Scattering transforms, recent applications in astrophysics and cosmology

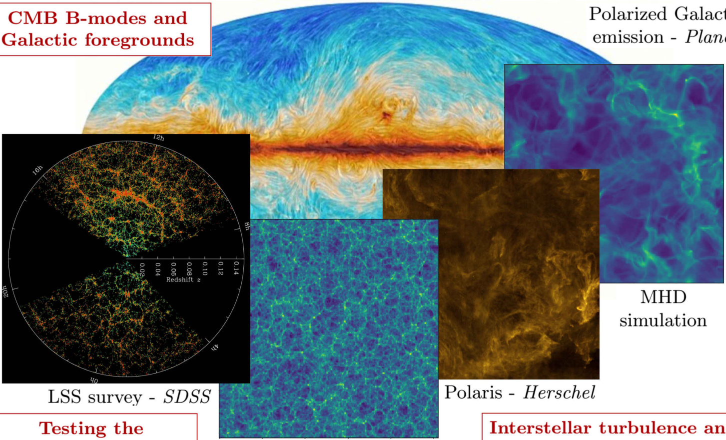
Erwan Allys
LPENS & ENS Center for Data Sciences, Paris

IFPU focus week
Triestre, June 29th 2022

Outline

CMB B-modes and Galactic foregrounds

Polarized Galactic emission - *Planck*



MHD simulation

LSS survey - *SDSS*

Polaris - *Herschel*

LSS simulation - *Quijote*

Testing the cosmological model

Interstellar turbulence and galactic cycle of matter

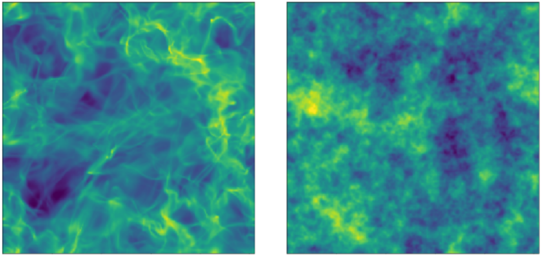
Common difficulty : non-linearity $\Rightarrow$ non-Gaussianity
$\rightarrow$ Important lever arm for a lot of astrophysical objectives

Which statistics for non-Gaussian astrophysics?

- **(Additional) Requirements for summary statistics**
  - Being expressive
    $\rightarrow$ To model/characterize a process (*sufficient statistics*)
    $\rightarrow$ To classify/infer parameters (e.g. *Fisher information*)
  - Low dimensional and low-variance
    $\rightarrow$ Expressivity from a single image?
    $\rightarrow$ Carefully concatenate $\neq$ statistics *(Park, EA+, 2022)*
  - Stability
    $\rightarrow$ To noise
    $\rightarrow$ To small deformations
  - Physically interpretable

$$\text{physical properties} \underset{\text{guide the construction of}}{\overset{\text{simplify characterization of}}{\begin{array}{c}\xleftarrow{\hspace{6em}}\\ \xrightarrow{\hspace{6em}}\end{array}}} \text{statistical description}$$

Which role for the power spectrum ?



Champs de même spectre de puissance

- **Power Spectrum: not sufficient, but essential**
  - Do not characterize structures
  - Still excellent physical tool to check
  - Provide benchmark and sandbox

Beyond PS $\not\Rightarrow$ Fourier analysis must be abandoned

Which statistics for non-Gaussian astrophysics?

- **Usual statistics reach their limits**
  - Very efficient for perturbative regime
  - Non-perturbative non-Gaussian regime:
    $\rightarrow$ High dimension/high variance
    $\rightarrow$ Lack of generic efficient methods

$\Rightarrow$ Active area of research
$\Rightarrow$ Increasing use of Machine Learning tools

Which statistics for non-Gaussian astrophysics?

- **Usual statistics reach their limits**
  - Very efficient for perturbative regime
  - Non-perturbative non-Gaussian regime:
    → High dimension/high variance
    → Lack of generic efficient methods

  ⇒ Active area of research
  ⇒ Increasing use of Machine Learning tools

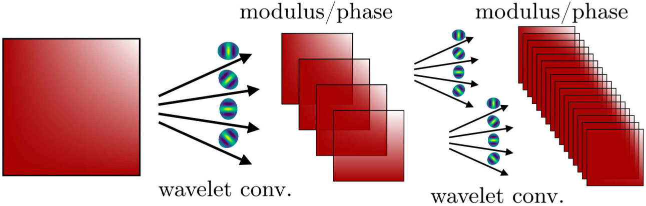
- **Machine Learning for astrophysics?**
  - Impressive successes with non-Gaussian fields
  - Require good quality dataset
    → often small amount of observational data
    → realistic simulations not always available
  - Physical interpretation is problematic

  ⇒ A lot of application are still very difficult

→ Can we find a solution in-between?

Scattering transform statistics

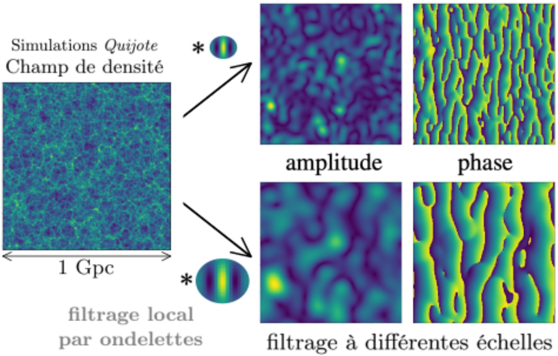
- **Scattering Transform (ST) statistics (Mallat+, 2010+)**
  - Developed in data sciences
  - Inspired from neural networks constructions
    $\rightarrow$ efficient characterization and reduced variance
  - Do not need any training stage
    $\rightarrow$ explicit mathematical form and interpretability



$\rightarrow$ Wavelet filters separating the different scales
$\rightarrow$ Coupling between scales with non-linearities

