PANCO2

The second Pipeline for the Analysis of NIKA2 Cluster Observations

Florian Kéruzoré, LPSC, Grenoble
mmUniverse@NIKA2, June 2021

financé par
IDEX Université Grenoble Alpes

Introduction

- Cluster physics & cosmology need ICM pressure profile measurements

  → High-resolution SZ observations *(see talk by L. Perotto)*

- NIKA2 LPSZ: 50 clusters to measure pressure profile at high-z with high resolution SZ

  → Need fast and accurate individual pressure profile evaluations

- This work: a new Python software to measure ICM pressure distribution

  - Using bayesian MCMC to fit a pressure profile on an SZ map
  - **Forward modeling approach** that takes into account mm data features
  - **Flexible user inputs:** radial binning, parameter priors, choice of center, …
  - Optimized: uses efficient Python numerical computation libraries, MCMC multithreading

- So far tested on simulations and working on NIKA2 data (see talks by E. Artis, M. Muñoz)

  → Generalization to any instrument in progress

Outline

Florian Kéruzoré
mmUniverse
June 2021

Algorithm

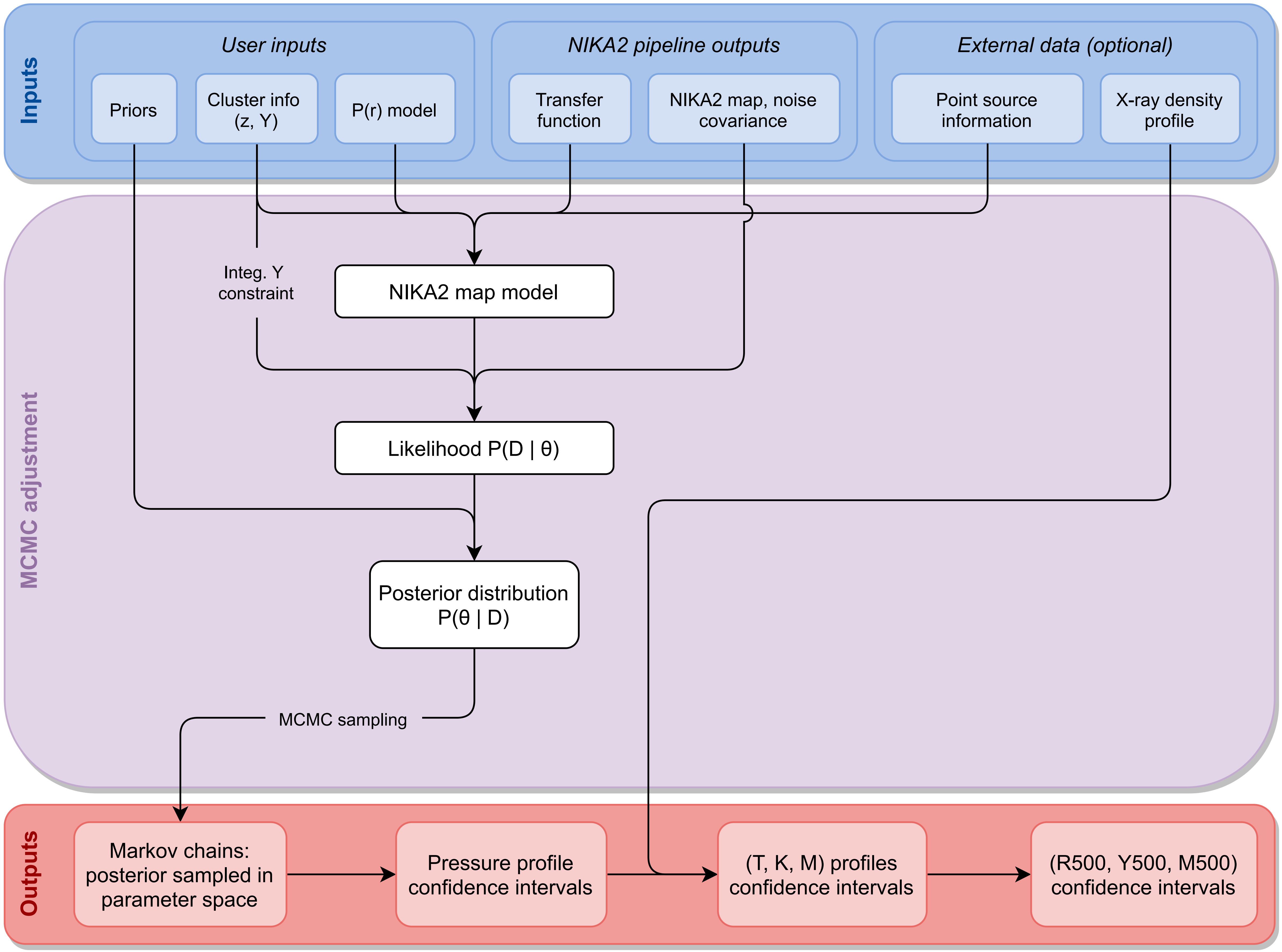
Validation on simulated input & results showcase

Validation on MUSIC hydrodynamical simulations

Conclusions & perspectives

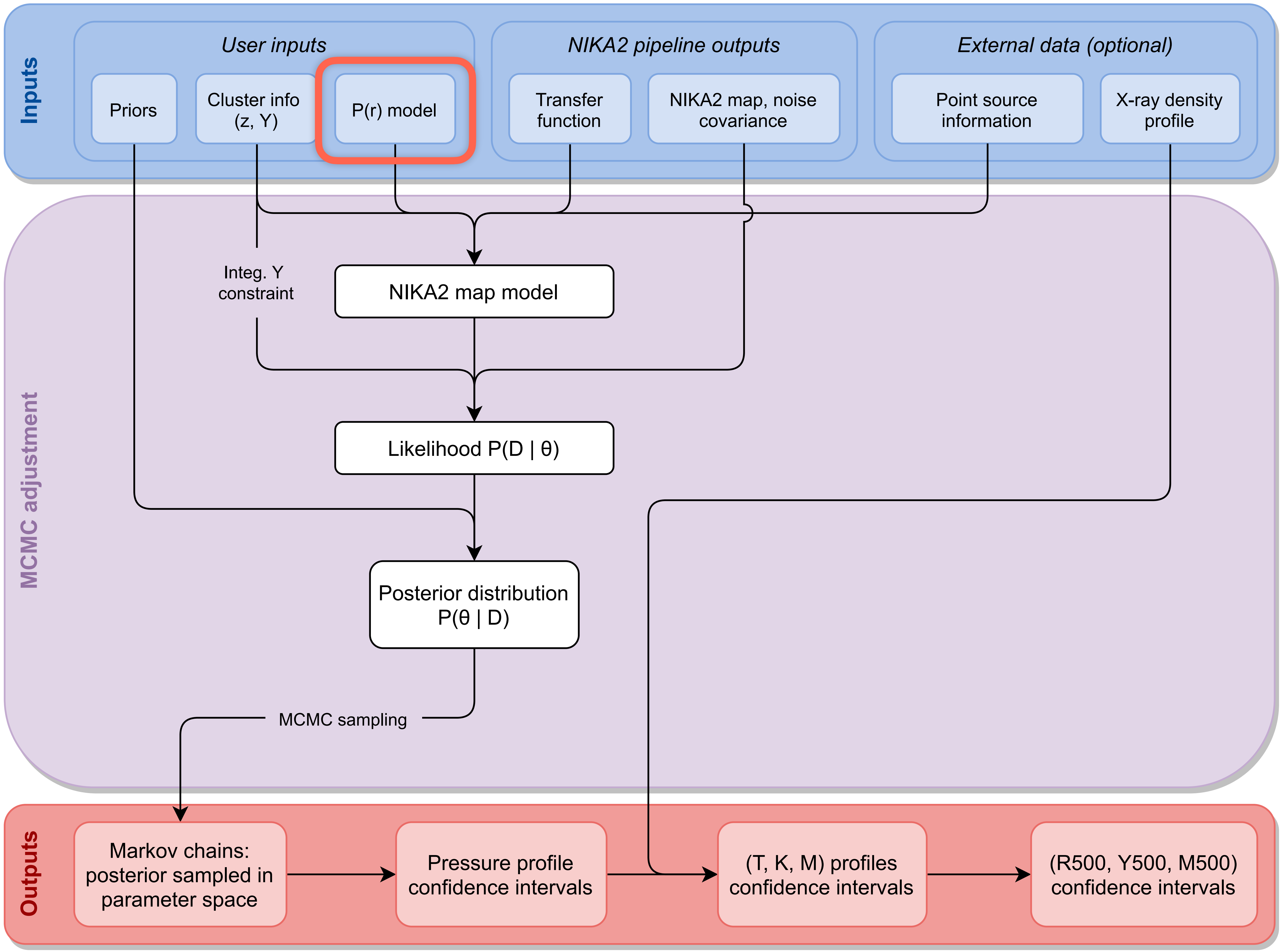
④ Algorithm flowchart

Florian Kéruzoré
mmUniverse
June 2021



④ Algorithm flowchart

Florian Kéruzoré
mmUniverse
June 2021



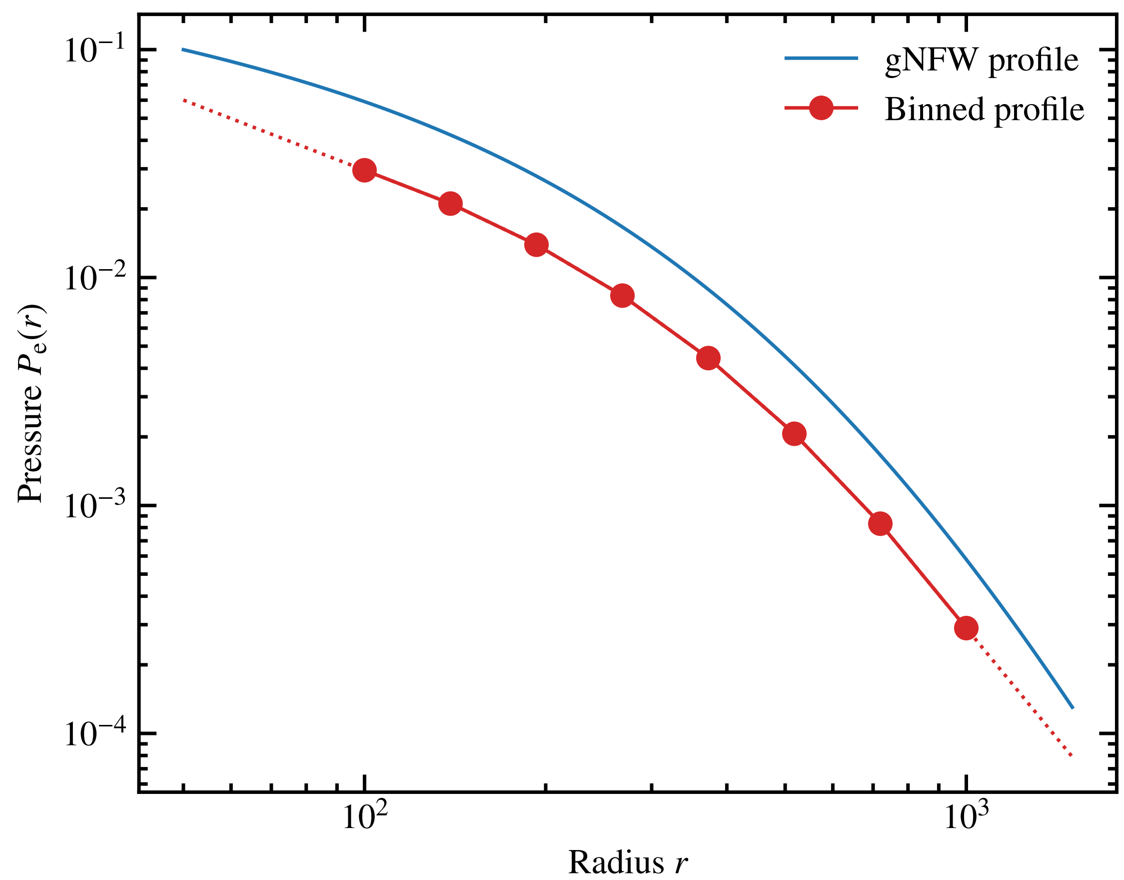
⑤ ICM pressure modeling

- Pressure distribution modeled by a pressure profile $P(r)$
- Two models implemented:
  - **gNFW:** $P(r) = P_0 \, (r/r_p)^{-c} \, [1 + (r/r_p)^a]^{\frac{c-b}{a}}$

