Cosmology with clusters of galaxies: from *Planck* to the Simons Observatory

Iñigo Zubeldia
Paris-Saclay Astroparticle Symposium
19/11/2024

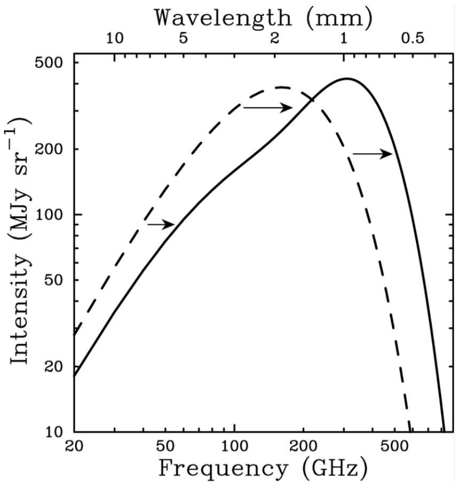
Outline of the talk

1. Introduction.
2. tSZ cluster finder: **`SZiFi`**.
3. Application to *Planck* data: *Planck* **`SZiFi`** catalogues and cosmological analysis.
4. Cluster number count likelihood code (**`cosmocnc`**).
5. Simulation-based inference (SBI) for cluster cosmology.

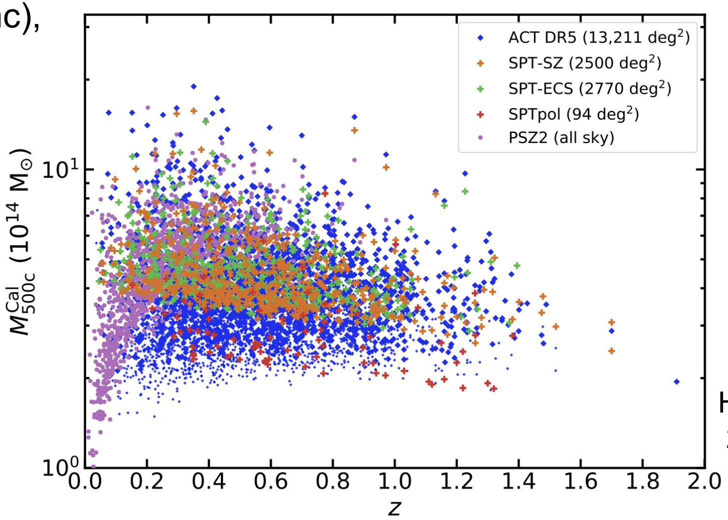
Towards precision tSZ cluster cosmology

- Number of clusters as a function of mass and redshift is a powerful cosmological probe (Ω_m, σ_8, w, Σm_ν).
- Clusters can be detected in X-ray, optical and mm.
- mm: Thermal Sunyaev-Zeldovich (tSZ) effect allows for cluster detection to high redshift with CMB data.
- These catalogues can be in turn used for cosmology (cnc), with the cluster tSZ signal as a mass proxy.
    - *Planck*: ~ 400 (SNR >6).
    - SPT-SZ/SPTpol: ~ 1000.
    - ACT: ~ 4000.

    - SPT-3G: ~ 20,000.
    - SO: ~ 20,000.
    - CMB-S4: ~100,000.



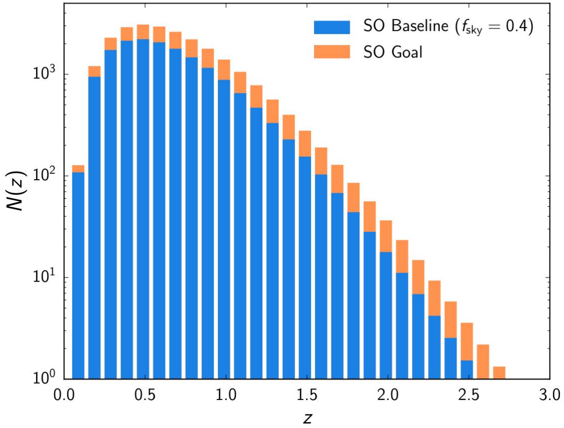
Sunyaev & Zeldovich 1970

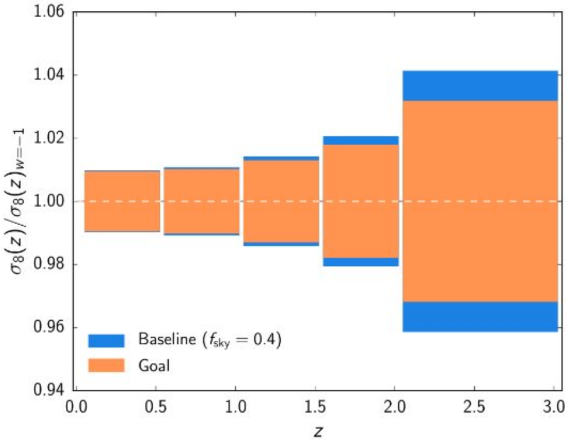


Hilton+ 2021

Towards precision SZ cluster cosmology

(Simons Observatory Collaboration forecast paper 2018)





Full analysis pipeline must be understood to unprecedented level if the statistical power of clusters from upcoming experiments such as SO is to be realised:

- Cluster detection: **SZiFi**.
- Accurate cosmological inference tools: **cosmocnc** (likelihood) and **SBI**.
- Calibration of scaling relations ("mass calibration").

