7-3-2014

A UV to Mid-Ir Study of AGN Selection

Sun Mi Chung
The Ohio State University

Christopher S. Kochanek
The Ohio State University

Roberto Assef
Diego Portales University

Michael J. I. Brown
California Institute of Technology

Daniel Stern
California Institute of Technology

See next page for additional authors

Authors

Sun Mi Chung, Christopher S. Kochanek, Roberto Assef, Michael J. I. Brown, Daniel Stern, Buell T. Jannuzi, Anthony H. Gonzalez, Ryan C. Hickox, and John Moustakas

A UV to Mid-IR Study of AGN Selection

Sun Mi Chung[1], Christopher S. Kochanek[1], Roberto Assef[2], Michael J. Brown[3], Daniel Stern[4], Buell T. Jannuzi[5], Anthony H. Gonzalez[6], Ryan C. Hickox[7], John Moustakas[8]

ABSTRACT

We classify the spectral energy distributions (SEDs) of 431,038 sources in the 9 sq. deg Boötes field of the NOAO Deep Wide-Field Survey (NDWFS). There are up to 17 bands of data available per source, including ultraviolet ($GALEX$), optical (NDWFS), near-IR (NEWFIRM), and mid-infrared (IRAC/MIPS) data, as well as spectroscopic redshifts for $\sim$20,000 objects, primarily from the AGN and Galaxy Evolution Survey (AGES). We fit galaxy, AGN, stellar, and brown dwarf templates to the observed SEDs, which yield spectral classes for the Galactic sources and photometric redshifts and galaxy/AGN luminosities for the extragalactic sources. The photometric redshift precision of the galaxy and AGN samples are $\sigma/(1+z) = 0.040$ and $\sigma/(1+z) = 0.169$, respectively, with the worst 5% outliers excluded. Based on the χ^2_ν of the SED fit for each SED model, we are able to distinguish between Galactic and extragalactic sources for sources brighter than $I = 23.5$. We compare the SED fits for a galaxy-only model and a galaxy+AGN model. Using known X-ray and spectroscopic AGN samples, we confirm that SED fitting can be successfully used as a method to identify large populations of AGN, including spatially resolved AGN with significant contributions from the host galaxy and objects with the emission line ratios of "composite" spectra. We also use our results to compare to the X-ray, mid-IR, optical color and emission line ratio selection techniques. For an F-ratio threshold of $F > 10$ we find 16,266 AGN candidates brighter than $I = 23.5$ and a surface density of ~ 1900 AGN deg^{-2}.

Subject headings: galaxies: active, quasars: general, galaxies: distances and redshifts

1. Introduction

Active Galactic Nuclei (AGN) present a wide variety of observational properties, which can be exploited to identify them in large surveys using a range of selection techniques. However, most selection methods

[1]Department of Astronomy, The Ohio State University, 140 W. 18th Ave., Columbus, OH 43210, USA

[2]Núcleo de Astronomía de la Facultad de Ingeniería, Universidad Diego Portales, Av. Ejército Libertador 441, Santiago, Chile

[3]School of Physics, Monash University, Clayton, Victoria 3800, Australia

[4]Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Mail Stop 169-221, Pasadena, CA 91109, USA

[5]Department of Astronomy and Steward Observatory, University of Arizona, Tucson, AZ 85721, USA

[6]Department of Astronomy, University of Florida, Gainesville, FL 32611, USA

[7]Department of Physics and Astronomy, Dartmouth College, 6127 Wilder Laboratory, Hanover, NH 03755, USA

[8]Department of Physics and Astronomy, Siena College, 515 Loudon Road, Loudonville, NY 12211, USA

are based on the characteristics in a limited wavelength region. This leads to selection methods that are incomplete for some AGN luminosities, classes, or redshifts. For example, the early single (optical) color searches have evolved into multi-color optical surveys such as the Sloan Digital Sky Survey (SDSS), which has successfully identified $\sim 10^5$ quasars based largely on their optical colors (Richards et al. 2002, 2005; Schneider et al. 2007, 2010). The color selection methods of Richards et al. (2002) achieved a completeness of $\sim$90% for unobscured, unresolved AGN with $i < 19$, $z < 5$, and emission dominated by the AGN, with an efficiency (number of quasars/number of candidates) of $\sim$65%. However, using $ugriz$ colors also presents difficulties in distinguishing between F stars and quasars at $z \sim 2$. Since quasar activity is known to peak near this redshift (Richards et al. 2006; Assef et al. 2011), this presents a problem for studying the accretion history of the universe. Using broadband optical colors to find AGN also relies on the roughly universal UV/optical power-law shape of direct emission from an accretion disk to distinguish AGN from the primarily photospheric emission of stellar populations in this wavelength range. Obscured, type-2 AGN are more difficult to identify using optical colors, due both to the distortion of the shape of the AGN continuum and the increasingly significant host galaxy contribution. Similarly, low accretion rate AGN are dominated by their host emission (e.g. Hopkins et al. 2009).

X-ray emission is sensitive to both unobscured and moderately obscured AGN, with the added advantage that X-ray emission almost always implies the presence of AGN except at low luminosities ($L_X \lesssim 10^{40}$ erg s^{-1}) where binary contributions become important (Shapley et al. 2001). However, *Chandra* and *XMM-Newton* surveys are biased against highly obscured AGN whose soft X-rays are absorbed by large columns of gas and dust (e.g., Polletta et al. 2006). Hard X-ray observations largely avoid this problem, but until the recent launch of *NuSTAR* (Harrison et al. 2013; Alexander et al. 2013), with its greater sensitivity to hard X-rays, there were only shallow surveys for the highly obscured, Compton-thick AGN, such as the *Swift* BAT Survey (Tueller et al. 2008). Radio surveys can also be used to search for AGN; however, typical radio surveys are limited to a very small population of AGN, since only $\sim$10%-20% are radio-loud (e.g. Kellermann et al. 1989; Urry & Padovani 1995; Stern et al. 2000; Ivezić et al. 2002).

Mid-infrared (mid-IR) surveys are sensitive to both type-1 and type-2 AGN. The accretion disk power law spectrum extends into the rest-frame near-IR and is less affected by extinction. Additionally, at rest-frame mid-IR and longer wavelengths, there is emission by hot dust in the "torus" where the UV and optical emission from the accretion disk is absorbed and reprocessed (e.g., Sanders et al. 1989; Netzer et al. 2007). These contributions roughly produce a power-law spectrum in the mid-IR that is easily distinguished from the infrared spectra of stars or relatively normal galaxies which are essentially falling Rayleigh-Jeans spectra regardless of stellar population or age at these wavelengths. At lower redshifts ($z \lesssim 1.5$) mid-IR selection triggers on dust emission, while at higher redshifts ($z \gtrsim 1.5$) it really triggers on the direct emission from the disk. As a result, many AGN occupy a region of mid-IR color space that is well separated from that of galaxies and stars (e.g., Lacy et al. 2004; Stern et al. 2005; Alonso-Herrero et al. 2006; Donley et al. 2007), which was initially demonstrated using data from the four bands (3.6μm to 8.0μm) of *Spitzer*/IRAC. The *Wide-Field Infrared Survey Explorer* (*WISE*; Wright et al. 2010) can be used to extend these methods over the full sky (Mateos et al. 2012; Stern et al. 2012; Assef et al. 2013; Yan et al. 2013). However, mid-IR color selection methods fail when the host galaxy contribution becomes large, or a strong Hα line lies in the 3.6μm band at $z \sim 4$ and distorts the colors (Assef et al. 2010). For quasars that are as faint as W2$\sim$15.5 in the WISE 4.6μm band, galaxy contamination becomes significant and mid-IR color selection becomes more difficult (Assef et al. 2013).

Each technique for identifying AGN has its own set of selection effects and it is important to understand the differences between AGN samples selected with the various methods. For example, AGN that are detected

using X-ray, radio, or mid-IR colors have different characteristic Eddington ratios, clustering properties, and host galaxy morphologies and host galaxy masses. Hickox et al. (2009) studied a sample of $\sim$600 AGN at $0.25 < z < 0.8$ with spectroscopic redshifts from the AGES and Galaxy Evolution Survey (AGES; Kochanek et al. 2012), and radio, X-ray, and mid-IR data from the Westerbork Synthesis Radio Telescope, the *Chandra* XBoötes Survey, and the *Spitzer* IRAC Shallow Survey, respectively. With modest overlap between the radio, X-ray, and mid-IR AGN samples, Hickox et al. (2009) found that radio AGN tend to be optically red and have massive host galaxies (i.e. red sequence galaxies), while X-ray AGN span a large range of host masses and colors but show a peak in the "green valley", and mid-IR AGN tend to be slightly optically bluer than the X-ray AGN. Griffith & Stern (2010) arrived at a similar conclusion by examining the host galaxy morphologies of radio, mid-IR, and X-ray selected AGN, and found that radio-selected AGN tend to have early-type hosts, while mid-IR and X-ray selected AGN are more likely to have disky host galaxies. This may represent a sequence in evolution of the quasar phase, which is thought to be triggered by a burst of star formation that funnels gas into the central engine, which ultimately quenches the star formation, transforming the optical colors from blue to red (Hopkins et al. 2007; Cardamone et al. 2010; Pović et al. 2012).

Traditionally, extragalactic surveys have either been wide and shallow, or narrow and deep. Wide/shallow surveys such as the 2 Micron All-Sky Survey (2MASS; Skrutskie et al. 2006), the Sloan Digital Sky Survey (SDSS; Abazajian et al. 2009), and the 2 Degree Field Galaxy Redshift Survey (2dF; Colless et al. 2001) have allowed studies of large numbers of nearby galaxies and contributed significant discoveries about galaxy properties in the low redshift universe and luminous AGN at all redshifts (e.g. Richards et al. 2005; Huchra et al. 2012). Examples of deep/narrow surveys include the Great Observatories Origins Deep Survey (GOODS; Dickinson et al. 2003) and the Cosmological Evolution Survey (COSMOS; Scoville et al. 2007). GOODS covers $\sim$300 arcmin^2 with deep observations from *Spitzer*, *Hubble*, *Chandra*, *Herschel* and *XMM-Newton*, as well as ground-based facilities. COSMOS is a wider survey, covering $\sim$2 deg^2 with imaging from both space-based and ground-based telescopes, including *Hubble*, *Spitzer*, *GALEX*, *Chandra*, *NuSTAR*, Subaru, VLT, and others. The relative rarity of AGN makes it difficult however to fully characterize AGN in deep fields or to fully account for faint AGN in shallow fields.

While these (and other) extragalactic surveys have provided a tremendous contribution to our knowledge of both nearby and distant galaxy populations, there have been few surveys on $\sim$10 deg^2 scales. The NOAO Deep Wide-Field Survey (NDWFS; Jannuzi & Dey 1999) in the 9 deg^2 Boötes field allows us to explore a cosmologically significant volume with data that is both deep and extensive in wavelength coverage. The large survey volume means that studies in the Boötes field are less affected by cosmic variance or small number statistics for rare, faint objects that would not be found in either deep/narrow or wide/shallow surveys. In addition to near-IR and optical data from NDWFS, there is a wealth of ancillary data in the Boötes field. This includes additional optical ground-based data (Cool 2007; Bian et al. 2013), X-ray data from the *Chandra X-ray Observatory* (Murray et al. 2005; Kenter et al. 2005; Brand et al. 2006), UV data from the *Galaxy Evolution Explorer* (*GALEX*; Hoopes 2004), and mid-IR data from the *Spitzer Space Telescope* (Eisenhardt et al. 2004; Ashby et al. 2009; Jannuzi et al. 2010). There is also spectroscopic data from the AGN and Galaxy Evolution Survey (AGES; Kochanek et al. 2012), which provides redshifts for $\sim$18,000 galaxies ($I < 20$) and $\sim$4,700 AGN candidates ($I < 22.5$).

In this paper, we examine and classify the UV to mid-IR spectral energy distributions (SEDs) of 431,038 sources in the Boötes survey region. By fitting galaxy, AGN, and stellar templates to the observed SEDs, we are able to isolate AGN candidates from galaxy and stellar populations using a much broader wavelength range than the common AGN selection methods discussed previously. In §2 we outline the data that are

used in this work. In §3.1 and §3.2 we compare photometric to spectroscopic redshifts and discuss in detail the SED fitting procedure and how our final sample of SED fits were assembled. In §3.3 we illustrate the separation of Galactic and extragalactic sources based on χ^2_ν statistics and we investigate the difficulty of classifying sources at faint magnitudes. Section §4 presents how AGN candidates may be selected based on the results of their SED fits as galaxy and galaxy+AGN models and the sensitivity of this SED selection to the relative strength of the underlying host galaxy. In §4.1, we compare the SED selection of AGN candidates to mid-IR and optical color selections. In §4.2 we compare the surface densities of different AGN samples. Finally in §5 we summarize our most important results and comment on future work. Throughout this work we assume a flat ΛCDM cosmology with $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_M = 0.28$, and $\Omega_\Lambda = 0.72$.

2. Data

The data used in this work consist of the extensive multi-wavelength imaging of the 9 deg^2 NDWFS Boötes field and optical spectroscopy from AGES (Kochanek et al. 2012) and Hickox et al. (private communication). The following sections provide brief descriptions of the various datasets used in this paper. Data from NDWFS (B_W, R, I, K) and NEWFIRM (J, H, K_s) are based on the Vega magnitude system, and data from *GALEX*, U_S, Y, and z bands are based on the AB system. All magnitudes reported throughout this paper are kept in their native zeropoint systems and refer to 6″ diameter aperture magnitudes unless otherwise stated. The 3σ magnitude limits quoted for each of the filters is based on calculating the signal-to-noise (S/N) ratio from the fractional flux error. The median S/N is computed in bins of 0.25 mag, and a simple linear interpolation is used to obtain the magnitude where S/N=3σ.

Objects were detected using SExtractor 2.3.3 (Bertin & Arnouts 1996) run on the I-band images from the NDWFS third data release. To obtain consistent aperture photometry across the full Boötes field, the B_W, R, I, Y, H and K_S images were smoothed to a common PSF of 1″35 FWHM while the U_S, z and J images were smoothed to a common PSF of 1″60 FWHM. We measured aperture photometry for each object using our own code. SExtractor segmentation maps were used to exclude flux associated with neighboring objects, and we corrected the photometry for missing pixels using the mean flux per pixel measured in a series of annuli surrounding each object. Uncertainties were determined by measuring photometry at $\simeq$100 positions within 2′ of the object position and finding the range that encompassed 68% of the measurements. Saturation mainly occurs for sources brighter than $R \simeq 17$ in typical NDWFS exposures (Jannuzi & Dey 1999), which corresponds roughly to $I \simeq 16.5$. For a more extensive description of the photometric catalogues we refer the reader to Brown et al. (2007).

2.1. X-ray Observations

XBoötes is an X-ray survey of the 9 deg^2 NDWFS Boötes field, with data taken from ACIS-I onboard the *Chandra X-ray Observatory* (Murray et al. 2005; Kenter et al. 2005; Brand et al. 2006). The integration time was 5 ks per position and the survey identified 3293 point sources with $\geq$4 counts. This corresponds to a flux of 7.8×10^{-15} ergs cm^{-2} s^{-1} in the 0.5-7 keV range assuming a standard unabsorbed AGN X-ray spectrum. However, most 4-count sources actually correspond to more numerous sources modestly fainter than this flux limit brought into the catalog by Poisson fluctuations in their counts (see Kenter et al. 2005).

2.2. UV Imaging

GALEX observed the field in its far-UV (FUV) and near-UV (NUV) bands with 3σ magnitude limits of 24.7 and 25.5 mag (AB), respectively.

2.3. Optical Imaging

The optical data consists of the NDWFS B_W, R, and I bands (Jannuzi & Dey 1999), the z band (zBoötes; Cool 2007), and the U_S band from the Large Binocular Telescope (LBT; Bian et al. 2013). The 3σ depths of the U_S, B_W, R, I, and z bands are 24.9 (AB), 25.2 (Vega), 23.9 (Vega), 22.9 (Vega), and 22.3 (AB) mag, respectively.

2.4. Near-Infrared

Near-IR data from the Infrared Boötes Imaging Survey (IBIS; Gonzalez et al. 2010) provide photometry in the J, H, and K_S bands with 3σ depths of 21.1, 20.1, 18.9 (Vega). There is also K-band data from the original NDWFS survey, though this is significantly shallower than the NEWFIRM data, with a 3σ limit of 17.4 mag (Vega). Finally, the Y-band data from the LBT (Bian et al. 2013) has a 3σ limit of 22.2 mag (AB).

2.5. Mid-Infrared

Mid-IR observations from the *Spitzer* Deep, Wide-Field Survey (SDWFS; Ashby et al. 2009) and the MIPS AGN and Galaxy Evolution Survey (MAGES; Jannuzi et al. 2010) provide data in the four IRAC channels and the 24μm MIPS channel. SDWFS includes data from the earlier IRAC Shallow Survey (Eisenhardt et al. 2010). The 3σ depths of the four IRAC bands ([3.6], [4.5], [5.8], [8.0]) and the MIPS 24μm band are 19.3, 18.7, 16.8, 16.1, and 11.8 mag, respectively.

