Draft version February 17, 2011
Preprint typeset using LaTeX style emulateapj v. 11/12/01

INSIGHTS ON THE FORMATION, EVOLUTION, AND ACTIVITY OF MASSIVE GALAXIES FROM ULTRA-COMPACT AND DISKY GALAXIES AT $Z = 2 - 3$

Tim Weinzirl[1], Shardha Jogee[1], Christopher J. Conselice[2], Casey Papovich[3], Ranga-Ram Chary[4], Asa Bluck[5], Ruth Grützbauch[2], Fernando Buitrago[2], Bahram Mobasher[6], Ray A. Lucas[7], Mark Dickinson[8], Amanda Bauer[9]

Draft version February 17, 2011

ABSTRACT

The assembly modes via which galaxies develop their present-day mass and structure remain hotly debated. We explore this issue using one of the largest samples of massive galaxies at $z = 1-3$ (166 with $M_\star \geq 5 \times 10^{10}$ $M_\odot$ and 73 with $M_\star \geq 10^{11}$ $M_\odot$), with NICMOS F160W observations from the GOODS NICMOS Survey (GNS), along with complementary ACS, *Spitzer* IRAC and MIPS, and Chandra X-ray data. *(1) Star formation and Cold Gas Content:* We estimate star formation rates (SFR_{IR}) using IR luminosities derived from *Spitzer* 24 μm data. Only 74/166 (44.6 $\pm$3.9%) of massive ($M_\star \geq 5 \times 10^{10}$ $M_\odot$) galaxies at $z = 1 - 3$ have SFR_{IR} above the 5σ detection limit. The mean SFR_{IR} rises with redshift over $z = 1-3$, going from ~ 44 to 419 $M_\odot$ yr^{-1} (after excluding AGN candidates). Assuming a Schmidt-Kennicutt law, we find that the massive non-AGN galaxies at $z \sim$ 2-3 have a cold gas fraction within r_e ranging from 6.5% to 71.0%. *(2) AGN activity:* Using a variety of techniques (based on X-ray, IR, and optical data) we classify 49/166 ($\sim$ 30%) of massive galaxies at $z = 1 - 3$ as AGN hosts. The AGN fraction rises from $\sim$ 18% at $z = 1-1.5$ to $\sim$ 40% at $z = 2-3$. *(3) Structure:* After fitting the 2D rest-frame-optical light traced by the NIC3/F160W images at $z = 1-3$, we find that the Sérsic index n and effective radius r_e of massive galaxies at $z = 2-3$ are strikingly offset toward lower values compared to $z \sim 0$ galaxies. Extensive fitting tests and artificial redshifting experiments show this offset is real and not driven by systematic effects . This offset translates into two interesting populations among massive galaxies at $z = 2 - 3$: ultra-compact galaxies and a large population of low $n \leq 2$ systems. *(4) Ultra-compact ($r_e \leq 2$ kpc) galaxies:* These make up $\sim$40% of the massive galaxies at $z = 2 - 3$, but less than 1% of massive galaxies at $z \sim 0$. The massive ultra-compact systems at $z = 2 - 3$ appear to have no analogs among $z \sim 0$ massive galaxies in terms of their size, structure, and optical surface brightness. The extreme rest-frame B-band surface brightness of the massive ultra-compact systems at $z = 2 - 3$ suggests that these galaxies formed through gas-rich wet major mergers, or other highly dissipative events associated with powerful starbursts. Compared to the whole sample of massive galaxies, the ultra-compact objects at $z = 2 - 3$ show a deficiency by a factor of ~ 3 in star-forming systems, and also have a lower average SFR_{IR}. We suggest that the radical transformation of massive ultra-compact objects at $z = 2-3$ into massive galaxies at $z \sim 0$ (e.g., E, S0s) can happen through dry major mergers, minor mergers, and cold mode accretion, with dry major mergers playing a particularly important role. *(5) A large population of disky (low $n \leq 2$) systems:* We find that systems with low $n \leq 2$ make up the majority ($\sim$ 60%) of massive galaxies at $z = 2 - 3$, compared to only $\sim$ 20% of massive galaxies at $z \sim 0$. SFR and structure are correlated: the majority of galaxies with high SFR_{IR} at $z = 1 - 3$ have $n \leq 2$ and are extended ($r_e > 2$ kpc). We present arguments suggesting that the massive galaxies with $n \leq 2$, particularly the extended ($r_e > 2$ kpc) ones, likely host a significant disk component (e.g., an outer disk or central disky pseudobulge). The large fraction of massive galaxies at $z = 2 - 3$ with low $n \leq 2$ disk components, and high SFR underscores the importance of cold mode accretion, in addition to gas-rich major mergers, in building such a population.

Subject headings: galaxies: bulges — galaxies: evolution — galaxies: formation — galaxies: fundamental parameters — galaxies: interactions — galaxies: structure

[1] Department of Astronomy, University of Texas at Austin, Austin, TX
[2] University of Nottingham, School of Physics & Astronomy, Nottingham NG7 2RD, U.K.
[3] George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, Department of Physics and Astronomy, Texas A&M University, 4242 TAMU, College Station, TX 77843
[4] U.S. Planck Data Center, MS220-6 Caltech, Pasadena, CA 91125
[5] Gemini Observatory, Hilo, HI 96720, USA
[6] Department of Physics and Astronomy, University of California, Riverside, CA 92521
[7] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD, 21218
[8] NOAO, 950 N. Cherry Avenue, Tucson, AZ 85719, USA
[9] Australian Astronomical Observatory, P.O. Box 296, Epping, NSW 1710, Australia

1. INTRODUCTION

Studies of high-redshift galaxies are essential for testing and constraining models of galaxy formation. Conventional wisdom suggests galaxies are assembled and shaped by a combination of mergers, smooth accretion, and internal secular evolution. Galaxies form inside cold dark matter halos that grow hierarchically through mergers with other halos and gas accretion (Cole et al. 2000; Steinmetz & Navarro 2002), while internal secular evolution (Kormendy & Kennicutt, 2004) redistributes accreted material. Within the paradigm of hierarchical assembly, a number of issues remain. It is not known when and how the main baryonic components of modern galaxies (bulges, disks, and bars) formed, but global stellar mass density rose substantially between $z \sim 1-3$, reaching $50-75\%$ of its present value by $z \sim 1$ (Dickinson et al. 2003b).

It is also not clear how high-redshift galaxies evolve into present-day galaxies. Mergers of various types (major, minor, gas-rich, and dry) have been proposed to play a role in mass and structural assembly. There is mounting evidence that cold mode accretion (Dekel et al. 2009; Keres et al. 2009; Brooks et al. 2009) is important for building star-forming galaxies. In cold mode accretion, gas from the intergalactic medium is channeled along cold filaments directly into star-forming regions without encountering a virial shock (e.g., Birnboim & Dekel 2003; Keres et al. 2005; Ocvirck, Pichon & Teyssier 2008; Brooks et al. 2009; Dekel et al. 2009; Keres et al. 2009). This process is particularly effective in galaxies with halos $\leq 10^{12}$ $M_{\odot}$ such that cold-mode accretion dominates the global growth of galaxies at high redshifts and the growth of lower mass objects at late times. Star formation depends on complex baryonic physics and processes like gas dissipation and feedback from supernovae or a central engine.

High-redshift galaxies are different from local galaxies. Within the framework of hierarchical assembly, early, high-redshift galaxies are expected to be smaller, at a given mass, than their present-day counterparts. The size difference is predicted to be a factor of a few at $z = 2-3$ (Loeb & Peebles 2003; Robertson et al. 2006; Khochfar & Silk 2006; Naab et al. 2007). Several recent studies using rest-frame optical data provide evidence for size evolution among massive galaxies (Daddi et al. 2005; Trujillo et al. 2006, 2007; Zirm et al. 2007; Toft et al. 2007; Longhetti et al. 2007; Cimatti et al. 2008; Buitrago et al. 2008; van Dokkum et al. 2008, 2010). Aside from size evolution, there is some evidence that the nature of the red sequence changes at higher redshift. At $z \lesssim 1$, the red sequence primarily consists of old, passively evolving galaxies (Bell et al. 2004). Among extremely red galaxies (EROs) at $z = 1-2$, less than 40% are morphologically early types (Yan & Thompson 2003; Moustakas et al. 2004). It is well known that star formation rates were more intense at higher redshift (Daddi et al. 2007; Drory & Alvarez 2008), and a link has been found between star formation, size, and morphology at $z \sim 2.5$. Toft et al. (2007) and Zirm et al. (2007) find from NICMOS rest-frame optical imaging that blue star-forming galaxies are significantly more extended than red quiescent galaxies. Additionally, examples of rapidly star-forming galaxies (SFR $\sim 50-200$ $M_{\odot}$ yr^{-1}) at $z \sim 2-3$ with rotating disk kinematics have been found by the SINS survey (Förster Schreiber et al. 2009; Genzel et al. 2008; Shapiro et al. 2008).

Progress on understanding the evolution of massive galaxies at high redshift has been hindered by significant observational challenges. The deep optical surveys carried out by *HST ACS*, such as the Hubble Ultra Deep Field (HUDF, Beckwith et al. 2006) and the Great Observatories Origins Deep Survey (GOODS, Giavalisco et al. 2004), trace rest-frame optical galaxy morphology only out to $z \sim 1$. At higher z, bandpass shifting effects cause filters to trace progressively bluer bands, and optical filters trace rest-frame UV at $z \gtrsim 2$. UV light traces massive young stars, but manages to set few constraints about the overall mass distribution, making it difficult to probe the structure and mass of galaxy components at early epochs. Without high-resolution, deep, rest-frame optical imaging, it is not possible to compare structural parameters in galaxies across redshift in a morphologically unbiased way. NIR imaging is required to probe the rest-frame optical at $z \sim 1-3$. High-quality, deep NIR imaging with *HST* can fill in the NIR gap of spectral energy distributions (SEDs) and sample rest-frame optical structures of galaxies out to $z \sim 3$. Unfortunately, deep NIR imaging with *HST* has been completed for a limited number of galaxies over relatively small fields and small volumes at $z > 1$, with most pointings being within the Hubble Deep Fields and the Hubble Ultra Deep Field due to the inefficiency of NICMOS in covering large areas (e.g., Dickinson et al. 2004; Thompson et al. 2005; Zirm et al. 2007; van Dokkum et al. 2008).

A large area, high-resolution, deep NIR survey would be bountiful for galaxy formation studies. The GOODS-NICMOS Survey (GNS; Conselice et al. 2010), covering 44 arcmin^2 of the GOODS fields with NIC3, is a strong first effort. The GOODS-North and GOODS-South are among the best-studied regions in the sky and are a natural choice for such a survey. The GOODS fields already have deep data from *HST ACS* (Giavalisco et al. 2004), *Spitzer* IRAC/MIPS (Dickinson et al. 2003a), and Chandra (Giacconi et al. 2002; Alexander et al. 2003; Lehmer et al. 2005; Luo et al. 2008), among others. GNS consists of 60 pointings centered on massive ($M_{\star} > 10^{11}$ $M_{\odot}$) galaxies at $z > 2$, observed to a depth of $H = 26.8$ magnitudes. The value of GNS lies in the fact that the target fields were optimally selected to include massive galaxies selected by multiple methods in order to create an unbiased sample (see Conselice et al. 2010). The targets include Distant Red Galaxies (DRGs, Papovich et al. 2006), Extremely Red Objects (EROs, Yan et al. 2004), and BzK-selected galaxies (Daddi et al. 2007). There are additional massive galaxies in each field beyond the 60 main targets, so that there are 73 $M_{\star} \geq 10^{11}$ $M_{\odot}$ galaxies at $z = 1-3$ across all pointings. *Thus, the GNS data contain one of the largest samples of very massive galaxies at high redshift with rest-frame optical imaging and robustly probe massive galaxies when the Universe was less than 1/3 of its current age, during the epoch of bulge and disk formation.*

The goal of this work is to investigate the evolution of massive galaxies over $z = 1-3$ with one of the largest samples at high-redshift with rest-frame optical imaging. We take advantage of the existing rich ancillary data to derive star formation rates (SFR) from 24 μm detections and look for AGN activity based on X-ray detections and mid-IR

SEDs. We correlate rest-frame optical structural parameters with SFR to gain insight into how massive galaxies are expected to evolve.

The plan of this paper is as follows. We discuss the data and sample properties in § 2. In §3 we describe the measurement of structural parameters, and in §3.3 we make a detailed comparison with $z \sim 0$ galaxies of similar stellar mass. A detailed artificial redshifting experiment is conducted in §3.4.4 to explore the impact of instrumental and redshift-dependent effects on structural parameters. In §4, we measure star formation properties based on *Spitzer* MIPS 24 μm detections and discuss star formation history. In §4.3 we consider the relationship between structural properties, stellar mass, and star formation. Estimates of the mass and fraction of cold gas in massive star-forming galaxies at $z = 2 - 3$ are presented in § 5. In §6, we use a variety of techniques (X-ray properties, IR power-law, and IR-to-optical excess) to identify AGN among the massive galaxies at $z = 1 - 3$. Finally, in §7 and §8, we discuss and summarize our results. All calculations assume a flat ΛCDM cosmology with $\Omega_\Lambda = 0.7$ and $H_0 = 70$ km s^{-1} Mpc^{-1}.

2. DATA AND SAMPLE

2.1. *Observations and Pointing Selections for GNS*

Our data comes from the GOODS-NICMOS Survey (GNS; Conselice et al. 2010). GNS is a deep, 180-orbit survey with the *HST* NICMOS-3 camera in the F160W (H) band that probes optical light from galaxies between $z \sim 1 - 3$. The coverage extends over both *ACS* GOODS fields and is divided between 60 pointings centered on massive $M_\star \geq 10^{11} M_\odot$ galaxies at $z > 2$. Each pointing cover51.″2 x 51.″2 and was observed to a depth of three orbits in nine exposures of ~ 900 seconds (~ 135 minutes per pointing). A total of ~ 8300 sources were detected across an effective area of ~ 44 arcmin^2. The 5σ limiting magnitude for an extended source with a 0.″7 diameter is H=26.8 (AB). The NIC-3 images were drizzled with a pixfrac of 0.7 and output platescale of 0.″1. The NIC3 camera is currently out of focus,and after detailed investigation (see §3.1), we find the Point Spread Function (PSF) spans a Full Width Half Maximum (FWHM) of 0.″26 – 0.″36 with a mean value of 0.″3.

The 60 GNS pointings were planned by identifying massive galaxies having a photometric redshift $1.5 < z < 2.9$ and stellar mass $M_\star > 10^{11}$ $M_\odot$ via three color selection criteria. The target galaxies include Distant Red Galaxies (DRGs, Papovich et al. 2006), Extremely Red Objects (EROs, Yan et al. 2004), and BzK-selected galaxies (Daddi et al. 2004). All of these methods are designed to select red dusty or red passively evolving galaxies. DRGs have evolved stellar populations that are identified with $J - K > 2.3$ (Vega mag). EROs are selected based on *Spitzer* and NIR data via $f_\nu(3.6\mu m)/f_\nu(z850) > 20$. This selection is sensitive to red populations that are either old or reddened, so EROs contain a mixture of young and old stellar populations. BzK galaxies are selected based on the quantity $BzK \equiv (z - K)_{AB} - (B - z)_{AB}$. Galaxies with $BzK > -0.2$ at $z > 1.4$ are identified as star-forming galaxies. Redder and possibly more evolved galaxies are identified with $BzK < -0.2$ and $(z - K)_{AB} > 2.5$. The final pointings were designed to include at least one red massive galaxy and to also maximize the total number of additional galaxies (e.g., Lyman-break galaxies and submm galaxies) within each pointing.

2.2. *Our Sample of Massive Galaxies at z=1-3*

The sample of massive galaxies that we work with in this paper is not limited to the original color-selected massive galaxies at $z > 1.5$ defining the original 60 GNS pointings. Instead, our sample of massive galaxies at $z = 1 - 3$ is derived from the set of all galaxies which are mapped with NIC3/F160W across the 60 fields, and for which a reliable stellar mass and photometric redshifts were estimated by Conselice et al. (2010), based on SED fits to the NIC3/F160W and optical imaging. A detailed description of how these quantities were estimated is in Conselice et al. (2010), and we only briefly summarize the methodology here, and describe how we selected our our final sample of massive galaxies at $z = 1 - 3$ in GNS.

The source extraction catalog for the NICMOS images across the 60 pointings of the GNS survey contains ~ 8300 sources with $H < 28$ and $V < 30$. For those galaxies detected in the *ACS* $BViz$ and NICMOS H bands, we use the available photometric redshifts and stellar masses from (Conselice et al. 2010). Photometric redshifts were determined by fitting template spectra to the $BVizH$ data. Stellar masses were measured by fitting the $BVizH$ magnitudes to a grid of SEDs generated from Bruzual & Charlot (2003) stellar population synthesis models, assuming a Salpeter IMF[10]. The grid includes different colors, ages of stellar populations, metallicities, dust content, and star formation histories as characterized by exponentially declining models. The typical uncertainty in stellar mass is a factor of ~ 2 across the sample.

In order to account for additional massive ($M_\star \geq 10^{10}$ $M_\odot$) red systems, which are undetected in the GOODS *ACS* BV and therefore do not have viable stellar masses from the above techniques, we use available masses and redshifts (Buitrago et al. 2008; Bluck et al. 2009) based on deep ground-based $RIJHK$ data along with ACS iz data, where available. Photometric redshifts are determined with a mixture of techniques (e.g., neural networks and Bayesian techniques) described more fully in Conselice et al. (2007). Stellar masses were measured from these data to uncertainties of a factor of ~ 2 with the multicolor stellar population fitting techniques from Conselice et al. (2007, 2008). As with the larger sample described above, a stellar mass is produced by fitting model SEDs to the observed SED for each galaxy. A Salpeter IMF is assumed, and the SED grids are constructed from Bruzual & Charlot (2003) stellar population synthesis models.

From the sample of galaxies with photometric redshifts and stellar masses determined as described above, we define the sample of massive galaxies used in this paper. We restrict our analysis to the redshift interval $z = 1 - 3$ over which our NIC3/F160W images probe the rest-frame optical light in order to minimize the effect of bandpass shifting. This ensures that we measure all structural parameters in the rest-frame optical across the entire redshift

[10] In §4.2 we use a Chabrier IMF for SFR estimates. Using a Chabrier IMF rather than a Salpeter IMF in estimating the stellar mass would lower the values by a factor of 0.25 dex or less.

range. Although the mass functions calculated for GNS by Mortlock et al. (submitted) show that the mass completeness limit is $\sim 3 \times 10^9$ $M_\odot$ at $z \sim 3$, we apply a higher mass cut of 5×10^{10} $M_\odot$ as our interest is specifically with the most massive galaxies.

Our final sample consists of the 166 (73) massive galaxies with $M_\star \geq 5 \times 10^{10}$ $M_\odot$ ($M_\star \geq 1 \times 10^{11}$ $M_\odot$) and $z = 1-3$. This is the largest HST-based dataset with rest-frame optical imaging of massive galaxies over $z = 1-3$. The $M_\star \geq 10^{11}$ $M_\odot$ galaxies from Buitrago et al. (2008) are a subset of our sample. The other previous HST NICMOS studies (e.g., Toft et al. 2007; Zirm et al. 2007; van Dokkum et al. 2008) each contain, at most, $10-20$ $M_\star \geq 10^{11}$ $M_\odot$ systems. The full distributions of apparent H and V magnitude, stellar mass, and redshift for this sample are shown in Figure 1. The bottom left panel of Figure 1 shows the mass function of our sample.

Spectroscopic redshifts are available for 44/166 ($26.5 \pm 3.4\%$) of the sample. For the remaining galaxies photometric redshifts are used. Based on the galaxies with spectroscopic redshifts, the median photometric redshift error is $\delta z/(1+z) = 0.071$ (Grützbauch et al. 2010). However, the galaxies with spectroscopic redshifts all have $V \leq 27$ with a median of 24.3, while there is a fair number of fainter ($V > 27$) sample galaxies with only photometric redshifts. For these fainter cases, the uncertainties in photometric redshifts will likely be larger than the afore-cited value.

2.3. *Properties and Selection Biases in the Sample*

We further discuss here additional properties and potential biases in our final sample , which consists of 166 (73) massive galaxies with $M_\star \geq 5 \times 10^{10}$ $M_\odot$ ($M_\star \geq 1 \times 10^{11}$ $M_\odot$) at $z = 1-3$.

We estimate the number density over $z = 2-3$ to be $\sim 5 \times 10^{-4}$ Mpc^{-3} (see Conselice et al. 2010 for a detailed discussion of the number density of massive galaxies in the GNS sample). The corresponding stellar mass density is $\sim 6 \times 10^7$ $M_\odot$ Mpc^{-3}. The massive GNS galaxies are collectively 10-100 times more abundant than SMGs, which have space densities of $10^{-5} - 10^{-6}$ Mpc^{-3} at $z \sim 2-3$ (Blain et al. 2002). Rather, the number density is in agreement with published values (Daddi et al. 2005; 2007) for other passively evolving and star-forming galaxies at $z \sim 2$.

How does our sample break down in terms of the typical color-selection methods, which are usually used to identify massive high redshift galaxies? About 63% (104/166) of this final sample is listed in existing catalogs for DRG (Papovich et al. 2006), BzK (Daddi et al. 2004), or ERO (Yan et al. 2004) galaxy populations. There are 8, 9, and 43 sources that are uniquely listed in one of the DRG, Bzk, or ERO galaxy catalogs, respectively. An additional 44 sources are listed in two or more of these catalogs. About 37% (62/166) sources were not previously identified as DRG, ERO, or BzK galaxies.

What are the selection biases impacting our sample? General biases in the selection of massive galaxies in the GNS survey have been discussed in Conselice et al. (2010), and we only discuss the points relevant for our sample here. The original 3 color methods use to identify the 60 high redshift massive galaxies, which are used to select the 60 GNS pointings, preferentially pick red massive galaxies, where the red color is indicative of red dusty systems or red evolved stellar populations. These color methods would miss blue galaxies, which could be a mix of lower redshift galaxies, but also potentially massive high redshift blue galaxies, although Conselice et al. (2010) argues that the latter are rare.

However, it is important to note here that our working sample of massive galaxies at $z = 1-3$ is not limited to the original color-selected preferentially red massive galaxies, and therefore do not suffer from the same bias. Since our sample of massive galaxies at $z = 1-3$ is derived from all galaxies that are mapped with NIC3/F160W and have a reliable stellar mass and photometric redshifts, they will only include galaxies bright enough to be detected in NIC3/F160W (which corresponds to rest-frame 5330 $\AA$ to 4000 $\AA$ at $z = 2-3$), and whose observed SED enables a reliable stellar mass and photometric redshift to be estimated. For instance, our sample will likely miss ultra-dusty galaxies whose photometric redshift would not be reliably measured from the afore described methods (Conselice 2010).

In terms of blue galaxies, our working sample is not as strongly biased against blue galaxies as samples of massive galaxies defined by using the three color methods for DRG, EROs, and BzK. Conselice et al. (2010) in fact shows that many galaxies with blue ($z-H$) colors, which would have been undetected by these color selections do get included in this final sample of massive galaxies for the GNS survey. Nearly all known Lyman Break Galaxies or BX/BM objects (Reddy et al. 2008) at $z = 2-3$ in the GNS fields are detected by the GNS NIC3/F160W imaging (Conselice et al. 2010). Furthermore, the galaxies detected by GNS at $z = 1-2$ or $z = 2-3$ include both intrinsically blue and red systems: the rest-frame $U-V$ color ranges from about -0.4 – 2.1 for galaxies in the stellar mass range $M_\star \sim 10^9$ – 10^{12} $M_\odot$ (Figure 2) The blue systems are preferentially at low masses, while GNS galaxies with $M_\star \geq 1 \times 10^{11}$ $M_\odot$ are preferentially red. As we will show in §4, many of these red systems are dusty systems with high levels of obscured star formation. If dusty starbursts blow out their gas and dust and go through a phase of blue colors, then one might have expected to see some blue massive galaxies. The fact that we do not see such systems in Figure 2, may be due to the deep potential of the massive galaxies or/and the short-lived nature of such a phase.

3. STRUCTURAL PROPERTIES OF MASSIVE GALAXIES

3.1. *PSF Modeling*

Knowledge of the PSF is important to assess data quality and for deriving structural parameters. NIC3 is out of focus, so the PSF can deviate from the theoretically expected one. PSF convolution with GALFIT is commonly performed with a user-provided bright, unsaturated star. Not all of the GNS tiles contain suitably bright, unsaturated stars. It is not advisable to adopt a set of PSF stars from a subset of pointings because the NIC3 PSF: 1) depends on position within the NIC3 field, and 2) is also subject to interpolation artifacts introduced by `drizzle` that are dependent on the adopted dither pattern (John Krist, private communication).

As a result, the best available option for handling PSF convolution is to make synthetic NIC3 PSFs with Tiny

Tim (Krist 1995). For each galaxy, Tiny Tim PSFs were generated for all the galaxy's positions in the individual, undrizzled exposures. Telescope breathing was accounted for with each PSF by refining the Tiny Tim parameters to match the Pupil Alignment Mechanism (PAM) value recorded in the headers of the undrizzled frames. Blank, zero-valued frames retaining the WCS information of the undrizzled frames were made. The synthetic PSFs were inserted into the blank frames precisely where each galaxy would be in the individual frames. The blank frames were drizzled together in the same way as the data with a pixfrac of 0.7 and a final output platescale of 0.″1/pixel. This process was repeated for all 166 $M_\star \geq 5\times10^{10}$ galaxies in our sample.

This approach accounts both for variation in PSF with position on the NIC3 field and for the dependence on the `drizzle` algorithm. The range of FWHM in the final drizzled synthetic PSFs is $\sim 0.''26 - 0.''36$, with a mean value of 0.″3. The range in PSF FWHM comes from differing positions in the NIC3 field and the PAM values used to create the synthetic PSFs. The mean PSF diameter of the science images (0.″3) is 2.5 kpc at $z = 2$, under the adopted cosmology.

3.2. *Structural Decomposition*

We characterize the massive GNS galaxies with structural decomposition. Ideally, one would like to fit multiple components (bulge, disk, bar, nuclear point source, etc.) in the decomposition, but the 0.″3 diameter (or full width half-maximum, FWHM) of the PSF (corresponding to ~ 2.4 kpc at $z = 1-3$) prevents such detailed decompositions[11]. Instead, we choose to fit only single Sérsic (1968) $r^{1/n}$ profiles to the 2D light distributions. It is clear that a single Sérsic profile is not a complete indicator of overall galaxy structure. For instance, in detailed images of nearby galaxies, the best-fit index n for a single Sérsic profile does not always correlate with the bulge Sérsic index obtained with 2D bulge-disk or bulge-disk-bar decomposition (Weinzirl et al. 2009). However, the single Sérsic index n is on average a good way to separate disk-dominated galaxies from the class of luminous spheroidal and bulged-dominated galaxies (see §3.4.4), and in studies of high-redshift galaxies the criterion $n \lesssim 2$ is often used to separate spirals or disk galaxies from ellipticals (e.g., Ravindranath et al. 2004; Bell et al. 2004; Jogee et al. 2004; Barden et al. 2005; Trujillo et al. 2007; Buitrago et al. 2008). Thus, we fit only single Sérsic (1968) $r^{1/n}$ profiles to the 2D light distributions to measure a Sérsic index (n) and half-light radius (r_e), while taking into account the PSF. Details of the fitting procedure are given below.

The original science frames were sky subtracted and converted to units of counts. Thumbnail images for each galaxy were extracted from the science frames. Each galaxy was centered in a thumbnail image of width and height corresponding to 120 kpc measured at the photometric redshift of the object.

The NIC3/F160W images of the 166 sample galaxies were fit with a single Sérsic component using GALFIT (Peng et al. 2002). In each image, objects that were near, but not blended with, the primary source were masked out. For the fraction ($\sim 15\%$) of the primary galaxies that were blended or overlapping with another galaxy identified in the source extraction catalog, the blended sources were each fitted simultaneously with a separate Sérsic profile. Some fraction of primary galaxies appeared visibly distorted ($\sim 8\%$, see Figure 3 and §2.2), but these were fitted with only a single Sérsic profile as they only counted as a single galaxy in the source extraction catalog.

GALFIT implements a Levenberg-Marquardt algorithm to minimize the χ^2 between the data and a PSF-convolved model. By convolving model profiles with a PSF, GALFIT can measure sizes smaller than the PSF. Of the 166 sample galaxies, $\sim 10\%$ have effective diameters below the FWHM (0.″3, corresponding to 2.5 kpc at $z \sim 2$) of the NIC3/F160W PSF and are formally unresolved. These constitute a small fraction of the sample and do not drive the main results in §3.3.

Another consideration is the effect of potential AGN on the structural fits. When fitting high resolution images of nearby galaxies, it is well known that fitting a galaxy that hosts a point source with a single Sérsic component will lead to an artificially high Sérsic index n (typically $n > 4$; e.g., Weinzirl et al. 2009; Pierce et al. 2010). If a point source is added to the Sérsic model, the index n of the Sérsic component falls to more reasonable values. In the case of the massive GNS galaxies at $z \sim 1$–3, we expect that the low resolution (0.″3, corresponding to 2.5 kpc at $z \sim 2$) of the NIC3/F160W images will reduce the effect of potential point sources on the structural decomposition. However, for completeness, we have fitted all the galaxies at $z \sim 1$–3 in which a potential AGN was identified via a variety of techniques (§6) with both a Sérsic component and a point source. The fractional luminosity of the point source components, or PSF/Total light ratio, ranges from 1-46%, with a median of 10%. Fig. 4 compares the (n, r_e) parameters of the Sérsic component with and without the inclusion of a point source. It shows, as expected, that including the point source produces generally small changes in (n, r_e) and goes in the direction of lowering n and enlarging r_e.

If a point source is included in the single Sérsic models, $\sim 27\%$ of the sources that had $n > 2$ in the single Sérsic fit have their n lowered to $n \leq 2$, and this will increase our fraction of massive galaxies with $n \leq 2$ from $\sim 65\%$ to $\sim 73\%$, thereby further *strengthening* our finding of a large population of low $n \leq 2$ massive galaxies. In terms of sizes, adding the point source raises r_e by a median of 10%. Only $\sim 8\%$ of sources with $r_e \leq 2$ kpc from the single Sérsic fit have $r_e > 2$ kpc with the point source. In the rest of the paper, we therefore choose to use the structural parameters for a single Sérsic component fit. The tests above clearly show that our results of a large fraction of systems with low n are not biased by the presence of AGN, since the latter would only decrease n.

3.3. *Derived Structural Properties at z ∼ 2-3*

The results of the structural fits to the NIC3/F160W images of the 166 sample galaxies are shown in Fig. 5 and Table 1. Figure 3 shows examples of massive ($M_\star \geq 5\times10^{10}$

[11] For the more extended galaxies multiple components (e.g., bulge and disk) decomposition was attempted with limited success and this is discussed in §7.1.

$M_\odot$) galaxies at $z \sim 2-3$ with different ranges of Sérsic index n and effective radius r_e. The majority (over 84%; Table 1) of the massive GNS galaxies have $r_e < 4$ kpc. In such systems, structural features are generally hard to discern due to resolution effects, so that systems appear fairly featureless (top 4 rows of Fig. 5). In the small fraction of massive galaxies at $z \sim 2-3$ with large $r_e \geq 4$ kpc, one can discern some structural features such as an elongated bar-like feature or a combination of a central condensation surrounded by a more extended lower surface brightness component, reminiscent of a bulge and disk (5th row). Row 6 contains irregular systems that appear disturbed. The fraction of these systems at $z = 2-3$ is small, only $\sim 8\%$.

One can gain useful insight by considering a combination of properties such as Sérsic index n, half light radius r_e, and surface brightness in the rest-frame optical band. For instance, the Sérsic index n characterizes both the steepness and wings of the surface brightness profile, and tends to be lower for a disky component formed in dissipative gas-rich events than for systems formed in dry major mergers or major mergers of moderate gas fraction. We discuss this point in greater detail in § 7. The surface brightness is also useful as a guide to the mass surface density, assuming a constant mass to light ratio. Very high values are typical of products of highly dissipative events.

The lower two rows of Fig. 5 shows the rest-frame optical Sérsic index n and effective radius r_e for the samples of massive galaxies ($M_\star \geq 5\times10^{10}$ $M_\odot$ and $M_\star \geq 10^{11}$ $M_\odot$), split in two separate redshift bins $z = 1-2$ and $z = 2-3$. The top row of Figure 5 also shows for comparison the rest-frame optical structural parameters for $z \sim 0$ galaxies of similar stellar mass taken from Allen et al. (2006), who performed a single component Sérsic fit to B-band images of galaxies in the Millennium Galaxy Catalogue (MGC), a large ground-based imaging and spectroscopic survey over 37.5 deg^2 (Liske et al. 2003; Driver et al. 2005).

It is strikingly clear from Fig. 5 that *the massive galaxies at $z \sim 2-3$ are strikingly offset toward lower (n, r_e) compared to the massive ~ 0 galaxies.* . This is further illustrated in quantitative terms in Table 1 and in Figure 6. Specifically, we find that *the majority ($64.9\% \pm 5.4\%$ for $M_\star \geq 5\times10^{10}$ $M_\odot$, and $59.4\%\pm8.7\%$ for $M_\star \geq 10^{11}$ $M_\odot$) of massive galaxies at $z = 2-3$ have low $n \leq 2$, while the fraction at $z \sim 0$ is three times lower.* (Table 1 and Figure 6).

In terms of sizes, the full sample of massive GNS galaxies at$z \sim 2-3$ also has lower r_e compared to massive galaxies at $z \sim 0$: *$39.0 \pm 5.6\%$ for $M_\star \geq 5\times10^{10}$ $M_\odot$ and $40.6 \pm 9.0\%$ for $M_\star \geq 1\times10^{11}$ $M_\odot$ have $r_e \leq 2$ kpc compared to less than a few percent at $z \sim 0$ (Table 1 and Figure 6).* How do these results compare with earlier studies? While many of the earlier studies focused on small samples, this work is a step forward because of the improved number statistics that comes with an unbiased sample of massive galaxies selected by multiple methods that is complete in mass. The observed *apparent* size evolution in our data generally agrees with results reported in other studies of massive galaxies (e.g., Daddi et al. 2005; Trujillo et al. 2007; Zirm et al. 2007; Toft et al. 2007; Buitrago et al. 2008; van Dokkum et al. 2008; 2010). The ratio in r_e of high-redshift galaxies with respect to $z \sim 0$ galaxies, or $r_e/r_{e,z\sim0}$, can be modeled as a power law in redshift of the form $\alpha(1+z)^\beta$, where α and β are constants. Using the $z \sim 0$ massive ($M_\star \geq 5\times10^{10}$ $M_\odot$) MGC galaxies as the normalization, we find for all 166 massive GNS galaxies that α and β are 1.09 and -1.45, respectively. This is comparable to the findings of Buitrago et al. (2008), who show for $M_\star \geq 10^{11}$ $M_\odot$ galaxies at $1.7 \leq z \leq 3$ that β varies from -0.8 for spheroidal galaxies to ~ -1.5 for disk-like galaxies.

However, in order to determine whether this apparent size evolution is real, one needs to address a number of systematic effects, as outlined in the next section.

It is more interesting to look at the combined properties such as Sérsic index n, half light radius r_e, and surface brightness in the rest-frame optical band. Figure 7 compares the effective radius r_e and optical surface brightness of the $z \sim$ 0 MGC galaxies to those of the the massive GNS galaxies at $z = 1-3$. Surface brightness was measured with B-band at $z \sim 0$ and in the NIC3/F160W band at $z = 2-3$ so that approximately the same rest-frame light is considered at both redshifts. For all galaxies, surface brightness is defined as $\mu_e = \mathrm{mag} + \mathrm{zeropoint} + 2.5\log_{10}(\pi r_e^2) - 10\log_{10}(1+z)$, with the last term being a surface brightness dimming correction.

