Likelihoods for cluster count cosmology

Constantin Payerne, Calum Murray, Céline Combet

Laboratoire de Physique Subatomique et Cosmologie, Grenoble, UGA

Rubin LSST France meeting November 22-24, 2021

Summary

1. Cluster count cosmology
2. Likelihoods for cluster count cosmology
   A. Standard likelihoods
   B. MPG likelihood
3. Framework for testing likelihood accuracies

Cosmology with galaxy clusters: Cluster abundance

Galaxy clusters:

- Are the largest gravitationally bound objects in the Universe
- Mass > 10^{14} solar masses

Cosmology with galaxy clusters: Cluster abundance

Galaxy clusters:

- Are the largest gravitationally bound objects in the Universe
- Mass > 10^{14} solar masses

Cluster abundance : Count the number of clusters in mass-redshift bins

A. Halo number density
B. Survey comoving volume

- Formation history of the Universe (amount of matter, $\boldsymbol{\Omega}_m$)
- Fluctuation of matter density field (fluctuation amplitude, $\boldsymbol{\sigma}_8$)


NASA, ESA, HST

Cosmology with galaxy clusters: Cluster abundance

Galaxy clusters:

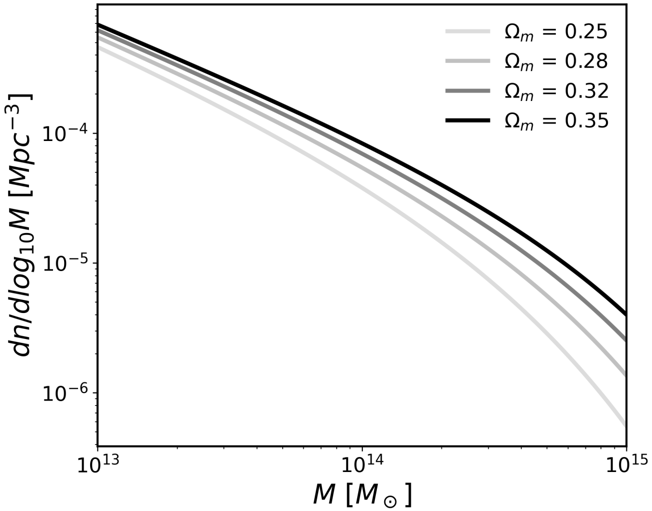
- Are the largest gravitationally bound objects in the Universe
- Mass > 10^{14} solar masses

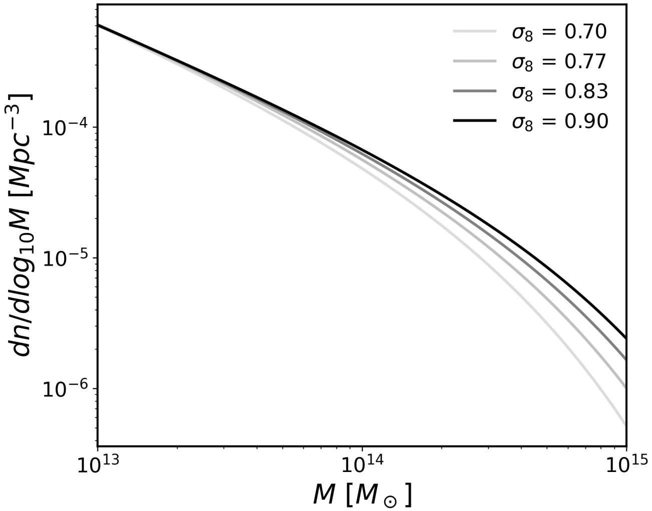
Cluster abundance : Count the number of clusters in mass-redshift bins

A. Halo number density
B. Survey comoving volume

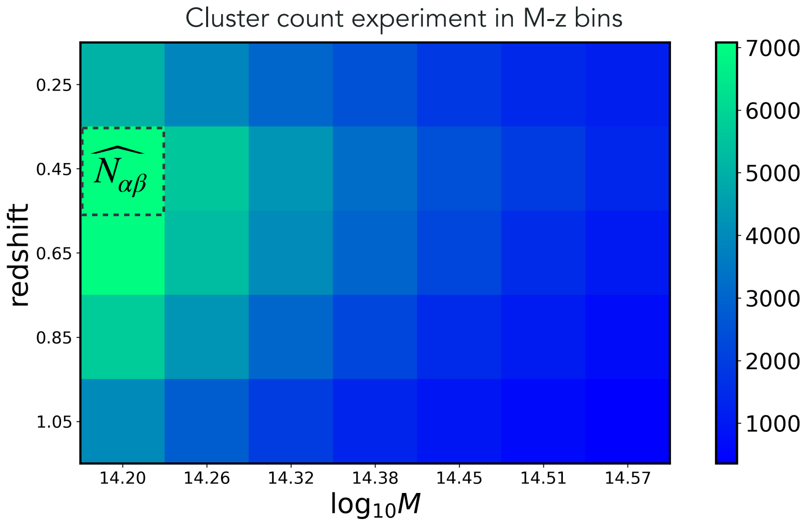
- Formation history of the Universe (amount of matter, $\boldsymbol{\Omega}_m$)
- Fluctuation of matter density field (fluctuation amplitude, $\boldsymbol{\sigma}_8$)

Impact of $\boldsymbol{\Omega}_m$ and $\boldsymbol{\sigma}_8$ on the halo mass function (z = 0)





Cosmology with galaxy clusters: Cluster abundance



Basic recipe for cluster abundance cosmology

a. From a galaxy cluster survey with known redshifts, masses
b. Count the galaxy clusters within several redshift and mass bins

$$\widehat{N_{\alpha\beta}} \longrightarrow N_{\alpha\beta}^{\text{predicted}} = \Omega_s \int_{z_\alpha}^{z_{\alpha+1}} \frac{d^2V(z)}{dz d\Omega} dz \int_{\log_{10} M_\beta}^{\log_{10} M_{\beta+1}} \frac{dn(M,z)}{d \log_{10} M} d \log_{10} M$$

