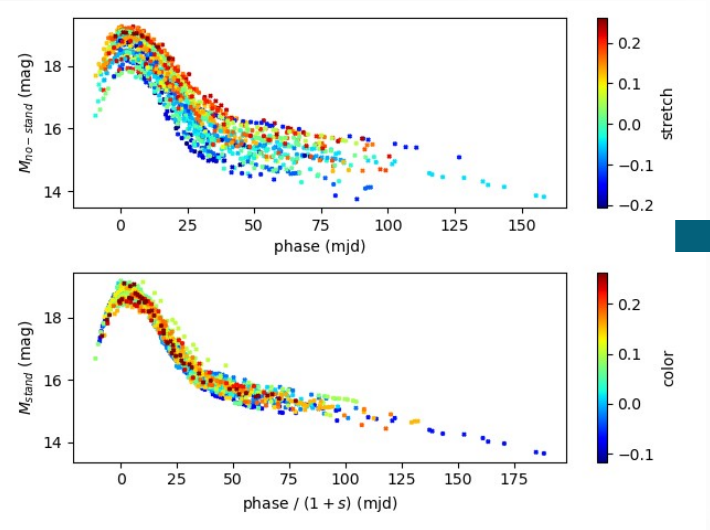
Distance estimation for truncated SNIa surveys

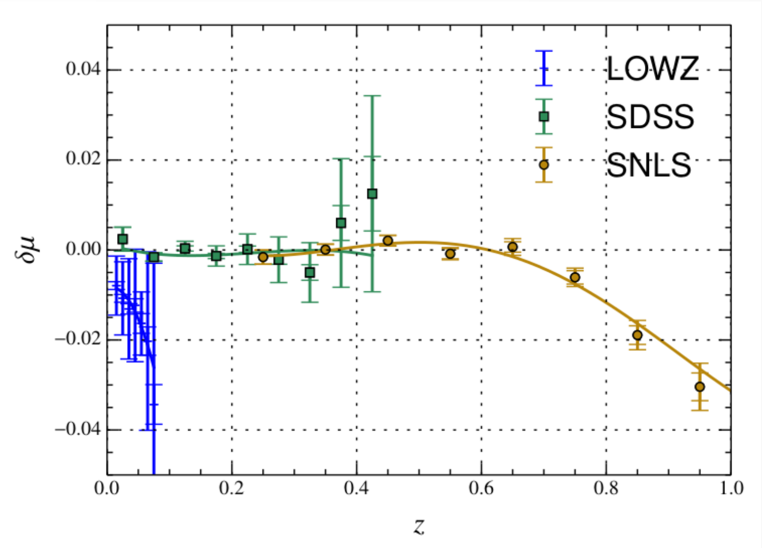
Dylan KUHN, Marc BETOULE, Nicolas REGNAULT
LPNHE Paris
September, 11th 2023

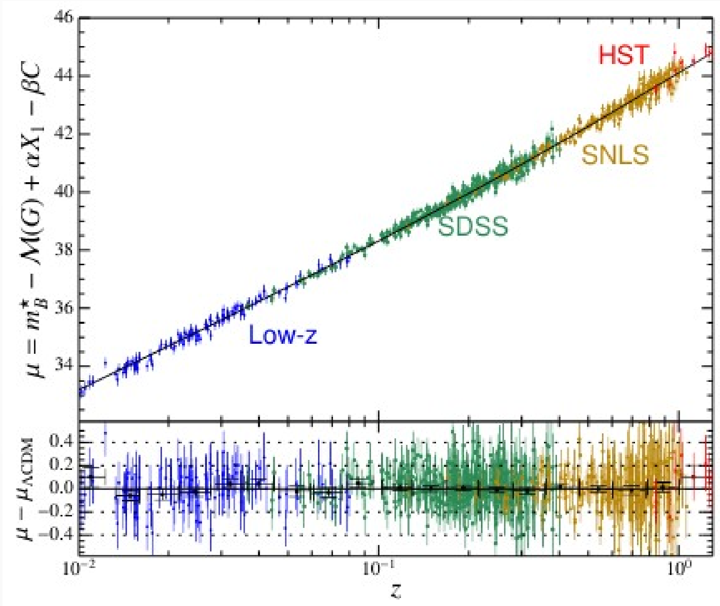
Integration of the work in the cosmological analysis



