cea

INSTITUT DE RECHERCHE sur les LOIS FONDAMENTALES de l'UNIVERS

SEDI

SAp Service d'Astrophysique

Missing data interpolation in Weak Lensing

Sandrine Pires

IRFU/Sedi-Sap/CEA Saclay, France

sandrine.pires@cea.fr

http://irfu.cea.fr/Pisp/4/sandrine.pires.html

Collaborators

- J-L. Starck, CEA-Saclay, France
- A. Réfrégier, CEA-Saclay, France
- A. Amara, University of Zurich, Switzerland
- R. Teyssier, ETH Zurich, Switzerland
- J. Fadili, Caen University, France

Outline

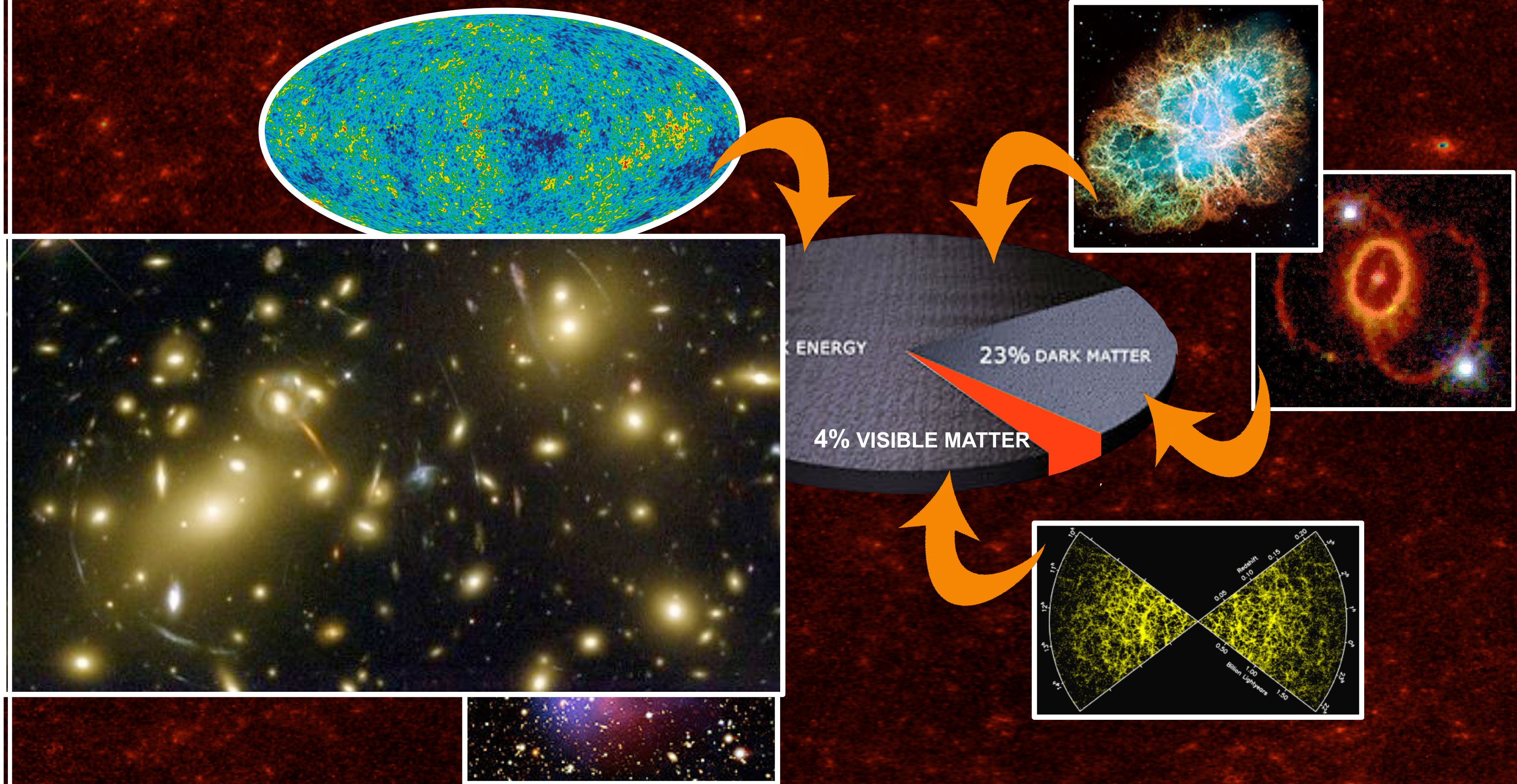
1- Introduction
- Introduction to Weak Lensing
- The Weak lensing data processing line

2- The missing data problem in Weak Lensing
- Weak Lensing analysis with missing data
- A new approach based on sparsity

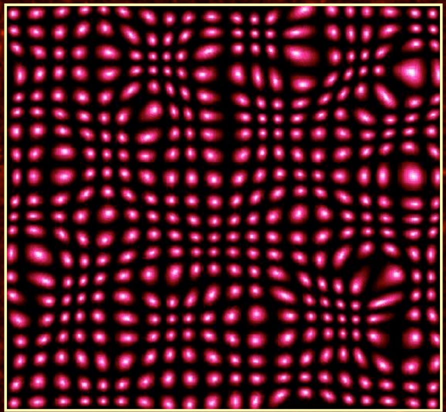
3- Application of Inpaiting to Weak Lensing data
- Results on Weak Lensing data
- The FASTLens software

From observations to cosmological model

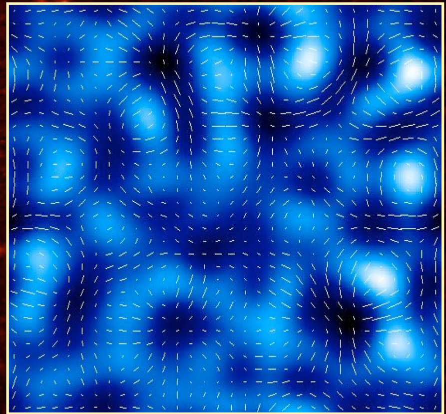
$$\mathcal{M}(\Omega_M, \Omega_\Lambda, \Omega_b, \sigma_8, \ldots)$$



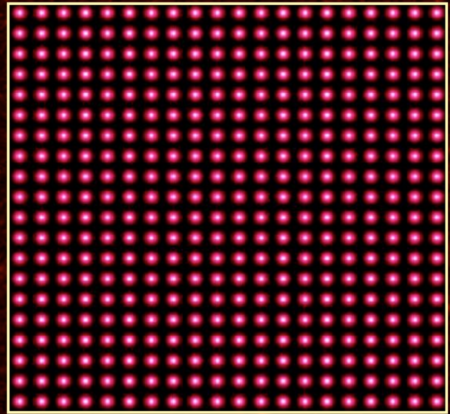
Weak Gravitational Lensing



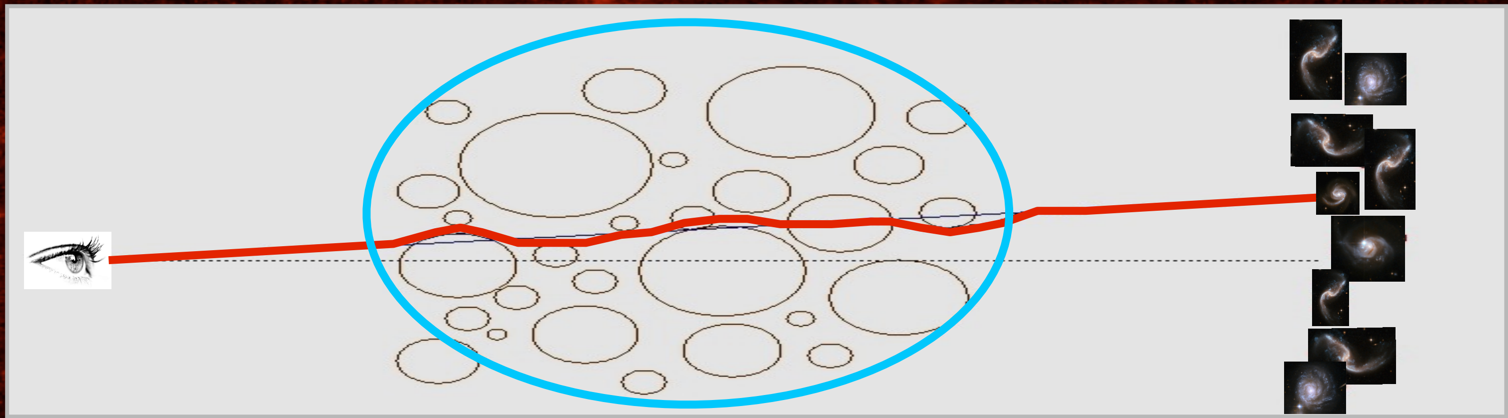
OBSERVER



GRAVITATIONAL LENS



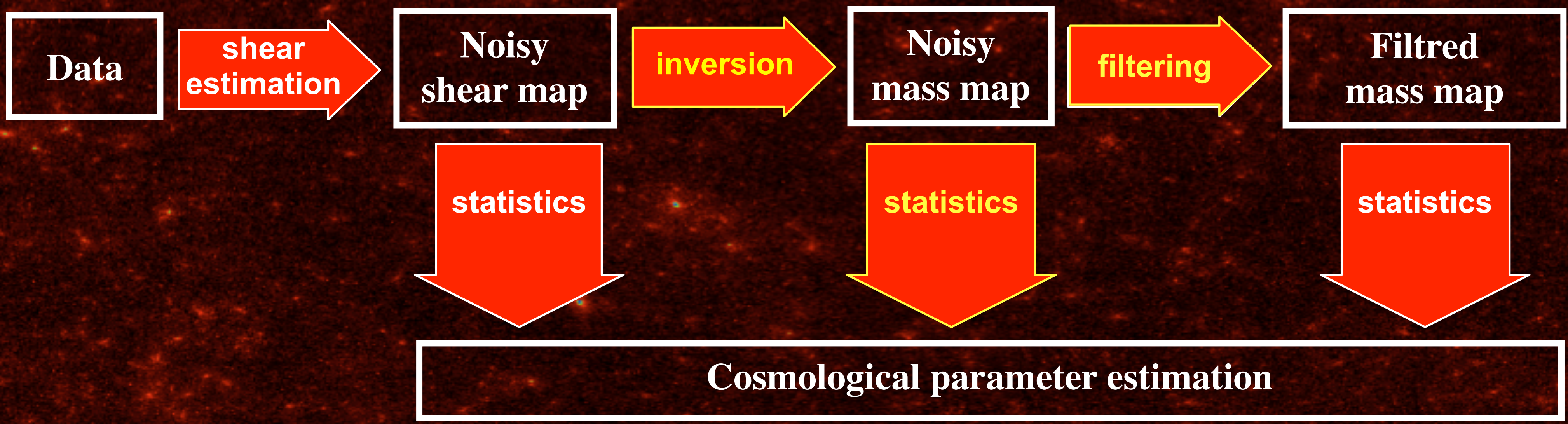
BACKGROUND GALAXIES

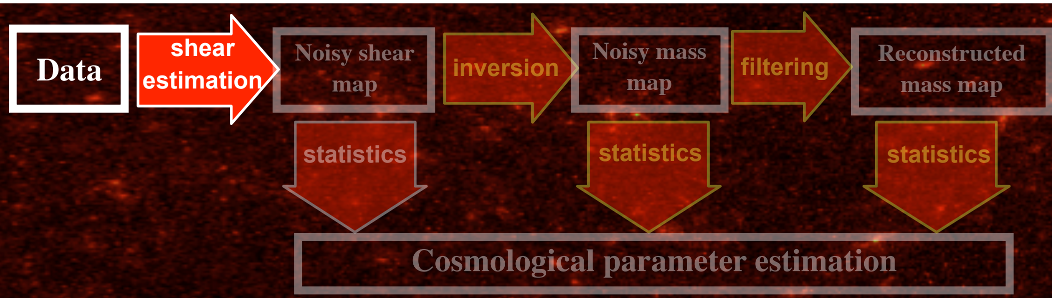


GRAVITATIONAL LENS

= GOAL =

Constrain cosmological parameters from weak lensing data

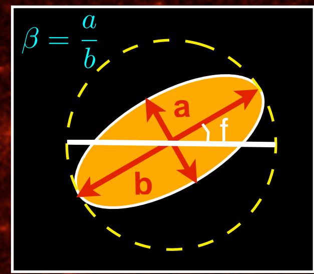




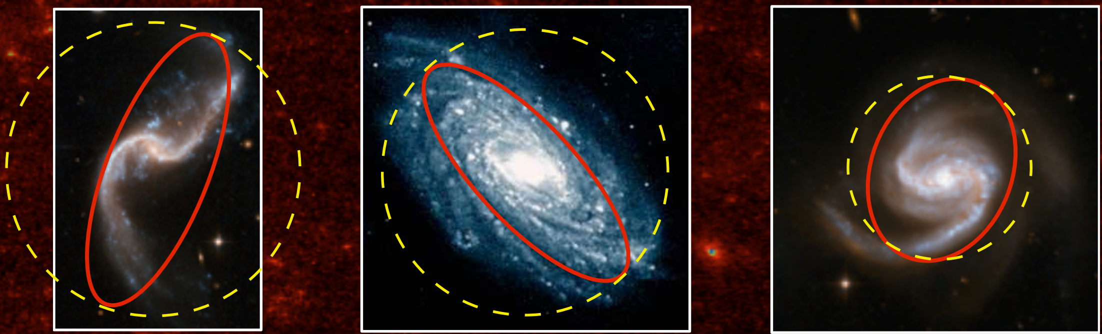
Shear estimation

✓ Shear estimation on each galaxy of the field

⇨ Ellipticity must be measured : $\begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix} = \frac{1-\beta}{1+\beta} \begin{pmatrix} \cos 2\phi \\ \sin 2\phi \end{pmatrix}$



