Space & Cosmology:
Tackling Big Data from the Sky

Jean-Luc Starck

http://jstarck.cosmostat.org

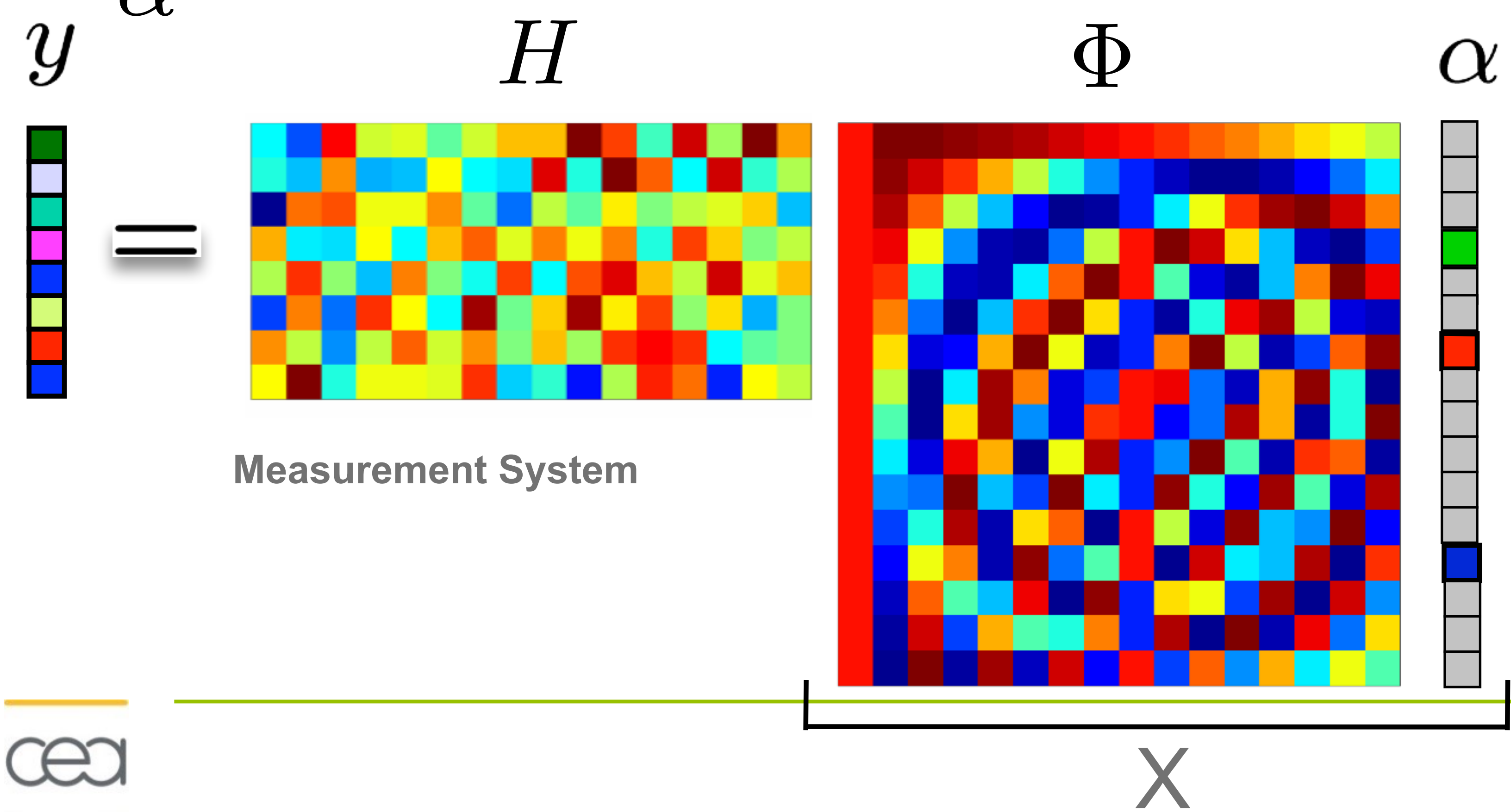
Sparse Recovery & Inverse Problems

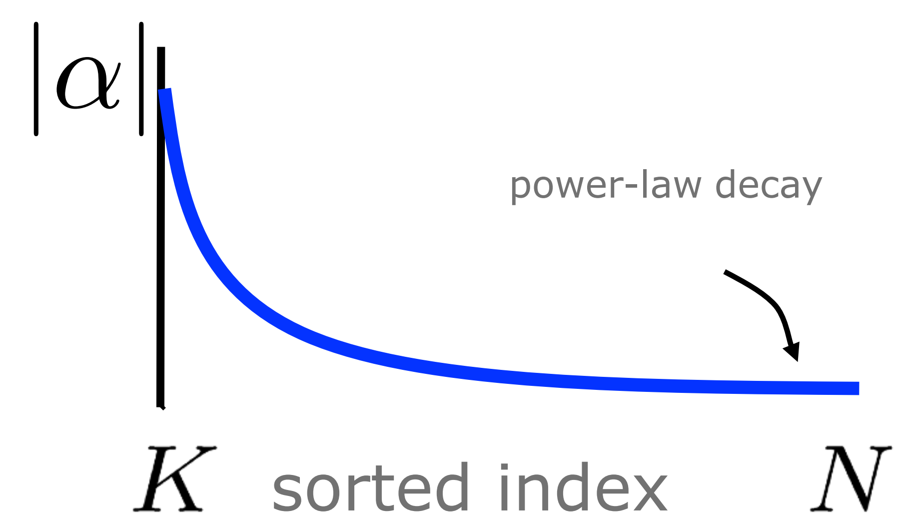
$$Y = HX + N$$

$X = \Phi\alpha$, and α is sparse

- Denoising
- Deconvolution
- Component Separation
- Inpainting
- Blind Source Separation
- Minimization algorithms
- Compressed Sensing

$$\min_{\alpha} \|\alpha\|_p^p \quad \text{subject to} \quad \|Y - H\Phi\alpha\|^2 \leq \epsilon$$





➡Part 1: Introduction to Accurate Space Cosmology

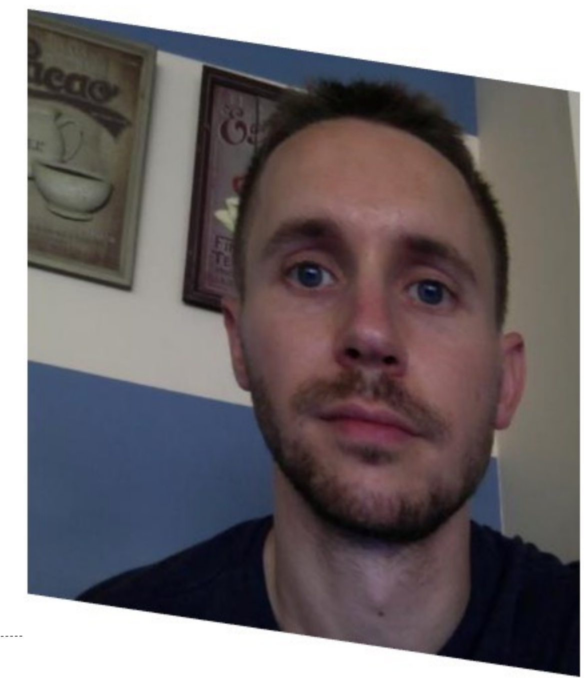
➡Part 2: Inverse Problems

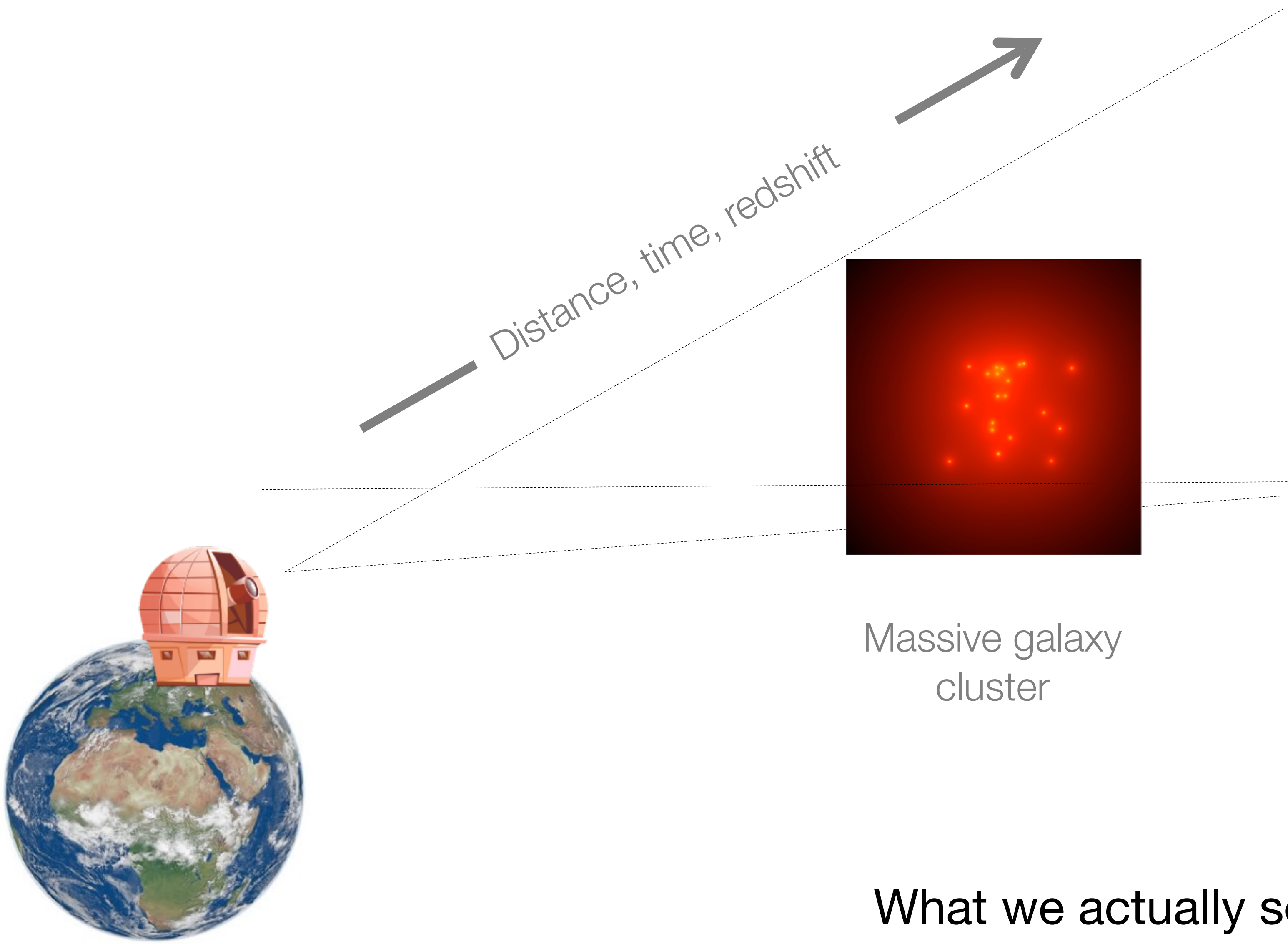
➡**Part 3: Euclid Weak Lensing**

- ➡**Introduction to Weak Lensing**
- ➡The Euclid space projet and its mathematical challenges
- ➡Advanced methodologies for Euclid

Gravitational lensing

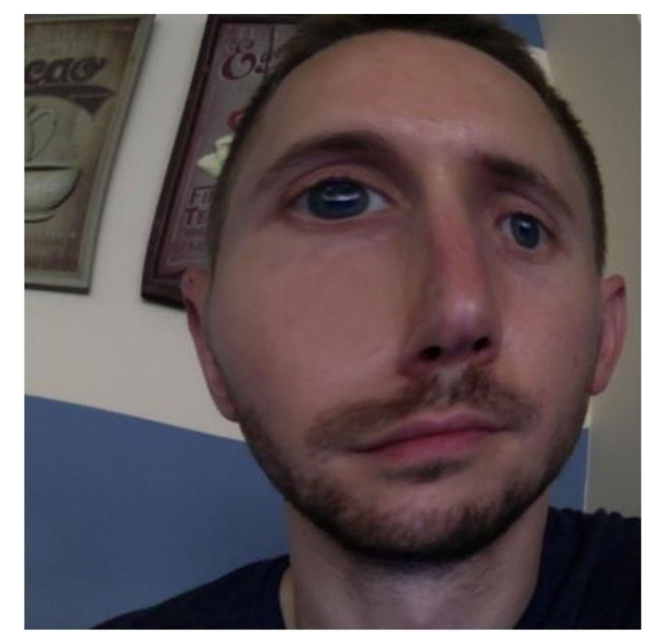
Distant source image





Observation point

What we actually see



Austin Peel

Strong Gravitational lensing

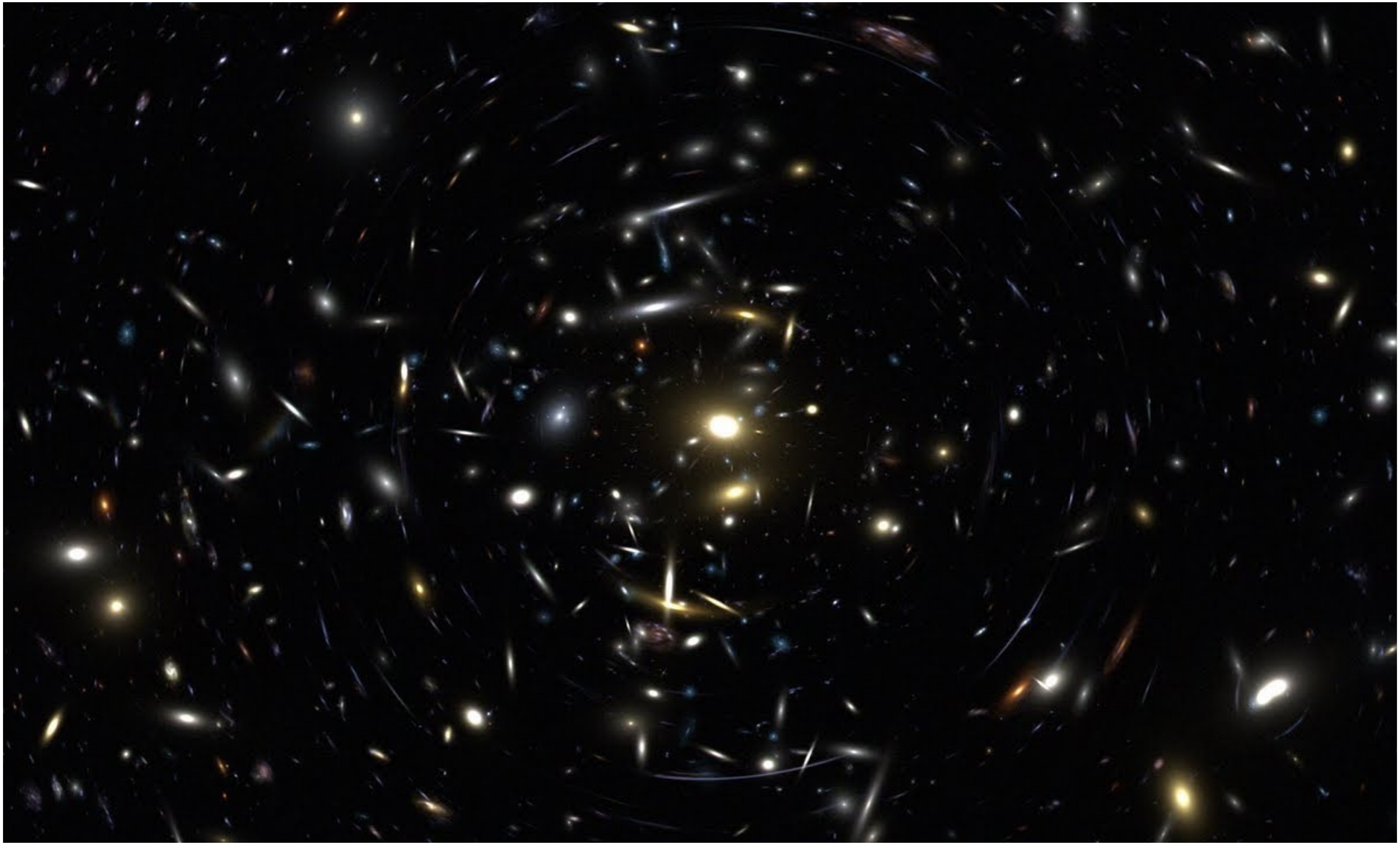
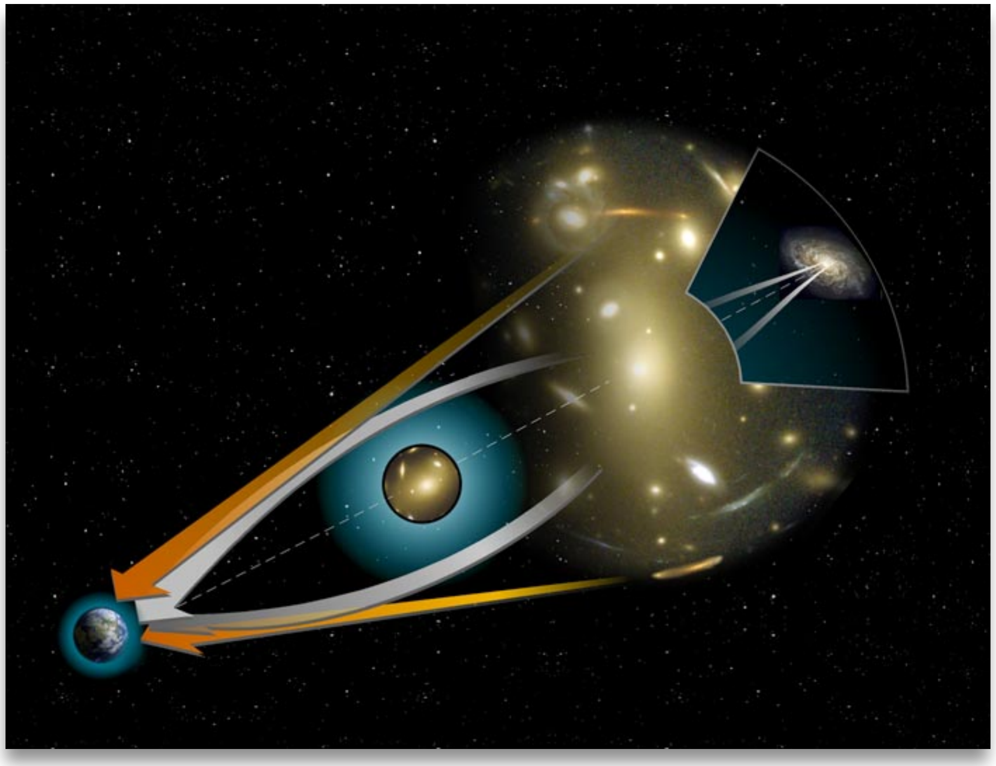


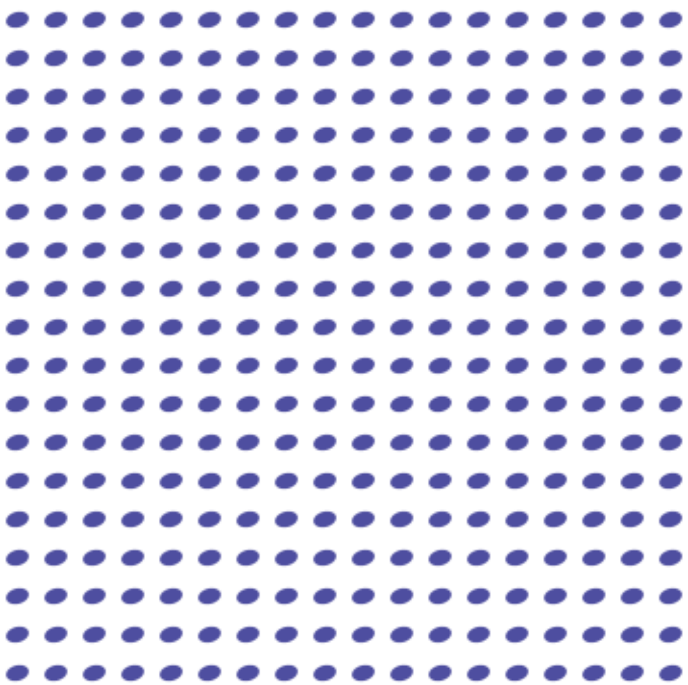
Image: nasa.gov

Weak Gravitational lensing

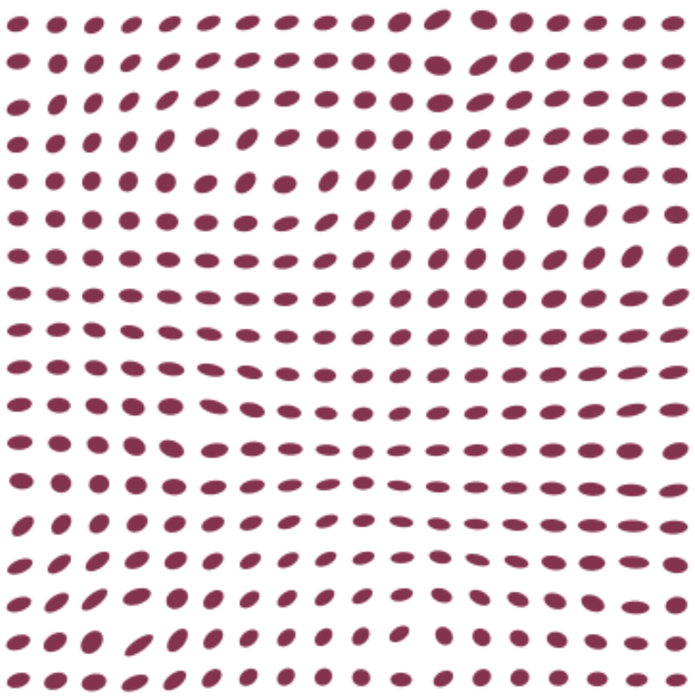
Very *slight* distortions of the measured shapes of distant galaxies

Order of a few percent effect





Unlensed sources



Weak lensing

Statistical shape correlations across angular scales carry cosmological information

Total matter content, clustering amplitude, dark energy equation of state, etc.