Figure 7 and Table 1 show that a large fraction ($39.0 \pm 5.6\%$ for $M_\star \geq 5\times10^{10}$ $M_\odot$ and $40.6 \pm 9.0\%$ for $M_\star \geq 1\times10^{11}$ $M_\odot$) of galaxies at $z = 2-3$ are in the form of ultra-compact ($r_e \leq 2$ kpc) galaxies, whose small sizes are unmatched among local massive galaxies, and whose surface brightness in the rest-frame optical can be 4-6 magnitudes brighter. Nearly half ($46.7 \pm 9.11\%$) of these ultra-compact systems at $z = 2-3$ have a Sérsic index $n \leq 2$ (Table 2). We henceforth refer to this group of objects as Class 1 or ultra-compact galaxies. The high surface brightness of the massive ultra-compact galaxies at $z = 2-3$ is suggestive of highly dissipative events playing an important role in their formation. In addition, the large difference they exhibit with respect to local massive galaxies in terms of surface brightness, $r-e$, and n implies that strong structural transformation must take place between $z = 2$ to $z = 0$ to convert them into their present-day descendants.

The remaining massive galaxies (61% for $M_\star \geq 5\times10^{10}$ $M_\odot$ and 59% for $M_\star \geq 1\times10^{11}$ $M_\odot$) at $z = 2-3$, have more extended r_e (2 to 8 kpc). We henceforth refer to these objects as Class 2. They overlap with the the lower range in r_e of local massive galaxies, but appear not to extend out to 16 kpc (Figure 7). Furthermore, at a given r_e, the surface brightness is 2.5 magnitudes higher. The vast majority ($76.6 \pm 6.18\%$) of these Class 2 objects have a Sérsic index $n < 2$ (Table 2). The Class 2 objects are not as radically offset from local massive galaxies as are the Class 1 ultra-compact objects.

We will discuss the formation and evolution of the Class 1 and 2 systems in more detail in § 7. Before doing so, we need to first investigate whether some of the observed differences between the massive galaxies at $z = 2-3$ (Class 1 and 2) and at $z \sim 0$ are due to redshift-dependent systematic effects.

3.4. *Impact of Systematic Effects on Structural Properties*

In the previous section we found that the massive galaxies at $z \sim 2-3$ are strikingly offset toward lower (n, r_e) compared to the massive ~ 0 galaxies. In particular, the fraction of massive galaxies at $z = 2-3$ with low $n \leq 2$ is three times higher than at $z \sim 0$. Furthermore, a large fraction (39.0±5.6% for $M_\star \geq 5\times10^{10}$ $M_\odot$ and 40.6±9.0% for $M_\star \geq 1\times10^{11}$ $M_\odot$) of the massive galaxies at $z = 2-3$ are in the form of ultra-compact ($r_e \leq 2$ kpc) galaxies, compared to less than a few percent at $z \sim 0$ (Table 1 and Figure 6). Nearly half of these ultra-compact systems at $z = 2-3$ have a Sérsic index $n < 2$. Among the massive galaxies which are not ultra-compact, most (76.6 ±6.18%) systems have a Sérsic index $n < 2$ (Table 2).

It is relevant to ask whether the large fraction of low (n, r_e) systems we observe among massive galaxies at $z \sim$2–3, compared to massive galaxies at $z \sim 0$ is real or due to a number of systematic effects. We will address those listed below:

1. Is it possible that the distribution of (n, r_e) for massive galaxies at $z \sim 0$ and at $z = 2-3$ is intrinsically similar, but that some selection effects at $z = 2.5$ is making us preferentially detect the compact low n systems, thereby causing an artificial excess of the latter? We argue that this is very unlikely because even if we take all the massive compact low n systems at $z \sim 0$, and appropriately scale them for the difference in number density between $z \sim 0$ and $z = 2.5$, we still would fall way short of reproducing the observed number densities of compact low n systems. The number density of massive ($M_\star \geq 1\times10^{11}$ $M_\odot$) galaxies at $z = 2.5$ is approximately 30% that at $z \sim 0$. If we take the most compact ($r_e \leq 2$ kpc) and low $n \leq 2$ systems at $z \sim$ 0, and scale this number by 30%, we find a much lower number density (7.0×10^{-6} gal Mpc^{-3}), than the observed no density (8.3×10^{-5} gal Mpc^{-3}) at $z = 2.5$ for such compact systems. In addition, the fact that the rest-frame optical surface brightness is radically different at $z \sim 0$ and $z = 2.5$ (Figure 7 supports the idea that these compact low n systems at z=2.5 have no local analogs.

2. How robust are our fits to the NIC3/F610W images of the $z = 2-3$ galaxies? Could some of the galaxies with a best-fit Sérsic index $n \leq 2$ have similarly good fits with much higher n? We address this issue in § 3.4.1 and § 3.4.2.

3. Can the offset in (n, r_e) between the $z = 2-3$ galaxies and the $z \sim 0$ galaxies be caused by systematic differences between the fitting techniques applied by us to the NIC3/F610W images of $z = 2-3$ galaxies and the fitting techniques used by Allen et al. (2006) on the B-band images of the massive $z \sim 0$ galaxies in the MGC catalogue? We address this issue in § 3.4.3.

4. Can redshift-dependent systematic effects cause structural parameters, such as the high Sérsic index n of massive galaxies at $z \sim 0$, to 'degrade' into the regime of low $n \leq 2$ values, measured in the $z = 2-3$ systems. We address this issue in §3.4.4, where we also demonstrate that the massive galaxies with $n \leq 2$ at $z = 2-3$ are more similar in terms of bulge-to-total ratio to massive local late-type disky spirals than to local massive ellipticals or bulge-dominated spirals.

3.4.1. *Tests on Robustness of Fits and Parameter Coupling*

How robust are the results that a dominant fraction of the massive galaxies at $z = 2-3$ have a low $n \leq 2$ and that a large fraction are ultra-compact? In particular, how non-degenerate are the fits? Could some of the galaxies with an $n \leq 2$ have similarly good fits with higher n?

First, one should note that the errors quoted by GALFIT on the structural parameters cannot be used to assess the robustness of the fits because the errors quoted by GALFIT underestimate the true parameter errors (Häussler et al. 2007), which are dominated by the systematics of galaxy structure, and in particular, by parameter coupling and degeneracy.

The task of assessing the coupling between model parameters is complicated when models have a large number of free parameters. The single Sérsic profiles fit to the NICMOS galaxy images have 7 free parameters (centroid, luminosity, r_e, n, axis ratio, and position angle). While GALFIT selects a best-fit by minimizing χ^2 for a given set of input guesses, it is not clear whether the minimized χ^2 is an absolute minimum or local minimum. Investigating the χ^2 values for all combinations of fit parameters over the full multi-dimensional parameter space is prohibitively time consuming and computationally expensive. Instead, we will adopt a simpler approach of focusing on strong coupling between r_e and n, and exploring how χ^2 varies as these parameters are moved away from the initial solution picked by GALFIT.

One important point should be noted when using changes in χ^2, or $\Delta\chi^2$, for fits to different models. When errors are normally distributed, the multi-dimensional ellipsoids for a given $\Delta\chi^2$ contour can be associated with a statistical confidence level (e.g., $\Delta\chi^2 \sim 1$ corresponds to a 68% confidence level). However, since the errors in the GALFIT models are not normally distributed, but are instead dominated by the systematics of galaxy structure, this means that we cannot a priori assign a confidence level to a given $\Delta\chi^2$. As outlined in the test below, we can still use the shape of $\Delta\chi^2$ as a function of n or r_e as a guide to the quality of fit in the sense that sharp rises in χ^2 as n is varied away from the best-fit value are taken as indicative of poorer fits. But, we cannot a priori say how much poorer the fits are in a statistical sense. This is a well-known and hard problem in structural fitting. We will return to this point in section § 3.4.2.

We carry out the test below for all galaxies in our sample We denote as $\chi^2_{\rm min,0}$, the value of χ^2 obtained when GALFIT fits the galaxy with n and r_e as free parameters. The associated best-fit parameters are $n_{\rm min,0}$ and $r_{e,\rm min,0}$. We then fit single Sérsic profiles with n fixed at discrete values (0.5-10), while allowing all other parameters to freely vary. The initial inputs to these fits were the same as those used to generate the model in which n is a free parameter. We

let GALFIT find the best-fit for each of these fixed n models by minimizing χ^2, and we record for each such best-fit the following quantities: the fixed value of n, the best-fit value of r_e, and the associated minimum in χ^2 called $\chi^2_{\rm min}$. We then evaluate how the difference $\chi^2_{\rm min}$ - $\chi^2_{\rm min,0}$ varies as a function of r_e and n, as we move to values away from $n_{\rm min,0}$ and $r_{e,{\rm min,0}}$.

The test was carried out for all galaxies. Figure 8 shows the results of the test for the four representative galaxies with $n \sim 1-4$ displayed in Figure 9. The first column of Figure 8 shows how ($\chi^2_{\rm min}$-$\chi^2_{\rm min,0}$) changes when n is varied away from $n_{\rm min,0}$ at discrete values (0.5-10) and GALFIT is allowed to vary all other parameters to get a best-fit that yields $\chi^2_{\rm min}$. The second column shows the corresponding best-fit r_e for that $\chi^2_{\rm min}$. Red stars in the plots denote $n_{\rm min,0}$ and $r_{e,{\rm min,0}}$, which are associated with $\chi^2_{\rm min,0}$. The shape of $\chi^2_{\rm min}$ - $\chi^2_{\rm min,0}$ is asymmetric for n and r_e. The coupling between n and r_e means ($\chi^2_{\rm min}$ - $\chi^2_{\rm min,0}$) varies in a similar way with both n and r_e.

We can see that in Figure 8, the absolute minimum χ^2 values occur at the $n_{\rm min,0}$ and $r_{e,{\rm min,0}}$ values, which GALFIT picked when it was allowed to freely fit the galaxies without fixing n. Shifting n away from $n_{\rm min,0}$ (denoted by the red stars) by ±1 can increase $\chi^2_{\rm min}$ by several 10s or 100s of χ^2 units. While only 4 representative galaxies are shown in Figure 8, we show results for the whole sample in Figure 10. This figure illustrates that the distribution of ($\chi^2_{\rm min}$ - $\chi^2_{\rm min,0}$) for $n_{\rm min,0}-1$ (top panel) and $n_{\rm min,0}+1$ (bottom panel), and demonstrates for the whole sample, $\chi^2_{\rm min}$ generally changes substantially when n is shifted away from $n_{\rm min,0}$. We draw two primary conclusions from Figure 8:

1. For galaxies with $n_{\rm min,0} > 2$ (rows 3 and 4), $\chi^2_{\rm min}$ - $\chi^2_{\rm min,0}$ rises sharply at lower $n < n_{\rm min,0}$, suggesting that lower n values are unlikely to yield a good fit for such systems. At $n > n_{\rm min,0}$, $\chi^2_{\rm min}$ - $\chi^2_{\rm min,0}$ rises less sharply, but the rise is still substantial as demonstrated in by the high-magnification inset plots in rows 3 and 4 of column 1.

2. The most important point to take from Figure 8 is that for galaxies with $n_{\rm min,0} < 2$ (as in rows 1 and 2), $\chi^2_{\rm min}$ - $\chi^2_{\rm min,0}$ rises rapidly at higher $n > n_{\rm min,0}$, thereby making it unlikely that a higher $n > 2$ would provide a similarly good fit. *Thus, we have a great degree of confidence that we are not highly overestimating the number of $n \leq 2$ galaxies in the sample.*

3.4.2. *Recovery of Parameters From Simulated Images*

Section §3.4.1 tests how well parameters are recovered in real galaxies, but we cannot a priori assign a confidence level to a given $\Delta\chi^2$ because the errors in the GALFIT models are not normally distributed. However, we can run an extra complementary test where we use simulated *idealized* galaxies whose (n, r_e) are a priori known. The drawback of using idealized galaxies as opposed to the real galaxies fitted in §3.2 is that the former lack the complexity of real galaxies, since they are simply generated from GALFIT models and exactly described by a functional form, such as a Sersic model with a specified (n, r_e). However, the advantage is that we do know the (n, r_e) values a priori and can therefore compare these values to those obtained once these idealized galaxies are inserted into frames with noise properties corresponding to the NIC3 GNS images of our sample galaxies at $z = 1-3$.

This test is performed by simulating 1000 galaxy images, each with a unique set of Sérsic parameters: surface brightness at the effective radius μ_e, effective radius r_e, Sérsic index n, axis ratio b/a, and position angle PA. The parameters are chosen randomly from uniform distributions spanning the parameter space of the observed galaxies. The ranges in μ_e, r_e, n, b/a, and PA are 18 to 35 mag/arcsec^2, $0\farcs05$ to $1\farcs0$, 0.5 to 10, 0.3 to 1.0, and -90 to 90 degrees, respectively. The chosen range in input μ_e mimics the effect of surface brightness dimming, and the range in r_e ensures the simulated objects span the angular size of the real GNS galaxies. The simulated galaxies were created with GALFIT and convolved with a drizzled PSF image. They were set within a sky background equivalent to the mean NIC3 sky background within GNS (0.1 counts/sec). Source noise, sky noise, and read noise (29 e^-) were added to the frames.

The simulated images were then re-fit with GALFIT to derive (n, r_e). Initial guess parameters for (μ_e, r_e, n, b/a, PA) were generated randomly from uniform distributions spanning ±1.5 mag/arcsec^2 in μ_e, $\pm0\farcs3$ in r_e, ±2 indices in n, 0.3 to 1 in b/a, and -90 to 90 degrees in PA. Figure 11 shows the recovery in (n, r_e) plotted against surface brightness. The dashed vertical lines represent the range in μ_e for the observed massive galaxies. Figure 11 shows (n, r_e) are well recovered across the full range observed μ_e. The recovery as a function of μ_e severely degrades only at ~ 2 mag/arcsec^2 fainter than the observed range in μ_e. Among all simulated images, the mean $\Delta n = n_{\rm output} - n_{\rm input}$ and $\Delta n = r_{e,{\rm output}} - r_{e,{\rm input}}$ are 0.9 and $0\farcs01$, respectively. In $\sim 70\%$ of cases, n and r_e are recovered to within 10% of their input values; for the μ_e range in the NIC3/F160W band for the massive galaxies in our sample, the rate increases to $\sim 90\%$.

3.4.3. *Tests on MGC Fits*

The structural parameters for the massive galaxies at $z \sim 0$, are derived by Allen et al. (2006) by using the GIM2D code (Simard et al. 2002) to fit single Sérsic component to the MGC B-band images. We derived the structural parameters for the massive galaxies at $z \sim 1$–3, by using the GALFIT code (Peng et al. 2002) on the NIC3/F610W images (§3.2). One might wonder whether the dramatic shift in Fig. 5 of the $z \sim$2–3 galaxies toward lower (n, r_e) compared to the the $z \sim 0$ MGC galaxies may be caused by systematic differences between the fitting techniques used by us versus those by Allen et al. (2006). This would be the case only if the fits by Allen et al. (2006) give systematically higher (n, r_e) than ours for the same galaxies. As we show below this is not the case.

In order to address this issue, we have applied GALFIT to a subset of B-band MGC images and compared our resulting structural parameters to the GIM2D-based results given in the MGC catalogue. The comparison (top row of Fig. 12) shows that the GIM2D-based fits of Allen et al. (2006) are not biased to higher (n, r_e) compared to our GALFIT-based fits for the $z \sim 0$ MGC galaxies. In fact,

for large r_e, the GIM2D-based values may even be lower. in many cases.

These results are consistent with extensive comparisons of single component Sérsic fits from GALFIT and GIM2D conducted by Häussler et al.(2007) on both simulated and real galaxy data. They concluded that both codes provide reliable fits with little systematic error for galaxies with effective surface brightnesses brighter than that of the sky, as long as one is not dealing with highly crowded fields.

Another possible source of difference between the structural parameters of the $z \sim 0$ and $z \sim 2$–3 massive galaxies might be the fact that Allen et al. (2006) fitted the $z \sim 0$ massive galaxies with only a single Sérsic component, and did not include an extra point source component in galaxies with evident nuclear sources. It seems unlikely that the much larger fraction of higher (n, r_e) systems at $z \sim 0$ in Fig. 5 is mainly driven by this effect. To illustrate this, we have fitted the $z \sim 0$ MGC galaxies in the top row of Fig. 12 with a combination of a single Sérsic component and a point source model using GALFIT. The bottom row of Fig. 12 shows the results. The values of r_e are not changed systematically. The Sérsic index is lowered by the addition of the point source, but only 20% of the sources with $n > 2$ in the single Sérsic fit have $n \leq 2$ after including the point source. Since not all $z \sim 0$ MGC galaxies in Fig. 5 will have nuclear sources, the fraction of sources impacted will be even less. We thus conclude that the presence of a point source in some of the $z \sim 0$ MGC galaxies and the inclusion of such a point source in the model fits, are not sufficient to shift the $z \sim 0$ MGC galaxies into the parameter space occupied by the $z \sim 2$–3 massive galaxies in Fig. 5.

3.4.4. *Artificial Redshifting*

We next investigate whether redshift-dependent systematic effects could potentially cause the offset in (n, r_e) shown in Fig. 5 between massive galaxies at $z \sim 0$ and $z \sim 2$–3, by causing the (n, r_e) of massive $z \sim 0$ galaxies to 'degrade' into the regime of low $n \leq 2$ and low r_e exhibited by the $z \sim 2$–3 systems.

Ideally one would investigate this question by artificially redshifting the entire MGC subsample of 740 massive $z \sim 0$ galaxies shown in Fig. 5 out to $z \sim 2.5$, and re-decomposing the redshifted galaxies. However, this is not feasible for two reasons. Firstly, working with the full sample of 740 galaxies would be extremely time-consuming and not practically viable. Secondly, the redshifting software (FERENGI; Barden et al. 2008) that we use works optimally on SDSS *ugriz* images, which are not available for all the 740 massive $z \sim 0$ MGC galaxies in Fig. 5.

We therefore decide to artificially redshift a *representative* subsample S1 of 152 $z \sim 0$ massive ($M_\star \geq 5 \times 10^{10}$ $M_\odot$) galaxies, which have SDSS imaging and whose (n, r_e) distribution matches that of full MGC $z \sim 0$ subsample of 740 galaxies. The good match in (n, r_e) is illustrated in Fig. 13, where the magenta points in the top row show the (n, r_e) of the subsample S1 selected for redshifting, while the black points represent the full MGC subsample. A more quantitative illustration is provided by Figure 14, which shows that the n and r_e distributions of the subsample S1 match those of the massive MGC galaxies within $\pm 10\%$ in a given bin. We also ensure that the morphological mix of subsample S1 is a good match to that of the 740 MGC massive $z \sim 0$ galaxies (58% E/S0, 38% Sabc, 4% Sd/Irr) shown in Table 1. Many of the galaxies in our subsample S1 are well studied and include massive, S0, and Sabc galaxies from Barden et al. (2008), massive E galaxies in Kormendy et al. (2009), as well as bulge-dominated spirals and S0s from Eskridge et al. (2002)

Artificial redshifting was performed with FERENGI (Barden et al. 2008) to mimic the effects of surface brightness dimming, (with the standard $(1+z)^{-4}$ dependence on redshift for bolometric luminosity and $(1+z)^{-3}$ dependence at fixed passband), instrumental resolution, transmission efficiency, and PSF effects. The geometrical effects of cosmological redshift are corrected for by appropriately re-binning input images for the desired redshift and platescale. When multi-band images are provided, FERENGI models rest-frame filter curves for every pixel that are interpolated to find the expected per-pixel flux for the chosen observing filter.

Galaxies were redshifted to $z = 2.5$, convolved with a PSF, and then re-observed in the NIC3 F160W (H) band at a three-orbit depth (8063 seconds) and a resolution of $0\farcs2$/pixel. A sky background equal to the mean sky background of the GNS NIC3 images (0.1 counts/second) was added to the redshifted images. Poisson noise, sky noise, and read noise (29 e^- for NIC3) were then added to the redshifted images.

Surface brightness evolution counters surface brightness dimming, and without it, artificially redshifted images can be overly pessimistic. It has been observed Galaxy surface brightness evolution has been observed in the sense that galaxies at higher redshifts have higher mean surface brightness. For instance, Barden et al. (2005) find from the GEMS ACS survey that galaxies with $M_V \lesssim -20$ show a brightening of ~ 1 magnitude in rest-frame V-band by $z \sim 1$. Lilly et al. (1998) find that surface brightness for disk-dominated galaxies of similar properties (e.g., bulge-total ratio, scalelength, and axial ratio) increases on average by 0.8 magnitudes by $z = 0.7$. However, there is no set rule for the amount of surface brightness evolution to apply when artificially redshfiting galaxies. One reason for this is that the structure of high redshift galaxies may be inherently very different from that of local galaxies, so that it is not straightforward how to translate the observed higher surface brightness of distant galaxies into a recipe for brightening the redshifts images of local galaxies We therefore performed a set of redshifting tests, where we tested the effect of applying a surface brightness evolution (brightening) of 0, 1.25, 2.5, and 3.75 magnitudes between $z \sim 0$ and $z = 2.5$. There is a discernible difference in the recovered morphology and structural parameters between 0 and 1.25 magnitudes of brightening, but less difference between 1.25, 2.5, or 3.75 magnitudes of brightening.

Among these 3 values, we chose to apply 2.5 magnitude of surface brightness evolution (brightening) between $z \sim 0$ and $z = 2.5$ in our final artificial redshifting of the subsample S1 for several reasons: 1) As shown in § 3.3 and Figure 7, the rest-frame B-band optical surface brightness of massive galaxies at $z = 2-3$ in the GNS sample is between 2.5 to 6 magnitudes brighter than that of massive galaxies at $z \sim 0$. A 2.5 magnitudes of assumed evolution in the redshifting experiment seems adequately con-

servative. 2) Using the Bruzual & Charlot (2003) models, one can show that the passive evolution of a single stellar population from $z = 2.5$ to $z = 0$, assuming an exponentially declining star formation history associated with an e-folding time of 100 Myr, will lead the rest-frame B luminosity to decline by 2.5 to 3 magnitudes, depending on the chosen metallicity. 3) The adopted 2.5 magnitudes of evolution out to $z = 2.5$ corresponds to one magnitude of brightening per unit redshift. This is comparable to previous measurements of the brightening for intermediate and high-redshift galaxies (Lilly et al. 1998; Barden et al. 2005).

After artificially redshifting the subsample S1 from $z \sim 0$ to $z = 2.5$, and assuming 2.5 magnitude of surface brightness evolution, we fit both the original galaxy images and their redshifted counterparts with single Sérsic profiles. We compare the rest-frame optical structural parameters in the original and redshifted images in order to assess the influence of redshift-dependent systematic effects (e.g., surface brightness dimming, loss of spatial resolution) and see how well the structural parameters are recovered. We also compare the redshifted distribution of (n, r_e) to the one actually observed in the GNS massive galaxies to assess whether they are similar. Note that the structural parameters are measured at $z \sim 0$, from g-band images, while at $z = 2.5$ they are measured from the artificially redshifted images in the NIC3/F160W band so that all parameters are measured in the rest-frame blue optical light, thereby avoiding bandpass shifting problems. Our main results are outlined below.

1. Fig. 13 shows the (n, r_e) distribution (blue points in rows 2 and 3) obtained by redshifting the subsample S1 (magenta points is row 1) of 152 $z \sim 0$ massive MGC galaxies out to $z \sim 2.5$. This redshifted distribution of (n, r_e) is still significantly offset from those observed in the massive GNS galaxies at $z \sim$ 2-3 (red points in row 3).

   This difference is shown more quantitatively in Figure 14 where results in discrete bins of n and r_e are compared. *The massive GNS galaxies at $z = 2 - 3$ (red line) are significantly offset towards lower n and r_e compared to the redshifted sample (blue line).* The massive GNS galaxies at $z = 2 - 3$ (red line) includes $64.9 \pm 5.4\%$ of systems with low $n \leq 2$ and $18.2 \pm 4.4\%$ sytems with $n > 3$, while the corresponding fractions for the redshifted sample (blue line) are $15.4 \pm 2.84\%$ and $59.9 \pm 3.85\%$, respectively. Similarly for the r_e distribution, of the massive GNS galaxies at $z = 2 - 3$, $39.0 \pm 5.6\%$ have $r_e \leq 2$ kpc and $18.2 \pm 4.4\%$ have $r_e > 4$ kpc, while the redshifted sample has 1.9 $\pm 1.06\%$ and 78.4 $\pm 3.23\%$, respectively.

   We therefore conclude that cosmological and instrumental effects are not able to account for the large offset shown in Fig. 13 and Figure 14 between the (n, r_e) distributions of the massive galaxies at $z = 2 - 3$ and those at $z \sim 0$.

2. It is very interesting to look at how the structural parameters of galaxies of different morphological types change during the redshifting. Figure 15 compares the rest-frame optical structural parameters in massive E, S0, and spirals (drawn from sample S1) at $z \sim 0$ to the structural parameters recovered after these galaxies were artificially redshifted to $z = 2.5$ and re-observed with NIC3/F160W.

   We first discuss the behavior of the half-light or effective radius r_e. From Figure 15 , one can see that r_e is recovered to better than a factor of 1.5 for the cast majority of redshifted E/S0 and spirals of early-to-late Hubble types. In the case of a few $z \sim 0$ galaxies with highly extended halos or disks and associated large r_e , the recovered r_e at $z = 2.5$ can be nearly a factor of two lower than the original r_e at $z \sim 0$. Inspection of the surface brightness profiles shows that this effect primarily happens because surface brightness dimming prevents the outer lower surface brightness components of the galaxies from being adequately recovered after redshifting. However, this happens for only a small fraction of cases and this effect is not enough to account for the large offset in r_e between the $z = 2-3$ and $z \sim 0$ galaxies, as clearly illustrated by Fig. 13 and Fig. 14.

   In fact, it is striking that even after redshifting out to $z = 2.5$, practically none of the massive $z \sim 0$ galaxies (whether E/S0 or spirals of early-to-late Hubble types), fall into the regime of $r_e \leq 2$ kpc (shows as shaded areas) inhabited by the ultra-compact systems, which make up $\sim 40 (39.0 \pm 5.6\%$ for $M_\star \geq 5 \times 10^{10}$ $M_\odot$ and $40.6 \pm 9.0\%$ for $M_\star \geq 1 \times 10^{11}$ $M_\odot$) of the massive galaxies at $z = 2 - 3$ (see § 3.3). *Thus, these massive ultra-compact ($r_e \leq 2$ kpc) systems at $z = 2 - 3$ appear to truly have no analogs among $z \sim 0$ massive galaxies, in terms of their size, structure, and optical surface brightness.*

   The top row of Figure 15 shows the distribution of Sérsic index n before and after redshifting out to $z = 2.5$ The recovered Sérsic index n can be lower or higher than the original n at $z \sim 0$, but is recovered to better than a factor of two in all cases. The shaded area in the plots represents the regime of $n \leq 2$ where the majority ($64.9 \pm 5.4\%$ for $M_\star \geq 5 \times 10^{10}$ $M_\odot$ and $59.4 \pm 8.7\%$ for $M_\star \geq 1 \times 10^{11}$ $M_\odot$) of massive GNS galaxies at $z = 2 - 3$ lie (Table 1). It is interesting to note that massive E and S0s, which are spheroid-dominated and bulge-dominated systems, do not typically lie in the $n \leq 2$ regime, before or after redshifting. In contrast, a large fraction of $z \sim 0$ spirals with intermediate-to-late Hubble types[12] populate the $n \leq 2$ regime, both before and after redshifting.

   In these $z \sim 0$ spirals, the low index ($n \leq 2$) of the single Sérsic component fitted to the whole galaxy is caused by the presence of an outer disk, as well as the presence of pseudobulges or disky bulges. Pseudobulges are central flattened disk-like components, which lie within the inner kiloparsec of galaxies and

[12] The Hubble types are based on the bulge-to-total light ratio (B/T), which we measured with bulge-disk and bulge-disk-bar decomposition of $z \sim 0$ g-band images.

are believed to be formed as a result of gas inflows driven by stellar bars and minor mergers (e.g., Kormendy & Kennicutt 2004; Jogee et al. 1999, 2005). Such bulges predominantly have $n \leq 2$ (e.g., Fisher & Drory 2008; Weinzirl et al. 2009; Graham & Worley 2008) and can lower the overall single Sérsic index n of a spiral galaxy. In other words, in these $z \sim 0$ spiral galaxies, disk features on large and small scales lead to an overall single Sérsic index $n \leq 2$ for the entire galaxy. *It is possible that similar disk features are responsible at least in part, for the low $n \leq 2$ values shown by the majority ($\sim 60\%$) of the massive GNS galaxies at $z = 2-3$.* We discuss this point further in § 7.

3. In order to visually illustrate the redshfiting, Figure 16 shows the rest-frame optical images before and after redshifting for a set of representative galaxies drawn from the subsample S1 of massive galaxies used for the redshifting experiment. The galaxies shown span a range of morphological types and includes E, S0s, and spirals with a range of bulge-to-total ratio (B/T) and a range of effective radii. The value of n and r_e before and after redshifting are shown.

   For spirals of small size ($r_e \lesssim 4$ kpc) features in the g-band images, such as bulges, disks, and bars are almost always washed out in the redshifted images due to degraded resolution. For larger systems, one can still discern some disk features, such as bars and spirals. Since most ($\sim 80\%$) of the massive GNS galaxies at $z \sim$ 2-3 have $r_e \leq 4$ kpc (Table 1 and Fig. 13), this implies that we would not expect to see prominent visible disk features in them even if they host a massive disk. As described in § 2.2, most of the GNS galaxies are indeed fairly featureless (top 4 rows of Fig. 3), and only a minority show bar-like, clumpy, disk-like features (e.g. row 5 of Fig. 3).

   Only a small fraction ($\sim 8\%$) of the massive GNS galaxies at $z = 2-3$ appear to be disturbed, but for completeness, we also show how the rest-frame optical images and Sérsic index n of local massive merging galaxies compare before and after redshifting in Fig. 17. These galaxies are not drawn from subsample S1, but are picked to illustrate the early to late phases of mergers. In the early phase, the individual merger progenitors can be fitted separately with a Sérsic index component, and the results after redshifting depend on the nature of the individual progenitors: the results will be the same as for E and spirals described above. In the case, where the merger is in a late phase, artificial redshifting tends to blend the structures of the progenitors together and increase the overall Sérsic index n.

4. As an extra test, we repeated the redshifting experiment out assuming zero surface brightness evolution rather than 2.5 magnitudes of brightening out to $z = 2.5$. Even in this case there is still a large offset in the (n, r_e) distributions of the redshifted sample S1 compared to the massive GNS galaxies. Specifically, the fraction of systems with low $n \leq 2$ ($30.7 \pm 3.8\%$) is still significantly less than that for massive GNS galaxies at $z = 2-3$ ($64.9 \pm 5.4\%$). Likewise, there are still few systems with $r_e \leq 2$ kpc ($0.67 \pm 0.67\%$) compared to the high fraction ($39.0 \pm 5.6\%$) found at $z = 2-3$. Thus, even without surface brightness evolution it is still true that cosmological and instrumental effects are not able to account for the large offset between massive galaxies at $z = 2-3$ versus $z \sim 0$.

We summarize the results on the structural properties of the massive GNS galaxies at $z = 2-3$ in §7.1 to §7.2, where we discuss the potential challenges that they present to hierarchical models of galaxy evolution.

4. STAR FORMATION ACTIVITY

4.1. *Matching GNS Galaxies to MIPS 24 μm Counterparts*

The Spitzer GOODS Legacy Program (Dickinson et al. 2003a; Dickinson et al. in prep) provides deep *Spitzer* MIPS 24 μm observations of the GOODS fields. The 5σ flux limit is ~ 30 μJy and we only consider MIPS 24 μm counterparts with $f_{24\mu m} \geq 30$ μJy in the discussion below. The MIPS images have a pixel scale of $1\farcs2$/pixel and a PSF diameter of $6''$, which corresponds to ~ 42 kpc at $z \sim 1-3$. In contrast the NIC3/F160W PSF is $0\farcs3$. MIPS 24 μm counterparts of the massive GNS galaxies were identified with a maximum matching radius of $1\farcs5$. There are 84/166 massive GNS galaxies with MIPS 24 μm counterparts having $f_{24\mu m} \geq 30$ μJy.

There are several potential problems with the above procedure of assigning a 24 μm flux to a massive GNS galaxy. Firstly, the above procedure of matching a GNS massive galaxy in the NIC3/F160W images to the closest MIPS 24 μm source within a radius of $1\farcs5$, potentially allows for the situation where a given MIPS 24 μm source could be matched to several massive GNS galaxies. This would happen if some massive GNS galaxies were crowded within a radius of a few arcseconds so that the MIPS source would be within $1\farcs5$ of all of them. However, in practice this situation occurs for the minority (2/84, $\sim 2\%$) of massive galaxies with a MIPS counterpart. We reject the two cases where the match is ambiguous. This yields 82 positional matches between the massive GNS galaxies and the MIPS 24 μm counterparts that are unique and secure.

A second possible caveat is that within the large MIPS 24 μm PSF diameter of $6''$, there may be several other NIC3/F160W sources, in addition to the main massive GNS galaxy to which the MIPS source is matched. These extra NIC3/F160W sources may be galaxies with masses below the cutoff value of $M_\star \geq 5 \times 10^{10}$ $M_\odot$ for our sample, or other massive sample GNS galaxies. In such a scenario, all the extra NIC3 sources could potentially contribute to the MIPS 24 μm flux, and the above-described procedure of assigning all the 24 μm flux of the MIPS counterpart to a given massive GNS sample galaxy via closest positional match would overestimate the 24 μm flux of this galaxy.

In order to assess the extent of this potential problem, we proceed as follows. For the MIPS 24 μm counterpart assigned previously to each massive GNS galaxy, we first determine how many extra NIC3/F160W sources with

$M_\star \geq 10^9\ M_\odot$, in addition to the massive GNS galaxy, lie within a circle of diameter 6″ (i.e., the PSF diameter) centered on the MIPS source. Of the 82 massive GNS galaxies with a secure MIPS 24 μm counterpart, we find that 30 involve cases where there are extra NIC3 sources, in addition to the massive GNS galaxy, within the MIPS PSF diameter.

Next, we estimate the relative expected contributions of the massive GNS galaxy and the extra NIC3/F160W sources to the overall 24 μm flux by using the stellar mass ratio scaled by a function that takes into account the different redshifts of the sources. For two sources with stellar mass $M_{\star 1}$ and $M_{\star 2}$ having redshifts z_1 and z_2, and luminosity distances D_{L1} and D_{L2}, the stellar mass ratio $M_{\star 1}/M_{\star 2}$ is scaled by $((1+z_2)\times D_{L2}^2)/((1+z_1)\times D_{L1}^2)$. In 15 of the 30 cases, the extra NIC3 sources have a cumulative redshift-scaled mass that is $> 10\%$ that of the main GNS galaxy. In 8 of 30 cases, the cumulative redshift-scaled mass is $> 20\%$ of the mass of the main GNS galaxy and spans $\sim$ 40% to $\sim$ 126%. This happens due to the presence of two or more extra NIC3/sources within the MIPS PSF diameter. We reject these latter 8 cases rather than try to correct for the contamination,which in all cases is distributed across two or more nearby galaxies. For the remaining 22 cases, the contamination by extra NIC3/F160W sources is $< 20\%$ and we deem that our afore-described procedure of assigning all the 24 μm flux of the MIPS counterpart to the massive GNS galaxy is reasonable.

