Higher Order Statistics for Neutral Hydrogen Intensity Mapping

Pauline Gorbatchev

Joint ARGOS-TITAN-TOSCA workshop 06/06/2024

Summary

Physical background and Motivations

Intensity Mapping and challenges

SKA and State of the Art

Tools for simulating 21cm signal maps

Bayesian Likelihood for Cosmological Parameter Inference

Utility of Higher Order Statistics

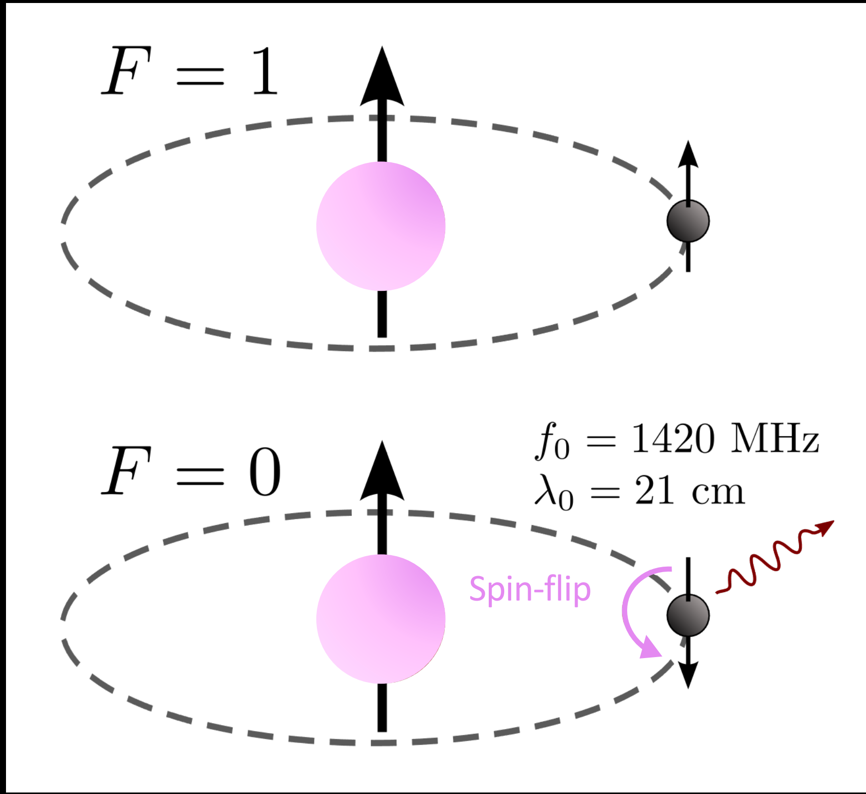
Prospective

TITAN ARTIFICIAL INTELLIGENCE IN ASTROPHYSICS
FORTH INSTITUTE OF COMPUTER SCIENCE
Signal Processing Laboratory
CosmoStat
cea
Funded by the European Union

Motivations

- Hydrogen: most abundant element in the Universe.
- After reionization, HI is located inside galaxies.

$\Longrightarrow$ **biased tracer** of the underlying **matter distribution** of the Universe.



HI hyperfine transition

Why using 21cm line ?

Benefits:

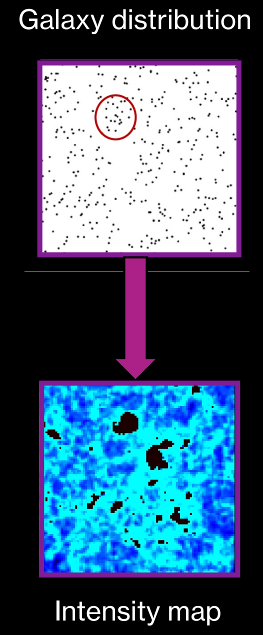
- The only astrophysical spectral feature in the L-band (GHz).
- Can be **measured from earth** (penetrates the atmosphere).
- thermal noise in HI surveys is less important than shot noise in galaxy surveys $\Longrightarrow$ HI analysis is **more constraining than galaxies.**

Uses:

- Probe the **Dark Ages** (future).
- Reconstruct **DM density fields.**
- **Map 3D Large Scale Structures** of the Universe.
- Complementary measurement to optical surveys to **constrain cosmological parameters.**

What is Intensity Mapping?

