The Astrophysical Journal, 671: L000–L000, 2007 December 20

PHOTOMETRIC REDSHIFTS AND PHOTOMETRY ERRORS

D. Wittman, P. Riechers, and V. E. Margoniner[1]
Physics Department, University of California, Davis, CA 95616; dwittman@physics.ucdavis.edu
Received 2007 September 20; accepted 2007 October 31; published 2007 November 16

ABSTRACT

We examine the impact of non-Gaussian photometry errors on photometric redshift performance. We find that they greatly increase the scatter, but this can be mitigated to some extent by incorporating the correct noise model into the photometric redshift estimation process. However, the remaining scatter is still equivalent to that of a much shallower survey with Gaussian photometry errors. We also estimate the impact of non-Gaussian errors on the spectroscopic sample size required to verify the photometric redshift rms scatter to a given precision. Even with Gaussian *photometry* errors, photometric redshift errors are sufficiently non-Gaussian to require an order of magnitude larger sample than simple Gaussian statistics would indicate. The requirements increase from this baseline if non-Gaussian photometry errors are included. Again the impact can be mitigated by incorporating the correct noise model, but only to the equivalent of a survey with much larger Gaussian photometry errors. However, these requirements may well be overestimates because they are based on a need to know the rms, which is particularly sensitive to tails. Other parameterizations of the distribution may require smaller samples.

Subject headings: galaxies: photometry — methods: statistical — surveys

Online material: color figures

1. INTRODUCTION

Photometric redshifts (Connolly et al. 1995; Hogg et al. 1998; Benítez 2000) are of increasing importance in observational tests of cosmology. Predicting photometric redshift performance has therefore become an important part of planning large optical surveys. There are two distinct aspects of performance to consider. First, there are straightforward goals of accuracy and precision. Second, to control systematic errors in the downstream science, one must be able to *know*, in some cases rather stringently, the accuracy and precision of the photometric redshifts in the actual survey (Ma et al. 2006; Huterer et al. 2006). Knowing the actual photometric redshift precision can be more important than maximizing the precision. For example, cosmic shear tomography calls for relatively wide redshift bins ($dz \sim 0.2$). Leakage between bins, to the extent that it is known, can be precisely incorporated into comparisons between models and data. This by itself is not very demanding in terms of photometric redshift precision. However, in a large survey with very small statistical errors, the leakage must be known very precisely to avoid nontrivial systematic errors. Ma et al. (2006) estimate that for cosmic shear tomography with next-generation surveys, the bias and rms scatter in each redshift bin must be known to ~0.003 to avoid degrading the shot-noise-limited constraints on dark energy.

To first order, photometric redshift performance depends on filter set, signal-to-noise ratio (S/N), and the desired range of redshifts and galaxy types. Here we wish to call attention to an often overlooked aspect: photometry errors. Photometric redshift simulations and real-life implementations typically assume Gaussian photometry errors. Real data are more complicated. As one anecdote, Cameron & Driver (2007) note that in one catalog of 42 galaxies with both photometric and spectroscopic redshifts, there were six outliers, all of which had questionable photometry due to saturation, neighbors, or multiple nuclei. In this Letter we show that knowing the true distribution of errors is important for optimizing photometric redshift precision. We also discuss how that in turn affects the size of the spectroscopic sample required to characterize the photometric redshift errors in a survey.

2. METHODS

We conduct four sets of simulations built around the following basic setup. We use the Bayesian Photometric Redshift (BPZ; Benítez 2000) code, which uses a set of template galaxy spectral energy distributions (SEDs) and a set of priors to help break degeneracies in color space. We chose the six SED templates and the HDF-N prior detailed in Benítez (2000). BPZ is representative of one of two types of methods in the photometric redshift community. We discuss possible impacts on the other type, training-set methods, in § 5. The choice of filter set is not important for this demonstration. We use the same filter set (F300W, F450W, F606W, F814W, *J*, *H*, *K*) used for the Hubble Deep Field–North (HDF-N) photometric redshifts discussed in Hogg et al. (1998), Benítez (2000), and Fernández-Soto et al. (1999, 2001).