`SZiFi`: the Sunyaev-Zeldovich iterative Finder

- Goal: carefully study systematics in cluster tSZ detection, tackle them if necessary, apply to *Planck* and SO.

- New implementation of multifrequency matched filter (MMF) cluster-finding approach (used in *Planck*, ACT and SPT).

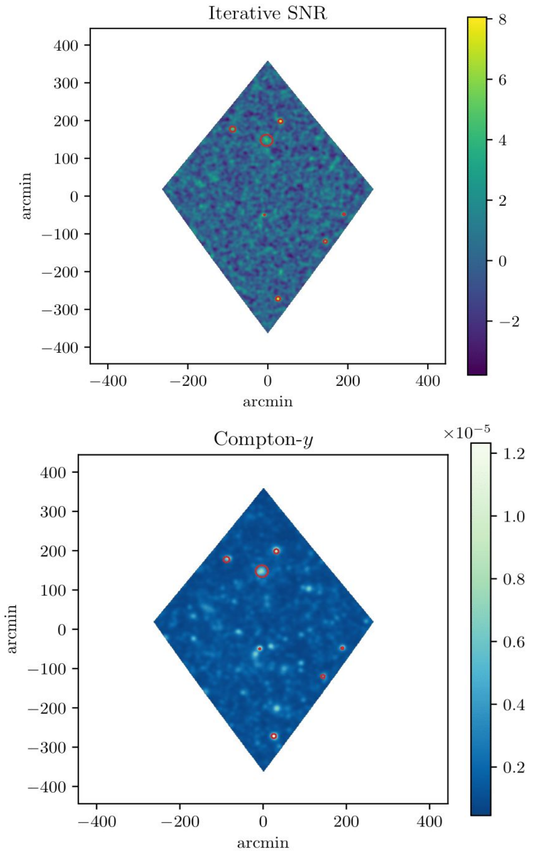
- Searches for clusters making use of both spectral and spatial information (tSZ SED and cluster Compton-y profile).

- Main observable: signal-to-noise (mass observable).

- Key novel features:

    - **Iterative noise** covariance estimation: boosts SNR and removes significant ILC bias from cluster observables.
    - **Foreground spectral deprojection** (e.g., CIB): can effectively remove bias from cluster-correlated CIB emission (with moment expansion). Analogous to cILC.
    - Can incorporate **relativistic corrections** to tSZ SED.

