Systematic search for systematic bias in SN Ia data

Valerio Marra

in collaboration with
L. Amendola & M. Quartin

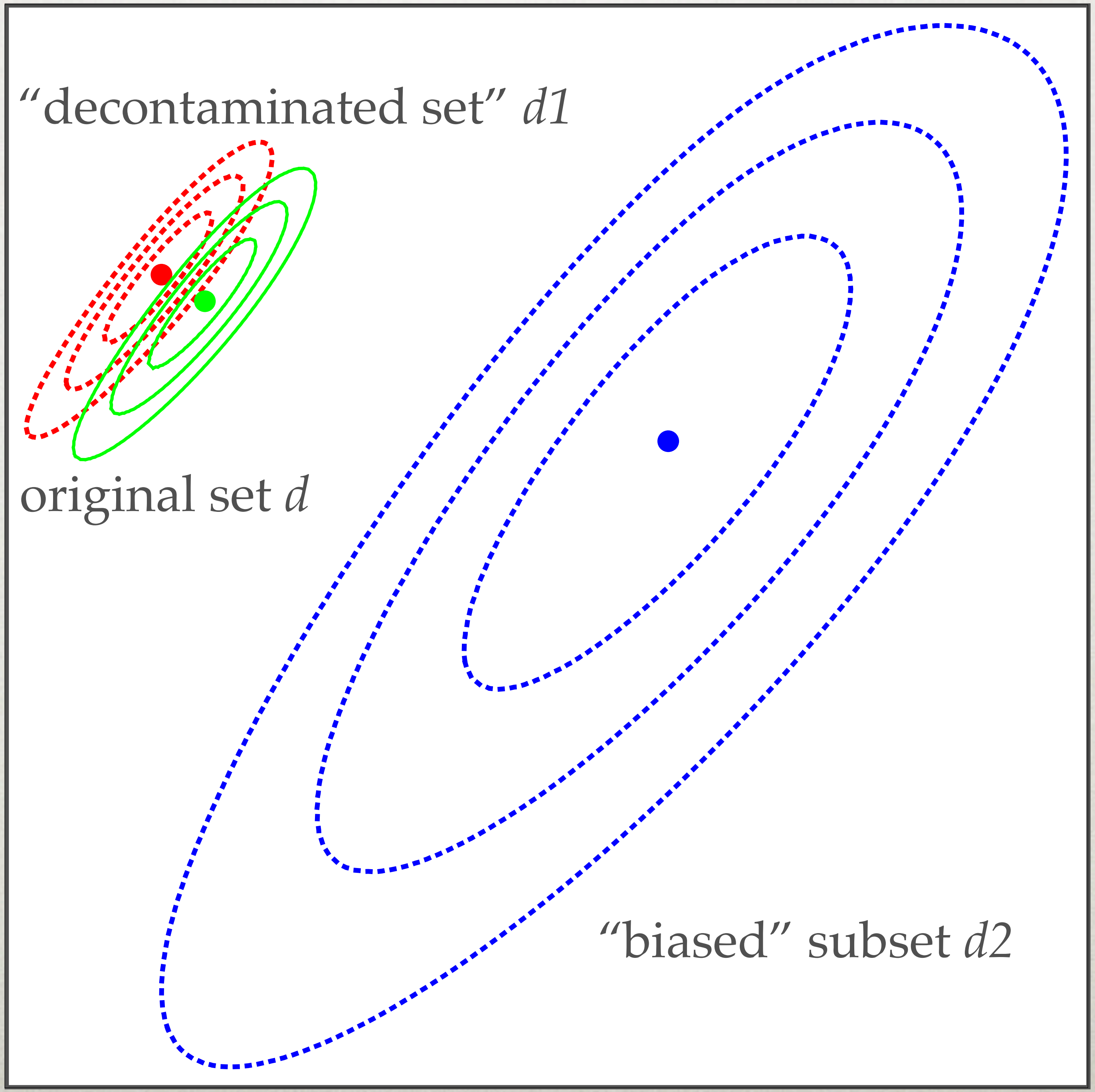
arXiv:1207.xxxx

Basic idea

look for "robustness" among subsets
in a systematic and blind way

$d1 \cup d2 = d$

d1 and *d2*
do not
overlap



Definition of (internal) robustness

$$\hat{R} = \frac{\text{experiment is internally robust}}{\text{experiment is } \textit{not} \text{ internally robust}}$$

$$= \frac{E_{\text{comb}}}{E_{\text{ind}}} = \frac{E(d; M_C)}{E(d_1; M_C)\, E(d_2; M_S)}$$

Robustness is an estimator "orthogonal" to FoM

see for more details
March, Amendola, Huterer & Trotta, MNRAS 2011

Bayesian evidence

$$E(\boldsymbol{x}; M) = \int f(\boldsymbol{x}; \theta_i^M) p(\theta_i^M)\, \mathrm{d}^n\theta_i^M$$

data

theoretical par.

likelihood

prior

SN magnitude

$$S_n = \sum_i \frac{m_i^n}{\sigma_i^2}$$

$$\chi^2 = S_2 - \frac{S_1^2}{S_0}$$

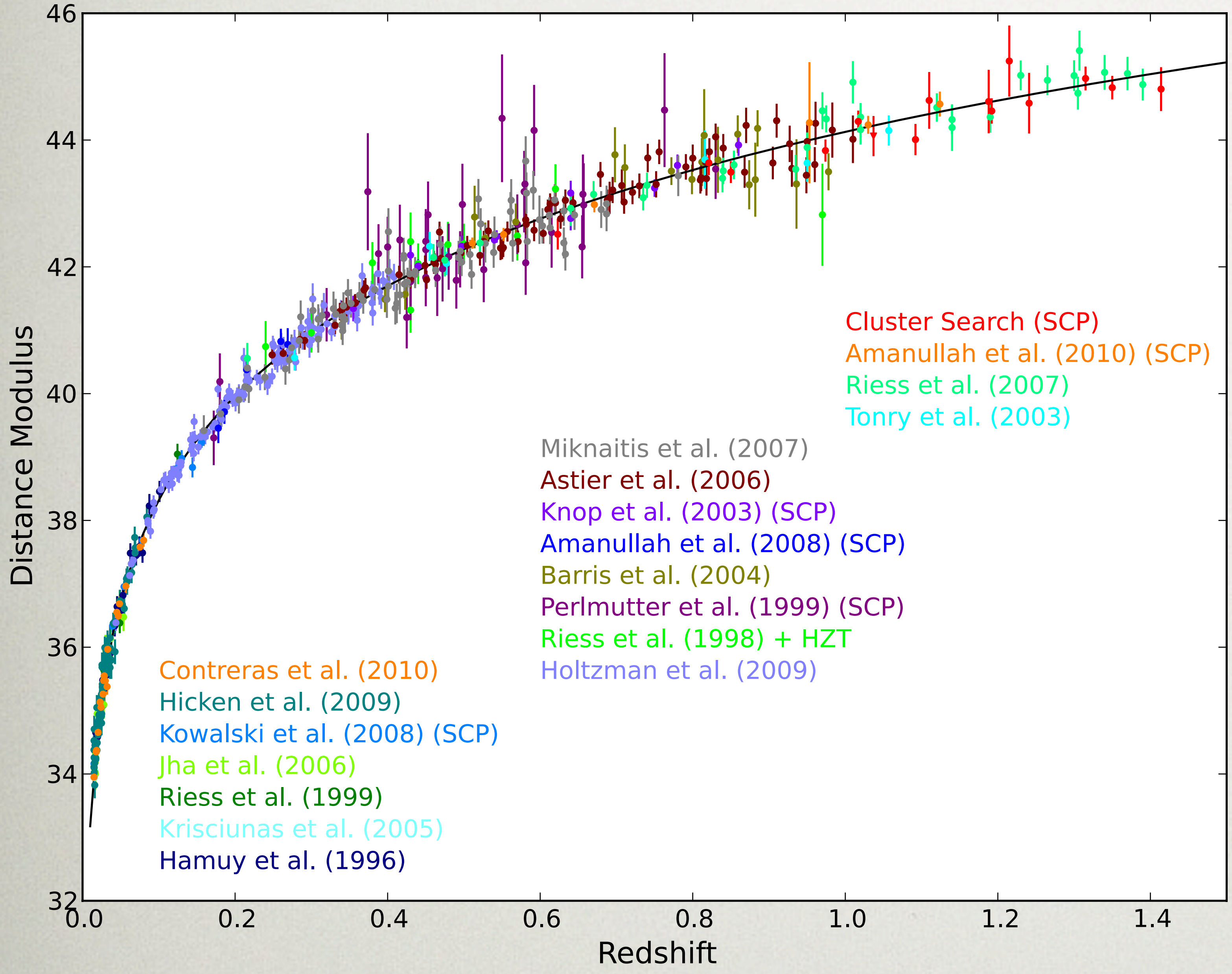
$$R \equiv \log \hat{R} = R_0 + \frac{1}{2} \log\left(\frac{S_{0,1} S_{0,S}}{S_0} \frac{|F_1 F_S|}{|F|} \right) - \frac{1}{2}(\chi_t^2 - \chi_1^2 - \chi_2^2)$$

