Capturing non-Gaussian cosmic information with Neutral Hydrogen one-point statistics

The Non-Gaussian Universe: Advanced Statistical Inference in the Era of Next-Generation Cosmological Surveys

Speaker: Bernhard Vos Ginés (Βερνάρδ), Postdoctoral researcher

Collaborators: Cora Uhlemann

Funded by the European Union

European Research Council

UNIVERSITÄT BIELEFELD

18/06/2026

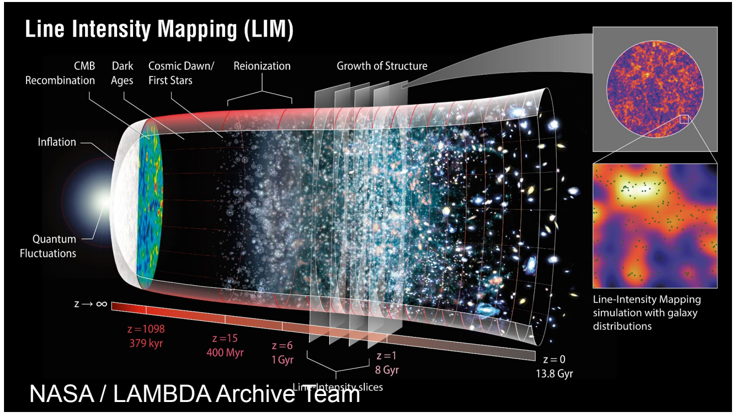
HI Intensity Mapping

The Intensity Mapping technique efficiently maps HI: no need to detect individual galaxies!

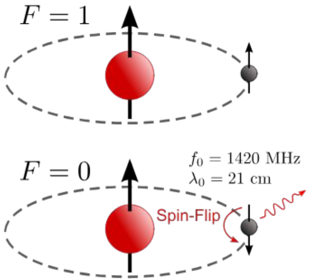
SKAO expects to map HI up to z = 6

Neutral hydrogen (HI) can be mapped through its 21 cm hyperfine transition

Suitable summary statistics & Need to take into account observational systematics



NASA / LAMBDA Archive Team



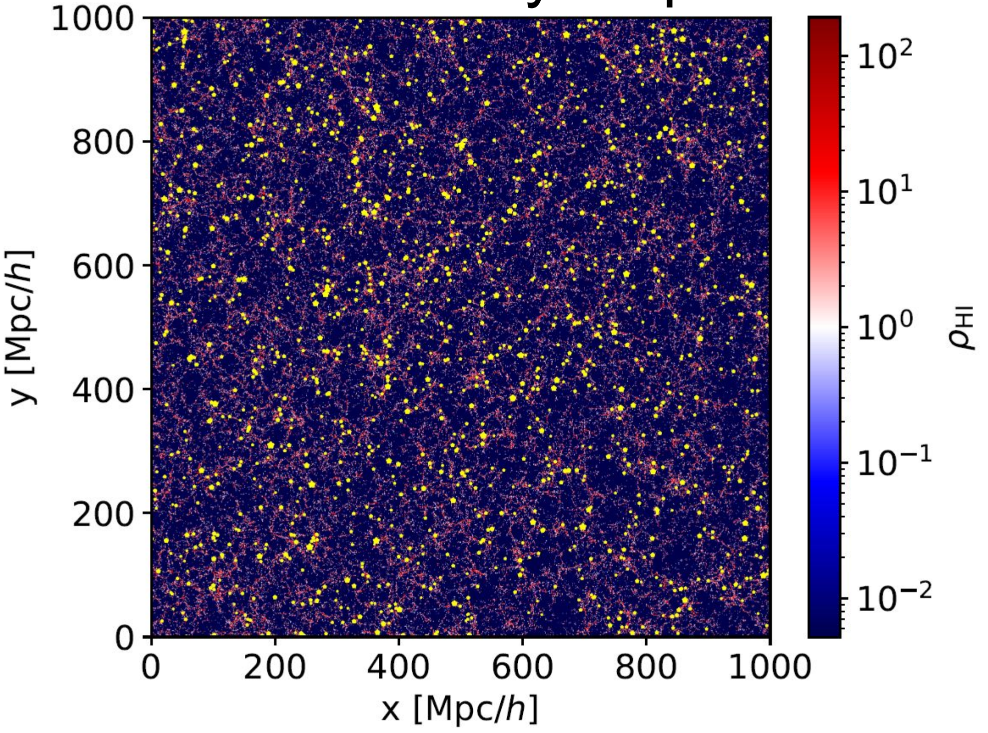
The input catalogues

- UNIT dark matter cosmological simulations Chuang, 2019
- Semi-Analytic Model of Galaxy Evolution (SAGE) Croton, 2016
- HI prescription as a function of the cold gas distribution Keres et al., 2003, Zwaan et al., 2005
- HI Intensity map -> SKA1MID-like

$$M_{\mathrm{H\,I}} = f_{\mathrm{H}} \cdot \left(1 - \frac{R_{\mathrm{mol}}}{R_{\mathrm{mol}} + 1}\right) \cdot M_{\mathrm{CG}}$$

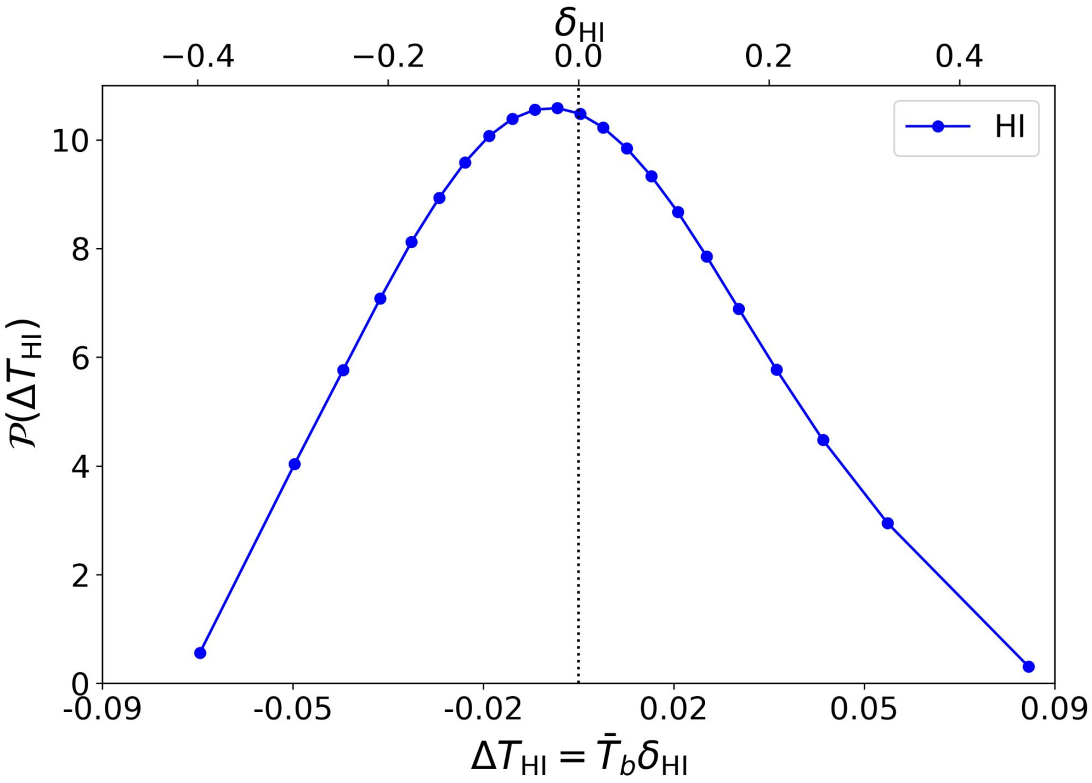
$$R_{\mathrm{mol}} = 0.4$$

HI Intensity Map



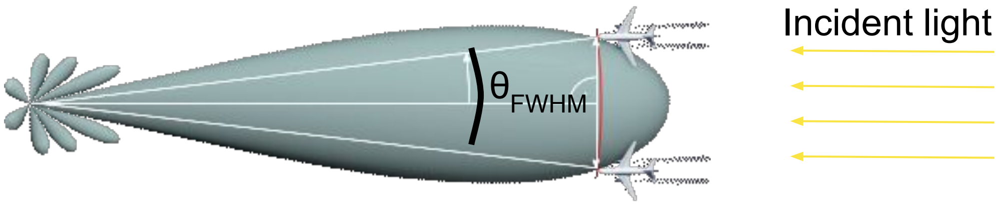
HI PDF without systematics

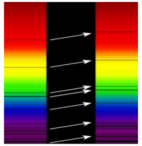
R = 31 Mpc/h



Telescope beam: angular smoothing

SKA beam effect





Angular gaussian smoothing simulated

$$B(k_\perp) = \mathcal{F}\left[\mathcal{N}(\mu, \sigma^2 = R^2_{\rm beam})\right](k_\perp) = \exp\left[-\frac{R^2_{\rm beam}k^2_\perp}{2}\right]$$

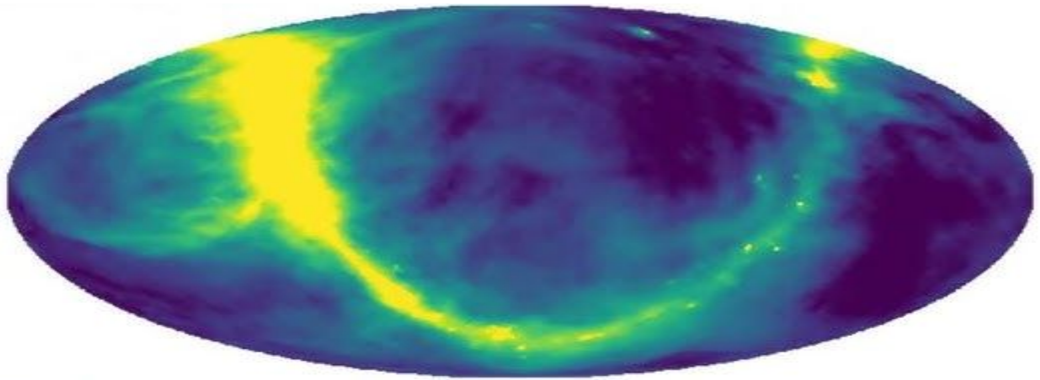
where R_{beam} is the transverse smoothing scale Villaescusa-Navarro, 2017

$$R_{\rm beam} = \; 38.45\, h^{-1}{\rm Mpc}$$

Foreground contamination

21-cm foregrounds

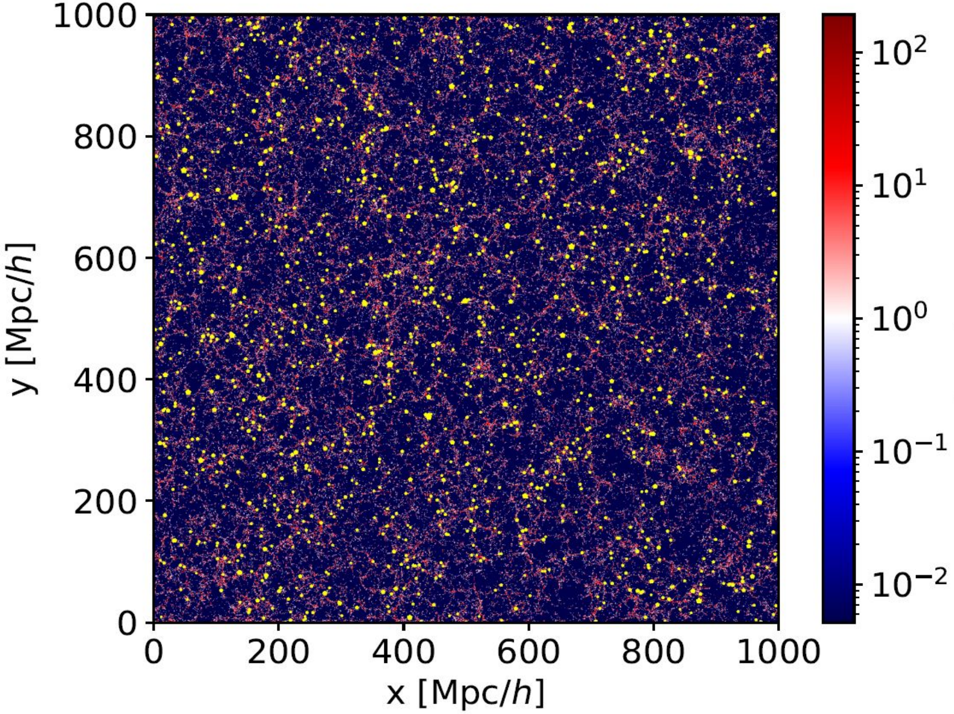
Radio signals several orders of magnitude higher than the cosmological signal but they are smooth in frequency! Soares, 2021

