A&A 541, A110 (2012)
DOI: 10.1051/0004-6361/201118599

Astronomy & Astrophysics

Type Ia supernova parameter estimation: a comparison of two approaches using current datasets

B. L. Lago[1,2], M. O. Calvão[1], S. E. Jorás[1], R. R. R. Reis[1], I. Waga[1], and R. Giostri[1,3]

[1] Instituto de Física, Universidade Federal do Rio de Janeiro, C.P. 68528, CEP 21941-972, Rio de Janeiro, RJ, Brasil
e-mail: [brunolz;orca;joras;ribamar;ioav;rgiostri]@if.ufrj.br
[2] Coordenação de Licenciatura em Física, Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, CEP 28635-000, Nova Friburgo, RJ, Brasil
[3] Departamento de Engenharia Rural, Universidade Federal do Espírito Santo, C.P. 16, CEP 29500-000, Alegre, ES, Brasil

Received 6 December 2011 / Accepted 21 March 2012

ABSTRACT

We compare the traditional χ^2 and complete-likelihood approaches for determining parameter constraints from type Ia supernovae (SNe Ia) when the magnitude dispersion is to be estimated as well. The dataset we used was sample combination "e" from Kessler (2009, ApJS, 185, 32), comprising the first-year SDSS-II Supernova Survey together with ESSENCE, SNLS, HST, and a compilation of nearby SNe Ia. We considered cosmological constant + cold dark matter (ΛCDM) and spatially flat, constant w dark energy + cold dark matter (FwCDM) cosmological models and show that, for current data, there is a small difference in the best-fit values and a difference of about 30% in confidence contour areas if the MLCS2k2 light-curve fitter is adopted. For the SALT2 light-curve fitter the differences in area are less significant ($\lesssim$13%). In both cases the likelihood approach gives more restrictive constraints. We argue for using the complete likelihood instead of the χ^2 approach when dealing with parameters in the expression for the variance.

Key words. cosmological parameters – supernovae: general – methods: statistical

1. Introduction

Observations of type I supernovae (SNe Ia) directly established in the late 1990s that the universe is accelerating (Riess et al. 1998; Perlmutter et al. 1999). This most remarkable result is possibly only surpassed by the discovery in the 1930s that the universe expands. The SNe Ia are still the backbone for the prevailing ΛCDM concordance model, which is also corroborated by the combination of other probes, such as cosmic microwave background anisotropies, baryon acoustic oscillations and galaxy clustering (Komatsu et al. 2011; Eisenstein et al. 2005; Percival et al. 2010; Vikhlinin et al. 2009; Mantz et al. 2010).

Already in the landmark papers from the two original surveying groups, a legitimate concern was expressed about not properly accounted for systematic effects. At that time, however, there were few (on the order of tens) observed SNe Ia in any of the available samples, so the uncertainties were statistically limited; the main task was gathering new, larger, and uniform datasets. For the recent compilations and surveys (Wood-Vasey et al. 2007; Kowalski et al. 2008; Hicken et al. 2009; Conley et al. 2011; Kessler et al. 2009; Amanullah et al. 2010), and more so for even future ones (Pan-STARRS; DES; LSST), we are no longer sample-limited; the urgency is again on the physics of the SNe Ia phenomenon and all the aspects that impact it (for review see, for instance, Kirshner 2010; Howell 2011; Goobar & Leibundgut 2011). Despite the major importance of investigating the systematics, here we focus on a distinct aspect, namely the statistically consistent determination of the parameters when the variance is to be estimated as well.

In contrast to the traditional χ^2 method in most of the literature, we here discuss a different approach to parameter fitting with unknown variance, which is based on the likelihood method. We recall that, if we aim to estimate both the covariance and the mean of a Gaussian process, the ordinary (uncorrected) χ^2 approach (or any iterative recipe thereof, for that matter) cannot be straightforwardly applied, because we might miss a nontrivial term in the objective function to be extremized. This is necessary when the covariance itself depends on free parameters of the underlying model. Recently, similar criticisms to the traditional χ^2 method were presented by Kim (2011, see also Vishwakarma & Narlikar 2010). Kim (2011) used a likelihood approach to simulated data for the distance moduli, and he focused on determining the intrinsic dispersion, without reference to a particular light-curve fitter. We reconsider the problem in the context of the MLCS2K2 and SALT2 light-curve fitters with real data. We point out that, in principle, with SALT2, additional difficulties may arise owing to the astrophysical parameters. We compare the results of the traditional χ^2 treatment to those of the likelihood method. We restrict ourselves to SNe Ia datasets only, without taking into account other important sources of information (such as CMB, BAOs or clusters), to make our point pristine, by avoiding the masking due to any other such probes.

The structure of the paper is as follows. In Sect. 2, we briefly recall the general procedure to obtain the final estimated parameters from the raw data. In Sect. 3, we describe the two most widely used "fitter pipelines" (MLCS2k2 and SALT2), which use the traditional χ^2 approach. In Sect. 4, we critically review the aforementioned traditional approach and present the likelihood-based method. In Sect. 5, the main numerical results are shown for both fitters for data from the first-year SDSS-II Supernova Survey together with other supernova data sets (ESSENCE, SNLS, HST, and a compilation of nearby type Ia supernovae), which constitute the sample combination "e" as described in Kessler et al. (2009). Finally, in Sect. 6, we end with some discussions and conclusions.

2. Light-curve fitting

As mentioned in Sect. 1, SNe Ia are standardizable candles; this means that there is a phenomenological recipe according to which the raw light curves, after being subjected to a transformation by an $N_{\rm st}$-parameter function, furnish a new set of so-called standardized light curves; in other words the dispersion in the new magnitudes is considerably lower than in the original input set. This phenomenological recipe is not unique at all: there are several light-curve fitters in the literature (Jha et al. 2007; Wang et al. 2003, 2005; Guy et al. 2005, 2007; Conley et al. 2008; Burns et al. 2011). Here we will exploit some features of the two most common ones: the multicolor light-curve shape (MLCS2k2) (Jha et al. 2007) one and the spectral adaptive light-curve template (SALT2) (Guy et al. 2007).