Scattering Transform statistics

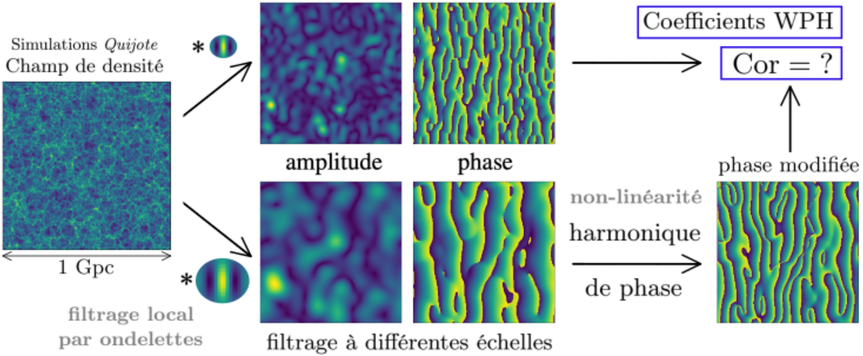
- Correlations/successive convolutions unusable



→ Need non-linearity to characterize scale interactions

Scattering Transform statistics

- **Wavelet Phase Harmonics (WPH) and phase alignment**



→ **Efficient characterization of scale interactions**

Scattering Transform statistics

- **Coming back at the requirements**
  - Being expressive
    $\rightarrow$ We'll see ! :-)
  - Low dimensional and low-variance
    $\rightarrow$ We'll see ! :-)
  - Stability
    $\rightarrow$ **Linearize noise and small deformations**
    $\rightarrow$ Logarithmic bins in Fourier + contractive operators
  - Physically interpretable
    $\rightarrow$ ***Natural* notion of interaction between scales**

Different scientific objectives

- **Different methodological objectives**
  (with increasing subjective difficulty...)
  - Estimate physical parameters
  - Model an astrophysical process
  - Separate different components
  - Constraint a physical model

Different scientific objectives

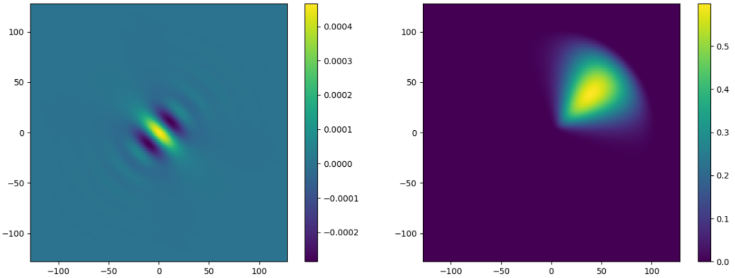
- **Different methodological objectives**
  (with increasing subjective difficulty...)
  - Estimate physical parameters
  - Model an astrophysical process
  - Separate different components
  - Constraint a physical model

$\rightarrow$ **ST reached or improved state-of-the-art on each topic**

Outline

Wavelet transform and multiscale analysis

- **Set of wavelets** $\psi_{j,\theta}(x, y)$
  - Localized wave packet, probing a given spectral domain
  - Set of wavelets obtained by discrete dilations and rotations



$\rightarrow$ $\psi_{j,\theta}$ probe the 2^j scale with orientation θ
$\rightarrow$ Spectral analysis of localized features

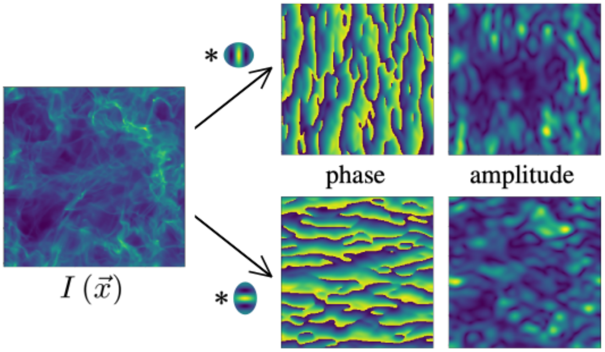
Wavelet transform and multiscale analysis

- **Convolution of a LSS density field $\rho(\vec{x})$ with $\psi_{j,\theta}(\vec{x})$:**

$$\left(\rho \star \psi_{j,\theta}\right)(\vec{x})$$

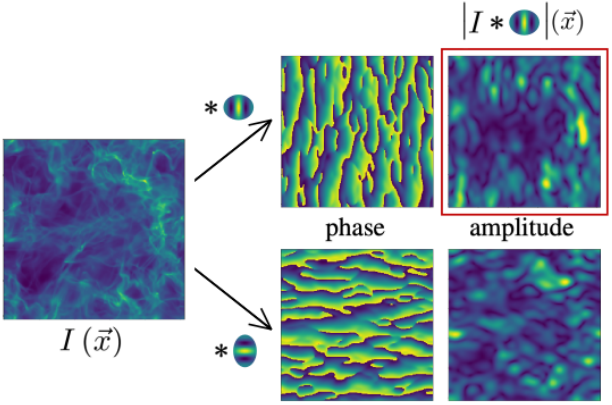
Scattering Transform and amplitude modulation

- **Wavelet Scattering Transform (WST) schematic computation**



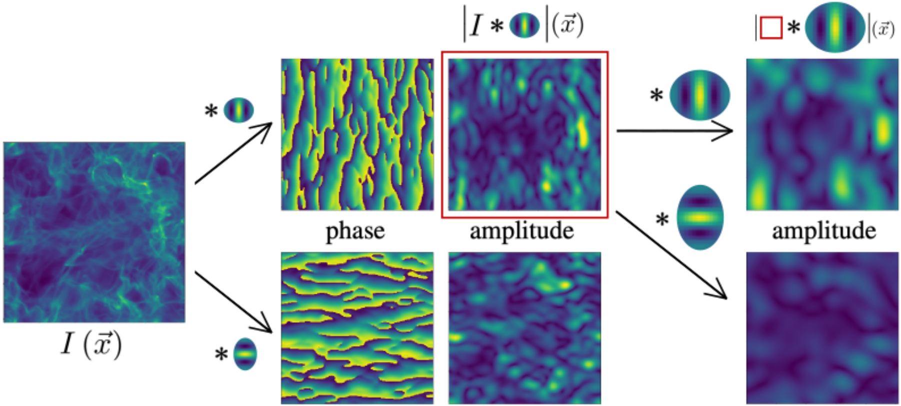
Scattering Transform and amplitude modulation

