Construction of C3VO COSMOS Master Spec-z Catalog

Ben Forrest

1 Catalog Entries

The final catalog contains 42701 objects, ordered by the ID of the COSMOS2020 best match. There are 30 entries for each galaxy, with -99 indicating a lack of information. In order:
0) COSMOS2020 matched object ID (C20),
1) Spectroscopic catalog object ID,
2) Spectroscopic survey,
3) Right ascension from the spectroscopic survey with the first-order astrometric correction,
4) C20 right ascension,
5) Declination from the spectroscopic survey with the first-order astrometric correction,
6) C20 declination,
7) C20 K-band AB magnitude (3" aperture corrected to total),
8) Difference between C20 and spectroscopic catalog K-band AB magnitude,
9) C20 i-band AB magnitude (3" aperture corrected to total),
10) Difference between C20 and spectroscopic catalog i-band AB magnitude,
11) Spectroscopic redshift,
12) C20 photometric redshift (EAZY or LePhare closest to $z_{\rm spec}$),
13) Spectroscopic quality flag in the style of VUDS,
14) Matching flag,
15-17) Median, 16th percentile, 84th percentile of stellar mass from LePhare,
18-20) Median, 16th percentile, 84th percentile of star formation rate from LePhare,
21-23) Median, 16th percentile, 84th percentile of age from LePhare,
24-28) Absolute NUV, u, r, V, and J magnitudes from LePhare,
29) χ^2 (not reduced) of best-fit model from LePhare,
30) Best-fit model index from LePhare,
31) E(B-V) of best-fit model from LePhare.

2 Versioning

v2	2022.01.31	VUDS now prioritized over DEIMOS10k. Only the new version of the zCOSMOS-DEEP catalog is now used. Issue with pulling in LePhare results fixed.
v1	2022.12.15	Corrected total magnitudes now used.
v0	2022.12.14	Initial version.

3 Construction Overview

The goal of this project was to compile a catalog of spectroscopically targeted objects in the COSMOS field from a number of surveys and match them to the COSMOS2020 photometric catalog (Weaver+2022). Here we use the CLASSIC_R1_v2.0 catalog. In order to do this, galaxies were sourced from the following catalogs.

3.1 Initial catalog

Brian Lemaux had constructed a similar catalog previously, matching objects to the COSMOS2015 catalog (Laigle+2016). This consisted mainly of spectra from VUDS and zCOSMOS. However, C3VO DEIMOS spectra of the $z = 4.57$ protostructure, as well as small numbers of galaxies from other sources were included too.

3.2 zCOSMOS-Deep

While the above catalog included spectroscopic data from both zCOSMOS-Bright and zCOSMOS-Deep, this catalog incorporates an updated zCOSMOS-Deep catalog from Kashino Daichi (proprietary). The primary change compared to the previous version is the quality flags of galaxies, with a general trend of decreasing confidence, though this is not universal. There is at least one case where the id's of two objects appear to have been switched. Regardless of any differences, the new values for a given ID were adopted.

3.3 DEIMOS-10k

The published catalog from the DEIMOS-10k survey (Hasinger+2018) is included.

3.4 C3VO

Spectra from the C3VO survey include two catalogs, one from DEIMOS spectra (Ekta Shah) and one from MOSFIRE spectra (Ben Forrest).

3.5 MAGAZ3NE

Not yet included.

Table 1: Positional offsets (arcseconds) for targets in each survey to the closest object in COSMOS2020.

Survey	$N_{\rm cat}$	Median($\Delta\alpha_{i,j0}$)	Median($\Delta\delta_{i,j0}$)
zCOSMOS	28628	0.08091	0.01123
DEIMOS10k	10770	0.01928	0.04672
VUDS	4613	0.00386	0.07289
C3VO	1204	0.08475	0.02163

Table 2: Distance (arcseconds) from corrected spec. coordinates to nearest phot. object.

