The Astronomical Journal, 167:31 (18pp), 2024 January https://doi.org/10.3847/1538-3881/acf7bc

The 50 Mpc Galaxy Catalog (50 MGC): Consistent and Homogeneous Masses, Distances, Colors, and Morphologies

David Ohlson[1], Anil C. Seth[1], Elena Gallo[2], Vivienne F. Baldassare[3], and Jenny E. Greene[4]
[1] University of Utah, 201 James Fletcher Building 115 S. 1400 E. Salt Lake City, UT 84112, USA; david.ohlson@utah.edu
[2] University of Michigan, USA
[3] Washington State University, USA
[4] Princeton University, USA

Received 2023 March 31; revised 2023 August 31; accepted 2023 September 5; published 2023 December 22

Abstract

We assemble a catalog of 15424 nearby galaxies within 50 Mpc with consistent and homogenized mass, distance, and morphological type measurements. Our catalog combines galaxies from HyperLeda, the NASA-Sloan Atlas, and the Catalog of Local Volume Galaxies. Distances for the galaxies combine best-estimates for flow-corrected redshift-based distances with redshift-independent distances. We also compile magnitude and color information for 11740 galaxies. We use the galaxy colors to estimate masses by creating self-consistent color—mass-to-light ratio relations in four bands; we also provide color transformations of all colors into Sloan *g–i* by using galaxies with overlapping color information. We compile morphology information for 13744 galaxies, and use the galaxy color information to separate early- and late-type galaxies. This catalog is widely applicable for studies of nearby galaxies and for placing these studies in the context of more distant galaxies. We present one application here: a preliminary analysis of the nuclear X-ray activity of galaxies. Out of 1506 galaxies within the sample that have available Chandra X-ray observations, we find that 291 have detected nuclear sources. Of the 291 existing Chandra detections, 249 have $\log(L_X) > 38.3$ and available stellar mass estimates. We find that the X-ray active fractions in early-type galaxies are higher than in late-type galaxies, especially for galaxy stellar masses between 10^9 and $10^{10.5}\ M_\odot$. We show that these differences may be due at least in part to the increased astrometric uncertainties in late-type galaxies relative to early types.

Unified Astronomy Thesaurus concepts: AGN host galaxies (2017); Catalogs (205); X-ray sources (1822); Supermassive black holes (1663); Dwarf galaxies (416); Galaxies (573)

Supporting material: machine-readable tables

1. Introduction

The galaxies in the nearby Universe provide our best laboratory for studying a wide range of scientific questions. Most fundamentally, the properties of nearby galaxies are the outcome of the process of galaxy evolution over the lifetime of our Universe. We can use our detailed view of nearby galaxies to quantify their morphology, kinematics, and star formation histories.

Our most complete census of galaxies comes from the local measurements. Only within the Local Group can the faintest galaxies be found (e.g., Willman et al. 2005; McConnachie 2012; Simon 2019), and even with modern surveys such as the Sloan Digital Sky Survey (SDSS), we can obtain a census of bright dwarf galaxies with masses ($M_\star$) $\sim 10^9\ M_\odot$ only out to ~50 Mpc (e.g., Blanton et al. 2005). The detailed structures within nearby galaxies can reveal many aspects of their history and evolution, including their faint stellar halos (e.g., Crnojević et al. 2016), star cluster populations (e.g., Krumholz et al. 2019; Hughes et al. 2021), dwarf galaxy satellites (e.g., Carlsten et al. 2022), central stellar structures (e.g., Erwin et al. 2015; Neumayer et al. 2020), and their supermassive black holes (e.g., Saglia et al. 2016). Nearby galaxies also provide a unique laboratory for stellar and high-energy astrophysics, such as quantifying rare stages of stellar evolution that are not represented in the Milky Way (e.g., Melbourne et al. 2012), finding and characterizing rare sources such as ultraluminous X-ray sources (Kaaret et al. 2017), and transient events including gravitational wave events (Abbott et al. 2017) and rare types of supernovae (e.g., Tartaglia et al. 2018). Finally, nearby galaxies provide a baseline for distance measurements in the Universe, and measurement of their distances is key for understanding cosmology (e.g., Freedman 2021).

A number of existing catalogs of nearby galaxies are regularly used in selecting sources for these science explorations. The Karachentsev et al. (2013) Local Volume Galaxy catalog[5] is a regularly updated and very complete catalog of the nearest galaxies with distances <11 Mpc and includes distance and photometric information. Beyond this distance, the most widely used sources of galaxy information are HyperLeda[6] (Makarov et al. 2014) and the NASA Extragalactic Database (NED).[7] In addition, many studies make use of catalogs of the nearest rich clusters of galaxies, Virgo (Binggeli et al. 1985; Kim et al. 2014). Distance measurements to nearby galaxies have been a subject of intense study, and much of this information for nearby galaxies has been compiled within a subsection of NED (Steer et al. 2017). All of these works combine heterogeneous photometric, distance, and morphological measurements, and make some attempt to synthesize and

---

[5] https://www.sao.ru/lv/lvgdb/introduction.php

[6] http://leda.univ-lyon1.fr/

[7] http://ned.ipac.caltech.edu/

homogenize them. Modern studies of galaxies typically characterize galaxies based on their stellar masses and colors (e.g., Baldry et al. 2004). Unfortunately, there are no uniform and homogeneous estimates of these quantities for nearby galaxies, and the varying photometric data as well as heterogeneous distance estimates available for these galaxies complicate their comparison to more distant galaxy samples.

The goal of this paper is to provide the first homogenized source of nearby galaxy masses, distances, colors, and morphologies. We have chosen a distance limit of ∼50 Mpc for this catalog, as many of the existing studies of nearby galaxies are limited to galaxies within this distance limit. We find in Section 4 that these galaxies have stellar masses that range from $\log(M_\star/M_\odot)$ of 6.6 to 10.9 (at the 99%ile). We also present one application that motivated the creation of this catalog: a comparison of nuclear X-ray activity between early- and late-type galaxies. These results are the first step to determining constraints on the occupation fraction of central massive black holes in nearby galaxies (Miller et al. 2015; Gallo & Sesana 2019), which will be presented in a follow-up paper (E. Gallo et al. 2023, in preparation).

This paper is organized as follows: We describe our sample selection and photometry in Section 2. We discuss distance estimates to the sample galaxies in Section 3, determine their masses based on consistent color-mass-to-light ratio measurements in Section 4, and describe their morphologies and group membership in Section 5. Finally, in Section 6, we apply the catalog by combining it with Chandra archival data to present our X-ray active fraction measurement before concluding. The full galaxy catalog is available in Appendix.[8] Throughout the paper, when we refer to column names without other clarifying information, we refer to the full catalog; columns are indicated through their font, i.e.,`column_font_format`.

2. Sample Selection and Photometry

In this section, we detail the construction of our sample from multiple sources as well as the compilation of photometry that enables luminosity and mass estimates for each galaxy.

2.1. Base Catalogs and Combination

Our catalog is formed from three base sources, the Local Volume Galaxy (LVG) catalog (Karachentsev et al. 2013), HyperLeda (Makarov et al. 2014), and the NASA-Sloan Atlas (NSA).[9] All three contain a wide variety of information, including redshifts and photometry. The strength of the NSA catalog is its uniform Sloan Digital Sky Survey (SDSS) photometry, which extends to quite faint sources, however, this photometry only extends over about one-third of the sky. The LVG and HyperLeda catalogs combine hetereogenous data sets and are continuously maintained and updated. The LVG catalog contains nearly complete measurements of galaxies brighter than $M_B \sim -12$ within 10 Mpc, but it lacks optical color information for the galaxies. Both NSA and HyperLeda extend to greater distances and have optical color information. Combined, these catalogs can provide the data needed to obtain distances, masses, and Hubble types for galaxies out to a maximum target distance of 50 Mpc.

From HyperLeda, we initially began with a subsample of 18726 objects designated as galaxies and with radial velocity <3500 km s^{-1}. We chose to remove the ∼20% of objects that were missing absolute B-band magnitude values in the column `mabs` from our sample, because the HyperLeda calculation of the absolute magnitude relied upon apparent B-band magnitude, distance modulus, and extinction. Removal of these objects from the sample ensured the inclusion of the most essential galaxy measurements. This cut the initial HyperLeda sample down to 14941 objects.

For the LVG subsample, we combined the tables from the "Catalog of Nearby Galaxies" and "Global Parameters of the Nearby Galaxies" that are available from the LVG catalog website.[10] The combined tables totaled 1240 galaxies.[11]

Finally, for the NSA data, we used the `nsa_v1_0_1` catalog and found 8242 objects with $cz < 3500$ km s^{-1}. We also added available SDSS spectral line flux data to these objects, matched by plate and fiber IDs. We note that the NSA catalog includes not just objects with SDSS redshift information, but also galaxies with redshifts measured from other sources (ALFALFA, NED, 6dF, 2dF, and ZCAT).

2.1.1. Combining the Sample

The three subsamples contain both duplicate objects and contaminants. Here, we describe how we combined the catalogs while eliminating duplicates and contaminants. We note that the presence of a galaxy in each input catalog is given by flags in our catalog: `hl_obj`, `lvg_obj`, and `nsa_obj`. If a galaxy is present in that catalog, the flag has a value of 1, while if it is not, it has a 0 value.

The HyperLeda sample included some duplicate galaxy entries. To remove them, we crossmatched each entry to its nearest neighbor within the sample. We followed up on galaxies with nearby neighbors within a 2′ separation radius for galaxies within 20 Mpc, and a 1′ separation for galaxies between 20 and 50 Mpc. For the 580 potential duplicate galaxies, we examined available data and flagged them as follows: 0: Individual galaxy with the best available data. 1: Duplicated galaxy with fewer available data. 2: Off-nuclear position in duplicate galaxy. 3: Image shows a bright or star-forming region within a larger galaxy. We include only objects flagged with 0, which removed 183/580 galaxies. This cut resulted in an initial HyperLeda sample of 14758 galaxies.

We first created a combined LVG and HyperLeda sample. We used `astropy.coordinates.SkyCoord` to conduct a coordinate-based crossmatch with a maximum separation distance of 30″. Of the 1240 galaxies in the LVG, 744 were found in HyperLeda. The 496 unmatched LVG galaxies almost all have a very low luminosity. This explains their exclusion, as these galaxies are likely to lack the photometric and redshift measurements required to have been included in our initial HyperLeda sample. When combining sources between these catalogs, we prioritize the LVG galaxy properties in almost all cases, i.e., radial velocities, B-band magnitudes, and distances (see Section 3).

We then crossmatched NSA catalog objects to our combined HyperLeda and LVG samples using a maximum separation distance of 30″. Of the 8242 NSA objects, 6966 matched the combined HyperLeda and LVG catalog. For the the 1276 unmatched objects, we found that many were contaminants. To

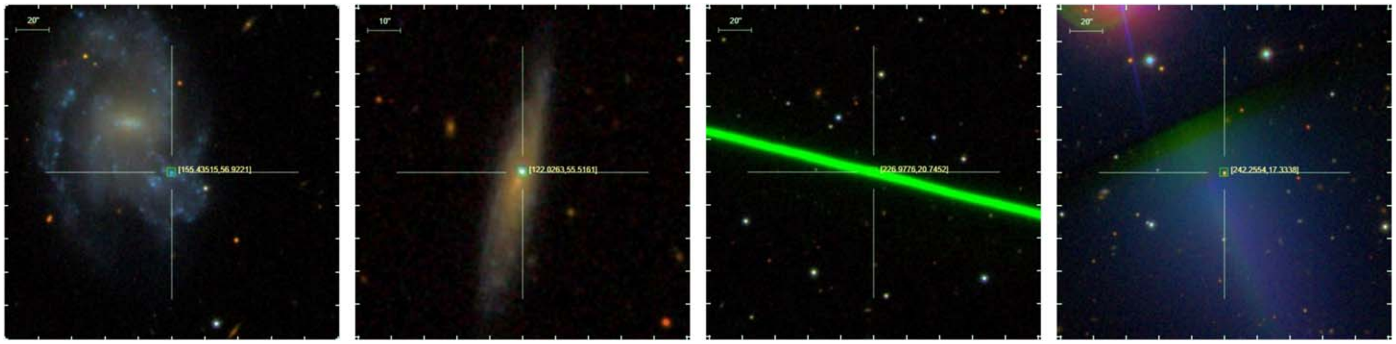
---

[8] This table is also available (and potentially updated) at https://github.com/davidohlson/50MGC.

[9] https://www.sdss.org/dr13/manga/manga-target-selection/nsa/

[10] http://www.sao.ru/lv/lvgdb/tables.php

[11] The version used was accessed 8/4/2020.



(a) (b) (c) (d)

Figure 1. Examples of SDSS imaging of NSA contaminant objects that are removed from our catalog. (a) Imaging shows a bright, star-forming region of a galaxy (NSAID: 177945). (b) A foreground star centered on a distant galaxy that causes incorrect redshift values for the galaxy (NSAID: 401970). (c) An imaging artifact treated as a galaxy by NSA (NSAID: 460798). (d) An artifact from a bright star in the top left corner surrounds a nearby star in the center, which causes both the classification as a galaxy and provides the redshift of the star (NSAID: 460798).