SNIa light-curves

$$\begin{pmatrix} m_B \\ x_1 \\ c \end{pmatrix} + \mathrm{Cov}(m_B, x_1, c)$$





My current work :
Preparation of an article (methodology)
→ how to do the last part in a single step ?

Determination of Malmquist bias & correction of distances

Credits : Marc Betoule et al., 2014

Thesis work
Development of a new distance estimator without Malmquist bias

Illustration of the bias on a toy model

- Toy model :

$$m_i^* = M^* + \mu_i + \epsilon_i \text{ with } \epsilon_i \sim \mathcal{N}(0, \sigma^2)$$

Information compressed in binned distance moduli

- The truncation is modelled as follows :

$$m_i = m_i^* \text{ if } m_i^* \leqslant m_{lim}$$

$$m_i \text{ is unobserved otherwise}$$

Useful relations :

$$\mu = \Xi\xi$$

$$\Phi(z) = \frac{1}{2}\left(1 + \text{erf}\left(\frac{z}{\sqrt{2}}\right)\right)$$

- The associated negative log-likelihood function is :

$$\Gamma = \sum_i 2\ln(\sigma) + \frac{1}{\sigma^2} r^\dagger r + \ln\left(\Phi\left(\frac{m_{lim} - M^* - \mu_i}{\sigma}\right)\right)$$

Related to the instrument

Intrinsic dispersion

Issues when dealing with real supernovae

- Type Ia supernovae are standardized :

$$m_i = M^* + \mu_i \longrightarrow m_i = M^* + \mu_i + \underbrace{\alpha_1 Y_1 + ... + \alpha_n Y_n}_{\alpha x_1 + \beta c}$$

- Additional noise :

$$\eta \sim \mathcal{N}(0, \underbrace{\text{Cov}(m, Y_1, ..., Y_n)}_{= C})$$

- Selection function depends also on the weather :

  → introduction of fluctuations on the limit magnitude σ_d

Model with latent parameters

EDRIS

French for ‘Distance Estimator for Incomplete Supernovae Surveys’

Model used for the EDRIS analysis

$$\begin{pmatrix} m^*{}_i \\ Y^*_{1\,i} \\ Y^*_{2\,i} \\ \vdots \\ Y^*_{n\,i} \end{pmatrix} = \begin{pmatrix} \mu_i(z,\theta) \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_n \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X^*_{1\,i} \\ X^*_{2\,i} \\ X^*_{3\,i} \\ \vdots \\ X^*_{n\,i} \end{pmatrix} + \begin{pmatrix} \epsilon_i \\ 0 \\ \vdots \\ 0 \end{pmatrix} \text{ with } \epsilon_i \sim \mathcal{N}(0,\sigma^2)$$

LCs parameters | Distances | Standardization coefficients | Latent parameters | Noise depending on the intrinsic dispersion

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

$$Y_i \text{ is unobserved otherwise}$$

Truncation

Negative log-likelihood function

$$\Gamma = -\ln(|W|) + r^\dagger W r + \sum_i \left[2\ln\left(\Phi\left(\frac{m_{lim} - \mu_i - \alpha_1 X^*_{1\,i} - \cdots - \alpha_n X^*_{ni}}{\sqrt{\sigma_d^2 + \sigma^2}}\right)\right) - 2\ln\left(\Phi\left(\frac{m_{lim} - m_i}{\sqrt{\sigma_d^2 + f(C_i)}}\right)\right)\right]$$

allows to estimate the intrinsic dispersion

classic chi2 term

takes into account the truncation effects

Parameters :

