

An Investigation of Lyman Continuum Emitting Green Peas Using the Limited Spatial Resolution of *HST*-COS

ALEX HAUGHTON AND BRIAN FLEMING

ABSTRACT

Green Pea (GP) type galaxies are potential analogs for dwarf starburst galaxies that could be sources of ionizing radiation during the Epoch of Reionization. However, their morphology and feedback mechanisms remain mysterious due to their compact nature. NGC 2366, a nearby dwarf starburst galaxy, suggests that GPs may have a brighter super star cluster surrounded by a dimmer diffuse galaxy. In this paper, we present a model of the spatial flaring of point and diffuse source spectra in *Hubble Space Telescope* - Cosmic Origins Spectrograph (*HST*-COS), using this model to pull spatial data from observations of Green Peas. We then analyze 14 confirmed Lyman continuum emitters from a recent *HST*-COS study and find that the *HST* data is consistent with a bright central cluster dominating any potential host galaxy flux in the FUV by several orders of magnitude. However, we do find that all the galaxies appear to have a more extended Lyman alpha line, consistent with an extended galaxy halo. Three of the galaxies show evidence of extended emission warranting a follow up investigation.

1. INTRODUCTION

During the Epoch of Reionization (EoR), between redshift 6-12, the hydrogen of the Universe transitioned from largely neutral to largely ionized (Gnedin & Ostriker 1997; Becker et al. 2001; Ciardi et al. 2003; Wise et al. 2014). Bounds on the end of the reionization period are set by the existence of a Gunn-Peterson trough (Gunn & Peterson 1965) near redshift $z \sim 6$ (Becker et al. 2001). However, the source of ionizing Lyman continuum (LyC) photons responsible for reionization is an ongoing scientific investigation. Two main types of galaxies are considered to be candidates for driving reionization: dim galaxies with Active Galactic Nuclei (AGN) (Volonteri & Gnedin 2009; Giallongo, E. et al. 2015; Madau & Haardt 2015), or dwarf starburst galaxies (Razoumov & Sommer-Larsen 2010; Alvarez et al. 2012; Wise et al. 2014; Paardekooper et al. 2015). Recently, the weight of the studies seem to favor dwarf star-

burst galaxies as the dominant contributors to ionizing radiation (Trebitsch et al. 2018; Finkelstein et al. 2019), although we do not yet have accurate counts of low-luminosity AGN at the EoR.

Studies have found that high z galaxies at the EoR need to average a LyC escape fraction (f_{esc}^{LyC}) of at least 10-20% to help drive reionization (Robertson et al. 2013; Dressler et al. 2015); f_{esc}^{LyC} is the fraction of LyC photons produced by a region that escape rather than being absorbed. While directly observing LyC photons from the EoR is currently impossible due to the neutral hydrogen filling the Intergalactic Medium (IGM) before and during reionization, there have been some recent observations of Lyman continuum emitters (LCEs) at intermediate redshifts near $z \sim 3$ with $f_{esc}^{LyC} > 0.5$ (Shapley et al. 2016; de Barros et al. 2016). As the furthest detectable LyC photons are found somewhere around $z \sim 3$ (Shapley et al. 2006;

Cooke et al. 2014) to $z \sim 4.5$ (Romano, M. et al. 2019), other research has focused on finding LCEs at lower redshifts.

Lower redshift galaxies are expected to have a smaller f_{esc}^{LyC} (Razoumov & Sommer-Larsen 2010; Yajima et al. 2011), and there is tentative evidence to support this (Inoue et al. 2006; Siana et al. 2010), but differences in methodology between low redshift ($z < 0.5$) observations and intermediate redshift ($z \sim 3$) observations make direct comparisons somewhat tenuous (Leitet, E. et al. 2013).

Until 2014, the only successfully identified Lyman continuum emitters at low redshifts were Haro 11 (Bergvall et al. 2006) and Tololo 1247-232 (Leitet, E. et al. 2013). Other searches (Leitherer et al. 1995; Deharveng et al. 2001) were unable to positively detect ionizing flux, leading to debate about the actual escape fraction limits and viability of these galaxies as analogs for ionizing galaxies at the EoR (Hurwitz et al. 1997; Grimes et al. 2007) as well as theories such as the Photon Underproduction Crisis (Kollmeier et al. 2014).

This paradigm changed with observations of “Green Pea” (GP) type galaxies, which are often found to be LCEs. GPs were first discovered as part of the Galaxy Zoo Citizen Science Project (Lintott et al. 2008) in 2009. Originally classified as stars by the photometric pipeline (Lupton et al. 2001) of the Sloan Digital Sky Survey (SDSS; York et al. 2000), they appeared green to volunteers due to dominating the SDSS r -band. Upon further examination, spectral features matched those of galaxies, not stars, and several common traits of GPs emerged. Cardamone et al. (2009) found that GPs have low mass, high star formation rates, low metallicity, high UV luminosity, and a strong [OIII] $\lambda 5007$ line responsible for the r -band dominance (see Table 1 for numbers).

Further investigations (Izotov et al. 2011; Jaskot & Oey 2013; Nakajima & Ouchi 2014;

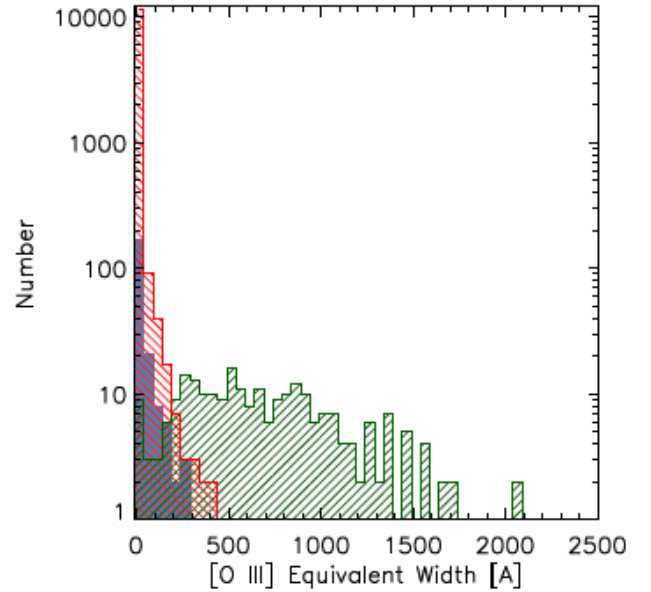


Figure 1. Figure from Cardamone et al. (2009). [OIII] equivalent widths are much higher for GPs (green histogram) than either $\sim 10,000$ comparison galaxies (red histogram) or ultraviolet luminous compact galaxies (blue histogram). All equivalent widths from SDSS pipeline.

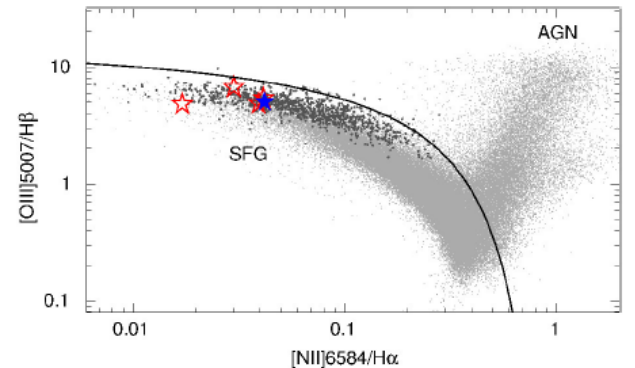


Figure 2. Figure from Izotov et al. (2016b). Baldwin-Phillips-Terlevich diagram (Baldwin et al. 1981) for classifying narrow line emission regions. The dark line separates star forming galaxies from active galactic nuclei. Luminous compact galaxies from Izotov et al. (2011) are small dark-grey circles. The GPs identified in Izotov et al. (2016a) are the red stars, with Izotov et al. (2016b) being the blue filled star.

Henry et al. 2015) have refined and expounded

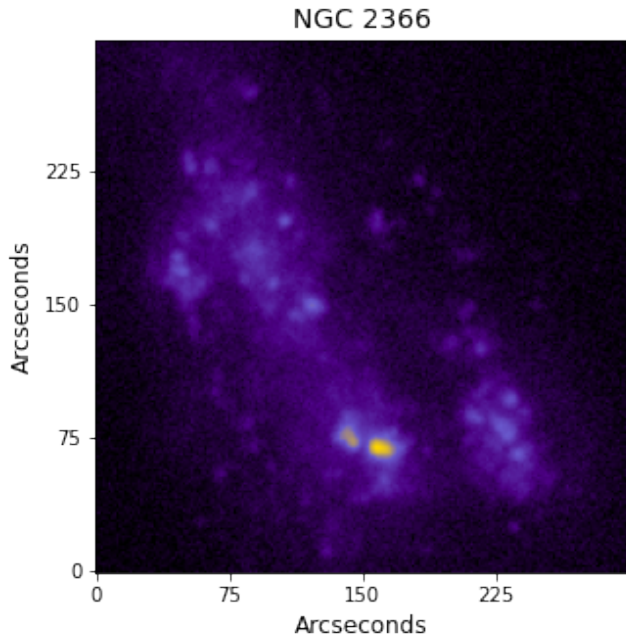


Figure 3. *Galaxy Evolution Explorer (GALEX, Morrissey et al. 2005) FUV image of dwarf galaxy NGC 2366, the nearest potential Green Pea analog. The bright spot at the bottom of the image is a pair of super star clusters known as Markarian 71.*

upon these properties: GPs also usually have narrowly separated double peaked Lyman alpha ($\text{Ly}\alpha$) lines, young burst ages (~ 4 Myr), large $\text{H}\beta$ equivalent widths, and high $[\text{OIII}]\lambda 5007/[\text{OII}]\lambda 3727$ ratios (Micheva et al. 2017). Most importantly, GPs are often found to be LCEs, which combined with other properties make them potential analogs for galaxies at the EoR (Izotov et al. 2016a,b, 2018a,b; Flury et al. 2022).

Various metrics have been used to search for LCEs and GPs. The $[\text{OIII}]\lambda 5007/[\text{OII}]\lambda 3727$ ratio, indicative of highly ionized gas (Nakajima & Ouchi 2014), has been the preferred search criteria in several instances (Izotov et al. 2016a, 2018a), but is believed to be a necessary but not sufficient condition for LyC leakage (Nakajima et al. 2020). Verhamme et al. (2017) found that LCEs usually have narrowly separated double $\text{Ly}\alpha$ peaks, with the velocity separation an indicator of $f_{\text{esc}}^{\text{LyC}}$; the smaller the velocity sep-

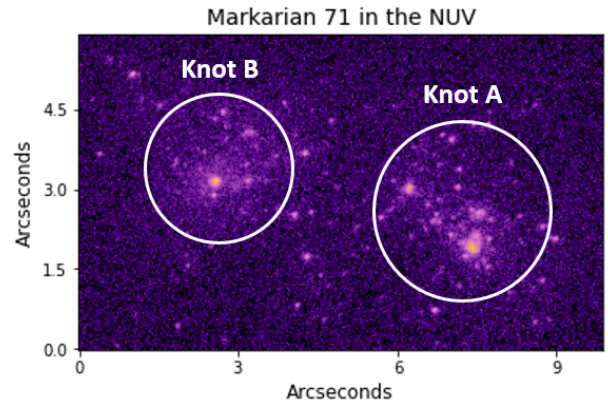


Figure 4. *HST-Space Telescope Imaging Spectrograph (Woodgate et al. 1998) NUV image of Markarian 71 using the F25CIII filter with a pivot wavelength of 1989 Å (Prichard et al. 2022). Knot A is on the right and Knot B is on the left.*

aration, the higher $f_{\text{esc}}^{\text{LyC}}$. High $\text{H}\beta$ equivalent widths (EW) should result from strong nebular emission associated with LyC leakage (Yamanaka et al. 2020), and combined with the UV spectral index β may form a good predictor of $f_{\text{esc}}^{\text{LyC}}$ (Zackrisson et al. 2013). β is the slope of a power law fit to flux in the UV continuum, $F(\lambda) \propto \lambda^\beta$, with $\beta \leq -2$ indicating a very blue ultraviolet color.

The Low-Redshift Lyman Continuum Survey (LzLCS, Flury et al. 2022) chose to broaden the range of the usual search criteria (they used high $[\text{OIII}]/[\text{OII}]$, high star formation rate, and spectral index β) in hopes of finding a more diverse set of LCEs. They were successful, with 35 detections above 2σ significance out of 66 surveyed galaxies covering a broader range of properties than previous LCEs, including 12 detections above 5σ significance.