Table 1
NSA Objects that Do Not Match Galaxies in the Hyperleda and LVG Catalogs

NSAID	R.A.	Decl.	Quality_FLAG	MATCH_RATIO
30129	343.8028213	−0.350540038	2	191.944
34438	13.23646298	−1.174127269	2	225.389
36914	24.69818421	1.071791959	2	371.004
38419	32.6382837	0.738166461	4	nan
41803	8.850357094	14.79454207	2	122.785
42629	14.93267592	14.56776754	2	331.128
46179	116.1573335	40.88379846	4	151.729
50066	129.4311915	51.64175054	1	0.752
51666	134.7746823	53.76023531	1	0.707
52302	138.5671796	57.04441951	4	318.963

(This table is available in its entirety in machine-readable form.)

distinguish these contaminants from real galaxies, we visually inspected the SDSS imaging of each galaxy and compared redshifts to those listed in the NED. At the lowest redshifts, the contaminants included foreground stars on top of image artifacts or distant galaxies. At larger distances, bright portions of galaxies with separate spectroscopic measurements from SDSS were often included as separate catalog entries. Some contaminants were also just image artifacts, and a small number contained no obvious image source whatsoever. Figure 1 shows examples of NSA contaminant objects. We flagged each of the 1276 unmatched NSA objects, based on our findings: 1: Image shows a bright or star-forming region within a larger galaxy. 2: Image shows no object, an imaging artifact, a star, or a foreground star over a high-redshift galaxy. 3: Duplicates of otherwise good galaxies, of which we flagged the entry with the most reliable redshift source to keep, in order, SDSS, NED, and ZCAT. 4: Good imaging of a galaxy with a redshift matching the redshift in NED. These qualtity flags are given for all 1276 galaxies in Table 1. Only objects in category 4 are considered as possible unique new NSA galaxies to add to the catalog.

Finally, to minimize the potential for duplicate galaxies, we wished to ensure that none of the unique NSA objects in category 4 were observations of the edge of a galaxy that was already included in our sample. To do this, we first needed uniform angular diameter measurements for each sample galaxy.

HyperLeda gives galaxy diameter measurements as the major axis using a standard 25 $\mathrm{mag\,arcsec^{-2}}$ isophote ($d25$), whereas LVG uses a diameter measurement corresponding to the Holmsberg isophote ($\sim$26.5 $\mathrm{mag\,arcsec^{-2}}$; which we call $d26$). We compared $d25$ HyperLeda diameters to $d26$ LVG diameters for galaxies found in both source catalogs to find their typical ratio, and then multiplied the LVG diameters by $\sim$0.838 to approximate $d25$ HyperLeda values for all the nearest-neighbor galaxies.

For each NSA object, we then divided the angular separation from its nearest-neighbor's angular diameter to create a match separation ratio, $\delta_{\mathrm{sep}}/d25$. Of the galaxies that were confirmed to be included in our sample, we only included those with a ratio $\delta \geqslant 0.75$ (i.e., lying beyond 1.5 galaxy radii). Of the 208 objects in category 4, 170 galaxies passed this cut and were added to our final sample. These galaxies have a median distance of 29.9 Mpc and a median mass of $8.08 \times 10^7\ M_{\odot}$ ($\log(M) \sim 8$), and thus they add significantly to our sample.

We note that many studies (e.g., Geha et al. 2012; Reines et al. 2013) have used the NSA catalog as a sample of nearby galaxies without removing contaminants. We find that $\sim$13% of the nearby NSA galaxies (1106/8242) are contaminants. Most contaminants show a low-mass estimate from NSA redshift distances and photometry, with a median mass of $2.6 \times 10^7\ M_{\odot}$. We therefore include the full list of NSA objects that do not match HyperLeda and LVG and the criteria we used to cut these down in Table 1 for future users.

2.1.2. Galaxy Coordinates

We wished the coordinates in our catalog to be centered as accurately as possible on the nuclear regions of the galaxies for the purpose of the selection and observation of active galactic nuclei (AGN). Our catalog combines multiple subsamples, each with slightly different astrometric methods and sources. HyperLeda applies a homogenized position computed by averaging published measurements weighted by an assigned quality flag. NSA provides coordinates of the measured object's photometric center in the SDSS. The source of LVG coordinates was not given. Each time we combined subsamples with different astrometric sources, we compared the coordinates for galaxies in the two catalogs. We visually inspected 150 galaxies with overlap between HyperLeda and either LVG (739 galaxies) or NSA (7136 galaxies). We randomly selected these galaxies to have a range of offsets between 1″ and 10″. We compared the galaxy positions to imaging with SDSS and 2MASS, both of which have good astrometry (Pier et al. 2003; Skrutskie et al. 2006). This visual inspection suggests that in general, the HyperLeda positions are better aligned with the photocenter than the LVG and NSA samples. We therefore use the following priority order for the `ra` and `dec` in our catalog: HyperLeda, LVG, and NSA.

In addition to the comparison and selection between position sources from multiple catalog sources, we also compared our selected R.A. and decl. values to the primary position provided by NED. We retrieved NED positions using a name-based query. After removing 25 matches with a separation >5′ due to misnamed entries in HyperLeda and NED, we reliably matched positions for 14412 catalog galaxies (∼93%). We used the visual inspection process described in the previous paragraph to determine whether NED coordinates are noticeably preferable to our catalog position. We found that our positions were generally better than those from NED, with a larger spread between our catalog positions and NED than between HyperLeda, LVG, and NSA. We find that 1680 galaxies have a separation >3″ between our catalog position and the NED position. We investigate the relation between coordinate offsets, mass, and galaxy type in more detail in Section 6. Because of the variation between positions from different sources, we choose to include columns for NSA and NED positions in the final catalog, in addition to our chosen compilation of coordinates described above.

2.2. Color and Luminosity

The luminosities and colors of our galaxies are two crucial parts of our catalog. In addition to being intrinsically useful, the colors are also important for calculating stellar masses (Section 4) and for assessing the galaxy type (early versus late; Section 5). To obtain these estimates, we use four different sources for colors and associated luminosities: (1) HyperLeda, (2) NSA, (3) the Siena Galaxy Atlas (SGA), and (4) the NASA/IPAC Extragalactic Database (NED). All luminosity calculations use the best distances described in Section 3.

Extinction Correction: For our foreground-extinction corrections, we uniformly apply the Schlafly & Finkbeiner (2011) corrections; all our sources have Schlegel et al. (1998) extinction estimates available, and we multiply these by 0.86 to convert them into Schlafly & Finkbeiner (2011) values.[12] For HyperLeda, we used B and V magnitudes, and scaled their A_B as described above. To determine A_V values, we assumed an $A_V/A_B = 0.769$ (Cardelli et al. 1989). For NSA, we used the provided Schlegel et al. (1998) extinction values for each band and scaled them by 0.86. We retrieved a table of location-based extinction values for our entire catalog from the IPAC Galactic DUST web service and added the table's Schlafly & Finkbeiner (2011) $E(B-V)$ values to the catalog as `EBV_irsa`. For SGA, we then created band-specific extinction arrays by multiplying `EBV_irsa` by the proper $A/E(B-V)$ values: $g = 3.303$, $r = 2.285$. Finally, for NED, we used B- and R-band magnitudes from the same source for consistency—these primarily come from the southern ESO Uppsala survey (Lauberts & Valentijn 1989); we used the HyperLeda B-band extinction, and assumed an $A_R/A_B = 0.628$ based on the extinction curves of Cardelli et al. (1989) and O'Donnell (1994), taken from the Padova CMD website.[13]

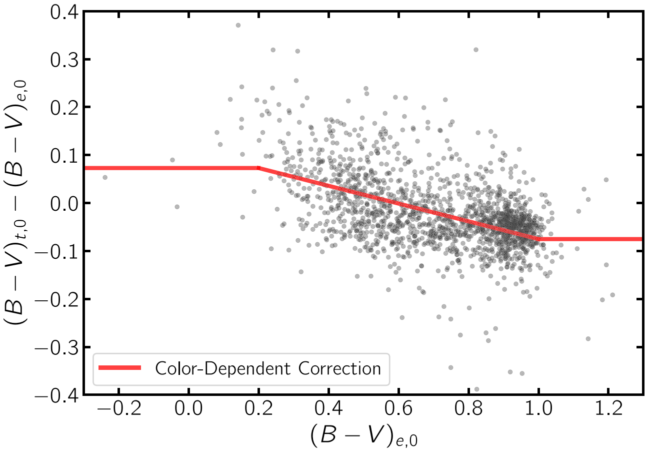


Figure 2. Difference between HyperLeda galaxies with overlapping total ($(B-V)_{t,0}$) and effective ($(B-V)_{e,0}$) colors vs. their $(B-V)_{e,0}$ value. The red line shows the linear correction applied to $(B-V)_{e,0}$ values to correct for the slight color-dependent offset to allow a combination of colors into a single field. Correction coefficients are based on a linear fit from $0.2 \leqslant (B-V)_{e,0} \leqslant 1.0$.

We did not apply any internal extinction corrections, thus the luminosities and colors presented here are only corrected for foreground extinction. Internal extinction values are available from both HyperLeda and LVG. However, we found no correlation between these two.

Luminosities and Colors from HyperLeda ($(B-V)_0$, L_V, L_B): HyperLeda provides two different $B-V$ colors: (1) The total asymptotic corrected color, `bvtc`, which is corrected for both internal and foreground extinction, and (2) the color measured within the effective aperture in which half the total B flux is emitted, `bve`, which is not corrected for extinction. Various HyperLeda galaxies have one or both measurements available. Our first step to obtain consistent values was to undo the HyperLeda extinction corrections on the `bvtc` column, and apply the Schlafly & Finkbeiner (2011) foreground-extinction correction to both the total and effective colors—we call these extinction-corrected quantities $(B-V)_{t,0}$ and $(B-V)_{e,0}$ respectively. Because we are using the colors to calculate total masses, we also attempt to correct the small offset between the $(B-V)_{t,0}$ and $(B-V)_{e,0}$—the difference between these colors for galaxies that have both values is shown in Figure 2. The offset is clearly color dependent, and the line in that figure shows the color-dependent correction applied to the $(B-V)_{e,0}$ to obtain our final $(B-V)_0$ values for galaxies that only had `bve` data available. We flagged these data using the `bvcolor_f` value in our table.

---

[12] https://irsa.ipac.caltech.edu/applications/DUST/

[13] http://stev.oapd.inaf.it/cgi-bin/cmd_3.4

To calculate luminosities and total magnitudes in B and V bands, we found that the given `bt` and `vt` columns were not consistent with the given colors (after correcting for the different extinction corrections applied)—this is likely due to the heterogeneity of the data sources used. Given the ubiquity of `btc` values, we first calculate the B total magnitude by removing the extinction corrections and applying Schlafly & Finkbeiner (2011) foreground-extinction corrections to obtain the total $B_{t,0}$ values. We then derived the total V-band magnitude by using our $B_{t,0}$ and subtracted the $(B-V)_{t,0}$ value to obtain a foreground-extinction V-band magnitude, $V_{t,0}$. We note the use of $B_{t,0}$ values throughout this paper when we plot galaxy luminosities. These are transformed into luminosities using the Vega-based absolute magnitudes of the Sun from http://mips.as.arizona.educnaw/Sun.html.

Luminosities and Colors from NSA ($(g-i)_0$, L_i, L_B): NSA provides Petrosian flux and extinction values for each *FNugriz* band. We calculate the Pogson g-, r-, and i-band magnitudes using the SDSS relation[14]: $m = 22.5–2.5\log_{10}f$. We then used extinction correction for all individual magnitudes as described above. To calculate galaxy masses (Section 4), we use the g_0 and i_0 magnitudes to calculate an extinction-corrected $(g-i)_0$ color. We also calculate an i-band luminosity for the galaxies (L_i) using the solar AB magnitude from http://mips.as.arizona.edu/cnaw/Sun.html. To analyze a single luminosity distribution for our entire sample, we also estimate the B-band luminosity L_B from the SDSS magnitudes using the stellar transformation $B_0 = g_0 + 0.39^*(g-r)_0 + 0.21$ from Jester et al. (2005).

Luminosities and Colors from SGA ($(g-r)_0$, L_r, L_B): A complementary source of photometry for primarily southern galaxies is the Siena Galaxy Atlas (SGA[15]). We matched our catalog to the SGA catalog after removing sources with known redshifts above $3500\ \mathrm{km\ s^{-1}}$ using a maximum separation distance of 30″. In cases of multiple galaxies within this threshold, we matched the galaxy with the smallest angular separation. This gave us $(g-r)_0$ data for 7681 galaxies, 1616 of which otherwise lacked color.

The SGA catalog contains a number of different magnitude measurements. To ensure consistency, we compared overlapping sources with NSA and found that `[G,R,Z]_MAG_SB26` matched the NSA magnitudes best. We then calculated reddenings and extinction using the IRSA online portal[16] to obtain Schlafly & Finkbeiner (2011) extinction values. We corrected for these values to calculate the $(g-r)_0$ color, L_r, and estimate the B-band luminosity L_B using the same Jester et al. (2005) equation as we applied to NSA photometry.

Luminosities and Colors from NED ($(B-R)_0$, L_R, L_B): To expand our mass estimations for southern galaxies, we used astroquery to request NED photometry for each sample object. We added available B and R data to the sample, specifically, objects with an observed passband of `B (Cousins) (B_25)` and `R (Cousins) (R_25)`. If multiple magnitude measurements were available for a single object, we included a mean value. Queries returned a uniform B and R magnitude error of 0.09 for all objects. We used these B- and R-band magnitudes, corrected for foreground extinction, to calculate the $(B-R)_0$ color, L_R, and L_B where available.

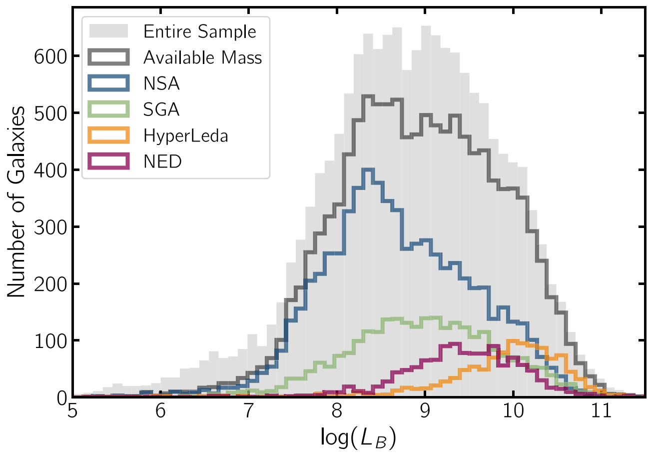


Figure 3. The distribution of the B-band luminosities log(L_B) for the entire catalog (shaded histogram), along with the subset that has available masses (dark gray line). The colored lines indicate the source of photometry used in estimating the galaxy masses (see Section 4.)

