ABSTRACT OF THE DISSERTATION

Insights on Galaxy Evolution from Studies of the Multiphase Interstellar Medium

By JOHN WU

Dissertation Director:
Andrew J. Baker

Modern galaxy formation and evolution models are able to match the statistical properties of galaxy populations across most of cosmic history. However, detailed observations and sophisticated analytical methods are still needed to test theoretical predictions in extreme scenarios, where the complex interplay of gas accretion, star formation, galaxy interactions, feedback, and other physical processes can have compounded impacts on interstellar medium properties. I discuss results from new far-infrared and submillimeter observations of star-forming galaxies in four massive galaxy clusters at redshifts 0.3 ≲ z ≲ 1.1. Despite being surrounded by hot intracluster plasma, some cluster members are able to retain their cold gas reservoirs, and have high dust-obscured star formation rates. I find that the prevalence of star-forming cluster galaxies increases with increasing redshift. I also present results from integral field spectroscopy, centered on the Paα line, targeting a sample of rare, low-redshift, compact starbursts that strongly resemble z ∼ 3 Lyman break galaxies (LBGs). H2 ro-vibrational emission is measured for the first time in either distant LBGs or their low-z analogs. Warm molecular gas and ionized gas properties suggest that star formation feedback is the dominant excitation mechanism in these systems. I also find that, compared to typical nearby star-forming galaxies, LBG analogs are characterized by high velocity dispersions and low ordered-to-disordered velocity ratios. LBG analogs deviate from kinematic scaling relations such as the stellar mass and baryonic Tully-Fisher relations; their anomalously low kinematic support may be attributed to their small physical sizes. Finally, I optimize a deep convolutional neural network (CNN) to predict the
gas-phase metallicities of typical low-z star-forming galaxies from three-band optical imaging. The trained CNN is not only able to accurately estimate metallicity, but also can reconstruct the empirical mass-metallicity relation with zero additional scatter. These results imply that multi-color morphological features are important for understanding the connection between galaxies’ stellar mass assembly and chemical enrichment histories.
Acknowledgments

I would like to acknowledge support from awards AST-0955810 and AST-1814421 from the National Science Foundation (NSF), award GSSP SOSPA2-018 from the National Radio Astronomy Observatory (NRAO), and a Research and Innovation Fellowship from the United States Agency for International Development (USAID). I thank Rutgers University for funding support through fellowship and TA positions. I also thank the Vatican Observatory, the University of Cape Town, South African Astronomical Observatory, and the South African Radio Astronomy Observatory for their hospitality.


My advisor, Professor Andrew Baker, has been instrumental in helping me become a better astronomer and a better person. I’m also grateful for his advice on matters of research, telescope observing, reading (works of fiction), writing, mentoring, awareness, public speaking, and life in general. Andrew has been an amazing professional and personal role model for me. I will always remember, and hope to emulate, his selflessness.

I would like to acknowledge Professors Eric Gawiser, Saurabh Jha, Jolie Cizewski, and Adam Leroy for taking the time to serve on my PhD committee, and for their helpful comments that have improved this dissertation. I am also very grateful for Eric’s advice and suggestions when I began my foray into machine learning.

I would also like to thank my undergraduate research advisor, Rachel Mandelbaum, for being an excellent mentor. I credit Rachel with getting me excited about research in astrophysics. Without her support and guidance, I don’t know if I would have gone to graduate school.
I am also indebted to many graduate students and postdocs in Rutgers Astronomy; our countless office discussions (including those of the non-research variety) have been invaluable to me. In particular, I want to thank Jesse Rivera for being a supportive friend and wonderful colleague in our research group. I have also found conversations with Kartheik Iyer to always be incredibly stimulating. Steven Boada has made significant and important contributions to Chapter 3, and this work would not be possible without his help. Matt Buckley and David Shih kindly allowed me to use their GPUs, for which I am very thankful.

I am especially grateful for my family and their support. I am indebted to my Mom and Dad, who have sacrificed far more than I can ever comprehend in order to see my brother and me succeed. My Dad always encouraged me to dream big, and I can still remember him reminding me, “When there’s a will, there’s a way.” Although he is no longer with me now, his unwavering hopefulness has and will always been a reservoir of strength for me.

Finally, and most importantly, I give my love and thanks to my wife, Priscilla. It is difficult to imagine that I managed to survive the first half of graduate school before meeting Priscilla, but meeting her was the best thing that could have happened to me. She is a constant source of support, wisdom, inspiration, and motivation. Priscilla is an oasis for me when work or life goes awry. Thank you for being by my side through this journey.
Dedication

To my Dad, Jianxin Jason Wu, who taught me curiosity, kindness, diligence, and love.
# Table of Contents

Abstract ................................................................. ii

Acknowledgments ......................................................... iv

Dedication ................................................................. vi

List of Tables ............................................................ x

List of Figures ............................................................ xi

1. Introduction .......................................................... 1
   1.1. Galaxy evolution ................................................. 1
   1.2. The baryon cycle ................................................ 3
   1.3. Galaxy scaling relations ....................................... 8
   1.4. Feedback processes in galaxies ................................. 12
   1.5. Effects of redshift and environment ........................... 15
   1.6. Some open questions in galaxy evolution ...................... 19

2. Herschel and ALMA observations of star-forming galaxies in massive clusters 23
   2.1. Galaxy evolution in massive clusters .......................... 23
   2.2. Observations of Sunyaev-Zel’dovich effect-selected clusters .... 26
   2.3. Herschel and ALMA detections ................................ 31
   2.4. The averaged properties of faint sources ....................... 36
   2.5. Galaxy evolution and the ISM properties of cluster galaxies .... 39
   2.6. Conclusions ...................................................... 46
   2.7. Appendix A: New ALMA and Herschel detections ................ 47
   2.8. Appendix B: Dusty star-forming galaxies ....................... 52

3. The star-forming interstellar medium of Lyman break galaxy analogs .......... 73
   3.1. Introduction ..................................................... 73
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2. Observations and data reduction</td>
<td>76</td>
</tr>
<tr>
<td>3.3. Ionized gas kinematics</td>
<td>77</td>
</tr>
<tr>
<td>3.4. The multiphase ISM</td>
<td>85</td>
</tr>
<tr>
<td>3.5. Warm and cool molecular gas</td>
<td>97</td>
</tr>
<tr>
<td>3.6. Resolved near-infrared line ratios</td>
<td>99</td>
</tr>
<tr>
<td>3.7. Conclusions</td>
<td>103</td>
</tr>
<tr>
<td>4. Kinematic scaling relations for Lyman break galaxy analogs</td>
<td>126</td>
</tr>
<tr>
<td>4.1. Introduction</td>
<td>126</td>
</tr>
<tr>
<td>4.2. Data</td>
<td>127</td>
</tr>
<tr>
<td>4.3. Kinematic scaling relations</td>
<td>129</td>
</tr>
<tr>
<td>4.4. Discussion</td>
<td>130</td>
</tr>
<tr>
<td>4.5. Conclusions</td>
<td>133</td>
</tr>
<tr>
<td>5. Predicting metallicity from three-band imaging using convolutional neural networks</td>
<td>138</td>
</tr>
<tr>
<td>5.1. Introduction: The growing role of machine learning in astronomy</td>
<td>138</td>
</tr>
<tr>
<td>5.2. The SDSS emission-line galaxy sample</td>
<td>139</td>
</tr>
<tr>
<td>5.3. Methodology for training the convolutional neural network</td>
<td>140</td>
</tr>
<tr>
<td>5.4. Predicting metallicity using only three-color imaging</td>
<td>141</td>
</tr>
<tr>
<td>5.5. Interpreting the CNN</td>
<td>144</td>
</tr>
<tr>
<td>5.6. The mass-metallicity relation</td>
<td>149</td>
</tr>
<tr>
<td>5.7. Summary</td>
<td>152</td>
</tr>
<tr>
<td>5.8. Appendix: Convolution neural network details</td>
<td>153</td>
</tr>
<tr>
<td>6. Summary</td>
<td>164</td>
</tr>
<tr>
<td>6.1. How do molecular gas, dust, and star formation properties evolve with redshift for galaxies in the most overdense environments?</td>
<td>164</td>
</tr>
<tr>
<td>6.2. What are the multiphase ISM properties of low-redshift LBG analogs?</td>
<td>166</td>
</tr>
<tr>
<td>6.3. How do LBG analogs compare to other systems in terms of kinematic scaling relations?</td>
<td>168</td>
</tr>
<tr>
<td>6.4. How do optical morphological features in local star-forming galaxies connect to stellar mass and gas-phase metallicity?</td>
<td>169</td>
</tr>
<tr>
<td>Bibliography</td>
<td>174</td>
</tr>
</tbody>
</table>
Appendix A. LADUMA and the search for neutral atomic hydrogen at high redshift

A.1. MeerKAT interferometric observations: calibration, imaging, and continuum subtraction .............................. 195
A.2. H I source finding ........................................................................................................................................... 201
List of Tables

2.1. Our sample of clusters ........................................... 54
2.2. ALMA Band 6 Observations ...................................... 55
2.3. Measurements of detected cluster members .................... 56
2.4. Derived properties of detected cluster members ............... 57
2.5. Numbers of SoFiA detections .................................... 58
2.6. Numbers of PACS detections ..................................... 58
2.7. PACS stacking 3 σ upper limits .................................. 58
2.8. ALMA continuum detections ..................................... 59
3.1. VLT/SINFONI observations .................................... 107
3.2. Ordered velocities and dispersions ............................... 108
3.3. Nebular and continuum color excesses ......................... 109
3.4. Recombination line fluxes ...................................... 109
3.5. Recombination line widths ...................................... 110
3.6. Ro-vibrational line fluxes ..................................... 111
3.7. Ro-vibrational line widths ..................................... 112
3.8. Warm and cool H$_2$ masses .................................... 112
A.1. LADUMA 64-antenna observations ............................ 207
A.2. LADUMA 64-antenna selfcal results ............................ 207
2.1. ALMA Band 6 continuum mosaic imaging of J0102 shown in grayscale. Confirmed cluster members are marked by circles: black circles are galaxies found via optical spectroscopy (S13); smaller green filled circles are galaxies with Herschel/PACS counterparts; larger red (blue) circles are galaxies with ALMA CO (4−3) ([C I] (3P_1 − 3P_0)) line detections (labeled by name). A purple cross marks the brightest cluster galaxy (also labeled). The large dashed circle encloses 0.5 × R_{200c} and is centered on the midpoint between the two peaks in the mass distribution (determined by weak lensing; Jee et al. 2014).

2.2. Chary & Elbaz (2001, CE01) and three example Kirkpatrick et al. (2015, K15) template SEDs with \( L_{\text{IR}} = 10^{11} \, L_\odot \), redshifted to \( z = 0.87 \). We show the sensitivities (3 σ upper limits) of our FIR/millimeter observations (point-source subtracted maps) in colored horizontal lines.

2.3. Histograms of bootstrapped stacked Herschel/PACS 100 µm and 160 µm flux densities for 5000 iterations. Outlying fluxes in excess of 4 MADs have been removed. Flux densities are taken from the “residual” point source-subtracted PACS maps. The black dashed vertical line signifies zero flux, and the solid black line indicates the mean flux of blank-sky pixels. We also report the number of galaxies per stack and \( p \)-values determined by computing the percentile of the blank-sky mean within the bootstrapped distribution of stacked flux densities.

2.4. (Left) Bootstrapped histograms of stacked ALMA dust continuum flux densities for galaxies in our two high-z clusters. (Right) Continuum fluxes converted to \( M_{\text{BH}} \) values and stacked in bins of clustercentric radius. 95% confidence intervals (computed by bootstrap resampling for 5000 iterations) are shown for each cluster. A dashed line marks zero. In both cases, 4-MAD outliers were removed.

2.5. [C I] (blue) and CO (red) stacked spectra from the J0102 and J0546 ALMA data cubes. Grey regions show 3 σ limits (centered around zero). All outlying flux densities in excess of 4 MADs were removed from calculation of stacked fluxes and uncertainties.
2.6. The ratio of [C I] and CO in integrated flux units plotted vs. IR luminosity (1 σ error bars shown). The fiducial literature-based flux ratio of 0.71 adopted in this paper is shown as a dashed line.

2.7. ISM mass (vertical) vs. molecular mass (horizontal) vs. IR luminosity (color). Error bars denote 1 σ, and leftward-pointing arrows indicate 3 σ upper limits on $M_{\text{ISM}}$.

2.8. FIR-radio correlation $q_{\text{IR}}$ parameter is plotted for 19 Herschel detections as a function of IR luminosity. The horizontal dashed line and shaded region show the median $q_{\text{IR}}$ value and ±1 σ uncertainties for 250 µm selected star-forming galaxies (Ivison et al. 2010), and the horizontal dotted line shows $q_{\text{IR}} = 1.8$ used to separate star-forming galaxies from AGNs. Error bars are ±1 σ; note that the error in $q_{\text{IR}}$ is correlated with errors in $L_{\text{IR}}$ and $L_{1.4 \text{ GHz}}$. Empty markers with upward-pointing arrows are 3 σ lower limits on $q_{\text{IR}}$ due to radio non-detections. Marker colors and symbols correspond to particular galaxy clusters and radio/optical spectral classifications, respectively. AGN-dominated systems as determined from best-fitting K15 IR spectral templates are marked with black stars.

2.9. IR luminosity due to star formation is plotted against clustercentric radius for 19 Herschel detections. Marker color and shape indicate cluster membership. Clustercentric radius is computed as the projected distance from the BCG for J0235, J0438, and J0546, and from the midpoint between weak lensing peaks for J0102 (Jee et al. 2014). Unambiguous radio AGNs ($q_{\text{IR}} < 1.8$) are shown with open markers. On the y-axis, a box plot shows the interquartile range in $L_{\text{IR}}^{\text{SF}}$.

2.10. Specific SFR (sSFR) is plotted against clustercentric radius for 19 Herschel detections. Colors and markers are the same as in Figure 2.9, where radio AGNs are shown with open markers. The dashed (solid) horizontal black line illustrates the measured SFR–$M_*$ correlation for $\log(M_*/M_\odot) = 11.0$ at $z \sim 1$ ($z \sim 0$) as reported by Elbaz et al. (2007). On the y-axis, a box plot shows the interquartile range in sSFR.
2.11. *Herschel* and ALMA contours overlaid on *HST*/ACS F625W + F775W + F850LP color images in 5″ cutouts, for J0102 detections (J010252.44-491531.2 (a); J010255.50-491416.2 (b); J010255.67-491556.7 (c)). Left panels: *Herschel*/PACS 160 μm (green; multiples of 1.5 σ) and ALMA dust continuum (purple; multiples of 1.5 σ) contours, with ALMA synthesized beam at lower left. Middle panels: CO (4 – 3) integrated intensity (red; multiples of 1.5 σ) contours, with synthesized beams at lower left and inset spectra showing ±1 σ (gray) and velocity range used for moment map (yellow). Right panel: same as middle panels for the [C I] (3P₁ – 3P₀) line. Red and blue crosses in left panels mark peaks of line emission in middle and right panels, respectively.

2.12. Spectral line sources in J0546 (J054627.43-534433.6 (a); J054638.87-534613.6 (b); J054644.15-534608.7 (c)). Notation as in Figure 2.11. Grayscale images are *HST* F814W; 160 μm contours for top panel (a) are multiples of 5 σ.

2.13. 16″ cutouts of *Herschel* and ALMA contours overlaid on *HST* images. (a) ALMA continuum contours (purple, ±2 σ levels) and PACS 160 μm contours (green; ±2 σ levels) overlaid on *HST*/ACS F625W + F775W + F850LP color imaging of a possible strongly lensed system, J010249.28-491506.8. Possible arcs of millimeter continuum emission are seen to the northeast and southwest, the latter of which is matched to a blue arc in *HST* imaging. The early-type galaxies at the center of the image and toward the southeast are confirmed cluster galaxies. (b) Two bright DSFGs, J010254.89-491514.6 and J010255.66-491509.0, which were detected as a blended source (and named “La Flaca”; Lindner et al. 2015), are shown with ALMA continuum contours (purple, ±8 σ levels) and PACS 160 μm contours (green; ±2 σ levels) overlaid on *HST* false-color imaging. No PACS emission is detected, and whereas no short-wavelength counterpart is detected for J010254.89-491514.6 (southwest), a *Spitzer* counterpart is found for J010255.66-491509.0 (northeast). (c) A strongly lensed galaxy, J054633.36-534548.6, is shown with ALMA continuum contours (purple, ±2 σ levels) and PACS 160 μm contours (green; ±2 σ levels) overlaid on *HST*/F814W grayscale imaging. No PACS emission is detected with the DSFG. (d) A bright submillimeter and FIR source, J054634.57-534552.3, is shown with ALMA continuum contours (purple, ±4 σ levels) and PACS 160 μm contours (green; ±2 σ levels) overlaid on *HST*/F814W grayscale imaging.
3.1. VLT/SINFONI channel maps of the continuum-subtracted data cubes for three representative LBAs in our sample, centered on the Paα line. Velocities are determined with respect to the Paα line centroid. The PSF is shown in the upper right corner of the top-right panel. The colorscale corresponds to log($F_\lambda$).  

3.2. In each set of panels, we show VLT/SINFONI Paα integrated flux maps (left), velocity maps (center), and dispersion maps (right), generated without clipping (upper) and with clipping of pixels below 4 $\sigma_{\text{rms, chan}}$ (lower). Unclipped velocity and dispersion maps have been masked at the 2 $\sigma$ integrated flux level only for visualization purposes. The PSF is shown at the lower left as a black ellipse.

3.2. (Continued from previous page.)

3.3. The NIR spectra for our sample extracted from a two-pixel radius aperture centered on the peak pixel of Paα integrated flux. The source spectrum is shown in blue, uncertainties are shown in gray, and the spectrum of the blank sky is shown in red.

3.3. (Continued on next page)

3.4. Recombination lines in J0150 for hydrogen (left) and He I(right) relative to systemic velocity. Strong Paα emission has been multiplied by 0.1 for better viewing. Typical uncertainty represented by a standard deviation is shown in the upper right side of the right panel.

3.4. Recombination lines (rows labeled at right) shown for each LBA (columns labeled at top). Spectra centered on the systemic velocity and extracted over a two-pixel radius aperture are shown in blue, and Gaussian model component fits are shown in red.

3.5. Recombination lines (rows labeled at right) shown for each LBA (columns labeled at top). Spectra centered on the systemic velocity and extracted over a two-pixel radius aperture are shown in blue, and Gaussian model component fits are shown in red.

3.5. Recombination lines (rows labeled at right) shown for each LBA (columns labeled at top). Spectra centered on the systemic velocity and extracted over a two-pixel radius aperture are shown in blue, and Gaussian model component fits are shown in red.

3.6. Dust-corrected hydrogen and He I recombination line fluxes, normalized to Brδ and He I 1869 respectively, shown for all LBAs in different colors and markers. Case B recombination line ratios are shown in black for hydrogen (solid; $T_e = 10^4$ K, $n_e = 10^3$ cm$^{-3}$) and helium (dashed; $T_e = 10^4$ K, $n_e = 10^4$ cm$^{-3}$). The He lines do not follow their Case B ratios, likely due to the combined effects of resonant scattering, dust or hydrogen absorption, and collisional population of the upper level for the 2.059 $\mu$m line (see Section 3.4.2 for more).

3.7. Observed dust-corrected He I-to-H I recombination line ratios (points) shown with Cloudy model predictions (curves). We show model results while varying effective temperature (colorbar) and logarithmic ionization parameter $\log U \in \{-4.0, -3.0, -2.0, -1.0\}$, as labeled.
3.8. Dust-corrected column density normalized by level degeneracy for H$_2$ $v = 1$ upper rotational levels plotted against energy in temperature units (color). A single best-fit excitation temperature ($T_{\text{ex}}$) is also plotted (in black) for each LBA component aside from J0212, whose ro-vibrational spectrum lacks high enough SNR to fit a temperature profile.

3.9. Continuum and line ratio maps for our LBA sample. Smoothed contours of the jointly-fit Pa$_\alpha$ flux are plotted at $\pm 2, 4, 8, \ldots \times \sigma_{\text{Pa}_\alpha, \text{rms}}$. Color scales show continuum flux or ratios of the indicated lines in logarithmic units. He lines are fit together using the same line width and with 1.869, 1.870, and 2.059 $\mu$m line ratios given by Benjamin et al. (1999) Case B calculations. The molecular hydrogen 1-0 S(1), S(2), S(3), and S(5) lines are jointly fit using the same line width and with fluxes determined from degeneracies and (unphysical) $T_{\text{ex}} = \infty$, which was chosen to maximize signal-to-noise ratio. These maps have not been corrected for dust attenuation.

3.9. (Continued from previous page.)

3.10. Hydrogen recombination (left) and H$_2$ ro-vibrational lines (right) for J1434. We show spectra from the main system (solid, $z = 0.180$) and companion (dashed, $z = 0.181$), extracted over two-pixel apertures centered on their respective stellar light peaks and shifted to the same velocity frame. The ratios of H$_2$ to Pa$_\alpha$ are clearly different for the main source and its companion. We have smoothed all spectra using a Gaussian filter with $\sigma = 1$ channel for visualization purposes only. Note also that H$_2$ 1-0 S(1) does not lie within our spectral range so it is not shown here.

4.1. Kinematic scaling relations for the LBG analogs (LBA) sample. Ordered rotation velocity $V_{\text{rot}}$, gas velocity dispersion $\sigma$, and combined kinematic estimator $S_{0.5} \equiv (0.5V_{\text{rot}}^2 + \sigma^2)^{1/2}$ are plotted against stellar mass $M_*$. Shown for comparison are the Di Teodoro et al. (2016) $z \sim 1$ sample (blue triangles), Simons et al. (2016) $z \sim 2$ sample (gray squares), and Straatman et al. (2017) $z \sim 2$ sample (purple diamonds).

4.2. $V_{\text{rot}}$ subtracted from the Reyes et al. (2011) TFR ridge line (i.e., TFR “residuals”) versus the $V_{\text{rot}}/\sigma$ ratio. Our LBG analogs are shown as filled black circles and with 1 $\sigma$ error bars, although most are too small to be seen. Also shown are TFR residuals for the Simons et al. (2015) $0.1 < z < 0.375$ sample (open green circles) and the same high-$z$ galaxy samples from Figure 4.1.
4.3. (Left) Ordered rotation velocity plotted against total baryonic mass. BTFR trends are also shown for $z \sim 0$ (red line; Lelli et al. 2016) and $z \sim 2$ samples (blue line; Übler et al. 2017). (Right) The $S_{0.5} - M_{\text{bar}}$ relation shown for our sample, and for the Price et al. (2016) $z \sim 2$ sample, normalized to the Kassin et al. (2007) $z \sim 0$ slope. LBG analogs’ gas mas fractions are illustrated via colors.

5.1. SDSS imaging with predicted and true metallicities from the test data set. Five examples are shown from each of the following categories: (a) lowest predicted metallicity, (b) lowest true metallicity, (c) highest predicted metallicity, (d) highest true metallicity, (e) most under-predicted metallicity, (f) most over-predicted metallicity, and (g) a set of randomly selected galaxies.

5.2. Distributions of the true (black) and predicted (red) galaxy metallicities. Note that the bin widths are different for the two distributions. See text for details.

5.3. Bivariate distribution of true galaxy metallicity ($Z_{\text{true}}$) and CNN prediction ($Z_{\text{pred}}$) is shown in the main panel. Overlaid are the median predicted metallicity (solid red line), RMSE scatter (dashed red lines), and NMAD scatter (dashed violet lines), in bins of $Z_{\text{true}}$. The solid black line shows the one-to-one relation. The distribution of residuals ($Z_{\text{pred}} - Z_{\text{true}}$) is shown in the inset plot. In the upper panel, we again show the binned scatter, where the size of each marker is proportional to the number of galaxies in that bin. Each horizontal line corresponds to the average scatter over the entire test data set (and the global value indicated in the upper panel legend).

5.4. The effects of image resolution and color on CNN performance. Red and violet circular markers indicate scatter in the residual distribution ($\Delta Z$) measured using RMSE and NMAD, respectively, for $gri$ imaging. (Each point is analogous to the horizontal lines shown in Figure 5.3.) We also show predictions from a random forest algorithm as open triangle markers, and constant $\langle Z_{\text{true}} \rangle$ predictions as open square markers. Large, labeled squares indicate the scatter for images consisting of only $r$-band imaging and a combination of $g$- and $r$-bands.

5.5. In the left panel, we plot the CNN predicted galaxy stellar mass against true stellar mass. Colors and marker or line styles are the same as in Figure 5.3. In the right panel, we compare the predicted stellar mass converted to metallicity, assuming the Tremonti et al. (2004) MZR, against the true metallicity. These findings indicate that using the empirical MZR and CNN-predicted $M_{\star, \text{pred}}$ yields poor results, unlike what we have observed in Figure 5.3.
5.6. In the main panel, the predicted MZR comparing true $M_\star$ against CNN-predicted $Z_{\text{pred}}$ is shown in grayscale. The running median (solid red) and scatter (dashed red and violet) are shown in 0.2 dex mass bins. For comparison, we also show the Tremonti et al. (2004) observed median and scatter (solid and dashed black lines, respectively), which are binned by 0.1 dex in mass. In the upper panel, we show the scatter in the predicted and empirical MZR. The standard deviation of the scatter for the empirical MZR is shown as a dashed black line, while the red and violet circles respectively show RMSE and NMAD for the predicted MZR. Marker sizes are proportional to the number of galaxies in each stellar mass bin for the test data set. Global scatter in the CNN-predicted MZR appears to be comparable to, or even lower than, scatter in the empirical MZR.

6.1. Sky emission lines (red; Hanuschik 2003) and SALT blocking filter transmission curves (blue; Rangwala et al. 2008) shown against observed wavelength. Redshifted [O II] observing windows for J0102 (at $\lambda \approx 7000$ Å) and J0546 (at $\approx 7700$ Å) are shown, each of which spans $\pm 2000$ km s$^{-1}$ in the cluster velocity frame.

6.2. SFR surface density ($\Sigma_{\text{SFR}}$) vs gas mass column density ($\Sigma_g$). Star formation law predictions from gravitational instability-driven + feedback (G+F) turbulence models, parameterized by orbital timescale $t_{\text{orb}}$ (Krumholz et al. 2018), are shown as solid green lines, and a star formation feedback-only model (Ostriker & Shetty 2011) is shown as a dashed black line. Global measurements for available LBG analogs (blue markers; Contursi et al. 2017), nearby low-$z$ disk galaxies (purple stars; Leroy et al. 2013), and high-$z$ star-forming galaxies (gray triangles; Genzel et al. 2010) are shown for comparison. Also shown are resolved $\Sigma_{\text{SFR}}$ and $\Sigma_g$ for J2104 ($=210358$) in red.

6.3. Velocity dispersion ($\sigma$) and star formation rate (SFR) for LBG analogs compared with Krumholz & Burkhart (2016) theoretical models. (Left) $\sigma$ measured from unresolved PdBI CO(1−0) observations (blue square) and resolved ALMA CO(3−2) observations (red square) of J2104, which has $f_{\text{gas}} = 0.5$. (Right) $\sigma$ measured from unresolved PdBI CO observations are shown for two systems, J1434 ($=143417; f_{\text{gas}} = 0.3$), which agrees with the gravity-driven turbulence model, and J0213 ($=021348; f_{\text{gas}} = 0.6$), which favors the feedback-driven turbulence model. For both panels, gravity-driven scaling relations are based on measured gas fractions and maximum velocities, and feedback-driven models assume a marginally unstable gas disk (for which the gas Toomre 1964 parameter $Q_g = 1.0$).
A.1. Comparisons between the SoFiA-generated first moment map (left), the “true” moment map (center), and difference between the two (right). No noise has been added. Velocity maps are shown in units of redshift. Note that the difference map, or residual map, is scaled over a range of $\sim 10^{-6}$. The size of the subcube used to generate the first moment maps is $200 \times 200 \times 1061$ voxels.

A.2. The integrated flux maps of subcubes in the no-noise simulated data cube is shown prior to stitching. Lighter colors correspond to higher flux.

A.3. $\text{H}\text{I}$ flux distribution recovered by SoFiA runs using different source-finding parameters, including permutations of minimum spatial size ($s_i, s_j$; columns), minimum spectral size ($s_k$; rows), and minimum spectral size for merging ($r_k$; color). In all cases, $r_i = r_j = 1$. The true flux distribution is shown in black.

A.4. $\text{H}\text{I}$ flux distribution recovered by SoFiA runs using different reliability thresholds ($R$; color). The hyperparameter values chosen ($s_i, s_j, s_k$) = (2, 2, 4) and ($r_i, r_j, r_k$) = (1, 1, 9) have been optimized for the noise-added simulation. The true flux distribution is shown in black.
Chapter 1

Introduction

1.1 Galaxy evolution

Galaxies are gravitationally bound structures composed of stars, gas, and dark matter. Astronomical observations have revealed detailed characteristics of galaxies, and how these properties correlate with each other. The diverse physical processes that influence galaxy properties can have complex interactions, often making it difficult to explain the growth of galaxies using only simple, idealized models. In the sections below, I will review the most important physical processes that govern the formation and evolution of galaxies and identify key open questions that this thesis aims to address.

1.1.1 Galaxy morphology

Galaxies can loosely be categorized into three morphological classes: disk/spiral, elliptical/spheroidal, and irregular (Hubble 1926). Disk galaxies are characterized by their flat, circular morphologies, and ordered rotation of stars and gas in spiral arms and/or bars in the planes of their disks. Many disk systems can also host bulge components at their centers. Elliptical galaxies are more symmetric and generally have featureless morphologies at optical wavelengths. Their stellar kinematics are described by high velocity dispersions, although some ellipticals are also characterized by rotation. Lenticular (or equivalently, S0) galaxies are hybrid systems that have flat disks and massive central bulges composed of stars, but their star formation and gas properties are more similar to those of ellipticals. Irregular galaxies are systems that do not fit into one of the above categories; merging galaxies, low-mass dwarfs, and clumpy early-Universe galaxies are all examples of irregular types.

Because a galaxy’s morphology is linked with its evolutionary history, morphological type is often correlated with other important properties (e.g., Dressler 1980). For example, many spiral and irregular galaxies are blue in color, which signifies that they contain populations of young, recently formed stars. Many elliptical galaxies are quiescent (based on stellar population synthesis models; e.g., Bruzual & Charlot 2003), meaning that they have not formed significant quantities of stars in a long time (> 1 Gyr; e.g., Trager et al. 2000), earning them the nickname “red and dead.” As is the case with all zeroth order approximations, this correlation can break down. Some spiral galaxies are
red because they contain substantial dust content, which reddens the stellar light (see, e.g., Calzetti 2001); others are red because they have truly shut off star formation (e.g., Larson et al. 1980; Wolf et al. 2009; Masters et al. 2010). There are also instances of elliptical galaxies that are able to produce new stars once more due to various physical processes (although it is possible that these galaxies are once again in the process of constructing new disks; e.g., Kannappan et al. 2009).

1.1.2 Dark matter structures

The physical processes that drive the growth of galaxies are dominated by gravity on large scales, e.g., > 1 Mpc (the characteristic separation between galaxies). On smaller scales associated with typical galaxy sizes (1 – 100 kpc) or distances that separate stars within galaxies (∼ 1 pc), thermodynamics, nuclear physics, dissipative hydrodynamics, radiative transfer, and magnetic fields also become important (e.g., Mo et al. 2010; Draine 2011; Somerville & Davé 2015). Before discussing the physics that determine the inner workings of individual galaxies, I will briefly describe the large-scale gravitational interactions that first allow a galaxy to form.

Dark matter accounts for most of the mass in the Universe (e.g., Peebles 1980). To our knowledge, it only interacts gravitationally with itself and with baryonic (i.e., non-dark) matter. Although little is known about the nature of dark matter, its gravitational effects have been empirically well studied and are essential to current models of galaxy evolution (e.g., Zwicky 1937; Babcock 1939; Rubin & Ford 1970; Rubin et al. 1980; Blumenthal et al. 1984; Buote et al. 2002; Clowe et al. 2006; Buckley & Peter 2018). Gravitational $N$-body simulations of collisionless dark matter have shown that overdensities first collapse into filamentary structures, creating a “cosmic web” that forms the backbone of large scale structure (see, e.g., Holmberg 1941; Bertschinger & Gelb 1991; Springel et al. 2005). Mass flows from lower- to higher-density regions and congregates at the intersections of cosmic filaments. When particles become self-gravitating, and eventually virialized, they organize into well-defined structures called dark matter halos (e.g., Press & Schechter 1974; van Albada 1982; Navarro et al. 1996; Jenkins et al. 2001). Gravitational instabilities on smaller scales result in the formation of lower mass substructures (or subhalos) that can exist within more massive halos. Baryonic gas can also condense toward the densest centers of dark matter halos or subhalos. When the gas density is high enough and the gas temperature is cool enough, stars can form, and eventually, a galaxy is born.

1 For massive galaxy clusters, which will be presented in Section 1.5, thermodynamics and magnetohydrodynamics can also impact galaxy evolution on ∼ 1 Mpc scales (Govoni & Feretti 2004; Fabian et al. 2006).

2 Non-dark matter is often referred to as “baryonic matter” by astronomers; this misnomer is due to the fact that most of the mass in non-dark matter can indeed be attributed to baryons. When discussing the baryonic content of galaxies, I am referring to protons and neutrons, as well as leptons such as electrons.
1.1.3 Hierarchical galaxy formation

Overdensities in the early Universe’s matter distribution will gravitationally grow more quickly than lower-density regions (e.g., Zel’dovich 1970). The gradual build-up of dark and baryonic matter is hierarchical, such that small halos and galaxies will coalesce and form more massive structures (White & Rees 1978; Baugh 2006). It is worth noting that “smooth” accretion of matter also fits into the hierarchical structure formation picture by way of a central halo merging with numerous lower-mass subhalos. Dynamical friction ensures that, over time, merging halos can relax, or virialize (e.g., Fukushige & Makino 2001).

In the hierarchical structure formation paradigm, lower-mass disk galaxies are likely to form in lower-density regions in the primordial matter distribution, whereas massive elliptical galaxies are more likely to form in higher-amplitude initial overdensities (e.g., Blumenthal et al. 1984). The clustering signal (often measured via a two-point correlation function; e.g., Landy & Szalay 1993) of galaxies is usually larger compared to that of dark matter halos (derived from simulations); the ratio of galaxy clustering to dark matter clustering is called “bias” (Kaiser 1984; Benson et al. 2000). Galaxy bias reflects the efficiency of galaxy formation as a function of the dark matter halo mass (see also Section 1.4 for more details).

1.2 The baryon cycle

Intergalactic gas can accrete into a dark matter halo, and subsequently, into the galaxy at its center. Parcels of gas may cool, become gravitationally unstable, and collapse into dense clouds within the galaxy. These dense, self-gravitating gas clouds are where stars are born. Not all of the gas that participates in star formation is forever stuck inside stellar interiors, however; some stars recycle significant fractions of their masses, enriched with freshly nucleosynthesized products, back into the galaxy and halo. This process of converting hot gas to cool gas, and then to stars, and then back into gas again, is called the **baryon cycle** (e.g., Davé et al. 2012; Lilly et al. 2013).

1.2.1 Warm and hot gas

Baryonic gas is primarily composed of hydrogen, which is by far the most abundant element in the Universe. Helium also makes up a sizeable fraction of the total gas mass, since it is the only other element produced in substantial quantity during Big Bang nucleosynthesis (see, e.g., Iocco et al. 2009; Cooke et al. 2014; Planck Collaboration et al. 2016a).[^1] At very early epochs of the Universe,

[^1]: Helium, deuterium, $^4$He, $^3$He, and minuscule quantities of tritium and $^7$Li are essentially the only elements produced in primordial nucleosynthesis. See references in text for details.
(i.e., at very high cosmological redshifts; \(z \sim 1100\)), electrons and nuclei recombine into neutral atomic gas, which proceeds to cool with the rapid expansion of the Universe. After stars begin to form, energetic radiation seeps into the interstellar and intergalactic medium (ISM and IGM, respectively), ionizing all of the neutral gas in between galaxies. This period of cosmic “reionization” occurs around a redshift of \(z \sim 6 - 10\), and represents an important transition in the early Universe (e.g., Bouwens et al. 2015; Robertson et al. 2015; Planck Collaboration et al. 2016b; Livermore et al. 2017).

Galaxies grow by accreting mass from the intergalactic medium, but after the Universe is a half billion years old (i.e., for \(z < 6\)), nearly all of the IGM gas is ionized. The majority of the present-day gas outside of galaxy halos is in a warm (\(\sim 10^5\) K) to hot (\(\sim 10^7\) K) plasma phase (see, e.g., Shull et al. 2012; McQuinn 2016). For massive halos (e.g., with dark matter mass \(M_h \gtrsim 10^{11.5} M_\odot\)), this gas is usually unable to cool as it flows into a central dark matter halo or galaxy due to its high (supersonic) infall velocities (e.g., “accretion shocks” as described by Binney 1977; Rees & Ostriker 1977; White & Frenk 1991; Okamoto et al. 2008). Most of the gas is shocked-heated to the virial temperature, which for halos of mass \(M_h \sim 10^{12} M_\odot\) well exceeds \(10^4\) K, and is unable to cool in the lifetime of the Universe (e.g., Somerville & Primack 1999; Birnboim & Dekel 2003; Mo et al. 2010). However, accreting gas is not distributed isotropically, and high-resolution simulations predict that cool gas can flow along cool filaments directly into the centers of dark matter halos (Kereš et al. 2005; Dekel et al. 2009).

Ionized gas can occupy a large volume while remaining at low densities if at high enough temperature. The circumgalactic medium (CGM) is a hot gaseous halo of mostly ionized, gravitationally bound baryons that live far (\(\sim 100\) kpc) from the galaxy centers (see, e.g., Tumlinson et al. 2017, and references within). Since the discovery of very distant quasars, whose bright continuum light can illuminate CGM gas via absorption lines along the line of sight (Bahcall & Spitzer 1969), significant progress has been made toward characterizing diffuse gas in the CGM and IGM (e.g., Gunn & Peterson 1965; Sargent et al. 1980). Observations and theory now agree that a considerable fraction of baryonic mass accreted into halos can be found in galaxy CGMs (Werk et al. 2014; Tumlinson et al. 2017; Bordoloi et al. 2018). The CGM gas can cool and precipitate back onto the galaxy, refueling future star formation events (Putman et al. 2012; Nelson et al. 2016).

1.2.2 Cool gas

The dense ISM, which refers to the gas-rich, dusty disks of star-forming spirals and mergers, is host to hot ionized plasma in addition to cooler and denser gas clouds. This stable coexistence of cold, warm, and hot gas is referred to as the multiphase (or three-phase) ISM, with each gaseous phase in pressure
equilibrium (McKee & Ostriker 1977). Each gas phase is defined by a combination of temperature, density, and volume; for example, a cold neutral medium with density $n \sim 30$ cm$^{-3}$, temperature $T \sim 100$ K, and volume filling factor \( \sim 1\% \) will be in pressure equilibrium with a warm neutral medium with density $n \sim 0.6$ cm$^{-3}$, temperature $T \sim 5 \times 10^3$ K, and volume filling factor \( \sim 40\% \) (Draine 2011). Thus, it is important to remember that physical phenomena such as gravitation, turbulent flows, star formation, dust-attenuated radiative transfer, supernovae, supersonic and super-Alfvénic shocks, and other atomic/molecular physics all occur in the context of the multiphase ISM.

Elements other than hydrogen and helium are important for physical processes such as radiative cooling and dust grain formation. In particular, the presence of dust — which is composed of carbon, silicon, and other heavier elements — can strongly impact ISM properties (e.g., Draine & Li 2007). Dust grains catalyze the formation of molecular hydrogen (H$_2$), attenuate the propagation of energetic photons, and shed energy at infrared (IR) wavelengths through thermal radiation. About half of the integrated energy in the history of the Universe has been emitted at IR and submillimeter wavelengths (see, e.g., Soifer et al. 1987; Chary & Elbaz 2001; Blain et al. 2002; Madau & Dickinson 2014). Dense gas and dust shielding enables the cores of giant molecular clouds (GMCs) to stay cold enough ($T \sim 10$ K) such that gravitational collapse can occur and form new stars.

Gas can cool via spectral line emission during the recombination of an ion or atomic de-excitation to a lower energy state. In the case of optically thin media, such as at the boundary of a dense gas cloud illuminated by a massive star, spectral line emission can escape into the ISM and, if not impeded by other opaque sources, out of its host galaxy. Although hydrogen recombination lines such as the Lyman (corresponding to electronic transitions $n_{\text{lower}} = 1$), Balmer ($n_{\text{lower}} = 2$), and Paschen ($n_{\text{lower}} = 3$) series have been observed extensively by UV, optical, and near-IR telescopes in order to trace ionized gas around newborn stars, these transitions are not responsible for most of the cooling in the ISM. The primary mechanism by which diffuse gas loses energy in star-forming galaxies is the [C II] 158 $\mu$m fine structure line at far-IR wavelengths (although [O I] 63 $\mu$m becomes the dominant coolant at $\gtrsim 100$ K in low-ionization gas; e.g., Field et al. 1969; Wolfire et al. 1995; Carilli & Walter 2013). Other notable transitions at long wavelengths (low frequencies) include the spin-flip hyperfine transition of atomic hydrogen at 21 cm (1.420 GHz; e.g., Ewen & Purcell 1951; Giovanelli et al. 2005; Blyth et al. 2016), and the rotational dipole transitions of the CO molecule at $J_{\text{upper}} \times 115.271$ GHz (e.g., Wilson et al. 1970; Young & Scoville 1991; Solomon et al. 1997; Bolatto

---

4 Giant molecular clouds (GMCs), in which hydrogen transitions from atomic to molecular phase, are the coldest ($\sim 10$ K) and densest ($\gtrsim 10^3$ cm$^{-3}$) ISM structures. GMCs are self-gravitating and thus not in pressure equilibrium with the rest of the multiphase ISM.

5 Gas cooling via IR fine-structure lines is only possible for chemically enriched gas, i.e., gas that carries carbon, oxygen, and nucleosynthesized elements. Gas that is of primordial composition relies on Ly$\alpha$ cooling at higher temperatures $T > 10^4$ K, and can continue to inefficiently cool via H$_2$ spectral lines between $10^2 \lesssim T/K \lesssim 10^4$ (see, e.g., Bromm et al. 2002).
et al. 2013). Collisionally excited “forbidden” lines such as \([\text{O II}]\ 3726,3729 \text{ Å}, \ [\text{O III}]\ 5007 \text{ Å},\) and \([\text{N II}]\ 6585 \text{ Å}\) at optical wavelengths can be found in addition to recombination lines in (star-forming) nebular regions and are useful for characterizing gas properties. At near-IR wavelengths, warm \((\gtrsim 2000 \text{ K})\) \(\text{H}_2\) gas can radiate via electric quadrupole transitions (e.g., Turner et al. 1977). Therefore multiwavelength observations that can target such lines are valuable for characterizing all phases of a galaxy’s ISM.

### 1.2.3 Star formation

Stars are born from self-gravitating molecular clouds. GMCs in typical low-z star-forming spirals have masses \(\gtrsim 10^5 \, M_\odot\) and are \(10 – 100 \text{ pc}\) in size (Kennicutt & Evans 2012). Molecular clouds form in the higher-density disks and/or spiral arms of galaxies, where pressure is higher and cooling more efficient. As gravitation or converging flows bring together molecular gas components, further collapse of the cloud may be halted by rotational motions (i.e., the conservation of angular momentum), gas pressure (or random motions), and/or magnetic pressure. The most simple toy model of cloud collapse is given by the Jeans (1902) instability, wherein the cloud’s self-gravitation exceeds its hydrostatic pressure. In more realistic ISMs, clouds are turbulent on large scales, and are thus resistant to widespread collapse. Moreover, cloud geometries are inhomogeneous and fragmentation is likely to occur (Hoyle 1953; Larson 1985). Progressively smaller and denser substructures within GMCs, termed clumps \((\sim 1 \text{ pc})\) and cores \((\sim 0.1 \text{ pc})\), can fragment from the larger complex.

Ultimately, stars are born from the densest cores that grow out of the fragmenting molecular cloud. The details of this final collapse, which transforms a protostellar object into a star, involve gas heating and ionization, formation of an accretion disk, and angular momentum shedding through bipolar jets or mass outflows (see, e.g., Shu et al. 1987; Lada & Lada 2003; McKee & Ostriker 2007). Due to the fragmentation process, molecular clouds give birth to clusters of hundreds or thousands of stars in a short span of time. It has been theorized that the mass distributions of newly formed stars follow a universal distribution (i.e., an initial mass function, or IMF; see Salpeter 1955; Kroupa 2001; Chabrier 2003), and observations appear to be in agreement (at least for stars with masses of a few solar masses; Bastian et al. 2010). The most massive stars have short lifespans, since their cores rapidly generate new elements via nuclear fusion (catalyzed by high densities and temperatures). For a massive star with 30 times the mass of our Sun (i.e., \(30 \, M_\odot\)), after about 10 Myr of stellar evolution, the core has fused into iron and degeneracy pressure is unable to prevent gravitational collapse, resulting in a core-collapse supernova (core-collapse supernovae are also labeled under Type Ib, Ic, and II supernovae). Low-mass mass stars undergo fusion at much lower rates, and thus can continue to burn hydrogen for durations exceeding the age of the Universe. Because of the strong
disparity in stellar lifetimes, stars of different masses in the same galaxy might have formed at
different times in the past, and may be described by disparate kinematics, elemental compositions,
and surrounding ISM properties.

1.2.4 Stellar evolution and nucleosynthesis

When hydrogen fusion begins at the center of a newly formed star, it is said to be a zero-age
main sequence (ZAMS) star. The IMF defines the probability distribution function of ZAMS stellar
masses: \(dN/dM \propto M^\alpha\), where observations have found \(\alpha = -2.35\) for stars more massive than \(\sim 0.5 \, M_\odot\) (e.g., Salpeter 1955; Kennicutt & Evans 2012). This implies that there are significantly more
low-mass stars than high-mass stars, although this power law relation physically and observationally
breaks down at low masses.\(^6\)

The majority of a galaxy’s optical and near-IR (NIR) light is emitted from its stars. One
consequence of the IMF shape is that most of a galaxy’s stellar mass is due to the great number of
low-mass stars. The bulk of its total starlight, however, is emitted from a tiny fraction of the galaxy’s
massive, hot, and short-lived stars. Most UV light can be attributed to such massive stars, whereas
optical and NIR radiation trace the less massive stars such as our Sun and lower-mass dwarfs. It
is worth noting that the NIR light also has strong contributions from thermally pulsing asymptotic
giant branch stars, too, and different treatments of this phase in stellar evolution models may yield
discrepant results (e.g., Maraston et al. 2006).

Ionizing, massive stars (e.g., classes O and B in the traditional categorization scheme) are par-
ticularly important for studying galaxy evolution. Their existence in any galaxy suggests that the
galaxy has formed stars “recently” (i.e., within the last \(10 - 100\) Myr). Since stars originate in
gas-rich, dusty environments, star formation can be traced by observations of the recombination
of hydrogen in natal gas clouds or thermal emission from hot dust near newborn stars. The pho-
toionized gas surrounding hot young stars constitutes “H II regions,” in which — as one traverses
radially away from the luminous sources — the gas density, neutral gas fraction, and optical depth
to ionizing photons increase (see, e.g., Strömgren 1939; Tielens & Hollenbach 1985).

As a star ages, its temperature, luminosity, and radius will change depending on its initial
mass. The highest initial mass stars have the shortest lifetimes, and in this brief amount of time
(\(< 10\) Myr) they rapidly fuse and expel heavy elements into the ISM through stellar winds or their
final supernovae. The process of populating the ISM with elements heavier than helium (defined

\(^6\)For objects with very low masses, \(\lesssim 0.08 \, M_\odot\), hydrogen cannot be fused in the core, so they are no longer stars
by definition. Stars with masses \(0.08 \lesssim M/M_\odot \lesssim 0.5\) may have truncated accretion or be ejected from their birth
clouds via n-body gravitational interactions (Krumholz et al. 2012).
by astronomers as \textit{metals}, which include, e.g., C, N, O, Si, and Fe) is called \textit{chemical enrichment}. Low- and intermediate-mass \((M/M_\odot \lesssim 8)\) stars are able to synthesize carbon through oxygen, and are abundant enough to dominate the chemical enrichment process for \(\alpha\) elements such as carbon, oxygen, neon, etc. Late in their evolution, as they finish burning helium in their cores, lower mass stars can convectively “dredge up” \(\alpha\) elements to the outer envelope; intense radiation pressure is able to generate stellar winds that cause stars to shed their enriched envelopes. High-mass \((M \gtrsim 8 M_\odot)\) stars are responsible for synthesizing nearly all of the heaviest elements such as silicon and iron, although only a small fraction is unbound during the death of the star (a core-collapse supernova) and ejected into the ISM. Type Ia supernovae, thermonuclear explosions in which at least one of the progenitor stars is an electron-degenerate carbon-oxygen white dwarf, also produce large quantities of \(^{56}\text{Ni}\) and \(^{56}\text{Co}\) that rapidly decay into \(^{56}\text{Fe}\); these are recycled directly into the ISM. Extensive reviews on the subject of stellar nucleosynthesis are given by Bethe (1939), Burbidge et al. (1957), Wallerstein et al. (1997).

The abundance of heavy elements measured in gas is called the \textit{gas-phase metallicity} (or nebular metallicity, since it is often derived from ratios of observed nebular spectral lines). Energetic stellar winds can return metals to the ISM with sufficient velocity to turbulently mix the enriched (and shock-heated) gas with the surrounding medium. In particularly energetic events, metal-laden gas may be launched out of the galaxy altogether (e.g., Somerville & Davé 2015; Prochaska et al. 2017). After some time, the enriched gas be able to cool and “rain” back down onto the galaxy disk. The galaxy may also gain pristine or low-metallicity gas from accretion events, which can depress the gas-phase metallicity (particularly near the outskirts of a disk; Tinsley & Larson 1978). Stars born out of enriched gas and/or accreted low-metallicity gas will have different \textit{stellar-phase metallicities} from stars born at earlier epochs.

### 1.3 Galaxy scaling relations

As galaxies accrete matter, their physically connected properties evolve according to various scaling relations. In other words, quantities relevant to a physical process should scale with each other (usually via a power law) in the energy or size regime where that physical process is the dominant effect. Intuitively, one may expect that as a galaxy grows in stellar mass, it should also become more luminous at optical and NIR wavelengths. A galaxy with more cold gas is expected to have higher star formation rate. If a disk galaxy is in virial equilibrium, then the velocities of stars moving along circular orbits should depend on its enclosed mass. These are just examples of galaxy scaling laws that characterize the statistical properties of galaxy populations through known physical relations.
1.3.1 The Tully-Fisher, Faber-Jackson, and $S_{0.5} - M_*$ relations

Tully & Fisher (1977) observed a correlation between a disk galaxy’s luminosity, $L$, and its maximum velocity, $V_{\text{max}}$: $L \propto V_{\text{max}}^\alpha$, where $\alpha$ is measured empirically to be about $3.5 - 4$. Luminosity is generally integrated over an optical or NIR wavelength band, and $V_{\text{max}}$ is measured from disk galaxy rotation curves at large radii, where the curve is typically observed to be near its maximum value. Sometimes the stellar mass is used directly instead of the optical/NIR luminosity, since the luminosity is meant to be a proxies for the stellar mass (Bell & de Jong 2001; Conselice et al. 2005); this variation is called the stellar mass TFR. The stellar mass TFR is a tighter correlation than the original, implying that $M_*$ is more physically connected to galaxy kinematics than luminosity. Another adaptation of the TFR is the baryonic TFR; McGaugh et al. (2000) found that an even tighter relationship exists between a galaxy’s total baryonic mass (equaling the summed stellar and gas masses) and $V_{\text{max}}$. Because the velocity depends on the gravitational potential, $V_{\text{max}}$ is a proxy for the total mass. Thus the original TFR and stellar mass TFR connect galaxies’ stellar mass and total mass, while the baryonic TFR connects baryonic mass to total mass.\(^7\)

A similar scaling relation exists for elliptical galaxies. By definition, elliptical systems do not have flat disks and are thus supported by hydrostatic equilibrium of their stars’ collisionless motions. Faber & Jackson (1976) find a strong correlation between the luminosity and velocity dispersions, with $L \propto \sigma^\alpha$ and $\alpha \sim 4$. The Faber-Jackson relation (FJR) is presented as evidence that both luminosity and velocity dispersion are controlled by the elliptical galaxy’s total mass, analogous to how luminosity and ordered rotation velocity are controlled by a spiral galaxy’s total mass.

From these scaling relations, it is evident that total mass (or gravitational potential energy) is the variable most closely correlated with a galaxy’s kinematics. Regularly rotating disk galaxies and dispersion-dominated spheroidal galaxies are simply instances of the same kinematic scaling relation. One example of a combined kinematic estimator is $S_{0.5} \equiv (0.5V_{\text{max}}^2 + \sigma^2)^{1/2}$ (e.g., Weiner et al. 2006; Kassin et al. 2007), which scales tightly with stellar mass for all galaxy populations according to the same power law index (i.e., log-log slope) as the TFR and FJR. The tight correlation between $S_{0.5}$ and $M_*$ suggests that the two quantities are good tracers of the underlying gravitational potential.

At $z \gtrsim 2$, star-forming galaxies emerge from collapsing clouds as clumpy systems with gas-rich, thick disordered disks (e.g., Lotz et al. 2006; Cresci et al. 2009; Förster Schreiber et al. 2009). As galaxies produce stars and consume gas over time, their disks become thinner and more kinematically ordered (rotation-dominated) (e.g., Cacciato et al. 2012; Wisnioski et al. 2015). Disk settling is

\(^7\)There are interesting implications for the extremely tight relationship between baryonic mass and total mass, including the possibility that baryonic matter accounts for the full gravitational potential and that modifications to Newtonian dynamics, rather than modifications to the non-dark matter paradigm, are needed (e.g., Milgrom 1983; McGaugh et al. 2016).
observed for blue, star-forming galaxies above some mass threshold $\log(M/M_\odot) > 9.5$ in the local Universe (see, e.g., Simons et al. 2015). The fraction of settled disk galaxies does not appear to evolve between $0 < z < 1.2$, despite the greater propensity for high-$z$ galaxies to be gas-rich (Kassin et al. 2012). However, galaxies at $z \sim 2$ (i.e., at the peak of Universal star formation history; Madau & Dickinson 2014) in the same stellar mass range are just beginning to assemble their disks, and thus exhibit lower rotation velocities than galaxies on the stellar mass TFR (Simons et al. 2016). These $z \sim 2$ systems are still well-characterized by the $S_{0.5} - M_*$ correlation, demonstrating that they still adhere to a Universal galaxy scaling relation even during mass accretion and star formation events.

1.3.2 The mass-metallicity relation

As discussed in Section 1.2.4, a galaxy’s stellar population will produce heavy elements over its lifetime, and return a fraction of these metals back into the ISM and IGM. For an individual star, nucleosynthetic yields are strongly dependent on its mass (and to a lesser extent, its rotation). A star’s lifetime is also effectively a function of its initial mass.\(^8\) Subsequent episodes of star formation will thus lead to chemical enrichment, after a time lag determined by the distribution of stellar masses, which is measurable and encoded in the IMF. Therefore, the metallicity of a galaxy is primarily sensitive to intermediate timescales (i.e., $\sim 0.1-1$ Gyr) of its star formation history (at least for relatively isolated systems — galaxies with recent merger events or substantial mass accretion and internal mixing can alter their chemical compositions). Contrast this with the instantaneous SFR which, by definition, traces short timescales (< 100 Myr) of the star formation history, and the total stellar mass, which traces the full time-integrated star formation history. It is therefore unsurprising that there exists a mass-metallicity relation (MZR; see, e.g., Tremonti et al. 2004), an empirical correlation between the stellar masses $M_*$ and metallicities $Z$ of galaxies.

Stellar mass can be inferred from fitting models to the spectral energy distribution (SED) of a galaxy (e.g., Conroy et al. 2009; Leja et al. 2017). For SED fitting, one compares observed broadband fluxes to modeled continuum light from stellar populations, which are parameterized by each galaxy’s stellar mass, dust content, recent SFR, and shape of the star formation history (to give a few examples). The gas-phase metallicity can be expressed as the oxygen abundance relative to hydrogen: $Z \equiv 12 + \log(O/H)$ (although note that other forms exist; e.g., Matteucci & Tornambe 1987). Numerous techniques using spectroscopic observations of the nebular gas emission lines are available for making this measurement (e.g., Kewley & Dopita 2002; Pettini & Pagel 2004; Tremonti

---

\(^8\)The chemical composition of a star when it is formed is also important for computing nucleosynthetic yields, stellar lifetimes, and perhaps even the IMF. For example, the first generation of stars that have formed from pristine, unenriched matter in the very early Universe ($z \sim 20-30$) may have been disproportionately high-mass compared to the $z \sim 0$ IMF (Bromm 2013). Since stars at these extreme redshifts are beyond the scope of this work, I will assume that a star’s elemental yield is only a function of its initial mass.
et al. 2004). Since emission lines trace star formation, MZR measurements of nebular metallicity are mostly available for star-forming galaxy samples.