The morphology of LCEs and the mechanisms that allow a galaxy to leak LyC photons remain unclear; the compact shape of GPs make them challenging subjects for study. Recently, the local irregular galaxy NGC 2366 has been discovered to be a potential low- z GP analog (Micheva et al. 2017). The traits of NGC 2366

Table 1. Properties of GPs, NGC 2366, and Mrk 71

Property	Average GP	NGC 2366	Mrk 71
Distance (Mpc)	1500	3.4	3.4
$r_{1/2}$ (kpc, optical)	1.0	2.7	...
$\log[M_\star]$ (M_\odot)	7.8-10.0	8.41	5.2
EW $[\text{H}\beta]$ (\AA)	9-295	44	127.6
$[\text{OIII}]/[\text{OII}]$	0.5-14	6.34	11.7
$12 + \log[\text{O}/\text{H}]$	7.82-8.54	7.89	7.89
$\log[L_{FUV}]$ (L_\odot)	10.5	8.34	...

NOTE—Data from [McQuinn et al. \(2015\)](#); [Cardamone et al. \(2009\)](#); [Lelli et al. \(2014\)](#); [Izotov et al. \(2011\)](#); [Micheva et al. \(2017\)](#); [James et al. \(2016\)](#); [Moustakas & Kennicutt \(2006\)](#) and [Tolstoy et al. \(1995\)](#).

compared to an average GP are found in Table 1. While the properties are largely similar, the stellar mass is on the lower side, and it has a lower luminosity than the average GP by two orders of magnitude. More importantly though, its distance of 3.4 Mpc ([Tolstoy et al. 1995](#)) makes it the best possible laboratory for studying a local GP analog at high angular resolution (see Figure 3).

Specifically, the GP-like traits of NGC 2366 are dominated by two super star clusters (SSCs) known as Knot A and Knot B in a 200 pc by 115 pc patch of sky known as Markarian 71 (Mrk 71, [James et al. 2016](#)). In comparison to NGC 2366 in general, Mrk 71 has a higher $\text{H}\beta$ equivalent width and $[\text{OIII}]/[\text{OII}]$ ratio (see Table 1). Knot A has the most extreme GP characteristics and displays signs of “catastrophic cooling,” a form of extreme radiative feedback where the superwinds normally found in super star clusters are suppressed ([Komarova et al. 2021](#)). Knot B, on the other hand, is a superbubble region where supernovae (SNe) have hollowed out the neutral gas, a process that takes several Myrs, thereby setting the star-formation timescale for that region. There have been several Wolf-Rayet (WR) stars confirmed

in Knot B, even though typical WR activity timescales are less than the SNe-driven superbubble timescale ([Drissen et al. 2000](#)).

If NGC 2366 is a good model of a GP, the traits of Mrk71 and its SSCs may hold insights into the mechanisms of LyC escape. Placed at a typical GP redshift of $z \sim 0.3$, Mrk71 would occupy an angular diameter of $0.04''$ and therefore appear as a point source to *Hubble Space Telescope*-Cosmic Origins Spectrograph (*HST*-COS, [Green et al. 2012](#)). The rest of NGC 2366 would appear as a larger diffuse region $1.2''$ in diameter around the bright point source. The LzLCS was performed with COS and found 14 new strongly detected LyC leakers ([Flury et al. 2022](#)). While *HST*-COS is nominally a point-source spectrograph, like all spectrographs there is a corresponding field shift as targets are moved around in the aperture, suggesting that some imaging resolution may be possible. This capability was demonstrated for the first time in [Yang et al. \(2018\)](#) to detect extended Ly α emission in local galaxies.

In Section 2, we discuss the construction and use of an *HST*-COS G140L CENWAVE 800 model to find which regions of the detector would be most likely to capture to a central

point source versus an extended diffuse region. In Section 3, we examine the 14 LCEs from the LzLCS with the highest f_{esc}^{LyC} , looking for evidence of a diffuse halo with different properties than a central bright superstar cluster. In Section 4, we present conclusions based on the research in Sections 2 and 3, along with future work to better understand the LzLCS galaxies and create a galaxy catalog for target selection.

2. *HST*-COS G140L CENWAVE 800 MODEL

HST-COS is the premier ultraviolet spectroscopy instrument for, among other things, studying formation and evolution of galaxies. While generally not used for imaging, there is spatial information in the cross dispersion direction of COS, depending on target, grating, CENWAVE, and other settings of the instrument (Soderblom 2021; James et al. 2022).

Many GP targets appear as bright point sources due to their redshifts and sizes (Izotov et al. 2016b, 2018b; Flury et al. 2022). If their morphology is similar to that of NGC 2366, it is possible the bright sources are simply a compact starburst region surrounded by a typical star-forming galaxy. To investigate if emission from the surrounding galaxy is hiding in the LzLCS data, it is necessary to fully understand *HST*-COS. The model can then be used to discover what different potential GP-like galaxies would look like if they were observed with COS and pick targets for further investigation.

2.1. *The Cross Dispersion Spread Function*

HST-COS has a circular 2.5 arcminute aperture, and counts at the top and bottom of the 2D spectrum correspond roughly to light entering the aperture near the top and bottom. However, since COS was not designed as an imaging spectrograph in order to maximize throughput at short wavelengths, the cross dispersion spread function (CDSF) is relatively broad. The CDSF depends on the Lifetime Position of the detector as well as the grating and central wave-

length used (James et al. 2022). For the LzLCS sample, all measurements were taken at Lifetime Position 4 with the G140L grating at CENWAVE 800.

The theoretical CDSF for the G140L grating (Fox 2018) shown in Figure 5 displays a “bowtie” shape; that is, the beam is tighter around 1200Å, where it has a full width at half maximum (FWHM) of about 5.8 y pixels, than it is at 800Å (FWHM = 11.4 y pixels) and especially 1800Å (FWHM = 18.9 y pixels). The pixel scale of COS in the cross dispersion direction is approximately 0.1” per pixel. The G140L configuration also has a greater effective area at 1200Å (1400 cm²) than at 1800Å (150 cm²). Examples of the CDSF at 1250Å and 1800Å are given in Figure 5; at 1800Å the CDSF is not only broader, but it features secondary peaks and is noticeably more asymmetrical by eye. In general, the wings of the CDSF spread more light to the low side of the detector. The smallest values given at the ends of the CDSF are on the order of 10⁻⁴ times the central peak.

Yang et al. (2018) performed a deconvolution on GP spectra to compare the spatial cross sections of Ly α and UV continuum emission, demonstrating that the spatial resolution of *HST*-COS can be used to carry out such a study. We push this idea further to investigate the continuum emission of a point source-like target that is orders of magnitude brighter than the surrounding diffuse host galaxy emission.

2.2. *Building the Model*

For our study we aim to cut horizontally along the spectrum to compare areas dominated by emission from an extended diffuse galaxy to areas dominated by a central point source. To start, we construct a model of the G140L CENWAVE 800 COS spectrum to better understand where photons fall on the detector.

Eventually we plan to build a catalog of galaxies to run through the detector, but we begin with a modified version of NGC 2366. Our test

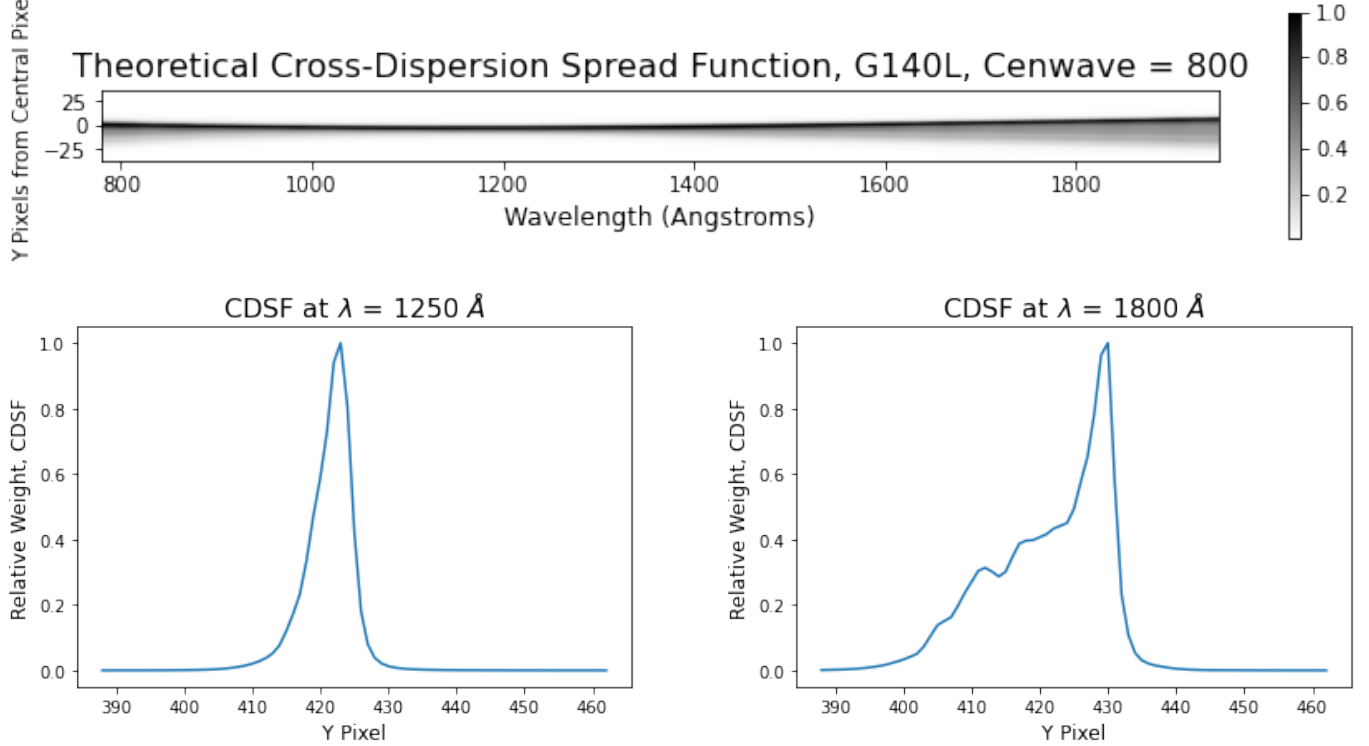


Figure 5. *Top: The Theoretical CDSF for G140L CENWAVE 800. Note the bowtie shape. Middle: The CDSF at 1250Å (left) and 1800Å (right), with HST-COS aiming such that pixel 425 is the central pixel. Note that the peak at 1250Å is slightly below the central pixel, and the peak at 1800Å is well above the central pixel.*

galaxy is modeled as one extremely bright point source (Mrk 71 would be small enough at a typical GP redshift of $z \sim 0.3$ to appear to COS as a point source) surrounded by a less bright diffuse galaxy large enough to overfill the COS aperture. We define the “Surface Brightness Ratio” (SBR) as the ratio of the point source surface brightness ($\text{erg cm}^{-2} \text{s}^{-1} \text{arcseconds}^{-2}$) to the diffuse galaxy surface brightness. For NGC 2366 this value is about 17. The diffuse and point source fluxes for the test galaxy also match those of NGC 2366, and the initial exposure time used was similar to that of a LzLCS galaxy (about 2000 seconds). A completely flat spectrum is assumed, with no variation in input by wavelength.

Photons entering COS first need to pass through the 2.5” aperture. Due to *HST*’s astigmatism, some light from beyond 2.5” is capable of leaking into the aperture - the transmission

function given by the COS Instrument Handbook drops from 100% transmission at a 1” diameter to zero at a diameter of 4” (James et al. 2022). The circular shape of the aperture is also projected onto the detector, so that a large central band falls in the middle of the detector and the throughput drops off with both the transmission function and the shape of the circle.

Photons then hit the selected grating in a quasi-Rowland circle mount; the gratings are designed to correct for the inherent aberrations in *HST* while diffracting the light. Finally, they hit the FUV cross delay line (XDL) detector, which has different efficiencies at different wavelengths. The model takes the image of the test galaxy and its fluxes and attenuates the photons at each step, accounting for the circular aperture, transmission function, effective area of the entire system, wavelengths of the photons, and ultimately the CDSF.

