SUPERNOVA ACCELERATION PROBE: INVESTIGATING PHOTOMETRIC REDSHIFT OPTIMIZATION

Tomas Dahlen[1], Bahram Mobasher[1], Stephanie Jouvel[2], Jean-Paul Kneib[2], Olivier Ilbert[2], Stephane Arnouts[2],
Gary Bernstein[3], and Jason Rhodes[4,5]

[1] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[2] Observatoire Astronomique de Marseille-Provence, F-13376 Marseille, France
[3] Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA
[4] California Institute of Technology, 1201E California Blvd., Pasadena, CA 91125, USA
[5] Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA 91109, USA

ABSTRACT

The aim of this paper is to investigate ways to optimize the accuracy of photometric redshifts for a Supernova Acceleration Probe (SNAP)-like mission. We focus on how the accuracy of the photometric redshifts depends on the magnitude limit and signal-to-noise ratio (S/N), wavelength coverage, and the number of filters and their shapes and observed galaxy type. We use simulated galaxy catalogs constructed to reproduce observed galaxy luminosity functions from the Great Observatories Origins Deep Survey and derive photometric redshifts using a template fitting method. By using a catalog that resembles real data, we can estimate the expected number density of galaxies for which photometric redshifts can be derived. We find that the accuracy of the photometric redshifts is strongly dependent on the S/N (i.e., S/N > 10 is needed for accurate photometric redshifts). The accuracy of the photometric redshifts is also dependent on galaxy type, with smaller scatter for earlier-type galaxies. Comparing results using different filter sets, we find that including the *U*-band and near-IR bands is important for decreasing the fraction of outliers, i.e., "catastrophic failures." Using broad overlapping filters with resolution ∼4 gives better photometric redshifts compared to narrower filters (resolution $\gtrsim$ 5) with the same integration time. We find that filters with square response curves result in a slightly higher scatter, mainly due to a higher fraction of outliers at faint magnitudes. We also compare a 9-filter set to a 17-filter set, where we assume that the available exposure time per filter in the latter set is half that of the first set. We find that the 9-filter set gives more accurate redshifts for a larger number of objects and reaches higher redshift, while the 17-filter set gives better results at bright magnitudes.

Key words: cosmology: observations – galaxies: distances and redshifts

Online-only material: color figures

1. INTRODUCTION

In recent years, there has been unprecedented progress in observational astronomy due, in large part, to the advent of large format and highly sensitive optical/infrared detectors. Installation of these cameras on 8 m ground-based telescopes and space-borne facilities has enabled planning of large and deep galaxy surveys, increasing the discovery space by over an order of magnitude. In particular, wide-area multi-waveband imaging from space, complemented by follow-up ground-based observations, has provided extremely valuable data sets for studying diverse topics in observational astronomy and cosmology. For example, installation of the Advance Camera for Survey (ACS) on the *Hubble Space Telescope* (*HST*) has resulted in multi-waveband surveys of galaxies, including the Great Observatories Origins Deep Survey (GOODS; Giavalisco et al. 2004), COSMOS (Scoville et al. 2007), and the Hubble Ultra-Deep Field (HUDF; Beckwith et al. 2006). These surveys provide deep multi-waveband data covering large areas, used to study a number of issues concerning formation and evolution of galaxies, including the study of rest-frame properties of different populations of galaxies (e.g., Bundy et al. 2005; Grogin et al. 2005; Dahlen et al. 2007), search for the highest redshift (Kneib et al. 2004; Bouwens et al. 2005) and new population of galaxies (Wiklind et al. 2008), mapping of the dark matter distribution in strong lensing clusters (Smith et al. 2005; Broadhurst et al. 2005; Limousin et al. 2007; Richard et al. 2007), cosmological constraints from weak lensing

(Massey et al. 2007a), clustering of galaxies (McCracken et al. 2007) and the three-dimensional large-scale structure dark matter distribution (Massey et al. 2007b).

Among the most important outcomes from these studies was the first ever space-borne search for Supernovae Type Ia (SNe Ia) in the GOODS fields (Riess et al. 2004). These observations have a darker sky background, leading to deeper images, and significantly narrower point-spread functions (PSFs), leading to better spatial resolution, and hence identification of more distant supernovae in galaxies, compared to ground-based images. This allowed the discovery of 23 high-redshift SNe Ia at z > 1, which includes almost all (but one) of the highest redshift SNe Ia known at the time (Strolger et al. 2004; Dahlen et al. 2008). Combining these high-z and nearby SNe Ia, the Hubble diagram was established, allowing significant constraints on the properties of dark energy and its equation of state (Riess et al. 2004, 2007). An essential component of this study was the measurement of photometric redshifts of the hosts of SN candidates to identify objects of the highest interest for follow-up with subsequent spectroscopic observations. This was possible due to the availability of multi-waveband data from space- and ground-based observations.

The SuperNova Acceleration Probe (SNAP[6]) mission is aimed at finding thousands of SNe of various types to redshift z ∼ 1.7, allowing a detailed study of reliability of SNe Ia as standard candles (i.e., dependence of their observed properties

---

[6] http://snap.lbl.gov/.

on the type of their host galaxy and its redshift, effect of dust, and their frequency of appearance in early-type galaxies). This proposed mission will improve the use of SNe Ia as standard candles and provide significant constraints on the properties of dark energy and its nature. Moreover, it will provide a large sample of SNe Type II, which are used as diagnostics for star-formation activity in galaxies, allowing a statistically large sample of these objects to examine the evolution of star formation and metallicity with redshift (Dahlen et al. 2004). As part of the SNAP's primary mission, a wide (at least 1000 deg² per year), multiband survey will be conducted. This survey will be optimized for the detection of weak gravitational lensing, a powerful probe of dark energy. Weak lensing provides a direct way for measuring the distribution of dark matter in the universe. The evolution of dark matter structures over cosmic time is governed by the nature of the dark energy. Thus, accurate photometric redshifts, which are required to measure the three-dimensional distribution of dark matter, are necessary to exploit weak lensing as a probe of dark energy. Furthermore, these multi-waveband deep data will be extremely useful in studying formation and evolution of galaxies, groups and clusters as a function of their redshift, morphology, environment, and color and star-formation properties. It will also be a unique mission to probe the first quasars and the first luminous galaxies in the universe, thus probing the epoch of cosmic reionization.

Future dark energy probes based on SN Ia and weak lensing ideally require accurate redshifts for individual galaxies. However, given the size of the planned surveys and their depth, it is not practical (nor feasible) to measure spectroscopic redshifts for all the observed galaxies. Therefore, a critical evaluation of the photometric redshift capabilities of any of the Joint Dark Energy Mission (JDEM) experiments is key to optimize the design of the respective mission. This requires an optimization of the number of filters used, their spectral resolution and shapes, and thus throughputs as well as the overall wavelength coverage, to allow most accurate measurement of photometric redshifts. To investigate these problems, we use observational data from the GOODS fields to create a mock galaxy catalog containing a large set of objects with known properties, i.e., redshift, spectral type, luminosity, and amount of internal extinction. We then run photometric redshift codes on the multi-waveband data for galaxies in the mock catalog. The aim of this investigation is to study how the accuracy of the galaxy photometric redshifts depends on a range of factors including redshift, signal-to-noise ratio (S/N), magnitude limits, galaxy spectral type, and, in particular, the number of filters, filter shapes and bandwidths. For an investigation on how to optimize filters for Type Ia cosmology investigations, see Davis et al. (2006). In a following paper (S. Jouvel et al. 2008, in preparation), we will investigate in further detail the impact of calibration, spectral energy distribution (SED) evolution, and size of the spectroscopic surveys on the photometric redshift determination. Here we will also discuss how shape and surface brightness information may improve photometric redshift, as previously discussed by, e.g., Kurtz et al. (2007) and Wray & Gunn (2008).

Throughout this paper we use $\Omega_\Lambda = 0.7$, $\Omega_M = 0.3$, and $H_0 = 70$ km s^{-1} Mpc^{-1}. Magnitudes are in the AB system.

