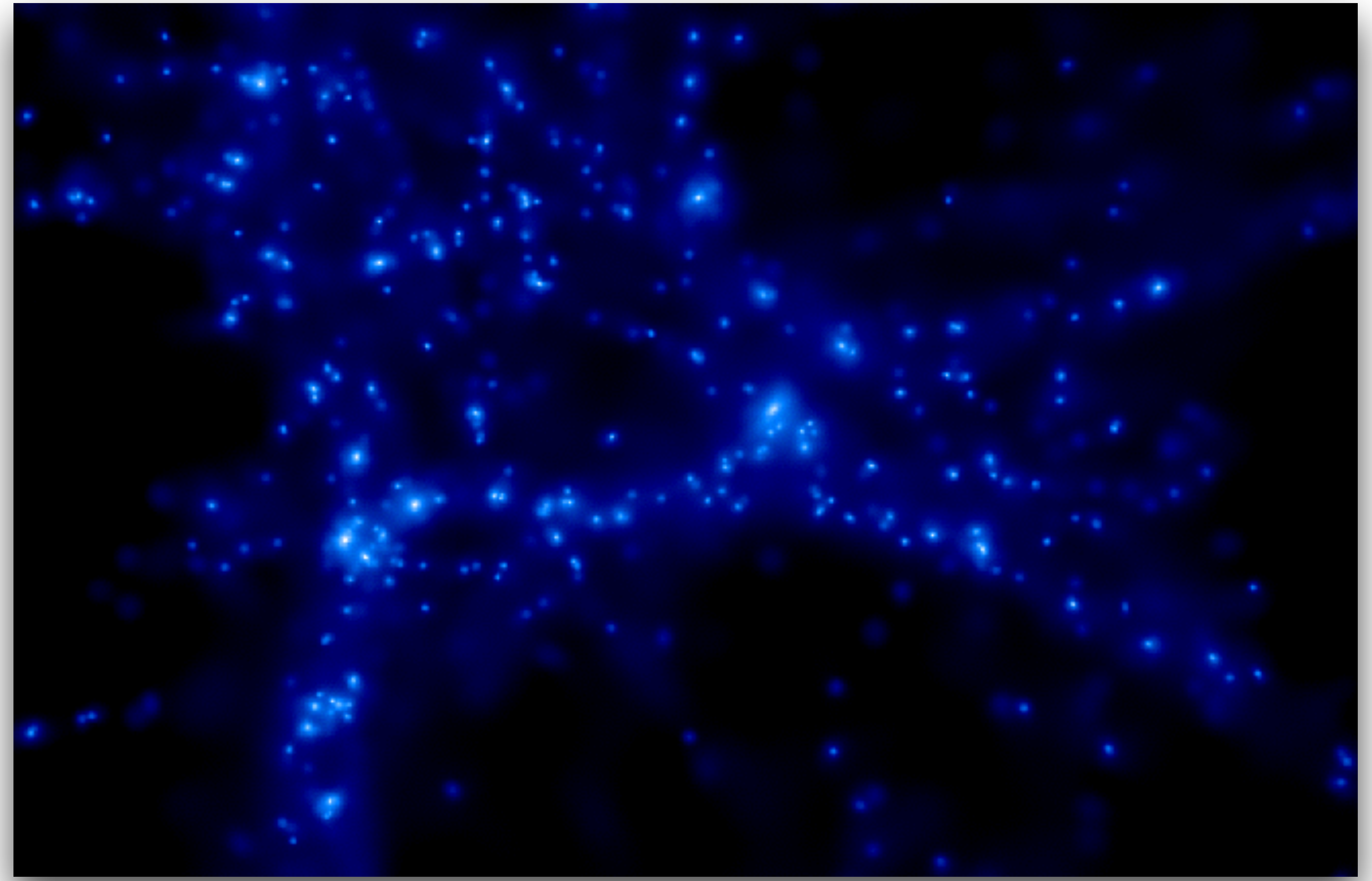
How to measure a foreground mask from a galaxy redshift survey

Pierluigi Monaco, Trieste University, INAF-OATs and INFN
with Enea Di Dio, Emiliano Sefusatti



Paving the way for the next generation of cosmological surveys, Sesto, July 2018

- Foregrounds induce spurious power:

$$1+\delta_{obs} \simeq (1+\varepsilon)\,(1+\delta_{true})$$

$$\langle \delta_{obs,1}\delta_{obs,2}\rangle = \langle \delta_{true,1}\delta_{true,2}\rangle + \langle \varepsilon_1\varepsilon_2\rangle + \langle \delta_{true,1}\delta_{true,2}\rangle\langle \varepsilon_1\varepsilon_2\rangle$$

where $\varepsilon = \delta L/L_0$ is a modulation of the survey depth

- But if the true measure gives $\langle \delta_{true,1}\delta_{true,2}\rangle=0$, then one can measure the foreground from $\langle \delta_{obs,1}\delta_{obs,2}\rangle$

- (Nearly) vanishing correlations are expected in angular cross-correlations of different redshift bins, except for the effect of lensing.

- Foregrounds induce spurious power:

$$1+\delta_{obs} \simeq (1+\varepsilon)\,(1+\delta_{true})$$

$$\langle\delta_{obs,1}\delta_{obs,2}\rangle \simeq \cancel{\langle\delta_{true,1}\delta_{true,2}\rangle} + \langle\varepsilon_1\varepsilon_2\rangle + \cancel{\langle\delta_{true,1}\delta_{true,2}\rangle\langle\varepsilon_1\varepsilon_2\rangle}$$

where $\varepsilon = \delta L/L_0$ is a modulation of the survey depth

- But if the true measure gives $\langle\delta_{true,1}\delta_{true,2}\rangle=0$, then one can measure the foreground from $\langle\delta_{obs,1}\delta_{obs,2}\rangle$

- (Nearly) vanishing correlations are expected in angular cross-correlations of different redshift bins, except for the effect of lensing.

Synergies vs Systematics

$$s^{\rm obs} = s + s^{\rm sys}$$

$$\langle s^{\rm obs} s^{\rm obs} \rangle = \langle ss \rangle + \cancel{2\langle s s^{\rm sys} \rangle} + \langle s^{\rm sys} s^{\rm sys} \rangle$$

$$\langle s^{\rm obs}_{(o)} s^{\rm obs}_{(r)} \rangle = \langle ss \rangle + \cancel{\langle s_{(r)} s^{\rm sys}_{(o)} \rangle} + \cancel{\langle s_{(o)} s^{\rm sys}_{(r)} \rangle} + \cancel{\langle s^{\rm sys}_{(o)} s^{\rm sys}_{(r)} \rangle}$$

- Foregrounds induce spurious power:

$$1+\delta_{\rm obs} \simeq (1+\varepsilon)\,(1+\delta_{\rm true})$$

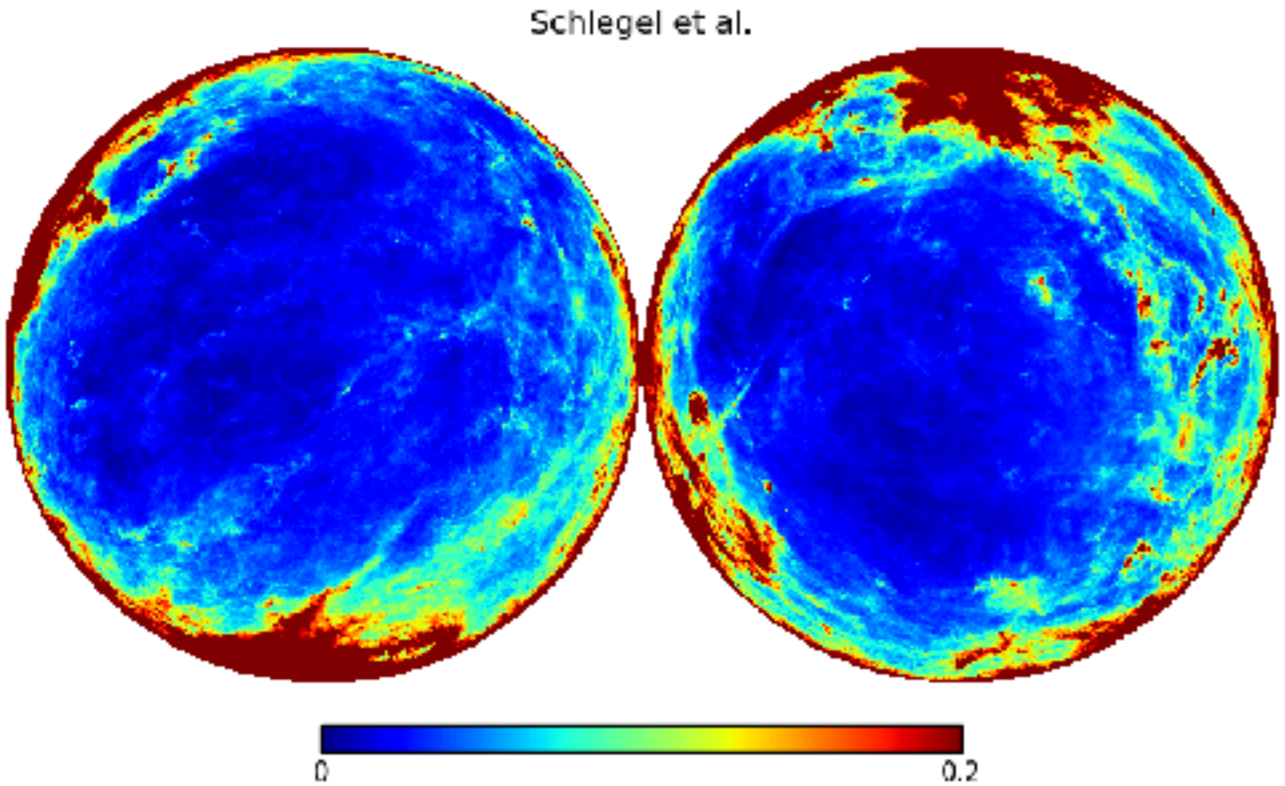
$$\langle \delta_{\rm obs,1}\delta_{\rm obs,2}\rangle \simeq \cancel{\langle \delta_{\rm true,1}\delta_{\rm true,2}\rangle} + \langle \varepsilon_1\varepsilon_2\rangle + \cancel{\langle \delta_{\rm true,1}\delta_{\rm true,2}\rangle\langle \varepsilon_1\varepsilon_2\rangle}$$

where $\varepsilon = \delta L/L_0$ is a modulation of the survey depth

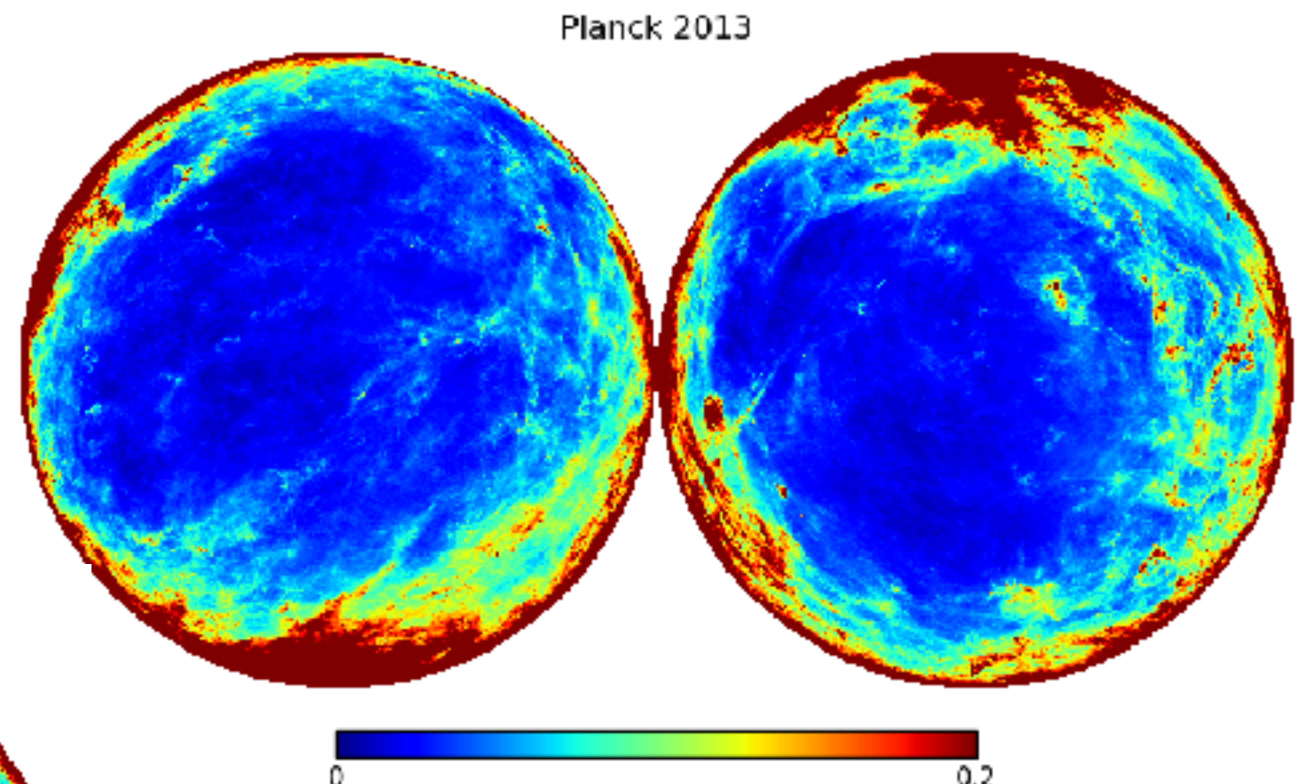
- Contribution to galaxy density cross-correlation:
    - vanishing correlations from large-scale structure
    - foreground contamination
    - gravitational lensing
    - catastrophic redshift errors