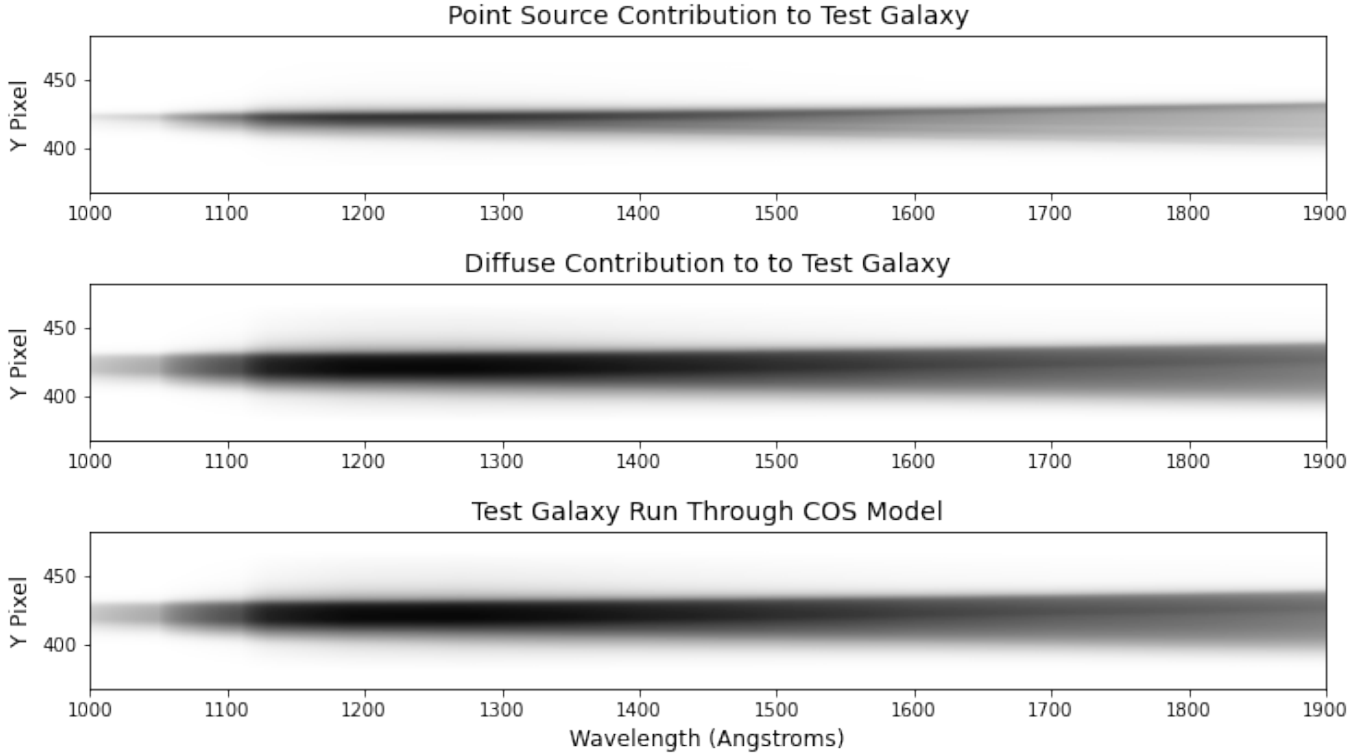


Figure 6. *Top: The contribution of a point source to the model galaxy. Middle: The contribution of a diffuse halo overfilling the COS aperture to the model galaxy. Bottom: The top two images coadded. The scale is logarithmic and the same for all three images. A flat spectrum is assumed, so the differences from left to right are a result of wavelength and effective area, while the differences from top to bottom are due to the transmission function and CDSF.*

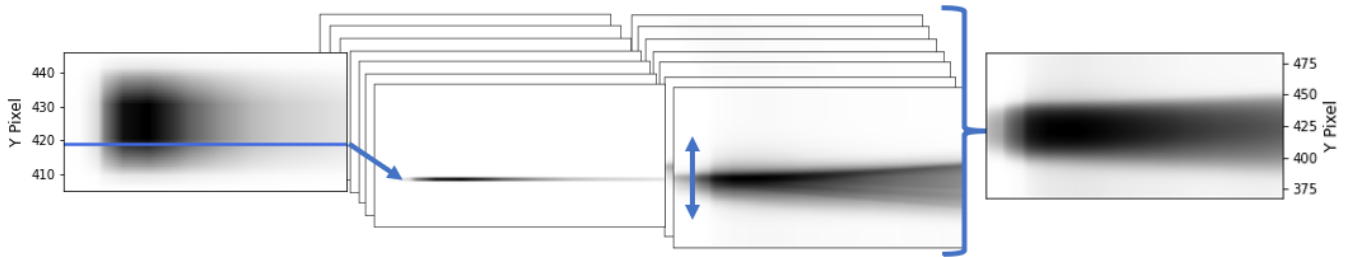


Figure 7. *To build an image of a diffuse source overfilling the slit, the model first generates how the detector would look with just the transmission function and effective area (Left). Next, it picks out one row of incoming photons (Second) and spreads them out with the CDSF (Third). Finally, it coadds all the resulting images to create the detector image (Final).*

To generate the point source contribution the model assumes all the light should be centered about the central pixel and then applies the CDSF. There are more steps to generate the contribution of the diffuse source; what the model does is shown in Figure 7. First, the

model calculates an intermediate image of what the detector would look like without the CDSF, adjusting for the aperture shape, transmission function, and effective area. Next, the model goes row by row in this image and spreads the light according to the CDSF, creating 39 im-

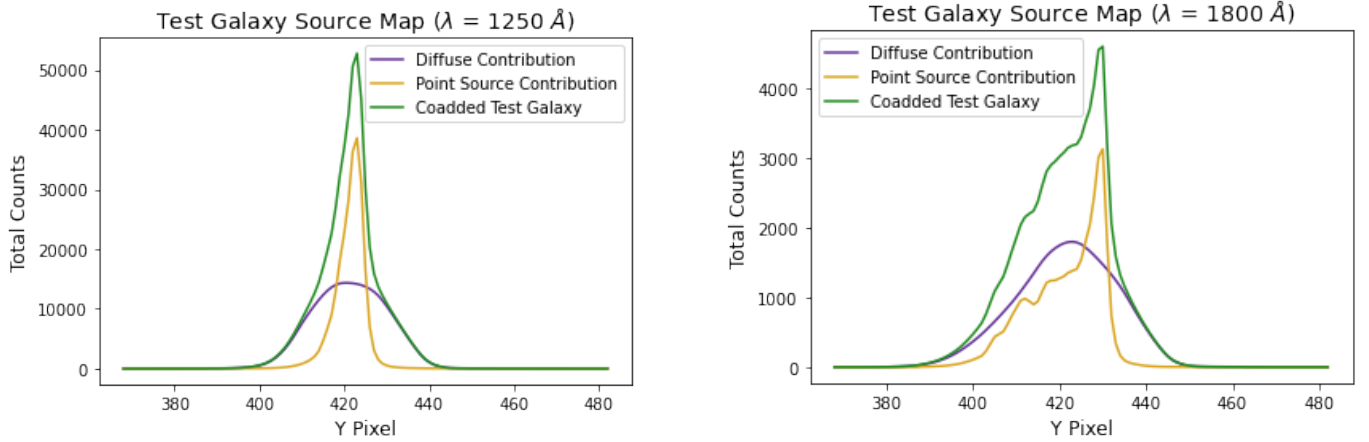


Figure 8. *Cross sections of the Test Galaxy Model at 1250 and 1800 angstroms. The diffuse emission smooths out the CDSF, while the point source emission copies it exactly.*

ages, one for each row. Finally, the detector coadds those images to create the final detector image. The contributions to and generation of the model galaxy can be seen in Figure 6. Notice that by eye, the test galaxy looks like it is dominated by the diffuse emission even with a SBR of 17.

2.3. Comparing Diffuse and Point Source Emission

With a first-order model constructed, the next step is to assess whether it is possible to bin the data in such a way that diffuse emission and point source emission can be separated out. Specifically, we look for rows dominated by photons from the point source and rows dominated by diffuse emission. Using the model, for each row we calculate how many counts come from which source and compare the two. The relative contribution of point versus diffuse emission at a given pixel varies with wavelength as well. We therefore determine this relative contribution at single wavelengths, beginning with 1250Å and 1800Å as test wavelengths.

While the cross sections shows smooth, continuous functions related to the CDSF (see Figure 8), taking a look at the percent counts from the point source leads to a discontinuity where the CDSF cutoff (y pixels 388 and 462 in top images of Figure 9). To resolve this, it is necessary

to add tails to the CDSF. With no physical reason to try a particular function, both Gaussian and exponential decay tails are checked. The resulting contribution curves are shown in Figure 9. The exponential decay tails show a greater degree of discontinuity than the Gaussian ones, so the Gaussian tails are chosen to create a modified CDSF.

2.4. Influence of Surface Brightness Ratio

Which pixels are dominated by emission from a diffuse source versus a point source changes based on the size and shape of the galaxy as well as the SBR. Altering the size and shape of the galaxy will be explored in a future phase of this project; for this investigation we focus on the SBR.

We define four categories of y pixels on the detector, outlined in Table 2. These regions are identified using our model for different SBRs ranging from 1:1 (equal point source/diffuse surface brightness) to 1000:1 (point source 1000x the surface brightness of the diffuse emission). Results are summarized in Figure 10.

For a galaxy like NGC 2366, where the SBR is about 17, the diffuse light would be the majority of light in most rows, with just a few in the center having the majority of light come from the point source.

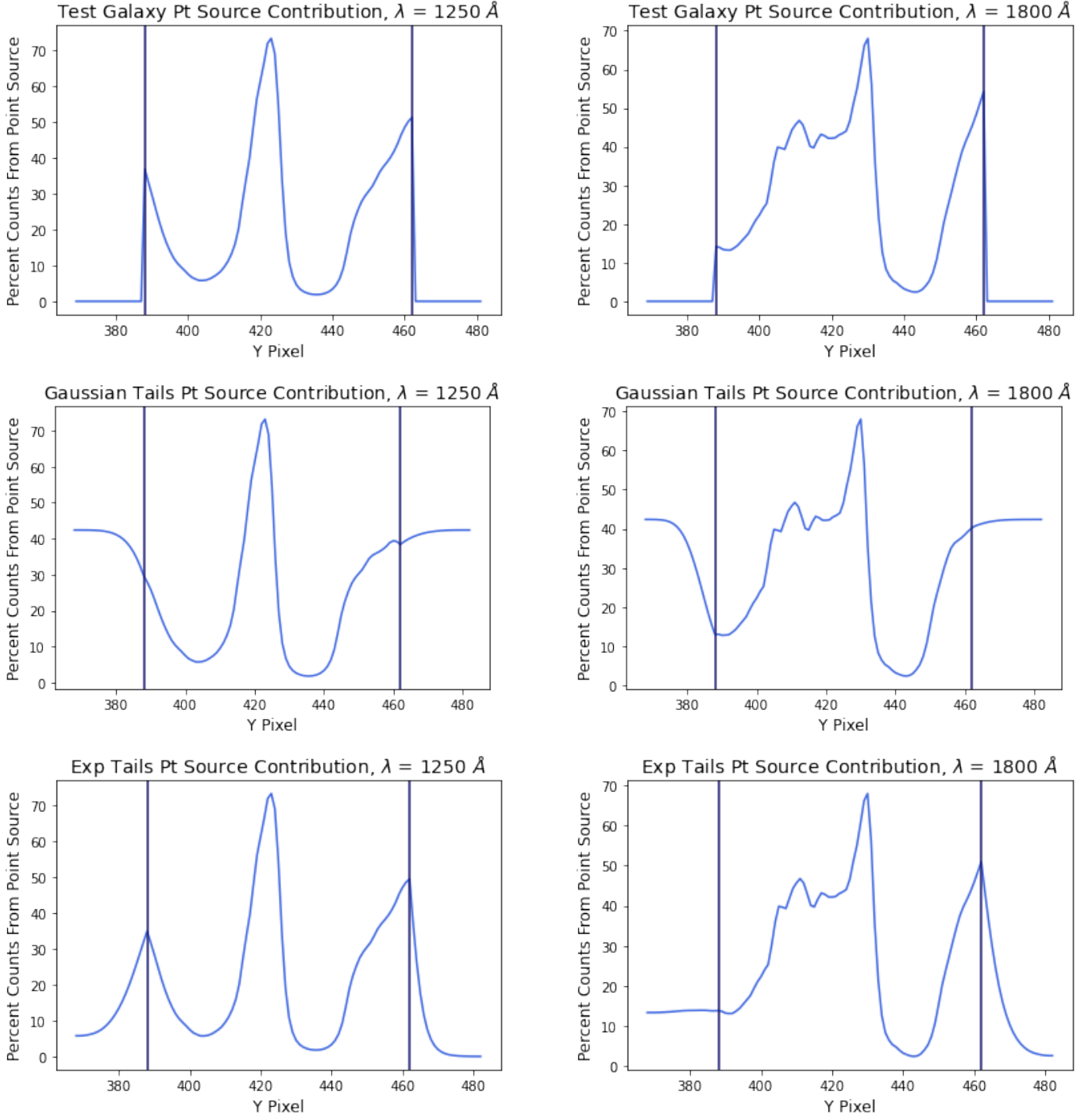


Figure 9. Point source contribution at different y pixels at both 1250 and 1800 angstroms for a test galaxy with SBR 17. The top shows the standard CDSF, the middle with Gaussian tails attached, and the bottom with exponentially decaying tails attached. Vertical lines indicate the edge of the book CDSF.