Combined Color and Luminosity Estimates: In Section 5 we describe color transformations between $(g-i)_0$ and other colors. We use these to derive an estimate of the $(g-i)_0$ color for all galaxies with colors. This is contained in column `gi_color`. For galaxies with multiple available colors, we give precedence to source colors in order: NSA $(g-i)_0$, and then synthesized $(g-i)_0$ values based on $(g-r)_0$, $(B-V)_0$, and $(B-R)_0$. We compile the L_B values from all sources in the column `B_lum`; most of these are taken directly from HyperLeda, while a handful are synthesized from other sources as described above, with the precedence order being HyperLeda, NED, NSA, and SGA. A vast majority of L_B estimates come directly from Hyperleda. The final distribution of L_B estimates, separated by preferred photometric source, is shown in Figure 3.

2.3. Cleaning the Photometric Measurements Used for Mass Estimates

Due to a number of objects with both NSA and HyperLeda data available, we wished to test which source provided more reliable photometry when both were available. We compared g versus B magnitudes, due to their similar wavelength, and investigated galaxies with significantly differing magnitudes to available NED values. Generally, by comparing to available photometry compiled in NED, we find that NSA photometry is more reliable in cases where it is discrepant from HyperLeda, and we therefore typically prioritize NSA data in estimating galaxy masses. An exception to this is for a selection of galaxies with NSA $g > 16$ and $B - g > 1$; for these galaxies, we preferentially use Hyperleda values, as the HyperLeda values matched other measurements from NED more closely. Based on this analysis, we created the `mag_flag` column to signify which source provided the more accurate measurement at a given magnitude—0: The difference between B and g are within one magnitude, so NSA values are used for calculating masses. 1: B and g are discrepant; NSA photometry is more reliable. 2: B and g are discrepant; HyperLeda photometry is more reliable, and is used to estimate masses.

There are 446 additional objects for which the NSA photometry is not used. These include objects with flags indicating unreliable photometry, specifically those with the

---

[14] https://www.sdss.org/dr12/algorithms/magnitudes/

[15] https://www.legacysurvey.org/sga/sga2020/

[16] https://irsa.ipac.caltech.edu/applications/DUST/

fifth or sixth bitwise flag set in the NSA `DFLAGS` field (M. Blanton, private communication). We also do not use NSA photometry for 20 objects with g or i petrosian flux $\leqslant 0$ nanomaggies. Of these 446 galaxies, 306 had other colors and luminosities available to calculate masses.

3. Deriving Best Distances

In this section, we discuss our derivation of the best available distance estimates across our full galaxy sample. These distance estimates fall into three broad categories: (i) redshift-based distances determined using the Cosmicflows-3 Distance–Velocity calculator (Kourkchi et al. 2020), (ii) redshift-independent distances taken from the LVG and from the NASA/IPAC Extragalactic Database (NED Steer et al. 2017; Steer 2020), and (iii) Virgo Cluster distances, for which objects matching Virgo samples and catalogs were assigned either measured distances or an appropriate mean distance. In general, we prioritize the redshift-independent and Virgo distances except in cases where the available redshift-based distances are more accurate. About 20% of the redshift distances are more accurate when the distances reach >30 Mpc and the error becomes a lower fraction of total distance.

3.1. Redshift-based Distances

Due to the peculiar motions of galaxies, redshift-based distances are not always accurate in the nearby Universe. However, this accuracy can be improved by using models of the local mass density based on the measured peculiar motions of galaxies (Courtois et al. 2012). We chose to use the Cosmicflows-3 Distance–Velocity calculator (Kourkchi et al. 2020) because it covered the full distance range of our sample using an up-to-date local mass density model (Graziani et al. 2019). We entered heliocentric velocities into the calculator to retrieve distance estimates that were corrected for the local mass density. Velocity measurements are given precedence in the order NSA, LVG, and HyperLeda where available.

To estimate the errors on the redshift-based distances, we first estimated a typical peculiar velocity for galaxies in the Local Volume. We based this on galaxies with redshift-independent measurements (see the next subsection) with errors <10%. We then used the CMB-relative velocity measurements from HyperLeda to calculate the peculiar velocity of each galaxy using $v_{\rm pec} = v_{\rm rad} - H_0 d$ assuming $H_0 = 70$. We then measured the width of this distribution using the 16th and 84th percentiles, taking half the width to obtain a typical peculiar velocity of ~ 400 km s^{-1}.

Our redshift-based distance errors were then calculated by adding and subtracting this typical peculiar velocity from the heliocentric velocity and recalculating the distance using the Cosmicflows-3 distance–velocity calculator. The final redshift-dependent distance error uses a symmetrization of the difference between upper- and lower-limit corrected distances. These distance errors are typically ~65.3% at 10 Mpc and ~17.4% at 40 Mpc.

3.2. Redshift-independent Distances

We use redshift-independent distances from two sources: (1) the updated LVG catalog distances, and (2) the compilation of redshift-independent distances in the NED distance database (NED-D; Steer et al. 2017). We chose to give the highest priority to the LVG catalog distances, as these distances typically represent the best available measurements and are kept up to date. We manually examined distance estimates where the values from LVG conflicted with those NED-D and NED-MED, and found LVG to typically be superior; for instance, the distances to the Maffei group galaxies were based on the recently updated TRGB distances presented in Anand et al. (2019).

For galaxies in NED-D, we chose our distances using the following list of primary indicators, in order of precedence based on the order given in Steer (2020): Cepheids, TRGB, RR Lyrae, Red Clump, supernova (SN) Ia, FGLR, Horizontal Branch, SBF, GCLF, CMD, Type II Cepheids, Miras, PNLF, asymmetric giant branch (AGB), carbon stars, and SNIa SDSS. We use the single most recent preferred distance estimate of the preferred estimate, except if two preferred measurements were from the same year, in which case, we apply an average of these preferred distances and errors. If NED-D provided more than two older measurements from the best primary indicator, we ensured that the newest distance was not greater than two standard deviations from the averaged older distances. For galaxies with large spreads in their preferred distances, we use the mean of all preferred distances with the best primary indicator. We encode this information in the field `'dist_nedd_flag'`—0: uses the single best NED-D measurement, 1: best NED-D distance averaged from multiple measurements, and 2: NED-D best indicator mean used due to discrepancy, as described previously.

3.3. Galaxies in Virgo

Due to the nearby Virgo Cluster's large velocity dispersion, redshift-based distances to Virgo are much more inaccurate than for galaxies in the rest of our sample. Because of this, we treated Virgo as a special case for distance purposes. We matched and compared our sample to galaxies from Mei et al. (2007) and the Extended Virgo Cluster Catalog (EVCC; Kim et al. 2014)—we included certain members from their catalog as special cases to have their distances revised. These galaxies have velocities that lie within the range of galaxies that are gravitationally bound to Virgo at a given projected radius.

For these galaxies, we first combined the separate redshift-based and redshift-independent distance lists, using whichever available option had the lowest fractional error. Galaxies with surface brightness fluctuation distances from Mei et al. (2007) are assigned that distance and error. Other Virgo galaxies with only redshift-based distances available are assigned the average distance of 16.5 ± 1.1 Mpc given by Mei et al. (2007). For galaxies with redshift-independent distances, we found more divergent values than expected, including some that did not agree with the Virgo distances. After careful evaluation of individual cases, we assigned the 16.5 ± 1.1 Mpc distance for most objects, keeping the redshift-independent distance only if $|D - 16.5| > 3\sigma$, where σ is the error on the distance; this is a total of 862 objects.

3.4. Final Compilation of Distances

For galaxies with multiple distance sources, our aim was to identify the more reliable and precise distance source. Our distance sources included Virgo distances, LVG distances, NED-D distances with the lowest error, and redshift-based distances from the Cosmicflows-3 calculator. We prioritized

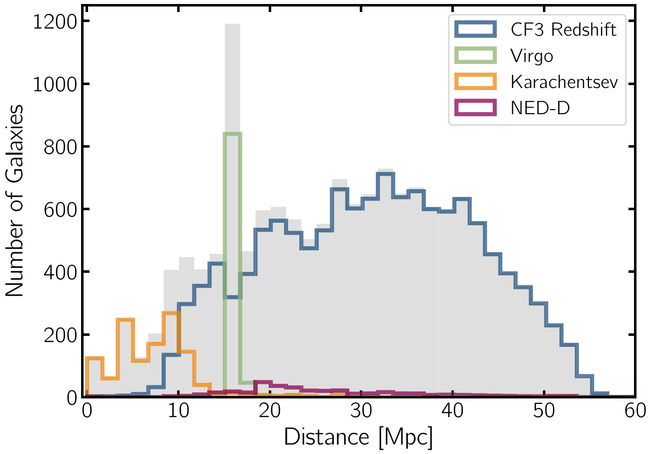


Figure 4. Distribution of our chosen best distances for the entire catalog (shaded histogram). The colored lines indicate the source (`bestdist_source`) used for the final distance compilation (see Section 3).

Virgo distances and otherwise chose the distances with the lowest error between the NED-D and CF3 z-distances.

Of the 55 objects that had no assigned distance from any of our sources, 29 are assigned the given distances and errors from HyperLeda. The remaining 26 are NSA-only objects for which velocity measurements returned a Cosmicflows-3 corrected distance $D \leqslant 0$; we do not supply a distance in this case.

As well as best distances (the `bestdist` column) and errors (`bestdist_error`), each galaxy in the catalog has a field for a specific distance indicator, if available. We have also included the field `bestdist_source`. The values of this are–0: Virgo, 1: LVG, 2: NED-D, 3: redshift based, and 4: HyperLeda, "nan": No Distance. The final distribution of galaxy distances, separated using these classifications, is shown in Figure 4.

4. Mass Estimation

Our goal in this section is to estimate consistent stellar masses across the full sample based on galaxy colors and luminosities. We use the *g*–*i* versus M/L_i color–*M/L* relation from Taylor et al. (2011) as the basis for our measurements and derive self-consistent relations in other colors using galaxies with overlapping measurements.

We use the *g*–*i* versus M/L_i relation as our preferred relation as previous work has found that it is accurate to ~0.1 dex due to the alignment of the extinction and stellar population effects (Zibetti et al. 2009; Taylor et al. 2011). We note that while near-IR (NIR) photometry is widely used to estimate stellar masses assuming a constant M/L (e.g., Meidt et al. 2014), this approach is compromised in galaxies with younger stellar populations by its sensitivity to difficult-to-model TP-AGB stars and hot dust emission (e.g., Kriek et al. 2010; Taylor et al. 2011; McGaugh & Schombert 2014; Telford et al. 2020).

One challenge in deriving stellar masses in our sample is the different photometry available for different galaxies. This creates two issues: first, McGaugh & Schombert (2014) have shown that individual color–*M/L* relations are not consistent across multiple bands, i.e., that using different colors in the Bell et al. (2003) or Zibetti et al. (2009) relations gives systematically different masses for observed galaxy colors. Second, many available relations are not available across both the Sloan and Johnson photometry that we are using for our galaxy sample (e.g., Roediger & Courteau 2015). To solve both of these issues, we therefore translate the Taylor et al. (2011) *g*–*i* versus M/L_i into the other colors we used ($(g-r)_0$, $(B-V)_0$, and $(B-R)_0$) using overlapping galaxies to ensure self-consistency.

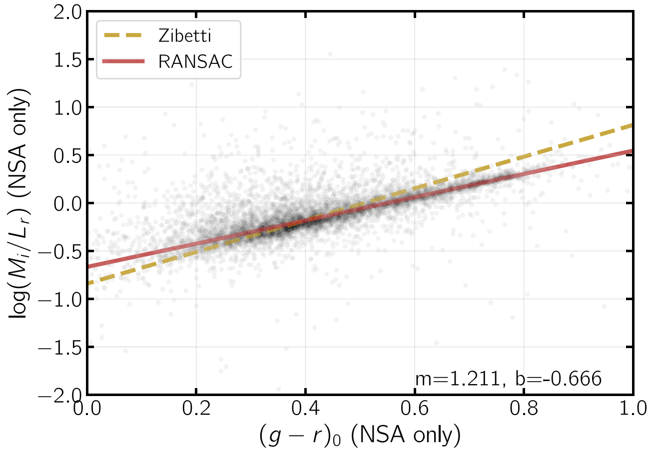


Figure 5. An example of how we transform the Taylor et al. (2011) $(g-i)_0$–M/L_i relation into $(g-r)_0$–M/L_r. The *y*-axis is determined by taking masses calculated from the NSA data using the $(g-i)_0$–M/L_i relation, and then dividing by the NSA L_r values. This is then plotted against NSA $(g-r)_0$ values and fit using a robust outlier-rejection algorithm, `RANSAC`. Also shown is the somewhat steeper relation from Zibetti et al. (2009). Our translation of the Taylor et al. (2011) relation into other colors ensures consistency between masses calculated in different bands. We note that although this $(g-r)_0$–M/L_r relation is used for galaxies with SGA data, we fit this relation with NSA photometry alone to ensure a large sample of galaxies and to minimize photometric outliers. A similar process was used to also create the other color–M/L relations given in Section 4.