Prototypical foreground: Milky Way extinction

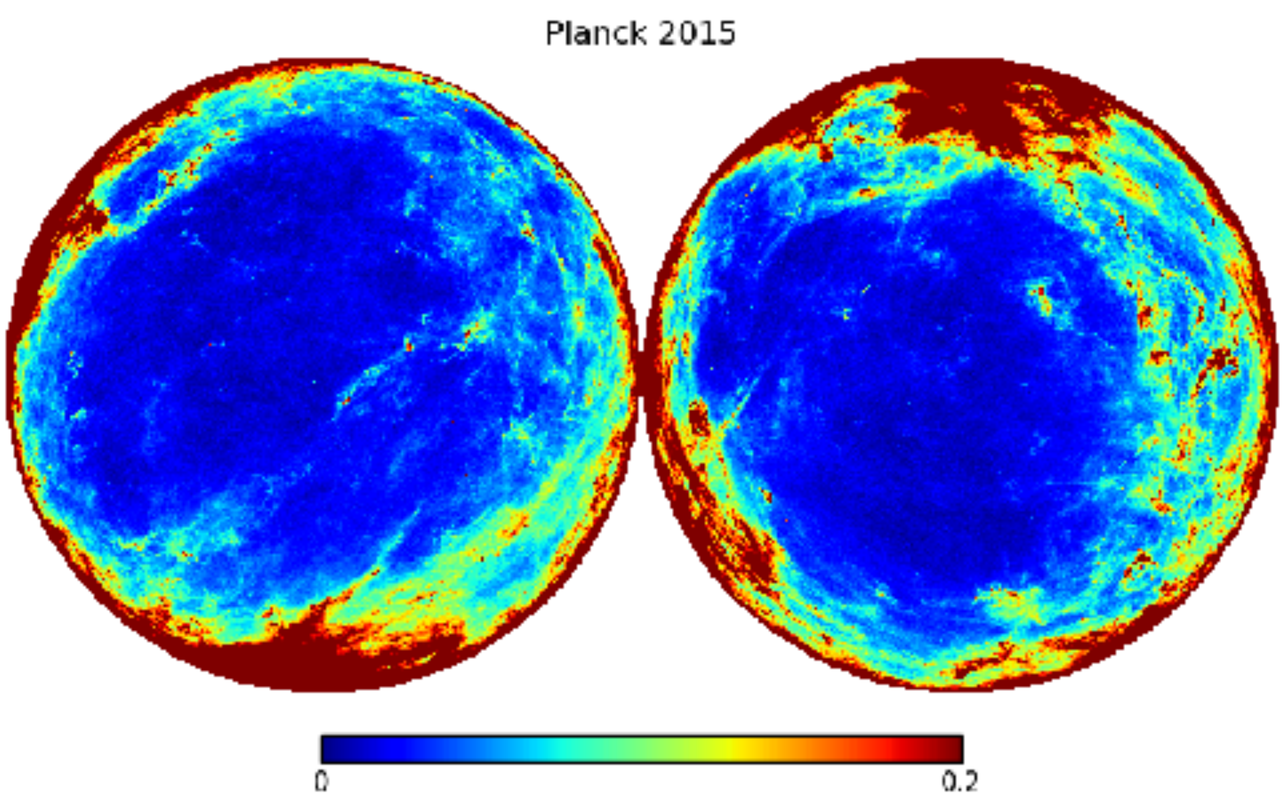
SFD



P13



P15

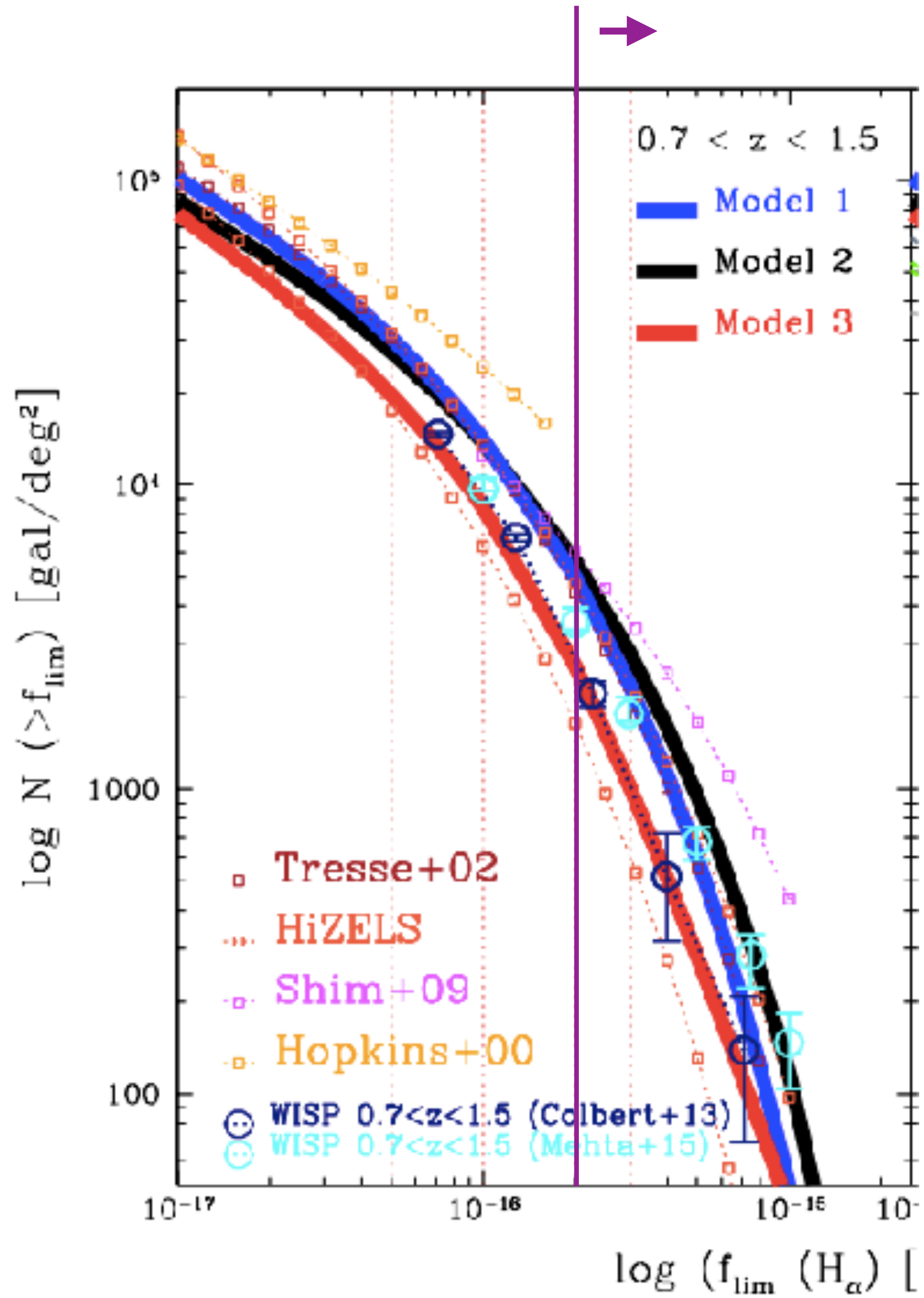


The observed number of galaxies:

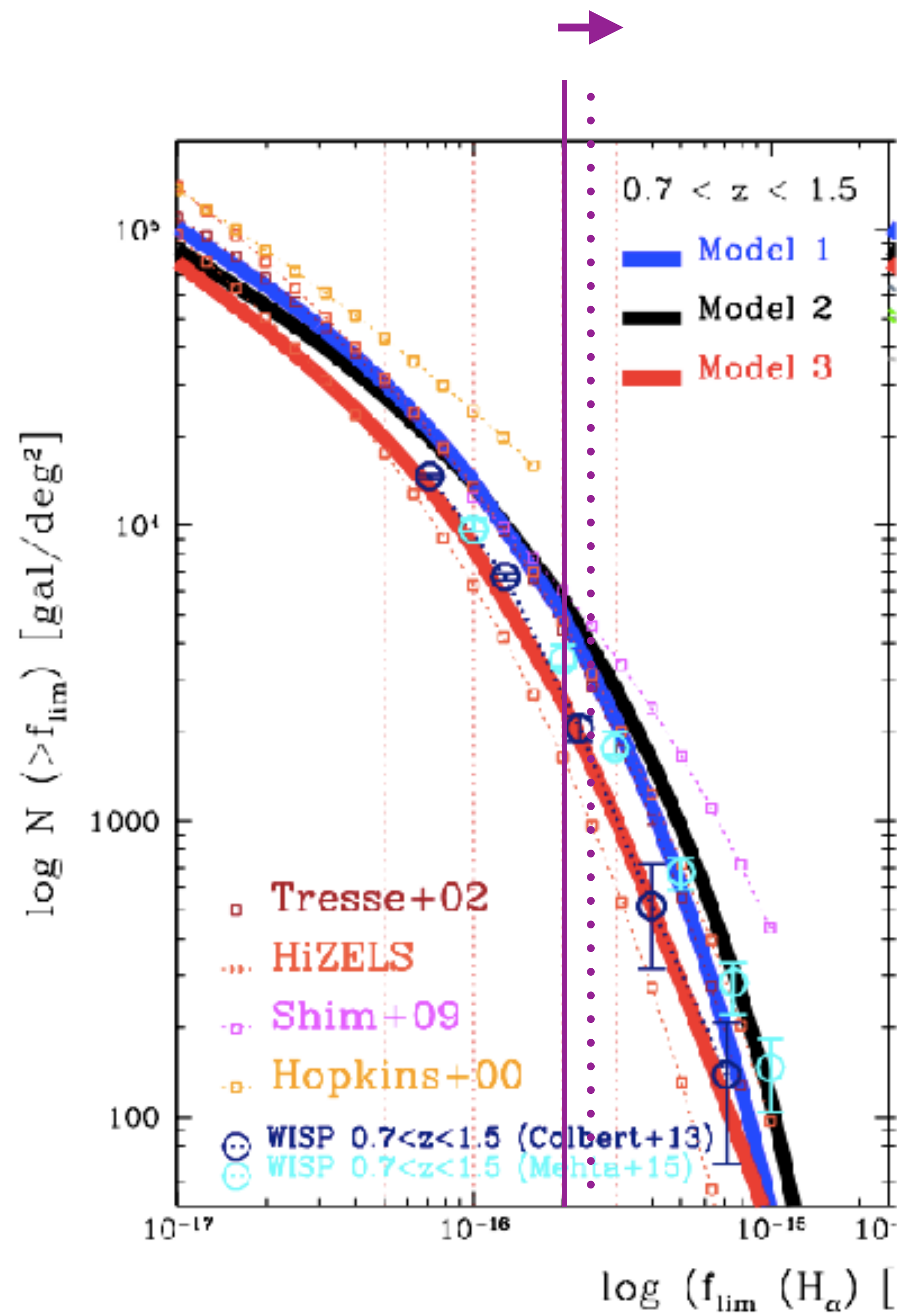
$$n_o(\mathbf{x}) = \int_{L_{\rm lim}}^{\infty} \Phi_{\rm local}(L|\mathbf{x}) dL$$

$$\Phi_{\rm local}(L|\mathbf{x}) = [1 + \delta_g(\mathbf{x})]\Phi(L|z)$$

under the assumption of a universal LF



Pozzetti et al. (2016)



The observed number of galaxies is subject to modulations of survey depth.

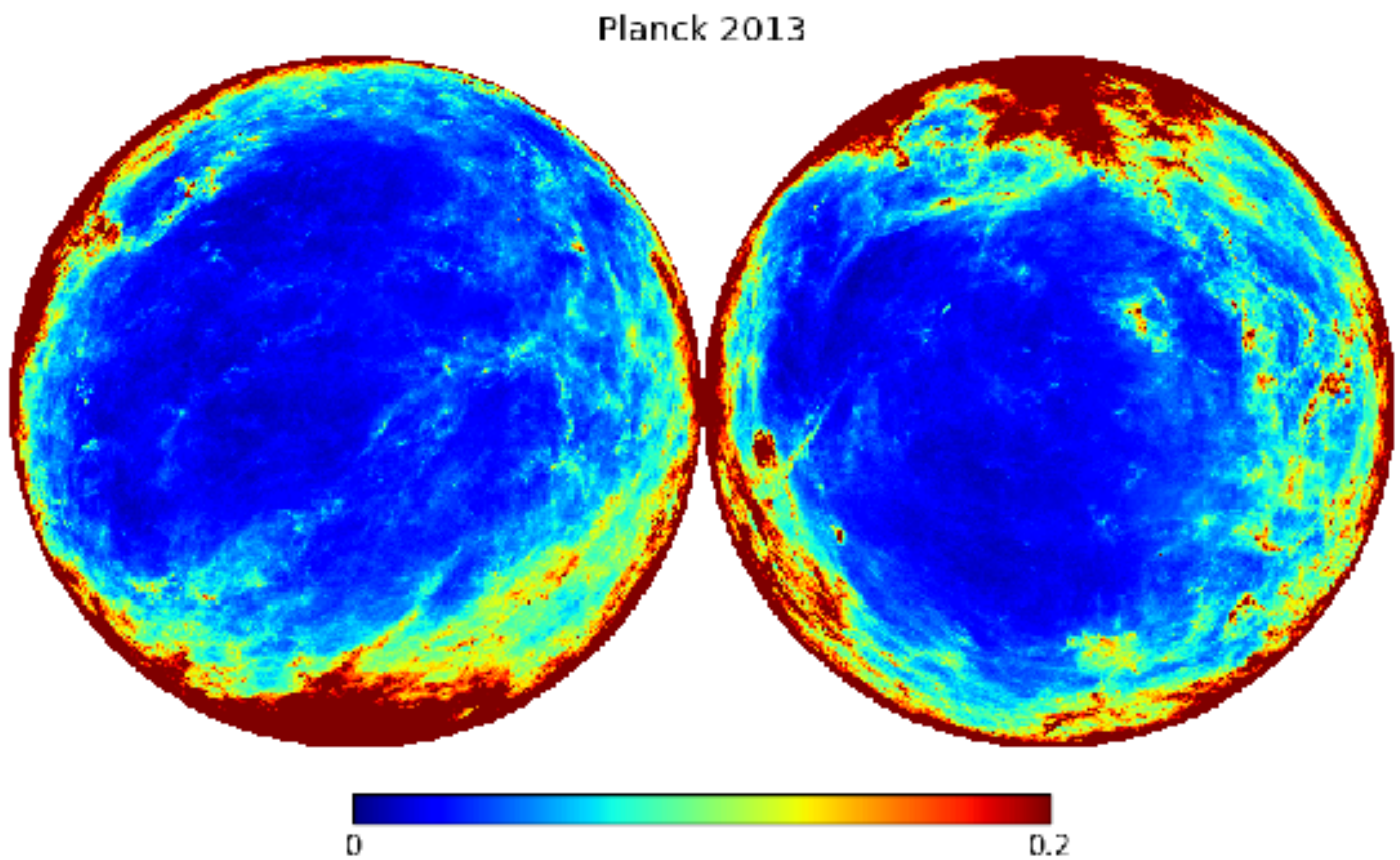
A second-order expansion in δL/L gives:

$$n_o(\mathbf{x}) = \int_{L_0+\delta L(\boldsymbol{\theta})}^{\infty} [1 + b_1(L)\delta(\mathbf{x})]\Phi(L)dL$$

$$\simeq \langle n\rangle[1 + \bar{b}_1(L_0)\delta(\mathbf{x})] - \boxed{\Phi(L_0)[1 + b_1(L_0)\delta(\mathbf{x})]\delta L} - \boxed{\frac{1}{2}\frac{d}{dL}[\Phi(L)(1 + b_1(L)\delta(\mathbf{x}))]_{L_0}(\delta L)^2}$$