ξ : distance moduli at the nodes of the spline

$(\alpha_1 \ \cdots \ \alpha_n)$: standardization coefficients

X^* : latent parameters

σ : intrinsic dispersion

m_{lim} : limit magnitude

σ_d : fluctuations of the limit magnitude

Useful relations :

$$\mu = \Xi\xi$$

$$\Phi(z) = \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{z}{\sqrt{2}}\right)\right)$$

Fast computation of the likelihood

- Inversion of the covariance matrix by the Schur complement technique :

$$W = \begin{pmatrix} C^{mm} + \sigma^2 I_N & C_1 \\ C_1^\dagger & C_2 \end{pmatrix}^{-1} \Rightarrow W = \begin{pmatrix} S^{-1} & -S^{-1}C_1C_2^{-1} \\ -C_2^{-1}C_1^\dagger S^{-1} & C_2^{-1} + C_2^{-1}C_1^\dagger S^{-1} C_1 C_2^{-1} \end{pmatrix}$$

$$\text{with } S = C^{mm} + \sigma^2 I_N - C_1 C_2^{-1} C_1^\dagger$$

- By writing :

$$C^{mm} - C_1 C_2^{-1} C_1^\dagger = Q\Lambda Q^\dagger$$

we obtain :

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

Fast computation of the likelihood

- We rewrite r to match the structure of W : $r = (r_1, r_2)$
- The chi2 term writes as follows :

$$r^\dagger W r = r_1^\dagger S^{-1} r_1 - 2 r_1 S^{-1} C_1 C_2^{-1} r_2 + r_2^\dagger C_2^{-1} + r_2^\dagger C_2^{-1} C_1^\dagger S^{-1} C_1 C_2^{-1} r_2$$

- The determinant of W writes as follows :

$$-\ln(|W|) = \ln(|C_2|) + \ln(|S|) = \ln(|C_2|) + \sum_i \ln(\Lambda_i + \sigma^2)$$

