Universiteit
Leiden
The Netherlands

# MAGAZ3NE: high stellar velocity dispersions for ultramassive quiescent galaxies at z ≳ 3

Forrest, B.; Wilson, G.; Muzzin, A.; Marchesini, D.; Cooper, M.C.; Marsan, Z.C.; ... ; Stefanon, M.

**Citation**
Forrest, B., Wilson, G., Muzzin, A., Marchesini, D., Cooper, M. C., Marsan, Z. C., … Stefanon, M. (2022). MAGAZ3NE: high stellar velocity dispersions for ultramassive quiescent galaxies at z ≳ 3. *The Astrophysical Journal*, *938*(2). doi:10.3847/1538-4357/ac8747

| | |
|---|---|
| Version: | Publisher's Version |
| License: | Creative Commons CC BY 4.0 license |
| Downloaded from: | https://hdl.handle.net/1887/3561986 |

**Note:** To cite this publication please use the final published version (if applicable).

The Astrophysical Journal, 938:109 (18pp), 2022 October 20 https://doi.org/10.3847/1538-4357/ac8747
© 2022. The Author(s). Published by the American Astronomical Society.
OPEN ACCESS

# MAGAZ3NE: High Stellar Velocity Dispersions for Ultramassive Quiescent Galaxies at $z \gtrsim 3$*

Ben Forrest[1,2], Gillian Wilson[2], Adam Muzzin[3], Danilo Marchesini[4], M. C. Cooper[5], Z. Cemile Marsan[3], Marianna Annunziatella[4,6], Ian McConachie[2], Kumail Zaidi[4], Percy Gomez[7], Stephanie M. Urbano Stawinski[5], Wenjun Chang[2], Gabriella de Lucia[8], Francesco La Barbera[9], Lori Lubin[1], Julie Nantais[10], Theodore Peña[11], Paolo Saracco[12], Jason Surace[13], and Mauro Stefanon[14]

[1] Department of Physics and Astronomy, University of California, Davis, One Shields Avenue, Davis, CA 95616, USA; bforrest@ucdavis.edu
[2] Department of Physics and Astronomy, University of California, Riverside, 900 University Avenue, Riverside, CA 92521, USA
[3] Department of Physics and Astronomy, York University, 4700, Keele Street, Toronto, ON MJ3 1P3, Canada
[4] Department of Physics and Astronomy, Tufts University, 574 Boston Avenue, Suite 304, Medford, MA 02155, USA
[5] Department of Physics and Astronomy, University of California, Irvine, 4129 Frederick Reines Hall, Irvine, CA, USA
[6] Centro de Astrobiología (CSIC-INTA), Ctra de Torrejón a Ajalvir, km 4, E-28850 Torrejón de Ardoz, Madrid, Spain
[7] W.M. Keck Observatory, 65-1120 Mamalahoa Highway, Kamuela, HI 96743, USA
[8] INAF—Astronomical Observatory of Trieste, via G.B. Tiepolo 11, I-34143 Trieste, Italy
[9] INAF—Osservatorio Astronomico di Capodimonte, sal. Moiariello 16, I-80131 Napoli, Italy
[10] Departamento de Ciencias Físicas, Universidad Andres Bello, Fernández Concha 700, Las Condes 7591538, Santiago, Región Metropolitana, Chile
[11] Department of Astronomy, University of Wisconsin-Madison, 475 N. Charter Street, Madison, WI 53706-1507, USA
[12] INAF—Osservatorio Astronomico di Brera, via Brera 28, I-20121 Milano, Italy
[13] IPAC, Mail Code 100-22, Caltech, 1200 E. California Boulevard, Pasadena, CA 91125, USA
[14] Leiden Observatory, Leiden University, 2300 RA Leiden, The Netherlands
*Received 2022 May 6; revised 2022 July 29; accepted 2022 August 3; published 2022 October 18*

## Abstract

In this work, we publish stellar velocity dispersions, sizes, and dynamical masses for eight ultramassive galaxies (UMGs; $\log(M_*/M_\odot) > 11$), $z \gtrsim 3$) from the Massive Ancient Galaxies At $z > 3$ NEar-infrared (MAGAZ3NE) Survey, more than doubling the number of such galaxies with velocity dispersion measurements at this epoch. Using the deep Keck/MOSFIRE and Keck/NIRES spectroscopy of these objects in the $H$ and $K$ bandpasses, we obtain large velocity dispersions of ∼400 km s$^{-1}$ for most of the objects, which are some of the highest stellar velocity dispersions measured and ∼40% larger than those measured for galaxies of similar mass at $z \sim 1.7$. The sizes of these objects are also smaller by a factor of 1.5–3 compared to this same $z \sim 1.7$ sample. We combine these large velocity dispersions and small sizes to obtain dynamical masses. The dynamical masses are similar to the stellar masses of these galaxies, consistent with a Chabrier initial mass function (IMF). Considered alongside previous studies of massive quiescent galaxies across $0.2 < z < 4.0$, there is evidence for an evolution in the relation between the dynamical mass–stellar mass ratio and velocity dispersion as a function of redshift. This implies an IMF with fewer low-mass stars (e.g., Chabrier IMF) for massive quiescent galaxies at higher redshifts in conflict with the bottom-heavy IMF (e.g., Salpeter IMF) found in their likely $z \sim 0$ descendants, though a number of alternative explanations such as a different dynamical structure or significant rotation are not ruled out. Similar to data at lower redshifts, we see evidence for an increase of IMF normalization with velocity dispersion, though the $z \gtrsim 3$ trend is steeper than that for $z \sim 0.2$ early-type galaxies and offset to lower dynamical-to-stellar mass ratios.

*Unified Astronomy Thesaurus concepts:* Galaxy evolution (594); High-redshift galaxies (734); Quenched galaxies (2016)

## 1. Introduction

Deep near-infrared photometric surveys of the last decade have suggested larger numbers of massive galaxies at high redshifts than predicted by cosmological galaxy simulations (e.g., Muzzin et al. 2013a; Straatman et al. 2014; Sherman et al. 2019; Marsan et al. 2022). More recent simulations have better agreement with observations, but the discrepancy is still a factor of a few to 10 at the highest masses (though see Donnari et al. 2021; Lustig et al. 2022). In the last several years, spectroscopic confirmation of a handful of galaxies with stellar masses of $\log(M_*/M_\odot) > 11$ and at redshifts of $z > 3$ have shown that such galaxies do indeed exist in nonnegligible numbers (Marsan et al. 2015, 2017; Glazebrook et al. 2017; Schreiber et al. 2018a; Tanaka et al. 2019; Forrest et al. 2020a, 2020b; Valentino et al. 2020), but a robust measurement of the number density of such galaxies is still lacking. This is largely due to the fact that the determination of stellar mass, particularly from photometry alone, requires a number of assumptions that introduce the possibility for significant error.

There are numerous programs that determine galaxy parameters via spectral energy distribution (SED) fitting. Nearly all require some assumptions about the geometry of dust and dust

* The spectra presented herein were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation.

Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

extinction, the initial mass function (IMF) of star formation, a parametric form of the star formation history, strength of emission lines, and choice of stellar population synthesis models, each of which plays a role in the determined stellar mass of a galaxy. For large populations of galaxies, the median mass determination appears sensitive to these choices with scatter ∼0.2 dex (e.g., Wuyts et al. 2009; Mobasher et al. 2015), though the choice of code can lead to systematic offsets of up to 0.3 dex (Muzzin et al. 2009; Leja et al. 2019). While these differences are perhaps tolerable, the differences for individual galaxies can greatly exceed these numbers in cases with significant flux contributions from strong emission lines (Stark et al. 2013; Salmon et al. 2015; Forrest et al. 2017) and active galactic nuclei (AGNs; Leja et al. 2018), as well as in outlier cases where photometric redshifts are highly discrepant from true redshifts, though this seems less common in massive galaxies even at $z > 3$ (Schreiber et al. 2018a; Forrest et al. 2020b).

As a result, probing stellar masses independently of the above assumptions is valuable. While the stellar velocity dispersion formally probes the total mass of a galaxy, the massive, high-redshift galaxies of interest here typically have small sizes and have central masses dominated by stars (e.g., van der Wel et al. 2014; Straatman et al. 2015a; Saracco et al. 2019). Locally, stellar velocity dispersion is well correlated with the luminosity and radius of elliptical galaxies (e.g., Faber & Jackson 1976; Djorgovski & Davis 1987; Dressler 1987; Shu et al. 2012), the mass of the central black hole (e.g., Gebhardt et al. 2000; Kormendy & Ho 2013), the mass-to-light ratio (e.g., Cappellari et al. 2006), and numerous other properties including galaxy color (Wake et al. 2012) and stellar mass (e.g., Zahid et al. 2016). Velocity dispersions have been studied out to higher redshifts as well, and many such correlations appear to hold for these data, though they may be offset from the local relations (e.g., van Dokkum et al. 2009b; Newman et al. 2010; Bezanson et al. 2012, 2013b; Thomas et al. 2013; van de Sande et al. 2013; Gargiulo et al. 2016; Hill et al. 2016; Belli et al. 2017). However, like the measurement of stellar masses, the measurement of stellar velocity dispersions holds the potential for systematic and statistical errors, the latter of which can of course be significant for low-signal-to-noise (S/N) spectra. The interpretation also requires careful analysis, as effects such as galaxy rotation and inclination can either increase or decrease measured velocity dispersions (Bezanson et al. 2018; Newman et al. 2018; Mendel et al. 2020).

Still, stellar velocity dispersions can be used in concert with structural measurements to calculate dynamical masses, which are sensitive to the gravitational potential of a galaxy and, therefore, to the contribution of dark matter as well as the contributions of dust, gas, and stars. This then provides an effective upper limit on the stellar mass of a galaxy, independent of the numerous assumptions intrinsic to the calculation of stellar masses via SED fitting, including the shape of the initial mass function (IMF).

Variability in the IMF, which traces the number of stars formed as a function of their mass in a star-forming molecular cloud, can contribute to nonnegligible differences in the determination of stellar mass as it sets the effective mass-to-light ratio. The IMF of many galaxies, particularly local massive early-type galaxies (ETGs), is inferred via spectral fitting or dynamical modeling to have a "heavy" mass-to-light ratio (with respect to the Milky Way distribution), such as that of the Salpeter (1955) IMF, which assumes a functional power law with index $x = -2.35$ (termed "heavy" due to the larger effective mass-to-light ratio). However, observations have suggested that the IMF is not universal (see Hopkins 2018 for a review) and can vary over cosmic time, between galaxies, or as a function of galaxy radius, metallicity, stellar mass, or star formation density (e.g., Cappellari et al. 2006; van Dokkum 2008; Conroy & van Dokkum 2012; Cappellari et al. 2013a, 2013b; Kroupa et al. 2013; van Dokkum et al. 2017; Villaume et al. 2017; La Barbera et al. 2019). As such, it is important to note that any measurement of the IMF in a galaxy is a measurement of the superposition of the IMF during any and all episodes of star formation in that galaxy.

Recently, Mendel et al. (2020) homogeneously analyzed 58 massive quiescent galaxies at $1.4 < z < 2.1$ and found that galaxies with higher stellar velocity dispersions at a given epoch prefer a heavier IMF such as that from Salpeter (1955), while galaxies with lower stellar velocity dispersions are better described by a lighter IMF such as the Chabrier (2003) IMF. This result agrees with the lower-redshift analysis from Posacki et al. (2015), though the higher-redshift galaxies have systematically higher velocity dispersions than lower-redshift galaxies with the same dynamical-to-stellar mass ratio.

Measurements of velocity dispersion require spectra with reasonable S/Ns, which are difficult to obtain for galaxies at earlier epochs. As such, only six massive galaxies with stellar masses $\log(M_*/M_\odot) \gtrsim 11$ at $z > 3$ have measured stellar velocity dispersions (Tanaka et al. 2019; Saracco et al. 2020; Esdaile et al. 2021). In this work, we measure velocity dispersions for eight additional massive galaxies at $z \gtrsim 3$ using the MOSFIRE (McLean et al. 2010, 2012) and NIRES (Wilson et al. 2004) instruments on Keck, more than doubling the size of the current sample in the literature—four out of eight of these galaxies are more massive than any in the $z > 3$ sample with velocity dispersions in the literature. Combined with size measurements for these galaxies and values from the literature, we perform the first statistical comparison of dynamical and stellar masses at this early epoch using 14 massive galaxies.

We present the data in Section 2, the velocity dispersion calculations and image analysis process in Section 3, and then a discussion of the results in Section 4, and the main conclusions in Section 5. All analysis here uses a ΛCDM cosmology with $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$ as well as the AB magnitude system (Oke & Gunn 1983). A Chabrier (2003) IMF is used for the calculation of stellar mass.

## 2. Data

### 2.1. Parent Photometric Catalogs

Targets selected for spectroscopic follow-up in the MAGAZ3NE survey were drawn from parent photometric catalogs in the UltraVISTA DR1 (Muzzin et al. 2013b), UltraVISTA DR3 (A. Muzzin et al. 2022, in preparation), and XMM-VIDEO (M. Annunziatella et al. 2022, in preparation) fields.

The UltraVISTA survey (McCracken et al. 2012) imaged over 1.62 deg$^2$ in the COSMOS field with deep near-infrared $Y$, $J$, $H$, and $K$ bandpasses. The first data release (DR1 catalogs; Muzzin et al. 2013b) combined additional photometry from 0.15–24 $\mu$m, yielding a total of 30 bandpasses with 90% completeness at $K_s = 23.4$ mag. Subsequent deep imaging over 0.84 deg$^2$ in the NIR furthered the value of the data set, with

DR3 (Marsan et al. 2022; A. Muzzin et al. 2022, in preparation) reaching deeper than DR1 by 1.1 mag in the $K_s$ band and $\sim$1.2 mag deeper in the IRAC 3.6 and 4.5 $\mu$m bandpasses (Ashby et al. 2018). A total of up to 49 bandpasses in DR3 allowed for highly accurate galaxy spectral energy distributions (SEDs) and photometric redshift determinations, as well as the detection of massive, quiescent galaxies at $z > 3$, which are too faint for accurate characterization with optical photometry alone.

The VISTA Deep Extragalactic Observations (VIDEO; Jarvis et al. 2013) survey similarly acquired deep NIR imaging over several fields, including IRAC data from SERVS (Mauduit et al. 2012) and the DeepDrill survey (Lacy et al. 2021). Catalogs used in this work are built from VIDEO DR4 data over 5.1 deg$^2$ in the XMM-Newton Large Scale Structure (XMM) field with up to 22 bandpasses from the $u$ band to IRAC 8.0 $\mu$m and a 5$\sigma$ depth of $K_s = 23.8$ mag (M. Annunziatella et al. 2022, in preparation). While this catalog is somewhat shallower in $K$-band depth, it covers a wider area, which is important for the detection of the rare massive, quiescent objects at these redshifts.