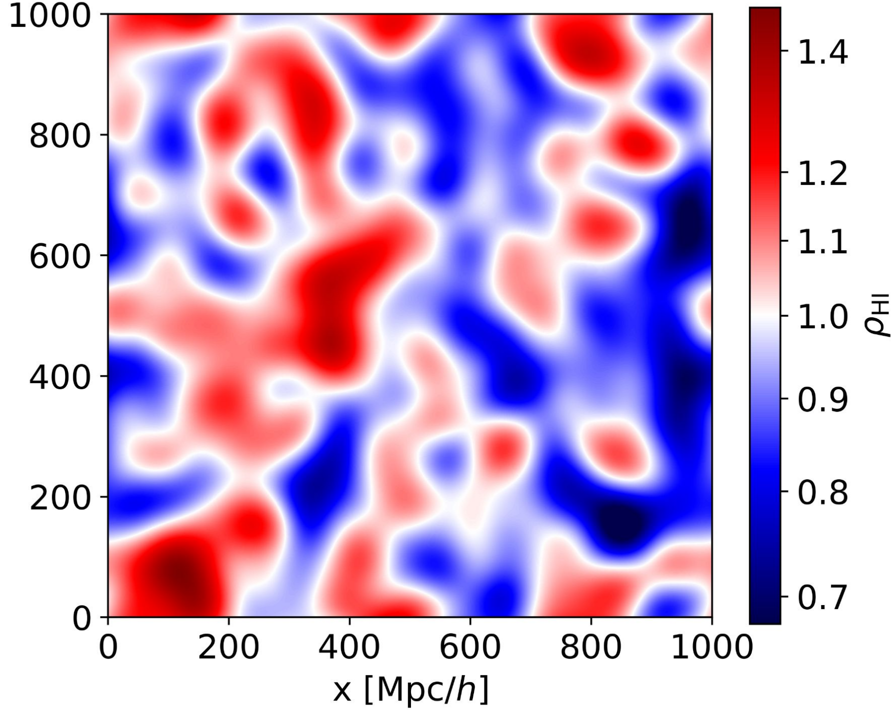


Spinelli, 2023

$$f(k_{||}) = 1 - \exp\left[-\left(\frac{k_{||}}{k_{\rm FG}}\right)^2\right]$$

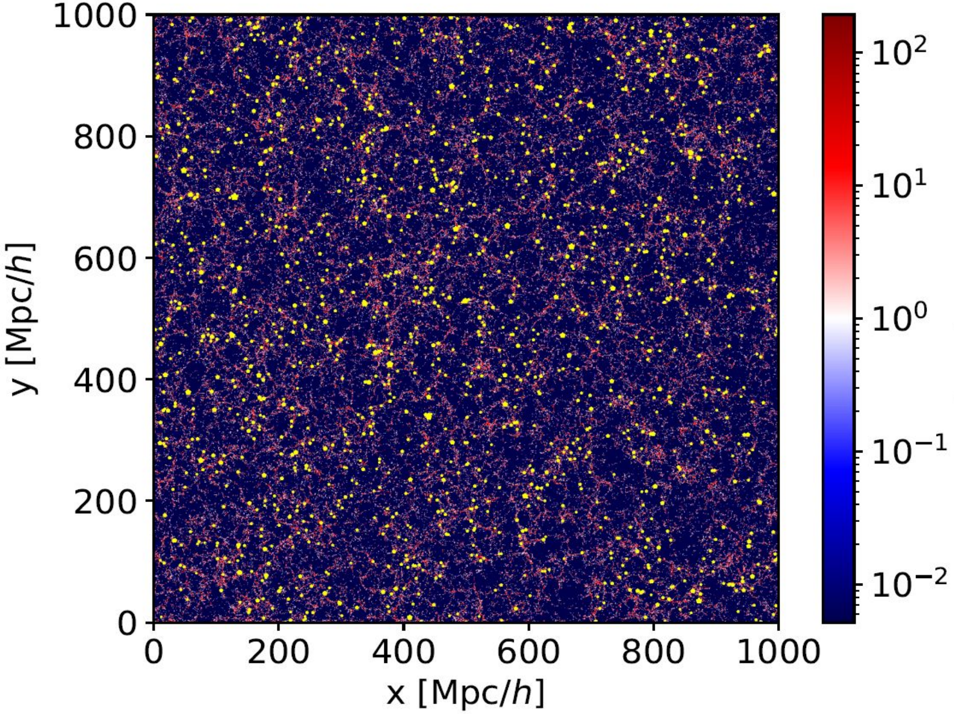
HI Intensity Maps



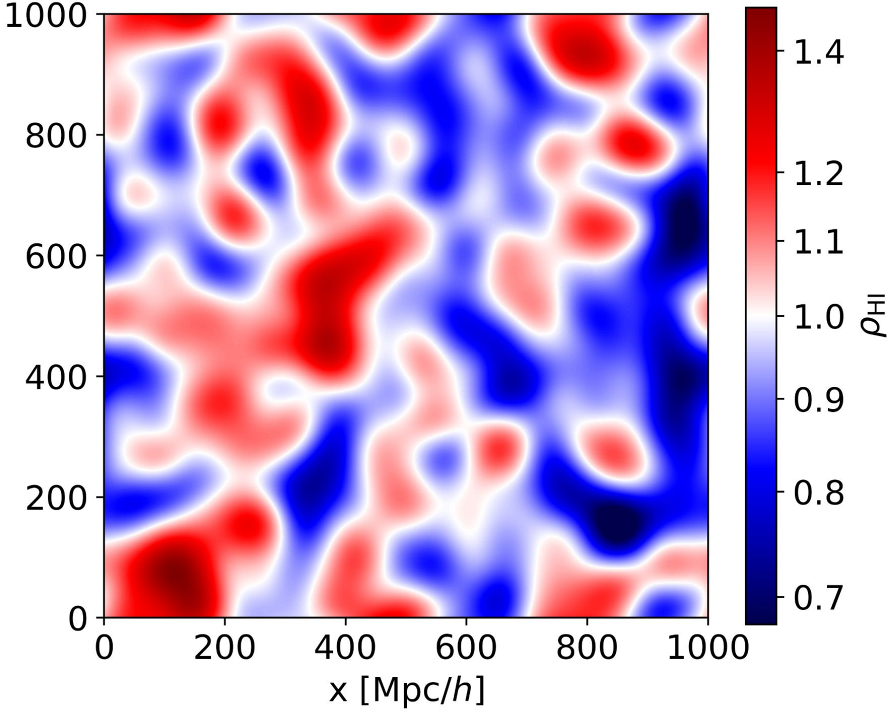


HI Intensity Maps

No systematics



With systematics

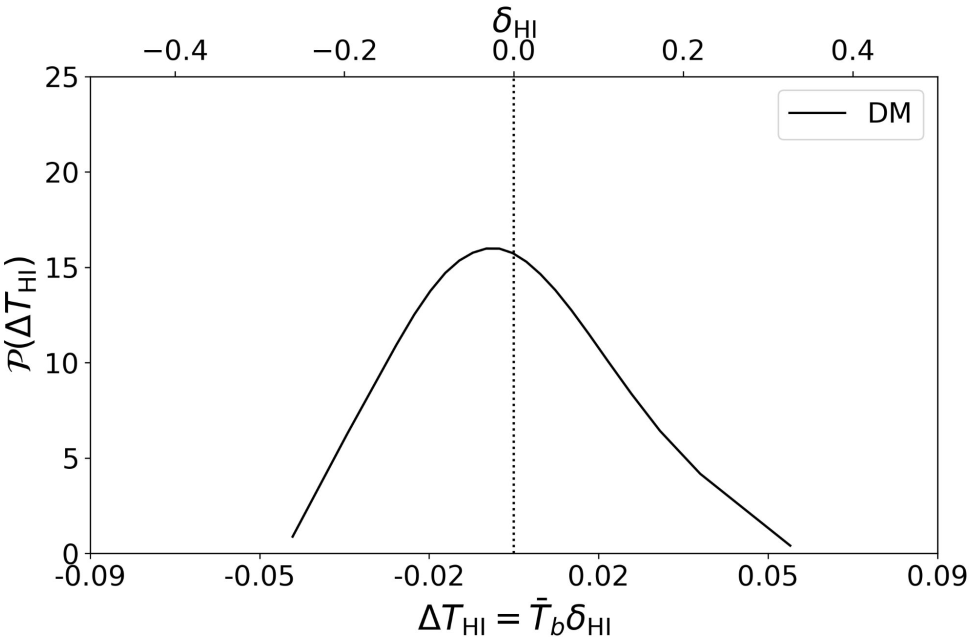


HI PDF modeling

1. We use CosMomentum to predict the matter PDF at scale R [Friedrich et al. 2026]

$$\Psi(\delta_m) = \frac{\sigma_{\mathrm{L}}^2(z,R)}{\sigma_{\mathrm{NL}}^2(z,R)} \frac{\delta_{\mathrm{L}}^{\mathrm{SC}}(\delta_m)^2}{2\sigma_{\mathrm{L}}^2(z,r)}$$

$$P_R(\delta_m) \propto e^{-\Psi(\delta_m;R)}$$



HI PDF modeling

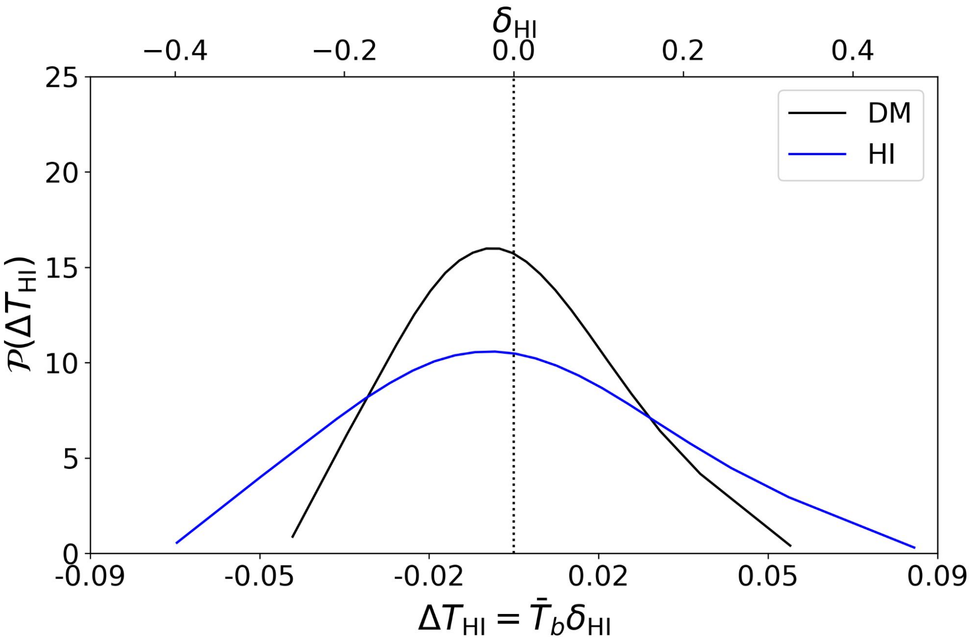
1. Matter PDF [Friedrich et al. 2026]
2. HI-matter connection through HI conditional mean and variance [McCarthy Gould et al. 2024]

$$\mathcal{P}(\delta_{\rm HI}) = \int \mathcal{P}(\delta_{\rm HI}, \delta_m)\, d\delta_m = \int \mathcal{P}(\delta_{\rm HI}|\delta_m)\, \mathcal{P}(\delta_m)\, d\delta_m$$

$$\mathcal{P}(N_{\rm HI}|\delta_m) = \mathcal{P}_{\rm P}\left(\tilde{N}_{\rm HI} = \frac{N_{\rm HI}}{\alpha(\delta_m)}; \overline{\tilde{N}}_{\rm HI} = \frac{\overline{N}_{\rm HI}(\delta_m)}{\alpha(\delta_m)}\right) \frac{1}{\alpha(\delta_m)}$$

$$\langle \delta_{\rm HI}|\delta_{\rm m}\rangle = b_1\delta_m + \frac{b_2}{2}\left(\delta_m^2 - \sigma_m^2\right)$$

$$\alpha(\delta_m) = \frac{\bar{N}_{\rm HI}\langle \delta_{\rm HI}^2|\delta_{\rm m}\rangle_c}{1 + \langle \delta_{\rm HI}|\delta_{\rm m}\rangle} = \alpha_0 + \alpha_1\delta_m + \alpha_2\delta_m^2$$