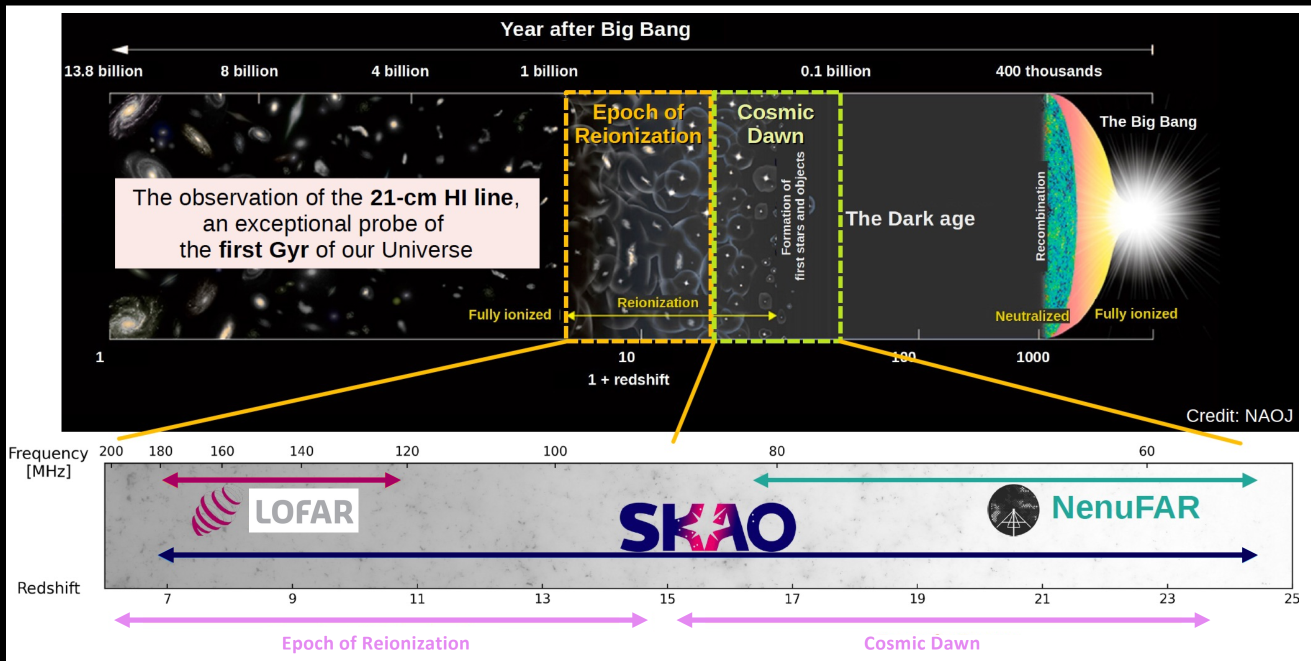
- Measurement of redshift and intensity of HI over the **whole sky.**
- Treats HI signal as a **diffuse background.**
- Large cosmological volume.
- Less costly, less time consuming.
- High spectral resolution ⟹ **high redshift resolution.**
- Individual galaxy detection not needed for LSS study.



