spectra: an integrated spectrum and spectra in two spatial bins (representing the core and outskirts). We followed a similar procedure as Cheng et al. (2024), but modified for our data. We describe our extraction routine here.

We obtained the flux profile of each galaxy by collapsing the 2D spectrum over the wavelength axis. We fit a Moffat (1969) profile to each flux profile, as in Cheng et al. (2024). We used the Moffat profile to identify spectral rows with significant flux (rows that are approximately between the third and 97th percentiles of the Moffat profile). We disregarded all rows outside of the third and 97th percentiles, as Cheng et al. (2024) find that using this region results in a minimum difference in S/N between the core and outskirt spectra. We also found that including these outer rows lowers the S/N.

To extract the integrated spectrum, we summed the rows between the third and 97th percentiles. For the spatially binned spectra, we took the brightest two rows to comprise the ‘core’ of the galaxy and summed these together. The remaining rows comprised the ‘outskirts’ of the galaxy and were similarly summed. We divided our spectra in this way to ensure that the two spatial bins had roughly similar and sufficient S/N. Additionally, we chose this division to avoid splitting our rows into sub-pixels, since the spatial undersampling of NIRSpec makes neighbouring rows largely (though not completely) independent. Table 4.1 shows that all core and outskirt bins have a median S/N $\gtrsim 20 \text{ \AA}^{-1}$, and that the S/N is roughly equally divided. Finally, we did not apply an optimal extraction, so as to not down weigh the lower-signal outer rows.

As the SUSPENSE galaxies were observed using two dither configurations (see Section 4.2 and Slob et al. 2024), there are two 2D spectra for each galaxy. We per-

⁴https://github.com/micheleperna/JWST-NIRSpec_wiggles .

⁵<https://github.com/antoinedumontneira/WICKED> .

formed the extraction routine described above on the 2D spectrum in each dither, giving us an integrated, core, and outskirt spectrum for each dither. We masked outliers in the flux due to cosmic ray impacts or bad pixels using the outlier detection algorithm from Slob et al. (2024). For seven galaxies, the dithers encompass the same spatial regions. In these cases, we mean-stacked the integrated, core, and outskirt spectra for each dither, respectively, leaving us with three 1D spectra (one integrated, core, and outskirt spectrum for each galaxy). As a consequence of the MSA geometry, the two dithers for the remaining galaxy (130040) were performed in different positions on the galaxy. This is because its centre was exactly between two shutters, and thus each dither encloses a different spatial region. Therefore, for 130040 only, we analysed the core and outskirt spectra for each dither separately. We found that the S/N in the outskirts of dither 1 was too low to robustly fit stellar population parameters (see Section 4.2 and Appendix 4.A). This lower S/N could be due to astrometric errors, implying that the slit in dither 2 was better centred on the galaxy than in dither 1 and than indicated by the image in Fig. 4.1. Thus, we only considered dither 2 for this galaxy. We show our defined core and outskirts regions for each galaxy in Fig. 4.1 (with only the dither 2 regions shown for 130040).

We show the stacked integrated, core, and outskirt spectra in Fig. 4.2. We stacked all of our extracted, continuum-normalised (where we divided by a seventh order polynomial), rest-frame integrated, core, and outskirt spectra, respectively, by resampling the spectra to the same wavelength array via SPECTRES (Carnall 2017), normalising them by the region between 5200 – 5600 Å, and smoothing them to the maximum velocity dispersion of our sample ($\sigma \sim 390 \text{ km s}^{-1}$). We also show the fractional differences in the bottom panel. The three spectra have visible variations relative to each other⁶. To emphasise these differences, we zoom in on the H β , Mgb, Fe52, and Fe53 strong absorption features. The four features demonstrate significant differences between the core and outskirts. The H β feature is largely sensitive to age, indicating that the average core stellar population has a different age compared to the average outskirt stellar population. The Fe52 and Fe53 features are mainly sensitive to [Fe/H] but are also affected by age due to the age-metallicity degeneracy (Worthey 1994; Bruzual & Charlot 2003; Gallazzi et al. 2005). Interestingly, the differences between the core and the outskirt spectra are opposite in these two features (i.e. Fe52 is stronger in the core while Fe53 is stronger in the outskirts). Meanwhile, the Mgb feature is primarily sensitive to [Mg/H], indicating that the Mg abundances of the core and outskirt stellar populations may be different. We discuss this further in Section 4.4.

4.3.3 Determining distances and radii

To present our spatially resolved measurements as a function of radius, we accounted for several factors. These included: (i) the convolved nature of the observations, (ii) the fact that our galaxies may be elliptical in shape, and (iii) the

⁶Interestingly, there are large differences between the spectra in some of the molecular bands between $\sim 7000 - 8000 \text{ \AA}$. It is beyond the scope of this work to discuss these bands in detail. This will be the subject of future work.

arbitrary location of the micro shutters on each galaxy.

First, our observations are blurred due to convolution with the NIRSpec PSF. To account for this effect and ensure consistency with the parameters derived from the PSF-convolved observations, we derived the convolved R_e ($R_{e,\text{conv}}$) for each galaxy. We used a similar method to those presented in Price et al. (2016) and Cheng et al. (2024), where the intrinsic R_e ($R_{e,\text{int}}$), measured by fitting images with GALFIT (Peng et al. 2010) is also convolved with the NIRSpec PSF. Specifically, we took the structural GALFIT parameters measured from *JWST*/NIRCam F150W images by Slob et al. (2025) (or those measured from COSMOS *HST*/ACS F814W images by Griffith et al. 2012 for 130040) and used these parameters to create an intrinsic galaxy image in GALFIT for each galaxy. We created these images at an increased spatial resolution ($0''.05$ per pixel) compared to our observations. We convolved each mock galaxy image with the NIRSpec-MSA PSF, generated with STPSF⁷ (Perrin et al. 2015), keeping all shutters open⁸. We performed aperture sums on a series of elliptical apertures with increasing major and minor axes that we placed on the convolved model image using PHOTUTILS (Bradley et al. 2024), maintaining the intrinsic axis ratio measured with GALFIT in Slob et al. (2025) (or Griffith et al. 2012 for 130040). We determined the major axis enclosing 50% of the light for the convolved model and took this to be our $R_{e,\text{conv}}$. We report $R_{e,\text{conv}}$ for each galaxy in Table 4.1.

To determine the average distance to the centre for our core and outskirts regions in units of $R_{e,\text{conv}}$, we also considered that our galaxies are elliptical in shape on the sky, and are not perfectly round. Additionally, we took into account the arbitrary placement of the micro shutters on each galaxy, which are typically not centred or aligned along the major axis. To achieve this, we first determined the location of each spectral row on the sky. We interpolated the coordinates of each row, by assuming each row was represented by a rectangle of width $0''.1$ (the MSA pixel scale) and height $0''.2$ (the open width of a micro shutter, Ferruit et al. 2022), calculating the distance between each corner of the rectangle, and dividing this distance into ~ 45000 subpixels. We found that this number of subpixels was approximately where our calculated distances stabilised, but we note that the precise number of subpixels does not affect our conclusions. We generated a series of ellipses with increasing $R_{e,\text{conv}}$ and b_{conv} centred on the target galaxy, maintaining the axis ratio. Note that we disregarded the effect of the PSF on the ellipticity, and that smaller, inner ellipses will become rounder than outer ellipses. We determined the $R_{e,\text{conv}}$ of each subpixel from the dimensions of the respective intersecting ellipse. The $R_{e,\text{conv}}$ of that pixel was then the median $R_{e,\text{conv}}$ of all of the subpixels in the row. To determine the average $R_{e,\text{conv}}$ of each spatial bin, we took the weighted mean of the $R_{e,\text{conv}}$ of the rows in each region (core and outskirts), using the flux of each row as the weights. Finally, we

⁷<https://stpsf.readthedocs.io/en/latest/>.

⁸Ideally, we would generate a PSF with the exact number of shutters open as in the SUSPENSE observations. However, STPSF currently only offers simulated PSFs with one shutter, three adjacent shutters, or all shutters open. We tested generating PSFs with three shutters open. Our resulting convolved R_e measurements were consistent with keeping all shutters open, and thus a simulated PSF with the same number of shutters open as the observations is not likely to impact our conclusions.

took the mean of the $R_{e,\text{conv}}$ in each of the two observed dithers to get the final de-projected distance to the core and outskirts in units of $R_{e,\text{conv}}$. The dithers are at approximately the same position on each galaxy (other than for 130040, see above), but we took the mean of the $R_{e,\text{conv}}$ in each dither to account for minor positional deviations. Throughout the rest of the paper, we note that while we refer to ‘core’ and ‘outskirt’ regions somewhat arbitrarily for convenience, we present our results in units of the $R_{e,\text{conv}}$ that we define here. Thus, ‘core’ regions can be taken to represent stellar populations at smaller distances (normalised to $R_{e,\text{conv}}$), while ‘outskirt’ regions can be taken to represent stellar populations at larger distances (normalised to $R_{e,\text{conv}}$).

4.3.4 Full spectrum fitting

We fit the spectra using ALF α ⁹ (Beverage 2024; Beverage et al. 2025), a Python version of the ABSORPTION LINE FITTER (ALF¹⁰, Conroy & van Dokkum 2012a; Conroy et al. 2018). The ALF and ALF α models are built on empirical simple stellar populations, constructed using the Mesa Isochrones and Stellar Tracks (MIST, Choi et al. 2016) and the Spectral Polynomial Interpolator¹¹ (SPI, Villaume et al. 2017b). ALF and ALF α make use of stellar libraries, including the Medium Resolution INT Library of Empirical Spectra (MILES, Sánchez-Blázquez et al. 2006), the Extended Infrared Telescope Facility stellar library (E-IRTF, Villaume et al. 2017b), the Mann et al. (2015) sample of M-dwarf spectra, and the theoretical C3K stellar library (see Conroy & van Dokkum 2012a). To fit the spectra, ALF α removes the continuum, by fitting a high-order Chebyshev polynomial to the ratio of the data to the model, and implements DYNESTY (Speagle 2020) or EMCEE (Foreman-Mackey et al. 2013) to sample the posteriors of 20 stellar population parameters. ALF and ALF α can fit spectra with wavelengths ranging from 3700 – 24000 Å and stellar populations older than 1 Gyr. For details, see Conroy & van Dokkum (2012a), Conroy et al. (2018), and Beverage et al. (2025). We note that ALF α has been thoroughly tested and shown to produce the same results as ALF (Beverage et al. 2025).

Prior to fitting, we smoothed the ALF α models to the instrumental resolution of the NIRSpec-MSA observations, where we made use of the instrumental resolution derived via MSAFIT (de Graaff et al. 2024b) by Slob et al. (2024). Slob et al. (2024) find that the wavelength dependence of the instrumental resolution is typically a factor of 1.3 better than pre-launch estimates from JDox. We assumed this corrected JDox curve when fitting. See Slob et al. (2024) and Beverage et al. (2025) for details. We also masked the NaD absorption feature (which can be affected by the interstellar medium, Conroy & van Dokkum 2012a), [OIII] lines, and H α + [NII] complex, where present, as in Beverage et al. (2025). Similar to Beverage et al. (2025), we excluded regions between 6400 - 8000 Å due to broad TiO absorption in the spectra. Thus, where possible, we fit the wavelength regions 3700 - 4700 Å, 4700 - 5100 Å, 5100 - 5800 Å, 5800 - 6400 Å, and 8000 - 8600 Å.

⁹<https://github.com/alizabeverage/alfalpha>.

¹⁰<https://github.com/cconroy20/alf/tree/master>.

¹¹https://github.com/AlexaVillaume/SPI_Utils.

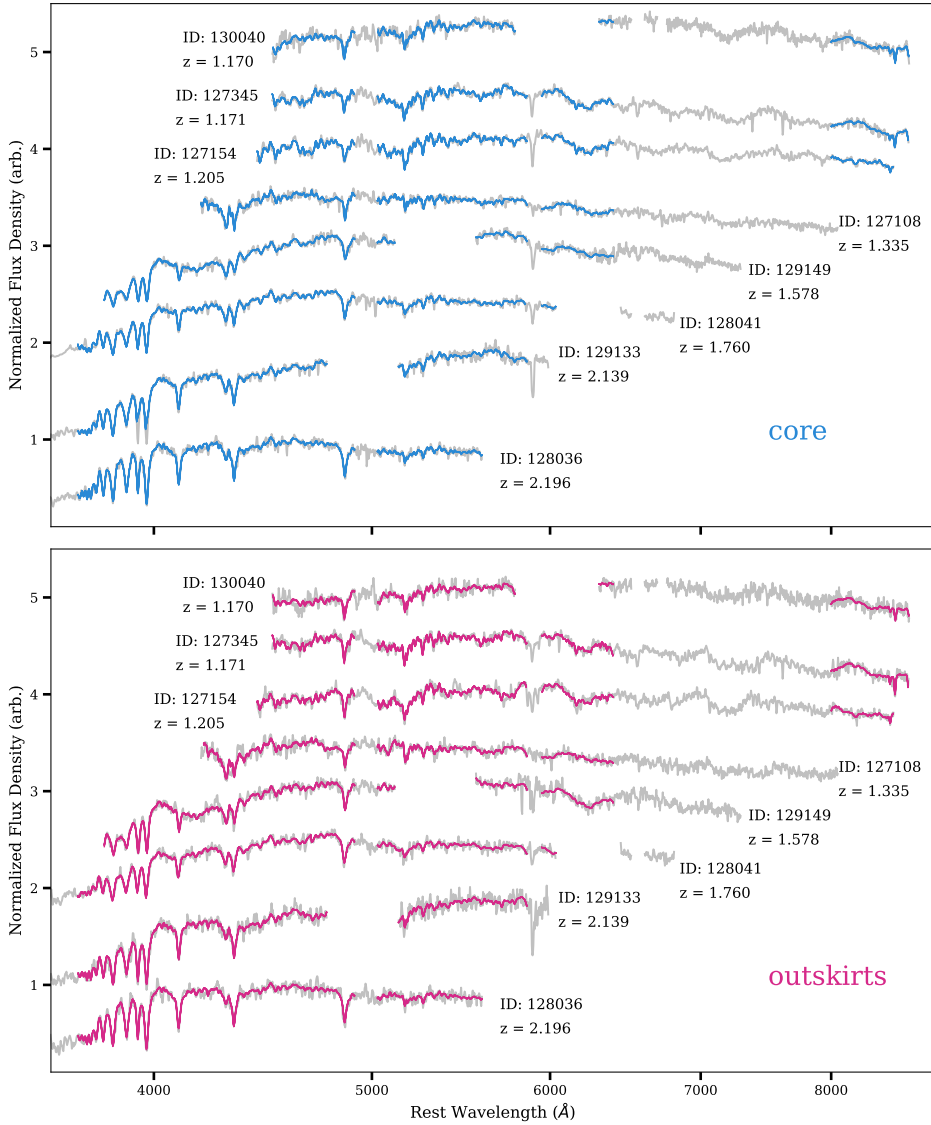


Figure 4.3: Best-fitting ALFA models to our quiescent galaxy spectra. The core spectra (grey lines) are shown in the top panel in order of increasing redshift from top to bottom, with their best-fit models overplotted (blue lines). The outskirt spectra (grey lines) and fits (magenta lines) are similarly shown in the bottom panel. We normalised each spectrum by its median value between $4480 - 4520 \text{ \AA}$ and arbitrarily offset them in the y -direction for visibility. We do not show emission lines. This figure illustrates that many absorption lines are robustly detected for both the core and outskirt regions.

Table 4.2: Results from our ALFA fits to the spatially resolved *JWST*-SUSPENSE spectra.

ID	log(Age/Gyr)			[Fe/H]			[Mg/Fe]			[Mg/H]		
	core	outsirt	gradient	core	outsirt	gradient	core	outsirt	gradient	core	outsirt	gradient
130040	0.50 ^{+0.19} _{-0.22}	0.43 ^{+0.22} _{-0.17}	-0.18 ^{+0.39} _{-0.53}	-0.35 ^{+0.19} _{-0.23}	-0.41 ^{+0.20} _{-0.22}	-0.16 ^{+0.63} _{-0.64}	-0.02 ^{+0.23} _{-0.17}	0.33 ^{+0.20} _{-0.22}	0.92 ^{+0.70} _{-0.72}	-0.37 ^{+0.23} _{-0.22}	-0.09 ^{+0.23} _{-0.27}	0.23 ^{+0.57} _{-0.57}
127345	0.54 ^{+0.15} _{-0.19}	0.35 ^{+0.10} _{-0.08}	-0.55 ^{+0.39} _{-0.38}	-0.08 ^{+0.14} _{-0.17}	-0.07 ^{+0.12} _{-0.22}	0.01 ^{+0.46} _{-0.49}	0.21 ^{+0.12} _{-0.13}	0.33 ^{+0.15} _{-0.19}	0.35 ^{+0.53} _{-0.54}	0.13 ^{+0.16} _{-0.23}	0.25 ^{+0.15} _{-0.32}	0.03 ^{+0.42} _{-0.42}
127154	0.66 ^{+0.12} _{-0.17}	0.13 ^{+0.11} _{-0.09}	-2.46 ^{+0.48} _{-0.46}	-0.06 ^{+0.09} _{-0.09}	0.14 ^{+0.15} _{-0.27}	0.96 ^{+0.76} _{-0.79}	-	-	-	-	-	-
127108	0.58 ^{+0.07} _{-0.13}	0.52 ^{+0.15} _{-0.14}	-0.41 ^{+0.45} _{-0.44}	-0.55 ^{+0.14} _{-0.15}	-0.39 ^{+0.18} _{-0.22}	0.54 ^{+0.95} _{-0.95}	0.26 ^{+0.20} _{-0.20}	0.59 ^{+0.17} _{-0.23}	1.27 ^{+1.02} _{-1.01}	-0.29 ^{+0.19} _{-0.20}	0.20 ^{+0.16} _{-0.26}	1.13 ^{+0.73} _{-0.79}
129149	0.24 ^{+0.08} _{-0.07}	0.21 ^{+0.09} _{-0.08}	-0.11 ^{+0.21} _{-0.25}	-0.12 ^{+0.12} _{-0.17}	-0.25 ^{+0.27} _{-0.23}	-0.38 ^{+0.61} _{-0.70}	-	-	-	-	-	-
128041	0.14 ^{+0.04} _{-0.06}	0.09 ^{+0.09} _{-0.09}	-0.41 ^{+0.44} _{-0.45}	-0.02 ^{+0.16} _{-0.16}	-0.08 ^{+0.23} _{-0.17}	-0.33 ^{+1.19} _{-1.19}	0.10 ^{+0.16} _{-0.14}	0.13 ^{+0.22} _{-0.21}	0.25 ^{+1.72} _{-1.76}	0.06 ^{+0.15} _{-0.20}	0.03 ^{+0.22} _{-0.29}	0.05 ^{+1.42} _{-1.47}
129133	0.15 ^{+0.05} _{-0.04}	0.03 ^{+0.06} _{-0.07}	-0.31 ^{+0.16} _{-0.15}	-0.20 ^{+0.11} _{-0.10}	-0.16 ^{+0.17} _{-0.29}	0.12 ^{+0.49} _{-0.52}	-	-	-	-	-	-
128036	0.09 ^{+0.04} _{-0.04}	0.00 ^{+0.07} _{-0.04}	-0.31 ^{+0.14} _{-0.13}	-0.27 ^{+0.10} _{-0.22}	-0.27 ^{+0.17} _{-0.30}	0.04 ^{+0.59} _{-0.62}	0.18 ^{+0.17} _{-0.17}	0.30 ^{+0.21} _{-0.25}	0.43 ^{+0.85} _{-0.86}	-0.09 ^{+0.14} _{-0.24}	0.03 ^{+0.20} _{-0.36}	0.08 ^{+0.64} _{-0.70}

Notes. For integrated stellar population parameters from SUSPENSE, see Beverage et al. (2025). For the ‘gradient’ columns, we report $\Delta\text{Parameter}/\Delta\log_{10}(R_{e,\text{conv}})$.

Finally, in general, NIRSpec-MSA noise has been found to be underestimated (see e.g. Maseda et al. 2023). Thus, we multiplied our noise spectra by the jitter term (an average factor of 2.1) fit in Beverage et al. (2025) for each SUSPENSE galaxy, prior to fitting.

In our implementation of ALF α , we used DYNESTY to sample the posteriors of the velocity offset, velocity dispersion, single stellar population-equivalent age, isochrone metallicity, ten elemental abundances (Fe, C, N, Mg, Na, Si, Ca, Ti, and Cr), Balmer emission line flux, emission line velocity and broadening, a shift in the T_{eff} of the fiducial isochrones, and an instrumental jitter term to account for over or underestimation of the observed uncertainties. We assumed a Kroupa (2001) IMF. We set the upper limit of the age prior to be the age of the Universe at the redshift of each galaxy, plus 2 Gyr to allow for uncertainties. We assumed the default priors for all other parameters. We examined the output ages and ensured that they are ≥ 1 Gyr, as ALF α is only suitable for stellar populations older than 1 Gyr.

We first fit the integrated spectrum, allowing for variation in all stellar parameters listed above. For the core and outskirt spectra, however, we could not constrain all of the elemental abundances at this S/N (we determined this by fitting mock observations similar to Cheng et al. 2024, see Appendix 4.A). Thus, as in Cheng et al. (2024), we fixed the values of all abundances other than Fe and Mg to the values from the integrated fits in our core and outskirt fits. In this way, we only allowed age, Fe, and Mg to vary so that we could accurately constrain these parameters. We visually inspected the results for each individual galaxy to ensure that the posterior distributions for age, Fe, and Mg were well-sampled. If these posteriors ran up against the priors, we extended the priors for $[Z/H]$, $[Fe/H]$, and $[Mg/H]$ as needed by up to 0.2 dex. If the re-fitted age, $[Z/H]$, or $[Fe/H]$ posteriors still ran up against the priors, we discarded these galaxies from our sample. For two galaxies (127154 and 129133), we found that the $[Mg/H]$ posteriors ran up against the extended priors in at least one of the three fits for these objects. Thus, for these objects only, we also fixed the Mg abundance in the core and outskirt fits to the value recovered in the integrated fits, and we do not report their Mg gradients. We also do not report the Mg gradient for 129149 as there are no detectable Mg features in the spectra of this galaxy. We show our core and outskirt ALF α fits in Fig. 4.3.

We compared our integrated fit results with those of Beverage et al. (2025) for the eight galaxies that we have in common (not shown). Our results are consistent within $\sim 0.2\sigma$, despite differences in the treatment of the data (i.e. the wiggle correction, see Section 4.3.1) and spectral extraction (see Slob et al. 2024).

4.4 Results

4.4.1 Spatially resolved ages and elemental abundances

In Fig. 4.4, we show the ages and elemental abundances in the cores and outskirts of eight massive quiescent galaxies at $1.2 \lesssim z \lesssim 2.2$, derived by fitting the spectra with ALF α . We show spatially resolved stellar population parameters as a function

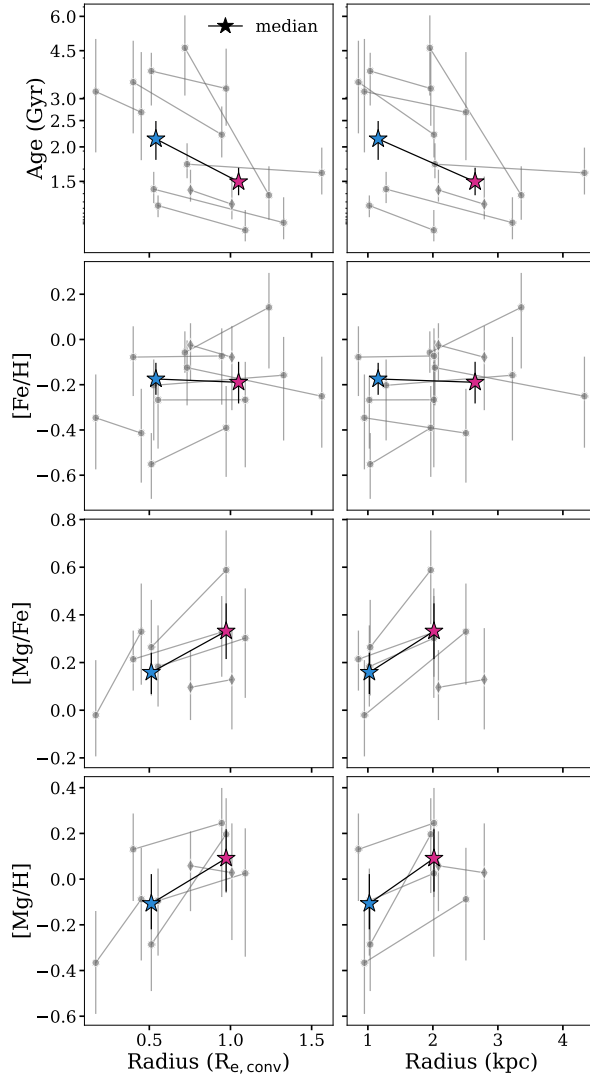


Figure 4.4: Spatially resolved stellar population parameters derived from our ALFA full spectrum fits. In the left column, we plot our measured parameters as a function of de-projected, convolved radius in units of $R_{e,conv}$ (determined as described in Section 4.3.3). In the right column, we plot our measured parameters as a function of de-projected, convolved radius in units of kpc. We show age in the top row, [Fe/H] in the second row, [Mg/Fe] in the third row, and [Mg/H] in the bottom row. In each panel, we show the measured parameters for each individual galaxy as points connected by lines. For galaxy 128041, we note that the spectrum representing its core may partially probe its outskirts regions (see Fig. 4.1 and Table 4.1) and show this galaxy as a diamond. We show the median parameters of all galaxies as stars connected by lines.

of de-projected radius from the centre of each galaxy. We report these values in Table 4.2. We show galaxy 128041 as a diamond, as the spectrum representing its core may also partially probe its outskirt regions. To estimate the uncertainties on the median values (stars), we performed a simple Monte Carlo simulation, where we took the median of the data points perturbed around the errors 1000 times, and we show the 16th and 84th percentiles of this distribution.

Fig. 4.4 indicates that massive quiescent galaxies at $1.2 \lesssim z \lesssim 2.2$ from SUSPENSE tend to have negative age gradients, with galaxies having older cores compared to their outskirts. Some galaxies may also have positive [Mg/H] gradients, with galaxy cores possibly being deficient in Mg compared to their outskirts (see Section 4.4.2 and Fig. 4.5 as well). We also find indications of positive [Mg/Fe] gradients, which may suggest that galaxy cores have longer SF timescales compared to galaxy outskirts (e.g. Matteucci 1994; Maiolino & Mannucci 2019, although note that the [Mg/H] and [Mg/Fe] gradients are overall consistent with being flat, within uncertainties). We discuss this further in Section 4.5.1.1. The individual SUSPENSE galaxies have diverse [Fe/H] gradients, although on average the [Fe/H] gradients are consistent with being flat, indicating that the cores and outskirts of these galaxies have similar metallicities. The fact that our fits prefer the stellar populations at smaller $R_{e,\text{conv}}$ to be older qualitatively agrees with the shallower H β feature that we see in the inset panel in Fig. 4.2. Similarly, the fact that we may observe stellar populations at smaller $R_{e,\text{conv}}$ to have less Mg qualitatively agrees with the shallower Mg b feature. Additionally, the contrasting differences in the core and outskirt spectra that we see in the Fe52 and Fe53 features could be contributing to our observed flat [Fe/H] gradient. We discuss the implications of these results for galaxy formation and assembly in Section 4.5.

It is interesting to note that our results differ from gradients measured at lower z . In particular, massive quiescent galaxies typically have flat age and [Mg/Fe] gradients, and negative metallicity gradients out to at least $z \sim 1$ (e.g. Mehlert et al. 2003; Greene et al. 2015; Martín-Navarro et al. 2018; Cheng et al. 2024; Parikh et al. 2024). This suggests that stellar population gradients in massive quiescent galaxies may evolve from what we find at $1.2 \lesssim z \lesssim 2.2$ to what has been found at $z \lesssim 1$. In this context, it is interesting to note that the smallest quiescent galaxies in the local Universe (i.e. relic galaxies), have been found to have positive [Mg/Fe] gradients (Martín-Navarro et al. 2018), potentially qualitatively similar to what we see here. We discuss this further in Section 4.5.

There are also a handful of other studies that have measured spatially resolved spectroscopic stellar population gradients in massive quiescent galaxies beyond $z \sim 1$. Specifically, Jafariyazani et al. (2020) measure a flat age and [Mg/Fe] gradient, and a negative [Fe/H] gradient in a massive lensed galaxy at $z = 1.98$, which is qualitatively consistent with results at $z \lesssim 1$. Akhshik et al. (2023) find diverse age gradients in eight lensed galaxies at $1.6 < z < 2.9$. Ditrani et al. (2022) measure age and metal gradients in four galaxies at $1.6 < z < 2.4$ and find flat age and negative metallicity gradients, again consistent with results at $z \lesssim 1$. Finally, Pérez-González et al. (2025) find a strong negative total metallicity gradient in an individual galaxy at $z \sim 3.7$ and D'Eugenio et al. (2024) measure a constant age as a function of radius in an individual galaxy at $z \sim 3$. It is not

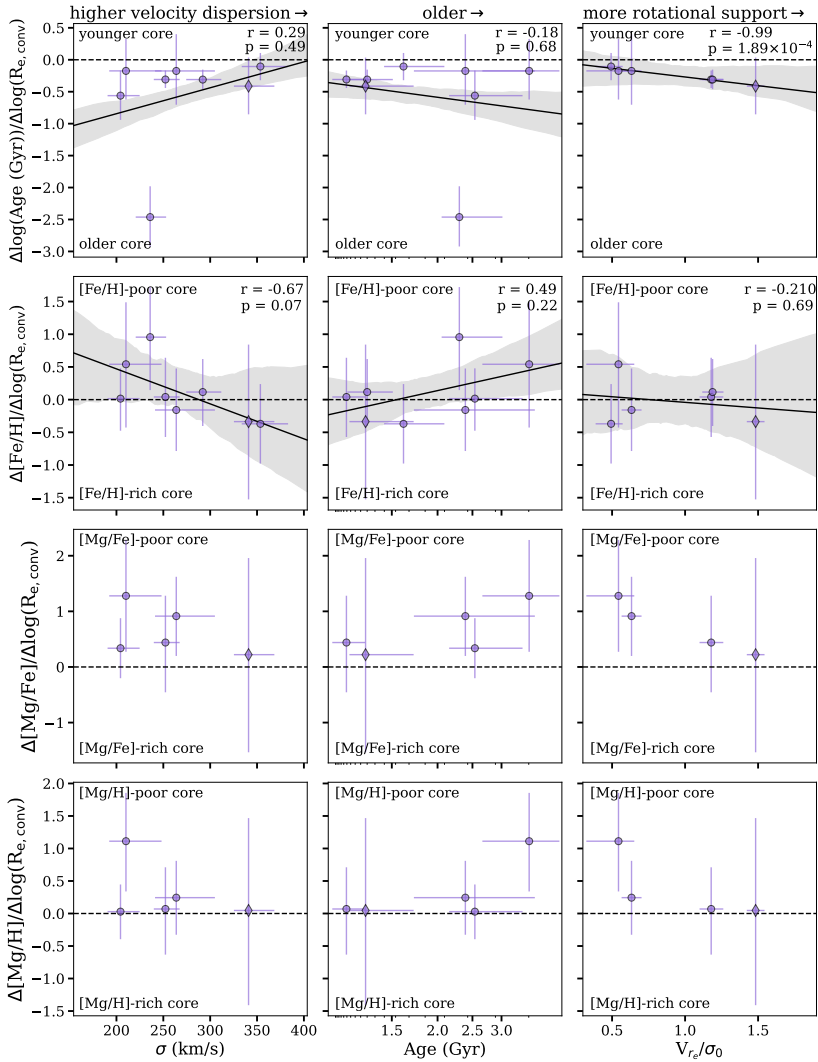


Figure 4.5: Spatially resolved stellar population gradients, normalised by $R_{e,conv}$ (see Section 4.3.3), as a function of galaxy parameters. We show the age gradients in the top row, the $[\text{Fe}/\text{H}]$ gradients in the second row, the $[\text{Mg}/\text{Fe}]$ gradients in the third row, and the $[\text{Mg}/\text{H}]$ gradients in the bottom row. We indicate 128041 as a diamond (as in Figure 4.4), as its core spectrum partially extends into its outskirts. The outlying object with a strongly negative age gradient is 127154. We show the integrated velocity dispersion (σ , from our ALFA fits to the integrated spectra) in the left column, the integrated age (from our integrated ALFA fits) in the middle, and V_{re}/σ_0 (measured in Slob et al. 2025) on the right. Horizontal dashed lines indicate where flat gradients would lie. We performed a linear fit to the points in the age and $[\text{Fe}/\text{H}]$ panels, shown by the solid lines, and 1σ bootstrapped uncertainties are indicated by the shaded regions.

surprising that different works find diverse gradients beyond $z \sim 1$ as different methods are used in each of these studies. Additionally, the sample sizes for which stellar population gradients can be measured at these redshifts (both in our work and in previous studies) are very small. Finally, previous work examines relatively younger quiescent galaxies, which could explain some of these differences. We discuss this further in Section 4.4.2.