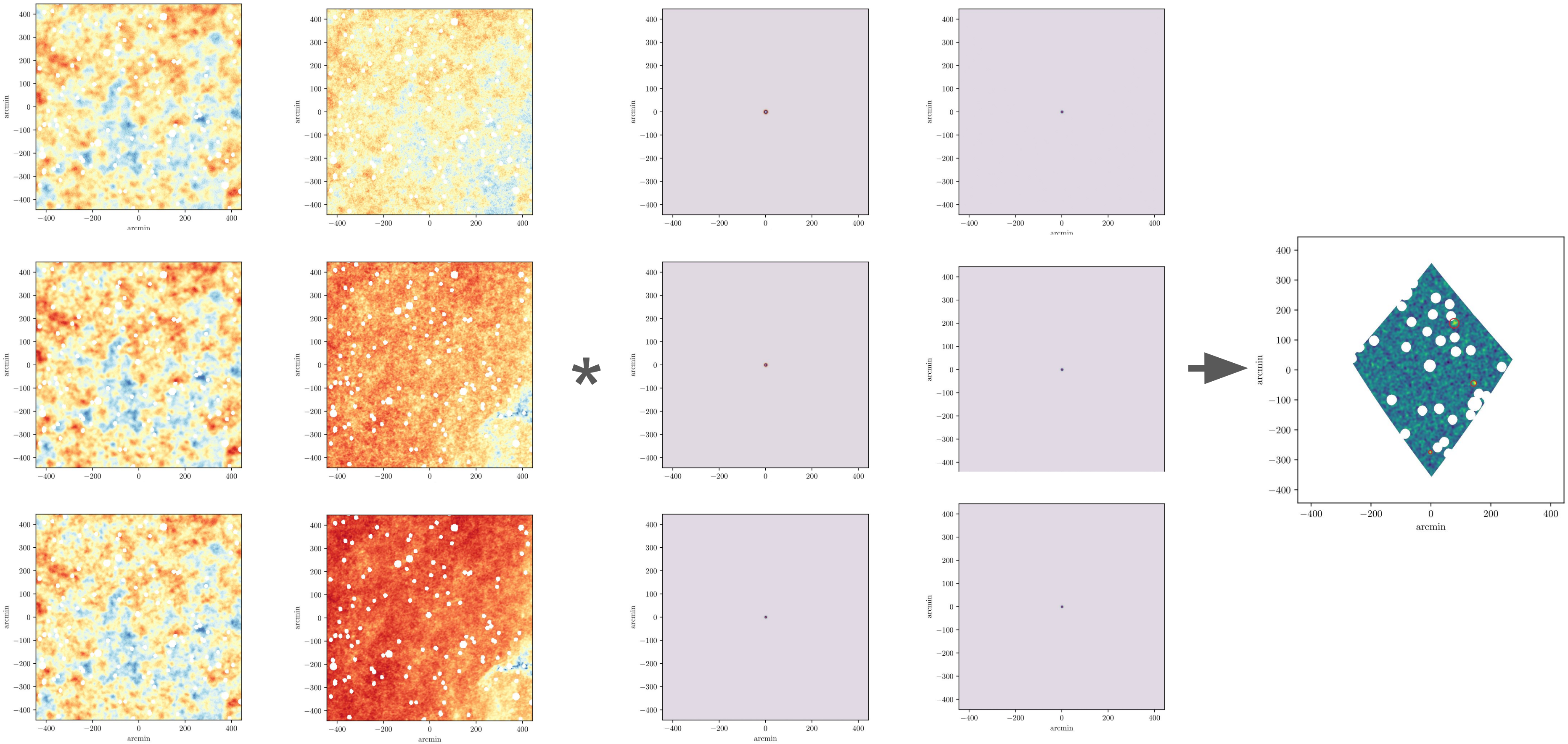
- Tested extensively on simulated *Planck* data, now also on SO.



IZ, Rotti, Chluba & Battye, 2204.13780
IZ, Chluba & Battye, 2212.07410

`github.com/inigozubeldia/szifi`

`SZiFi`: the Sunyaev-Zeldovich iterative cluster Finder



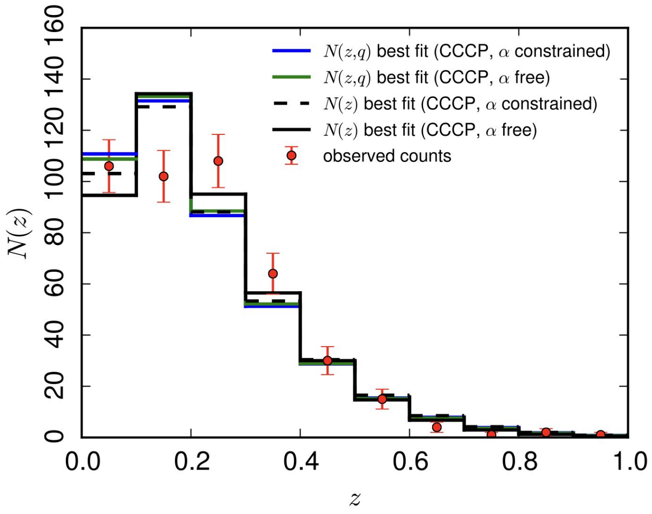
github.com/inigozubeldia/szifi

Planck cluster cosmology: and end-to-end reanalysis

with Jean-Baptiste Melin, Jens Chluba, Richard Battye,
Joe Mohr, Sebastian Bocquet, Aditya Singh, Anthony Challinor, Boris Bolliet

Goal: **improved *Planck* cluster catalogues**, **cosmology** (and learning for SO!)