Ellipticity as a measure of shear

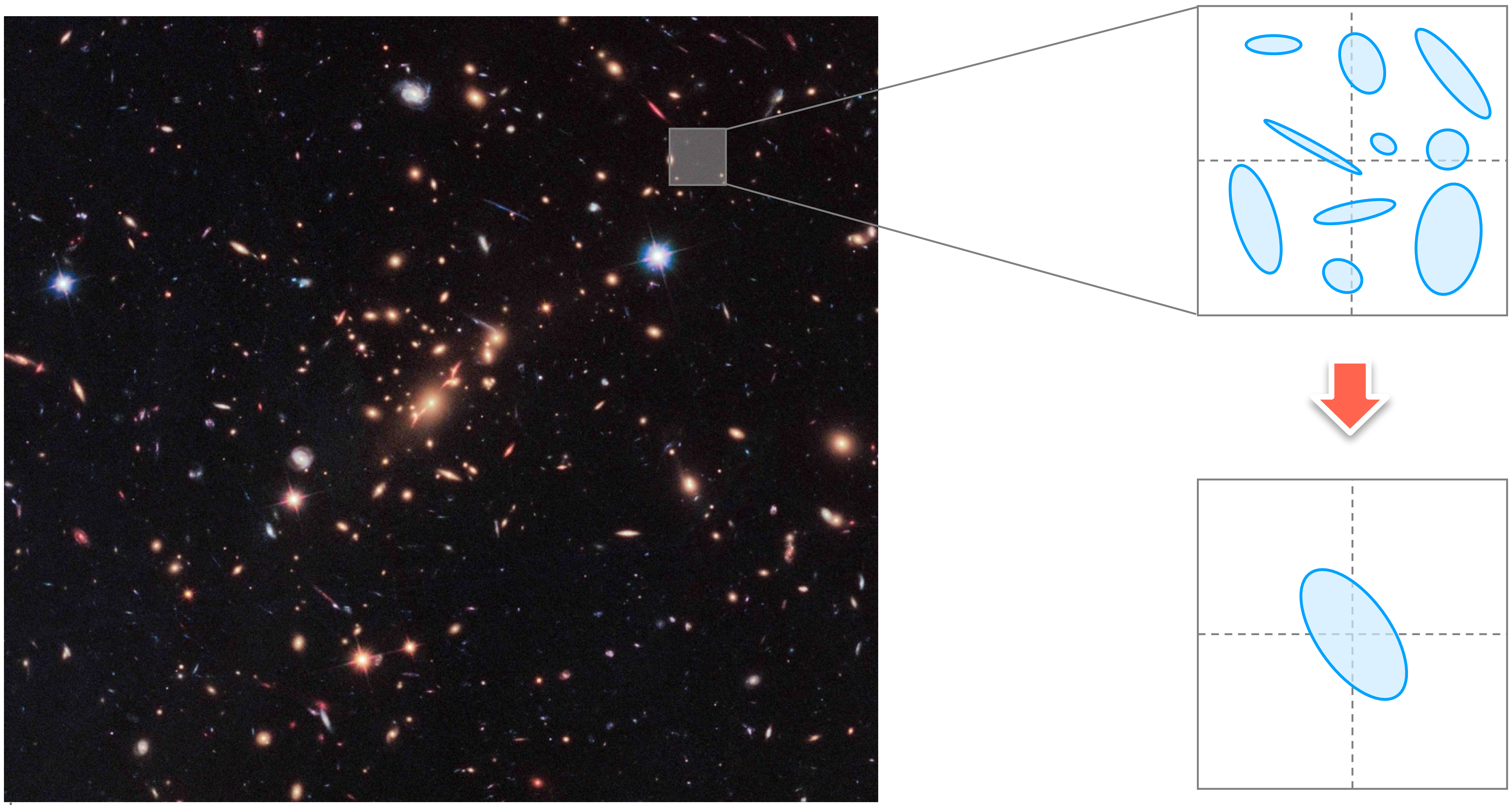


Image: nasa.gov

mean shape in this patch

WL assumption < ellipticity > ≈ shear

Shape measurement

Real galaxy on the sky

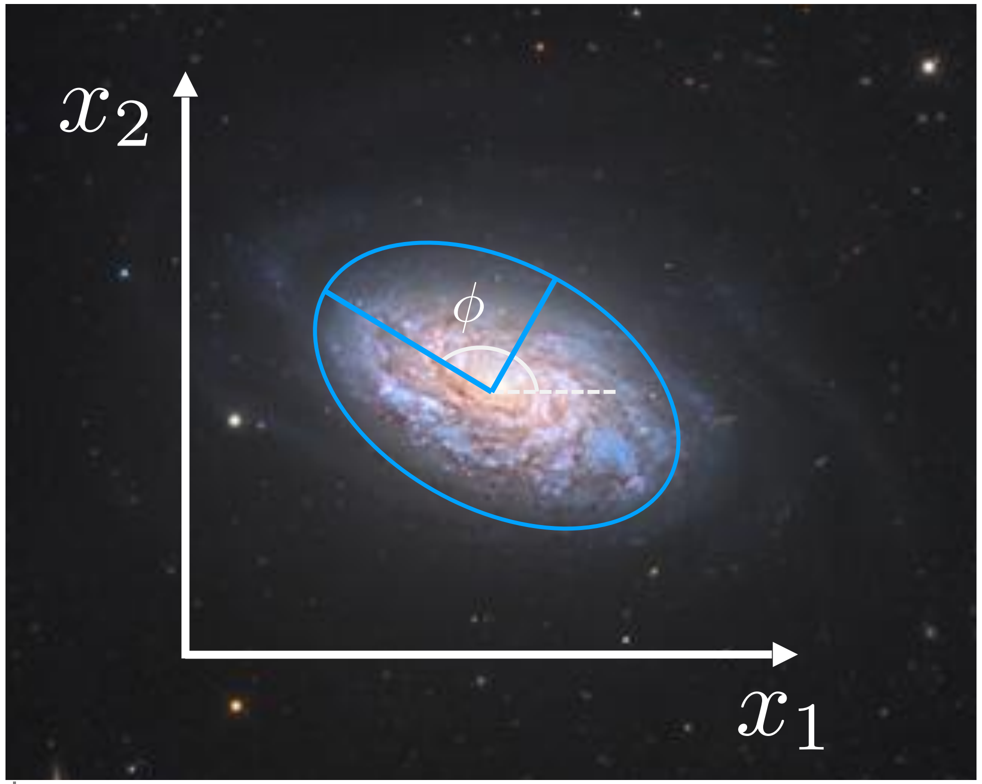
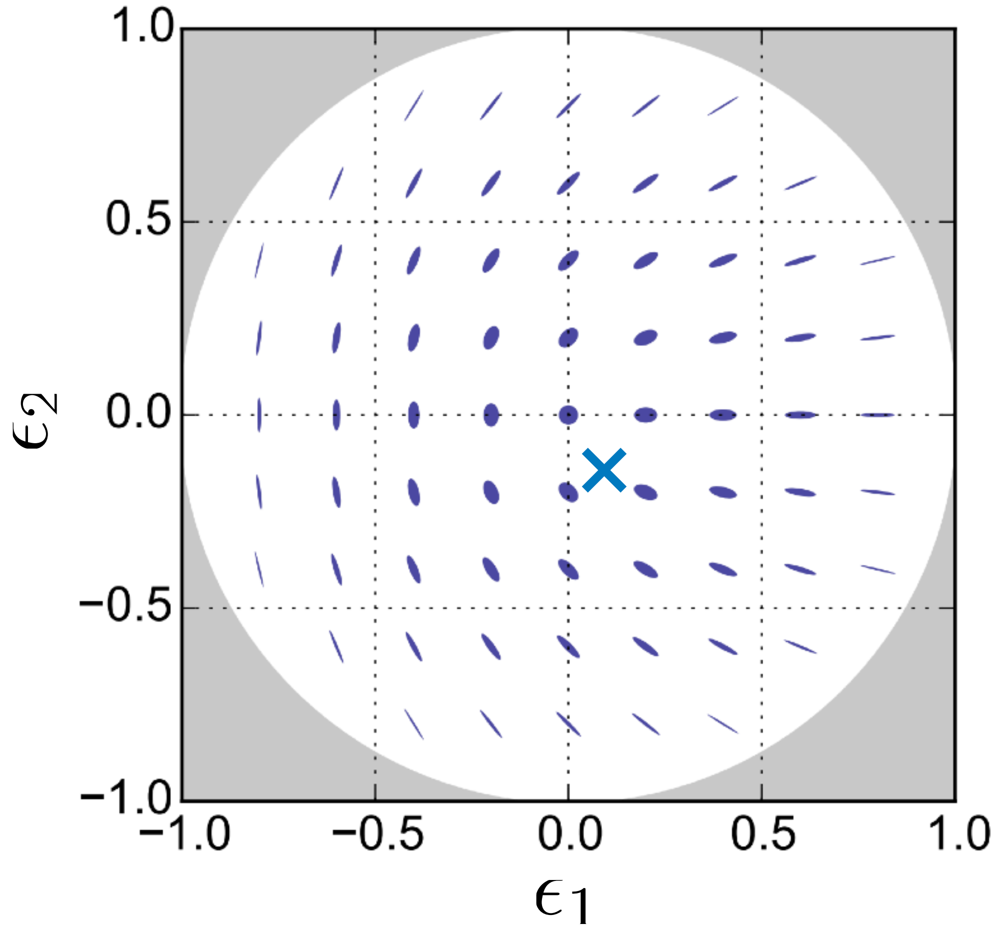


Image: nasa.gov

$$\epsilon(\boldsymbol{x}) \equiv \epsilon_1 + \mathrm{i}\epsilon_2 = |\epsilon| e^{2\mathrm{i}\phi}$$

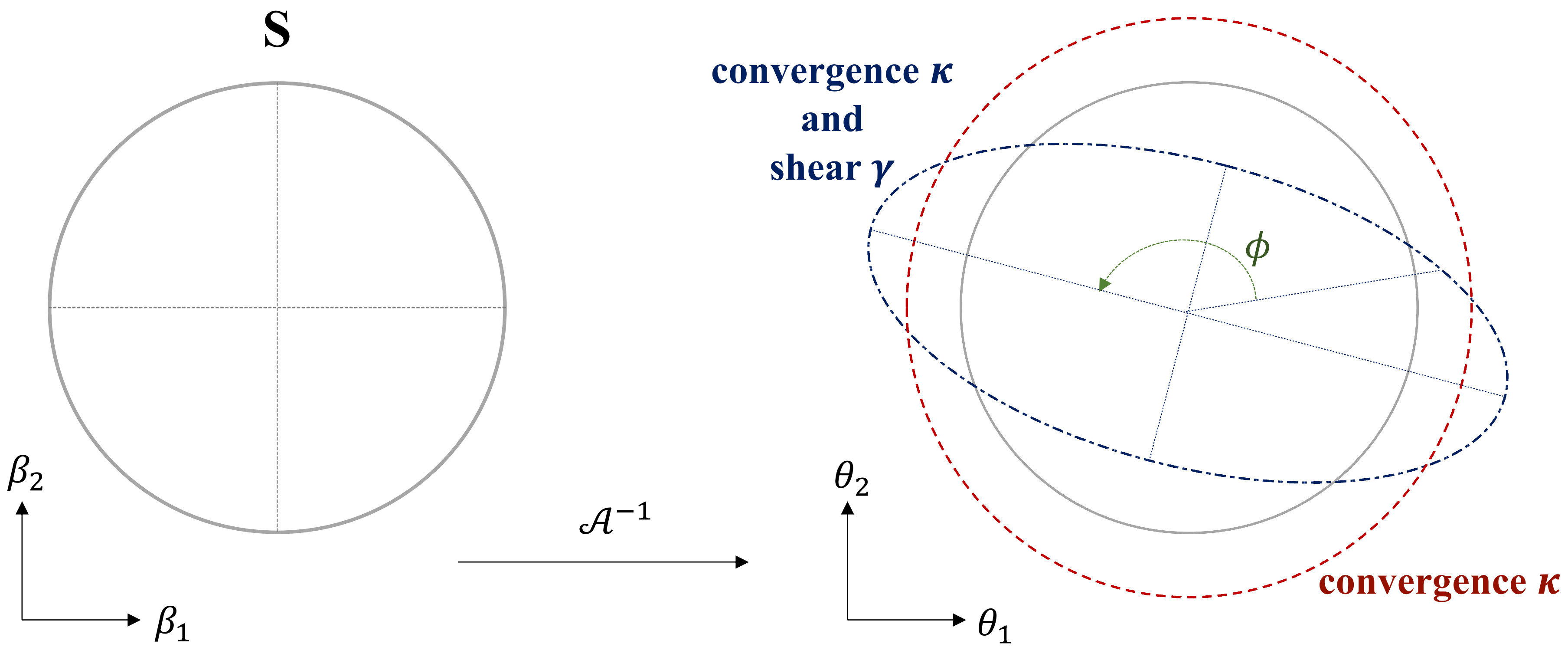
Ellipticity component space



Ellipticity and shear are spin-2 quantities characterised as a complex field

Shear and convergence

$\kappa << 1, \quad |\gamma| << 1$

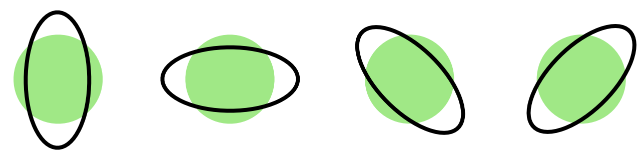


- $\gamma = \gamma_1 + i\gamma_2$ **shear**
- κ **convergence**

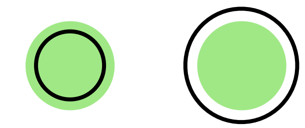
$$A = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix}$$

Shear and convergence

galaxy shape catalog

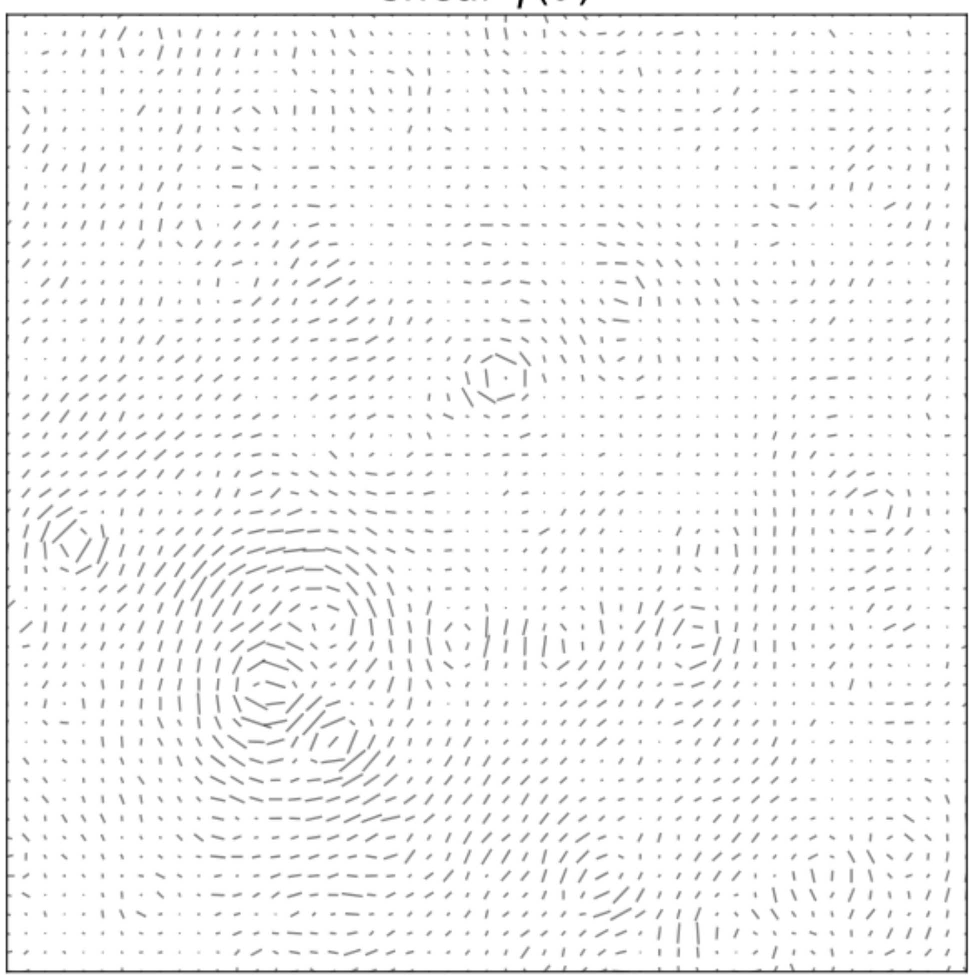


$$\gamma = \gamma_1 + \mathrm{i}\gamma_2$$

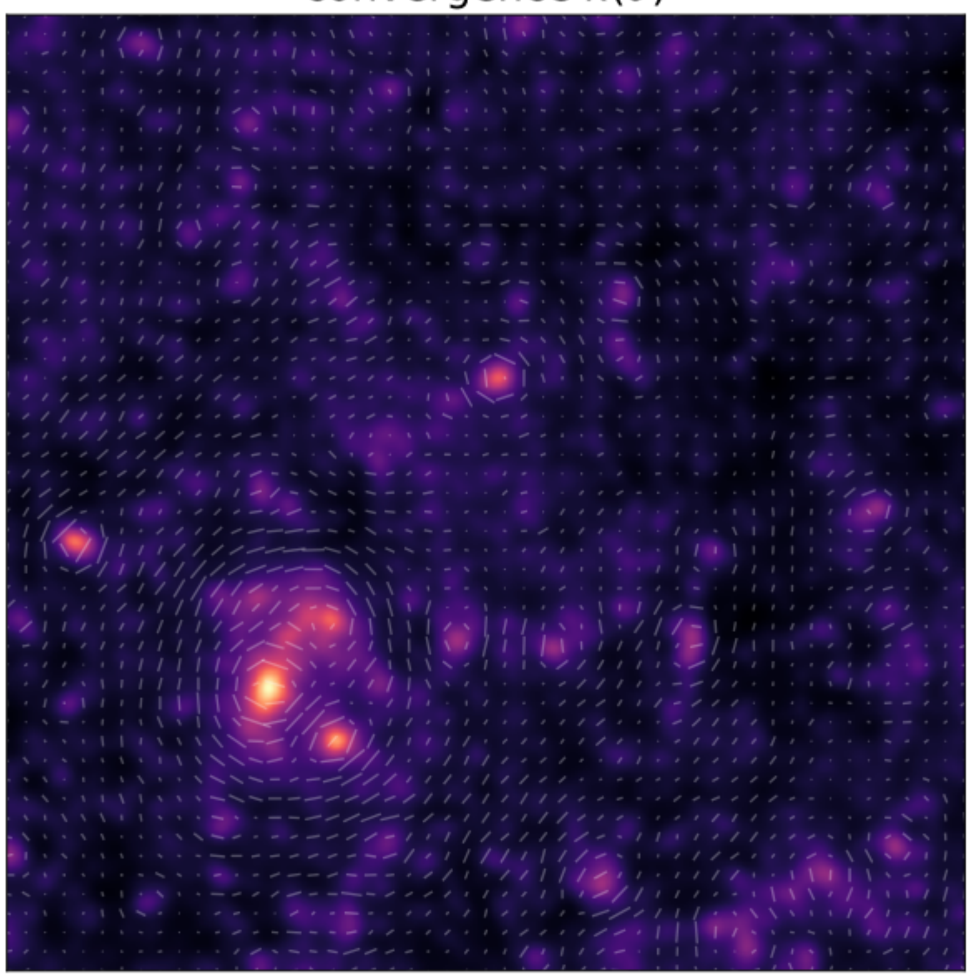


$$\kappa$$

shear $\gamma(\vec{\theta})$



convergence $\kappa(\vec{\theta})$



Convergence by shear inversion

Shear and convergence are derivable from a lensing potential $\psi(\boldsymbol{\theta})$

$$\gamma_1 = \frac{1}{2}\left(\partial_1^2 - \partial_2^2\right)\psi \qquad \gamma_2 = \partial_1\partial_2\psi \qquad \kappa = \frac{1}{2}\left(\partial_1^2 + \partial_2^2\right)\psi$$

Kaiser & Squires (1993)

$$\kappa(\boldsymbol{\theta}) - \kappa_0 = \frac{1}{\pi}\int_{\mathcal{R}^2} \mathrm{d}^2\theta' \, \mathcal{D}^*(\boldsymbol{\theta} - \boldsymbol{\theta}') \, \gamma(\boldsymbol{\theta}')$$

$$\mathcal{D}(\boldsymbol{\theta}) = \frac{\theta_2^2 - \theta_1^2 - 2\mathrm{i}\theta_1\theta_2}{|\boldsymbol{\theta}|^4}$$

