

Photometric Redshift Survey

Physics Department, University of California, Santa Barbara, CA 93106

(Dated: June 13, 2023)

We compare the effectiveness of using two different methods for finding the photometric redshift, z_{phot} , of 6 galaxies. One method utilizes the public domain code HyperZ, while the other utilizes Nearest Neighbor machine learning. Of the six galaxies, images and data of four were obtained using the LCO telescope system, while the other two were obtained from SIMBAD astronomical database for testing the methods. One galaxy, 2dFGRS TGN132Z253 (referred to as galaxy D), has data both from SIMBAD and from experimental data. For six galaxies, our resulting percent differences from the accepted redshift for each using HyperZ for were 12900%, 353%, 801%, 274%, 42.5%, 6038%, and 18500%. The percent differences using the Nearest Neighbor method were 9430%, 439%, 781%, 174%, 10.6%, 43.2% and 85.9%. These results show the difficulty in capturing photometric redshift using a groundbased telescope but also indicate great possibility for photometric redshift, overall.

I. INTRODUCTION

Photometric redshift is the technique of using the differences in the properties of a stellar object or galaxy, called a source throughout, when viewed through different filters to estimate the source's redshift. This is useful, as it can be done for many sources in one image, reducing the amount of telescope time needed. It also allows for estimates to be made before the use of a spectroscope is needed. To understand what these terms mean and the impact of this technique, we need to understand the theory.

A. Redshift and Hubble's Law

Edwin Hubble discovered that the universe is expanding, as he observed that all galaxies seem to be moving away from us and each other. The speed at which they move away from us seemed to dependent on distance, as shown in Hubble's Law,

$$v = H_0 D, \quad (1)$$

Where v is the recessional speed, H_0 is the Hubble constant, and D is the distance between us and the galaxy (also referred to as the proper distance). H_0 is currently agreed (under the Benchmark Model) to have a value of 68 ± 2 $\text{kms}^{-1}\text{Mpc}^{-1}$. (Ryden (2017))

As the galaxies move away from us (and we move away from them), the light waves emitted by these galaxies become longer due to the increased distance. This effect is similar to the Doppler effect, where soundwaves produced by moving sources have a lower frequency. The relationship between the original unshifted wavelength and the stretched wavelength is called redshift, and is given by the formula,

$$z = \frac{\lambda_{obs} - \lambda_{em}}{\lambda_{em}} \quad (2)$$

Where λ_{obs} is the observed wavelength and λ_{em} is the emitted, unshifted wavelength. Hubble's law can be reformatted to be in terms of redshift, if we view it as similar to a Doppler shift (Ryden (2017)),

$$z = \frac{H_0}{c} r \quad (3)$$

Where r is the proper distance to the galaxy and c is the speed of light. This is usually what is referred to as Hubble's Law (or the Hubble-Lemaitre Law), since it is much more useful overall.

Since redshift is a phenomena of light, we use spectroscopy (the study of spectrums of light) to analyze the light from a galaxy and find its redshift.

B. Absorption Spectra and Spectroscopy

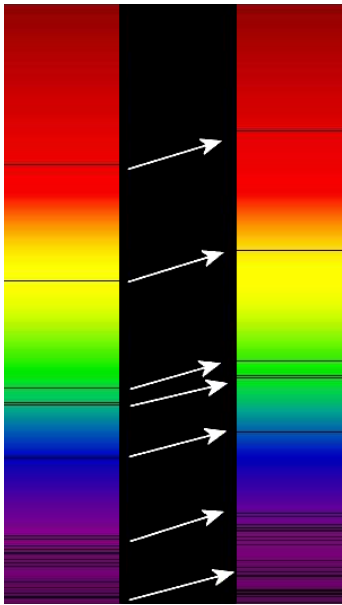


FIG. 1: This illustration (Stokes, accessed June 2023) shows the redshift of spectral features for a generic cluster of galaxies.

and helium produced in the atmosphere create absorption spectra. Thus, one way to view the redshift of a star is to take a spectrograph of the star's light, look for absorption lines, and find how much they have been shifted towards the redder end of the spectrum (Lena et al. (2008)).

Absorption spectra are usually only visible using spectrometers, which are designed to scan through wavelengths and graph them. In the occasion that a telescope lacks spectrometers, we can rely on photometry instead.

C. Vega Magnitude versus Asinh Magnitude

Traditionally, the magnitude at wavelength λ_0 is defined as

$$m_{\lambda_0} = -2.5 \log \left(\frac{e(\lambda_0)}{e_0} \right) = -2.5 \log e(\lambda_0) + q_0 \quad (4)$$

When an electron gains or loses energy, it produces a photon of a specific wavelength depending on the transition. Thus, when the spectrum of light from a heated gas is viewed using a spectrometer, there will be intensity peaks corresponding to the light with these frequencies. These are called emission lines, as they usually form vertical lines on a spectrograph. In addition, if light passes through a gas that usually produces a specific emission spectra, there will be dark lines where photons with the corresponding emission wavelengths are absorbed by the electrons, creating absorption lines. See Fig. 1 for an example of a light spectra with absorption lines.