This involves luminosity dependence of bias

$$\mathcal{M}(\boldsymbol{\theta}) = E(B-V)$$



Planck maps have been resampled to the same resolution of SFD

The (purely angular) mask has an impact that depends on redshift

$$\boxed{C(z) = 0.4 \ln 10\, R(z)}$$

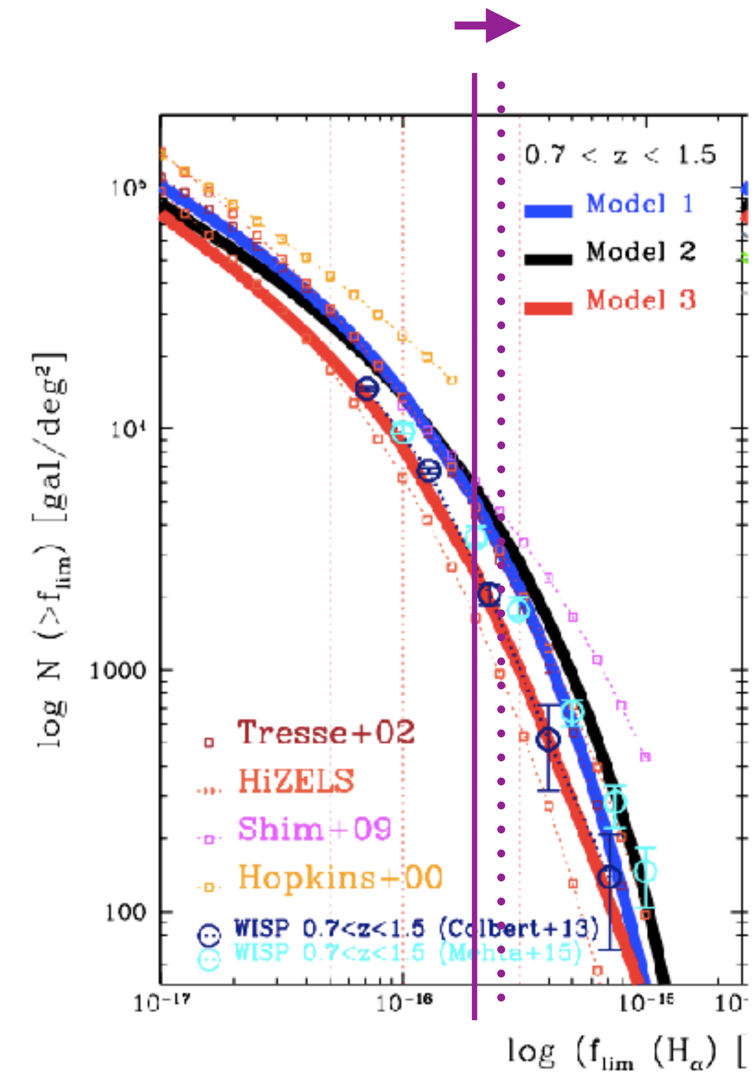
this is proportional to the extinction curve

$$\epsilon(\boldsymbol{\theta}) = \exp[C(z)\,\mathcal{M}(\boldsymbol{\theta})] - 1 \simeq C(z)\mathcal{M}(\boldsymbol{\theta}) + \frac{1}{2}C^2(z)\mathcal{M}^2(\boldsymbol{\theta})$$

But the mask changes the average density

$$\langle n\rangle(z) = B_1(z)\Phi(L_0|z)L_0(z) \qquad B_2(z) = \frac{d\Phi}{dL}(L_0|z)\frac{L_0(z)}{2\Phi(L_0|z)} + \frac{1}{2}$$

these functions depend on the shape of the luminosity function



$$\frac{\langle n_o\rangle}{\Phi L_0} = B_1 - \langle\epsilon\rangle - \left(B_2 - \frac{1}{2}\right)\langle\epsilon^2\rangle$$

$$\delta_o = \frac{n_o}{\langle n_o\rangle} - 1 \simeq \boxed{\frac{B_1\bar{b}_1 - C\mathcal{M}b_1 - B_2C^2\mathcal{M}^2b_1 - \left(B_2 - \frac{1}{2}\right)C^2\mathcal{M}^2b_1'\frac{\Phi}{\Phi'}}{B_1 - C\langle\mathcal{M}\rangle - B_2C^2\langle\mathcal{M}^2\rangle}} \times \delta$$

$$\boxed{-\frac{C\left(\mathcal{M} - \langle\mathcal{M}\rangle\right) + B_2C^2\left(\mathcal{M}^2 - \langle\mathcal{M}^2\rangle\right)}{B_1 - C\langle\mathcal{M}\rangle - B_2C^2\langle\mathcal{M}^2\rangle}}$$

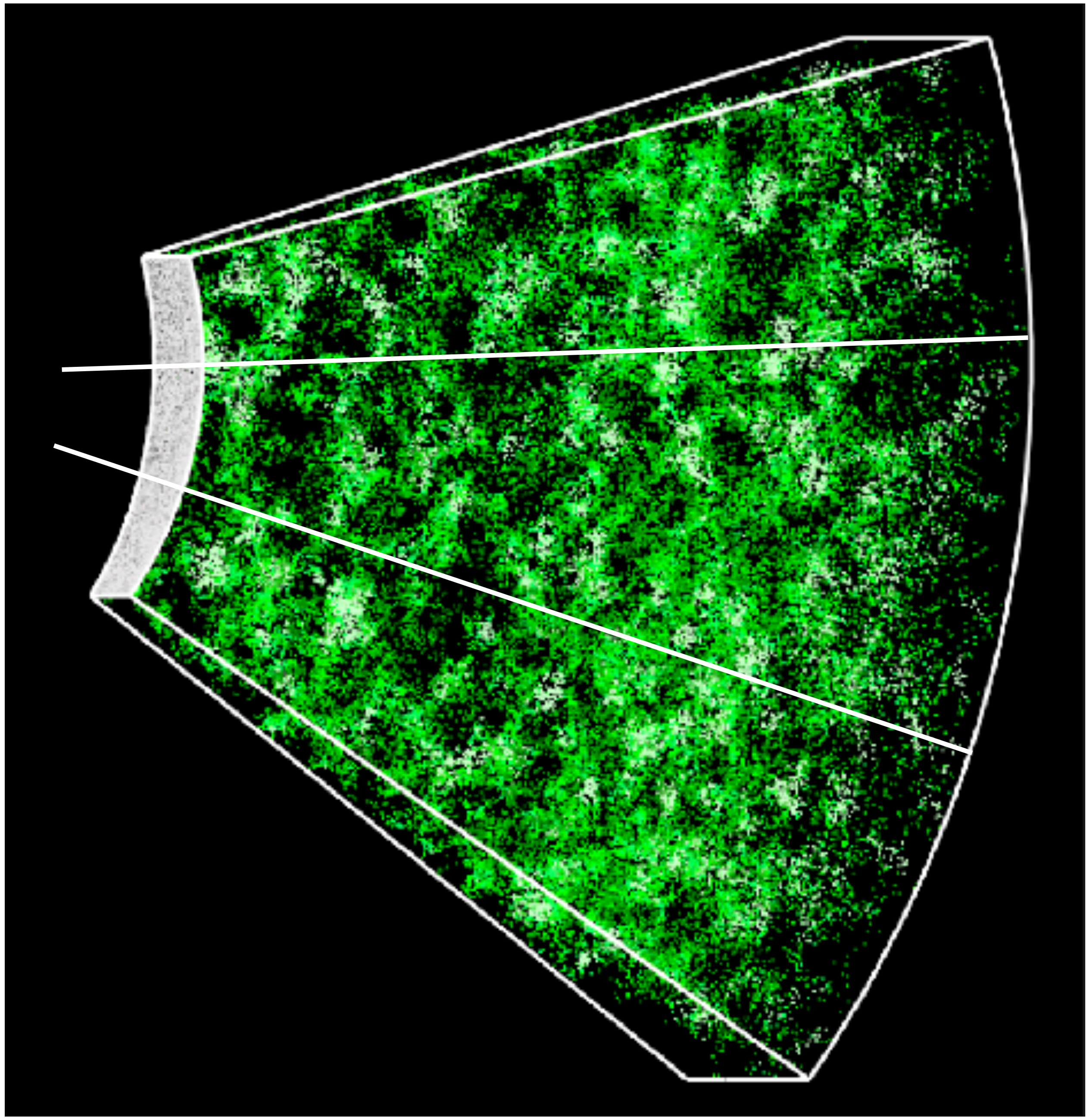
this is the cosmological signal
and it averages out for our estimators

this is the effect of the mask

Define estimators that are sensitive to the "mask" M:

$$Avz(\boldsymbol{\theta}) \equiv \frac{1}{N_z} \sum_i \delta_o([z_i, \boldsymbol{\theta}])$$

$$Ccz(\boldsymbol{\theta}) \equiv \frac{1}{N_p} \sum_i \sum_{j>i} \delta_o([z_i, \boldsymbol{\theta}])\delta_o([z_j, \boldsymbol{\theta}])$$



If prior knowledge of <M> and <M^2> is assumed:

$$\bar{\delta}_o = \frac{n_o}{\langle n \rangle} - 1 \simeq \frac{B_1 \bar{b}_1 - C\mathscr{M} b_1 - B_2 C^2 \mathscr{M}^2 b_1 + \left(B_2 - \frac{1}{2}\right) C^2 \mathscr{M}^2 b_1' \frac{\Phi}{\Phi'}}{B_1} \times \delta - \frac{C\mathscr{M} + B_2 C^2 \mathscr{M}^2}{B_1}$$

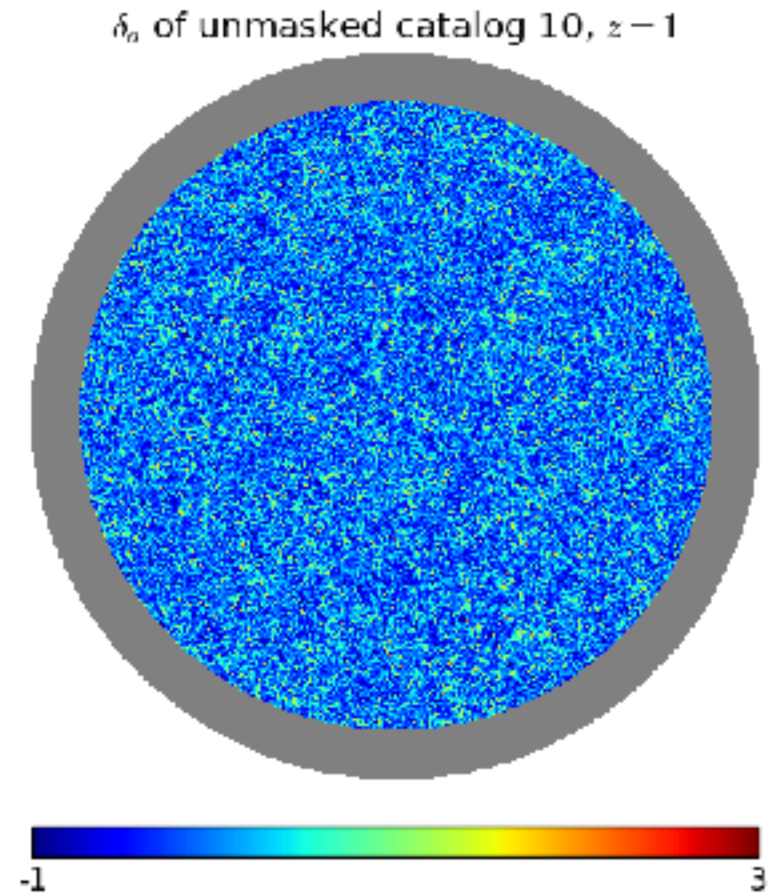
$$Avz(\boldsymbol{\theta}) \simeq \boxed{-\frac{1}{N_z} \sum_i \frac{C_i}{B_{1i}}} \mathscr{M} - \boxed{\frac{1}{N_z} \sum_i \frac{B_{2i} C_i^2}{B_{i1}}} \mathscr{M}^2 .$$

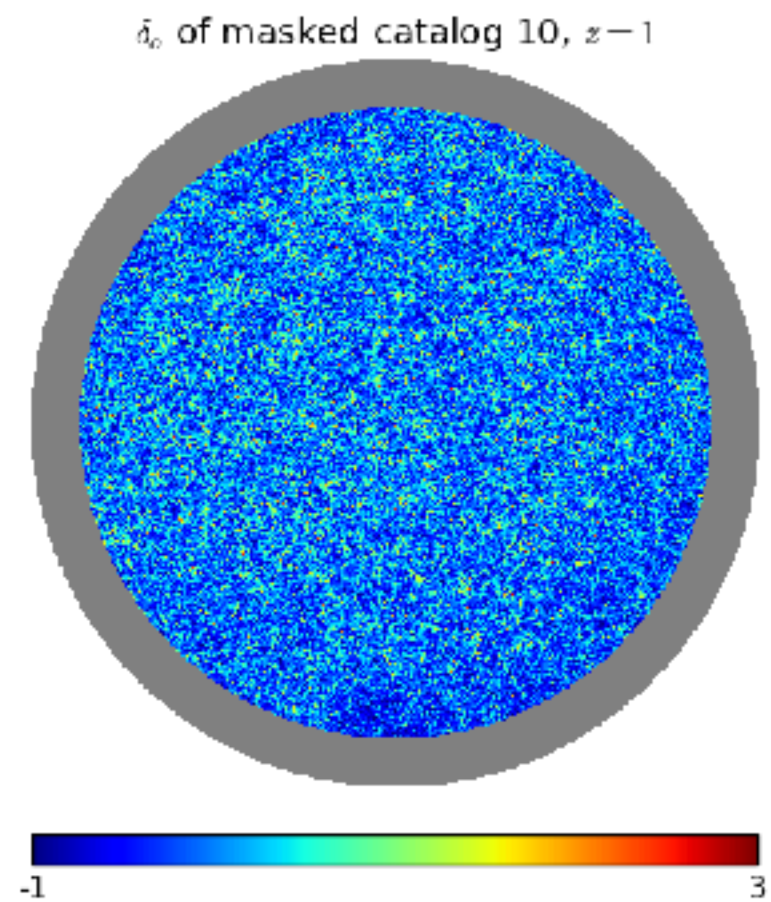
$$Ccz(\boldsymbol{\theta}) \simeq \boxed{\frac{1}{N_p} \sum_i \sum_{j>i} \frac{C_i C_j}{B_{1i} B_{1j}}} \mathscr{M}^2 + \boxed{\frac{1}{N_p} \sum_i \sum_{j>i} \frac{C_i C_j (B_{2i} C_i + B_{2j} C_j)}{B_{1i} B_{1j}}} \mathscr{M}^3 + \boxed{\frac{1}{N_p} \sum_i \sum_{j>i} \frac{B_{2i} B_{2j} C_i^2 C_j^2}{B_{1i} B_{1j}}} \mathscr{M}^4 .$$

these coefficients require knowledge
of the luminosity function
and of the redshift dependence of the mask impact

