Development of an ultra-fast, likelihood-based, distance inference framework for the next generation of Type Ia Supernova Surveys

Dylan KUHN, Marc BETOULE, on behalf of the LEMAÎTRE collaboration Sorbonne Université, CRNS, Université de Paris, LPNHE, 75252 Paris Cedex 05, France

In this work, we present EDRIS (French for Distance Estimator for Incomplete Supernova Surveys), a cosmological inference framework tailored to reconstruct unbiased cosmological distances from type Ia supernovae light-curve parameters. This goal is achieved by including data truncation directly in the statistical model which takes care of the standardization of luminosity distances. It allows us to build a single-step distance estimate by maximizing the corresponding likelihood, free from the biases the survey detection limits would introduce otherwise. Moreover, we expect the current worldwide statistics to be multiplied by O(10) in the upcoming years. This provides a new challenge to handle as the cosmological analysis must stay computationally towable. We show that the optimization methods used in EDRIS allow for a reasonable time complexity of $O(N^2)$ resulting in a very fast inference process (O(10s) for 1500 supernovae).

1 Introduction

The main goal of EDRIS is to handle the instrumental bias known as Malmquist bias. The observable magnitude limitation inherent to each survey induces a preferential observation of the intrinsically brighter objects and, therefore, a negative bias on the distance estimator⁷. The usual way to prevent this estimator from being biased is to run extensive simulations to compute a bias correction which is propagated to the reconstructed distances afterwards^{2,4}. However, this method does not scale well with the number of data and we expect the current worldwide statistics to increase tenfold in the upcoming years. To tackle both the bias issue and the scaling issue, we propose an innovative approach based on a truncated likelihood minimization. This allows to integrate a modeling of the selection effect in the distance standardization model. The presented method is implemented in the context of a large data combination named LEMAÎTRE. The LEMAÎTRE analysis is an end-to-end cosmology analysis using three unpublished SNe Ia samples (ZTF, SNLS 5y, HSC/Subaru), totalizing O(4000) supernovae. All data are processed with a common pipeline going from pixels to cosmological inference, this last role being filled by EDRIS.

2 Modeling of the Malmquist bias

The common model used to describe the behaviour of type Ia supernovae is the Tripp mode[§] (see equation 1). The usual standardization process involves two parameters (color and stretch) but this model can easily be generalized to consider an arbitrary number of standardization parameters. The * symbols indicate latent parameters we need to introduce to properly account for the fact that the c and x_1 measurements are affected with uncertainties that cannot be neglected.

$$Y_{i} = \begin{pmatrix} m_{i} \\ x_{1,i} \\ c_{i} \end{pmatrix} = \begin{pmatrix} M + \mu_{i}(z,\theta) + \alpha x_{1,i}^{*} + \beta c_{i}^{*} \\ x_{1,i}^{*} \\ c_{i}^{*} \end{pmatrix} + \begin{pmatrix} \epsilon_{i} \\ 0 \\ 0 \end{pmatrix} \text{ with } \epsilon_{i} \sim \mathcal{N}(0,\sigma_{int}^{2})$$
(1)

 $m_i, x_{1,i}, c_i$ and μ_i respectively refer to the apparent magnitude of the supernova *i*, its stretch parameter, its color parameter and its distance modulus. *M* is the absolute magnitude of the supernovae and σ_{int} the absolute magnitude dispersion of the SNe Ia population. The $(m_i, x_{1,i}, c_i)$ parameters are correlated and we denote the full covariance matrix *C*. The truncation effect can then be written as:

$$Y_i^{obs} = Y_i + \eta_i \text{ if } m_i \leqslant m_{lim} + \kappa_i \text{ with } \eta_i \sim \mathcal{N}(0, C_i) \text{ and } \kappa_i \sim \mathcal{N}(0, \sigma_d^2)$$

$$Y_i^{obs} \text{ is unobserved otherwise}$$
(2)

where η_i is the measurement noise, m_{lim} is the absolute magnitude of the survey and κ_i is its fluctuation due to the variability in observation conditions σ_d . The negative log-likelihood function associated to the model described above is based on the standard likelihood associated to the multivariate normal distribution. The novelty lies in the sum in equation 3. Starting from the Bayes theorem, we derived two new terms for the likelihood. Both depend on the cumulative distribution function of the normal distribution and the selection function parameters, allowing to take into account the truncation of data⁶.

$$\Gamma = -\ln(|W|) + r^{\dagger}Wr + \sum_{i} 2\ln\left(\Phi\left(\frac{m_{lim} - M - \mu_{i} - \alpha x_{1,i}^{*} - \beta c_{i}^{*}}{\sqrt{\sigma_{int}^{2} + \sigma_{d}^{2}}}\right)\right) - 2\ln\left(\Phi\left(\frac{m_{lim} - m_{i}^{obs}}{\sqrt{\sigma_{d}^{2} + f(C_{i})}}\right)\right)$$
(3)

with
$$\Phi(z) = \frac{1}{2} \left(1 + \operatorname{erf}\left(\frac{z}{\sqrt{2}}\right) \right), W = C^{-1} \text{ and } r = \begin{pmatrix} m^{oos} \\ x_1^{obs} \\ c^{obs} \end{pmatrix} - \begin{pmatrix} M^* - \mu - \alpha x_1^* - \beta c^* \\ x_1^* \\ c^* \end{pmatrix}.$$

The modeling of the Malmquist bias described is equivalent to the modeling of the selection functions by three independant sigmoids (one per survey). Each sigmoid is characterized by a central value (m_{lim}) and a width (σ_d) .