The Sloan Digital Sky Survey (SDSS; York et al. 2000; Strauss et al. 2002) enabled gas-phase oxygen abundance and stellar mass to be efficiently measured for hundreds of thousands of emission-line galaxies in the nearby Universe (e.g., at $z < 0.3$). Tremonti et al. (2004) find that the MZR is extremely tight ($\sigma \approx 0.10$ dex) over about three orders of magnitude in $M_\star$. Gas-phase metallicity flattens out for galaxies above $M_\star \sim 3 \times 10^{10} M_\odot$, indicating that (at present time in the Universe), even the most massive galaxies are only able to enrich the ISM to $Z \approx 9.3$. Lara-López et al. (2010) and Mannucci et al. (2010) later found that scatter in the MZR decreased further to $\sim 0.05$ dex when galaxies’ SFR is taken into account (the fundamental metallicity relation, or FMR). The stellar-phase metallicity $Z_\star$ also correlates with $M_\star$, although the scatter intrinsic to this relationship is larger, particularly for lower-mass galaxies which are subject to more stochastic star formation (Gallazzi et al. 2005).

1.3.3 The Kennicutt-Schmidt relation

An empirical scaling relation exists between the gas surface density ($\Sigma_{\text{gas}}$) and the SFR surface density ($\Sigma_{\text{SFR}}$), called the Kennicutt-Schmidt law (KS law; Schmidt 1959; Kennicutt 1998b). The KS law is the projected version of the intuitive volume density relationship between gas mass and SFR. For unresolved observations of nearby disk galaxies, a tight scaling relation $\Sigma_{\text{SFR}} \propto \Sigma_{\text{gas}}^N$ can be measured, where $N = 1.4$ (Kennicutt 1998b). The KS law can also be observationally resolved, and comparisons between resolved $\Sigma_{\text{SFR}}$ and $\Sigma_{\text{gas}}$ showed that star formation tightly correlated with the molecular gas (rather than, e.g, the atomic H I; Wong & Blitz 2002). Hence, star formation is more commonly found at toward smaller radii in a galaxy’s disk, where molecular gas is generally most abundant (e.g., Helfer et al. 2003), rather than toward the outskirts, where diffuse H I sometimes extends to larger radii than the stars (Bosma 1981). These findings are expected for a theory in which stars form out of the densest gas (where hydrogen is predominantly molecular). Unsurprisingly, even denser molecular gas tracers such as HCN are more tightly correlated than CO with SFR (Gao & Solomon 2004).

Leroy et al. (2013) have found that, for a sample of nearby star-forming spirals observed at 1 kpc resolution, the relationship between gas mass and SFR surface densities is not only described by a power law, but also varies as a function of both global and local properties in galaxies. For example, deviations from the KS law may be attributed to different star-forming conditions in galaxies’ nucleus.

HCN traces the high-density ($\gtrsim 3 \times 10^4$ cm$^{-3}$) cores of molecular clouds, whereas CO probes lower-density gas ($\sim 2.2 \times 10^3$ cm$^{-3}$; e.g, Draine 2011; Shirley 2015).
and disk regions (see, e.g., Genzel et al. 2015), a gas-to-dust ratio that alters the conversion between CO and H$_2$ molecular gas masses (e.g., Bolatto et al. 2013), or structural processes that impact the molecular gas depletion timescale, $\tau_{\text{dep}} \equiv \Sigma_{\text{mol}}/\Sigma_{\text{SFR}}$, on scales larger than a single cloud (e.g., such as spiral arms). On scales comparable to the sizes of GMCs ($\lesssim 100$ pc), the KS law completely breaks down because recently formed stars drift away from, or dissipate their parent clouds (e.g., Onodera et al. 2010).

1.4 Feedback processes in galaxies

Physical processes in galaxies are generally complex and difficult to model from first principles, hence astronomers rely on empirical relationships and effective toy models to describe observed properties. These complex processes do appear to converge toward attractor solutions, or regularities, as few physical phenomena engage in runaway positive feedback loops. In the context of galaxy evolution, a balance between star formation and feedback results in a remarkably tight correlation between the SFR and $M_*$ of galaxies spanning more than three orders of magnitude in stellar mass and out to high redshifts (e.g., Daddi et al. 2007; Elbaz et al. 2007; Noeske et al. 2007; Whitaker et al. 2014; Iyer et al. 2018).

Feedback mechanisms are also needed from the perspective of numerical simulations. Early models of gas cooling struggled to control the rate of star formation, i.e. the overcooling problem, since simulated dense gas would cool rapidly with little thermal or kinetic resistance (e.g., Moore et al. 1999). The additions of star formation feedback through radiation pressure, winds, and supernova feedback, and luminous active galactic nucleus (AGN) feedback via energy-driven and momentum-driven winds, are able to remedy theoretical models and allow simulations to generate more realistic galaxy populations (e.g., Benson et al. 2002; Governato et al. 2007; Somerville et al. 2008; Choi et al. 2015). One of the major implications of including such physical processes in galaxy evolution models is the effect on the stellar-to-halo mass ratios $M_*/M_h$ of galaxies. Theoretical models and observations broadly agree that $M_*/M_h$ for galaxies are at least an order of magnitude lower than the cosmic baryon-to-dark matter mass ratio, $M_{\text{baryon}}/M_{\text{dark matter}} \sim 0.15$; powerful feedback processes ensure that at least some fraction of the cool gas is not converted into stars (e.g., Guo et al. 2010; Behroozi et al. 2013b; Wechsler & Tinker 2018). Moreover, the lowest-mass and

10 In astronomy, when feedback is discussed, it is nearly always implied that the feedback is negative and regulates the original physical process.

11 An ionizing UV background at high redshifts is effective at suppressing the formation of galaxies in low-mass ($\sim 10^8 M_\odot$) halos, since the typical velocities of ionized gas exceed their escape velocities (see, e.g., Somerville 2002). Because our Universe is hierarchically assembled, the absence of high-$z$ low-mass galaxies will also impact higher-mass systems at present day. Since observational biases permit us to more easily study the evolution of higher-mass systems, I will solely focus on star formation and AGN feedback.
highest-mass galaxies are the most susceptible to star formation and AGN feedback, respectively in line with observational evidence that dwarf galaxies are extremely dark matter dominated, and that the most massive galaxy clusters contain only 1% of their total mass in the form of stars. Both star formation and AGN feedback will now be discussed in more detail.

1.4.1 Star formation feedback

Stars output radiative energy and mass back into the galaxy over the course of their lifetimes, which can alter the properties of nearby gas and subsequently affect future star formation.\textsuperscript{12} Some massive stars have high enough radiation pressure that they blow away a significant portion of their outer layers (or photospheres), ejecting their synthesized heavy elements into the surrounding ISM. These stellar winds, and also the final supernovae of massive stars, can carve out bubbles and reshape the ionization structure of the ISM (e.g., Stinson et al. 2006). Supernovae are particularly effective for regulating star formation activity and reshaping the matter density profiles in lower-mass galaxies (e.g., Dalla Vecchia & Schaye 2012; Governato et al. 2012; Brooks & Zolotov 2014). In general, the stellar mass losses from winds inject most of the mass (and metals), and supernovae inject most of the kinetic energy back into the ISM.

Star formation feedback tends to be most effective at suppressing future star formation in dwarf galaxies (Brooks et al. 2013). Lower-mass galaxies are unable to retain gas that has been blown to high ($\gtrsim 40 \text{ km s}^{-1}$) velocities, and ram pressure and/or tidal stripping from a central galaxy’s disk can remove cool gas. Moreover, dwarf galaxies cannot efficiently cool their lower-density, lower-metallicity gas. In this way, star-forming dwarf galaxies may achieve one dramatic episode of stellar mass assembly, but consequently eject or superheat their remaining gas reservoirs for the rest of their lifetimes (at least, to the current age of the Universe). The inclusion of these feedback processes in numerical simulations represents a major success in reconciling theoretical and observational astrophysics (see, e.g., Munshi et al. 2013; Bullock & Boylan-Kolchin 2017). Since galaxy formation is hierarchical, any reduction in lower-mass galaxy formation would leave imprints on the mass assembly histories of higher-mass galaxies in the present-day Universe (e.g., Somerville & Davé 2015).

1.4.2 Active galactic nuclei

Nearly all non-dwarf galaxies (e.g., those with stellar masses $M_\star \gtrsim 10^{9.5} M_\odot$) are known to have a central supermassive black hole (SMBH; Kormendy & Ho 2013); the Milky Way is no exception.

\textsuperscript{12}Substantial energy is also released through neutrino emission too (for a review, see, e.g., Bethe 1990), but the effects on galaxy evolution and formation are negligible given their diminutive cross sections.
(e.g., Gravity Collaboration et al. 2018). For typical star-forming spiral galaxies, these SMBHs range in mass from $M_\bullet = 10^6 - 10^9 M_\odot$, where the black hole mass $M_\bullet$ scales with the velocity dispersion (or mass) of the bulge component (Ferrarese & Merritt 2000; Gebhardt et al. 2000). Massive elliptical galaxies and their likely progenitors — extreme starbursting mergers at higher redshifts (e.g., Hopkins et al. 2008a; Kormendy et al. 2009) — can harbor SMBHs with even higher masses (up to $\sim 10^{10} M_\odot$). For most galaxies, the central bulge containing the SMBH is also packed with stars and globular clusters, and sometimes dense gas and dust. Gas may be located at very small radii ($\ll 1$ pc) in an accretion disk surrounding the SMBH. If an actively accreting SMBH exhibits broad non-thermal radiation (e.g., at X-ray through radio wavelengths), (usually) bright spectral line emission with velocities ranging between a few hundred to thousands of km s$^{-1}$, and/or variability in continuum or spectral line emission, then it is said to be an AGN.

A galaxy’s AGN is sometimes apparent from morphological features such as (bipolar) jets or radio wavelength-bright bubbles being launched from the galaxy nucleus (e.g., Ho 2008). Energetic broadband radiation due to synchrotron emission and inverse Compton scattering from relativistic electrons can dominate the non-thermal spectrum of AGN. Some AGN are buried within layers of dense gas and dust, and their existences are inferred solely from unusually bright near- and mid-IR radiation. In addition, many AGN are observed to have strong emission lines, sometimes from highly ionized species, which may be extremely broad (i.e., well over $\sim 1000$ km s$^{-1}$) or narrow (a few hundred km s$^{-1}$; for more observed properties, see Heckman & Best 2014). Because gas dynamics near the SMBH primarily depend on $M_\bullet$, broad-line emission can be mapped to high-velocity clouds $\sim 1$ pc away from, and narrow-line emission to lower-velocity gas clouds $\sim 10 - 100$ pc away from the SMBH. Unified AGN models suggest that both types of line emission are common, but that (depending on inclination effects) the broad-line region may be obscured from view and thus undetected (e.g., Antonucci 1993; Netzer 2015).

AGN are fueled by the infall of gaseous material, torqued toward the center of a galaxy by gravitational instabilities. The simplest models assume spherically symmetric accretion (Bondi 1952), although realistic accretion models for high inflow rates (and thus high gas densities) include a geometrically thin and radiatively efficient accretion disk (i.e., resulting in an AGN; Shakura & Sunyaev 1973). For lower gas densities, much of the energy can be transported inwards toward the SMBH via radiatively inefficient flows (as appears to be the case for the center of the Milky Way, which does not host a detectable AGN; e.g., Pringle 1981; Narayan & Yi 1995; Yuan & Narayan 2014).

Radiation from an AGN may outshine the integrated luminosity of the host galaxy by a thousandfold (see, e.g., Kormendy & Ho 2013). In massive galaxies, AGN feedback can halt star formation through radiative and mechanical feedback (Choi et al. 2015). Strong UV and X-ray radiation will
heat and ionize gas in the surrounding medium, and also drive out material through outflows as the radiation pressure couples to dense gas and dust. Both photoionization and gas outflows can deplete cold gas reservoirs, inhibiting future star formation events. Thus, AGN play a strong role in influencing the evolution of galaxies, particularly the most massive systems (e.g., Kormendy & Ho 2013).

1.5 Effects of redshift and environment

Galaxies in the local Universe are characterized by lower SFRs and less AGN activity than at higher redshifts (e.g., at $z = 1$ Madau & Dickinson 2014). Changes to the Universe’s large-scale properties over cosmic time can impact galaxy evolution: cold gas mass fractions $f_{\text{gas}} \equiv M_{\text{gas}}/M_*$, gas-phase metallicities, halo masses, and ISM physical conditions in galaxies differ with redshift (e.g., Erb et al. 2006a; Tacconi et al. 2013; Springel et al. 2018; Strom et al. 2018). Galaxy interactions, such as in mergers or in dense galaxy group/cluster environments, can also influence their growth. The subsections below discuss several key differences between $z \sim 0$ and $z \gtrsim 1$ galaxy evolution, and highlight a population of rest-frame UV-selected galaxies that have been vital for studying the distant Universe.

1.5.1 Galaxy mergers

In nearby merging systems, morphological and spectroscopic signatures of bright AGN and prodigious starbursts can be seen in both the optical and far-IR (FIR) wavelengths (e.g., Sanders & Mirabel 1996; Scoville et al. 1997; Veilleux et al. 2002). For a major merger event, i.e., when the mass ratio of two colliding galaxies is less than 3:1, strong torques and tides can trigger gas inflows from the two progenitor disks toward the center of the galaxy merger. Compressive turbulence can additionally promote disk fragmentation and drive up star formation or nuclear activity. For low-$z$ systems, major merger events mark a departure from the SFR-$M_*$ correlation that describes typical star-forming galaxies (e.g., Barnes & Hernquist 1991).

Galaxy mergers are observed more frequently at $z \sim 1–2$ than in the local Universe, although it appears that the overall increase in SFR at higher redshifts cannot be solely attributed to an increase in mergers (e.g., Fensch et al. 2017). There exists a population of dust-rich, major mergers with SFRs in the range of $\sim 1000 \ M_\odot \ \text{yr}^{-1}$ called submillimeter galaxies (SMGs), which are generally among the brightest observed dusty star-forming galaxies (DSFGs; e.g., Chapman et al. 2005; Casey et al. 2014). SMGs, which were first detected by submillimeter-wavelength surveys (e.g., Blain et al. 2002; Smail et al. 2002), are dust-rich, optically obscured systems that constant higher flux
densities with increasing redshifts due to the negative $k$-correction in the Rayleigh-Jeans tail of thermal emission. SMGs may be the progenitors of giant elliptical galaxies today, which are observed to have star formation histories consistent with a burst of star formation at $z \sim 3$ and little remaining cold gas (Hopkins et al. 2006; Tacconi et al. 2008). Other gas-rich mergers, while not as intense in their star formation activity, may also have transformed into ellipticals or other quiescent galaxies (Hopkins et al. 2008b). The current evidence suggests that while major mergers are interesting laboratories for studying extreme galaxy evolution, and are essential for producing observed populations of ultraluminous infrared galaxies and SMGs (e.g., Casey et al. 2014), they are not dominant contributors to the total cosmic star formation history (e.g., Rodighiero et al. 2011; Schreiber et al. 2015; Somerville & Davé 2015).

1.5.2 Galaxy clusters

Another important effect in galaxy evolution is the assembly of large scale structure later in the Universe. The merging of dark matter halos into groups and clusters allows hot gas halos to combine and become more extended. Although it may initially seem likely that the elevated density in such environments will allow gas to cool and precipitate newborn stars, AGN heating is usually powerful enough to offset this instability (e.g., Fabian 1994; McNamara & Nulsen 2007). In a massive galaxy cluster, defined to be a self-gravitating group of hundreds or thousands of galaxies, the intracluster medium (ICM) of ionized gas can easily reach temperatures of $> 10^7$ K for clusters with mass $> 10^{14} M_\odot$ (Bryan & Norman 1998; Kravtsov & Borgani 2012). The impact of a diffuse plasma medium is visually seen for star-forming galaxies falling through the hot ICM, where the cold ISM in the disk experiences ram pressure $P \approx \rho_{\text{ICM}} v^2$ that scales with the ICM gas density $\rho_{\text{ICM}}$ and projected velocity $v$ (Gunn & Gott 1972). ICM-ISM interactions can efficiently strip away the more diffuse H I disk at larger radii from the galaxy center before impacting the denser molecular gas disk (e.g., Crowl et al. 2005; Chung et al. 2009). With the heating and loss of cold material, star-forming galaxies in clusters, and sometimes even in groups, are unable to continue forming stars for long (Peng et al. 2010).

At higher redshifts, the cluster ICMs have had less time to build up, which is partially evidenced by the dearth of X-ray detections in high-$z$ clusters (although the lack of detections is exacerbated by surface brightness dimming; e.g., Planck Collaboration et al. 2011). Moreover, velocity dispersions in distant clusters are substantially lower ($\sim 700$ km s$^{-1}$; e.g., Brodwin et al. 2011) than the $\gtrsim 13$

---

13 For telescope observations at a given wavelength, the redshifted emission may originate from a fainter portion of the spectrum, e.g., as UV photons shift to optical wavelengths, or from a brighter portion of the spectrum, as far-IR photons shift into submillimeter wavelengths. These corrections to the flux density at an observed wavelength, based on a model template SED, are termed positive and negative $k$-corrections, respectively. SMGs selected at 850 $\mu$m have comparable brightnesses at $1 < z < 6$. 
1000 km s$^{-1}$ typical for massive $z \sim 0$ clusters (Struble & Rood 1999), such that ram pressure forces are less effective at stripping gas-rich disks. Star formation is able to proceed at higher rates in such (proto-)clusters (e.g., Brodwin et al. 2013; Alberts et al. 2016). In these high-density environments (although comparatively less dense than $z \sim 0$ clusters), tidal interactions between pairs of galaxies and between the galaxy and cluster halos become more significant, galaxies are closer together and interact at lower velocities (see, e.g., Toomre & Toomre 1972). Tidal forces can both accelerate star formation by driving gas to galaxy centers, and also diminish it by unbinding gas through disk heating. The net effect of these interactions over billions of years is that cluster galaxies will eventually deplete all of their cold gas, and without any cooling or cold gas accretion, they will cease star formation activity.

1.5.3 Gas fraction

There is an observed redshift evolution in gas mass fraction, $f_{\text{gas}} \equiv M_{\text{gas}}/M_* \approx 0.1(1 + z)^2$ (Carilli & Walter 2013), and star forming galaxies at $z = 2$ host an order of magnitude higher gas mass fractions than typical $z = 0$ spirals (e.g., Daddi et al. 2010; Tacconi et al. 2010; Magdis et al. 2012). Some of this effect is due to the fact that galaxies have lower stellar masses at higher redshifts, but observations also agree that gas masses at higher-$z$ are significantly larger than at lower-$z$. In particular, the molecular gas fractions appear to be elevated in the distant galaxy samples relative to local Universe (see, e.g., Casoli et al. 1998; Helfer et al. 2003; Pelupessy & Papadopoulos 2009; Bauermeister et al. 2010). For example, typical CO-derived molecular gas masses at $z \sim 2$ range between $(3 - 10) \times 10^{10} M_\odot$, which is over an order of magnitude greater than the Milky Way’s H$_2$ mass (Kalberla & Kerp 2009).

The available gas mass for star-forming systems depends on the rates at which cold gas is consumed, accreted, heated, and cooled. Gas cooling is more efficient at high redshifts, since the cooling rate depends on the gas density. Thus the cooling timescale decreases with redshift like $(1 + z)^{-3}$, which at high redshifts is shorter than the dynamical timescale $\propto \rho^{-1/2} \propto (1 + z)^{-3/2}$ (e.g., Mo et al. 2010). The massive gas disks in high-$z$ galaxies provide more fuel for star formation. However, since molecular gas consumption timescales $\tau_{\text{dep}}$ can be as short as 10 Myr at high redshifts, the ISM must be continuously resupplied with gas from the IGM in order to match the observed populations of high-$z$ star-forming galaxies (e.g., Bauermeister et al. 2010).

It has been theorized that “cold-mode” accretion, where streams of $\sim 10^4$ K intergalactic gas can flow through dark matter halos without being stalled by shock, are important for fueling galaxy

\footnote{A proto-cluster is not particularly well defined, and is usually taken to mean a system that will later form a virialized cluster. Galaxy groups in the local Universe may be considered proto-clusters.}
growth throughout cosmic time (e.g., Kereš et al. 2005; Dekel et al. 2009; Kereš et al. 2009; Elmegreen & Burkert 2010; van de Voort et al. 2012). Some models predict that the bulk of high-z star formation, as well as the formation of giant elliptical galaxies is due to cold-mode accretion (Hopkins et al. 2010; Sales et al. 2012). Although such cold mode filaments are difficult to detect (Ribaudo et al. 2011; Kacprzak et al. 2012; Crighton et al. 2013; Fumagalli et al. 2016), current and upcoming H I surveys will be able to impose stronger constraints on neutral gas accretion. For example, CHILES (the COSMOS H I Large Extragalactic Survey; Fernández et al. 2013) has detected neutron hydrogen in an individual $z \approx 0.4$ star-forming galaxy for the first time (Fernández et al. 2016). Looking At the Distant Universe with the MeerKAT Array (LADUMA; Blyth et al. 2016) will push the redshift limit to $z < 1.45$ (i.e., the past nine billion years) before the survey is complete. With neutral atomic and molecular gas masses to be measured for the majority of cosmic history, all of the cold gas in star-forming galaxies can finally be accounted for.

1.5.4 Lyman break galaxies

Much of the known statistics for high-redshift ($z \gtrsim 2$) galaxy populations result from efficient short-wavelength color selection of distant galaxies (e.g., Madau & Dickinson 2014). Steidel et al. (1996) pioneered a method to identify $z \sim 3$ systems that had their Lyman continuum spectral breaks (e.g., substantially reduced flux to the blue of $\lambda_{\text{rest}} < 912$ Å) positioned between two broadband filters. This “Lyman break” technique has been applied using various combinations of filters, yielding color-selected samples of galaxy candidates to $z \lesssim 10$ (Steidel et al. 2003; Ouchi et al. 2004; Bouwens et al. 2007; Oesch et al. 2010; Bernard et al. 2016).

Lyman break galaxies (LBGs) are diverse systems with typical SFRs of $10 - 100 M_\odot$ yr$^{-1}$ and masses of $\sim (1 - 5) \times 10^{10} M_\odot$ (see, e.g., Shapley 2011). Large surveys and detailed follow-up observations of these systems have provided valuable insights into the high-redshift Universe. The luminosity function, or number density distribution of galaxies at a given luminosity, is helpful for constraining dust-unobscured SFRs or stellar masses at different redshifts (Hopkins & Beacom 2006; Madau & Dickinson 2014). High-resolution imaging (such as with the Hubble Space Telescope, or HST) reveals that high-z galaxies are much more compact than low-z systems, demonstrating the need for significant cosmic structural evolution (Kriek et al. 2009). Kinematic evolution of galaxies since $z \sim 2$ also appears to be necessary, given the observed co-evolution in dispersion-to-rotation velocity ratios (e.g., Law et al. 2009).

Although rest-frame UV-selected systems have been extremely fruitful for studying the distant Universe, they cannot be exclusively relied upon. LBGs are not representative of all galaxy populations at their redshifts, and can have very different properties from those selected using other
observational methods (e.g., SMGs; Chapman et al. 2005, which generally have higher dust content than LBGs). For example, although cold dust and molecular gas masses can now be measured for large samples of SMGs (Tacconi et al. 2008; Genzel et al. 2015), surveys of molecular gas in unlensed LBGs usually yield non-detections (but see also Magdis et al. 2012; Tan et al. 2013). Thus, it is not surprising that some of the ISM properties of LBGs are not well known, e.g., their (resolved) cold molecular gas and dust properties, H I content, and warm molecular gas properties and heating mechanisms. LBG populations are often not well-characterized by the Calzetti et al. (2000) dust attenuation law that describes normal z ∼ 0 star-forming galaxies (e.g., Salmon et al. 2016). Other discrepancies can be found when UV dust attenuation and IR dust emission are compared (see, e.g., Meurer et al. 1999); empirical models severely overpredict the FIR luminosities of some LBGs (e.g., Baker et al. 2001; Reddy et al. 2010). These issues suggest that cold gas and dust in LBGs — which are integral to understanding star formation properties — are still poorly understood.

1.6 Some open questions in galaxy evolution

Galaxy formation and evolution models can now explain many of the statistical properties of galaxy populations, as we have discussed above. Physical processes such as gas accretion, star formation, galaxy and cluster interactions, and feedback can produce realistic looking galaxies in simulations (Governato et al. 2007; Hopkins et al. 2014; Pillepich et al. 2019). However, the growth of galaxies is determined by the complex interplay of many physical processes, and simple model prescriptions used to resolve individual problems may cause other ones. Empirical tests of galaxy evolution models in extreme scenarios, or in the context of multiple evolutionary variables, become necessary for refining theoretical understanding. Detailed, multiwavelength observations can then offer valuable constraints for properties sensitive to such effects. In the subsections below, I will introduce the central questions to be addressed in my thesis.

1.6.1 How do molecular gas, dust, and star formation properties evolve with redshift for galaxies in the most overdense environments?

Galaxies residing in the most overdense environments, i.e., massive (≥ 10^{15} M_\odot) clusters, are subject to extreme physical processes that inhibit their star formation activity. Tidal interactions between the central cluster and galaxy halos, ICM ram pressure stripping, gas depletion due to evaporation from the hot (≥ 10^7 K) ICM, and rapid flyby events (galaxy “harassment”) have shut off nearly all star formation in massive clusters at z ∼ 0 (e.g., Dressler 1980). At higher redshifts, Butcher & Oemler (1978) find evidence of increased star formation activity for clusters, which suggests
that cluster galaxies at $z \gtrsim 0.3$ still harbor enough cold gas to form stars despite their extreme environmental conditions.

What are the cold gas and dust properties of galaxies in massive higher-$z$ clusters? On one hand, typical high-$z$ systems are usually more gas rich than at $z = 0$ (e.g., Tacconi et al. 2010). On the other hand, higher-$z$ galaxy clusters tend to have lower masses, and many of the mechanisms responsible for extinguishing star formation also become less effective at lower mass (e.g., the ICM temperature scales with cluster mass). Therefore it is difficult to separate the effects of increasing redshift (which promote star formation) and with increasing environmental density (which reduce star formation). Massive clusters at $z \sim 1$ are extraordinarily rare, with only a few having been detected via the Sunyaev & Zel’dovich (1972) effect (e.g., Menanteau et al. 2010; Marriage et al. 2011). In Chapter 2, I will introduce new FIR and submillimeter observations of four extremely massive $0.3 \gtrsim z \gtrsim 1.1$ clusters, which reveal the presence of (unexpectedly) gas- and dust-rich star-forming galaxies. This work has been published in The Astrophysical Journal, Volume 853, Issue 2, article id. 195, 25 pp. (2018).

1.6.2 What are the multiphase ISM properties of low-redshift LBG analogs?

LBG samples have revolutionized our understanding of galaxy evolution in the distant Universe. Despite this progress, many of the ISM properties of LBGs remain poorly understood, as their small angular sizes (see, e.g., Grazian et al. 2011) and faint dust emission (e.g., Capak et al. 2015) generally prevent them from being studied in detail. In particular, no observations have been able to probe the warm molecular phases of their ISMs, or study their dust attenuation properties from (rest-frame) UV to NIR wavelengths. Until the James Webb Space Telescope launches (Gardner et al. 2006), the rest-frame NIR properties of high-$z$ LBGs will stay redshifted out of view.

These observations are sorely needed for testing galaxy evolution models in the parameter space for recently formed, $M_\ast \sim 10^{10} M_\odot$, starbursting galaxies. Fortunately, Heckman et al. (2005) and Hoopes et al. (2007) have identified a sample of luminous and compact UV-bright galaxies in the local Universe that resemble distant LBGs in terms of UV-optical color, dust extinction, metallicity, SFR, physical size, and emission-line velocity width. These rare $z < 0.3$ LBG analogs can be used to address key questions that are difficult or impossible to tackle directly at $z \gtrsim 3$. In Chapter 3, I introduce a sample of seven LBG analogs targeted with for NIR integral field spectrograph. I discuss their ionized gas kinematics, multiwavelength dust attenuation properties, H$_2$ ro-vibrational line spectra (tracing warm molecular gas), and spatially resolved line ratios and excitation mechanisms. This work has been submitted for publication in The Astrophysical Journal.
1.6.3 How do LBG analogs compare to other systems in terms of kinematic scaling relations?

Whether or not galaxies adhere to scaling laws is an important observational constraint for theoretical models to match. However, it is not actually clear from observations of LBG analogs that such systems should even follow known scaling relations. LBGs and their analogs are characterized by high gas dispersions and disordered, thick disks. LBG analogs also show signatures of very recent starbursts, which may imply (temporarily) high $M_\star$ relative to total mass and result in departures from, e.g., the stellar mass TFR. In Chapter 4, I will present an investigation of how their kinematics scale with stellar (and baryonic) masses, and how they compare to other low- and high-$z$ galaxies in terms of the TFR and other generalized kinematic scaling relations. This work will be submitted for publication as Wu, Baker, et al. (2019).

1.6.4 How do optical morphological features in local star-forming galaxies connect to stellar mass and gas-phase metallicity?

The MZR and FMR are well-characterized and very low in scatter at low redshifts, implying that galaxies’ gas-phase metallicities, stellar masses, and SFRs scale in a predictable way (e.g., Tremonti et al. 2004; Lara-López et al. 2010; Mannucci et al. 2010). These connections can be particularly valuable for spectroscopically limited surveys, whereby the stellar mass and SFR can be well-modeled via SED fitting (e.g., Conroy 2013), but the gas-phase metallicity remains fairly unconstrained solely from broadband photometry. Additional information can be gleaned directly from the imaging data, such as galaxy sizes, morphologies, and color gradients, which may be physically relevant for predicting the gas-phase metallicity when spectroscopic information is not available.

Morphological properties are often encoded using simple features or empirical heuristics (e.g., Doi et al. 1993; Conselice 2003; Abraham et al. 2003). Convolutional neural networks (which fall under the category of machine learning; e.g., LeCun et al. 1998; Krizhevsky et al. 2012) may be able to automatically identify the morphological features necessary for predicting metallicity, given that they have been able to classify galaxies by morphology as well as human “citizen scientists” (Dieleman et al. 2015). In Chapter 5, I discuss a method for training deep convolutional neural networks to accurately predict gas-phase metallicity using only three-channel color images. Correlations between learned morphological features and metallicity can be productively used in two ways: (1) metallicity can be directly predicted from color imaging when spectroscopic observations are absent, and (2) the small amount of variance in the MZR or FMR may be reduced further by controlling morphological parameters. This work has been published in *Monthly Notices of the Royal Astronomical Society,*
1.6.5 How can the LADUMA survey deliver maximum scientific impact?

The LADUMA survey will be the first to detect H I in emission out to \( z \gtrsim 0.4 \) (Holwerda et al. 2012; Blyth et al. 2016). When the MeerKAT interferometer is fully commissioned and ready to take observations at frequencies \( \nu < 800 \) MHz, LADUMA will be able to detect neutral atomic hydrogen out to \( z < 1.45 \). The H I/H\(_2\) mass ratio is expected to evolve strongly over this cosmic interval (e.g., Obreschkow et al. 2009), although few observational constraints are presently available. Before the survey begins, the LADUMA collaboration will need to organize ancillary data sets, develop radio interferometry calibration pipelines, optimize continuum subtraction techniques, and test H I source extraction strategies. In the Appendix, I will discuss some of my efforts as part of the LADUMA collaboration in preparation for the LADUMA survey.
Chapter 2

Herschel and ALMA observations of star-forming galaxies in massive clusters


2.1 Galaxy evolution in massive clusters

Galaxy clusters are the most massive virialized structures in the universe. Due to interactions with the hot intracluster medium (ICM) and with other galaxy members, cluster galaxies are likely to evolve differently from field galaxies. As a result, nearby galaxy clusters are almost entirely devoid of the gas-rich, star-forming galaxies that we often see in the field; instead, they are full of quiescent early-type galaxies, particularly in their dense cores (see, e.g., Dressler 1980).

The impacts of cluster environment on star formation properties have been studied extensively in the local universe (see, e.g., Caldwell et al. 1993; Lewis et al. 2002; Kauffmann et al. 2004; Koopmann & Kenney 2004b; Boselli & Gavazzi 2006). One such example is the Virgo Cluster, in which H I-rich galaxies are still making their first passes through the cluster (Kenney et al. 2004; Koopmann & Kenney 2004b; Chung et al. 2009). Ongoing ram-pressure stripping by the hot ICM truncates infalling galaxy cold gas disks (Gunn & Gott 1972), and stripping of hot gas in their surrounding halos depletes gas reservoirs that would otherwise cool and replenish their disks (starvation/strangulation; Larson et al. 1980). These mechanisms catalyze the evolution of cluster members, leaving them with old stellar populations after their cold gas components are exhausted. Other effects, such as collisional interactions between galaxies (Moore et al. 1996; Mihos 2004), tidal interactions (Bekki 1999), viscous or turbulent stripping (Nulsen 1982), or thermal evaporation (Cowie & Songaila 1977), may play a part in their evolution as well. Most $z \sim 0$ virialized clusters reflect these quenching processes and are full of “red and dead” elliptical galaxies.

Observations of clusters at increasing redshifts reveal that their populations are more likely to include blue, star-forming galaxies (Butcher-Oemler effect; Butcher & Oemler 1978). Star formation
rate (SFR) trends with redshift have also been measured at infrared wavelengths in intermediate-redshift clusters (see, e.g., Haines et al. 2009; Finn et al. 2010). At $z \gtrsim 1.4$, cluster galaxies appear to be forming new stars at exceptionally high rates ($\sim 100 M_\odot$ yr$^{-1}$; Brodwin et al. 2013), typical of obscured field galaxies at $z \sim 2$ (i.e., Magnelli et al. 2013). Furthermore, the well-known local morphology-density relation reverses at high redshifts; in fact, galaxy SFRs in high-$z$ cluster cores exceed field galaxy SFRs at the same epochs (Tran et al. 2010; Hilton et al. 2010; Hayashi et al. 2010; Brodwin et al. 2013; Alberts et al. 2014, but see also, e.g., Ziparo et al. 2014; Popesso et al. 2015). The fraction of active galactic nuclei (AGNs) in clusters also appears to evolve significantly with redshift (analogous to the Butcher-Oemler effect; Galametz et al. 2009; Martini et al. 2013). Galametz et al. (2009) find that the prevalence of X-ray-selected AGNs increases by at least a factor of 3 from clusters at $0.5 < z < 1$ to clusters at $1 < z < 1.5$, and Hickox et al. (2009) and Alberts et al. (2016) find similar results with AGNs selected at other wavelengths.

Brodwin et al. (2013) suggest that clusters undergo an epoch of galaxy merger-driven starbursts and AGN activity at $z \gtrsim 1.4$. In this scenario, abundant cool gas in high-$z$ cluster galaxies, carried by in-falling galaxies and filaments, induces star formation activity. As stellar mass assembly ramps up in galaxies replete with dense gas, tidal forces and galaxy mergers also promote gas accretion onto the super-massive black holes at the centers of galaxies, causing AGN feedback. AGNs provide efficient heating mechanisms to quench star formation and mature post-starburst galaxies to quiescent early-types (e.g., Hopkins et al. 2008b; Somerville et al. 2008). As cluster galaxies grow in stellar mass and cold gas reservoirs are depleted, they also become quiescent, and SFRs plummet.

How does galaxy density or cluster halo mass shape the star formation histories of cluster galaxies, and how do these histories compare to those of field galaxies? Chung et al. (2011) find that, in nearby clusters, the specific SFR ($sSFR \equiv SFR/M_*$) and fraction of star-forming galaxies increase with projected distance from the cluster center — i.e., SFR anticorrelates with density even when stellar mass is taken into account. However, they find that cluster halo masses do not correlate with integrated sSFR, indicating that ram pressure stripping and galaxy harassment — both of which scale with cluster halo mass — are not important mechanisms for galaxy evolution in low-$z$ clusters. Such conclusions may not be applicable to star-forming populations in more massive clusters. For example, Boselli & Gavazzi (2006) find that ram pressure stripping and other ICM processes are most relevant to the evolution of late-type cluster galaxies in local massive clusters. Although tidal forces and galaxy-galaxy interactions induce star-formation and AGN activity and eventual gas consumption in higher-$z$ clusters with lower velocity dispersions, ICM processes such as ram pressure stripping and thermal evaporation are thought to be the main quenching mechanisms for galaxies in massive, evolved clusters.
In addition to environmental variables such as density and cluster halo mass, how do redshift and cluster dynamical state (which correspond to larger scales) impact AGN prevalence and SFR? Dwarakanath & Owen (1999) study two $z \sim 0.25$ clusters with different dynamical states in order to understand their star-forming and AGN populations as traced by radio observations. The cluster with significant substructure, indicative of an ongoing merger, contains a larger fraction of star-forming galaxies, as traced by blue optical color or low radio luminosity, than the other virialized, passive cluster. The authors propose that the different dynamical states in the two clusters are responsible for the differences in star-forming populations. Coia et al. (2005) find an anomalously high fraction of mid-infrared sources in a $z \sim 0.5$ cluster, which they attribute to the cluster’s recent merger. By juxtaposing the populations of obscured star-forming sources of this and another cluster at similar redshift, Geach et al. (2006) conclude that a combination of ICM effects and dynamical state are responsible for triggering recent star formation. Similar examples are also seen in the local universe, such as the still-merging Virgo cluster and the evolved Coma cluster. Abundant signs of recent galaxy evolution can be found in the former (Chung et al. 2009), but few are seen in the latter (e.g., Poggianti et al. 2004).

Detailed observations and analyses are necessary for understanding the growth of cluster galaxies during the transition phase at intermediate ($0.3 \lesssim z \lesssim 1.4$) redshifts in the mass regime where environmental effects are most pronounced. The existence of cold gas reservoirs, which may be traced by molecular lines or dust emission, is vital for continued star formation in clusters. One might ask how gas and star formation properties of cluster galaxies are impacted by an extreme cluster environment (i.e., how galaxies might evolve in the cores versus near the outskirts of clusters). Other properties, such as cluster dynamical state (whether a cluster is merging, dynamically disturbed, or virialized) and total cluster mass, are worth investigating in order to explore their impacts on star formation. Questions of how cluster galaxies evolve at intermediate redshifts can be comprehensively answered by studying both their star formation and their gas/dust properties.

We present a study of galaxies in four $0.3 \lesssim z \lesssim 1.1$ massive clusters selected via the Sunyaev-Zel’dovich effect (SZE; Sunyaev & Zel’dovich 1972). The clusters are selected from the LABOCA/ACT Survey of Clusters at All Redshifts (LASCAR; Lindner et al. 2015) sample. Our sample of clusters ranges in mass from $(5 - 13) \times 10^{14} \, M_\odot$, nearly an order of magnitude greater than the masses of clusters used in previous studies (e.g., the IRAC Shallow Cluster Survey; Eisenhardt et al. 2008; Brodwin et al. 2013; Alberts et al. 2014). By targeting the most massive clusters, we can see how the most extreme environments affect their galaxies – and because many clusters in the LASCAR sample are not dynamically relaxed, we can additionally study the effects of cluster mergers on galaxy properties. We focus on infrared observations that trace obscured star formation, as well as
(in the case of our two \( z \approx 1 \) clusters) the dust and cold gas content that are detectable at millimeter wavelengths.

In Section 2.2, we introduce new Herschel and Atacama Large Millimeter/submillimeter Array (ALMA) Band 6 observations, and present observations at other wavelengths used in the analysis. We describe new detections in Section 2.3 and the results of a stacking analysis in Section 2.4. In Section 2.5, we discuss the implications of these findings and compare with other results in the literature. Detailed descriptions of individual galaxy detections are presented in the Appendix.

We assume a flat, concordance \( \Lambda \)CDM cosmology with \( H_0 = 70 \) km s\(^{-1}\) Mpc\(^{-1}\), \( \Omega_\Lambda = 0.7 \), and \( \Omega_M = 0.3 \). All magnitudes are reported in the AB system.

### 2.2 Observations of Sunyaev-Zel’dovich effect-selected clusters

We use multiple datasets to study the star formation, dust, and cold gas properties of galaxies in our cluster sample. In the following subsections, we describe the relevant datasets and all steps needed to clean data and generate data products for our analysis. In Figure 2.1, we show an example ALMA continuum image of J0102 with Herschel and ALMA spectral line detections of spectroscopically confirmed cluster members. We show in Figure 2.2 fiducial SEDs from the Chary & Elbaz (2001) and Kirkpatrick et al. (2015) template libraries after redshifting to \( z = 0.87 \) in order to compare to our J0102 observations. Herschel and ALMA continuum sensitivities (3 \( \sigma_{\text{rms}} \) values) of J0102 are also shown for comparison. Details about new ALMA Band 6 observations are listed in Table 2.2.

#### 2.2.1 The sample of massive clusters

The Atacama Cosmology Telescope (ACT; Swetz et al. 2011) observed at 148 GHz a 455 deg\(^2\) patch of the southern sky in a region spanning right ascension 00\(^{\text{h}}\)12\(^{\text{m}}\) to 07\(^{\text{h}}\)08\(^{\text{m}}\) and declination \(-56^\circ\) 11\('\) to \(-49^\circ\) 00\('\), identifying 23 SZE decrements in the cosmic microwave background (CMB) as cluster candidates (Marriage et al. 2011). Optical and X-ray follow-up observations confirmed these SZE detections to be rich clusters (Menanteau et al. 2010), and spectroscopy for 16 clusters provided precise redshifts and dynamical mass estimates (Sifón et al. 2013, hereafter S13). Sifón et al. (2016) found that the median dynamical mass of the 16 clusters was \( M_{200c} \approx 8.2 \times 10^{14} M_\odot \), consistent with expectations that the ACT SZE survey would detect the most massive clusters in its volume.

The LASCAR sample was selected from the highest signal-to-noise ratio (SNR) of these SZE decrements that had not yet been observed at submillimeter wavelengths as of 2011. Nine out of

\[ M_{200c} \equiv 200(4\pi/3)\rho_c r_{200c}^3, \]  

where \( \rho_c \) is the critical density of the universe at the redshift of the cluster, and \( r_{200c} \) is the radius enclosing a density of 200 times \( \rho_c \).
ten LASCAR clusters were undiscovered before the ACT or SPT surveys (Menanteau et al. 2010; Marriage et al. 2011; Vanderlinde et al. 2010). These clusters have redshifts $z \approx 0.3 - 1.1$ and masses $M_{200c} \sim (5 - 13) \times 10^{14} M_\odot$ (Sifón et al. 2016).

We focus our analysis on a subsample of four LASCAR clusters: two are at low redshift ($z \sim 0.3$) and two are at high redshift ($z \sim 1$). ACT-CL J0102-4915 (J0102 for short; also known as “El Gordo”; Menanteau et al. 2012) appears to be caught in the midst of a spectacular merger, and ACT-CL J0546-5345 (J0546) is the highest-redshift cluster in the LASCAR sample. The two clusters at lower redshift are of comparable mass. All four are considered “disturbed” on the basis of their galaxy dynamics (see Section 4.2 of S13). These clusters are listed in Table 2.1.

### 2.2.2 Ground-based optical observations

S13 selected galaxies visually by color and brightness (from optical $gri$ imaging by Menanteau et al. 2010) as targets for optical spectroscopy. Their precise ($\Delta z/z \lesssim 0.005$) spectroscopic observations led to the identification of a few dozen galaxies per cluster via emission-line and/or absorption-line features; these objects form our sample of galaxies. Most galaxies in this catalog are red-sequence, absorption-line systems. It so happens that all S13 emission-line galaxies in the two higher-redshift clusters (J0102 and J0546) also show Ca $\Pi$ (K,H) 3950 Å absorption features. In Section 2.5.4, we discuss possible biases and resulting effects from the optical selection.

Sifón et al. (2016) produced revised catalogs based on new cluster membership criteria inspired by simulation results and redshifts corrected to the heliocentric frame. We use updated cluster properties in our analysis (e.g., dynamical mass and $R_{200c}$), but base our analysis on the original, larger spectroscopic catalog of cluster members in order to study a larger sample of star-forming galaxies and AGNs. For spectral line stacking (Section 2.4.4), we repeat our analysis with both catalogs and find that the results do not change.

### 2.2.3 Hubble Space Telescope observations

We make use of $HST$/ACS F625W, F775W, F850LP images of J0102 (Program ID: 12755, PI: Hughes; Zitrin et al. 2013) and $HST$/ACS F606W and F814W imaging of J0546 (Program ID: 12477, PI: High; Bleem et al. 2015). $HST$ images in the paper use all three bands ($rgb$) for J0102 and F814W for J0546. There is a small offset between the S13 and $HST$ astrometry, so we manually re-registered the former to align with the latter, and with Spitzer images for galaxies outside $HST$ coverage. The average corrections are about $\Delta \theta \sim 0''.7$, and maximum offsets were $\Delta \theta \sim 1''.6$. The difference between the imprecise positions used for long-slit spectroscopy (which were rounded to the nearest 0:1 $\approx 1''0$ in RA; S13) and the high-resolution $Hubble$ and $Spitzer$ observations accounts
for the astrometric shift. We then compare the new positions to interferometric observations (ATCA mapping, see §2.2.5) to verify revised S13 catalog positions by matching with radio-loud AGN and other radio sources.

### 2.2.4 Spitzer Space Telescope observations

For our analysis, we also include Spitzer/IRAC 3.6 and 4.5 µm observations to complement optical photometry and spectroscopy (Hilton et al. 2013). Spitzer catalogs are produced using the SExtractor software (Bertin & Arnouts 1996). Because rest-frame near-infrared (NIR) emission traces total stellar mass, the Spitzer observations are useful for cross-matching detections at long wavelengths, and particularly helpful for crude estimation of photometric redshifts. Spitzer detections of optical-wavelength dropouts are excluded as high-redshift galaxies, i.e., contaminants in our study of star-forming cluster members (see Section 2.8 for results on high-redshift submillimeter galaxies).

Hilton et al. (2013) have estimated stellar mass \(M_\ast\) from 3.6 µm photometry by employing a Bruzual & Charlot (2003) \(\tau = 0.1\) Gyr burst model beginning at a formation time \(z_f = 3\), and assuming a Salpeter (1955) IMF (which we convert to a Chabrier 2003 IMF by multiplying the final masses by 0.61; see Figure 4 of Madau & Dickinson 2014). Uncertainties are estimated by computing \(M_\ast\) for a range of formation redshifts from \(z_f = 2 - 5\) and measuring the resulting scatter (additional systematic errors from choice of IMF and star formation history are neglected). The same process is used for new detections found via ALMA spectral line emission. We choose not to employ optical and NIR spectral energy distribution (SED) fitting because our optical wavelength observations do not sufficiently cover all cluster galaxies, potentially leading to a bias in our stellar mass estimates as a function of clustercentric distance.

### 2.2.5 ATCA observations

We use Australia Telescope Compact Array (ATCA) continuum observations centered at 2.1 GHz as presented by Lindner et al. (2015). Radio flux densities and uncertainties are measured using the IMFIT task in the Common Astronomy Software Applications package (CASA; McMullin et al. 2007). Re-registered galaxy positions agree with bright radio sources to sub-pixel (1″) precision.

### 2.2.6 Herschel PACS and SPIRE observations

_Herschel Space Observatory_ (Pilbratt et al. 2010) observations using the Photoconductor Array Camera and Spectrometer (PACS; Poglitsch et al. 2010) and Spectral and Photometric Imaging
Receiver (SPIRE; Griffin et al. 2010) instruments were obtained for our clusters through Program ID OT2_abaker01_2 (PI: Baker). PACS 100 and 160 \( \mu m \) images were produced using version 10.3.0 of HIPE (Ott 2010) and have diffraction-limited beamwidths of 7\( '' \)2 and 11\( '' \)5, respectively. SPIRE 250, 350, and 500 \( \mu m \) observations of our clusters were introduced in Lindner et al. (2015).

Point sources with SNR \( > 4.0 \) are extracted and catalogued from each of the PACS maps using a matched-filter algorithm (e.g., Serjeant et al. 2003; Lindner et al. 2015). In Section 2.3.3, we show that the number of false positives is expected to be very low. 100 \( \mu m \) and 160 \( \mu m \) point source-subtracted maps are generated by subtracting model images, constructed from the point source catalogs, from the observed sky maps. The point source-subtracted maps contain residual positive signal due to incomplete subtraction, and we consider its effects on our stacking analysis in Section 2.4.1. Rms values in point source-subtracted maps range from 13 – 15 \( \mu Jy \) pixel\(^{-1} \) (100 \( \mu m \)) and 8 – 12 \( \mu Jy \) pixel\(^{-1} \) (160 \( \mu m \)).

### 2.2.7 ALMA Band 6 observations

We have obtained new ALMA Cycle 2 observations (Program 2013.1.01358.S; PI: Baker) of the two highest redshift clusters in our sample, J0102 and J0546. Band 6 observations were taken between 2015 January 3 and 2015 April 6. Table 2.2 summarizes the observing frequencies and dates for these observations. We show a continuum image of J0102, including the positions of cluster members detected via ALMA spectral lines, in Figure 2.1.

**Observing strategy**

For J0102, 35 available antennas provided baseline lengths ranging from 42 m to 1082 m. For J0546, around 30 antennas were consistently unflagged, providing baseline lengths ranging from 36 m to 859 m. We requested two adjacent spectral windows (1.875 GHz usable bandwidth at 15.625 MHz resolution) in each of the lower and upper sidebands for each of the two clusters. The spectral window pairs spanned 3.75 GHz and were centered at 246.6 GHz and 262.6 GHz for J0102, allowing the detection of redshifted CO \((4 - 3) \ (\nu_{\text{rest}} = 461.04 \text{ GHz}) \) and \([\text{C I}] \ (^{3}P_{1} - ^{3}P_{0})\) \((\nu_{\text{rest}} = 492.16 \text{ GHz})\) emission, respectively. Similarly, spectral window pairs were centered on 223.2 GHz and 238.2 GHz in order to detect the same lines in J0546 cluster galaxies. For J0102, we chose to image a 2.5' \( \times \) 2.0' rectangular mosaic angled 55\( ^{\circ} \) east of north, covering 41 spectroscopically confirmed cluster members, one of which is the brightest cluster galaxy (BCG). The 150 mosaic pointings are arranged in a hexagonal grid to maximize sensitivity. For J0546, we chose to image a 3.0' \( \times \) 1.7' rectangular mosaic, angled 35\( ^{\circ} \) east of north, with 142 pointings in a hexagonal grid that covers 40 cluster galaxies (including the BCG). Each pointing received \( \sim 40 \) seconds of integration.
The mosaic centers and orientations were chosen to maximize the numbers of spectroscopically confirmed members they contained.

**Calibration and imaging**

We manually reduced the J0102 data using CASA version 4.3.1 under the guidance of ALMA data analysts/NRAO staff in Charlottesville, VA. The J0546 data were automatically reduced by the ALMA pipeline, which used CASA version 4.2.2, although we tweaked some default settings for both clusters.\(^2\) To calibrate the J0102 data, the quasars J0334-4008 and J2357-5311 were used for bandpass calibration, and the latter was used for gain calibration at \(\sim 6 - 10\) minute intervals. Uranus was observed for flux calibration (Griffin & Orton 1993; Perley & Butler 2013). For the J0546 observations, the quasars J0538-4405, J0519-454, J0854+2006, and J1107-4449 were used for bandpass calibrations, J0549-5246 was used for gain calibration, and Ganymede and Callisto were used for flux calibration. System temperatures for both datasets were \(\sim 80\) K, except in the vicinity of narrow atmospheric features at 237.2 GHz, 239.1 GHz, 248.2 GHz, and 263.7 GHz, where they increased to 160 – 200 K.

Continuum images and data cubes centered on the CO and \(|\text{C I}|\) lines were produced using the standard deconvolution task in CASA, `clean`. Both clusters were continuum imaged with natural weighting, which yielded a synthesized beam of \(\sim 1.58'' \times 0.98''\) (1.85'' \times 1.08'') and a 1 \(\sigma\) continuum sensitivity of 0.11 mJy beam\(^{-1}\) (0.09 mJy beam\(^{-1}\)) for J0102 (J0546). The center frequency of the J0102 (J0546) continuum map is 254.585 GHz (230.683 GHz). Continuum-subtracted spectral line cubes were also imaged with natural weighting and with 50 MHz channel widths. The synthesized beam size is 1.5'' \times 1.0'' (1.9'' \times 1.1'') for J0102 (J0546), and the line cube has 1.5 mJy beam\(^{-1}\) (1.0 mJy beam\(^{-1}\)) 1 \(\sigma\) sensitivity per channel, apart from the narrow frequency intervals noted in the preceding paragraph.

**Astrometry**

Our interferometric source centroids are in agreement with optical and NIR imaging to within one Spitzer/IRAC 3.6 \(\mu\)m pixel (0.6') for the limited number of objects detected in both images. Most long-wavelength counterparts of optical/NIR sources are also found in the ATCA maps (pixel size = 1''). If we only examine comparisons between high-resolution observations, i.e., ALMA

\(^2\)Standard calibration settings automatically flag eight channels at both ends of each 128-channel spectral window; because we separated spectral windows by 1.875 GHz, this flagging left 125 MHz-wide gaps at the centers of our combined spectral windows. To remedy the problem, we flagged four rather than eight edge channels and included the noisier edge-channel data. In those channels, sensitivities are a factor of \(\sim 1.2 - 1.5\) worse than those in the rest of the cube.
CO (4 − 3) sources (pixel size = 0′′.15) with HST (pixel size = 0′′.04) counterparts, we find offsets of 0′′.34 (six sources). When we compare ALMA continuum sources and their HST counterparts, we measure a mean offset of 0′′.26 (six sources, three of which do not have CO counterparts). Dunlop et al. (2017) report offsets of ∼ 0′′.3 − 0′′.6 between deep HST and ALMA imaging of the Hubble Ultra Deep Field. We conclude that the astrometry in HST observations is in agreement with that of our interferometric images.

2.3 Herschel and ALMA detections

We find six ALMA CO (4 − 3) line detections corresponding to cluster galaxies, of which four are accompanied by [C I] (3P1 − 3P0) detections. Three ALMA Band 6 continuum sources are matched with cluster members in J0546; no counterparts are found in J0102. We also catalog 19 Herschel/PACS counterparts to cluster members. In Table 2.3, we summarize measurements at infrared and radio wavelengths for cluster galaxies with Herschel/PACS detections. Herschel and ALMA continuum observations are fit to Kirkpatrick et al. (2015, hereafter K15) infrared SED templates in order to estimate IR luminosity and identify AGN. In Table 2.4, we report astrophysical properties derived from NIR, FIR, submillimeter, and radio observations for cluster members with Herschel detections.

In the following subsections, we present our strategies for finding long-wavelength sources and our methods of converting measurements (flux and flux density) into physical quantities (mass and luminosity). Additionally, Appendices 2.7.1, 2.7.2, and 2.7.3 present detailed descriptions of ALMA line and continuum and Herschel sources, including information on radio-wavelength and morphological properties when available. The six ALMA line sources are displayed in Figures 2.11 and 2.12.

2.3.1 Dust continuum sources

The Rayleigh-Jeans tail of the modified blackbody emission can reliably trace cluster galaxy dust mass while also serving as a proxy for total (H I + molecular) gas mass, assuming a constant dust-to-gas ratio (Scoville et al. 2016). We use a dust-to-gas scaling law (equation 8.26) from Scoville (2013) to convert the ALMA continuum flux density into a total gas mass, $M_{\text{ISM}}$:

$$
\frac{M_{\text{ISM}}}{M_\odot} = 1.12 \times 10^{10} \left( \frac{S_\nu}{\text{mJy}} \right)^{3+\beta} \left( \frac{1}{1+z} \right)^{3+\beta} \left( \frac{350 \mu m}{\nu_{\text{obs}}} \right)^{2+\beta} \left( \frac{d_L}{\text{Gpc}} \right)^2 ,
$$

(2.1)
where $S_\nu$ is the ALMA continuum flux density, $\beta = 1.8$ parameterizes the dust absorption coefficient $\kappa \propto \nu^{-\beta}$, $\nu_{\text{obs}}$ is the observed wavelength (i.e., the continuum map central frequency), and $d_L$ is the luminosity distance. Equation 2.1 is approximately equal to equation (A14) in Scoville et al. (2016) for galaxies at $z = 1$ with dust temperature $T_d = 20$ K (see their Figure A1), assuming a Galactic $X_{\text{CO}} = 2 \times 10^{20}$ cm$^{-2}$ (K km s$^{-1}$)$^{-1}$ conversion factor (Bolatto et al. 2013).

We search for ALMA Band 6 continuum sources in the J0102 and J0546 fields by identifying $> 4 \sigma$ pixels in the ALMA maps and checking for HST or Spitzer counterparts. 23 ALMA sources are found this way, of which three are cluster members and three are later determined to be contaminants near cluster galaxies on the sky. These detections are described in Appendices 2.7.1 and 2.7.2. The remaining sources are not detected at optical wavelengths. We use the CASA IMFIT task to measure flux densities from mosaicked ALMA continuum maps. In Table 2.3, we list the dust flux densities or 3 $\sigma$ upper limits for all cluster members that have PACS counterparts. Additional ALMA continuum detections that are not associated with known cluster galaxies are described in Section 2.8. These (sub)millimeter sources are described in Table 2.8, and some examples are shown in Figure 2.13.

### 2.3.2 Millimeter line detections

We use the Source Finding Application package (SoFiA; Serra et al. 2015) to search for spectral line sources in the ALMA data cubes. We use a SNR threshold of 5.0 and smooth over 3 channels and 3 spatial pixels with Gaussian kernels. Detections without optical counterparts are disregarded. Additionally, we roughly estimate the number of false detections by extracting “sources” from versions of each data cube that have been multiplied by $-1$. The numbers of SoFiA detections found in true data cubes and in the negative versions (before assessment of counterparts) are shown in Table 2.5.