- **Wavelet Scattering Transform (WST) schematic computation**



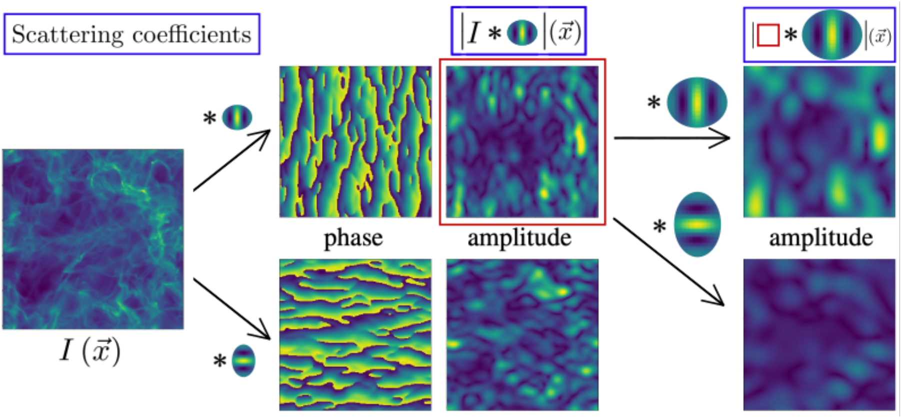
Scattering Transform and amplitude modulation

- **Wavelet Scattering Transform (WST) schematic computation**



Scattering Transform and amplitude modulation

- **Wavelet Scattering Transform (WST) schematic computation**



Scattering Transform and amplitude modulation

- **Computation of the scattering coefficients**
  - From an image $I(\boldsymbol{r})$ $\rightarrow$ $\boxed{S_0 = \int I(\boldsymbol{r})\ \mathrm{d}^2 r}$

Scattering Transform and amplitude modulation

- **Computation of the scattering coefficients**
  - From an image $I(\boldsymbol{r})$ $\rightarrow$ $\boxed{S_0 = \int I(\boldsymbol{r})\ \mathrm{d}^2 r}$
  - Filtering at a first scale with ψ_{j_1,θ_1}

$$|I * \psi_{j_1,\theta_1}|\,(\boldsymbol{r})$$

$\rightarrow$ one image for each (j_1, θ_1) $\rightarrow$ $\boxed{S_1(j_1,\theta_1) = \int |I * \psi_{j_1,\theta_1}|\,(\boldsymbol{r})\ \mathrm{d}^2 r}$

Scattering Transform and amplitude modulation

- **Computation of the scattering coefficients**
  - From an image $I(\boldsymbol{r})$ $\rightarrow$ $\boxed{S_0 = \int I(\boldsymbol{r}) \, \mathrm{d}^2 r}$
  - Filtering at a first scale with ψ_{j_1,θ_1}

$$|I * \psi_{j_1,\theta_1}|(\boldsymbol{r})$$

$\rightarrow$ one image for each (j_1, θ_1) $\rightarrow$ $\boxed{S_1(j_1,\theta_1) = \int |I * \psi_{j_1,\theta_1}|(\boldsymbol{r}) \, \mathrm{d}^2 r}$

  - Filtering at a second scale with ψ_{j_2,θ_2}

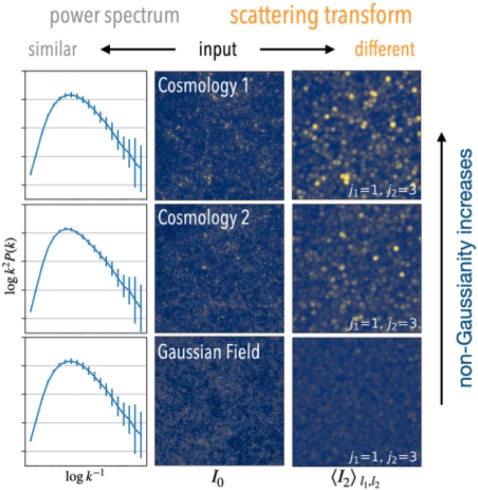
$$\big| |I * \psi_{j_1,\theta_1}| * \psi_{j_2,\theta_2} \big|(\boldsymbol{r})$$

$\rightarrow$ one image for each $(j_1, \theta_1, j_2, \theta_2)$

$$\rightarrow \boxed{S_2(j_1,\theta_1,j_2,\theta_2) = \int \big| |I * \psi_{j_1,\theta_1}| * \psi_{j_2,\theta_2} \big|(\boldsymbol{r}) \, \mathrm{d}^2 r}$$

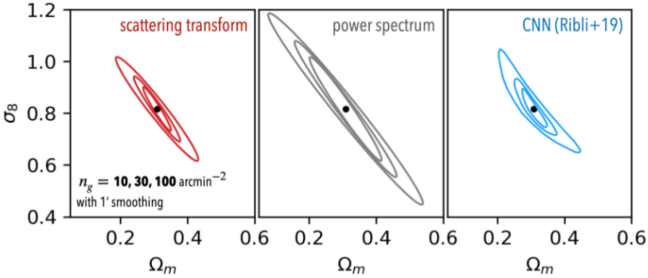
Application to weak lensing convergence fields

- Cosmological inference and weak lensing *(Cheng+, 2020)*



Application to weak lensing convergence fields

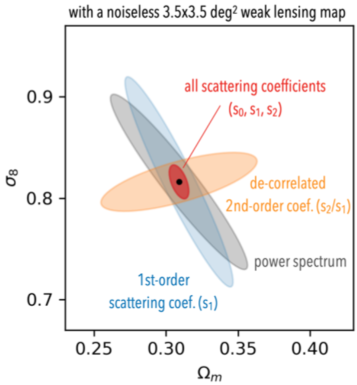
- **Cosmological inference and weak lensing** *(Cheng+, 2020)*
  - Simulated weak-lensing convergence maps
  - Power spectrum, WST (∼100 coeffs), and neural networks
  - Ω_m and σ_8, from Fisher estimates