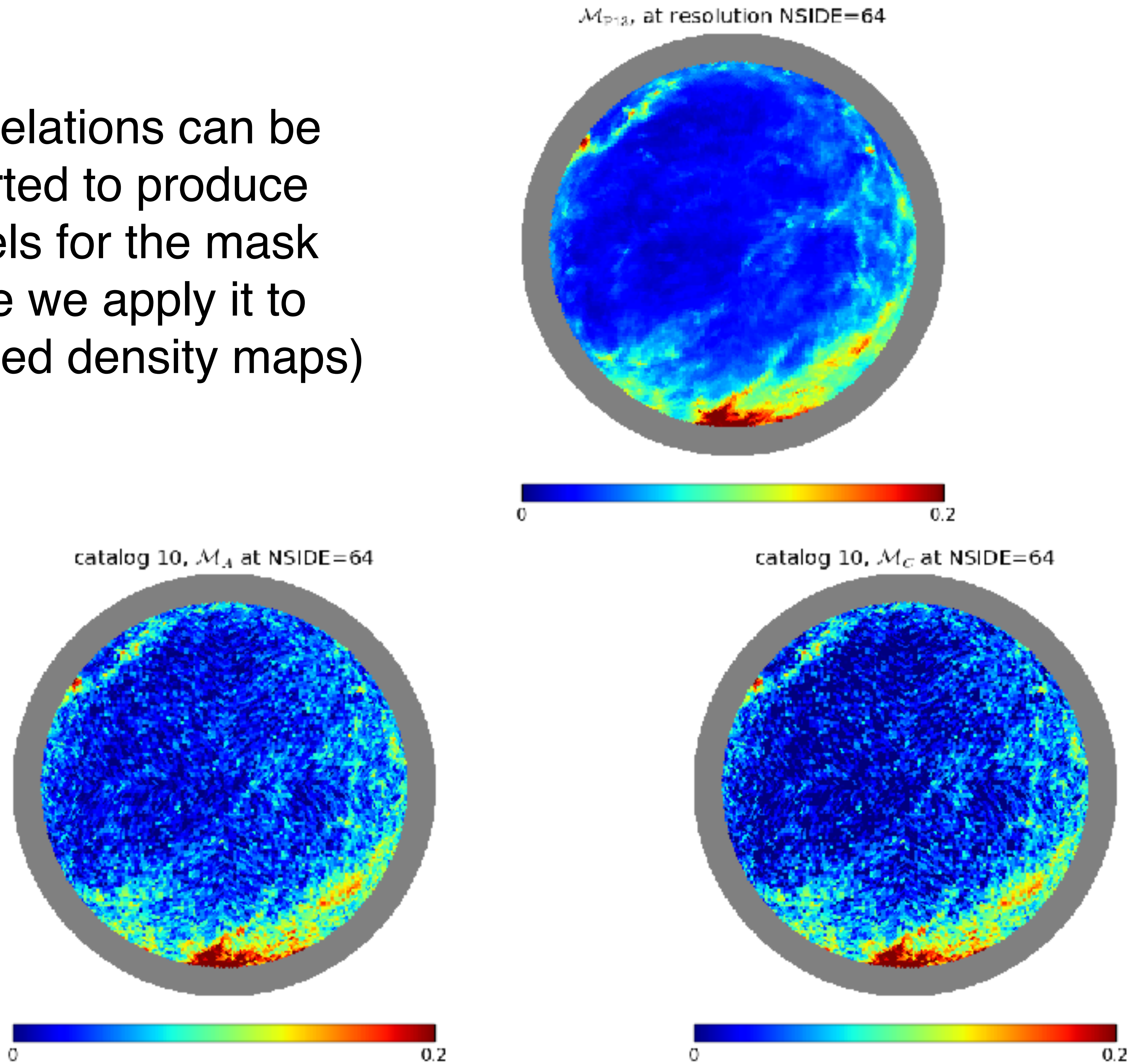
Mocks

- Generate 20 light cones with PINOCCHIO:
  - 3.2 Gpc/h box sampled with 4096^3 particles,
  - smallest halo 7.5 10^{11} M_{sun}/h (20 particles),
  - light cone from z=2.5 to z=0, covering 1/4 of the sky;
- abundance matching of halos with LF of Hα emitters, model 1 of Pozzetti et al. (2016);
- **“shuffled” masses** to **remove luminosity-dependent bias**;
- flux limit of 2 10^{-16} erg s^{-1} cm^{-2}, complete from z=0.8 to z=2.5;
- apply galactic extinction using P13 map;
- create density maps on the sky with healpy;
- redshift bins of delta z=0.1, no redshift error;

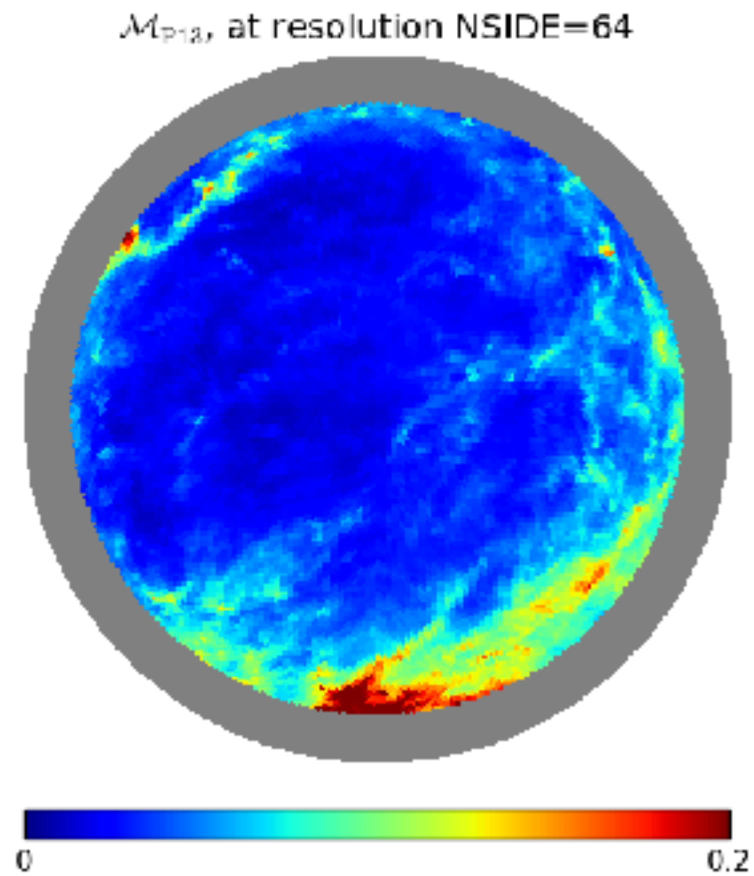


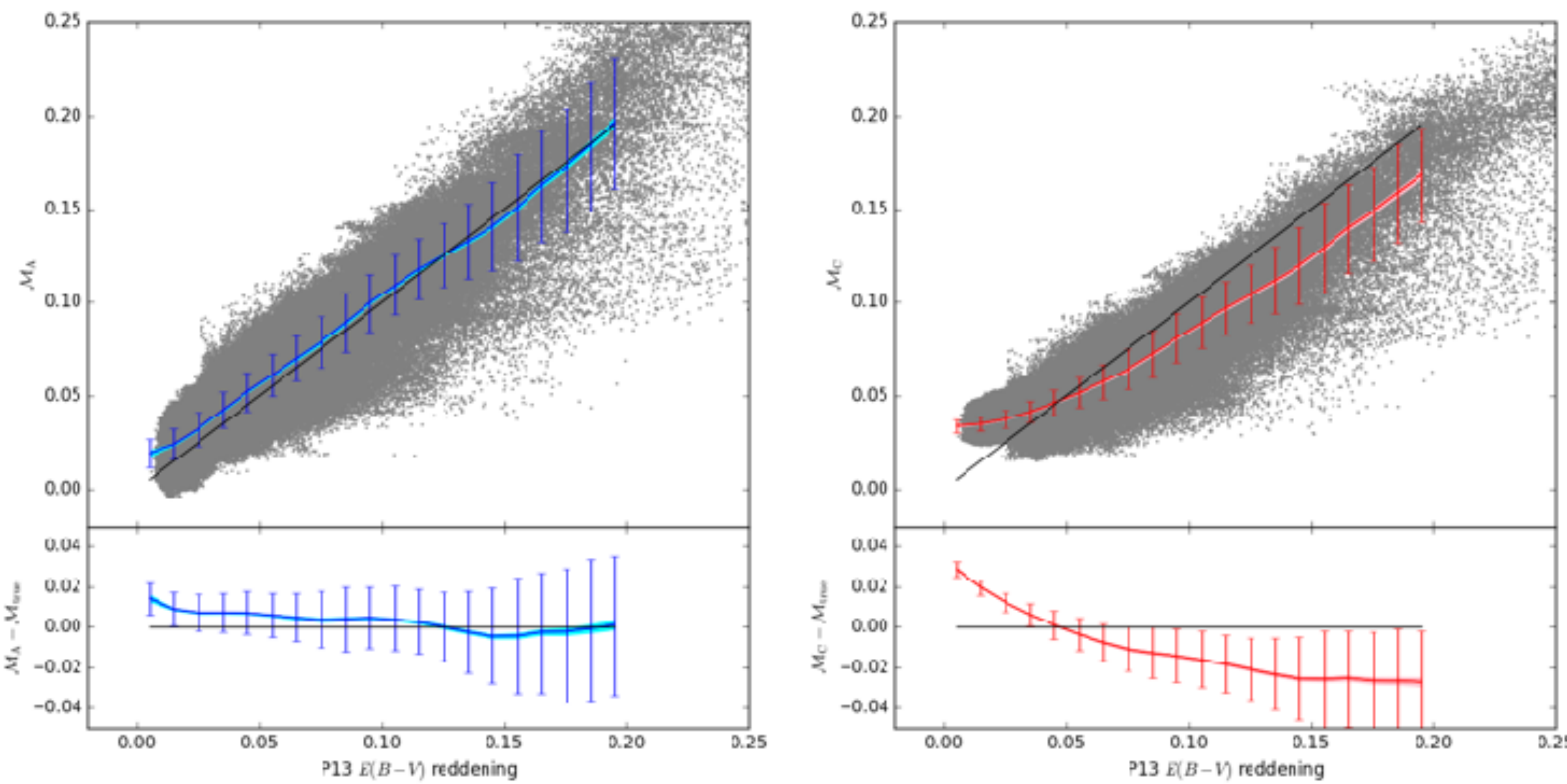


The relations can be inverted to produce models for the mask (here we apply it to smoothed density maps)



The relations can be inverted to produce models for the mask (here we apply it to smoothed density maps)

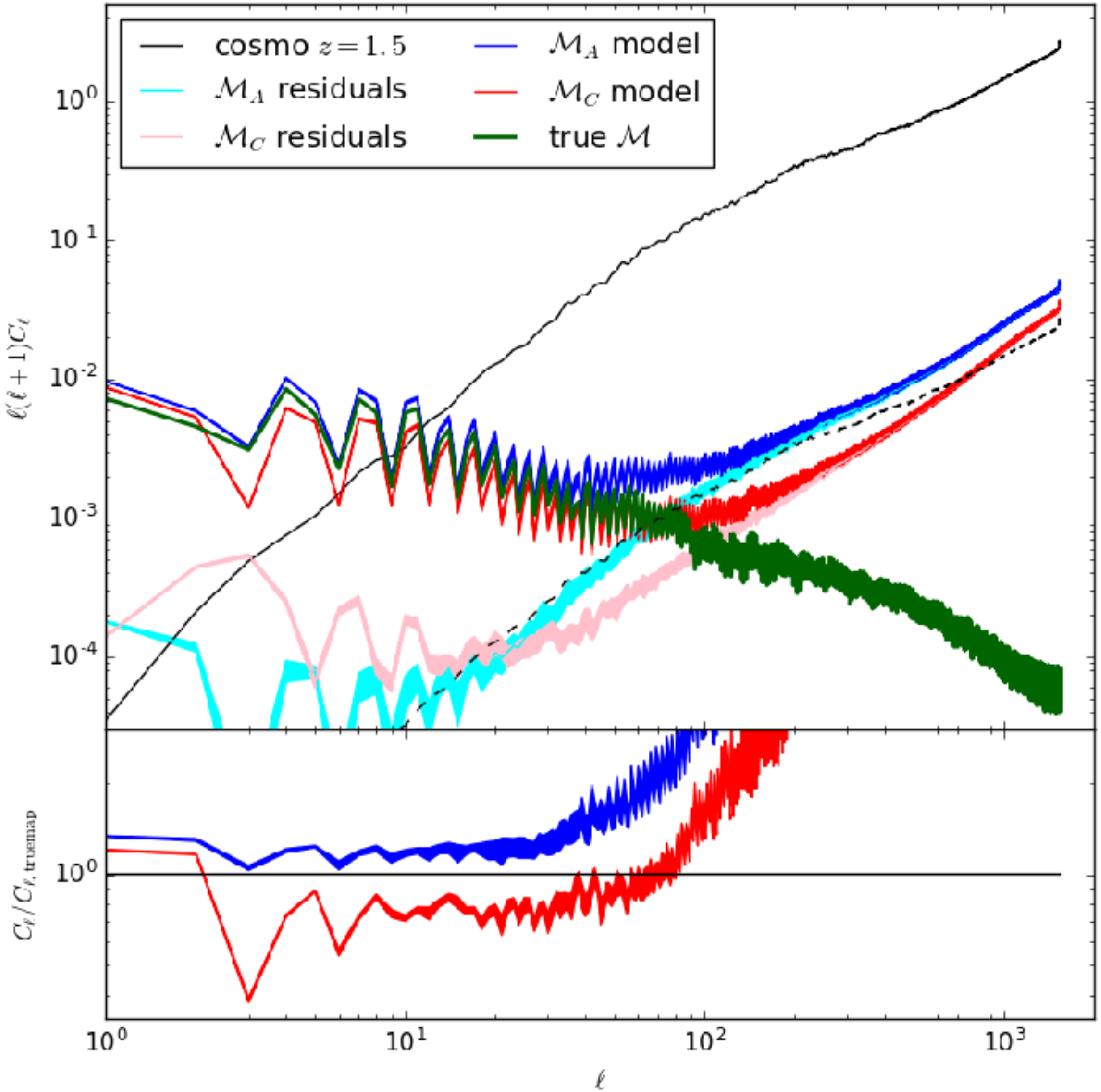




Here we compute M_A and M_C on the full resolution density maps

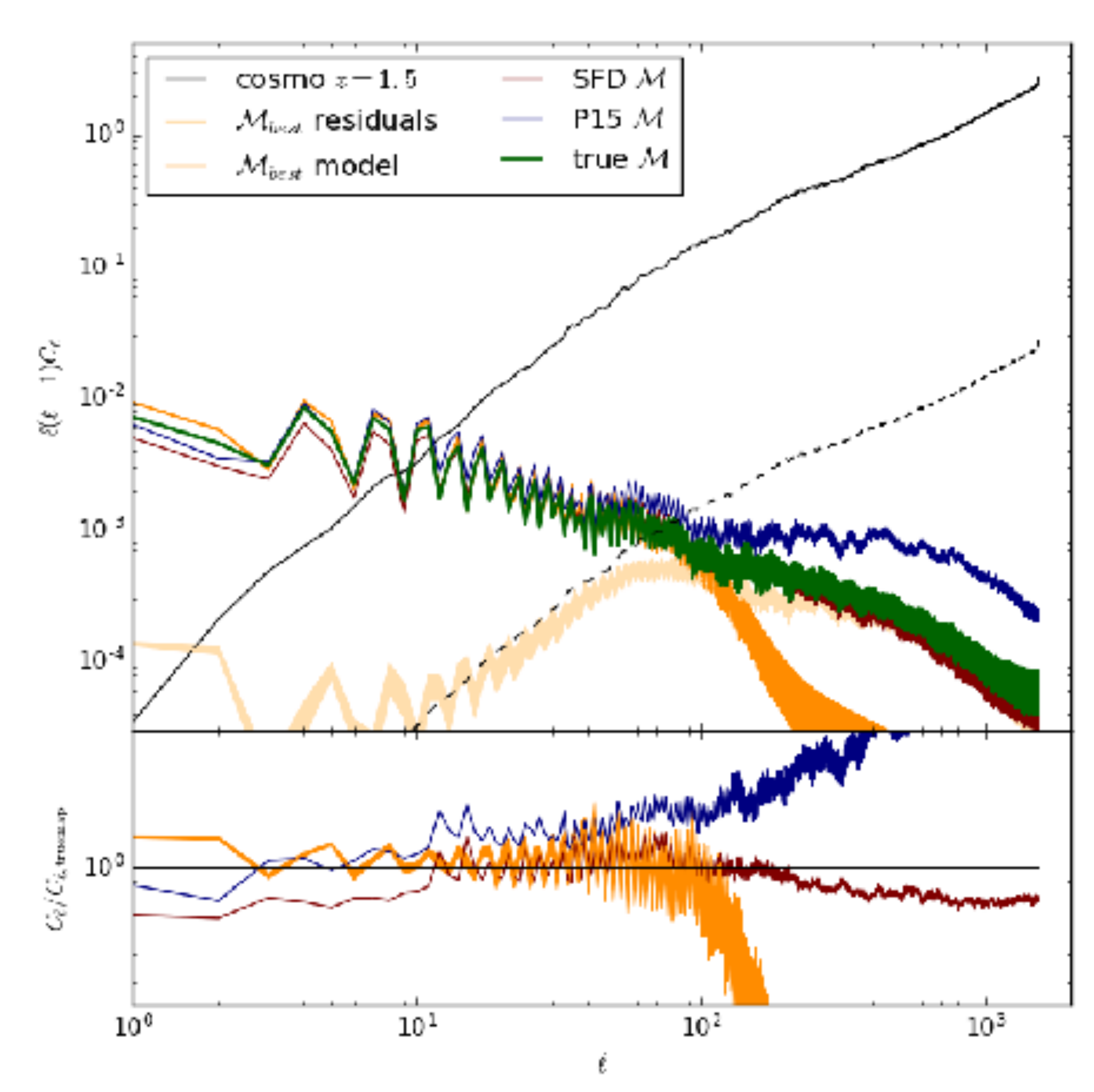
Residuals with respect to the true mask correlate with the cosmic signal