Given these results, we expect greater fidelity for the SoFiA CO line detections relative to the [C I] sources, which are less likely to be real. We indeed find that CO line detections are more likely to have short-wavelength counterparts. None of the negative CO line cube detections align with any optical or NIR sources, and only one [C I] negative line detection is within 0''.5 of a faint HST source in J0546.

For J0102, two CO and one [C I] lines detected by SoFiA have optical counterparts; however, further assessment suggest that the two CO lines are spurious (Section 2.7.1). We also manually search for ALMA spectral line counterparts of S13 cluster members and find that one [O II] emitter has CO emission at its systemic velocity. A third source is found by examining the ALMA spectra of bright optical sources with Herschel detections. In J0546, three SoFiA-detected CO lines have optical counterparts.

---

3We neglect uncertainties in the CO-to-gas mass conversion factor – although we note that adopting the Solomon et al. (1997) value, $\alpha_{\text{CO}} = 0.8$ $M_\odot$ (K km s$^{-1}$ pc$^2$)$^{-1}$, results in five-fold smaller gas mass estimates. The same is true for computation of gas mass in Section 2.3.2.
counterparts, one of which (J054638.87-534613.6) is also the lone optical and NIR counterpart of a [C I] line detection. No additional ALMA sources are found as counterparts to S13 cluster members or to other bright optical and Herschel sources. For the [C I] SoFiA line detection in J0102, we search for and find CO emission at the same redshift and position. We similarly find positive [C I] line emission for four out of the five CO-selected sources; for the final CO source, redshifted [C I] emission would lie beyond frequency range of the spectral cube, so no [C I] data are available. All detections, including spurious sources, are described in detail in Section 2.7.1.

Once detected sources have been validated, they are examined to determine the spectral extent of their line emission, and moment zero (integrated flux) maps are produced using channels with spatially coherent emission. We use IMFIT to measure flux and estimate uncertainties. If IMFIT fails to converge on a solution, we use the source’s peak flux and compute uncertainties using the standard deviation of spatially-adjacent pixels (off-galaxy) in the moment zero map. The upper limit of a non-detection is computed by multiplying three times the rms in a moment zero map generated assuming a fiducial line width of \( \Delta v = 300 \, \text{km} \, \text{s}^{-1} \) (e.g., similar to CO line widths found in Virgo spirals, low-z ULIRGs, and high-z quasar hosts; Hafok & Stutzki 2003; Solomon et al. 1997; Carilli & Wang 2006). We report CO and [C I] flux measurements in Table 2.3.

We derive molecular gas masses, \( M_{\text{mol}} \), whenever CO (4\text{--}3) is detected using a Galactic \( X_{\text{CO}} \) factor and a CO (4\text{--}3)/CO (1\text{--}0) brightness temperature ratio \( r_{4,1} = 0.4 \) (see, e.g., Carilli & Walter 2013; Topal et al. 2014) to convert CO line flux into molecular gas mass (Bolatto et al. 2013):

\[
\frac{M_{\text{mol}}}{M_{\odot}} = 1.05 \times 10^4 \left( \frac{I_{\text{CO} \,(1\text{--}0)}}{\text{Jy} \, \text{km} \, \text{s}^{-1}} \right) \left( \frac{d_L}{\text{Mpc}} \right)^2 \left( \frac{1}{1 + z} \right) .
\] (2.2)

An additional 0.3 dex uncertainty as recommended by Bolatto et al. (2013) is included in all of this paper’s calculations that depend on \( M_{\text{mol}} \).

Along with CO rotational transitions, the[C I] line can also be used as a tracer of dense gas. Gerin & Phillips (2000) analyze local galaxies to derive a line luminosity ratio of \( L'([\text{C I}] \, (^{3}P_1 - ^{3}P_0))/L'(\text{CO}(1 - 0)) = 0.2 \pm 0.2 \). Weiß et al. (2005a) find ratios between 0.15 – 0.32 at \( z \sim 2.5 \), and Walter et al. (2011) find a ratio of 0.29 ± 0.12 in their sample of \( z > 2 \) SMGs and quasar host galaxies. Using a fiducial ratio of \( L'([\text{C I}] \, (^{3}P_1 - ^{3}P_0))/L'(\text{CO}(1 - 0)) = 0.25 \) and \( r_{4,1} = 0.4 \), we expect that \( I_{\text{[CI]}} = 0.71 \times I_{\text{CO}(4\text{--}3)} \). We can thus estimate a molecular gas mass by measuring the neutral atomic carbon line flux and using Equation 2.2 as a conversion. However, because the CO emission is generally stronger than [C I], and traces a denser phase of the gas (molecular versus atomic), we use CO measurements to characterize the molecular gas reservoirs that fuel star formation.

The gas fraction, \( f_{\text{gas}} \equiv M_{\text{mol}}/M_\star \), quantifies a galaxy’s available gas for forming new stars.
relative to its existing stellar mass. The gas depletion time, $\tau_{\text{dep}} \equiv M_{\text{mol}}/\text{SFR}$ (proportional to $L'_{\text{CO}}/L_{\text{IR}}$), is another metric of how much gas can fuel star formation at its current rate. In Table 2.4, we report molecular gas masses, gas fractions, and gas depletion times computed from CO line fluxes for J0102 and J0546 cluster members.

### 2.3.3 Far-infrared continuum sources

We catalog Herschel/PACS SNR $> 4.0$ sources using a matched-filter algorithm as described in Section 2.2.6. We cross-match the 100 $\mu$m and 160 $\mu$m catalogs with the S13 confirmed cluster members and new ALMA line sources using search radii of 3$''$6, which allows the inclusion of sources that may have been assigned incorrect centroid positions. For example, Lindner et al. (2011) used a FWHM $= 11''$ matched filter to recover artificially injected sources and found $1 - 3''$ offsets between injected and recovered source centroids. We remove contaminants such as strongly lensed PACS sources or neighboring sources of emission, which in some cases are blended with the true counterparts. PACS sources that are closer in projection to other optical or NIR neighbors than to cluster members are considered contaminants. For cluster members without matches, we search for 4 $\sigma$ peaks relative to the local noise. After rejecting contaminants, we find 14 PACS 100 and/or 160 $\mu$m counterparts to S13 cluster members, and PACS counterparts to all five new ALMA CO line sources. The majority of matched PACS counterparts are offset from optical centroids by less than 2$''$, and none are offset by more than 3$''$. All 19 Herschel/PACS detections and 3 $\sigma$ upper limits are shown in Table 2.3 (not including two contaminants, which we describe in the Appendix).

We also extract sources from a negative version of the PACS images in order to check the fidelity of our catalogs. If negative “detections” are purely due to noise, and the noise profile is centered around zero, then we can use the negative catalogs to quantify the likelihood of false positives. However, we have found that the noise profile is not centered around zero, and in fact contains excess positive signal. Whether or not the signal is astrophysical in nature, it will bias our fidelity estimate of the PACS catalogs; we therefore add uniform signal to the negative PACS image such that the mean value is equal to that in the original (positive) image. We extract negative sources using the same matched-filter technique. Only a single 4 $\sigma$ detection is found in the negative J0438 100 $\mu$m map, and it is not near any known cluster members. Based on the results of this test, the probability of contaminants appearing in our catalog is small (i.e., the number of contaminants is likely less than one). The numbers of both positive and negative sources found using the matched-filter algorithm on the PACS maps are shown in Table 2.6.

Elbaz et al. (2010) have shown that the total infrared luminosity in dusty galaxies can be derived from a single PACS band at intermediate redshifts ($z \sim 0.5 - 1$) using the Chary & Elbaz (2001)
library of template SEDs. However, it has also been reported that the Chary & Elbaz (2001) SED templates, which are primarily assembled from observations of local (U)LIRGs, underestimate the cold dust content for $z \sim 1 - 2$ star-forming galaxies (see, e.g., Rieke et al. 2009; Kirkpatrick et al. 2012). Because the Chary & Elbaz (2001) templates are constructed from a disproportionately large fraction of local major mergers, they generally suffer from a dearth of cold dust emission at $\lambda_{\text{rest}} > 200$ µm. After fitting 160 µm PACS flux densities to these redshifted template SEDs, we indeed find that the Chary & Elbaz (2001) SED templates are inconsistent with observations at long wavelengths based on comparison of ALMA continuum measurements against SED predictions.4

We also use the K15 comprehensive library of infrared SED templates, which are empirically constructed from observations of $0.3 < z < 2.8$ (U)LIRGs, a population that includes ordinary star-forming galaxies, AGNs, and merger-induced starbursts. This library contains star-forming galaxies (sfg), AGNs (agn) and combinations of the two (COMPOSITE); each of these categories are further divided by luminosity and redshift. To compare monochromatic fits to Chary & Elbaz (2001) SEDs, we also fit 160 µm flux densities using only the sfg subset of templates in the K15 library. In both cases, 160 µm flux density is a better observational constraint for fitting because it lies near the SED peak, and is detected for all 19 cluster members. In order to find the best fitting template, $\chi^2$ is minimized while the overall SED normalization is allowed to vary as a free parameter. We find that $L_{\text{IR}}$ derived using Chary & Elbaz (2001) templates is systematically 0.3 dex lower compared to $L_{\text{IR}}$ computed from K15 star-forming galaxy templates. Thus, we proceed with fitting SEDs using the comprehensive K15 library, using $\chi^2$ to discriminate among templates. The best-fit template for each detection is reported in Table 2.4.

Once the best-fit template is known, we use emcee, a Markov chain Monte Carlo (MCMC) code (Foreman-Mackey et al. 2013a), to sample the posterior probability distribution of the normalization free parameter. For each galaxy, we verify that the samples follow a normal distribution. Uncertainties in $L_{\text{IR}}$ are estimated by taking a standard deviation of the distribution.

Finally, we follow Kennicutt & Evans (2012) in deriving SFR from IR luminosity assuming a Chabrier initial mass function (Chabrier 2003):

$$\log \left( \frac{\text{SFR}}{\text{M}_\odot \text{ yr}^{-1}} \right) = \log \left( \frac{L_{\text{THIR}}(3 - 1100 \mu \text{m})}{\text{erg s}^{-1}} \right) - 43.41$$

(2.3)

In this paper, we approximate $L_{\text{THIR}}(3 - 1100 \mu \text{m}) \approx 1.1 \times L_{\text{IR}}(8 - 1000 \mu \text{m})$ (see, e.g., section 2.5 of Rosario et al. 2016, who find a 4 – 15% difference between the two IR luminosity conventions).

---

4 Note that we ported the Chary & Elbaz (2001) IDL routine to Python, which uses a different interpolation routine for numerical integration. Computed $L_{\text{IR}}$ values differ randomly by about 0.03% (1 σ scatter). The code can be found at http://github.com/jwuphysics/chary_elbaz_python.
In practice, we use the conversion \( \text{SFR} = 1.65 \ M_\odot \text{yr}^{-1} \times (L_{\text{IR}}/10^{10} \ L_\odot) \). This approach, although limited, is justified because we have no mid-IR observational constraints.

### 2.3.4 Radio-bright sources

The hot and warm dust components of AGN are bright in the PACS wavebands and can contaminate any study of star-forming cluster galaxies. Out of all the S13 galaxies in our sample, only one (J010243.99-491744.4) was found to match an optical AGN spectral template. However, it seems likely that our sample of clusters contains more AGNs than the one singled out by optical-wavelength observations. We therefore search through the catalogs of 4 \( \sigma \) ATCA point sources to identify radio-loud AGN (Lindner et al. 2015). For counterparts to cluster members, we use IMFIT to measure 2.1 GHz flux density. Contamination rates are expected to be very low based on the fact that we do not find any 4 \( \sigma \) negative detections near cluster galaxies. ATCA flux densities or upper limits are shown in Table 2.3.

We also calculate the (rest-frame) 1.4 GHz luminosity by measuring \( S_{2.1\text{GHz,obs}} \) the observed 2.1 GHz flux density, and extrapolating to a rest frequency of 1.4 GHz by assuming \( S_\nu \propto \nu^{-\alpha} \) with spectral index \( \alpha = 0.8 \):

\[
L_{1.4\text{GHz}} = (4\pi d_L^2) \times S_{2.1\text{GHz,obs}} \left(\frac{1.4\text{GHz}}{2.1\text{GHz}}\right)^{-\alpha} (1 + z)^\alpha, \tag{2.4}
\]

where \( d_L \) is the luminosity distance. We estimate the AGN fraction in our sample based on the FIR-radio correlation, parameterized by \( q_{\text{IR}} \), in Section 2.5.6. Both \( L_{1.4\text{GHz}} \) and \( q_{\text{IR}} \) are reported in Table 2.4.

### 2.4 The averaged properties of faint sources

#### 2.4.1 Herschel/PACS continuum stacking

By stacking a collection of galaxies, we trade off information about individual properties for more sensitive average measurements of otherwise undetected galaxies. We stack, or co-add, flux densities of non-detections in order to reduce noise, which should decrease as \( N_{\text{gal}}^{-1/2} \) barring correlations between pixels. For 271 out of 285 cluster galaxies in our optically selected sample, no PACS emission was individually detected. We use 100 \( \mu \text{m} \) and 160 \( \mu \text{m} \) PACS “residual” point source-subtracted maps to extract fluxes from pixels centered at the locations of these undetected cluster members. Outliers have been removed by using a cut of 4 MADs, where the median absolute
deviation (MAD) of a dataset, $\bar{x}$, is defined

$$
\text{MAD} = \text{median}(|\bar{x} - \text{median}(\bar{x})|).
$$

(2.5)

We use bootstrapping (with replacement) to calculate stacking uncertainties or confidence intervals. Histograms of PACS 100 $\mu$m and 160 $\mu$m stacks (with 5000 resampling iterations) are shown in Figure 2.3 in color.

Separate stacks of pixels are also selected at random from the “blank-sky” PACS images, which are the point source-subtracted maps with (1) all *Spitzer* detections masked, in order to remove FIR contaminants, and (2) all regions beyond the projected radius of the furthest cluster galaxy masked, to exclude noisy regions in the maps. We then measure the mean blank-sky flux density after removing outliers using the 4 MAD cut. In Figure 2.3, bootstrapped distributions of the stacked galaxy flux density (colored histograms) are compared to the mean blank-sky flux densities (vertical black lines).

We see that, although the stacked galaxy signal is inconsistent with zero at the $2 - 3 \sigma$ level, it is not significantly greater than the blank-sky mean flux density. In each panel of Figure 2.3, we also report the $p$-value of the stack relative to the mean blank-sky value; the $p$-value is simply the (fractional) percentile rank of the blank-sky flux compared to the bootstrapped distribution of stacked flux. In every case, $p > 0.1$, signifying that the stacked fluxes do not statistically exceed the blank-sky values in the point source-subtracted PACS maps. Excluding the BCG from each stack does not significantly change our results. Additionally, stacking only [O II] emitters undetected by *Herschel* yields non-detections at both 100 $\mu$m and 160 $\mu$m for each cluster. Table 2.7 shows the PACS $3 \sigma$ upper limits for non-detection along with $L_{\text{IR}}$ computed using a monochromatic fit to the sfg1 star-forming galaxy K15 template. We find that the $4 \sigma$ flux limits correspond to luminosities below $\log(L_{\text{IR}}/L_\odot) < 10.6$, which we adopt as the luminosity completeness limit.

We also divide galaxies into bins of clustercentric radius, where the radii are normalized by $R_{200c}$ and centered on the BCG (or, for J0102, the midpoint of the two merging mass distributions). For each radial bin, we compute the bootstrapped distribution of stacked galaxy flux densities and compare to the blank-sky flux densities at each radius. Although we are stacking on smaller samples of galaxies than before, the goal is to examine star formation trends over a range in galaxy density (which decreases with distance from the cluster center). We still do not find any significant differences between the clustercentric stacks and the positive excess blank-sky flux densities.
2.4.2 ALMA continuum stacking

We stack the ALMA continuum flux densities of J0102 and J0546 cluster galaxies, and display them in the left panel of Figure 2.4. Bootstrapping results show that the stacked flux is consistent with zero. Blank-sky flux densities are not displayed in the figure because the mean values are very near zero. ISM masses derived from continuum fluxes binned by clustercentric radius as in Section 4.1 (right panel of Figure 2.4) are also consistent with zero at the 3 $\sigma$ level.

Our 3$\sigma$ limits on ISM mass are $\log(M_{\text{ISM}}/M_\odot) < 10.0$ for J0102 and $\log(M_{\text{ISM}}/M_\odot) < 9.9$ for J0546. Intriguingly, the J0102 stacks in two radial bins are positive at the $\sim 2\sigma$ level: we measure $\log(M_{\text{ISM}}/M_\odot) = 9.5$ ($p = 0.022$) for galaxies at $r < 0.25 \, R_{200c}$ and $\log(M_{\text{ISM}}/M_\odot) = 9.6$ ($p = 0.036$) for galaxies at $r > 0.5 \, R_{200c}$. We calculate $p$-values using bootstrapping.

2.4.3 Radio continuum stacking

We stack the observed 2.1 GHz flux densities in our ATCA maps, but bootstrapping reveals that the mean fluxes are consistent with zero. We also consider stacks in bins of clustercentric distance, but again find no evidence for non-zero stacks at any radius.

2.4.4 ALMA spectral stacking

Spectral line stacking uses positional and velocity information for confirmed (cluster) galaxies to measure their averaged spectral line properties (see, e.g., Chengalur et al. 2001). From the ALMA data cubes, we extract individual spectra positionally centered on confirmed cluster members (not including ALMA detections) and spectrally centered on their $[^3\text{P}_1 - ^3\text{P}_0]$ and CO ($4 - 3$) lines. We shift the spectra to a common velocity frame and co-add them in order to produce stacked spectra for each cluster and line. Outliers are removed at each channel using a 4 MAD cut. Because we include in the analysis galaxies that are near the frequency edges of the data cubes, with spectra that are incomplete over the velocity range of the stack, different velocity channels in the stacked spectrum have different sensitivities (scaling as $\sigma_{\text{channel rms}} = N_{\text{channel}}^{-1/2} \sigma_{\text{cube rms}}$). The number of channels going into each stacked spectrum lies in the range $23 - 31$ ($27 - 29$) in the J0102 (J0546) $[^3\text{P}_1 - ^3\text{P}_0]$ data cube, and in the range $28 - 34$ ($28 - 29$) in the J0102 (J0546) CO cube.

The stacked spectra and 3 $\sigma$ limits are shown in Figure 2.5. Stacked $[^3\text{P}_1 - ^3\text{P}_0]$ and CO ($4 - 3$) line emission are consistent with zero flux in both clusters. We stack galaxies in the Sifón et al. (2016) catalogs, which are presented in the heliocentric frame, and estimate $\pm 3\sigma$ confidence intervals based on the sensitivities at the corresponding positions and frequencies. (Stacks based on the earlier S13 catalog also yield non-detections.) It is possible that we excluded legitimate line
detections by excising 4 MAD outliers, but visual inspection reveals that all outlier fluxes were due to large primary-beam corrections (near the mosaicked image edges) or random noise.

CO non-detections offer more stringent constraints on the total gas mass, so we start from the CO channel rms $\sigma_{\text{channel rms}}$ to calculate upper limits on $M_{\text{mol}}$. We assume a conservative fiducial line width, $\Delta v = 300$ km s$^{-1}$, and use $r_{4,1} = 0.4$ as before. We find $\log(M_{\text{mol}}/M_\odot) < 9.9$ for average galaxies in both J0102 and J0546. We also stack the CO and [C I] non-detections together, assuming a constant flux ratio of $I_{\text{[CI]}} (3P_1 - 3P_0)/I_{\text{CO}} (4 - 3) = 0.71$ (see Section 2.3.2), to find upper limits: $\log(M_{\text{mol}}/M_\odot) < 9.8$ in both J0102 and J0546.

### 2.5 Galaxy evolution and the ISM properties of cluster galaxies

#### 2.5.1 Atomic carbon abundance

Among our six CO detections, we are able to measure four [C I] line fluxes. We compare integrated line fluxes in Figure 2.6. We find a mean (median) flux density ratio of $I_{\text{[CI]}}/I_{\text{CO}} (4 - 3) = 0.54 \pm 0.20$ (0.55), and an upper limit of 0.70 in one galaxy, which roughly agree with the fiducial [C I]-to-CO (4 − 3) ratio of 0.71 used in Section 2.3. In temperature units, the mean line luminosity ratio is $L'_{\text{[CI]}} (3P_1 - 3P_0)/L'_{\text{CO}(4-3)} = 0.48 \pm 0.17$. Converted to a [C I] $(3P_1 - 3P_0)/\text{CO} (1 - 0)$ ratio (again using $r_{4,1} = 0.4$), we find a mean $L'_{\text{[CI]}} (3P_1 - 3P_0)/L'_{\text{CO}(1-0)} = 0.19 \pm 0.07$. Our results agree with the line ratios found in samples of nearby (i.e., $0.2 \pm 0.2$; Gerin & Phillips 2000) and high-z (i.e., $0.29 \pm 0.12$; Walter et al. 2011) field galaxies, indicating a consistent carbon abundance and similar excitation state in our star-forming cluster members and AGNs.

#### 2.5.2 Gas mass, infrared luminosity, and gas depletion timescale

We plot all ALMA detections in Figure 2.7 along with their IR luminosities. For the three galaxies that have both ALMA CO and continuum detections, all of which belong to J0546, we have used conversion factors to estimate molecular gas mass and ISM (dust/molecular + atomic gas) mass, respectively (Bolatto et al. 2013; Scoville 2013). The mean $M_{\text{mol}}/M_{\text{ISM}}$ ratio in J0546 is $0.55^{+0.58}_{-0.34}$. For the three CO sources in J0102, none of which have matching dust continuum detections, we find a nominal $M_{\text{mol}}/M_{\text{ISM}} > 1.33$ based on 3 $\sigma$ ISM mass upper limits. For comparison, the ratio of molecular gas mass to total ISM mass predicted by simulations is $\sim 2/3$ for massive halos ($M_{\text{vir}} \geq 10^{12} M_\odot$) at $z = 1$ (i.e., Popping et al. 2015).

Disagreement between the mass ratios in the two clusters is possibly a result of small sample size, as Scoville et al. (2016) find that this ratio is constant over a wide range in mass and redshift. However, it is also plausible that physical differences between the ISM of galaxies in J0102 and
J0546 may impact the gas-to-dust ratio. Notably, J0102 is actively undergoing a merger, which may affect its galaxy members by increasing their CO excitation and thus appearing to elevate $M_{\text{mol}}$ (as calculated from a mid-$J$ CO line flux) relative to $M_{\text{ISM}}$ at an unphysical level. A higher ratio of $r_{4,1} \sim 0.8$, as is seen in local (U)LIRGs (e.g., M 82; Mao et al. 2000; Weiß et al. 2005b; Papadopoulos et al. 2012) or in QSOs (Carilli & Walter 2013), is able to resolve the conflict. Analyses of these ULIRGs have found a lower $X_{\text{CO}}$ conversion factor than what we have used (Downes & Solomon 1998), implying less massive gas reservoirs in ULIRG-like objects. Depending on the conversion factor used, gas masses may be lowered by a factor of 2—a few (see, e.g., Bolatto et al. 2013; Vantyghem et al. 2017). The ratio of $M_{\text{mol}}/M_{\text{ISM}}$ will not be further affected by choice of $X_{\text{CO}}$, since both mass estimates depend on it.

CO-detected galaxies in J0102 have star formation rates nearly an order of magnitude lower than those in J0546. As a result, we find longer $\tau_{\text{dep}}$ for J0102 galaxies in which molecular gas is measured. The average depletion time for our sample is $\tau_{\text{dep}} \sim 2$ Gyr. However, for one CO non-detection in J0546, we find $\tau_{\text{dep}}$ as low as 0.5 Gyr. Our CO sources are characterized by longer depletion timescales than normal star-forming field galaxies at $z \sim 1$ (e.g., Carilli & Walter 2013; Genzel et al. 2015), for which $\tau_{\text{dep}} \sim 0.1 - 1$ Gyr. We compare our sources to two CO detections of luminous infrared galaxies in the outskirts of a rich $z = 0.4$ cluster (Geach et al. 2009, 2011). They find gas depletion timescales of $\sim 300 - 900$ Myr (calculated using a Galactic CO-to-H$_2$ conversion factor), which are on par with only the most efficiently star-forming members of J0546. Similar gas depletion timescales are found in three gas-rich galaxies within the virial radius of two $z \sim 0.5$ clusters (with $M_{200c} \sim 4 \times 10^{14}$ $M_\odot$; Jablonka et al. 2013), and in a $z \sim 1$ massive ($\gtrsim 10^{14}$ $M_\odot$) $z \sim 1$ cluster (Wagg et al. 2012). At higher redshifts ($z \sim 1.6$), however, Noble et al. (2017) and Rudnick et al. (2017) have found longer $\tau_{\text{dep}} \sim 0.7 - 3.0$ Gyr in cluster galaxies, which agree well with the timescales in our two massive clusters.

2.5.3 Substructure: planes of infrared-bright galaxies?

The 4 $\sigma$ detection threshold in our PACS maps corresponds to galaxy luminosities of $\log(L_{\text{IR}}/L_\odot) \approx 9.5$ for our two lower-$z$ clusters and to $\log(L_{\text{IR}}/L_\odot) \approx 10.6$ for our two higher-$z$ clusters. For convenience, we define an infrared-bright galaxy (IRBG) as one with IR luminosity in excess of $\log(L_{\text{IR}}/L_\odot) = 10.6$ in order to facilitate a complete census of PACS sources across all clusters. In both J0102 and J0546, IRBGs appear to lie along planes comprising five and six galaxies, with those in J0102 appearing to lie along a plane perpendicular to its merging axis. Although the planes look significant by eye, they turn out to be fairly common occurrences by chance selection of galaxies. For each cluster, we estimate significance by (1) summing distances (i.e., computed residuals) from
the IR-bright galaxies to a best-fit plane, and (2) selecting similar subsamples of random cluster members and calculating residuals to their respective best-fit planes. Our planes of IRBGs were less significant than randomly-chosen cluster members the majority of the time \( (p = 0.88 \text{ for } J0102 \) and \( p = 0.63 \text{ for } J0546) \). Therefore, chance alignments in our higher-\( z \) clusters are likely to form planes of IRBGs (or in other words, a handful of galaxies in a cluster with \( 50-90 \) members can easily trace a statistically insignificant “plane” configuration).

### 2.5.4 Biases from optical selection of cluster members

The original targets for the optical spectroscopy we have used to identify cluster members were visually selected, on the basis of their \( gri \) colors and brightnesses, to favor redder, brighter galaxies (S13). Thus, it is no surprise that cluster members with \([\text{O II}]\) emission lines comprise only a small fraction of the original sample (17/285 \( \approx 6\% \)). For comparison, the Gemini Observations of Galaxies in Rich Early Environments survey (GOGREEN; Balogh et al. 2017) have also observed J0546; their fraction of \([\text{O II}]\) emitters found so far is 7/28 (M. Balogh, private communication), which is about twice the fraction in our sample (6/49). Since detection of \([\text{O II}]\) serves as a viable indicator of \textit{Herschel} emission,\(^5\) we have grounds for concern that our optical selection induces a general bias in our search for dust-obscured galaxies, and thus, in our stacking analysis. As an extreme example, in J0235, our lowest-\( z \) cluster, we find a low fraction of \([\text{O II}]\) emitters (1/82 members) relative to the other clusters, with all \textit{Herschel} detections in that cluster having absorption-line counterparts. It is important to note that we needed to perform an astrometric shift of \( \sim 0\prime 7 \) in order to align S13 positions with \textit{HST} imaging when available (i.e., for our higher-\( z \) clusters, Section 2.2.3). If the spectroscopic slits were not originally centered on the nuclear regions of cluster galaxies, then the observations may have systematically missed spectral line emission from star formation, resulting in our low \([\text{O II}]\) fractions in J0235 and J0546 (and perhaps in our other two clusters as well).

An additional population of optically faint or obscured sources, some of which are detectable via dust emission, may lurk below the \([\text{O II}]\) detection threshold. The \([\text{O II}]\) completeness limit in S13 is not provided, but observations to similar depth \( (g < 22.8) \) at \( z \sim 0.3 \) yield detections down to \( \log(L_{[\text{OII}]}/\text{erg s}^{-1}) \sim 41 \) (see, e.g., the SCUSS+SDSS sample used in Comparat et al. 2015). We would thus expect to detect star-forming galaxies with \textit{unobscured} SFR \( \sim 1 \, M_\odot \, \text{yr}^{-1} \) (e.g., Kennicutt 1998a; Kewley et al. 2004). We can compare this with our J0235 detection threshold for \textit{Herschel}/PACS-derived \textit{obscured} SFR, which is approximately 0.6 \( M_\odot \, \text{yr}^{-1} \). If we make the broad assumption that unattenuated and attenuated SFR components are roughly equal, our

\(^5\)6/14 \textit{Herschel} sources in the S13 sample are \([\text{O II}]\) emitters, and 6/17 \([\text{O II}]\) emitters have secure \textit{Herschel} counterparts.
Herschel-selected sample should probe slightly lower SFRs than the [O II]-selected sample. This asymmetry would disappear if star-forming cluster members were dust-poor, but would be more pronounced for heavily obscured galaxies. If such observational systematics have left emission-line galaxies underrepresented, it may not be surprising that Herschel counterparts have been found for cluster members lacking [O II] emission. This appears to be the case for Ca II absorbers with significant Herschel/PACS emission in J0235; fits to K15 suggest that their average SFR is a modest $1.4 \, M_\odot \, \text{yr}^{-1}$, which is only just above the expected [O II] detection threshold.

Part of the selection bias can be remedied with our new ALMA observations, which reveal dusty and gas-rich sources. Our analysis is insensitive to the most heavily obscured sources because we require all cluster members to be detected at optical wavelengths. However, six ALMA sources selected via CO emission are found to be cluster members (and have optical counterparts), and only one of them was previously identified in S13 as an [O II] emitter.\(^6\) The small samples of CO and [O II] detections mean that we can not securely conclude whether or not the ALMA observations correct for the selection effects on our optical sample. Nonetheless, their relatively disjoint populations indicate that long-wavelength spectroscopy is important for facilitating a more complete census of star-forming cluster members.

The members detected in CO and [C I] have rich molecular gas reservoirs with masses $M_{\text{mol}} \sim (2 - 6) \times 10^{10} \, M_\odot$. Other measurements in the literature (e.g., Geach et al. 2011; Wagg et al. 2012; Jablonka et al. 2013), as well as our own stacking results of [O II]-detected galaxies, show that these gas-rich sources are rare in massive clusters at intermediate redshifts. Future (sub)millimeter surveys will likely identify star-forming galaxies based on optical (emission line) priors, as we have seen. Otherwise, FIR or millimeter spectroscopy is needed to discover cluster galaxies that are optically faint due to heavy dust obscuration (see, e.g., Rudnick et al. 2017).

### 2.5.5 Redshift evolution of infrared-bright galaxies

For our massive clusters, we find that the fraction of galaxies that are IRBGs increases with redshift: 0/82 in J0235, 1/65 in J0438, 6/91 in J0102, and 6/52 in J0546. We consider IRBGs (rather than all Herschel counterparts) because they are a luminosity-limited population. For a fairer comparison, we restrict our analysis to galaxies within $r < 0.5 \, R_{200c}$ so that the higher-$z$ clusters can be compared to lower-$z$ clusters (for which the same FOV corresponds to a smaller physical area). The number of IRBGs then rises from 0/82 in J0235, to 1/63 in J0438, to 4/45 in J0102, and to 4/29.

\(^6\)We find that J010255.50-491416.2, the galaxy detected in both CO and [O II] line emission, has the lowest IR luminosity of all higher-$z$ cluster members detected by Herschel. J010255.50-491416.2 appears to be a face-on galaxy, for which the line-of-sight dust column is expected to be minimized, making it more likely that bluer light (and therefore, the [O II] doublet emission) would be detected. For more discussion and HST imaging of this object, see Section 2.7.1.
in J0546. To assess the significance of this apparent redshift evolution, we calculate 68% binomial confidence intervals for the IRBG fraction in each cluster: $f_{\text{IRBG}} = 0^{+0.012}_{-0.009}, 0.016^{+0.025}_{-0.009}, 0.089^{+0.052}_{-0.034}$, and $0.138^{+0.076}_{-0.052}$. Using these uncertainties, we minimize $\chi^2$ to find the maximum a posteriori power-law model fitting $f_{\text{IRBG}}$ to redshift. A multiplicative coefficient and power-law spectral index are allowed to vary as free parameters and we force the model always to output a positive IRBG fraction. We sample the marginalized distribution of the power-law index parameter using emcee. Despite the small number of IRBGs in our galaxy sample, we find that $p < 0.005$ for a zero or negative slope, indicating a definite redshift evolution in the IRBG population of our massive clusters. This evolution is consistent with a picture in which the prevalence of obscured, strongly star-forming and accreting systems increases with redshift (e.g., in the field; Le Floc’h et al. 2005; Rujopakarn et al. 2010; and in cluster environments; Haines et al. 2009).

2.5.6 The FIR-radio correlation and AGN fraction

The far-infrared and radio emission of star-forming galaxies are tightly correlated over many orders of magnitude in luminosity and flux density (de Jong et al. 1985; Helou et al. 1985; Condon et al. 1991). Following Helou et al. (1985), we define the logarithmic ratio between the FIR flux and non-thermal radio ($\nu_{\text{rest}} = 1.4$ GHz) flux density as

$$q_{\text{IR}} = \log \left( \frac{S_{\text{IR}}}{(3.75 \times 10^{12} \text{ W m}^{-2})} \right) \left( \frac{S_{\nu, 1.4 \text{ GHz}}}{(\text{W m}^{-2} \text{ Hz}^{-1})} \right),$$

(2.6)

where the FIR flux is $S_{\text{IR}} = L_{\text{IR}}/4\pi d_L^2$ (and $d_L$ the luminosity distance), and $S_{\nu, 1.4 \text{ GHz}}$ is the $k$-corrected ATCA flux density. The radio detection limit in our high-$z$ clusters is $\log(L_{1.4 \text{ GHz}}/\text{W Hz}^{-1}) \approx 23.5$. At these radio luminosities, AGNs exceed star-forming galaxies by at least an order of magnitude in the field (Condon et al. 2002; Mauch & Sadler 2007; Ivison et al. 2010). However, our sample of Herschel and/or ALMA-detected cluster galaxies are selected by infrared rather than radio fluxes. We find that six cluster members have measurable radio emission and 13 do not (although in some cases, we see $\sim 3 \sigma$ peaks in the radio surface brightness; see Section 2.7 for details on individual galaxies).

We calculate $q_{\text{IR}}$ values, or assign $3 \sigma$ lower limits when no radio counterpart is detected, and plot them against $L_{\text{IR}}$ in Figure 2.8. In line with expectation, the optically confirmed AGN in the S13 catalog has the highest radio luminosity and the lowest $q_{\text{IR}} = 1.3 \pm 0.3$ within our sample. For the other radio detections, we find a mean $q_{\text{IR}} = 1.54 \pm 0.22$. Previous studies have established $q \approx 2.3$ for star-forming galaxies (Yun et al. 2001; Condon et al. 2002; Bell 2003). Ivison et al. (2010) find a median $q_{\text{IR}} = 2.40 \pm 0.24$ when studying a sample of Herschel 250 $\mu$m-selected galaxies (we show...
this as a dashed black line in Figure 2.8). Additionally, a separator of \( q_{\text{IR}} \approx 1.8 \) is sometimes used to differentiate between star-forming galaxies and AGNs (see, e.g., Condon et al. 2002).

Four out of the 19 Herschel detections in our study have \( q_{\text{IR}} < 1.8 \) and appear to be AGNs, in addition to all four being IRBGs. This AGN fraction \( f_{\text{AGN}} \gtrsim 0.20 - 0.30 \) is large compared to values found in other studies of IR-selected AGN in \( z \sim 1 \) clusters (e.g., \( f_{\text{AGN}} \sim 0.1 \) in Martini et al. 2013; Alberts et al. 2016). One cluster member, J054638.87-534613.6, is a new CO detection and we cannot examine its optical spectrum to check for emission lines. Its IR SED is best fit by an AGN template, although this classification may be incorrectly driven by its bright dust continuum emission. A star-forming galaxy template with less warm dust would imply a lower \( L_{\text{IR}} \) and thereby drive down \( q_{\text{IR}} \), which would reinforce its classification as a radio AGN. The other three radio-loud AGNs have both [O II] and radio emission. Note, however, that the uncertainties on \( q_{\text{IR}} \) are large, mainly reflecting scatter in the IR luminosity fit. The two AGNs in J0102 appear to be secure classifications, but the other two are less certain since they have \( q_{\text{IR}} \) within 1 \( \sigma \) of the dividing line.

### 2.5.7 Star formation rate as a function of environment?

In Figure 2.9, we plot the luminosity contribution from dust-obscured star formation, \( L_{\text{IR}}^{\text{SF}} \), against clustercentric radius for all cluster galaxies with Herschel detections. We also include sources that are labeled as AGNs （\( q_{\text{IR}} < 1.8 \); shown as open circles); all infrared luminosities have been corrected for the contributions from AGN implied by the best-fit K15 templates. For J0102, we compute clustercentric radius using the midpoint between its mass peaks determined via weak lensing measurements (Jee et al. 2014). For the other clusters, we use the BCG as the cluster center. There appears to be a modest trend of increasing \( L_{\text{IR}}^{\text{SF}} \) with radius, from which we might be tempted to infer a relationship between SFR and environmental galaxy density. However, the decrease in SFR with clustercentric distance could also be a reflection of (or driven by) the SFR–\( M_* \) correlation (see, e.g., Noeske et al. 2007; Elbaz et al. 2007; Daddi et al. 2007). Stellar mass tends to increase with density, although the effects of mass and of environment should be separable (see, e.g., Peng et al. 2010). Other authors have argued that the correlation between SFR and environment is preserved even when stellar mass is controlled for (Koyama et al. 2013; Guglielmo et al. 2015): the cluster environment affects the fraction of star-forming galaxies, but has little effect on an individual galaxy’s star formation rate, which is instead correlated with that galaxy’s halo and stellar properties.

In order to separate the effects of cluster environment and galaxy mass, we compute specific SFR （sSFR = SFR/\( M_* \)). Stellar mass is derived as described in Section 2.2.4; the average (median) \( \log(M_*/M_\odot) = 11.0 \) (11.2). Both \( M_* \) and sSFR are included in Table 2.4. For our sample, we compute a median \( \log(\text{sSFR}/\text{yr}^{-1}) = -10.07 \), and average \( \log(\text{sSFR}/\text{yr}^{-1}) = -10.36 \pm 0.57 \). We
compare our sample to those found at $z \sim 1$ in the GOODS field and at $z \sim 0$ in the Sloan Digital Sky Survey (SDSS; Elbaz et al. 2007): $\text{SFR}_{z=1} = 7.2^{+7.2}_{-3.6} \times (M_* / 10^{10} \, M_\odot)^{0.9} \, M_\odot \, \text{yr}^{-1}$, and $\text{SFR}_{z=0} = 8.7^{+7.4}_{-3.4} \times (M_* / 10^{11} \, M_\odot)^{0.77} \, M_\odot \, \text{yr}^{-1}$. The SFR–$M_*$ correlations at $z \sim 1$ and $z \sim 0$ for $\log(M_*/M_\odot) = 11.0$ are respectively shown as dashed and solid horizontal lines in Figure 2.10, in which we show the sSFR plotted against clustercentric radius. The relationship between star formation and radius seen in Figure 2.9 has completely vanished once we normalize the SFR by mass. (These results do not change for J0102 even if we consider the BCG to be its cluster center.)

sSFR appears to increase with cluster redshift, similar to the results seen in Section 2.5.5. We fit a straight-line model to the cluster-averaged log sSFR against redshift, varying the slope and intercept as free parameters. We determine the best-fit slope to be positive at $p < 0.001$ significance using emcee to sample its posterior distribution. Therefore, the redshift evolution in sSFR is still evident, and can be seen in Figure 2.10. It appears that the SFR-radius trend seen before is driven by a radial gradient in stellar mass. We do not see a difference in the effects of environment on sSFR between the cluster center and 0.8× the virial radius. However, our results are consistent with a scenario in which the cluster environment depresses sSFRs below the SFR–$M_*$ correlation measured for field galaxies at similar redshift.

We have also examined the radial distributions of the gas depletion time and gas fraction in the two higher-$z$ clusters. Remarkably, the galaxy with highest gas fraction, J010252.44-491531.2, is located near ([projected distance $< 0.1 \, R_{200c}$]) the J0102 cluster merging center. However, with only six CO detections, we do not find any measurable correlations between cold gas properties and environment. The mean gas fraction is $f_{\text{gas}} = 0.22$.

Most luminous star-forming galaxies live in the outskirts of intermediate redshift clusters (Haines et al. 2009). CO has likewise been detected at large clustercentric distances, $\sim 1 - 3 \times$ the virial radius, in other intermediate-redshift clusters (Geach et al. 2011; Wagg et al. 2012; but see also Jablonka et al. 2013). If we assume that gas fraction decreases as we move toward the cluster center, then it would not be surprising if the average gas fraction $f_{\text{gas}} \sim 0.2 - 0.3$ found in the field at $z \sim 1$ (see, e.g., Tacconi et al. 2013; Genzel et al. 2015) is also characteristic of cluster galaxies at large clustercentric distances. Even if cold gas reservoirs are not preferentially less massive near cluster cores, the same stellar mass gradient that dominated our SFR-environment trend might be expected to reduce gas fractions at small radii. However, we have found six CO detections with $f_{\text{gas}} \sim 0.1 - 0.3$ located within $R_{200c}$ of their cluster centers. It is noteworthy that these gas-rich, star-forming cluster members, with gas fractions comparable to those of $z \sim 1$ field galaxies, reside in the cores of the most massive galaxy clusters at the same redshifts.
2.6 Conclusions

We have presented new *Herschel* PACS 100/160 µm and ALMA Band 6 observations of four massive, SZE-selected galaxy clusters. Based on our analysis, we conclude the following about the star formation, cold gas, and dust properties of our galaxy sample:

1. We detect 19 *Herschel*/PACS counterparts to galaxies in our sample of four massive SZE-selected clusters. Five are newly confirmed as cluster members by CO (4−3) and/or [C I] (3P_1−3P_0) detections in ALMA data cubes.

2. We detect dust emission in three cluster galaxies, which we use to calculate ISM (H_2 + H I) masses. All detections are from galaxies in the J0546 cluster. ALMA-traced gas and dust masses correlate as expected with *Herschel*-derived IR luminosities. While ISM and molecular gas masses largely agree, the mass ratios differ between clusters: \( M_{\text{mol}}/M_{\text{ISM}} > 1.33 \) for J0102 and \( M_{\text{mol}}/M_{\text{ISM}} = 0.55^{+0.58}_{-0.34} \) for J0546. This apparent discrepancy may be a result of enhanced CO excitation in the violent merger in J0102, or may simply reflect the small size of the detected sample.

3. We find a mean [C I]-to-CO (4−3) ratio of \( 0.54 \pm 0.20 \) in flux units, corresponding to a line ratio of \( L'_\text{[CI]}(3P_1-3P_0)/L'_\text{CO}(1−0) = 0.19 \pm 0.07 \). Our results agree with previous literature measurements for both local and high-z field galaxies. We also find that cluster-averaged sSFR increases significantly with redshift.

4. We find strong evidence for an increase in the prevalence of infrared-bright galaxies (IRBGs; \( \log(L_\text{IR}/L_\odot) > 10.6 \)) with redshift, for the inner regions of our cluster sample.

5. We place upper limits on dust and dense gas mass via stacking ALMA continuum and line non-detections in the higher-z clusters. Our 3 \( \sigma \) limits on ISM mass are \( \log(M_{\text{ISM}}/M_\odot) < 10.0 \) for J0102, and \( \log(M_{\text{ISM}}/M_\odot) < 9.9 \) for J0546. By stacking CO and [C I] lines together, we constrain the mean molecular gas mass for galaxies in our sample: \( \log(M_{\text{mol}}/M_\odot) < 9.8 \) for both J0102 and J0546.

6. By using radio interferometric observations, we estimate the FIR-radio correlation parameter, \( q_{\text{IR}} \), for our 19 IR-detected cluster members. Five have both *Herschel*/PACS and ATCA counterparts, including one optical AGN, and we find that the mean \( \langle q_{\text{IR}} \rangle = 1.54 \pm 0.24 \). Four galaxies are clearly below \( q_{\text{IR}} = 1.8 \) – a threshold commonly used to separate AGNs from star-forming galaxies – all of which are also IRBGs. This lower limit on the AGN fraction,
$f_{\text{AGN}} \gtrsim 0.2$, is approximately twice the fraction seen for FIR-selected galaxies in less-massive, $z \sim 1$ clusters.

7. We find modest radial trends in SFR, with decreased SFR at small distances from the cluster center. However, no significant correlation is found when we examine sSFR binned by radial distance. The mean (median) $\log(\text{sSFR/yr}^{-1})$ of our sample is $-10.36 \pm 0.57$ ($-10.07$), consistent with the SFR–$M_*$ correlation seen at low redshift.

8. For CO detections, we find an average gas fraction $f_{\text{gas}} \approx 0.2$ in the cores of massive $z \sim 1$ clusters, consistent with those of field galaxies at the same redshift. However, the average gas depletion timescale, $\tau_{\text{dep}} \approx 2$ Gyr, is long relative to those of $z \sim 1$ field galaxies.

2.7 Appendix A: New ALMA and Herschel detections

We describe new galaxies in our high-$z$ clusters that have ALMA and/or Herschel source detections in Sections 2.7.1, 2.7.2, and 2.7.3. All new line detections are shown in Figures 2.11 (J0102) and 2.12 (J0546). In some cases, cluster galaxies are at small angular separations from background or foreground galaxies that are bright at long wavelengths, and it becomes challenging to deblend the sources of emission. Cluster galaxy names preceded by an asterisk (*) are those that have PACS or ALMA emission dominated by such contaminants, or are otherwise spurious, and are excluded from the stacking analysis of Section 2.4.

2.7.1 ALMA line detections

**J010252.44-491531.2** (Figure 2.11a) is a relatively compact galaxy at redshift $z = 0.8610$. CO (4–3) is detected at about 6 $\sigma$ significance, and we see a 3 $\sigma$ peak to its south, suggesting that its molecular gas reservoir is large in size. (Our [C I] spectral coverage does not extend to this redshift, so we cannot use it for comparison.) No radio continuum emission is detected, giving a lower limit $q_{\text{IR}} > 1.7$.

**J010252.56-491400.5** is a SoFiA-detected CO line redshifted to $z = 0.8735$, with an integrated flux $I_{\text{CO}} = 0.82 \pm 0.33$ Jy km s$^{-1}$. It has an optical counterpart, a red elliptical absorption-line cluster galaxy ($z = 0.8817$), which is separated by a rest-frame velocity $\Delta v \sim 2500$ km s$^{-1}$. We do not detect Herschel, dust continuum, or [C I] line emission, so this detection is likely spurious and is excluded from the list of detections.
*J010255.27-491441.8* is a SoFiA-detected CO line at \( z = 0.8678 \) with a spheroidal counterpart seen in *HST* imaging. Upon visual inspection, none of the channels exceed 2 \( \sigma \) significance, and the emission centroids are not aligned across channels. Moreover, we do not detect *Herschel*, dust continuum, or [C I] line emission, so this detection is likely spurious and is excluded from the list of detections.

**J010255.50-491416.2** (Figure 2.11b) is a face-on spiral galaxy detected in [O II] emission by S13. A nearby source 7\" to the north has bright emission at \( \lambda = 100 - 350 \) \( \mu m \), which blends in with the galaxy. We measure the FIR flux density from the peak surface brightness. Radio emission might also be blended with that of the nearby FIR source. Based on the \( \sim 2\" \) separation between the radio and optical/NIR/FIR emission, the ATCA source is not associated with the cluster galaxy. Only a CO line is detected by SoFiA in the ALMA cubes; closer inspection reveals positive emission in the [C I] line cube, but because it is not spatially coherent we only report an upper limit.

**J010255.67-491556.7** (Figure 2.11c) is a galaxy with a redder bulge-like component surrounded by bluer spiral arm features or a ring feature with two bright knots, with a nearby red spheroidal source. Neither the spiral nor the red neighbor have optical redshifts, but CO and [C I] lines are marginally detected in the ALMA data cubes consistent with a redshift of \( z = 0.8678 \). Only the [C I] emission is detected by SoFiA. The CO zeroth moment map peaks about 0\'.7 to the east of the spiral galaxy, and the [C I] zeroth moment map peaks between the spiral and the red galaxy (about 0\'.7 to the northwest of the spiral galaxy). No ALMA or radio continuum is detected, although there exists a faint 3 \( \sigma \) peak in the radio surface brightness.

**J054627.43-534433.6** (Figure 2.12a) is a fairly compact, star-forming galaxy at redshift \( z = 1.0566 \). Due to the source's position near the edge of our mosaic, line fluxes are subject to large primary-beam corrections. CO is unambiguously detected by SoFiA and [C I] is present at 3 \( \sigma \) significance. This is the brightest galaxy at IR wavelengths (and the only ULIRG) in our sample, and it also has the highest ISM mass. We find \( q_{IR} = 1.9 \pm 0.4 \), above the \( q_{IR} = 1.8 \) threshold that labels it a star-forming galaxy.

**J054638.87-534613.7** (Figure 2.12b) is a galaxy with a few near-UV-bright knots on its north and east sides. Both CO and [C I] lines are detected by SoFiA. The dust continuum and CO integrated flux appear to be marginally resolved. The CO (4 – 3) line is at \( z = 1.0810 \), and the [C I] line is offset \( \sim 400 \) km s\(^{-1}\) redward of the CO redshift. The difference in systemic velocity suggests that we may have observed a separate gas concentration, or (more likely) the [C I] detection
is unreliable. Removing the [C I] detection from our sample does not change our results on the [C I]-to-CO ratio (Section 2.5.1). We detect an ATCA source and compute $q_{IR} = 1.7 \pm 0.4$, implying the galaxy’s bolometric luminosity is dominated by AGN emission.

**J054644.15-534608.7** (Figure 2.12c) is compact and does not possess many distinguishing optical features aside from a lone star-forming clump on the east side of its nucleus. We have detected a CO line in the ALMA data cube consistent with J0546 membership ($z = 1.0649$) using SoFiA; this galaxy has the highest CO luminosity within our sample. The CO flux extends toward the west side of the galaxy, possibly indicating the presence of cold gas in a tidal feature or faint neighbor. [C I] emission is measured from a 3.6 $\sigma$ peak in the zeroth moment map. The dust emission is extended to the south. Because no radio continuum is detected, we find that $q_{IR} > 2.1$ and conclude that this galaxy’s luminosity is dominated by star formation activity.

### 2.7.2 ALMA dust-only detections

**J054633.40-534454.4** is a Ca II absorption-line galaxy with an elongated southern tail in *HST* imaging. We detect a faint ALMA source separated by a distance of 0\arcsec5 from the *HST* centroid. The dust detection, if associated with the cluster galaxy, is not at odds with our upper limits on CO and [C I] emission, which constrains the molecular mass to $\lesssim 3.0 \times 10^{10} M_\odot$, since molecular gas comprises half to two-thirds of a galaxy’s ISM mass (see, e.g., Scoville 2013). At *Herschel* wavelengths, we see unevenly-distributed positive emission totaling $S_{160 \mu m} \sim 1.5 \pm 0.4$ mJy, but no single pixel exceeds 2.3 $\sigma$ significance; therefore, we do not claim detection of a FIR source. No radio emission is detected ($S_{2.1 \text{ GHz}} < 24 \mu$Jy). We categorize the ALMA source as a dusty star-forming galaxy (DSFG) rather than a cluster member and rename it J054633.40-534454.5 in Table 2.8 to reflect its dust continuum centroid.

**J054636.61-534405.9** has an *HST* morphology suggesting that it has several clumpy regions of star formation, but it is labeled as a Ca II absorber in S13. No ALMA line emission is detected. There is a faint ALMA continuum source displaced to the southeast of the optical light by 1\arcsec1, which is significantly larger than the offsets normally seen in *HST*-ALMA comparisons (e.g., $\sim 0\arcsec3 - 0\arcsec6$; Dunlop et al. 2017). In Table 2.8, it is labeled as J054636.61-534405.9 along with its submillimeter properties.

**J054642.12-534543.9** is an [O II]-emitting galaxy that is compact and relatively featureless. A *Herschel*/PACS peak lies 1\arcsec6 to the northwest of the optical centroid, which itself lies 0\arcsec7 to
the northwest of a Spitzer source. The Spitzer NIR source agrees to sub-pixel accuracy with an ALMA continuum detection and a possible 3 σ radio source. No CO or [C I] emission is detected, implying either that a large fraction of the gas reservoir is atomic (M_{ISM} = (5.5±0.8) \times 10^{10} M_\odot vs. M_{mol}(CO) < 3.0 \times 10^{10} M_\odot at the cluster redshift), or that the dust emission is not hosted by the cluster galaxy. We assume the latter, because the PACS source is separated by 2\farcs3 in projection from the ALMA source, and attribute the ALMA emission to an optically-obscured background galaxy centered on the Spitzer source, and match the PACS source to the cluster galaxy. Bright SPIRE emission, unlikely to be from a cluster galaxy, is also detected at the level of S_{250 \mu m} = 23.8±4.0 mJy, which strengthens our case for a dusty background galaxy. We label the ALMA and SPIRE source an DSFG rather than a cluster member and rename it J054642.09-534544.3 in Table 2.8 to reflect its position centroid.

2.7.3 Additional Herschel detections

*J010243.11-491408.6* is a Ca II absorber with the appearance of a lenticular galaxy, and is coincident with faint Herschel/PACS emission. It is bright at SPIRE wavelengths (S_{250 \mu m} = 14.3 ± 3.5 mJy; S_{350 \mu m} = 13.3 ± 4.0 mJy) It lies outside our ALMA coverage, and no significant radio emission is found. Given the lack of [O II] emission and peak flux at SPIRE wavelengths, it is likely that a background dusty star-forming galaxy is responsible for the Herschel emission.

J010243.99-491744.4 lies outside our HST and ALMA field of view (FOV), is bright at Spitzer wavelengths (and a LIRG), and is identified as the lone confirmed AGN in our sample on the basis of optical spectroscopy (S13). Deep Chandra X-ray observations of J0102 (Hughes et al., in prep.) show an intrinsically absorbed with column density of N_H = 2.0 ± 0.5 \times 10^{23} H atoms cm^{-2} at the galaxy’s redshift of z = 0.87535) power-law spectrum with a neutral Fe fluorescence line with fitted source-frame energy of E_{Fe} = 6.37 ± 0.08 keV) and a 2 – 10 keV band unabsorbed luminosity of 1.2 \times 10^{44} erg s^{-1}, which confirms our identification of this source as an AGN.

J010247.68-491635.7 is an [O II] emission-line galaxy lying outside our HST and ALMA FOV. It is a LIRG and also an AGN due to its low q_{IR} = 1.3 ± 0.4.

*J010257.79-491519.0* is a Ca II absorber 1\farcs1 away from an optical point source (perhaps a quasar). We find that the PACS flux density is S_{160 \mu m} = 3.7 ± 0.2 mJy and identify it with faint (2 – 3 σ) double-lobed peaks in the radio surface brightness. The cluster member is separated from the PACS source (1\farcs8) and ATCA centroid (1\farcs4), as well as from the weak Herschel/SPIRE emission
(\sim 0.5 \text{ pix} = 3''\) and the NIR centroid (1''3). We conclude that the long-wavelength emission does not come from the cluster galaxy.

**J010302.99-491458.4** has a disturbed optical morphology suggesting that this \([\text{O II}]\)-emitting galaxy is undergoing a merger and a burst of star formation. It is also detected in the 250 \(\mu\text{m}\) SPIRE band (and has weak positive emission in the 350 \(\mu\text{m}\) band), but these appear to suffer from source confusion. No ALMA or ATCA emission is detected.

**J023540.10-512255.5** is a \(\text{Ca II}\) absorber detected by \textit{Herschel}/PACS. Based on its 160 \(\mu\text{m}\) flux density and radio non-detection, we conclude that its bolometric luminosity is dominated by star formation \((q_{\text{IR}} > 1.9)\).

**J023542.4-512101.0** is a \(\text{Ca II}\) absorber detected by \textit{Herschel}/PACS. No ATCA emission is found.

**J023547.60-512029.4** is a \(\text{Ca II}\) absorber with a neighboring cluster member 2''5 to its northwest. Its \textit{Herschel}/PACS 100/160 \(\mu\text{m}\) counterparts suffer from source blending, so we measure the PACS flux density from the brightest pixel in its surface brightness distribution. We do not attempt to deblend the \textit{Herschel} emission. No radio counterpart is found.

**J023549.20-511920.6** is a \(\text{Ca II}\) absorber located \sim 5'' from a thin, extended source that looks like it may be a lensed background galaxy. Its \textit{Herschel} emission is blended, so we measure its PACS flux density from its peak surface brightness. It is possible that the nearby source dominates the FIR emission, although we find a slight peak in the radial profile of the \textit{Herschel}/PACS emission coincident with the cluster member. No radio counterpart is found.

**J023557.20-511820.8** is a \(\text{Ca II}\) absorption-line galaxy with bright \textit{Herschel}/PACS emission. There is a 3 \(\sigma\) peak in the radio surface brightness, but \texttt{IMFIT} does not converge on a fit, so we report no ATCA detection.