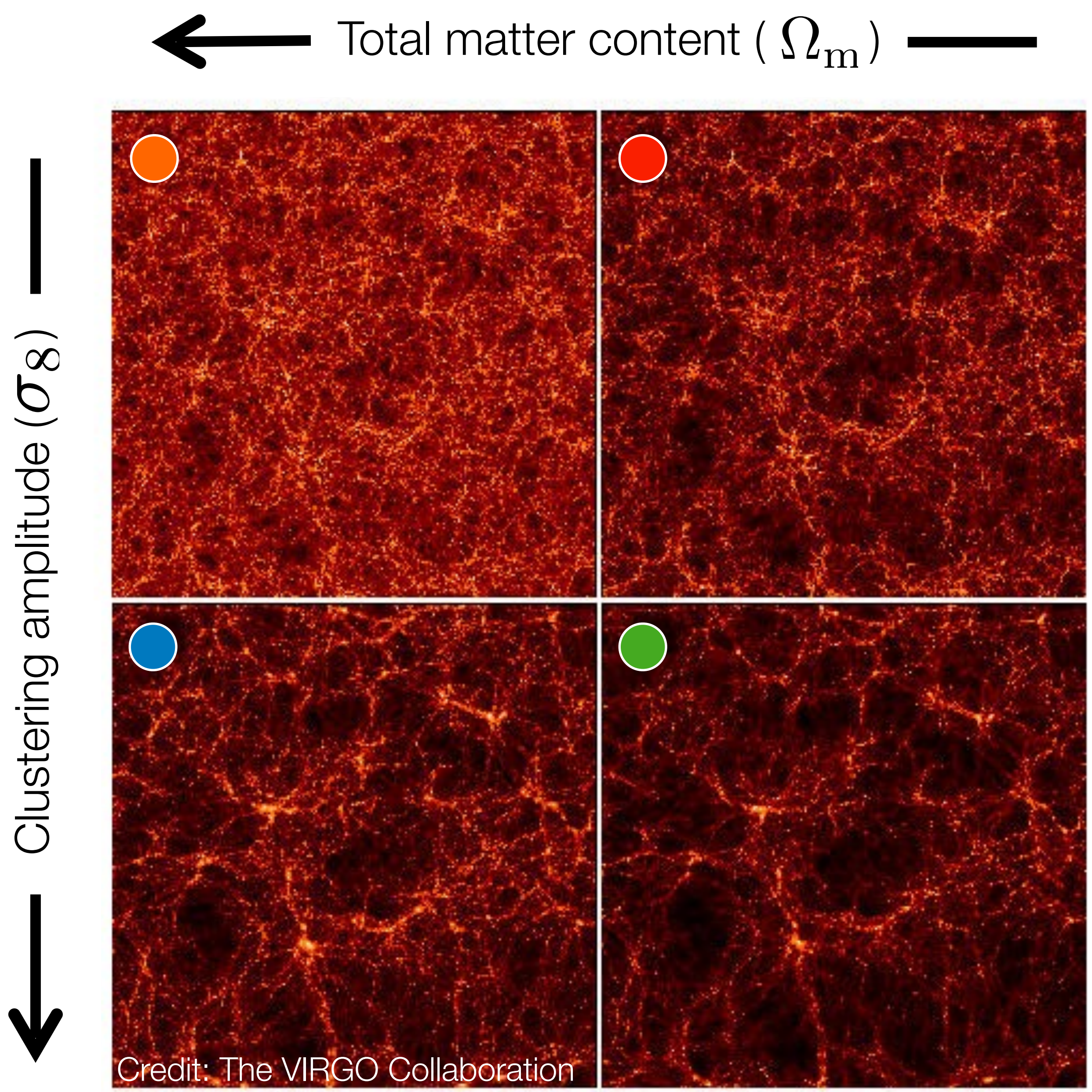
Easy in Fourier space

$$\hat{\kappa}(\boldsymbol{\ell}) = \hat{\mathcal{D}}^*(\boldsymbol{\ell}) \, \hat{\gamma}(\boldsymbol{\ell}) \qquad (\boldsymbol{\ell} \neq 0)$$

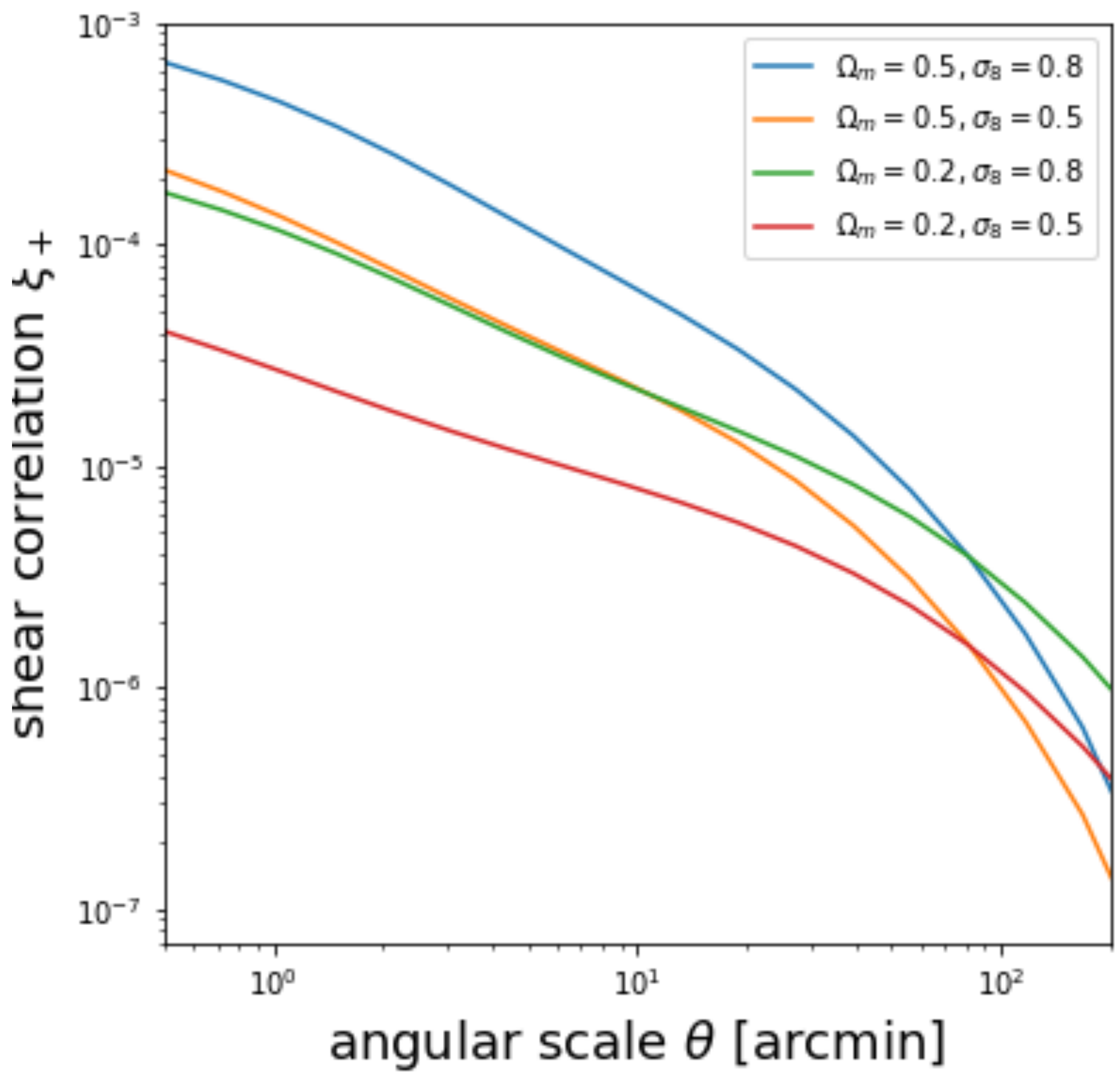
$$\hat{\mathcal{D}}(\boldsymbol{\ell}) = \frac{\ell_1^2 - \ell_2^2 + 2\mathrm{i}\ell_1\ell_2}{|\boldsymbol{\ell}|^2}$$

Regularize shear values
to a grid and use FFTs !

Weak lensing sensitivities



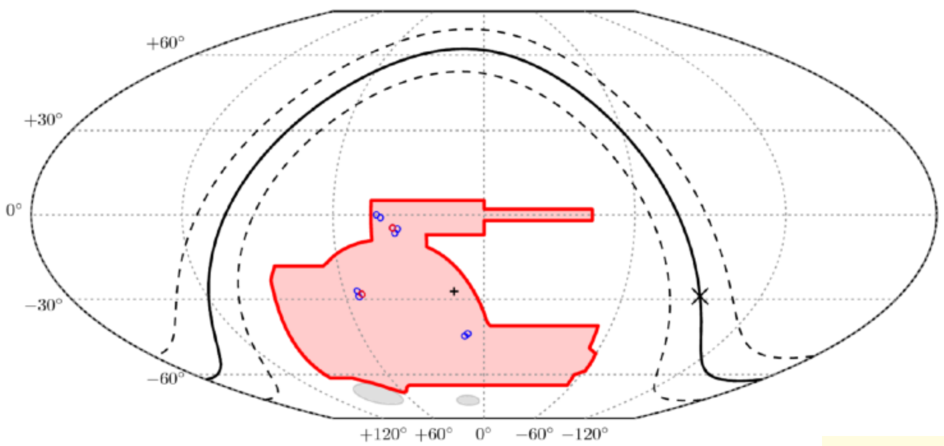
Simulated large-scale structure



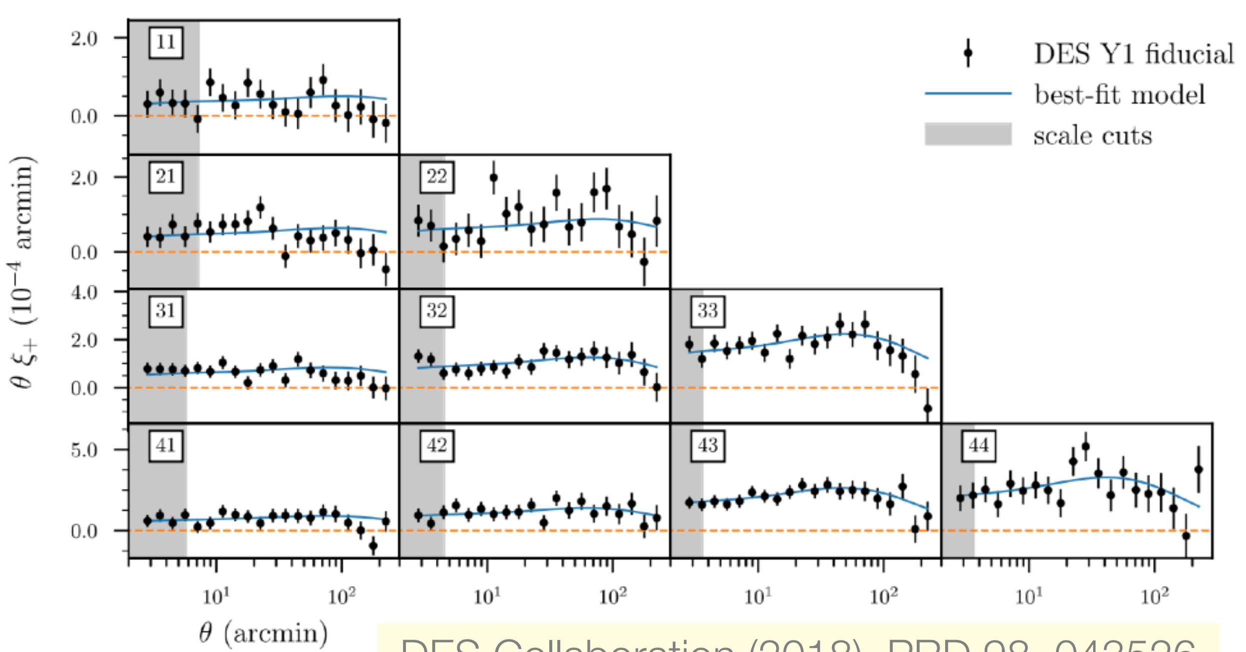
Real data and summary statistics

Dark Energy Survey (DES)

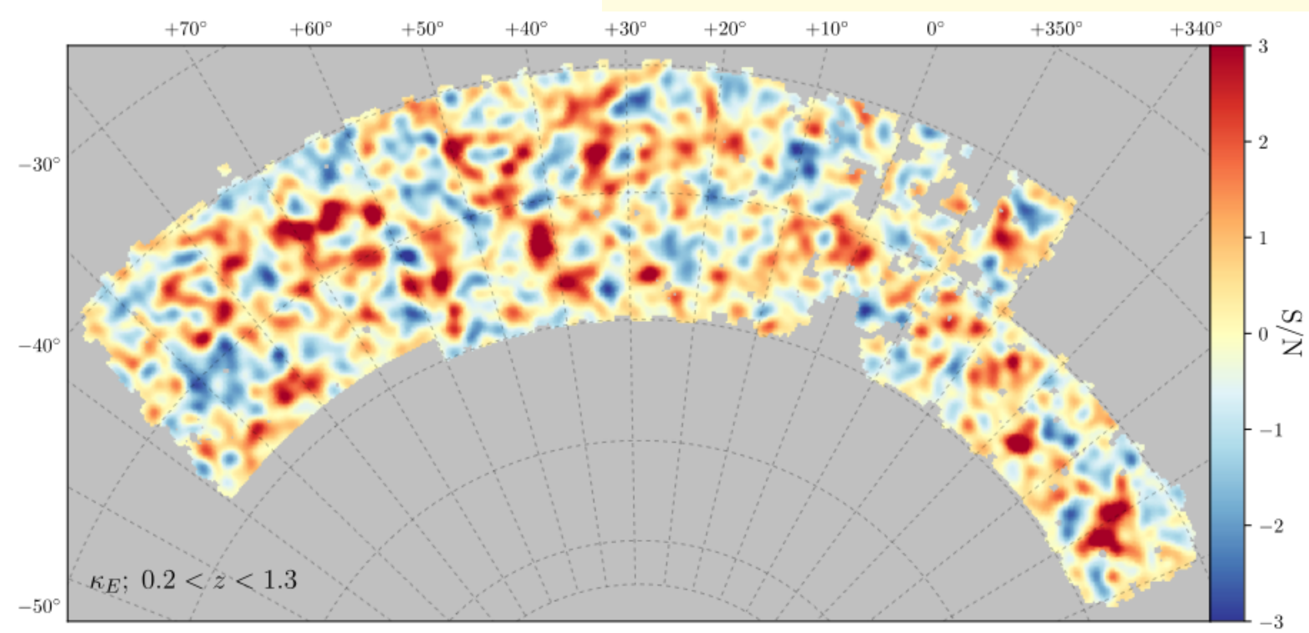
darkenergysurvey.org



Wallis et al. (2017), arXiv:1703.09233



DES Collaboration (2018), PRD 98, 043526



Chang et al. (2018), MNRAS 475, 3165

Cosmological constraints

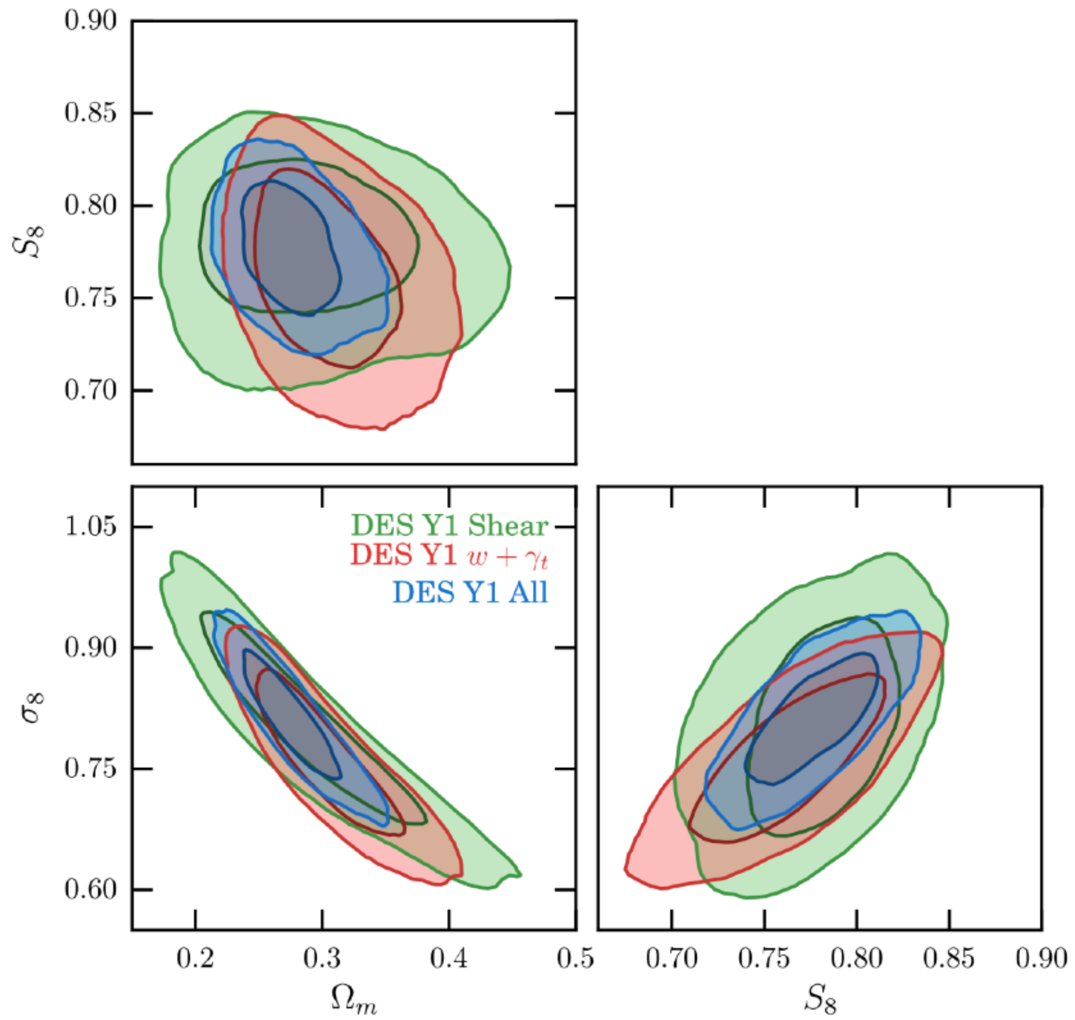
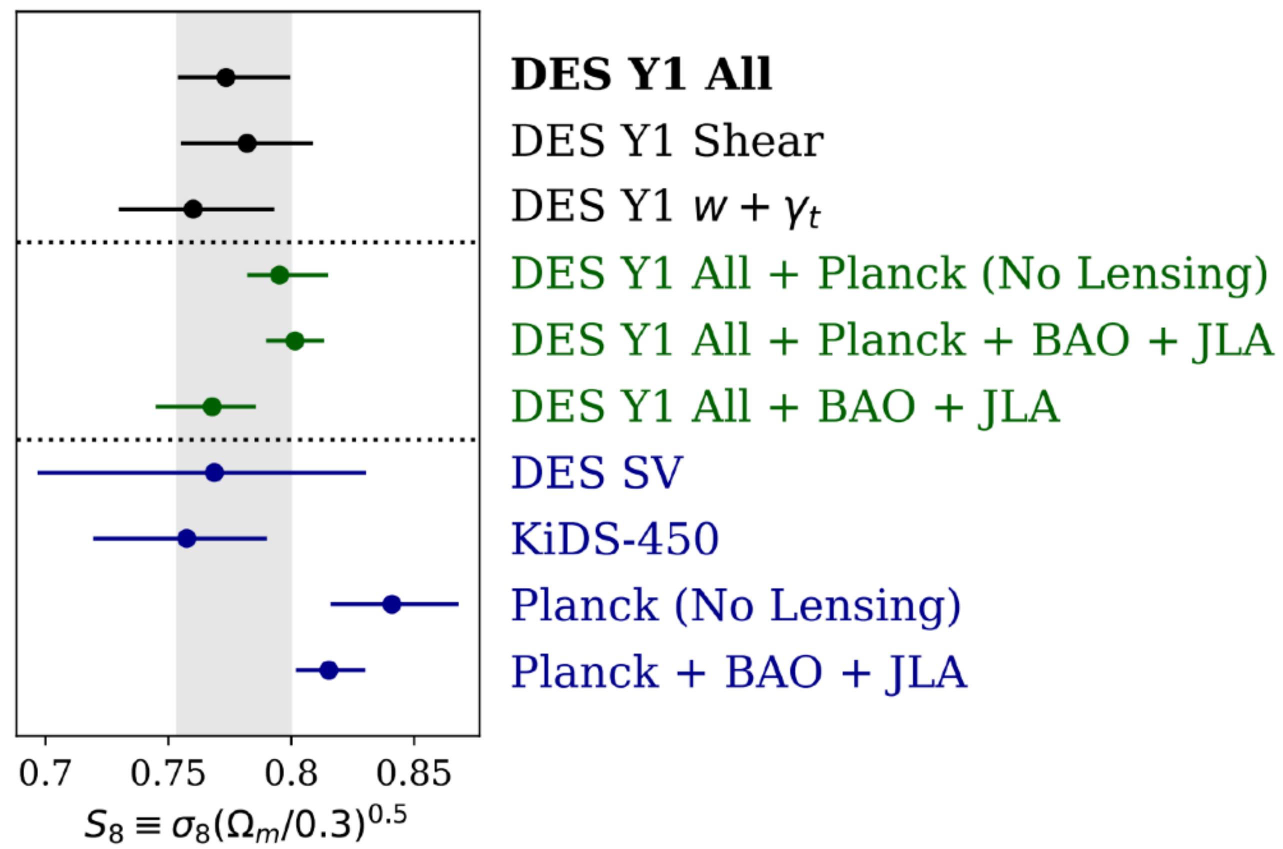


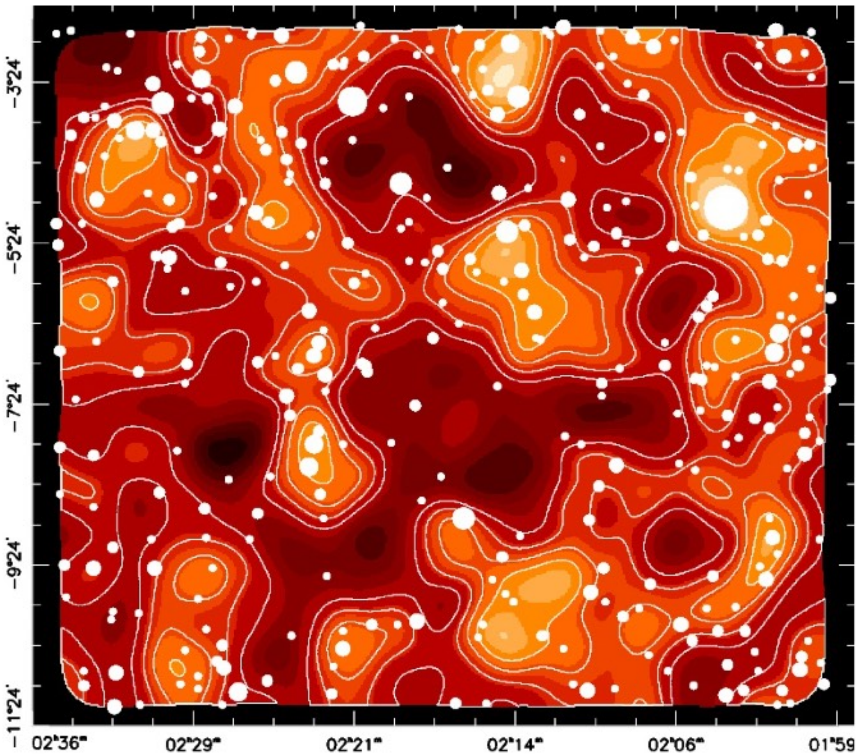
FIG. 5. ΛCDM constraints from DES Y1 on Ω_m, σ_8, and S_8 from cosmic shear (green), REDMAGIC galaxy clustering plus galaxy-galaxy lensing (red), and their combination (blue). Here, and in all such 2D plots below, the two sets of contours depict the 68% and 95% confidence levels.