In a nutshell:

- The MLCS2k2 fitting model (Jha et al. 2007) describes the variation among SNe Ia light curves with a single parameter (Δ). The MLCS2k2 theoretical magnitude, observed in an arbitrary filter Y, at an epoch γ, is given by (Kessler et al. 2009)

$$m^{\rm th}_{Y,\gamma} = M_{Y',\gamma} + p_{Y',\gamma}\Delta + q_{Y',\gamma}\Delta^2 + K_{Y'Y,\gamma} + \mu + X^{\rm host}_{Y',\gamma} + X^{\rm MW}_{Y}, \quad (1)$$

where $Y' \in \{U, B, V, R, I\}$ is one of the supernova rest-frame filters for which the model is defined, Δ is the MLCS2k2 shape-luminosity parameter that accounts for the correlation between peak luminosity and the shape/duration of the light curve, $X^{\rm host}_{Y',\gamma}$ is the host-galaxy extinction, $X^{\rm MW}_{Y}$ is the Milky Way extinction, $K_{Y'Y,\gamma}$ is the K-correction between rest-frame and observer-frame filters, and μ is the distance modulus. The coefficients $M_{Y',\gamma}$, $p_{Y',\gamma}$, and $q_{Y',\gamma}$ are model vectors that have been evaluated using nearly 100 well observed low-redshift SNe Ia as a training set. Above, $\gamma = 0$ labels quantities at the B-band peak magnitude epoch.
Fitting the model to the magnitudes of each SN Ia, usually fixing R_V, gives μ, Δ, A_V and t_0, the B-band peak magnitude epoch; for more details, see (Kessler et al. 2009).
- The SALT2 fitter (Guy et al. 2007) uses a two-dimensional surface in time and wavelength that describes the temporal evolution of the rest-frame spectral energy distribution (SED) or specific flux for SNe Ia. The original model was trained on a set of combined light-curves and 303 spectra, not only from (very) nearby but also medium- and high-redshift SNe Ia.
In SALT2, the rest-frame specific flux at wavelength λ and phase (time) p ($p = 0$ at B-band maximum) is modeled by

$$\phi(p, \lambda; x_0, x_1, c) = x_0[M_0(t, \lambda) + x_1 M_1(t, \lambda)] \exp[cC(\lambda)] \quad (2)$$

and does depend through the parameters x_0, x_1, and c on the particular type Ia supernova. $M_0(t, \lambda)$, $M_1(t, \lambda)$, and $C(\lambda)$ are determined from the training process described in Guy et al. (2007). $M_0(t, \lambda)$, $M_1(t, \lambda)$ are the zeroth and the first moments of the distribution of training sample SEDs. One might consider adding moments of higher order to Eq. (2).
To compare with photometric SNe Ia data, the observer-frame flux in passband Y is calculated as

$$F^Y(p(1+z)) = (1+z) \int {\rm d}\lambda' \left[\lambda' \phi(p, \lambda') T^Y(\lambda'(1+z))\right], \quad (3)$$

where $T^Y(\lambda)$ defines the transmission curve of the observer-frame filter Y, and z is the redshift.
Each SN Ia light curve is fitted separately using Eqs. (2) and (3) to determine the parameters x_0, x_1, and c with corresponding errors. However, the SALT2 light-curve fit does not yield an independent distance modulus estimate for each SN Ia. As we will see in the next section, the distance moduli are estimated as part of a global parameter fit to an ensemble of SN Ia light curves in which cosmological parameters and global SN Ia properties are also obtained.

In the next section we will discuss how to obtain constraints on cosmological parameters using MLCS2k2 and SALT2 output quantities as our data.

3. The traditional χ^2 approach

The prevailing SNe Ia cosmological analysis is based on the χ^2 function:

$$\chi^2 := \boldsymbol{X}^{\rm T}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}, \quad (4)$$

where $\boldsymbol{X} := (\boldsymbol{\mu} - \boldsymbol{\mu}_{th}(\boldsymbol{z}, \boldsymbol{\Theta}))$, $\boldsymbol{\mu}$ is the set of distance moduli derived from the light-curve fitting procedure for each SNe Ia event, at redshifts given by $\boldsymbol{z}$, $\boldsymbol{\mu}_{th}(\boldsymbol{z}, \boldsymbol{\Theta})$ is the theoretical prediction for them, given in terms of a vector $\boldsymbol{\Theta}$ of parameters and $\boldsymbol{\Sigma}$ is the covariance matrix of the events.

3.1. The χ^2 approach from SALT2 output

The SALT2 light-curve fitter gives three quantities, with corresponding errors, to be used in the analysis of cosmology:

$$m^*_B := -2.5 \log\left[x_0 \int {\rm d}\lambda' M_0(p = 0, \lambda') T^B(\lambda')\right], \quad (5)$$

to be interpreted as the peak rest-frame magnitude in the B band, x_1, a parameter related to the stretch of the light-curve and c, related to the color of the supernova alongside the redshift z of the supernova. The distance modulus is modeled, in this context, as a function of (m^*_B, x_1, c) and two new parameters, $\boldsymbol{\delta} := (\alpha, \beta)$, plus the peak absolute magnitude, in B band, M_B. Defining the corrected magnitude as

$$m^{\rm corr}_B(\boldsymbol{\delta}) := m^*_B + \alpha x_1 - \beta c, \quad (6)$$

we can write

$$\mu(\boldsymbol{\delta}, M_B) = m^{\rm corr}_B(\boldsymbol{\delta}) - M_B. \quad (7)$$