✓ Galaxies have an intrinsic ellipticipty

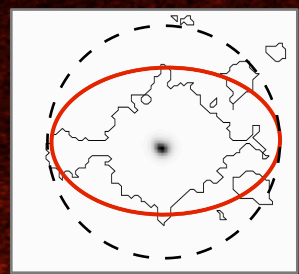


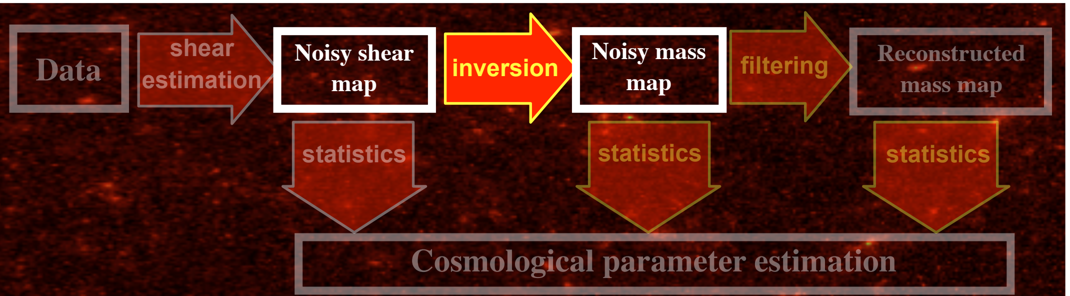
⇨ ellipticity must be averaged over several nearby galaxies :

$$< \epsilon_i > \approx \gamma_i$$

✓ Galaxies are convolved by an asymetric PSF

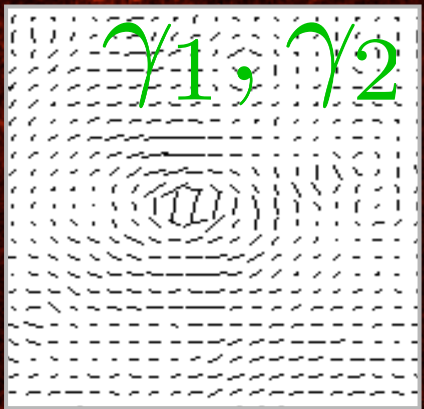
⇨ PSF have to be estimated and deconvolved





Inversion equations

SIMULATED SHEAR MAP



$$\gamma_1 = \frac{1}{2}\left(\partial_1^2 - \partial_2^2\right)\psi$$

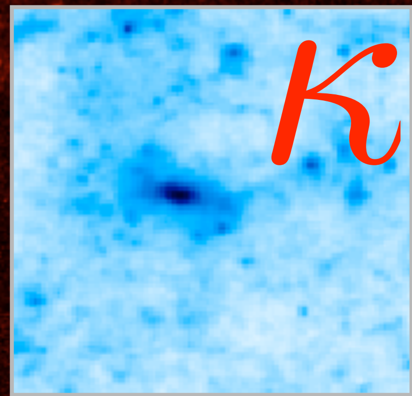
$$\gamma_2 = \partial_1 \partial_2 \psi$$

LENSING POTENTIAL

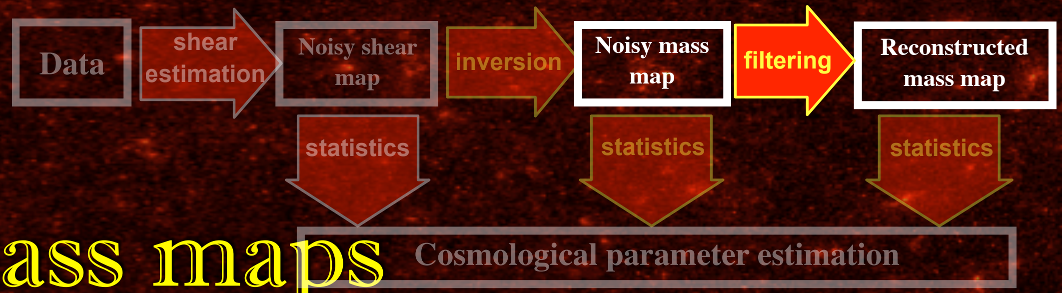
$$\psi$$

$$\frac{1}{2}\left(\partial_1^2 + \partial_2^2\right)\psi = \kappa$$

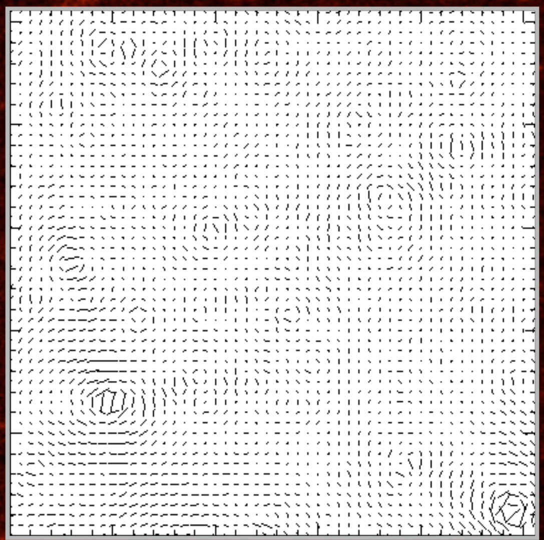
SIMULATED MASS MAP
(Vale & White, 2003)



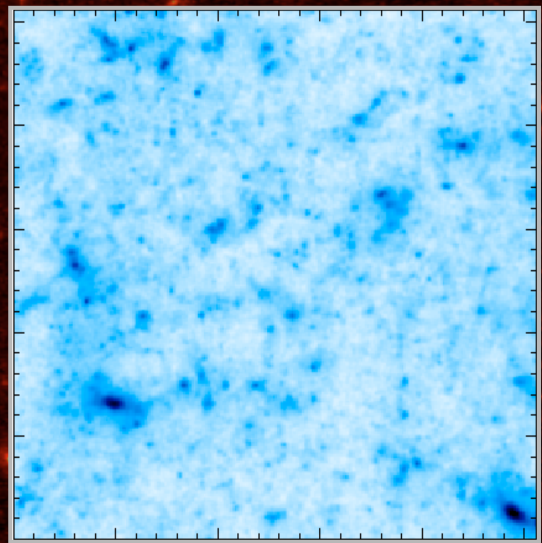
$$\gamma_1, \gamma_2 \rightleftharpoons \kappa$$



Noisy mass maps

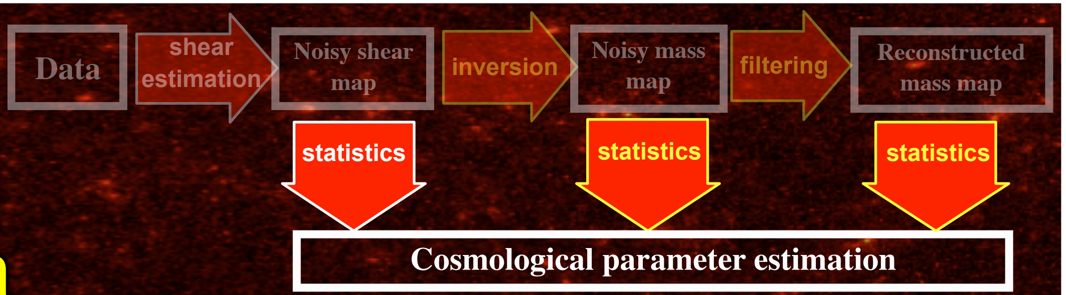


SHEAR MAP

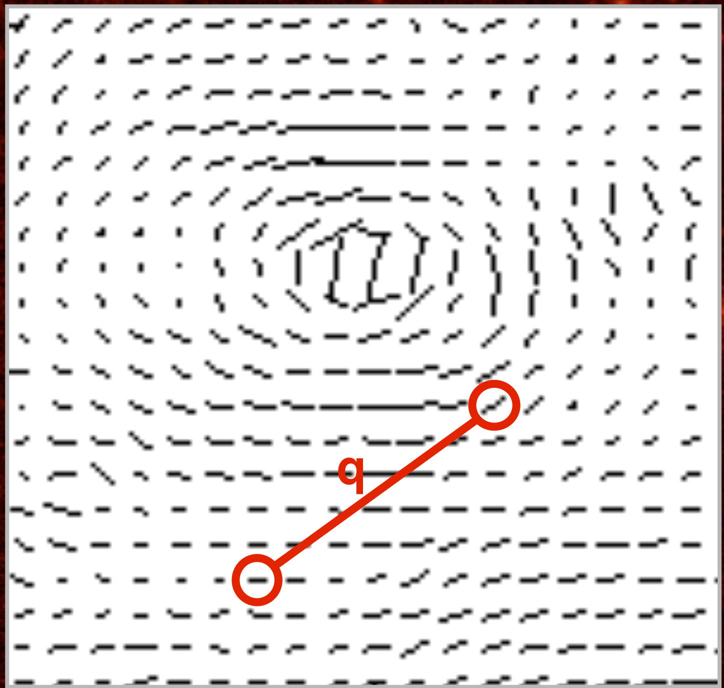


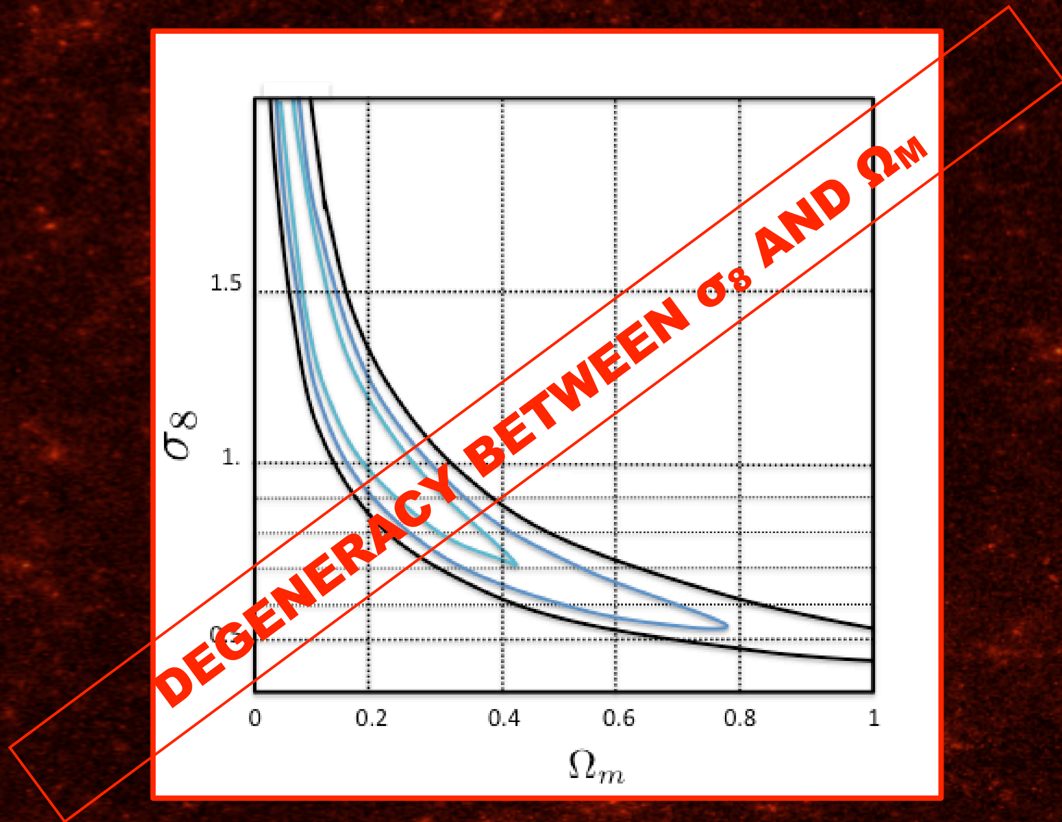
ORIGINAL MASS MAP
(Vale & White, 2003)