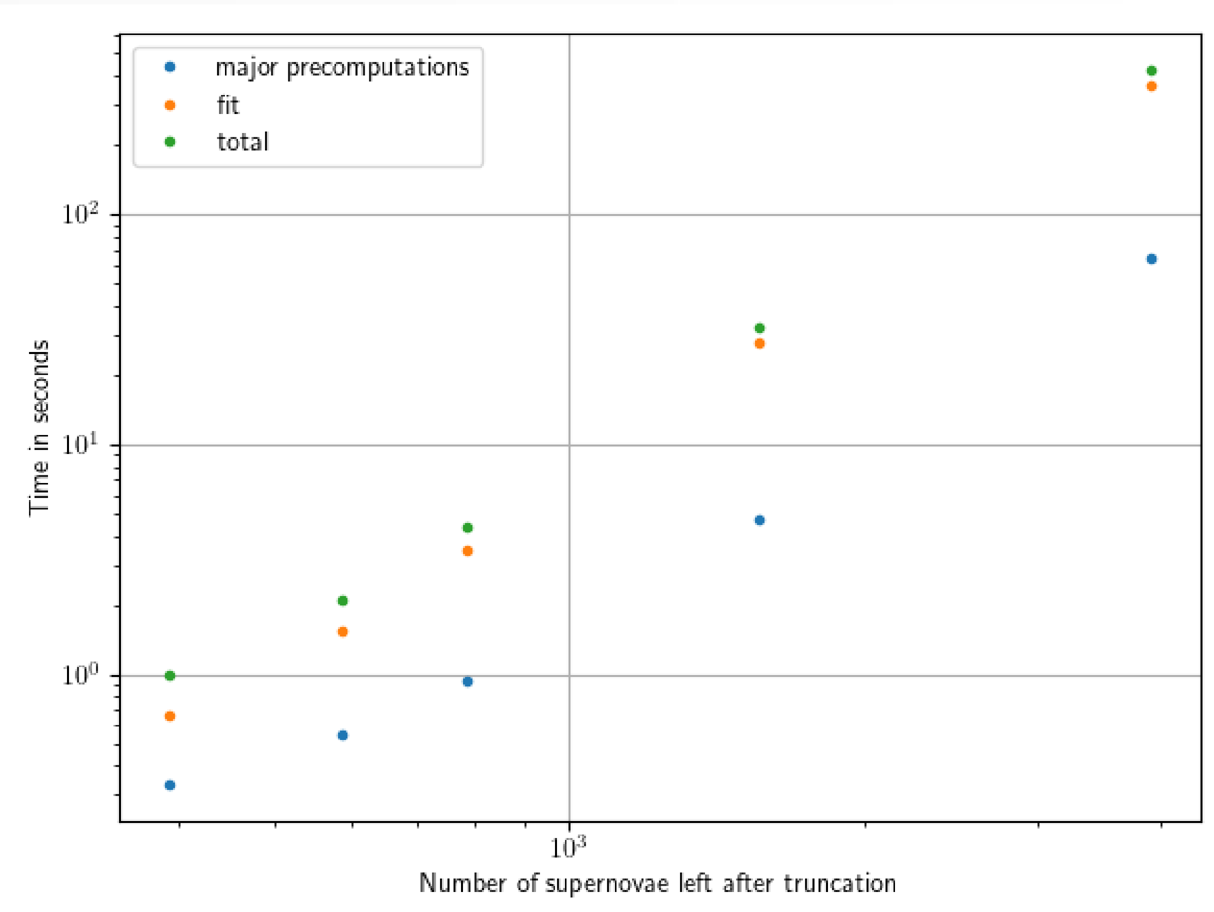
Only matrix-to-vector products

- At the end of the day, for one iteration we only need to compute :

$$\left(r \quad (\Lambda + \sigma^2 I_N)^{-1} \quad \sum_i \ln(\Lambda_i + \sigma^2)\right)$$

Computation in $O(N^2)$

Time scaling



Fit steps :

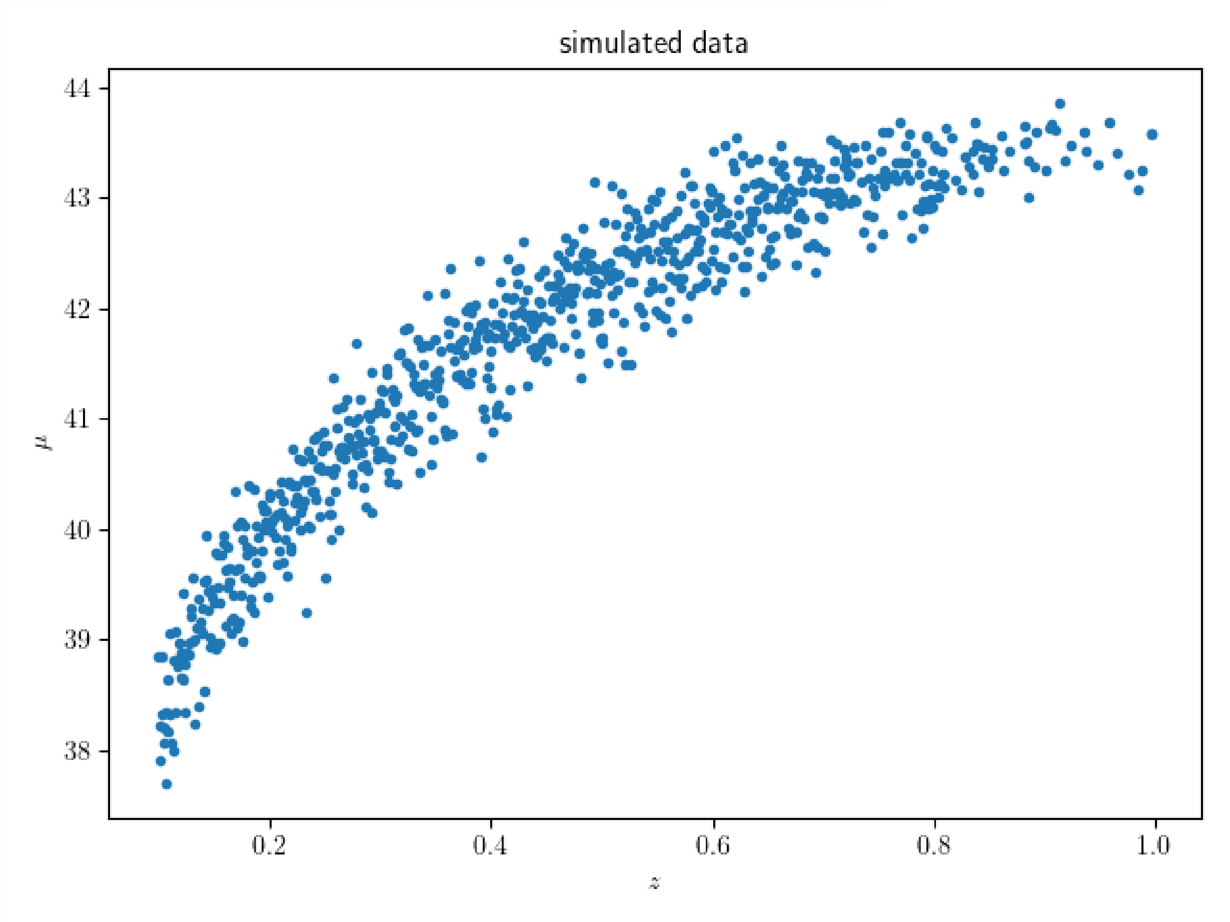
→ blue dots : major precomputations (see maths)

