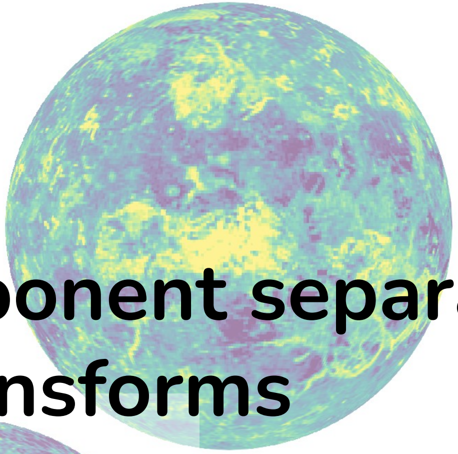
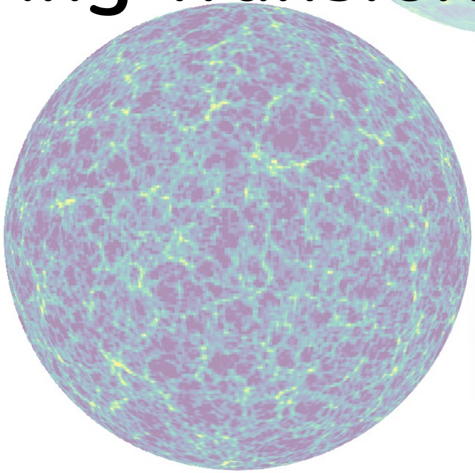
WG - Dark Energy
Université de Montpellier

November 6, 2025



Generative models and component separation with Scattering Transforms

Louise Mousset
E. Allys and all the Scattering Team

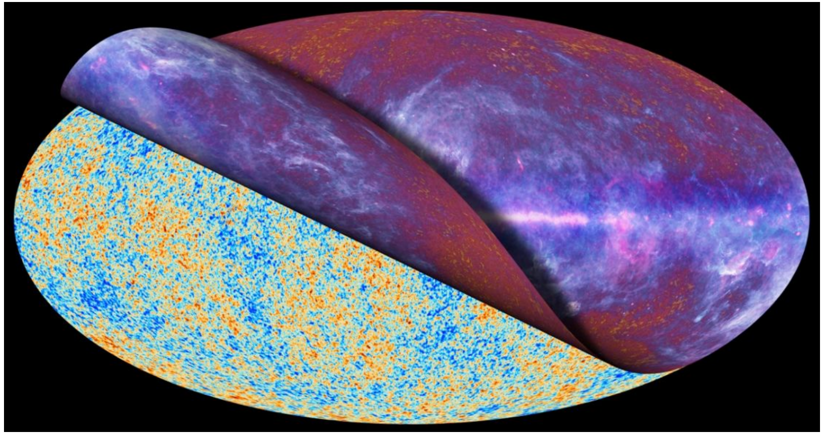


IAS Orsay

LPENS
LABORATOIRE DE PHYSIQUE DE L'ÉCOLE NORMALE SUPÉRIEURE

Motivations

- We observe a **single** multi-frequency sky => a single realisation.
- **Mixture** of components
- Non linear processes => Non gaussian structures
- Often non complete physical or numerical model, and can be computationally expensive



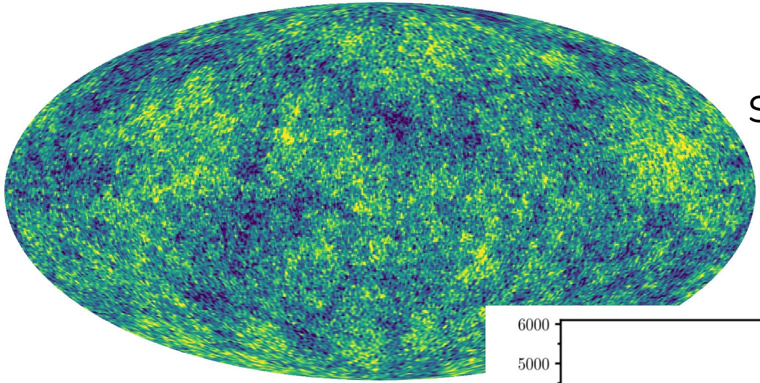
Goal : Develop a statistical model of the components from observation only.

The power spectrum

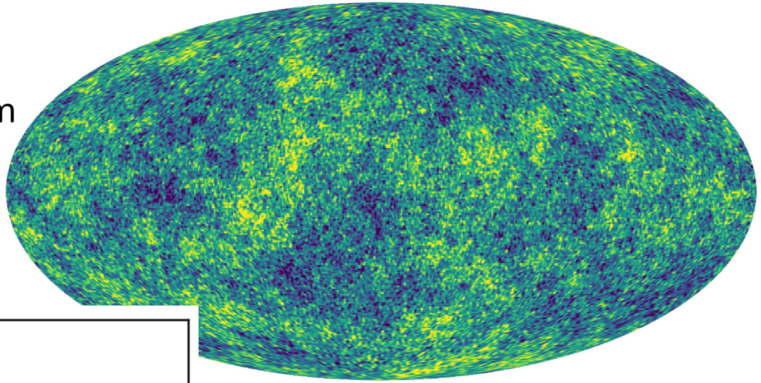
Linear processes → Fourier modes are independent, no coupling between scales → Power spectrum contains all the statistical information

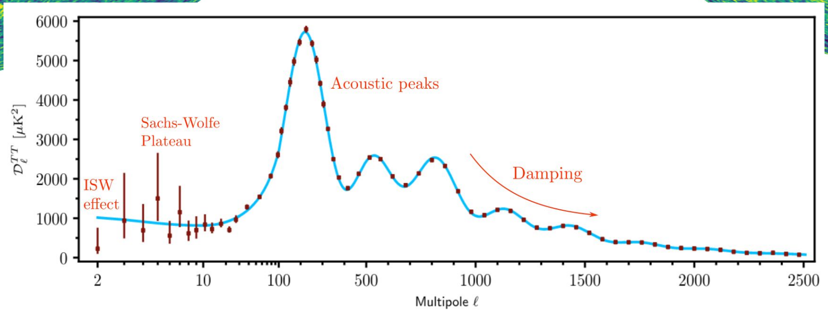
The power spectrum

Linear processes → Fourier modes are independent, no coupling between scales → Power spectrum contains all the statistical information



Same power spectrum





Limitations of the power spectrum

Non linear physical processes: gravity, magneto-hydrodynamics...

→

Coupling between scales, Non Gaussian structures (ex: filaments)

→

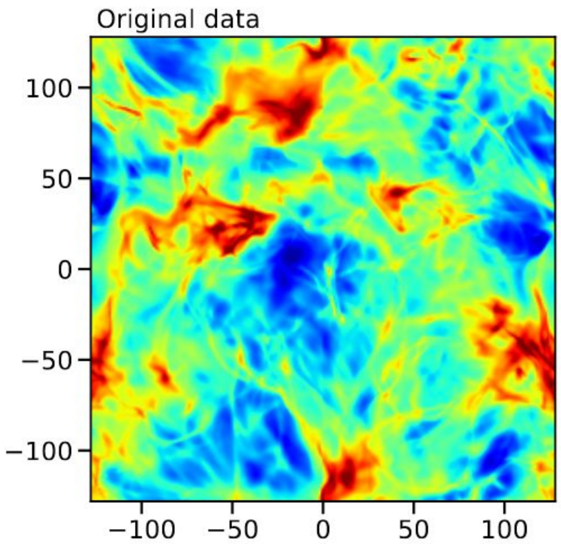
Power spectrum is not sufficient

Limitations of the power spectrum

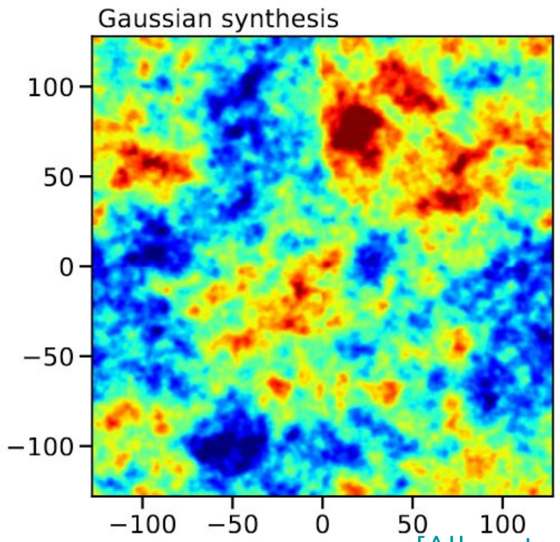
Non linear physical processes: gravity, magneto-hydrodynamics...

Coupling between scales, Non Gaussian structures (ex: filaments)

Power spectrum is not sufficient



Same power spectrum



[Allys et al. 2019]

=> Need higher order non-Gaussian statistics for proper characterisation

Scattering Transforms

Scattering Transforms are [Bruna et al. 2013, Allys et al. 2019]

- a set of non-Gaussian summary statistics
- inspired from neural networks, but can be computed without a training stage, even from a single image.
- **directional wavelet filters** to separate the different scales and angles

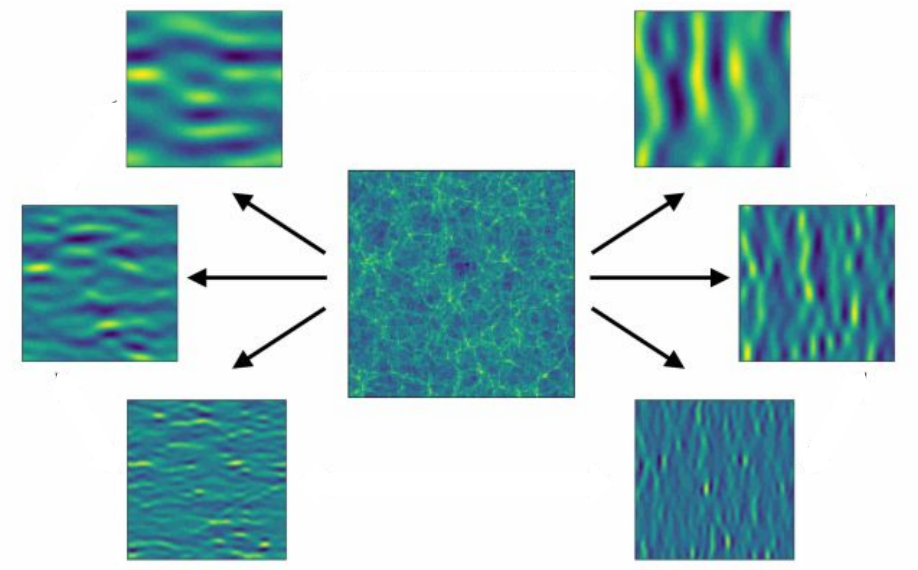


Image from E. Allys

Scattering Transforms

Scattering Transforms are [Bruna et al. 2013, Allys et al. 2019]

- a set of non-Gaussian summary statistics
- inspired from neural networks, but can be computed without a training stage, even from a single image.
- **directional wavelet filters** to separate the different scales and angles
- **non linearities** to extract the couplings between scales