Assuming that all SN Ia events are independent, one can rewrite Eq. (4) as

$$\chi^2_{\rm SALT2}(\boldsymbol{\theta}, \boldsymbol{\delta}, \mathcal{M}(M_B, h)) = \sum_{i=1}^{N} \frac{[\mu_i(\boldsymbol{\delta}, M_B) - \mu_{\rm th}(z_i; \boldsymbol{\theta}, h)]^2}{\sigma_i^2(\boldsymbol{\delta}) + \sigma^2_{\rm int}}, \quad (8)$$

where N is the number of SNe Ia in the sample, $\boldsymbol{\theta}$ denotes the cosmological parameters other than h, with the present value of the Hubble parameter given by $H_0 = 100\, h$ km s^{-1} Mpc^{-1}. The theoretical distance modulus is given by

$$\mu_{\rm th}(z; \boldsymbol{\theta}, h) = 5 \log[\mathcal{D}_{\rm L}(z; \boldsymbol{\theta})] + \mu_0(h), \quad (9)$$

with

$$\mu_0(h) := 5 \log\left(\frac{10^3 c/({\rm km\, s^{-1}})}{h}\right) \simeq 42.38 - 5 \log h. \quad (10)$$

The dimensionless luminosity distance (in units of the Hubble distance today), $\mathcal{D}_\mathrm{L}$, for comoving observers in a Robertson-Walker universe is given by

$$\mathcal{D}_\mathrm{L}(z;\boldsymbol{\theta}) = \begin{cases} (1+z)\left(\frac{1}{\sqrt{\Omega_{k0}}}\right)\sinh\left(\sqrt{\Omega_{k0}}\int_{z'=0}^{z}\frac{1}{E(z';\boldsymbol{\theta})}\mathrm{d}z'\right), & \text{if } \Omega_{k0}>0,\\ (1+z)\int_{z'=0}^{z}\frac{1}{E(z';\boldsymbol{\theta})}\mathrm{d}z', & \text{if } \Omega_{k0}=0,\\ (1+z)\left(\frac{1}{\sqrt{-\Omega_{k0}}}\right)\sin\left(\sqrt{-\Omega_{k0}}\int_{z'=0}^{z}\frac{1}{E(z';\boldsymbol{\theta})}\mathrm{d}z'\right), & \text{if } \Omega_{k0}<0, \end{cases} \quad (11)$$

where Ω_{k0} is the "curvature density parameter", such that it is proportional to the three-curvature, and

$$E(z;\boldsymbol{\theta}) := H(z;\boldsymbol{\theta},h)/H_0 \quad (12)$$

is the dimensionless Hubble parameter. As indicated in Eq. (8), the χ^2 function depends on the parameters M_B and h only through their combination

$$\mathcal{M}(M_B,h) := M_B + \mu_0(h). \quad (13)$$

It may thus be thought of as effectively directly dependent on only the parameters $\boldsymbol{\theta}$, $\boldsymbol{\delta}$ and $\mathcal{M}$.

A floating dispersion term, σ_int, is added in quadrature to the distance modulus dispersion, which is given by

$$\sigma_i^2(\boldsymbol{\delta}) = \sigma^2_{m_B^*,i} + \alpha^2\sigma^2_{x_1,i} + \beta^2\sigma^2_{c,i} + 2\alpha\sigma_{m_B^*,x_1,i} - 2\beta\sigma_{m_B^*,c,i} - 2\alpha\beta\sigma_{x_1,c,i} + \sigma^2_{\mu,z,i}, \quad (14)$$

where $\sigma^2_{\mu,z,i}$ is the contribution to the distance modulus dispersion owing to redshift uncertainties from peculiar velocities and also from the measurement itself. Following Kessler et al. (2009), we will model this contribution for simplicity with the distance-redshift relation for an empty universe, which gives

$$\sigma_{\mu,z,i} = \sigma_{z,i}\left(\frac{5}{\ln 10}\right)\frac{1+z_i}{z_i(1+z_i/2)}, \quad (15)$$

with

$$\sigma^2_{z,i} = \sigma^2_{\mathrm{spec},i} + \sigma^2_\mathrm{pec}, \quad (16)$$

where $\sigma_{\mathrm{spec},i}$ is the redshift measurement error, and $\sigma_\mathrm{pec} = 0.0012$ is the redshift uncertainty due to peculiar velocity.

As advocated by some groups (Astier et al. 2006), minimizing Eq. (8) gives rise to a bias towards increasing values of α and β. To circumvent this feature, an iterative method is employed (Astier et al. 2006).

In this iterative method, the χ^2 presented in Eq. (8) is replaced by

$$\chi^2_\mathrm{SALT2}(\boldsymbol{\theta},\boldsymbol{\delta},\mathcal{M}) = \sum_{i=1}^{N}\frac{[\mu_i(\boldsymbol{\delta},M_B)-\mu_\mathrm{th}(z_i;\boldsymbol{\theta},h)]^2}{\sigma_i^2(\boldsymbol{\eta})+\sigma^2_\mathrm{int}}. \quad (17)$$

In this expression, $\boldsymbol{\eta}$ is not a parameter of the χ^2_SALT2. To obtain the best-fit values for the parameters, $\boldsymbol{\eta}$ is given initial values and the optimization is performed on $\boldsymbol{\theta}, \boldsymbol{\delta}$ and $\mathcal{M}$. After this step, $\boldsymbol{\eta}$ is updated with the best-fit value of $\boldsymbol{\delta}$ and the optimization is performed again. The process continues until a convergence is achieved, which means that $\boldsymbol{\eta}$ does not change under the required precision.

During this process σ_int is not considered as a free parameter to be optimized, being determined rather by the following procedure: start with a guess value (usually $\sigma_\mathrm{int} = 0.15$). Perform the iterative procedure described above. The value of σ_int is then obtained by fine-tuning it so that the reduced χ^2 equals unity (with all the other parameters fixed on their best-fit values). The iterative procedure is repeated once more with this new value and the final best-fit values are obtained. It is important to note that the value of σ_int affects both the best fit and the confidence levels of the parameters, since it changes the weight given to each supernova in the χ^2 (cf. Eq. (17)).

3.2. The χ^2 approach from MLCS2k2 output

The MLCS2k2 light-curve fitter gives a direct estimation of the distance modulus. In this context, the analogue of Eq. (8) is

$$\chi^2_\mathrm{MLCS2k2}(\boldsymbol{\theta},h) = \sum_{i}^{N}\frac{[\mu_i-\mu_\mathrm{th}(z_i;\boldsymbol{\theta},h)]^2}{\sigma_i^2+\sigma^2_\mathrm{int}+\sigma^2_{\mu,z,i}}, \quad (18)$$

where σ_i is the distance modulus dispersion as given by MLCS2k2.