4.4.2 Trends between gradients and galaxy parameters

As we have age and [Fe/H] gradients for our whole sample, we examined if there are any correlations with global velocity dispersions, integrated ages, and spatially resolved stellar kinematics. These parameters are commonly used as measures of evolutionary stage. In Fig. 4.5, we show the gradient slopes as a function of different galaxy parameters. We fit a line to the points in the age and [Fe/H] panels using the non-linear least squares algorithm implemented in the `SCIPY curve_fit` function, in the `optimize` module (Virtanen et al. 2020). We computed the Pearson- r correlation coefficient and corresponding p -value, which we state in each panel. We note that we also show our [Mg/Fe] and [Mg/H] gradient slopes in the bottom two rows, but we did not fit a line to these points as we only have Mg abundances for 5 galaxies.

In the left column, there is a negative correlation with [Fe/H] gradient. Thus, galaxies with higher velocity dispersions have more negative [Fe/H] gradients. We note, however, that further investigation is required to understand the significance of this apparent trend. This tentative trend is in contrast with results at low- z (see, e.g. Sánchez-Blázquez et al. 2007; Kuntschner et al. 2010; Spolaor et al. 2010; Pastorello et al. 2014; González Delgado et al. 2015; Greene et al. 2015; Martín-Navarro et al. 2018; Ferreras et al. 2019; Santucci et al. 2020. For example, Spolaor et al. (2010) and Ferreras et al. (2019) find positive correlations between metallicity gradients and mass/velocity dispersion, while Sánchez-Blázquez et al. (2007), Pastorello et al. (2014), González Delgado et al. (2015), and Greene et al. (2015) find no dependence of metallicity gradients on mass.

In the middle column of Fig. 4.5, there is a mild positive trend between [Fe/H] gradients and age, although this trend is also statistically insignificant. Thus, younger quiescent galaxies may have more negative [Fe/H] gradients, while older quiescent galaxies may have more positive [Fe/H] gradients. Interestingly, this tentative trend is qualitatively similar to the results of Cheng et al. (2024), who find a positive trend between [Fe/H] gradients and age in quiescent galaxies at $0.6 \lesssim z \lesssim 1.0$. However, we note that Cheng et al. (2024) also find a negative trend between age gradients and age, which is inconsistent with our findings here (although we sample a slightly younger age range).

In this context, the tentative trends that we find between [Fe/H] gradients and age may be qualitatively consistent with previous spatially resolved studies at $z > 1$. In particular, Jafariyazani et al. (2020), Pérez-González et al. (2025), and D'Eugenio et al. (2020) all examine relatively young galaxies and find flat/slightly negative age gradients and negative metallicity gradients. We may be seeing a similar effect in our youngest galaxies (middle column of Fig. 4.5).

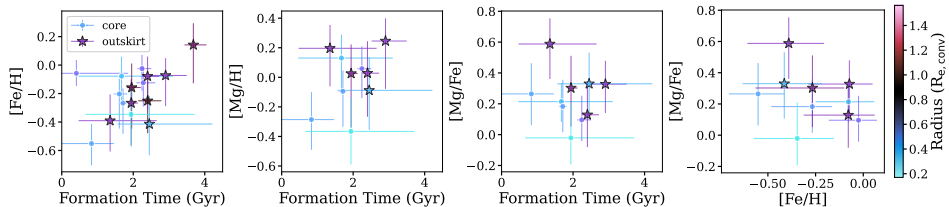


Figure 4.6: Individual core and outskirts elemental abundances. In the first three panels, we show the formation time on the x -axis, which we computed by correcting the stellar age (from our ALFA fits to the integrated spectra) by the age of the Universe at the redshift of each galaxy. We show our $[\text{Fe}/\text{H}]$ abundances in the first panel, $[\text{Mg}/\text{H}]$ abundances in the second panel, and $[\text{Mg}/\text{Fe}]$ abundances in the third panel. In the fourth panel, we show $[\text{Mg}/\text{Fe}]$ as a function of $[\text{Fe}/\text{H}]$. In each panel, core values are shown as points, and outskirts values are shown as stars, and we colour-coded the symbols by their average distance to the centre of the galaxy in units of $R_{e,conv}$.

Finally, in the right column of Fig. 4.5, we compare our gradients to spatially resolved stellar kinematics measured in Slob et al. (2025). In particular, Slob et al. (2025) derive their kinematic measurements from the NIRSpect-MSA spectra using a forward modelling approach. We show our stellar population gradients as a function of their V_{r_e}/σ_0 , where a higher value of V_{r_e}/σ_0 indicates more rotational support. We find a significant negative correlation between our age gradients and V_{r_e}/σ_0 . Thus, galaxies with more rotational support may have more strongly negative age gradients (with their cores being much older than their outskirts). On the other hand, Greene et al. (2019) find no significant correlations between stellar population gradients and the ratio of rotational to dispersion support in massive local elliptical galaxies. We discuss our trends in Section 4.5.

4.5 Discussion

4.5.1 Implications for galaxy evolution

Below $z \sim 1$, resolved studies have revealed that massive quiescent galaxies tend to have flat age and α -element abundance gradients and negative metallicity gradients (e.g. Mehlert et al. 2003; Greene et al. 2015; Martín-Navarro et al. 2018; Cheng et al. 2024). These findings are consistent with inside-out growth, where outskirts are built up by the accretion of low-mass, low-metallicity satellites (minor mergers; e.g. Bezanson et al. 2009; Naab et al. 2009; van de Sande et al. 2013; Haryana et al. 2025). Alternatively, the observed evolution in galaxy properties over cosmic time can be due to progenitor bias. In this scenario, where the quiescent population grows by the addition of larger galaxies at later times (van Dokkum & Franx 2001; Carollo et al. 2013; Poggianti et al. 2013), gradients may come directly from the star-forming progenitors (Cheng et al. 2024, see Jones et al. 2015; Tripodi et al. 2024; Shen et al. 2024; Ju et al. 2025), or could be produced by the quenching mechanism. Here, we measured resolved stellar popu-

lation gradients out to $z \sim 2.2$. Our galaxies were observed early, in the peak of the quenching and assembly era ($z \sim 1 - 3$; e.g. Oser et al. 2010; Whitaker et al. 2012; Rodriguez-Gomez et al. 2016), and are more compact compared to galaxies at $z \lesssim 1$. These characteristics suggest that they are relatively unpolluted by potential late-time mergers compared to nearby quiescent galaxies. Their gradients may thus provide interesting clues to quenching processes at $z \sim 2$ and subsequent assembly processes towards $z \lesssim 1$.

4.5.1.1 Quenching at $z \sim 2$

The negative age gradients that we find may be the imprint of inside-out quenching, where star formation first occurs and/or terminates in the central regions of galaxies, followed by later-time star formation and/or quenching in the outskirts (e.g. Dekel & Burkert 2014; Tacchella et al. 2015a,b, 2016a, 2018; Zolotov et al. 2015; Zibetti et al. 2020). This scenario may be supported by our [Mg/H] gradients, which may be regulated by whether and how quickly the gas supply in the centres of the galaxies was depleted during the quenching process (see e.g. Ellison et al. 2018; Spilker et al. 2019; Trussler et al. 2020; Beverage et al. 2021). In particular, Beverage et al. (2021) compare different chemical evolution models from Spitoni et al. (2017) and find that abrupt gas expulsion prevents the galaxy from further enriching, leading to lower metallicities for stellar populations that quench earlier, in agreement with observations (e.g. Kriek et al. 2016; Kriek et al. 2019; Jafariyazani et al. 2025; Beverage et al. 2023, 2024, 2025; Carnall et al. 2022; Zhuang et al. 2023). This scenario may thus explain our observed [Mg/H] abundances if the cores quenched more abruptly (perhaps due to AGN outflows; e.g. Croton et al. 2006; Bluck et al. 2014; Zolotov et al. 2015; Tacchella et al. 2015a, 2016a; Belli et al. 2024; Park et al. 2024, although we note that our [Mg/H] gradients are also consistent with being flat within uncertainties).

The inside-out quenching scenario presented here is reinforced by Fig. 4.6, where we examine the general spatially resolved stellar populations in SUSPENSE as a function of formation time. First, similar to Beverage et al. (2025), we find that earlier-forming stellar populations may have slightly lower [Fe/H] and [Mg/H]. Moreover, stellar populations at smaller average distances from the centre of the galaxy tend to have formed earlier than stellar populations at larger average distances and may be less enriched by the time they stop forming stars, with tentative trends between [Fe/H] and distance from the centre. Thus, this Figure reinforces the idea that the inner regions of the SUSPENSE galaxies may have formed earlier and also quenched more efficiently at earlier times.

Although the scenario described above can explain the age gradients and perhaps some of the [Mg/H] gradients as well, it is puzzling that we find flat [Fe/H] gradients, as we would also expect more iron-deficient cores in this scenario. Additionally, some of our [Mg/Fe] gradients may suggest that the star-formation timescale in the core is longer than in the outskirts (see e.g. Matteucci 1994; Maiolino & Mannucci 2019). This is not what we expect from inside-out quenching, in the case that both populations started forming at the same time. However, it is possible that star formation in the outskirts also started much later and oc-

curred over a shorter time compared to the core. It is also important to note that our measurements are luminosity-weighted, and thus our ages and [Mg/Fe] ratios, which may be driven by different wavelengths, may reflect slightly different stellar populations (although this effect may be minor). Moreover, [Mg/Fe] is also influenced by the star-formation efficiency, the IMF, outflows, and late-time mergers (see e.g. Tinsley 1979; Zolotov et al. 2010; Martín-Navarro 2016; Andrews et al. 2017; Sybilska et al. 2018). In any case, our sample is small, and larger samples are required to come to definitive conclusions.

Finally, we note that our results are inconsistent with a scenario in which galaxies quench their star formation via wet, gas-rich, major mergers. Major mergers are expected to result in outside-in quenching, by funnelling gas to the galaxy’s centre and triggering a central starburst. This would result in a positive age gradient (Hopkins et al. 2008; Snyder et al. 2011; Wellons et al. 2015; Pathak et al. 2021), which is generally inconsistent with our findings. Nonetheless, it is possible that different quenching mechanisms operate at different redshifts. For example, Cheng et al. (2024) find positive age gradients in their youngest massive quiescent galaxies at $0.6 \lesssim z \lesssim 1.0$, consistent with central starbursts triggered by wet mergers.

4.5.1.2 Galaxy evolution towards $z \lesssim 1$

Assuming that the SUSPENSE galaxies are representative of the progenitors of $z \lesssim 1$ massive quiescent galaxies, it is interesting to compare our results with those from lower redshift studies (e.g. Mehlert et al. 2003; Greene et al. 2015; Martín-Navarro et al. 2018; Cheng et al. 2024). In contrast to the flat age gradients at $z \lesssim 1$, we find negative age gradients at $1.2 \lesssim z \lesssim 2.2$. In addition, while negative [Fe/H] gradients are found at lower redshifts, we find flat [Fe/H] gradients. Finally, we find hints of positive [Mg/Fe] gradients, in contrast with the flat [Mg/Fe] gradients at $z \lesssim 1$. Thus, age gradients may evolve from negative to flat, [Fe/H] gradients may evolve from flat to negative, and [Mg/Fe] gradients could evolve from positive to flat between $z \sim 2$ and $z \sim 1$. This evolution could be due to inside-out growth via minor mergers, if the accreted galaxies are slightly older or have comparable ages to the central regions of distant quiescent galaxies, have lower [Fe/H], and possibly have lower [Mg/Fe]. In this scenario, the stars of these satellite galaxies will primarily be deposited in the outskirts and may flatten the age and [Mg/Fe] gradients that may have been generated by inside-out quenching, while simultaneously steepening the [Fe/H] gradients in the negative direction (see also Zibetti et al. 2020). We note that this scenario may not be fully consistent with the age-mass relation, where lower-mass galaxies are expected to be younger (e.g. Kauffmann et al. 2003). On the other hand, this expectation of lower-mass galaxies being younger may not hold for satellite galaxies (see, e.g. Pasquali et al. 2010; Gallazzi et al. 2021; Oyarzún et al. 2023).

This minor merger picture, which can also explain the difference in size between our galaxies and those at lower- z (e.g. Trujillo et al. 2004; Franx et al. 2008; Bezanon et al. 2009; van Dokkum et al. 2010; van der Wel et al. 2014) may be supported by some of the trends that we see with galaxy properties in Fig. 4.5. Specifically,

in the right column, we compare our resolved stellar population gradients to the six galaxies that we have in common with Slob et al. (2025), who measure spatially resolved stellar kinematics in 15 SUSPENSE galaxies. Interestingly, we find a significant negative trend between our age gradients and V_{re}/σ_0 , where galaxies with more rotational support also have stronger negative age gradients. This supports the scenario presented in Slob et al. (2025), where minor mergers gradually destroy rotational support over time while simultaneously flattening the age gradients (again, assuming that the accreted galaxies are older than the galaxy core). However, larger sample sizes are required to constrain any trends between [Fe/H] gradients and V_{re}/σ_0 , and between age and [Fe/H] gradients and galaxy age, and accordingly assess whether these are also consistent with this picture (see also D’Eugenio et al. 2024).

Alternatively, progenitor bias (van Dokkum & Franx 2001; Carollo et al. 2013; Poggianti et al. 2013) can also explain the evolution in gradients (as well as size) between the distant quiescent galaxies in our sample and those at $z < 1$. In this scenario, the galaxies that quench at later times would have the stronger negative [Fe/H] and flat age gradients that we see at lower z . These gradients could have been directly inherited from the star-forming progenitors and/or could indicate that different quenching processes operate at different times (see Section 4.5.1.1).

For cases where galaxies are not affected by minor mergers, the high- z , quiescent galaxies must survive as they are until the present day. Relic galaxies may represent this unique subpopulation, in which the early-forming cores of massive quiescent galaxies have passively evolved since $z \sim 2$ without accreting *ex situ* material (Trujillo et al. 2009; Quilis & Trujillo 2013; Stringer et al. 2015). Thus, they may represent the descendants of distant quiescent galaxies, such as those in SUSPENSE. Interestingly, and possibly similar to our findings, Martín-Navarro et al. (2018) find positive [Mg/Fe] and flat metallicity gradients in a sample of nearby relic galaxies, which could have been directly inherited from the population of distant quiescent galaxies studied in this work. On the other hand, our age gradients may be inconsistent with the finding of flat age gradients from Martín-Navarro et al. (2018). Thus, it is likely that multiple mechanisms (i.e. progenitor bias and minor mergers) are playing a role (see also van der Wel et al. 2008; Hopkins et al. 2010; Newman et al. 2012; Oser et al. 2012).

4.5.2 Caveats

In this work, we presented age and elemental abundance gradients in eight massive quiescent galaxies at $1.2 \lesssim z \lesssim 2.2$. We emphasise that this is the first study of its kind, with the largest sample size to date. Thus, we have demonstrated that it is possible to measure stellar population gradients in galaxies beyond $z \sim 1$ using the *JWST*/NIRSpec-MSA. However, there are some caveats that should be taken into consideration.

There are a few observational factors affecting our data, including the limited spatial resolution and convolved nature of the MSA spectra, and the fact that the arbitrary position of the micro shutters does not always allow us to capture each galaxy’s precise centre. Despite our efforts to determine the exact regions of the

galaxies contained within each micro shutter (see Section 4.3.3), these factors likely result in some contamination from the outskirts in our core spectra and vice versa. However, any contamination is likely minor, as we detect gradients in our galaxies. This detection suggests that our measured gradients are real, as contamination from different regions of the galaxy will serve to weaken gradients (see Cheng et al. 2024). High spatial resolution follow-up with current or planned IFU instruments (i.e. the NIRSPEC-IFU, ERIS on the *VLT*, HARMONI on the *ELT*) may provide a more detailed view of assembly histories by allowing measurements of detailed stellar population profiles (see e.g. Oyarzún et al. 2019) and avoiding MSA slitlet offsets.

Our fitting technique could also be biasing our results, as we only rely on one code (*ALF α*). However, the stacked spectra that we show in Fig. 4.2 are indicative of negative age and positive [Mg/H] gradients, with a shallower core H β feature indicating an older age and a shallower Mgb feature potentially suggesting less [Mg/H] (although we note that Mgb is also sensitive to age). Additionally, the fact that the Fe52 and Fe53 features show opposing core and outskirt depths may support a flat [Fe/H] gradient. Thus, while the absolute values of our elemental abundances could be affected by our fitting technique, our qualitative gradients and our conclusions about them are presumably robust. We note that we do not consider dust in our analysis. However, in contrast to traditional stellar population synthesis (SPS) fitting results in the literature, our study does not suffer from degeneracies between dust and age/metallicity (Worthey 1994; Bell & de Jong 2001; Bruzual & Charlot 2003; Gallazzi et al. 2005), as dust does not affect the depth of absorption features and our measurements are based on continuum-normalised spectra.

It is also possible that the IMF is contributing to the potential [Mg/Fe] gradient as elemental abundances and abundance ratios are dependent on the IMF (e.g. Martín-Navarro 2016; Martín-Navarro et al. 2018; De Lucia et al. 2017; Fontanot et al. 2017). In this work, we assume a Kroupa (2001) IMF, but there are indications that the low-mass end (e.g. van Dokkum et al. 2017; Gu et al. 2022; Cheng et al. 2026) and perhaps also the high-mass end (e.g. Fontanot et al. 2017, see also van Dokkum & Conroy 2024) of the IMF in the SUSPENSE galaxies may deviate from the Kroupa (2001) form (see also Slob et al. 2025). In particular, a variable IMF can affect the ratio of Mg to Fe by changing the relative number of core-collapse supernovae (which primarily produce Mg) compared to Type Ia supernovae (which primarily produce Fe, Matteucci 1994; Thomas et al. 1999). In general, it is challenging to constrain the shape of the IMF, as well as its impact on elemental abundances. In the future, a robust characterisation of the slope of the IMF in distant quiescent galaxies, as well as detailed chemical evolution modelling to understand the effects of both star-formation timescales and the IMF on elemental abundance patterns, are required (see also Beverage et al. 2025; Cheng et al. 2026).

Finally, while this is the largest sample of $z \gtrsim 1$ galaxies for which gradients have been derived using this method so far, our sample size is still small, which makes it challenging to draw strong conclusions about the population of galaxies at these redshifts. Thus, in the future, it will be informative to make the same

measurements in a statistically significant sample of galaxies, perhaps with similar follow-up NIRSpec-MSA observations.

4.6 Summary and conclusions

Spatially resolving the stellar populations of massive quiescent galaxies at early times can reveal crucial information about how galaxies quenched their star formation and assembled their stellar populations. Previous work beyond the local Universe has been based on photometry or very small spectroscopic sample sizes. In this work, we leveraged the capabilities of the *JWST*/NIRSpec-MSA and presented age, [Fe/H], [Mg/H], and [Mg/Fe] gradients in eight massive quiescent galaxies at $1.2 \lesssim z \lesssim 2.2$ from *JWST*-SUSPENSE, an ultra-deep spectroscopic survey (Slob et al. 2024). These gradients were derived by fitting the central and outskirts spectra of each galaxy with ALFA, a flexible full-spectrum stellar population synthesis model (Conroy & van Dokkum 2012a; Conroy et al. 2018; Beverage et al. 2024, 2025). Our study accounted for the fact that the NIRSpec-MSA micro shutters were not centred or aligned on the galaxies in our sample as well as for spatial undersampling of the *JWST* PSF. In summary, we found the following:

- Massive quiescent galaxies at $1.2 \lesssim z \lesssim 2.2$ from SUSPENSE tend to have negative age gradients and flat [Fe/H] gradients, and they may have positive [Mg/H] and [Mg/Fe] gradients (although we note that the last two properties are also consistent with being flat within uncertainties). In other words, the cores of the SUSPENSE galaxies are generally older, possibly Mg enhanced, and have approximately the same [Fe/H] compared to their outskirts.
- The negative age gradients suggest inside-out quenching (e.g. Tacchella et al. 2016a, 2018), where star formation terminated early in galaxy cores while galaxy outskirts continued to form stars until later times. The possible low central [Mg/H] may indicate rapid central gas expulsion (e.g. Beverage et al. 2021, 2025), further supporting the inside-out quenching scenario. Thus, galaxy cores formed earlier and may have stopped their star formation more abruptly, preventing further enrichment. Meanwhile, galaxy outskirts formed later and may have experienced more enrichment before quenching. However, it is unclear how our [Fe/H] and [Mg/Fe] gradients fit into this picture.
- Our gradients differ from those at lower- z , where studies typically find flat age and [Mg/Fe] gradients and negative [Fe/H] gradients (e.g. Greene et al. 2015; Martín-Navarro et al. 2018; Cheng et al. 2024). Thus, gradients appear to evolve over cosmic time.
- We examined the correlations between our gradients and galaxy properties including global velocity dispersion, integrated galaxy age, and V_{r_e}/σ_0 . In particular, we found that age gradients are stronger in galaxies that have more rotational support. This trend and the fact that gradients may evolve over time suggest that galaxy growth may continue post-quenching via minor

mergers (e.g. Bezanson et al. 2009; Naab et al. 2009; van de Sande et al. 2013). In this scenario, the accretion of old low-mass and low-metallicity satellites could act to flatten age and $[\text{Mg}/\text{Fe}]$ gradients and build up negative $[\text{Fe}/\text{H}]$ gradients while simultaneously increasing galaxy sizes (e.g. Trujillo et al. 2004; van Dokkum et al. 2010; van der Wel et al. 2014) and decreasing their rotational support (Slob et al. 2025).

- The observed ‘evolution’ in gradients may also be explained by different quenching mechanisms occurring at different times, with early-time quenchers having negative age and flat $[\text{Fe}/\text{H}]$ gradients and late-time quenchers having negative $[\text{Fe}/\text{H}]$ and flat age gradients.

In this work, we probed the peak of the star-formation quenching and galaxy assembly era with the first detailed stellar population gradients derived in a sample of distant massive quiescent galaxies using ultra-deep *JWST*/NIRSpec-MSA spectra. This study has allowed us to drive closer to the detailed formation histories of such galaxies, with our results suggesting that their cores formed earlier and quenched more abruptly than their outskirts. However, there are still a number of open questions. For example, it is not clear why we find indications of a positive $[\text{Mg}/\text{H}]$ gradient but no gradient in $[\text{Fe}/\text{H}]$. It is also unclear how the IMF behaves in these galaxies and how it may affect the derived stellar population gradients. In the future, larger sample sizes, increased spatial resolution, and a robust characterisation of the stellar IMF will paint an even clearer picture of the physical processes at play in the evolution of these galaxies.

Acknowledgments

We thank the anonymous referee for taking the time to give us useful feedback that improved this manuscript. We thank Antoine Dumont, Ciarán Rogers, and Colin Yip for useful conversations. We thank the *JWST*/NIRSpec helpdesk for their support. This work is based on observations made with the NASA/ESA/CSA *JWST*. The data were obtained from the Mikulski Archive for Space Telescopes (MAST) at the Space Telescope Science Institute (STScI), which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-03127 for *JWST*. These observations are associated with program *JWST*-GO-2110. Support for program *JWST*-GO-2110 was provided by NASA through a grant from the STScI. Some of the data products presented herein were retrieved from the Dawn *JWST* Archive (DJA). DJA is an initiative of the Cosmic Dawn Center (DAWN), which is funded by the Danish National Research Foundation under grant D NRF140. This work was performed using the compute resources from the Academic Leiden Interdisciplinary Cluster Environment (ALICE) provided by Leiden University. This work also used the Dutch national e-infrastructure with the support of the SURF Cooperative using grant no. EINF-10017 which is financed by the Dutch Research Council (NWO). MK acknowledges funding from the NWO through the award of the Vici grant VI.C.222.047.

Appendix

4.A Mock observations

To determine our S/N limits (see Section 4.2) and fitting strategy (see Section 4.3.4) we performed test fits with simulated *JWST*-SUSPENSE observations, as discussed in Section 4.2. These tests are very similar to the ones performed in Cheng et al. (2024). We generated mock spectra using the model grids in ALF α . We used the integrated fit results from Beverage et al. (2025) to set the values of the stellar population parameters (age, velocity dispersion, and elemental abundances) to the values of an example galaxy (127154) from our sample. We then scaled these mock spectra to have different S/N, using the noise spectrum from the real observation of 127154. We scaled the noise to the desired S/N between 4520–4820 Å (the common wavelength region covered by our entire sample, see Section 4.2) and added the noise that we randomly sampled from a Gaussian distribution to the mock flux.

The purpose of these tests was twofold: (i) we tested the S/N that we needed to recover ages, Fe-, and Mg-abundances, and (ii) we tested whether we needed to fix the remaining elemental abundances to their integrated values, or whether we could leave them free. We show the results of these tests in Figure 4.7. We generated 10 mocks in each of six S/N bins and fit each individual mock with ALF α . In each panel of Figure 4.7, we took the difference between the recovered and input parameters from each fit and show the medians and 1σ errors. In the top row, we show a test where we leave all stellar population parameters free in the fits. In the bottom row, we show a test where we fix all elemental abundances other than Fe and Mg to their integrated values from Beverage et al. (2025). This test demonstrates that we require a $S/N \gtrsim 20 \text{ \AA}^{-1}$ between 4520 – 4820 Å in order to reliably recover age, Fe, and Mg. Additionally, the recovery of these three parameters is much more accurate when we fix the elemental abundances to their integrated values. In the top panel, it is difficult to recover the input parameters when all parameters are left free, even with very high S/N data.

4.B Wiggle correction algorithm

As discussed in Section 4.3.1, spatially resolved *JWST* observations are affected by undersampling of the PSF. Algorithms to correct these wiggles in NIRSpec IFU data have been introduced in Perna et al. (2023) and in the WICKED Python package (Dumont et al. 2025). However, at the time of writing, this study is the first one that tries to correct the wiggles in spatially resolved NIRSpec-MSA spectra. Thus, we implemented an algorithm that is heavily based on the ones presented in Perna et al. (2023) and WICKED, but we combined and customised these routines for our specific case. We summarise the steps in our algorithm here.

We first fit the wiggles using the spectrum extracted from the row with the highest S/N. To obtain the shape of the wiggle, we masked strong emission lines. Then, taking the integrated spectrum as a reference spectrum, we divided the

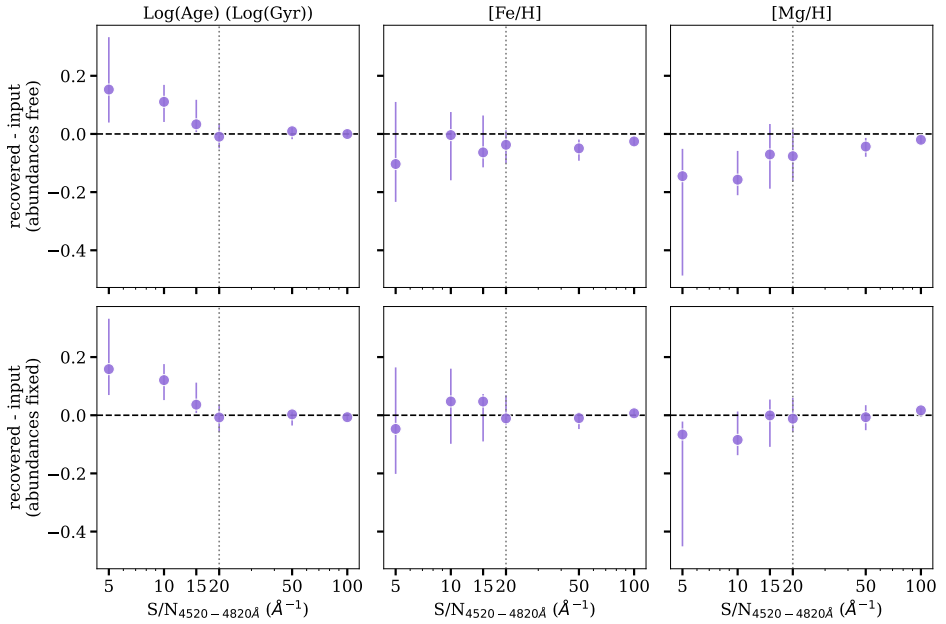


Figure 4.7: Tests with mock *JWST*-SUSPENSE observations. In each panel, we show the difference between the recovered and input parameters. We show true recovery (a difference of zero) with the black, dashed line and we show the S/N limit that we use in the body of the paper as the vertical, dotted, grey line. We show $\log(\text{Age})$ in the left column, $[\text{Fe}/\text{H}]$ in the middle column, and $[\text{Mg}/\text{H}]$ in the right column. In the top row, we left all stellar population parameters free in the ALFA fits. In the bottom row, we fixed all elemental abundances other than $[\text{Fe}/\text{H}]$ and $[\text{Mg}/\text{H}]$. All three parameters (age, Fe, and Mg) can be reliably recovered at a $\text{S/N} \gtrsim 20 \text{ \AA}$ between $4520 - 4820 \text{ \AA}$. Moreover, the recovery is much more accurate when the remaining elemental abundances are fixed. Thus, the fitting strategy shown in the bottom row is the one that we use throughout the paper.

brightest row by this reference, with both the reference and individual row spectra normalised by the value between $4480 - 4520 \text{ \AA}$ (see Slob et al. 2024). We broke the wobble spectrum up into small regions (non-uniform in length) that covered the entire wavelength range and used a sinusoidal function to model the wobble in each chunk. We iterated this process to find both the best chunk sizes and best-fitting sinusoidal models by minimising the χ^2 value over the entire wobble spectrum. We divided this wobble model from the spectral row to produce a corrected spectrum in the brightest row. For objects with a detector gap, we fit the wobble on each side of the gap independently (if the spectrum was long enough for a wobble to be detected visually).

For the subsequent spectral rows, we implemented the Fourier transform algorithm from WICKED to identify which rows contained a wobble. In particular, for rows near the edges, the spectrum is dominated by noise and the S/N is too low to detect a wobble. Additionally, there are certain brighter rows that are not as

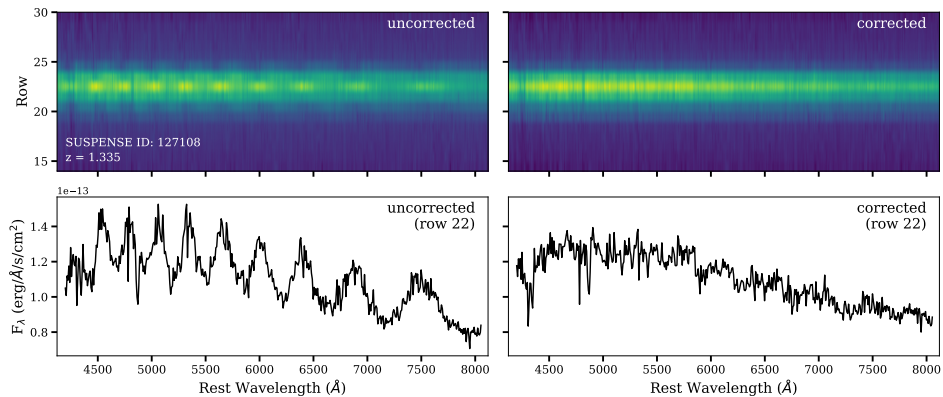


Figure 4.8: Example of the effect of our resampling noise correction for one object in our sample (object ID 127108). *Top:* Two-dimensional spectra. The uncorrected spectrum is shown on the left, where bright variations in the flux can be seen across the wavelength range. The corrected spectrum is shown on the right, where the flux is much more uniform across the wavelength range. *Bottom:* Extracted 1D rows, median-binned by 3 pixels. In the left panel, we extracted the brightest row from the uncorrected 2D spectrum. The bright variations in the flux seen in the top panel manifest as a sinusoidal flux variation in the 1D row. In the right panel, we extracted the brightest row from the corrected 2D spectrum, where this sinusoidal flux variation has been removed.

strongly affected by the undersampling and spectral rectification, and thus do not require wiggle correction. For the rows that contained wiggles, we used the chunks and wiggle model fit in the brightest row as a prior to constrain the wiggle in the remaining rows. We visually checked the fits in each row. In cases where these priors did not produce a good fit, we re-fit the rows without a prior. Similar to the brightest row, we divided the wiggle model for each row out of each respective spectral row to obtain a corrected spectrum. We ensured that flux was conserved during the wiggle correction procedure. We show an example of the wiggles and our correction in Fig. 4.8.

