# PAPER I.

Marat Musin [1] [2], Jiasheng Huang [1], Haojing Yan [2],
*Draft version July 5, 2019*

## ABSTRACT

Context: What can be the use of this catalog?
Aims: We present a catalog of XX galactic and 14 million extragalactic sources with photometric redshifts in Stripe 82 with deep data from SDSS and WISE.
Methods: In order to get consistent photometry, optical sources are convolved to the same PSF for the WISE data we perform "template fitting" - technique that uses high resolution optical sources as a prior to obtain total flux of the low resolution near-IR sources.
Results: T-PHOT code that does "template fitting" uses high resolution and low resolution PSFs to extract robust flux even when source have little flux in near-IR and would not appear in WISE catalog or when sources are blended. Morphological star-galaxy separation is performed and then for each source four photometric redshifts are calculated using SED fitting code (with and without near-IR bands) EAZY and machine learning code ANNz (using two different training set). Available data allows to retrieve photometric redshift up to z∼0.8 with the average scatter XX.
Conclusions: Our final catalog with photometry and appropriate errors, CLASS_STAR indices, photometric redshift and cross-matched spectroscopic data within Stripe 82 is publicly available at XXX.XXX.com
*Keywords:* infrared: galaxies — submillimeter: galaxies — galaxies: starburst — methods: data analysis

## 1. INTRODUCTION

### 1.1. *Lilly-Madau formalism*

### 1.2. *SED fitting as a standard technique of mass and redshift estimation*

SED fitting is now a standard technique of deriving stellar mass and photometric redshifts for a large set of galaxies. In this method multi-band photometry for a given galaxy is fitted to a series of a templates predicted by a certain stellar population synthesis (SPS) model. The best-fit template gives the parameters of the galaxy, including its redshift and mass. Historically, SPS models were using restframe optical photometry. One caveat is the degeneracy between the dust extinction and age of the stellar population, as both make the color of galaxy red, i.e. galaxy can be red because it is intrinsically red with no young massive star and ongoing star-formation, or it can be very dusty, or it can be metal-rich and metals effectively absorb light in the bluer bands. Solution to this is to implement restframe near-IR where light suffers much less extinction (comparing to restframe UV and optical) and thus the degeneracy can be broken. We aim to build the largest sample of galaxies with optical and near-IR photometry over a large sky area. The natural choice for us then is to use optical Sloan Digital Sky Survey (SDSS) and IR all-sky data from Wide-Field Infrared Survey Explorer (WISE).

### 1.3. *problems associated with construction of the catalog*

[1] Chinese Academy of Sciences South America Center for Astronomy (CASSACA), National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, China
[2] Department of Physics & Astronomy, University of Missouri, Columbia, MO 65211, USA

Blending, poor spatial resolution in IR our method – template fitting

### 1.4. *Goal of this paper*

In this paper we present our technique for construction a catalog of galaxies with reliable SED data in optical and near-IR in Stripe 82 field. We discuss data selection, sources identification in different bands and problems associated with it.

## 2. DATA DESCRIPTION

### 2.1. *SDSS and Stripe 82*

The imaging component of the SDSS, which was done in five broad bands $(u'g'r'i'z')$, has covered 14,555 deg$^2$. In most area, the SDSS only scanned for one pass at an exposure time of 53.9 seconds per band, and thus is rather shallow (for example, the $r'$-band 5 $\sigma$ limiting magnitude is 22.2 mag). For this reason, in most cases the SDSS can only probe the normal galaxy population up to $z \approx 0.4$. However, the Stripe 82 region, which is a long stripe along the equator that spans $20^h < RA < 4^h$ and $-1.26^o < Dec < 1.26^o$, totalling in ≈300 deg$^2$ is the exception. It was repeatedly scanned (∼70-90 times, depending on RA) for calibration purpose during the survey (Adelman-McCarthy et al. 2007), and thus the combined scans can reach much better sensitivities.

A number of teams have created deep Stripe 82 stacks and made them available to public. The first such stacks were produced by Annis et al. (2014) based on the data obtained up to December 2005 (20-35 runs), which achieved 1-2 magnitude deeper limits than the single-pass SDSS images. Several other teams (e.g., Jiang et al. 2009; Huff et al. 2014) produced different stacks using different procedures to optimize the image qualities.

Jiang et al. (2014; hereafter J14) released a new version of stacks using only the images that were taken under

the best weather conditions. These stacks are $\sim 0.2$ mag deeper than those produced by Annis et al. (2014), reaching $5\sigma$ limits of 23.9, 25.1, 24.6, 24.1, 22.8 mag in $u'g'r'i'z'$, respectively, and also have better PSF characteristics. We adopt these stacks in our work.

## 2.2. *Structure of SDSS Stripe 82 files*

We use description from J14 to present the structure of optical data. An SDSS run (strip) consists of six parallel scanlines, identified by camera columns (Figure **??**). The scanlines are 13.5 arcmin wide, with gaps of roughly the same width, so two interleaving strips make a stripe that consists of total 12 scanlines (columns).

The size of each co-added SDSS image is 2854 x 2048 pixels, or roughly 18.8' x 13.5' (RA x Dec), with a pixel size of $0.396''$ and an average full width at half maximum (FWHM) of $\sim 1.5''$ in u-band, $\sim 1.3''$ in g-band, and $\sim 1''$ in r-, i-, and z-bands. In total there are 401 SDSS images in each column and overall $12 \cdot 401 \cdot 5 = 24,060$ SDSS images in all 5 bands. Each SDSS image has a corresponding *weight.fits* image, that records relative weights at individual pixels.