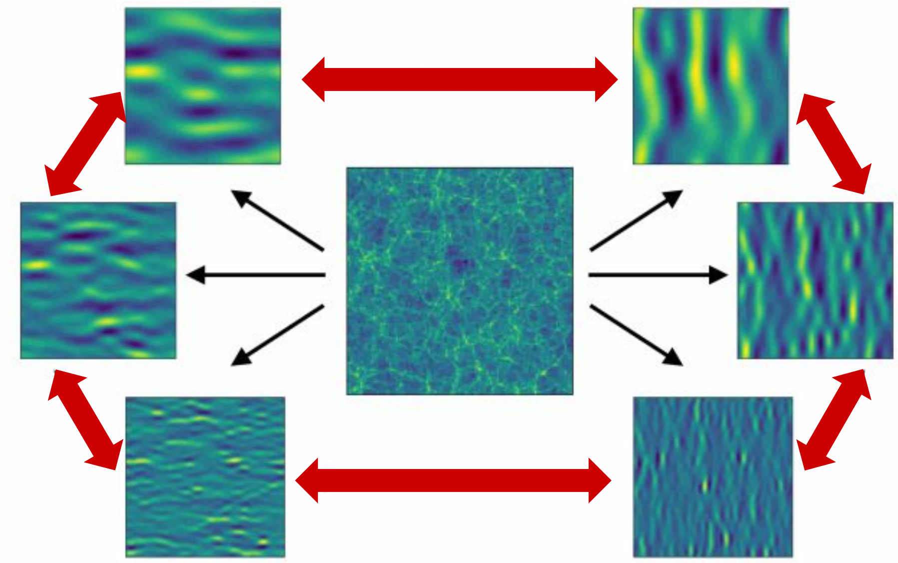
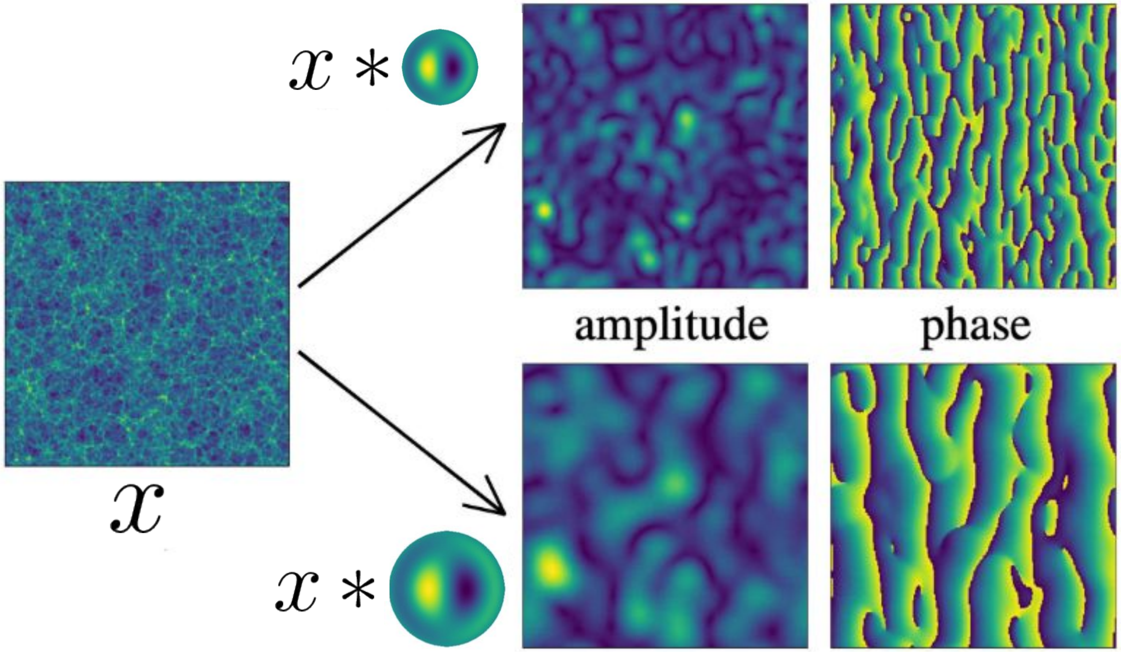
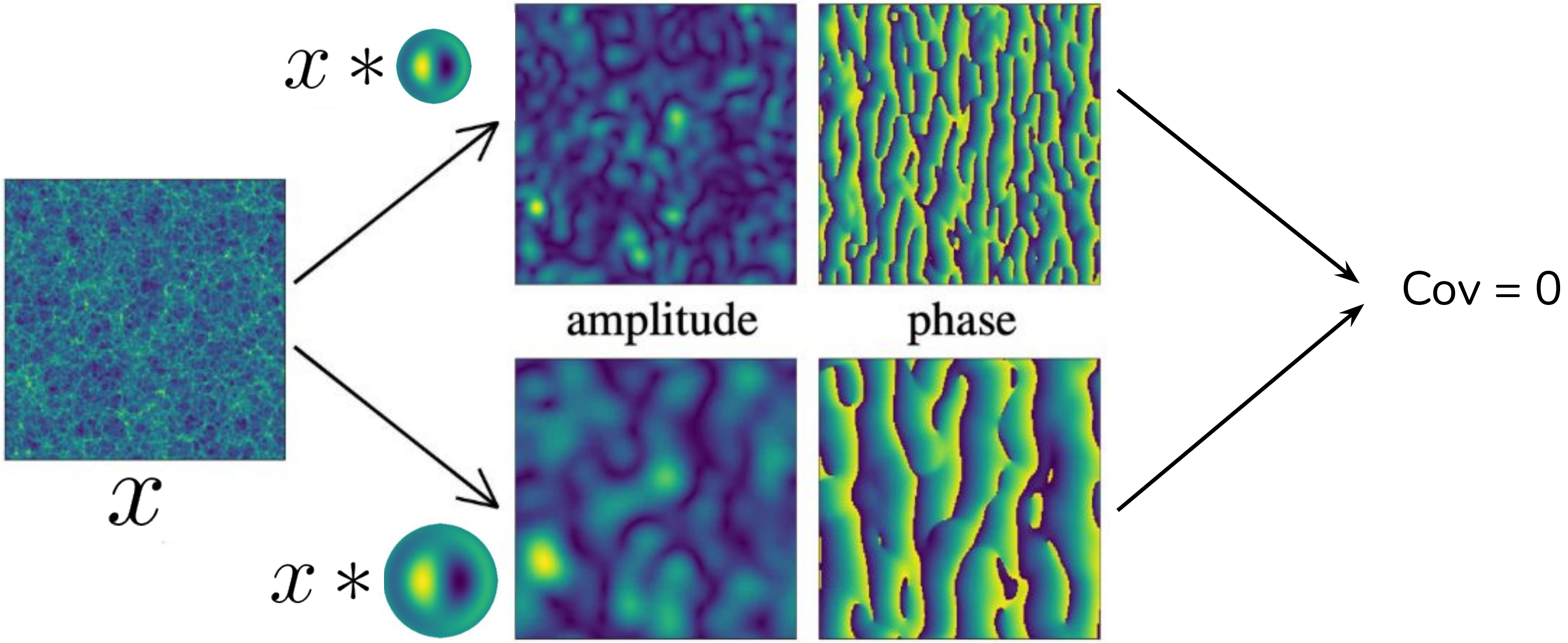


Image from E. Allys

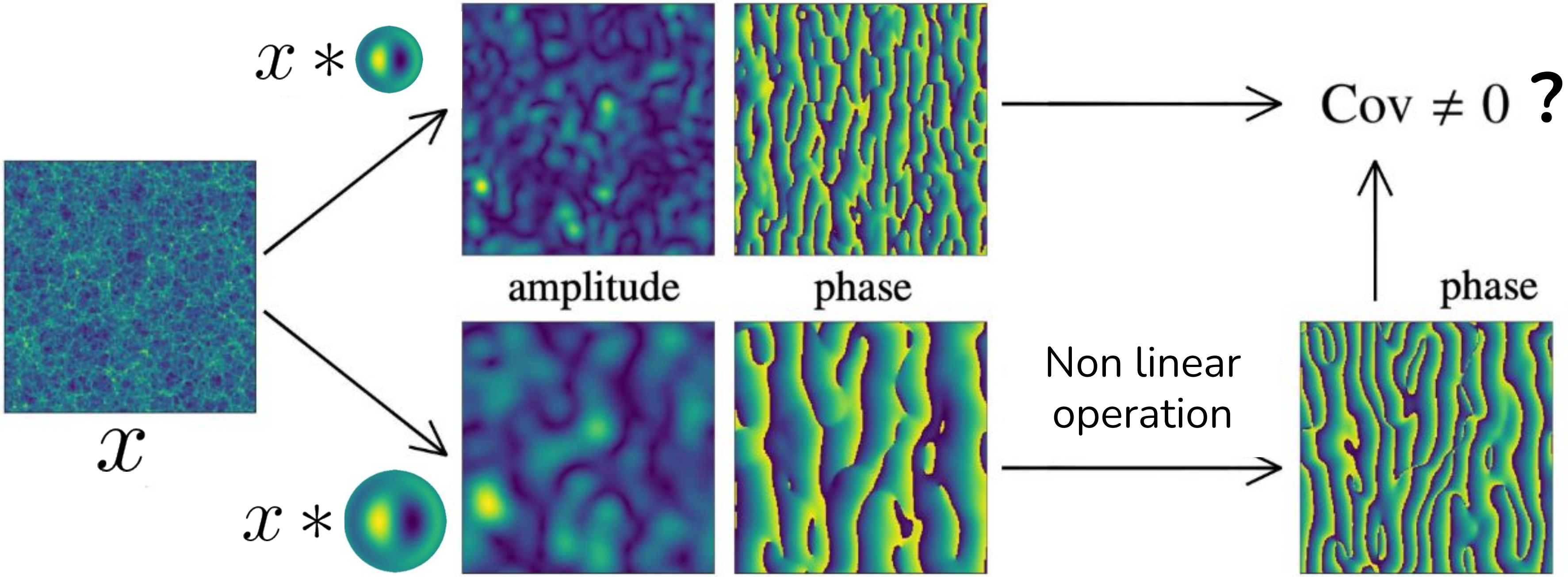
Extract coupling between scales



Extract coupling between scales



Extract coupling between scales



This is for one pair of scales.

=> Following the same for all scales we get a family of coefficients.

Generative models with Scattering Transforms

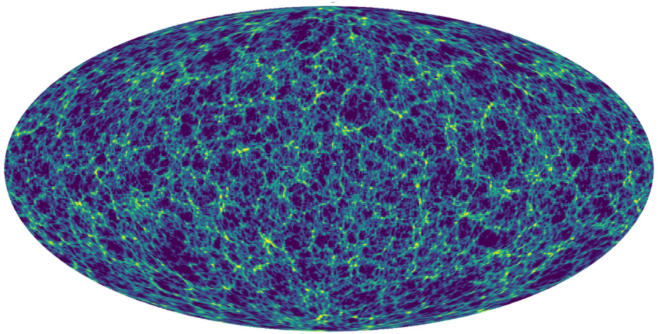
Maximum entropy generative model

[Allys et al. 2020]

Summary statistics:

$$\phi(x) = \{\langle x \rangle, \mathrm{Var(x)}, \mathrm{Scattering\ coefficients}\}$$

Target x_t



$$\phi(x_t)$$

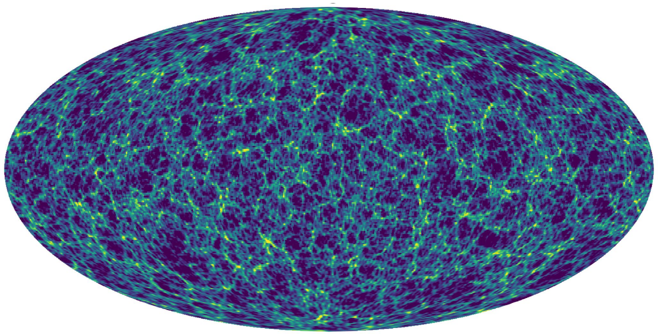
[Allys et al. 2020]

Maximum entropy generative model

Summary statistics:

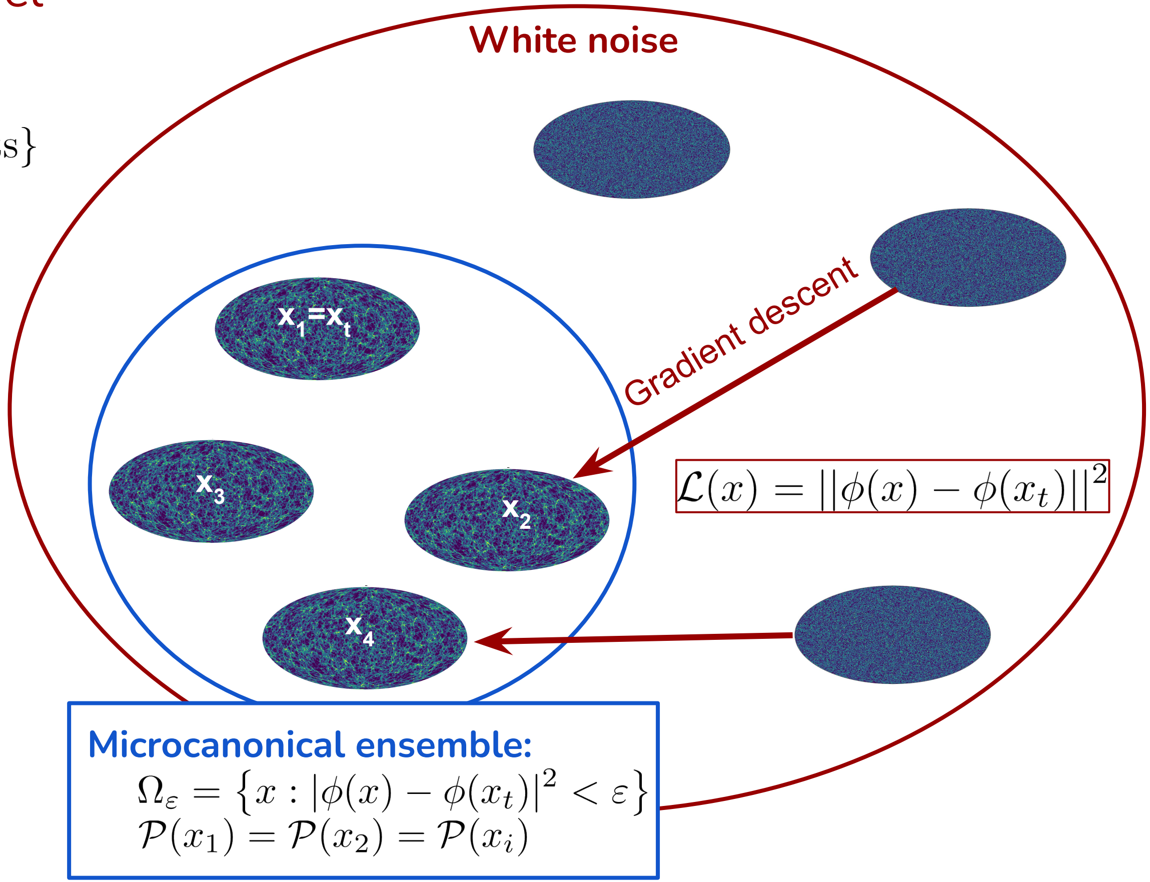
$$\phi(x) = \{\langle x \rangle, \mathrm{Var(x)}, \mathrm{Scattering\ coefficients}\}$$

