The Kinematics of Massive Quiescent Galaxies at $1.4 < z < 2.1$: Dark Matter Fractions, IMF Variation, and the Relation to Local Early-type Galaxies[*]

J. Trevor Mendel[1,2,3,4], Alessandra Beifiori[1,2], Roberto P. Saglia[1,2], Ralf Bender[1,2], Gabriel B. Brammer[5],
Jeffrey Chan[6], Natascha M. Förster Schreiber[2], Matteo Fossati[1,2,7], Audrey Galametz[8], Ivelina G. Momcheva[9],
Erica J. Nelson[2,10], David J. Wilman[1,2], and Stijn Wuyts[11]
[1] Universitäts-Sternwarte München, Scheinerstr. 1, D-81679 München, Germany; trevor.mendel@anu.edu.au
[2] Max-Planck-Institut für Extraterrestrische Physik, Giessenbachstr. 1, D-85748 Garching, Germany
[3] Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2611, Australia
[4] ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D)
[5] Cosmic Dawn Center, Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, DK-2100 Copenhagen, Denmark
[6] Department of Physics and Astronomy, University of California, Riverside, CA 92521, USA
[7] Dipartimento di Fisica G. Occhialini, Università degli Studi di Milano-Bicocca, Piazza della Scienza 3, I-20126 Milano, Italy
[8] Department of Astronomy, University of Geneva, 1205, Versoix, Switzerland
[9] Space Telescope Science Institute, Baltimore, MD 21218, USA
[10] Harvard-Smithsonian Center for Astrophysics, Cambridge, USA
[11] Department of Physics, University of Bath, Claverton Down,Bath, BA2 7AY, UK

Abstract

We study the dynamical properties of massive quiescent galaxies at $1.4 < z < 2.1$ using deep Hubble Space Telescope WFC3/F160W imaging and a combination of literature stellar velocity dispersion measurements and new near-infrared spectra obtained using the K-band Multi Object Spectrograph (KMOS) on the ESO Very Large Telescope. We use these data to show that the typical dynamical-to-stellar mass ratio has increased by ~ 0.2 dex from $z = 2$ to the present day, and we investigate this evolution in the context of possible changes in the stellar initial mass function (IMF) and/or fraction of dark matter contained within the galaxy effective radius, $f_{DM}[<r_e]$. Comparing our high-redshift sample to their likely descendants at low redshift, we find that $f_{DM}[<r_e]$ has increased by a factor of more than 4 since $z \approx 1.8$, from $f_{DM}[<r_e] = 6.6\% \pm 1.0\%$ to $\sim 24\%$. The observed increase appears robust to changes in the methods used to estimate dynamical masses or match progenitors and descendants. We quantify possible variation of the stellar IMF through the offset parameter α, defined as the ratio of dynamical mass in stars to the stellar mass estimated using a Chabrier IMF. We demonstrate that the correlation between stellar velocity dispersion and α reported among quiescent galaxies at low redshift is already in place at $z = 2$, and we argue that subsequent evolution through (mostly minor) merging should act to preserve this relation while contributing significantly to galaxies' overall growth in size and stellar mass.

Unified Astronomy Thesaurus concepts: Galaxy kinematics (602); Galaxy evolution (594); High-redshift galaxies (734)

Supporting material: extended figure

1. Introduction

Spectroscopic surveys of the high-redshift universe have shown that well-known scaling relations such as the fundamental and mass planes were already in place by at least $z = 2$ (e.g., Toft et al. 2012; Bezanson et al. 2013b; van de Sande et al. 2014; Beifiori et al. 2017; Prichard et al. 2017), despite the fact that individual galaxies appear to evolve significantly from the time they join the passive population to the present day. The most conspicuous signature of this evolution is seen in galaxy sizes, where massive quiescent galaxies at $z > 1$ are significantly smaller than their local counterparts at fixed stellar mass (e.g., Daddi et al. 2005; Trujillo et al. 2006; van Dokkum et al. 2008; Cimatti et al. 2012; van der Wel et al. 2014; Chan et al. 2016, 2018, but see also Carollo et al. 2013), but it is also apparent in measurements of galaxy stellar velocity dispersions and surface brightness profiles (e.g., Kriek et al. 2009; Cenarro & Trujillo 2009; van der Wel et al. 2011; Chang et al. 2013; van de Sande et al. 2013). However, the exact degree to which

individual galaxies change as they evolve is still unclear: although some amount of inferred evolution can be explained by a bias in the matching of progenitor and descendant populations (progenitor bias, e.g., van Dokkum & Franx 1996; Saglia et al. 2010; Valentinuzzi et al. 2010; Keating et al. 2015), some evolution is still required to reproduce properties of the full population (e.g., Belli et al. 2015).

Guided by the intrinsically hierarchical assembly of structure in ΛCDM models, the most attractive explanation for the continued structural evolution of quiescent galaxies is by gas-poor merging after the cessation of star formation. Both major (mass ratio $\mu_* > 0.25$) and minor ($\mu_* < 0.25$) mergers can significantly alter galaxy light profiles, leading to a disproportionate increase in (half-light/half-mass) size relative to stellar mass (e.g., Oser et al. 2010; Hilz et al. 2012, 2013), which seems all but demanded in the most compact, massive high-z galaxies (e.g., Damjanov et al. 2011). Detailed photometric and kinematic analyses of nearby passive galaxies appear to support the idea of a "two-phase" formation scenario characterized by early, rapid formation and subsequent assembly through repeated mergers (e.g., Arnold et al. 2011, 2014; de la Rosa et al. 2016; Foster et al. 2016). But while it appears that

mergers with $0.1 < \mu_* < 1$ can account for the evolution of galaxy sizes and velocity dispersions since $z \sim 1$, they have more difficulty explaining the dramatic increase in average sizes at earlier epochs (e.g., Newman et al. 2012), suggesting that other mechanisms such as stellar mass loss or feedback from active galactic nuclei (AGNs) may also play some role (e.g., Fan et al. 2008; Damjanov et al. 2009; Fan et al. 2010).

While different evolutionary scenarios predict different physical characteristics for the resulting galaxy population, the persistence of the fundamental plane, mass plane, and other scaling relations over time limits the parameter space available to models describing the evolution of galaxy properties. The existence of a fundamental plane for quiescent galaxies can be understood as a manifestation of the virial relation, where for relaxed systems the dynamical mass $M_{\rm dyn} \propto \sigma_*^2 r_{\rm e}$, with σ_* and $r_{\rm e}$ the stellar velocity dispersion and half-light size, respectively. Given measurements of σ_* and $r_{\rm e}$, the remaining unknown is the dynamical mass-to-light ratio, $M_{\rm dyn}/L$. Following Graves et al. (2009), $M_{\rm dyn}/L$ can be rewritten in terms of its underlying physical dependencies as

$$\frac{M_{\rm dyn}}{L} = \frac{M_{\rm dyn}}{M_{\rm tot}} \times \frac{M_{\rm tot}}{M_*} \times \frac{M_*}{M_{*,\rm IMF}} \times \frac{M_{*,\rm IMF}}{L}. \qquad (1)$$

The first and last terms, $M_{\rm dyn}/M_{\rm tot}$ and $M_{*,\rm IMF}/L$, depend on our ability to model certain galaxy properties: the former encapsulates offsets between the derived dynamical mass $M_{\rm dyn}$ and the true total mass of the system $M_{\rm tot}$, while the latter is the stellar mass-to-light ratio for some fiducial stellar initial mass function (IMF), usually obtained by modeling multiband photometric data. The fact that dynamical studies of nearby early-type galaxies can recover the virial relation suggests that, given appropriate assumptions, $M_{\rm dyn}/M_{\rm tot} \approx 1$ (e.g., Hyde & Bernardi 2009; Cappellari et al. 2013b). Uncertainties in the derivation of $M_{*,\rm IMF}/L$ from multiband photometry, on the other hand, can be significant (of order 0.1–0.2 dex) depending on the treatment of star formation history (SFH), metallicity, and dust (e.g., Leja et al. 2019). The short formation timescales and low attenuation generally inferred for passive galaxies help to reduce these uncertainties considerably (e.g., Pforr et al. 2012), but the extent to which these assumptions remain valid at higher redshift remains to be seen.

The remaining terms of Equation (1), $M_{\rm tot}/M_*$ and $M_*/M_{*,\rm IMF}$, encapsulate the relationship between different physical components of the galaxy and are the most likely to be affected by evolutionary processes. $M_{\rm tot}/M_*$ is the ratio of total to stellar mass and is related to the balance of baryonic and dark matter (DM) within a given aperture—typically the effective radius, $r_{\rm e}$—while $M_*/M_{*,\rm IMF}$ accounts for differences between the assumed and true stellar IMF. Variation of the IMF might be expected owing to the evolution of interstellar medium (ISM) properties with redshift and stellar mass, but there is no clear theoretical consensus as to how these changes might manifest in the observed galaxy population (see, e.g., Chabrier et al. 2014; Krumholz 2014, and references therein).

In nearby galaxies, deep photometric and spectroscopic data can be used to study the relationship between galaxies, their stellar populations, and the properties of their dark matter halos in great detail. Van Dokkum & Conroy (2010) used stellar population models to show that massive early-type galaxies host a large population of low-mass stars in their cores ($\lesssim r_{\rm e}/8$), suggesting a very bottom-heavy IMF compared to the Milky

Way (MW) and other nearby star-forming galaxies. These results were consistent with a complementary analysis of strong-lensing systems by Treu et al. (2010), who additionally found evidence for *systematic* variation of the IMF from MW-like at low stellar velocity dispersions to Salpeter (1955) or heavier in the most massive galaxies. Cappellari et al. (2012) obtained similar results based on modeling the spatially resolved stellar kinematics of galaxies in the ATLAS$^{\rm 3D}$ survey. Stellar population results from studies like van Dokkum & Conroy (2010) are uniquely sensitive to a galaxy's stellar content, but dynamical IMF constraints cannot necessarily distinguish between IMF variation and changes in the central DM fraction. Cappellari et al. (2013b) showed that the typical dark matter fraction within $r_{\rm e}$, $f_{\rm DM}[<r_{\rm e}]$, is relatively low (9%–17%) and, while $f_{\rm DM}[<r_{\rm e}]$ tends to increase with increasing galaxy mass, this variation cannot account for the observed trends in *total* M/L, supporting their conclusion of a systematically varying IMF (Cappellari et al. 2013a); unfortunately, the picture becomes complicated if there is no clear distinction between the baryonic and dark matter distributions (e.g., Thomas et al. 2011). Even though there is no consensus on the exact correlations between IMF normalization (or shape) and observed galaxy properties, variability of the IMF is now supported by a number of different studies using a wide range of stellar population, lensing, and dynamical techniques (e.g., Thomas et al. 2011; Conroy & van Dokkum 2012; Dutton et al. 2012; Cappellari et al. 2013a; Conroy et al. 2013; Ferreras et al. 2013; Spiniello et al. 2014; Martín-Navarro et al. 2015a; Parikh et al. 2018; but see also Smith et al. 2015).

At intermediate redshift, Tortora et al. (2018) used data from the Kilo Degree Survey (KiDS) and Sloan Digital Sky Survey (SDSS) to show that the locally observed correlations between stellar mass, dynamical mass, stellar velocity dispersion, and structural parameters are already in place by $z \sim 0.65$, but that high-redshift quiescent galaxies likely have lower $f_{\rm DM}[<r_{\rm e}]$ at fixed stellar velocity dispersion than nearby galaxies (see also Beifiori et al. 2014; Tortora et al. 2014). Shetty & Cappellari (2015) found a similar decrease in the central dark matter fraction for massive galaxies at $z \approx 0.8$, while at the same time reporting a Salpeter-like IMF consistent with massive galaxies at $z = 0$ (see also Shetty & Cappellari 2014; Sonnenfeld et al. 2015; Martín-Navarro et al. 2015b). Extending such studies of kinematic scaling relations beyond $z > 1$ remains challenging. While an abundance of massive, compact red galaxies have been identified using deep Hubble Space Telescope (HST) and ground-based imaging (e.g., Cimatti et al. 2004; Daddi et al. 2005; Whitaker et al. 2011, 2013), kinematic data for individual galaxies have been notoriously difficult to obtain (e.g., Kriek et al. 2009). The development of efficient, highly multiplexed near-infrared spectrographs such as MOSFIRE at Keck (McLean et al. 2012) and the K-band Multi Object Spectrograph (KMOS) at the ESO Very Large Telescope (Sharples et al. 2012, 2013) has led to rapid growth in the number of kinematic measurements at $z > 1.4$, but individual samples remain relatively small and have been analyzed using a wide variety of methods that make combining the results from different surveys difficult.

In this paper we undertake a homogeneous reanalysis of currently available kinematic data at high redshift in order to study the key parameters governing the behavior of Equation (1) over cosmic time, namely, the central dark matter fraction and normalization of the stellar IMF. Our sample

comprises 58 quiescent galaxies at $1.4 \leqslant z \leqslant 2.1$ with stellar velocity dispersion measurements and high-resolution HST/ WFC3 imaging available. These data include 17 new stellar velocity dispersion measurements obtained as part of the VLT IR IFU Absorption Line Survey (VIRIAL; Mendel et al., in preparation), in addition to measurements from a variety of samples in the literature. We derive dynamical properties based on both a straightforward application of the virial theorem and more complex dynamical models, allowing us to test the influence of different assumptions about galaxy structure on the study of high-redshift stellar kinematics.

The outline of this paper is as follows: In Section 2 we describe the compilation of high-redshift galaxies, along with a comparison sample at $z = 0$. In Section 3 we discuss our modeling of galaxy surface brightness profiles and the calculation of dynamical masses. The main results of this work—the relationship between dynamical and stellar masses, central dark matter fraction, and dynamical constraints on the normalization of the stellar IMF—are presented in Section 4. In Section 5 we discuss our results in the context of the high- and low-redshift galaxy populations. We summarize our conclusions in Section 6.

Throughout this paper we use AB magnitudes (Oke & Gunn 1983) and adopt a flat ΛCDM cosmology with $\Omega_\Lambda = 0.7$, $\Omega_M = 0.3$, and $H_0 = 70 \, \mathrm{km \, s^{-1} \, Mpc^{-1}}$.

2. Samples and Data

2.1. KMOS Observations at 1.5 < z < 2.0

Our analysis includes new spectroscopic data for 17 galaxies in the redshift range $1.5 < z < 2.0$ observed as part of the VIRIAL GTO survey (Mendel et al., in preparation) using KMOS (Sharples et al. 2012, 2013). These galaxies were selected from 3D-HST (Brammer et al. 2012; Skelton et al. 2014; Momcheva et al. 2016) in the COSMOS, GOODS-S, and UDS fields to have $m_{\mathrm{F140W}} \leqslant 22.5$ mag and be classified as quiescent according to their rest-frame $U - V$ and $V - J$ colors using the criteria described by Whitaker et al. (2011; see also Williams et al. 2009), shown in the top panel of Figure 1. Their general properties are given in Table 1.

2.1.1. Observations and Data Reduction

Observations of VIRIAL galaxies were carried out between 2014 and 2016 using the KMOS YJ band (1–1.36 μm). Data were taken using a standard object-sky-object pattern with individual exposure times of 300 s. Each science exposure was offset by between 0.″1 and 0.″6 in order to avoid bad pixels in the final extracted spectra. Along with our science targets, we assigned one IFU from each of the three KMOS spectrographs to a reference star that we used to monitor the ambient conditions (seeing, atmospheric transmission, etc.), pointing accuracy, and point-spread function (PSF) shape. Due to the relatively small angular size of the KMOS IFUs (2.″8 × 2.″8), sky exposures were taken nodding completely off-source. Total on-source integration times range from 440 to 740 minutes (see Table 1).

Data were reduced using a combination of the Software Package for Astronomical Reductions with KMOS pipeline tools (SPARK; Davies et al. 2013) and custom Python scripts. In the following we briefly outline the steps used to produce calibrated one-dimensional spectra. Details of the VIRIAL reduction will be described in a future paper (Mendel et al., in preparation). Calibration exposures (dark, arc, and flat) were
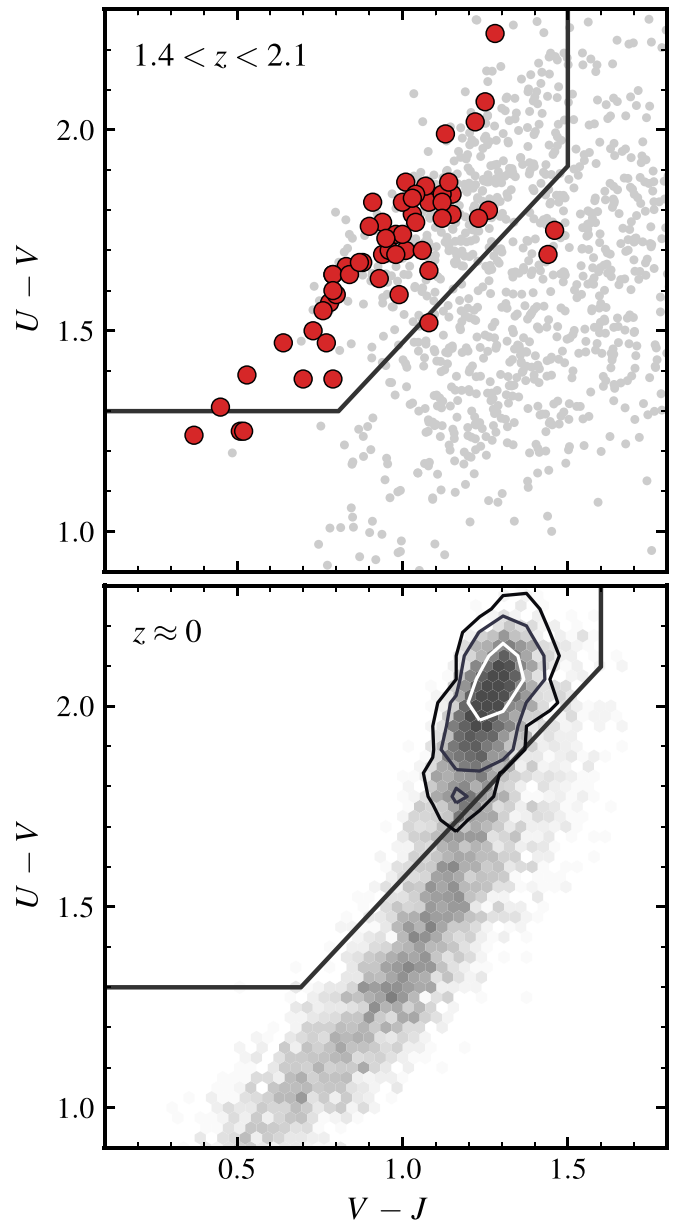

Figure 1. Color−color selection used to identify quiescent galaxies for the high-redshift sample (top panel) and the low-redshift GAMA/SDSS data (bottom panel). The *UVJ* selection window is taken from Whitaker et al. (2011). In the top panel, small (gray) points show the underlying distribution of galaxies with $1.4 < z < 2.1$ from 3D-HST, while large filled (red) circles indicate the distribution of high-redshift data discussed in Sections 2.1 and 2.2. Note that although some galaxies in the high-redshift sample fall outside of the *UVJ* selection window, we nevertheless include them in our analysis based on the presence of strong absorption features and the relative lack of emission lines in their spectra. Contours in the bottom panel show the distribution of $U - V$ and $V - J$ colors for galaxies with mass-weighted stellar ages older than 9 Gyr.

reduced using standard SPARK routines to produce flat-field, wavelength, and spatial calibration frames. When processing science frames, we first corrected each raw image for a readout-channel-dependent bias term estimated from reference pixels around the perimeter of each detector. We then adjusted the wavelength and spatial illumination calibrations for each exposure based on the positions and relative flux of bright sky lines before subtracting the object and sky images. The brightness of atmospheric OH lines can vary significantly

Table 1
General Properties of 17 VIRIAL Targets

Field	ID	R.A. (J2000)	Decl. (J2000)	m_{F160W} (mag)	$(U - V)_{\mathrm{rf}}$	$(V - J)_{\mathrm{rf}}$	Exposure (min.)
UDS	22480	34.3353	−5.2017	20.85	1.83	1.03	670
UDS	24891	34.4458	−5.1940	21.37	1.65	1.08	735
GOODS-S	39364	53.0628	−27.7265	20.99	1.70	1.06	475
GOODS-S	42113	53.1279	−27.7189	20.95	1.99	1.13	495
GOODS-S	43548	53.1294	−27.7073	21.82	1.47	0.77	505
COSMOS	6977	150.0695	2.2500	21.62	1.73	0.95	650
UDS	22802	34.4469	−5.2007	21.05	1.70	0.96	635
UDS	29352	34.4696	−5.1786	21.44	1.69	0.94	740
UDS	10237	34.3148	−5.2433	20.75	1.78	1.12	440
COSMOS	7411	150.1770	2.2552	21.37	1.82	1.00	630
UDS	35111	34.4536	−5.1589	21.63	1.67	0.87	740
UDS	32892	34.3896	−5.1681	21.17	1.55	0.76	660
UDS	38073	34.3365	−5.1490	21.30	1.38	0.79	635
COSMOS	6396	150.1728	2.2441	21.89	1.69	0.98	615
COSMOS	9227	150.0618	2.2737	21.47	1.60	0.79	620
COSMOS	7391	150.0773	2.2548	22.01	1.39	0.53	650
COSMOS	2816	150.1411	2.2085	21.43	1.84	1.04	650

between object and sky exposures (∼10% on 5- to 10-minute timescales; Ramsay et al. 1992; Davies 2007), often leading to significant systematic residuals in the initial sky-subtracted frames. In order to limit the impact of these systematics on our final spectra, we performed a second-order correction to the sky for each IFU using residuals measured in other IFUs in the same detector, excluding the IFU of interest.

One-dimensional spectra were extracted directly from the flat-fielded, illumination-corrected, and sky-subtracted detector frames for each exposure separately. Since VIRIAL targets are typically undetected in individual 300 s exposures, we used the available 3D-HST/CANDELS F125W imaging (Grogin et al. 2011; Koekemoer et al. 2011; Skelton et al. 2014) to model the source flux distribution and mask neighboring objects in the optimal extraction. The HST images were convolved to match the KMOS PSF measured from the reference stars in each exposure, which were also used to adjust for changes in transmission between exposures. Individual optimally extracted spectra were then corrected for telluric absorption using synthetic atmospheric models computed with MOLECFIT (Kausch et al. 2014) and combined using inverse variance weights. Uncertainties on the output spectra were estimated using bootstrap combines of the individual 1D spectra for each object. The typical spectral resolution in the extracted 1D spectra (as measured from sky lines) ranges from $R = 3000$ to 3500 ($\sigma_{\mathrm{inst}} \approx 36$–$42$ km s^{-1}) depending on arm and detector (see also Wisnioski et al. 2019).

2.1.2. Stellar Masses and Velocity Dispersions

We estimated stellar velocity dispersions for VIRIAL galaxies using a simultaneous fit to the observed KMOS spectrum and multiband photometry from 3D-HST (Skelton et al. 2014). We generated model spectral energy distributions (SEDs) using FSPS v2.4 (Conroy et al. 2009; Conroy & Gunn 2010) assuming a lognormal SFH with

$$\mathrm{SFR}(t) = \begin{cases} \dfrac{1}{t\sqrt{2\pi\tau^2}} e^{-\frac{(\ln t - \ln t_0)^2}{2\tau^2}} & \text{if } t \leqslant t_{\mathrm{trunc}} \\ 0 & \text{if } t > t_{\mathrm{trunc}}, \end{cases} \quad (2)$$

where t is the age of the universe, t_0 is the delay time, and τ controls the width of the distribution (see also Gladders et al. 2013). The additional parameter t_{trunc} allows for star formation to be abruptly truncated and provides added flexibility when modeling the SFHs of quiescent galaxies. We stress that our adoption of a lognormal SFH is motivated by its flexibility compared to more commonly used τ or delayed-τ models, rather than an assumption that galaxies' SFHs are intrinsically lognormal (e.g., Gladders et al. 2013; Abramson et al. 2016; Diemer et al. 2017). In Appendix A we show that our derived velocity dispersions are not biased by the use of a parametric SFH. We modeled the effects of dust using a two-component extinction law that includes a foreground screen and additional attenuation toward young stellar populations ($<10^7$ yr), which are assumed to remain embedded within their birth clouds (see, e.g., Charlot & Fall 2000). We used the reddening curve of Calzetti et al. (2000) and, following Wuyts et al. (2013), adopted a relationship between the total V-band extinction A_V and the additional extinction toward young stellar population A_{extra} such that $A_{\mathrm{extra}} = 0.9A_V - 0.15A_V^2$. For simplicity we assume a fixed solar metallicity; in Appendix A we show that changing the metallicity by ±0.2 dex leads to systematic shifts in the derived velocity dispersion of $\lesssim2\%$.