We tested the impact of the wiggles and the wiggle correction on our results by extracting and fitting the uncorrected spectra in the same way as the corrected spectra for the galaxies in our sample. In Figure 4.9, we show one-to-one plots comparing our fitted ages (left column), $[\text{Fe}/\text{H}]$ values (middle column), and $[\text{Mg}/\text{Fe}]$ values (right column) for our integrated (top row), core (middle row), and outskirt (bottom row) spectra. The integrated fits are largely consistent, showing that the wiggles do not have a significant impact on integrated spectra, although we note that there is still some scatter. The core and outskirt fits are also relatively consistent but have more scatter around the one-to-one relation compared to the integrated fits. We also note that Dumont et al. (2025) show that this wiggle correction algorithm does not affect the true equivalent widths of absorption features, and only acts to correct the broad, sinusoidal artefact across broad wavelength regions.

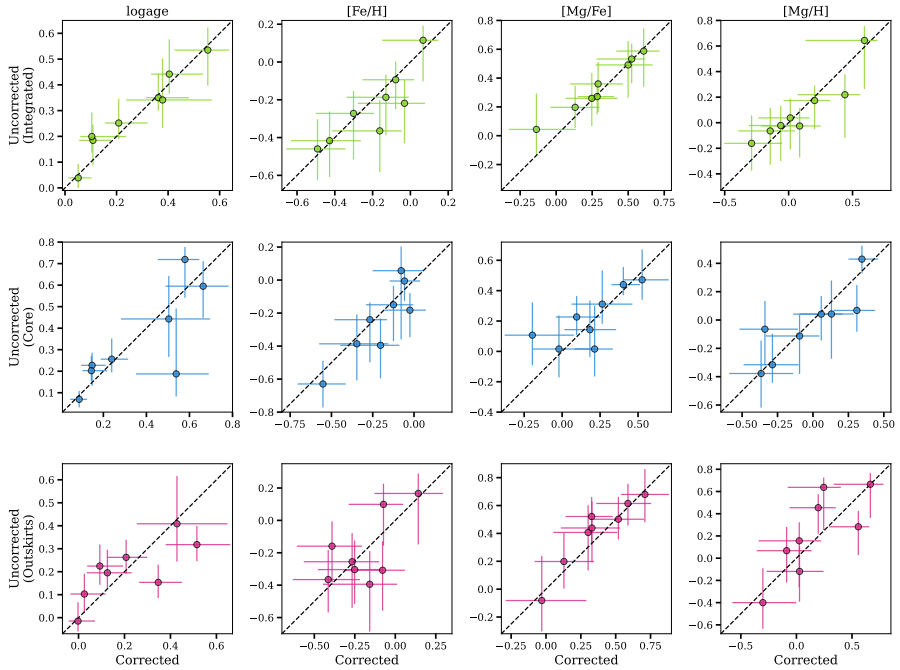


Figure 4.9: One-to-one plots comparing our fiducial (wiggle-corrected) fits to fits where we do not correct the wiggles. The integrated fits are shown in the top row, the core fits are shown in the middle row, and the outskirts fits are shown in the bottom row. We compare age in the left column, $[\text{Fe}/\text{H}]$ in the second, $[\text{Mg}/\text{Fe}]$ in the third column, and $[\text{Mg}/\text{H}]$ in the right column. Uncorrected fit results are shown on the y -axes and fiducial fit results are shown on the x -axes.

5 |

HIDDEN MASS IN EARLY GALAXIES
REVEALED BY BOTTOM-HEAVY
INITIAL MASS FUNCTIONS

ABSTRACT

JWST observations have revealed that massive galaxies formed and evolved faster than predicted by galaxy formation models, with many having already assembled a large mass in stars ~ 12 billion years ago. However, masses of distant galaxies are uncertain, as they assume a distribution of stellar birth masses (the initial mass function [IMF]) similar to that in the Milky Way (MW). Specifically, the contribution from low-mass stars, which make up the bulk of stellar mass, is not directly observed, but inferred based on an extrapolation of the MW IMF. Here, we provide the first robust constraints on the low-mass IMF beyond the local Universe from full-spectrum models. Using ultra-deep spectra of nine massive quiescent galaxies at $z \sim 0.7$ from *JWST*-IMFERNO, extended to bluer wavelengths with deep *VLT*/LEGA-C spectra, we find that the most massive galaxies have excess low-mass stars. Remarkably, our oldest galaxy ($z_{\text{form}} > 5$) has the most bottom-heavy IMF. This galaxy may be a descendant of *JWST*'s 'impossibly early' galaxies, implying that the latter may have had similarly bottom-heavy IMFs increasing their masses by a factor of $\sim 4 \pm 1$. Our findings may thus amplify the tension with galaxy formation models.

Chloe M. Cheng, Martje Slob, Mariska Kriek, Aliza G. Beverage, Pieter G. van Dokkum, Rachel Bezanson, Gabriel Brammer, Charlie Conroy, Anna de Graaff, Elham Eftekhari, Robert Feldmann, Wout M. Goesaert, Meng Gu, Joel Leja, Brian Lorenz, Pavel E. Mancera Piña, Ignacio Martín-Navarro, Andrew B. Newman, Sedona H. Price, Alice E. Shapley, Piyush Sharda, Katherine A. Suess, Arjen van der Wel, and Daniel R. Weisz

2026, *Nature Astronomy*, under review

5.1 Article

The IMF is an essential ingredient in galaxy formation models. It is required to infer fundamental properties from observed galaxy light, including stellar masses and star formation rates. Historically, the IMF has been considered to be ‘universal’, and assumed to take the same form as the MW IMF everywhere. However, in agreement with theoretical predictions (e.g. Schwarzschild & Spitzer 1953; Kroupa 2001; Krumholz et al. 2011; Hopkins 2012; Chabrier et al. 2014), observational evidence that the IMF varies in diverse stellar systems in the nearby Universe (e.g. Spiniello et al. 2012; Cappellari et al. 2013a; Geha et al. 2013; La Barbera et al. 2013; McDermid et al. 2014; Martín-Navarro et al. 2015b, 2021; Hallakoun & Maoz 2021; Cheng et al. 2023, see also Smith 2020) has been mounting. In particular, the cores of present-day massive early-type galaxies have been found to host bottom-heavy IMFs (e.g. Conroy & van Dokkum 2012b; van Dokkum et al. 2017; Gu et al. 2022). Despite these indications, studies of distant massive quiescent galaxies – the likely progenitors of nearby early-type cores – still largely adopt a single, MW IMF. Thus, the assumed shape of the IMF represents a major source of systematic uncertainty in our mass estimates. With recent *James Webb Space Telescope* (*JWST*) observations revealing significant tensions between the inferred stellar masses of early galaxies and galaxy formation models (e.g. Steinhardt et al. 2016; Carnall et al. 2023a,b, 2024; Antwi-Danso et al. 2025; Baker et al. 2025; de Graaff et al. 2025; Ito et al. 2025; Stevenson et al. 2026), it is imperative that we constrain the IMF at early times.

In this work, we constrain the low-mass end of the stellar IMF in nine massive quiescent galaxies at $z \sim 0.7$, observed via the *JWST*-IMFERN0 program. Targets were selected from the LEGA-C survey (van der Wel et al. 2021), on the basis of their quiescent stellar populations, redshifts, *H*-band magnitudes, and availability of COSMOS-Web photometry (Casey et al. 2023; Brammer 2023; Valentino et al. 2023, see Methods). The IMFERN0 spectra were obtained with the *JWST*-NIRSpec/MSA, using the G140M-F100LP disperser-filter combination. Combined with the bluer LEGA-C spectra, the average rest-frame wavelength coverage is $\sim 3700 - 10800 \text{ \AA}$. To robustly constrain the low-mass end of the IMF, we require a median S/N $\gtrsim 15 \text{ \AA}^{-1}$ for the LEGA-C spectra and a median S/N $\gtrsim 70 \text{ \AA}^{-1}$ for the IMFERN0 spectra, which we determine via mock spectra tests similar to Cheng et al. (2024). On the LEGA-C side, this rest-frame S/N allows us to constrain ages and elemental abundance patterns required to break strong degeneracies between abundance and IMF variations (Conroy & van Dokkum 2012a,b; Gu et al. 2022). On the IMFERN0 side, we require higher rest-frame S/N as we target much fainter spectral features, sensitive to the relative fraction of red dwarfs to red giants. We show the nine IMFERN0 spectra in Figure 5.1, and the corresponding LEGA-C spectra in Extended Data Figure 5.1. We show an example combined LEGA-C + IMFERN0 spectrum in Extended Data Figure 5.2. Our sample properties are reported in Extended Data Table 5.1.

To constrain the low-mass end of the IMF, we simultaneously fit the entire rest-frame wavelength range covered by LEGA-C and IMFERN0 with the ABSORPTION LINE FITTER (ALF), a flexible, full-spectrum stellar population synthesis (SPS)

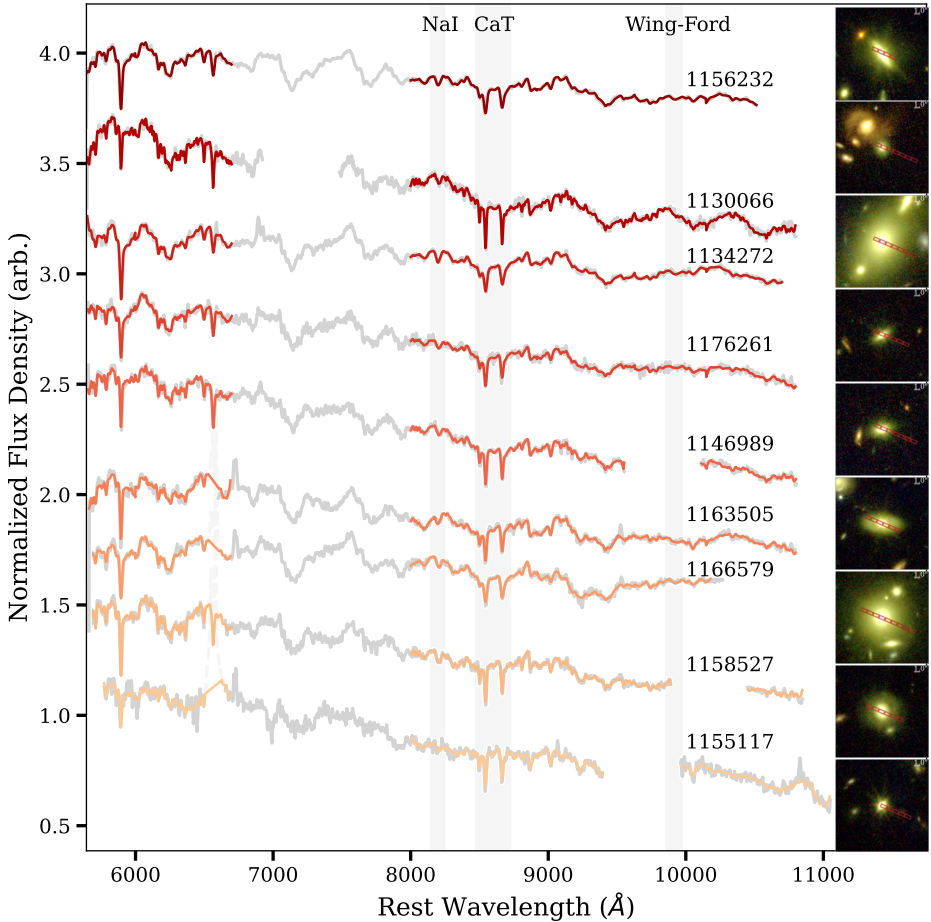


Figure 5.1: JWST-NIRSpec/MSA spectra with best-fit stellar population models and image cutouts. The NIRSpec/MSA spectra with best-fit stellar population models over plotted are shown for our sample of 9 massive quiescent galaxies at $z \sim 0.7$, observed via the IMFERNO program. Spectra are presented in order of increasing redshift (bottom to top). We normalize each spectrum by the median flux between rest-frame 7500 – 7600 Å, arbitrarily offsetting them in the y -direction for visibility. We indicate key IMF-sensitive absorption features with shaded regions. We do not fit the region between 7000 – 8000 Å due to broad TiO absorption (see Methods). We also show the RGB image cutout ($8''.5 \times 8''.5$) for each galaxy next to its corresponding spectrum, where we have combined the F115W, F277W, and F444W COSMOS-Web images (Casey et al. 2023; Brammer 2023; Valentino et al. 2023). We display the NIRSpec/MSA shutters on top of each image.

code (Conroy & van Dokkum 2012a; Conroy et al. 2018). ALF is unique amongst SPS codes in that it can account for variability in abundance patterns and the low-mass shape of the IMF. This flexibility, along with our broad wavelength range,

allows us to disentangle degeneracies between elemental abundance and low-mass IMF variations (Conroy & van Dokkum 2012a,b; Gu et al. 2022). We adopt a single-age stellar population, where age, 18 elemental abundances, and the two low-mass slopes of a double broken power-law IMF can vary freely (see Methods and Extended Data Figure 5.3). We show representative corner plots of key recovered parameters for two galaxies in our sample in Extended Data Figure 5.4. We find that we are able to disentangle well-known degeneracies between the low-mass IMF and elemental abundance variations (see Conroy & van Dokkum 2012a,b; Gu et al. 2022 and Methods). To enable comparison with previous work, we also fit the spectra with a Kroupa (2001) IMF, referred to as the Milky Way (MW) IMF going forward. We test adopting a two-component star-formation history (SFH) comprised of two bursts of star formation with free ages and a parameter representing the relative mass fraction of the young component. The results from this more complex SFH are consistent within uncertainties with those of the single-age model. Thus, we present only the single-age model.

We show the fits to the IMFERNO spectra in Figure 5.1, along with galaxy RGB images from COSMOS-Web (Casey et al. 2023; Brammer 2023; Valentino et al. 2023). We present the LEGA-C fits in Extended Data Figure 5.1, although note that we fit the LEGA-C and IMFERNO spectra for each galaxy simultaneously. We also show an example full spectrum and best-fitting model in Extended Data Figure 5.2.

Our ALF fit results are summarized in Figure 5.2 and reported in Extended Data Table 5.2. We show the IMF mismatch parameter (α_{IMF}), which is the ratio between the mass-to-light (M/L) ratio where we allow the low-mass end of the IMF to vary freely and the M/L ratio where we adopt a MW IMF. As the luminosity is the same for each IMF characterisation, α_{IMF} represents the stellar mass excess compared to the MW. We derive M/L ratios in the rest-frame *HST*/ACS-F814W band. The dashed line represents where galaxies with a MW IMF would lie, with anything above this line considered to be bottom-heavy. We colour-code each point by virial mass ($\log(M_{\text{vir}}/M_{\odot})$, see van der Wel et al. 2022 and Methods). We compare α_{IMF} 's to integrated stellar velocity dispersions (σ_v , see Methods) and [Fe/H] ratios (from our ALF fits). For comparison, we show analogous literature measurements from nearby massive elliptical galaxies (grey, Conroy & van Dokkum 2012b; van Dokkum et al. 2017; Gu et al. 2022; Cheng et al. 2023) and lensing/dynamical measurements from nearby massive ellipticals (blue, Treu et al. 2010; Cappellari et al. 2013a; Newman et al. 2017).

Figure 5.2 demonstrates that two of our galaxies have α_{IMF} 's that are > 2 standard deviations above 1, indicating that they have an excess of low-mass stars compared to the MW (see Extended Data Figure 5.3). In other words, their IMFs are bottom-heavy. When considering the full sample, we find that α_{IMF} ranges from 1.1 to 4.0, with a median of 1.4. We primarily find bottom-heavy IMFs in galaxies with higher σ_v , [Fe/H], and $\log(M_{\text{vir}}/M_{\odot})$. Qualitatively similar trends were found in the nearby Universe, using similar and complementary methods (Treu et al. 2010; Conroy & van Dokkum 2012b; Cappellari et al. 2013a; La Barbera et al. 2013; McDermid et al. 2014; van Dokkum et al. 2017, see also Smith 2020). Thus, we show that these trends were already in place at $z \sim 0.7$, 6.5 billion

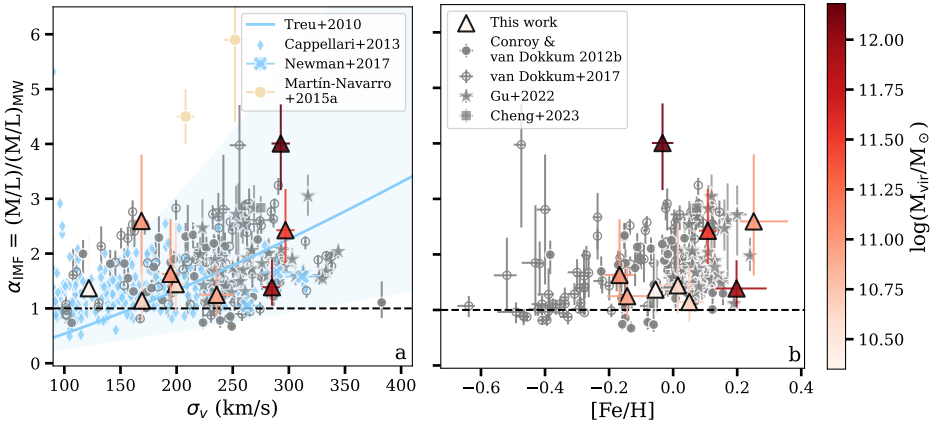


Figure 5.2: The IMF mismatch parameter compared to different galaxy properties. The IMF mismatch parameter (α_{IMF}) is defined as the ratio between the mass-to-light (M/L) ratio where we allow the low-mass IMF to vary freely and the M/L ratio where we assume a MW IMF. α_{IMF} is shown as a function of stellar velocity dispersion (σ_v ; panel a) and iron abundance ($[\text{Fe}/\text{H}]$; panel b). The IMFERNO galaxies are shown as triangles, colour-coded by their virial masses ($\log(M_{\text{vir}}/M_{\odot})$). Small open and closed circles, stars, and squares represent measurements for massive elliptical galaxies in the nearby Universe, from Conroy & van Dokkum (2012b) (re-derived in Cheng et al. 2023), van Dokkum et al. (2017), Gu et al. (2022), and Cheng et al. (2023). In panel a, the small diamonds, cross, and solid line represent lensing or dynamical measurements in the nearby Universe, from Treu et al. (2010), Cappellari et al. (2013a), and Newman et al. (2017). Large circles in panel a represent spectral index constraints from stacked spectra at $z \sim 1.1$, from Martín-Navarro et al. (2015a). Error bars and uncertainty bands represent statistical uncertainties at the level of 1 standard deviation.

years ago. Our results are consistent with weak lensing (Fujikawa & Oguri 2026) and spectroscopic indications (Martín-Navarro et al. 2015a) beyond the nearby Universe. In particular, Martín-Navarro et al. (2015a) constrained the IMF in two galaxy stacks at $z \sim 1.1$ using the IMF-sensitive TiO_2 index (beige circles in Figure 5.2). While our individual-galaxy constraints from full-spectrum fitting are more robust, it is encouraging that distant galaxy stacks exhibit qualitatively similar IMF behaviour.

In our analysis we have only studied the stellar components of our galaxies, and it is important to consider whether our inferred IMFs are realistic when accounting for non-baryonic matter. Thus, to assess the plausibility of our findings, we compare virial (see Methods) and stellar masses in Figure 5.3. We show stellar masses, estimated using MAGPHYS (da Cunha et al. 2008) and assuming a MW IMF (de Graaff et al. 2021), as grey circles. We correct the MAGPHYS masses for our best-fit IMF by multiplying them by α_{IMF} , and show these corrected masses as triangles colour-coded by α_{IMF} .

Figure 5.3 shows that, when assuming a MW IMF, the offset between virial and stellar masses increases towards higher mass. Similar results have been found

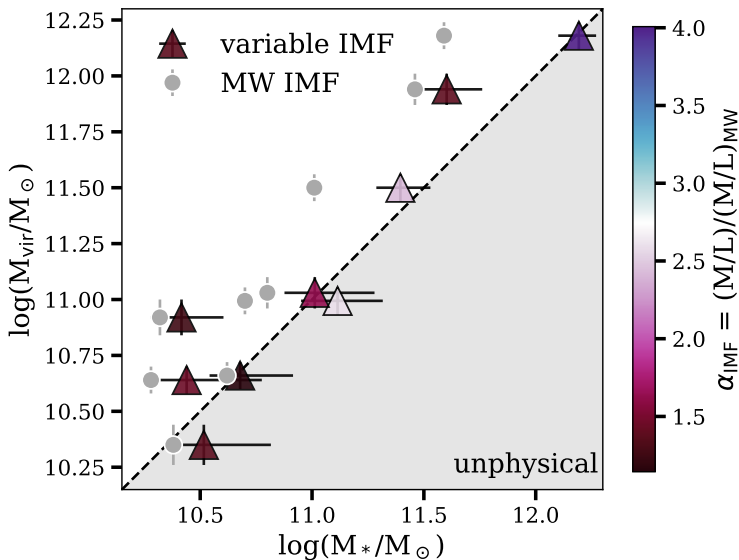


Figure 5.3: Comparison between stellar and virial masses for different low-mass IMF parameterizations. Virial masses ($\log(M_{\text{vir}}/M_{\odot})$) are shown on the y -axis (see Methods). Stellar masses, derived using MAGPHYS (da Cunha et al. 2008; de Graaff et al. 2021) and assuming a MW IMF, are shown on the x -axis (circles). We correct these masses for our best-fit low-mass IMF by multiplying them with α_{IMF} . We show the corrected masses as triangles, colour-coded by α_{IMF} . Error bars represent statistical uncertainties at the level of 1 standard deviation. The shaded region represents an unphysical region, where stellar masses are larger than virial masses. The dashed line represents a one-to-one comparison.

by van de Sande et al. (2013), Belli et al. (2017), Mendel et al. (2020), Kriek et al. (2024), and Slob et al. (2025). Interestingly, the stellar masses increase when applying our variable IMF, with the galaxies with the largest virial-to-stellar mass offsets shifting the most. This stellar mass increase generally brings the stellar and virial masses into agreement. In particular, the stellar masses do *not* exceed the virial masses within uncertainties, and are consistent with the physical plausibility of bottom-heavy IMFs in these galaxies (see also Lyubenova et al. 2016).

The excellent consistency between the virial and IMF-corrected stellar masses in Figure 5.3 leaves little room for dark matter and gas. However, this is not unexpected as these galaxies have small effective radii (median ~ 5.4 kpc) and thus the virial masses primarily capture their star-dominated central regions (see, e.g. de Graaff et al. 2024a). Additionally, little interstellar gas is expected to be present in massive quiescent galaxies at these redshifts (e.g., Spilker et al. 2018). Moreover, the virial masses may be underestimated as we do not fully correct σ_v for slit-loss effects (see Methods and Slob et al. 2025). In addition, these galaxies may be rotating (Bezanson et al. 2018b; van Houdt et al. 2021), in which case their σ_v 's may not be an ideal tracer of their potential wells, contributing to the

possible underestimation of the virial masses. This underestimation would allow for more non-baryonic matter. Detailed dynamical modeling is required to fully account for these effects (Slob et al. 2025, Slob et al. in prep.).

Our finding of bottom-heavy IMFs at early times is consistent with the two-phase formation model for massive galaxies, in which the cores of massive quiescent galaxies are in place at early times (with their central densities staying constant), with their outskirts subsequently building up via minor mergers (Bezanson et al. 2009; Naab et al. 2009; Feldmann et al. 2010; Oser et al. 2010; Rodriguez-Gomez et al. 2016; Suess et al. 2020). Moreover, a few of our galaxies may be even more bottom-heavy than nearby galaxy cores. Thus, in the two-phase scenario, the mixing of stellar populations by merging activity may reduce the bottom-heavy nature of the IMF between $z \sim 0.7$ and $z \sim 0$ in such galaxies (see also Chabrier et al. 2014).

We examine the galaxies for which we find the most bottom-heavy IMFs (i.e. inconsistent with a MW IMF by > 2 standard deviations based on statistical uncertainties) in detail. Galaxy 1134272 has the oldest age and earliest formation time of our sample, implying a formation redshift of $z_{\text{form}} \gtrsim 5$. In general, we would expect older galaxies to have slightly higher α_{IMF} 's, as an excess of low-mass stars fractionally contributes more to the mass for an older stellar population (in which more of the high-mass stars have died). However, this effect would be $< 10\%$ for the ages covered by our sample and thus cannot explain the high α_{IMF} . Assuming a typical star formation timescale of $\sim 100 - 200$ million years (Kriek et al. 2016; Beverage et al. 2025), this oldest galaxy was quiescent by $z \sim 4.6$, and thus may be a descendant of the ‘impossibly early’ galaxies discovered with *JWST* (Carnall et al. 2023a,b, 2024; Valentino et al. 2023; de Graaff et al. 2025; Ito et al. 2025). We note that we do not exclude the possibility that the remaining galaxies in our sample are also descendants of these early, massive galaxies, as their relatively younger ages could be driven by late-time accretion or star formation. In any case, the formation redshift of 1134272 as the oldest and most bottom-heavy galaxy implies that *JWST*'s distant, massive galaxies may have also had bottom-heavy IMFs at early times. If similarly bottom-heavy IMFs were already in place when these galaxies formed, this would make them a factor of $\sim 4 \pm 1$ *more massive* (considering statistical uncertainties alone) than originally reported. These larger masses would suggest an even greater tension with galaxy formation models (Steinhardt et al. 2016; Carnall et al. 2023a,b, 2024; Valentino et al. 2023; Antwi-Danso et al. 2025; de Graaff et al. 2025; Ito et al. 2025).

In this context, we note that we have applied a single α_{IMF} to each galaxy, implicitly assuming that the low-mass IMFs (and other stellar population parameters) measured within the NIRSspec-MSA slitlets are representative of each galaxy as a whole. However, present-day massive early-type galaxies – the likely descendants of our distant systems – are known to exhibit IMF gradients, with bottom-heavy cores (within $0.4 R_e$) and MW-like outskirts (e.g. Conroy & van Dokkum 2012b; van Dokkum et al. 2017; Gu et al. 2022, though see Alton et al. 2017). On average, our spectra probe the inner half-light radius of our galaxies (see Figure 5.2), meaning that the outer regions of our galaxies covered by our slitlets could plausibly host a different IMF. In a scenario where the low-mass IMF becomes fully

MW-like beyond $1 R_e$, our correction factors should actually be $\alpha_{\text{IMF}}/2$. However, given the substantial size growth observed over cosmic time (a factor of $\gtrsim 2$ in half-light radius between $z \sim 5$ and $z \sim 0.7$; e.g. Bezanson et al. 2009; Valentino et al. 2023; Baker et al. 2025; Stevenson et al. 2026) and assuming limited stellar mixing, the central regions of our galaxies were likely already in place at higher redshifts. Under this assumption, applying a uniform α_{IMF} to the full galaxies at $z \sim 4 - 5$ would be justified. Nevertheless, our low-mass IMF constraints here should be taken to be aperture-specific. Spatially resolved IMF measurements in the IMFERNO galaxies will be the subject of future work.

In this work, we have presented novel measurements of the low-mass end of the IMF in the early Universe, using state-of-the-art *JWST* data and sophisticated full-spectrum modeling methods. However, some limitations should be considered when interpreting our results. First, we note that our sample size is small, and that our galaxies are associated with an over density (see Methods). These factors may limit the generalisability of our findings. In the future, larger and more diverse sample sizes are required to confirm our findings.

Additionally, our findings are largely based on our ALF fits to the IMFERNO spectra. While ALF is currently the only code capable of constraining the shape of the low-mass IMF from continuum-normalized spectra while also accounting for variable abundance patterns, there may be some level of systematic uncertainty as a result of our modeling choices (i.e. stellar evolution assumptions, model and prior choices within the available parameter space, continuum-normalisation, line-spread function modeling, see e.g. Conroy 2013; Byrne & Stanway 2023; Bellstedt & Robotham 2025). In particular, different spectral libraries predict different absorption feature strengths in the rest-frame optical at the same age and metallicity (Byrne & Stanway 2023; Bellstedt & Robotham 2025; Jones et al. 2025). Encouragingly, Byrne & Stanway (2023) found the C3K stellar library (used by ALF) to be the most robust out of the libraries that they tested. Additionally, previous work has shown that Lick index and full-spectrum modeling methods generally produce the same results in local galaxies within uncertainties (e.g. Villaume et al. 2017a; Lonoce et al. 2021; Cheng et al. 2023). We have also performed several sensitivity tests (see Methods), and find that our constraints are generally robust to our precise fitting setup (i.e. IMF low-mass cutoff, masking of IMF sensitive features, and continuum normalisation). Furthermore, we note that our comparison between virial and stellar masses in Figure 5.3 supports the plausibility of our results.

We also fix the slope of the IMF above $1 M_\odot$, as our method is only sensitive to the relative fraction of low-to-high-mass stars (see Conroy & van Dokkum 2012a). However, if this slope is shallower than in the MW, the stellar masses at high redshifts can be reduced. We emphasize that our finding of bottom-heavy IMFs does not negate the possibility that the slope of the IMF beyond $1 M_\odot$ also differed from that of the MW at formation. In particular, the IMF across different stellar mass regimes may respond differently to the physical conditions of star formation, which could affect inferred stellar masses and their implications for galaxy formation models (e.g. Larson 1998; Fontanot et al. 2018; La Barbera et al. 2019; Martín-Navarro et al. 2021). For example, previous studies have proposed

IMFs that are simultaneously top- and bottom-heavy (e.g. Arrigoni et al. 2010; Coulter et al. 2017; den Brok et al. 2024; van Dokkum & Conroy 2024), time-varying (e.g. Vazdekis et al. 1997; Martín-Navarro 2016; Jeřábková et al. 2018), or metallicity-dependent (e.g. Kroupa 2001; Yan et al. 2024).

Though our spectra have an unprecedented depth at these redshifts and our IMF measurements are supported by the comparison with virial masses in Figure 5.3, our results are still uncertain. In an upcoming study, we will perform detailed anisotropic Jeans modeling of the IMFERNO galaxies to independently probe our conclusions (Slob et al. in prep.). Additionally, by extending our spectra to longer wavelengths and examining redder IMF-sensitive spectral features, the IMF can be more precisely constrained. Redder IMF constraints will also require further development of stellar population models, as they are not currently able to fully reproduce such spectra of nearby galaxies (Eftekhari et al. 2021, 2022). This spectral extension and development of the near-infrared models will be the subject of an upcoming approved JWST Cycle 5 program to obtain redder spectra of the current IMFERNO sample (PID: 10660).

In summary, for the first time we robustly measured the low-mass end of the IMF in a sample of massive quiescent galaxies beyond the nearby Universe, using full-spectrum modeling of extremely deep *JWST*/NIRSpec-MSA data. Our finding of bottom-heavy IMFs at $z \sim 0.7$ within the apertures probed by the spectra, particularly in our oldest galaxy, suggests that the most massive quiescent galaxies discovered at high redshifts (stellar masses $\gtrsim 10^{11} M_{\odot}$, Steinhardt et al. 2016; Carnall et al. 2023a,b, 2024; de Graaff et al. 2025; Ito et al. 2025) may have even higher masses than originally reported. Our results could additionally imply that the masses of their star-forming progenitors may also be underestimated, further heightening the tension with galaxy formation models. Looking forward, IMF measurements at higher redshifts, even closer to the epoch when these ‘impossibly early’ galaxies formed, will clarify this apparent tension with galaxy formation models. Future extremely deep *JWST*/NIRSpec observations are required to make these constraints possible.

Data Availability

This work makes use of data from the LEGA-C survey. The reduced data can be obtained from the ESO Science Archive Facility (http://archive.eso.org/eso/eso_archive_main.html). The reduced spectra and catalogue have been released by ESO (<http://archive.eso.org/cms/eso-archive-news/Third-and-final-release-of-the-Large-Early-Galaxy-Census-LEGA-C-Spectroscopic-Public-Survey-published.html>) and are also available here: <https://users.ugent.be/~avdrwel/research.html#legac>. For details, see van der Wel et al. (2021).

The *JWST* data analysed in this work were obtained from the Mikulski Archive for Space Telescopes (MAST) at the Space Telescope Science Institute (STScI). The specific observations can be accessed via 10.17909/k0zz-f371 (Cheng, Chloe 2025).

Our sample properties and fit results are published in Extended Data Tables 5.1 and 5.2. Files containing the spectra as they are fit with ALF (including the resolution curves and wavelength windows) and the `alf.f90` configuration files for each galaxy are available as Supplementary Data.