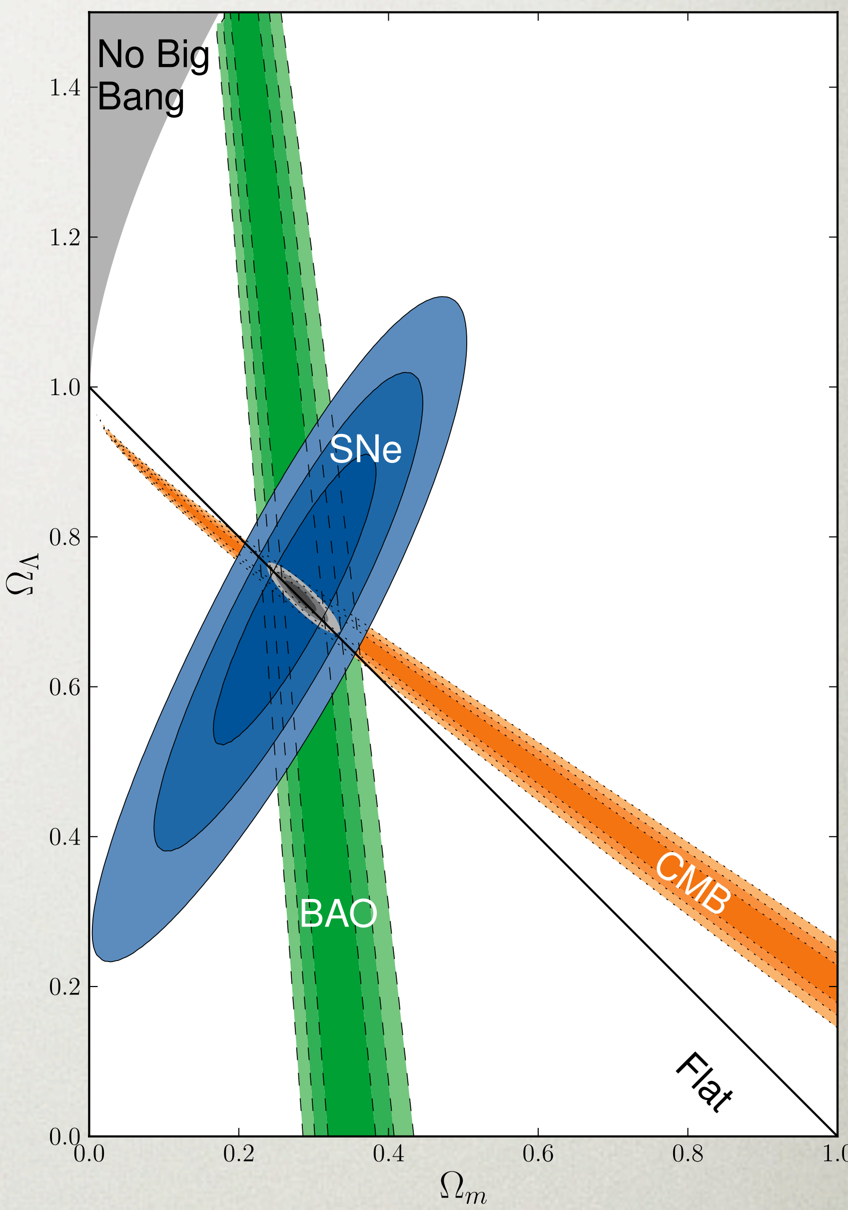
Fisher approximation

Occam's factor

degree of overlapping/compatibility

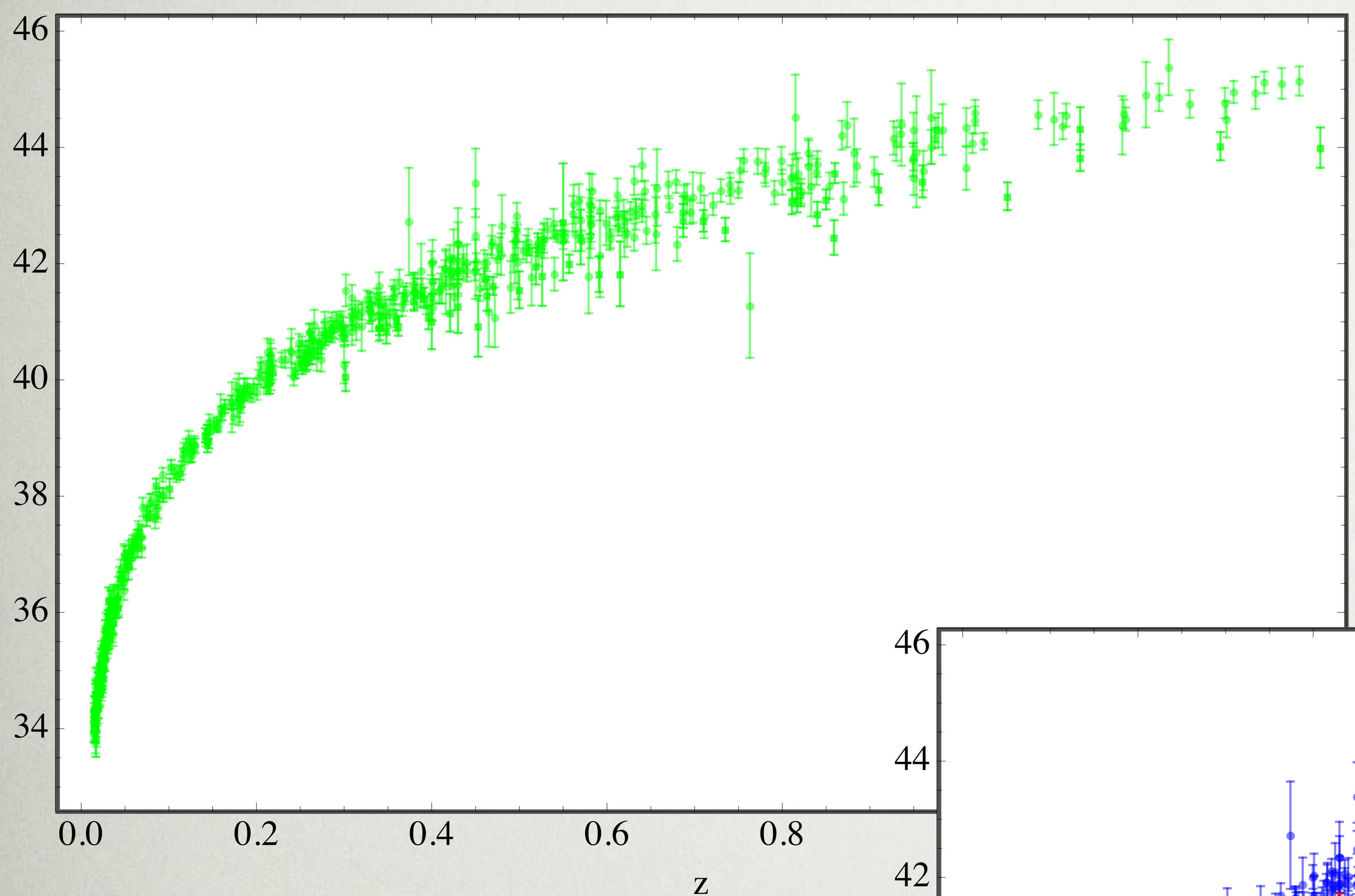
Union2.1 SNIa dataset

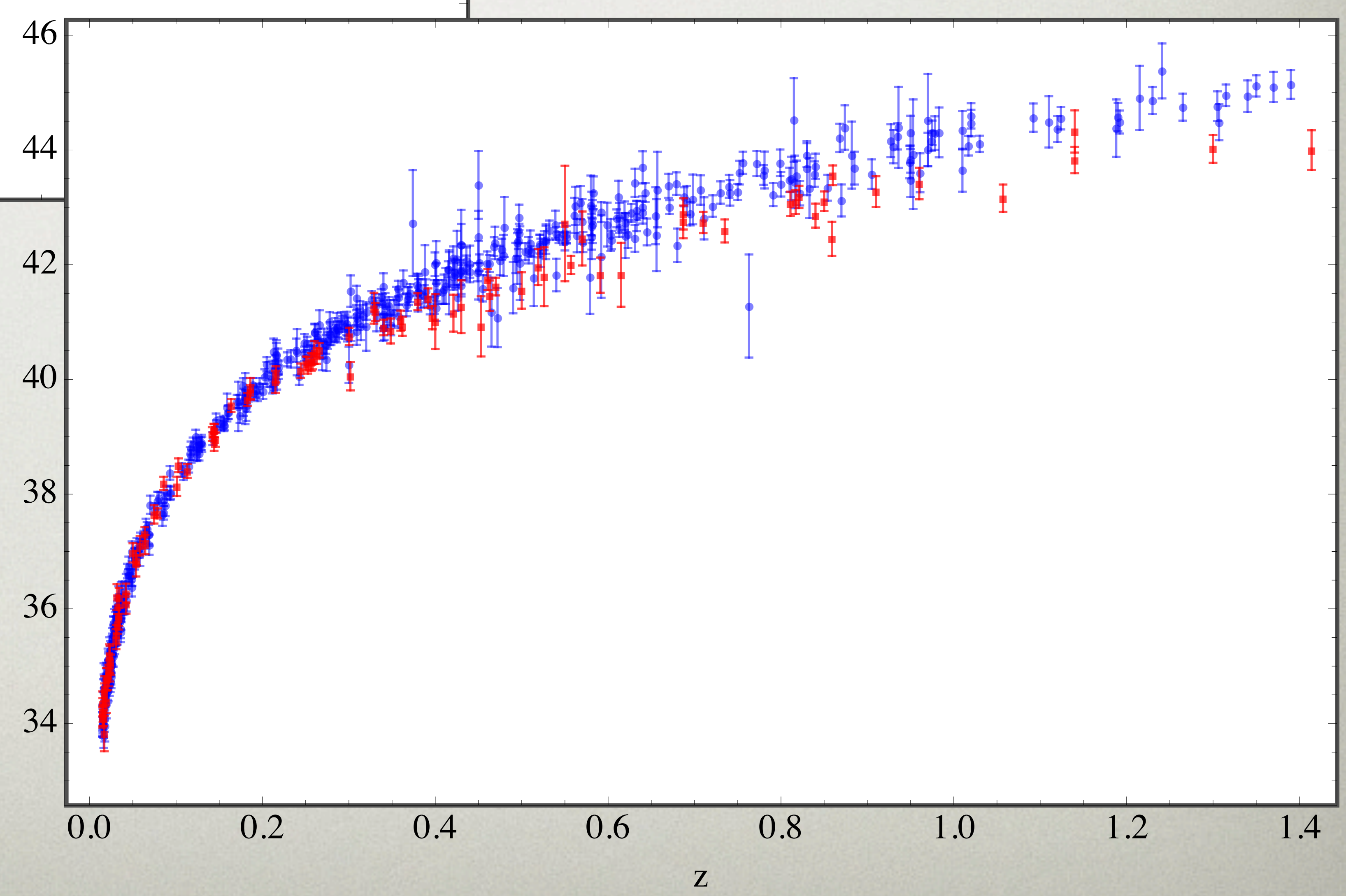




Suzuki etal, ApJ 2012

Try and find the biased subset





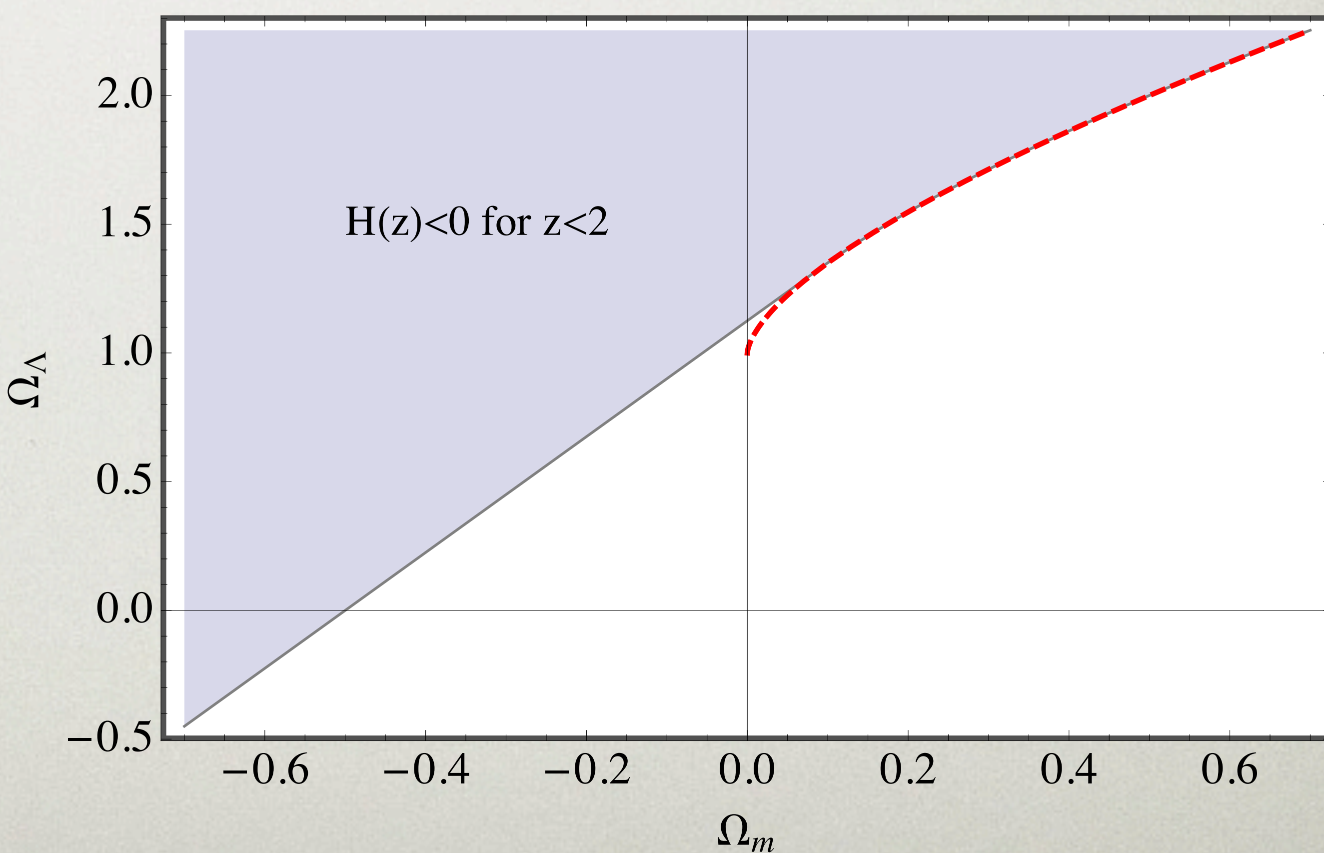
Systematic parameters

We use the cosmological parameters Ω_m and Ω_Λ to parametrize the (possibly cosmology unrelated) systematic parameters.