## 2.3. *WISE and unWISE*

WISE (Wright et al. 2010) is a near-to-mid IR space telescope launched in 2009 and has performed an all-sky imaging survey in four bands at 3.4, 4.6, 12, and 22 $\mu$m (denoted as W1, W2, W3, and W4, respectively). During its original mission phase from 2010 January 7 to 2010 August 6 (the "4-band Cryogenic" phase), WISE surveyed the entire sky 1.2 times in all four bands simultaneously until the solid hydrogen coolant in the outer cryogen tank was depleted. It then entered the "3-band Cryogenic" phase for the next 54 days, during which time it mapped an additional 30% of the sky in W1, W2 and W3. When the coolant in the inner tank was also depleted by 2010 September, only W1 and W2 are operational. The NEOWISE project took over the mission on 2010 October 1 and brought it into the four-month "Post-Cryo" phase to survey the sky in these two bands for near-earth objects until 2011 February 1 (see Mainzer et al. 2014). The telescope was then put into hibernation for the next 35 months as the funding stopped. The extended NEOWISE project reactivated it in 2013 December to continue the two-band observations ("NEOWISE Reactivation") through today.

The WISE team made three data releases separately for the 4-band Cryogenic, the 3-band Cryogenic and the NEOWISE Post-Cryo phases in 2012 March, 2012 June and 2013 May, respectively. To take the advantage of these repeated observations, the WISE team also made the "AllWISE Data Release" in 2013 November by combining all the WISE data available till then (see Cutri et al. 2013; for details). The included image products, known as the "Atlas Images" reach the nominal 5 $\sigma$ limits of 0.054, 0.071, 0.73, and 5.0 mJy in the four bands, respectively.

To optimize the detection of isolated sources, the WISE team has been using a special treatment when combining images, namely, the single-exposure images are convolved with the individual point spread function (PSF) during the stacking process. However, this operation has the drawback that it reduces the spatial resolution of the fi-

nal stacks, which is not desirable in many applications. To deal with this problem, Lang (2014; hereafter L14) reprocessed all the WISE images independently without the PSF convolutions, and produced the stacks that preserve the original WISE spatial resolutions. These image products of L14, dubbed as the "unWISE" images, have the PSF full-width at half maximum (FWHM) values of $6''$ in W1, W2 and W3 and $12''$ in W4. Sensitivity of bands W3 and W4 is $\sim$10 times worse than that of bands W1 and W2 and is not sufficient so that at least 10% of the optical sources have secure W3 and W4 fluxes. We use unWISE W1 and W2 images for this work.

## 2.4. *Structure of unWISE files - once again maybe I need to omit it in the paper*

The unWISE coadds are on the same tile centers as the WISE tiles with 18,240 images per band, 1.56 x 1.56 degrees each. The tiles are named by their RA, Dec center: tile "0591p530" is at RA = 59.1, Dec = +53.0 degrees; i.e., the first four digits of the tile name is $int(RA \cdot 10)$, then "p" for +Dec and "m" for -Dec, then three digits of $int(abs(Dec)\cdot 10)$. For each tile and band W1-w4, several images are produced, we shall list only the ones that we make use of:

– unwise-0000p000-W1-img-m.fits - "Masked" image, 2048 x 2048 pixels, TAN projected at 2.75"/pixel. Background-subtracted, in units of "Vega nanomaggies" per pixel: $mag = -2.5 \cdot (log_{10}(flux) - 9)$. This is the science image, the word "masked" means that some pixels have no unmasked pixels and no measurement at all: pixel value 0 and infinite uncertainty.

– unwise-0000p000-W1-std-m.fits - Sample standard deviation (scatter) of the individual-exposure pixels contributing to this coadd pixel.

Three unWISE images centered at the same RA cover the whole width of Stripe 82 in Dec ($-1.26^o$ to $+1.26^o$). We shall call three such unWISE images a frame. There may be up to 72 SDSS images within one frame.

## 2.5. *Spectroscopic sample*

Spectroscopic redshifts are used to perform star/galaxy separation and calibrate photometric redshift estimation: determine a set of templates and photometric offsets for the template-fitting code and train machine learning algorithms. We shall call such sample of sources a training set. Originally, we constructed our training set using spectroscopic data from the SDSS DR14 (Bolton et al. (2012)). Spectroscopic redshifts and classification into galaxies, stars and QSOs for the SDSS DR14 catalog are calculated using principal component analysis (PCA). The software `idlspec2d` is used to perform, at each potential redshift, a least-squares fit to each spectrum, using a fairly general set of models, for galaxies, for stars, for cataclysmic variables, and for quasars. The best fit model and redshift is chosen and assigned for the object.

Their spectroscopic sample consists of a wide variety of galaxies, stars and QSO with no cuts on color, although it is rather limited in terms of redshift (Strauss et al. (2002). Stripe 82 data are 1-2 mag deeper than single-pass images and therefore potentially has a lot of sources at higher redshifts. We decided to extend our training set by cross-matching galaxies in our photometric catalog with spectroscopic measurements of other, publicly

available surveys, such as 6dF, WiggleZ, DEEP2, VVDS and VIPERS.

In Tab.1 we list auxiliary catalogs with selection criteria, median redshifts for galaxies and QSO and total number of sources that were added to our training set. For each source, we use published redshift quality flag to select trustworthy sources, but also add stars when possible (which sometimes have special flag or assigned with negative redshift).

We cross-matched the galaxies from the training set with our photometric catalog using 2 asec as maximum separation in RA and Dec.

## 3. OVERVIEW OF METHODS FOR ANALYSIS

The most critical factor in SED fitting is consistent photometry in the involved bands, i.e., the photometry should include the same fraction of light across all bands so that the colors are defined in a consistent manner. This is challenging in our case because the spatial resolutions of WISE are at least $6\times$ worse than that of the SDSS. For this reason, the objects detected in WISE often suffer blending. Even for relatively isolated WISE sources, the photometric apertures appropriate for the (low resolution) WISE images cannot guarantee the same fraction of light being included as what is done in the (high resolution) SDSS images. Such a systematic offset, which is different for every galaxy, severely skews the SED fitting.

To best address this problem, we opt to use the T$-$PHOT software (**?**), which recently emerged as a robust and flexible tool to perform "template fitting". The basic idea is to use a high-resolution image (here an image from the SDSS) as the prior to build the morphological template of the source under question, convolve this template with the PSF of the low-resolution image (here the corresponding image from the unWISE), and fit this degraded template to the low-resolution image to obtain the total flux that is within the aperture as defined by the high-resolution image. In this way, we get reliable color information (i.e., flux ratio) in the most consistent manner.