→ orange dots : minimization, dominated by the construction of the hessian matrix at each step (likelihood and gradient in $O(N^2)$)

(Plot regularly updated with each new optimization)

Characterization of the estimator with simulations

Exemple of simulated data



Parameters of the simulation :

$$N_{SN} = 1000$$
$$N_{bins} = 30$$
$$m_{lim} = 24.5$$
$$\sigma_d = 0.2$$
$$\sigma = 0.1$$
$$x_1 \sim \mathcal{N}(0,1)$$
$$c \sim \mathcal{N}(0,0.1)$$
$$\text{FLCDM : } H_0 = 70 \ \& \ \Omega_m = 0.3$$

Monte-Carlo simulation

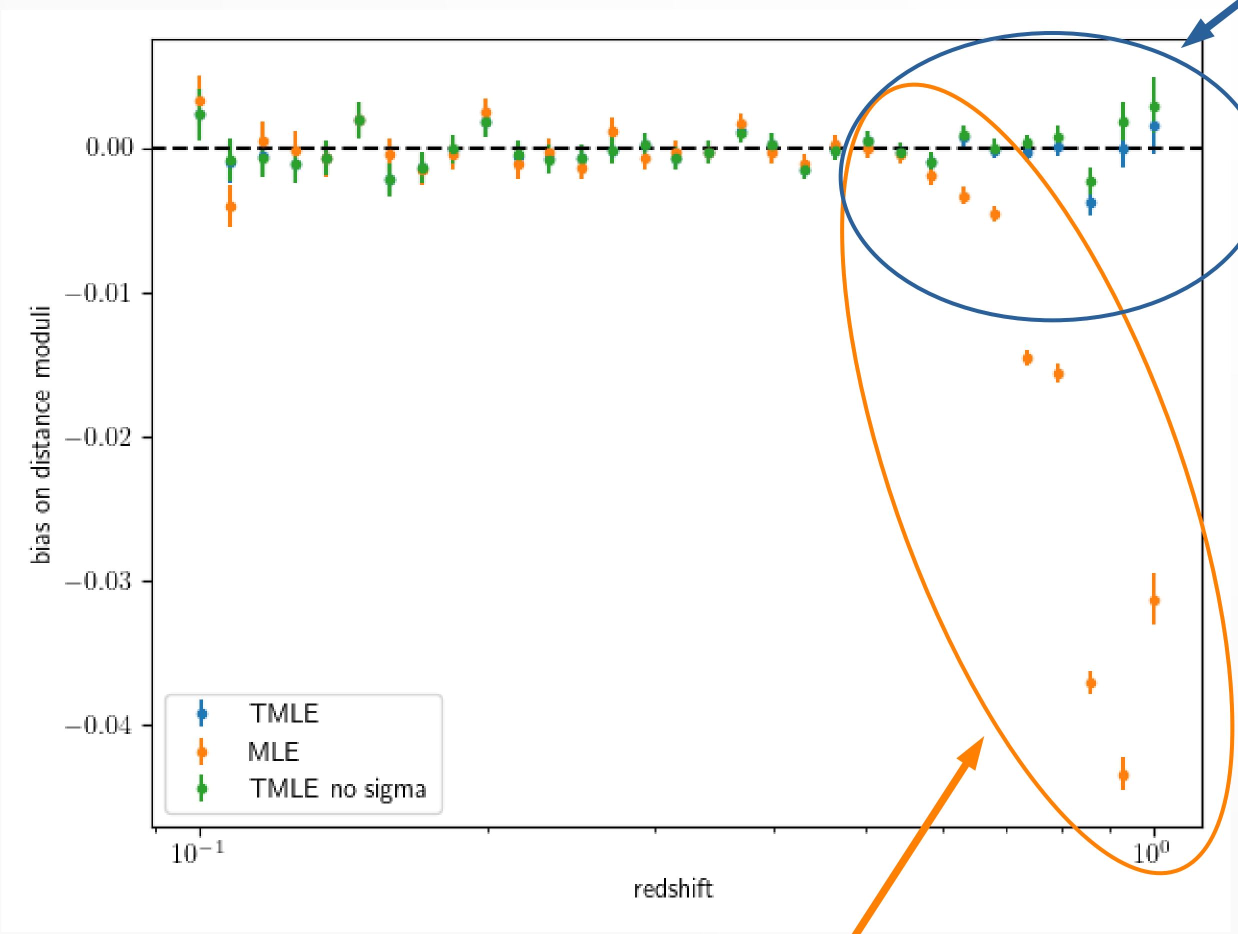
- We compare the bias of two estimators with a Monte-Carlo simulation :

  → Estimator 1 (MLE) : classic maximum likelihood estimator associated with the following negative log-likelihood function :

$$\Gamma = -\ln(|W|) + r^{\dagger} W r$$

  → Estimator 2 (TMLE) : new maximum likelihood estimator with takes into account the truncation

Results when SN are very well measured



corrected Malmquist bias

apparent Malmquist bias

Fitted parameters for the MLE :