Survey	Median($r_{\rm match}$)	$r_{3\sigma}$	$f_{r<+3\sigma}$
zCOSMOS	0.071	0.492	0.950
DEIMOS10k	0.132	1.346	0.927
VUDS	0.097	0.701	0.910
C3VO	0.071	0.512	0.928

4 Initial Object Matching

4.1 Astrometric Correction

For each spectroscopic survey, each spectroscopic entry (s_i) was matched to the nearest photometric catalog member (p_{j0}). The coordinates for spectroscopic entries were then updated based on the median positional offsets between the data sets, median($\Delta\alpha_{i,j0}$) and median($\Delta\delta_{i,j0}$), to account for different astrometric frames in the catalogs (Table 1).

4.2 Positional Threshold

Catalog matching was performed again to find the nearest photometric catalog member (p_j) using the updated spectroscopic coordinates. We then used the distribution of distances between matches to calculate our matching tolerance, considering that this distribution is composed of a combination of correct matches and random nearest matches in cases where the spectroscopic target is not in the photometric catalog. These two components become clear by analyzing the histogram of logarithmic distance separations.

We fit a gaussian to each of these distributions. The 3σ upper limit for the main peak (assumed to be correct matches) is then taken to be the distance threshold within which to search for a given spectroscopic survey (Figure 1, Table 2). Alternatively taking the 3σ limit from fitting a single gaussian to the entire distribution or by choosing the distance at which the contribution of the two gaussians are equal, while changing the number of galaxies matched, does not result in significant differences.

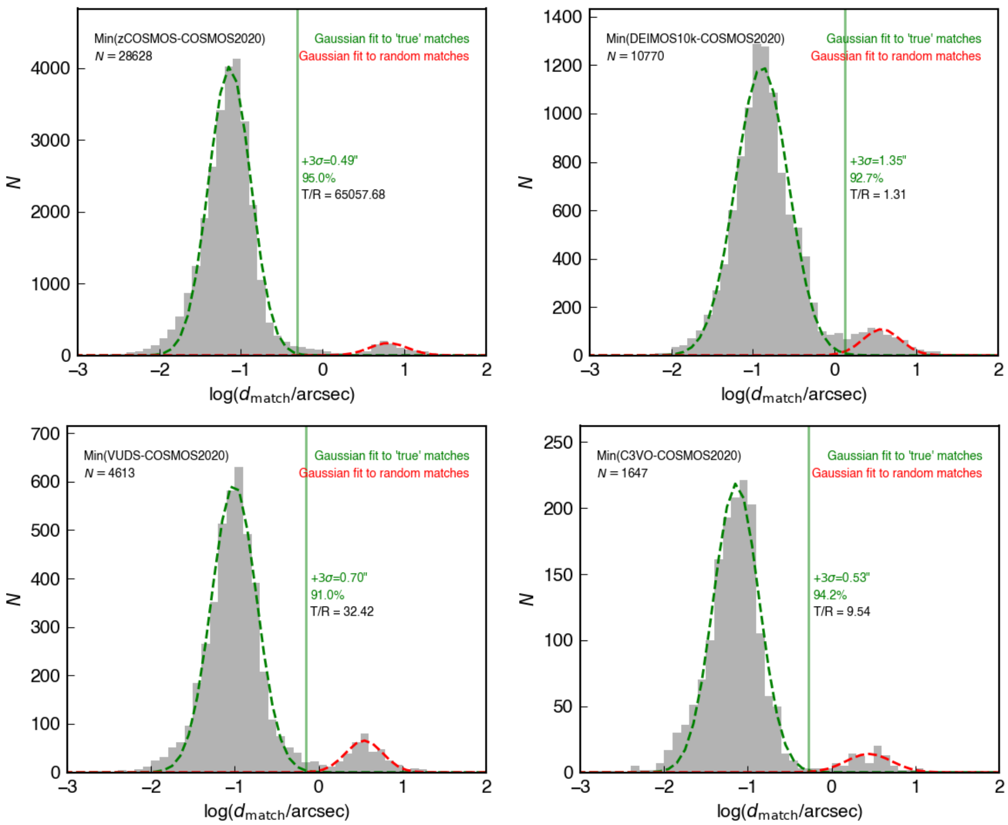

Figure 1: The distribution of positional offsets between spectroscopic entries and the nearest COSMOS2020 object. The bimodal distribution is fit by two gaussians, assuming a component of 'true', accurate matches (green) and a component of 'random', incorrect matches (red). The $+3\sigma$ value of the 'true' match fit is considered to be the reliable matching radius for the survey. The percentage of spectroscopic entries with matches within that radius and the ratio of the 'true' to 'random' galaxies at that radius is also given.