In stars, when the light from a star passes through the star's outer atmosphere, the hydrogen, nitrogen

Where $e(\lambda)$ is the monochromatic flux from the star. e_0 is the flux from a reference object and is absorbed into the constant, q_0 , which defines an object of magnitude zero. Since astronomers usually calibrate around the star Vega, this is generally referred to as the Vega magnitude system. Telescope filters all take in light flux from stars based on their sensitivity functions, $t_0(\lambda)$. Thus, the apparent magnitude of a star through a filter is given by (Lena et al. (2008), Eq. 106),

$$m_{\lambda_0} = -2.5 \log \int_0^\infty t_0(\lambda) e(\lambda) d\lambda + 2.5 \log \int_0^\infty t_0(\lambda) d\lambda + q_{\lambda_0} \quad (5)$$

The transmission functions for the different filters used by LCO telescopes can be easily found on their website (Las Cumbres Observatory, accessed June 2023) and are shown in this paper in Fig. 3.

Shortly after the creation of the SDSS filters, many researchers noted that were differences between the expected and measured magnitudes viewed through SDSS filters for different locations, as detailed in the Sloan Digital Sky Survey Early Data Release (Stoughton et al. (2002)). This required the invention of another magnitude system which would reduce inaccuracies between the different systems that used SDSS filters (namely the USNO 40 in (1m) Telescope, the SDSS 2.5m Telescope, and the SDSS PT system). Now, while most systems use the Vega magnitude system detailed above, many databases use the "asinh magnitude" system as a way to make the magnitudes of fainter objects more accurate. It was first derived and documented by Lupton et al., in Sept. 1999, in their paper, "A Modified Magnitude System that Produces Well-Behaved Magnitudes, Colors, and Errors Even for Low Signal-to-Noise Ratio Measurements." Instead of the log based Vega system, it's based around the asinh function. Using the flux of an object, f , the flux of a 0 magnitude object seen through the filter, f_0 , and the softening parameter, b , the asinh magnitude is given by,

$$m = -\frac{2.5}{\ln(10)} \left[\operatorname{asinh} \left(\frac{(f/f_0)}{(2b)} \right) + \ln(b) \right] \quad (6)$$

The softening parameter depends on the filter and was found by setting the parameter to be equal to 1 σ of the sky noise as part of the calculations for the Sloan Digital Sky Survey, as shown in Table I. Note the values for $m(f/f_0 = 10b)$, as the asinh and log magnitudes above these values differ by less than 1%.

Now that we have the magnitude, we can start to look for the redshift using the difference in the magnitudes of the same object viewed through different filters. This is known as the color index. This is simply $m_{\lambda_1} - m_{\lambda_2}$ using the magnitudes given by Eq. 6. Since this correlates to the temperature and size of a star, most tables are organized using the stars classification, ranging from hottest (O) to coolest (M), the color index can be seen in Table II.

Band	b	Zero-Flux Magnitude $[m(f/f_0 = 0)]$	$m(f/f_0 = 10b)$
u	1.4×10^{-10}	24.63	22.12
g	0.9×10^{-10}	25.11	22.60
r	1.2×10^{-10}	24.80	22.29
i	1.8×10^{-10}	24.36	21.85
z	7.4×10^{-10}	22.83	20.32

TABLE I: This table, from the Sloan Digital Sky Survey: Early Data Release (Stoughton et al. (2002)), shows the softening parameter, b , for each filter band, as well as the Zero-Flux magnitude, and the $f = f_0 10b$ magnitude value.

Class	B-V	U-B	V-R	R-I	T (K)
O	-0.33	-1.19	-0.15	-0.32	42,000
B	-0.30	-1.08	-0.13	-0.29	30,000
A	-0.02	-0.02	0.02	-0.02	9,790
F	0.30	0.03	0.30	0.17	7,300
G	0.58	0.06	0.50	0.31	5,940
K	0.81	0.45	0.64	0.42	5,150
M	1.40	1.22	1.28	0.91	3,840

TABLE II: This table, from Zombeck (1990), shows the star class, the color indices for each star between two filters, and the temperatures of the stars they correspond to.

Using the magnitudes and color indices of our galaxies, as well as their spectral features, we can use two methods to find the redshift of a source. This is referred to as photometric redshift. In this paper, we use two methods that utilize open source codes. We're referring to the two methods as the HyperZ method and the Nearest Neighbor Method. These are discussed below.

II. MATERIALS AND METHODS

A. Using the Las Cumbres Observatory Telescope System

To take an image of a target using the Las Cumbres Observatory (LCO) Telescope System, it requires the target identifier, the coordinates of the target, filter chosen, the exposure time, how many exposures and the time the image will be taken. All images taken during this paper utilized LCO's 0.4 m SBIG Telescope, as well as the u'g'r'i' filters. The required variables are straightforward, but we include them for convenience:

The coordinates are given in Right Ascension and Declination, also known as the equatorial coordinate system. All coordinates we used are under the 2000 epoch. The filter is the lens on the telescope and determines what wavelengths of light will be captured from the galaxy. See the SDSS filters section for more information about the filters used for this paper. The exposure time determines how long the aperture will be open. The time must be chosen using LCO's visibility tool to make sure the target will be visible to the telescopes during the chosen time period. Finally, the number of exposures determines how many times the same image is taken using the above exposure time and filter.

To obtain a full color picture of a galaxy, one must use multiple filters that cover the visible wavelength. These can then be used either for photometry, as we did in this project, or for astrophotography, amongst many other applications.

B. BANZAI

The images that come from the LCO telescopes must first undergo a cleaning up process to make the images more readable, remove noise, etc. This is referred to as the BANZAI pipeline (an abbreviation of "Beautiful Algorithms to Normalize Zillions of Astronomical Images"). It is coded by LCO scientists and stored in a github on their website. All images used in this report were processed using the BANZAI program, created by McCully et al. (2018). The processes are as follow:

Bad-pixel masking compensates for dead pixels or other systematic errors in the telescope's sensor. As telescopes are used over time, specific sensors in the camera can be damaged. These could be caused by overexposure to light, electrical problems, or by neutrinos or electrons interfering with the device. These create faulty pixels on the final image, as well as noise. Bad-pixel masking compensates for any noise or errors created by these. In the BANZAI pipeline, it uses a technique called sigma clipping.