    Normalisation P_0, two slopes (b, c), transition radius r_p and sharpness a
  - **Binned:** $P(R_i < r < R_{i+1}) = P_i \, (r/R_i)^{-\alpha_i}$

    For n bins, n pressure P_i at radii R_i, and $\alpha_i = -\log(P_{i+1}/P_i) \Big/ \log(R_{i+1}/R_i)$

    Inner and outer slopes (outside bins) are extrapolated



⑤ ICM pressure modeling

- Pressure distribution modeled by a pressure profile $P(r)$
- Two models implemented:
  - **gNFW:** $P(r) = P_0\,(r/r_p)^{-c}\,[1 + (r/r_p)^a]^{\frac{c-b}{a}}$
    Normalisation P_0, two slopes (b, c), transition radius r_p and sharpness a
  - **Binned:** $P(R_i < r < R_{i+1}) = P_i\,(r/R_i)^{-\alpha_i}$
    For n bins, n pressure P_i at radii R_i, and $\alpha_i = -\log(P_{i+1}/P_i)\;\big/\;\log(R_{i+1}/R_i)$
    Inner and outer slopes (outside bins) are extrapolated

- Advantages and drawbacks:

gNFW

- widely used
  → easy to compare with literature
- smooth
  → easy to differentiate and extrap.
- restricted shape
  → cannot identify features
- strongly correlated parameters
  → slow MCMC

Binned

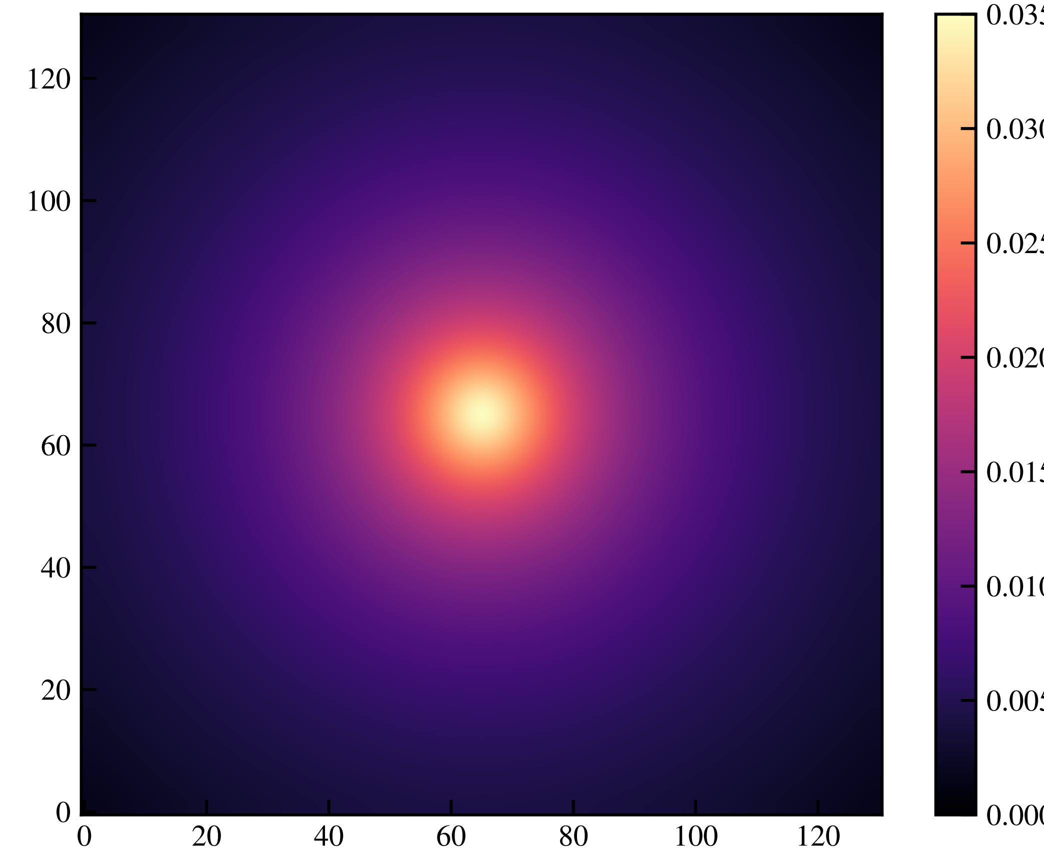
- less restricted shape
  → can identify features
- lower parameter correlations
  → faster MCMC
- not smooth
  → trickier to differentiate and extrap.

Best model depends on analysis goals

⑥ SZ signal modeling

- Pressure profile integrated along the line of sight
  Numerical integration for gNFW, analytical for binned

  → Compton parameter y map



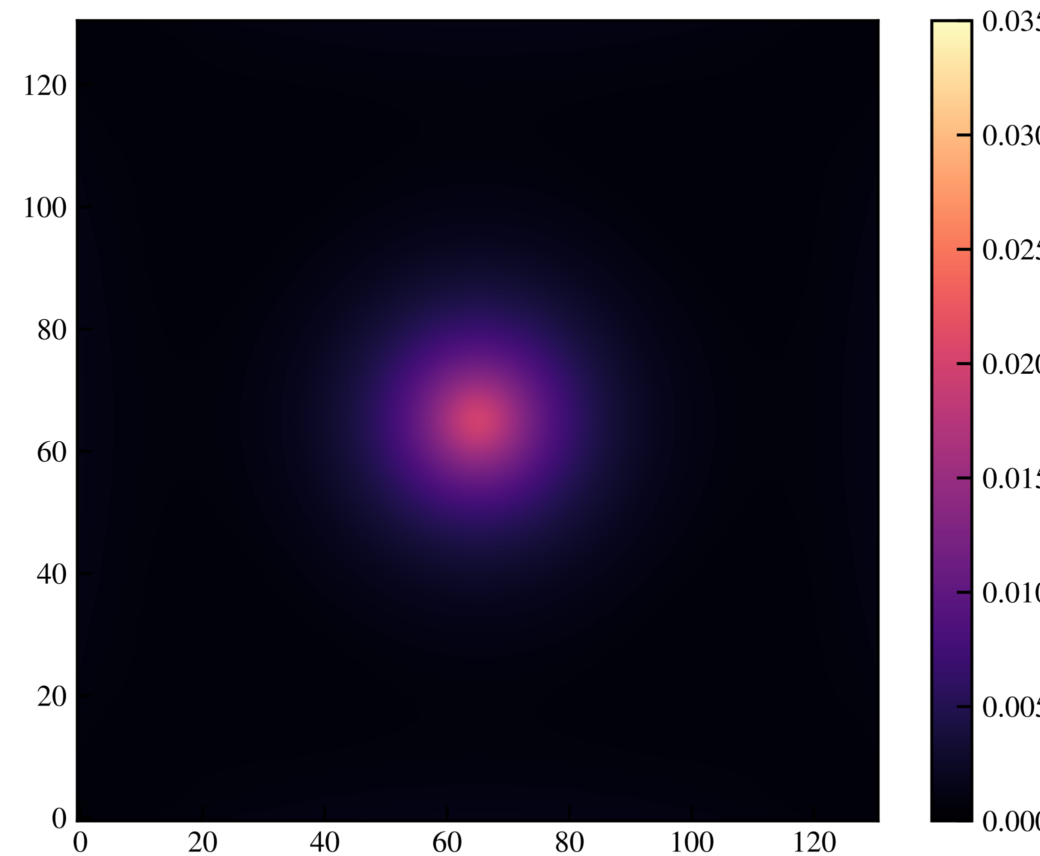
6 SZ signal modeling

- Pressure profile integrated along the line of sight
  Numerical integration for gNFW, analytical for binned

  → **Compton parameter y map**

- Convolved by the **NIKA2 beam** (PSF smearing) and **transfer function** (pipeline filtering)

  → **Filtered (data-like) y map**



⑥ SZ signal modeling

- Pressure profile integrated along the line of sight
  Numerical integration for gNFW, analytical for binned

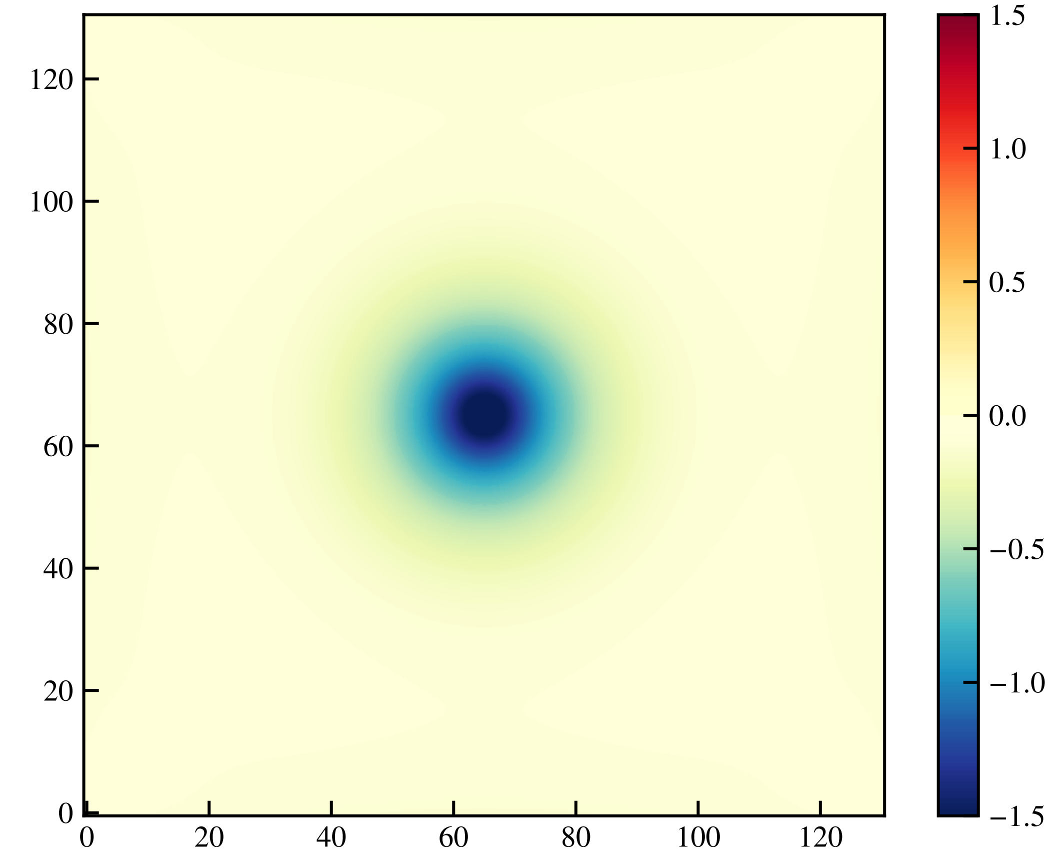
  → Compton parameter y map

- Convolved by the **NIKA2 beam** (PSF smearing) and **transfer function** (pipeline filtering)

  → Filtered (data-like) y map

- Conversion to surface brightness units
  Coefficient taken in input, treated as nuisance parameter

  → Filtered, calibrated SZ map



⑥ SZ signal modeling

- Pressure profile integrated along the line of sight
  Numerical integration for gNFW, analytical for binned