Consequently we consider a larger parameter space than the usual physical one

$$-10 < \Omega_m < 10$$

$$-20 < \Omega_\Lambda < 10.$$



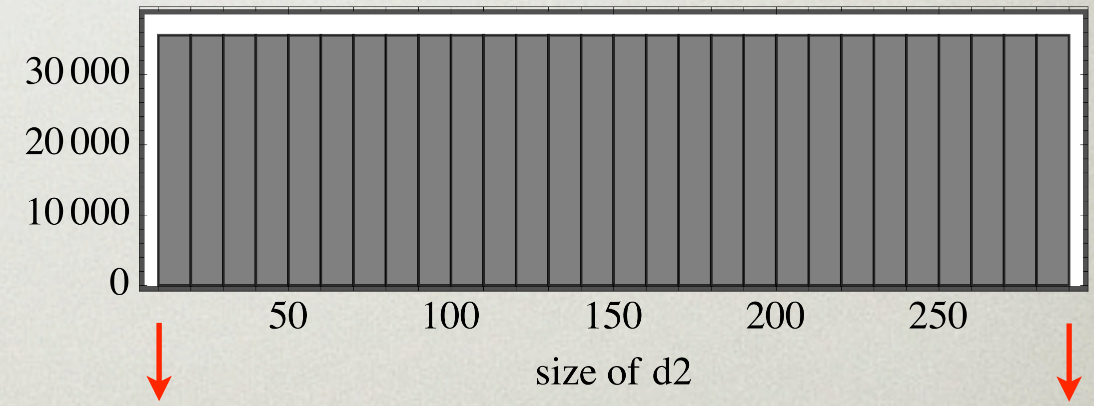
Scanning the subsets

The robustness is a statistical quantity, which has to be studied by building a PDF

There are $\approx 10^{0.3\,N} = 10^{174}$ possible partitions: a complete scan is impossible

We need to define a strategy Ξ in selecting the subsets, the PDF will depend on it.