Target x_t



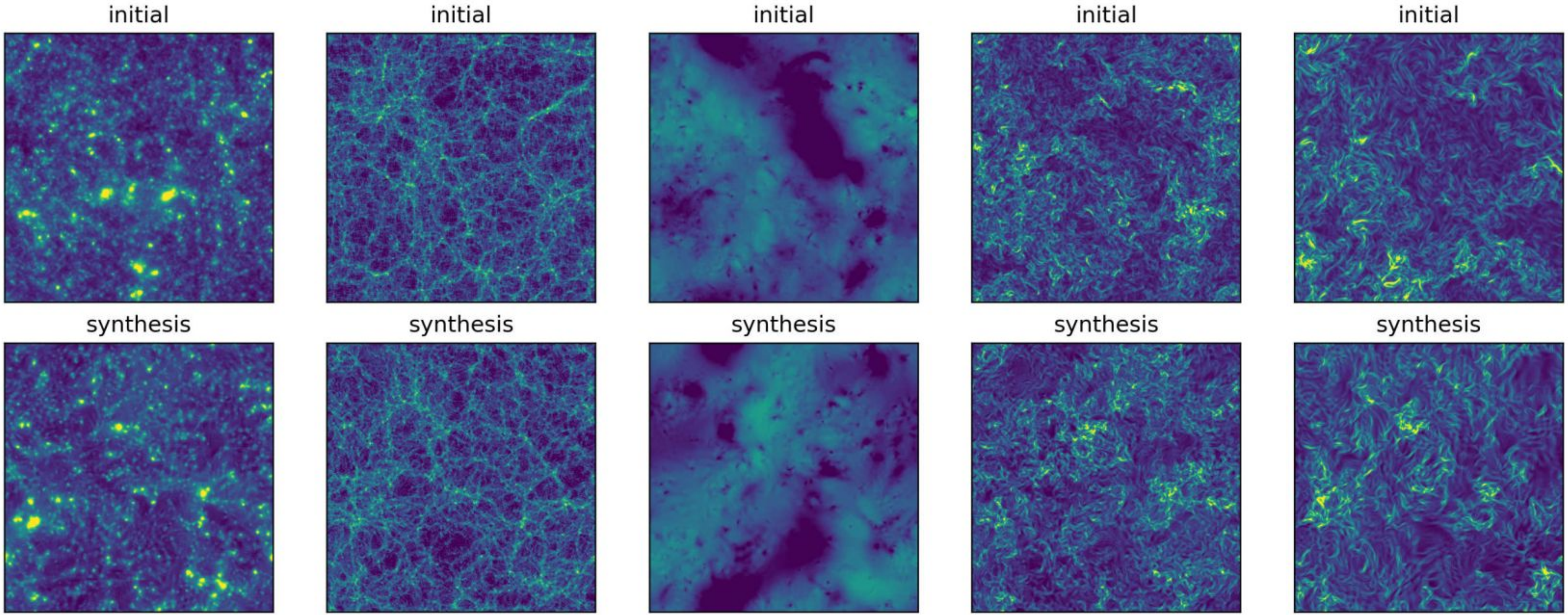
$$\downarrow$$

$$\phi(x_t)$$



2D planar data with various textures

[Cheng et al. 2024]



Quantitatively realistic generative models from a few 100 of coefficients only.

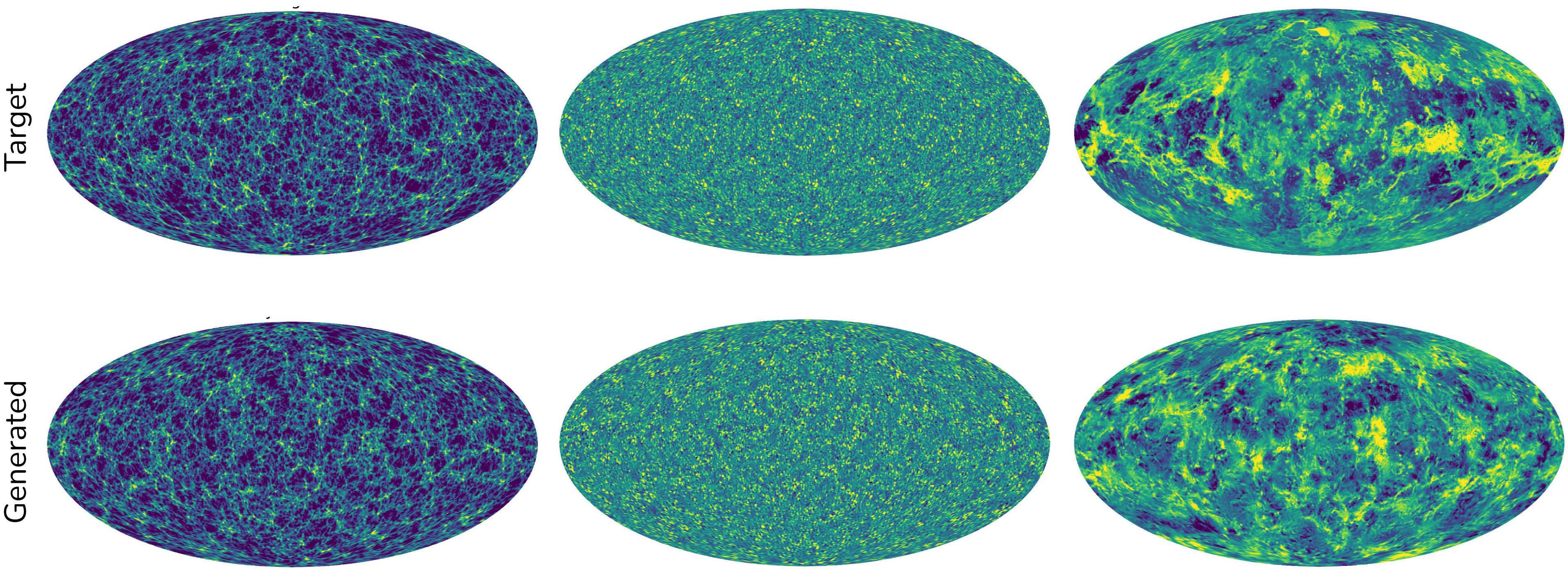
Spherical 2D maps

[Mousset et al. 2024]

Large Scale Structures
N-body simulation [Kacprzak et al. 2022]

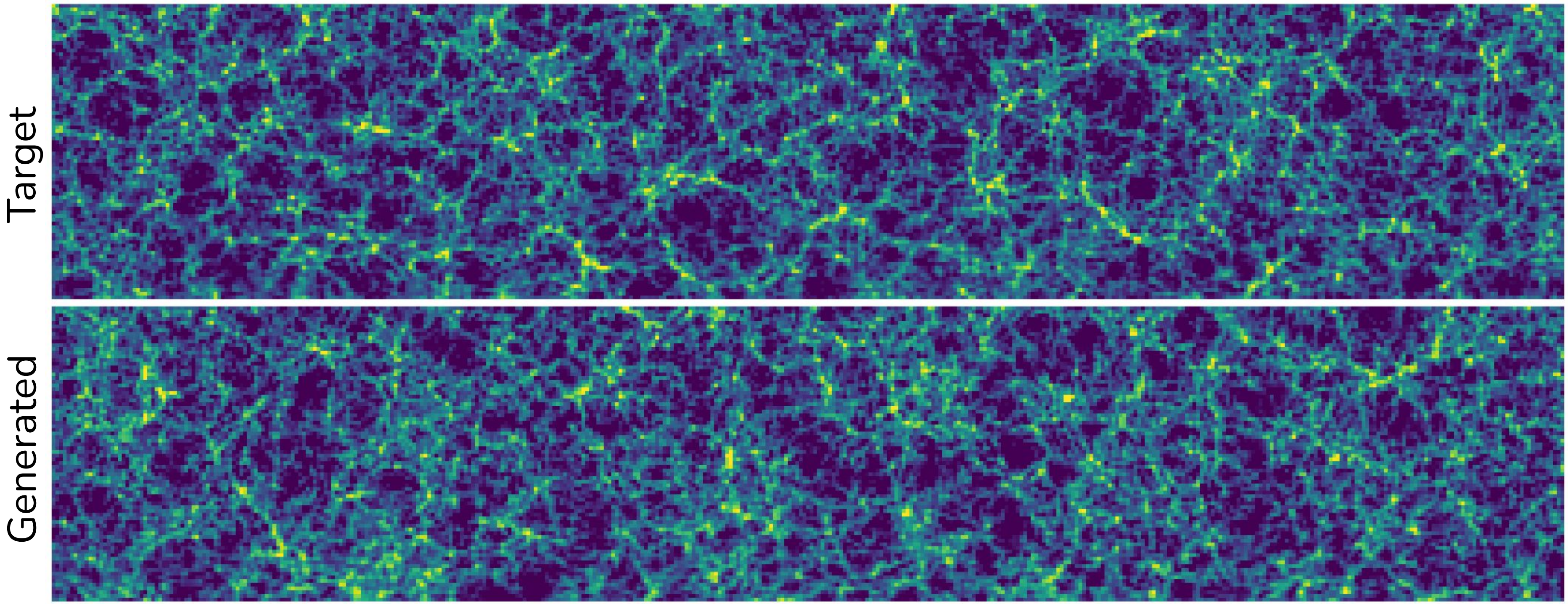
Thermal Sunyaev-Zeldovitch effect (tSZ) [Ade et al. 2019]