2.6. Spectroscopic Data

In addition to the photometric catalogs, we use spectroscopic data from AGES (Kochanek et al. 2012), which was carried out with MMT/Hectospec (Roll et al. 1998; Fabricant et al. 2005). The AGES galaxy samples were designed to sample the broad range of galaxy colors with well-defined samples at all wavelengths from B_W to 24μm. It is complete for $I < 18.5$ and sparsely sampled for $18.5 < I < 20$ galaxies, with redshifts obtained for $\gtrsim 30\%$ of all $I < 20$ galaxies. The AGES limiting magnitude for AGN was significantly fainter, reaching $I < 22.5$ mag for point-like sources in the I-band. The AGN targets were selected based on several different criteria, including point-like sources with MIPS emission, mid-IR colors, and X-ray or radio emission. A total of $\sim$18,000 redshifts were collected for $I < 20$ galaxies and $\sim$4,700 redshifts for $I < 22.5$ AGN candidates. We extend this spectroscopic sample with an additional $\sim$1000 Hectospec redshifts from Hickox et al. (private communication). The magnitude of a field $L_\star$ galaxy at $z \sim 1$ is $i' \sim 21.2$ (Gabasch et al. 2006), roughly corresponding to $I \sim 20.8$ (Vega). This is fainter than the limiting magnitude of the AGES and Hickox galaxy samples and we can consider essentially all $z \gtrsim 1$ spectroscopic sources to be AGN.

We start with ∼two million sources that have either $I < 23.5$ or $[4.5] < 18$. These two limits are broadly comparable for typical sources, but by using both we avoid biases against optically faint sources such as high redshift galaxies/AGN or brown dwarfs. The optical limiting magnitude corresponds to a galaxy luminosity of $0.1L^*$ at $z = 1$ (e.g., Gabasch et al. 2006). We next require that each source has data in at least ten different filters, with a minimum of five detections at S/N>3. Since we fit SED models to the fluxes, negative flux estimates are simply included as measurements, which is more statistically correct than including them as one-sided upper bounds. Photometric uncertainties that are too small compared to the systematic residuals typical of the SED fits will bias the fits. Therefore we assign a minimum photometric error of 0.05 mag. Because the NDWFS I-band subfields overlap, there can be up to 4 detections for individual objects. To select the best detection, we reject objects that are outside nominal subfield boundaries and then (if needed) select the detection with the highest S/N. There are regions within the fields that have poor data (e.g., bad pixels, saturated pixels) or increased backgrounds (e.g., bright objects halos). We excluded objects in these regions based on "flagdeep", which identifies objects with good faint object photometry. These criteria left us with a sample of 431,038 sources which we fit with galaxy, AGN, and stellar templates. Among the ∼1.5 million sources that we exclude from our sample, roughly 25% are lost due to "regional" issues associated with "flagdeep", while the remaining sources are lost due to not having enough bands that meet the S/N criterion.

3.   Analysis

This section discusses the SED fitting procedure and compares the resulting photometric redshifts with spectroscopic redshifts for ∼20,000 sources. We present the χ^2 statistic per degree of freedom (χ^2_ν) of the galaxy+AGN and stellar/brown dwarf SED models and demonstrate that examining χ^2_ν is an effective way to separate Galactic and extragalactic sources. The results of the final fits are reported in Table 2.

3.1.   SED fitting

We use the empirically derived SED templates of Assef et al. (2010) to fit the sources. The templates extend from 0.03 to 30μm. There are three galaxy templates, corresponding to "elliptical" (old stellar population), "spiral" (on-going star formation), and "irregular" galaxies (starburst population), and a single AGN template. The AGN template is fit with a variable amount of internal reddening and both galaxy and AGN templates are fit with an IGM absorption model that is a fixed function of redshift (Stengler-Larrea et al. 1995; Fan et al. 2006). We do not fit the IGM absorption as a free parameter because this worsens the photometric redshifts (Assef et al. 2010). In addition to the galaxy and AGN templates, we fit all sources with the stellar atmosphere models of Castelli & Kurucz (2004) and as brown dwarfs with effective temperatures of 2500K, 2000K, 1500K, and 1000K from the Allard et al. (2007) models. We linearly interpolate between stellar templates of adjacent temperature/spectral class to create a one parameter sequence in stellar temperature for both dwarf and giant spectral classes.

The templates are fit to the data using the publicly available code of Assef et al. (2010), which fits non-negative linear combinations of the templates to the data. An R-band galaxy luminosity function based on the Las Campanas Redshift Survey (Lin et al. 1996) is used as a luminosity prior on the galaxy templates in order to avoid unlikely luminosities. No such prior is applied to the luminosity of the AGN component in the SED. The data are fit once with only the three galaxy templates, then again with the galaxy templates

and the AGN component, and a final time with only the stellar and brown dwarf templates. The redshift parameter of the fits is allowed to vary between 0 and 3 in steps of 0.01 when fitting for galaxies and AGN, and is fixed at $z = 0$ for stars and brown dwarfs.

The goodness-of-fit is measured by χ^2_ν which is simply χ^2 divided by the degrees of freedom (DOF or ν). A good model fit should yield $\chi^2_\nu \simeq 1 \pm \sqrt{2/\nu}$. Values significantly smaller than unity indicate that the data are being over-fit, while values much larger than unity signify a poor model fit. For N bands of data, including upper limits, the galaxy, galaxy+AGN, and stellar models have $\nu = N - 4$, $N - 6$, and $N - 2$ degrees of freedom, respectively. The galaxy model parameters are the three template luminosities and the redshift. The galaxy+AGN model adds the luminosity and extinction of the AGN component. The stellar models depend on the total flux and the "temperature" parameter.

We fit each object once using only the three galaxy templates, then a second time adding the AGN template. We then determine whether the inclusion of the AGN template significantly improves the fit. The χ^2 value cannot be used to distinguish whether an SED is better fit with or without the AGN component because the two additional model parameters will always result in a "better" fit with a smaller χ^2. In order to test whether a significant improvement has occurred, we calculate the F-ratio

$$F = \frac{\chi^2_{\mathrm{Gal}} - \chi^2_{\mathrm{Gal+AGN}}}{\chi^2_{\mathrm{Gal+AGN}}} \cdot \frac{\nu_{\mathrm{Gal+AGN}}}{\nu_{\mathrm{Gal}} - \nu_{\mathrm{Gal+AGN}}}, \tag{1}$$

which compares the change in χ^2 to the change in degrees of freedom. For Gaussian uncertainties, the probability distribution of F is known and we can assign a probability to the improvement in the fit being significant. For example, if $\nu_{\mathrm{Gal}} = 11$ (implying that $\nu_{\mathrm{Gal+AGN}} = 9$) and $F = 10$, there is only a 0.5% probability that the χ^2 improvement is a chance occurrence, and we have high confidence that the χ^2 improvement is due to the presence of an AGN. In practice, we must be more careful because the residuals of the fits are almost certainly not simple Gaussians.

While the F-ratio and associated probability distribution (or the F-test) quantifies the probability that the SED of a particular object is better fit with an AGN component, we must also take into consideration the relative rarity of AGN compared to galaxies when constructing statistical samples. Suppose $P(F)$ is the probability that the source is better fit as a galaxy rather than as a quasar by chance. In a sample of N_G galaxies and N_Q quasars, we would expect to find $P(F)N_G$ galaxies that are falsely characterized as quasars and $(1 - P(F))N_Q$ quasars. The number of sources above and below a given F value can be estimated as $N(< F) = (1 - P(F))N_G + P(F)N_Q$ and $N(> F) = P(F)N_G + (1 - P(F))N_Q$, respectively. A crude estimate of the quasar fraction is then

$$\frac{N_Q}{N(> F)} \simeq 1 - \frac{P}{(1 - 2P)}\left(\frac{N(< F)}{N(> F)} - 1\right). \tag{2}$$

For example, based only on the probability of a false positive associated with the F-ratio, one might naively expect that for $F(> 3)$ (or $P = 0.055$), there is an implied purity of $N_Q/(N(> F)) \simeq 95\%$. However, the ratio of sources above and below $F = 3$ is $N(> F)/N(< F) \simeq 19\%$, and Equation (2) would suggest an expected purity of $N_Q/N(> F) \simeq 75\%$. Even for $F > 8$ (naively 99.9% pure), Equation (2) suggests a purity of 98%. This somewhat oversimplifies the problem, but it emphasizes the consequence of searching for relatively rare sources in large populations.

There are $\sim$20% of sources in our sample that have $F < 0$ because the data are slightly worse fit by the Galaxy+AGN model than by the galaxy-only model. This occurs when the luminosity of the galaxy-only model is too high and there is a large penalty from the luminosity prior. To solve this problem, the code

includes an AGN component in order to decrease the galaxy luminosity, even at the expense of having a slightly worse χ^2 for the fit to the SED because we optimize on the χ^2+prior, rather than just the χ^2. Of the $F < 0$ SED fits, more than 99% have $\chi^2_{\mathrm{Gal}} - \chi^2_{\mathrm{Gal+AGN}} > -5$, which means the addition of an AGN component has not significantly improved the overall fit. We interpret these sources as pure galaxies or galaxy-dominated SEDs with little to no AGN contribution.

3.2. Anomalous Data Points

One of the challenges of working with a large dataset is to confirm the fidelity of the SED fits for all sources. With photometry for ∼430,000 sources from multiple telescopes, filters, and epochs, it is inevitable that some fraction of objects have badly fit SEDs due to inconsistent or bad photometry in one or more bands. Problematic photometry can arise from either instrumental (e.g. stray light, cosmic rays), or astrophysical causes (e.g. variability, supernova). Source confusion can also be an issue, especially for the *GALEX* bands where the PSF is ∼ 6″. One way to minimize the impact of outliers is to sequentially exclude data from a single filter and re-fit the SED to the remaining data. If the results of the original fit are stable, then dropping any single band should not dramatically alter the SED fit, redshift, or the χ^2 goodness-of-fit. If, however, a single band is inconsistent with the data from all the other filters and is driving the results, then we can identify the problematic filter and use the results excluding the problematic filter. Figure 1 illustrates this with a galaxy that is badly fit by all templates due to an anomalous Y-band measurement. The SED fits in the left panel use all the bands, including the Y-filter, with the results for the galaxy templates, galaxy+AGN, and stellar templates shown from top to bottom. In this case, dropping the Y-band data point greatly improves the fit (from χ^2_ν=19.4 to χ^2_ν=0.9) and the resulting SED fits are shown in the right panel of Figure 1. Data points with small error bars tend to be dropped more frequently because they cause large changes in χ^2. While in the example shown in Figure 1, the SED model changes significantly due to the dropped filter, this is not always the case. Sometimes a filter with a small error bar is dropped, thereby significantly improving the χ^2 of the fit, yet the final SED model remains approximately the same.

Figure 2 shows the distribution of the dropped filters with band pass that produce a greater than 3σ improvement in the χ^2 of the SED fits, where $\sigma = \sqrt{2\nu}$. We show the results for the galaxy, AGN, and stellar samples. The histogram for each sample is normalized to the total number of sources that are dropped. The stellar sample includes sources that are better fit by stellar SED models rather than galaxy or AGN models. The remaining extragalactic sources are separated into "galaxy" and "AGN" samples, where the AGN sample is comprised of sources with SED fits that show an improvement of $\Delta\chi^2> 20$ after an AGN component is introduced to the fit. While this is neither a pure nor a complete AGN sample, it is sufficient for the purpose of comparing the distribution of dropped filters among galaxy and AGN-dominated systems.

The distribution of dropped filters for the galaxy and AGN sample are nearly identical. This indicates that the distribution of the dropped filters is likely a good representation of the distribution of problems in the data rather than being due to systematic problems in the SED models. In both samples, the most frequently dropped filter is the Y-band, which is most likely due to a zeropoint calibration offset, which we have adjusted for on average by examining the SED residuals among optically bright, well-fit sources and applying the average offset between the SED model and the Y-band data. Even with this zeropoint correction, there are still enough objects for which the Y-band magnitude is offset from the rest of the SED, and has a small photometric error, that dropping it will cause a significant improvement in the overall SED fit. The next most commonly dropped filters are the NUV and U_S band filters due to the low levels of UV emission from most galaxies. In the case of AGN, the FUV emission is weak because it is blueward of
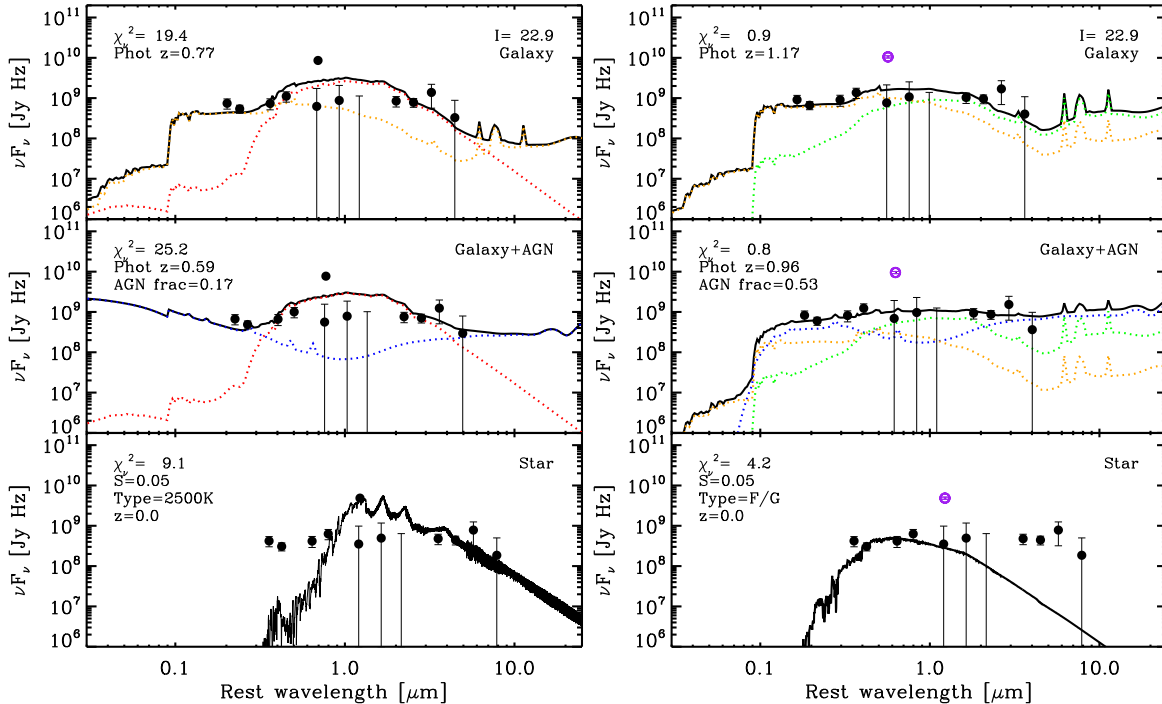
Fig. 1.— Left: SED of a galaxy with an anomalous data point in the Y-band, with coordinates α=216.40562 deg and δ=32.32618 deg. Due to the small photometric error of the Y-band measurement relative to the other bands, the SED fit does not characterize the true shape of the galaxy SED, leading to the large χ^2_ν in the left panels. The top, middle, and bottom panels are fit with galaxy templates, galaxy+AGN templates, and a stellar or brown dwarf template, respectively. The χ^2_ν photometric redshift, AGN fraction (ratio of AGN luminosity to total luminosity), stellarity index (S), and star or brown dwarf template type, are included. The dotted blue, red, green, and orange lines show the AGN, elliptical, spiral, and irregular galaxy SED models, respectively. Right: The Y-band measurement (purple open circle) is excluded from the fit, which results in a significantly improved fit, as illustrated by the large improvements in χ^2_ν.

the Lyman limit for $z \gtrsim 0.2$. The [8.0] and 24μm bands are also frequently dropped for both galaxy and AGN samples. The distribution of dropped filters for the stellar population is somewhat different from the distribution for galaxies and AGN. The most commonly dropped filters are the U_S, B_W, [3.6], and [4.5] bands.

We limit the total number of dropped filters such that only 1% of the photometry is excluded from the SED fits. This leads to 18% of the sources having a single excluded filter. For each filter, we only allow up to 1% of the data for that filter to be dropped. If more than 1% of the sources show an improvement greater than 3σ, we drop the 1% which lead to the greatest χ^2 improvement. For a typical source with 15 bands of photometry being fit with the galaxy+AGN models, we require a $\Delta\chi^2$ improvement of at least 12.7 due to the exclusion of a single filter. Different filters can be dropped for the same source depending on whether it is being modelled as a galaxy, AGN, or star/brown dwarf. However, the total number of bands used in the SED fit must be the same for any given object, regardless of which templates are being used. The final results are all based on SED fits that exclude this worst 1% of the photometry. The choice to exclude 1% of the photometry is somewhat arbitrary. However, changing the exact choices of these parameters (e.g. increasing the threshold of improvement to 5σ or excluding a larger or smaller fraction of the data) does not significantly change any of the subsequent results.

