Impact of the Tinker mass function on galaxy cluster constraints for cosmology in Rubin/LSST-DESC

T. Colin[1] M. Aguena Da Silva[2] S. Mei[3]

[1]Ecole Normale Supérieure Paris

[2,3]Astroparticule et Cosmologie

Rubin-LSST France meeting

Table of Contents

Table of Contents

Mass function

Define the mass function : $\frac{d\bar{n}(M,z)}{d\ln M} = \frac{\bar{\rho}_m}{M}\frac{d\ln\sigma^{-1}}{d\ln M}f(\sigma)$

- Press-Schechter, 1974 : $f(\sigma) = \sqrt{\frac{2}{\pi}}\frac{\delta_c}{\sigma}\exp\left(-\frac{\delta_c^2}{2\sigma^2}\right)$
- Jenkins, 2001 : $f(\sigma) = a\exp\left(-|b - \ln\sigma|^c\right)$
- Tinker, 2008 : $f(\sigma) = A\left[\left(\frac{\sigma}{b}\right)^{-a} + 1\right]\exp\left(-c/\sigma^2\right)$
  - A : overall amplitude of the mass function
  - a : slope of the low-mass power law
  - b : amplitude of the low-mass power law
  - c : cutoff scale at which the abundance of halos exponentially decreases

Halo number counts

The mean number density of halos in a mass bin α :

$$\bar{n}_\alpha(z) \equiv \int_{M_\alpha}^{M_{\alpha+1}} \frac{dM}{M} \frac{d\bar{n}}{d\ln M}$$

The mean value of dark matter halo number counts in a redshift bin i and mass bin α :

$$\bar{m}_{\alpha,i} = \int dV_i \bar{n}_\alpha(z) = \Delta\Omega \int_{z_i}^{z_{i+1}} dz \frac{D_A^2(z)}{H(z)} \bar{n}_\alpha(z)$$

Table of Contents

Bayes inference

Bayes' rule : posterior $= \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}, \mathcal{P}(\theta|Y) = \frac{\mathcal{L}(Y|\theta)\Pi(\theta)}{p(Y)}$

- Likelihood : $\ln\mathcal{L}(Y|\theta) = -\frac{1}{2}\sum_{ij}(y_i - m_i)C_{ij}^{-1}(y_j - m_j) + \text{cst}$
- Prior : Uniform
- Evidence : Do not contribute

Parameter space sampled using Markov Chain Monte Carlo.

Table of Contents

cosmoDC2 simulation

- Sky area : 439.8 $\deg^2$
- 454k dark matter halos, $\log(M/M_\odot) \in [13.0, 15.3]$, Redshift $\in [0.015, 3.04]$

Simulation parameters

- 12×12 Bins $= \begin{cases} \text{logM} & \in [13.5, 15.0], \text{bin size} = 0.1, \text{last bin} = 0.4 \\ \text{z} & \in [0.10, 1.60], \text{bin size} = 0.1, \text{last bin} = 0.4 \end{cases}$
- Number of dark matter halos considered = 59k
- Overdensity $\Delta = 200$
- MCMC: emcee (24 walkers, 300-500 steps)

Modeling done with a modification on CCL (Cluster Cosmology Library, from DESC) to change halo mass function parameters.

Impact of interpolation

Time: for 12 logM and 12 bins z it takes 200s for full computation and 40s for interpolated computation (1700 interpolation points), 5 times faster!

Precision: Using DC2 cosmology $\{\Omega_c = 0.22, \sigma_8 = 0.8\}$, full computation is $\sim$ 1.3% higher than in the interpolated.

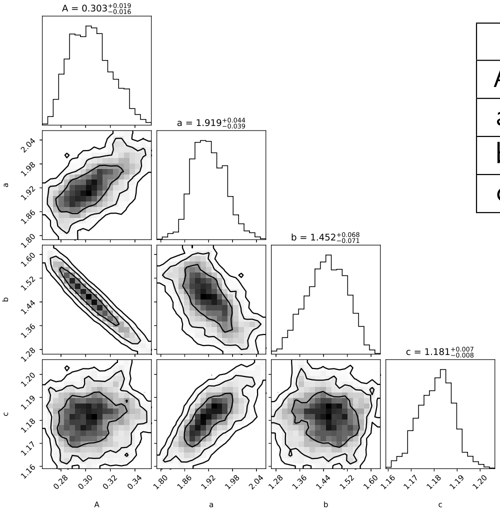
Cosmological constraints :

