1. INTRODUCTION

The cosmological applications of luminosity–distance measurements to Type Ia supernovae (SNeIa) are now well known (Riess et al. 1998; Perlmutter et al. 1999; Leibundgut 2001). While the current sample of SNeIa-based distances are limited to a few hundred SNe (Astier et al. 2006; Wood-Vasey et al. 2007; Kowalski et al. 2008; Hicken et al. 2009; Kessler et al. 2009), future surveys are now planned to increase the sample size to a few thousand or more that could potentially allow a few percent accurate dark energy equation of state measurements in several redshift bins between \(0 < z < 1\) (see, e.g., Howell et al. 2009; Sarkar et al. 2008). The main challenge for constructing large samples is likely to be spectroscopic follow-up measurements to identify if each SN detected in a photometric monitoring campaign is Type Ia and to establish the redshift of that SN.

In addition to the planned space-based programs such as the Joint Dark Energy Mission (JDEM), we are in the near future, there will also be several ground-based photometric surveys for cosmological measurements and other astronomical studies. These include the Dark Energy Survey (DES), and the Pan-Starrs survey, and ultimately the Large Synoptic Survey Telescope (LSST), which plan to monitor a large area of the sky every few days leading to a large sample of transient sources including SNe. Given the large size of the samples of SNe expected, it is highly unlikely to have spectroscopic follow-ups for all or even a large fraction of them. Due to this limitation it appears challenging to obtain cosmological measurements with the SNeIa seen by LSST. Since LSST is likely to detect a few hundred thousand or more SNe per year, it would be highly desirable to identify whether a given SN as Type Ia or not, and to extract useful parameters such as redshift and luminosity with photometric data alone. If reliable techniques could be established, then even with a large degradation in accuracy for individual data compared with the case where spectroscopic data are also available, given the large number of statistics expected, one could still aim to achieve a good measurement of cosmological parameters.

In this spirit, we pursue a study to establish the extent to which photometric data from a survey like LSST can be used to identify SNeIa and to measure the cosmological parameters. We do this by fitting the photometric light curve data with sampling and errors consistent with LSST. Our mock SNe samples also include core-collapse SNe and we vary the fractions expected based on the current rate estimates of various types of SNe. Our Markov Chain Monte Carlo (MCMC) analysis is focused on a joint parameter estimation including the redshift, apparent magnitude, stretch factor, and the phase of the SNe. We find that the model fitting works adequately well when the true SNe redshift is below 0.5, while at \(z < 0.2\) the accuracy of the photometric data is almost comparable with spectroscopic measurements of the same sample.

We discuss the contamination of Type Ib/c (SNIb/c) and Type II supernova (SNII) on the SNeIa data set. We find that it is easy to distinguish the SNII through the large \(\chi^2\) mismatch when fitting to photometric data with Ia light curves. This is not the case for SNIb/c. We implement a statistical method based on the Bayesian estimation in order to statistically reduce the contamination from SNIb/c for cosmological parameter measurements from the whole SNe sample. The proposed statistical method also evaluates the fraction of the SNIa in the total SNe data set, which provides a valuable guide to establish the degree of contamination.

Key words: cosmology; theory – distance scale – large-scale structure of universe – supernovae; general

Online-only material: color figures
to Ia measurements. In addition to a cut in $\chi^2$ values, we implement a statistical Bayesian estimation method to reduce the contamination from SN Ib/c in the subset of SNe sample selected for cosmological measurements. This technique also establishes statistically the fraction of the SN Ia in the total SNe data set, which provides a valuable guide to the degree of contamination from Ib/c’s.

We employ the filter functions as currently publicized by the LSST team in addition to survey parameters outlined in Ivezic et al. (2008). We note that while our work is focused toward a survey like LSST, others have also considered the use of photometric data alone for SNe distance measurements (Johnson & Crotts 2006; Sullivan et al. 2006; Kuznetsova & Connolly 2007; Wang 2007; Kim & Miquel 2007; Zentner & Bhattacharya 2009).

The discussion is organized as follows. In the next section, we describe our procedure to simulate SNe data in a survey like LSST and move on to discuss our six-parameter model fits to the multiwavelength light curves from a large mock sample using a MCMC analysis. In Section 4, we discuss the contamination from Type II and Ib/c SNe to Ia photometric samples and a way to statistically reduce the contamination from Ib/c’s using a technique that implements the Bayes theorem.

2. SIMULATING SNIa OBSERVED FLUX DATA

In this section, we describe the process to generate the various SNe data. We first discuss the observed SNIa mock flux data.

