Marginal evidence of cosmic acceleration from Supernovæ of type Ia

Lucas Barbier

Department of Physics, Linnéuniversitetet, Sweden
Universita degli Studi dell'Insubria, Italy

September 2022

1 Introduction

Until the early 1990, a cosmological model of Friedmann type with either radiations or dust as main energetic components was thought to be the standard model in cosmology. It is however common in cosmology that when new data are available, we shall adjust some parameters or modify the model itself. The era of a model with a Λ-constant, or dark energy, finds its origin in 1994 with the measurement of the Hubble constant H_0 by Freedman et al.(1994) [1] and Pierce et al.(1994) [5]. This, with the flateness imposed by the theory of inflation resulted in an age problem for the universe, and required a Λ-term. However the idea of a non-zero Λ-term has always been around and its apparition did not redefine the paradigm of the model. Only later when the analysis of Supernovae of Type Ia (SNIa) from Perlmutter [4] et Riess [6] came, it has been realized that a large enough Λ-term was needed to cause an acceleration of the expansion of the universe.

Nowadays, the ΛCDM has become the standard cosmological model. Λ for the cosmological constant, and CDM for *Cold Dark Matter*. Although able to reproduce with high precision numerous observations, this concordance model implies some *dark components* representing up to 95% of the energy content of the universe and, from which we ignore properties.

Recently there have been some important discussions on the capability of SNIa data alone to prove the acceleration of the expansion. In this work we try to determine if the available data permit to conclude with certainty to a cosmic acceleration. In Sec. 2 we display the properties of SNIa and introduce some notions of cosmology. In Sec. 3 following Nielsen et al. [3] we present the model used for the analysis. Results are discussed in Sec. 4 and compared to older analyses in Sec. 5. We conclude in Sec. 6.

2 Supernovae of type Ia and the SALT2 model

Type Ia supernovae[1] (SNIa) are thought to be the result of the explosion of a carbon-oxygen white dwarf in a binary system as it goes over the Chandrasehkar limit. The phenomenon is either due to accretion from a donor or mergers. They are the brightest of all supernovae with an absolute magnitude of $M_B \approx -19.5$ at maximum light, occur in all galaxy types, and are characterised by a silicon absorption feature (rest wavelength = 6355 angstroms) in their maximum light spectra.

[1] This part is extracted from https://astronomy.swin.edu.au/cosmos/t/Type+Ia+Supernova.

They can eject material at speeds of the order of 10,000 km/s and outshine an entire galaxy at their peak brightness.

Originally thought to be standard candles where every SNIa had the same peak brightness, it has been shown that this is close to the truth, but not quite. SNIa exhibit brightnesses at maximum that range from about $+1.5$ to -1.5 magnitudes around a typical SNIa.

2.1 How to standardise SNIa?

We use in this work the Joint Lightcurve Analysis (JLA) catalogue, which combines data from the Sloan Digital Sky Survey (SDSS), the SuperNova Legacy Survey (SNLS), some supernovae from the Hubble Space Telescope (HST), and a few other supernovae selected among other surveys. The data are treated with the SALT2 [2] (Spectral Adaptative Lightcurve Template 2) approach where SNIa are standardised by adjusting their lightcurves to an empirical templates. The parameters found with this fit are then used in the cosmological analysis.

In this analysis we assign three parameters to each SNIa: m_B^* the apparent magnitude at maximum (in the rest frame "B-band"), x_1 the light curve shape and c the colour correction.

The distance modulus is then taken to be :

$$\mu_{SN} = m_\beta^* - M + \alpha x_1 - \beta_c$$

M is the absolute magnitude, α and β are unknown constants for all SNIas. These global constants are fitted along the cosmological parameters. The SNIa distance modulus is then compared to the expectation in the standard ΛCDM cosmological model :

$$\mu = 25 + 5\log_{10}(d_L/Mpc)$$

with

$$d_L = (1+z)\frac{d_H}{\sqrt{\Omega_k}} \cdot \sinh\left(\sqrt{\Omega_k}\int_0^z \frac{H_0 dz'}{H(z)}\right)$$

Here $d_H = c/H_0$, $\mathrm{H_0}$= 100h $\cdot$ km $\cdot$ s^{-1}/Mpc and $H(z) = \sqrt{\Omega_m(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda}$.