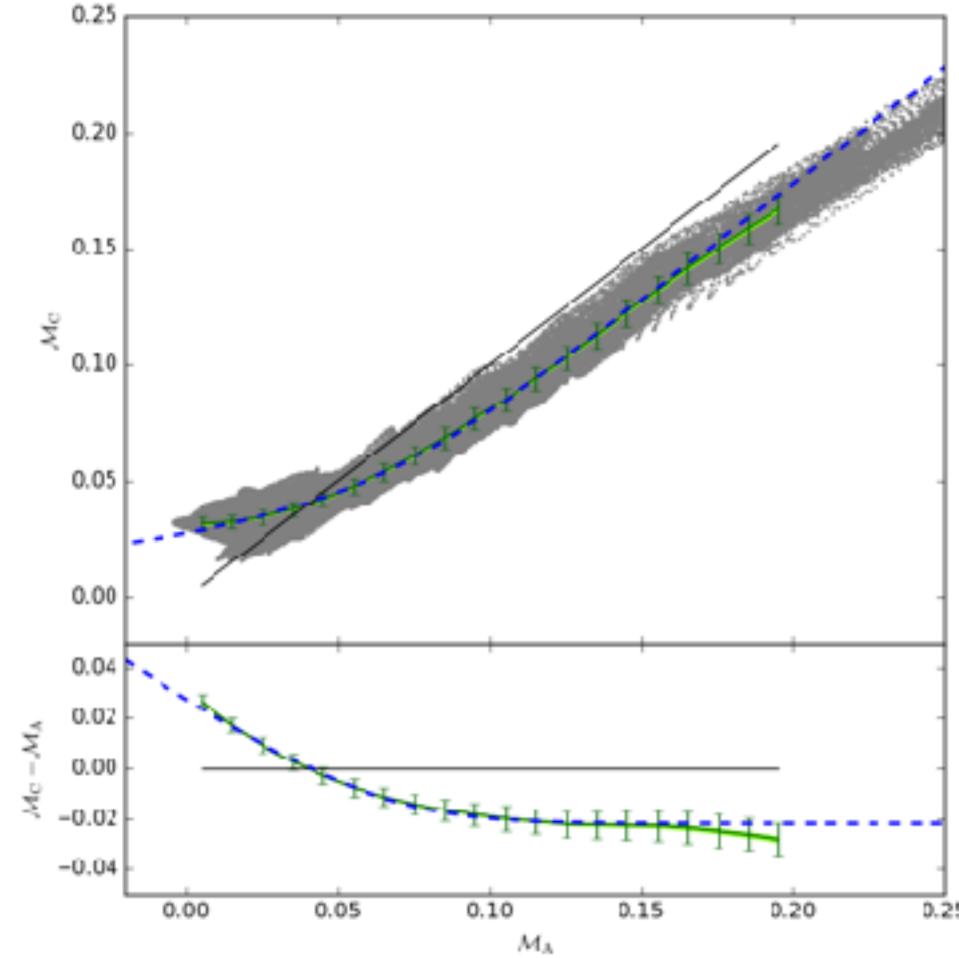
The model M_C based on Ccz has lower residuals

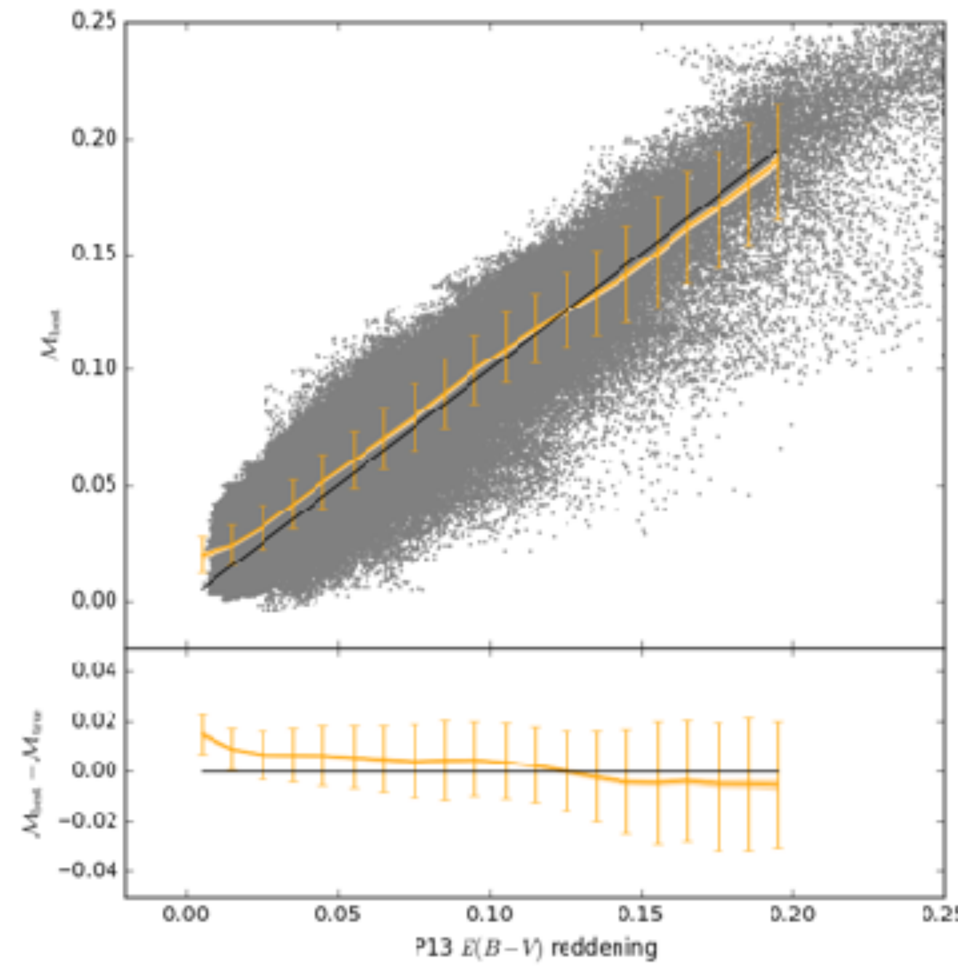


- smooth both models to fiter residuals out
- find the relation $f(M_A)$ between M_A and M_C
- combine them as:

$$\mathcal{M}_{\text{best}} = \mathcal{M}_C - (f(\mathcal{M}_A) - \mathcal{M}_A)$$

