Title

Cosmology requirements on supernova photometric redshift systematics for the Rubin LSST and Roman Space Telescope

Journal

Physical Review D, 103(2)

ISSN

2470-0010

Authors

Mitra, Ayan
Linder, Eric V

Publication Date

2021-01-15

DOI

10.1103/physrevd.103.023524

Peer reviewed

Cosmology requirements on supernova photometric redshift systematics for the Rubin LSST and Roman Space Telescope

Ayan Mitra*
School of Engineering and Digital Sciences, Nazarbayev University, Nur-Sultan 010000, Kazakhstan

Eric V. Linder
Berkeley Center for Cosmological Physics & Berkeley Lab, University of California, Berkeley, California 94720, USA and Energetic Cosmos Laboratory, Nazarbayev University, Nur-Sultan 010000, Kazakhstan

Some million type Ia supernovae (SN) will be discovered and monitored during upcoming wide area time domain surveys such as the Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST). For cosmological use, accurate redshifts are needed among other characteristics; however the vast majority of the SN will not have spectroscopic redshifts, even for their host galaxies, only photometric redshifts. We assess the redshift systematic control necessary for robust cosmology. Based on the photometric vs true redshift relation generated by machine learning applied to a simulation of 500,000 galaxies as observed with LSST quality, we quantify requirements on systematics in the mean relation and in the outlier fraction and deviance so as not to bias dark energy cosmological inference. Certain redshift ranges are particularly sensitive, motivating spectroscopic followup of SN at $z \lesssim 0.2$ and around $z \approx 0.5$–0.6. Including Nancy Grace Roman Space Telescope near infrared bands in the simulation, we reanalyze the constraints, finding improvements at high redshift but little at the low redshifts where systematics lead to strong cosmology bias. We identify a complete spectroscopic survey of SN host galaxies for $z \lesssim 0.2$ as a highly favored element for robust SN cosmology.

I. INTRODUCTION

Type Ia supernovae (SN) are incisive probes of the cosmic expansion history, giving the tightest constraints on dark energy properties of any probe for a given distance precision. Wide field surveys can multiplex the observations, i.e., measure many SN at once in a given survey field, and time domain surveys that revisit fields on a cadence compatible with SN rise and fall times can monitor them throughout their lightcurve. Upcoming surveys such as ZTF [1] and LSST [2,3] are time domain surveys that are wide field (as well as deep, measuring SN over a significant range of redshifts) and with multiple wavelength bands, and thus will produce thousands to of order a million SN lightcurves.

For best use as a cosmological probe, these sources need not only to be observed in multiple bands with good photometry for a significant part of their lightcurve (for accurate fitting and color corrections), as these surveys will provide, but also be classified as type Ia supernovae to ensure a pure sample, and ideally subtyped as normal type Ia. The last two characteristics are most robustly established through spectroscopy. The SN redshift must also be determined, e.g., through the redshift of the host galaxy, either through spectroscopy or photometry.

This final step of the redshift determination for the distance-redshift relation to measure the cosmic expansion is what we focus on here, specifically the requirement on photometric redshift ("photo-z") accuracy. Statistical uncertainties in the redshift will propagate through to increasing the cosmological parameter uncertainties (see [4] for detailed calculations), but systematic errors will bias the cosmology. We extend the analysis of [5], which employed analytic toy models for redshift systematic uncertainties, to use data simulated to reflect LSST observing characteristics, i.e., filters, exposure depths, etc.

In Sec. II we review the methodology for propagating systematic redshift errors into cosmological parameter bias. The simulated data and the derived photo-z vs true redshift mapping is discussed in Sec. III, with the results presented in Sec. IV, along with the requirements necessary to avoid significant bias. We explore the effects of external near infrared data (NIR) from the Roman Telescope in Sec. V, and conclude in Sec. VI.

*ayan.mitra@nu.edu.kz

II. METHOD

To propagate an observational systematic to bias in cosmology inference, a well used and convenient technique is the Fisher information bias formalism [6,7]. Specifically we follow the approach in [5], with a parameter set including the supernova absolute magnitude parameter $\mathcal{M}$, the present matter density Ω_m as a fraction of the critical density, and the dark energy equation of state parameters w_0 and w_a giving its present value and a measure of its time variation. We include the effects of a misestimated redshift z on both the distances and the lightcurve width-luminosity relation (but not extinction, which is expected to be a small effect).

The final relation, as in [5], for the apparent magnitude offset Δm is

$$\Delta m = \frac{5}{\ln 10}\ln\left[\frac{D_L(z+\delta z)}{D_L(z)}\right] + 1.4\frac{\delta z}{1+z}, \qquad (1)$$

where δz is the redshift systematic at redshift z and D_L is the luminosity distance. This misestimation Δm then biases the cosmology parameters. Imposing constraints on the degree of cosmology bias in turn propagates back to requirements on redshift systematics, as a function of type, degree, and redshift at which they occur. We will present the confidence contours in the dark energy w_0–w_a plane (marginalized over the other parameters), and require that the bias shift the cosmology by less than 1σ, i.e., staying within the 68.3% joint confidence level contour: $\Delta\chi^2 < 2.3$.

In [5] we focused on toy models for additive and multiplicative systematics, and outliers. Here we use the photometric vs true redshift mapping derived from simulations, as described in the next section.

III. DATA

To obtain the redshift of a SN, the most common method is to measure the host galaxy redshift. This either exists in previous catalogs or can be determined by the survey itself, possibly after the SN has faded. Less common is getting an estimate of the redshift from the SN colors (flux differences between wavelength bands) [8]; it is also possible that this could be useful in avoiding catastrophic outliers in the galaxy photometric redshift [9–11]. However this has not been fully tested for side effects and confirmed, and here we consider only host galaxy redshifts.