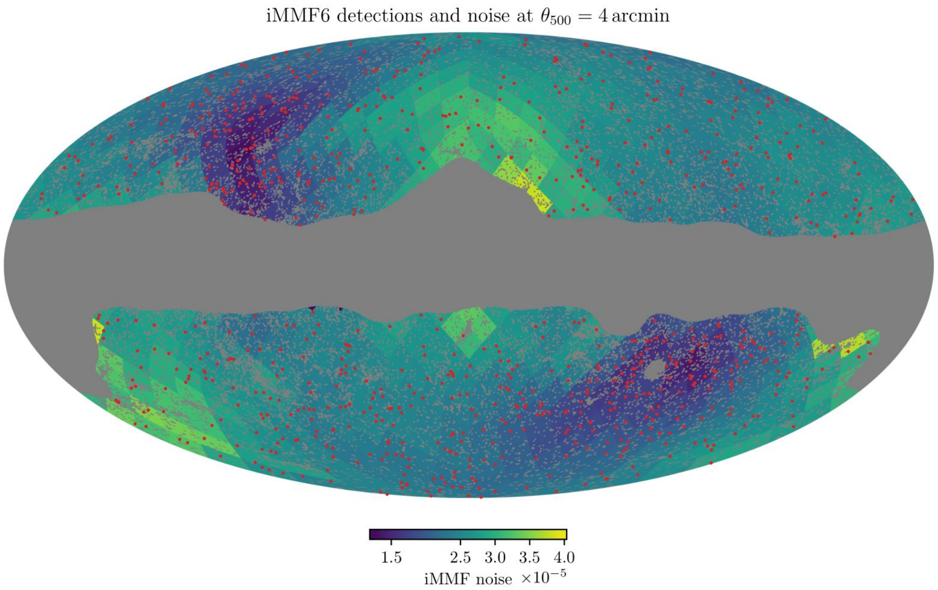
- Official Planck analysis down to SNR = 6 (439 clusters). We believe we can go down to **SNR = 5**.
- Catalogue obtained with all 6 HFI channels (100-857 GHz). No consistency checks between channels and **no assessment of foreground contamination**, e.g., from CIB.
- Noise covariance estimated non-iteratively.
- Cosmological analysis done in an "old-fashioned" way, calibrating the SZ-mass scaling relation with X-ray and lensing data separately. Potentially affected by selection effects (e.g., IZ & Challinor 2019). **We want to forward model all the observables (SZ, lensing)**.
- Poor goodness of fit.



Planck SZiFi cluster catalogues

with Jean-Baptiste Melin, Jens Chluba, Richard Battye

- New set of **10 tSZ-detected *Planck* cluster catalogues**, obtained with **SZiFi**:
  - Spanning 3 *Planck* frequency channel combinations.
  - Deprojecting CIB (7 catalogues, 4 with moments).
- Baseline catalogue, iMMF6: 833 detections with SNR > 5.
- **Good consistency** between different channel and deprojection combinations: first time checked for *Planck*.
- Impact of the **cluster-correlated CIB** on catalogues **probably negligible**: first time checked.
- Impact of the **relativistic corrections negligible** for main cluster observable, SNR: first time checked.



IZ, Melin, Chluba & Battye, 2408.06189

github.com/inigozubeldia/szifi

New *Planck* cluster count cosmological analysis

Use SZiFi catalogues down to SNR 5.

Empirical calibration of SZ-mass scaling relation ("mass calibration"):

- ***Planck* CMB lensing** mass calibration for most of the clusters (with Anthony Challinor). Total SNR about 5-6.

- **Optical weak lensing** profiles from DES Y3 for 165 clusters (by group at LMU Munich: Joe Mohr, Sebastian Bocquet, and Aditya Singh). Total SNR = 33, same model as used in latest SPT analysis (Bocquet et al. 2024).

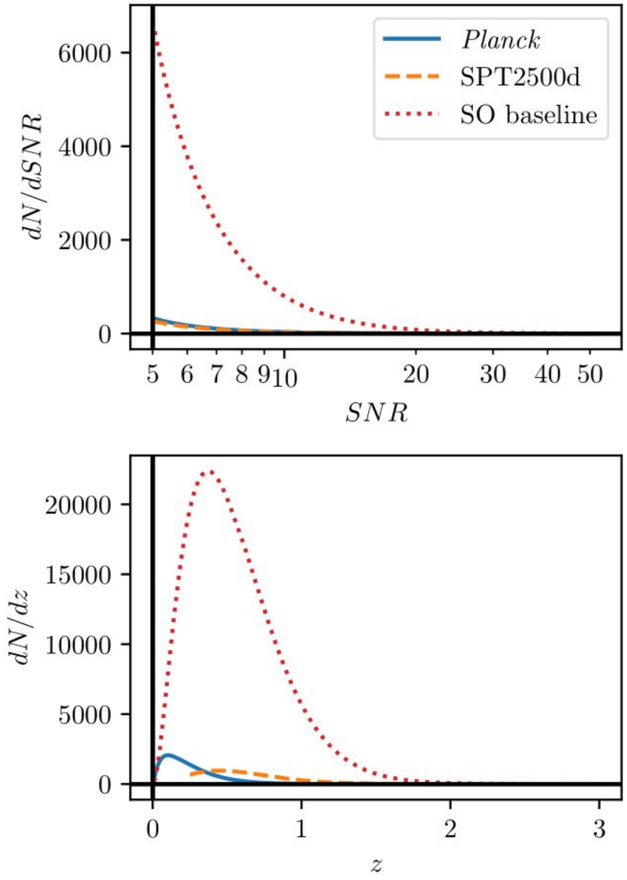
Likelihood: **`cosmocnc`**, forward modelling all observables (redshifts, tSZ, lensing), with mass calibration performed jointly with the number count likelihood analysis.