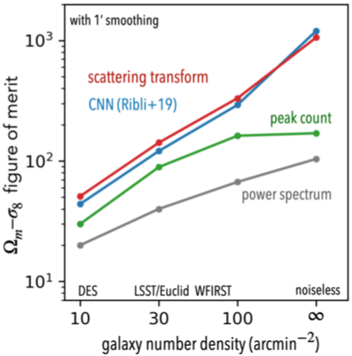


→ **Results on par with state-of-the-art neural networks**

Application to weak lensing convergence fields

- Cosmological inference and weak lensing *(Cheng+, 2020)*

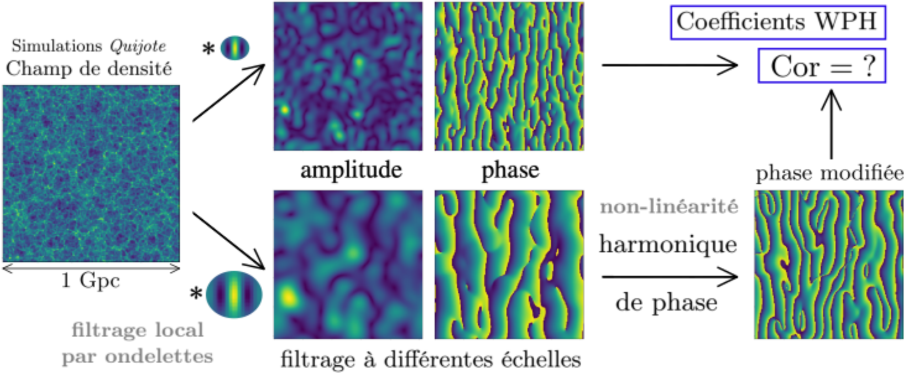




→ Non-Gaussian features identification + stable under noise

Wavelet Phase Harmonics and phase alignment

- **Wavelet Phase Harmonics (WPH) schematic computation**



Wavelet Phase Harmonics and phase alignment

- **Coherent structures from wavelet transform $\rho \star \psi_{\vec{k}}$**
  - Covariance of wavelet transforms:

$$C_{\vec{k}_1,\vec{k}_2} = \mathrm{Cov}\left[\rho \star \psi_{\vec{k}_1}(\vec{x}), \rho \star \psi_{\vec{k}_2}(\vec{x})\right]$$

  - Vanishing if $\vec{k}_1 \neq \vec{k}_2$ since maps band-passed at $\neq$ frequencies

- **Non-linearities required to couple different wavelet bands**

Wavelet Phase Harmonics and phase alignment

- **Phase harmonics of wavelet coefficients**
  - Take $z = |z|\mathrm{e}^{i\cdot\arg(z)} \in \mathbb{C}$, its p^{th} phase harmonic is:

$$[z]^p = |z|\mathrm{e}^{i\cdot p\cdot\arg(z)}$$

Wavelet Phase Harmonics and phase alignment

- **Phase harmonics of wavelet coefficients**
  - Take $z = |z|\mathrm{e}^{i \cdot \arg(z)} \in \mathbb{C}$, its p^{th} phase harmonic is:

$$[z]^p = |z|\mathrm{e}^{i \cdot p \cdot \arg(z)}$$

  - $\rightarrow \rho \star \psi_{\vec{k}_1}$ contains the $\vec{k}_1$ spatial frequency
    $\rightarrow \left[\rho \star \psi_{\vec{k}_1}\right]^p$ contains the $p \cdot \vec{k}_1$ spatial frequency

Wavelet Phase Harmonics and phase alignment

- **Phase harmonics of wavelet coefficients**
  - Take $z = |z|\mathrm{e}^{i\cdot\arg(z)} \in \mathbb{C}$, its p^{th} phase harmonic is:

$$[z]^p = |z|\mathrm{e}^{i\cdot p\cdot\arg(z)}$$

  - $\to \rho \star \psi_{\vec{k}_1}$ contains the $\vec{k}_1$ spatial frequency
    $\to \left[\rho \star \psi_{\vec{k}_1}\right]^p$ contains the $p \cdot \vec{k}_1$ spatial frequency
  - $\Rightarrow$ Non-vanishing covariance terms with:

$$\mathrm{Cov}\left(\left[\rho \star \psi_{\vec{k}_1/2}\right]^2, \rho \star \psi_{\vec{k}_1}\right)$$

    $\to$ Statistics of relative phase shifts between wavelet bands

Wavelet Phase Harmonics and phase alignment

- **Constructing WPH statistics**

$$\boxed{C_{\vec{k}_1,p_1,\vec{k}_2,p_2} = \mathrm{Cov}\left(\left[\rho \star \psi_{\vec{k}_1}(\vec{x})\right]^{p_1}, \left[\rho \star \psi_{\vec{k}_2}(\vec{x})\right]^{p_2}\right)}$$

"WPH moments"

- One moment for each $(\vec{k}_1, \vec{k}_2, p_1, p_2)$ set
- Different couplings between pairs of scales
- WPH statistics built from a set of WPH moments

Wavelet Phase Harmonics and phase alignment

- **Constructing WPH statistics**

$$\boxed{C_{\vec{k}_1,p_1,\vec{k}_2,p_2} = \mathrm{Cov}\left(\left[\rho \star \psi_{\vec{k}_1}(\vec{x})\right]^{p_1}, \left[\rho \star \psi_{\vec{k}_2}(\vec{x})\right]^{p_2}\right)}$$

"WPH moments"

  - One moment for each $(\vec{k}_1, \vec{k}_2, p_1, p_2)$ set
  - Different couplings between pairs of scales
  - WPH statistics built from a set of WPH moments

- **Naturally extended to cross-statistics**

$$C_{\vec{k}_1,p_1,\vec{k}_2,p_2}(\mathbf{A}, \mathbf{B}) = \mathrm{Cov}\left(\left[\mathbf{A} \star \psi_{\vec{k}_1}(\vec{x})\right]^{p_1}, \left[\mathbf{B} \star \psi_{\vec{k}_2}(\vec{x})\right]^{p_2}\right)$$

Wavelet Phase Harmonics and phase alignment

- **Constructing WPH statistics**

$$\boxed{C_{\vec{k}_1,p_1,\vec{k}_2,p_2} = \mathrm{Cov}\left(\left[\rho \star \psi_{\vec{k}_1}(\vec{x})\right]^{p_1}, \left[\rho \star \psi_{\vec{k}_2}(\vec{x})\right]^{p_2}\right)}$$

"WPH moments"

  - One moment for each $(\vec{k}_1, \vec{k}_2, p_1, p_2)$ set
  - Different couplings between pairs of scales
  - WPH statistics built from a set of WPH moments

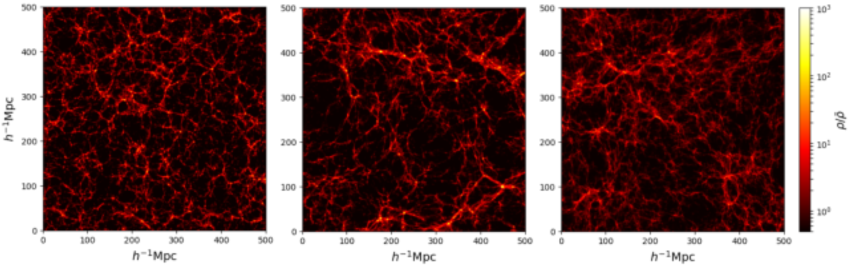
- **Naturally extended to cross-statistics**

$$C_{\vec{k}_1,p_1,\vec{k}_2,p_2}(\mathbf{A}, \mathbf{B}) = \mathrm{Cov}\left(\left[\mathbf{A} \star \psi_{\vec{k}_1}(\vec{x})\right]^{p_1}, \left[\mathbf{B} \star \psi_{\vec{k}_2}(\vec{x})\right]^{p_2}\right)$$