The procedure to obtain σ_int is similar to the one described in the previous subsection, however, here we use only a subsample of nearby SNe Ia and not the full one, as for the SALT2 analysis. After setting up the value of σ_int, we minimize the χ^2_MLCS2k2 using the full SNe Ia sample to obtain the best-fit values for $\boldsymbol{\theta}$ and h.

4. The proposed likelihood approach

Considering the SNe Ia light-curve fitting parameters as Gaussian distributed random variables, we propose to take as starting point the likelihood

$$L = \frac{1}{\sqrt{(2\pi)^N\det\boldsymbol{\Sigma}}}\exp(-\boldsymbol{X}^\mathrm{T}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}/2), \quad (19)$$

which is related to the χ^2 in Eq. (4) by

$$\mathcal{L} := -2\ln L = \chi^2 + \ln\det\boldsymbol{\Sigma} + N\ln(2\pi). \quad (20)$$

Equations (19) and (20) are the single basis upon which our whole statistical procedure lies. When the full covariance of the problem is known, minimizing χ^2 is completely equivalent to minimizing $\mathcal{L}$. However, this is not the case for current SNe Ia observations and neglecting the last but one term in Eq. (20) would, in principle, lead to a biased result. Our proposal is to minimize the following functions for each case discussed in the previous section

$$\mathcal{L}_\mathrm{SALT2}(\boldsymbol{\theta},\boldsymbol{\delta},\mathcal{M},\sigma_\mathrm{int}) = \chi^2_\mathrm{SALT2}(\boldsymbol{\theta},\boldsymbol{\delta},\mathcal{M},\sigma_\mathrm{int}) + \sum_{i}^{N}\ln(\sigma_i^2(\boldsymbol{\delta})+\sigma^2_\mathrm{int}) \quad (21)$$

and

$$\mathcal{L}_\mathrm{MLCS2k2}(\boldsymbol{\theta},h,\sigma_\mathrm{int}) = \chi^2_\mathrm{MLCS2k2}(\boldsymbol{\theta},h,\sigma_\mathrm{int}) + \sum_{i}^{N}\ln(\sigma_i^2+\sigma^2_\mathrm{int}), \quad (22)$$

where we neglected parameter-independent terms. $\chi^2_\mathrm{SALT2}(\boldsymbol{\theta},\boldsymbol{\delta},\mathcal{M},\sigma_\mathrm{int})$ and $\chi^2_\mathrm{MLCS2k2}(\boldsymbol{\theta},h,\sigma_\mathrm{int})$ are given, respectively, by Eqs. (8) and (18) now considering σ_int *also as a free parameter*. With this procedure, we can directly obtain unbiased probability distributions functions for all parameters, including σ_int and $\boldsymbol{\delta}$.

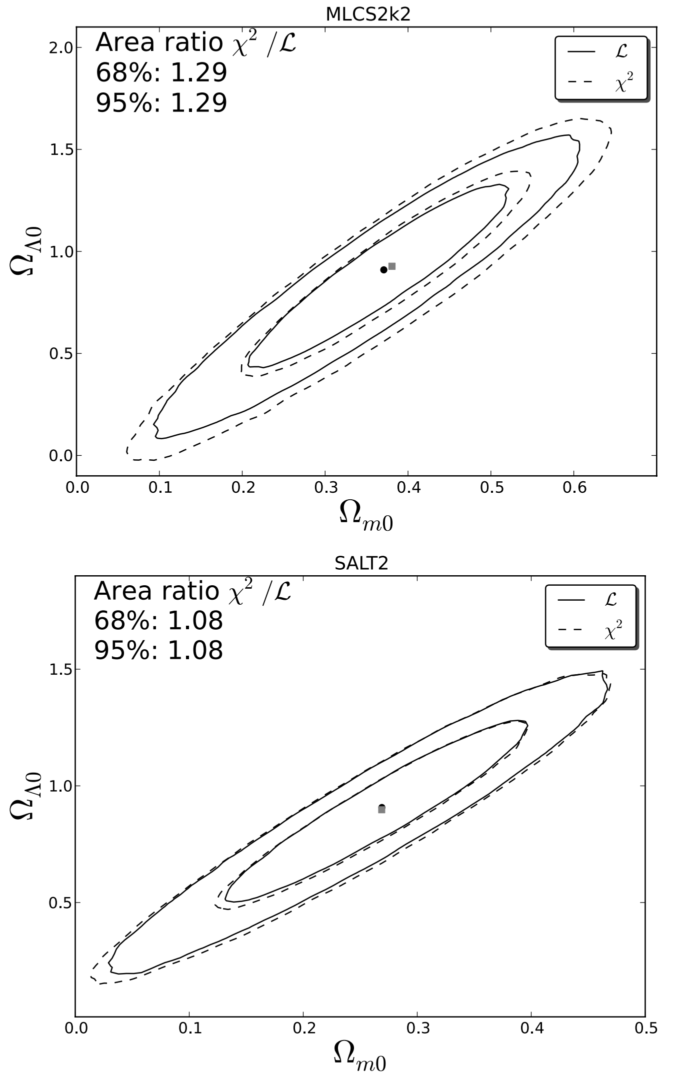

Fig. 1. Contours of 68% and 95% confidence level in the plane $\Omega_{m0} \times \Omega_{\Lambda 0}$ for the ΛCDM model. We marginalized over all other parameters with flat prior. The solid (dashed) lines are the results for the likelihood (χ^2) approach. The black circle (gray square) is the best fit value for the likelihood (χ^2). *Left panel*: results for the MLCS2k2 version of the SDSS compilation. *Right panel*: results for the SALT2 version of the SDSS compilation.