2.1. The Mock Light Curve

The apparent observed flux from a SN at $z$ can be written as the convolution of the spectral energy distribution (SED) and the transmission function of the telescope,

$$f_{\text{obs}} = \int T_X(\lambda_{\text{obs}}) \text{SED}(\lambda_{\text{obs}}, t_{\text{obs}}, s, E_{B-V}^{\text{host}}, E_{B-V}^{\text{MW}}) d\lambda_{\text{obs}},$$

where $T_X(\lambda_{\text{obs}})$ is the filter response for band $X$, $\lambda_{\text{obs}}$ is the observed wavelength, $t_{\text{obs}}$ is the observation date, $s$ is the stretch factor, and $E_{B-V}^{\text{host}}$ and $E_{B-V}^{\text{MW}}$ are the color excess for the host galaxy and the Milky Way, respectively.

The transmission functions used in our analysis for 5-bands of LSST are shown in Figure 1 (LSST filter draft 2005). We also plot rest-frame SEDs for Type Ia SNe at $z = 0.2, 0.5$, and 0.8. These SED templates are from Nugent et al. (2002) and they cover the spectral wavelength from 1000 to 25000 Å in rest-frame days from -20 to 70 with respect to the $B$-band maximum light day. The SNe flux for the epochs before -20 are set to be zero.

There are now several techniques to parameterize the SNIa light curves, such as the 15 day decline after the $B$-band maximum light day $\Delta m_{15}$ (Phillips 1993) and the multicolor light curve shape method (MLCS and the update version MLCS2k2: Riess et al. 1996; Jha et al. 2007). In this paper, we calibrate the SNIa light curve with the timescale and stretch factor relation following the works of Perlmutter et al. (1997, 1999). By stretching and compressing the time axis around the rest frame $B$-band maximum light day, this method can fit the observed light curve very well using the light curve template (Goldhaber et al. 2001). Then this SED can be re-scaled by the apparent, unextincted $B$-band peak magnitude

$$m_B = M_B + 5 \log_{10} d_L(z, \theta) + 25 - \alpha(s - 1) + \Delta m,$$

where $M_B$ is the $B$-band absolute peak magnitude, $d_L$ is the luminosity distance which is a function of the redshift $z$ and a broad set of cosmological parameters denoted by $\theta$, and $\alpha$ is the coefficient of the relation between $s$ and $m_B$. Here, we take $M_B = -19.3, \alpha = 1.5$ (Knop et al. 2003; Astier et al. 2006), and the set of cosmological parameters $\theta$ with $\Omega_{m0} = 0.27, \Omega_{\Lambda 0} = 0.73$, and $h_0 = 0.7$ (Komatsu et al. 2009), where $\Omega_{m0}$ and $\Omega_{\Lambda 0}$ have the usual meaning with the present-day matter and dark energy density parameters and $h_0$ is the dimensionless Hubble constant. Besides, we also consider the dispersion $\Delta m$ of the rest-frame $B$-band peak magnitude after the calibration of the stretch factor. The $B$-band filter we use is from the Johnson–Morgan system (Bessell 1990, 2005). Also, the timescale of the SED is calibrated by the stretch factor $s$, which is assumed to be available from -15 to 35 around the $B$-band maximum luminosity day (Astier et al. 2006). We note that there is also an intrinsic color scatter, $\sigma_{\text{int}}^{B-V}$ (standard deviation), that should be taken into account when producing mock light curves. Based on prior work, we find this uncertainty to be small with a value of $\sim 0.05$ mag (Phillips et al. 1999; Jha et al. 2007). Hence, it would not affect our results much, so that for simplicity we don’t consider it here.

During the transit, the SNe light will be partly absorbed by the dust of the host galaxy. We employ the reddening law of Cardelli et al. (1989) with $R_V = 3.1$, from infrared to far-ultraviolet ($0.3 \mu m^{-1} \lesssim x \lesssim 10 \mu m^{-1}$, where $x = 1/\lambda$). For the optical to near-ultraviolet wavelength range ($1.1 \mu m^{-1} \lesssim x \lesssim 3.3 \mu m^{-1}$), we use an updated version for extinction given by O’Donnell (1994). The latter uses the same analytical form for extinction as Cardelli et al. (1989) but with values of the fitting parameters revised slightly from the previous version.
The level of extinction we assume here is consistent with the one measured recently by Menard et al. (2009) corresponding to large angular scales based on galaxy–quasi-stellar object cross-correlation in Sloan Digital Sky Survey. Since the SNe are at a different redshift than the observer, the spectrum is redshifted for both wavelength and the phase, i.e., \( \lambda' = \lambda(1 + z) \) and \( t' = t(1 + z) \). We also apply an extinction associated with dust in the Milky Way (Burstein & Heiles 1982; Schlegel et al. 1998), and assume that we have a perfect measurement of \( E_{B-V}^{MW} \). This assumption has no effect on our final conclusions. While the extinction of the Milky Way has different values for different sky regions, we do not have any information on the exact field selection of future SNe surveys from ground. Thus, we do not account for sky variation of extinction and simplify by just taking an average value with \( E_{B-V}^{MW} \approx 0.03 \) and \( R_V = 3.1 \) for the extinction law. Finally, the spectrum is integrated with the LSST filters to get the mock light curve sampling in each of the five LSST filters.