Due to degeneracy between M_0 and H_0, we fix $h = 0.7$, which is consistent with results from other methods.

3 Statistical Analysis

In order to perform this type of analysis we first need to specify our model for the data. For a given SNIa, the true data (m_β^*, x_1, c) are drawn from some global distribution, and the values are contaminated by various sources of noise, yielding the observed values $(\hat{m_\beta^*}, \hat{x_1}, \hat{c})$. When experimental uncertainties are of the same order of the intrinsic variance, the observed value is *not* a good estimate of the true value. For this analysis we choose to use the maximum likelihood estimator; for a given dataset and probability model, makimum likelihood finds values of the model parameters that give the observed data the highest probability. In more technical terms, the likelihood is defined as the probability density function (pdf) of the data $\hat{X}$ given a model θ, and can be written as

$$\mathscr{L} = f\left(\hat{X}|\theta\right)$$

3.1 Maximum likelihood estimator

Parametrizing our cosmological model by θ, the likelihood function can be written as:

$$\mathscr{L} = p\left[(m^*_\beta, x_1, c)|\theta\right] = \int p\left[(m^*_\beta, x_1, c)|(M, x_1, c), \theta\right] p\left[(M, x_1, c)|\theta\right] dM dx_1 dc \qquad (1)$$

In this expression the first part of the integral represents the experimental uncertainties while the second one is the variance of intrinsic distributions.

One can observe on Fig. 1 the corrective parameters x_1 and c of the SALT2 fitting. There is no theoretical model to describe the distribution of these lightcurve parameters. For simplicity we adopt global and independent Gaussian distributions for every parameter (M, x_1, c), which according to Fig. 1 seem to be a decent approximation. Their probability density is therefore:

$$p\left[(M, x_1, c)|\theta\right] = p(M|\theta)p(x_1|\theta)p(c|\theta)$$

with $p(i|\theta) = \left(2\pi\sigma_{i_0}^2\right)^{-1/2} \exp\left(-\left[\frac{i - i_0}{\sqrt{2}\sigma_{i_0}}\right]^2\right)$, $i \in \{M, x_1, c\}$. The six free parameters $\{M_0, \sigma_{M_0}, x_{1,0}, \sigma_{x_{1,0}}, c_0, \sigma_{c_0}\}$ are fitted along the cosmological parameters and are included in θ.

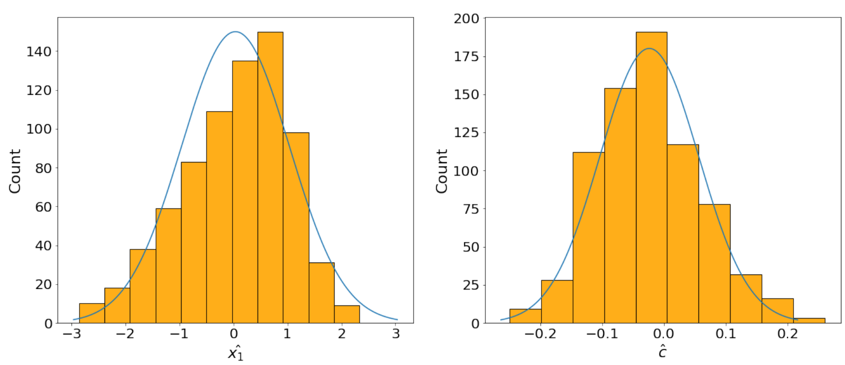

Figure 1: Distribution of the parameters x_1 and c, with a superimposed Gaussian.