Venus surface



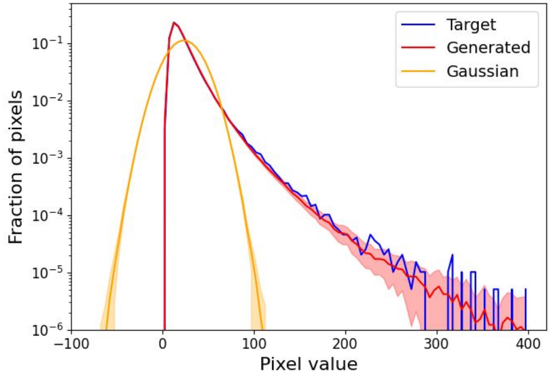
~ few seconds per map at Nside=128

Zoom on a patch - LSS map

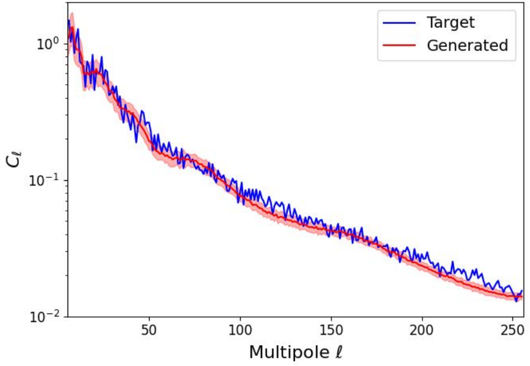


Statistical validation, example on the LSS map

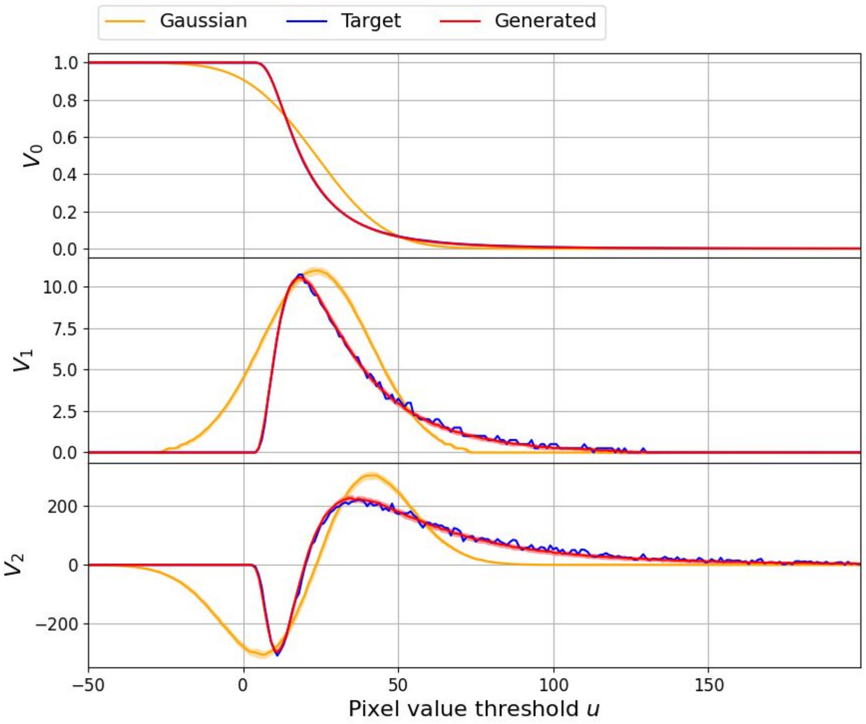
PDF



Power spectrum

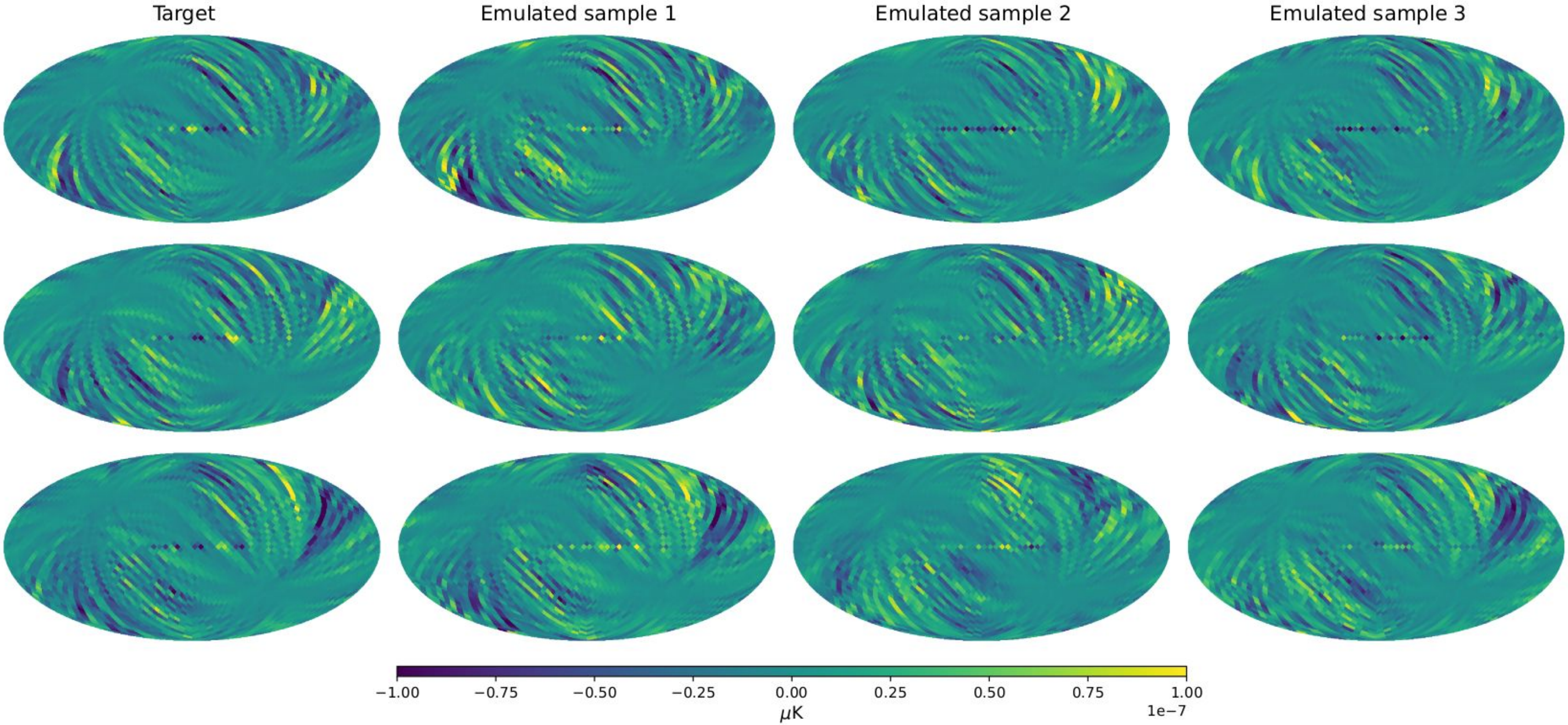


Minkowski functionals



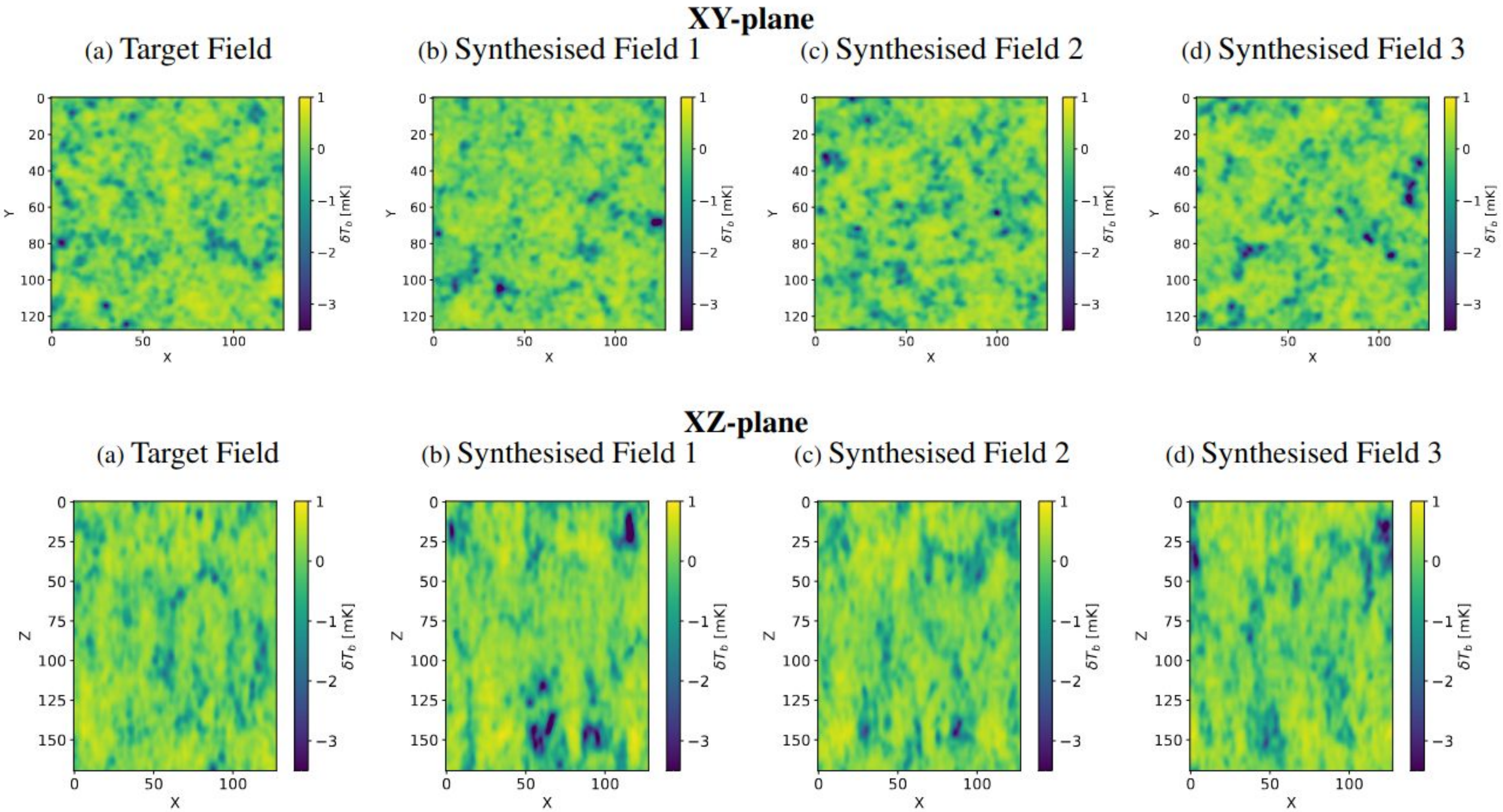
Data augmentation: LiteBIRD noise simulations

[Campeti et al. 2025]



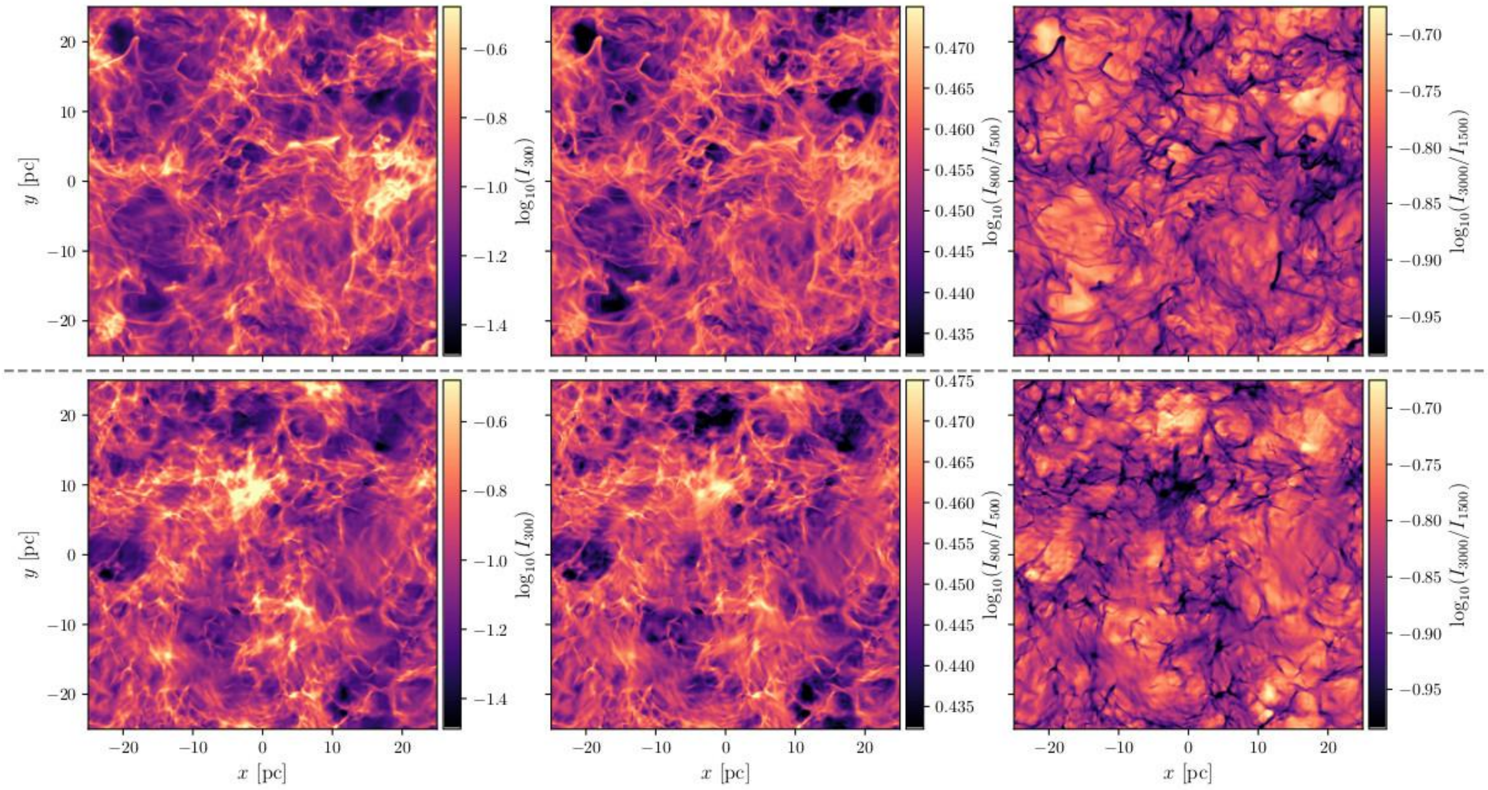
Spectroscopic data: 21cm signal from epoch of reionization

[Hothi et al. 2025]



Multi-frequency data : dust intensity with cross statistics

[Regaldo et al. 2022]



Dust intensity at 300/500/800/1300/1500 GHz with cross statistics

Component separation with Scattering Transforms

Component separation: general framework

$$data = signal + contaminant$$
$$d = s + c$$

We assume we have realisations of the contaminant $\{c_i\}$ => we know $\phi(c)$

Goal : recover the signal s

Method :

- From the data we can compute $\phi(d)$
- Gradient descent : find a map $\tilde{s}$ such that

$$\langle \phi(\tilde{s} + c_i) \rangle_i \simeq \phi(d)$$

$$\phi(\tilde{c}) \simeq \phi(c) \text{ with } \tilde{c} = d - \tilde{s}$$

- Start for example the first iteration with

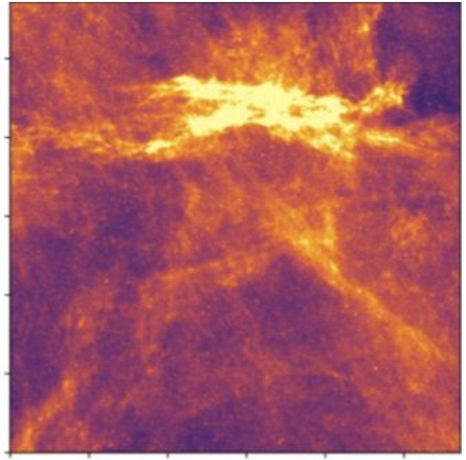
$$\tilde{s}^0 = d$$

Component separation in Hershel data

[Auclair et al. 2024]

2 deg on the sky at 250 μm

$$\textcolor{magenta}{data} = \textcolor{blue}{dust} + \textcolor{red}{CIB}$$



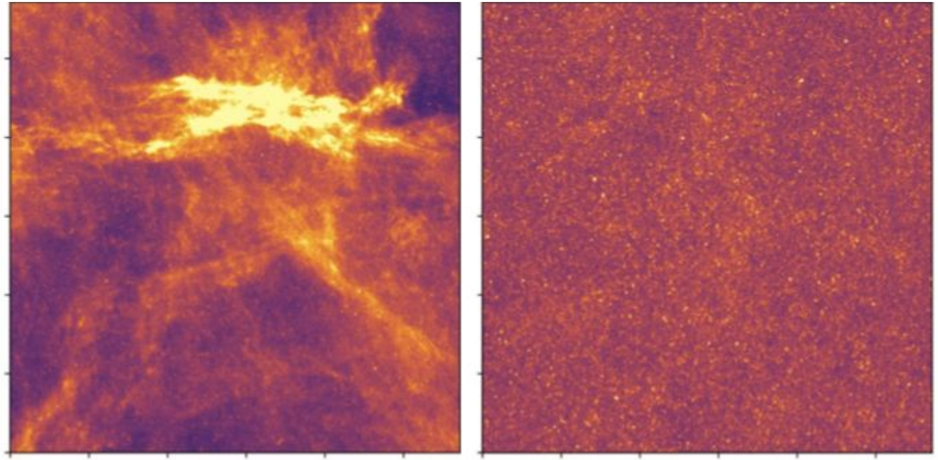
Component separation in Hershel data

[Auclair et al. 2024]

2 deg on the sky at 250 μm

$$data = dust + CIB$$

Isolated CIB observation



Compute ST statistics

$$\phi(c)$$

Component separation in Hershel data

[Auclair et al. 2024]