  → **Compton parameter y map**

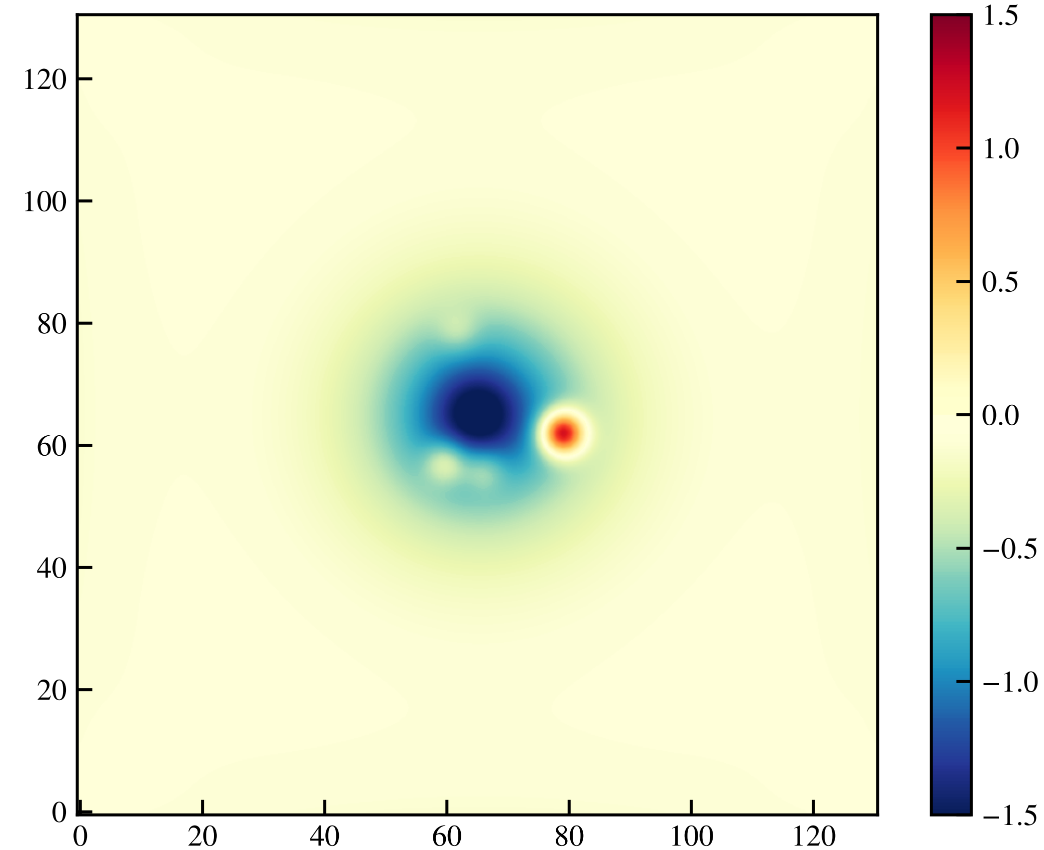
- Convolved by the **NIKA2 beam** (PSF smearing) and **transfer function** (pipeline filtering)

  → **Filtered (data-like) y map**

- Conversion to surface brightness units
  Coefficient taken in input, treated as nuisance parameter

  → **Filtered, calibrated SZ map**

- (Optional) Point source contamination: point sources with variable fluxes can be part of the model
  Treated as nuisance parameters; see Keruzore+20



⑥ SZ signal modeling

- Pressure profile integrated along the line of sight
  Numerical integration for gNFW, analytical for binned

  → **Compton parameter y map**

- Convolved by the **NIKA2 beam** (PSF smearing) and **transfer function** (pipeline filtering)

  → **Filtered (data-like) y map**

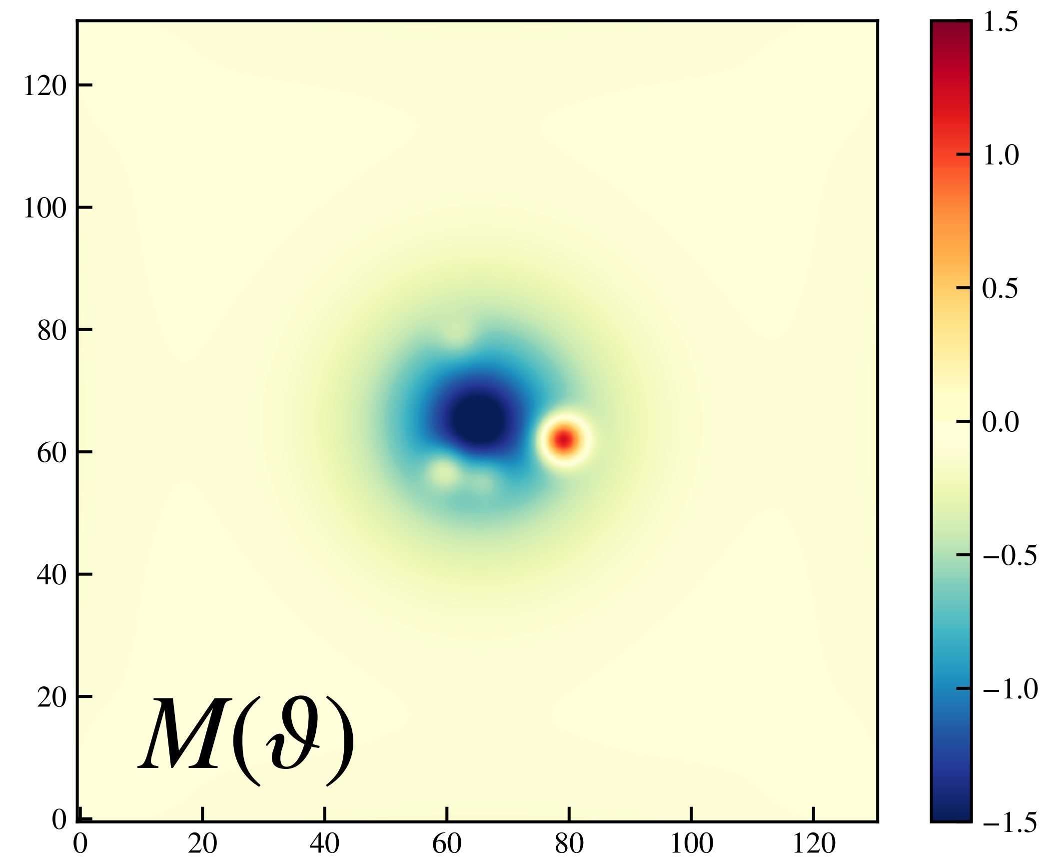
- Conversion to surface brightness units
  Coefficient taken in input, treated as nuisance parameter

  → **Filtered, calibrated SZ map**

- (Optional) Point source contamination: point sources with variable fluxes can be part of the model
  Treated as nuisance parameters; see Keruzore+20

- Summary — Parameters of the model:
  - Pressure profile parameters
  - Conversion coefficient
  - (optional) Point source fluxes
  - (optional) Map zero level

  $\Big\}\ \vartheta$



⑦ Model fitting

- Model map from previous slide $M(\vartheta)$ fitted on the data with likelihood function:

$$-2\log\mathscr{L}(\vartheta) = \sum_{\text{pixels}} \left(\frac{D_{\text{NIKA2}} - M(\vartheta)}{\sigma_{\text{NIKA2}}}\right)^2 + \left(\frac{Y_\Delta^{\text{input}} - Y_\Delta(\vartheta)}{\delta Y_\Delta^{\text{input}}}\right)^2$$

Comparison between NIKA2 map D and model map $M(\vartheta)$ (with noise rms map σ)

Constraint on integrated SZ signal from input survey (Planck, ACT)

$$Y_\Delta(\vartheta) \propto \int_0^{R_\Delta} P_e(r;\vartheta)\, r^2 \mathrm{d}r$$

- A noise covariance matrix can also be included for correlated noise
- Priors on parameters defined by the user → posterior distribution
- MCMC sampling of the posterior distribution
  - Convergence check based on Gelman-Rubin and autocorrelation
  - User-defined analysis parameters: # of chains, burn-in length, convergence check parameters
- Once convergence is reached:
  - remove the chains considered unconverged by the convergence check
  - (optional) thinning: keep one point every autocorrelation length

→ Final chains

Outline

Algorithm

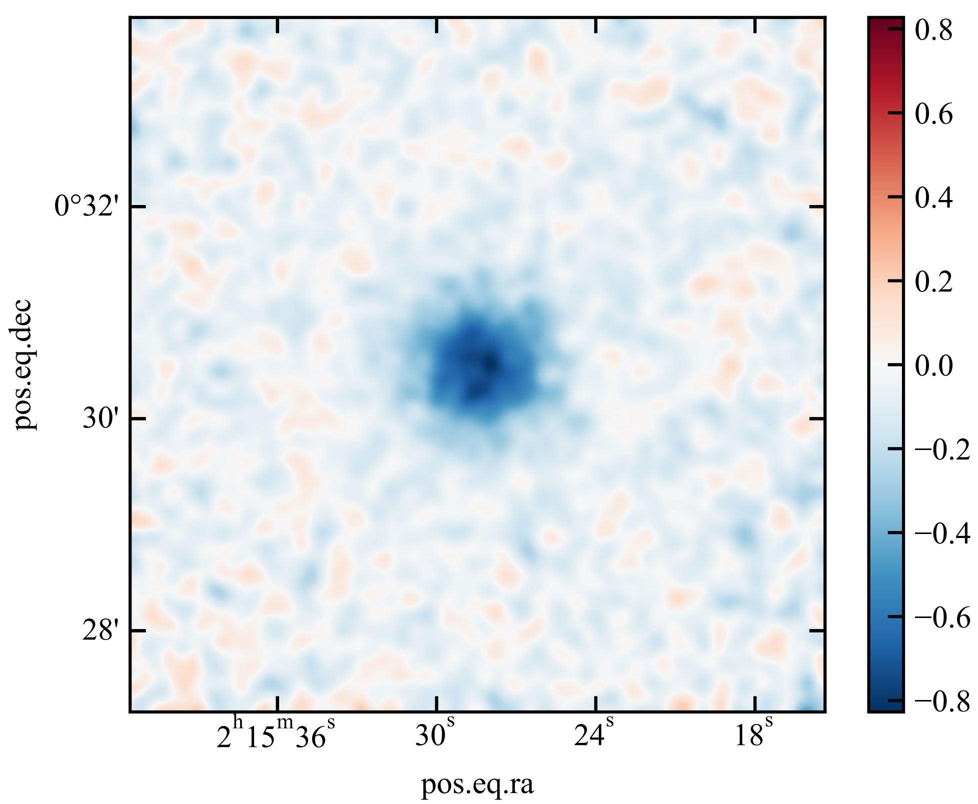
Validation on simulated input & results showcase