By introducing the vector $Y = \{M_1, x_{11}, c_1, \ldots, M_N, x_{1N}, c_N\}$, the zero point Y_0 and, the matrix $\Sigma_l = diag(\sigma_{M_0}, \sigma_{x_{1,0}}, \sigma_{c_0})$, the probability density of the true parameters writes:

$$p(Y|\theta) = |2\pi\Sigma_l|^{-1/2} \exp\left[\frac{-\left(Y - Y_0\right)\Sigma_l^{-1}\left(Y - Y_0\right)^T}{2}\right]$$

We now have an expression for the intrinsic variance of the distributions, i.e. the right term of the integral in (1). The next step is now to specify a model for the uncertainty on data. We introduce another set of vectors: the true values $X = \{m^*_\beta, x_{11}, c_1, \ldots\}$, the observed values $\hat{X}$, and the estimated covariance matrix Σ_d which includes both systematic and statistical errors. Given these vectors, the probability density of the data given some set of true parameters is :

$$p(\hat{X}|X, \theta) = |2\pi\Sigma_d|^{-1/2} \exp\left[\frac{-\left(\hat{X} - X\right)\Sigma_d^{-1}\left(\hat{X} - X\right)^T}{2}\right]$$

We have described both terms of expression (1) and can now compute it. In order to facilitate the computation we first introduce the vector $\hat{Z} = \{m^*_\beta, -\mu_1, \hat{x}_{11}, \hat{c}_1, \dots\}$ and the block diagonal matrix

$$A = \begin{pmatrix} 1 & 0 & 0 & \dots \\ -\alpha & 1 & 0 & 0 \\ \beta & 0 & 1 & 0 \\ \dots & 0 & \dots & \dots \end{pmatrix}$$

With these, we have $\hat{X} - X = (\hat{Z}A^{-1} - Y)A$ and therefore $p(\hat{X}|X,\theta) = p(\hat{Z}|Y,\theta)$. The likelihood can be integrated analytically:

$$\mathscr{L} = |2\pi(\Sigma_d + A^T\Sigma_l A)|^{-1/2} \cdot \exp\left[\frac{-(\hat{Z} - Y_0 A)(\Sigma_d + A^T\Sigma_l A)^{-1}(\hat{Z} - Y_0 A)^T}{2}\right]$$

This is the likelihood introduced at the beginning and that takes the simple Gaussian model into account. This likelihood is the quantity we maximize to deduce confidence intervals. The 10 parameters we fit are:

$$\{\Omega_m, \Omega_\Lambda, \alpha, x_{1,0}, \sigma_{x_{1,0}}, \beta, c_0, \sigma_{c_0}, M_0, \sigma_{M_0}\}$$

4 Results of the main fit

We present in this section the main results[2] as well as the confidence zones of the fit.

[2] Details of the analysis can be found at https://github.com/lukbrb/academic-physics/blob/master/SupernovaeAnalysis/Analysis.ipynb