Here we choose a uniform distribution in the size of d2

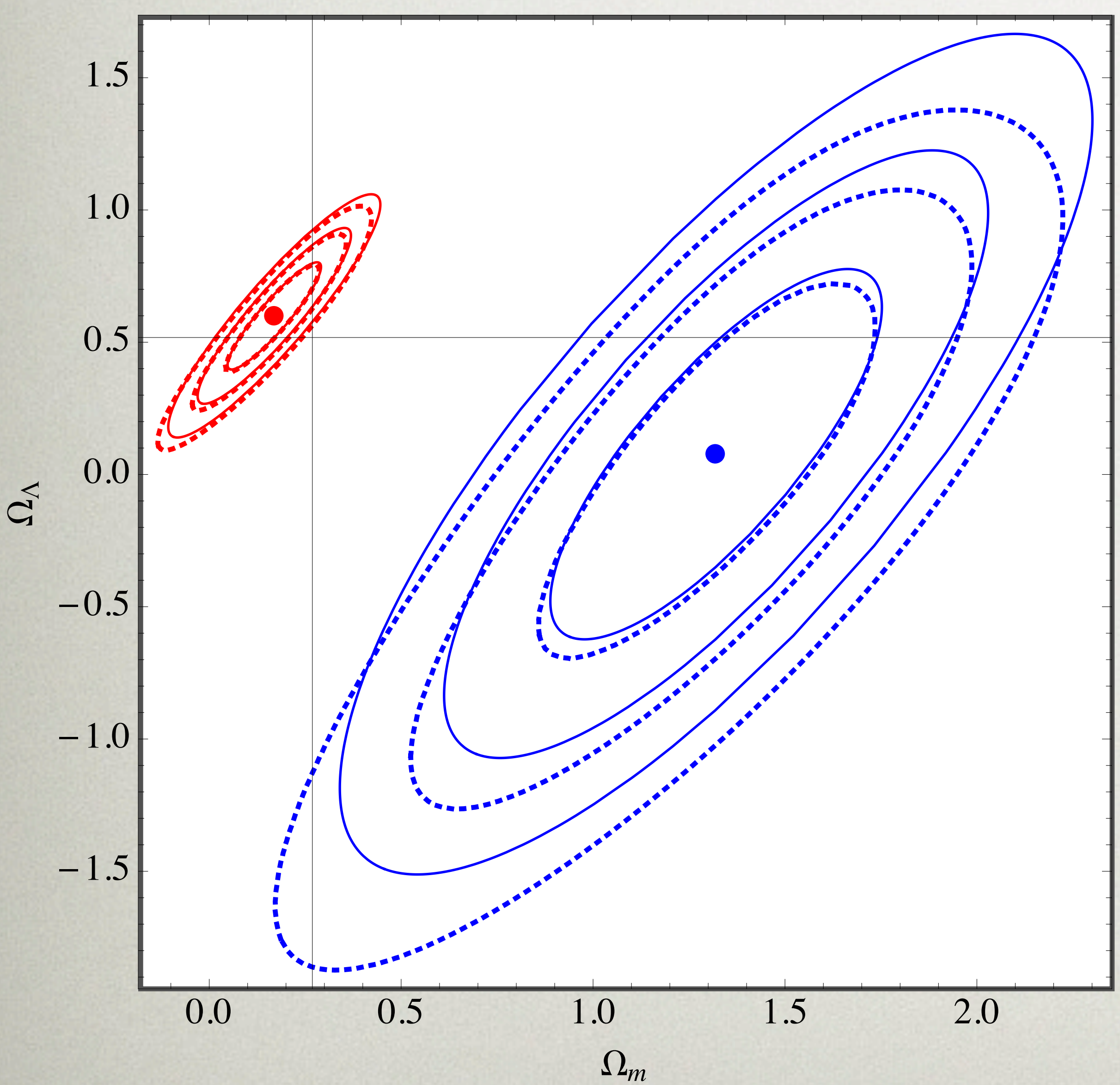


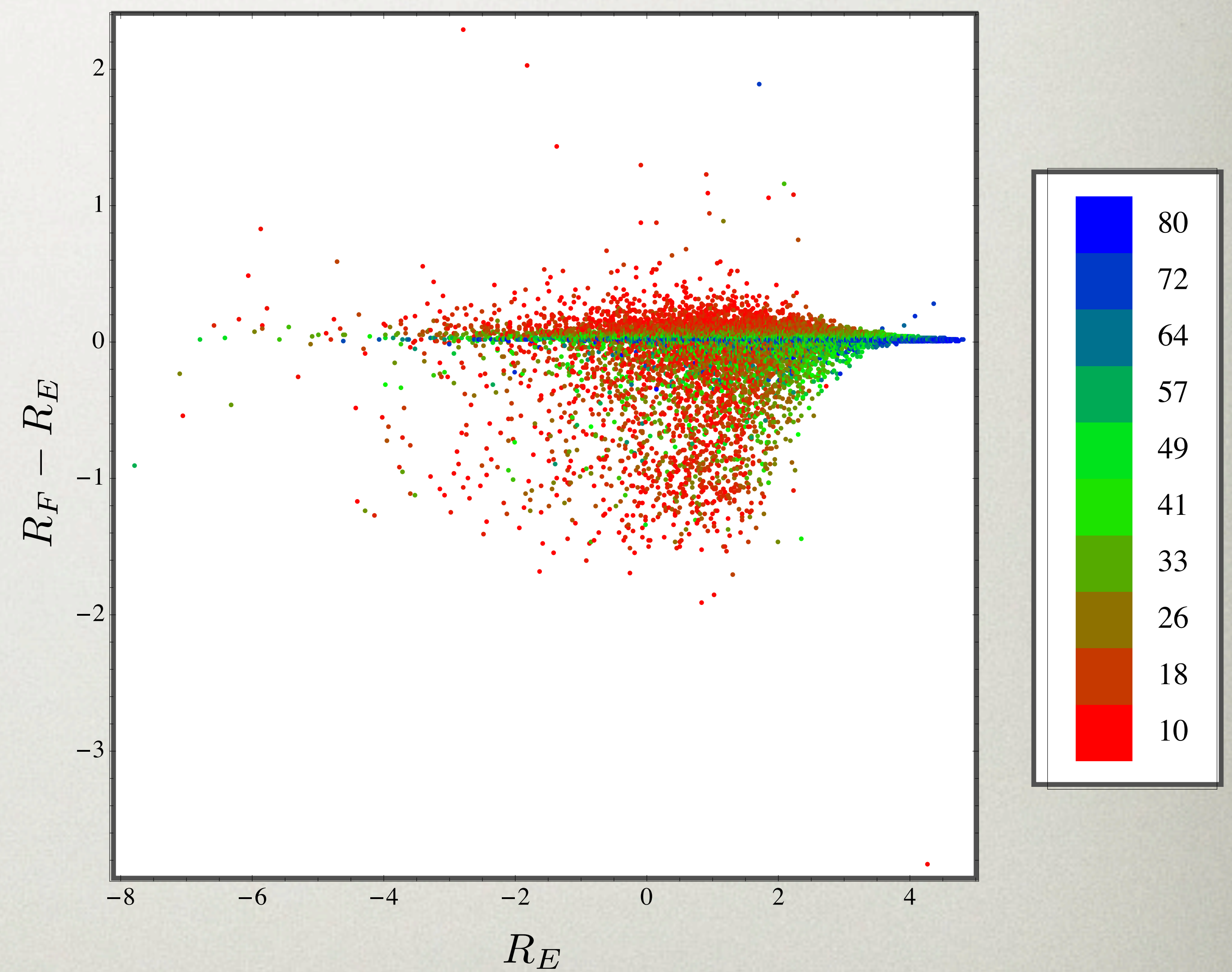
10
so that the likelihood of d2 has support within the parameter space considered

N/2=290
because of symmetry

Fisher or not to Fisher...