Each simulation generates a synthetic catalog of 6000 galaxies evenly spread throughout the F814W magnitude range 20–26. This and other aspects of the simulations are not realistic, but are adopted to facilitate analysis by covering parameter space evenly. The results presented here therefore do not apply quantitatively to any real survey, but they demonstrate the issues. The simulator uses each galaxy's magnitude to choose a random type and redshift following the distributions described by the priors. The functional form of the priors is (adapting Benítez 2000 eqs. [22]–[24])

$$p(t|m) = f_t \exp\,[-k_t(m-20)],$$

$$p(z|t,\,m) \propto z^{\alpha_t} \exp\left\{-\left[\frac{z}{z_{0t}+k_{mt}(m-20)}\right]^{\alpha_t}\right\},$$

where t is the type, f_t is the fraction of bright galaxies consisting of that type, k_t describes the type mix evolution with magnitude,

[1] Current address: Physics Department, California State University, Sacramento, CA 95819.

TABLE 1
SIMULATION PARAMETERS

NAME	INPUT NOISE PARAMETERS			
	σ (%)	A	B	$\delta z_{\rm rms}$
SIM1	5	0	...	0.026
SIM2	10	0	...	0.070
SIM3	5	0.1 (0[a])	0.15	0.092
SIM4	5	0.1	0.15	0.072
SDSS-like	5	0.1	0.06	0.031

[a] In SIM3, the input noise was purposely modeled incorrectly, with $A = 0$.

α_t parameterizes the high-redshift cutoff for that type, and the denominator inside the square brackets corresponds roughly to the median redshift for that type as a function of magnitude. The values of the parameters are given in Table 1 of Benítez (2000), except that the values of k_t for ellipticals and spirals were inadvertently switched in that table (Margoniner et al. 2007). As in Benítez (2000), for this purpose spirals are considered a single type ($t = 2$), and their numbers are evenly split between the Sbc and Scd SEDs. Similarly, galaxies chosen to be "irregular" ($t = 3$) are randomly assigned (with equal probability) the Im, SB2, or SB3 SED.

Having assigned a redshift and SED template, the galaxy is then assigned "observed" fluxes in the seven filters by adding noise (the character of which varies with the simulation) to the synthetic observer-frame fluxes of that template at that redshift:

$$f_{{\rm obs},i} = f_{{\rm templ},i} + \delta f,$$

where i indexes the filters. An unrealistic aspect of the noise in all simulations is that it is a fixed percentage of the model flux. That is, every galaxy is observed at the same S/N, regardless of magnitude, redshift, or filter. This is another analysis convenience. The effect of varying S/N was explored in one specific case by Margoniner & Wittman (2007), and will have to be customized to each survey.

We then run the catalogs through BPZ, with the HDF-N prior turned on, and analyze the performance in terms of $\delta z \equiv (z_{\rm phot} - z_{\rm spec})/(1 + z_{\rm spec})$, specifically the bias δz and the scatter $\delta z_{\rm rms}$.

3. REALIZATIONS

As baselines, we do two simulations with Gaussian noise: SIM1 with 5% noise (S/N = 20) and SIM2 with 10% noise (S/N = 10). Table 1 lists input parameters and summary output statistics for the different simulation sets. These photometry error distributions are shown in Figure 1. The resulting δz distributions are shown in Figure 2. In both cases, the bias is small (0.003 or less in absolute value) and not inconsistent with zero. The scatter depends strongly on S/N: $\delta z_{\rm rms} = 0.026$ for S/N of 20, increasing to 0.070 for S/N of 10. We also did a run with S/N = 100, not shown in the figures: $\delta z_{\rm rms} = 0.004$. This is extremely tight because the quoted S/N is achieved in *each* band for *each* galaxy.