In essence, sigma clipping assumes a gaussian distribution of the noise, finds the sigma value, termed the robust

standard deviation (rstd), using the median absolute deviation (since the mean can be affected heavily by outliers, the median is more appropriate here) of the distribution of all pixels' flux values in the image,

$$\sigma \approx 1.4826 \times \text{mad} \quad (7)$$

and masks any pixels that fall outside of 3 rstd of the median flux value. This process is detailed in the latest BANZAI release, McCully et al. (2019).

Moving on, all cameras, telescopes, and other optical devices have readout noise. This is caused by either imperfect transfer between the sensor and the computer, or by imperfect capture by the sensor. This can cause the images to have light or dark spots, static, and problems around the edges of the image. To avoid this, one can take a bias image (or many bias images and stack them), usually an image where the telescope is covered, and subtract it from the final image. This is called bias subtraction. Similar to bias subtraction is dark subtraction.

Where bias subtraction covers readout noise from capturing the image, dark subtraction helps deal with heat generated over long exposures that can cause the image to have light spots. These light spots are caused by the sensors registering the heat from the instrument as a photon activating the silicon sensors. To correct for this, long exposure images are taken with similar time to the bias images, and subtracted from the final product.

LCO takes bias and dark images every morning and night of various exposure times to ensure that they can obtain the most accurate images post BANZAI. They average the morning and night images to make a master bias that they subtract from their processed images. They choose bias and dark images that are appropriate for the exposure times of the images being put through the pipeline (See McCully et al. (2019) for more detail on this process).

In addition, they perform flat field correction, which corrects flaws in the image caused by the lens' curvature. This curvature can create an effect of darkening around the edges of the image. This can also include correcting specific flaws from the lens (such as dust, scratches, etc.). LCO takes flat field images each day at twilight to ensure accurate reduction. Since twilight is not long enough to take photos using all of the filters, they alternate every day.

Finally, one of the most complex processes in the BANZAI pipeline is source extraction. SEP, the "Python and C library for Source Extraction and Photometry", is a python library derived from a command line program named SExtractor, or Source Extractor. The command line program was published in Bertin and Arnouts (1996), while the Python library was adapted much later in Barbary (2016) for ease of use. SEP performs many functions, but the main function is source extraction, finding the stars or galaxies in the image and coding their coordinates (in pixels) in the image. It also finds fluxes of these sources using kron apertures (See below in the Magnitude section for more about this). The BANZAI pipeline processes the information from the telescopes using SEP and produces a FITS file, which

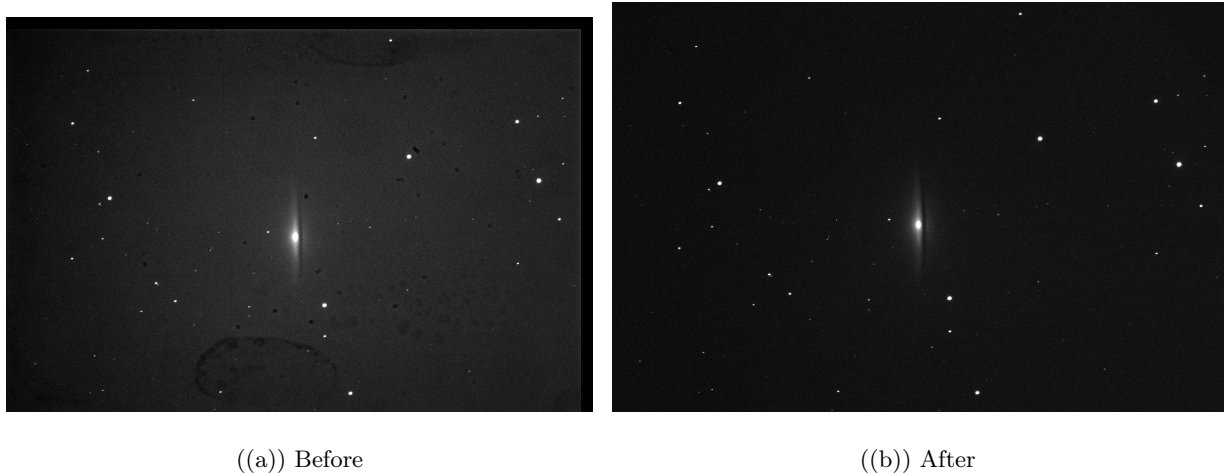


FIG. 2: These figures show an image, taken using the SDSS g' filter, of the Sombrero Galaxy, galaxy A in Table ??, before and after the BANZAI process.

contains the image, as well as a "header" – a text file that contains technical information useful for astronomers, such as the fluxes of each object in the image, the filter that was used to take the image, the coordinates of the sources in the image and many other useful properties.

While the pixel coordinates of the sources are useful on their own, having the Right Ascension and Declination attached to each source provides much more ease of reference for the astronomer. Astrometric calibration adds the equatorial coordinates to the header of the fits file so that the coordinates of each point on the image are accurate and the sources can be matched to a database. The LCO telescope system uses *astrometry.net* to calibrate the coordinates of each image after they are taken (See Lang et al. (2010) for more detail about how *astrometry.net* functions).