MASS MAP
(SPACE OBSERVATIONS)

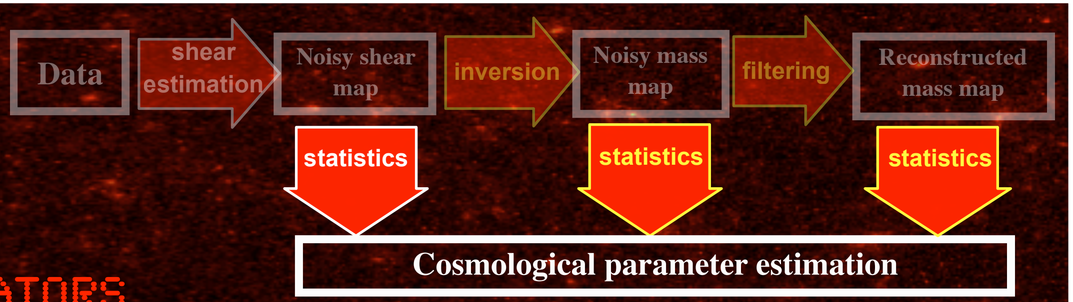


Statistic estimation



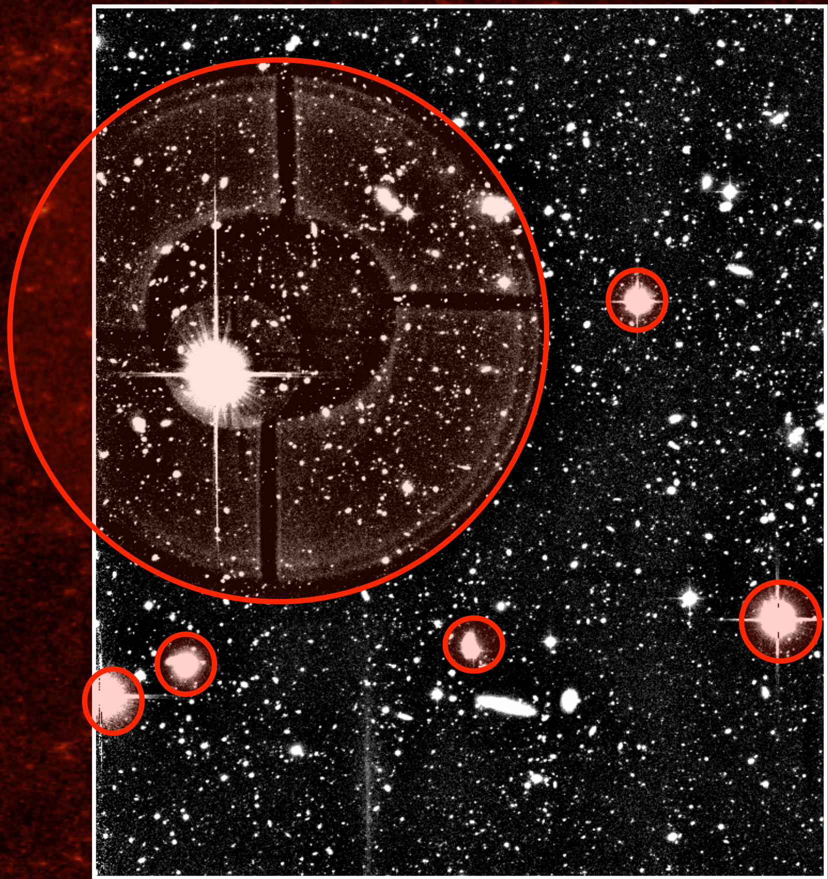


Statistic estimation



GAUSSIAN ESTIMATORS	NON-GAUSSIAN ESTIMATORS	
TWO-POINT STATISTICS	THREE-POINT STATISTICS	FOUR-POINT STATISTICS
$\sigma^2 = \sum_1^N (\kappa_i - \bar{\kappa})^2$ VARIANCE	$S = \frac{\sum_1^N (\kappa_i - \bar{\kappa})^3}{N\sigma^3}$ SKEWNESS	$K = \frac{\sum_1^N (\kappa_i - \bar{\kappa})^4}{N\sigma^4} - 3$ KURTOSIS
$\xi_{i,j} =< \kappa(\theta_i)\kappa(\theta_j) >$ TWO-POINT CORRELATION FUNCTION	$\xi_{i,j,k} =< \kappa(\theta_i)\kappa(\theta_j)\kappa(\theta_k) >$ THREE-POINT CORRELATION FUNCTION	$\xi_{i,j,k,l} =< \kappa(\theta_i)\kappa(\theta_j)\kappa(\theta_k)\kappa(\theta_l) >$ FOUR-POINT CORRELATION FUNCTION
$P_\kappa =< \hat{\kappa}(\theta_i)\hat{\kappa}(\theta_j) >$ POWER SPECTRUM	$B_\kappa =< \hat{\kappa}(\theta_i)\hat{\kappa}(\theta_j)\hat{\kappa}(\theta_k) >$ BISPECTRUM	$T_\kappa =< \hat{\kappa}(\theta_i)\hat{\kappa}(\theta_j)\hat{\kappa}(\theta_k)\hat{\kappa}(\theta_l) >$ TRISPECTRUM

Weak lensing missing data



Outline

1- Introduction
- Introduction to Weak Lensing
- The Weak lensing data processing line

2- The missing data problem in Weak Lensing
- Weak Lensing analysis with missing data
- A new approach based on sparsity

3- Application of Inpaiting to Weak Lensing data
- Results on Weak Lensing data
- The FASTLens software

Inversion methods

- ✓ Global inversion
  - ✓ pros: less noisy
  - ✓ cons: boundary effects

$$\kappa = P_1 * \gamma_1 + P_2 * \gamma_2$$

$$\hat{P}_1(k) = \frac{k_1^2 - k_2^2}{k^2}$$

$$\hat{P}_2(k) = \frac{2k_1 k_2}{k^2}$$

- ✓ Local inversions
  - ✓ pros: unbiased by missing data

$$\nabla[log(1-\kappa)] \equiv \frac{-1}{1-|g|^2} \begin{pmatrix} 1-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1+\gamma_1 \end{pmatrix} \begin{pmatrix} \partial_1\gamma_1 + \partial_2\gamma_2 \\ \partial_1\gamma_2 - \partial_2\gamma_1 \end{pmatrix}$$

Statistical estimation with missing data

- Estimation of N-point correlation functions in direct space by avoiding the points falling in gaps
  - pros: unbiased by missing data
  - cons: time comsuming
- Estimation of the power spectrum in Fourier space by applying a mask correction
  - pros: fast estimation with FFT
  - cons:
    - stability depending on the shape of the mask
    - estimation of the mask correction can be long

Weak lensing inpainting formalism

$$\gamma_i^{obs} \rightarrow \min_{\kappa} \| \Phi^t \kappa \|_{l_0} \text{ subject to } \sum_i \| \gamma_i^{obs} - M(P_i * \kappa) \|_{l_2}^2 \leq \varepsilon \rightarrow \kappa$$

↑ Physical priors

$$\gamma_i^{obs} = M.\gamma_i$$

$$\kappa = P_1 * \gamma_1 + P_2 * \gamma_2$$

Φ^t is the DCT

Weak lensing inpainting algorithm

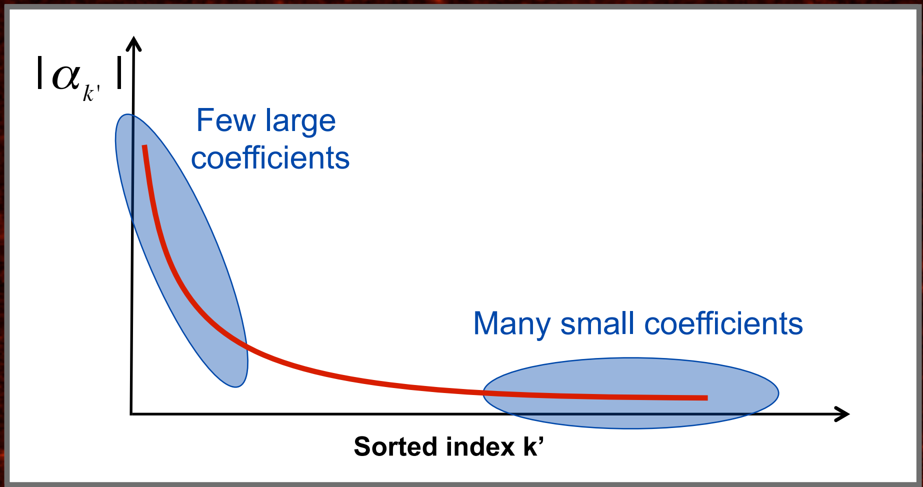
- ✓ If $\Phi^t \kappa$ sparse enough, the l_0-norm can be replaced by the convex l_1-norm
- ✓ The solution of the previous minimisation can be obtained by an iterative thresholding (MCA by Elad, 2005)

$$X^{n+1} = \Delta_{\Phi,\lambda_n}(X^n + M(Y - X^n))$$

- ✓ Décompose the signal on the dictionary Φ
- ✓ Threshold the coefficients α with a threshold λ_n
- ✓ Reconstruct from $\tilde{\alpha}$

What is sparsity ?

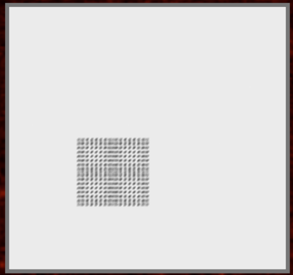
A signal S is sparse in a basis Φ if most of the coefficients α are equal to zero or closed to zero: $\min_S ||\phi^t S||_0^2$



Looking for Adapted representations

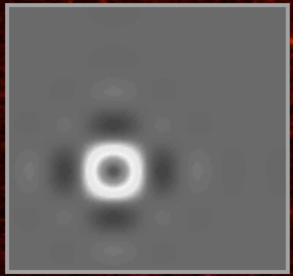
✓ Local DCT:

- Stationary textures
- Locally oscillatory



✓ Wavelet transform:

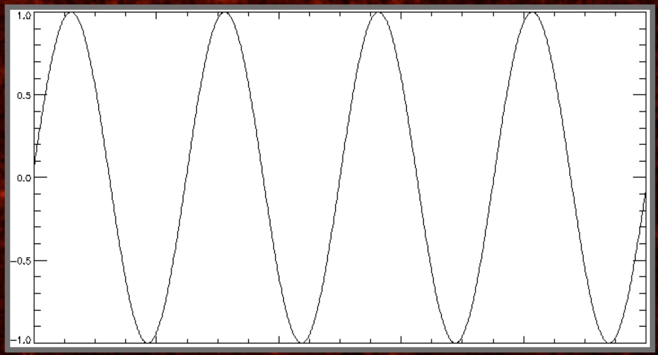
- Piecewise smooth
- Isotropic structures



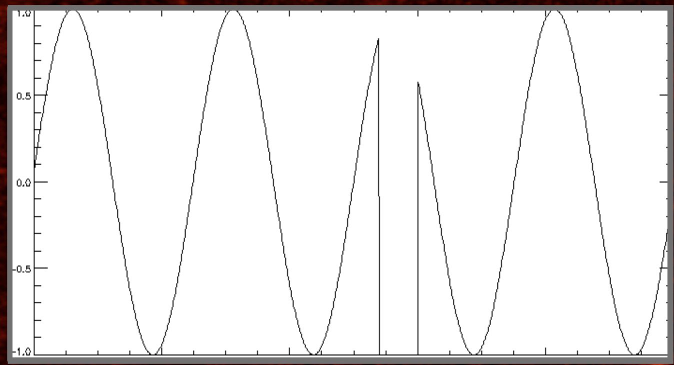
✓ Curvelet transform:

- Piecewise smooth
- Edge structures

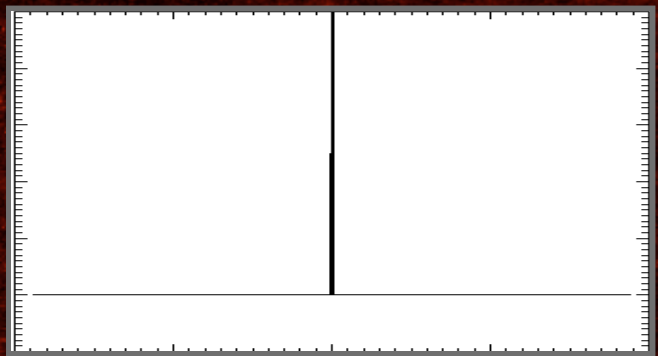
Inpainting based on sparse representation of data



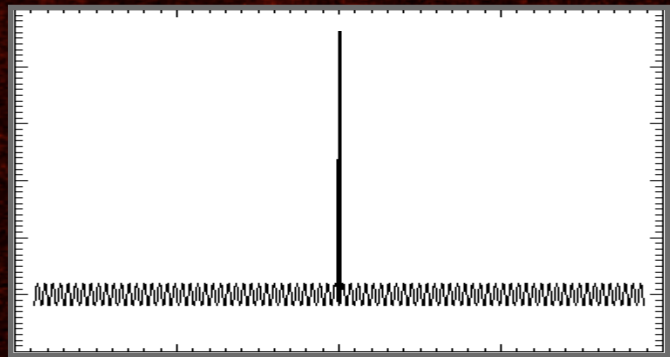
SINE CURVE



TRUNCATED SINE CURVE

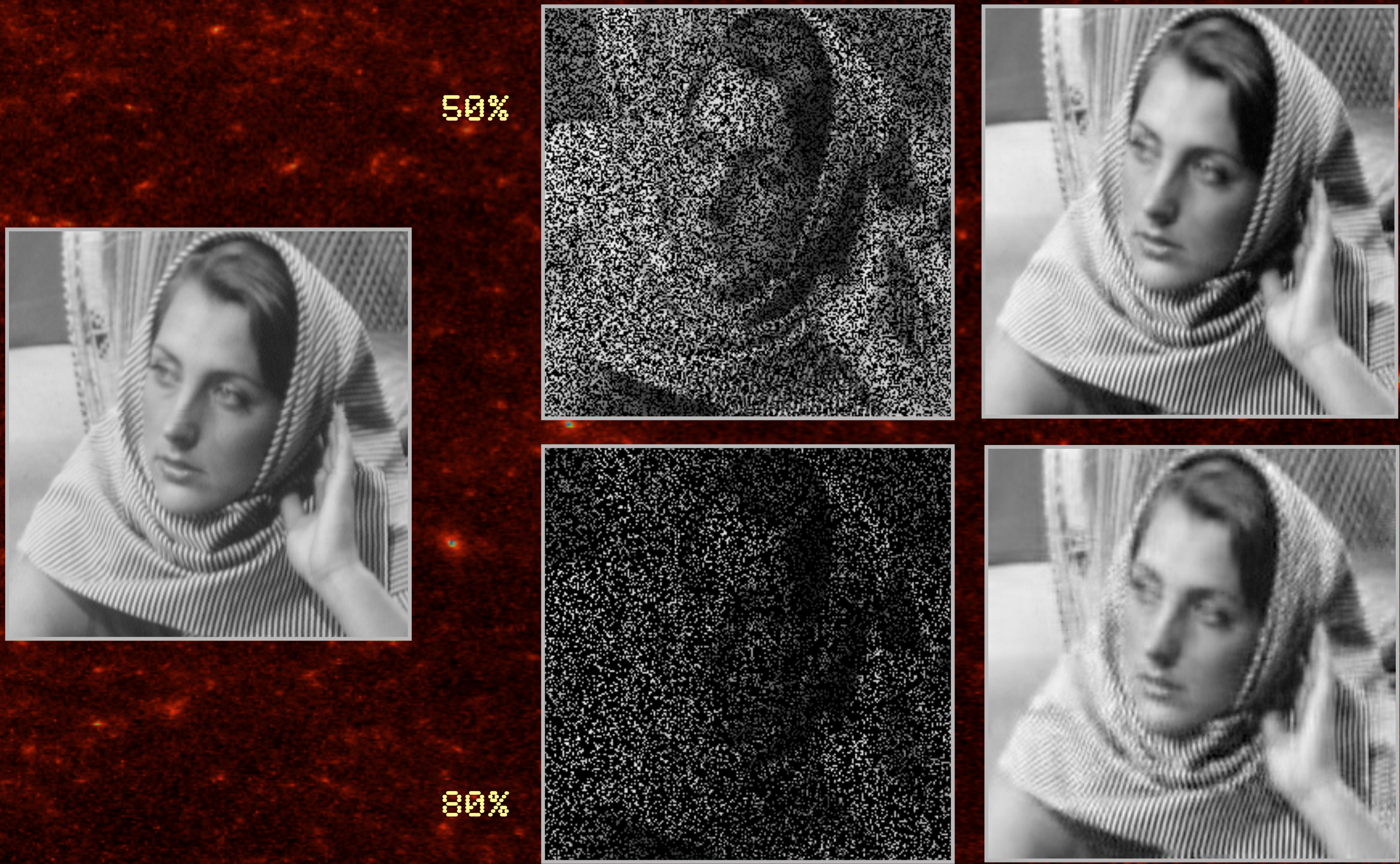


TF OF A SINE CURVE

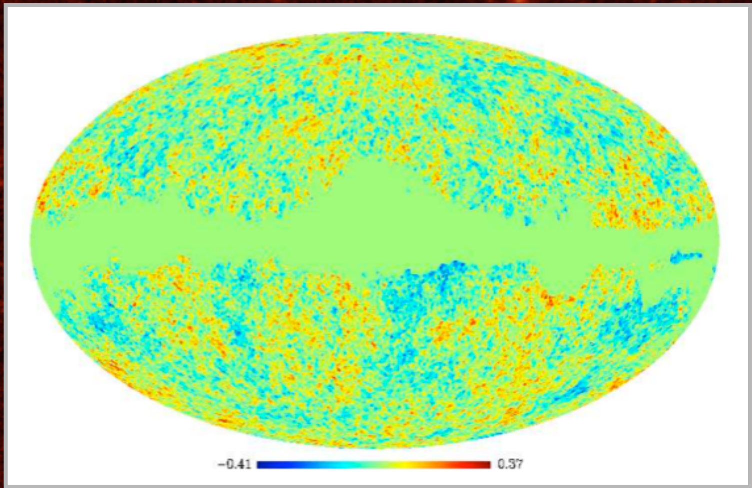


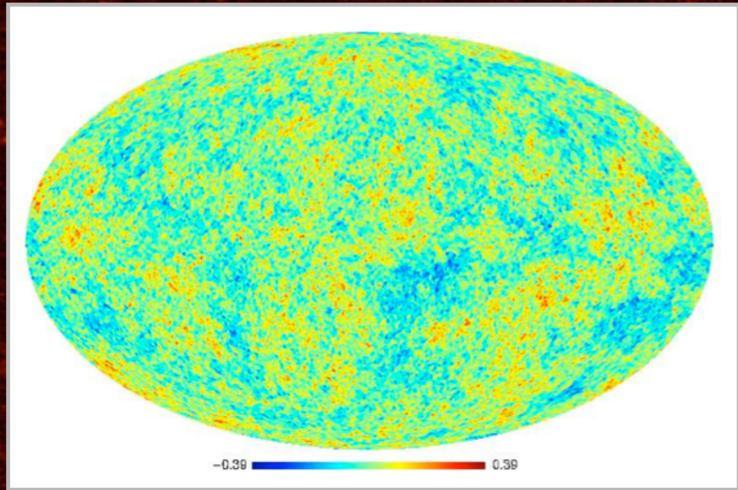
TF OF A TRUNCATED SINE CURVE

Inpainting
randomly distributed missing data



Inpainting on WMAP data





WMAP 3 YEARS

INPAINTED MAP
(COURTESY P. ABRIAL)

Outline

1- Introduction

- Introduction to Weak Lensing
- The Weak lensing data processing line

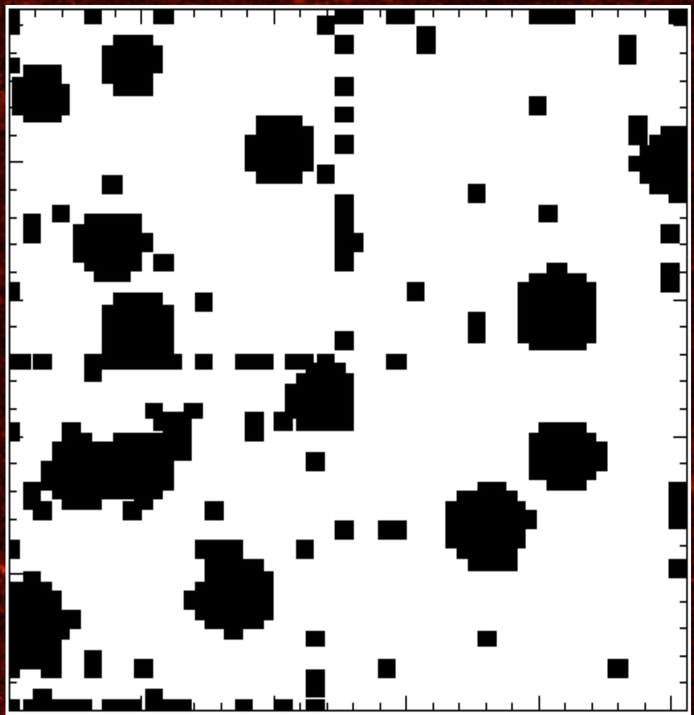
2- The missing data problem in Weak Lensing

- Weak Lensing analysis with missing data
- A new approach based on sparsity

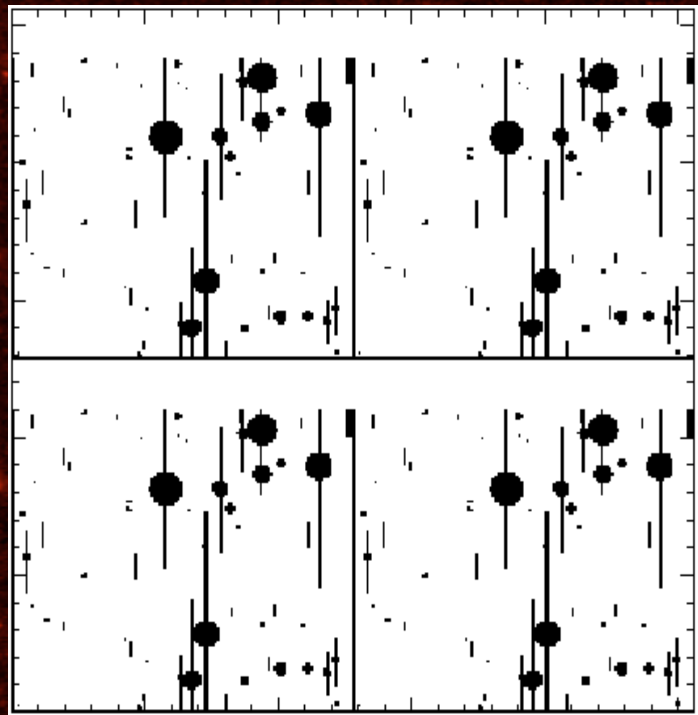
3- Application of Inpaiting to Weak Lensing data