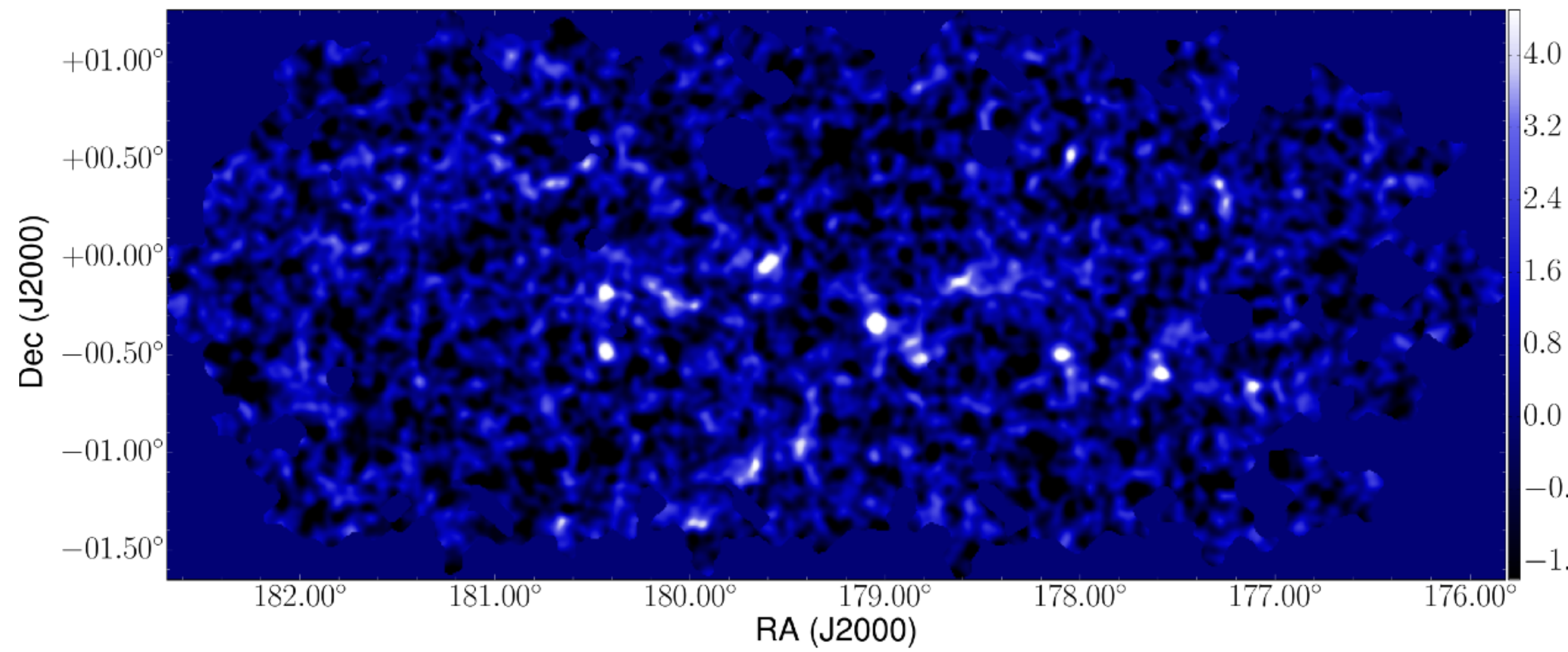
DES Collaboration (2018), PRD 98, 043526

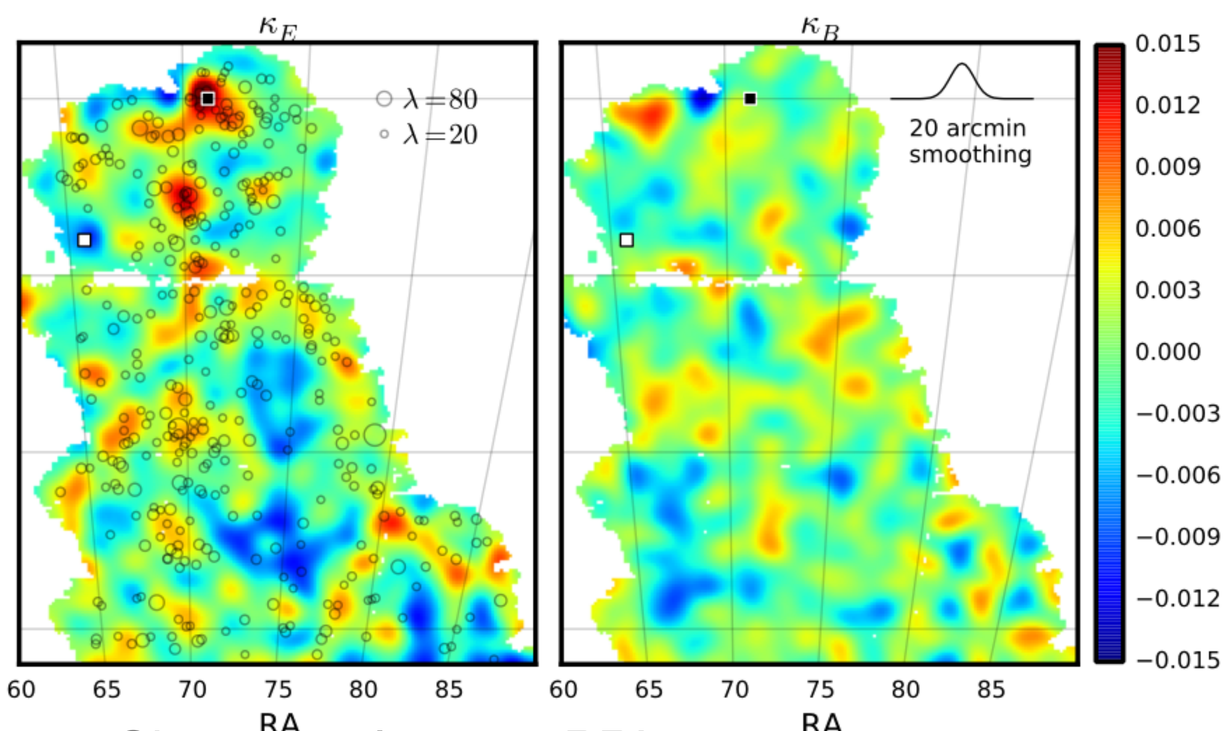
Mass Mapping

Van Waerbeke et al. 2013, MNRAS 433, 3373

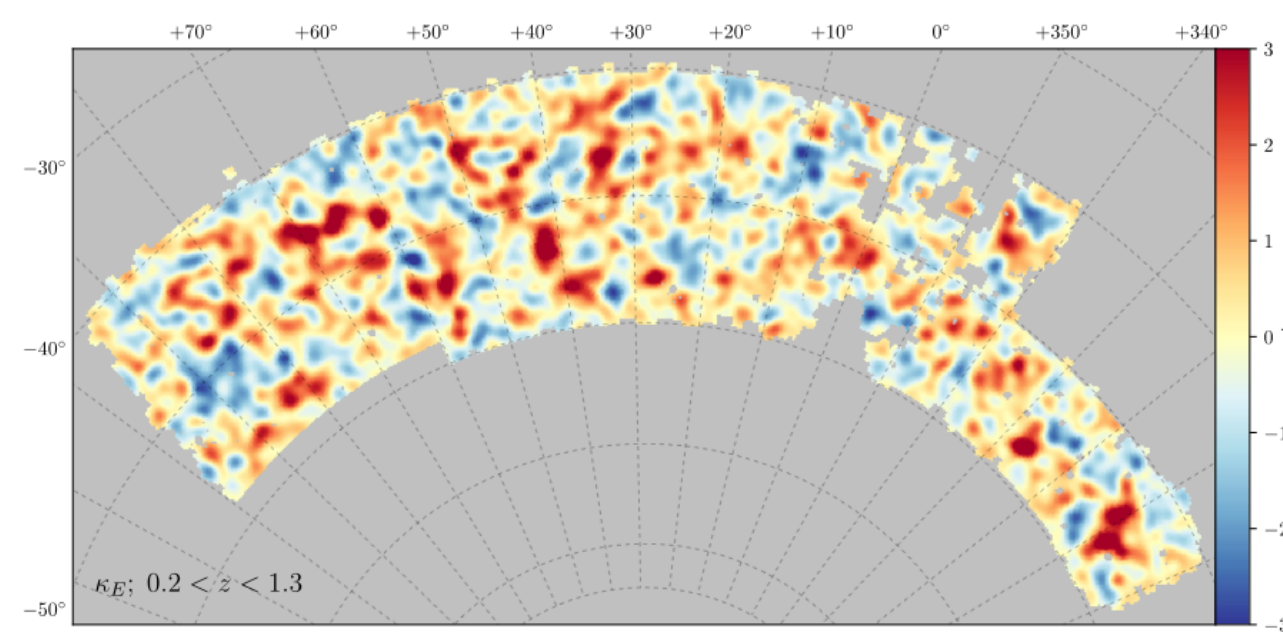


Oguri et al. 2018, PASJ 70, S26





Chang et al. 2015, PRL 115, 051301



Chang et al. 2018, MNRAS 475, 3165

Bullet Cluster

The main signature of a merger between clusters is the **dissociation** of the **intracluster gas** from the **dark matter** and the galaxies,

Galaxies interact principally via the tidal gravitational fields

➡ essentially pass through each other.

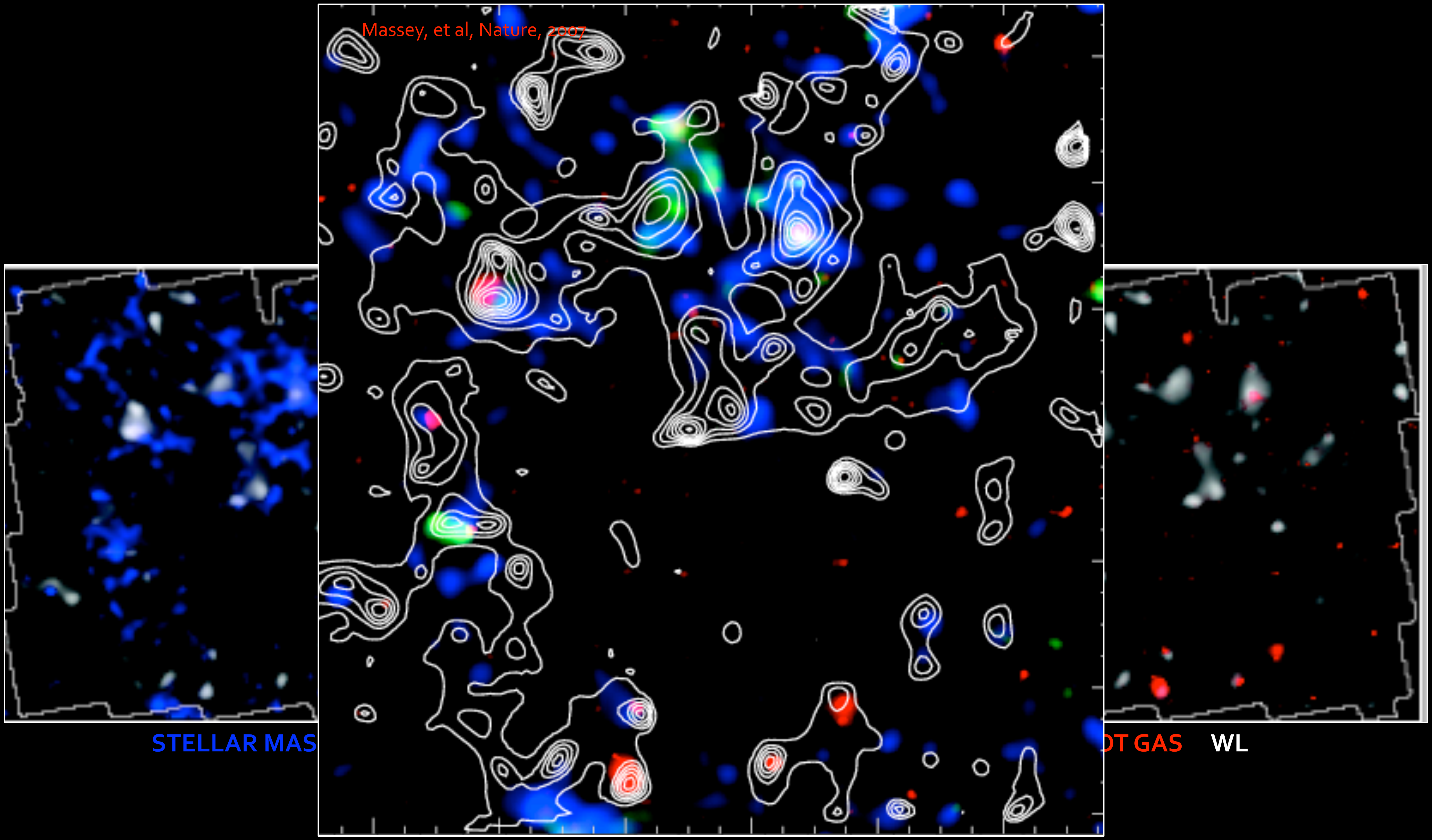
Ionized intracluster plasma clouds have a ram pressure

➡ slows them down during crossing.

➡ leaves an overdensity of X-ray emitting gas between the luminous subclusters along the merger axis.

Clear separation of dark matter and gas clouds is considered as a **direct evidence** that dark matter exists

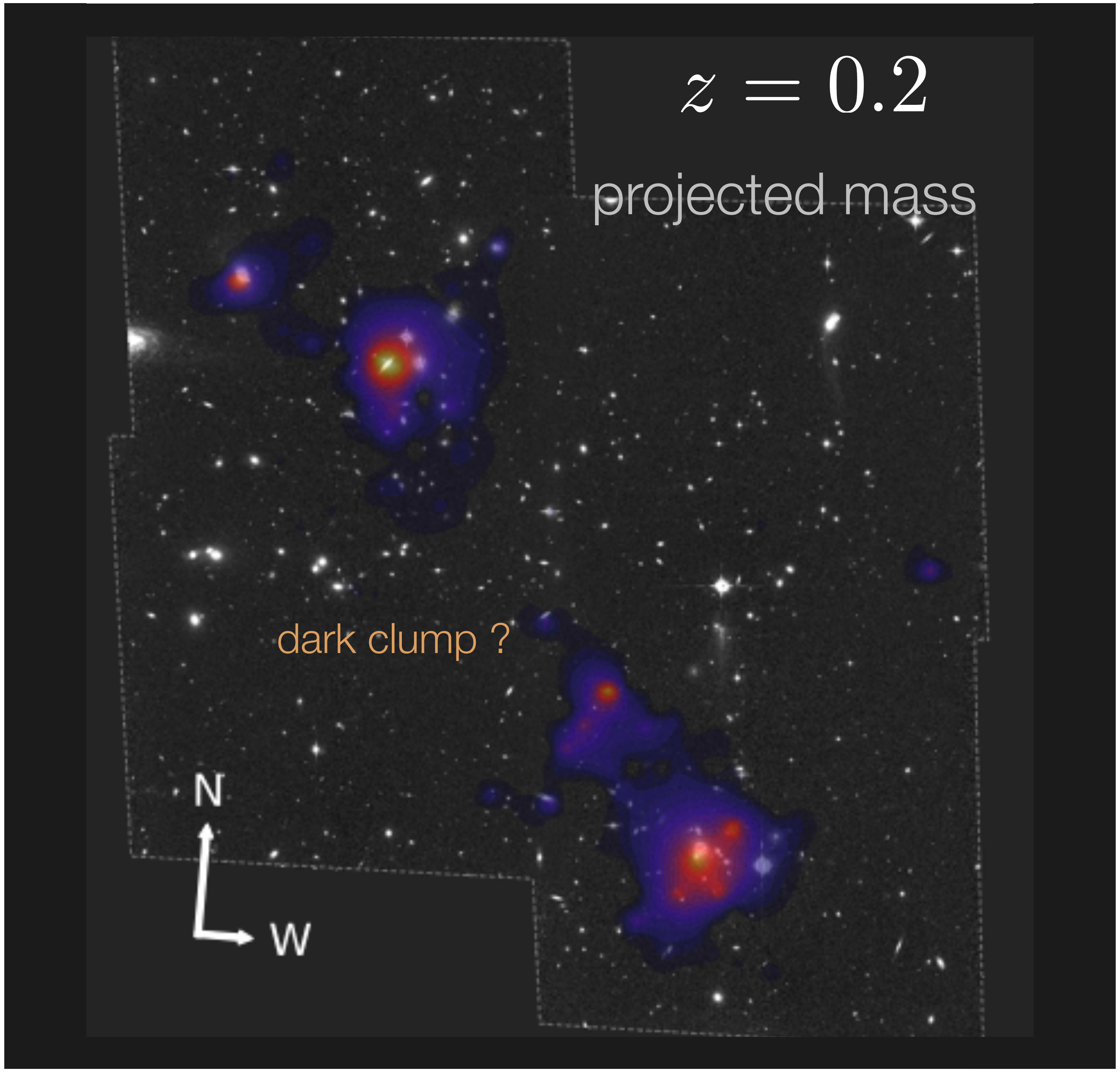
Spatial Correlation with other probes



Cluster A520: the puzzling dark core

The Abell 520 system (MS 0451+02, z = 0.2, Abell et al. (1989)) exhibits complex structure and offers a possible c**ounterexample to the collisionless dark matter scenario**.

Detection of a **dark core (**Mahdavi et al, 2007), labeled as P3, using data from the Canada-France-Hawaii Telescope(CFHT) and Subaru, which coïncides spatially with the peak of the X-ray emission.