Next, we add non-Gaussian tails to the photometry error distribution. We adopt a functional form

$$p(\delta f) = \frac{1}{\sigma\sqrt{2\pi} + AB}\left\{\exp\left[-\frac{(\delta f)^2}{2\sigma^2}\right] + A\exp\left(-\frac{|\delta f|}{B}\right)\right\},$$

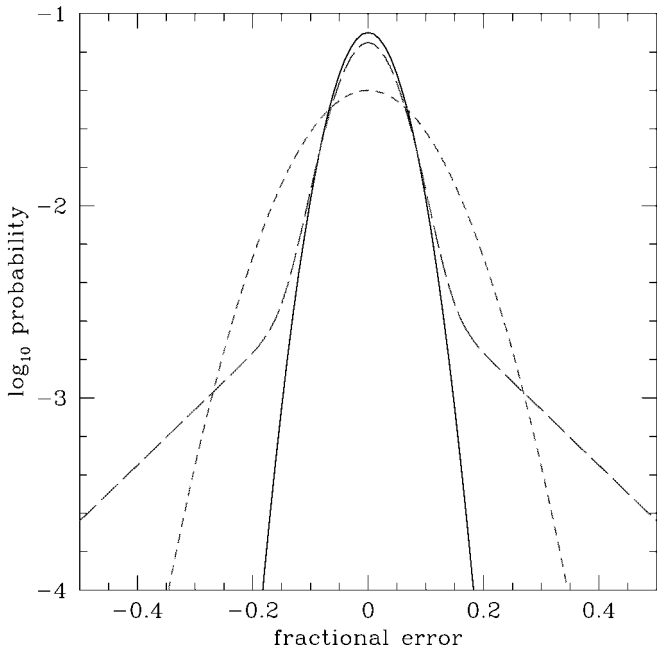

FIG. 1.—Photometry error distributions. SIM1, 5% Gaussian (*solid curve*); SIM2, 10% Gaussian (*short-dashed curve*); SIM3 and SIM4, 5% Gaussian with exponential tails (*long-dashed curve*). [*See the electronic edition of the Journal for a color version of this figure.*]

where δf is the flux error, σ describes the width of the Gaussian core, and the parameters A and B describe the tails. For a given σ, the fraction of galaxies in the tails is sensitive to changes in the product AB but relatively insensitive to changes in A and B as long as the product is held constant. There is little published data on realistic values of A and B. Margoniner & Wittman (2007) briefly describe photometry simulations in which synthetic galaxies are added to real images from the Deep Lens Survey (DLS; Wittman et al. 2002). We roughly

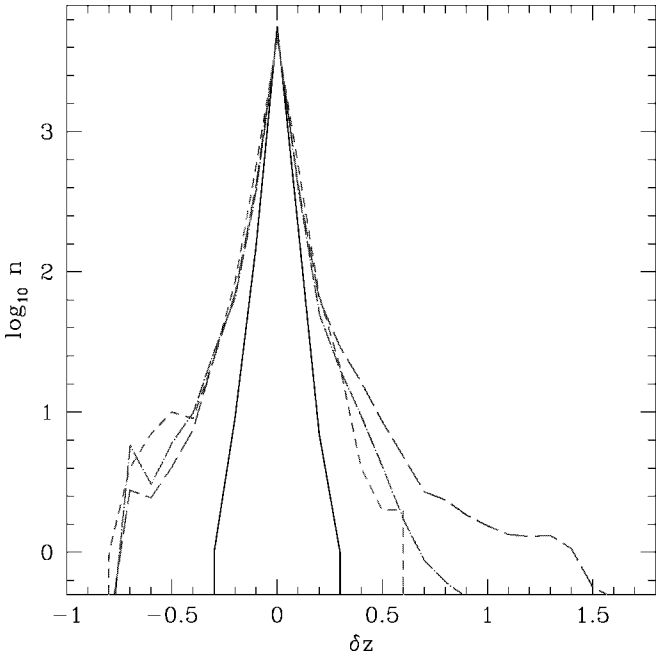

FIG. 2.—Distributions of δz. Colors and line types are as in Fig. 1, with the addition of SIM4 (*dot-dashed curve*), which uses the non-Gaussian noise model in the photometric redshift estimation. [*See the electronic edition of the Journal for a color version of this figure.*]