**J043810.40-542008.1** is an \([\text{O II}]\) emitter. It is one of the brightest \textit{Herschel}/PACS sources in the map of J0438 and the only IRBG in either of our low-z clusters. We also find bright SPIRE emission \(S_{250 \mu\text{m}} = 25.6 \pm 1.6 \text{ mJy}\) (and a debatable \(S_{350 \mu\text{m}} \sim 8.4 \pm 5.0 \text{ mJy}\)). We also find a radio counterpart and measure \(q_{\text{IR}} = 1.6 \pm 0.3\).
**J043824.40-541716.0** is a Ca II absorber. The *Herschel* emission is blended, but we are able to measure the flux density by using the matched-filter extractor at 100 $\mu$m and by measuring a 4 $\sigma$ peak in the lower-resolution 160 $\mu$m map. It also has a slightly fainter neighbor to its south that is not a confirmed cluster member at a projected angular distance of 2$''$.4. A radio source is also detected, but it is closer on the sky to the neighboring galaxy.

**J054635.39-534541.2** is an [O II] emitter that appears to be an inclined star-forming disk galaxy with a partially stripped tail to the west. The bulk of the *Herschel* emission is centered slightly to the west and offset by about 2$''$. We see a faint ($\sim 3\sigma$) radio peak centered $\sim 1''$5 to the west of the optical centroid, but are unable to reliably measure its flux density.

### 2.8 Appendix B: Dusty star-forming galaxies

In Table 2.8, we list all new ALMA Band 6 continuum sources detected at 4 $\sigma$ significance, and their millimeter (observed frame) flux densities, as well as a brief description of the objects. ALMA continuum counterparts to cluster galaxies, or sources separated by small projected distances to cluster galaxies, are also included in the table. CASA IMFIT is used to measure flux densities and uncertainties unless otherwise stated. Because their redshifts are not known, we cannot determine their dust masses or IR properties. In Figure 2.13, we show some examples of these DSFGs along with *Herschel* contours and *HST* imaging. Aguirre et al., in preparation will discuss the DSFG content of these and other cluster fields in greater detail.

**Acknowledgments**

We thank the anonymous referee for detailed and useful comments that have greatly improved this paper. J.F.W. and A.J.B. thank Amanda Kepley for valuable support and guidance during the ALMA data reduction process. P.A. acknowledges support from CONICYT through grant FONDECYT Iniciación 11130590. J.P.H. acknowledges the hospitality of the Flatiron Institute which is supported by the Simons Foundation. This paper makes use of the following ALMA data: ADS/JAO.ALMA#2013.1.01358.S. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada), MOST and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. This work has been supported by (i) an award issued by JPL/Caltech in
association with *Herschel*, which is a European Space Agency Cornerstone Mission with significant participation by NASA, (ii) the National Science Foundation through grant AST-0955810 and award GSSP SOSPA2-018 from the National Radio Astronomy Observatory, which is operated under cooperative agreement by Associated Universities, Inc., and (iii) the National Aeronautics and Space Administration (NASA) through Chandra Award numbered GO2-13156X issued to Rutgers University by the Chandra X-ray Observatory Center, which is operated by the Smithsonian Astrophysical Observatory for and on behalf of NASA under contract NAS8-03060.
### Table 2.1. Our sample of clusters

<table>
<thead>
<tr>
<th>Cluster name</th>
<th>$z^a$</th>
<th>$N_{\text{gal}}^a$</th>
<th>$M_{200c}^b$ [10$^{14}$ $M_\odot$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACT-CL J0235-5121</td>
<td>0.2777</td>
<td>82</td>
<td>8.0 ± 2.9</td>
</tr>
<tr>
<td>ACT-CL J0438-5419</td>
<td>0.4214</td>
<td>65</td>
<td>12.9 ± 3.2</td>
</tr>
<tr>
<td>ACT-CL J0102-4915</td>
<td>0.8701</td>
<td>89</td>
<td>11.3 ± 2.9</td>
</tr>
<tr>
<td>ACT-CL J0546-5345</td>
<td>1.0663</td>
<td>49</td>
<td>5.5 ± 2.3</td>
</tr>
</tbody>
</table>

$^a$Spectroscopic redshifts and numbers of spectroscopically confirmed members ($N_{\text{gal}}$) from Sifón et al. (2013).

$^b$Dynamical mass estimates ($M_{200c}$) from Sifón et al. (2016).
<table>
<thead>
<tr>
<th>Cluster (Phase center)</th>
<th>Observing frequency [GHz]</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Continuum</td>
<td>CO (4 – 3)</td>
</tr>
<tr>
<td>ACT-CL J0102-4915</td>
<td>254.6</td>
<td>244.7 – 248.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACT-CL J0546-5345</td>
<td>230.7</td>
<td>221.3 – 224.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2.3. Measurements of detected cluster members

<table>
<thead>
<tr>
<th>Object</th>
<th>RA</th>
<th>Dec</th>
<th>$z$</th>
<th>Spectral features&lt;sup&gt;a&lt;/sup&gt;</th>
<th>$S_{100}$ μm [mJy]</th>
<th>$S_{160}$ μm [mJy]</th>
<th>$S_{CO (4-3)\Delta v}$ [Jy km s&lt;sup&gt;-1&lt;/sup&gt;]</th>
<th>$S_{[CII]^{3}P_{1}^{-3}P_{0}\Delta v}$ [Jy km s&lt;sup&gt;-1&lt;/sup&gt;]</th>
<th>$S_{ALMA}$ [mJy]</th>
<th>$S_{2.1 \text{ GHz}}$ [μJy]</th>
</tr>
</thead>
<tbody>
<tr>
<td>J010243.99-491744.4</td>
<td>15.68328</td>
<td>-49.29565</td>
<td>0.87535</td>
<td>AGN</td>
<td>2.58 ± 0.41</td>
<td>3.77 ± 0.45</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>256 ± 13</td>
</tr>
<tr>
<td>J010247.68-491635.7</td>
<td>15.69867</td>
<td>-49.27659</td>
<td>0.88977</td>
<td>[O II], Ca II</td>
<td>1.22 ± 0.19</td>
<td>1.54 ± 0.26</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>119 ± 21</td>
</tr>
<tr>
<td>J010252.44-491531.2</td>
<td>15.71849</td>
<td>-49.25867</td>
<td>0.86100</td>
<td>…</td>
<td>0.87 ± 0.16</td>
<td>1.80 ± 0.24</td>
<td>0.97 ± 0.26</td>
<td>…</td>
<td>&lt; 0.36</td>
<td>&lt; 27</td>
</tr>
<tr>
<td>J010255.50-491416.2</td>
<td>15.73126</td>
<td>-49.23782</td>
<td>0.87117</td>
<td>[O II], Ca II</td>
<td>&lt; 0.78</td>
<td>1.38 ± 0.34&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.16 ± 0.31</td>
<td>&lt; 0.81</td>
<td>&lt; 0.36</td>
<td>&lt; 24</td>
</tr>
<tr>
<td>J010255.67-491556.7</td>
<td>15.73198</td>
<td>-49.26574</td>
<td>0.86780</td>
<td>…</td>
<td>0.74 ± 0.16</td>
<td>1.17 ± 0.23</td>
<td>0.95 ± 0.25</td>
<td>0.58 ± 0.11&lt;sup&gt;c&lt;/sup&gt;</td>
<td>&lt; 0.33</td>
<td>&lt; 24</td>
</tr>
<tr>
<td>J010302.99-491458.4</td>
<td>15.76246</td>
<td>-49.24956</td>
<td>0.87433</td>
<td>[O II], Ca II</td>
<td>0.64 ± 0.18</td>
<td>1.37 ± 0.27</td>
<td>&lt; 0.72</td>
<td>&lt; 0.78</td>
<td>&lt; 0.39</td>
<td>&lt; 30</td>
</tr>
<tr>
<td>J025540.10-512255.5</td>
<td>38.91708</td>
<td>-51.38208</td>
<td>0.27350</td>
<td>Ca II</td>
<td>1.04 ± 0.21</td>
<td>2.15 ± 0.50&lt;sup&gt;b&lt;/sup&gt;</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>&lt; 33</td>
</tr>
<tr>
<td>J025542.40-512101.0</td>
<td>38.92667</td>
<td>-51.35028</td>
<td>0.27277</td>
<td>Ca II</td>
<td>0.80 ± 0.18</td>
<td>1.82 ± 0.34&lt;sup&gt;b&lt;/sup&gt;</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>&lt; 36</td>
</tr>
<tr>
<td>J025547.60-512029.4</td>
<td>38.94833</td>
<td>-51.34150</td>
<td>0.28608</td>
<td>Ca II</td>
<td>1.04 ± 0.18</td>
<td>1.46 ± 0.29&lt;sup&gt;b&lt;/sup&gt;</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>&lt; 33</td>
</tr>
<tr>
<td>J025549.20-511920.6</td>
<td>38.95500</td>
<td>-51.32329</td>
<td>0.27379</td>
<td>Ca II</td>
<td>0.95 ± 0.20</td>
<td>1.89 ± 0.41&lt;sup&gt;b&lt;/sup&gt;</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>&lt; 33</td>
</tr>
<tr>
<td>J025557.20-511820.8</td>
<td>38.98333</td>
<td>-51.30578</td>
<td>0.28160</td>
<td>Ca II</td>
<td>2.21 ± 0.50&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2.89 ± 0.71</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>&lt; 36</td>
</tr>
<tr>
<td>J043810.40-542068.1</td>
<td>69.54333</td>
<td>-54.33558</td>
<td>0.41830</td>
<td>[O II], Ca II</td>
<td>4.66 ± 0.22</td>
<td>6.76 ± 0.52</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>130 ± 27</td>
</tr>
<tr>
<td>J043824.40-541716.0</td>
<td>69.60167</td>
<td>-54.28778</td>
<td>0.41748</td>
<td>Ca II</td>
<td>1.24 ± 0.26</td>
<td>2.00 ± 0.49&lt;sup&gt;b&lt;/sup&gt;</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>&lt; 39</td>
</tr>
<tr>
<td>J054627.43-534433.6</td>
<td>86.61429</td>
<td>-53.74267</td>
<td>1.05660</td>
<td>…</td>
<td>4.24 ± 0.17</td>
<td>7.03 ± 0.30</td>
<td>1.37 ± 0.35&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.72 ± 0.24&lt;sup&gt;d&lt;/sup&gt;</td>
<td>1.21 ± 0.18</td>
<td>101 ± 13</td>
</tr>
<tr>
<td>J054635.39-534541.1</td>
<td>86.64745</td>
<td>-53.76143</td>
<td>1.07116</td>
<td>[O II], Ca II</td>
<td>0.76 ± 0.16</td>
<td>1.10 ± 0.22</td>
<td>&lt; 0.42</td>
<td>&lt; 0.42</td>
<td>&lt; 0.30</td>
<td>&lt; 24</td>
</tr>
<tr>
<td>J054636.61-534405.9</td>
<td>86.65252</td>
<td>-53.73499</td>
<td>1.05774</td>
<td>Ca II</td>
<td>0.90 ± 0.16</td>
<td>1.47 ± 0.28</td>
<td>&lt; 0.45</td>
<td>&lt; 0.63</td>
<td>&lt; 0.30</td>
<td>&lt; 27</td>
</tr>
<tr>
<td>J054638.87-534613.6</td>
<td>86.66196</td>
<td>-53.77045</td>
<td>1.08130</td>
<td>…</td>
<td>1.13 ± 0.17</td>
<td>1.70 ± 0.27</td>
<td>0.84 ± 0.17&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.49 ± 0.13&lt;sup&gt;d&lt;/sup&gt;</td>
<td>1.08 ± 0.15</td>
<td>51 ± 11</td>
</tr>
<tr>
<td>J054642.12-534543.9</td>
<td>86.67550</td>
<td>-53.76220</td>
<td>1.05240</td>
<td>[O II], Ca II</td>
<td>&lt; 0.69</td>
<td>2.15 ± 0.27</td>
<td>&lt; 0.39</td>
<td>&lt; 0.48</td>
<td>&lt; 0.42</td>
<td>&lt; 30</td>
</tr>
<tr>
<td>J054644.15-534608.7</td>
<td>86.68937</td>
<td>-53.76098</td>
<td>1.06590</td>
<td>…</td>
<td>1.47 ± 0.17</td>
<td>2.18 ± 0.28</td>
<td>1.47 ± 0.24&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.67 ± 0.14</td>
<td>0.81 ± 0.11</td>
<td>&lt; 24</td>
</tr>
</tbody>
</table>

Note. — Thresholds for detection are 3 $\sigma$ for ALMA line sources and 4 $\sigma$ for Herschel, ALMA, and ATCA continuum sources. For all non-detections, we report 3 $\sigma$ upper limits.

<sup>a</sup> Optical spectral features reported by S13.

<sup>b</sup> Measured from peak brightness because no detection was found by our matched-filter source extractor.

<sup>c</sup> Identified using SoFiA (Serra et al. 2015).

<sup>d</sup> Measured from peak brightness because D3M did not converge on a solution.
Table 2.4. Derived properties of detected cluster members

<table>
<thead>
<tr>
<th>Object</th>
<th>r</th>
<th>K15 IR template</th>
<th>log (L_{\text{IR}}) ([L_\odot])</th>
<th>log (L_{\text{H}2}^\text{P}) ([L_\odot])</th>
<th>log (M_\ast) ([M_\odot])</th>
<th>sSFR ([\text{yr}^{-1}])</th>
<th>log (M_{\text{mol}}) ([M_\odot])</th>
<th>(f_{\text{esc}})</th>
<th>log (\tau_{\text{dep}}) ([\text{yr}])</th>
<th>log (M_{\text{ISM}}) ([M_\odot])</th>
<th>log (L_{\text{1.4 GHz}}) ([\text{W} \text{Hz}^{-1}])</th>
<th>qH</th>
</tr>
</thead>
<tbody>
<tr>
<td>J010243.99-491744.4</td>
<td>0.76</td>
<td>AGN2</td>
<td>11.64 ± 0.13</td>
<td>11.42 ± 0.07</td>
<td>12.17 ± 0.09</td>
<td>−10.53 ± 0.11</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>24.34 ± 0.02</td>
<td>1.3 ± 0.3</td>
</tr>
<tr>
<td>J010247.68-491635.7</td>
<td>0.38</td>
<td>AGN1</td>
<td>11.32 ± 0.17</td>
<td>10.96 ± 0.12</td>
<td>11.03 ± 0.09</td>
<td>−9.85 ± 0.14</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>24.03 ± 0.08</td>
<td>1.3 ± 0.4</td>
</tr>
<tr>
<td>J010252.44-491531.2</td>
<td>0.08</td>
<td>COMPOSITE2</td>
<td>11.02 ± 0.19</td>
<td>10.92 ± 0.06</td>
<td>10.85 ± 0.09</td>
<td>−9.71 ± 0.10</td>
<td>10.4 ± 0.3</td>
<td>0.36</td>
<td>9.3</td>
<td>&lt; 10.3</td>
<td>&lt; 23.35</td>
<td>&gt; 1.7</td>
</tr>
<tr>
<td>J010255.50-491416.2</td>
<td>0.45</td>
<td>COMPOSITE1</td>
<td>10.85 ± 0.15</td>
<td>10.82 ± 0.14</td>
<td>11.46 ± 0.09</td>
<td>−10.42 ± 0.16</td>
<td>10.5 ± 0.3</td>
<td>0.11</td>
<td>9.5</td>
<td>&lt; 10.3</td>
<td>&lt; 23.31</td>
<td>&gt; 1.6</td>
</tr>
<tr>
<td>J010255.67-491556.7</td>
<td>0.13</td>
<td>COMPOSITE3</td>
<td>10.91 ± 0.22</td>
<td>10.72 ± 0.10</td>
<td>11.27 ± 0.09</td>
<td>−10.34 ± 0.13</td>
<td>10.4 ± 0.3</td>
<td>0.14</td>
<td>9.5</td>
<td>&lt; 10.3</td>
<td>&lt; 23.30</td>
<td>&gt; 1.6</td>
</tr>
<tr>
<td>J010302.99-491458.4</td>
<td>0.52</td>
<td>COMPOSITE2</td>
<td>10.92 ± 0.19</td>
<td>10.82 ± 0.09</td>
<td>11.02 ± 0.09</td>
<td>−9.98 ± 0.12</td>
<td>&lt; 10.3</td>
<td>&lt; 0.19</td>
<td>&lt; 9.3</td>
<td>&lt; 10.4</td>
<td>&lt; 23.41</td>
<td>&gt; 1.5</td>
</tr>
<tr>
<td>J023540.10-512255.5</td>
<td>0.29</td>
<td>AGN1</td>
<td>10.00 ± 0.18</td>
<td>9.64 ± 0.16</td>
<td>11.49 ± 0.03</td>
<td>−11.63 ± 0.16</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>&lt; 22.11</td>
<td>&gt; 1.9</td>
</tr>
<tr>
<td>J023542.40-512101.0</td>
<td>0.07</td>
<td>AGN1</td>
<td>9.92 ± 0.18</td>
<td>9.56 ± 0.15</td>
<td>11.43 ± 0.03</td>
<td>−11.65 ± 0.15</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>&lt; 22.15</td>
<td>&gt; 1.8</td>
</tr>
<tr>
<td>J023547.60-512029.4</td>
<td>0.10</td>
<td>COMPOSITE1</td>
<td>9.76 ± 0.10</td>
<td>9.72 ± 0.06</td>
<td>10.87 ± 0.03</td>
<td>−10.94 ± 0.07</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>&lt; 22.16</td>
<td>&gt; 1.6</td>
</tr>
<tr>
<td>J023549.20-511920.6</td>
<td>0.27</td>
<td>AGN1</td>
<td>9.96 ± 0.18</td>
<td>9.61 ± 0.16</td>
<td>9.77 ± 0.03</td>
<td>−9.94 ± 0.16</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>&lt; 22.11</td>
<td>&gt; 1.9</td>
</tr>
<tr>
<td>J023557.20-511820.8</td>
<td>0.48</td>
<td>SFG1</td>
<td>10.00 ± 0.15</td>
<td>9.98 ± 0.08</td>
<td>11.15 ± 0.03</td>
<td>−10.95 ± 0.08</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>&lt; 22.18</td>
<td>&gt; 1.8</td>
</tr>
<tr>
<td>J043810.40-542008.1</td>
<td>0.23</td>
<td>SFG1</td>
<td>10.78 ± 0.12</td>
<td>10.76 ± 0.02</td>
<td>11.50 ± 0.04</td>
<td>−10.53 ± 0.04</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>23.17 ± 0.09</td>
<td>1.6 ± 0.3</td>
</tr>
<tr>
<td>J043824.40-541716.0</td>
<td>0.40</td>
<td>AGN1</td>
<td>10.51 ± 0.18</td>
<td>10.15 ± 0.16</td>
<td>11.01 ± 0.04</td>
<td>−10.64 ± 0.17</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>&lt; 22.65</td>
<td>&gt; 1.9</td>
</tr>
<tr>
<td>J054627.43-534433.6</td>
<td>0.78</td>
<td>AGN1</td>
<td>12.11 ± 0.16</td>
<td>11.75 ± 0.03</td>
<td>11.30 ± 0.09</td>
<td>−9.33 ± 0.09</td>
<td>10.7 ± 0.3</td>
<td>0.27</td>
<td>8.8</td>
<td>11.0 ± 0.3</td>
<td>24.17 ± 0.06</td>
<td>1.9 ± 0.4</td>
</tr>
<tr>
<td>J054635.39-534541.1</td>
<td>0.16</td>
<td>AGN3</td>
<td>11.27 ± 0.25</td>
<td>10.85 ± 0.18</td>
<td>11.13 ± 0.09</td>
<td>−10.07 ± 0.20</td>
<td>&lt; 10.2</td>
<td>&lt; 0.13</td>
<td>&lt; 9.2</td>
<td>&lt; 10.4</td>
<td>&lt; 23.57</td>
<td>&gt; 1.7</td>
</tr>
<tr>
<td>J054636.61-534405.9</td>
<td>0.62</td>
<td>COMPOSITE3</td>
<td>11.26 ± 0.22</td>
<td>11.06 ± 0.09</td>
<td>11.26 ± 0.09</td>
<td>−9.98 ± 0.13</td>
<td>&lt; 10.3</td>
<td>&lt; 0.10</td>
<td>&lt; 9.0</td>
<td>&lt; 10.4</td>
<td>&lt; 23.60</td>
<td>&gt; 1.7</td>
</tr>
<tr>
<td>J054638.87-534613.6</td>
<td>0.31</td>
<td>AGN1</td>
<td>11.58 ± 0.17</td>
<td>11.22 ± 0.10</td>
<td>11.45 ± 0.09</td>
<td>−10.01 ± 0.14</td>
<td>10.5 ± 0.3</td>
<td>0.12</td>
<td>9.1</td>
<td>11.0 ± 0.3</td>
<td>23.91 ± 0.10</td>
<td>1.7 ± 0.4</td>
</tr>
<tr>
<td>J054642.12-534543.9</td>
<td>0.30</td>
<td>COMPOSITE1</td>
<td>11.33 ± 0.10</td>
<td>11.30 ± 0.07</td>
<td>10.74 ± 0.09</td>
<td>−9.23 ± 0.11</td>
<td>&lt; 10.2</td>
<td>&lt; 0.28</td>
<td>&lt; 8.7</td>
<td>&lt; 10.6</td>
<td>&lt; 23.64</td>
<td>&gt; 1.7</td>
</tr>
<tr>
<td>J054644.15-534608.7</td>
<td>0.49</td>
<td>AGN1</td>
<td>11.67 ± 0.17</td>
<td>11.31 ± 0.08</td>
<td>11.24 ± 0.09</td>
<td>−9.72 ± 0.12</td>
<td>10.8 ± 0.3</td>
<td>0.34</td>
<td>9.2</td>
<td>10.9 ± 0.3</td>
<td>&lt; 23.56</td>
<td>&gt; 2.1</td>
</tr>
</tbody>
</table>

Note. — For all non-detections, we report 3 \(\sigma\) upper limits.
Table 2.5. Numbers of SoFiA detections

<table>
<thead>
<tr>
<th>Cluster + line</th>
<th>cube (true)</th>
<th>$-1 \times$ cube (false)</th>
</tr>
</thead>
<tbody>
<tr>
<td>J0102 [C I]</td>
<td>36</td>
<td>35</td>
</tr>
<tr>
<td>J0102 CO</td>
<td>17</td>
<td>11</td>
</tr>
<tr>
<td>J0546 [C I]</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>J0546 CO</td>
<td>8</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 2.6. Numbers of PACS detections

<table>
<thead>
<tr>
<th>Cluster</th>
<th>100 $\mu$m map (true)</th>
<th>$-1 \times$ 100 $\mu$m map (false)</th>
<th>160 $\mu$m map (true)</th>
<th>$-1 \times$ 160 $\mu$m map (false)</th>
</tr>
</thead>
<tbody>
<tr>
<td>J0102</td>
<td>62</td>
<td>0</td>
<td>46</td>
<td>0</td>
</tr>
<tr>
<td>J0235</td>
<td>88</td>
<td>0</td>
<td>44</td>
<td>0</td>
</tr>
<tr>
<td>J0438</td>
<td>49</td>
<td>1</td>
<td>24</td>
<td>0</td>
</tr>
<tr>
<td>J0546</td>
<td>65</td>
<td>0</td>
<td>54</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2.7. PACS stacking 3 $\sigma$ upper limits

<table>
<thead>
<tr>
<th>Cluster</th>
<th>$S_{100 \mu m}$ [mJy]</th>
<th>$\log L_{100 \mu m}^{IR}$</th>
<th>$S_{160 \mu m}$ [mJy]</th>
<th>$\log L_{160 \mu m}^{IR}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>J0235</td>
<td>&lt; 0.08</td>
<td>&lt; 8.6</td>
<td>&lt; 0.18</td>
<td>&lt; 8.8</td>
</tr>
<tr>
<td>J0438</td>
<td>&lt; 0.10</td>
<td>&lt; 9.1</td>
<td>&lt; 0.15</td>
<td>&lt; 9.1</td>
</tr>
<tr>
<td>J0102</td>
<td>&lt; 0.09</td>
<td>&lt; 10.0</td>
<td>&lt; 0.12</td>
<td>&lt; 9.9</td>
</tr>
<tr>
<td>J0546</td>
<td>&lt; 0.12</td>
<td>&lt; 10.5</td>
<td>&lt; 0.16</td>
<td>&lt; 10.2</td>
</tr>
</tbody>
</table>
Table 2.8. ALMA continuum detections

<table>
<thead>
<tr>
<th>Object</th>
<th>RA [°]</th>
<th>Dec [°]</th>
<th>$S_{230 \text{ GHz}}$ [mJy]</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>J010249.26-491438.1</td>
<td>15.705244</td>
<td>-49.243920</td>
<td>1.07 ± 0.12</td>
<td>DSFG</td>
</tr>
<tr>
<td>J010249.28-491506.8</td>
<td>15.705319</td>
<td>-49.251876</td>
<td>3.97 ± 0.39</td>
<td>Strongly lensed DSFG; measured manually</td>
</tr>
<tr>
<td>J010250.46-491541.6</td>
<td>15.710257</td>
<td>-49.261566</td>
<td>0.67 ± 0.19</td>
<td>DSFG</td>
</tr>
<tr>
<td>J010250.78-491409.3</td>
<td>15.711571</td>
<td>-49.235906</td>
<td>1.61 ± 0.19</td>
<td>DSFG</td>
</tr>
<tr>
<td>J010251.06-491538.8</td>
<td>15.712751</td>
<td>-49.260790</td>
<td>3.48 ± 0.18</td>
<td>Bright DSFG</td>
</tr>
<tr>
<td>J010254.58-491519.9</td>
<td>15.727436</td>
<td>-49.255337</td>
<td>0.80 ± 0.19</td>
<td>DSFG</td>
</tr>
<tr>
<td>J010254.89-491514.6</td>
<td>15.728714</td>
<td>-49.249054</td>
<td>4.00 ± 0.20</td>
<td>Bright DSFG; star nearby; La Flaca$^a$</td>
</tr>
<tr>
<td>J010255.66-491509.0</td>
<td>15.731921</td>
<td>-49.252502</td>
<td>9.50 ± 0.27</td>
<td>Bright DSFG; La Flaca$^a$</td>
</tr>
<tr>
<td>J010258.13-491456.2</td>
<td>15.742222</td>
<td>-49.248937</td>
<td>0.82 ± 0.18</td>
<td>DSFG</td>
</tr>
<tr>
<td>J054627.42-534433.6</td>
<td>86.614257</td>
<td>-53.742676</td>
<td>1.21 ± 0.18</td>
<td>Cluster galaxy (CO)</td>
</tr>
<tr>
<td>J054629.26-534456.9</td>
<td>86.621929</td>
<td>-53.749146</td>
<td>1.17 ± 0.19</td>
<td>DSFG</td>
</tr>
<tr>
<td>J054633.40-534454.5</td>
<td>86.639162</td>
<td>-53.748474</td>
<td>0.38 ± 0.09</td>
<td>DSFG; cluster galaxy (Ca II) nearby$^b$</td>
</tr>
<tr>
<td>J054633.36-534548.6</td>
<td>86.638985</td>
<td>-53.763494</td>
<td>0.51 ± 0.16</td>
<td>Strongly lensed DSFG</td>
</tr>
<tr>
<td>J054634.57-534552.3</td>
<td>86.644030</td>
<td>-53.764525</td>
<td>2.84 ± 0.17</td>
<td>Bright DSFG</td>
</tr>
<tr>
<td>J054636.65-534406.5</td>
<td>86.652722</td>
<td>-53.735132</td>
<td>0.34 ± 0.08</td>
<td>Cluster galaxy (Ca II) nearby$^b$</td>
</tr>
<tr>
<td>J054638.46-534553.6</td>
<td>86.660242</td>
<td>-53.768475</td>
<td>0.89 ± 0.13</td>
<td>DSFG</td>
</tr>
<tr>
<td>J054638.87-534613.8</td>
<td>86.661961</td>
<td>-53.770508</td>
<td>1.08 ± 0.15</td>
<td>Cluster galaxy (CO)</td>
</tr>
<tr>
<td>J054639.19-534519.2</td>
<td>86.663296</td>
<td>-53.755322</td>
<td>0.95 ± 0.16</td>
<td>DSFG; near foreground spiral</td>
</tr>
<tr>
<td>J054639.69-534602.6</td>
<td>86.665388</td>
<td>-53.767383</td>
<td>1.02 ± 0.12</td>
<td>DSFG</td>
</tr>
<tr>
<td>J054640.59-534600.5</td>
<td>86.669128</td>
<td>-53.766816</td>
<td>0.64 ± 0.05</td>
<td>DSFG</td>
</tr>
<tr>
<td>J054641.59-534628.5</td>
<td>86.673299</td>
<td>-53.774584</td>
<td>0.89 ± 0.11</td>
<td>No counterpart at optical or NIR wavelengths</td>
</tr>
<tr>
<td>J054642.09-534544.3</td>
<td>86.675374</td>
<td>-53.762302</td>
<td>0.59 ± 0.09</td>
<td>DSFG; cluster galaxy ([O II]) nearby$^b$</td>
</tr>
<tr>
<td>J054644.14-534608.9</td>
<td>86.683928</td>
<td>-53.769137</td>
<td>0.81 ± 0.11</td>
<td>Cluster galaxy (CO)</td>
</tr>
</tbody>
</table>

$^a$La Flaca is the nickname of two blended millimeter sources in the J0102 field of view detected by APEX/LABOCA (see Section 4.1.1 of Lindner et al. 2015).

$^b$See Section 2.7.2
Figure 2.1 ALMA Band 6 continuum mosaic imaging of J0102 shown in grayscale. Confirmed cluster members are marked by circles: black circles are galaxies found via optical spectroscopy (S13); smaller green filled circles are galaxies with Herschel/PACS counterparts; larger red (blue) circles are galaxies with ALMA CO \((4 - 3)\) ([C I] \((^3P_1 - ^3P_0)\)) line detections (labeled by name). A purple cross marks the brightest cluster galaxy (also labeled). The large dashed circle encloses \(0.5 \times R_{200c}\) and is centered on the midpoint between the two peaks in the mass distribution (determined by weak lensing; Jee et al. 2014).
Figure 2.2 Chary & Elbaz (2001, CE01) and three example Kirkpatrick et al. (2015, K15) template SEDs with $L_{IR} = 10^{11} \ L_\odot$, redshifted to $z = 0.87$. We show the sensitivities (3 $\sigma$ upper limits) of our FIR/millimeter observations (point-source subtracted maps) in colored horizontal lines.
Figure 2.3 Histograms of bootstrapped stacked *Herschel/PACS* 100 µm and 160 µm flux densities for 5000 iterations. Outlying fluxes in excess of 4 MADs have been removed. Flux densities are taken from the “residual” point source-subtracted PACS maps. The black dashed vertical line signifies zero flux, and the solid black line indicates the mean flux of blank-sky pixels. We also report the number of galaxies per stack and *p*-values determined by computing the percentile of the blank-sky mean within the bootstrapped distribution of stacked flux densities.
Figure 2.4 (Left) Bootstrapped histograms of stacked ALMA dust continuum flux densities for galaxies in our two high-$z$ clusters. (Right) Continuum fluxes converted to $M_{\text{ISM}}$ values and stacked in bins of clustercentric radius. 95% confidence intervals (computed by bootstrap resampling for 5000 iterations) are shown for each cluster. A dashed line marks zero. In both cases, 4-MAD outliers were removed.
Figure 2.5 [C I] (blue) and CO (red) stacked spectra from the J0102 and J0546 ALMA data cubes. Grey regions show 3 $\sigma$ limits (centered around zero). All outlying flux densities in excess of 4 MADs were removed from calculation of stacked fluxes and uncertainties.
Figure 2.6 The ratio of [C I] and CO in integrated flux units plotted vs. IR luminosity (1 σ error bars shown). The fiducial literature-based flux ratio of 0.71 adopted in this paper is shown as a dashed line.
Figure 2.7 ISM mass (vertical) vs. molecular mass (horizontal) vs. IR luminosity (color). Error bars denote 1 σ, and leftward-pointing arrows indicate 3 σ upper limits on $M_{\text{ISM}}$. 
Figure 2.8 FIR-radio correlation $q_{\text{IR}}$ parameter is plotted for 19 Herschel detections as a function of IR luminosity. The horizontal dashed line and shaded region show the median $q_{\text{IR}}$ value and $\pm 1\sigma$ uncertainties for 250 $\mu$m selected star-forming galaxies (Ivison et al. 2010), and the horizontal dotted line shows $q_{\text{IR}} = 1.8$ used to separate star-forming galaxies from AGNs. Error bars are $\pm 1\sigma$; note that the error in $q_{\text{IR}}$ is correlated with errors in $L_{\text{IR}}$ and $L_{1.4\,$GHz}. Empty markers with upward-pointing arrows are $3\sigma$ lower limits on $q_{\text{IR}}$ due to radio non-detections. Marker colors and symbols correspond to particular galaxy clusters and radio/optical spectral classifications, respectively. AGN-dominated systems as determined from best-fitting K15 IR spectral templates are marked with black stars.
Figure 2.9 IR luminosity due to star formation is plotted against clustercentric radius for 19 Herschel detections. Marker color and shape indicate cluster membership. Clustercentric radius is computed as the projected distance from the BCG for J0235, J0438, and J0546, and from the midpoint between weak lensing peaks for J0102 (Jee et al. 2014). Unambiguous radio AGNs ($q_{\text{IR}} < 1.8$) are shown with open markers. On the y-axis, a box plot shows the interquartile range in $L_{\text{IR}}^\text{SF}$. 
Figure 2.10 Specific SFR (sSFR) is plotted against clustercentric radius for 19 Herschel detections. Colors and markers are the same as in Figure 2.9, where radio AGNs are shown with open markers. The dashed (solid) horizontal black line illustrates the measured SFR–$M_*$ correlation for $\log(M_*/M_\odot) = 11.0$ at $z \sim 1$ ($z \sim 0$) as reported by Elbaz et al. (2007). On the $y$-axis, a box plot shows the interquartile range in sSFR.
Figure 2.11 Herschel and ALMA contours overlaid on HST/ACS F625W + F775W + F850LP color images in 5′ cutouts, for J0102 detections (J010252.44-491531.2 (a); J010255.50-491416.2 (b); J010255.67-491556.7 (c)). Left panels: Herschel/PACS 160 μm (green; multiples of 1.5 σ) and ALMA dust continuum (purple; multiples of 1.5 σ) contours, with ALMA synthesized beam at lower left. Middle panels: CO (4 – 3) integrated intensity (red; multiples of 1.5 σ) contours, with synthesized beams at lower left and inset spectra showing ±1 σ (gray) and velocity range used for moment map (yellow). Right panel: same as middle panels for the [C I] (3P1 – 3P0) line. Red and blue crosses in left panels mark peaks of line emission in middle and right panels, respectively.
Figure 2.12 Spectral line sources in J0546 (J054627.43-534433.6 (a); J054638.87-534613.6 (b); J054644.15-534608.7 (c)). Notation as in Figure 2.11. Grayscale images are HST F814W; 160 µm contours for top panel (a) are multiples of 5 σ.
Figure 2.13 16” cutouts of Herschel and ALMA contours overlaid on HST images. (a) ALMA continuum contours (purple, ±2 σ levels) and PACS 160 µm contours (green; ±2 σ levels) overlaid on HST/ACS F625W + F775W + F850LP color imaging of a possible strongly lensed system, J010249.28-491506.8. Possible arcs of millimeter continuum emission are seen to the northeast and southwest, the latter of which is matched to a blue arc in HST imaging. The early-type galaxies at the center of the image and toward the southeast are confirmed cluster galaxies. (b) Two bright DSFGs, J010254.89-491514.6 and J010255.66-491509.0, which were detected as a blended source (and named “La Flaca”; Lindner et al. 2015), are shown with ALMA continuum contours (purple, ±8 σ levels) and PACS 160 µm contours (green; ±2 σ levels) overlaid on HST false-color imaging. No PACS emission is detected, and whereas no short-wavelength counterpart is detected for J010254.89-491514.6 (southwest), a Spitzer counterpart is found for J010255.66-491509.0 (northeast). (c) A strongly lensed galaxy, J054633.36-534548.6, is shown with ALMA continuum contours (purple, ±2 σ levels) and PACS 160 µm contours (green; ±2 σ levels) overlaid on HST/F814W grayscale imaging. No PACS emission is detected with the DSFG. (d) A bright submillimeter and FIR source, J054634.57-534552.3, is shown with ALMA continuum contours (purple, ±4 σ levels) and PACS 160 µm contours (green; ±2 σ levels) overlaid on HST/F814W grayscale imaging.
Chapter 3
The star-forming interstellar medium of Lyman break galaxy analogs


3.1 Introduction

Lyman break galaxies (LBGs; Steidel et al. 1996) are among the brightest and best-studied systems at high \((z > 3)\) redshift. Efficient selection of LBGs has unveiled a wealth of information about the cosmic star formation rate (SFR) history (Madau & Dickinson 2014) and imposed constraints on reionization (e.g., Atek et al. 2015). Despite this recent progress, a number of the interstellar medium (ISM) properties of LBGs remain poorly understood, as their small angular sizes (see, e.g., Grazian et al. 2011) and generally faint dust emission (e.g., Capak et al. 2015) are impediments to detailed study except in the rare cases of strongly lensed systems (e.g., Baker et al. 2004; Coppin et al. 2007) or exceptionally massive, resolved systems (e.g., Magdis et al. 2012). In particular, little is known from long-wavelength observations about the physical processes that drive evolution in LBGs.

For high-\(z\) galaxies, stellar mass is assembled through in-situ star formation and accretion/merger activity (e.g., Conselice 2014; Madau & Dickinson 2014; Somerville & Davé 2015; Scoville et al. 2017). Interactions between galaxies can lead to enhanced star formation, an effect compounded by generally higher gas mass fractions and shorter free-fall times earlier in the Universe (e.g., McKee & Ostriker 2007; Krumholz 2014; Utomo et al. 2018; Krumholz et al. 2018). The same gravitational instabilities that promote star formation can lead to an active galactic nucleus (AGN). Both AGN feedback (driven by accretion onto the supermassive black hole) and star formation feedback (supernovae or stellar winds) can inject energy and shock-heat the ISM gas (see, e.g., Heckman & Best 2014; Thompson et al. 2016; Li et al. 2017; Richings & Faucher-Giguère 2018). For high-redshift samples, Lyman and Balmer lines are shifted into optical and near-infrared (NIR) wavelengths respectively,
enabling the characterization of the ionized gas phase of their ISMs. Observations of these lines have revealed strong outflows, mergers, and gas disks supported by rotation or turbulence/dispersion (Steidel et al. 1996; Shapley et al. 2003; Förster Schreiber et al. 2009; Wisnioski et al. 2015; Wilson et al. 2018). Other phases of high-z galaxies’ ISMs are not as well understood, although recent instrumentation advances at millimeter wavelengths have allowed for more sensitive probes of the cool atomic and molecular gas in high-z systems (e.g., Carilli & Walter 2013; Genzel et al. 2015; Scoville et al. 2016). Warm, potentially post-shock molecular gas in LBGs has never been detected owing to the cosmological redshifting of the relevant spectral lines from rest-frame NIR to mid-infrared wavelengths. Warm H$_2$ has been studied in detail for low-z galaxy samples, but the lack of such data at high redshifts hinders our comprehension of how feedback processes affect different phases of the ISM at earlier epochs.

The existence of LBG analogs (LBAs) in the local Universe gives us the opportunity to study rest-frame UV-selected galaxies in far greater detail than is generally possible at high redshift. From the Sloan Digital Sky Survey (SDSS) and Galaxy Evolution Explorer (GALEX) All-Sky Imaging Survey, Heckman et al. (2005) have identified a sample of $z < 0.3$ starbursts with unusually high FUV luminosities ($\log(L_{\text{FUV}}/L_\odot) > 10.3$) and high FUV surface brightnesses. This sample was later refined to include only the most “supercompact” UV-luminous galaxies with $\log(I_{\text{FUV}}/L_\odot \text{kpc}^{-2}) > 9$ (Hoopes et al. 2007). Overzier et al. (2009) report follow-up observations of the supercompact LBA subsample with Hubble Space Telescope (HST) UV and optical imaging, revealing clumpy star formation and evidence for frequent mergers/interactions. Although morphological signatures of mergers are absent from the majority of high-z LBGs, implying that rapid gas accretion must be fueling their high SFRs, Overzier et al. (2010) show that LBAs also lose their asymmetric merger-like features when artificially redshifted to $z \sim 2 - 4$. Without rest-UV observations sensitive to low surface brightnesses, it is difficult to conclude solely via morphology whether LBGs predominantly grow through merging activity or gas accretion. However, resolved kinematics from integral field spectroscopy can provide additional constraints on their evolution.

Low-z LBAs resemble the LBG population in many other ways; the populations have similar UV-optical colors, dust extinctions, metallicities, SFRs, physical half-light radii, and emission-line velocity widths. Basu-Zych et al. (2007) investigate the radio continuum properties of supercompact LBAs using Very Large Array (VLA) observations, finding that they are characterized by low dust attenuations and a depressed radio to UV ratio. Using CO$(1-0)$ spectral line observations from the Combined Array for Research in Millimeter-wave Astronomy (CARMA), Gonçalves et al. (2014) find that LBAs adhere to the low-redshift Kennicutt-Schmidt (Kennicutt 1998b) law and consume their gas on short ($0.1 - 1$ Gyr) timescales, while harboring large ($0.2 - 0.6$) gas mass fractions. Contursi
et al. (2017) use Herschel PACS and IRAM Plateau de Bure Interferometer (PdBI) observations to provide additional evidence that physical conditions in the ISMs of LBAs are more similar to those of high-redshift sources than local star-forming galaxies.

Gonçalves et al. (2010) investigate the Paα (ionized gas) dynamics for a sample of 19 super-compact LBAs by using Keck/OSIRIS observations with adaptive optics (AO). Their high angular resolution (∼0.1") allows them to probe kinematics on ∼200 pc scales without artificially smearing velocity structure over a large beam, but may also miss faint extended emission. Nevertheless, velocity dispersions are very high, indicating that LBAs are dynamically young systems characterized by strong starbursts and merger activity. By simulating observations of these LBAs at \(z = 2.2\), the authors find strong morphological resemblances to high-redshift LBGs, and show that they can resolve complex kinematic structures that would otherwise be smoothed out at lower physical resolutions.

We have selected for followup six LBAs drawn from the original Heckman et al. (2005) sample, of which four satisfy the Hoopes et al. (2007) supercompact definition (and thus also have HST imaging available). These six LBAs had been targeted for earlier, long-wavelength followup observations on the basis of their rest-UV luminosities and colors (or equivalently, attenuations), which predicted that their far-IR luminosities would be among the highest in the parent sample (e.g., Meurer et al. 1999), and indeed all but one\(^1\) yielded (low-S/N) detections when IRAS 60 \(\mu\)m scans were reprocessed using the SCANPI tool (Alexov et al. 2009). Contursi et al. (2017) later confirmed their high dust-obscured SFRs using Herschel observations. We have obtained VLT/SINFONI integral field spectroscopy at NIR wavelengths for all six targets, providing spatially resolved views of hydrogen and helium recombination in addition to \(\text{H}_2\) ro-vibrational lines. In Table 3.1, we list some of our sample’s properties while the new observations and data reduction are described in Section 3.2. In Section 3.3, we assess ionized gas dynamics using the bright Paα line. In Section 3.4, we use multi-component nebular lines to characterize the global dust attenuations and effective temperatures of star-forming clumps in the ISM. Additionally, we report results on the warm molecular ISM phase for the first time in LBAs. In Section 3.5, we compare the cool and warm molecular gas phases of the ISMs of LBAs, and in Section 3.6, we investigate spatially resolved differences in line flux ratios. Throughout the paper, we assume a concordance ΛCDM cosmology \((H_0 = 70\text{ km s Mpc}^{-1}, \Omega_M = 0.3, \text{ and } \Omega_\Lambda = 0.7)\), such that 1" corresponds to a projected distance of 3.3 kpc at \(z = 0.2\).

\(^1\)J021225 = 0212
3.2 Observations and data reduction

3.2.1 VLT/SINFONI observations

VLT/SINFONI observations of our targets were obtained in service mode between 2008 June 16 and 2008 August 11 via ESO programme 081.B-0947 (PI: Baker). We used SINFONI (Eisenhauer et al. 2003; Bonnet et al. 2004) in its seeing-limited mode, with a 0.25'' pixel^{-1} image scale delivering an 8'' × 8'' field of view. (Most of our targets were sufficiently compact to lie well within this field of view; for J2116, which comprises a pair of objects, our pointing center lay closer to J2116a, with the result that J2116b ended up falling partly outside the field.) For the redshifts of our sample, observing the Paα line required observations in K band. Each observation began with the acquisition of a bright (11.8 ≤ K ≤ 15.4), nearby star, which was also used as a PSF reference, before the telescope was slewed to the science target. Using the SINFONI_ifs_obs_GenericOffset observing template, we then proceeded to take six 600s exposures switched between object (O) and sky (S) positions in an O-S-O-O-S-O sequence; offsets to sky positions were ∼ 20 arcsec, and additional sub-arcsecond dithers between exposures were introduced to facilitate data processing. The choice of a 600s exposure time was motivated by our interest in enabling reliable background subtraction while remaining in the background-limited regime for K band. LBAs J0212, J0213, and J2104 were observed once apiece, i.e., giving 2400s on-source; J0150 and J2116 were observed twice apiece, giving 4800s on-source; and J1434 was observed 1.5 times, giving 3600s on-source. Weather conditions during the observations were generally good, with clear skies and K-band seeing of 0.8'' or better.

3.2.2 Data reduction

Data reduction was completed with the custom software package SPRED (Abuter et al. 2006). SPRED performs the typical data reduction steps for NIR spectra with additional routines necessary to reconstruct the data cube. The final pixel scale is 0.125'' for all data cubes. We measure the PSF by fitting an elliptical Gaussian to the PSF reference source, which has been averaged over all channels. The PSF is typically characterized by FWHM ∼ 0.5 – 0.7'' for our observations. In addition, the residuals from night sky emission lines are minimized using the methods outlined in Davies (2007) through careful wavelength registration and background subtraction, which exploits the sky frames interspersed with on-source exposures. Telluric correction and flux calibration were performed using A- and B-type stars, with flux calibration estimated to be accurate to within 10%.
3.3 Ionized gas kinematics

3.3.1 Moment maps

Using the VLT/SINFONI data cubes, we produce velocity moment maps of the Paα line to study the ionized gas. We begin by subtracting the median brightness at every volumetric pixel (voxel) over the spectral axis in order to produce a line-only data cube. The median subtraction does not leave an imprint on the line-only data cube, as we inspect channel maps (e.g., Figure 3.1) and verify that the edge channels centered around Paα do not contain residual emission. For J0150, certain channels show strong noise features due to lowered sensitivity where a night sky line was subtracted.

We fit spectral lines with single Gaussian components using the `curve_fit` routine in SciPy (van der Walt et al. 2011). Because Paα is the brightest spectral line in our SINFONI observations, we report velocity with respect to the systemic velocity determined from the best-fit Paα centroid. We produce a zeroth moment map of integrated flux by selecting line-only channels and summing over velocity: $M_0 \equiv \int_{\text{line}} I_\nu \, dv$. To create a velocity map, we sum the brightness-weighted velocity over channels with line emission and normalize by the integrated flux at each pixel: $M_1 \equiv (\int_{\text{line}} v I_\nu \, dv)/M_0$. To create a velocity dispersion map, we sum the brightness-weighted squared velocity difference from the mean velocity over channels with line emission, divide by the integrated flux, and take the square root at each pixel: $M_2 \equiv (\int_{\text{line}} (v - M_1)^2 I_\nu \, dv)/M_0^{1/2}$. These moment maps are shown in the upper panels of Figure 3.2, where integrated flux, velocity, and dispersion maps are ordered left to right. For the velocity field and dispersion maps, we have masked pixels that are below the 2 $\sigma$ level in the corresponding zeroth moment maps for visual clarity. No masking is applied to the integrated flux maps in order to provide a realistic estimate of uncertainties due to noise.

We produce a separate set of moment maps in which low signal-to-noise ratio (SNR) pixels are masked, a method sometimes called thresholding or clipping. We estimate the channel rms noise, $\sigma_{\text{rms, chan}}$, by removing 3$\sigma$ outliers with a median absolute deviation (MAD) cut and taking a standard deviation for each channel; outliers are only removed for the purpose of estimating $\sigma_{\text{rms, chan}}$. Subsequently, pixels at each channel with intensity less than 4 $\sigma_{\text{rms, chan}}$ are removed in order to generate a signal-only cube. Our approach implicitly assumes that the astronomical signal significantly exceeds the level of noise and artifacts. These clipped moment maps are shown in the lower panels of Figure 3.2.

---

2The MAD is a more outlier-robust statistic than the root mean square (rms) for estimating noise and, in the case of normally distributed data, is related as $1 \sigma \approx 1.4826 \text{MAD}$.

3If the noise is normally distributed and uncorrelated, then for each $\sim 5000$-pixel image, we would expect of order $\sim 3$ noisy pixels to lie randomly above this threshold. We expect that Paα can thus be cleanly selected at the 4 $\sigma$ level, even if some low-surface brightness emission may also be masked.
Clipped and unclipped moment maps

We can immediately spot visual differences between moment maps generated using the two methods. Unclipped moment maps sometimes wash out coherent motions due to integration of noise, and look more smooth in terms of kinematics. This erasure of velocity structure can be most easily seen by comparing clipped and unclipped first moment maps for J0150 and J2104, i.e., the LBAs with the largest ordered velocity gradients. In their unclipped maps, velocity fields appear tattered and noisy, whereas for their clipped versions, signatures of rotating disks appear very strong. For J0213 and J0212, which have small circular velocities, we also find smoother velocity fields in the clipped moment maps, although now with the opposite effect: rather than being washed out, clumpy velocity gradients emerge in the unclipped velocity maps. These clumps moving at different velocities can manifest because low-brightness flux at extreme velocities (i.e., the most highly blueshifted or redshifted emission) is included, rather than masked as in the clipped case. Gonçalves et al. (2010) find similar lumpy morphologies in their LBA sample and note that such structures appear to be more common in low-mass systems, consistent with our finding that this effect is strongest in the LBAs with the weakest rotational support.

We also find that in J0150, ionized gas bridging the main system with its northwestern companion appears only in the unclipped maps. A similar bridge between components may be present for J2116 as well. In general, low-surface brightness emission is often omitted from the clipped maps, tending to cause Paα flux to be underestimated (although we will later measure line fluxes by modeling them with Gaussian components). For the bright rotating disk seen in J2104, we find that the exclusion of low-surface brightness blueshifted emission toward the south gives the appearance of a stronger redshifted rotating component there (compared to the unclipped version). For J0212, the faintest Paα emitter in our sample, and more importantly, a morphologically irregular merger, clipping prevents faint extreme-velocity flux from being included in moment maps, e.g., as can be seen by comparing J0212’s southeastern edge in the clipped and unclipped versions. Clipped velocity dispersion maps are characterized by lower values for the same reason, particularly near the edges of Paα sources where, most likely, emission from only a single channel was included, corresponding to a dispersion of zero. While the clipped second moment maps likely underpredict dispersions, we can also see that the unclipped maps overpredict them by including noise. These artifacts can manifest as $\sigma \gtrsim 100 \text{ km s}^{-1}$ pixels, which may be clearly rejected as outliers (e.g., for J0212 or J1434); however, in less extreme cases, it is more difficult to determine their validity.
Multi-component systems

We use the $4\sigma$ clipped moment maps to evaluate whether a system comprises multiple components or a single (or heavily blended) source. The unclipped moment maps are too noisy to be used for image segmentation. This choice may bias our estimates of ordered velocity and average dispersion, but the effects are generally small as both quantities are flux-weighted; that is, any low-dispersion pixel near the source edges is likely also to be faint, and therefore contribute only marginally to that pixel’s flux-weighted velocity moment.

In order to mitigate source overlap from galaxy interactions, we use image segmentation functions provided in the `photutils` package (Bradley et al. 2016). Individual components are identified from the clipped integrated flux map profiles. We impose a 40 pixel minimum area requirement to prevent the inclusion of correlated noise artifacts. Distinct galaxy components are labeled and separated in the clipped moment maps, and the segmented masks are also used for determining the kinematics of the unclipped moment maps. Below, we comment on individual LBA systems:

**J0150** We identify and isolate the main system from its blueshifted neighbor to the northwest. From the moment maps, it appears that both sources are supported by rotation. In the unclipped moment map, there appears to be a bridge between the two components, although the channel maps suggest that they are distinct sources.

**J0212** This appears to be a two-component system, although the components are in very close proximity and cannot be separated on the basis of their flux distributions. The two components have some degree of counter-rotation as they collide in what appears to be a merger; unfortunately, no HST observations are available to determine whether these are merging systems or a close pair. Since the VLT observations indicate a level of blending, we treat the J0212 system as a single source.

**J0213** This system is isolated and compact in Pa$_\alpha$ emission. We do not spatially resolve J0213, so its ordered rotation velocity is likely underestimated due to beam-smearing. Based on its UV-optical HST morphology (Overzier et al. 2009), it is possible that low inclination explains J0213’s particularly low ordered rotation velocity.

**J1434** The main LBA has a faint companion to the northwest at higher recessional velocity. The companion appears to be co-rotating with the main galaxy, such that Pa$_\alpha$ emission is both

---

$^4$An area of 40 pixels is slightly larger than the PSF size of $\sim 30$ pixels. In practice, however, this size threshold is needed to reject noise and artifacts cleanly (particularly near the detector edges), while being sufficiently small to recover all main Pa$_\alpha$-emitting sources. All analyzed LBA components are detected by the segmentation program, and the choice of minimum area only affects visualization.
spatially and spectrally continuous between star-forming components. Our image segmentation procedure does not separate these components, so we treat them as a single blended source.

**J2104** This LBA is the most luminous Paα emitter in our sample with line emission extending over 1000 km s\(^{-1}\). Blueshifted Paα emission begins to blend with a pair of He I recombination lines at 1.869 and 1.870 µm. From the channel maps, we see three marginally resolved neighbors to the south of the main system; one of these has detectable UV emission in HST imaging (Overzier et al. 2009). The neighbors can be visually identified in the unclipped integrated flux map, but only the southernmost one is shown in the clipped moment map because of the 40-pixel minimum size cut. As they are too faint to be detected in other emission lines, we restrict our analysis to the main component.

**J2116** We find two distinct components separated by a few arcsec. We measure \(v/\sigma\) for both J2116a, the system at the center of the SINFONI imaging, and J2116b, the brighter neighbor towards the east. A small portion of J2116b’s redder emission extends off the edge of the detector, which will cause its ordered velocity to be underestimated. We therefore choose to define its rotational velocity as its peak blueshifted velocity, assuming that its velocity profile is approximately symmetric about the systemic velocity.

**Kinematic comparisons using \(v/\sigma\)**

We compute the ratio of ordered rotation velocity to velocity dispersion, \(v/\sigma\), where \(v \equiv (v_{\text{max}} - v_{\text{min}})/2\) is half the range in the flux-weighted velocity field,\(^5\) and \(\sigma \equiv \text{mean}[(\int v^2 I_d dv/M_0)^{1/2}]\) is the Paα flux-weighted velocity dispersion averaged over all pixels. For ordinary star-forming galaxies, \(v/\sigma\) measures the degree of rotational support. Strong turbulent processes, such as the injection of energy from star formation or AGN feedback, will physically puff up galaxy disks and inflate \(\sigma\). For interacting systems, \(v/\sigma\) can skew low because \(\sigma\) is elevated by disks with multiple inclinations or even counterrotation, relative to a single gas disk. Such galaxy interactions can also promote star formation and escalate feedback processes.

In Table 3.2, we list each LBA and its ordered velocity, dispersion, and \(v/\sigma\) ratio. We have not applied inclination corrections to our sample due to the relatively low resolution of our observations, and also because of the irregularity of most sources. On average, this effect would increase \(v\) by a factor of \(4/\pi \approx 1.3\) (see, e.g., the Appendix of Law et al. 2009). From the clipped maps, we measure average \(\langle v/\sigma \rangle = 1.70 \pm 0.74\), indicating that our LBA sample is characterized by modest

\(^5\) \(v_{\text{max}}\) and \(v_{\text{min}}\) are the maximum and minimum velocities in the first moment map, respectively.
rotational support (as opposed to kinematic support from random motions). The projected ordered velocity spans a large range, from $35 - 250$ km s$^{-1}$, signifying a diverse set of velocity structures and/or inclinations in our sample. We find that velocity dispersions are high across our sample, and our mean dispersion $\sigma = 76 \pm 45$ km s$^{-1}$ (or median $\sigma = 57$ km s$^{-1}$) greatly exceeds the $\sigma \approx 5 - 15$ km s$^{-1}$ characteristic of low-$z$ disk galaxies (Dib et al. 2006). All LBAs have significantly smaller $v/\sigma$ than the values commonly found in local spirals ($\sim 10$).

It is possible that resolution effects may artificially increase $\sigma$ via beam smearing, which is most problematic at small radii where the ordered velocity gradient is large, i.e., near the center of a rotating disk. However, from our moment maps in Figure 3.2, we see that the dispersions near the LBAs’ nuclei generally appear flat (with the exceptions of J0213 and J2104). If we repeat our calculations of $v/\sigma$ after masking the central beam/PSF, the ordered velocity does not change at all (since the most extreme velocities tend to be located at the edges of rotating disks), and the updated median dispersion $\sigma = 54$ km s$^{-1}$ is still high relative to those of low-$z$ galaxies. For all LBAs except J0213 and J2104, we find that the reduction in $\sigma$ is less than 10%, and there is a small increase in $v/\sigma = 1.87 \pm 0.73$. We will see in Section 3.4.2 that J0213 and J2104 both have genuine broad-line Pa$\alpha$ components, which are expected to be strongest in their central regions if driven by nuclear starbursts or AGN; thus, it is possible that beam smearing causes their flux-weighted dispersions to be overestimated.