Therefore, in summary, 74/166 ($44.6 \pm 3.9\%$) massive ($M_\star \geq 5\times 10^{10}\ M_\odot$) GNS galaxies have a reliable MIPS 24 μm counterpart (with $f_{24\mu m} \geq 30$ μJy) whose entire flux is assigned to the massive GNS galaxy. In contrast, 82/166 ($49.4 \pm 3.9\%$), massive GNS galaxies do not have a MIPS counterpart with $f_{24\mu m} \geq 30$ μJy and in these cases we can only measure upper limits on their SFR. Table 4 lists the fraction of massive GNS galaxies with a MIPS 24 μm counterpart as a function of redshift. The top left panel in Figure 18 shows the 24 μm distribution for the 74 massive galaxies with a secure MIPS match and a 24 μm flux above the limiting flux of 30 μJy. There are 23/74 sources with $30 \leq f_{24\mu m} < 100$ μJy, 29/74 with $100 \leq f_{24\mu m} < 200$ μJy, and 22/74 with $f_{24\mu m} \geq 200$ μJy.

4.2. *Star Formation Measures*

In order to estimate the SFR, the total IR luminosity (L_{IR}) over 8–1000 μm is first estimated from the observed 24 μm flux (corresponding to rest-frame wavelengths of 6–12 μm over $z = 1-3$) by using SED templates from Chary & Elbaz (2001). Using solely 24 μm flux density to measure L_{IR} works well for inferred $L_{IR} \lesssim 10^{12}\ L_\odot$ galaxies at $z \sim 2$, but L_{IR} is overestimated by a factor of ~ 3 in more luminous galaxies (e.g., Papovich et al. 2007). Early results from Herschel (e.g., Elbaz et al. 2010; Nordon et al. 2010; D. Lutz, private communication) suggest that at $z > 1.5$, the SFRs extrapolated from 24 μm fluxes may overestimate the true SFR, typically by a factor of 2 to 4. This overestimate could be due to a rise in the strength of PAH features, changes in the SEDs, or AGN contamination at $z > 1.5$. Murphy et al. (2009) find that estimates of L_{IR} from 24 μm flux density alone are incorrect because the templates used are based on local galaxies with smaller PAH equivalent widths than galaxies of similar luminosity at high-redshift. We account for this discrepancy by making a correction for galaxies with inferred $L_{IR} > 6\times 10^{11}$ $L_\odot$ using $\log_{10}(\mathrm{L_{IR}}) = 0.59\times\log_{10}(\mathrm{L_{IR}^{24}}) + 4.8$, where L_{IR}^{24} is the infrared luminosity inferred solely from 24 μm flux density (R., Chary, private communication).

We investigate how the derived L_{IR} changes when using SED templates from Dale & Helou (2002) versus Chary & Elbaz (2001). Figure 19 compares the L_{IR} obtained from the Dale & Helou (2002) templates to the L_{IR} obtained from the Chary & Elbaz (2001) templates before and after the correction for $L_{IR} > 6\times 10^{11}$ $L_\odot$. Below $L_{IR} \sim 10^{11}$ $L_\odot$ the Dale & Helou values are larger but agree to within factor of ~ 1.5. From $L_{IR} \sim 10^{11}-10^{12}$ $L_\odot$, the templates give very similar results. Above this limit, the Chary & Elbaz templates give substantially larger L_{IR} before the correction for $L_{IR} > 6\times 10^{11}$ $L_\odot$ is applied; the deviation is a factor of ~ 7 at $L_{IR} \sim 10^{13}$ $L_\odot$ and a factor of ~ 20 at $\sim 10^{14}$ $L_\odot$. The discrepancy is reduced to a factor of $\lesssim 3$ after applying the correction, but the Chary & Elbaz L_{IR} are still higher overall.

The obscured star formation rate can be calculated using the expression $\mathrm{SFR_{IR}} = 9.8\times 10^{-11} L_{\mathrm{IR}}$ (Bell et al. 2007). This expression is based on a Chabrier IMF (Chabrier 2003) and assumes that the infrared emission is radiated by dust that is heated primarily by massive young stars. Uncertainties in the SFR estimates are a factor of ~ 2 or higher for individual galaxies. Using the Dale & Helou (2002) templates instead of the Chary & Elbaz (2001) templates would make the inferred $\mathrm{SFR_{IR}}$ higher by a factor of ~ 1.5 at $L_{IR} \lesssim 10^{11}$ $L_\odot$ and lower by a factor of $\lesssim 3$ for $L_{IR} > 6\times 10^{11}$ $L_\odot$.

If an AGN is present, then $\mathrm{SFR_{IR}}$ only gives an upper limit to the true SFR. In §6, we adopt several techniques (direct X-ray detections, IR power-law emission, and IR-to-optical excess) to identify AGN candidates in the sample and estimate the mean SFR for galaxies with and without a candidate AGN (see Table 4 and below). The upper right panel of Figure 18 shows the distribution of L_{IR} inferred from the 24 μm flux using the Chary & Elbaz (2001) templates and including the correction for $L_{IR} > 6\times 10^{11}$ $L_\odot$. The infrared luminosity spans 10^{11} to 10^{14} $L_\odot$, and the sample galaxies separate into LIRGs[13] (30/74), ULIRGs[14] (32/74), and a few (11/74) HyLIRGs[15]. The most extreme L_{IR} in the HyLIRGs are possibly due to IR emission from AGN, and in §6, we indeed find that 10 of these 11 HyLIRGs host an AGN. The lower right panel plots $\mathrm{SFR_{IR}}$ versus mass with the SFR per unit mass (SSFR) rates overlaid. The potential AGN candidates identified in §6 are coded separately. About $\sim 41\%$ (31/74) of sources have $\mathrm{SFR_{IR}} < 100$ $M_\odot$ yr^{-1}. Another $\sim 34\%$ (25/74) have $\mathrm{SFR_{IR}}$ between 100 and 500 $M_\odot$ yr^{-1}. There are 11/74 ($\sim 15\%$) sources with $\mathrm{SFR_{IR}} > 1000$ $M_\odot$ yr^{-1}, and 10 of these are AGN candidates.

For massive ($M_\star \geq 5\times 10^{10}\ M_\odot$) GNS galaxies, Fig-

[13] LIRGs are defined to have $10^{11} \leq L_{IR} < 10^{12}\ L_\odot$

[14] ULIRGs are defined to have $10^{12} \leq L_{IR} < 10^{13}\ L_\odot$

[15] HyLIRGs are defined to have $L_{IR} \geq 10^{13}\ L_\odot$

ure 20 plots SFR_{IR} versus $M_\star$ in three redshift bins ($z = 1 - 1.5$, $z = 1.5 - 2$, and $z = 2 - 3$) spanning cosmic time intervals of $\sim 1 - 1.4$ Gyr. Only galaxies with 24 μm flux above the limiting flux of 30 μJy are plotted. This limit corresponds to $\sim$ 4.3, 12.4, and 47.2 $M_\odot$ yr^{-1} in the three respective redshift bins (Table 4). For each redshift bin shown in the top and bottom left panels of Figure 20, the mean SFR_{IR} and the minimum detectable SFR are marked as dotted and solid lines. Table 4 and the bottom right panel of Figure 20 show how the mean SFR_{IR} varies with redshift before excluding potential AGN candidates. The mean SFR_{IR} rises from $\sim$ 64 to 222 to 1146 $M_\odot$ yr^{-1} across the three redshift bins, when no correction for AGN contamination is made. Once the AGN candidates identified in §6 are excluded, the mean SFR_{IR} changes to $\sim$ 44, 128, and 419 $M_\odot$ yr^{-1}, respectively. Once AGN candidates are excluded, the SFRs at $z \sim 1-3$ in Figure 20 are in good agreement with those reported for star-forming galaxies (e.g., see Fig. 1 of Genzel et al. 2010).

It would thus appear from Figure 20 that the mean SFR_{IR} of massive ($M_\star \geq 5 \times 10^{10}$ $M_\odot$) galaxies increases from $z = 1$ to $z = 3$. Part of this effect could in principle be due to the fact that the minimum detectable SFR is also rising rapidly with redshift. However, this cannot be the case as mean SFR_{IR} still rises with redshift when the limiting SFR_{IR} for $z = 2-3$ is applied to all redshift bins. Choosing the Chary & Elbaz (2001) templates versus the Dale & Helou (2002) templates for computing L_{IR} does not affect the trend we find in SFR_{IR} versus redshift. It is clear that the massive ($M_\star \geq 5 \times 10^{10}$ $M_\odot$) galaxies at $z = 2 - 3$ are actively forming stars at a high rate. Even if one scales down the SFR_{IR} by a factor of three, one still gets a SFR of at least several tens of $M_\odot$ yr^{-1}. Without any down-scaling, the SFR can be several hundreds of $M_\odot$ yr^{-1} at $z = 2 - 3$ (Figure 20).

4.3. *Relation Between Structure and Star-Formation*

Figures 21 and 22 summarize the structural properties (r_e and Sérsic index n) of galaxies with SFR_{IR} above the detection limit, as well as below the detection limit. The potential AGN candidates identified in § 6 are coded separately as $\Sigma_{SFR_{IR}}$ is likely overestimating the true SFR in the galaxy (see § 4).

Figure 21 shows the distribution of SFR_{IR} among systems of different n. On the LHS panel, galaxies with SFR_{IR} below the 5σ detection limit are shown as downward pointing arrows. For the histograms on the RHS panel, the y-axis shows the fraction of massive GNS galaxies in each redshift bin, while on the x-axis, we plot the actual value of SFR_{IR} for systems with SFR_{IR} above the 5σ detection limit (indicated by the black vertical line), and the upper limit for the other systems. AGN hosts are removed from the histograms since SFR_{IR} likely overestimates their true SFRs (see § 4).

Among our sample of massive ($M_\star \geq 5 \times 10^{10}$ $M_\odot$) galaxies at $z = 1 - 3$, the fraction of galaxies with SFR_{IR} above the 5σ detection limit is 44.6 ±3.9%, while the corresponding fraction among galaxies with $n \leq 2$ is 58.3 ±5.03%. In other words, compared to our full sample, the galaxies with $n \leq 2$ show an excess by a factor of $\sim$ 1.3 in the fraction of systems with SFR_{IR} above the 5 σ detection limit.

Figure 22 illustrates this further. Among the massive ($M_\star \geq 5{\times}10^{10}$ $M_\odot$) galaxies at $z = 2{-}3$ with SFR_{IR} above the 5σ detection limit, the majority (81.8 ±6.71%) have $n \leq 2$, a significant fraction (15.2 ±6.24%) have $2 < n \leq 4$, and only a minority (3.0 ±2.98%) have $n > 4$. The corresponding numbers across the lower redshift bin $z = 1 - 2$ are 70.7 ±7.11%, 26.8 ±6.92%, and 2.4 ±2.41%. Furthermore, the RHS panel of Figure 21 shows that the high SFR tail in each redshift bin is populated primarily by $n \leq 2$ systems. While the $n \leq 2$ disky systems have a wide range of SFR_{IR}, they include the systems of highest SFR. At $z = 2 - 3$, the mean SFR_{IR} after excluding AGN candidates is 216.3 $M_\odot$ yr^{-1} for systems with $n \leq 2$, compared to 19.9 $M_\odot$ yr^{-1} for systems with $n > 2$.

Next we look at the relationship between SFR and effective radius r_e. The distribution of SFR_{IR} for different r_e ranges is shown in Figure 22. The same convention as for Figure 21 is adopted, with upper limits being plotted for galaxies with SFR_{IR} below the 5σ detection limit, and only non-AGN systems being plotted on the RHS panel.

Among our sample of massive ($M_\star \geq 5 \times 10^{10}$ $M_\odot$) galaxies at $z = 1 - 3$, the fraction of galaxies with SFR_{IR} above the 5σ detection limit is 44.6 ±3.9%, while the corresponding fraction among the massive Class 1 or ultra-compact objects (defined in § 3.3 to be systems with $r_e \leq 2$ kpc), is only 14.8 ±4.83% (8/54). Thus, the ultra-compact massive galaxies shows a deficiency by a factor of $\sim$ 3 of systems with SFR_{IR} above the detection limit, compared to the whole sample of massive galaxies. Figure 22 further shows that among ultra-compact objects with SFR_{IR} above the 5σ detection limit, the mean SFR_{IR} for ultra-compact objects is 8.0 $M_\odot$ yr^{-1} at $z = 1 - 2$ and 104 $M_\odot$ yr^{-1} at $z = 2 - 3$. In contrast, the corresponding mean SFR_{IR} for the non-ultra-compact object ($r_e > 2$ kpc) are 60.0 $M_\odot$ yr^{-1} at $z = 1 - 2$ and 216.6 $M_\odot$ yr^{-1} at $z = 2 - 3$. Thus, at $z = 2 - 3$, the ultra-compact objects show a paucity of systems with high SFR_{IR}, and on average have lower SFR_{IR} than more extended systems.

Combining the information of n and r_e suggests that at $z \sim 2 - 3$, compared to the general sample, it is the extended ($r_e > 2$ kpc) galaxies with a low $n \leq 2$, that show the largest fraction (66.7 ±7.86%) of systems with high SFR_{IR} above the detection limit. These systems are reminiscent of the galaxies in the SINS survey (Förster Schreiber et al. 2009), which appear to have high SFR and rotating disk kinematics (Shapiro et al. 2008). Our SFRs are higher or comparable to the SINS galaxies at $z \sim 2$. Figure 23 compares the distributions of $M_\star$ and SFR for the SINS and massive GNS galaxies. Most (75%) SINS galaxies are less massive than 5×10^{10} $M_\odot$, while the GNS galaxies include the most massive galaxies at $z \sim 2$ with $M_\star \geq 10^{11}$ $M_\odot$. The median SFR of the SINS galaxies is $\sim$ 72 $M_\odot$ yr^{-1}, while the massive GNS galaxies without AGN have a median SFR_{IR} of $\sim$ 78 $M_\odot$ yr^{-1}.

4.4. *Comparison of SFR With Other Studies*

Our results suggest a decline in star formation over time. In Figure 20, the mean SFR falls as a function of redshift, reflecting the lower normalization in SFR with redshift. There is enough scatter in each bin that a well-defined SFR-mass sequence is not visible. The apparent flatness in SFR versus mass could be due to the the relatively

small dynamic range in stellar mass as well an abundance of IR-excess sources at $z \sim 2$. We demonstrate that the SFR derived at $z = 2-3$ are in agreement with the SFR from Daddi et al. (2007). The latter derive a SFR using the 1500 $\mathring{A}$ flux and estimate dust reddening using a Calzetti et al. (2000) reddening law. The lower-left panel of Figure 20 shows as black dots the SFR from UV measurements for galaxies at $z \sim 2$. The green line represents the SFR-mass correlation derived by Daddi et al. (2007). The GNS galaxies are scattered above and below the relation. There are several extreme outliers above the line that are due to AGN contamination.

Drory & Alvarez (2008) measure the SFR-mass correlation over a wide range in stellar mass ($M_\star \sim 10^9 - 10^{12}$ $M_\odot$) and redshift ($z \sim 0-5$). They measure SFR from UV data and use the dust-corrected continuum emission around 2800 $\mathring{A}$ to estimate the SFR. They parameterize the resulting SFR as a function of mass using a power-law with an exponential cutoff at high mass. At $z \geq 3$, average SFR is a power-law in stellar mass, but SFR drops at high mass starting at $z \sim 3$. The break mass separating the power-law from the exponential cutoff is a smoothly decreasing function of redshift. Between $z = 1-3$, the cutoff mass is $\sim 1-5\times10^{11}$ $M_\odot$, which is near the upper end of the GNS galaxy mass range.

In our work, the lower left panel of Figure 20 does not show a recognizable trend between SFR and mass within the $z = 2-3$ bin. This is probably because the narrow mass range ($5\times10^{10} \leq M_\star \leq 3\times10^{11}$ $M_\odot$) is very close to the cutoff mass where the SFR-mass correlation is flat. In the lower-right panel of Figure 20, the black line shows average SFR versus redshift for a 5×10^{10} $M_\odot$ galaxy as calculated by Drory & Alvarez (2008). The break in the average SFR is at $z \sim 2$, so average SFR falls rapidly at lower redshifts. Average $\mathrm{SFR_{IR}}$ for all massive GNS galaxies and the non-AGN hosts are displayed separately. The former deviates strongly from the calculations of Drory & Alvarez (2008), but the non-AGN values are quite consistent at $z < 2$.

Kurczynski et al. (2010) derive SFR by stacking ~ 600 star-forming BzK galaxies at $0.9 < z < 3.2$. From average stacked galaxy SEDs spanning UV to radio at median redshifts of 1.4, 1.8, and 2.2, they measure total combined SFR from IR and UV (SFR_{IR+UV}) of 58 ± 7, 100 ± 14, 130 ± 28 $M_\odot$ yr^{-1}. Their results for $z < 2$ are quite close to the $\mathrm{SFR_{IR}}$ we present in Figure 20 and Table 4; however, our $\mathrm{SFR_{IR}}$ measurement at $z = 2-3$ is a factor of ~ 3.2 higher. Kurczynski et al. (2010) caution against using only the 24 μm band to derive $\mathrm{SFR_{IR}}$ at $z > 2$, with which they find $\mathrm{SFR_{IR}}$ at median redshift 2.2 to be 760 $M_\odot$ yr^{-1}. This is larger by a factor of 1.8 than our measurement at $z = 2-3$ because they did not apply a correction to account for the intrinsically wider PAH equivalent widths of high-redshift galaxies compared to their templates.

Bauer et al. (2011) measure dust-corrected UV-based SFR ($\mathrm{SFR_{UV,corr}}$) for galaxies in GNS with $M_\star \geq 3\times10^9$ $M_\odot$ over $1.5 < z < 3$. Their $\mathrm{SFR_{UV,corr}}$ are based on the GOODS ACS z-band fluxes and assume a Salpeter IMF. The dust corrections are made with a UV spectral slope measured from SEDs spanning observer-frame optical through the mid-IR. For the massive ($M_\star \geq 5\times10^{10}$ $M_\odot$) galaxies the $\mathrm{SFR_{UV,corr}}$ of Bauer et al. (2011) can be higher by a factor of 10 than $\mathrm{SFR_{IR}}$ at low $\mathrm{SFR_{IR}} < 40$ $M_\odot$ yr^{-1}. At higher $\mathrm{SFR_{IR}}$ the IR-based values tend to be within a factor of $\sim 2-3$ of the UV-dust-corrected values, excluding the AGN candidates. Ultimately, the relationship between $\mathrm{SFR_{UV,corr}}$ and $\mathrm{SFR_{IR}}$ is non-monotonic and impacted by the uncertainties in a UV dust correction and the uncertainties in the templates for computing L_{IR} from 24 μm.

5. CONSTRAINTS ON COLD GAS CONTENT

The high estimated $\mathrm{SFR_{IR}}$ found in §4 suggest that copious cold gas reservoirs are present to fuel the star formation. For the 74 massive GNS galaxies with $\mathrm{SFR_{IR}}$ measurements above the 5 σ detection limit, we assume half of $\mathrm{SFR_{IR}}$ lies within the circularized half light-radius $\left(r_c = r_e \times \sqrt{b/a}\right)$ from single component Sérsic fits, and thereby estimate that the de-projected SFR per unit area as $\left(\Sigma_{\mathrm{SFR_{IR}}} = \frac{0.5\times\mathrm{SFR_{IR}}}{\pi\times r_c^2}\right)$. In galaxies which are potentially AGN, $\Sigma_{\mathrm{SFR_{IR}}}$ is likely overestimating the true SFR in the galaxy (see § 4). If potential AGN candidates are included, $\Sigma_{\mathrm{SFR_{IR}}}$ ranges from $\sim 0.10 - 360.80$ $M_\odot$ yr^{-1} kpc^{-2}, with a mean value of ~ 19.4 $M_\odot$ yr^{-1} kpc^{-2}. After excluding the potential AGN candidates $\Sigma_{\mathrm{SFR_{IR}}}$ ranges from $\sim 0.24-360.80$ $M_\odot$ yr^{-1} kpc^{-2}, with a mean value of ~ 14.8 $M_\odot$ yr^{-1} kpc^{-2}. This range is comparable to that seen in BzK/normal galaxies, ULIRGS and submillimeter galaxies (e.g., see Daddi et al. 2010b).

The issue of whether there is a Universal relation between molecular gas surface density and SFR surface density, which applies to different types of star-forming systems over a broad range of redshifts is not yet settled. Gnedin & Kravtsov (2010) find from simulated galaxies at $z \sim 3$ that the Schmidt-Kennicutt relation (Kennicutt et al. 1998) is steeper and has a lower amplitude than at $z \sim 0$. Studies based on direct CO observations suggest that different gas-SF laws may apply to starbursts or mergers, as compared to normal or non-merging systems, with the derived law being dependent on the assumed CO luminosity to gas mass conversion factors (Daddi et al. 2010b; Genzel et al. 2010). Given the lack of a universally agreed-upon gas-SF law, and given that we are only interested in estimating the cold gas surface density within a factor of a few, we decided to use the standard Schmidt-Kennicutt law with a power-law index of 1.4. If potential AGN candidates are included, the resulting implied cold gas surface density $\left(\Sigma_{\mathrm{gas}} = \left[\frac{10^4\times\Sigma_{\mathrm{SFR_{IR}}}}{2.5}\right]^{1/1.4}\right)$ ranges from $\sim 73-25,091$ $M_\odot$ pc^{-2}, with a median value of ~ 908 $M_\odot$ pc^{-2} (Figure 24). The corresponding values after excluding AGN candidates are $\sim 136 - 25091$ $M_\odot$ pc^{-2}, with a median value of ~ 607 $M_\odot$ pc^{-2} (Figure 24). These values are again comparable to those observed in BzK/normal galaxies, ULIRGS and submillimeter galaxies (e.g., see Daddi et al. 2010b).

In the subsequent discussion, we only cite values obtained after excluding AGN candidates, but Figure 24 also shows the values for the full sample of 74 galaxies. Next we estimate the cold gas fraction relative to the baryonic mass within r_c. For each galaxy, we use the above cold gas surface density within the circularized effective radius to estimate the total cold gas mass $\left(M_{\mathrm{gas}} = \Sigma_{\sigma_{\mathrm{gas}}} \times \pi \times r_c^2\right)$.

$M_{\rm gas}$ ranges from $3.4 \times 10^9 - 1 \times 10^{11}$ $M_\odot$, with a mean value of 1.9×10^{10} $M_\odot$ (Figure 24). The baryonic mass ($M_{\rm Baryon}$) within r_c is taken to be the sum of cold gas mass and stellar mass, and we assume that the latter term is half of the total stellar mass of the galaxy. The cold gas fraction ($f_{\rm gas} \equiv M_{\rm gas}/\left[M_{\rm gas} + M_\star\right]$) within r_c ranges from $6.5 - 71.0\%$, with a mean of $\sim 24\%$ (Figure 24). Thus, most of the massive galaxies at $z \sim$ 2-3 appear to be more gas-rich than comparably massive systems at $z \sim 0$, where the typical cold gas fraction ranges from 10% to 20%.

Figure 25 shows how gas fraction varies as a function of stellar mass and redshift, both with and without the AGN candidates. The mean gas fraction is shown in three redshift bins for galaxies with 5×10^{10} $M_\odot \leq M_\star < 10^{11}$ $M_\odot$ and $M_\star \geq 10^{11}$ $M_\odot$. For 5×10^{10} $M_\odot \leq M_\star < 10^{11}$ $M_\odot$ galaxies, the mean gas fraction (without AGN candidates) rises from $\sim 19\%$ to $\sim 25\%$ to $\sim 44\%$ across the three redshift bins. In comparison, for $M_\star \geq 10^{11}$ $M_\odot$ galaxies, the gas fraction changes little and stays between $\sim 14\%$ and $\sim 18\%$. On the mean, the less massive galaxies have gas fractions larger by a factor of $\sim 1.3 - 2.4$. The large 1σ error bars indicate considerable overlap in each bin, particularly at the highest redshift. Still, at a given redshift, the greatest gas fractions are found among the less massive galaxies.

Previous direct measurements of gas fraction in high-redshift galaxies are consistent with our inferred cold gas fractions for the massive GNS galaxies at $z = 1-3$. Daddi et al. (2008, 2010a) report gas fractions of 50-65% in massive ($M_\star \sim 4 \times 10^{10} - 1 \times 10^{11}$ $M_\odot$) IR-selected BzK galaxies at $z \sim 1.5$. Tacconi et al. (2010) also measure cold gas fraction from CO observations of high-redshift galaxies at $z = 1.1 - 2.4$. For stellar masses spanning $M_\star \sim 3 \times 10^{10} - 3.4 \times 10^{11}$ $M_\odot$, they find cold gas fractions in the range of $\sim 14 - 78\%$.

6. AGN IN MASSIVE GALAXIES AT $z = 1 - 3$

Active Galactic Nuclei (AGN) are characterized by significant emission across a range in wavelengths from the radio to the X-ray or even gamma rays. Accurately measuring the number of AGN candidates at high redshift requires a variety of selection techniques. Selection based on X-ray emission alone may fail at high redshift in the case of Compton-thick AGN where much of the soft X-ray emission is Compton scattered or absorbed by thick columns of gas ($N_H \gg 10^{24}$ cm^{-2}; Brandt et al. 2006). Such obscured AGN may radiate significantly in the mid-IR and submillimeter due to re-processing of absorbed UV, optical, and X-ray radiation. In order to identify AGN among the massive GNS galaxies at $z = 1-3$, we use a variety of techniques (X-ray properties, IR power-law, IR-to-optical excess, and mid-IR colors). § 6.1–6.4 describe each method and the number of AGN it identifies. Table 5 and § 6.5 summarize the overall frequency of AGN hosts in massive galaxies at $z = 1-3$.

6.1. *AGN Identification Via X-ray Properties*

X-ray counterparts to the massive GNS sources were searched for in the CDF-N and CDF-S catalogs of Alexander et al. (2003) and Luo et al. (2008), as well as the ECDF-S catalogs of Lehmer et al. (2005). A total of 33/166 ($19.9 \pm 3.1\%$) massive GNS galaxies had counterparts across all catalogs. The mean positional offset of the Chandra detections from the NICMOS positions is $\sim 0\farcs50$, slightly less than the mean error on the Chandra coordinates of $\sim 0\farcs65$, with a 1σ scatter of $0\farcs35$. The 30/33 objects with positional offsets $< 1''$ are considered secure matches. The three matches with offsets $1'' - 1\farcs5$ were visually inspected to ensure they were not ambiguous matches with other nearby NICMOS detections.

Not all of the 33 X-ray detections are necessarily AGN as intense X-ray emission can also be produced by high mass X-ray binaries, supernova remnants, hot gas in starburst winds, cataclysmic variables, etc. We therefore resort to additional X-ray properties, such as X-ray luminosity and X-ray spectral slope index, to estimate how many of the X-ray detections are likely AGN candidates.

The X-ray spectral slope index ($\Gamma \equiv \alpha + 1$) can provide additional insight on the populations represented by the X-ray detections. X-ray spectral slopes are provided in the CDF-N and CDF-S catalogs of Alexander et al. (2003) and Luo et al. (2008). As discussed by Alexander et al. (2005) and references therein, unobscured AGN and most star-forming galaxies have $\Gamma \approx 2$. However, some types of star-forming galaxies (e.g., SMGs) with a potential large population of high-mass X-ray binaries producing intrinsically hard X-ray emission may have flatter slopes ($\Gamma \sim 0.5 - 1$). Obscured AGN have flatter slopes ($\Gamma < 1$) due to energy-dependent photoelectric absorption of X-rays by gas columns. A value of $\Gamma < 0.5$ is taken as an unambiguous case of an obscured AGN, while $\Gamma \sim 0.5 - 1$ is possibly indicative of an obscured AGN.

We find that 13/166 ($7.8 \pm 2.1\%$) of the massive galaxies over $z = 1 - 3$ or 13/33 ($39.4 \pm 8.5\%$) of the X-ray detections have $\Gamma \leq 0.5$ and can therefore be considered as unambiguous obscured AGN. The remaining X-ray detections could have X-ray emission from star formation or/and AGN. Specifically, the 3/166 ($1.8 \pm 1.0\%$) massive galaxies with $0.5 < \Gamma < 1$ could be attributed to high-mass X-ray binaries or/and obscured AGN, while 17/166 ($10.2 \pm 2.4\%$) with $\Gamma > 1$, could be typical star-forming galaxies or/and unobscured AGN.

Figure 26 plots the X-ray luminosities $L_{\rm X}$ versus the the X-ray spectral index Γ. $L_{\rm X}$ is defined as $4\pi D_L^2 f_x \left(1+z\right)^{\Gamma-2}$, where D_L is the luminosity distance, f_x the full-band (0.5–8 keV) X-ray flux, and $(1+z)^{\Gamma-2}$ is the appropriate k-correction for a power-law spectrum. The 33 X-ray sources have full-band fluxes which are between 10^{-16} and 10^{-14} erg cm^{-2} s^{-1} and X-ray luminosities $L_{\rm X}$ between 10^{41} and 10^{45} erg s^{-1}. Typically AGN have X-ray luminosities $\geq 10^{42}$ erg s^{-1} (Donley et al. 2008), and we note that all except one of the sources with $\Gamma \leq 0.5$ (considered as unambiguous cases of obscured AGN) satisfy this criterion. Some of the sources with $\Gamma > 0.5$ have the highest X-ray luminosities ($> 10^{43}$ erg s^{-1}), suggesting that they might indeed be of mixed nature and host an AGN.

6.2. *AGN Identification Via Infrared Power-law SEDs*

The SEDs of pure star-forming galaxies are dominated by the 1.6 μm stellar bump and longer-wavelength dust-emission features, but when a sufficiently strong central engine is present, the gap between the stellar emission and the dust emission is filled by the mid-IR emission resulting

from the reprocessing of UV, optical, and X-ray radiation. The resulting composite SED may be well-fit with a power-law over $z = 1-3$ (e.g., Alonso-Herrero et al. 2006; Donley et al. 2008).

Following Alonso-Herrero et al. (2006) and Donley et al. (2008), we look for AGN power-law emission over $z = 1-3$ using SEDs from the IRAC bands at 3.6, 4.5, 5.8, and 8.0 μm. Over $z = 1-2$, IRAC bands primarily trace the rest-frame NIR and a small part of the rest-frame mid-IR SED out to 4 μm. Over $z = 2-3$, they primarily trace the rest-frame NIR out to 2.0 μm. All 166 of the massive galaxies were detected in all four IRAC bands. The massive GNS galaxies were matched against IRAC catalogs of the GOODS fields. The IRAC SEDs were fit with a power-law SED ($f_\nu \propto \nu^\alpha$) by minimizing the χ^2 between the observed and model SED. Following Alonso-Herrero et al. (2006), sources with power-law index $\alpha \leq -0.5$ are considered obscured AGN candidates, and we refer to these as power-law galaxies (PLGs). Only a small fraction (3/166 or $1.8 \pm 1.0\%$) of the massive galaxies over $z = 1-3$ are PLGs. The other sources are labeled as non-PLGs: one source has $-0.5 < \alpha < 0$, and all remaining sources have $\alpha \geq 0$.

We consider the three PLGs as likely AGN candidates. As shown on Fig. 26, one of the PLGs was already identified as an unambiguous obscured AGN because its X-ray spectral index is $\Gamma < 0.5$. The other two PLGs have $\Gamma > 0.5$ and were considered in § 6.1 as ambiguous cases where the X-ray emission can be due to SF and/or AGN activity. Thus, the power-law method adds two new likely AGN to our sample.

It is however striking that of the 13 sources which were identified as unambiguous obscured AGN § 6.1 due to their X-ray spectral index $\Gamma < 0.5$, only one is a PLG. This point is shown in both Figure 26 and Figure 27, which compares the X-ray spectral index Γ to the the IR power-law index α. The absence of more PLGs may be due to the AGN being intrinsically faint and outshone by the host galaxy in the IR.

6.3. *AGN Identification Via IR-to-Optical Excess*

Heavily obscured AGN may be present in highly reddened, IR-excess galaxies. Fiore et al. (2008) identify obscured AGN candidates in IR-bright, optically faint, red galaxies over $z = 1.2-2.6$ using the criteria $f_{24\mu m}/f_R \geq 1000$ and $R - K > 4.5$. We search for such IR-bright, optically faint systems with $f_{24\mu m}/f_R > 1000$ in our sample of massive galaxies. R-band flux is determined by linear interpolation between the *ACS V* and i-band fluxes.

We find 25 sources meeting this criteria, of which 16 are new AGN candidates not identified via the methods used so far, namely X-ray detections (§ 6.1) and IR power-law SEDs (§ 6.2). The remaining 9 sources have X-ray detections and are shown on Figure 27. They include 4 of the unambiguous obscured AGN with X-ray spectral slope index $\Gamma \leq 0.5$ (§ 6.1), and the 3 PLGs with mid-IR power-law index $\alpha \leq -0.5$ (§ 6.2).

6.4. *AGN identification via mid-IR colors*

Lacy et al. (2004) and Stern et al. (2005) separately formulate mid-IR color selection criteria for the *Spitzer* IRAC (Fazio et al. 2004) bands, 3.6 μm, 4.5 μm, 5.8 μm, and 8.0 μm. They show that AGN are likely to lie within specific regions of color-color plots based on IRAC photometry. The Lacy et al. (2004) sample involves galaxies at $z < 0.7$, while the Stern et al. (2005) sample consists primarily of galaxies at $z = 0.5-2$. However, when dealing with high redshift galaxies at $z \geq 2$, these criteria and in particular the Lacy criteria, are susceptible to contamination from high-redshift star-forming galaxies (e.g., Donley et al. 2008).

For completeness, we investigated both of these methods for our sample of 166 galaxies and find 22 potential AGN candidates. Of these 22 sources, 12 were identified already through their X-ray detections (§ 6.1), IR power-law SEDs (§ 6.2), and IR-to-Optical Excess (§ 6.3). Specifically, 12 have X-ray detections, six have an X-ray spectral slope index $\Gamma \leq 0.5$, and three are PLGs. The remaining 10 are new AGN candidates, which were not identified so far via any of these methods. However, due to the contamination problem from high-redshift star-forming galaxies, we do not consider the 10 new AGN candidates identified by these two methods to be reliable, and exclude them from our further analysis.

6.5. *Frequency of AGN*

In summary, we have used a variety of techniques (X-ray properties, IR power-law, and IR-to-optical excess) to identify AGN among the 166 massive GNS galaxies at $z = 1-3$. We find that 33/166 ($19.9 \pm 3.1\%$)of the galaxies have a Chandra X-ray detection and are likely AGN. Of the X-ray sources, 13 have X-ray spectral index $\Gamma \leq 0.5$ and can be considered as unambiguous obscured AGN. Among the remaining 20 X-ray sources, the X-ray emission can come from star formation or/and AGN: specifically three have $0.5 < \Gamma < 1$ and could be due to intense starbursts (e.g., SMGs) or/and obscured AGN, while 17 have $\Gamma > 1$ and could be star-forming galaxies or/and unobscured AGN. Furthermore, among the 33 X-ray detections, nine sources show IR-to-optical excess and three sources are power-law galaxies with index $\alpha \leq -0.5$ (usually indicative of an obscured AGN). Finally, in addition to the 33 X-ray detections, 16 more AGN candidates are identified based on their IR-to-optical excess, bringing the total number of AGN candidates to 49 among the 166 massive galaxies at $z = 1-3$. The corresponding AGN fraction is 49/166 or $29.5 \pm 3.5\%$. When the results are broken down in terms of redshift, the AGN fraction rises with redshift, increasing from $17.9 \pm 14.5\%$ at $z = 1-1.5$ to $40.3 \pm 8.8\%$ at $z = 2-3$. Table 5 summarizes the AGN selected by each method as a function of redshift.