$$\hat{R} = \frac{E(d; M_C)}{E(d_1; M_C)\, E(d_2; M_S)}$$





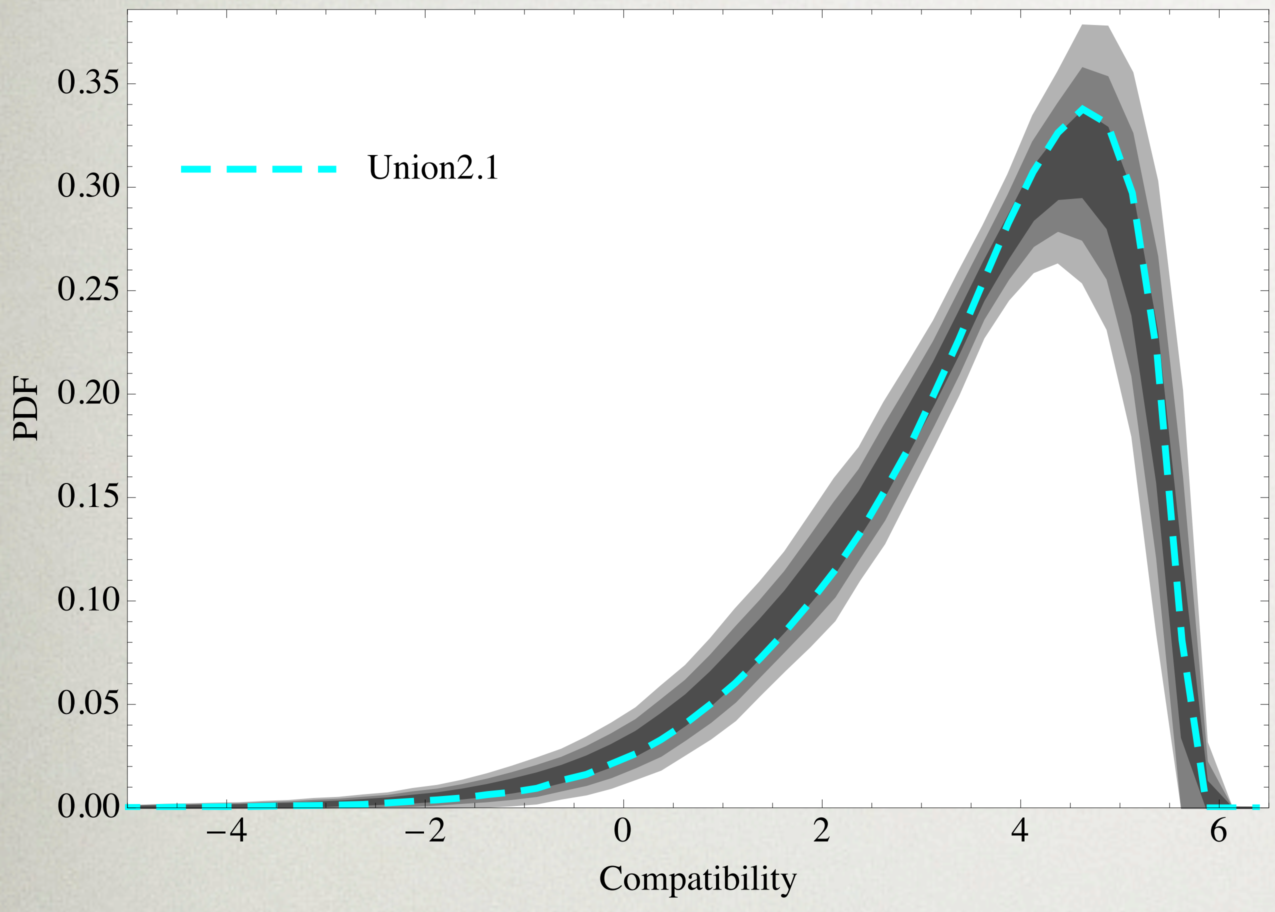
We switch to Fisher for size of d2 > 90

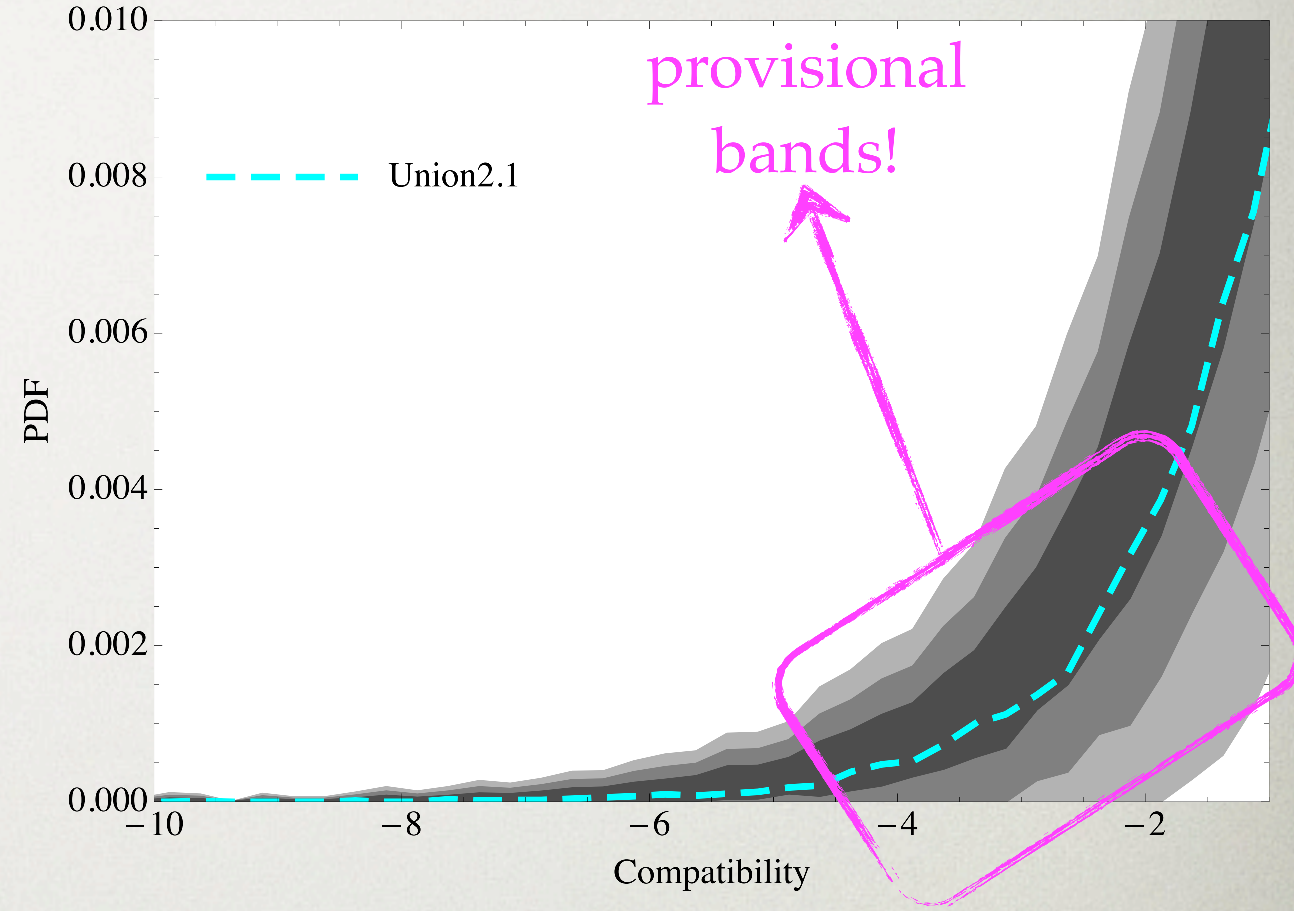
Necessity for mock unbiassed catalogues

$$R = R_0 + \frac{1}{2}\log\left(\frac{S_{0,1}S_{0,S}}{S_0}\frac{|F_1 F_S|}{|F|}\right) - \frac{1}{2}(\chi_t^2 - \chi_1^2 - \chi_2^2)$$

the full R-PDF is a highly nontrivial object: we need to compare with mock unbiased data, which we generate randomizing the magnitudes only, with the best-fit model of Union2.1 as fiducial model

(very preliminary!) results: Union2.1





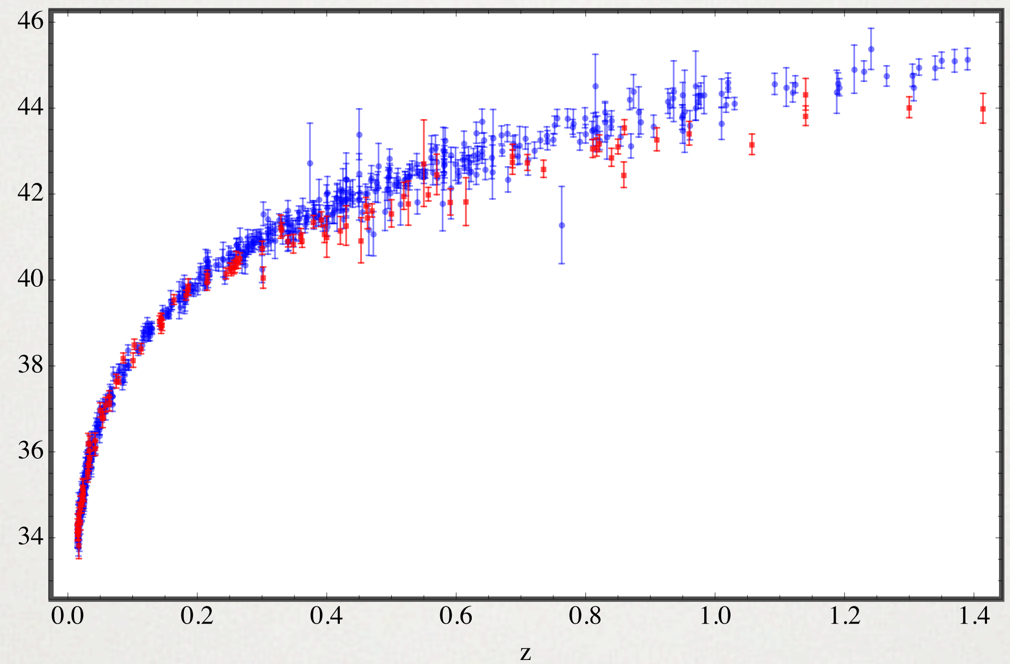
- Union2.1: 500k statistics.
- Mocks: 100 catalogues from which the (100) PDFs are generated with 30k statistics

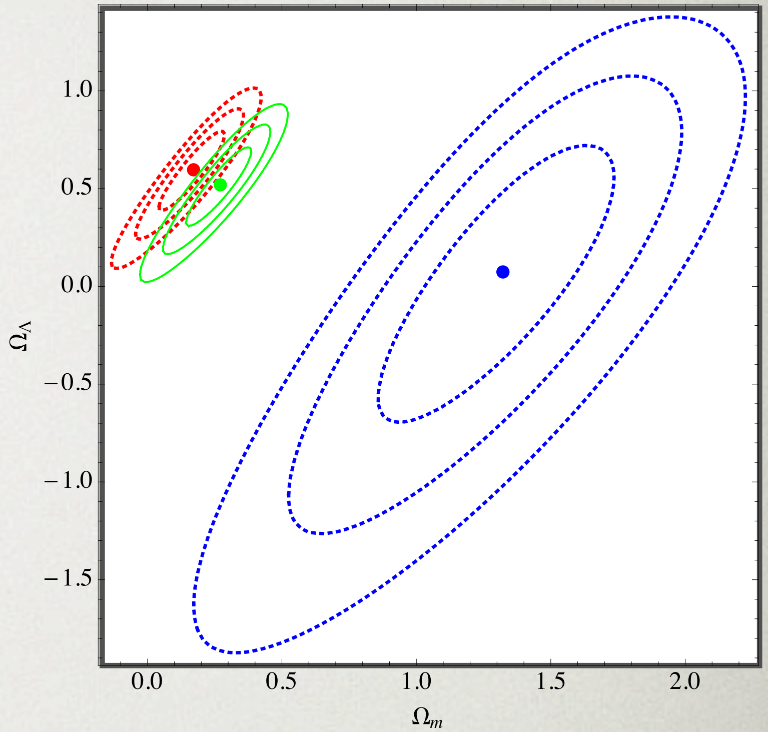
For the comparison with mocks, the "cosmology" fluctuations are more important than the "sampling" fluctuations