### 2.2. Near-infrared Spectroscopy

For this work, we analyze $H$- and $K$-band spectroscopic observations from Keck-MOSFIRE (McLean et al. 2010, 2012) taken as part of the MAGAZ3NE survey (Forrest et al. 2020b), and details of the spectroscopic target selection are provided therein. The general survey observing strategy called for targeting ultramassive galaxies (UMGs) in the $K$ band, where the strong emission features [OIII]$\lambda$5007 and H$\beta$ fall at the redshift of the sample. On-the-fly reduction was used, and once a redshift was confirmed, observation of a UMG was stopped. As such, UMGs with strong emission lines and only faint detection of the continuum have insufficient S/N to calculate a stellar velocity dispersion.

However, 6 of the 16 confirmed UMGs from Forrest et al. (2020b) have MOSFIRE observations in the $H$ band, where a greater number of spectral features lie (e.g., $D_n(4000)$, Ca H&K, and higher-order Balmer absorption features), enabling a more reliable velocity dispersion calculation. Since the publication of Forrest et al. (2020b), a redshift has also been obtained for an additional UMG, COS-DR1-99209, at $z = 2.983$ observed in both $H$ and $K$ bands with MOSFIRE.

For these seven galaxies, the MOSFIRE DRP was used to reduce the raw spectroscopy to 2D spectra. From there, a custom code written by one of us (B.F.) was used to optimally extract a 1D spectrum and perform telluric corrections using stars observed on the same masks, modeled with the PHOENIX stellar models (Husser et al. 2013). A more detailed description is provided in Forrest et al. (2020b). Galaxy photometry is fit in conjunction with a single-band spectrum using FAST++ (Schreiber et al. 2018b) to obtain the relative scaling of different spectral bandpasses.

In the case of XMM-VID1-2075, the only MAGAZ3NE UMG in this work without $H$-band MOSFIRE spectroscopy, $J$-, $H$-, and $K$-band NIRES spectroscopy were independently obtained (PI: Gomez; P. Gomez et al. 2022, in preparation). A comparison of the two (very similar) $K$-band spectra is presented in Appendix A. The NIRES data were reduced using Pypeit (version 1.0.4; Prochaska et al. 2020). Pypeit flat fields the science data, performs wavelength calibration, models and subtracts the sky background, and performs a flux calibration. A telluric correction was also calculated using Molecfit (Smette et al. 2015; Kausch et al. 2015).

In total, we thus present new velocity dispersions for eight MAGAZ3NE UMGs in this work: COS-DR1-99209, COS-DR3-84674, COS-DR3-111740, COS-DR3-201999, COS-DR3-202019, XMM-VID1-2075, XMM-VID3-1120, and XMM-DR3-2457. We also include a ninth UMG from the MAGAZ3NE sample, COS-DR3-160748, which has a velocity dispersion from a high-S/N spectrum taken with the LBT published in Saracco et al. (2020) as C1-23152.

## 3. Analysis

### 3.1. Velocity Dispersions

Absorption feature stellar velocity dispersions were calculated using the Penalized Pixel-Fitting method (pPXF; Cappellari & Emsellem 2004; Cappellari 2017) in conjunction with the UMG spectra. This maintains consistency with the analysis of other $z > 3$ massive galaxies (Tanaka et al. 2019; Saracco et al. 2020; Esdaile et al. 2021).

#### 3.1.1. Inputs for pPXF Velocity Dispersion Calculations

When available, spectroscopy from both the $H$ and $K$ bands was used. Observed spectra were logarithmically rebinned and corrected for instrumental resolution, and templates were resampled to match this resolution. In cases where the resolution for a template was less than that of the observed spectra, the spectra were binned using inverse variance weighting to match the model resolution. Numerous runs were performed for each galaxy using a variety of spectral template libraries, wavelength-masking strategies, and a range of additive Legendre polynomial orders to limit the effects of template mismatch and telluric correction inaccuracies. The variety of inputs also allows us to characterize the systematic error on the velocity dispersion, which exceeds the statistical error provided by pPXF.

Extensive testing of pPXF on a sample of five massive quiescent galaxies $1.4 < z < 2.1$ was performed in van de Sande et al. (2013). In their Appendix A, they test the dependence of the velocity dispersions output by pPXF on various inputs, including template choice, polynomial degree, and stellar population models. We do similar testing when fitting the MAGAZ3NE sample, which is described in more detail in Appendix B.

Briefly, we ran pPXF using the templates from Bruzual & Charlot (2003, BC03), SSPs constructed from the MILES library (Sánchez-Blázquez et al. 2006; Vazdekis et al. 2010), and the Indo-US library (Valdes et al. 2004). These libraries provide a sufficient variety of spectral templates to fit the observed spectra well. However, pPXF does not incorporate galaxy photometry into the fit, and failure to do this can result in underestimating the velocity dispersions (Mendel et al. 2020; though results are often consistent within the errors). As such, we also use FAST++ (Schreiber et al. 2018b) with the Bruzual & Charlot (2003, BC03) templates to jointly fit the observed photometry and spectroscopy and obtain a best-fit template, subsequently using pPXF with that template choice fixed—we designate these runs as BC03++.

Runs with each of these four template sets (BC03++, BC03, MILES, and Indo-US) were also done with an additive Legendre polynomial from order $0 \leqslant d \leqslant 50$. Such a polynomial corrects for differences in the template and observed spectral shape as can result from, e.g., telluric correction

inaccuracies and helps avoid template mismatch. The effect of adding a polynomial of a very high order is to perturb a template to fit all the noise features in an observed spectrum, and thus, a somewhat low-order polynomial is preferred. The choice of polynomial order varies in the literature: van de Sande et al. (2013) use $d \sim 17$, with velocity dispersions only showing a small dependence on this choice from $0 \leqslant d \leqslant 50$; Mendel et al. (2020) use $d = 9$; Saracco et al. (2020) use $d = 4$; Tanaka et al. (2019) use $d = 1$; and Esdaile et al. (2021) do not use an additive polynomial (effectively $d = 0$). In general, we find that the velocity dispersion varies the least over the range $10 < d < 20$ for the UMGs in this sample.

Finally, we also choose various methods of masking the spectral wavelengths used in the fit. We test pPXF while masking all observed emission lines as well as (1) all Balmer features, (2) the H$\beta$ feature, (3) no other wavelengths, and (4)–(6) wavelengths in 1–3 plus sky lines. Exclusion of the Balmer features can result in a more stable velocity dispersion (van de Sande et al. 2013) and remove any degeneracy between small-scale emission and template choice, but also remove a strong constraint on the velocity dispersion for spectra with low-S/N as is typical for galaxies at these redshifts (Tanaka et al. 2019; Esdaile et al. 2021). Masking only the H$\beta$ feature in these quiescent galaxies strongly mitigates the emission issue.

#### 3.1.2. Measured Velocity Dispersions

The resultant best-fit templates from each run were visually inspected and also compared to the galaxy photometry, with results involving clearly incorrect templates discarded (these were uncommon, on the order of a few percent). Our galaxies have sufficient S/N such that the results of the many runs form a distribution with a clear mode for each galaxy, which we use as the velocity dispersion in subsequent analysis. The (asymmetrical) spread of the distribution of results is used to derive errors on the velocity dispersion, which can differ from the output error of pPXF by up to a factor of $\sim$2. Median values of the fitted velocity dispersion distributions and averages weighted by the reduced $\chi^2$ and reported error are all statistically consistent with the mode of the distribution. Models with the best-fit velocity dispersions are shown in Figure 1. Plots showing the dependencies on the choice of input parameters, as well as a more complete discussion, are included in Appendix B.

#### 3.1.3. Aperture Correction of the Measured Velocity Dispersions

For comparison with other measurements in the literature, we correct the measured velocity dispersions to velocity dispersions at the effective radius, $\sigma_e$ (size calculations are described in Section 3.2). This removes instrumental dependence and accounts for the effects of seeing. Such a correction is dependent upon the size and shape of the spectral aperture, the observing conditions (i.e., seeing), and the size of the target. The MOSFIRE aperture size of interest, $r_{\rm aperture}$, is the distance along the slit over which the 1D spectrum was optimally extracted and is thus a function of both intrinsic size and seeing conditions which varies for different masks on which the same object is located. In theory, this could also be affected by the length of a slit if it was insufficiently long to cover the entire object (minimum MOSFIRE slit length is $7\farcs1$), though this would only be a concern for very large objects or extremely poor conditions, which does not affect this sample.

Extensive modeling in Appendix B of van de Sande et al. (2013) shows that for a rectangular aperture with weighted extraction, this correction factor is quite flat as a function of $r_{\rm aperture}/r_{\rm eff}$, when the PSF is taken into account. Indeed, the correction factors for our velocity dispersions calculated following van de Sande et al. (2013) range from 1.048 to 1.058, though the small differences in this correction are far exceeded by the errors in the measured velocity dispersions. The corrected values are shown in Table 1 and used for the remainder of this analysis.

### 3.2. Sizes

GALFIT (Peng et al. 2002, 2010) was used to model the $K_S$-band images of all objects, and the HST/*WFC*3/ F160W images of COS-DR3-201999, COS-DR3-202019, and COS-DR3-84674. For the sources in the COSMOS field, the UltraVISTA DR4 $K_S$ mosaic with a pixel scale of $0\farcs15$ per pixel and FWHM $= 0\farcs78$ (McCracken et al. 2012) was adopted. For the sources in the XMM field, the VIDEO DR4 $K_S$ mosaic with a pixel scale of $0\farcs2$ per pixel and FWHM $= 0\farcs82$ (Jarvis et al. 2013) was adopted. The fitting process was similar for all the galaxies. A small cutout centered on the relevant galaxy was created, making sure to include the central object and any nearby objects along with enough empty region for the sky background calculation. In most cases, the central galaxy was fitted simultaneously with the neighboring objects. In a few cases, the neighboring objects were not fitted if they were far enough from the UMG that their light was not contaminating the objects. In this case, the neighboring objects were only masked out in the GALFIT fitting. All objects were fitted with a single Sérsic profile. The free parameters had the following fitting constraints: the centroid of the object was allowed to vary at most by 2 pixels in each direction from the initial coordinates: $0.05 \leqslant r_e[''] \leqslant 1$, $0.2 \leqslant n \leqslant 7$, and $0.1 \leqslant q \leqslant 1$. We allowed GALFIT to fit a constant sky background as a free parameter. Previous studies have shown this to be the preferred choice and that GALFIT performs significantly better when allowed to internally measure a sky background, as opposed to being provided a fixed background (Haussler et al. 2007; Cutler et al. 2022). Furthermore, the convolution box was allowed to span the whole cutout.

For each $K_S$-band object fit, two to three nearby, unsaturated, uncontaminated, and background-subtracted stars were used as point-spread functions (PSFs) for model convolution. We also adopted as the model PSF a high-S/N PSF constructed using 10 different nearby stars, stacking the corresponding sky-subtracted stamps after masking any nearby objects, recentering the stars, and normalizing the integrated flux. Utilizing different stars/PSFs allows for a more realistic estimate of the size measurement error, which is generally underestimated by GALFIT. For the *WFC*3/F160W images, a position-dependent PSF model was created using `grizli` (https://grizli.readthedocs.io) to shift and drizzle HST empirical PSFs (Anderson et al. 2015) at the position of the UMGs. Cutler et al. (2022) showed that there is no significant difference in GALFIT structural measurements between galaxies fit with position-dependent PSFs and those with PSFs determined over a larger area of the mosaic, even at $z > 2$. While most of the galaxies are not resolved in the ground-based imaging, GALFIT can still recover fits down to FWHM/2 (Häußler et al. 2013; Nedkova et al. 2021), although Ribeiro et al. (2016) suggest such

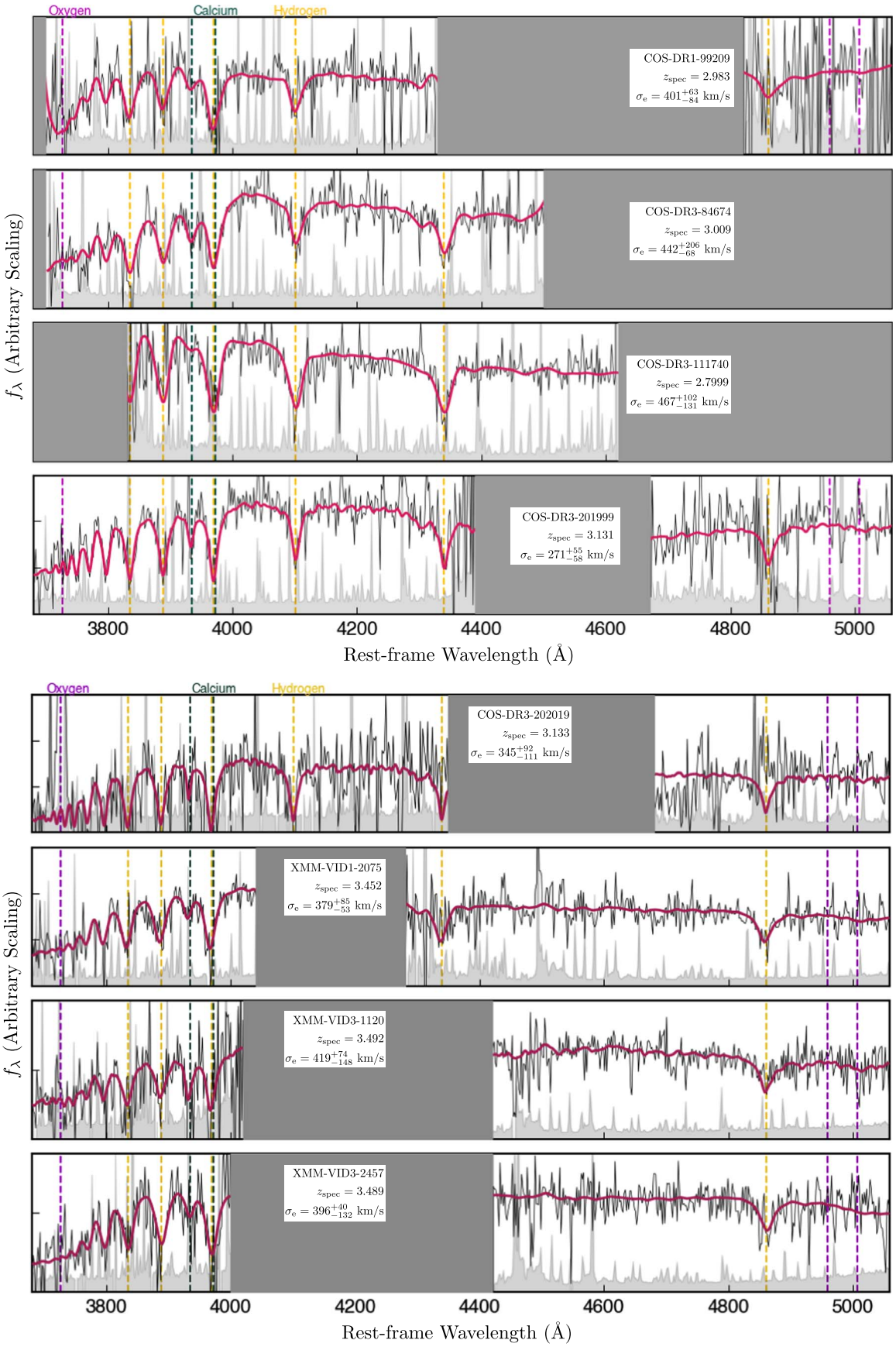


**Figure 1.** The observed *H*- and *K*-band spectra for the MAGAZ3NE UMGs (black) and associated error spectra (gray). The best-fit pPXF model is shown in red. The wavelengths of prominent features from oxygen (magenta), calcium (green), and hydrogen (gold) are indicated as vertical dashed lines. When emission lines are present, we mask these features in the velocity dispersion fit.