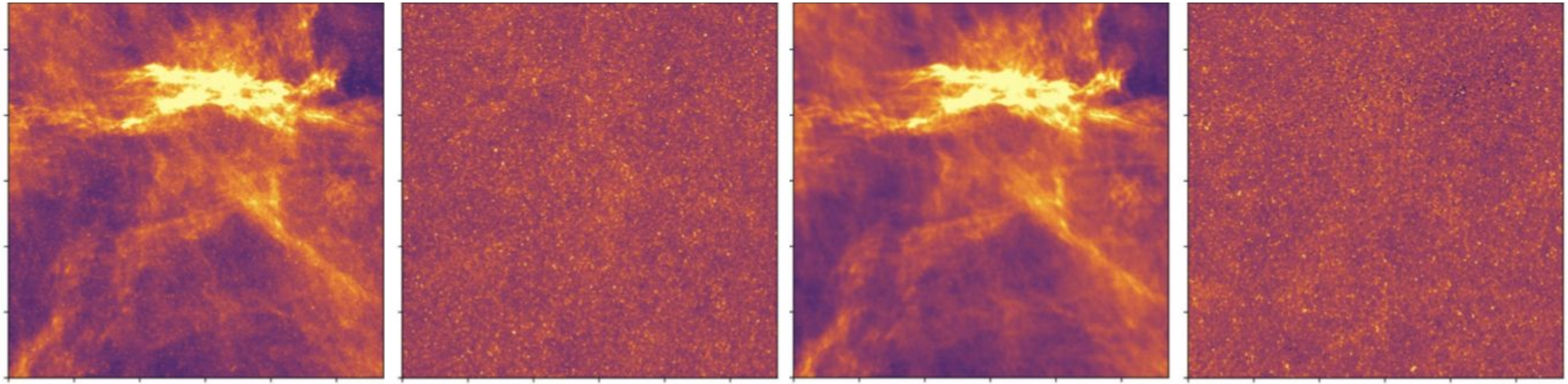
2 deg on the sky at 250 μm

$$data = dust + CIB$$

Isolated CIB observation

Separated Dust

Separated CIB



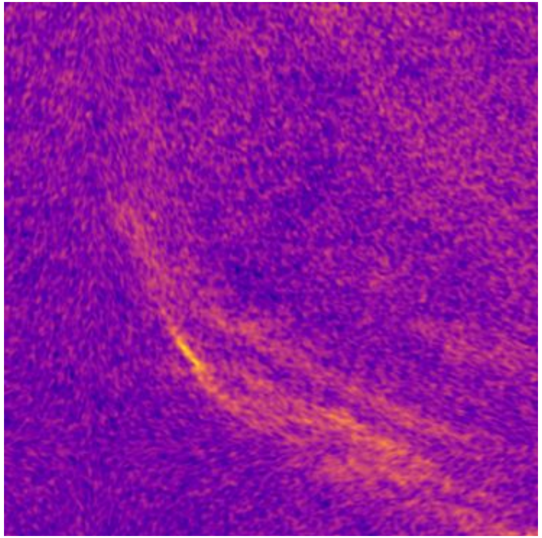
Component separation using only non Gaussian spatial information.
Using only 2 observations.

Planck polarization maps

[Tsouros et al. in prep]

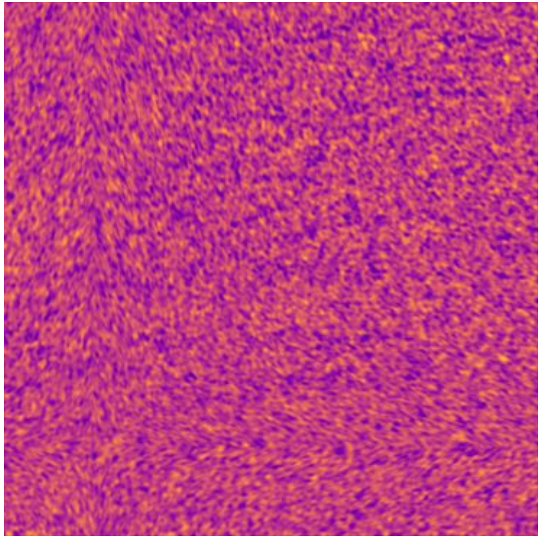
Polarization Q at 353 GHz

$$data = dust + CMB + Noise$$

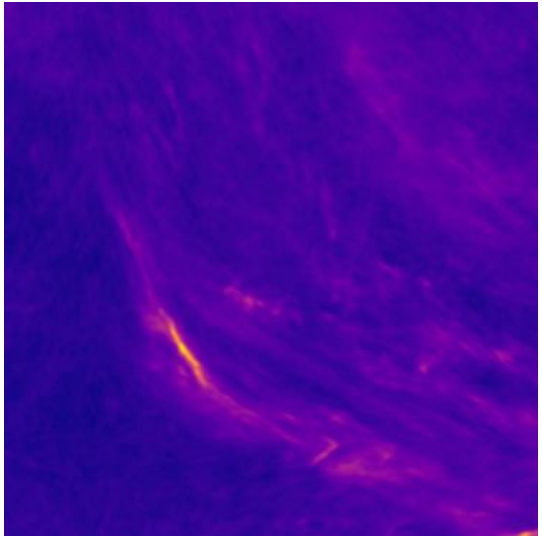


Contaminant: CMB + Noise

$$\{c_i\}$$



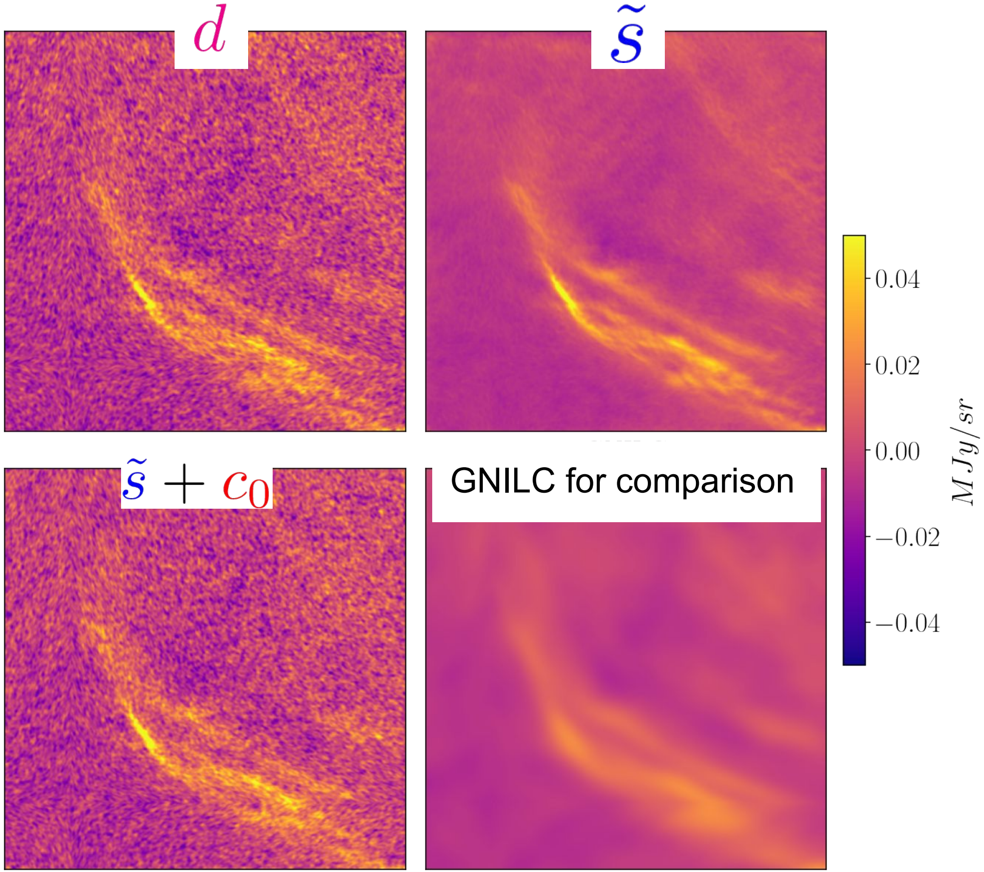
Additional information: Intensity I at 857 GHz



Slide from A. Tsouros

Planck polarization maps

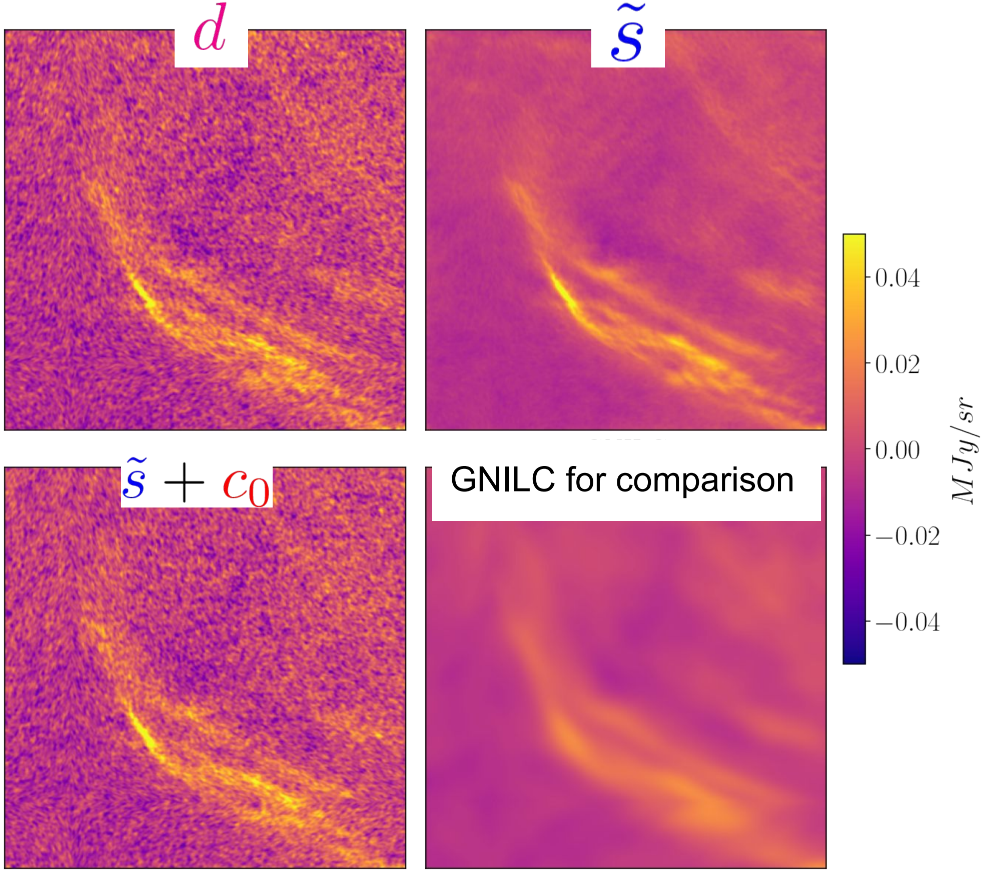
[Tsouros et al. in prep]

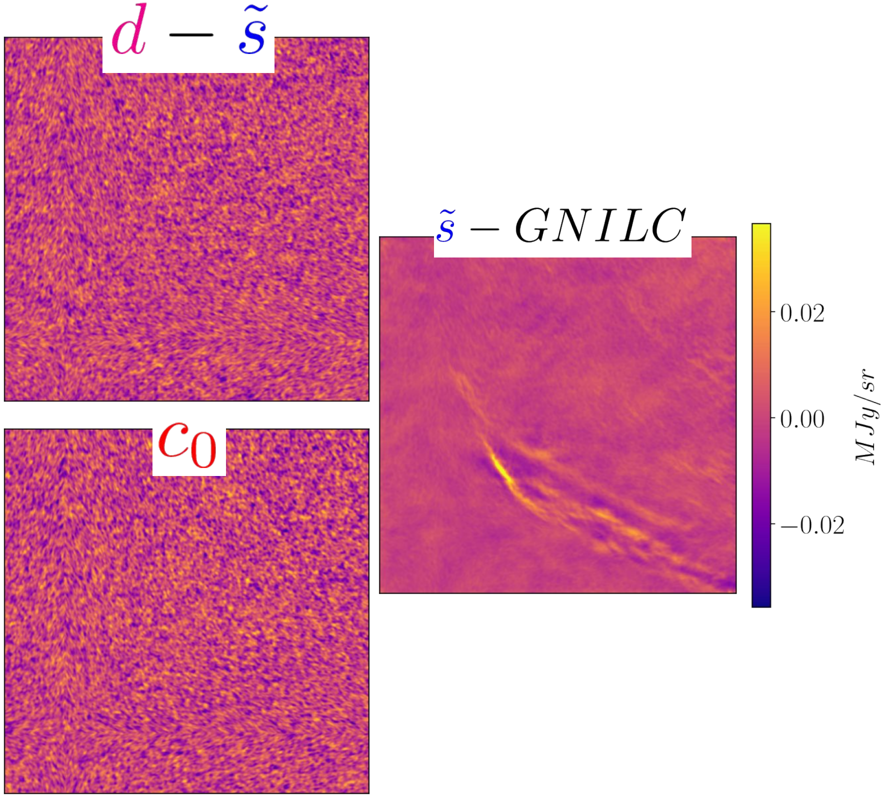


Images from A. Tsouros

Planck polarization maps

[Tsouros et al. in prep]