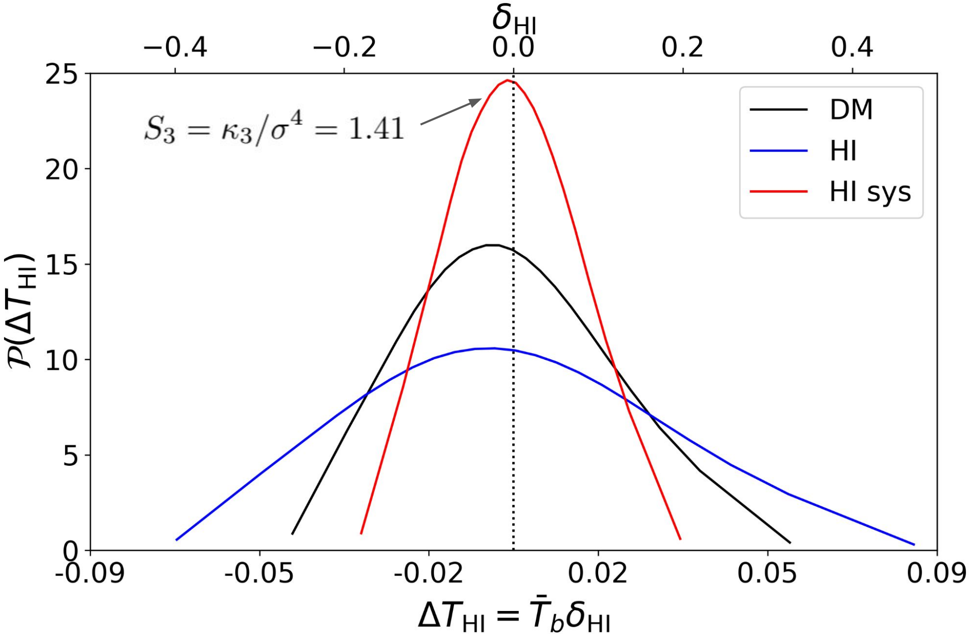
HI PDF modeling

1. We use CosMomentum to predict the matter PDF [Friedrich et al. 2026]
2. DM -> HI [McCarthy Gould et al. 2024]
3. Introduce systematics as variance rescaling. Telescope beam noticeably affects the PDF shape

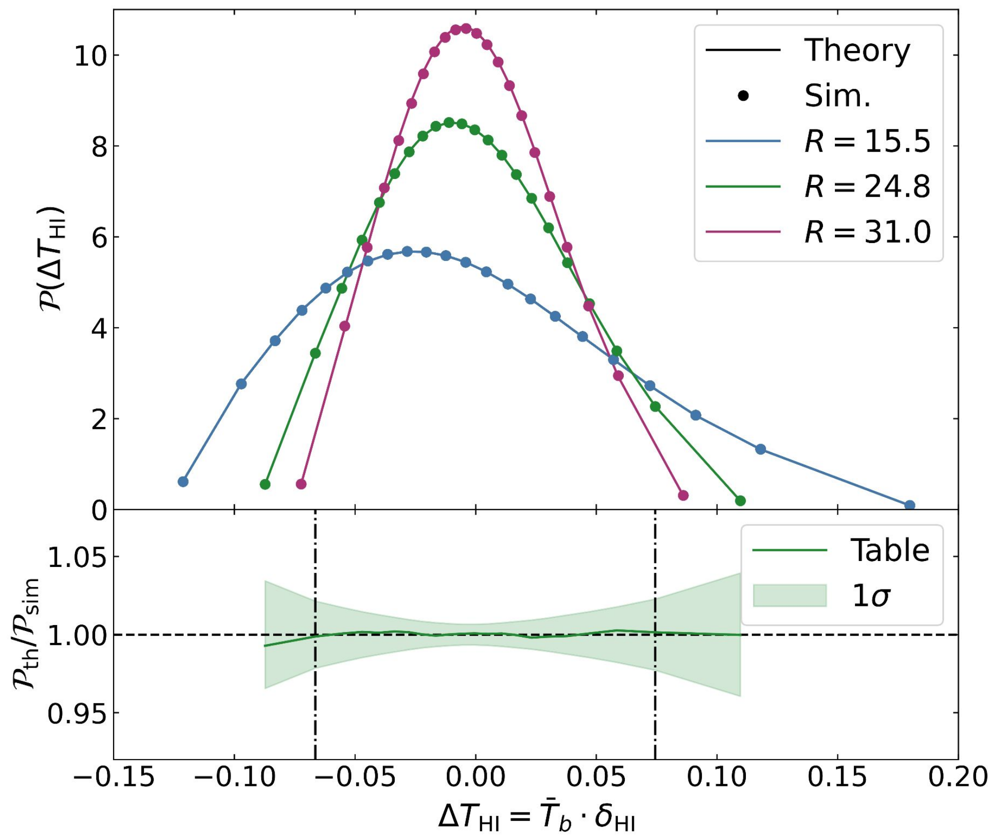
$$\sigma^2_{\mathrm{HI(m)}}(R) = \frac{1}{2\pi^2}\int_{\mathrm{k_F}}^{\mathrm{k_{nyq}}} dk\; k^2 P_{\mathrm{theory(m)}}(k) W^2_{\mathrm{TH}}(kR)\,,$$

$$\sigma^2_{\mathrm{HI(m),sys}}(R) = \frac{1}{2\pi^2}\int_{\mathrm{k_F}}^{\mathrm{k_{nyq}}} dk\; k^2 P_{\mathrm{theory(m)}}(k) W^2_{\mathrm{TH}}(kR) S_{\mathrm{eff}}(k)$$

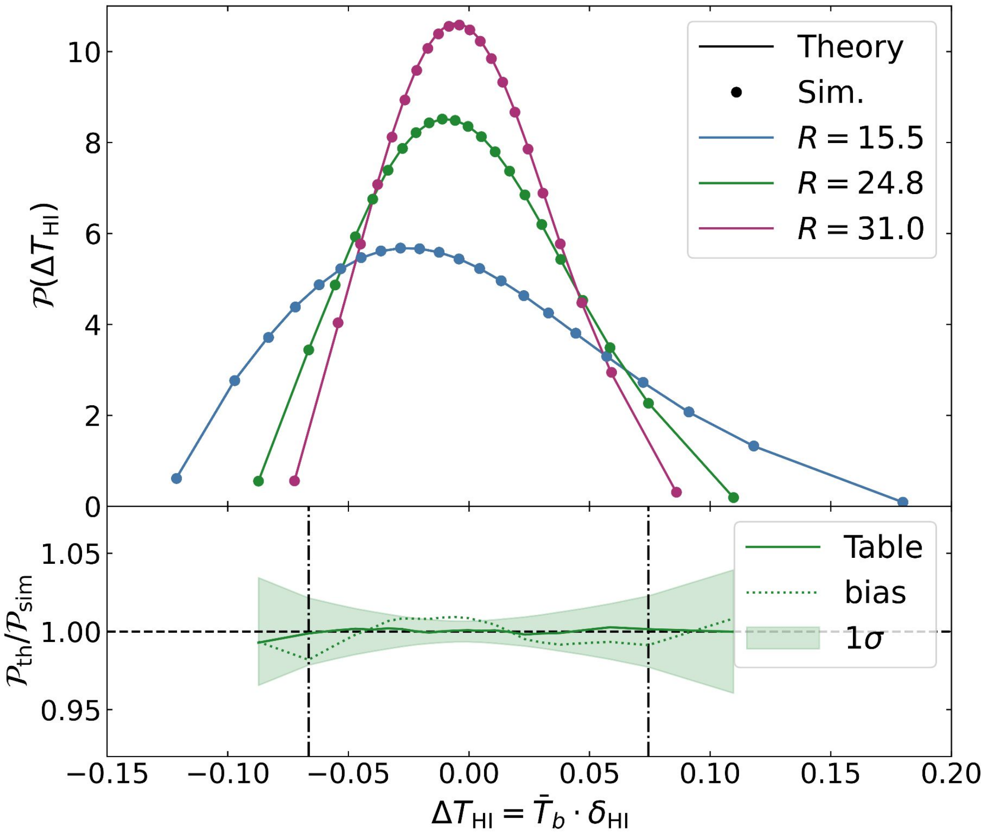
$$S_{\mathrm{eff}}(k) = \frac{1}{2}\int_{-1}^{1} B^2(k\sqrt{1-\mu^2}) f^2(k\mu) d\mu$$



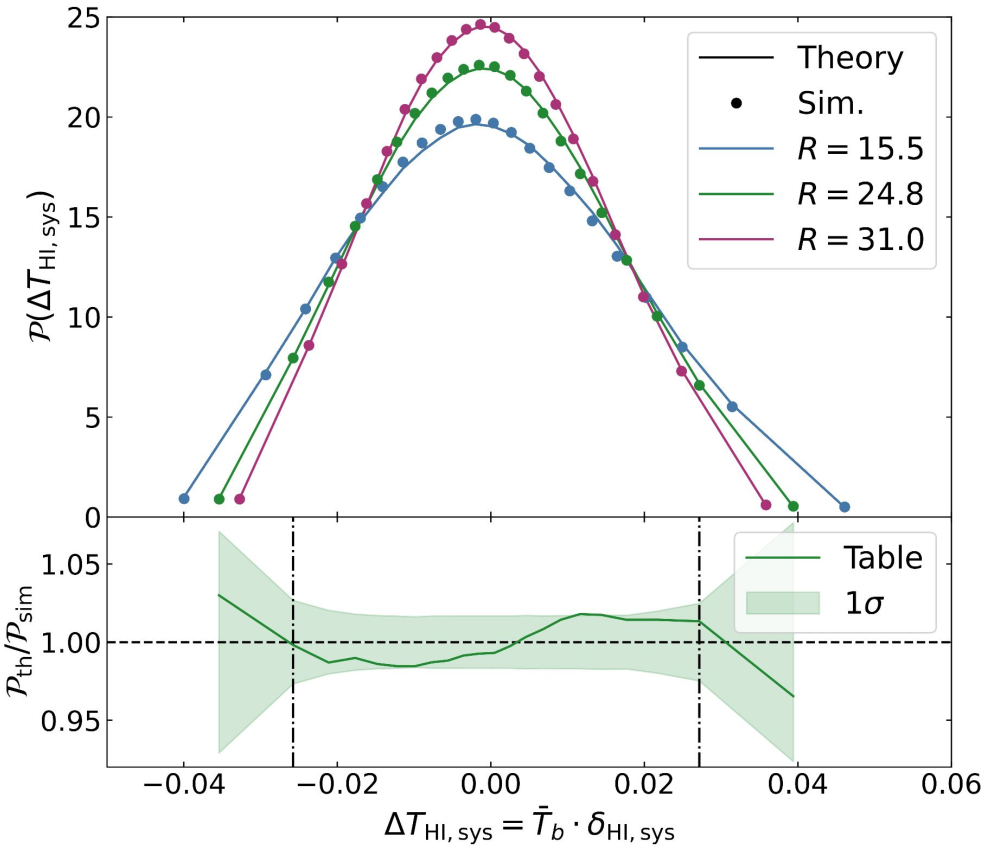
HI PDF (theory vs simulation)



HI PDF (theory vs simulation)

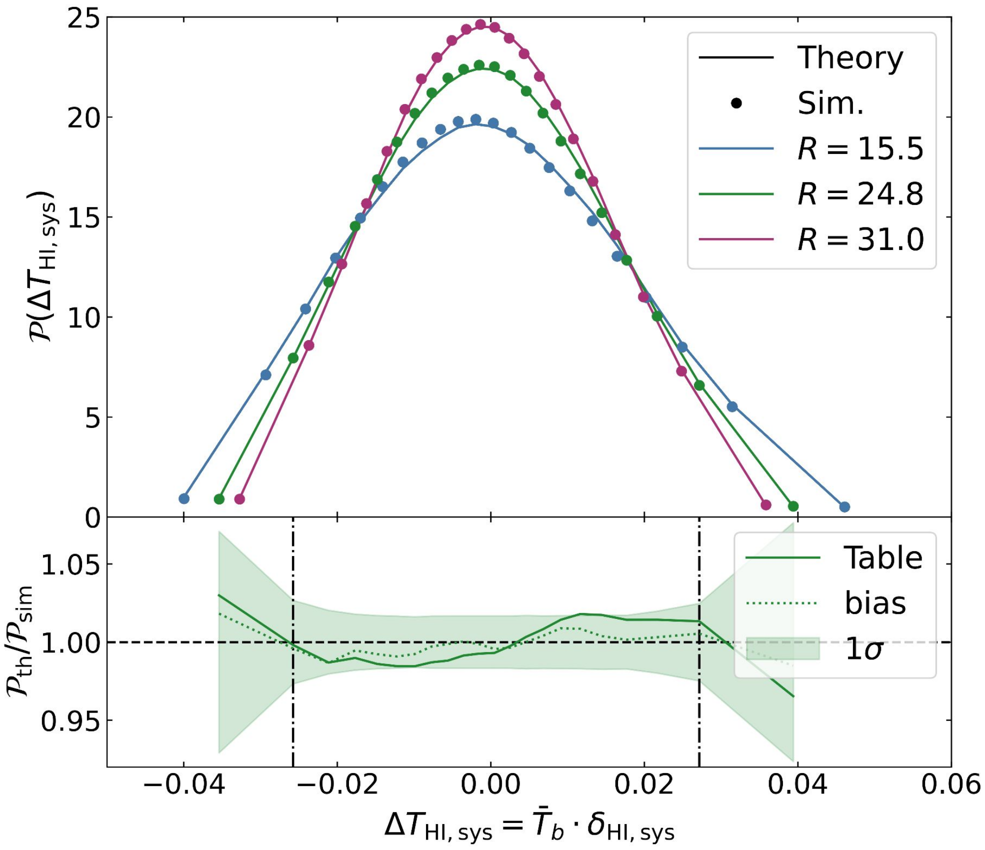


HI PDF with systematics (theory vs simulation)