Before fitting, templates were smoothed to match the wavelength-dependent KMOS resolution measured from sky lines in extracted 1D spectra. The final matched templates include an additional (constant) offset of $\sigma_{\mathrm{offset}} = 65$ km s^{-1} to account for the resolution difference between KMOS ($\sigma_{\mathrm{inst}} \approx 40$ km s^{-1}) and the adopted MILES spectral library ($\sigma_{\mathrm{MILES}} \approx 75$ km s^{-1}; Beifiori et al. 2011). This effectively sets a floor for our velocity dispersion measurements of 65 km s^{-1}. We limited our fits to the wavelength range from 3750 to 5300 Å and included a ninth-order additive polynomial—corresponding to ∼1 order per 10,000 km s^{-1}—to minimize the effects of template mismatch on our final velocity dispersion measurements. We verified that our results are not sensitive to the adoption of an additive, as opposed to multiplicative, polynomial. In the end our model has a total of seven free

Table 2
Redshift, Velocity Dispersion, and Stellar Masses of KMOS Galaxies

Field	ID	z	$\log(M_{SPS}/M_\odot)$	$\sigma_e^{[a]}$ (km s^{-1})
UDS	22480	1.5288	11.08	323 ± 42
UDS[b]	24891	1.6031	10.99	146 ± 32
GOODS-S	39364	1.6118	11.10	203 ± 42
GOODS-S	42113	1.6140	11.20	362 ± 65
GOODS-S	43548	1.6143	10.64	169 ± 43
COSMOS	6977	1.6424	10.86	187 ± 32
UDS[b]	22802	1.6660	11.13	337 ± 28
UDS[b]	29352	1.6886	10.91	277 ± 46
UDS	10237	1.7664	11.38	233 ± 23
COSMOS	7411	1.7808	11.09	186 ± 28
UDS	35111	1.8217	10.95	228 ± 36
UDS	32892	1.8243	11.02	206 ± 27
UDS	38073	1.8245	10.94	194 ± 49
COSMOS	6396	1.8364	10.90	169 ± 33
COSMOS	9227	1.8604	10.98	273 ± 41
COSMOS	7391	1.8681	10.54	145 ± 38
COSMOS	2816	1.9230	11.26	297 ± 49

Notes. The formal statistical uncertainties on stellar masses derived from our SED fitting are of order 0.02 dex. Where relevant we include an additional 0.15 dex uncertainty on $\log(M_*/M_\odot)$ in quadrature to account for systematic uncertainties in the determination of stellar masses (see, e.g., Conroy et al. 2009; Mendel et al. 2014).
[a] Velocity dispersion corrected to r_e following the procedure described by van de Sande et al. (2013).
[b] These galaxies are in common with the Belli et al. (2017) sample. See discussion in Section 2.2.

parameters: redshift, z; stellar mass, M_{SPS};[12] three parameters that describe the SFH, τ, t_0, and t_{trunc}; absolute V-band extinction, A_V; and stellar velocity dispersion, σ_*. Samples from the posterior distribution were generated using emcee (Foreman-Mackey et al. 2013), and our final estimates of velocity dispersion and stellar mass were taken as the medians of their respective marginal posterior distributions, with 1σ uncertainties estimated from the 16th and 84th percentiles. We have confirmed that the derived stellar masses do not change significantly if we refit objects using only the available photometric data (i.e., excluding spectra). The final redshifts, stellar masses, and velocity dispersions are provided in Table 2.[13] One-dimensional spectra and the corresponding best-fit models are shown in Figure 2.

In Figure 3 we show a comparison of stellar velocity dispersions obtained with and without the inclusion of photometric data in the fit. The two estimates are generally consistent within their quoted uncertainties, though there is a clear systematic offset in the sense that spectra-only fits return velocity dispersions that are ∼5% lower on average than those that also incorporate photometric data. This stems from the fact that the photometric data generally down-weight the youngest spectral templates—as one might expect from our a priori

selection of galaxies based on their $U − V$ and $V − J$ colors— preferring instead solutions with smaller contributions from (rapidly rotating) early stellar types.

2.2. Literature Data at 1.4 < z < 2.1

In addition to our KMOS data, we have compiled a sample of quiescent galaxies with $1.4 < z < 2.1$ from the literature where HST/WFC3 F160W imaging, multiwavelength photometric catalogs, and stellar velocity dispersion measurements were available. A full accounting of the literature data is given in Table 3, along with a few general galaxy properties. In Figure 4 we show the redshift distribution of our full high-redshift galaxy sample.

This literature sample includes 15 galaxies from Newman et al. (2010), Bezanson et al. (2013a), and Belli et al. (2014a) observed with Keck LRIS at $1.4 < z < 1.6$, as well as two galaxies from the GMASS spectroscopic sample (Cimatti et al. 2008) with velocity dispersions published by Cappellari et al. (2009). Belli et al. (2014a) incorporate the LRIS data from Newman et al. (2010) in their analysis, and there is one galaxy in common between Belli et al. (2014a) and Cappellari et al. (2009). Velocity dispersion measurements derived using near-IR spectroscopy are available for 29 additional galaxies with $1.4 < z < 2.1$ from Toft et al. (2012) and van de Sande et al. (2013), obtained using VLT XShooter, and from Belli et al. (2014b), Barro et al. (2016), and Belli et al. (2017) using Keck MOSFIRE. The sample of Barro et al. (2016) includes one galaxy in common with the LRIS sample of Newman et al. (2010) and Belli et al. (2014a), and there are several galaxies in common between Toft et al. (2012), van de Sande et al. (2013), Belli et al. (2014b), and Belli et al. (2017). See Table 3 for details.

There are three galaxies in common between Belli et al. (2017) and the KMOS sample described in Section 2.1—UDS 24891, UDS 29352, and UDS 22802—which are highlighted in Tables 2 and 3. For UDS 22802, the two independent velocity dispersion measurements are in relatively good agreement (337 ± 28 vs. 291 ± 31 km s^{-1}); however, for the other two galaxies the discrepancy is larger: 277 ± 46 km s^{-1} versus 146 ± 31 km s^{-1} for UDS 29352 (2.4σ offset) and 146 ± 32 km s^{-1} versus 391 ± 71 km s^{-1} for UDS 24891 (3.1σ offset). Although we are not in a position to assess which of these measurements are "correct," we note that adopting σ_e as measured by Belli et al. (2017) for these galaxies results in large offsets between their dynamical and stellar masses (see Section 4.1), such that UDS 29352 (UDS 24891) would have the highest (lowest) dynamical-to-stellar mass ratio in the sample. Nevertheless, in the absence of additional data we adopt a final σ_e for these objects based on an error-weighted average of the quoted measurements, with an increased uncertainty to reflect the large discrepancy between quoted values; in Section 3.1 we describe in more detail how we combine data for galaxies with multiple velocity dispersion measurements.

In order to ensure that our high-redshift sample is as homogeneous as possible, we remeasured stellar masses for all galaxies using the SED fitting procedure described in Section 2.1.2. In most cases, multiwavelength photometric catalogs were available from either the Newfirm Medium Band Survey (NMBS; Whitaker et al. 2011) or 3D-HST (Skelton et al. 2014). Several galaxies in the UDS field—UDS 55531 and UDS 53937 from Bezanson et al. (2013a), as well as UDS
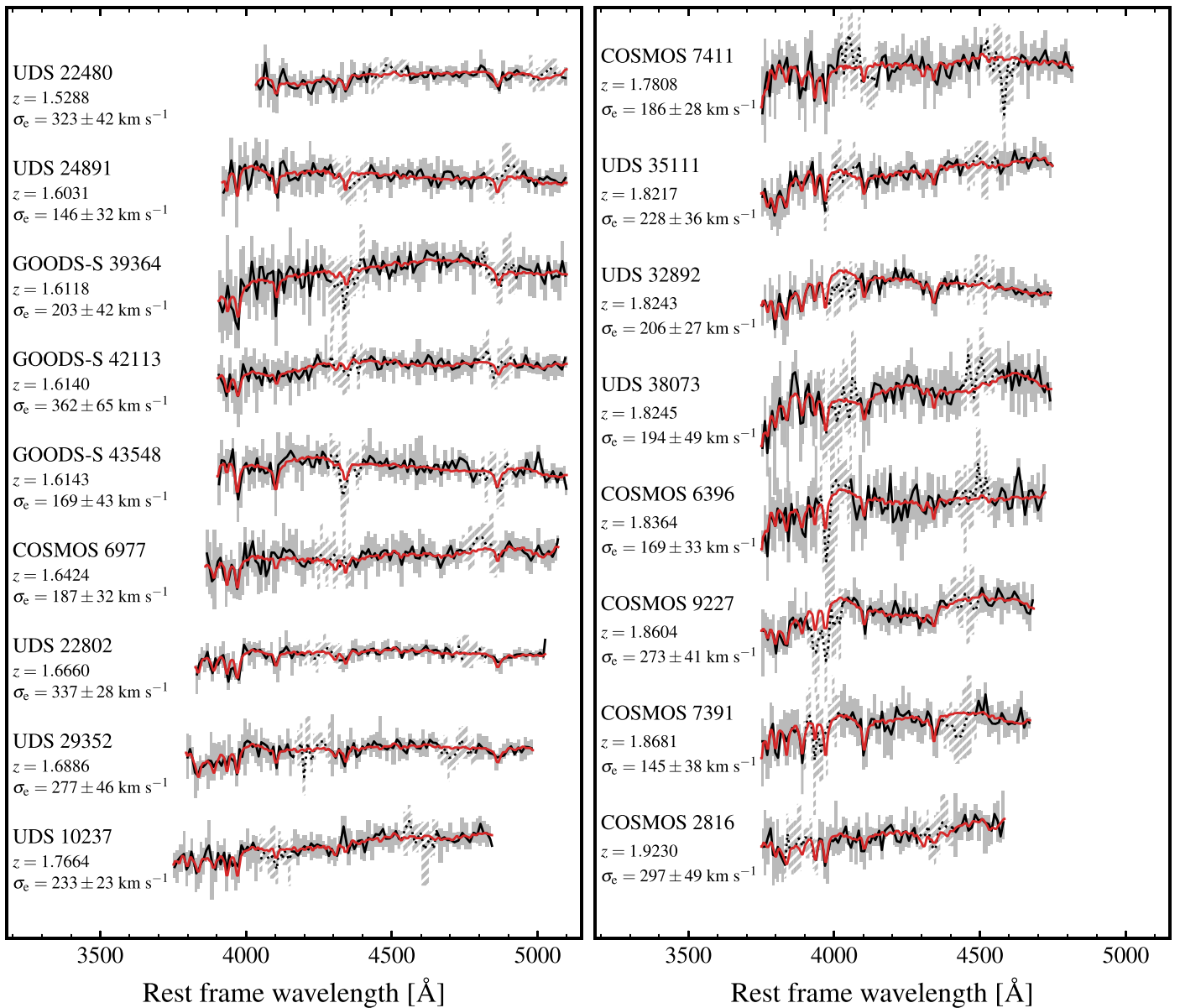
---

[12] For clarity we will refer to stellar masses derived via SED fitting as M_{SPS} in order to distinguish them from those derived using dynamical methods. In the context of Equation (1) these represent $M_{*,IMF}$, the stellar mass derived using a fiducial, in this case Chabrier (2003), IMF.
[13] The dispersions quoted in Table 2 have been corrected for the effects of seeing and scaled to the luminosity-weighted mean within the half-light radius following the procedure outlined by van de Sande et al. (2013). The derived corrections range between 1.02 and 1.1 and are consistent with similar corrections derived directly from the dynamical modeling discussed in Section 3.3.2.

Figure 2. KMOS spectra for the 17 galaxies described in Section 2.1. The extracted one-dimensional spectra (black) and uncertainties (gray) have been median rebinned in a 15-pixel (∼10 Å rest-frame) moving window for display purposes. The best-fit model is overplotted in red. Dotted lines and hatching indicate regions of the spectra that are significantly contaminated by sky emission and absorption features.

19627 from Toft et al. (2012) and van de Sande et al. (2013)—fall outside of the 3D-HST footprint, and for these objects we adopted the combined Subaru/XMM-Newton Deep Survey (SXDS; Furusawa et al. 2008) and UKIRT Infrared Deep Sky Survey (Lawrence et al. 2007) catalogs described by Simpson et al. (2012). We supplemented these data with deep Spitzer/IRAC 3.6 and 4.5 μm flux measurements from Ashby et al. (2013), which were corrected to match the 3″ apertures used by Simpson et al. (2012) using the UKIDSS K-band mosaics. Stellar masses derived for the literature sample are provided in Table 3.

2.3. Comparison Sample at z ≈ 0

We identified a comparison sample of quiescent galaxies at low redshift from the SDSS Legacy Survey (Sloan Digital Sky Survey; Abazajian et al. 2009) using the same color-based selection criteria as at high redshift. We select galaxies with $0.02 \leqslant z \leqslant 0.2$ and that also have stellar velocity dispersions measured by the Portsmouth group (see Thomas et al. 2013) using pPXF (Cappellari & Emsellem 2004). In order to avoid potential biases in the SED fitting between our high- and low-redshift data, we limit our selection to galaxies in the GAMA DR2 survey area (Galaxy and Mass Assembly; Driver et al. 2011; Liske et al. 2015), where Wright et al. (2016) provide aperture-matched photometric catalogs covering from the ultraviolet to infrared.

We computed rest-frame colors for these galaxies using EAZY (Brammer et al. 2008), and the resulting distribution is shown in the bottom panel of Figure 1, where we again adopt the color criteria of Whitaker et al. (2011) to select quiescent
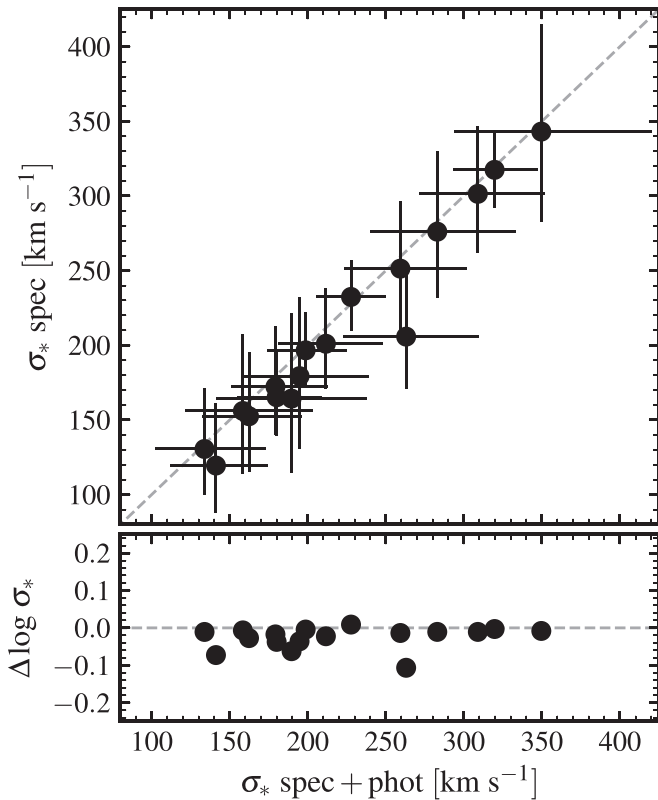
Figure 3. Comparison of velocity dispersions derived from fits to the KMOS spectra alone to those that include multiband photometry in the fits. Uncertainties are $\pm 1\sigma$ estimated from the marginalized posterior distribution described in Section 2.1.2. The dashed line marks a one-to-one correlation between the two measurements. There is a clear systematic offset between the two measurements such that the spectra-plus-photometry fits predict slightly higher velocity dispersions on average.

galaxies. Stellar masses for the 4546 galaxies satisfying this selection were estimated from fits to their far-UV to K-band photometry using the procedures described in Sections 2.1.2 and 2.2. Based on these data, we derive a redshift-dependent stellar mass limit following the approach of Sohn et al. (2017) and Zahid et al. (2019), such that our final sample of 3108 galaxies is mass complete at the 97.5% level.

The passive galaxy population appears to grow significantly from $z = 2$ to the present day, suggesting that our low-redshift data contain galaxies that are too young to be descendants of the galaxies in our high-z sample. While there is no consensus on the magnitude of such progenitor bias effects (e.g., Carollo et al. 2013; Belli et al. 2015; Fagioli et al. 2016), it is nevertheless important to account for them in our analysis. We use here the mass-weighted stellar ages derived by Comparat et al. (2017) for SDSS galaxies using FIREFLY (Wilkinson et al. 2017), and we select those galaxies with ages older than 9 Gyr as the most likely descendants of our high-z sample. This identifies a subsample of 792 galaxies, or $\sim$27% of the full quiescent sample. In the following we will discuss results for both the full and age-selected samples.

3. Dynamical Modeling

The main focus of this work is a discussion of the dynamical constraints afforded by current high-redshift quiescent galaxy samples and a comparison with low-redshift data. In this section we describe the key quantities required for this analysis —stellar velocity dispersions, structural parameters (sizes, Sérsic indices, etc.)—as well as our estimates of dynamical masses and their related quantities.

3.1. Stellar Velocity Dispersions

Galaxies in our high-redshift sample have stellar velocity dispersions derived within a range of apertures and are based on data obtained with a variety of instruments and extraction methods. These measurements therefore require some degree of homogenization in order to be meaningfully combined. In many cases authors quote velocity dispersions corrected such that they represent the luminosity-weighted mean with one effective radius, σ_e, and we adopt these values when available. Where velocity dispersions are quoted within a different aperture—as is the case for Cappellari et al. (2009), Newman et al. (2010), Toft et al. (2012), Bezanson et al. (2013a), and Barro et al. (2016)—we correct the quoted velocity dispersion to one effective radius following the procedure outlined in van de Sande et al. (2013).

In cases where multiple velocity dispersion measurements were available we used an inverse-variance-weighted average of the published dispersions, after correcting them to a common r_e aperture. In addition to their propagated uncertainties, we included a term (in quadrature) to account for large offsets between quoted dispersions, taken as half of the range of dispersion measurements. The one exception to this procedure is GOODS-N 17678, where the velocity dispersion measured by Newman et al. (2010) differs significantly from the measurements of Belli et al. (2014b) and Barro et al. (2016); for this object we used an average of only the Belli et al. (2014b) and Barro et al. (2016) dispersions.

In our low-redshift sample all dispersions were measured from spectra within a common $3''$ aperture, corresponding to the SDSS fiber diameter. Where we quote individual stellar velocity dispersions, these aperture values have been corrected to one effective radius, again following the procedure described by van de Sande et al. (2013) and using structural parameters described below. However, in our dynamical modeling (see Section 3.3) we fit directly to model dispersions computed within the $3''$ (fiber) aperture, accounting for seeing effects. Although the physical scale subtended by the $3''$ SDSS fibers increases dramatically over the redshift range of our low-z sample, the physical quantities derived from our dynamical models are *independent* of redshift at fixed stellar mass, suggesting that the use of aperture measurements does not bias our results.

3.2. Structural Parameters

We adopted two different approaches to measuring structural properties for our galaxy samples: first using galfit (Peng et al. 2002) to model their two-dimensional surface brightness distributions using a single Sérsic profile (Sérsic 1963), and second using the Multi-Gaussian Expansion (MGE) approach described by Emsellem et al. (1994; see also Cappellari 2002).

Table 3
Properties of the High-z Literature Sample

Field	3D-HST ID	z	$\log(M_{SPS}/M_\odot)$[a]	$(U-V)$	$(V-J)$	Ref. ID	σ_e (km s^{-1})	Reference
COSMOS	30145	1.4010	10.90	1.84	1.09	19498	250 ± 39	Belli et al. (2014a)
AEGIS	5087	1.4060	11.00	1.84	1.15	42109	369 ± 48	Belli et al. (2014a)
						E9	295 ± 69[b]	Newman et al. (2010)
GOODS-S	40623	1.4149	10.89	2.07	1.25	2239	116 ± 36[b]	Cappellari et al. (2009)
GOODS-S	42466	1.4150	11.07	1.82	1.08	5020	181 ± 54	Belli et al. (2014a)
						2470	147 ± 27[b]	Cappellari et al. (2009)
GOODS-S	43042	1.4190	11.32	2.24	1.28	4906	298 ± 26	Belli et al. (2014a)
AEGIS	⋯	1.4235	11.26	1.57	0.78	A17300	276 ± 7[b]	Bezanson et al. (2013a)
COSMOS	21628	1.4320	10.82	1.79	1.15	13880	169 ± 70	Belli et al. (2014a)
COSMOS[d]	31780	1.4390	10.78	1.75	1.46	20841	267 ± 52	Belli et al. (2014a)
COSMOS	31136	1.4420	10.93	1.86	1.07	20275	221 ± 70	Belli et al. (2014a)
UDS	1854	1.4560	11.49	1.74	0.98	29410	355 ± 98	van de Sande et al. (2013)
UDS[d]	⋯	1.4848	11.53	1.52	1.08	U55531	260 ± 24[b]	Bezanson et al. (2013a)
COSMOS	⋯	1.5222	11.34	1.77	0.94	C20866	284 ± 24[b]	Bezanson et al. (2013a)
COSMOS	⋯	1.5223	11.26	1.66	0.83	C21434	229 ± 17[b]	Bezanson et al. (2013a)
COSMOS	17364	1.5260	11.02	1.84	1.12	17364	168 ± 84	Belli et al. (2017)
COSMOS	17361	1.5270	10.86	1.63	0.93	17361	169 ± 43	Belli et al. (2017)
COSMOS	17641	1.5280	10.79	1.70	1.01	17641	142 ± 54	Belli et al. (2017)
COSMOS	17089	1.5280	11.37	2.02	1.23	17089	348 ± 57	Belli et al. (2017)
AEGIS	17926	1.5730	11.14	1.79	1.03	17926	231 ± 39	Belli et al. (2017)
AEGIS	22719	1.5790	11.13	1.87	1.14	22719	262 ± 51	Belli et al. (2017)
COSMOS	28523	1.5825	11.38	1.82	0.91	34265	377 ± 54	Belli et al. (2014a)
						18265	400 ± 72	van de Sande et al. (2013)
AEGIS	⋯	1.5839	11.24	1.47	0.64	A21129	275 ± 10[b]	Bezanson et al. (2013a)
GOODS-N	17678	1.5980	11.00	1.59	0.80	2653	174 ± 27	Belli et al. (2014a)
						GN5	245 ± 37[b]	Newman et al. (2010)
						12632	187 ± 36[b]	Barro et al. (2016)
UDS[c]	24891	1.6035	10.99	1.65	1.08	24891	391 ± 71	Belli et al. (2017)
UDS	35616	1.6090	11.19	1.64	0.79	35616	198 ± 49	Belli et al. (2017)
UDS	30737	1.6200	11.37	1.77	1.04	30737	307 ± 82	Belli et al. (2017)
UDS	⋯	1.6210	10.93	1.31	0.47	U53937	251 ± 21[b]	Bezanson et al. (2013a)
UDS	43367	1.6240	11.26	1.80	1.26	43367	299 ± 74	Belli et al. (2017)
UDS	30475	1.6330	10.83	1.38	0.70	30475	296 ± 109	Belli et al. (2017)
UDS	32707	1.6470	11.25	1.82	1.12	32707	174 ± 30	Belli et al. (2017)
COSMOS	16629	1.6570	10.67	1.64	0.79	16629	358 ± 76	Belli et al. (2017)
UDS	37529	1.6650	11.13	1.78	1.23	37529	232 ± 60	Belli et al. (2017)
UDS[c]	22802	1.6665	11.13	1.70	0.96	22802	291 ± 31	Belli et al. (2017)
GOODS-N[d]	11470	1.6740	10.77	1.25	0.51	8231	221 ± 36[b]	Barro et al. (2016)
GOODS-N	24033	1.6740	10.80	1.64	0.84	17360	155 ± 31[b]	Barro et al. (2016)
GOODS-N	3604	1.6750	10.69	1.67	0.88	2617	317 ± 118[b]	Barro et al. (2016)
UDS[c]	29352	1.6895	10.91	1.69	0.94	29352	146 ± 31	Belli et al. (2017)
COSMOS	19958	1.7220	10.75	1.50	0.73	19958	169 ± 87	Belli et al. (2017)
COSMOS	17255	1.7390	10.97	1.74	1.00	17255	147 ± 40	Belli et al. (2017)
AEGIS	25526	1.7520	10.84	1.59	0.99	25526	134 ± 36	Belli et al. (2017)
COSMOS[d]	⋯	1.8000	11.31	1.24	0.37	7447	287 ± 53	van de Sande et al. (2013)
UDS[d]	⋯	2.0360	11.20	1.25	0.52	19627	304 ± 41	van de Sande et al. (2013)
						19627	335 ± 56[b]	Toft et al. (2012)
COSMOS	13083	2.0880	11.10	1.76	0.90	13083	197 ± 52	Belli et al. (2017)
COSMOS	11494	2.0920	11.58	1.87	1.01	7865	446 ± 57	van de Sande et al. (2013)
						31719	358 ± 30	Belli et al. (2014b)
						11494	319 ± 26	Belli et al. (2017)
COSMOS[d]	12020	2.0960	11.34	1.69	1.44	31769	312 ± 65	Belli et al. (2014b)

Notes.