Validation on MUSIC hydrodynamical simulations

Conclusions & perspectives

⑨ Results on simulated input

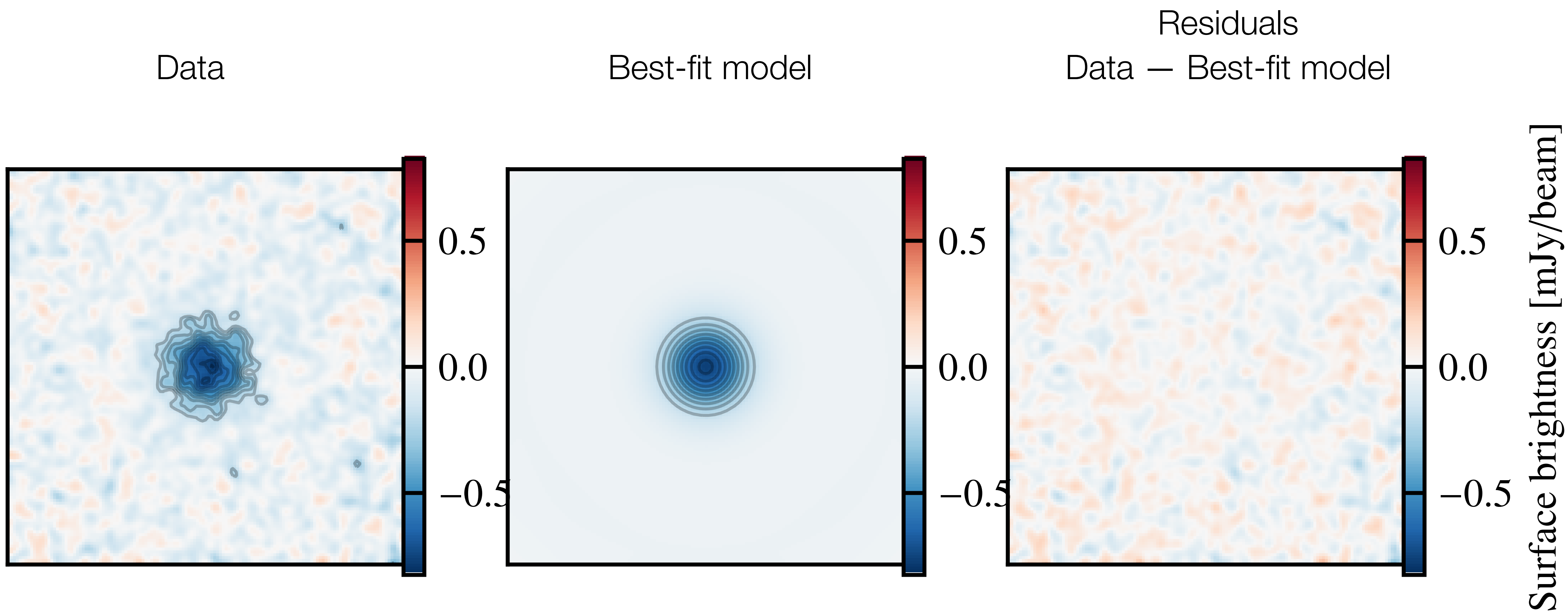
- Test — run the fit:
  - Of a simulated cluster map
  - With white noise, realistic filtering
  - With constraints on Y_{500}
  - With a binned pressure profile
  - Combined with X-ray density profile
- Convergence reached in <10 minutes with 30 chains running on 30 threads
- Example will be included in the public release




```
⌥⌘3                    IPython: keruzore/panco2

==> Initialization
    No covariance or noise simulations: considering white noise
    Large scale signal constrained by Y_500 = (38.42  +/- 10.67) kpc2
    TF convention: NYQUIST
    Loaded X-ray profiles from ./Demo/Data/parprod4.json
    MCMC starting point:
        P0 = 0.08005177270475204
        P1 = 0.0250456576742367
        P2 = 0.011697007853108513
        P3 = 0.004320671079528028
        P4 = 0.0012254141649320678
        P5 = 0.00010317434803111635
        calib = -11.9
        zero = 0.0

==> MCMC sampling...
    Convergence will be checked every 1000 steps
  1%|▌                                        | 999/100000 [01:35<2:33:32, 10.75it/s]
    1000 iterations
  2%|▊                                        | 1999/100000 [03:10<2:37:33, 10.37it/s]
    2000 iterations
  3%|█                                        | 3000/100000 [04:46<2:36:30, 10.33it/s]
    3000 iterations
    Removed 28 chains because of too long autocorrelation
    Max autocorrelation length per chain: 56 +/- 2
R_hat = [1.001 1.009 1.009 1.003 1.017 1.045 1.016 1.006]
  4%|█▍                                       | 4000/100000 [06:22<2:40:28,  9.97it/s]
    4000 iterations
    Removed 17 chains because of too long autocorrelation
    Max autocorrelation length per chain: 82 +/- 13
R_hat = [1.008 1.01  1.012 1.008 1.012 1.012 1.009 1.006]
  5%|█▊                                       | 5000/100000 [08:00<2:41:53,  9.78it/s]
    5000 iterations
    Removed 16 chains because of too long autocorrelation
    Max autocorrelation length per chain: 93 +/- 21
R_hat = [1.009 1.009 1.006 1.01  1.009 1.012 1.008 1.008]
  6%|██                                       | 6000/100000 [09:36<2:25:47, 10.75it/s]
    6000 iterations
    Removed 9 chains because of too long autocorrelation
    Max autocorrelation length per chain: 112 +/- 32
R_hat = [1.01  1.008 1.009 1.01  1.007 1.009 1.011 1.005]
  6%|██                                       | 6000/100000 [09:36<2:30:34, 10.41it/s]
    MCMC running time: 00h 09m 37s

==> Managing Markov chains
    Removed 0 chains because of bad posterior values
    Removed 9 chains because of too long autocorrelation
    Max autocorrelation length per chain: 112 +/- 32
    Keep one point every autocorrelation length
    -> Total number of accepted independent points: 1033
    Best fit params:
       P0 = 0.035512528165889984
       P1 = 0.02940393485735021
       P2 = 0.012154193690797732
       P3 = 0.004851537647741591
       P4 = 0.0014813747673666307
       P5 = 0.00012207074455567458
       calib = -11.84930060146621
       zero = -5.490897554209001e-05

==> Thermodynamical properties computation

==> Plotting things
    Thermodynamical profiles...
    Parameter distributions...
    Data / model / residuals...

==> Integrated values from your fit
    R_500 = 803.92 +/- 32.60 kpc
    M_500 = 3.90 +/- 0.48 e14 Msun
    Y_500 = 41.42 +/- 4.93 kpc2

==> End of program.
    Your results are stored in: ./Results/demo_binned/
```

Results: Data, model, residuals

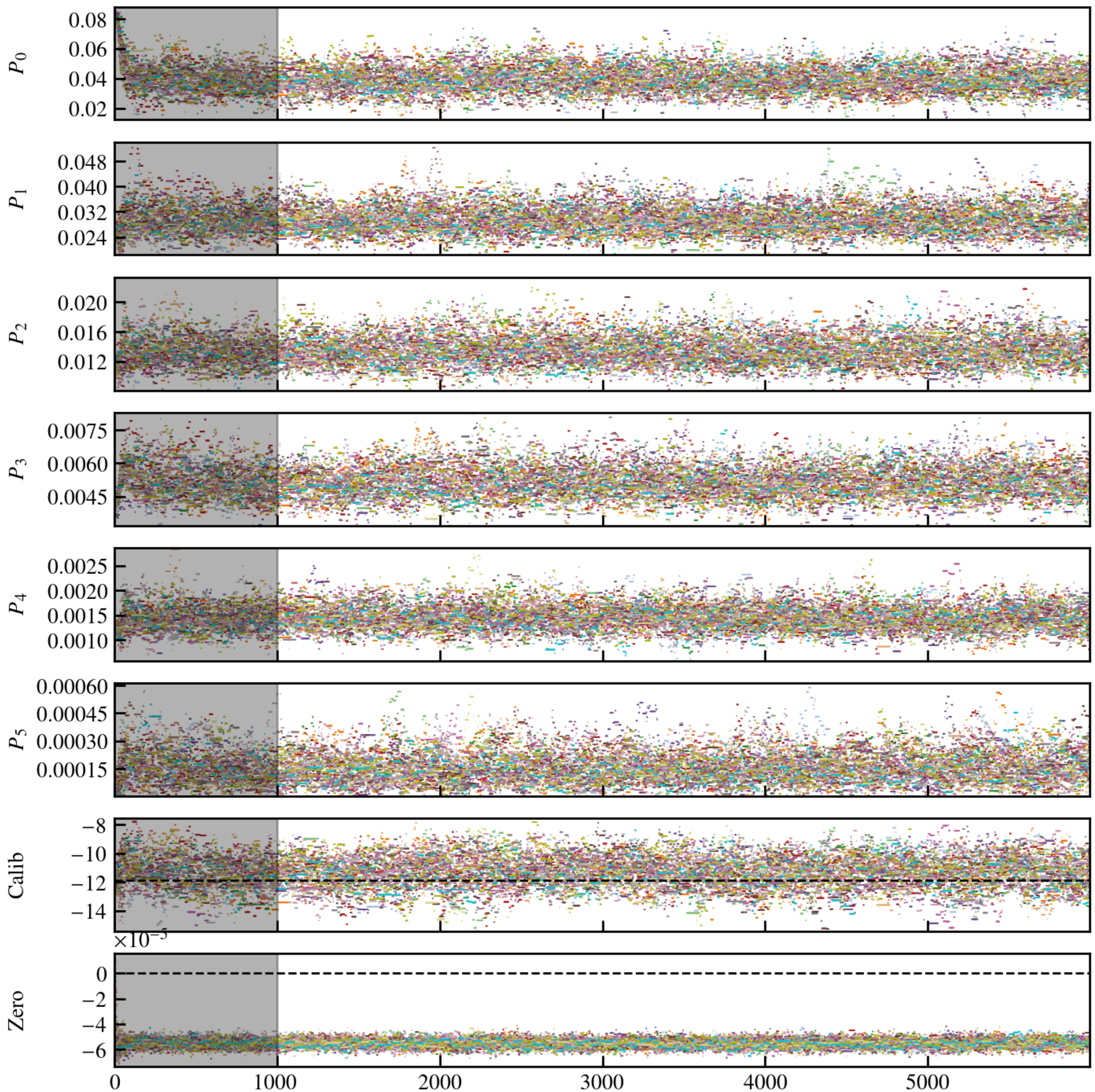
Florian Kéruzoré
mmUniverse
June 2021



Residuals compatible with noise

Results: Markov chains

Florian Kéruzoré
mmUniverse
June 2021



$P_0 \dots P_n$: pressure at radial bins;

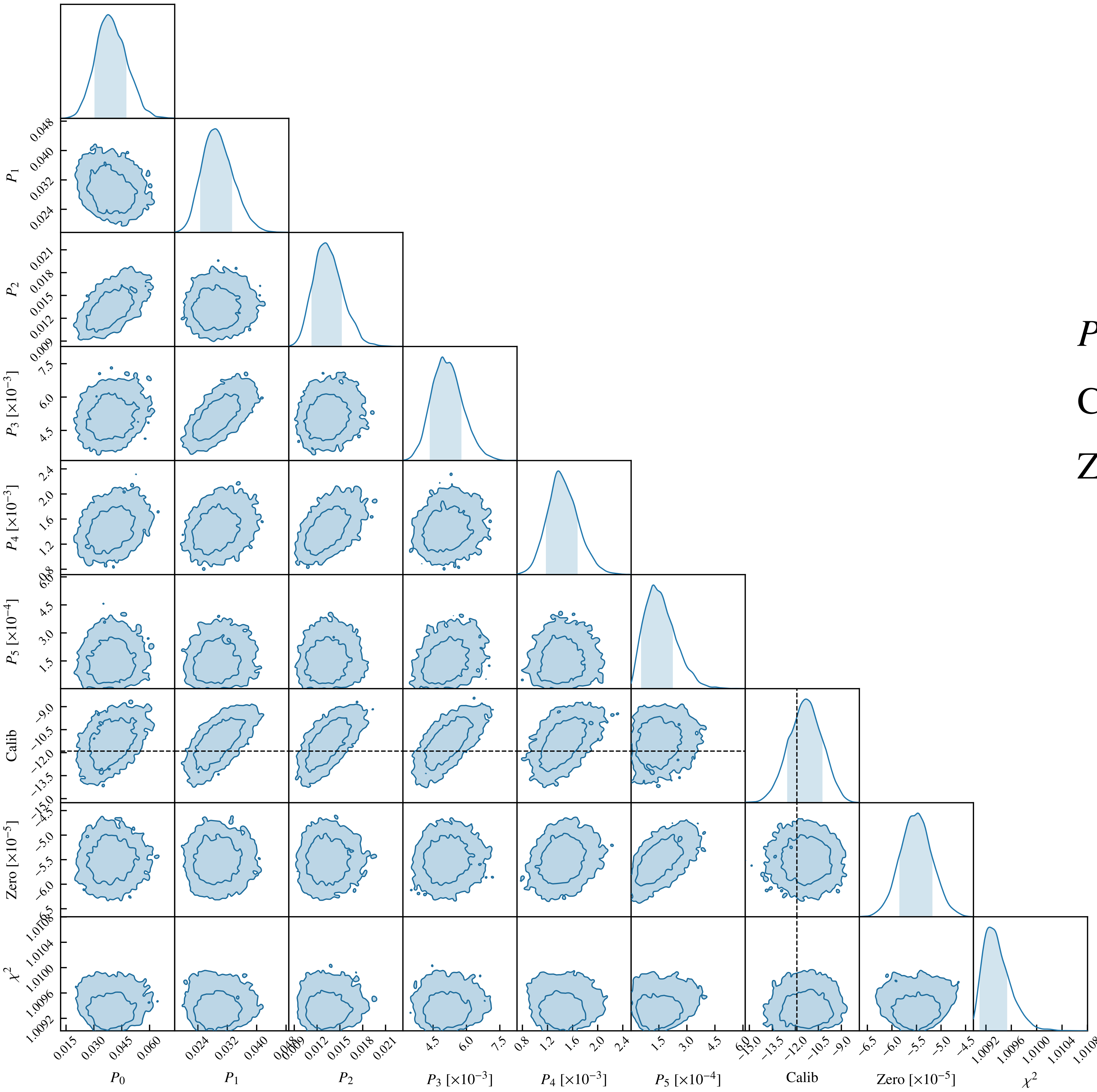
Calib : Conversion coefficient;

Zero : map zero level

Chains visually converged

Florian Kéruzoré
mmUniverse
June 2021

⑫ Results: parameter distributions



$P_0\ \ldots\ P_n$: pressure at radial bins;