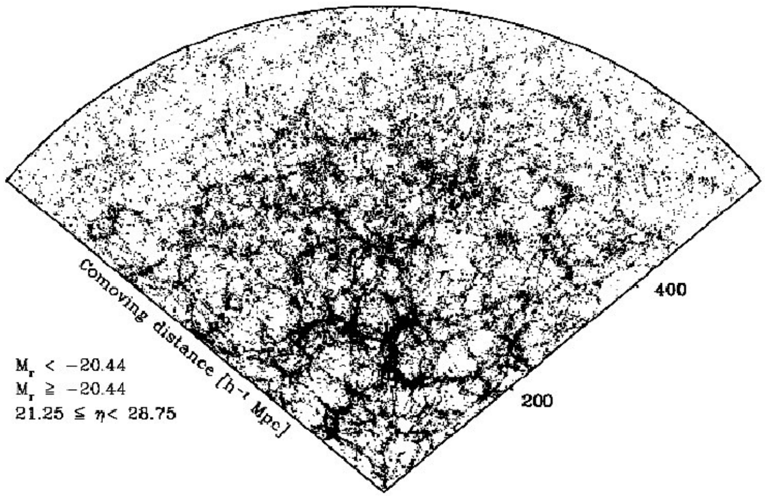
Observed

differential comoving volume ($\Omega_{\rm m}$)

Halo mass function ($\Omega_{\rm m}, \sigma_8$)

Cosmology with galaxy clusters: Variance

- From the dark matter halo distribution in the Universe → derive a statistical model for describing what we observe!
- Know the statistical properties of cluster abundance
- Cluster abundance as a Poisson variable?
  - Poisson counting experiment : discrete, un-correlated random count
  - Poisson variance $\sigma^2(N) = N$
  - N : average abundance over many realisations of the same cosmology
- Additional variance σ^2_{sample}, from fluctuations of the matter density field



Sloan Digital Sky Survey, Park et al. 2005

$$\sigma^2(N) = N + \sigma^2_{\text{sample}}(N)$$

Poisson variance/Shot noise → N

Sample variance → $\sigma^2_{\text{sample}}(N)$

σ^2_{sample} depends on:

- matter power spectrum $P_{\text{mm}}(k)$
- mass-redshift range considered
- Survey geometry

→ increases with the number of halos N per M-z bins

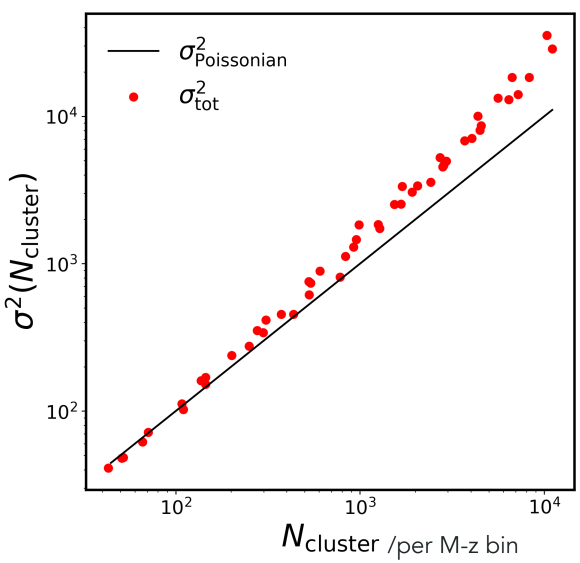
Covariance matrix for cluster count

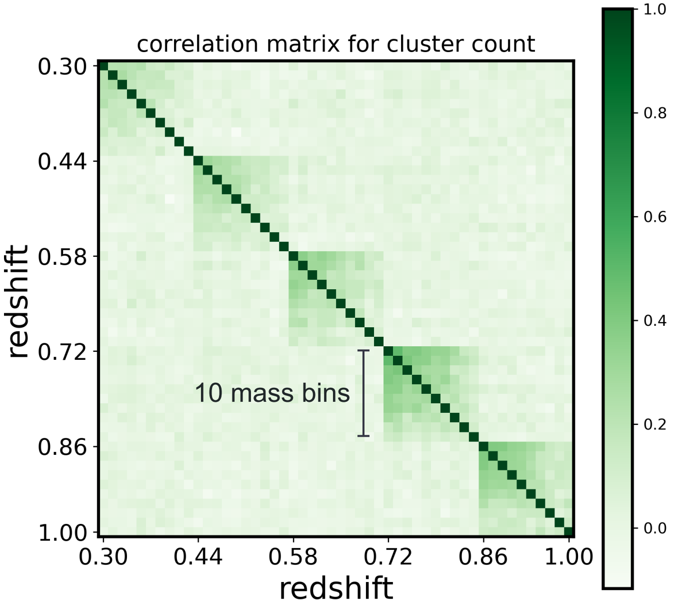
Example: dark matter halo catalog from simulation

- Mass-redshift catalog
- Binning the catalog in several mass-redshift bins
- Estimation of the covariance matrix

$$\sigma^2(N) = N + \sigma^2_{\text{sample}}(N)$$

Diagonal element of the covariance matrix

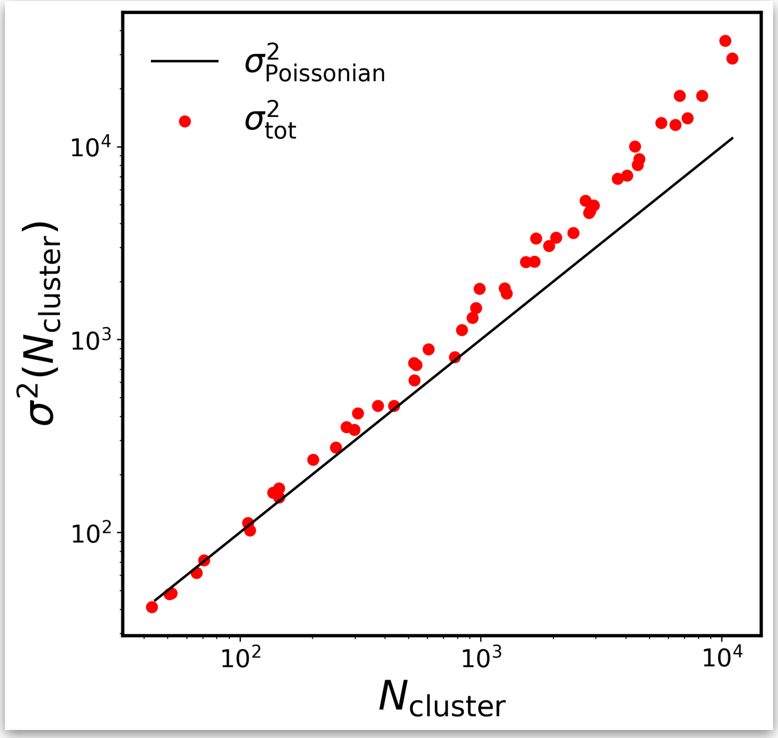




- Deviation from sample noise when N/bin is large
- Off-diagonal terms in correlation matrix