[a] Stellar masses are rederived in this work following the method described in Section 2.1.2.

[b] Dispersions corrected to R_e following the Appendix B of van de Sande et al. (2013).

[c] These galaxies are in common with the KMOS sample. See discussion in Section 2.2.

[d] These galaxies fall outside of the UVJ quiescent selection defined by Whitaker et al. (2011); however, their spectra show strong absorption features characteristic of post-starburst galaxies, as well as weak or absent [O II] emission, so we include them in our analysis.
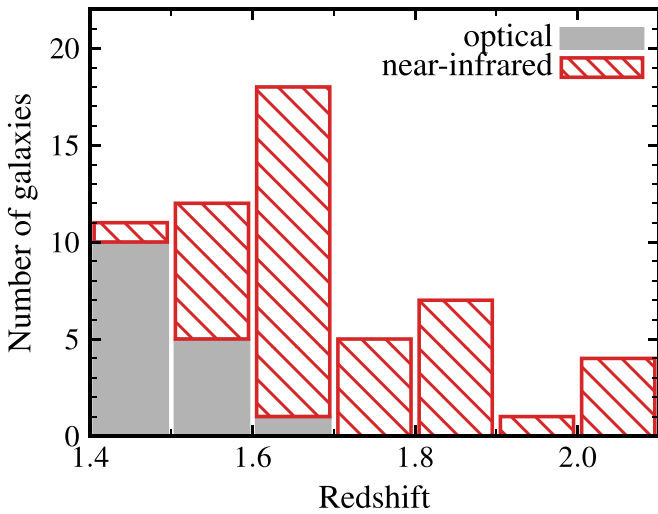
Figure 4. Redshift distribution of our passive galaxy sample. The filled gray histogram shows the distribution of data observed using red-sensitive optical detectors, while the hatched red distribution shows the contribution of near-infrared observations.

3.2.1. Sérsic Profile Fits

In our high-redshift sample 50/58 galaxies fall within the HST WFC3/F160W imaging footprint of the CANDELS survey (Grogin et al. 2011; Koekemoer et al. 2011), and for these objects we used the mosaics and composite PSFs described by Skelton et al. (2014).[14] The remaining eight galaxies were observed separately using HST WFC3/F160W as part of HST-GO-12167 (PI: Franx; AEGIS 17300, AEGIS 21129, COSMOS 21434, COSMOS 20866, COSMOS 07447, UDS 53937, and UDS 55531) and HST-GO-13002 (PI: Williams; UDS 19627) for a single orbit each with total exposure times of 2611 or 2411 s, respectively. Level 2 data products were retrieved from the Hubble Legacy Archive (HLA),[15] and we constructed empirical PSFs for these objects by stacking the images of bright unsaturated stars in each combined frame. We generated segmentation maps for the HLA images using SExtractor (Bertin & Arnouts 1996) with parameters similar to those given by Skelton et al. (2014) for 3D-HST.

We used galfit to model the surface brightness distribution of the primary galaxy, while also including in the fit neighboring galaxies with $m_{\rm F160W} < 25$ and projected separations $r_{\rm p} < 5$ ($r_{\rm primary} + r_{\rm neighbor}$). Initial estimates of the galaxy sizes, i.e., $r_{\rm primary}$ and $r_{\rm neighbor}$, were taken from the SExtractor output. The local sky background for each object was estimated using the full image by first masking all pixels within 3 Kron radii of nearby sources using the ellipse parameters produced by SExtractor. We then identified the nearest 10,000 unmasked pixels as sky. An initial estimate of the background was taken as the mode of these sky pixels, which was then iteratively refined to obtain our final estimates of the local sky background. Postage stamps for individual objects were then extracted and the local sky background removed; the background level was subsequently held fixed during fitting. The structural parameters derived in this way are consistent with those available in the literature; a direct comparison with literature values is given in Appendix B.1.

An example of our photometric modeling for COSMOS 30145 is shown in Figure 5, with figures for the remaining galaxies included in Appendix C.

Although there are numerous existing catalogs of structural parameter measurements for the SDSS and GAMA (e.g., Simard et al. 2002; Kelvin et al. 2012; Meert et al. 2015), for consistency with our high-z data we chose to rederive these quantities using the methodology described above. We retrieved "corrected" r-band images from the SDSS Data Archive Server (DAS), along with their associated mask and PSF files. We then used SExtractor to generate segmentation images following the procedures described by Simard et al. (2011), and the local sky background for each source was estimated using the method described above. Individual postage stamps and PSFs[16] for each galaxy were then extracted and the background removed. galfit was used to simultaneously fit the primary galaxy and any neighboring sources with $m_r < 22$ and $r_{\rm p} < 5$ ($r_{\rm primary} + r_{\rm neighbor}$). All other sources were masked during the fit. A comparison of our measurements with several different literature catalogs can be found in Appendix B.2.

At both high and low redshift we derive sizes in fixed photometric bands (HST WFC3/F160W and SDSS r band, respectively), which probe different *rest-frame* wavelengths at different redshifts. In the presence of strong color gradients this shift in rest-frame wavelength can systematically bias our size measurements and must be taken into account. Following van der Wel et al. (2014), we define the corrected, in this case r-band, semimajor axis size $r_{\rm e}^{\rm sma}$ as

$$r_{\rm e}^{\rm sma} = r_{\rm e,obs}^{\rm sma}\left(\frac{1+z}{1+z_p}\right)^{\frac{\Delta \log r_{\rm e}}{\Delta \log \lambda}}, \qquad (3)$$

where $r_{\rm e,obs}^{\rm sma}$ is the measured half-light size in either the WFC3/F160W or SDSS r-band filter and z_p is the "pivot redshift." $z_p = 0$ by definition for the GAMA/SDSS sample, as we are correcting to the rest-frame r-band size, while for F160W imaging $z_p = 1.49$. Kelvin et al. (2012) used GAMA data to show that $\Delta \log r_{\rm e}/\Delta \log \lambda = -0.3$ for early-type galaxies on average. Chan et al. (2016) and van der Wel et al. (2014) derive similar values based on their analyses of quiescent galaxies' high redshift, and we therefore adopt $\Delta \log r_{\rm e}/\Delta \log \lambda = -0.3$ for all galaxies in our sample. The typical correction derived in this way is of order 2%–3%, and we adopt these corrected r-band sizes for the remainder of this work.

3.2.2. Multi-Gauss Expansion Fits

While the single-component Sérsic fits described in Section 3.2.1 provide a straightforward summary of the overall surface brightness profile, Sérsic models have several drawbacks that complicate their use in constructing dynamical models. As well as providing a poor description of multi-component profiles (e.g., bulge + disk), the coupling between inner and outer profile shapes makes the Sérsic models extremely sensitive to sky background: over-/undersubtraction of the sky level can significantly affect the inferred inner profile shape. In addition, with the exception of a few special cases,

---

[14] http://3dhst.research.yale.edu/Data.php
[15] http://hla.stsci.edu/

[16] Source-specific PSFs were extracted from the SDSS drField files using the read_PSF routine described at http://classic.sdss.org/dr7/products/images/read_psf.html.
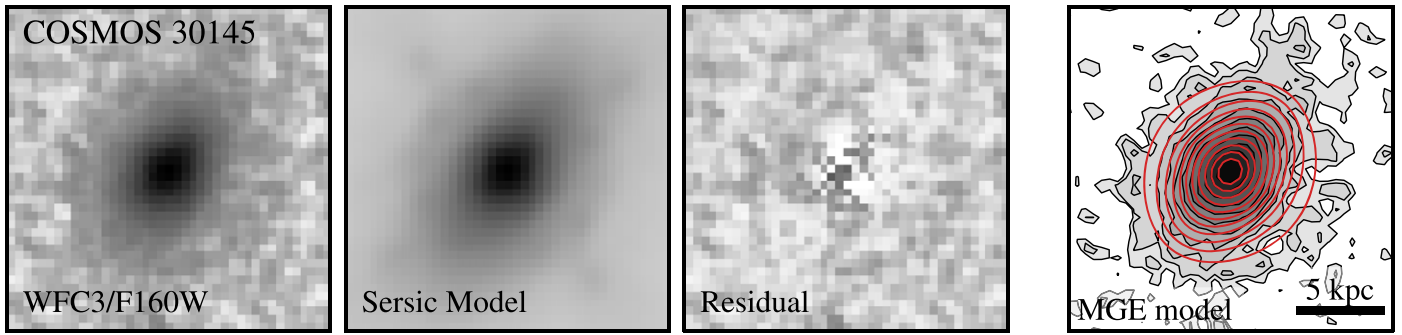
Figure 5. Example of the photometric models adopted for galaxies in our high-redshift sample. From left to right panels show the observed HST WFC3/F160W image from CANDELS/3D-HST (Skelton et al. 2014), the best-fit galfit model, and the resulting image residual. The rightmost panel shows contours from the best-fit MGE model (red) overlaid on the WFC3/F160W image. In all cases the images are plotted in surface brightness units, and contours are evenly spaced in steps of 0.5 mag arcsec^{-2}. Models for all galaxies in our high-redshift sample are shown in Appendix C.

Sérsic profiles cannot be deprojected analytically, making their use for constructing dynamical models computationally expensive compared to simpler functional forms. In this context, modeling galaxies as a sum of individual Gaussian components—so-called multi-Gaussian expansion (MGE; Emsellem et al. 1994; Cappellari 2002)—provides a flexible description of surface brightness profiles that does not require any extrapolation of the profile to large radii, can accommodate multiple photometric components, and can be easily deprojected to obtain an estimate of the three-dimensional luminosity density (see Section 3.3.2).

The starting points for our MGE models were the background-subtracted postage stamps produced as described in Section 3.2.1. We used the results of our Sérsic model fits to subtract neighboring sources before identifying the primary object and producing binned two-dimensional surface brightness measurements using the find_galaxy and sectors_photometry routines described by Cappellari (2002).[17] A model of the surface photometry in terms of nested Gaussians was then derived using the mge_fit_sectors method of Cappellari (2002). For high-redshift galaxies we constructed MGE-based PSF models per field using either composite PSFs provided by Skelton et al. (2014) for galaxies within the CANDELS/3D-HST footprint or else the stacked images of bright stars within the same field for stand-alone observations. In the right panel of Figure 5 we show a comparison of the observed and MGE-derived surface brightness contours for one object, COSMOS 30145.

MGE PSF models for the low-redshift SDSS/GAMA data were constructed on a galaxy-by-galaxy basis using the PSF extracted from the SDSS drField files. In all cases—that is, both high and low redshift—we tied the ellipticity of the fitted Gaussian components together to avoid large variations in the derived axis ratios for low surface brightness components; however, we confirmed that our results are not qualitatively sensitive to this assumption.

3.3. Dynamical Masses

The final piece of information we require is an estimate of total galaxy mass, including both stellar and dark matter components. For this work we investigate two broad approaches to estimating dynamical masses in order to test their sensitivity to underlying assumptions: the first is based on

a simple application of the virial theorem and scaling relations derived for nearby galaxies, while the second relies on more detailed dynamical modeling of the stellar density profile and velocity dispersion.

3.3.1. Virial Mass Estimates

As outlined in Section 1, the tight relationship between size, stellar velocity dispersion, and mass for nearby early-type galaxies can be understood as a consequence of virial equilibrium, where for a pressure-supported system the total mass is given by

$$M_{\rm vir} = \kappa(n) \frac{\sigma_{\rm e}^2 r_{\rm e}^{\rm sma}}{G}. \qquad (4)$$

Here $\kappa(n)$ is the so-called virial coefficient, and in this case it is taken as an analytic function of Sérsic index that encapsulates the effects of structural and orbital nonhomology (e.g., Bertin et al. 2002; Cappellari et al. 2006). We adopt the relation derived by Cappellari et al. (2006) based on spherical, isotropic models,

$$\kappa(n) = 8.87 - 0.831n + 0.0241n^2, \qquad (5)$$

which has been shown to provide a reliable estimate of the total mass for nearby early-type galaxies in the SAURON and ATLAS$^{\rm 3D}$ samples (e.g., Cappellari et al. 2006, 2013b). Note that in Equation (4) we used the semimajor axis size, $r_{\rm e}^{\rm sma}$, following the discussion of Cappellari et al. (2013a, their Figure 14). The semimajor axis size is expected to be more robust to systematic changes in galaxy shapes than the harmonic mean size (e.g., $\sqrt{ab}$, where a and b are the semimajor- and semiminor-axis sizes, respectively), especially for (thin) disk galaxies where the observed b/a is an indicator of inclination rather than intrinsic shape.

3.3.2. Jeans Models

The assumptions of spherical symmetry and isotropy discussed above appear to be reasonable at low redshift; however, high-z quiescent galaxies are known to be flatter on average—that is, have intrinsically lower b/a—than their low-redshift counterparts (e.g., van der Wel et al. 2011; Chang et al. 2013), leading to a possible bias in their derived masses when using Equation (5). We therefore consider an alternative approach to computing dynamical masses based on the Jeans

---

[17] Available at http://purl.org/cappellari/software.

Anisotropic MGE (JAM) method discussed by Cappellari (2008), which allows us to relax these assumptions. The modeling requires as input the MGE-derived surface brightness profile described in Section 3.2.2 and a measurement of the stellar velocity dispersion (see Section 2.1.2 and 2.2).

Following Cappellari (2002), the deprojected luminosity density can be computed from the best-fit MGE decomposition given assumptions about the inclination, which is related to the *intrinsic* axis ratio of an oblate ellipsoid q_{int} by

$$\cos i = \sqrt{\frac{q_{obs}^2 - q_{int}^2}{1 - q_{int}^2}}, \qquad (6)$$

where i is the inclination and q_{obs} is the observed axis ratio. Since in this work we are concerned with the sensitivity of our dynamical mass estimates to possible changes of the intrinsic axis ratio, we computed JAM models over a grid of q_{int} from $0.05 \leqslant q_{int} \leqslant \min(0.95, q_{obs})$ in steps of 0.05; unless otherwise stated, our results are based on marginalizing over q_{int}. In our default modeling we assumed that the velocity ellipsoid is marginally anisotropic with an anisotropy parameter $\beta \equiv 1 - \sigma_z^2/\sigma_R^2 = 0.2$ (where z and R define directions parallel and perpendicular to the symmetry axis for an axisymmetric system) based on local early-type galaxies (e.g., Cappellari et al. 2007; Thomas et al. 2009). We explored possible systematic effects over a range of anisotropies from $0 \leqslant \beta \leqslant 0.8$ and found that they resulted in variations of the derived dynamical masses of at most a few percent, consistent with previous results (e.g., Wolf et al. 2010; Dutton et al. 2013); all results are therefore quoted adopting our fiducial value of $\beta = 0.2$.

We adopt two different implementations of the JAM modeling procedure distinguished by their treatment of baryonic and dark matter components. In the first instance we assume that the total mass is proportional to the light at all radii, i.e., mass-follows-light (MFL). This provides a self-consistent estimate of the dynamical mass-to-light ratio $(M/L)_{MFL}$. MFL models have been shown to reliably recover the total mass within relatively small apertures ($\lesssim r_e$), even in the presence of multiple mass components (e.g., Cappellari et al. 2006; Williams et al. 2010), and provide a baseline comparison for dynamical masses computed following Equation (5). A similar approach was used by Shetty & Cappellari (2014) to study quiescent galaxies at $z \sim 0.8$ in the DEEP2 survey. In the MFL case the best-fitting value of $(M/L)_{MFL}$ for a given combination of q_{int} and β is simply given by $(M/L)_{MFL} = (\sigma_{e,obs}/\sigma_{e,model})^2$, where $\sigma_{e,obs}$ is the observed aperture velocity dispersion and $\sigma_{e,model}$ is the model prediction assuming $(M/L)_{MFL} = 1$. Instead, our second implementation includes an explicit dark matter component described by a spherical NFW halo profile (Navarro et al. 1996). With sufficient sampling of the velocity field it is possible to independently constrain the stellar mass-to-light ratio, $(M/L)_{*,NFW}$, and properties of the dark matter halo (e.g., Cappellari et al. 2013b; Übler et al. 2018). However, the aperture velocity dispersions used here cannot be used to break the degeneracy between stellar and dark matter components, leading us to impose additional constraints on the properties of the dark matter halo. Starting from our photometric estimates of galaxy stellar mass, we assigned dark matter halo masses based on the evolving stellar-to-halo mass relation derived by

Moster et al. (2013). We then used the calculations of Diemer & Kravtsov (2015) to assign a halo concentration. This halo profile was then fed back into the JAM modeling procedure along with the MGE-based stellar density profile, and a grid search was used to determine the mass-to-light ratio of the stellar component, $(M/L)_{*,NFW}$, as a function of q_{int} and β.

In the explicit DM halo case we obtain an estimate of the dark matter fraction within r_e, $f_{DM}[<r_e]$, defined as

$$f_{DM}[<r_e] = \frac{M_{DM}}{M_{*,NFW} + M_{DM}}. \qquad (7)$$

We compute $f_{DM}[<r_e]$ within a volume defined by $4\pi r_e^3/3$, where for consistency with the literature r_e is the circularized half-light radius ($\equiv r_e^{sma} \times \sqrt{q_{obs}}$), and the relevant masses are computed using the derived M/L values and deprojected MGE luminosity densities. For all galaxies we use the rest-frame r-band sizes computed following Equation (3).

4. Results

In this section we present the main results of this work, which are focused in two areas: the relationship between dynamical and stellar masses, and the interplay between dark matter content and the IMF at high redshift. Derived quantities for our high-redshift sample are provided in Table 4, and they are described in more detail in Sections 2 and 3.

4.1. The Relationship between Dynamical and Stellar Mass

In Figure 6 we show a comparison of dynamical and stellar masses for the different dynamical mass estimates described in Section 3.3. Two features are apparent. First, fixed stellar mass high-redshift galaxies appear to have dynamical masses that are ~0.20 dex lower on average than their low-redshift counterparts. This offset appears regardless of the dynamical mass estimate used (i.e., M_{MFL} vs. M_{vir}). Second, the correlation between dynamical and stellar mass is superlinear regardless of redshift, in the sense that the ratio of dynamical to stellar mass increases with increasing stellar mass. Such a "tilt" in the relationship between dynamical and stellar mass has been studied extensively at low redshift and has commonly been interpreted as variation of the central dark matter fraction and/or stellar IMF (e.g., Renzini & Ciotti 1993; Cappellari et al. 2013a; Dutton et al. 2013); in the following sections we will consider evidence for changes in the dark matter fraction and stellar IMF among high-redshift galaxies in more detail.

Finally, there are a number of galaxies in Figure 6 with stellar mass estimates formally larger than their derived dynamical masses. While this cannot physically be the case, there a number of factors that influence the apparent trend, particularly at low stellar masses. Observational uncertainties at $\log(M_{SPS}/M_\odot) < 11$ increase dramatically, driven primarily by the increased uncertainty on galaxy size as one pushes down the size−mass relation. While these increased uncertainties cannot in and of themselves explain the apparent shift toward low dynamical masses, when combined with the tilt of the relation described above they can nevertheless increase the fraction of galaxies with low dynamical-to-stellar mass ratios. In addition, we will show in Section 4.1.2 that our dynamical modeling likely underestimates the dynamical mass for galaxies that are intrinsically flat. Enforcing a flat structure for face-on galaxies can increase dynamical mass estimates by as much as ~0.2 dex. Indeed, nearly 59% of galaxies with