Images from A. Tsouros

Sampling of the posterior of scattering statistics

Bayesian approach and posterior sampling

[Pierre et al. in prep]

$$data = f(signal)$$

Toy model: noise + PSF + masks

Goal : sample the posterior distribution $P(\phi(s)|\phi(d))$

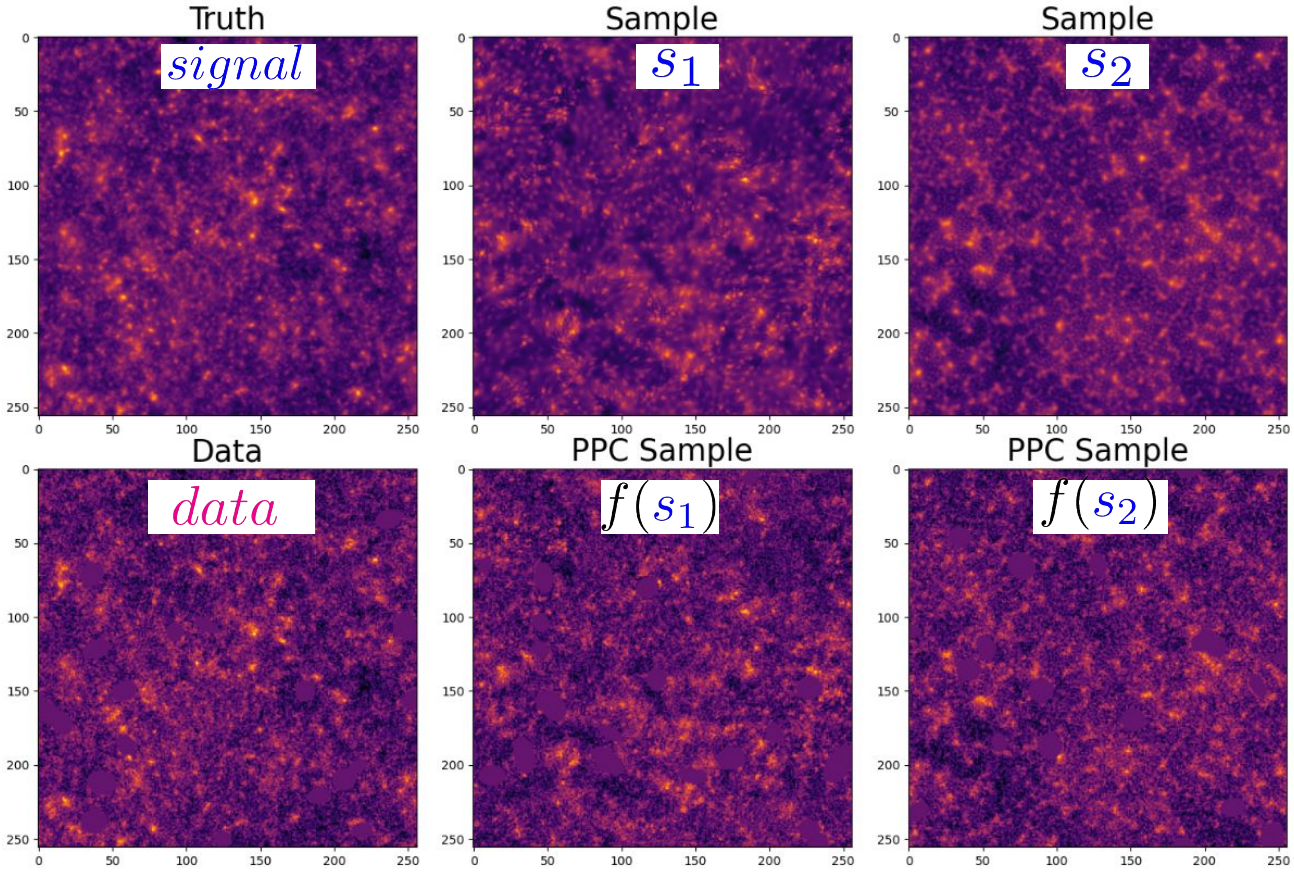
Then, from each $\phi(s)$ you can sample the posterior distribution $P(s)$ with a generative model

Hypothesis :

- A single observed data
- Probabilistic forward model known
- No prior on the signal

Bayesian approach and posterior sampling

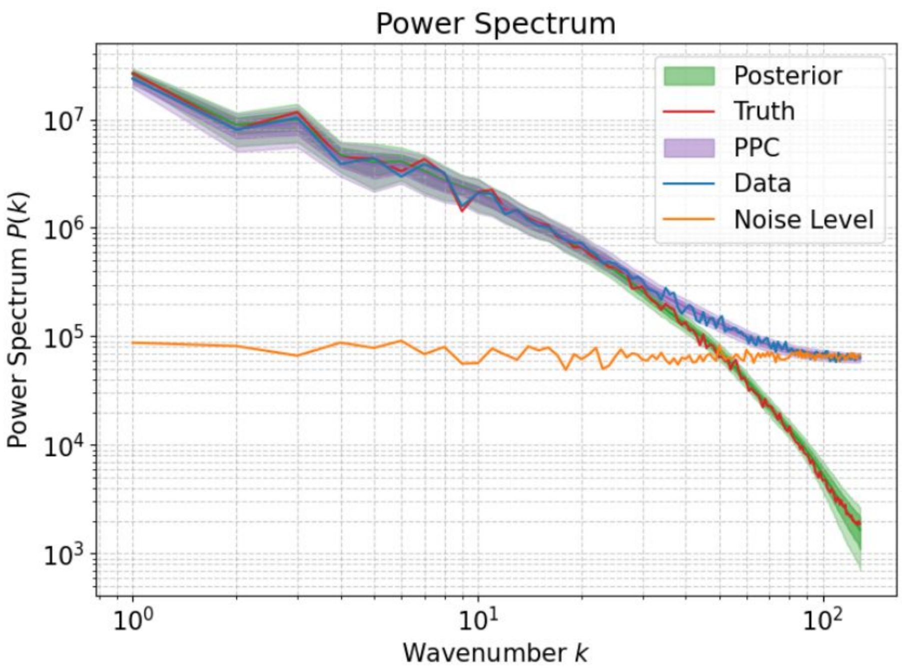
[Pierre et al. in prep]

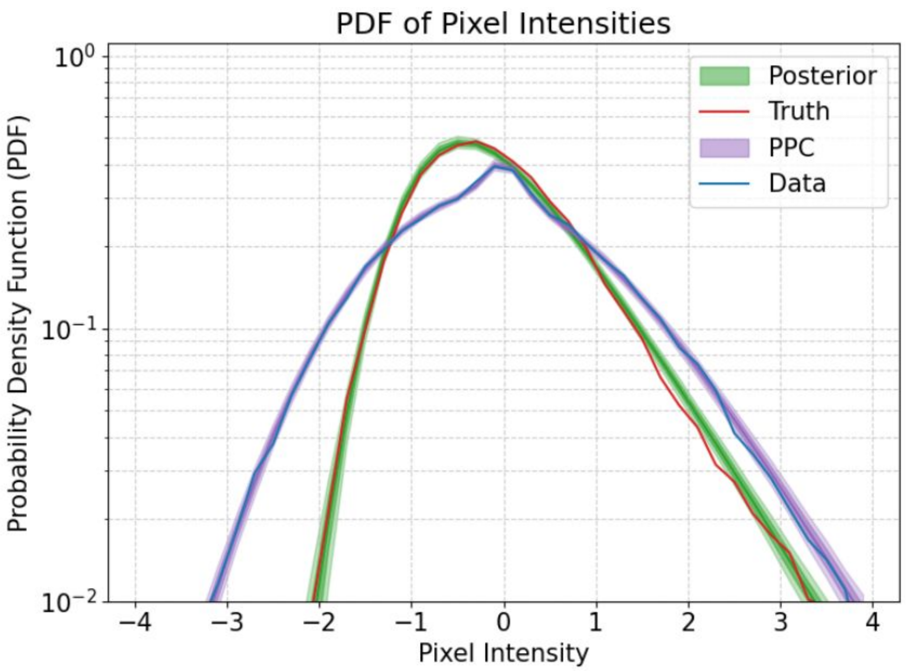


PPC = Posterior Predictive Check

Bayesian approach and posterior sampling

[Pierre et al. in prep]





Power spectrum and PDF recovered

Summary

- **Scattering transforms are efficient low variance summary statistics to characterize non-Gaussian fields.**
- **Generative models with scattering transforms have been developed on various data formats**
- **Large potential for component separation**
- **A bayesian approach is under development**

Thank you for your attention!

Back-up

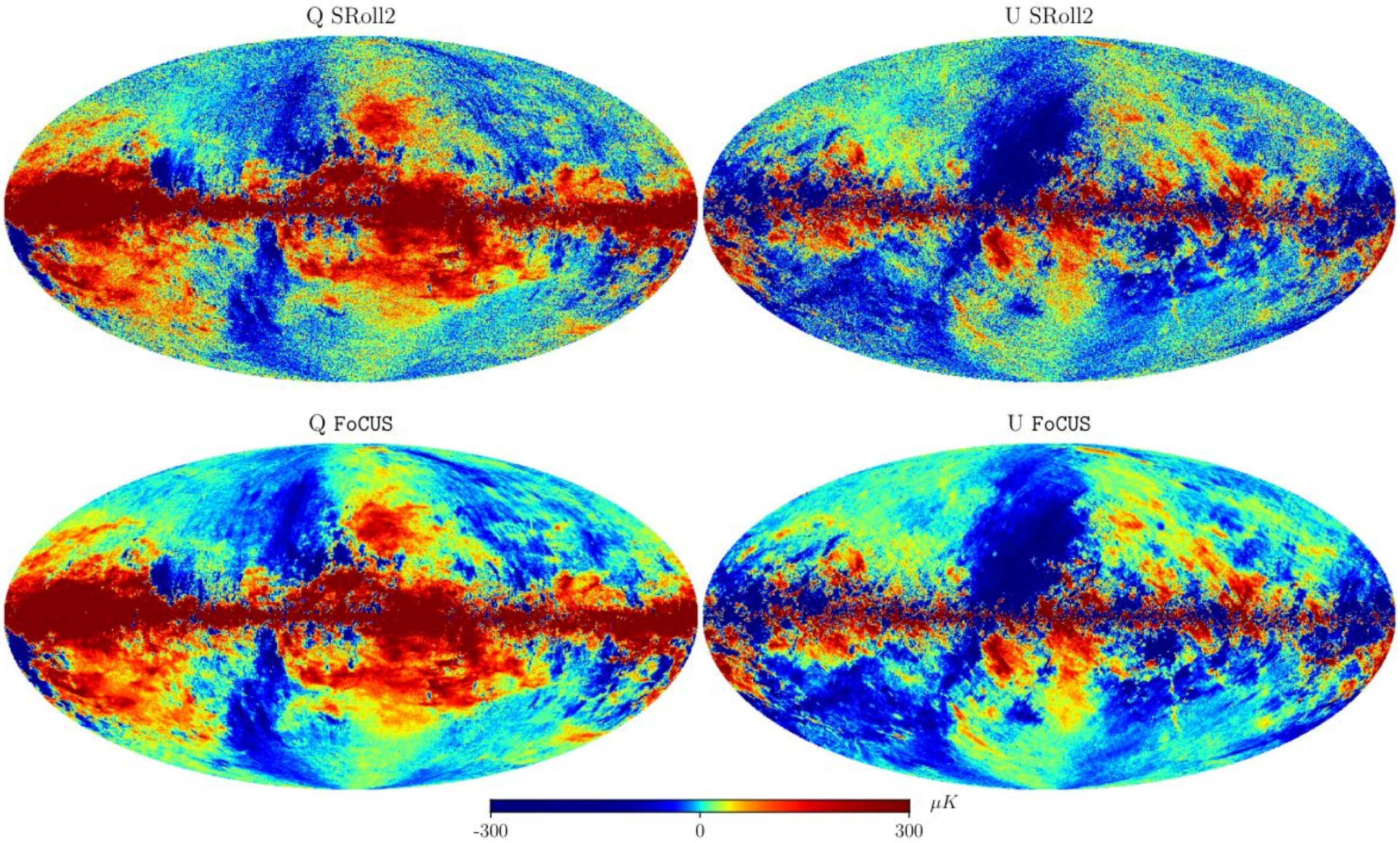
Statistical denoising of the Planck maps

[Delouis et al. 2023]

Input: Noise templates for Planck maps.

Deterministic denoising up to SNR≃ 0.1.

Statistical denoising up to SNR≃0.01.



Power spectra

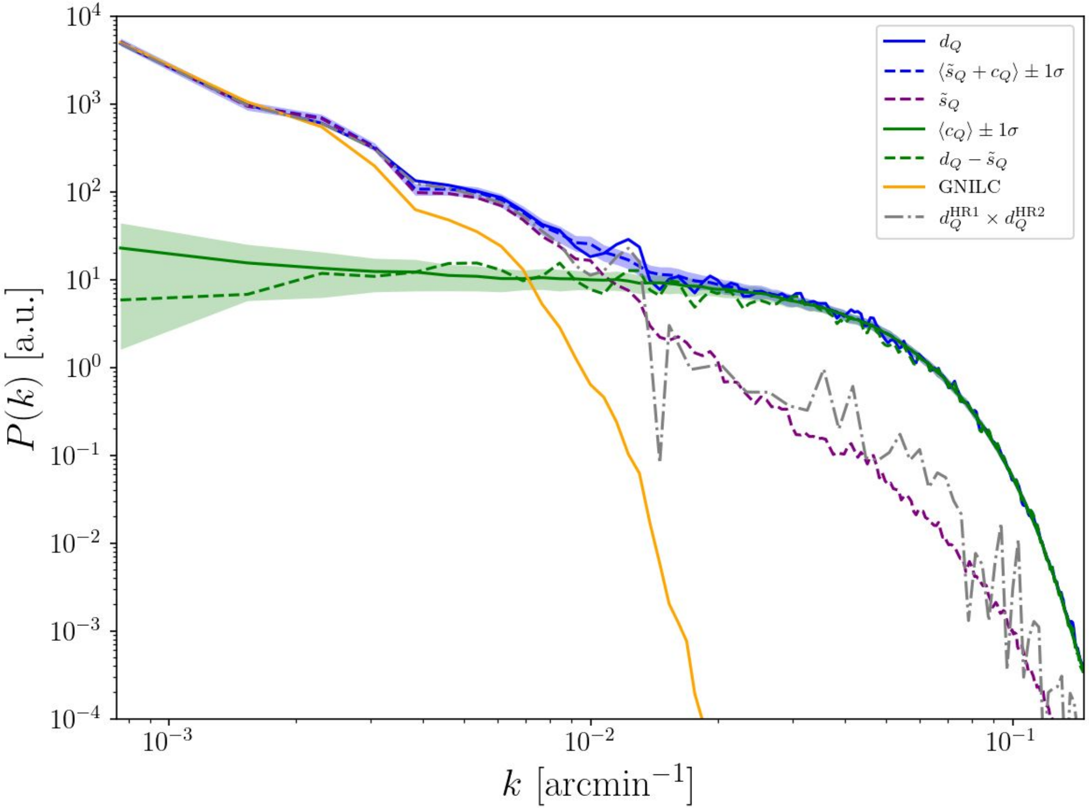


Image from A. Tsouros

Constraints

What we require

$$\Phi(d_Q) \simeq \left\langle \Phi(\tilde{s}_Q + c_Q) \right\rangle_{c_Q}$$

$$\Phi(d_Q - \tilde{s}_Q) \simeq \left\langle \Phi(c_Q) \right\rangle_{c_Q}$$

$$\Phi(d_Q, d_I) \simeq \left\langle \Phi(\tilde{s}_Q + c_Q, d_I) \right\rangle_{c_a}$$