Likelihoods: Poisson vs Gaussian case

Likelihood : links statistical properties of the observables to the data



Likelihoods: Poisson vs Gaussian case

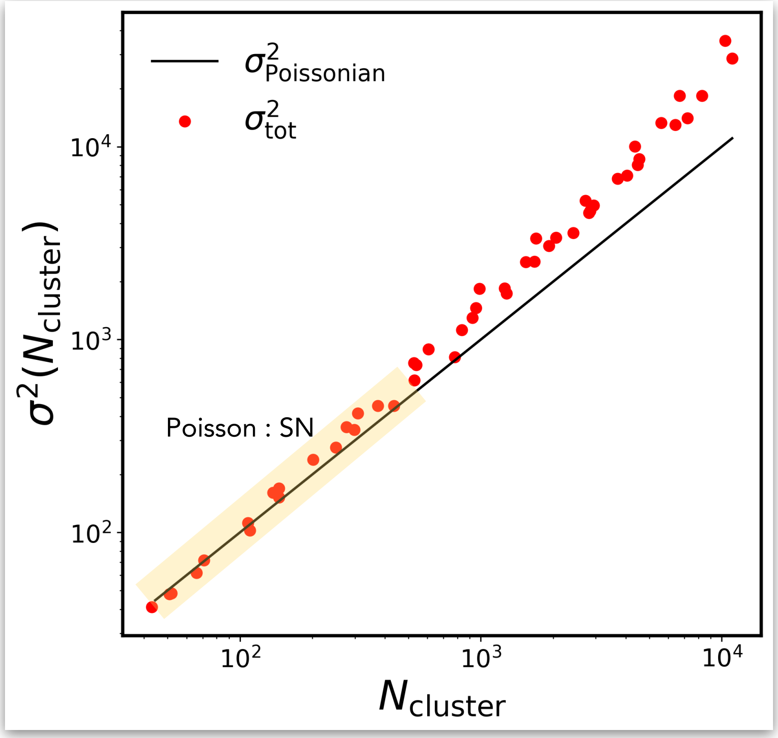
Likelihood : links statistical properties of the observables to the data

Poisson case (example : SZ clusters, Planck 2015)

- When the shot noise is dominant, Poisson counting

$$P(\widehat{N} \mid \vec{\theta}) = \frac{N(\vec{\theta})^{\widehat{N}} \, e^{-N(\vec{\theta})}}{\widehat{N}\,!}$$

- Binned Approach : Count clusters in M-z bins
- Pros : Unbinned approach → Consider clusters at given M, z
- Use more information!
- Cons : Neglect sample variance



Likelihoods: Poisson vs Gaussian case

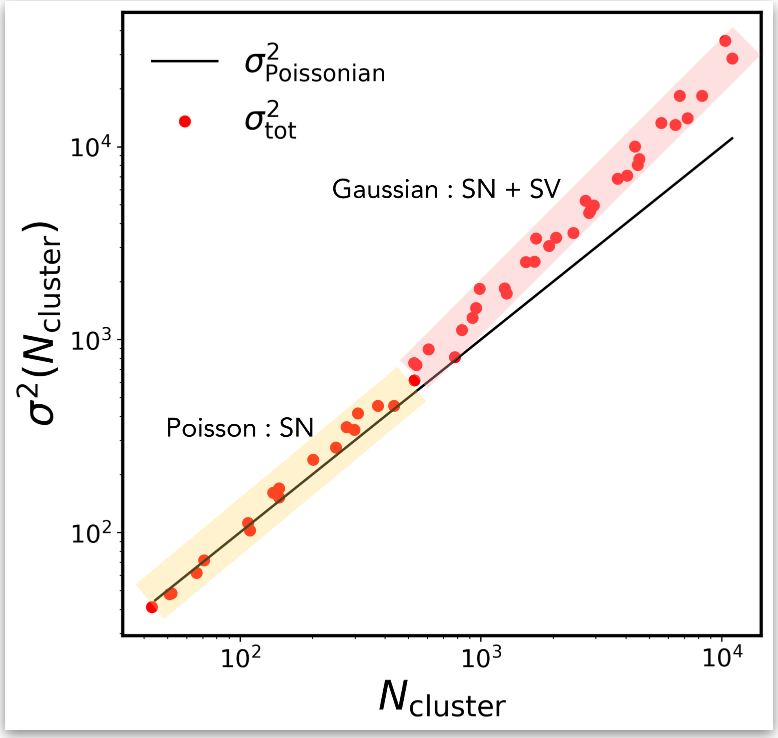
Likelihood : links statistical properties of the observables to the data

Poisson case (example : SZ clusters, Planck 2015)

- When the shot noise is dominant, Poisson counting

$$P(\widehat{N} \mid \vec{\theta}) = \frac{N(\vec{\theta})^{\widehat{N}} e^{-N(\vec{\theta})}}{\widehat{N}\,!}$$

- Binned Approach : Count clusters in M-z bins
- Pros : Unbinned approach → Consider clusters at given M, z
- Use more information!
- Cons : Neglect sample variance



Gaussian case (example : optical clusters, DES 2020)

- Sample variance is not negligible
- Gaussian approximation of the Poisson case (N >> 1)

$$P(\overrightarrow{\widehat{N}} \mid \theta) \propto \exp -\frac{1}{2}[\overrightarrow{\widehat{N} - N(\vec{\theta})}]^T \Sigma^{-1} [\overrightarrow{\widehat{N} - N(\vec{\theta})}]$$