We obtain both photometric redshifts and "bolometric" luminosities of the individual template com-
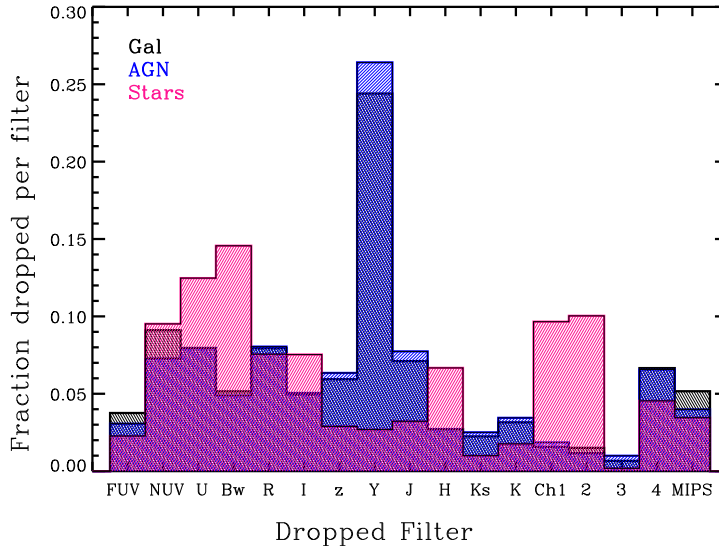
Fig. 2.— Distribution of dropped filters for the galaxy, AGN, and stellar populations (black, blue, and dark pink histograms, respectively). Note that this shows the distribution of dropped filters among the 18% of sources in our sample that have a single filter excluded from their final SED fits.

ponents, where "bolometric" means the integrated luminosity from 0.03 to 30μm for galaxy templates and 0.1216 to 30μm for the AGN template. The AGN template luminosity is not integrated for wavelengths shorter than Lyα (0.1216 μm) because the template is not well constrained at these wavelengths due to absorption by the IGM (see Assef et al. 2010). We note that while the galaxy luminosities are integrated from 0.03 to 30μm, this is effectively the same as integrating from 0.1216 to 30μm because there is very little far-UV flux in the galaxy templates.

In Figure 3 we compare the photometric and spectroscopic redshifts for a sample of 20,726 galaxies and AGN, primarily from the AGES survey with an additional $\sim$1000 redshifts from Hickox et al. (private communication). Most of the sources in Figure 3 have bolometric luminosities that are dominated by galaxy templates, with 85% of the galaxies having $<$ 20% of an AGN component. This is consistent with the AGES sample, which is comprised of $\sim$80% galaxies and $\sim$20% AGN candidates. The photometric redshift dispersion is calculated as

$$\frac{\sigma}{(1+z)} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{z_{\mathrm{phot}} - z_{\mathrm{spec}}}{1 + z_{\mathrm{spec}}}\right)^2}. \tag{3}$$

The photometric redshifts of the entire sample shown in Figure 3 have a dispersion of $\sigma/(1+z) = 0.125$ and a median offset $<$ 0.001. As is typical of photometric redshifts, the dispersion is dominated by the tails of the distribution. If we drop the worst 5% of the outliers (largest offsets in $|z_p - z_s|$), then the dispersion for the remaining 95% of the sources goes down to $\sigma/(1+z)$=0.061. Both the dispersion and accuracy of the photometric redshifts are significantly worse for the AGN because AGN-dominated SEDs have few distinct features compared to galaxies (Rowan-Robinson et al. 2008; Salvato et al. 2009; Assef et al. 2010). In many cases, as the AGN component starts to dominate the bolometric luminosity, the SED becomes increasingly flat, leading to a large increase in the photometric redshift uncertainties.

If we separately consider sources that are likely to be AGN-dominated versus those that are galaxy-dominated based on their F-ratios, we find that the redshift dispersions are much higher for the AGN-
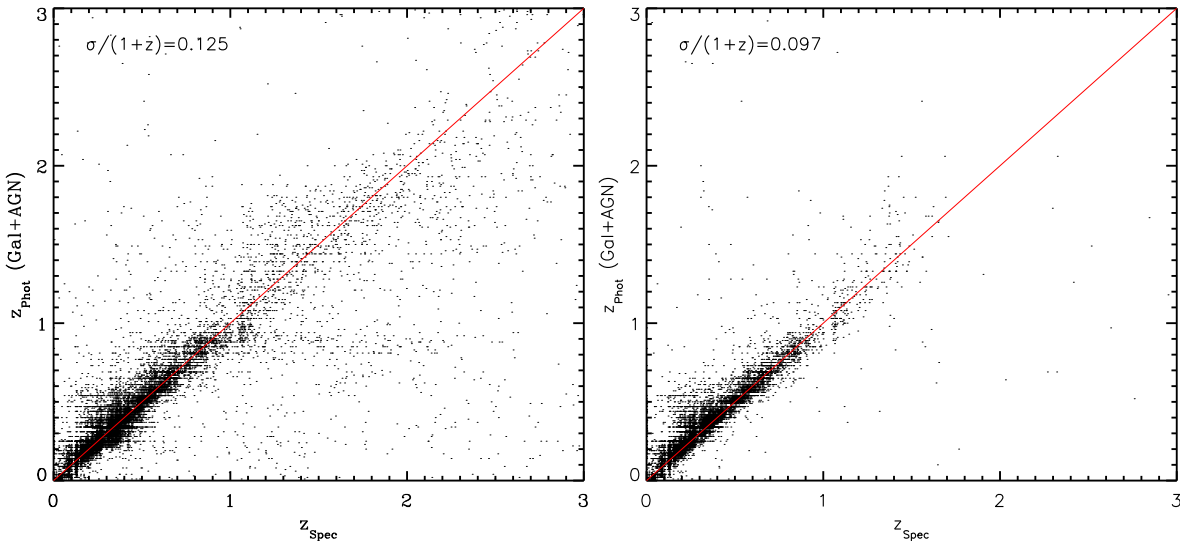
Fig. 3.— Left: Photometric versus spectroscopic redshifts for the entire spectroscopic sample in the range $0 < z < 3$, with the red line denoting a one-to-one relation. The photometric redshift is derived from the galaxy+AGN SED models. Right: Photometric versus spectroscopic redshifts for sources that are well fit by galaxy-only templates with $F < 1$.

dominated objects, as expected. We assign the $F < 1$ sources to the galaxy sample, and the $F > 10$ sources to the AGN sample, thus avoiding the more ambiguous, composite AGN/galaxy sources. After rejecting the worst 5% of each sample, the redshift dispersion of the galaxy and AGN samples are $\sigma/(1 + z) = 0.040$ and $\sigma/(1 + z) = 0.169$, respectively. These dispersions are also consistent with the photometric dispersions of sources targetted as galaxies or AGN candidates in the AGES survey. The $z_s < 1$ (primarily galaxies) and $z_s > 1$ (AGN) have photometric dispersions of $\sigma/(1 + z) = 0.044$ and 0.204, after rejecting the worst 5% outliers. One advantage of our study is that AGES supplies spectroscopic redshifts for most of the luminous AGN where photometric redshifts are most problematic. As we examine AGN that have a stronger host galaxy contribution, the photometric redshifts become increasingly robust. Our photometric redshift dispersions are comparable to the outlier-excluded dispersions from Assef et al. (2010) for galaxy and point-source AGN of $\sigma/(1 + z) = 0.041$ and $\sigma/(1 + z) = 0.184$, respectively. Brodwin et al. (2006) also calculated photometric redshifts for $\sim$200,000 sources in the Boötes field using a hybrid technique combining a template fitting algorithm with artificial neural nets (Collister & Lahav 2004), using optical, near-infrared, and IRAC data from NDWFS (Jannuzi & Dey 1999), the FLAMEX survey (Elston et al. 2006), and the IRAC shallow survey (Eisenhardt et al. 2004) but not the $GALEX$, U_S, z, Y, SDWFS, NEWFIRM, or MIPS data included here. Using the template fitting algorithm, Brodwin et al. (2006) obtained photometric redshift dispersions slightly higher than ours, with $\sigma/(1 + z) = 0.061$ and 0.341 for the galaxy and AGN samples, respectively. When Brodwin et al. (2006) use the hybrid technique, their photometric redshift dispersions decrease to $\sigma/(1 + z) = 0.047$ and 0.120 for the galaxy and AGN samples, respectively.

3.3. Galactic versus Extragalactic Sources

An important step in sorting through the different types of sources in the Boötes field is to separate the Galactic from the extragalactic sources. To do this, we examine the χ^2_ν of the stellar templates as compared to AGN+galaxy templates, as shown in Figure 4. The sample is split into extended and point-like

sources, and further separated into bins of I-band magnitude. Extended and point-like source definitions are determined from the SExtractor stellarity index S, which has values from 1 (point source) to 0, measured on the NDWFS I-band images. We consider sources with $S < 0.7$ as extended and $S \geq 0.7$ as point-like because there is a minimum in the distribution of stellarity indices at $S \sim 0.7$. Based purely on χ_ν^2, sources above the blue diagonal line have SEDs which are better fit by the galaxy+AGN templates, while sources below the line are better fit as stars or brown dwarfs. We will refer to those regions in χ_ν^2 space as the "extragalactic" and the "stellar" zones, respectively. Because the distinction between stars and extragalactic sources is not a continuum, we do not use F statistics in this separation.

As expected, most of the extended sources shown in Figure 4 (left) lie above the blue line, indicating that they are better fit by the galaxy and AGN templates than the stellar or brown dwarf templates. In the optically brightest bin ($17 < I < 20.5$), nearly all the sources lie above the blue line and are better fit as extragalactic sources. As the sources become optically fainter, the cloud moves to smaller values of both $\chi_\nu^2(\mathrm{Gal+AGN})$ and $\chi_\nu^2(\mathrm{Star})$ as the photometric uncertainties increase, hitting minimum χ_ν^2 values in the $22.5 < I < 23.5$ bin. In the optically faintest $I > 23.5$ bin, χ_ν^2 increases slightly, reflecting the increased uncertainty in the SED fits due to fewer available high S/N data points. In the brightest bins, systematic errors (e.g., offsets between filter zeropoints, photometric anomalies, problems in the templates) dominate the true total uncertainty in the SED fits rather than the photometric errors, leading to the high values of χ_ν^2. In the $17.0 < I < 20.5$ bin, the extended and point-like sources have median χ_ν^2 values of $\langle \chi_\nu^2(\mathrm{Gal+AGN}) \rangle \simeq 3.5$ and $\langle \chi_\nu^2(\mathrm{Star}) \rangle \simeq 3.5$. In comparison, the χ_ν^2 of extended and point-like sources in the fainter $21.5 < I < 22.5$ bin is lower, with $\langle \chi_\nu^2(\mathrm{Gal+AGN}) \rangle \simeq 1.0$ and $\chi_\nu^2(\mathrm{Star}) \simeq 1.8$. Note that the $\langle \chi_\nu^2(\mathrm{Star}) \rangle \simeq 1.8$ of $21.5 < I < 22.5$ point-like sources is misleadingly high due to the presence of contaminating galaxies and AGN. If we consider only the point-like sources that are best fit by stellar templates in the $21.5 < I < 22.5$ bin, $\chi_\nu^2(\mathrm{Star})$ decreases to $\simeq 0.8$. The overall shifts of χ_ν^2 suggest that the photometric error bars at the brightest magnitudes should be broadened by closer to 0.10 mag rather than 0.05 mag in order to encompass systematic uncertainties.

The right side of Figure 4 shows the same sequences for point-like sources. There are three classes of point-like sources in Figure 4: Galactic stars and brown dwarfs, unresolved high-luminosity AGN, and galaxies misclassified as point-like sources due to their apparent compactness. Among the optically bright, point-like sources, there is an elongated horizontal cloud of sources at $\chi_\nu^2(\mathrm{Star}) \sim 100$, which is mainly populated by bright AGN. The remainder of the sources in the Galaxy/AGN region are largely comprised of lower luminosity AGN with a more significant host component. Hosts generally have more "stellar" SEDs leading to lower values of $\chi_\nu^2(\mathrm{Star})$ than for the high-luminosity AGN. However, the dominant population of optically bright point-like sources are clearly best fit by stellar SEDs. While the $\chi_\nu^2(\mathrm{Star})$ are sometimes high for these sources, it is not driven by saturation in the $17.0 < I < 20.5$ magnitude range, but rather due to the simple suite of stellar template models (e.g. only solar metallicity) and to the sometimes overly small photometric uncertainties.

Figure 5 shows the SEDs of two optically bright point-like sources. The source on the left is clearly best fit by a stellar template. On the right is a point-like, luminous X-ray source, best fit by an AGN. The SED of a cool star has a shape that is somewhat similar to the SED shape of an elliptical galaxy longward of the UV wavelengths. Nonetheless, in most cases, the fits are able to easily distinguish between stars and elliptical galaxies. The χ_ν^2 of the star shown in Figure 5 (left) is significantly better fit with a stellar model ($\chi_\nu^2 \sim 1$) than a galaxy or galaxy+AGN model ($\chi_\nu^2 \sim 20$). The SEDs of luminous AGN are very non-stellar (more so than the SEDs of non-AGN galaxies), which means they are well separated from the stellar regions of Figure 4. The SED shown in Figure 5 (right) is an example from the elongated horizontal cloud of sources at
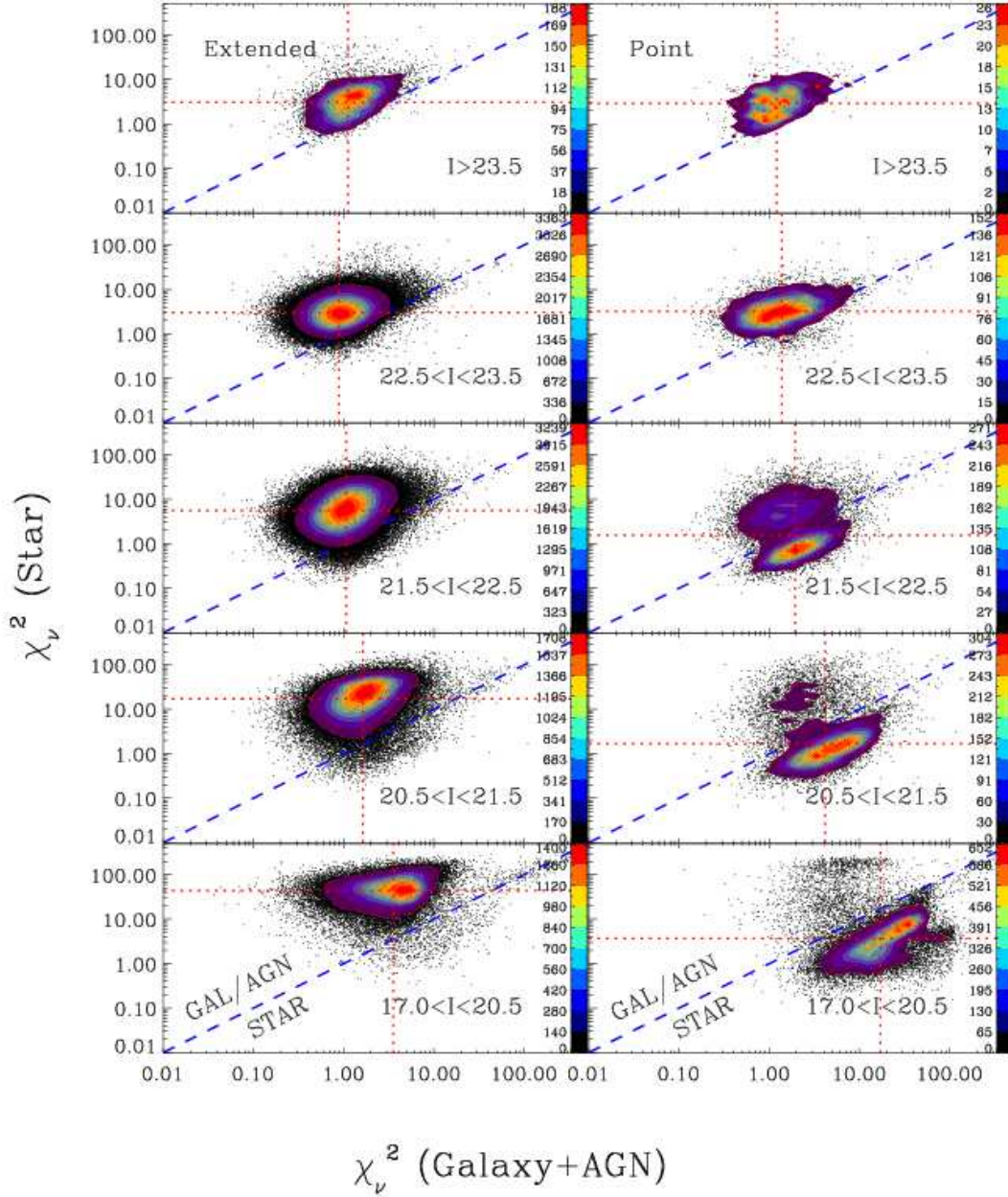
– 13 –