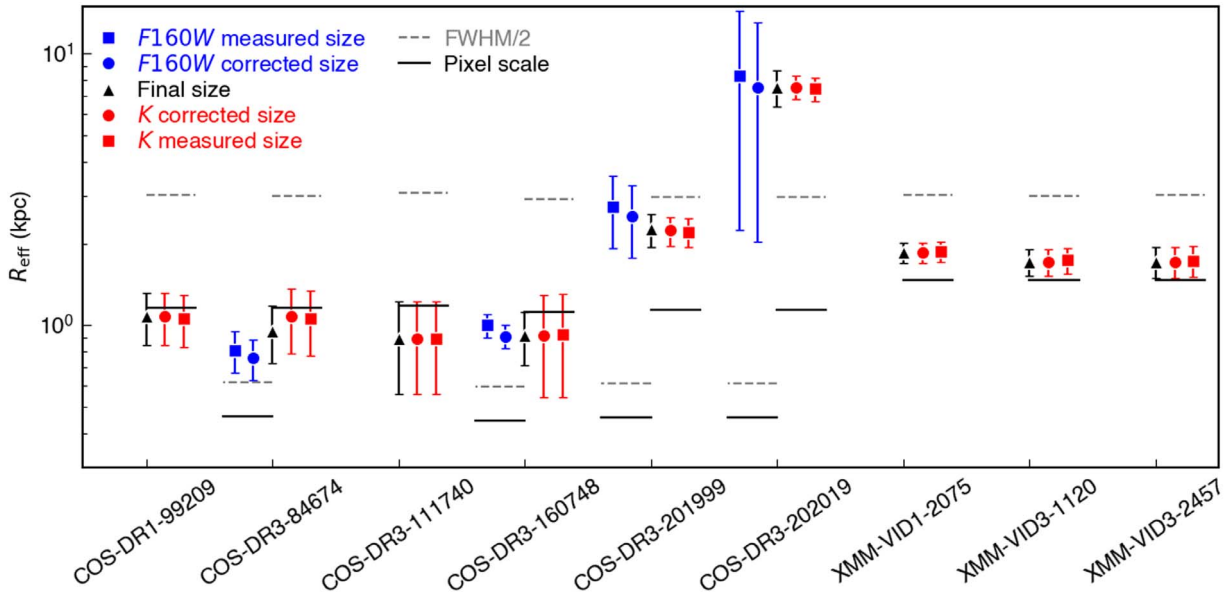


**Figure 2.** Effective radii for the UMGs. Radii measured directly from the $K$ band and $F160W$ imaging are shown as squares colored red and blue, respectively. Correcting these sizes to 5000 Å radius results in circles of the same colors. The weighted average of the corrected measurements (the same as the $K$-band value if $F160W$ data are unavailable) is shown as a black triangle. The pixel size of the imager at the redshift of each target is shown as a black line, while the FWHM/2 is a gray dashed line.

**Table 1**
Properties of Massive Quiescent Galaxies in the $z \gtrsim 3$ Sample Discussed in This Work

| UMG | $z_{\rm spec}$ | $\log(M_*/M_\odot)$ | $\log(M_{\rm dyn}/M_\odot)$ | $\sigma_e$ (km s$^{-1}$) | $r_{\rm eff,5000Å}$ (kpc) | Reference |
|---|---|---|---|---|---|---|
| COS-DR1-99209 | $2.9834^{+0.0023}_{-0.0028}$ | $11.22^{+0.05}_{-0.06}$ | $11.31^{+0.12}_{-0.20}$ | $401^{+63}_{-84}$ | $1.08 \pm 0.24$ | This work |
| COS-DR3-84674 | $3.0094^{+0.0015}_{-0.0011}$ | $11.25^{+0.01}_{-0.02}$ | $11.33^{+0.23}_{-0.14}$ | $442^{+206}_{-68}$ | $0.95 \pm 0.23$ | This work |
| COS-DR3-111740 | $2.7988^{+0.0013}_{-0.0011}$ | $10.98^{+0.01}_{-0.00}$ | $11.02^{+0.13}_{-0.24}$ | $467^{+102}_{-131}$ | $0.89 \pm 0.33$ | This work |
| COS-DR3-201999 | $3.1313^{+0.0014}_{-0.0012}$ | $11.40^{+0.03}_{-0.01}$ | $11.28^{+0.15}_{-0.24}$ | $271^{+55}_{-58}$ | $2.26 \pm 0.31$ | This work |
| COS-DR3-202019 | $3.1326^{+0.0021}_{-0.0011}$ | $11.67^{+0.04}_{-0.05}$ | $12.00^{+0.14}_{-0.27}$ | $345^{+92}_{-111}$ | $7.54 \pm 1.16$ | This work |
| XMM-VID1-2075 | $3.4520^{+0.0014}_{-0.0017}$ | $11.52^{+0.00}_{-0.05}$ | $11.49^{+0.12}_{-0.11}$ | $379^{+85}_{-53}$ | $1.85 \pm 0.16$ | This work |
| XMM-VID3-1120 | $3.4919^{+0.0018}_{-0.0029}$ | $11.47^{+0.02}_{-0.03}$ | $11.54^{+0.10}_{-0.31}$ | $419^{+74}_{-148}$ | $1.71 \pm 0.19$ | This work |
| XMM-VID3-2457 | $3.4892^{+0.0032}_{-0.0024}$ | $11.26^{+0.02}_{-0.03}$ | $11.49^{+0.07}_{-0.29}$ | $396^{+40}_{-132}$ | $1.71 \pm 0.22$ | This work |
| ZF-COS-20115 | 3.715 | $11.06^{+0.06}_{-0.04}$ | $10.86^{+0.14}_{-0.20}$ | $283^{+52}_{-52}$ | $0.66 \pm 0.08$ | Esdaile et al. (2021) |
| 3D-EGS-40032 | 3.219 | $11.31^{+0.03}_{-0.03}$ | $11.41^{+0.11}_{-0.16}$ | $275^{+56}_{-56}$ | $2.40 \pm 0.19$ | Esdaile et al. (2021) |
| 3D-EGS-18996 | 3.239 | $10.99^{+0.02}_{-0.03}$ | $10.56^{+0.13}_{-0.19}$ | $196^{+48}_{-48}$ | $0.63 \pm 0.05$ | Esdaile et al. (2021) |
| 3D-EGS-31322 | 3.434 | $10.99^{+0.05}_{-0.04}$ | $10.85^{+0.20}_{-0.39}$ | $201^{+119}_{-119}$ | $0.61 \pm 0.05$ | Esdaile et al. (2021) |
| C1-23152[a] | $3.352^{+0.002}_{-0.002}$ | $11.30^{+0.19}_{-0.13}$ | $11.34^{+0.07}_{-0.09}$ | $409^{+60}_{-60}$ | $1 \pm 0.1$ | Saracco et al. (2020) |
| SXDS-27434 | $4.0127^{+0.0005}_{-0.0005}$ | $11.06^{+0.04}_{-0.04}$ | $<11.32$ | $268^{+59}_{-59}$ | $<1.3$ | Tanaka et al. (2019) |

**Note.**
[a] This object is renamed COS-DR3-160748 in the MAGAZ3NE Survey.

measurements tend to be underestimates. Additionally, the sizes derived from the unresolved $K_S$ band and the resolved $WFC3$/F160W GALFIT modelings are consistent with each other, as shown in Figure 2.

The S/N of the images is not sufficiently high to obtain a reliable value of the Sérsic index. As the size and Sérsic index are covariant in the fitting process, we also use GALFIT to perform fits with the Sérsic index fixed to $n = 1$, 3, 4, and 6 and compare to the reported best fit, in which $n$ is allowed to vary, to discern another source of possible error on the size measurement. In some cases, these fits do not converge, and in some, the reported fit is clearly incorrect upon visual inspection. Ignoring these cases, we find that objects with a best-fit Sérsic index $2 < n < 4$ from ground-based $K$-band imaging show size variations on the order of 10% in these tests. In the other cases, variation on the order of up to 20% is seen. For all UMGs, these variations are smaller than the reported errors based on different characterizations of the PSF.

For galaxies with imaging in both bandpasses, the measured sizes are consistent within the errors. However, the two bandpasses are probing different wavelengths, which can be on opposite sides of the $D_n(4000)$ feature. To avoid any issues on this front, we convert all measured sizes to rest-frame 5000 Å sizes following van der Wel et al. (2014) as

$$r_{\rm eff,5000\ \text{Å}} = r_{\rm eff,\lambda_{obs}} \left( \frac{1+z}{1+z_{\rm pivot}} \right)^{\frac{\Delta \log(r_{\rm eff})}{\Delta \log \lambda}}, \quad (1)$$

where

$$z_{\rm pivot} = \lambda_{\rm obs}/5000\text{Å} - 1, \quad (2)$$

$$\frac{\Delta \log(r_{\rm eff})}{\Delta \log \lambda} = -0.35 + 0.12z - -0.25 \log\left(\frac{M_*}{10^{10} M_\odot}\right). \quad (3)$$

For rest-frame optical sizes (5000 Å), $z_{\rm pivot}(F160W) = 2.2$ and $z_{\rm pivot}(K) = 3.3$. While the $\frac{\Delta \log(r_{\rm eff})}{\Delta \log \lambda}$ relation from van der Wel et al. (2014) was derived using less massive galaxies at $z < 2$, we note that the corrections here are considerably smaller than the errors on the size measurements. Given the consistency of all these half-light radii for a given galaxy, in what follows we use a weighted average of the corrected size measurements in all available bands for the determination of the dynamical mass. This size is listed in Table 1.

We also note that the morphology of COS-DR3-201999 was analyzed in Lustig et al. (2021), with the id 252568, which returned $r_{\rm e}(5000\ \text{Å})/\text{kpc} = 2.37^{+0.58}_{-0.37}$, a size fully consistent with the analysis herein.

### 3.3. Dynamical Masses

The velocity dispersion and effective radius measurements can be used to calculate dynamical masses for the UMGs in this sample,

$$M_{\rm dyn}(< r_{\rm e}) = \kappa_{\rm e} \frac{\sigma_{\rm e}^2 r_{\rm e}}{G}, \quad (4)$$

where $\kappa_{\rm e}$ is a virial coefficient that depends upon the (an) isotropy of the stellar velocities and the intrinsic mass profile of the galaxy. This value has been calibrated using lower-redshift ellipticals, as such determinations for high-redshift, compact quiescent galaxies have not been done due to their small sizes and faint magnitudes. The typical value used for $z \sim 2$ quiescent galaxies is $\kappa_{\rm e} = 2.5$ (Newman et al. 2012; Barro et al. 2014). The resultant value of $M_{\rm dyn}(< r_{\rm e})$ is then doubled to estimate the total $M_{\rm dyn}$, which is then compared to the total stellar mass.

Cappellari et al. (2006) also published an analytical estimator, which folds in both the virial coefficient and the correction to total mass,

$$M_{\rm dyn} = \beta(n) \frac{\sigma_{\rm e}^2 r_{\rm e}}{G}, \quad (5)$$

$$\beta(n) = 8.87 - -0.831n + 0.0241n^2. \quad (6)$$

For a sample of massive, quiescent galaxies at $z \sim 2$, a typical value of $\beta(n) \sim 5$ is found, which is equivalent to the choice of $\kappa_{\rm e} = 2.5$ (van de Sande et al. 2013; Belli et al. 2014a). Previous samples of UMGs at $z \gtrsim 3$ (Esdaile et al. 2021; Saracco et al. 2020) have used this estimator and returned values in the range of $5.4 < \beta(n) < 6.4$, while Tanaka et al. (2019) also adopt $\beta(n) = 5$ due to a lack of a confident measure of the Sérsic index.

In this work we also adopt the value of $\beta(n) = 5$, as the S/N of the images used for size calculations is not sufficiently high to obtain a reliable value of the Sérsic index. Results of these calculations are provided in Table 1.

## 4. Results and Discussion

We compare our results to massive, quiescent galaxies at a range of redshifts. The first sample, from Posacki et al. (2015), is a reanalysis of 55 massive early-type galaxies at $z \sim 0.2$ from SLACS (Treu et al. 2010) and a subset of 223 H$\beta$ massive absorption line galaxies in the local volume from ATLAS$^{3D}$ (Cappellari et al. 2013b). Galaxies selected from SDSS with velocity dispersions $\sigma > 350$ km s$^{-1}$ at similar redshifts were also compared in an attempt to mitigate progenitor bias (Bernardi et al. 2006; Saracco et al. 2020). Mendel et al. (2020) compiled and reanalyzed spectra from early-type galaxies at $1.4 < z < 2.1$, including spectra presented in Cappellari et al. (2009), Newman et al. (2010), Toft et al. (2012), Bezanson et al. (2013a), van de Sande et al. (2013), Belli et al. (2014a, 2014b), Barro et al. (2016), and Belli et al. (2017). In addition to our eight $z \gtrsim 3$ UMGs, we fold in six previously published $z > 3$ UMGs with velocity dispersion measurements: SXDS-27434 (Tanaka et al. 2019), C1-23152 (published by members of our group in Saracco et al. (2020) and subsequently renamed as COS-DR3-160748 in the context of the MAGAZ3NE Survey; Forrest et al. 2020b), ZF-COS-20115, 3D-EGS-40032, 3D-EGS-18996, and 3D-EGS-31322 (Esdaile et al. 2021).

Galaxies in these works have also been studied spectroscopically in Valentino et al. (2020); Marsan et al. (2015) and Forrest et al. (2020b); and Glazebrook et al. (2017) and Schreiber et al. (2018a), respectively. For the most part, the massive galaxies at $z > 1.4$ were selected for spectroscopic follow-up via a combination of magnitude/stellar mass, color/SFR, and photometric redshift cuts. Nonetheless, it is important to keep in mind that these cuts are not identical given the different survey depths, photometric wavelength coverage, and photometric redshift tools. Thus, it is possible that studies are selecting different subpopulations of massive quiescent galaxies.