3 Estimation of distances and cosmological parameters

In statistics, the estimator of the variance (including σ_{int} in our case) is biased. As a consequence, EDRIS is intrinsically biased. To quantify this effect, we performed 100 Monte-Carlo simulations. The simulated SNe Ia light-curves and redshifts are generated from observation logs with SKYSURVEY^a. For this analysis, we consider a simplified case with only one standardization parameter (colour). This allows to fit the light-curves with a 1D version of SALT³. Thus, we keep the complexity of our problem (the light-curve model is still trained) while running the analysis faster. Simulations parameters are presented in table 1 while results are presented in figure 1. The left panel shows the mean of the difference between the reconstructed distance and the simulation input in 30 bins logarithmically distributed in redshift. The blue points correspond to the classic maximum likelihood estimator for multivariate normal distribution and show a strong negative bias on reconstructed distances due to the selection effect for each survey. The orange points correspond to the distance estimator including the modeling of the selection effect and show no significant bias on Ω_m (see table 2).

^ahttps://skysurvey.readthedocs.io/en/latest/



Figure 1: *Left panel:* Bias on the reconstructed binned distances moduli for each estimator. *Right panel:* Histogram of the reconstructed matter energy density parameter for each estimator.

Table 1:	Simulations	parameters	for the	he Monte-Carlo	analysis.
----------	-------------	------------	---------	----------------	-----------

Survey	Redshift range	β	M	Cosmology	m_{lim}	σ_d
ZTF	[0.0, 0.2]	3.15	-19.0	Planck 2018 ¹	18.59	0.1
SNLS 5y	[0.0, 1.07]	3.15	-19.0	Planck 2018 ¹	24.38	0.25
HSC/Subaru	[0.0, 1.6]	3.15	-19.0	Planck 2018 ¹	25.2	0.07

Table 2: Bias and uncertainty on Ω_m derived from Monte-Carlo simulations for both estimators.

Estimator	Bias on Ω_m	Mean of reconstructed Ω_m	$\sigma(\Omega_m)$
classic	0.0093 ± 0.0007	0.309	0.007
EDRIS	-0.0010 ± 0.0007	0.299	0.007

4 Acceleration of the computation

As the covariance matrix of the observations depends on the estimated σ_{int} parameter, we need to invert it at each step of the likelihood minimization. However, this dependency is simple enough to take advantage of the Schur complement technique⁵. The first step consists in taking into account the block structure of C (see equation 4) and writing r as (r_1, r_2) to match this structure.

$$W = \begin{pmatrix} C_{mm} + \sigma_{int}^2 I_N & C_1 \\ C_1^{\dagger} & C_2 \end{pmatrix}^{-1}$$
(4)

Then, we can compute and diagonalize (see equation 5) the Schur complement $S = C_{mm} + \sigma_{int}^2 I_N - C_1 C_2^{-1} C_1^{\dagger}$ of the lower right block C_2 in C.

$$S^{-1} = Q(\Lambda + \sigma_{int}^2 I_N)^{-1} Q^{\dagger}$$
⁽⁵⁾

The two first terms of the likelihood can then be written:

$$-\ln(|W|) = \sum_{i} \ln(\Lambda_i + \sigma_{int}^2) + \ln(|C_2|)$$
(6)

$$r^{\dagger}Wr = r_{1}^{\dagger}S^{-1}r_{1} - 2r_{1}^{\dagger}S^{-1}C_{1}C_{2}^{-1}r_{2} + r_{2}^{\dagger}C_{2}^{-1}r_{2} + r_{2}^{\dagger}C_{2}^{-1}C_{1}^{\dagger}S^{-1}C_{1}C_{2}^{-1}r_{2}$$
(7)

The last step consists in precomputing every constant matrix-matrix products in equation 7. Thus, only matrix-vector products remain, which allows to scale the likelihood computation in $O(N^2)$ instead of $O(N^3)$ when working with naive determinant computation and matrix inversion. To push the optimization even further, we decided to implement a hessian-free minimization method using JAX^b. The total computing time is presented in figure 2. We manage to reach a minimization in O(10s) for 1500 supernovae, including the standardization of the magnitudes, the estimation of σ_{int} and the correction of the Malmquist bias.



Figure 2: Time scaling of the likelihood minimization as a function of the number of simulated supernovae. The blue, orange and green lines respectively account for the precomputation of constant terms, the cosmological inference and the total time.

References

 Aghanim, N., Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, et al. Astronomy and Astrophysics, vol. 641, September 2020, p. A6.

^bhttps://jax.readthedocs.io/en/latest/

- 2. Betoule, M., et al. Astronomy and Astrophysics, vol. 568, August 2014, p. A22.
- 3. Hazenberg, François. PhD thesis, Sorbonne Université, 2019.
- 4. Kessler, Richard, et Dan Scolnic. The Astrophysical Journal, vol. 836, no 1, February 2017, p. 56.
- Ouellette, Diane Valérie. Linear Algebra and its Applications, vol. 36, March 1981, p. 187-295.
- 6. Rubin, David, et al. The Astrophysical Journal, vol. 813, no 2, November 2015, p. 137.
- 7. Teerikorpi, P. Astronomy and Astrophysics, vol. 339, November 1998, p. 647-57.
- 8. Tripp, Robert. Astronomy and Astrophysics, vol. 331, March 1998, p. 815-20.