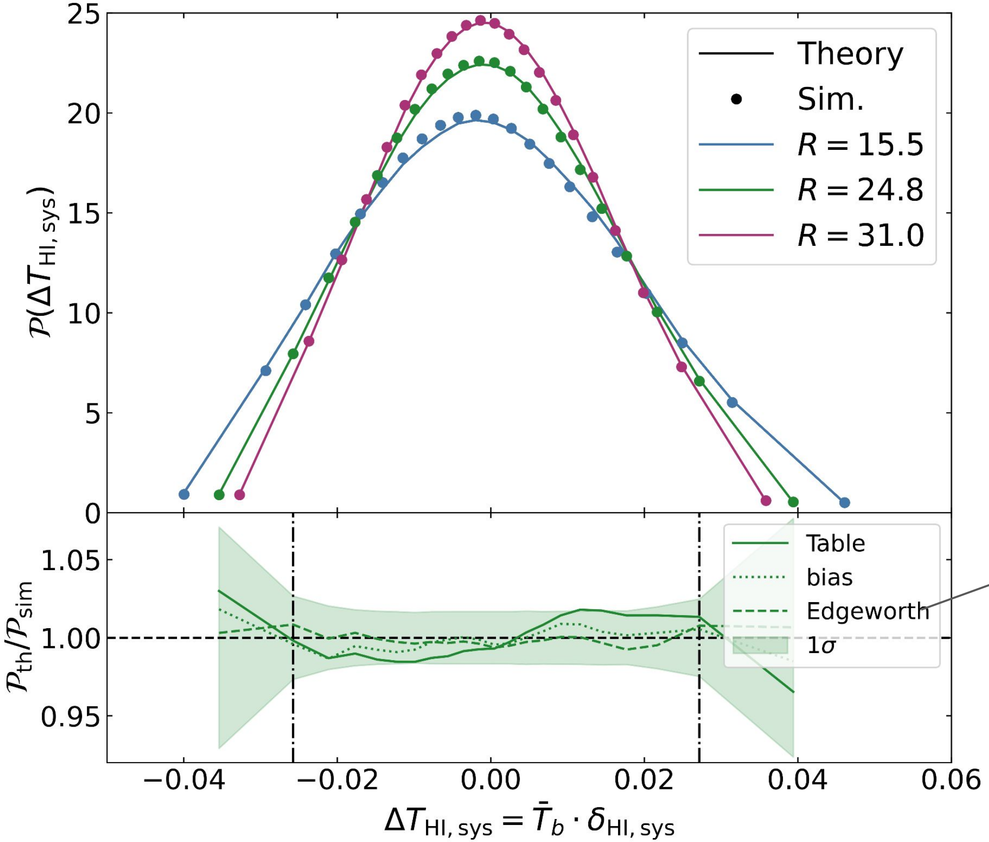


$$\Psi(\delta_m) = \frac{\sigma_{\mathrm{L}}^2(z,R)}{\sigma_{\mathrm{NL}}^2(z,R)} \frac{\delta_{\mathrm{L}}^{\mathrm{SC}}(\delta_m)^2}{2\sigma_{\mathrm{L}}^2(z,r)}$$

HI PDF with systematics (theory vs simulation)



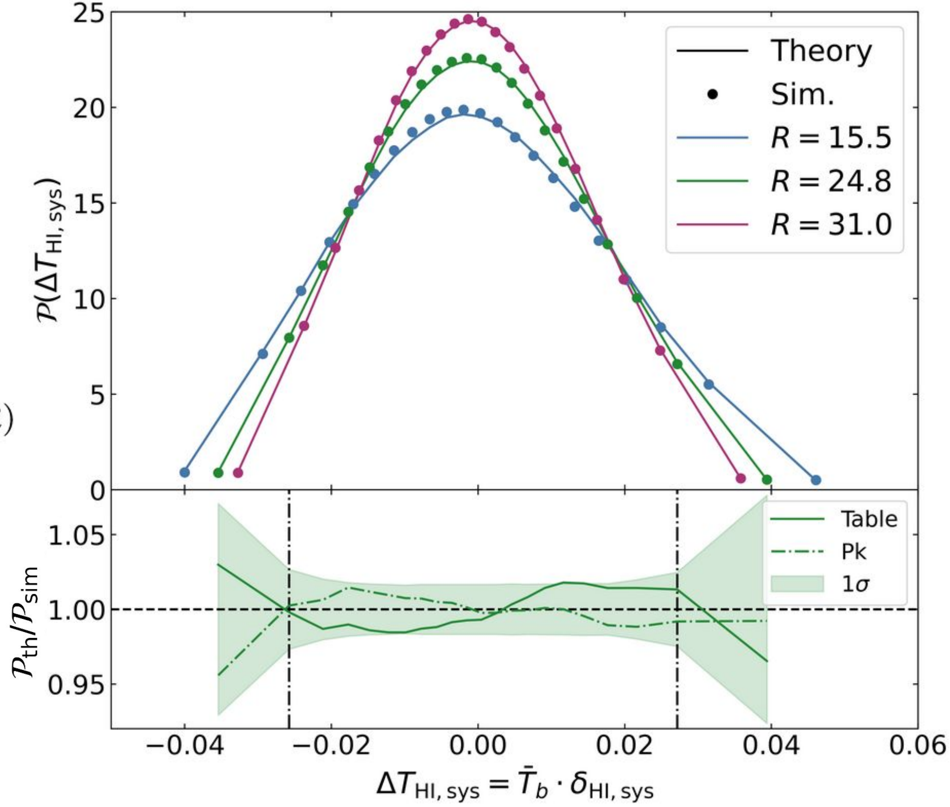
HI PDF with systematics (theory vs simulation)



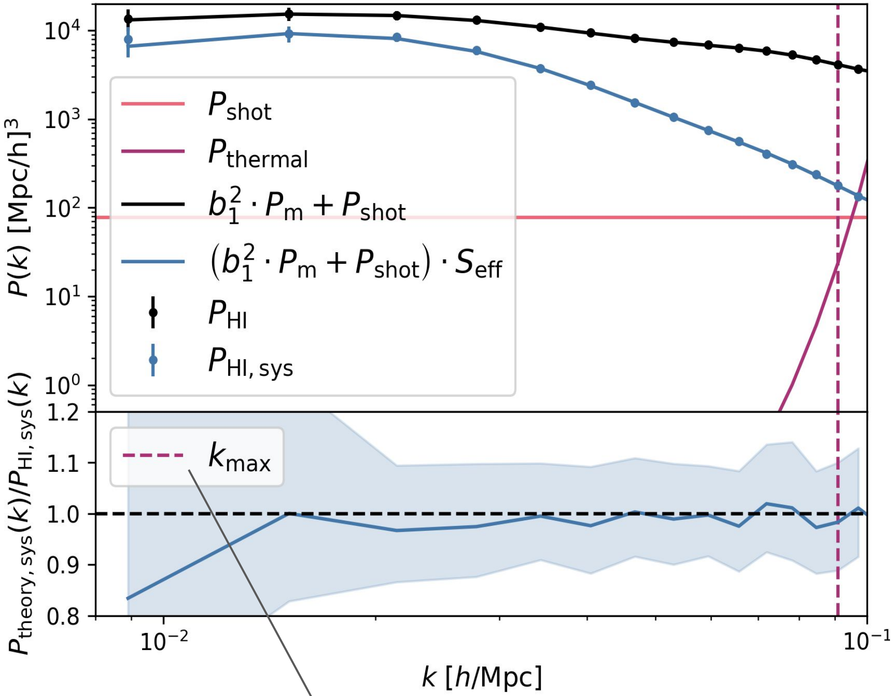
$$\mathcal{P}(\delta_{\mathrm{HI}}) = \int \mathcal{P}(\delta_{\mathrm{HI}}|\delta_m)\,\mathcal{P}(\delta_m)\,d\delta_m$$

HI PDF with systematics (theory vs simulation)

$$P_R(\delta_m) \propto e^{-\Psi(\delta_m;R)}$$



Power spectrum: simulations vs modelling



$$S_{\rm eff}(k) = \frac{1}{2}\int_{-1}^{1} B^2(k\sqrt{1-\mu^2}) f^2(k\mu) d\mu$$

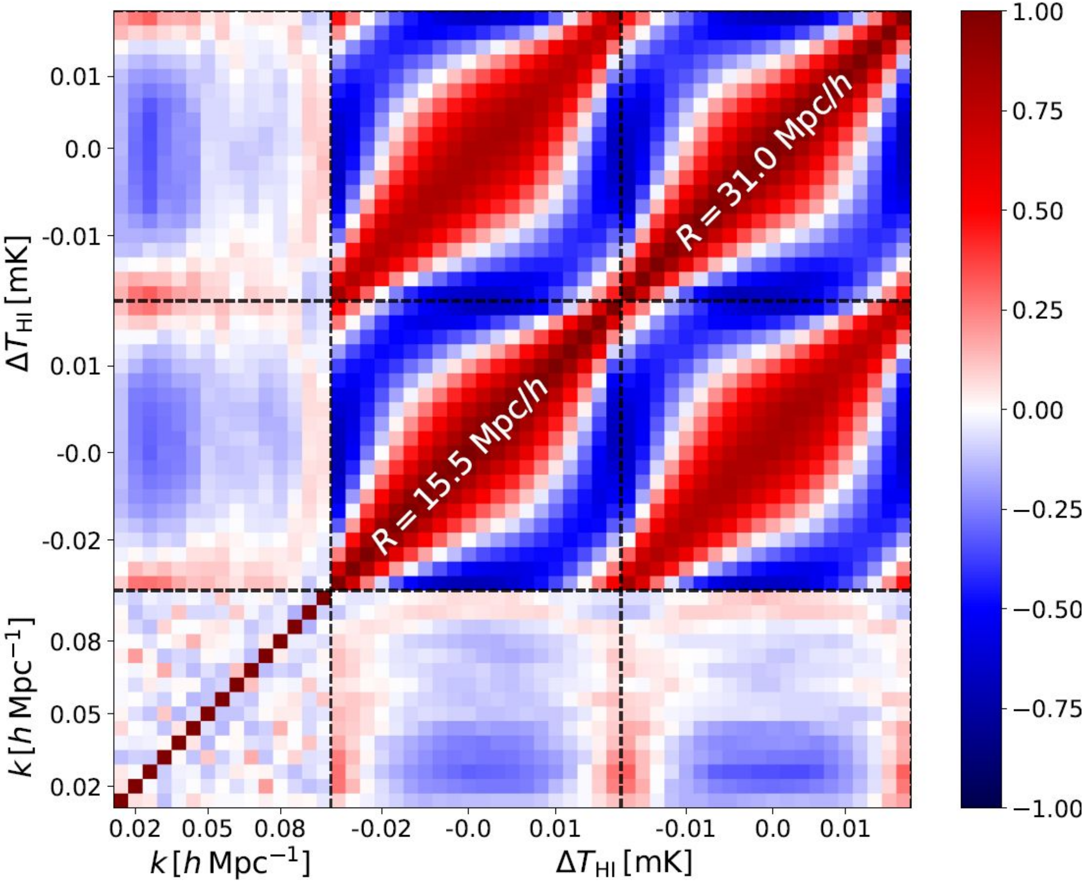
$$k_{\rm max} = \max\{k \in \mathbb{R}^+ | 3P_{\rm thermal}(k) < P_{\rm HI,sys}(k)\} = 0.091\ h/{\rm Mpc}\,.$$