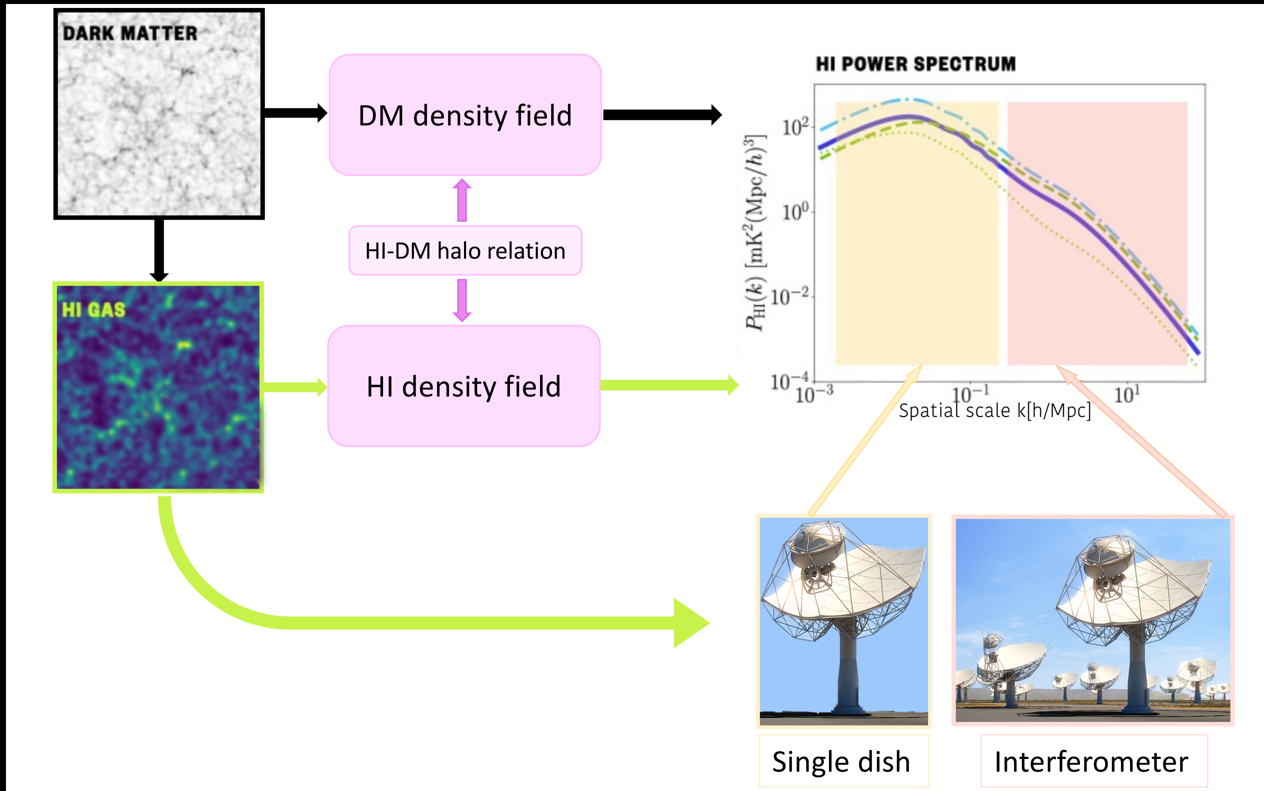
TITAN ARTIFICIAL INTELLIGENCE IN ASTROPHYSICS · FORTH INSTITUTE OF COMPUTER SCIENCE · Signal Processing Laboratory · CosmoStat · cea · Funded by the European Union

State of the Art

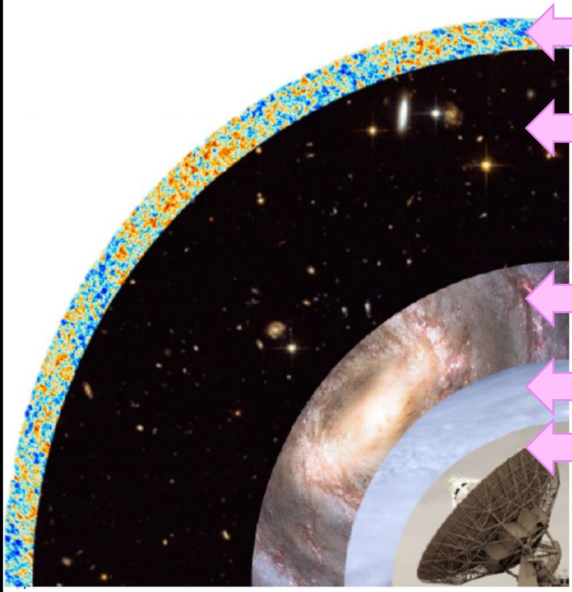
- For now detection by **cross-correlation between galaxy and 21 cm.**
- **Not yet possible** to obtain a measurement of the **21cm auto-Power Spectrum.**



Inferring the Power Spectrum



Measurements scheme



- 21 cm signal
- Extra galactic emission (AGNs)
- Galactic emission (Milky Way synchrotron)
- Ionosphere
- Instrumental effects