We also repeat our analysis with the unclipped moment maps to examine how our methodology changes $v/\sigma$. However, several complications arise due to the noisy nature of the unclipped maps. Because of the way that we have defined ordered rotation velocity $v \equiv (v_{\text{max}} - v_{\text{min}})/2$, noisy extremal values can completely corrupt estimates of this quantity. Therefore, we remove outlying velocities via a $\pm 3 \sigma$ (MAD) cut before selecting $v_{\text{max}}$ and $v_{\text{min}}$. Even after this step, ordered velocities tend to be extremely and unphysically high. Image segmentation on the basis of unclipped integrated flux maps becomes difficult because, despite our outlier rejection, noise from different velocities can still be included in the maps. We thus apply the same segmentation masks from the clipped maps to the unclipped maps, and then measure $v$ and $\sigma$. Using this methodology, we find that the ordered velocities increase for fainter systems (e.g., J0212, J0213, and J2116a) while the ordered velocities decrease for brighter systems (e.g., J0150, J2104, and J2116b). The mean (median) dispersive velocity remains high: $89 \pm 42$ km s$^{-1}$ (73.9 km s$^{-1}$). For unclipped moment maps, we measure average $\langle v/\sigma \rangle = 1.41 \pm 0.53$, in agreement with (but slightly lower than) the value determined using the clipped data. These results are also shown in Table 3.2.

Such low values of $v/\sigma$ generally indicate that LBAs are subject to intense feedback processes that elevate dispersions in their ISMs. Another interpretation is that these LBAs are dynamically young
and are characterized by recent star formation or even merger activity, which would lower ordered velocities while raising random gas motions. These findings are not surprising, given that LBAs resemble high-z galaxies and are dominated by young stellar populations that can drive significant feedback. Interacting systems seem to be common in our sample as well, although most systems seem to still possess a main rotating disk (with the notable exception of J0213).

3.3.2 Disk modeling

As discussed above, it is possible that velocity dispersions measured from moment maps can be artificially boosted by beam smearing (see, e.g., Davies et al. 2011). By dropping the central regions, we have found that this effect is relatively small for most LBAs; however, we must still verify that beam smearing is not responsible for transforming rotational kinematics into dispersive motions in the outer disks. These concerns are warranted in particular at higher redshifts, where the angular sizes of distant sources become comparable to the observed resolutions. In such cases, it is often advantageous to use sophisticated models that can factor in the PSF and instrument’s spectral broadening while constraining the disk’s intrinsic kinematics (e.g., Cresci et al. 2009; Kassin et al. 2012; Newman et al. 2013; Genzel et al. 2014; Simons et al. 2015, 2016, 2017; Johnson et al. 2018).

We use the GalPaK3D software (Bouché et al. 2015) to model the kinematics of our LBA Paα line emission directly from the SINFONI data cubes. GalPaK3D can account for seeing effects and instrumental line broadening, which allows us to extract not only kinematic quantities (v, σ, and velocity turnover radius), but also geometric and shape parameters (inclination, position angle, and half-light radius). We use the default exponential surface brightness model with an arctan rotation curve and thick disk (see, e.g., Binney & Tremaine 2008; Genzel et al. 2008).

In general, we do not impose any priors other than the default uniform priors, although for J2116a and J2116b we also implement uniform priors on the galaxy center’s coordinates (restricted to a range of ±2 pixels in x and y coordinates). For J1434 we impose priors on v (uniform between 100 and 200 km s\(^{-1}\)) and on σ (uniform between 40 and 100 km s\(^{-1}\)), both of which were determined from the moment map analysis. We extract median-subtracted subcubes from the SINFONI data cubes centered on the Paα lines and ensure that line-free channels are included so that GalPaK3D can determine the noise. For J1434, J2116a, and J2116b, we mask out companion sources in order to ensure good fits to the data. For each source, we run the Markov Chain Monte Carlo (MCMC) routine until convergence for each LBA, which is between 3000 to 4000 steps; we then use the final 700 steps for estimating the mean and standard deviation for v and σ. Both v and σ are well-constrained by the model, but the turnover radius can take on a larger range of values. The modeled half-light radius generally agrees with the observed radius from SDSS optical imaging. In
Table 3.2, we show each object’s kinematically modeled $v$ and $\sigma$, as well as $v/\sigma$.

For a few of our LBAs (J0213, J2104, and J2116b), the extracted half-light radii for Pa$\alpha$ emission are small compared to the seeing ($\sim 0.7''$), and thus the shape and inclination parameters for these objects may be unreliable. Bouché et al. (2015) find that in such cases, the modeled rotational velocity also tends to be biased high (e.g., see their Figure 8), although $\sigma$ can still be extracted with no apparent bias. Thus, we will treat $v/\sigma$ for these three objects as overestimates of their true rotation-to-dispersion ratios.

From GalPaK3D modeling results, our LBAs have an unweighted mean and median velocity dispersions of 82.1 km s$^{-1}$ and 87.4 km s$^{-1}$, respectively. Most kinematic model dispersions are higher than those inferred from moment maps, confirming that our observations are not biased by beam smearing effects (with the exception of J2104). All sources are described by $v/\sigma < 3$, and our LBA sample’s unweighted mean $\langle v/\sigma \rangle$ is $1.20 \pm 0.84$. The averaged model kinematics are broadly in agreement with the results from our moment map analysis (now corrected for inclination): $\langle v/\sigma \rangle = 1.80 \pm 0.67$ (unclipped) and $2.16 \pm 0.94$ (4 $\sigma$ clipped). Therefore, we can confirm using both analyses that the LBAs have high velocity dispersions in their disks, and have generally low ordered rotation-to-dispersion ratios.

### 3.3.3 Comparison to previous results

We can compare our results with those of Basu-Zych et al. (2009) and Gonçalves et al. (2010), which are based on 19 LBAs observed with Keck/OSIRIS using adaptive optics (AO). They find low overall $\langle v/\sigma \rangle = 1.22 \pm 0.50$ (where we have adjusted the average by a factor of $4/\pi$ to correct for average inclination). J0150, J1434, and J2104 from our sample also belong to theirs, so we can compare the kinematics derived from our lower-spatial resolution observations with their AO observations. (J0213 is also a member of their sample, but they do not measure a velocity gradient, so they exclude it from their analysis.) Despite the potential for spatial blending of pixels to masquerade as lower ordered rotational velocity, we find higher ordered velocity than Gonçalves et al. (2010). In J2104, for example, we find an ordered velocity of $\sim 200 - 250$ km s$^{-1}$ (in projection), higher than the 145 km s$^{-1}$ reported by Basu-Zych et al. (2009). Since our values of $\sigma$ from moment maps agree, we can be confident that the large $\langle v \rangle$ in our sample is not an artifact of our lower resolution. The disagreement can be explained if we are able to detect extended emission that traces material at larger radii and rotation velocities because of better surface brightness sensitivity. In addition to observing for longer durations (e.g., 2400 – 4800 s compared to 1500 – 2700 s for Gonçalves et al. 2010), our larger pixel scales imply that our surface brightness sensitivities are $\sim 4 \times$ better. Our LBAs have optical half-light radii of $\sim 2.5 - 5$ kpc and are multiple beams across, indicating that
beam smearing is unlikely to cause statistical mischaracterization of the velocity dispersion (e.g., Newman et al. 2013); the kinematic modeling results also support our conclusion that beam smearing is a minor issue.

The SINS survey has obtained ionized gas dynamics for higher-redshift star-forming galaxies (Cresci et al. 2009; Förster Schreiber et al. 2009). The authors infer velocity and dispersion terms for 18 $z \sim 2$ sources using dynamical modeling of H$\alpha$ emission. They estimate $\sigma \sim 30 - 80$ km s$^{-1}$ from the outer regions of their sources, which are similar to our measured values. However, compared to our sample, theirs is characterized by stronger signatures of (inclination-corrected) ordered rotation, with $v_{\text{max}} \sim 150 - 300$ km s$^{-1}$. This result is not surprising given that their sources are physically larger and generally more massive than LBAs or LBGs. For the Förster Schreiber et al. (2009) inclination-corrected sample, $\langle v/\sigma \rangle \sim 4.4$, although approximately a third are dispersion-dominated ($v/\sigma < 1$).

We also compare results with results from KMOS$^{3D}$, which surveyed $0.7 < z < 2.7$ objects with the VLT/KMOS instrument (Wisnioski et al. 2015). These authors report mean velocity dispersions extracted along each galaxy’s major axis: $\sigma = 25$ km s$^{-1}$ at $z \sim 1$ (for which the mean stellar mass is $\log(M_*/M_\odot) = 10.65$) and $\sigma = 48$ km s$^{-1}$ at $z \sim 2$ ($\log(M_*/M_\odot) = 10.86$). Wisnioski et al. (2015) describe the majority of objects in both redshift bins as velocity-supported ($v/\sigma > 1$), and also find that inclination-corrected rotation-to-dispersion velocity ratio decreases from $\langle v/\sigma \rangle \sim 5$ at $z \sim 1$ to $\langle v/\sigma \rangle \sim 3$ at $z \sim 2$.

The MOSDEF sample comprises $z \sim 1 - 3$ galaxies with $\log(M_*/M_\odot) \gtrsim 9$ and diverse star formation and dust properties (observed with Keck/MOSFIRE; Kriek et al. 2015). Price et al. (2016) investigate the dynamical properties of a subsample of MOSDEF objects, finding median velocity dispersions $\sigma = 70$ km s$^{-1}$ at $z \sim 1.5$ and $\sigma = 80$ km s$^{-1}$ at $z \sim 2.3$. They also report that the median integrated H$\alpha$ rotation-to-dispersion velocity ratio (corrected for inclination) is $v/\sigma \sim 2.7$, with a modest anticorrelation between $v/\sigma$ and sSFR that may be due to disk settling (see also, e.g., Wisnioski et al. 2015; Price et al. 2019). Although our sample is similar to the KMOS$^{3D}$ and MOSDEF objects in terms of stellar mass, the LBAs appear to be characterized by lower $v/\sigma$. The relatively high level of velocity dispersion in LBAs/LBGs may be due to more recent star formation activity, which prevents the disk from becoming kinematically settled.

Law et al. (2009) present kinematics for $z \sim 2 - 3$ star-forming galaxies using AO-assisted Keck/OSIRIS observations. Most of these sources have been selected via a generalized Lyman-break technique, although a few submillimeter and Ly$\alpha$-emitting galaxies are also included in their sample. This paper’s targets are generally less massive and have smaller half-light radii than the SINS sample, and thus appear to be better suited for comparison to our LBA sample than the SINS targets. Law
et al. (2009) find very high velocity dispersions, with \( \sigma \sim 60 - 100 \) km s\(^{-1}\) and extremely low average velocity-to-dispersion ratio (i.e., inclination corrected \( \langle v/\sigma \rangle \sim 0.3 \)). In summary, we find qualitative agreement between LBAs and the Law et al. (2009) sample, as opposed to the stronger rotational support found in more massive, non-UV-selected high-z galaxies and in typical local spirals.

### 3.4 The multiphase ISM

The SINFONI IFU data reveal a wealth of information about multiple phases of the ISM, which we can use to understand physical processes that drive galaxy evolution. For each LBA, we extract a spectrum from a two-pixel radius aperture centered on the \( \text{Pa}_\alpha \) emission peak. In these spectra, shown in Figure 3.3, we find bright recombination line features, primarily from neutral hydrogen (H I) and neutral helium (He I). From recombination line strengths and their ratios, we are able to infer properties of the ionized (H II) regions, such as the nebular dust attenuation and the effective temperature of young stars. We also detect molecular hydrogen ro-vibrational lines (\( \text{H}_2 \) 1-0 S(·) series) in our spectra. \( \text{H}_2 \) ro-vibrational line ratios are indicators of the gas excitation temperature and provide unique insight into feedback processes in the ISMs of LBAs.

#### 3.4.1 Dust attenuation

We model the effects of dust by assuming a uniform attenuating screen described by the Calzetti et al. (1994, 2000) attenuation curve, \( k(\lambda) \). This empirical curve relates the observed attenuation to the color excess, e.g., in the optical \( V \)-band: 

\[
A_V = k(V) \times E(B-V)
\]

We characterize the color excess in nebular regions by using pairs of the brightest available hydrogen recombination lines (H\( \beta \), H\( \alpha \), and Pa\( \alpha \)).

We have obtained H\( \alpha \) and H\( \beta \) line fluxes by querying the SDSS DR14 SkyServer (Abolfathi et al. 2018). Tremonti et al. (2004) and Brinchmann et al. (2004) describe the process by which these line fluxes are extracted from optical spectra over 3\" SDSS fiber apertures. For all LBAs except J2116, which has no H\( \alpha \) flux available, we find \( E(B-V)_{\text{neb}} \) using the Balmer decrement (assuming the attenuation curve from Calzetti et al. 2000). The average nebular color excess for our sample is 

\[
\langle E(B-V)_{\text{H}\alpha+\text{H}\beta} \rangle = 0.256 \pm 0.041
\]

Our new SINFONI observations allow for additional estimates of the dust attenuation. Using Pa\( \alpha \) in the NIR data cubes along with H\( \alpha \) and H\( \beta \) lines enables robust estimation of the attenuation curve over a large range in wavelengths. We extract Pa\( \alpha \) over a 3\" diameter aperture centered on the continuum emission, matching the SDSS spectroscopic fiber placement, and fit a single Gaussian component to each median-subtracted spectrum. Integrated flux uncertainties are estimated by
taking the variance of 10,000 MCMC samples, excluding a burn-in of 10,000 steps. Using the Calzetti et al. (2000) attenuation curve and intrinsic line ratios from Hummer & Storey (1987), we compute unweighted averages for the LBA sample: \( \langle E(B - V)_{\text{Pa}\alpha + \text{H}3} \rangle = 0.316 \pm 0.113 \) and \( \langle E(B - V)_{\text{Pa}\alpha + \text{H}\alpha} \rangle = 0.349 \pm 0.166 \). In Table 3.3, we show \( E(B - V)_{\text{neb}} \) color excesses derived using these line ratios.

We find that there is some variation in nebular color excesses computed using different pairs of recombination lines, although the statistical significance is low due to large uncertainties.\(^6\) The average nebular color excesses are related as \( \langle E(B - V)_{\text{Ho} + \text{H}\beta} \rangle < \langle E(B - V)_{\text{Pa}\alpha + \text{H}3} \rangle < \langle E(B - V)_{\text{Pa}\alpha + \text{H}\alpha} \rangle \), which is in line with expectations given that bluer photons are likely emitted from shallower regions in the case of a mixture of dust, gas, and stars. \( \text{Pa}\alpha + \text{H}\beta \)-based attenuation spans the largest wavelength range and thereby offers the most sensitivity to differential attenuation. Since \( \text{Pa}\alpha \) and \( \text{H}\beta \) lines are well-measured for all six LBAs, we use \( E(B - V)_{\text{neb}} \) estimated from \( \text{Pa}\alpha + \text{H}\beta \) to correct VLT/SINFONI nebular emission for dust attenuation. This correction, using \( \langle E(B - V)_{\text{Pa}\alpha + \text{H}3} \rangle = 0.32 \pm 0.11 \), is relatively low. For our sample, we find that correction for attenuation increases the NIR nebular line fluxes by 12–25%. The dust corrections do not strongly impact line ratios; for example, the ratio of He I 1869/Br\( \delta \) is only increased by 2%.

The continuum attenuation at 1530 Å, \( A_{1530} \), has been inferred from stellar population synthesis model fits to SDSS and GALEX photometry (see Salim et al. 2005 for details). Such a measurement depends most strongly on near- and far-UV observations, from which a UV slope can be derived (commonly parameterized by \( \beta \); see, e.g., Meurer et al. 1999; Reddy et al. 2015; Salmon et al. 2016). For our sample of LBAs, Heckman et al. (2005) measure \( A_{1530} \sim 2 \). Individual uncertainties are not provided for model parameters, but the total scatter in \( A_{1530} \) is about 30% (i.e., about 0.6 for our level of attenuation; Salim et al. 2005). Although some of this scatter is likely due to intrinsic scatter in a diverse sample of galaxies, we include an uncertainty of 0.5 in \( A_{1530} \). The far-UV attenuation can be converted to a continuum color excess, \( E(B - V)_{\text{cont}} \), assuming an attenuation curve similar to the one used for nebular regions (with normalization \( R_V = 4.05 \); Calzetti et al. 2000). In Table 3.3, we show \( E(B - V)_{\text{cont}} \) derived from \( A_{1530} \) and these assumptions.

We find that the ratio of UV continuum to nebular color excess is

\[
E(B - V)_{\text{cont}} = (0.82 \pm 0.19) \times E(B - V)_{\text{Ho} + \text{H}\beta}, \tag{3.1}
\]

\(^6\)We have reported unweighted averages for \( E(B - V)_{\text{neb}} \) thus far, because one source with unusually low uncertainties would otherwise dominate the average. Inclusion of J0212 in the weighted average would bias the sample \( E(B - V)_{\text{neb}} \) in an unrepresentative manner: 0.24 ± 0.04, 0.26 ± 0.09, and 0.24 ± 0.11 for \( \text{Ho} + \text{H}\beta \), \( \text{Pa}\alpha + \text{H}\beta \), and \( \text{Pa}\alpha + \text{H}\alpha \) respectively. When J0212 is excluded, the weighted average nebular excesses are 0.27 ± 0.02, 0.32 ± 0.03, and 0.35 ± 0.04 respectively.
where in this case we show the nebular color excess derived from $H\alpha$ and $H\beta$ lines. We can also compute

$$E(B - V)_{\text{cont}} = (0.62 \pm 0.17) \times E(B - V)_{\text{Pa}\alpha + H\beta},$$

$$E(B - V)_{\text{cont}} = (0.54 \pm 0.19) \times E(B - V)_{\text{Pa}\alpha + H\alpha},$$

using weighted averages of the color excess ratios. In Table 3.3, we show the continuum-to-nebular color excess ratio, hereafter defined as $f$, computed using $\text{Pa}\alpha$ and $H\beta$ for all six LBAs.

It is unsurprising that nebular light is more reddened than continuum light, as we have found here, because the ionized gas traces massive OB stars that tend to be embedded in high-attenuation birth clouds, while the continuum light originates from less massive stars that may have outlived or dissipated their birth clouds (see, e.g., Charlot & Fall 2000). Simple models account for this effect by requiring two components of dust, one that is diffusely spread out across the entire ISM, and one that only heavily redens or obscures the star-forming nebular regions. If the two dust components are well-mixed throughout the galaxy, or at least in the regions traced by UV and emission lines, then the nebular and continuum color excesses should be equal to one another (e.g., Calzetti 2001). Further considerations such as complex geometries or morphological line-of-sight effects can be added to this model, but the overall intuition remains the same (see, e.g., Wild et al. 2011; Kriek & Conroy 2013; Price et al. 2014).

For low-redshift, typical star-forming spirals, the ratio of continuum-to-nebular color excess is $f = 0.44$ (Calzetti et al. 2000). At higher redshifts, $f$ ranges from about one-half (e.g., Wuyts et al. 2013; Price et al. 2014) to near unity (e.g., Kashino et al. 2013; Reddy et al. 2015). The Calzetti et al. (2000)-like low values of $f$ can be physically interpreted using the previously mentioned two-component dust model, where the attenuated UV emission is not co-spatial with dustier nebular regions. Systems for which $f$ is higher may host a more homogeneous mixture of star-forming clouds and older stars, such that both UV-bright and ionizing stars are effectively attenuated by the same dust screen.

For our sample of LBAs, we find that $f_{H\alpha + H\beta} > f_{\text{Pa}\alpha + H\beta} > f_{\text{Pa}\alpha + H\alpha}$ (although uncertainties are large). This ordering of color excess ratios is expected because the color excess derived using longer-wavelength lines (i.e., $\text{Pa}\alpha$) is probing deeper into attenuated regions, whereas the shorter-wavelength Balmer lines can only probe shallower surfaces of star-forming clouds. $A_V$ probed by $\text{Pa}\alpha$ is nearly an order of magnitude larger than for $H\beta$. At the optical depths probed by $H\alpha$ emission, the ionizing stars are nearly as obscured as the non-ionizing stars that dominate the UV emission, based on the measured $f_{H\alpha + H\beta} \sim 0.8$. At the greater optical depths probed by $\text{Pa}\alpha$, the more
embedded ionizing stars are more highly obscured than the non-ionizing populations, as evidenced from the more Calzetti-like $f_{\text{Pa} \alpha + \text{H} \alpha} \sim 0.5$.

Other studies of LBAs (see, e.g., Basu-Zych et al. 2007) have found that Balmer decrement-derived line attenuations are low (i.e., $A_{V,\text{neb}} \sim 1$), but that continuum-based attenuation is comparable to that seen in typical low-$z$ star-forming galaxies. Implied nebular excess ratios are therefore high, and comparable to our measurement of $f_{\text{H} \alpha + \text{H} \beta} \sim 0.8$. These previous findings are consistent with the interpretation that LBA stellar populations traced by Balmer lines are still fully embedded in their birth clouds.

At higher redshifts, there is evidence that galaxies have been forming stars at a higher clip, and also undergo mergers more frequently. Both high-$z$ LBGs and our sample of LBAs exhibit high velocity dispersions, which may be a consequence of the recent and ongoing star formation. The combination of recent merger and star-formation activity could then explain a high degree of mixing between dusty clouds hosting ionizing stars and the more diffuse dust associated with populations of less massive stars (e.g., Reddy et al. 2015; Pannella et al. 2015; Puglisi et al. 2016).

Not all high-redshift studies have found evidence for highly mixed populations of ionizing and non-ionizing stars. Theios et al. (2018) report $f_{\text{H} \alpha + \text{H} \beta} \sim 0.75$ using SED fits to the continuum color excess (as do Steidel et al. 2016), although the authors note that the ratio falls to $\sim 1/3$ when the SMC attenuation curve is applied (as described in Gordon et al. 2003; Reddy et al. 2016) instead of the Calzetti et al. (2000) curve. The SMC curve is steeper at shorter wavelengths, and appears to better match the canonical conversion factor of $f = 0.44$, particularly for lower-mass systems. The higher-mass SINS sample (Förster Schreiber et al. 2009) is characterized by a lower, Calzetti-like $f$, which the authors cite as evidence for an additional level of attenuation toward nebular regions. Wild et al. (2011) are able to explain this effect using a strong correlation between $f$ and sSFR. We can physically interpret these trends by positing an inverse relationship between the age of the starburst and the degree of mixing for younger and older stellar populations.

Our results are also consistent with an empirical “double Calzetti” law (see, e.g., Charlot & Fall 2000; Wuyts et al. 2013; Price et al. 2014; Lo Faro et al. 2017), which combines a steep Calzetti et al. (1994) attenuation curve for nebular regions at shorter wavelengths (e.g., for Balmer lines), and a shallower curve at longer wavelengths (e.g., for Pa$\alpha$). At longer wavelengths, the nebular color excess appears higher because the optical depth reaches unity deeper into the dusty star-forming clouds where stars form. The extra attenuation toward the most embedded clouds accounts for why ionized gas traced by Pa$\alpha$ is closer to the canonical Calzetti et al. (2000) ratio of $f \sim 0.44$. 
3.4.2 The near-infrared recombination spectrum

In addition to strong Pa\(\alpha\) lines, we find hydrogen Brackett series lines and He I 2.059 \(\mu\)m and 1.869 \(\mu\)m doublet lines present in all LBA spectra. These lines arise from the H and He I recombination cascades in gas photoionized by star formation or other physical processes. In J2116, multiple lines have been detected unambiguously in both components, so we report results for the two separately.

Measuring line fluxes and kinematics

We measure the properties of recombination lines over a smaller aperture (\(\sim 0.6''\), or about the size of the PSF) in order to accurately determine kinematics near the LBAs’ central regions. Spectra are extracted from a 2-pixel radius aperture centered on the brightest Pa\(\alpha\) pixel as determined from the zeroth moment map. For Pa\(\alpha\), Br\(\delta\), Br\(\epsilon\), Br10, and He I 2.059 \(\mu\)m, we shift spectral lines to a common velocity frame and extract the spectra between \(-1500 \text{ km s}^{-1} < v < 1500 \text{ km s}^{-1}\). Br11 is observable for J0213 and J1434, but the line is also very faint, so we do not include it in our analysis. For the case of the He I doublet comprising lines at 1.869 and 1.870 \(\mu\)m, due to strong Pa\(\alpha\) emission at slightly longer wavelengths, we use the spectrum from \(-2500 \text{ km s}^{-1} < v < 500 \text{ km s}^{-1}\). For each spectral “cutout” over its limited velocity range, we estimate the median absolute deviation (MAD) and convert to a standard deviation, \(\sigma\), assuming Gaussian statistics. We subtract the local continuum by taking a median of the (mostly) line-free channels \(|v| > 500 \text{ km s}^{-1}\) for each spectral cutout. In Figure 3.4, we show recombination lines from an example median-subtracted, dust-uncorrected spectrum.

Each line is fit to a three-parameter Gaussian curve using SciPy optimization, and we sample parameter uncertainties using MCMC (the emcee package; Foreman-Mackey et al. 2013b) with 10,000 steps (after 10,000 burn-in steps). He I 1869+1870 lines are blended, so we model them jointly with a double Gaussian at a single velocity and with fixed line ratios determined by Benjamin et al. (1999), such that only three parameters are free to vary. For Pa\(\alpha\), we find that a single Gaussian does not always adequately fit the data, as residuals (data − model) sometimes show signs of another Gaussian component. We therefore also jointly fit two Gaussian components to Pa\(\alpha\), and select between one-component and two-component models on the basis of a lower Akaike Information Criterion (AIC; Akaike 1974). We have implemented the following priors: (1) the line width\(^8\) is assumed to be distributed \(\sim \mathcal{N}(\text{FWHM}_{\text{Pa}\alpha},[50 \text{ km s}^{-1}]^2)\) (i.e., centered on the Pa\(\alpha\) FWHM line width measured

\(^7\)Hereafter, we refer to He I 2.059 \(\mu\)m as He I 2059, and the He I doublet at 1.869 and 1.870 \(\mu\)m as He I 1869, or He I 1869+1870 when space permits.

\(^8\)The FWHM line width is related to the standard deviation assuming a Gaussian distribution: \(\text{FWHM} = 2\sqrt{2\ln 2}\sigma\).
from the clipped moment maps in Table 3.2, and with standard deviation 50 km s\(^{-1}\)), and (2) the velocity center is assumed to be distributed \(\sim \mathcal{N}(v_{\text{sys}}, [50 \text{ km s}^{-1}]^2)\), where \(v_{\text{sys}}\) is the systemic velocity corresponding to the redshift. For Pa\(\alpha\), we allow parameters of the second component to freely vary, but with only a prior on the velocity center \(\sim \mathcal{N}(v_{\text{sys}}, [150 \text{ km s}^{-1}]^2)\).

In Figure 3.5, we compare best-fit model recombination line components (shown in red) with the data (blue). For high-SNR lines, the Gaussian models fit the data well, and the model scatter is consistent with uncertainties in the data. Some lines are fit with considerable uncertainty, such as the fainter lines from J0212 and J2116a. For J0150, J2104, and J2116b, all available recombination lines are well-described by the model.

**Hydrogen recombination lines**

Best-fit line fluxes and widths are determined from the medians of posterior MCMC samples, and uncertainties from the standard deviations of these distributions. In Table 3.4, we show the total recombination line fluxes. The values for Pa\(\alpha\) are different from the fluxes in Table 3.3, which were used for computing dust attenuation and meant to be compared directly with SDSS-measured H\(\alpha\) and H\(\beta\) lines (Tremonti et al. 2004). In this case, we have fit Pa\(\alpha\) with either a single or double Gaussian line profile depending on which fit minimizes the AIC. We note again that these spectra are extracted from a 2-pixel (or \(\sim 0'.6\)) aperture, whereas in Section 3.4.1, the 3' SDSS matched-aperture spectra capture ionized gas dynamics over larger fractions of the galaxies and may include the outer disk or neighboring components. The present apertures are also now centered on the Pa\(\alpha\) integrated flux maxima, rather than the continuum brightness peaks.

In Table 3.5, we show the recombination line widths. We also list individual fluxes, line centers, and line widths for Pa\(\alpha\) components in order to compare dynamics. For the two LBAs with the faintest Pa\(\alpha\) emission, J0212 and J2116a, only a single kinematic model component is favored, but the others require two components. For most LBAs, the line width of the more luminous Pa\(\alpha\) component is similar to those of the other recombination lines. This more luminous component can be interpreted as tracing the star-forming ISM in each galaxy, with a velocity width that depends on the host’s individual dynamics. The fainter component can be attributed to ionized gas being heated and driven out of the galaxy disk, perhaps by supernovae, stellar winds, or AGN feedback, albeit with different degrees of confidence for different systems. For J2104, the most Pa\(\alpha\)-luminous LBA, we find evidence for blueshifted (\(v \sim 300 \text{ km s}^{-1}\)) ionized gas with significant Pa\(\alpha\) flux, consistent with expectations for an energetic outflow. For J0150, J0213, and J1434, the second Pa\(\alpha\) component has a broader line width than the first. For LBAs in general, we would expect the rotating disk component to dominate over emission from velocity-broadened material in the Pa\(\alpha\) line. However,
this is not the case for J0213 and J1434, where the broad-component fluxes are similar to those of the narrow components. For J0213, despite the fact that both components are centered at its systemic velocity, the Paα profile cannot be fit by a single Gaussian model without broad “wings” showing up in the residuals. Finally, while J2116b appears to have a bright blueshifted component, from inspecting Figure 3.2, we determine that a significant portion of the redshifted emission is actually off the detector, so the asymmetric line profile is not due to any physical phenomenon.

In Figure 3.6, we compare observations to the dust-free emissivity ratios provided by Hummer & Storey (1987) and Benjamin et al. (1999) for Case B recombination line ratios at $T_e = 10^4$ K for hydrogen ($n_e = 10^3 \text{ cm}^{-3}$) and neutral helium ($n_e = 10^4 \text{ cm}^{-3}$) respectively. In the figure, fiducial hydrogen lines fluxes are normalized to the observed Brδ flux, and helium lines to He I 1869+1870. The recombination spectrum is generally not sensitive to the electron temperature or density of the ionized medium. We find that case B predictions generally agree with the data.

Brϵ and Br10 appear to be under-luminous for J0150, J0213, and J1434, whose data otherwise seem to match the Case B models. Three possible explanations can conceivably contribute to this phenomenon. First, the Brδ flux may be overestimated due to spectral features (of unknown origin) that artificially diminish the continuum (as can be seen in Figure 3.5), and thereby elevate Brδ flux. If this is indeed the case, then all lines would appear to be lowered with respect to Brδ, except Paα, which is fit using a two-component model and may include emission not captured by the other single-Gaussian fits. Second, the nebular dust attenuation as estimated from a global integration over the large aperture might be considerably lower than attenuation extracted over only the central regions. If this were the sole reason for the low Brϵ and Br10 fluxes, then Paα, which is blueward of Brδ, should also be slightly biased low; however, we can see that this is not the case in Figure 3.6. Third, night sky lines (whose corrections are evident in Figure 3.5 as ringing patterns in the spectra) may add noise to the line fits, although we do not see any reason why the fluxes of only Brϵ and Br10 would be biased low.

A combination of the first and second points can explain why Paα and Brδ are apparently more luminous relative to Brϵ and Br10 than expected for Case B recombination. Our method of estimating the continuum generally produces valid results, but it is possible that unknown absorption lines exist near Brδ. Moreover, Paα properties cannot be directly compared to those of other hydrogen recombination lines, due to the former’s different fitting procedure. Finally, radial gradients in metallicity may correlate with decreased attenuation in the outer regions of the LBAs (see, e.g., Cresci et al. 2010). This trend can lead to overly low estimates of $E(B - V)$ in central regions, and thus to underprediction of the unattenuated fluxes of bluer lines (e.g., Br10 and Brϵ).
He I recombination

We can constrain the spectral hardness of ionizing radiation by comparing line ratios across the He I and hydrogen recombination cascades. Such an approach has been shown to be an useful diagnostic of the effective temperature \( T_{\text{eff}} \); see, e.g., Rigby & Rieke 2004) of ionizing star populations. The emission line strength of He I 2059, associated with the \( 2^1P \to 2^1S \) transition, in conjunction with an assumed or measured helium abundance can be used to probe the ionizing continuum hardness of OB stars with \( T_{\text{eff}} \lesssim 40,000 \) K (Shields 1993). For example, He I 2059/Br\( \gamma \) has been used to measure an average effective blackbody temperature of young O and B stars in star-forming clouds (e.g., Geballe et al. 1984; Doyon et al. 1992; Doherty et al. 1995). However, even in the context of a full photo-ionization model, the He I 2059/Brackett recombination line ratio cannot be reliably predicted and therefore is not advisable (in isolation) for determining the effective temperature in starbursts (Lumsden et al. 2003).

Förster Schreiber et al. (2001) use helium/hydrogen and mid-IR nebular line ratios to measure stellar \( T_{\text{eff}} \approx 37,400 \) K in M82. Their methodology uses two He I/He I recombination line ratios (He I 2059/Br\( \gamma \) and He I 1.701 \( \mu m \)/Br10) to constrain the effective temperature. Despite the non-monotonic scaling of He I 2059/Br\( \gamma \) with \( T_{\text{eff}} \) (see, e.g., Shields 1993), the second He I/H I ratio allows them to break degeneracies and constrain model parameters.

We estimate \( T_{\text{eff}} \) using a method similar to the one employed by Förster Schreiber et al. (2001), whereby multiple He I and H I line ratios are compared to photoionization model predictions. Calculations were performed with version 17.00 of Cloudy, most recently described by Ferland et al. (2017). Our model assumes a spherical geometry of only hydrogen and helium (with solar He/H ratio) surrounding a central star with effective temperature in the range 30,000 K \( \leq T_{\text{eff}} \leq 40,000 \) K.

We also vary the logarithmic ionization parameter, \(-4.0 \leq \log U \leq 0\) (where \( U \) is the ratio of ionizing photon density to hydrogen density), and assume Case B recombination. Cloudy reports the ratios of He I (1.869, 1.870, and 2.059 \( \mu m \) lines) to H I (Pa\( \alpha \) and Br\( \delta \)) line fluxes for comparison to observed values. Since the He I 1869 and 1870 lines are blended, we compare the sum of modeled components with the observed doublet. Line ratio data and Cloudy results are shown in Figure 3.7.

We find that He I 2059/Br\( \delta \) is extremely sensitive to the helium and hydrogen ionization fractions, which in turn depend on the ionization parameter and spectral hardness. From Cloudy predictions, we learn that this line ratio decreases with increasing temperature at 32,000 K \( < T_{\text{eff}} < 38,000 \) K for \( \log U = -1.0 \), whereas lower ionization parameter allows the line ratio to scale with effective temperature (Figure 3.7, right panel). For \( \log U \approx -2 \), an ionization parameter similar to those of some \( z \approx 2-3 \) star-forming galaxies (see, e.g., Nakajima & Ouchi 2014; Sanders et al. 2016), we find
that He I 2059 μm/Brδ for our sample is too low compared to Cloudy predictions. An ionization parameter of log $U \approx -3$, characteristic of typical low-$z$ star-forming galaxies, still does not reconcile differences between observations and model predictions. High-SNR LBA detections such as J0150 or J1434 have He I 2059/Brδ significantly lower than even the predicted log $U = -4.0$ ratios.

In an ionized, dust-free medium, resonant pumping of the $1^1S \rightarrow 2^1P$ transition via an optically thick 584 Å line amplifies the production of He I 2059 photons. However, dust attenuation and neutral hydrogen absorption in realistic nebulae reduce fluorescent He I 2059 emission. Collisional depopulation from $2^3S \rightarrow 2^1P$ and ionization fractions can also impact the production of He I 2059 photons (see Osterbrock 1989; Shields 1993 for more detailed discussions). For our sample of LBAs, it is unlikely that the ionization parameter is below log $U = -4$. Therefore, we conclude that dust and gas opacity is responsible for attenuating the resonant transition up to the $2^1P$ state and depressing the He I 2059/Brδ ratio relative to Cloudy predictions.

In the left panel of Figure 3.7, we see that the He I 1869+1870/Pα and He I 1869+1870/Brδ ratios are better behaved and increase monotonically with effective temperature. Therefore, we primarily rely on this diagnostic to constrain $T_{eff}$. J2104, J0150, and J2116b, the strongest Pα emitters in our sample, are also characterized by elevated He I 1869+1870 emission relative to hydrogen lines. This result suggests a high fraction of massive, young stars, perhaps due to ongoing vigorous starbursts, or even contributions from AGN. Among these three LBAs, He I 1869+1870/hydrogen line ratios for J0150 are in agreement with $T_{eff} \sim 40,000$ K for log $U = -3$. J2104 and J2116b exhibit $\sim 1 \sigma$ tension with model predictions for log $U = -3$.

There is some evidence that the ionization parameters of LBGs may be up to an order of magnitude higher than those of local galaxies (e.g., Nakajima & Ouchi 2014), although others have found that high-$z$ systems have log $U \sim -3$, comparable to typical nearby star-forming galaxies (e.g., Strom et al. 2017, 2018). We can use [O III]/[O II] flux ratios (Tremonti et al. 2004) to obtain an order-of-magnitude estimate of the ionization parameter for our sample (see comparisons to, e.g., Stasińska et al. 2015); the LBAs are characterized by [O III]/[O II] $\sim 1$, which implies a lower value of log $U \sim -3$. Recent works have reported that the effective temperature of newly formed stars in higher-$z$ systems may be much higher than in the nearby Universe, with $T_{eff} \sim 50,000 - 60,000$ K (see, e.g., Steidel et al. 2014). Such extreme conditions may be applicable to J2104 and J2116b, which reside in a region of He I/H I parameter space that seems to be described by higher effective temperatures than the ones we have modeled. We repeat our comparison for He I and H I lines extracted over a larger aperture (5 pixel radius, or $\sim 1.5''$), and find that J2104 has anomalously high He I/H I line ratios as before (implying $T_{eff} > 40,000$ K), whereas J2116b becomes more consistent with $T_{eff} = 40,000$ K.
Overzier et al. (2009) find that J2104 hosts a dominant compact object (DCO), defined as a single massive ($\sim 10^9 M_\odot$) compact (radius $< 100$ pc) clump. They propose that DCOs are massive starbursts that will eventually form the cuspy light profiles found in local ellipticals. J0213 and J1434 also host DCOs, but their He I/H I ratios do not indicate enhanced $T_{\text{eff}}$ like that of J2104. However, we will see that J0213 and J2104 have relatively large warm H$_2$ masses (see Section 3.5), which are consistent with strong (but not molecule-dissociating) starbursts powering the fluorescent excitation of molecular material. The relative dearth of warm H$_2$ in J2104 may suggest the opposite: the same energetic processes driving high-velocity Pa$\alpha$ wings and likely dissociating a significant fraction of H$_2$ may be responsible for the extremely high stellar photospheric temperatures as well. Such effects may indeed be due to activity stemming from accretion onto the central supermassive black hole (despite no direct evidence of AGN found in J2104; Overzier et al. 2009).

Alternatively, a top-heavy IMF, in which the proportion of newly formed massive and high-$T_{\text{eff}}$ stars is higher than traditional models (such as Salpeter 1955; Kroupa 2001; Chabrier 2003) may account for the increased $T_{\text{eff}}$ seen in J2104 and J2116b. A top-heavy IMF in LBGs is also consistent with theoretical work supporting the formation of very massive ($> 100 M_\odot$) stars in the early Universe when few or no metals are present (see, e.g., Bromm 2013; Mebane et al. 2018). A related possibility is that J2104 hosts a significant fraction of massive star binaries, similar to stellar populations that are theorized to reside in high-z star-forming galaxies but are noticeably absent in low-z systems (see, e.g., Steidel et al. 2016). Such massive star binaries could also help explain some of the Wolf-Rayet features found in J2104’s optical spectrum (Brinchmann et al. 2008). The stellar evolution of massive binaries may include an extended ($\sim 100$ Myr) phase during which stellar $T_{\text{eff}}$ can reach as high as $\sim 100,000$ K (to the point of doubly ionizing helium; Eldridge & Stanway 2012; Eldridge et al. 2017). Our results for J2104 are consistent with both the top-heavy IMF and massive binary population interpretations of high-z galaxy evolution, as well as nuclear activity.

We caution that our findings are based on a simple photoionization model with no dust or metals, simulated for ideal Case B conditions. The He I 2059 $\mu$m line strength is notoriously difficult to predict when the physical conditions (i.e., density and geometry) are not known. However, our results agree qualitatively with the basic model of He I 1869+1870 flux density ratios with Pa$\alpha$ and Br$\delta$ and are consistent with $T_{\text{eff}} \sim 32,000 - 40,000$ K for LBAs other than J2104 and J2116b, which appear to require modestly higher $T_{\text{eff}}$ (and are consistent with 40,000 K at the $\sim 1 \sigma$ level). J0150, J2116b, and J2104 are the three strongest Pa$\alpha$ emitters in our LBA sample, yet their He I 1869 + 1870/Pa$\alpha$ flux ratios are also the three highest, demonstrating that they likely have higher average effective temperatures than other LBAs, e.g., J0213 (for which $T_{\text{eff}} \sim 32,000$ K).
\[ \text{[Fe II] recombination} \]

We at least marginally detect [Fe II] 1.810 \( \mu m \) line emission for all LBAs except J0212. [Fe II] traces low-ionization gas, since the electronic potential is only 7.9 eV. For most nearby galaxies, [Fe II] 1.257 or 1.644 \( \mu m \) emission can be used to characterize shocks driven by supernovae (Oliva et al. 1989; Alonso-Herrero et al. 2003) or AGN processes in Seyfert galaxies (see, e.g., Storchi-Bergmann et al. 2009). We detect [Fe II] 1.810 \( \mu m \) at > 3 \( \sigma \) significance for only J0150, and at \( \sim 1.5 \sigma \) for J0213, J2104, and J2116b. Due to its faint nature, we do not include it in our tables and figures.

### 3.4.3 H\(_2\) ro-vibrational spectrum

Molecular hydrogen can emit ro-vibrational lines in the NIR when heated or fluorescently excited. In the local Universe, these H\(_2\) emission lines are often attributed to shocks driven by young stars, or to photon-dominated regions (PDRs) near UV radiation (or in extreme cases, in X ray-dominated regions; XDRs). Gas can be heated by collisional excitation in both PDRs and shocks, while PDRs can also be responsible for fluorescent excitation of H\(_2\) (see, e.g., Sternberg & Dalgarno 1989). The properties of warm H\(_2\) emission are important for understanding star formation feedback and its effects on the ISM, which can impact the host galaxy’s evolution. However, H\(_2\) ro-vibrational emission lines are redshifted into the mid-infrared at high redshifts, and thus remain poorly studied for traditional samples of LBGs. Our SINFONI observations provide us with the unique opportunity to measure multiple H\(_2\) ro-vibrational lines for LBAs, and the potential to transfer this knowledge to their higher-redshift cousins.

We use the same single Gaussian fitting procedure as before to measure ro-vibrational line fluxes and width, and run 10,000-step MCMC simulations to estimate uncertainties. Extracted H\(_2\) line fluxes and line widths are shown in Tables 3.6 and 3.7. For the NIR recombination lines, we corrected for extinction using the Calzetti et al. (2000) nebular gas prescription as described in Section 3.4.1, as appropriate for star-forming clouds.

The column density of molecular hydrogen in each upper ro-vibrational state is proportional to the observed line flux, and from the distribution of lines a gas excitation temperature, \( T_{\text{ex}} \), can be inferred. Following Beckwith et al. (1978), we compute the column density of molecular hydrogen by measuring surface brightness \( I = (1/\Omega) \int S_\lambda d\lambda \) using a circular aperture with two-pixel radius, with

\[
N_{\text{H}_2} = \frac{4\pi I}{A_{ul} h\nu},
\]

where \( A_{ul} \) is the Einstein coefficient (provided by Turner et al. 1977) and \( h\nu \) is the photon energy. We divide the column density by the appropriate rotational level statistical weight and compare
with upper ro-vibrational level energies (as measured by Dabrowski 1984).

We then fit a single-temperature model to the logarithmic column density and estimate uncertainties using a MCMC routine as before. Our fit assumes that the H$_2$ gas is optically thin and adequately described by a single temperature. H$_2$ gas becomes optically thick at column density $\sim 10^{26}$ cm$^{-2}$ (Beckwith et al. 1978); our LBAs are well below this threshold ($N_{H_2} \lesssim 10^{15}$ cm$^{-2}$). Other studies have found that multiple temperatures are sometimes needed to fit the ro-vibrational spectrum (see, e.g., Krabbe et al. 2000), especially when higher energy transitions reveal the need for a hotter component. For the range of energies probed by our SINFONI observations, single-temperature fits describe the H$_2$ data well. Figure 3.8 shows the ro-vibrational lines and best-fit models assuming a 3:1 ortho-to-para H$_2$ ratio (see also Sternberg & Neufeld 1999).

Ratios between the H$_2$ 1-0 S(1) and 2-1 S(1) ro-vibrational lines are useful for distinguishing between collisionally excited and fluoresced ro-vibrational transitions (see, e.g., Black & van Dishoeck 1987; Wolfire & Konigl 1991). Although observed values of 2-1 S(1)/1-0 S(1) = 0.5 have traditionally been cited as evidence for UV fluorescence (e.g., Black & van Dishoeck 1987; Hollenbach & McKee 1989), and lower ratios of $\sim 0.1$ pointed to as evidence for shock heating (e.g., Shull & Beckwith 1982), Sternberg & Dalgarno (1989) have showed that low 2-1 S(1)/1-0 S(1) ratios are not smoking-gun indicators of shocks as previously claimed; rather, such low ratios can also be attributed to (dense) PDRs (see also, e.g., Burton et al. 1990; Hollenbach & Tielens 1997). The possibility that dense PDRs might appear thermal must be taken into account before we attempt to discriminate between PDRs and shock heating as the primary source of H$_2$ emission.

We do not detect any H$_2$ 2-1 S(·) emission lines in our spectra. H$_2$ 2-1 S(1) is always redshifted out of our spectral range, but the 2-1 S(3) line (and/or other bright v = 2 lines) should be detectable in at least some of our LBAs even if the v = 2 − 1 transitions are not probing even warmer H$_2$. Thus, our non-detection of H$_2$ 2-1 S(·) lines is consistent with both UV heating and shocks as origins for the warm H$_2$ emission.

Bright [Fe II] emission has been invoked as evidence for shocks (e.g., Moorwood & Oliva 1988; Hollenbach & McKee 1989; Oliva et al. 2001), but it requires extremely widespread shocks with high pre-shock densities. Maloney et al. (1996) find that observed [Fe II] emission may not require shocks after all, but may instead be straightforwardly produced in XDRs (e.g., due to irradiation from massive stars or AGN; Meijerink & Spaans 2005). XDR models have been able to successfully reproduce both the [Fe II] and H$_2$ ro-vibrational emission observed in nearby AGN. For our LBA sample, [Fe II] 1.810 µm emission lines are detected at $> 1.5 \sigma$ in 4/7 spectra, which lends credence

---

9Because we have used a logarithmic scale for the column density of molecular gas, $N_{mol}$, we approximate uncertainties as $\sigma_{\ln N_{mol}} \approx \sigma_{N_{mol}}/N_{mol}$. This linear approximation breaks down when $\sigma_{N_{mol}} > N_{mol}$ and may cause bad fits for low-SNR lines, but we find that our fits are qualitatively consistent with the data.
to the idea that our LBA sample’s H$_2$ lines represent X ray-heated gas emission. In Section 3.6, we will see that both [Fe II] and H$_2$ emission are faint relative to hydrogen recombination lines, supporting a scenario in which recently formed massive stars (and not, e.g., AGN) are responsible for exciting the molecular gas.

Using a sample of LBAs similar to ours (with four overlapping members), Contursi et al. (2017) use Herschel/PACS spectroscopy to measure [O I] 63 $\mu$m and [C II] 158 $\mu$m lines in order to probe physical conditions. Contursi et al. (2017) favor PDRs over shocks for all detected LBAs in their sample on the basis of low [O I] 63 $\mu$m/[C II] 158 $\mu$m ratios (Hollenbach & McKee 1989). Izotov & Thuan (2016) also come to the conclusion that, for a sample of low-metallicity blue compact dwarf galaxies, the dominant mechanism for exciting H$_2$ is fluorescence from strong UV radiation fields. For our sample of LBAs, we find that the H$_2$ ro-vibrational emission is consistent with gas heating and line pumping from UV photons (and X rays, in some cases). Our observations are not inconsistent with molecular shocks, but we also do not find strong evidence to warrant the invocation of shocks. Therefore, we favor the Contursi et al. (2017) interpretation of PDRs near star-forming regions as the primary locus of H$_2$ excitation.

### 3.5 Warm and cool molecular gas

By comparing H$_2$ ro-vibrational lines with tracers of other phases of the ISM, we can better understand the formation of warm molecular clouds in PDRs, winds, and/or outflows. For example, Tadhunter et al. (2014) find evidence of AGN-driven warm molecular outflows in the nearby Seyfert galaxy IC 5063. The post-shock gas cools and forms warm ($10^4$ K) gas, which allows for recombination cooling (e.g., through the P$_\alpha$ or Brackett lines) and subsequently the formation of molecules. Warm H$_2$ can then cool via ro-vibrational lines at temperatures of $\sim$ few $\times$ $10^3$ K. Thompson et al. (2016) present related analytic work on the acceleration, destruction, and re-formation of neutral gas clouds. According to their model, hot ($10^7 – 10^8$ K) outflows powered by starbursts can drive high-velocity gas to large distances into the dark matter halo. Observations by Oonk et al. (2010) also support a scenario of AGN-driven fast winds that later cool and condense to form warm molecular clouds in cool-core clusters.

An alternative scenario is that cold molecular gas becomes entrained and accelerated to large velocities (see, e.g., Veilleux et al. 2009; Gaspari & Sądowski 2017; Gronke & Oh 2018). Keck/OSIRIS AO observations of a nearby buried quasar, F08572+3915:NW, reveal a strong molecular outflow bright in H$_2$ ro-vibrational transitions (which also shows evidence of dust entrainment; Rupke & Veilleux 2013). The H$_2$ 1-0 S(3)/S(1) line ratio is significantly higher in the outflow than in the disk.
region, and the authors find outflowing H$_2$ $T_{\text{ex}} \approx 2370$ K comparable to our sample (Section 3.4.3). ALMA compact array observations of another ULIRG/quasar, F11119+3257, detect CO (1-0) emission extending to $\pm 1000$ km s$^{-1}$ (Veilleux et al. 2017). These authors derive cool H$_2$ outflow rates consistent with those predicted for momentum-conserving winds entraining molecular gas.

In the context of these previous studies, it is instructive to compare the masses of LBAs’ cool molecular gas, such as the $\sim 30$ K gas most easily traced by CO rotational observations, and warm gas traced by H$_2$ ro-vibrational emission. Large variations in this ratio may support scenarios in which cool molecular gas is dissociated and re-forms over longer timescales. Correlations between the cool/warm ratio and driving mechanism (e.g., modest star-formation vs. vigorous starbursts vs. AGN) are also informative for discriminating between models. If we find little variation throughout our sample, then it is possible that non-dissociative processes such as PDR heating or slow entrainment in winds are exciting the H$_2$ gas.

Scoville et al. (1982) consider thermally populated warm emitting H$_2$, and estimate its mass assuming optically thin emission. Storchi-Bergmann et al. (2009, and citations within their Section 4.8) provide the formula for determining the warm H$_2$ mass at an assumed $T_{\text{ex}} = 2000$ K using H$_2$ 1-0 S(1) flux:

$$M_{\text{H}_2} = 5.078 \times 10^{13} \ M_\odot \left( \frac{F_{\text{H}_2 \ 1-0 \ S(1)}}{\text{erg s}^{-1} \text{cm}^{-2}} \right) \left( \frac{D_L}{\text{Mpc}} \right)^2 .$$

(3.5)

From Figure 3.8 we see that $T_{\text{ex}} = 2000$ K is within 1 $\sigma$ of the best fits for most LBAs (although the H$_2$ excitation temperature is unconstrained for J0212).

In Table 3.8 we compare warm and cool H$_2$ gas masses, where cool H$_2$ masses are calculated using low-resolution radio observations of the CO(1-0) line (Contursi et al. 2017), assuming $X_{\text{CO}} = 2 \times 10^{20}$ cm$^{-2}$ (K km s$^{-1}$)$^{-1}$ following Bolatto et al. (2013). Contursi et al. (2017) detect significant CO (1-0) flux using the IRAM PdBI for four LBAs in our sample. For J0150, we have used the CO (1-0) flux obtained from CARMA observations with longer integration times (Gonçalves et al. 2014). We determine warm H$_2$ masses using Equation 3.5. For LBAs where H$_2$ 1-0 S(1) is redshifted out of the $K$-band, we compute the flux $F_{\text{H}_2 \ 1-0 \ S(1)} \approx 1.9 \times F_{\text{H}_2 \ 1-0 \ S(3)}$, where H$_2$ 1-0 S(3) is the second-brightest observed H$_2$ line and the line ratio has been derived from a $T_{\text{ex}} = 2000$ K ro-vibrational spectrum.

J0213 appears to host the largest mass fraction of warm molecular gas, even higher than the vigorously star-forming J2104 (although the two agree to within 1 $\sigma$ uncertainties). Could this indicate that a larger fraction of the molecular ISM is excited, due to either shocks or PDRs/XDRs? Both J0213 and J2104 are composite AGN/star-forming galaxies according to their emission line ratios (Overzier et al. 2009) using the Baldwin et al. (1981, ; BPT) classification. However, LBAs
(and LBGs) generally occupy a locus of BPT parameter space that is consistent with higher ionization state, even if they show no signs of any AGN. Overzier et al. (2009) investigate complementary optical emission-line diagnostics, following Veilleux & Osterbrock (1987) and Kewley et al. (2006), but they find no evidence in support of obscured or unobscured AGN in J0213 or J2104.

We compare our results with those for three prototypical low-z objects that have been well-studied in the literature: M82, a starburst galaxy; F11119+3257, a ULIRG hosting a Type 1 quasar; and IC 5063, a radio galaxy hosting a Type 2 Seyfert. Their warm and cool molecular masses are shown in Table 3.8. For each system, we report the total CO-derived cool H\(_2\) mass; although it is possible to separate a broad CO component from a rotating disk for these nearby, bright, and well-resolved systems, we show the global masses in order to facilitate comparisons with our LBA sample. Cool H\(_2\) masses have been estimated using a Galactic CO-to-H\(_2\) conversion factor, \(\alpha_{\text{CO}}\), for IC 5063 (Morganti et al. 2013). A ULIRG-like \(\alpha_{\text{CO}}\) factor is appropriately adopted for F11119+3257 (Cicone et al. 2014), and a combination of starburst and Galactic \(\alpha_{\text{CO}}\) factors are used for different gas components of M82 (Walter et al. 2002).

We find that IC 5063 best matches the LBA sample in terms of warm/cool H\(_2\) mass ratio. It is worth noting that the ionized gas (Br\(\gamma\)), warm molecular gas (H\(_2\) 1-0 S(1)), and cool molecular gas (CO(2-1)) components of IC 5063 are each kinematically distinct (Tadhunter et al. 2014), not unlike the multiphase gas components in our own sample. IC 5063 is also notable among our three comparison objects in having a morphologically early type, in which one might expect star formation and thus warm gas to be centrally concentrated. For M82, which is undergoing a galaxy-wide starburst driving a large-scale wind, and F11119+3257, in which a merger-induced starburst is likely contributing to strong H\(_2\) emission across the system, the global warm-to-cool molecular gas mass ratios may be elevated in a way that is only the case for the center of IC 5063.

If we could observe the LBAs’ CO gas at higher angular resolution, then we might be able to investigate how the local warm-to-cool molecular gas mass ratio varies spatially. Future higher-resolution studies of the cool gas may be useful for determining whether the warm/cool molecular mass ratio is elevated or diminished near the sites of vigorous star formation (e.g., at the centers of DCOs). It is conceivable that the warm-to-cool mass ratio could increase by an order of magnitude, from \(\sim 10^{-6}\) to \(\sim 10^{-5}\), and locally match that seen in M82 or F11119+3257.

### 3.6 Resolved near-infrared line ratios

We jointly fit the continuum and all recombination and ro-vibrational lines in the NIR spectra using the same methodology as in Section 3.4, except now at *every pixel* and with only a single Pa\(\alpha\)
Gaussian component. From Figure 3.9, we can compare spatial distributions of continuum, Brδ, He I, and H2 line ratios relative to Paα. Paα moment map contours are also included for visual comparison. We have masked the edge pixels in all maps, and also applied a 2σ cut in Paα flux to all line ratio maps to prevent noisy pixels from cluttering the plots.

For H I recombination, we choose to show Brδ rather than another Brackett series line because it generally has the highest SNR and best-determined fit parameters (see Figure 3.5). For the He I/Paα ratio, we fit lines using a fixed He I 1869/He I 2059 ratio assuming Case B recombination. For H2, we fit 1-0 S(1), S(2), S(3), and S(5) ro-vibrational lines assuming a $T_{ex} \to \infty$ thermally populated spectrum (i.e., we are effectively averaging all of the lines using a 3:1 ortho-to-para ratio in order to maximize the SNR), and plot a single H2 1-0 S(·)/Paα ratio.