Table 4
Derived Properties of High-redshift Galaxies

Field	ID[a]	r_e^{smac} (kpc)	n	q_{obs}	$\log(L_r)$ ($L_\odot$)	σ_* (km s^{-1})	$\log(M/L_r)_{VIR}$ ($M_\odot/L_\odot$)	$\log(M/L_r)_{MFL}$ ($M_\odot/L_\odot$)	$\log(M/L_r)_{*,NFW}$ ($M_\odot/L_\odot$)	$\log(f_{DM})$	Reference
COSMOS	30145	1.25 ± 0.26	4.31	0.42	11.28 ± 0.01	250 ± 39	−0.30 ± 0.17	−0.32 ± 0.14	$-0.33^{+0.11}_{-0.18}$	$-1.77^{+0.18}_{-0.11}$	2
AEGIS	05087	1.18 ± 0.26	2.53	0.41	11.36 ± 0.01	345 ± 54	−0.02 ± 0.17	−0.12 ± 0.14	$-0.12^{+0.11}_{-0.18}$	$-2.02^{+0.18}_{-0.11}$	2, 3
GOODS-S	40623	2.29 ± 0.27	2.98	0.88	11.18 ± 0.01	116 ± 36	−0.52 ± 0.27	−0.50 ± 0.27	$-0.61^{+0.12}_{-1.51}$	$-0.63^{+0.34}_{-0.11}$	4
GOODS-S	42466	2.53 ± 0.30	6.42	0.91	11.45 ± 0.01	154 ± 30	−0.75 ± 0.18	−0.51 ± 0.17	$-0.61^{+0.13}_{-0.45}$	$-0.62^{+0.21}_{-0.12}$	2, 4
GOODS-S	43042	3.48 ± 0.30	5.62	0.63	11.68 ± 0.01	298 ± 26	−0.20 ± 0.09	−0.11 ± 0.08	$-0.19^{+0.07}_{-0.11}$	$-0.72^{+0.08}_{-0.07}$	2
AEGIS	A17300[b]	2.93 ± 0.29	5.48	0.68	11.89 ± 0.01	276 ± 7	−0.54 ± 0.06	−0.47 ± 0.02	$-0.52^{+0.02}_{-0.03}$	$-0.92^{+0.02}_{-0.02}$	5
COSMOS	21628	1.09 ± 0.26	3.07	0.70	11.18 ± 0.01	169 ± 70	−0.52 ± 0.37	−0.48 ± 0.36	$-0.50^{+0.16}_{-1.35}$	$-1.42^{+0.73}_{-0.20}$	2
COSMOS	31780	2.45 ± 0.27	1.28	0.32	11.11 ± 0.01	267 ± 52	0.39 ± 0.18	0.30 ± 0.17	$0.29^{+0.13}_{-0.25}$	$-1.74^{+0.24}_{-0.13}$	2
COSMOS	31136	2.01 ± 0.28	4.24	0.45	11.33 ± 0.01	221 ± 70	−0.24 ± 0.28	−0.29 ± 0.28	$-0.31^{+0.16}_{-0.78}$	$-1.36^{+0.50}_{-0.17}$	2
UDS	01854	2.40 ± 0.27	2.69	0.49	11.99 ± 0.01	355 ± 98	−0.32 ± 0.24	−0.39 ± 0.24	$-0.43^{+0.15}_{-0.69}$	$-1.12^{+0.40}_{-0.16}$	6
UDS	U55531[b]	8.01 ± 0.40	3.90	0.75	11.94 ± 0.01	260 ± 24	−0.09 ± 0.08	−0.05 ± 0.08	$-0.43^{+0.10}_{-0.56}$	$-0.22^{+0.08}_{-0.07}$	5
COSMOS	C20866[b]	2.71 ± 0.27	3.45	0.66	11.84 ± 0.01	284 ± 24	−0.35 ± 0.09	−0.38 ± 0.07	$-0.43^{+0.07}_{-0.10}$	$-0.90^{+0.08}_{-0.06}$	5
COSMOS	C21434[b]	1.99 ± 0.26	3.43	0.72	11.87 ± 0.01	229 ± 17	−0.70 ± 0.09	−0.69 ± 0.06	$-0.74^{+0.06}_{-0.08}$	$-0.97^{+0.07}_{-0.06}$	5
COSMOS	17364	2.86 ± 0.28	2.77	0.48	11.37 ± 0.01	168 ± 84	−0.28 ± 0.44	−0.35 ± 0.43	$-0.42^{+0.13}_{-1.88}$	$-0.90^{+0.57}_{-0.26}$	7
COSMOS	17361	1.91 ± 0.26	2.13	0.70	11.45 ± 0.01	169 ± 43	−0.49 ± 0.23	−0.58 ± 0.22	$-0.61^{+0.15}_{-0.51}$	$-1.07^{+0.34}_{-0.15}$	7
COSMOS	17089	5.65 ± 0.40	4.58	0.86	11.85 ± 0.01	348 ± 57	0.06 ± 0.15	0.12 ± 0.14	$0.01^{+0.12}_{-0.34}$	$-0.54^{+0.16}_{-0.11}$	7
COSMOS	17641	1.19 ± 0.27	5.32	0.92	11.26 ± 0.02	142 ± 54	−0.86 ± 0.35	−0.68 ± 0.33	$-0.70^{+0.16}_{-1.20}$	$-1.14^{+0.57}_{-0.19}$	7
UDS	22480	1.84 ± 0.26	4.54	0.55	11.62 ± 0.01	323 ± 42	−0.26 ± 0.13	−0.24 ± 0.11	$-0.25^{+0.09}_{-0.15}$	$-1.54^{+0.14}_{-0.09}$	1
AEGIS	17926	5.45 ± 0.43	4.19	0.72	11.61 ± 0.02	231 ± 39	−0.04 ± 0.15	−0.05 ± 0.15	$-0.20^{+0.12}_{-0.11}$	$-0.47^{+0.17}_{-0.11}$	7
AEGIS	22719	2.25 ± 0.30	6.08	0.94	11.54 ± 0.01	262 ± 51	−0.39 ± 0.18	−0.26 ± 0.17	$-0.29^{+0.13}_{-0.30}$	$-0.88^{+0.22}_{-0.12}$	7
COSMOS	28523	1.88 ± 0.26	2.93	0.25	11.95 ± 0.01	385 ± 45	−0.32 ± 0.12	−0.43 ± 0.10	$-0.46^{+0.09}_{-0.13}$	$-1.49^{+0.12}_{-0.09}$	2, 6
AEGIS	A21129[b]	1.96 ± 0.26	7.15	0.49	11.97 ± 0.01	275 ± 10	−0.94 ± 0.08	−0.64 ± 0.03	$-0.66^{+0.03}_{-0.04}$	$-1.40^{+0.03}_{-0.03}$	5
GOODS-N	17678	1.29 ± 0.26	8.00	0.70	11.61 ± 0.01	179 ± 23	−1.25 ± 0.16	−0.91 ± 0.11	$-0.94^{+0.10}_{-0.15}$	$-1.13^{+0.13}_{-0.09}$	2, 3, 8
UDS	24891	1.88 ± 0.26	2.45	0.88	11.48 ± 0.01	187 ± 126	−0.46 ± 0.59	−0.43 ± 0.58	$-0.45^{+0.16}_{-1.92}$	$-1.10^{+0.74}_{-0.31}$	1, 7
UDS	35616	4.47 ± 0.48	6.07	0.65	11.82 ± 0.02	198 ± 49	−0.61 ± 0.23	−0.50 ± 0.21	$-0.67^{+0.12}_{-1.31}$	$-0.50^{+0.25}_{-0.16}$	7
GOODS-S	39364	1.58 ± 0.26	2.97	0.95	11.64 ± 0.01	203 ± 42	−0.64 ± 0.19	−0.56 ± 0.18	$-0.61^{+0.13}_{-0.34}$	$-1.07^{+0.26}_{-0.13}$	1
GOODS-S	42113	1.96 ± 0.27	6.53	0.78	11.66 ± 0.01	362 ± 65	−0.33 ± 0.17	−0.16 ± 0.16	$-0.19^{+0.12}_{-0.23}$	$-1.17^{+0.20}_{-0.12}$	1
GOODS-S	43548	0.94 ± 0.25	3.91	0.62	11.30 ± 0.01	169 ± 43	−0.74 ± 0.25	−0.66 ± 0.22	$-0.68^{+0.15}_{-0.42}$	$-1.57^{+0.37}_{-0.15}$	1
UDS	30737	3.32 ± 0.27	2.71	0.51	11.82 ± 0.01	307 ± 82	−0.13 ± 0.24	−0.21 ± 0.23	$-0.30^{+0.13}_{-1.21}$	$-0.77^{+0.35}_{-0.16}$	7
UDS	U53937[b]	0.63 ± 0.25	3.78	0.78	11.74 ± 0.01	251 ± 21	−1.01 ± 0.19	−0.86 ± 0.07	$-0.87^{+0.07}_{-0.08}$	$-1.91^{+0.08}_{-0.06}$	5
UDS	43367	2.69 ± 0.28	5.18	0.53	11.60 ± 0.01	299 ± 74	−0.18 ± 0.22	−0.16 ± 0.21	$-0.23^{+0.14}_{-0.73}$	$-0.83^{+0.32}_{-0.15}$	7
UDS	30475	0.98 ± 0.25	3.04	0.75	11.55 ± 0.01	296 ± 109	−0.44 ± 0.34	−0.39 ± 0.32	$-0.39^{+0.17}_{-0.94}$	$-1.91^{+0.72}_{-0.18}$	7
COSMOS	06977	1.49 ± 0.25	1.43	0.79	11.41 ± 0.01	187 ± 32	−0.43 ± 0.17	−0.48 ± 0.15	$-0.51^{+0.12}_{-0.23}$	$-1.23^{+0.20}_{-0.12}$	1
UDS	32707	1.75 ± 0.26	3.62	0.25	11.70 ± 0.01	174 ± 30	−0.83 ± 0.16	−0.92 ± 0.15	$-0.99^{+0.12}_{-0.29}$	$-0.96^{+0.22}_{-0.12}$	7
COSMOS	16629	0.74 ± 0.25	2.40	0.72	11.32 ± 0.01	358 ± 76	−0.13 ± 0.24	−0.07 ± 0.18	$-0.08^{+0.13}_{-0.28}$	$-2.29^{+0.27}_{-0.13}$	7
UDS	37529	2.32 ± 0.28	3.83	0.64	11.52 ± 0.01	232 ± 60	−0.29 ± 0.23	−0.28 ± 0.23	$-0.33^{+0.14}_{-0.68}$	$-0.98^{+0.35}_{-0.15}$	7
UDS	22802	1.50 ± 0.25	2.33	0.36	11.66 ± 0.01	316 ± 31	−0.27 ± 0.11	−0.23 ± 0.09	$-0.24^{+0.07}_{-0.10}$	$-1.85^{+0.10}_{-0.07}$	1, 7
GOODS-N	11470	2.85 ± 0.29	4.00	0.71	11.60 ± 0.01	221 ± 36	−0.33 ± 0.15	−0.31 ± 0.14	$-0.35^{+0.12}_{-0.21}$	$-1.15^{+0.19}_{-0.11}$	8
GOODS-N	24033	1.09 ± 0.25	3.20	0.72	11.43 ± 0.01	155 ± 31	−0.84 ± 0.20	−0.85 ± 0.17	$-0.88^{+0.13}_{-0.28}$	$-1.25^{+0.24}_{-0.13}$	8
GOODS-N	03604	0.82 ± 0.25	2.55	0.25	11.31 ± 0.01	317 ± 118	−0.19 ± 0.35	−0.25 ± 0.32	$-0.25^{+0.17}_{-0.92}$	$-2.47^{+0.82}_{-0.17}$	8
UDS	29352	1.08 ± 0.25	4.79	0.77	11.52 ± 0.01	187 ± 70	−0.87 ± 0.34	−0.73 ± 0.33	$-0.75^{+0.16}_{-1.05}$	$-1.41^{+0.65}_{-0.18}$	1, 7
COSMOS	19958	2.62 ± 0.29	2.94	0.84	11.45 ± 0.01	169 ± 87	−0.39 ± 0.45	−0.39 ± 0.45	$-0.46^{+0.13}_{-1.86}$	$-0.91^{+0.59}_{-0.26}$	7

Table 4
(Continued)

Field	ID[a]	r_e^{smac} (kpc)	n	q_{obs}	$\log(L_r)$ ($L_\odot$)	σ_* (km s^{-1})	$\log(M/L_r)_{VIR}$ ($M_\odot/L_\odot$)	$\log(M/L_r)_{MFL}$ ($M_\odot/L_\odot$)	$\log(M/L_r)_{*,NFW}$ ($M_\odot/L_\odot$)	$\log(f_{DM})$ ($M_\odot/L_\odot$)	Reference
COSMOS	17255	1.54 ± 0.26	2.93	0.60	11.46 ± 0.01	147 ± 40	−0.75 ± 0.25	−0.76 ± 0.24	$-0.81^{+0.15}_{-0.69}$	$-1.05^{+0.40}_{-0.16}$	7
AEGIS	25526	0.83 ± 0.25	2.48	0.51	11.37 ± 0.01	134 ± 36	−0.98 ± 0.27	−0.99 ± 0.23	$-1.02^{+0.15}_{-0.52}$	$-1.39^{+0.41}_{-0.15}$	7
UDS	10237	3.07 ± 0.29	4.07	0.67	11.88 ± 0.01	233 ± 23	−0.54 ± 0.10	−0.58 ± 0.09	$-0.79^{+0.10}_{-0.20}$	$-0.39^{+0.09}_{-0.07}$	1
COSMOS	07411	2.04 ± 0.27	3.95	0.85	11.64 ± 0.01	186 ± 28	−0.66 ± 0.14	−0.69 ± 0.13	$-0.77^{+0.11}_{-0.23}$	$-0.73^{+0.16}_{-0.10}$	1
COSMOS	C07447[b]	1.67 ± 0.25	5.58	0.68	12.19 ± 0.01	287 ± 53	−1.03 ± 0.18	−0.87 ± 0.16	$-0.92^{+0.13}_{-0.29}$	$-0.95^{+0.22}_{-0.12}$	6
UDS	35111	0.78 ± 0.25	2.56	0.31	11.56 ± 0.01	228 ± 36	−0.74 ± 0.19	−0.75 ± 0.14	$-0.76^{+0.11}_{-0.18}$	$-2.02^{+0.18}_{-0.11}$	1
UDS	32892	1.63 ± 0.26	4.05	0.87	11.73 ± 0.01	206 ± 27	−0.76 ± 0.14	−0.68 ± 0.11	$-0.72^{+0.10}_{-0.16}$	$-1.12^{+0.14}_{-0.09}$	1
UDS	38073	2.78 ± 0.29	7.06	0.84	11.68 ± 0.02	194 ± 49	−0.78 ± 0.23	−0.54 ± 0.22	$-0.59^{+0.14}_{-0.62}$	$-0.84^{+0.31}_{-0.15}$	1
COSMOS	06396	1.44 ± 0.25	1.38	0.86	11.47 ± 0.01	169 ± 33	−0.58 ± 0.19	−0.61 ± 0.17	$-0.65^{+0.13}_{-0.31}$	$-1.09^{+0.25}_{-0.13}$	1
COSMOS	09227	1.16 ± 0.25	2.75	0.68	11.65 ± 0.01	273 ± 41	−0.51 ± 0.16	−0.55 ± 0.13	$-0.56^{+0.11}_{-0.17}$	$-1.60^{+0.17}_{-0.15}$	1
COSMOS	07391	0.69 ± 0.25	8.00	0.80	11.40 ± 0.01	145 ± 38	−1.48 ± 0.29	−0.94 ± 0.23	$-0.96^{+0.15}_{-0.45}$	$-1.56^{+0.39}_{-0.15}$	1
COSMOS	02816	1.93 ± 0.25	1.82	0.24	11.74 ± 0.01	297 ± 49	−0.25 ± 0.15	−0.30 ± 0.14	$-0.33^{+0.12}_{-0.22}$	$-1.40^{+0.20}_{-0.11}$	1
UDS	U19627[b]	2.08 ± 0.25	3.48	0.48	12.06 ± 0.01	315 ± 37	−0.57 ± 0.12	−0.65 ± 0.10	$-0.69^{+0.09}_{-0.14}$	$-1.15^{+0.12}_{-0.09}$	6, 9
COSMOS	13083	1.38 ± 0.24	3.43	0.88	11.75 ± 0.01	197 ± 52	−0.84 ± 0.24	−0.73 ± 0.23	$-0.78^{+0.14}_{-0.73}$	$-1.02^{+0.40}_{-0.16}$	7
COSMOS	11494	2.70 ± 0.28	4.68	0.80	12.19 ± 0.01	348 ± 66	−0.58 ± 0.17	−0.53 ± 0.17	$-0.66^{+0.13}_{-0.53}$	$-0.62^{+0.22}_{-0.13}$	6, 7, 10
COSMOS	12020	1.98 ± 0.26	3.87	0.59	11.70 ± 0.01	312 ± 65	−0.26 ± 0.19	−0.38 ± 0.18	$-0.46^{+0.13}_{-0.46}$	$-0.70^{+0.23}_{-0.13}$	10

References. (1) this work; (2) Belli et al. 2014a; (3) Newman et al. 2010; (4) Cappellari et al. 2009; (5) Bezanson et al. 2013a; (6) van de Sande et al. 2013; (7) Belli et al. 2017; (8) Barro et al. 2016; (9) Toft et al. 2012; (10) Belli et al. 2014b.

[a] Unless otherwise noted, IDs correspond to those provided by Skelton et al. (2014) for galaxies in the 3D-HST fields.

[b] These galaxies fall outside the 3D-HST footprint; IDs therefore correspond to those given in their originating publications.

[c] Sizes have been corrected to the rest-frame r band according to Equation (3).
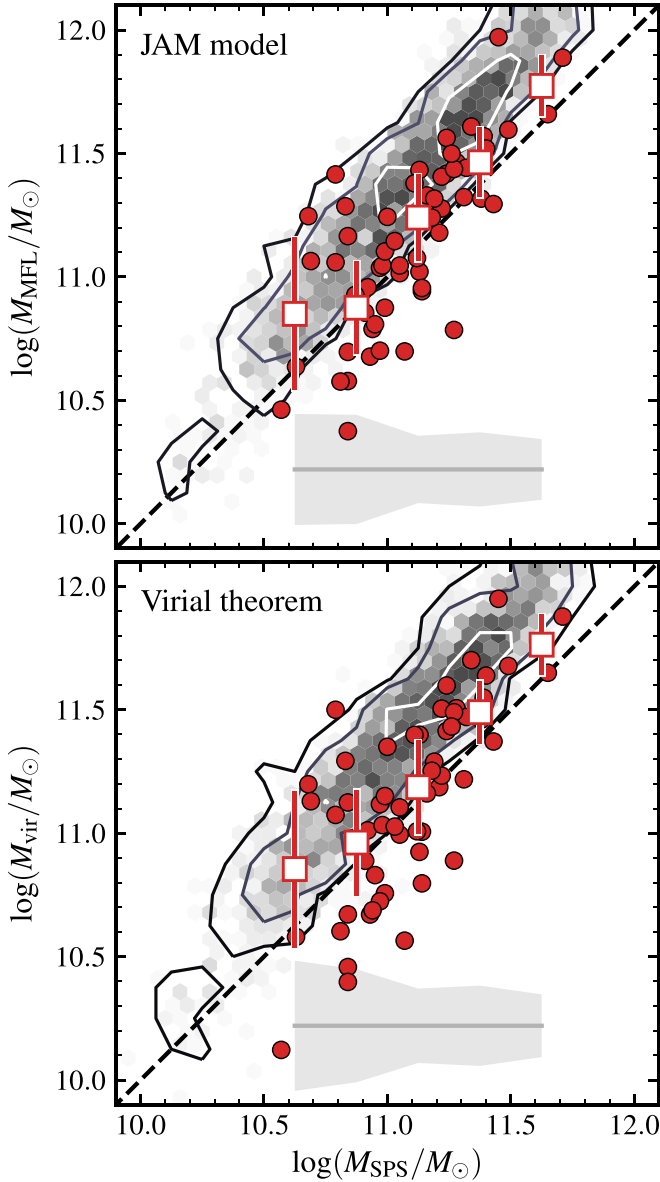
Figure 6. Dynamical vs. stellar masses for two different dynamical mass estimates based on either JAM models (top panel; see Section 3.3.2) or a simple n-dependent virial coefficient (bottom panel; Equations (4) and (5)). Quiescent galaxies at $1.4 < z < 2.1$ are shown as red circles, with large red squares indicating their (binned) median and scatter. The gray shading in the bottom of each panel shows the average uncertainty of individual dynamical mass estimates. The distribution of galaxies in our low-redshift SDSS/GAMA sample is indicated by the background shading. Contours show the 30th, 60th, and 90th percentile distributions of dynamical and stellar masses for local galaxies with ages >9 Gyr, i.e., old enough to be the descendants of galaxies in our high-redshift sample.

$q_{\rm obs} \geqslant 0.7$ have dynamical masses smaller than their derived stellar mass, compared to only 24% for galaxies with $q_{\rm obs} < 0.7$.

4.1.1. Central Dark Matter Fractions

The tendency for high-redshift quiescent galaxies to have lower dynamical-to-stellar mass ratios compared to low redshift has been reported in a number of previous studies (e.g., Toft et al. 2012; van de Sande et al. 2013; Belli et al. 2017) and is generally interpreted as reflecting a systematic decrease in the central dark matter fraction, $f_{\rm DM}[<r_{\rm e}]$. This decline in

dynamical-to-stellar mass ratio appears to occur relatively smoothly with increasing redshift, as shown by a number of studies based on large spectroscopic surveys at $z < 1$ (e.g., Beifiori et al. 2014; Tortora et al. 2014, 2018). Figure 7 shows the cumulative distribution of dynamical-to-stellar mass ratio in the samples considered here. We find an offset in the mean dynamical-to-stellar mass ratio of -0.20 dex when moving to high redshift—$\log(M_{\rm MFL}/M_*) = 0.29 \pm 0.01$ at $z = 0$ compared to 0.09 ± 0.05 at $1.4 \leqslant z \leqslant 2.1$—which is consistent with the results of previous studies. The magnitude of this offset is independent of the dynamical mass estimator used (e.g., $M_{\rm MFL}$ vs. $M_{\rm vir}$) and does not change when considering only the oldest galaxies at $z = 0$ (shown as contours in Figure 6 and light-gray lines in Figure 7). The right panel of Figure 7 shows the same distribution of dynamical-to-stellar mass ratios as the left panel, but with individual measurements normalized by their uncertainties. These can be compared to the dotted–dashed (black) line, which shows the prediction for a standard normal distribution. Although nearly 40% of galaxies in the high-redshift sample have photometrically derived stellar masses that exceed their dynamical masses, given measurement uncertainties the overall distribution is consistent with a positive (albeit small) dynamical-to-stellar mass ratio on average.

We can examine the evolution of $f_{\rm DM}[<r_{\rm e}]$ more directly using our dynamical models that include an explicit dark matter component, where dark matter fractions are computed following Equation (7). In Figure 8 we show $f_{\rm DM}[<r_{\rm e}]$ as a function of $M_{*,\rm NFW}$, the dynamical mass of the stellar component. While there is significant uncertainty in the individual measurements of $f_{\rm DM}[<r_{\rm e}]$ at high redshift, the overall trends support our interpretation of Figures 6 and 7 in terms of an evolution in the central dark matter fraction: galaxies at $1.4 < z < 2.1$ have a mean $f_{\rm DM}[<r_{\rm e}] = 6.6 \pm 1.0\%$, a factor of >2 lower than galaxies of a similar mass in our SDSS/GAMA sample at $z = 0$ ($f_{\rm DM}[<r_{\rm e}] \approx 16.3 \pm 0.3\%$; cf. 17% from Cappellari et al. 2013b). Furthermore, our low-redshift $f_{\rm DM}[<r_{\rm e}]$ measurements are consistent with the values derived by Thomas et al. (2011) and Cappellari et al. (2013b) based on more detailed dynamical modeling of low-redshift galaxies, suggesting that the observed offset in $f_{\rm DM}[<r_{\rm e}]$ between different redshifts is unlikely to be due to differences in the modeling approach. The above comparison between high- and low-redshift galaxies at fixed mass must nevertheless be made with some caution, as individual galaxies are expected to evolve from $z = 2$ to 0; we will revisit the evolution of $f_{\rm DM}[<r_{\rm e}]$ using more carefully matched progenitor and descendant samples in Section 5.

Using data from the SINS survey, Förster Schreiber et al. (2009) found that star-forming galaxies at $z \sim 2$ are strongly baryon dominated, even for a Chabrier (2003) IMF, suggesting little room for either a bottom-heavy Salpeter IMF or significant dark matter. These results have recently been supported by kinematic data for hundreds of early star-forming disks in the KMOS$^{\rm 3D}$ (Wisnioski et al. 2015, 2019) and MOSDEF (Kriek et al. 2015) surveys (e.g., Price et al. 2016; Wuyts et al. 2016; Lang et al. 2017; Übler et al. 2017), as well as the detailed analysis of outer rotation curves for individual high-redshift disks (Genzel et al. 2017; Genzel et al. 2020; but see also Tiley et al. 2019). For comparison, in Figure 8 we show the dark matter fractions derived by Genzel et al. (2017, shown as squares and upper limits for a subsample of high-redshift, star-forming galaxies), which are in good agreement with the $f_{\rm DM}[<r_{\rm e}]$ measurements derived here.
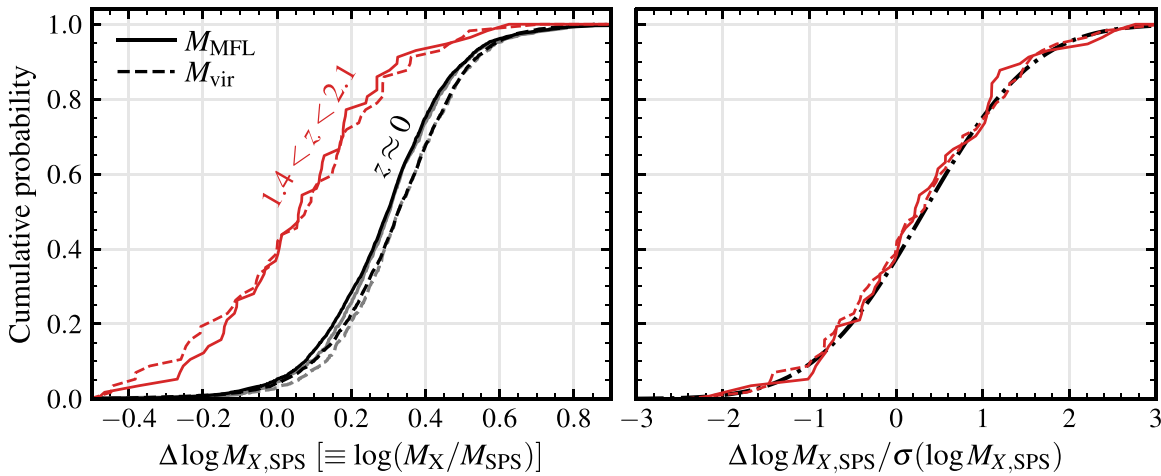
Figure 7. Cumulative distribution of the dynamical-to-stellar mass ratio, comparing different redshifts and dynamical mass estimators. In the left panel, red curves show high-redshift data, while the $z = 0$ GAMA/SDSS data are shown in black. The light-gray curves show the results for galaxies in the GAMA/SDSS sample with mass-weighted stellar ages >9 Gyr (i.e., accounting for possible progenitor bias effects). Solid and dashed curves indicate dynamical masses derived using JAM models assuming MFL and a simple virial estimator, respectively. The median dynamical-to-stellar mass ratio is ~0.2 dex lower at high redshift compared to $z = 0$ regardless of the adopted mass estimator. In the right panel we show the error-normalized distribution of dynamical-to-stellar mass ratios. Red curves again show high-redshift data, while the dotted–dashed curve shows the expectation based on a standard normal distribution with the same mean as the observed data. A number of galaxies have stellar masses apparently larger than their associated dynamical mass estimates; however, the overall population is consistent (within uncertainties) with a positive dynamical-to-stellar mass ratio on average.
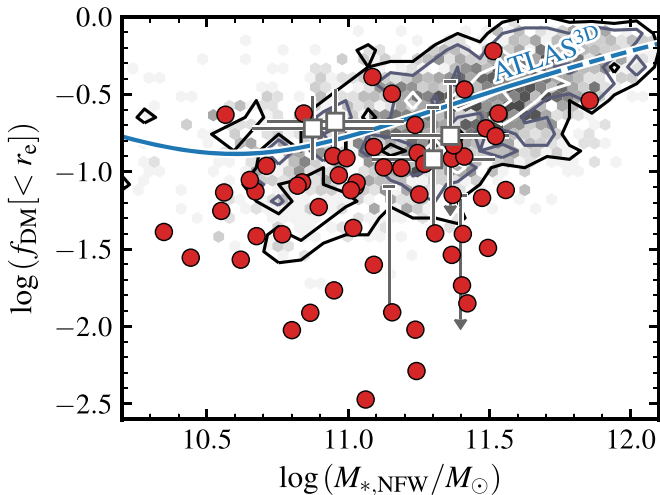
Figure 8. Estimated dark matter fraction within the half-light radius, $f_{DM}[<r_e]$, as a function of the dynamical mass in the stellar component $M_{*,NFW}$. Circles (red) show the measured dark matter fractions for our sample of quiescent galaxies at $1.4 < z < 2.1$. The blue curve shows the results from Cappellari et al. (2013b) derived using IFU data from the local ATLAS³ᴰ sample, while background shading shows the distribution of $f_{DM}[<r_e]$ for galaxies in our GAMA/SDSS at $z = 0$. Squares and upper limits represent dark matter fractions measured by Genzel et al. (2017) for a sample of massive disk galaxies at $z > 1$, which are consistent with our quiescent galaxy data at high redshift. Overall, the dark matter fractions in high-redshift quiescent galaxies appear lower by a factor of >2 on average compared to galaxies of the same stellar mass at $z = 0$.