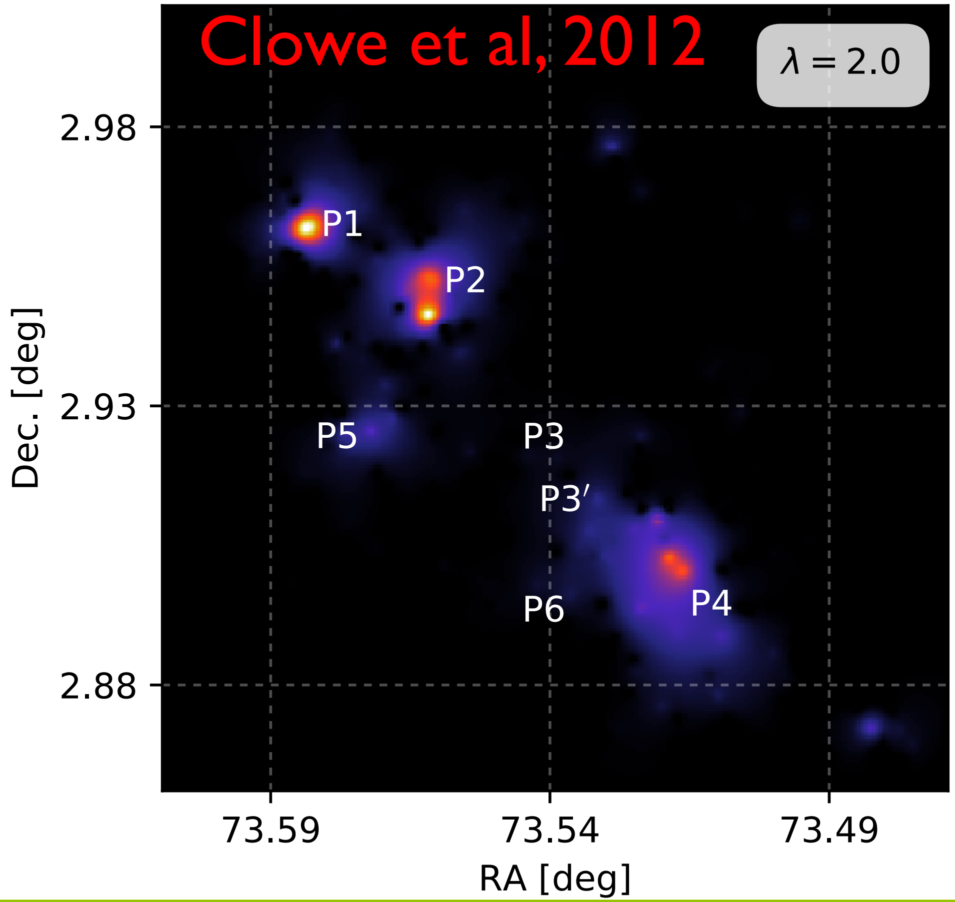


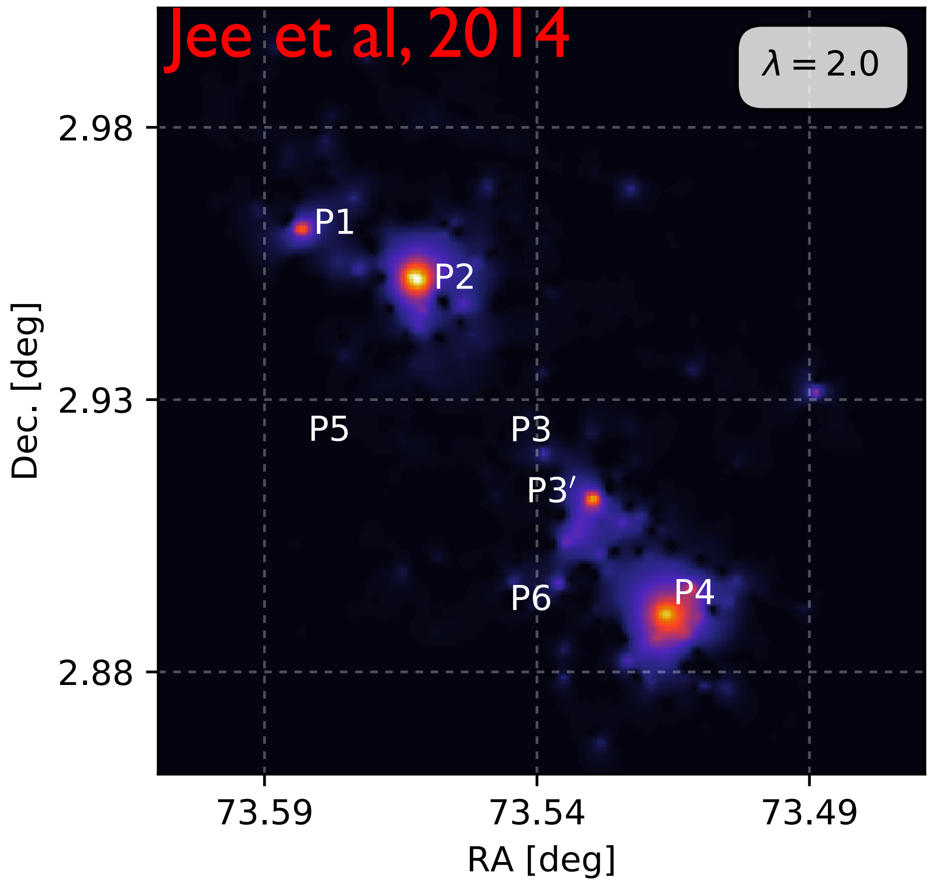
A520: the puzzling dark core

Dark matter particles possessing an **appreciable self-interaction cross-section** has been suggested as a possible explanation.

- ➡ **Confirmed** in Jee, Mahdavi, Hoekstra; 2012) using HST Wide Field Planetary Camera 2 (WFPC2) => 10 sigma détection
- ➡ **Confirmed** in Jee et al. (2014) , confirms a very high mass-to-light ratio, but not exactly at the same location. P3′ was shifted from the former by about 1 arcmin southwest toward the largest mass sub- structure P4.

Clowe et al. (2012) does **NOT** confirmed it, using ground-based Magellan observations combined with mosaic images from the HST Advanced Camera for Surveys (ACS).





Space & Cosmology

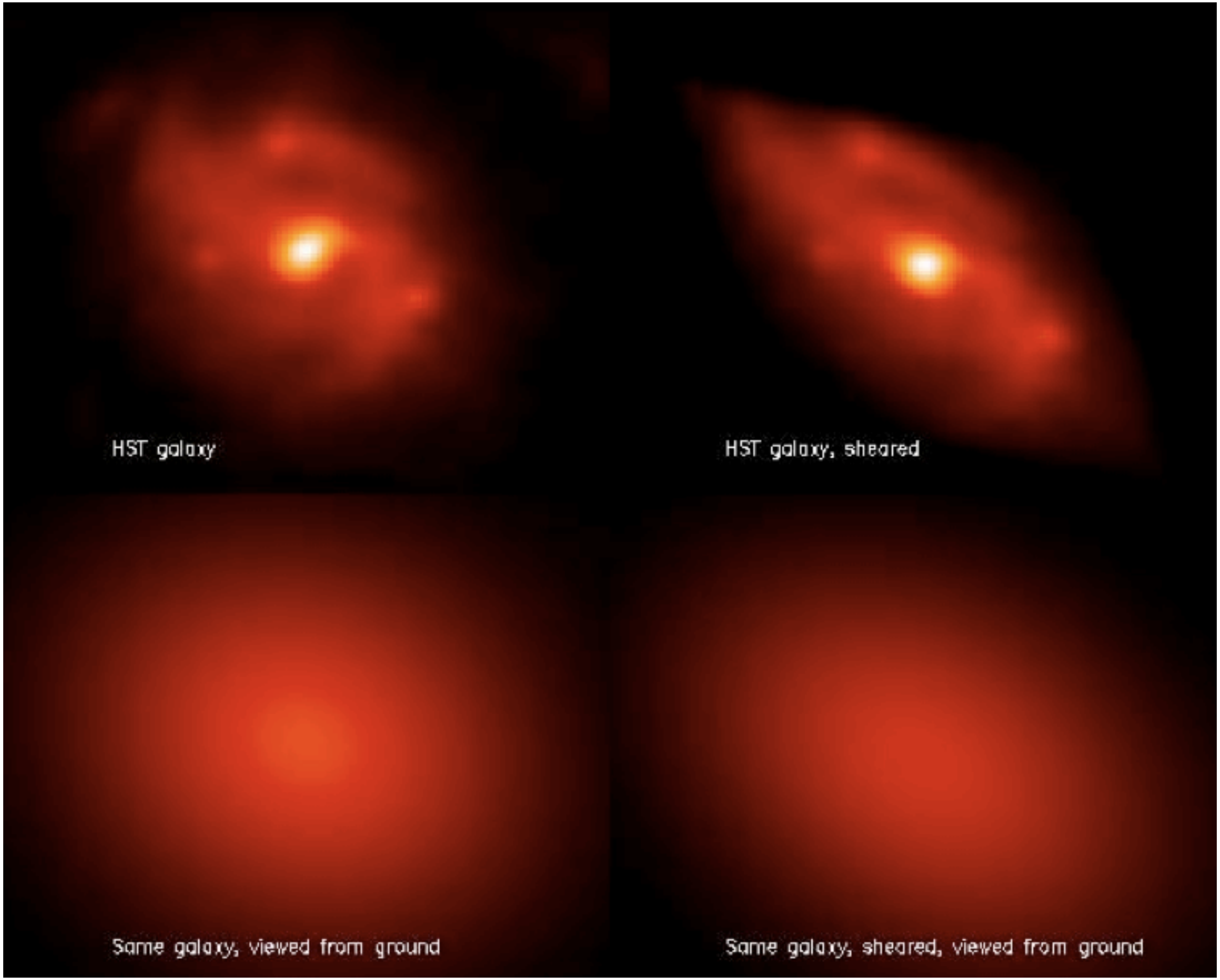
➡Part 1: Introduction to Accurate Space Cosmology

➡Part 2: Inverse Problems

➡**Part 3: Euclid Weak Lensing**

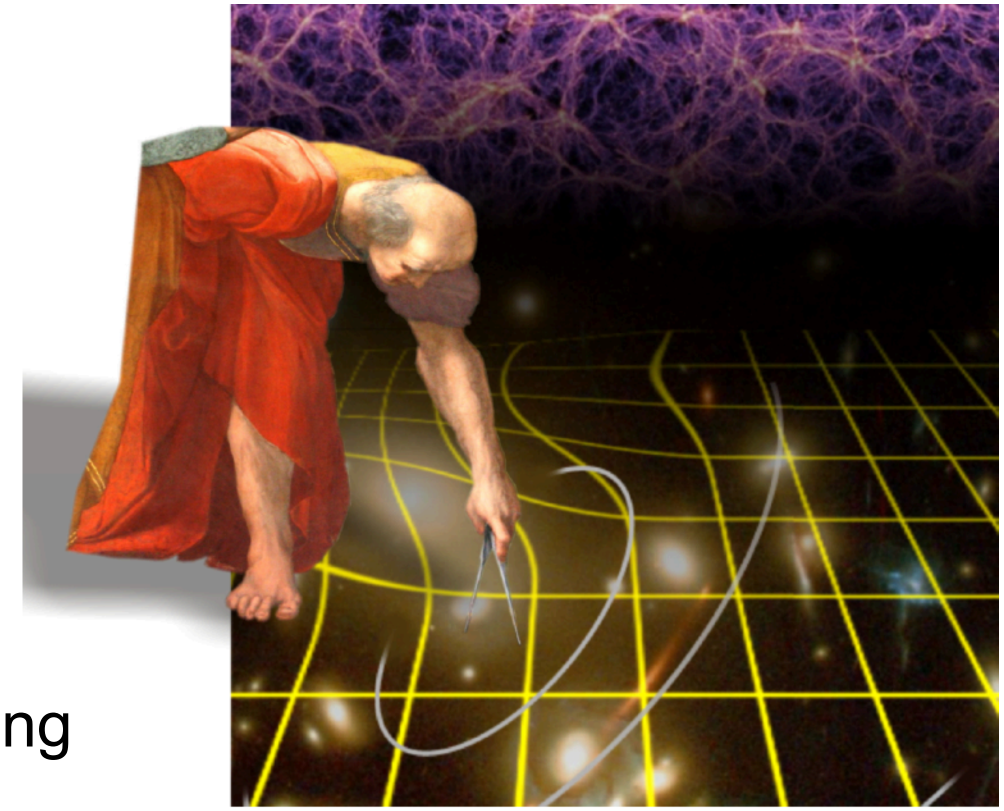
- ➡Introduction to Weak Lensing
- ➡**The Euclid space projet and its mathematical challenges**
- ➡Advanced methodologies for Euclid

Motivation for spatial observations



Euclid ESA Space Mission

Understand the origin of the Universe's accelerating expansion:



→probe the properties and nature of *dark energy, dark matter, gravity* and distinguish their effects decisively

→by tracking their observational signatures on the

- geometry of the universe:
  Weak Lensing + Galaxy Clustering

- cosmic history of structure formation: WL, z-space distortion, clusters of galaxies

→ **Controling systematic residuals to an unprecedented level of accuracy, that cannot be reached by any other competing missions/telescopes**

Gains in space:

Stable data: homogeneous data set over the whole sky

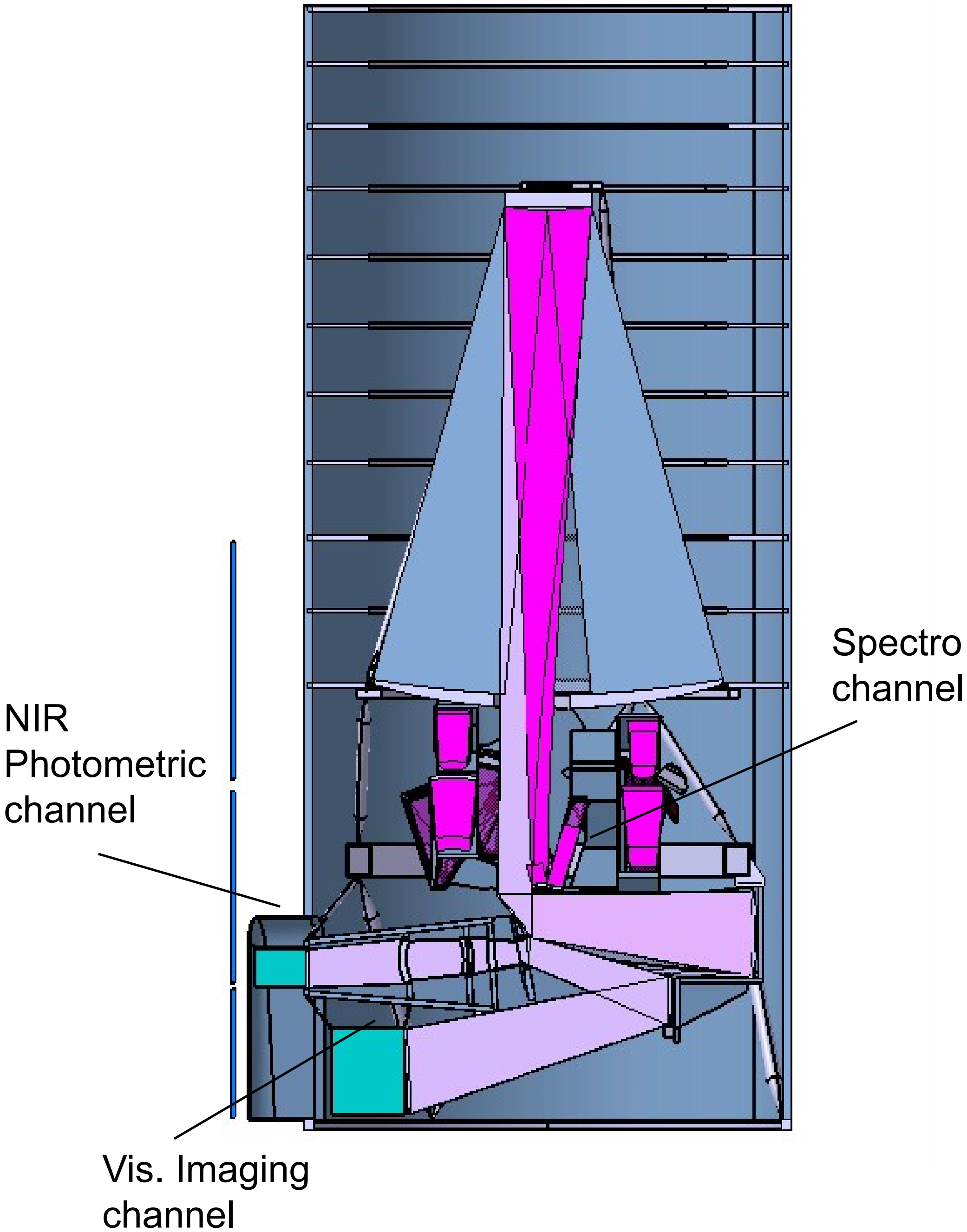
→**Systematics** are small, understood and controlled

→Homogeneity : Selection function perfectly controlled

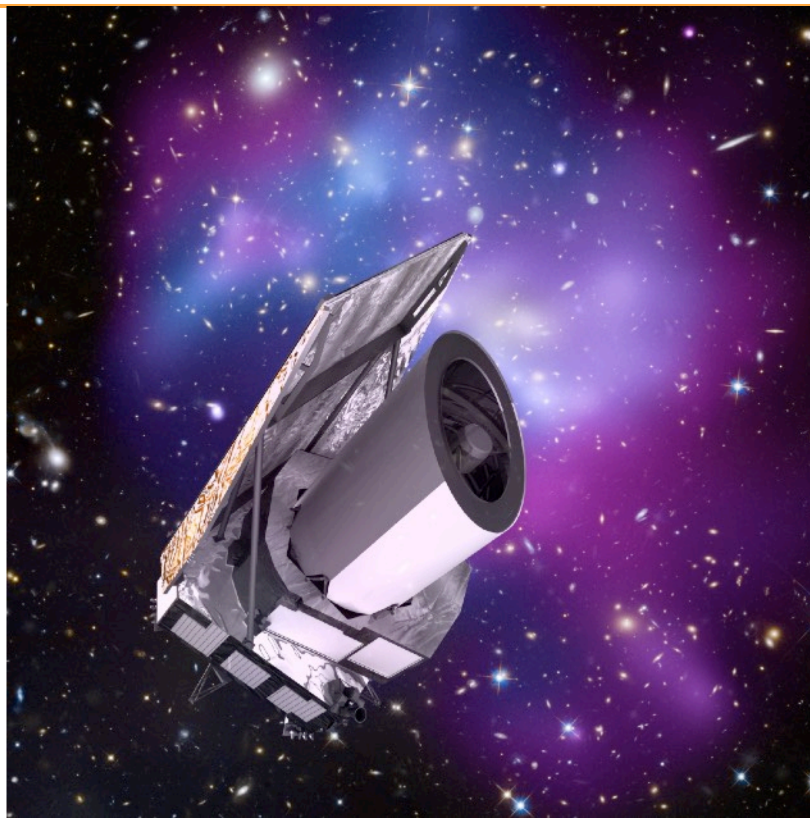
Euclid mission element

Euclid Consortium

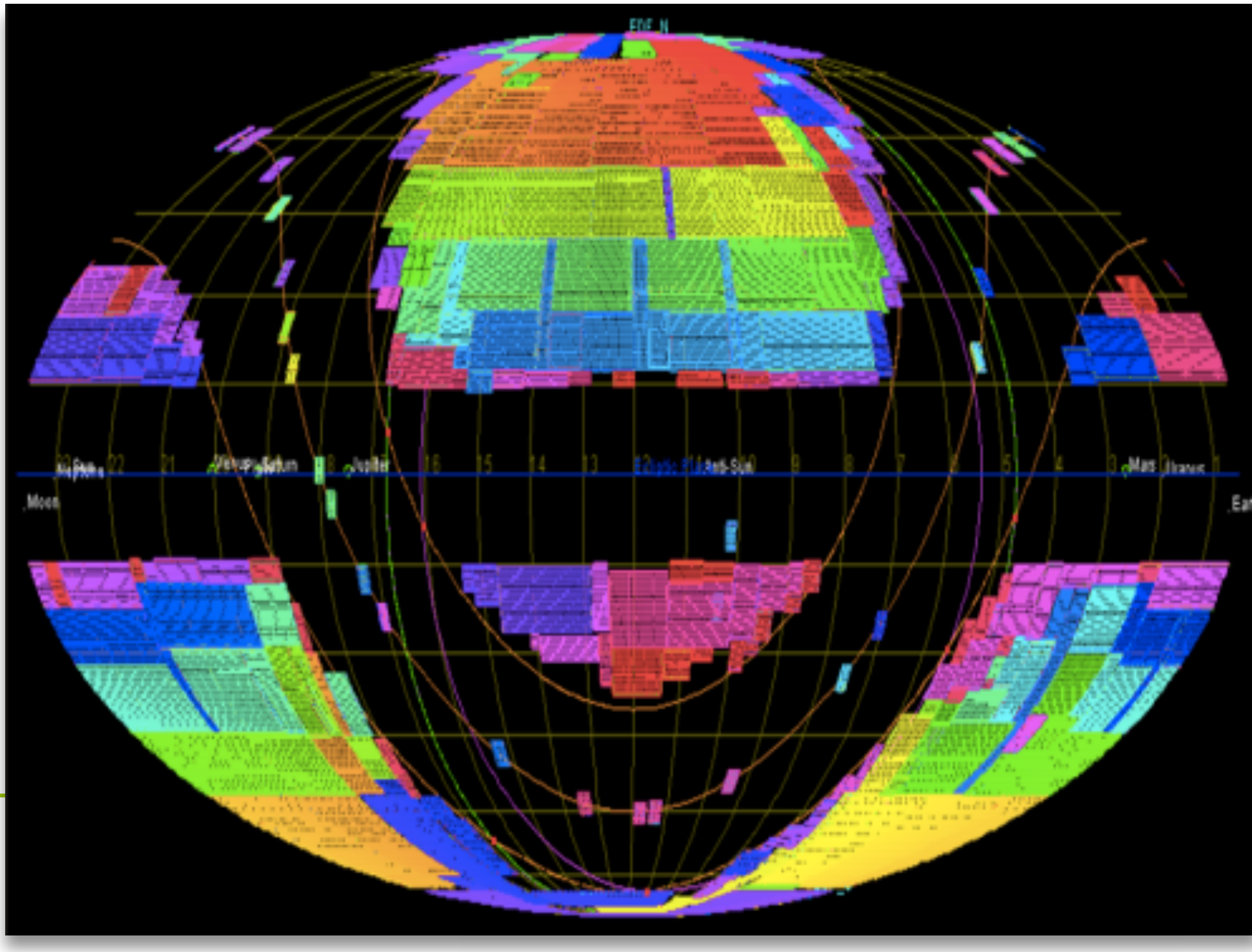
- Launch Soyuz, in 2023, L2 Orbit
- 6 years mission
- Telescope: 1.2 m
- Instruments:
- **VIS**: Visible imaging channel:
  - 0.54 deg^2, 0.10'' pixels, 0.16'' PSF FWHM,
  - 1 broad band R+I+Z (0.55-0.92mu),
  - 36 CCD detectors, galaxy shapes
- **NISP**: NIR photometry channel:
  - 0.54 deg^2, 0.3'' pixels,
  - 3 bands Y,J,H (1.0-1.7mu),
  - 16 HgCdTe detectors, photo-z's
- **NISP**: NIR Spectroscopic channel:
  - 0.54 deg^2,
  - R(mean)=250,
  - 0.9-1.7mu, slitless, spectro redshifts