- N°1 : Tinker08 Full computation
  $\rightarrow \quad \{\Omega_{c,1} = 0.2039^{+0.0007}_{-0.0008}, \sigma_{8,1} = 0.8268^{+0.0007}_{-0.0007}\}$
- N°2: Tinker08 Interpolation
  $\rightarrow \quad \{\Omega_{c,2} = 0.2068^{+0.0007}_{-0.0008}, \sigma_{8,2} = 0.8265^{+0.0008}_{-0.0008}\}$

Same relative difference between $\Omega_{c,1}$ and $\Omega_{c,2}$

3rd case : Calibrating the mass function

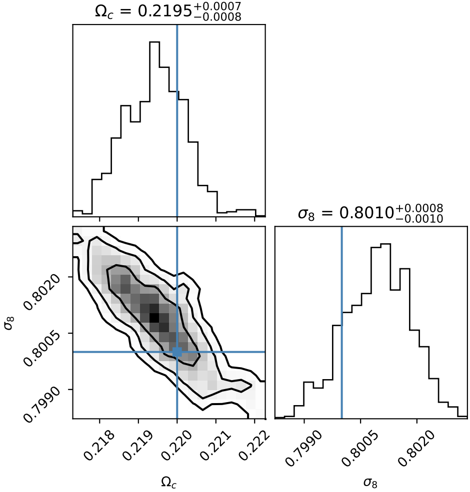
Using the original values $\Omega_c = 0.22$ and $\sigma_8 = 0.8$ from the DC2 simulation, we get new Tinker08 halo mass function parameters :



	Tinker08	Calibrated mf
A	0.186	$0.3030^{+0.019}_{-0.016}$
a	1.47	$1.919^{+0.044}_{-0.039}$
b	2.57	$1.452^{+0.068}_{-0.071}$
c	1.19	$1.181^{+0.007}_{-0.008}$

4th case : Consistency check

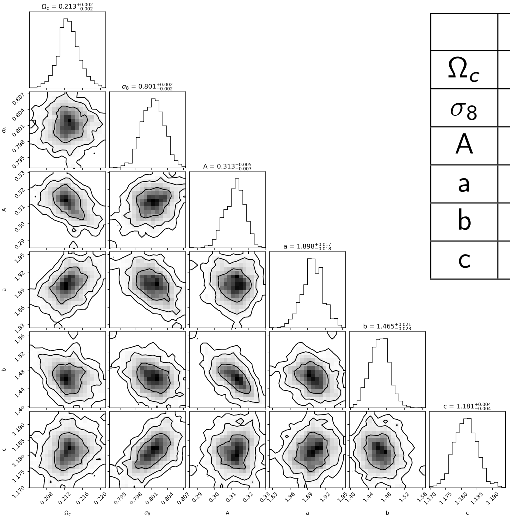
Take the previous maximum posterior values of {A, a, b, c} and try to find back the original imput parameters $\{\Omega_c = 0.22, \sigma_8 = 0.8\}$



	3rd	4th
Ω_c	0.220	$0.2195^{+0.0007}_{-0.0008}$
σ_8	0.800	$0.8010^{+0.0008}_{-0.0010}$
A	$0.3030^{+0.019}_{-0.016}$	0.3030
a	$1.919^{+0.044}_{-0.039}$	1.918
b	$1.452^{+0.068}_{-0.071}$	1.452
c	$1.181^{+0.007}_{-0.008}$	1.182

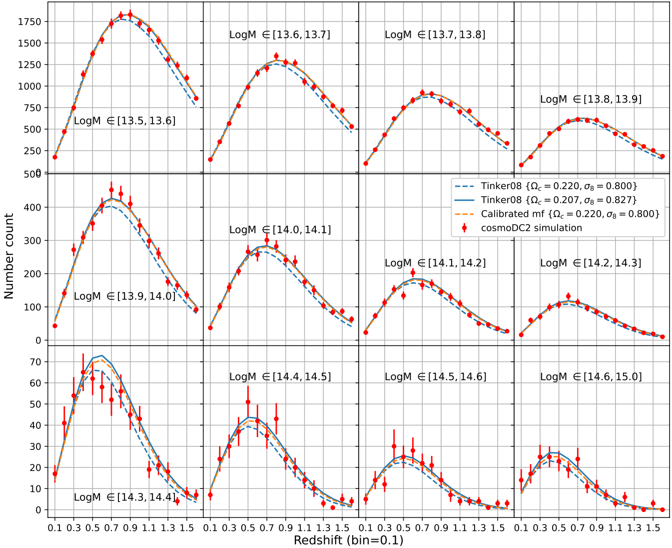
5th case : Cosmology and HMF

Approximate the previous posterior distributions by Gaussians, reuse them as new priors and leave the 6 parameters $\{\Omega_c, \sigma_8, A, a, b, c\}$ free.