### 3.6.1 Constraining excitation mechanisms using Paα, H2, and [Fe II] emission lines

Piqueras López et al. (2012) measure resolved Brγ, H2, and [Fe II] 1.644 μm emission for a sample of nearby LIRGs and ULIRGs. They find average flux ratios log(H2/Paα) = −1.5 for the LIRG sample, and log(H2/Paα) = −2.1 for the ULIRGs.11 For a subsample of these galaxies, Colina et al. (2015) are able to cleanly separate AGN, supernovae, and young stars as the dominant excitation mechanisms in their sample of (U)LIRGs by comparing resolved [Fe II]/Brγ and H2/Brγ line ratios. Emission lines associated with massive, young stars uniquely populate a locus of low H2/Brγ and low Fe II/Brγ line ratios, i.e., $-2.3 \leq \log(H2/Paα) \leq -1.2$ and $-2.2 \leq \log([Fe II] 1.810 \mu m/Paα) \leq -1.4$.

For our LBA sample, [Fe II] 1.810 μm emission is faint and generally not detected aside from near the Paα peaks, so we do not show it in Figure 3.9. We measure positive Fe II 1.810 μm emission at > 1.5σ significance for J0150, J0213, J2104, and J2116b. For the central regions of these systems, we find log([Fe II] 1.810 μm/Paα) $\sim -1.7$. Despite the large uncertainties, we find that [Fe II] is too low relative to Paα for supernovae or AGN to be its primary excitation mechanism (e.g., using the line ratios described by Colina et al. 2015, which are also similar to the results from Riffel et al. 2013). Our well-detected H2/Paα line ratios reinforce this scenario: overall, we find that log(H2/Paα) $\lesssim -1.5$, which is in the locus for young, star-forming regions.

---

10We attempted to fit the broad Gaussian components as well, but many pixels have low SNR for the Paα line, which then biases any line ratios that depend on it. Thus, we continue using only a single-component Paα fit. We also attempted to fit spectra extracted from $2 \times 2$ and $3 \times 3$ pixel apertures, but the results were no better than single-pixel spectral fitting.

11We assume Case B recombination in order to convert between Brγ and Paα ($T_e = 10^4$ K and $n_e = 10^{3}$ cm$^{-3}$; Hummer & Storey 1987). We also convert [Fe II] 1.644 μm to [Fe II] 1.810 μm using a constant branching ratio of 0.205 (which has no temperature dependence because both transitions have a common upper state; Verner et al. 1999; Ferland et al. 2017).
3.6.2 Radial gradients in line ratios

For J2104 (and to an extent, the other LBAs), H$_2$/Pa$_\alpha$ line ratios are depressed toward the Pa$_\alpha$ emission peak, relative to their values farther from the peak. We would naively expect Pa$_\alpha$ and H$_2$/Pa$_\alpha$ to be anticorrelated, but there is an additional more subtle effect at play here. Specifically, we have only fit single Gaussian profiles to both lines, which may adequately describe the fainter H$_2$ emission (whose broad wings are low in flux, and therefore do not matter much to the model) while failing to fit the brighter Pa$_\alpha$ line (whose broad wings force strong constraints on the Gaussian model parameters), such that the single-component Pa$_\alpha$ flux is overestimated. Precisely for this reason, we have used double Gaussian models to measure the total Pa$_\alpha$ flux in our previous analysis (e.g., in Section 3.4.2). We also considered fitting the H$_2$ lines pixel-by-pixel using two-component models, but discovered that the other lines are too faint/noisy to be adequately fit with multiple Gaussians.

Given these difficulties, we avoid comparing line flux ratios for the centers of LBAs, where broad ionized gas components contribute the most to Pa$_\alpha$ flux. We thus interpret the line ratios shown in Figure 3.9 conservatively. For example, there is an apparent dip in H$_2$/Pa$_\alpha$ toward the center of J2104, but we attribute this to the broad, Pa$_\alpha$-bright wind that we have reported in Section 3.4.2. We also find that Br$_\delta$/Pa$_\alpha$ and He I/Pa$_\alpha$ exhibit similar depressions toward the LBAs’ central regions, although the effect seems to be less pronounced.

A high H$_2$/Pa$_\alpha$ ratio for peripheral regions of J2104 may indicate that heated molecular gas is prevalent in its outskirts. It is possible that the ISM conditions are similar to those of the spatially offset H$_2$/Pa$_\alpha$ peak in J1434, which we discuss in detail in Section 3.6.3. For the other LBAs, H$_2$ emission is detected throughout the star-forming, Pa$_\alpha$-bright systems, but we find no gradient in the H$_2$/Pa$_\alpha$ line ratio once we account for broad-line ionized gas components.

3.6.3 H$_2$ spatially offset from ionized gas for J1434

The distribution of H$_2$ 1-0 ro-vibrational emission in J1434 is concentrated toward the northwest source where the Pa$_\alpha$ line is weak. Near the location of the northwest H$_2$/Pa$_\alpha$ peak, we also find a stellar continuum peak. For the more luminous central Pa$_\alpha$ source in J1434, the H$_2$ lines are only weakly detected, implying that its higher rate of star formation may have dissociated H$_2$. We find log(H$_2$/Pa$_\alpha$) $\sim$ –2 for the Pa$_\alpha$-bright nucleus, and log(H$_2$/Pa$_\alpha$) $\sim$ –1 for the companion system. It is possible that the bright Pa$_\alpha$-emitting source is powered by intense young star formation activity, while its neighbor to the northwest is affected by AGN or supernova feedback. Unfortunately,
$H_2/Pa\alpha$ cannot uniquely identify whether massive stars, supernovae, or AGN dominates the excitation. Colina et al. (2015) use $[\text{Fe} \, \text{II}]$ emission to break this degeneracy for their sample of local LIRGs and Seyferts, but we are unable to detect the $[\text{Fe} \, \text{II}]$ line at high significance for either component in J1434. We place 3 $\sigma$ upper limits of $\log([\text{Fe} \, \text{II}] \, 1.810 \, \mu\text{m}/Pa\alpha) < -1.4$ for the main system, and $< -0.8$ for the companion. These line ratios imply that the Pa$\alpha$-bright system is mostly excited by young stars; for the $H_2$-bright companion, low-SNR line ratios are consistent with all excitation mechanisms.

We fit Gaussians to the Pa$\alpha$ and $H_2$ lines for both spatial components, and show the smoothed profiles in Figure 3.10. We find that, although the $H_2$ 1-0 S(3) flux toward the companion is higher $((4.5 \pm 1.5) \times 10^{-17} \, \text{erg} \, s^{-1} \, \text{cm}^{-2} \, \mu\text{m}^{-1})$ than toward the main system $((2.8 \pm 1.7) \times 10^{-17} \, \text{erg} \, s^{-1} \, \text{cm}^{-2} \, \mu\text{m}^{-1})$, the Pa$\alpha$ emission from the northwest companion is lower. The $H_2$ line width is similar to that of the ionized gas for the companion, suggesting that this source is accelerating warm molecular gas to velocities comparable to those of the ionized gas. However, the same is not true for the main component, where we find the existence of a broader Pa$\alpha$ component line width ($245 \pm 23 \, \text{km} \, s^{-1}$) not seen in $H_2$ ($154 \pm 49 \, \text{km} \, s^{-1}$; see Tables 3.5 and 3.7). It appears that molecular hydrogen is not being accelerated or is being dissociated by the energetic activity that is responsible for driving Pa$\alpha$ to high velocities.

3.6.4 Spatial variations in $H_2$ temperature

We have tried mapping the $H_2$ 1-0 S(3)/S(1) ratio for various LBAs; however, the SNRs are too low for the excitation temperature to be constrained in individual pixels. Indeed, from Figure 3.8, we find that uncertainties for the LBAs with even the highest ro-vibrational line SNRs result in $\sim 10\%$ systematic uncertainty for fitted $T_{\text{ex}}$. Those spectra were extracted over approximately one PSF (two-pixel radius), meaning that the expected scatter per average individual pixel in the brightest regions will yield excitation temperature fits with $\gtrsim 50\%$ scatter.

However, the dual $H_2$-peaked nature of J1434 allows for comparison of excitation temperatures between the Pa$\alpha$-bright main component and its $H_2$-bright, Pa$\alpha$-deficient companion. The two brightest available ro-vibrational lines are S(3) and S(5), so we translate the two fluxes to a $T_{\text{ex}}$ in accordance with the methods introduced in Section 3.4.3. For the companion we find $T_{\text{ex}} = 1390^{+1190}_{-490} \, \text{K}$, compared with $1510^{+1160}_{-1020} \, \text{K}$ for the main J1434 source. The warm $H_2$ temperatures are in good agreement with each other. Although subject to large uncertainties, $T_{\text{ex}}$ fits are also consistent with those in other LBAs (fit using their entire ro-vibrational spectra), as well as the fiducial 2000 K value that we have assumed in previous sections. Therefore, we find no significant difference in excitation temperatures between the main and companion components of J1434.
3.7 Conclusions

We have obtained \( \sim 0.7'' \) resolution integral field spectroscopy for a sample of six UV-bright galaxies at \( z = 0.1 - 0.3 \) selected from the Heckman et al. (2005) SDSS sample, which resemble high-redshift Lyman break galaxies in terms of SFRs, stellar masses, metallicities, and rest-frame UV sizes. We therefore consider our sample to be low-redshift Lyman break galaxy analogs (LBAs). One LBA system, J2116, consists of two bright sources with rotating gas disks. Our conclusions are as follows:

1. Using the Pa\( \alpha \) line, we have traced the kinematics of the ionized gas medium and produced moment maps. We measure inclination-corrected ordered velocity-to-dispersion ratios \( \langle v/\sigma \rangle \approx 1.8 - 2.2 \) from moment maps, and \( \sim 1.2 \) by kinematically modeling the 3d data cube, finding values that are much lower than typical \( v/\sigma \) for low-\( z \) star-forming spirals. Mean velocity dispersions in the LBA sample are very high: \( \langle \sigma \rangle = 75 - 90 \text{ km s}^{-1} \), depending on how they are measured.

2. We use H\( \alpha \), H\( \beta \), and Pa\( \alpha \) line ratios to measure the nebular color excess under the assumption of a Calzetti attenuation curve. For our sample, we derived nebular color excesses \( \langle E(B-V)_{\text{neb}} \rangle = 0.26 \pm 0.04 \), \( 0.32 \pm 0.11 \), and \( 0.35 \pm 0.17 \), using the H\( \alpha \) + H\( \beta \), Pa\( \alpha \) + H\( \beta \), and Pa\( \alpha \) + H\( \alpha \) line ratios, respectively.

3. We find that the average continuum-to-nebular color excess ratio changes as a function of the optical depth (or recombination line) used to probe H\( \Pi \) regions: \( f_{\text{H}\alpha + H\beta} = 0.82 \pm 0.19 \), whereas \( f_{\text{Pa}\alpha + H\beta} = 0.62 \pm 0.17 \) and \( f_{\text{Pa}\alpha + H\alpha} = 0.54 \pm 0.19 \). The observed attenuation is consistent with a well-mixed model of ionizing and non-ionizing stars (e.g., closer to a single effective screen) when we use only the bluer Balmer lines. When we use the Pa\( \alpha \) line, the color excess ratio is more similar to a lower, Calzetti-like \( f = 0.44 \), indicating an extra layer of attenuation toward nebular regions.

4. We measure over a central two-pixel (~0.6'') radius circular aperture the narrow and broad velocity components of Pa\( \alpha \), which is detected at high SNR for all LBAs in our sample, and well-fit by a double Gaussian model in most cases. For J0213, J1434, and J2104, we find ionized gas components with different dynamics from their rotating disks. For J2104, the broad (FWHM = 700 \( \pm \) 172 km s\(^{-1} \)) Pa\( \alpha \) velocity centroid is blueshifted relative to the main component by nearly 300 km s\(^{-1} \). We interpret this secondary component as a powerful ionized gas outflow.

5. We detect Brackett series hydrogen recombination and He I helium recombination lines at 1.869+1.870 \( \mu \text{m} \) (blended) and 2.059 \( \mu \text{m} \). We model the effective blackbody temperature of
an ensemble of young stars and the ionization parameter $\log U$ and compare with ratios of the Pa$\alpha$, Br$\delta$, He I 1869+1870, and He I 2059 lines. We find that most objects in our sample are consistent with $32,000 \, \text{K} < T_{\text{eff}} < 40,000 \, \text{K}$ and $\log U = -3$, with the three strongest Pa$\alpha$ emitters also characterized by the highest average effective blackbody temperatures. J2104 appears to require a higher $T_{\text{eff}}$, which, in conjunction with its broad Pa$\alpha$ line width and relatively low fraction of molecular gas mass in the warm phase, points to the possibility of an AGN or an overabundance of very massive stars (i.e., a top-heavy IMF).

6. We detect H$_2$ 1-0 S(1) through S(7) ro-vibrational lines and compare the velocity profiles to those of the ionized gas. We apply a Calzetti nebular-like rather than stellar-like attenuation correction because the distribution of H$_2$ emission spatially correlates more strongly with the recombination lines than with the stellar continuum. We measure rotational excitation temperatures in the warm molecular gas: $T_{\text{ex}} \sim 1600 - 2500 \, \text{K}$ for all LBAs aside from J0212. The $T_{\text{ex}} \sim 2000 \, \text{K}$ single-temperature H$_2$ gas reservoirs, non-detection of H$_2$ 2-1 S($\cdot$) lines, presence of [Fe II] emission in some cases, and low [O I]63 $\mu$m/[C II]158 $\mu$m line ratios in others (Contursi et al. 2017) together suggest that H$_2$ ro-vibrational emission originates from UV- and X ray-irradiated (dense) PDRs rather than shocks.

7. We use H$_2$ 1-0 S(1) (or S(3), when S(1) is unavailable) line flux and the assumption of $T_{\text{ex}} \sim 2000 \, \text{K}$ thermally populated ro-vibrational levels to estimate the mass of the warm molecular ISM. The warm molecular masses (near the central 1″ of peak Pa$\alpha$ emission) range from $(1.9 \pm 38) \times 10^3 \, M_\odot$, where J2104 and J0213 have the highest warm gas masses of our sample. Taking into account the two components of warm H$_2$ for J1434, our LBA sample’s global average rises to $(8.0 \pm 1.1) \times 10^{-7}$. We estimate the cool H$_2$ mass using previously measured CO (1-0) fluxes (Gonçalves et al. 2014; Contursi et al. 2017) and a Galactic $X_{\text{CO}} = 2 \times 10^{20} \, \text{cm}^{-2} \, (\text{K} \, \text{km} \, \text{s}^{-1})^{-1}$, and find global warm-to-cool molecular mass gas ratios between $(6.2 \sim 15.1) \times 10^{-7}$. The warm-to-cool molecular mass ratio for J0213 is marginally consistent with that of a nearby Type 2 Seyfert, IC 5063, but the sample’s median and mean ratios are significantly lower than those found in M82 and a nearby Type 1 quasar/ULIRG.

8. We fit the spectra at every pixel for each LBA, allowing for investigation of the spatially resolved distributions of stellar continuum, ionized gas (using helium and multiple hydrogen lines), and warm molecular gas. The H$_2$ ro-vibrational lines are fit with $T_{\text{ex}} \rightarrow \infty$, chosen to maximize SNR. We find that the H$_2$/Pa$\alpha$ ratio appears to be depressed toward the centers of some LBAs, but this trend can be attributed to Pa$\alpha$ emission associated with the broad ionized component. We measure $\log (\text{H}_2/\text{Pa}\alpha) \lesssim -1.5$ for most regions in the sample, which
is consistent with massive, recently formed stars as the primary excitation mechanism of the nebular gas.

9. For J1434, the H$_2$ lines are stronger for the northwest companion source, which has its own peak in continuum and Pa$\alpha$ emission, than for the main system. We propose that (merger-driven?) star formation is responsible for exciting warm molecular gas in the companion. It is also possible that nuclear activity in J1434’s center may be dissociating molecular gas, which could explain its low warm-to-cool H$_2$ mass ratio. The mass ratio of warm to cool H$_2$ for the combined system is \( \sim 7 \times 10^{-7} \), which is closer to the ratio in the main component than to the mass ratios found for the other LBAs.

These results blaze a trail for future rest-NIR observations of \( z \sim 3 \) LBGs in the *James Webb Space Telescope* era, which will allow us to assess how far we can extend the analogy with LBAs in terms of the latter population’s diverse but overlapping properties.

**Acknowledgments**

The authors thank Natascha Förster Schreiber and Sabine Mengel for their help preparing the SINFONI observations and the staff of the ESO Paranal Observatory for carrying them out. The authors also thank Whitney Kropat and Kelsey Funkhouser for their contributions to the early stages of this analysis, and Jane Rigby and Amiel Sternberg for useful discussions. JFW and AJB acknowledge significant support for this work from the U.S. National Science Foundation through grant AST-0955810.

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is www.sdss.org.

SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State
University, New York University, University of Notre Dame, Observatório Nacional / MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.
Table 3.1. VLT/SINFONI observations

<table>
<thead>
<tr>
<th>(1) Object</th>
<th>(2) Right Ascension</th>
<th>(3) Declination</th>
<th>(4) z</th>
<th>(5) log($M_*/M_\odot$)</th>
<th>(6) 12+log(O/H)</th>
<th>(7) log(SFR/$M_\odot$ yr$^{-1}$)</th>
<th>(8) log(sSFR/Gyr$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>015028 (J0150)</td>
<td>01$^{h}$50$^{m}$28$^{s}$39</td>
<td>13$^{\circ}$08'58''4</td>
<td>0.147</td>
<td>10.3</td>
<td>8.39</td>
<td>1.71</td>
<td>0.41</td>
</tr>
<tr>
<td>021225 (J0212)</td>
<td>02$^{h}$12$^{m}$25$^{s}$82</td>
<td>-08$^{\circ}$01'22''8</td>
<td>0.114</td>
<td>9.9</td>
<td>8.44</td>
<td>1.15</td>
<td>0.25</td>
</tr>
<tr>
<td>021348 (J0213)</td>
<td>02$^{h}$13$^{m}$48$^{s}$53</td>
<td>12$^{\circ}$59'51''5</td>
<td>0.219</td>
<td>10.5</td>
<td>8.74</td>
<td>1.55</td>
<td>0.05</td>
</tr>
<tr>
<td>143417 (J1434)</td>
<td>14$^{h}$34'17'15</td>
<td>02$^{\circ}$07'42''5</td>
<td>0.180</td>
<td>10.7</td>
<td>8.65</td>
<td>1.30</td>
<td>-0.40</td>
</tr>
<tr>
<td>210358 (J2104)</td>
<td>21$^{h}$03'58''74</td>
<td>-07$^{\circ}$28'02''4</td>
<td>0.137</td>
<td>10.9</td>
<td>8.70</td>
<td>2.03</td>
<td>0.14</td>
</tr>
<tr>
<td>211531 (J2116)</td>
<td>21$^{h}$15'31''01</td>
<td>-07$^{\circ}$27'39''1</td>
<td>0.138</td>
<td>10.7</td>
<td>...</td>
<td>1.26</td>
<td>-0.44</td>
</tr>
</tbody>
</table>

Note. — Ancillary data from Heckman et al. (2005) and Overzier et al. (2009). Columns are (1) object name, (2) Right Ascension, (3) Declination, (4) spectroscopic redshift, (5) stellar mass, (6) nebular-phase metallicity, (7) SFR from Hα and 24 µm flux, and (8) specific SFR $\equiv$ SFR/$M_\ast$. The median SFR for our sample is 28 $M_\odot$ yr$^{-1}$, and the median sSFR is 1.2 Gyr$^{-1}$. 
Table 3.2. Ordered velocities and dispersions

<table>
<thead>
<tr>
<th>LBA</th>
<th>4σ clip</th>
<th>No clip</th>
<th>Modeled</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$v$ [km s$^{-1}$]</td>
<td>$\sigma$ [km s$^{-1}$]</td>
<td>$v/\sigma$</td>
</tr>
<tr>
<td>J0150</td>
<td>161.3</td>
<td>75.2</td>
<td>2.14</td>
</tr>
<tr>
<td>J0212</td>
<td>63.6</td>
<td>36.0</td>
<td>1.77</td>
</tr>
<tr>
<td>J0213</td>
<td>35.4</td>
<td>49.9</td>
<td>0.71</td>
</tr>
<tr>
<td>J1434†</td>
<td>182.5</td>
<td>56.5</td>
<td>3.23</td>
</tr>
<tr>
<td>J2104</td>
<td>251.4</td>
<td>177.9</td>
<td>1.41</td>
</tr>
<tr>
<td>J2116a</td>
<td>56.8</td>
<td>43.1</td>
<td>1.32</td>
</tr>
<tr>
<td>J2116b</td>
<td>128.5</td>
<td>94.3</td>
<td>1.36</td>
</tr>
</tbody>
</table>

Note. — Kinematics measured from 4σ clipped and unclipped Paα moment maps (uncorrected for inclination). Image segmentation from the photutils package has been applied to isolate emission from individual components. Kinematics are also inferred from GalPaK3D modeling, where inclination, PSF, and instrumental broadening effects have been taken into account.

† Kinematic information shown for the full blended system. For only the main central component, isolated using a manual mask applied to the clipped moment maps, we find $v = 99.5$ km s$^{-1}$, $\sigma = 57.2$ km s$^{-1}$, and $v/\sigma = 1.74$. 
and weighted by the inverse variance, respectively.


\[ \text{P}\alpha \quad \text{H}\beta \]

recombination line. Uncertainties have been determined through sampling using MCMC. For

\[ \text{J0150} \quad 0.144 \quad 0.251 \pm 0.003 \quad 0.331 \pm 0.006 \quad 0.366 \pm 0.006 \quad 0.43 \pm 0.15 \]
\[ \text{J0212} \quad 0.183 \quad 0.182 \pm 0.003 \quad 0.132 \pm 0.006 \quad 0.111 \pm 0.005 \quad 1.39 \pm 0.38 \]
\[ \text{J0213} \quad 0.293 \quad 0.271 \pm 0.012 \quad 0.521 \pm 0.038 \quad 0.630 \pm 0.040 \quad 0.56 \pm 0.10 \]
\[ \text{J1434} \quad 0.243 \quad 0.273 \pm 0.006 \quad 0.299 \pm 0.011 \quad 0.312 \pm 0.010 \quad 0.81 \pm 0.17 \]
\[ \text{J2104} \quad 0.201 \quad 0.303 \pm 0.004 \quad 0.317 \pm 0.009 \quad 0.325 \pm 0.009 \quad 0.63 \pm 0.16 \]
\[ \text{J2116} \quad 0.203 \quad \cdots \quad 0.295 \pm 0.015 \quad \cdots \quad 0.69 \pm 0.17 \]

Average \quad 0.211 \quad 0.256 \pm 0.041 \quad 0.316 \pm 0.113 \quad 0.349 \pm 0.166 \quad 0.62 \pm 0.17

Note. — \( E(B-V)_{\text{cont}} \) is computed from UV continuum attenuation (Salim et al. 2005; Heckman et al. 2005) under the assumption of a Calzetti et al. (2000) attenuation curve with \( R_V = 4.05 \), and uniform uncertainty of 0.049. \( E(B-V)_{\text{neb}} \) is computed from pairs of recombination lines, where \( \text{H}\alpha \) and \( \text{H}\beta \) line fluxes and uncertainties are obtained from the Tremonti et al. (2004) GalSpecLine catalog using the SDSS DR14 SkyServer, and \( \text{Pa}\alpha \) fluxes are fit to SINFONI spectra extracted from 3\arcsec apertures to match SDSS fibers. We also show the ratio of continuum-to-nebular (\( \text{Pa}\alpha + \text{H}\beta \)) color excess for each LBA. Averages and standard deviations for the LBA sample are also included, where the \( E(B-V)_{\text{neb}} \) and \( f \) averages are unweighted and weighted by the inverse variance, respectively.

Table 3.4. Recombination line fluxes

<table>
<thead>
<tr>
<th>LBA</th>
<th>( \text{Pa}\alpha )</th>
<th>( \text{Br}\delta )</th>
<th>( \text{Br}\epsilon )</th>
<th>( \text{Br}10 )</th>
<th>( \text{He I 2059} )</th>
<th>( \text{He I 1869} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>J0150</td>
<td>9.53 \pm 0.98</td>
<td>0.56 \pm 0.05</td>
<td>0.28 \pm 0.05</td>
<td>0.22 \pm 0.04</td>
<td>0.52 \pm 0.05</td>
<td>0.54 \pm 0.05</td>
</tr>
<tr>
<td>J0212</td>
<td>1.43 \pm 0.07</td>
<td>0.04 \pm 0.06</td>
<td>0.04 \pm 0.08</td>
<td>0.00 \pm 0.23</td>
<td>0.03 \pm 0.05</td>
<td>0.06 \pm 0.04</td>
</tr>
<tr>
<td>J0213</td>
<td>4.01 \pm 0.64</td>
<td>0.25 \pm 0.07</td>
<td>0.03 \pm 0.09</td>
<td>0.02 \pm 0.08</td>
<td>\cdots</td>
<td>0.13 \pm 0.07</td>
</tr>
<tr>
<td>J1434</td>
<td>3.34 \pm 0.57</td>
<td>0.15 \pm 0.04</td>
<td>0.08 \pm 0.04</td>
<td>0.06 \pm 0.03</td>
<td>0.14 \pm 0.05</td>
<td>0.14 \pm 0.05</td>
</tr>
<tr>
<td>J2104</td>
<td>24.07 \pm 1.57</td>
<td>1.03 \pm 0.28</td>
<td>0.66 \pm 0.25</td>
<td>0.40 \pm 0.22</td>
<td>1.42 \pm 0.29</td>
<td>1.96 \pm 0.29</td>
</tr>
<tr>
<td>J2116a</td>
<td>2.05 \pm 0.11</td>
<td>0.07 \pm 0.06</td>
<td>0.03 \pm 0.06</td>
<td>0.05 \pm 0.05</td>
<td>0.13 \pm 0.07</td>
<td>0.06 \pm 0.06</td>
</tr>
<tr>
<td>J2116b</td>
<td>10.29 \pm 1.21</td>
<td>0.30 \pm 0.11</td>
<td>0.14 \pm 0.10</td>
<td>0.21 \pm 0.11</td>
<td>0.43 \pm 0.13</td>
<td>0.56 \pm 0.14</td>
</tr>
</tbody>
</table>

Note. — Dust-corrected flux (10\(^{-15}\) erg s\(^{-1}\) cm\(^{-2}\)) fit using Gaussian models for each recombination line. Uncertainties have been determined through sampling using MCMC. For cases in which the central wavelengths are shifted out of the LBA spectrum, no data are shown (\( \cdots \)).

\(^a\)\( \text{Pa}\alpha \) line fit using one or two summed Gaussian components.

\(^b\)\( \text{He I 1869+1870} \) blended lines fit using two Gaussians with fixed flux ratio, relative velocity, and equal line width.
Table 3.5. Recombination line widths

<table>
<thead>
<tr>
<th>LBA</th>
<th>Pa(\alpha) (1)</th>
<th></th>
<th></th>
<th>Pa(\alpha) (2)</th>
<th></th>
<th></th>
<th>Br(\delta)</th>
<th></th>
<th></th>
<th>Br(\epsilon)</th>
<th></th>
<th></th>
<th>Br(\iota)</th>
<th></th>
<th></th>
<th>He I 2059</th>
<th></th>
<th></th>
<th>He I 1869</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Flux</td>
<td>Center</td>
<td>Width</td>
<td>Flux</td>
<td>Center</td>
<td>Width</td>
<td>237 ± 22</td>
<td>212 ± 30</td>
<td>221 ± 33</td>
<td>220 ± 24</td>
<td>278 ± 27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J0150</td>
<td>8.0 ± 0.7</td>
<td>6</td>
<td>255 ± 6</td>
<td>1.6 ± 0.7</td>
<td>−30</td>
<td>178 ± 15</td>
<td>98 ± 52</td>
<td>104 ± 57</td>
<td>72 ± 59</td>
<td>111 ± 52</td>
<td>120 ± 50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J0212</td>
<td>1.4 ± 0.1</td>
<td>−8</td>
<td>131 ± 5</td>
<td>⋯</td>
<td>⋯</td>
<td>⋯</td>
<td>123 ± 28</td>
<td>119 ± 52</td>
<td>129 ± 52</td>
<td>⋯</td>
<td>139 ± 45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J0213</td>
<td>1.8 ± 0.5</td>
<td>2</td>
<td>123 ± 16</td>
<td>2.2 ± 0.4</td>
<td>−3</td>
<td>312 ± 46</td>
<td>166 ± 34</td>
<td>154 ± 41</td>
<td>161 ± 45</td>
<td>185 ± 48</td>
<td>188 ± 39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J1434</td>
<td>2.0 ± 0.4</td>
<td>34</td>
<td>249 ± 21</td>
<td>1.3 ± 0.4</td>
<td>12</td>
<td>130 ± 13</td>
<td>214 ± 45</td>
<td>210 ± 50</td>
<td>233 ± 49</td>
<td>431 ± 42</td>
<td>483 ± 42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J2104</td>
<td>19.8 ± 1.2</td>
<td>39</td>
<td>391 ± 13</td>
<td>4.3 ± 1.0</td>
<td>−235</td>
<td>700 ± 172</td>
<td>414 ± 45</td>
<td>421 ± 50</td>
<td>423 ± 49</td>
<td>431 ± 42</td>
<td>483 ± 42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J2116a</td>
<td>2.0 ± 0.1</td>
<td>8</td>
<td>169 ± 6</td>
<td>⋯</td>
<td>⋯</td>
<td>⋯</td>
<td>180 ± 47</td>
<td>171 ± 48</td>
<td>190 ± 49</td>
<td>192 ± 47</td>
<td>184 ± 45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J2116b</td>
<td>5.9 ± 0.8</td>
<td>−21</td>
<td>179 ± 12</td>
<td>4.4 ± 0.8</td>
<td>138</td>
<td>188 ± 17</td>
<td>196 ± 42</td>
<td>186 ± 45</td>
<td>206 ± 43</td>
<td>196 ± 38</td>
<td>254 ± 41</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. — FWHM line widths in km s\(^{-1}\) for each recombination line. For Pa\(\alpha\), the dust-corrected flux (10\(^{−15}\) erg s\(^{-1}\) cm\(^{-2}\)), central velocity (km s\(^{-1}\) from the systemic velocity), and FWHM (km s\(^{-1}\)) are shown for each component. For lines that are shifted out of the LBA spectra, or if no second Pa\(\alpha\) component is fit, we report no data (⋯).
Table 3.6. Ro-vibrational line fluxes

<table>
<thead>
<tr>
<th>LBA</th>
<th>S(1)</th>
<th>S(2)</th>
<th>S(3)</th>
<th>S(4)</th>
<th>S(5)</th>
<th>S(6)</th>
<th>S(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>J0150</td>
<td>29.3 ± 5.3</td>
<td>11.3 ± 4.4</td>
<td>31.9 ± 4.9</td>
<td>4.5 ± 3.6</td>
<td>13.1 ± 4.2</td>
<td>3.3 ± 3.4</td>
<td>−35.5 ± 6.6</td>
</tr>
<tr>
<td>J0212</td>
<td>0.8 ± 8.5</td>
<td>1.2 ± 3.6</td>
<td>1.1 ± 3.7</td>
<td>0.2 ± 17.7</td>
<td>1.1 ± 4.8</td>
<td>2.3 ± 24.4</td>
<td>4.3 ± 7.0</td>
</tr>
<tr>
<td>J0213</td>
<td>···</td>
<td>···</td>
<td>24.2 ± 9.0</td>
<td>3.7 ± 5.8</td>
<td>10.2 ± 5.7</td>
<td>6.6 ± 8.8</td>
<td>5.7 ± 5.4</td>
</tr>
<tr>
<td>J1434</td>
<td>···</td>
<td>5.0 ± 3.5</td>
<td>5.0 ± 3.7</td>
<td>1.9 ± 5.0</td>
<td>1.8 ± 3.3</td>
<td>−1.6 ± 5.8</td>
<td>1.7 ± 3.0</td>
</tr>
<tr>
<td>J2104</td>
<td>177.7 ± 27.9</td>
<td>66.7 ± 27.0</td>
<td>168.6 ± 29.2</td>
<td>45.9 ± 24.7</td>
<td>86.4 ± 24.9</td>
<td>19.3 ± 23.6</td>
<td>39.5 ± 23.8</td>
</tr>
<tr>
<td>J2116a</td>
<td>8.3 ± 7.2</td>
<td>5.1 ± 6.9</td>
<td>11.4 ± 6.5</td>
<td>−1.8 ± 5.9</td>
<td>5.5 ± 5.8</td>
<td>−0.4 ± 5.7</td>
<td>1.3 ± 5.1</td>
</tr>
<tr>
<td>J2116b</td>
<td>41.8 ± 13.7</td>
<td>19.5 ± 11.6</td>
<td>33.5 ± 12.3</td>
<td>5.7 ± 9.7</td>
<td>11.5 ± 9.4</td>
<td>6.0 ± 9.8</td>
<td>9.5 ± 9.5</td>
</tr>
</tbody>
</table>

Note. — Dust-corrected flux ($10^{-17}$ erg s$^{-1}$ cm$^{-2}$) fit using a single Gaussian model for each H$_2$ 1 − 0 ro-vibrational line. Uncertainties have been determined through sampling using MCMC. No data are shown (···) if a line is shifted out of the LBA spectrum.
Table 3.7. Ro-vibrational line widths

<table>
<thead>
<tr>
<th>LBA</th>
<th>S(1)</th>
<th>S(2)</th>
<th>S(3)</th>
<th>S(4)</th>
<th>S(5)</th>
<th>S(6)</th>
<th>S(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>J0150</td>
<td>235 ± 31</td>
<td>189 ± 41</td>
<td>243 ± 33</td>
<td>184 ± 49</td>
<td>218 ± 42</td>
<td>183 ± 48</td>
<td>?</td>
</tr>
<tr>
<td>J0212</td>
<td>98 ± 54</td>
<td>99 ± 49</td>
<td>95 ± 51</td>
<td>94 ± 60</td>
<td>88 ± 55</td>
<td>83 ± 59</td>
<td>120 ± 56</td>
</tr>
<tr>
<td>J0213</td>
<td>...</td>
<td>...</td>
<td>188 ± 50</td>
<td>131 ± 52</td>
<td>124 ± 42</td>
<td>131 ± 53</td>
<td>143 ± 47</td>
</tr>
<tr>
<td>J1434</td>
<td>...</td>
<td>136 ± 56</td>
<td>154 ± 49</td>
<td>146 ± 53</td>
<td>147 ± 48</td>
<td>?</td>
<td>143 ± 47</td>
</tr>
<tr>
<td>J2104</td>
<td>407 ± 45</td>
<td>428 ± 45</td>
<td>422 ± 45</td>
<td>426 ± 49</td>
<td>423 ± 46</td>
<td>423 ± 50</td>
<td>426 ± 49</td>
</tr>
<tr>
<td>J2116a</td>
<td>181 ± 48</td>
<td>186 ± 51</td>
<td>188 ± 47</td>
<td>?</td>
<td>188 ± 48</td>
<td>?</td>
<td>181 ± 51</td>
</tr>
<tr>
<td>J2116b</td>
<td>204 ± 43</td>
<td>193 ± 49</td>
<td>200 ± 44</td>
<td>190 ± 49</td>
<td>189 ± 47</td>
<td>184 ± 51</td>
<td>186 ± 50</td>
</tr>
</tbody>
</table>

Note. — H$_2$ 1 – 0 ro-vibrational line widths (km s$^{-1}$). No data are shown (⋯) if a line is shifted out of the LBA spectrum, and a question mark (?) is shown if the best-fit flux is negative.

Table 3.8. Warm and cool H$_2$ masses

<table>
<thead>
<tr>
<th>Object</th>
<th>$M_{H_2}^{\text{warm}}$ [$10^3$ $M_\odot$]</th>
<th>$M_{H_2}^{\text{cool}}$ [$10^{10}$ $M_\odot$]</th>
<th>$M_{H_2}^{\text{warm}}$/$M_{H_2}^{\text{cool}}$ [$10^{-7}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>J0150</td>
<td>7.2 ± 1.3</td>
<td>1.2 ± 0.1$^b$</td>
<td>6.2 ± 1.7</td>
</tr>
<tr>
<td>J0212</td>
<td>⋯</td>
<td>⋯</td>
<td>⋯</td>
</tr>
<tr>
<td>J0213</td>
<td>27.5 ± 10.2$^a$</td>
<td>1.9 ± 0.2$^e$</td>
<td>15.1 ± 5.9</td>
</tr>
<tr>
<td>J1434</td>
<td>3.7 ± 2.7$^a$</td>
<td>1.3 ± 0.4$^e$</td>
<td>2.8 ± 2.3</td>
</tr>
<tr>
<td>J1434 (total)</td>
<td>9.4 ± 3.0$^a$</td>
<td>1.3 ± 0.4$^e$</td>
<td>7.2 ± 3.2</td>
</tr>
<tr>
<td>J2104</td>
<td>37.4 ± 5.9$^a$</td>
<td>4.1 ± 0.3</td>
<td>9.3 ± 1.6</td>
</tr>
<tr>
<td>J2116a</td>
<td>1.8 ± 1.6</td>
<td>⋯</td>
<td>⋯</td>
</tr>
<tr>
<td>J2116b</td>
<td>8.9 ± 2.9</td>
<td>⋯</td>
<td>⋯</td>
</tr>
<tr>
<td>IC 5063</td>
<td>0.82 ± 0.12$^d$</td>
<td>0.050 ± 0.005$^e$</td>
<td>16 ± 3</td>
</tr>
<tr>
<td>M82</td>
<td>12$^f$</td>
<td>0.13$^g$</td>
<td>92</td>
</tr>
<tr>
<td>F11119+3257</td>
<td>52$^h$</td>
<td>0.15$^i$</td>
<td>340</td>
</tr>
</tbody>
</table>

Note. — Total warm and cool H$_2$ masses and mass ratios shown for our sample of LBAs and three nearby objects. For J1434, we show H$_2$ derived from only its main component, and also from the sum of its main and companion as described in Section 3.6.3. (a) In some cases, H$_2$ 1-0 S(1) is unavailable, so we approximate the S(1) line flux with the S(3) flux using the method described in the text. Measurements are taken from (b) Gonçalves et al. (2014), (c) Contursi et al. (2017), (d) Tadhunter et al. (2014), (e) Morganti et al. (2013), (f) Veilleux et al. (2009), (g) Walter et al. (2002), (h) Rupke & Veilleux (2013), and (i) Cicone et al. (2014).
Figure 3.1 VLT/SINFONI channel maps of the continuum-subtracted data cubes for three representative LBAs in our sample, centered on the Pa$\alpha$ line. Velocities are determined with respect to the Pa$\alpha$ line centroid. The PSF is shown in the upper right corner of the top-right panel. The colorscale corresponds to $\log(F_\lambda)$. 
Figure 3.2 In each set of panels, we show VLT/SINFONI Paα integrated flux maps (left), velocity maps (center), and dispersion maps (right), generated without clipping (upper) and with clipping of pixels below $4 \sigma_{\text{rms, chan}}$ (lower). Unclipped velocity and dispersion maps have been masked at the $2 \sigma$ integrated flux level only for visualization purposes. The PSF is shown at the lower left as a black ellipse.
Figure 3.2 (Continued from previous page.)
Figure 3.3 (Continued on next page)
Figure 3.3 The NIR spectra for our sample extracted from a two-pixel radius aperture centered on the peak pixel of Paα integrated flux. The source spectrum is shown in blue, uncertainties are shown in gray, and the spectrum of the blank sky is shown in red.
Figure 3.4 Recombination lines in J0150 for hydrogen (left) and He I (right) relative to systemic velocity. Strong Paα emission has been multiplied by 0.1 for better viewing. Typical uncertainty represented by a standard deviation is shown in the upper right side of the right panel.
Figure 3.5 Recombination lines (rows labeled at right) shown for each LBA (columns labeled at top). Spectra centered on the systemic velocity and extracted over a two-pixel radius aperture are shown in blue, and Gaussian model component fits are shown in red.
Figure 3.6 Dust-corrected hydrogen and He I recombination line fluxes, normalized to Brδ and He I 1869 respectively, shown for all LBAs in different colors and markers. Case B recombination line ratios are shown in black for hydrogen (solid; $T_e = 10^4$ K, $n_e = 10^3$ cm$^{-3}$) and helium (dashed; $T_e = 10^4$ K, $n_e = 10^4$ cm$^{-3}$). The He I lines do not follow their Case B ratios, likely due to the combined effects of resonant scattering, dust or hydrogen absorption, and collisional population of the upper level for the 2.059 µm line (see Section 3.4.2 for more).
Figure 3.7 Observed dust-corrected He I-to-H I recombination line ratios (points) shown with Cloudy model predictions (curves). We show model results while varying effective temperature (colorbar) and logarithmic ionization parameter \( \log U \in \{-4.0, -3.0, -2.0, -1.0\} \), as labeled.
Figure 3.8 Dust-corrected column density normalized by level degeneracy for $\text{H}_2$ $v = 1$ upper rotational levels plotted against energy in temperature units (color). A single best-fit excitation temperature ($T_{\text{ex}}$) is also plotted (in black) for each LBA component aside from J0212, whose ro-vibrational spectrum lacks high enough SNR to fit a temperature profile.
Figure 3.9 Continuum and line ratio maps for our LBA sample. Smoothed contours of the jointly-fit \( \text{Pa}\alpha \) flux are plotted at \( \pm 2, 4, 8, \ldots \times \sigma_{\text{Pa}\alpha,\text{rms}} \). Color scales show continuum flux or ratios of the indicated lines in logarithmic units. He I lines are fit together using the same line width and with 1.869, 1.870, and 2.059 \( \mu m \) line ratios given by Benjamin et al. (1999) Case B calculations. The molecular hydrogen 1-0 S(1), S(2), S(3), and S(5) lines are jointly fit using the same line width and with fluxes determined from degeneracies and (unphysical) \( T_{\text{ex}} = \infty \), which was chosen to maximize signal-to-noise ratio. These maps have not been corrected for dust attenuation.
Figure 3.9 (Continued from previous page.)
Figure 3.10 Hydrogen recombination (left) and H$_2$ ro-vibrational lines (right) for J1434. We show spectra from the main system (solid, $z = 0.180$) and companion (dashed, $z = 0.181$), extracted over two-pixel apertures centered on their respective stellar light peaks and shifted to the same velocity frame. The ratios of H$_2$ to Pa$\alpha$ are clearly different for the main source and its companion. We have smoothed all spectra using a Gaussian filter with $\sigma = 1$ channel for visualization purposes only. Note also that H$_2$ 1-0 S(1) does not lie within our spectral range so it is not shown here.
Chapter 4

Kinematic scaling relations for Lyman break galaxy analogs

4.1 Introduction

The Tully-Fisher relation (TFR; Tully & Fisher 1977) describes the fundamental scaling relation between luminosity and rotation velocity for disk galaxies (see also, e.g., Dalcanton et al. 1997). Following its initial formulation, more recent works have found tighter relationships between ordered rotation velocity and stellar mass \( M_\star \) (see, e.g., Bell & de Jong 2001; Reyes et al. 2011) and total baryonic mass (McGaugh 2005). Kassin et al. (2007) have revealed tight correlations between the stellar mass and \( S_{0.5} \), a kinematic estimator constructed using both ordered rotation and dispersion velocities, where \( S_{0.5} = (0.5V_{\text{rot}}^2 + \sigma^2)^{1/2} \) traces the underlying dark matter potential well. \( S_{0.5} \) encompasses both the rotational support of disk systems and the pressure support of dispersion-dominated systems (e.g., Faber & Jackson 1976).

Kassin et al. (2012) find that at fixed stellar mass, star-forming galaxies decrease in velocity dispersions and increase in ordered rotation velocities over cosmic time. Joint evolution of stellar mass and \( V_{\text{rot}}/\sigma \) results in a “kinematic downsizing” between \( z \sim 1 \) and \( z \sim 0 \). Kinematic downsizing is used to describe the observed propensity for higher-mass galaxies to form disks at earlier epochs (e.g., Cowie et al. 1996; Kassin et al. 2012). A transition mass in the TFR, \( \log(M_\star/M_\odot) = 9.5 \), has been measured for star-forming systems at \( z \sim 0.2 \) (i.e., the mass of disk formation; Simons et al. 2015). In this framework, star-forming galaxies transition from turbulent, dispersion-dominated systems to regularly rotating disks, with the more massive galaxies settling into disks at higher redshifts (see, e.g., Förster Schreiber et al. 2009; Cresci et al. 2009; Law et al. 2009; Cappellari et al. 2013; Krick et al. 2015; Livermore et al. 2015; Wisnioski et al. 2015; Ceverino et al. 2017; Tapia et al. 2017; Übler et al. 2017; Alcorn et al. 2018). These studies demonstrate that disk kinematics of typical star-forming systems in the local Universe are different from their progenitors at \( z \gtrsim 1 \).

At the peak of cosmic star formation (Madau & Dickinson 2014), the progenitors of present-day massive disk galaxies are assembling stellar mass at high rates. Elevated star formation rates (SFRs) and specific SFRs (sSFR = SFR/\( M_\star \)) at \( z \gtrsim 1 \) may be due to large gas fractions (e.g., Tacconi et al. 2014).
2018; Isbell et al. 2018), high rates of interactions and mergers (e.g., Lotz et al. 2011; Cibinel et al. 2019), and turbulent disk instabilities (e.g., Goldbaum et al. 2016). The transition in modes of mass assembly, e.g., from chaotic starbursts at \( z > 1 \) to calmer star formation in the galaxy disks at \( z \sim 0 \), also coincides with the decrease in violent mergers and disruptive interactions and (re-)formation of gas disks (e.g., Ceverino et al. 2017). These findings suggest that the physical processes that regulate star formation at lower redshifts are also responsible for promoting rotational support (e.g., Simons et al. 2017).

Starbursting galaxies and gas-rich mergers can also be found in the local Universe (e.g., Heckman et al. 2005; Buat et al. 2007; Ellison et al. 2013), and it is not clear whether these extreme low-redshift systems also abide by the kinematic downsizing trends over cosmic time, or if they remain disordered. In order to understand the kinematics of extreme star-forming galaxies at low redshifts, we select a sample of \( z \lesssim 0.2 \) starbursts that resemble Lyman break galaxies (LBGs; Steidel et al. 1996) at high redshifts. These low-
\( z \) analogs are well-suited for exploring how the stellar mass TFR and other kinematic scaling relations depend on redshift when other galaxy properties are held constant. Whereas other studies (e.g., Simons et al. 2017) have used abundance matching models to investigate the expected kinematics of the same galaxy populations evolving over time, we now consider kinematic evolution for high- and low-
\( z \) populations with similar SFR, stellar mass, optical morphology, and other galaxy properties.

The structure of this paper is as follows. In Section 4.2, we describe the sample of LBG analogs, and how kinematic quantities are measured. In Section 4.3, we compare \( V_{\text{tot}}, \sigma, \) and \( S_{0.5} \) for star-forming galaxies at low and high redshifts. In Section 4.4, we interpret the results in context of ancillary data sets and current literature. The conclusions are given in Section 4.5. We assume a \( \Lambda \)CDM concordance cosmology (\( H_0 = 70 \) km s\(^{-1}\) Mpc\(^{-1}\), \( \Omega_m = 0.3, \Omega_\Lambda = 0.7 \)). All uncertainties are at the 1 \( \sigma \) level unless otherwise stated.

4.2 Data

4.2.1 Lyman break galaxy analog sample

From the Sloan Digital Sky Survey (SDSS; York et al. 2000) and \textit{GALEX} All-Sky Imaging Survey (Martin et al. 2005), Heckman et al. (2005) have identified a sample of low-redshift starbursts with unusually high FUV luminosities (\( L_{\text{FUV}} > 10^{10.3} L_\odot \)) and high FUV surface brightnesses (later broadened to focus on the most “ultracompact” UV-luminous galaxies with \( I_{\text{FUV}} > 10^9 L_\odot \) kpc\(^2\); Hoopes et al. 2007). This sample resembles the high-
\( z \) LBG population in many ways; the populations have similar UV-optical color, dust extinction, metallicity, SFR, physical size, and emission-line
velocity width (e.g., Overzier et al. 2009). We have obtained VLT/SINFONI integral field spectroscopy of the Paα line for six LBG analogs (Wu et al. 2019). These systems are shown to be moderately dusty on the basis of far-UV attenuations and Herschel imaging (Heckman et al. 2005; Contursi et al. 2017). One member of our sample, 211531, is resolved by our SINFONI observation to be a close pair of similarly sized, orderly rotating galaxies, so we report their individual kinematics in the following analysis. Wu et al. (2019) present details of the multiphase gas, dust, and star-formation properties of these rare z ≤ 0.2 systems, and we briefly discuss below how stellar mass and ionized gas kinematics are measured.

4.2.2 Kinematic measurements

Ionized gas kinematics are determined from ∼ 0.6″ resolution VLT/SINFONI integral field spectroscopy centered on the Paα spectral line (Wu et al. 2019) using the GalPaK3D program (Bouché et al. 2015). The dynamical model accounts for PSF and instrumental broadening, allowing for accurate estimates of the inclination-corrected maximum velocity $V_{\text{rot}}$ and gas velocity dispersion $\sigma$, in addition to other shape and kinematic quantities. The LBG analogs are fit using an arctan rotation curve and an exponential surface brightness profile with a thick disk component. Parameter uncertainties are computed from converged Markov Chain Monte Carlo chains (i.e., the final 700-1000 steps in a 3000-4000 step chain). The ordered velocity for 021348 is not well-constrained since it appears to be a high-inclination system; we find that removing it from our analysis does not change any conclusions. Uncertainties for $V_{\text{rot}}$ and $\sigma$ are low (≈ 5%) for the rest of our sample.

4.2.3 Stellar mass measurements

Stellar masses are estimated from GALEX and SDSS observations using spectral energy distribution (SED) fitting (Heckman et al. 2005; Salim et al. 2005). The seven photometric bands are fit to Bruzual & Charlot (2003) population synthesis templates with exponential declining star formation histories and random star formation bursts superposed (see, e.g, Kauffmann et al. 2003). Stellar masses for our sample range from $9.90 \leq \log(M_*/M_\odot) \leq 10.68$. For 211531, the system comprising J2116a and J2116b, we assign 1/3 and 2/3 of the total stellar mass according to their respective SINFONI $K_\text{s}$-band continuum fluxes (Wu et al. 2019). Systematic uncertainties in $M_*$ are about 0.1 dex (Salim et al. 2005), and broadly agree with dynamical masses (Overzier et al. 2009). We adopt the Heckman et al. (2005) $M_*$ measurements with a conservative 0.2 dex uncertainty to account for measurement and systematic effects. All stellar masses are computed using the Chabrier (2003) initial mass function (i.e., scaled by 0.63 and 0.94 when converting from Salpeter 1955 and Kroupa 2001 IMFs respectively).
4.3 Kinematic scaling relations

We compare $V_{\text{rot}}$, $\sigma$, and $S_{0.5}$ against stellar mass in Figure 4.1. In the top panel, we show the empirical stellar mass TFR for Di Teodoro et al. (2016) at $z \sim 1$, and the Simons et al. (2016) and Straatman et al. (2017) samples at $z \sim 2$. Also overlaid is the Reyes et al. (2011) ridge line representing the tight observed correlation at $z \sim 0$. Our sample lies along the low-$V_{\text{rot}}$ envelope of the high-$z$ data and is highly inconsistent with the Reyes et al. (2011) TFR ridge line.

Redshift evolution in the stellar mass TFR relation suggests that disk galaxies are still in the process of forming their disks at earlier epochs of the Universe (e.g., Conselice et al. 2005; Kassin et al. 2007; Wisnioski et al. 2015; Straatman et al. 2017). Gas that has been recently accreted or influenced by mergers tends to be more kinematically disordered, causing two primary deviations from the stellar mass TFR: galaxy disks are characterized by lower $V_{\text{rot}}$ at fixed stellar mass and higher redshifts, and at fixed redshift and lower masses. If a galaxy’s potential does not allow it to form a disk within the age of the Universe at a given redshift, then it is expected to fall below the stellar mass TFR.

The characteristic mass below which disks cannot efficiently form is observed to evolve with redshift. At $z \sim 0.2$, Simons et al. (2015) find that their sample of galaxies with $\log(M_*/M_\odot) \gtrsim 9.5$ always lie on the stellar mass TFR. At $z \sim 2$, however, disks are able to efficiently form only at $\log(M_*/M_\odot) \gtrsim 10.2$ (Simons et al. 2016). Conversion of gas to stars (i.e., the reduction in gas mass fraction $f_{\text{gas}} \equiv M_{\text{gas}}/M_*$) over cosmic time may be responsible for driving the redshift evolution (e.g., Straatman et al. 2017; Price et al. 2019). We find that LBA disks are not yet settled, despite the fact that they are at low redshift and have $\log(M_*/M_\odot) \gtrsim 10$.

The middle panel of Figure 4.1 shows the gas velocity dispersion compared to stellar mass for the same data sets. The $z \sim 0$ Faber-Jackson relation between $M_*$ and stellar dispersion for spheroidal systems is also indicated (Gallazzi et al. 2006). Note that the Faber-Jackson line is shown for reference only, as the gas and stellar velocity dispersions can respond differently to the same physical processes and should not be equated. We find that LBG analogs occupy the same locus in $M_*-\sigma$ space as the $z \sim 2$ samples, but have higher dispersions than Di Teodoro et al. (2016) $z \sim 1$ star-forming galaxies. All data sets are characterized by elevated gas dispersions relative to typical $z \sim 0$ star-forming spirals (e.g., 5 - 15 km s$^{-1}$; Epinat et al. 2008; Wilson et al. 2011).

In the bottom panel, we compare the $S_{0.5} - M_*$ relation for our LBG analogs and other $z \lesssim 2$ samples. Measurements of $S_{0.5}$ for LBG analogs qualitatively agree with those for higher-$z$ systems, populating the lower envelope of the Simons et al. (2016) sample. However, the $S_{0.5} - M_*$ slope for our sample appears to be steeper than the $z \sim 0$ ridge line. At lower stellar mass, $S_{0.5}$ systematically
falls short of the TFR ridge line, and also the Kassin et al. (2007) $S_{0.5} - M_*$ relation (not shown).

Because the gas kinematics trace the total gravitational potential, (generalized) TFR scaling relations describe the collective growth of stellar and halo masses (Bell & de Jong 2001). When there is an observed decrease in kinematic support (as is seen for some systems at lower mass), we can interpret this as an elevated stellar-to-total mass ratio. Statistics from numerical simulations suggest that the stellar-to-halo mass ratio peaks around $\log(M_*/M_\odot) = 10.5$ and evolves little with redshift (Moster et al. 2010; Behroozi et al. 2013a). Reduced kinematic support for our lower-$M_*$ LBG analogs may be expected if recent physical processes temporarily boost the star formation efficiency and enables mass assembly at rates higher than expected for a given halo mass. Because this explanation is also consistent with the young starburst ages and high fraction of mergers in our sample, we will discuss the gas content of LBG analogs in Section 4.4.2. The observed lack of kinematic support might also be due to the possibility that LBG analogs are not virialized, which would increase the apparent stellar-to-total mass ratio. However, Overzier et al. (2009) find that stellar masses and dynamical masses (computed from H$\alpha$ velocity dispersions and optical sizes) largely agree, which implies that the LBAs’ kinematics are consistent with equilibrium dynamics.

4.4 Discussion

4.4.1 Departures from the stellar mass TFR

Kannappan et al. (2002) find that bright, nearby galaxies with rotation velocities that are too low for the TFR tend to exhibit signs of recent star formation, i.e., they are bluer, stronger in H$\alpha$ emission, and more kinematically asymmetric. Similarly, Flores et al. (2006) and Weiner et al. (2006) report that galaxies with irregular morphologies and disturbed kinematics strongly deviate from the stellar mass TFR. In such systems, the lack of rotation support is compensated with pressure support in the form of dispersive motions. Simons et al. (2015) find that galaxies with stellar mass below the disk formation mass threshold are morphologically compact and asymmetric in $V$-band imaging, and are more likely to be characterized by $V_{\text{rot}}/\sigma < 1$. Low-$z$ LBAs follow this trend: the majority are described by irregular optical morphologies indicative of galaxy interactions and/or dense nuclear starbursts (see, e.g., Overzier et al. 2009), in addition to their high gas dispersion velocities.

Simons et al. (2015) compare the differences between the Reyes et al. (2011) TFR ridge line and the ordered rotation velocities (i.e., the TFR residuals) with $V_{\text{rot}}/\sigma$ for their $z \sim 0.2$ sample (which spans a wide range in stellar mass, $8 \lesssim \log(M_*/M_\odot) \lesssim 11$; see also Kassin et al. 2012). In Figure 4.2, we show the TFR residuals for our LBAs corresponding to a similar redshift range, and for a few high-$z$ galaxy samples (Di Teodoro et al. 2016; Simons et al. 2016; Straatman et al. 2017).
It is immediately apparent that the high-$z$ and low-$z$ systems adhere to a similar relationship between TFR residual and $V_{\text{rot}}/\sigma$. The more disordered a galaxy’s gas motions are, the further it deviates from the Reyes et al. (2011) stellar mass TFR. At $z \sim 0.2$, the trend describes our sample of compact, starbursting LBG analogs, in addition to morphologically diverse field galaxies spanning over two orders of magnitude in $M_*$ (Simons et al. 2015). At $z \sim 2$, galaxies in the Simons et al. (2016) and Straatman et al. (2017) samples follow the same relationship.