Fig. 4.— χ_ν^2 distributions for the stellar/brown dwarf templates and galaxy+AGN SED models, for extended (left; $S < 0.7$) and point-like (right; $S \geq 0.7$) sources, in bins of I-band magnitude. Ten density contours are overlaid with the colorscale shown on the right of each panel. The one-to-one, dashed blue line represents the separation between the stellar/brown dwarf population and the galaxy/AGN population. The red dotted lines show the median χ_ν^2 values for a source with the typical number of degrees of freedom.
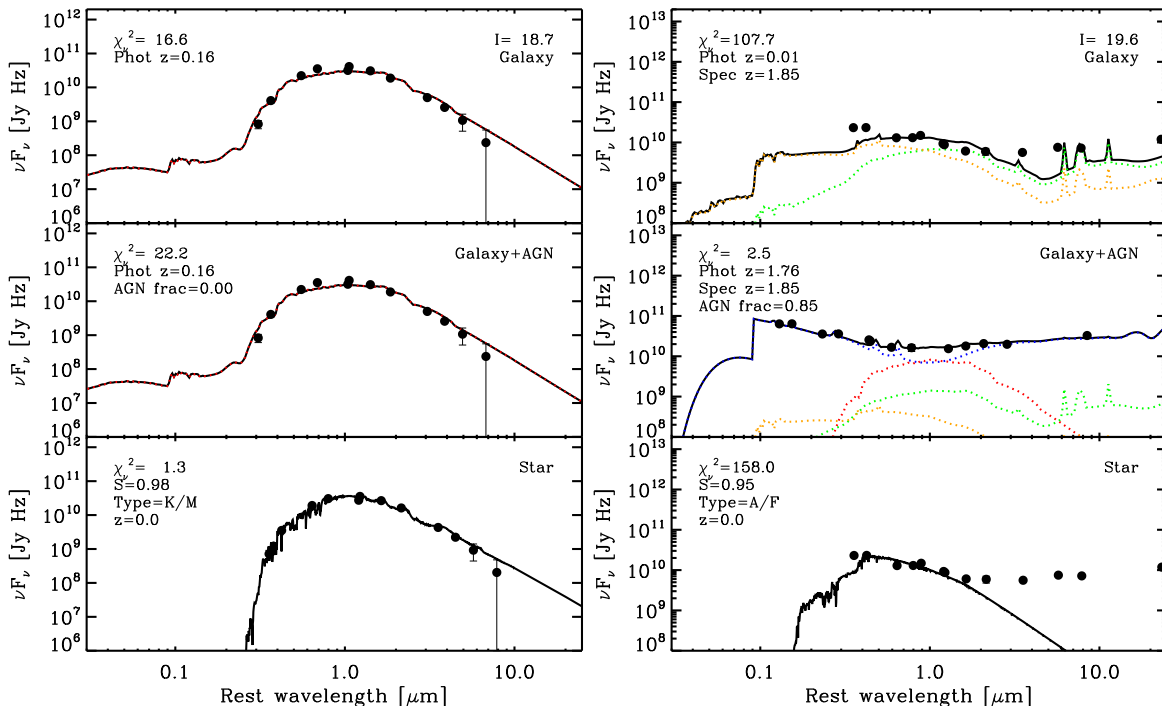
Fig. 5.— Left: SED of a K/M star fit with galaxy (top), galaxy+AGN (middle) and stellar (bottom) templates. Note that while the χ^2_ν has increased for the galaxy+AGN model, this is only due to the change in ν. Right: SED of a known optically bright, point-like, X-ray AGN.

$\chi^2_\nu(\mathrm{Star}) \sim 100$. The photometric redshift of this AGN is $z_p = 1.76$ and the measured spectroscopic redshift is $z_s = 1.85$, confirming that the SED fit has correctly fit the source as AGN. As we consider sources with less luminous AGN and a stronger host galaxy component, the composite SED becomes more "star-like" in structure and therefore less separated in χ^2_ν space from the stellar regions.

As the point-like sources become optically fainter, the χ^2_ν distribution becomes bi-modal between the stellar and the extragalactic regions of the figure, as seen in the $20.5 < I < 21.5$, $21.5 < I < 22.5$, and $22.5 < I < 23.5$ panels. The point-like sources that fall into the galaxy/AGN region of χ^2_ν space are a mix of point-like AGN and compact galaxies with a low stellarity index. In these three magnitude bins, $\sim 13\%$, 17%, and 24% of the extragalactic SEDs have $F > 3$ for the galaxy+AGN fits over the galaxy-only fits. This nominally corresponds to a $\sim 90\%$ probability that an AGN component is required modulo the caveats discussed in §3.1. As we examine fainter sources, the morphological star-galaxy separation fails more frequently and as a result, the distributions of χ^2_ν for extended and point-like sources look increasingly similar in the $22.5 < I < 23.5$ and $I > 23.5$ panels. Finally, by the time we reach the optically faintest bin, the distribution of χ^2_ν for point-like sources looks nearly identical to that of extended sources. The vast majority of these are extragalactic, and the stellar population is naturally excluded in this magnitude bin due to the criteria that optically faint $I > 23.5$ sources are required to have mid-IR detections with $[4.5] < 18$.

Figure 6 shows the expected and observed integrated number counts of the stellar and extragalactic populations. The expected number counts of galaxies are calculated using Ellis & Bland-Hawthorn (2007) and the stellar number counts are from the Besançon stellar population synthesis model (Robin et al. 2003). The observed counts simply use the χ^2_ν separation illustrated in Figure 4. The resulting number counts
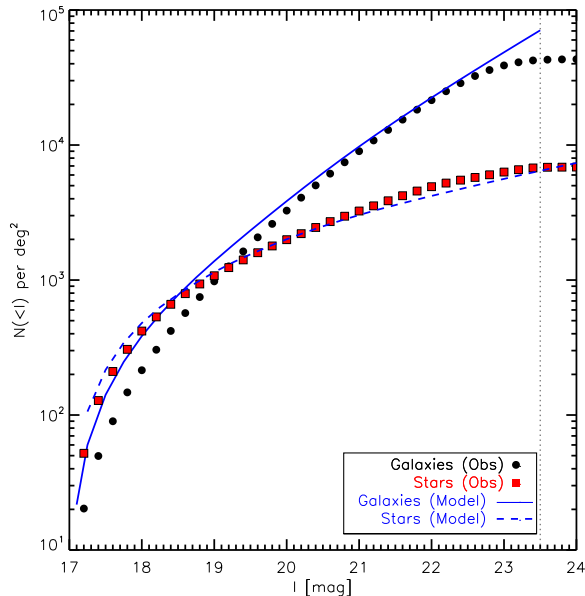
Fig. 6.— Integrated number of expected and observed ($I > 17$) stars and galaxies as a function of I magnitude. The dotted vertical line indicates the faint optical limit of our sample. Since we are using aperture magnitudes, the observed galaxy counts fall below the model at bright magnitudes.

for both the galaxy and stellar populations match the models quite well, although the completeness of the galaxy sample begins to drop rapidly for $I \gtrsim 23$. The mismatch for bright galaxies arises from using fixed $6\rlap{.}''0$ diameter aperture magnitudes rather than integrated Kron magnitudes. This underestimates the luminosities of the brighter galaxies and shifts the observed number counts to lie to the left of the models.

3.3.1. Outliers and Degeneracies

As discussed above and shown in Figure 4, the χ^2_ν separation of extragalactic and Galactic sources works well for the vast majority of our sample. However, as with any large sample based on automatic fits to data, there are bound to be some failures. In some cases this is due to a combination of inadequate data, and/or degeneracies between the galaxy and stellar templates, while in other cases it is due to one or more deviant data points that skew the fit.

Figure 7 explores possible degeneracies between the stellar templates and galaxies by showing the distribution of χ^2_ν(Star) as a function of the photometric redshift for the galaxy+AGN SED models as applied to the non-stellar sources (anything classified as a galaxy or AGN based on the χ^2_ν of their SED fits; see Figure 4). We do not include objects better fit as stars or brown dwarfs in Figure 7. The data are again split into extended (left) and point-like (right) sources and then further into I-band magnitude bins. The different colors represent best fit stellar/brown dwarf templates split into groups by temperature. Sources below the horizonal line at χ^2_ν(Star)=1.0 are galaxies or AGN that are nominally well-fit by a stellar template.

Among the optically bright galaxies and AGN, the best fitting stellar templates are overwhelmingly the cool stars, with some contribution from intermediate temperature stars, essentially because the SEDs of hotter stars peak at too short a wavelength. There is also a well-defined lower envelope to the χ^2_ν(Star) and z_p(Gal+AGN) distribution for the optically brighter galaxies and AGN. As the galaxy/AGN redshift
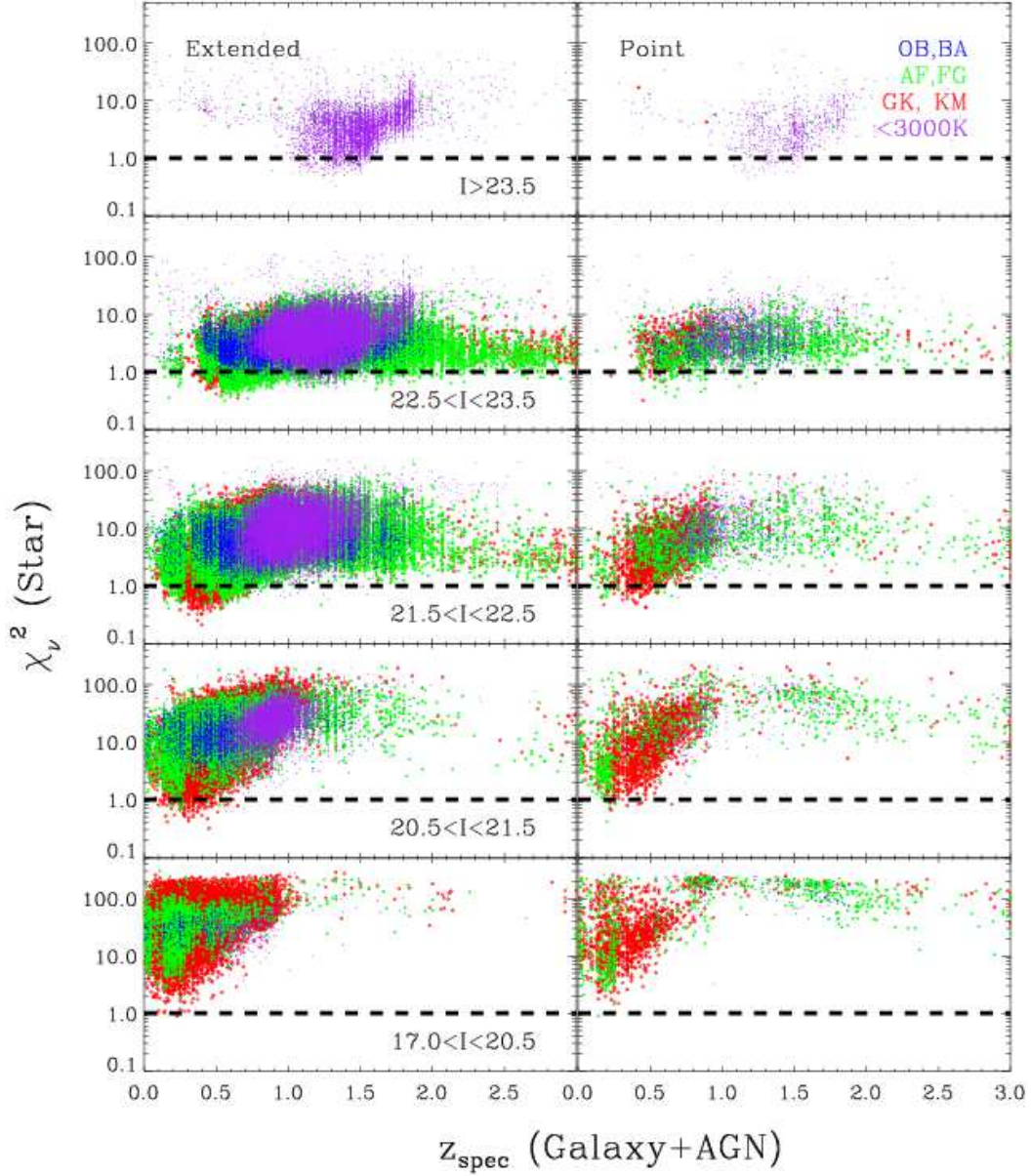
Fig. 7.— Distribution of non-stellar sources (as defined by the χ^2_ν separation in Figure 4) for the best fit stellar/brown dwarf template as a function of photometric redshift from the galaxy+AGN SED models, separated into extended (left) and point-like (right) sources, in bins of I-band magnitude. The color-coding indicates which of the stellar/brown dwarf templates produced the lowest χ^2_ν roughly split into "hot" (O/B and B/A), "intermediate" (A/F, F/G), "cool" (G/K, K/M), and "cold" ($< 3000K$) "stars", in blue, green, red, and purple, respectively.

increases, the χ_ν^2(Star) values rise sharply – optically bright $z \gtrsim 0$ galaxies and AGN are poorly fit by all stellar/brown dwarf templates.

As sources become fainter than $I \sim 21.5$ (top three panels), there tends to be more of a degeneracy between galaxy and stellar templates. This is reflected by the increasing number of sources for which stellar templates can produce reasonable values of $\chi_\nu^2 \sim 1$. The range of galaxy redshifts that are compatible with stellar/brown dwarf templates generally increases with decreasing optical magnitude (except in the case of $I > 23.5$, due to the [4.5] constraint). Also, as the galaxies become optically fainter, the cool star/brown dwarf templates (purple) become more prevalent, and overlap with higher redshift galaxies as well as for a wider range of galaxy redshifts. The intermediate and cool star templates (green and red) can fit some galaxies with a range of redshifts from $0.1 \lesssim z_p \lesssim 2.0$ with reasonable χ_ν^2 values in the $21.5 < I < 22.5$ magnitude bin. In the $22.5 < I < 23.5$ bin, the intermediate and cool stellar templates that are fixed at $z = 0$, can fit galaxies in almost the entire range of photometric redshifts and still yield χ_ν^2 values close to 1. In the $I > 23.5$, most sources are best fit with the $T_{\rm eff} < 3000$K templates. We did attempt to identify brown dwarfs using these fits but it works poorly because they are usually detected in only a few bands.

4.   The Search for AGN

While AGN emit energy in a broad range of wavelengths, most existing AGN selection techniques utilize only a narrow slice of the AGN spectrum such as the sources' optical or mid-IR color. Each of these AGN diagnostics is sensitive to a particular type or "viewing angle" in the unified picture of AGN and the balance between AGN and host emission (e.g. Hickox et al. 2009; Mendez et al. 2013). However, we have the advantage of having up to 17 bands that span the wavelength range from the near-UV to the mid-IR, so fitting galaxy and AGN templates to these data should allow for more complete AGN selection than using only a few colors within a limited wavelength window.

Figure 8 shows the distribution of sources in stellarity as a function of the F-ratio between the galaxy and galaxy+AGN templates, in bins of I-band magnitude for all sources except those classified as stars based on the χ_ν^2 separation shown in Figure 4. The horizontal line at $S = 0.7$ is the division we use between extended ($S < 0.7$) and point-like ($S \geq 0.7$) sources. The vertical line at $F = 3$ corresponds to a $\sim 90\%$ probability for the existence of an AGN component in the SED for the typical number of degrees of freedom (see §3.1). As discussed earlier, $z_s > 1$ sources are all AGN, so we use this as one comparison sample. As a second comparison sample we use XBoötes sources with at least 4 X-ray counts. The right panels of Figure 8 show the distribution of F-ratios for the entire non-stellar sample (galaxies and AGN) in black, and the X-ray and $z_s > 1$ AGN non-stellar samples in red and blue, respectively, where "non-stellar" is purely defined by having χ_ν^2(Galaxy+AGN)$<\chi_\nu^2$(Star). The panels are again separated into I-band magnitude bins, and the histograms are normalized to a peak of unity. The two AGN samples clearly tend to reside at higher F-ratios relative to the entire galaxy/AGN population, particularly for the brightest optical magnitudes.