A. Photo-z catalog

The leading current simulations of LSST galaxy photometric redshift distributions use color matched nearest neighbors (CMNN) estimators [12,13]. This incorporates the expected photometric data quality and survey characteristics. The photo-z estimator is based on the location of a test galaxy in color space, identifying it to the nearest color matched galaxy from the spectroscopic training data set. Minimization on the distance metric is done via χ^2 to estimate the photo-z of the test galaxy. The χ^2 (Mahalanobis) distance D is computed as

$$D = \sum_1^N \frac{(c_{\rm train} - c_{\rm test})^2}{(\delta c)^2}, \qquad (2)$$

where c is the color, N is the total number of colors, and δc quantifies the measurement error in color. A test galaxy needs to be detected with at least $N = 3$ colors to be assigned a photo-z.

The training set is analogous to the spectroscopic galaxy sample while the test set is composed of galaxies with simulated colors, for which the photo-z are estimated. For training the CMNN estimator we have simulated a larger data set than in [13], of 500,000 galaxies and for testing a sample size of 90,000. As both are simulated, the true redshifts are known and the distribution of the derived photo-z relative to the true redshift can be mapped.

This estimator is designed such that the accuracy and the precision of the photo-z are directly related to the precision of the survey's photometry. The training and testing sets were drawn entirely from the simulated photometry data catalog, with both having the same distribution of redshift and magnitude (flux). The simulated galaxy catalog is based on the Millennium simulation [14] and uses realistic photometric characteristics. Details on the galaxy catalog construction are described in [15,16].

The CMNN photo-z estimator was designed to model the optical (*ugrizy*) and NIR (*YJHK*) properties of galaxies. For the NIR, note that from the figures in [12] the NIR bands from the Euclid satellite [17] do not significantly impact the photo-z systematic uncertainties at redshifts $z \lesssim 1$ where we observe SN with LSST (note this refers to the much tighter photo-z requirements for supernovae; Euclid will be quite valuable for weak lensing photo-z constraints), though those from the Nancy Grace Roman Space Telescope [18] could. For the purposes of this work we use only the 10 year LSST projections, with the optical filters' 5σ detection limits as tabulated in Table I, as the main input. In Sec. V we extend this to include Roman *YJH* bands.

B. Photo-z systematics

Given the catalog of photo-z's and true redshifts, one can carry out various statistical analyses to assess robustness of the distribution. Reference [13] presented different

TABLE I. Summary of the 5σ limiting magnitudes for each filter for a LSST 10 year forecast, based on the LSST simulation software package [19].

u	g	r	i	z	y
26.1	27.4	27.5	26.8	26.1	24.9

statistical measures of the photo-z quality based on the results from the CMNN estimators on the mock galaxy catalog. In this analysis, we will follow these definitions, with alterations as described in Sec. IV. From these we will derive the quantity needed for the cosmology requirements on systematics, $\delta z(z)$ in Eq. (1).

To model the LSST-like uncertainty, Gaussian random photometric scatter was added to the simulated observed apparent magnitudes from the true catalog with a standard deviation equal to the predicted magnitude error for each galaxy. The magnitude error was modeled as for LSST, based on the description in [20]. For computing the cosmology bias we use two types of photo-z systematic offsets for δz, the bias in the core of the photo-z distribution (referred to in [13] as "robust bias") and in the outlier distribution.

1. Robust bias

If the mean photo-z relation not only scatters about the true redshift but is biased from it, generally by different amounts at different redshifts, this will lead to a cosmology bias. This core offset, or robust bias, is defined as the mean bias in the interquartile range (IQR; defined as including 50% of the galaxies) of the galaxy photo-z error [13]. For a photo-z error defined as

$$\Delta z_{(1+z)} = \left(\frac{z_{\rm true} - z_{\rm phot}}{1 + z_{\rm phot}}\right), \tag{3}$$

where $z_{\rm true}$ is the true or spectroscopic catalog redshift and $z_{\rm phot}$ is the estimated photometric redshift, the robust bias is taken to be the mean over the interquartile range

$$\text{Robust Bias} = \overline{\Delta z_{(1+z),\rm IQR}}. \tag{4}$$

We refer to this as $\delta z_{\rm core}$.

2. Outlier bias

Photo-z errors sometimes have large departures from the true values, lying outside the core. There are two types of these: outliers and catastrophic outliers. Here we need to know not only the degree of offset but the fraction of photo-z's that are outliers.

Catastrophic outliers are defined as $|z_{\rm true} - z_{\rm phot}| > 1.5$ and it is highly unlikely that a SN with such a mistaken redshift would go unrecognized and be placed on a Hubble diagram (its peak magnitude, lightcurve width, etc. would lie well off expectation for a reasonable variation of cosmology). Therefore we do not consider catastrophic outliers.

We classify outliers as photo-z's that lie outside the core, defined as $\Delta z_{(1+z)} > 3\sigma(z)$ where $\sigma(z)$ is the standard deviation for the IQR galaxies. Furthermore we require $\Delta z_{(1+z)} > 0.06$.

3. Sample selection

To select our data set for the analysis, we modified and reran the CMNN simulation with two alterations relative to [13]:

(i) $0 < z_{\rm true} \leq 1.2$, as the LSST supernova survey is expected to cover this range, i.e., SN at higher true redshifts would likely be too faint for the survey magnitude limits. This restricts the redshift range on the high end relative to [13], but expands it on the low end since [13] did not identify outliers for $z \leq 0.2$.

(ii) $|\delta z| < 1.5$, i.e., we distinguish between outliers and catastrophic outliers.