We find pixels that are diffuse dominant for all models up to an SBR of ~ 50 ; we also find pixels that remain point source dominant at all SBRs greater than 200. From these, we can define re-

gions of the COS detector that are most-likely to be diffuse or point source dominated for an unknown SBR. We can then mask the detector and compare these two regions to attempt

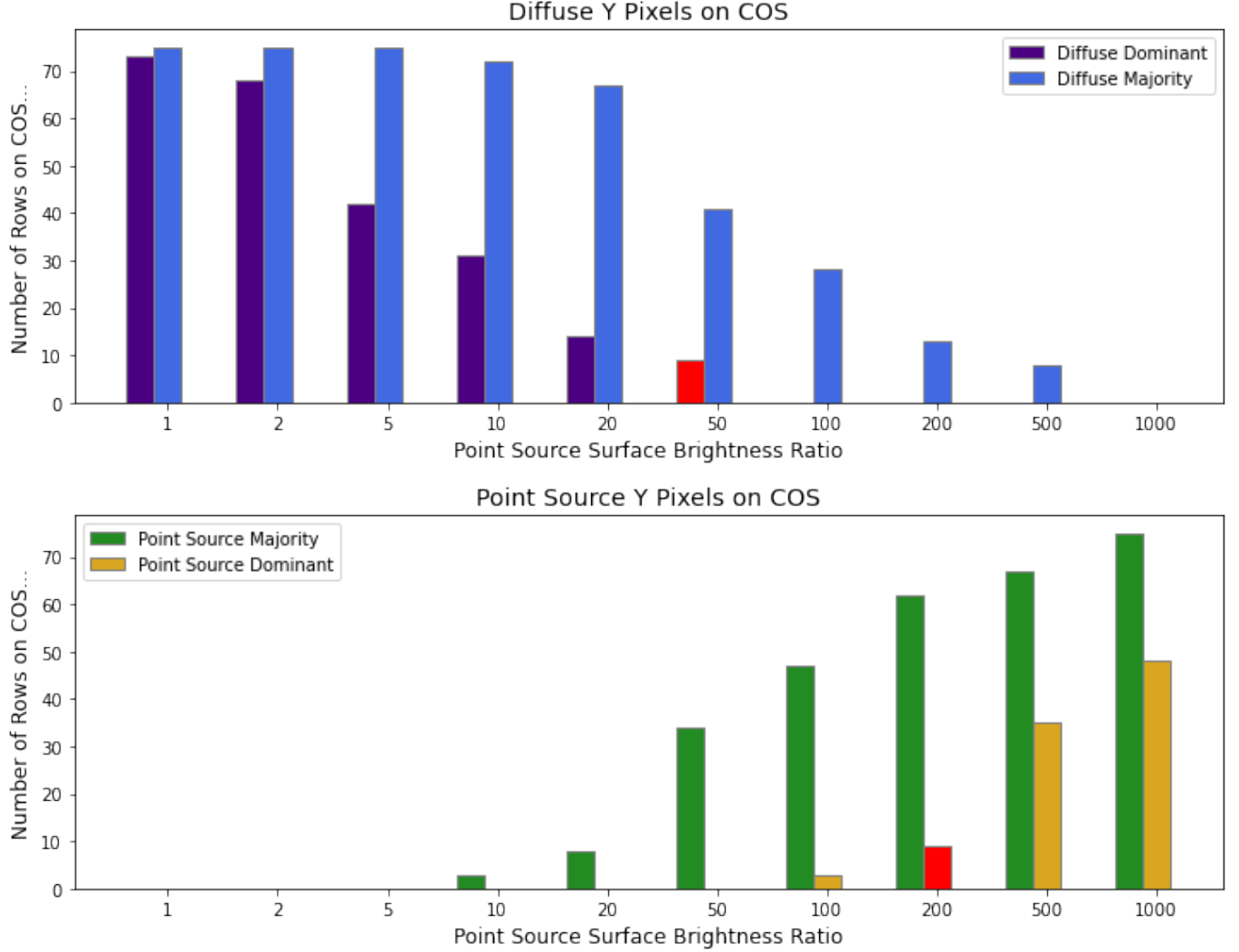


Figure 10. Charts showing number of y pixels with over 90% diffuse counts, a majority of diffuse counts, a majority of point source counts, and over 90% point source counts, dependent on the SBR. The pixels used to create the masks are highlighted in red.

Table 2. Categories of Y Pixels

Name	Description
Diffuse Dominant	> 90% diffuse counts
Diffuse Majority	> 50% diffuse counts
Pt Source Majority	> 50% pt source counts
Pt Source Dominant	> 90% pt source counts

to find a contribution from an extended host galaxy. The masks can be seen overlaid on one of the LzLCS galaxies in Figure 11.

2.5. Validity of Model

To test the validity of the model, we use the Ly α λ 1216 and OI λ 1304 geocoronal lines from LzLCS galaxy J103344+635317. Since both of these lines are from emission in Earth's atmosphere, the source is diffuse and overfilling the aperture. For Ly α , the counts are enough to completely overwhelm any from the target galaxy, meaning the cross section should match the model run with an entirely diffuse source. For OI, the geocoronal counts outnumber the galaxy counts, but the central galaxy does pro-

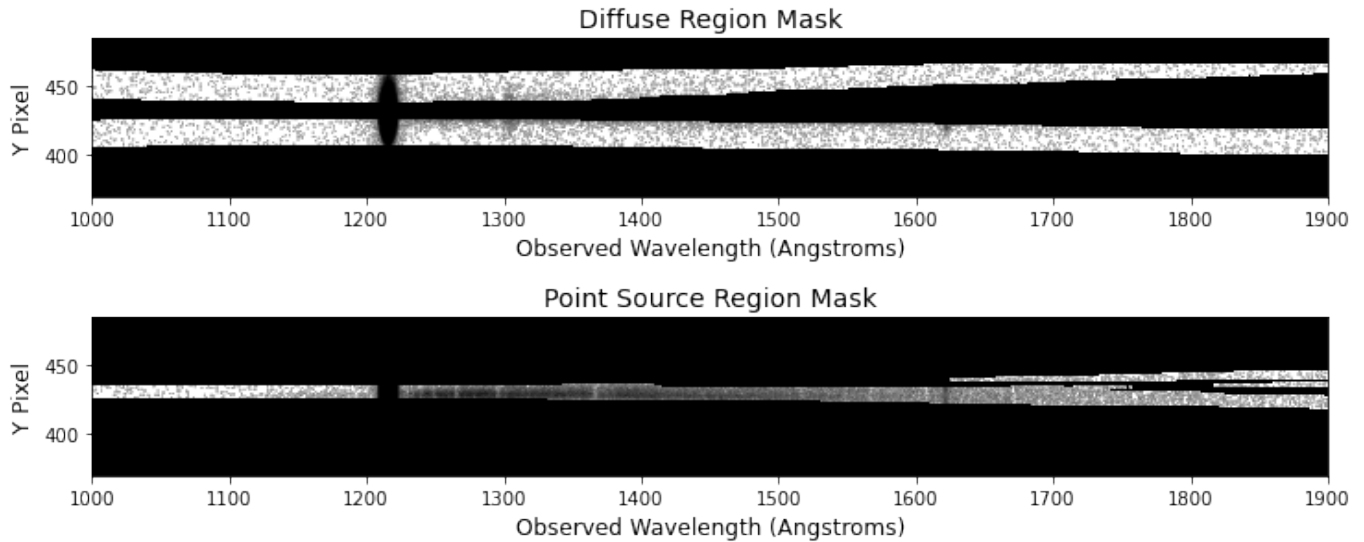


Figure 11. Masks showing the regions for capturing potential diffuse emission and potential point source emission. If the SBR is somewhere between 50-100, the spectra produced by these regions of the detector may show substantial differences.

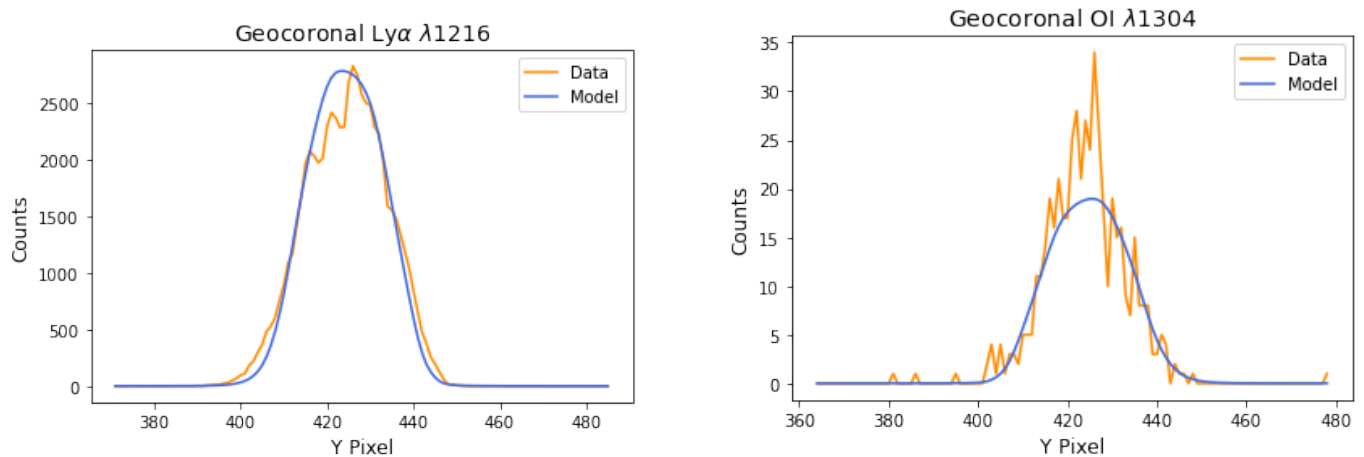


Figure 12. Checking the model predictions for a diffuse source overfilling the slit at 1216 \AA and 1304 \AA versus geocoronal measurements of $\text{Ly } \alpha$ and OI lines. The model fits roughly, but misses some intricacies of the instrument.

vide some additional counts. Therefore, $\text{Ly}\alpha$ is the preferred test case for this study.

Fitting the model to the data shows a correlation in the size and shape of the cross sections (see Figure 12). The error has shape and varies depending on the location on the detector with the highest percent error (around 50%) in the tail regions around y pixels 400 and 447. There is also an increase in error (to around 20%) in the center where the data has several

sharp peaks, perhaps a sign of gain sag (there is no reason for structure in the $\text{Ly}\alpha$, so it must come from *HST-COS*). Outside of those two regions the error is less than 5%. The shape suggests more extended wings than the book measurement of the CDSF. Fits to the OI emission feature are less correlated; there is a fair bit of noise due to the galaxy counts proving significant in comparison to the geocoronal counts for

this line. Overall, the model fits the diffuse geocoronal feature to first-order.

3. EXAMINATION OF LZLCS GALAXIES

The LzLCS published tables with data on all 66 observed galaxies in the sample. To choose our sample, we sort by the probability that the counts observed in the LyC spectrum for a galaxy could be caused by the background and choose those with the lowest probability. The lowest the LzLCS table goes is a probability of less than 1×10^{-8} , and there are 14 new observations in the LzLCS data that meet this criteria, ranging from redshift $z=0.243$ to $z=0.357$. [Flury et al. \(2022\)](#) examines many of the properties of these galaxies, but doesn't attempt any sort of spatial analysis, outside of a brief examination of the LyC profile in comparison to the non-LyC starlight profile and a calculation of the half-light radius from the NUV acquisition images. A typical calculated half-light radius is on the order of 0.5 kpc, which at redshift $z=0.3$ would be 0.22 arcseconds. At this angular size half of the light would fall on the two central y pixels of the detector before taking into account the CDSF.

Using GALEX data, the half light radius of NGC 2366 in the NUV is 1.3 kpc, which is at the larger end of the half light radii for the LzLCS GPs. This is approximately half the half light radius of NGC 2366 in the optical (2.7 kpc, [Hunter et al. 2001](#)). Notably, the LzLCS galaxies appear to be on average more compact than NGC 2366, which supports the idea of a super star forming region acting as a point source. Of course, as the galaxies are more compact, it is also less likely that a diffuse galaxy halo overfills the slit, making the diffuse regions generated by the model more likely to have a higher percentage of point source light.

3.1. Unpacking the Data

The LzLCS galaxies all fall under the same proposal ID 15626. The `corrtag_a` and `flt-`

sum files of the relevant galaxies are used. The corrected time tag files account for detector walk, thermal and geometric distortion, the Doppler shift caused by the movement of *HST*, and any optical select mechanism motions, and the image files are flat field corrected and linearized ([Soderblom 2021](#)). The spectra are produced using the G140L grating in the CEN-WAVE 800 Å setting and the FUV-XDL detector, Segment A. The acquisition images use MIRRORA and the NUV multi-anode micro-channel array detector.

From the `corrtag_a` files the YFULL and WAVELENGTH values are pulled for each incoming photon in order to create a 2D image of the detector (see Figure 13). WAVELENGTH is used instead of XFULL as sub-Angstrom resolution spectra are not required for the purposes of this paper. Since each galaxy has multiple exposures, the different exposures are coadded to create a full dataset for the galaxy. Exposure time is also pulled from the `corrtag_a` files.