2. THE MOCK GALAXY CATALOG

To investigate the expected behavior of the photometric redshifts for an SNAP-like mission, we create a mock galaxy catalog. In the mock catalog, we assign to each galaxy a redshift, a spectral type, an absolute luminosity, and a value for extinction.

The redshift, spectral type, and absolute luminosity are drawn from a distribution according to the observed type-specific luminosity functions (LFs) derived from GOODS (Dahlen et al. 2005). First, a redshift is assigned by the redshift-dependent LF. Second, at the assigned redshift, an absolute magnitude in the range $-24 < M_B < -13$ is given to the object according to the LF at that redshift. Finally, the galaxy type is assigned to the object with dependences on both redshift and absolute magnitude. Redshifts are distributed in the range $0 < z < 6$. The spectral templates used cover types E, Sbc, Scd, Im (Coleman et al. 1980), and two starbursts from Kinney et al. (1996; templates SB2 and SB3). The templates are extended in ultraviolet (UV) and near-IR (NIR) wavelengths as described in Mobasher et al. (2007). The template set used is shown in Figure 1. To get a continuous set of templates, we make random linear interpolations between adjacent templates when assigning type to the mock galaxies. A random internal extinction is also assigned to each galaxy with a maximum value $E_{B-V} = 0.10$ for early types and $E_{B-V} = 0.30$ for starbursts. For star-forming galaxies, we use a Calzetti et al. (2000) extinction law, while for later-type galaxies we assume a Galactic extinction law (Cardelli et al. 1989). We hereafter refer to the redshift in the mock catalog as spectroscopic redshifts.

Using the template SEDs, extinction values, and the response functions for the filter set, we calculate the K-corrections corresponding to the spectroscopic redshift for each galaxy. The K-corrections together with the absolute magnitudes and distance moduli give us the set of apparent magnitudes in each band. To each of these "incident" magnitudes, we add an error to derive the actual "measured" magnitudes that go into the mock galaxy catalog. This "statistical" error is derived from each "incident" magnitude using the S/N at this magnitude (where S/N = 10 corresponds to a magnitude error $\sigma_m \sim 0.10$). We thereafter add the error to the magnitude, assuming that errors have a Gaussian probability distribution with σ_m as dispersion. This gives the "measured" magnitude. Furthermore, an additional error of 1% of the flux is also added in quadrature. This accounts for, e.g., zero-point uncertainties and photometry uncertainty due to nonperfect image reductions. The default filter set is shown in the top panel of Figure 2. This consists of six optical and three NIR filters, indexed 0–8. Besides the standard set, we also include a "U-band" filter set shown in the middle panel of Figure 2. Here we have stretched the standard filter set into the U-band so that the bluest filter has an effective wavelength $\lambda_{eff} = 3910$ Å, compared to the standard filter set which has $\lambda_{eff} = 4750$ Å for the bluest filter. We include this set to examine the importance of the U-band when it comes to photo-z accuracy and the minimization of the outlier fraction. Furthermore, we include a 17-filter set which consists of the same filters as the U-band set with the addition of eight intermediate filters (bottom panel of Figure 2). The working concept for the SNAP focal plane is to distribute the six optical detectors and three infrared (IR) detectors in multiple squares with sizes 6×6 and 3×3, respectively, where each detector has a fixed filter (see Figure 15 in Aldering et al. 2004). This design allows an efficient scanning of all objects in all nine filters. Therefore, it is not possible to add the U-band as a 10th filter in an efficient way, instead we have here adjusted the throughput of the existing filters. For the 17-filter set, the concept is to keep the nominal number of detectors for one of the optical filters, while replacing half of the remaining detectors with the new intermediate filters. This will allow the same scanning advantages, with the exception that 16 of the 17
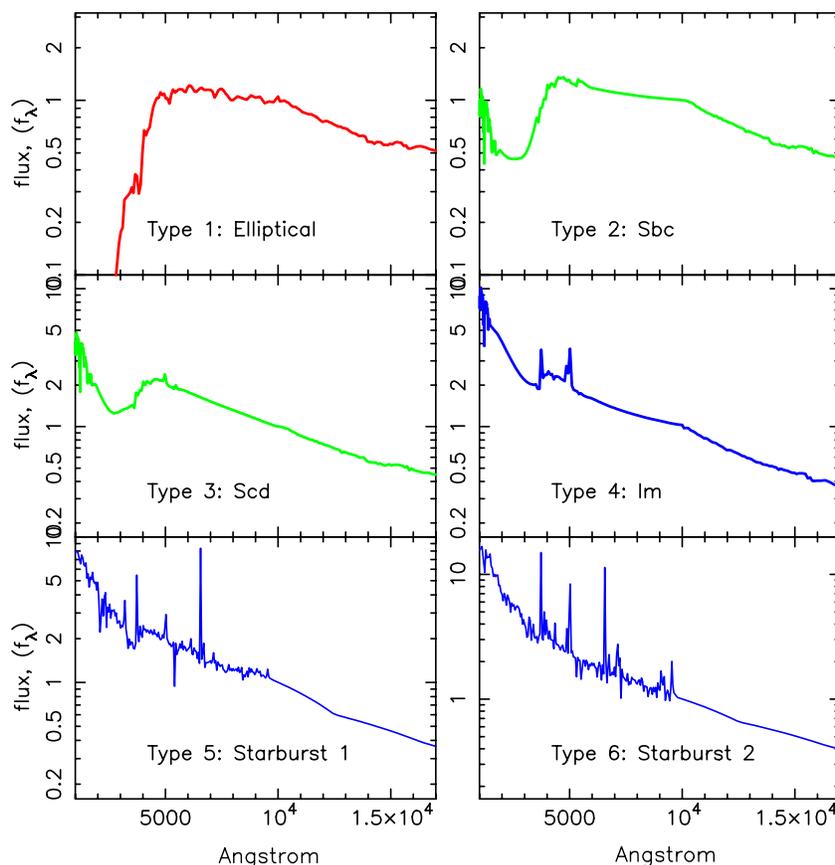
Figure 1. Six template SEDs that are used in constructing the mock galaxy catalog. Templates 1–4 are taken from Coleman et al. (1980), while the two starbursts, templates 5–6, are taken from Kinney et al. (1996). Templates are extended into UV and NIR as described in Mobasher et al. (2007).

(A color version of this figure is available in the online journal.)

filters will be less deep by a factor of 2 compared to the 9-filter option.

In Table 1, we give the limiting magnitudes for S/N = 10 and filter characteristics corresponding to the filter sets in Figure 2. These are based, with some adjustments, on values given in Aldering et al. (2004), but with fainter limits in the IR wavelengths, reflecting the recent increase in the quantum efficiency for these detectors. To get a U-band S/N limit, we assume that the detector quantum efficiency integrated over the U-band is ∼65% of that of the B-band. Note that due to the planned design of SNAP, the effective exposure time for each of the NIR bands is twice that of the optical bands; therefore, a similar depth (in AB magnitude) is reached over the whole wavelength range. We assume the same total exposure times for all the different scenarios listed in Table 1. Therefore, for the cases with more filters, the exposure time per filter is lower, leading to the lower S/N ratios. This allows a more meaningful comparison between different scenarios. For the detection filter in the 17-filter set (filter #11, corresponding to filter #5 in the 9-filter sets), the exposure time is kept the same as for the other filter sets. This means that a similar number of objects should be detected in all sets when using this band as the detection band. Note that the filter shapes, as well as S/N values, should be viewed as representative for an SNAP-like survey, and that the latter do not include details about the detective quantum efficiency (DQE) shapes, mirror reflectivities, etc. This will be addressed in S. Jouvel et al. (2008, in preparation).