Figure 1 shows the result of the updated simulation analysis. Core points are in green, outliers are in blue if they fall within the $z_{\rm true}$ cut and red if they are at higher true redshift, while gold points are catastrophic outliers and not included in the cosmology analysis. Green core points out to $z_{\rm true} \leq 1.2$ determine $\delta z_{\rm core}(z_{\rm true})$ while blue points determine $\delta z_{\rm out}(z_{\rm true})$ and the fraction of outliers $f_{\rm out}(z_{\rm true})$, equaling the ratio of the number of blue points to blue + green points at that redshift. We use a binning on the statistics of $\Delta z = 0.05$, much finer than the original $\Delta z = 0.3$ in [13]. To train the CMNN estimator, we simulated 500,000 galaxies for training and 90,000 galaxies for testing (out to $z_{\rm true} = 3$).

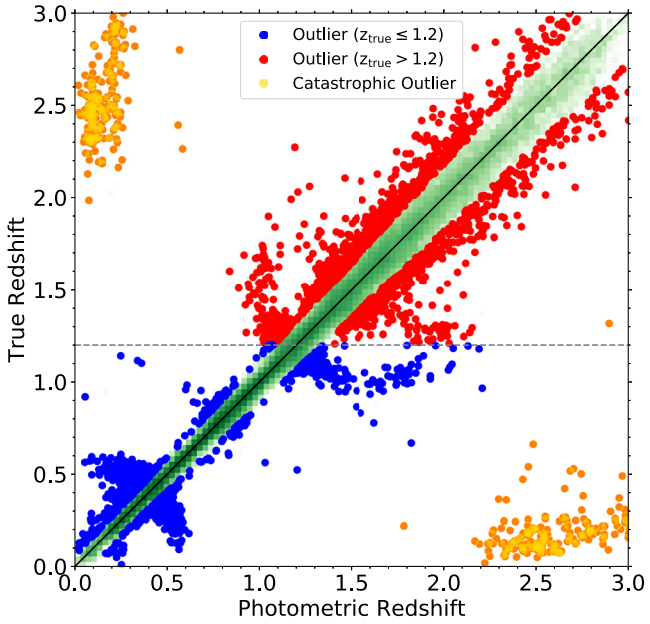

FIG. 1. Photo-z's vs true redshifts resulting from the CMNN galaxy simulation [13], with modified selection criteria. Galaxies are classified as lying in the core of the distribution (green), outliers (blue if $z_{\rm true} \leq 1.2$, red if $z_{\rm true} > 1.2$), or catastrophic outliers (gold). The grey horizontal dashed line shows the $z_{\rm true} \leq 1.2$ cut and the black diagonal line shows a perfect survey, with $z_{\rm phot} = z_{\rm true}$. Any deviation in the mean from the diagonal gives a bias.

IV. ANALYSIS

We propagate the quantities $\delta z_{\rm core}$, $\delta z_{\rm out}$, and $f_{\rm out}$ derived from analysis of Fig. 1 into the cosmology analysis, i.e., Eq. (1) and then the Fisher information analysis. The quantity $f_{\rm out}$ enters there since only a fraction $f_{\rm out}$ of the SN have the ensuing outlier bias Δm. Thus, both $\delta z_{\rm out}$ and $f_{\rm out}$ are important: if the redshift offset $\delta z_{\rm out}$ is high, but happens only rarely (low $f_{\rm out}$), this will give a small cosmology bias, as will a large outlier fraction $f_{\rm out}$ but with only a small offset $\delta z_{\rm out}$.

For the cosmology calculation we follow the analysis described in [5], with SN distributed over the range $z = [0\text{–}1.2]$. We evaluate δz in bins of width $\Delta z = 0.05$ (see discussion below). We treat the core bias and outlier bias separately, for clarity.

Figure 2 shows the bias in cosmology as a result of the core bias redshift systematic. The input cosmology has $(w_0, w_a) = (-1, 0)$ and the ellipse shows the 1σ (68.3%) joint confidence contour, marginalized over the other parameters (matter density and SN absolute magnitude $\mathcal{M}$), for a rough approximation of the LSST SN sample plus a Planck CMB prior on the distance to last scattering.

Applying the $\delta z_{\rm core}(z)$ systematic to the redshift bins one by one, the corresponding shift in the cosmology is shown by the individual square boxes, beginning with the orange box ($z = [0, 0.05]$), and ending with the blue box ($z = [1.15, 1.2]$). The red line connecting the boxes traces the locus of the cosmology bias with increasing bin redshift. The green arrow shows the net effect (vector sum) for the systematic present on all the redshift bins (zoomed-out insets are used to show this and any individual

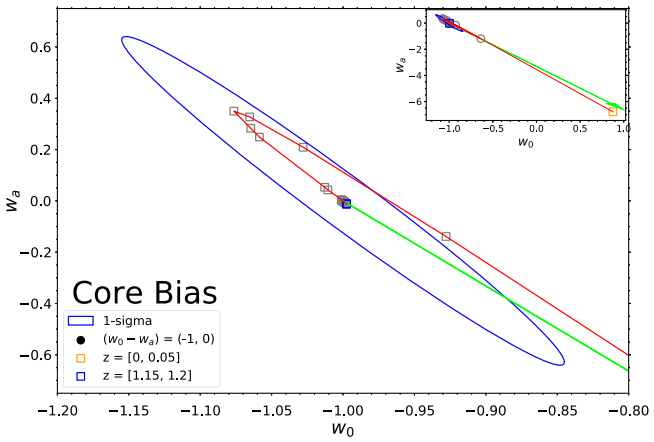

FIG. 2. Cosmology bias in the dark energy w_0–w_a plane coming from core bias redshift systematics is shown by the red curve tracing the effect from one redshift bin systematic $\delta z_{\rm core}(z)$ at a time, marked by squares (circles in the zoomed-out inset plot) from the lowest (orange; only visible in the inset plot) to highest redshift (blue), for bins of $\Delta z = 0.05$. The green arrow gives the vector sum over all redshifts, with the inset showing the full extent of the bias. For reference, the statistical uncertainty is shown as the 1σ (68.3%) joint confidence contour (blue ellipse).

shifts when they extend outside the main figure—e.g., the orange square from the lowest redshift bin is only visible in the inset, and intermediate bins are shown as circles in the inset to distinguish them from the main figure). Note the direction of bias from some redshifts is such that it can diminish bias from another redshift. Cosmology is most sensitive to the low redshift systematics, with systematics in the three lowest bins (two of which are only visible in the inset) biasing cosmology outside the 1σ contour. This can be a hopeful sign in that these can be the most easily addressed with supplementary observations.