In Figure 3, we show the rest-frame colors, stellar masses, and star formation rates of the objects in the $z \gtrsim 3$ sample. Most of the galaxies are consistent with recently quenched post-starburst galaxies, as they lie in the lower left of the *UVJ* quiescent wedge or slightly blueward of it and show SFRs significantly below the main sequence for their mass at this redshift. The two notable exceptions to this are COS-DR3-202019 and SXDS-27434 (Tanaka et al. 2019). The former has SFR $= 82\ M_\odot$ yr$^{-1}$ and is the reddest and most massive of the sample, consistent with a dusty star-forming galaxy (Forrest et al. 2020b), while the latter has SFR $= 24\ M_\odot$ yr$^{-1}$ and has the bluest $(U-V)_{\rm REST}$ color of the sample (Valentino et al. 2020).

### 4.1. Large Velocity Dispersions

The best-fit velocity dispersions for the MAGAZ3NE sample are very large, at ~400 km s$^{-1}$. Nonetheless, several galaxies at $z \sim 2$ have previously been measured with similarly high velocity dispersions (van Dokkum et al. 2009a; van de Sande et al. 2013; Belli et al. 2014b, 2017). These velocity dispersions confirm the large stellar masses of these objects while being independent of the various problems intrinsic to SED fitting such as choice of IMF and contamination by emission lines (see Section 4.3.4).

A positive correlation between the stellar velocity dispersion and stellar mass is expected as the mass within the small sizes over which we probe the stellar velocity dispersion is dominated by stars. At $1.4 < z < 2.1$, the data compiled in Mendel et al. (2020) show a positive correlation between the two, though individual galaxies show significant scatter. A least-squares regression to the entire $z \gtrsim 3$ sample shows a vertical offset toward larger velocity dispersions at a given stellar mass, but a similar slope to both the $1.4 < z < 2.1$

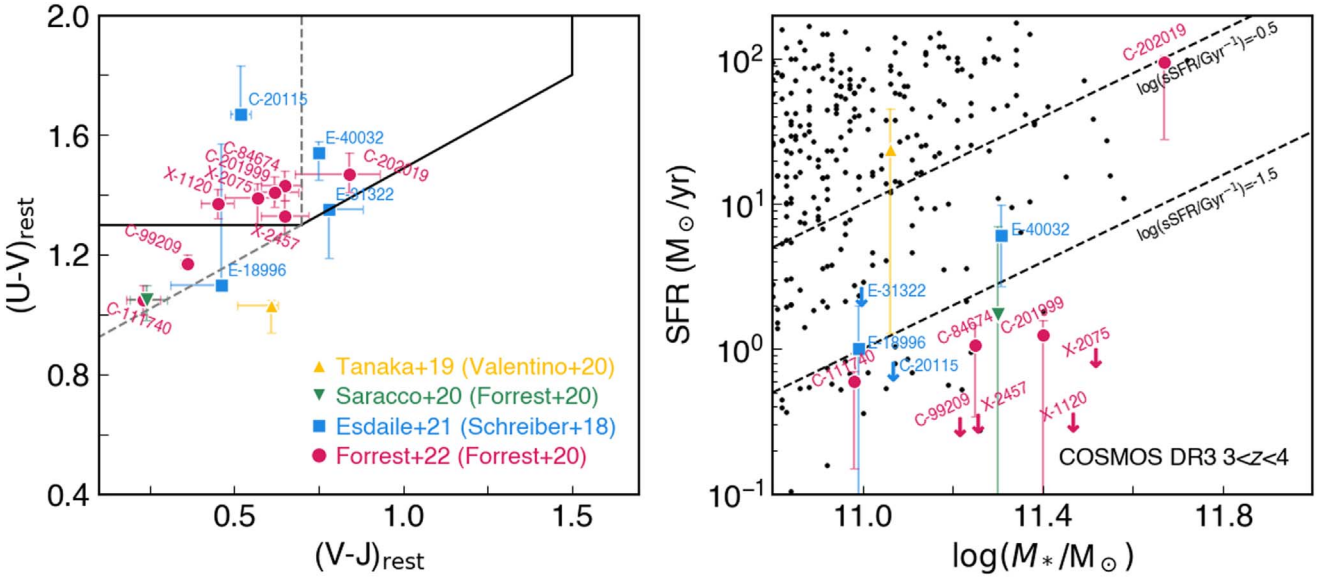


**Figure 3.** Properties of the massive quiescent galaxies in the $z \gtrsim 3$ sample that have published stellar velocity dispersions. Left: the rest-frame color *UVJ* diagram. Galaxies are labeled here and in subsequent plots by the reference containing stellar velocity dispersion measurements. Rest-frame colors, star formation rates, and/or stellar masses are often from separate publications, also noted. Data at $z \gtrsim 3$ are shown as a gold upward-facing triangle (Tanaka et al. 2019), a seagreen downward-facing triangle (Saracco et al. 2020), azure squares (Esdaile et al. 2021), and raspberry circles (this work). Right: the relation between the star formation rate and stellar mass. Galaxies with photometric redshifts $3 < z < 4$ from the COSMOS-UltraVISTA DR3 catalog are shown as black points, and dashed lines are plotted showing a constant specific SFR ($-1.5$ and $-0.5$ Gyr$^{-1}$), corresponding roughly to the main sequence and 1 dex below it.

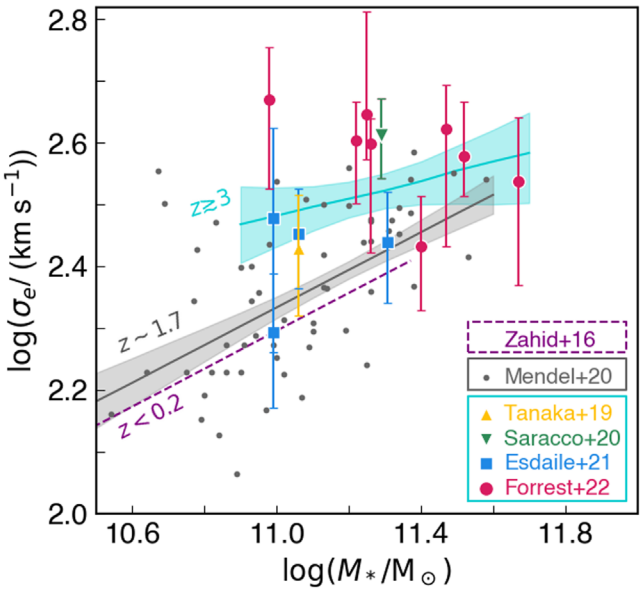


**Figure 4.** Stellar velocity dispersions and stellar masses of the massive quiescent galaxies. The colors and markers used for the $z \gtrsim 3$ sample are identical to those in Figure 3, while data from Mendel et al. (2020) at $1.4 < z < 2.1$ are shown in gray. The median error bar for the Mendel et al. (2020) data set is ∼0.1 dex. Fits to both sets of data are shown as turquoise and gray lines, respectively, with shaded error regions representing the 16th to 84th percentile range from Monte Carlo resampling. The fit to low-redshift data from Zahid et al. (2016) is shown in purple.

sample and a sample of massive quiescent galaxies from SDSS (Zahid et al. 2016), shown in Figure 4.

### 4.2. The Size–Stellar Mass Relation

At a given epoch, the effective radius and stellar mass of a galaxy are also correlated, though quiescent and star-forming galaxies tend to follow different relations, and those relations evolve with time to smaller sizes for a given stellar mass at earlier times (e.g., van der Wel et al. 2014; Straatman et al. 2015b; Mowla et al. 2019; Marsan et al. 2019).

Indeed, the $1.4 < z < 2.1$ sample from Mendel et al. (2020) is in agreement with the $z \sim 1.75$ relation for early-type galaxies presented in van der Wel et al. (2014) using data from 3D-HST:

$$r_e/\mathrm{kpc} = 1.23 \times (M_*/5 \times 10^{10} M_\odot)^{0.76},$$

or equivalently,

$$\log(r_e/\mathrm{kpc}) = -8.04 + 0.76 \log(M_*/M_\odot).$$

From Monte Carlo resampling of the $z \gtrsim 3$ galaxies with SFR $< 3\ M_\odot\ \mathrm{yr}^{-1}$, we find

$$\log(r_e/\mathrm{kpc}) = -9.73(\pm 1.50) + 0.87(\pm 0.15) \log(M_*/M_\odot),$$

that is, smaller sizes for a given stellar mass showing a statistically consistent, but perhaps slightly steeper, relation with stellar mass (see Figure 5).

Relative to the $z \sim 1.75$ relation, this $z \gtrsim 3$ fit shows smaller sizes by a factor greater than 3 at $\log(M_*/M_\odot) \sim 11$ and a factor of $\gtrsim 2$ at $\log(M_*/M_\odot) \sim 11.5$, which also agrees with the redshift–size evolution shown in Straatman et al. (2015b). Limiting the $z \gtrsim 3$ sample to quiescent galaxies with HST/*WFC3* imaging does not significantly change the best-fit relation, though including the galaxies with SFR $> 3\ M_\odot\ \mathrm{yr}^{-1}$ does result in a steeper slope.

### 4.3. Comparison of Dynamical Mass and Stellar Mass

The dynamical and stellar masses for the $z \gtrsim 3$ sample are listed in Table 1 and shown in Figure 6. For massive, quiescent galaxies with little gas or dust and small sizes, the dynamical and stellar masses are expected to be quite similar because the central regions are dominated by baryons with little dark matter contribution. The most obvious exception to this in the

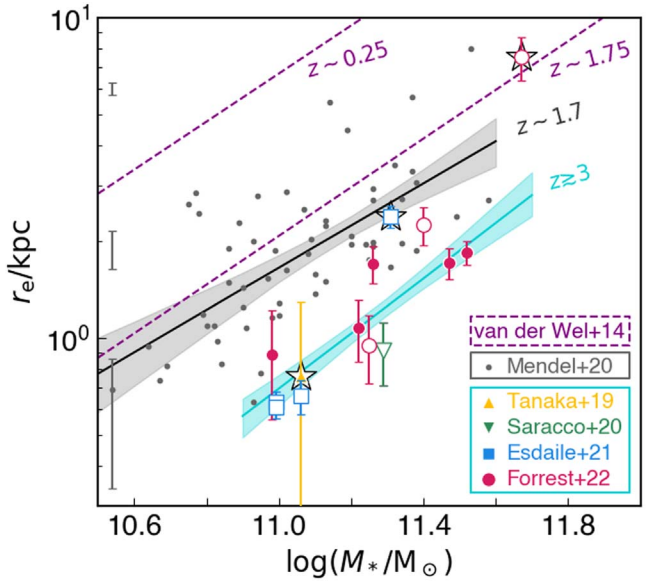


**Figure 5.** Effective radii and stellar masses of the massive galaxies. Colors and markers remain the same as in Figure 4, and objects with sizes derived from high-resolution HST imaging have white centers. Representative error bars for the Mendel et al. (2020) data set are shown on the left. The three galaxies with SFR $> 3\ M_\odot\ \mathrm{yr}^{-1}$ are marked with a black star and excluded from the $z \gtrsim 3$ fit. The size–mass relations for early-type galaxies at $z \sim 0.25$ and $z \sim 1.75$ from van der Wel et al. (2014) are shown as purple dashed lines, while fits to the Mendel et al. (2020) data set and the $z \gtrsim 3$ data set are a black line and a solid turquoise line with the range of Monte Carlo fits from 16%–84% shaded accordingly.

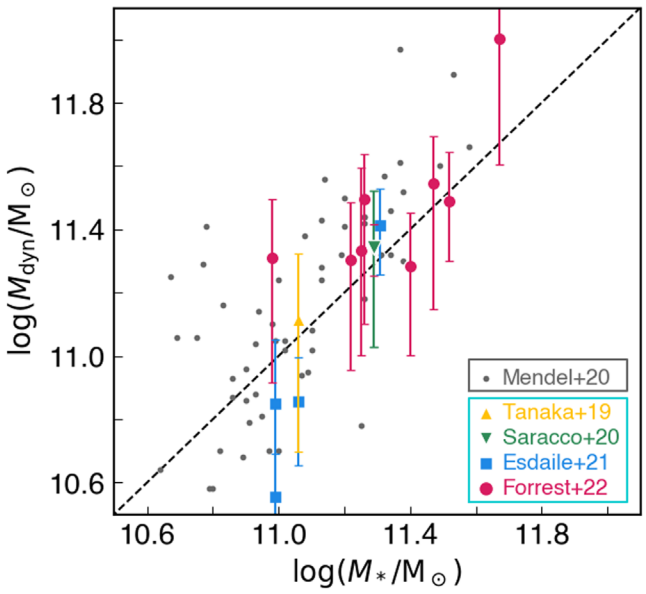


**Figure 6.** Dynamical and stellar masses of the massive quiescent galaxies. The colors and markers used are identical to those in Figure 5. The black dashed line represents a ratio of unity, corresponding to a Chabrier IMF.

MAGAZ3NE sample is COS-DR3-202019 (the most massive galaxy in the sample), which has a radius ∼3× larger than any other galaxy in the sample and is also the only one that shows evidence of ongoing star formation (see Figure 3) but is still consistent with a one-to-one ratio between stellar and dynamical mass within $1\sigma$. The consistency of the $z \gtrsim 3$ sample's ratios of dynamical to stellar mass, $M_{\rm dyn}/M_*$, with unity suggests that the Chabrier IMF used to derive the stellar masses for these objects is in general reasonable.

While similarly massive galaxies at lower redshifts appear to prefer heavier IMFs (e.g., Conroy & van Dokkum 2012; Cappellari et al. 2013a; Zahid & Geller 2017), at $z \sim 1.7$, Mendel et al. (2020) also find that a lighter IMF such as the Chabrier IMF is required to prevent stellar masses from exceeding dynamical masses. Dynamical masses in significant excess of stellar masses would be expected if either the choice of IMF is incorrect or if there is an appreciable fraction of dark matter in the galaxy. We note that the contribution of dark matter for similar galaxies at lower redshift, ∼5%–20% (Cappellari et al. 2013b; Mendel et al. 2020), is too small to be quantified here given the observational errors involved.

That said, a comparison of $M_{\rm dyn}/M_*$ to stellar velocity dispersion can still yield important insights. For instance, high-redshift quiescent galaxies have lower ratios of $M_{\rm dyn}/M_*$ for a given velocity dispersion than galaxies at lower redshifts (van de Sande et al. 2013; Hill et al. 2016; Belli et al. 2017; Mendel et al. 2020; Esdaile et al. 2021), which is suggestive of a preference for a lighter IMF in such systems in early times. While our data do not allow for significant constraints on dark matter content or IMF form for individual galaxies, a combination of the eight new MAGAZ3NE galaxies presented here with the four UMGs from Esdaile et al. (2021), one from Saracco et al. (2020), and one from Tanaka et al. (2019) allows a first look at these properties using a statistical sample at $z \gtrsim 3$, shown in Figure 7.