- Results on Weak Lensing data
- The FASTLens software

Masked masks

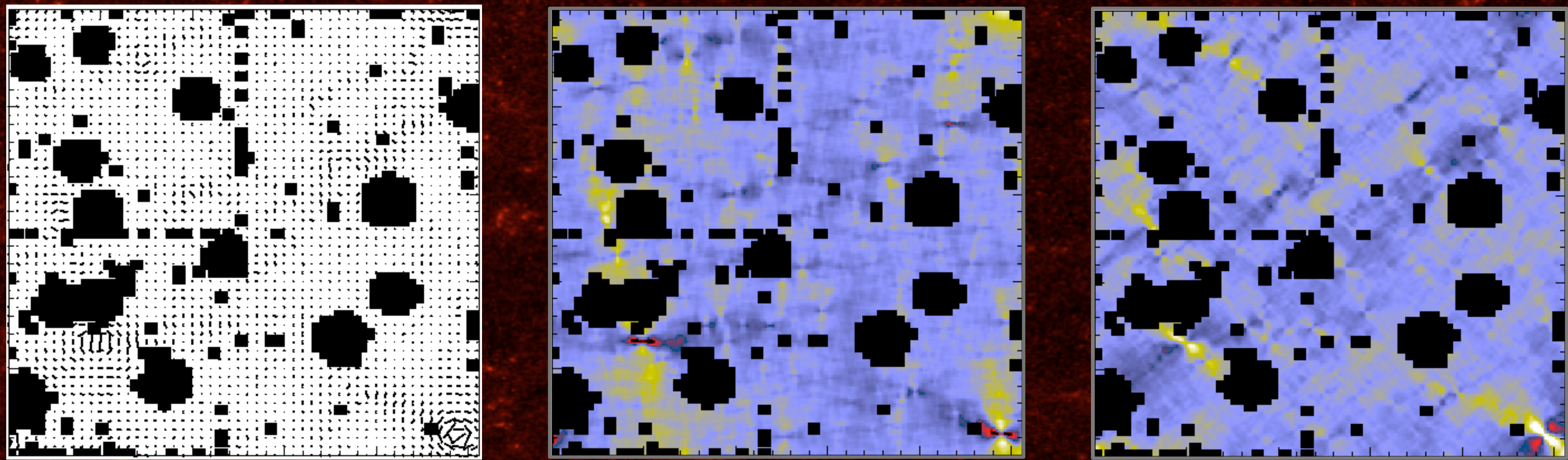


MASK PATTERN OF CFHTLS SURVEY
ON 1° X 1° FIELD (COURTESY J. BERGE)



MASK PATTERN OF SUBARU
SURVEY ON 1° X 1° FIELD

Inpainting from shear maps

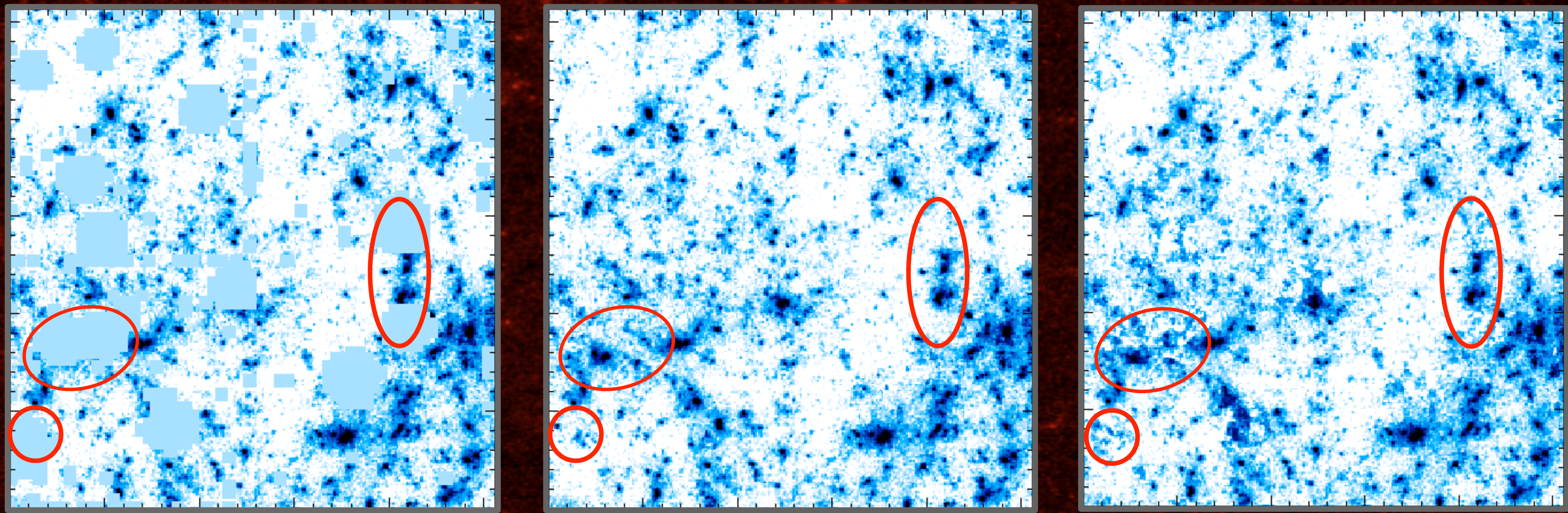


γ_1^{obs} γ_2^{obs}

$$\gamma_i^{obs} \rightarrow \min_\kappa \| \Phi^t \kappa \|_{l_0} \text{ subject to } \sum_i \| \gamma_i^{obs} - M(P_i * \kappa) \|_{l_2}^2 \leq \varepsilon \rightarrow \kappa$$

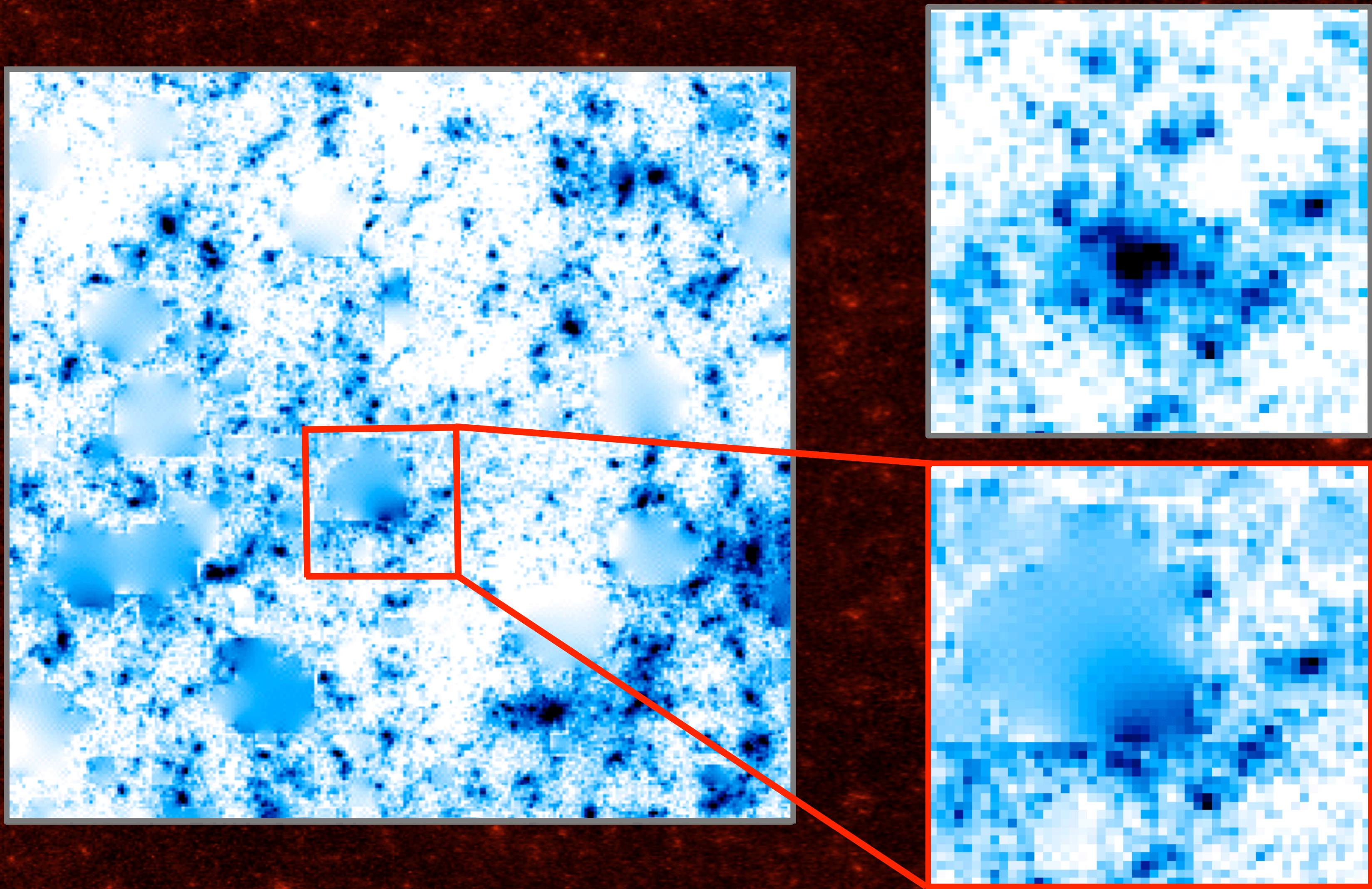
Physical priors

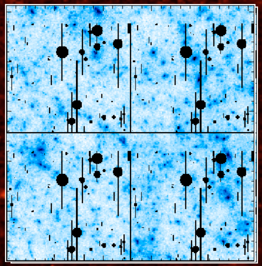
Inpainting on simulated weak lensing data



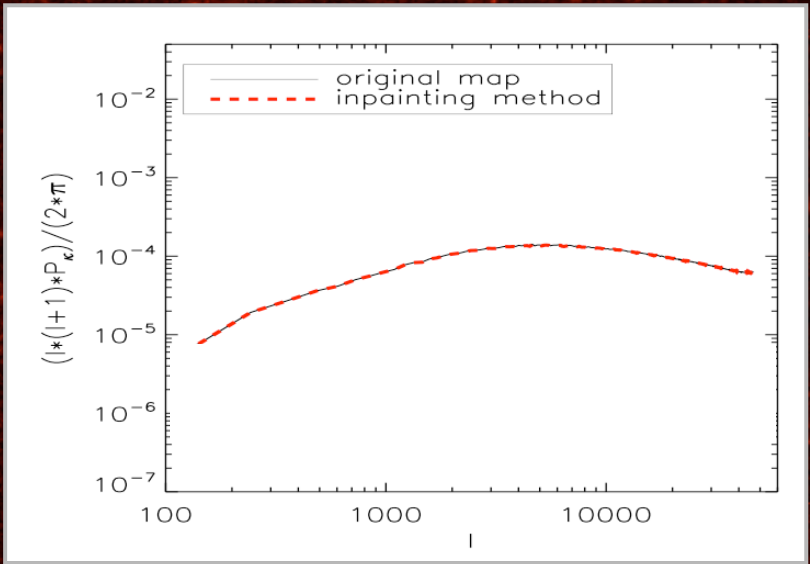
WHICH IMAGE IS THE ORIGINAL ONE ?

Inpainting on simulated weak lensing data



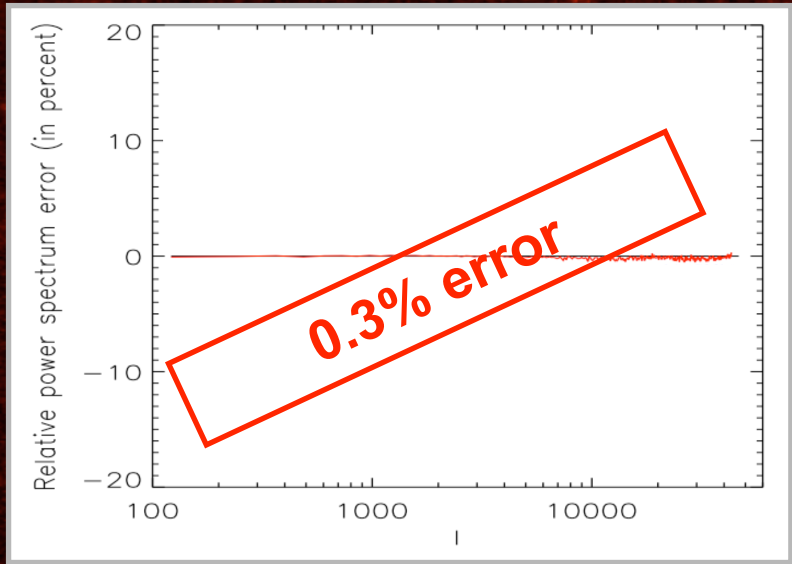


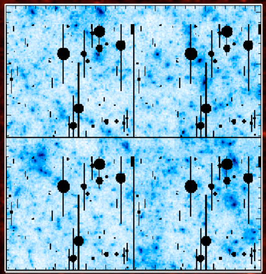
Power spectrum estimation



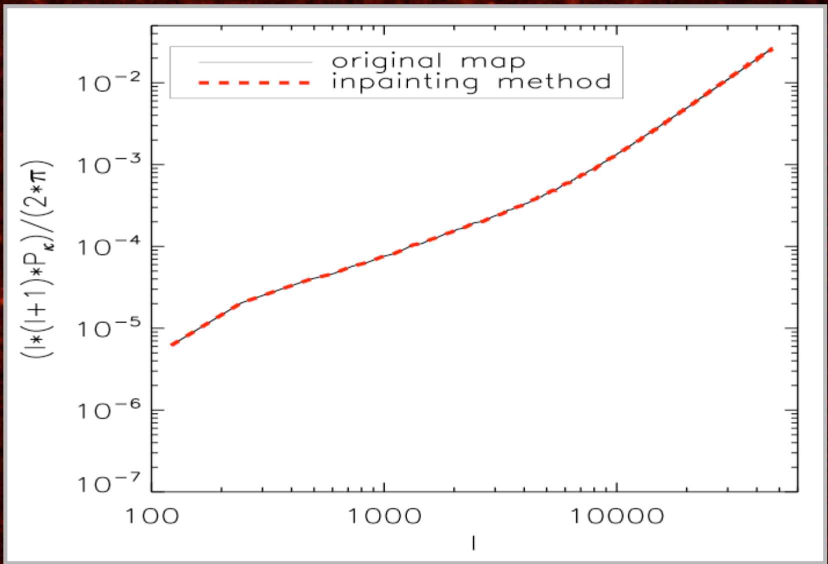
MEAN POWER SPECTRUM COMPUTED FROM
- 100 COMPLETE MASS MAPS (BLACK)
- 100 INPAINTED MAPS (RED).