	3rd	5th
Ω_c	0.220	$0.2128^{+0.0024}_{-0.0021}$
σ_8	0.800	$0.8013^{+0.0021}_{-0.0022}$
A	$0.3030^{+0.019}_{-0.016}$	$0.3127^{+0.0067}_{-0.0052}$
a	$1.919^{+0.044}_{-0.039}$	$1.898^{+0.017}_{-0.018}$
b	$1.452^{+0.068}_{-0.071}$	$1.465^{+0.021}_{-0.023}$
c	$1.181^{+0.007}_{-0.008}$	$1.181^{+0.004}_{-0.004}$

Difference between models



Simulation set

Recap :

N°	Method	Ω_c	σ_8	{A,a,b,c}
1	Full computation	$0.2039^{+0.0007}_{-0.0008}$	$0.8268^{+0.0007}_{-0.0007}$	{0.186, 1.47, 2.57, 1.19}
2	Interpolated model	$0.2068^{+0.0007}_{-0.0008}$	$0.8265^{+0.0008}_{-0.0008}$	{0.186, 1.47, 2.57, 1.19}
3	Interpolated model	0.220	0.800	{0.303, 1.92, 1.45, 1.18}
4	Interpolated model	$0.2195^{+0.0007}_{-0.0008}$	$0.8010^{+0.0008}_{-0.0010}$	{0.303, 1.92, 1.45, 1.18}
5	Interpolated model	$0.2128^{+0.0024}_{-0.0021}$	$0.8013^{+0.0021}_{-0.0022}$	{0.316, 1.89, 1.46, 1.18}

Table of Contents

Conclusions and perspectives

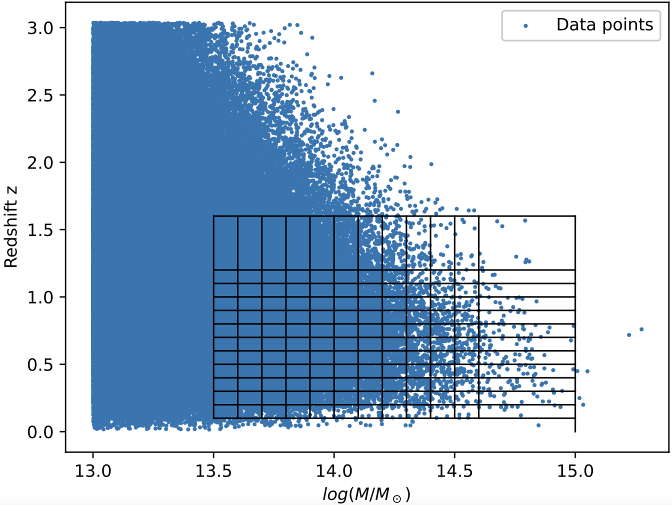
Conclusions

- We reduce computation time of Tinker08 model with interpolation.
- Need to find a compromise between precision and computation time.
- We find new Tinker08 parameters that fit cosmoDC2 better.
- **But** when all parameters are free, the function calibrated on the Tinker08 model does not provide the best prediction for cosmoDC2.

Perspectives

- Improve the interpolated function.
- Test new sets of data, see if the interpolated model is still consistent.
- Implement in the LSST pipeline a way to modify easily the Tinker08 halo mass function parameters.
- Test other halo mass function.