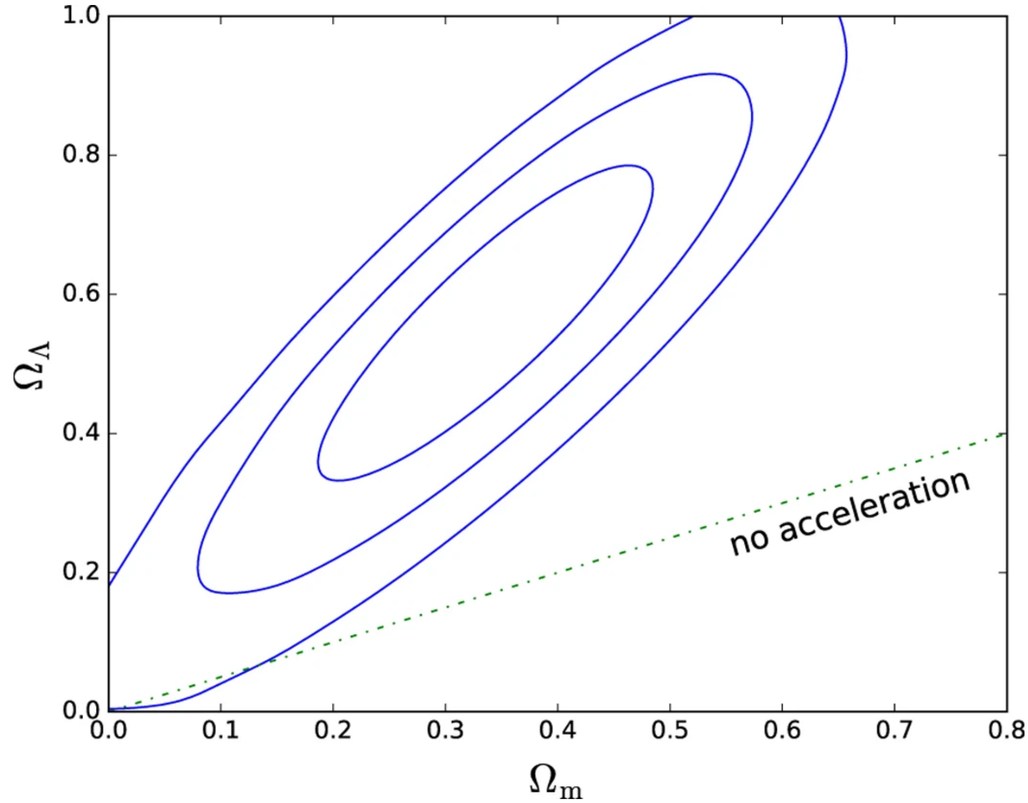

Figure 2: Confidence contours at 1, 2 and 3σ of the Ω_m, Ω_Λ profile likelihood. Figure from [3].

The parameters of the fits performed for different cosmological models can be found in Table 4.

Constraint	$\Delta\chi^2$	Ω_m	Ω_Λ	α	$x_{1,0}$	σ_{x1}	β	c_0	σ_{c0}	M_0	σ_{M_0}
None (best fit)	0.000	0.341	0.569	0.134	0.038	0.932	3.059	-0.016	0.071	-19.052	0.108
Flat geometry	0.147	0.376	0.624	0.135	0.039	0.932	3.060	-0.016	0.071	-19.055	0.108
Empty universe	11.900	0.000	0.000	0.133	0.034	0.932	3.051	-0.015	0.071	-19.014	0.109
Non-accelerating	11.000	0.068	0.034	0.132	0.033	0.931	3.045	-0.013	0.071	-19.006	0.109
Matter-less universe	10.400	0.000	0.094	0.134	0.036	0.932	3.059	-0.017	0.071	-19.032	0.109
Einstein-deSitter	221.970	1.000	0.000	0.123	0.014	0.927	3.039	0.009	0.072	-18.839	0.125

Are presented on Fig. 2 the contours at 1, 2 and 3σ of the Ω_m, Ω_Λ profile likelihood projected on the $(\Omega_m, \Omega_\Lambda)$ plane.

We can already observe from Fig. 2 and the data from Table 4 that the Milne universe is rejected at only 3σ, a marginal evidence. In particle physics a discovery is claimed at 5σ.

4.1 Testing the results

We now wish to verify the goodness of the obtained results. A simple Gaussian model has been supposed, and we now determine at what extent it corresponds to the data. We also verify the validity of our statistical approximations. Although the model is *not* a linear one, we use Wilk's theorem to determine the confidences zones. Figure 3 presents the *pulls*, defined as the normalised residuals to the combine expected error. pull= $(\hat{Z} - Y_0 A)U^{-1}$.