C. SDSS Filters, What are they?

The Sloan Digital Sky Survey was a project started in the year 2000 to record a large portion of the sky in detail. For the project, they invented SDSS $u'g'r'i'z'$ photometric system of filters, where each filter covers a specific part of the light spectrum and blocks any light outside that spectrum. Their creation, as well as plans for the survey, are detailed in Fukugita et al. (1996). These filters are still used today throughout astronomy due to their well documented transmission functions. Using these filters in combination, one can obtain nearly the full spectrum of a galaxy's light from ultraviolet (u') to near infrared (z'). The transmission functions for the SDSS filters used on the Las Cumbres Observatory's (LCO) 0.4m telescope, which determine what light passes through the filter for each filter, are shown in Fig. 3. These filters (except z' , which is not equipped on the 0.4m) were used throughout this

paper for any experimental data, as detailed in Table IV.

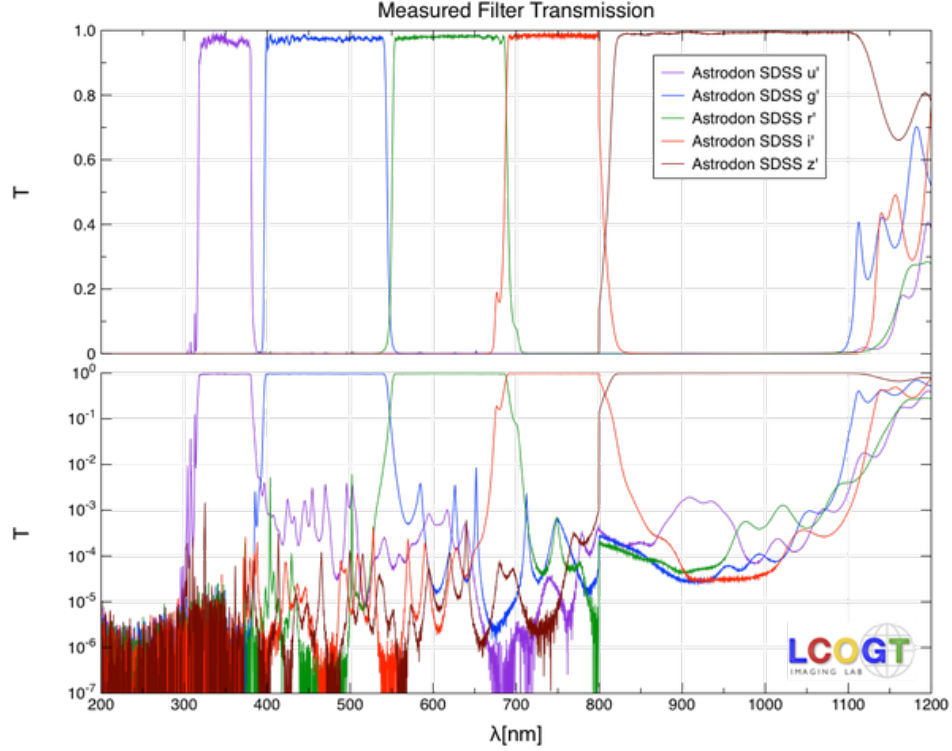


FIG. 3: This diagram shows the transmission functions for the SDSS u'g'r'i'z' filters. The filters available to us were u'g'r'i'. This graph is sourced from the Las Cumbres Observatory website (Las Cumbres Observatory, accessed 6/11/2023).

D. Finding Magnitudes using FITS Files

As detailed in the BANZAI section, all FITS files provide the flux of each source in the image. Extracting these are necessary to be able to calculate the redshifts in both the HyperZ method and the Nearest Neighbor method. The flux is recorded by taking the several Kron apertures around a source. These are named after Kron (1978), whose algorithm allowed the flux of a star to be estimated within rings of a certain radius, r_{kron} . In a circular aperture, ϵ , this is given by,

$$r_{kron} = \frac{\sum_{i \in \epsilon} r_i p_i^{(d)}}{\sum_{i \in \epsilon} p_i^{(d)}} \quad (8)$$

Where $p_i^{(d)}$ is the pixel value in the detection image at pixel i (the image centered around the source but not the

entire image) and r_i is essentially the distance from the centroid of the source (see Bertin and Arnouts (1996) for more details on the "reduced pseudo-radius").

According to Kron (1978), 90% of the flux of an object is captured inside a circle of radius $2r_{kron}$. Interestingly, this flux is independent of the magnitudes of our values. Thus, the total flux is given by the sum of the pixel values, $p_i^{(d)}$, inside the kron ellipse of radius $2r_{kron}$.

These kron magnitude values are programmed into the FITS header for each source in each image using SEP during the BANZAI pipeline. Using Astropy, a python library specifically designed for astronomical analysis by Astropy Collaboration et al. (2022), we were able to extract the flux for each source in our images, provided the source was bright enough for SEP to detect. To find the flux of a specific filter after we had taken multiple exposures, we summed the flux from each exposure and divided by the sum of the exposure time. Finally, we used this summed flux, the reference fluxes from Lupton et al. (1999), and b values from Table I to find the asinh magnitudes of each our sources using Eq. 6.

E. HyperZ

HyperZ works by comparing the Spectral Energy Distributions (SEDs) of a galaxy as seen through different filters to template SEDs. Like I mentioned earlier in the redshift section, the SEDs are picked to bracket important features, such as prominent Balmer Lines (the line from a Hydrogen atom shifting from state n to state 2) or the Lyman break at 3600 \AA , which corresponds to the maximum energy a hydrogen atom can absorb before it is ionized. A spectral energy distribution is a graph of the energy versus frequency (or wavelength). See Fig.4 for some SED examples.