$$(\xi \quad \alpha_1 \quad \cdots \quad \alpha_n \quad \sigma \quad X^*)$$

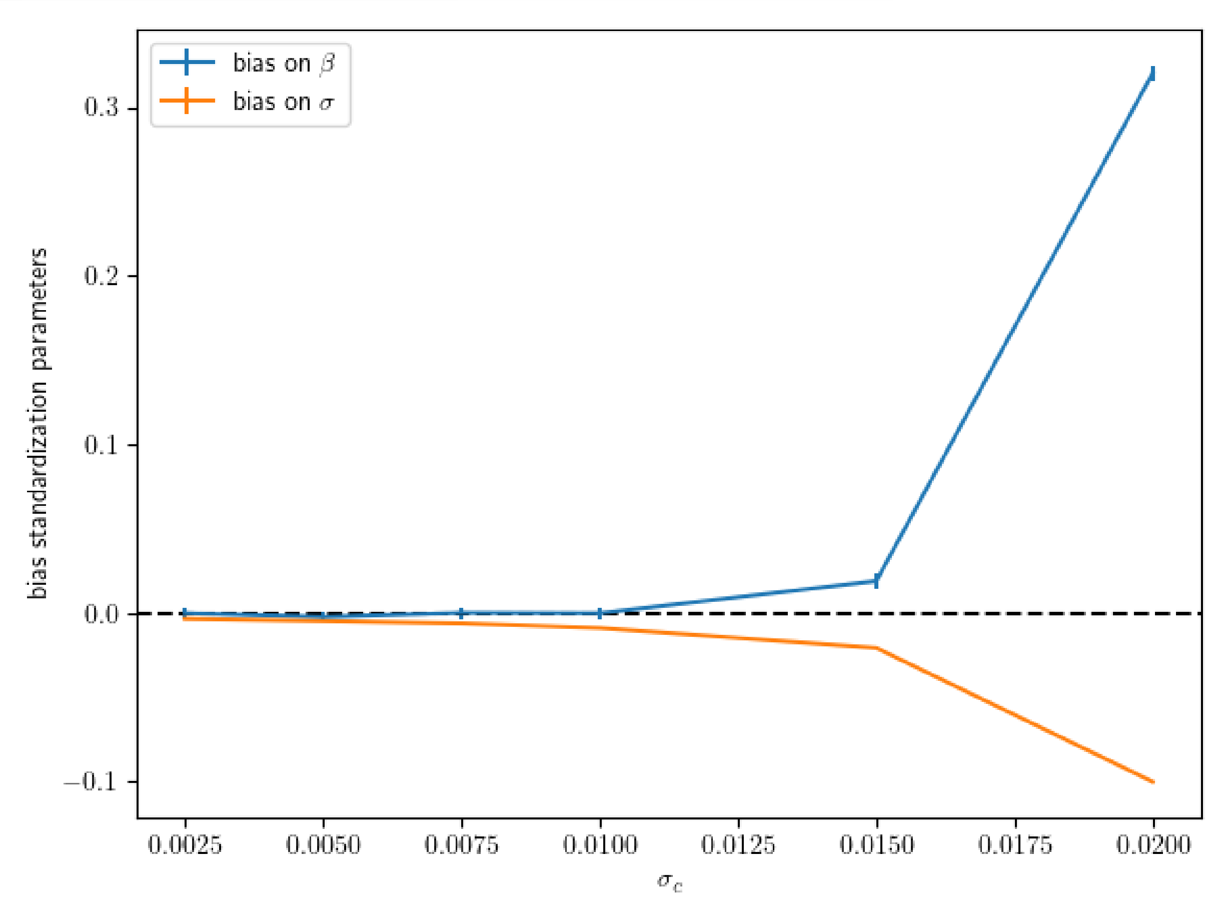
Fitted parameters for the TMLE :

$$(\xi \quad \alpha_1 \quad \cdots \quad \alpha_n \quad \sigma \quad X^*)$$

Fixed parameters for the TMLE :