The percentage of AGN among all massive GNS galaxies ($29.5 \pm 3.5\%$) is higher than at $z \sim 1$, where it is reported that less than 15% of the total 24 μm emission at $z < 1$ is in X-ray luminous AGN (e.g., Silva et al. 2004; Bell et al. 2005; Franceschini et al. 2005; Brand et al. 2006). Bluck et al. (2011) measure the fraction of X-ray detected AGN with $L_X > 2.35 \times 10^{43}$ erg s^{-1} in galaxies with $M_\star > 3 \times 10^{10}$ $M_\odot$. They find the fraction increases with redshift, reaching $7.4 \pm 2.0\%$ at $2.5 < z < 3.0$. Bluck et al. (2011) require a high X-ray luminosity ($L_X > 2.35 \times 10^{43}$ erg s^{-1}), and this results in a lower measured fraction relative to this work.

6.6. *Relation Between AGN Activity, Structure, and SFR*

We summarize the properties of the AGN host candidates and discuss their implications in terms of host galaxy structure, and inferred SFR.

Table 5 summarizes the mean $U-V$ color, Sérsic n, and r_e of the AGN as a function of redshift. Mean $U-V$ becomes progressively redder at earlier epochs when dust was more ubiquitous. On average, the AGN hosts are more compact at $z=2-3$ (mean $r_e\sim3$ kpc) than at $z=1-1.5$ (mean $r_e\sim4.4$ kpc). Sérsic index is ~2 in all redshift bins. Figure 28 shows the single Sérsic n versus r_e. Across the broad redshift range $z=2-3$, most ($81.58\pm6.29\%$) of the AGN hosts have relatively extended sizes with $r_e>2$ kpc. This fraction exceeds the fraction among all massive ($M_\star\geq5\times10^{10}$ $M_\odot$) GNS galaxies with $r_e>2$ kpc ($61.0\pm5.56\%$). A significant fraction of the AGN ($28.95\pm7.36\%$) have disky ($n\leq2$) morphologies. This fraction is much less than the fraction among the massive $z=2-3$ GNS galaxies with $n\leq2$ ($64.9\pm5.44\%$). Almost than one-quarter ($23.68\pm6.90\%$) of the AGN candidates are both disky ($n\leq2$) and extended ($r_e>2$ kpc). This is also smaller than the fraction among $z=2-3$ massive GNS galaxies ($46.7\pm9.11\%$) with $n\leq2$ and $r_e>2$ kpc.

If the disky extended AGN host candidates host massive black holes, it would imply that massive black holes are present in galaxies, which are not dominated by a massive spheroid, but have a fairly disky structure. In the local Universe, nearly all massive galaxies are believed to host a central supermassive black hole (Kormendy 1993; Magorrian et al. 1998; Ferrarese & Merritt 2000; Gebhardt et al. 2000; Marconi & Hunt 2003), and the black hole mass is tightly related to the bulge stellar velocity dispersion (Ferrarese & Merritt 2000; Gebhardt et al. 2000). This has led to the suggestion that the black hole and bulge or spheroid probably grew in tandem (e.g., Cattaneo & Bernardi 2003; Hopkins et al. 2006). The presence at $z=2-3$ of luminous and potentially massive black holes in high mass galaxies that do not seem to have a prominent bulge or spheroid may be at odds with this picture.

We can deduce simple constraints on the accretion rates of the black holes based on X-ray luminosity. Among the AGN with X-ray detections, the Chandra rest-frame X-ray luminosities are $\sim10^{42-43}$ erg s^{-1}. Luminous quasars have X-ray luminosities $>10^{44}$ erg s^{-1} (e.g., Szokoly et al. 2004) and typical accretion rates of 10-100 $M_\odot$ yr^{-1} (Jogee 2006). Thus, most AGN hosts appear to be less powerful than quasars and probably have slower accretion rates.

As discussed in § 4, if an AGN is present, then the IR luminosity L_{IR} and star formation rate ($\mathrm{SFR_{IR}}$) estimated from the 24 μm flux overestimate the true value. Among the 74 massive galaxies with a secure 24 μm flux above the limiting flux of $\sim$30 μJy, L_{IR} spans 10^{11} to 10^{14} $L_\odot$, giving LIRGs (30/74), ULIRGs (32/74), and a few (11/74) HyLIRGs. Ten of the HyLIRGs turn out to host an AGN. As shown earlier in Figure 20 and Table 4, the mean SFR is higher by a factor of $\sim2-2.5$ when AGN candidates are included (see § 4 for details).

7. DISCUSSION: FORMATION AND EVOLUTION OF MASSIVE GALAXIES

7.1. *Ultra-Compact Massive Galaxies at $z=2-3$*

We have found that massive GNS galaxies at $z\sim2-3$ are greatly offset toward lower (n, r_e) compared to $z\sim0$ galaxies of similar or higher stellar mass (Fig. 13 and Table 1). In particular, we find that approximately 40% ($39.0\pm5.6\%$ for $M_\star\geq5\times10^{10}$ $M_\odot$ and $40.6\pm9.0\%$ for $M_\star\geq1\times10^{11}$ $M_\odot$) of the massive galaxies at $z=2-3$ are in the form of ultra-compact ($r_e\leq2$ kpc) galaxies, compared to less than 1% at $z\sim0$ (Table 1). These ultra-compact galaxies have small sizes, which are practically unmatched among $z\sim0$ massive galaxies, and their surface brightness in the rest-frame optical can be 4-6 magnitudes brighter (Figure 7). Nearly half ($46.7\pm9.11\%$) of these ultra-compact systems at $z=2-3$ have a Sérsic index $n\leq2$.

Through the extensive tests we conducted in § 3.4.1 to § 3.4.4, we conclude that the offset in (n, r_e) between massive galaxies at $z\sim2-3$ and $z\sim0$ is real. It is not primarily caused by systematic effects related to the fitting techniques (e.g., use of GALFIT *versus* GIM2D, or the inclusion of point sources in model fits), instrumental effects, or redshift-dependent effects (e.g., cosmological surface brightness dimming and the loss of spatial resolution.) *Thus, the massive ultra-compact ($r_e\leq2$ kpc) systems at $z=2-3$ appear to have no analogs among $z\sim0$ massive galaxies, in terms of their size, structure, and optical surface brightness* .

How do these massive ultra-compact at $z=2-3$ form? The high surface brightness in the rest-frame B-band of the massive ultra-compact galaxies at $z=2-3$ implies a large mass surface density of young-to-intermediate-age stars, and therefore suggests that highly dissipative gas-rich events and powerful starbursts (associated with a large gas and SFR density) played an important role in their formation. We suggest that gas-rich or wet major mergers at $z>3$ are important for the buildup of the ultra-compact massive galaxies at $z=2-3$, since they produce large gas concentrations and intense starbursts, leading to dense stellar remnants. This suggestion is partly supported by recent reports that the merger fraction for massive ($M_\star\geq10^{11}$ $M_\odot$) galaxies continues to rise out to at least $z\sim3$ (Bluck et al. 2009).

However, we also find that the massive ultra-compact objects at $z=2-3$ show a paucity of systems with high $\mathrm{SFR_{IR}}$ with respect to the general population of massive galaxies, and on average have lower $\mathrm{SFR_{IR}}$ than more extended systems. These properties imply that the intense starburst phase at $z>3$ building these ultra-compact dense system must have ended by $z=2-3$, implying that either the cold gas has been ejected or heated rapidly through winds (associated with supernovae or AGN), despite the deep potential well of these massive galaxies.

Next we consider the question of how will the massive ultra-compact objects at $z=2-3$ evolve into $z\sim0$ massive galaxies, such as E and S0s? As illustrated by Figure 7, *this evolution would have to be a radical one, requiring r_e to change by a factor of 2-6, rest-frame surface brightness μ_e to dim by at least 5 magnitudes, and n to increase to $n>2$* in a large fraction of ultra-compact massive galaxies. An increase in (n, r_e, μ_e) can be achieved in via several pathways. One such pathway might involve a series of minor mergers, as well as cold mode accretion, which cu-

mulatively add mass to the outskirts of the ultra-compact galaxy (e.g., Naab et al. 2009). While this scenario may lead to the formation of S0s or lower mass ellipticals by $z \sim 0$, it seems unlikely to produce the really large changes in μ_e, n, r_e need to produce massive ellipticals by $z \sim 0$.

Another natural pathway to produce large changes in (n, r_e, μ_e) is a dry or gas-poor major merger of two disk systems, where violent relaxation leads to a remnant with $n \sim 4$, and a lower surface brightness than the progenitors. The importance of gas-poor or dry major mergers for making the most massive $z \sim 0$ ellipticals is also supported by other evidence. Firstly, the most massive local ellipticals are found to harbor cores (missing light), which are believed to be scoured by binary black holes that form in dry major mergers (Kormendy et al. 2009). Secondly, from a study of the tidal features associated with bulge-dominated early-type galaxies van Dokkum (2005) concludes that today's most luminous ellipticals form through mergers of gas-poor, bulge-dominated systems. Thirdly, over $z = 0 - 3$, Bluck et al. (2009) estimate that a galaxy with $M_\star \geq 10^{11}$ $M_\odot$ will experience 1.7 ± 0.5 major mergers. Finally, Robaina et al. (2010) find that the factor of ~ 4 increase in number density of red $M_\star \geq 1 \times 10^{11}$ $M_\odot$ galaxies between $z = 1$ and $z = 0$ can be accounted for by major mergers among $M_\star \geq 5 \times 10^{10}$ $M_\odot$ progenitors.

7.2. *A Large Population of Massive Galaxies at $z = 2 - 3$ With $n \leq 2$*

Another interesting result is our finding that a large fraction of massive galaxies at $z = 2 - 3$ have a low Sérsic index $n \leq 2$ (§ 3.3). Indeed, the majority (64.9%±5.4% for $M_\star \geq 5 \times 10^{10}$ $M_\odot$, and $59.4\% \pm 8.7\%$ for $M_\star \geq 10^{11}$ $M_\odot$) of massive galaxies at $z = 2 - 3$ have low $n \leq 2$, while the fraction at $z \sim 0$ is three times lower. We discuss the implication of this result on the structure of massive galaxies at $z = 2 - 3$, and the associated constraints on the assembly modes of galaxies.

It is well known that an idealized pure disk with an exponential surface brightness profile yields a Sérsic index $n = 1$, while an idealized spheroid with a de Vaucouleurs profile yields $n = 4$. In the latter case, the large $n = 4$ value is driven by the large extended wings in the surface brightness profile, combined with the steeply declining central surface brightness profile. Some insight into the interpretation of $n \leq 2$ values can be gleaned by considering massive galaxies at $z \sim 0$. As discussed in § 3.4.4 and illustrated in Figure 15, massive E and S0s, which are spheroid-dominated and bulge-dominated systems, are predominantly associated with $n > 2$, both at $z \sim 0$ and after artificially redshifting to $z = 2.5$. In contrast, spiral galaxies of intermediate to low bulge-to-total ratios, often have an overall low Sérsic index $n \leq 2$ (Figure 15) because they have a disk component (e.g. an outer disk or or a central disky pseudobulge), which contributes significantly to the total blue light of the galaxy.

In practice, the situation can be complicated by the fact that AGN or bright star-forming regions can drive the Sérsic index n of a galaxy to higher values. Galaxies (e.g., massive ellipticals) with an already high intrinsic $n = 4$ and extended wings are particularly susceptible to this n-pumping effect, but it affects all galaxies. Thus, in the presence of AGN and bright star-forming regions, even disk-dominated galaxies can end up with $n \gg 2$. Nonetheless, since these complications go in the direction of *raising* n, we can still make the case that for a massive galaxy at $z \sim 0$, a low Sérsic index $n \leq 2$ usually indicates the presence of a disk component (e.g. an outer disk or or a central disky pseudobulge)

An extension of these arguments to $z = 2-3$ would suggest that the large fraction (64.9%±5.4% for $M_\star \geq 5\times10^{10}$ $M_\odot$, $59.4\% \pm 8.7\%$ for $M_\star \geq 10^{11}$ $M_\odot$) of massive galaxies at $z = 2 - 3$ with low $n \leq 2$ is driven, at least partially, by the presence of a disk component (e.g an outer disk or central disky pseudobulge), which contributes significantly to the total rest-frame blue light of the galaxies. We next look into further tests of this possibility.

Figure 29 shows the observed axial ratio b/a of the massive galaxies at $z = 2 - 3$ as a function of the structural parameters n and r_e. We compare these distributions of b/a to those of galaxies at $z \sim 0$, considering spheroid-dominated systems such as ellipticals, bulge-dominated systems, such as S0s, and systems with more significant disk components, such as spirals. The observed distribution of b/a for a large number of nearby galaxies (2135 ellipticals, 4782 S0s, and 13482 spirals) is given in Binney & Merrifield (1998). For $z \sim 0$ elliptical galaxies, as well as S0s, the distribution of b/a peaks at $b/a \sim 0.75$ and declines rapidly at $b/a < 0.4$. Only a small fraction ($\sim 11\%$) of bulge-dominated systems have $b/a < 0.4$ (Binney & Merrifield 1998). In contrast, among $z \sim 0$ spiral galaxies, the distribution of b/a is a fairly flat for $b/a \sim 0.3 - 0.9$, and falls rapidly at $b/a < 0.3$. The latter low b/a values correspond to highly inclined spirals with an outer disk inclination larger than 72 degrees. Among massive galaxies at $z = 2 - 3$, the distribution of b/a for systems with low $n \leq 2$ (Figure 29, Row 1, left panel) or $r_e > 2$ kpc (Figure 29, Row 2, left panel) display a much stronger similarity to those of nearby spirals than to those of bulge-dominated S0 and ellipticals. In particular, the distribution of b/a does not show a strong decline at $b/a < 0.75$ and in fact peaks at $b/a \sim 0.4 - 0.5$. In contrast, among massive galaxies at $z = 2 - 3$ with $n > 2$ (Figure 29, Row 1, right panel), the distribution of b/a peaks at 0.75. There is also a larger fraction of systems with low $b/a < 0.4$ among massive galaxies with $n \leq 2$ (Row 1, left panel) than among those with $n > 2$ (Row 1, right panel). Thus, the observed axial ratios among massive galaxies with $n \leq 2$ or $r_e \leq 2$ kpc (Row 2, right panel) is consistent with the presence of a disk component among such galaxies, analogous to nearby spiral galaxies.

It is also interesting to note that most ($\sim 77\%$ for $M_\star \geq 5 \times 10^{10}$ $M_\odot$ see Table 3) of these massive galaxies at $z = 2 - 3$ with low $n \leq 2$ are extended ($r_e > 2$ kpc) rather than ultra-compact systems. This is in itself, does not prove that disk components exist in low $n \leq 2$ systems, but it is generally consistent with such a picture. Furthermore, we found in § 4.3 that at $z = 2 - 3$, the mean $\mathrm{SFR_{IR}}$ is significantly higher (216.3 $M_\odot$ yr^{-1}) for systems with $n \leq 2$ than for those with $n > 2$ (19.9 $M_\odot$ yr^{-1}). This result is generally consistent with the idea that the systems with $n \leq 2$ are actively star-forming and host copious amounts of gas, which tend to settle in disk-like configurations.

If disk components are indeed present in many massive

galaxies at $z = 2-3$ with $n \leq 2$ and $r_e > 2$ kpc, one may ask why disk features such as bars, spiral arms, and star-forming regions, are not evident in in the NIC3/F160W images (Fig. 3). Our artificial redshifting experiments (§ 3.4.4) and Fig. 16 show that for spirals with $r_e \lesssim 4$ kpc, features in the g-band images, such as bulges, disks, and bars are almost always washed out in the redshifted images due to degraded resolution. Since most ($\sim 80\%$) of the massive GNS galaxies at $z \sim$ 2-3 have $r_e \leq 4$ kpc (Table 1 and Fig. 13), this implies that we would not expect to see prominent visible disk features in them even if they host a massive disk. In the small fraction of massive galaxies at $z \sim 2-3$ with large $r_e \geq 4$ kpc (see row 5 of Fig. 3), the NIC3/F160w images do actually show some structural features such as an elongated bar-like feature or a central condensation surrounded by a more extended lower surface brightness feature.

In this work (§ 3.2), we fitted the NIC3/F160W images of the massive galaxies at $z = 2-3$ with single Sérsic components, rather than separate bulge and disk components because the low resolution (PSF FWHM of $0\farcs3$ corresponding to ~ 2.4 kpc at $z = 1-3$) of the images prevent reliable multiple component decomposition for all the galaxies, particularly the fairly compact ones. However, for the galaxies with large $r_e \geq 4$ kpc we attempted a bulge-plus-disk decomposition following the techniques outlined in Weinzirl et al. (2009). The decomposition was reliable only for the more extended systems within this group and yielded bulge-to-total light ratio below 0.4, indeed suggesting the presence of a significant disk component.

For completeness, we note that the presence of a disk component is not the only way to produce a low Sérsic index $n \leq 2$ in massive galaxies at $z = 2-3$. In particular, for the ultra-compact ($r_e \leq 2$ kpc) massive galaxies with $n \leq 2$, it has been argued that such systems could be somewhat like a massive elliptical, which has a bright high surface brightness central component surrounded by a very extended low surface brightness envelope. This system inherently has a Sérsic index $n = 4$, but if the low surface brightness envelope is somehow not detected by the NIC3/F160W images, then the latter could yield a lower $n \leq 2$, as the wings of the surface brightness profile would be effectively clipped. While such a scenario is in principle possible, our artificial redshifting experiments (§ 3.4.4) suggest it is unlikely. Figure 15 show that the Sérsic index n of massive local ellipticals does not fall in the $n \leq 2$ regime after they are redshifted to $z = 2.5$, assuming 2.5 magnitude of surface brightness evolution. The results hold even for 1.25 magnitudes of evolution, which is already a very conservative assumption, given that the rest-frame B-band optical surface brightness of massive galaxies at $z = 2-3$ in the GNS sample is between 2.5 to 6 magnitudes brighter than that of massive galaxies at $z \sim 0$ (Figure 7).

In summary, based on all the above tests and arguments, we conclude that many of the massive galaxies at $z = 2-3$ with $n \leq 2$, particularly the extended systems with $r_e > 2$ kpc, likely host a disk component (e.g., an outer disk or central disky pseudobulge), which contributes significantly to the total rest-frame blue light of the galaxies. The presence of a large population of massive galaxies at $z = 2-3$ with low $n \leq 2$ and disk components underlies the importance of cold mode accretion (Dekel et al. 2009; Keres et al. 2009; Brooks et al. 2009), in addition to major mergers, for building star-forming galaxies.

Major mergers of low-to-moderate gas fraction (e.g., $\leq 30\%$) will typically produce merger remnants with a de Vaucouleurs type profile and a Sérsic index $n \geq 4$. For major mergers whose progenitors have a higher gas fraction ($\sim$ 40% to 60%), the remnant is more disky with $2.5 < n < 6$, with *most having* $n > 3$ (e.g., Naab et al. 2006; see Fig. 14 of Hopkins et al. 2009; see Fig. 22 in Weinzirl et al. 2009). Thus, even with very gas-rich major mergers, it is challenging to have a scenario where the fraction of massive merger remnants with $n \leq 2$ is as high as $\sim 60\%$, to match the fraction of low $n \leq 2$ systems present among the massive galaxies at $z = 2-3$. However, in a cosmological context, as remnants at $z = 2-3$ acquire gas via cold mode accretion, a gas disk is expected to form. Depending on the angular momentum of the accreted gas, it can settle into a compact disk component (akin to a disky pseudobulge) within the remnant, or into an outer extended disk around the remnant. The formation of such a gas disk via cold mode accretion and its subsequent formation of a stellar disk, would lower the overall Sérsic index of the massive galaxies at $z = 2-3$, making them more in line with the observed values.

In cold mode accretion, gas from the intergalactic medium is channeled along cold filaments directly into star-forming regions without encountering a virial shock (e.g., Birnboim & Dekel 2003; Keres et al. 2005; Ocvirck, Pichon & Teyssier 2008; Brooks et al. 2009; Dekel et al. 2009; Keres et al. 2009). This process is particularly effective in galaxies with halos $\leq 10^{12}$ $M_\odot$, such that cold-mode accretion dominates the global growth of galaxies at high redshifts (e.g., $z > 2$), while at later times it mainly shapes the growth of lower mass objects. The halo mass of 10^{12} $M_\odot$ corresponds to a baryonic mass of $1.4 \times 10^{11} M_\odot$, which is right in the middle of the stellar mass range ($5.0 \times 10^{10} \leq M_\star/M_\odot \leq 6.0 \times 10^{11}$) of our sample. In semi-analytic models (SAMs) or non-cosmological simulations of major mergers, gas accretion from a spherical halo is included, but the efficient form of cold mode accretion, namely the accretion of gas along cold filaments, is not included. The challenge of producing a large population of disky ($n \leq 2$) systems with high SFRs from isolated major mergers (even gas-rich ones) may be an indication that the accretion of cold gas along cosmological filaments may be particularly important in the build-up of massive galaxies at $z = 2-3$, especially its large population of low $n \leq 2$ systems.

8. SUMMARY

We present a study of the structure, activity, and evolution of massive galaxies using one of the largest complete samples of massive galaxies at $z \sim 1$–3, with 166 systems having $M_\star \geq 5 \times 10^{10}$ $M_\odot$ and 73 with $M_\star \geq 10^{11}$ $M_\odot$. The target galaxies include Distant Red Galaxies (DRGs), Extremely Red Objects (EROs), and BzK-selected galaxies. The galaxies have NICMOS F160W observations from the GOODS NICMOS Survey (GNS), along with complementary ACS, *Spitzer* IRAC and MIPS, and Chandra X-ray data, and available stellar masses and photometric

redshifts. Our main results are summarized below.

1. Structural properties: We analyze the structure of these galaxies by fitting single Sérsic profiles to the 2D rest-frame optical light distribution. We find that the Sérsic index n and effective radius r_e of massive galaxies at $z \sim 2-3$ are strikingly offset toward lower values compared to $z \sim 0$ galaxies (Table 1 and Fig. 5). We describe this offset by highlighting in (2) and (3) two unusual populations of massive galaxies at $z = 2 = 3$: massive ultra-compact ($r_e \leq 2$ kpc) systems and the large population of galaxies with low $n \leq 2$.

2. Massive ultra-compact ($r_e \leq 2$ kpc) galaxies at $z = 2-3$: We find that approximately 40% ($39.0 \pm 5.6\%$ for $M_\star \geq 5 \times 10^{10}$ $M_\odot$ and $40.6 \pm 9.0\%$ for $M_\star \geq 1 \times 10^{11}$ $M_\odot$) of the massive galaxies at $z = 2-3$ are in the form of ultra-compact ($r_e \leq 2$ kpc) galaxies, compared to less than 1% at $z \sim 0$ (Table 1). These ultra-compact galaxies have small sizes, which are practically unmatched among $z \sim 0$ massive galaxies, and their surface brightness in the rest-frame optical can be 4-6 magnitudes brighter (Figure 7). Nearly half ($46.7 \pm 9.11\%$) of these ultra-compact systems at $z = 2-3$ have a Sérsic index $n < 2$.

Through the extensive fitting tests and artificial redshifting experiments we conclude that the offset in (n, r_e) between massive galaxies at $z \sim 2-3$ and $z \sim 0$ is real and not caused systematic effects related to the fitting techniques instrumental effects, or redshift-dependent effects (e.g., cosmological surface brightness dimming and the loss of spatial resolution.) *Thus, the massive ultra-compact ($r_e \leq 2$ kpc) systems at $z = 2-3$ appear to have no analogs among $z \sim 0$ massive galaxies, in terms of their size, structure, and optical surface brightness .*

The very high rest-frame B-band surface brightness of the massive ultra-compact galaxies at $z = 2-3$, implies a high stellar density and suggests that these systems formed through gas-rich wet major mergers or other classes of highly dissipative gas-rich events associated with powerful starbursts.

In order for massive ultra-compact ($r_e \leq 2$ kpc) objects at $z = 2-3$ to evolve into massive galaxies at $z \sim 0$, such as E and S0s, they will need to undergo a radical transformation, requiring r_e to change by a factor of 2-6, rest-frame surface brightness μ_e to dim by at least 5 magnitudes, and n to increase to $n > 2$. We suggest that this can happen effectively through dry major mergers, minor mergers and cold mode accretion, with dry major mergers playing a particularly important role.

3. A large population of massive galaxies at $z = 2-3$ with low $n \leq 2$: We find that the majority ($64.9\% \pm 5.4\%$ for $M_\star \geq 5 \times 10^{10}$ $M_\odot$, and $59.4\% \pm 8.7\%$ for $M_\star \geq 10^{11}$ $M_\odot$) of massive galaxies at $z = 2-3$ have low $n \leq 2$, while the corresponding fraction among massive systems at $z \sim 0$ is three times lower. Most ($\sim 67\%$) of these masive galaxies at $z = 2-3$ with low $n \leq 2$ have $r_e > 2$ kpc and therefore complement the ultra-compact galaxies. Based on artificial redshifting experiments and various properties (e.g., SFR, axial ratios, morphologies, r_e) of these systems, we argue that the massive galaxies at $z = 2-3$ with $n \leq 2$, particularly those with $r_e > 2$ kpc, likely host a disk component (e.g., an outer disk or central disky pseudobulge), which contributes significantly to the total rest-frame blue light of the galaxies.

The large fraction of massive galaxies at $z = 2-3$ with low $n \leq 2$, disk components, and high SFR, underscores the importance of cold mode accretion, in addition to gas-rich major mergers, in building such a population. Even with very gas-rich major mergers, it is challenging to have a scenario where the fraction of massive merger remnants with $n \leq 2$ is as high as $\sim 60\%$, to match the fraction of low $n \leq 2$ systems present among the massive galaxies at $z = 2-3$. However, the inclusion of cold mode accretion in models helps to alleviate this problem and may be particularly important in the build-up of massive galaxies at $z = 2-3$, especially its large population of low $n \leq 2$ systems.

4. Star formation rates: We estimate star formation rates using IR luminosities (8-1000 μm) derived from the *Spitzer* 24 μm flux for massive GNS galaxies, which have a secure MIPS 24 μm counterpart and a 24 μm flux exceeding the 5σ detection limit of 30 μJy. Among our sample of massive ($M_\star \geq 5 \times 10^{10}$ $M_\odot$) galaxies at $z = 1-3$, only $44.6 \pm 3.9\%$ of systems have $\mathrm{SFR_{IR}}$ above the 5σ detection limit, and in these cases we can only measure upper limits on SFR.

The mean $\mathrm{SFR_{IR}}$ of massive ($M_\star \geq 5 \times 10^{10}$ $M_\odot$) galaxies increases from $z = 1$ to $z = 3$. Across $z = 1-1.5$, $z = 1.5-2$, and $z = 2-3$, it rises from ~ 64 to 222 to 1146 $M_\odot$ yr^{-1} when no correction for AGN contamination is made (Table 4 and Figure 20). Once AGN candidates are excluded, the mean $\mathrm{SFR_{IR}}$ changes to $\sim$44, 128, and 419 $M_\odot$ yr^{-1}, respectively.

The massive ultra-compact objects at $z = 2-3$ are less likely by a factor of ~ 3 to have $\mathrm{SFR_{IR}}$ above the detection limit, compared to the whole sample of massive galaxies. On average they also have lower mean $\mathrm{SFR_{IR}}$ with values of 8.0 $M_\odot$ yr^{-1} at $z = 1-2$, and 104 $M_\odot$ yr^{-1} at $z = 2-3$.

Furthermore, the structure and SFR are correlated: across $z = 1-3$, the majority of galaxies with high $\mathrm{SFR_{IR}}$ are disky ($n \leq 2$; Figure 21) and extended ($r_e > 2$ kpc). At $z \sim 2-3$, compared to the general sample, it is the extended ($r_e > 2$ kpc) galaxies with a low $n \leq 2$, that show the largest fraction ($66.7 \pm 7.86\%$) of systems with high $\mathrm{SFR_{IR}}$ above the detection limit.

5. Cold gas fractions: The high estimated $\mathrm{SFR_{IR}}$ suggests that copious cold gas reservoirs are present. Assuming a Schmidt-Kennicutt law (Kennicutt et al. 1998) with a power-law index of 1.4, we estimate that the average cold gas surface density in non-AGN hosts ranges from ~ 136 to $\sim 25,091$ $M_\odot$ pc^{-2}, with a median value of ~ 607 $M_\odot$ pc^{-2} (Figure 24). The associated estimates of the cold gas mass ranges from 3.4×10^9 to 1×10^{11} $M_\odot$, with a mean value of 1.9×10^{10} $M_\odot$ (Figure 24). The implied cold gas fraction (ratio of cold gas mass to baryonic mass) within r_e ranges from 6.5% to 71.0%, with a mean of $\sim 24\%$ (Figure 24). Thus, most of the massive galaxies at $z \sim$ 2-3 appear to be more gas-rich than comparably massive systems at $z \sim 0$.

6. AGN activity: Using a variety of techniques (X-ray properties, IR power-law, and IR-to-optical excess) to

identify AGN, we find that 49/166 (29.5±3.5%) of the massive galaxies at $z = 1 - 3$ are AGN candidates. The AGN fraction rises with redshift, increasing from $17.9 \pm 14.5\%$ at $z = 1 - 1.5$ to $40.3 \pm 8.8\%$ at $z = 2 - 3$ (Table 5). Many ($85.0 \pm 5.0\%$) of the AGN candidates are extended with $r_e > 2$ kpc or have disky ($n \leq 2$) morphologies ($42.3 \pm 7.1\%$). More than one-third ($38.8 \pm 7.0\%$) of the AGN candidates are both disky ($n \leq 2$) and extended ($r_e > 2$ kpc). If these AGN candidates host massive black holes, then it would imply that such black holes are present in galaxies, which are not dominated by a massive spheroid, but have a fairly disky structure. These systems would diverge from the tight correlation observed between bulge stellar velocity dispersion and black hole mass in the local Universe.

S.J. and T.W. acknowledge support from the National Aeronautics and Space Administration (NASA) LTSA grant NAG5-13063, NSF grant AST-0607748, and *HST* grants GO-10395 and GO-10861 from STScI, which is operated by AURA, Inc., for NASA, under NAS5-26555. CJC acknowledges support from STFC and the Leverhulme Foundation. We thank Knud Jahnke and Marco Barden for technical assistance with the operation of FERENGI. We acknowledge the usage of the HyperLeda database (http://leda.univ-lyon1.fr). Some/all of the data presented in this paper were obtained from the Multimission Archive at the Space Telescope Science Institute (MAST). STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. Support for MAST for non-HST data is provided by the NASA Office of Space Science via grant NAG5-7584 and by other grants and contracts. The Millennium Galaxy Catalogue consists of imaging data from the Isaac Newton Telescope and spectroscopic data from the Anglo Australian Telescope, the ANU 2.3m, the ESO New Technology Telescope, the Telescopio Nazionale Galileo and the Gemini North Telescope. The survey has been supported through grants from the Particle Physics and Astronomy Research Council (UK) and the Australian Research Council (AUS). The data and data products are publicly available from http://www.eso.org/ jliske/mgc/ or on request from J. Liske or S.P. Driver.