Euclid ESA Space Mission

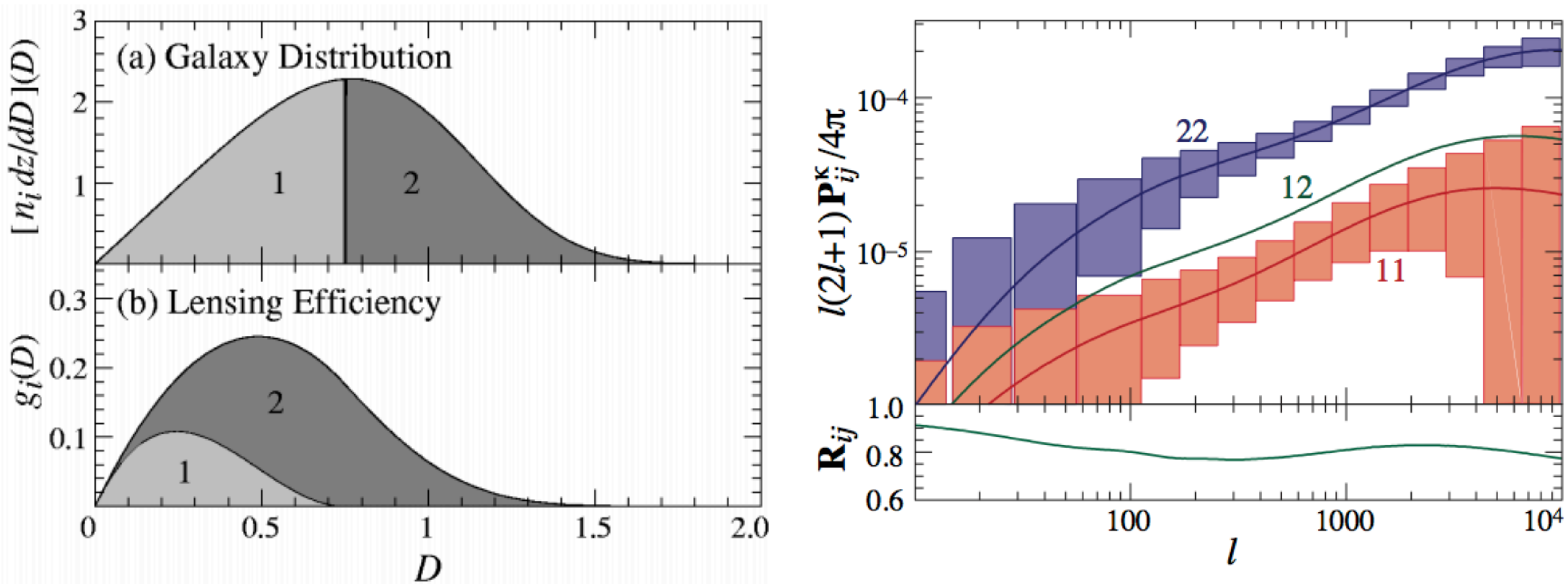


Euclid Mission
• Medium-class mission in ESA's Cosmic Vision program
• 6 year survey, launch in Q4 2020

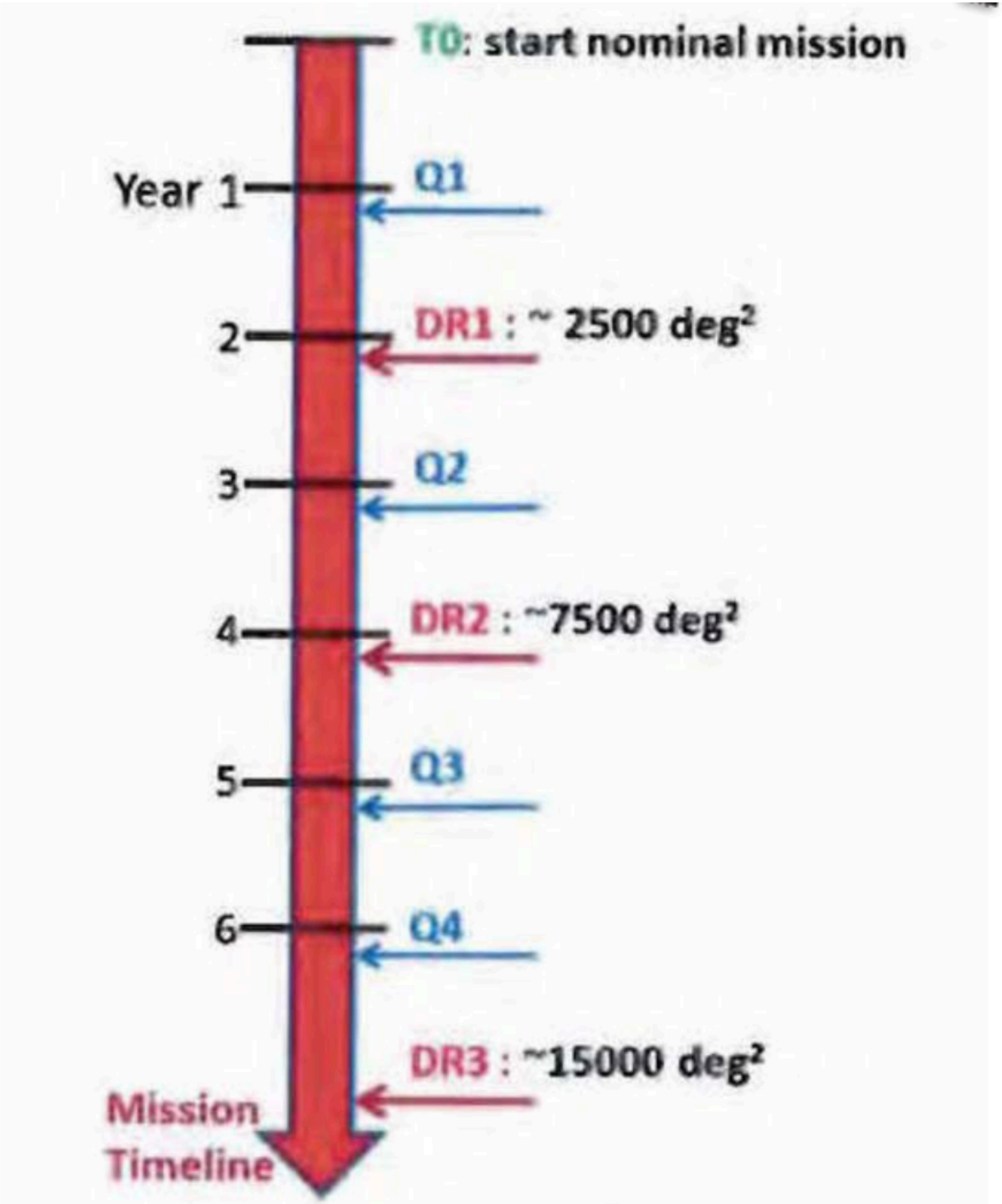


Optimized for weak lensing (and galaxy clustering)
— **15,000 deg2** survey area
— over **1.5 billion galaxies**
— redshifts out to **z = 2**

Tomographic Weak Lensing



The power spectra of two slices, their cross power spectrum, and their correlation coefficient (Hu, ApJ, 1999).



9

Allemagne

gen - Pays-Bas

e

TAT

Euclid Figure of Merit

Euclid Consortium

	Modified Gravity	Dark Matter	Initial Conditions	Dark Energy		
Parameter	γ	m_ν/eV	f_{NL}	w_p	w_a	*FoM*
Euclid Primary (WL+GC)	0.010	0.027	5.5	0.015	0.150	430
Euclid All (WL+GC)+CL+ISW	0.009	0.020	2.0	0.013	0.048	1540
Euclid+Planck (Euclid All)	0.007	0.019	2.0	0.007	0.035	4020
Current	0.200	0.580	100	0.100	1.500	~10
Improvement Factor	**30**	**30**	**50**	**>10**	**>50**	**>300**

High Accuracy Requirements

$$\gamma = \gamma_1 + i\gamma_2 = |\gamma| e^{i2\Phi}$$

$$\tilde{\gamma} = (1+m)\gamma + c$$

Multiplicative bias (→ m)

Additive bias (→ c)