- Pros : include sample variance
- Cons : only binned approach

Likelihoods: Poisson vs Gaussian case

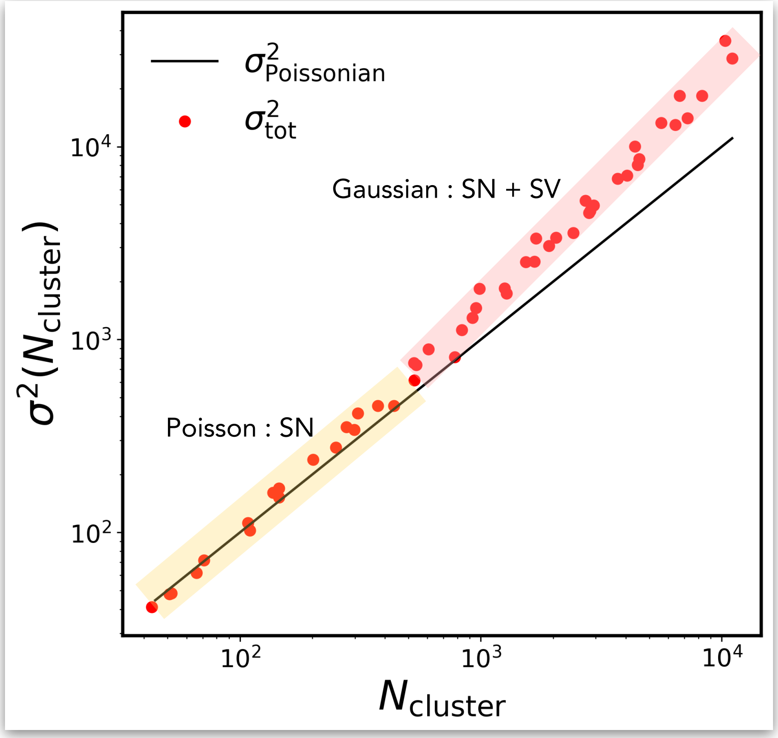
Likelihood : links statistical properties of the observables to the data

Poisson case (example : SZ clusters, Planck 2015)

- When the shot noise is dominant, Poisson counting

$$P(\widehat{N} \mid \vec{\theta}) = \frac{N(\vec{\theta})^{\widehat{N}}\, e^{-N(\vec{\theta})}}{\widehat{N}\,!}$$

- Binned Approach : Count clusters in M-z bins
- Pros : Unbinned approach → Consider clusters at given M, z
- Use more information!
- Cons : Neglect sample variance



Gaussian case (example : optical clusters, DES 2020)

- Sample variance is not negligible
- Gaussian approximation of the Poisson case (N >> 1)

$$P(\overrightarrow{\widehat{N}} \mid \theta) \propto \exp -\frac{1}{2}[\overrightarrow{\widehat{N} - N(\vec{\theta})}]^T \Sigma^{-1} [\overrightarrow{\widehat{N} - N(\vec{\theta})}]$$

- Pros : include sample variance
- Cons : only binned approach

GOAL : derive likelihood that satisfies the unbinned approach and includes the sample variance

Likelihoods: Multivariate Poisson Gaussian case (MPG)

Likelihoods: Multivariate Poisson Gaussian case (MPG)

The cluster count is the result of a Gaussian realisation followed by a Poisson realisation:

Likelihoods: Multivariate Poisson Gaussian case (MPG)

The cluster count is the result of a Gaussian realisation followed by a Poisson realisation:

1. The cluster abundance $\widehat{N}$ is a Poisson realisation of an underlying mean x
2. The local cluster count x is a realisation of a Gaussian variable with mean $N(\Omega_m, \sigma_8)$

Likelihoods: Multivariate Poisson Gaussian case (MPG)

The cluster count is the result of a Gaussian realisation followed by a Poisson realisation:

1. The cluster abundance $\widehat{N}$ is a Poisson realisation of an underlying mean x
2. The local cluster count x is a realisation of a Gaussian variable with mean $N(\Omega_m, \sigma_8)$

$$P_{\mathrm{MPG}}(\{\widehat{N}_i\}_{1\leq i\leq n} | \theta) = \int dx^n \underbrace{\mathcal{G}[\overrightarrow{x} \,|\, \overrightarrow{N}(\theta), \Sigma_{\mathrm{sample}}]}_{\text{1. Gaussian}} \times \underbrace{\prod_{k=1}^{n} \mathcal{P}[\widehat{N}_k | x_k]}_{\text{2. Poisson}}$$

*J. Aitchison & C. H. Ho (1989)

Likelihoods: Multivariate Poisson Gaussian case (MPG)

The cluster count is the result of a Gaussian realisation followed by a Poisson realisation:

1. The cluster abundance $\widehat{N}$ is a Poisson realisation of an underlying mean x
2. The local cluster count x is a realisation of a Gaussian variable with mean $N(\Omega_m, \sigma_8)$

$$P_{\mathrm{MPG}}(\{\widehat{N}_i\}_{1\leq i\leq n}|\theta) = \int dx^n \underbrace{\mathcal{G}[\vec{x}\,|\,\vec{N}(\theta), \Sigma_{\mathrm{sample}}]}_{\text{1. Gaussian}} \times \underbrace{\prod_{k=1}^{n} \mathcal{P}[\widehat{N}_k|x_k]}_{\text{2. Poisson}}$$

*J. Aitchison & C. H. Ho (1989)

Single-variate distribution : "analytical" form

$$P_{\mathrm{SPG}}(\widehat{N}\,|\theta) = \frac{K(N, \sigma^2_{\mathrm{sample}})}{\hat{N}!\sqrt{2\pi\sigma^2_{\mathrm{sample}}}} e^{-\frac{\mu^2}{2\sigma^2_{\mathrm{sample}}}} \frac{1}{2} c^{-\frac{a}{2}-1} \left(\sqrt{c}\,\Gamma\left(\frac{a+1}{2}\right)\, {}_1F_1\left(\frac{a+1}{2}; \frac{1}{2}; \frac{b^2}{4c}\right) - b\Gamma\left(\frac{a}{2}+1\right)\, {}_1F_1\left(\frac{a}{2}+1; \frac{3}{2}; \frac{b^2}{4c}\right)\right)$$

with:

$$\begin{cases} a = \hat{N} \\ b = 1 - \dfrac{N}{\sigma^2_{\mathrm{sample}}} \\ c = \dfrac{1}{2\sigma^2_{\mathrm{sample}}} \end{cases}$$

Likelihoods: Multivariate Poisson Gaussian case (MPG)

The cluster count is the result of a Gaussian realisation followed by a Poisson realisation:

1. The cluster abundance $\widehat{N}$ is a Poisson realisation of an underlying mean x
2. The local cluster count x is a realisation of a Gaussian variable with mean $N(\Omega_m, \sigma_8)$

$$P_{\mathrm{MPG}}(\{\widehat{N}_i\}_{1\leq i\leq n}|\theta) = \int dx^n \underbrace{\mathcal{G}[\overrightarrow{x}\,|\,\overrightarrow{N}(\theta), \Sigma_{\mathrm{sample}}]}_{\text{1. Gaussian}} \times \underbrace{\prod_{k=1}^{n} \mathcal{P}[\widehat{N}_k|x_k]}_{\text{2. Poisson}}$$

*J. Aitchison & C. H. Ho (1989)

Single-variate distribution : "analytical" form

$$P_{\mathrm{SPG}}(\widehat{N}|\theta) = \frac{K(N,\sigma^2_{\mathrm{sample}})}{\hat{N}!\sqrt{2\pi\sigma^2_{\mathrm{sample}}}} e^{-\frac{\mu^2}{2\sigma^2_{\mathrm{sample}}}} \frac{1}{2} c^{-\frac{a}{2}-1} \left(\sqrt{c}\,\Gamma\left(\frac{a+1}{2}\right) {}_1F_1\left(\frac{a+1}{2};\frac{1}{2};\frac{b^2}{4c}\right) - b\Gamma\left(\frac{a}{2}+1\right) {}_1F_1\left(\frac{a}{2}+1;\frac{3}{2};\frac{b^2}{4c}\right)\right) \text{ with: } \begin{cases} a = \hat{N} \\ b = 1 - \dfrac{N}{\sigma^2_{\mathrm{sample}}} \\ c = \dfrac{1}{2\sigma^2_{\mathrm{sample}}} \end{cases}$$

Pros (Both Poisson and Gaussian advantages):

- Effect of sample variance
- Can be used in a binned and un-binned framework

Likelihoods: Multivariate Poisson Gaussian case (MPG)

The cluster count is the result of a Gaussian realisation followed by a Poisson realisation:

1. The cluster abundance $\widehat{N}$ is a Poisson realisation of an underlying mean x
2. The local cluster count x is a realisation of a Gaussian variable with mean $N(\Omega_m, \sigma_8)$

$$P_{\mathrm{MPG}}(\{\widehat{N}_i\}_{1\leq i\leq n} | \theta) = \int dx^n \, \underbrace{\mathscr{G}[\vec{x} \,|\, \vec{N}(\theta), \boxed{\Sigma_{\mathrm{sample}}}]}_{\text{1. Gaussian}} \times \underbrace{\prod_{k=1}^{n} \mathscr{P}[\widehat{N}_k | x_k]}_{\text{2. Poisson}}$$

*J. Aitchison & C. H. Ho (1989)

Single-variate distribution : "analytical" form

$$P_{\mathrm{SPG}}(\widehat{N} | \theta) = \frac{K(N, \sigma^2_{\mathrm{sample}})}{\hat{N}!\sqrt{2\pi\sigma^2_{\mathrm{sample}}}} e^{-\frac{\mu^2}{2\sigma^2_{\mathrm{sample}}}} \frac{1}{2} c^{-\frac{a}{2}-1} \left(\sqrt{c}\,\Gamma\left(\frac{a+1}{2}\right) {}_1F_1\left(\frac{a+1}{2}; \frac{1}{2}; \frac{b^2}{4c}\right) - b\Gamma\left(\frac{a}{2}+1\right) {}_1F_1\left(\frac{a}{2}+1; \frac{3}{2}; \frac{b^2}{4c}\right)\right) \text{ with: } \begin{cases} a = \hat{N} \\ b = 1 - \frac{N}{\sigma^2_{\mathrm{sample}}} \\ c = \frac{1}{2\sigma^2_{\mathrm{sample}}} \end{cases}$$

Pros (Both Poisson and Gaussian advantages):

- Effect of sample variance
- Can be used in a binned and un-binned framework

Sample variance is negligible → Poisson case

Likelihoods: Multivariate Poisson Gaussian case (MPG)

The cluster count is the result of a Gaussian realisation followed by a Poisson realisation:

1. The cluster abundance $\widehat{N}$ is a Poisson realisation of an underlying mean x
2. The local cluster count x is a realisation of a Gaussian variable with mean $N(\Omega_m, \sigma_8)$

$$P_{\mathrm{MPG}}(\{\widehat{N}_i\}_{1\leq i\leq n}|\theta) = \int dx^n \underbrace{\mathscr{G}[\vec{x}\,|\,\vec{N}(\theta), \boxed{\Sigma_{\mathrm{sample}}}]}_{\text{1. Gaussian}} \times \underbrace{\prod_{k=1}^{n} \mathscr{P}[\widehat{N}_k|x_k]}_{\text{2. Poisson}}$$

*J. Aitchison & C. H. Ho (1989)

Single-variate distribution : "analytical" form

$$P_{\mathrm{SPG}}(\widehat{N}\,|\theta) = \frac{K(N,\sigma^2_{\mathrm{sample}})}{\hat{N}!\sqrt{2\pi\sigma^2_{\mathrm{sample}}}} e^{-\frac{\mu^2}{2\sigma^2_{\mathrm{sample}}}} \frac{1}{2} c^{-\frac{a}{2}-1} \left(\sqrt{c}\,\Gamma\left(\frac{a+1}{2}\right) {}_1F_1\left(\frac{a+1}{2};\frac{1}{2};\frac{b^2}{4c}\right) - b\Gamma\left(\frac{a}{2}+1\right) {}_1F_1\left(\frac{a}{2}+1;\frac{3}{2};\frac{b^2}{4c}\right)\right) \text{ with: } \begin{cases} a = \hat{N} \\ b = 1 - \dfrac{N}{\sigma^2_{\mathrm{sample}}} \\ c = \dfrac{1}{2\sigma^2_{\mathrm{sample}}} \end{cases}$$

Pros (Both Poisson and Gaussian advantages):

- Effect of sample variance
- Can be used in a binned and un-binned framework

Sample variance is negligible → Poisson case

N>>1 and sample variance is no more negligible → Gaussian case

Testing likelihood accuracy

- Problem: We can, using a given likelihood (Poissonian, Gaussian, MPG) estimate our cosmological parameters ($\boldsymbol{\Omega}_m$,$\boldsymbol{\sigma}_8$) and their errors (posterior distribution) but how do we know the errors are correct?