Next we turn to the outlier bias. For this we find the fine redshift bins of width $\Delta z = 0.05$ to be important to treat properly the sharp outlier features, especially at low redshift. Quantitatively, the cosmology parameter bias is too coarsely estimated (relative to a numerically intensive unbinned analysis) by using bins of $\Delta z = 0.1$ rather than 0.05, by $\sim$0.03 in w_0 and 0.13 in w_a only at $z < 0.2$; above this the maximum errors are $\lesssim$0.015 and 0.06 respectively, and generally much less. Recall that for redshift systematics due to outliers we care about both the offset and the fraction of galaxies exhibiting the systematics. Except at low redshift (where dm/dz is steep), the cosmology bias is basically proportional to their product, $f_{\rm out}\delta z$.

Figure 3 shows the simulation results for $\delta z(z)$, $f_{\rm out}(z)$, and their product. Although the product looks quite small, one must propagate it to the cosmology bias to determine its impact. Even a small systematic can have a significant effect on a high precision survey such as LSST.

Figure 4 shows the resulting cosmology bias from the outlier systematics. While the effect on the first redshift bin is less than from the core bias, systematics from the remainder of the redshift range have comparable,

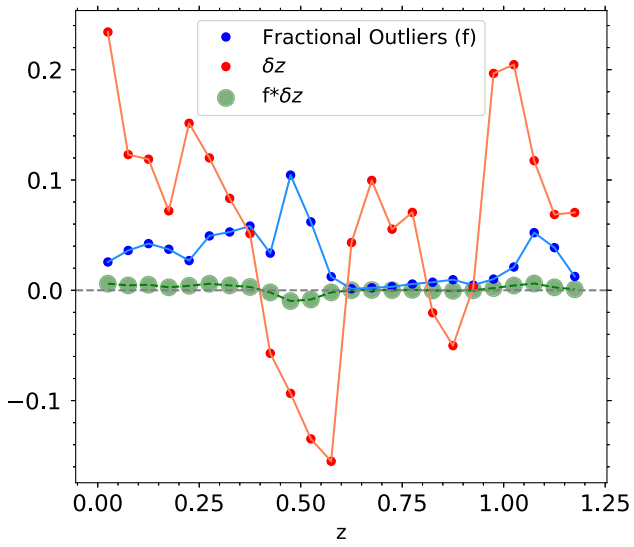

FIG. 3. Simulation output for outlier redshift systematics magnitude $\delta z_{\rm out}(z)$, fraction of outliers $f_{\rm out}(z)$, and their product. Sharp features in the outlier systematics require fine binning in redshift.

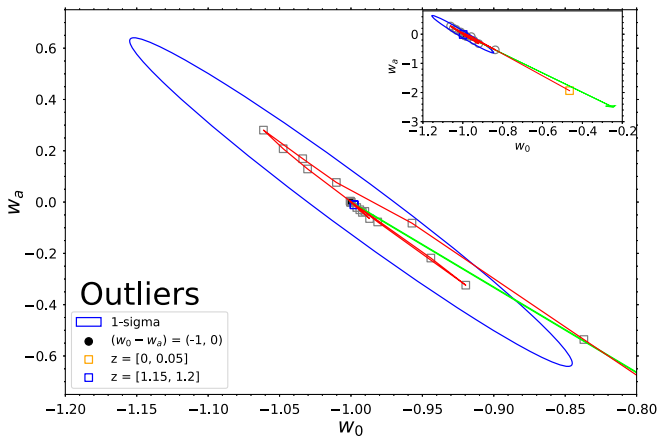

FIG. 4. As Fig. 2, but for outlier bias. Again, low redshift bins $z \lesssim 0.2$ are particularly affected.

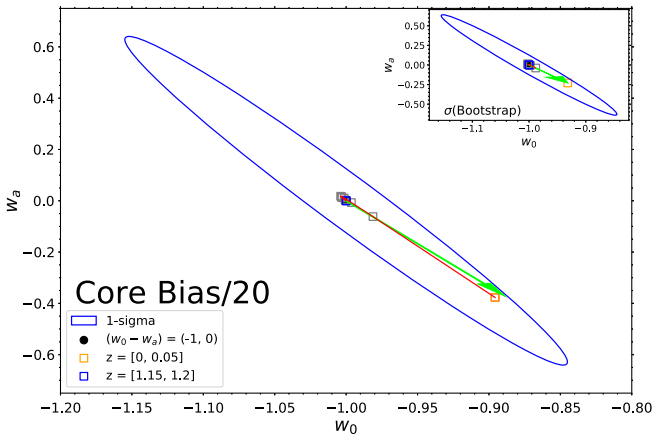

FIG. 5. As Fig. 2, but with the core bias now scaled down by a factor 20, mimicking modeling of the systematic to leave a small residual. At this level of control, the net bias is contained within the 1σ statistical confidence contour. The inset, rather than being a zoom-out like in the other similar figures, shows the result if the systematics were modeled down to the bootstrap sampling error from 1000 simulations.

significant effects. Of course both types of systematics will be present, and their induced bias goes in the same direction for distorting the cosmology inference.