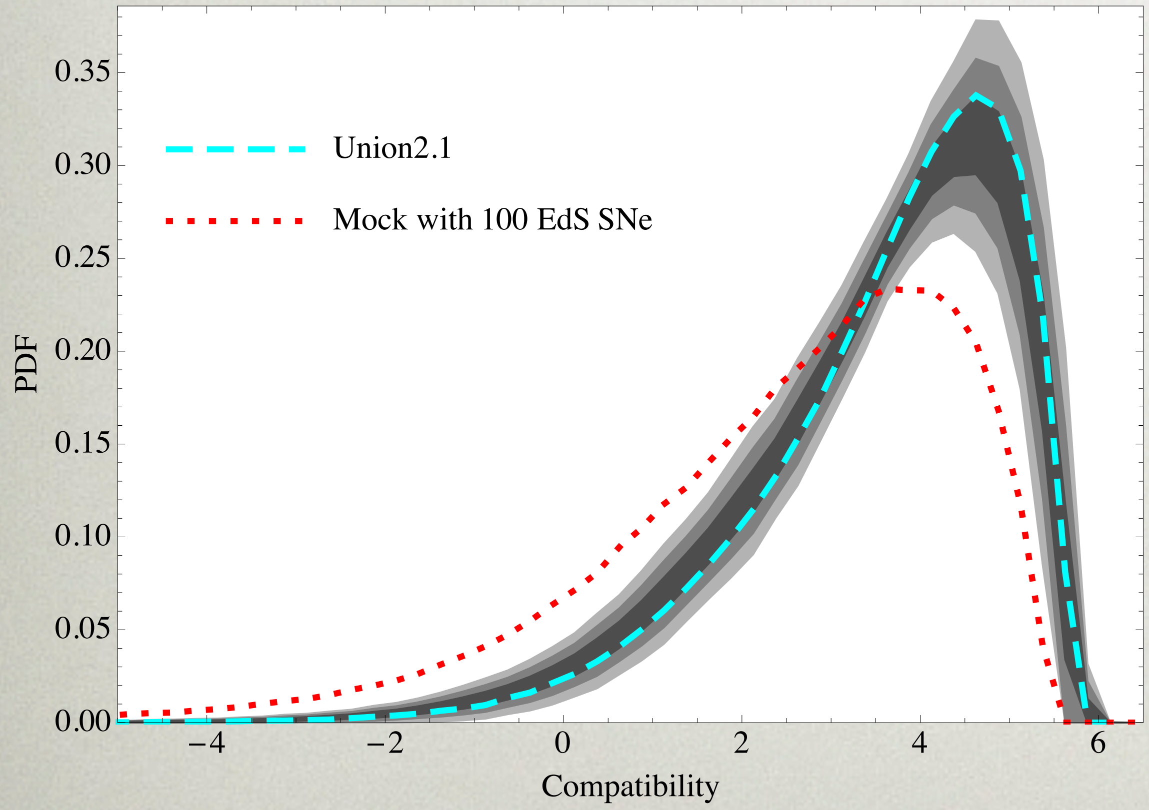
Average CPU time per partition is ~3 s. Luckily it can be easily parallelized

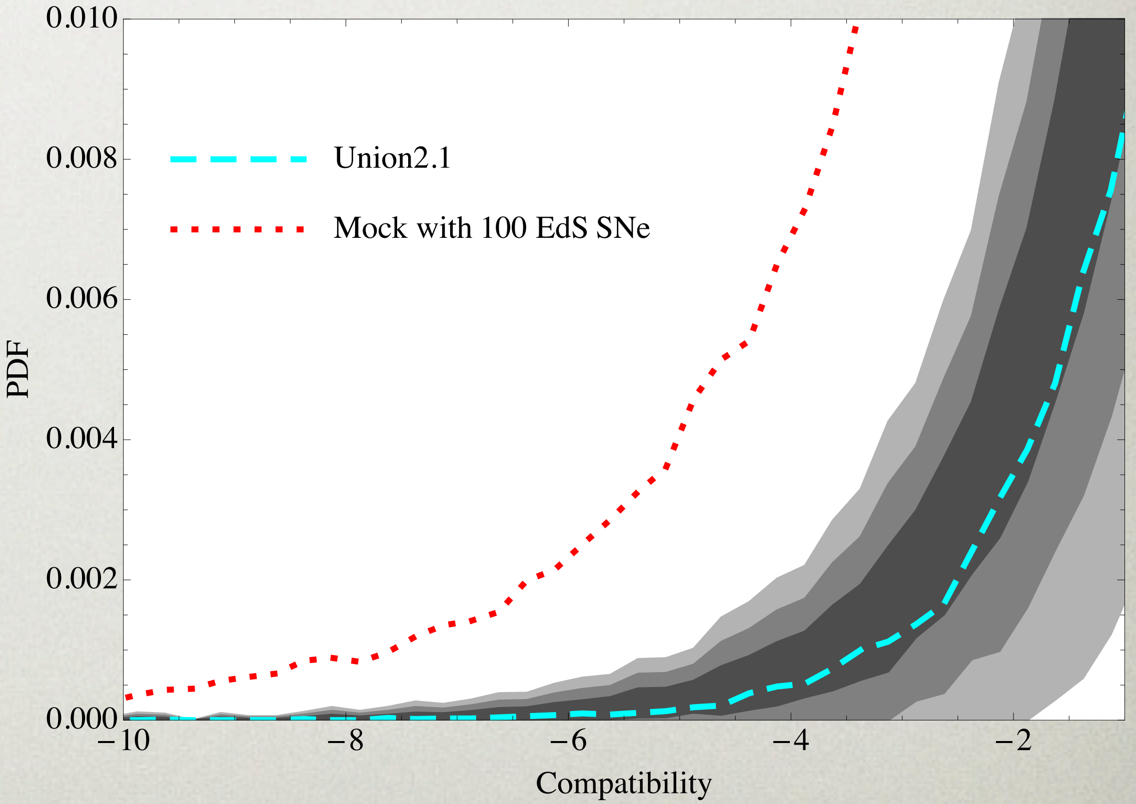
Biased mock (1)

Mock catalogue with 100 SNe drawn from EdS as fiducial model (and not the best-fit model of Union 2.1)