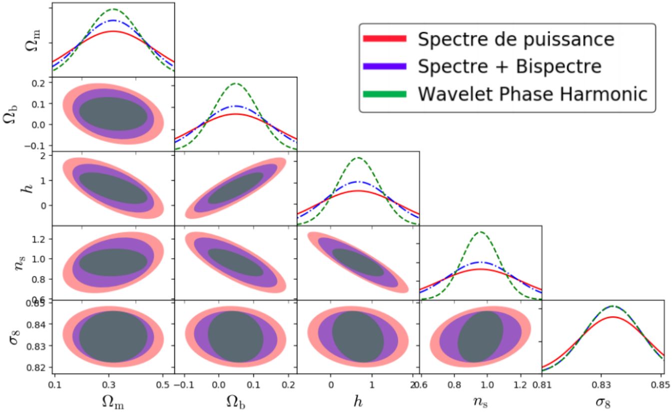
Efficient multiscale interaction language: ***(e.g., EA+, 2019)***
$\rightarrow$ **Scales selection, equivariance, symmetries, regularities...**

Application to LSS matter density field

- **Cosmological inference and weak lensing** *(EA+, 2020)*
  - Quijote 2D matter density field, integrated thin slices
  - P_k (127 coeffs), $P_k + B_k$ (313 coeffs), WPH statistics (327 coeffs)
  - 5 cosmological parameters, from Fisher estimates



- **Constraints from Fisher matrix** *(EA+, 2020)*



→ **Better that joint power spectrum and bispectrum**

Other applications to LSS

- **3D cosmological matter density field (*Quijote*)**
  - Wavelet Scattering Transform (WST) ***(Valogiannis+, 2021)***
    $\rightarrow$ comparable to Marked PS
  - Wavelets moments ***(Eickenberg, EA+, 2022)***
    $\rightarrow$ comparable to Marked PS

Other applications to LSS

- **3D cosmological matter density field (*Quijote*)**
  - Wavelet Scattering Transform (WST) ***(Valogiannis+, 2021)***
    $\rightarrow$ comparable to Marked PS
  - Wavelets moments ***(Eickenberg, EA+, 2022)***
    $\rightarrow$ comparable to Marked PS

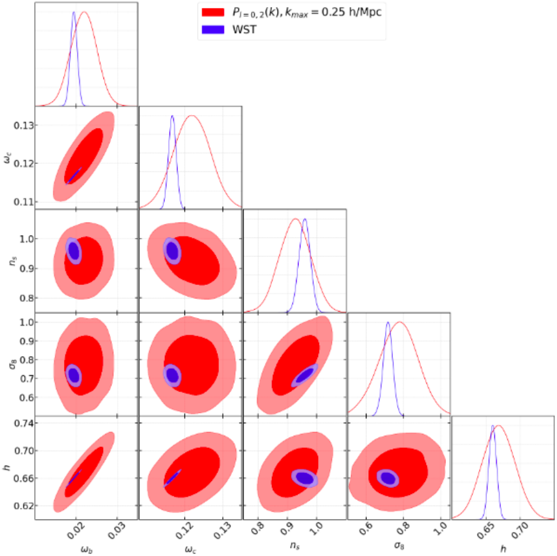
- **Actual galaxy surveys**
  - WST on BOSS data ***(Valogiannis+, 2022)***
    $\rightarrow$ substantial improvement w.r.t. PS

Other applications to LSS

- **3D cosmological matter density field (*Quijote*)**
  - Wavelet Scattering Transform (WST) *(Valogiannis+, 2021)*
    → comparable to Marked PS
  - Wavelets moments *(Eickenberg, EA+, 2022)*
    → comparable to Marked PS
- **Actual galaxy surveys**
  - WST on BOSS data *(Valogiannis+, 2022)*
    → substantial improvement w.r.t. PS
- **Examples on other fields**
  - Classification of MHD simulations *(EA+ 2019, Saydjari+, 2020)*
  - 21cm Epoch or Reionization spectroscopic data *(Greig+, 2022)*

→ **Very informative tool (sometimes on par with CNN !)**
→ **Wide range of applicability (generic, training-less)**

- **Constraints from BOSS survey** *(Valogiannis+, 2022)*



Outline

Maximum entropy generative models

- **Generative model from ST statistics** *(Bruna, Mallat, 2019)*
  - Maximum entropy model under statistical constraints
  - New realizations of a process from its ST statistics
  - Non-gaussian properties quantitatively reproduced

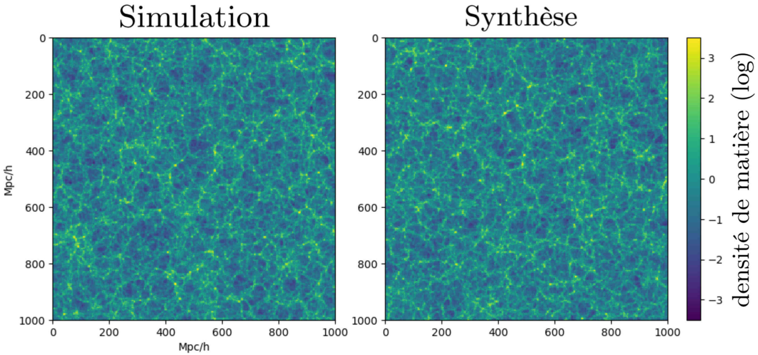
Maximum entropy generative models

- **Generative model from ST statistics** *(Bruna, Mallat, 2019)*
  - Maximum entropy model under statistical constraints
  - New realizations of a process from its ST statistics
  - Non-gaussian properties quantitatively reproduced

- **Practical implementation**
  - Constraints $\Phi(x)$ from a (set of) realization(s) x
  - Sampled with a gradient-descent algorithm
    $\rightarrow$ from a white noise
    $\rightarrow$ optimizing $\tilde{x}$ such that $\Phi(\tilde{x}) \simeq \Phi(x)$
  - 10-30 seconds on a GPU for a 256^2 map