Among the brightest sources, the $z_s > 1$ F-ratio distribution is roughly log normal and centered at $F \sim 50$, with most of the contribution coming from point-like sources. The X-ray sample similarly peaks at F$\sim$40, but with a much broader, assymetric distribution that extends towards low F-ratios, presumably because X-ray selection is more sensitive to host-dominated, spatially extended AGN than the selection methods that dominate the $z_s > 1$ AGN sample (Hickox et al. 2009; Mendez et al. 2013). However, for the $I \gtrsim 20$ AGN candidates in AGES, there are complex sampling biases. For example, point-like sources were selected more broadly than extended sources (e.g., all 24μm quasars were required to be optically point-like
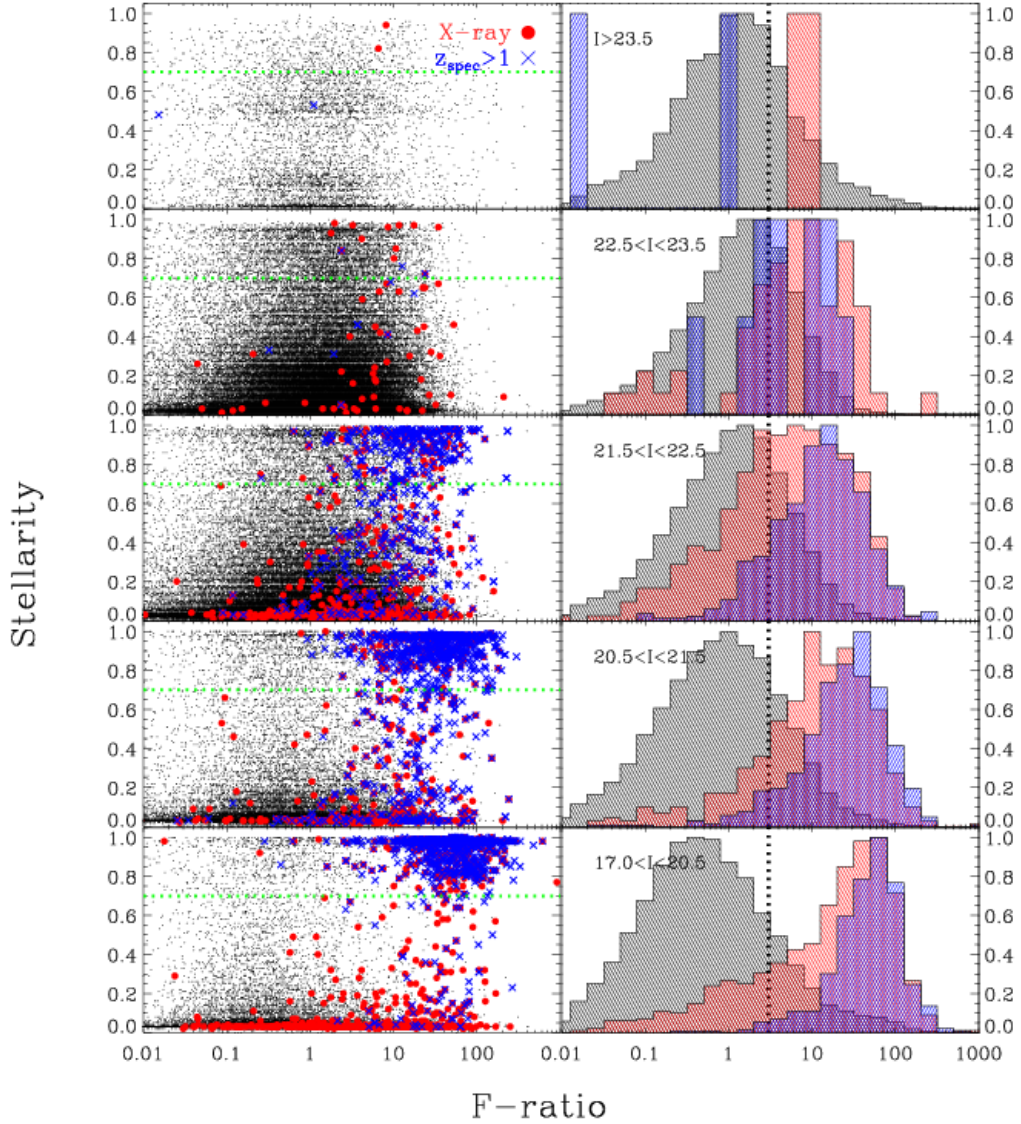
Fig. 8.— The left panels show stellarity (S) as a function of the F-ratio between the galaxy and galaxy+AGN models, in bins of I-band magnitude, excluding sources that are best fit by the stellar templates. The dotted horizontal line (left) at $S = 0.7$ shows our adopted separation between extended ($S < 0.7$) and point-like ($S \geq 0.7$) sources. The X-ray and $z_s > 1$ samples are shown using red circles and blue crosses, respectively. The right panels show the normalized distribution of F-ratios for the full, X-ray, and $z_s > 1$ samples in black, red, and blue histograms, respectively. The dotted vertical line at $F = 3$ is meant to provide only a visual aid, and corresponds to ∼90% AGN probability based purely on the F-ratio distribution (see caveats in §3.1. Note that the F-ratio distributions for the X-ray and $z_s > 1$ samples peak at larger F-ratios in all magnitude bins with the exception of the faintest bin, where there are only 2 X-ray and 2 $z_s > 1$ sources.

– see Kochanek et al. 2012). These sampling bias effects will be minimized for the $I < 20.5$ bin, and even here the X-ray sample shows a broader F-ratio distribution. It is unlikely to be an effect of obscuration since mid-IR SEDs are fairly immune to moderate levels of dust ($A_v \lesssim 10$, Assef et al. see 2011).

The only non-AGN that should appear in the X-ray samples are very low redshift galaxies, where X-ray emission by the integrated binary populations can dominate, and X-ray active stars. Out of the 136 X-ray sources with $F < 1$ in the optically bright bin of Figure 8, 113 have spectroscopic redshifts. Among these, nine sources have $z_s < 0.1$, four of which are stars ($z = 0$) and five which have SEDs that are consistent with star-forming galaxies or host-galaxy-dominated AGN. We also checked for a correlation between the off-axis angle of the source in the XBoötes observations and the F-ratios of the optically bright and extended sample. Due to the increasing size of the *Chandra*/ACIS point spread function (PSF) with distance from the center of the image, X-ray detections made at larger off-axis angles have an increased likelihood of being matched to the wrong optical counterpart. However, we find no relation between the low F-ratio X-ray sources and their X-ray off-axis angles, indicating that there are likely to be few false positives, and the tail of low F-ratio X-ray sources in Figure 8 is primarily from extended host-galaxy-dominated AGN. The spatially extended X-ray AGN tend to have higher F-ratios than the overall galaxy sample in all magnitude bins.

The left panels of Figure 8 also show that as the sources become optically faint, the median F-ratios of the point-like sources and the extended sources start to converge. In the optically brightest bin, the point-like sources are clustered at high F-ratios, while the extended sources cluster at significantly lower F-ratios. Yet in the faintest bin, the distinction between the extended and point-like sources in terms of their F-ratio values is less clear, presumably due to contamination by compact and/or high-redshift galaxies. In the fainter bins, we also have fewer X-ray and $z_s > 1$ sources due to the X-ray and redshift survey flux limits.

Figure 9 shows the F-test probability for the galaxy and galaxy+AGN models as a function of the host galaxy luminosity fraction $F_{\rm gal} = L_{\rm gal}/L_{\rm total}$ for extended (left) and point-like (right) sources in bins of I-band magnitude. Note that Assef et al. (2010) found that $F_{\rm gal}$ was relatively accurate even for sources with poor photometric redshift estimates. As explained in §3.1, low probabilities imply that the AGN component is unlikely to be improving the fits by chance. The dotted vertical line at $F_{\rm gal}$=0.5 indicates where the luminosity contribution from the AGN is equal to the contribution from the host galaxy, with the arrow pointing in the direction of increasing AGN fraction. The X-ray and $z_s > 1$ samples are overplotted in large red and blue filled circles, respectively. We will refer to AGN candidates with $F_{\rm gal} = L_{\rm gal}/L_{\rm total} > 0.5$ as "host-dominated".

As expected, we see that the extended-source AGN tend to be more host-dominated than their point-source counterparts. This is reassuring since the SED fits have no knowledge of the morphological information. Among the X-ray AGN sample, 74% of extended AGN are host-dominated, while only 23% of point-source AGN are host-dominated. For the $z_s > 1$ AGN sample, 33% and 10% of extended and point-source AGN are host-dominated, respectively. The difference between extended and point source AGN is less dramatic for the $z_s > 1$ sample compared to the X-ray sample, but this again cannot be trivially interpreted because morphology played a role in the selection process. Not surprisingly, there is a strong correlation between the F-test probability for the significance of an AGN component and the AGN luminosity fraction ($F_{\rm AGN} = 1 - F_{\rm gal}$) at bright magnitudes. In this regime ($I < 21.5$), it should be straightforward to select AGN, even when the total luminosity is dominated by the host galaxy. This continues to be true for point-like sources until $21.5 < I < 22.5$, but in the fainter bins it is unclear how well such a selection method would work given the available photometry.
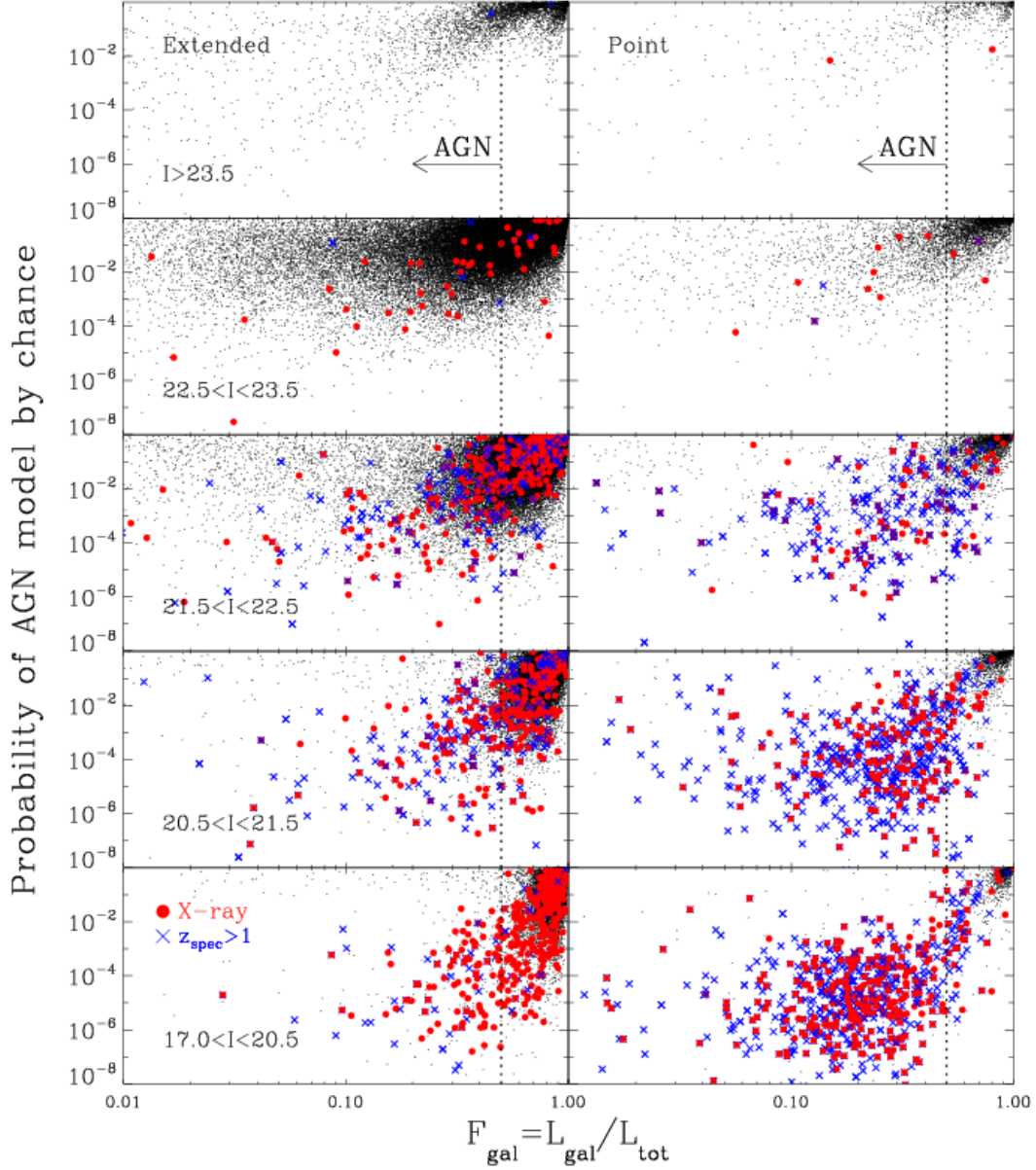
Fig. 9.— F-test probability for the significance of an AGN component as a function of host galaxy luminosity fraction (L_{gal}/L_{total}) for extended ($S < 0.7$, left) and point-like ($S \geq 0.7$, right) extragalactic sources in bins of I-band magnitude. Small probabilities mean that the improvement in the fit from adding the AGN component is unlikely to occur by chance. The vertical dotted line at $L_{gal}/L_{tot} = 0.5$ marks where the AGN and host contribution are equal, with increasing AGN fraction in the direction of the arrow. The X-ray and $z_s > 1$ samples are shown by the red circles and blue crosses, respectively.
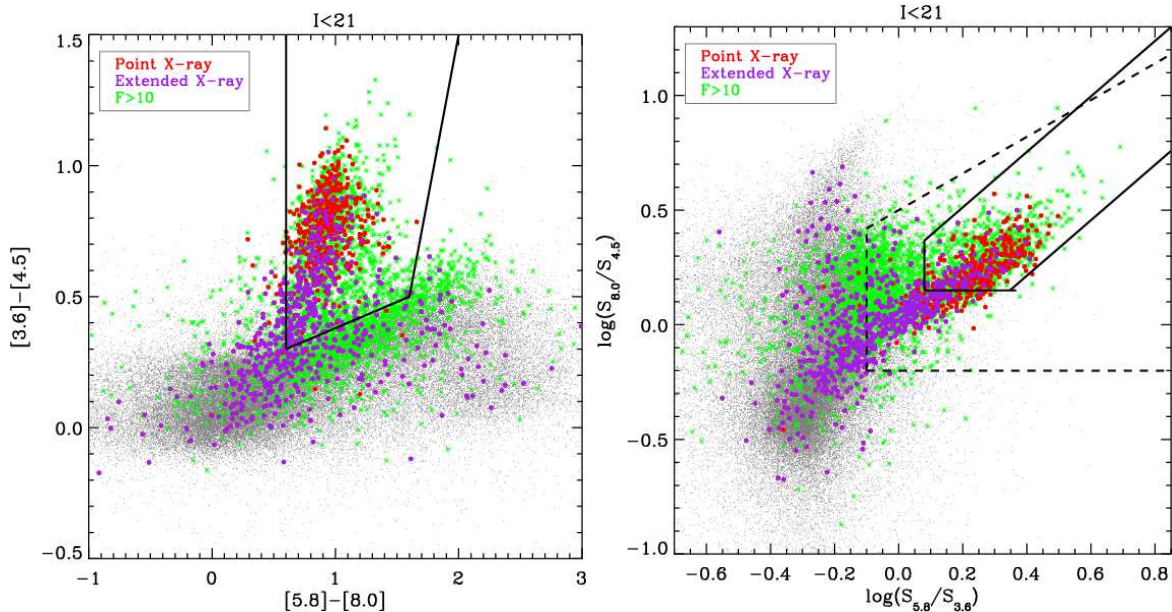
Fig. 10.— IRAC color-color diagram with X-ray and $F > 10$ sources overplotted in red and green symbols, respectively. Left: IRAC color criteria from Stern et al. (2005) shown in solid black. Right: IRAC color criteria from Lacy et al. (2004) and Donley et al. (2012) shown in dashed and solid black lines, respectively. The samples are limited to sources with $I < 21$.

4.1. Comparison to mid-IR and optical color selection

In the previous sections, we have shown that there is significant overlap between known samples of X-ray and $z_s > 1$ AGN to sources selected as AGN candidates based on the SED fitting and F-ratios. In this section, we will examine how the AGN sample selected from SED fitting compares to mid-IR color and optically selected AGN samples.

4.1.1. Mid-IR

As discussed in §1, the mid-IR SEDs of AGN are distinct from the SEDs of normal galaxies due to emission from the accretion disk (higher redshifts) or dust (lower redshift). If the luminosity of the AGN is at least comparable to that of the host galaxy, the composite SED in the infrared will resemble a power-law. Thus, mid-IR colors can be used to identify AGN candidates depending on the depth of the IR data and the relative strength of the AGN and the host component. In this section we investigate how the F-ratio values compare to the "standard" mid-IR color selections of Lacy et al. (2004), Stern et al. (2005), and Donley et al. (2012).