For the template SEDs, Bolzonella et al. chose to use the GISEL98 (Galaxy Isochrone Synthesis Spectral Evolution Library) library from Bruzual A. and Charlot (1993) because of its accessibility and accu-

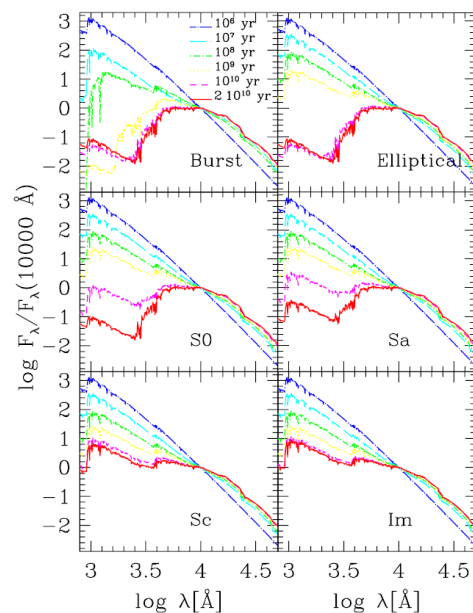


FIG. 4: This diagram from Bolzonella et al. (1999) shows the evolution of different SEDs over time using several models obtained from Bruzual A. and Charlot (1993).

racy at the time.

HyperZ matches the inputted data to the template SEDs by minimizing the equation,

$$\chi^2(z) = \sum_{i=1}^{N_{Filters}} \left[\frac{F_{obs,i} - b \times F_{temp,i}(z)}{\sigma_i} \right]^2 \quad (9)$$

Where $F_{obs,i}$ is the observed SED, $F_{temp,i}$ is the template, and its uncertainty in filter i , σ_i . b is a normalization constant and should not be confused with the softening parameter from the asinh equation. To make sure the redshift is accurate, it takes several other factors into account that could increase or decrease the redshift unduly as they find SEDs from inputted magnitudes. First, they consider star evolution and, using data from Bruzual A. and Charlot (1993), how the mass of the galaxy evolves over time. Next, they consider reddening from gas clouds absorbing light and increasing the redshift of the light emitted from the galaxy. Then, they consider metallicity of the stars in the galaxy, which can shift the spectrums redder, throwing it off the templates. In Fig. 4, we can see several SED evolutions. Using the magnitudes, their appearances in different filters, and Eqn. 9, it runs everything through a "hypercube" (a type of code too advanced for me to understand) that compares the properties of our sources and filters with the template SEDs. Notably, it requires us to input our filter's transmission functions, as shown Fig. 3, to get an accurate SED.

After running, it outputs a file that contains several guesses at the best photometric z , with probabilities attached to each number. In this paper, we chose to present the photometric redshifts with the highest probability.

F. Nearest Neighbor model

For the nearest neighbor model, we used the k-Nearest Neighbor (kNN) algorithm by Taunk et al. (2019) in conjunction with techniques presented in Beck et al. (2016), in which they detail how they used Nearest Neighbor algorithms and spectroscopic SDSS data to add photometric redshift data to the SDSS catalogue.

In essence, the algorithm finds the color indices as compared to magnitudes of galaxies with a recorded spectroscopic redshift. It then defines this quantity as part of a colour-magnitude space. It groups galaxies with similar redshifts near each other in this colour magnitude space. When a galaxy is then inputted into the program, it looks for galaxies with similar color-magnitude space coordinates. It then makes an estimate of the redshift based off of the closest matched galaxy to the input galaxy – the nearest neighbor, perchance.

Like in the HyperZ method, it does this numerically by attempting to minimize the equation. It starts with defining

$$z_{phot} = c_i + \mathbf{a}_i \mathbf{d}_i \quad (10)$$

Where c_i is an offset term, \mathbf{a}_i is a vector of linear coefficients, and \mathbf{d}_i is our input galaxy's coordinate vector in the color-magnitude space. The parameters for the equation above are found by minimizing the equation (from Beck et al. (2016)) below,

$$\chi^2(z) = \sum_{j \in NN} \frac{(z_j - c_i - \mathbf{a}_i \mathbf{d}_j)^2}{w_j} \quad (11)$$

Where the symbols indexed i represents our inputted galaxies and those with j represent our data set, in this case from Stoughton et al. (2002). z_j indicates the recorded redshift of template galaxies, c_i is an offset and \mathbf{a}_i is the vector of linear coefficients of our z_i coordinate while, finally, \mathbf{d}_j is the coordinate of our known galaxies in the color-magnitude space. w is the weight from the template galaxies, which depends on the situation but is usually the standard deviation of their distribution.

Since the above χ^2 is a little abstract, We simplified our distance metric to be

$$D = \sum_{j \in NN} |\mathbf{d}_i - \mathbf{d}_j| = \sum_{j \in NN} (r_i + g - r_i + r - i_i) - (r_j + g - r_j + r - i_j) \quad (12)$$

Where r , $g - r$, and $r - i$ are the color indices of the input galaxy, indicated by index i , and the template galaxies are indicated by j . This was minimized as our χ^2 .

Thus, the error in our redshift is given by,

$$\delta z_{phot,i} \approx \sqrt{\frac{\sum_{j \in NN} (z_j - z_{predicted})^2}{k}} \quad (13)$$

Where k is the total number of galaxies in our training set and $z_{predicted}$ is the average of the redshifts in the training set.

III. RESULTS

In Table III, we can see the galaxies we chose to observe using the LCO telescope system, galaxies A-D, as well as several galaxies whose data we sourced from the SIMBAD Astronomical Database (Wenger et al. (2000)) to test the accuracy of our methods as they have accepted redshifts.