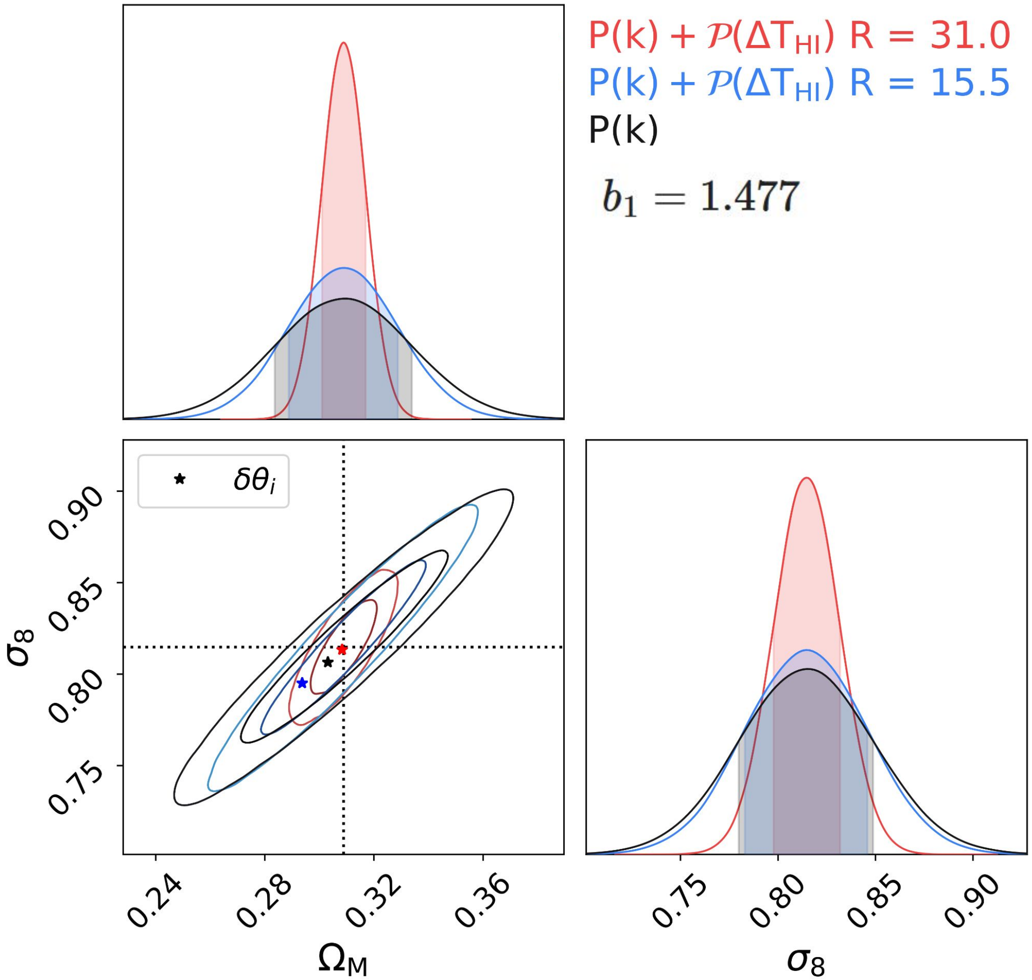
Jackknife Correlation matrix

4 boxes

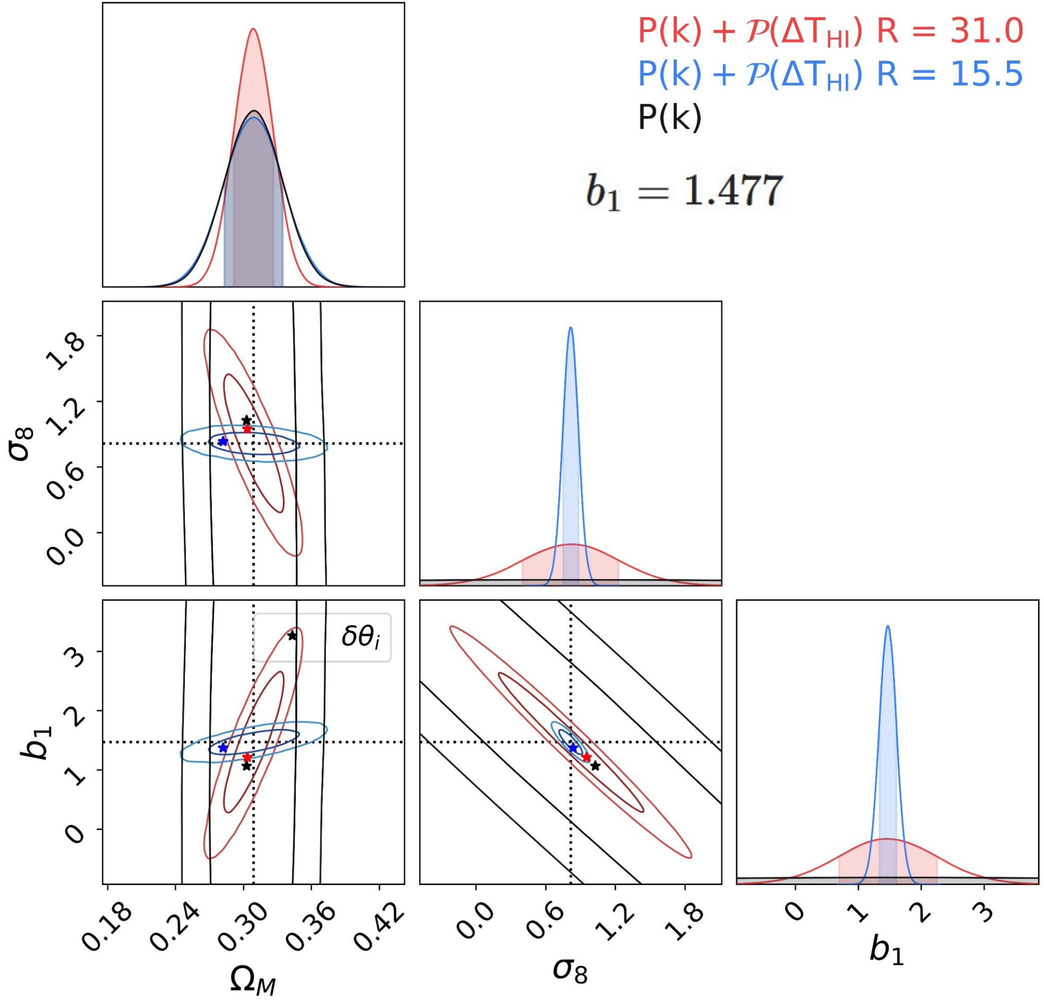
125 regions

$$\mathcal{P}_i = \mathcal{P}_i\left(\frac{\Delta T_{\mathrm{HI}}+1}{\langle \Delta T_{\mathrm{HI}}\rangle_i + 1} - 1\right)$$





$P(k) + \mathcal{P}(\Delta T_{HI})$ R = 31.0
$P(k) + \mathcal{P}(\Delta T_{HI})$ R = 15.5
P(k)
$b_1 = 1.477$
$\delta\theta_i$
0.90
0.85
0.80
0.75
σ_8
0.24
0.28
0.32
0.36
Ω_M
0.75
0.80
0.85
0.90
σ_8



P(k) + $\mathcal{P}(\Delta T_{HI})$ R = 31.0
P(k) + $\mathcal{P}(\Delta T_{HI})$ R = 15.5
P(k)
$b_1 = 1.477$
$\delta\theta_i$
σ_8
b_1
Ω_M

Conclusions

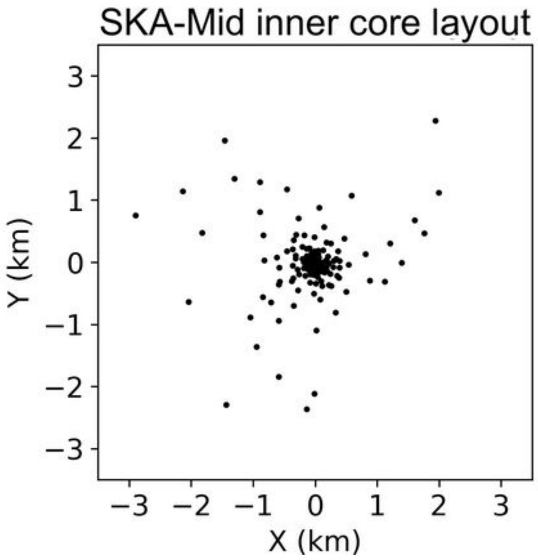
- We produce SKA1-MID-like HI Intensity Maps on a high-resolution simulation (UNIT) including telescope beam and foreground-removal observational systematics + thermal noise.
- We provide a theoretical model for the HI PDF.
- Observational systematics affect the 1-point statistics (PDF) shape, specially the telescope beam.
- The PDF helps to reduce uncertainties and degeneracies when is added to the P(k) for HI Intensity Mapping surveys.

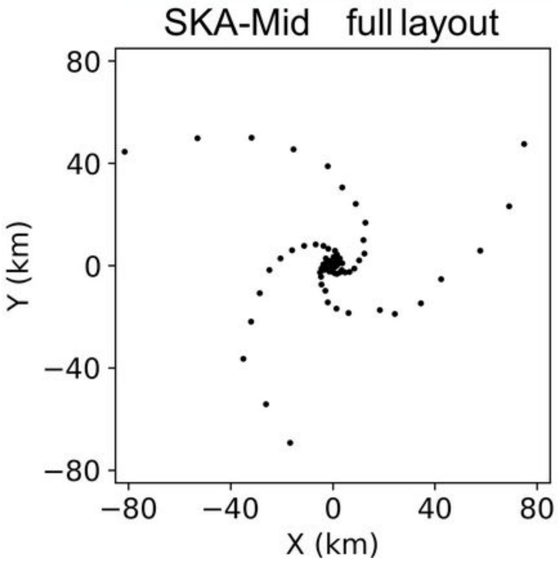

ευχαριστώ!

Back-up slides

SKA-MID

133 15-m SKA dishes and 64 13.5-m Meerkat dishes at the Karoo Radio Astronomy Reserve in the Northern Cape of South Africa. The core will be composed of around 50% of the dishes, randomly distributed within 2 km. The remaining dishes are spaced logarithmically on three spiral arms providing baselines out to 150 km.





Foregrounds

For the relevant radio frequencies it should be possible to describe GSE as a power-law in frequency with a spectral index beta = -2.8

Furthermore, GSE is **partially linearly polarised.**

Its polarised part will be affected by the Faraday effect, a rotation of the polarization angle caused by the Galactic magnetic field and the optically thick interstellar medium.

Foreground removal: you assume your observed data is a **linear mixture** of several independent sources:

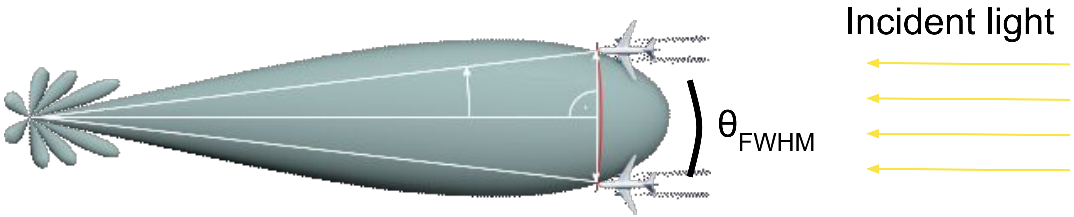
X=AS

- **X** → observed data (e.g., sky maps at different frequencies, EEG channels)
- **S** → unknown independent source signals (cosmological signal + foregrounds + noise)
- **A** → unknown mixing matrix

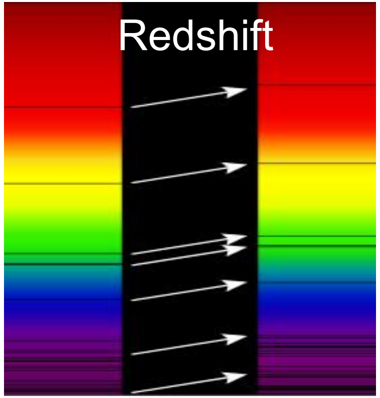
ICA tries to recover **S** and **A**, given only **X**