To check how well our code simulates real observed galaxy samples, we produce an additional mock galaxy catalog using the GOODS filter set and compare with the observational data from GOODS-South. Since the mock galaxy catalog is produced using absolute magnitude, spectral type, and redshift distributions derived from the GOODS data, a comparison between the observed and simulated galaxy number counts primarily examines whether the mock catalog is consistently produced. In Figure 3, we show B-, R-, and J-band number counts both from the GOODS observations and the mock galaxy catalog. We find that the observed and simulated number counts are in good agreement in all bands. While the mock galaxy catalog is produced from the rest-frame B-band LF, the good agreement in observed B-, R-, and J-bands assures us that the mock galaxy colors well represent the colors of the GOODS galaxies. The GOODS data are not as deep as the mock galaxy catalog, making the GOODS counts incomplete at faint magnitudes. To further investigate how the mock galaxy catalog reproduces adequate colors at different magnitudes, we plot in Figure 4 the observed and mock galaxy $B − R$ colors in two different magnitude bins. We find a good agreement as well as that the brighter sample has redder colors as expected.

In total, there are ∼1.6 × 10^6 galaxies over an area of 1000 arcmin² within the specified absolute magnitude and redshift ranges described above. What is more interesting is the number of galaxies that have an apparent magnitude brighter than a given magnitude limit. In Table 2, we give the total number of galaxies as well as the number in ten equally spaced redshift bins to $z < 3$. We also give the predicted numbers to $z = 6$ in coarser bins. Results are given after applying different S/N cuts for the U-band filter set. The least restrictive cut is
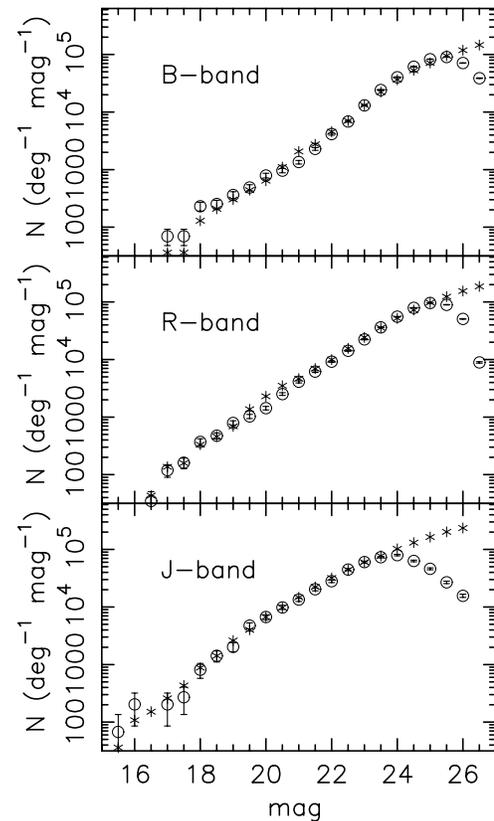
Table 1
S/N = 10 Limiting Magnitudes

Filter	Standard	λ_{eff}	$\Delta\lambda$	U-band Set	λ_{eff}	$\Delta\lambda$	17-band Set	λ_{eff}	$\Delta\lambda$
0	26.8	4750	1070	26.6	3950	990	26.2	3950	990
1	26.7	5450	1240	26.7	4730	1180	26.3	4340	1090
2	26.6	6270	1440	26.6	5610	1410	26.3	4730	1180
3	26.6	7200	1630	26.6	6640	1660	26.2	5170	1300
4	26.6	8270	1870	26.6	7820	1950	26.2	5610	1410
5	26.6	9610	2190	26.6	9270	2320	26.2	6130	1540
6	26.7	10960	2500	26.7	10730	2680	26.2	6640	1660
7	26.7	12570	2820	26.7	12460	3110	26.2	7230	1810
8	26.7	14450	3300	26.7	14450	3620	26.2	7820	1950
9							26.2	8540	2140
10							26.6	9270	2320
11							26.3	10000	2500
12							26.3	10730	2680
13							26.3	11590	2900
14							26.3	12460	3110
15							26.3	13450	3360
16							26.3	14450	3620

Notes. Indicative S/N = 10 limits for SNAP filters based on values from Aldering et al. (2004). The U-band filter set is constructed by stretching the standard filter set to shorter wavelengths, while the 17-filter set is based on the U-band set with eight additional intermediate filters. Due to the stretching, the latter two sets have somewhat wider filters (lower resolution) compared to the standard set. Filter sets used are shown in Figure 2.
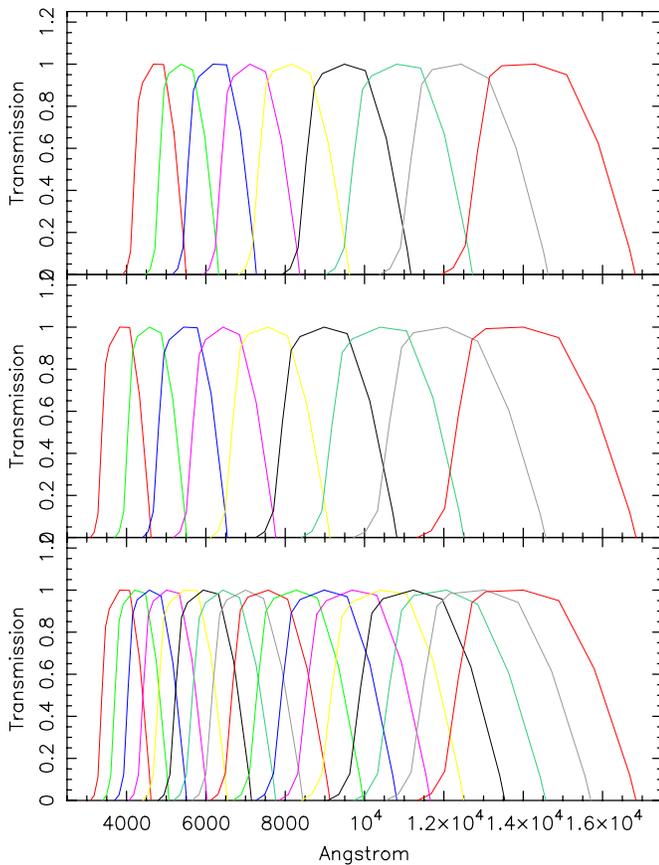

Figure 2. The top panel shows the default filter set used consisting of six optical and three NIR filters. Filters are indexed 0–8. The middle panel shows the U-band filter set constructed by stretching the standard filter set into the U-band. A 17-filter set (bottom panel) is constructed from the U-band filter set by adding eight intermediate filters. All filters have the peak transmission normalized to unity.

Figure 3. Number counts in B-, R-, and J-band for data from GOODS (open circles with statistical error bars) and a mock galaxy catalog produced using the GOODS filter response functions (asterisk).

S/N > 10 in-any-filter. Here, it suffices that the magnitude in at least one of the filters is brighter than the S/N = 10 limits in Table 1. The remaining selections are based on the I-band (filter 5), where we choose limits $m_5 < 26.6$ (corresponding to S/N = 10), $m_5 < 25.6$, and $m_5 < 24.6$. Note here that the field size chosen results in statistical errors in the number of objects (and in all bins) that are insignificant compared to
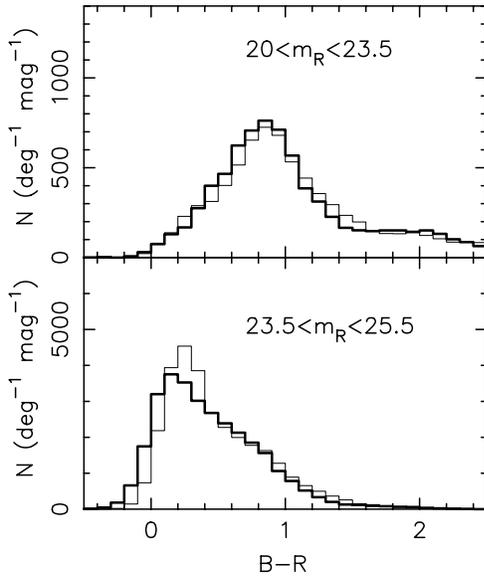
Figure 4. Observed $B-R$ for data from GOODS (thick line) and from the mock galaxy catalog (thin line). Note the good agreement at both bright and faint magnitudes.