We first used Equation (7) from Taylor et al. (2011),

$$\log(M_{\star,i}/L_i) = -0.68 + 0.70(g-i), \tag{1}$$

to calculate galaxy stellar mass estimates ($\log(M_{\star,i}/L_i)$) using $(g-i)_0$ and L_i derived from NSA data. To translate this relation into other bands, we used overlapping samples. First, for 6677 NSA galaxies with both $(g-i)_0$ and $(g-r)_0$ color available, we compared the stellar masses calculated from the $i-$band to the $r-$band luminosities, i.e., $\log(M_{\star,i}/L_r)$. We then used this to fit the $M_{\star,r}/L_r$ versus $(g-r)_0$ relation using the `RANSAC` fitting from `sklearn.linear_model` to create a linear fit to this comparison while ignoring outliers, as shown in Figure 5. We followed the same process with 986 overlapping galaxies between NSA and HyperLeda to translate the color–*M/L* relation into $(B-V)_0$. Our sample lacks significant overlap between galaxies with both $(g-i)_0$ and $(B-R)_0$ because the latter primarily come from southern sky observations. We therefore used the same method on 822 galaxies with $(g-r)_0$ and $(B-R)_0$ to bootstrap an appropriate and self-consistent M/L relation. The resulting equations from these transformations are

$$\log(M_{\star,r}/L_r) = -0.66 + 1.20(g-r)_0. \tag{2}$$

$$\log(M_{\star,V}/L_V) = -0.72 + 1.07(B-V)_0 \tag{3}$$

$$\log(M_{\star,R}/L_R) = -0.51 + 0.51(B-R)_0. \tag{4}$$

We multiplied the color-based *M/L* ratios by their corresponding luminosity to calculate a stellar mass estimate. We created separate columns for mass estimates from each color: *g*–*i*, $g-r$, $B-V$, and $B-R$. To compile our final column of

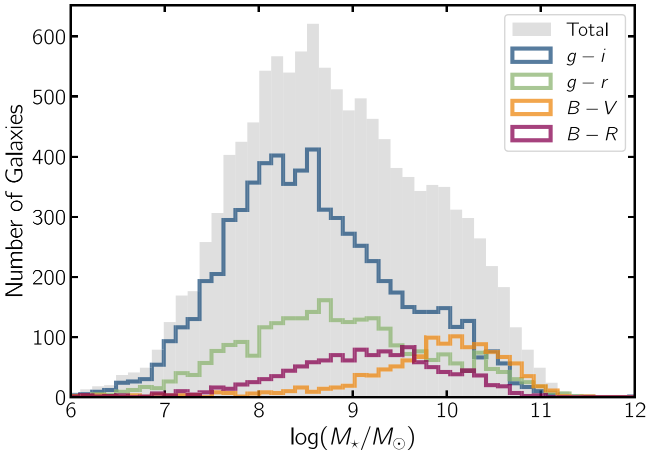


Figure 6. Distribution of log($M_\star/M_\odot$) estimates for the entire catalog (shaded histogram). The colored lines indicate which colors are used to estimate the stellar mass using color-mass-to-light ratio relations and precedence as described in Section 4.

the best-mass estimates for 11740 galaxies, we gave precedence when available to first: 6626 log($M_{\star,i}$) values calculated using NSA data using Equation (1). Second: 2705 log($M_{\star,r}$) values from SGA data using Equation (2). Third: 1182 log($M_{\star,V}$) values from HyperLeda using Equation (3). Fourth: 1227 log($M_{\star,R}$) values from retrieved NED photometry using Equation (4). Figure 6 shows the distribution of log($\mathrm{M}\star/\mathrm{M}_\odot$) estimates for the catalog, separated by which color is used to estimate the stellar mass.

To obtain errors on the mass estimates, we use the scatter from multiple estimates for the same objects. To do this, we take the sample of 986 galaxies with both $(B-V)_0$ and $(g-i)_0$ colors, and compare the masses derived with these two separate colors. This comparison includes scatter in the mass estimates due to different bandpass information and independent photometric measurements. The 68 percentile of the mass differences is 0.136. We use this value as an assumed floor on the log($M_\star/M_\odot$) error. We also propagated distance errors by calculating high and low masses by adding and subtracting `bestdist_error`. For our final `logmass_error`, we use the larger of these two error estimates.

4.1. Mass Comparisons

Our mass estimates are based on those of Taylor et al. (2011) and thus should be consistent with the widely used GAMA mass functions (Baldry et al. 2012; Kelvin et al. 2014; Wright et al. 2017; Driver et al. 2022), which provide the deepest mass function estimates for galaxies in the local Universe. The Taylor et al. (2011) mass estimates are based on the Bruzual & Charlot (2003) stellar population models and assume a Chabrier (2003) mass function and the dust extinction law from Calzetti et al. (2000). There is an extensive comparison of how this color–*M/L* relation compares to earlier color–*M/L* relations in Taylor et al. (2011). In particular, their Figure 13 shows that their relation is ~0.2 dex below the Bell et al. (2003) relation (due to differences in the initial mass function; IMF) and a similar normalization but significantly different slopes from the Zibetti et al. (2009) relation due to differences in the weighting of stellar population models. Here we compare our derived masses against dynamical mass estimates from ATLAS^{3D} in Cappellari et al. (2011) and IR-based mass estimates from the S4G survey (Sheth et al. 2010).

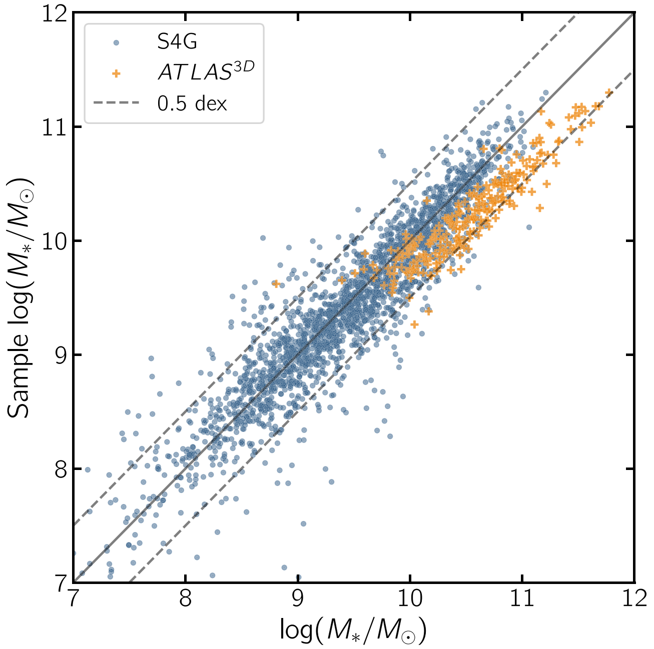


Figure 7. Position offsets and their potential effects on active fraction as a function of galaxy mass and type. Left–Comparison between the X-ray active fractions calculated using different separation radii to crossmatch with Chandra observations. The solid line shows the active fraction using a matching radius of 1″, the dashed lines use a radius of 2 ″, and the dotted lines use a radius of 3″. The red and blue lines show early- and late-type galaxies, respectively. Right–Fraction of catalog galaxies with large position offsets between our catalog and NED as a function of mass and type. The solid line shows the fraction with an offset >1″, the dashed lines show the fraction with an offset >2″, and the dotted lines show the fraction with an offset >3″. Large offsets are seen much more commonly in late-type galaxies than in early-type galaxies, especially at lower masses.

4.1.1. ATLAS^{3D} Dynamical Mass Comparison

ATLAS^{3D} is a volume-limited sample of 260 nearby early-type galaxies in Cappellari et al. (2011) within 42 Mpc. Dynamical masses were derived using Jeans anisotropic modeling of the integral field kinematics in Cappellari et al. (2013a) and Cappellari et al. (2013b). We compare our masses to those derived from the stellar-only dynamical $(M/L)_{\mathrm{stars}}$ estimates from Cappellari et al. (2013b), shown in Figure 7. The ATLAS^{3D} masses are mostly higher than our masses, with a median offset of $\tilde{0}$.34 dex. The offset is near zero in the lowest-mass galaxies (log($M_\star/M_\odot$)<10) and is highest (nearly 0.5 dex) at the highest masses. These offsets are consistent with the findings of Cappellari et al. (2012), who found that the vast majority of ATLAS^{3D} galaxies had dynamical mass estimates that were higher than stellar population model estimates based on a Chabrier (2003) IMF, with the typical offset roughly a factor of 2 (0.3 dex) higher. They found that this offset strongly depends on the observed velocity dispersion, with the highest dispersion and mass galaxies having the largest offsets, just as observed here. This means that overall, this comparison suggests that our mass estimates are accurate for their assumption of a Chabrier IMF, but this assumption may be a poor one especially for the most massive early-type galaxies, and our stellar mass estimates may significantly underestimate the masses of these galaxies.

4.1.2. S4G IR-based Mass Comparison

The Spitzer Survey of Stellar Structure in Galaxies (S4G) is a volume-, magnitude-, and size-limited survey of over 2300 nearby galaxies at 3.6 and 4.5 μm (Sheth et al. 2010). The available mass estimates are based on IR luminosities using the relation of Eskew et al. (2012). This relation uses the resolved stellar mass estimates assuming a Salpeter IMF of the Large Magellanic Cloud (LMC) from Harris & Zaritsky (2009) and combines this with IR fluxes 3.6 and 4.5 μm to convert the luminosity to stellar mass. This relation was applied to all the S4G galaxies to derive stellar masses available in the v. 2 S4G catalog.[17] Using these values, we match 2346/2352 galaxies (99.7%) and plot our derived masses versus the S4G Eskew et al. (2012) masses (after subtracting 0.24 dex to account for the difference expected in mass between a Chabrier and Salpeter IMF; Cappellari et al. 2012) in Figure 7. After the correction for the Salpeter IMF, we see very close agreement between our masses, with a bias of just 0.03 dex (with S4G masses being very slightly more massive). The one-to-one relation extends over a wide range of galaxy masses, and the 1σ scatter around this relation is just 0.24 dex. Overall, we find good agreement between the S4G masses and the masses we derive after accounting for different IMF assumptions.

4.2. Mass Completeness

Many science applications benefit from complete, volume-limited samples of galaxies. To gauge the completeness of our sample, we compare our galaxy catalog to mock galaxy samples created from galaxy stellar mass functions. Above, we have derived masses for our galaxies based on the color–stellar mass relations used in the GAMA survey (Baldry et al. 2010; Taylor et al. 2011); we therefore use the GAMA galaxy stellar mass functions to evaluate the completeness of our catalog.

For the full sample of galaxies, we use the most recent stellar mass function published by GAMA from Driver et al. (2022), which extends down to $M_\star \sim 10^7 M_\odot$. We simulate a galaxy population with a volume of $4/3\pi(50\,\mathrm{Mpc})^3$ based on the Driver et al. (2022) mass function. We then compare the number of galaxies with available masses in our catalog to this mock galaxy sample in 0.2 dex mass bins to calculate the fraction of the expected number of galaxies that we detect. We then translate this fraction to estimate the volume-completeness radius, which we define as the radius out to which we would have a complete sample based on the total number of galaxies in our sample. For instance, if we detect 50% of the number of expected galaxies, this translates into a volume-completeness radius of ~40 Mpc. Figure 8 shows that the catalog is >50% complete at $M_\star \gtrsim 10^{8.5} M_\odot$, which corresponds to an effective volume with radius $R \sim 40$ Mpc. At lower masses, the completeness drops significantly. At the highest masses, the sample also appears to become less complete. We return to this point below.

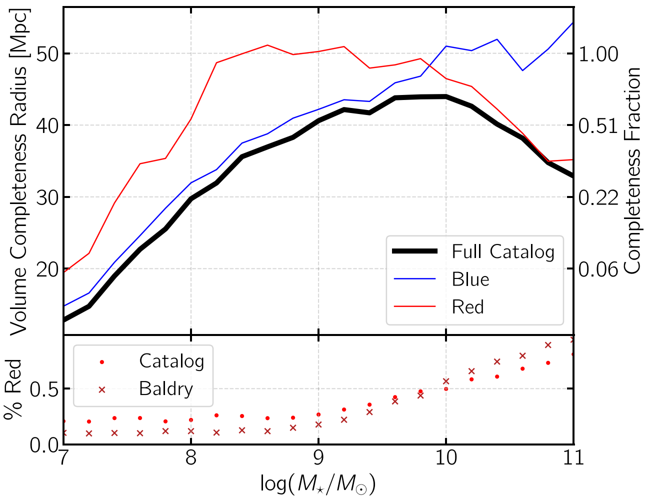


Figure 8. Comparison of our mass estimates (`logmass_best`) against those from S4G (blue; Sheth et al. 2010; Eskew et al. 2012) and ATLAS^{3D} dynamical stellar mass estimates (red; Cappellari et al. 2011). The S4G masses are corrected for a Salpeter to Chabrier IMF by subtracting 0.24 dex for a consistent comparison to our catalog. The solid line shows the one-to-one relation, and the dashed lines indicate an 0.5 dex offset.

The Driver et al. (2022) paper does not have a comparable morphology or color separation to what we have done here, so we use an earlier GAMA paper by Baldry et al. (2012), which is more directly comparable to our color-based morphology separation (i.e., the `color_type` in our sample). We note that the color separation used here and in Baldry et al. (2012) are similar, but not identical. Figure 8 shows that for the late-type galaxies, the volume-completeness radius rises to the expected 50 Mpc at stellar masses of $\sim 10^{10} M_\odot$. The early types achieve a similar completeness at much lower masses, but then the completeness declines at the highest masses. Values greater than one in this figure could be due to (1) a real difference in the number density of galaxies in the local Universe than in the regions used for the Baldry et al. (2012) mass function, or (2) differences in the color separation between our study and theirs. Note that a direct comparison using identical color separations is not possible due to their use of $u - r$ colors in separating their populations and the lack of $u - r$ colors for many galaxies in our sample. The bottom panel of Figure 8 shows the fraction of early-type galaxies as a function of stellar mass in our sample compared to the Baldry et al. (2012) mass function. This shows that the fraction of red galaxies is higher than expected in the local Universe at low masses, but lower at higher masses.

The lack of high-mass galaxies overall, and specifically, high-mass early-type galaxies, is a surprising result because we expect to be complete at the highest masses. This therefore implies that these galaxies are underdense in the local Universe relative to the volumes probed in the GAMA survey. We repeat that our stellar mass estimates should be consistent with GAMA, thus making the difference more likely to represent a real deficit of massive early-type galaxies locally relative to the GAMA volume.