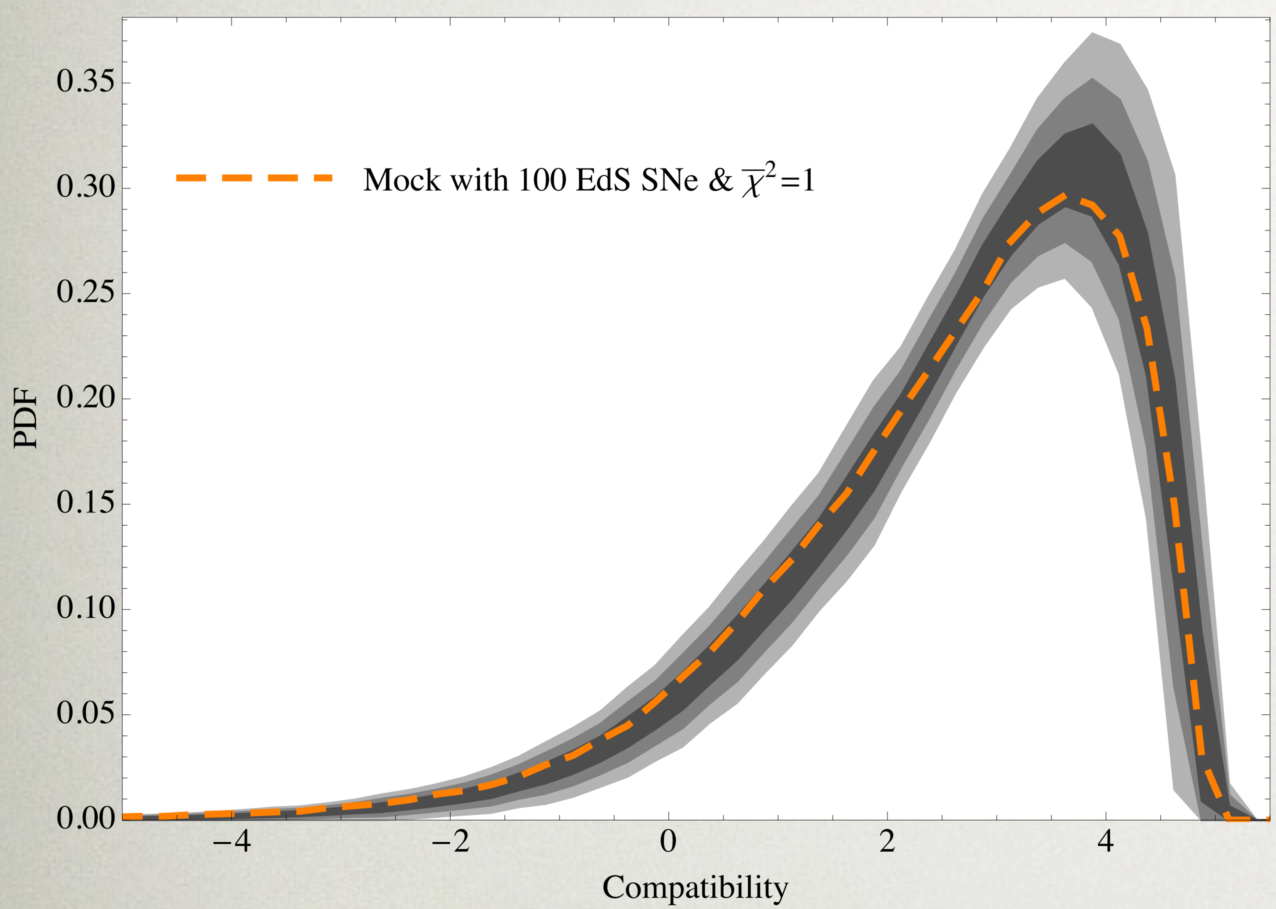


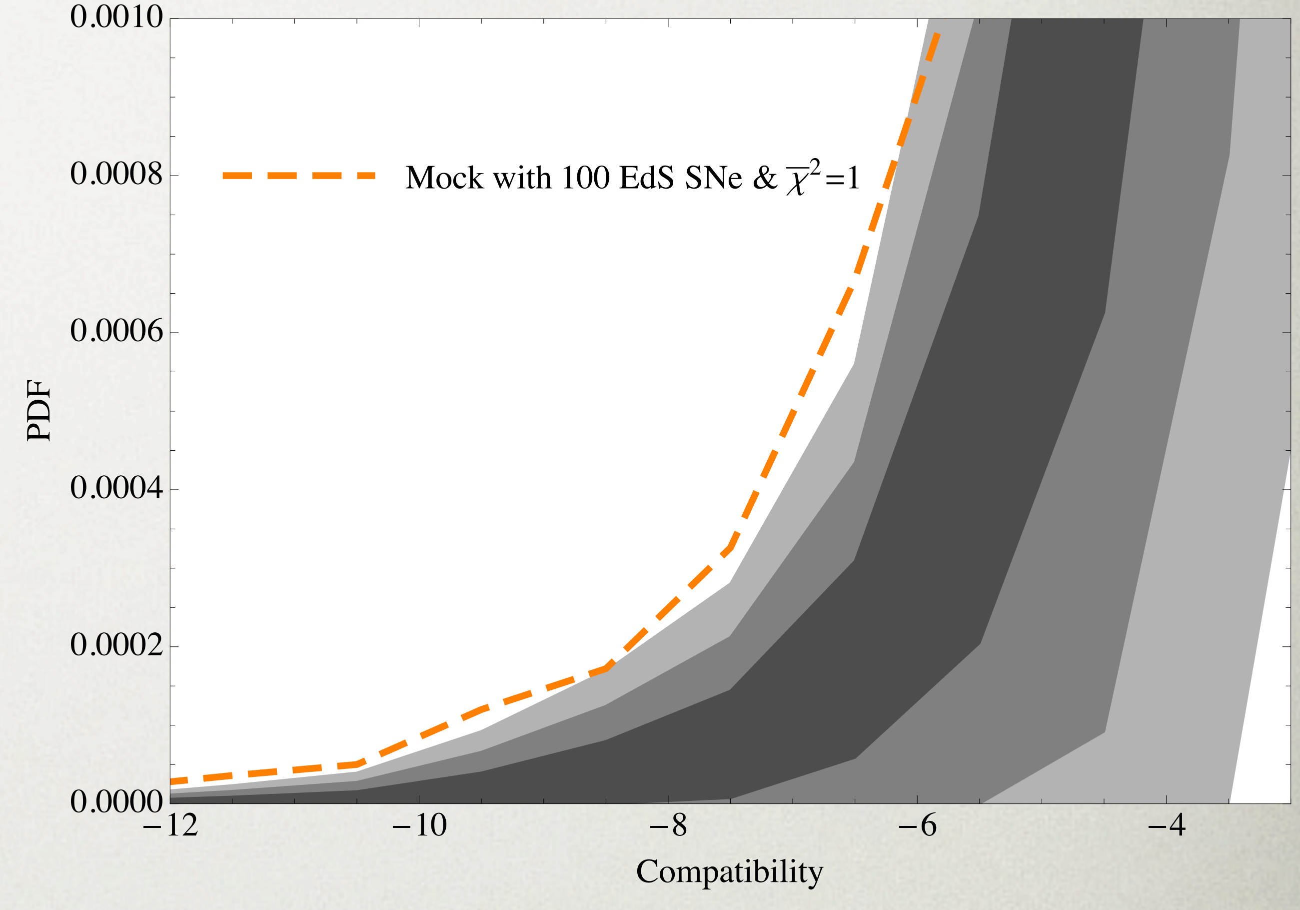




However, $\bar{\chi}^2 = 1.40$: can we still see the signal if we impose $\bar{\chi}^2 = 1$?

Biased mock (2)