REFERENCES

Alexander, D. M., Bauer, F. E., Chapman, S. C., Smail, I., Blain, A. W., Brandt, W. N., & Ivison, R. J. 2005, ApJ, 632, 736
Alexander, D. M., et al. 2003, AJ, 126, 539
Allen, P. D., Driver, S. P., Graham, A. W., Cameron, E., Liske, J., & de Propris, R. 2006, MNRAS, 371, 2
Alonso-Herrero, A., et al. 2006, ApJ, 640, 167
Baldry, I. K., Glazebrook, K., Brinkmann, J., Ivezić, Ž., Lupton, R. H., Nichol, R. C., & Szalay, A. S. 2004, ApJ, 600, 681
Barden, M., Jahnke, K., Haüßler, B. 2008, ApJS, 175, 105
Barden, M., et al. 2005, ApJ, 635, 959
Barnes, J. E., & Hernquist, L. E. 1991, ApJL, 370, L65
Beckwith, S. V. W., et al. 2006, AJ, 132, 1729
Bell, E. F., Zheng, X. Z., Papovich, C., Borch, A., Wolf, C., & Meisenheimer, K. 2007, ApJ, 663, 834
Bell, E. F., et al. 2004, ApJL, 600, L11
Bell, E. F., et al. 2005, ApJ, 625, 23
Binney, J., & Merrifield, M. 1998, Galactic astronomy / James Binney and Michael Merrifield. Princeton, NJ: Princeton University Press, 1998. (Princeton series in astrophysics) QB857 .B522 1998
Blain, A. W., Smail, I., Ivison, R. J., Kneib, J.-P., & Frayer, D. T. 2002, Phys. Rep., 369, 111
Bluck, A. F. L., Conselice, C. J., Almaini, O., Laird, E. S., Nandra, K., & Grützbauch, R. 2011, MNRAS, 410, 1174
Bluck, A. F. L., Conselice, C. J., Bouwens, R. J., Daddi, E., Dickinson, M., Papovich, C., & Yan, H. 2009, MNRAS, 394, L51
Borch, A., et al. 2006, A&A, 453, 869
Bournaud, F., et al. 2008, A&A, 486, 741
Brand, K., et al. 2006, ApJ, 644, 143
Brandt et al. 2006, in Physics of Active Galactic Nuclei at all Scales, ed. D. Alloin, R. Johnson, & P. Lira (Berlin: Springer), 185
Brooks, A. M., Governato, F., Quinn, T., Brook, C. B., & Wadsley, J. 2009, ApJ, 694, 396
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Buitrago, F., Trujillo, I., Conselice, C. J., Bouwens, R. J., Dickinson, M., & Yan, H. 2008, ApJL, 687, L61
Cattaneo, A., & Bernardi, M. 2003, MNRAS, 344, 45
Chabrier, G. 2003, ApJL, 586, L133
Chary, R., & Elbaz, D. 2001, ApJ, 556, 562
Cimatti, A., et al. 2008, A&A, 482, 21
Cole, S., Lacey, C. G., Baugh, C. M., & Frenk, C. S. 2000, MNRAS, 319, 168
Conselice, C. J. 2006, MNRAS, 373, 1389
Conselice, C. J., Rajgor, S., & Myers, R. 2008, MNRAS, 386, 909
Conselice, C. J., et al. 2010, MNRAS submitted (arXiv:1010.1164)
Conselice, C. J., et al. 2007, MNRAS, 381, 962
Daddi, E., Dannerbauer, H., Elbaz, D., Dickinson, M., Morrison, G., Stern, D., & Ravindranath, S. 2008, ApJL, 673, L21
Daddi, E., et al. 2005, ApJ, 626, 680
Daddi, E., et al. 2007, ApJ, 670, 156
Daddi, E., et al. 2010a, ApJ, 713, 686
Daddi, E., et al. 2010b, ApJL, 714, L118
Dale, D. A., & Helou, G. 2002, ApJ, 576, 159
Dekel, A., Sari, R., & Ceverino, D. 2009, ApJ, 703, 785
Dickinson, M., Giavalisco, M., & GOODS Team 2003a, in The Mass of Galaxies at Low and High Redshift, ed. R. Bender & A. Renzini (Heidelberg: Springer-Verlag), 324
Dickinson, M., Papovich, C., Ferguson, H. C., & Budavári, T. 2003b, ApJ, 587, 25
Dickinson, M., et al. 2004, ApJL, 600, L99
Donley, J. L., Rieke, G. H., Pérez-González, P. G., & Barro, G. 2008, ApJ, 687, 111
Driver, S. P., Liske, J., Cross, N. J. G., De Propris, R., & Allen, P. D. 2005, MNRAS, 360, 81
Driver, S. P., et al. 2006, MNRAS, 368, 414
Drory, N., & Alvarez, M. 2008, ApJ, 680, 41
Elbaz, D., et al. 2010, A&A, 518, L29
Eskridge, P. B., et al. 2002, ApJS, 143, 73
Fazio, G. G., et al. 2004, ApJS, 154, 10
Ferrarese, L., & Merritt, D. 2000, ApJL, 539, L9
Fiore, F., et al. 2008, ApJ, 672, 94
Fisher, D. B., & Drory, N. 2008, AJ, 136, 773
Förster Schreiber, N. M., et al. 2009, ApJ, 706, 1364
Franceschini, A., et al. 2005, AJ, 129, 2074
Gebhardt, K., et al. 2000, ApJL, 539, L13
Genzel, R., et al. 2008, ApJ, 687, 59
Genzel, R., et al. 2010, MNRAS, 407, 2091
Giacconi, R., et al. 2002, ApJS, 139, 369
Giavalisco, M., et al. 2004, ApJL, 600, L93
Gnedin, N. Y., & Kravtsov, A. V. 2010, ApJ submitted (arXiv:1004.0003)
Graham, A. W., & Worley, C. C. 2008, MNRAS, 752
Grützbauch, R., Chuter, R. W., Conselice, C. J., Bauer, A. E., Bluck, A. F. L., Buitrago, F., & Mortlock, A. 2010, arXiv:1011.4846
Häussler, B., et al. 2007, ApJS, 172, 615
Hopkins, P. F., Cox, T. J., Dutta, S. N., Hernquist, L., Kormendy, J., & Lauer, T. R. 2009a, ApJS, 181, 135
Hopkins, P. F., & Elvis, M. 2010, MNRAS, 401, 7
Hopkins, P. F., Hernquist, L., Cox, T. J., Robertson, B., & Springel, V. 2006, ApJS, 163, 50
Jogee, S. 2006, in Physics of Active Galactic Nuclei at all Scales, ed. D. Alloin, R. Johnson, & P. Lira (Berlin: Springer), 1
Jogee, S., Kenney, J. D. P., & Smith, B. J. 1999, ApJ, 526, 665
Jogee, S., Scoville, N., & Kenney, J. D. P. 2005, ApJ, 630, 837
Jogee, S., et al. 2004, ApJL, 615, L105
Kennicutt, R. C., Jr. 1998, ApJ, 498, 541
Kereš, D., Katz, N., Fardal, M., Davé, R., & Weinberg, D. H. 2009, MNRAS, 395, 160
Khochfar, S., & Silk, J. 2006, ApJL, 648, L21
Kormendy, J., Fisher, D. B., Cornell, M. E., & Bender, R. 2009, ApJS, 182, 216
Kormendy, J. in the Nearest Active Galaxies (eds Beckman, J., Colina, L., & Netzer, H.) 197-218 (Consejo Superior de Investigaciones Cientificas, Madrid, 1993)
Kormendy, J., & Kennicutt, R. C., Jr. 2004, ARAA, 42, 603
Krist, J. 1995, Astronomical Data Analysis Software and Systems IV, 77, 349
Kurczynski, P., et al. 2010, arXiv:1010.0290
Lacy, M., et al. 2004, ApJS, 154, 166
Lehmer, B. D., et al. 2005, ApJS, 161, 21
Lilly, S., et al. 1998, ApJ, 500, 75
Liske, J., Lemon, D. J., Driver, S. P., Cross, N. J. G., & Couch, W. J. 2003, MNRAS, 344, 307
Loeb, A., & Peebles, P. J. E. 2003, ApJ, 589, 29
Longhetti, M., et al. 2007, MNRAS, 374, 614
Luo, B., et al. 2008, ApJS, 179, 19
Magorrian, J., et al. 1998, AJ, 115, 2285
Marconi, A., & Hunt, L. K. 2003, ApJL, 589, L21
Mihos, J. C., & Hernquist, L. 1996, ApJ, 464, 641
Moustakas, L. A., et al. 2004, ApJL, 600, L131
Murphy, E. J., Chary, R.-R., Alexander, D. M., Dickinson, M., Magnelli, B., Morrison, G., Pope, A., & Teplitz, H. I. 2009, ApJ, 698, 1380
Naab, T., Johansson, P. H., Ostriker, J. P., & Efstathiou, G. 2007, ApJ, 658, 710
Naab, T., Johansson, P. H., & Ostriker, J. P. 2009, ApJL, 699, L178
Nordon, R., et al. 2010, A&A, 518, L24
Ocvirk, P., Pichon, C., & Teyssier, R. 2008, MNRAS, 390, 1326
Papovich, C., et al. 2006, ApJ, 640, 92
Papovich, C., et al. 2007, ApJ, 668, 45
Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, AJ, 124, 266
Pierce, C. M., et al. 2010, MNRAS, 405, 718
Ravindranath, S., et al. 2004, ApJL, 604, L9
Reddy, N. A., Steidel, C. C., Erb, D. K., Shapley, A. E., & Pettini, M. 2006, ApJ, 653, 1004
Robaina, A. R., Bell, E. F., van der Wel, A., Somerville, R. S., Skelton, R. E., McIntosh, D. H., Meisenheimer, K., & Wolf, C. 2010, ApJ, 719, 844
Robertson, B., Cox, T. J., Hernquist, L., Franx, M., Hopkins, P. F., Martini, P., & Springel, V. 2006, ApJ, 641, 21
Schulze, A., & Wisotzki, L. 2010, arXiv:1004.2671
Sersic, J. L. 1968, Cordoba, Argentina: Observatorio Astronomico, 1968
Shapiro, K. L., et al. 2008, ApJ, 682, 231
Shen, S., Mo, H. J., White, S. D. M., Blanton, M. R., Kauffmann, G., Voges, W., Brinkmann, J., & Csabai, I. 2003, MNRAS, 343, 978
Silva, L., Maiolino, R., & Granato, G. L. 2004, MNRAS, 355, 973
Simard, L., et al. 2002, ApJS, 142, 1
Steinmetz, M., & Navarro, J. F. 2002, NewA, 7, 155
Stern, D., et al. 2005, ApJ, 631, 163
Szokoly, G. P., et al. 2004, ApJS, 155, 271
Tacconi, L. J., et al. 2006, ApJ, 640, 228
Tacconi, L. J., et al. 2008, ApJ, 680, 246
Tacconi, L. J., et al. 2010, Nature, 463, 781
Thompson, R. I., et al. 2005, AJ, 130, 1
Toft, S., et al. 2007, ApJ, 671, 285
Trujillo, I., Conselice, C. J., Bundy, K., Cooper, M. C., Eisenhardt, P., & Ellis, R. S. 2007, MNRAS, 382, 109
Trujillo, I., et al. 2006, ApJ, 650, 18
van Dokkum, P. G. 2005,
van Dokkum, P. G., et al. 2008, ApJL, 677, L5
van Dokkum, P. G., et al. 2010, ApJ, 709, 1018
Weinzirl, T., Jogee, S., Khochfar, S., Burkert, A., & Kormendy, J. 2009, ApJ, 696, 411
Yan, H., et al. 2004, ApJ, 616, 63
Yan, L., & Thompson, D. 2003, ApJ, 586, 765
Zirm, A. W., et al. 2007, ApJ, 656, 66

TABLE 1
REST-FRAME OPTICAL SÉRSIC INDEX n AND r_e IN MASSIVE ($M_\star \geq 5 \times 10^{10}\ M_\odot$) GALAXIES

z	Morphology	$n \leq 2$	$n > 2$	$n > 3$
		$M_\star \geq 5 \times 10^{10}\ M_\odot$		
$z = 2 - 3$ ($N = 77$)	All	**64.9** ±5.44%	35.1 ±5.44%	18.2 ±4.40%
$z = 1 - 2$ ($N = 89$)	All	**49.4** ±5.30%	50.6 ±5.30%	30.3 ±4.87%
$z \sim 0$ ($N = 740$)	All	**18.1** ±1.42%	81.9 ±1.42%	66.6 ±1.73%
	E/S0	0.70 ±0.30%	57.6 ±1.82%	50.5 ±1.84%
	Sabc	14.5 ±1.29%	23.6 ±1.56%	15.7 ±1.34%
	Sd/Irr	2.97 ±0.62%	0.68 ±0.30%	0.41 ±0.23%
		$M_\star \geq 1 \times 10^{11}\ M_\odot$		
$z = 2 - 3$ ($N = 32$)	All	**59.4** ±8.68%	40.6 ±8.68%	18.8 ±6.90%
$z = 1 - 2$ ($N = 41$)	All	**34.1** ±7.41%	65.9 ±7.41%	43.9 ±7.45%
$z \sim 0$ ($N = 263$)	All	**12.2** ±2.02%	87.8 ±2.02%	77.2 ±2.59%
	E/S0	1.10 ±0.65%	69.2 ±2.85%	63.9 ±2.96%
	Sabc	9.51 ±1.81%	17.1 ±2.32%	12.5 ±2.04%
	Sd/Irr	1.52 ±0.75%	1.52 ±0.75%	0.76 ±0.54%
z	Morphology	$r_e \leq 2$ kpc	$2 < r_e \leq 4$ kpc	$r_e > 4$ kpc
		$M_\star \geq 5 \times 10^{10}\ M_\odot$		
$z = 2 - 3$ ($N = 77$)	All	**39.0** ±5.56%	42.9 ±5.64%	18.2 ±4.40%
$z = 1 - 2$ ($N = 89$)	All	**24.7** ±4.57%	48.3 ±5.30%	27.0 ±4.70%
$z \sim 0$ ($N = 740$)	All	**0.54** ±0.27%	7.8 ±0.99%	91.6 ±1.02%
	E/S0	0.41 ±0.23%	7.0 ±0.94%	50.8 ±1.84%
	Sabc	0.00 ±0.00%	0.8 ±0.33%	38.1 ±1.79%
	Sd/Irr	0.14 ±0.14%	0.00 ±0.00%	3.51 ±0.68%
		$M_\star \geq 1 \times 10^{11}\ M_\odot$		
$z = 2 - 3$ ($N = 32$)	All	**40.6** ±8.68%	43.8 ±8.77%	15.6 ±6.42%
$z = 1 - 2$ ($N = 41$)	All	**22.0** ±6.46%	56.1 ±7.75%	22.0 ±6.46%
$z \sim 0$ ($N = 263$)	All	**0.38** ±0.38%	1.90 ±0.84%	97.7 ±0.92%
	E/S0	0.00 ±0.00%	1.9 ±0.84%	68.4 ±2.87%
	Sabc	0.00 ±0.00%	0.0 ±0.00%	26.6 ±2.73%
	Sd/Irr	0.38 ±2.73%	0.00 ±0.00%	2.66 ±0.99%

Note. — Rows 1-12: For a given redshift (Column 1), morphology (Column 2), and stellar mass range, Columns 3, 4, and 5 list the fraction of galaxies in three separate bins of Sérsic index n. Rows 13-24: Same as the above except that Columns 3, 4, and 5 reflect bins of half-light radius r_e.

TABLE 2
r_e DISTRIBUTION OF n AMONG MASSIVE GALAXIES WITH $r_e \leq 2$ KPC AND $r_e > 2$ KPC

z	r_e (kpc)	No.	Fraction w/ $n < 2$	Fraction w/ $2 \leq n < 4$	Fraction w/ $n > 4$
			$M_\star \geq 5 \times 10^{10}\ M_\odot$		
$z = 2 - 3$ (GNS)	$r_e \leq 2$	30	46.7 ±9.11%	36.7 ±8.80%	16.7 ±6.80%
$z = 1 - 2$ (GNS)	$r_e \leq 2$	22	31.8 ±9.93%	45.5 ±10.62%	22.7 ±8.93%
$z \sim 0$ (MGC)	$r_e \leq 2$	4	25.0 ±21.65%	50.0 ±25.00%	25.0 ±21.65%
$z = 2 - 3$ (GNS)	$r_e > 2$	14	76.6 ±6.18%	21.3 ±5.97%	2.1 ±2.10%
$z = 1 - 2$ (GNS)	$r_e > 2$	67	55.2 ±6.08%	34.3 ±5.80%	7.5 ±3.21%
$z \sim 0$ (MGC)	$r_e > 2$	736	18.1 ±1.42%	38.5 ±1.79%	43.5 ±1.83%
			$M_\star \geq 1 \times 10^{11}\ M_\odot$		
$z = 2 - 3$ (GNS)	$r_e \leq 2$	13	46.2 ±13.83%	38.5 ±13.49%	15.4 ±10.01%
$z = 1 - 2$ (GNS)	$r_e \leq 2$	9	11.1 ±10.48%	66.7 ±15.71%	22.2 ±13.86%
$z \sim 0$ (MGC)	$r_e \leq 2$	1	100.0 ±0.00%	0.0 ±0.00%	0.0 ±0.00%
$z = 2 - 3$ (GNS)	$r_e > 2$	19	68.4 ±10.66%	31.6 ±10.66%	0.0 ±0.00%
$z = 1 - 2$ (GNS)	$r_e > 2$	32	40.6 ±8.68%	46.9 ±8.82%	12.5 ±5.85%
$z \sim 0$ (MGC)	$r_e > 2$	262	50.8 ±3.09%	31.7 ±2.87%	56.5 ±3.06%

Note. — For a given redshift (Column 1), half-light radius r_e (Column 2), and stellar mass range, Column 3 lists the total number of matching galaxies. Columns 4, 5, and 6 list the fraction of galaxies in separate bins of Sérsic index n.

TABLE 3
DISTRIBUTION OF r_e AMONG MASSIVE GALAXIES WITH $n \leq 2$ AND $n > 2$

z	n	No.	Fraction w/ $r_e < 2$ kpc	Fraction w/ $2 \leq r_e < 4$ kpc	Fraction w/ $r_e > 4$ kpc
			$M_\star \geq 5 \times 10^{10}\ M_\odot$		
$z = 2-3$ (GNS)	$n \leq 2$	50	28.0 ±6.35%	46.0 ±7.05%	22.0 ±5.86%
$z = 1-2$ (GNS)	$n \leq 2$	44	15.9 ±5.51%	43.2 ±7.47%	40.9 ±7.41%
$z \sim 0$ (MGC)	$n \leq 2$	134	0.7 ±0.74%	1.5 ±1.05%	97.8 ±1.28%
$z = 2-3$ (GNS)	$n > 2$	27	59.3 ±9.46%	37.0 ±9.29%	3.7 ±3.63%
$z = 1-2$ (GNS)	$n > 2$	45	33.3 ±7.03%	53.3 ±7.44%	13.3 ±5.07%
$z \sim 0$ (MGC)	$n > 2$	606	0.5 ±0.29%	9.2 ±1.18%	90.3 ±1.20%
			$M_\star \geq 1 \times 10^{11}\ M_\odot$		
$z = 2-3$ (GNS)	$n \leq 2$	19	31.6 ±10.66%	47.4 ±11.45%	21.1 ±9.35%
$z = 1-2$ (GNS)	$n \leq 2$	41	7.1 ±6.88%	57.1 ±13.23%	35.7 ±12.81%
$z \sim 0$ (MGC)	$n \leq 2$	32	3.1 ±3.08%	6.2 ±4.28%	90.6 ±5.15%
$z = 2-3$ (GNS)	$n > 2$	13	53.8 ±13.83%	38.5 ±13.49%	7.7 ±7.39%
$z = 1-2$ (GNS)	$n > 2$	27	29.6 ±8.79%	55.6 ±9.56%	14.8 ±6.84%
$z \sim 0$ (MGC)	$n > 2$	231	0.0 ±0.00%	1.3 ±0.74%	98.7 ±0.74%

Note. — For a given redshift (Column 1), Sérsic index n (Column 2), and stellar mass range, Column 3 lists the total number of matching galaxies. Columns 4, 5, and 6 list the fraction of galaxies in separate bins of half-light radius r_e.

TABLE 4
FRACTION OF MASSIVE ($M_\star \geq 5 \times 10^{10}\ M_\odot$) GALAXIES WITH 24 μM DETECTIONS

z	SFR_{min} ($M_\odot$ yr^{-1})	Fraction With $f_{24\mu m} > 30\mu$Jy (%)	Mean SFR (All) ($M_\odot$ yr^{-1})	Mean SFR (No AGN) ($M_\odot$ yr^{-1})
(1)	(2)	(3)	(4)	(5)
$z = 1 - 1.5$	4.29	43.6 $\pm$7.9%	63.8	44.0
$z = 1.5 - 2$	12.4	48.0 $\pm$7.1%	222.8	128.9
$z = 2 - 3$	47.2	42.9 $\pm$5.6%	1145.6	418.8

Note. — Column 2 estimates the detection limit on SFR given the 5σ limit on $f_{24\mu m}$ of 30 μJy. The expected $\mathrm{SFR_{IR}}$ at 30 μJy is determined by linear regression of the distribution of $f_{24\mu m}$ versus $\mathrm{SFR_{IR}}$ in each redshift bin. Column 3 lists the percentage of massive GNS galaxies with $f_{24\mu m} > 30\mu$Jy. Column 4 shows the mean SFR among all galaxies having $f_{24\mu m} > 30$ μJy. Column 5 shows the mean SFR among all galaxies without any evidence for AGN activity (see §6.5).

TABLE 5
SUMMARY OF AGN DETECTION AND MEAN PROPERTIES

z	X-ray AGN	IR Excess AGN	AGN Fraction	Mean n	Mean r_e (kpc)	Mean $U-V$ (Vega mag)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
$z = 1 - 1.5$	7	0	17.9 $\pm$14.5%	2.48	4.38	1.14
$z = 1.5 - 2$	6	5	22.0 $\pm$12.5%	1.89	3.98	1.72
$z = 2 - 3$	20	11	40.3 $\pm$8.81%	2.34	2.98	2.40

Note. — Columns 2 and 3 list the AGN selection criteria in order of precedence. Galaxies that are selected based on mid-IR excess are not X-ray detected. In §6.2, we state that three galaxies show mid-IR power-law emission. The power-law galaxies are not shown in a separate column because they are all classified as AGN based on their X-ray properties.

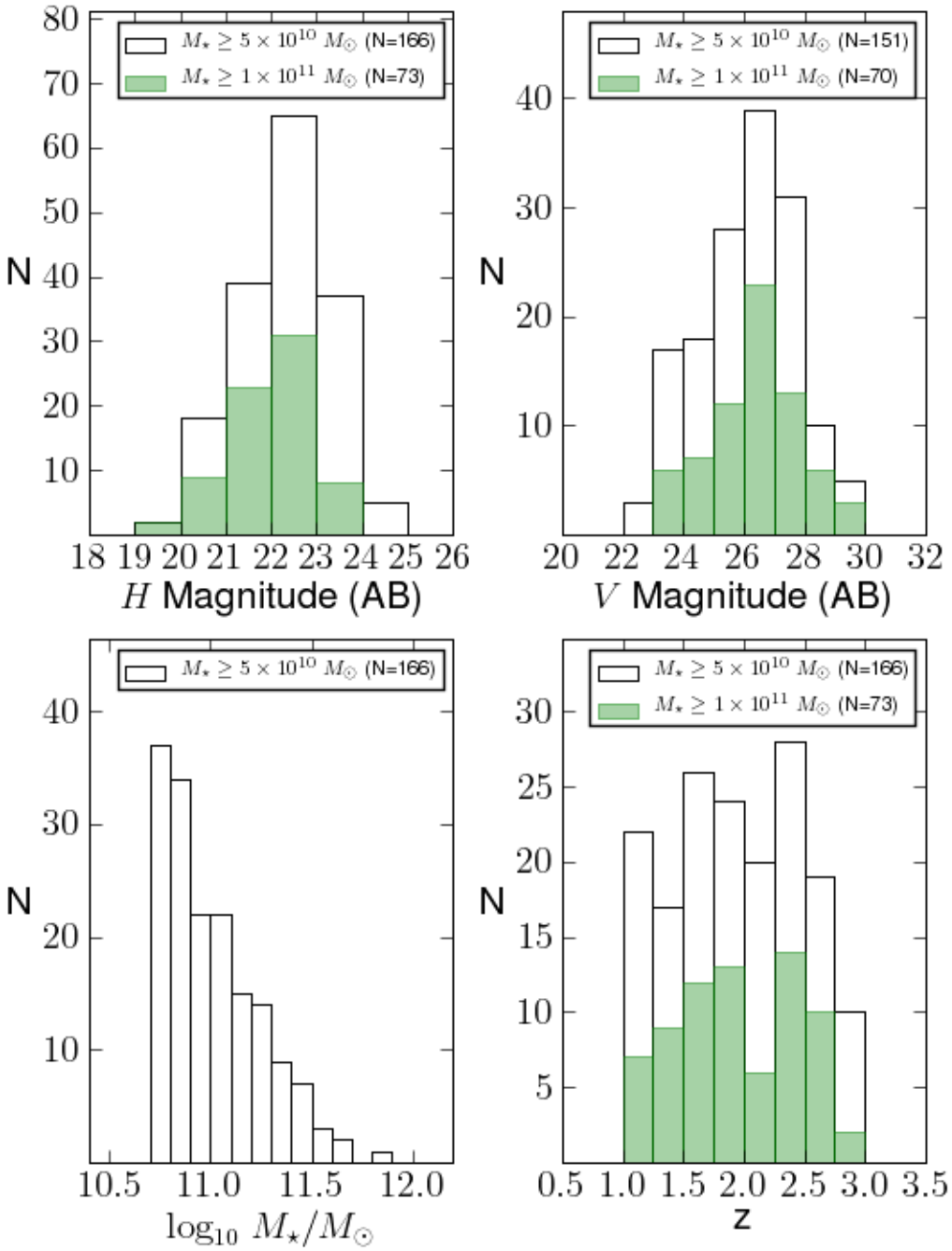

FIG. 1.— The distribution of apparent H (F160W), V apparent magnitude, stellar mass, and redshiftfor the final, complete sample of 166 galaxies with $M_\star \geq 5 \times 10^{10}\ M_\odot$ and redshift $z = 1 - 3$.

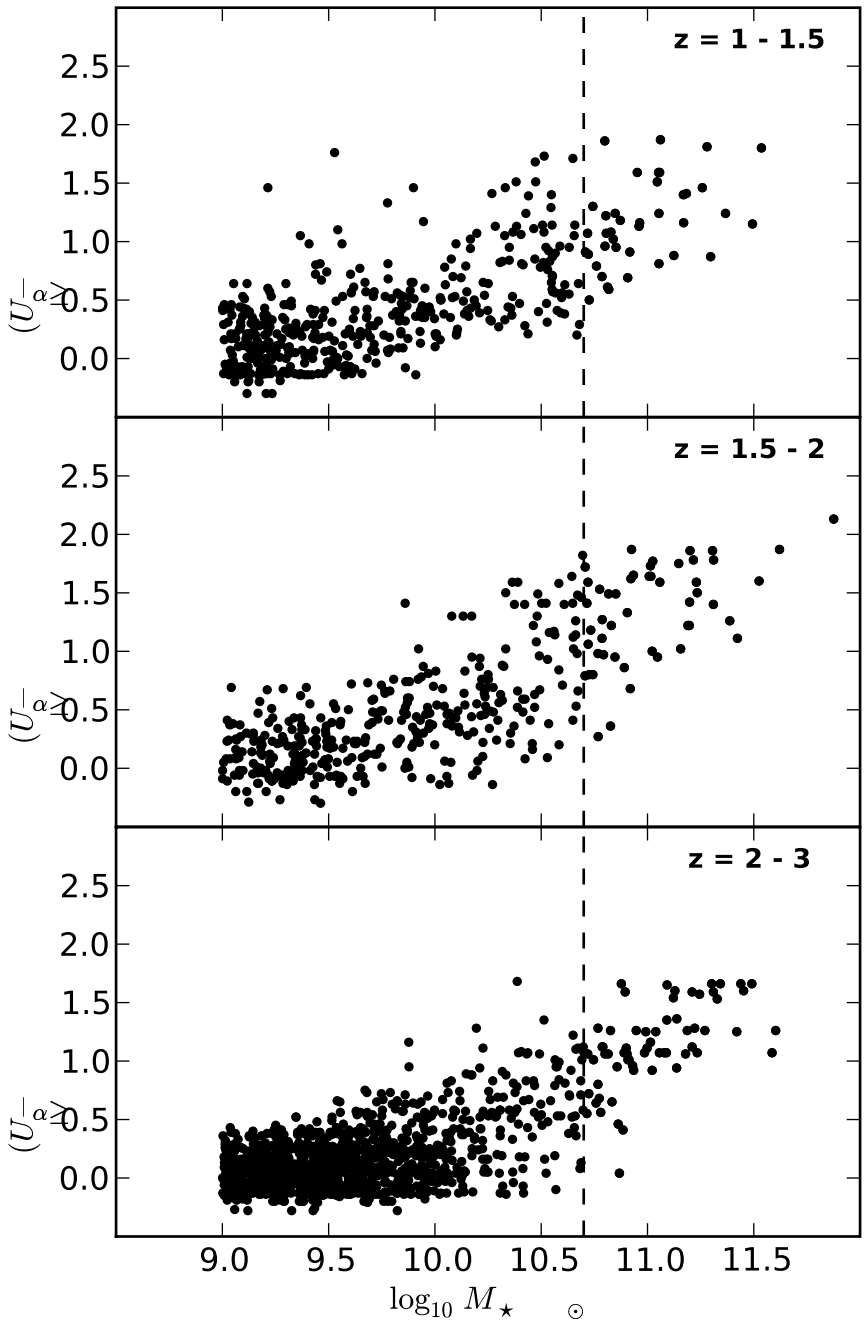

FIG. 2.— For all galaxies detected in the GOODS-NICMOS Survey (GNS) over $z = 1-3$, rest-frame $U-V$ is plotted against $M_\star$ for different redshift bins. Blue systems are preferentially at low masses, while the most massive ($M_\star \geq 1 \times 10^{11}\ M_\odot$) galaxies with are preferentially red. The vertical line denotes $M\star = 5 \times 10^{10}\ M_\odot$, the mass cut we adopt for our final sample of 166 galaxies.

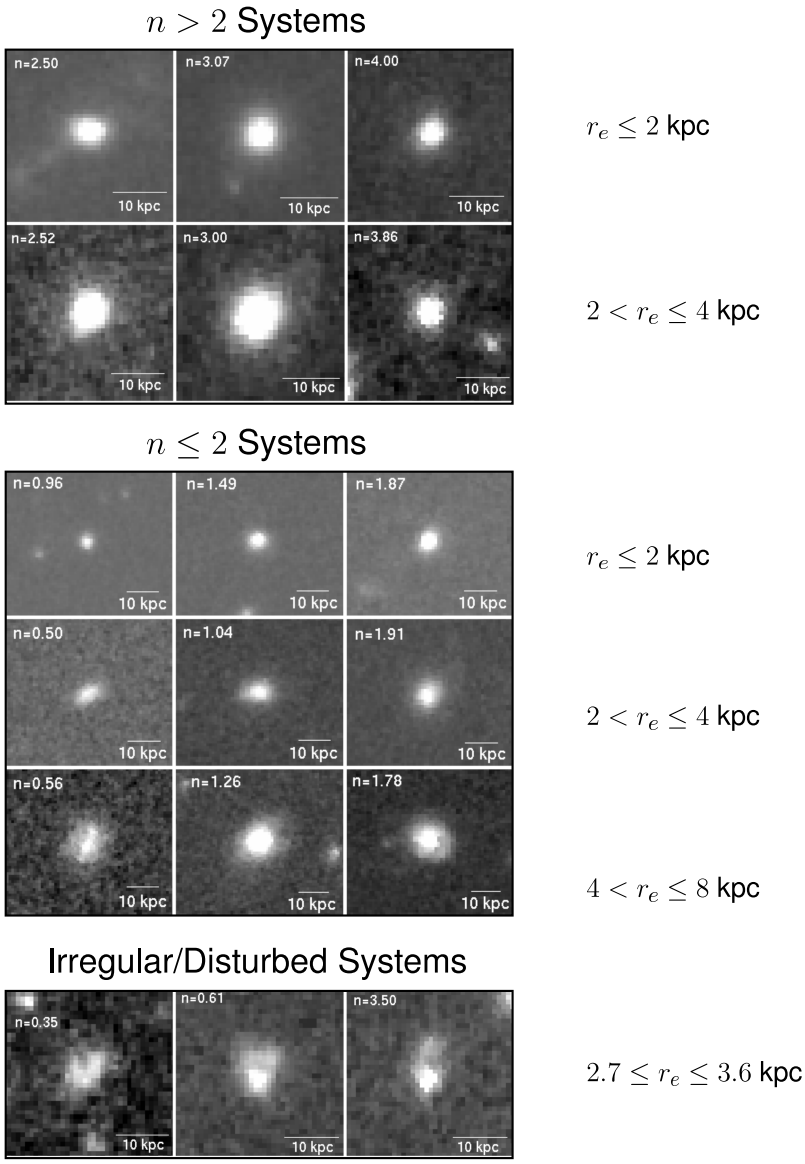

FIG. 3.— NICMOS F160W images for representative GNS galaxies with $M_\star \geq 5 \times 10^{10}\ M_\odot$ at $z = 2 - 3$. The Sérsic index n and effective radius r_e referenced here are based on fitting single Sérsic components to the NICMOS images, as described in § 3. The top panel shows example systems with Sérsic index $n > 2$ and half-light radii $r_e \leq 4$ kpc. The middle panel shows examples with $n \leq 2$ and $r_e \leq 8$ kpc. The majority (over 84%; Table 1) of the massive GNS galaxies have $r_e < 4$ kpc. In such systems, structural features are generally hard to discern due to resolution effects, so that systems appear fairly featureless (top 4 rows). In the small fraction of massive galaxies at $z \sim 2 - 3$ with large $r_e \geq 4$ kpc, one can discern some structural features such as an elongated bar-like feature or a combination of a central condensation surrounded by a more extended lower surface brightness component, reminiscent of a bulge and disk (row 5). The bottom panel (row 6) contains irregular systems that appear disturbed.

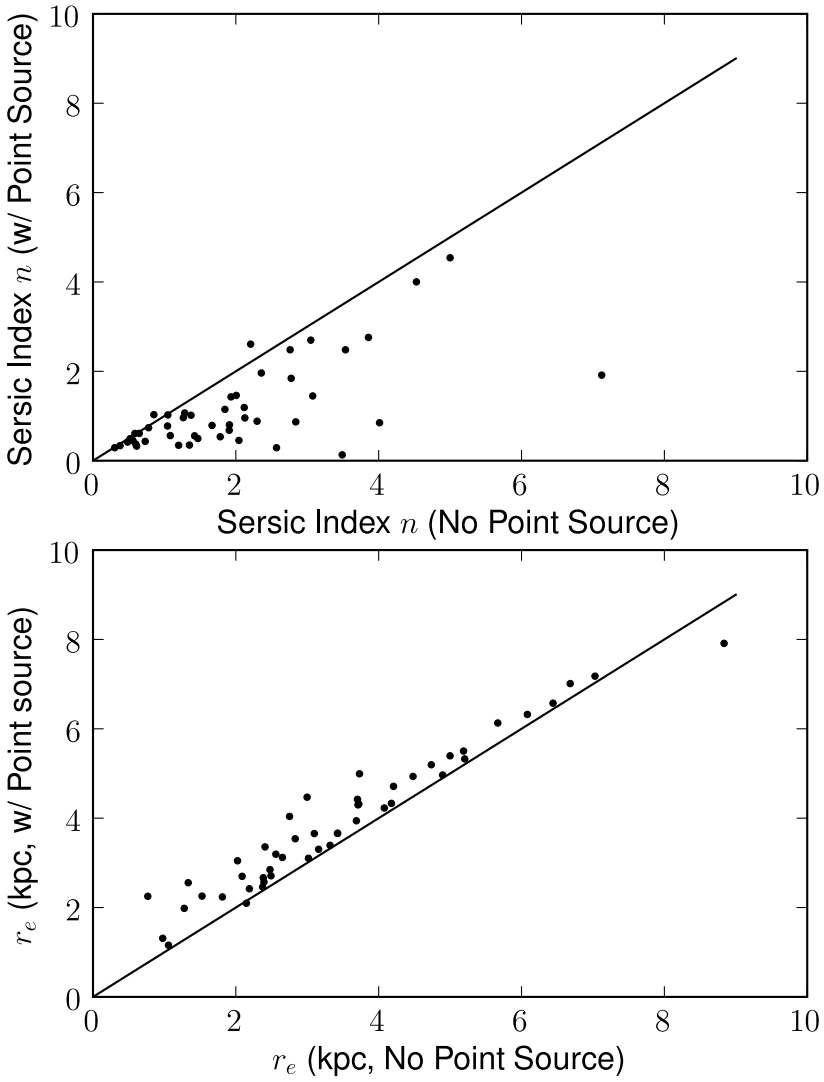

FIG. 4.— The structural parameters (n, r_e) of massive galaxies identified as AGN hosts are shown for a single Sérsic profile fit with and without the inclusion of a point source component. Including the point source produces generally small changes in (n, r_e) and goes in the direction of lowering n and enlarging r_e.