Table 2
Galaxy Number Counts

Redshift	S/N > 10 in-any[a]	$m_5 < 26.6$	$m_5 < 25.6$	$m_5 < 24.6$
	Number of galaxies arcmin^{-2}			
$0.0 < z < 3.0$	156	110	67	36
$0.0 < z < 0.3$	5.0	5.0	4.8	3.7
$0.3 < z < 0.6$	15	15	12	8.0
$0.6 < z < 0.9$	21	20	13	8.0
$0.9 < z < 1.2$	23	21	12	6.9
$1.2 < z < 1.5$	23	17	9.2	4.8
$1.5 < z < 1.8$	24	13	6.5	2.8
$1.8 < z < 2.1$	15	7.3	3.6	1.3
$2.1 < z < 2.4$	14	5.4	2.2	0.54
$2.4 < z < 2.7$	9.8	4.1	1.4	0.27
$2.7 < z < 3.0$	7.0	3.1	0.96	0.14
$3.0 < z < 4.0$	16	7.3	1.4	0.14
$4.0 < z < 5.0$	12	4.6	0.41	0.01
$5.0 < z < 6.0$	4.2	1.3	0.10	0.00

Notes. [a] The selection requires that the magnitude in at least one filter is brighter than the S/N values given in Table 1 for the U-band filter set. Other filter sets should have comparable numbers, assuming that S/N in the detection band is the same. Clustering variance may add an uncertainty of at least ~20%. Numbers at $z > 3$ are more uncertain since they are derived using extrapolations of lower-redshift LFs.

the uncertainty due to cosmic variance. We estimate that the uncertainty in the GOODS LF due to cosmic variance results in an uncertainty in the numbers in Table 2 of ~20% (Dahlen et al. 2005).

3. PHOTOMETRIC REDSHIFTS

To derive photometric redshifts, we use the template fitting method (e.g., Gwyn 1995; Mobasher et al. 1996). This method compares the observed and template SEDs in redshift intervals and assigns photometric redshift and spectral type to individual galaxies by minimizing the χ^2 values:

$$\chi^2 = \sum_{i=1}^{n} \left(\left[F_{\mathrm{obs}}^i - \alpha F_{\mathrm{template}}^i \right] / \sigma^i \right)^2, \quad (1)$$

where the summation is taken over the n filters available and F_{obs}^i and F_{template}^i are the observed and template fluxes in band i, respectively. Here F_{template}^i includes information on the template SEDs for different galaxy types, absorption values, and redshifts, as well as the response curves for the filters. We assume that the latter is well known and do not introduce any additional error exceeding the already included extra errors of 1%. Finally, α is a normalization constant and σ^i is the flux error in band i. We use the six template SEDs described above, together with two interpolations between each template going from early to later types, making the full set consisting of 16 discrete templates (in contrast to the continuous set in the mock galaxy catalog). Each template SED is redshifted in the range $0 < z < 6$ in steps $\Delta z = 0.01$. We include an LF Bayesian prior in the photo-z fitting. For the prior, we calculate the absolute magnitude the galaxy would have at each tested redshift and compare this with an input LF. If the absolute magnitude corresponding to a particular redshift is improbable, i.e., significantly brighter than M^*, then this redshift is disfavored. The input LF used here is independently chosen and is not the same as the LF from which the mock catalogs are generated. The template fitting method and priors we use are further described in Dahlen et al. (2005). We use the same template set for constructing the mock galaxy catalog and calculating photometric redshifts, although the mock galaxies are drawn from a continuous set of galaxies and have had their photometry adjusted by extinction as well as statistical error. Therefore, we can investigate how the photometric redshifts depend on various parameters such as S/N, filter shapes, and wavelength coverage, without introducing any bias due to the choice of the templates. However, as a consistency check, we also construct a mock catalog from an alternative set of template SEDs and derive photometric redshifts using the original templates in the fitting. This is described in Section 5.4.

4. RESULTS

We present results on the accuracy of the photometric redshifts after applying different magnitude cuts as discussed above. The accuracy of the photometric redshifts, σ_z, is defined as

$$\sigma_z \equiv \mathrm{rms}[(z_{\mathrm{spec}} - z_{\mathrm{phot}})/(1 + z_{\mathrm{spec}})]. \quad (2)$$

For the results presented here we focus on a redshift range $0 < z_{\mathrm{phot}} < 3$, while the redshift range of the mock galaxy catalog is $0 < z_{\mathrm{spec}} < 6$. Therefore, any galaxy with true redshift $z > 3$ that is scattered into the $z_{\mathrm{phot}} < 3$ range will be included. This is important for deriving the correct redshift errors for the photometric redshift selected sample. Note that at $z > 6$, the galaxy light is redshifted longward of 8000 Å and galaxies will only be detected in the four reddest filters and will therefore not be part of the primary lensing catalog. Furthermore, the shapes of these galaxies will not be well measured because of the coarser pixels of the NIR detectors and because they are expected to be extremely faint. We present results for the full sample of galaxies in the mock catalog to specified magnitude limits, together with results after rejecting outliers with large errors, the so-called catastrophic failures. The overall accuracy of the photometric redshifts is often dramatically increased after excluding outliers. We therefore also quote the results after excluding outliers, together with the fraction of these objects. Outliers are defined here as objects with $|(z_{\mathrm{spec}} - z_{\mathrm{phot}})/(1 + z_{\mathrm{spec}})| > 0.3$. Note, however, that in a real situation it is not possible to know which galaxies are

Table 3
Results

Sample	σ_z	σ_z Without Outliers	Fraction Outliers	N arcmin^{-2}
	Standard filter set, $0 < z < 3$			
S/N > 10 in-any-filter	0.135	0.073	0.0361	156
$m_5 < 26.6$	0.096	0.060	0.0183	113
$m_5 < 25.6$	0.053	0.044	0.0032	70
$m_5 < 24.6$	0.041	0.038	0.0017	39
$m_5 < 25.6$, early	0.040	0.039	0.0003	7
$m_5 < 25.6$, late	0.042	0.039	0.0009	38
$m_5 < 25.6$, starburst	0.068	0.051	0.0075	25
	U-band filter set, $0 < z < 3$			
S/N > 10 in-any-filter	0.140	0.071	0.0320	156
$m_5 < 26.6$	0.090	0.055	0.0120	110
$m_5 < 25.6$	0.041	0.037	0.0007	67
$m_5 < 24.6$	0.033	0.032	0.0003	36
$m_5 < 25.6$, early	0.031	0.028	0.0003	6
$m_5 < 25.6$, late	0.041	0.038	0.0005	37
$m_5 < 25.6$, starburst	0.042	0.038	0.0011	24
	U-band filter set, $3 < z < 6$			
S/N > 10 in-any-filter	1.410	0.062	0.480	33
$m_5 < 26.6$	1.273	0.036	0.311	13
$m_5 < 25.6$	0.455	0.026	0.040	2
$m_5 < 24.6$	0.036	0.024	0.006	0.2
	U-band filter set using D95 method, $0 < z < 3$			
S/N > 10 in-any, D95 < 0.40	0.049	0.045	0.0015	103
S/N > 10 in-any, D95 < 0.30	0.042	0.039	0.0005	93
S/N > 10 in-any, D95 < 0.25	0.039	0.036	0.0004	86
$m_5 < 26.6$, D95 < 0.40	0.046	0.043	0.0012	95
$m_5 < 26.6$, D95 < 0.30	0.041	0.039	0.0005	89
$m_5 < 26.6$, D95 < 0.25	0.039	0.036	0.0004	83
	17-band filter set, $0 < z < 3$			
S/N > 10 in-any-filter	0.153	0.061	0.0214	133
$m_5 < 26.6$	0.144	0.052	0.0167	110
$m_5 < 25.6$	0.052	0.033	0.0010	67
$m_5 < 24.6$	0.025	0.025	0.0001	36
$m_5 < 25.6$, early	0.024	0.024	0.0000	6
$m_5 < 25.6$, late	0.039	0.037	0.0007	37
$m_5 < 25.6$, starburst	0.071	0.028	0.0019	24

Notes. $\sigma_z \equiv \mathrm{rms}[(z_{\mathrm{spec}} - z_{\mathrm{phot}})/(1 + z_{\mathrm{spec}})]$. Outliers are defined as objects with $|(z_{\mathrm{spec}} - z_{\mathrm{phot}})/(1 + z_{\mathrm{spec}})| > 0.3$. See Section 5.1 for the definition of D95.

outliers. Later we discuss methods for identifying galaxies with a reasonable probability of being outliers.