These regular departures from the stellar mass TFR may lead to a steeper $V_{\text{rot}} - M_*$ (or shallower $M_* - V_{\text{rot}}$) correlation, as has been seen at higher redshifts (e.g., Flores et al. 2006; Kassin et al. 2007; Cresci et al. 2009; Tiley et al. 2016). Based on Figure 4.2, we find that $v/\sigma$ may be a better control parameter for evolution in the stellar mass TFR. Previous reports of redshift evolution can be attributed to the redshift evolution in $V_{\text{rot}}/\sigma$ for galaxies at a given mass. Substantial fractions of the lower-mass, $z \sim 0$ and higher-mass, $z \sim 2$ star-forming populations are not disk-dominated (e.g., they have $V_{\text{rot}}/\sigma < 3$; Simons et al. 2017). Our sample of LBG analogs are unique because they are at low redshifts, high stellar masses, and yet are still not disk-dominated.

It is worth noting that nearly all systems in the Di Teodoro et al. (2016) and Straatman et al. (2017) samples are characterized by $V_{\text{rot}}/\sigma > 1$. The relatively high values of $V_{\text{rot}}/\sigma$ for the $z \sim 2$ Straatman et al. (2017) sample may also explain why these systems are in better agreement with the Reyes et al. (2011) stellar mass TFR than the Simons et al. (2016) sample at similar redshifts. Similarly, the $z \sim 1$ galaxies are highly rotation-dominated ($V_{\text{rot}}/\sigma > 3$) relative to their parent sample (see, e.g., Stott et al. 2016). Di Teodoro et al. (2016) select non-interacting, moderately inclined, $z \sim 1$ galaxies with H$\alpha$ spread over eight channels. Although it is difficult to quantify exactly how these criteria would bias kinematic scaling relations, it is plausible that the selection effects can account for differences between this and the other $0 \lesssim z \lesssim 2$ samples (e.g., Tiley et al. 2016). The fact that these galaxies do not follow the same correlation in TFR residuals versus $V_{\text{rot}}/\sigma$ is interesting, but may arise from selection biases and small sample size.

4.4.2 Cold gas and the Baryonic TFR

Cold gas masses from IRAM/PdBI CO(1–0) observations are available for 4/7 LBG analogs (Gonçalves et al. 2014; Contursi et al. 2017). For J2104, we use CO(1–0) flux measured from higher-resolution PdBI observations. Because our sample is characterized by moderately high SFRs ($\sim 10 - 100 \, M_\odot \, \text{yr}^{-1}$) and only modestly sub-solar metallicities ($0.3 < Z/Z_\odot < 0.7$), we assume a Galactic $X_{\text{CO}}$ factor to convert CO(1–0) line flux to $M_{\text{gas}}$ (which does not include corrections for helium or heavier elements; e.g., Bolatto et al. 2013). We also adopt a 0.3 dex uncertainty for the gas mass. Gas masses for the four LBG analogs are between $(1.1 - 2.1) \times 10^{10} \, M_\odot$. 
For the remaining three systems, we use the SFR, size, an assumed gas depletion time $t_{\text{dep}}$, and a Kennicutt (1998b) star-formation law with $N = 1.4$ in order to compute the total gas mass (e.g., Erb et al. 2006b). SFRs for J2116a and J2116b have previously been combined because they are estimated (partly) from blended $\sim 5''$-resolution GALEX observations (Salim et al. 2005). We assume that the total SFR can be divided between J2116a and J2116b according to their relative contributions to the summed Pa$\alpha$ flux, i.e., 0.17 and 0.83 respectively (Wu et al. 2019). For two lensed $z \sim 3$ LBGs, gas depletion timescales are about $t_{\text{dep}} \sim 100 - 200$ Myr (when normalized to the Galactic $X_{\text{CO}}$ conversion; see, e.g., Baker et al. 2004; Coppin et al. 2007; Riechers et al. 2010). Under the assumption of the same conversion factor, gas depletion timescales for our CO-detected LBG analogs are somewhat longer: $t_{\text{dep}} \sim 200 - 650$ Myr, with a mean value of $\sim 420$ Myr. The difference in gas depletion times between the two samples may not be due to redshift effects; one possibility is that the selection of IR-luminous LBG analogs biases our sample to be more likely to be dusty and thus gas-rich (even given comparable SFRs; e.g., Scoville et al. 2016).

Using combined gas and stellar masses, we can examine whether the LBG analog sample follows the baryonic TFR (BTFR). In the left panel of Figure 4.3, we plot $V_{\text{rot}}$ against total baryonic mass $M_{\text{bar}} \equiv M_{\text{gas}} + M_{\star}$. LBG analogs are color-coded by their gas mass fractions, which are in the range $0.10 \leq f_{\text{gas}} \leq 0.73$ (and $0.26 \leq f_{\text{gas}} \leq 0.59$ for the four systems with directly measured CO($1 - 0$) fluxes). For comparison, we also show linear fits to the BTFR measured at $z \sim 0$ (Lelli et al. 2016) and at $z \sim 1.5$ (Übler et al. 2017).\(^1\) In the right panel of Figure 4.3, we also show $S_{0.5}$ against total baryonic mass. The same relation for a sample of star-forming galaxies at $z \sim 2$ (Price et al. 2016, who have adopted the $S_{0.5} - M_{\text{bar}}$ slope from Kassin et al. 2007) is shown for comparison.

From Figure 4.3, we see that our sample of LBG analogs does not agree with the low- or high-$z$ literature relations (even if we remove J0213, the galaxy with anomalously low-$V_{\text{rot}}$). LBG analogs are shifted to higher baryonic masses at lower rotation velocities or $S_{0.5}$. These results suggest that the LBG analogs’ excursions from the stellar mass TFR cannot be attributed to extreme starbursts that temporarily increase $M_{\star}$ before the rotation velocity (and total potential) can catch up, since total baryonic mass should be insensitive to such processes. Similarly, the hint of a steeper slope in the $S_{0.5} - M_{\star}$ relation appears to also hold true for the $S_{0.5} - M_{\text{bar}}$ relations.

Galaxies tend to be physically larger at lower redshifts (e.g., van der Wel et al. 2014; Shibuya et al. 2015), such that their gas kinematics are determined by higher enclosed masses that are increasingly dark matter-dominated at larger radii. Conversely, distant ($z > 2$) galaxies are more

---

\(^1\)Übler et al. (2017) measure the BTFR at $z \sim 1.5$, but fix the slope to the Lelli et al. (2016) relation measured at $z \sim 0$. Fixing the slope parameter is a common procedure at higher redshifts due to the lack of available data, but it can also hide some of the evolution in the BTFR. When Übler et al. (2017) allow the slope to vary (e.g., their “free-slope” fit) and adopt baryonic mass uncertainties of $\sim 0.3 - 0.4$ (see their Appendix C), the $z \sim 1.5$ BTFR becomes steeper, as we show in Figure 4.3.
likely to baryon-dominated (especially so for compact systems; see, e.g., Wuyts et al. 2016; Price et al. 2019). LBG analogs are, by selection (Heckman et al. 2005; Hoopes et al. 2007), extremely compact low-
\( z \) systems characterized by high SFR surface densities (and thus, high gas surface densities). Part of the reason why LBG analogs do not fall on the stellar mass or baryonic TFR may be due to their small sizes, such that their kinematics do not include significant contributions from the subdominant dark matter halo. Their compact natures also explain why they systematically populate lower \( V_{\text{rot}} \) and \( S_{0.5} \) for both the stellar mass and baryonic kinematic scaling relations.

Other physical and observational effects may complicate this interpretation. Dissipative baryonic processes (Blumenthal et al. 1986; Lovell et al. 2018), disk turbulence (Teklu et al. 2018; Price et al. 2019), or feedback (Brooks & Zolotov 2014; Weinberger et al. 2017) may also perturb the dark matter potential in complex ways, resulting in departures from kinematic scaling relations. The 0.6᾽-resolution NIR integral field spectroscopy used for modeling disks may artificially transfer ordered rotation into dispersive motions, but Wu et al. (2019) show that kinematics are mostly in agreement with previous AO observations (Gonçalves et al. 2010; and in the case of J2104, we find strong agreement with ALMA CO\((3-2)\) kinematics modeled in the same way).

4.5 Conclusions

We have studied the kinematic scaling relations for a sample of seven compact, \( z < 0.2 \), star-forming galaxies that resemble Lyman break galaxy systems in the distant Universe, i.e., they are nearby LBG analogs. These systems are described by low ordered rotation and high dispersion velocities, and do not follow the stellar mass TFR derived from more typical low-
\( z \) star-forming galaxies (e.g., Reyes et al. 2011). By comparing \( S_{0.5} \) to stellar mass (Kassin et al. 2007), we find better agreement between LBG analogs and high-
\( z \) samples, although disagreements with low-
\( z \) scaling relations are still present.

We examine how the ordered-to-disordered motions, parameterized by \( V_{\text{rot}}/\sigma \), correlates with offsets from the \( z \approx 0 \) stellar mass TFR (i.e., TFR residuals). Unsurprisingly, LBG analogs with very low rotation support deviate strongly, and higher-\( V_{\text{rot}}/\sigma \) systems lie on the stellar mass TFR. This trend appears to describe star-forming galaxies for a wide range of redshifts, \( 0 \lesssim z \lesssim 2 \), and only breaks down for samples that are already biased to conform to the stellar mass TFR (e.g., those marked by high \( V_{\text{rot}} \)). Because the relation between TFR residuals and \( V_{\text{rot}}/\sigma \) is otherwise constant for most of cosmic history, the ordered-to-disordered velocity ratio may be useful as a proxy for understanding when the stellar mass TFR breaks down independently of redshift.

We also investigate the total baryonic content, \( M_{\text{bar}} = M_\star + M_{\text{gas}} \), for our sample. We find that
the LBG analogs’ kinematics disagree with the BTFR measured at both low and high redshifts, requiring an even steeper $V_{\text{rot}} - M_{\text{bar}}$ slope than the relation measured at high redshifts. The same is true when comparing our sample with the empirical $S_{0.5} - M_{\text{bar}}$ relation.

We ascribe the anomalously low kinematic support of LBG analogs to their small physical sizes. Gas in compact galaxies have their dynamics determined mostly by the galaxy rather than the more extended dark matter halo. Because $V_{\text{rot}}$ or $S_{0.5}$ probe only a truncated portion of the halo, where dark matter does not substantially contribute to the gravitational potential, the usual kinematic scaling relations break down.

These trends strongly depend on the lowest and highest mass LBG analogs. Our sample is also small, comprising just seven objects, so we caution against any over-interpretation of these results. Additional observations for larger samples are required to further test the results of this work.
Figure 4.1 Kinematic scaling relations for the LBG analogs (LBA) sample. Ordered rotation velocity $V_{\text{rot}}$, gas velocity dispersion $\sigma$, and combined kinematic estimator $S_{0.5} \equiv (0.5V_{\text{rot}}^2 + \sigma^2)^{1/2}$ are plotted against stellar mass $M_\star$. Shown for comparison are the Di Teodoro et al. (2016) $z \sim 1$ sample (blue triangles), Simons et al. (2016) $z \sim 2$ sample (gray squares), and Straatman et al. (2017) $z \sim 2$ sample (purple diamonds).
Figure 4.2 $V_{\text{rot}}$ subtracted from the Reyes et al. (2011) TFR ridge line (i.e., TFR “residuals”) versus the $V_{\text{rot}}/\sigma$ ratio. Our LBG analogs are shown as filled black circles and with 1 $\sigma$ error bars, although most are too small to be seen. Also shown are TFR residuals for the Simons et al. (2015) $0.1 < z < 0.375$ sample (open green circles) and the same high-z galaxy samples from Figure 4.1.
Figure 4.3 (Left) Ordered rotation velocity plotted against total baryonic mass. BTFR trends are also shown for $z \sim 0$ (red line; Lelli et al. 2016) and $z \sim 2$ samples (blue line; Übler et al. 2017). (Right) The $S_{0.5} - M_{\text{bar}}$ relation shown for our sample, and for the Price et al. (2016) $z \sim 2$ sample, normalized to the Kassin et al. (2007) $z \sim 0$ slope. LBG analogs’ gas mas fractions are illustrated via colors.
Chapter 5

Predicting metallicity from three-band imaging using convolutional neural networks


5.1 Introduction: The growing role of machine learning in astronomy

Large-area sky surveys, both on-going and planned, are revolutionizing our understanding of galaxy evolution. The Dark Energy Survey (DES; The Dark Energy Survey Collaboration 2005) and upcoming Large Synoptic Survey Telescope (LSST; LSST Dark Energy Science Collaboration 2012) will scan vast swaths of the sky and create samples of galaxies of unprecedented size. Spectroscopic follow-up of these samples will be instrumental in order to understand their properties. Previously, the Sloan Digital Sky Survey (SDSS; York et al. 2000) and its spectroscopic campaign enabled characterization of the mass-metallicity relation (hereafter MZR; Tremonti et al. 2004) and the fundamental metallicity relation, (hereafter FMR; e.g., Mannucci et al. 2010). As future surveys are accompanied by larger data sets, individual spectroscopic follow-up observations will become increasingly impractical.

Fortunately, the large imaging data sets to be produced are ripe for application of machine learning (ML) methods. ML is already showing promise in studies of galaxy morphology (e.g., Dieleman et al. 2015; Huertas-Company et al. 2015; Beck et al. 2018; Dai & Tong 2018; Hocking et al. 2018), gravitational lensing (e.g., Hezaveh et al. 2017; Lanusse et al. 2018; Petrillo et al. 2017, 2018), galaxy clusters (e.g., Ntampaka et al. 2015, 2016), star-galaxy separation (e.g., Kim & Brunner 2017), creating mock galaxy catalogs (e.g., Xu et al. 2013), asteroid identification (e.g., Smirnov & Markov 2017), and photometric redshift estimation (e.g., Hoyle 2016; D’Isanto & Polsterer 2018; Pasquet et al. 2019), among many others. ML methods utilizing neural networks have grown to prominence in recent years. While neural networks are a relatively old technique (e.g., LeCun et al. 1989), their recent increase in popularity is driven by the widespread availability of affordable graphics processing
units (GPUs) that can be used to do general purpose, highly parallel computing. Also, unlike more “traditional” ML methods, neural networks excel at image classification and regression problems.

Inferring spectroscopic properties from the imaging taken as part of a large-area photometric survey is, at a basic level, an image regression problem. These problems are most readily solved by use of convolutions in multiple layers of the network (see, e.g., Krizhevsky et al. 2012). Convolutional neural networks (CNNs, or convnets) efficiently learn spatial relations in images whose features are about the same sizes as the convolution filters (or kernels) that are to be learned through training. CNNs are considered deep when the number of convolutional layers is large. Visualizing their filters reveals that increased depth permits the network to learn more and more abstract features (e.g., from Gabor filters, to geometric shapes, to faces; Zeiler & Fergus 2014).

In this work, we propose to use supervised ML by training CNNs to analyze pseudo-three color images and predict the gas-phase metallicity. We use predicted metallicities to recover the empirical Tremonti et al. (2004) MZR. This paper is organized as follows: In Section 5.2, we describe the acquisition and cleaning of the SDSS data sample. In Section 5.3, we discuss selection of the network’s hyperparameters and outline training the of network. We present the main results in Section 5.4. In Section 5.5, we interpret the CNN’s performance and discuss our findings in the context of current literature. In Section 5.6, we characterize the MZR using the metallicity predicted by our CNN. We summarize our key results in Section 5.7.

Unless otherwise noted, throughout this paper, we use a concordance cosmological model ($\Omega_{\Lambda} = 0.7$, $\Omega_m = 0.3$, and $H_0 = 70$ km s$^{-1}$ Mpc$^{-3}$), assume a Kroupa initial mass function (Kroupa 2001), and use AB magnitudes (Oke 1974).

5.2 The SDSS emission-line galaxy sample

To create a large training sample, we select galaxies from the Sloan Digital Sky Survey (SDSS; York et al. 2000) DR7 MPA/JHU spectroscopic catalog (Kauffmann et al. 2003; Brinchmann et al. 2004; Tremonti et al. 2004; Salim et al. 2007). The catalog provides spectroscopically derived properties such as stellar mass ($M_*$) and gas-phase metallicity ($Z$) estimates (Tremonti et al. 2004). We select objects with low reduced chi-squared of model fits ($r$Chi2 < 2), and median $Z$ estimates available (oh_p50). We supplement the data from the spectroscopic catalog with photometry in each of the five SDSS photometric bands ($u, g, r, i, z$), along with associated errors from SDSS DR14 (Abolfathi et al. 2018).

We require that galaxies magnitudes are $10 < ugriz < 25$ mag, in order to avoid saturated and low signal-to-noise detections. We enforce a color cut, $0 < u - r < 6$, in order to avoid extremely
blue or extremely red objects, and require objects to have spectroscopic redshifts greater than
$z = 0.02$ with low errors ($z_{\text{err}} < 0.01$). The median redshift is 0.07 and the highest-redshift object
has $z = 0.38$. We also require that the $r$-band magnitude measured inside the Petrosian radius
($\text{petroMag}_r$; Petrosian 1976) be less than 18 mag, corresponding to the spectroscopic flux limit.
With these conditions we construct an initial sample of 142,182 objects (there are four objects with
duplicate SDSS DR14 identifiers). We set aside 25,000 objects for later testing, and use the rest for
training and validation.

We create RGB image cutouts of each galaxy with the SDSS cutout service\(^1\), which converts
gri bands to RGB channels according to the algorithm described in Lupton et al. (2004) (with
modifications by the SDSS SkyServer team). Since images are not always available, we are left with
116,429 SDSS images with metallicity measurements, including 20,466/25,000 of the test subsample.
We create $128 \times 128$-pixel JPG images with a pixel scale of 0.296″, which corresponds to $38″ \times 38″$
on the sky. We do not further preprocess, clean, or filter the images before using them as inputs to
our CNN.

5.3 Methodology for training the convolutional neural network

Before the CNN can be asked to make predictions, it must be trained to learn the relationships
between the input data (the images described above) and the desired output (metallicity). The
CNN makes predictions using the input images, and the error (or loss) is determined based on the
differences between true and predicted values. The CNN then updates its parameters, or weights,
in a way that minimizes the loss function. We use the root mean squared error loss function:

$$\text{RMSE} \equiv \sqrt{\langle (y_{\text{true}} - y_{\text{pred}})^2 \rangle},$$

where $y_{\text{true}}$ is the “true” and $y_{\text{pred}}$ is the predicted value, and $y$ represents the target quantity.

It is worth emphasizing that the $Z_{\text{true}}$ is the metallicity estimated by model fits to strong emission
lines in the SDSS spectra. Tremonti et al. (2004) determine a likelihood distribution of metallicities
based on the model fits, and we define their 50th percentile metallicity estimates to be the true
metallicity ($Z_{\text{true}}$) for the purpose of training our network. The typical systematic uncertainty in
their metallicity model fits is about 0.03 dex.

We randomly split our training sample of $\sim 96,953$ images into 80% (76,711) training and 20%
(19,192) validation data sets, respectively. The test data set of 20,466 images is isolated for now,\(^1\)

\(^1\)http://skyserver.sdss.org/dr14/en/help/docs/api.aspx
and is not accessible to the CNN until all training is completed. Images and $Z_{\text{true}}$ answers are given to the CNN in “batches” of 256 at a time, until the full training data set has been used for training. Each full round of training using all of the data is called an epoch, and we compute the loss using the validation data set at the end of each epoch. We use gradient descent for each batch to adjust weight parameters, and each weight’s fractional contribution of loss is determined by the backpropagation algorithm (LeCun et al. 1989), during which finite partial derivatives are computed and propagated backwards through layers (i.e., using the chain rule for derivatives).

We use a 34-layer residual CNN architecture (He et al. 2015) initialized to weights pre-trained on the ImageNet data set, which consists of 1.7 million images belonging to 1000 categories of objects found on Earth (e.g., cats, horses, cars, or books; Russakovsky et al. 2014). The CNN is trained for a total of 10 epochs. For more details about the CNN architecture, transfer learning, hyperparameter selection, data augmentation, and the training process, see Section 5.8. In total, our training process requires 25-30 minutes on our GPU and uses under 2 GB of memory.

We evaluate predictions using the RMSE loss function, which approaches the standard deviation for Gaussian-distributed data. We also report the NMAD, or the normal median absolute deviation (e.g., Ilbert et al. 2009; Dahlen et al. 2013; Molino et al. 2017):

$$\text{NMAD}(x) \approx 1.4826 \times \text{median}(|x - \text{median}(x)|),$$

where for a Gaussian-distributed $x$, the NMAD will also approximate the standard deviation, $\sigma$. NMAD has the distinct advantage in that it is insensitive to outliers and can be useful for measuring scatter. However, unlike the RMSE, which quantifies the typical scatter distributed about a center of zero, NMAD only describes the scatter around the (potentially non-zero) median.

## 5.4 Predicting metallicity using only three-color imaging

### 5.4.1 Example predictions

In Figure 5.1, we show examples of $128 \times 128$ pixel $gri$ SDSS images that are evaluated by the CNN. Rows (a) and (b) depict the galaxies with lowest predicted and lowest true metallicities, respectively. The CNN associates blue, edge-on disk galaxies with low metallicities, and is generally accurate in its predictions. In rows (c) and (d), we show the galaxies with highest predicted and highest true metallicities, respectively. Here we find that red galaxies containing prominent nuclei are predicted to be high in metallicity, and that their predictions generally match $Z_{\text{true}}$.

Galaxies predicted by our CNN to have high metallicities ($Z_{\text{pred}} > 9.0$) tend to be characterized
by high $Z_{\text{true}}$, and the equivalent is true for low-metallicity galaxies. Conversely, galaxies with the highest (lowest) $true$ metallicities in the sample are also predicted to have high (low) metallicities. Note that inclined galaxies tend to be lower in metallicity whereas face-on galaxies appear to be higher in metallicity. Tremonti et al. (2004) explain this correlation by suggesting that the SDSS fiber aperture captures more column of a projected edge-on disk, allowing the metal-poor, gas-rich, and less-extincted outer regions to more easily be detected and depress the integrated $Z_{\text{true}}$.

We will now consider examples of the most incorrectly predicted galaxies. In rows (e) and (f), we show instances in which the CNN predicted too low metallicity and too high metallicity, respectively. The two galaxies with the most negative residuals $\Delta Z \equiv Z_{\text{pred}} - Z_{\text{true}}$ (i.e., most under-predicted metallicities) suffer from artifacts that cause unphysical color gradients, and/or are labeled as quasars on the basis of their SDSS spectra (for which we expect $Z_{\text{true}}$ to be biased). It is not unsurprising that the CNN has made mistakes in some of these cases, since they go against astronomers’ usual heuristics: blue, disk-dominated sources are generally thought of as lower in metallicity, and redder, more spheroidal objects tend to be higher in metallicity.

In the bottom row (g) of Figure 5.1, we show five randomly selected galaxies. The random SDSS assortment consists of lenticular, spiral, and possibly even an interacting pair of galaxies. Residuals are low (below 0.15 dex), and we again find that the CNN predictions track with human visual intuition.

5.4.2 Comparing predicted and true metallicities

In Figure 5.2, we show histograms of the true and predicted metallicities in black and red, respectively. The histogram bin sizes are chosen according to the Freedman & Diaconis (1981) rule for each distribution. The discreet striping of the Tremonti et al. (2004) and Brinchmann et al. (2004) metallicity estimator appears in the $Z_{\text{true}}$ distribution but does not appear in our CNN predictions. This striping should increase the scatter in our distribution of residuals.

The range of $Z_{\text{pred}}$ is more limited than the range of $Z_{\text{true}}$, which can also be seen from Figure 5.1 for extreme values of $Z_{\text{true}}$. Too narrow a domain in $Z_{\text{pred}}$ will lead to systematic errors, as the CNN will end up never predicting very high or very low metallicities. Although the two distributions are qualitatively consistent with each other at low metallicities (e.g., $Z < 8.5$), the fraction of galaxies with high $Z_{\text{true}} > 9.1$ (2573/20466 = 12.6%) is higher than the fraction with high $Z_{\text{pred}} > 9.1$ (1174/20466 = 5.7%).

We find that the mode of the binned predicted metallicity distribution is higher than that of $Z_{\text{true}}$. This result may be a consequence of the CNN overcompensating for its systematic under-prediction of metallicity for galaxies with $Z_{\text{true}} > 9.1$. However, its effect on the entire distribution
is small, and may be remedied simply by increasing the relative fraction of very high-$Z_{\text{true}}$ objects. We find overall good qualitative agreement between the $Z_{\text{pred}}$ and $Z_{\text{true}}$ distributions.

5.4.3 Scatter in $Z_{\text{pred}}$ and $Z_{\text{true}}$

In Figure 5.3, we compare the distributions of $Z_{\text{true}}$ and $Z_{\text{pred}}$ using a two-dimensional histogram (shown in grayscale in the main, larger panel). We also show the median predictions varying with binned $Z_{\text{true}}$ (solid red line), in addition to the scatter in RMSE (dashed red) and NMAD (dashed violet), and also the one-to-one line (solid black). The running median agrees well with the one-to-one line, although at low metallicity we find that the CNN makes makes overpredictions.

A histogram of metallicity residuals is shown in the inset plot of the Figure 5.3 main panel. The $\Delta Z$ distribution is characterized by an approximately normal distribution with a heavy tail at large positive residuals; this heavy tail is likely due to the systematic over-prediction for low-$Z_{\text{true}}$ galaxies. There is also an overabundance of large negative $\Delta Z$ corresponding to under-predictions for high $Z_{\text{true}}$, although this effect is smaller. We do not find significant correlations between $\Delta Z$ and galaxy observables including spectroscopic redshift, any combination of photometric color (including $u$ and $z$ bands), emission line signal-to-noise ratios, observed $gri$-magnitudes, or axis ratios.

We now turn our attention to the upper panel of Figure 5.3, which shows how the scatter varies with spectroscopically derived metallicity. The RMSE scatter and outlier-insensitive NMAD are both shown. Marker sizes are proportional in area to the number of samples in each $Z_{\text{true}}$ bin, and the horizontal lines are located at the average loss (RMSE or NMAD) for the full test data set.

Predictions appear to be both accurate and low in scatter for galaxies with $Z_{\text{true}} \approx 9.0$, which is representative of a typical metallicity in the SDSS sample. Where the predictions are systematically incorrect, we find that the RMSE increases dramatically. However, the same is not true for the NMAD; at $Z_{\text{true}} < 8.5$, it asymptotes to $\sim 0.10$ dex, even though the running median is incorrect by approximately the same amount! This discrepancy is because the NMAD determines the scatter about the median and not $\Delta Z = 0$, and thus, this metric becomes somewhat unreliable when the binned samples do not have a median value close to zero. Fortunately, the global median of $\Delta Z$ is $-0.006$ dex, or less than 10% of the RMSE, and thus the global NMAD $= 0.067$ dex is representative of the outlier-insensitive scatter for the entire test data set.

This effect partly explains why the global NMAD (0.067 dex) is higher than the weighted average of the binned NMAD ($\sim 0.05$ dex). Also, each binned NMAD is computed using its local scatter, such that the outlier rejection criterion varies with $Z_{\text{true}}$. To illustrate this effect with an example: $\Delta Z \approx 0.2$ dex would be treated as an $3\sigma$ outlier at $Z_{\text{true}} = 9.0$, where the CNN is generally accurate, but the same residual would not be rejected as an outlier using NMAD for $Z_{\text{true}} = 8.5$. Since the
binned average NMAD depends on choice of bin size, we do not include those results in our analysis and only focus on the global NMAD. RMSE is a robust measure of both local and global scatter (although it becomes biased high by outliers).

5.5 Interpreting the CNN

5.5.1 Uncertainty in the scatter

It is worth examining how reliable our estimate of $\text{RMSE} = 0.085$ dex is. Because we have a large data set, we can calculate uncertainties on the RMSE through multiple training/test realizations. We divide our full data set into cross-validation and nested cross-validation splits in order to see how the RMSE varies. There also exists an additional contribution to the scatter due to imaging noise, which we do not account for in our estimate of the uncertainties.

For the first method (5-fold cross-validation), we take the entire data set from Section 5.2 and split it into five 80%/20% training/test subsets, each of which is optimized independently. We then compute the mean and standard deviation of the five test samples’ $Z_{\text{pred}}$, and find that the RMSE = $0.0836 \pm 0.0005$. Because there are more training examples here than in our original training set, the mean RMSE is lower than what we have previously found in Section 5.4.

For the second method (nested cross-validation), we split the full data set into five 80%/20% training/validation-test splits, and then further divide the training/validation data sets into 75%/25% cross-validation splits. We compute the mean and standard deviation of $Z_{\text{pred}}$ for the ensemble of 5-fold test splits. When we select the model with the best cross-validation score, the RMSE is $0.0823 \pm 0.0009$. The unweighted average of all training/validation models is $\text{RMSE} = 0.0831 \pm 0.0011$.

5.5.2 Impact of artificially increasing scatter

In order to simulate additional uncertainty that may arise from noisier measurements of spectral lines, we add normally distributed scatter to the $Z_{\text{true}}$ values. We train our CNN as before (in Section 5.4), except that we add random $\sigma = \{0.03, 0.05, 0.10, 0.20\}$ dex of scatter to the target $Z_{\text{true}}$. The smallest value, 0.03 dex, is the same as the systematic uncertainty in the Tremonti et al. (2004) measurements, and 0.20 dex represents the standard deviation for the entire $Z_{\text{true}}$ distribution. We compare CNN-predicted $Z_{\text{pred}}$ with the original $Z_{\text{true}}$ (i.e., the underlying values without artificial scatter introduced) using the RMSE metric as before (Equation 5.1). For all values of additional scatter, the CNN is able to estimate $Z_{\text{pred}}$ to $\text{RMSE} = \{0.0851, 0.0851, 0.0869, 0.0882\}$ dex respectively. These results show that the CNN is robust to extra scatter added in an unbiased way.

As a second test, we include random scatter drawn from a normal distribution centered at 0 and
with standard deviation equal to the galaxy’s redshift. This simple model can test how the CNN responds to redshift-dependent scatter in $Z_{\text{true}}$. We note that this toy model disproportionately impacts higher-metallicity galaxies because our sample of lower-mass (and thus, lower-metallicity) galaxies is less complete at higher redshifts. After training on the original images with these modified $Z_{\text{true}}$ values, we find that the resulting $Z_{\text{pred}}$ are not strongly affected: RMSE = 0.0862 dex.

These tests demonstrate that our CNN is robust to normally distributed scatter in $Z_{\text{true}}$. Our results show initial promise for cases in which training data are more uncertain, such as at higher redshift. However, more testing is necessary to understand the effects of biased or correlated sources of uncertainty, and to account for the evolving relationships between metallicity and other observed properties (e.g., Zahid et al. 2013; Salim et al. 2015).

### 5.5.3 Resolution and color effects

Because our methodology is so computationally light, we can run the same CNN training and test procedure on images scaled to different sizes in order to understand the effects of image resolution. Our initial results use SDSS 38″ × 38″ cutouts resized to 128 × 128 pixels, and we now downsample the same images to 64 × 64, 32 × 32, · · ·, 2 × 2, and even 1 × 1 pixels via re-binning. All images retain their three channels, so the smallest 1 × 1 image is effectively the pixels in each of the gri bands averaged together with the background and possible neighboring sources.

In Figure 5.4, we show the effects of image resolution by measuring the global scatter in $\Delta Z$ using the RMSE and NMAD metrics (shown in red and violet circular markers, respectively). Also shown is the scatter in $\Delta Z$ if we always predict the mean value of $Z_{\text{true}}$ over the data set (shown using a square marker). This constant prediction effectively delivers the worst possible scatter, and the Tremonti et al. (2004) systematic uncertainty in $Z_{\text{true}}$ of $\sim 0.03$ dex yields the best possible scatter. We find that both RMSE and NMAD decrease with increasing resolution, as expected if morphology or color gradients are instrumental to predicting metallicity.

There appears to be little improvement in scatter going from 1 × 1 to 2 × 2 pixel images. 1 × 1 three-color images contain similar information to three photometric data points (although because background and neighboring pixels are averaged in, they are less information-dense than photometry), which can be used to perform a crude spectral energy distribution (SED) fit. Therefore it unsurprising that the 1 × 1 CNN predictions perform so much better than the baseline mean prediction. A 2 × 2 three-color image contains four times as many pixels as a 1 × 1 image, but because the object is centered between all four pixels, information is still averaged among all available pixels. Therefore, the scatter does not improve appreciably going from 1 × 1 to 2 × 2 resolution.²

²There is extra information in the 2 × 2 pixel images in non-circularly symmetric cases. For an inclined disk, it is
The scatter is a strong function of resolution as the images are resolved from $2 \times 2$ to about $32 \times 32$ pixels. With further increasing resolution, improvement is still evident, although the scaling with scatter is noticeably weaker. Because the angular size of each image cutout stays the same, the pixel scale changes from $1.184'' \text{ pix}^{-1}$ for $32 \times 32$ images, to $0.592'' \text{ pix}^{-1}$ for $64 \times 64$ images, to $0.296'' \text{ pix}^{-1}$ for $128 \times 128$ images. The native SDSS pixel resolution is $0.396'' \text{ pix}^{-1}$, such that the $64 \times 64$ and $128 \times 128$ resolutions result in the oversampling of each image. Thus, scatter is expected to plateau for images larger than $128 \times 128$. It is worth noting, however, that the CNN attempts to learn filters that depend on the size of the input image, so smaller images may result in the CNN training filters that are too low in resolution to be completely effective for prediction. Therefore, it is also not surprising that the CNN makes incremental gains for images with increasing resolution beyond $64 \times 64$ pixels.

We also train the CNN to predict metallicity using only the central $16 \times 16$-pixel regions of each SDSS $gri$ image. We find that the network is able to predict metallicity to within $\text{RMSE} = 0.0965$ dex. Because this value is higher than the scatter found when using the full-sized images, we conclude that the CNN loses valuable information when only the central regions are considered, and that relevant information for predicting global metallicity can be found in the galaxies’ outer regions that are not probed by the 3'' SDSS spectroscopic fibers.

As a way of testing how the CNN responds to reduced color information, we have also repeated our training and testing routines using $r$-band and $gr$-band $128 \times 128$ images. In order to make use of our pretrained network, we modify the original, three color JPG images to correspond to either one- or two-band SDSS images. For the single-band imaging, we duplicate the $r$-band data into the blue and red JPG channels. For the $gr$-band images, the blue and red channels correspond to the $g$ and $r$ filters, while the green channel is the mean of the two.

The large squares labeled “$r$” in Figure 5.4 show that the network trained and tested on single-band images performs relatively poorly ($\text{RMSE} = 0.1381$ dex) compared to $gri$ imaging even at low resolution. The addition of a second color improves CNN performance significantly. When a second band is added to the training images (box labeled “$gr$” in Figure 5.4), the RMSE improves to a level similar (0.0915 dex) to that of the original three-color images (0.0851 dex). This enhancement may be due to extra information that the bluer $g$-band provides about younger stellar populations. In both examples, the CNN is able to utilize spatial information about a galaxy to improve metallicity estimates.

possible to roughly determine the orientation in the sky plane, but this information is not very useful. In the case of a major merger or interacting companion, the $2 \times 2$ images may be more powerful than $1 \times 1$ images.
5.5.4 Random forest predictions for metallicity

We also construct a random forest (RF) of decision trees in order to predict metallicity using the implementation from scikit-learn (Pedregosa et al. 2012). Hyperparameters are selected according to the optimal RF trained by Acquaviva (2016). We use exactly the same data labels (i.e., galaxies) to train/validate or test the RF that we have used for training and testing the CNN, so that our measurements of scatter can be directly compared. However, we have used the gri three-band photometry data (given in magnitudes) to train and predict metallicity. Since each galaxy only has three pieces of photometric information, it can be compared to the 1 × 1 three-band “images” processed by our CNN.

The RF predicts metallicity with RMSE = 0.130 dex, which is superior to our CNN trained and tested on 1 × 1 and 2 × 2 images. This result is unsurprising because the RF is supplied aperture-corrected photometry, whereas the CNN is provided 1 × 1 gri “images” whose features have been averaged with their backgrounds. 2 × 2 images are only marginally more informative. When the resolution is further increased to 4 × 4 images, then the CNN can begin to learn rough morphological features and color gradients, which is already enough to surpass the performance (measured by both RMSE and NMAD) of the RF. This result suggests that the CNN is able to learn a nontrivial representation of gas-phase metallicity based on three-band brightness distributions, even with extremely low-quality data.

5.5.5 Comparisons to previous work

CNNs have been used for a wide variety of classification tasks in extragalactic astronomy, including morphological classification (e.g., Dieleman et al. 2015; Huertas-Company et al. 2015; Simmons et al. 2017), distinguishing between compact and extended objects (Kim & Brunner 2017), selecting observational samples of rare objects based on simulations (Huertas-Company et al. 2018; Lanusse et al. 2018), and visualizing high-level morphological galaxy features (Dai & Tong 2018). These works seek to improve classification of objects into a discreet number of classes, i.e., visual morphologies. Our paper uses CNNs to tackle the different problem of regression, i.e., predict values from a continuous distribution.

Examples of regressing stellar properties in the astronomical ML literature (e.g., Bailer-Jones 2000; Fabbro et al. 2018) train on synthetic stellar spectra and test on real data. Their predicted measurements of stellar properties, e.g., stellar effective temperature, surface gravity, or elemental abundance, can be derived from the available training data set. Our work is novel because we predict metallicity, a spectroscopically determined galaxy property, using only three-color images. Said
another way, it is not necessarily the case that \( Z \) can be predicted from our training data. However, we find that galaxy shape supplements color information in a way that is useful for predicting metallicity.

A study similar to this work is that of Acquaviva (2016), who uses a variety of machine learning methods including RFs, extremely random trees (ERTs), boosted decision trees (AdaBoost), and support vector machines (SVMs) in order to estimate galaxy metallicity. The Acquaviva (2016) data set consisted of a \( z \sim 0.1 \) sample (with \( \sim 25,000 \) objects) and a \( z \sim 0.2 \) sample (with \( \sim 3,000 \) objects), each of which has five-band SDSS photometry (ugriz) available as inputs. These samples are sparsely populated at low metallicities, and they contain smaller fractions of objects with \( Z_{\text{true}} < 8.5 \) than our sample, but are otherwise similarly distributed in \( Z_{\text{true}} \) to ours. Our samples have different sizes because we require SDSS objects to have imaging available, whereas the Acquaviva (2016) criteria impose stronger spectroscopic redshift constraints.

We will first compare RF results, since this technique is common to both of our analyses, and they reveal important differences in our training data. Because outliers are defined differently in both works, we will use the RMSE metric to compare scatter between the two. Acquaviva (2016) obtain RMSE of 0.081 and 0.093 dex when using RFs on the five-band photometry for the \( z \sim 0.1 \) and 0.2 subsamples. Using exactly the same RF approach on a larger sample, while working with only three bands of photometric information, we find RMSE = 0.130 dex. Our scatter is larger than the value reported by Acquaviva (2016) by a factor of \( \sim 1.5 \). This result may partly be explained by the fact that Acquaviva (2016) \( Z_{\text{true}} \) distribution is narrower than for our training data set, or the fact that our data set spans a broader range in galaxy redshift; however, some of this advantage is offset by our larger sample size. Ultimately, it appears that the extra \( u \) and \( z \) bands supply machine learning algorithms with valuable information for predicting metallicity.

Indeed, the \( u \) and \( z \) bands convey information about a galaxy’s SFR and stellar mass (see, e.g., Hopkins et al. 2003). For this reason, it is possible that the RF trained on five-band photometry can estimate \( Z_{\text{true}} \) down to the limit of the FMR, which has very small scatter \( (\sim 0.05 \, \text{dex}) \) at fixed \( M_\star \) and SFR. The \( g, r, \) and \( i \) bands are less sensitive to the SFR, but can still provide some information about the stellar mass, and so our RF and CNN results are more linked to the MZR rather than the FMR.

Regardless of these limitations, our CNN is able to estimate metallicity with \( \Delta Z = 0.085 \, \text{dex} \), which is comparable to the scatter in residuals using the best algorithms from Acquaviva (2016). There is evidence that the morphological information provided by using images rather than photometric data is helping the CNN perform so well: (1) the RMSE scatter decreases with increasing
image resolution, and (2) it identifies edge-on galaxies as lower-$Z_{\text{pred}}$ and face-on galaxies as higher-$Z_{\text{pred}}$ (consistent with observational bias). Gradients in color, or identification of mergers (e.g., Ackermann et al. 2018) may also be helpful for predicting metallicity.

5.6 The mass-metallicity relation

The MZR describes the tight correlation between galaxy stellar mass and nebular gas-phase metallicity. Scatter in this correlation is approximately $\sigma \approx 0.10$ dex in $Z_{\text{true}}$ over the stellar mass range $8.5 < \log(M/\text{M}_\odot) < 11.5$ (Tremonti et al. 2004), where $\sigma$ is the standard deviation of the metallicity and is equivalent to the RMSE for a normal distribution. The MZR at $z = 0$ can be characterized empirically using a polynomial fit:

$$Z = -1.492 + 1.847 \log(M/\text{M}_\odot) - 0.08026 [\log(M/\text{M}_\odot)]^2. \quad (5.3)$$

The physical interpretation of the MZR is that a galaxy’s stellar mass strongly correlates with its chemical enrichment. Proposed explanations of this relationship’s origin include metal loss through blowout (see, e.g., Garnett 2002; Tremonti et al. 2004; Brooks et al. 2007; Davé et al. 2012), inflow of pristine gas Dalcanton et al. (2004), or a combination of the two (e.g., Lilly et al. 2013); however, see also Sánchez et al. (2013). Although the exact physical process responsible for the low (0.10 dex) scatter in the MZR is not known, its link to SFR via the FMR is clear, as star formation leads to both metal enrichment of the interstellar medium and stellar mass assembly.

The FMR connects the instantaneous ($\sim 10$ Myr) SFR with the gas-phase metallicity ($\sim 1$ Gyr timescales; see, e.g., Leitner & Kravtsov 2011) and $M_*$ (i.e., the ~13 Gyr integrated SFR). Our CNN is better suited for predicting $M_*$ rather than SFR, using the $gr$ bands, which can only weakly probe the blue light from young, massive stars. Therefore, we expect the scatter in CNN predictions to be limited by the MZR (with scatter $\sigma \sim 0.10$ dex) rather than the FMR ($\sigma \sim 0.05$ dex). It is possible that galaxy color and morphology, in tandem with CNN-predicted stellar mass, can be used to roughly estimate the SFR, but in this paper we will focus on only the MZR.

5.6.1 Predicting stellar mass

Since galaxy stellar mass is known to strongly correlate with metallicity, and is easier to predict (than, e.g., SFR) from $gr$ imaging, we consider the possibility that the CNN is simply predicting stellar mass ($M_{*,\text{pred}}$) accurately and then learning the simple polynomial transformation in order to estimate metallicity. We can simulate this scenario by training the CNN on $M_{*,\text{true}}$ and then
converting the stellar mass predictions to metallicities using Equation 5.3.

We re-run the CNN methodology to train and predict $M_*$ using the 116,394 available images (out of the 142,145/142,186 original objects that have stellar mass measurements). These results are shown in the left panel of Figure 5.5. From the same subsample as before (minus three objects that do not have $M_*$ estimates), we verify that $M_{*,\text{true}}$ median agrees with the median of $M_{*,\text{true}}$ for values between $9.0 \lesssim \log M_*/M_\odot \lesssim 10.5$. The RMSE scatter in the $M_*$ residuals is $\sim 0.22$ dex, and the NMAD is $\sim 0.20$ dex. The slope of the empirical MZR at $\log(M_*/M_\odot) \sim 10$ is $(0.4 \text{ dex in } Z)/(1.0 \text{ dex in } M_*)$, implying that the CNN might be able to leverage the MZR and predict metallicity to $\sim 0.08$ dex (plus any intrinsic scatter in the MZR, in quadrature).

We use Equation 5.3 and $M_{*,\text{pred}}$ to predict metallicity, which we call $Z_{\text{MZR}}$. In the right panel of Figure 5.5, we compare $Z_{\text{MZR}}$ against $Z_{\text{true}}$. The scatter in residuals $Z_{\text{MZR}} - Z_{\text{true}}$ is 0.12 dex, which is significantly higher than the 0.085 dex scatter reported in Section 5.4. If the MZR alone were mediating the CNN’s ability to estimate from gri imaging, then we would expect the scatter for $Z_{\text{pred}}$ to be greater than for $Z_{\text{MZR}}$; instead we find that the opposite is true. This evidence suggests that the CNN has learned to determine metallicity in a more powerful way than by simply predicting $M_{*,\text{pred}}$ and then effectively applying a polynomial conversion.

### 5.6.2 An unexpectedly strong CNN-predicted mass-metallicity relation

The RMSE = 0.085 dex difference between the true and CNN-predicted metallicities can be interpreted in one of two ways: (1) the CNN is inaccurate, and $Z_{\text{pred}}$ deviates randomly from $Z_{\text{true}}$, or (2) the CNN is labeling $Z_{\text{pred}}$ according to some other hidden variable, and $\Delta Z$ residuals represent non-random shifts in predictions based on this variable. If the first scenario were true, we would expect the random residuals to increase the scatter of other known correlations such as the MZR when predicted by the CNN. If the second were true, we would expect the scatter of such correlations to remain unchanged or shrink. We can therefore contrast the MZR constructed from $Z_{\text{pred}}$ and from $Z_{\text{true}}$ in order to test these interpretations.

In the main panel of Figure 5.6, we plot CNN-predicted metallicity versus true stellar mass. For comparison, we also overlay the Tremonti et al. (2004) MZR median relation and its $\pm 1 \sigma$ scatter (which is $\sim 0.10$ dex). Their empirical median relation (solid black) matches our predicted MZR median (solid red), and the lines marking observed scatter (dashed black) match $Z_{\text{pred}}$ scatter as well (dashed red and violet). Over the range $9.5 \leq \log(M_{*,\text{true}}/M_\odot) \leq 10.5$, the RMSE scatter in $Z_{\text{pred}}$ appears to be even tighter than the observed $\pm 1 \sigma$ (dashed black). The same is true for the NMAD, which is even lower over the same interval.

In the upper panel of Figure 5.6, we present the scatter in both predicted and Tremonti et al.
MZR binned by mass. We confirm that the CNN predicts a MZR that is at most equal in scatter than one constructed using the true metallicity. The stellar mass bins for which our CNN found tighter scatter than the empirical MZR are the same bins that happen to contain the greatest number of examples (9.5 ≤ log(M⋆,true/M⊙) ≤ 10.5); thus, the strong performance of our network at those masses may be due to a wealth of training examples. If our data set were augmented to include additional low- and high-M⋆,true galaxies, then the scatter in the predicted MZR could be even lower overall.

The fact that a CNN trained on only gri imaging is able to predict metallicity accurately enough to reproduce the MZR in terms of median and scatter is not trivial. The error budget is very small: σ = 0.10 dex affords only, e.g., 0.05 dex of scatter when SFR is a controlled parameter plus a 0.03 dex systematic scatter in Ztrue measurements, leaving only ∼ 0.08 dex remaining for CNN systematics, assuming that these errors are not correlated and are added in quadrature. This remaining error budget may barely be able to accommodate our result of RMSE(ΔZ) = 0.085. Interpreting the MZR scatter as the combination of intrinsic FMR scatter, Ztrue systematics, and ΔZ systematics cannot be correct since it assumes that the CNN is recovering the FMR perfectly. As we have discussed previously, it is highly unlikely that the CNN is sensitive to the SFR, and therefore cannot probe the MZR at individual values of the SFR.

If we assume that the error budget for the MZR is not determined by the FMR, then the error “floor” should be 0.10 dex. This is immediately exceeded, as we have found RMSE ≈ 0.10 dex for the predicted MZR without accounting for the fact that Zpred and Ztrue differ by RMSE = 0.085 dex! Consider the case in which all Ztrue values are shifted randomly by a Gaussian noise distribution with σ = 0.085 dex. These shifted values should not be able to reconstruct a correlation without introducing additional scatter unless the shifts were not arbitrary to begin with.

We thus find more evidence that the CNN has learned something from the SDSS gri imaging that is different from, but at least as powerful as, the MZR. One possible explanation is that the CNN is measuring some version of metallicity that is more fundamentally linked to the stellar mass, rather than Zpred as derived from oxygen spectral lines. Another possibility is that the MZR is a projection of a correlation between stellar mass, metallicity, and a third parameter, perhaps one that is morphological in nature. If this is the case, then the Tremonti et al. (2004) MZR represents a relationship that is randomly distributed in the yet unknown third parameter, while our CNN would be able to stratify the MZR according to this parameter (much as the FMR does with the SFR). We are unfortunately not able to identify any hidden parameter using the current CNN methodology, but we plan to explore this topic in a future work.
5.7 Summary

We have trained a deep convolutional neural network (CNN) to predict galaxy gas-phase metallicity using only $128 \times 128$-pixel, three-band ($gri$) JPG images obtained from SDSS. We characterize CNN performance by measuring scatter in the residuals between predicted ($Z_{\text{pred}}$) and true ($Z_{\text{true}}$) metallicities. Our conclusions are as follows:

1. By training for a half-hour on a GPU, the CNN can predict metallicity well enough to achieve residuals characterized by RMSE = 0.085 dex (or outlier-insensitive NMAD = 0.067 dex). These findings may be promising for future large spectroscopy-limited surveys such as LSST.

2. We find that the residual scatter decreases in an expected way as resolution or number of channels is increased, suggesting that the CNN is leveraging both the spatial information about a galaxy’s light distribution and the color in order to predict metallicity.

3. The CNN outperforms a random forest trained on $gri$ photometry if provided images larger than $4 \times 4$ pixels, and is as accurate as a random forest trained on $ugriz$ photometry when given $128 \times 128$ pixel $gri$ images.

4. We find that scatter in the mass-metallicity relation (MZR) constructed using CNN-predicted metallicities is as tight as the empirical MZR ($\sigma = 0.10$ dex). Because predicted metallicities differ from the “true” metallicities by RMSE = 0.085 dex, the only way that the predicted MZR can have such low scatter is if the CNN has learned a connection to metallicity that is more strongly linked to the galaxies’ light distributions than their nebular line emission.

All of the code used in our analysis and for making the figures can be accessed at https://github.com/jwuphysics/galaxy-cnns.

Acknowledgments

SB is supported by NSF Astronomy and Astrophysics Research Program award number 1615657. The authors thank the anonymous referee for useful comments that have improved this paper. The authors thank Andrew Baker, Eric Gawiser, and John Hughes for helpful comments and discussions, and also thank David Shih and Matthew Buckley for the use of their GPU cluster in the Rutgers University Experimental High Energy Physics department. JW thanks Jeremy Howard, Rachel Thomas, and the development team for creating the fastai on-line courses and deep learning library.\footnote{https://github.com/fastai/fastai} JW also thanks Florian Peter for valuable assistance with using the fastai library. This
research made use of the IPython package (Pérez & Granger 2007) and matplotlib, a Python library for publication quality graphics (Hunter 2007).

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is www.sdss.org.

SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatório Nacional / MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

5.8 Appendix: Convolution neural network details

5.8.1 Residual neural network architecture

CNNs are divided into “layers” that compute the convolutions of filters, or kernels, with each of the inputs. A Rectified Linear Unit (ReLU) activation function is applied to the convolved output (ReLUs have been shown to propagate information about the relative importances of different features, and are effective for training deep neural networks; Nair & Hinton 2010). In a residual CNN, multiple convolutional layers containing small (e.g., 3 × 3) filters are arranged sequentially, and a final “shortcut connection” adds the first layer, unaltered, to the final output (before the final ReLU activation; see, e.g., Figure 2 of He et al. 2015). Such combinations of convolutions, activations, and shortcuts are called residual building blocks.

We use a 34-layer residual convolutional neural network with the architecture described by He
et al. (2015), and implemented using PyTorch (version 0.3.1; Paszke et al. 2017) provided by the fastai framework (version 0.7; Howard et al. 2018). A full description of the architecture’s layers can be found in the online PyTorch documentation, but we also provide a brief overview below.

The resnet can be separated into three “layer groups” that roughly correspond to the levels of abstraction able to be learned by the network. Once an image is fed into the network, e.g., a three-channel $128 \times 128$ SDSS image, it is effectively converted into activation maps that depend on how well the filters match the input image. These maps are further convolved with the next layer of filters, and this process continues until the last layer group is reached. The activation maps are periodically downsampled, max pooled, or average pooled, which effectively halve the map sizes in each spatial dimension (for more about pooling layers in CNNs, see Scherer et al. 2010). The first two layer groups comprise multiple residual building blocks, and the final layer group consists of two fully connected linear layers, with a ReLU activation after the first and no activation after the second. The last fully connected layer does not have an activation function because we are working on a regression problem, and so the weights trained in that layer should be tuned to predict metallicity in the desired range.

5.8.2 Adaptive learning rates

Neural network performance tends to depend dramatically on choice of hyperparameters. After an image is fed forward and the residual (= prediction − true) value is computed, relative contributions of error are propagated backward through the network, starting from the final layer and ending at the first layer. Using gradient descent of the loss (in our case, the root mean squared error), the network layers’ weights are adjusted according to their error contributions multiplied by the learning rate. The process of computing errors from known images and metallicities and updating weights is called training, and when all of the training data set has been used to adjust network weights, a training epoch is completed.

The learning rate can be thought of as the step size during each weight update. A high learning rate allows the network to improve quickly, but at some point the large step size may become too coarse for additional optimization; conversely, a low learning rate might allow the network to traverse every bump and wiggle in the error landscape, but might also take a very long time to reach convergence (or get stuck indefinitely in a local minimum). We first select a learning rate by using the method described by Smith (2015). Over a number of epochs, the learning rate is reduced (or annealed) as the network needs to make more fine-tuned updates in order to achieve better accuracy. We use a method called cosine annealing, during which the learning rate is annealed with the cosine

---

function continuously over individual (or batches of) training examples. It has been shown that if the learning rate is annealed and then restarted after one or more epochs, the network is less likely to get caught in local minima and overall accuracy is improved. We refer to Loshchilov & Hutter (2016) for details about employing cyclical learning rates and gradient descent with restarts, which are implemented in our CNN.

5.8.3 Optimization techniques and preventing overfitting

Losses are computed for small “batches” of training examples at a time. Gradients that minimize each batch are expected to be noisier than gradients that are computed to optimize the entire training data set loss. This technique of stochastic gradient descent helps prevent the CNN from overfitting training data, which is a possibility given the huge number of parameters in a deep CNN. We also use weight decay, another commonly used regularization technique, which adds a decay term proportional to each layer weight during the update step of training (e.g., Krogh & Hertz 1992).

As the learning rate is annealed with increasing numbers of batches, the weight updates are also expected to diminish. The Adam optimizer adaptively smooths the gradient descent in a way that depends on previous gradients (Kingma & Ba 2014). Adam is analogous to rolling downhill with gravitational potential, momentum, and friction terms (whereas gradient descent would be analogous to movement dependent only on the potential at its given time step). For caveats about combining weight decay and Adam, see Loshchilov & Hutter (2017), whose updated algorithm is implemented in fastai.

We implement batch normalization (BN), a technique developed to fix a problem that previously caused deep networks to train extremely slowly (Ioffe & Szegedy 2015). To briefly recap the issue: updates to the layer weights depend on the contribution of the backpropagated error, but when the number of layers is large (i.e., in a deep CNN), the contribution becomes vanishingly small. BN is simply the rescaling of each input to the nonlinear activation so that it has mean of zero and standard deviation of unity (i.e., subtract the mean and divide by the standard deviation). A new choice of hyperparameter is the batch size, or the number of training examples from which the mean and variance are calculated; we choose 256 based on tests of performance in ten training epochs.

Dropout is a method of disabling a random subset of connections after linear layers in a network in order to improve the network’s generalizability (Hinton et al. 2012). The ensemble of learned gradients is less prone to overfit the training data set because the network is forced to discard random (and potentially valuable) information. The resulting network is better able to, e.g., learn subtle differences in the data that would otherwise be ignored when more obvious features dominate the gradient descent process. We apply dropout layers only to the final fully connected layers in our
5.8.4 Training the network

We initialize the network using weights that have been pretrained on the 1.7 million example ImageNet data set (which contains 1000 classes of objects; Russakovsky et al. 2014). The network should more quickly optimize toward the global minimum loss through transfer of low-level features already learned in earlier layers of the network (known as transfer learning; see, e.g., Pan & Yang 2010).

We train only the final layer group for the first two epochs, which can be accomplished by not updating weights in the first two layer groups. The learning rate is initially set to 0.1 and then annealed according to a cosine schedule over an epoch (and then restarted to 0.1 at the beginning of the following epoch). We then allow the updating of weights in all layer groups while setting the learning rates to 0.01, 0.03, and 0.1 for the first, second, and last layer groups, respectively. This approach allows the final group of fully connected layers to respond strongly to different types of training examples (e.g., galaxies that appear very different in gri imaging) while the earlier layers are trained very slowly in order to preserve their more general features. Using these layered learning rates, we train the full network using a cosine annealing schedule that spans one, one, two, and then four epochs (where the different learning rates are annealed by the same amount). Using this combination of learning rate schedules, we find that our network quickly achieves low training losses (RMSE $\sim 0.085$ on validation data sets). Altogether, only ten epochs of training are needed, which takes under 30 minutes on a GPU. We find that further training does yield some gains, but this improvement plateaus around RMSE $\sim 0.083$ and takes many more hours.

5.8.5 Data augmentation

Nearly all neural networks benefit from larger training samples because they help prevent overfitting. Beyond the local Universe, galaxies are seen at nearly random orientation; such invariance permits synthetic data to be generated from rotations and flips of the training images (see, e.g., Simonyan & Zisserman 2014). Each image is fed into the network along with four augmented versions, thus increasing the total training sample by a factor of five.