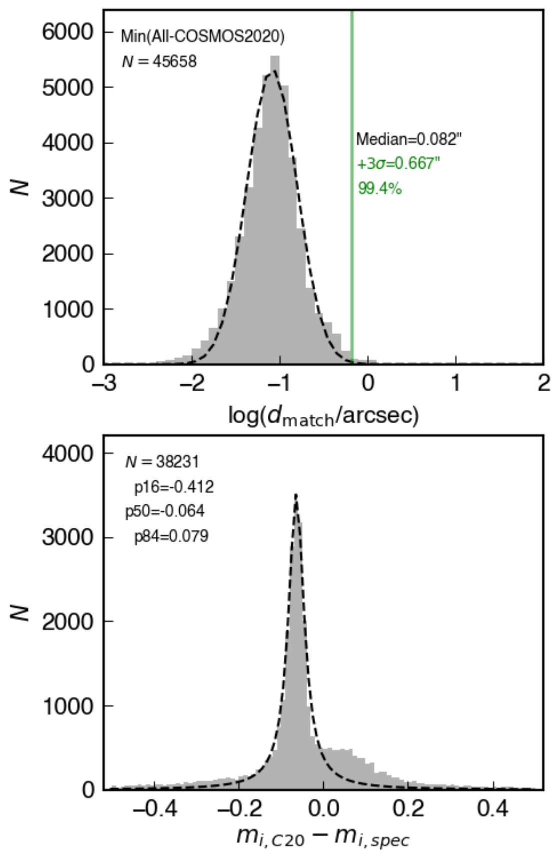

Figure 2: The difference between galaxy properties in the spectroscopic and photometric catalogs. The fits to the position and redshift distributions are gaussians, while the fits to the magnitude differences are Lorentzian in form.

Table 3: Distribution of number of matches to each spec. target.

Survey	$N_{\rm cat}$	$N_{\rm 1match}$	$N_{>1\rm match}$	$N_{\rm nomatch}$
zCOSMOS	28628	27156	33	1439
DEIMOS10k	10770	9720	421	629
VUDS	4613	4182	15	416
C3VO	1647	1087	464	96
Total	45658	42145	933	2580
Percentage	100	92.31	2.04	5.65

5 Resolving Multiple Matches

Between all the spectroscopic catalogs considered, there are 45658 entries. Of these, 42145 (92.31%) have a single match within the matching radius considered (matching flag = SINGLEMATCH). Meanwhile, 933 (2.04%) have multiple such matches, while 2580 (5.65%) have none (matching flag = NOMATCHES; see Table 3 for a breakdown by survey). We note that our matching radius is somewhat conservative and that relaxing this cut would result in more matches (and also more cases of multiple matches).

In the case of multiple matches, we compare the photometric and spectroscopic redshifts, and both the i-band and K-band AB magnitudes reported in the spectroscopic and photometric catalogs in addition to the positional offset (see Figure 2). For the COSMOS2020 catalog, we look at both the EAZY (ez_z_phot) and LePhare (lp_zPDF) photometric redshifts and choose the one that best matches the spectroscopic redshift of the target. For magnitudes, we convert the 3" aperture magnitudes to total magnitudes by correcting the aperture size and accounting for extinction, as done in the code released with the catalogs. For those spectroscopic catalogs which source their photometry from the COSMOS2015 catalog (Laigle+2016) we similarly correct to the total aperture magnitude.

$\sim 2\%$ of the compiled spectroscopic entries have more than one photometric catalog member which satisfies the distance cut above. For each potential photometric match to a spectroscopic target, we analyze the four parameters above to select the best match. Clearly if one photometric object best matches the spectroscopic target in all four parameters, this is deemed the best match and the matching flag is set to MULTIMATCH-Q4.