Code Availability

The ALF code on which this work is based is publicly available via <https://github.com/cconroy20/alf> (commit c21ccc5). See Conroy & van Dokkum (2012a); Conroy et al. (2018) for details. The procedure used to rescale the spectra to the photometry is available via https://github.com/Mslob/JWST_IMFERNO_phot_rescaling.

Acknowledgments

We thank the anonymous referees for taking the time to give us useful feedback that improved this manuscript. We thank the LEGA-C team for making their dataset public. This work is based on observations made with the NASA/ESA/CSA *JWST*. The data were obtained from the Mikulski Archive for Space Telescopes (MAST) at the Space Telescope Science Institute (STScI), which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-03127 for *JWST*. These observations are associated with program *JWST*-GO-5629. This work used the Dutch national e-infrastructure with the support of the SURF Cooperative using grants no. EINF-10017 and EINF-16504, which are financed by the Dutch Research Council (NWO). Some of the data products presented herein were retrieved from the Dawn *JWST* Archive (DJA). DJA is an initiative of the Cosmic Dawn Center (DAWN), which is funded by the Danish National Research Foundation under grant DNRF140.

Funding Statement

MK acknowledges funding from the NWO through the award of the Vici grant VI.C.222.047. Support for program *JWST*-GO-5629 was provided by NASA through a grant from the STScI. AGB acknowledges support from NASA through the NASA Hubble Fellowship grant #HST-HF2-51571 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. AdG acknowledges support from a Clay Fellowship awarded by the Smithsonian Astrophysical Observatory. PEMP acknowledges the support from the NWO through the Veni grant VI.Veni.222.364. PS is supported by the Leiden University Oort Fellowship and the IAU Gruber Foundation Fellowship.

Author Contributions

MK, AGB, and CMC wrote the primary *JWST* proposal. MK was the primary supervisor of this project and critically edited the text. MS, CMC, MK, and AGB designed the *JWST* observing plan. MS reduced the data and wrote the data reduction text. CMC fit the spectra with ALF, performed the velocity dispersion correction, wrote the remaining text, and created the figures. CMC, MK, and PGvD led the interpretation. All authors contributed to the analysis and interpretation.

Competing Interests

The authors declare no competing interests.

5.2 Methods

5.2.1 Observing Strategy and Data Reduction

In order to robustly constrain the low-mass IMF, we require extremely deep spectra covering rest-frame wavelengths between $\sim 4000 - 10000 \text{ \AA}$. To achieve this beyond the local Universe, we take advantage of the optical spectra from the third data release of the Large Early Galaxy Astrophysics Census (LEGA-C), a European Southern Observatory (ESO) Public Spectroscopic survey of 3600 galaxies at $0.6 \lesssim z \lesssim 1.0$. These deep (20 hr integration), $R \sim 3500$ spectra were collected over 128 nights using the VIMOS spectrograph on the ESO *Very Large Telescope* (VLT), resulting in an average S/N $\sim 20 \text{ \AA}^{-1}$ (see van der Wel et al. 2021 for details).

At $z \sim 0.7$, the LEGA-C observing strategy results in a rest-frame wavelength range of $\sim 3700 - 5100 \text{ \AA}$, capturing several Balmer features and metal lines that are sensitive to age, [Fe/H], and several elemental abundances. However, features that are sensitive to the ratio of dwarf to giant stars, which can be used to constrain the low-mass IMF (Conroy & van Dokkum 2012a), are located between $\sim 8100 - 10000 \text{ \AA}$. To capture these features, we extend the LEGA-C spectra with extremely deep Near Infrared Spectrograph micro shutter array (NIRSpec/MSA, Ferruit et al. 2022) spectra from the *JWST*-IMFERNO program (Initial Mass Function of Early Red NIRSpec Objects, PID 5629, PIs: Kriek, Beverage, and Cheng), executed on May 3-4 and 25, 2025.

To select our targets for *JWST*-IMFERNO, we identified quiescent targets from the LEGA-C galaxy sample with the *UVJ* classification from Muzzin et al. (2013a), using the UltraVISTA (McCracken et al. 2012) DR3 K_s -band catalogue (Muzzin et al. 2013b). In order to achieve our S/N requirements, we required an *H*-band magnitude < 21.5 . We additionally required the NIRSpec/MSA spectra to capture IMF sensitive features including NaI ($\sim 8180 - 8200 \text{ \AA}$), CaT ($\sim 8475 - 8725 \text{ \AA}$), and the Wing-Ford band ($\sim 9905 - 9945 \text{ \AA}$, Wing & Ford 1969) where possible. This was feasible for galaxies at $z \sim 0.7$ using the G140M-F100LP disperser-filter combination, allowing us to obtain spectra covering $\sim 9700 - 18400$

\AA , with a spectral resolution of $R \sim 1300$ (e.g. de Graaff et al. 2024b; Slob et al. 2024). This wavelength range corresponds to a rest-frame wavelength coverage of $\sim 5700 - 10800 \text{\AA}$ on average. We identified one extraordinary pointing for which we observed 13 massive ($10.28 \lesssim \log(M_*/M_\odot) \lesssim 11.59$, assuming a Kroupa 2001 IMF, van der Wel et al. 2021) quiescent galaxies at $z \sim 0.7$ which have LEGA-C optical spectra and COSMOS-Web imaging Casey et al. (2023). Nine of the 13 galaxies have sufficient S/N to simultaneously constrain ages, elemental abundances, and low-mass IMF variations. We note that all of our targets fit in one pointing as most are part of the ‘COSMOS Wall’ structure at $z \sim 0.73$ (Scoville et al. 2007a). Thus, our targets are also located in an old region of the $z \sim 0.7$ Universe, reinforcing the connection between our targets and massive galaxies in the early Universe.

The galaxies were observed using a custom two-point nod pattern with a two-micro-shutter ($1''.06$) offset in the cross-dispersion direction, following the *JWST*-SUSPENSE program (Slob et al. 2024). With this nodding pattern and large offset, we mitigate the self-subtraction of continuum emission for our extended targets in the data reduction process. To correct for detector and micro shutter artifacts we observed our galaxies in two nearly identical MSA configurations, offset by 3 micro shutters in the dispersion direction. Our nod-dither pattern resulted in four observing positions. While it was not possible to observe all filler targets in all four positions, seven of our nine primary targets presented in this work were covered by all positions resulting in a total on-source integration time of 31.2 hrs. Two primary targets (1155117 and 1158527) were only fully observed in three nod-dither positions due to slit failures during one visit. However, these two targets still have high enough S/N to be included in our final sample. To ensure sufficient empty sky for background calibration, we observed the primary targets with slitlets consisting of five to nine shutters.

We reduce the data to 2D flux calibrated spectra using the *JWST* Science Calibration Pipeline (Bushouse et al. 2023) v1.20.2 and version 1464 of the Calibration Reference Data System (CRDS). We add a $1/f$ correlated vertical readout noise correction to this pipeline, as described in Slob et al. (2024). After generating the 2D spectra via the pipeline, we use an optimal extraction algorithm with a Moffat profile (Moffat 1969) to extract the 1D spectra. Since the two dithers in our observations have slightly different MSA configurations, we reduce the dithers separately and combine the resulting 1D spectra. Finally, to ensure the LEGA-C and IMFERNO spectra are on the same flux scale, we re-scale the 1D spectra from each survey to the overlapping COSMOS2025 (Shuntov et al. 2025) photometry.

5.2.2 Full-spectrum fitting

To measure ages, elemental abundances, and low-mass IMF slopes, we fit the combined LEGA-C and IMFERNO spectra of 9 massive, quiescent galaxies at $z \sim 0.7$ with the ABSORPTION LINE FITTER (ALF), a full spectrum stellar population synthesis (SPS) model (Conroy & van Dokkum 2012a; Conroy et al. 2018). The ALF models are built on empirical simple stellar populations (SSPs), created using the MIST isochrones (Choi et al. 2016) and the Spectral Polynomial Interpolator (SPI,

Villaume et al. 2017b). We use the MILES empirical spectral library (Sánchez-Blázquez et al. 2006), the E-IRTF stellar library (Villaume et al. 2017b), and a large sample of M-dwarf spectra (Mann et al. 2015) with SPI. We additionally use a theoretical stellar library (C3K, see Conroy & van Dokkum 2012a) to ensure the quality of interpolation at the boundaries of the empirical parameter space. The ALF models allow for variable abundance patterns by differentially including theoretical element response functions. We explore an empirical parameter space spanning $-2.0 \lesssim [\text{Fe}/\text{H}] \lesssim 0.5$ and $3.5 \lesssim \log(T_{\text{eff}}/\text{K}) \lesssim 3.9$, which is set by the combined E-IRTF and Mann et al. (2015) samples. With these components, ALF develops stellar spectra as a function of T_{eff} , surface gravity, and metallicity from a data-driven model.

We parameterize the IMF as a double broken power-law with break points at $m = 0.5 M_{\odot}$ and $1 M_{\odot}$, similar to the Kroupa (2001) IMF, and a fixed low-mass cutoff (m_c) at $0.08 M_{\odot}$. From $1.0 M_{\odot} < m < 100 M_{\odot}$, the IMF slope is assumed to have the Salpeter (1955) value of 2.35:

$$\frac{dN}{dm} = \begin{cases} k_1 m^{-\alpha_1}, & 0.08 M_{\odot} < m < 0.5 M_{\odot} \\ k_2 m^{-\alpha_2}, & 0.5 M_{\odot} < m < 1.0 M_{\odot} \\ k_3 m^{-2.35}, & 1.0 M_{\odot} \leq m < 100 M_{\odot} \end{cases} \quad (5.1)$$

For a MW IMF, we use the Kroupa (2001) values of $\alpha_1 = 1.3$ and $\alpha_2 = 2.3$. This IMF parameterization is shown in Extended Data Figure 5.3, for a Kroupa IMF, a Salpeter IMF, and the best-fit IMFs of the IMFERNO sample.

We note that fixing the upper mass limit to $100 M_{\odot}$ does not significantly affect the M/L , especially for the old stellar populations that we are examining here. In this case, stars near the upper mass limit have already died and evolved into remnants. Thus, changing this limit primarily affects how much mass was lost due to stellar evolution, rather than the present-day light. As a result, increasing the limit (i.e. to $150 M_{\odot}$) lowers the current stellar mass by $\lesssim 2\%$ while leaving the luminosity unchanged.

Similarly, we fix the low-mass cutoff (m_c) to the hydrogen-burning limit at $0.08 M_{\odot}$. We perform a preliminary test of the sensitivity of our results to this choice by constraining m_c to a small range around $0.1 M_{\odot}$ and find that our results are robust to the assumed m_c . However, we emphasise that we are unlikely to be able to constrain m_c well, as this requires much higher S/N ($\gtrsim 300 \text{ \AA}^{-1}$, see Conroy et al. 2017; Newman et al. 2017).

To derive stellar population parameters, ALF fits a high-order Chebyshev polynomial to the ratio of the data to the model, to continuum-normalize the target spectrum. We note that our constraints are robust to the choice of polynomial order. We test this by fixing the order of the polynomial to $n = 7$ across the entire spectrum of each of our galaxies and refitting them. We find that the results are consistent with our fiducial fits within uncertainties. See Conroy & van Dokkum (2012b) and Conroy et al. (2018) for additional ALF continuum-normalization tests.

We also note that NIRSpec/MSA spectra have been found to be affected by undersampling of the NIRSpec point-spread function, resulting in broad variations in flux as a function of wavelength, known as ‘wiggles’ (Cheng et al. 2025b). As we

continuum-normalise our spectra, these broad features should not have a significant effect on our fitting results. Moreover, these artifacts have only been found to significantly affect spatially resolved spectra. Nevertheless, we test correcting these wiggles in our IMFERNO spectra using the algorithm presented in Cheng et al. (2025b) and find that the wiggle correction does not significantly affect our results or change our conclusions.

Using a FORTRAN implementation of the Markov chain Monte Carlo algorithm (MCMC), `emcee` (Foreman-Mackey et al. 2013), ALF samples the posteriors of 46 different stellar parameters, allowing for arbitrary variation in stellar age, elemental abundances, and the two low-mass slopes of the IMF (see Equation 5.1). It also fits for systematic parameters to characterize observed errors, including a jitter term. Additionally, it calculates M/L ratios by computing the mass in stars and remnants (using the prescription in Renzini & Ciotti 1993), assuming an IMF, and dividing this by the integral of the spectrum over the filter of interest (*HST*/ACS-F814W in our case). M/L ratios are calculated assuming the best-fit free IMF and Kroupa IMF for each fit. For details, see Conroy & van Dokkum (2012a) and Conroy et al. (2018). Based on our assumptions, we note that all uncertainties stated throughout the text are statistical, and do not include systematics.

We smooth the ALF models to the instrumental resolution of the LEGA-C and NIRSpec/MSA observations before fitting. We determine the instrumental resolution for the bluer sides of the spectra from LEGA-C as in Cheng et al. (2024) and Cheng et al. (2025a), by modeling 40 skylines across the wavelength range of each LEGA-C spectrum as Gaussians, and converting the median Gaussian σ to a velocity dispersion in km/s as a function of wavelength. For the redder sides of the spectra, since the instrumental resolution of the NIRSpec/MSA depends strongly on the source morphology and position in the slit (de Graaff et al. 2024b; Slob et al. 2024), we derive the instrumental resolution for each individual galaxy using MSAFIT (de Graaff et al. 2024b) and COSMOS-Web F150W NIRCам imaging (Casey et al. 2023), using a double Sérsic profile. The derived instrumental resolution for each galaxy differs from pre-launch estimates (<https://jwst-docs.stsci.edu/jwst-near-infrared-spectrograph/nirspec-instrumentation/nirspec-dispersers-and-filters>) for the G140M disperser, by a wavelength independent factor. We thus use the pre-launch estimated resolution curve, multiplied by the derived offset (factors ranging from 1.2 to 1.4), as the final instrumental resolution for each galaxy. We note that we test switching the resolution curves for two of the galaxies in our sample, and find that our constraints are robust to the precise modeling of the resolution curve.

Following the methodology of Cheng et al. (2024) and Cheng et al. (2025a), we fit each spectrum using 1024 walkers, 20000 burn-in steps, and a 1000 step production run. We examine 500 MCMC chains per fit and visually confirm that they are converged. We fit a single age, and avoid the walkers getting trapped at an unrealistically high age by initializing the age of each galaxy with a random value drawn from a uniform distribution centred at 3 Gyr. Additionally, we set the upper limit of the age prior to be the age of the Universe at the redshift of each galaxy, plus 2 Gyr to allow for uncertainties. We note that ALF can fit spectra

between 3700 – 24000 Å and can be used for stellar populations that are older than 1 Gyr. In Extended Data Figure 5.4, we display representative corner plots for one bottom-heavy galaxy from our sample (1134272) and one Milky-Way-like galaxy from our sample (1146989). We show our constraints on age, $[Z/H]$, $[Fe/H]$, $[Na/H]$, $[Ca/H]$, and the two IMF slopes (α_1 and α_2) for each galaxy.

Our fiducial results consist of fits to the full spectra, fitting the LEGA-C and IMFERNO spectra simultaneously (although note that we additionally fit the LEGA-C and IMFERNO spectra separately for testing purposes, see the next Section). We mask strong emission lines, including [OIII], [SII], and the $H\alpha$ + [NII] complex, when they are present. Where possible, we fit the wavelength ranges 3700 – 4700 Å, 4700 – 5200 Å, 5600 – 6000 Å, 6000 – 6700 Å, 8000 – 8920 Å, 8920 – 9630 Å, 9630 – 10150 Å, and 10150 – 10800 Å, excluding the regions between 7000 – 8000 Å due to broad TiO absorption (this is in line with previous work, see, e.g. Conroy & van Dokkum 2012b; Cheng et al. 2023; Beverage et al. 2025; Cheng et al. 2025b). We also test fitting the spectra including the region between 7000 – 8000 Å and find that while the results are relatively consistent with our fiducial fit, including this TiO absorption introduces large errorbars for some of our α_{IMF} parameters.

In our fiducial fit, we include the NaD absorption feature near ~ 5900 Å, which is important for constraining the upper limit of the IMF-degenerate Na-abundance (Conroy & van Dokkum 2012a). However, since NaD can be affected by the interstellar medium (Conroy & van Dokkum 2012a), we also test masking this feature. Similar to our TiO test, we find that while the results are consistent, some of the errorbars on α_{IMF} are inflated. We show the effect of this test on our sample in Supplementary Information Figure 5.1. We note that, independent of the IMF, our constraints on $[Na/H]$ come from a combination of the NaD and NaI features, as well as the entire spectrum. This is due to the fact that Na has an indirect effect on the whole spectrum as a result of the free electron abundance in stellar atmospheres. On the other hand, $[Ca/H]$ has been found to be constrained well by the CaI feature near 4227 Å. See Conroy & van Dokkum (2012b) for more details.

Moreover, we note that for some of the IMFERNO spectra, the IMF-sensitive Wing-Ford band (Wing & Ford 1969) falls in the detector gap (see Figure 5.1). Nonetheless, it has been shown that while all three features (NaI, CaT, and the Wing-Ford band) are important for robustly constraining the low-mass IMF, the exact shape of the low-mass IMF and trends with bottom-heaviness are not sensitive to any one feature. In other words, the results remain consistent when excluding individual features, as IMF variations affect the entire spectrum (see Conroy & van Dokkum 2012b; Cheng et al. 2023). We perform similar tests (i.e. masking each of NaI, CaT, and the Wing-Ford band) and find that our constraints are similarly robust to excluding individual IMF-sensitive features. We show our final fits to the LEGA-C spectra in Extended Data Figure 5.1 and our final fits to the IMFERNO spectra in Figure 5.1, but again note that we fit the two sides of each spectrum simultaneously.

Finally, for each galaxy we calculate α_{IMF} (the ratio between the M/L where we allow the low-mass IMF to vary and the M/L where we fix a MW IMF) and use this

α_{IMF} value to correct the stellar mass. These masses were originally constrained by fitting the photometric spectral energy distributions (from the UltraVISTA K_s -band catalogue, Muzzin et al. 2013b) of the galaxies with the MAGPHYS code (da Cunha et al. 2008). They have been scaled to the total stellar mass using the total luminosity of the best-fit Sérsic profile (see de Graaff et al. 2021). We correct the MAGPHYS masses with our best-fit α_{IMF} values instead of extracting masses from our ALF fits as MAGPHYS considers star-formation histories and dust more rigorously than ALF (which only has two simple star-formation history options and does not account for dust).

5.2.3 Stellar velocity dispersions and virial masses

The bluer LEGA-C and redder IMFERNO data were obtained with two different instruments. Thus, we must consider the fact that we are not examining exactly the same spatial areas of each galaxy in the blue and red regions of the spectra, due to differences in apertures, point-spread functions, slit orientations, and slit offsets with respect to galaxy centres. In particular, our measurements may be affected by stellar population gradients (Cheng et al. 2024, 2025b; Jafariyazani et al. 2020) and galaxy dynamics (Bezanson et al. 2018a; Slob et al. 2025). Stellar population gradients in LEGA-C galaxies have been found to be relatively weak (Cheng et al. 2024). In contrast, the velocity dispersions may be affected by differences in slit orientation, as the LEGA-C slits are oriented in the north-south direction and the IMFERNO slitlets have a position angle of ~ 247 degrees. This is especially the case if the galaxies are rotating (Bezanson et al. 2018a, Slob et al. in prep.).

To assess – and, if needed, correct for – velocity dispersion differences, we fit the IMFERNO spectra in isolation (i.e. without combining them with the bluer LEGA-C spectra), and compare the recovered velocity dispersions ($\sigma_v, \text{IMFERNO}$) to those from the LEGA-C survey, measured with PPXF (Cappellari 2023, $\sigma_v, \text{LEGA-C}$) (van der Wel et al. 2021; Bezanson et al. 2018b). For galaxies whose σ_v 's are offset by more than 1 standard deviation considering statistical uncertainties (four galaxies), we broaden the side of the spectrum (i.e. the LEGA-C or IMFERNO side) with the lower σ_v . We convolve the respective side of the spectrum with a Gaussian kernel with a width equal to the quadrature difference between $\sigma_v, \text{LEGA-C}$ and $\sigma_v, \text{IMFERNO}$. We finetune the precise width of the Gaussian until the two σ_v 's are consistent. We re-fit the corrected full spectra (i.e. LEGA-C + IMFERNO, shown in the body of the paper) and the individual sides of the spectra to assess the effect of our correction. We show this effect in Supplementary Information Figure 5.2. We note that this correction does not strongly affect our conclusions.

For the three galaxies for which $\sigma_v, \text{IMFERNO}$ is greater than $\sigma_v, \text{LEGA-C}$ (likely due to different slit alignments along the kinematic major axes), we also correct their virial masses. In particular, we multiply the virial masses from van der Wel et al. (2022) by a factor of $\sigma_{v, \text{IMFERNO}}^2 / \sigma_{v, \text{LEGA-C}}^2$. These corrected virial masses are shown throughout the paper. We additionally use the maximum velocity dispersion between LEGA-C and IMFERNO as the fiducial σ_v throughout the paper, as this is more representative of the galaxy's velocity dispersion. This is

due to the fact that the lower σ_v may neglect part of the kinematics if the slit is oriented along the minor axis.

5.3 Extended Data

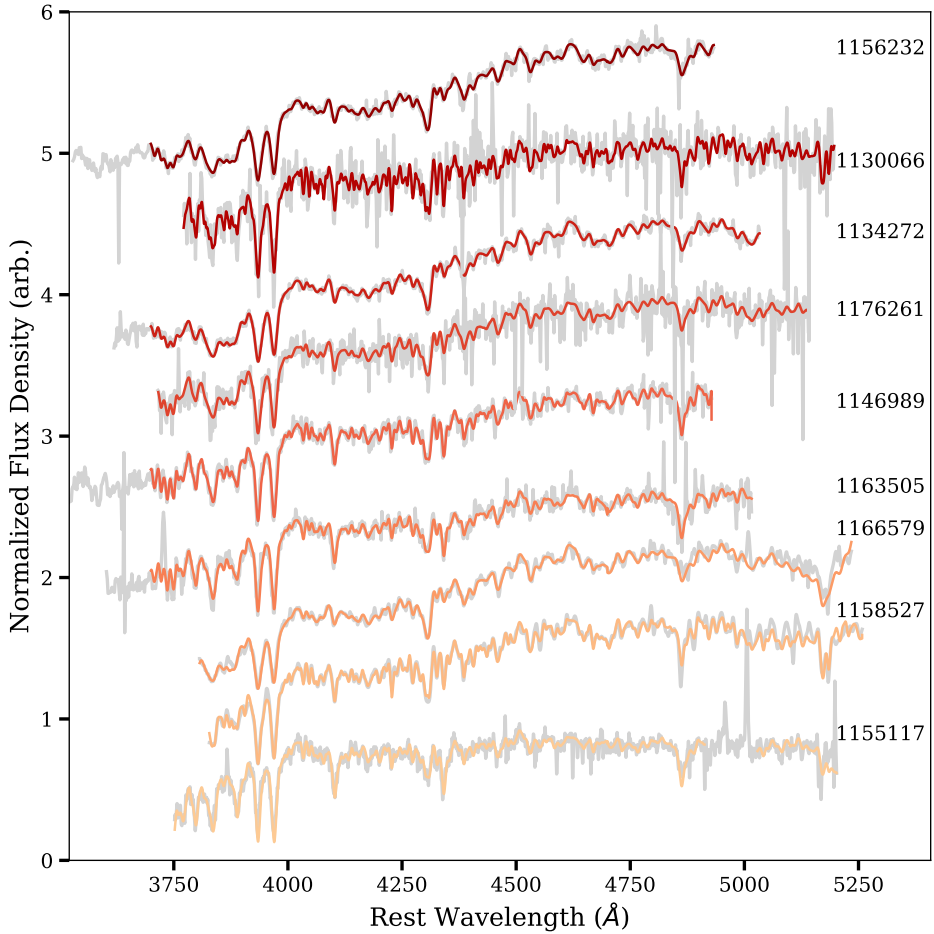
Extended Data Table 5.1: Properties of our sample of nine galaxies from *JWST*-IMFERNO. ¹Derived from running SOURCEEXTRACTOR (Bertin & Arnouts 1996) on the COSMOS-Web F150W images (Casey et al. 2023). ²From the LEGA-C survey (van der Wel et al. 2021). ³From the UltraVISTA K-band photometric catalogue (Muzzin et al. 2013b).

ID	LEGA-C ID	RA ¹ (hh:mm:ss)	Dec ¹ (dd:mm:s)	z_{spec} ²	H ³ (mag)
1155117	209642	09:59:40.85	+02:30:51.2	0.6984	20.5
1158527	209933	09:59:38.92	+02:30:59.7	0.7281	19.9
1166579	210564	09:59:39.77	+02:31:16.7	0.7286	18.6
1163505	239944	09:59:30.41	+02:31:07.4	0.7303	19.8
1146989	238668	09:59:33.55	+02:30:20.2	0.7324	20.2
1176261	211337	09:59:39.27	+02:32:04.2	0.7327	20.8
1134272	207735	09:59:46.85	+02:29:08.4	0.7340	18.4
1130066	207657	09:59:36.93	+02:29:20.3	0.7347	21.2
1156232	239469	09:59:32.99	+02:30:47.2	0.7357	19.2

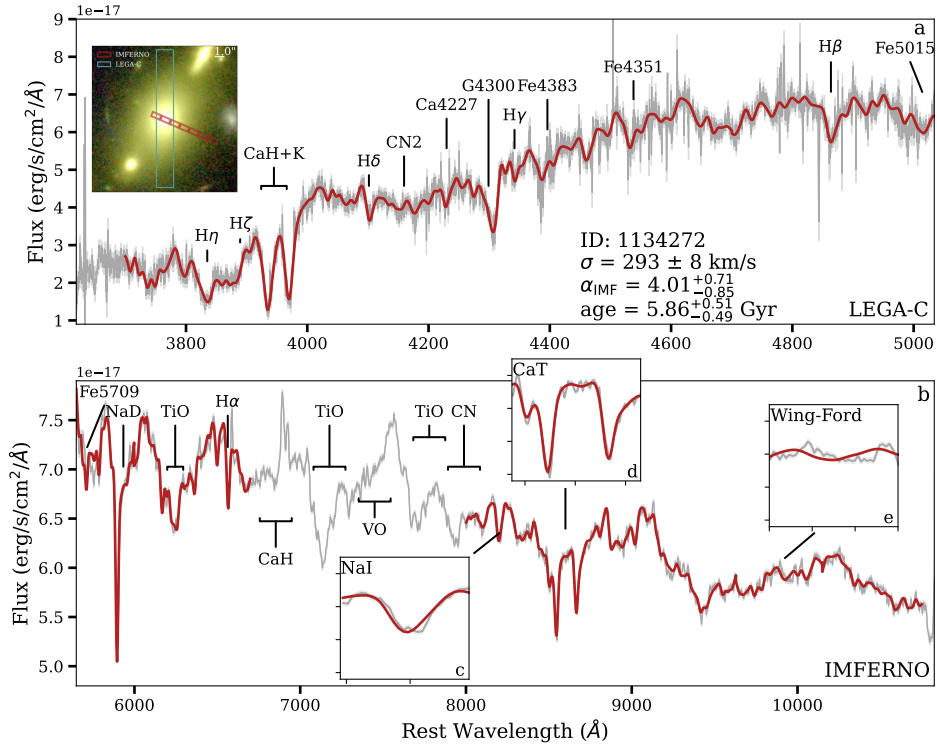
Extended Data Table 5.2: Results from our ALF fits to the combined LEGA-C and JWST-IMFERRNO spectra. σ_v represents the largest σ_v between the LEGA-C reported value and the value that we obtain from fitting the IMFERRNO spectra in isolation. M/L ratios are derived in the *HST*/ACS-F814W band.

ID	$\log(M_{\text{vir}})$ (M_{\odot})	$\log(M_*)$ (free, M_{\odot})	$\log(M_*)^1$ (MW, M_{\odot})	σ_v (km/s)	Age (Gyr)	[Fe/H]	M/L (free, M_{\odot}/L_{\odot})	M/L (MW, M_{\odot}/L_{\odot})	α_{IMF}
1155117	10.64 ± 0.06	$10.44^{+0.34}_{-0.25}$	10.28	$199.0^{+8.0}_{-6.0}$	$1.34^{+0.06}_{-0.05}$	$0.01^{+0.05}_{-0.06}$	$9.76^{+7.19}_{-12.34}$	$6.73^{+0.14}_{-0.13}$	$1.44^{+1.12}_{-0.21}$
1158527	10.99 ± 0.06	$11.11^{+0.16}_{-0.16}$	10.70	$169.0^{+7.0}_{-6.0}$	$2.40^{+0.07}_{-0.07}$	$0.25^{+0.06}_{-0.04}$	$28.00^{+10.67}_{-10.67}$	$10.65^{+0.17}_{-0.17}$	$2.60^{+0.98}_{-0.51}$
1166579	11.94 ± 0.07	$11.60^{+0.11}_{-0.11}$	11.46	$285.0^{+8.0}_{-8.0}$	$3.64^{+0.21}_{-0.20}$	$0.20^{+0.09}_{-0.07}$	$19.74^{+6.52}_{-4.38}$	$14.35^{+0.58}_{-0.52}$	$1.39^{+0.34}_{-0.34}$
1163505	11.03 ± 0.07	$11.01^{+0.27}_{-0.24}$	10.80	195.0 ± 8.0	$1.95^{+0.07}_{-0.07}$	$-0.17^{+0.04}_{-0.05}$	$14.79^{+9.09}_{-4.69}$	$8.85^{+0.15}_{-0.15}$	$1.63^{+1.06}_{-0.51}$
1146989	10.66 ± 0.06	$10.68^{+0.24}_{-0.14}$	10.62	$169.0^{+6.0}_{-6.0}$	$2.10^{+0.08}_{-0.09}$	$0.05^{+0.05}_{-0.04}$	$11.00^{+5.86}_{-3.92}$	$9.38^{+0.15}_{-0.15}$	$1.14^{+0.62}_{-0.36}$
1176261	10.92 ± 0.08	$10.42^{+0.19}_{-0.08}$	10.32	236.0 ± 16.0	$4.49^{+0.11}_{-0.51}$	$-0.14^{+0.06}_{-0.03}$	$17.63^{+7.78}_{-13.63}$	$13.99^{+1.39}_{-0.85}$	$1.25^{+0.34}_{-0.29}$
1134272	12.18 ± 0.06	$12.19^{+0.09}_{-0.09}$	11.59	293.0 ± 8.0	$5.86^{+0.25}_{-0.25}$	$-0.03^{+0.03}_{-0.03}$	$64.56^{+13.88}_{-13.88}$	$16.48^{+0.76}_{-0.76}$	$4.01^{+0.79}_{-0.85}$
1130066	10.35 ± 0.09	$10.52^{+0.30}_{-0.16}$	10.38	$122.0^{+7.0}_{-8.0}$	$2.95^{+0.14}_{-0.14}$	$-0.05^{+0.05}_{-0.05}$	$16.25^{+11.19}_{-6.01}$	$11.39^{+0.48}_{-0.37}$	$1.37^{+0.95}_{-0.51}$
1156232	11.50 ± 0.06	$11.39^{+0.13}_{-0.11}$	11.01	297.0 ± 8.0	$3.61^{+0.24}_{-0.22}$	$0.11^{+0.03}_{-0.03}$	$31.42^{+9.07}_{-7.83}$	$12.93^{+0.39}_{-0.37}$	$2.43^{+0.75}_{-0.60}$

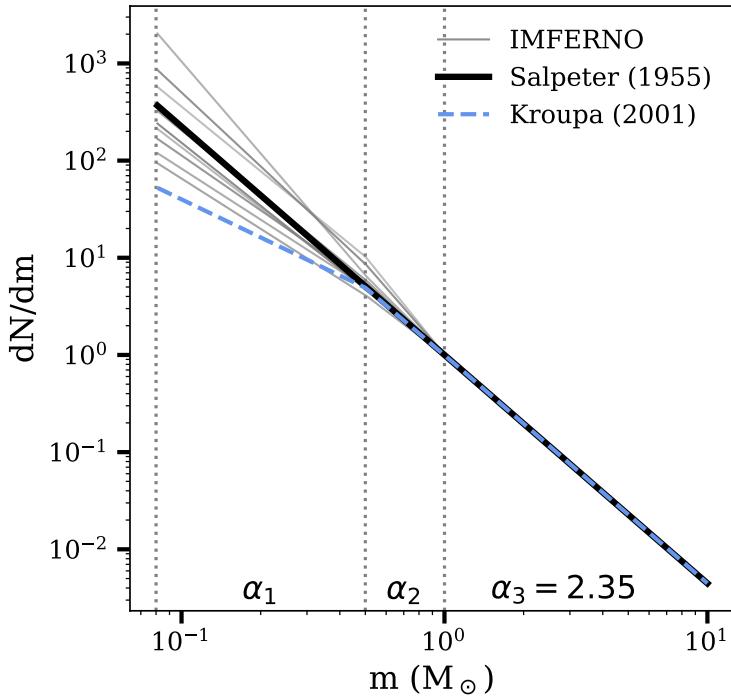
^aFrom the LEGA-C survey (van der Wel et al. 2021), measured with ppxf (Cappellari 2023).