Since the mean redshift of the SNe detections with LSST main survey is expected to be about 0.5 and the deeper, but smaller, survey can potentially detect SNe out to \( \sim 1 \), we choose the redshift range from 0.01 to 1.1 when making mock SNe samples.

When creating large samples, we assume the flat \( \Lambda \)CDM model with \( \Omega_{\text{m0}} = 0.27 \) and \( h_0 = 0.71 \). Then, the redshift \( z \), stretch factor \( s \), the extinction of the host galaxy \( E_{B-V}^{\text{host}} \), and the magnitude dispersion \( \Delta m \) are generated from the Gaussian distribution with truncated tails as follows: 0.01 \( \leq z \leq 1.1 \) with \( \bar{z} = 0.5 \) and \( \sigma_z = 0.4, 0.6 \leq s \leq 1.4 \) with \( \bar{s} = 1 \) and \( \sigma_s = 0.3, -0.1 \leq E_{B-V}^{\text{host}} \leq 0.3 \) with \( E_{B-V}^{\text{host}} = 0.0 \) and \( \sigma_E = 0.2 \) and \( -0.3 \leq \Delta m \leq 0.3 \) with \( \bar{\Delta m} = 0.0 \) and \( \sigma_{\Delta m} = 0.17 \). This extra dispersion acts as an extra source of noise in our mock data (Sullivan et al. 2006; Hamuy et al. 1995, 1996; Phillips et al. 1999; Guy et al. 2005).

2.2. The Photometric Error and The Cadence

The photometric error we use for LSST comes from Ivezic et al. (2008) and takes the form of

\[
\sigma^2_{\text{phos}} = \sigma^2_{\text{sys}} + \sigma^2_{\text{zero}} + \sigma^2_{\text{rand}},
\]

where \( \sigma_{\text{sys}} \) is the systematic photometric error which is designed to be very small (< 0.005 mag), \( \sigma_{\text{zero}} \) is the absolute photometric error that we set to be \( \sigma_{\text{zero}} = 0.02 \) mag (Astier et al. 2006; Sullivan et al. 2006). We note that, in practice, there is only one zero-point realization in any given experiment that is applied to all SNe. This would result in a non-diagonal covariance matrix for the distance modulus (Kim & Miquel 2006). However, since the inclusion or non-inclusion of this covariance does not change the principles of our methodology, for simplicity, we ignore this correlation here. We do suggest that it must be considered in an analysis of real data.

In Equation (3), \( \sigma^2_{\text{rand}} \) is the random photometric error for point sources given by

\[
\sigma^2_{\text{rand}} = (0.04 - \gamma)\lambda + \gamma x^2.
\]

Here, \( \gamma \) is a parameter related to the sky brightness and readout noise, among others, and \( x = 10^{0.4(m_m - m_s)} \), where \( m \) is the magnitude and \( m_s \) is 5\( \sigma \) depth for a detection of a point source in each of LSST bands. The \( m_s \) is a function of the sky brightness, the seeing, the exposure time, atmospheric extinction, the air mass, and the overall throughput of the instrument. All of the values of these parameter can be found in Table 2 of Ivezic et al. (2008).

We randomly generate the first observational day from -20 to 35 rest-frame days to ensure that we always have enough data to establish the stretch factor. We next randomly select the data point to occur every 3 or 4 days based on the cadence of the LSST (Ivezic et al. 2008). Finally, about 1000 mock SNIa flux data are generated for each of the five filters.

In Figure 2, we show the examples of the mock light curves in \( g, r, \) and \( i \) bands at different redshifts. The mock flux is created from the Gaussian distribution with the mean on the light curve. Here, we set the first observe day \( t^0_{\text{obs}} = -10 \) in the observer frame.

3. FITTING THE LIGHT CURVE

There are six light curve parameters that we hope to extract from multiwavelength light curve fitting. These parameters are the \( z, m_B, s, E_{B-V}^{\text{host}}, \Delta m, \) and \( t^0_{\text{rest}} \) (i.e., the rest-frame date for the first observe day). The \( \chi^2 \) statistical method is employed here with

\[
\chi^2 = \sum_{i=1}^{t} \sum_{j=1}^{\text{bands}} \frac{(f^\text{obs}_{ij} - f^\text{th}_{ij}(T_j; z, m_B, s, E_{B-V}^{\text{host}}, \Delta m, t^0_{\text{rest}}))}{\sigma^2_{\text{obs}}},
\]

where \( f^\text{obs}_{ij}, f^\text{th}_{ij}, \) and \( \sigma^2_{\text{obs}} = \sigma^2_{\text{phos}} \) are the observed, theoretical flux, and observed error for the observe day \( t_i \) and band \( j \), and \( T_j \) is the transmission of band \( j \). The summation goes through all bands and days with observed samplings of the light curves.