To find our chosen galaxies, we used a database to look for galaxies that (with help from Professor Lubin and our TA, Jeongwha) would be visible from our telescopes, with a magnitude less than approximately 16, and have a moderately high redshift in comparison. We used the 6dF Galaxy Survey (Jones et al. (2004)) to choose our galaxies as it had distance, magnitude, and redshifts for each galaxy in the survey.

Galaxy Letter	Galaxy Name	Accepted Redshift
A	NGC 4594 (Sombrero Galaxy)	0.003416 ± 0.000017
B	2MASX J12275231-1120370	0.08504 ± 0.00015
C	6dFGS gJ124555.9-424016	0.04768 ± 0.00015
D	2dFGRS TGN132Z253	0.11333 ± 0.00003
E	M101 (Pinwheel Galaxy)	0.000811 ± 0.000016
F	NGC 4151	0.003262 ± 0.000067

TABLE III: This table shows the galaxies with their accepted values of redshift, as well as their designated letters for ease of reference. Images of galaxies A-D were obtained using the LCO telescope system and E,F were taken from SIMBAD (Wenger et al. (2000)) for testing purposes.

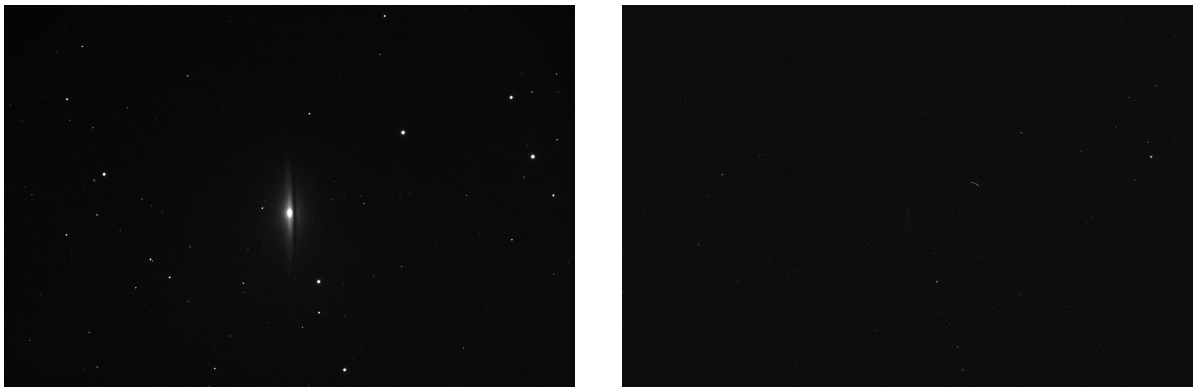


FIG. 5: These two images show the Sombrero Galaxy through the r' filter (left) and the u' filter (right), post BANZAI.

Since galaxy A also doubled as an experiment in astrophotography, we used four filters with five exposures on each. It also has a very small redshift, so we didn't initially plan to use it for this paper, but we included it to show how our programs worked with very small redshifts. The exact amount of exposures, what filters were used, and their exposure times, can be seen in Table IV. Otherwise, exposures of the galaxies taken with the u' filter were too faint to be detected by SEP, so we only used filters g' , r' and i' . As such, they only have color indices $g-r$ and $r-i$ (see Table II). As can be seen in Fig. 5, even the Sombrero galaxy, which has a magnitude of 8 (Jones et al. (2004)), is not picked up in the u' filter. Of the galaxies we took images of using LCO, the Sombrero galaxy had the brightest magnitude, meaning that our other galaxies were not visible at all through the u' filter. The magnitudes of the galaxies in each filter is featured in Table V. These were found using the technique outlined earlier.

Galaxy	RA (HH:MM:SS)	Dec(HH:MM:SS)	SDSS Filters	Exposures per Filter	Exposure Time (s)
A_{exp}	+12 : 39 : 59.43	−11 : 37 : 22.97	u',g',r',i'	5	100
B_{exp}	+12 : 27 : 52.30	−11 : 20 : 37.1	g',r',i'	10	100
C_{exp}	+12 : 46 : 20.79	−42 : 41 : 38.0	g',r',i'	10	100
D_{exp}	+12 : 58 : 22.25	−03 : 06 : 05.7	g',r',i'	10	100

TABLE IV: Observational methods. This table displays the coordinates where the telescope was aimed towards for each galaxy and which filters were used to capture the galaxy.

	r' magnitude	r' error	g' magnitude	g' error	i' magnitude	i' error
A_{exp}	13.47746163	75.87	14.79941838	53.91	13.56564163	58.78
B_{exp}	20.5571003	2.68	21.70406992	2.22	20.96777631	2.41
C_{exp}	20.41815517	2.56	22.09272542	1.59	20.55157546	2.51
D_{exp}	19.38510287	3.44	20.70239522	3.02	19.61346501	3.05
D_{acc}	17.511	0.007	18.441	0.01	17.051	0.007
E_{acc}	11.529	0.002	12.219	0.002	11.237	0.002
F_{acc}	11.613	0.006	12.284	0.006	11.081	0.007

TABLE V: This table shows the asinh magnitudes used to find the results presented in Tables VII and VI.

A_{exp} - D_{exp} were found using the technique outlined in the "Finding Magnitudes using FITS Files" section.

Unfortunately, the creators of HyperZ did not include a way to find the error of the measurement in their documentation and we were unable to derive the error ourselves. As such, we included the probability measurement instead. This acts as a measure of confidence in the accuracy of the program's fit for the redshift. These can be seen in table VII, while the results for the Nearest Neighbor method can be seen in VI.