5. Galaxy Morphologies and Group Membership

5.1. Galaxy Morphologies

Many galaxies in the nearby Universe have been morphologically classified using the Hubble and de Vaucouleurs classification systems. Galaxies can be separated into two broad categories: Early-type galaxies are bulge-dominated, elliptical, or lenticular in shape, and host little to no current star formation. Late-type galaxies have a spiral-disk shape with smaller central bulges and undergo current star formation. The numerical Hubble T classification ranges from −6 to 10, with negative values corresponding to early-type galaxies and

---

[17] Available at https://vizier.u-strasbg.fr/viz-bin/VizieR?-source=J/PASP/122/1397/s4g.

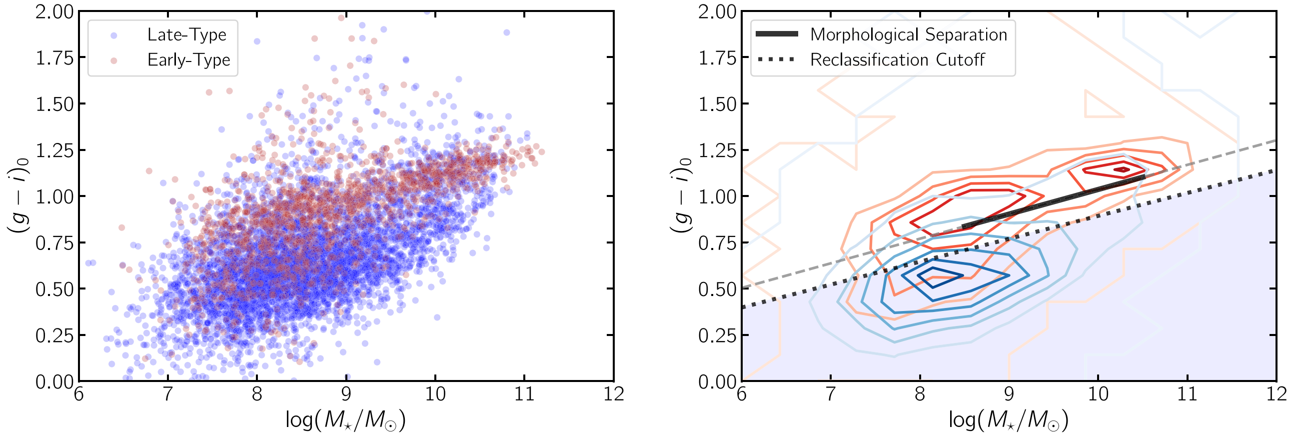


Figure 9. Completeness of our catalog as a function of stellar mass. Top–Mass completeness of our catalog, using 0.2 dex bins. The *Y*-axes show the radius of the effective volume, left, and the completeness fraction, right. The black line compares all catalog galaxies against a <50 Mpc mock sample created using the Driver et al. (2022) galaxy stellar mass function. The colored lines show comparisons between `color_type` separated catalog galaxies against the color-separated mock sample created using the Baldry et al. (2012) galaxy stellar mass functions. Bottom–Percent of galaxies classified as red by our catalog and the Baldry et al. (2012) sample in each bin.

positive values to late types. In addition, early- and late-type galaxies are often separated using galaxy color–magnitude or color-mass diagrams. At a given stellar mass, early-type galaxies are typically associated with redder colors, while late types are associated with bluer colors. In this section, we discuss the source of our morphological type information. We then also describe how we translate this into a color-based separation of early- and late-type galaxies. We then evaluate the X-ray active fraction as a function stellar mass and morphology/color in Section 6.

We use the numerical *T*-type translation of the morphological classifications in this work. Both HyperLeda and LVG provide numerical *T*-types for 13744 of the 15611 (89%) galaxies in our catalog. Of the remaining galaxies, 603 sample galaxies have available color information. Furthermore, as shown in the left panel of Figure 9, the early- and late-type galaxies (also known as the red sequence and blue cloud) show considerable overlap, especially for low-mass ($<10^9\,M_\odot$) galaxies. In order to give all sample objects an appropriate morphology, we worked to minimize the potential for morphological misclassification and calculated a color-based type for all galaxies with available color information.

To separate galaxies into early and late types using the color-mass diagram, we first calculated a morphological separation line by minimizing the misclassification of galaxies with both morphological *T*-types and NSA colors. We separated galaxies into early and late type based on their *T*-type (with $T>0$ being late type), and we fit a line to the color-mass diagram that minimizes the number of misclassifications, i.e., late types redder than the line, and early types bluer than the line. To fit this line, we first divided the data into 0.25 dex mass bins and calculated the fraction of misclassifications as a function of $(g-i)_0$ color. We then plotted the $(g-i)_0$ value with the minimum misclassification in each bin and found the best-fit line to establish our color-based morphological separation line. We note that we only fit the line between $\log(M_\star/M_\odot)$ of 8.5 to 10.5, where galaxy mass estimates are more robust. The equation for this line is

$$(g-i)_0 = 0.133{*}\log(M_\star/M_\odot) - 0.294. \tag{5}$$

This equation appears as the upper dashed line in the right panel of Figure 9.

Visual inspection of SDSS imaging of the misclassified galaxies using our morphological separation line suggested that while most red late-type galaxies showed distinct signatures of being dusty late-type galaxies (i.e., spiral arms or clumpy regions of active star formation), many blue early-type galaxies were in fact not properly classified and also showed these signatures of ongoing star formation. This was especially true for early-type galaxies that had much bluer colors than our morphological separation line. To better separate early- and late-type galaxies, we decided to reclassify very blue early-type galaxies. We created a second reclassification cutoff line below which we reclassified these galaxies in our `best_type` column (as described below). To create this line, we sorted the early-type contaminants by color in 0.25 dex mass bins. For each mass bin, we visually classified these contaminants as early or late type and then calculated the threshold color such that legitimate misclassification was minimized within the cutoff. Then, as above, we fit the reclassification cutoff line to the colors determined in each bin to obtain

$$(g-i)_0 = 0.124 * \log(M_\star/M_\odot) - 0.344. \tag{6}$$

This equation appears as the lower dotted line in the right panel of Figure 9.

To apply our color-based reclassification across our full sample, we needed to transform our morphological separation and reclassification cutoff lines into $(g-r)_0$, $(B-V)_0$, and $(B-R)_0$. Note that to derive the $(B-R)_0$, we added $(B-R)_0$ data from Cook et al. (2014) for the galaxies in our sample. Using overlapping galaxies and using `RANSAC` as described above, we used the following color transformations:

$$(g-r)_0 = 0.562 * (g-i)_0 + 0.109, \tag{7}$$

$$(B-V)_0 = 0.625 * (g-i)_0 + 0.131, \tag{8}$$

$$(B-R)_0 = 0.898 * (g-i)_0 + 0.133. \tag{9}$$

Applying these color-transformations, we obtain the following morphological separation lines:

$$(g-r)_0 = 0.075 * \log(M_\star/M_\odot) - 0.056, \tag{10}$$

$$(B-V)_0 = 0.083 * \log(M_\star/M_\odot) - 0.052, \tag{11}$$

$$(B-R)_0 = 0.119 * \log(M_\star/M_\odot) - 0.131, \tag{12}$$

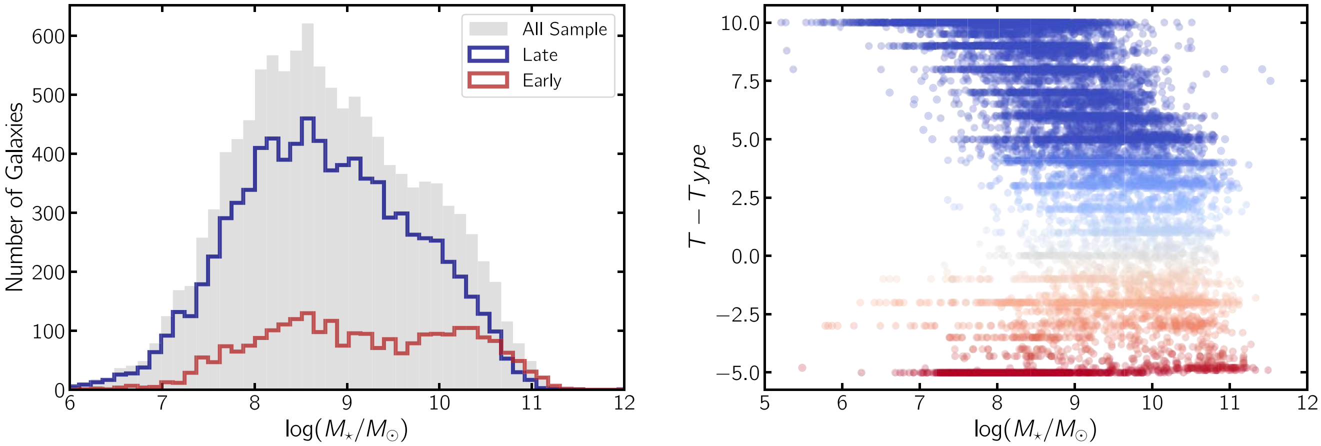


Figure 10. Morphological classifications of our galaxies in the color–mass plane. Left–A color–mass diagram of 6677 galaxies with *g*–*i* data from NSA and morphologies from HyperLeda and LVG. Right–The same color–mass distribution (early-type galaxies in red contours, and late-type galaxies in blue), with the black line showing our linear fit to most cleanly separate early- and late-type galaxies. This fit is extrapolated to higher and lower masses (gray line). The dotted line and the shaded blue regions show where we reclassify blue early-type galaxies as late-type galaxies based on visual inspection of these galaxies (see Section 5).

and the reclassification lines

$$(g - r)_0 = 0.070 * \log(M_\star/M_\odot) - 0.126, \tag{13}$$

$$(B - V)_0 = 0.062 * \log(M_\star/M_\odot) - 0.098, \tag{14}$$

$$(B - R)_0 = 0.111 * \log(M_\star/M_\odot) - 0.242. \tag{15}$$

For our final catalog, we provide several morphological types. In addition to our catalog columns for numerical *T*-types and their source, we created two type columns that broadly categorize galaxies into "early" or "late." Our `color_type` classifies a galaxy as early if its color lies above the morphological separation line and as late if it lies below the line. Our recommended `best_type`column follows this process:

1. Most galaxies are classified based on their *T*-type from HyperLeda or Karachentsev et al. (2013) with $t \leqslant 0$ = "early" and $t > 0$ = "late."
2. The 387 early-type galaxies with bluer colors than the reclassification cutoff are reclassified as late-type galaxies.
3. 603 galaxies without a morphology are assigned their color-based type. We reclassify each galaxy based on the color used to estimate the mass in the final catalog.

In total, we have `best_type` values for 14347 out of 15611 galaxies and `color_type` values for 11740 galaxies. We show the masses of galaxies separated into early and late type in the left panel of Figure 10, and we also include the masses of all galaxies with *T*-types in the right panel.

5.2. Galaxy Environments: Group Catalog Assignment

The galaxies in the local Universe span a wide range of environments, from dense clusters such as the Virgo galaxies to galaxies in voids. One common way of identifying galaxies in dense environments is through group catalogs. To enable easy identification of galaxies in groups in our survey, we use the Lambert et al. (2020) group catalog based on the 2MRS survey. This group catalog identifies 3022 groups from the 2MRS catalog (which extends well beyond our 50 Mpc limit) using a friends-of-friends algorithm.

To match our galaxy catalog, we first used Table 3 from Lambert et al. (2020), which gives individual galaxies that are part of a group or subgroup using the friends-of-friends algorithm described in the paper. We crossmatched our sample to galaxies with $v_{\rm cmb} < 3867$ km s^{-1} (based on our highest-redshift galaxy), resulting in 2463/2864 unique matches within 30″. We wished to ensure that we also matched low-mass sample galaxies that might not be included in the Lambert et al. (2020) catalog due to the magnitude limit of 2MRS. To do this, we used Table 1 from Lambert et al. (2020), which includes group positions, comoving distances, velocity dispersion, and radius/R_{200} estimates. We crossmatched our entire sample to find the nearest group for each galaxy, considering sample galaxies to be part of a group when they fulfilled two requirements: (i) the velocity for our sample galaxy must be within twice the velocity dispersion of the group, and (ii) the angular separation between between the galaxy and its nearest group matches was smaller than the angular radius of that group, ($d \leqslant R_{200}$). Using this three-dimensional approximation, we flagged an additional 790 sample galaxies as likely group members. As expected, the median $\log(M_\star/M_\odot)$ for the distinct group-method matches is ∼8.4, compared to ∼10 for the direct galaxy matches.

We assigned Lambert et al. (2020) group IDs to 3253 sample galaxies, giving precedence to those with direct galaxy matches. This simultaneously flags galaxies as being part of a group, and easily allows users to seek further group information (including group radius and velocity dispersion) by cross correlating our catalog with the catalogs available from Lambert et al. (2020) on the publisher's website.

The Virgo cluster is indicated in the Lambert et al. (2020) catalog as cluster 2987. Of the 347 galaxies with this group identification, we find that 276 of them (79%) are included in Virgo based on the distance sources in Section 3. This is a small fraction of the 907 total galaxies that we identify as being part of Virgo in our catalog because the Lambert et al. (2020) catalog is based on 2MASS and thus misses many fainter Virgo galaxies.