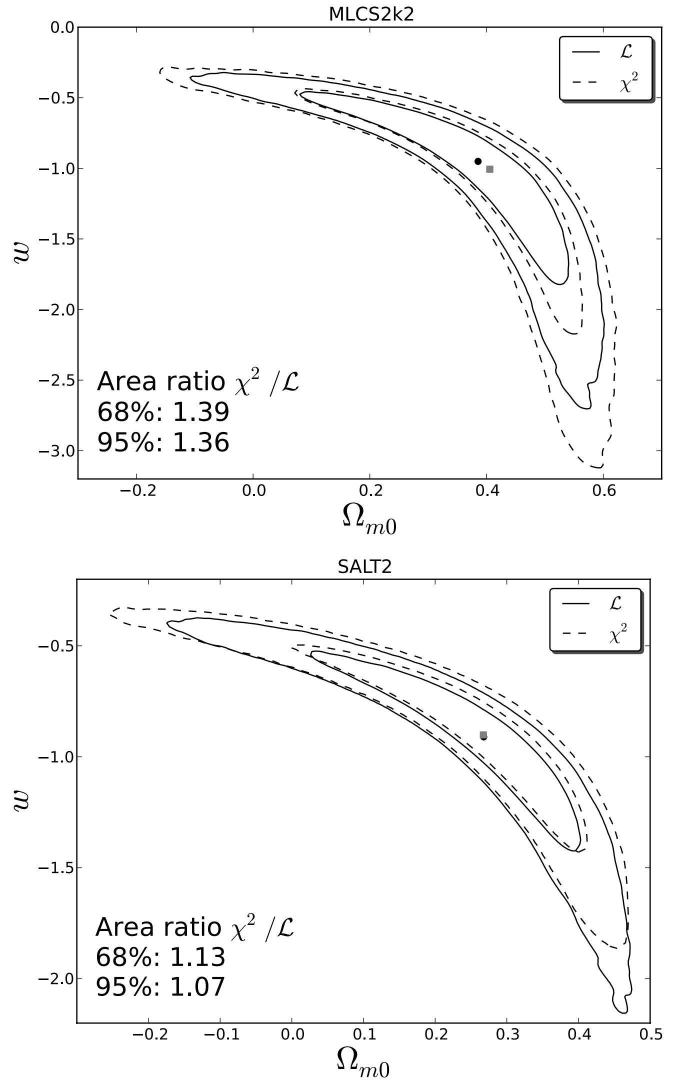

Fig. 2. Contours of 68% and 95% confidence level in the plane $\Omega_{m0} \times w$ for the FwCDM model. We marginalized over all other parameters with flat prior. The solid (dashed) lines are the results for the likelihood (χ^2) approach. The black circle (gray square) is the best fit value for the likelihood (χ^2). *Left panel*: results for the MLCS2k2 version of the SDSS compilation. *Right panel*: results for the SALT2 version of the SDSS compilation.

5. Results

In this section we compare the results obtained from the χ^2 and the likelihood approaches, as described in Sects. 3 and 4, using real data from the first year SDSS-II Supernova Survey together with other supernova data sets (ESSENCE, SNLS, HST, and a compilation of nearby supernovae) as described in Kessler et al. (2009).

To perform the comparison, we considered the following cosmological models:

– ΛCDM, in which we can write the Friedmann equation in terms of the parameters $\boldsymbol{\theta} = (\Omega_{m0}, \Omega_{k0})$ as

$$E^2(z;\boldsymbol{\theta}) = \Omega_{m0}(1+z)^3 + \Omega_{k0}(1+z)^2 + (1-\Omega_{m0}-\Omega_{k0}). \quad (23)$$

– FwCDM, described by $\boldsymbol{\theta} = (\Omega_{m0}, w)$ and

$$E^2(z;\boldsymbol{\theta}) = \Omega_{m0}(1+z)^3 + (1-\Omega_{m0}-\Omega_{k0})(1+z)^{3(1+w)}. \quad (24)$$

We chose the ΛCDM and FwCDM models to directly compare the best-fit and the 68% and 95% confidence contours for the parameters $\boldsymbol{\delta}$, $\sigma_{\rm int}$ and $\boldsymbol{\theta}$, for both SALT2 and MLCS2k2 data. The best-fit values were obtained with the MIGRAD minimization of the Minuit (James & Roos 1975) implementation in ROOT (Antcheva et al. 2009) and the probability distributions were obtained with Monte Carlo Markov Chains (MCMC). We considered as the probability distributions, in the context of χ^2 approach, the following functions:

$$\mathcal{P}_{\rm SALT2}(\boldsymbol{\theta},\boldsymbol{\delta},\mathcal{M}) = N_{\rm SALT2}\exp\left[-\chi^2_{\rm SALT2}(\boldsymbol{\theta},\boldsymbol{\delta},\mathcal{M},\sigma_{\rm int})/2\right], \quad (25)$$

$$\mathcal{P}_{\rm MLCS2k2}(\boldsymbol{\theta},h) = N_{\rm MLCS2k2}\exp\left[-\chi^2_{\rm MLCS2k2}(\boldsymbol{\theta},h,\sigma_{\rm int})/2\right], \quad (26)$$

where the normalization factors $N_{\rm SALT2}$ and $N_{\rm MLCS2k2}$ are independent of the parameters to be estimated. Note that for the traditional χ^2 method $\sigma_{\rm int}$ is fixed so the probability distribution does not depend on it.