- Answer: We estimate ($\boldsymbol{\Omega}_m$,$\boldsymbol{\sigma}_8$) with many different realisation of the universe, and look at the distributions
  - This tests the accuracy of the errors
  - Tests the bias of our likelihood estimation

Testing likelihood accuracy: 1000 simulations

1000 simulations, P. Monaco et al., 2002 (Fumagalli et al. Euclid collaboration):

- Euclid-like light-cone (¼ sphere), can be used for Rubin survey
- $V = (3800\ \mathrm{Mpc})^3$ $\sim 10^5$ halos/simulation
- $z < 2.5$ and $M_{halo} > 2.43.10^{14}\ M_{sun}$

Testing likelihood accuracy: 1000 simulations

1000 simulations, P. Monaco et al., 2002 (Fumagalli et al. Euclid collaboration):

- Euclid-like light-cone (¼ sphere), can be used for Rubin survey
- V = (3800 Mpc)^3 ~10^5 halos/simulation
- z < 2.5 and M_{halo} > 2.43.10^{14} M_{sun}

- **Low abundance and high abundance M-z bins**
- → Histogram of observed abundance over the 1000 simulations

Low abundance bin $\langle N \rangle \approx 2$

$$0.2 < z < 0.25$$
$$14.5 < \log_{10} M < 14.501$$

High abundance bin $\langle N \rangle \approx 2500$

$$0.2 < z < 0.25$$
$$14.16 < \log_{10} M < 14.5$$

Testing likelihood accuracy: 1000 simulations

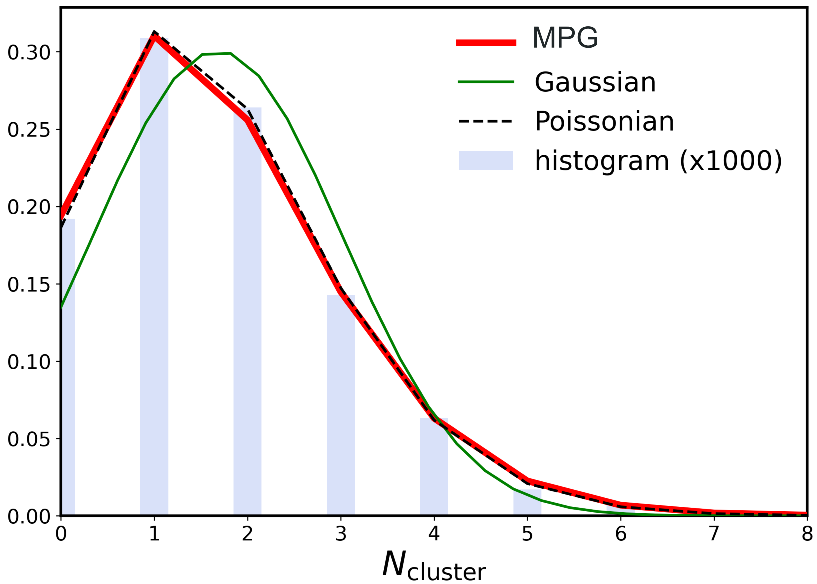
1000 simulations, P. Monaco et al., 2002 (Fumagalli et al. Euclid collaboration):

- Euclid-like light-cone (¼ sphere), can be used for Rubin survey
- V = (3800 Mpc)^3 ~10^5 halos/simulation
- z < 2.5 and M_{halo} > 2.43.10^{14} M_{sun}

- **Low abundance and high abundance M-z bins**
- → Histogram of observed abundance over the 1000 simulations

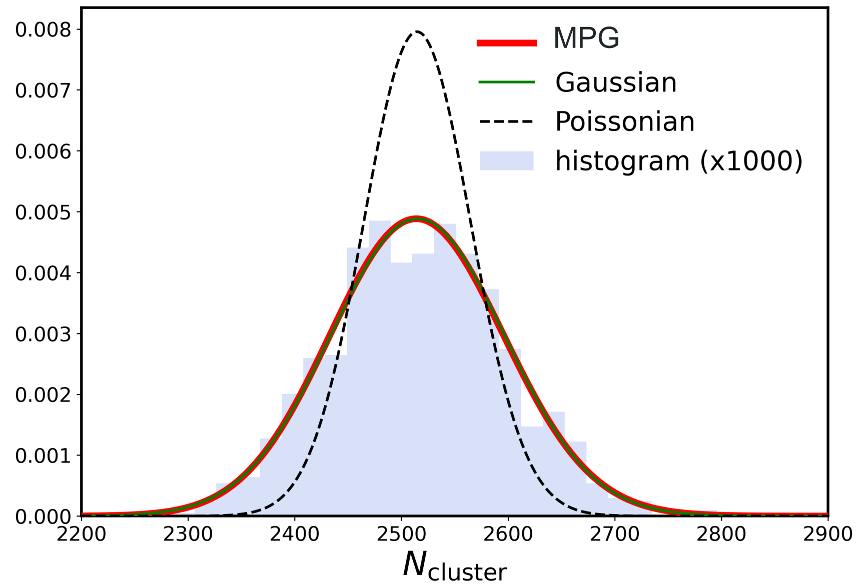
Low abundance bin $\langle N \rangle \approx 2$