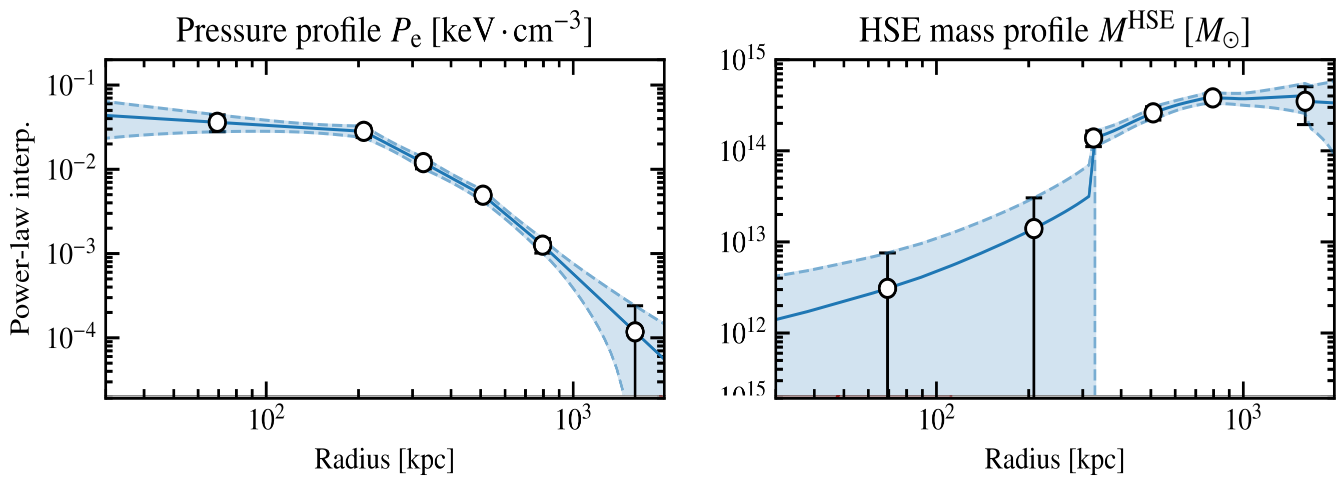
Calib : Conversion coefficient;

Zero : map zero level

Pressure profile interpolations

Florian Kéruzoré
mmUniverse
June 2021

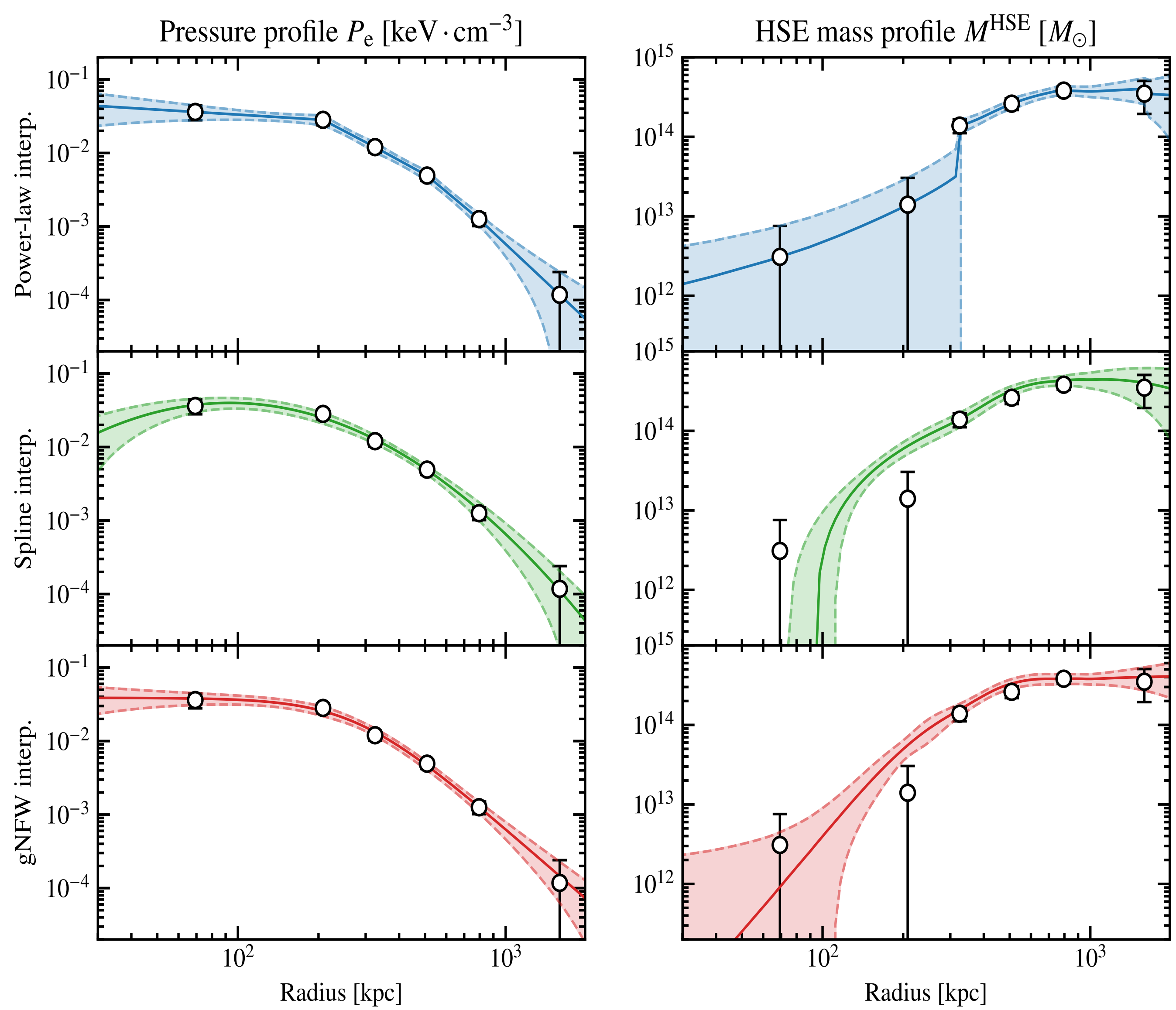
- For binned profiles, we need to interpolate the pressure profile between the radial bins.
- Interpolation scheme:
  - Perform interpolation on each pressure profile in the Markov chains
  - Use each profile to estimate quantities of interest (mass, temperature, etc)
  - Infer confidence intervals
- Power-law interpolation gives bumpy profiles, which leads to discontinuities in the mass profiles since $M^{\mathrm{HSE}}(r) \propto \dfrac{\mathrm{d}P}{\mathrm{d}r}$



Pressure profile interpolations

Florian Kéruzoré
mmUniverse
June 2021

- For binned profiles, we need to interpolate the pressure profile between the radial bins.
- Interpolation scheme:
  - Perform interpolation on each pressure profile in the Markov chains
  - Use each profile to estimate quantities of interest (mass, temperature, etc)
  - Infer confidence intervals
- Power-law interpolation gives bumpy profiles, which leads to discontinuities in the mass profiles since $M^{\mathrm{HSE}}(r) \propto \dfrac{\mathrm{d}P}{\mathrm{d}r}$
- 3 interpolation methods implemented:
  - Power-law interpolation (linear in log-log)
    → bumpy profile
  - Spline interpolation (in log-log)
    → smooth but can give weird extrap.
  - gNFW fit on each MCMC sample
    → smooth and "physics motivated"