In Figs. 1 and 2 we show the confidence contours for the parameters $\boldsymbol{\theta}$ for ΛCDM and FwCDM models, respectively. For

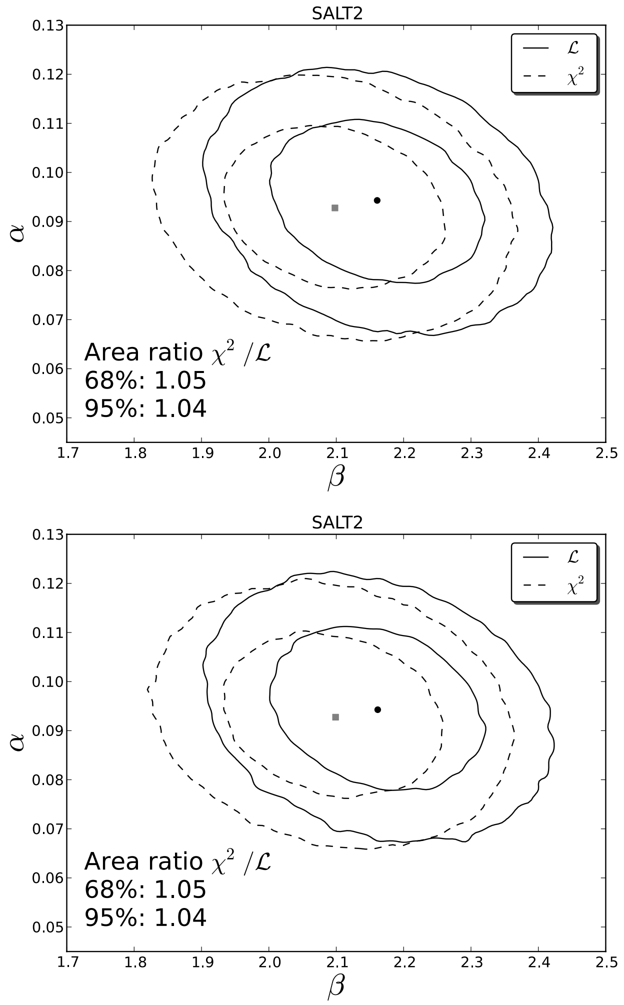

Fig. 3. Contours of 68% and 95% confidence level in the plane $\beta \times \alpha$. We marginalized over all other parameters with flat prior. The solid (dashed) lines are the results for the likelihood (χ^2) approach. The black circle (gray square) is the best fit value for the likelihood (χ^2). *Left panel*: results for the ΛCDM model with the SDSS compilation. *Right panel*: results for the FwCDM model with the SDSS compilation.

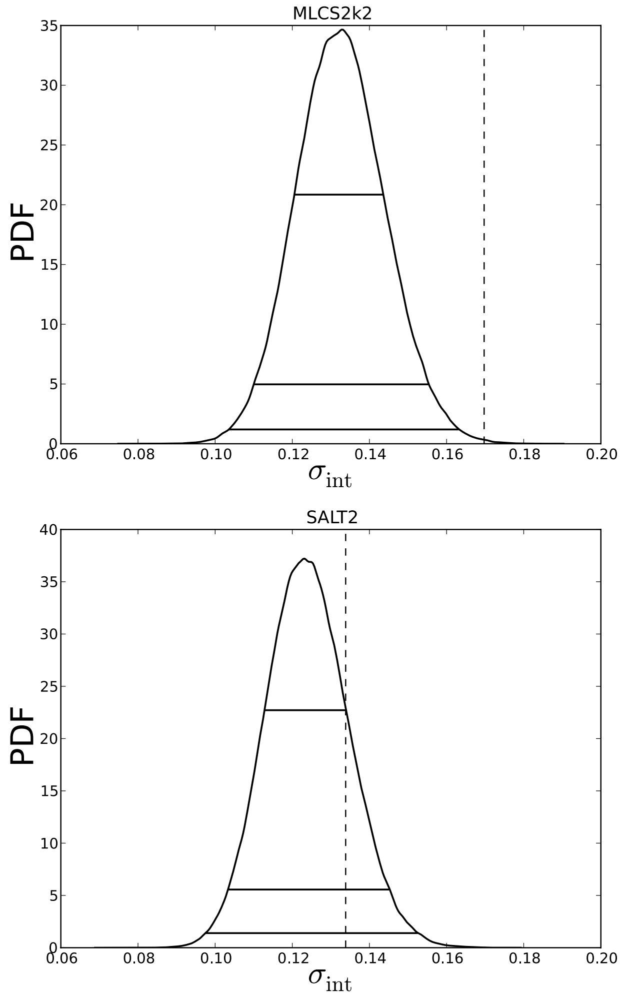

Fig. 4. Distributions for $\sigma_{\rm int}$ for the ΛCDM model from the likelihood approach. We marginalized over all other parameters with flat prior. The horizontal lines depict the 68%, 95% and 99% confidence intervals. The dashed line is the result for the χ^2 approach. *Left panel*: results for the MLCS2k2 version of the SDSS compilation. *Right panel*: results for the SALT2 version of the SDSS compilation.

these models, we note that the differences between the best fit and the area of the contours for the χ^2 and likelihood approaches are more significant when the MLCS2k2 fitter is used. Indeed, for the SALT2 fitter, the differences are not significant (less than 13%) – see also Fig. 4 and discussion below. Whether this is a general feature or depends on the models or dataset used has to be further investigated.

In Fig. 3 we show the confidence contours for the SALT2 parameters $\boldsymbol{\delta}$ for both ΛCDM and FwCDM models. We can see that there is no significant difference in the constraints for α. For β we find a bias that is, however, small compared to the 68% confidence interval for this parameter.

In Fig. 4 we show the probability distributions for $\sigma_{\rm int}$, given by the likelihood approach, for MLCS2k2 and SALT2 data. The traditional χ^2 approach gives only one value for this parameter without uncertainty and we represent it by the dashed vertical line in the figure. We can see that the discrepancy between the value obtained from the χ^2 approach and the best-fit value obtained from the likelihood approach is greater for the MLCS2k2 data. The results are incompatible at more than 99% confidence interval, which does not happen for the SALT2 data. This can possibly be due to the fact that $\sigma_{\rm int}$ is obtained using only a nearby sample in the χ^2 approach for MLCS2k2, while the whole sample is used in the likelihood approach. This issue deserves additional investigation and will be the subject of future work.

We also allow for a possible variation of the parameters α, β, $\mathcal{M}$ and $\sigma_{\rm int}$ with redshift, in the context of the SALT2 data. To perform such analysis the dataset was divided into redshift bins and the cosmological parameters were fixed at the best-fit values obtained from the global fit, then releasing α, β, $\mathcal{M}$ and $\sigma_{\rm int}$ to be determined in each bin. The results are shown in Fig. 5. We found evidence of evolution for the parameter β, in agreement with Kessler et al. (2009); furthermore we also found evidence of evolution for $\sigma_{\rm int}$, which might support the use of a variable $\sigma_{\rm int}$ instead of a constant one.