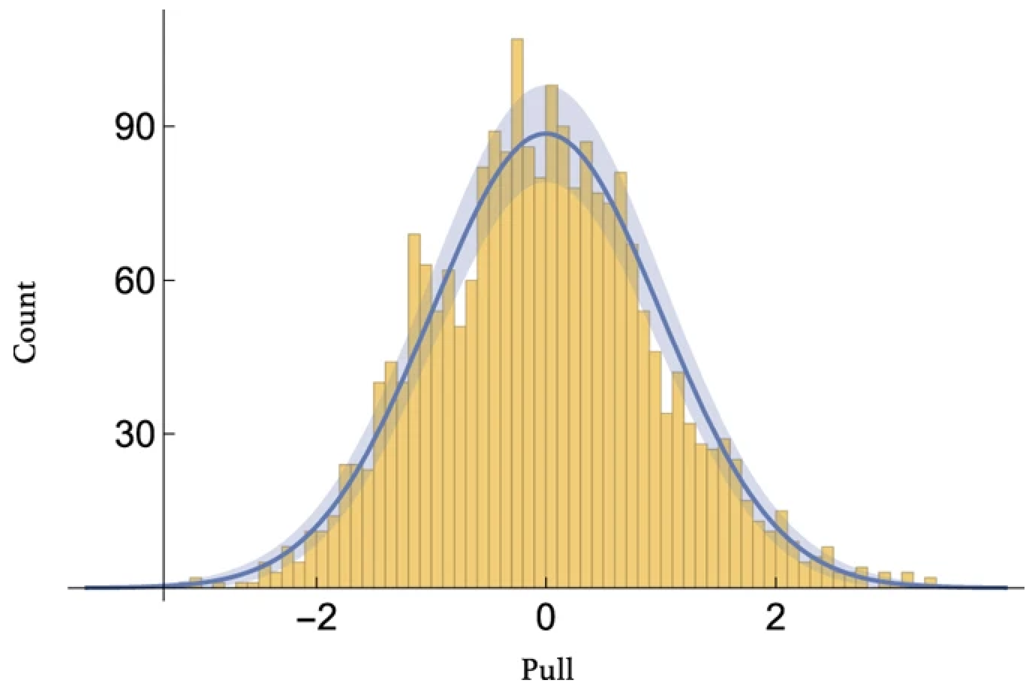

Figure 3: Distribution of pulls for the best-fit model compared to a normal distribution. Figure from [3].

It is clear from Fig. 3 that although the description is not perfect, the result is not invalid. A K-S *goodness-of-fit* test of the pull distribution yields a p-value of 0.1389

In order to verify the confidence zones set by Wilk's theorem, we perform a Monte-Carlo test. To do so, we suppose our model correct- we take the parameters of our "best-fit" for the model- and we simulate 10 000 datasets. For each of those, we keep the values of z but for M, x_1 and c, new values are sampled following the predicted distributions. For each dataset we compute the maximum likelihood, and keep the best parameters of this fit. According to Wilk's theorem, the quantity $-2\log \mathscr{L}_{true}/\mathscr{L}_{max}$ must follow a χ^2_{10} distribution. This is confirmed to high precision by the Fig. 4, which shows the χ^2_{10} distribution superimposed on the $-2\log \mathscr{L}_{true}/\mathscr{L}_{max}$ distribution. We can therefore trust the confidence intervals determined by the likelihood ration method.