Since the cosmology bias from the redshift core systematics and the outlier systematics, and their sum, are well beyond the desired 1σ statistical confidence contour, it must be reduced for useful cosmology estimation from LSST photometric supernovae. We present three possible strategies for amelioration: (1) modeling, (2) select spectroscopy, (3) external imaging data.

The redshift systematics found by the simulations is a raw systematic, without remediation. One could attempt to model the systematics and correct for them, up to the fidelity of the modeling process, leaving a smaller, residual systematic. Figure 5 shows that if the residual core bias is scaled down by a factor of 20, i.e., leaving only 5% of the simulation systematic, then the cosmology bias lies within the 1σ confidence contour. (In fact, one should not take the factor 20 too precisely: the Fisher bias formalism is valid for small changes in the observable, i.e., Δm, so the effect of large changes at the lowest redshifts is possibly exaggerated.) The inset shows the case for the systematics reduced all the way to the bootstrap sampling error[1] [13]. A similar process could potentially be applied to the outlier bias.

The second strategy involves targeting the systematics in particular redshift ranges. We have seen that the low redshift systematics produce the highest cosmology bias. This is expected as the SN apparent magnitude m on the Hubble diagram starts off steeply varying with redshift, roughly $\sim 1/z$, and then flattens at higher redshift. So a small δz at low redshift has a significant effect. This is fortunate in that the low redshift region is most amenable to use of spectroscopy to determine the SN redshift (e.g., through targeting its host galaxy).

Figure 6 shows the application of this approach to the case of outlier systematics. The three arrows correspond to the total cosmology bias that ensues from SN over the full range $z = [0, 1.2]$ if those (and only those) SN at $z < z_\star$ get spectroscopic redshifts, and so there are no outliers there: $f_{\rm out}(z < z_\star) = 0$. Spectroscopic redshifts for $z \lesssim 0.2$ SN bring the bias under control. One possibility for carrying this out is through the secondary target program of the DESI Bright Galaxy Survey [21], or other multifiber spectrographs [22,23]. Such a low redshift spectroscopic SN sample is quite interesting scientifically as it also has power as a cosmic probe of gravity through peculiar velocities [24,25].

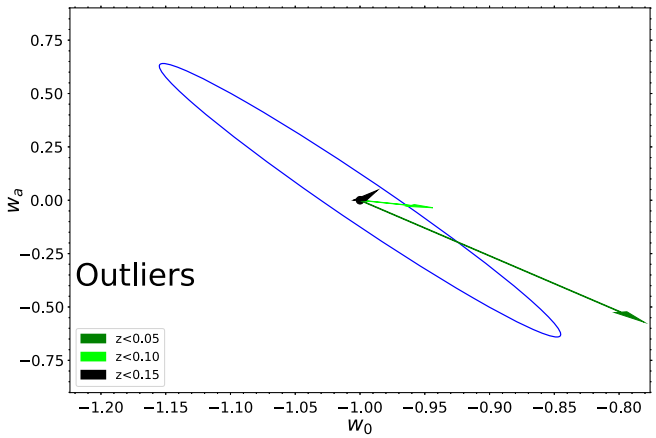

FIG. 6. As Fig. 4, but with spectroscopic redshifts, and hence $f_{\rm out} = 0$, for $z < z_\star$, for three different values of $z_\star$.

[1] The bootstrap error in the core bias is derived by randomly drawing galaxy subsets with replacement and recalculating the statistics 1000 times, and then using the standard deviation of all recalculations as the error.

We consider the third strategy of utilization of external imaging data to improve redshift systematics in the next section.

V. LSST + ROMAN SPACE TELESCOPE

During the LSST, supplementary imaging data that can help constrain the photometric redshift estimation will be available, notably from the Euclid satellite and the Nancy Grace Roman Space Telescope, both with near infrared wavelength measurements. As discussed in Section III A, Euclid's leverage comes at redshifts higher than those whose photometric systematics most significantly impact the supernova cosmology. Roman, however, extends to lower redshifts and we consider the benefit from adding YJH[2] imaging data with exposure depths comparable to those from Roman to the photometric redshift determination (26.7, 26.9, 26.0 mag respectively for 5σ limiting depths). We follow [13] for the joint analysis photo-z catalogs. While Roman will of course have its own highly incisive sample of spectroscopic SN, here we consider only its effect on LSST photo-z.

We carry out the joint analysis in two distinct ways. First, we consider the effect of the added Roman information on the outlier systematics analysis of Sec. IV, comparing LSST alone with LSST + Roman. However, since the addition of Roman data can decrease the photo-z standard deviation, the outlier fraction can actually increase since it is defined in terms of galaxy photo-z's deviating from $z_{\rm true}$ by 3σ. (Also, LSST catastrophic outliers can become LSST + Roman regular outliers.) Therefore we also discuss the full systematics—from outlier galaxies (outside 3σ) and inlier galaxies (within 3σ; note that the core bias of the previous section, as defined by [13], is restricted to the 1σ core: here the sum of the outliers and inliers contains all the non-catastrophic galaxies)—in a second cosmology analysis.

A. LSST + Roman outlier analysis

The color matching nearest neighbors (CMNN) algorithm of [13] works somewhat differently when combining external data with LSST, so the interested reader should consult that paper for a full discussion. The main point to note here is that because including new filters not only adds information but also degrees of freedom, i.e., fit parameters, in the algorithm, if the additional NIR filter does not carry clear photo-z information (as can occur at, say, low redshift) then the combination can actually give worse results than LSST alone. Future work could consider how to treat this, either by cutting or tapering multiband information in such cases, or adjusting the CMNN algorithm.