Maximum entropy generative models

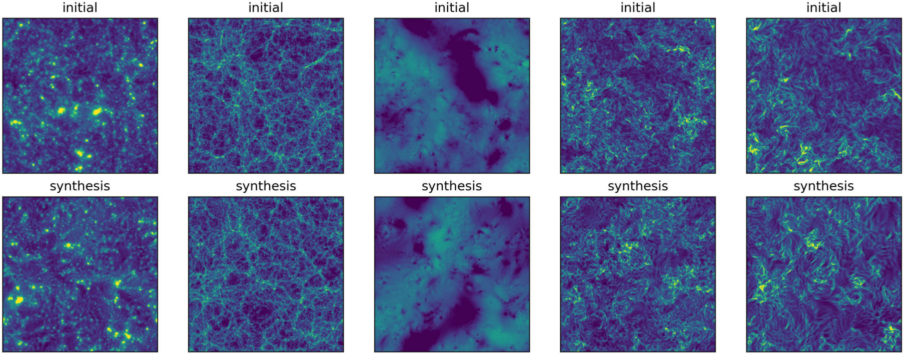
- **Quantitative validation of syntheses** *(EA+, 2020)*
  - Wavelet Phase Harmonics (WPH - 5k coefficients)
  - Matter density field of the large scale structures (*Quijote*)



→ Usual statistics very well reproduced (up to 1-10 %)

Maximum entropy generative models

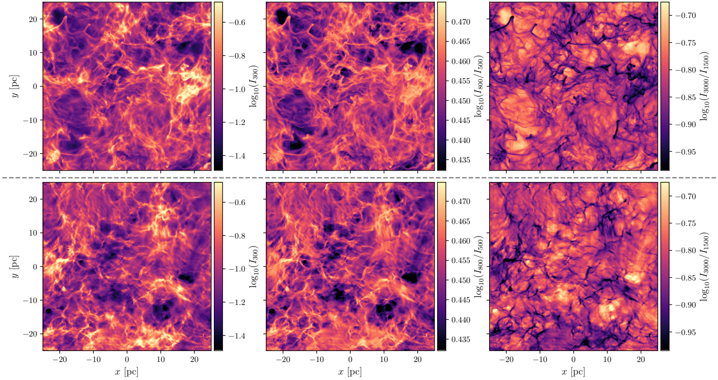
- **Syntheses from a single image** *(Cheng, EA+, in prep.)*
  - Scattering covariances + dimensionality reduction (up to few hundreds!)



→ **Realistic syntheses from a single input image!**

Maximum entropy generative models

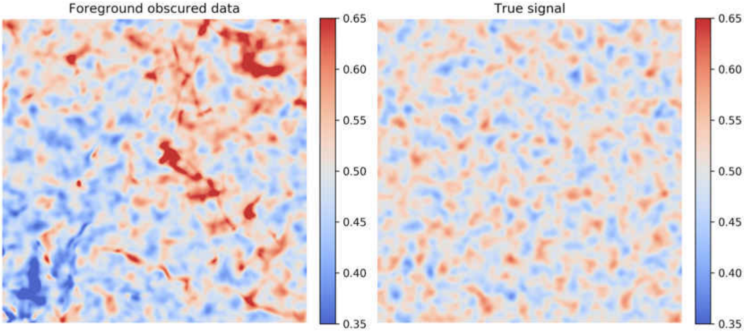
- **Multi-channel syntheses** *(Regaldo, EA+, in prep.)*
  - From cross Wavelet Phase Harmonics



$\rightarrow$ Multifrequency generative models can be constructed!

Components separation from synthesized data

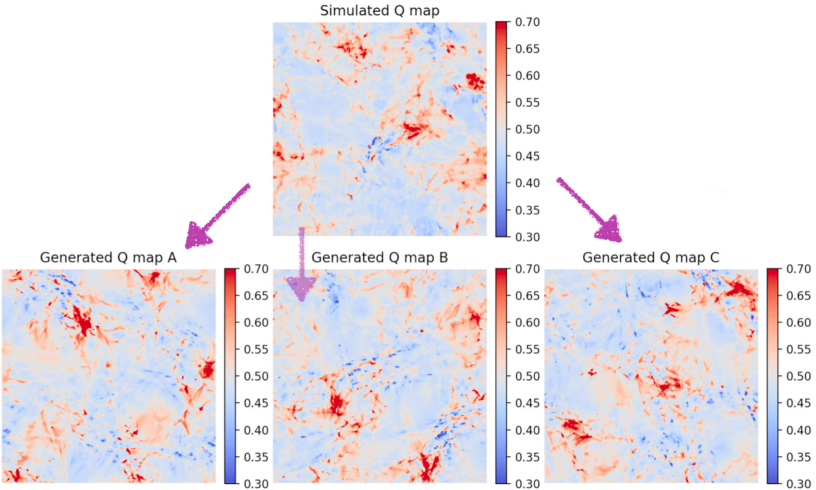
- **CMB B-mode and galactic foregrounds** *(Jeffrey, EA+, 2021)*



1. WPH generative model from one dust map
2. Training the moment network for foreground removal

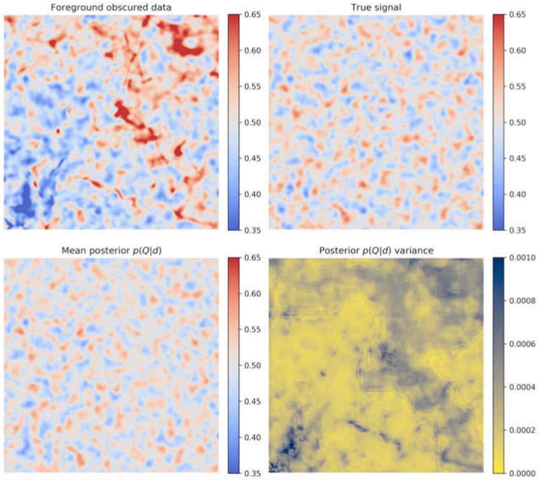
Components separation from synthesized data