Figure 10 shows the IRAC colors for 77,277 non-stellar, $I < 21$ sources in our sample, as compared to 1053 X-ray sources and 3114 $F > 10$ sources. Of the 1053 X-ray sources, there are 413 point-like ($S \geq 0.7$) and 640 extended ($S < 0.7$) sources. The left panel shows the IRAC AGN selection criteria of Stern et al. (2005), while the right panel shows the AGN selection criteria of Lacy et al. (2004) and Donley et al. (2012). The Stern et al. (2005) and Lacy et al. (2004) criteria are broadly similar, with the Stern et al. (2005) wedge including 62% of all the X-ray AGN and the Lacy et al. (2004) wedge including 70%. The X-ray AGN that fall outside the Stern et al. (2005) and Lacy et al. (2004) wedge are mostly extended sources that have a
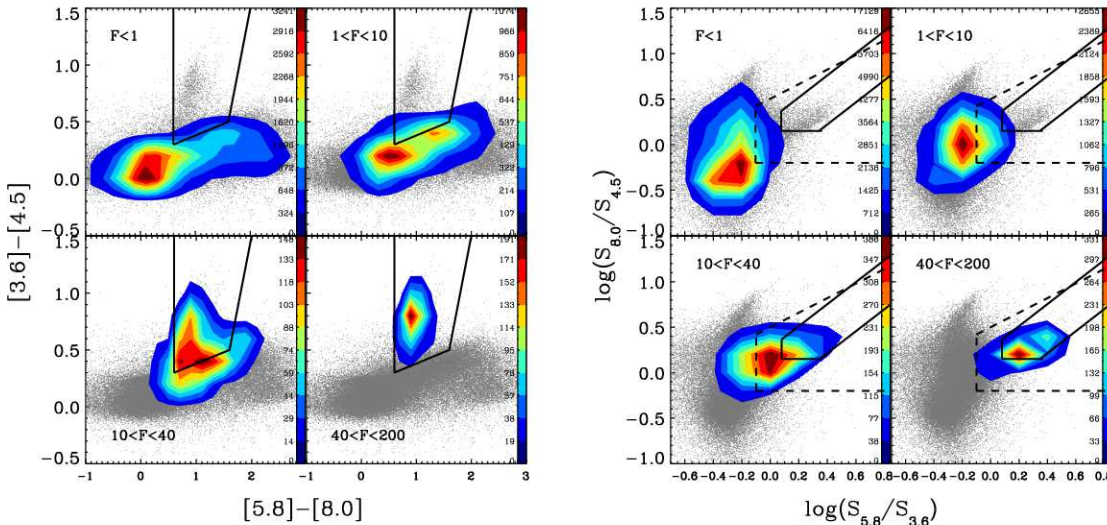
Fig. 11.— IRAC color-color diagram with contours showing the distribution of sources in increasing bins of F-ratio. Black lines illustrate mid-IR selection criteria, as decsribed in Figure 10. The grey cloud shows all non-stellar sources with $I < 21$.

significant luminosity contribution from their underlying host galaxies. The X-ray sources that are AGN-dominated (i.e., optical point sources) are nearly all identified as mid-IR AGN, with $\sim95\%$ falling within the Stern et al. (2005) and Lacy et al. (2004) wedges. This is consistent with the results of Gorjian et al. (2008), Cardamone et al. (2008), and Mendez et al. (2013), who find that mid-IR selection can miss a large fraction of X-ray identified AGN, especially moderate-luminosity or low accretion rate X-ray AGN. The Donley et al. (2012) color criteria for AGN selection is more strict, and includes only 35% of the overall X-ray AGN sample and 70% of the point source X-ray AGN sample shown in Figure 10. The fractions of $F > 10$ sources included in the Stern et al. (2005), Lacy et al. (2004), and Donley et al. (2012) AGN selection criteria are 57%, 75%, and 32%, respectively. The $F > 10$ sample generally spans a wider range of IRAC colors than the X-ray sources.

To better understand how F-ratio is correlated with mid-IR color, we examine the IRAC color distribution of four different F-ratio samples, as shown in Figure 11. Sources with low F-ratios ($F < 1$) are concentrated near $[3.6] - [4.5] \sim 0$ and $[5.8] - [8.0] \sim 0$ and form a tail out to redder $[5.8] - [8.0]$ colors. As the F-ratio increases, the distribution moves towards redder colors in both $[3.6] - [4.5]$ and $[5.8] - [8.0]$, until at the highest F-ratio bin ($F > 40$), the sources are almost exclusively located within the Stern et al. (2005) wedge.

Figure 11 shows that the F-ratios are qualitatively consistent with the mid-IR color selection of AGN, in the sense that low F-ratios have colors consistent with quiescent galaxies, whereas high F-ratios have colors consistent with mid-IR AGN. However, there is also an "intermediate" F-ratio population, many of which have bluer [3.6]-[4.5] colors. A similar trend is seen in the right panel of Figure 11, where IRAC colors become redder with increasing F-ratio, and the $F > 40$ sample lying almost exclusively within the Lacy et al. (2004) wedge.
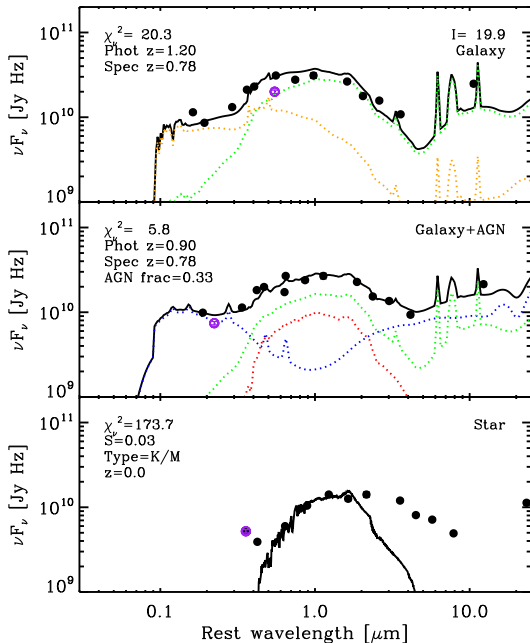
Fig. 12.— SED of an $F = 10$ source ($\alpha = 216.26442$ deg, $\delta = 33.62389$ deg) with IRAC colors that place it just outside the Stern et al. (2005) wedge. It is significantly better fit by a galaxy+AGN model (middle panel) than by a galaxy-only model (top) and its AGES spectrum contains a broad MgII emission line, confirming the presence of an AGN.

Figure 12 shows an example of an $F = 10$ SED for a source with IRAC colors that lie just blueward of the Stern et al. (2005) and Lacy et al. (2004) wedges. The data are clearly better fit when an AGN component is included in the SED model. The AGN component in this case contributes 33% of the bolometric luminosity, while the rest of the luminosity is attributed to the star-forming and starburst galaxy SED models. This is an example of an AGN with a host component that is strong enough that the mid-IR colors alone could not be used to identify this source as an AGN. There is however, an AGES spectrum of this source, and the presence of a broad Mg II λ2800 emission line confirms it as an AGN. While this source does have an X-ray counterpart, it is not included in our X-ray sample because it does not meet the ≥ 4 counts criterion.

There are 86 sources with AGES spectra that have SED fits with $10 < F < 15$, and IRAC colors that are $\sim$0.2 mag bluer than the bottom edge of the Stern et al. (2005) wedge. Of these 86 sources, only four were best-fit with a quasar template spectrum during the redshift cross-correlation procedure. However, many composite AGN/star-forming galaxies would not be best fit by the quasar template because they lack strong, broad emission lines. Using emission line ratio diagnostics, we can gain insight into the nature of the intermediate F-ratio SEDs that lie just outside the mid-IR AGN wedges.

Figure 13 again shows the IRAC mid-IR colors for non-stellar $I < 21$ sources, with the various AGN wedges indicated in solid and dashed lines. Objects that have been classified as either star-forming galaxies, AGN, or composite sources based on their Baldwin-Phillips-Terlevich (BPT; Baldwin et al. 1981) emission line ratios (Moustakas et al. 2011) are shown with the larger, colored symbols. Note that the BPT diagnostic can only be used for low redshift $z_s \lesssim 0.35$ narrow-line sources because the Hα emission line at λ6563Å is redshifted out of the AGES optical spectrum at $z \sim 0.4$. There are 3694, 920, and 420 star-forming galaxies, composite galaxies, and AGN shown in Figure 13. Among this sample of narrow emission line $z \lesssim 0.35$ AGES galaxies and AGN, we see that the star-forming galaxies and the AGN tend to occupy different,
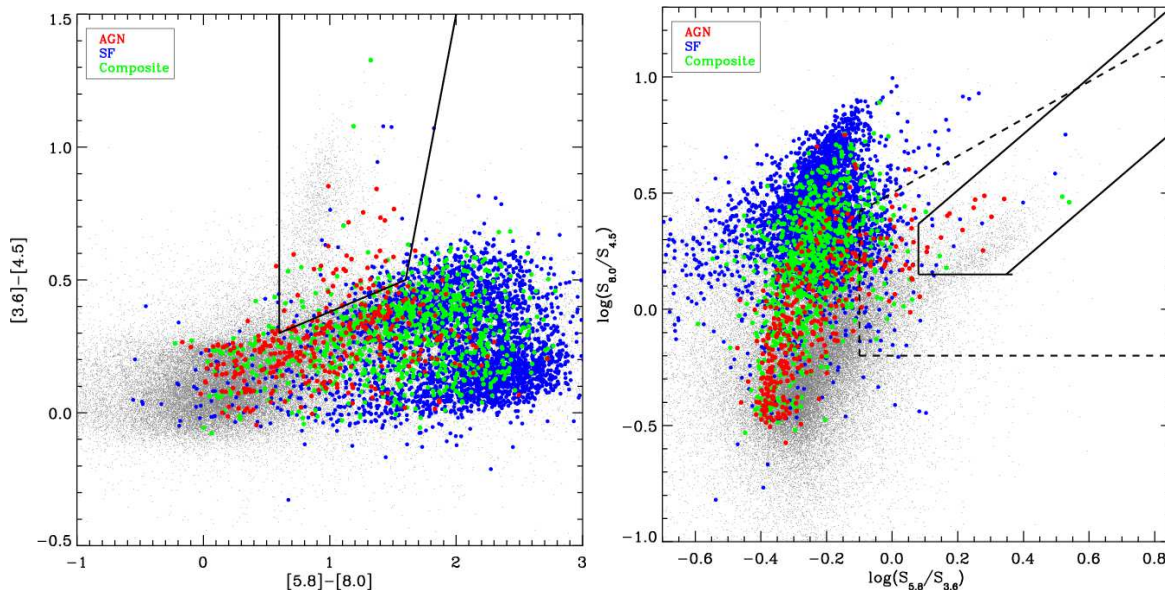
Fig. 13.— IRAC color-color diagrams with sources classified as star-forming galaxies, AGN, or composite galaxies based on BPT diagnostic from emission line ratios, overplotted in blue, red, and green symbols, respectively. Black lines illustrate mid-IR selection criteria, as described in Figure 10.

though overlapping, regions of IRAC color space, with the composite sources being clustered in between the two populations. The BPT AGN and composite sources have a color distribution roughly parallel to the blue edge of the Stern et al. (2005) wedge. This is similar to the color distribution of the $F > 10$ sources that extend beyond the Stern et al. (2005) wedge, shown in Figure 11 and Figure 13, strongly suggesting that the SED fits are identifying the narrow-line AGN and composite population. Also note that the mid-IR colors of the narrow line AGN/composite sources are different from that of the X-ray sources.

We examined the emission line ratios of 4,030 sources separated into bins of F-ratio in order to examine how the SED fitting technique compares to the BPT classification of narrow-line AGN, composite sources, and star-forming galaxies. Figure 14 shows the line ratio distribution of sources in F-ratio bins. The dotted and dashed lines demarcate regions on the BPT diagram where the emission line ratios can be explained by either H II regions (i.e., star-forming galaxies) or AGN (Kewley et al. 2006). Composite sources lie in the area between the two lines. Figure 14 does not show sources with $F > 20$ because we want to focus on the "ambiguous" cases.

Table 1 reports the distribution of these 4,030 sources in F-ratio and BPT classification. For $F < 1$, there are 2,973 sources and most of these (76%) are classified as H II regions and few (6%) as AGN. The AGN fraction then steadily increases as the F-ratio increases. However, even for the $10 < F < 20$ bin, only 44% of sources are classified as either a composite source or AGN. This is an underestimate of the true AGN fraction because the BPT sample explicitly excludes higher redshift ($z \gtrsim 0.35$), broad-lined AGN (Gaussian $\sigma > 500$ km s^{-1}).
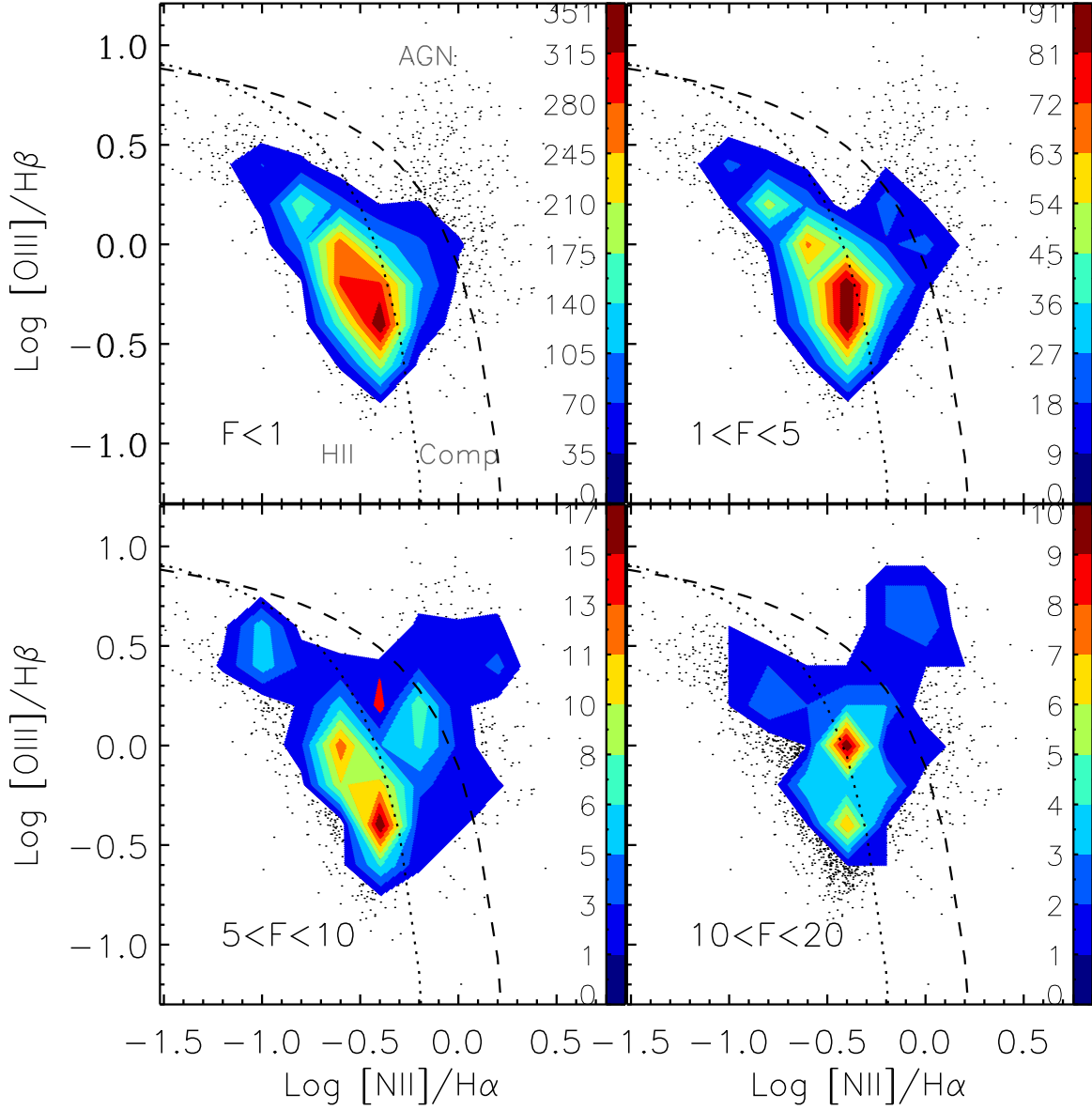
Fig. 14.— BPT diagram showing the emission line ratios of [OIII]/Hβ and [NII]/Hα of 4,030 sources. Each panel shows the density contours of sources in different *F*-ratio bins. The dotted and dashed lines demarcate regions dominated by star-forming galaxies (HII), composite sources, and narrow-line AGN. As the *F*-ratio increases, there is a general shift towards the composite and narrow-lined AGN regions of the diagram.

Table 1: BPT Classifications

	$F < 1$ (2973)	$1 < F < 5$ (828)	$5 < F < 10$ (158)	$10 < F < 20$ (71)
HII	76%	67%	59%	56%
Composite	18%	20%	22%	18%
AGN	6%	13%	19%	26%

Note. – The entries are the percentages of sources in F-ratio bins with HII region, composite, or AGN emission line ratio classifications. The total number of sources in each F-ratio bin is shown in parenthesis.

4.1.2.   SDSS colors

In this section we examine the distribution of the sources in their optical colors. Since we lack deep photometric data in the SDSS filters, we produced synthetic SDSS colors from the best fit SED models and the SDSS filter response curves. Figure 15 shows the synthetic SDSS colors of the 383,604 extragalactic sources and 47,434 stars, where the Galactic sources are defined by χ^2_ν(Galaxy+AGN)$>\chi^2_\nu$(Star). The X-ray, $z_s > 1$, and $F > 10$ samples are overplotted in red, blue, and pink, respectively. The stars and brown dwarfs are shown in green.