While T$-$PHOT is much more user-friendly as compared to its predecessors (i.e. TFIT **?**), running this software is still non-trivial. It not only requires careful tuning of parameters but also several tedious preparatory steps with both the high- and the low-resolution images. Here we detail our procedures.

### 3.1. *Initial preparation of unWISE and SDSS images*

T-PHOT requires that the low- and the high-resolution images have the same orientation and the same World Coordinate System (WCS) reference position, the latter of which is defined by the FITS keywords (CRVAL1, CRVAL2). It also requires that their pixel scale ratio must be an integer. To meet these prerequisites, we carried out the following procedures utilizing the SWarp software (Bertin et al. 2002), which can subsample or bin an image to any pixel scale and then re-project to an arbitrary orientation at any tangential point.

We first rescaled the unWISE images from $2.750''$/pix to $2.772''$/pix. As the scale of an SDSS image is $0.396''$/pix, this makes the ratio of their pixel scales an integer ($2.772/0.396 = 7$). We kept the same orientation, which is always North-up and East-left, and the

same reference position for each unWISE image. This process was done for both the W1 and the W2 unWISE images, which are always aligned.

For a given SDSS image, we oriented it to North-up and East-left, and re-projected it at the tangential point as defined by the reference position of the unWISE images that it lies within. In other words, the FITS keywords (CRVAL1, CRVAL2) of the re-projected SDSS image are the same as those of the unWISE images. As an unWISE image covers much larger area and thus encompasses multiple SDSS footprints, the tangential projection point of a reprojected SDSS image is often outside of its coverage. In the extreme cases, it can be as far as $0.7°$ outside of the image itself.

About 30% of the SDSS images' footprints lie across two adjacent unWISE fields, and therefore need to be treated separately. If an SDSS image has more than 60 arcmin$^2$ belonging to adjacent unWISE fields, such image is duplicated, and each copy is reprocessed with respect to the appropriate unWISE field as described above. This increases the number of SDSS images from 4,812 to 5,556 per band. All SDSS images in col02 and col11 have $\sim$2.9' overlap with unWISE images centered at Dec=-1.7 and Dec=1.5 respectively. Processing of such small region requires construction of $\sim$930 PSFs per band and 2,000 hours of CPU time and is unviable. This excludes 11.34 deg$^2$ from the total area for which catalog is constructed.

We note that the above procedures were done for both the science images and the standard deviation (for the unWISE) or the weight (for the SDSS) images. After the subsampling and re-projection, the standard deviation or the weight value per pixel no longer preserves the absolute scale. In other words, the value of a given pixel on an unWISE (SDSS) reprojected standard deviation (weight) image no longer reflects the true standard deviation (weight) on that pixel. Fortunately, this does not affect the performance of T$-$PHOT, as it only uses these values in a relative sense (i.e., the absolute scale does not matter).

*** *describe unWISE cutouts for each SDSS image?* ***

We also note that one special treatment needs to be done for the saturated pixels in the unWISE standard deviation images. They are all assigned with zero values in the unWISE release, which is invalid for T$-$PHOT. We therefore use the IRAF/imcalc task to set such values to "9999" before reprocessing.

### 3.2. *Input SDSS source catalog*

J14 produced object catalogs from their stacked images using SExtractor (Bertin & Arnouts 1996). While it is tempting to use them directly as the input source catalogs for T$-$PHOT on the unWISE images, several caveats prevented us from taking this approach. For example, these catalogs are not cleaned of duplicated sources from the overlapped areas between adjacent images; they are not matched among the bands and thus have different number of sources for each band; the object detection threshold was set too high and many faint objects were excluded; bright and saturated objects have clusters of false detection; etc. For these reasons, we constructed our own input source catalogs.

**Table 1**
Spectroscopic training samples.

| Survey Name | Selection criteria | References | median redshift | Number of sources | Comments |
|---|---|---|---|---|---|
| SDSS DR14 | zErr < 0.1; zWarning = 0 | Bolton et al. (2012) | 0.339 | 233,123 | |
| 6dF | Q = 3, 4, 6 | Jones et al. (2004), Jones et al. (2009) | 0.054 | 123,135 | |
| WiggleZ | Qop = 4, 5 | Drinkwater et al. (2010) | 0.579 | 8,477 | no zErr provided |
| DEEP2 | ZQUALITY = 3, 4 and zErr < 0.01 or CLASS = STAR | c | 0.943 | 17,284 | |
| VVDS | 23.113231 | Le Fevre, 2013 | a | a | no zErr provided |
| VIPERS | 23.113231 | c | a | a | no zErr provided |

### 3.2.1. *Rationale*

In addition to providing input source information to T−PHOT, our SDSS catalog also provides optical SEDs for all the detected sources. For the latter, it is critical that the colors of a given source are measured consistently across all five bands, or in other words, the photometric aperture of a source must include the same fraction of total light in any given band. To achieve this goal, we took the standard approach by performing "PSF-matching" of the SDSS images. For a given SDSS field, we first matched the size of the PSF in g, r, i, and z-bands ($\sim 1.0$—$1.3''$) to that of u-band, which always has the largest PSF ($\sim 1.5''$; see J14). We then ran matched-aperture photometry using r-band as the reference band, which detects more sources than any other band.

The PSF-matching step was the most tedious part of the process, which required a lot of human intervention. To derive the convolving kernels between two images, we first must obtain the PSFs of both images. As the SDSS PSF varies from image to image, we had to construct it for each image individually. In total, we built 27,780 SDSS PSFs (5,556 per band).

### 3.2.2. *PSF construction and PSF-matching*

An empirical PSF is best derived by combining a large number of bright, isolated point sources (also known as the "PSF stars") distributed over the entire image, given that PSF is constant across the image. The most robust method to derive a PSF is to run the IRAF task "psf" interactively on a list of candidate PSF stars and to retain only the best ones in the construction. Given the huge number of images involved, however, this was not practical. After extensive tests, we settled on an approach that would result in reliable PSF stars, which would then allow us to run the "psf" task non-interactively in most cases.