This technique is called data augmentation, and is particularly helpful for the network to learn uncommon truth values (e.g., in our case, very metal-poor or metal-rich galaxies). Each augmented image is fed-forward through the network and gradient contributions are computed together as part
of the same batch. A similar process is applied to the network during predictions, which is known as test-time augmentation (TTA), whereby synthetic images are generated according to the same rules applied to the training data set. The CNN predicts an ensemble average over the augmented images, which tends to further improve RMSE by a few percent. We use the default hyperparameters in the fastai library.
Figure 5.1 SDSS imaging with predicted and true metallicities from the test data set. Five examples are shown from each of the following categories: (a) lowest predicted metallicity, (b) lowest true metallicity, (c) highest predicted metallicity, (d) highest true metallicity, (e) most under-predicted metallicity, (f) most over-predicted metallicity, and (g) a set of randomly selected galaxies.
Figure 5.2 Distributions of the true (black) and predicted (red) galaxy metallicities. Note that the bin widths are different for the two distributions. See text for details.
Figure 5.3 Bivariate distribution of true galaxy metallicity ($Z_{\text{true}}$) and CNN prediction ($Z_{\text{pred}}$) is shown in the main panel. Overlaid are the median predicted metallicity (solid red line), RMSE scatter (dashed red lines), and NMAD scatter (dashed violet lines), in bins of $Z_{\text{true}}$. The solid black line shows the one-to-one relation. The distribution of residuals ($Z_{\text{pred}} - Z_{\text{true}}$) is shown in the inset plot. In the upper panel, we again show the binned scatter, where the size of each marker is proportional to the number of galaxies in that bin. Each horizontal line corresponds to the average scatter over the entire test data set (and the global value indicated in the upper panel legend).
Figure 5.4 The effects of image resolution and color on CNN performance. Red and violet circular markers indicate scatter in the residual distribution ($\Delta Z$) measured using RMSE and NMAD, respectively, for $gr$ imaging. (Each point is analogous to the horizontal lines shown in Figure 5.3.) We also show predictions from a random forest algorithm as open triangle markers, and constant $\langle Z_{true} \rangle$ predictions as open square markers. Large, labeled squares indicate the scatter for images consisting of only $r$-band imaging and a combination of $g$- and $r$-bands.
Figure 5.5 In the left panel, we plot the CNN predicted galaxy stellar mass against true stellar mass. Colors and marker or line styles are the same as in Figure 5.3. In the right panel, we compare the predicted stellar mass converted to metallicity, assuming the Tremonti et al. (2004) MZR, against the true metallicity. These findings indicate that using the empirical MZR and CNN-predicted $M_{\star, \text{pred}}$ yields poor results, unlike what we have observed in Figure 5.3.
Figure 5.6 In the main panel, the predicted MZR comparing true $M_*$ against CNN-predicted $Z_{\text{pred}}$ is shown in grayscale. The running median (solid red) and scatter (dashed red and violet) are shown in 0.2 dex mass bins. For comparison, we also show the Tremonti et al. (2004) observed median and scatter (solid and dashed black lines, respectively), which are binned by 0.1 dex in mass. In the upper panel, we show the scatter in the predicted and empirical MZR. The standard deviation of the scatter for the empirical MZR is shown as a dashed black line, while the red and violet circles respectively show RMSE and NMAD for the predicted MZR. Marker sizes are proportional to the number of galaxies in each stellar mass bin for the test data set. Global scatter in the CNN-predicted MZR appears to be comparable to, or even lower than, scatter in the empirical MZR.
Chapter 6
Summary

Galaxy evolution models are becoming more sophisticated and can explain a wide range of observed phenomena. Now that numerical simulations and semi-analytic models can match the mean relations and statistical properties of observed galaxies, the challenge becomes for theory to explain galaxy evolution in more extreme scenarios. Rare galaxy populations, such as low-z LBG analogs or galaxies belonging to massive, distant clusters, are valuable laboratories for testing the limits of theoretical understanding. Also worth investigating is whether or not machine learning algorithms can be valuable additions to theoretical models and/or observational data sets. In the sections below, I reiterate the questions that my thesis work has begun to address. I provide brief summaries for each chapter, and also offer potential avenues for future work.

6.1 How do molecular gas, dust, and star formation properties evolve with redshift for galaxies in the most overdense environments?

Conclusions. In Chapter 2 (Wu et al. 2018), I introduced Herschel FIR photometric observations for probing dust-obscured star formation in four massive, SZE-selected galaxy clusters between $0.3 \lesssim z \lesssim 1.1$. The number of Herschel detections (crossmatched to cluster members confirmed via optical spectroscopy) in each cluster increases with redshift, as do the star-forming fraction of cluster members and the star-forming fraction within half of each cluster’s virial radius. The SFR appears to depend on environmental density, i.e., SFR increases with increasing clustercentric radius. Furthermore, these SFR-density trends are stratified by redshift, such that the highest-redshift cluster (J0546) hosts galaxies with the highest SFRs, and the lowest-redshift cluster (J0235) hosts galaxies with the lowest SFRs. After the cluster members’ SFRs are normalized by stellar mass, yielding the sSFR, this radial correlation disappears. However, the redshift evolution persists, and supports the existence of a Butcher-Oemler effect in massive clusters at intermediate redshifts (Butcher & Oemler 1978; Haines et al. 2009).

---

1AGN and dust-obscured star formation are jointly modeled from the Herschel and ALMA photometry using Kirkpatrick et al. (2015) empirical template SEDs.
The two more distant clusters are also spectroscopically targeted with ALMA in order to simultaneously detect CO($4-3$) and [C I] line emission as well as rest-frame $\lambda = 600 \mu$m dust continuum emission. Six (five) cluster members are (newly) detected from ALMA line emission, including four that are detected in both CO and [C I]. These systems host massive molecular gas reservoirs, log$(M_{\text{mol}}/M_\odot) = 10.4 - 10.8$, assuming a Milky Way-like CO-to-$H_2$ conversion factor (Bolatto et al. 2013), and large molecular gas mass fractions, $f_{\text{gas}} \approx 0.1 - 0.3$, that are more comparable to $z \sim 1$ field galaxies than other cluster galaxies (e.g., Tacconi et al. 2013). Dust continuum emission is only detected for three galaxies in the highest-$z$ cluster, and implies total cold ISM (i.e., combined dust, atomic gas, and molecular gas) masses of $\sim 10^{11} M_\odot$ under the assumption of a constant dust-to-gas conversion (Scoville et al. 2016). On average, gas depletion timescales are relatively long $\tau_{\text{dep}} \approx 2$ Gyr, signifying that mass assembly in our $z \sim 1$ cluster members proceeds inefficiently compared to star-forming field galaxies at the same redshifts ($\tau_{\text{dep}} \lesssim 1$ Gyr; e.g., Genzel et al. 2015).

Future work. The optical spectroscopy selection criteria used to confirm cluster galaxy redshifts was identified to be a source of potential bias to the study. Sifón et al. (2013) selected red and bright cluster galaxies for spectroscopic follow-up, resulting in a low overall fraction ($= 17/285$) of emission-line (and thus optically unobscured star-forming) systems. In order to alleviate the under-representation of line emitters in the sample, I have proposed for optical spectroscopic observations with the Fabry-Pérot tuneable filter in the Southern African Large Telescope (SALT) Robert Stobie Spectrograph (RSS; Rangwala et al. 2008). By virtue of being tightly packed in physical and redshift space, many cluster galaxies can be efficiently observed with the Fabry-Pérot instrument over the 8′ SALT FOV. I have obtained 22 240 second exposures of the J0102 field, spanning 5 Å intervals between 6930 and 7035 Å, and similar observations between 7660 and 7765 Å for the J0546 field. Fabry-Pérot observations were taken in the low-spectral resolution mode between 2015 December 5 and 2016 October 27. In Figure 6.1, I show the SALT Fabry-Pérot observing windows for J0102 and J0546. Night-sky emission lines at the observed wavelengths (e.g., Hanuschik 2003) appear as concentric rings due to the geometry of the interferometer and make data reduction challenging. In order to extract [O II] spectral line detections, I will need to modify the Python-based SALT Fabry-Pérot data reduction pipeline (which is best suited for medium-resolution observations; Mitchell et al. 2015). The next step will be to constrain wavelength and astrometric solutions for the spectral line cubes and then subtract the night-sky line artifacts. Observations will be normalized and approximately flat-fielded (using stars in the FOV) so that channel-to-channel variations do not incur

---

2I also have obtained 11 200 second exposures of J0102 in 10 Å increments between 6930 and 7030 Å. This pilot run was taken as part of a Rutgers SALT director’s discretionary time allocation. A broad-line AGN belonging to the J0102 cluster is quickly detected in the pilot study, although a full analysis has not been done on the data set.
spurious signals. I plan to extract sources using, e.g., a targeted search (at the positions of *Herschel* or ALMA detections), or a full search over the entire data cube (using automated software; e.g., Serra et al. 2015).

As a new direction for research, I will investigate other populations of star-forming cluster galaxies that may have been missed in previous studies. Luminous compact blue galaxies (LCBGs) are relatively low-mass ($M_\star \sim 5 \times 10^9 M_\odot$; Guzmán et al. 2003) star-forming galaxies that are occasionally found in the field (e.g., Koo et al. 1994) and commonly sighted in massive clusters (Koo et al. 1997; Crawford et al. 2006). Although the field LCBGs have diverse cold gas properties and can evolve along different evolutionary trajectories (Pisano et al. 2001), cluster LCBGs are proposed to be the progenitors of quiescent dwarf elliptical galaxies often found in present-day overdense environments (Crawford et al. 2016). Visual inspection of the J0102 *HST* imaging reveals numerous blue, compact galaxies, many of which may satisfy the LCBG criteria and belong to the massive, merging cluster. LCBG cluster member candidates can also be crossmatched with *Herschel* or SALT [O II]-emitting sources. Because LCBGs may be important contributors to the cosmic SFR at redshifts $0.4 < z < 1$ (Guzmán et al. 1997), it would be extremely important to account for them in any studies of star formation in massive clusters. LCBGs in $z \sim 0.5$ clusters are distributed in a shell-like structure, suggesting recent infall (Crawford et al. 2006). I propose to investigate the projected distribution of (candidate) LCBGs in J0102 in order to determine what (e.g., morphological) impacts are felt by compact star-forming systems near the merging front, compared to those at the cluster outskirts. Another future direction would be to obtain high-resolution (adaptive optics-assisted, or *HST*) imaging for J0546. LCBG samples could then be compared for both cluster environments, and any overall differences in the two populations (i.e., in number density) might be attributed to the clusters’ different dynamical states.

6.2 What are the multiphase ISM properties of low-redshift LBG analogs?

**Conclusions.** In Chapter 3 (Wu et al. 2019), I reported VLT/SINFONI integral field spectroscopy centered on the NIR Paα spectral line for six LBG analogs, including one that consists of two interacting systems (J2116a and J2116b). Ionized gas dynamics measured from moment maps and three-dimensional kinematic modeling are characterized by high velocity dispersions ($\sigma \sim 80 \text{ km s}^{-1}$) and low $v/\sigma$, both of which are much more common in the distant Universe than at low redshifts (e.g., Wisnioski et al. 2015). For the optical depths probed by Hα and Hβ, the dust attenuation is consistent with a well-mixed model of ionizing (responsible for hydrogen recombination lines) and non-ionizing stars (responsible for most of the UV continuum); however, Paα measurements support
a nebular-to-continuum color excess ratio closer to the Calzetti et al. (2000) value of 0.44, indicating an additional component of dust toward H II regions. Observed hydrogen and helium recombination line ratios are compared to photoionization models in order to constrain the effective blackbody temperatures of ionizing stars, which mostly lie in the standard range for typical massive stellar populations \((32,000 \leq (T_{\text{eff}}/K) \leq 40,000)\); e.g., Förster Schreiber et al. 2001), but in the notable case of J2104, higher stellar \(T_{\text{eff}}\) may indicate the presence of massive star binaries or a top-heavy IMF.

Multiple \(H_2\) 1-0 S ro-vibrational transitions are detected in emission at high SNR for 6/7 galaxies, marking the first time that warm \((T_{\text{ex}} \approx 2000\) K) \(H_2\) has been detected for LBGs or their low-z analogs. The combination of warm \(H_2\) properties, occasional presence of [Fe II], and low \([O I]/[C II]\) FIR line ratios (Contursi et al. 2017) evidences that such emission likely originates from photon-dominated regions (PDRs). Resolved \(H_2/P\alpha\) line ratios also suggest that star formation feedback is the primary excitation mechanism of nebular gas for the LBG analog sample (a conclusion further supported by relatively low warm-to-cold molecular gas masses; see, e.g., Section 3.5). Finally, for an interacting system (J1434), \(H_2\) ro-vibrational emission is enhanced in the companion source characterized by weak \(P\alpha\), opening up the possibility that tidal shocks or merger-driven starbursts are responsible for heating the molecular gas.

**Future work.** Although this previous work details the resolved properties of LBG analogs' warm and ionized ISM phases, high-resolution observations of the cold gas in LBG analogs (or their high-z cousins) are lacking. I have proposed to observe CO(3 – 2) emission on sub-kpc (< 0.2") scales using ALMA for our sample of LBG analogs.\(^3\) The immediate benefits of these high-resolution observations are twofold. First, the global and resolved Kennicutt-Schmidt relations can be measured for LBG analogs, which can be compared to other low-z and high-z observations (e.g., Leroy et al. 2013; Tacconi et al. 2013) and strongly constrain theoretical models that purport to explain turbulent gas collapse and star formation efficiency (e.g., Krumholz et al. 2018). Figure 6.2 shows theoretical predictions compared to resolved and global Kennicutt-Schmidt relations, where J2104’s resolved CO observations have been included to highlight the constraining power of such observations. Second, measurements of CO velocity dispersion can be used to distinguish between physical processes that drive turbulence in galaxies. Krumholz & Burkhart (2016) report that if galaxy-scale turbulence is primarily driven by star formation feedback, then SFR is expected to scale strongly with velocity dispersion independent of the gas fraction: \(\text{SFR} \propto \sigma^2\). If gravitational

\(^3\)There are ALMA archival observations at this resolution for J2104 (Program ID 2016.1.01265.S, PI: Gonçalves.), but not for any of the other LBG analogs.
instabilities are most important for driving turbulence, then SFR scales more weakly with dispersion and does depend on gas fraction: SFR $f_{\text{gas}}\sigma$. Figure 6.3 shows that which scenario holds can be seen most easily in systems with moderate rates of star formation (SFR $\sim 10 - 100 M_\odot \text{yr}^{-1}$) and high velocity dispersions ($\sigma > 50 \text{ k s}^{-1}$) — precisely the region of parameter space that hosts LBGs and LBG analogs.

### 6.3 How do LBG analogs compare to other systems in terms of kinematic scaling relations?

**Conclusions.** I have investigated the stellar mass Tully-Fisher relation (TFR) and other kinematic scaling relations for the LBG analog sample. For an ordinary star-forming disk galaxy, the stellar mass TFR connects $M_*$ with the total mass probed by $V_{\text{rot}}$. For systems with significant disordered motions, an analogous scaling relation between $S_{0.5} \equiv (0.5V_{\text{rot}} + \sigma^2)^{1/2}$ and $M_*$ exists. Indeed, the kinematics of LBG analogs are not adequately described by the $z \sim 0$ stellar mass TFR; instead, they occupy a similar locus as $z \sim 2$ systems (albeit the ones with less rotation support). Departures from the stellar mass TFR correlate strongly with $V_{\text{rot}}/\sigma$ independent of redshift, implying that some of the previously reported TFR redshift evolution may in fact be controlled by $V_{\text{rot}}/\sigma$. LBG analogs also deviate toward lower $V_{\text{rot}}$ and $S_{0.5}$ at given baryonic mass $M_{\text{bar}} \equiv M_* + M_{\text{gas}}$, when compared to the baryonic TFR measured for $0 \lesssim z \lesssim 2$ star-forming galaxy samples. This result suggests that the compact nature of LBG analogs is linked to their lower kinematic support, and that the enclosed mass probed by $V_{\text{rot}}$ and $S_{0.5}$ is more baryon-dominated than the case for more typical $z \sim 0$ galaxies.

**Future work.** The proposed ALMA observations mentioned in the previous section would also be valuable for studying kinematic scaling relations. High resolution CO(1 − 0) observations can be useful for deriving the spatially resolved CO TFR (up until now, only the integrated CO TFR has been measured; e.g., Tiley et al. 2016; Isbell et al. 2018). Molecular gas tends to be more settled and kinematically ordered in disks than, e.g., H I (which is more likely to be disturbed; Verheijen 2001), or ionized gas (e.g., Cresci et al. 2009; Miller et al. 2012). Moreover, CO gas is less disrupted by galaxy interactions or mergers while still being detectable out to large radii (e.g., Pereira-Santaella et al. 2018; Xue et al. 2018). Therefore, comparisons between LBG analogs and other low-$z$ samples using the CO TFR may be more robust than via measurements of other emission lines.
6.4 How do optical morphological features in local star-forming galaxies connect to stellar mass and gas-phase metallicity?

Conclusions. Convolutional neural networks (CNNs) are supervised machine learning algorithms that can be trained to process images and make predictions or classifications. In Chapter 5 (Wu & Boada 2019), I have trained a deep (or many-layered) CNN in order to accurately predict gas-phase metallicity using only three-band (gri) SDSS imaging. After splitting the parent SDSS gri imaging crossmatched with spectroscopic data into train/validation/test data subsets (of approximate sizes 75000/20000/20000), I optimized a deep CNN to predict metallicity $Z_{\text{pred}}$ derived from optical emission-lines (referred to as $Z_{\text{true}}$; Brinchmann et al. 2004; Tremonti et al. 2004) to within $\text{RMSE} = 0.085$ dex. Various tests suggest that this trained CNN is robust at estimating $Z_{\text{pred}}$. Mixing up the train/test splits or artificially adding unbiased Gaussian-distributed scatter to $Z_{\text{true}}$ does not corrupt training or final results. Ablation studies, in which pixels are binned together to reduce image resolution, or color channels sequentially removed, demonstrate that both colors and image resolution are necessary for accurate predictions. The deep CNN also outperforms another commonly used machine learning algorithm, a random forest, by a wide margin.

A MZR reconstructed using $Z_{\text{pred}}$ and independently derived stellar masses (Kauffmann et al. 2003) yields no more scatter than is in the original empirical MZR (i.e., 0.10 dex; Tremonti et al. 2004). The extremely low scatter in the $Z_{\text{pred}}$-constructed MZR cannot solely be due to connections between morphological features and $M_{\text{star}}$, as shown in Figure 5.5. Therefore, the multi-color morphological features seen in optical imaging$^4$ may provide information relevant for predicting metallicity beyond what can be probed via optical emission lines or via stellar mass modeling.

Future work. The success of using CNNs to directly predict metallicities shows that deep learning can complement and multiply scientific returns of spectroscopically limited surveys such as LSST. One method for expediting the optimization process on other data sets is to use transfer learning, where a model trained for one particular task (e.g., using SDSS images to predict gas-phase metallicity) is re-trained on another data set. A transfer learning extension to this work might be to re-train the CNN to output $Z_{\text{pred}}$ using images from Pan-STARRS (Panoramic Survey Telescope and Rapid Response System; Chambers et al. 2016) where the SDSS and Pan-STARRS surveys overlap. The upgraded model can then produce $Z_{\text{pred}}$ catalogs for star-forming galaxies across the entire Pan-STARRS $3\pi$ steradian survey field (i.e., three quarters of the sky). A similar process can be repeated with the deeper (and ongoing) Dark Energy Survey imaging (Abbott et al. 2018), which

---

$^4$ Examples of relevant morphological features could include dust lanes in spiral arms, color gradients throughout the disk, or blue star-forming knots.
overlaps with the SDSS Stripe 82.

Another promising extension is to use optical imaging to predict other kinds of spectroscopic quantities. For example, CNNs may be able to powerfully estimate H I masses ($M_{\text{HI}}$), since cold gas properties appear to correlate with morphological and color information (e.g., Kannappan et al. 2013; Bahé et al. 2016). After crossmatching SDSS optical imaging with extragalactic H I catalogs (e.g., from ALFALFA, the Arecibo Legacy Fast ALFA Survey; Haynes et al. 2018), I can train CNNs to learn H I masses directly from optical imaging. This study can also help identify the impact of environmental variables on H I properties, since optical image cutouts can include interacting systems or (projected) neighboring sources. In the near future, CNNs can be trained using the rich optical imaging data sets that will supplement the LADUMA survey (see Appendix A for details), in order to build morphological priors for high-z LADUMA detections. Trained CNN models can assist future H I surveys by identifying which optical sources are most likely to be H I-rich, have high $f_{\text{gas}}$, or even have high ratios of H I-to-H$_2$ mass.
Figure 6.1 Sky emission lines (red; Hanuschik 2003) and SALT blocking filter transmission curves (blue; Rangwala et al. 2008) shown against observed wavelength. Redshifted [O II] observing windows for J0102 (at $\lambda \approx 7000$ Å) and J0546 (at $\lambda \approx 7700$ Å) are shown, each of which spans $\pm 2000$ km s$^{-1}$ in the cluster velocity frame.
Figure 6.2 SFR surface density ($\Sigma_{\text{SFR}}$) vs gas mass column density ($\Sigma_g$). Star formation law predictions from gravitational instability-driven + feedback (G+F) turbulence models, parameterized by orbital timescale $t_{\text{orb}}$ (Krumholz et al. 2018), are shown as solid green lines, and a star formation feedback-only model (Ostriker & Shetty 2011) is shown as a dashed black line. Global measurements for available LBG analogs (blue markers; Contursi et al. 2017), nearby low-$z$ disk galaxies (purple stars; Leroy et al. 2013), and high-$z$ star-forming galaxies (gray triangles; Genzel et al. 2010) are shown for comparison. Also shown are resolved $\Sigma_{\text{SFR}}$ and $\Sigma_g$ for J2104 ($-210358$) in red.
Figure 6.3 Velocity dispersion ($\sigma$) and star formation rate (SFR) for LBG analogs compared with Krumholz & Burkhart (2016) theoretical models. (Left) $\sigma$ measured from unresolved PdBI CO(1−0) observations (blue square) and resolved ALMA CO(3−2) observations (red square) of J2104, which has $f_{\text{gas}} = 0.5$. (Right) $\sigma$ measured from unresolved PdBI CO observations are shown for two systems, J1434 ($=143417$; $f_{\text{gas}} = 0.3$), which agrees with the gravity-driven turbulence model, and J0213 ($=021348$; $f_{\text{gas}} = 0.6$), which favors the feedback-driven turbulence model. For both panels, gravity-driven scaling relations are based on measured gas fractions and maximum velocities, and feedback-driven models assume a marginally unstable gas disk (for which the gas Toomre 1964 parameter $Q_g = 1.0$).
Akaike, H. 1974, IEEE Transactions on Automatic Control, 19, 716
Babcock, H. W. 1939, Lick Observatory Bulletin, 19, 41
Baugh, C. M. 2006, Reports on Progress in Physics, 69, 3101
Bentley, J. L. 1975, Commun. ACM, 18, 509
Bertschinger, E., & Gelb, J. M. 1991, Computers in Physics, 5, 164
Bethe, H. A. 1939, Physical Review, 55, 434
—. 1990, Reviews of Modern Physics, 62, 801
Bromm, V. 2013, Reports on Progress in Physics, 76, 112901
Burbidge, E. M., Burbidge, G. R., Fowler, W. A., & Hoyle, F. 1957, Reviews of Modern Physics, 29, 547
Dabrowski, I. 1984, Canadian Journal of Physics, 62, 1639
Draine, B. T. 2011, Physics of the Interstellar and Intergalactic Medium


Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013a, PASP, 125, 306

—. 2013b, PASP, 125, 306


Griffin, M. J., & Orton, G. S. 1993, Icarus, 105, 537
Hoyle, B. 2016, Astronomy and Computing, 16, 34
Hunter, J. D. 2007, Computing In Science & Engineering, 9, 90
Jeans, J. H. 1902, Philosophical Transactions of the Royal Society of London Series A, 199, 1
Kennicutt, Jr., R. C. 1998a, ARA&A, 36, 189


McQuinn, M. 2016, ARA&A, 54, 313


Mo, H., van den Bosch, F. C., & White, S. 2010, Galaxy Formation and Evolution
Morganti, R., Frieswijk, W., Oonk, R. J. B., Oosterloo, T., & Tadhunter, C. 2013, A&A, 552, L4
Osterbrock, D. E. 1989, Astrophysics of gaseous nebulae and active galactic nuclei


Pan, S. J., & Yang, Q. 2010, IEEE Trans. on Knowl. and Data Eng., 22, 1345


Peebles, P. J. E. 1980, The large-scale structure of the universe


Pérez, F., & Granger, B. E. 2007, Computing in Science and Engineering, 9, 21


Scherer, D., Müller, A., & Behnke, S. 2010, in Artificial Neural Networks–ICANN 2010 (Springer), 92–101
Serra, P., Jurek, R., & Flöer, L. 2012, PASA, 29, 296
Sunyaev, R. A., & Zel’dovich, Y. B. 1972, Comments on Astrophysics and Space Physics, 4, 173


Appendix A

LADUMA and the search for neutral atomic hydrogen at high redshift

Looking At the Distant Universe with the MeerKAT Array (LADUMA) is an extragalactic radio-wavelength survey that aims to detect neutral atomic hydrogen H I between $0 \leq z \leq 1.45$ (Blyth et al. 2016). The survey uses the MeerKAT interferometer, which consists of 64 antennas that are 13.5 m in diameter each. LADUMA requires two sets of correlators on MeerKAT: the L-band receiver is suited for spectroscopic observations between 900 – 1670 MHz, and the UHF receiver is sensitive in the frequency range 580 – 1015 MHz. Because a telescope array’s field of view (FOV) increases with observed wavelength $\propto \lambda_{\text{obs}}^2$, the LADUMA FOV will grow with decreasing frequency of observation. The MeerKAT FOV subtends 0.9 deg$^2$ at 1420 MHz, 2.2 deg$^2$ at 900 MHz, and 5.4 deg$^2$ at 580 MHz, which correspond to H I detections at $z = 0, 0.58, \text{and } 1.45$ respectively.

LADUMA is expected to detect thousands of H I sources with its 333 hr L-band and 3091 hr UHF-band time allocation (e.g., Obreschkow et al. 2009; Maddox et al. 2016). In the context of current H I surveys, such as the COSMOS H I Large Extragalactic Survey (CHILES; Fernández et al. 2013) sensitive to H I at $z < 0.5$ (e.g., Fernández et al. 2016), or the HIGHz Arecibo survey (which probes $z < 0.25$; Catinella & Cortese 2015), LADUMA breaks new ground by exploiting next-generation instruments and by being the first survey to characterize neutral hydrogen gas emission to $z \sim 1.4$.

New information about the H I content in galaxies over the past nine billion years of cosmic evolution will be key for understanding galaxy evolution and formation. LADUMA will measure the H I masses of galaxies in a broad range of environments and redshifts in the the survey field (pointing center $03^h32^m30.4s -28:07:57$), providing constraints on how neutral atomic gas content varies with external properties. Comparison with rich ancillary data sets in the target field (e.g., Jarvis et al. 2013; Vaccari et al. 2016) will be valuable for understanding the joint evolution of $M_*$, SFR, metallicity, molecular gas, and ISM properties. H I sources will also constrain the baryonic

---

1The full LADUMA survey volume will take on a vuvuzela shape due to its quadratically growing FOV with redshift. See Holwerda et al. (2012) and Blyth et al. (2016) for details about the survey strategy.
Tully-Fisher relation (whereby kinematics are compared against combined gas and stellar mass; e.g., McGaugh et al. 2000) and offer insights on the dark matter-baryon connection in galaxies.

In order to achieve these ambitious science goals, high-quality spectral line data cubes must be constructed from the wide-field, wide-band MeerKAT observations. A spectral line cube requires precise continuum subtraction, which in turn depends on reliable calibration. Once H I cubes are generated, it remains a computational challenge to search for and extract spectral lines from the massive survey volume. In Section A.1, I will discuss a MeerKAT/LADUMA calibration pipeline, and also present work on testing continuum subtraction techniques. The calibration and continuum-subtraction work was done during my visit to South Africa in mid-2018. In Section A.2, I will present a parallelized approach toward H I extraction using a suite of source finding algorithms. This segment of work began during my visit to South Africa in late 2016 and was supported by a United States Agency for International Development (USAID) Research and Innovation Fellowship.

A.1 MeerKAT interferometric observations: calibration, imaging, and continuum subtraction

A.1.1 An overview of radio interferometry calibration

Radio interferometry relies on synthesizing a large-aperture “telescope” using multiple smaller telescope array elements in order to achieve high angular resolutions when observing an astronomical source (see, e.g., excellent overviews by Taylor et al. 1999 and Condon & Ransom 2016). Interferometers correlate received signals in order to map complex visibilities (composed of amplitude and phase components). In other words, they map the Fourier transformed instrumental response to the astronomical source. Two telescope receivers pointed to the same source may record different signals because of differences in geometric path length (the known vector separation between antennas projected along the line of sight — i.e., toward the phase center) and because of other effects that corrupt measurements. It is not trivial to recover the original source properties, since the interferometer only samples a limited portion of the uv plane (i.e., the Fourier transformed sky plane). Corruptions can occur when the signals propagate through a turbulent atmosphere or are distorted by instrumental errors from telescope receivers.

A radio astronomer’s job is to recover the original source properties as accurately as possible. Although the atmosphere and telescope’s electronics impose a thermal noise floor, in practice, interferometer imaging is dominated by other sources of noise. For example, incorrectly assumed antenna
positions can cause errors in Fourier inversion, particularly as the Earth rotates beneath the array and changes the geometry over time. Phase and amplitude distortions because of fluctuations in the sky, terrestrial/satellite radio frequency interference (RFI), and mischaracterization of the telescope primary beam\(^2\) across the sky can result in a Fourier transformed image dominated by systematic effects rather than thermal noise.

To recover the source’s correct phase and amplitude (which can be done by characterizing the array’s complex gain), and spectral variations (bandpass), calibrator sources with constant, known properties can be observed. By comparing the known calibrator signal with observed output, gain and bandpass solutions can be derived, along with any geometric delay errors. Calibrator solutions can be directly applied to the target field or source, and this process is called *cross-calibration*. Calibrator sources should be observed near on sky to the scientific target, such that the telescope is probing approximately the same patch of atmosphere, and near in time, so that solutions remain valid for when the target is observed. If a calibrator is observed multiple times, gain and bandpass solutions may also be interpolated.

The cross-calibration steps that I have outlined above are generally sufficient for producing coherent images of the scientific target field. In practice, sidelobes — due to incomplete *uv* sampling — disperse flux across the image and create artificial ripples; these sidelobe errors are usually well above the thermal noise limit of modern radio telescopes. The Fourier transform of incompletely sampled visibilities is called the *dirty image*, since it represents the true sky convolved with the Fourier transform of the *uv* sampling function (i.e., the *dirty beam*). To remedy this problem, various algorithms such as *CLEAN* (Högbom 1974; Schwab 1984) have been developed to iteratively reconstruct interferometric images (by estimating the unmeasured Fourier components) under some assumptions about source compactness, smoothness, and sparsity.\(^3\) *CLEAN* algorithms, or “cleaning,” can isolate signals from the true astronomical source by modeling the brightest components as point sources. The clean model can be convolved with the dirty beam and then subtracted from the dirty map, yielding a residual map; more clean components can iteratively be identified in the residual map, added to the model, and subtracted after convolution with the dirty beam. The final clean model is convolved a *clean beam* (a two-dimensional Gaussian fit to the main lobe of the dirty beam) in order to generate the restored image.

---

2The primary beam is the telescope’s sensitivity pattern in the source plane.

3Note that there are other methods than *CLEAN*, although I will only use *CLEAN* in this section. For example, wavelet methods assume different spatial basis functions than *CLEAN*, which operates using a Dirac $\delta$ basis. Maximum entropy methods assume priors regarding the sky brightness distribution.
In order to further improve image fidelity, *self-calibration* is used as a second step beyond cross-calibration (e.g., Cornwell & Wilkinson 1981; Pearson & Readhead 1984). The self-calibration (or selfcal) method can be summarized as follows. First, a set of cross-calibrated visibilities are Fourier transformed to generate a map of the sky brightness. From this sky brightness distribution, a *model* sky is generated (e.g., using CLEAN), usually containing only the brightest unresolved sources that can be confidently attributed to genuine sky emission. Differences between the sky model and inverted visibilities are interpreted to be calibration errors, and the selfcal model is used to update calibration solutions. A new sky brightness can be derived from visibilities after selfcal solutions are applied, and the selfcal cycle is repeated until the noise is sufficiently diminished.

Cross-calibration and self-calibration methods can alleviate insufficient uv coverage and propagation errors to an extent. However, these methods are fundamentally limited because they do not account for *direction-dependent effects* (DDEs). The observed visibilities across different baselines may not be identical to each other or constant in time, as is often assumed for simplicity. The primary beam can vary across the field of view and change with the rotating Earth differently for every pair of antennas. For source positions far away from the phase center, the beam may be severely mischaracterized at any given time; these types of corruptions are just one example of DDEs (see, e.g., Smirnov 2011). For MeerKAT, a wide-band and wide-field telescope array, cross-calibration and self-calibration already require new strategies for synthesis imaging. Additional challenges are anticipated for LADUMA’s massive data volumes (e.g., ~3 TB of data for an 8 hr observation), which will require scalable and efficient software pipelines. Therefore, I will not discuss DDE algorithms, which are being improved at rapid pace by LADUMA team members and other research groups (e.g., Smirnov & Tasse 2015; van Weeren et al. 2016; Tasse et al. 2018), and I will instead describe the process of cross-calibrating, self-calibrating, and subtracting continuum emission from LADUMA test data sets.

### A.1.2 LADUMA data and computational resources

The Inter-University Institute for Data-Intensive Astronomy (IDIA) manages compute and storage servers for MeerKAT observations. The compute server can be accessed in two ways: via the interactive single hello node, or via the parallelized slurm node. The slurm node makes use of the Slurm scheduling software, and can support MPICASA, the parallelized distribution of CASA (Common Astronomy Software Applications; McMullin et al. 2007) which has been available since

---

4https://slurm.schedmd.com/overview.html

5MPI, or Message Passing Interface, is a design standard for parallelized software.
version 4.5. In this work, I use CASA and MPICASA pre-release version 5.3.0-115 supplied in a Singularity container image.\textsuperscript{6}

I have made use of two sets of LADUMA L-band observations, one of which was taken using the 16-antenna array, and the second of which was taken with the 64-antenna full array. The 16-antenna data set was taken in October 2017, and relied on now-deprecated ROACH2 correlators. The 64-antenna commissioning data set was taken in October 2018 using the newer SKARAB correlators. Although both data sets are valuable for testing the data reduction, calibration, imaging, and continuum subtraction strategies, only the 64-antenna results will be directly applicable for future studies due to the change in back-end correlators. For the 64-antenna data, I utilized the \texttt{helo} node on the IDIA clusters for computation. The same data reduction steps have also been performed for the 16-antenna data using both the \texttt{slurm} and \texttt{helo} cluster nodes, but I will not discuss these data any further.

LADUMA requires MeerKAT operations in 32k mode with 32,768 channels in order to meet its science goals, but this observing mode is not yet available as of the time of writing. The 64-antenna data set consists of an 8 hr track in 4,096 channel (4k) spectral mode, time-averaged to 8 seconds. Each channel is 208.984 kHz wide, and the center frequency is 1283.8955 MHz. Two bandpass calibrator sources and one gain (phase) calibrator are observed in addition to the LADUMA field. In Table A.1, I provide an overview of the pointing targets for the 64-antenna data set. The full data set including calibrator observations is over 800 GB in size. For testing the software pipeline, I use the CASA task \texttt{partition} to isolate a fraction of the observations between 1400 to 1440 MHz (channels 2600-2800).

A.1.3 Cross-calibrating the LADUMA field

The first step in calibration is to remove (or flag) the outlier data, most of which are due to RFI. In general, RFI dominates over true astronomical signals, and cannot be corrected via calibration. I begin by flagging autocorrelations, and then continue with standard automatic flagging by using the \texttt{tfcrop} algorithm. After flagging, I verify that visibility amplitudes are consistent with baseline length.

I perform an initial bandpass calibration (i.e., spectrally varying flux calibration) using observations of J1939-6342 and its modeled spectrum (using the 2010 standard; Perley & Butler 2013). I

\textsuperscript{6}Containers are fully self-contained software packages that ensure all dependencies are met and remain constant. Popular container systems include Docker (\url{https://docs.docker.com/}), Kubernetes (\url{https://kubernetes.io/}), and Singularity (\url{https://www.sylabs.io/docs/}).
proceed with cross-calibrating the phase solutions using J1939-6342, after which the antenna phase delays can also be solved using the bandpass calibrator as reference. The phase delay solutions can be used to correct the bandpass again since more of the flux is now appropriately centered. The bandpass calibrator can also be used to “bootstrap” solutions for the gain calibrator’s flux, since CASA does not have any flux model available for J0240-2309. Using reference gains from J1939-6342, I find that the 1400 MHz flux density of J0240-2309 is 6.038 Jy. I then solve for J0240-2309 amplitude and phase solutions. Phase and amplitude solutions are applied to the bandpass calibrator, the gain calibrator, and the LADUMA field. I then split cross-calibrated LADUMA visibilities for self-calibration.

A.1.4 Self-calibrating the LADUMA field

An overview of the selfcal loop is described in Section A.1.1. At each iteration, I save a copy of the calibration tables, flagged data, and model visibilities. I first generate a continuum dirty image (i.e., the Fourier inverted cross-calibrated visibilities with zero cleaning). In recent versions of (MPI)CASA, Fourier inversion and deconvolution are performed using the tclean algorithm. I specify a $3840 \times 3840$ pixel grid with $1.5''$ pixel$^{-1}$ scale and run tclean in multifrequency synthesis mode (appropriate for continuum imaging) with a robust = 0.5 Briggs weighting scheme.\footnote{More details on Briggs weighting can be found in the Ph.D. thesis of D. S. Briggs: http://www.aoc.nrao.edu/dissertations/dbriggs/}

The total continuum emission flux in the dirty map is about 67 mJy. To enforce a conservative model that only contains some of the brightest sources (and thus, accurate phase information), I run the Högbom (1974) CLEAN algorithm with a threshold of 20 mJy non-interactively for 100 iterations. Briggs weighting with low robustness (on the same image grid as before) is used in order to make sources appear more point-like. The continuum model, comprising a small number of clean components, is stored virtually for fast access. It can then be used for constraining phase solutions, which are valid for 15 second to 60 second time-averaged intervals and averaged over all polarizations.

Prior to the selfcal steps, dirty map imaging was characterized by a noise rms of $\sim 0.17$ mJy. After two rounds of selfcal using uniform weighting and 60 s solution intervals, the LADUMA field rms decreases first to 38 nJy and then to 32 nJy. Additional rounds of selfcal do not decrease the noise further. I have also experimented with changing the weighting scheme and decreasing solution intervals, but the rms noise remains high. A few selfcal experiment results are shown in Table A.2.
A.1.5 Continuum subtraction in the LADUMA field

For the calibration steps, I have used multi-measurement sets in order to run tasks in parallel. However, due to a bug discovered in the CASA uvs sub method related to using multi-measurement sets, I convert the self-calibrated data into single measurement sets. While using mstransform to convert the measurement set, I also add a weight spectrum (for estimating uncertainties). I repeat the mstransform conversion using the built-in capability to perform first-order continuum subtraction (uvs contsub with fitorder=1).

I test two strategies for continuum subtraction: polynomial fits to the uv data (called uvs contsub in CASA or UVLIN in other software, and a uv continuum model-based subtraction (called uvs sub in CASA). The uvs contsub task can be run directly, or as part of the mstransform task; I find that in both cases the results are the same. Polynomial order can be supplied, but I have found that higher than first-order fits produce artifacts in the continuum subtraction. Line-free channels can also be specified, which is more useful for individual sources at known redshifts; however, since the LADUMA field is large and line emission can originate from sources in a wide range of unknown redshifts, I do not mask any channels from the fit.

The model-based subtraction, uvs sub, requires a continuum model to be provided before it can be run. I clean non-interactively for 5000 iterations and a threshold of 1 mJy in order to generate the model. Given a measurement set with model and corrected visibilities in the MODEL_DATA and CORRECTED_DATA columns, respectively, the uvs sub CASA task subtracts the model and overwrites the CORRECTED_DATA column with the residuals, i.e., the continuum-free data.

I have generated dirty cubes (e.g., prior to continuum subtraction and cleaning) and line-only cleaned cubes using the two continuum subtraction methods. The method uvs contsub with a first-order polynomial does not perform well, leaving bright peaks in some locations and a negative bowl/sidelobe pattern. Strong sinusoidal patterns in the spectra are present near sources. uvs sub, which works only with single MS data, appears to yield qualitatively better results. Model-based subtraction results in shallower holes than uvs contsub over the image. Also, spectral variations are about $2 - 5\times$ smaller than the artificial ripples in line cubes produced with uvs contsub continuum subtraction. Therefore, it appears that in the regime of many compact sources — including many that may be far from the phase center — uvs sub is the superior option for continuum subtraction. Although the analysis presented above is by no means exhaustive, it provides a valuable first step toward determining continuum subtraction and calibration strategies for the LADUMA survey. A

---

8This bug has been confirmed by NRAO and should be resolved in future versions of CASA.
number of additional tests would be helpful (and will soon be feasible), including the following:

1. Subtract the continuum over the full L-band frequency range.
2. Test newer versions of MPICASA for parallelized uvsub.
3. Test both continuum subtraction methods after correcting for DDEs.
4. Test both methods on the UHF-band observations.
5. Test both methods using the MeerKAT 32k mode.
6. Test continuum subtraction routines using non-CASA software.\(^9\)

### A.2 H I source finding

Once continuum-subtracted MeerKAT line cubes are generated, astronomers will need to rely on robust source-finding software in order to extract astronomical sources of emission (and absorption). One such program is SoFiA, a set of Source Finding Applications that allows users to employ different strategies for detecting different types of sources in data cubes (Serra et al. 2015). Popping et al. (2012) compare the performances of some SoFiA methods on Australian SKA Pathfinder (ASKAP) model cubes, finding that applying a 2D-1D wavelet filter and extracting on the basis of a simple threshold can reliably (and with high completeness) recover unresolved sources at narrow linewidths. I test other promising strategies, including the “S+C” source finding algorithm and reliability-based methods, using SoFiA version 1.0.\(^{10}\)

In preparation for analysis of MeerKAT early science datasets, I have developed a simple parallelized pipeline for benchmarking different SoFiA algorithms. I test for accuracy, parameterization, source deblending, robustness to error, and computational efficiency. The suite of tests includes a flexible way to generate summary plots as well as intermediate data products such as moment maps, subcubes, and even visualizations of the source parameter space. The benchmarking pipeline can be run exclusively within Jupyter notebooks (Kluyver et al. 2016).

#### A.2.1 Identifying H I sources

I have initially tested SoFiA on a simulated 3D dataset of realistic H I profiles that has been convolved with a circular 2D gaussian with FWHM = 15” (for simulation details, see Elson et al.\(^9\))

\(^9\)It is worth noting that CASA does not officially support wide-field (https://casa.nrao.edu/casadocs/casa-5.5.0/imaging/synthesis-imaging/wide-field-imaging-full-primary-beam) or wide-band imaging (https://casa.nrao.edu/casadocs/casa-5.5.0/imaging/synthesis-imaging/wide-band-imaging).

\(^{10}\)https://github.com/SoFiA-Admin/SoFiA/releases/tag/v1.0.0
202

2016). The cube’s pixel size is 3′ and no primary beam correction or other simulated artifacts (aside from uniform noise) have been added. It subtends a primary beam of 0.7 deg$^2$ and spans the frequency range equivalent to H I redshifted to 0.7−0.758 (835.533 MHz to 807.973 MHz in −26 kHz channels). The velocity width per channel ranges between ∆v = 9.33 − 9.65 km s$^{-1}$; I assume a constant ∆v = 9.49 km s$^{-1}$ in this section. The full data cube is $1680 \times 1680 \times 1061$ voxels (where a voxel is a volumetric pixel).

The first challenge is to use SoFiA on a very small machine: a laptop with an Intel Core M 0.8 GHz CPU and 8 GB of RAM. The data cube is 12 GB in size and cannot be directly loaded into memory. I therefore opt to split the data cube into 16 subcubes, each of which spans the full spectral range and 1/4 of the full spatial extent in both RA and Dec. These subcubes overlap each other by 20 pixels, which should be more than enough to prevent unresolved sources from being cleft by subcube splitting. Subcubing is implemented in SoFiA and done concurrently with source finding, so I supply subcube boundaries along with other inputs by procedurally generating parameter files.

In the first run, I set a SNR > 4 threshold and run SoFiA with the default 3D smoothing kernel on a cube with no added noise (no-noise cube). I repeat the process on a cube with uniform noise of rms = 16 µJy beam$^{-1}$ (noisy cube). I specify values of $s_i, s_j, s_k$ (minimum sizes) and $r_i, r_j, r_k$ (merge radii) as SoFiA input parameters. Sources are considered bona fide detections if at least a $s_i \times s_j$ group of pixels and $s_k$ consecutive channels are above the signal-to-noise threshold. Detections merged together as a single source if they spatially border each other by $r_i$ or $r_j$ pixels, or if they are within $r_k$ spectral channels of each other. The SoFiA reliability calculation is not used to filter out results at this point, but I will discuss it as another test case for later benchmarking (see Section A.2.4).

I first run SoFiA with parameters $s_i = s_j = 4, s_k = 6, r_i = r_j = 1$, and $r_k = 8$. These parameters were chosen based on the examples provided in the SoFiA tutorial. On the small machine, a full run of the entire cube (divided into 16 subcubes) takes ~ 90 minutes of compute time. The pipeline then produces, for each subcube, moment 0 and 1 maps (i.e., integrated flux and velocity maps) and catalogs with extracted sources. For the no-noise runs, SoFiA-generated integrated flux and velocity maps accurately matched the “true” maps created by collapsing the cube; see Figure A.1 for a comparison. I show the velocity maps in units of redshift, and find that the SoFiA velocity map and the true velocity map agree to within ∆z/z < 10$^{-5}$. The distribution of residual fluxes, or the differences between the two velocity maps, is super-gaussian (i.e., it has fat tails, similar to a Student’s $t$-distribution).
I stitch the moment map images together by using the following procedure:

1. A 1680 \times 1680 \times 16 array of zeros is created, where 1680 is the number of pixels along one spatial axis of the original cube, and 16 is the number of subimages.

2. Each subcube moment map (subimage) is correctly positioned in space and padded by zeros, sequentially along the third axis of the array.

3. A maximum filter is applied along the third dimension, so that maximum-valued pixel is kept in the stitched image.

In Figure A.2, I show the zeroth moment (integrated flux) map of all 16 no-noise subcubes prior to stitching. Careful examination of the overlap regions (20 pixels for bordering subcubes) confirms that source fluxes are recovered consistently. Although not illustrated here, the stitched zeroth moment map does not deviate from the true zeroth moment map by more than $10^{-12}$ Jy, except for two point sources that deviate in flux by $\sim 0.1$ mJy. The two erroneous sources are not in the overlap regions between subcubes, so they likely result from the few instances in which SoFiA fails to accurately recover flux in the no-noise cube. I also find that $\Delta z \sim 10^{-7}$, except for 0.19% of cases (a few dozen sources) in which the residual $\Delta z > 0.001$. Again, these velocity map errors do not preferentially occur in regions of overlap, but actually appear more numerous in locations of high projected density. Therefore, the stitching strategy faithfully re-builds the flux and velocity maps, although different strategies must be employed in order to stitch together images with negative flux (e.g., H I absorbers).

**A.2.2 Removing duplicates**

SoFiA source catalogs are produced by combining all sources together and then removing duplicated sources. As a consequence of the subcubing strategy, not only are there duplicate sources in regions of subcube overlap, but also “quadruplicates” where the corners of four subcubes overlap. I therefore use the following method to reject duplicate sources:

1. Set a *metric* equal to the median H I profile size used to generate the data cubes: (1, 1, 20) in units of voxels (Elson et al. 2016).

2. Scale all SoFiA source sizes and positions by the metric, converting the $i, j, k$ dimensions to approximately equal spatial scales (assuming that H I profiles are randomly inclined in the cube).
3. Create a \(k\)-d tree (Bentley 1975) of all the spatial positions of the SoFiA sources and produce a list of each all nearest-neighbor (NN) distances.

4. Cut SoFiA catalog by restricting sources to a SNR threshold, and repeat for each remaining source’s NN.

5. Take the H I flux ratio of each source to its NN, ensuring that the value is always greater than unity (i.e., by inverting the ratio if not).

6. Only keep sources that have NN distances above some distance threshold, or have NN flux ratio above some flux ratio threshold. The logical or statement here is essential for rejecting only NNs that are very nearby and also have a very similar flux to the source in question.

This procedure is run once in regions of duplicate overlap, and again in only the regions of quadruplicate overlap.

The distance and flux ratio thresholds in step (6) have been chosen in order to prevent rejecting true sources rather than maximizing purity. The threshold values, distance < 1.0 normalized voxels, and flux ratio < 1.15, have been individually tested on the catalog of “true” sources, yielding false rejection rates of 7.4% and 3.0% respectively. The combination of the two thresholds, however, drops the false rejection rate to zero (but it should be noted that the purity cannot be estimated via testing on the true catalog). I cut the SoFiA catalogs by using the same criteria. Any sources marked as duplicates in regions of non-overlap must have been falsely rejected, and so I find that the rate of false rejections is 3.68%. In the case of the no-noise cube, the duplicate-rejection strategy reduced the number of total sources from 45055 to 43955 (a 2.4% reduction), although \(\sim 100\) sources were likely rejected falsely.

For the noisy cube, the duplicate-rejection method reduced the number of sources from 3242 to 3140 (a 3.1% reduction), and it is likely that a few were falsely rejected as well. The dramatic drop in number of sources as with the introduction of noise is expected because the H I mass function (or nearly-equivalently, the flux distribution) continues to rise at fluxes below the completeness limit introduced along with the simulated noise. Likewise, the noisy cube moment maps appear very different from the true moment maps because the they are sparsely populated with recovered sources.
A.2.3 Parallelizing SoFiA

Next, SoFiA is set up to run on an IDIA cluster. The system runs on Linux (Ubuntu 14.04) and has 48 GB of RAM and 8 (single-threaded) Intel Haswell 2.5 GHz cores. Because SoFiA requires $3 - 4 \times$ the input cube size in memory (Serra et al. 2015), it is still not feasible to run the pipeline without splitting the full data cube ($\sim 12$ GB) into subcubes. However, I can leverage the powerful computing capabilities by parallelizing some of the processes within the pipeline. Eight instances of SoFiA are called in parallel using the multiprocessing package in Python, such that a full data cube requires up to $\sim 24$ GB of memory, and takes $3 - 4$ minutes to be processed (i.e., a factor of $\sim 25$ speed-up). The compute time for stitching together images and catalogs is far less than actually running SoFiA, so I do not parallelize the other steps in the analysis.

A.2.4 Results

Optimizing parameters with the S+C method

Now that an efficient pipeline for generating SoFiA data products has been created, I can afford to explore the source-finding parameter space. I use the SoFiA “Smooth + Clip” finder (S+C) and smooth with the default smoothing kernel. All combinations of $s_i = s_j \in \{2, 3, 4\}$, $s_k \in \{4, 5, 6, 7, 8, 9, 10\}$, $r_i = r_j = 1$, and $r_k \in \{2, 3, 4, 5, 6, 7, 8, 9\}$ are explored. Subcube images and catalogs are stitched together according to the strategy described previously.

Recovered H I flux distributions, analogous to the H I mass function, are shown in Figure A.3. One of the limitations of using the H I flux distribution as a diagnostic is that impurity and incompleteness errors can slightly cancel each other out. Finding sources with a large minimum source size (e.g., $s_i = s_j = 4$) results in an excess number of detections at $\log(F_{HI}/\text{Jy km s}^{-1}) \sim -1.8$. When $s_k$ is increased, this problem is further subject to incompleteness, which gives the appearance of alleviating the purity problem.

It appears that small source sizes ($s_i = s_j = 2$, $s_k \leq 6$) and large spectral merge radii ($r_k \approx 9$) lead to the best match with the true flux distribution over the largest range. Low values of $s_i, s_j, s_k$ lead to accurate number counts near the bright end of the flux distribution. It is also apparent that large values of $s_i, s_j$ and $s_k$ cause the completeness to decrease toward the faint end. Whereas the distribution is nearly complete at $\log(F_{HI}/\text{Jy km s}^{-1}) = -2$ for $(s_i, s_j, s_k) = (2, 2, 4)$, it is only complete at $\log(F_{HI}/\text{Jy km s}^{-1}) = -1.5$ for $(s_i, s_j, s_k) = (4, 4, 10)$.

---

11 The original cluster was named ARCADE, although it has since been upgraded and no longer exists.
Reliability-based source finding

Serra et al. (2012) have introduced a method of isolating true, positive-flux sources by assuming that false detections due to noise have symmetric flux about zero, and using negative detections to parameterize the reliability, $R$, of positive detections. This strategy is best-suited for isolating detections that are significantly different from noise peaks in the source parameter space (e.g., integrated flux, number of pixels, peak flux). I benchmark the source-finding process using $R \in \{0.5, 0.75, 0.8, 0.85, 0.9, 0.95, 0.97, 0.99\}$ while constraining the source parameters to the optimized values $(s_i, s_j, s_k) = (2, 2, 4)$ and $(r_i, r_j, r_k) = (1, 1, 9)$.

From the recovered H I flux distribution, shown in Figure A.4, it can be seen that very few sources are detected when the reliability criterion is used. Almost all detections are bright (i.e., $\log(F_{\text{HI}}/\text{Jy km s}^{-1}) \geq -1.5$), and sources fainter than $\log(F_{\text{HI}}/\text{Jy km s}^{-1}) \approx -1$ are incomplete. It is no surprise, in hindsight, that the reliability parameterization does not help source-finding: nearly all simulated galaxies are unresolved and fall squarely in the low-SNR regime. Given these stipulations, I would still expect SoFiA reliability calculations to be effective for finding sources at lower redshifts. At higher redshifts, it would be more advantageous to disable the reliability parameter.

Acknowledgments I thank Brad Frank, Sarah Blyth, and Ian Heywood for valuable conversations and assistance with data reduction. I also thank the University of Cape Town (UCT), the Inter-University Institute for Data Intensive Astronomy (IDIA), the South African Radio Astronomy Observatory (SARAO), and the South African Astronomical Observatory (SAAO) for hosting me. This research would not be possible without support from the United States Agency for International Development (USAID), and the Rutgers Centers for Global Advancement and International Affairs (Rutgers GAIA Centers).
Table A.1. LADUMA 64-antenna observations

<table>
<thead>
<tr>
<th>Field</th>
<th>Object</th>
<th>Coordinates (J2000)</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>J1939-6342</td>
<td>19:39:25.03 −63.42.45.6</td>
<td>Bandpass calibrator</td>
</tr>
<tr>
<td>1</td>
<td>J0240-2309</td>
<td>02:40:08.14 −23.09.16.0</td>
<td>Gain calibrator</td>
</tr>
<tr>
<td>2</td>
<td>J0332-2807</td>
<td>03:32:30.40 −28.07.57.0</td>
<td>LADUMA field</td>
</tr>
<tr>
<td>3</td>
<td>J0521+1638</td>
<td>05:21:09.89 +16.38.22.1</td>
<td>Bandpass calibrator</td>
</tr>
</tbody>
</table>

Table A.2. LADUMA 64-antenna selfcal results

<table>
<thead>
<tr>
<th>Selfcal step</th>
<th>Threshold (mJy)</th>
<th>Weighting</th>
<th>Solution interval (s)</th>
<th>Final rms (nJy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dirty map</td>
<td></td>
<td>robust = 0.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>20</td>
<td>uniform</td>
<td>60</td>
<td>170</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>uniform</td>
<td>60</td>
<td>32</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>uniform</td>
<td>40</td>
<td>36</td>
</tr>
<tr>
<td>1</td>
<td>20</td>
<td>robust = −2</td>
<td>60</td>
<td>58</td>
</tr>
</tbody>
</table>

Note. — The threshold is the non-interactive cleaning flux density threshold used for generating the selfcal model. Solution intervals are for gain solutions, assuming a minimum SNR of 3 and four valid baselines per antenna. The final rms is calculated after a round of non-interactive clean.
Figure A.1 Comparisons between the SoFiA-generated first moment map (left), the “true” moment map (center), and difference between the two (right). No noise has been added. Velocity maps are shown in units of redshift. Note that the difference map, or residual map, is scaled over a range of $\sim 10^{-6}$. The size of the subcube used to generate the first moment maps is $200 \times 200 \times 1061$ voxels.
Figure A.2 The integrated flux maps of subcubes in the no-noise simulated data cube is shown prior to stitching. Lighter colors correspond to higher flux.
Figure A.3 H I flux distribution recovered by SoFiA runs using different source-finding parameters, including permutations of minimum spatial size \((s_i, s_j); \text{columns}\), minimum spectral size \((s_k); \text{rows}\), and minimum spectral size for merging \((r_k); \text{color}\). In all cases, \(r_i = r_j = 1\). The true flux distribution is shown in black.
Figure A.4 H I flux distribution recovered by SoFiA runs using different reliability thresholds ($R$; color). The hyperparameter values chosen ($s_i, s_j, s_k$) = (2, 2, 4) and ($r_i, r_j, r_K$) = (1, 1, 9) have been optimized for the noise-added simulation. The true flux distribution is shown in black.