6. X-Ray Active Fraction Measurement

6.1. Background

The initial motivation for creating this sample was to quantify the demographics of central black holes in nearby lower-mass galaxies. These demographics can provide constraints on the formation mechanisms of central massive black holes (see the recent review by Greene et al. 2020). X-ray measurements of AGN in nearby galaxies $\lesssim$50 Mpc can provide deep and clean enough observations to constrain the occupation fraction of black holes in galaxies below $10^{10}\,M_\odot$. Results on large samples of early-type galaxies (Miller et al. 2015; Gallo & Sesana 2019) suggest a high occupation fraction for 10^9–$10^{10}\,M_\odot$ galaxies, in agreement with dynamical measurements in a handful of nearby early-type galaxies (Nguyen et al. 2018, 2019). However, a majority of galaxies below $10^{10}\,M_\odot$ are late-type galaxies (Baldry et al. 2012), and thus far, no suitable galaxy sample exists to compare the X-ray active fraction (and thus occupation fraction) of late-type galaxies to the studies of early-type galaxies at masses below $M_\star \sim 10^{10}\,M_\odot$. Recent compilations of archival X-ray data in nearby galaxy samples cut out most low-mass nearby galaxies by requiring redshift-independent distance measurements (She et al. 2017a, 2017b), or by including only brighter samples dominated by massive galaxies (Bi et al. 2020). The X-ray active fractions of more distant dwarfs (between z of 0.08 and 2.4) have also been studied by Mezcua et al. (2018) in the COSMOS field. Our new 50 Mpc Galaxy Catalog provides the best starting point for constraining the local X-ray active fraction and the occupation fraction of central black holes. We specifically focus on trying to obtain a sample that is as large and complete as possible down to lower masses (i.e., $10^8\,M_\odot$), with stellar mass estimates and Hubble-type measurements based on color index. Below, we quantify the X-ray active fraction using our new catalog.

The detection of a nuclear X-ray source above 10^{42} erg s^{-1} (the "classical" AGN threshold) is invariably associated with accretion-powered emission from a massive (i.e., nonstellar) black hole. Below this threshold, and even more so, below 10^{40} erg s^{-1}, emission from bright X-ray binaries could contribute to the detected signal. The expected luminosity from low-mass X-ray binaries, in particular, scales with stellar mass (Lehmer et al. 2019), which makes spatial resolution the key for discriminating between massive black holse with a low Eddington ratio versus contaminants. Chandra is thus preferable to XMM-Newton for searching and identifying massive black holes with luminosities lower than traditional AGN because its resolution is higher by a factor of 4.

We present the Chandra-based X-ray active measurements of our galaxy catalog here, without an attempt to model the distribution or account for contamination from X-ray binaries. However, in a subsequent paper, we will use these data to constrain the L_X–$M_\star$ relation and the occupation fraction as a function of galaxy stellar mass, including estimates of contamination from X-ray binaries (E. Gallo et al. 2023, in preparation).

6.2. 50 Mpc Active Fraction Measurement

We determined which galaxies in our catalog have available Chandra X-ray data using the Chandra Source Catalog 2.0 (Evans et al. 2010, and I. N. Evans 2023, in preparation), which includes Chandra observations made public before the end of 2014. We searched for X-ray sources coincident with the nuclei of our galaxies using their catalog `ra` and `dec`. Note that our choice of coordinates has been discussed in Section 2.1.2, and the effect of uncertainties on these central coordinates is considered in more detail in Section 6.2.2. Using Chandra's CSCView,[18] we retrieved two tables from a crossmatch with our catalog: (1) a table of limiting sensitivities for all 1506 objects with Chandra data, and (2) a table of detections within a given angular separation (1″–3″) of the galaxy `ra` and `dec`. The second table of detections contains 291 unique galaxies with matching detections when using a 1″ search radius (note that 20 objects returned multiple sources). We used our distance estimates (`best_dist`) to calculate X-ray luminosities from the 0.5 to 7 keV fluxes. These have a median $\log(L_X/\mathrm{erg\,s^{-1}})$ of 39.21 (hereafter, we just refer to this as $\log(L_X)$), with a full range of $\log(L_X)$ from 33.19 to 42.26. While some of these sources are likely contaminants from X-ray binaries and not accreting massive black holes, we attempted to minimize this contamination in two ways: (1) We restricted our X-ray activity analysis below to the 249 (85%) of sources that have available $\log(M_\star/M_\odot)$ and $\log(L_X) > 38.3$ (as in Miller et al. 2015; Gallo & Sesana 2019), and (2) we used a 1″ match as our default matching radius. We note that previous work on X-ray binary contamination in similar Chandra observations has shown that only a small fraction (∼10%) of nuclear X-ray detections with $\log(L_X) > 38.3$ are likely to be contaminants (e.g., Gallo et al. 2010; Miller et al. 2012; Foord et al. 2017). We include the columns `chandra_observation` and `chandra_detection` to flag matched catalog galaxies using our default 1″ matching, and we also include a `chandra_detection_3arcsec` for galaxies with sources within 3″ (see Section 6.2.2); the X-ray luminosity is given in the `log_lx` column.

Of the 249 galaxies we used for our active fraction analysis, 159 have available SDSS imaging. To visualize the X-ray sources in our sample, we show 42 of these galaxies in Figure 11. In the first three rows of the figure, we show the 18 lowest-mass SDSS galaxies, and we present a sampling of galaxies at higher masses in the bottom half. The 18 lowest-mass galaxies have $\log(M_\star/M_\odot)$ from 7.99 to 9.61; an additional 4 galaxies in this mass range lack available SDSS imaging.

6.2.1. Local Active Fraction as a Function of Galaxy Mass and Type

We calculate the X-ray active fraction by dividing the number of galaxies with Chandra detections by the number of galaxies with Chandra observations. We use a 1″ matching radius to limit the impact of X-ray binary contamination and take advantage of Chandra's excellent resolution (e.g., Gallo et al. 2010), however, we examine the results of larger matching radii in the next subsection. The results of this X-ray active fraction binned by mass and separated by morphological type are shown in Figure 12. Although some detections may be caused by nuclear X-ray binaries, this fraction serves as an upper limit of AGN occupation for our catalog.

We observe an apparent difference in detection fraction between early- and late-type galaxies, particularly within the mass range $9 \lesssim \log(M_\star/M_\odot) \lesssim 10.5$. We examine the significance of this difference following the methods of Hoyer et al. (2021). We calculate p-values under the null hypothesis that both early- and late-type galaxies have the same X-ray active fraction. We split the data into 0.25 dex bins from $7.75 \leqslant \log(M_\star/M_\odot) \leqslant 11.25$. We use the total X-ray active fraction in each bin as an estimator of the binomial success probability. Using this fraction, we draw n samples from a binomial distribution, where n is the number of galaxies per bin in the early- and late-type subsamples. We repeat this exercise 10^6 times and use the fraction of estimates that match or exceed our observed difference between early- and late-type galaxies

---

[18] http://cda.cfa.harvard.edu/cscview/

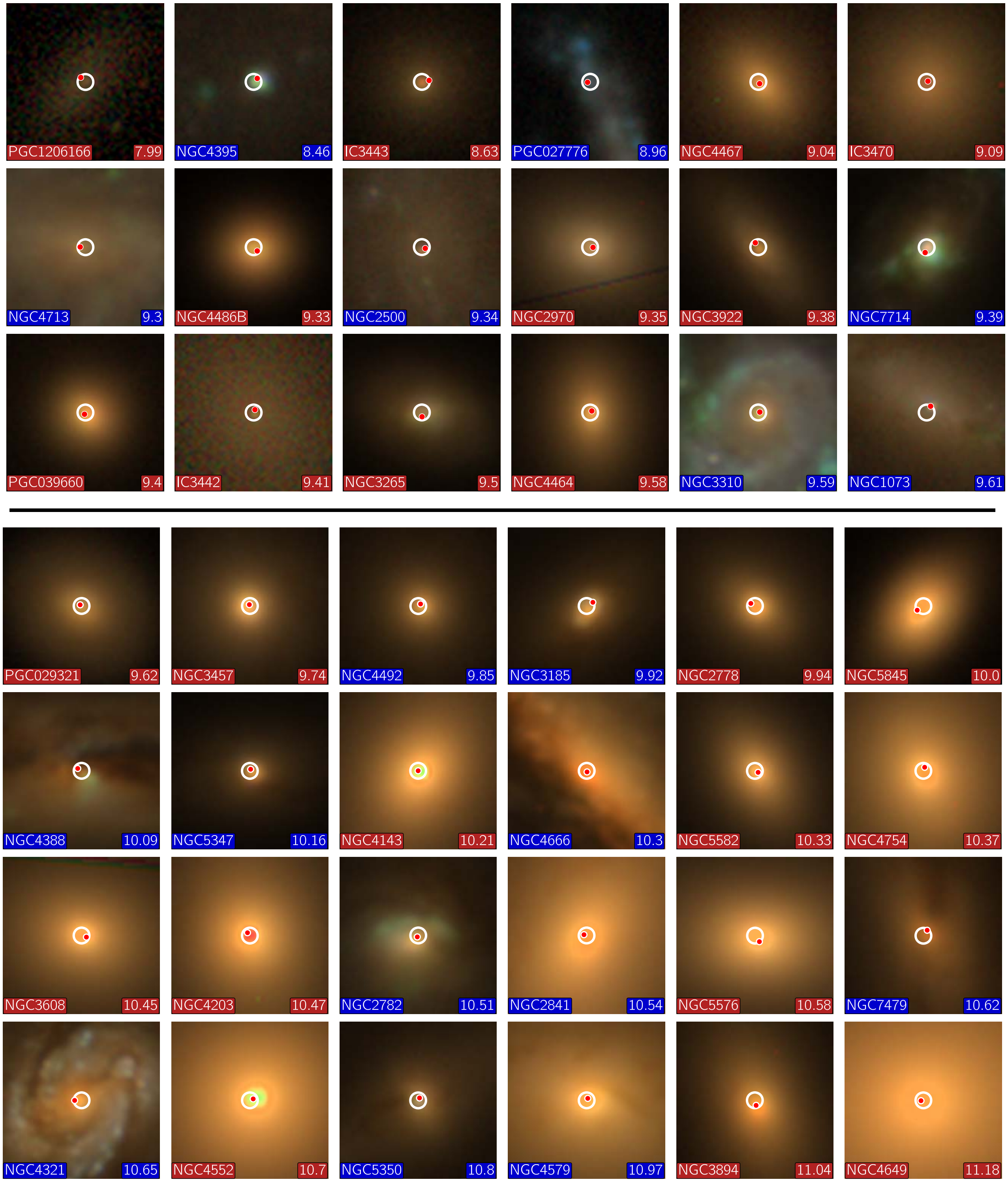
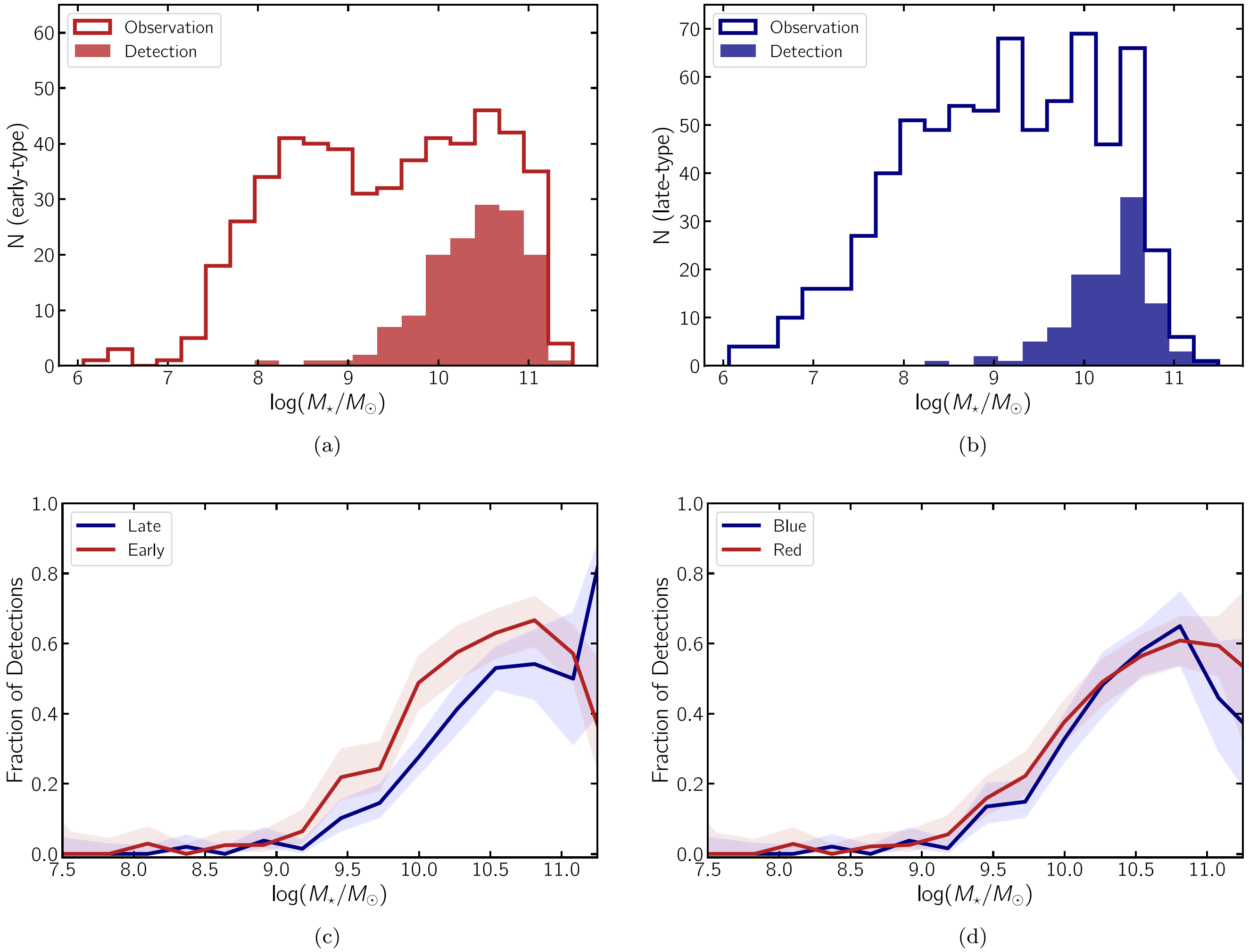


Figure 11. Mass distribution of galaxies as a function of their morphology. Left–Distribution of $\log(M_\star/M_\odot)$ for the full catalog (shaded histogram), using `best_type` as described in Section 5.1 to separate into early (red) and late types (blue). Right–Comparison of numerical T-Type vs. $\log(M_\star/M_\odot)$ for all catalog galaxies with an available `t_type`. The color gradient is included for visual representation of the galaxy type and coincides with the T-Type value.

to estimate a *p*-value for each bin. Using the Fisher method (Fisher 1992), we combine all *p*-values into a single parameter under the assumption that the null hypothesis is true and that the data in each mass bin are independent. The final *p*-value is 0.0004 over the entire mass range, thus suggesting that the observed enhancement of early-type galaxies is significant at a level $>3\sigma$. This significance increases, for instance, when we consider just the galaxies between $9 \leqslant \log(M_\star/M_\odot) \leqslant 10.5$. Thus, it appears there is a significant enhancement of X-ray detections in early-type versus late-type galaxies. We examine belwo whether this result is due to a higher AGN fraction in early-type galaxies.