3.1. The Markov Chain Monte Carlo Technique

The best-fit value for each light curve parameter usually can be found using the nonlinear least-squares fitting technique (e.g., Sullivan et al. 2006). Here, considering the number of the
parameters, the efficiency, and the accuracy, we would like to employ the MCMC technique to perform the fitting process. This method does not require to assume the Gaussian distribution for the likelihood, and it is easy to perform the marginalization over other parameters when quoting error for one parameter. Most importantly, it is very efficient for the multiparameter fitting (Neil1993; Lewis & Bridle 2002; MacKay 2003; Doran & Mueller 2004; Gong & Chen 2007; Trotta 2008).

Our purpose is to estimate the posterior probability \( P(\theta|D) \) for the parameter set \( \theta \) given the observational data set \( D \). Based on the Bayes theorem

\[
P(\theta|D) = \frac{\mathcal{L}(D|\theta)P(\theta)}{P(D)},
\]

where \( \mathcal{L}(D|\theta) \sim e^{-\chi^2/2} \) is the likelihood which denotes the probability to get \( D \) given the parameters \( \theta \), \( P(\theta) \) is the prior probability for \( \theta \), and \( P(D) \) is the normalization factor which would not affect our analysis here.

The Metropolis–Hastings algorithm is applied in our MCMC technique to decide if a new point should be accepted by an acceptance probability:

\[
a(\theta_{n+1}|\theta_n) = \min \left\{ \frac{P(\theta_{n+1}|D) q(\theta_n|\theta_{n+1})}{P(\theta_n|D) q(\theta_{n+1}|\theta_n)} \right\},
\]

where \( q(\theta_{n+1}|\theta_n) \) is the proposal density to propose a new point \( \theta_{n+1} \) given a current point \( \theta_n \) in the chain. Here, we assume uniform prior probabilities for the parameters which are canceled in Equation (8). If \( a = 1 \), the new point \( \theta_{n+1} \) is accepted; otherwise, the new point is accepted with probability \( a \). This process is repeated until a new point is accepted, and then we set \( \theta_{n+1} = \theta_{n+1} \). Also, we set a uniform Gaussian-distributed proposal density for every point, so that it is independent of the position on the chain, i.e., \( q(\theta_{n+1}|\theta_n) = q(\theta_n|\theta_{n+1}) \), we then have

\[
a(\theta_{n+1}|\theta_n) = \min \left\{ \frac{\mathcal{L}(D|\theta_{n+1}) q(\theta_n|\theta_{n+1})}{\mathcal{L}(D|\theta_n)} \right\}.
\]

Since the proposal density determines the step size of the MCMC process, it is closely related to the convergence and mixing of the chain. Here, we adopt the adaptive step size Gaussian sampler given by Doran & Mueller (2004). The criterion of the convergence we use was described in Gelman & Rubin (1992), and after convergence we freeze the step size (Doran & Mueller 2004).

The ranges of the parameters when we run the MCMC are set as follows: \( z \in (0, 2) \), \( m_B \in (10, 30) \), \( s \in (0.5, 1.5) \), \( E_{\text{host}}^{\text{rest}} \in (-0.5, 0.5) \), \( \Delta m \in (-0.5, 0.5) \), and \( t^\text{rest} \in (-20, 40) \). For each mock SNeIa, we take about 10,000 chain points to illustrate the probability distribution of the parameters after the burn-in and thinning process.

### 3.2. The Light Curve Fitting Results

In Figure 3, we compare the input redshift of each of our 1000 SNeIa in the simulation with the redshift obtained from MCMC fitting of SEDs to the multiwavelength light curves. We find that when \( z < 0.2 \) the SNeIa light curves are adequately sampled with enough accuracy to allow good redshift estimates along with other parameters, with uncertainties as small as 0.001. Such an error is comparable with the spectroscopic measurements, and even if the spectroscopic measurements could provide a higher precision on the measurement of the redshift, in any case the unknown bulk flows (Cooray & Caldwell 2006; Zhang & Chen 2008) would produce an error on the redshift at this level. For the medium redshift \( 0.2 < z < 0.5 \), the estimated redshift is still useful but the uncertainty is about 0.1. For \( z > 0.5 \), the apparent magnitude becomes large, and since \( \sigma_{\text{phot}} \sim 1^{\text{phot}} \), the redshift errors increase quickly with increasing redshift and can reach \( \sim 1 \). The limitation at high redshift is also due to lack of near-IR photometric coverage and addition of IR bands beyond the \( z \) band will improve photometric determinations when \( z > 0.5 \).

The residuals and 1σ errors for the total six fitting parameters in the MCMC analysis are shown in Figure 4. Similar to the redshift, the multiwavelength light curve model fitting leads to parameter accuracies that are remarkably accurate when \( z < 0.2 \), except for \( \Delta m \) as it acts as an extra source of noise independent of the redshift. As shown in Figure 5, over the whole redshift range studied out to \( z \) of 1.1, the dispersion of the fitting \( z, m_B, s, E_{\text{host}}^{\text{rest}}, \Delta m, \) and \( t^\text{rest} \) are mainly less than ±0.3, ±0.4, ±0.2, ±0.2, ±0.4, and ±4, respectively.