We perform a linear regression between the logarithm of $M_{\rm dyn}/M_*$ and the logarithm of the velocity dispersion at the effective radius for our sample, as well as those at $z \sim 0.2$ and $z \sim 1.7$. Additionally, we use Monte Carlo resampling (accounting for the correlated errors) to characterize the uncertainties on the resulting best fits:

$$\log(M_{\rm dyn}/M_*)_{z\sim0.2} = (0.29 \pm 0.02) + (0.40 \pm 0.05) \times \log(\sigma_e/350), \quad (7)$$

$$\log(M_{\rm dyn}/M_*)_{z\sim1.7} = (0.30 \pm 0.06) + (1.25 \pm 0.20) \times \log(\sigma_e/350), \quad (8)$$

$$\log(M_{\rm dyn}/M_*)_{z\gtrsim3} = (0.03 \pm 0.04) + (1.29 \pm 0.36) \times \log(\sigma_e/350). \quad (9)$$

The best-fit slope at $z \gtrsim 3$ ($1.29 \pm 0.36$) is consistent with that of the fit at $z \sim 1.7$ ($1.25 \pm 0.20$) and significantly steeper than the low-redshift relation ($0.40 \pm 0.05$). Additionally, the $z \gtrsim 3$ sample is offset to lower $M_{\rm dyn}/M_*$ by ∼0.3 dex relative to the $z \sim 1.7$ sample and ∼0.5 dex relative to the low-redshift sample for a given velocity dispersion. This means that while the $z \gtrsim 3$ sample shows the same trend of preferring a heavier IMF at higher velocity dispersions relative to lower velocity dispersions, many of the highest velocity dispersion objects prefer a Chabrier IMF (or an IMF lighter than Chabrier) to a Salpeter IMF (see Figure 7). In order for high velocity dispersion galaxies to prefer a bottom-heavy IMF such as Salpeter or even heavier (e.g., Conroy & van Dokkum 2012), at least one of several parameters must be systematically incorrect and provide a 0.2 dex (∼60%) gain in $M_{\rm dyn}/M_*$, addressed below.

#### 4.3.1. Are the High-velocity Dispersions Too Low?

The reported velocity dispersions herein are some of the largest measured (see also van Dokkum et al. 2009a;

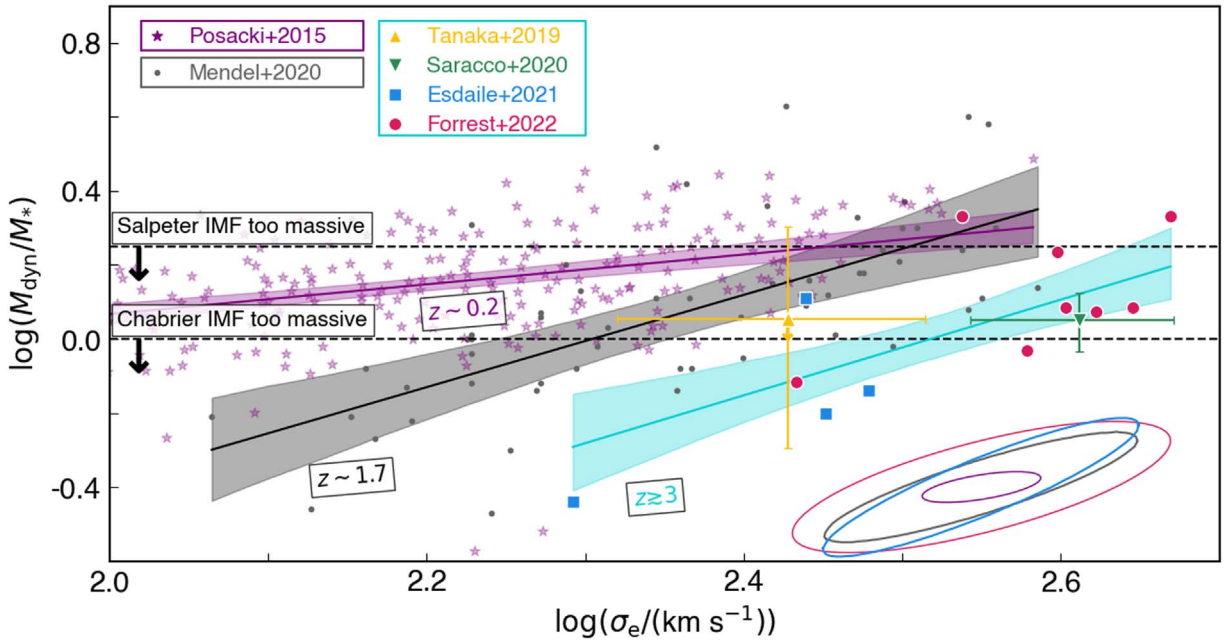


**Figure 7.** Ratio of dynamical-to-stellar mass (IMF mismatch parameter) plotted against velocity dispersion at the half-light radius. The best fit to the data from Posacki et al. (2015) at $z \sim 0.2$ is shown in purple, while the best fit to the data from Mendel et al. (2020) at $1.4 < z < 2.1$ is shown in gray. Data at $z > 3$ are shown in the same colors as the previous plot, while the fit is in turquoise. Average error ellipses for the data sets from Posacki et al. (2015), Mendel et al. (2020), Esdaile et al. (2021), and this work are shown in the bottom right. Horizontal dashed lines show the result when using a Salpeter or Chabrier IMF, below which using the respective IMF results in a stellar mass greater than the dynamical mass.

van de Sande et al. 2013; Saracco et al. 2020 for other galaxies with $\sigma > 400$ km s$^{-1}$). To reach agreement with a Salpeter IMF, the velocity dispersions would have to be even higher by $\sim$100 km s$^{-1}$ for the highest velocity dispersion objects (and $\sim$500 km s$^{-1}$ for those galaxies with lower velocity dispersions). This increase is perhaps not unrealistic for some galaxies here, given the errors on the measured velocity dispersions. Intriguingly, this is in line with the large velocity dispersion of 510 km s$^{-1}$ measured for a massive compact galaxy at $z = 2.2$ in van Dokkum et al. (2009b). Of course, while we have performed a robust investigation into the possible systematics involved in the calculation of velocity dispersions for this sample (see Appendix B), the fact remains that the systematics may contribute to the results.

Another complicating factor here is the possibility of significant rotation in these systems, which would make the use of the measured velocity dispersion in the calculation of dynamical mass incorrect. Several massive, quiescent galaxies at $z \sim 2$ are disk dominated and have been confirmed to have significant rotation, thanks to gravitational lensing (Toft et al. 2017; Newman et al. 2018). Resolving rotation is not possible with our data. Measured velocity dispersions could be inflated by a rotational component if a spectral slit is oriented with the major axis of the disk or could be underestimated if the spectral slit is misaligned. Our sample is not large enough to claim that these effects cancel each other out on average.

#### 4.3.2. Are the Size Measurements Too Small?

The GALFIT package used is widely used and appears to be generally accurate in calculating sizes. Several of the objects in the $z \gtrsim 3$ sample are not resolved in ground-based imaging, which can lead to incorrect size estimates below FWHM/2, particularly if the PSF is not well determined (Häußler et al. 2013; Nedkova et al. 2021). Fortunately, the agreement between sizes calculated from HST and ground-based imaging indicates that the sizes are reliable. To find agreement with the $z \sim 1.75$ relations from van der Wel et al. (2014) and the Mendel et al. (2020) data set, the sizes must be 2×–4× larger than measured. To improve consistency with a Salpeter IMF, the sizes must be underestimated by $\sim$30%, which is considered unlikely as objects that are barely resolved are more likely to have their sizes overestimated.

#### 4.3.3. Are the Dynamical Masses Calculated Appropriately?

The calculation of dynamical mass, in addition to relying upon accurate size and stellar velocity dispersion measurements, also contains a factor to account for the distribution of mass in the system. The standard transformation used is a function of Sérsic index, $n$. While the imaging used to calculate sizes is not deep enough to reliably recover a Sérsic index (this usually requires S/N $\sim$3× deeper than that required for size measurements; Haussler et al. 2007; Häußler et al. 2013), the assumption of $\beta(n) = 5$ corresponds to $n \sim 5.5$. An increase of $\sim$60% in this factor would require $n \sim 1.2$, typically seen in larger galaxies with well-developed disks, which remains a possibility.

The correction factor $\beta(n)$ was originally derived using low-redshift elliptical galaxies, and while it shows great precision across $2 < z < 10$ (Bertin et al. 2002; Cappellari et al. 2006), it is possible such a transformation is not accurate for these galaxies for some reason. Deeper, higher-resolution imaging as may be obtained with JWST may allow for insights into this possibility.

#### 4.3.4. Are the Stellar Masses Accurate?

Mass inaccuracies can be caused by the presence of strong emission lines, which can cause an overestimate of up to $\sim$0.5 dex (e.g., Stark et al. 2013; Salmon et al. 2015; Forrest et al. 2017).

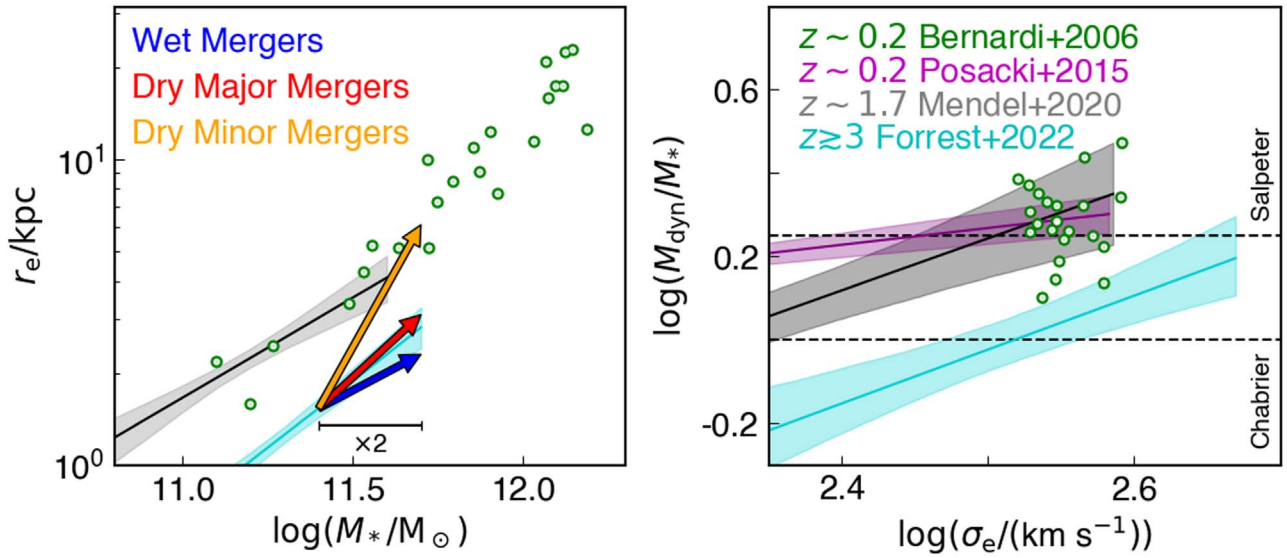


**Figure 8.** Similar to Figures 5 and 7, with additional low-redshift, high velocity dispersion ETGs (green circles; Bernardi et al. 2006). The arrows on the size–stellar mass relation (left) show the effect on those properties via various merger scenarios, normalized to a doubling of the stellar mass.

However, Forrest et al. (2020b) model stellar masses for the MAGAZ3NE galaxies in this sample after correcting broadband photometry for any strong emission lines seen in the spectra ([OII], [OIII], H$\beta$), though of the sample here only COS-DR3-202019 shows significant emission. The only other strong line that could normally be an issue is H$\alpha$, though at the redshifts of the sample, this line falls in between the $K$ band and the IRAC 3.6$\mu$m bandpass, and so should not affect the photometry either.

It is also known that the choice of modeling parameters and program can lead to differences in stellar mass calculations of around 0.2 dex (e.g., Mobasher et al. 2015). Leja et al. (2019) compare stellar masses for objects in the 3D-HST study derived using FAST (Kriek et al. 2009) and Prospector-$\alpha$ (Leja et al. 2017), and while they find a systematic offset of up to 0.4 dex in stellar mass, these differences appear to be $<$0.1 dex for high-mass galaxies at high redshifts, as is the case for our sample. Regardless, the Prospector-$\alpha$ code outputs higher stellar masses than FAST, and thus, any such offset would only increase the tension with, e.g., a Salpeter IMF.

The possibility of young stars outshining older populations in a spectrum and leading to a light-weighted stellar mass different from the true stellar mass would similarly result in an underestimate of the stellar mass.

Of course, the calculation of stellar masses rests upon the modeling of stellar populations, often based on the spectra of local stars. It is possible that these model populations are not applicable to stellar populations in the early universe. A test of this possibility would require high-resolution stellar spectra at high redshifts and is not currently technologically feasible.

### 4.4. Evolutionary Insights

The high velocity dispersions presented here for the $z \gtrsim 3$ sample support the large stellar masses calculated through SED modeling for massive, high-redshift galaxies and suggest that SED modeling of large photometric samples can be trusted to first order, outliers notwithstanding. We also note, however, that the velocity dispersions do not support much more mass than the stellar mass, implying that the contribution of dark matter at the centers of these compact galaxies is small.

#### 4.4.1. Progenitor Bias

Galaxies with velocity dispersions such as those measured for some of our UMGs are exceedingly rare in the local universe. While analyzing the apparent trends seen in previous figures, we must carefully consider factors such as progenitor bias as well as the possibility that descendants of the rare $z \gtrsim 3$ UMGs do not exist in the limited local volume. In an attempt to mitigate these effects, we compare the UMGs in the $z \gtrsim 3$ sample to an additional sample of massive low-redshift ETGs, which are among the most massive galaxies in SDSS and which are not actively forming stars (Bernardi et al. 2006; Saracco et al. 2020). We correct the published velocity dispersions, originally corrected to $r_e/8$, and transform them to $r_e$ using the relation from Jorgensen et al. (1995),

$$\frac{\sigma_{ap}}{\sigma_e/8} = \left(\frac{r_{aper}}{r_e/8}\right)^{-0.04}, \quad (10)$$

which was used for the original correction in Bernardi et al. (2006; though see the discussion about issues with this method for ETGs in La Barbera et al. 2019). This corresponds to a correction factor of $8^{-0.04} = 0.92$. Galaxies in this sample have larger stellar masses, velocity dispersions, and dynamical-to-stellar mass ratios than most of the $z \sim 0.2$ sample from Posacki et al. (2015; Figure 8). Velocity dispersion is known to correlate well with age for SDSS ETGs (e.g., Van Der Wel et al. 2009; Zahid & Geller 2017), and we thus assume that the stellar populations of these galaxies are also quite old. We note that making cuts to the galaxy samples herein by stellar mass or velocity dispersion do not result in qualitative changes to our conclusions.