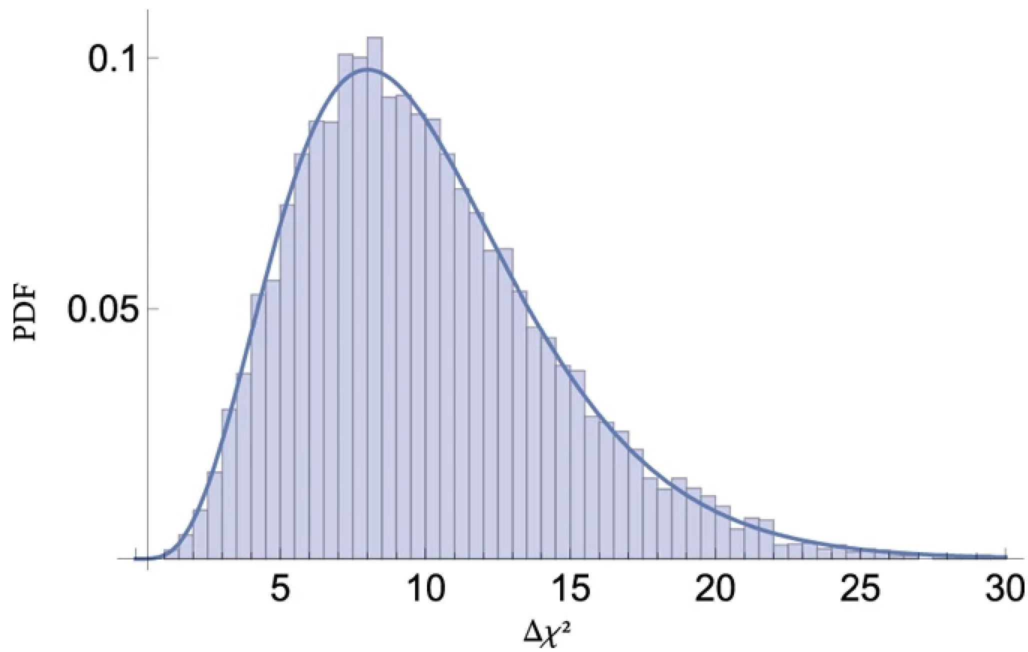

Figure 4: The distribution of the likelihood ratio from Monte Carlo, with a χ^2 distribution with 10 degrees of freedom superimposed. Figure from [3].

5 Comparison with old analyses

In order to highlight the strength of the analysis hereby performed, it is instructive to look at its differences with older methods. Precedent analyses can be put in two categories: the likelihood based one, which we argue is not a good fit, and a residual based one, which in particular is not a likelihood maximization.

We first detail these two methods, and we come back later on the comparison with the approach proposed in this work.

5.1 Residual based method

This method is the most common one for analyses using the SALT method, or any similar lightcurve fitting method. The exact procedure varies somehow but the principal idea is that we consider a quantity f (often called χ^2) such that

$$f \approx \sum \frac{\Delta\mu^2}{\sigma_\mu^2 + \sigma_{int}^2}$$

The most recent versions of this type of analysis tries to determine σ_{int} independently, but most of older analyses proceeded in a peculiar way. First of all f was minimized using a plausible value of σ_{int}, often relying on previous analyses. Then, once the minimum is found, one tunes σ_{int} such that f equals the degree of freedom. However as noted by [3] the likelihood estimate does not exactly satisfy this condition. Although this method *is not* derived from a likelihood, confidence zones are derived using Wilk's theorem, which is a nonsense in this context.

However since the estimation of f is an educated one, the limits determined in this frame are not totally incorrect.

5.2 Simple likelihood

This part is extracted from Nielsen et al. [3]

One other method maximises a likelihood, which is written in the case of uncorrelated magnitudes as:

$$\mathcal{L} \propto \sqrt{2\pi(\sigma_\mu^2 + \sigma_{int}^2)} \cdot \exp\left(-\frac{1}{2}\frac{\Delta\mu_{SN}}{\sigma_\mu^2 + \sigma_{int}^2}\right)$$

so it integrates over μ_{SN} to unity and can be used for model comparison. From Equation (1) we see that this corresponds to assuming flat distributions for x_1 and c. However the actual distributions of $\hat{x_1}$ and $\hat{c}$ are close to Gaussian, as seen in Fig. 1.

5.3 Comparison with previous analysis

The important parameter that distinguishes the two analyses are the terms linking the residuals of $\hat{x}_1$, $\hat{c}$ to the ones of $\hat{m}_B^*$, the anti-diagonal of $A^{T-1}\Sigma_d A^{-1} + \Sigma_l$. Since the other analyses do not consider the residuals of these terms, they obviously cannot include these corrections.

However, it does not mean that the results obtained from the residuals method are wrong. It simply has to be determined by another way than with the MLE. In particular, since we cannot conduct analytical calculation, we have to make use of numerical simulations. Yet to perform a simulation, one needs to know the distribution from which we sample x_1 and c. In the idea of pursuing the comparison between the two methods we adopt the distributions and results from the previous Monte-Carlo simulation. For the term σ_{int} we utilize the predetermined value from the MLE.

On Fig. 5 we can observe the prediction differences for each dataset. This Monte-Carlo study is then compared to the values of the real dataset. By fitting the real dataset with the residuals method, the best fit gives:

$$\{\Omega_m, \Omega_\Lambda, \alpha, \beta, M_0\} = \{0.200, 0.591, 0.134, 3.08, -19.07\}$$