The dark count rate for *HST*-COS varies over time, and the number of x pixels per angstrom also changes over the detector, so the noise is calculated by summing counts in a region in the same x position on the detector but over y pixels 575 to 625 (the galaxies are centered near y pixel 425). The noise is binned by angstrom to reflect the different numbers of x pixels per angstrom.

The location of the brightest y pixel compared to the central y pixel of the spectrum changes by wavelength due to the CDSF. To find the central y pixel of the spectrum, the data in a five angstrom bin from 1248 Å to 1252 Å is summed and the y pixel corresponding to the peak of the cross-dispersion profile is found. The pixel two pixels higher than this pixel is assumed to be the central pixel, which matches the difference given by the CDSF at this wavelength.

The counts from a 115 y pixel bin are summed to create the full spectrum. This is a much larger region than most papers use to bin ([Flury](#)

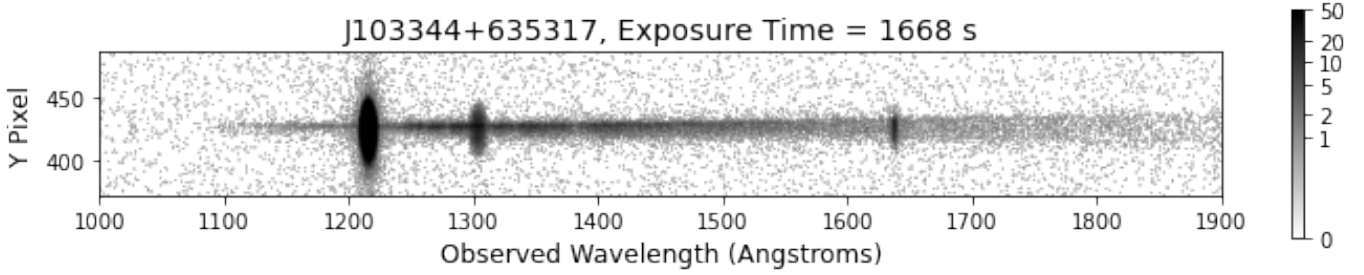


Figure 13. An example image of the COS detector for LzLCS galaxy J103344+635317. Geocoronal $\text{Ly}\alpha$, OI $\lambda 1304$, and source $\text{Ly}\alpha$ are all clearly visible. Counts are capped at 50 to cutoff geocoronal $\text{Ly}\alpha$. The image looks much closer to a point source than a diffuse source overfilling the aperture.

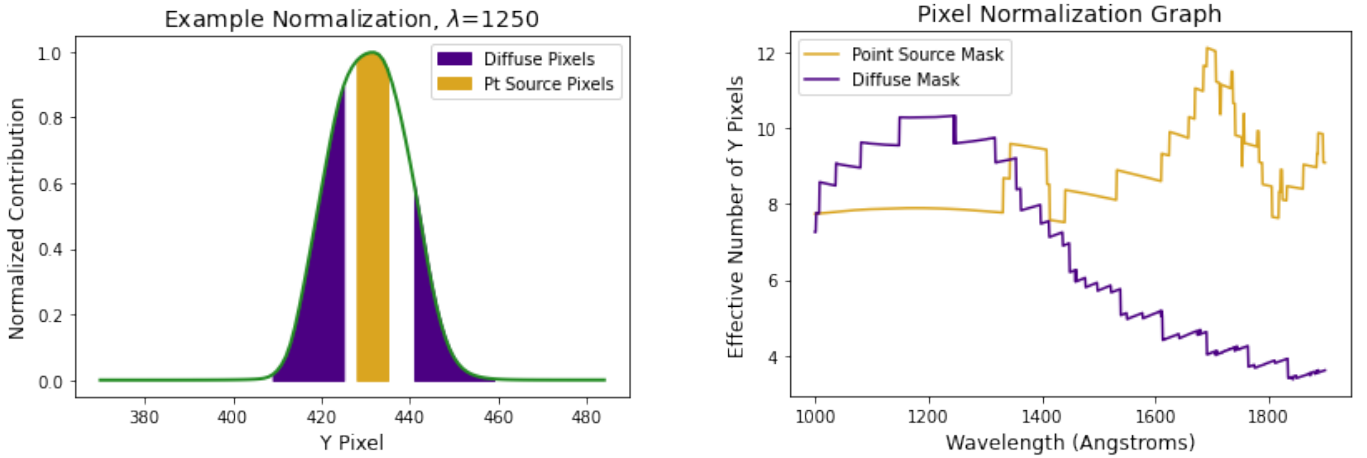


Figure 14. Left: Calculating the effective pixels at $\lambda=1250 \text{ \AA}$. The curve represents the result of diffuse source overfilling the slit. The area under the curve is the effective number of y pixels for the masks at this wavelength. Right: The effective pixels for the point source and diffuse masks as a function of wavelength. The jumps indicate pixels being added or taken away, and the curve shows the underlying width of the CDSF.

et al. (2022) uses a 25 pixel bin corresponding to the $2.5''$ aperture); the size indicates the furthest edges of the combined transmission function and book CDSF. We use a larger range to try and capture potential diffuse counts. Next, the two masks (see Figure 11) are applied to create two additional spectra, with one expected to have the largest contribution from a diffuse source based on Figures 9 and 10, and the other expected to have the largest contribution from a point source. The mean background noise per angstrom per pixel is then scaled appropriately to the number of pixels in each spectra, subtracted from all three spectra, and any values that fall below zero counts are masked.

The effective number of pixels at each wavelength is calculated (see Figure 14) for both the diffuse and point masks, and the spectra are then divided by that number so they can be directly compared. Finally, the spectra are converted to flux units with the effective area from the COS Handbook (James et al. 2022). The full spectrum flux generated from this data reduction reproduces the spectra presented in the LzLCS study.

The Signal-to-Noise Ratio (SNR) of the diffuse and point source spectra is also calculated assuming a Poisson distribution for the detector dark counts. In general, the higher wavelengths have a small SNR for both detector regions. For

the example spectra, the main graphs include the $\text{SNR} = 1$ line.

3.2. Additional Calculations

The finder scope NUV images are also pulled for each galaxy to give a rough image of the galaxy in the NUV. The galaxies in question are expected to be dimmer in the NUV due to the young ages of the starbursts, when O and B stars still dominate the spectrum, but the shape of the finder scope images may give some information on potential extended emission (these images were used by [Flury et al. \(2022\)](#) to calculate the half light radius).

Finally, a number of additional calculations are done to try and identify the galaxies most likely to have a diffuse halo. The percentage of counts falling in the diffuse and point source regions is calculated for a couple of large bins across the wavelength range. This gives a quick check to see if a significant amount of light is falling in the diffuse region. Next, cross sections of the model are fit to cross sections of the data to estimate the SBR of the data. A general ‘continuum cross section’ is generated by binning all the light from observed $\lambda = 1240 \text{ \AA}$ to just below source $\text{Ly}\alpha$, cutting out the area around $\text{OI } \lambda 1304$. In addition, a comparison is done with the source $\text{Ly}\alpha$ cross section to see if it is wider than the rest of the continuum (similar to the work in [Yang et al. \(2018\)](#)).

3.3. Results

All 14 LyC emitters examined have cross sections more similar to point sources than extended objects. A comparison of the detector image in [Figure 13](#) and the model galaxy generation in [Figure 6](#) shows a spectrum that looks much closer to a point source, and this is true for all the galaxies. As the measured NUV half-light radii of the galaxies average around $0.22''$, this is not surprising. While the point source region produces spectra that resemble young starburst galaxies, the diffuse region often has

a low SNR, consistent with no appreciable diffuse emission. For 5 of the 14 galaxies the median SNR in the wavelength range analyzed (observed wavelengths between 1150 \AA and 1900 \AA) was zero, meaning over half of the wavelength bins were below the mean background noise. Only one galaxy had a median SNR above 1 in the diffuse region.

The flux corresponding with the mean background noise changes with both exposure time and effective area; thus, the SNR is used as the main measurement in this report. Typical flux values for noise run from on the order of $10^{-19} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$ per normalized y pixel around 1200 \AA to on the order of $10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$ per normalized y pixel around 1800 \AA .

There were two areas of the spectrum that are more likely to generate a higher SNR: source $\text{Ly}\alpha$ and between the $\lambda 1212$ and $\lambda 1304$ geocoronal lines. Source $\text{Ly}\alpha$ produces over an order of magnitude more counts than the surrounding continuum, and the area between the geocoronal lines has an effective area of over 1000 cm^2 . Of these, the $\text{Ly}\alpha$ line is visible in the diffuse region in all of the galaxies and the area between the geocoronal lines has a median SNR greater than 1 in the diffuse region for 8 of 14 galaxies.

Comparing the number of counts that fall in the diffuse and point source regions to the full detector area reaffirms the compact nature of the targets - most of the wavelengths have less than 10% of their counts falling in the diffuse region (see [Figure 16](#)). The LyC bin tends to measure the highest percentage of counts in the diffuse region (over 20% on 10 of the 14 galaxies), although we suspect this is due to the small sample size in that area of the detector.

The SBR fit calculation for the continuum bin results in a ratio of 1000 to 1 for all of the galaxies, which is the highest ratio we tested for and consistent with point-source domination.

For the source $\text{Ly}\alpha$ line, on the other hand, the model tends to predict a SBR of around 20

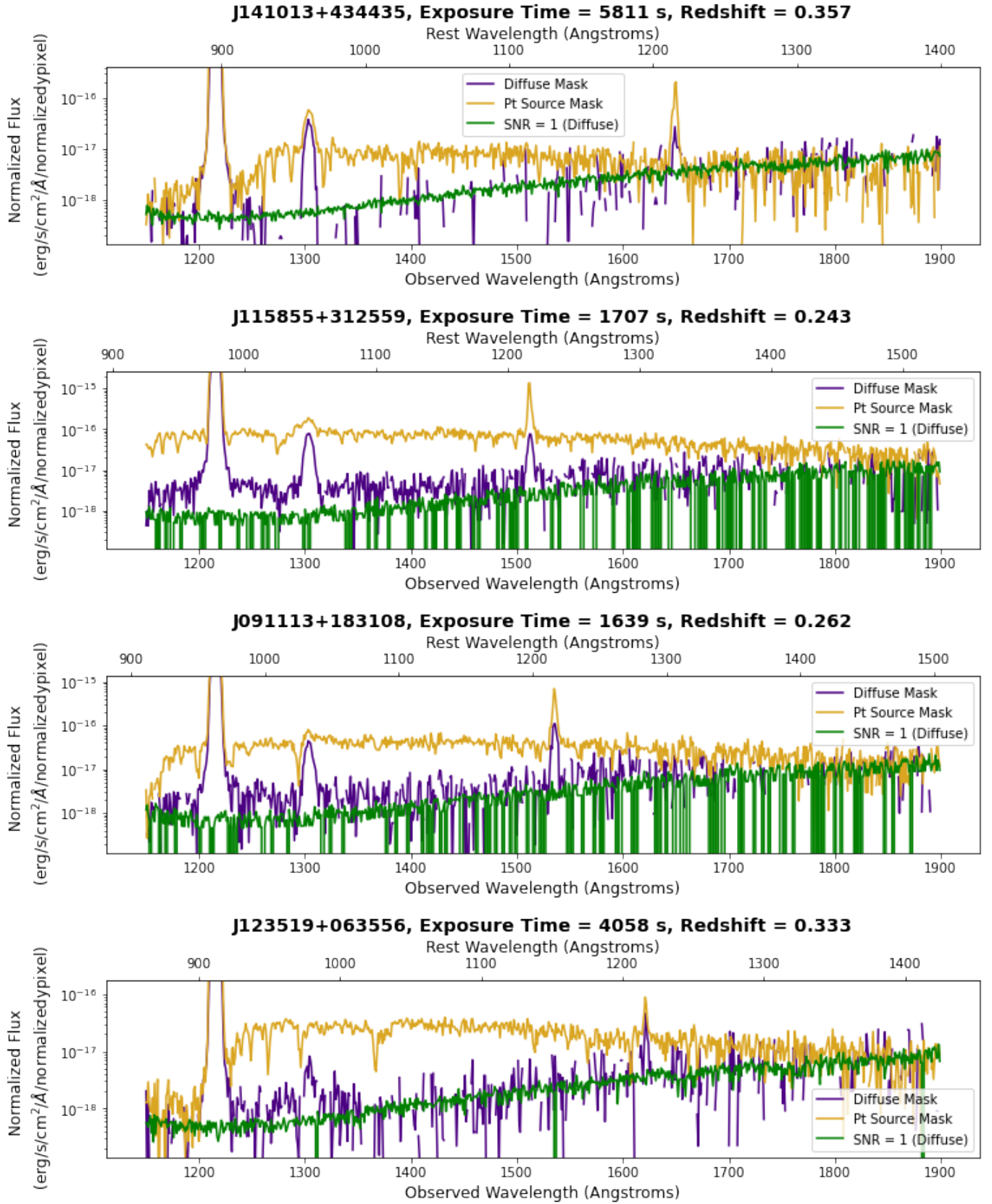


Figure 15. Four of the LzLCS galaxies chosen for discussion.