(a) (b)

(c) (d)

Figure 12. Images of catalog galaxies with Chandra sources within 1″ of their centers that have $\log(L_X) > 38.3$ and have available SDSS imaging. The white circle is centered on the catalog's galaxy coordinates and shows the 1″ radius used to match to Chandra detections, while the red dot marks the Chandra source position. Each image has a width of 20″. The galaxies are arranged in order of increasing $\log(M_\star/M_\odot)$. The top three rows show the 18 galaxies with the lowest $\log(M_\star/M_\odot)$ and represent 18/22 lowest-mass galaxies with X-ray sources. The bottom four rows show galaxies at higher masses sampled evenly in $\log(M_\star/M_\odot)$. The image labels provide the catalog `objname`, `logmass`, and `best_type` (red are early types, and blue are late types).

6.2.2. Local Active Fraction Uncertainties and Systematics

One significant potential uncertainty is the locations of galaxy centers, and whether our catalog coordinates accurately reflect the true center of each galaxy. To gauge the impact of this uncertainty, we conducted the same analysis using a 2″ and 3″ separation constraint for Chandra detection matching. As shown in the left panel of Figure 13, larger matching radii increase the X-ray active fraction for both types overall. The difference between the early- and late-type active fraction depicted in Figure 12 still exists for larger separation constraints, but the statistical significance decreases to $p = 0.0019$ for 2″ and $p = 0.0104$ for 3″.

The decreasing difference in X-ray active fractions between early and late types with larger matching radii could be explained if the centers of late-type galaxies were less certain due to their lower surface brightness and more complicated morphology. The central surface brightness does depend on galaxy mass and luminosity (e.g., Blanton et al. 2003; Ferrarese et al. 2006), with decreasing surface brightness in lower-mass galaxies below roughly the luminosity and mass of the Milky Way. Late-type galaxies generally also have a lower surface brightness than early-type galaxies (e.g., Blanton et al. 2003). In addition, many of our late-type galaxies are irregulars; these dominate our late-type galaxy number counts below $\log(M_\star/M_\odot) = 8.0$. Finally, distance plays a role in our ability to accurately locate the angular center of our galaxies, with more distant galaxies likely having more accurate angular coordinates.

As noted in Section 2.1.1, we see significant discrepancies between our catalog positions (mostly from HyperLeda) and the best available positions from NED. These offsets can give us a sense of how uncertain the nuclear positions of each galaxy are. In the right panel of Figure 13, we show the fraction of galaxies with offsets larger than 1″, 2″, and 3″ as a function of galaxy type (early versus late) and stellar mass. The fraction of galaxies with significant position differences between NED and our catalog is noticeably larger for late-type galaxies than for early-type galaxies in the mass range $7.5 \lesssim \log(M_\star/M_\odot) \lesssim 9.5$. For a 1″ position match, the differences are large enough (>20% at some masses between early and late types) to explain the lower active fractions we see in late-type galaxies. These larger positional uncertainties are likely tied to the lower surface brightness and irregular morphology among our late-type galaxies. However, at higher masses ($9.5 \lesssim \log(M_\star/M_\odot) \lesssim 10.5$), the positional uncertainty differences become smaller than the active fraction differences.

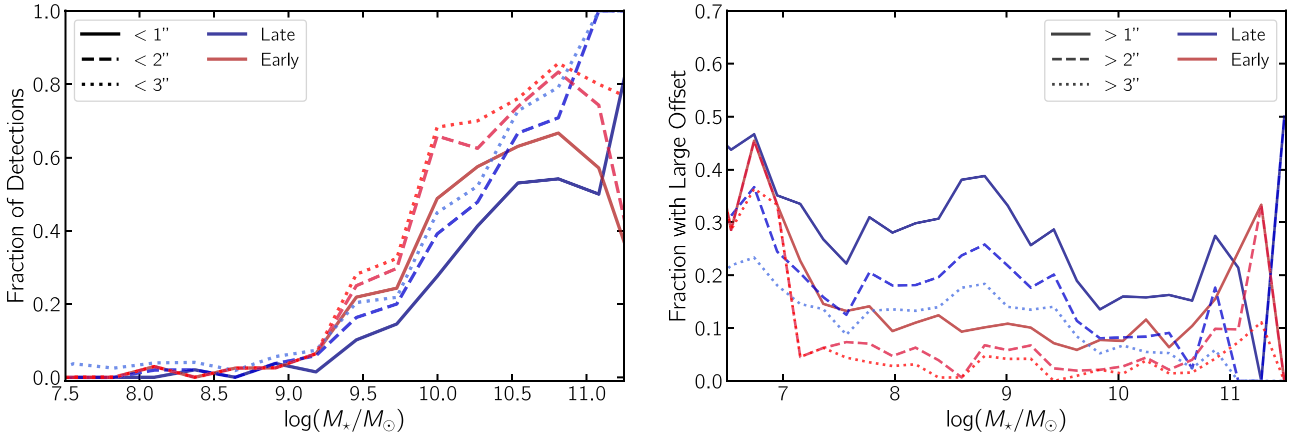


Figure 13. X-ray active fraction of catalog galaxies as a function of galaxy mass and type, using a 1″ matching radius. Top row—The $\log(M_\star/M_\odot)$ distribution of catalog galaxies with Chandra data. The solid line shows galaxies with Chandra observations, and the filled histogram shows galaxies with Chandra detections with $\log(L_X) > 38.3$ for (a) early types and (b) late types. Panel (c)–X-ray detection fraction as a function of $\log(M_\star/M_\odot)$ and separated by `best_type`, a combination of morphological and color-based types (see Section 5). Panel (d)–X-ray detection fraction as a function of $\log(M_\star/M_\odot)$ and separated using `color_type`. The shaded areas in the bottom panels show Agresti-Coull binomial confidence intervals.

To further examine the possibility that uncertain nuclear coordinates impact our active fraction measurements, we visually examined all 17 galaxies (16 of which are late-type galaxies) with $\log(M_\star/M_\odot) \lesssim 9.5$ that had X-ray sources within 3″ of their centers, but not within 1″. In 11 of these 17 galaxies, the X-ray sources are clearly offset from the galaxy centers. In another 5 galaxies, it was ambiguous whether the X-ray source was nuclear or not, while in only one galaxy (NGC 598/M33), it was clear that the X-ray source was nuclear, and our nuclear position is offset by ∼1″ from the true center of the galaxy (this galaxy has a known nuclear star cluster that is visible in SDSS imaging). This shows that accurate nuclear positions are an important factor to consider when examining the presence of AGN in nearby galaxies, especially in lower-mass late-type galaxies. However, errors in nuclear positions apparently do not fully account for the difference we see between early- and late-type galaxy active fractions.

6.2.3. Local Active Fraction as a Function of Galaxy Mass and Color

We also repeated the process to calculate X-ray active fractions using only color information to separate early- and late-type galaxies (using the `color_type` described in Section 5.1). This is shown in the lower right panel of Figure 12. When we classify type based on color alone, the X-ray active fractions appear similar in all mass bins. This change from our primarily morphology-based types results from a suppression of the X-ray active fraction in red galaxies relative to early-type galaxies. This is primarily caused by galaxies that are classified as late-type, but that have a red color. These galaxies lie above our color-based type separation (Figure 9). Visual inspection shows that the majority of these red late-type galaxies are dusty spiral galaxies, many of which are viewed at high inclinations. This suggests that despite the potential issues with visual morphological classification (see Section 5.1), they provide valuable information about the galaxy type over and above the color information. This is particularly true for red late-type galaxies, which are quite numerous and have accurate morphological classifications.

The active fraction measurement we present here is the first step in a more sophisticated occupation fraction analysis (including accounting for XRB contamination) that will be presented in a follow-up paper (E. Gallo et al. 2023, in preparation).

7. Conclusions

In this paper, we present a robust catalog of 15424 nearby galaxies within 50 Mpc with self-consistent measurements of distance, mass, and morphological type. This catalog was constructed primarily to enable measurements of the X-ray active fraction and occupation fraction, but the utility of this catalog spans a much broader range of possible applications. In particular, it provides a more complete catalog of lower-mass galaxies than were included in recent similar studies (She et al. 2017a; Bi et al. 2020).

Our catalog combines galaxies from HyperLeda, the NASA-Sloan Atlas, and the Local Volume Galaxy catalog. We also include additional photometry from the Siena Galaxy Atlas and from NED. For galaxies with data from multiple sources, we compared the values to the literature and included the best possible measurements in the catalog. We extensively cleaned the catalog, including individual investigation of extreme outliers. Through individual investigation of SDSS imaging for NSA nonmatches, we discovered that ∼13% of the nearby NSA galaxies are contaminants.

We compile extinction-corrected magnitude, luminosity, and color information for 11674 galaxies (76% of our full sample). We use the galaxy colors to estimate masses by creating self-consistent color mass-to-light ratio relations in four bands. We also provide errors on our mass estimates that include both the scatter around our color-M/L relations and distance errors. We also compare our masses to the photometric masses from S4G and to the dynamical masses from ATLAS^{3D}.

Galaxy morphologies are available for 13744 galaxies. We combine this information with galaxy masses and colors to fit a line that optimizes the separation of early- and late-type galaxies. Through visual inspection, we find that early-type galaxies bluer than this line are frequently incorrectly classified, and we therefore reclassify those that fall more than ∼0.13 magnitudes below our morphological separation line. Red late-type galaxies are typically truly late-type galaxies; many of these are edge-on dusty galaxies.

In addition, we

1. combine best estimates for flow-corrected redshift-based distances with redshift-independent distances, with a focus on minimizing the distance error, and including special treatment for galaxies in the Virgo Cluster;
2. use galaxies with overlapping color information to calculate and provide empirical transformations of $B - V$, $B - R$, and $g - r$ colors into Sloan g–i; and
3. identify galaxies in the catalog that belong to dense groups with the help of the Lambert et al. (2020) group catalog.

Last, we present a preliminary analysis of X-ray sources and find that 249 out of the 1506 galaxies with existing Chandra observations have X-ray detections with $\log(L_{\rm X}) > 38.3$ and available stellar mass estimates. We present evidence that the X-ray active fractions differ for early- and late-type galaxies. In particular, we find that the active fraction is higher for early types within the mass range $9 < \log(M_\star/M_\odot) < 10.5$. We show that increased astrometric uncertainties in late-type galaxies than early types may partially be responsible for this difference in active fractions by galaxy type. In an upcoming paper (E. Gallo et al. 2023, in preparation), we use this catalog to conduct the first robust analysis of the AGN occupation fraction down to $\log(M_\star/M_\odot) = 8$ both overall and separated by morphological type.

Acknowledgments

We would like to thank Michael Blanton for his useful clarifications on NSA photometry. We also thank Ian Steer for helpful correspondence and assistance concerning NED-D. A. C.S. acknowledges support from NSF grant AST-2108180.

This research has made use of the NASA/IPAC Extragalactic Database (NED), which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

We acknowledge the usage of the HyperLeda database (leda.univ-lyon1.fr).

We acknowledge the usage of the Siena Galaxy Atlas, based on the Legacy Survey. The Legacy Surveys consist of three individual and complementary projects: the Dark Energy Camera Legacy Survey (DECaLS; Proposal ID #2014B-0404; PIs: David Schlegel and Arjun Dey), the Beijing-Arizona Sky Survey (BASS; NOAO Prop. ID #2015A-0801; PIs: Zhou Xu and Xiaohui Fan), and the Mayall z-band Legacy Survey (MzLS; Prop. ID #2016A-0453; PI: Arjun Dey). DECaLS, BASS and MzLS together include data obtained, respectively, at the Blanco telescope, Cerro Tololo Inter-American Observatory, NSFs NOIRLab; the Bok telescope, Steward Observatory, University of Arizona; and the Mayall telescope, Kitt Peak National Observatory, NOIRLab. Pipeline processing and analyses of the data were supported by NOIRLab and the Lawrence Berkeley National Laboratory (LBNL). The Legacy Surveys project is honored to be permitted to conduct astronomical research on Iolkam Duag (Kitt Peak), a mountain with particular significance to the Tohono Oodham Nation.[19]

The Siena Galaxy Atlas was made possible by funding support from the U.S. Department of Energy, Office of Science, Office of High Energy Physics under Award Number DE-SC0020086 and from the National Science Foundation under grant AST-1616414.

Appendix
Full Catalog

Table A1 shows an example of important catalog data, including the names and descriptions of all the columns.