4.1.2. The Effects of Unresolved Rotation

One of our goals in comparing multiple dynamical mass estimators is to assess the impact and importance of different modeling assumptions on the inferred properties of high-redshift galaxies. In that regard, the main distinctions between M_{vir} and M_{MFL} are the assumptions of spherical symmetry and isotropy inherited through the application of Equations (4) and (5).

In practice, the dynamical masses derived here are only weakly dependent on changes of the anisotropy, β, at fixed q_{int} for a range of values consistent with local early-type galaxies ($0 \leqslant \beta \leqslant 0.6$; Cappellari et al. 2007). It is therefore unlikely that the assumption of isotropy has a significant impact on the results presented in Figures 6–8, particularly given the typical uncertainties on measurements of σ_e (20%–30%). On the other hand, changes in assumed galaxy structure—for example, from spherically symmetric to oblate and axisymmetric—can systematically bias dynamical mass estimates depending on the degree of intrinsic flattening and relative importance of rotation versus pressure support.

Crucially, there is growing observational evidence that quiescent galaxies at high redshift may indeed be rotationally flattened, violating the assumption of spherical symmetry inherent in Equation (5). Bezanson et al. (2018) showed that passive galaxies at $z \sim 1$ have on average a higher proportion of rotational support (higher V/σ) than galaxies of the same mass at low redshift (see also van der Wel & van der Marel 2008). These results are consistent with the observed evolution of photometrically derived axis ratios over the same redshift range, which favor a significant portion of the quiescent galaxy population having $0.2 \leqslant q_{int} \leqslant 0.3$ (van der Wel et al. 2011; Chang et al. 2013; Hill et al. 2019). Belli et al. (2017) argued that the dynamical masses of quiescent galaxies at $z > 1.5$ are statistically consistent with a factor of ~2 increase in V/σ compared to $z = 0$ based on their correlation with observed axis ratios. More directly, a handful of strongly lensed passive galaxies at $z > 2$ have resolved kinematic profiles that are consistent with being rotationally flattened disks (e.g., Newman et al. 2015; Toft et al. 2017; Newman et al. 2018).

In the case of integrated (as opposed to resolved) absorption-line kinematics, rotation is expected to manifest as a dependence of the measured velocity dispersion—and, by extension, dynamical mass—on galaxy inclination. For an oblate model observed at inclination i with no azimuthal variation of the velocity ellipsoid (i.e., $\sigma_\phi/\sigma_r = 1$, with σ_ϕ and σ_r describing velocity dispersion in the azimuthal and radial

directions, respectively), the second moment of the velocity distribution $\sigma_{\rm obs}$ can be written as

$$\sigma_{\rm obs}^2 = \sigma^2(1 - \beta \cos^2 i)$$
$$\times \left[1 + \left(\frac{V}{\sigma}\right)_e^2 \frac{\sin^2 i}{(1 - \beta \cos^2 i)^2} \right], \qquad (8)$$

where V and σ are the flux-weighted mean circular velocity and velocity dispersion within the effective radius for the edge-on case ($i = 90°$) with $(V/\sigma)_e^2 \equiv V_e^2/\sigma_e^2$, and β is the anisotropy parameter as defined in Section 3.3.2. Following Belli et al. (2017), substituting Equation (8) into Equation (4) and normalizing by the dynamical mass predicted for the face-on case (i.e., $i = 0°$) gives

$$\frac{M_{\rm vir}(i)}{M_{\rm vir}(i = 0°)} = \frac{1 - \beta \cos^2 i}{1 - \beta}$$
$$\times \left[1 + \left(\frac{V}{\sigma}\right)_e^2 \frac{\sin^2 i}{(1 - \beta \cos^2 i)^2} \right], \qquad (9)$$

with the relationship between $q_{\rm int}$, $q_{\rm obs}$, and i given by Equation (6). In the isotropic case where $\beta = 0$, Equation (9) reduces to Equation (5) of Belli et al. (2017) modulo a factor $\gamma = 1.51$.[18]

In Figure 9 we show the dynamical-to-stellar mass ratio as a function of $q_{\rm obs}$ for both $M_{\rm vir}$ and $M_{\rm MFL}$ estimates. In the virial theorem case we find evidence for a weak negative correlation between $M_{\rm vir}$ and $q_{\rm obs}$, with a Spearman rank coefficient $\rho = -0.23 \pm 0.10$ ($p = 0.007$), while for the MFL models the correlation is not significant ($\rho = -0.14 \pm 0.10$; $p = 0.068$). Individual galaxies are color-coded according to their Sérsic indices as derived from the profile fits described in Section 3.2.1. In contrast to Belli et al. (2017), we find no significant dependence of the dynamical-to-stellar mass ratio on Sérsic index in either case with $\rho = -0.20 \pm 0.10$ ($p = 0.04$) and -0.08 ± 0.10 ($p = 0.44$), which appears to preclude a simple exclusion of disk-dominated systems based on their structure and motivates a more detailed examination of the correlation between $q_{\rm obs}$ and dynamical-to-stellar mass ratio.

Lines in the top panel of Figure 9 show predicted behavior of the dynamical-to-stellar mass ratio for a galaxy with $q_{\rm int} = 0.25$ observed at different inclinations as given by Equation (9). We set $\beta = 0.7(1 - q_{\rm int})$ based on the results of Cappellari et al. (2007) and Emsellem et al. (2011) for nearby fast-rotating early-type galaxies. In the case of an oblate system, V/σ and anisotropy are related by (Binney & Tremaine 1987; Binney 2005)

$$\beta = 1 - \frac{1 + (V/\sigma)^2}{1 - \alpha(V/\sigma)^2} \left(\frac{W_{zz}}{W_{xx}}\right),$$

where α is a dimensionless number that quantifies the contribution of streaming motions to the line-of-sight velocity dispersion and (W_{zz}/W_{xx}) is a shape parameter related to the

---

Figure 9. Dynamical-to-stellar mass ratio as a function of $q_{\rm obs}$, the observed axis ratio. The top and bottom panels show results for two different estimates of the dynamical mass based on the virial theorem (i.e., Equations (4) and (5); top panel) or JAM models (bottom panel). Individual points are color-coded according to their Sérsic indices as indicated by the color bar at the top of the figure. In the top panel, thin gray lines show the expected dependence of the dynamical-to-stellar mass ratio on $q_{\rm obs}$ for an anisotropic model with $q_{\rm int} = 0.25$ and $\beta = 0.7(1 - q_{\rm int})$ following Equation (9). In the bottom panel, gray vertical lines show the effect of adopting a Gaussian prior on the intrinsic axis ratio such that $q_{\rm int} = N(0.25, 0.05)$. In both cases the apparent anticorrelation between dynamical-to-stellar mass ratio and $q_{\rm obs}$ is consistent with some portion of the population harboring significant rotational support.

---

intrinsic axis ratio q as

$$\left(\frac{W_{zz}}{W_{xx}}\right) = \frac{2(q\sqrt{1 - q^2} - q^2 \arccos q)}{\arccos q - q\sqrt{1 - q^2}}.$$

We adopt a value of $\alpha = 0.15$, which provides a good description for nearby galaxies (Cappellari et al. 2007). Models are offset in $M_{\rm vir}/M_{\rm SPS}$ to reflect a range of dynamical-to-stellar mass ratios. The predicted trends qualitatively reproduce the observed correlation between mass ratio and $q_{\rm obs}$, supporting previous statistical evidence of rotational support among a fraction of high-redshift quiescent galaxies (e.g., Belli et al. 2017).

In the bottom panel of Figure 9 the correlation between $q_{\rm obs}$ and dynamical-to-stellar mass ratio for MFL models is notably weaker than in the virial theorem case, both visually and as measured by Spearman ρ, though galaxies with higher $q_{\rm obs}$ still tend toward lower dynamical-to-stellar mass ratios. Unlike the virial theorem case, we can explicitly test the impact of intrinsic structure on our MFL mass estimates through application of a

---

[18] Belli et al. (2017) adopt the value of γ determined by Cappellari et al. (2013b), which relates the measured second moment σ_e to the circular velocity at $r_e^{\rm sma}$. Here we instead define V/σ in terms of the flux-weighted mean *within* r_e, so that all of $\sigma_{\rm obs}$, V, σ, and $(V/\sigma)_e$ are defined over the same aperture.

prior on q_{int} in our modeling.[19] The vertical lines in the bottom panel of Figure 9 show the effect of assuming that galaxies are intrinsically flat, with $q_{int} = 0.25$, as opposed to the default case where we adopt a uniform prior on q_{int}. For galaxies with low q_{obs}, the effect of assuming a different intrinsic structure is minimal, but for galaxies with $q_{obs} > 0.6$ $(i < 55°)$ the estimated dynamical mass can increase by as much as $\sim65\%$ (0.22 dex), with a median increase of $\sim15\%$ (0.06 dex). The resulting correlation between dynamical-to-stellar mass ratio and q_{obs} is also flatter, with $\rho = 0.03 \pm 0.06$ $(p = 0.95)$. Furthermore, assuming an intrinsically flat structure for these objects reduces the number of galaxies with dynamical masses significantly lower than their photometrically derived stellar mass M_{SPS}.

In summary, the data considered here support the conclusions of previous studies suggesting that rotational support is prevalent among quiescent galaxies at high redshift (e.g., Chang et al. 2013; Newman et al. 2015; Belli et al. 2017; Toft et al. 2017; Newman et al. 2018; Hill et al. 2019). While we expect rotational flattening to have a minimal impact on dynamical mass estimates for galaxies with $q_{obs} < 0.6$, galaxies with high q_{obs} can have their dynamical masses underestimated by 0.2 dex or more depending on their intrinsic structure (i.e., if they are intrinsically spherical vs. flattened systems viewed face-on). As mentioned in Section 4.1, such a discrepancy between intrinsic and assumed structure can at least partially explain those galaxies in our sample that have dynamical masses formally less than their photometrically derived stellar mass, though it may not be the only factor affecting this comparison. Finally, we note that enforcing an intrinsically flat structure for all galaxies in our sample (e.g., $q_{int} = 0.25$) shifts the results presented in Section 4.1.1 toward *lower* central dark matter fractions and cannot explain the apparent evolution of $f_{DM}[<r_e]$ without also appealing to significant changes in the stellar initial mass function (see Section 4.2).

4.2. The Normalization of the Stellar IMF at $1.4 \leqslant z \leqslant 2.1$

In the case that we include an explicit dark matter component in our dynamical models, then we obtain an independent estimate of the *stellar* dynamical mass, $M_{*,NFW}$, that can be used to diagnose changes in the normalization of the stellar IMF. A similar approach has been used to highlight possible IMF variation in low- and intermediate-redshift galaxies through the IMF offset parameter $\alpha \equiv M_{*,dyn}/M_{*,IMF}$, where $M_{*,IMF}$ is the stellar mass computed for some default IMF (e.g., Treu et al. 2010; Thomas et al. 2011; Cappellari et al. 2013b; Conroy et al. 2013; Dutton et al. 2013; Spiniello et al. 2014; Smith et al. 2015). In our case α is measured with respect to the Chabrier (2003) IMF used in our SPS models (i.e., $M_{*,IMF} \equiv M_{SPS}$). While in principle α does not rely on any assumptions about *how* the IMF varies, significant deviations from a Salpeter-like IMF above 1–2 $M_\odot$ are difficult to reconcile with observations of color and luminosity evolution for elliptical galaxies (e.g., Tinsley 1978; van Dokkum 2008). We therefore assume that any variation in the IMF occurs at stellar masses that contribute very little to the overall luminosity of the population, i.e., well below the MS turnoff,
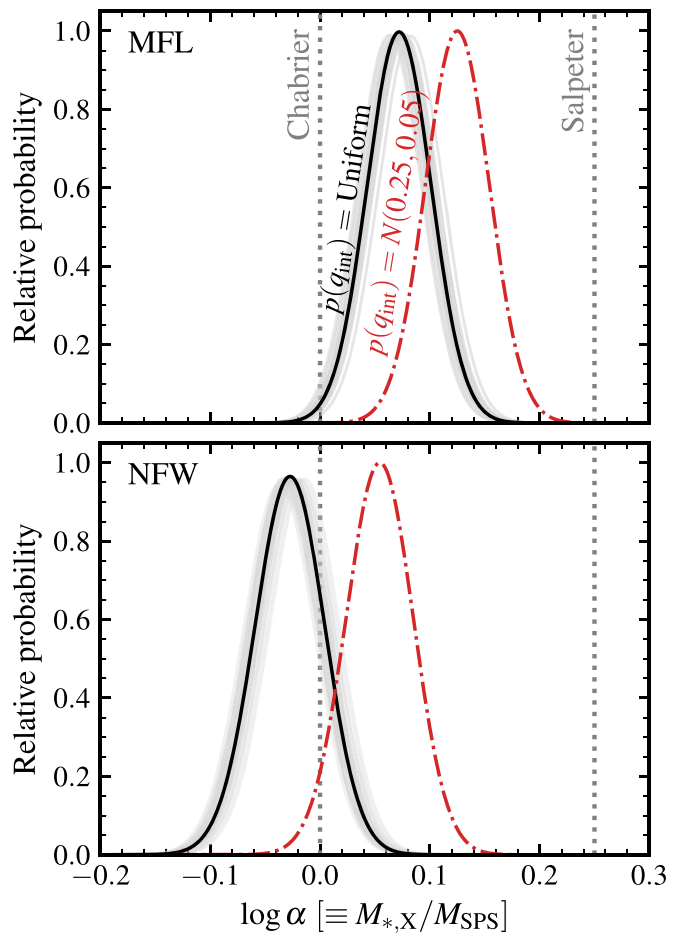
Figure 10. Constraints on the IMF normalization parameter α derived by combining the posterior PDFs of individual galaxies in our $1.4 < z < 2.1$ sample. The top panel shows the case where mass follows the observed light profile (MFL), while the bottom panel shows results when explicitly including an NFW-like dark matter halo. Solid curves are derived assuming a uniform prior on galaxies' intrinsic axis ratio q_{int} $(\equiv b/a)$, with light curves indicating variations derived from a jackknife analysis. Dot-dashed curves show the effect of assuming that high-redshift galaxies are intrinsically flat with an axis ratio of $q_{int} = 0.25 \pm 0.05$, consistent with the values derived by Chang et al. (2013). Vertical dotted lines indicate the expected values of α for different IMFs as indicated.

which is $<2M_\odot$ for stellar populations older than ~1 Gyr (the typical age for galaxies in our high-redshift sample; see, e.g., Mendel et al. 2015).

In Figure 10 we consider two limiting cases for the derivation of α: one where total mass follows the light profile and $f_{DM}[<r_e] = 0$ (top panel), and a second where we include a static NFW dark matter halo following the procedure outlined in Section 3.3 (bottom panel). In each case we show the combined constraint obtained from stacking individual posterior probability distribution functions (PDFs) for galaxies in our high-redshift sample.

We find that the high-redshift data prefer an overall normalization of the IMF that is lighter than reported for nearby early-type galaxies of a similar mass $(\log(M_*/M_\odot) \sim 11)$, which tend to favor Salpeter (1955) or heavier IMFs (Conroy & van Dokkum 2012; Conroy et al. 2013; Cappellari et al. 2013a; Li et al. 2017; but see also Smith et al. 2015). There is an offset between the MFL and NFW models such that models including an explicit dark matter halo predict $\log\alpha = -0.03 \pm 0.03$, consistent with a Chabrier IMF, while MFL models prefer a

---

[19] Functionally speaking, we enforce different intrinsic axis ratios by deprojecting our MGE models at inclination i given q_{obs} and q_{int} by inverting Equation (6).

slightly heavier IMF normalization with $\log \alpha = 0.07 \pm 0.03$. There is little evidence for the bottom-heavy IMFs that have been reported in the central regions ($\lesssim 1/8 r_e$) of massive nearby early-type galaxies (e.g., van Dokkum et al. 2017; Parikh et al. 2018), which one might expect if *all* quiescent galaxies seen at $z > 1$ are the seeds of local massive ellipticals. We will discuss this further in Section 5.2.

As highlighted by Section 4.1.2, systematic differences in galaxy structure can influence the derivation of dynamical masses, and by extension our inferences about the IMF. In order to estimate the magnitude of this effect, we recomputed the stacked α PDFs, imposing a Gaussian prior on $q_{\rm int} = N(0.25, 0.05)$ following the result of Chang et al. (2013); the results are shown as dotted–dashed lines in Figure 10. Assuming an intrinsically flattened structure *for all galaxies* results in a slightly heavier overall normalization of IMF, such that for the MFL (NFW) case $\log \alpha = 0.12 \pm 0.03$ (0.05 ± 0.03). It therefore seems unlikely that structural evolution *alone* can account for the apparent IMF differences between galaxies in our high-redshift sample and the cores of local early-type galaxies.

4.2.1. The Degeneracy between Central Dark Matter Fraction and IMF Normalization

One of the key assumptions in computing $M_{*,\rm NFW}$ is that the dark matter halo component is well represented by an NFW profile, with no accounting for the possible influence of baryons on the dark matter profile shape. However, if the timescale for galaxy formation is long compared to the halo dynamical time, then the halo is expected to contract adiabatically as a result of baryonic collapse (e.g., Blumenthal et al. 1986; Gnedin et al. 2004). Dutton et al. (2016) argue that the dark matter halo can contract or expand depending on the relative balance of inflows, outflows, and feedback (see also Lovell et al. 2018), suggesting that our assumption of a static halo may bias the derived values of $M_{*,\rm NFW}$ and, by extension, α. In this section we therefore explore a broader set of dynamical models that explicitly probe the effect of a variable dark matter halo response on our results.

In the case of spherical symmetry and circular dark matter particle orbits, the adiabatic invariant is given by $r M_{\rm tot}(r)$—where $M_{\rm tot}(r)$ is the total (baryonic plus dark matter) mass within radius r—so that $r_{\rm f}/r_{\rm i} = M_{\rm tot,i}(r_i)/M_{\rm tot,f}(r_f)$. Therefore, given an initial mass distribution $M_{\rm tot,i}(r)$ and a final baryonic mass profile $M_{\rm bar,f}(r)$, we can derive the final dark matter profile $M_{\rm DM,f}(r)$. Here we assume that the initial dark matter distribution is described by an NFW profile with mass and concentration parameter set by the scaling relations adopted in Section 3.3, and that the baryonic mass is distributed in the same way, i.e., $M_{\rm bar,i}(r) = f_{\rm bar} M_{\rm tot,i}(r)$, with $f_{\rm bar}$ set by the stellar-to-halo mass relation. The final baryonic profile $M_{\rm bar,f}(r)$ is given by the deprojected MGE luminosity density scaled to match the galaxy stellar mass. We note that this assumes that star formation is distributed throughout the halo and is therefore likely an upper limit on the expected contraction.

If we assume no shell crossing of the dark matter, then $M_{\rm DM,i}(r_i) = M_{\rm DM,f}(r_f)$, and the final mapping between r_f and r_i is given by

$$\Gamma = M_{\rm DM,i}(r_i)/[M_{\rm bar,f}(r_f) + (1 - f_{\rm bar})M_{\rm DM,i}(r_i)], \quad (10)$$
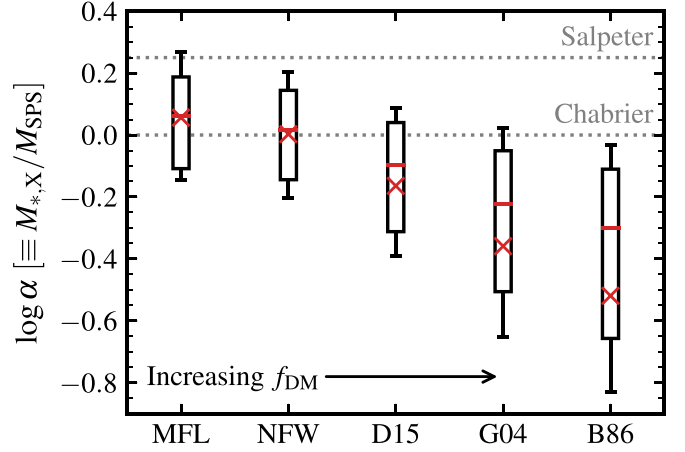
Figure 11. IMF normalization α ($\equiv M_{*,\rm dyn}/M_{\rm Chabrier}$) for different prescriptions describing the dark matter halo response to galaxy formation. Models are arranged from left to right in order of increasing $f_{\rm DM}[<r_e]$: MFL, NFW, Dutton et al. (2015), Gnedin et al. (2004), and Blumenthal et al. (1986). Red horizontal lines and crosses show the median and mean values, respectively. Boxes indicate the interquartile range, while error bars show the 16th/84th percentiles of the observational data. Prescriptions that predict stronger contraction of the dark matter halo lead to higher central dark matter fractions and correspondingly lighter stellar IMFs.

with $\Gamma = (r_f/r_i)^\nu$ following the generalized contraction formula suggested by Dutton et al. (2007). In this framework $\nu = 1$ reproduces the standard adiabatic contraction derived by Blumenthal et al. (1986), while $\nu = 0.8$ reproduces the modified contraction scenario described by Gnedin et al. (2004). $\nu = 0$ is equivalent to an unmodified NFW profile. We also include a milder model for the halo response derived from the NIHAO simulations discussed by Dutton et al. (2015), such that

$$r_{\rm f}/r_{\rm i} = 0.5 + 0.5(M_{\rm tot,i}/M_{\rm tot,f})^2. \quad (11)$$

For each contraction model we solve for the mapping between $r_{\rm f}$ and $r_{\rm i}$, and we use this modified dark matter profile as input to the JAM modeling procedure. While we do not explicitly include any models for halo *expansion* (i.e., $\nu < 0$), our default MFL models can be interpreted as the extreme case where dark matter is completely evacuated within r_e, setting an upper limit for the dynamical effects of an expanded halo.

In Figure 11 we show the distributions of α derived for these different models of halo response, along with the standard MFL and NFW cases described in the previous section. The expected trend of a decreasing stellar contribution—that is, a lighter overall IMF normalization—with increasingly strong halo contraction is clearly visible, with pure adiabatic contraction (e.g., Blumenthal et al. 1986) predicting stellar masses that are a factor of >3 lighter than those obtained assuming a Chabrier (2003) IMF. Even the mildest model for halo response, Dutton et al. (2015), predicts values of α that are $\sim 60\%$ lighter than Chabrier, and all of the contraction models considered here predict IMF normalizations that are lighter than observed or inferred for nearby stellar systems (e.g., Chabrier 2003; Bastian et al. 2010). Taken together, the results in Figure 11 suggest that any contraction of the dark matter halo due to gas inflow should be counterbalanced by equally violent outflows during the formation process.