RELATIVE POWER SPECTRUM ERROR, I.E. THE NORMALIZED DIFFERENCE BETWEEN THE TWO UPPER CURVES OF THE LEFT PANEL.



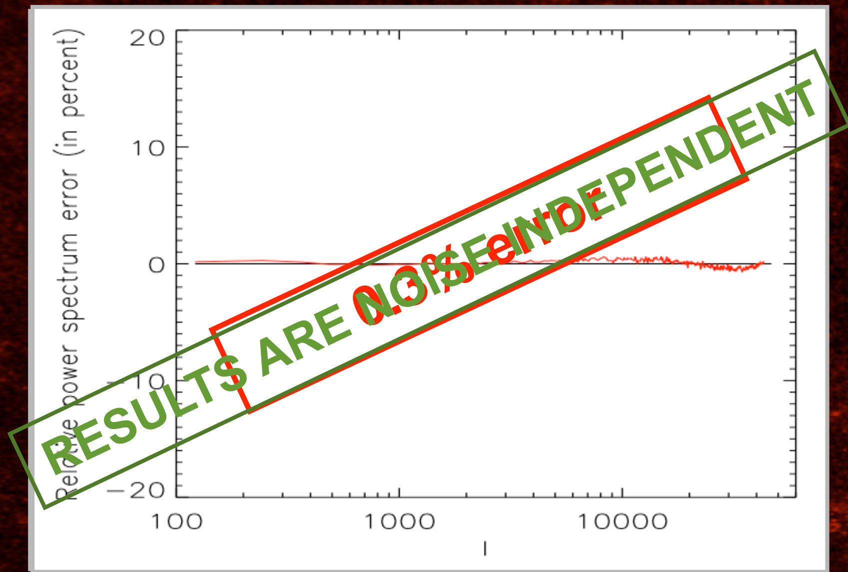


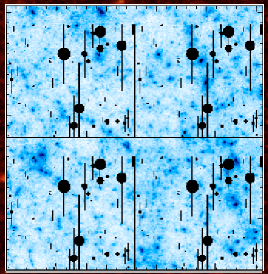
Noisy power spectrum estimation



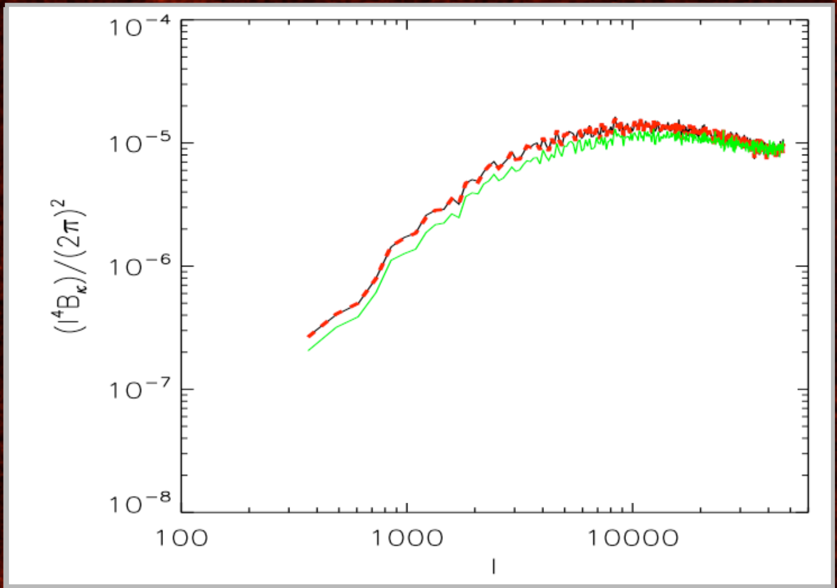
MEAN NOISY POWER SPECTRUM COMPUTED FROM
- 100 COMPLETE MASS MAPS (BLACK)
- 100 INPAINTED MAPS FROM INCOMPLETE SHEAR MAPS (RED)..

RELATIVE NOISY POWER SPECTRUM ERROR, I.E. THE NORMALIZED DIFFERENCE BETWEEN THE TWO CURVES OF THE LEFT PANEL.





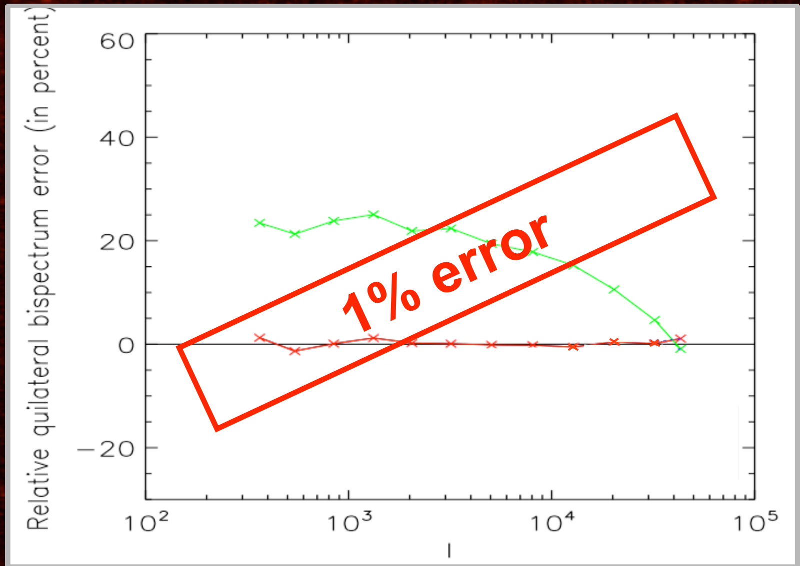
Equilateral bispectrum estimation



MEAN BISPECTRUM COMPUTED FROM

- 100 COMPLETE MASS MAPS (BLACK)
- 100 INPAINTED RECONSTRUCTED MAPS (RED)
- 100 INCOMPLETE MASS MAPS (GREEN).

RELATIVE BISPECTRUM ERROR, I.E. THE NORMALIZED DIFFERENCE BETWEEN THE TWO UPPER CURVES OF THE LEFT PANEL.



FASTLens

(FAst STatistics for weak Lensing)

http://www-irfu.cea.fr/Ast/fastlens_software.php

FASTLENS (FAST STATISTICS FOR WEAK LENSING) :
FAST METHOD FOR WEAK LENSING STATISTICS AND MAP MAKING

S. Pires, J.L. Starck, A. Amara, A. Réfrégier and J. Fadili

The analysis of weak lensing data requires to account for missing data such as masking out of bright stars. To date, the majority of lensing analyses uses the two point-statistics of the cosmic shear field. These can either be studied directly using the two-point correlation function, or in Fourier space, using the power spectrum. The two-point correlation function is unbiased by missing data but its direct calculation will soon become a burden with the exponential growth of astronomical data sets. The power spectrum is fast to estimate but a mask correction should be estimated. Others statistics can be used but these are strongly sensitive to missing data.

The solution that is proposed by FASTLens is to properly fill-in the gaps with only NlogN operations, leading to a complete weak lensing mass map from which we can compute straight forwardly and with a very good accuracy any kind of statistics like power spectrum or bispectrum. We propose also a new method to compute fastly and accurately the power spectrum and the bispectrum with a polar FFT algorithm.

Fig. 1: Simulated weak lensing mass map for a ΛCDM cosmological model with σ8= 0.9 and Ωm = 0.3. The region shown is 1° x 1°.

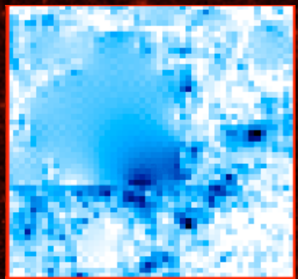
FASTLens (FAst STatistics for weak Lensing)

(Pires 2009, MNRAS)

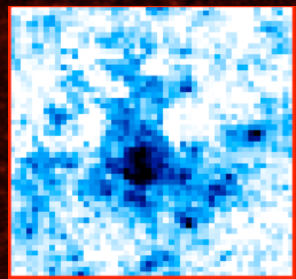
http://www-irfu.cea.fr/Ast/fastlens_software.php

✓ Inpainting method:

- ✓ Estimation of a complete dark matter mass map from incomplete shear maps

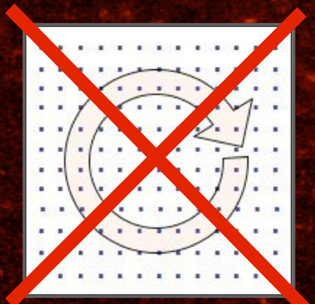


INPAINTING

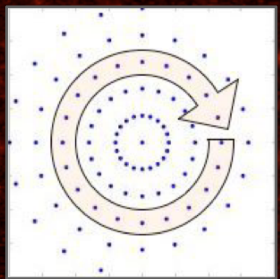


✓ Polar FFT code:

- ✓ Fast and Exact estimation of the power spectrum and the bispetrum



CARTESIAN FFT



POLAR FFT

STATICTIC ESTIMATION

$$P_\kappa =< \hat{\kappa}(\theta_i)\hat{\kappa}(\theta_j) >$$

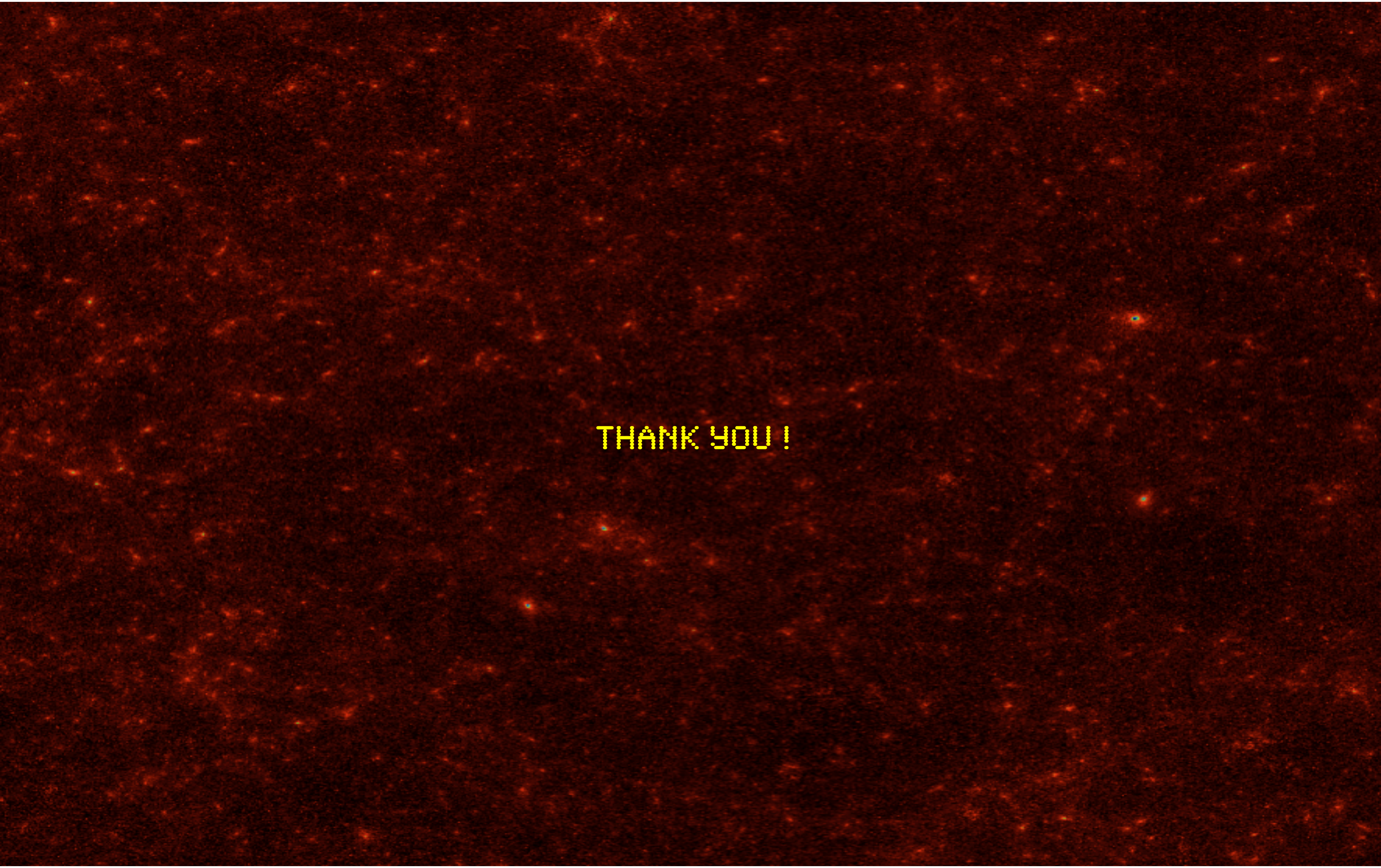
$$B_\kappa =< \hat{\kappa}(\theta_i)\hat{\kappa}(\theta_j)\hat{\kappa}(\theta_k) >$$

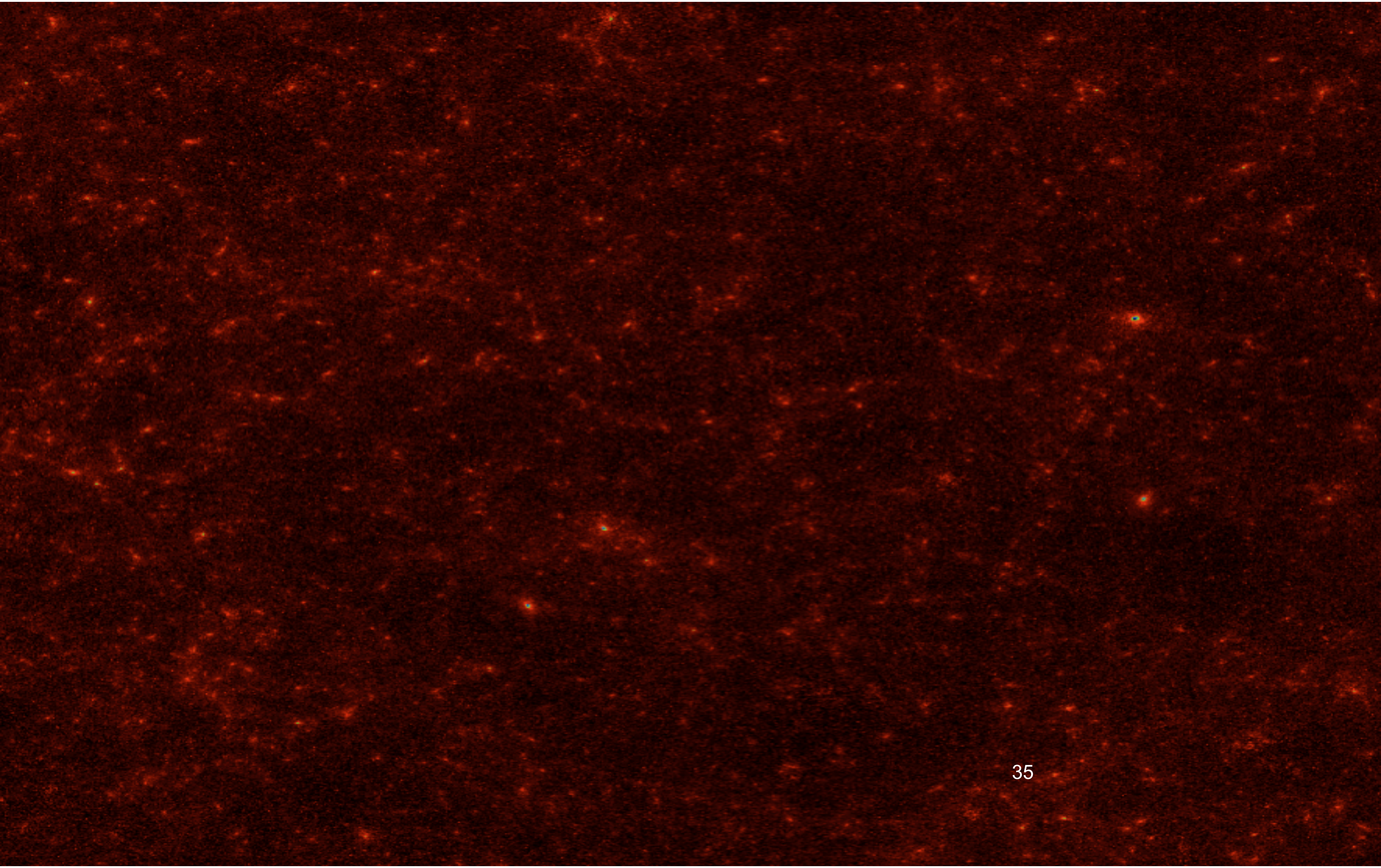
...

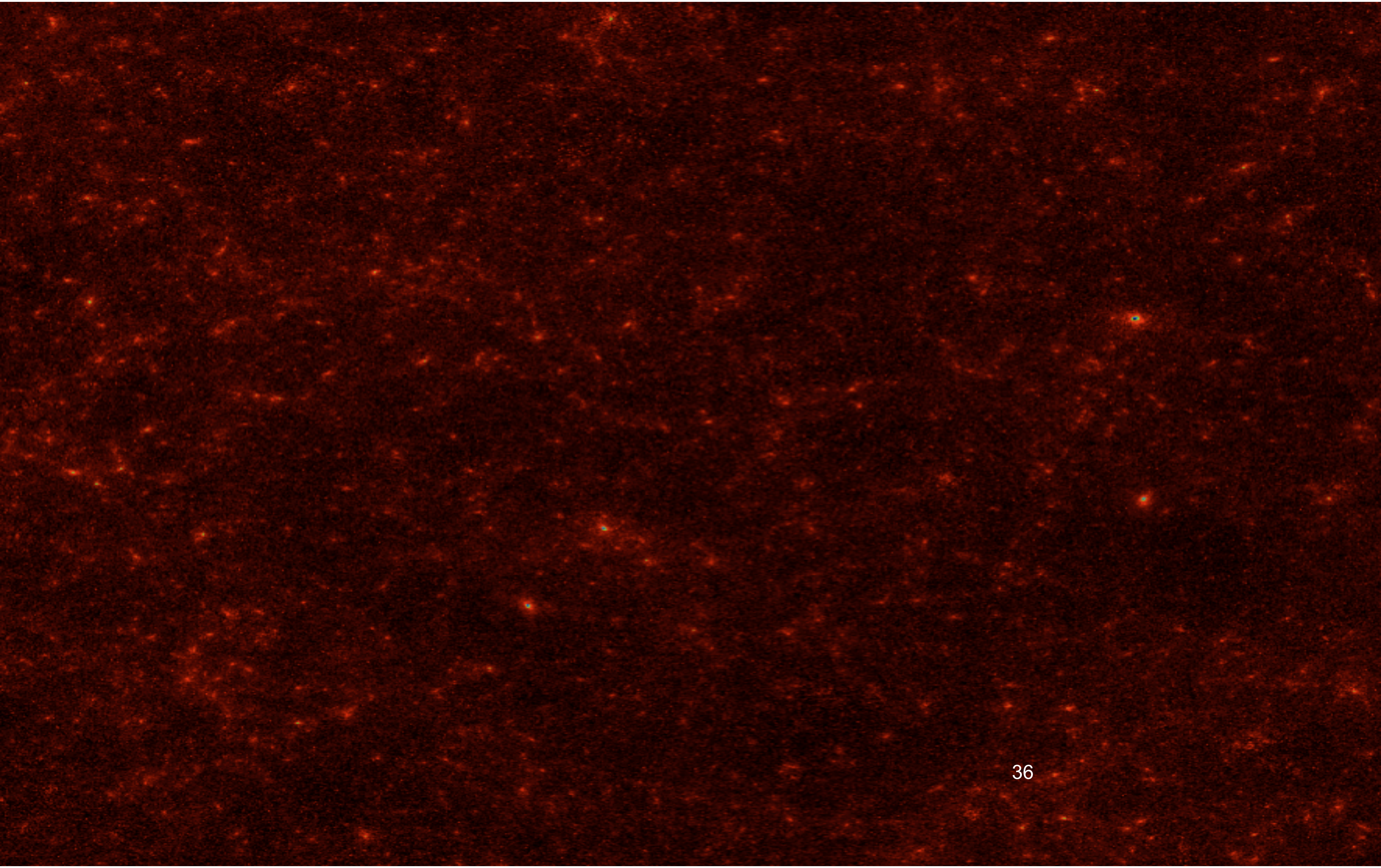
Conclusion

A method to reconstruct a full Weak Lensing mass map from incomplete shear maps has been developed (FASTLens software)

- ✓ Make faster the estimation of statistics:
  - ✓ The maximum error on power spectrum estimation is 1%
  - ✓ The maximum error on bispectrum estimation is 3%
- ✓ Enables estimation of many statistics:
  - ✓ Power spectrum, Bispectrum, Trispectrum...
  - ✓ Dark matter statistics (cluster abundance, cluster correlations...)
- ✓ Enables unbiased filtering


THANK YOU !





Inpainting algorithm from shear maps

$$
\begin{aligned}
Y &= P_1 * \gamma_1^{obs} + P_2 * \gamma_2^{obs} \\
I_{max} &= 100 \\
\kappa^0 &= 0 \\
R^0 &= Y \\
\lambda_{max} &= max(|\alpha = \Phi^T Y|) \\
\lambda_{min} &= 0
\end{aligned}
$$

$for\ n = 0\ to\ I_{max}\ do\ begin$

$U = \kappa^n + MR^n(\gamma^{obs})\ et$

$R^n(\gamma^{obs}) = P_1 * (\gamma_1^{obs} - P_1 * \kappa_n) + P_2 * (\gamma_2^{obs} - P_2 * \kappa_n)$

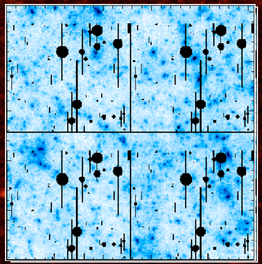
Digital Cosine Transform (DCT) of U: $\alpha = \Phi^T U$

Threshold determination: λ_n

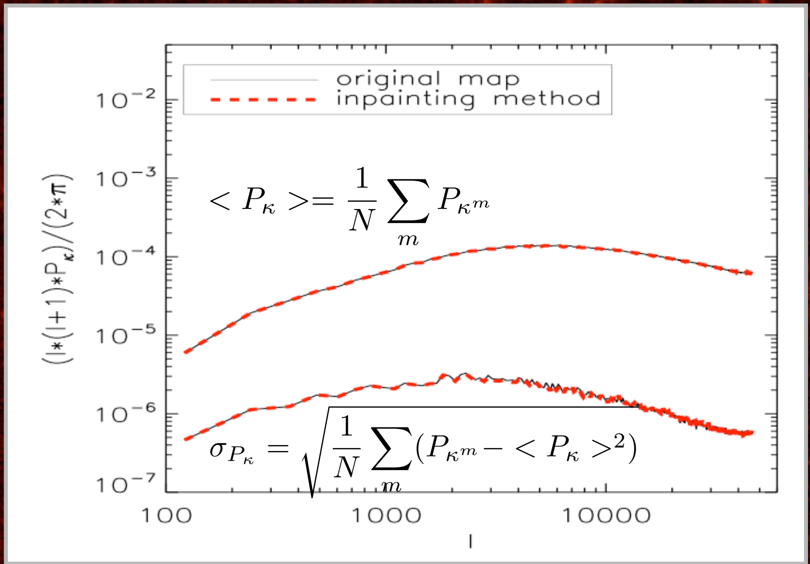
Hard-thresholding of α *with* $\alpha_n : \tilde{\alpha} = S_{\lambda_n} \alpha$