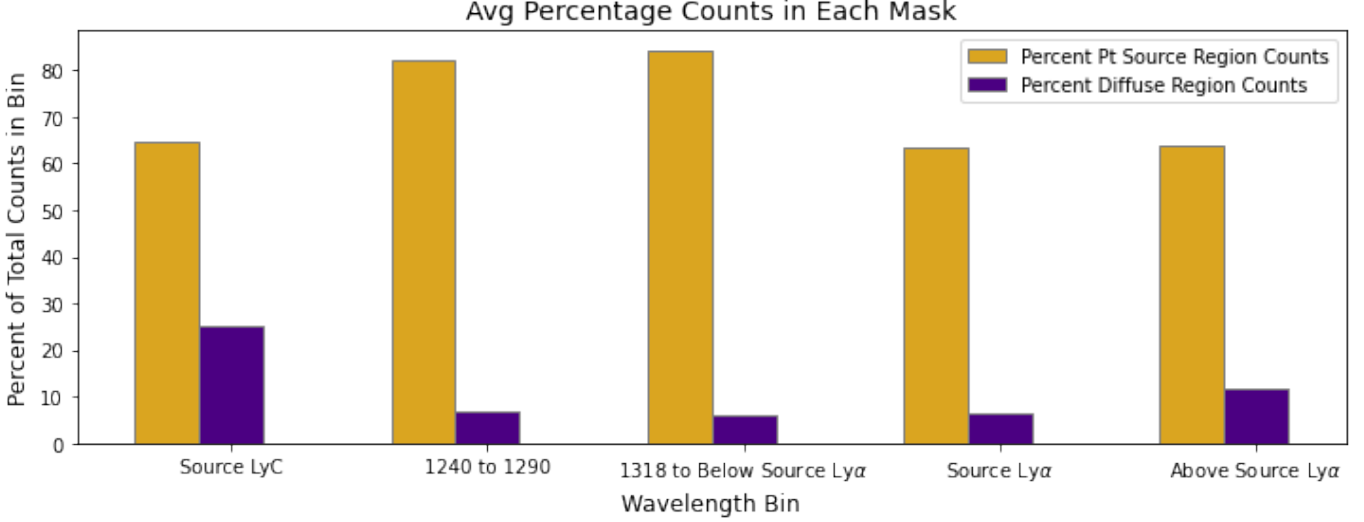


Figure 16. The average percentage counts in the diffuse and point source regions compared to the full detector area, sorted into five different wavelength bins.

to 1 - in other words, the cross section here is wider and indicates some level of diffuse emission not from the center of the aperture. This changes from galaxy to galaxy, with ratios from 12 (the most extended Ly α cross section) to 22 (the least extended Ly α cross section).

3.4. Examples of Processed Data

There are a couple of galaxies that pop out as examples of the various results, with spectra shown in Figure 15 and NUV images shown in Figure 17. The first is J141013+434435, which has the least number of counts in the diffuse region - only the geocoronal lines and source Ly α generate enough counts to rise strongly above SNR of 1. It is probably the most ‘point source-like’ galaxy in the sample, which is also apparent in its NUV image, although some of that may be due to having the largest redshift at $z=0.357$.

The next is the galaxy with the most flux in the diffuse region, J115855+312559. This galaxy has a median SNR before the source Ly α line of 3.0, and it has the most extended source Ly α line with a calculated SBR of 12. It is both the brightest galaxy (has the highest flux) and closest (lowest redshift). Because the continuum SBR is still consistent with point source

domination and the point source and diffuse region spectra are similar, it is likely that most of the diffuse counts are simply point source counts spread into the diffuse region by the CDSF.

The third is the galaxy with the second most flux in the diffuse bin, J09113+183108. It has a SBR that indicates the second most extended source Ly α line. Supporting this, it has the third most percent counts in the source Ly α bin (11.1). The same caveats as J115855+312559 apply: it is also at a closer redshift and slightly brighter than most of the other galaxies.

The final galaxy, J1213519+063556, shows some evidence of diffuse emission at a higher redshift. It has the third strongest SNR in the diffuse region, and the second most percent counts in the source Ly α bin (14.3). On the other hand, the cross section of the source Ly α line indicates a SBR of 20, which is about average for this set of galaxies. However, this is not one of the closer galaxies, and appears slightly broader in the NUV image.

4. CONCLUSION AND FUTURE WORK

4.1. Conclusion

The 14 LzLCS galaxies examined in this study closely resemble point sources; pulling differ-

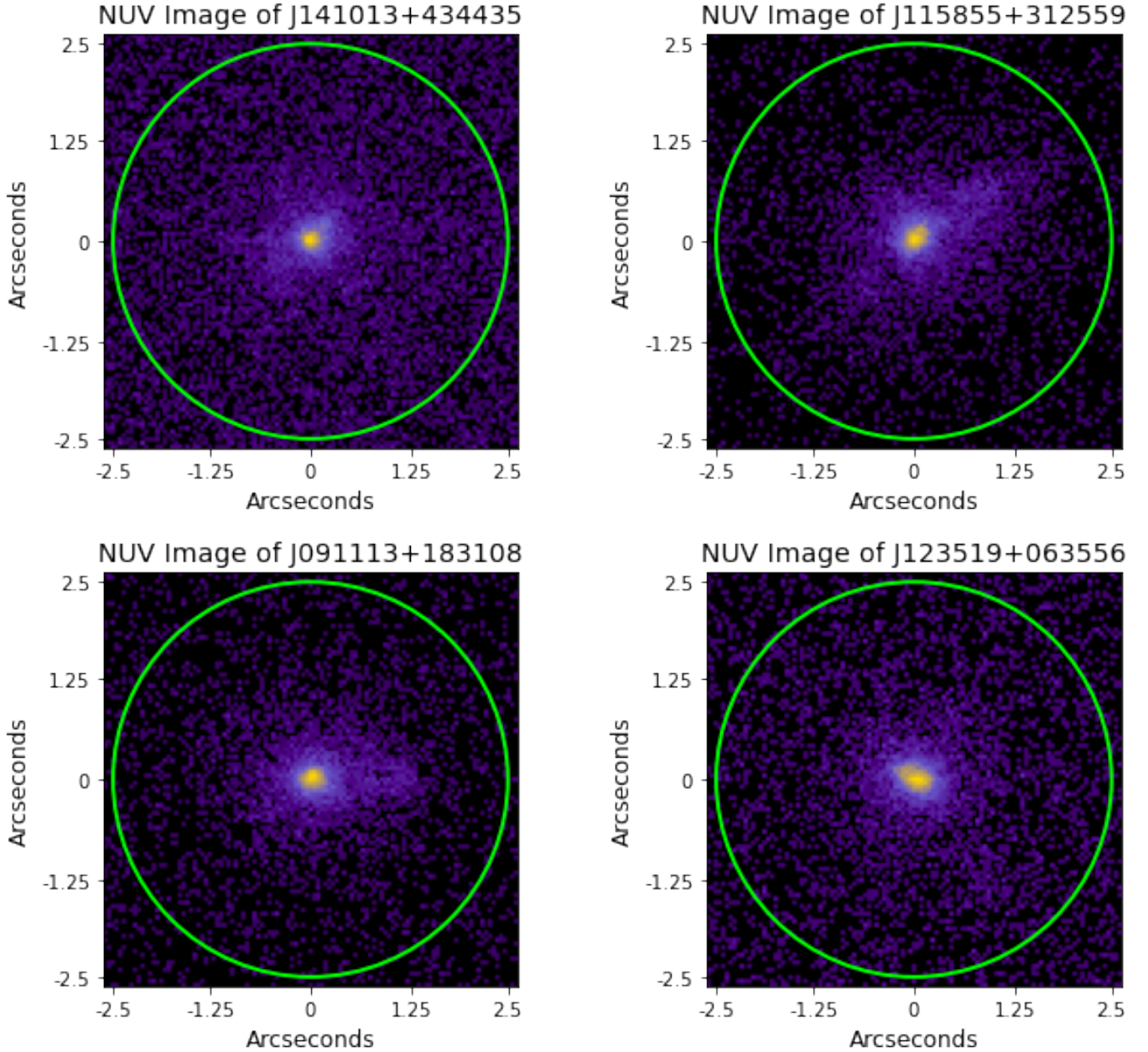


Figure 17. *NUV acquisition images of the example galaxies. The green circle indicates the 2.5'' aperture.*

ent spectra from different portions of the detector was not possible. COS is, in general, ill-disposed to this task due to the breadth of the CDSF and its changes over the different wavelengths.

The source Ly α line for every galaxy does indicate a larger Ly α emitting galaxy halo, and there does appear to be a difference in angular size of this halo from galaxy to galaxy.

The two closest galaxies, J115855+312559 and J091113+183108, have the widest halos and may be suitable for follow up, and one other galaxy, J123519+063556, has enough SNR in the diffuse region of the detector that it may also be worthy of follow up.

4.2. Future Work

Additional analysis is needed on the galaxies in question to see if they are good targets for

follow up. In particular, for the closer, brighter galaxies, it needs to be checked to see if the counts falling in the diffuse mask can be readily explained simply as a function of the CDSF and brightness of the galaxy. As the two best ones were at the lowest redshift, perhaps an even closer galaxy with GP characteristics might give more insight as to the morphology and feedback mechanisms in these galaxies.

A deconvolution of the CDSF at the $\text{Ly}\alpha$ line is an additional method to measure line cross section and should be used to check the results against the SBR calculated by the model. The LyC region should also be checked for shape using the SBR and deconvolution methods to see how it compares to the main continuum and source $\text{Ly}\alpha$ cross sections.

The model should be checked against a full spectrum diffuse source overfilling the aperture. Some exist for the G130M grating, which has a different CDSF, but the same methodology can be used to construct a model for G130M to check.

Most importantly, the model parameters should be expanded to include a diffuse region of different sizes (instead of just overfilling the aperture), different locations in the aperture for the bright region, and different redshifts and luminosities of the galaxies. This will allow generation of a galaxy catalog that will indicate what sort of galaxy could be a good candidate to distinguish between a bright central region and a less bright surrounding region on the COS detector. This may change target selection metrics for follow up studies.

After the further analysis is done, either the galaxies mentioned here or other targets found from the additional analysis will lead to a proposal for imaging using either *HST* for the ultraviolet or a ground based telescope such as Keck for optical. The higher angular resolution will allow us to study the morphology of these important galaxies.

Finally, we will launch the INtegral-field Far-Ultraviolet Spectrographic Experiment sounding rocket in Fall 2024 to study NGC 2366 and Mrk71.