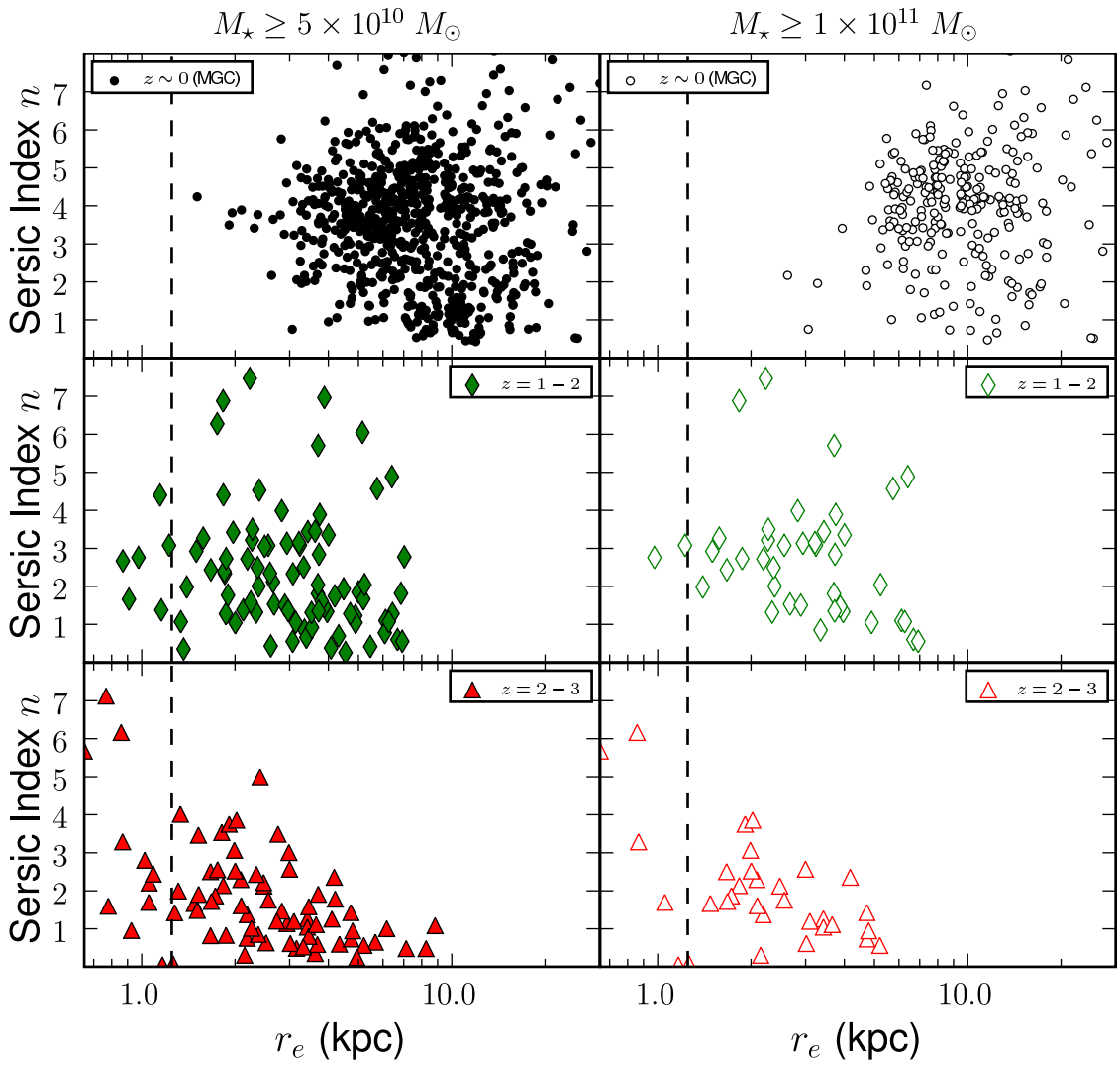

FIG. 5.— The B-band Sérsic index n and effective radius r_e derived from single Sérsic profile fits to massive ($M_\star \geq 5 \times 10^{10}\ M_\odot$) galaxies are plotted for the three redshift bins listed in Table 1. In the top row, the black points represent fits to $z \sim 0$ galaxies by Allen et al. (2006) on B-band images of galaxies from the Millennium Galaxy Catalog (Liske et al. 2003). The lower two rows are based on our fits to the NIC3 F160W images of massive GNS galaxies at $z = 1 - 2$ and $z = 2 - 3$. Note that the massive galaxies at $z \sim 2 - 3$ are strikingly offset toward lower (n, r_e) compared to the massive ~ 0 galaxies, and have 3 times more low $n \leq 2$ disky systems (see also Figure 6). The black dashed line represents the typical half-width half maximum of the NICMOS3 PSF at $z = 1 - 3$ of ~ 1.2 kpc.

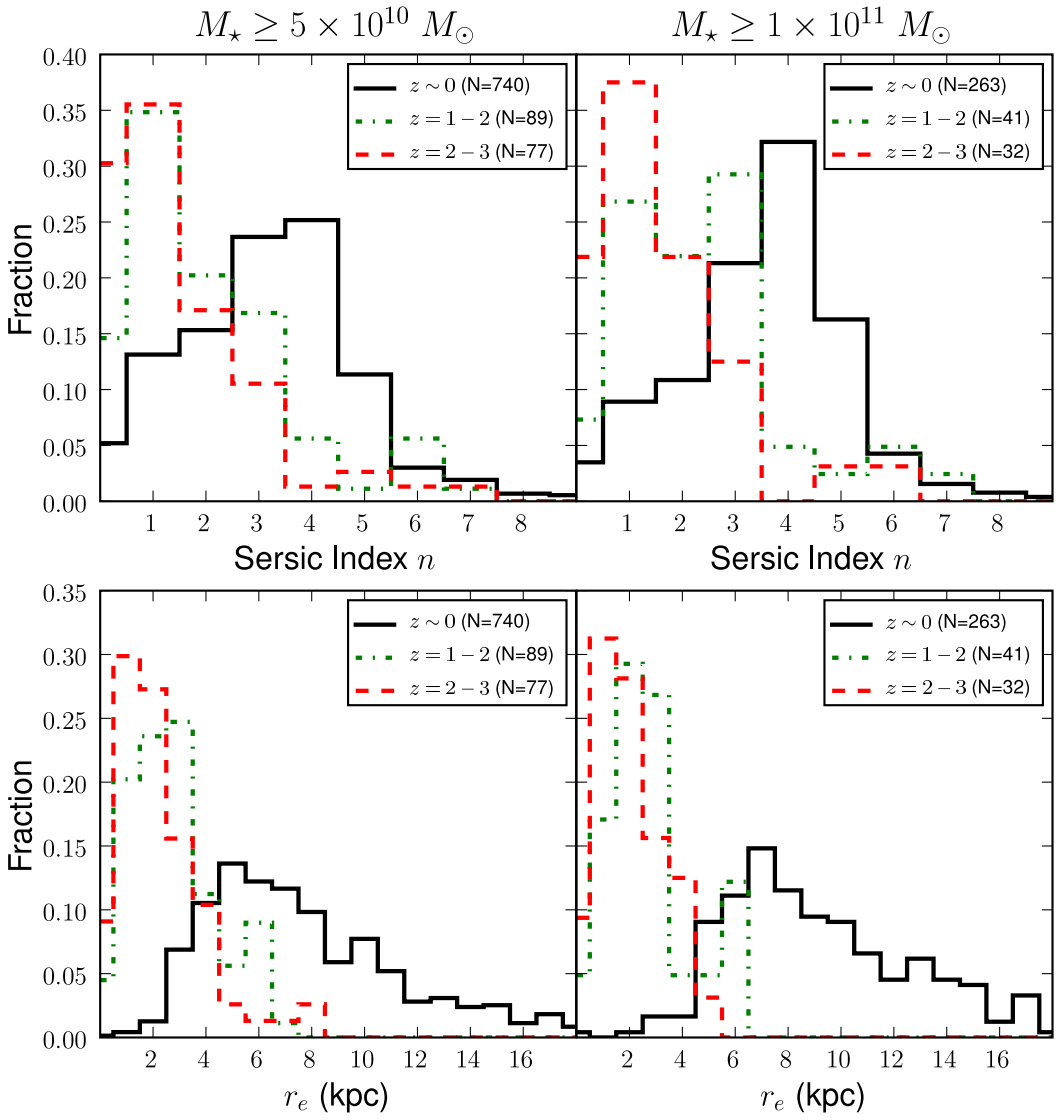

FIG. 6.— Left column: The distributions of rest-frame optical Sérsic index and effective radius r_e based on single Sérsic profile fits to massive ($M_\star \geq 5 \times 10^{10}\ M_\odot$) galaxies are plotted for the three redshift bins listed in Table 1: at $z \sim 0$ (solid line), based on the fits of Allen et al. (2006) on B-band images of galaxies from the Millennium Galaxy Catalog (Liske et al. 2003), and at $z = 1-2$ (dash-dotted line) and $z = 2-3$ (dashed line), based on our fits to the NIC3 F160W images of massive GNS galaxies. Note also that a significant fraction ($39.0 \pm 5.56\%$) of massive ($M_\star \geq 5 \times 10^{10}\ M_\odot$) galaxies at $z = 2-3$ have $r_e \leq 2$ kpc, compared to only $0.54 \pm 0.27\%$ at $z \sim 0$. Right column: Same as left column but for the mass range $M_\star \geq 1 \times 10^{11}\ M_\odot$.

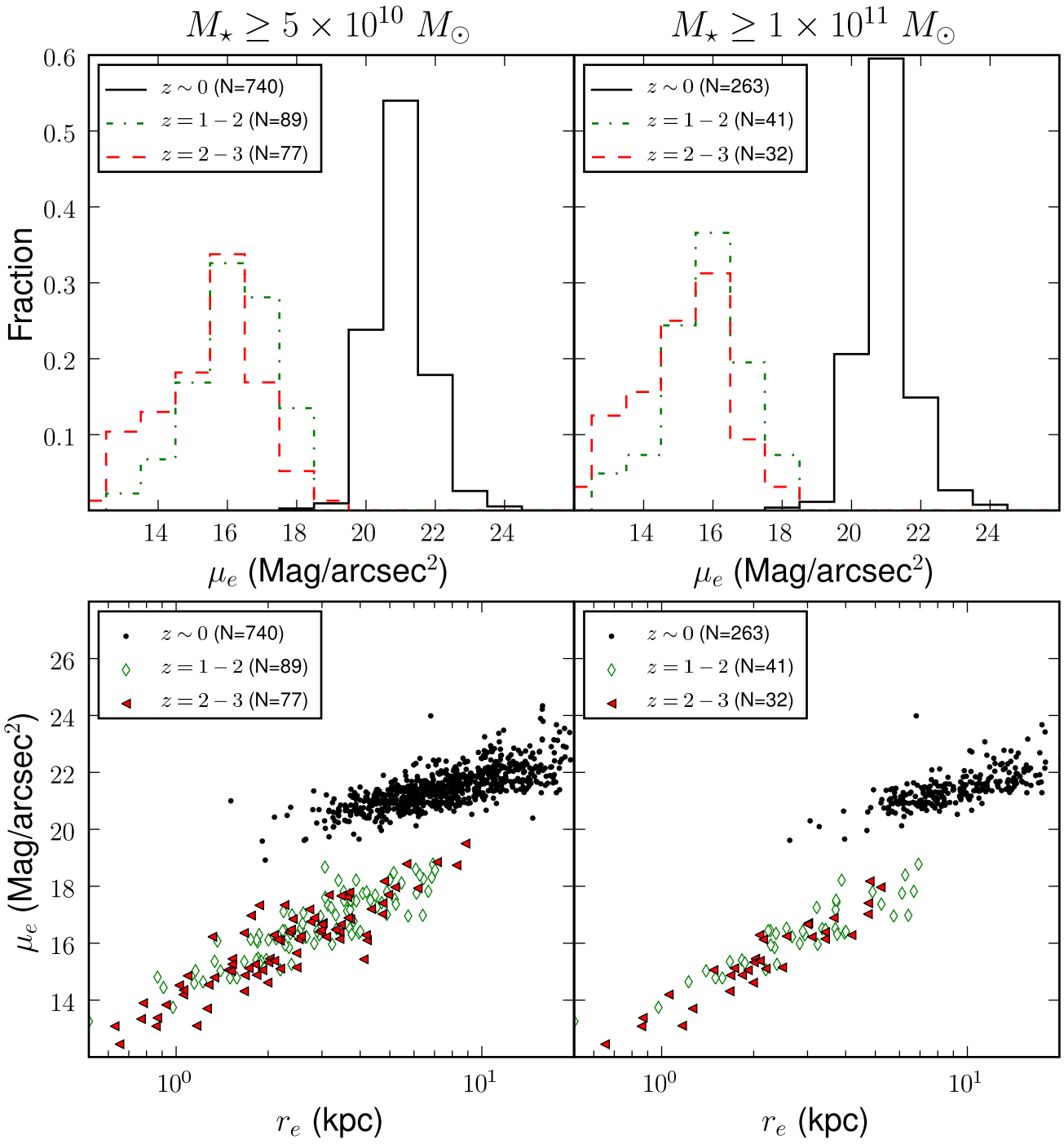

FIG. 7.— Left column: The top panel shows rest-frame B-band surface brightness at the effective radius (μ_e) for massive ($M_\star \geq 5 \times 10^{10}$ $M_\odot$) galaxies for the three redshift bins listed in Table 1. The solid line is for $z \sim 0$ MGC galaxies. The dash-dotted line ($z = 1-2$) and the dashed line ($z = 2-3$) are based on our fits to the NIC3 F160W images of massive GNS galaxies. The GNS galaxies at $z = 2-3$ have a mean surface brightness of 16.0 mag/arcsec^2 and are systematically brighter than the $z \sim 0$ MGC galaxies, which have a mean surface brightness of 21.3 mag/arcsec^2. In the bottom panel, surface brightness at the effective radius is plotted against effective radius r_e for the same redshift bins. Right column: The same plots are repeated for galaxies with $M_\star \geq 1 \times 10^{11}$ $M_\odot$. Surface brightness is defined as $\mu_e = \mathrm{mag} + \mathrm{zeropoint} + 2.5\log_{10}(\pi r_e^2) - 10\log_{10}(1+z)$. Surface brightness was measured with B-band for $z \sim 0$ and H-band for $z = 2-3$ so that approximately the same rest-frame light is considered at both redshifts.

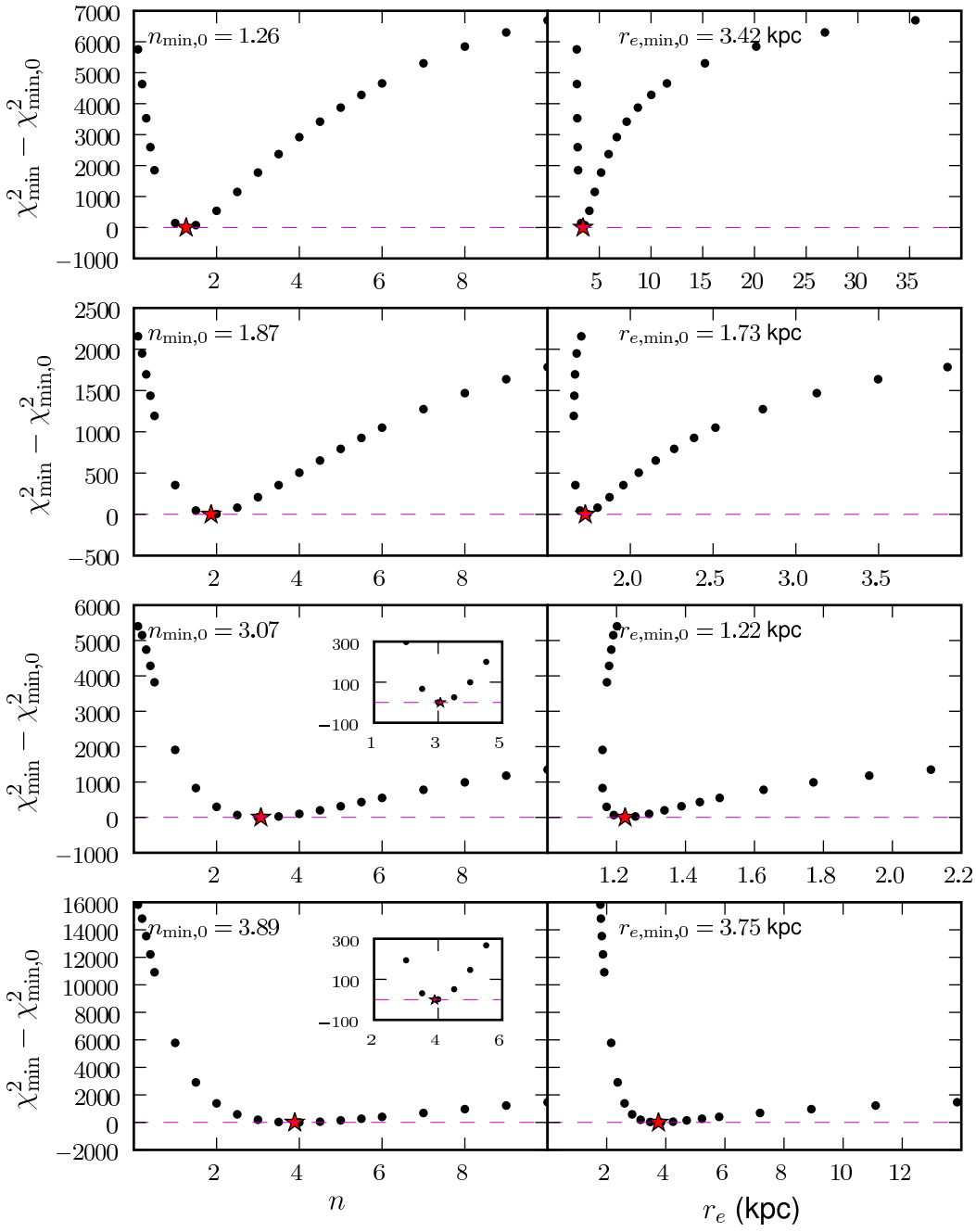

FIG. 8.— For the galaxies pictured in Figure 9, the first and second columns show the difference $\chi^2_{\rm min} - \chi^2_{\rm min,0}$ versus n and r_e, respectively. $\chi^2_{\rm min,0}$ is the minimum χ^2 obtained when all parameters are freely fit, and $\chi^2_{\rm min}$ is the minimum χ^2 when n is held at discrete values (0.5-10). The r_e in the second column are the best-fit results for a given n and $\chi^2_{\rm min}$. The red stars mark the best-fit $n_{\rm min,0}$ and $r_{e,\rm min,0}$ corresponding to $\chi^2_{\rm min,0}$. The insets in rows 3 and 4 of column 1 show a magnified view around the minimum in $\chi^2_{\rm min} - \chi^2_{\rm min,0}$. Note that for galaxies with $n_{\rm min,0} < 2$ (rows 1 and 2), $\chi^2_{\rm min}$ - $\chi^2_{\rm min,0}$ rises sharply at higher $n > n_{\rm min,0}$, thereby making it unlikely that a higher $n > 2$ would provide a similarly good fit.

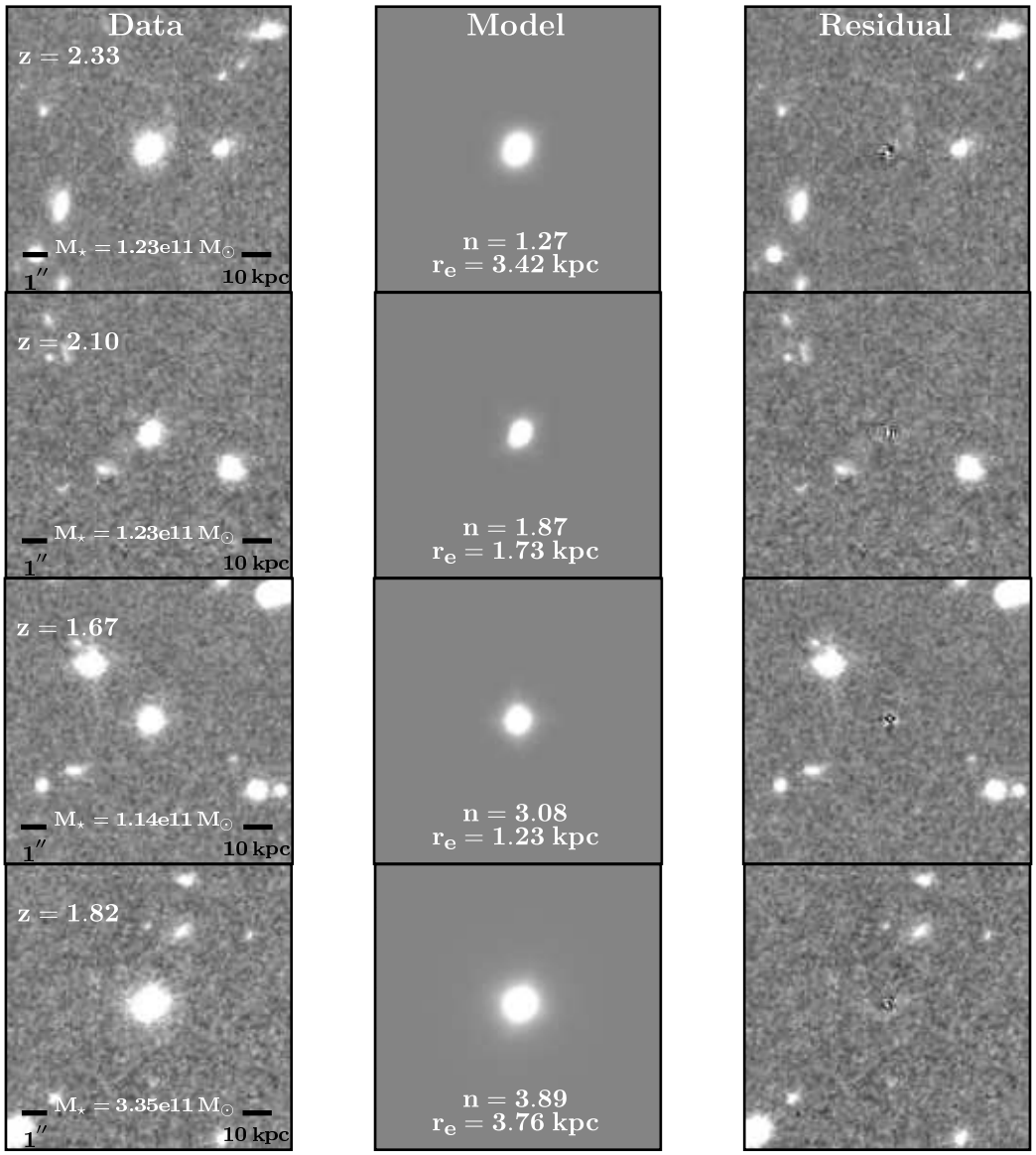

FIG. 9.— The first column shows NICMOS F160W images for representative galaxies with $M_\star \geq 1 \times 10^{11}\ M_\odot$ and Sérsic index $n \sim 1-4$ over $z = 1-3$. The first column shows the redshift of each source and the lengths corresponding to $1''$ and 10 kpc at the redshift of the source. The second column contains the single Sérsic profiles created by GALFIT, with the best-fit n and r_e indicated. The third column contains the residual images.

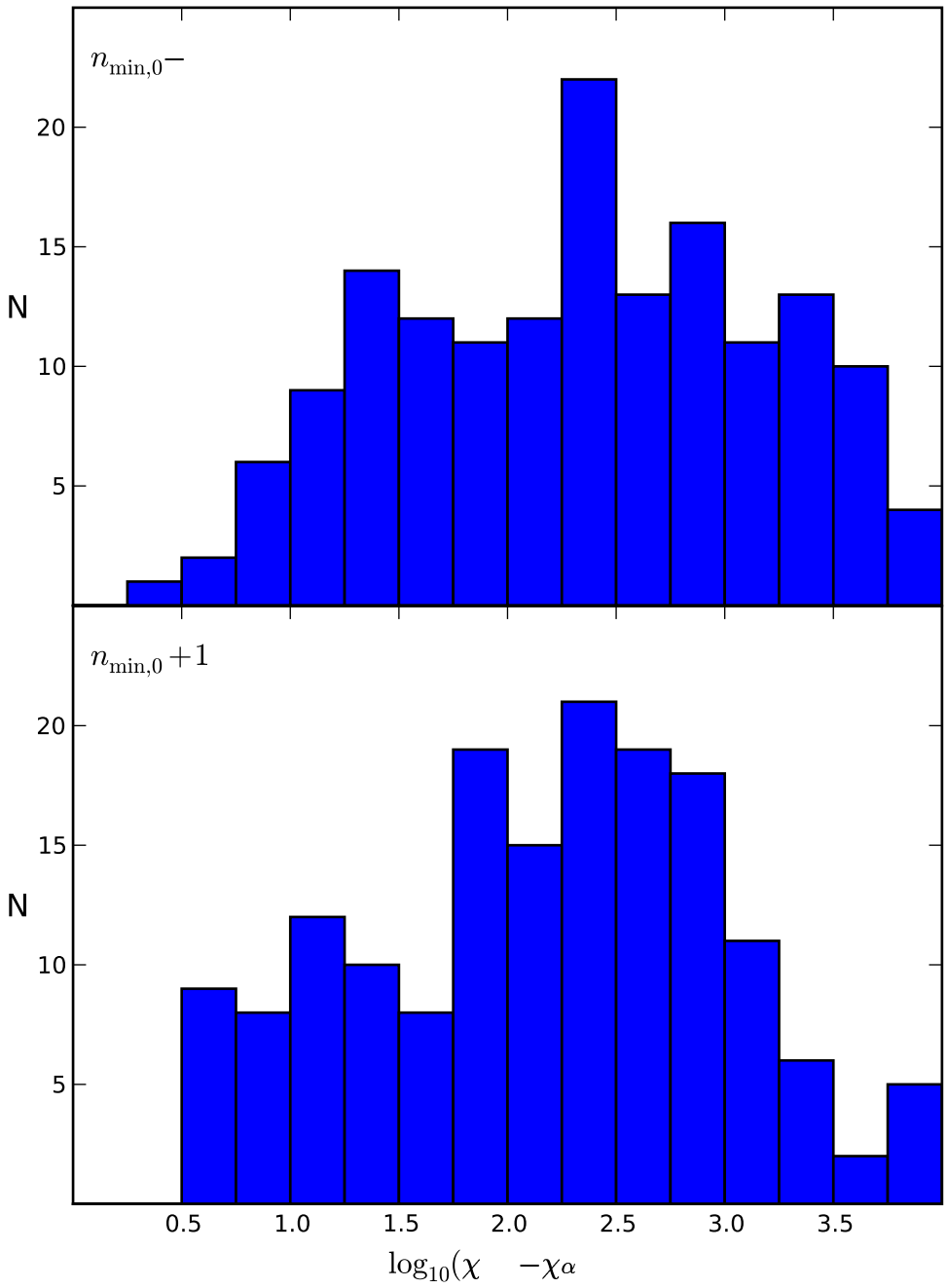

FIG. 10.— The quantity $\chi^2_{\rm min} - \chi^2_{\rm min,0}$ from Figure 8 is shown for all massive GNS galaxies well fitted with a single Sérsic profile. The top panel evaluates $\chi^2_{\rm min} - \chi^2_{\rm min,0}$ at $n_{\rm min,0} - 1$, and the bottom panel evaluates $\chi^2_{\rm min} - \chi^2_{\rm min,0}$ at $n_{\rm min,0} + 1$, where $n_{\rm min,0}$ is the best-fit Sérsic index corresponding to $\chi^2_{\rm min,0}$.

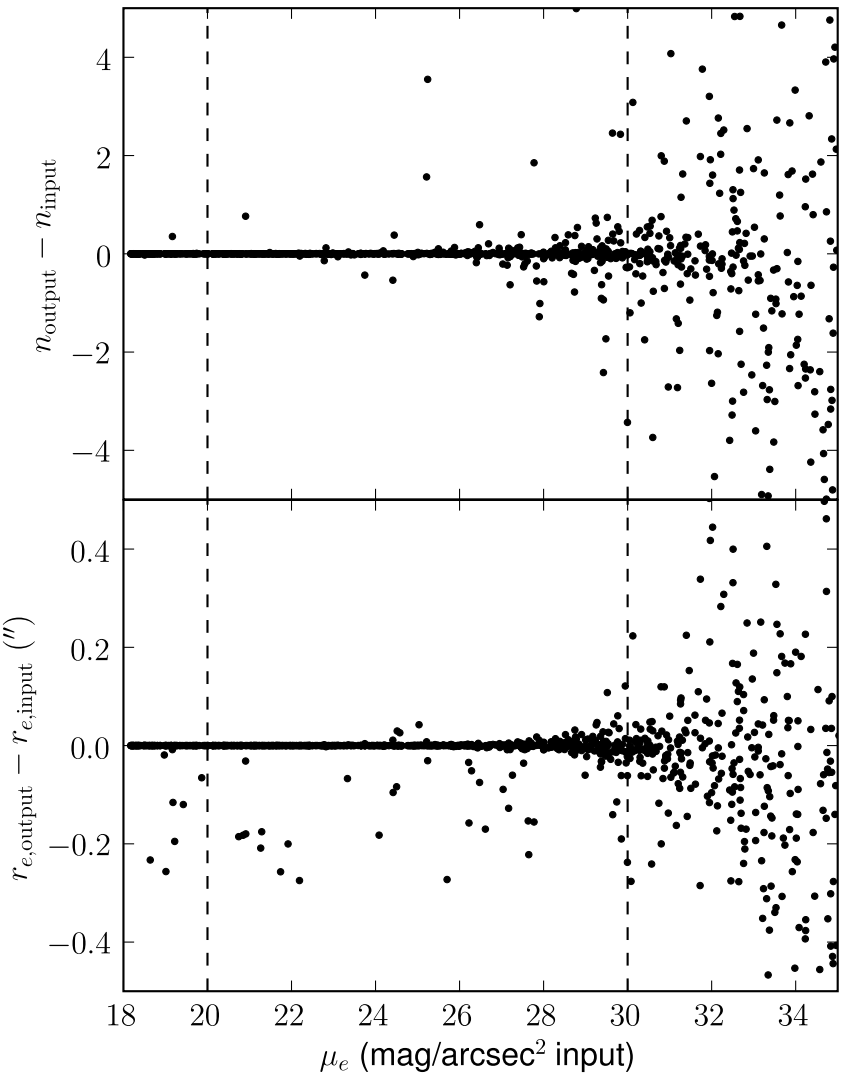

FIG. 11.— For the simulations described in §3.4.2, the difference between input and output Sérsic index n and effective radius r_e are plotted against effective surface brightness μ_e, the surface brightness at r_e. The vertical lines correspond to the range in μ_e in the NIC3/F160W band for the massive galaxies at $z \sim$1-3 in our sample.

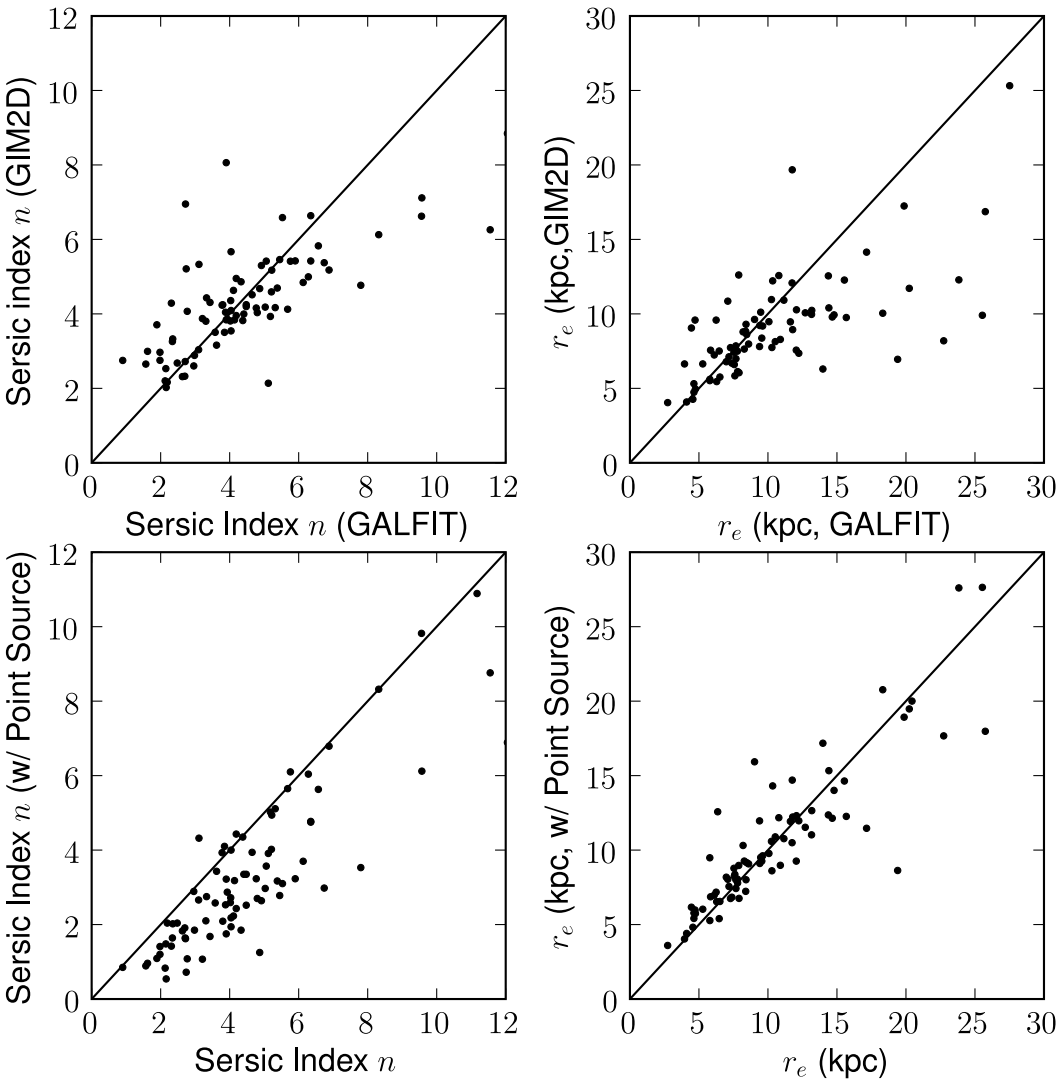

FIG. 12.— Top row: We demonstrate for a subset of $z \sim 0$ galaxies in the MGC catalog that the GIM2D-based (n, r_e) values from Allen et al. (2006) are not biased to higher values compared to our GALFIT-based fits for the same galaxies. All fits are performed on the B-band images from MGC. Bottom row: We show the effects of adding a point source in the GALFIT models fitted to the $z \sim 0$ MGC galaxies in Fig. 12. The values obtained using a model made of a Sérsic component plus a point source are plotted along the y-axis, while the x-axis shows the values obtained with a single Sérsic component. The values of r_e are not changed systematically. The Sérsic index is lowered by the addition of the point source, but only 20% of sources with $n > 2$ in the single Sérsic fit have $n \leq 2$ after including the point source.

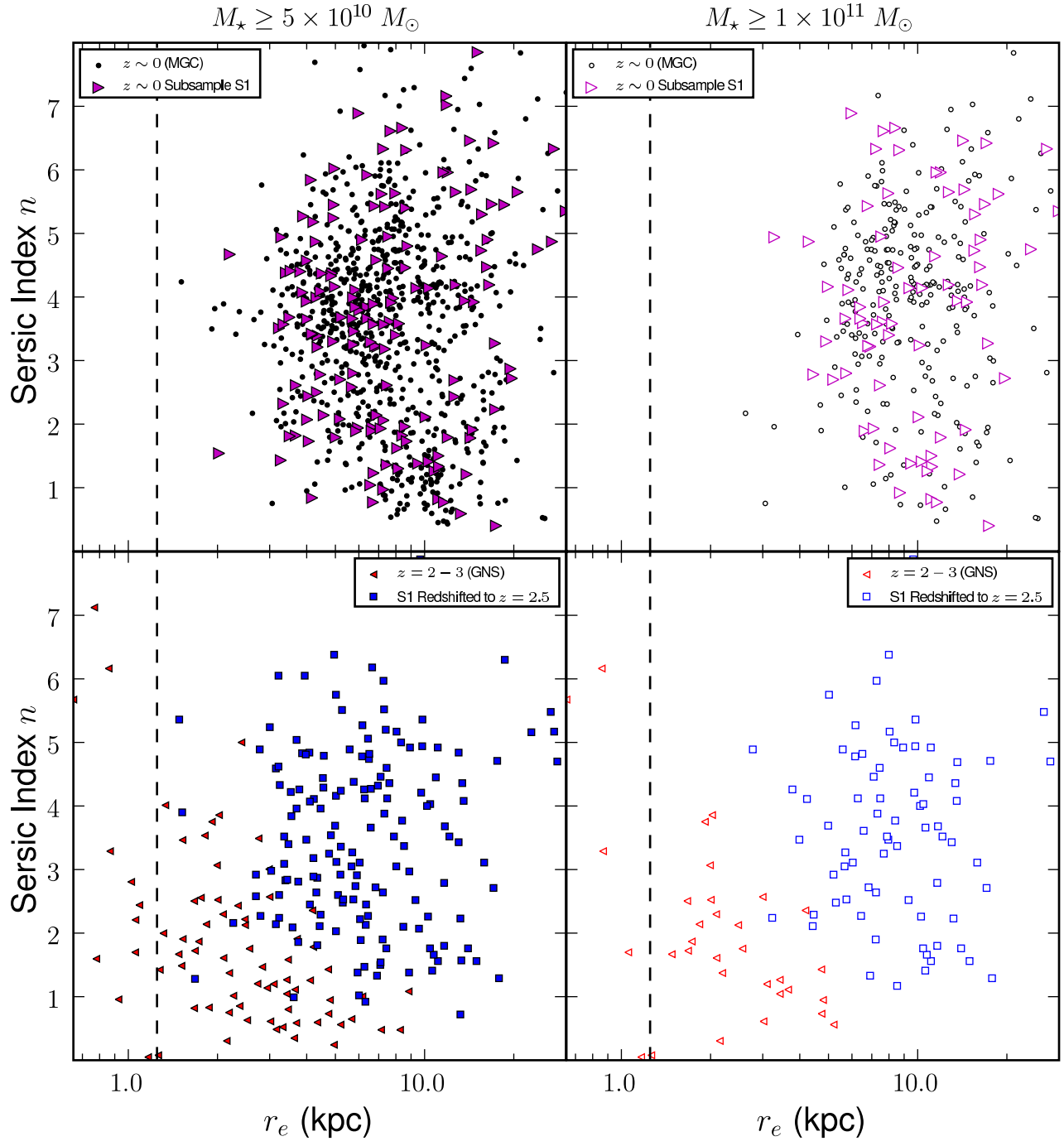

FIG. 13.— Top row: The black points show the MGC sample of 740 massive $z \sim 0$ galaxies, described earlier in Fig. 5. The magenta points denote the representative subsample S1 of 152 massive $z \sim 0$ galaxies used in the redshifting experiment. Note that the (n, r_e) distribution of S1 is representative of that of the MGC sample. This is also shown quantitatively in Figure 14. Row 2: We show as blue squares the (n, r_e) distribution obtained after redshifting subsample S1 to $z \sim 2.5$ and 're-observing' it with NIC3/F16W as in the GNS survey. We assume a surface brightness evolution of 2.5 magnitudes and brighten each redshifted galaxy by this amount. The actual observed (n, r_e) distributions of the massive galaxies at $z = 2 - 3$ in the GNS survey are significantly offset toward lower values compared to the redshifted galaxies. The black dashed line represents the typical Half-Width Half Max of the NICMOS3 PSF at $z = 1 - 3$ of ~ 1.2 kpc.