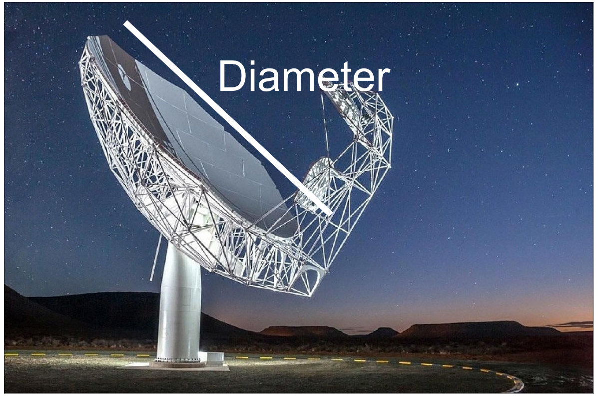
Telescope beam: angular smoothing

SKA telescope beam effect



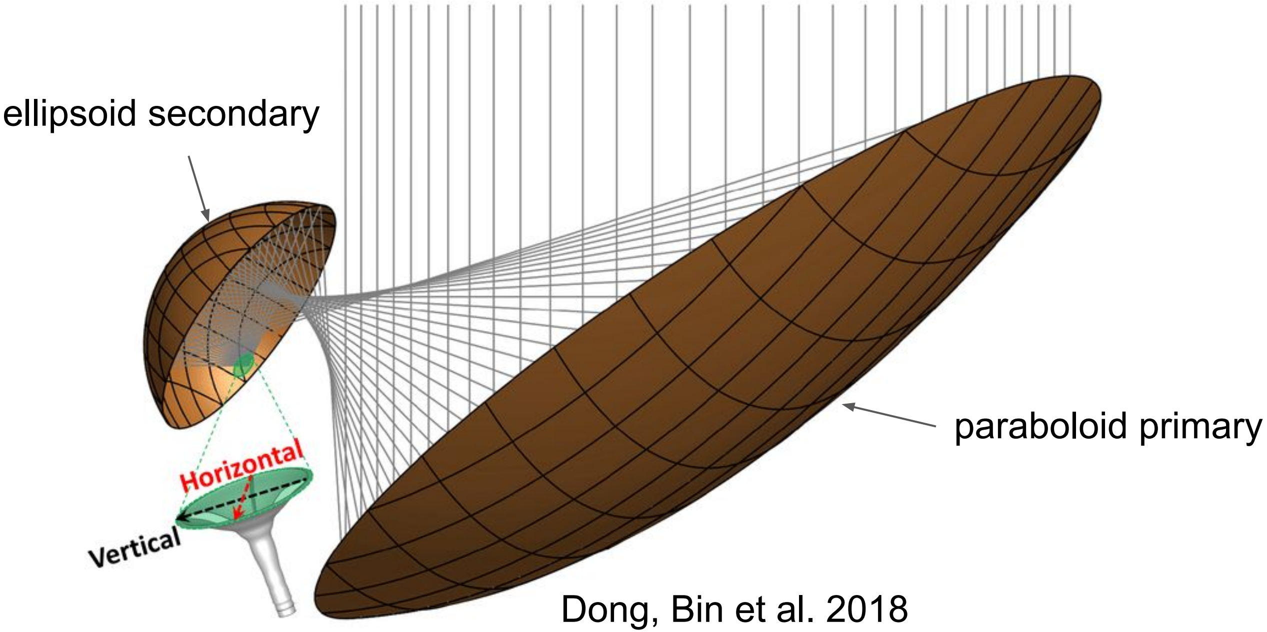
$$\theta_{\mathrm{FWHM}} = \lambda / D_{\mathrm{dish}} \simeq 0.8(1+z)\,\mathrm{deg}$$





SKA1-MID parabolic antennas

The SKA dish geometry: an offset-Gregorian dual-reflector antenna



Dong, Bin et al. 2018

1. Beam characterization

They measure the full beam pattern (main lobe + sidelobes), often using:

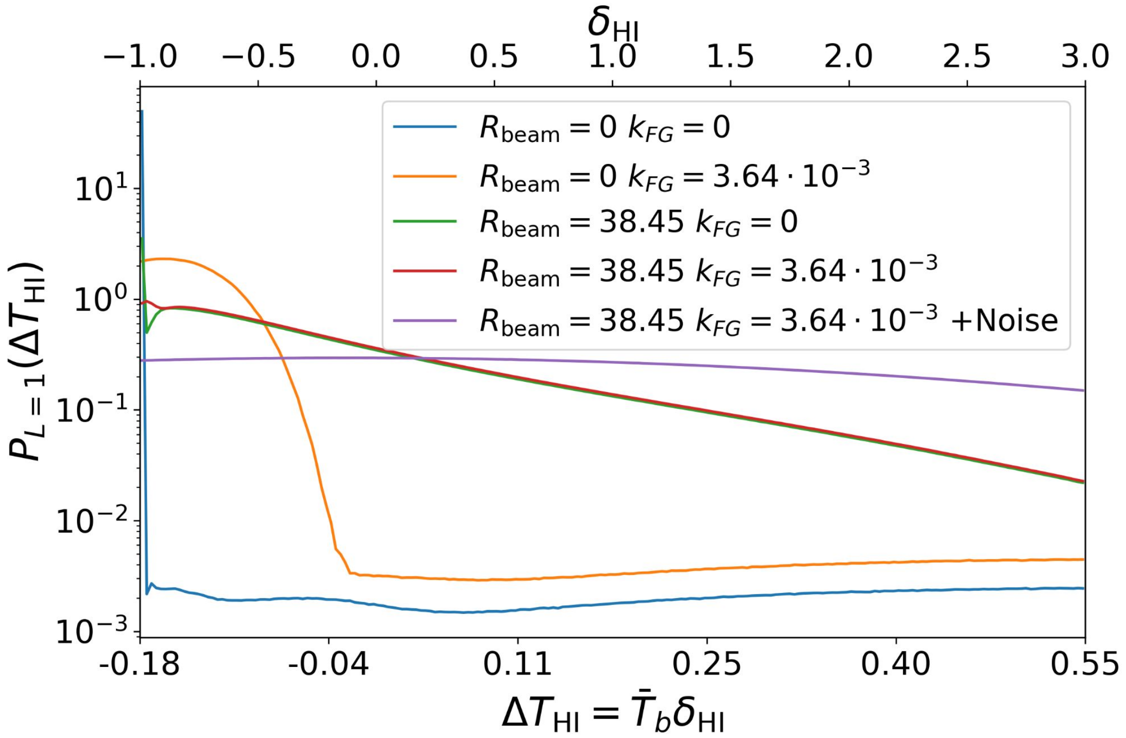
- Bright point sources
- Planets (e.g., Jupiter, Saturn)

This gives a **beam model**.

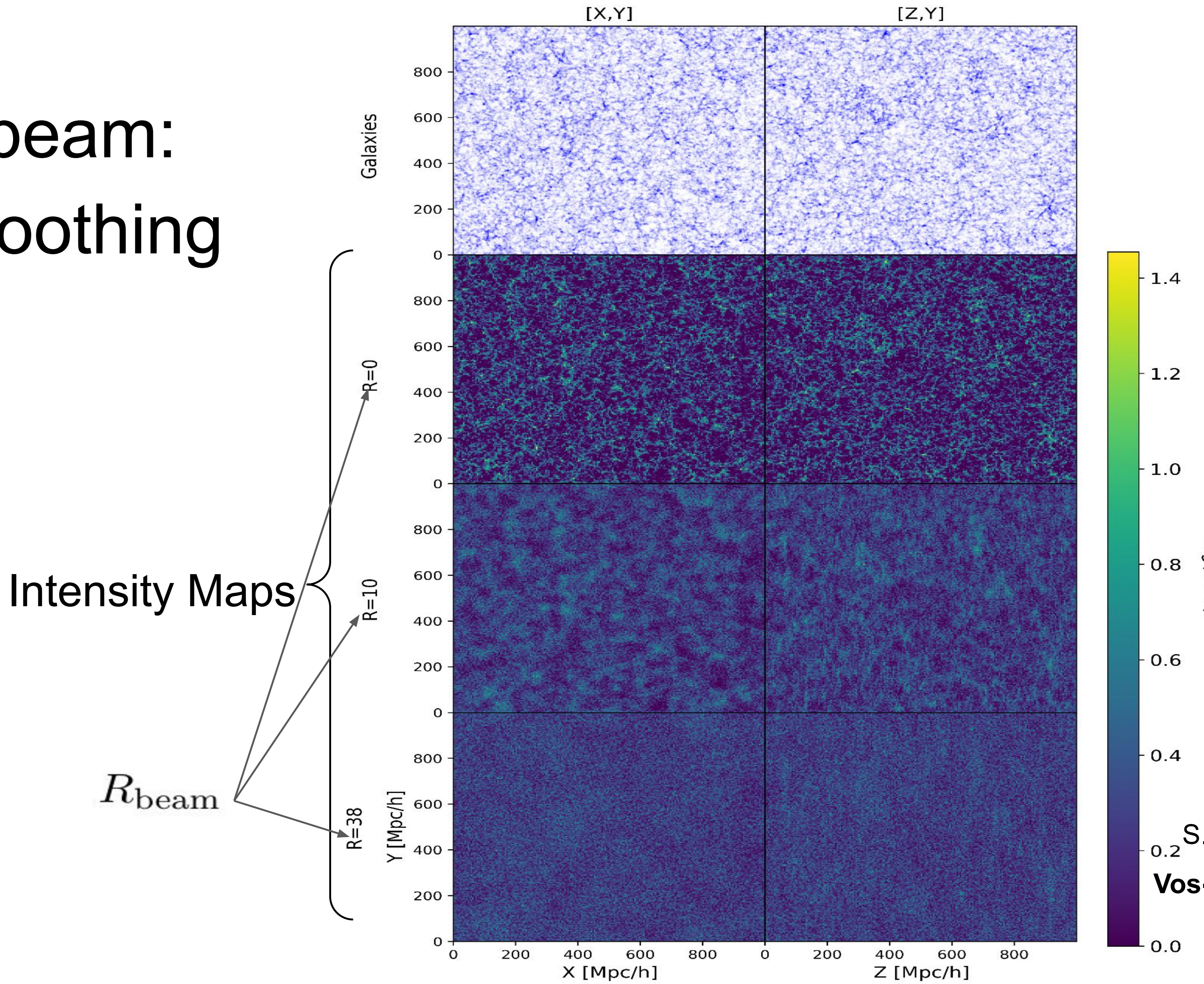
PDF of **unsmoothed** HI Intensity Maps

L_{cell} = 1 Mpc/h

Thermal noise dominates



Telescope beam: angular smoothing



S. Avila & **B. Vos-Ginés**, 2021

$\theta_{FWHM}=\lambda_{HI}\,(1+z)/D_{dish}$,z=1.321

Adding noise in the Intensity Maps

34.75 K

$$\sigma_{\rm pix} = \frac{T_{\rm sys}}{\sqrt{\Delta f\, t_{\rm total}(\Omega_{\rm pix}/S_{\rm area})N_{\rm dishes}N_{\rm beams}}},$$

Δf = 25.0 MHz: which corresponds to a radial dist. of 1 Mpc/h.

(Not proper SKA channel width,would be Δf =13.4 KHz)

10000 hours

Ω_{pix}=1.28·10^{-7} steradians which corresponds to (1 Mpc/h)^2.