APPENDIX

A. SPECTRA AND IMAGES OF OTHER ANALYZED GALAXIES

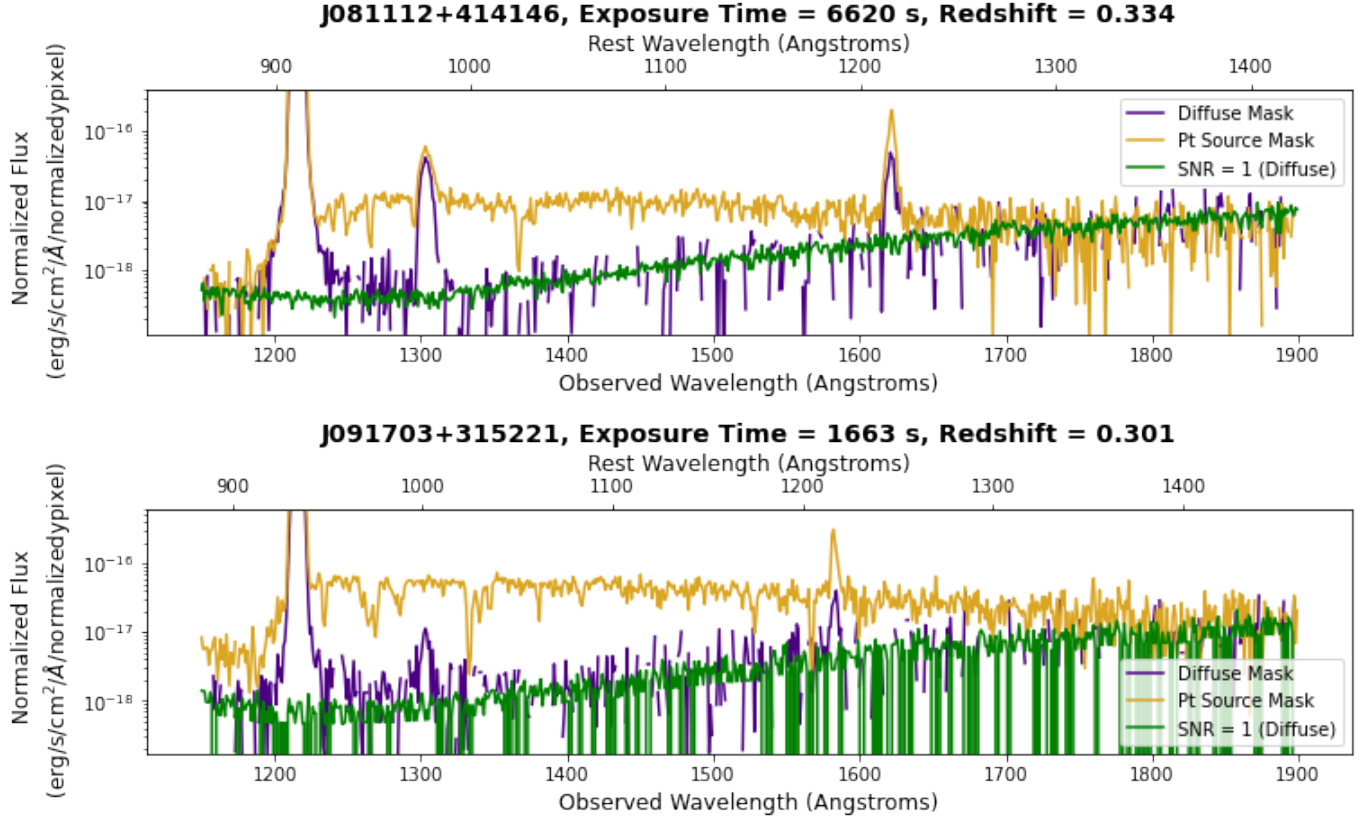


Figure 18. *Two processed spectra of the LzLCS galaxies.*

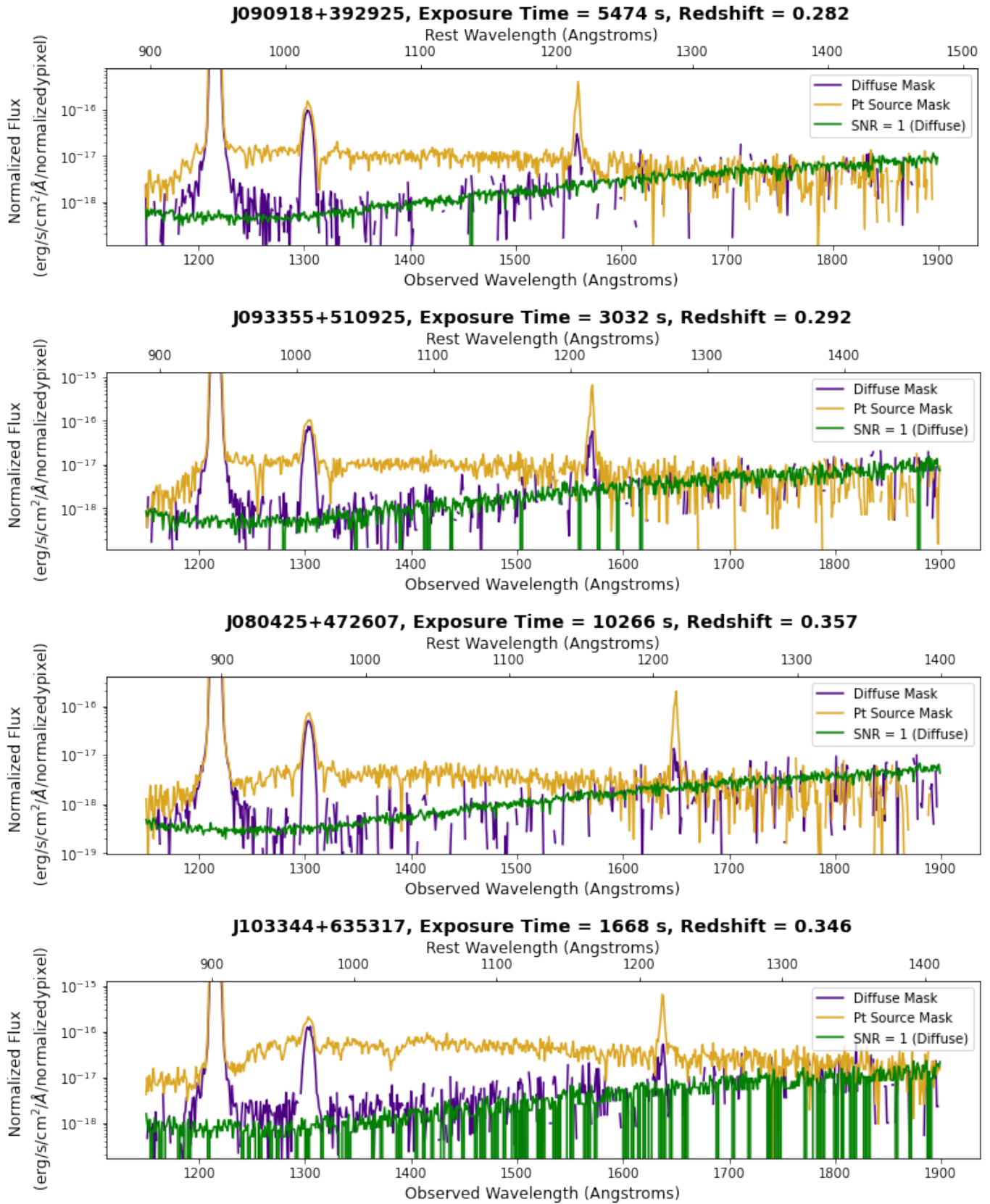


Figure 19. Four processed spectra of the LzLCS galaxies.

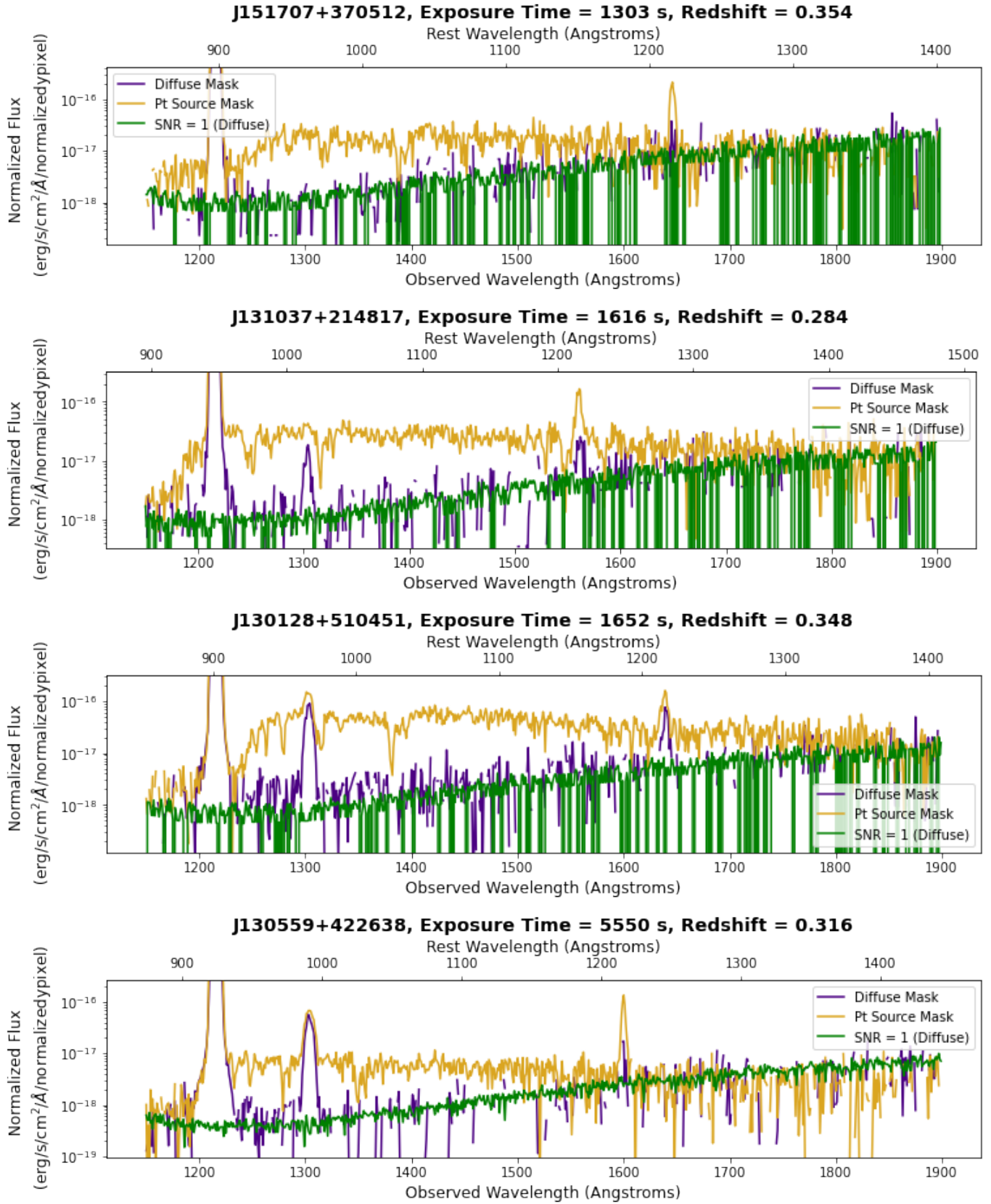


Figure 20. Four processed spectra of the LzLCS galaxies.