[2]The K filter, really F184, is useful at much higher redshift than we consider here. See Fig. 10 and Sec. IV in [13] for discussion.

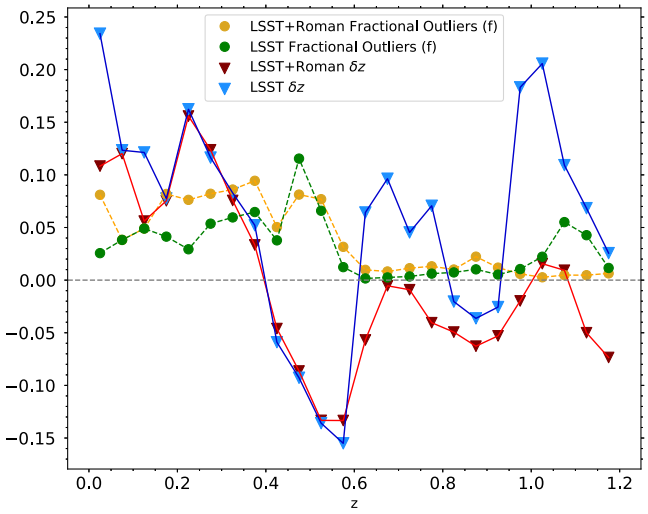

FIG. 7. Comparison between LSST + Roman and LSST simulation output for outlier redshift systematics magnitude $\delta z_{\rm out}(z)$ and the corresponding fraction of outliers $f_{\rm out}(z)$.

Figure 7 compares outlier systematics in terms of $\delta z_{\rm out}$ and $f_{\rm out}$ for LSST and LSST + Roman. We see that Roman does help to tame the excess systematic in the lowest redshift bin, although it actually has a higher fraction of outliers there. Over the range $z = [0.2, 0.6]$ there is little impact on $\delta z_{\rm out}$ from the NIR bands, but Roman data improves the photo-z significantly for $z \gtrsim 0.6$. For $z \gtrsim 1$, the outlier fraction when including Roman is strongly reduced, so what deviations $\delta z_{\rm out}$ do exist only affect a small number of galaxies.

The outlier systematics at $z \lesssim 0.2$ and around $z \approx 0.5$ still remain, so the cosmology bias issue is not solved. Figure 8 shows the cosmological effects of the photo-z systematic. By comparing to Fig. 4 we see qualitatively similar behavior, and the quantitative aspects are also not very different. The improvement in photo-z at high redshifts from Roman does not result in a large impact because high

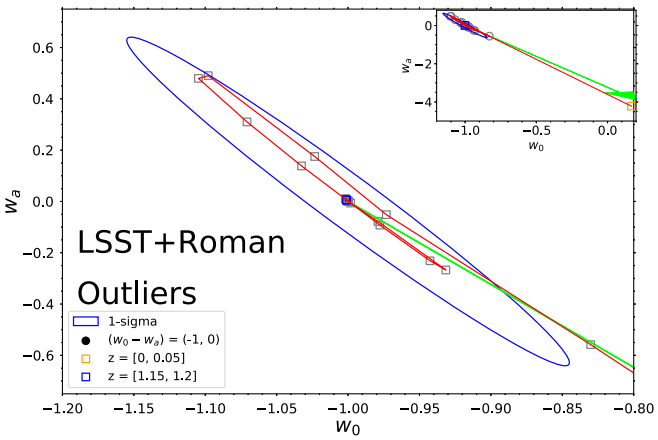

FIG. 8. As Fig. 4, but cosmology bias from the outlier systematics in the LSST + Roman case.

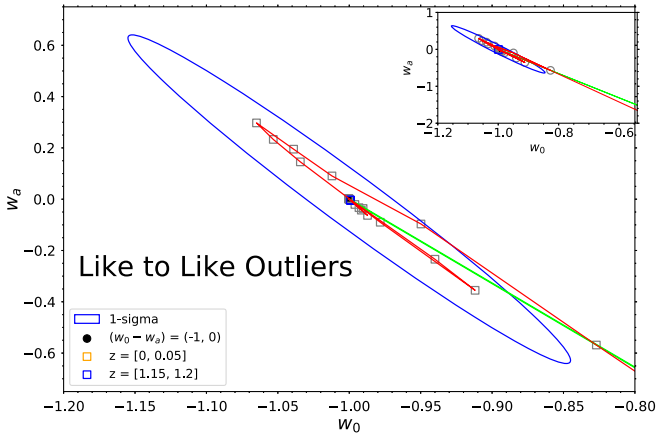

FIG. 9. As Fig. 8, but for a like to like treatment where the LSST + Roman outliers are restricted to galaxies that were also LSST alone outliers. Cosmology bias is reduced from LSST alone (Fig. 4) and from LSST + Roman general outliers (Fig. 8).

redshifts give a fairly modest contribution to the cosmology bias, which is much more sensitive to low redshifts. Thus, the need for low redshift spectroscopy to remove the photo-z outlier systematic remains strong.

As mentioned, adding Roman NIR data to LSST changes which galaxies are considered outliers, and so the ameliorating effects of the extra data are somewhat obscured. We can consider only those galaxies classified as outliers using just LSST, and then add Roman data to those alone and examine the properties of those that remain outliers: this is a "like to like" comparison. Figure 9 shows the cosmology bias for this case, and we see that indeed significant improvement is evident.