### 3.3. The Constraints on Cosmology

To establish the overall effect of the uncertainty from photometric redshift for cosmological studies, we also generate 1000 SNeIa with spectroscopic redshifts \( z_{\text{spec}} \) with the LSST photometric error \( \sigma_{\text{phot}} \), i.e., we just fix the redshift and only model fit the other five parameters. The Hubble diagram for the...
spectroscopic and photometric cases is shown in Figure 6. Only one-sixth of the whole data are shown in each figure.

We use the MCMC approach to fit the cosmological parameters from the two Hubble diagrams. Two cosmological scenarios are considered, that first one is $\Lambda$CDM with non-flat geometry and the second is $w$CDM with the time-evolved equation of state for the dark energy with $w(z) = w_0 + w_1 z/(1 + z)$.

In Figure 7, we show the contour maps of $\Omega_{m0}$ versus $\Omega_{\Lambda0}$ and $w_0$ versus $w_1$ with 1$\sigma$ and 2$\sigma$ errors. As can be seen, the 1$\sigma$

Figure 4. Residuals for the six parameters in the MCMC analysis with 1$\sigma$ errors for the 1000 SNeIa in the mock sample. The results are pretty good for $z < 0.2$, except for $\Delta m$ since it can be seen as the noise and is independent of the redshift.

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

Figure 5. Distribution of the fitting value minus the actual value for each light curve parameter. The number has been normalized.

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

Figure 6. Hubble diagram for the $z_{\text{spec}}$ and $z_{\text{phot}}$ simulations. The data points on each figure are just one-sixth of the whole data sets.

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

Figure 7. Contour maps for $\Omega_{m0}$ vs. $\Omega_{\Lambda0}$ (left) and $w_0$ vs. $w_1$ (right). The 1$\sigma$ and 2$\sigma$ errors are shown. The red solid and blue dotted contours are for photometric and spectroscopic redshift simulations, respectively.

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

contours using $z_{\text{phot}}$ and its error in cosmological parameter fits nearly overlap with the 2$\sigma$ contours of the case where redshift is known precisely using $z_{\text{spec}}$. Also, we find little deviation for the directions of the main axis of the contours for the two cases. Thus, for a survey such as those planned for LSST, we effectively find a factor of $\sim 2$ degradation in parameter uncertainties when using the SNIa sample with only photometric redshifts compared with one with also spectroscopic redshifts.

In terms of the dark energy figure of merit that involves the inverse area of the $w_0$ versus $w_1$ ellipse, photometric
SNe samples lead to a factor of 4 degradation compared to spectroscopic sample. This difference, however, is likely to be a minor issue: compared with the planned SNeIa surveys that will involve spectroscopic measurements of a few thousand SNe per year, photometric-only surveys such as the one with LSST will produce a sample of a few hundred thousand SNe. Moreover, we have to note that for a real survey the sample is always magnitude limited, so that some \( z > 1.1 \) objects could contaminate the \( z < 1.1 \) sample and lead to a bias in the fitting results. Given that we cannot quantify this bias fraction, we don not include this effect in our analysis. In an upcoming paper, we hope to implement a new technique to account for such biases in large SNe samples.

### 4. Removing the Contamination from SNIb/c and SNII

In a pure photometric survey such as the one with LSST without spectroscopic measurements to identify if each of the SNe is Type Ia or not, in addition to the error in the measurement of redshift, the photometric SNe samples would also be contaminated by core-collapse SNe. Here, we consider the contamination from Type Ib/c (SNIb/c) and SNII.

#### 4.1. Estimating the Contamination

To estimate the level of contamination, we create mock samples of light curve data for SNIb/c and SNII. The spectral templates for SNIb/c are from Levan et al. (2005), and for SNII we use the templates of SNIIP and SNIIIL given by Gilliland et al. (1999) and Baron et al. (2004). Here, for simplicity, we consider primarily the SNIIP and SNIIIL. The SNIIn which have “unusual” progenitors (Mobberley 2007) may be an important contamination for the SNIa (Poznanski et al. 2007a). We may discuss these objects in future work. We set the percentage of SNIIP and SNIIIL as 50% and 50%, respectively, for the SNII sample. The mock flux data of the SNIb/c and SNII samples are generated with the same procedure as those used to generate the SNIa mock data in Section 2.

Since the core-collapse SNe are intrinsically fainter than the SNeIa and have no magnitude–phase relation, we take the absolute peak magnitude from Richardson et al. (2002) and the Gaussian distribution with \( \bar{z} = 0.4 \) for SNIb/c and SNII, and then set \( s = 1 \) when mimicking their \( B \)-band peak magnitude. Also, given that the core-collapse SNe are usually found in star-forming regions, they are expected to suffer more extinction from the host galaxy. We set \( -0.2 < E_{B-V}^{\text{host}} < 0.4 \) with \( E_{B-V} = 0.0 \) and \( \sigma_E = 0.3 \). Once simulated, we continue to use the SNIa SED light curves to fit them just as we did in Section 3.