$$0.2 < z < 0.25$$
$$14.5 < \log_{10} M < 14.501$$



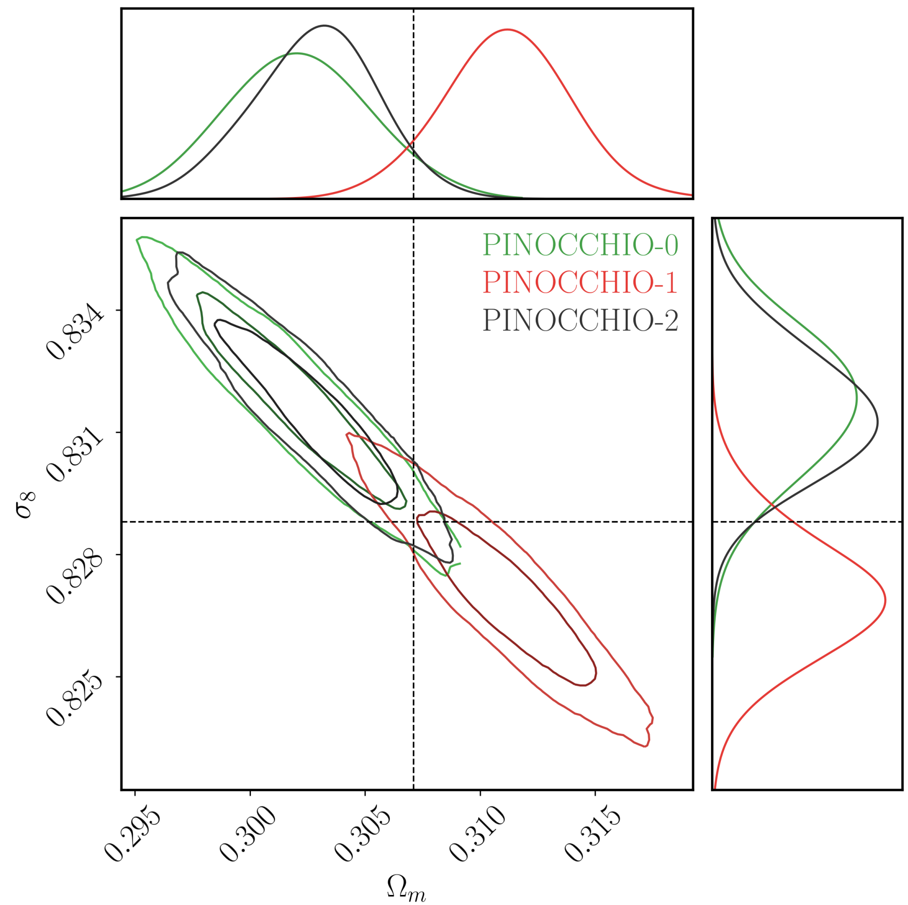
High abundance bin $\langle N \rangle \approx 2500$

$$0.2 < z < 0.25$$
$$14.16 < \log_{10} M < 14.5$$



Preliminary results: Binned Gaussian likelihood

- For each simulation, access the posterior for ($\boldsymbol{\Omega}_m$,$\boldsymbol{\sigma}_8$)



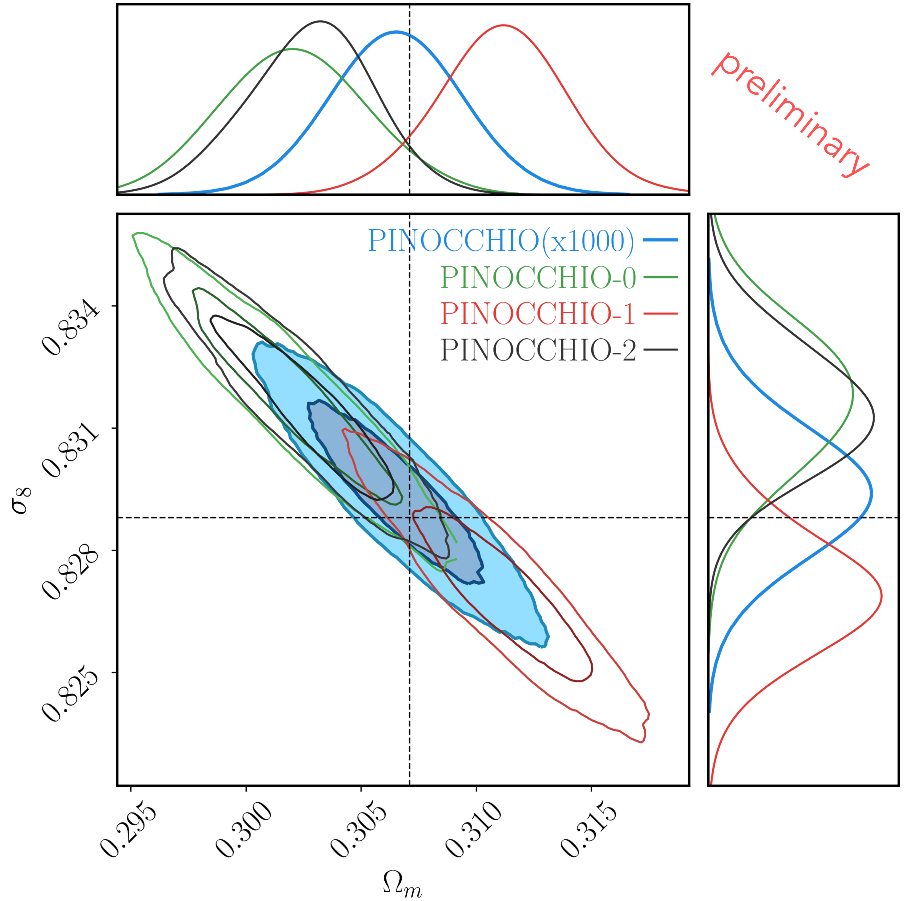
Un-filled contours : Posterior distributions contours for ($\boldsymbol{\Omega}_m$,$\boldsymbol{\sigma}_8$)

Preliminary results: Binned Gaussian likelihood

- For each simulation, access the posterior for $(\boldsymbol{\Omega}_m, \boldsymbol{\sigma}_8)$

Repeat 1000 times over the 1000 simulations:

- Estimates of $(\boldsymbol{\Omega}_m, \boldsymbol{\sigma}_8)$
- Look at the distribution



Un-filled contours : Posterior distributions contours for $(\boldsymbol{\Omega}_m, \boldsymbol{\sigma}_8)$

Filled contour : Histogram of the 1000 $(\boldsymbol{\Omega}_m, \boldsymbol{\sigma}_8)$ individual best fits