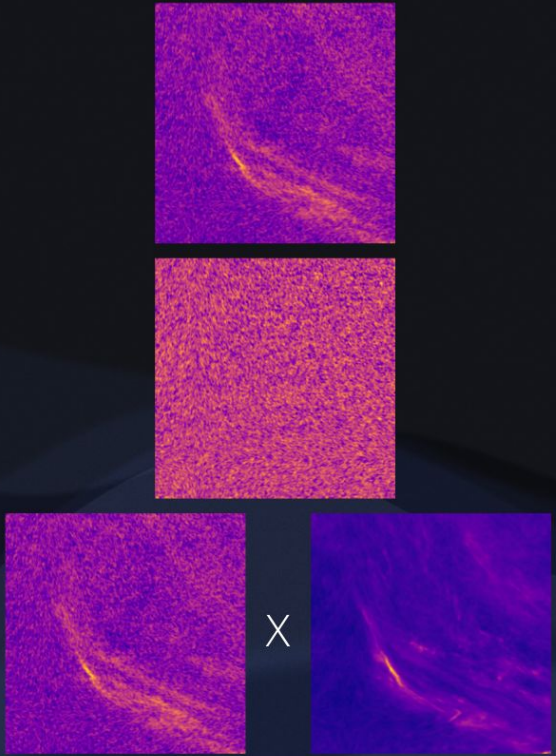


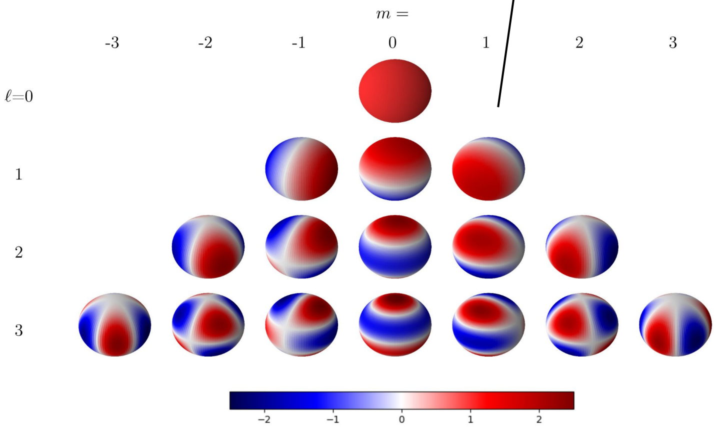
Image from A. Tsouros

Wavelet filters on the sphere

[Price & McEwen 2023]

Spherical harmonic transform:

$$\Psi(\theta,\varphi) = \sum_{\ell m} \Psi_{\ell m} Y_{\ell m}(\theta,\varphi)$$



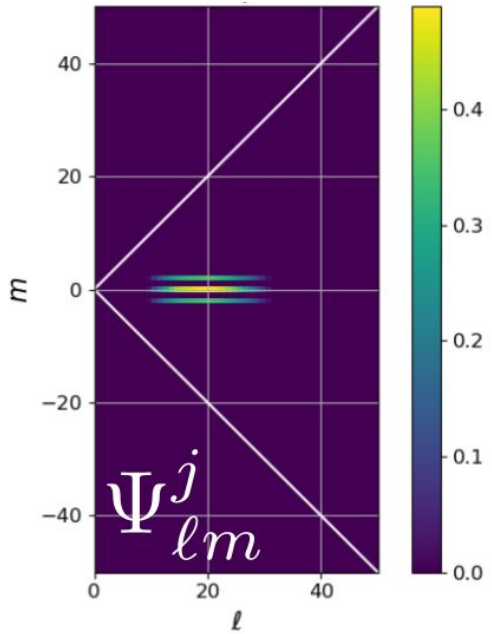
$$-\ell \leq m \leq \ell \qquad \ell \sim \frac{\pi}{\theta}$$

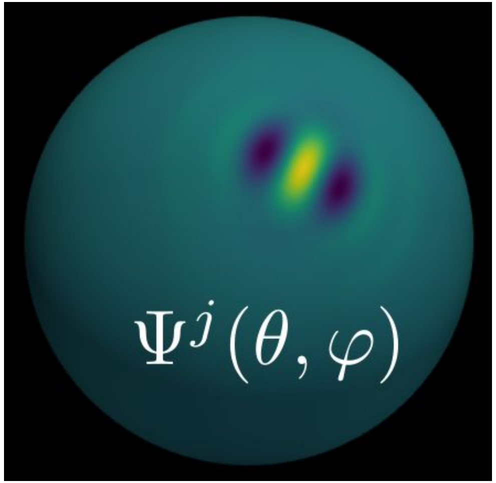
Wavelet filters:

Filter scale

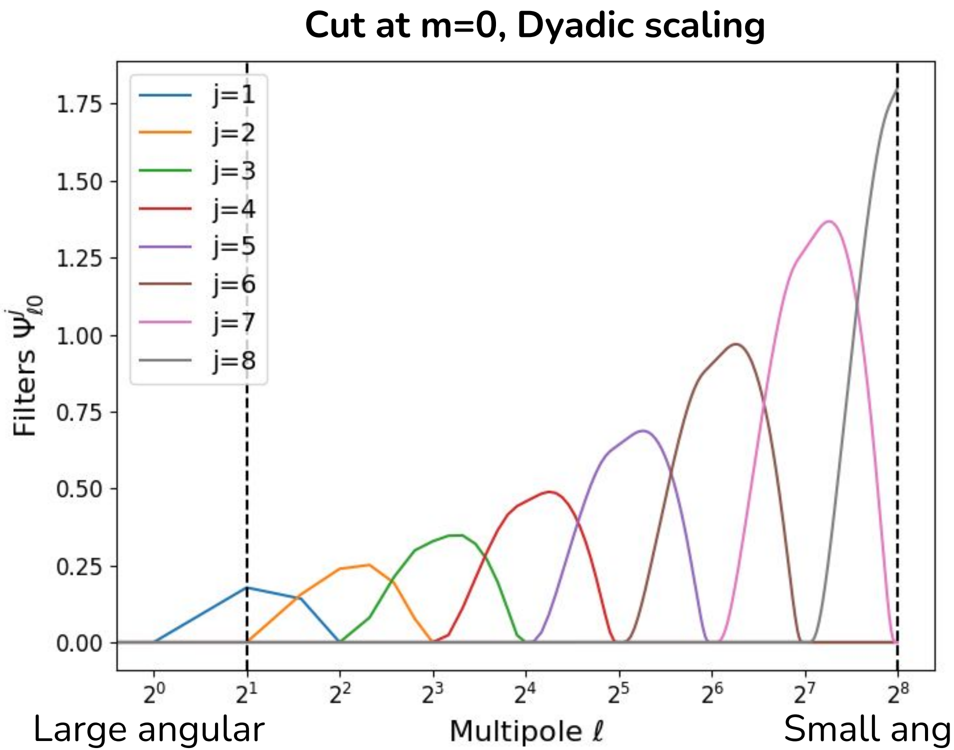
$$\Psi^j_{\ell m} = \sqrt{\frac{2\ell+1}{8\pi^2}}\,\kappa^j_\ell\,\zeta_{\ell m}$$

Directionality



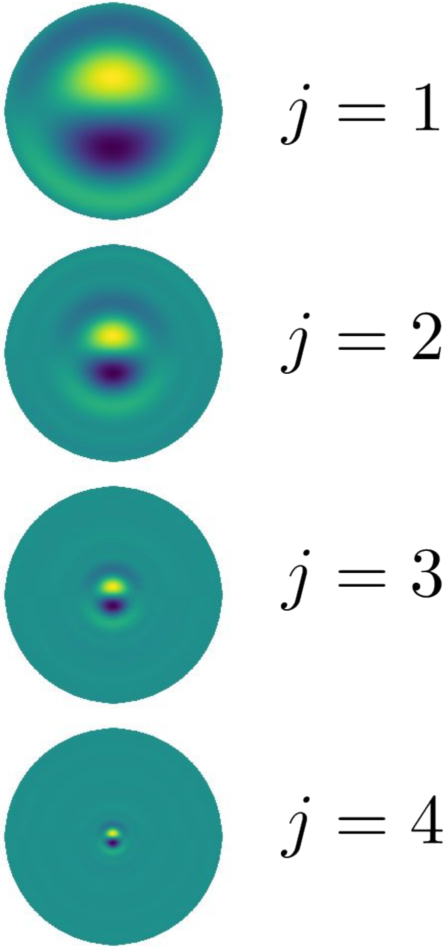


Filter set scaling



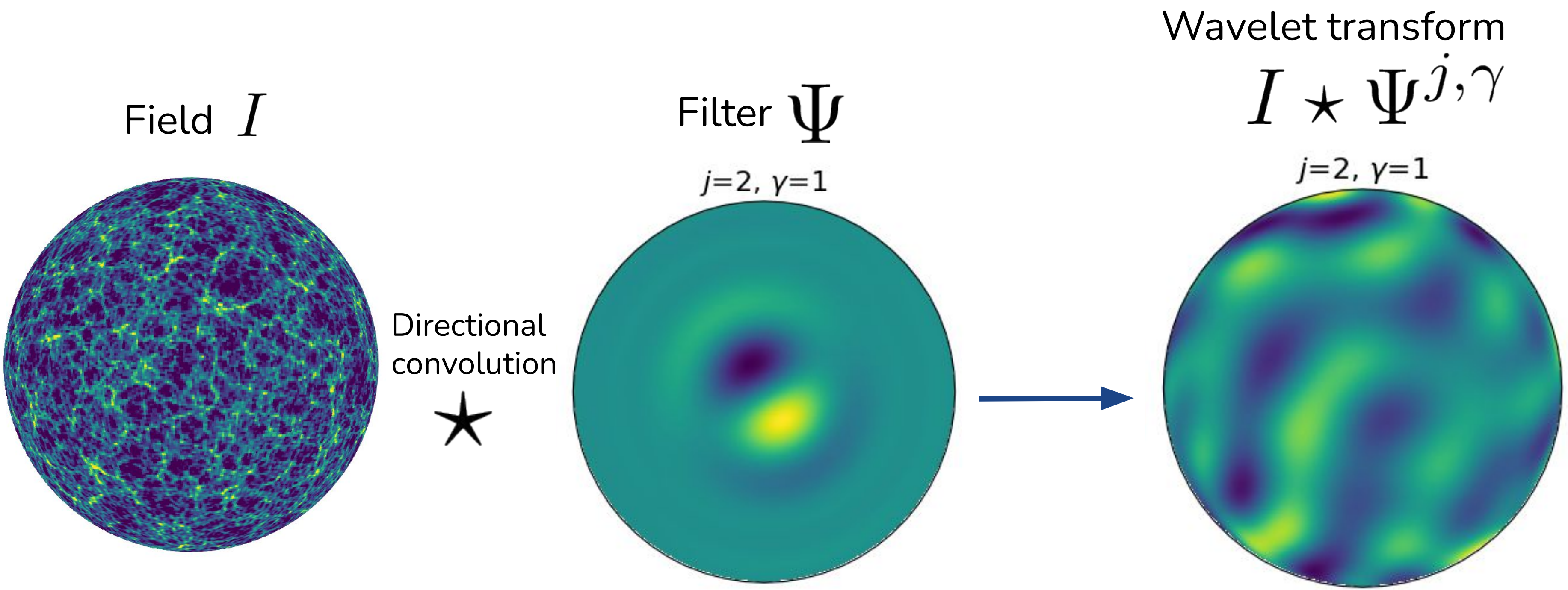
Large angular scales

Small angular scales



Wavelet transform

[McEwen et al. 2015, Price & McEwen 2023]
S2FFT Python package



Directional convolution performed in harmonic space.