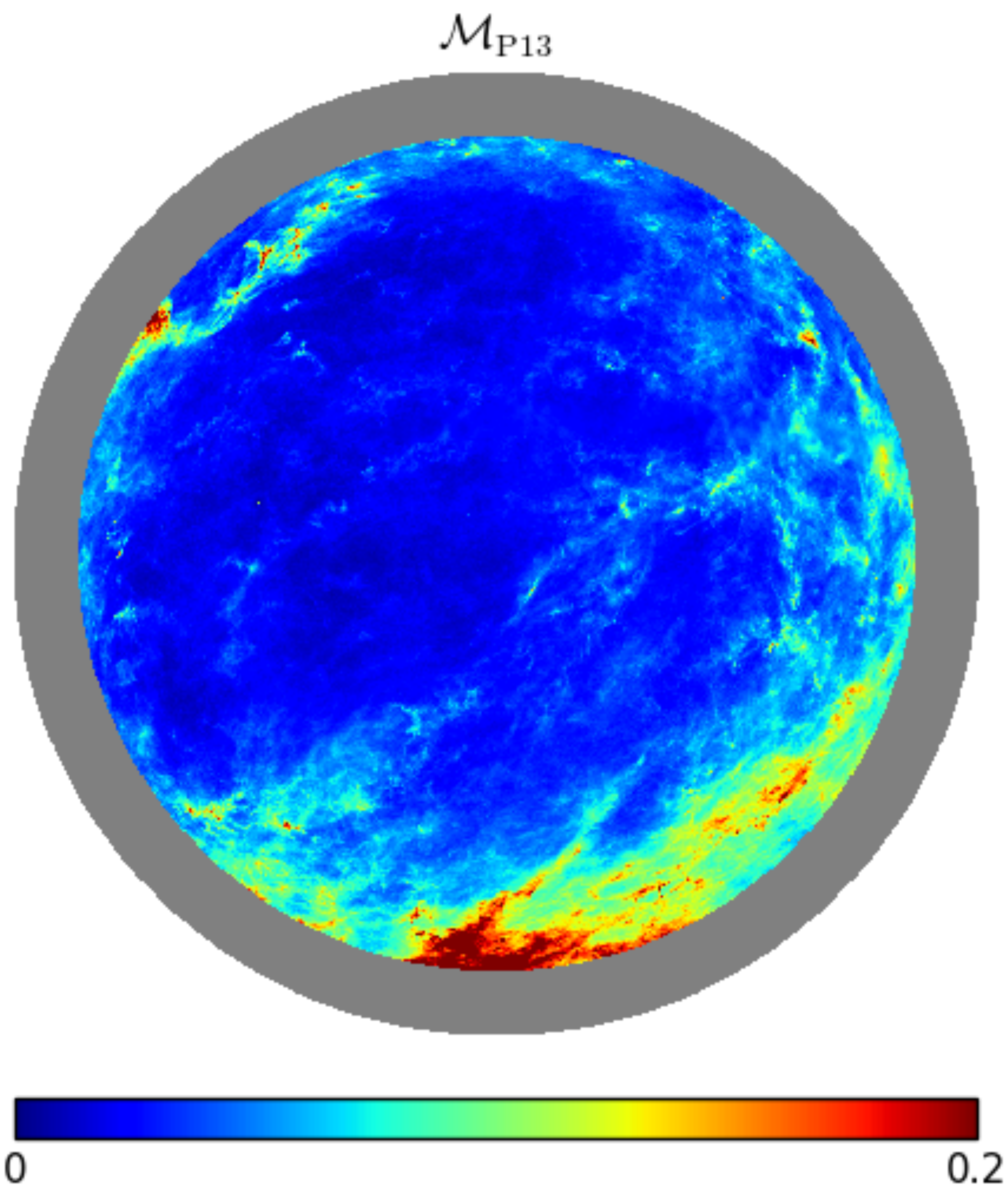


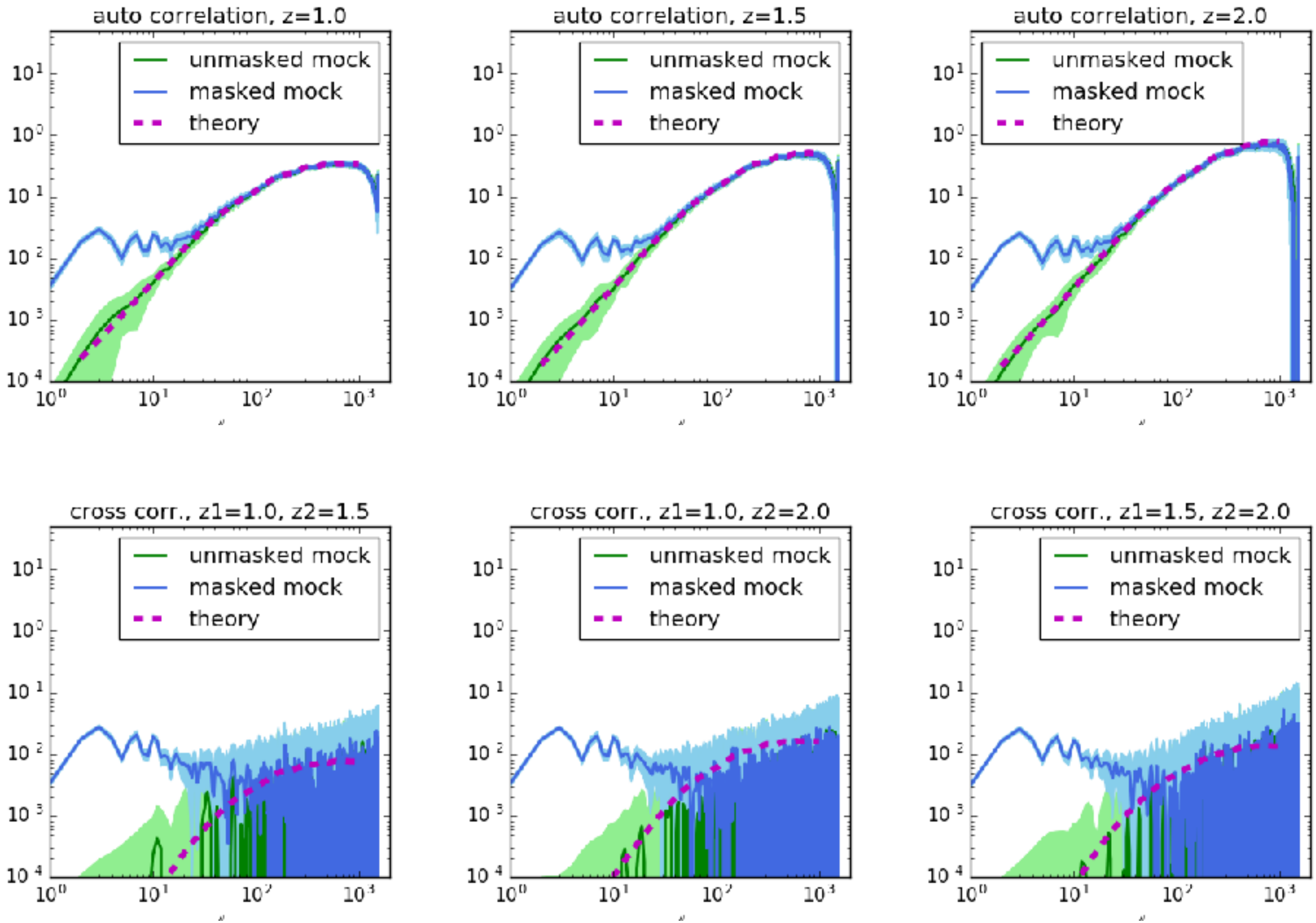




e next

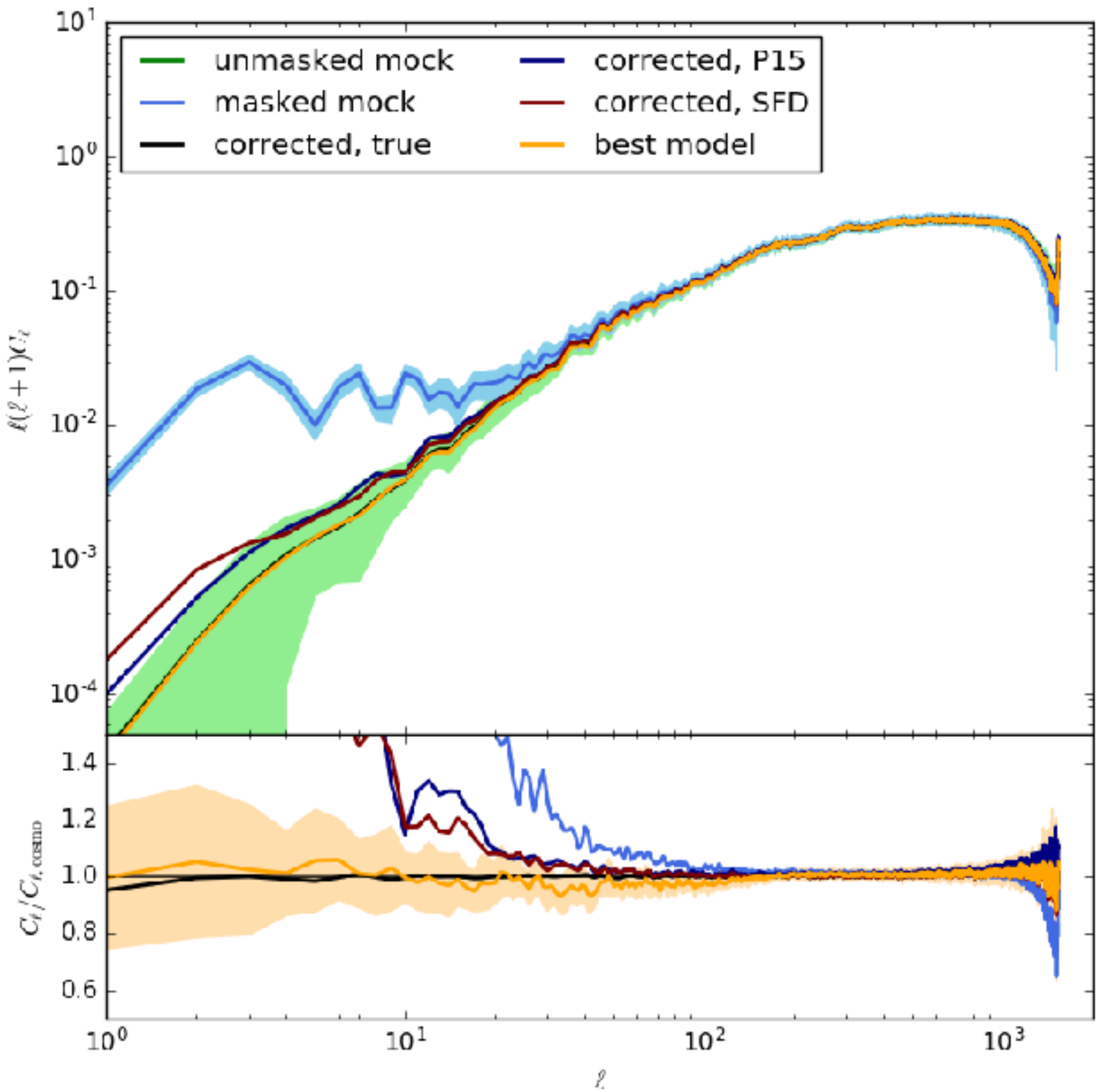
2018


$\mathcal{M}_{\rm P13}$
catalog 10, $\mathcal{M}_{\rm best}$
0
0.2
0
0.2

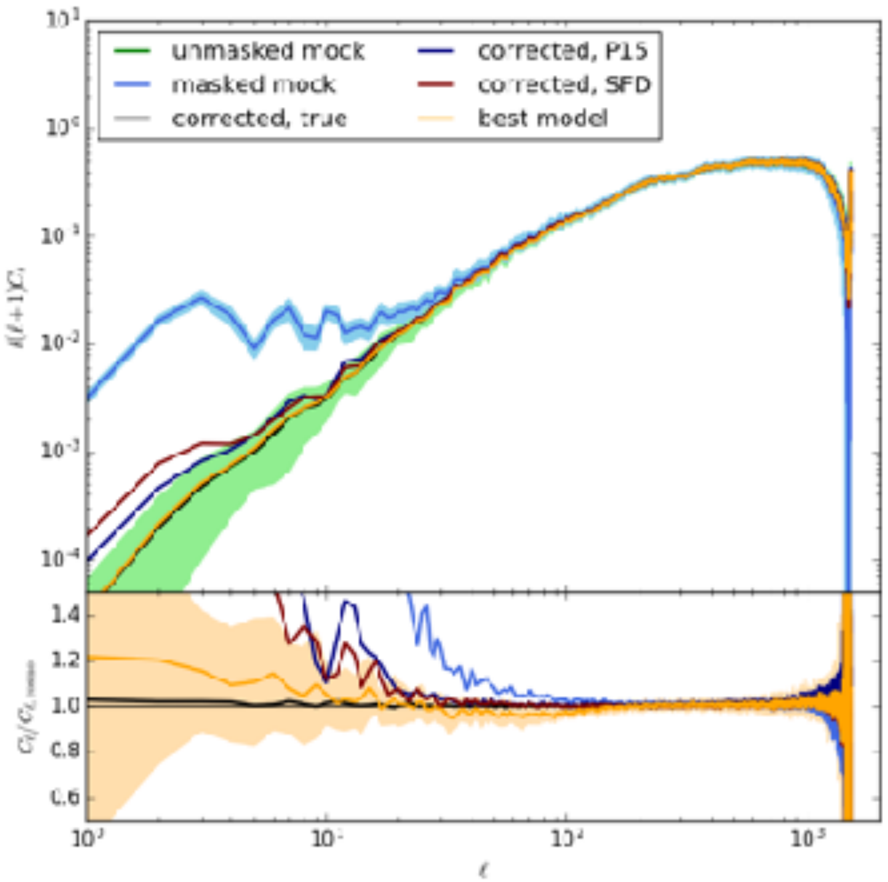

auto correlation, z=1.0
auto correlation, z=1.5
auto correlation, z=2.0
cross corr., z1=1.0, z2=1.5
cross corr., z1=1.0, z2=2.0
cross corr., z1=1.5, z2=2.0
unmasked mock
masked mock
theory

$$\delta_\mathrm{g} = \frac{\left(B_1 - C\langle \mathscr{M} \rangle - B_2 C^2 \langle \mathscr{M}^2 \rangle\right)\delta_\mathrm{o} + C(\mathscr{M} - \langle \mathscr{M} \rangle) + B_2 C^2 (\mathscr{M}^2 - \langle \mathscr{M}^2 \rangle)}{B_1 - C\mathscr{M} - B_2 C^2 \mathscr{M}^2}$$

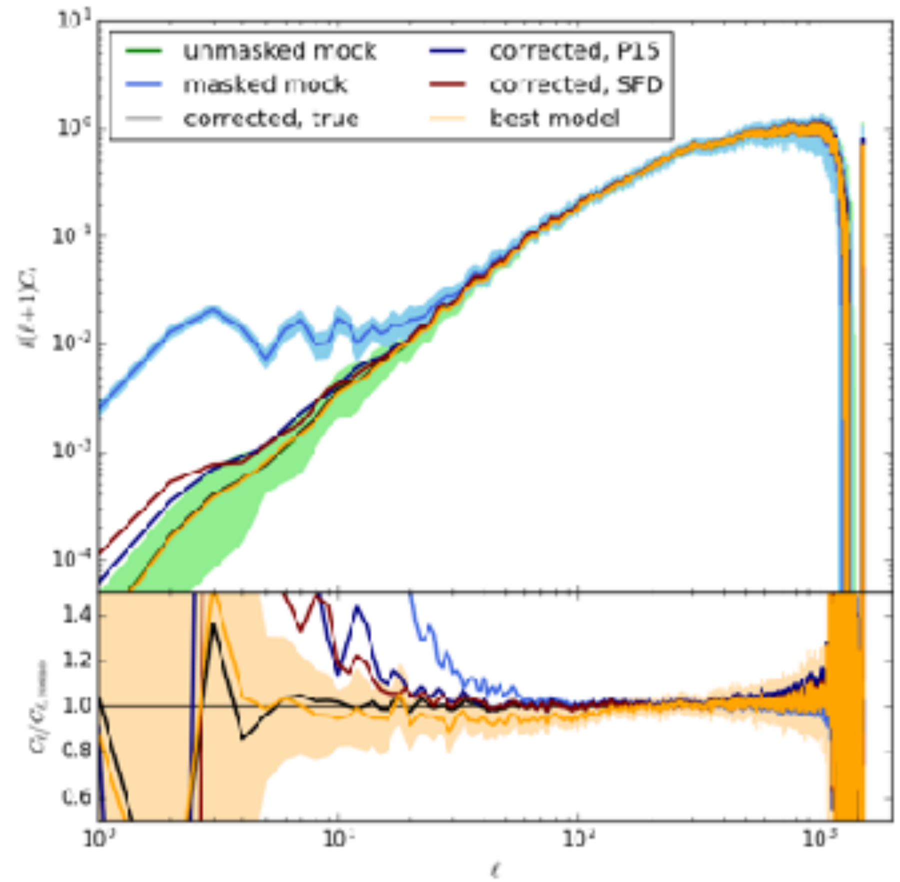
z=1



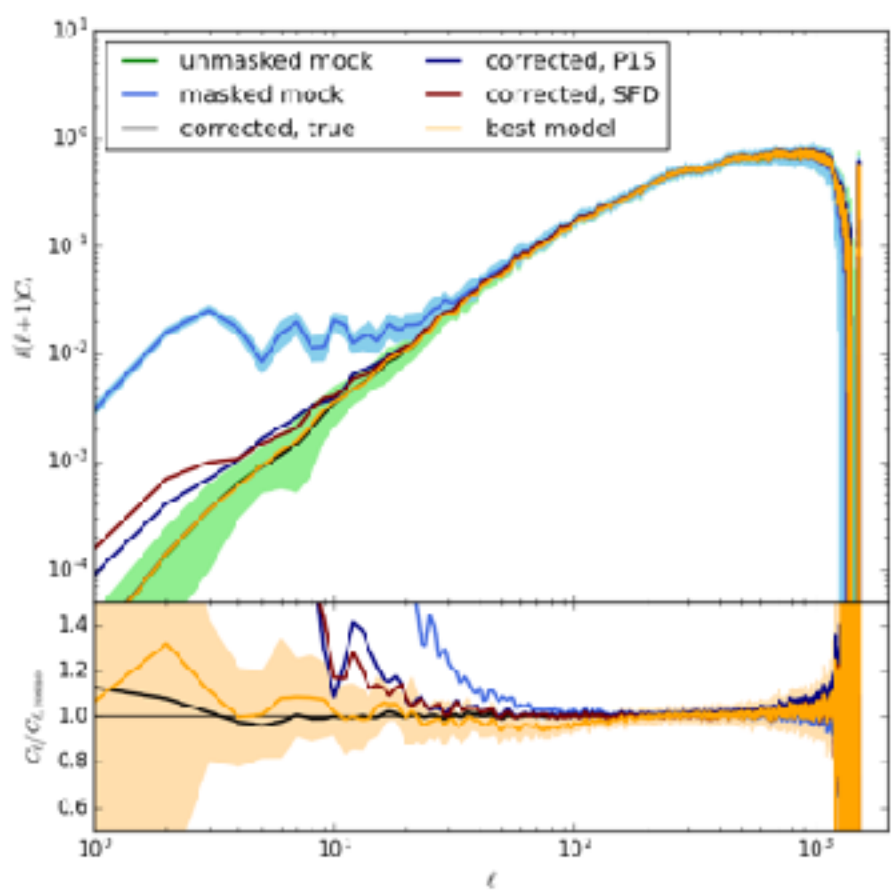
z=1.5

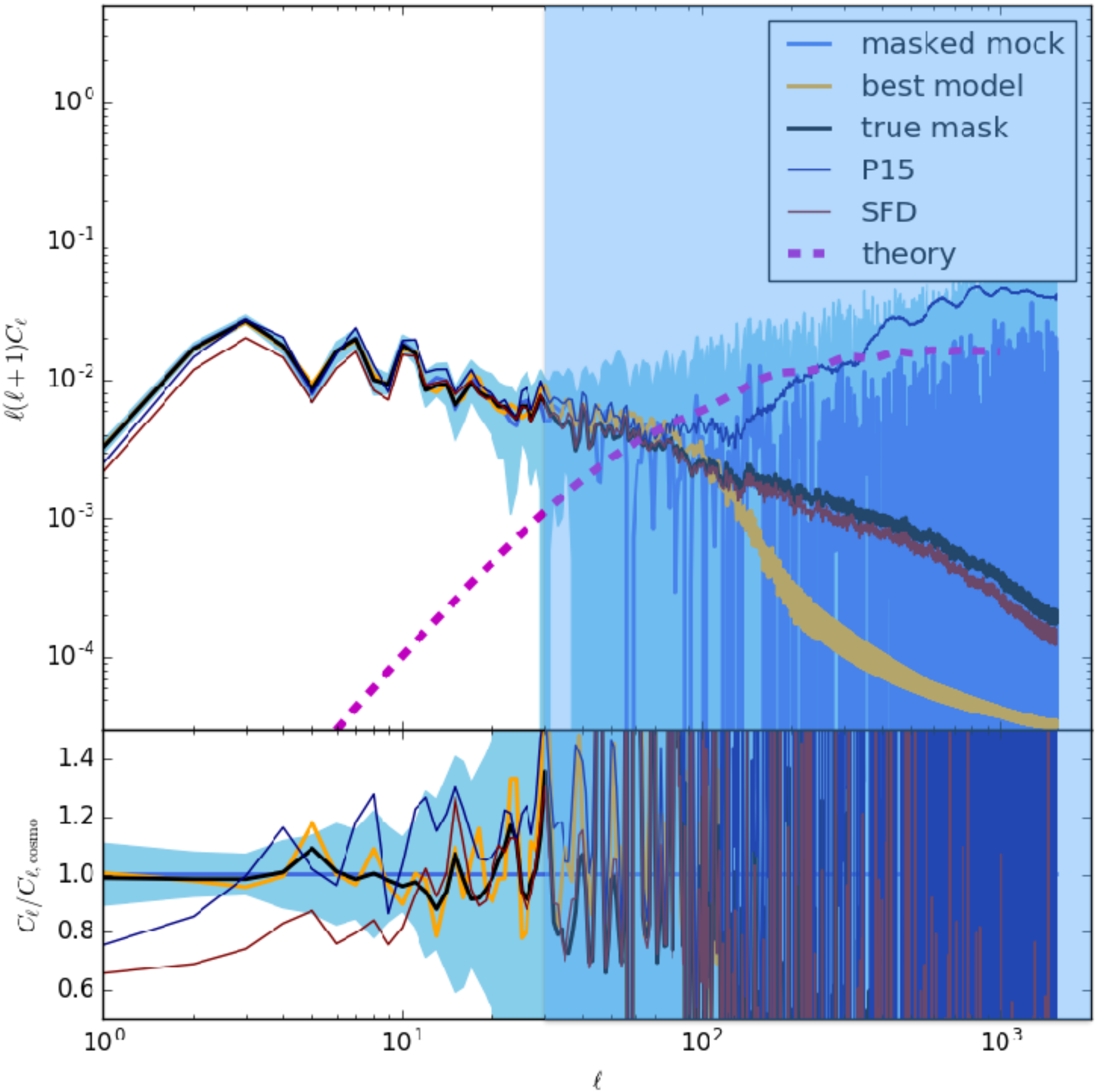

unmasked mock
masked mock
corrected, true
corrected, P15
corrected, SFD
best model

z=2.0


unmasked mock
masked mock
corrected, true
corrected, P15
corrected, SFD
best model

z=2.4


unmasked mock
masked mock
corrected, true
corrected, P15
corrected, SFD
best model