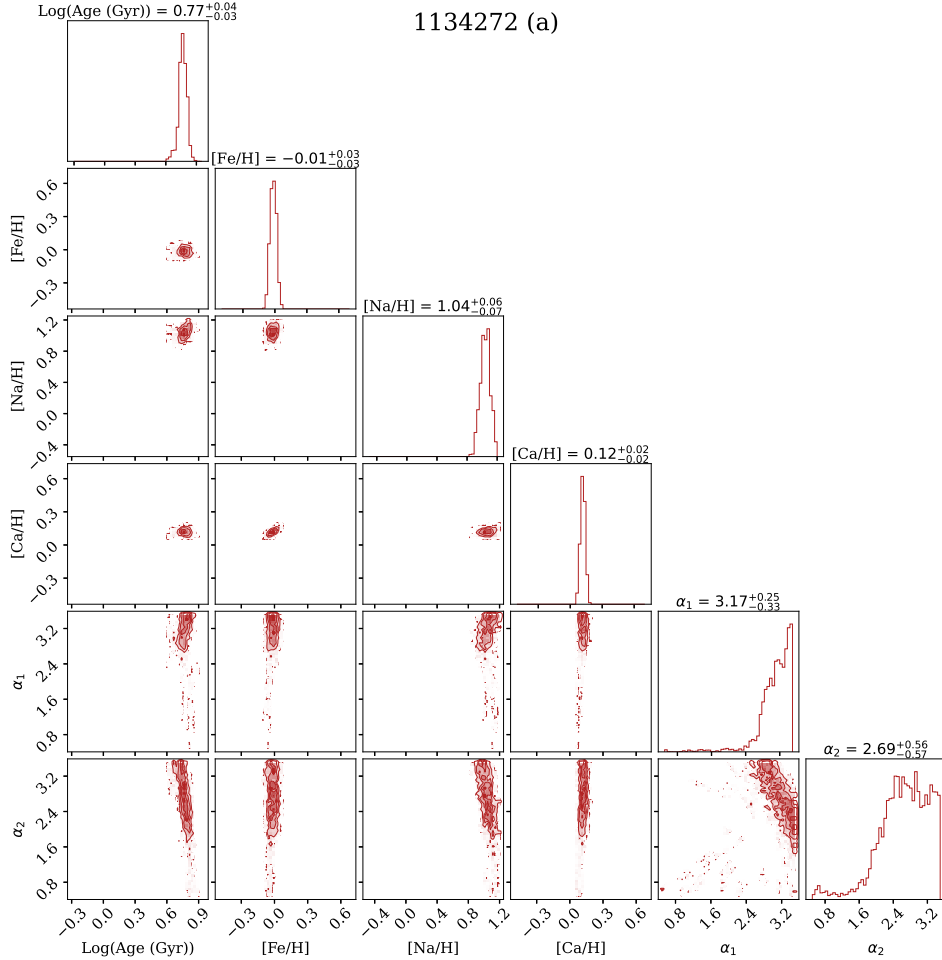
Extended Data Figure 5.1: Best fits to the bluer sides of the spectra from LEGA-C. Spectra are shown from bottom to top in order of increasing redshift. The best-fitting models are over plotted. We normalize each spectrum and fit by the median value of the flux between 4400 – 4500 Å. We arbitrarily offset the spectra in the y -direction. For the 6 spectra which we do not broaden, we median bin the data in bins of 5 pixels for visibility.



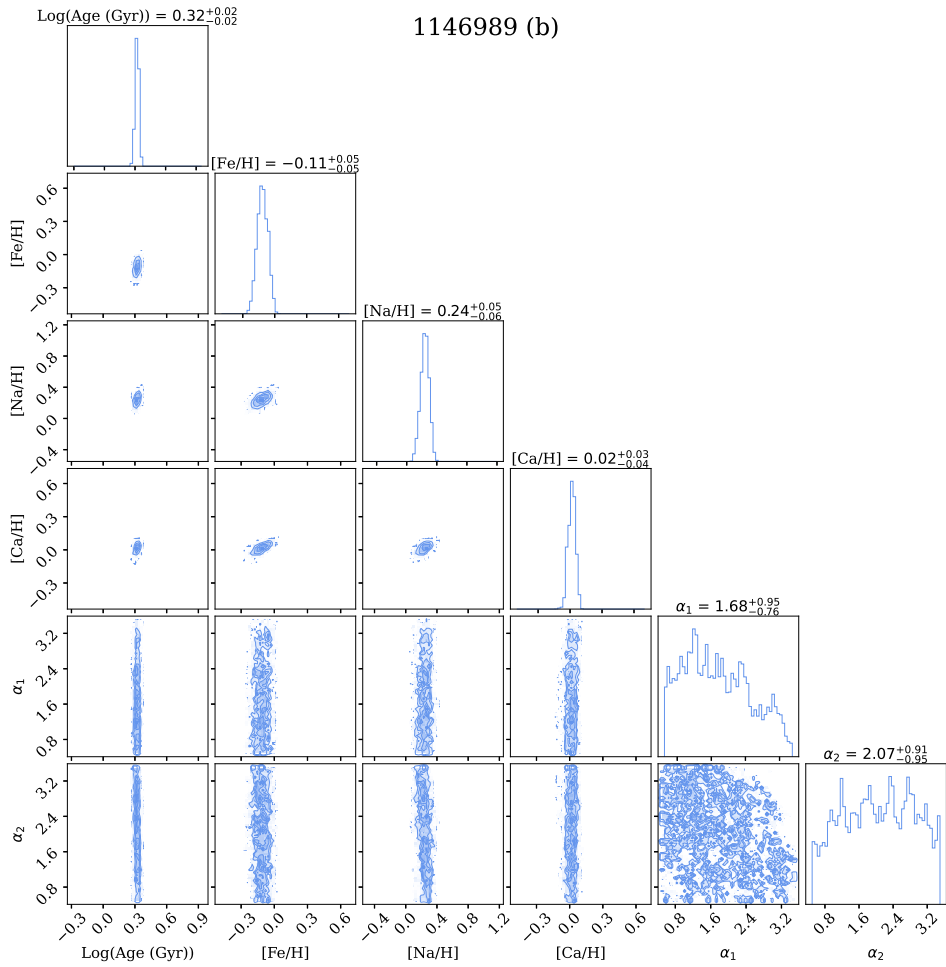
Extended Data Figure 5.2: Example ALF fit to a bottom-heavy galaxy in our sample. The data are shown with shaded regions indicating one standard deviation on the flux, and the best-fit model overlaid. We show the spectrum from LEGA-C in panel a and the IMFERNO spectrum in panel b, but we note that we fit the entire wavelength range simultaneously. In the inset panels (panels c-e), we zoom in on three strong absorption features which are sensitive to low-mass IMF variations. We continuum-normalize the features in the inset panels, but include the continuum in the main panels.



Extended Data Figure 5.3: IMFERNO IMF shapes compared to Kroupa and Salpeter. The IMF shapes for the IMFERNO sample are shown as thin, solid lines, and are compared to Kroupa (2001) (dashed line) and Salpeter (1955) (thick, solid line) IMFs. The double broken power law IMF is parameterized as in Equation 5.1 and shown on the y -axis for a range of stellar masses (x -axis). We indicate the breakpoints of the IMF with vertical dotted lines.

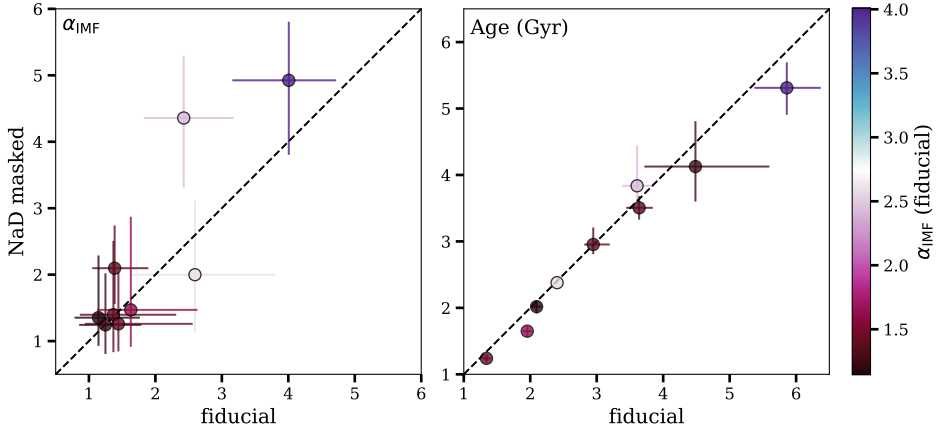


Extended Data Figure 5.4: Corner plots for representative galaxies in our sample. For each galaxy, we show contours for $\log(\text{Age (Gyr)})$, $[\text{Z}/\text{H}]$, $[\text{Fe}/\text{H}]$, $[\text{Na}/\text{H}]$, $[\text{Ca}/\text{H}]$, α_1 , and α_2 (the two low-mass IMF slopes). We show a bottom-heavy galaxy (1134272) in panel a and a Milky-Way-like galaxy (1146989) in panel b. Figure adapted for this thesis, continued on the next page.

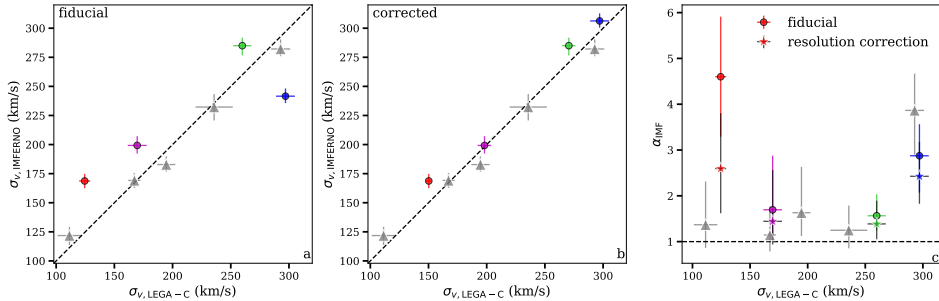


Extended Data Figure 5.5: Continued from Extended Data Figure 5.4. Figure adapted for this thesis.

5.4 Supplementary Information



Supplementary Information Figure 5.1: The effect of masking the NaD feature for our sample. We compare α_{IMF} (panel a) and age (panel b) from our fiducial fits and fits where we mask NaD. We colour code each point by the α_{IMF} that we obtain from our fiducial fits presented in the body of the paper. Error bars represent statistical uncertainties at the level of 1 standard deviation.



Supplementary Information Figure 5.2: Demonstration of our velocity dispersion correction to four galaxies in our sample. The triangles show the galaxies that do not need to be corrected, and all other points show the galaxies for which we perform the correction (i.e. broaden the side of the spectrum with lower σ_v). Error bars represent statistical uncertainties at the level of 1 standard deviation. In panel a, we show the fiducial σ_v values from LEGA-C (Bezanson et al. 2018b; van der Wel et al. 2021) and from fits to the individual IMFERNO spectra. In panel b, we show the corrected σ_v 's. In panel c, we show the effect that the correction has on our derived α_{IMF} .

BIBLIOGRAPHY

- Abdurro'uf, Coe, D., Jung, I., et al. 2023, *ApJ*, 945, 117
- Abdurro'uf, Lin, Y.-T., Wu, P.-F., & Akiyama, M. 2021, *ApJS*, 254, 15
- Abramson, L. E., Kelson, D. D., Dressler, A., et al. 2014, *ApJ*, 785, L36
- Acquaviva, V., Gawiser, E., & Guaita, L. 2011, *ApJ*, 737, 47
- Akhshik, M., Whitaker, K. E., Leja, J., et al. 2023, *ApJ*, 943, 179
- Akins, H. B., Narayanan, D., Whitaker, K. E., et al. 2022, *ApJ*, 929, 94
- Almaini, O., Wild, V., Maltby, D. T., et al. 2017, *MNRAS*, 472, 1401
- Alongi, M., Bertelli, G., Bressan, A., et al. 1993, *A&AS*, 97, 851
- Alton, P. D., Smith, R. J., & Lucey, J. R. 2017, *MNRAS*, 468, 1594
- Alvarez, R., Lançon, A., Plez, B., & Wood, P. R. 2000, *A&A*, 353, 322
- Andersen, M., Gennaro, M., Brandner, W., et al. 2017, *A&A*, 602, A22
- Anderson, J. 2016, *Empirical Models for the WFC3/IR PSF*, Instrument Science Report WFC3 2016-12, 42 pages
- Andrews, B. H., Weinberg, D. H., Schönrich, R., & Johnson, J. A. 2017, *ApJ*, 835, 224
- Anglés-Alcázar, D., Faucher-Giguère, C.-A., Kereš, D., et al. 2017, *MNRAS*, 470, 4698
- Antwi-Danso, J., Papovich, C., Esdaile, J., et al. 2025, *ApJ*, 978, 90
- Arrigoni, M., Trager, S. C., Somerville, R. S., & Gibson, B. K. 2010, *MNRAS*, 402, 173
- Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, *ARA&A*, 47, 481
- Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al. 2022, *ApJ*, 935, 167
- Avila-Reese, V., Ibarra-Medel, H., Lacerna, I., et al. 2023, *MNRAS*, 523, 4251

- Baade, W. 1944, *ApJ*, 100, 137
- Baker, W. M., Valentino, F., Lagos, C. d. P., et al. 2025, *A&A*, 702, A270
- Balogh, M. L., Morris, S. L., Yee, H. K. C., Carlberg, R. G., & Ellingson, E. 1999, *ApJ*, 527, 54
- Barone, T. M., D'Eugenio, F., Colless, M., et al. 2018, *ApJ*, 856, 64
- Barone, T. M., D'Eugenio, F., Scott, N., et al. 2022, *MNRAS*, 512, 3828
- Barro, G., Faber, S. M., Dekel, A., et al. 2016a, *ApJ*, 820, 120
- Barro, G., Faber, S. M., Koo, D. C., et al. 2017, *ApJ*, 840, 47
- Barro, G., Faber, S. M., Pérez-González, P. G., et al. 2013, *ApJ*, 765, 104
- Barro, G., Faber, S. M., Pérez-González, P. G., et al. 2014a, *ApJ*, 791, 52
- Barro, G., Kriek, M., Pérez-González, P. G., et al. 2016b, *ApJ*, 827, L32
- Barro, G., Trump, J. R., Koo, D. C., et al. 2014b, *ApJ*, 795, 145
- Bell, E. F. 2008, *ApJ*, 682, 355
- Bell, E. F. & de Jong, R. S. 2000, *MNRAS*, 312, 497
- Bell, E. F. & de Jong, R. S. 2001, *ApJ*, 550, 212
- Bell, E. F., McIntosh, D. H., Katz, N., & Weinberg, M. D. 2003, *ApJS*, 149, 289
- Bell, E. F., van der Wel, A., Papovich, C., et al. 2012, *ApJ*, 753, 167
- Belli, S., Bugiani, L., Park, M., et al. 2025, arXiv e-prints, arXiv:2510.11775
- Belli, S., Newman, A. B., & Ellis, R. S. 2015, *ApJ*, 799, 206
- Belli, S., Newman, A. B., & Ellis, R. S. 2017, *ApJ*, 834, 18
- Belli, S., Newman, A. B., & Ellis, R. S. 2019, *ApJ*, 874, 17
- Belli, S., Newman, A. B., Ellis, R. S., & Konidaris, N. P. 2014, *ApJ*, 788, L29
- Belli, S., Park, M., Davies, R. L., et al. 2024, *Nature*, 630, 54
- Bellstedt, S. & Robotham, A. S. G. 2025, *MNRAS*, 540, 2703
- Bernardi, M., Domínguez Sánchez, H., Sheth, R. K., Brownstein, J. R., & Lane, R. R. 2023, *MNRAS*, 518, 4713
- Bertin, E. & Arnouts, S. 1996, *A&AS*, 117, 393
- Bevacqua, D., Saracco, P., La Barbera, F., et al. 2025, *A&A*, 699, A203
- Beverage, A. 2024, alizabeverage/alfalpa: Initial Release alongside Beverage et al. 2024

- Beverage, A. G., Kriek, M., Conroy, C., et al. 2021, *ApJ*, 917, L1
- Beverage, A. G., Kriek, M., Conroy, C., et al. 2023, *ApJ*, 948, 140
- Beverage, A. G., Kriek, M., Suess, K. A., et al. 2024, *ApJ*, 966, 234
- Beverage, A. G., Slob, M., Kriek, M., et al. 2025, *ApJ*, 979, 249
- 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
- Bica, E. & Alloin, D. 1986, *A&AS*, 66, 171
- Bluck, A. F. L., Maiolino, R., Sánchez, S. F., et al. 2020, *MNRAS*, 492, 96
- Bluck, A. F. L., Mendel, J. T., Ellison, S. L., et al. 2014, *MNRAS*, 441, 599
- Bower, R. G., Benson, A. J., Malbon, R., et al. 2006, *MNRAS*, 370, 645
- Bradley, L., Sipőcz, B., Robitaille, T., et al. 2024, *astropy/photutils*: 1.12.0
- Brammer, G. 2023, *grizli*
- Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, *ApJ*, 686, 1503
- Breda, I., Papaderos, P., Gomes, J. M., et al. 2020, *A&A*, 635, A177
- Bressan, A., Fagotto, F., Bertelli, G., & Chiosi, C. 1993, *A&AS*, 100, 647
- Bruzual, G. & Charlot, S. 2003, *MNRAS*, 344, 1000
- Budavári, T., Wild, V., Szalay, A. S., Dobos, L., & Yip, C.-W. 2009, *MNRAS*, 394, 1496
- Burgarella, D., Buat, V., & Iglesias-Páramo, J. 2005, *MNRAS*, 360, 1413
- Burstein, D., Faber, S. M., Gaskell, C. M., & Krumm, N. 1984, *ApJ*, 287, 586
- Bushouse, H., Eisenhamer, J., Dencheva, N., et al. 2023, *JWST Calibration Pipeline*
- Byrne, C. M., Eldridge, J. J., & Stanway, E. R. 2025, *MNRAS*, 537, 2433
- Byrne, C. M. & Stanway, E. R. 2023, *MNRAS*, 521, 4995
- Caldwell, N. & Rose, J. A. 1998, *AJ*, 115, 1423
- Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, *ApJ*, 533, 682
- Cannon, A. J. & Pickering, E. C. 1924, *Annals of Harvard College Observatory*, 99, 1

- Cappellari, M. 2023, MNRAS, 526, 3273
- Cappellari, M., McDermid, R. M., Alatalo, K., et al. 2013a, MNRAS, 432, 1862
- Cappellari, M., McDermid, R. M., Alatalo, K., et al. 2012, Nature, 484, 485–488
- Cappellari, M., Scott, N., Alatalo, K., et al. 2013b, MNRAS, 432, 1709
- Carlberg, R. G. 1984, ApJ, 286, 403
- Carnall, A. C. 2017, arXiv e-prints, arXiv:1705.05165
- Carnall, A. C., Cullen, F., McLure, R. J., et al. 2024, MNRAS, 534, 325
- Carnall, A. C., Leja, J., Johnson, B. D., et al. 2019, ApJ, 873, 44
- Carnall, A. C., McLeod, D. J., McLure, R. J., et al. 2023a, MNRAS, 520, 3974
- Carnall, A. C., McLure, R. J., Dunlop, J. S., & Davé, R. 2018, MNRAS, 480, 4379
- Carnall, A. C., McLure, R. J., Dunlop, J. S., et al. 2022, ApJ, 929, 131
- Carnall, A. C., McLure, R. J., Dunlop, J. S., et al. 2023b, Nature, 619, 716
- Carollo, C. M., Bschorr, T. J., Renzini, A., et al. 2013, ApJ, 773, 112
- Carollo, C. M., Danziger, I. J., & Buson, L. 1993, MNRAS, 265, 553
- Carroll, B. W. & Ostlie, D. A. 2017, An introduction to modern astrophysics, Second Edition
- Casey, C. M., Kartaltepe, J. S., Drakos, N. E., et al. 2023, ApJ, 954, 31
- Cenarro, A. J., Cardiel, N., Gorgas, J., et al. 2001, MNRAS, 326, 959
- Chabrier, G. 2003, PASP, 115, 763
- Chabrier, G., Hennebelle, P., & Charlot, S. 2014, ApJ, 796, 75
- Chan, J. C. C., Beifiori, A., Mendel, J. T., et al. 2016, MNRAS, 458, 3181
- Chauke, P., van der Wel, A., Pacifici, C., et al. 2018, ApJ, 861, 13
- Cheng, C. M., Kriek, M., Beverage, A. G., et al. 2025a, MNRAS, 540, 1527
- Cheng, C. M., Kriek, M., Beverage, A. G., et al. 2024, MNRAS, 532, 3604
- Cheng, C. M., Slob, M., Kriek, M., et al. 2025b, arXiv e-prints, arXiv:2509.12316
- Cheng, C. M., Slob, M., Kriek, M., et al. 2026, arXiv e-prints, arXiv:2601.20864
- Cheng, C. M., Villaume, A., Balogh, M. L., et al. 2023, MNRAS, 526, 4004
- Cheng, Chloe. 2025, JWST-IMFERNO
- Chevallard, J. & Charlot, S. 2016, MNRAS, 462, 1415

- Chilingarian, I. V., Prugniel, P., Sil'Chenko, O. K., & Afanasiev, V. L. 2007, *MNRAS*, 376, 1033
- Choi, J., Conroy, C., & Johnson, B. D. 2019, *ApJ*, 872, 136
- Choi, J., Conroy, C., Moustakas, J., et al. 2014, *ApJ*, 792, 95
- Choi, J., Dotter, A., Conroy, C., et al. 2016, *ApJ*, 823, 102
- Cid Fernandes, R., Mateus, A., Sodré, L., Stasińska, G., & Gomes, J. M. 2005, *MNRAS*, 358, 363
- Ciocca, F., Saracco, P., Gargiulo, A., & De Propris, R. 2017, *MNRAS*, 466, 4492
- Clausen, M., Momcheva, I. G., Whitaker, K. E., et al. 2025, *ApJ*, 993, 106
- Cleveland, W. S. & Devlin, S. J. 1988, *Journal of the American Statistical Association*, 83, 596
- Coelho, P., Bruzual, G., Charlot, S., et al. 2007, *MNRAS*, 382, 498
- Coelho, P. R. T., Bruzual, G., & Charlot, S. 2020, *MNRAS*, 491, 2025
- Connolly, A. J., Szalay, A. S., Bershadsky, M. A., Kinney, A. L., & Calzetti, D. 1995, *AJ*, 110, 1071
- Conroy, C. 2013, *ARA&A*, 51, 393
- Conroy, C., Graves, G. J., & van Dokkum, P. G. 2014, *ApJ*, 780, 33
- Conroy, C. & Gunn, J. E. 2010, *ApJ*, 712, 833
- Conroy, C., Gunn, J. E., & White, M. 2009, *ApJ*, 699, 486
- Conroy, C. & van Dokkum, P. 2012a, *ApJ*, 747, 69
- Conroy, C. & van Dokkum, P. G. 2012b, *ApJ*, 760, 71
- Conroy, C., van Dokkum, P. G., & Villaume, A. 2017, *ApJ*, 837, 166
- Conroy, C., Villaume, A., van Dokkum, P. G., & Lind, K. 2018, *ApJ*, 854, 139
- Conselice, C. J., Mundy, C. J., Ferreira, L., & Duncan, K. 2022, *ApJ*, 940, 168
- Conselice, C. J., Rajgor, S., & Myers, R. 2008, *MNRAS*, 386, 909
- Cook, B. A., Conroy, C., Pillepich, A., Rodriguez-Gomez, V., & Hernquist, L. 2016, *ApJ*, 833, 158
- Cooper, A. P., D'Souza, R., Kauffmann, G., et al. 2013, *MNRAS*, 434, 3348
- Cortese, L., Catinella, B., & Smith, R. 2021, *PASA*, 38, e035
- Coulter, D. A., Lehmer, B. D., Eufrazio, R. T., et al. 2017, *ApJ*, 835, 183

- Crain, R. A. & van de Voort, F. 2023, *ARA&A*, 61, 473
- Croton, D. J., Springel, V., White, S. D. M., et al. 2006, *MNRAS*, 365, 11
- Cushing, M. C., Rayner, J. T., & Vacca, W. D. 2005, *ApJ*, 623, 1115
- Cutler, S. E., Whitaker, K. E., Mowla, L. A., et al. 2022, *ApJ*, 925, 34
- da Cunha, E., Charlot, S., & Elbaz, D. 2008, *MNRAS*, 388, 1595
- Da Rio, N., Robberto, M., Hillenbrand, L. A., Henning, T., & Stassun, K. G. 2012, *ApJ*, 748, 14
- Daddi, E., Renzini, A., Pirzkal, N., et al. 2005, *ApJ*, 626, 680
- Daddi, E., Röttgering, H. J. A., Labbé, I., et al. 2003, *ApJ*, 588, 50
- Damian, B., Jose, J., Samal, M. R., et al. 2021, *MNRAS*, 504, 2557
- Damjanov, I., Abraham, R. G., Glazebrook, K., et al. 2011, *ApJ*, 739, L44
- Damjanov, I., McCarthy, P. J., Abraham, R. G., et al. 2009, *ApJ*, 695, 101
- Damjanov, I., Sohn, J., Geller, M. J., Utsumi, Y., & Dell’Antonio, I. 2023, *ApJ*, 943, 149
- Damjanov, I., Zahid, H. J., Geller, M. J., et al. 2019, *ApJ*, 872, 91
- Davies, R. L., Sadler, E. M., & Peletier, R. F. 1993, *MNRAS*, 262, 650
- de Graaff, A., Bezanson, R., Franx, M., et al. 2021, *ApJ*, 913, 103
- de Graaff, A., Pillepich, A., & Rix, H.-W. 2024a, *ApJ*, 967, L40
- de Graaff, A., Rix, H.-W., Carniani, S., et al. 2024b, *A&A*, 684, A87
- de Graaff, A., Setton, D. J., Brammer, G., et al. 2025, *Nature Astronomy*, 9, 280
- De Lucia, G., Fontanot, F., & Hirschmann, M. 2017, *MNRAS*, 466, L88
- Dekel, A. & Birnboim, Y. 2006, *MNRAS*, 368, 2
- Dekel, A. & Burkert, A. 2014, *MNRAS*, 438, 1870
- den Brok, M., Krajnović, D., Emsellem, E., et al. 2024, *MNRAS*, 530, 3278
- D’Eugenio, F., Pérez-González, P. G., Maiolino, R., et al. 2024, *Nature Astronomy*, 8, 1443
- D’Eugenio, F., van der Wel, A., Wu (), P.-F., et al. 2020, *MNRAS*, 497, 389
- Di Matteo, T., Springel, V., & Hernquist, L. 2005, *Nature*, 433, 604
- Díaz-García, L. A., Cenarro, A. J., López-Sanjuan, C., et al. 2019, *A&A*, 631, A156

- Ditrani, F. R., Andreon, S., Longhetti, M., & Newman, A. 2022, *A&A*, 660, A132
- Dotter, A. 2016, *ApJS*, 222, 8
- Dressel, L., Hodge, P., & Barrett, P. 2007, wx2d: A PyRAF Routine to Resample Spectral Images, Instrument Science Report STIS 2007-04, 20 pages
- Dumont, A., Neumayer, N., Seth, A. C., et al. 2025, *A&A*, 703, A54
- Eftekhari, E., La Barbera, F., Vazdekis, A., Allende Prieto, C., & Knowles, A. T. 2022, *MNRAS*, 512, 378
- Eftekhari, E., Vazdekis, A., & La Barbera, F. 2021, *MNRAS*, 504, 2190
- Ellison, S. L., Sánchez, S. F., Ibarra-Medel, H., et al. 2018, *MNRAS*, 474, 2039
- Eminian, C., Kauffmann, G., Charlot, S., et al. 2008, *MNRAS*, 384, 930
- Fagotto, F., Bressan, A., Bertelli, G., & Chiosi, C. 1994a, *A&AS*, 104, 365
- Fagotto, F., Bressan, A., Bertelli, G., & Chiosi, C. 1994b, *A&AS*, 105, 29
- Falcón-Barroso, J., Sánchez-Blázquez, P., Vazdekis, A., et al. 2011, *A&A*, 532, A95
- Fan, L., Lapi, A., De Zotti, G., & Danese, L. 2008, *ApJ*, 689, L101
- Fanelli, M. N., O’Connell, R. W., Burstein, D., & Wu, C.-C. 1992, *ApJS*, 82, 197
- Fang, J. J., Faber, S. M., Koo, D. C., & Dekel, A. 2013, *ApJ*, 776, 63
- Feldmann, R., Carollo, C. M., Mayer, L., et al. 2010, *ApJ*, 709, 218
- Ferreras, I., Charlot, S., & Silk, J. 1999, *ApJ*, 521, 81
- Ferreras, I., La Barbera, F., de La Rosa, I. G., et al. 2013, *MNRAS*, 429, L15
- Ferreras, I., Scott, N., La Barbera, F., et al. 2019, *MNRAS*, 489, 608
- Ferruit, P., Jakobsen, P., Giardino, G., et al. 2022, *A&A*, 661, A81
- Fontanot, F., De Lucia, G., Hirschmann, M., et al. 2017, *MNRAS*, 464, 3812
- Fontanot, F., La Barbera, F., De Lucia, G., Pasquali, A., & Vazdekis, A. 2018, *MNRAS*, 479, 5678
- Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, 125, 306
- Franx, M. & Illingworth, G. 1990, *ApJ*, 359, L41
- Franx, M., Labbé, I., Rudnick, G., et al. 2003, *ApJ*, 587, L79
- Franx, M., van Dokkum, P. G., Förster Schreiber, N. M., et al. 2008, *ApJ*, 688, 770

- French, K. D. 2021, *PASP*, 133, 072001
- Fujikawa, M. & Oguri, M. 2026, *Open J. Astrophys*, 9, 61580
- Fumagalli, M., Franx, M., van Dokkum, P., et al. 2016, *ApJ*, 822, 1
- Gallazzi, A. & Bell, E. F. 2009, *ApJS*, 185, 253
- Gallazzi, A., Charlot, S., Brinchmann, J., White, S. D. M., & Tremonti, C. A. 2005, *MNRAS*, 362, 41
- Gallazzi, A. R., Pasquali, A., Zibetti, S., & Barbera, F. L. 2021, *MNRAS*, 502, 4457
- Gargiulo, A., Saracco, P., Longhetti, M., La Barbera, F., & Tamburri, S. 2012, *MNRAS*, 425, 2698
- Geha, M., Brown, T. M., Tumlinson, J., et al. 2013, *ApJ*, 771, 29
- Genzel, R., Förster Schreiber, N. M., Lang, P., et al. 2014, *ApJ*, 785, 75
- Girardi, L., Bressan, A., Chiosi, C., Bertelli, G., & Nasi, E. 1996, *A&AS*, 117, 113
- Glazebrook, K., Nanayakkara, T., Schreiber, C., et al. 2024, *Nature*, 628, 277
- Glazebrook, K., Schreiber, C., Labbé, I., et al. 2017, *Nature*, 544, 71
- Gobat, R., Daddi, E., Magdis, G., et al. 2018, *Nature Astronomy*, 2, 239
- Goddard, D., Thomas, D., Maraston, C., et al. 2017, *MNRAS*, 465, 688
- Gonzalez, J. J., Faber, S. M., & Worthey, G. 1993, in *American Astronomical Society Meeting Abstracts*, Vol. 183, *American Astronomical Society Meeting Abstracts*, 42.06
- González Delgado, R. M., Cid Fernandes, R., Pérez, E., et al. 2016, *A&A*, 590, A44
- González Delgado, R. M., García-Benito, R., Pérez, E., et al. 2015, *A&A*, 581, A103
- Gonzalez-Perez, V., Castander, F. J., & Kauffmann, G. 2011, *MNRAS*, 411, 1151
- Gottumukkala, R., Barrufet, L., Oesch, P. A., et al. 2024, *MNRAS*, 530, 966
- Graves, G. J. & Schiavon, R. P. 2008, *ApJS*, 177, 446
- Gray, R. O. & Corbally, J., C. 2009, *Stellar Spectral Classification*
- Greene, J. E., Janish, R., Ma, C.-P., et al. 2015, *ApJ*, 807, 11
- Greene, J. E., Murphy, J. D., Graves, G. J., et al. 2013, *ApJ*, 776, 64
- Greene, J. E., Veale, M., Ma, C.-P., et al. 2019, *The Astrophysical Journal*, 874, 66