$\kappa^{n+1} = \Phi \tilde{\alpha}$

$n = n + 1 \text{ if } n < I_{max}(2)$

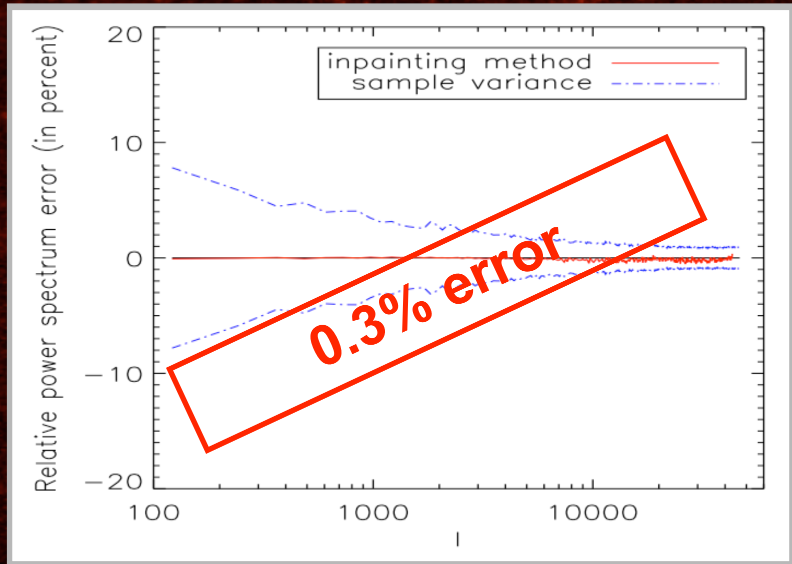


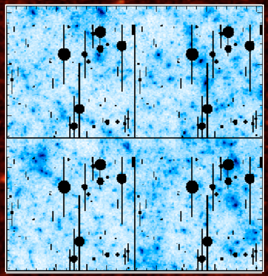
Power spectrum estimation



MEAN POWER SPECTRUM COMPUTED FROM
- 100 COMPLETE MASS MAPS (BLACK)
- 100 INPAINTED MAPS (RED).

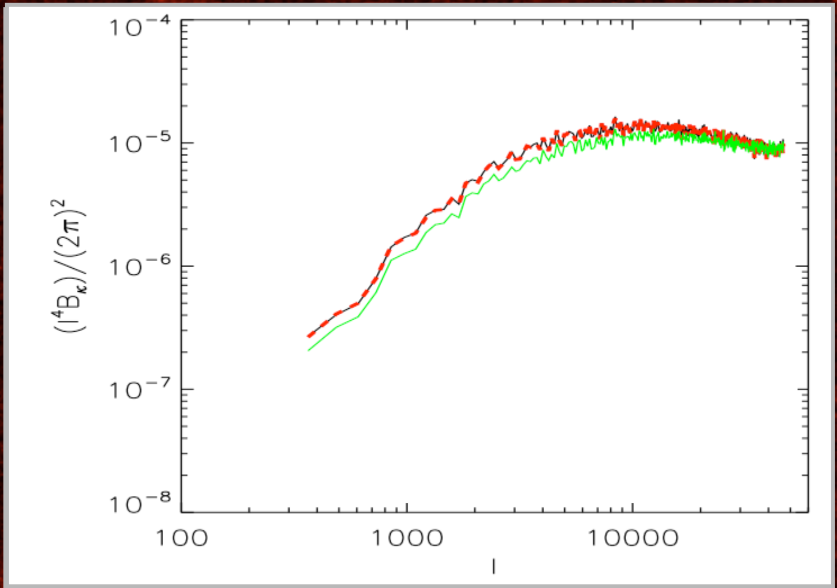
RELATIVE POWER SPECTRUM ERROR, I.E. THE NORMALIZED DIFFERENCE BETWEEN THE TWO UPPER CURVES OF THE LEFT PANEL.





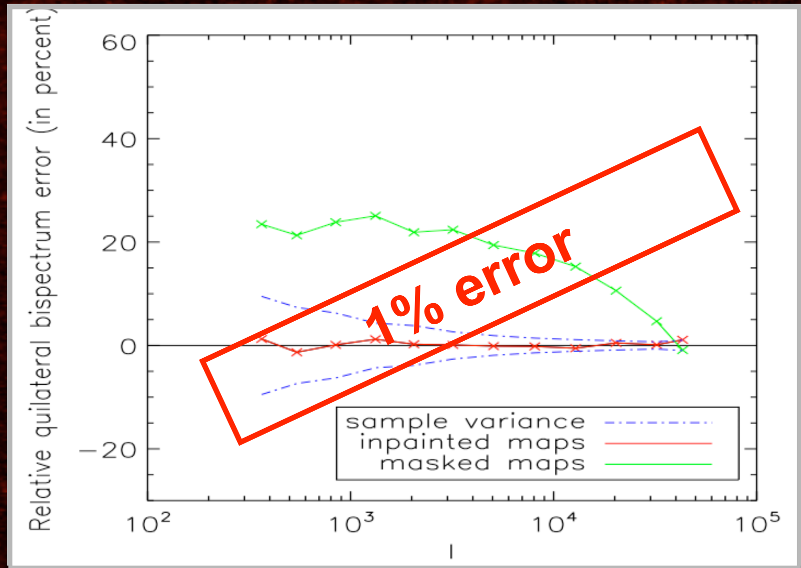
Equilateral bispectrum estimation

FASTLens, Pires et al 2008a

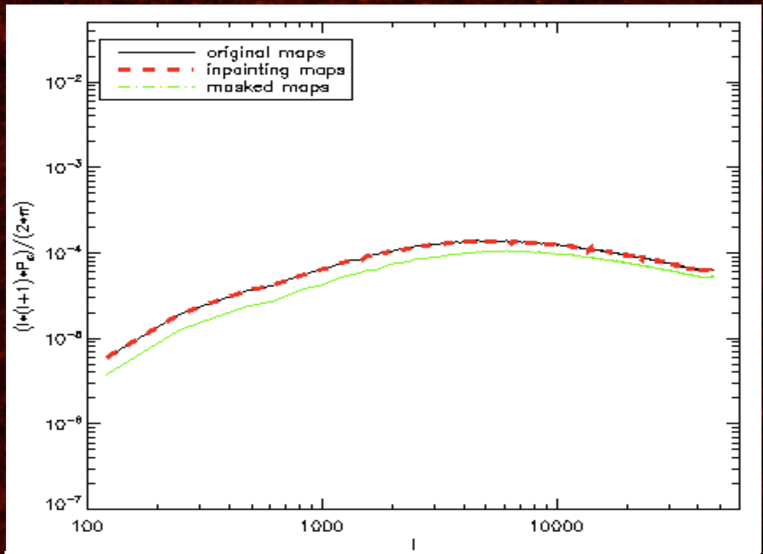


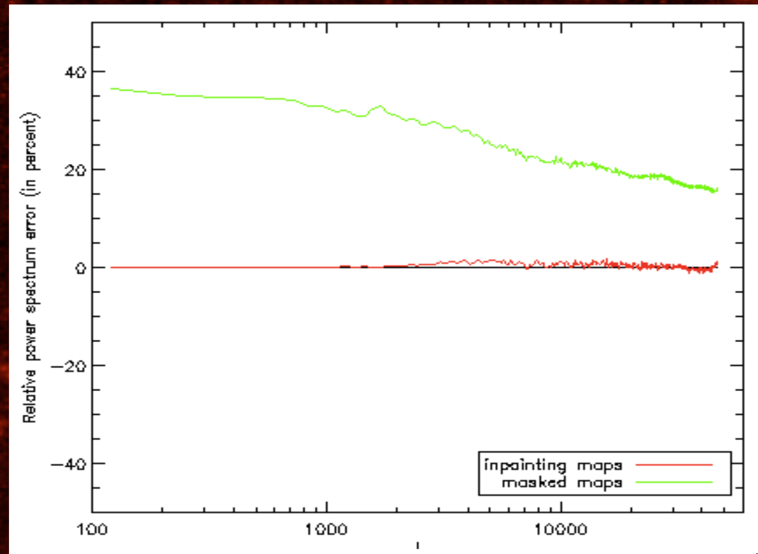
MEAN BISPECTRUM COMPUTED FROM
- 100 COMPLETE MASS MAPS (BLACK),
- 100 INPAINTED RECONSTRUCTED MAPS (RED)
- 100 INCOMPLETE MASS MAPS (GREEN) .

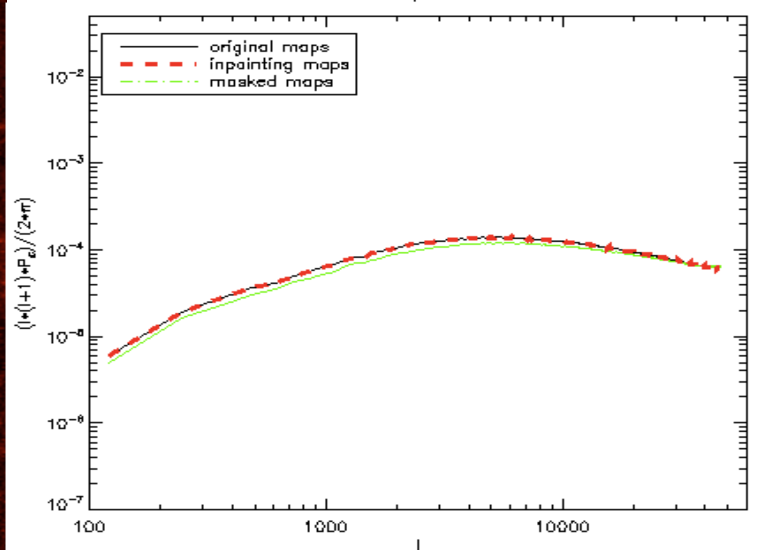
RELATIVE BISPECTRUM ERROR, I.E. THE NORMALIZED DIFFERENCE BETWEEN THE TWO UPPER CURVES OF THE LEFT PANEL.

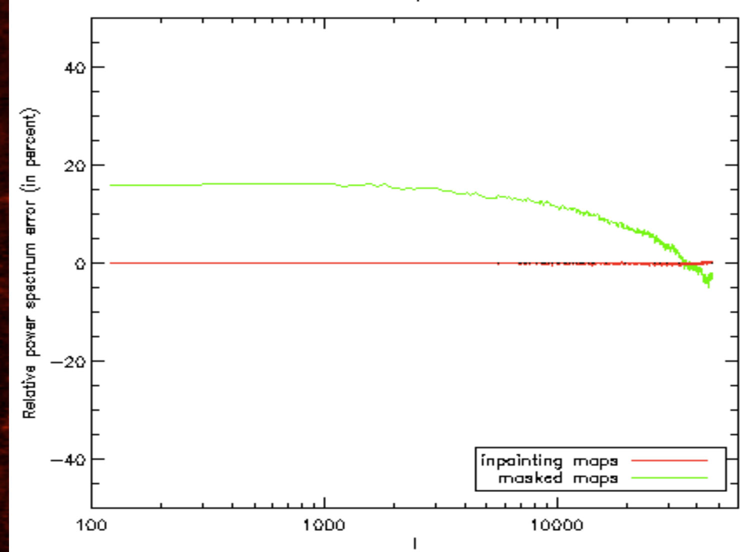


Power spectrum estimation









Simulations numériques

- ✓ 3D N-body simulation by solving the hydrodynamic equations on a AMR grid (Ramses code)
- ✓ Dark matter mass maps simulation by projeting the density along the line of sight (using the Born approximation):

$$\kappa_e \approx \frac{3H_0^2 {}_m L}{2c^2} \sum_i \frac{\chi_i(\chi_0 - \chi_i)}{\chi_0 a(\chi_i)} \left(\frac{n_p R^2}{N_t s^2} - \Delta r_{f_i} \right)$$