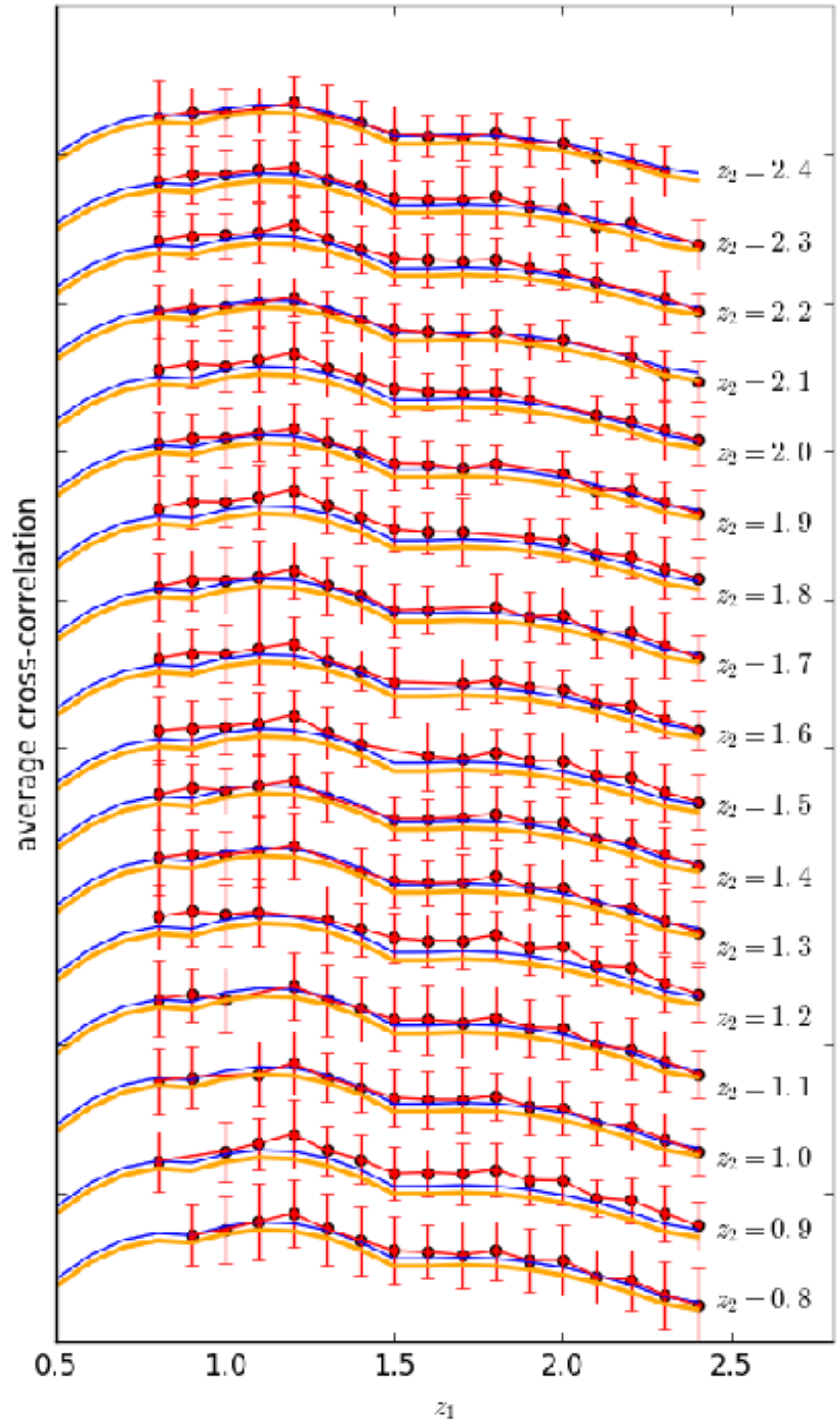

masked mock
best model
true mask
P15
SFD
theory
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
$\ell(\ell+1)C_\ell$
1.4
1.2
1.0
0.8
0.6
$C_\ell/C_{\ell,\,\mathrm{cosmo}}$
10^0
10^1
10^2
10^3
ℓ

$$\delta_o = \frac{n_o}{\langle n_o \rangle} - 1 \simeq$$

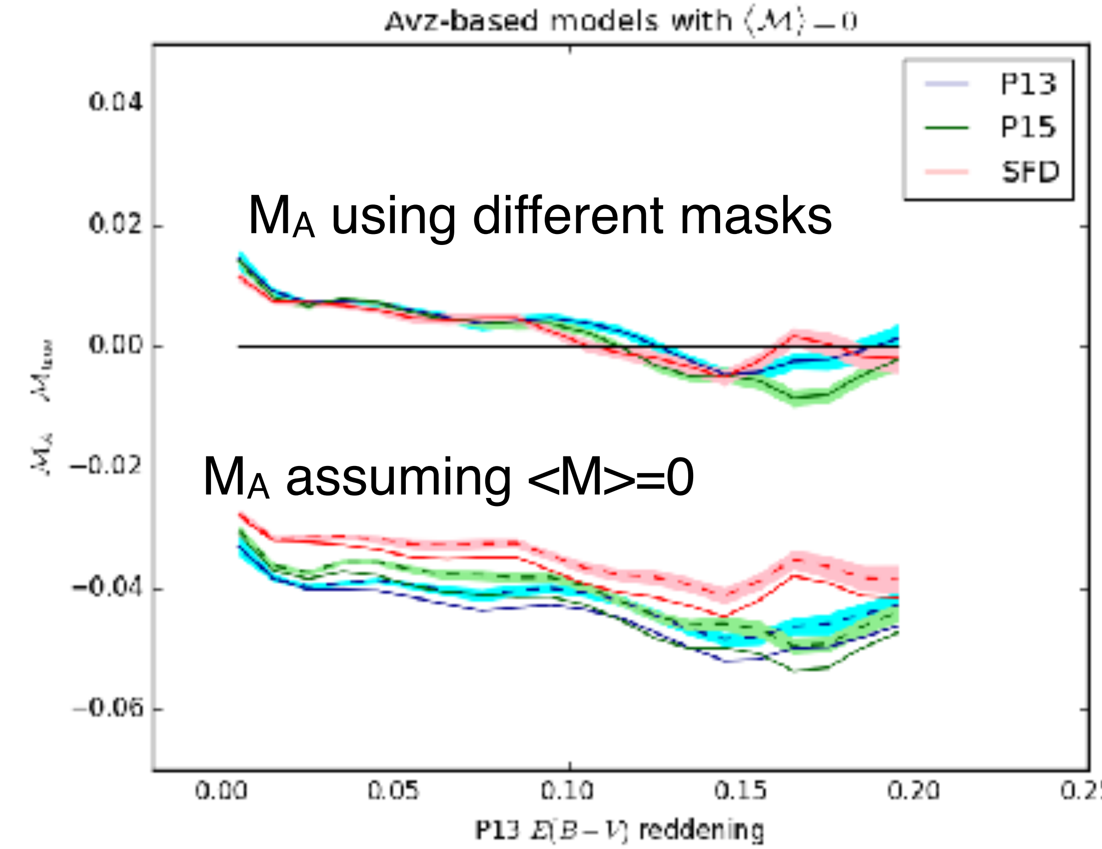
$$- \frac{C\left(\mathscr{M} - \langle \mathscr{M} \rangle\right) + B_2 C^2 \left(\mathscr{M}^2 - \langle \mathscr{M}^2 \rangle\right)}{B_1 - C\langle \mathscr{M} \rangle - B_2 C^2 \langle \mathscr{M}^2 \rangle}$$

average cross correlation as a function of redshift: a strong constraint to the mask model

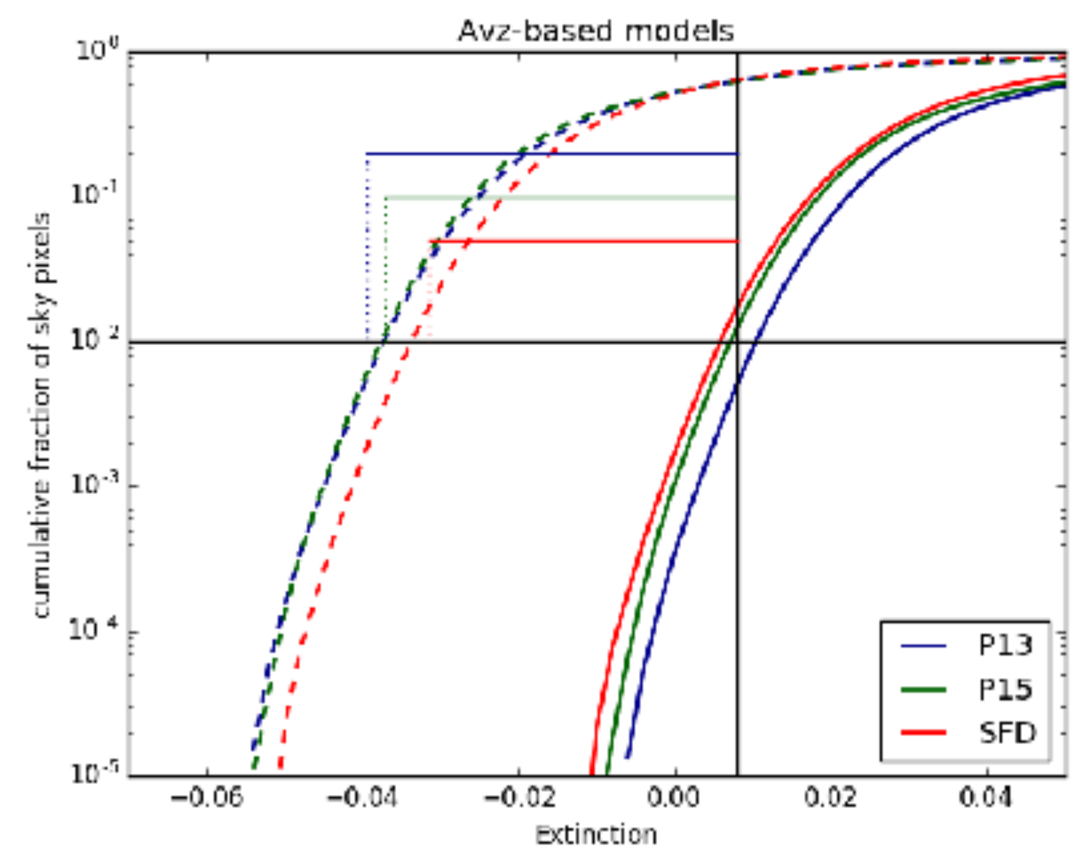
- points: measured mean Cl for masked mock
- errorbars: sample variance over the 20 mocks
- blue curve: predictions with true mask
- orange curve: prediction with best model



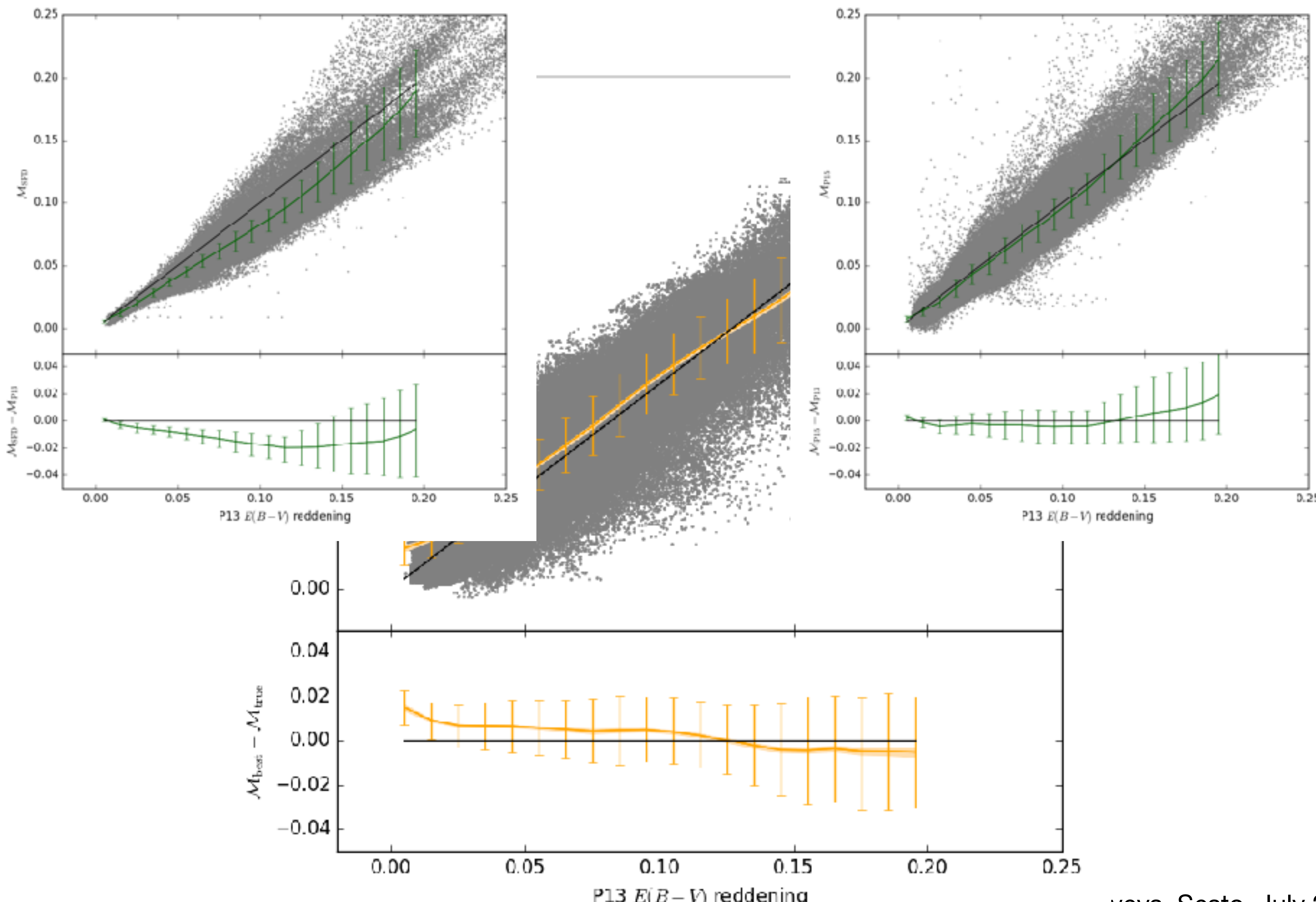
Can we assume we know <M> and <M²>?