Signal

↓

Analysis pipeline

- Foreground cleaning
- Noise cleaning

↓

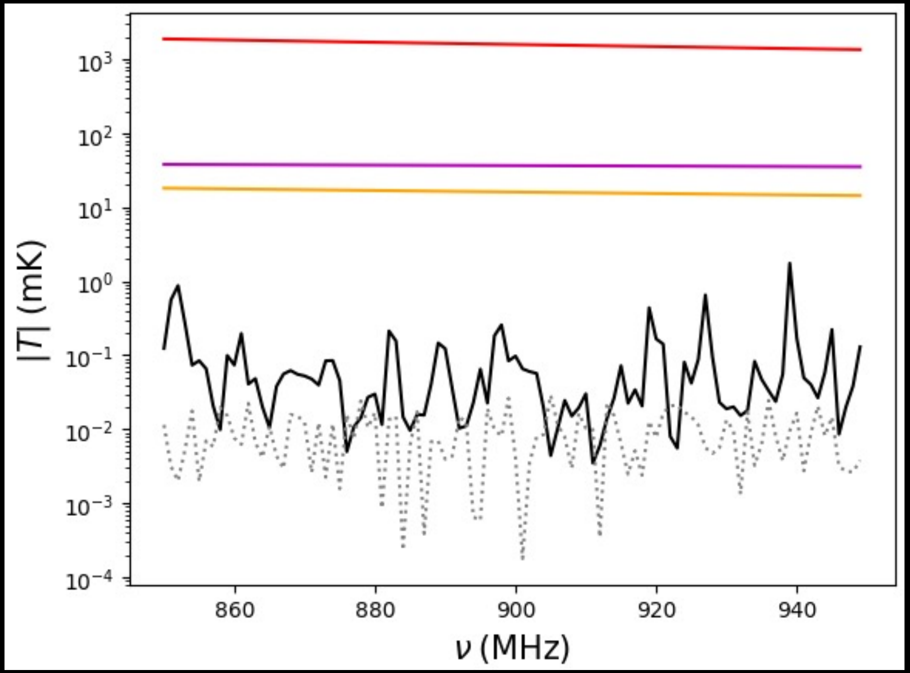
Power spectrum

↓

Cosmological parameters constraints

TITAN ARTIFICIAL INTELLIGENCE IN ASTROPHYSICS
FORTH INSTITUTE OF COMPUTER SCIENCE
Signal Processing Laboratory
CosmoStat
cea
Funded by the European Union

Challenges of Intensity Mapping



https://github.com/spinemart/Tonale2021

Foregrounds:

- Milky way synchrotron emission (high energy electrons accelerated by magnetic fields).
- Extra-galactic point sources (Active Galactic Nuclei).
- Galactic/extra galactic free-free emissions (electrons scattered by ions).

4 orders of magnitude higher than the signal !

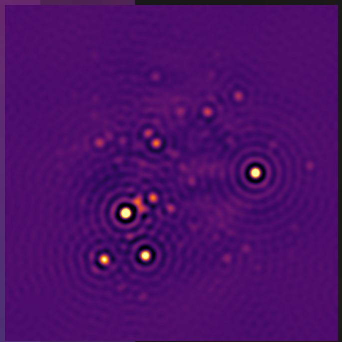
⟹ Foreground removal needed.

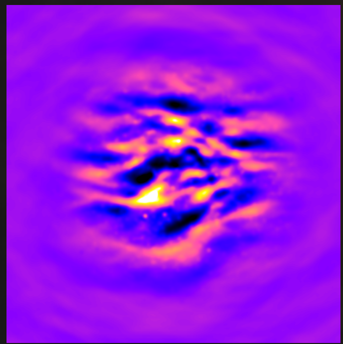
Simulation tools

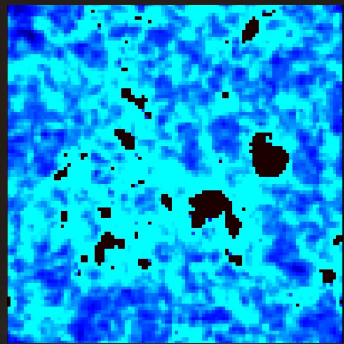
- **CAMB** (Code for Anisotropies in the Microwave Background):
  - Provides matter power spectrum and transfer functions essential for large-scale structure studies.
  - Takes cosmological parameters (e.g., Hubble constant, matter density, dark energy parameters) as input.
- **21cmFAST**:
  - A semi-numerical simulation code used to generate large-scale 21 cm signal maps.
  - Models the cosmic 21 cm signal from the epoch of reionization (EoR) and cosmic dawn.