Spatially resolved studies of local massive ETGs have shown that stellar populations at their cores appear to be older than stars on the outskirts, as well as being regions with higher velocity dispersions (e.g., van Dokkum et al. 2017; La Barbera et al. 2019), consistent with the bulk of star formation occurring at $z \gtrsim 2$, followed by passive evolution via gas-poor (dry) mergers. Dry major mergers, having a mass ratio between the two galaxies close to unity, increase both stellar mass and

radius at similar rates with minimal new star formation (i.e., retaining an old stellar age) and without much change in velocity dispersion (e.g., Hopkins et al. 2009). Dry minor mergers on the other hand are expected to increase the effective radius approximately twice as fast as the stellar mass while also decreasing velocity dispersion slightly, though the cores of these galaxies could still retain high velocity dispersions (e.g., Bezanson et al. 2009; Saracco et al. 2020).

As seen in the left panel of Figure 8, the passive evolution of the $z \gtrsim 3$ UMGs via dry minor mergers could lead to galaxies with sizes, stellar masses, and velocity dispersions of some of the most massive galaxies in SDSS (Bernardi et al. 2006). While the $z \gtrsim 3$ UMGs could evolve into galaxies at the massive end of the $z \sim 1.7$ sample via dry minor mergers, those galaxies in the $z \sim 1.7$ sample with lower stellar masses and velocity dispersions descend from galaxies with different properties than the $z \gtrsim 3$ UMGs. In particular, these progenitors have lower stellar masses and are possibly still forming stars at $z \sim 3$. Similarly, the $z \sim 1.7$ sample could plausibly evolve into galaxies in the Bernardi et al. (2006) sample, but only those with larger velocity dispersions.

Most of the $z \gtrsim 3$ UMGs herein are compact, post-starburst galaxies. A gas-rich (wet) merger may have triggered such a burst of star formation in situ, thus boosting the stellar mass significantly while keeping the effective radius small in contrast to the dry merger scenarios above (e.g., Hopkins et al. 2009). Major mergers, wet or dry, are expected to be few in number for massive galaxies, and it is perhaps the case that the $z \gtrsim 3$ UMGs have simply undergone additional major mergers relative to the progenitors of the lower-mass half of the $z \sim 1.7$ sample. If so, the possibility exists that further major mergers would evolve these galaxies into systems more massive than any in the local volume.

Regardless, while the evolution of sizes, stellar masses, and velocity dispersions can be explained with mergers, the dynamical-to-stellar mass ratio is less easily explained. If dynamical mass is calculated as $M_{\rm dyn} \propto r_{\rm e}\sigma_e^2$, then $\Delta \log(M_{\rm dyn}/M_*) < 0.05$ for all three merger cases described above. Instead of an offset in dynamical-to-stellar mass at a given velocity dispersion, it may be that the increase in velocity dispersion from wet mergers is causing an offset in velocity dispersion at fixed dynamical-to-stellar mass, and we simply do not have any galaxies with large dynamical-to-stellar mass ratios in our $z \gtrsim 3$ sample. Major mergers can also introduce rotation into the system, making the dynamical mass calculations incorrect.

#### 4.4.2. IMF

Detailed studies of absorption lines in local massive quiescent galaxies have suggested that the cores of these galaxies require a bottom-heavy "super-Salpeter" IMF (e.g., Läsker et al. 2013; Saulder et al. 2015; Conroy et al. 2017). The higher inferred mass-to-light ratios associated with such a bottom-heavy IMF also correlate with velocity dispersion (e.g., Conroy et al. 2013; Cappellari et al. 2013b). This is not only seen in samples of galaxies but also within individual nearby galaxies, where the central cores have larger velocity dispersions and heavier inferred IMFs than the outskirts (e.g., La Barbera et al. 2019). While the sample of $z \gtrsim 3$ galaxies in this study do show a similar trend toward a heavier IMF with higher velocity dispersion, none of the galaxies in our sample show evidence for a "super-Salpeter" IMF and most require a Chabrier IMF in order for the stellar mass to remain below the dynamical mass, despite having very large velocity dispersions. This creates a conundrum, as massive compact systems at high redshift, such as the $z \gtrsim 3$ sample, are generally thought to be the progenitors of the low-redshift, high-mass sample such as that from Bernardi et al. (2006), growing largely through mergers as described above (e.g., Bezanson et al. 2009; Van Der Wel et al. 2009; Saracco et al. 2020; Mendel et al. 2020). Such a picture does not offer a way to significantly change the observed IMF from high-redshift progenitors to the cores of local, massive ETGs.

An alternative possibility is that the IMF is determined by metallicity (e.g., Köppen et al. 2007), which shows a close positive correlation with the inferred IMF slope for local ETGs from IFU data in the CALIFA survey (Martín-Navarro et al. 2015). In this view, massive galaxies at early times undergo gas-rich mergers and form substantial fractions of their stars with gas containing a significant amount of metals from previous generations of stars. This causes new star formation at high metallicity in the $z \gtrsim 3$ UMGs (Saracco et al. 2020), which occurs with a bottom-heavier IMF. Meanwhile, less massive galaxies, being located in less massive halos, are more likely to build up their stellar mass not through merger-induced star formation but by inflows of pristine gas. Further, due to their lower masses, these galaxies lose many of the metals they produce via galactic outflows. This then creates a lower-metallicity environment for star formation, which generates a bottom-lighter IMF. This picture is also consistent with the mass–metallicity relationship seen for star-forming galaxies out to $z > 3$ (e.g., Tremonti et al. 2004; Lian et al. 2018; Sanders et al. 2021). Subsequent growth of massive galaxies via minor mergers then deposits stars from the lower-mass galaxies in the outskirts of the massive galaxy, producing the radial IMF trends seen in spatially resolved data (e.g., van Dokkum et al. 2017; La Barbera et al. 2019).

## 5. Conclusions

We have calculated stellar velocity dispersions and sizes for eight UMGs at $z \gtrsim 3$, more than doubling the sample at this epoch. The high dispersions, on the order of $\sim 400\,{\rm km\,s^{-1}}$, are some of the largest measured, about $1.5\times$ those of galaxies at $z \sim 1.7$ of similar stellar mass. They also agree with the large stellar masses derived from SED fitting, supporting the conclusion that ultramassive quiescent galaxies at $z > 3$ do exist in nonnegligible numbers. Size measurements for these objects additionally show a continuation of the evolution to smaller sizes at higher redshifts, with galaxies of similar stellar mass being about one-third the size of their $z \sim 1.7$ counterparts.

We have used these size and stellar velocity dispersion measurements to calculate the dynamical mass. The ratio of the dynamical-to-stellar mass for these objects shows a trend with velocity dispersion as seen at lower redshifts, though it is offset to higher velocity dispersions/lower mass-to-light ratios. This favors a Chabrier (or even bottom-lighter) IMF for most of the sample and is in tension with the "super-Salpeter" IMFs seen in the cores of the most massive galaxies in the local universe.

The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain. This work also uses data products from observations made with ESO Telescopes at

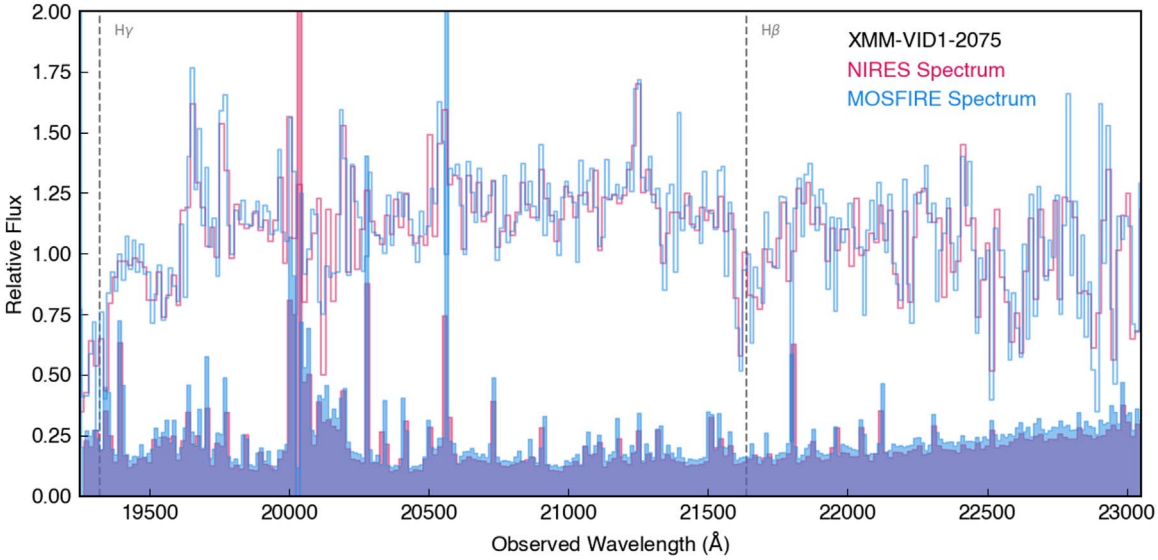


**Figure A1.** Observed *K*-band spectra for XMM-VID1-2075 from NIRES (red) and MOSFIRE (blue). Spectra are binned to ∼3 Å in the rest frame. Errors are represented by the shaded regions at the bottom, and the H$\beta$ and H$\gamma$ absorption features at $z = 3.45$ are labeled.

the La Silla Paranal Observatory under ESO program ID 179. A-2005 and on data products produced by CALET and the Cambridge Astronomy Survey Unit on behalf of the Ultra-VISTA consortium.

This material is based upon work supported by the National Science Foundation under Cooperative Agreement No. AST-2009442. We gratefully acknowledge support from the NASA Astrophysics Data Analysis Program (ADAP) through grant numbers 80NSSC17K0019 and NNX16AN49G, the National Science Foundation through grants AST-1517863 and AST-2205189 and HST program numbers GO-15294 and GO-16300 provided by NASA through grants from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Incorporated, under NASA contract NAS5-26555. Data presented herein were obtained using the UCI Remote Observing Facility, made possible by a generous gift from John and Ruth Ann Evans. Some of the material presented in this paper is based upon work supported by the National Science Foundation under grant No. 1908422. P.S. acknowledges support by the grant PRIN-INAF-2019 1.05.01.85.11.

B.F. also thanks B. Lemaux, A. Pillepich, and L. Lewis for helpful discussions and input, as well as the authors of the codes referenced below, upon which this work has relied heavily. Thanks to the anonymous referee as well, whose comments improved the manuscript.

*Software:* Astropy (Robitaille et al. 2013; Price-Whelan et al. 2018), FAST++ (Schreiber et al. 2018b), `grizli` (ascl.net/1905.001), IPython (Pérez & Granger 2007), Matplotlib (Hunter 2007), Molecfit (Smette et al. 2015; Kausch et al. 2015), MosfireDRP (ascl.net/1908.007), NumPy (Harris et al. 2020), pPXF (Cappellari 2017), PyPeit (Prochaska et al. 2020).

## Appendix A
## Comparison of MOSFIRE and NIRES Spectra for XMM-VID1-2075

One of the UMGs in this work, XMM-VID1-2075, has both a MOSFIRE *K*-band spectrum and an NIRES spectrum, which also includes both *K*-band and *H*-band data with some S/N. The spectra appear quite similar (Figure A1). We fit the *K*-band spectrum from each instrument with the galaxy photometry using FAST++ and compare the results. The redshifts from the two fits are very similar, with $z_{\rm MOSFIRE} = 3.4523$ and $z_{\rm NIRES} = 3.4482$, a difference of ∼0.1%. In both cases, the best fit indicates a galaxy with $\log(M_*/M_\odot) \sim 11.5$, $A_V \sim 0.3$, and age ∼300–500 Myr.

However, including the NIRES *H*-band data while performing the fit results in a slightly older, less massive, less dusty galaxy ($\log(M_*/M_\odot) \sim 11.3$, $A_V \sim 0$, age ∼800 Myr). When each spectrum is fit with pPXF with a set of inputs and assuming the best-fit redshift to that spectrum, the results are statistically consistent. In this work we use the values from the fit to the entire *H*- and *K*-band NIRES spectrum, as this provides a greater number of features for the determination of the velocity dispersion.

## Appendix B
## Dependence of Velocity Dispersions on pPXF Inputs

Due to the low S/N of our spectra (order ∼1/pixel) compared to those on which pPXF was originally tested (order ∼100/pixel), the resultant velocity dispersions can be sensitive to various parts of the fitting mechanism, including choice of templates, additive polynomial order, and wavelength range, among others. Extensive tests along these lines have been performed by van de Sande et al. (2013) for a sample of galaxies at $z \sim 2$, some of which we reproduce for our sample.