Stay tuned!

Likelihood code: `cosmocnc` (with **Boris Bolliet**)

Cluster number count likelihood code, general and flexible:

- **Binned**, **unbinned**, and extreme value likelihoods.
- Arbitrary number of mass observables (useful for **mass calibration**), including vector observables (e.g., lensing profiles).
- **Stacked** observables.
- Allows for **correlated scatter** between mass observables.
- Redshift measurement uncertainties.
- **Unconfirmed** detections.
- Written in Python, interfaced with **Cobaya**.
- Pretty fast and accurate.



IZ & B Bolliet 2024 2403.09589

`github.com/inigozubeldia/cosmocnc`

Likelihood code: `cosmocnc`

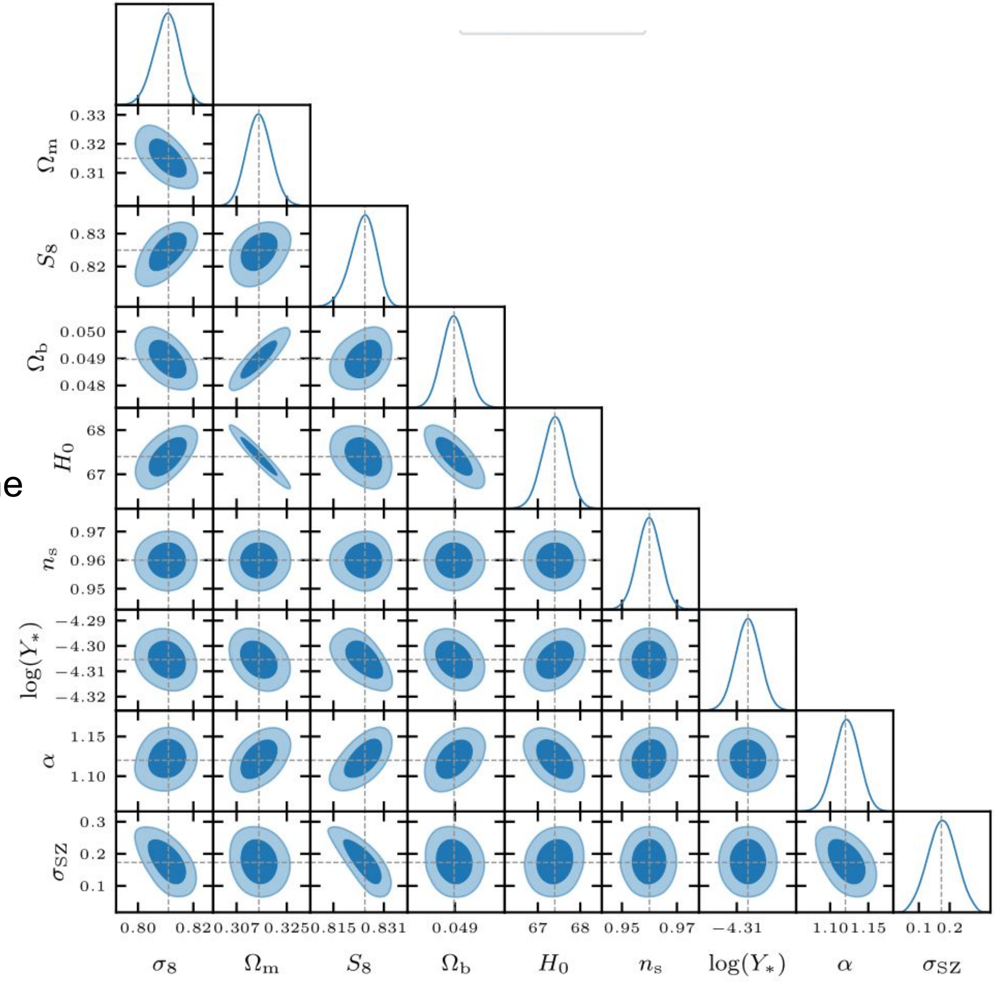
Validated with 32 SO-like synthetic catalogues produced at same point in parameter space (**SZ + cluster CMB lensing**, with lensing both cluster-by-cluster and stacked).

Synthetic catalogues (same assumptions as likelihood):

- Sample hmf x dV/dz: realisation of clusters in the Universe.
- Apply scaling relations.
- Impose selection criterion.

Biases less than ~0.15*σ* for all cosmological parameters!

Likelihood is **SO-ready** (modulo sample variance). Also tested for ***Planck***.