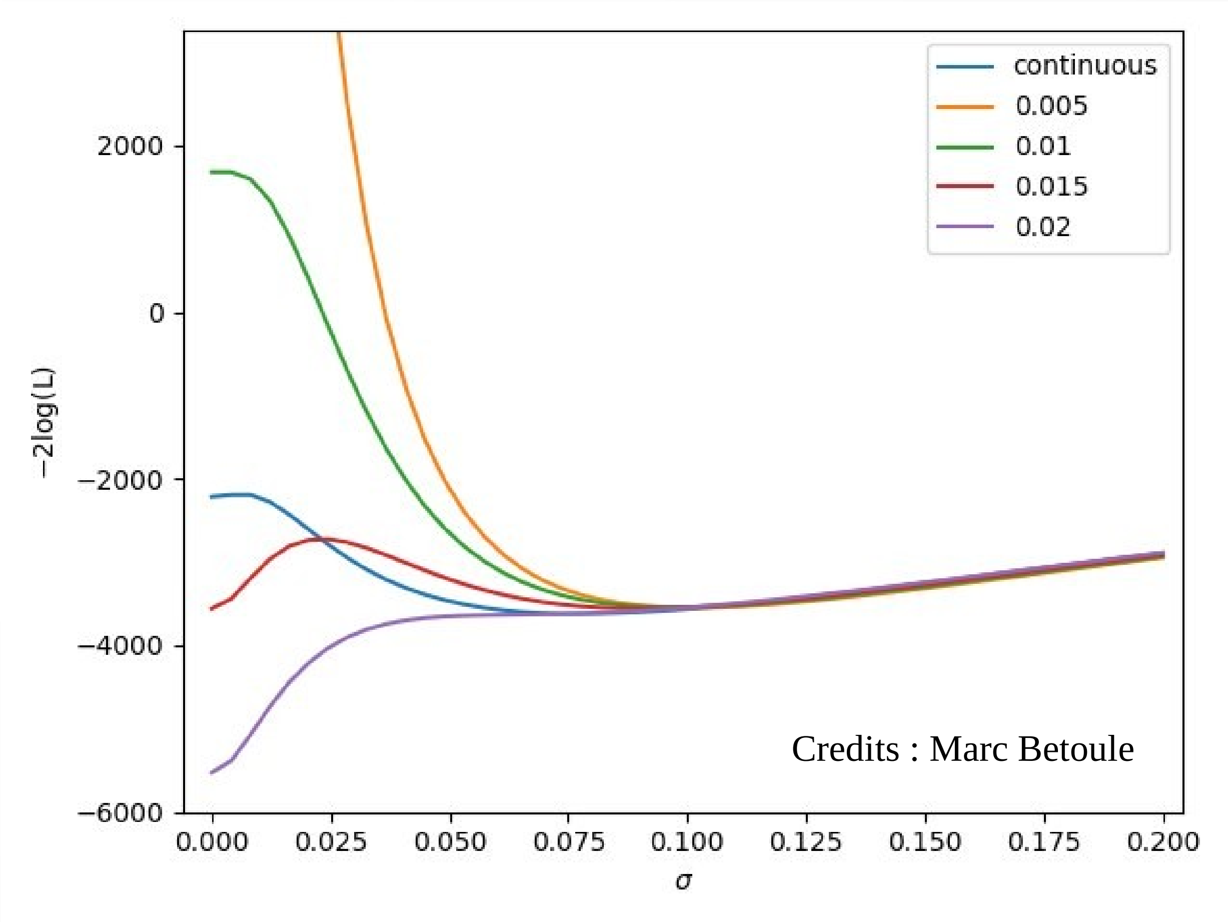
$$(m_{lim} \quad \sigma_d)$$

Bias of the TMLE when SN are less well measured



The TMLE is biased when

$$\beta\sigma_c \sim \sigma$$



Similar effect when we increase the error on the stretch

Restricted Maximum Likelihood Estimator (WIP)

- When using the MLE, denoting n the number of data and k the number of degrees of freedom

$$\mathbb{E}(\hat{\sigma}) = \frac{n-k}{n}\sigma$$

→ thus the variance estimator is biased

- Restricted Maximum Likelihood Estimator (ReMLE) allows to unbias the variance estimator by reducing the number of degrees of freedom
- Implemented on a simple toy model :

  → only color for standardization & no truncation
- Seems to work pretty well for now

Conclusions

Work in progress

- Optimizing the code
- For the paper :

1°) Finding a solution to correct bias on standardization parameters and intrinsic dispersion :

Merge TMLE with ReMLE

2°) Estimating the parameters of the selection function :

Limit magnitude & fluctuations

3°) Writing a full simulation pipeline based on *skysurvey* to study the behaviour of the estimator when we deviate from initial hypothesis :

Selection applied on observed magnitudes instead of reconstructed magnitudes, selection depending on the wavelenght/filter, training of SALT2/NaCl with a truncated dataset, ...

Thanks for listening

Do not hesitate to ask questions