1. **WPH generative model from one single map**



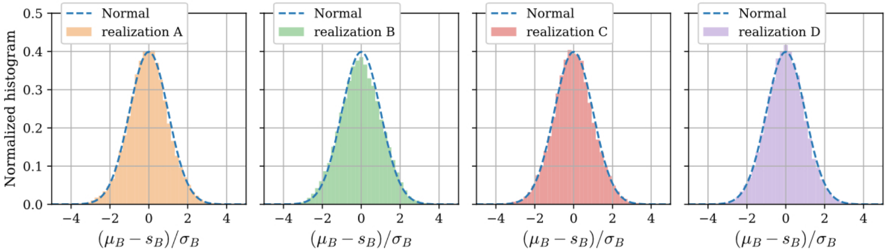
Components separation from synthesized data

2. Train moment network for foreground removal



Components separation from synthesized data

2. Train moment network for foreground removal *(Jeffrey, EA+, 2021)*
→ Validation for all pixels: (mean posterior - truth)/(std posterior)



→ **Foreground removal from one single dust image**
→ **Validates the ST generative model**

Component separation from ST statistics

- **Difficulty with astrophysical observations**
  - Process x of interest rarely isolated
  - Mixed with other components y

  $\rightarrow$ How to proceed without specific prior on x ?

Component separation from ST statistics

- **Difficulty with astrophysical observations**
  - Process x of interest rarely isolated
  - Mixed with other components y

  $\rightarrow$ How to proceed without specific prior on x ?

- **General case study**
  - We observe a mixture $d = x + y$
  - Use knowledge on y to recover x

    $\rightarrow$ Statistically : ST or other statistics
      * allows for generative model

    $\rightarrow$ Deterministically : More difficult *a priori*
      * other methods available if ST known

  $\Rightarrow$ Synthesis from indirect constraints

Synthesis from indirect constraints

- **Direct constraint case : x available**
  - $\Phi(x)$ known
  - Generate $\tilde{x}$ such that

$$\Phi(\tilde{x}) \simeq \Phi(x)$$

Synthesis from indirect constraints

- **Direct constraint case : x available**
  - $\Phi(x)$ known
  - Generate $\tilde{x}$ such that

$$\Phi(\tilde{x}) \simeq \Phi(x)$$

- **Indirect constraint case : $d = x + y$ available**
  - $\Phi(d)$ and $\Phi(y)$ known
  - Generate $\tilde{x}$ such that (for instance)

$$\left\langle \Phi(\tilde{x} + y_i) \right\rangle_i \simeq \Phi(d) \quad \text{and} \quad \Phi(d - \tilde{x}) = \Phi(y)$$

$\rightarrow$ Do we recover a map $\tilde{x}$ with correct statistics ?

Polarized dust/noise component separation

- **Separation of dust polarized microwave emission and noise**
  - Allows for an accurate model of polarized dust emission
  - Important for separation CMB/foreground

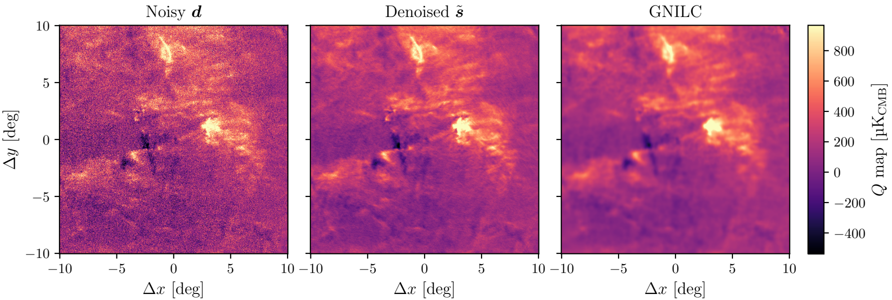
Polarized dust/noise component separation

- **Separation of dust polarized microwave emission and noise**
  - Allows for an accurate model of polarized dust emission
  - Important for separation CMB/foreground
- **First conceptual application** **(Regaldo, EA+ 21)**
  - $d = s + n$
  - 300 realistic noises including scanning strategy
  - One single indirect constraint

$$\left\langle \Phi(u + n_i) \right\rangle_i \simeq \Phi(d)$$

Separation of polarized dust emission and noise

- **Results on observational data** (Régaldo, EA+ 21)



→ Transition btw. deterministic and statistical
→ Conceptual validation of the method
→ Clear room for improvement

Polarized dust/noise component separation

- **With cross-statistics on the sphere** **(Delouis, EA+, in prep.)**
  - Cross statistics between maps $\Phi(x_1, x_2)$
  - Set of data available:
    $\rightarrow$ $d = s + n$ full observation
    $\rightarrow$ $d_i = s + n_i$ half missions
    $\rightarrow$ $\{\tilde{n}, \tilde{n}_1, \tilde{n}_2\}$ 300 noise realisations

Polarized dust/noise component separation

- **With cross-statistics on the sphere** **(Delouis, EA+, in prep.)**
  - Cross statistics between maps $\Phi(x_1, x_2)$
  - Set of data available:
    $\rightarrow$ $d = s + n$ full observation
    $\rightarrow$ $d_i = s + n_i$ half missions
    $\rightarrow$ $\{\tilde{n}, \tilde{n}_1, \tilde{n}_2\}$ 300 noise realisations

- **Set of constraints**

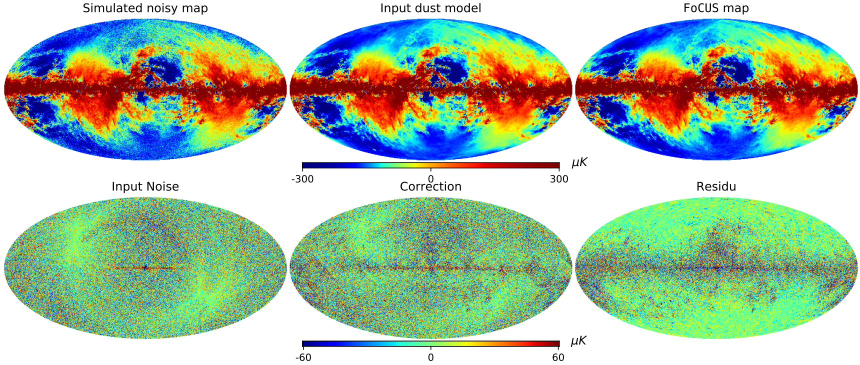
$$\Phi(d_1, d_2) \simeq \left\langle \Phi(u + \tilde{n}_1, u + \tilde{n}_2) \right\rangle_{\tilde{n}}$$

$$\Phi(d, u) \simeq \left\langle \Phi(u + \tilde{n}, u) \right\rangle_{\tilde{n}}$$