Simulation-based inference (SBI) for cluster cosmology

Preliminary!

With **Will Handley** and **Boris Bolliet** (Cambridge)

cosmocnc likelihood validated with synthetic catalogues.

Can we use those catalogues for cosmological inference, without the likelihood?

Answer: **yes! SBI** Why? See later.

Basic idea:

- Simulate parameter-data pairs.
- Learn their joint pdf.
- Condition pdf on data -> posterior on parameters given the data.

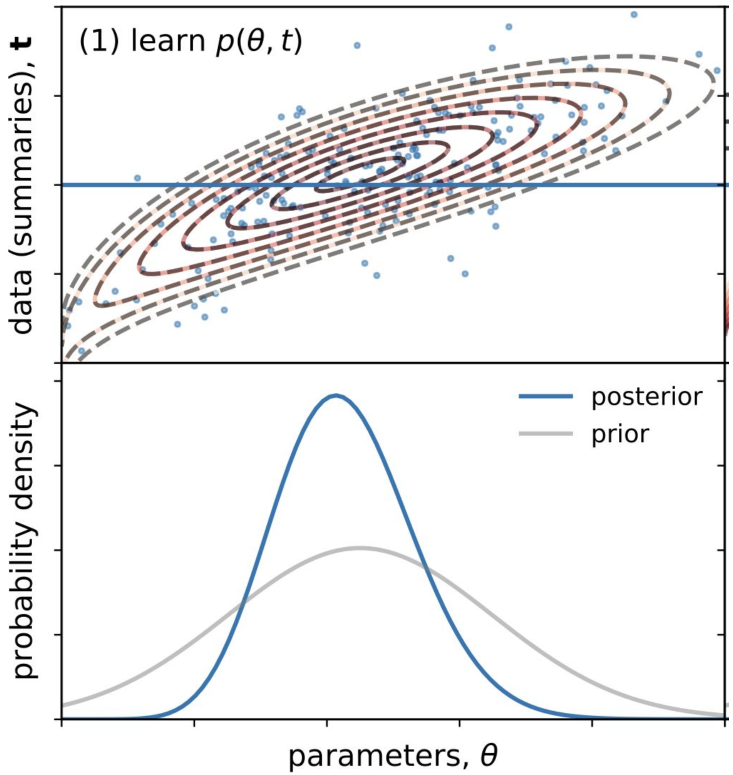


Figure credit: Alsing et al. 2019

Simulation-based inference (SBI) for cluster cosmology

With **Will Handley** and **Boris Bolliet** (Cambridge)

- Doing this for SO-like sample (~16000 clusters with 2 mass observables + redshift each) . Data vector very high-dimensional (16000 x 3).

- Compress catalogue into the **score function** (optimal data compression; Alsing & Wandelt 2017).

$$\mathcal{C} \rightarrow \mathbf{t} = \nabla_{\boldsymbol{\theta}} \log \mathcal{L}(\mathcal{C})$$

- Likelihood for compression only needs to be approximate (SNR loss, but not a bias).

- Evaluated at fiducial point in parameter space.

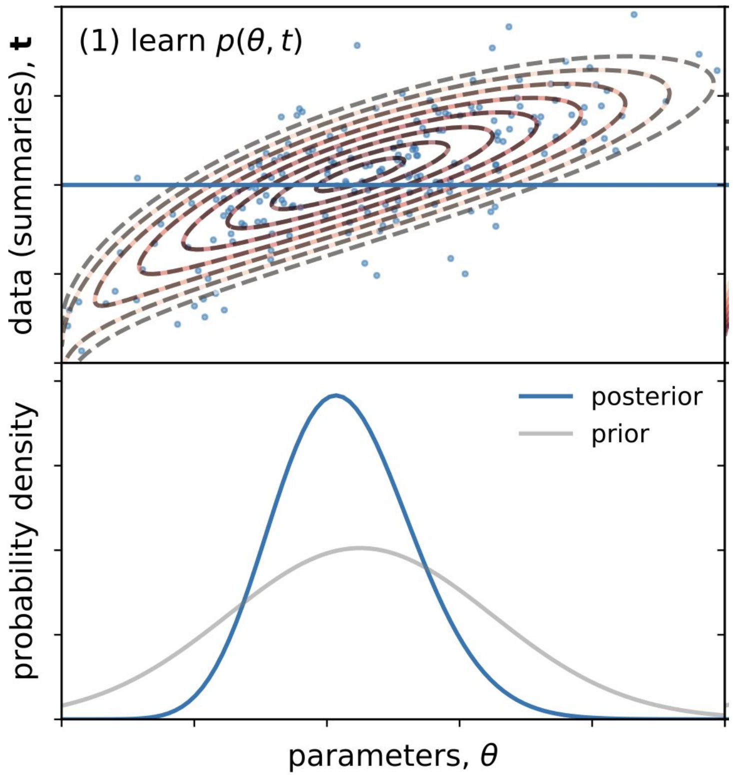


Figure credit: Alsing et al. 2019

Simulation-based inference (SBI) for cluster cosmology

With **Will Handley** and **Boris Bolliet** (Cambridge)

Algorithm:

- **Sample points in parameter space from the prior** (10 parameters, cosmology + scaling relation)**.**

- **Generate catalogues** at those points in param. space

- **Compress** them into score function (10-dimensional data vectors).