(Not proper SKA pixel area, which would be calculated as $\Omega_{pix}=1.13\theta_{FWHM}^{2}$)

196

1

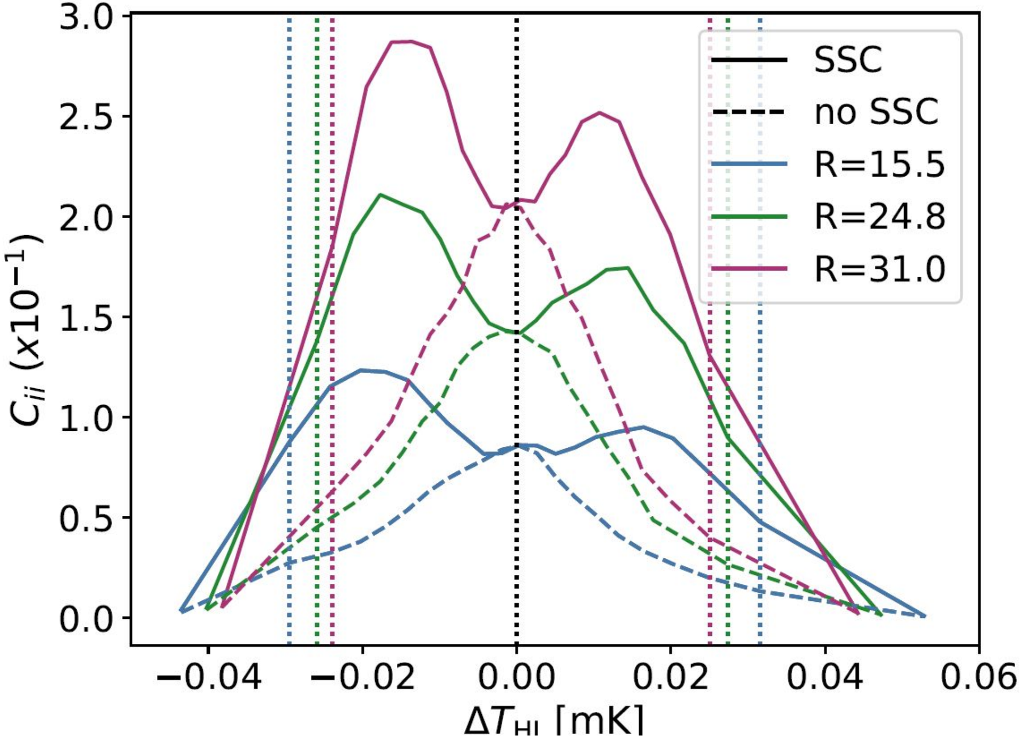
σ_{pix} = 0.449 mK for an IM of 1 Mpc/h side pixels!

fsky = 0.3 -> steradians

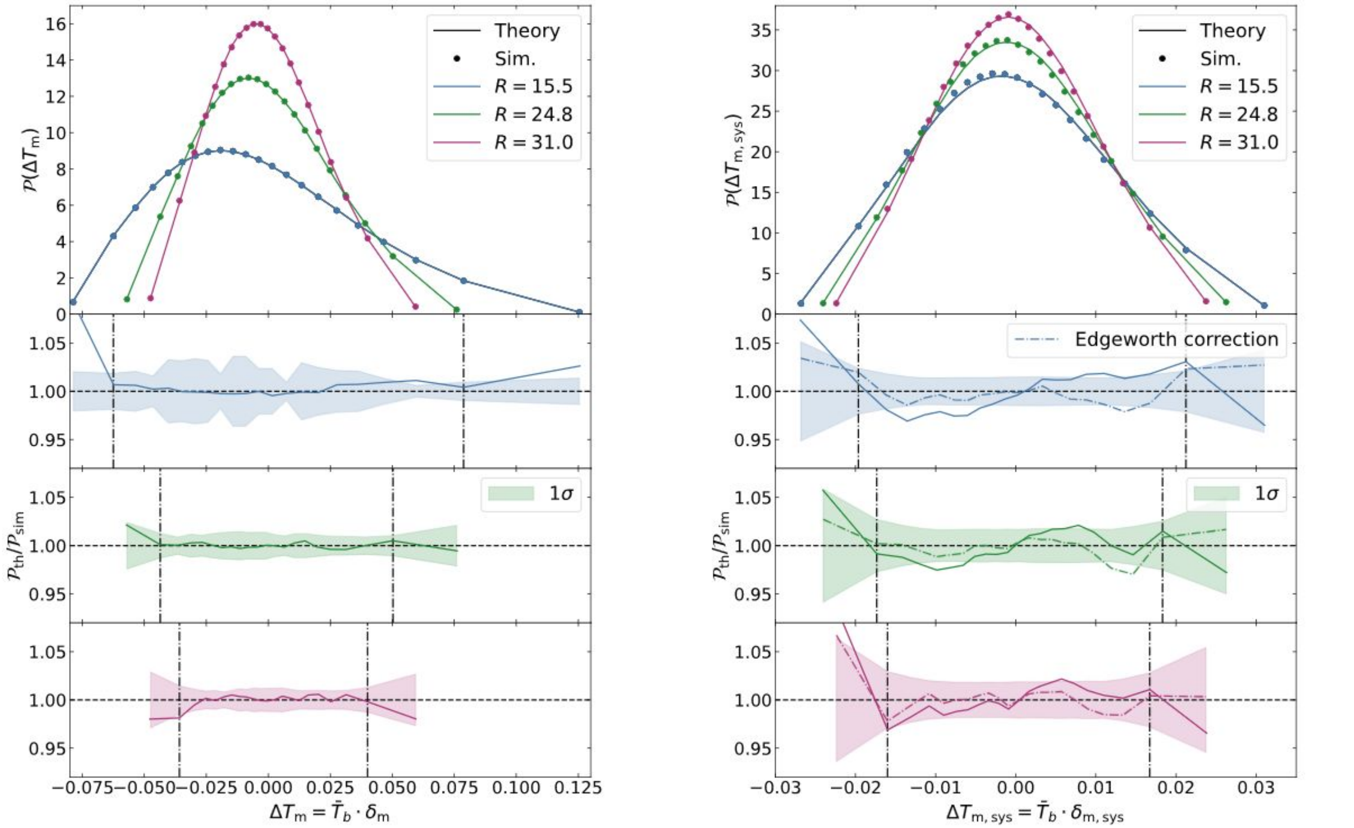
Jackknife Covariance for the PDF

- Using 4 UNITsims
- Each sim divided in 125 subboxes
- Hartlap factor $h_P = 0.91$

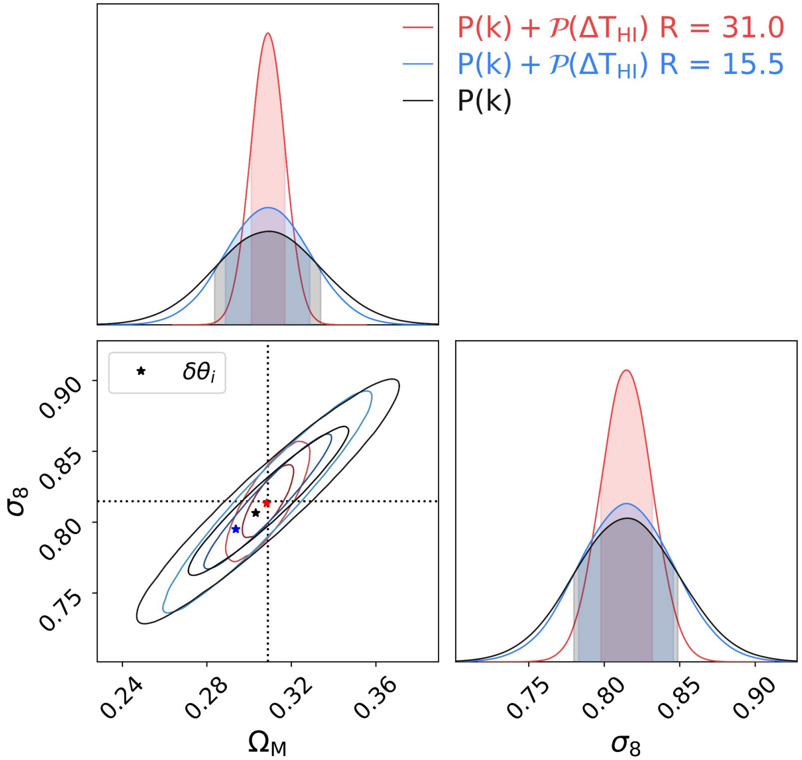
Diagonal amplitudes



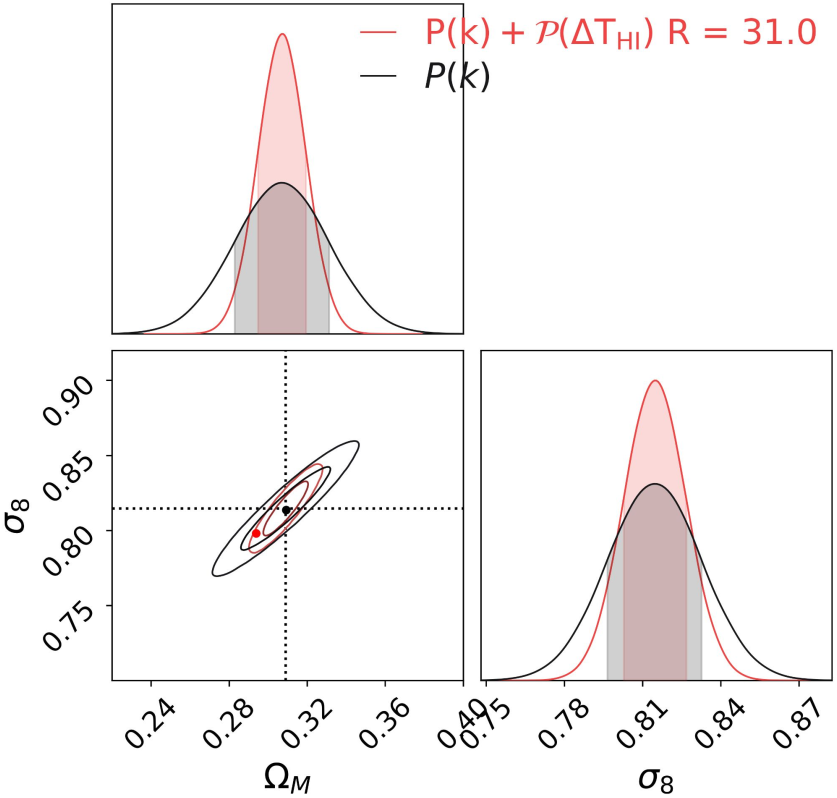
Dark matter PDF (model vs simulation)