The key in the PSF star selection is to determine whether an object is a point source of high image quality. We first removed the objects that are saturated or too faint. The magnitude cuts are typically $\sim xx$ and $\sim xx$ mag at the bright and the faint ends, respectively. We then used the stellarity identifier produced by SExtractor, "CLASS_STAR", to select the potential point sources. The output value of this parameter ranges from 0 (most likely non-stellar) to 1 (most likely stellar), and we rejected the objects that have CLASS_STAR < 0.85. The survived candidates were fur-

ther refined by comparing their "core" magnitudes and the total magnitudes. The smaller the difference between the two means the more compact the object is. The size of the aperture depends on the particular band, and is $\sim 3.56''$ in general. We adopted the criterion of $< 0.xx$ mag for this difference in this refinement. Finally, we only retained the stellar objects that are not close to the image edge ($> xx$ pixels to the boundary) and do not have a neighbor within $XX''$.

The retained point sources were then supplied to the IRAF/psf task for PSF construction. Typically, each SDSS image had more than 20 point sources in the end to build its PSF. Lower value appeared in only a few images, most of which are in the less sensitive u-band. On the other hand, there are plenty of images that have hundreds of qualified point sources, and using all of them would actually make the PSF construction exceedingly time-consuming. Therefore, we applied an upper limit of 160 point sources per image when running IRAF/psf.

*** Parameters, Moffat function ***

The output from this task was then convereted by using IRAF/seepsf to the format that can be further used. At this point, we visually examined all the resulted PSFs. If a PSF showed some undesirable features indicative of contaminations (for examples, elongation, gradient in the wing, neighbors not identified previously), we would manually reject the problematic point sources and re-build the PSF. Finally, all the PSF images were normalized to the total count of unity using wcstools/sumpix (Mink 1998) and IRAF/imarith tasks.

Before the PSF matching could be carried out, we would need the transferring kernal between the two PSFs under question. This was done by suppling the PSFs to IRAF/lucy that uses the algorithm developed by Richardson (1972) and Lucy (1974). Finally, the PSF-matchinng to the u-band was performed by IRAF/psfmatch that used the g'r'i'z' images and the relevant kernels as the inputs. For the sake of simplicity, we shall refer to these u-band PSF-matched images as the "g2u", "r2u", "i2u", and "z2u" images, respectively.

A subtle problem with PSF matching is that the convolution introduces extra correlation among nearby pixels, which artificially suppresses the noise. In other words, the noise of the PSF-matched images would be underestimated if it were measured by the pixel-to-pixel background fluctuation. We will return to this point in the next section.
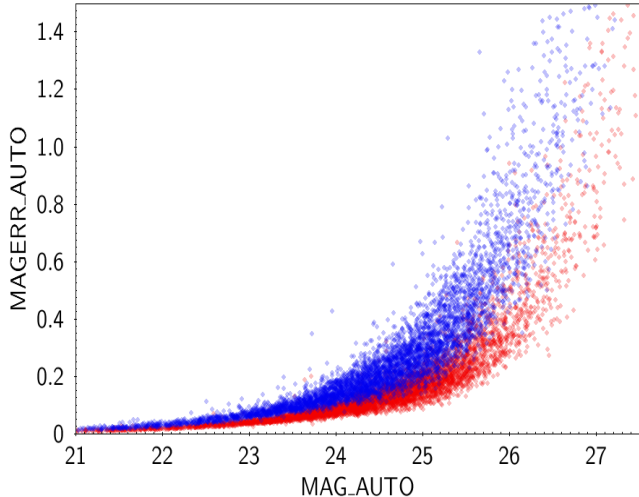
**Figure 1.** Magnitude error vs. absolute g-band magnitude for one SDSS image. Mmagnitude errors in the g2u image(red) are underestimated as compared to the original SDSS g-band magnitude errors (blue).

### 3.2.3. *Optical catalog construction*

The catalogs were produced by running SExtractor in the dual-image mode, where the r2u images were set as the detection images and the photometry was run in turn on the u, g2u, r2u, i2u, and z2u images.

*** You're missing a lot of details here!!! You should provide these:
  - What is the filter?
  - What is the detection threshold?
  - What is the number of connecting pixels (DE-TECT_MINAREA)?
  - What flavors of magnitude are derived?
***

The image weighting scheme was set to "MAP_WEIGHT", and we used the original weight maps from J14. Strictly speaking, these weight maps no longer reflect the pixel-to-pixel weights because the science images have been convolved. Nonetheless, this does not significantly impact the source detection because the general trend of the weight is still preserved (e.g., it is unlikely that the convolution would transform a low-weight pixel into a high-weight one).

However, as the correlation of pixels due to the PSF matching is not taken into account by these weight maps, the photometric errors reported in our `ISExtractor` runs are underestimated (except in u-band because the u-band images were not convolved). This is demonstrated in the left panel of Figure 1, where we show the MAGERR_AUTO vs. MAG_AUTO diagram for a randomly chosen g-band image. The blue and the red points represent the photometry using the original g-band image and the g2u image, respectively. The same weight image was used in both cases. Clearly, the red points are below the blue ones, or in other words, the photometric errors on the PSF-matched image are significantly underestimated.

### 3.3. *Kernels for T-PHOT*

Running `T-PHOT` requires the convolution kernel between the Hi-res and the Lo-res PSFs, which in our case are the PSFs of the SDSS "r matched to u" images and

the unWISE W1 and W2 images, respectively. While in theory the PSF of a "r matched to u" image should be the same as that of the u-band image, in reality the matching is never exact. Therefore, we could not use the u-band PSFs but had to construct the PSFs of the "r matched to u" images to optimize the performance of `T-PHOT`. We already had a list of point sources when constructing the original r-band PSFs (see §3.2.2), which could be used here again. We visually examined them on the "r matched to u" images, removed all those that are contaminated by neighbors due to the larger PSF in "r matched to u", and used the refined list of point sources to derive the "r matched to u" PSFs in the same way as described in §3.2.2.