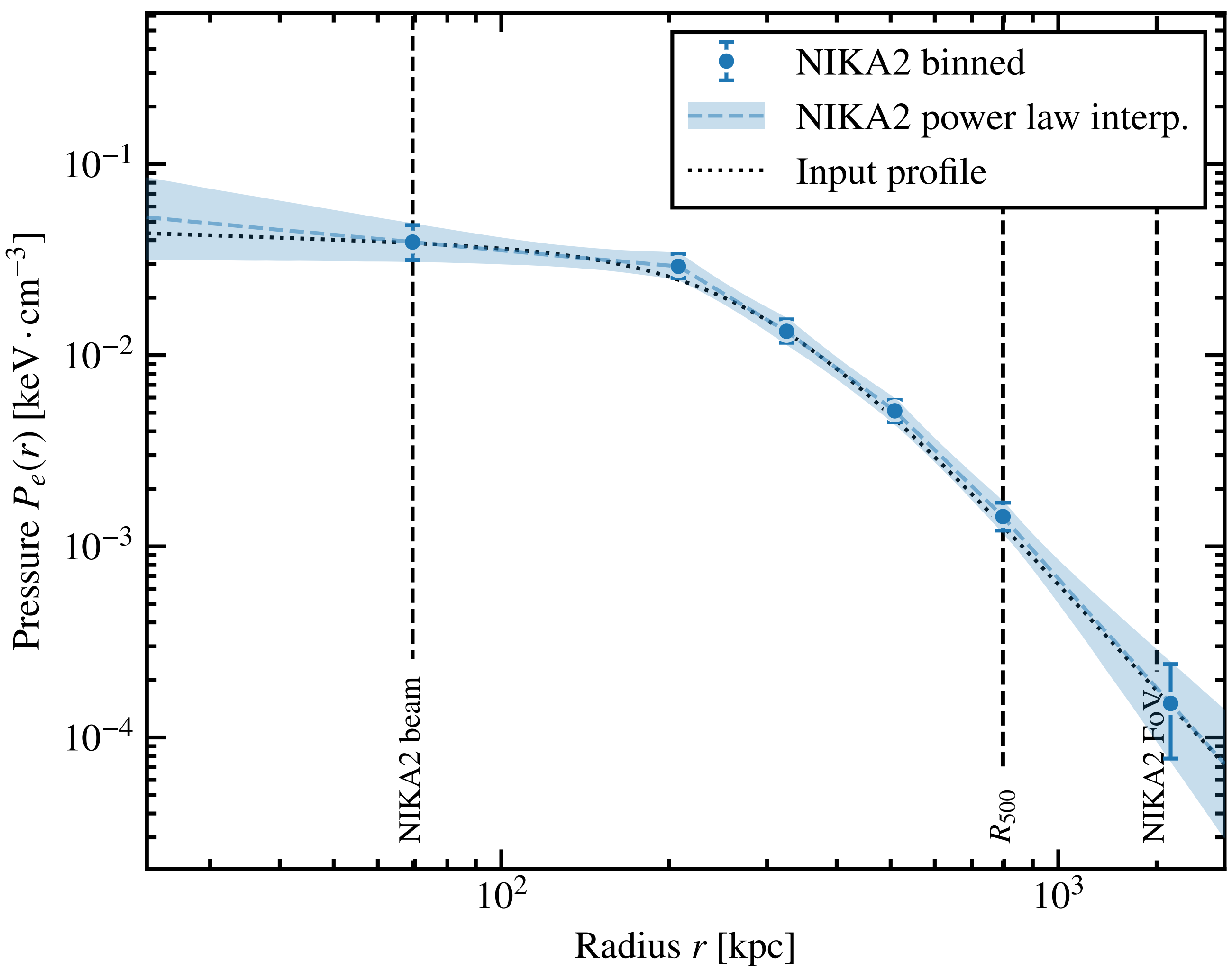


Best method depends on analysis goals

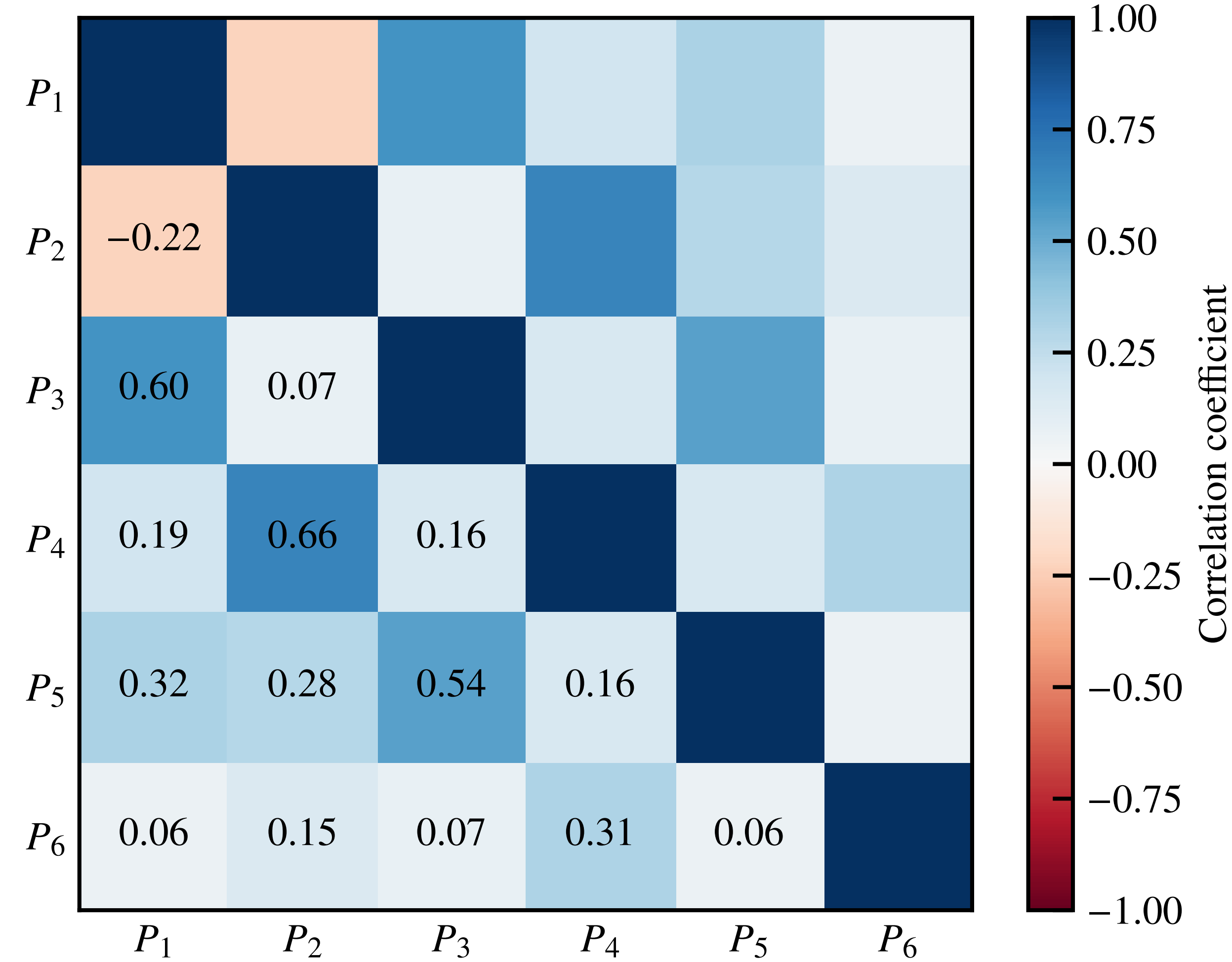
Results: pressure profile

Florian Kéruzoré
mmUniverse
June 2021

Pressure profile



If binned, correlation matrix of the pressure bins



	P_1	P_2	P_3	P_4	P_5	P_6
P_1						
P_2	−0.22					
P_3	0.60	0.07				
P_4	0.19	0.66	0.16			
P_5	0.32	0.28	0.54	0.16		
P_6	0.06	0.15	0.07	0.31	0.06	

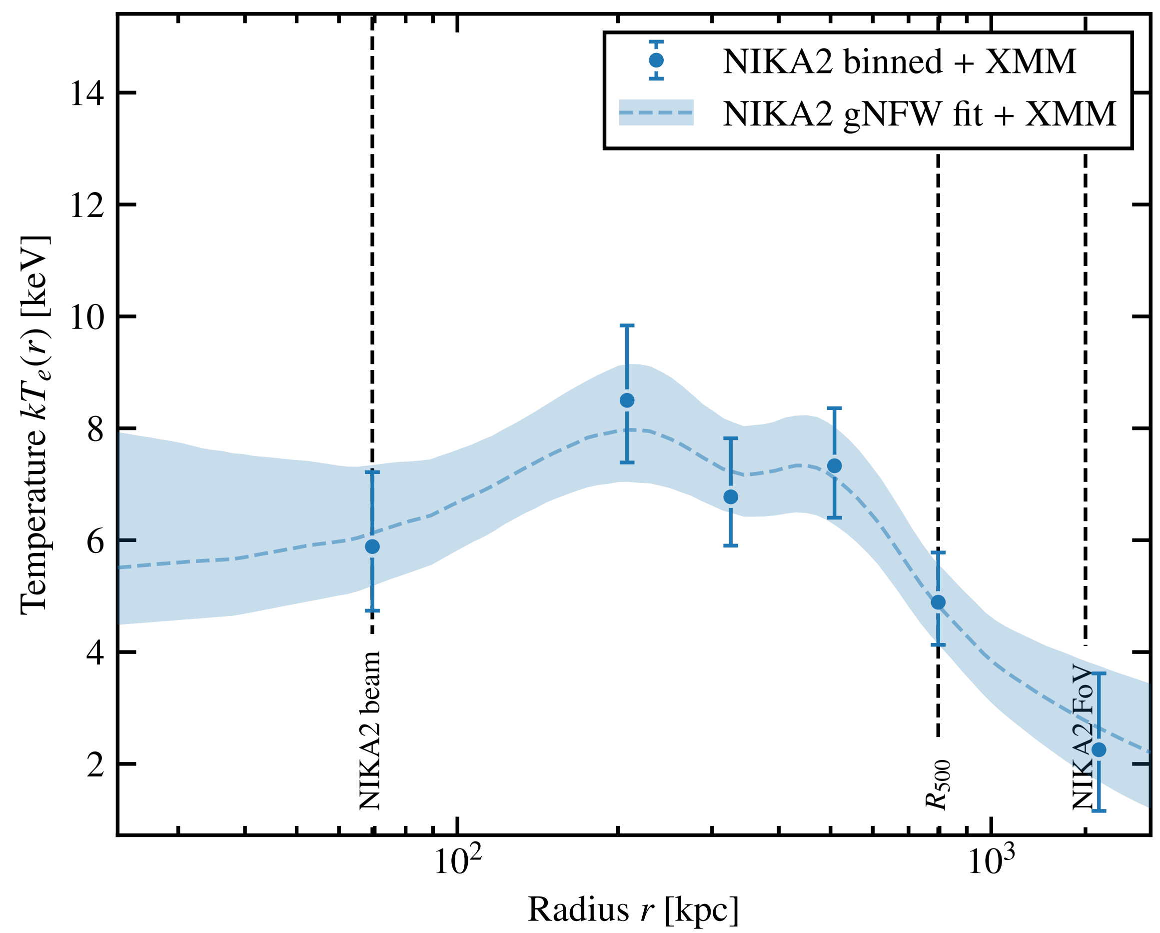
Recovered pressure profile in excellent agreement with the true input profile

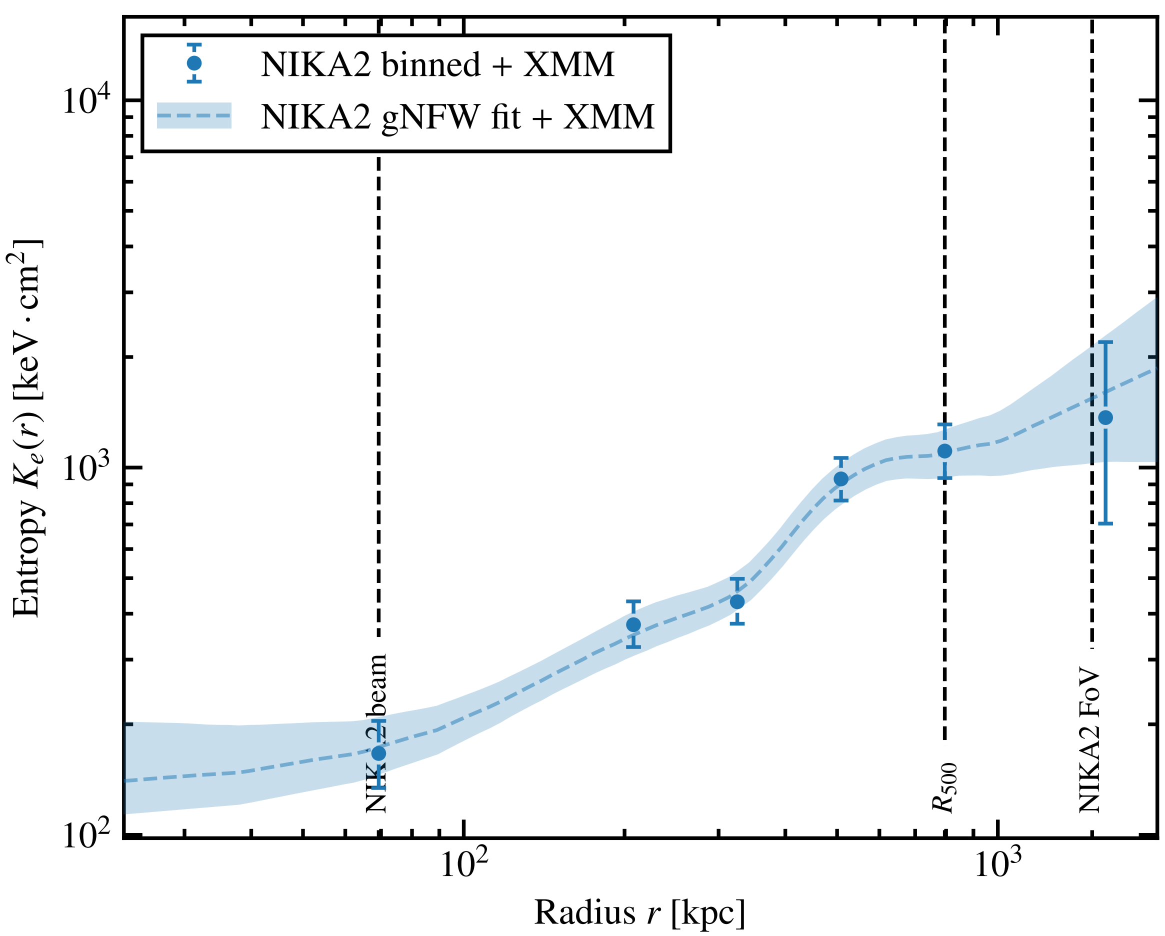
Results: Entropy, temperature, mass

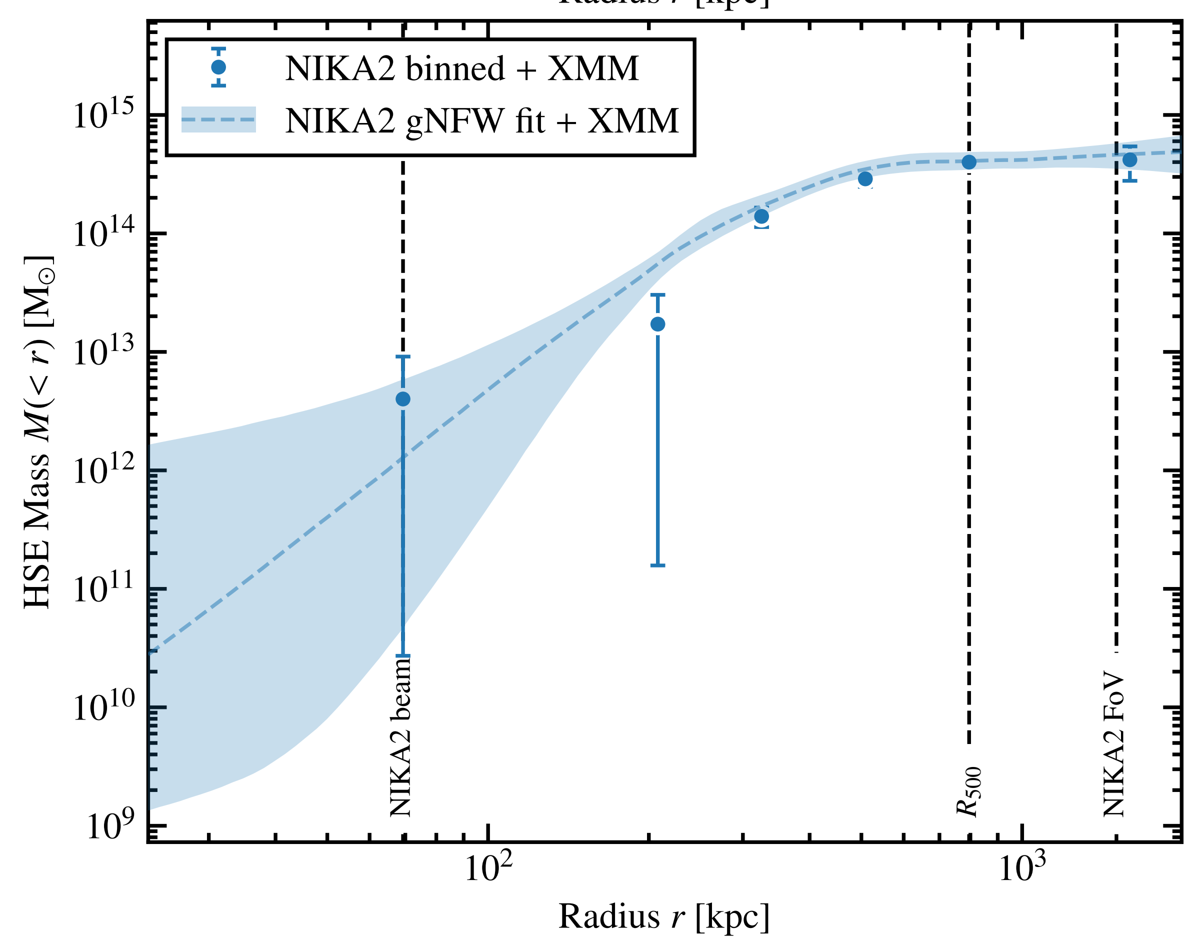
If available X-ray density:

X-ray + SZ profiles combination for further thermodynamical properties:

- Temperature $kT_\mathrm{e} = P_\mathrm{e}/n_\mathrm{e}$
- Entropy $K_\mathrm{e} = P_\mathrm{e} n_\mathrm{e}^{-5/3}$
- HSE mass $M^\mathrm{HSE}(r) = -\dfrac{1}{G\mu m_\mathrm{p}} \dfrac{r^2}{n_\mathrm{e}} \dfrac{\mathrm{d}P_\mathrm{e}}{\mathrm{d}r}$

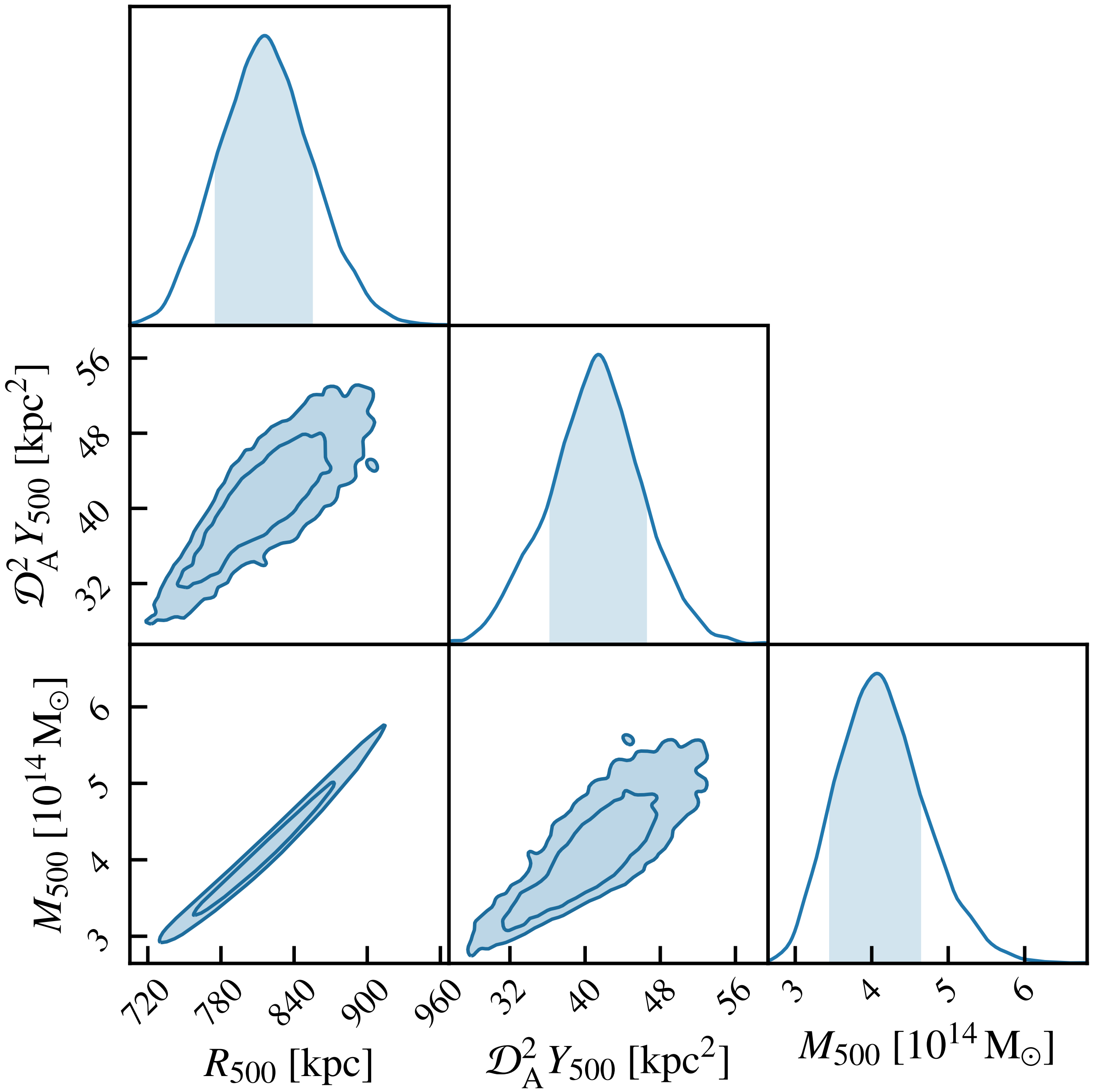






Results: integrated quantities

Florian Kéruzoré
mmUniverse
June 2021



If available X-ray density:

1. Compute overdensity profiles $\delta(r)$ from each mass profile from MCMC chains
2. Solve each profile for $\delta(r) = 500$
   → R_{500} value for each MCMC sample
   → probability distribution for R_{500}
3. For each sample, compute $M(<R_{500}), Y(<R_{500})$
   → M_{500}, Y_{500} values for each MCMC sample
   → probability distribution for M_{500}, Y_{500}

Outline

Florian Kéruzoré
mmUniverse
June 2021

Algorithm

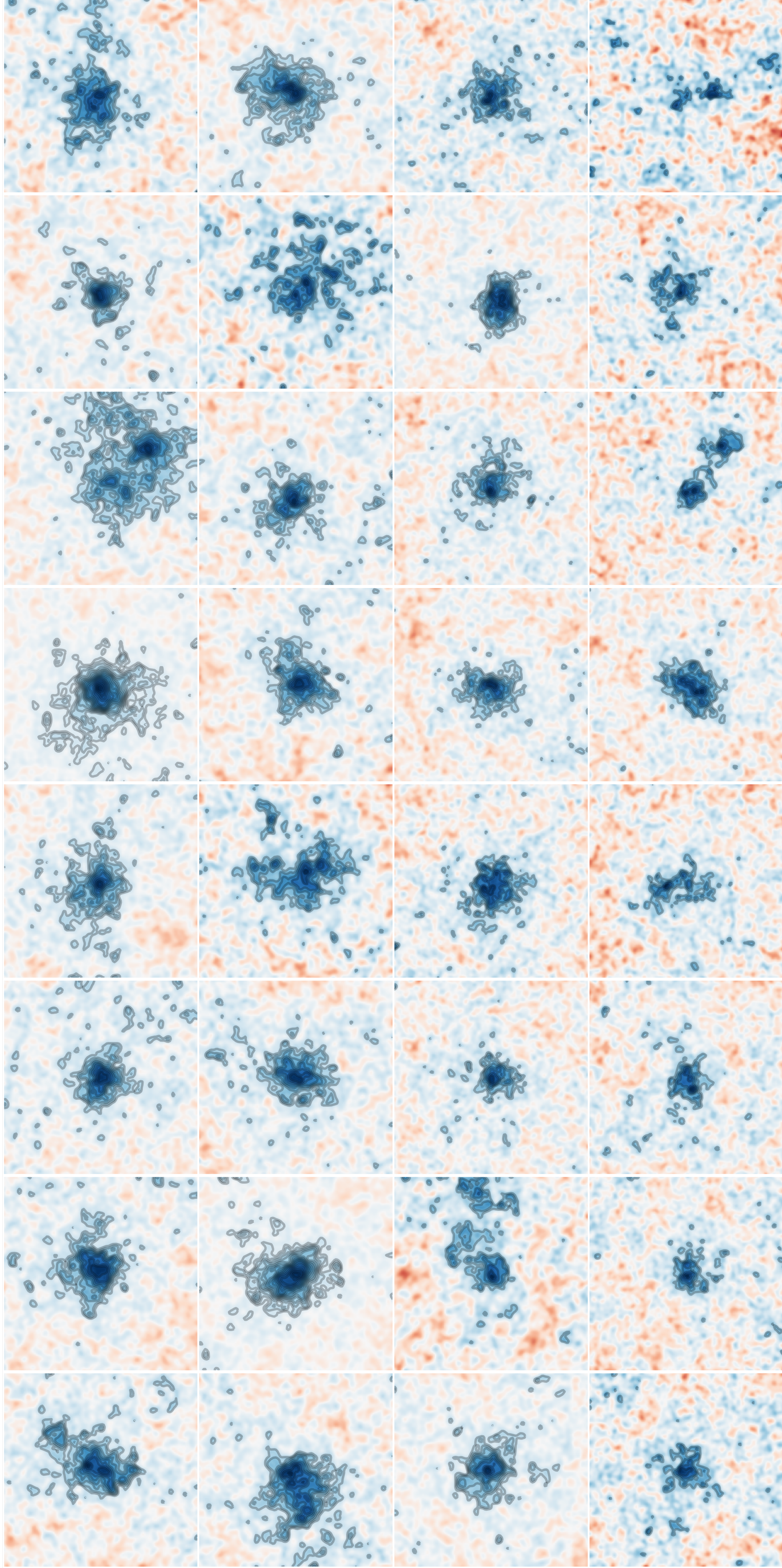
Validation on simulated input & results showcase

Validation on MUSIC hydrodynamical simulations

Conclusions & perspectives

⑱ Scope

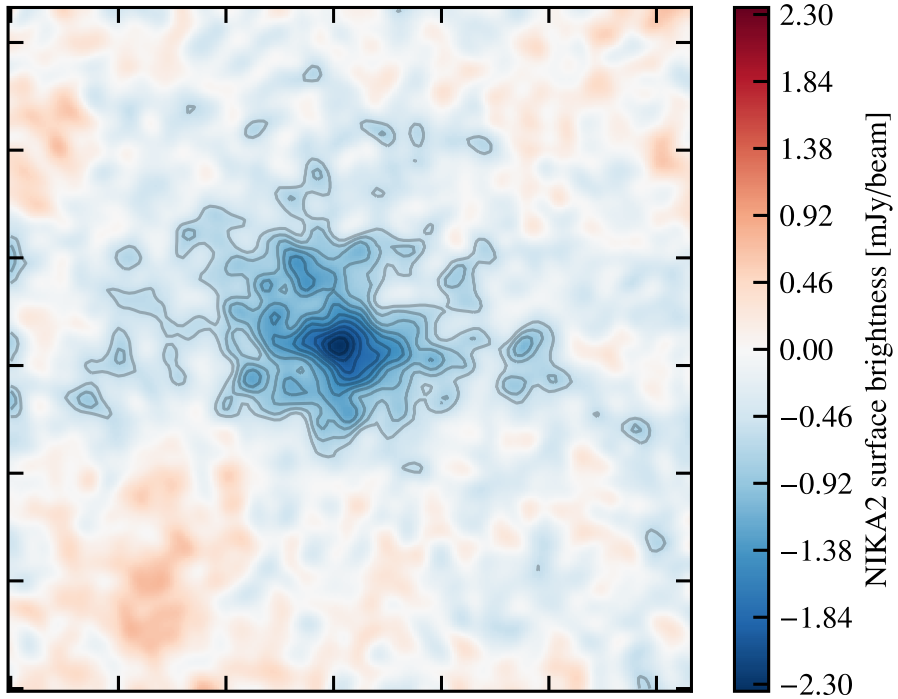
- **Goal:** testing on "real life" clusters
- **Data used:** simulated NIKA2 observations of clusters from hydrodynamical simulations
  - Clusters from the MUSIC simulation (Sembolini+13)
  - 32 clusters forming a sample similar to the NIKA2 LPSZ
  - Maps from Ruppin+19: include cluster SZ signal, PSF filtering, transfer function filtering, correlated noise
- **PANCO2 analysis:** fit the pressure profile of the ICM
  - Binned pressure profile
  - From the center of the map (some clusters are off-centered)
  - Taking into account noise covariance matrix