In Table 3, we present results using four different magnitude cuts and the three filter sets discussed. The quoted values for σ_z and outlier fraction are subject to a statistical uncertainty depending on the mock catalog sample size. Using simulations we find that the uncertainties in σ_z and the outlier fractions are <0.2% and <0.04%, respectively. The uncertainty in the number densities is $\gtrsim$20%, mainly caused by cosmic variance.

As expected, the accuracy in the photometric redshifts increases when using brighter magnitude cuts. At the same time, this naturally also reduces the number of objects for which photometric redshifts are calculated (Table 3). The increased scatter at fainter magnitude cuts is also evident in Figure 5, where we plot distributions of the photometric minus spectroscopic redshifts for the U-band filter set in different magnitude bins. Going from the brightest magnitude bins (left panels) to successively fainter cuts (right panels) we note an increased scatter. There is also a slight increase in scatter at higher redshifts (comparing top panels with bottom panels), but not as large as the dependence on magnitude. Furthermore, the figure clearly illustrates the non-Gaussian shape of the error distribution of the photometric redshifts.

Comparing the results for the different filter sets in Table 3 reveals that the accuracy increases when including the U-band. However, the most important difference is the clear reduction in the outlier fraction when including the U-band. For example, at $m_5 < 26.6$, the outlier fraction decreases by a third when including the U-band, while at $m_5 < 25.6$ the difference is more than a factor of 4. This can be attributed to the ability of the U-band to better probe the 4000 Å break at low redshifts and the Lyman break at redshifts close to $z \sim 3$. The B-band probes the Lyman break at $z \gtrsim 3$ and is therefore not helpful in the redshift range investigated here.

Since the numbers of detected objects are similar for the standard and the U-band filter sets, there are only advantages in extending the wavelength coverage to the U-band, assuming that the S/N values, in particular for filter #0, are representative. Therefore, we hereafter concentrate on the U-band filter set. The resulting root mean square (rms) for the U-band filter set is in the range $\sigma_z \sim 0.03$–0.14, depending on selection. However, most of the spread, especially at the fainter cut, is mainly due to a few objects enhancing the errors. As can be seen from the table, after excluding a few percent or less of the objects classified as outliers, the rms drops to $\sigma_z \sim 0.03$–0.07. Below, we discuss methods for identifying objects that are outlier candidates.
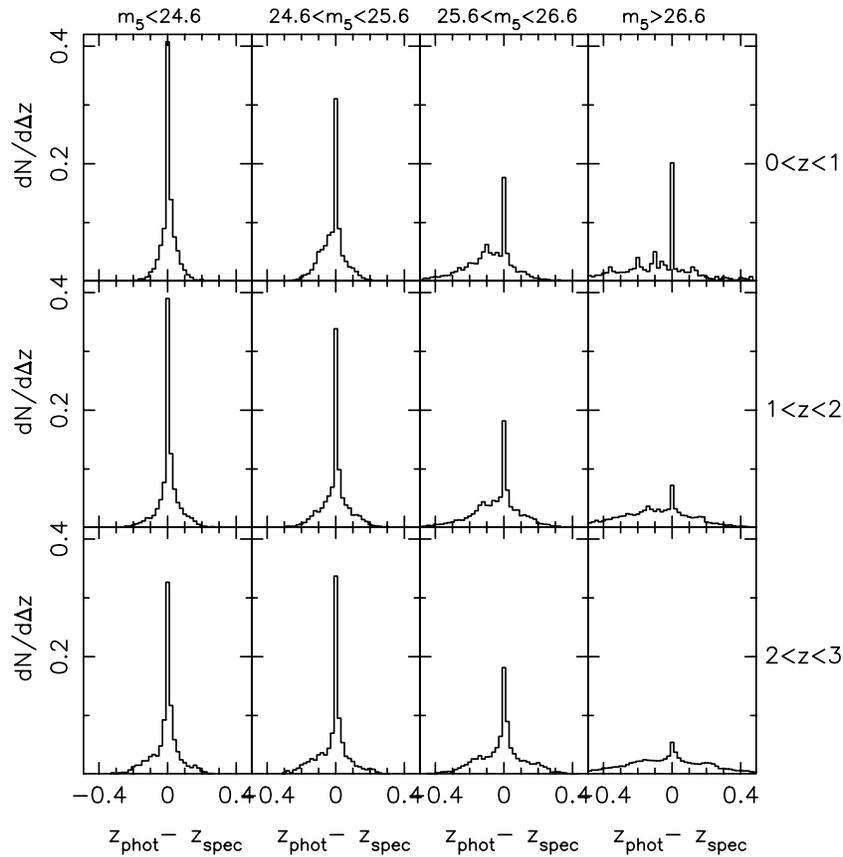
Figure 5. Distributions of photometric minus simulated "spectroscopic" redshifts for different magnitude bins. Going from left to right, the cuts are $m_5 < 24.6$, $24.6 < m_5 < 25.6$, $25.6 < m_5 < 26.6$, and $m_5 > 26.6$ combined with the requirement that S/N > 10 in at least one of the remaining bands. Distributions are normalized to unity.
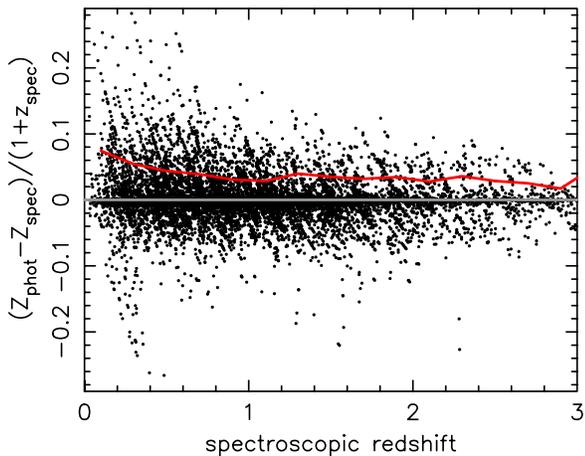
Figure 6. Scatter between photometric redshift and spectroscopic redshift as a function of spectroscopic redshift, using the U-band filter set and $m_5 < 25.6$. The red line shows changes in the rms. A flat change in rms with redshift indicates that photometric redshift errors scale proportional to $(1 + z)$.

(A color version of this figure is available in the online journal.)

We emphasize that the inclusion of the NIR filters is essential for the accuracy of the photometric redshifts and the minimization of the outlier fraction, especially when aiming at redshifts $z > 1$, where the rest frame 4000 Å break moves out of the optical bands. We discuss this further in Section 5.3.4.

Next we investigate in more detail how the accuracy of the photometric redshifts depends on redshift. Figure 6 shows the normalized difference between photometric and spectroscopic redshifts ($[z_{\mathrm{phot}} - z_{\mathrm{spec}}]/[1 + z_{\mathrm{spec}}]$) to redshift $z = 3$. The red line shows changes in rms with redshift. The small variation of the rms with redshift shows that expected redshift errors scale as $\sim(1 + z)$, confirming that using normalized errors gives a fair "redshift-independent" measurement of the scatter. In Figure 7, we divide the sample into different spectral types as given by the best-fitting template SEDs. The figures clearly show that the early-type galaxies have a smaller fraction of galaxies with high scatter. This is also shown in Table 3 where earlier-type galaxies have lower scatter and outlier fractions. We expect this behavior since earlier types have the strongest 4000 Å break, the most important spectral feature for determining photometric redshifts at $z < 3$.