match the fraction of objects in that tail, but with two symmetric tails and $\sigma = 0.05$ as in SIM1, by setting $A = 0.1$ and $B = 0.15$ or 3 σ. For this choice of A and B, used in SIM3 and SIM4 and shown as the long-dashed curve in Figure 1, the tails begin to dominate over the Gaussian core at 2.51 times the rms of the Gaussian core, and 9.4% of the galaxies are "in" the tails, compared to 1.2% falling outside 2.51 σ for a pure Gaussian. The rms of the distribution is 0.103, very close to that of SIM2.

As a comparison, the photometry error distribution for bright, unresolved objects in the Sloan Digital Sky Survey (SDSS) is published in Figure 3 of Ivezić et al. (2003), who state that 0.9% of objects lie outside of ± 3 σ (where $\sigma = 0.02$), versus 0.3% for a pure Gaussian. This observation, and the figure, are reasonably approximated by $A = 0.1$ and $B = 0.0235$ or 1.2 σ. These tails are much smaller than used in SIM3 and SIM4, which have 7.3% of their galaxies outside ± 3 σ. However, the available SDSS data are for *bright* ($g < 20.5$) *point sources*. Photometry is notably more difficult for extended sources and for faint sources. In the DLS simulations, A is consistent with zero for bright ($20 < R < 22$) galaxies, and grows steadily with magnitude. Of course, most of the galaxies in a deep survey are at the faint end. Therefore, while noting the near-Gaussianity of the SDSS bright point-source photometry, we believe that heavier tails are currently more appropriate for faint galaxies in deep ground-based surveys.

We attribute the Gaussian cores of these distributions to photon statistics, which is the nominal error reported by most photometry packages, and the tails to other effects such as crowding (photometry biased by the presence of neighboring galaxies). For space-based data, crowding is less important, but there may be different sources of non-Gaussianity, as Figure 10 of Benítez et al. (2004) certainly appears to be non-Gaussian. The tails in this Letter are meant to emulate ground-based surveys as described above. We quantify their impact by estimating redshifts in SIM3 using the nominal Gaussian photometry error as input to BPZ. Averaged over 100 realizations, δz remained small (0.0038), but $\delta z_{\rm rms}$ increased to 0.092. The distribution is shown as the long-dashed curve in Figure 2.

Clearly, these tails are very harmful. Adding them to the S/N = 20 distribution more than doubled $\delta z_{\rm rms}$. In fact, *doubling* the Gaussian photometry noise had less impact on $\delta z_{\rm rms}$ than did adding these tails. Surveys will have to control the tails of their photometry error distributions if they are to reach the photometric redshift performance expected based on their filter set and S/N. Modern surveys do recognize this and work to reduce the tails, but tails will always be present at some level. Legacy surveys may have non-Gaussian errors frozen into their data, and new surveys will find it expensive to eliminate all non-Gaussian sources of error. Therefore, we investigate the extent to which knowledge of these errors can render them less damaging to photometric redshifts.

4. LIVING WITH NON-GAUSSIAN ERRORS

Accounting for these errors is straightforward. In the BPZ code, the probability of observing colors C given a model SED type T and redshift z, $p(C|T, z)$, is simply a Gaussian of width set by the nominal photometry errors for that galaxy. In SIM4, we use the same input photometry as in SIM3 but replace that noise model with the full heavy-tailed distribution used to the generate the catalog. The resulting δz distribution is shown in Figure 2 as the dot-dashed curve. The outliers in δz which appeared in SIM3 have now largely disappeared, and $\delta z_{\rm rms}$ is down to 0.072. This is comparable to $\delta z_{\rm rms}$ in SIM2, which had twice the simulated sky noise, but no tails.