In Fig. 6, we can the results show in Tables VII and VI graphed against each other as well as against the accepted redshifts of each galaxy in our survey.

Galaxy	Accepted Redshift	Nearest Neighbors Redshift	Percent Error
A_{exp}	0.00342 ± 0.00002	0.326 ± 0.449	9430%
B_{exp}	0.0850 ± 0.0002	0.458 ± 0.269	439%
C_{exp}	0.0477 ± 0.0002	0.420 ± 0.205	781%
D_{exp}	0.113 ± 0.00003	0.310 ± 0.105	174%
D_{acc}	0.113 ± 0.00003	0.125 ± 0.122	10.6%
E_{acc}	0.000811 ± 0.000016	0.000461 ± 0.000000	43.2%
F_{acc}	0.00326 ± 0.00007	0.000461 ± 0.000000	85.9%

TABLE VI: This table shows the redshifts obtained using the Nearest Neighbors method as compared to the accepted redshifts, as taken from SIMBAD. The redshifts were calculated using magnitude values obtained either using the technique detailed in the "Finding Magnitudes" section or magnitudes obtained from SIMBAD. Galaxies A-C only have magnitudes obtained from experimental data while galaxies E and F only have database magnitudes.

D has both, which allows for greater comparison of the efficiency of our technique.

Galaxy	Accepted Redshift	HyperZ Redshift	Probability (%)	Percent Error
A_{exp}	0.00342 ± 0.00002	0.443	100.00	12900%
B_{exp}	0.0850 ± 0.0002	0.385	99.40	353%
C_{exp}	0.0477 ± 0.0002	0.430	99.33	801%
D_{exp}	0.113 ± 0.00003	0.423	99.81	274%
D_{acc}	0.113 ± 0.00003	0.065	97.63	42.5%
E_{acc}	0.000811 ± 0.000016	0.066	97.58	8038%
F_{acc}	0.00326 ± 0.00007	0.607	96.53	18500%

TABLE VII: This table shows the redshifts obtained using the HyperZ method. Like in Table VI, galaxies A-C lack accepted values of magnitude and D-F were obtained from SIMBAD.

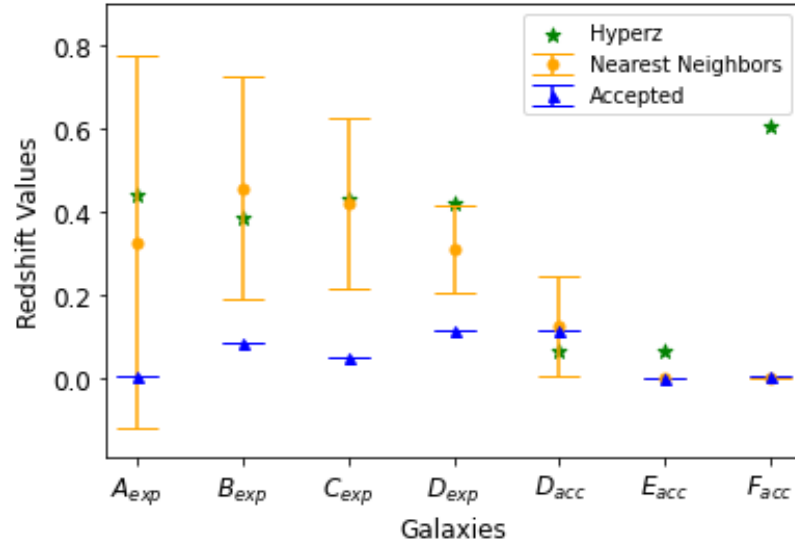


FIG. 6: This graph shows the redshift values obtained using the HyperZ method, the Nearest Neighbors method, and the accepted values for each galaxy.

IV. DISCUSSION

As discussed by the creators of HyperZ, sources with lower redshift are much more likely to be inaccurate. This is because the spectrums of sources with lower redshift don't display more prominent or recognizable lines like those produced the 4000 Å break, the Lyman Forest, or longer wavelength Balmer lines (see Bolzonella et al. (2000) for more details). They note that when $z \leq 0.2$, the galaxies do not show up on the u' filter and that the other filters are unable to detect any of the stronger spectral features mentioned above. We can see this inaccuracy in galaxy A, which has an accepted redshift of 0.00342 and obtained a percent error in our measurements of 9430% and 12900%, which is at least two orders of magnitude great than the rest of the galaxies in the Nearest Neighbor method and almost all of the galaxies in the HyperZ method.

The 0.4 m telescopes that we used for this project are only able to see galaxies up to approximately 1. Including the near-infrared z' filter can greatly increase accuracy but the 0.4m telescopes do not come equipped with the z' filter.

In addition, the conversion between the source fluxes and the asinh magnitudes detailed in the introduction is not exactly for the SDSS u'g'r'i'z' system we are using. The conversion rate was originally designed for the 2.5m SDSS ugriz filter system. Due to differences between the filter constructions (the primed filter system was in air for instance, while the unprimed filter system was in vacuum), there is about a 2.5 to 10% difference in flux depending on the filter (u, g, r, i, or z). See Stoughton et al. (2002) for more details about the differences. This may have manifested in our errors in our magnitudes, as shown in Table V.

We suspect that the above problems with magnitude may have affected our redshifts, as the trials done using the accepted magnitudes from databases were very close to the recorded redshifts of the galaxies : 42.5% for D_{acc} in the HyperZ method, and 10.6%, 43.2% and 85.9% for D_{acc} , E_{acc} and F_{acc} , respectively.