5. Evolutionary Trends

The data presented in Figures 6–8 suggest an evolution of quiescent galaxy properties from high to low redshift, with high-redshift galaxies having on average lower central dark matter fractions and/or a lighter overall IMF compared to their low-redshift counterparts. These trends appear to persist even when considering only the oldest galaxies at low redshift. However, some care must be taken when making such comparisons, as individual galaxies are expected to evolve over the ~8 Gyr separating our two samples. The challenge of connecting progenitor and descendant populations is therefore that it relies on having at least some a priori knowledge of *how* this evolution proceeds. Here we consider two different methods for connecting galaxies across redshift based on either their number density or central stellar velocity dispersion.

The stochastic nature of dark matter halo assembly in a ΛCDM cosmology leads to a broad range of plausible descendants for any single galaxy at high redshift. This diversity of assembly histories can bias inferred evolutionary trends when considering samples selected at either a single fixed or evolving (median) cumulative number density (e.g., Mundy et al. 2015; Wellons & Torrey 2017). As an alternative, one can identify descendants based on the full number density PDF as opposed to a single value, which we do here using NDPredict[20] as described by Wellons & Torrey (2017). For each galaxy in our high-redshift sample, NDPredict provides an estimate of the likely descendant stellar mass distribution based on both expected evolution of the median number density *and its scatter*. This distribution is then used as a weight to select probable descendants from our low-redshift sample. We use the stellar mass functions published by Muzzin et al. (2013) to translate between stellar mass and number density at any given redshift, but we have verified that our results are insensitive to the particular choice of mass function.

As an alternative to the number density matching described above, we also construct a population of descendants matched at fixed central stellar velocity dispersion. Numerical studies have shown that central stellar velocity dispersion is relatively insensitive to assembly via dissipationless mergers (Hopkins et al. 2009b; Nipoti et al. 2012; Oser et al. 2012), which is expected to be the main growth channel for quiescent galaxies from high redshift. Hopkins et al. (2009b) argue that repeated dry mergers tend to decrease stellar velocity dispersions by at most 30%, as any increase in size leads to a corresponding increase in the dark matter content. To first order this assumption is consistent with the global trends shown in Figure 8. Based on these arguments, we identify likely descendants as those with σ_e within 0.05 dex of galaxies in our high-redshift sample, allowing for replacement—that is, the same low-redshift galaxies can be matched to multiple galaxies in our high-redshift sample.

5.1. Galaxy Evolution in Size, Stellar Mass, and σ_e

Assuming that high-redshift galaxies evolve to reproduce the typical properties of their matched descendants, we can use the relative difference in r_e^{sma}, M_{SPS}, and σ_e to study the evolutionary processes acting on galaxies over their lifetimes. In the context of ΛCDM, individual quiescent galaxies are expected to evolve from $z \sim 2$ to the present day through

---

Table 5
Inferred Evolution of Stellar Mass, Velocity Dispersion, and Size since $1.4 < z < 2.1$

Parameter	Evol. Number Density	Fixed σ_e
$\Delta \log(M_{SPS}/[M_\odot])$	0.24 ± 0.01	0.22 ± 0.03
$\Delta \log(\sigma_e/[\mathrm{km\ s^{-1}}])$	-0.05 ± 0.02	$\cdots$
$\Delta \log(r_e^{sma}/[\mathrm{kpc}])$	0.68 ± 0.02	0.63 ± 0.03

---

continued mergers after quenching. Bezanson et al. (2009) used scaling relations based on the virial theorem to argue that minor, gas-poor mergers can efficiently increase half-light sizes and decrease stellar velocity dispersions, a result that has been reiterated in a number of theoretical studies (e.g., Naab et al. 2009; Hopkins et al. 2009b; Oser et al. 2010, 2012; Hilz et al. 2013; but see also Nipoti et al. 2012). These predictions are supported by results that show that minor merging can explain the observed evolution of galaxy properties since $z \approx 2$ (e.g., Damjanov et al. 2011; van de Sande et al. 2013; Belli et al. 2014a, 2017).

For each high-redshift galaxy we compute the average difference between its properties and those of its matched descendants. In Table 5 we quote the median evolution of M_{SPS}, σ_e, and r_e^{sma} derived in this way for the full sample, with uncertainties estimated by jackknife resampling. There is little dependence of the inferred parameters on the method used to define descendant populations, with both the number density and velocity dispersion matched samples pointing toward a dramatic increase in size relative to stellar mass (a factor of ~4.5 in r_e^{sma} compared to only ~1.5 in M_{SPS}). For samples matched on number density we can additionally infer the evolution of σ_e, which appears to decrease by only 12% on average from $z > 1.4$ to the present day, in good agreement with the predictions of numerical simulations (e.g., Hopkins et al. 2009b; Hilz et al. 2012).

Figure 12 shows the pairwise distributions of M_{SPS}, σ_e, and r_e^{sma}, where the observed evolution can be compared to simple energetic arguments for major and minor merging as in Bezanson et al. (2009, shown as inset coordinate arrows in each panel). The factor of 4–5 evolution in r_e^{sma} discussed above is immediately apparent, as is the comparably milder evolution in M_{SPS} and σ_e. These results are quantitatively similar regardless of how we choose to identify low-redshift descendants (i.e., evolving number density or fixed σ_e), and they are consistent with the simple predictions for evolution by predominantly minor merging, where $\Delta r_e \propto \Delta M_*^\alpha$ with $\alpha \approx 2$ (cf. $\alpha \approx 1$ for major mergers). In contrast, the distributions of both r_e^{sma} versus σ_e and σ_e versus M_{SPS} suggest a more complicated interpretation, whereby those galaxies with $\log(M_{SPS}/M_\odot) > 11.2$ or $\log \sigma_e \gtrsim 2.4$ evolve relatively *more* in M_{SPS} and/or σ_e than those with lower masses or velocity dispersions. The inclusion of velocity dispersion also complicates our otherwise straightforward interpretation of M_{SPS} and r_e^{sma} evolution in terms of minor merging, although we note that more detailed numerical simulations find a range of behaviors depending on the properties of galaxies' host dark matter halos (e.g., Oser et al. 2012; Nipoti et al. 2012; Hilz et al. 2012, 2013). We will discuss these trends further in Section 5.2.

In order to investigate these differences further, in Figure 13 we show the same galaxy samples as in Figure 12, however this
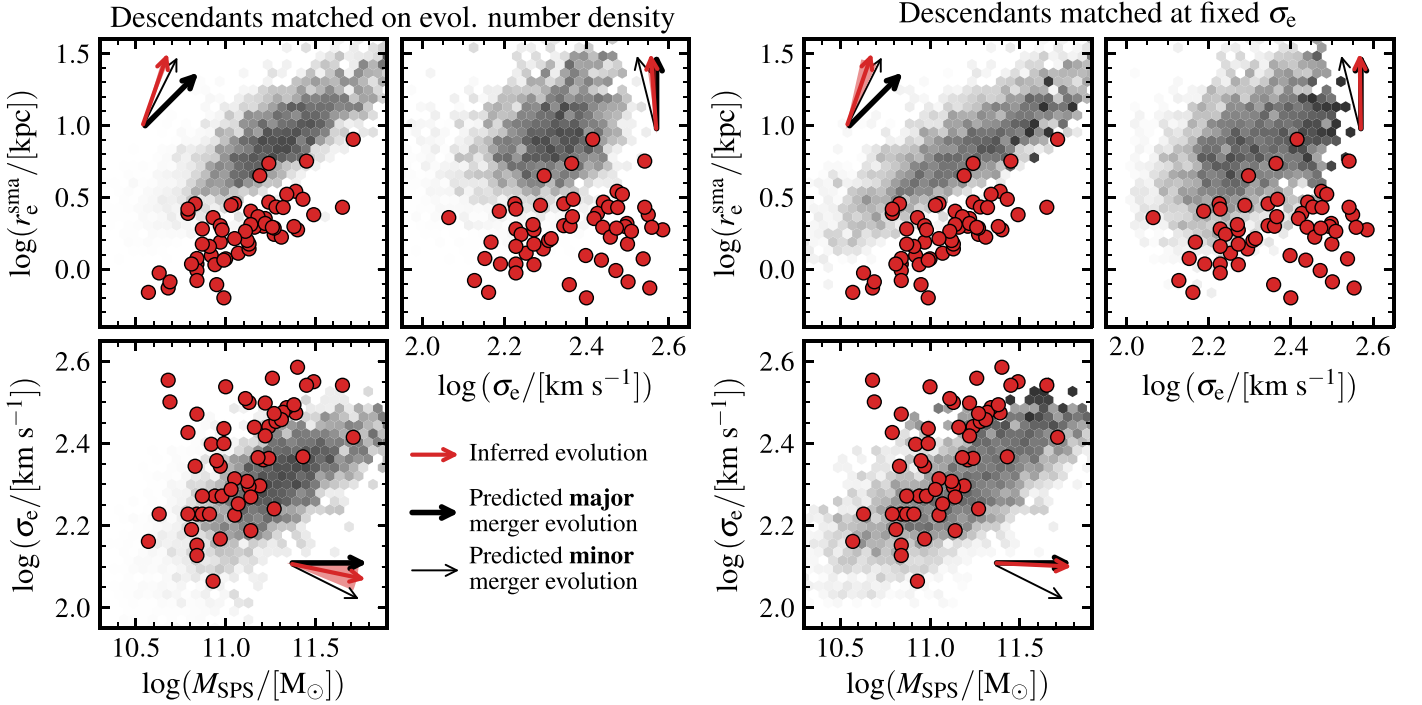

Figure 12. Evolutionary trends in half-light size, stellar velocity dispersion, and stellar mass. Circles (red) show measurements for individual quiescent galaxies at $1.4 < z < 2.1$. Shading shows the distribution of $z \approx 0$ descendants matched on either evolving number density or stellar velocity dispersion (left and right panels, respectively) as described in Section 5. In each panel, black arrows indicate the predicted evolution for each pair of parameters based on simple energetic arguments (e.g., Bezanson et al. 2009). Red arrows and shading show the measured trends and their $\pm 1\sigma$ uncertainties derived from jackknife resampling. Note that by definition σ_e does not evolve for descendants matched on σ_e.
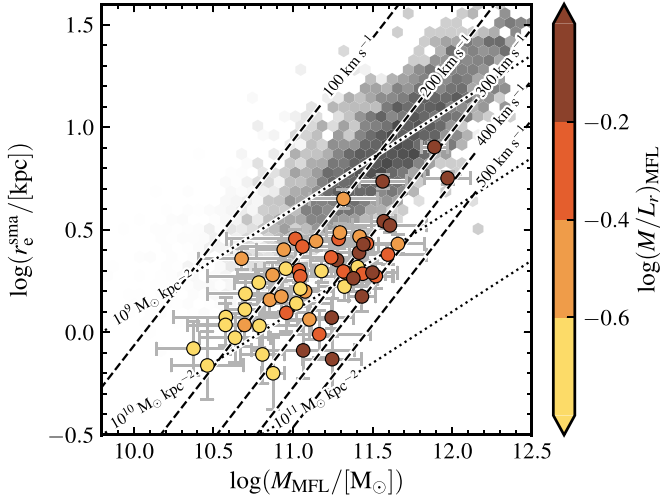
Figure 13. Half-light size vs. JAM-derived dynamical mass for high-redshift galaxies and their (evolving number density matched) descendants. High-redshift galaxies, shown as circles, are colored according to their dynamical mass-to-light ratio as indicated by the scale on the right. Shading shows the distribution of $z \approx 0$ descendants derived following Section 5. Dashed lines show tracks of constant $\sigma_e = 100$, 200, 300, 400, and 500 km s^{-1} assuming the scalar virial relation (Equation (4)) with $\kappa = 5$. Dotted lines instead show tracks of constant stellar surface density, with $\Sigma_e = 10^9$, 10^{10}, and 10^{11} $M_\odot$ kpc^{-2}. Galaxies with high dynamical mass-to-light ratios are relatively well aligned with tracks of constant velocity dispersion, while lower-M/L galaxies follow more closely tracks of constant stellar mass density.

time in terms of r_e^{sma} and dynamical mass, often referred to as the mass plane. Galaxies have additionally been color-coded according to their dynamical mass-to-light ratio, M/L. Given the low dark matter fractions inferred for a majority of galaxies

in our high-redshift sample, we expect that M/L_{MFL} is primarily a tracer of stellar population variations. The mass plane provides a useful parameter space within which to understand systematic variations of galaxy properties and their evolution and has been used extensively to study the interrelationship between dynamics, structural properties, and stellar populations (e.g., Cappellari et al. 2013a; McDermid et al. 2015; Cappellari 2016; Belli et al. 2017; Scott et al. 2017).

We find that those galaxies with the highest M/L follow lines of roughly constant σ_e (dashed lines in Figure 13), while at lower M/L galaxies more closely follow lines of constant stellar mass surface density, Σ_e (dotted lines in Figure 13). Under the assumption that variations in $(M/L)_{MFL}$ are primarily driven by changes in the mean stellar age, these differences are in stark contrast to observational results at low redshift, where σ_e is by far the best predictor of galaxy stellar populations (e.g., Graves et al. 2009; McDermid et al. 2015; Scott et al. 2017). Nevertheless, numerous studies have shown that stellar surface density is a strong predictor of galaxy quiescence at low redshift (e.g., Cheung et al. 2012; Fang et al. 2013; Woo et al. 2015; but see also Wake et al. 2012), and high densities appear to be a necessary (if not sufficient) condition for quenching at high redshift (e.g., Barro et al. 2013; van Dokkum et al. 2015; Barro et al. 2017). It is therefore unsurprising to see such a correlation borne out in our M/L measurements: at any given epoch galaxies are likely added to the quiescent population as a function of their stellar surface density. Subsequent merger-driven assembly will then tend to move galaxies along lines of constant velocity dispersion (or steeper; see, e.g., Nipoti et al. 2012; Hilz et al. 2013).
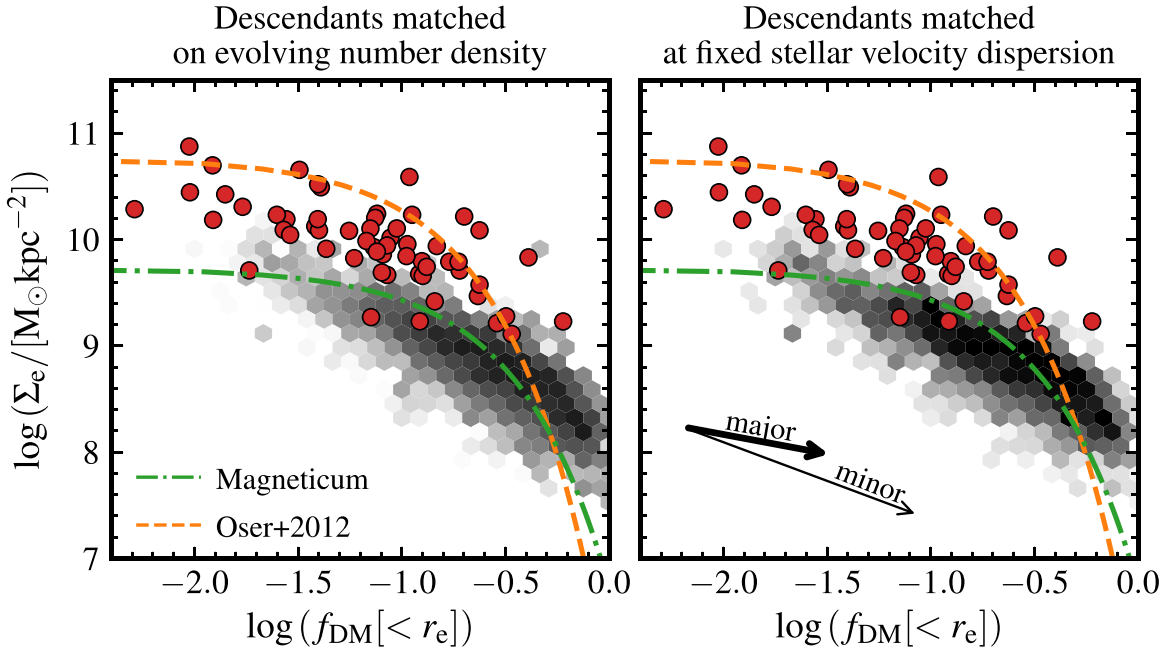
Figure 14. Stellar mass surface density, $\Sigma_{\rm e}$, as a function of central dark matter fraction for high- and low-redshift galaxies. Filled circles indicate measurements for quiescent galaxies at $1.4 < z < 2.1$, while the background shading shows the distribution of $z \approx 0$ descendants matched by either evolving number density (left panel) or central stellar velocity dispersion (right panel; see Section 5 for details). Dashed and dotted–dashed curves show the trends derived by Remus et al. (2017) for simulated early-type galaxies from Oser et al. (2012) and Magneticum (Hirschmann et al. 2014, Dolag et al., in preparation). In the right panel, arrows show the predicted trajectories for evolution driven by major and minor mergers described in Section 5.1.1. In both cases our simple model fails to match the required evolution in $\Sigma_{\rm e}$ and $f_{\rm DM}[<r_{\rm e}]$.

5.1.1. Size Evolution as the Primary Driver of f$_{DM}$

We showed in Section 4.1.1 and Figure 8 that, at fixed stellar mass, $f_{\rm DM}[<r_{\rm e}]$ increases by a factor of ~ 2 from high redshift to the present day. However, based on the matched progenitor and descendant samples described in Section 5, we find that the growth of $f_{\rm DM}[<r_{\rm e}]$ for individual galaxies is likely even larger, reaching 0.64 ± 0.05 dex $(0.52 \pm 0.07$ dex) on average for progenitors and descendants matched on evolving stellar mass (fixed $\sigma_{\rm e}$).

Just as major and minor mergers are expected to have different effects on a given galaxy's evolution in size, stellar mass, and stellar velocity dispersion (e.g., Figure 13), they also have a distinct influence on the evolution of $f_{\rm DM}$. Hilz et al. (2012, 2013) show that minor mergers can dramatically increase $f_{\rm DM}[<r_{\rm e}]$, so that a factor of 2 growth in stellar mass can nearly double the central dark matter fraction. In contrast, a single equal-mass (major) merger may only increase $f_{\rm DM}[<r_{\rm e}]$ by 50%. This difference is driven by the relatively efficient growth of sizes in minor compared to major mergers; for an NFW-like halo the dark matter mass within a given (small) radius r scales roughly with the virial mass of the halo as $M_{\rm DM}(r) \propto M_{\rm vir} r^2$ (Boylan-Kolchin et al. 2005). In an equal-mass merger where both $r_{\rm e}$ and the stellar mass within $r_{\rm e}$ double, the dark matter mass within $r_{\rm e}$ can increase by up to a factor of 8 (assuming that the stellar mass and halo mass increase by a similar amount). By comparison, a similar doubling of stellar mass through multiple minor mergers can increase $r_{\rm e}$ by a factor of 4 and the enclosed dark matter mass by a factor of $\gtrsim 30$.

These differences are particularly apparent when viewed in terms of stellar mass surface density, $\Sigma_{\rm e}$ ($\equiv M_{\rm SPS}/2\pi r_{\rm e}^2$), and $f_{\rm DM}[<r_{\rm e}]$, which we show in Figure 14. We find that $\Sigma_{\rm e}$ and $f_{\rm DM}[<r_{\rm e}]$ are anticorrelated, in the sense that those galaxies

with the highest stellar mass surface densities have the lowest $f_{\rm DM}[<r_{\rm e}]$; a similar anticorrelation has been demonstrated previously for both late- and early-type galaxies (e.g., McGaugh 2005; Sonnenfeld et al. 2015). Figure 14 additionally shows that $\Sigma_{\rm e}$ decreases by ~ 1 dex on average from $z > 1.5$ to $z \approx 0$, which is a direct consequence of the size and stellar mass evolution discussed in Section 5.1 (see also Table 5 and Figure 12). This can be compared to the results of Remus et al. (2017), who show that simulated galaxies follow well-defined tracks in $\Sigma_{\rm e}$–$f_{\rm DM}$ regardless of redshift (shown as dashed and dotted–dashed lines in Figure 14). Nevertheless, the simulation results roughly reproduce the observational trends at any given epoch, with the primary difference between the two simulations discussed by Remus et al. (2017) being their treatment of black hole feedback.

We can use simple scaling relations to better understand galaxies' expected evolution in Figure 14 given various assumptions. Following Figure 12, we adopt a simple model for size growth during mergers such that $\Delta r_{\rm e} \propto \Delta M_*^\alpha$, with $\alpha = 1$ or 2 for major and minor mergers, respectively. We additionally assume that the enclosed dark matter mass scales with the total (virial) mass of the halo and radius as $M_{\rm DM}(r) \propto M_{\rm vir} r^2$, and that the stellar and halo mass grow at the same rate ($\Delta \log M_{\rm vir} = \Delta \log M_*$). Arrows in the right panel of Figure 14 show the predicted evolution for a galaxy doubling its stellar mass through either a single major merger or successive 10:1 (minor) mergers. As expected, the efficient size growth associated with minor mergers in our toy model drives rapid evolution in both $\Sigma_{\rm e}$ and $f_{\rm DM}[<r_{\rm e}]$. However, while minor mergers can explain the observed decrease in $\Sigma_{\rm e}$ from high to low redshift, our toy model overpredicts the increase of $f_{\rm DM}[<r_{\rm e}]$. Tortora et al. (2018) report a similar result for galaxies at $z \approx 0.7$ and suggest that allowing for variation in the stellar-to-halo mass ratio of accreted galaxies
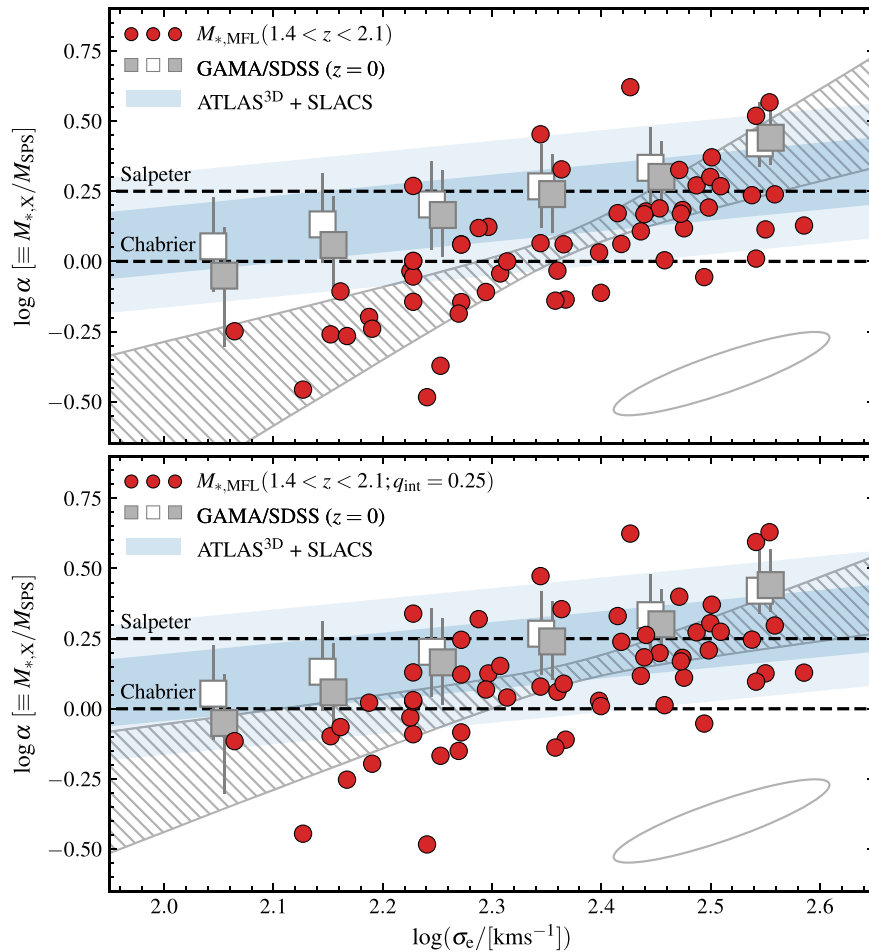
Figure 15. IMF offset parameter, α, as a function of stellar velocity dispersion for different samples. In the top panel, filled circles show data at $1.4 < z < 2.1$ assuming MFL and a uniform prior on the intrinsic axis ratio, q_{int}. In the bottom panel we show the same data, but adopting a prior such that $q_{int} = 0.25$. The error ellipse in the lower right corner shows the typical uncertainty for individual galaxies, while the hatched regions illustrate the 16th/84th percentile confidence interval based on a linear fit to the high-redshift data. Large squares represent the median α derived for GAMA/SDSS data matched in either evolving number density (filled squares) or stellar velocity dispersion (open squares) as described in Section 5. Error bars indicate the 16th/84th percentile uncertainties on the binned values. Shaded bands show the α–σ_e correlation based on a joint analysis of ATLAS^{3D} and SLACS data by Posacki et al. (2015), with dark and light light regions indicating the 1σ and 2σ bounds, respectively. There is good agreement between the matched GAMA/SDSS data and derived best fit from Posacki et al., both of which show evidence for a positive correlation between α and σ_e. High-redshift galaxies show evidence for a similar correlation, albeit with a steeper slope such that objects with $\log \sigma_e \gtrsim 2.45$ have α and σ_e comparable to low-redshift galaxies, while at lower σ_e further evolution is required to reproduce the low-redshift correlation.