The scatter in δz increases to 0.082 if one uses the unmodified BPZ code assuming Gaussian errors, but with an rms of 0.1 instead of 0.05, to roughly approximate the wider distribution of photometry errors. As another comparison case for incorrect noise models, we estimated redshifts from a SIM2 realization using the SIM1 noise model. In this case, $\delta z_{\rm rms}$ changed by only 0.003, which was not quite significant given the sample size. Thus, it appears that if the photometry errors are Gaussian, knowing the width of that Gaussian is not very critical. We see from Figure 2 that it is the 1 in ~500 outlier that is responsible for the poor performance of SIM3. SIM2 lacks extreme outliers, so qualitatively, its better performance makes sense despite its broader core. Yet this degree of insensitivity to the Gaussian width is somewhat surprising.

For comparison, we perform a version of SIM4 in which the tails are much less prominent, as in the SDSS bright point-source photometry: $\sigma = 0.05$, $A = 0.1$, and $B = 0.06$ (1.2 σ). We find that $\delta z_{\rm rms} = 0.031$, with the noise model affecting only the fourth decimal place. The photometry tails are apparently small enough that including them in the noise model is not very helpful, but overall performance is still significantly worse than with no tails at all. (SIM1 had $\delta z_{\rm rms} = 0.026$, while the variation from realization to realization is ~0.001 and these numbers are quoted after averaging over 100 realizations.) This indicates that even small photometry tails can have a significant impact on photometric redshift performance.

5. DISCUSSION

It is not surprising that tails in the photometry error distribution can cause outliers in the δz distribution. However, a number of points are worth remarking:

1. Adding heavy tails (comprising ≤10% of the galaxies) caused more increase in $\delta z_{\rm rms}$ than did *doubling* the Gaussian photometry error. In other words, the photometric redshift performance of a survey with large tails could be worse than that of a survey with *half* the S/N but with no tails. Surveys should therefore pay close attention to reducing the tails of the color errors. This is not the same as reducing the tails of the flux errors. As an extreme example, if an equal fraction of light is lost in all filters, the colors are unaffected.
2. Assuming that non-Gaussian errors can never be entirely eliminated, the effect of the tails on photometric redshift performance can be mitigated by including an accurate noise model in the photometric redshift process. This will in turn require extensive Monte Carlo simulations which include all important sources of non-Gaussian errors, such as crowding and complex galaxy morphology. In addition, the importance of the tails is likely to vary with magnitude, seeing, etc.
3. No clear rule is evident for required accuracy of the noise model. Photometric redshift precision was not significantly affected when errors and model were both Gaussian but the rms was wrong by a factor of 2. When errors were heavy-tailed, approximating them with a Gaussian of the same rms won back about half of the precision that could be won back with the fully correct noise model.
4. Even very small tails have a measurable impact on $\delta z_{\rm rms}$, but in this case the noise model made no measurable difference.

The tails also have a disproportionate impact on the problem of knowing $\delta z_{\rm rms}$ precisely for each redshift bin, whereas pre-

cision on δz did not suffer substantially. If the δz distribution is Gaussian, the spectroscopic sample size required to calibrate $\delta z_{\rm rms}$ to a desired accuracy $\sigma_{\rm cal}$ is $\sim(\delta z_{\rm rms})^2/2\sigma^2$ (this of course assumes that the spectroscopic sample is representative of the photometric sample). For $\sigma_{\rm cal} = 0.003$ and a class of sources with $\delta z_{\rm rms} = 0.026$ as in SIM1, only ~40 galaxies per redshift bin would be required. However, bootstrap resampling of SIM1 shows that 7 times more galaxies are required to know $\delta z_{\rm rms}$ to the same accuracy, due to its non-Gaussian tails (which stem from the properties of galaxies in color space, not from the photometry). For SIM2, the factor is 13, presumably because the greater noise in SIM2, although still Gaussian, allows more near-degeneracies in color space to come into play. For SIM3 with its heavy photometry tails, the factor is ~50. However, this can be much reduced simply by incorporating the correct noise model into the photometric redshift estimation. SIM4 requires "only" ~25 times as many galaxies as the Gaussian prediction would suggest, and the Gaussian prediction is itself ~2 times smaller than for SIM4, because of the smaller $\delta z_{\rm rms}$. Of course, it would be preferable to reduce non-Gaussian tails in the underlying photometry as much as possible, as dramatically illustrated by the large remaining differences between SIM4 and either SIM1 or the simulation with SDSS-like tails.