$\rightarrow$ with 5 different masks to deal with inhomogeneity

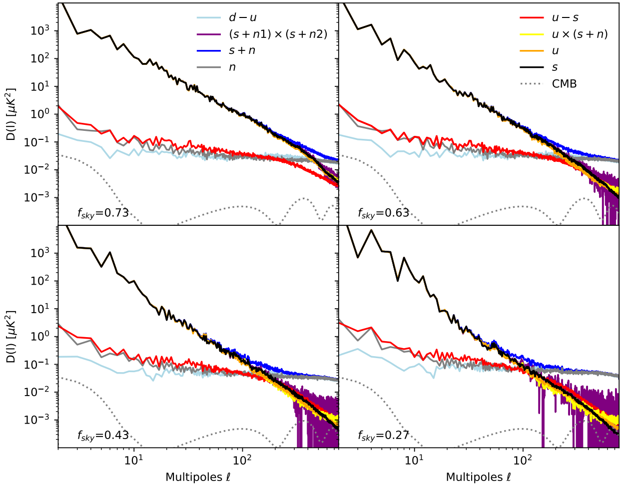
Polarized dust/noise component separation

- **Results on a test map**



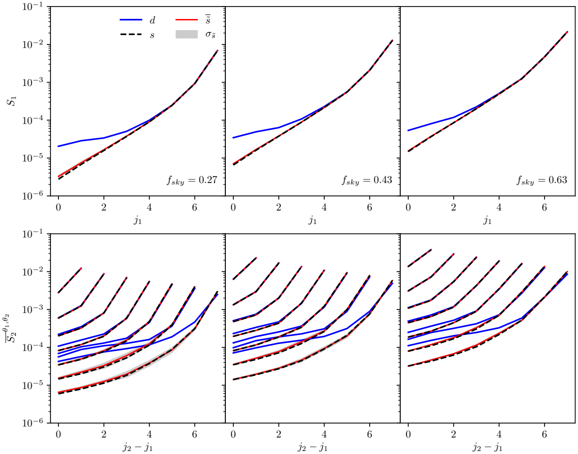
Polarized dust/noise component separation

- **Transition between deterministic and statistics**



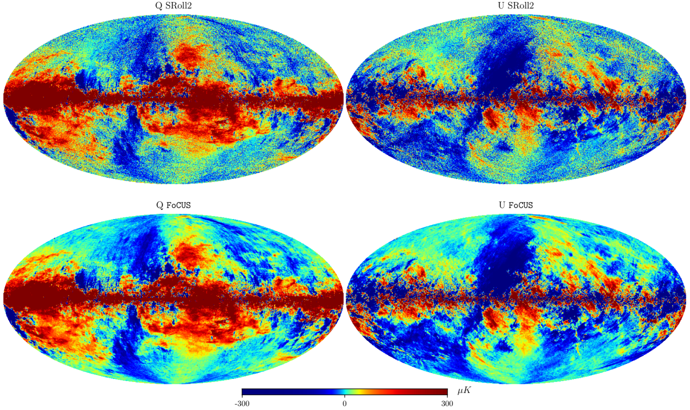
Polarized dust/noise component separation

- **Non-Gaussian features well recovered**



Polarized dust/noise component separation

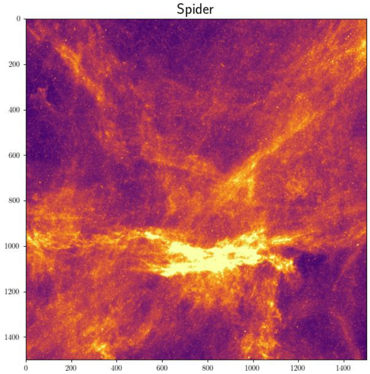
- Results on *Planck* observational data (Delouis, EA+, in prep)



→ **Cross-ST simple language for physical constraints**
→ **ST statistics and cross-statistics well recovered**

Dust in intensity/CIB separation

- *Herschel* infrared noisy dust observation with CIB



→ No realistic noise or CIB models are available
→ Can we work only from obs. data?

Dust in intensity/CIB separation

- **Recovering the components** **(Auclair, EA+, in prep.)**
  - $d = s + \mathrm{cib} + n$
  - CIB from *Lockman Hole* field
  - Noise from half-missions observations

  → Generative model for the contamination!

Dust in intensity/CIB separation

- **Recovering the components (Auclair, EA+, in prep.)**
  - $d = s + \mathrm{cib} + n$
  - CIB from *Lockman Hole* field
  - Noise from half-missions observations

  → Generative model for the contamination!

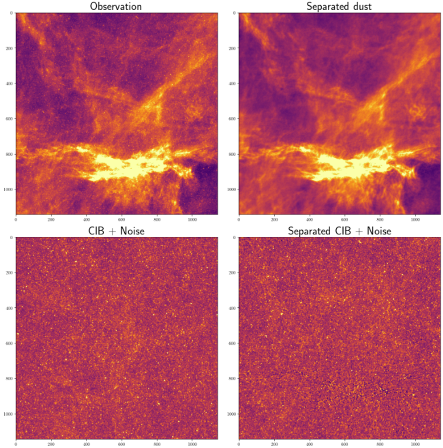
- **Set of constraints**

$$\left\langle \Phi(u + \tilde{n} + \widetilde{\mathrm{cib}}) \right\rangle_{\tilde{n}, \widetilde{\mathrm{cib}}} \simeq \Phi(d)$$

$$\Phi(d - u) \simeq \Phi(n + \mathrm{cib})$$

Dust in intensity/CIB separation

- Application on the *spider* field (Auclair, EA+, in prep)



→ Components separation solely from obs. data!

Conclusion

- **Scattering Transform statistics**
  - Very generic and efficient
  - Allows for new applications
    $\rightarrow$ Realistic syntheses from a single image
    $\rightarrow$ Components separation from limited data
  - Check our software:
    $\rightarrow$ PyWST
    $\rightarrow$ PyWPH

Conclusion

- **Scattering Transform statistics**
  - Very generic and efficient
  - Allows for new applications
    $\rightarrow$ Realistic syntheses from a single image
    $\rightarrow$ Components separation from limited data
  - Check our software:
    $\rightarrow$ PyWST
    $\rightarrow$ PyWPH

Thanks for your attention!