---

[19] Full acknowledgment available at https://www.legacysurvey.org/acknowledgment/#siena-galaxy-atlas

Table A1

Stub of Catalog, Shortened to Include Only the Most Important Columns

`objname`	`ra` (deg)	`dec` (deg)	`v_h` (km/s)	`t_type`	`best_type`	`bestdist` (Mpc)	`bestdist_error` (Mpc)	`logmass` ($M_\odot$)	`logmass_error` ($M_\odot$)	`logmass_src`
NGC2788B	135.898976	−67.969798	1444.9	3.2	late	17.451	6.903	9.1114	0.3634	*B-V*
NGC2793	139.194264	34.431685	1724.49	8.7	late	26.016	7.16	9.1861	0.2454	*g-i*
NGC2798	139.345353	41.99981	1726.0	1.1	late	26.745	6.563	9.983	0.2176	*g-i*
NGC2799	139.379054	41.994172	1870.86	8.5	late	30.689	6.964	9.3076	0.2006	*g-i*
NGC2805	140.084902	64.102968	1730.35	6.9	late	27.77	6.506	9.8968	0.2073	*g-i*
NGC2810	140.518791	71.843972	3479.7	−4.9	early	51.345	6.415	10.8048	0.1354	*g-r*
NGC2811	139.046301	−16.312679	2355.6	1.1	late	28.1	1.29	10.617	0.1354	*B-V*
NGC2814	140.297846	64.25236	1592.0	3.1	late	25.786	6.638	9.1279	0.2287	*g-r*
NGC2815	139.082217	−23.633299	2545.3	2.9	late	36.977	5.217	10.702	0.1354	*B-V*
NGC2820	140.441039	64.257886	1575.01	5.3	late	25.507	6.642	9.7143	0.2315	*g-i*

Notes. `objname`—Common galaxy name; `ra`—R.A. from combined sources; `dec`—decl. from combined sources; `v_h`—heliocentric radial velocity; see Section 3.1; `t_type`—numerical Hubble T-type; `best_type`—combined galaxy type; see Section 5.1; `bestdist`—our chosen best-distance estimate; `bestdist_error`—error for the best-distance estimate; `logmass`—compiled best log ($M_*/M\odot$) estimate; `logmass_error`—error on logmass; `logmass_src`—color used for the best-mass estimate: g–i, $g-r$, $B-V$, $B-R$; . Columns found in the full catalog, available from https://github.com/davidohlson/50MGC: `pgc`—PGC number (from HyperLeda); `nsa_id`—NASA-Sloan Atlas NSAID number; `group_id`—galaxy is a member of a group, the number is based on the Lambert et al. (2020) group catalog; `ra_nsa`—R.A. provided by NSA; `dec_nsa`—decl. provided by NSA; `ra_ned`—primary R.A. provided by NED; `dec_ned`—primary decl. provided by NED; `d25`—apparent diameter from Karachentsev and HyperLeda; `v_cmb`—radial velocity with respect to CMB radiation; `v_source`—original source for the compiled velocity: HyperLeda, LVG, NSA; `hl_obj`—true for objects in HyperLeda; `lvg_obj`—true for objects in Karachentsev's Catalog of Local Volume Galaxies; `nsa_obj`—true for objects in the NASA-Sloan Atlas; `sga_obj`—true for objects in the Siena Galaxy Atlas; `color_type`—color-based type; see Section 5.1; `a_B_leda`—*B*-band extinction from HyperLeda multiplied by 0.86 to translate into Schlafly, $A_V = 0.769{*}a_{B_{\rm leda}}$ and $A_R = 0.629{*}a_{B_{\rm leda}}$; `a_g_nsa`—*g*-band extinction from the NASA-Sloan Atlas, the *i*-band extinction used is $=0.550{*}a_{g_{\rm nsa}}$; `EBV_irsa`—$E(B-V)$ value from the IRSA dust website; Schlafly values for all galaxies. Used only for the Siena Galaxy Atlas sources, $A_g/{\rm EBV}_{\rm sga} = 3.303$, $A_r/{\rm EBV}_{\rm sga} = 2.285$; `Bt0_leda`—extinction-corrected total *B*-band magnitude from HyperLeda; `BV_color_leda`—$(B-V)$ color from HyperLeda; `B_lum`—*B*-band luminosity derived for non-HyperLeda sources as described in Section 2.2;`gi_color_nsa`—extinction-corrected $(g-i)$ color from the NASA-Sloan Atlas; `i_lum_nsa`—*i*-band luminosity, calculated using $M_{i,\odot} = 4.53$; `gr_color_sga`—extinction-corrected $(g-r)$ color from the Siena Galaxy Atlas; `r_lum_sga`—*r*-band luminosity, calculated using $M_{r,\odot} = 4.65$; `BR_color_ned`—extinction-corrected $(B-R)$ color from NED; `R_lum_ned`—*R*-band luminosity from NED, calculated using $M_{R,/{\rm odot}} = 4.60$; `BMag`—estimated absolute *B*-band magnitude for all galaxies; `gi_color`—estimated g–i color for all galaxies with color measurements; `mag_flag`—flags galaxies with magnitude preference exceptions as described in Section 2.3; `cf3_dist`—distances from the Cosmicflows-3 calculator; `cf3_dist_error`—Error on cf3_dist; `zind_dist`—redshift-independent distances; `zind_dist_error`—Error on zind_dist; `zind_indicator`—redshift-independent distance indicator listed in NEDD; `bestdist_method`—general method used for the best distances: CF3-Z, Karachentsev, NED-D, Mei, Cantiello, EVCC, HyperLeda; `bestdist_source`—source/reference for bestdist; see Section 3; `dist_ned_flag`—flags galaxies with NED-D best-distance exceptions as described in Section 3.2; `logmass_gi`—log ($M_*/M\odot$) from $(g-i)$; `logmass_gr`—log ($M_*/M\odot$) from $(g-r)$; `logmass_BV`—log ($M_*/M\odot$) from $(B-V)$; `logmass_BR`—log ($M_*/M\odot$) from $(B-R)$; `chandra_observation`—true if CSCView crossmatch returned a limiting sensitivity at a matching position as described in Section 6; `chandra_detection`—true if CSCView crossmatch returned flux information within a 1″ matching radius as described in Section 6; `log_lx`—X-ray luminosity $\log(L_{\rm X}/{\rm erg\ s}^{-1})$ calculated from the CSC 0.5–7 keV flux, using a 1″ radius; `chandra_detection_3arcsec`—true if CSCView crossmatch returned flux information within a 3″ matching radius as described in Section 6; `log_lx_3arcsec`—X-ray luminosity $\log(L_{\rm X}/{\rm erg/sec})$ calculated from the CSC 0.5–7 keV flux, using a 3″ match radius.

(This table is available in its entirety in machine-readable form.)

ORCID iDs

David Ohlson https://orcid.org/0009-0004-9457-2495
Anil C. Seth https://orcid.org/0000-0003-0248-5470
Elena Gallo https://orcid.org/0000-0001-5802-6041
Vivienne F. Baldassare https://orcid.org/0000-0003-4703-7276
Jenny E. Greene https://orcid.org/0000-0002-5612-3427

References

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017, PhRvL, 119, 161101
Anand, G. S., Tully, R. B., Rizzi, L., & Karachentsev, I. D. 2019, ApJL, 872, L4
Baldry, I. K., Driver, S. P., Loveday, J., et al. 2012, MNRAS, 421, 621
Baldry, I. K., Glazebrook, K., Brinkmann, J., et al. 2004, ApJ, 600, 681
Baldry, I. K., Robotham, A. S. G., Hill, D. T., et al. 2010, MNRAS, 404, 86
Bell, E. F., McIntosh, D. H., Katz, N., & Weinberg, M. D. 2003, ApJS, 149, 289
Bi, S., Feng, H., & Ho, L. C. 2020, ApJ, 900, 124
Binggeli, B., Sandage, A., & Tammann, G. A. 1985, AJ, 90, 1681
Blanton, M. R., Hogg, D. W., Bahcall, N. A., et al. 2003, ApJ, 594, 186
Blanton, M. R., Lupton, R. H., Schlegel, D. J., et al. 2005, ApJ, 631, 208
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Cappellari, M., Emsellem, E., Krajnović, D., et al. 2011, MNRAS, 413, 813
Cappellari, M., McDermid, R. M., Alatalo, K., et al. 2012, Natur, 484, 485
Cappellari, M., McDermid, R. M., Alatalo, K., et al. 2013b, MNRAS, 432, 1862
Cappellari, M., Scott, N., Alatalo, K., et al. 2013a, MNRAS, 432, 1709
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Carlsten, S. G., Greene, J. E., Beaton, R. L., Danieli, S., & Greco, J. P. 2022, ApJ, 933, 47
Chabrier, G. 2003, PASP, 115, 763
Cook, D. O., Dale, D. A., Johnson, B. D., et al. 2014, MNRAS, 445, 881
Courtois, H. M., Hoffman, Y., Tully, R. B., & Gottlöber, S. 2012, ApJ, 744, 43
Crnojević, D., Sand, D. J., Spekkens, K., et al. 2016, ApJ, 823, 19
Driver, S. P., Bellstedt, S., Robotham, A. S. G., et al. 2022, MNRAS, 513, 439
Erwin, P., Saglia, R. P., Fabricius, M., et al. 2015, MNRAS, 446, 4039
Eskew, M., Zaritsky, D., & Meidt, S. 2012, AJ, 143, 139
Evans, I. N., Primini, F. A., Glotfelty, K. J., et al. 2010, ApJS, 189, 37
Ferrarese, L., Côté, P., Jordán, A., et al. 2006, ApJS, 164, 334
Foord, A., Gallo, E., Hodges-Kluck, E., et al. 2017, ApJ, 841, 51
Freedman, W. L. 2021, arXiv:2106.15656
Gallo, E., & Sesana, A. 2019, ApJL, 883, L18
Gallo, E., Treu, T., Marshall, P. J., et al. 2010, ApJ, 714, 25
Geha, M., Blanton, M. R., Yan, R., & Tinker, J. L. 2012, ApJ, 757, 85
Graziani, R., Courtois, H. M., Lavaux, G., et al. 2019, MNRAS, 488, 5438
Greene, J. E., Strader, J., & Ho, L. C. 2020, ARA&A, 58, 257
Harris, J., & Zaritsky, D. 2009, AJ, 138, 1243
Hoyer, N., Neumayer, N., Georgiev, I. Y., Seth, A. C., & Greene, J. E. 2021, MNRAS, 507, 3246
Hughes, A. K., Sand, D. J., Seth, A., et al. 2021, ApJ, 914, 16
Jester, S., Schneider, D. P., Richards, G. T., et al. 2005, AJ, 130, 873
Kaaret, P., Feng, H., & Roberts, T. P. 2017, ARA&A, 55, 303
Karachentsev, I. D., Makarov, D. I., & Kaisina, E. I. 2013, AJ, 145, 101
Kelvin, L. S., Driver, S. P., Robotham, A. S. G., et al. 2014, MNRAS, 444, 1647
Kim, S., Rey, S.-C., Jerjen, H., et al. 2014, ApJS, 215, 22
Kourkchi, E., Courtois, H. M., Graziani, R., et al. 2020, AJ, 159, 67
Kriek, M., Labbé, I., Conroy, C., et al. 2010, ApJL, 722, L64
Krumholz, M. R., McKee, C. F., & Bland-Hawthorn, J. 2019, ARA&A, 57, 227
Lambert, T. S., Kraan-Korteweg, R. C., Jarrett, T. H., & Macri, L. M. 2020, MNRAS, 497, 2954
Lauberts, A., & Valentijn, E. A. 1989, The Surface Photometry Catalogue of the ESO-Uppsala Galaxies (Garching: European Southern Observatory)
Lehmer, B. D., Eufrasio, R. T., Tzanavaris, P., et al. 2019, ApJS, 243, 3
Makarov, D., Prugniel, P., Terekhova, N., Courtois, H., & Vauglin, I. 2014, A&A, 570, A13
McConnachie, A. W. 2012, AJ, 144, 4
McGaugh, S. S., & Schombert, J. M. 2014, AJ, 148, 77
Mei, S., Blakeslee, J. P., Côté, P., et al. 2007, ApJ, 655, 144
Meidt, S. E., Schinnerer, E., van de Ven, G., et al. 2014, ApJ, 788, 144
Melbourne, J., Williams, B. F., Dalcanton, J. J., et al. 2012, ApJ, 748, 47
Mezcua, M., Civano, F., Marchesi, S., et al. 2018, MNRAS, 478, 2576
Miller, B., Gallo, E., Treu, T., & Woo, J.-H. 2012, ApJ, 747, 57
Miller, B. P., Gallo, E., Greene, J. E., et al. 2015, ApJ, 799, 98
Neumayer, N., Seth, A., & Böker, T. 2020, A&ARv, 28, 4
Nguyen, D. D., Seth, A. C., Neumayer, N., et al. 2018, ApJ, 858, 118
Nguyen, D. D., Seth, A. C., Neumayer, N., et al. 2019, ApJ, 872, 104
O'Donnell, J. E. 1994, ApJ, 422, 158
Pier, J. R., Munn, J. A., Hindsley, R. B., et al. 2003, AJ, 125, 1559
Reines, A. E., Greene, J. E., & Geha, M. 2013, ApJ, 775, 116
Roediger, J. C., & Courteau, S. 2015, MNRAS, 452, 3209
Saglia, R. P., Opitsch, M., Erwin, P., et al. 2016, ApJ, 818, 47
Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
She, R., Ho, L. C., & Feng, H. 2017a, ApJ, 835, 223
She, R., Ho, L. C., & Feng, H. 2017b, ApJ, 842, 131
Sheth, K., Regan, M., Hinz, J. L., et al. 2010, PASP, 122, 1397
Simon, J. D. 2019, ARA&A, 57, 375
Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163
Steer, I. 2020, AJ, 160, 199
Steer, I., Madore, B. F., Mazzarella, J. M., et al. 2017, AJ, 153, 37
Tartaglia, L., Sand, D. J., Valenti, S., et al. 2018, ApJ, 853, 62
Taylor, E. N., Hopkins, A. M., Baldry, I. K., et al. 2011, MNRAS, 418, 1587
Telford, O. G., Dalcanton, J. J., Williams, B. F., et al. 2020, ApJ, 891, 32
Willman, B., Dalcanton, J. J., Martinez-Delgado, D., et al. 2005, ApJL, 626, L85
Wright, A. H., Robotham, A. S. G., Driver, S. P., et al. 2017, MNRAS, 470, 283
Zibetti, S., Charlot, S., & Rix, H.-W. 2009, MNRAS, 400, 1181