- Griffith, R. L., Cooper, M. C., Newman, J. A., et al. 2012, *ApJS*, 200, 9
- Gu, M., Conroy, C., Law, D., et al. 2018, *ApJ*, 859, 37
- Gu, M., Greene, J. E., Newman, A. B., et al. 2022, *ApJ*, 932, 103
- Guo, Y., Giavalisco, M., Cassata, P., et al. 2011, *ApJ*, 735, 18
- Hallakoun, N. & Maoz, D. 2021, *MNRAS*, 507, 398
- Hamilton, D. 1985, *ApJ*, 297, 371
- Haryana, N. S., Akiyama, M., Abdurro'uf, et al. 2025, *ApJ*, 994, 215
- Haskell, P., Das, S., Smith, D. J. B., et al. 2024, *MNRAS*, 530, L7
- Heavens, A. F., Jimenez, R., & Lahav, O. 2000, *MNRAS*, 317, 965
- Hennebelle, P. & Chabrier, G. 2008, *ApJ*, 684, 395
- Hill, A. R., Muzzin, A., Franx, M., et al. 2017, *ApJ*, 837, 147
- Hopkins, P. F. 2012, *MNRAS*, 423, 2037
- Hopkins, P. F. 2013, *MNRAS*, 433, 170
- Hopkins, P. F., Cox, T. J., Dutta, S. N., et al. 2009a, *ApJS*, 181, 135
- Hopkins, P. F., Cox, T. J., Kereš, D., & Hernquist, L. 2008, *ApJS*, 175, 390
- Hopkins, P. F., Croton, D., Bundy, K., et al. 2010, *ApJ*, 724, 915
- Hopkins, P. F., Hernquist, L., Cox, T. J., Keres, D., & Wuyts, S. 2009b, *ApJ*, 691, 1424
- Ito, K., Valentino, F., Brammer, G., et al. 2025, arXiv e-prints, arXiv:2506.22642
- Jafariyazani, M., Newman, A. B., Mobasher, B., et al. 2025, *ApJ*, 986, 148
- Jafariyazani, M., Newman, A. B., Mobasher, B., et al. 2020, *ApJ*, 897, L42
- Jeřábková, T., Zonoozi, A. H., Kroupa, P., et al. 2018, *A&A*, 620, A39
- Johnson, B., Foreman-Mackey, D., Sick, J., et al. 2023, *dfm/python-fsps: v0.4.6*
- Johnson, B. D., Leja, J., Conroy, C., & Speagle, J. S. 2021, *ApJS*, 254, 22
- Jones, G. T., Byrne, C. M., & Stanway, E. R. 2025, *MNRAS*, 543, 167
- Jones, G. T., Stanway, E. R., & Carnall, A. C. 2022, *Monthly Notices of the Royal Astronomical Society*, 514, 5706
- Jones, T., Ellis, R., Jullo, E., & Richard, J. 2010, *ApJ*, 725, L176
- Jones, T., Ellis, R. S., Richard, J., & Jullo, E. 2013, *ApJ*, 765, 48

- Jones, T., Wang, X., Schmidt, K. B., et al. 2015, *AJ*, 149, 107
- Ju, M., Wang, X., Jones, T., et al. 2025, *ApJ*, 978, L39
- Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003, *MNRAS*, 341, 54
- Kaushal, Y., Nersesian, A., Bezanson, R., et al. 2024, *ApJ*, 961, 118
- Kaviraj, S., Kirkby, L. A., Silk, J., & Sarzi, M. 2007, *MNRAS*, 382, 960
- Keating, S. K., Abraham, R. G., Schiavon, R., et al. 2015, *ApJ*, 798, 26
- Khochfar, S. & Silk, J. 2006, *ApJ*, 648, L21
- Kim, T., Kim, M., Ho, L. C., et al. 2025, *AJ*, 169, 44
- Knowles, A. T., Sansom, A. E., Allende Prieto, C., & Vazdekis, A. 2021, *MNRAS*, 504, 2286
- Knowles, A. T., Sansom, A. E., Vazdekis, A., & Allende Prieto, C. 2023, *MNRAS*, 523, 3450
- Kobayashi, C. 2004, *MNRAS*, 347, 740
- Kobayashi, C., Karakas, A. I., & Lugaro, M. 2020, *ApJ*, 900, 179
- Koekemoer, A. M., Aussel, H., Calzetti, D., et al. 2007, *ApJS*, 172, 196
- Koleva, M., Prugniel, P., Bouchard, A., & Wu, Y. 2009, *A&A*, 501, 1269
- Koleva, M., Prugniel, P., De Rijcke, S., & Zeilinger, W. W. 2011, *MNRAS*, 417, 1643
- Kormendy, J. & Djorgovski, S. 1989, *ARA&A*, 27, 235
- Kriek, M., Beverage, A. G., Price, S. H., et al. 2024, *ApJ*, 966, 36
- Kriek, M., Conroy, C., van Dokkum, P. G., et al. 2016, *Nature*, 540, 248
- Kriek, M., Labbé, I., Conroy, C., et al. 2010, *ApJ*, 722, L64
- Kriek, M., Price, S. H., Conroy, C., et al. 2019, *The Astrophysical Journal*, 880, L31
- Kriek, M., van der Wel, A., van Dokkum, P. G., Franx, M., & Illingworth, G. D. 2008a, *ApJ*, 682, 896
- Kriek, M., van Dokkum, P. G., Franx, M., et al. 2008b, *ApJ*, 677, 219
- Kriek, M., van Dokkum, P. G., Franx, M., et al. 2006, *ApJ*, 649, L71
- Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2009, *ApJ*, 700, 221

- Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2018, FAST: Fitting and Assessment of Synthetic Templates, Astrophysics Source Code Library, record ascl:1803.008
- Kroupa, P. 2001, MNRAS, 322, 231
- Krumholz, M. R., Klein, R. I., & McKee, C. F. 2011, ApJ, 740, 74
- Kuntschner, H., Emsellem, E., Bacon, R., et al. 2006, MNRAS, 369, 497
- Kuntschner, H., Emsellem, E., Bacon, R., et al. 2010, MNRAS, 408, 97
- La Barbera, F. & de Carvalho, R. R. 2009, ApJ, 699, L76
- La Barbera, F., de Carvalho, R. R., Gal, R. R., et al. 2005, ApJ, 626, L19
- La Barbera, F., Ferreras, I., Vazdekis, A., et al. 2013, MNRAS, 433, 3017
- La Barbera, F., Vazdekis, A., Ferreras, I., et al. 2019, MNRAS, 489, 4090
- Labbé, I., Huang, J., Franx, M., et al. 2005, ApJ, 624, L81
- Lacerna, I., Ibarra-Medel, H., Avila-Reese, V., et al. 2020, A&A, 644, A117
- Lagos, C. d. P., Valentino, F., Wright, R. J., et al. 2025, MNRAS, 536, 2324
- Lapiner, S., Dekel, A., Freundlich, J., et al. 2023, MNRAS, 522, 4515
- Larson, R. B. 1974, MNRAS, 166, 585
- Larson, R. B. 1986, MNRAS, 218, 409
- Larson, R. B. 1998, MNRAS, 301, 569
- Larson, R. B. & Tinsley, B. M. 1978, ApJ, 219, 46
- Law, D. R., E. Morrison, J., Argyriou, I., et al. 2023, AJ, 166, 45
- Lee, H., Worthey, G., Dotter, A., et al. 2009, The Astrophysical Journal, 694, 902
- Lee, H.-c., Worthey, G., Trager, S. C., & Faber, S. M. 2007, ApJ, 664, 215
- Lee, J. H., Kim, M., Kim, T., et al. 2025, AJ, 169, 185
- Lee, M. M., Steidel, C. C., Brammer, G., et al. 2024, MNRAS, 527, 9529
- Lee, Y. H., Hwang, H. S., Hwang, N., Lee, J. C., & Kim, K.-B. 2023, ApJ, 953, 88
- Leethochawalit, N., Kirby, E. N., Ellis, R. S., Moran, S. M., & Treu, T. 2019, ApJ, 885, 100
- Leja, J., Carnall, A. C., Johnson, B. D., Conroy, C., & Speagle, J. S. 2019a, ApJ, 876, 3
- Leja, J., Johnson, B. D., Conroy, C., et al. 2019b, ApJ, 877, 140

- Leja, J., Johnson, B. D., Conroy, C., van Dokkum, P. G., & Byler, N. 2017, *ApJ*, 837, 170
- Leja, J., Tacchella, S., & Conroy, C. 2019c, *ApJ*, 880, L9
- Li, C., Wang, E., Lin, L., et al. 2015, *ApJ*, 804, 125
- Li, H., Mao, S., Cappellari, M., et al. 2018, *MNRAS*, 476, 1765
- Liao, L.-W. & Cooper, A. P. 2023, *MNRAS*, 518, 3999
- Lin, L., Belfiore, F., Pan, H.-A., et al. 2017, *ApJ*, 851, 18
- Liu, F. S., Jiang, D., Faber, S. M., et al. 2017, *ApJ*, 844, L2
- Lonoce, I., Feldmeier-Krause, A., & Freedman, W. L. 2021, *ApJ*, 920, 93
- Lower, S., Narayanan, D., Leja, J., et al. 2020, *ApJ*, 904, 33
- Lu, S., Daddi, E., Maraston, C., et al. 2025, *Nature Astronomy*, 9, 128
- Lupton, R., Blanton, M. R., Fekete, G., et al. 2004, *PASP*, 116, 133
- Lyubenova, M., Martín-Navarro, I., van de Ven, G., et al. 2016, *MNRAS*, 463, 3220
- MacArthur, L. A., Courteau, S., Bell, E., & Holtzman, J. A. 2004, *ApJS*, 152, 175
- Madau, P. & Dickinson, M. 2014, *ARA&A*, 52, 415
- Maiolino, R. & Mannucci, F. 2019, *A&A Rev.*, 27, 3
- Maltby, D. T., Almaini, O., Wild, V., et al. 2018, *MNRAS*, 480, 381
- Mann, A. W., Feiden, G. A., Gaidos, E., Boyajian, T., & von Braun, K. 2015, *ApJ*, 804, 64
- Maraston, C. 1998, *MNRAS*, 300, 872
- Maraston, C. 2005, *MNRAS*, 362, 799
- Maraston, C., Daddi, E., Renzini, A., et al. 2006, *ApJ*, 652, 85
- Maraston, C. & Strömbäck, G. 2011, *MNRAS*, 418, 2785
- Martig, M., Bournaud, F., Teyssier, R., & Dekel, A. 2009, *ApJ*, 707, 250
- Martín-Navarro, I. 2016, *MNRAS*, 456, L104
- Martín-Navarro, I., Pérez-González, P. G., Trujillo, I., et al. 2015a, *ApJ*, 798, L4
- Martín-Navarro, I., Pinna, F., Coccato, L., et al. 2021, *A&A*, 654, A59
- Martín-Navarro, I., Vazdekis, A., Falcón-Barroso, J., et al. 2018, *MNRAS*, 475, 3700

- Martín-Navarro, I., Vazdekis, A., La Barbera, F., et al. 2015b, *ApJ*, 806, L31
- Martorano, M., van der Wel, A., Gebek, A., et al. 2026, *A&A*, 705, A236
- Maseda, M. V., Lewis, Z., Matthee, J., et al. 2023, *ApJ*, 956, 11
- Matteucci, F. 1994, *A&A*, 288, 57
- McConnell, N. J., Lu, J. R., & Mann, A. W. 2016, *ApJ*, 821, 39
- McCracken, H. J., Milvang-Jensen, B., Dunlop, J., et al. 2012, *A&A*, 544, A156
- McDermid, R. M., Alatalo, K., Blitz, L., et al. 2015, *MNRAS*, 448, 3484
- McDermid, R. M., Cappellari, M., Alatalo, K., et al. 2014, *ApJ*, 792, L37
- McGrath, E. J., Finkelstein, S. L., Barro, G., et al. 2026, *ApJ*, 999, L6
- Mehlert, D., Thomas, D., Saglia, R. P., Bender, R., & Wegner, G. 2003, *A&A*, 407, 423
- Mendel, J. T., Beifiori, A., Saglia, R. P., et al. 2020, *ApJ*, 899, 87
- Mendel, J. T., Saglia, R. P., Bender, R., et al. 2015, *ApJ*, 804, L4
- Miller, T. B., van Dokkum, P., & Mowla, L. 2023, *ApJ*, 945, 155
- Miller, T. B., Whitaker, K. E., Nelson, E. J., et al. 2022, *ApJ*, 941, L37
- Milone, A. D. C., Sansom, A. E., & Sánchez-Blázquez, P. 2011, *MNRAS*, 414, 1227
- Mo, H., van den Bosch, F. C., & White, S. 2010, *Galaxy Formation and Evolution*
- Moffat, A. F. J. 1969, *A&A*, 3, 455
- Moretti, L., Belli, S., Rudie, G. C., et al. 2026, *A&A*, 707, A146
- Morgan, W. W. & Keenan, P. C. 1973, *ARA&A*, 11, 29
- Morgan, W. W., Keenan, P. C., & Kellman, E. 1943, *An atlas of stellar spectra, with an outline of spectral classification*
- Mosleh, M., Tacchella, S., Renzini, A., et al. 2017, *ApJ*, 837, 2
- Moster, B. P., Naab, T., & White, S. D. M. 2013, *MNRAS*, 428, 3121
- Mowla, L. A., van Dokkum, P., Brammer, G. B., et al. 2019, *ApJ*, 880, 57
- Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013a, *ApJ*, 777, 18
- Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013b, *ApJS*, 206, 8
- Muzzin, A., Marchesini, D., van Dokkum, P. G., et al. 2009, *ApJ*, 701, 1839

- Naab, T., Johansson, P. H., & Ostriker, J. P. 2009, *ApJ*, 699, L178
- Naab, T., Johansson, P. H., Ostriker, J. P., & Efstathiou, G. 2007, *ApJ*, 658, 710
- Nanayakkara, T., Glazebrook, K., Schreiber, C., et al. 2025, *ApJ*, 981, 78
- Nelson, E. J., Tacchella, S., Diemer, B., et al. 2021, *MNRAS*, 508, 219
- Nelson, E. J., van Dokkum, P. G., Förster Schreiber, N. M., et al. 2016, *ApJ*, 828, 27
- Nersesian, A., van der Wel, A., Gallazzi, A., et al. 2024, *A&A*, 681, A94
- Nersesian, A., van der Wel, A., Gallazzi, A. R., et al. 2025, *A&A*, 695, A86
- Newman, A. B., Ellis, R. S., Bundy, K., & Treu, T. 2012, *ApJ*, 746, 162
- Newman, A. B., Gu, M., Belli, S., et al. 2025, arXiv e-prints, arXiv:2503.17478
- Newman, A. B., Smith, R. J., Conroy, C., Villaume, A., & van Dokkum, P. 2017, *ApJ*, 845, 157
- Nipoti, C., Treu, T., Leauthaud, A., et al. 2012, *MNRAS*, 422, 1714
- Nomoto, K., Kobayashi, C., & Tominaga, N. 2013, *ARA&A*, 51, 457
- Ocvirk, P., Pichon, C., Lançon, A., & Thiébaud, E. 2006, *MNRAS*, 365, 46
- Oke, J. B. & Gunn, J. E. 1983, *ApJ*, 266, 713
- Onodera, M., Carollo, C. M., Renzini, A., et al. 2015, *ApJ*, 808, 161
- Oort, J. H. 1958, *Ricerche Astronomiche*, 5, 415
- Oser, L., Naab, T., Ostriker, J. P., & Johansson, P. H. 2012, *ApJ*, 744, 63
- Oser, L., Ostriker, J. P., Naab, T., Johansson, P. H., & Burkert, A. 2010, *ApJ*, 725, 2312
- Oyarzún, G. A., Bundy, K., Westfall, K. B., et al. 2019, *ApJ*, 880, 111
- Oyarzún, G. A., Bundy, K., Westfall, K. B., et al. 2023, *ApJ*, 947, 13
- Pacifici, C., Iyer, K. G., Mobasher, B., et al. 2023, *ApJ*, 944, 141
- Pacifici, C., Kassin, S. A., Weiner, B. J., et al. 2016, *ApJ*, 832, 79
- Papovich, C., Dickinson, M., & Ferguson, H. C. 2001, *ApJ*, 559, 620
- Parikh, T., Saglia, R., Thomas, J., et al. 2024, *MNRAS*, 528, 7338
- Parikh, T., Thomas, D., Maraston, C., et al. 2021, *MNRAS*, 502, 5508
- Park, M., Belli, S., Conroy, C., et al. 2024, *ApJ*, 976, 72

- Park, M., Conroy, C., Johnson, B. D., et al. 2025, *ApJ*, 994, 165
- Pasquali, A., Gallazzi, A., Fontanot, F., et al. 2010, *MNRAS*, 407, 937
- Pastorello, N., Forbes, D. A., Foster, C., et al. 2014, *MNRAS*, 442, 1003
- Pathak, D., Belli, S., & Weinberger, R. 2021, *ApJ*, 916, L23
- Peña Ramírez, K., Béjar, V. J. S., Zapatero Osorio, M. R., Petr-Gotzens, M. G., & Martín, E. L. 2012, *ApJ*, 754, 30
- Peletier, R. F., Davies, R. L., Illingworth, G. D., Davis, L. E., & Cawson, M. 1990a, *AJ*, 100, 1091
- Peletier, R. F. & Valentijn, E. A. 1989, *AP&SS*, 156, 127
- Peletier, R. F., Valentijn, E. A., & Jameson, R. F. 1990b, *A&A*, 233, 62
- Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2010, *AJ*, 139, 2097
- Peng, Y., Maiolino, R., & Cochrane, R. 2015, *Nature*, 521, 192
- Percival, S. M., Salaris, M., Cassisi, S., & Pietrinferni, A. 2009, *ApJ*, 690, 427
- Pérez-González, P. G., D'Eugenio, F., Rodríguez del Pino, B., et al. 2025, *Nature Astronomy*, 9, 1240
- Perna, M., Arribas, S., Marshall, M., et al. 2023, *A&A*, 679, A89
- Perrin, M. D., Long, J., Sivaramakrishnan, A., et al. 2015, *WebbPSF: James Webb Space Telescope PSF Simulation Tool*, *Astrophysics Source Code Library*, record ascl:1504.007
- Pipino, A., D'Ercole, A., Chiappini, C., & Matteucci, F. 2010, *MNRAS*, 407, 1347
- Pipino, A., D'Ercole, A., & Matteucci, F. 2008, *A&A*, 484, 679
- Pizzardo, M., Damjanov, I., Sohn, J., & Geller, M. J. 2026, *A&A*, 705, A30
- Plat, A., Charlot, S., Bruzual, G., et al. 2019, *MNRAS*, 490, 978
- Poggianti, B. M., Moretti, A., Calvi, R., et al. 2013, *ApJ*, 777, 125
- Popping, G., Decarli, R., Man, A. W. S., et al. 2017, *A&A*, 602, A11
- Posacki, S., Cappellari, M., Treu, T., Pellegrini, S., & Ciotti, L. 2015, *MNRAS*, 446, 493
- Pracy, M. B., Croom, S., Sadler, E., et al. 2013, *MNRAS*, 432, 3131
- Price, S. H., Kriek, M., Shapley, A. E., et al. 2016, *ApJ*, 819, 80
- Pulsoni, C., Gerhard, O., Arnaboldi, M., et al. 2021, *A&A*, 647, A95
- Quilis, V. & Trujillo, I. 2013, *ApJ*, 773, L8

- Rayner, J. T., Cushing, M. C., & Vacca, W. D. 2009, *ApJS*, 185, 289
- Reddy, N. A., Erb, D. K., Pettini, M., Steidel, C. C., & Shapley, A. E. 2010, *ApJ*, 712, 1070
- Reddy, N. A., Pettini, M., Steidel, C. C., et al. 2012, *ApJ*, 754, 25
- Renzini, A. 2006, *ARA&A*, 44, 141
- Renzini, A. 2009, *MNRAS*, 398, L58
- Renzini, A. & Buzzoni, A. 1986, in *Astrophysics and Space Science Library*, Vol. 122, *Spectral Evolution of Galaxies*, ed. C. Chiosi & A. Renzini, 195–231
- Renzini, A. & Ciotti, L. 1993, *ApJ*, 416, L49
- Roche, N., Bernardi, M., & Hyde, J. 2010, *MNRAS*, 407, 1231
- Rodriguez-Gomez, V., Pillepich, A., Sales, L. V., et al. 2016, *MNRAS*, 458, 2371
- Rose, J. A., Arimoto, N., Caldwell, N., et al. 2005, *AJ*, 129, 712
- Sagar, R., Munari, U., & de Boer, K. S. 2001, *MNRAS*, 327, 23
- Saglia, R. P., Maraston, C., Greggio, L., Bender, R., & Ziegler, B. 2000, *A&A*, 360, 911
- Salpeter, E. E. 1955, *ApJ*, 121, 161
- San Roman, I., Cenarro, A. J., Díaz-García, L. A., et al. 2018, *A&A*, 609, A20
- Sánchez, S. F., Avila-Reese, V., Hernandez-Toledo, H., et al. 2018, *Rev. Mexicana Astron. Astrofis.*, 54, 217
- Sánchez, S. F., Kennicutt, R. C., Gil de Paz, A., et al. 2012, *A&A*, 538, A8
- Sánchez-Blázquez, P., Forbes, D. A., Strader, J., Brodie, J., & Proctor, R. 2007, *MNRAS*, 377, 759
- Sánchez-Blázquez, P., Peletier, R. F., Jiménez-Vicente, J., et al. 2006, *MNRAS*, 371, 703
- Sandage, A. 1972, *ApJ*, 176, 21
- Santucci, G., Brough, S., Scott, N., et al. 2020, *ApJ*, 896, 75
- Sawicki, M. & Yee, H. K. C. 1998, *AJ*, 115, 1329
- Schwarzschild, M. & Spitzer, L. 1953, *The Observatory*, 73, 77
- Scoville, N., Aussel, H., Benson, A., et al. 2007a, *ApJS*, 172, 150
- Scoville, N., Aussel, H., Brusa, M., et al. 2007b, *ApJS*, 172, 1
- Searle, L., Sargent, W. L. W., & Bagnuolo, W. G. 1973, *ApJ*, 179, 427

- Serven, J., Worthey, G., & Briley, M. M. 2005, *ApJ*, 627, 754
- Setton, D. J., Bezanson, R., Suess, K. A., et al. 2020, *ApJ*, 905, 79
- Setton, D. J., Khullar, G., Miller, T. B., et al. 2024, *ApJ*, 974, 145
- Shen, L., Papovich, C., Matharu, J., et al. 2024, *ApJ*, 963, L49
- Shuntov, M., Akins, H. B., Paquereau, L., et al. 2025, *A&A*, 704, A339
- Siegel, J. C., Setton, D. J., Greene, J. E., et al. 2025, *ApJ*, 985, 125
- Slob, M., Kriek, M., Beverage, A. G., et al. 2024, *ApJ*, 973, 131
- Slob, M., Kriek, M., de Graaff, A., et al. 2025, *A&A*, 702, A110
- Smail, I., Owen, F. N., Morrison, G. E., et al. 2002, *ApJ*, 581, 844
- Smith, J. D. T., Armus, L., Dale, D. A., et al. 2007, *PASP*, 119, 1133
- Smith, R. J. 2020, *ARA&A*, 58, 577
- Smith, R. J., Lucey, J. R., & Hudson, M. J. 2009, *MNRAS*, 400, 1690
- Snyder, G. F., Cox, T. J., Hayward, C. C., Hernquist, L., & Jonsson, P. 2011, *ApJ*, 741, 77
- Somerville, R. S. & Davé, R. 2015, *ARA&A*, 53, 51
- Speagle, J. S. 2020, *MNRAS*, 493, 3132
- Spergel, D. N., Bean, R., Doré, O., et al. 2007, *ApJS*, 170, 377
- Spilker, J., Bezanson, R., Barišić, I., et al. 2018, *ApJ*, 860, 103
- Spilker, J. S., Bezanson, R., Weiner, B. J., Whitaker, K. E., & Williams, C. C. 2019, *ApJ*, 883, 81
- Spiniello, C., Trager, S. C., Koopmans, L. V. E., & Chen, Y. P. 2012, *ApJ*, 753, L32
- Spitler, L. R., Straatman, C. M. S., Labbé, I., et al. 2014, *ApJ*, 787, L36
- Spitoni, E., Vincenzo, F., & Matteucci, F. 2017, *A&A*, 599, A6
- Spolaor, M., Kobayashi, C., Forbes, D. A., Couch, W. J., & Hau, G. K. T. 2010, *MNRAS*, 408, 272
- Springel, V., Pakmor, R., Pillepich, A., et al. 2018, *MNRAS*, 475, 676
- Steinhardt, C. L., Capak, P., Masters, D., & Speagle, J. S. 2016, *ApJ*, 824, 21
- Stevenson, S. D., Carnall, A. C., Leung, H.-H., et al. 2026, *MNRAS*, 545, staf2087
- Straatman, C. M. S., Labbé, I., Spitler, L. R., et al. 2014, *ApJ*, 783, L14

- Straatman, C. M. S., Spitler, L. R., Quadri, R. F., et al. 2016, *ApJ*, 830, 51
- Straatman, C. M. S., van der Wel, A., Bezanson, R., et al. 2018, *The Astrophysical Journal Supplement Series*, 239, 27
- Stringer, M., Trujillo, I., Dalla Vecchia, C., & Martinez-Valpuesta, I. 2015, *MNRAS*, 449, 2396
- Suárez, G., Downes, J. J., Román-Zúñiga, C., et al. 2019, *MNRAS*, 486, 1718
- Suess, K. A., Kriek, M., Price, S. H., & Barro, G. 2019a, *ApJ*, 877, 103
- Suess, K. A., Kriek, M., Price, S. H., & Barro, G. 2019b, *ApJ*, 885, L22
- Suess, K. A., Kriek, M., Price, S. H., & Barro, G. 2020, *ApJ*, 899, L26
- Suess, K. A., Kriek, M., Price, S. H., & Barro, G. 2021, *ApJ*, 915, 87
- Suess, K. A., Leja, J., Johnson, B. D., et al. 2022, *ApJ*, 935, 146
- Suess, K. A., Williams, C. C., Robertson, B., et al. 2023, *ApJ*, 956, L42
- Suh, H., Jeong, H., Oh, K., et al. 2010, *ApJS*, 187, 374
- Sybilska, A., Kuntschner, H., van de Ven, G., et al. 2018, *MNRAS*, 476, 4501
- Szomoru, D., Franx, M., & van Dokkum, P. G. 2012, *ApJ*, 749, 121
- Szomoru, D., Franx, M., van Dokkum, P. G., et al. 2013, *ApJ*, 763, 73
- Tacchella, S., Carollo, C. M., Förster Schreiber, N. M., et al. 2018, *ApJ*, 859, 56
- Tacchella, S., Carollo, C. M., Renzini, A., et al. 2015a, *Science*, 348, 314
- Tacchella, S., Conroy, C., Faber, S. M., et al. 2022, *ApJ*, 926, 134
- Tacchella, S., Dekel, A., Carollo, C. M., et al. 2016a, *MNRAS*, 458, 242
- Tacchella, S., Dekel, A., Carollo, C. M., et al. 2016b, *MNRAS*, 457, 2790
- Tacchella, S., Diemer, B., Hernquist, L., et al. 2019, *MNRAS*, 487, 5416
- Tacchella, S., Lang, P., Carollo, C. M., et al. 2015b, *ApJ*, 802, 101
- Tadaki, K.-i., Kodama, T., Nelson, E. J., et al. 2017, *ApJ*, 841, L25
- Thomas, D., Greggio, L., & Bender, R. 1999, *MNRAS*, 302, 537
- Thomas, D., Maraston, C., & Bender, R. 2003, *MNRAS*, 339, 897
- Thomas, D., Maraston, C., Bender, R., & Mendes de Oliveira, C. 2005, *ApJ*, 621, 673
- Thomas, D., Maraston, C., Schawinski, K., Sarzi, M., & Silk, J. 2010, *MNRAS*, 404, 1775

- Tinsley, B. M. 1972, *A&A*, 20, 383
- Tinsley, B. M. 1979, *ApJ*, 229, 1046
- Tinsley, B. M. 1980, , 5, 287
- Toft, S., van Dokkum, P., Franx, M., et al. 2007, *ApJ*, 671, 285
- Tojeiro, R., Heavens, A. F., Jimenez, R., & Panter, B. 2007, *MNRAS*, 381, 1252
- Tortora, C., Napolitano, N. R., Cardone, V. F., et al. 2010, *MNRAS*, 407, 144
- Tortora, C., Napolitano, N. R., Romanowsky, A. J., et al. 2011, *MNRAS*, 418, 1557
- Tortorelli, L., McCullough, J., & Gruen, D. 2024, *A&A*, 689, A144
- Trager, S. C., Faber, S. M., Worthey, G., & González, J. J. 2000, *AJ*, 120, 165
- Trager, S. C. & Somerville, R. S. 2009, *MNRAS*, 395, 608
- Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, *ApJ*, 613, 898
- Treu, T., Auger, M. W., Koopmans, L. V. E., et al. 2010, *ApJ*, 709, 1195
- Tripodi, R., D'Eugenio, F., Maiolino, R., et al. 2024, *A&A*, 692, A184
- Trujillo, I., Aguerri, J. A. L., Cepa, J., & Gutiérrez, C. M. 2001, *MNRAS*, 328, 977
- Trujillo, I., Cenarro, A. J., de Lorenzo-Cáceres, A., et al. 2009, *ApJ*, 692, L118
- Trujillo, I., Förster Schreiber, N. M., Rudnick, G., et al. 2006, *ApJ*, 650, 18
- Trujillo, I., Rudnick, G., Rix, H.-W., et al. 2004, *ApJ*, 604, 521
- Trussler, J., Maiolino, R., Maraston, C., et al. 2020, *MNRAS*, 491, 5406
- Valentino, F., Brammer, G., Gould, K. M. L., et al. 2023, *ApJ*, 947, 20
- Valentinuzzi, T., Fritz, J., Poggianti, B. M., et al. 2010, *ApJ*, 712, 226
- 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 den Bosch, F. C., Aquino, D., Yang, X., et al. 2008, *MNRAS*, 387, 79
- van der Wel, A., Bezanson, R., D'Eugenio, F., et al. 2021, *The Astrophysical Journal Supplement Series*, 256, 44
- van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, *ApJ*, 788, 28
- van der Wel, A., Holden, B. P., Zirm, A. W., et al. 2008, *ApJ*, 688, 48

- van der Wel, A., Martorano, M., Häußler, B., et al. 2024, *ApJ*, 960, 53
- van der Wel, A., Noeske, K., Bezanson, R., et al. 2016, *The Astrophysical Journal Supplement Series*, 223, 29
- van der Wel, A., van Houdt, J., Bezanson, R., et al. 2022, *ApJ*, 936, 9
- van Dokkum, P. & Conroy, C. 2024, *ApJ*, 973, L32
- van Dokkum, P., Conroy, C., Villaume, A., Brodie, J., & Romanowsky, A. J. 2017, *ApJ*, 841, 68
- van Dokkum, P. G., Bezanson, R., van der Wel, A., et al. 2014, *ApJ*, 791, 45
- van Dokkum, P. G. & Franx, M. 2001, *ApJ*, 553, 90
- van Dokkum, P. G., Franx, M., Kriek, M., et al. 2008, *ApJ*, 677, L5
- van Dokkum, P. G., Whitaker, K. E., Brammer, G., et al. 2010, *ApJ*, 709, 1018
- van Houdt, J., van der Wel, A., Bezanson, R., et al. 2021, *The Astrophysical Journal*, 923, 11
- Vazdekis, A. 2001, *AP&SS*, 276, 921
- Vazdekis, A., Coelho, P., Cassisi, S., et al. 2015, *MNRAS*, 449, 1177
- Vazdekis, A., Koleva, M., Ricciardelli, E., Röck, B., & Falcón-Barroso, J. 2016, *MNRAS*, 463, 3409
- Vazdekis, A., Peletier, R. F., Beckman, J. E., & Casuso, E. 1997, *ApJS*, 111, 203
- Vernet, J., Dekker, H., D’Odorico, S., et al. 2011, *A&A*, 536, A105
- Villaume, A., Brodie, J., Conroy, C., Romanowsky, A. J., & van Dokkum, P. 2017a, *ApJ*, 850, L14
- Villaume, A., Conroy, C., Johnson, B., et al. 2017b, *ApJS*, 230, 23
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *Nature Methods*, 17, 261
- Walcher, C. J., Böker, T., Charlot, S., et al. 2006, *ApJ*, 649, 692
- Walcher, J., Groves, B., Budavári, T., & Dale, D. 2011, *AP&SS*, 331, 1
- Weibel, A., de Graaff, A., Setton, D. J., et al. 2025, *ApJ*, 983, 11
- Weinmann, S. M., Kauffmann, G., van den Bosch, F. C., et al. 2009, *MNRAS*, 394, 1213
- Wellons, S., Torrey, P., Ma, C.-P., et al. 2016, *MNRAS*, 456, 1030
- Wellons, S., Torrey, P., Ma, C.-P., et al. 2015, *MNRAS*, 449, 361
- Wetzell, A. R., Tinker, J. L., & Conroy, C. 2012, *MNRAS*, 424, 232

- Whitaker, K. E., Kriek, M., van Dokkum, P. G., et al. 2012, *ApJ*, 745, 179
- Whitaker, K. E., van Dokkum, P. G., Brammer, G., et al. 2010, *ApJ*, 719, 1715
- Whitaker, K. E., van Dokkum, P. G., Brammer, G., et al. 2013, *ApJ*, 770, L39
- White, S. D. M. 1980, *MNRAS*, 191, 1P
- White, S. D. M. & Rees, M. J. 1978, *MNRAS*, 183, 341
- Whitler, L., Stark, D. P., Endsley, R., et al. 2023, *MNRAS*, 519, 5859
- Wing, R. F. & Ford, Jr., W. K. 1969, *PASP*, 81, 527
- Worthey, G. 1994, *ApJS*, 95, 107
- Worthey, G., Faber, S. M., & Gonzalez, J. J. 1992, *ApJ*, 398, 69
- Worthey, G., Faber, S. M., Gonzalez, J. J., & Burstein, D. 1994, *ApJS*, 94, 687
- Worthey, G. & Ottaviani, D. L. 1997, *ApJS*, 111, 377
- Wu, P.-F. 2021, *ApJ*, 913, 44
- Wu, P.-F., van der Wel, A., Bezanson, R., et al. 2020, *ApJ*, 888, 77
- Wuyts, S., Cox, T. J., Hayward, C. C., et al. 2010, *ApJ*, 722, 1666
- Wuyts, S., Franx, M., Cox, T. J., et al. 2009, *ApJ*, 696, 348
- Wuyts, S., Labbé, I., Franx, M., et al. 2007, *ApJ*, 655, 51
- Xiao, M., Oesch, P. A., Elbaz, D., et al. 2024, *Nature*, 635, 311
- Yan, Z., Li, J., Kroupa, P., et al. 2024, *ApJ*, 969, 95
- Yano, M., Kriek, M., van der Wel, A., & Whitaker, K. E. 2016, *ApJ*, 817, L21
- Yoon, Y., Ko, J., & Kim, J.-W. 2023, *ApJ*, 946, 41
- Zheng, Z., Li, C., Mao, S., et al. 2019, *ApJ*, 873, 63
- Zhuang, Z., Leethochawalit, N., Kirby, E. N., et al. 2023, *ApJ*, 948, 132
- Zibetti, S., Charlot, S., & Rix, H.-W. 2009, *MNRAS*, 400, 1181
- Zibetti, S., Gallazzi, A. R., Hirschmann, M., et al. 2020, *MNRAS*, 491, 3562
- Zick, T. O., Kriek, M., Shapley, A. E., et al. 2018, *ApJ*, 867, L16
- Zieleniewski, S., Houghton, R. C. W., Thatte, N., Davies, R. L., & Vaughan, S. P. 2017, *MNRAS*, 465, 192
- Zolotov, A., Dekel, A., Mandelker, N., et al. 2015, *MNRAS*, 450, 2327
- Zolotov, A., Willman, B., Brooks, A. M., et al. 2010, *ApJ*, 721, 738

ENGLISH SUMMARY

In order to understand the formation and evolution of structure in our Universe, we first need to ask how galaxies and their host dark matter haloes, the Universe's building blocks, have grown and changed over time. Massive, quiescent galaxies hold the answers as some of the most extreme systems in existence. These ancient cosmic fossils, which have ceased forming stars, currently account for most of the Universe's stellar mass. Their properties challenge theories of cosmic structure formation, making them unique laboratories for probing the fundamental physics on which our Universe operates.