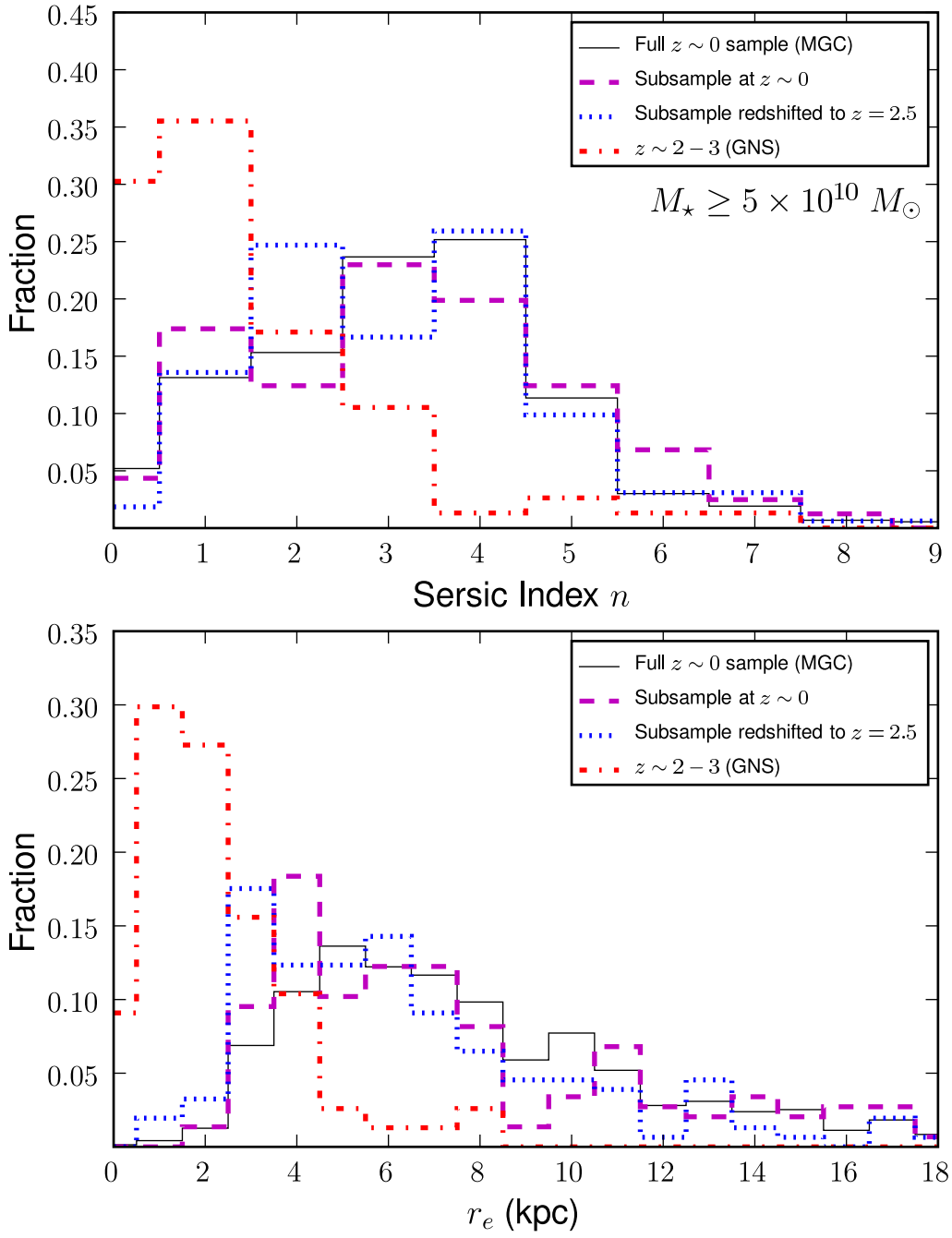

FIG. 14.— This figure illustrates the same information as in Fig. 13 but in more quantitative terms. It shows that the n and r_e distributions for the full MGC sample of 740 massive $z \sim 0$ galaxies (black line), and for the representative subsample S1 of 152 MGC galaxies used in the redshifting experiment (magenta line). Note that there the subsample S1 does a good job of matching the full MGC sample, typically within $\pm 10\%$ in a given bin. We also contrast the $(n,\ r_e)$ values after redshfiting S1 to $z \sim 2.5$, brightening by 2.5 magnitudes in surface brightness, and 're-observing' it with NGC3/F160W as in the GNS survey (blue line) with the actual distribution observed in the massive the GNS galaxies at $z = 2-3$ (red line): they are significantly offset from each other. The results shown here are for galaxies with $M_\star \geq 5 \times 10^{10}$ $M_\odot$, but a similar result is obtained for $M_\star \geq 1 \times 10^{11}$ $M_\odot$.

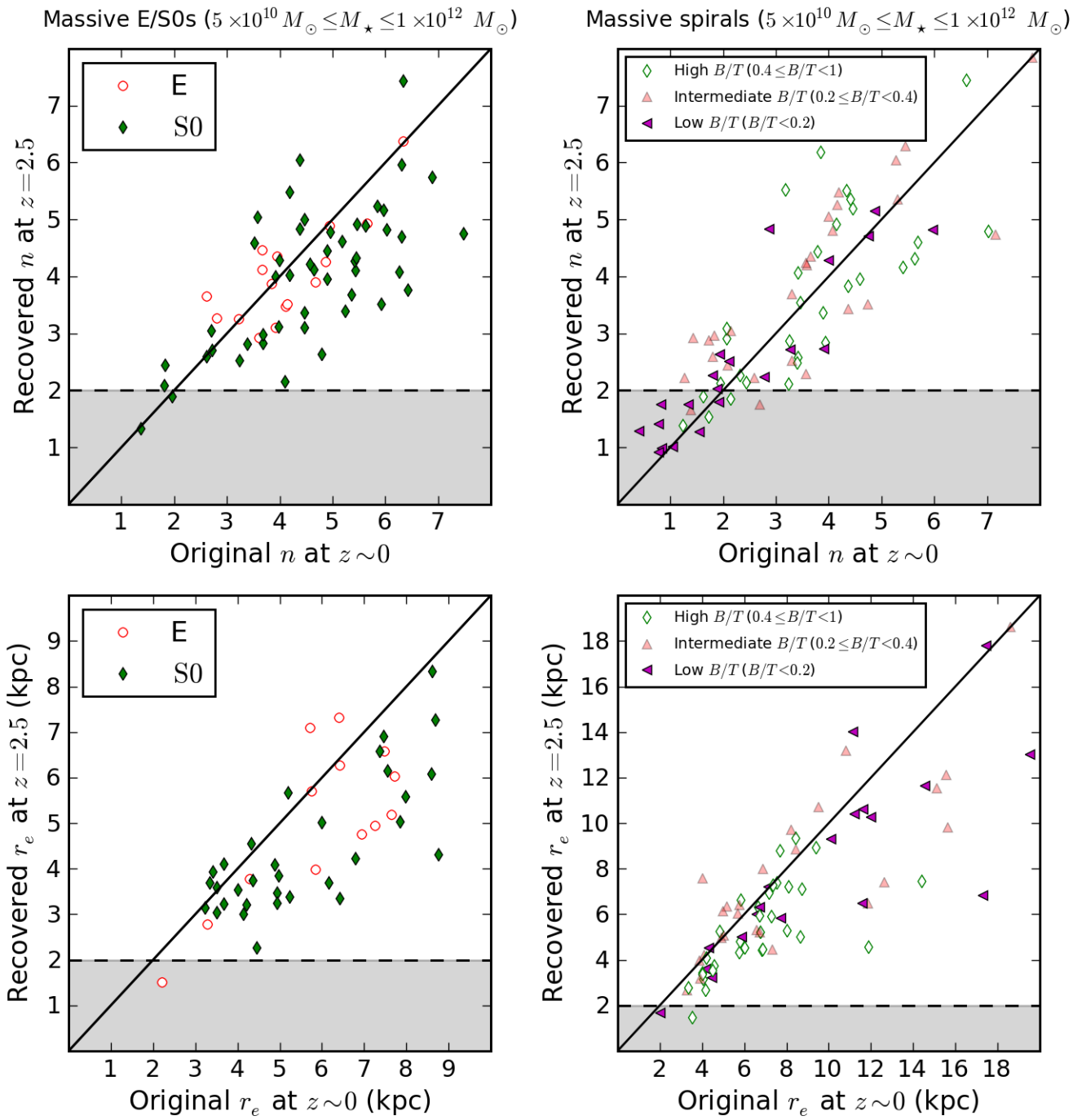

FIG. 15.— Left column: The panels compare the rest-frame optical structural parameters (Sérsic index n and effective radius r_e) of massive ($M_\star \geq 5\times10^{10}\ M_\odot$) elliptical and S0 galaxies at $z\sim0$ to the structural parameters recovered after these galaxies were artificially redshifted to $z=2.5$, brightened by 2.5 magnitudes in surface brightness, and re-observed with NIC3/F160W. At $z\sim0$, the structural parameters were measured from g-band images, while at $z=2.5$ they are measured from the artificially redshifted images in the NIC3/F160W band, so that all parameters are measured in the rest-frame blue optical light. The black lines represent equality, while the shaded area represents the regime of $n\leq2$ and $r_e\leq2$ kpc, where $64.9\pm5.4\%$ and $39.0\pm5.6\%$, respectively, of massive GNS galaxies at $z=2-3$ lie (Table 1 and Fig. 13). The plots show that the Sérsic index n and effective radius r_e of the massive $z\sim0$ E and S0s may be lower or higher after redshifting out to $z=2.5$, but they do not, in general, drop to values as low as $n\leq2$ and $r_e\leq2$ kpc, and avoid the shaded area. Right column: Same as left column, but this time for massive ($M_\star \geq 5\times10^{10}\ M_\odot$) $z\sim0$ spiral galaxies. The galaxies are coded by bulge-to-total light ratio (B/T). B/T was measured with bulge-disk and bulge-disk-bar decomposition of the $z\sim0$ g-band images. The top plot shows that it is mainly massive $z\sim0$ late-type spirals of low B/T which yield Sérsic index n as low as $n\leq2$ after redshifting, and populate the shaded area where $64.9\pm5.4\%$ of massive GNS galaxies at $z=2-3$ lie. The lower plot shows that massive $z\sim0$ spirals do not tend to have $r_e\leq2$ kpc after artificial redshfiting and avoid the shaded area where $39.0\pm5.6\%$ of the massive GNS galaxies at $z=2-3$ lie.

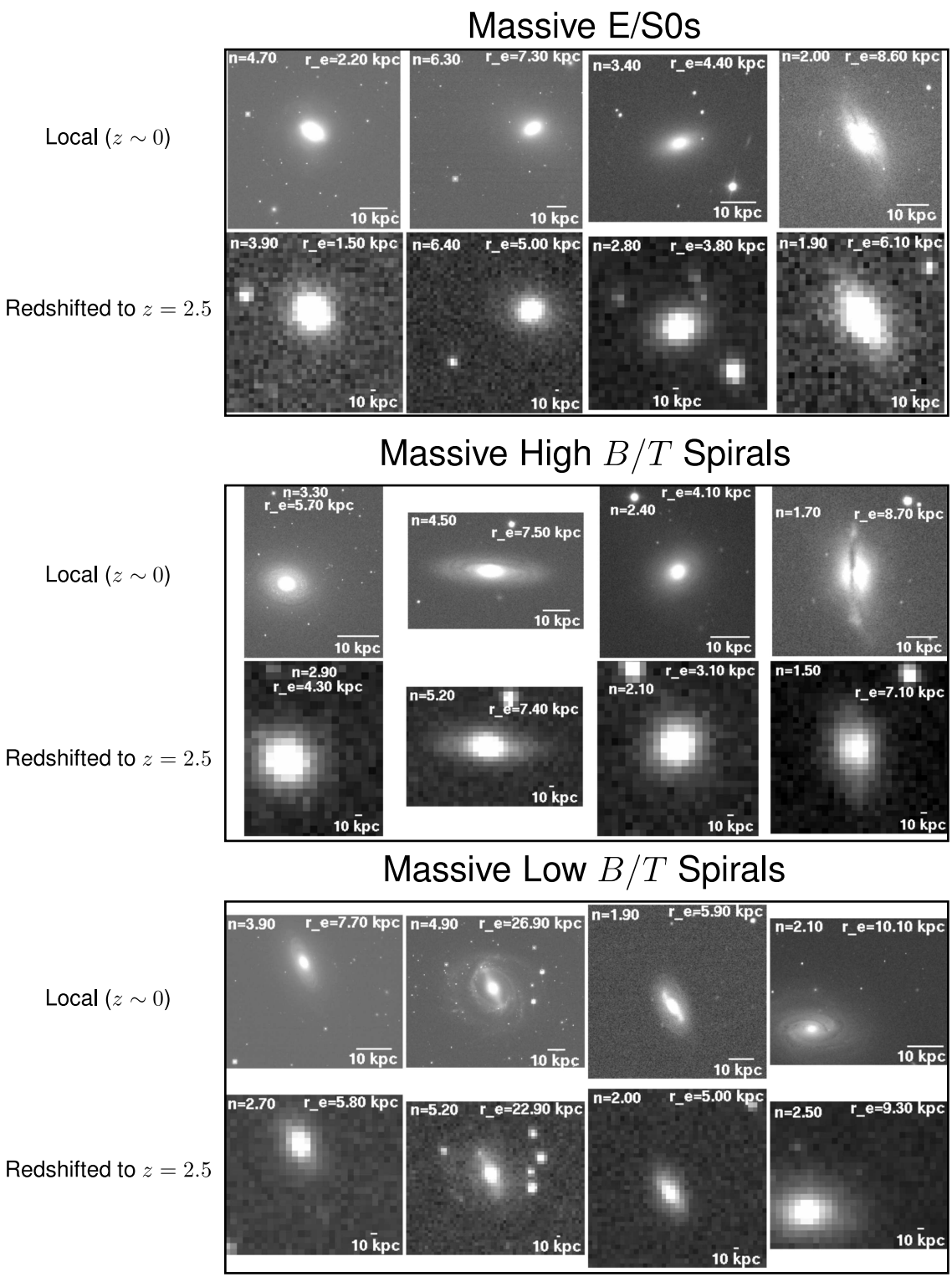

FIG. 16.— The top row of each panel shows g-band SDSS images of nearby massive ($M_\star \geq 5 \times 10^{10}\ M_\odot$) E/S0 or spiral galaxies of various bulge-to-total light ratios (B/T). The bottom row of each panel shows how the galaxy looks after being redshifted to $z = 2.5$, re-observed in F160W with NIC3 to the same depth as the GNS observations, and distorted with noise. We assume a surface brightness evolution of 2.5 magnitudes and brighten each redshifted galaxy by this amount. The images for a given object show the same field. Printed on all plots are the best-fit Sérsic index n and half-light radius r_e from a Sérsic profile plus point source or single Sérsic profile fit. The stars visible in all images are masked during the fitting and do not influence the derived structural parameters.

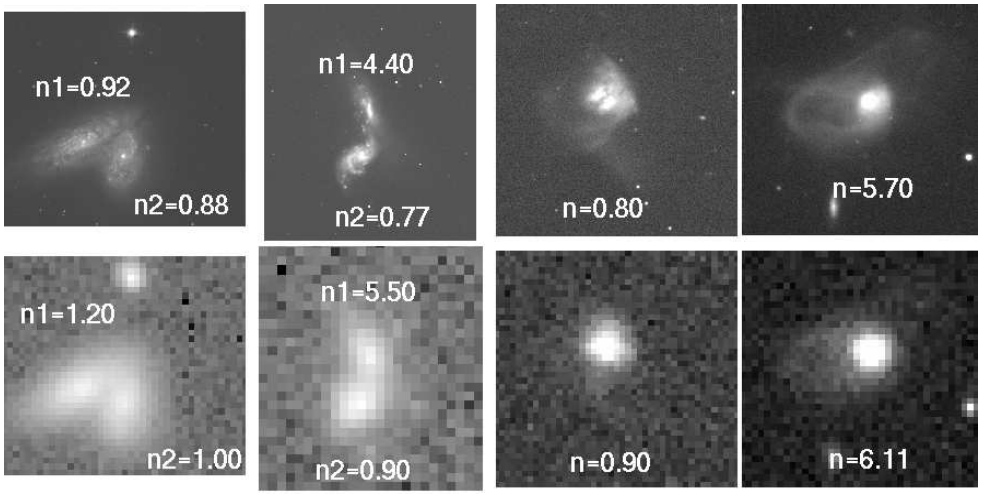

FIG. 17.— The top row of each panel shows g-band SDSS images of nearby merging systems, in different phases of a merger. From left to right the systems are NGC 4568, NGC 3396, Arp 220, and NGC 3921. The range in stellar mass is $\sim 1 \times 10^{10} \leq M_{\star} \leq 2 \times 10^{11}\ M_{\odot}$. The bottom row of each panel shows how the galaxy looks after being redshifted to $z = 2.5$, re-observed in F160W with NIC3 to the same depth as the GNS observations, and distorted with noise. We assume a surface brightness evolution of 2.5 magnitudes and brighten each redshifted source by this amount. The images for a given object show the same field. Printed on all plots are the best-fit Sérsic index n from a single Sérsic profile fit. The stars visible in all images are masked during the fitting and do not influence the derived structural parameters.

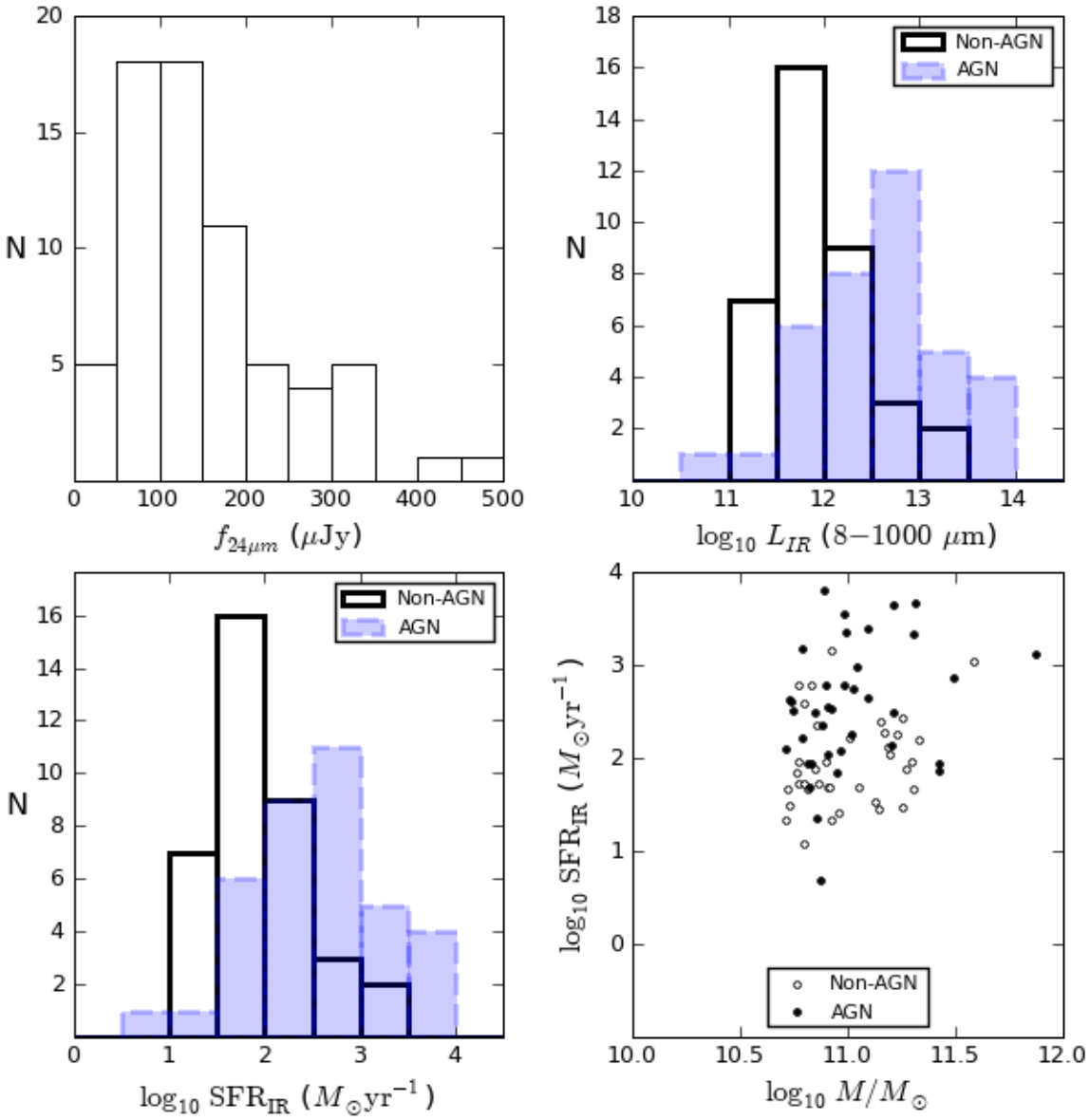

FIG. 18.— Top left: The $f_{24\mu m}$ distribution for the massive ($M_\star \geq 5 \times 10^{10}$ $M_\odot$) GNS galaxies with reliable MIPS 24 μm counterpart. Upper right: The inferred L_{IR} distribution over 8–1000 μm. Lower left: The inferred SFR$_{\rm IR}$ distribution based on L_{IR}, which is estimated using the Chary & Elbaz (2001) templates, with a correction at $L_{IR} > 6 \times 10^{11}$ $L_\odot$. Lower right: SFR$_{\rm IR}$ versus $M_\star$. For sources containing an AGN, the measured L_{IR} and SFR$_{\rm IR}$ are upper limits. The upper right and bottom panels use different coding for sources identified in §6 as hosting an AGN.

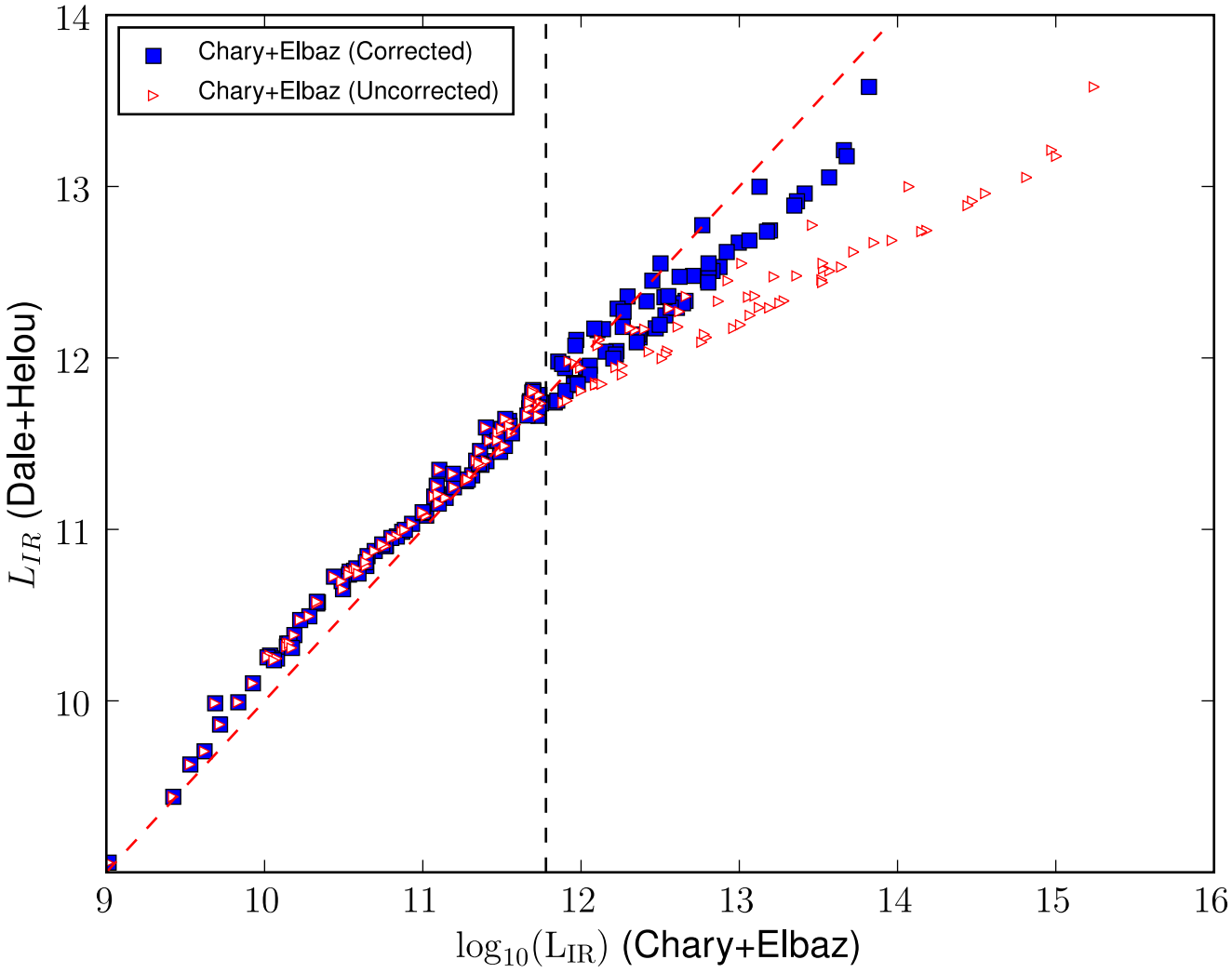

FIG. 19.— The L_{IR} obtained from 24 μm flux density using the Dale & Helou (2002) templates is compared to the Chary & Elbaz (2001) templates. The two sets of values shown for the Chary & Elbaz (2001) templates correspond to L_{IR} obtained before and after a correction is applied at $L_{IR} > 6 \times 10^{11}$ $L_{\odot}$. The correction takes into account the fact that the templates used are based on local galaxies with smaller PAH equivalent widths than galaxies of similar luminosity at high-redshift.

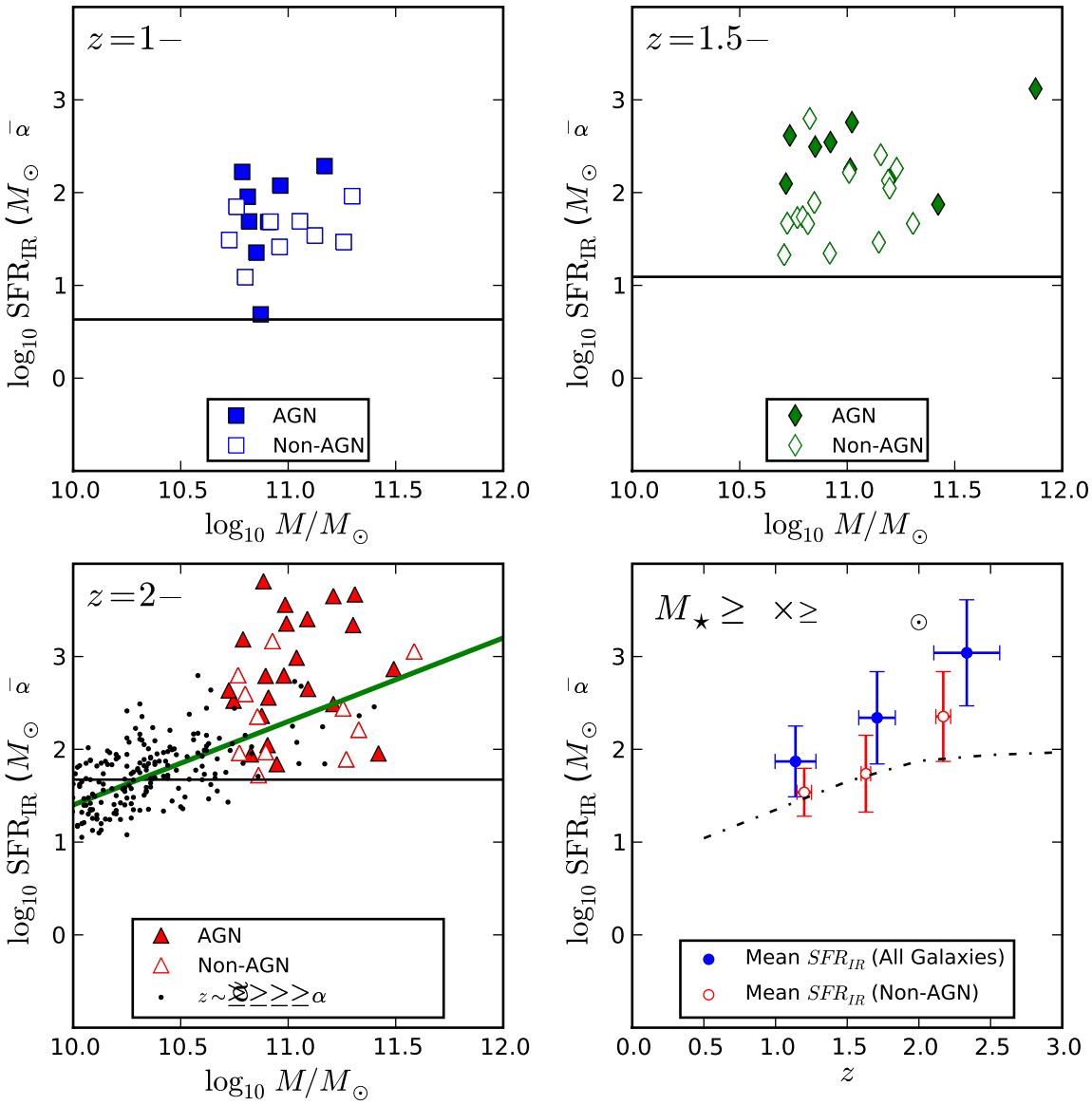

FIG. 20.— The top left, top right, and bottom left show SFR$_{\rm IR}$ versus $M_\star$, for $M_\star \geq 5 \times 10^{10}$ $M_\odot$, in different redshift bins, each corresponding to $\sim 1-1.4$ Gyr in cosmic time. The values of SFR$_{\rm IR}$ (see §4.2) are based on the assumption that L_{IR} comes from dust heated primarily by young stars. In these three plots, the AGN candidates are coded as filled symbols, and their SFR$_{\rm IR}$ likely overestimate their true SFR. The completeness limit in SFR$_{\rm IR}$ (corresponding to the limiting 24 μm flux of $\sim$30 μJy) changes with redshift (see Table 4) and is shown as a black solid line in each panel. In the lower-left panel, the black dots represent SFR from UV measurements by Daddi et al. (2007) for $z \sim 2$; the diagonal green line is the corresponding SFR-mass correlation at $z \sim 2$. The bottom right panel shows mean SFR$_{\rm IR}$ in the different redshift bins. The error bars are the 1σ standard deviation around the mean. The black line shows average UV-based SFR versus redshift for a 5×10^{10} $M_\odot$ galaxy as calculated by Drory & Alvarez (2008).

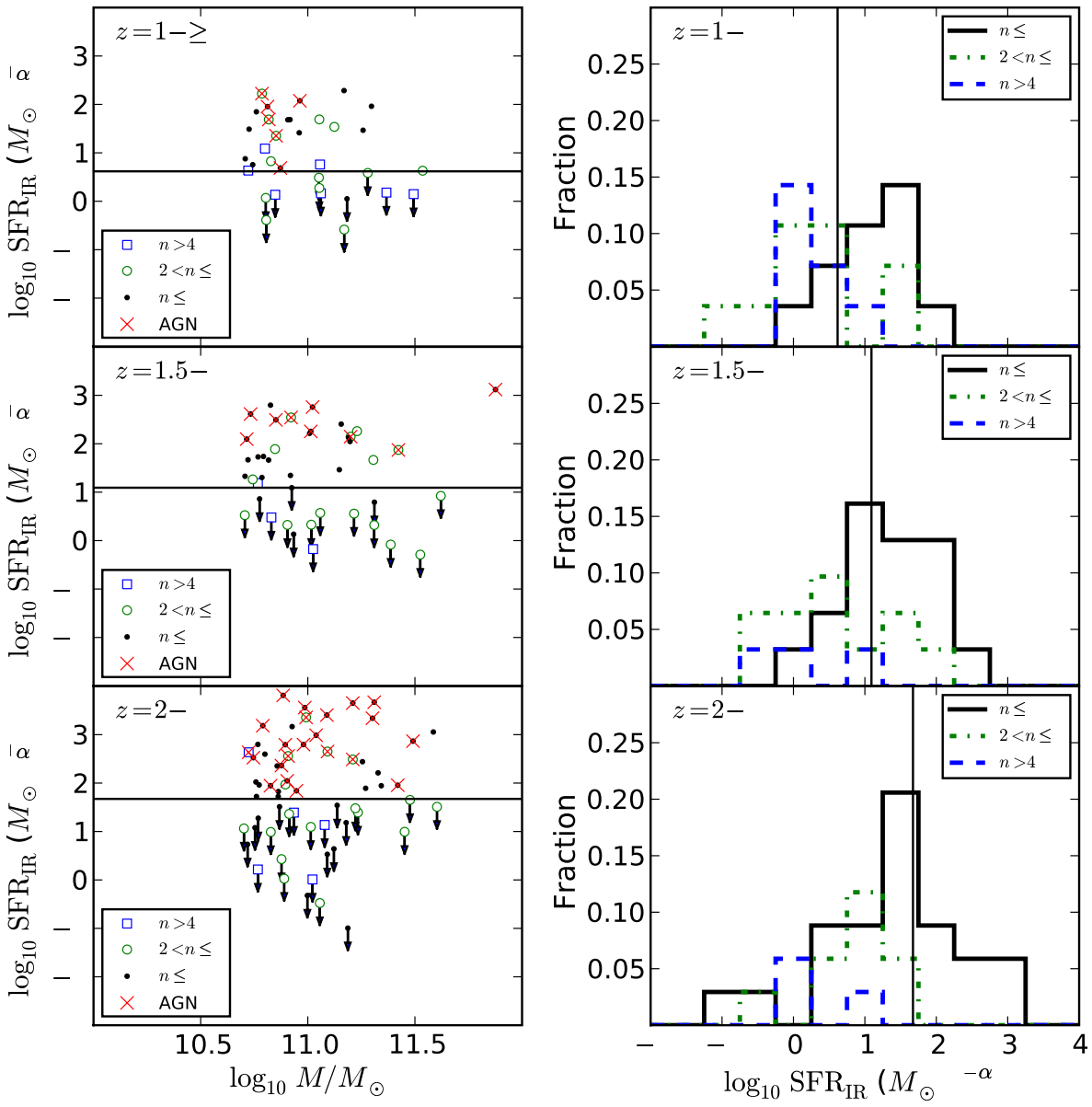

FIG. 21.— Left column: SFR_{IR} is plotted versus $M_\star$, for all galaxies with $M_\star \geq 5 \times 10^{10}\ M_\odot$, in different redshift bins spanning 1-1.4 Gyr in cosmic time. Data are sorted according to the Sérsic index n calculated in §3.2. AGN candidates (see §6) are labeled with red x's. Galaxies with SFR_{IR} below the detection limit (shown as a horizontal line) are shown with downward pointing arrows because they are upper limits. Right column: For non-AGN sources, histograms show *the fraction of massive galaxies* in each redshift bin with a given SFR_{IR} for separate ranges of n. The vertical black lines mark the SFR_{IR} detection limit. For sources to the left of the line, we plot upper limits for SFR_{IR}. The majority of galaxies with SFR_{IR} above the detection limit have $n \leq 2$, and in fact, the high SFR tail in each redshift bin is populated primarily by $n \leq 2$ systems.