In addition to the SN Ia sample combination "e" from Kessler et al. (2009), we also used the Union2 compilation (Amanullah et al. 2010) and the SNLS third-year data (Guy et al. 2010). However, these references do not provide the covariances between the SALT2 parameters (m_B^*,x_1,c). To make a fair comparison with the SDSS data, we repeated the analyses for all three samples, this time neglecting the cross terms in Eq. (14). Without taking into account the covariances, the differences in

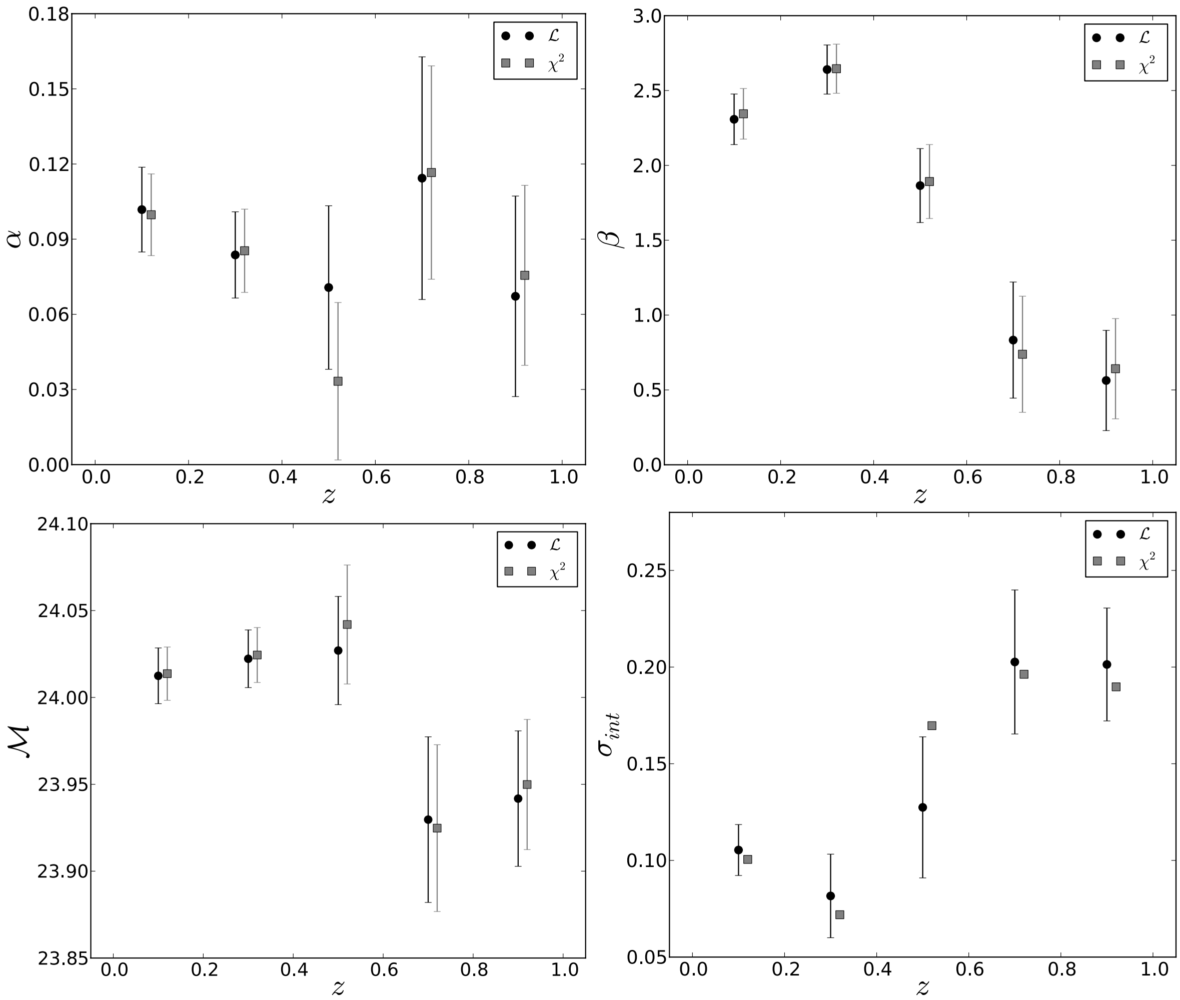

Fig. 5. Evolution of SALT2 parameters α, β, $\mathcal{M}$, and $\sigma_{\rm int}$ with redshift for the SDSS compilation for model ΛCDM. The cosmological parameters were kept fixed at their global best-fit values. The black circles and the gray squares are the results for the likelihood and the χ^2 approaches, respectively. The error bars represent the 68% confidence intervals, marginalizing over all other parameters with flat priors.

the areas for the SDSS compilation are essentially the same. The main difference is that, for the FwCDM model, the constraints are tighter for the χ^2 approach than for the likelihood one. For the Union 2 compilation, the difference in the areas ranged from 13% to 14%, while the SNLS data showed the greater discrepancies, reaching up to ~53%.

6. Conclusion

In this work we considered data from the first-year SDSS-II Supernova Survey together with other supernova data sets (ESSENCE, SNLS, HST, and a compilation of nearby supernovae) and compared two different approaches (traditional χ^2 and likelihood) to determine constraints from SNe Ia processed by two of the most used light-curve fitters in the literature, the MLCS2k2 and the SALT2. The MLCS2k2 gives an estimation of the distance modulus for each SN Ia, with its corresponding variance, which can be directly compared with the model prediction for this quantity. With SALT2 we can only have an estimate of the distance modulus depending on parameters to be obtained simultaneously with the cosmological ones. Furthermore, in the SNe Ia analysis, it is common to introduce a residual, unknown, contribution $\sigma_{\rm int}$ to the variance which, in the traditional approach, is determined by imposing that the reduced χ^2 be unity, when considering the full sample, in the SALT2 case, or only nearby SNe Ia, in the MLCS2k2 case.