Overall, our two methods were generally close to each other, as seen in Fig. 6, but were not very accurate, as indicated by the large percent differences in Tables VII and VI. However, the two methods, if deployed with more filters, more access to data such as weather conditions, etc, could produce much more accurate results.

V. REFERENCES

- Barbara Ryden. *Introduction to Cosmology*. Cambridge University Press, New York, NY, 2 edition, 2017.
- Harold T. Stokes. Open access redshift diagram. URL https://stokes.byu.edu/teaching_resources/computer_resources.html.
- Pierre Lena, Daniel Rouan, Francois Lebrun, Francois Mignard, and Didier Pelat. *Observational Astrophysics*. Springer, 3 edition, 2008.
- Las Cumbres Observatory. Transmission functions. URL <https://lco.global/observatory/instruments/filters/>.
- Chris Stoughton, Robert H. Lupton, Mariangela Bernardi, Michael R. Blanton, Scott Burles, et al. Sloan Digital Sky Survey: Early Data Release. *The Astronomical Journal*, 123(1):485–548, January 2002. doi: 10.1086/324741. URL <https://iopscience.iop.org/article/10.1086/324741/pdf>.
- Robert H. Lupton, James E. Gunn, and Alexander S. Szalay. A Modified Magnitude System that Produces Well-Behaved Magnitudes, Colors, and Errors Even for Low Signal-to-Noise Ratio Measurements. *The Astronomical Journal*, 118(3):1406–1410, September 1999. doi: 10.1086/301004.
- Martin V. Zombeck. *Handbook of space astronomy astrophysics*. Cambridge University Press, New York, NY, 2 edition, 1990.
- Curtis McCully, Monica Turner, N Volgenau, Daniel Harbeck, et al. Lcogt/banzai: Initial release (0.9.4). 2018. doi: 10.5281/zenodo.1257560. URL <https://doi.org/10.5281/zenodo.1257560>.
- Curtis McCully, Monica Turner, David Collom, and Matt Daily. *BANZAI Documentation, Release 0.23.0*, 5 2019.
- E. Bertin and S. Arnouts. SExtractor: Software for source extraction. *Astronomy and Astrophysics Supplement*, 117:393–404, June 1996. doi: 10.1051/aas:1996164.
- Kyle Barbary. Sep: Source extractor as a library. *Journal of Open Source Software*, 1(6):58, 2016. doi: 10.21105/joss.00058. URL <https://doi.org/10.21105/joss.00058>.
- Dustin Lang, David W. Hogg, Keir Mierle, Michael Blanton, and Sam Roweis. Astrometry.net: Blind astrometric calibration of arbitrary astronomical images. *The Astronomical Journal*, 139(5):1782–1800, March 2010. doi: 10.1088/0004-6256/139/5/1782.
- M. Fukugita, T. Ichikawa, J. E. Gunn, M. Doi, K. Shimasaku, and D. P. Schneider. The Sloan Digital Sky Survey Photometric System. *Astronomical Journal*, 111:1748, April 1996. doi: 10.1086/117915.
- R. G. Kron. *Photometry of a Complete Sample of Faint Galaxies*. PhD thesis, University of California, Berkeley, December 1978.
- Astropy Collaboration, Adrian M. Price-Whelan, Pey Lian Lim, Nicholas Earl, Nathaniel Starkman, et al. The Astropy

- Project: Sustaining and Growing a Community-oriented Open-source Project and the Latest Major Release (v5.0) of the Core Package. *Astrophys. J.* , 935(2):167, August 2022. doi: 10.3847/1538-4357/ac7c74.
- Micol Bolzonella, Joan-Marc Miralles, and Roser Pell’o. *HyperZ v1.2 User’s Manual*, 1999. URL http://www.bo.astro.it/~micol/Hyperz/hyperz_manual1.2.pdf.
- Gustavo Bruzual A. and Stephane Charlot. Spectral Evolution of Stellar Populations Using Isochrone Synthesis. *Astrophysical Journal*, 405:538, March 1993. doi: 10.1086/172385.
- Kashvi Taunk, Sanjukta De, Srishti Verma, and Aleena Swetapadma. A brief review of nearest neighbor algorithm for learning and classification. In *2019 International Conference on Intelligent Computing and Control Systems (ICCS)*, pages 1255–1260, 2019. doi: 10.1109/ICCS45141.2019.9065747.
- Róbert Beck, László Dobos, Tamás Budavári, et al. Photometric redshifts for the SDSS Data Release 12. *Monthly Notices of the Royal Astronomical Society*, 460(2):1371–1381, 04 2016. ISSN 0035-8711. doi: 10.1093/mnras/stw1009. URL <https://doi.org/10.1093/mnras/stw1009>.
- M. Wenger, F. Ochsenbein, D. Egret, P. Dubois, F. Bonnarel, et al. The SIMBAD astronomical database. The CDS reference database for astronomical objects. , 143:9–22, April 2000. doi: 10.1051/aas:2000332.
- D. Heath Jones, Will Saunders, Matthew Colless, Mike A. Read, Quentin A. Parker, et al. The 6dF Galaxy Survey: samples, observational techniques and the first data release. *Monthly Notices of the Royal Astronomical Society*, 355(3):747–763, December 2004. doi: 10.1111/j.1365-2966.2004.08353.x.
- Micol Bolzonella, Joan-Marc Miralles, and Roser Pell’o. Photometric redshifts based on standard sed fitting procedures. *Astronomy and Astrophysics*, 363:476–492, October 2000.