The massive galaxy formation case file

The Universe began about 13.8 billion years ago, in an extremely hot and dense state, with the Big Bang (see Figure 1). In an instant, space itself began to expand and cool, collapsing into larger and larger structures over time, from particles, to atoms, to stars and galaxies. In this way, structure in the Universe forms hierarchically. This model of the formation of the Universe is called Λ CDM cosmology.

Galaxies are huge collections of stars, gas, dust, and dark matter held together by gravity. They grow via infalling gas, by forming new stars from cold gas, and by merging with other, nearby galaxies. Eventually, many galaxies stop forming stars, a process known as quenching. In our current galaxy formation picture, galaxies may have formed in a hierarchical manner similar to the larger cosmic

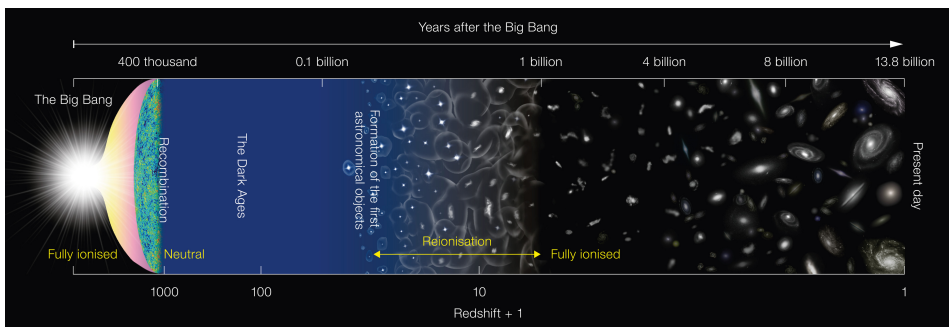


Figure 1: A diagram depicting the history of the Universe, beginning with the Big Bang. Image reproduced from ESO/NAOJ.

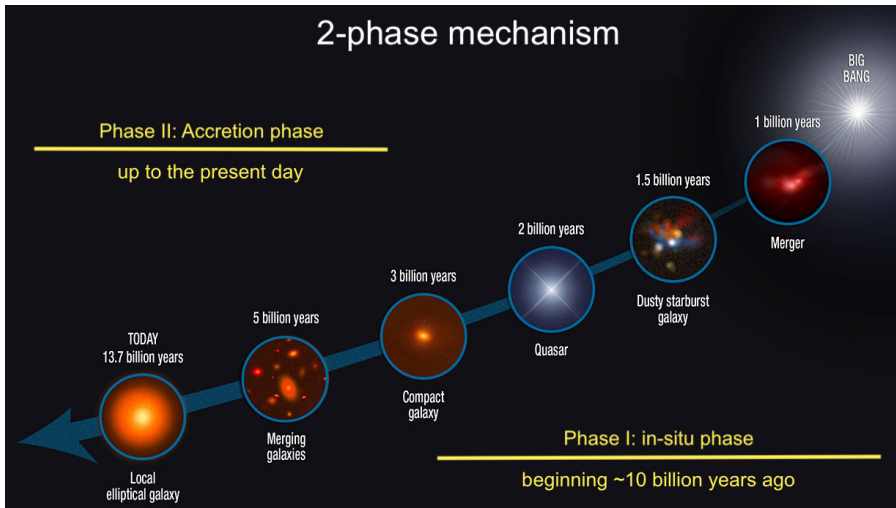


Figure 2: A diagram illustrating the two-phase model of massive galaxy formation. Figure adapted from NASA/ESA/Sune Toft.

structure of the Universe. Specifically, they may have formed in two phases (see Figure 2): their dense, compact centres likely formed early, about 10 billion years ago, during a period of rapid star formation. After star formation ended, they may have continued to grow, building up their outer regions by merging with tiny, neighbouring galaxies. This picture is supported by observations that distant, quiescent galaxies are more compact compared to nearby galaxies.

However, alternative galaxy formation scenarios are also possible. For example, it may be that individual galaxies are not growing at all. Instead, galaxies that quench earlier may be smaller, while galaxies that quench closer to the present day may be larger. Additionally, it is unclear how and why galaxies stop forming stars and become quiescent. There are a variety of processes that can allow galaxies to achieve this quenched state, by removing, using up, or heating the reservoir of cold gas needed to form stars. In reality, a combination of processes are likely occurring. Nonetheless, isolating the mechanisms by which the majority of galaxies form, quench, and subsequently evolve is crucial to our broader understanding of structure formation in the Universe.

To achieve this understanding, we can leverage cutting-edge observatories, such as the James Webb Space Telescope (*JWST*). Commencing operations in 2021 after one of the most challenging launches ever achieved, *JWST* is an unparalleled international collaboration between the US, Europe, and Canada. With a cost of \$10 billion USD and ongoing operations of \$10 000/hour, this investment has transformed science. Specifically, *JWST* has facilitated the discovery of massive, quiescent galaxies at much earlier times than expected, when the Universe was about 2 billion years old. These findings suggest that galaxies began and ended star formation faster and earlier than predicted by Λ CDM cosmology, putting pressure on galaxy assembly theories and our cosmological framework.

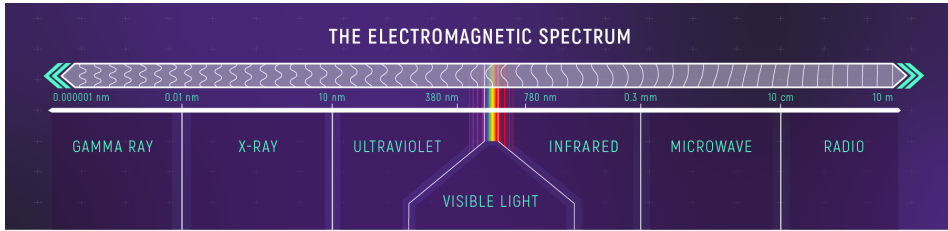


Figure 3: An illustration of the electromagnetic spectrum of light. We can disperse this light into broad colours spanning a range of wavelengths (photometry), or into individual wavelengths (spectroscopy). Image reproduced from NASA/ESA/Leah Hustak (STScI).

However, this interpretation relies on limited data and models. In particular, we rely on fitting stellar population models to a handful of photometric data points. These models may predict biased properties due to broad assumptions about galaxies' underlying chemistry and physics. Moreover, our data are limited in that they depict only 2% of the stars (the brightest and most massive) that we can see in these very distant galaxies. Finally, they represent tiny galaxy samples. As a result, these major claims about galaxy formation may be incorrect. In the era of *JWST*, traditional methods are insufficient.

Our sleuthing kit for investigating galaxy histories

The components of a galaxy, including its stellar population properties, provide clues to its evolution. This information is encoded in the chemical makeup of its stars. While we can never directly observe a galaxy's full history, we can study its past history indirectly via its stellar light. Measuring the total brightness of a galaxy in a small number of broad, photometric colours provides an initial view of its stellar content. By further dispersing galaxy light into individual wavelengths, we can unpack galaxy assembly signatures imprinted in this spectrum (see Figures 3 and 4).

We can use these data to unlock insights into galaxy formation, by inferring the properties of galaxies' constituent stellar populations. Interpreting the light from these groups of stars requires a deep understanding of the ingredients on which they are built. In particular, each star is born with a certain mass. We need to know the distribution of these birth stellar masses (called the initial mass function) to describe stellar populations. The initial mass function may vary with environment, but this is still a major open area of investigation. Additionally, we need to be able to describe how stars of different masses and chemical compositions evolve over time. Finally, we need to understand what kinds of spectra stars produce, as each type of star can change both the overall shape of a galaxy's spectrum, as well as the presence of certain dips in a spectrum (spectral lines). By studying atoms in the lab, scientists have calculated how each element absorbs light at specific wavelengths, with each element having its own spectral fingerprint. For example, hydrogen atoms absorb light at well-known wavelengths (see Figure 5), so when

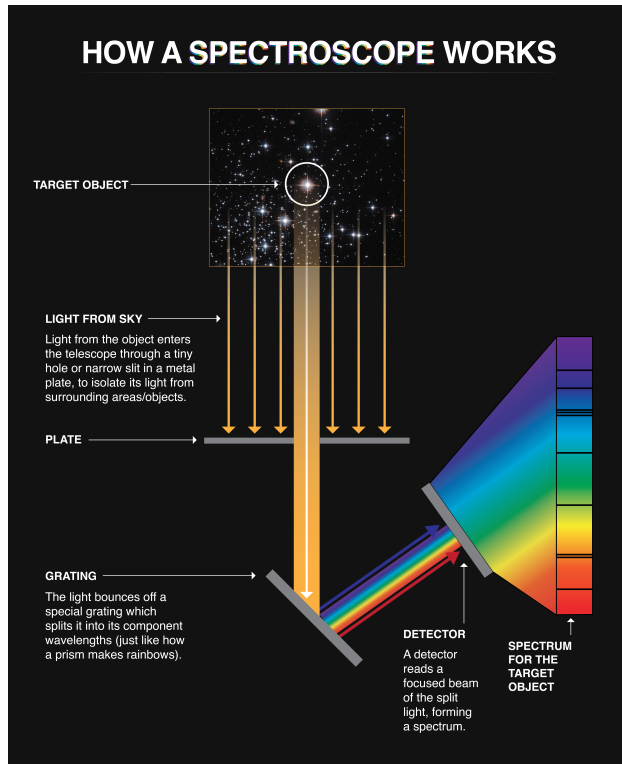


Figure 4: A diagram showing how we can obtain spectra of galaxies using a spectrograph on a telescope. Image reproduced from NASA Goddard/Shireen Dooling.

we observe these same dips or lines in a galaxy spectrum, we know that hydrogen is present. We can then quantify the amount of hydrogen based on the depths of these features.

By understanding these ingredients, we can infer the ages and detailed chemistry of stellar populations, which provide a window into the processes that drove a galaxy's evolution. In particular, a galaxy's past star formation and interactions can be inferred from its overall chemistry. Additionally, different types of stars forge different elements on varying timescales as they evolve and die. The ratios of elements can give us insight into the chemical evolution and star formation histories of a galaxy. However, interpreting observed galaxy data is complicated by the fact that different properties can have similar observed effects on the shapes of spectra and the depths of spectral lines. For example, changes in stellar age and chemistry can have nearly identical spectral impacts. This is known as the age-metallicity degeneracy. Additionally, variations in the shape of the initial mass function can have similar effects as variations in chemistry. Thus, correctly disentangling the specific age, chemical composition, and initial mass function of a galaxy in order to understand its place in evolutionary history is extremely challenging.

We can apply these physical principles to real data, in order to interpret galax-

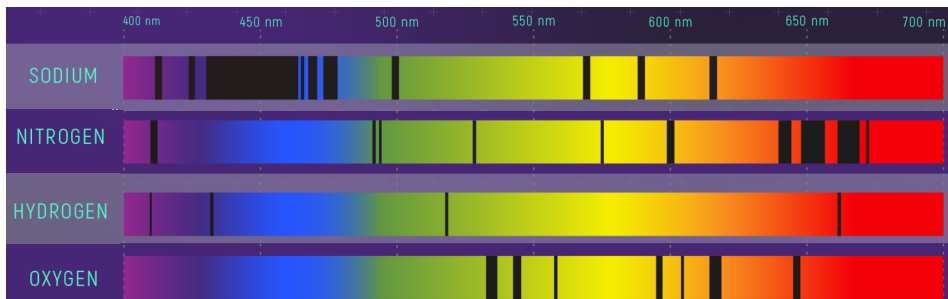


Figure 5: Absorption spectra for various elements. Image adapted from NASA/ESA/Leah Hustak (STScI).

ies' constituent stellar populations. To disentangle galaxy ages and detailed chemical compositions, and to constrain the initial mass function in diverse galaxies, we can fit models to spectra. This can be done by examining individual spectral features, which are primarily sensitive to certain properties. For example, the strengths of hydrogen lines are primarily sensitive to the age of a galaxy, while iron lines are primarily sensitive to the amount of iron in a galaxy.

In practice, however, individual elements and galaxy properties contribute to the signal of the entire spectrum. We can thus gain more information and more robustly constrain galaxy properties by modeling the entire spectrum of a galaxy simultaneously. In this way, we can also constrain a galaxy's initial mass function, which presents itself as extremely subtle variations across the whole spectrum.

Dusting for galaxy formation fingerprints

The tools described above have been applied to data of a wide variety of galaxies. In particular, by looking at light averaged over whole galaxies, we have learned that massive galaxies in the nearby Universe are also the oldest objects and are rich in elements like iron and magnesium. These findings suggest that these galaxies both formed their stars and quenched early and quickly. Moreover, we have learned that nearby, massive galaxies have an excess of low-mass stars in their centres compared to the Milky Way, indicating that their initial mass functions are different.

These measurements have guided our understanding of the global properties of galaxies, but we can unlock more detailed insights by spatially resolving a galaxy's constituent stellar populations. Stellar population gradients encode stellar mass build-up, allowing us to differentiate possible quenching and assembly mechanisms. Spatially resolved data have shown us that nearby, massive galaxies tend to have redder colours in their centres and bluer colours in their outskirts, which are driven by the fact that galaxy centres are richer in iron compared to their outskirts. These findings reinforce the idea that galaxies formed hierarchically.

While we have learned a lot about galaxy evolution from studying nearby, massive galaxies, it is important to keep in mind that in the two-phase model, we expect these galaxies to be in a relatively late stage of evolution. In particular,

their outskirts may have already been built up by merging with tiny, neighbouring galaxies. To truly understand the formation of massive, quiescent galaxies, we must look back in time, and study their more pristine stellar populations, closer to the epoch when they formed. Because light takes time to travel, observing galaxies at greater distances allows us to see them as they were earlier in the Universe's history. As a result of the expansion of the universe, this time is quantified by redshift, with higher redshift corresponding to earlier cosmic times.

Observing these distant, high-redshift galaxies is more challenging than studying nearby galaxies, as they are much dimmer due to their extreme distance from us. Additionally, their element-sensitive absorption features are very faint and redshifted into wavelength regions where our telescopes are less sensitive. Thus, we must observe individual distant galaxies for tens to hundreds of hours to measure detailed ages and chemical compositions.

With innovative technological advancements in astronomical observations over the last ~ 15 years, we have begun to survey massive, quiescent galaxies out to very early times. Surprisingly, our picture of galaxy formation is actually becoming muddier as we reach back towards the advent of the Universe. In particular, it is not clear how these high-redshift, quiescent galaxies evolve into those we see in the nearby Universe (if they are even related). For example, as discussed above, *JWST* has facilitated the discovery of massive, quiescent galaxies at much earlier times than predicted by Λ CDM simulations. The formation and evolution of massive, quiescent galaxies thus remains an open question. To make progress in isolating the dominant mechanisms of stellar mass assembly and star formation quenching, and to confirm findings of 'impossibly early' galaxies, we require large samples of distant, massive galaxies with high-quality and spatially resolved data.

This thesis: identifying the culprits

In this thesis, we provide some of the most detailed constraints to date on galaxy formation out to ~ 10 billion years ago. Through flexible modeling of the highest quality integrated and spatially resolved spectra, we present novel measurements of age and chemical gradients in distant galaxies, reveal discrepancies in model predictions, and make the first robust measurements of the initial mass function in the distant Universe.

In **Chapter 2**, we measure gradients in a large sample of distant, massive, quiescent galaxies, using data from the LEGA-C survey. These galaxies existed ~ 7 billion years ago, when the Universe was about half of its current age. We find that the stars in the centres of these galaxies have the same age and magnesium content as in their outskirts, but have higher amounts of iron. Our findings are consistent with a scenario in which these galaxies underwent a large amount of star formation in their centres before they quenched, followed by their outskirts being built up by small, neighbouring galaxies with lower amounts of iron. However, alternative scenarios can also explain our results.

In **Chapter 3**, we combine our detailed spectroscopic measurements with independent photometric colours to challenge predictions made by stellar population

models. We measure ages and chemical compositions of a large sample of massive, quiescent galaxies, again from the LEGA-C survey. We compare the measured colours of our galaxies to the predicted colours based on our measured ages and chemical compositions. We find that stellar population models are not able to reproduce the age-colour and age-chemical composition trends that we find in the data. Our results demonstrate that commonly used stellar population models are incorrect, illustrating that the underlying physical assumptions on which these models are built must urgently be re-assessed.

In **Chapter 4**, we measure age and chemical gradients in 8 massive, quiescent galaxies that existed ~ 10 billion years ago, close to the epoch when these galaxies formed and quenched. Using data from the *JWST*-SUSPENSE program, we find that the stars in the centres of these galaxies are older and richer in magnesium compared to their outskirts, but have the same amount of iron. These findings suggest a picture where the centres of galaxies quenched their star formation first, followed by their outskirts. Towards the present day, these galaxies may have built up their outskirts through merging with small, neighbouring galaxies to produce the gradients that we found in Chapter 2. On the other hand, it may be that different quenching mechanisms are occurring at different times and result in different gradients.

Finally, in **Chapter 5**, we present the first robust measurements of the initial mass function, in a sample of 9 massive, quiescent galaxies beyond the nearby Universe. This study is based on data from our *JWST*-IMFERN0 program. We find that the most massive, distant galaxies have excess low-mass stars compared to the Milky Way. In combination with previous findings that the centres of nearby, massive, quiescent galaxies also have excess low-mass stars, our results are consistent with the two-phase model. Our oldest and most massive galaxy is a likely descendant of the distant, massive, ‘impossibly early’ galaxies found with *JWST*, implying that they may have also had this excess of low-mass stars. Our constraints suggest that their stellar masses are even greater than originally reported, heightening the tension between observations of massive, early galaxies and current galaxy formation models.

The case continues...

This work has contributed substantially to our understanding of massive galaxy formation over the last ~ 10 billion years, however questions still remain. To identify the dominant galaxy formation pathways, we must examine larger samples of galaxies with even higher spatial resolution, taking advantage of the next generation of modern telescopes (including the *Extremely Large Telescope*), and simultaneously developing more sophisticated modeling methods. These ways forward will allow us to paint a more complete picture of galaxy formation over cosmic time.

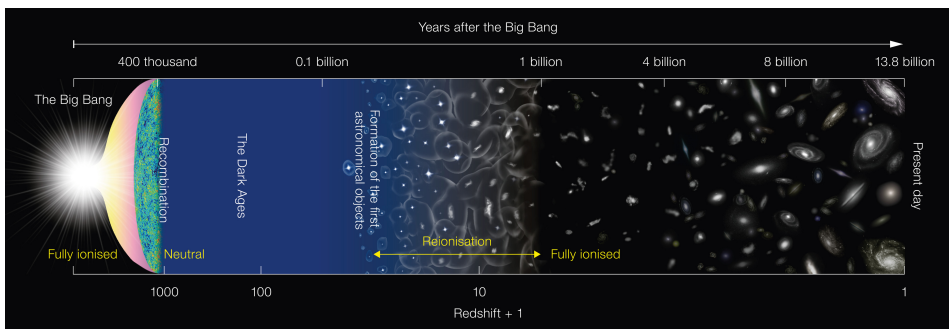
NEDERLANDSE SAMENVATTING

Om de vorming en evolutie van structuur in ons heelal te begrijpen, moeten we eerst de vraag stellen hoe sterrenstelsels en hun omringende halo's van donkere materie — de bouwstenen van het heelal — in de loop van de tijd zijn gegroeid en veranderd. Zware sterrenstelsels met oude sterpopulaties bieden de antwoorden, aangezien zij tot de meest extreme systemen behoren die bestaan. Deze oude kosmische fossielen, die zijn gestopt met stervorming, bevatten het grootste deel van de stellaire massa in het heelal. Hun eigenschappen dagen theorieën over kosmische structuurvorming uit en maken ze tot unieke laboratoria om de fundamentele fysica te onderzoeken waarop ons heelal berust.

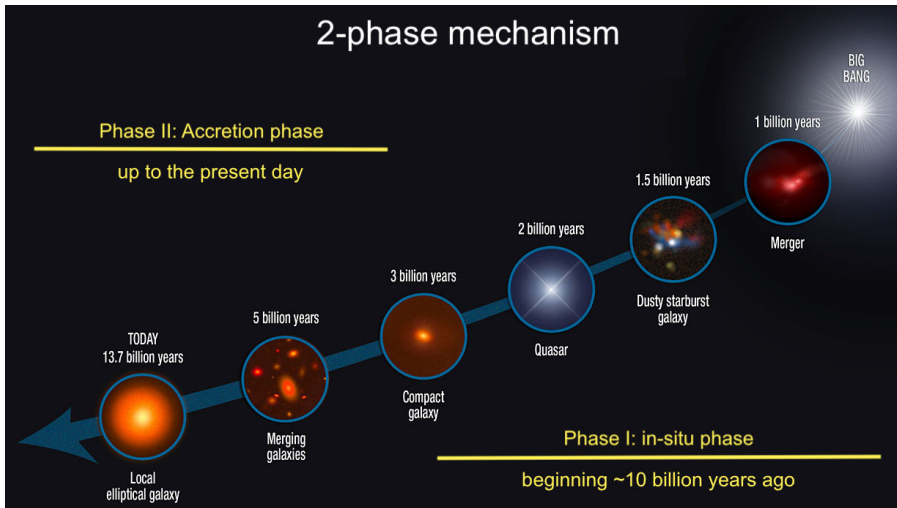
Het dossier van zware sterrenstelselvorming

Het heelal begon ongeveer 13,8 miljard jaar geleden in een extreem hete en dichte toestand, met de oerknal (zie Figuur 1). In een oogwenk begon de ruimte zelf uit te zetten en af te koelen, waarbij zich in de loop van de tijd steeds grotere structuren vormden: van deeltjes naar atomen, naar sterren en sterrenstelsels. Op deze manier vormt structuur in het heelal zich hiërarchisch. Dit model voor de vorming van het heelal wordt de Λ CDM-kosmologie genoemd.

Sterrenstelsels zijn enorme verzamelingen van sterren, gas, stof en donkere materie die door zwaartekracht bij elkaar worden gehouden. Ze groeien via invallend gas, door nieuwe sterren te vormen uit koud gas, en door samen te smelten



Figuur 1: Een diagram dat de geschiedenis van het heelal weergeeft, beginnend met de oerknal. Afbeelding overgenomen van ESO/NAOJ.



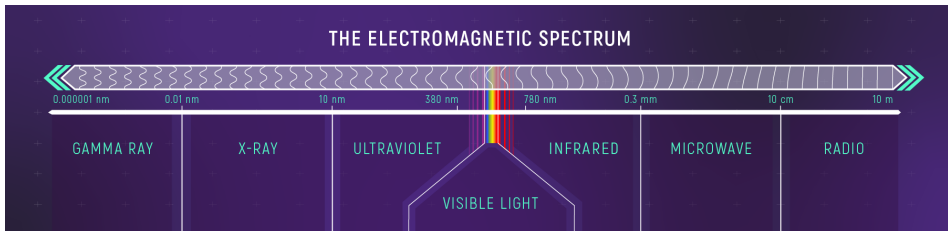
Figuur 2: Een diagram dat het tweefasenmodel van zware sterrenstelselvorming illustreert. Figuur aangepast van NASA/ESA/Sune Toft.

met andere, nabije sterrenstelsels. Uiteindelijk stoppen veel sterrenstelsels met stervorming, een proces dat bekend staat als ‘quenching’. In ons huidige beeld van sterrenstelselvorming hebben sterrenstelsels zich op een hiërarchische manier gevormd, vergelijkbaar met de grotere kosmische structuur van het heelal.

De vorming van zware sterrenstelsels gebeurde waarschijnlijk in twee fasen (zie Figuur 2): hun dichte, compacte centra zijn waarschijnlijk vroeg gevormd, ongeveer 10 miljard jaar geleden, tijdens een periode van snelle stervorming. Nadat de stervorming stopte, zijn ze mogelijk blijven groeien. Dat zou vooral aan de buitenkant gebeurd zijn door botsingen met kleine naburige sterrenstelsels. Dit beeld wordt ondersteund door waarnemingen die laten zien dat verre, passieve sterrenstelsels compacter zijn dan nabije sterrenstelsels.

Er zijn echter ook alternatieve scenario’s mogelijk. Zo kan het zijn dat individuele sterrenstelsels helemaal niet groeien. In plaats daarvan zou de evolutie in afmeting het gevolg kunnen zijn van het feit dat sterrenstelsels die vroeger stoppen met het vormen van sterren kleiner zijn dan stelsel die later stoppen. Bovendien is het nog onduidelijk hoe en waarom sterrenstelsels stoppen met stervorming. Er zijn verschillende processen die dit kunnen veroorzaken, bijvoorbeeld door het koude gas — dat nodig is voor stervorming — te verwijderen, op te gebruiken of te verhitten. In werkelijkheid spelen waarschijnlijk meerdere processen tegelijk een rol. Toch is het cruciaal om de mechanismen te bepalen waarmee het merendeel van de sterrenstelsels zich vormt, ‘quencht’ en vervolgens evolueert, om zo ons bredere begrip van structuurvorming in het heelal te verbeteren.

Om de vorming van zware sterrenstelsels te begrijpen, kunnen we gebruikmaken van geavanceerde observatoria zoals de James Webb Space Telescope (*JWST*). In 2021 begon *JWST* het heelal waar te nemen, na een van de meest uitdagende lanceringen ooit. De *JWST* vormt een ongeëvenaarde internationale samenwerking



Figuur 3: Een illustratie van het elektromagnetische spectrum van licht. We kunnen dit licht opsplitsen in brede kleuren (fotometrie) of in individuele golflengten (spectroscopie). Afbeelding overgenomen van NASA/ESA/Leah Hustak (STScI).

tussen de VS, Europa en Canada. Met een kostprijs van \$10 miljard en operationele kosten van \$10.000 per uur heeft deze investering de wetenschap al ingrijpend veranderd. *JWST* heeft met name de ontdekking mogelijk gemaakt van zware sterrenstelsels met oude sterpopulaties op veel vroegere tijdstippen dan verwacht, toen het heelal ongeveer 2 miljard jaar oud was. Deze resultaten suggereren dat sterrenstelsels veel sneller en eerder met stervorming begonnen en stopten dan voorspeld door de Λ CDM-kosmologie, wat problemen veroorzaakt voor theorieën over sterrenstelselvorming en ons kosmologisch raamwerk.

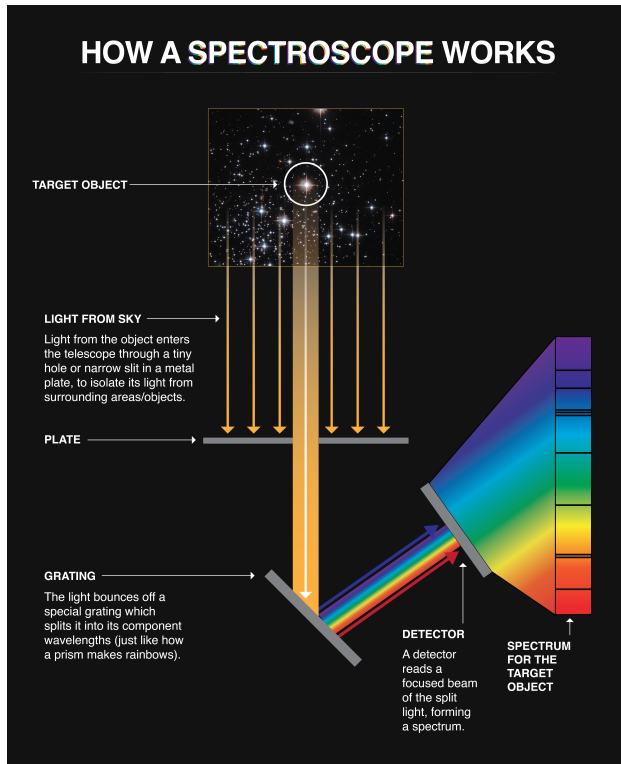
Deze interpretatie is echter gebaseerd op beperkte data en modellen. Dit werk is met name gebaseerd op het fitten van sterpopulatiemodellen aan een klein aantal fotometrische datapunten. Deze modellen kunnen vertekende eigenschappen voorspellen door brede aannames over de onderliggende chemische samenstelling en fysica van sterrenstelsels. Bovendien zijn onze data beperkt: ze tonen slechts ongeveer 2% van de sterren (de helderste en meest zware) die we in deze zeer verre sterrenstelsels kunnen waarnemen. Ten slotte gaat het om kleine steekproeven. Hierdoor kunnen belangrijke conclusies over sterrenstelselvorming onjuist zijn. In het tijdperk van *JWST* schieten traditionele methoden tekort.