Quasar selection in SDSS is based on examining the colors of $i \lesssim 20$ point sources and excluding known stellar regions of color space (Richards et al. 2002). The blue and cyan boxes show the white dwarf and A star exclusion regions, and the green box indicates an inclusion region for $z \sim 2.7$ quasars which also have some optical colors similar to A stars. In the upper left panel of Figure 15 there is also a quasar inclusion region for sources bluer than $u-g = 0.6$, but outside the white dwarf exclusion box. This color cut is roughly equivalent to earlier UV excess (UVX) methods (e.g., Boyle et al. 1990).

The SDSS stellar locus (black asterisks) is very similar to the synthesized SDSS colors of our stellar sample (green curves), confirming that the extragalactic/Galactic separation based on χ^2_ν of the SED models is generally a success. The synthesized SDSS colors of the stellar templates (green) form two distinct curves for dwarfs and giants, respectively, with each curve formed by the sequence of spectral types.

In the upper right and lower panels of Figure 15, the $z_s > 1$ sample of AGN are roughly centered in the $z \sim 2.7$ quasar inclusion region from Richards et al. (2002), except in the $g - r$, $u - g$ color space where the $z_s > 1$ and X-ray AGN do not coincide with the $z \sim 2.7$ inclusion box. Figure 16 shows the $u - g$ color distribution among the extended and point-like sources for our non-stellar sample, including the $z_s > 1$, $F > 10$, and X-ray sources. The point-like AGN candidate samples show a relatively narrow $u - g$ color distribution in comparison to the extended sources, which is not surprising because point-like AGN have colors that are less contaminated by their host galaxies. Among the extended sources, it is clear that both the X-ray and $F > 10$ samples have $u - g$ colors that extend significantly beyond the $u - g < 0.6$ UVX color criterion, demonstrating the potential of the F-ratio method to find a diverse sample of AGN.

4.2.   AGN Surface Density

We can also examine the success of SED fitting as an AGN selection method by comparing the surface density of AGN candidates found from the F-test, optical, mid-IR, and X-ray selection methods. Figure 17 shows the integrated number of non-stellar sources in total and with $F > 5$, $F > 10$, and $F > 20$ along with
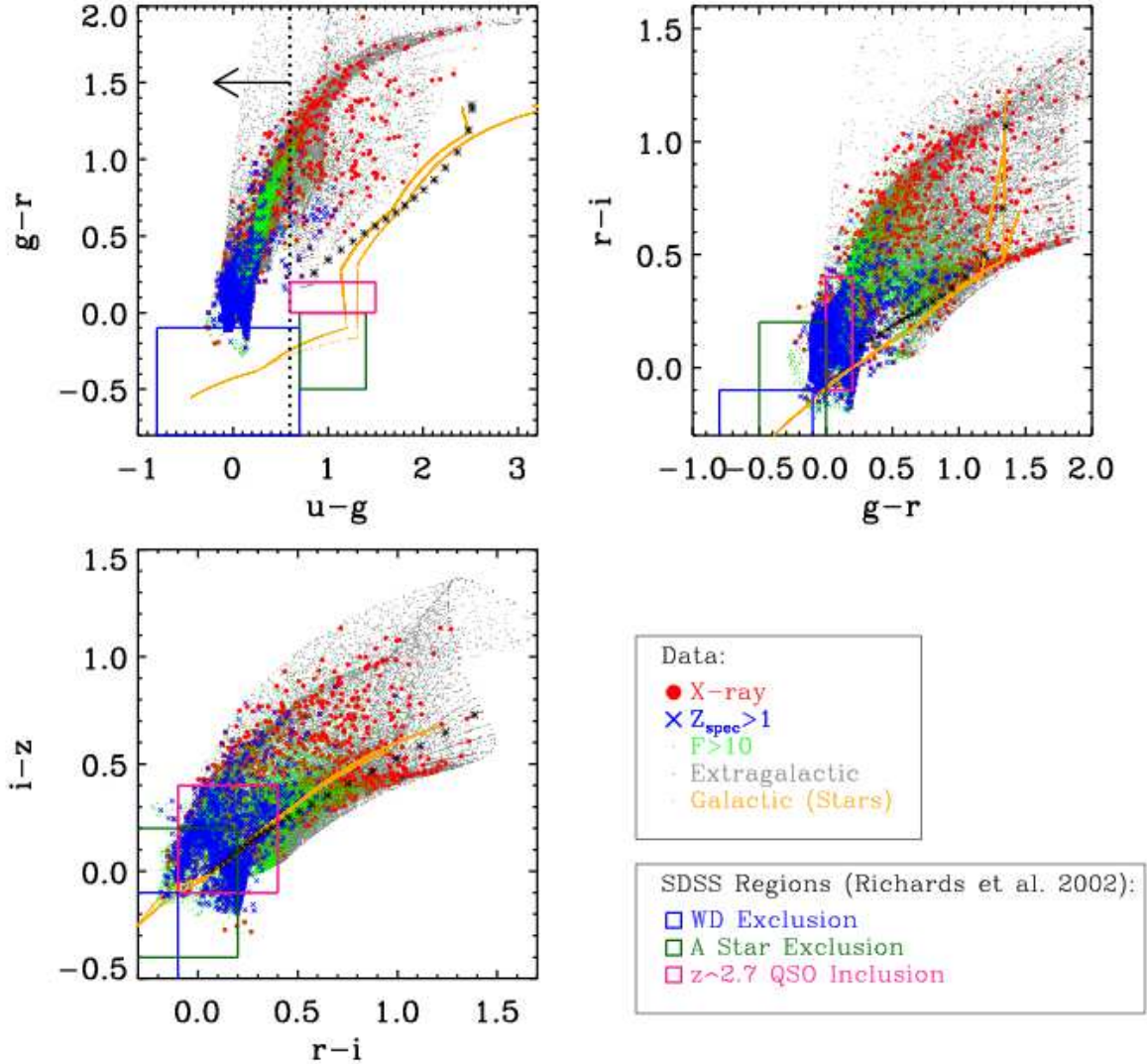
Fig. 15.— Synthesized SDSS colors for our sample of 383,604 extragalactic (gray dots) and 47,434 Galactic (orange dots) sources. Because we synthesize the SDSS colors from the template fits, the orange stellar source sequences trace out two curves formed by the giant and dwarf templates. The X-ray, $z_s > 1$, and $F > 10$ AGN samples are shown in red, blue, and green, respectively. The SDSS exclusion and inclusion zones adopted from Richards et al. (2002) are shown, with the white dwarf and A star exclusion zones shown as blue and dark green rectangles, and the $z \sim 2.7$ quasar inclusion zone is shown with the magenta rectangle. Black asterisks show the SDSS stellar locus (Richards et al. 2002). In the $g - r$, $u - g$ color space, the black dotted vertical line shows the equivalent UVX color criterion. Sources that avoid stellar exclusion zones and are blueward of $u - g = 0.7$ are considered quasar candidates. Most of the AGN and AGN candidates from our sample meet the UVX color criterion.
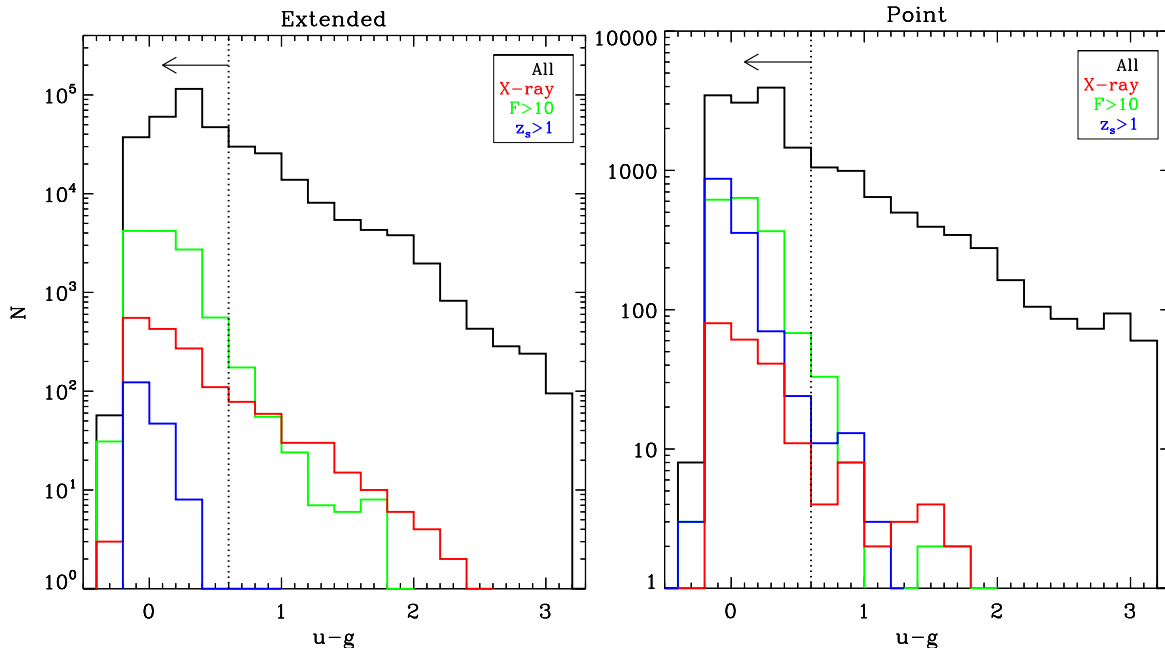
Fig. 16.— Distribution of $u-g$ color among extended (left) and point-like (right) non-stellar sources. The X-ray, $F > 10$, and $z_s > 1$ AGN candidate samples are shown in red, green, and blue histograms, respectively. The vertical dotted line at $u - g = 0.6$ highlights the SDSS UVX color criterion for AGN.

AGN surface densities from optical and X-ray surveys. We also show the number of SDWFS sources that satisfy the Stern et al. (2005) mid-IR selection criteria. As is typical of AGN distributions, the number rises steeply and then flattens at $I \sim 20$. The further flattening at fainter magnitudes is due to completeness and the effects of steadily increasing photometric errors on F.

The surface density of point-like, optically selected quasars from SDSS and the 2dF/6dF QSO Redshift Survey (2QZ/6QZ) adopted from Richards et al. (2006) is shown by a solid green curve in Figure 17. These quasars were selected based on their optical (SDSS) colors and are also required to be point sources in the magnitude range of $18.0 < g < 21.85$. We also show the number of $z_s > 1$ AGES sources, which are dominated by point-like, mid-IR selected AGN with $I < 21$ (Kochanek et al. 2012). While different in selection methods, these sources are largely broad line quasars similar to what is found in the SDSS or 2QZ surveys.

In Figure 11 we showed that the Stern et al. (2005) mid-IR selection region typically contained sources with $F \gtrsim 10$, and in Figure 17 we see that the surface density of $F > 10$ sources is very similar to that of sources which satisfy the Stern et al. (2005) mid-IR selection criteria. Mid-IR selection is relatively impervious to dust extinction and is not limited to optically point-like sources, so it is not surprising that it lies well above the SDSS/2QZ sample densities. At faint magnitudes, the mid-IR sample is limited by the need for $5.8\mu m$ and $8.0\mu m$ detections, and the method fails as the AGN becomes similar in luminosity to its host, so it is still an incomplete inventory of AGN. At least in AGES, there was very little contamination of the mid-IR sample by non-AGN, although this is likely a greater problem at $I > 21$ where there begin to be galaxies that can match the $[3.6] - [4.5]$ criteria and the $[5.8] - [8.0]$ color is becoming noisier.

Finally, we can compare to the shallow XBoötes and deep *Chandra* Deep Field-North (CDF-N; Brandt et al.
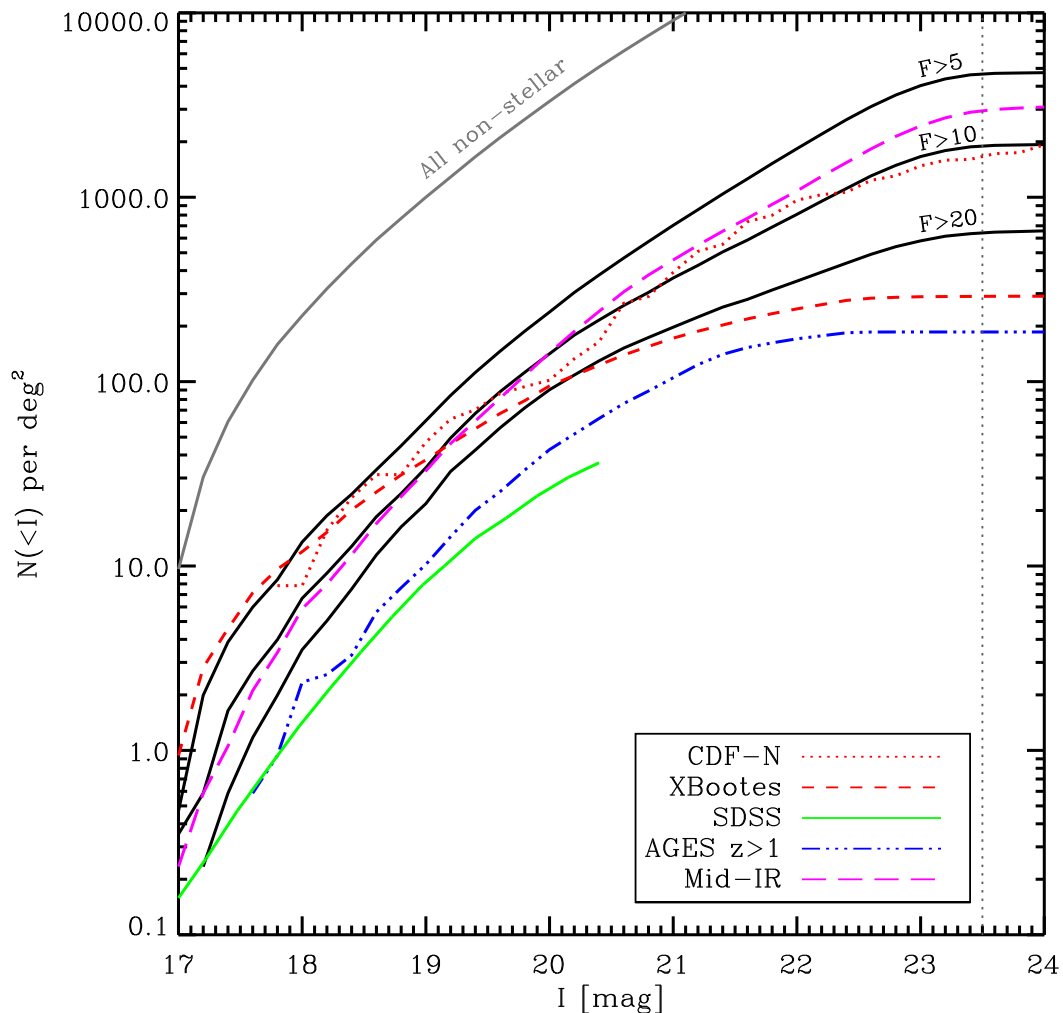
Fig. 17.— Integrated surface density of AGN candidates with $I > 17$ (number per deg^2) as a function of I magnitude. The three curves in solid, dotted, and dashed black lines show the surface density of AGN candidates from the $F > 5$, $F > 10$, and $F > 20$ samples. For comparison, we show the surface densities of AGN candidates selected from X-ray, optical, and mid-IR surveys in red, green, and magenta, respectively. The blue AGES sample is a spectroscopic sample of $z_s > 1$ AGN. The vertical dotted line at $I = 23.5$ shows the optical magnitude limit imposed on our sample. The F-ratio selection of AGN yields higher AGN surface densities than the optically selected samples.

2001) X-ray selected samples. These should identify all AGN other than the highly obscured, Compton thick population. For the CDF-N sample, we exclude sources with low X-ray to optical flux ratios (i.e. $\log(f_X/f_R) < -2$) because the X-ray emission of these sources is likely of stellar origin, and any objects that have been spectroscopically identified as stars. Due to its small area, the CDF-N has few bright AGN, but for $18 \lesssim I \lesssim 20$, the XBoötes and CDF-N surface densities are comparable. The shallow XBoötes survey is increasingly incomplete for $I > 20$. For $I \lesssim 19$, the density of the $F > 5$ sample is similar to that of X-ray sources but then rises to be a factor of ~ 2 higher at fainter magnitudes. The $F > 10$ sample is lower than the X-ray samples at bright magnitudes but has comparable number counts at the faint end.

These comparisons of surface densities are consistent with the results in §4.1. The $F \gtrsim 20$ sample will yield large numbers of luminous, broad-line AGN similar to those found in optical surveys. The sources with $F \gtrsim 5 - 10$ will be similar to many of those found in X-ray surveys but with an increasing false positive rate for lower values of F. The $10 \lesssim F \lesssim 20$ sources appear to track the composite population found by emission line diagnostics. We find a total of 16,266 sources with $F > 10$ at $I < 23.5$, which yields a surface density of 1904 AGN deg^{-2}.