For the unWISE images, we opted to derive their PSFs independently in all tiles instead of using the WISE PSFs as published by the WISE team. While the WISE PSFs were rather stable throughout the mission, they should not be used directly for our purpose because the image stacks are the combination of multiple single-exposure images of varying orientations. As the result, the PSF of a WISE image stack is no longer that of a single exposure, and also changes significantly from tile to tile.

There are 240 unWISE image tiles within the Stripe 82 footprint, and thus we derived 480 unWISE PSFs (240 each in W1 and W2, respectively) using point sources that were derived in the same way as described in §3.2.2. All PSFs were constructed by running `IRAF/psf` interacively.

All the unWISE PSFs were normalized to unity total counts using `IRAF/imarith` and sub-sampled in size by the factor of 7 using `IRAF/imlintran` to match the pixel scale ratio between the unWISE and the SDSS images (see §3.1). Finally, we used `IRAF/lucy` to generate the individual convolution kernels between the "r matched to u" images and the unWISE images. In total there are 11,112 kernels: 5,556 as a result of convolving "r matched to u" to W1 band and 5,556 - convolving "r matched to u" to W2 band.

### 3.4. *Running T-PHOT*

With the r2u source catalogs (§3.2.3) and the transformation kernels between the r2u and the unWISE images, we performed `T-PHOT` on each r2u/unWISE pair. This was by far the most computing-intensive step, which was carried out at the Lewis cluster of the High-Performance Computer (HPC) at the University of Missouri.

Following the recommendation of **?**, we ran `T-PHOT` for two passes. The first pass performs template fitting "globally" for the Hi-res sources on the Lo-res image at the exact source locations as provided by the input Hi-res source catalog. In the meantime, it also allows the fitting position of each source to shift by a small amount around the prior to find the most optimal position that would result in even better fit than at the prior position. The second pass then uses these refined positions to perform the template fitting again. In all our testing cases, the results from the second pass were significantly improved as compared to those from the first pass, which justified spending extra computing time to finish two passes.

As mentioned in §3.1, there are 5,556 SDSS images per band in the Stripe 82. Therefore, there were 5,556 r2u/W1 pairs and the same number of r2u/W2 pairs. Each T-PHOT pass took about 3 hours of comput-

ing time at the cluster, and the whole process totaled $\sim 66,700$ hours.

The output `T-PHOT` catalogs in W1 and W2 were merged with the SDSS optical catalogs for the future use. The sensitivity of WISE is not sufficient to detect all the SDSS sources. In total, 11,367,420 (39%) and 12,868,350 (44%) r2u sources are not detected in W1 and W2, respectively, and `T-PHOT` resulted in zero or negative fluxes for them. For the sake of completeness, these sources are still kept in the final catalogs and are assigned magnitude of -99.0 in W1 and/or W2.

***** These are numbers before masking, I am not sure what you mean by "cleaning" *****

### 3.5. *Catalog post-processing*

The merged photometric catalog from the above contains 29,046,660 entries, which would need to go through a few steps before it can be used further. In particular, our major goal is to produce a catalog of galaxies that is suitable for future applications. These steps are detailed below

#### 3.5.1. *Removing duplicate sources*

Adjacent SDSS images have 25" and 28" wide overlapping regions in RA and Dec, respectively. Similarly, adjacent unWISE images have 186" and 180" wide overlapping regions regions in RA and Dec, respectively. Sources in these regions appear in the catalog more than once and thus the duplications should be removed. Internal match on the catalog is performed and 1,423,979 duplicate sources with lower SNR in r-band are removed using 1.2 asec matching radius.

We performed an internal matching within the catalog, using a matching radius of $1.2''$. We found 1,423,979 duplications, which were subsequently removed from the catalog based on the SNR in the r-band (i.e. the source with the lower SNR is removed, while the source with the higher SNR is kept n the catalog).

#### 3.5.2. *Masking around bright stars*

Extremely bright objects severely contaminate their vicinities with various artifacts, such as bright halos, stripes of saturated pixels, diffraction spikes, ghost images, etc. As the sources found in such regions are often spurious and their photometry highly unreliable, we opted to remove them entirely from our catalog by masking out such regions. We used the SAO Bright Star Catalog (URL) and the Bright IR Stars Compilation (URL) as the basis of masking. There are around 400 (exact number) stars in both catalogs within Stripe 82, and our visual inspection of the T-PHOT residual images identified 300 (exact number?) additional objects (stars and very bright galaxies). We applied circular masks centered on these 705 sources, where the masking sizes depend on their V-band magnitudes if exist or our visual estimate of the halo sizes otherwise. The masking radius ranges from 50 to 600 arcsec. We removed the objects detected within the masked regions from our catalog. The maksed area totals 3.866 $\deg^2$, which reduces the total sky area of the survey to 288.212 $\deg^2$.

#### 3.5.3. *Star-Galaxy separation*

Our catalog includes both galaxies and Galactic stars, the latter of which are not the subjects of our study and therefore should be flagged. There are two major ways to separate stars and galaxies, namely, morphological classification and photometric separation. The former distinguishes point-like sources (mostly stars, but also including QSOs) from extended ones, while the latter uses some certain color criteria to separate stars (and QSOs) from galaxies in the color space. We chose to use the morphological method as the primary criterion and the follow up cleaning will be done by the gentle photometric separation using r-i vs. z-w1 color diagram. This was again to utilize the SExtractor stellarity parameter CLASS_STAR, which we used previously in selecting point sources for the SDSS PSF derivation (see §3.2). We ran `SExtractor` in the dual-image mode again, this time using the r2u images as the detection images and measuring the CLASS_STAR parameter on the original g', r' and i' images. An accurate derivation of CLASS_STAR needs an accurate input of the seeing FWHM value for each image, which we obtained based on the PSFs derived earlier. With the CLASS_STAR values, we then used the spectroscopic sample of 254,190 sources in Stripe 82 classified by the SDSS team (Bolton et al. (2012)) to determine the optimal range of CLASS_STAR. We found that it must simultaneously be less than 0.47, 0.43 and 0.75 in g', r' and i', respectively, to have the lowest fraction of missed galaxies (1.76%) and contamination from stars and QSO (2.39%).