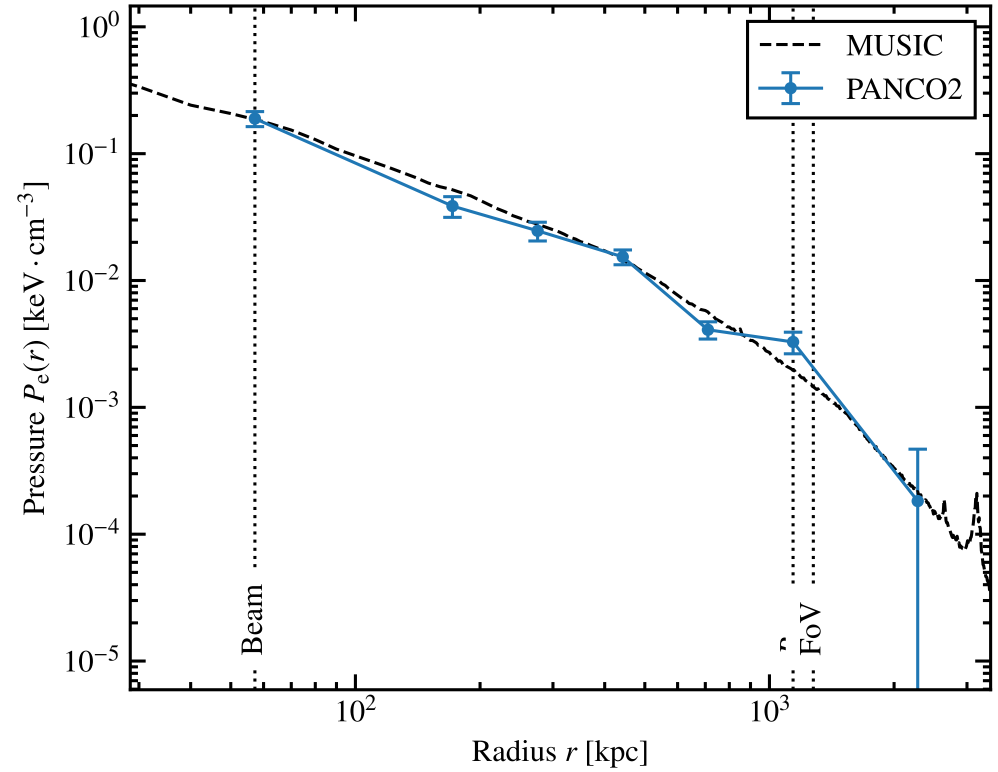


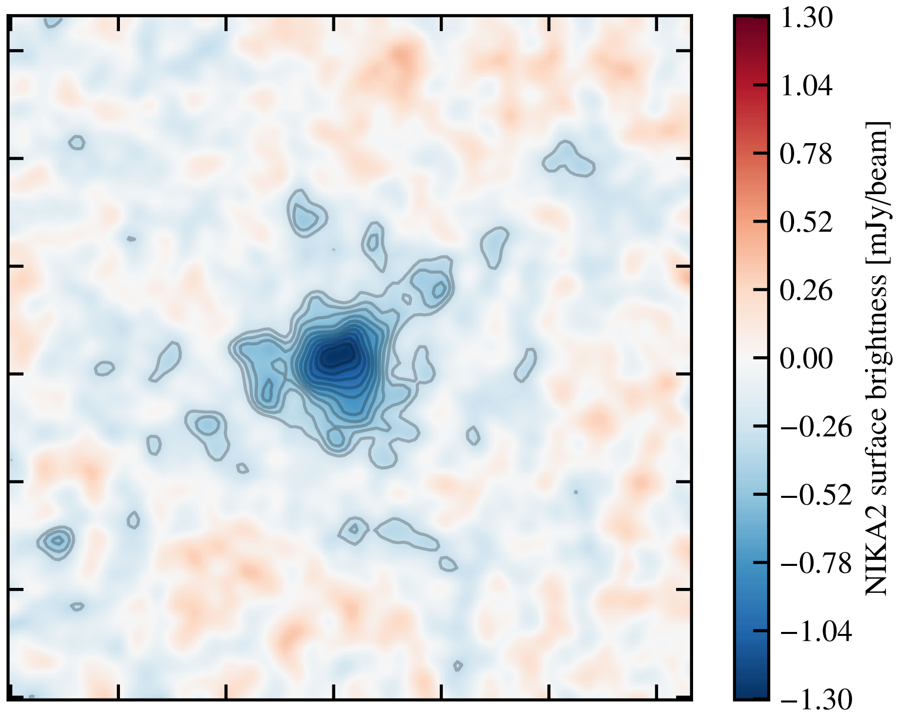
19

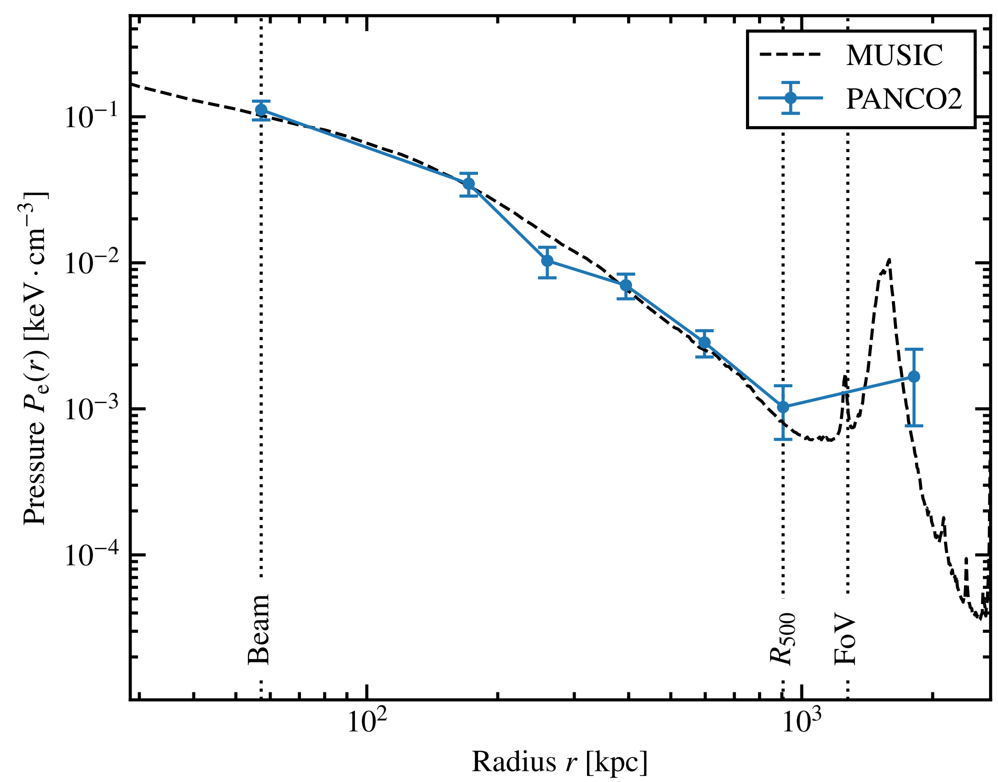
Results examples: $z = 0.54$

Florian Kéruzoré
mmUniverse
June 2021





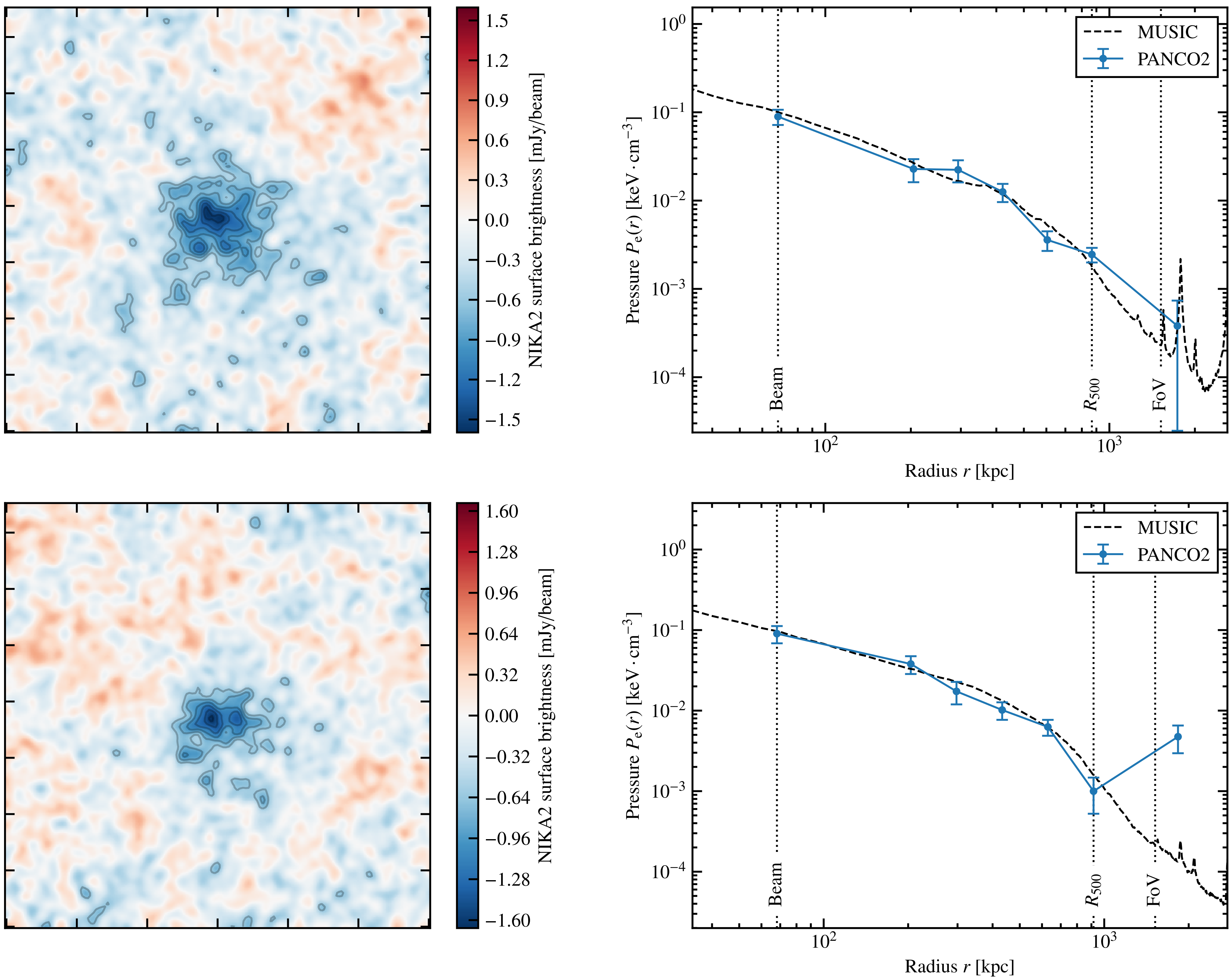




Recovered profiles in agreement with true profiles

Results examples: $z = 0.82$

Florian Kéruzoré
mmUniverse
June 2021



Recovered profiles in agreement with true profiles

Outline

Florian Kéruzoré
mmUniverse
June 2021

Algorithm

Validation on simulated input & results showcase

Validation on MUSIC hydrodynamical simulations

Conclusions & perspectives

Summary

- **A new software** to fit pressure profiles on SZ maps
  - Takes into account mm data features through forward-modeling
  - Fit done using bayesian MCMC for inference on ICM physical properties
  - Highly customizable analyses
  - Very efficient: most simple analysis ~10 minutes
- **Validated** on simulations
  - Excellent agreement between truth and results for simple simulations
  - Same for complex hydrodynamical simulations
- **Official pipeline** for the NIKA2 SZ Large Program
  - Routinely used for LPSZ analyses
  - Efficient enough for cluster sample analyses
- **Public release:** stay tuned!
  - Generalization to any other instrument in progress
  - Will come with extensive documentation and an accompanying paper

Many thanks to all beta-testers for their reports and the resulting improvements