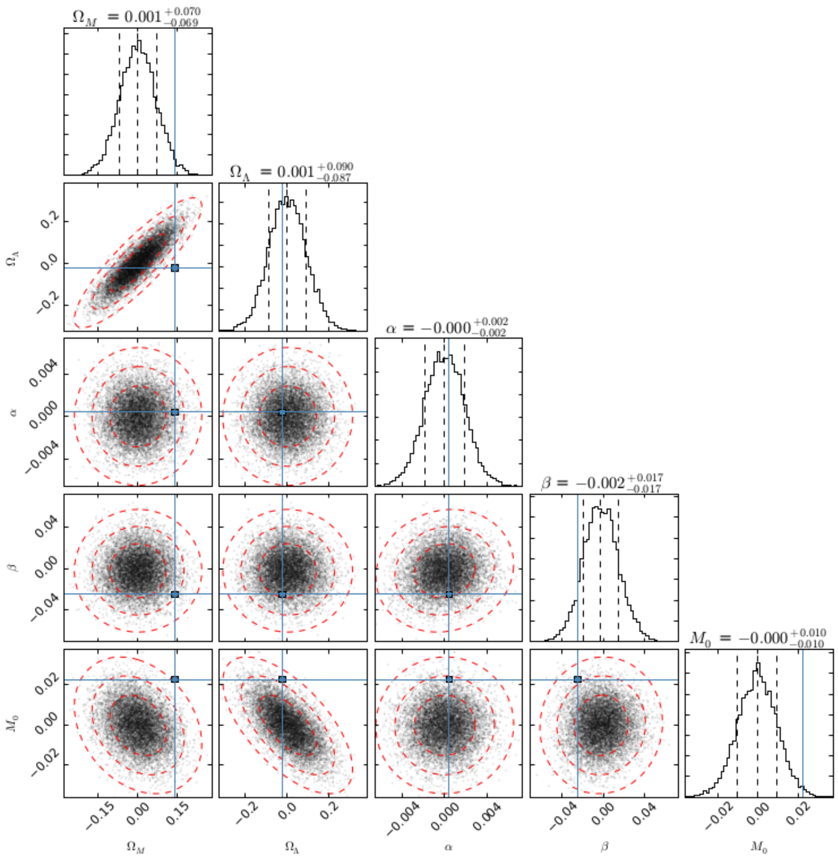

Figure 5: Correlations between MLE and residual method parameters, for the relevant parameters. Plotted is MLE residual estimate. The dashed ellipses show approximate 1, 2, 3σ 2d profile contours. The blue markers show the value obtained from the real data. We see that in particular the Ω_m , Ω_Λ point is off. In total, there are 9945 simulated datasets plotted. Figure from [3].

Both methods agree on average. This simply means that for *this* model both methods more or less agree - see the spread on Fig. 5. Looking at the real values we observe that they don't agree so well. In order to quantify this slight disagreement,we first calculate the covariance matrix of Monte-Carlo values, $\mathscr{F}$. If we consider the points distribution of the Monte-Carlo method as an estimate of the probability density function (pdf), $\mathscr{F}$ is the covariance of a 5-dimensional Gaussian distribution. One can then compute the χ^2 of the difference observed in real data:

$$\Delta\chi^2 = \Delta\theta_0 \cdot \mathscr{F}^{-1} \Delta\theta_0 \approx 22.73$$

where θ_0 is the set of parameters. For a χ^2 distribution with 5 degrees of freedom, it yields approximately 3.6σ. In other words, considering the two fits, it is pretty unlikely that they differ this much, if the Gaussian model is correct. What is important to note here is that the results obtained from the residuals method can only be validated by a Monte-Carlo study. This does not prevent a good estimation, but we do not have any information on the parameters underlying the distribution. We loose the capacity of evaluating the correction parameters, which nevertheless has to be supposed in order to validate the method. By performing the MLE, only standard assumptions have been made.

6 Conclusion

This part is strongly inspired by [3]. Using the JLA catalogue processed by the SALT2 method, we have confirmed by a statistically principled analysis that the SN Ia Hubble diagram appears consistent with an uniform rate of expansion. We find a marginal evidence for the widely accepted claim that the expansion of the universe is presently accelerating.

Although the Gaussian model we used is not perfect, it appears to be an adequate first step towards understanding SN Ia standardisation. One might be concerned that different effects, such as Malmquist bias, might affect the data. Such effects might not be appropriate to the method used in this work, and are better treated in a Bayesian framework. This work solely intends to perform a analysis statistically sound to highlight the differences from previous analyses with the same data.