By comparing the results obtained from the traditional χ^2 approach with those of the likelihood method, we showed that for current data and chosen cosmological models, there is a small difference in the best-fit values and confidence contours (~30% difference in area) (cf. Figs. 1 and 2) if the MLCS2k2 fitter is adopted. For SALT2 the difference is less significant (≲13% difference in areas). In both cases the likelihood approach gives more restrictive constraints. We can understand these results by observing, from Fig. 4, that the estimated value of $\sigma_{\rm int}$ in the traditional approach, is higher than the peak of the $\sigma_{\rm int}$ distribution obtained with the likelihood method. Indeed, for MLCS2k2 the $\sigma_{\rm int}$ value obtained with the traditional χ^2 approach is outside the 99% confidence level of the distribution obtained with the likelihood method. For SALT2 it is on the order of 68% confidence level and this might explain why for SALT2 the differences between these two approaches are less significant. We also remark that the covariance between the SALT2 parameters (m_B^*, x_1, c) has an important role in the above result since we obtained, for the 68% confidence level contour, an area ratio of 1.13, when considering the covariance, and of 0.95, when neglecting the covariance.

We also studied the possible evolution of the SALT2 parameters α, β and $\mathcal{M}$ with the redshift, splitting the samples in redshift bins and performing the fit separately for each one. In this case, we found evidence of evolution in the parameter β and also in $\sigma_{\rm int}$.

While this paper was in preparation two articles appeared that addressed the issue of using χ^2 with unknown variances. The first (Kim 2011) focused on the question of determining $\sigma_{\rm int}$ by applying the likelihood approach to simulated data for the distance modulus, in which case there is no need to specify the light-curve fitter. In contrast, we used real data here and compared the constraints arising from the two most common light-curve fitters. The second paper (March et al. 2011) employed SALT2 and proposed a more sophisticated, Bayesian analysis instead of the likelihood method proposed here and also by Kim (2011). Furthermore, we did not find in our work, with the SDSS first year compilation, any kind of catastrophic biases for parameter estimation when adopting the likelihood approach, a possibility suggested in March et al. (2011).

In summary, we used current data to compare the traditional χ^2 and the likelihood approaches to determine best fit and confidence regions from SNe Ia. We argued that when the variance is not completely known, minimizing the traditional χ^2 is not formally equivalent to maximizing the likelihood function, since the normalization of the likelihood, assumed to be Gaussian, is also a function of parameters to be determined. We conclude by suggesting to adopt the likelihood framework instead of the traditional χ^2 one, since it is more straightforward, numerically more efficient, and self-consistent.

Acknowledgements. We thank Masao Sako for helpful discussions and suggestions. B.L.L. and I.W. also thank CNPq, Brazil, and S. E. J. also thanks ICTP, Italy, for support.

References

Amanullah, R., Lidman, C., Rubin, D., et al. 2010, ApJ, 716, 712
Antcheva, I., Ballintijn, M., Bellenot, B., et al. 2009, Comput. Phys. Commun., 180, 2499
Astier, P., Guy, J., Regnault, N., et al. 2006, A&A, 447, 31
Burns, C. R., Stritzinger, M., Phillips, M. M., et al. 2011, AJ, 141, 19
Conley, A., Sullivan, M., Hsiao, E. Y., et al. 2008, ApJ, 681, 482
Conley, A., Guy, J., Sullivan, M., et al. 2011, ApJS, 192, 1
DES: http://www.darkenergysurvey.org/
Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. 2005, ApJ, 633, 560
Goobar, A., & Leibundgut, B. 2011, Annu. Rev. Nucl. Part. Sci., 61, 251
Guy, J., Astier, P., Nobili, S., Regnault, N., & Pain, R. 2005, A&A, 443, 781
Guy, J., Astier, P., Baumont, S., et al. 2007, A&A, 466, 11
Guy, J., Sullivan, M., Conley, A., et al. 2010, A&A, 523, A7
Hicken M., Wood-Vasey, W. M., Blondin, S., et al. 2009, ApJ, 700, 1097
Howell, D. A. 2011, Nat. Commun. 2, 350
James, F., & Roos, M. 1975, Comput. Phys. Commun., 10, 343
Jha, S. Riess, A. G., & Kirshner, R. P. 2007, ApJ, 659, 122
Kessler, R., Becker, A. C., Cinabro, D. et al. 2009, ApJS, 185, 32
Kim, A. G. 2011, PASP, 123, 230
Kirshner, R. P. 2010, in Dark energy, Observational and theoretical approaches, ed. P. Ruiz-Lapuente [arXiv:0910.0257]
Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, ApJS, 192, 18
Kowalski, M., Rubin, D., Aldering, G., et al. 2008, ApJ, 686, 749
LSST, http://www.lsst.org/
Mantz, A., Allen, S. W., Rapetti, D., & Ebeling, H. 2010, MNRAS, 406, 1759
March, M. C., Trotta, R., Berkes, P., Starkman, G. D., & Vaudrevange, P. M. 2011, MNRAS, 418, 2308
Pan-STARRS, http://pan-starrs.ifa.hawaii.edu/
Percival, W. J., Reid, B. A., Eisenstein, D. J., et al. 2010, MNRAS, 401, 2148
Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, ApJ, 517, 565
Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, ApJ, 116, 1009
Vikhlinin, A., Kravtsov, A. V., Burenin, R. A., et al. 2009, ApJ, 692, 1060
Vishwakarma, R. G., & Narlikar, J. V. 2010, Res. Astron. Astrophys., 10, 1195
Wang, L., Goldhaber, G., Aldering, G., & Perlmutter, S. 2003, ApJ, 590, 944
Wang, X., Wang, L., Zhou, X., Lou, Y.-Q., & Li, Z. 2005, ApJ, 620, L87
Wood-Vasey, W. M., Miknaitis, G., Stubbs, C. W., et al. 2007, ApJ, 666, 694