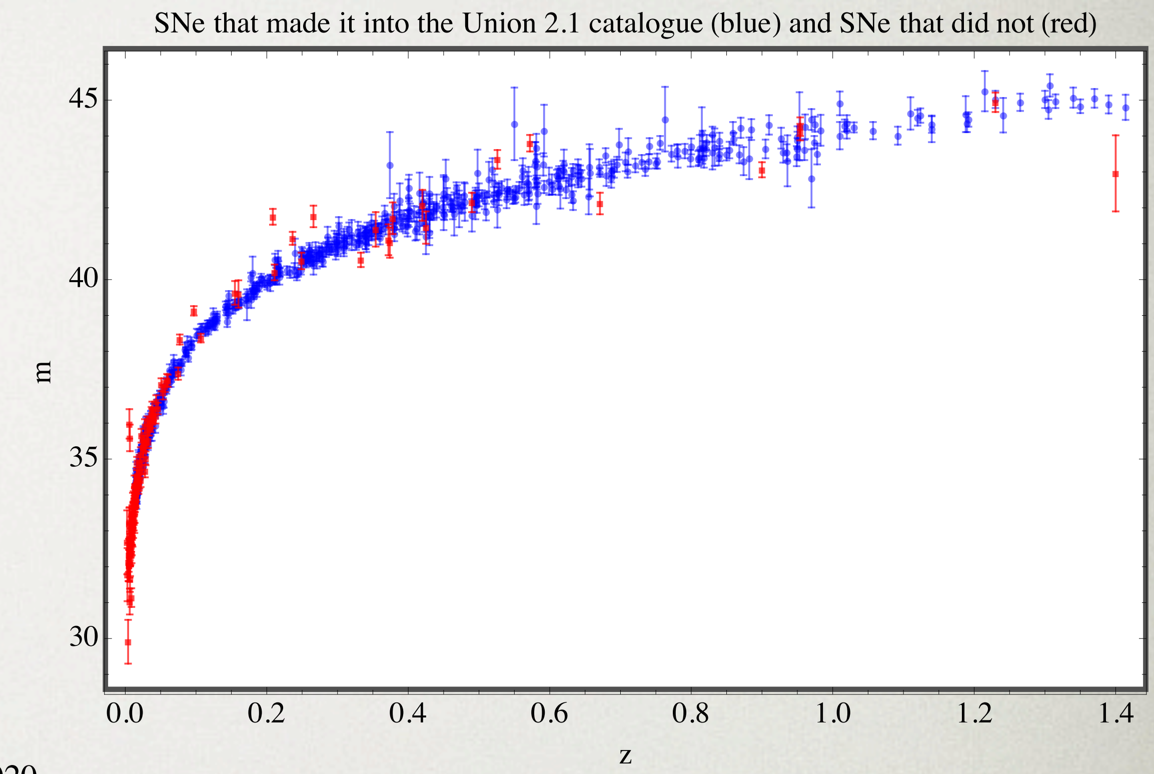


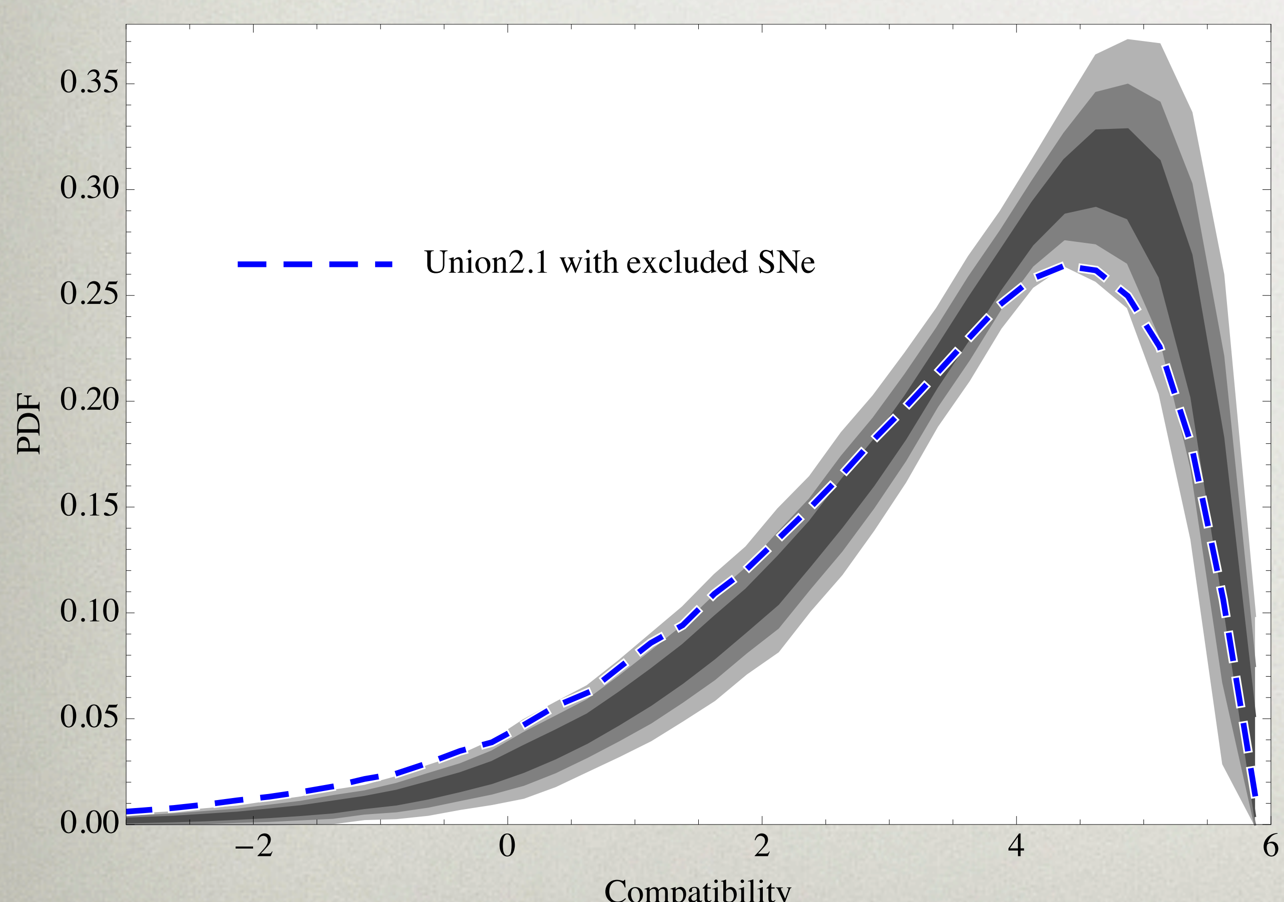


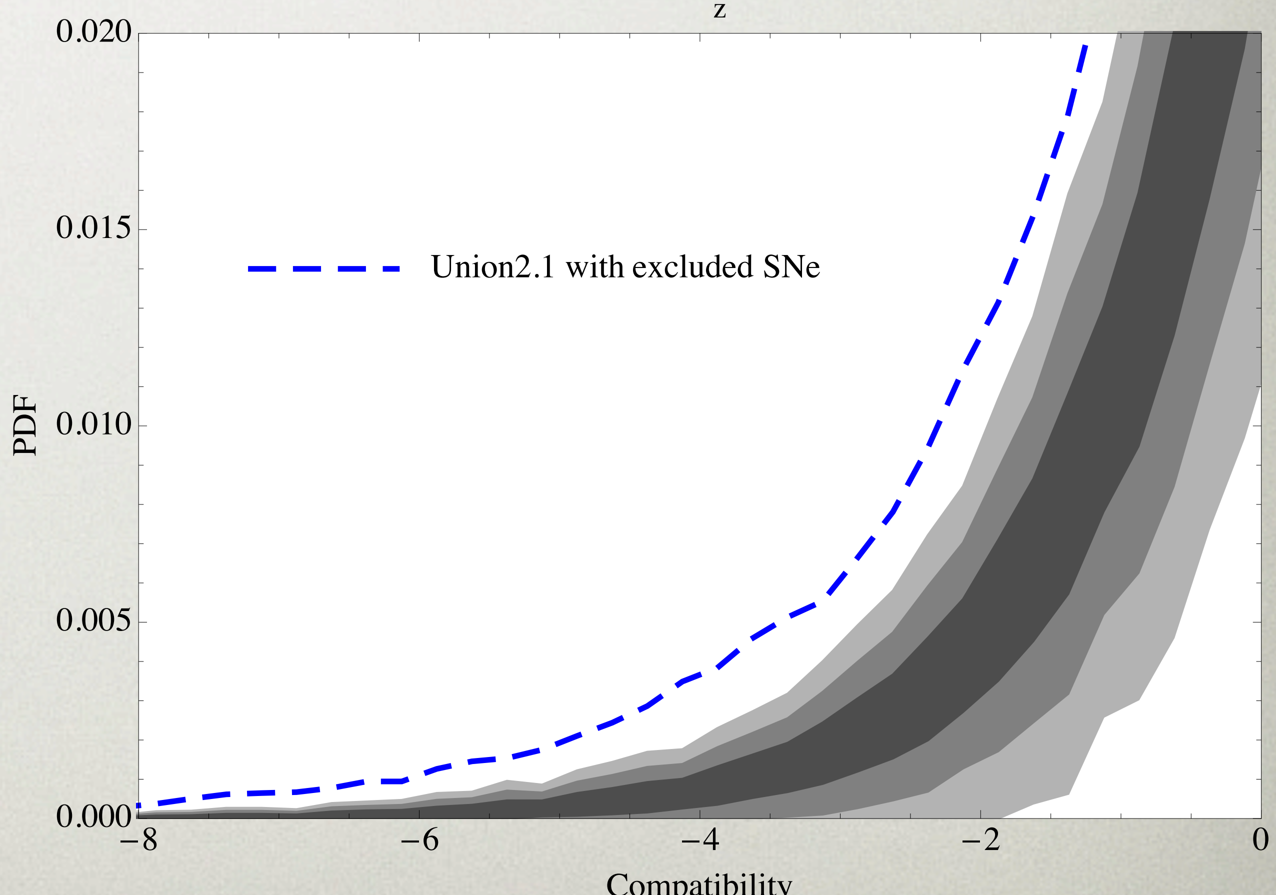
It looks as we can still detect the signal, which is now focused on the low-robustness tail of the PDF. In order to assess the significance of the signal we have to properly treat the bin statistics of the mocks... **coming soon!**

Union2.1 including previously excluded SNe (1)

135 SNe did not pass the quality cuts and make it into the final Union2.1. Can we confirm that these SNe are dominated by systematics?

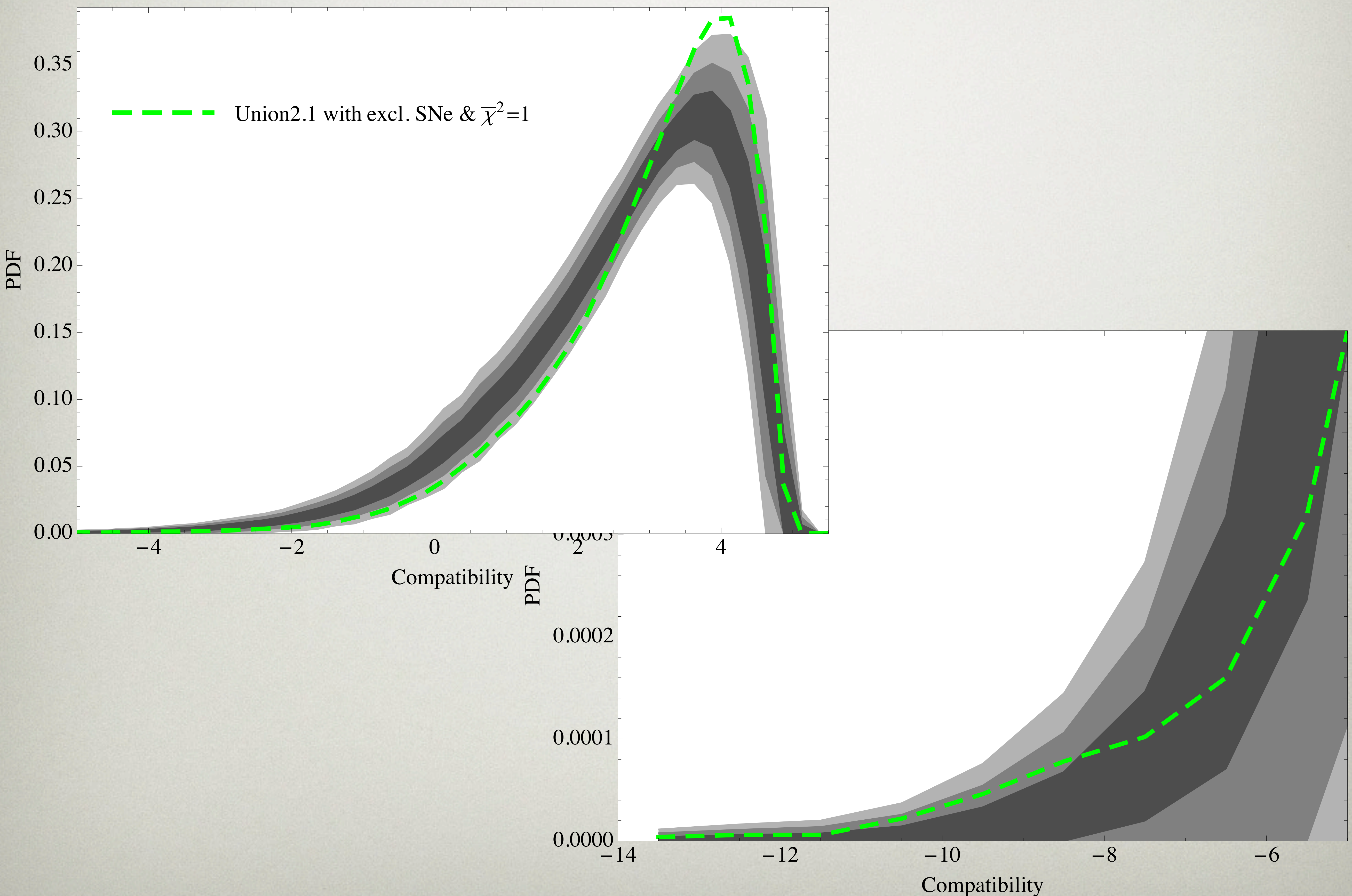






Again, $\bar{\chi}^2 = 1.73$: can we still see the signal if we impose $\bar{\chi}^2 = 1$?

Union2.1 including previously excluded SNe (2)



Next...

- as the signal is clearly in the low-robustness tail, try another scanning strategy so as to focus on the tail
- try to speed up the calculation of R
- try to find part of the systematics-driven SNe
- apply to other datasets
- use different sets of parameters for cosmology/systematics
- apply to (many) other observables
- ...

THANKS