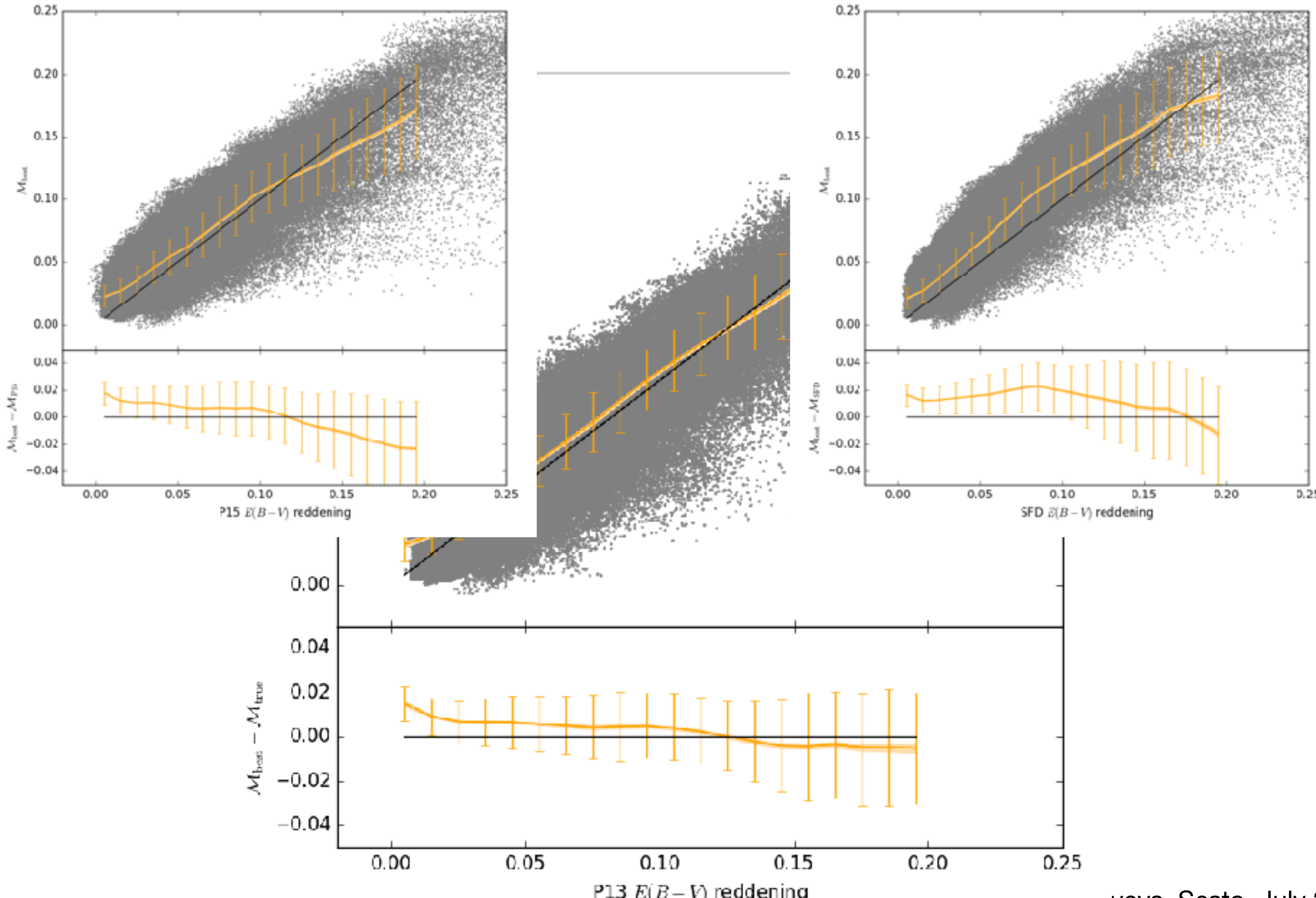
- Using values from P15 and SFD gives very similar results
- <M> can be estimated by using M_A computed for <M>=0
- <M²> can be estimated as the mean of the square of the best model computed with <M²>=0
- In any case, one can calibrate <M> to reproduce the cross correlations



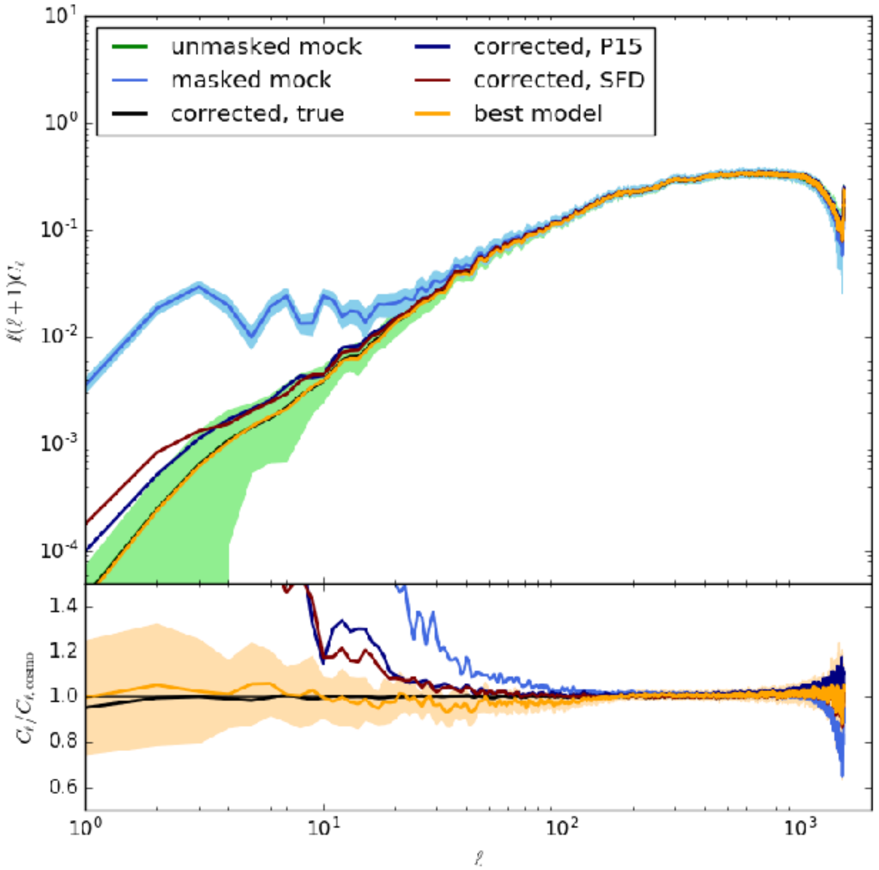
Can we distinguish between reddening models?

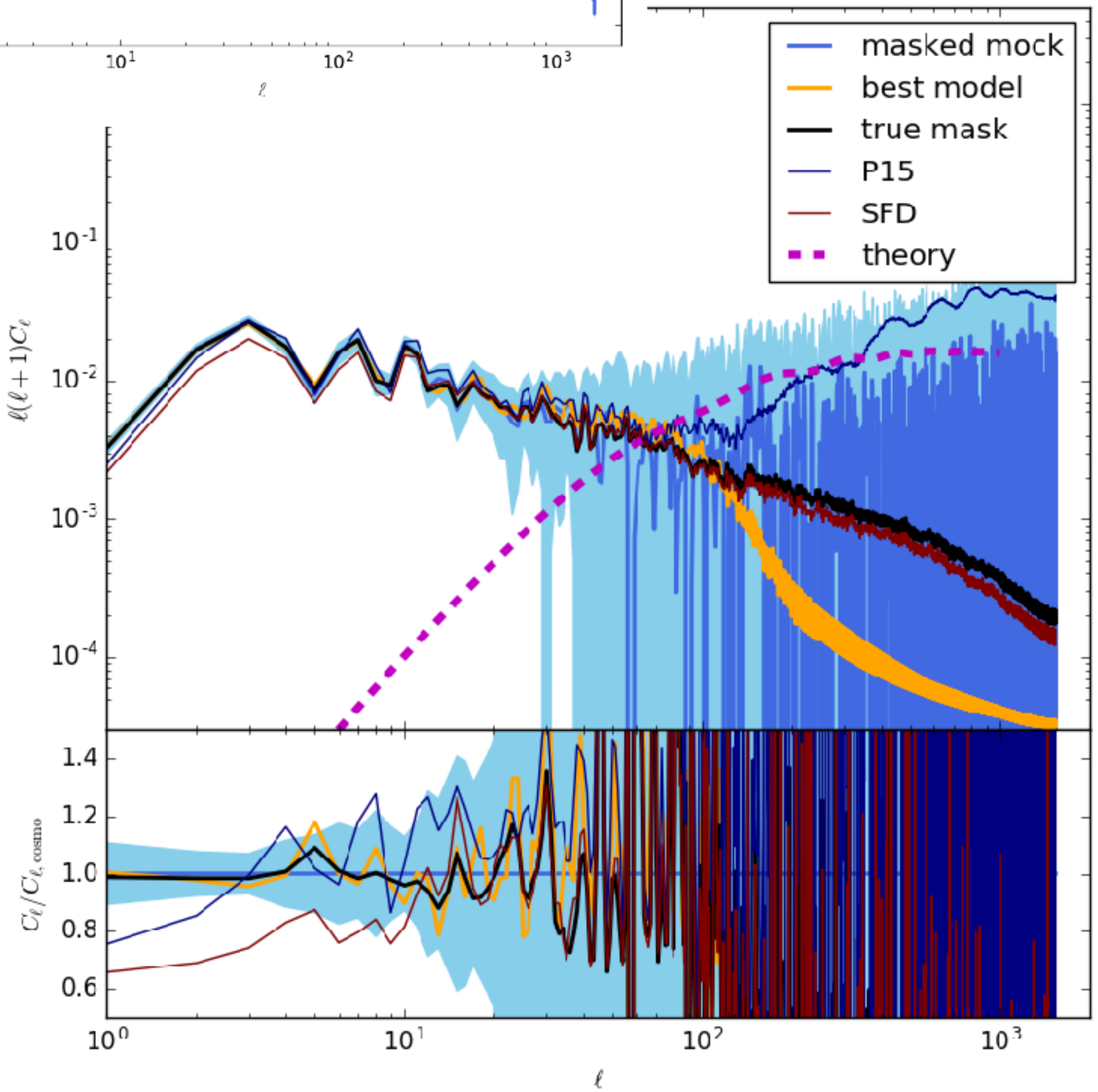


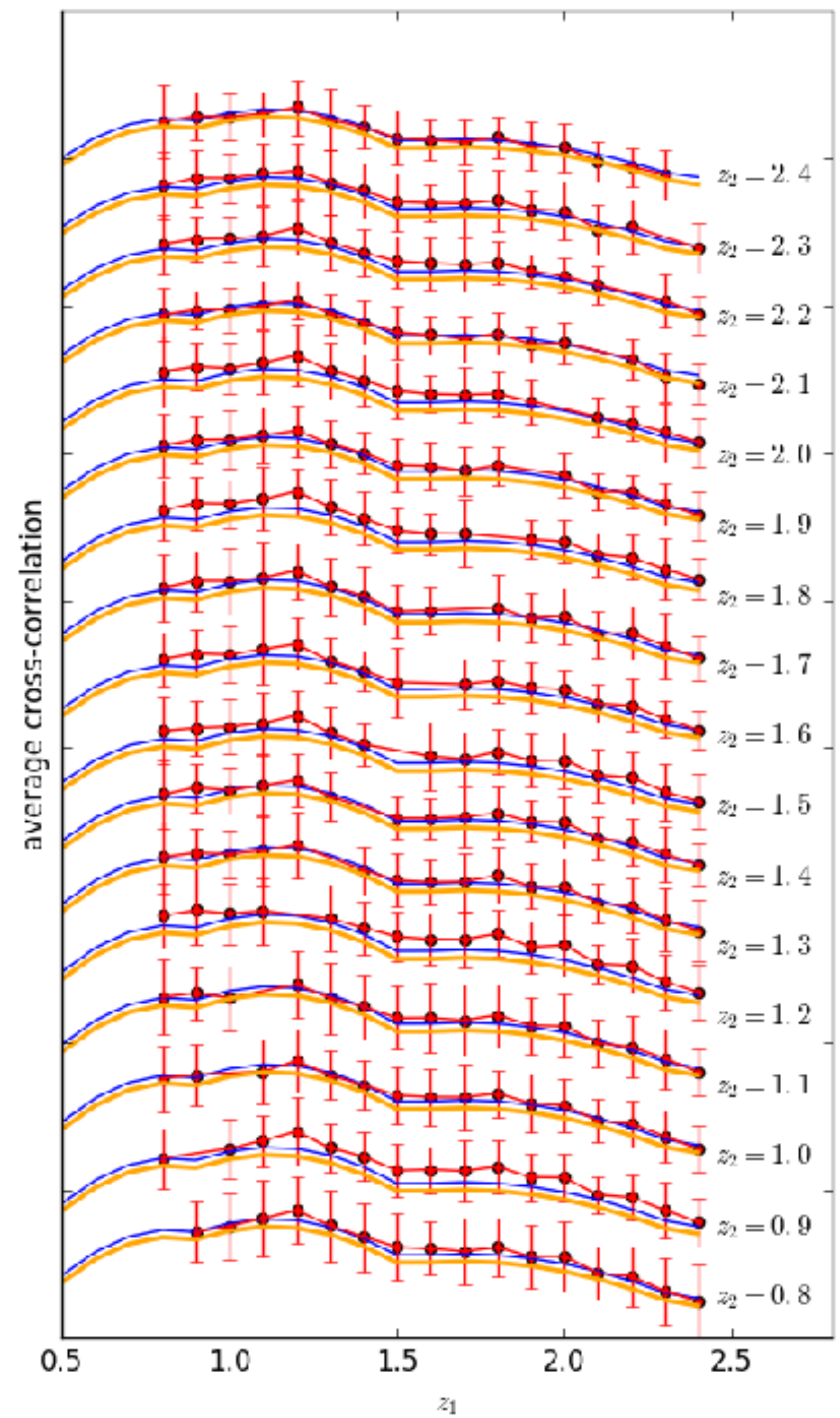
Can we distinguish between reddening models?



Can we distinguish between reddening models?







g

Main uncertainties:

- every "foreground", including observing biases, will get into the mask - we need to model the sum of different foregrounds
- catastrophic redshift errors will create features in cross correlations
- we need to take lensing properly into account - easy to model
- we need to assume a universal galaxy LF, environmental dependence will result as an effective foreground term.
- $B_1(z)$ and $B_2(z)$: need to propagate uncertainties of the galaxy LF
- we must know luminosity-dependent bias and propagate its uncertainty

...a more sophisticated statistical approach?

Two possible strategies to use these results:

Use cross-correlations as a diagnostic for foreground contamination

or try a more aggressive program:

☞ apply some correction for MW extinction

☞ model residual foreground mask from cross correlations

☞ create a random

☞ use it to measure the 2D clustering

   ☞ compute covariances using many mock catalogues