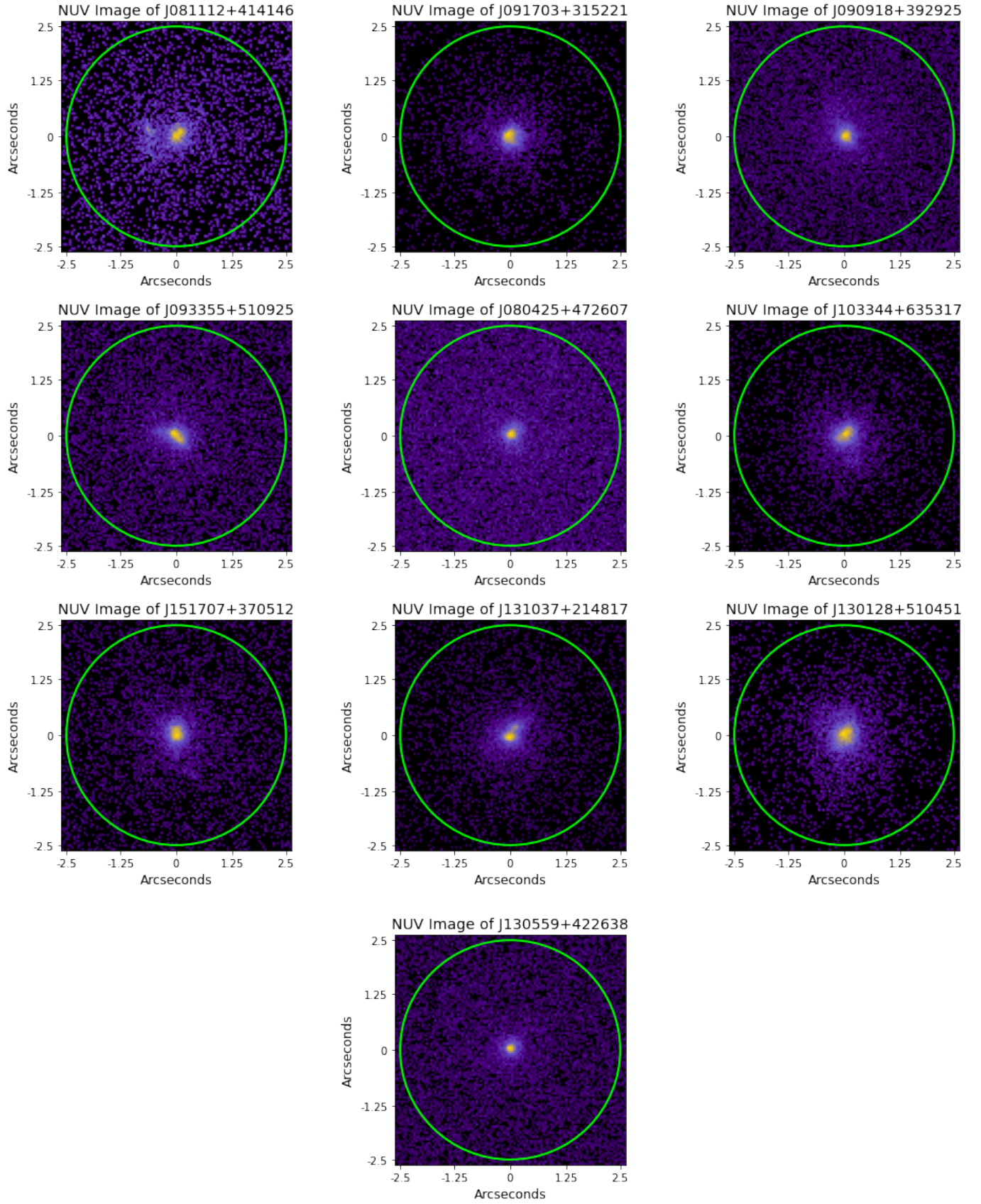


Figure 21. *NUV acquisition images of the analyzed LzLCS galaxies.*

REFERENCES

- Alvarez, M. A., Finlator, K., & Trenti, M. 2012, *ApJ*, 759, L38, doi: [10.1088/2041-8205/759/2/L38](https://doi.org/10.1088/2041-8205/759/2/L38)
- Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, *PASP*, 93, 5, doi: [10.1086/130766](https://doi.org/10.1086/130766)
- Becker, R. H., Fan, X., White, R. L., et al. 2001, *AJ*, 122, 2850, doi: [10.1086/324231](https://doi.org/10.1086/324231)
- Bergvall, N., Zackrisson, E., Andersson, B.-G., et al. 2006, *A&A*, 448, doi: [10.1051/0004-6361:20053788](https://doi.org/10.1051/0004-6361:20053788)
- Cardamone, C., Schawinski, K., Sarzi, M., et al. 2009, *MNRAS*, 399, 1191, doi: [10.1111/j.1365-2966.2009.15383.x](https://doi.org/10.1111/j.1365-2966.2009.15383.x)
- Ciardi, B., Ferrara, A., & White, S. D. M. 2003, *MNRAS*, 344, L7, doi: [10.1046/j.1365-8711.2003.06976.x](https://doi.org/10.1046/j.1365-8711.2003.06976.x)
- Cooke, J., Ryan-Weber, E. V., Garel, T., & Díaz, C. G. 2014, *MNRAS*, 441, 837, doi: [10.1093/mnras/stu635](https://doi.org/10.1093/mnras/stu635)
- de Barros, S., Vanzella, E., Amorín, R., et al. 2016, *A&A*, 585, A51, doi: [10.1051/0004-6361/201527046](https://doi.org/10.1051/0004-6361/201527046)
- Deharveng, J. M., Buat, V., Le Brun, V., et al. 2001, *A&A*, 375, 805, doi: [10.1051/0004-6361:20010920](https://doi.org/10.1051/0004-6361:20010920)
- Dressler, A., Henry, A., Martin, C. L., et al. 2015, *ApJ*, 806, 19
- Drissen, L., Roy, J.-R., Robert, C., & Devost, D. 2000, *AJ*, 119, 688
- Finkelstein, S. L., D’Aloisio, A., Paardekooper, J.-P., et al. 2019, *ApJ*, 879, 36, doi: [10.3847/1538-4357/ab1ea8](https://doi.org/10.3847/1538-4357/ab1ea8)
- Flury, S. R., Jaskot, A. E., Ferguson, H. C., et al. 2022, *ApJS*, 260, 1, doi: [10.3847/1538-4365/ac5331](https://doi.org/10.3847/1538-4365/ac5331)
- Fox, A. 2018, *STScI*, 2018, 08
- Giallongo, E., Grazian, A., Fiore, F., et al. 2015, *A&A*, 578, A83, doi: [10.1051/0004-6361/201425334](https://doi.org/10.1051/0004-6361/201425334)
- Gnedin, N. Y., & Ostriker, J. P. 1997, *ApJ*, 486, 581, doi: [10.1086/304548](https://doi.org/10.1086/304548)
- Green, J. C., Froning, C. S., Osterman, S., et al. 2012, *ApJ*, 744, 60
- Grimes, J. P., Heckman, T., Strickland, D., et al. 2007, *ApJ*, 668, 891, doi: [10.1086/521353](https://doi.org/10.1086/521353)
- Gunn, J. E., & Peterson, B. A. 1965, *ApJ*, 142, 1633, doi: [10.1086/148444](https://doi.org/10.1086/148444)
- Henry, A., Scarlata, C., Martin, C. L., & Erb, D. 2015, *ApJ*, 809, 19, doi: [10.1088/0004-637x/809/1/19](https://doi.org/10.1088/0004-637x/809/1/19)
- Hunter, D. A., Elmegreen, B. G., & van Woerden, H. 2001, *ApJ*, 556, 773, doi: [10.1086/321611](https://doi.org/10.1086/321611)
- Hurwitz, M., Jelinsky, P., & Dixon, W. V. D. 1997, *ApJ*, 481, L31, doi: [10.1086/310649](https://doi.org/10.1086/310649)
- Inoue, A. K., Iwata, I., & Deharveng, J.-M. 2006, *MNRAS*, 371, L1, doi: [10.1111/j.1745-3933.2006.00195.x](https://doi.org/10.1111/j.1745-3933.2006.00195.x)
- Izotov, Y. I., Guseva, N. G., & Thuan, T. X. 2011, *ApJ*, 728, 161, doi: [10.1088/0004-637x/728/2/161](https://doi.org/10.1088/0004-637x/728/2/161)
- Izotov, Y. I., Orlitova, I., Schaerer, D., et al. 2016a, *Nat*, 529, 178
- Izotov, Y. I., Schaerer, D., Thuan, T. X., et al. 2016b, *MNRAS*, 461, 3683
- Izotov, Y. I., Schaerer, D., Worseck, G., et al. 2018a, *MNRAS*, 474, 4514
- Izotov, Y. I., Worseck, G., Schaerer, D., et al. 2018b, *MNRAS*, 478, 4851
- James, B. L., Auger, M., Aloisi, A., Calzetti, D., & Kewley, L. 2016, *ApJ*, 816, 40
- James, B. L., Hirschauer, A. S., Dashtamirova, S., & Fischer, W. 2022, in *COS Instrument Handbook v. 14.0 (STScI)*, 290
- Jaskot, A. E., & Oey, M. S. 2013, *ApJ*, 766, 91, doi: [10.1088/0004-637X/766/2/91](https://doi.org/10.1088/0004-637X/766/2/91)
- Kollmeier, J. A., Weinberg, D. H., Oppenheimer, B. D., et al. 2014, *ApJ*, 789, L32, doi: [10.1088/2041-8205/789/2/L32](https://doi.org/10.1088/2041-8205/789/2/L32)
- Komarova, L., Oey, M. S., Krumholz, M. R., et al. 2021, *ApJL*, 920, L46
- Leitet, E., Bergvall, N., Hayes, M., Linné, S., & Zackrisson, E. 2013, *A&A*, 553, A106, doi: [10.1051/0004-6361/201118370](https://doi.org/10.1051/0004-6361/201118370)
- Leitherer, C., Ferguson, H. C., Heckman, T. M., & Lowenthal, J. D. 1995, *ApJ*, 454, doi: [10.1086/309760](https://doi.org/10.1086/309760)
- Lelli, F., Verheijen, M., & Fraternali, F. 2014, *A&A*, 566, A71
- Lintott, C. J., Schawinski, K., Slosar, A., et al. 2008, *MNRAS*, 389, 1179, doi: [10.1111/j.1365-2966.2008.13689.x](https://doi.org/10.1111/j.1365-2966.2008.13689.x)

- Lupton, R., Gunn, J. E., Ivezić, Z., Knapp, G. R., & Kent, S. 2001, in *Astronomical Society of the Pacific Conference Series*, Vol. 238, *Astronomical Data Analysis Software and Systems X*, ed. J. Harnden, F. R., F. A. Primini, & H. E. Payne, 269.
<https://arxiv.org/abs/astro-ph/0101420>
- Madau, P., & Haardt, F. 2015, *ApJ*, 813, L8, doi: [10.1088/2041-8205/813/1/L8](https://doi.org/10.1088/2041-8205/813/1/L8)
- McQuinn, K. B. W., Mitchell, N. P., & Skillman, E. D. 2015, *ApJS*, 218, 29, doi: [10.1088/0067-0049/218/2/29](https://doi.org/10.1088/0067-0049/218/2/29)
- Micheva, G., Oey, M. S., Jaskot, A. E., & James, B. L. 2017, *ApJ*, 845, 165, doi: [10.3847/1538-4357/aa830b](https://doi.org/10.3847/1538-4357/aa830b)
- Morrissey, P., Schiminovich, D., Barlow, T. A., et al. 2005, *ApJL*, 619, L7, doi: [10.1086/424734](https://doi.org/10.1086/424734)
- Moustakas, J., & Kennicutt, Robert C., J. 2006, *ApJS*, 164, 81
- Nakajima, K., Ellis, R. S., Robertson, B. E., Tang, M., & Stark, D. P. 2020, *ApJ*, 889, 161, doi: [10.3847/1538-4357/ab6604](https://doi.org/10.3847/1538-4357/ab6604)
- Nakajima, K., & Ouchi, M. 2014, *MNRAS*, 442, 900, doi: [10.1093/mnras/stu902](https://doi.org/10.1093/mnras/stu902)
- Paardekooper, J.-P., Khochfar, S., & Vecchia, C. D. 2015, *MNRAS*, 451, 2544, doi: [10.1093/mnras/stv1114](https://doi.org/10.1093/mnras/stv1114)
- Prichard, L., Welty, D., & Jones, A. 2022, in *STIS Instrument Handbook for Cycle 29 v. 21*, Vol. 20 (STScI), 20
- Razoumov, A. O., & Sommer-Larsen, J. 2010, *ApJ*, 710, 1239, doi: [10.1088/0004-637x/710/2/1239](https://doi.org/10.1088/0004-637x/710/2/1239)
- Robertson, B. E., Furlanetto, S. R., Schneider, E., et al. 2013, *ApJ*, 768, 71
- Romano, M., Grazian, A., Giallongo, E., et al. 2019, *A&A*, 632, A45, doi: [10.1051/0004-6361/201935550](https://doi.org/10.1051/0004-6361/201935550)
- Shapley, A. E., Steidel, C. C., Pettini, M., Adelberger, K. L., & Erb, D. K. 2006, *ApJ*, 651, 688, doi: [10.1086/507511](https://doi.org/10.1086/507511)
- Shapley, A. E., Steidel, C. C., Strom, A. L., et al. 2016, *ApJL*, 826, L24, doi: [10.3847/2041-8205/826/2/L24](https://doi.org/10.3847/2041-8205/826/2/L24)
- Siana, B., Teplitz, H. I., Ferguson, H. C., et al. 2010, *ApJ*, 723, 241, doi: [10.1088/0004-637x/723/1/241](https://doi.org/10.1088/0004-637x/723/1/241)
- Soderblom, D. R. 2021, in *COS Data Handbook v. 5.0* (STScI), 208
- Tolstoy, E., Saha, A., Hoessel, J. G., & McQuade, K. 1995, *AJ*, 110, 1640
- Trebitsch, M., Volonteri, M., Dubois, Y., & Madau, P. 2018, *MNRAS*, 478, 5607, doi: [10.1093/mnras/sty1406](https://doi.org/10.1093/mnras/sty1406)
- Verhamme, A., Orlitová, I., Schaerer, D., et al. 2017, *A&A*, 597, A13, doi: [10.1051/0004-6361/201629264](https://doi.org/10.1051/0004-6361/201629264)
- Volonteri, M., & Gnedin, N. Y. 2009, *ApJ*, 703, 2113, doi: [10.1088/0004-637x/703/2/2113](https://doi.org/10.1088/0004-637x/703/2/2113)
- Wise, J. H., Demchenko, V. G., Halicek, M. T., et al. 2014, *MNRAS*, 442, 2560, doi: [10.1093/mnras/stu979](https://doi.org/10.1093/mnras/stu979)
- Woodgate, B. E., Kimble, R. A., Bowers, C. W., et al. 1998, *PASP*, 110, 1183, doi: [10.1086/316243](https://doi.org/10.1086/316243)
- Yajima, H., Choi, J.-H., & Nagamine, K. 2011, *MNRAS*, 412, 411, doi: [10.1111/j.1365-2966.2010.17920.x](https://doi.org/10.1111/j.1365-2966.2010.17920.x)
- Yamanaka, S., Inoue, A. K., Yamada, T., et al. 2020, *MNRAS*, 498, 3095, doi: [10.1093/mnras/staa2507](https://doi.org/10.1093/mnras/staa2507)
- Yang, H., Malhotra, S., Rhoads, J. E., et al. 2018, *ApJ*, 838, 4, doi: [10.3847/1538-4357/aa6337](https://doi.org/10.3847/1538-4357/aa6337)
- York, D. G., Adelman, J., John E. Anderson, J., et al. 2000, *AJ*, 120, 1579, doi: [10.1086/301513](https://doi.org/10.1086/301513)
- Zackrisson, E., Inoue, A. K., & Jensen, H. 2013, *ApJ*, 777, 39, doi: [10.1088/0004-637X/777/1/39](https://doi.org/10.1088/0004-637X/777/1/39)