Onze speurderskit voor het ontrafelen van de geschiedenis van sterrenstelsels

De componenten van een sterrenstelsel, inclusief de eigenschappen van zijn sterpopulaties, geven aanwijzingen over zijn evolutie. Deze informatie ligt opgeslagen in de chemische samenstelling van de sterren. Hoewel we de volledige geschiedenis van een sterrenstelsel nooit direct kunnen observeren, kunnen we het verleden indirect bestuderen via het licht van de sterren.

Door de totale helderheid van een sterrenstelsel in een klein aantal brede kleuren te meten, krijgen we een eerste indruk van de sterinhoud. Door het licht verder op te splitsen in individuele golflengten, kunnen we de opbouwgeschiedenis van sterrenstelsels ontleden (zie Figuren 3 en 4).

Om deze data te interpreteren, moeten we begrijpen hoe sterpopulaties zijn opgebouwd. Elke ster wordt geboren met een bepaalde massa, en we moeten de



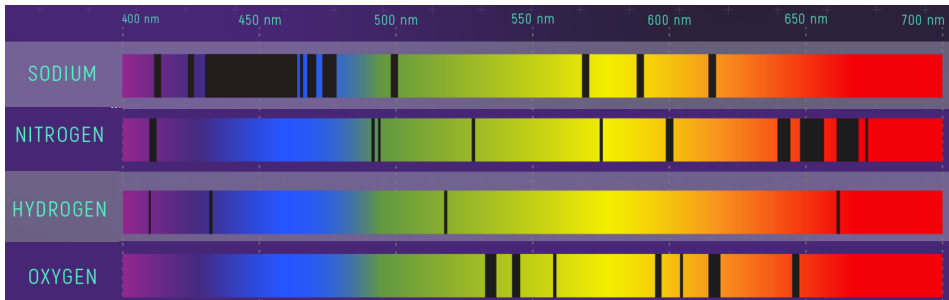
Figuur 4: Een diagram dat laat zien hoe we spectra van sterrenstelsels verkrijgen met een spectrograaf op een telescoop. Afbeelding overgenomen van NASA Goddard/Shireen Dooling.

verdeling van deze massa's kennen (de initiële massafunctie) om sterpopulaties te beschrijven. Deze kan variëren met omgeving, maar dit is nog een belangrijk open vraagstuk. Daarnaast moeten we begrijpen hoe sterren met verschillende massa's en chemische samenstellingen evolueren. Ook moeten we weten welke spectra verschillende typen sterren produceren, omdat zij zowel de vorm van het spectrum als de aanwezigheid van spectrale lijnen beïnvloeden.

Door atomen in het laboratorium te bestuderen, hebben wetenschappers berekend hoe elk element licht absorbeert op specifieke golflengten — elk element heeft zijn eigen spectrale vingerafdruk. Zo absorbeert waterstof licht op bepaalde golflengten (zie Figuur 5), waardoor we bij het zien van deze lijnen in een spectrum weten dat waterstof aanwezig is.

Door deze ingrediënten te begrijpen, kunnen we leeftijden en gedetailleerde chemie afleiden, wat inzicht geeft in de evolutie van sterrenstelsels. Verschillende typen sterren produceren verschillende elementen op uiteenlopende tijdschalen, waardoor elementverhoudingen informatie geven over chemische evolutie en sterformingsgeschiedenis.

Het interpreteren van spectra is echter complex, omdat verschillende eigen-



Figuur 5: Absorptiespectra voor verschillende elementen. Afbeelding aangepast van NASA/ESA/Leah Hustak (STScI).

schappen vergelijkbare effecten kunnen hebben. Zo kunnen leeftijd en chemische samenstelling bijna identieke spectrale veranderingen veroorzaken (de leeftijd-metalliciteit degeneratie). Ook variaties in de initiële massafunctie kunnen lijken op chemische variaties. Het correct ontleden van deze effecten is daarom bijzonder uitdagend.

We kunnen deze fysische principes toepassen op echte data door modellen te fitten aan spectra. Individuele spectrale kenmerken zijn gevoelig voor specifieke eigenschappen: waterstoflijnen voor leeftijd en ijzerlijnen voor zowel leeftijd als het ijzergehalte.

In de praktijk dragen echter alle eigenschappen bij aan het volledige spectrum. Door het hele spectrum tegelijk te modelleren, verkrijgen we meer informatie en robuustere resultaten, inclusief subtiele bepalingen van de initiële massafunctie.

Op zoek naar de vingerafdrukken van sterrenstelselvorming

Met deze technieken hebben we geleerd dat zware sterrenstelsels in het nabije heelal ook tot de oudste objecten behoren en rijk zijn aan elementen zoals ijzer en magnesium. Dit suggereert dat zij hun sterren snel vormden en vroeg hun stervorming eindigden. Daarnaast hebben nabijgelegen zware sterrenstelsels in hun centra een overschot aan lage-massa sterren ten opzichte van de Melkweg, wat wijst op een andere initiële massafunctie.

Door ook de ruimtelijke verdeling van sterpopulaties mee te nemen, verkrijgen we meer detail. Gradiënten in sterpopulaties geven informatie over massa-opbouw en helpen verschillende vormingsmechanismen te onderscheiden. Nabije zware sterrenstelsels hebben vaak rodere centra en blauwere buitengebieden, veroorzaakt door hogere ijzergehaltes in de centra — in lijn met hiërarchische vorming.

Om de vorming van zware oude sterrenstelsels echt te begrijpen, moeten we echter terugkijken in de tijd en hun jongere fasen bestuderen. Door de eindige lichtsnelheid zien we verre sterrenstelsels zoals ze vroeger waren. Dit wordt gekwantificeerd met roodverschuiving: hoe hoger de roodverschuiving, hoe vroeger

de kosmische tijd.

Het observeren van deze verre sterrenstelsels met hoge roodverschuiving is uitdagender dan het bestuderen van nabije sterrenstelsels, omdat ze door hun enorme afstand veel zwakker zijn. Bovendien zijn hun elementgevoelige absorptiekenmerken zeer zwak en verschoven naar golflengtegebieden waarin onze telescopen minder gevoelig zijn. Daarom moeten we individuele verre sterrenstelsels tientallen tot honderden uren observeren om gedetailleerde leeftijden en chemische samenstellingen te meten.

Dankzij innovatieve technologische vooruitgang in de astronomische waarnemingen in de afgelopen ~ 15 jaar zijn we begonnen zware sterrenstelsels met oude sterpopulaties tot zeer vroege tijden te onderzoeken. Verrassend genoeg wordt ons beeld van sterrenstelselvorming juist minder duidelijk naarmate we dichter bij het begin van het heelal komen. Het is bijvoorbeeld onduidelijk hoe deze oude sterrenstelsels evolueren tot de sterrenstelsels die we in het nabije heelal zien (als ze al verwant zijn). Zoals hierboven besproken heeft *JWST* de ontdekking mogelijk gemaakt van zware sterrenstelsels met oude sterpopulaties op veel vroegere tijden dan voorspeld door Λ CDM-simulaties. De vorming en evolutie van zware sterrenstelsels met oude sterpopulaties blijft daarom een open vraag. Om vooruitgang te boeken in het identificeren van de dominante mechanismen van ster-massa-opbouw en het stoppen van stervorming, en om de vondsten van “onmogelijk vroege” sterrenstelsels te bevestigen, hebben we een groter aantal nodig van verre, zware sterrenstelsels met diepe en hoge ruimtelijke resolutie waarnemingen.

Dit proefschrift: het identificeren van de ouders

In dit proefschrift presenteren we enkele van de meest gedetailleerde metingen tot nu toe van sterrenstelselvorming tot ~ 10 miljard jaar geleden. Met de flexibele modellering van hoogwaardige geïntegreerde en ruimtelijk opgeloste spectra presenteren we nieuwe metingen van leeftijds- en chemische gradiënten in verre sterrenstelsels, onthullen we discrepanties in modelvoorspellingen, en voeren we de eerste robuuste metingen uit van de initiële massafunctie in het vroege heelal.

In **Hoofdstuk 2** meten we gradiënten in een grote steekproef van verre zware sterrenstelsels met oude sterpopulaties met data van de LEGA-C survey. Deze sterrenstelsels bestonden ~ 7 miljard jaar geleden, toen het heelal ongeveer half zo oud was als nu. We vinden dat de sterren in de centra dezelfde leeftijd en hetzelfde magnesiumgehalte hebben als in de buitengebieden, maar hogere ijzergehaltes. Onze resultaten zijn consistent met een scenario waarin deze sterrenstelsels eerst intense stervorming in hun centra ondergingen voordat zij hun stervorming eindigden, waarna hun buitengebieden werden opgebouwd door fusies met kleine, naburige sterrenstelsels met lagere ijzergehaltes. Alternatieve scenario's kunnen onze resultaten echter ook verklaren.

In **Hoofdstuk 3** combineren we onze gedetailleerde spectroscopische metingen met onafhankelijke fotometrische kleuren om voorspellingen van sterpopulatiemodellen te testen. We meten leeftijden en chemische samenstellingen van een grote steekproef van zware sterrenstelsels met oude sterpopulaties, opnieuw afkomstig

uit de LEGA-C survey. We vergelijken de gemeten kleuren van onze sterrenstelsels met de voorspelde kleuren op basis van onze gemeten leeftijden en chemische samenstellingen. We vinden dat sterpopulatiemodellen de leeftijd-kleur- en leeftijd-chemische-samenstelling-relaties die we in de data vinden niet kunnen reproduceren. Onze resultaten tonen aan dat veelgebruikte sterpopulatiemodellen onjuist zijn en laten zien dat de onderliggende fysische aannames waarop deze modellen zijn gebaseerd dringend opnieuw moeten worden geëvalueerd.

In **Hoofdstuk 4** meten we leeftijds- en chemische gradiënten in 8 zware sterrenstelsels met oude sterpopulaties die ~ 10 miljard jaar geleden bestonden, dicht bij het tijdperk waarin deze sterrenstelsels gevormd werden en hun stervorming eindigden. Met behulp van data van het *JWST*-SUSPENSE-programma vinden we dat de sterren in de centra van deze sterrenstelsels ouder zijn en rijker aan magnesium dan in de buitengebieden, maar dezelfde hoeveelheid ijzer bevatten. Deze resultaten suggereren een scenario waarin de centra van sterrenstelsels eerst stopten met stervorming, gevolgd door de buitengebieden. In de richting van het huidige tijdperk zouden deze sterrenstelsels hun buitengebieden kunnen hebben opgebouwd door fusies met kleine, naburige sterrenstelsels, wat heeft geleid tot de gradiënten die we in Hoofdstuk 2 hebben gevonden. Aan de andere kant is het mogelijk dat verschillende quenching-mechanismen op verschillende tijdstippen optreden en leiden tot verschillende gradiënten.

Ten slotte presenteren we in **Hoofdstuk 5** de eerste robuuste metingen van de initiële massafunctie, in een steekproef van 9 zware sterrenstelsels met oude sterpopulaties buiten het nabije heelal. Deze studie is gebaseerd op data van ons *JWST*-IMFERNO-programma. We vinden dat de meest zware, verre sterrenstelsels een overschot aan lage-massa sterren hebben ten opzichte van de Melkweg. In combinatie met eerdere bevindingen dat de centra van nabije zware sterrenstelsels met oude sterpopulaties ook een overschot aan lage-massa sterren bevatten, zijn onze resultaten consistent met het tweefasenmodel. Ons oudste en meest zware sterrenstelsel is waarschijnlijk een afstammeling van de verre, zware, “onmogelijk vroege” sterrenstelsels die met *JWST* zijn ontdekt, wat impliceert dat zij mogelijk ook dit overschot aan lage-massa sterren hadden. Onze beperkingen suggereren dat hun stellaire massa’s nog groter zijn dan oorspronkelijk gerapporteerd, wat de spanning tussen waarnemingen van zware, vroege sterrenstelsels en huidige modellen voor sterrenstelselvorming verder vergroot.

Het dossier gaat verder...

Hoewel dit werk ons begrip van sterrenstelselvorming in de afgelopen ~ 10 miljard jaar aanzienlijk heeft verbeterd, blijven er belangrijke vragen bestaan. Om de dominante vormingsroutes te identificeren, moeten we grotere steekproeven bestuderen met nog hogere ruimtelijke resolutie, gebruikmakend van de volgende generatie telescopen (zoals de *Extremely Large Telescope*), en tegelijkertijd geavanceerdere modelleringstechnieken ontwikkelen. Deze stappen zullen ons in staat stellen een completer beeld te schetsen van sterrenstelselvorming door de kosmische tijd heen.

PUBLICATIONS

First Author

* = Part of This Thesis

1. *Hidden mass in early galaxies revealed by bottom-heavy initial mass functions**
Cheng, C. M., Slob, M., Kriek, M., Beverage, A. G., van Dokkum, P. G., Bezanson, R., Brammer, G., Conroy, C., de Graaff, A., Eftekhari, E., Feldmann, R., Goesaert, W. M., Gu, M., Leja, J., Lorenz, B., Mancera Piña, P. E., Martín-Navarro, I., Newman, A. B., Price, S. H., Shapley, A. E., Sharda, P., Suess, K. A., van der Wel, A., Weisz, D. R. 2026, *Nature Astronomy*, under review.
2. *Clues to inside-out quenching in quiescent galaxies at $1.2 \lesssim z \lesssim 2.2$: Age, Fe-, and Mg-abundance gradients from JWST-SUSPENSE**
Cheng, C. M., Slob, M., Kriek, M., Beverage, A. G., Barro, G., Bezanson, R., de Graaff, A., Förster Schreiber, N. M., Lorenz, B., Marchesini, D., Martín-Navarro, I., Muzzin, A., Newman, A. B., Price, S. H., Suess, K. A., van der Wel, A., van de Sande, J., van Dokkum, P. G., & Weisz, D. R. 2026, *A&A*, in press. DOI: 10.1051/0004-6361/202557254.
3. *Ages and metallicities of quiescent galaxies: confronting broad-band (UVJ) colours with stellar absorption lines**
Cheng, C. M., Kriek, M., Beverage, A. G., Slob, M., Bezanson, R., Franx, M., Leja, J., Mancera Piña, P. E., Suess, K. A., van der Wel, A., van de Sande, J., & van Dokkum, P. G. 2025, *MNRAS*, 540, 1527. DOI: 10.1093/mnras/staf806.
4. *Age and metal gradients in massive quiescent galaxies at $0.6 \lesssim z \lesssim 1.0$: implications for quenching and assembly histories**
Cheng, C. M., Kriek, M., Beverage, A. G., van der Wel, A., Bezanson, R., D'Eugenio, F., Franx, M., Mancera Piña, P. E., Nersesian, A., Slob, M., Suess, K. A., van Dokkum, P. G., Wu, P.-F., Gallazzi, A., & Zibetti, S. 2024, *MNRAS*, 532, 3604. DOI: 10.1093/mnras/stae1739.
5. *Initial mass function variability from the integrated light of diverse stellar systems.*
Cheng, C. M., Villaume, A., Balogh, M. L., Brodie, J. P., Martín-Navarro,

I., Romanowsky, A. J., & van Dokkum P. G. 2023, MNRAS, 526, 4004. DOI: 10.1093/mnras/stad2967.

6. *Testing the chemical homogeneity of chemically tagged dissolved birth clusters*

Cheng, C. M., Price-Jones, N., & Bovy, J. 2021, MNRAS, 506, 5573. DOI: 10.1093/mnras/stab2106.

Contributing Author

1. *Fast Rotators at Cosmic Noon: Stellar Kinematics for 15 Quiescent Galaxies from JWST-SUSPENSE*

Slob, M., Kriek, M., de Graaff, A., **Cheng, C. M.**, Beverage, A. G., Bezanson, R., Förster Schreiber, N. M., Lorenz, B., Mancera Piña, P. E., Marchesini, D., Muzzin, A., Newman, A. B., Price, S. H., Suess, K. A., van de Sande, J., van Dokkum, P., Weisz, D. R. 2025, A&A, 702, A110. DOI: 10.1051/0004-6361/202555812.

2. *Carbon and Iron Deficiencies in Quiescent Galaxies at $z = 13$ from JWST-SUSPENSE: Implications for the Formation Histories of Massive Galaxies*

Beverage, A. G., Slob, M., Kriek, M., Conroy, C., Barro, G., Bezanson, R., Brammer, G., **Cheng, C. M.**, de Graaff, A., Förster Schreiber, N. M., Franx, M., Lorenz, B., Mancera Piña, P. E., Marchesini, D., Muzzin, A., Newman, A. B., Price, S. H., Shapley, A. E., Stefanon, M., Suess, K. A., van Dokkum, P., Weinberg, D., Weisz, D. R. 2024, ApJ, 979, 249. DOI: 10.3847/1538-4357/ad96b6.

3. *The JWST-SUSPENSE Ultradeep Spectroscopic Program: Survey Overview and Star-Formation Histories of Quiescent Galaxies at $1 < z < 3$*

Slob, M., Kriek, M., Beverage, A. G., Suess, K. A., Barro, G., Bezanson, R., Brammer, G., **Cheng, C. M.**, Conroy, C., de Graaff, A., Förster Schreiber, N. M., Franx, M., Lorenz, B., Mancera Piña, P. E., Marchesini, D., Muzzin, A., Newman, A. B., Price, S. H., Shapley, A. E., Stefanon, M., van Dokkum, P., Weisz, D. R. 2024, ApJ, 973, 131. DOI: 10.3847/1538-4357/ad65ff.

4. *Low-density star cluster formation: discovery of a young faint fuzzy on the outskirts of the low-mass spiral galaxy NGC 247*

Romanowsky, A. J., Larsen, S. S., Villaume, A., Carlin, J. L., Janz, J., Sand, D. J., Strader, J., Brodie, J. P., Chakrabarti, S., **Cheng, C. M.**, Crnojević, D., Forbes, D. A., Garling, C. T., Hargis, J. R., Karunakaran, A., Martín-Navarro, I., Olsen, K. A. G., Rider, N., Salimkumar, B., Santhanakrishnan, V., Spekkens, K., Tang, Y., van Dokkum, P. G., Willman, B. 2023, MNRAS, 518, 3164. DOI: 10.1093/mnras/stac2898.

Non-Astronomy

1. *First Evidence of Axial Shape Asymmetry and Configuration Coexistence in*

⁷⁴Zn: *Suggestion for a Northern Extension of the N = 40 Island of Inversion*

Rocchini, M., Garrett, P. E., Zielińska, M., Lenzi, S. M., Dao, D. D., Nowacki, F., Bildstein, V., MacLean, A. D., Olaizola, B., Ahmed, Z. T., Andreoiu, C., Babu, A., Ball, G. C., Bhattacharjee, S. S., Bidaman, H., **Cheng, C.**, Coleman, R., Dillmann, I., Garnsworthy, A. B., Gillespie, S., Griffin, C. J., Grinyer, G. F., Hackman, G., Hanley, M., Illana, A., Jones, S., Laffoley, A. T., Leach, K. G., Lubna, R. S., McAfee, J., Natzke, C., Pannu, S., Paxman, C., Porzio, C., Radich, A. J., Rajabali, M. M., Sarazin, F., Schwarz, K., Shadrack, S., Sharma, S., Suh, J., Svensson, C. E., Yates, D., Zidar, T. 2023, Phys. Rev. Lett., 130, 122502. DOI: 10.1103/PhysRevLett.130.122502.

2. *EEG features of spontaneous recurrent seizures in a mouse model of extended hippocampal kindling*

Liu, H., Hameed, A. Z., Chow, J., Sivanenthiran, N., **Cheng, C.**, Liu, Y., Cheung, P., Lim, S., Jin, Y., Lin, Y. Q., Mao, M., Wu, C., Carlen, P. H., Eubanks, J. H., Song, H., Zhang, L. 2021, Clinph, 132(9), e2, DOI: 10.1016/j.clinph.2021.03.028.

3. *Electrographic Features of Spontaneous Recurrent Seizures in a Mouse Model of Extended Hippocampal Kindling*

Liu, H., Tufa, U., Zahra, A., Chow, J., Sivanenthiran, N., **Cheng, C.**, Liu, Y., Cheung, P., Lim, S., Jin, Y., Mao, M., Sun, Y., Wu, C., Wennberg, R., Bardakjian, B., Carlen, P. L., Eubanks, J. H., Song, H., Zhang, L. 2021, TexCom, 2(1), DOI: 10.1093/texcom/tgab004.

4. *High-precision branching ratio measurement and spin assignment implications for ⁶²Ga superallowed β decay*

MacLean, A. D., Laffoley, A. T., Svensson, C. E., Ball, G. C., Leslie, J. R., Andreoiu, C., Babu, A., Bhattacharjee, S. S., Bidaman, H., Bildstein, V., Burbadge, C., Bowry, M., **Cheng, C.**, Cross, D. S., Diaz-Varela, A., Dillmann, I., Dunlop, M. R., Dunlop, R., Evitts, L. J., Finlay, P., Gillespie, S., Garnsworthy, A. B., Garrett, P. E., Gopaul, E., Griffin, C. J., Grinyer, G. F., Hackman, G., Henderson, J., Jigmeddorj, B., Leach, K. G., Kasanda, E., McAfee, J., Moukaddam, M., Natzke, C., Nittala, S., Olaizola, B., Park, J., Paxman, C., Pore, J. L., Porzio, C., Radich, A. J., Ruot-salainen, P., Saito, Y., Sharma, S., Smallcombe, J., Smith, J. K., Sultana, R., Turko, J., Williams, J., Yates, D., Zidar, T. 2020, Phys Rev C, 102(5), DOI: 10.1103/physrevc.102.054325.

5. *Impaired Spatial Learning and Memory in Middle-Aged Mice with Kindling-Induced Spontaneous Recurrent Seizures*

Liu, H., Stover, K. R., Sivanenthiran, N., Chow, J., **Cheng, C.**, Liu, Y., Lim, S., Wu, C., Weaver, D. F., Eubanks, J. H., Song, H., Zhang, L. 2019, Front. Pharmacol., 10, 1077, DOI: 10.3389/fphar.2019.01077.

6. *Effects of Antiepileptic Drugs on Spontaneous Recurrent Seizures in a Novel*

Model of Extended Hippocampal Kindling in Mice

Song, H., Tufa, U., Chow, J., Sivanenthiran, N., **Cheng, C.**, Lim, S., Wu, C., Feng, J., Eubanks, J. H., Zhang, L. 2018, *Front. Pharmacol.*, 9, 451, DOI:

10.3389/fphar.2018.00451.

CURRICULUM VITAE

On February 12, 1998, I was born in Toronto, Ontario, Canada. Three months later, my parents, Joan and Sydney, moved us to Arcata, California, where we lived for three years, and where my brother, Ethan, was born. The US ultimately did not live up to my parents' expectations. So, in 2001 we moved back to Canada, settling in Markham, Ontario, where my youngest brother, Jordan, was born a few years later.

As a child, I was extremely fortunate in many ways, but especially in that my parents encouraged me to try everything. I dabbled in a variety of sports¹ – from soccer, to snowboarding, to figure skating – in several musical instruments – including piano, flute, violin, and classical singing – and a multitude of creative hobbies – such as knitting, crocheting, and sewing. Additionally, as physicians, my parents fostered in me a love for learning, and a need to seek answers to unsolved problems.

I attended Roy H. Crosby and Unionville Meadows Public Schools, where I continued to develop my interests in athletics, the arts, and science. At this point, I decided that my primary passion was for music, so I auditioned for an arts program at Unionville High School. I was successful, and began my secondary school career as a classical voice major. These four years allowed me to hone my musicianship and performance skills. Simultaneously, however, the majority of my peers were interested in science, and this influenced me to continue to take science courses. I quickly realized that science was the optimal way to sate my need to solve problems, so at around age 15, I decided that I wanted to be a doctor, like my parents. When I was in Grade 11, I took my first physics course, which opened my eyes to an analytical, quantitative way of understanding the world that really tickled my brain (this was ironic as I had always hated math). Around this time, NASA and ESA released the sharpest ever view of the Andromeda Galaxy, taken with the Hubble Space Telescope². Until this point, I had had zero interest in space, but I found this technological feat that gave us the ability to see millions of tiny stars from such a great distance truly inspiring. While I was still keen on studying medicine, I had been bitten by the astronomy bug.

When it came time to start university, I applied for undergraduate programs widely, keeping doors open for music, medicine, and physics. I decided to study science at the University of Toronto as a member of Trinity College, where I completed an Astronomy and Physics Specialist and Mathematics Minor. For the first

¹Other prominent athletic endeavours included: skiing, swimming, horseback riding, gymnastics, taekwondo, and running.

²<https://esahubble.org/images/heic1502a/> .

few years of my Bachelor's, I was still interested in medicine. As a result, I worked in a fundamental neurobiology lab for two summers at the Toronto Western Hospital, supervised by Dr. Liang Zhang. This was a unique experience, where I had the opportunity to participate in hands-on mouse experiments to study epilepsy. In the summer of my third year of my undergrad, however, I decided it was time to commit to physics. I did a summer internship at TRIUMF, a particle accelerator in Vancouver, British Columbia, where, supervised by Dr. Adam B. Garnsworthy, I worked on nuclear physics experiments. In the fourth year of my undergrad, I finally stopped beating around the bush. I completed a research thesis project in galaxy evolution, where I worked with Professor Jo Bovy and was primarily supervised by Dr. Natalie Price-Jones, his senior grad student at the time. I had the opportunity to assess chemical homogeneity in chemically tagged birth clusters in the Milky Way. Through this experience, I learned that I had a keen interest in spectroscopy, galactic archaeology, and statistical and computational methods. With the support of my supervisors, I was able to publish my first astronomy paper. I received my Honors Bachelor of Science with Distinction from the University of Toronto in 2020. I also kept up my other interests during this time – I played quidditch for the University of Toronto Centaurs and continued to participate in classical singing competitions.

Towards the end of my Bachelor's I began to apply for direct-entry PhD programs, as my goal was to stay at the University of Toronto, where there was no Master's program. Instead, I was accepted for a Master's program at the University of Waterloo, where I was supervised by Professor Michael Balogh. With Michael's support, I worked on measuring the stellar initial mass function (IMF) in compact stellar systems, and through this research I became an expert on stellar population modeling and the IMF, resulting in my second publication. I received my Master of Science in 2022.

This period was also challenging for me, in that I completed my entire Master's degree remotely due to the COVID-19 pandemic, and that I experienced some difficult interpersonal conflicts. I seriously considered leaving the field. However, I realized that I thoroughly enjoyed the work that I was doing, but needed to be around the right people. With this in mind, I decided to apply for PhD programs, but executed a strategy of requesting Zoom meetings with potential supervisors before even applying. I was fortunate in that most people agreed. One of these people was Professor Mariska Kriek, who thankfully accepted my invitation to meet (something that she later told me she never does). I had thought that Mariska was at the University of California Berkeley, but in reality she had just moved to Leiden. I had never heard of Leiden, and did not know much about The Netherlands, but I admired Mariska's style and work, so I decided to apply for her position. In 2022, I was accepted to the PhD program at Leiden Observatory to work with Mariska. Accepting her offer was the best decision I could have made. I have transitioned from studying Milky Way stars in my Bachelor's, to nearby stellar systems in my Master's, to distant, massive, quiescent galaxies in my PhD, but have continued to apply the same principles of examining the fossil records of galaxies with detailed modeling throughout.

In addition to building up my scientific profile, I have gained a lot of other in-

credible experiences during my PhD. I had the rewarding opportunity to supervise two Master's students at Leiden. I have also had the chance to travel and present my work widely, nationally and internationally, at conferences in Malaysia, Spain, Italy, Switzerland, Ireland, and Japan. It has additionally been important to me to contribute to the Observatory, and I have participated in the EDI, Borrel, Social, and PhSki Committees. Finally, I have been involved in observing programs on major telescopes, and have even had the opportunity to act as Co-PI and PI on two JWST programs, respectively.

I am very thankful that my academic journey will not end here. While I have experienced a tough job application cycle, I am happy to say that I will be starting a postdoctoral position with Professor Allison Man at the University of British Columbia in Vancouver, Canada. I am excited to move back to my home country, and to continue to advance the expertise that I have developed during my PhD while expanding my experience in galaxy evolution research.

ACKNOWLEDGMENTS

To quote Fall Out Boy, Thanks fr th Mmrs.

First, I want to extend my heartfelt thanks to the Observatory support staff, especially to Nancy, Alina, Robin, Bram, Huib, David, and Erik, for your invaluable help with borrels, administrative issues, and computing questions.

To my supervisor, Mariska, for your guidance, mentorship, and faith in me. Every moment of your time, scientific or otherwise, has been a privilege. Thank you for taking a chance on a random kid from Canada. Thank you also to Arjen, for your support, and Pieter, for your investment in my science and well-being from even before my PhD. Thank you to my past supervisors, Dr. Zhang, Adam, Jo and Michael, for setting me up well for this degree.

Thank you to my academic pals. Aliza, thank you for your advice and emotional support. I am so grateful that I got to work closely with you, even though we were usually 8900 km (that's 5500 miles) apart. Martje, I agonized over whether to make this unserious or heartfelt, and unfortunately for you I chose the latter. In you I have found not only a colleague, but also a dear friend and even a sister. Thank you for everything. Wren, thank you for your scientific insight, and for the crazy treks. To Brian, Jesse, Wout, and Elham – I have truly valued working with you. To my students, Ying and Andreea, thank you for your hard work and trust.

In Leiden, I have been extremely lucky to find some incredible friends. Lucie, thanks for being my partner in crime and for the fabulous date nights. Ivana, thank you for always listening, and for making me re-appreciate my Tumblr era. To Amy, twin, thank you for trusting me to be your co-pilot. Verónica, thank you for your take-charge attitude, may we never go to Lowlands. To Marta, thank you for your kindness and for carrying me through our scuba course. Anniek, thank you for your advice, and to you and Tom, thank you for trusting us with your cat. And to Koetje, I couldn't have done it without your fluffiness and cuddles.

Special shout out to the Pedro Pancakes peeps for all of the media marathons. Ciarán, buddy, thank you for all of the great chats and advice, scientific or otherwise. Lekker biertjes on me next time. To Sill, Josh, Nicco, Nicole, and Roi, I'll always treasure our in-depth GoT/LotR/etc discussions. Merci Thomas, for your friendship and the awesome parties, you are too cool for me.

A few more Observatory friends to thank: Christian, your cat reels have gotten me through some hard times. Lars, I hope that you'll finally be able to get some work done now. Jeger and Timo, coming up with creative borrel themes with you was the best. To Christiaan and Thijs, for being the best PhSki dads. Finally, thank you to Victorine, Piyush, Gijs, Sai, Markus, Tong, Alessia, Andrew, Jinyi,

Dario, Elia, Silvia, Aniruddh, Beth, Bianca, Brigitte, Jelle, Kevin, Osmar, Paola, Roland, Pavel, Jessie, Sebastian, Aaron, Corné, Sid, Erik, Scott, Zephyr, Kostas, Yuming, and anyone I may have missed. And of course, to Big Dipper – we laughed, we cried, we had a blast.

I also have to thank my fantastic support network at home. Natalie, thank you for your mentorship and friendship. I would not be where I am today without you. Youssef and Joe, I know you'll always be my teammates. Michelyn and Wriggs, thank you for teaching me discipline and performance, I use these skills every day. Mr. Dietrich, thank you for inspiring me to pursue this path. Andrew, thanks for kicking my butt.

To my family: I know these years living apart have been tough, so thank you for bearing with me. Mom, you are my role model today and every day. I want to be just like you when I grow up. Mada, thank you for your constant support and advice. To my brothers – Ethan, I truly value your logic and honesty, and Jordan, your creativity forever inspires me. Never forget that I'm the strong one, though. To Ah-ma and Ah-gong for being my number one cheerleaders, always. This PhD is for you.

And finally, to Colin, for your unconditional love and support over these past 11 years. Thank you for agreeing to go on this grand adventure with me, for your patience, and for growing together with me. I could not have made this crazy move or finished my PhD without you, and I am so grateful for all of the experiences that we've gotten to share. Who knew that those kids teaching swimming at Centennial would end up here? The future may look uncertain, but I know we'll be able to get through it together. I don't know how I got so lucky. I love you.