(i.e., $\Delta \log M_{vir} \neq \Delta \log M_*$) can help to lessen the tension between predicted and observed evolution in f_{DM}. This is particularly likely for the massive quiescent galaxies in our high-redshift sample, which are expected to trace the most massive halos at their respective redshifts (e.g., Lin et al. 2019).

5.2. Evolution of the IMF at Fixed σ_e

In Section 4.2 we showed that the kinematics of massive quiescent galaxies at $z > 1.5$ are consistent with an MW-like IMF on average. In contrast, a number of studies based on stellar absorption features, lensing, and dynamical modeling have shown that a Salpeter or even heavier IMF may be more typical in the inner 0.1–$0.2\ R_e$ of present-day early-type galaxies (e.g., Thomas et al. 2011; Cappellari et al. 2012, 2013a; Conroy & van Dokkum 2012; Spiniello et al. 2014; Conroy et al. 2017; van Dokkum et al. 2017; but see also Smith et al. 2015), suggesting some tension between our high-redshift results and those in the nearby universe.

Figure 15 shows a comparison of the IMF offset parameter, α, as a function of stellar velocity dispersion, σ_e, at different redshifts. We include in this figure constraints derived from

low-redshift galaxies in the ATLAS^{3D} and SLACS samples by Posacki et al. (2015), as well as binned results for the matched GAMA/SDSS samples described in Section 5. This comparison shows that the apparent offset in mean α between high and low redshift depends on σ_e, such that galaxies with the highest velocity dispersions generally have α values consistent with observations at low redshift, while lower velocity dispersion galaxies are offset toward lower α—that is, toward a "lighter" IMF—at fixed stellar velocity dispersion. Assuming an intrinsically flat axis ratio for high-redshift galaxies (e.g., $q_{int} = 0.25$; bottom panel of Figure 15) reduces the apparent offset of α at low σ_e, though the qualitative trend remains unchanged. We note that this offset is unlikely to be the result of (correlated) uncertainties in σ_e and α, shown by the error ellipse in the lower right corner of Figure 15, which tend to scatter galaxies along the low-redshift α–σ_e correlation rather than away from it. It may be that the differential evolution seen between high- and low-σ_e galaxies is tied to the overall buildup of the velocity dispersion function (VDF) over time. Bezanson et al. (2012) show that galaxies with $\log \sigma_e \gtrsim 2.4$ form early and their number density has changed little since at least

$z \approx 1.5$, while the population at lower σ_e evolves significantly owing to the addition of newly quenched galaxies. A more detailed dissection of the α–σ_e relation in the context of galaxy SFHs will be the subject of future work.

An interesting implication of Figure 15 is that the α–σ_e relation found at low redshift is established early on, and any scenario invoked to explain quiescent galaxies' subsequent evolution in $r_e^{\rm sma}$, σ_e, and $M_{\rm SPS}$ (e.g., Figure 12) should largely preserve the underlying correlation; this is especially true at the highest stellar velocity dispersions. Using high signal-to-noise ratio, long-slit spectra of six nearby early-type galaxies, van Dokkum et al. (2017) found evidence for strong radial gradients in α that increased from MW-like at $R > 0.4R_e$ to Salpeter or heavier at $0.1R_e$ (see also Martín-Navarro et al. 2015b; Lyubenova et al. 2016). Under the assumption that gas-poor mergers primarily contribute to the buildup of an extended stellar envelope (e.g., Hopkins et al. 2009a; van Dokkum et al. 2010; Karademir et al. 2019), then we would expect that the remnants of our high-redshift sample should survive in the cores of massive nearby ellipticals. Indeed, the typical half-light sizes of galaxies in our sample (1–1.5 kpc) are comparable to the physical extent over which the IMF is found to vary significantly in van Dokkum et al. (2017).

Finally, although we have tried to be comprehensive in our measurement of stellar masses, we cannot rule out that at least some part of the evolutionary trends implied by Figure 15 could be the result of redshift-dependent systematic uncertainties in our derivation of $M_{\rm SPS}$. These could be, for example, due to our adopted SFH or stellar population synthesis models (e.g., Pforr et al. 2012; Leja et al. 2019). Based on the analysis of mock galaxy spectra generated from semianalytic models, Pforr et al. (2012) showed that stellar masses for passive galaxies can be recovered to better than ~ 0.05 dex for a wide range of SFHs, suggesting that our results are unlikely to be due *entirely* to the details of our SFH modeling. Furthermore, while Pforr et al. (2012) show that mismatches in metallicity can lead to systematic offsets of up to 0.2–0.3 dex, direct observations at $z > 1.5$ support our adoption of a solar-metallicity template library (e.g., Onodera et al. 2015; Kriek et al. 2016).

6. Conclusions

We present an analysis of 58 massive quiescent galaxies at $1.4 < z < 2.1$ with measured stellar velocity dispersions and deep near-infrared HST imaging. We use these data to study the evolution of dynamical masses and the dynamical-to-stellar mass ratio, which are sensitive to the central dark matter fraction and normalization of the stellar IMF. We find the following:

(i) The median dynamical-to-stellar mass ratio of quiescent galaxies is lower by ~ 0.2 dex at $1.4 < z < 2.1$ compared to low redshift. In Figures 6 and 7 we showed that this offset appears to be independent of the method used to derive dynamical masses (e.g., Jeans models vs. virial mass estimates). The observed evolution is consistent with a decrease in the fraction (by mass) of dark matter within the galaxy effective radius, $f_{\rm DM}[<r_e]$, which is lower by a factor of ~ 2 in our high-redshift sample compared to nearby galaxies in the SDSS/GAMA surveys ($f_{\rm DM}[<r_e] = 6.6\% \pm 1.0\%$ at $z \approx 1.8$ vs. $16.3\% \pm 0.3\%$ at $z = 0$) at fixed stellar mass. These measurements appear consistent with recent results based

on the rotation curves of high-redshift star-forming galaxies. Based on the matching of progenitor and descendant populations, we argue in Section 5.1.1 that the evolution of *individual* galaxies is likely even larger, with $f_{\rm DM}[<r_e]$ increasing by a factor of 4–5 from high redshift to the present day.

(ii) Under the assumption that central dark matter fractions are intrinsically low in high-redshift galaxies, the dynamical-to-stellar mass ratio can be used as a probe of the stellar IMF. For MFL models, we find that high-redshift data are consistent with a Kroupa-like IMF on average, while models including an explicit NFW dark matter halo are consistent with a Chabrier IMF (see Figure 10). We find a correlation between stellar velocity dispersion and the IMF offset parameter, α, at high redshift that is consistent with low-redshift data, suggesting that any subsequent evolution should act to preserve this underlying correlation (see Figure 15). We argue that minor mergers are the most likely drivers of galaxies' growth in $r_e^{\rm sma}$ and $M_{\rm SPS}$, as they primarily add material at large radii while preserving the stellar populations in the inner regions.

(iii) Simple models for the contraction of dark matter halos in response to baryonic collapse predict high central dark matter fractions. In the most extreme case of pure adiabatic contraction, such models require IMF normalizations a factor of ~ 3 *lighter* than Chabrier to explain the observed kinematics (see Figure 11). Significant contraction of the dark matter halo is difficult to accommodate given current observational constraints *unless* other baryonic process (e.g., outflows or AGN feedback) act to reduce the central dark matter content after collapse.

(iv) A comparison of kinematics and structural properties between high-redshift galaxies and their likely descendants at low redshift supports minor merging as the dominant evolutionary pathway after quenching; however, those galaxies with the highest stellar masses and/or stellar velocity dispersions appear to evolve relatively more than lower-mass/dispersion objects. This separation is apparent for descendants matched both on evolving number density and at fixed σ_e (see Figure 12).

(v) In the two-dimensional parameter space of size and dynamical mass—the so-called "mass plane" (Figure 13)—high-redshift galaxies both are smaller and have lower dynamical masses than their low-redshift descendants. Separating the galaxy population in terms of total mass-to-light ratio (M/L), which we interpret here as a proxy for stellar population age, galaxies with the highest M/L follow lines of roughly constant σ_e, while those with lower M/L follow more closely lines of constant stellar surface density. We interpret these differences as being driven by two separate phases of passive galaxy formation, whereby galaxies first quench as a strong function of their stellar mass surface density and their subsequent evolution on the mass plane is driven by minor merging.

Taken together, our results point toward an evolutionary scenario for massive quiescent galaxies that sees their formation occurring rapidly at $z > 2$. Subsequent quenching of star formation appears to preserve the disk-like structural and kinematic signatures associated with massive star-forming

galaxies at those redshifts (e.g., Wuyts et al. 2011; Wisnioski et al. 2015, 2019), as well as their dark matter properties (e.g., Wuyts et al. 2016; Genzel et al. 2017, 2020; Lang et al. 2017). Subsequent evolution through continued merging is then required to transform both their structural and kinematics properties to reproduce the massive, predominantly slowly rotating galaxies that constitute their likely descendants at $z = 0$ (e.g., Veale et al. 2017). Based on galaxies' evolution in size, stellar mass, stellar velocity dispersion, and central dark matter fraction, it appears that the most likely channel for this evolution is through the accretion of relatively lower mass galaxies (i.e., minor mergers). While these results are based on the best currently available kinematic and photometric data, future spectroscopic observations with the James Webb Space Telescope and 30+ m ground-based observatories will enable systematic surveys of high-redshift stellar kinematics.

Appendix A
Additional Tests of Velocity Dispersion Measurements

In this appendix we explore possible systematics in our derived velocity dispersions stemming from assumptions made in our spectrophotometric modeling approach. There are two effects in particular that we are concerned with: the impact of assuming a fixed solar metallicity for our SED template library, and the effect of our adopted (parametric) SFH.

A.1. Metallicity Effects

While the majority of low-redshift massive early-type galaxies are consistent with solar metallicity or higher (e.g.,
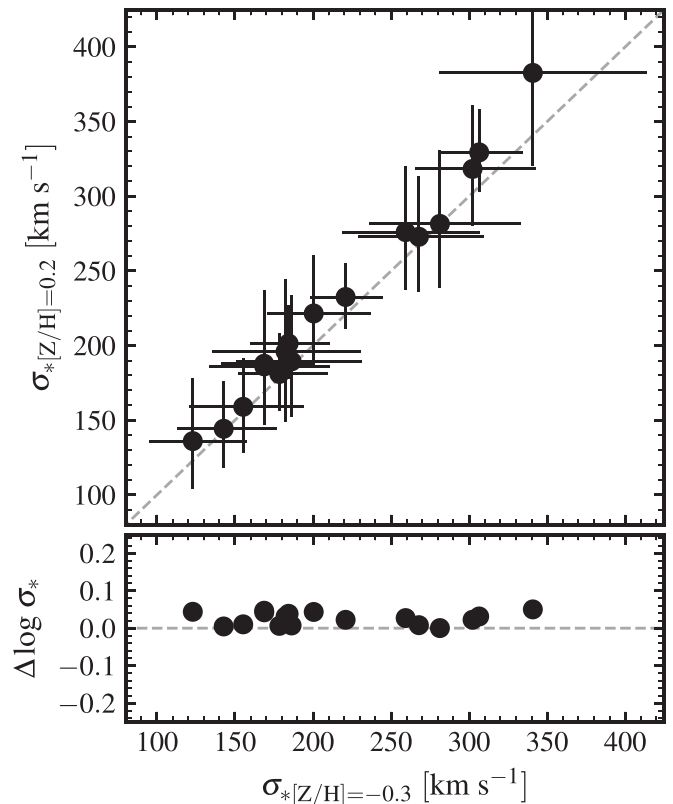
Figure A1. Effects of changing template metallicity on the derived stellar velocity dispersion. In this case we consider the effects of changing the assumed metallicity by a factor of ~3, from $[Z/H] = -0.3$ to 0.2 dex. The overall impact is small, with the mean dispersion increasing by less than 2% when going from low to high metallicity.
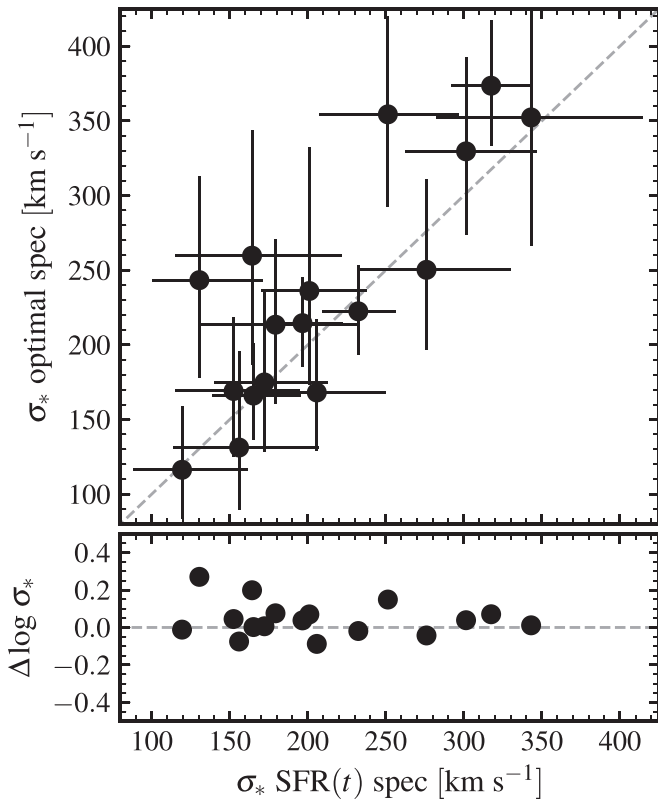
Figure A2. Impact of using "optimal" vs. parametric template models on the derived stellar velocity dispersion. The optimal templates are derived from a linear combination of single stellar population models, while the parametric models assume an explicit description of star formation rate as a function of cosmic time as given by Equation (2).

Gallazzi et al. 2006; Thomas et al. 2010), the picture at high redshift is still uncertain, with different results suggesting variations in total metallicity of up to 0.8 dex (e.g., Lonoce et al. 2015; Onodera et al. 2015; Kriek et al. 2016; Morishita et al. 2018). Given our adoption of a fixed solar metallicity for our template grid, it is therefore worthwhile to investigate the possible impact of this assumption on our derived velocity dispersions. In Figure A1 we show a comparison of stellar velocity dispersions obtained with templates a factor of $\sim$2 higher or lower in metallicity. While there is a clear systematic shift in the resulting values of σ_*, the offsets are of order a few percent, significantly smaller than the typical 20% uncertainties on our measurements of σ_*, and are therefore unlikely to bias our results.

A.2. Star Formation History Effects

While the adoption of a parametric SFH to describe the observed photometry is common practice, the use of such

models in performing kinematic measurements is less common and bears further investigation. In brief, we modified the Markov Chain Monte Carlo (MCMC) fitting code described in Section 2.1.1 to construct the best-fit template from a linear combination of simple stellar population templates, as opposed to the parametric SFHs adopted previously. This approach mimics the internal fitting procedure adopted in well-known fitting codes such as pPXF (Cappellari 2002), while at the same time providing improved handling of low signal-to-noise data thanks to the MCMC sampling of the line-of-sight velocity distribution. We show the results of this refitting in Figure A2.

There is significant scatter between the two velocity dispersion estimates, but in nearly all cases the derived σ_* values are consistent within uncertainties. The values derived in our default modeling are larger by $\sim$0.06 dex on average; however, this does not significantly affect our results.

Appendix B
Comparisons with Previous Structural Parameter Measurements

In this appendix we present a comparison of the galaxy structural parameters derived here with those available in the literature (where available).

B.1. High Redshift

The comparison of apparent magnitude, Sérsic index, and size for galaxies in our high-redshift sample is shown in Figure B1. For most galaxies these quantities were taken from van der Wel et al. (2014), which are based on the same CANDELS/3D-HST WFC3/F160W imaging. As discussed in Section 2.2, several galaxies from Bezanson et al. (2013a) and van de Sande et al. (2013) fall outside the CANDELS/3D-HST footprint, and there our measurements are based on pipeline-processed data retrieved from the HLA. In most cases the agreement between different measurements is excellent.

B.2. Low Redshift

Figures B2 and B3 show a comparison of apparent r-band magnitude, Sérsic index, and size for galaxies in our low-redshift SDSS/GAMA sample with measurements from Meert et al. (2015) and Simard et al. (2011), respectively. While there is generally good agreement between measurements in the different structural catalogs, there is significant scatter driven by differences in approach used to mask/model neighboring objects, the size of fitted images, and the method for measuring the sky background. In most cases any systematic biases are relatively small, $\lesssim$20%, and do not affect the conclusions of this work.
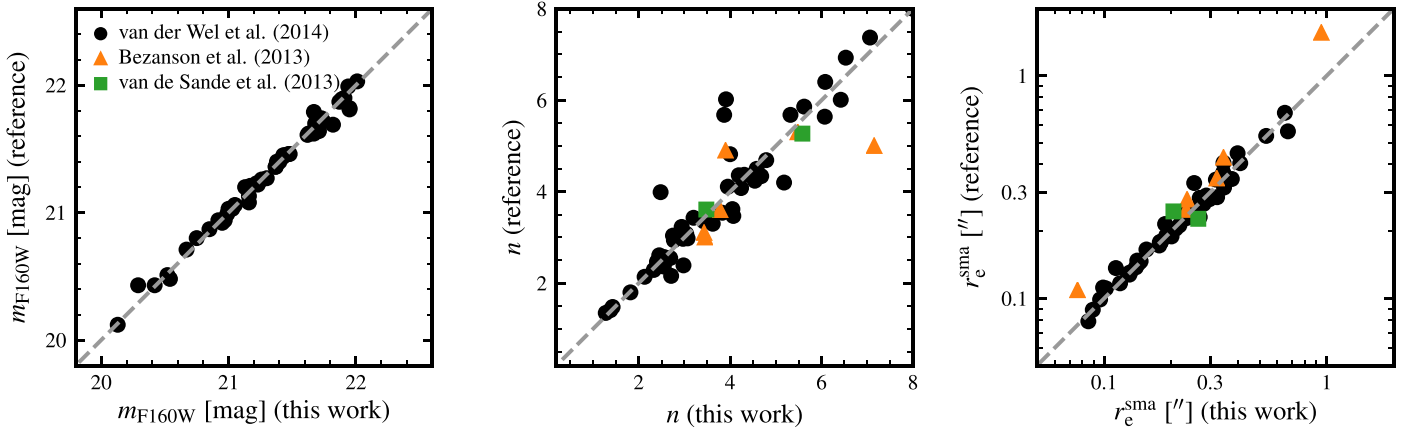
Figure B1. Comparison of total magnitude m_r, Sérsic index n, and semimajor axis size r_e^{sma} for galaxies in our high-redshift sample from different sources.
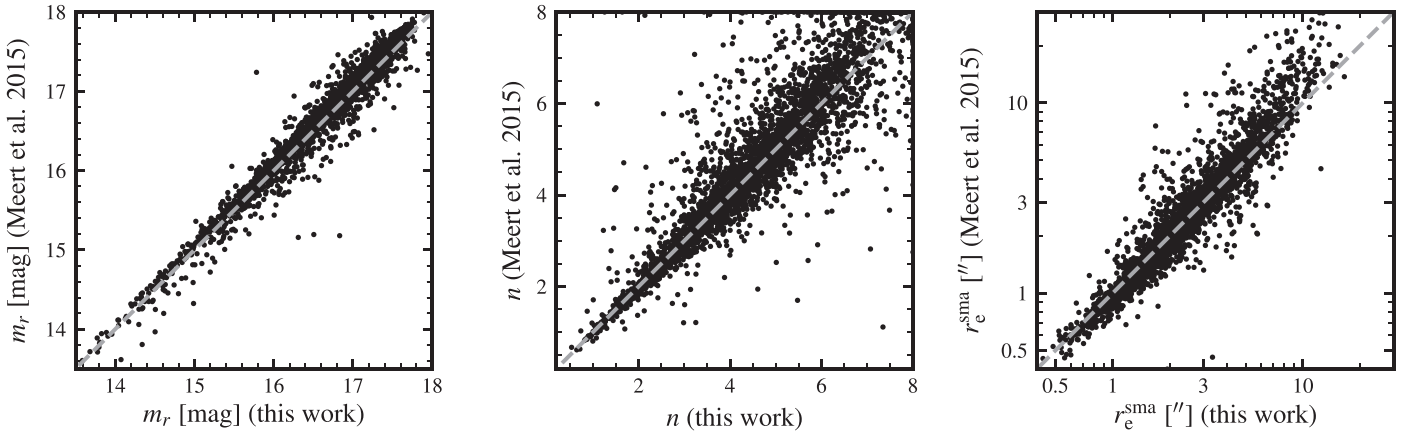

Figure B2. Comparison of total magnitude m_r, Sérsic index n, and semimajor axis size r_e^{sma} for low-redshift galaxies derived in this work vs. Meert et al. (2015).
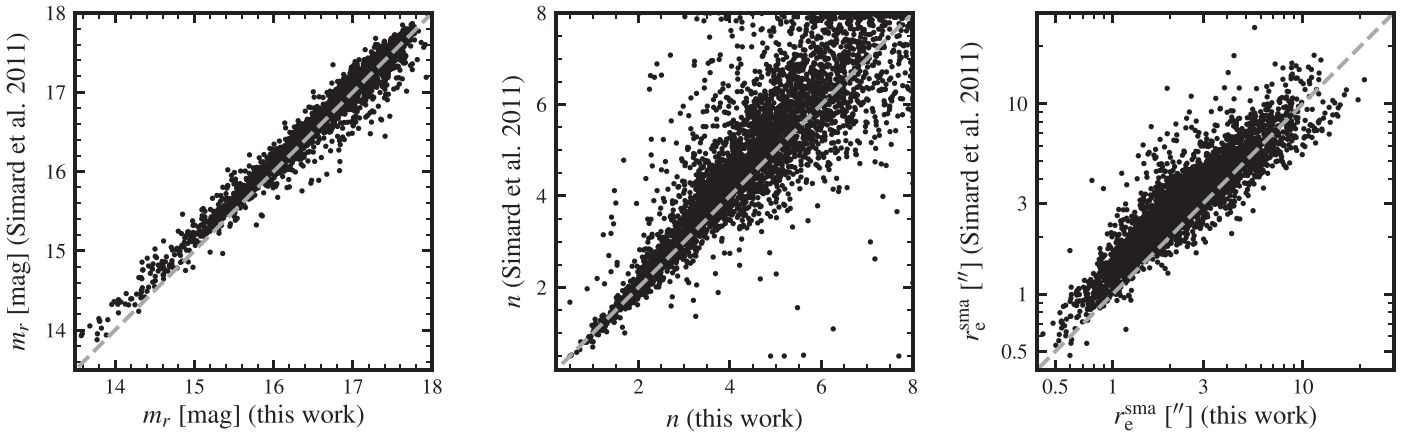

Figure B3. Comparison of total magnitude m_r, Sérsic index n, and semimajor axis size r_e^{sma} for low-redshift galaxies derived in this work vs. Simard et al. (2011).