- **Learn posterior** (**sbi** package, neural-net density estimator).

- **Sample from learnt posterior** using MCMC for 100 data points generated at same point in parameter space ("truth"): assess biases and precision.

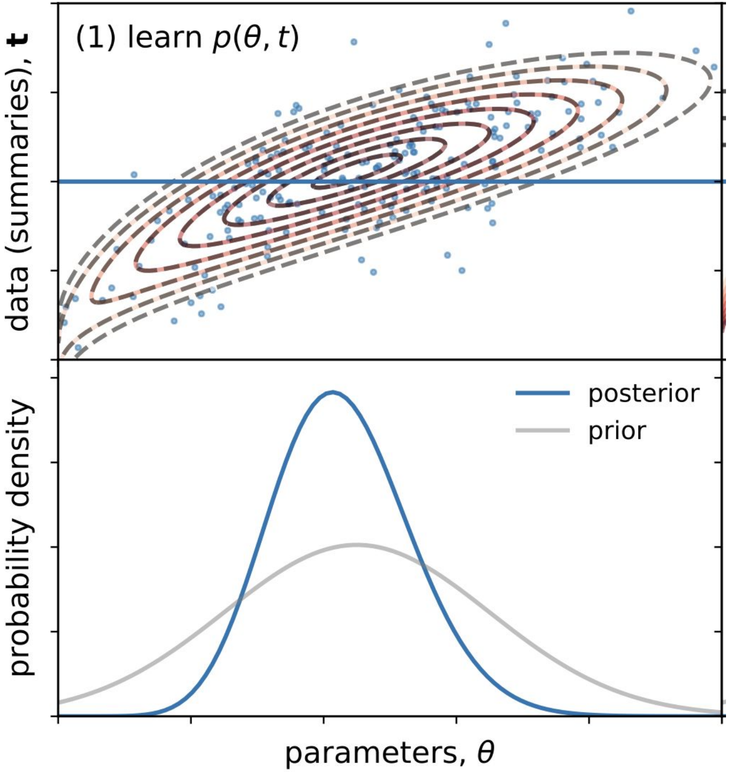


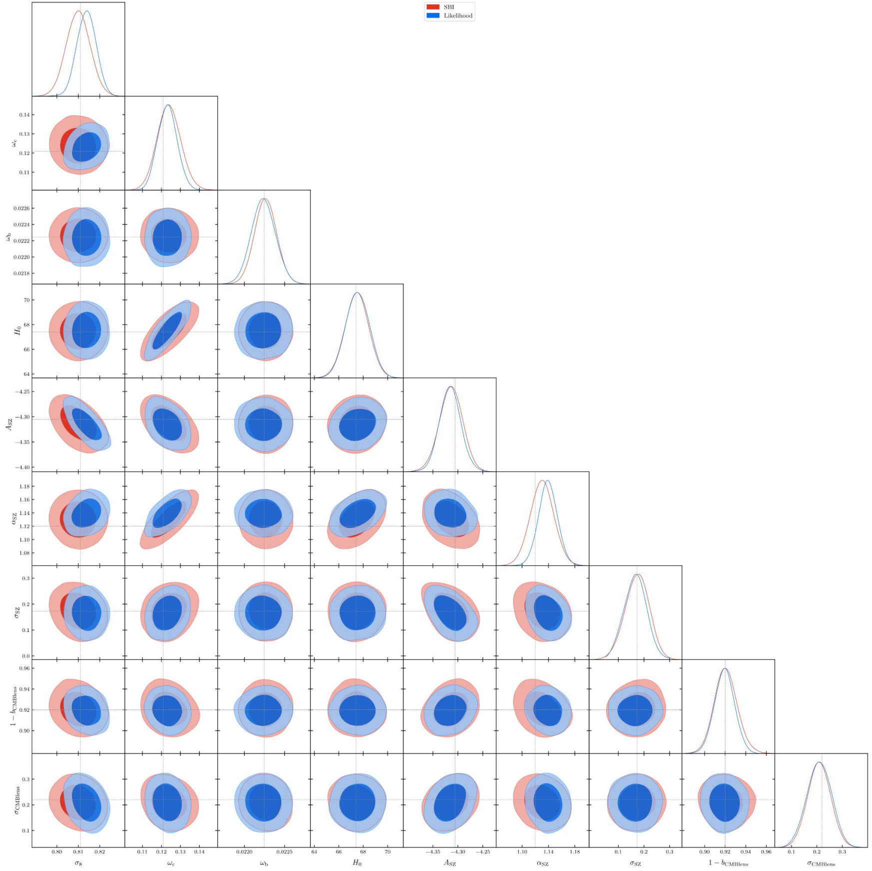
Figure credit: Alsing et al. 2019

Simulation-based inference (SBI) for cluster cosmology

Preliminary!