### *B.1. Age and Metallicity Template Dependence*

As the spectra herein have low (S/N)/pixel, slightly different templates can yield similar fits to the spectra alone. In particular, the degeneracy between age and metallicity can affect line widths and depths in ways that are difficult to disentangle using a low-S/N spectrum alone. These difficulties can be somewhat alleviated by taking into account the broadband photometry of a galaxy. Similar to van de Sande et al. (2013) and Hill et al. (2016), we test model dependency in this regime using the BC03 models due to their extended

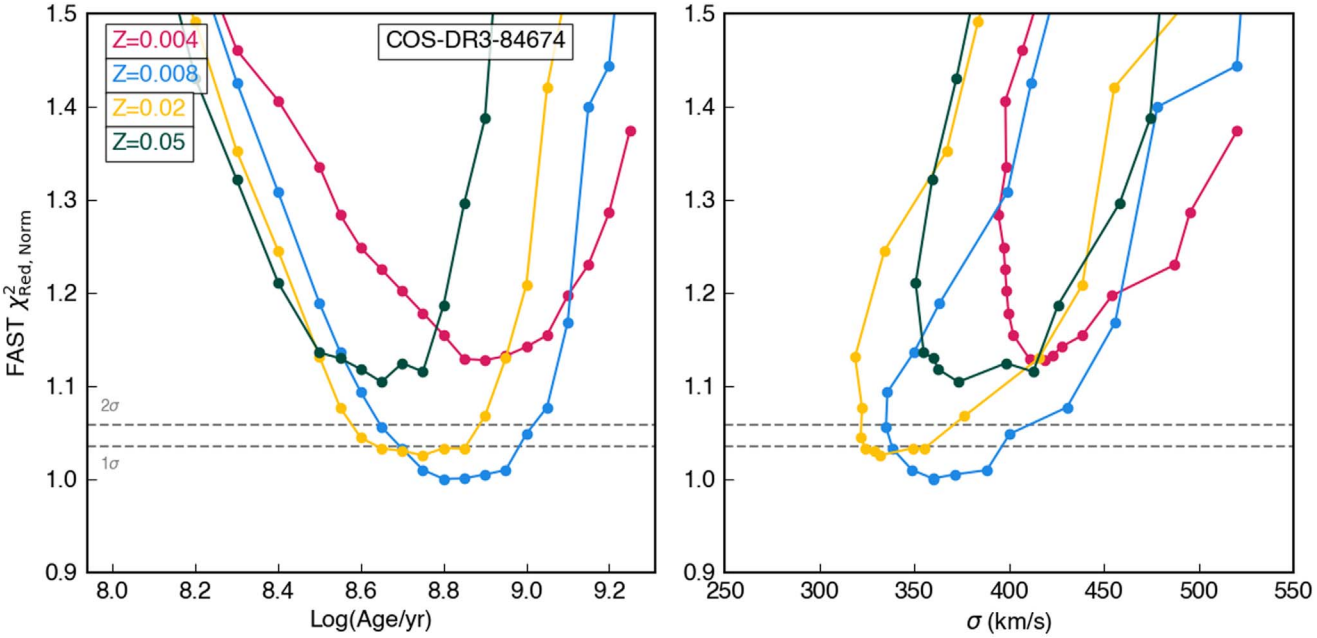


**Figure B1.** The reduced $\chi^2$ for COS-DR3-84674 compared to velocity dispersion for BC03 templates demonstrating the age–metallicity degeneracy. The $\chi^2$ values are taken from the FAST++ joint fit to both photometric and spectroscopic data and are normalized to the lowest value by scaling the input spectral errors. The $1\sigma$ and $2\sigma$ significance given the number of degrees of freedom are indicated by horizontal lines. Each colored line represents templates with a set metallicity, while each point is a different age template.

wavelength coverage. We use FAST++ to fit the spectra and photometry in combination with age and metallicity fixed over a range of values (each combination of $\log(\mathrm{Age/yr}) = [8.0, 8.1, 8.2, 8.3, 8.4, 8.5, 8.55, 8.6, 8.65, 8.7, 8.75, 8.8, 8.85, 8.9, 8.95, 9.0, 9.05, 9.1, 9.15, 9.2, 9.25]$ and $Z = [0.004, 0.008, 0.02$ (solar), $0.05]$). We then use pPXF to fit the velocity dispersion of the galaxy using the best-fit template from each combination of age and metallicity, and compare to the reduced $\chi^2$ value from the FAST++ fit. An example of the results is shown for COS-DR3-84674 in Figure B1. In all cases, the models show clear minima for each choice of metallicity, though in some cases a particular metallicity is not statistically favored. The model with the lowest reduced $\chi^2$ was used for this paper and in subsequent tests. Importantly, this choice is independent of pPXF and therefore also independent of additive polynomial order and spectral wavelength range (see following sections).

### *B.2. Dependence on Additive Polynomial Order*

The pPXF program allows for the addition of a $d$-dimensional Legendre polynomial to a template in order to better match the observed spectrum. This provides better fits to lower-S/N features in the observed spectral line profiles. A choice of polynomial order that is too low can fail to accurately match the template and observed spectrum, while excessively large-order polynomials end up perturbing a template to match observational noise, which often yields nonsensical results. We test the dependence of output velocity dispersion on polynomial order by forcing pPXF to fit the observed spectrum with the single best-fit BC03 template as determined above with order fixed to each $d = [1, 2, 3, ..., 50]$. Example results are shown in Figure B2. For the most part, we see the greatest variability in output velocity dispersion at $d > 20$, as well as some at $d < 5$, while between these values the output velocity dispersion appears generally stable.

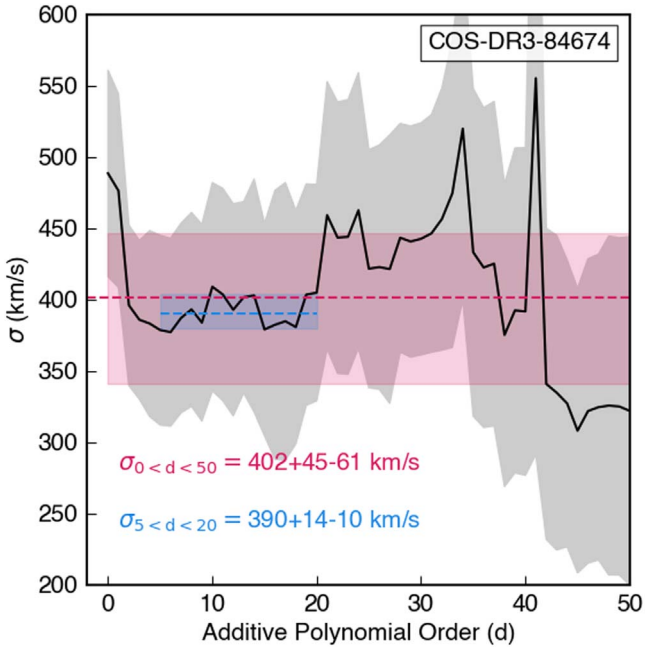


**Figure B2.** Velocity dispersion dependence upon choice of additive polynomial order for COS-DR3-84674. Results are calculated for $0 \leqslant d \leqslant 50$, with the gray shaded region indicating the uncertainties returned by pPXF for each measurement. The red dashed line indicates the median value across all choices of $d$, the shading between the 16th and 84th percentiles. The blue dashed line is the median value over $5 \leqslant d \leqslant 20$.

### *B.3. Dependence on Spectral Wavelength Range*

Velocity dispersion fits are also dependent upon the wavelengths available in the observed spectrum, where the inclusion or exclusion of specific spectral features can alter results. We refit truncated spectra using a range of starting wavelengths from $3200 < \lambda_{\rm rest,blue}/\text{Å} < 5000$ and ending

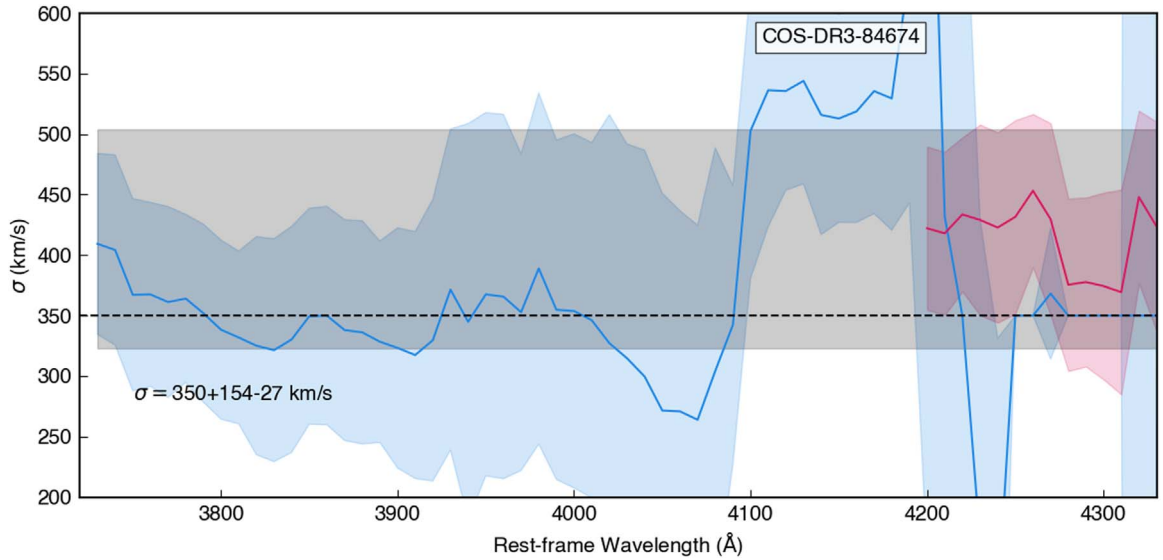


**Figure B3.** Velocity dispersion fit as a result of trimming the observed spectra. Blue lines represent the velocity dispersion returned when the spectrum is fit from the rest-frame wavelength on the abscissa to the reddest wavelength available. Red lines represent the velocity dispersion returned when the spectrum is fit from the bluest wavelength available to the wavelength on the abscissa. Shaded regions show the uncertainties returned by pPXF for each measurement. An additive polynomial of order $d = 10$ is used for this example.

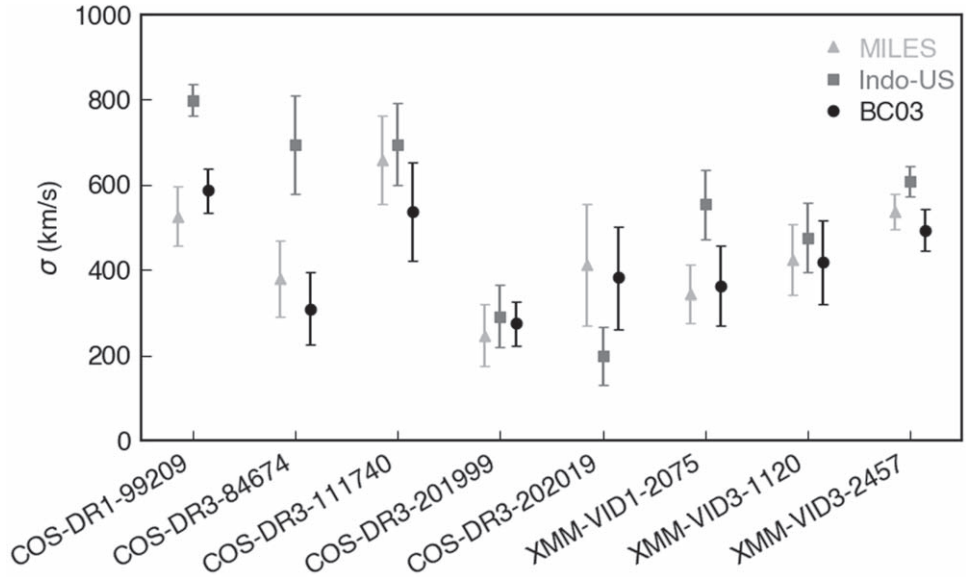


**Figure B4.** Velocity dispersion fit as a result of choice of template library with all other parameters fixed. Templates from the MILES library (light gray triangles), Indo-US library (dark gray squares), and Bruzual & Charlot models (black circles) are shown for each galaxy.

wavelengths $4200 < \lambda_{\rm rest,red}/\text{Å} < 6000$ and analyze the results (Figure B3). The most apparent result is that when the spectrum includes the Ca H&K lines, the velocity dispersion results are significantly more stable. In many cases, there also appears to be variability with the inclusion or exclusion of H$\delta$. Notably, we tend not to see much dependence on the H$\beta$ feature, which suggests that there is little line infilling. Further insights are difficult due to the small sample, low S/N of the spectra, and dependence of results on polynomial order.

Given the strong dependence of the results on the inclusion of Ca H&K, we also fit the spectra over the narrow range of $3900 < \lambda_{\rm rest}/\text{Å} < 4000$ so as to isolate these features. However, doing so precludes the use of the higher-order polynomials discussed above, as the narrow wavelength range means each order has outsize effects on the result. Nevertheless, the results of this fit are statistically consistent within $1\sigma$ for five of the galaxies. The remaining three (COS-DR1-99209, COS-DR3-111740, and COS-DR3-202019) showed significant deviations at $d = 0$ when testing the polynomial order above and so this discrepancy is not surprising.

### *B.4. Dependence on the Template Library*

In this work we use the BC03 template library due to its longer wavelength coverage, which allows joint fitting with photometry using FAST++. However, the velocity dispersions from pPXF can be highly dependent upon the availability of templates which are appropriate to the data. As such, we analyze results from using solely pPXF with three libraries: the BC03 library, the Indo-US stellar library templates (Valdes et al. 2004), and SSPs from the MILES stellar library

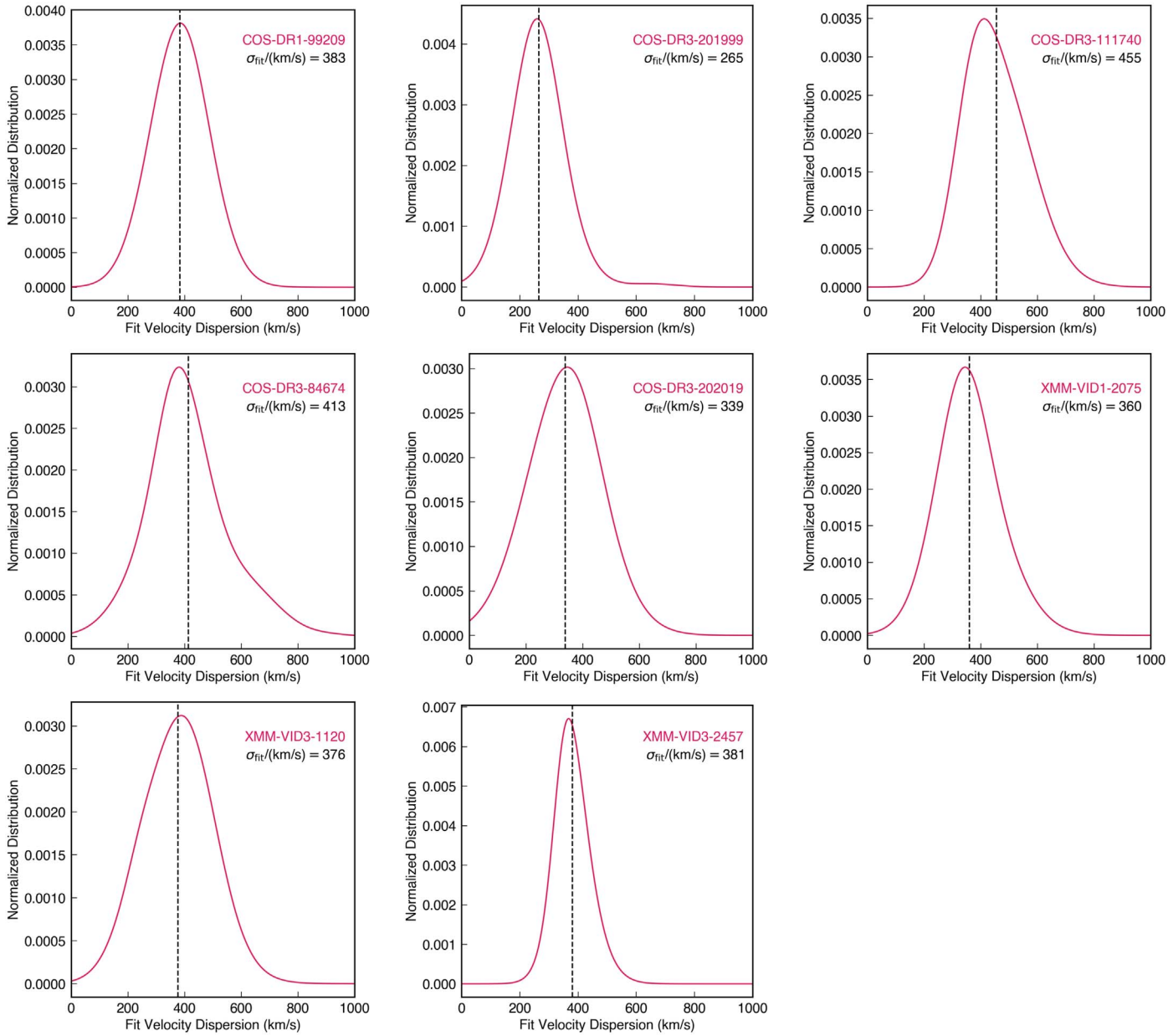


**Figure B5.** The distribution of measured velocity dispersions returned by pPXF for each of the MAGAZ3NE objects presented in this work. Here, each run of pPXF is convolved with a Gaussian kernel with a width equivalent to the uncertainty reported by pPXF and additionally weighted by the reduced $\chi^2$ of the fit. The weighted average is also labeled. This is the value transformed to $\sigma_e$ by performing an aperture correction.