The photometric redshift code returns the best-fitting spectral template for each galaxy, and since we know the input spectral type, we can estimate the accuracy in deriving galaxy spectral types. Figure 8 shows the scatter between the input and derived spectral types at different magnitude cuts, with the spectral types numbered from 1 to 6, according to Figure 1. As expected, there is a better agreement between the input and derived spectral type for brighter magnitudes. Also, the scatter is larger for later-type galaxies, mainly due to the relative similarity between the colors of types 4–6. For the $m_5 < 26.6$ and $m_5 < 24.6$ cuts, we find that respectively 92% and 94% of the galaxies are given a spectral type within ± 1 from the input type. Note that the quantized structure of Figure 8 is caused by the discrete set of 16 template SEDs used in the photometric redshift method.
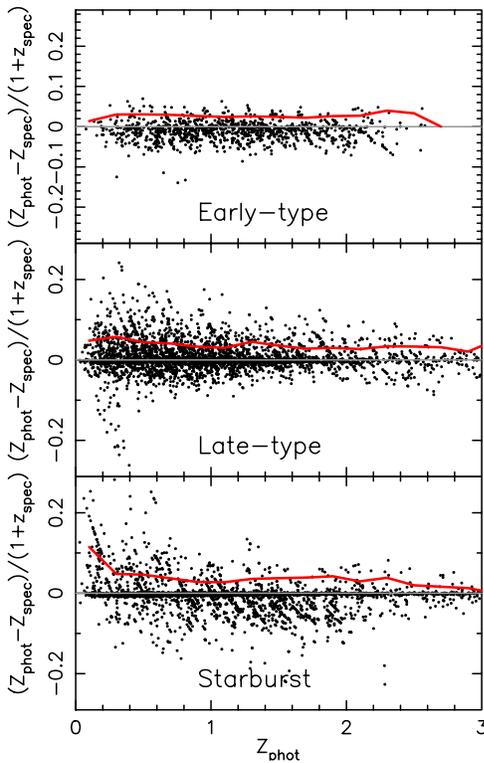
Figure 7. Scatter between photometric redshift and spectroscopic redshift as a function of spectroscopic redshift, using the U-band filter set and $m_5 < 25.6$. The top panel shows results for early-type galaxies, while the middle and bottom panels show late-type galaxies and starbursts, respectively. The red lines show changes in the rms.

(A color version of this figure is available in the online journal.)

4.1. Redshift $3 < z < 6$ Interval

While the SNAP mission is foremost focused on the redshift range $z < 3$, a large number of higher-redshift objects will be detected in any survey reaching depths comparable to those discussed here. In Table 2, we give the predicted number counts to $z = 6$; however, since these counts are derived from extrapolations of lower-redshift LFs, the uncertainty is high. Investigating the photometric redshifts using the U-band filter set, we find a significant increase in the outlier fraction at $z_{phot} > 3$, in particular at faint magnitudes. At $m_5 < 26.6$, we estimate ~30% outliers, decreasing to ~4% and ~1% at $m_5 < 25.6$ and $m_5 < 24.6$ in the redshift range $3 < z < 6$, respectively. Excluding the outliers, the accuracy is comparable to the lower-redshift case. These results are presented in Table 3.

5. DISCUSSION

5.1. Reducing the Fraction of Outliers: D95 Method

As already noted, a few outliers with "catastrophic redshifts" are often responsible for a large part of the estimated scatter between spectroscopic and photometric redshifts. Therefore, it is valuable to have methods for finding and flagging objects that may be outliers, which thereafter could be excluded. A successful method should identify as many outliers as possible, while keeping the total number excluded at a minimum. In Mobasher et al. (2007), we discussed and proved the utility of the so-called D95 method. We define D95 as the width of the 95% confidence interval derived from the photometric redshift
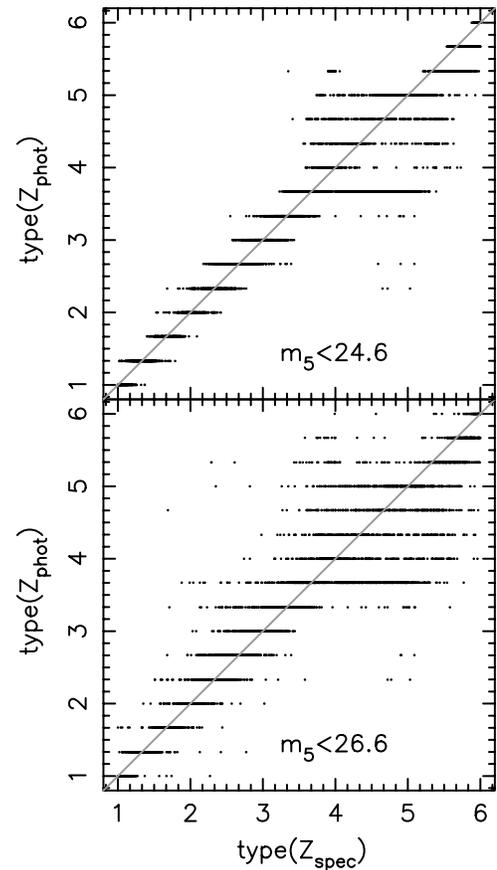
Figure 8. Comparison between input "spectroscopic" spectral type and derived photometric spectral type for two different magnitude limits.

fitting divided by one plus the photometric redshift:

$$D95 = 95\% \text{ confidence interval}/(1 + z_{phot}). \qquad (3)$$

Large D95 values are assigned to galaxies with wide 95% confidence intervals. These could be due to a broad peak in the redshift probability distribution caused by large photometric errors, or they could reflect a double peak in the probability distribution. The latter case happens when there is a confusion between the Lyman break and the 4000 Å break. Therefore, when D95 is large, the uncertainty in the photometric redshift will also be large, increasing the probability that the galaxy is an outlier.

In Table 3, we give results on the photometric redshift accuracy after applying different cuts in D95 for the U-filter set. The results show that it is possible to significantly reduce the fraction of outliers, while only decreasing the total fraction of objects by a small amount. For example, for the "S/N > 10 in-any-filter" selection and a D95 < 0.40 cut, the fraction of outliers decreases by 95%, while the total number of objects decreases by 34%. For the $m_5 < 26.6$ selection, a cut D95 < 0.40 decreases the number of outliers by 90%, while only decreasing the total number of objects by 14%.

We have previously shown that applying a brighter magnitude cut also decreases the number of outliers and, at the same time, decreases the number of galaxies in the sample. Inspecting Table 3 shows that better results are obtained using the D95 method when requiring a particular number density of galaxies. For example, using the "S/N > 10 in-any-filter" selection together with a D95 < 0.40 cut compared to using the $m_5 < 26.6$
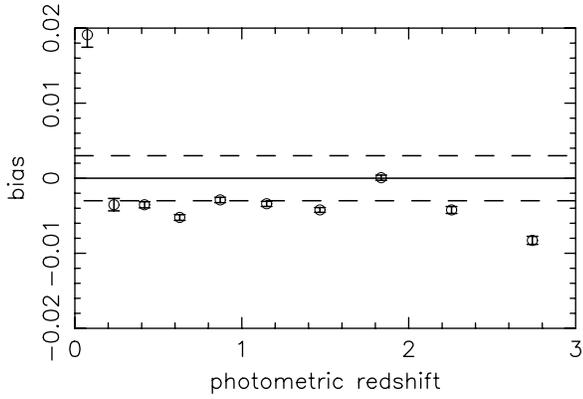
Figure 9. Normalized bias (mean offset between photo-z and spec-z divided by $(1 + z)$) in ten redshift bins with widths $\ln(1 + z) = 0.14$. The sample shown is selected by $m_5 < 26.6$ and D95 < 0.3 and has a galaxy density $\sim$90 arcmin^{-2} to $z < 3$. The dashed lines show the bias $< 0.003(1 + z)$, which is the limit taken from Huterer et al. (2006).

selection without a D95 cut results in a similar number of objects but with better photometric redshifts and significantly fewer outliers for the former selection. It therefore seems more efficient to use D95 as the primary criterion when making a cut in the galaxy sample to improve the photometric redshift accuracy compared to using a magnitude cut.

