Some insights into the LSS likelihood from N-body simulations?

F. Elsner, F. Schmidt, M. Nguyen, J. Jasche

08/11/17

Recent deveolopments

Since our last meeting, we have implemented into BORG

- the LSS bias model to second order
- the negative binomial likelihood

Analyzing N-body simulations, we failed to reproduce bias parameter measurements obtained from the 2-point statistics.[1]

To do tests much more quickly, we developed a simplified MLE outside of BORG.

[1] Now understood as not necessarily being equivalent

A simplified maximum likelihood estimator

To quickly test different bias models and likelihoods, we compute maximum likelihood estimates for a fixed density field,

$$P(\{b_i\},\, \delta_m \,|\, \delta_h^{\mathrm{N-body}})$$

$$\downarrow$$

$$P(\{b_i\} \,|\, \delta_h^{\mathrm{N-body}},\, \delta_m^{\mathrm{N-body}})$$

$$\downarrow$$

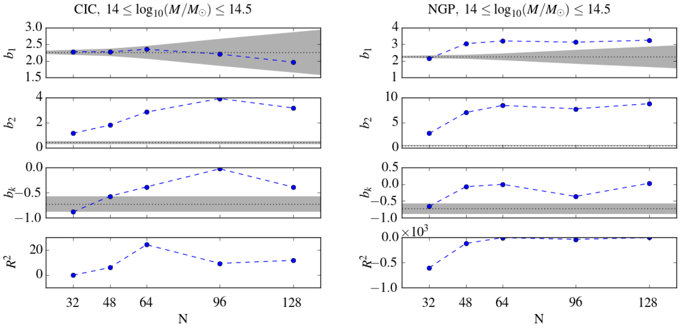
$$\underset{\{b_i\}}{\operatorname{argmax}}[P(\{b_i\} \,|\, \delta_h^{\mathrm{N-body}},\, \delta_m^{\mathrm{N-body}})]\,,$$

where we used a halo catalog as input data ($\delta_h^{\mathrm{N-body}}$), computed from a given dark matter density field ($\delta_m^{\mathrm{N-body}}$).

MLE results vs. previous measurements I

Results from > 500 test are quite discouraging, e.g.,

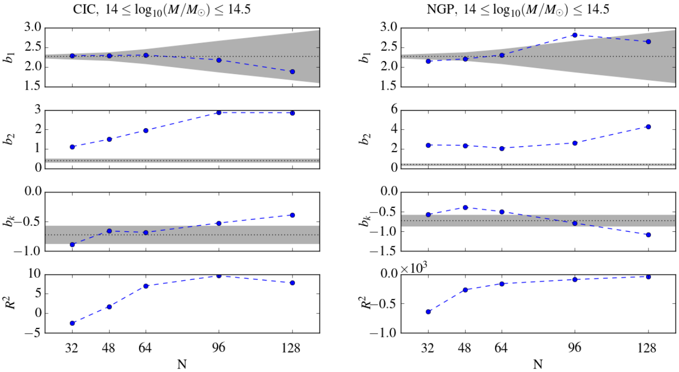
Negative binomial likelihood:



MLE results vs. previous measurements II

Results from > 500 test are quite discouraging, e.g.,

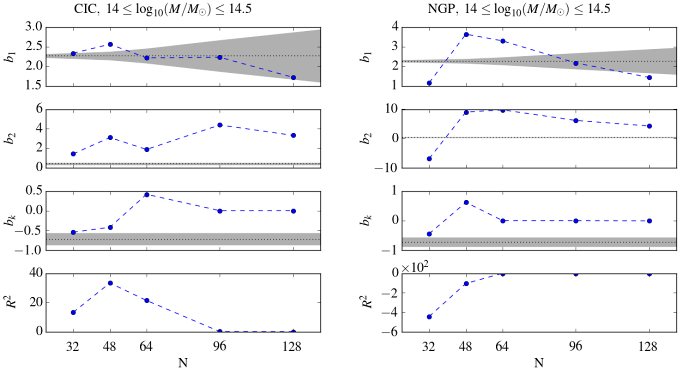
Gaussian likelihood:



MLE results vs. previous measurements III

Results from > 500 test are quite discouraging, e.g.,

Log-normal likelihood:



MLE results summary

We find evidence for

- systematically different results of NGP and CIC data projections
- disagreement in the recovered parameters even at the lowest resolution runs – in particular for b_2
- a noticeable dependence of the MLE on the likelihood used

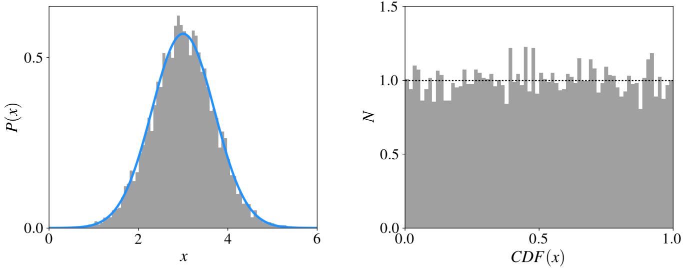
Can we asses the shortcomings of the likelihood quantitatively?

Quantile analysis I

We start by observing that

- we have a large data vector, with many entries
- the likelihood is different in every pixel, dependent on δ_m, ...

For any given likelihood, we can compute the quantile at each grid point. Its distribution should be flat:

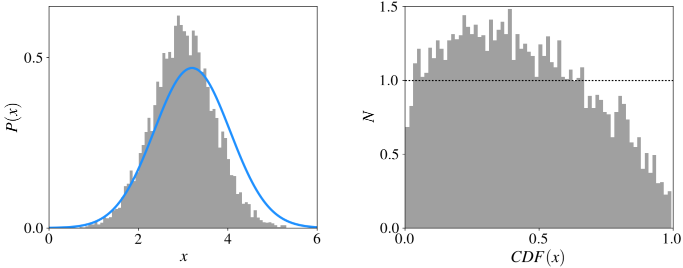


Quantile analysis I

We start by observing that

- we have a large data vector, with many entries
- the likelihood is different in every pixel, dependent on δ_m, ...

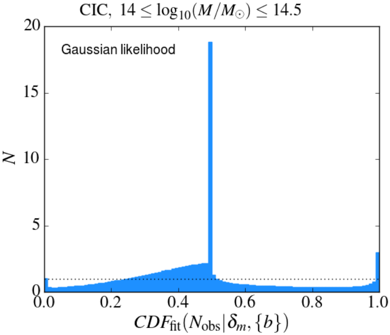
For any given likelihood, we can compute the quantile at each grid point. Its distribution should be flat:



Quantile analysis II

How does the LSS quantile distibution looks like if we fit the data with a simple Gaussian?

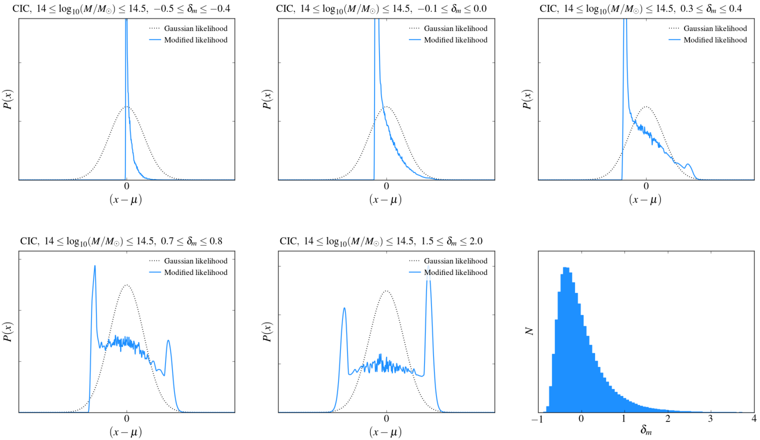
It's not flat at all:



How do we have to modify the shape of the likelihood for the quantiles to become flat?

LSS likelihood reconstruction

Since we expect the likelihood to be a function of matter density, we should do this analysis in bins in δ_m:



Current status

In our tests, we have found that

- typical LSS likelihoods proposed in literature fail to capture "second order" effects
- a quantile analysis can provide us with insights into the actual shape of the likelihood – it is very complex

Important questions remain open:

- How well do we have to reproduce the likelihood shape?
- Can we recover unbiased *cosmological* parameters using more standard analytical likelihoods?