The main goal of this project is to create a clean sample of galaxies. After morphological star/galaxy separation was done we determined a locus on the r-i vs z-w1 (?) plane that is mostly populated by spectroscopically confirmed stars that are still present in our sample. We used the color cut XXXXXX to perform the second order star-galaxy separation and it additionally removed XX% of the stars and XX% of galaxies from the spectroscopic sample. After performing star-galaxy separation our catalog contains 15,842,283 galaxies (old data).

–add figure with star/galaxy separation based on the r-i vs z-w1 colors

## 4. PHOTOMETRIC REDSHIFT ESTIMATION

A key goal of this work is to derive photometric redshifts ($z_{ph}$) for the galaxies in our sample. To this end, we tested a number of softwares widely used in the literature. There are two broad categories of $z_{ph}$ tools, one being based on SED fitting and the other being based on machine learning. In the SED fitting category, we tested HyperZ (citation), EAZY (citation) and LePhare (citation). In the machine learning category, we only used ANNZ (citation), which takes a neural network approach. We compiled a large list of spectroscopic redshifts ($z_{spec}$) in Stripe 82, and compared the performance of the different $z_{ph}$ tools by comparing the derived $z_{ph}$ and the availavle $z_{spec}$, quantified as $\Delta z = z_{ph} - z_{spec}$. The purpose of our test was to find the most appropriate tool(s) that would produce the smallest dispersion in terms of $\Delta z/(1 + z)$ and the least number of outliers.

### 4.1. *Running EAZY*

EAZY comes with several sets of SED templates, both theoretical and empirical, and they do not necessarily

perform equally well with different data. After extensive testing, we chose the CWW + KIN set (**??**) extended by Arnouts et al) and complemented it with Dusty model and Erb model (citation) because it performs the best according to the criteria outlined above.

### 4.1.1. *Redshift probability distribution*

SED fitting method suffers from the color-redshift degeneracy - redshift probability may have several peaks at significantly different values. One such example could be relatively featureless blue SED that can be fit well at z∼0 and z∼3 as templates can not distinguish between Lyman or Balmer breaks. Ideally one would want to include more bands to break this degeneracy, but as this is usually not possible, EAZY follows **?** and uses statistical approach to constrain photometric estimates. Bayesian prior is used to assign different probabilities to different redshifts for a given source with apparent magnitude $m$: low probability is assigned to low redshift due to smaller volume and high redshift due to the lower probability to find extremely bright galaxies. Ideally, redsift distribution should be determined from observed data, but observed samples are generally small and incomplete at high redshifts and for faint sources, so we use prior probability distribution derived with synthetic photometry of the semi-analytic model lightcone catalog described in §2.2 in Brammer et al. (2008) and implemented in EAZY.

### 4.1.2. *Magnitude zeropoint offset*

The largest advantage of the template fitting method is that derived photometric redshifts can be compared against the full spectroscopic sample. If any systematic offset is present, then it is possible to calibrate either the template library or the input magnitudes. As one of the output EAZY produces observed magnitudes - flux from the best fit template is convolved with the input filter response curve. In all bands we find the non-zero mean difference between input and fitted magnitudes for the sources whose photometric redshift was within $1\sigma$ from the spectroscopic one.

*I can add plot "mag_diff vs $z_{spec}$" or just a histogram of mag_diff*

While generally such offsets in different bands can be redshift- and/or magnitude-dependent, we choose to take it as a constant and apply it to the whole input catalog. The offsets of the AB mag XX, XX, XX, XX, XX, XX, XX for $u'g'r'i'z'W1W2$, respectively are supplied to EAZY and do not modify our original catalog.

Application of the magnitude offset reduces standard deviation of the subsample of the galaxies with spectroscopic redshifts from 0.XX to 0.XX (here and below we use the term "standard deviation" for the function $\Delta z/(1+z)$ for the subsample of the galaxies for which 0 < $z_{ph}$ and 0 < $z_{spec}$ < 1.5). (p32 of 1008.0395.pdf)

### 4.1.3. *EAZY with and without unWISE bands*

While it is generally accepted that near-IR bands YJHK improve photometric redshift estimation, especially beyond z∼1.3, where the 4000Å break goes out of the z band and the Lyman break is not yet detectable in the u band, using WISE bands may not be so efficient for two reasons: very few templates extend their models to 3-5$\mu$m and errors associated with the W1 and

W2 photometry are much larger than that of the optical bands, hence not giving strict constraints on the choice of the template. Nevertheless, our tests reveal that using W1 and W2 bands moderately improves photometry as compared to using only optical bands ($\sigma = XX$ and $\sigma = YY$, respectively).

As was mentioned above, only 39% and 44% of sources have W1 and W2 photometry, respectively, hence, there is some inconsistency in the input data for the photometric redshift estimation. In addition to that there is a population of sources with z-W1 > XX and W1 - W2 > XX for which photometric redshift is always underestimated. After visual inspection we confirm that though some of them are intrinsically very red sources, majority of them are associated with the bad flux subtraction made by T-PHOT. Those sources were assigned with the magnitude of -95.0 in W1 and W2. For these two reasons we present two photometric redshifts in the catalog - with and without unWISE W1 and W2 bands. Besides the number of the input filters, the rest of the parameters are identical:

TEMPLATES_FILE    CWW_KIN.spectra.param    # Template definition file
TEMPLATE_COMBOS    a    #    Template combination option
APPLY_IGM y # Apply Madau 1995 IGM absorption
N_MIN_COLORS 3 # Require N_MIN_COLORS to fit
APPLY_PRIOR y # Apply apparent magnitude prior
PRIOR_FILE    templates/prior_K_extend.dat    #    File containing prior grid
Z_MIN 0.01 # Minimum redshift
Z_MAX 1.0 # Maximum redshift
Z_STEP 0.01 # Redshift step size
Z_STEP_TYPE 1 # 0 = Z_STEP, 1 = Z_STEP*(1+z)
GET_ZP_OFFSETS y # Look for zphot.zeropoint file and compute zeropoint offsets

Derived photometric redshifts are published with the associated $1\sigma$ confidence intervals and redshift quality parameter "odds" that represents the fraction of the total probability that redshift lies within z±0.2 of the $z_{ph}$ estimate, and is designed to identify sources that have broad and/or multi-modal probability distributions.