B. LSST + Roman all galaxy analysis

Adding NIR not only affects the photo-z outliers, but changes the size of the core and hence the dividing line between outliers and inliers. To take into account all the

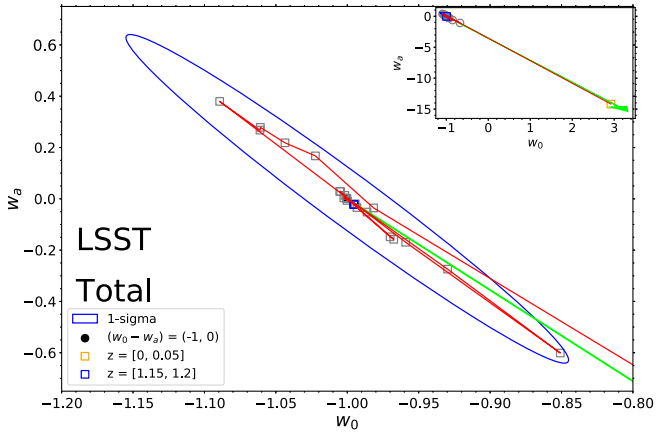

FIG. 10. As Fig. 2, but cosmology bias from the total outlier and inlier photo-z systematics for LSST.

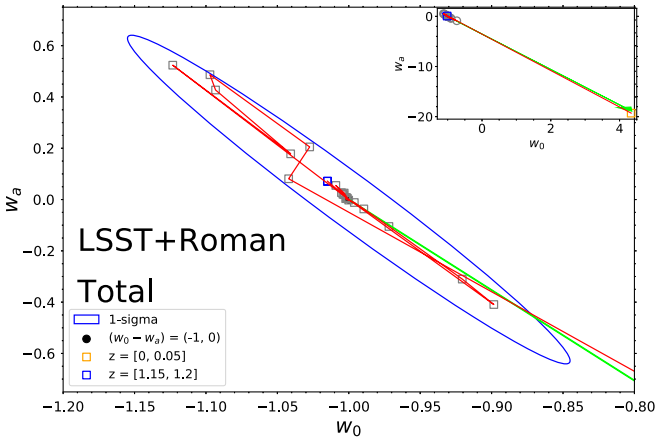

FIG. 11. As Fig. 10, but for LSST + Roman.

effects on the photo-z systematics of adding NIR data, we consider here all the galaxies together, outliers and inliers, and compute the total cosmology bias.

Figures 10 and 11 show the total cosmology bias from all galaxy photo-z systematics for LSST alone and for LSST + Roman. Again, qualitatively they are similar and quantitatively there is not a large difference. While cosmology bias is significantly reduced around $z \approx 0.5$ (by at least $\delta w_a > 0.15$), due to a combined reduction in photo-z outlier and inlier systematics there, and photo-z systematics is improved at $z \gtrsim 1$—but systematics there causes relatively little cosmology bias—the low redshift systematics remains, and this substantially drives the cosmology bias.

Thus, the addition of NIR data does not obviate the need for spectroscopic redshifts for SN host galaxies at $z \lesssim 0.2$, and further improvements around $z \approx 0.5$ would be useful as well.

VI. CONCLUSIONS

Supernovae continue to be one of the most incisive probes of cosmic acceleration. With forthcoming surveys the data available will vastly increase, limited however by spectroscopic followup. We consider the use of type Ia supernovae with only photometric redshifts, assessing the impact of systematics in the redshift estimation through large simulations of host galaxy colors. This builds on the analytic work of [5] by using mock data meant to emulate LSST observations to derive the systematics requirements for controlling bias in the cosmological parameter estimation, particularly for dark energy.

Simulating half a million galaxies provides us with good statistics on the redshift systematics, categorized into outlier and core systematics, as a function of true redshift. We propagate this to cosmological parameter bias for a LSST-like survey, showing the impact of each individual redshift range as well as the full sum. Our results show that for both outliers and inliers, the redshift systematic

requirement is reduction by an order of magnitude—principally in the lower redshift range—for the bias not to exceed the 68% confidence statistical uncertainty in the w_0–w_a plane.

The low redshift photometric systematics are the most dangerous. Fortunately, they are also easiest to mitigate with further observations. The favored situation would be to use the photometric supernova for cosmology only for $z > 0.3$, and obtain spectroscopic followup for all supernovae at $z \lesssim 0.2$–0.3 (see below).

We explore three potential methods for controlling systematics: modeling, select spectroscopy, and external imaging data. Fully successful modeling, i.e., a residual at only the level of the bootstrap uncertainty on a large suite of simulations (that accurately capture the survey photo-z characteristics), would be ideal, while a residual an order of magnitude below base is necessary. Improvement in understanding host galaxy properties (for both the core bias and the outlier bias) is desirable, e.g., are the spectroscopic catalogs used for training the CMNN estimators—and the resulting outliers—representative, in particular of type Ia SN host galaxies. Machine learning algorithms to estimate the redshifts will likely continue to become better; Ref. [26] has summarized the accuracy for several different algorithms, and while the CMNN estimator fares quite well in comparison to most of the other machine learning estimators, the authors of [13] state that it is not optimized for the absolute best fit, but rather aims to assess differences in survey strategy.

Select spectroscopy is an attractive and highly practical solution. This would involve a multiobject spectrograph survey to obtain SN host galaxy redshifts out to $z \lesssim 0.2$, where the systematics have the greatest impact on cosmology bias. Such data has numerous other astrophysical applications, including for probing gravity with peculiar velocities [24,25].

External imaging data will be available from NIR surveys such as the Nancy Grace Roman Space Telescope and the Euclid satellite. We carry out an analysis including Roman NIR filter constraints on galaxy photometric redshifts with LSST mock data in the simulations. We consider the impact on systematics from photo-z outliers, inliers (the complement of outliers), the total set, and a special like to like comparison where we choose the same set of galaxies from both LSST and LSST + Roman. Especially for the like to like comparison, Roman helps in controlling the systematics. However, overall the main redshift estimation improvement is at high redshifts where the cosmology bias is less. The sensitive low redshifts show relatively little gain. So even in the era of LSST, Euclid, and Roman, low redshift spectroscopy of SN host galaxies will be quite important for enabling the full potential of dark energy constraints from supernova cosmology through time domain surveys.