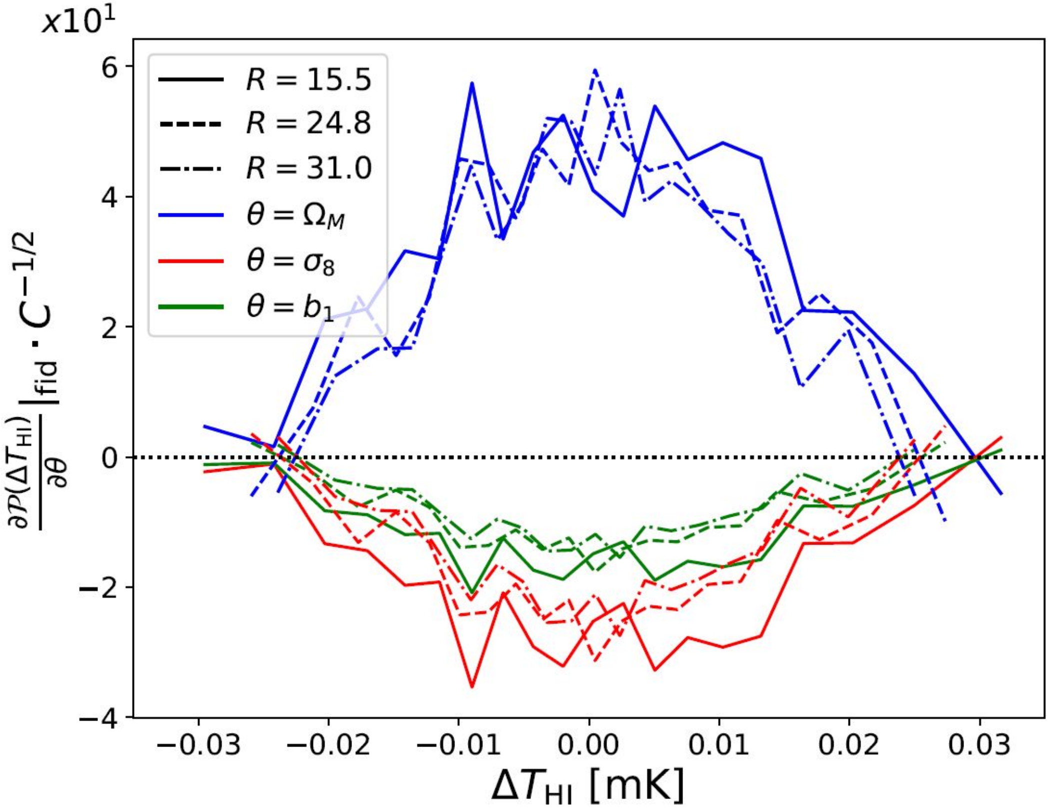
HI systematics



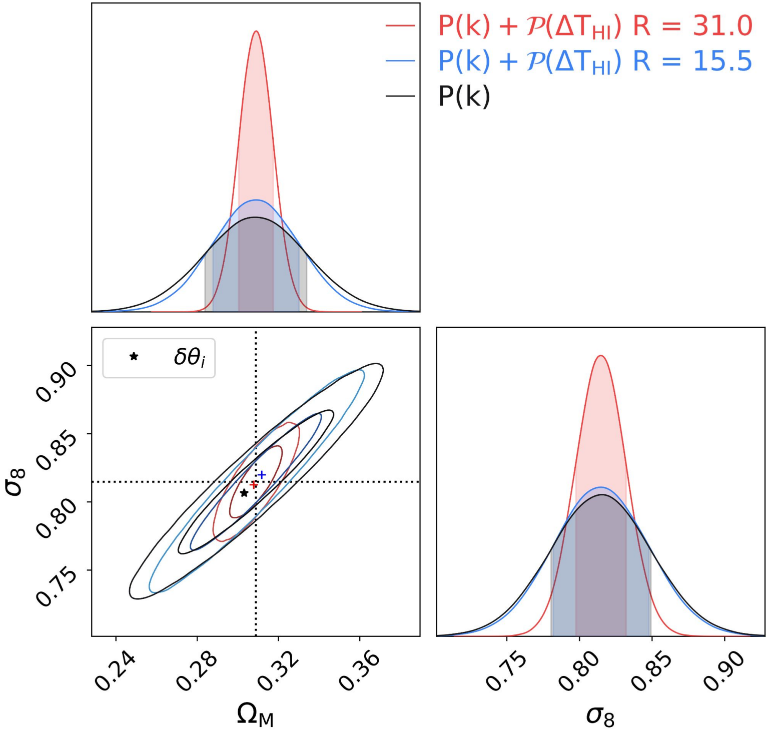
HI NO systematics

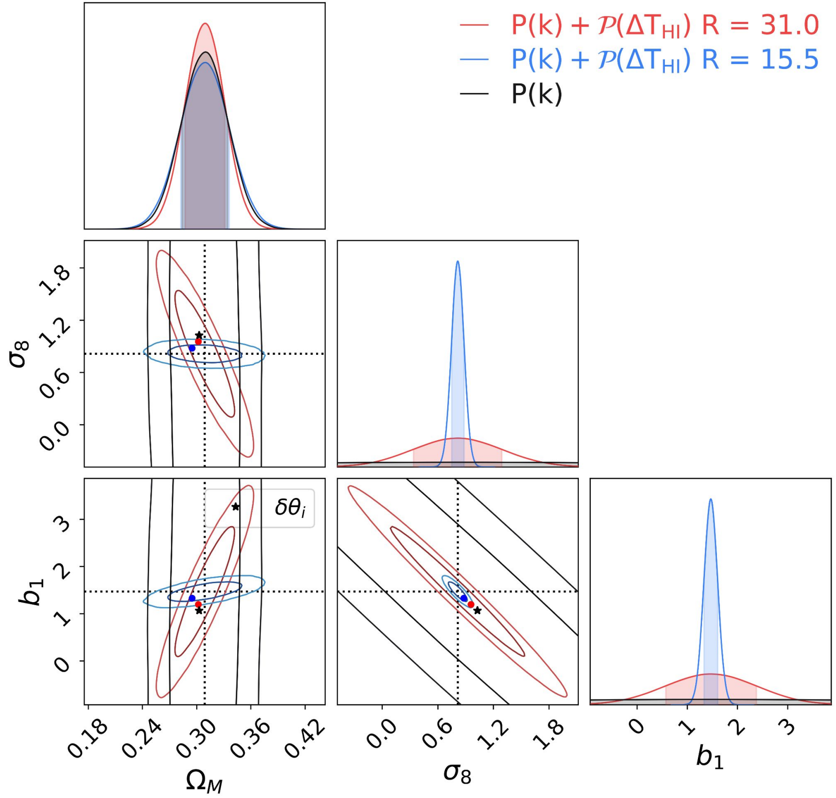


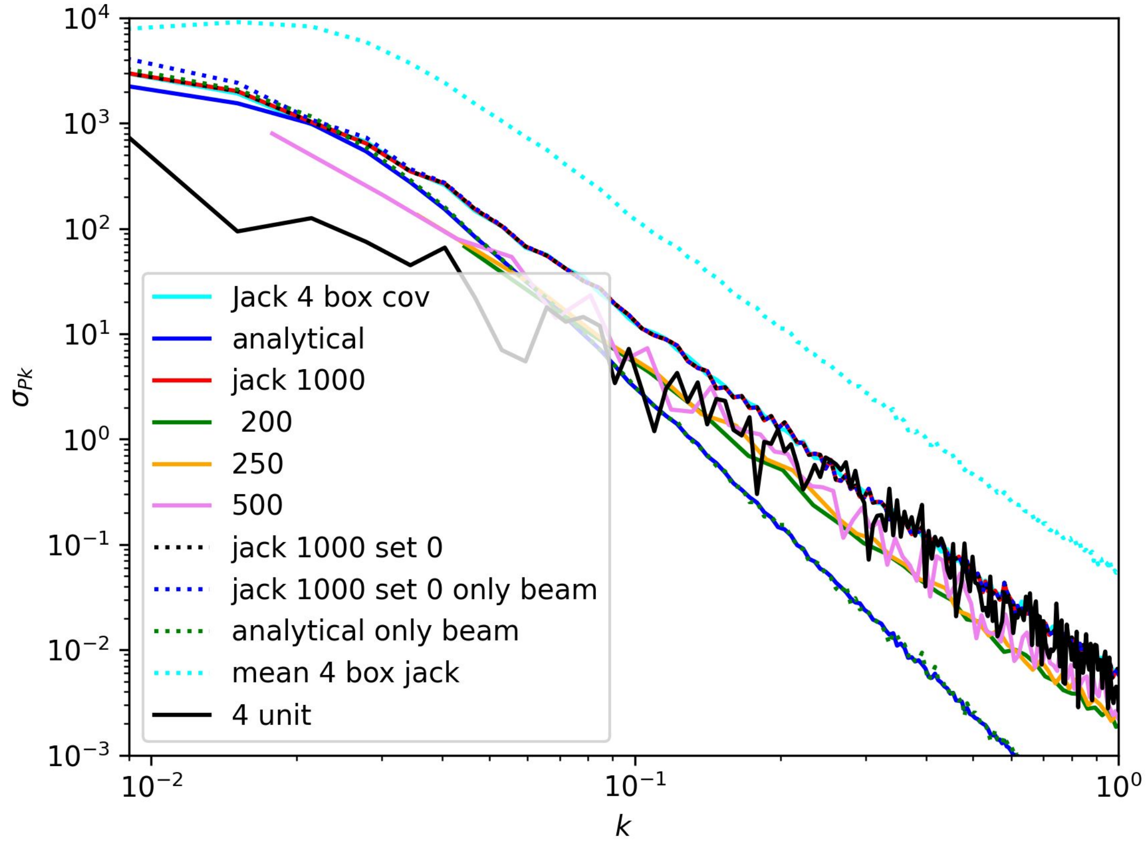
PDF derivatives

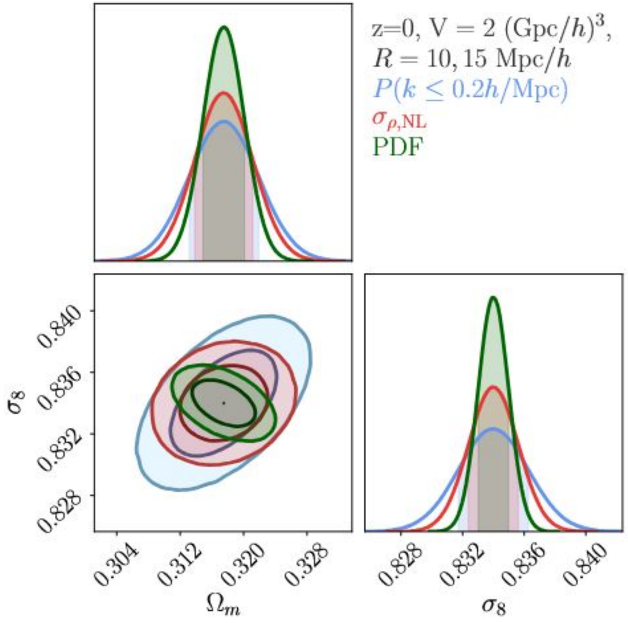


Not taking into account super-sample covariance on 1-point statistics

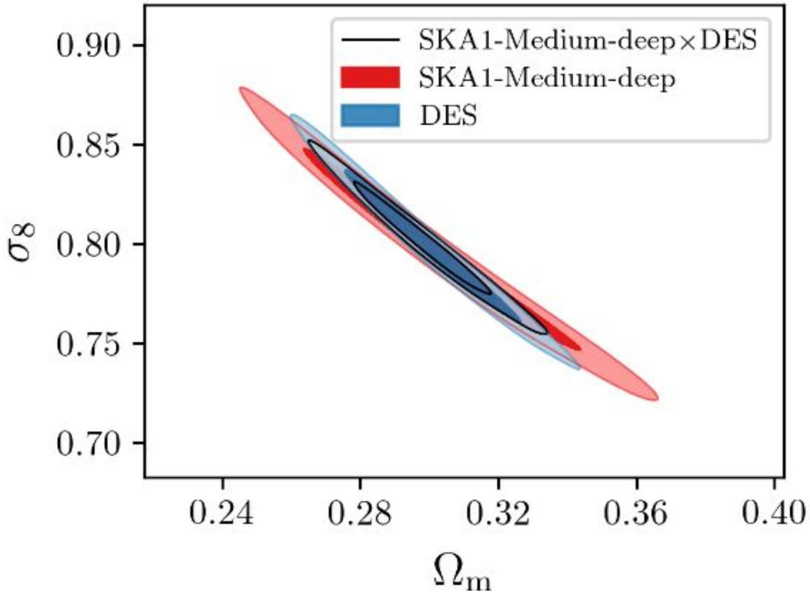




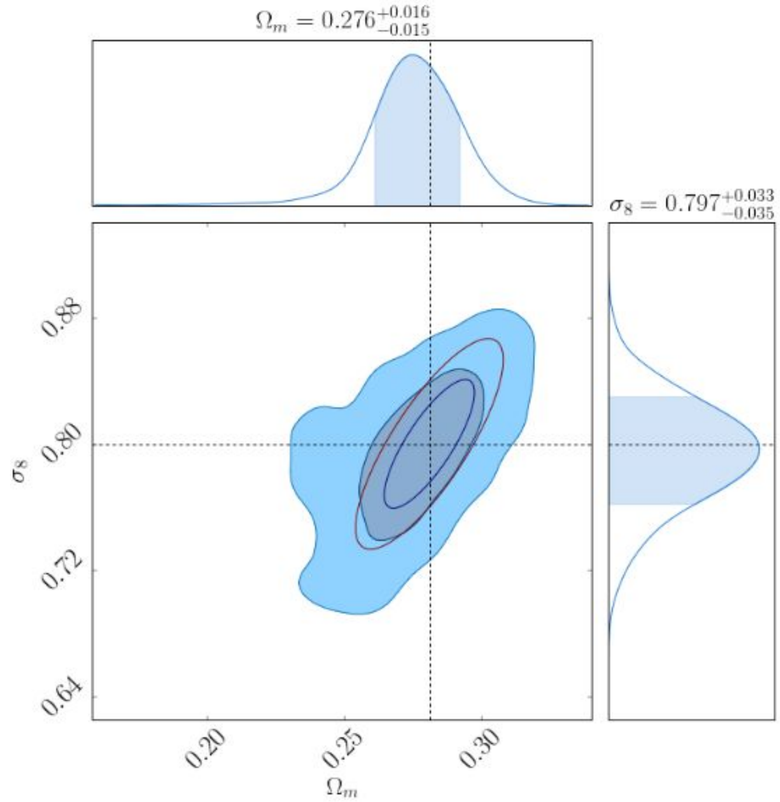

10^4
10^3
10^2
10^1
10^0
10^{-1}
10^{-2}
10^{-3}
σ_{Pk}
10^{-2}
10^{-1}
10^0
k
Jack 4 box cov
analytical
jack 1000
200
250
500
jack 1000 set 0
jack 1000 set 0 only beam
analytical only beam
mean 4 box jack
4 unit



Ω_M - σ_8 Fisher Forecast for P(k) and **matter** PDF at **z=0 combining R = 10, 15 Mpc/h**. [Uhlemann et al., 2019](#)



Forecast constraints for weak lensing (two-point statistics): [SKA Cosmology Sciene Working Group, 2018](#)



$$\sigma^2_{\mathrm{HI(m)}}(R) = \frac{1}{2\pi^2}\int_{\mathrm{k_F}}^{\mathrm{k_{max}}} dk\; k^2 P_{\mathrm{theory(m)}}(k) W^2_{\mathrm{TH}}(kR)\,,$$

$$\sigma^2_{\mathrm{HI(m),sys}}(R) = \frac{1}{4\pi^2}\int_{\mathrm{k_F}}^{\mathrm{k_{max}}} dk\; k^2 P_{\mathrm{theory(m)}}(k) W^2_{\mathrm{TH}}(kR)\int_{-1}^{1} d\mu f^2(k\mu) B^2(k\sqrt{1-\mu^2})$$

SKA-MID1 using angular power spectrum C_l, l_{max} = 1000. z = 0.082. Solid lines: Fisher, contours: MCMC. Padmanabhan et al, 2019