We caution that this procedure may substantially overestimate spectroscopic sample requirements. They are based on the Gaussian model of photometric redshift errors employed by Ma et al. (2006), who derived a prescription for precision of our knowledge of $\delta z_{\rm rms}$. But the rms of a distribution is driven by its tails, so that the tails seem to be all-important here. If the photometric redshift error model used in the cosmological parameter estimation were modeled differently, the tails could assume a more proportional influence, and fewer spectroscopic redshifts would be required to characterize their effect. Mandelbaum et al. (2007) discuss some related aspects in the context of galaxy-galaxy lensing.

The applicability of this work to training-set methods depends on the details of the method. An advantage of training-set methods is that they may "learn" the correct noise model automatically, and therefore should not require any modification to reach optimum performance (which is presumably still much reduced compared to the no-tails case). But for this to happen, the training set must be sufficiently large to encompass the non-Gaussian features of the photometry. This may require a rather larger training set than would otherwise be required, and it also requires a training set that is not cleaner than the full data set. However, it may be possible to build a hybrid approach in which detailed knowledge of photometry error distributions from large sets of Monte Carlos is combined with a modest spectroscopic sample to train the algorithm.

Non-Gaussian photometry errors may not be a substantial source of catastrophic outliers in current surveys. The SIM3/SIM4 tails may be unrealistically heavy, as there is scant published data on the size of the non-Gaussian tails for faint galaxy photometry. Furthermore, catastrophic outliers exist even with purely Gaussian photometry errors, due to color-space degeneracies. However, real-world experience such as that of Cameron & Driver (2007) and, in a different context, Bolton et al. (2004), suggests that non-Gaussian errors are often not negligible. Color-space degeneracies are usually *near*-degeneracies, and galaxies become much more likely to scatter across a near-degeneracy if the photometry has non-Gaussian tails.

Our example started from an unrealistically good baseline of S/N = 20 in each of seven filters and $\delta z_{\rm rms} = 0.026$, so the effect of the tails was particularly dramatic. Surveys starting from a more realistic performance baseline will not see such a large fractional increase in scatter, but may still see the effect of tails in the overall error budget. Limiting the tails of the photometry error distribution and using an accurate error model will reduce photometric redshift scatter and greatly reduce the size of the spectroscopic sample required to calibrate the scatter.

REFERENCES

Benítez, N. 2000, ApJ, 536, 571
Benítez, N., et al. 2004, ApJS, 150, 1
Bolton, A. S., Burles, S., Schlegel, D. J., Eisenstein, D. J., & Brinkmann, J. 2004, AJ, 127, 1860
Cameron, E., & Driver, S. P. 2007, MNRAS, 377, 523
Connolly, A. J., Csabai, I., Szalay, A. S., Koo, D. C., Kron, R. G., & Munn, J. A. 1995, AJ, 110, 2655
Fernández-Soto, A., Lanzetta, K. M., Chen, H.-W., Pascarelle, S. M., & Yahata, N. 2001, ApJS, 135, 41
Fernández-Soto, A., Lanzetta, K. M., & Yahil, A. 1999, ApJ, 513, 34
Hogg, D. W., et al. 1998, AJ, 115, 1418
Huterer, D., Takada, M., Bernstein, G., & Jain, B. 2006, MNRAS, 366, 101
Ivezić, Ž., et al. 2003, Mem. Soc. Astron. Italiana, 74, 978
Ma, Z., Hu, W., & Huterer, D. 2006, ApJ, 636, 21
Mandelbaum, R., et al. 2007, MNRAS, submitted (arXiv:0709.1692)
Margoniner, V. E., & Wittman, D. M. 2007, ApJ, submitted (arXiv:0707.2403)
Wittman, D. M., et al. 2002, Proc. SPIE, 4836, 73