The source is assigned the $z_{ph}$ = -99 when it has input data in less than 3 bands. Number of such sources is low - 2,946 and 11,101 (update after color-color cut) for the EAZY runs with and without unWISE data, respectively.

Comparison of the two sets of $z_{ph}$ shows similar behavior of the $\Delta z/(1+z)$ with marginally smaller scatter and more linear trend around z≈0 for the data with seven input bands as compared to the no unWISE data.

### 4.2. *Running ANNz*

ANNz requires a subsample of galaxies with known spectroscopic redshifts as a training set to build up an empirical relation between measurable parameters and redshifts and another subsample of galaxies to validate derived relation. Then it is applied to the rest of the sample to derive their photometric redshifts (Firth et al. 2003; Brodwin et al. 2006). Magnitude, surface brightness, color, size, and morphology (i.e. Sersic index) can be used as an input node for ANNz, but the most common values are magnitudes, colors or combination of both.

Photometric redshifts derived by ANNz are only reliable for the range of magnitudes or colors of the training set, so the input values should be chosen based on the rate of the overlap between the training set and the full input catalog. Our catalog is not magnitude limited and contains a large fraction of faint sources in all bands except for the r-band, which is used as a base of our catalog. Thus the training set inevitably covers only part of the plane taken by the full catalog, no matter what input values to choose. We randomly divided a sample with spectroscopic redshifts into three subsamples - to train ANNz, validate relations and test the output redshift and performed a set of tests to find the best input node. We tried all seven magnitudes, only optical magnitudes, seven non-redundant colors, combination of seven magnitudes and seven colors, seven magnitudes and three optical colors (except for the u-r) and seven colors with gri magnitudes. We found that seven magnitudes gives the best result and use it as the input node.

Architecture of the ANNz network can be described as $N_{input} : N_1 : N_2 : ... : N_m : N_{output}$, where $N_{input} = 7$ is the number of input magnitudes for each object, $N_1, N_2, ...,$ and $N_m$ are the number of nodes in each ANNz hidden layer for a total of $m$ layers, and $N_{output}$ is the number of sets for output photometric redshift, usually set to be 1. We again randomly divided the spectroscopic catalog into three subsamples and tested ANNz for the best combination of the nodes and hidden layers. We found that two nodes with ten layers in each gives the best performance of ANNz with our data. The final redshifts use 7:10:10:1 architecture, meaning seven input magnitudes, two hidden layers with ten nodes in each layer, and one output.

Quality of the photometric redhifts crucially depends on the training sets. While all spectroscopically confirmed galaxies are usually detected in all bands with sufficient SNR, it is faint sources that lack training sets. There are no sufficiently deep and wide spectroscopic surveys to date, so the redshift distribution of such faint sources is unknown and it will be greatly affected by the choice of the training set. We chose to present 2 flavors of the photometric redshifts - one derived using the training set with only sources with SDSS $z_{spec}$ (spec_SDSS_only) and the one derived using the training set with all available information: SDSS, 6dF, WiggleZ, DEEP2, VVDS and VIPERS (spec_ALL).

Depending on the particular seed for the random number generator, number of iterations and certainly, the fraction of the spectroscopic sample that is used for used for the weights, the training process usually converges to different redshift local minima. However, it is not necessary to select one best network - in fact, ANNz can use the sub-optimal networks to improve overall accuracy: the mean of the individual outputs from a group of networks usually gives a better estimate for the photometric redshift than the outputs of any one network alone. We create four networks spec_SDSS_only subsamples that contains of XXX,XXX sources in the following way: each time the subsample is randomly divided into two equal parts and ANNz is trained using 500 iterations, then the network is saved. The same approach is used for the spec_ALL subsample that contains YYY,YYY sources.

Usually one wants to estimate the accuracy of pho-

tometric redshifts using an independent subsample with spectroscopic redshifts, none of which is used for either training or validation. We follow recommendations of Firth et al. 2003 to increase the number of sources in the training sample and opted out to create such independent subsample.

*I rewrote till this phrase*

- training set with high-z galaxies (all available sources with $z_{spec}$) outliers were removed and ANNz was trained for the second times. (?) 4 weight files were supported

- comparing results - drastic difference in the redshift distribution, but very good agreement to the spec-z results

### 4.3. *Redshift distribution*

- We consider only sources with SNR larger 5 in bands gri to be trustworthy and in subsequent calculation we use only such objects.

- maximum redshift at which scatter is smalller than $3\sigma$ and also behavior of the sample around z∼0.4 and around z∼0

- shape of the distribution: where it peaks and why, comparison to the known redshift distributions from CANDELS and CFHTLS with appropriate magnitude cuts

- a lot of high-z sources are in the catalog (almost 80 percent of high-z sources from the spec-z catalogs are cross-macthed with our catalog), but our 7 bands are not sufficient to assign correct redshift to such sources using any method, so we discuss, how this can influence the GSMD estimation. Such high-z sources cannot be removed based on the color-color criteria.

- redshift bins: 20 Gpc$^3$ using standard cosmology; number of galaxies per bin; last 5th bin is not used in estimation of the GSMD, but to reject all high-z galaxies

- completeness of the catalog - as seen from the distribution, catalogs are incomplete after z=0.45. Thus Stripe 82 does not produce comlete catalog at significantly deeper redshifts as was anticipated, but still has much more data at higher redshifts.