Preliminary results: Binned Gaussian likelihood

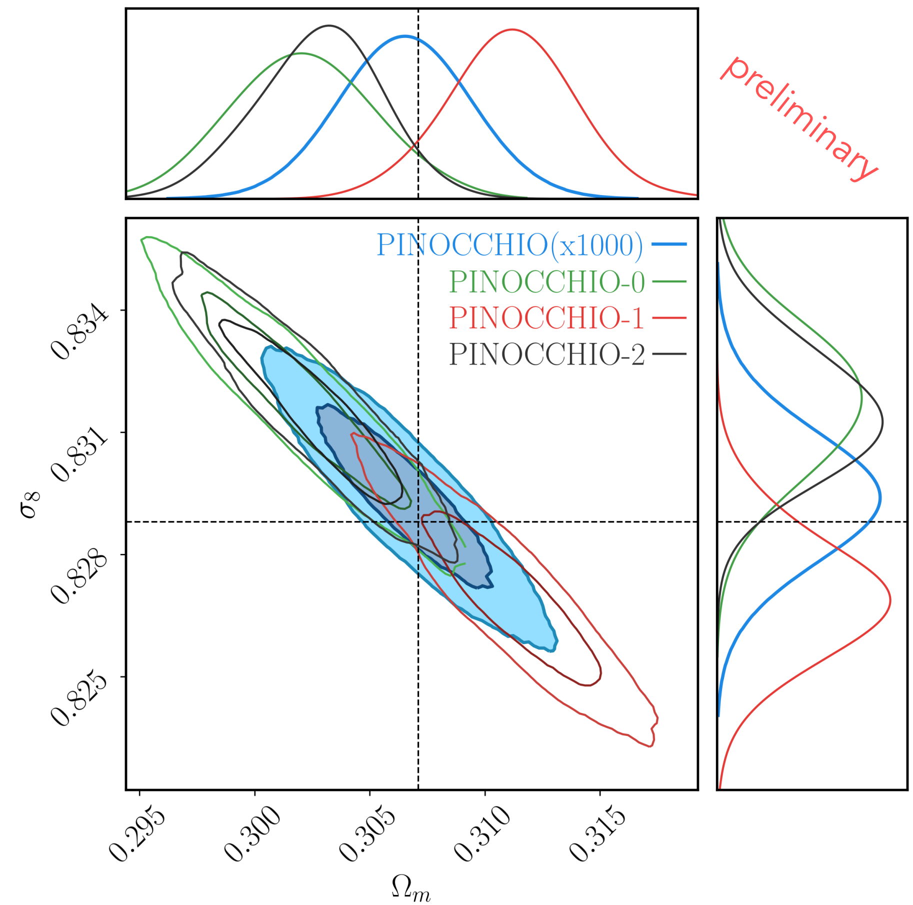
- For each simulation, access the posterior for ($\boldsymbol{\Omega}_m$,$\boldsymbol{\sigma}_8$)

Repeat 1000 times over the 1000 simulations:

- Estimates of ($\boldsymbol{\Omega}_m$,$\boldsymbol{\sigma}_8$)
- Look at the distribution

Perspectives :

- Test the accuracy of the individual errors
- Tests the bias of our likelihood estimation
- Do the same with
    1. Poisson binned/un-binned
    2. MPG binned/un-binned
- Compare likelihood accuracies



Un-filled contours : Posterior distributions contours for ($\boldsymbol{\Omega}_m$,$\boldsymbol{\sigma}_8$)

Filled contour : Histogram of the 1000 ($\boldsymbol{\Omega}_m$,$\boldsymbol{\sigma}_8$) individual best fits

Conclusions

- Cluster count : specific count experiment under matter density fluctuations
- LSST, Euclid Survey : $o(10^5)$ detected clusters, sample variance no more negligible
- Improvement of cluster likelihoods must be considered (unbinned approach including sample variance)
- We propose to use a new Poissonian/Gaussian mixture likelihood combining both advantages
- 1000 simulations → high statistics to determine likelihood accuracies

Thank you for your attention!

Cluster count - covariance matrix

- Consider a volume V
- A mean number density n with local estimate $\hat{n}(\vec{x}) = \sum_i \delta^D(\vec{x} - \vec{x}_i)$
- The instantaneous count of particles within this volume is given by

Window function (1 in V, 0 elsewhere)

$$N_V = \int_V d^3\vec{x} \sum_i \delta^D(\vec{x} - \vec{x}_i) = \int W_V(\vec{x}) d^3\vec{x} \sum_i \delta^D(\vec{x} - \vec{x}_i)$$

$$\mathrm{Cov}(N_{\alpha_1}, N_{\alpha_2}) = N_{\alpha_1}\delta_K^{\alpha_1,\alpha_2} + \bar{n}^2 \int \frac{d^3k}{(2\pi)^3} P_{\mathrm{hh}}(\vec{k}) W^*_{\alpha_1}(\vec{k}) W_{\alpha_2}(\vec{k}) = \Sigma_{\mathrm{Poissonian}} + \Sigma_{\mathrm{sample}}$$

Power spectrum

$$W_V(\vec{k}) = \int d^3x W_V(\vec{x}) e^{-i\vec{k}\cdot\vec{x}}$$

2 integrals

$$N_{\alpha\beta}^{\text{predicted}} = \Omega_s \int_{z_1}^{z_2} dz \int_{\log_{10} M_1}^{\log_{10} M_2} d\,\log_{10} M \,\frac{dn(M,z)}{d\,\log_{10} M}\frac{d^2V(z)}{dz d\Omega}$$

3 integrals

$$N_{\alpha\beta}^{\text{predicted}} = \Omega_s \int_{z_1}^{z_2} dz \int_{\lambda_1}^{\lambda_2} d\lambda \int_{\log_{10} M_{min}}^{\log_{10} M_{max}} d\,\log_{10} M \,\frac{dn(M,z)}{d\,\log_{10} M}\frac{d^2V(z)}{dz d\Omega} P_{M-\lambda}(\lambda\,|\,M,z)$$

3 integrals + selection function

$$N_{\alpha\beta}^{\text{predicted}} = \Omega_s \int_{z_1}^{z_2} dz \int_{\lambda_1}^{\lambda_2} d\lambda \int_{\log_{10} M_{min}}^{\log_{10} M_{max}} d\,\log_{10} M \,\frac{dn(M,z)}{d\,\log_{10} M}\frac{d^2V(z)}{dz d\Omega} S(M,z) P_{M-\lambda}(\lambda\,|\,M,z)$$

3 integrals + selection function + uncertainties on cluster redshifts and proxies

$$N_{\alpha\beta}^{\text{predicted}} = \Omega_s \int_{z_1}^{z_2} dz \int_{\lambda_{obs,1}}^{\lambda_{obs,2}} d\lambda_{obs} \int_{\lambda_1}^{\lambda_2} d\lambda \int_{\log_{10} M_{min}}^{\log_{10} M_{max}} d\,\log_{10} M \,\frac{dn(M,z)}{d\,\log_{10} M}\frac{d^2V(z)}{dz d\Omega} P_{M-\lambda}(\lambda\,|\,M,z) P(\lambda_{obs}\,|\,\lambda)$$