5. Summary and Future Work

Using photometric data that ranges from the far-UV to the mid-IR, we fit galaxy, AGN, stellar, and brown dwarf SED models to 431,038 sources in the Boötes NDWFS field. The photometric separation of stars and galaxies compares well with morphological separation and star and galaxy surface densities at all magnitudes. Comparing to the over 20,000 available spectroscopic redshifts, we find photometric redshift dispersions of $\sigma/(1+z)$=0.040 and $\sigma/(1+z) = 0.169$ for the galaxy and AGN samples, respectively, after clipping the worst 5% of sources. In practice, the AGES survey (Kochanek et al. 2012) obtained spectroscopic redshifts for a large fraction of the most problematic sources, the luminous broad line AGN, so the photometric redshift dispersion of the sources without spectra should be lower. We estimated the likelihood of an AGN component based on F-test comparisons of the fits. The results for all the sources are reported in Table 2.

When we examine the distribution of F-ratios as a function of morphology and host luminosity fraction, we see the expected trends that high F-ratio sources tend to be more point-like and have smaller host galaxy contributions to their SEDs. X-ray sources show a broader distribution of F-ratios and host galaxy contributions. We also examined the distribution of sources relative to the Stern et al. (2005), Lacy et al. (2004) and Donley et al. (2012) mid-IR selection criteria. As a comparison, 62%, 70% and 35% of the XBoötes X-ray sources satisfy these criteria. There is a clear trend of the sources moving into these selection regions as F increases, with 57%, 75% and 32% of the sources with F-ratios $F > 10$ for adding an AGN component falling in these mid-IR selection regions. Sources with $F > 40$ lie almost exclusively in these regions. The $F \sim 10$ sources have significant host contributions, and many lie just bluewards of these selection regions in their $[3.6] - [4.5]$ colors. They partially overlap the mid-IR color distribution of X-ray sources, but they extend towards redder $[5.8] - [8.0]$ colors. Their mid-IR color distribution is very similar to that of narrow line sources with "composite" line ratios indicative of a mixture of star formation and AGN activity, which is also consistent with their redder $[5.8] - [8.0]$ colors. Similarly, we can use the F-ratio distribution to examine how AGN extend outwards from the SDSS optical color selection regions as the host galaxy contribution becomes more important.

These results strongly suggest that a sample of $F > 10$ sources can provide a more complete inventory

of AGN activity than any of the methods restricted to limited wavelength regimes. Like any method, it is not perfect. There are clearly a minority of X-ray AGN with such low F-ratios that they cannot be identified based on their UV-IR SEDs with an acceptable false positive rate. What needs to be calibrated at this point are the false positive rates as a function of the apparent significance of the AGN contribution. The surface densities of $F > 10$ sources as compared to all extragalactic sources and other AGN samples suggests that the false positive rate should be acceptably low, but this needs to be tested with spectroscopy. For $F \sim 5$, the false positive rate is likely unacceptably high. The problem for spectroscopy is that many redshift ranges will lack the emission line diagnostics needed to classify the nature of the source.

Most surveys of evolution separate the study of galaxies and AGN because they have difficulties tracking the populations which are strong mixtures of both. This large scale decomposition of the SEDs into host and AGN components provides a means of studying the co-evolution of these populations with relatively reliable estimates of the two components separately. Even when the models have only photometric redshifts, Assef et al. (2010) found that the estimates of the host galaxy luminosity fraction were relatively robust. In particular, it should be feasible to examine the duty cycle of AGN activity in galaxies as a function of redshift. For example, if the false positive rates for AGN activity can be calibrated as a function of F using modest spectroscopic samples, then the full survey sample can be used in studies of the evolution of galaxies and AGN.

These very broad baseline SED models should be comparably "stable" to surveys using large numbers of narrower filters over smaller wavelength ranges (e.g., Geach et al. 2008; Abramo et al. 2012). Instead of trying to better identify spectral breaks or strong emission lines, the broad structure of galaxy SEDs with a peak in the near-IR allows robust photometric redshifts and the structure of the UV and mid-IR tails of the distribution provides a robust probe of star formation and AGN activity. This approach should be particularly valuable for integrating wide area, multi-wavelength surveys in the ultraviolet (*GALEX*), optical (SDSS, DES[1], Pan-STARRS[2], LSST[3]), near-IR (VISTA, Emerson et al. (2006), *EUCLID*[4], *WFIRST*[5]) and mid-IR (*WISE*, Wright et al. (2010), *Akari* (Murakami et al. 2007)).

---

[1]http://www.darkenergysurvey.org/

[2]http://pan-starrs.ifa.hawaii.edu/public/

[3]http://www.lsst.org/

[4]http://sci.esa.int/euclid/45403-mission-status/

[5]http://wfirst.gsfc.nasa.gov/

Table 2. Model results

RA [deg]	Dec [deg]	I [mag]	S	z_s	N	z_p	χ^2_ν	L_{Ell}	L_{Sbc}	L_{Irr}	z_p	χ^2_ν	L_{AGN}	L_{Ell}	L_{Sbc}	L_{Irr}	F	E(B-V)	χ^2_ν	Stellar Type
								Galaxy Only						Galaxy+AGN						
217.43651	34.12883	17.70	0.99	...	13/1	0.47	36.08	1.416	Zero	Zero	0.47	45.10	Zero	1.416	Zero	Zero	0.00	0.00	9.50	K/M
219.23635	34.12883	22.68	0.42	...	14/2	0.60	0.98	Zero	-1.463	-0.148	0.55	0.79	-0.802	Zero	-0.861	-0.366	2.42	0.00	2.12	f/g
219.53921	34.12883	20.77	0.03	...	12/3	0.25	1.52	-1.114	Zero	-0.407	0.25	1.86	Zero	-1.114	Zero	-0.407	0.00	0.00	14.10	f/g
218.86312	34.12883	22.27	0.02	...	15/0	1.80	1.02	Zero	1.191	1.313	1.50	1.21	1.275	0.133	1.053	0.233	0.15	0.08	28.95	f/g
219.12657	34.12884	20.00	0.03	...	14/1	0.54	3.24	0.598	0.474	Zero	0.53	3.84	0.027	0.711	-0.617	-0.594	0.14	0.40	28.61	2000K
218.70934	34.12884	22.97	0.37	...	14/0	1.54	1.17	0.043	0.259	1.042	1.53	0.57	1.074	0.652	Zero	-0.954	6.23	0.04	6.18	a/f
218.04767	34.12886	21.91	0.73	...	12/3	0.24	0.89	-0.918	Zero	Zero	0.24	1.09	Zero	-0.918	Zero	Zero	0.00	0.00	0.59	K/M
216.28490	34.12886	19.51	0.98	...	14/0	0.39	25.40	0.551	Zero	Zero	0.39	32.66	Zero	0.551	Zero	Zero	0.00	0.00	7.02	K/M
217.74110	34.12886	22.52	0.17	...	12/2	1.93	2.06	Zero	Zero	1.367	1.55	2.29	1.55	1.671	0.827	Zero	2.68	0.05	3.96	f/g
219.01851	34.12886	21.68	0.16	...	13/2	0.61	1.24	Zero	0.479	Zero	0.61	1.52	Zero	Zero	0.479	Zero	0.00	0.00	5.65	K/M
216.70422	34.12887	21.90	0.03	...	12/2	0.73	1.33	Zero	-0.789	0.319	0.65	1.45	-0.533	-0.802	Zero	0.121	0.60	0.00	4.71	f/g

Note. — RA/Dec are the source coordinate in decimal degrees, I is the $6''0$ diameter aperture magnitude, S is the SExtractor stellarity index, z_s is the AGES spectroscopic redshift if available and N is the number of bands with detections/limits used in the fits. For the Galaxy and Galaxy+AGN fits, z_p is the photometric redshift, χ^2_ν is the goodness of fit, L_{Ell}, L_{Sbc}, L_{Irr} and L_{AGN} are the luminosities of the template components in units of $\log(L/10^{10} L_\odot)$. Luminosities are calculated for the fit at the spectroscopic redshift if known, and an entry of "Zero" means that the best fit included no contribution from that template. The Galaxy+AGN section also include the reddening $E(B-V)$ applied to the AGN template and the F-ratio value compared to the Galaxy-only fit. The stellar fits include χ^2_ν and the best fit template, where K/M means the best fit lay between the K and M stellar templates and 2000 K means the best fit was the 2000 K brown dwarf template. Uppercase letters refer to giant spectral templates and lowercase letters refer to dwarf spectral templates. If used in other contexts, the spectroscopic redshifts and photometry should be referenced to Kochanek et al. (2012) and Brown et al. (2007), respectively.

REFERENCES

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, ApJS, 182, 543

Abramo, L. R., Strauss, M. A., Lima, et al. 2012, MNRAS, 423, 3251

Alexander, D. M., Stern, D., Del Moro, A., et al. 2013, ApJ, 773, 125

Allard, F., Allard, N. F., Homeier, D., Kielkopf, J., McCaughrean, M. J., & Spiegelman, F. 2007, A&A, 474, L21

Alonso-Herrero, A., Pérez-González, P. G., Alexander, D. M., et al. 2006, ApJ, 640, 167

Ashby, M. L. N., Stern, D., Brodwin, M., et al. 2009, ApJ, 701, 428

Assef, R. J., Kochanek, C. S., Ashby, M. L. N., et al. 2011, ApJ, 728, 56

Assef, R. J., Kochanek, C. S., Brodwin, M., et al. 2010, ApJ, 713, 970

Assef, R. J., Stern, D., Kochanek, C. S., et al. 2013, ApJ, 772, 26

Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, PASP, 93, 5

Bertin, E. & Arnouts, S. 1996, A&AS, 117, 393

Bian, F., Fan, X., Jiang, L., et al. 2013, ArXiv e-prints

Boyle, B. J., Fong, R., Shanks, T., & Peterson, B. A. 1990, MNRAS, 243, 1

Brand, K., Brown, M. J. I., Dey, A., et al. 2006, ApJ, 641, 140

Brandt, W. N., Alexander, D. M., Hornschemeier, A. E., et al. 2001, AJ, 122, 2810

Brodwin, M., Brown, M. J. I., Ashby, M. L. N., et al. 2006, ApJ, 651, 791

Brown, M. J. I., Dey, A., Jannuzi, B. T., et al. 2007, ApJ, 654, 858

Cardamone, C. N., Urry, C. M., Damen, M., et al. 2008, ApJ, 680, 130

Cardamone, C. N., Urry, C. M., Schawinski, K., Treister, E., Brammer, G., & Gawiser, E. 2010, ApJ, 721, L38

Castelli, F. & Kurucz, R. L. 2004, ArXiv Astrophysics e-prints

Colless, M., Dalton, G., Maddox, S., et al. 2001, MNRAS, 328, 1039

Collister, A. A. & Lahav, O. 2004, PASP, 116, 345

Cool, R. J. 2007, ApJS, 169, 21

Dickinson, M., Giavalisco, M., & GOODS Team. 2003, in The Mass of Galaxies at Low and High Redshift, ed. R. Bender & A. Renzini, 324

Donley, J. L., Koekemoer, A. M., Brusa, M., et al. 2012, ApJ, 748, 142

Donley, J. L., Rieke, G. H., Pérez-González, P. G., Rigby, J. R., & Alonso-Herrero, A. 2007, ApJ, 660, 167

Eisenhardt, P. R., Stern, D., Brodwin, M., et al. 2004, ApJS, 154, 48

Eisenhardt, P. R. M., Griffith, R. L., Stern, D., et al. 2010, AJ, 139, 2455

Ellis, S. C. & Bland-Hawthorn, J. 2007, MNRAS, 377, 815

Elston, R. J., Gonzalez, A. H., McKenzie, E., et al. 2006, ApJ, 639, 816

Emerson, J., McPherson, A., & Sutherland, W. 2006, The Messenger, 126, 41

Fabricant, D., Fata, R., Roll, J., et al. 2005, PASP, 117, 1411

Fan, X., Strauss, M. A., Becker, R. H., et al. 2006, AJ, 132, 117

Gabasch, A., Hopp, U., Feulner, G., et al. 2006, A&A, 448, 101

Geach, J. E., Smail, I., Best, P. N., et al. 2008, MNRAS, 388, 1473

Gonzalez, A. H., Brodwin, M., Brown, M. J. I., et al. 2010, in American Astronomical Society Meeting Abstracts, Vol. 216, American Astronomical Society Meeting Abstracts 216, 415.13

Gorjian, V., Brodwin, M., Kochanek, C. S., et al. 2008, ApJ, 679, 1040

Griffith, R. L. & Stern, D. 2010, AJ, 140, 533

Harrison, F. A., Craig, W. W., Christensen, F. E., et al. 2013, ApJ, 770, 103

Hickox, R. C., Jones, C., Forman, W. R., et al. 2009, ApJ, 696, 891

Hoopes, C. G. 2004, in Bulletin of the American Astronomical Society, Vol. 36, American Astronomical Society Meeting Abstracts 204, 746

Hopkins, P. F., Bundy, K., Hernquist, L., & Ellis, R. S. 2007, ApJ, 659, 976

Hopkins, P. F., Hickox, R., Quataert, E., & Hernquist, L. 2009, MNRAS, 398, 333

Huchra, J. P., Macri, L. M., Masters, K. L., et al. 2012, ApJS, 199, 26

Ivezić, Ž., Menou, K., Knapp, G. R., et al. 2002, AJ, 124, 2364

Jannuzi, B., Weiner, B., Block, M., et al. 2010, in Bulletin of the American Astronomical Society, Vol. 42, American Astronomical Society Meeting Abstracts 215, 470.01

Jannuzi, B. T. & Dey, A. 1999, in Astronomical Society of the Pacific Conference Series, Vol. 193, The Hy-Redshift Universe: Galaxy Formation and Evolution at High Redshift, ed. A. J. Bunker & W. J. M. van Breugel, 258

Kellermann, K. I., Sramek, R., Schmidt, M., Shaffer, D. B., & Green, R. 1989, AJ, 98, 1195

Kenter, A., Murray, S. S., Forman, et al. 2005, ApJS, 161, 9

Kewley, L. J., Groves, B., Kauffmann, G., & Heckman, T. 2006, MNRAS, 372, 961

Kochanek, C. S., Eisenstein, D. J., Cool, R. J., et al. 2012, ApJS, 200, 8

Lacy, M., Storrie-Lombardi, L. J., Sajina, A., et al. 2004, ApJS, 154, 166

Lin, H., Yee, H. K. C., Carlberg, R. G., & Ellingson, E. 1996, JRASC, 90, 337

Mateos, S., Alonso-Herrero, A., Carrera, F. J., et al. 2012, MNRAS, 426, 3271

Mendez, A. J., Coil, A. L., Aird, J., et al. 2013, ApJ, 770, 40

Moustakas, J., Zaritsky, D., Brown, M., et al. 2011, ArXiv e-prints

Murakami, H., Baba, H., Barthel, P., et al. 2007, PASJ, 59, 369

Murray, S. S., Kenter, A., Forman, W. R., et al. 2005, ApJS, 161, 1

Netzer, H., Lutz, D., Schweitzer, M., et al. 2007, ApJ, 666, 806

Polletta, M. d. C., Wilkes, B. J., Siana, B., et al. 2006, ApJ, 642, 673

Pović, M., Sánchez-Portal, M., Pérez García, A. M., et al. 2012, A&A, 541, A118

Richards, G. T., Croom, S. M., Anderson, S. F., et al. 2005, MNRAS, 360, 839

Richards, G. T., Fan, X., Newberg, H. J., et al. 2002, AJ, 123, 2945

Richards, G. T., Strauss, M. A., Fan, X., et al. 2006, AJ, 131, 2766

Robin, A. C., Reylé, C., Derrière, S., & Picaud, S. 2003, A&A, 409, 523

Roll, J. B., Fabricant, D. G., & McLeod, B. A. 1998, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 3355, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, ed. S. D'Odorico, 324–332

Rowan-Robinson, M., Babbedge, T., Oliver, S., et al. 2008, MNRAS, 386, 697

Salvato, M., Hasinger, G., Ilbert, O., et al. 2009, ApJ, 690, 1250

Sanders, D. B., Phinney, E. S., Neugebauer, G., Soifer, B. T., & Matthews, K. 1989, ApJ, 347, 29

Schneider, D. P., Hall, P. B., Richards, G. T., et al. 2007, AJ, 134, 102

Schneider, D. P., Richards, G. T., Hall, P. B., et al. 2010, AJ, 139, 2360

Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1

Shapley, A., Fabbiano, G., & Eskridge, P. B. 2001, ApJS, 137, 139

Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163

Stengler-Larrea, E. A., Boksenberg, A., Steidel, C. C., et al. 1995, ApJ, 444, 64

Stern, D., Assef, R. J., Benford, D. J., et al. 2012, ApJ, 753, 30

Stern, D., Djorgovski, S. G., Perley, R. A., de Carvalho, R. R., & Wall, J. V. 2000, AJ, 119, 1526

Stern, D., Eisenhardt, P., Gorjian, V., et al. 2005, ApJ, 631, 163

Tueller, J., Mushotzky, R. F., Barthelmy, S., et al. 2008, ApJ, 681, 113

Urry, C. M. & Padovani, P. 1995, PASP, 107, 803

Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868

Yan, L., Donoso, E., Tsai, C.-W., et al. 2013, AJ, 145, 55