5.2. Bias

Weak-lensing studies show that a small photometric redshift bias is important for accurate results. The bias is here defined as the mean offset between the photometric redshifts and the true "spectroscopic" redshift. Typically, a bias $< 0.003(1 + z)$ in each of ten redshift bins to $z = 3$ is desirable, assuming that the survey reaches a depth where at least 100 galaxies arcmin^{-2} have determined photometric redshifts (Ma et al. 2006; Huterer et al. 2006).

In Figure 9, we plot the bias in ten redshift bins to $z = 3$ for a magnitude limit $m_5 < 26.6$ and D95 < 0.3, including ~ 90 galaxies arcmin^{-2}. Bin size is chosen so that $\ln(1 + z) =$ constant. The figure shows that the results are near the value discussed in Huterer et al. (2006) in most bins. This is reassuring since the aim of this investigation is not to minimize bias. In a real situation, using a training set of galaxies with spectroscopic redshifts should allow us to minimize the bias. What the figure shows is that the bias offsets may be a problem at the very lowest and highest redshifts, which could be due to low statistics and relatively high outlier fractions.

5.3. Alternative Filter Sets

5.3.1. A 17-filter Set

In our simulations we also include a 17-filter set. The rationale behind this is that with a narrower spacing in wavelength between filters, it should be easier to pick up the location of the redshifted 4000 Å break, and therefore determine the redshift with higher accuracy. The drawback is that with a fixed amount of observing time available, the S/N in each individual filter decreases. In our simulations, we have kept the nominal exposure time in the detection filter, while decreasing the exposure time in the remaining 16 filters by a factor of 2 (consistent with the detector configuration discussed in Section 2).

Results using the 17-filter set are presented in Table 3. Compared to the 9-filter sets, the 17-filter set has a larger scatter except at the brightest magnitudes. However, after excluding outliers, the 17-filter set shows smaller scatter at all magnitudes. Both sets have a comparable number of outliers. Inspecting how the results depend on galaxy type in Table 3 shows that the 17-filter set is more efficient for normal early-type/late-type galaxies, but that the scatter is higher for starburst galaxies This should be a consequence of the weak spectral breaks in the starburst SEDs. Due to this, having additional filters, but with larger photometric errors, will not improve the ability to constrain these weak spectral features and determine the photometric redshift.

In addition to the overall increase in scatter at faint magnitudes in the 17-filter set, the objects will start to drop out of filters in this set earlier than in the 9-filters set (the 17-filter set reaches ~ 0.4 mag less deep with only half the exposure time available per filter). Therefore, instead of having a photometric point, only an upper limit can be used in the photometric redshift fitting. This loss of information contributes to the increased errors. Since both sets do collect the same number of photons, and have the same read noise, one could imagine adding adjacent filters in the 17-filter set to recover a 9-filter set with the same S/N in each filter as the original 9-filter set. From this one could expect the same photometric redshift accuracy for both sets. However, there are differences between the sets, in particular the "effective" filter transmission function for the added filters will be broader than the original filters, which will make them less efficient at locating the spectral breaks. Also, the overall filter shape will be different, with a more "peaked" shape for the added filters.

We therefore conclude that the 9-filter set is preferred except at the brightest magnitudes, but with a method that efficiently excludes outliers, the 17-filter set should be comparable to, or even better, than the 9-filter set.

5.3.2. Less Wide Broadbands

The filters used so far are fairly broad with significant overlap between them (see Figure 2). As an alternative, we also construct a set that has the same effective wavelengths as the U-band filter set, but with filter widths being only 75% of the original. This is approximately equivalent of changing the resolution from ~ 4 to ~ 5.3 (numbers are somewhat filter dependent). The reason for testing this alternative filter set is to investigate whether the increased resolution will make it easier to locate the redshifted spectral breaks in the galaxies' SEDs and therefore decrease the scatter in the photometric redshifts. Our results show that there is no gain in the accuracy due to the resolution, but instead, the scatter in the photometric redshifts increases due to the larger errors caused by lower counts in the narrower filters. At the faintest limits ($m_5 < 26.6$), both the scatter and fraction of outliers are twice that of the full width broadband filter set. At bright magnitudes ($m_5 < 25$), both sets give comparable results. Also, with the narrower filters, the number of objects with S/N > 10 decreases by $\sim 14\%$.

5.3.3. Square Filters

Finally we include a filter set with square transmission functions. These filters have, by construction, the same area (i.e., integral of transmission over wavelength) as the U-band filter set and are centered on the effective wavelengths of those filters. This leads to a filter set with similar resolution compared to the U-band set. The resulting photometric redshifts
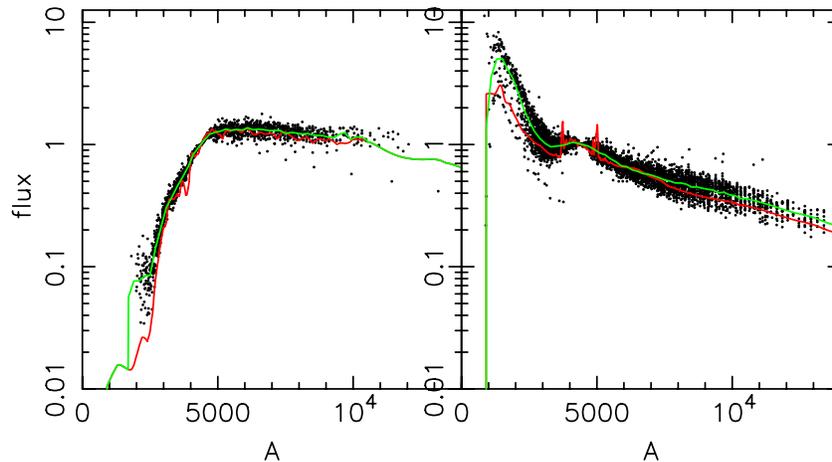
Figure 10. Correcting template SEDs using spectroscopic sample. For each galaxy, the observed flux is normalized to rest frame 4400 Å. The red lines show the input template SED, while the green lines show the corrected SED after fitting to the data. The left panel shows an early-type galaxy, while the right panel shows an Im-type galaxy.

show three differences compared to the U-band filter set. First, the scatter (including outliers) increases for all magnitude selections. Second, the fraction of outliers increases by a factor of $\sim$2 at the faintest selections, while being similar at brighter magnitudes ($m_5 < 25.6$). Finally, the scatter in the photometric redshifts decreases after excluding the outliers (down $\sim$ 15–30%) at all selections. Overall, this means that the photometric redshifts for most objects are somewhat improved, but for a few objects there is a large degradation in the accuracy.

The increase in the outlier fraction at faint magnitudes when using the square filters can be attributed to the sharp gradients in K-corrections (i.e., colors) with redshift as spectral features such as breaks and emission lines promptly move in and out of filters. In particular at faint magnitudes with large photometric errors, the photometric redshift code may misidentify these features and therefore assign a "catastrophic" redshift.

We conclude that for the best overall accuracy, the standard filter set is preferred compared to the square filter set; however, the difference is quite marginal.

5.3.4. Importance of Near-IR Filters

It is well established that the inclusion of NIR filters is important for determining accurate photometric redshifts (e.g., Hogg et al. 1998; Rudnick et al. 2001). The reason for this is that at $z > 1$ the 4000 Å break moves longward of the optical filters and there is no sharp spectral break at observed optical wavelengths from which a photometric redshift can be accurately derived. At $z > 3$, the Lyman break moves into the optical filter and once again there is a sharp feature to detect. But to detect a break at $1 \lesssim z \lesssim 3$, NIR filters are required.

We have tested the importance of the NIR filters on our mock galaxy catalog by deriving photometric redshift using all nine filters as well as using only the six optical filters. As expected, the number of outliers significantly increases after excluding the NIR filters. At $m_5 < 26.6$, we find an increase by a factor of $\sim$10 in the number of outliers as well as a $\sim$50% increase in scatter even after excluding outliers. Furthermore, even if we replace the NIR detectors with optical detectors, effectively doubling the exposure time in each optical filter, there is still a factor of $\gtrsim$5 more outliers when only having the optical filters. In the latter case, there is also a $\sim$33% increase in scatter after excluding outliers. We therefore conclude that

NIR photometry is highly important for the photometric redshift accuracy.