With **Will Handley** and **Boris Bolliet** (Cambridge)

- Works pretty well!
- Constraints on Ω_m, σ_8 that are 20-30% worse than likelihood analysis, but **unbiased**.
- No reason why constraints should be worse:
  - Score compression with finite differences. We're now writing the likelihood in **JAX** to get automatic differentiation (99% there).
  - Used 2500 training data points. Investigate if improves with more.
  - Play more with sbi hyperparameters.



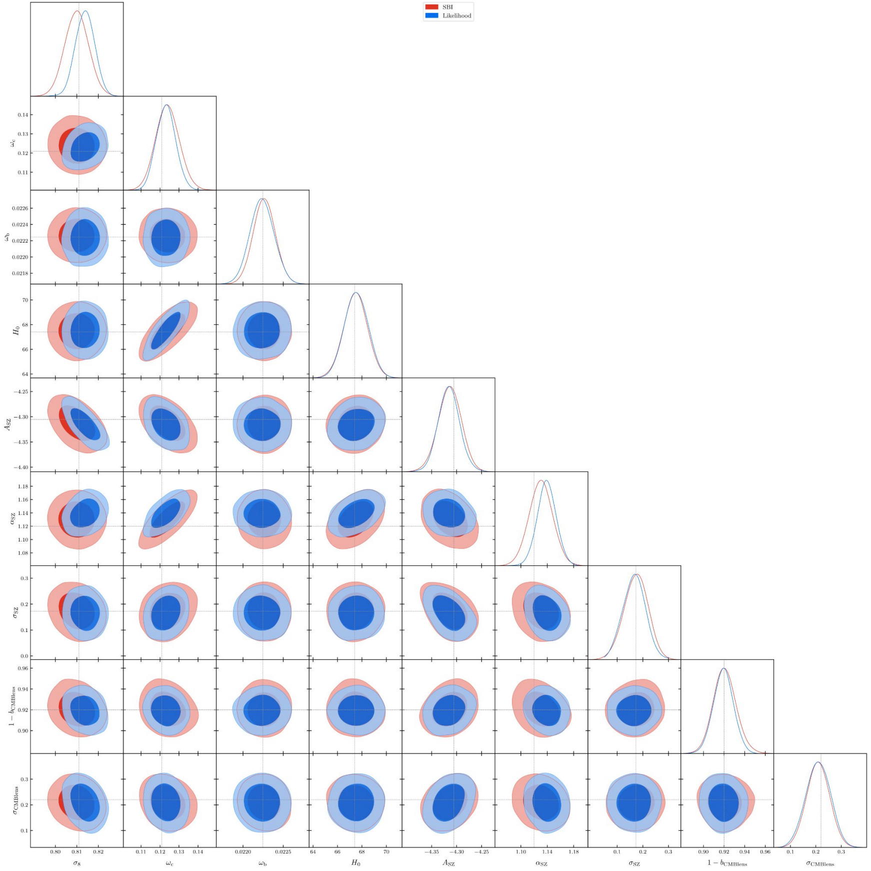
Simulation-based inference (SBI) for cluster cosmology

Preliminary!

With **Will Handley** and **Boris Bolliet** (Cambridge)

Why?

- For many clusters and many mass observables, **likelihood can be slow** to evaluate, especially if there is correlation in the scatter between observables. **Generating catalogues barely scales with number of observables** and whether their scatter is correlated.
- Likelihood for score compression only approximate: precision parameters can be small (i.e., fast) and correlation can be ignored (fast).
- Catalogue generation is trivially **parallelisable.**
- May be easier to include **sample variance**. We plan to explore this.



Summary

- **`SZiFi`**:
  - Iterative noise estimation: removes ILC bias, boosts SNR.
  - Spectrally constrained MMF: useful to tackle CIB contamination.
  - Relativistic corrections.
- **`cosmocnc`**: New fast and flexible cluster number count likelihood code, S3-ready.
- Both applied to *Planck* data:
  - **New *Planck* catalogues**.
  - **Upcoming cosmological analysis**, with CMB lensing and DES Y3 lensing mass calibration (by LMU group). Stay tuned!
- SBI for cluster cosmology: promising novel technique!