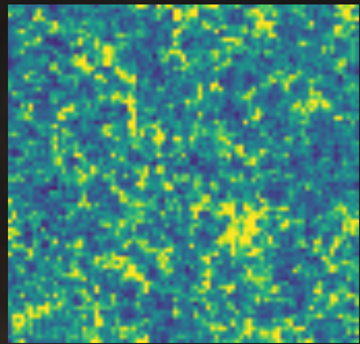
TITAN Artificial Intelligence in Astrophysics · FORTH Institute of Computer Science · Signal Processing Laboratory · CosmoStat · cea · Funded by the European Union

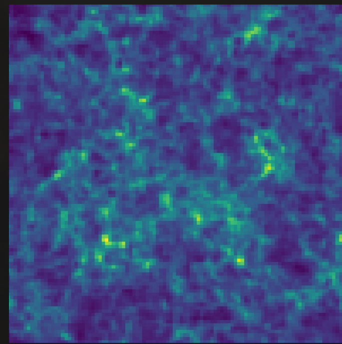
From observations to cosmological information











1 Signal with bright sources

2 Signal without bright sources

3 Brightness temperature map

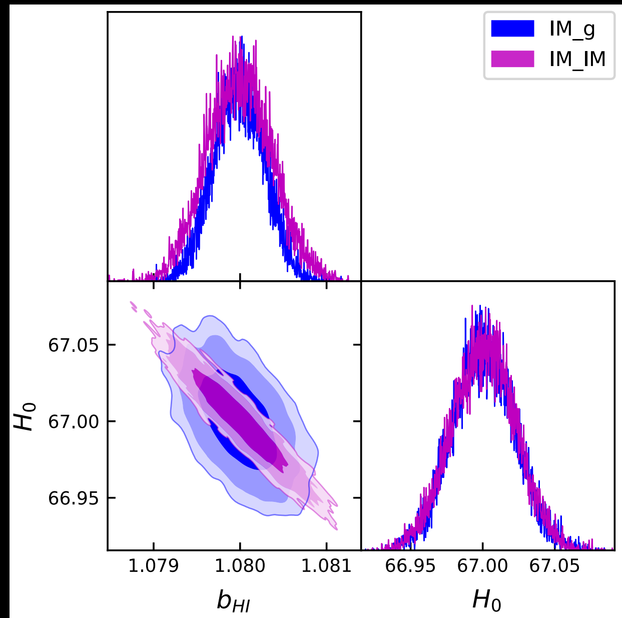
4 HI density field

5 Underlying dark matter distribution

Funded by the European Union

Cosmological parameter inference

- Use CAMB to generate theoretical power spectra based on a given set of cosmological parameters.
- Compare with observed data (Simulated from CAMB) to infer the posterior distribution of the model parameters using MCMC sampling for Bayesian likelihood inference.



$$P_{\mathrm{IM}}(z,k) = \left[189\, h \, \frac{(1+z)^2 H_0}{H(z)} \, \Omega_{\mathrm{HI}}(z)\, \mathrm{mK} \right]^2 \frac{[d_{\mathrm{A,ref}}(z)]^2 \; H(z)}{d_{\mathrm{A}}^2(z)\, H_{\mathrm{ref}}(z)} \; \exp\left[- \frac{k^2\,(1-\mu_\theta^2)\, r^2(z)\, \theta_{\mathrm{pb}}^2(z)}{16 \ln 2} \right]^2$$

$$\times \frac{1}{1 + k^2 \mu_\theta^2 \frac{1}{6\pi^2} \int \mathrm{d}k\, P_{\delta\delta}(k,z) f^2(k,z)} \left[b_{\mathrm{HI}}(z)\, \sigma_8(z) + f(z,k)\, \sigma_8(z)\, \mu_\theta^2 \right]^2 \times \frac{P_{\mathrm{dw}}(z,k,\mu_\theta)}{\sigma_8^2(z)}$$

$$P_{\mathrm{IM,g}}(z,k,\mu_\theta) = r_{\mathrm{HI,g}} \frac{[d_{\mathrm{A,ref}}(z)]^2 \; H(z)}{d_{\mathrm{A}}^2(z)\, H_{\mathrm{ref}}(z)} \; 189\, h \, \frac{(1+z)^2 H_0}{H(z)} \, \Omega_{\mathrm{HI}}(z)\, \mathrm{mK} \left[b_{\mathrm{g}}(z)\, \sigma_8(z) + f(z,k)\, \sigma_8(z)\, \mu_\theta^2 \right]$$