5.4. Investigating an Alternative Set of SEDs

So far in this investigation, we have used the same set of template SEDs when creating both the mock galaxy catalog as we did when deriving photometric redshifts. This allows us to concentrate on how the photometric redshift accuracy depends on S/N and filter choices without adding biases that can be introduced if different galaxy sets are used for creating catalogs and deriving redshifts. However, the real case will be different from this investigation in the sense that it will not be known *a priori* if the template set used for calculating the photometric redshifts represents the actual distribution of SEDs for the real galaxies.

To investigate how well we can derive photometric redshifts in a situation where we do not know the shape of the observed galaxies' SEDs, we made a new set of simulations with the intent of calibrating our template SEDs using spectroscopic redshifts, an approach shown successfully by Ilbert et al. (2006). We create an alternative mock galaxy catalog using a different set of galaxy template SEDs. This second set of template SEDs is based on the PEGASE galaxy models (Fioc & Rocca-Volmerange 1997) and includes six templates from elliptical to starburst. We also added systematic offsets to the galaxy magnitudes (of the order of a few 0.01 mag) to mimic zero-point calibration uncertainties and dust extinction. For a subset "training sample" of 10^4 galaxies, the true spectroscopic redshift was given to test and calibrate the photometric redshifts. Photometric redshifts were thereafter derived using the first set of template SEDs and compared to the spectroscopic sample. The first run produced a fairly large scatter in the redshifts due to both the zero-point offsets and SED mismatches. To decrease scatter we use two approaches. First, we add offsets to the catalog magnitudes and rerun the photometric redshift code minimizing the scatter between the photometric redshifts and the spectroscopic redshifts in the training sample. Second, a new set of modified template SEDs was created. To make this, we first divided the spectroscopic sample into six types using the template set from the photometric redshift code (without corrections for dust extinction). For each type, we thereafter plotted the flux of all objects at the rest-frame wavelength of each filter normalized to 4400 Å. Figure 10 shows the case for early-type galaxies (left) and

Im-type galaxies (right). The red lines show the original template SED, while the green lines show a fit to the data. We adopt these fits as a new set of corrected template SEDs. We finally recalculate the photometric redshifts using both zero-point corrections and corrected template SEDs. The results are consistent with the results given in Table 3 both in terms of scatter and fraction of outliers, assuring us that the size of the scatter presented here will not dramatically change even though the SEDs of the actual galaxies observed are not the same as the assumed set of template SEDs. Note, however, that the accuracy of the photometric redshifts will depend on the spectroscopic sample and the diversity of the SED population. The more diverse the SEDs of the true galaxy population are, the larger the number of spectroscopic redshifts needed. This will be further investigated by S. Jouvel et al. (2008, in preparation).

6. CONCLUSIONS AND SUMMARY

We have used simulations to investigate how the behavior of photometric redshifts for a SNAP-like mission depends on, e.g., magnitude limit, filter choices, wavelength coverage, and galaxy types. We have also discussed methods for decreasing the expected fraction of outliers, i.e., galaxies with significant disagreement between the photometric and spectroscopic redshifts. We note that our investigation is not primarily focused on getting exact numbers for, e.g., the photometric redshift accuracy and outlier fractions, but to investigate how these diagnostics are affected when changing, e.g., S/N and filter sets. Our main conclusions are as follows.

1. We find that including the U-band significantly decreases the fraction of outliers and results in an increase in the photometric redshift accuracy.
2. A 17-filter set results in larger scatter compared to a 9-filter set except at bright magnitudes. However, after excluding outliers, the 17-filter set gives more accurate redshifts at all magnitudes.
3. A 9-filter set with narrower filter functions (resolution ∼5 instead of ∼4) results in an increase in the scatter of the photometric redshifts. This is caused by the larger photometric errors when fewer photons are detected.
4. Using a filter with a square transmission curve decreases the scatter in the photometric redshifts for the majority of the objects. At the same time, however, the fraction of outliers is doubled. Therefore, if possible outliers could be efficiently flagged, the square filter set would be preferred.
5. The accuracy of the photometric redshifts depends on both magnitude (or efficiently the S/N) and galaxy spectral type, with better results at high S/N and for earlier-type galaxies.
6. Using the D95 method can significantly decrease the number of outliers, while only decreasing the total number

of objects moderately. Using this method is therefore preferred compared to using only the S/N as a cut to decrease scatter.

REFERENCES

Aldering, G., et al. 2004, (arXiv:astro-ph/0405232)
Beckwith, S. V. W., et al. 2006, AJ, 132, 1729
Bouwens, R. J., Illingworth, G. D., Thompson, R. I., & Franx, M. 2005, ApJ, 624, L5
Broadhurst, T., et al. 2005, ApJ, 621, 53
Bundy, K., Ellis, R. S., & Conselice, C. J. 2005, ApJ, 625, 621
Calzetti, D., Armus, L., Bohlin, R. C., Kinney, A. L., Koornneef, J., & Storchi-Bergmann, T. 2000, ApJ, 533, 682
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Coleman, G. D., Wu, C.-C., & Weedman, D. W. 1980, ApJS, 43, 393
Dahlen, T., Mobasher, B., Dickinson, M., Ferguson, H. C., Giavalisco, M., Kretchmer, C., & Ravindranath, S. 2007, ApJ, 654, 1727
Dahlen, T., Mobasher, B., Somerville, R. S., Moustakas, L.A., Dickinson, M., Ferguson, H. C., & Giavalisco, M. 2005, ApJ, 631, 126
Dahlen, T., Strolger, L.-G., & Riess, A. G. 2008, ApJ, 681, 462
Dahlen, T., et al. 2004, ApJ, 613, 189
Davis, T. M., Schmidt, B. P., & Kim, A. G. 2006, PASP, 118, 205
Fioc, M., & Rocca-Volmerange, B. 1997, A&A, 326, 950
Giavalisco, M., et al. 2004, ApJ, 600, L93
Grogin, N. A., et al. 2005, ApJ, 627, L97
Gwyn, S. 1995, Master's thesis, Univ. Victoria
Hogg, D. W., et al. 1998, AJ, 115, 1418
Huterer, D., Takada, M., Bernstein, G., & Jain, B. 2006, MNRAS, 366, 101
Ilbert, O., et al. 2006, A&A, 457, 841
Kinney, A. L., Calzetti, D., Bohlin, R. C., McQuade, K., Storchi-Bergmann, T., & Schmitt, H. R. 1996, ApJ, 467, 38
Kneib, J.-P., Ellis, R. S., Santos, M. R., & Richard, J. 2004, ApJ, 607, 697
Kurtz, M. J., Geller, M. J., Fabricant, D. G., Wyatt, W. F., & Dell'Antonio, I. P. 2007, AJ, 134, 1360
Limousin, M., et al. 2007, ApJ, 668, 643
Ma, Z., Hu, W., & Huterer, D. 2006, ApJ, 636, 21
Massey, R., et al. 2007a, ApJS, 172, 239
Massey, R., et al. 2007b, Nature, 445, 286
McCracken, H. J., et al. 2007, ApJS, 172, 314
Mobasher, B., Rowan-Robinson, M., Georgakakis, A., & Eaton, N. 1996, MNRAS, 282, L7
Mobasher, B., et al. 2007, ApJS, 172, 117
Richard, J., et al. 2007, ApJ, 662, 781
Riess, A. G., et al. 2004, ApJ, 607, 665
Riess, A. G., et al. 2007, ApJ, 659, 98
Rudnick, G., et al. 2001, AJ, 122, 2205
Scoville, N., et al. 2007, ApJ, in press
Smith, G. P., Kneib, J.-P., Smail, I., Mazzotta, P., Ebeling, H., & Czoske, O. 2005, MNRAS, 359, 417
Strolger, L.-G., et al. 2004, ApJ, 613, 200
Wiklind, T., Dickinson, M., Ferguson, H. C., Giavalisco, M., Mobasher, B., Grogin, N. A., & Panagia, N. 2008, ApJ, 676, 781
Wray, J. J., & Gunn, J. E. 2008, ApJ, 678, 144