$$m = m_1 + i m_2 = |m| e^{i2\Phi_m}$$

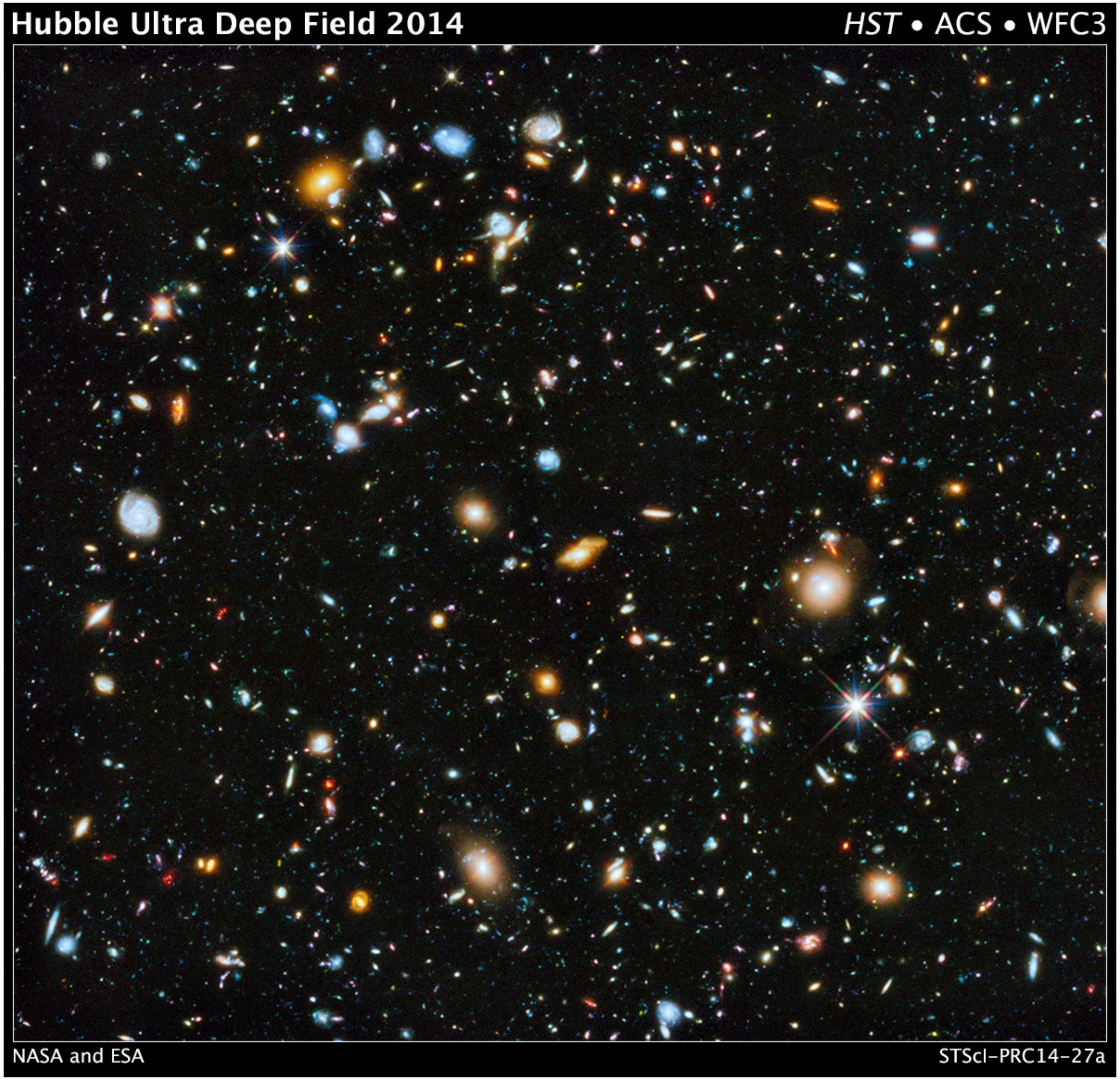
$$c = c_1 + i c_2 = |c| e^{i2\Phi_c}$$

Requirements for the ESA Euclid Space Telescope

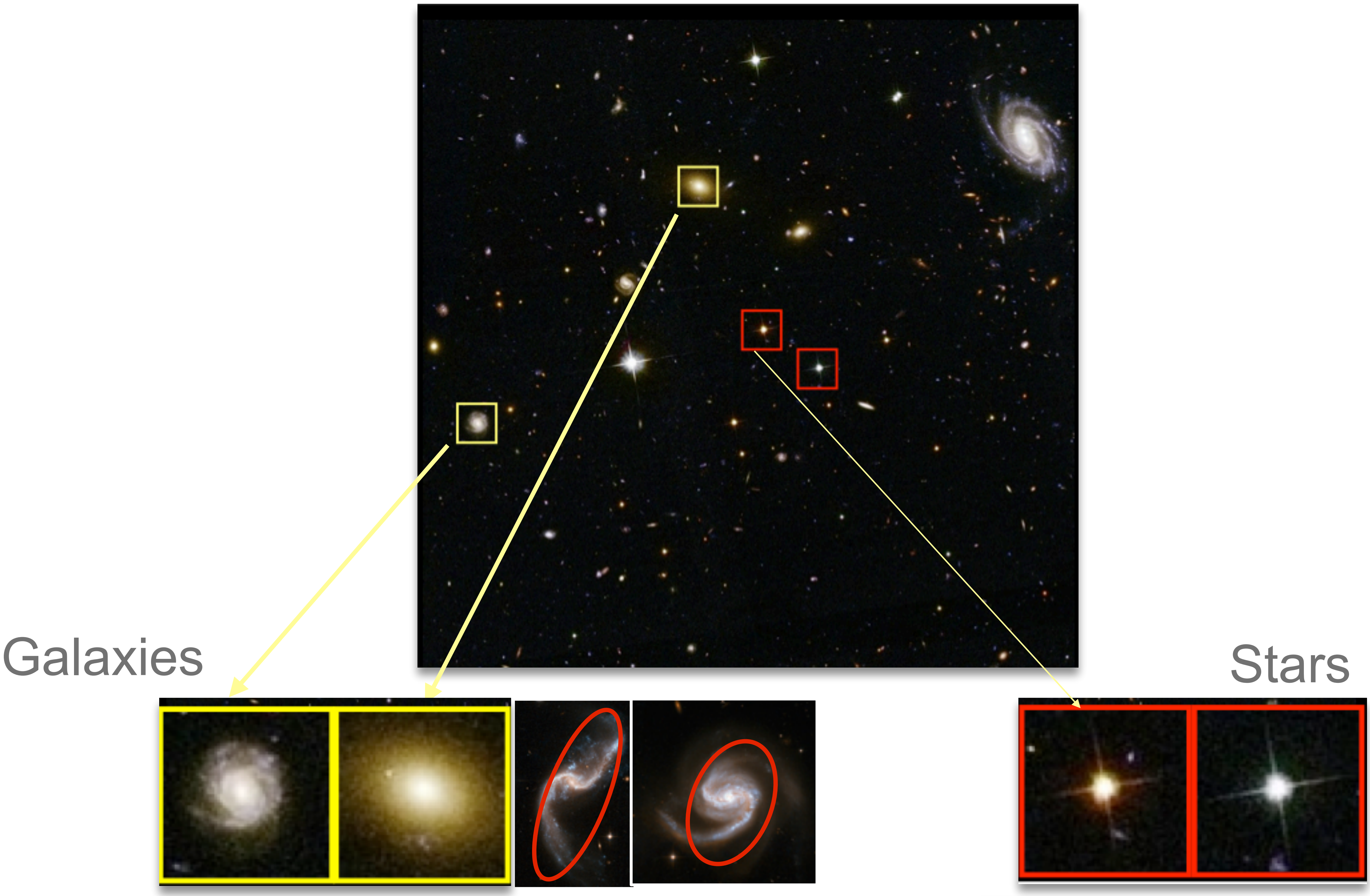
$$\sigma_m < 2 \times 10^{-3}$$

$$\sigma_c < 5 \times 10^{-4}$$

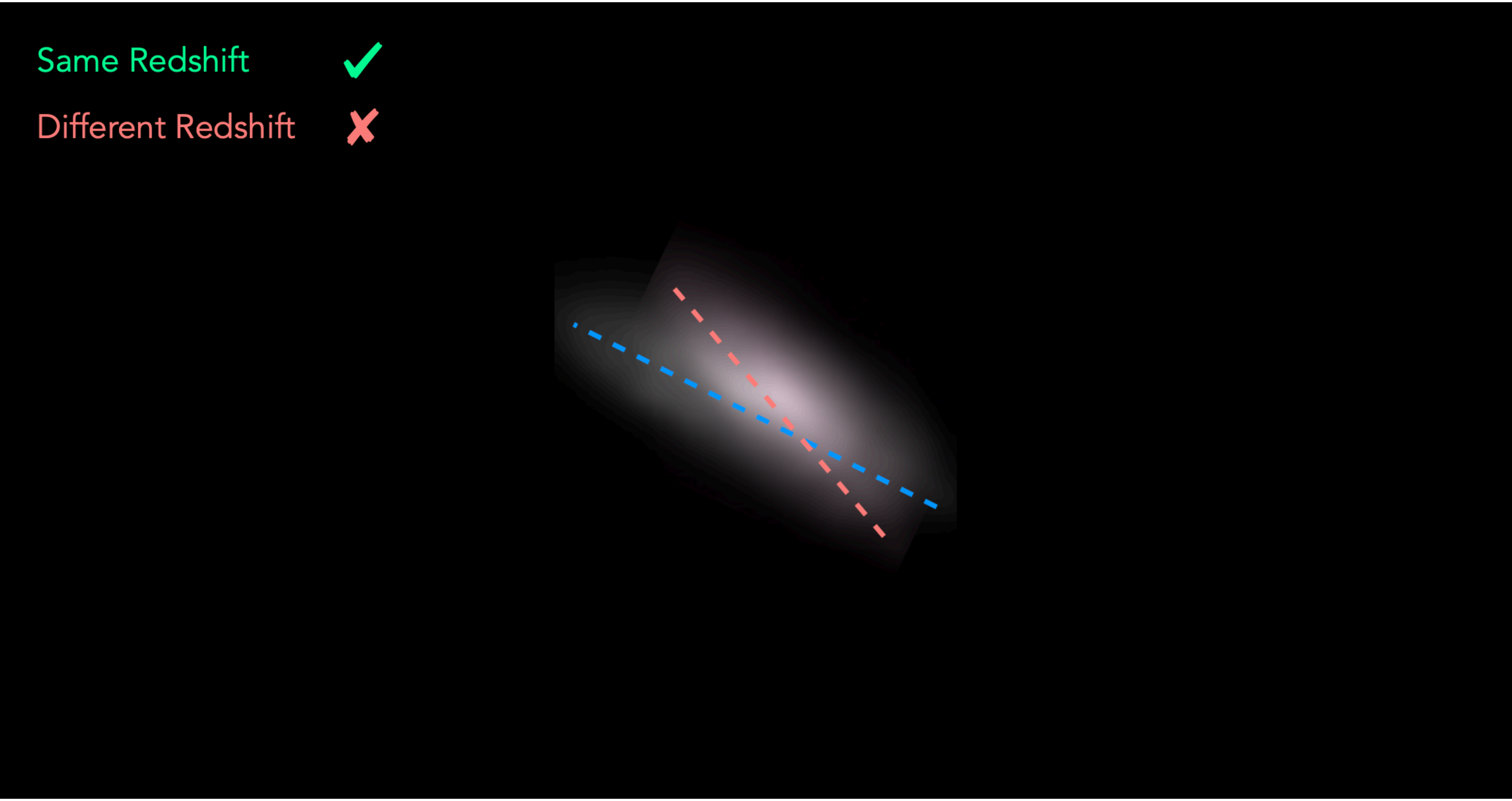
Galaxies


Hubble Ultra Deep Field 2014
HST • ACS • WFC3
NASA and ESA
STScI-PRC14-27a

Detection + Classification stars/galaxies



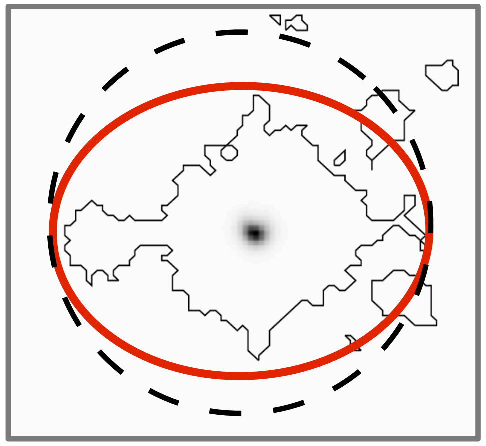
Problem: Blended Objects



Shape Measurements

Galaxies are convolved by an **asymetric** PSF

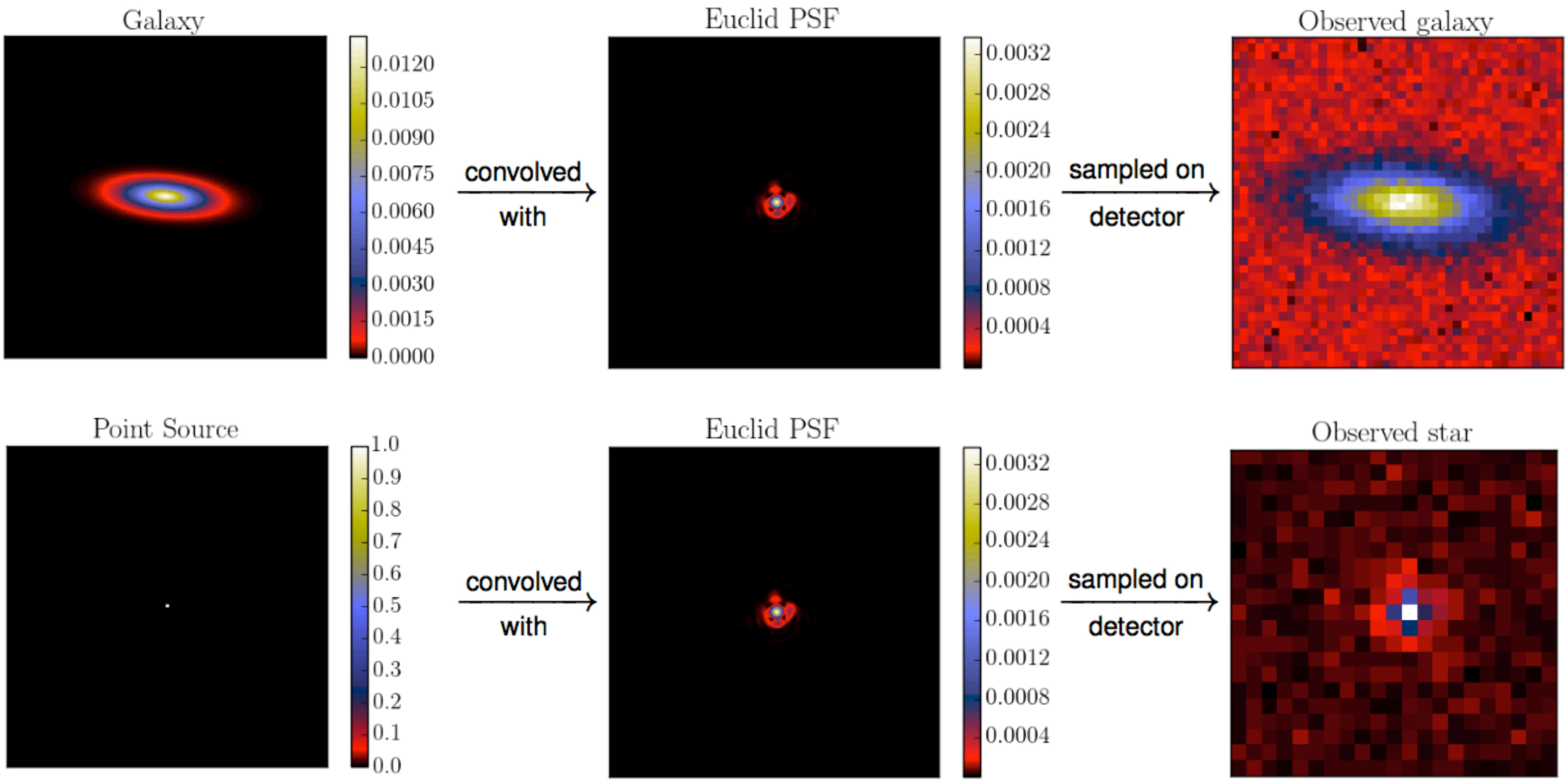
\+ Images are **undersampled**



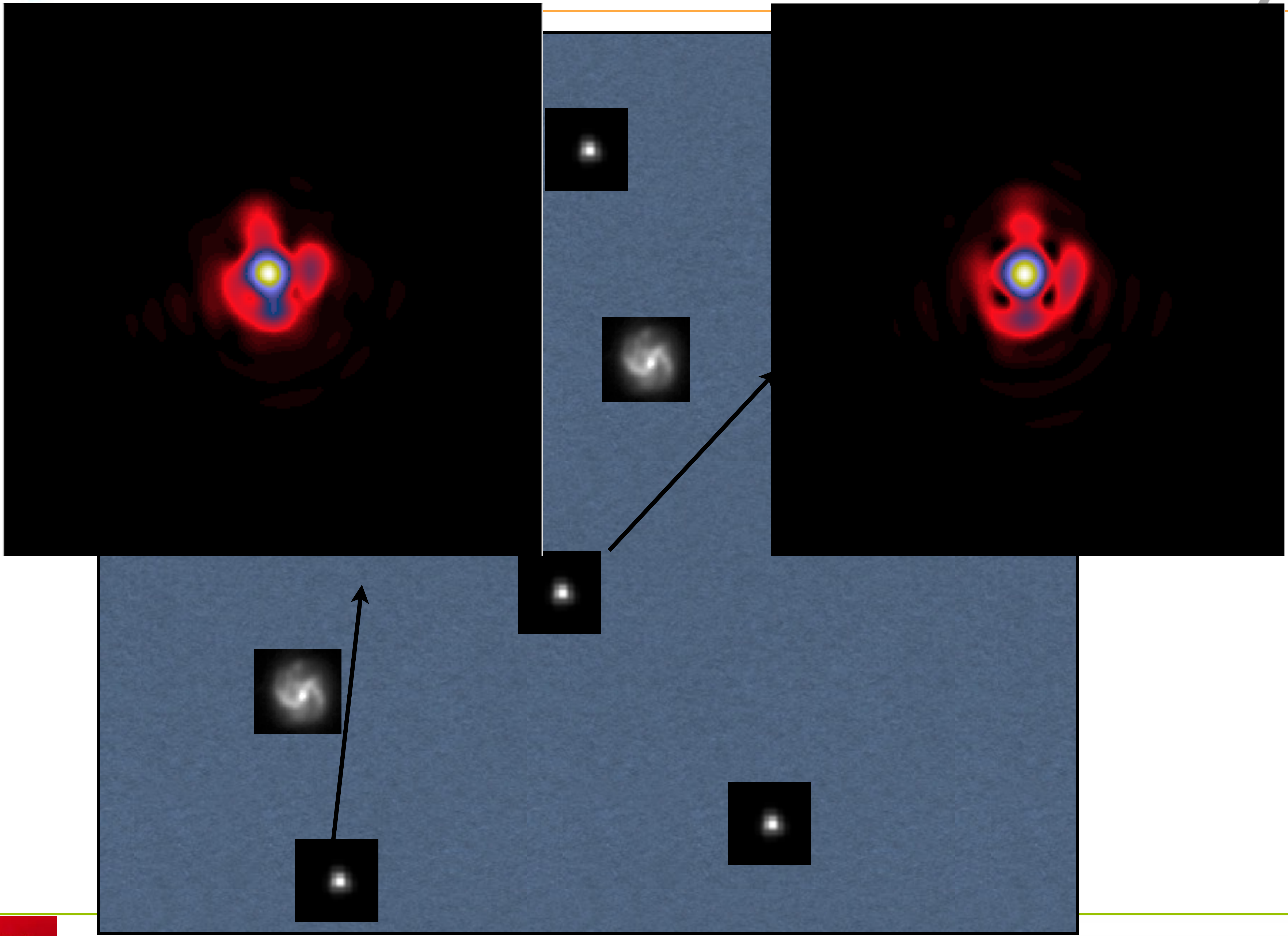
Shape measurements must be deconvolved

Methods: Moments (KSB), Shapelets, **Forward-Fitting**, Bayesian estimation, metacal, etc

Convolution Operator + Sampling

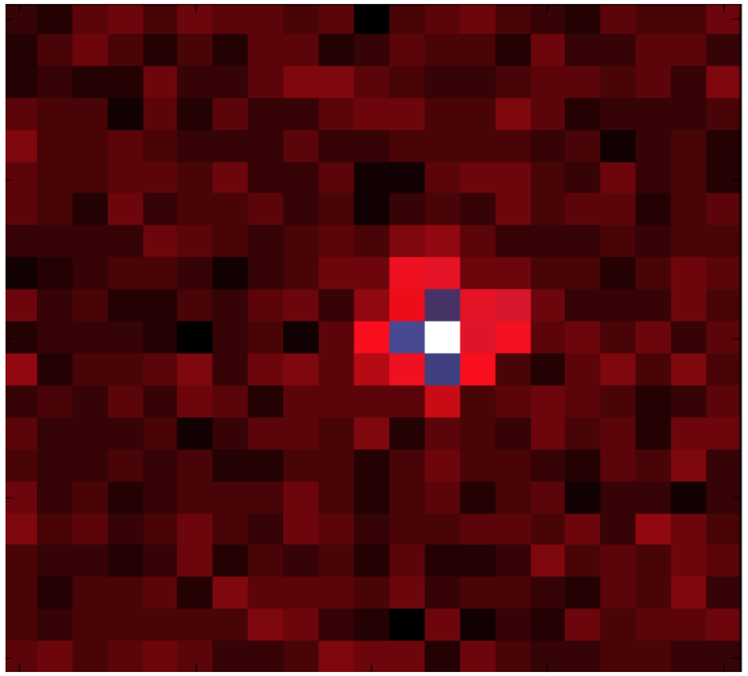


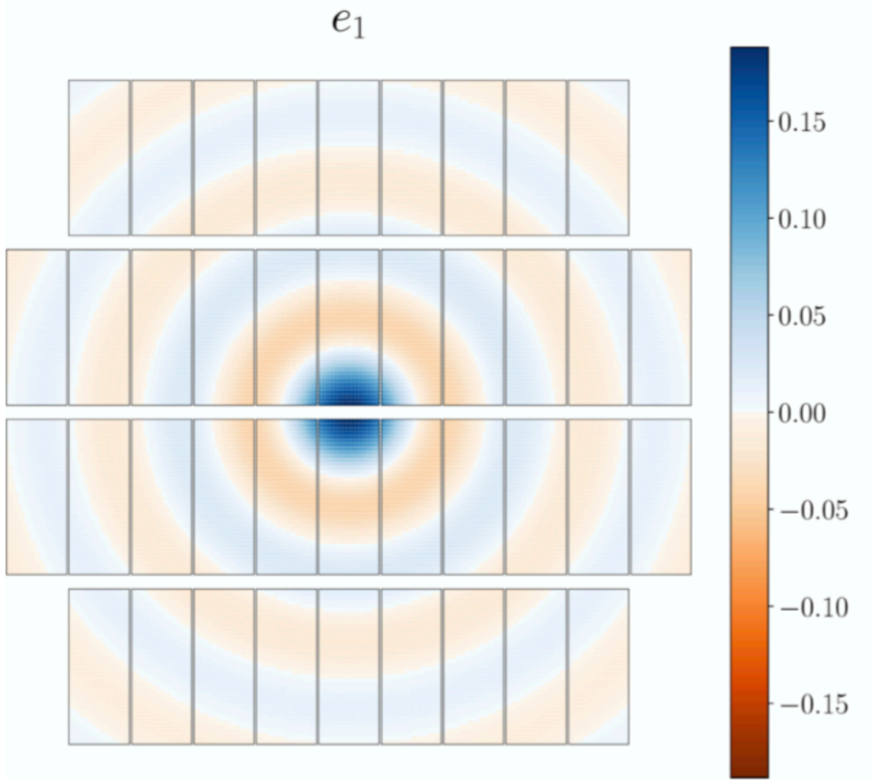
Space Variant PSF



Euclid PSF Modeling

- PSF Modeling
  - ➡Undersampling
  - ➡Space dependency
  - ➡ Time dependency
  - ➡Wavelength dependency
  - ➡Multi-CCD





State of the art: PSF Modeling

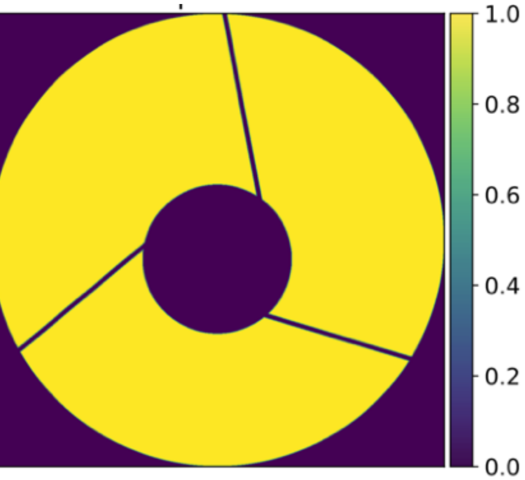
1. **Forward Modelling** approach (FM) leaded by Lance Miller
   - Model the exit pupil using pre-defined Zemax modes.
   - Fit to all stars.

➡ PRO: No other **existing** method can achieve the very strong requirement on the PSF.

➡ CON: But hard to validate since i) the simulations are done with the same model, and ii) the model may change (vibrations at launch).

2. Need a **data driven** approach
   - Validate the FM solution on real data.
   - All surveys ((KIDS, CFHTLenS, DES) have used the data driven approach.
   - HST: TinyTim HST modelling software (Krist 1995) for the Hubble Space Telescope not as good as a data driven approach (Jee et al, PASP, 2007; Hoffmann and Anderson, Instrument Science Report ACS 2017-8, 2017).
   - Combination of both approaches could lead to the optimal solution.

PSF Modeling



$$\hat{\mathcal{H}}^{\lambda}_{u_i} \propto \mathcal{D}_{\theta} \left\{ P^{\lambda}_{u_i} \odot \exp\left[2\pi i \left(\frac{\Phi_{u_i}}{\lambda} \right) \right] \right\}$$

Monochromatic PSF reconstruction

Diffraction operator + Downsampling

Obscurations

WFE
Parametric Component

Star observation

$$\bar{\mathcal{H}}_{u_i} \approx \sum_{b=1}^{N_\lambda} S_{u_i}(\lambda_b)\, \mathcal{H}^{\lambda_b}_{u_i}$$

- Star considered as point source
- Discretization of the integral into N wavelength bins
- **Star SED information known**

CosmoStat

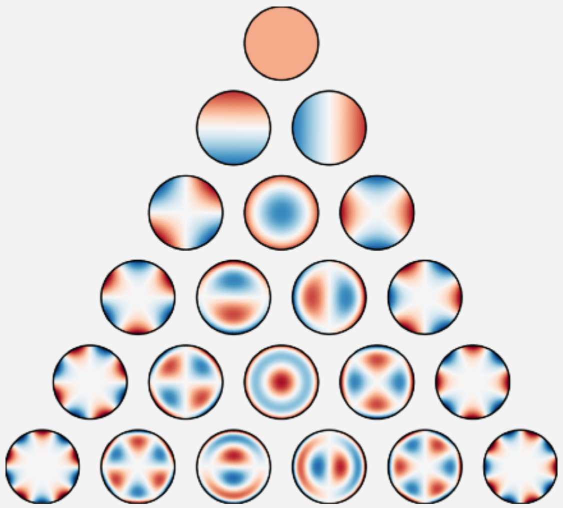
WFE Parametric Component

Model: $$\hat{\mathcal{H}}^{\lambda}_{u_i} \propto \mathcal{D}_{\theta}\left\{ P^{\lambda}_{u_i} \odot \exp\left[2\pi i \left(\frac{\Phi_{u_i}}{\lambda} + C^{\lambda}_{u_i}\right)\right]\right\}$$

$$\Phi_{u_i}[x,y] = \sum_{j=1}^{N_Z} a_{j,u_i} Z_l[x,y]$$

$$\text{e.g. } a_{j,u_i} = c^j_0 + c^j_1 u_i[0] + c^j_2 u_i[1]$$

Zernike polynomials



- Based on Zernike polynomials up to mode Nz.
- Polynomial spatial variation of Zernike coefficients.
- Chromatic variations follow the 1/lambda dependence of diffraction.
- Small number of parameters to represent all the variability.

State of the art: PSF Modeling

1. **Forward Modelling** approach (FM) leaded by Lance Miller
   - Model the exit pupil using pre-defined Zemax modes.
   - Fit to all stars.

➡ PRO: No other **existing** method can achieve the very strong requirement on the PSF.

➡ CON: But hard to validate since i) the simulations are done with the same model, and ii) the model may change (vibrations at launch).

2. Need a **data driven** approach
   - Validate the FM solution on real data.
   - All surveys ((KIDS, CFHTLenS, DES) have used the data driven approach.
   - HST: TinyTim HST modelling software (Krist 1995) for the Hubble Space Telescope not as good as a data driven approach (Jee et al, PASP, 2007; Hoffmann and Anderson, Instrument Science Report ACS 2017-8, 2017).
   - Combination of both approaches could lead to the optimal solution.

Wavefront Error Modeling

- The model is now build on the wavefront error (WFE) and not on the pixel images.

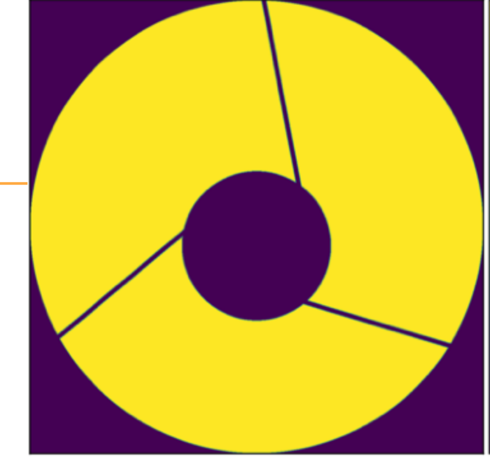
WFE = Parametric Part + Non Parametric Part

- The non parametric to correct **the mismatch** between the model and the truth.
- Easier include the **dichroic coating effect.**
- **RCA regularization** technique on the Non Parametric Part.

GPU + TensorFlow automatic differentiation

- Uses a forward model, WFE → pixels
    - Includes diffraction phenomena, obscuration, downsampling, etc..
- **End-to-end differentiable!**
    - Based on an automatic differentiation framework → TensorFlow.
    - Fast computations on GPU.

WF-RCA Modeling



Forward model: $$\hat{\mathcal{H}}^{\lambda}_{u_i} \propto \mathcal{D}_{\theta} \left\{ P^{\lambda}_{u_i} \odot \exp\left[2\pi i \left(\frac{\Phi_{u_i}}{\lambda} + C^{\lambda}_{u_i} \right) \right] \right\}$$

Monochromatic PSF reconstruction

Diffraction operator + Downsampling

Obscurations

WFE
Parametric part

WFE
Non-parametric part

Wavefront PSF model

Star observation

$$\bar{\mathcal{H}}_{u_i} \approx \sum_{b=1}^{N_\lambda} S_{u_i}(\lambda_b)\, \mathcal{H}^{\lambda_b}_{u_i}$$

- Star considered as point source
- Discretization of the integral into N wavelength bins
- Star SED information known

WFE Non-parametric Part

Model: $$\hat{\mathcal{H}}_{u_i}^{\lambda} \propto \mathcal{D}_{\theta}\left\{ P_{u_i}^{\lambda} \odot \exp\left[2\pi i \left(\frac{\Phi_{u_i}}{\lambda} + C_{u_i}^{\lambda} \right) \right] \right\}$$

- Factorisation scheme.

$$C_{u_i}^{\lambda} = \frac{1}{\lambda} \sum_{i_1=1}^{r} (\Pi_{u_i} \alpha)_{i_1} S_{i_1, i_2, i_3}$$

Number of PSF features

Fixed spatial variations dictionary

Sparse weights

Tensor of wavefront PSF features

- Sparsity constraint on alpha weights.
- Spatial variations based on the graph Laplacian eigenvectors as in RCA [Schmitz et al 2020].
- Diffraction based wavelength dependence.
  - We can add more sophisticated wavelength dependencies (dichroic coating).

Numerical experiment: Set-up

We compare two approaches:

- Fitting the Parametric model
  - Using **45 modes Zernike polynomial** (***i.e. the true underlying model***)
  - Coefficients vary 2D polynomial of degree 2 of FoV positions.
  - **Same model** as the one used to generate the data.

- Using the semi-parametric PSF model (WF-RCA)
  - Parametric part: **15 modes Zernike** polynomial with degree 2 polynomial variations.
    i.e. the parametric part is only an approximation of the underlying true model.
  - Non-parametric part with 20 PSF features.

- The SED of the observed and test stars is given as input.

Incomplete parametric model in the proposed WF-RCA.

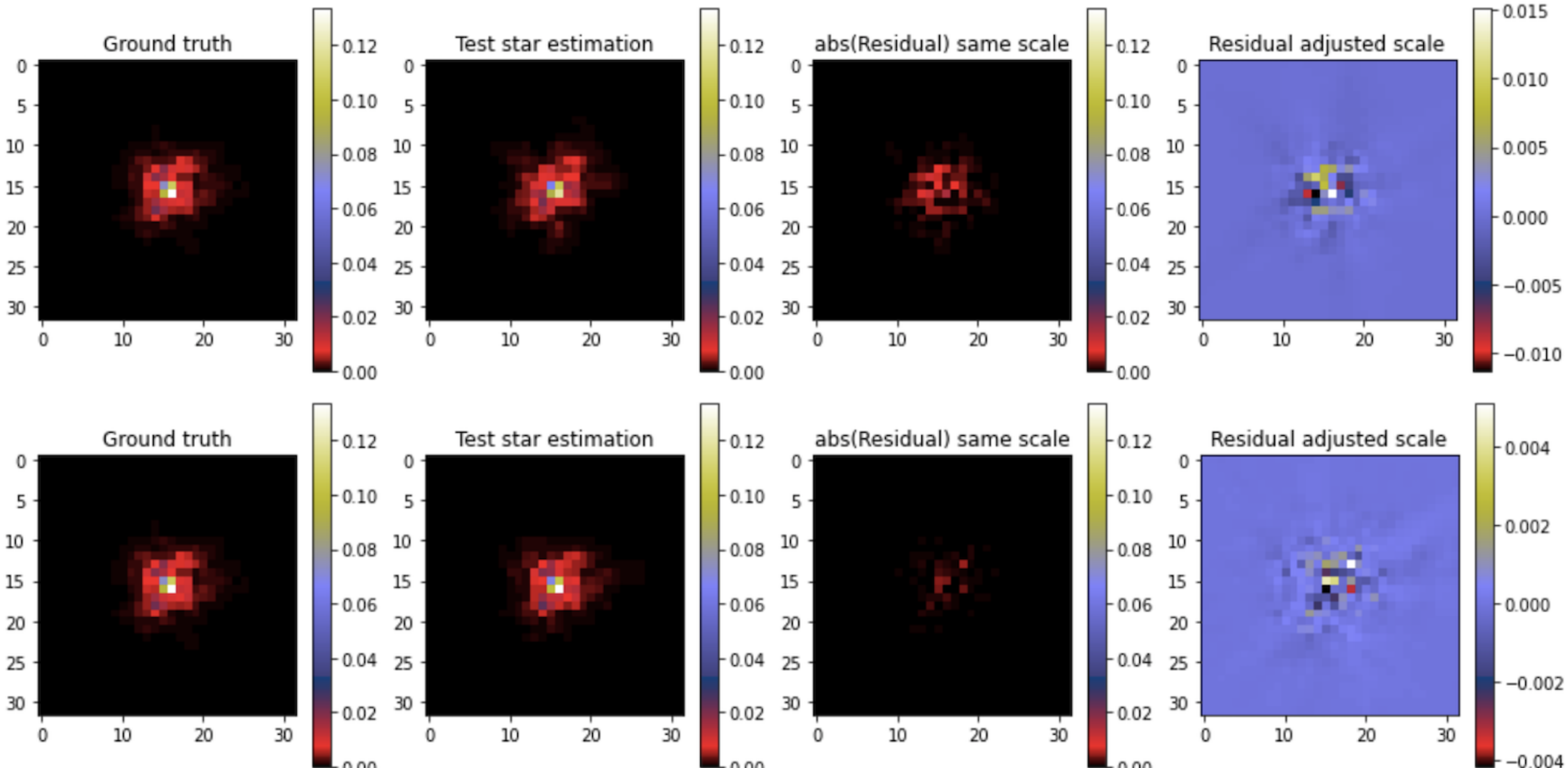
Experiment objective

- Estimate the test stars (30%) while using the observed stars (70%) to train the model.