(Sánchez-Blázquez et al. 2006; Vazdekis et al. 2010). For half of the galaxies, all three libraries yield statistically similar results with other parameters fixed (Figure B4), while in the other half the MILES and BC03 outputs are similar and the Indo-US library produces discrepant results.

### *B.5. Overall Distribution of Velocity Dispersions*

As mentioned in the text, we perform a large number of fits with pPXF for each galaxy. Due to the variety of results and uncertainties associated with any particular fit, we instead use the distribution of results as a whole to determine stellar velocity dispersion for a particular galaxy. Each fit was convolved with a Gaussian kernel with a standard deviation equal to the reported uncertainty on the velocity dispersion and additionally weighted by the reduced $\chi^2$ of the fit. The normalized summation of these fits for the eight MAGAZ3NE UMGs are shown in Figure B5.

## ORCID iDs

Ben Forrest https://orcid.org/0000-0001-6003-0541
Gillian Wilson https://orcid.org/0000-0002-6572-7089
Adam Muzzin https://orcid.org/0000-0002-9330-9108
Danilo Marchesini https://orcid.org/0000-0001-9002-3502
M. C. Cooper https://orcid.org/0000-0003-1371-6019
Z. Cemile Marsan https://orcid.org/0000-0002-7248-1566
Marianna Annunziatella https://orcid.org/0000-0002-8053-8040
Ian McConachie https://orcid.org/0000-0002-2446-8770
Percy Gomez https://orcid.org/0000-0003-0408-9850
Stephanie M. Urbano Stawinski https://orcid.org/0000-0001-8169-7249

Wenjun Chang https://orcid.org/0000-0003-2144-2943
Francesco La Barbera https://orcid.org/0000-0003-1181-6841
Lori Lubin https://orcid.org/0000-0003-2119-8151
Julie Nantais https://orcid.org/0000-0002-7356-0629
Theodore Peña https://orcid.org/0000-0002-0033-5041
Paolo Saracco https://orcid.org/0000-0003-3959-2595
Jason Surace https://orcid.org/0000-0001-7291-0087
Mauro Stefanon https://orcid.org/0000-0001-7768-5309

## References

Anderson, J., Bourque, M., Sahu, K., Sabbi, E., & Viana, A. 2015, A Study of the Time Variability of the PSF in F606W Images taken with WFC3/UVIS, Instrument Science Report WFC3 2015-08, Space Telescope Science Institute
Ashby, M. L. N., Caputi, K. I., Cowley, W., et al. 2018, ApJS, 237, 39
Barro, G., Kriek, M., Pérez-González, P. G., et al. 2016, ApJ, 827, L32
Barro, G., Trump, J. R., Koo, D. C., et al. 2014, ApJ, 795, 145
Belli, S., Newman, A. B., & Ellis, R. S. 2014a, ApJ, 783, 117
Belli, S., Newman, A. B., & Ellis, R. S. 2017, ApJ, 834, 18
Belli, S., Newman, A. B., Ellis, R. S., & Konidaris, N. P. 2014b, ApJ, 788, L29
Bernardi, M., Sheth, R. K., Nichol, R. C., et al. 2006, AJ, 131, 2018
Bertin, G., Ciotti, L., & Del Principe, M. 2002, A&A, 386, 149
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, ApJ, 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, ApJ, 779, L21
Bezanson, R., van der Wel, A., Pacifici, C., et al. 2018, ApJ, 858, 60
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Cappellari, M. 2017, MNRAS, 466, 798
Cappellari, M., Bacon, R., Bureau, M., et al. 2006, MNRAS, 366, 1126
Cappellari, M., di Serego Alighieri, S., Cimatti, A., et al. 2009, ApJ, 704, L34
Cappellari, M., & Emsellem, E. 2004, PASP, 116, 138
Cappellari, M., McDermid, R. M., Alatalo, K., et al. 2013b, MNRAS, 432, 1862
Cappellari, M., Scott, N., Alatalo, K., et al. 2013a, MNRAS, 432, 1709
Chabrier, G. 2003, PASP, 115, 763
Conroy, C., Dutton, A. A., Graves, G. J., Mendel, J. T., & Van Dokkum, P. G. 2013, ApJL, 776, L26
Conroy, C., & van Dokkum, P. G. 2012, ApJ, 760, 71
Conroy, C., van Dokkum, P. G., & Villaume, A. 2017, ApJ, 837, 166
Cutler, S. E., Whitaker, K. E., Mowla, L. A., et al. 2022, ApJ, 925, 34
Djorgovski, S., & Davis, M. 1987, ApJ, 313, 59
Donnari, M., Pillepich, A., Nelson, D., et al. 2021, MNRAS, 506, 4760
Dressler, A. 1987, ApJ, 317, 1
Esdaile, J., Glazebrook, K., Labbé, I., et al. 2021, ApJ, 908, L35
Forrest, B., Annunziatella, M., Wilson, G., et al. 2020a, ApJ, 890, L1
Forrest, B., Marsan, Z. C., Annunziatella, M., et al. 2020b, ApJ, 903, 47
Forrest, B., Tran, K.-V. H., Broussard, A., et al. 2017, ApJ, 838, L12
Faber, S. M., & Jackson, R. E. 1976, ApJ, 204, 668
Gargiulo, A., Saracco, P., Tamburri, S., Lonoce, I., & Ciocca, F. 2016, A&A, 592, A132
Gebhardt, K., Bender, R., Bower, G., et al. 2000, ApJ, 539, L13
Glazebrook, K., Schreiber, C., Labbé, I., et al. 2017, Natur, 544, 71
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Natur, 585, 357
Haussler, B., McIntosh, D. H., Barden, M., et al. 2007, ApJS, 172, 615
Häußler, B., Bamford, S. P., Vika, M., et al. 2013, MNRAS, 430, 330
Hill, A. R., Muzzin, A., Franx, M., & van de Sande, J. 2016, ApJ, 819, 74
Hopkins, A. M. 2018, PASA, 35, e039
Hopkins, P. F., Lauer, T. R., Cox, T. J., Hernquist, L., & Kormendy, J. 2009, ApJS, 181, 486
Hunter, J. D. 2007, CSE, 9, 90
Husser, T.-O., Wende-von Berg, S., Dreizler, S., et al. 2013, A&A, 553, A6
Jarvis, M. J., Bonfield, D. G., Bruce, V. A., et al. 2013, MNRAS, 428, 1281
Jorgensen, I., Franx, M., & Kjaergaard, P. 1995, MNRAS, 276, 1341
Kausch, W., Noll, S., Smette, A., et al. 2015, A&A, 576, A78
Köppen, J., Weidner, C., & Kroupa, P. 2007, MNRAS, 375, 673
Kormendy, J., & Ho, L. C. 2013, ARA&A, 51, 511
Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2009, ApJ, 700, 221
Kroupa, P., Weidner, C., Pflamm-Altenburg, J., et al. 2013, in Planets, Stars and Stellar Systems, ed. T. D. Oswalt & G. Gilmore, Vol. 5 (Dordrecht: Springer), 115
La Barbera, F., Vazdekis, A., Ferreras, I., et al. 2019, MNRAS, 489, 4090
Lacy, M., Surace, J. A., Farrah, D., et al. 2021, MNRAS, 501, 892
Läsker, R., van den Bosch, R. C., van de Ven, G., et al. 2013, MNRAS, 434, L31
Leja, J., Johnson, B. D., Conroy, C., & van Dokkum, P. 2018, ApJ, 854, 62
Leja, J., Johnson, B. D., Conroy, C., van Dokkum, P. G., & Byler, N. 2017, ApJ, 837, 170
Leja, J., Johnson, B. D., Conroy, C., et al. 2019, ApJ, 877, 140
Lian, J., Thomas, D., & Maraston, C. 2018, MNRAS, 39, 6027
Lustig, P., Strazzullo, V., D'Eugenio, C., et al. 2021, MNRAS, 501, 2659
Lustig, P., Strazzullo, V., Remus, R.-S., et al. 2022, arXiv:2201.09068
Marsan, Z. C., Marchesini, D., Brammer, G. B., et al. 2017, ApJ, 842, 21
Marsan, Z. C., Marchesini, D., Brammer, G. B., et al. 2015, ApJ, 801, 133
Marsan, Z. C., Marchesini, D., Muzzin, A., et al. 2019, ApJ, 871, 201
Marsan, Z. C., Muzzin, A., Marchesini, D., et al. 2022, ApJ, 924, 25
Martín-Navarro, I., Vazdekis, A., La Barbera, F., et al. 2015, ApJL, 806, L31
Mauduit, J.-C., Lacy, M., Farrah, D., et al. 2012, PASP, 124, 1135
McCracken, H. J., Milvang-Jensen, B., Dunlop, J., et al. 2012, A&A, 544, A156
McLean, I. S., Steidel, C. C., Epps, H., et al. 2010, Proc. SPIE, 7735, 77351E
McLean, I. S., Steidel, C. C., Epps, H., et al. 2012, Proc. SPIE, 8446, 84460J
Mendel, J. T., Beifiori, A., Saglia, R. P., et al. 2020, ApJ, 899, 87
Mobasher, B., Dahlen, T., Ferguson, H. C., et al. 2015, ApJ, 808, 101
Mowla, L. A., van Dokkum, P., Brammer, G. B., et al. 2019, ApJ, 880, 57
Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013a, ApJ, 777, 18
Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013b, ApJS, 206, 8
Muzzin, A., Wilson, G., Yee, H. K. C., et al. 2009, ApJ, 698, 1934
Nedkova, K. V., Häußler, B., Marchesini, D., et al. 2021, MNRAS, 506, 928
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, ApJ, 717, L103
Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713
Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, AJ, 124, 266
Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H. W. 2010, AJ, 139, 2097
Pérez, F., & Granger, B. E. 2007, CSE, 9, 21
Posacki, S., Cappellari, M., Treu, T., Pellegrini, S., & Ciotti, L. 2015, MNRAS, 446, 493
Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., et al. 2018, AJ, 156, 123
Prochaska, J., Hennawi, J., Westfall, K., et al. 2020, JOSS, 5, 2308
Ribeiro, B., Le Fèvre, O., Tasca, L. A. M., et al. 2016, A&A, 593, A22
Robitaille, T. P., Tollerud, E. J., Greenfield, P., et al. 2013, A&A, 558, A33
Salmon, B., Papovich, C., Finkelstein, S. L., et al. 2015, ApJ, 799, 183
Salpeter, E. E. 1955, ApJ, 121, 161
Sánchez-Blázquez, P., Peletier, R. F., Jiménez-Vicente, J., et al. 2006, MNRAS, 371, 703
Sanders, R. L., Shapley, A. E., Jones, T., et al. 2021, ApJ, 914, 19
Saracco, P., La Barbera, F., Gargiulo, A., et al. 2019, MNRAS, 484, 2281
Saracco, P., Marchesini, D., Barbera, F. L., et al. 2020, ApJ, 905, 40
Saulder, C., Van Den Bosch, R. C., & Mieske, S. 2015, A&A, 578, A134
Schreiber, C., Glazebrook, K., Nanayakkara, T., et al. 2018a, A&A, 618, A85
Schreiber, C., Labbé, I., Glazebrook, K., et al. 2018b, A&A, 611, A22
Sherman, S., Jogee, S., Florez, J., et al. 2019, MNRAS, 491, 3318
Shu, Y., Bolton, A. S., Schlegel, D. J., et al. 2012, AJ, 143, 90
Smette, A., Sana, H., Noll, S., et al. 2015, A&A, 576, A77
Stark, D. P., Schenker, M. a., Ellis, R., et al. 2013, ApJ, 763, 129
Straatman, C. M., Labbé, I., Spitler, L. R., et al. 2015a, ApJL, 808, L29
Straatman, C. M. S., Labbé, I., Spitler, L. R., et al. 2014, ApJ, 783, L14
Straatman, C. M. S., Labbé, I., Spitler, L. R., et al. 2015b, ApJ, 808, L29
Tanaka, M., Valentino, F., Toft, S., et al. 2019, ApJ, 885, L34
Thomas, D., Steele, O., Maraston, C., et al. 2013, MNRAS, 431, 1383
Toft, S., Gallazzi, A., Zirm, A., et al. 2012, ApJ, 754, 3
Toft, S., Zabl, J., Richard, J., et al. 2017, Natur, 546, 510
Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, ApJ, 613, 898
Treu, T., Auger, M. W., Koopmans, L. V., et al. 2010, ApJ, 709, 1195
Valdes, F., Gupta, R., Rose, J. A., Singh, H. P., & Bell, D. J. 2004, ApJS, 152, 251
Valentino, F., Tanaka, M., Davidzon, I., et al. 2020, ApJ, 889, 93
van de Sande, J., Kriek, M., Franx, M., et al. 2013, ApJ, 771, 85
Van Der Wel, A., Bell, E. F., Van Den Bosch, F. C., Gallazzi, A., & Rix, H. W. 2009, ApJ, 698, 1232
van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, ApJ, 788, 28
van Dokkum, P., Conroy, C., Villaume, A., Brodie, J., & Romanowsky, A. J. 2017, ApJ, 841, 68

van Dokkum, P. G. 2008, ApJ, 674, 29
van Dokkum, P. G., Kriek, M., & Franx, M. 2009a, Natur, 460, 717
van Dokkum, P. G., Labbé, I., Marchesini, D., et al. 2009b, PASP, 121, 2
Vazdekis, A., Sánchez-Blázquez, P., Falcón-Barroso, J., et al. 2010, MNRAS, 404, 1639
Villaume, A., Brodie, J., Conroy, C., Romanowsky, A. J., & van Dokkum, P. 2017, ApJ, 850, L14
Wake, D. A., van Dokkum, P. G., & Franx, M. 2012, ApJ, 751, L44
Wilson, J. C., Henderson, C. P., Herter, T. L., et al. 2004, Proc. SPIE, 5492, 1295
Wuyts, S., Franx, M., Cox, T. J., et al. 2009, ApJ, 696, 348
Zahid, H. J., & Geller, M. J. 2017, ApJ, 841, 32
Zahid, H. J., Geller, M. J., Fabricant, D. G., & Hwang, H. S. 2016, ApJ, 832, 203