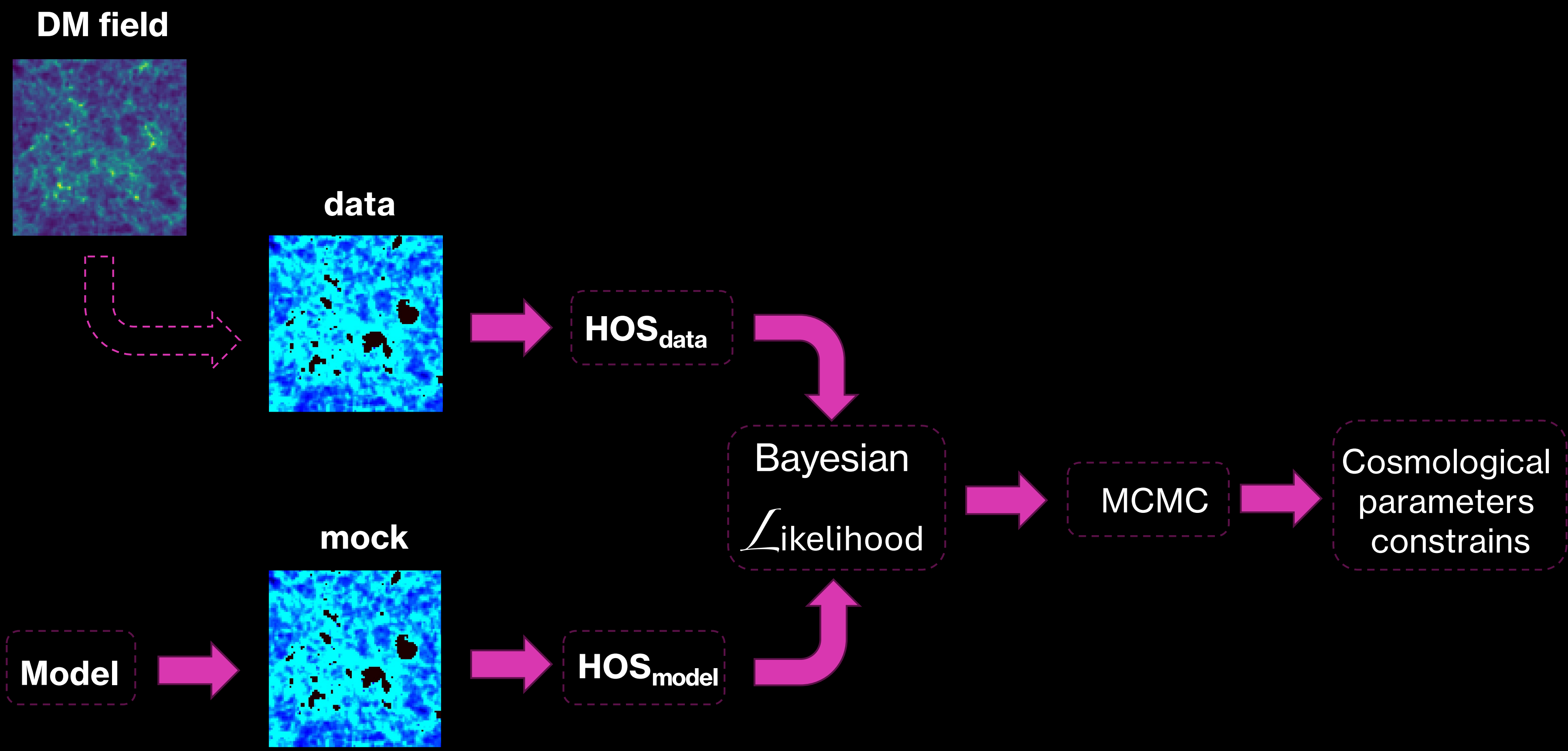
$$\times \left[b_{\mathrm{HI}}(z)\, \sigma_8(z) + f(z,k)\, \sigma_8(z)\, \mu_\theta^2 \right] \frac{1}{1 + k^2 \mu_\theta^2 \frac{1}{6\pi^2} \int \mathrm{d}k\, P_{\delta\delta}(k,z) f^2(k,z)} \times \frac{P_{\mathrm{dw}}(z,k,\mu_\theta)}{\sigma_8^2(z)} \; \exp\left[- \frac{k^2\,(1-\mu_\theta^2)\, r^2(z)\, \theta_{\mathrm{pb}}^2(z)}{16 \ln 2} \right]$$