The distribution of the difference of \( \chi^2 \) (i.e., relative \( \chi^2 \)) for the SNIa, SNIb/c, and SNII is shown in Figure 8. We find that when fitted with Ia SED light curves, the \( \chi^2 \) values for SNeII are so large that they are easily distinguished from the SNIa even with photometric data alone. However, for SNeIb/c the \( \chi^2 \) peak overlaps with that of the SNeIa, so they are the cause of primary contamination to the total sample.

To obtain a less-contaminated sample, as a first cut we note that the \( \chi^2 \) distribution of the SNeIb/c has a long tail which can extend to tens of thousands, and some of the SNIb/c can be removed by an overall restriction on the \( \chi^2 \) values in the fitting to SNIa light curve template. If the selection is restricted to \( \chi_{\text{min}}^2 < 20 \), keeping all real Type Ia’s, this results in a removal of about 40% of the SNeIb/c’s. Since the ratio of the rate of SNIa to SNIb/c out to \( z \sim 1 \) is about 10 to 7 (Calura & Matteucci 2006; de Plaa et al. 2007; Sato et al. 2007; Poznanski et al. 2007b; Eldridge et al. 2008; Smartt et al. 2009; Georgy et al. 2009), we expect about 250 SNeIb/c to remain and contaminate a sample that contains 1000 SNeIa selected photometrically. Note that here we just use a simplified assumption with the ratio of the rate of SNIa to SNIb/c to be redshift independent. In the next section, we discuss a statistical method to further reduce the contamination of Ib/c during model fits to the total sample.

#### 4.2. The Bayesian Statistical Method

We employ the Bayesian estimation method proposed by Press (1996) and Kunz et al. (2007) to further reduce the contamination from SNIb/c. We note that our proposed statistical method cannot distinguish each SNIb/c from a Type Ia individually, but statistically it reduces the overall contamination and the associated bias in cosmological parameters. As we illustrate here, the same method also allows us to jointly estimate the fraction of the SNIa (or Ib/c’s) within the whole SNe sample used for cosmology.

We take the case that the observational sample of supposedly Type Ia’s \( \mathbf{D} \) contains a mixture of true SNeIa data \( \mathbf{D}_0 \) and SNeIb/c’s \( \mathbf{D}_b \) which mimics Ia. We define a vector \( \mathbf{v} \) of length the total number of SNeIa \( N \) with the value of \( v_i \) taking either 1 or 0 if \( D_i \) is or not a SNIa. We also define a quantity \( p \) to account for the total fraction of the true SNeIa in the total SNe data set \( \mathbf{D} \). Using \( \mathbf{v} \) and \( p \), the posterior probability can be written as

\[
P(\theta | \mathbf{D}) = \sum_{p, \mathbf{v}} P(\theta, \mathbf{v}, p | \mathbf{D})
\]

\(^{10}\) http://supernova.lbl.gov/~nugent/nugent_templates.html
\[ \propto \sum_{p,n} \mathcal{L}(\theta, v, p) P(\theta, v, p) \]  

(11)

\[ \propto \sum_{p,n} \mathcal{L}(\theta, v, p) P(p) P(\theta|p) P(v|\theta, p). \]  

(12)

In Equation (10), the sum over \( p \) will be the integration if the value of \( p \) is continuous, and the sum of \( v \) goes through all \( 2^N \) possible values of \( v \). Equation (11) is derived from the Bayes theorem, and \( \mathcal{L}(\theta, v, p) \) is the likelihood. The \( P(\theta|p) \) in Equation (12) can be reduced to \( P(\theta) \) since there is no reason to believe the parameter \( p \) affect the cosmological evolution of the universe (it is not a cosmological parameter).

Thus, we can simplify to

\[ P(\theta|\mathbf{D}) \propto P(\theta) \sum_{p} P(p) \sum_{v} \mathcal{L}(\theta, v, p) P(v|\theta, p). \]  

(13)

For any value of \( p \), \( P(v|\theta, p) \) is 0 when \( v_i \) involving the \( i \)th datum is a SNIb/c. When normalized, \( P(v_i = 1|\theta, p) = p \) and \( P(v_i = 0|\theta, p) = 1 - p \). Therefore, we find

\[ P(\theta|\mathbf{D}) \propto P(\theta) \sum_{p} P(p) \sum_{v} \left[ \prod_{v_i=1} \mathcal{L}_i^1 p + \prod_{v_i=0} \mathcal{L}_i^0 (1 - p) \right]. \]  

(14)

Here \( \mathcal{L}_i^1 \) is the likelihood that the \( i \)th SN is a Ia and this is taken to be

\[ \mathcal{L}_i^1 = \frac{1}{\sqrt{2\pi} \sigma_i} e^{-\chi_i^2/2}, \]  

(15)

where \( \sigma_i \) is the error and \( \chi_i^2 = (\mu_i^{\text{obs}} - \mu_i^{\text{th}})^2/\sigma_i^2 \), and the \( \mu_i^{\text{obs}} \) and \( \mu_i^{\text{th}} \) are the observational and theoretical distance moduli, respectively.