In the event that there is a clear correct match, the matching flag is set to MULTIMATCH-Q3.#-XX(X). Here, # is equivalent to $3 \times N_{\rm Agree} - N_{\rm Non-entry} - 2 \times N_{\rm Disagree}$, and XX(X) denotes the parameters that best agree between the spectroscopic target and the best match, with P=position, Z=redshift, K=K-band magnitude, and I=i-band magnitude.

So for instance, a flag of MULTIMATCH-Q3.8-PIK means that one photometric object best matches the spectroscopic target in position, i-band magnitude, and K-band magnitude and that the spectroscopic target does not have a reported spectroscopic redshift for comparison. Similarly, a flag of MULTIMATCH-Q3.7-ZIK means that one photometric object best matches the spectroscopic target in redshift, i-band magnitude and K-band magnitude. While a different photometric object is a better match to the galaxy position, both photometric objects are close enough to the spectroscopic coordinates to be considered potentially the same object.

There are also objects with MULTIMATCH-Q2.2-XX and MULTIMATCH-Q1. The former are less secure matches, with the position and one other parameter of the selected best match photometric object and spectroscopic target agreeing and the other parameters either being a better match to a different photometric object or not having entries. The latter indicates that redshift, i-band magnitude and K-band magnitude were not available for comparison, and in this case the closest photometric object (in position) was taken as the correct match.

6 Removing Duplicates

At this point, 45685 spectroscopic entries have been matched to the correct photometric catalog object, while 2580 (5.65%) are considered to not have a match. Due to natural overlap between different spectroscopic surveys, photometric catalog objects with multiple spectroscopic matches exist.

If only one entry has a quality flag of 3, 4, or 9, we keep that entry and remove the other(s). In the event that multiple spectroscopic entries for a photometric object have these high-quality flags, or in the event that none of the multiple spectroscopic entries for a photometric object have high-quality flags, we choose the entry to keep based on survey (and effectively instrument) priority. In order of preference, these are C3VO (MOSFIRE), MAGAZ3NE, C3VO (DEIMOS), VUDS, DEIMOS-10k, zCOSMOS. Eliminating these duplicates results in 42701 entries in the catalog of which 34999 have a spectroscopic redshift.

7 Fitting with LePhare

There are 40121 entries in the catalog matched to a COSMOS2020 object, of which 32958 (82.1%) have a measured redshift $0 < z < 9$. For all such objects, we run LePhare with the redshift fixed to the spectroscopic redshift and using the COSMOS2020 photometry corrected to total as described above to obtain absolute magnitudes and physical parameters.

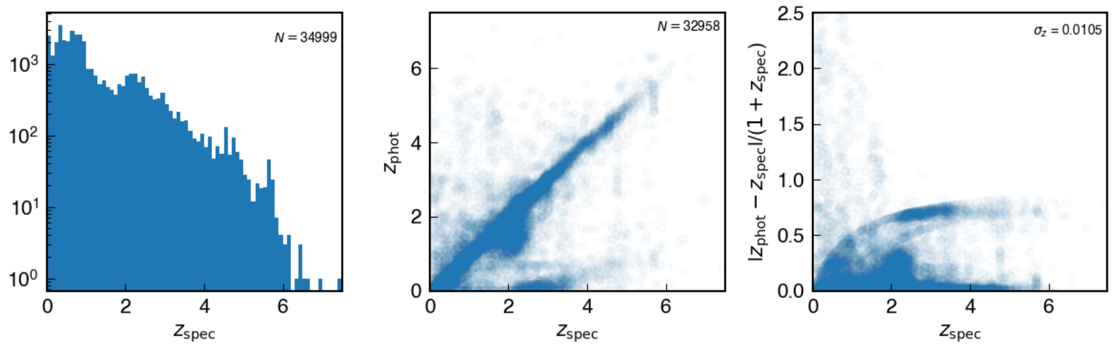

Figure 3: Redshift comparison from the final catalog. There are 34999 measured spectroscopic redshifts, with 32958 matched to a COSMOS2020 object. For these, $\sigma_z = 0.0105$, showing excellent photometric redshift accuracy.