Limitations of the power spectrum

- **Gaussian Assumption**:
  - The power spectrum is most effective for Gaussian random fields.
- **Non-Gaussian Features**:
  - The universe exhibits non-Gaussian features due to non-linear growth of structures and primordial non-Gaussianities.
- **Loss of Information**:
  - Higher order interactions and complex structures are not captured by the power spectrum.
  - Important information about the morphology and connectivity of cosmic structures is lost.

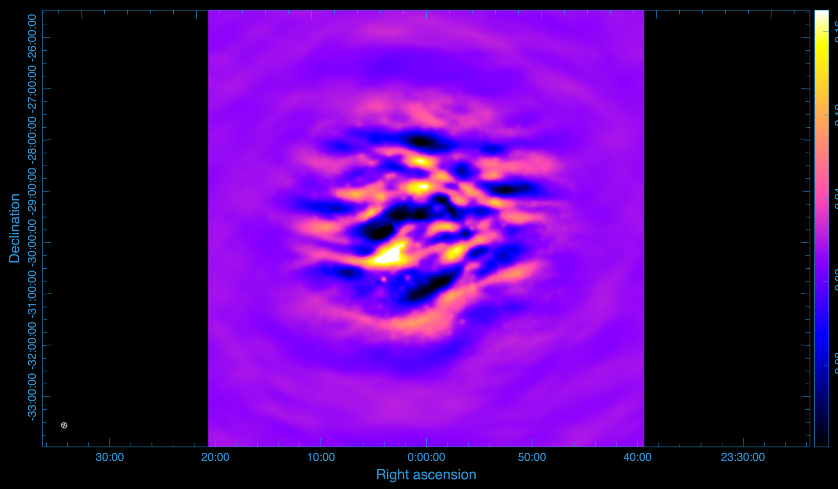
Make use of Higher order statistics which are sensitive to the non-Gausianities.

Bayesian cosmological parameter inference



Participation in SKA Data-Challenge

- Team Eos (Goddess of Dawn).
- Visibilities (points in the Fourier plane (u,v))~ 8 Tb.
- Cube images (ponderal and uniform) computed from the visibilities.
- Goal: compute the power spectrum of the 21 cm signal.
- Available codes : ps_eor (it separates the foreground and the signal using GPR, reconstruct the power spectrum from images).



https://gitlab.com/flomertens

References

- Cosmology with the SKA Observatory, Marta Spinelli , Observatoire de la Cote d'Azur 2024.
- Reza Ansari, "Current status and future of cosmology with 21 cm Intensity Mapping," 2022.
- Soares, P., et al. "Gaussian Process Regression for foreground removal in HI Intensity Mapping experiments," in *Monthly Notices of the Royal Astronomical Society*, vol. 510, no. 4, pp. 5872–5890, 2021.
- M. Irfan, P. Bull. "Cleaning foregrounds from single-dish 21 cm intensity maps with Kernel principal component analysis," in *Monthly Notices of the Royal Astronomical Society*, vol. 508, no. 3, pp. 3551–3568, 2021.
- HI Intensity Mapping with the MIGHTEE Survey, Aishrila Mazumder, University of Machester, 2024.
- The MeerKLASS survey: updates, Mário G. Santos, University of the Western Cape, 2024.
- Foreground Leakage from Calibration Errors inInterferometric MeerKAT 21 cm Observations, Zhaoting Chen, University of Edinburgh, 2024.

All pictures were modified following my needs for the presentation