The \( \mathcal{L}_i^0 \) is the likelihood for the SNIb/c samples though we don not know a priori the exact distribution. We use two parameters \( b \) and \( \sigma_0 \) belonging to the parameter set \( \theta \) to describe \( P(\theta) \) as

\[ \mathcal{L}_i^0 = \frac{1}{\sqrt{2\pi} \sigma_0} e^{-\chi_i^2/2}, \]  

(16)

where \( \chi_i^2 = (\mu_i^{\text{obs}} - \mu_i^{\text{th}} - b)^2/\sigma_0^2 \).

We can simplify Equation (14) further by noting that the \( 2^N \) summation term can be written as the product of \( N \) terms. We finally get

\[ P(\theta|\mathbf{D}) \propto P(\theta) \sum_{p} P(p) \prod_{i} \left[ \mathcal{L}_i^1 p + \mathcal{L}_i^0 (1 - p) \right]. \]  

(17)

The sum over \( p \) is easily performed with MCMC runs, and we assume \( P(p) \) is a uniform distribution. When analyzing our mock samples we take the ranges for \( p, b \), and \( \sigma_0 \) of \( p \) \( \in \) (0.5, 1), \( b \) \( \in \) (−20, 20), and \( 1/\sigma_0 \) \( \in \) (0, 1000).

4.3. The Results

Extending the discussion in Section 4.1, we add 250 SNIb/c data with \( \chi_{\theta}^2 < 20 \) to the 1000 SNIa data, and extract cosmological constraints on the time-evolving equation of state of dark energy with an analysis which implements the Bayesian estimation method described above, in addition to a method where all data are analyzed with a MCMC run without making an attempt to account for Ib/c contamination to the total sample.

We show the contour maps of \( w_0 \) versus \( w_1 \) in Figure 9. The red solid and blue dotted contours are the fitting results with and without the Bayesian estimation, respectively. As can be seen in Figure 9, the constraint on \( w_0 \) and \( w_1 \) is completely wrong when we ignore the Ib/c contamination in the total sample and just do direct fitting to the Hubble diagram. This is caused by the large differences of the distribution between the SNIa and the SNIb/c data and MCMC chains are easily trapped in a wrong likelihood value. The \( \chi^2 \) value for the overall best fit in this case is also very large reaching as high as 2000.

When we implement the Bayesian estimation method, the result is improved significantly. Although there is a difference between the best-fit and the actual (fiducial) value, the fiducial value of the cosmological parameter set \((w_0, w_1) = (-1, 0)\) lies safely within the 2\( \sigma \) contour around the best fit. Also, as discussed, we also jointly estimate the fraction of SNeIa in the total data set. We plot the likelihood for \( P(p) \) in Figure 10. The fraction of the SNIas in this particular mock data should be 80\% while the fitting leads to the result of \( p = 0.75^{+0.05}_{-0.05} \) with errors at 1\( \sigma \). While there still remains a bias associated with the contaminating Type Ib/c’s, we have reduced this bias to the level of a few percent. Also, if the distributions of \( b \) and \( \sigma_0 \) are better measured in the future, this method will get better typing result.

We believe that the Bayesian estimation method provides a useful statistical tool to reduce contamination and to evaluate the fraction of the contaminating SNe in the LSST photometric SNIa survey. Of course, to identify whether each individual SN is a core-collapse one or a Ia, the method discussed above is inadequate, but the Bayesian statistical analysis can be a valuable guide for further advanced study (Poznanski et al. 2002, 2007a; Gal-Yam et al. 2004; Johnson & Crotts 2006; Kuznetsova & Connolly 2007).
In this paper, we explore the ability to determine the redshift and other parameters useful to construct the Hubble diagram with light curve for the LSST SNIa photometric measurements. Using a SNIa SED template and the expected photometric error of the LSST, we first simulate the observed flux data of 1000 SNIa in each of five LSST filters, and then apply the MCMC technique to fit the redshift, stretch factor, apparent magnitude, and the phase of the SNIa, among others. We find that when $z < 0.2$, these parameters can be determined accurately at a level comparable to the case where spectroscopic redshift is known. At higher redshifts, the uncertainty in photometric redshift goes up quickly since $\sigma_{\text{phot}} \sim 10^{0.4m}$, but the photometric data are still very useful when $0.2 < z < 0.5$. To illustrate the effect of the uncertainty of the photometric redshift on the fitting of the cosmological parameters, we also extract cosmological constraints using parameters of the SNIa light curves with and without spectroscopic redshifts. Using the fitting results of the two cases, we constrain the cosmological parameters for $\Omega_{m0}$ and $\Omega_{\Lambda0}$ in the $\Lambda$CDM model and $w_0$ and $w_1$ in the time-evolved wCDM model. We find that for the same number of SNIa data, the cosmology fitting with only the photometric data leads to a factor of 2 degradation in error of cosmological parameters or a factor of 4 in the figure of merit of dark energy equation of state (i.e., the inverse area of the $w_0 - w_1$ ellipse) compared with the case of fitting with spectroscopic data. However, as the number of photometric-only data far exceeds that with spectroscopic data, the overall statistical uncertainty in the former would still be smaller.