This part is extracted from[3].

Whether the expansion rate is accelerating or not is a kinematic test and it is only for ease of comparison with previous results that we have chosen to show the impact of doing the correct statistical analysis in the ΛCDM framework. In particular the "Milne model" refers here to an equation of state $p = \rho/3$ and should not be taken to mean an empty universe. For example the deceleration due to gravity may be countered by bulk viscosity associated with the formation of structure, resulting in expansion at approximately constant velocity even in an universe containing matter but no dark energy. Such a cosmology is not prima facie in conflict with observations of the angular scale of fluctuations in the cosmic microwave background or of baryonic acoustic oscillations, although this does require further investigation. In any case, both of these are geometric rather than dynamical measures and do not provide compelling direct evidence for a cosmological constant — rather its value is inferred from the assumed "cosmic sum rule": $\Omega_\Lambda = 1 - \Omega_m + \Omega_k$. This would be altered if e.g. an additional term due to the 'back reaction' of inhomogeneities is included in the Friedmann equations.

References

[1] Wendy L Freedman, Barry F Madore, Jeremy R Mould, Robert Hill, Laura Ferrarese, Robert C Kennicutt, Abhijit Saha, Peter B Stetson, John A Graham, Holland Ford, et al. Distance to the virgo cluster galaxy m100 from hubble space telescope observations of cepheids. *Nature*, 371(6500):757–762, 1994.

[2] J. Guy, P. Astier, S. Baumont, D. Hardin, R. Pain, N. Regnault, S. Basa, R. G. Carlberg, A. Conley, S. Fabbro, D. Fouchez, I. M. Hook, D. A. Howell, K. Perrett, C. J. Pritchet, J. Rich, M. Sullivan, P. Antilogus, E. Aubourg, G. Bazin, J. Bronder, M. Filiol, N. Palanque-Delabrouille, P. Ripoche, and V. Ruhlmann-Kleider. SALT2: using distant supernovae to improve the use of type ia supernovae as distance indicators. *Astronomy &amp Astrophysics*, 466(1):11–21, feb 2007.

[3] J. T. Nielsen, A. Guffanti, and S. Sarkar. Marginal evidence for cosmic acceleration from type ia supernovae. *Scientific Reports*, 6(1):35596, Oct 2016.

[4] S. Perlmutter, G. Aldering, G. Goldhaber, R. A. Knop, P. Nugent, P. G. Castro, S. Deustua, S. Fabbro, A. Goobar, D. E. Groom, I. M. Hook, A. G. Kim, M. Y. Kim, J. C. Lee, N. J. Nunes, R. Pain, C. R. Pennypacker, R. Quimby, C. Lidman, R. S. Ellis, M. Irwin, R. G. McMahon, P. Ruiz-Lapuente, N. Walton, B. Schaefer, B. J. Boyle, A. V. Filippenko, T. Matheson,

A. S. Fruchter, N. Panagia, H. J. M. Newberg, W. J. Couch, and The Supernova Cosmology Project. Measurements of {*upOmega* and {*upLambda* from 42 high-redshift supernovae. *The Astrophysical Journal*, 517(2):565–586, jun 1999.

[5] Michael J. Pierce, Douglas L. Welch, Robert D. McClure, Sidney van den Bergh, Renè Racine, and Peter B. Stetson. The hubble constant and virgo cluster distance from observations of cepheid variables. *Nature*, 371(6496):385–389, Sep 1994.

[6] Adam G. Riess, Alexei V. Filippenko, Peter Challis, Alejandro Clocchiatti, Alan Diercks, Peter M. Garnavich, Ron L. Gilliland, Craig J. Hogan, Saurabh Jha, Robert P. Kirshner, B. Leibundgut, M. M. Phillips, David Reiss, Brian P. Schmidt, Robert A. Schommer, R. Chris Smith, J. Spyromilio, Christopher Stubbs, Nicholas B. Suntzeff, and John Tonry. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *The Astronomical Journal*, 116(3):1009–1038, sep 1998.