### 4.4. *Combining results of photometric redshifts*

median - as there's no difference as described in Dahlen et al 2013 and also I can't compare odds (from ANNz) and chi2 (from EAZY).

### 4.5. *Outliers*

Check my notes p8. and section 4.8 of Dahlen et al 2013

Majority of outliers $(z_{ph} - z_{spec})/(1 + z_{spec})>3 \sigma$ are diiferent for different codes, 1250 of them are outliers in all 4 codes and also outliers in SCUSS, Reis, CFHTLS and SDSS DR15 (which are photo-z catalogs). 460 of them have zWarning=4 from SDSS DR14 (chi-squared of best fit is too close to that of second best), so it is not necessary that our photo-z are wrong. So it is possible that the real fraction of outliers is even smaller than reported here.

### 4.6. *testing against available photo-z catalogs*

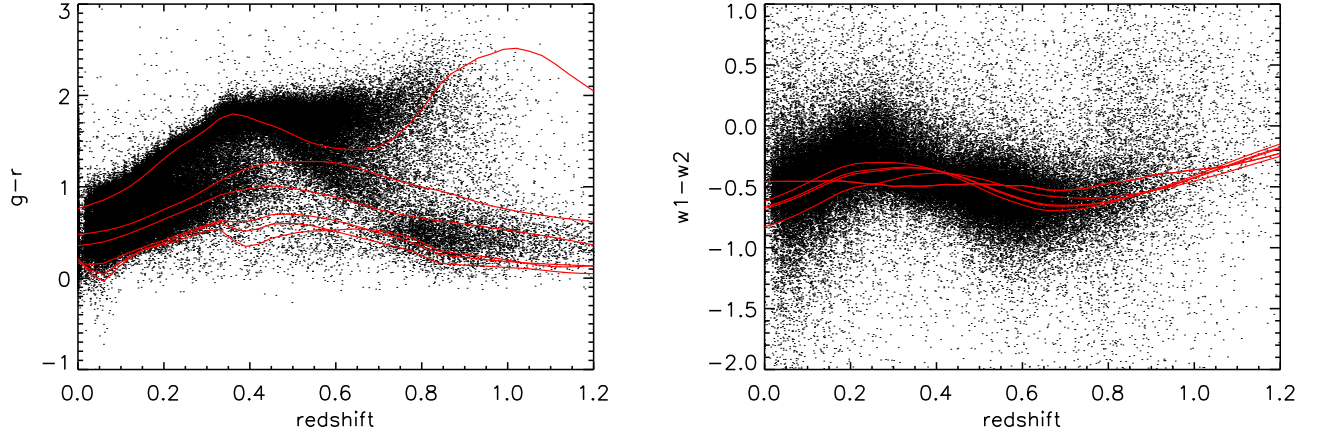SCUSS, SDSS DR15, Reis et al, CFHTLS

## 5. GALACTIC CATALOG

**Figure 2.** Color vs. spectroscopic redshift for the whole set of the photometric catalog sources (colored black) with available spectra. Colors of the six CWW+KIN templates as a function of redshift templates are plotted as red lines. These templates alone cover sufficient area on the plane and their linear combination returns reliable photometric redshift

That many stars based on CLASS_STAR in gri and color-color diagram. all have SNR > 5 in r-band.

## 6. RESULTS

## 7. DISCUSSION

## 8. SUMMARY

### REFERENCES

Adelman-McCarthy, J., M.A., A., S.S., A., et al. 2007, VizieR Online Data Catalog, 2276

Annis, J., Soares-Santos, M., Strauss, M. A., et al. 2014, ApJ, 794, 120

Arnouts, S., & Ilbert, O. 2011, LePHARE: Photometric Analysis for Redshift Estimate, Astrophysics Source Code Library, ascl:1108.009

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

Bertin, E., Mellier, Y., Radovich, M., et al. 2002, in Astronomical Society of the Pacific Conference Series, Vol. 281, Astronomical Data Analysis Software and Systems XI, ed. D. A. Bohlender, D. Durand, & T. H. Handley, 228

Bolton, A. S., Schlegel, D. J., Aubourg, É., et al. 2012, AJ, 144, 144

Bolzonella, M., Miralles, J.-M., & Pelló, R. 2011, Hyperz: Photometric Redshift Code, Astrophysics Source Code Library, ascl:1108.010

Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503

Cutri, R. M., Wright, E. L., Conrow, T., et al. 2013, Explanatory Supplement to the AllWISE Data Release Products, Tech. rep.

Huff, E. M., Hirata, C. M., Mandelbaum, R., et al. 2014, MNRAS, 440, 1296

Jiang, L., Fan, X., Bian, F., et al. 2009, AJ, 138, 305

—. 2014, ApJS, 213, 12

Jones, D. H., Saunders, W., Colless, M., et al. 2004, MNRAS, 355, 747

Jones, D. H., Read, M. A., Saunders, W., et al. 2009, MNRAS, 399, 683

Lang, D. 2014, AJ, 147, 108

Lucy, L. B. 1974, AJ, 79, 745

Mainzer, A., Bauer, J., Cutri, R., et al. 2014, in Lunar and Planetary Inst. Technical Report, Vol. 45, Lunar and Planetary Science Conference, 2724

Merlin, E., Bourne, N., Castellano, M., et al. 2016, Astronomy and Astrophysics, 595, A97

Mink, D. J. 1998, in Bulletin of the American Astronomical Society, Vol. 30, AAS/Division of Dynamical Astronomy Meeting, 1144

Pollo, A., Rybka, P., & Takeuchi, T. T. 2010, A&A, 514, A3

Richardson, W. H. 1972, Journal of the Optical Society of America (1917-1983), 62, 55

Staff, S. 1966, Smithsonian publ. 4652

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

Strauss, M. A., Weinberg, D. H., Lupton, R. H., et al. 2002, AJ, 124, 1810

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

# **APPENDIX**

APPENDIX

## ACKNOWLEDGMENTS