<div align="center">

Appendix C
Photometric Fits for High-redshift Galaxies

</div>

In Figure C1 we show the Sérsic and MGE fits derived for each galaxy in the high-redshift sample.
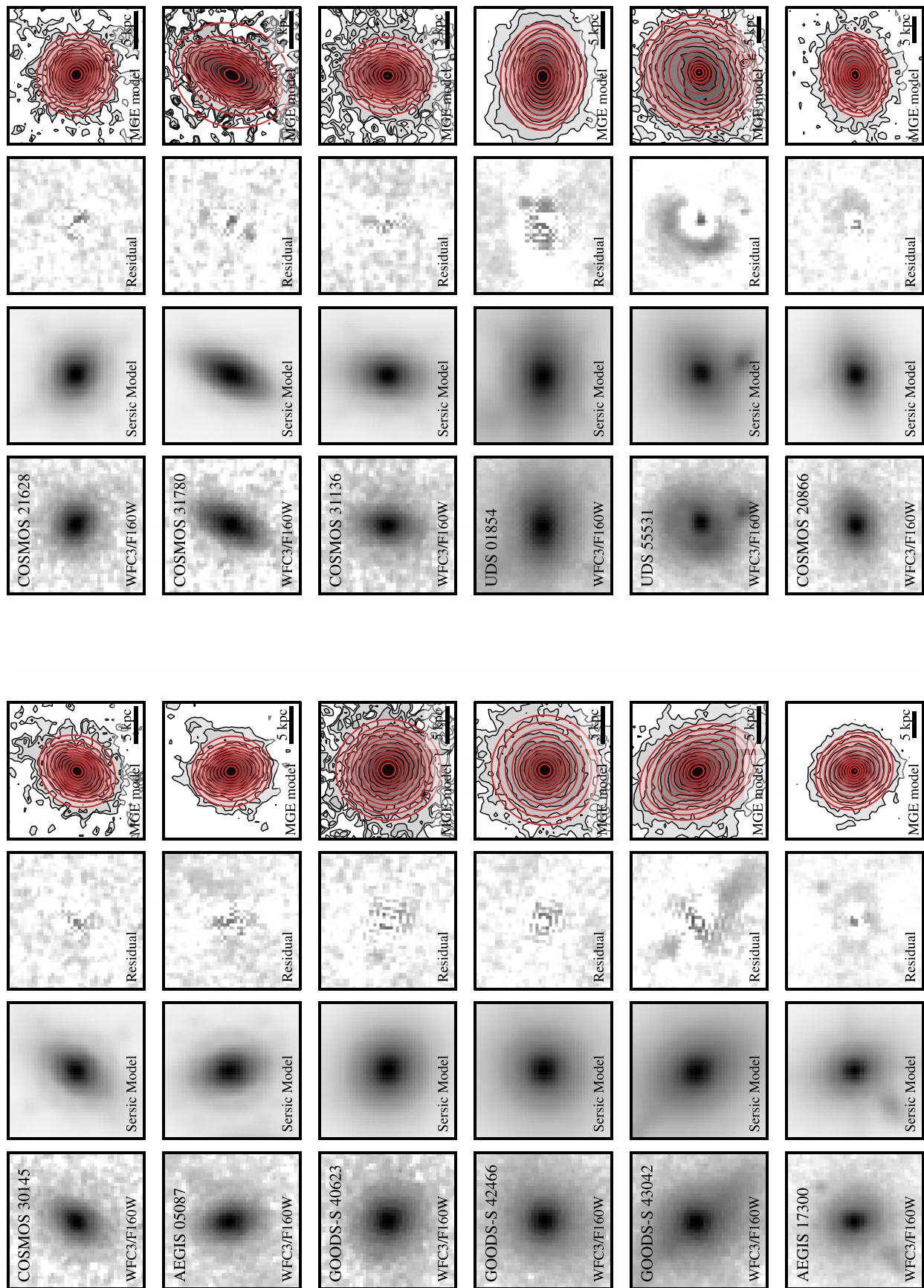
Figure C1. Individual galaxy images and fits, following the format of Figure 5. From left to right, panels show the observed HST WFC3/F160W image, the best-fit Sérsic model derived using `galfit`, residual maps, and an overlay of MGE contours on the observed galaxy image. Images are plotted in surface brightness units, and contours are evenly spaced in steps of 0.5 mag arcsec^{-2}.

(An extended version of this figure is available.)

ORCID iDs

J. Trevor Mendel ◉ https://orcid.org/0000-0002-6327-9147
Alessandra Beifiori ◉ https://orcid.org/0000-0001-8017-6097
Roberto P. Saglia ◉ https://orcid.org/0000-0003-0378-7032
Ralf Bender ◉ https://orcid.org/0000-0001-7179-0626
Gabriel B. Brammer ◉ https://orcid.org/0000-0003-2680-005X
Jeffrey Chan ◉ https://orcid.org/0000-0001-6251-3125
Natascha M. Förster Schreiber ◉ https://orcid.org/0000-0003-4264-3381
Matteo Fossati ◉ https://orcid.org/0000-0002-9043-8764
Ivelina G. Momcheva ◉ https://orcid.org/0000-0003-1665-2073
Erica J. Nelson ◉ https://orcid.org/0000-0002-7524-374X
David J. Wilman ◉ https://orcid.org/0000-0002-1822-4462
Stijn Wuyts ◉ https://orcid.org/0000-0003-3735-1931

References

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, ApJS, 182, 543
Abramson, L. E., Gladders, M. D., Dressler, A., et al. 2016, ApJ, 832, 7
Arnold, J. A., Romanowsky, A. J., Brodie, J. P., et al. 2011, ApJL, 736, L26
Arnold, J. A., Romanowsky, A. J., Brodie, J. P., et al. 2014, ApJ, 791, 80
Ashby, M. L. N., Willner, S. P., Fazio, G. G., et al. 2013, ApJ, 769, 80
Barro, G., Faber, S. M., Dekel, A., et al. 2016, 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
Bastian, N., Covey, K. R., & Meyer, M. R. 2010, ARA&A, 48, 339
Beifiori, A., Maraston, C., Thomas, D., & Johansson, J. 2011, A&A, 531, A109
Beifiori, A., Mendel, J. T., Chan, J. C. C., et al. 2017, ApJ, 846, 120
Beifiori, A., Thomas, D., Maraston, C., et al. 2014, ApJ, 789, 92
Belli, S., Newman, A. B., & Ellis, R. S. 2014a, ApJ, 783, 117
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., & Konidaris, N. P. 2014b, ApJL, 788, L29
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Bertin, G., Ciotti, L., & Del Principe, M. 2002, A&A, 386, 149
Bezanson, R., van der Wel, A., Pacifici, C., et al. 2018, ApJ, 858, 60
Bezanson, R., van Dokkum, P., & Franx, M. 2012, ApJ, 760, 62
Bezanson, R., van Dokkum, P., van de Sande, J., Franx, M., & Kriek, M. 2013a, ApJL, 764, L8
Bezanson, R., van Dokkum, P. G., Tal, T., et al. 2009, ApJ, 697, 1290
Bezanson, R., van Dokkum, P. G., van de Sande, J., et al. 2013b, ApJL, 779, L21
Binney, J. 2005, MNRAS, 363, 937
Binney, J., & Tremaine, S. 1987, Galactic Dynamics (Princeton, NJ: Princeton Univ. Press)
Blumenthal, G. R., Faber, S. M., Flores, R., & Primack, J. R. 1986, ApJ, 301, 27
Boylan-Kolchin, M., Ma, C.-P., & Quataert, E. 2005, MNRAS, 362, 184
Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503
Brammer, G. B., van Dokkum, P. G., Franx, M., et al. 2012, ApJS, 200, 13
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Cappellari, M. 2002, MNRAS, 333, 400
Cappellari, M. 2008, MNRAS, 390, 71
Cappellari, M. 2016, ARA&A, 54, 597
Cappellari, M., Bacon, R., Bureau, M., et al. 2006, MNRAS, 366, 1126
Cappellari, M., di Serego Alighieri, S., Cimatti, A., et al. 2009, ApJL, 704, L34
Cappellari, M., & Emsellem, E. 2004, PASP, 116, 138
Cappellari, M., Emsellem, E., Bacon, R., et al. 2007, MNRAS, 379, 418
Cappellari, M., McDermid, R. M., Alatalo, K., et al. 2012, Natur, 484, 485
Cappellari, M., McDermid, R. M., Alatalo, K., et al. 2013a, MNRAS, 432, 1862
Cappellari, M., Scott, N., Alatalo, K., et al. 2013b, MNRAS, 432, 1709
Carollo, C. M., Bschorr, T. J., Renzini, A., et al. 2013, ApJ, 773, 112
Cenarro, A. J., & Trujillo, I. 2009, ApJL, 696, L43
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
Chan, J. C. C., Beifiori, A., Saglia, R. P., et al. 2018, ApJ, 856, 8
Chang, Y.-Y., van der Wel, A., Rix, H.-W., et al. 2013, ApJ, 773, 149
Charlot, S., & Fall, S. M. 2000, ApJ, 539, 718
Cheung, E., Faber, S. M., Koo, D. C., et al. 2012, ApJ, 760, 131
Cimatti, A., Cassata, P., Pozzetti, L., et al. 2008, A&A, 482, 21
Cimatti, A., Daddi, E., Renzini, A., et al. 2004, Natur, 430, 184
Cimatti, A., Nipoti, C., & Cassata, P. 2012, MNRAS, 422, L62
Comparat, J., Maraston, C., Goddard, D., et al. 2017, arXiv:1711.06575
Conroy, C., Dutton, A. A., Graves, G. J., Mendel, J. T., & van Dokkum, P. G. 2013, ApJL, 776, L26
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. G. 2012, ApJ, 760, 71
Conroy, C., van Dokkum, P. G., & Villaume, A. 2017, ApJ, 837, 166
Daddi, E., Renzini, A., Pirzkal, N., et al. 2005, ApJ, 626, 680
Damjanov, I., Abraham, R. G., Glazebrook, K., et al. 2011, ApJL, 739, L44
Damjanov, I., McCarthy, P. J., Abraham, R. G., et al. 2009, ApJ, 695, 101
Davies, R. I. 2007, MNRAS, 375, 1099
Davies, R. I., Agudo Berbel, A., Wiezorrek, E., et al. 2013, A&A, 558, A56
de la Rosa, I. G., La Barbera, F., Ferreras, I., et al. 2016, MNRAS, 457, 1916
Diemer, B., & Kravtsov, A. V. 2015, ApJ, 799, 108
Diemer, B., Sparre, M., Abramson, L. E., & Torrey, P. 2017, ApJ, 839, 26
Driver, S. P., Hill, D. T., Kelvin, L. S., et al. 2011, MNRAS, 413, 971
Dutton, A. A., Macciò, A. V., Dekel, A., et al. 2016, MNRAS, 461, 2658
Dutton, A. A., Macciò, A. V., Mendel, J. T., & Simard, L. 2013, MNRAS, 432, 2496
Dutton, A. A., Macciò, A. V., Stinson, G. S., et al. 2015, MNRAS, 453, 2447
Dutton, A. A., Mendel, J. T., & Simard, L. 2012, MNRAS Letters, 422, L33
Dutton, A. A., van den Bosch, F. C., Dekel, A., & Courteau, S. 2007, ApJ, 654, 27
Emsellem, E., Cappellari, M., Krajnović, D., et al. 2011, MNRAS, 414, 888
Emsellem, E., Monnet, G., & Bacon, R. 1994, A&A, 285, 723
Fagioli, M., Carollo, C. M., Renzini, A., et al. 2016, ApJ, 831, 173
Fan, L., Lapi, A., Bressan, A., et al. 2010, ApJ, 718, 1460
Fan, L., Lapi, A., De Zotti, G., & Danese, L. 2008, ApJL, 689, L101
Fang, J. J., Faber, S. M., Koo, D. C., & Dekel, A. 2013, ApJ, 776, 63
Ferreras, I., La Barbera, F., de la Rosa, I. G., et al. 2013, MNRAS, 429, L15
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Förster Schreiber, N. M., Genzel, R., Bouché, N., et al. 2009, ApJ, 706, 1364
Foster, C., Pastorello, N., Roediger, J., et al. 2016, MNRAS, 457, 147
Furusawa, H., Kosugi, G., Akiyama, M., et al. 2008, ApJS, 176, 1
Gallazzi, A., Charlot, S., Brinchmann, J., & White, S. D. M. 2006, MNRAS, 370, 1106
Genzel, R., Price, S. H., Übler, H., et al. 2020, arXiv:2006.03046
Genzel, R., Schreiber, N. M. F., Übler, H., et al. 2017, Natur, 543, 397
Gladders, M. D., Oemler, A., Dressler, A., et al. 2013, ApJ, 770, 64
Gnedin, O. Y., Kravtsov, A. V., Klypin, A. A., & Nagai, D. 2004, ApJ, 616, 16
Graves, G. J., Faber, S. M., & Schiavon, R. P. 2009, ApJ, 698, 1590
Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35
Hill, A. R., van der Wel, A., Franx, M., et al. 2019, ApJ, 871, 76
Hilz, M., Naab, T., Ostriker, J. P., et al. 2012, MNRAS, 425, 3119
Hilz, M., Naab, T., & Ostriker, J. P. 2013, MNRAS, 429, 2924
Hirschmann, M., Dolag, K., Saro, A., et al. 2014, MNRAS, 442, 2304
Hopkins, P. F., Bundy, K., Murray, N., et al. 2009a, MNRAS, 398, 898
Hopkins, P. F., Hernquist, L., Cox, T. J., Keres, D., & Wuyts, S. 2009b, ApJ, 691, 1424
Hyde, J. B., & Bernardi, M. 2009, MNRAS, 396, 1171
Karademir, G. S., Remus, R.-S., Burkert, A., et al. 2019, MNRAS, 487, 318
Kausch, W., Noll, S., Smette, A., et al. 2014, in ASP Conf. Ser. 485, Astronomical Data Anaylsis Softward and Systems XXIII, ed. N. Manset & P. Forshay (San Francisco, CA: ASP), 403
Keating, S. K., Abraham, R. G., Schiavon, R., et al. 2015, ApJ, 798, 26
Kelvin, L. S., Driver, S. P., Robotham, A. S. G., et al. 2012, MNRAS, 421, 1007
Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36
Kriek, M., Conroy, C., van Dokkum, P. G., et al. 2016, Natur, 540, 248
Kriek, M., Shapley, A. E., Reddy, N. A., et al. 2015, ApJS, 218, 15
Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2009, ApJ, 700, 221
Krumholz, M. R. 2014, PhR, 539, 49
Lang, P., Förster Schreiber, N. M., Genzel, R., et al. 2017, ApJ, 840, 92
Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, MNRAS, 379, 1599
Leja, J., Johnson, B. D., Conroy, C., et al. 2019, ApJ, 877, 140
Li, H., Ge, J., Mao, S., et al. 2017, ApJ, 838, 77
Lin, X., Fang, G., Cai, Z.-Y., et al. 2019, ApJ, 875, 83

Liske, J., Baldry, I. K., Driver, S. P., et al. 2015, MNRAS, 452, 2087
Lonoce, I., Longhetti, M., Maraston, C., et al. 2015, MNRAS, 454, 3912
Lovell, M. R., Pillepich, A., Genel, S., et al. 2018, MNRAS, 481, 1950
Lyubenova, M., Martín-Navarro, I., van de Ven, G., et al. 2016, MNRAS, 463, 3220
Martín-Navarro, I., La Barbera, F., Vazdekis, A., Falcón-Barroso, J., & Ferreras, I. 2015a, MNRAS, 447, 1033
Martín-Navarro, I., Pérez-González, P. G., Trujillo, I., et al. 2015b, ApJL, 798, L4
McDermid, R. M., Alatalo, K., Blitz, L., et al. 2015, MNRAS, 448, 3484
McGaugh, S. S. 2005, ApJ, 632, 859
McLean, I. S., Steidel, C. C., Epps, H. W., et al. 2012, Proc. SPIE, 8446, 84460J
Meert, A., Vikram, V., & Bernardi, M. 2015, MNRAS, 446, 3943
Mendel, J. T., Saglia, R. P., Bender, R., et al. 2015, ApJL, 804, L4
Mendel, J. T., Simard, L., Palmer, M., Ellison, S. L., & Patton, D. R. 2014, ApJS, 210, 3
Momcheva, I. G., Brammer, G. B., van Dokkum, P. G., et al. 2016, ApJS, 225, 27
Morishita, T., Abramson, L. E., Treu, T., et al. 2018, ApJL, 856, L4
Moster, B. P., Naab, T., & White, S. D. M. 2013, MNRAS, 428, 3121
Mundy, C. J., Conselice, C. J., & Ownsworth, J. R. 2015, MNRAS, 450, 3696
Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, ApJ, 777, 18
Naab, T., Johansson, P. H., & Ostriker, J. P. 2009, ApJL, 699, L178
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, ApJ, 462, 563
Newman, A. B., Belli, S., & Ellis, R. S. 2015, ApJL, 813, L7
Newman, A. B., Belli, S., Ellis, R. S., & Patel, S. G. 2018, ApJ, 862, 126
Newman, A. B., Ellis, R. S., Bundy, K., & Treu, T. 2012, ApJ, 746, 162
Newman, A. B., Ellis, R. S., Treu, T., & Bundy, K. 2010, ApJL, 717, L103
Nipoti, C., Treu, T., Leauthaud, A., et al. 2012, MNRAS, 422, 1714
Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713
Onodera, M., Carollo, C. M., Renzini, A., et al. 2015, ApJ, 808, 161
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
Parikh, T., Thomas, D., Maraston, C., et al. 2018, MNRAS, 477, 3954
Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H. 2002, AJ, 124, 266
Pforr, J., Maraston, C., & Tonini, C. 2012, MNRAS, 422, 3285
Posacki, S., Cappellari, M., Treu, T., Pellegrini, S., & Ciotti, L. 2015, MNRAS, 446, 493
Price, S. H., Kriek, M., Shapley, A. E., et al. 2016, ApJ, 819, 80
Prichard, L. J., Davies, R. L., Beifiori, A., et al. 2017, ApJ, 850, 203
Ramsay, S. K., Mountain, C. M., & Geballe, T. R. 1992, MNRAS, 259, 751
Remus, R.-S., Dolag, K., Naab, T., et al. 2017, MNRAS, 464, 3742
Renzini, A., & Ciotti, L. 1993, ApJL, 416, L49
Saglia, R. P., Sánchez-Blázquez, P., Bender, R., et al. 2010, A&A, 524, A6
Salpeter, E. E. 1955, ApJ, 121, 161
Scott, N., Brough, S., Croom, S. M., et al. 2017, MNRAS, 472, 2833
Sérsic, J. L. 1963, BAAA, 6, 41
Sharples, R., Bender, R., Agudo Berbel, A., et al. 2012, Proc. SPIE, 8446, 84460K
Sharples, R., Bender, R., Agudo Berbel, A., et al. 2013, Msngr, 151, 21
Shetty, S., & Cappellari, M. 2014, ApJL, 786, L10
Shetty, S., & Cappellari, M. 2015, MNRAS, 454, 1332
Simard, L., Mendel, J. T., Patton, D. R., Ellison, S. L., & McConnachie, A. W. 2011, ApJS, 196, 11
Simard, L., Willmer, C. N. A., Vogt, N. P., et al. 2002, ApJS, 142, 1

Simpson, C., Rawlings, S., Ivison, R., et al. 2012, MNRAS, 421, 3060
Skelton, R. E., Whitaker, K. E., Momcheva, I. G., et al. 2014, ApJS, 214, 24
Smith, R. J., Lucey, J. R., & Conroy, C. 2015, MNRAS, 449, 3441
Sohn, J., Zahid, H. J., & Geller, M. J. 2017, ApJ, 845, 73
Sonnenfeld, A., Treu, T., Marshall, P. J., et al. 2015, ApJ, 800, 94
Spiniello, C., Trager, S., Koopmans, L. V. E., & Conroy, C. 2014, MNRAS, 438, 1483
Thomas, D., Maraston, C., Schawinski, K., Sarzi, M., & Silk, J. 2010, MNRAS, 404, 1775
Thomas, D., Steele, O., Maraston, C., et al. 2013, MNRAS, 431, 1383
Thomas, J., Jesseit, R., Saglia, R. P., et al. 2009, MNRAS, 393, 641
Thomas, J., Saglia, R. P., Bender, R., et al. 2011, MNRAS, 415, 545
Tiley, A. L., Swinbank, A. M., Harrison, C. M., et al. 2019, MNRAS, 485, 934
Tinsley, B. M. 1978, ApJ, 222, 14
Toft, S., Gallazzi, A., Zirm, A., et al. 2012, ApJ, 754, 3
Toft, S., Zabl, J., Richard, J., et al. 2017, Natur, 546, 510
Tortora, C., Napolitano, N. R., Roy, N., et al. 2018, MNRAS, 473, 969
Tortora, C., Napolitano, N. R., Saglia, R. P., et al. 2014, MNRAS, 445, 162
Treu, T., Auger, M. W., Koopmans, L. V. E., et al. 2010, ApJ, 709, 1195
Trujillo, I., Feulner, G., Goranova, Y., et al. 2006, MNRAS, 373, L36
Übler, H., Förster Schreiber, N. M., Genzel, R., et al. 2017, ApJ, 842, 121
Übler, H., Genzel, R., Tacconi, L. J., et al. 2018, ApJL, 854, L24
Valentinuzzi, T., Poggianti, B. M., Saglia, R. P., et al. 2010, ApJL, 721, L19
van de Sande, J., Kriek, M., Franx, M., et al. 2013, ApJ, 771, 85
van de Sande, J., Kriek, M., Franx, M., Bezanson, R., & van Dokkum, P. G. 2014, ApJL, 793, L31
van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, ApJ, 788, 28
van der Wel, A., Rix, H.-W., Wuyts, S., et al. 2011, ApJ, 730, 38
van der Wel, A., & van der Marel, R. P. 2008, ApJ, 684, 260
van Dokkum, P., Conroy, C., Villaume, A., Brodie, J., & Romanowsky, A. 2017, ApJ, 841, 68
van Dokkum, P. G. 2008, ApJ, 674, 29
van Dokkum, P. G., & Conroy, C. 2010, Natur, 468, 940
van Dokkum, P. G., & Franx, M. 1996, MNRAS, 281, 985
van Dokkum, P. G., Franx, M., Kriek, M., et al. 2008, ApJL, 677, L5
van Dokkum, P. G., Nelson, E. J., Franx, M., et al. 2015, ApJ, 813, 23
van Dokkum, P. G., Whitaker, K. E., Brammer, G., et al. 2010, ApJ, 709, 1018
Veale, M., Ma, C.-P., Thomas, J., et al. 2017, MNRAS, 464, 356
Wake, D. A., van Dokkum, P. G., & Franx, M. 2012, ApJL, 751, L44
Wellons, S., & Torrey, P. 2017, MNRAS, 467, 3887
Whitaker, K. E., Labbé, I., van Dokkum, P. G., et al. 2011, ApJ, 735, 86
Whitaker, K. E., van Dokkum, P. G., Brammer, G., et al. 2013, ApJL, 770, L39
Wilkinson, D. M., Maraston, C., Goddard, D., Thomas, D., & Parikh, T. 2017, MNRAS, 472, 4297
Williams, M. J., Bureau, M., & Cappellari, M. 2010, MNRAS, 409, 1330
Williams, R. J., Quadri, R. F., Franx, M., van Dokkum, P., & Labbé, I. 2009, ApJ, 691, 1879
Wisnioski, E., Förster Schreiber, N. M., Fossati, M., et al. 2019, ApJ, 886, 124
Wisnioski, E., Förster Schreiber, N. M., Wuyts, S., et al. 2015, ApJ, 799, 209
Wolf, J., Martinez, G. D., Bullock, J. S., et al. 2010, MNRAS, 406, 1220
Woo, J., Dekel, A., Faber, S. M., & Koo, D. C. 2015, MNRAS, 448, 237
Wright, A. H., Robotham, A. S. G., Bourne, N., et al. 2016, MNRAS, 460, 765
Wuyts, S., Förster Schreiber, N. M., Nelson, E. J., et al. 2013, ApJ, 779, 135
Wuyts, S., Förster Schreiber, N. M., van der Wel, A., et al. 2011, ApJ, 742, 96
Wuyts, S., Förster Schreiber, N. M., Wisnioski, E., et al. 2016, ApJ, 831, 149
Zahid, H. J., Geller, M. J., Damjanov, I., & Sohn, J. 2019, ApJ, 878, 158