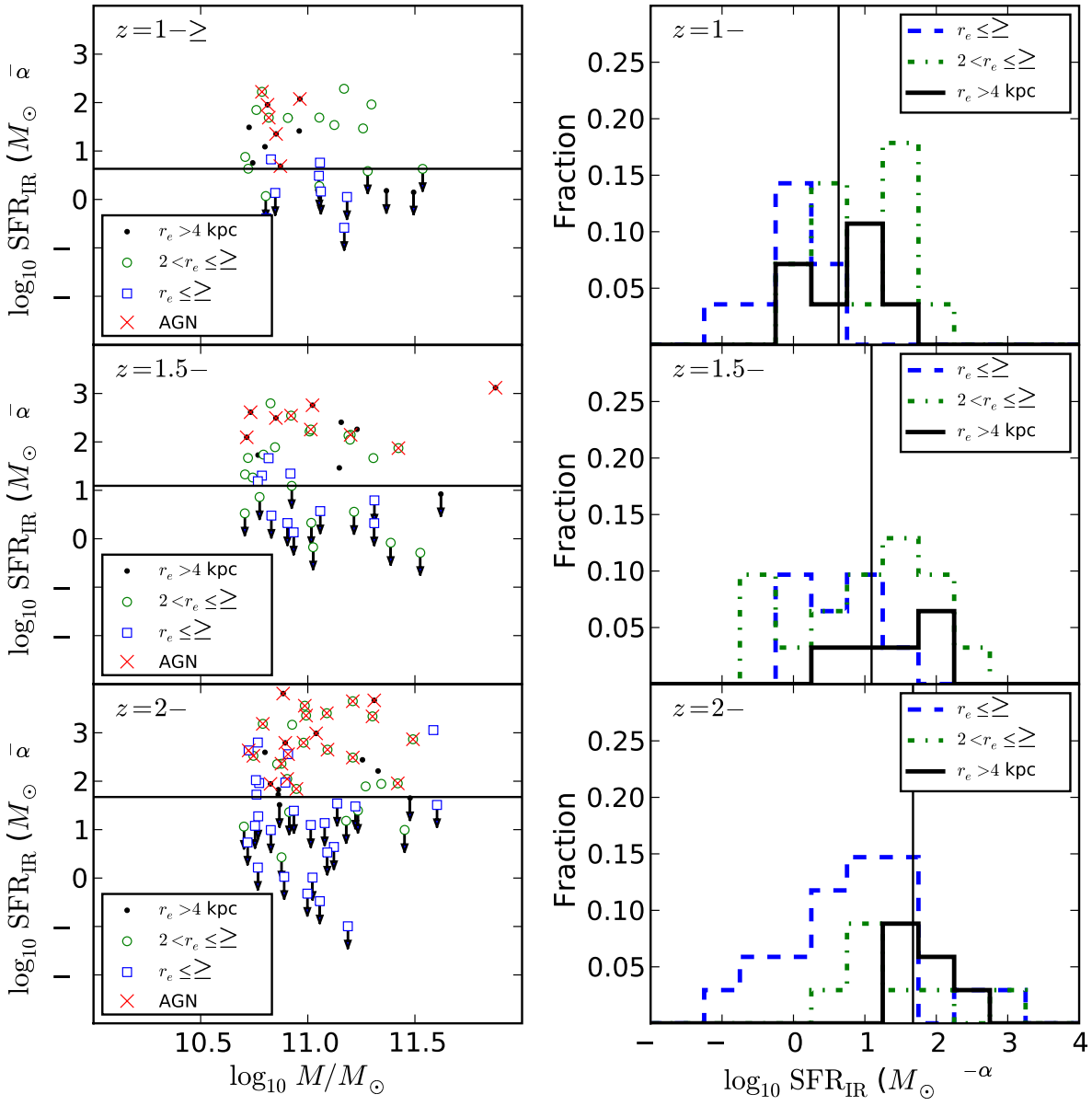

FIG. 22.— Same as Figure 21, but now the data are sorted by half-light radius r_e. Note that only a small fraction of the ultra-compact ($r_e \leq 2$ kpc) galaxies have SFR$_{\rm IR}$ above the 5σ detection limit and have on average lower SFR than more extended systems.

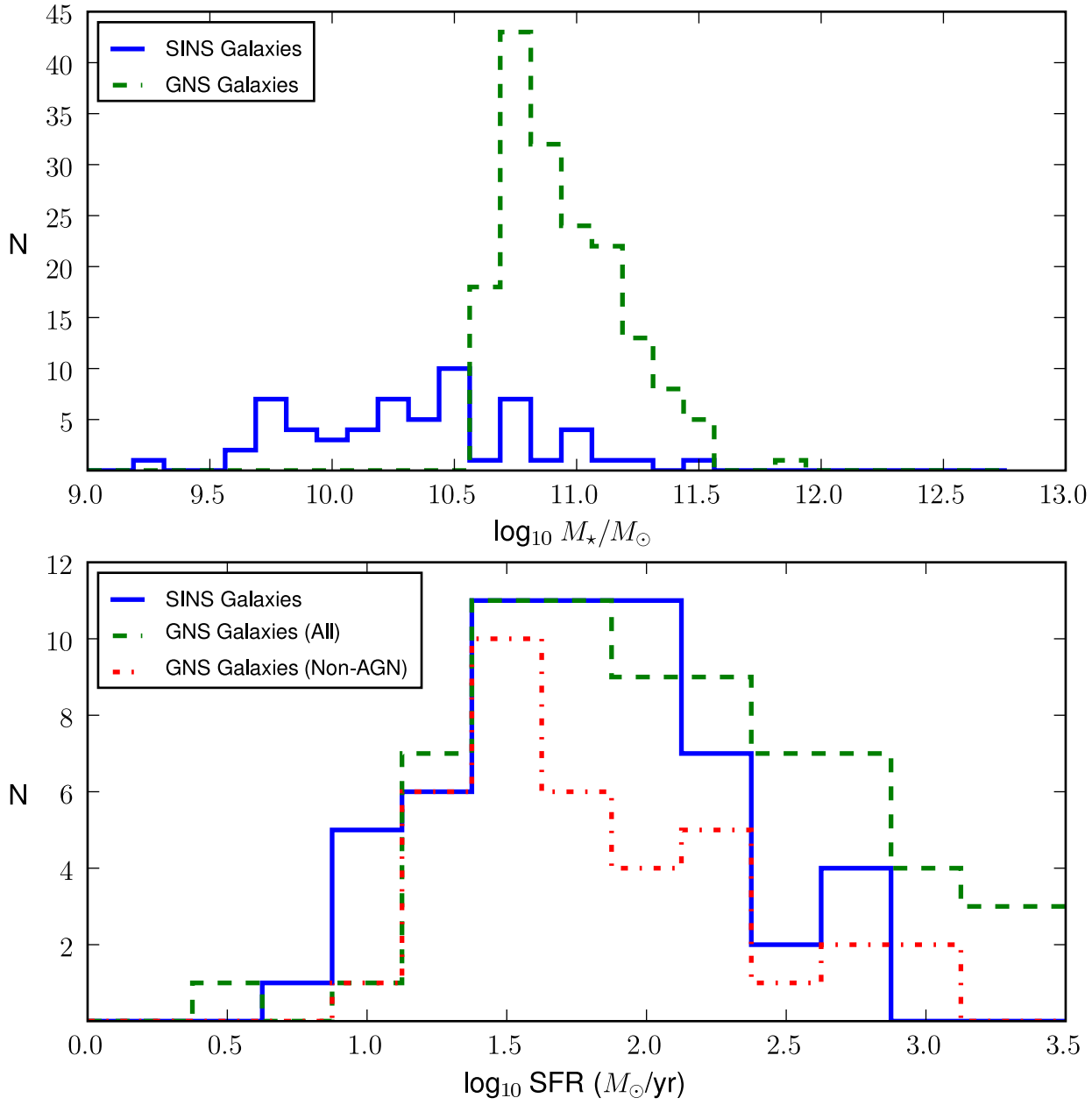

FIG. 23.— The top panel compares the stellar mass from the SINS galaxies (Förster Schreiber et al. 2009) to the massive GNS galaxies. The bottom panel shows the distribution in SFR_{IR} for each sample. For the GNS galaxies, SFR_{IR} is shown for all sources with secure 24 μm detections and the subset of those without evidence for AGN contamination (see §6).

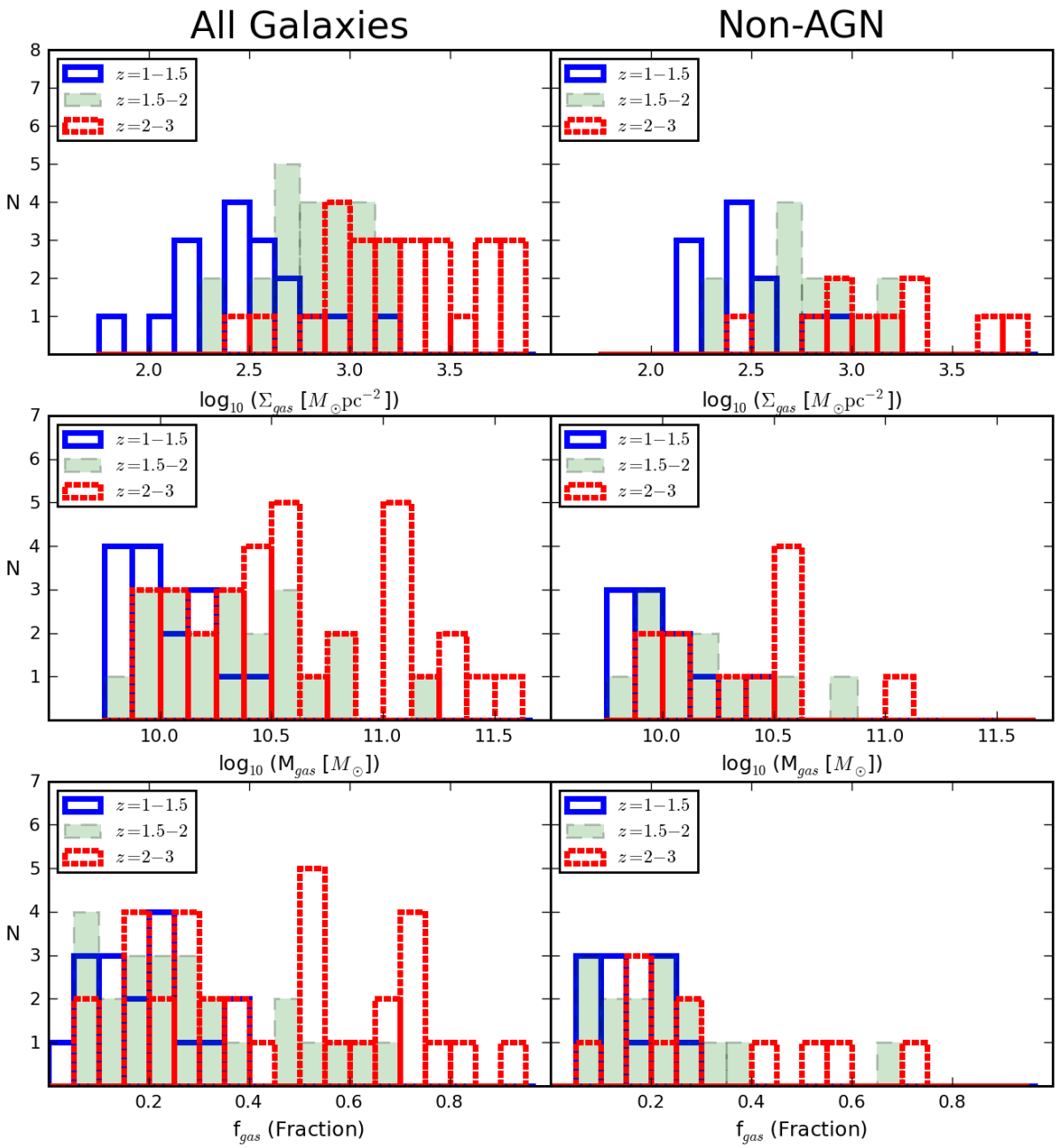

FIG. 24.— Left column: The distributions of cold gas surface density ($\Sigma_{\rm gas}$), cold gas mass $M_{\rm gas}$, and cold gas fraction ($f_{\rm gas}$) within r_e are shown for all galaxies across different redshift ranges. $\Sigma_{\rm gas}$ is calculated using a Schmidt-Kennicutt law with power-law index 1.4. Cold gas fraction is calculated relative to the total baryonic mass within r_e ($f_{\rm gas} \equiv M_{\rm gas}/(M_{\rm gas} + M_\star)$). Right column: Same as left column except that only non-AGN sources are shown.

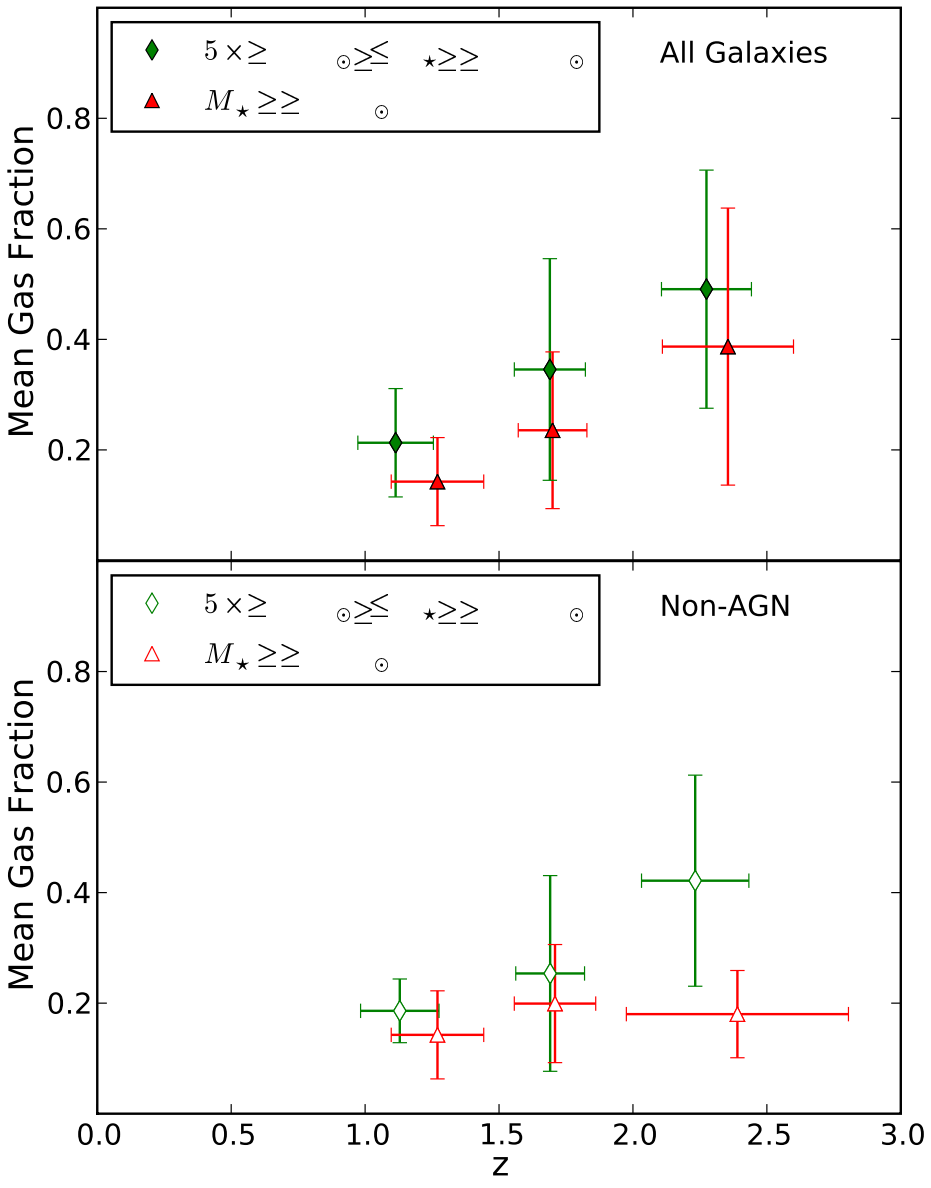

FIG. 25.— Top panel: The mean gas fraction is shown in three redshift bins for all galaxies with $5 \times 10^{10}\ M_{\odot} \leq M_{\star} < 10^{11}\ M_{\odot}$ and $M_{\star} \geq 10^{11}\ M_{\odot}$. The error bars indicate the 1σ scatter in gas fraction and redshift. Bottom panel: Same as the top except that only non-AGN sources are shown.

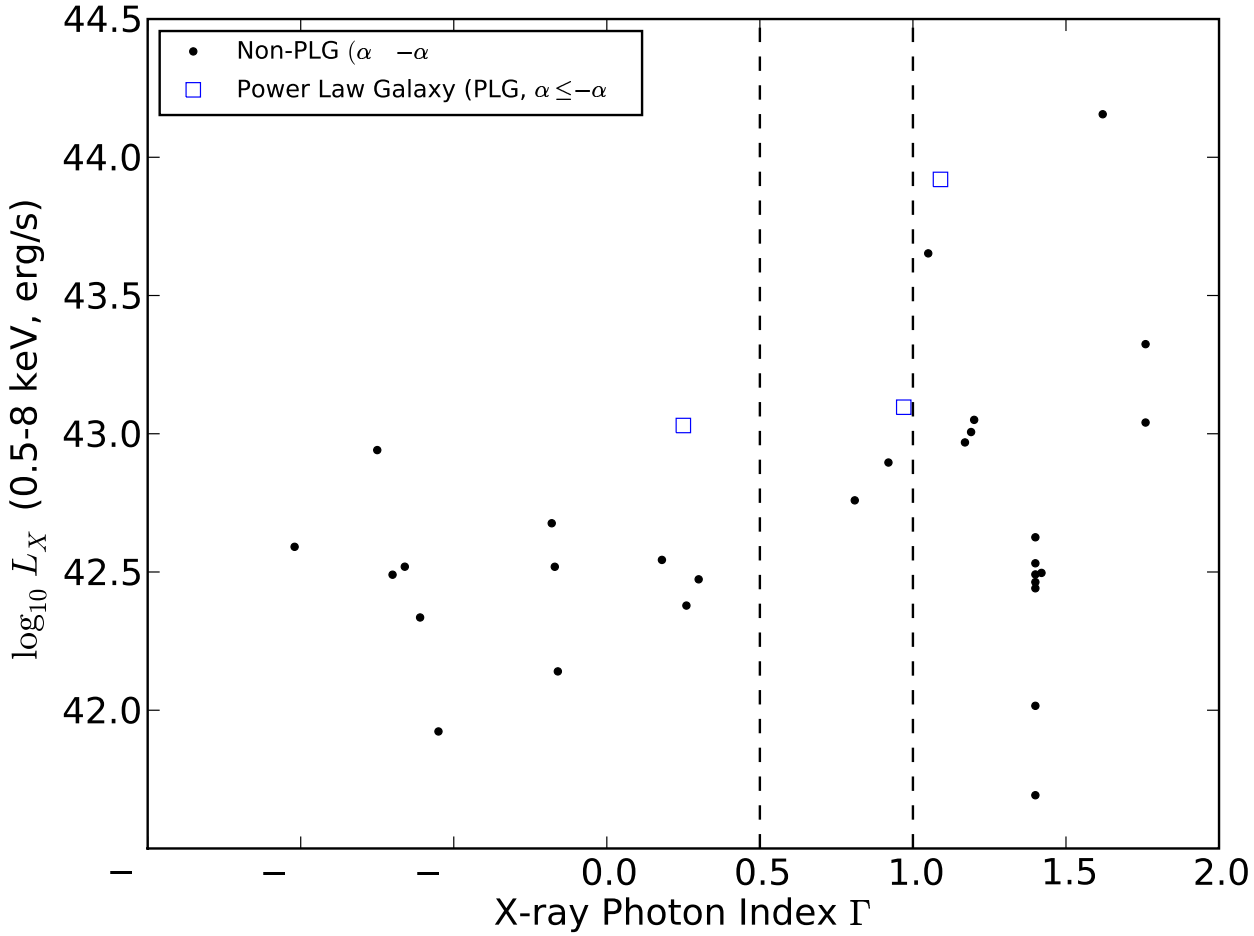

FIG. 26.— Rest-frame Chandra luminosity (L_X, 0.5-8 keV) is plotted against X-ray photon index Γ. Typically AGN have X-ray luminosities $\geq 10^{42}$ erg s^{-1}. Unobscured AGN and typical star-forming galaxies tend to have $\Gamma \approx 2$. Obscured AGN tend to have $\Gamma < 1$. A value of $\Gamma < 0.5$ is usually taken as an unambiguous obscured AGN, while $\Gamma \sim 0.5 - 1$ is indicative of either an obscured AGN or an abundance of high-mass X-ray binaries (see § 6.1 for details). For comparison, we also show sources with mid-IR power-law index $\alpha \leq -0.5$ (identified in § 6.2). These are usually considered as obscured AGN and are marked as power-law galaxies (PLGs).

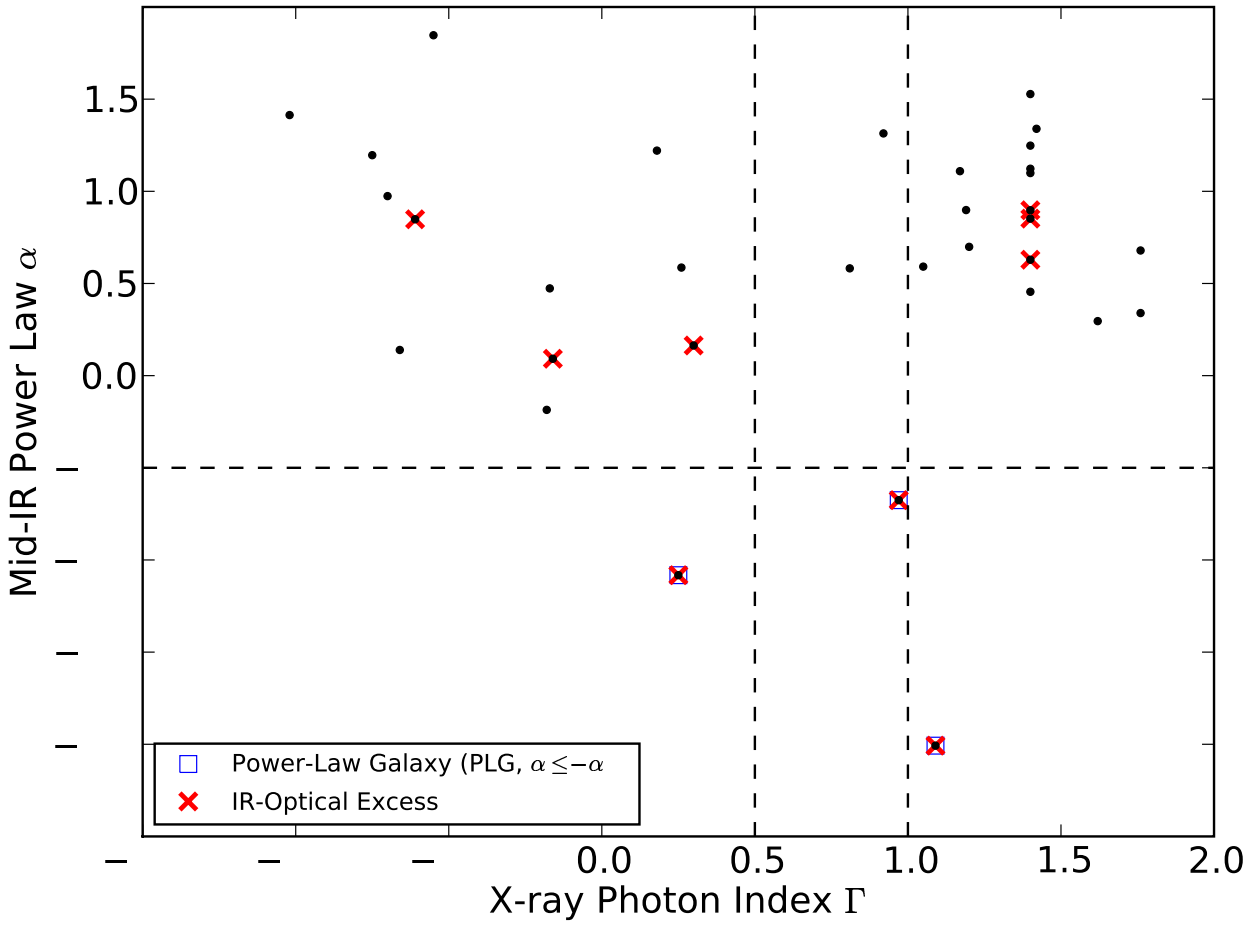

FIG. 27.— IR power-law index α determined from the IRAC bands is plotted against X-ray spectral index Γ. The horizontal and vertical lines denote $\alpha = -0.5$, $\Gamma = 0.5$, and $\Gamma = 1.0$. Sources with mid-IR power-law index $\alpha \leq -0.5$ are usually considered as obscured AGN and are marked as power-law galaxies (PLGs) with blue squares; three such candidates are found. Sources with X-ray spectral index $\Gamma < 0.5$ are usually considered as unambiguous obscured AGN and 13 such candidates are found. The regime $\Gamma \sim 0.5 - 2.0$ can be populated by both AGN and starbursts. It is noteworthy that of the 13 sources which were identified as unambiguous obscured AGN due to their $\Gamma < 0.5$ values, only one is a PLG. The red x's represent the subset of X-ray sources that are also identified to show an IR-to-optical emission excess ($f_{24\mu m}/f_R \geq 1000$). Note that three of these cases are also PLGs.

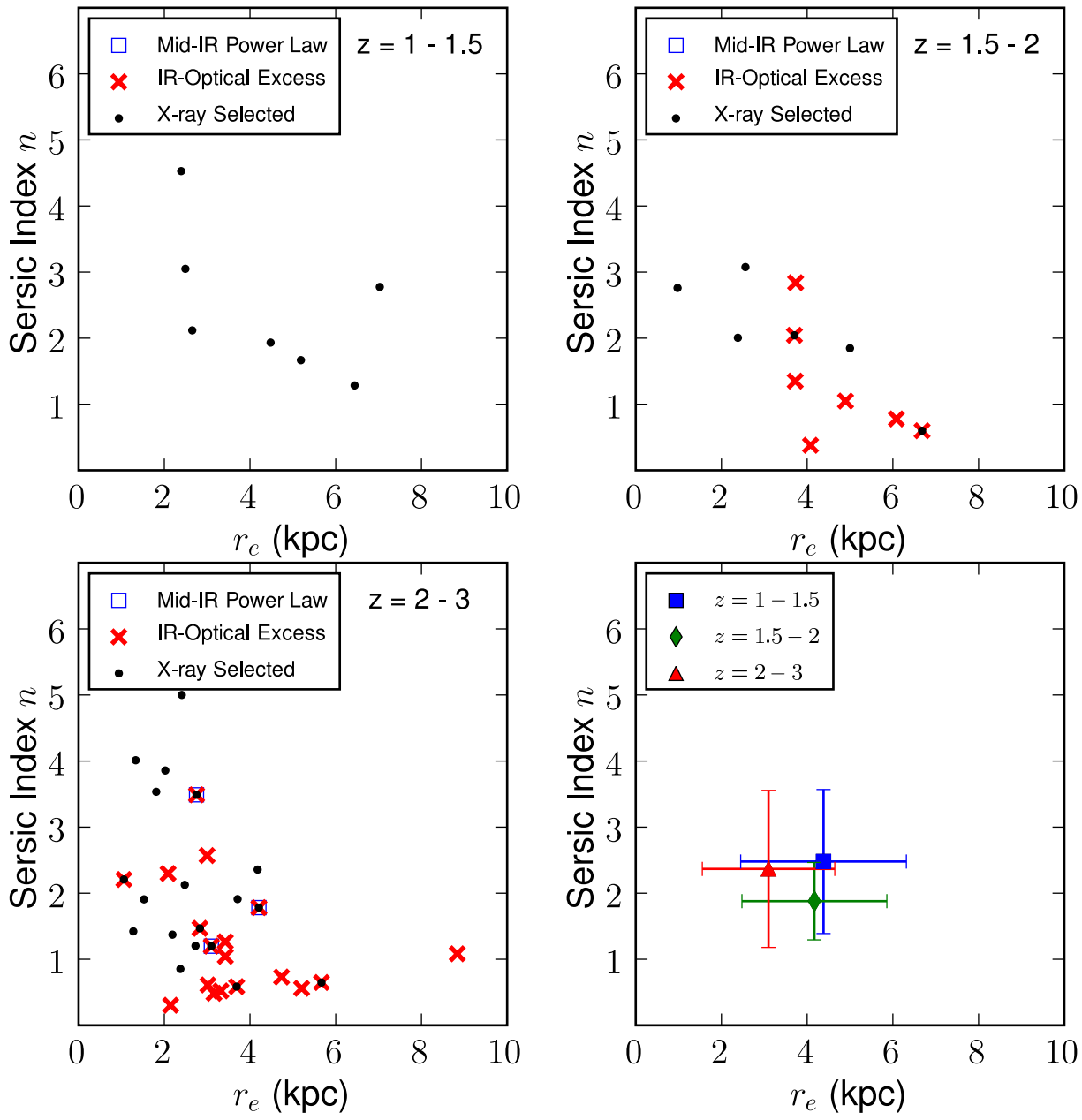

FIG. 28.— The upper and lower left panels show rest-frame single Sérsic index versus effective radius r_e for the 49 AGN candidates selected either based on X-ray properties, mid-IR power-law, or IR-to-optical excess. The lower right panel shows the mean Sérsic index and r_e in each redshift bin. The error bars in the last panel show the 1σ scatter in Sérsic index and r_e for each redshift bin.

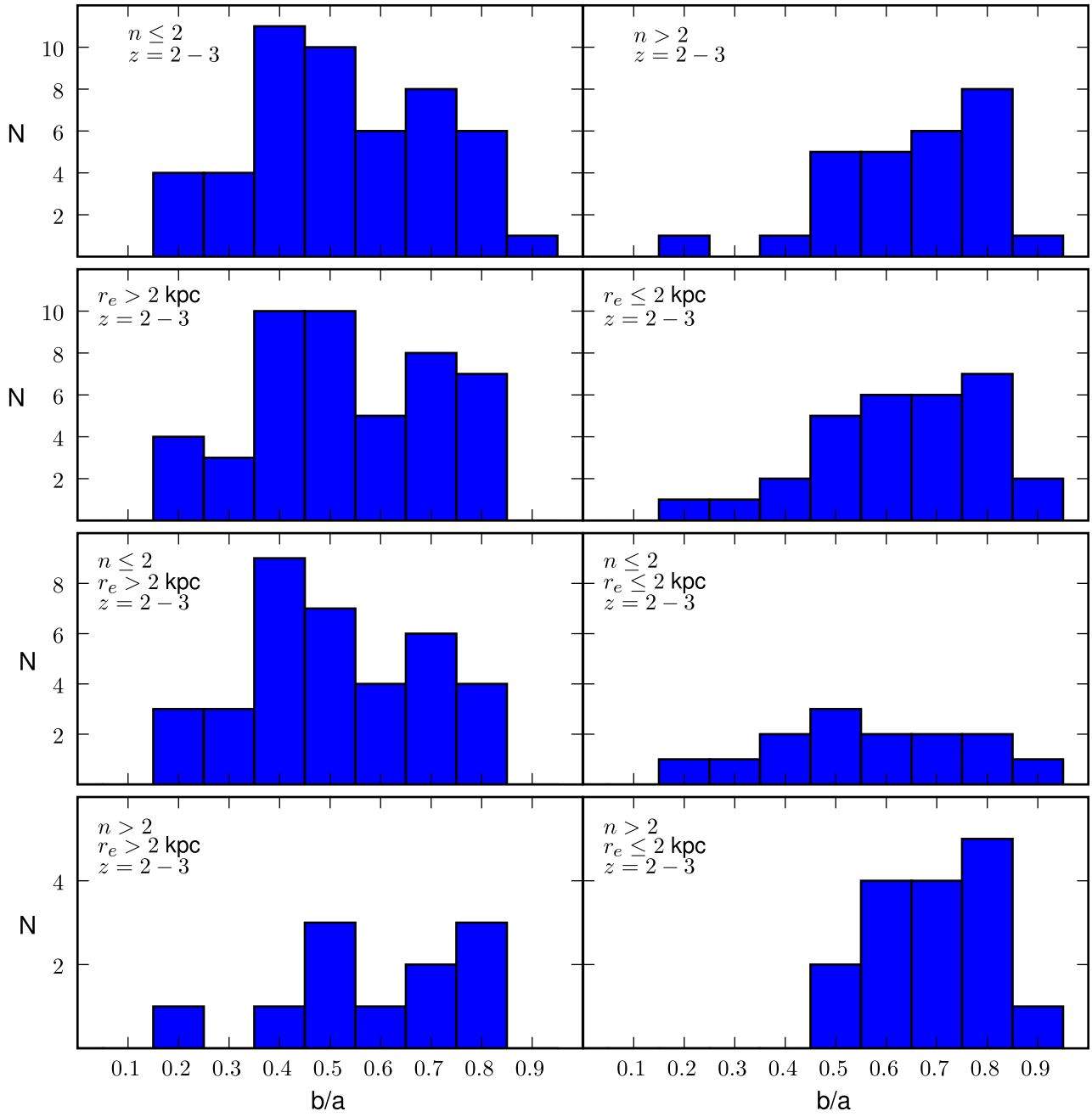

FIG. 29.— Axial ratio (b/a) is shown for massive ($M_\star \geq 5 \times 10^{10}\ M_\odot$) GNS galaxies at $z = 2 - 3$. Each panel shows b/a for a different combinations of structural parameters (n,r_e).