Finally, we discuss the contamination on the SNIa data from core-collapse SNe involving types II and Ib/c, and the feasibility of using the Bayesian estimation statistical method to reduce the overall contamination. Similar to SNIa mock samples, we generate the mock flux data for the SNIb/c and SNII based on their spectral templates, and use the SNIa fitting process to fit them. We find that the SNIbII are easily distinguished from SNIa because there is an apparent mismatch (large $\chi^2$) when fitting with the SNIa templates. However, this is not the case for Type Ib/c’s. The peak of its $\chi^2$ distribution is overlapping with that of the SNIa and presents a significant contamination of any photometric selected supposedly SNIa samples, even if a conservative cut is applied in the $\chi^2$ values for selection. To further account for this contamination, at least statistically when doing cosmological model fits, we employ Bayes theorem. Our suggested method could reduce the contamination down to a few percent level, leading to estimates of cosmological parameters that are biased within 1σ errors. The method also establishes the fraction true SNIa in the total photometric SNe data set. Nevertheless, we must note that this method cannot distinguish if an individual SN is whether Type Ia or not. We will need an extended analysis complemented with additional observations if we are required to recognize the type of an individual SN. 

5. SUMMARY

This research is supported by the NSF under CAREER AST-0645427 at UCI, by the NSFC under the Distinguished Young Scholar grant 10525314, the Key Project grant 10533010, by the Chinese Academy of Sciences under grant KJCX3-SYW-N2, and by the Ministry of Science and Technology of China under the National Basic Science program (project 973) grant 2007CB815401. X.C. also acknowledges the hospitality of the Moore Center of Theoretical Cosmology and Physics at Caltech, where part of this research is performed.

REFERENCES

Astier, P., et al. 2006, A&A, 447, 31
Baron, E., et al. 2004, ApJ, 616, 91
Bessell, M. S. 1990, PASP, 102, 1181
Bessell, M. S. 2005, ARA&A, 43, 293
Burstein, D., & Heiles, C. 1982, AJ, 87, 1165B
Calura, F., & Matteucci, F. 2006, ApJ, 652, 889
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Coe, R., & Caldwell, R. R. 2006, Phys. Rev. D, 73, 103002
de Plaa, J., et al. 2007, A&A, 465, 345
Doran, M., & Mueller, C. M. 2004, J. Cosmol. Astropart. Phys., JCAP09(2004)003
Eldridge, J. J., Izzard, R. G., & Tout, C. A. 2008, in IAU Symp. 250, Massive Stars as Cosmic Engines, ed. F. Bresolin, P. A. Crowther, & J. Puls (Cambridge: Cambridge Univ. Press), 179
Fal-Yam, A., et al. 2004, PASP, 116, 597
Gelman, A., & Rubin, D. 1992, Stat. Sci., 7, 457
Goldhaber, G., et al. 2001, ApJ, 558, 457
Gong, Y., & Chen, X. 2007, Phys. Rev. D, 76, 123007
Guy, J., Astier, P., Nobili, S., Regnault, N., & Pain, R. 2005, A&A, 443, 781
Hamuy, M., Phillips, M. M., Suntzeff, N. B., Schommer, R. A., & Aviles, R. 1995, AJ, 109, 1
Hamuy, M., Phillips, M. M., Suntzeff, N. B., Schommer, R. A., & Aviles, R. 1996, AJ, 112, 2391
Hicken, M., et al. 2009, ApJ, 700, 1097
Howell, D. A., et al. 2008, Type Ia Supernova Science 2010–2020, White Paper, submitted (arXiv:0903.1086)
Ivezić, Z., et al. 2008, AJ, 176, 1
Jha, S., Riess, A. G., & Kirshner, R. P. 2007, ApJ, 659, 122
Johnson, B. D., & Crotts, A. P. S. 2006, AJ, 132, 756
Kessler, R., et al. 2009, ApJS, 185, 32
Kim, A., & Miquel, R. 2006, Astropart. Phys., 24, 451
Kim, A., & Miquel, R. 2007, ApJ, 663, 100
Kowalski, M., et al. 2008, ApJ, 686, 749
Kunz, M., Bassett, B. A., & Hlozek, R. 2007, Phys. Rev. D, 75, 103508
Kuznetsova, N. V., & Connolly, A. J. 2007, ApJ, 659, 122
Leibundgut, B. 2001, ARA&A, 39, 67
Levan, A., et al. 2005, ApJ, 624, 880