cosmoDC2 simulation : data points

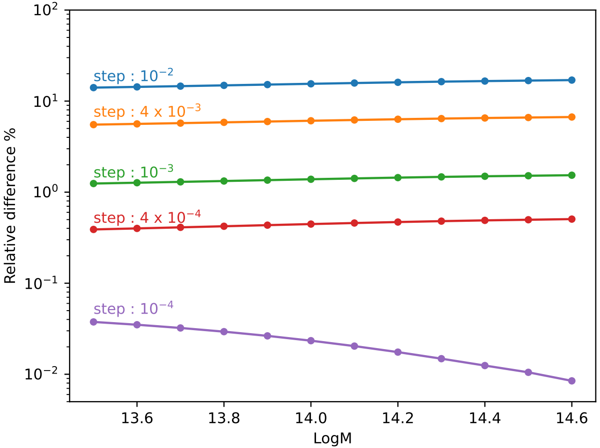


Add Poisson cases

N°	Method	Data Points	$\mathcal{L}$	Ω_c	σ_8	{A,a,b,c}
1	Full computation	DC2 simulation	FC	$0.2039^{+0.0007}_{-0.0008}$	$0.8268^{+0.0007}_{-0.0007}$	{0.186, 1.47, 2.57, 1.19}
2	Interpolated model	DC2 simulation	FC	$0.2068^{+0.0007}_{-0.0008}$	$0.8265^{+0.0008}_{-0.0008}$	{0.186, 1.47, 2.57, 1.19}
3	Interpolated model	DC2 simulation	FC	0.220	0.800	{0.303, 1.92, 1.45, 1.18}
4	Interpolated model	DC2 simulation	FC	$0.2195^{+0.0007}_{-0.0008}$	$0.8010^{+0.0008}_{-0.0010}$	{0.303, 1.92, 1.45, 1.18}
5	Interpolated model	DC2 simulation	FC	$0.2128^{+0.0024}_{-0.0021}$	$0.8013^{+0.0022}_{-0.0022}$	{0.316, 1.89, 1.46, 1.18}
6	Full computation	Gaussian distribution with Poisson error	P	$0.2238^{+0.0013}_{-0.0017}$	$0.7975^{+0.0019}_{-0.0016}$	{0.186, 1.47, 2.57, 1.19}
7	Interpolated model	Only Poisson error	P	0.2200	0.8000	{0.186, 1.47, 2.57, 1.19}
8	Interpolated model	Gaussian distribution with Poisson error	P	$0.2267^{+0.0016}_{-0.0021}$	$0.7972^{+0.0016}_{-0.0016}$	{0.186, 1.47, 2.57, 1.19}
9	Interpolated model	DC2 simulation	P	$0.2095^{+0.0015}_{-0.0015}$	$0.8200^{+0.0016}_{-0.0018}$	{0.186, 1.47, 2.57, 1.19}

Relative difference between Tinker08 and Tinker08 interp in function of the interpolation step of redshift

Sum over the redshift $\in [0.1, 1.6]$



Comparison 6th and 8th cases : Method

For the same initial parameters, the number of clusters in the Tinker08 model is approximately 1.3% higher than in the interpolated Tinker08.

Same data points generated by a Gaussian perturbation of the theoretical models $\{\Omega_c = 0.22, \sigma_8 = 0.8\}$ with Poisson likelihood :

- N°6 : Tinker08 Full computation
  $\rightarrow \quad \{\Omega_{c,6} = 0.2238^{+0.0017}_{-0.0013}, \sigma_{8,6} = 0.7975^{+0.0016}_{-0.0019}\}$
- N°8 : Tinker08 Interpolation
  $\rightarrow \quad \{\Omega_{c,8} = 0.2267^{+0.0021}_{-0.0016}, \sigma_{8,8} = 0.7972^{+0.0016}_{-0.0016}\}$

Same relative difference between $\Omega_{c,5}$ and $\Omega_{c,7}$.

Comparison 2nd and 9th cases : Likelihood

Use cosmoDC2 simulation with interpolated Tinker08 :

- N°2 : Full covariance Likelihood
  $\rightarrow \quad \{\Omega_{c,2} = 0.2068^{+0.0007}_{-0.0008}, \sigma_{8,2} = 0.8265^{+0.0008}_{-0.0008}\}$
- N°9 : Poisson Likelihood
  $\rightarrow \quad \{\Omega_{c,9} = 0.2095^{+0.0015}_{-0.0015}, \sigma 8, 9 = 0.8200^{+0.0016}_{-0.0018}\}$

Tinker08 interpolation

- To reduce the computation time, let's interpolate the Tinker08 mass function:

$$\bar{m}_{\alpha,\mathcal{I}} = \int_{\mathcal{I}} \mathrm{d}z \left(\int_{\alpha} \mathrm{dlog}M \frac{d\bar{n}}{dM}(\mathrm{cosmo}) \right) \mathrm{d}V(z)$$

- For a given $\alpha \equiv \log M$ and $z_i \in [0.0, 1.7]$, step of 0.001 :

$$g^{\alpha}(z_i) = \left(\int_{\log M_{\alpha}}^{\log M_{\alpha} + \mathrm{dlog}M} \mathrm{dlog}M \frac{d\bar{n}}{dM}(\mathrm{cosmo}) \right) \mathrm{d}V(z_i)$$

- Introduce $g^{\alpha}_{\mathrm{interp}} = \mathrm{Interp}(z_i, g^{\alpha}(z_i))$
- It finally gives : $\bar{m}_{\alpha,\mathcal{I}} = \int_{\mathcal{I}} \mathrm{d}z \; g^{\alpha}_{\mathrm{interp}}$

Time rule

- Full computation : 1 point (logM,z) = 1.4s
- Interpolation : 1 vector (logM,$\forall z$) x 1000 interpolation points = 2s

Example : Our grid : 12 logM, 12z and 1700 interpolation points :

- Full computation : 12 x 12 x 1.4 = 200s
- Interpolation : 12 x 1.7 x 2 = 40s

$\longrightarrow$ Here interpolation model is 5 times faster.