Scattering covariance coefficients S_1, S_2, S_3, S_4

[Morel et al. 2023, Cheng et al. 2023]

Coefficients associated to a single scale and a single angle:

$$S_1^{\lambda_1} = \langle |I \star \Psi^{\lambda_1}| \rangle$$

$$S_2^{\lambda_1} = \langle |I \star \Psi^{\lambda_1}|^2 \rangle$$

Averages over pixels

$$\lambda = (j, \gamma)$$

Scale

Angle

Coefficients to probe the coupling between scales:

$$S_3^{\lambda_1,\lambda_2} = \mathrm{Cov}\left[I \star \Psi^{\lambda_1}, |I \star \Psi^{\lambda_2}| \star \Psi^{\lambda_1}\right]$$

$$S_4^{\lambda_1,\lambda_2,\lambda_3} = \mathrm{Cov}\left[|I \star \Psi^{\lambda_3}| \star \Psi^{\lambda_1}, |I \star \Psi^{\lambda_2}| \star \Psi^{\lambda_1}\right]$$

For instance, with $j = [1, 8]$ and $\gamma = [1, 5]$

$$\Rightarrow \sim 10^3 \text{ coefficients.}$$

Directional convolution on the sphere

[McEwen et al. 2015, Price & McEwen 2023]
S2FFT Python package

$$W^{j}_{\ell mn} = \frac{8\pi^2}{2\ell+1} I_{\ell m} \Psi^{j*}_{\ell n}$$

Inverse Wigner Transform

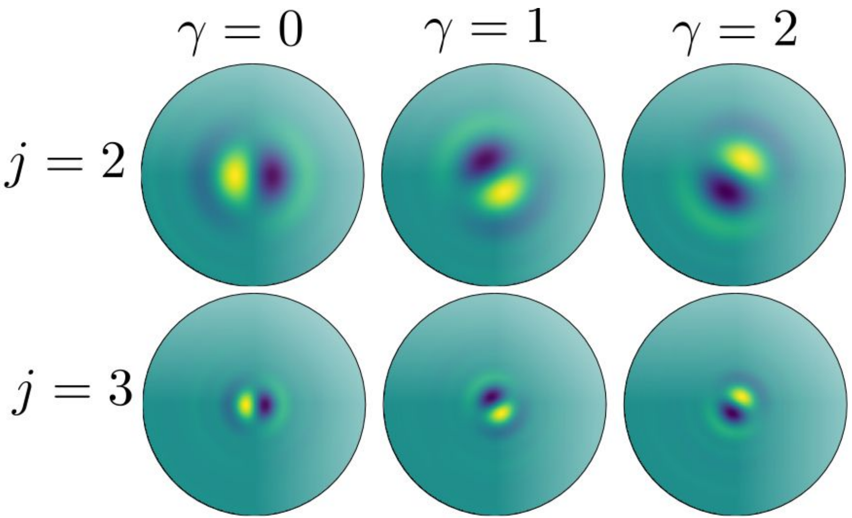
$$W^{j}(\alpha, \beta, \gamma) \rightarrow W^{j\gamma}(\theta, \varphi)$$

Euler angles

Scale

Angle

Convolution of a Dirac map:



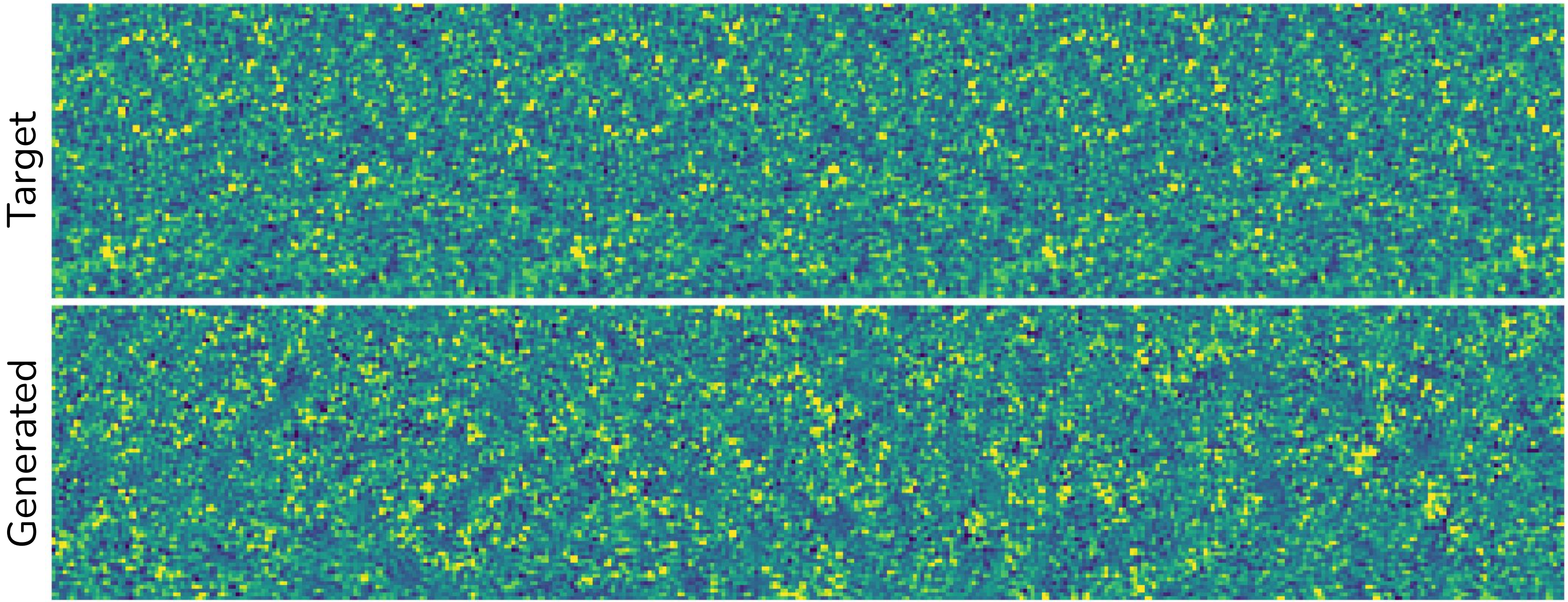
Computational benchmarking

Table 1. Computational benchmarking.

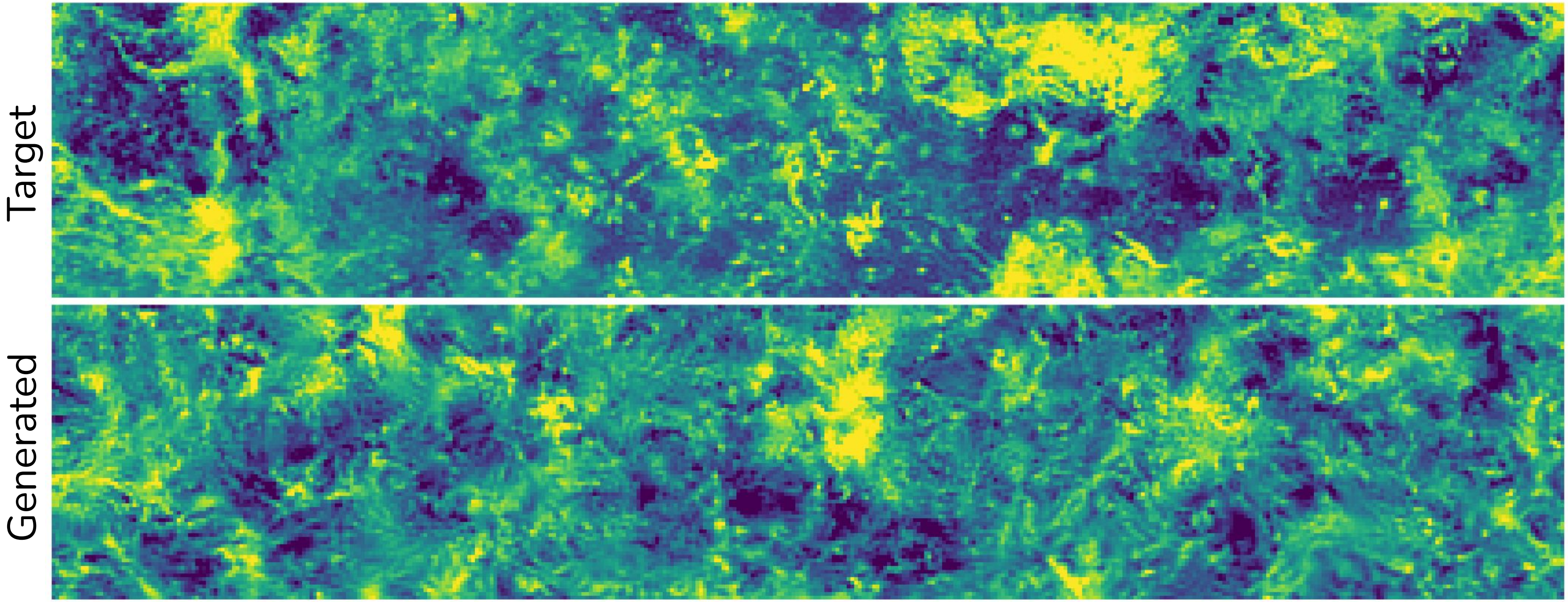
Pre-compute Mode			
Bandlimit	Forward	Gradient	JIT Compilation
256	15 ms	30 ms	20 s
512	100 ms	200 ms	25 s
Recursive Mode			
Bandlimit	Forward	Gradient	JIT Compilation
256	120 ms	300 ms	90 s
1024	5 s	10 s	3 m
2048	20 s	50 s	6 m

Notes. Results of the SC transform provided by `s2scat`. These results were recovered on a single NVIDIA A100 40GB GPU, although it is possible to run across multiple GPUs. In our analysis we generate spherical images through 400 iterations to be conservative. In practice, however, we find that $\sim$ 100 iterations is typically sufficient, in which case an image at $L = 256$ can be generated in $\sim$ 4s. Furthermore, batched generation can dramatically decrease per sample compute time. For example, 20 images at $L = 256$ can be generated in $\sim$ 12s, corresponding to $\sim$ 0.5s per sample.

Zoom on a patch - tSZ



Zoom on a patch - Venus



Comparison with a Gaussian generative model

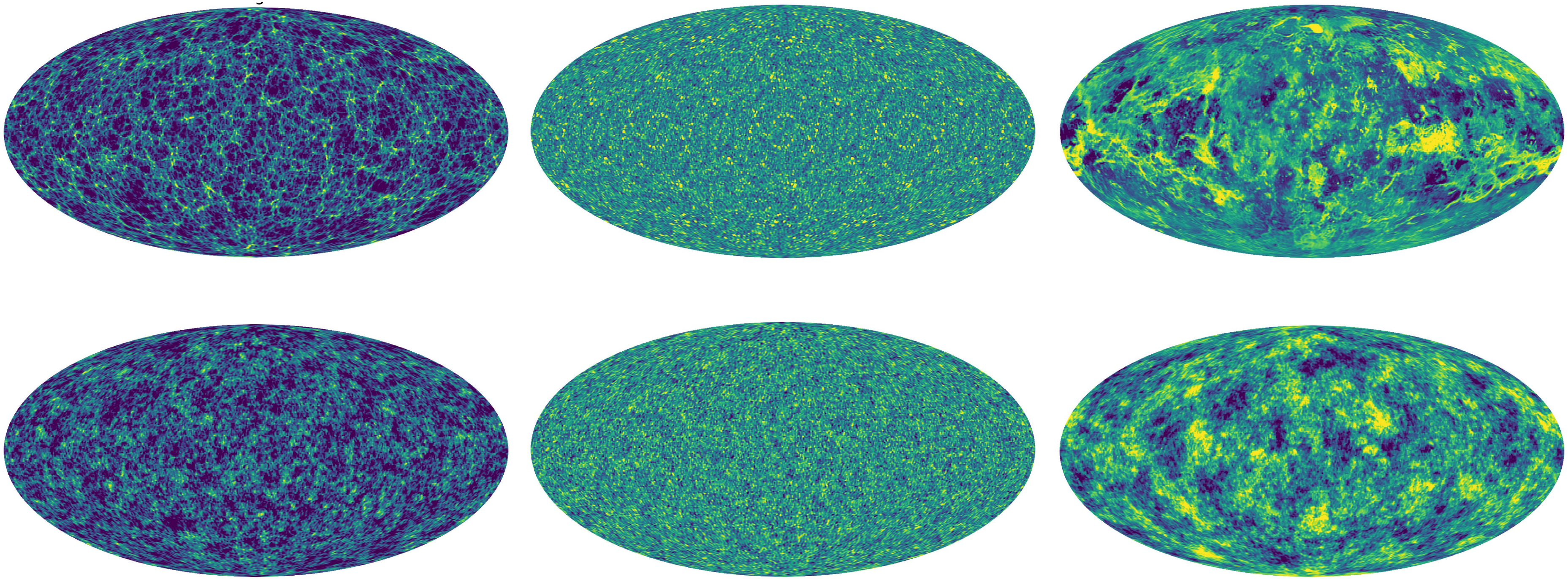
Large Scale Structures
N-body simulation [Kacprzak et al. 2022]

Thermal Sunyaev-Zeldovich effect (tSZ) [Ade et al. 2019]

Venus surface

Target

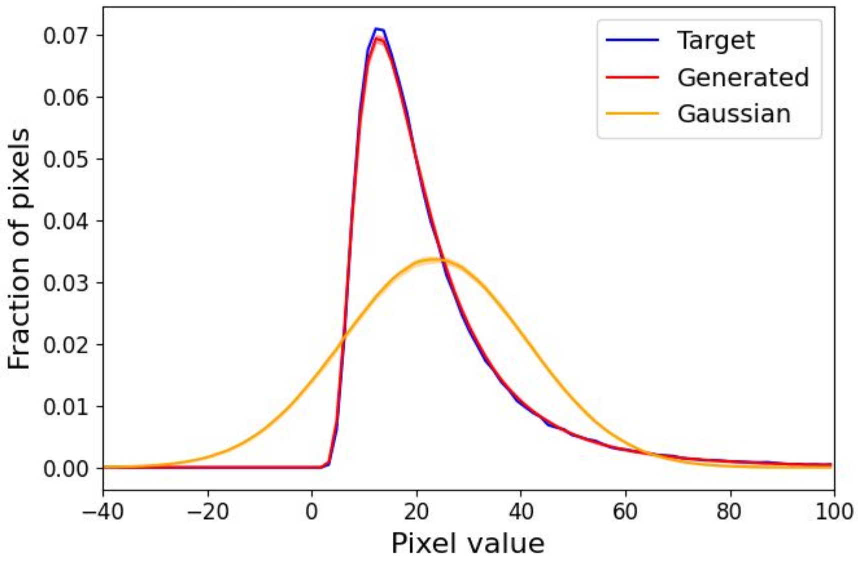
Generated Gaussian



Statistical validation - Probability Density Function (PDF)

LSS

y-axis linear scaling



y-axis log scaling