ACKNOWLEDGMENTS

We are very grateful to Melissa Graham for making her data and code public and for many useful tips on how to adapt them. We also thank Rick Kessler, Alex Kim, Gautham Narayan, and the LSST DESC Supernova and Photometric Redshift working groups for comments and suggestions. A. M. acknowledges the support of ORAU Grant No. 110119FD4534. E. L. is supported in part by the U.S. Department of Energy, Office of Science, Office of High Energy Physics, under Contract No. DE-AC02-05CH11231, by NASA ROSES Grant No. 12-EUCLID12-0004, and by the Energetic Cosmos Laboratory.

[1] E. C. Bellm *et al.*, The Zwicky transient facility: System overview, performance, and first results, Publ. Astron. Soc. Pac. **131**, 018002 (2019).
[2] The LSST Dark Energy Science Collaboration, The LSST Dark Energy Science Collaboration (DESC) science requirements document, arXiv:1809.01669.
[3] H. Zhan and J. Anthony Tyson, Cosmology with the large synoptic survey telescope: An overview, Rep. Prog. Phys. **81**, 066901 (2018).
[4] D. Huterer, A. Kim, L. M. Krauss, and T. Broderick, Redshift accuracy requirements for future Supernova and number count surveys, Astrophys. J. **615**, 595 (2004).
[5] E. V. Linder and A. Mitra, Photometric supernovae redshift systematics requirements, Phys. Rev. D **100**, 043542 (2019).
[6] L. Knox, R. Scoccimarro, and S. Dodelson, The Impact of Inhomogeneous Reionization on Cosmic Microwave Background Anisotropy, Phys. Rev. Lett. **81**, 2004 (1998).
[7] E. V. Linder, Biased cosmology: Pivots, parameters, and figures of merit, Astropart. Phys. **26**, 102 (2006).
[8] R. Kessler *et al.*, Photometric estimates of redshifts and distance moduli for type Ia Supernovae, Astrophys. J. **717**, 40 (2010).
[9] R. Kessler *et al.*, Testing models of intrinsic brightness variations in type Ia supernovae and their impact on measuring cosmological parameters, Astrophys. J. **764**, 48 (2013).
[10] M. Dai, S. Kuhlmann, Y. Wang, and E. Kovacs, Photometric classification and redshift estimation of LSST Supernovae, Mon. Not. R. Astron. Soc. **477**, 4142 (2018).

[11] E. Roberts, M. Lochner, J. Fonseca, B. A. Bassett, P.-Y. Lablanche, and S. Agarwal, zBEAMS: A unified solution for supernova cosmology with redshift uncertainties, J. Cosmol. Astropart. Phys. 10 (2017) 036.

[12] M. L. Graham, A. J. Connolly, Ž. Ivezić, S. J. Schmidt, R. L. Jones, M. Jurić, S. F. Daniel, and P. Yoachim, Photometric redshifts with the LSST: Evaluating survey observing strategies, Astron. J. **155**, 1 (2018).

[13] M. L. Graham *et al.*, Photometric redshifts with the LSST. II. The impact of near-infrared and near-ultraviolet photometry, Astron. J. **159**, 258 (2020).

[14] V. Springel *et al.*, Simulations of the formation, evolution and clustering of galaxies and quasars, Nature (London) **435**, 629 (2005).

[15] V. Gonzalez-Perez, C. G. Lacey, C. M. Baugh, C. D. P. Lagos, J. Helly, D. J. R Campbell, and P. D. Mitchell, How sensitive are predicted galaxy luminosities to the choice of stellar population synthesis model?, Mon. Not. R. Astron. Soc. **439**, 264 (2014).

[16] A. I. Merson *et al.*, Lightcone mock catalogues from semi-analytic models of galaxy formation—I. Construction and application to the BzK colour selection, Mon. Not. R. Astron. Soc. **429**, 556 (2013).

[17] R. Laureijs *et al.*, Euclid definition study report, arXiv: 1110.3193.

[18] O. Dore *et al.*, WFIRST: The essential cosmology space observatory for the coming decade, Bull. Am. Astron. Soc. **51**, 341 (2019).

[19] A. J. Connolly *et al.*, An end-to-end simulation framework for the Large Synoptic Survey Telescope, edited by G. Z. Angeli and P. Dierickx, in *Modeling, Systems Engineering, and Project Management for Astronomy VI*, Vol. 9150 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series (2014), p. 915014.

[20] Ž. Ivezić *et al.*, LSST: From science drivers to reference design and anticipated data products, Astrophys. J. **873**, 111 (2019).

[21] O. Ruiz-Macias *et al.*, Preliminary target selection for the DESI bright galaxy survey (BGS), Res. Notes Am. Astron. Soc. **4**, 187 (2020).

[22] R. Mandelbaum and LSST Dark Energy Science Collaboration, Wide-field multi-object spectroscopy to enhance dark energy science from LSST, Bull. Am. Astron. Soc. **51**, 363 (2019).

[23] M. Graham *et al.*, Discovery frontiers of explosive transients: An ELT and LSST perspective, Bull. Am. Astron. Soc. **51**, 339 (2019).

[24] A. G. Kim and E. V. Linder, Complementarity of peculiar velocity surveys and redshift space distortions for testing gravity, Phys. Rev. D **101**, 023516 (2020).

[25] A. G. Kim *et al.*, Testing gravity using type Ia supernovae discovered by next-generation wide-field imaging surveys, Bull. Am. Astron. Soc. **51**, 140 (2019).

[26] Z. Ansari, A. Agnello, and C. Gall, Mixture models for photometric redshifts, arXiv:2010.07319.