Numerical experiment: Results

Pixel errors with respect to noiseless ground truth polychromatic PSF at Euclid resolution.

Model	Train stars RMSE	Test stars RMSE
Parametric	1.1087e-03 (14.16%)	1.0950e-03 (14.45%)
WF-RCA	2.4031e-04 (3.07%)	2.7927e-04 (3.68%)



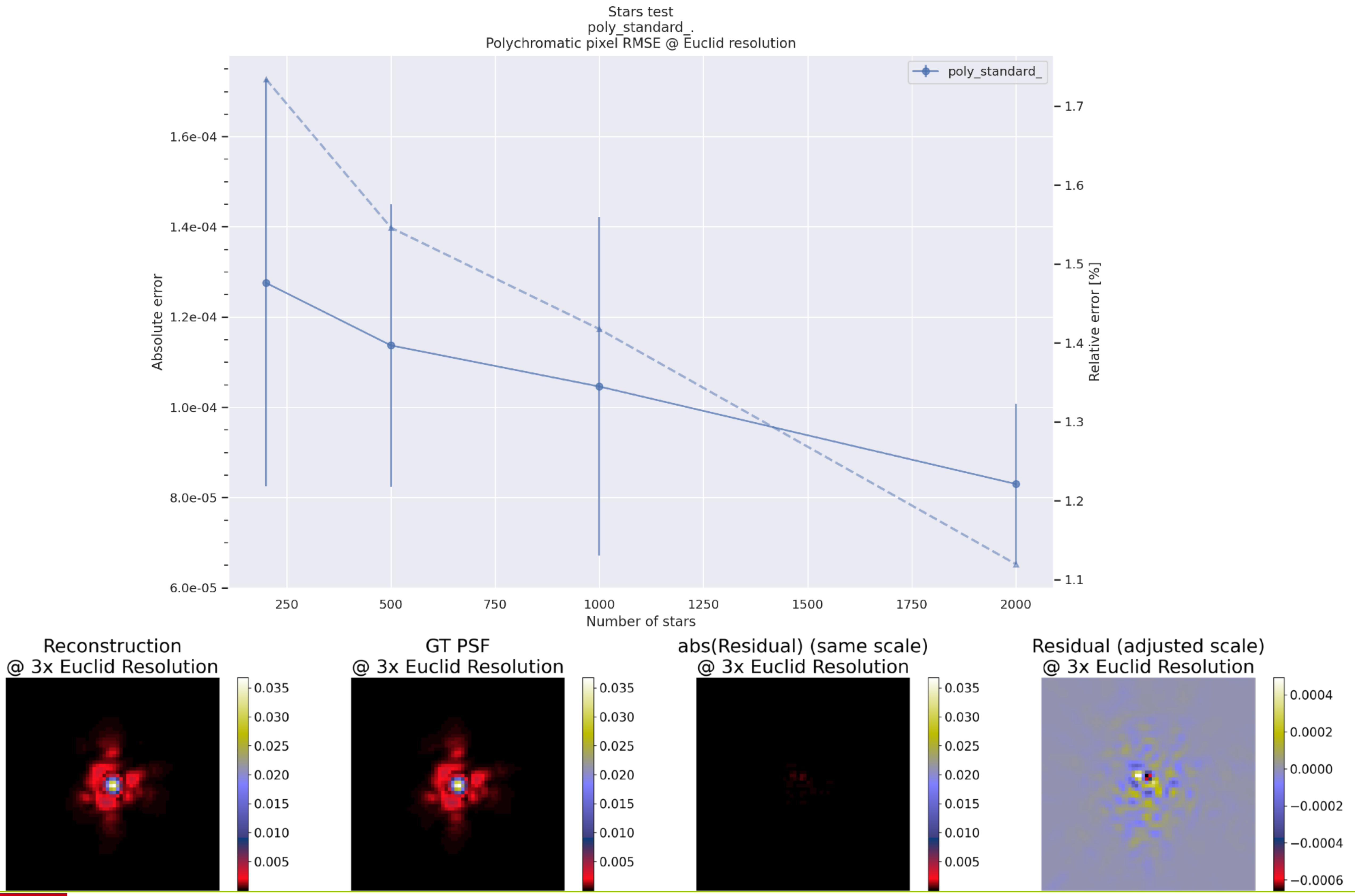
Parametric model

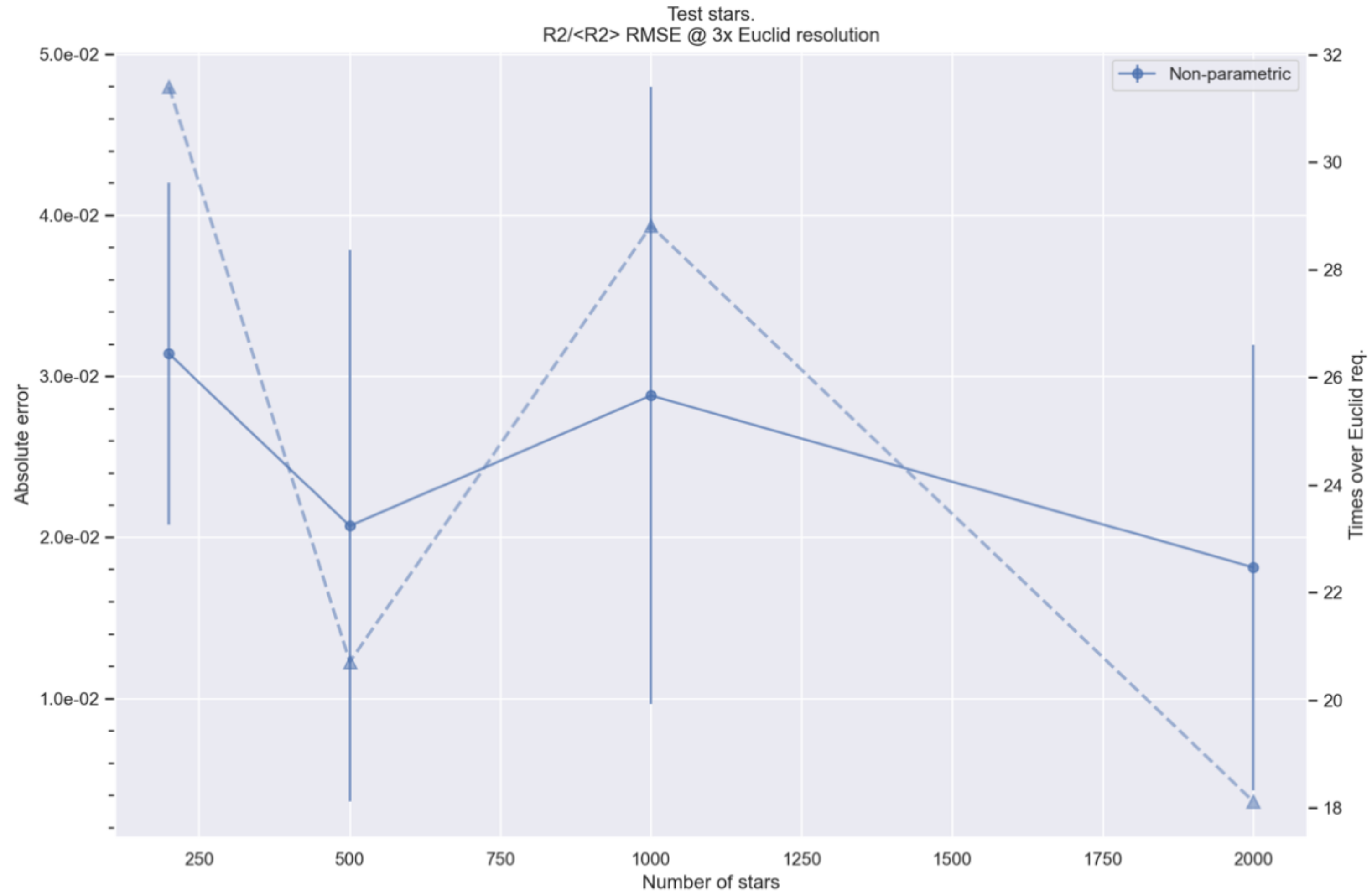
WF-RCA model

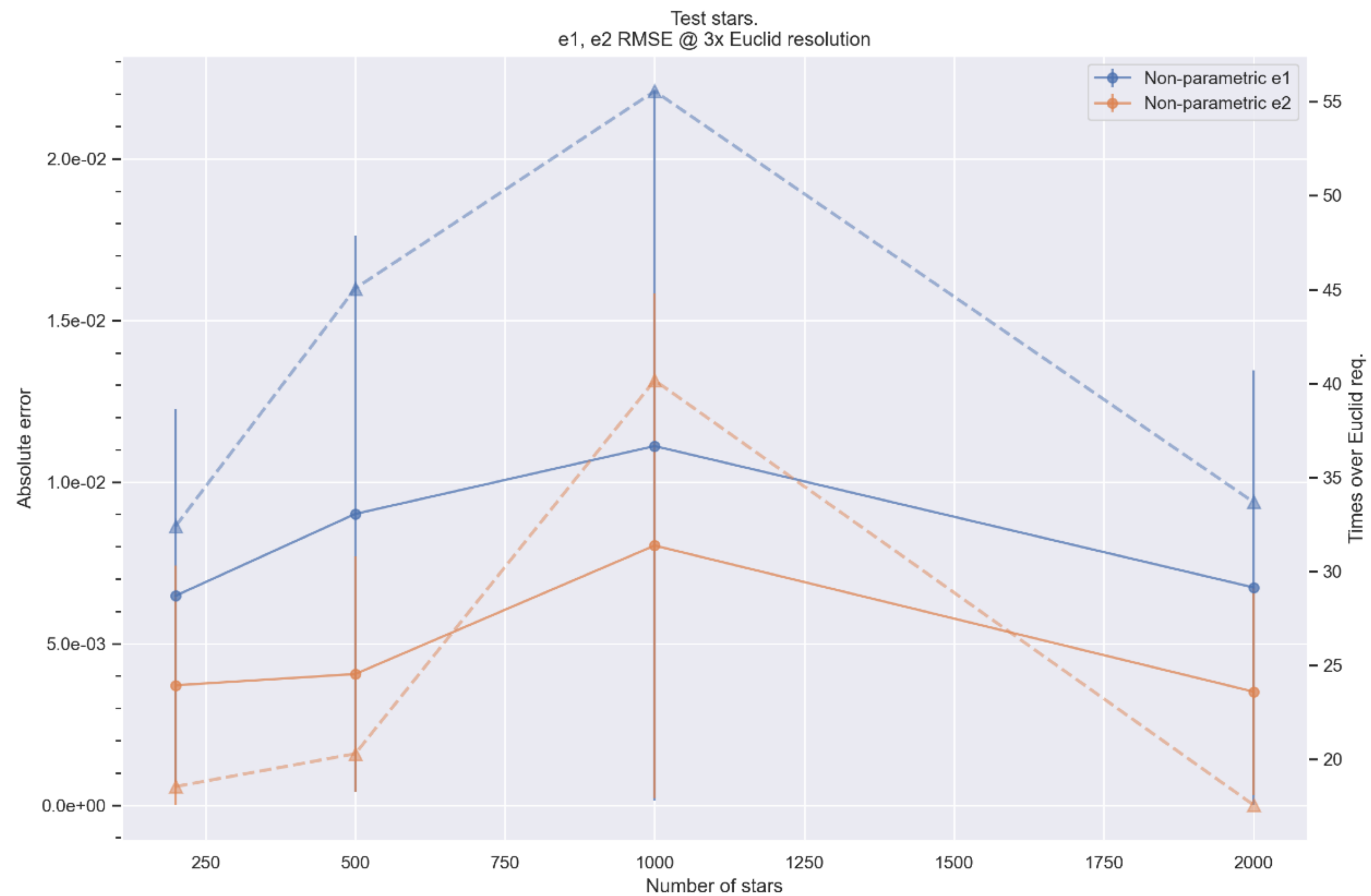
→ WF-RCA improves by a factor of 4 the RMSE compared to the parametric model which uses the ground truth underlying model.

CosmoStat

Numerical experiment: Results

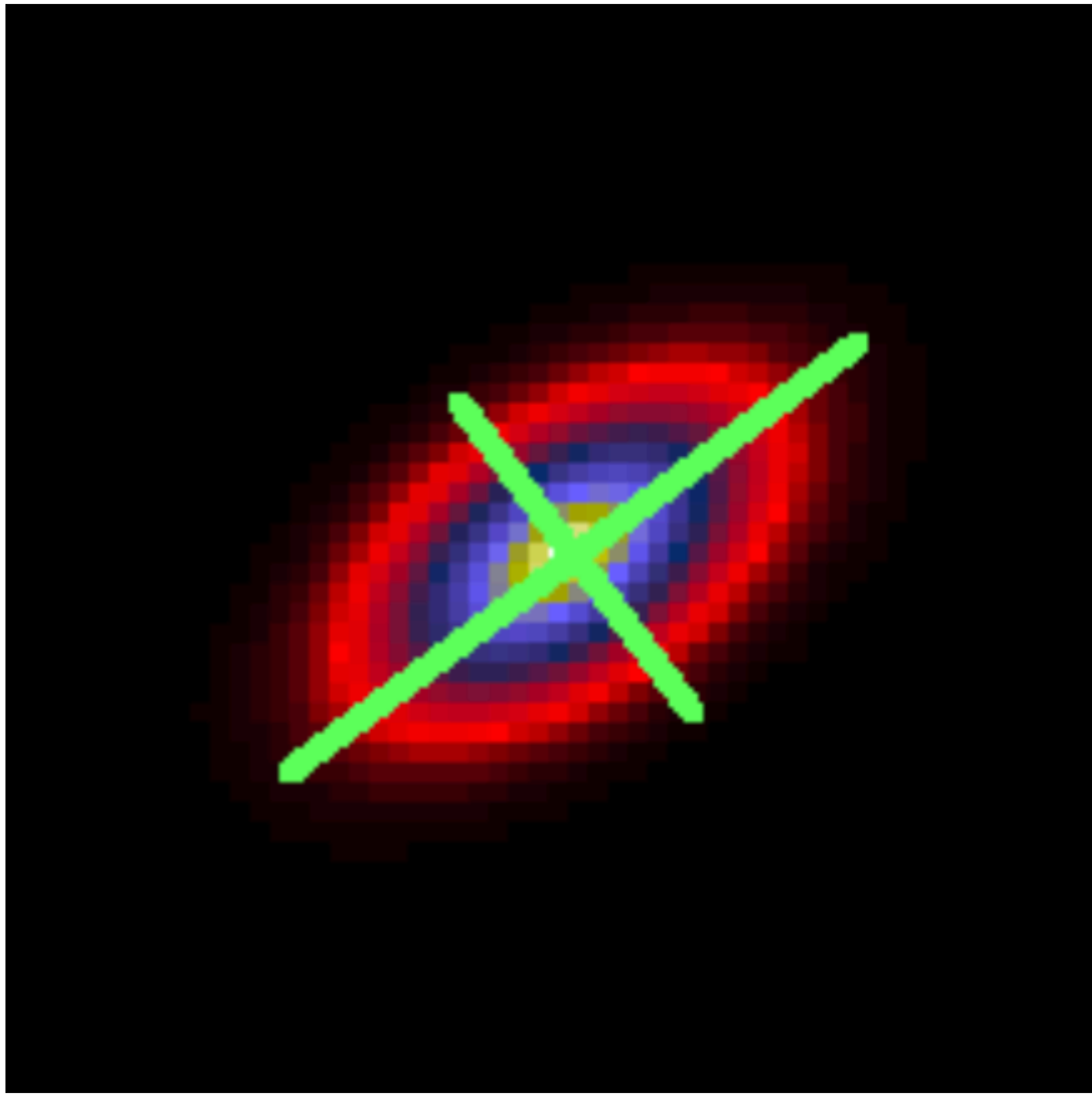



Test stars.
R2/<R2> RMSE @ 3x Euclid resolution
Non-parametric
Absolute error
Times over Euclid req.
Number of stars

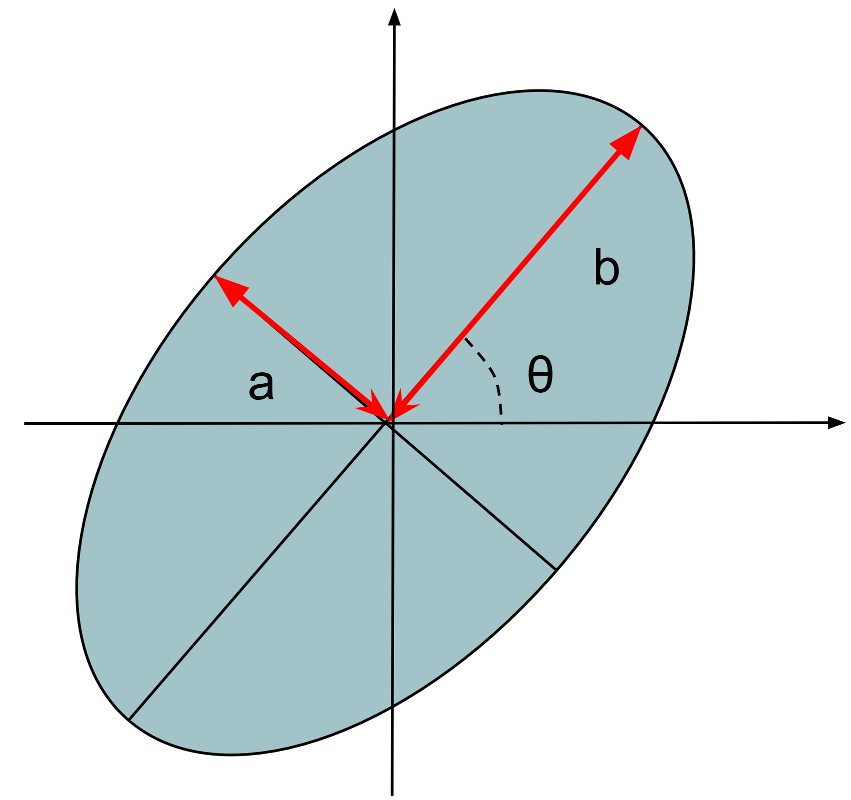

Test stars.
e1, e2 RMSE @ 3x Euclid resolution
Non-parametric e1
Non-parametric e2
Absolute error
Times over Euclid req.
Number of stars

Measuring Galaxies shapes

The complex ellipticity of a galaxy image uses quadrupole moments



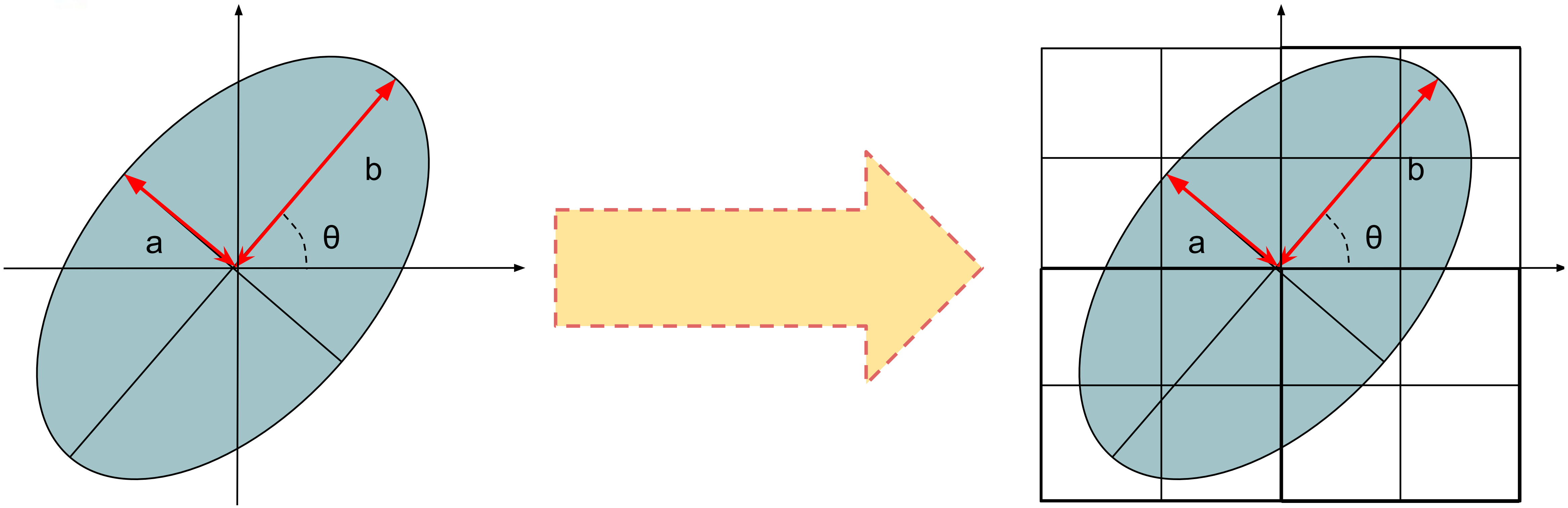
X : **Galaxy**



$$e(X) = \frac{1 - \left(\frac{b}{a}\right)^2}{1 + \left(\frac{b}{a}\right)^2} \exp(2i\theta) = e_1(X) + ie_2(X)$$

$e = \gamma + e_g$ and $< e > \simeq \gamma$

From Ellipticity to Moments



$$X \in \mathcal{M}_n(\mathbb{R})$$

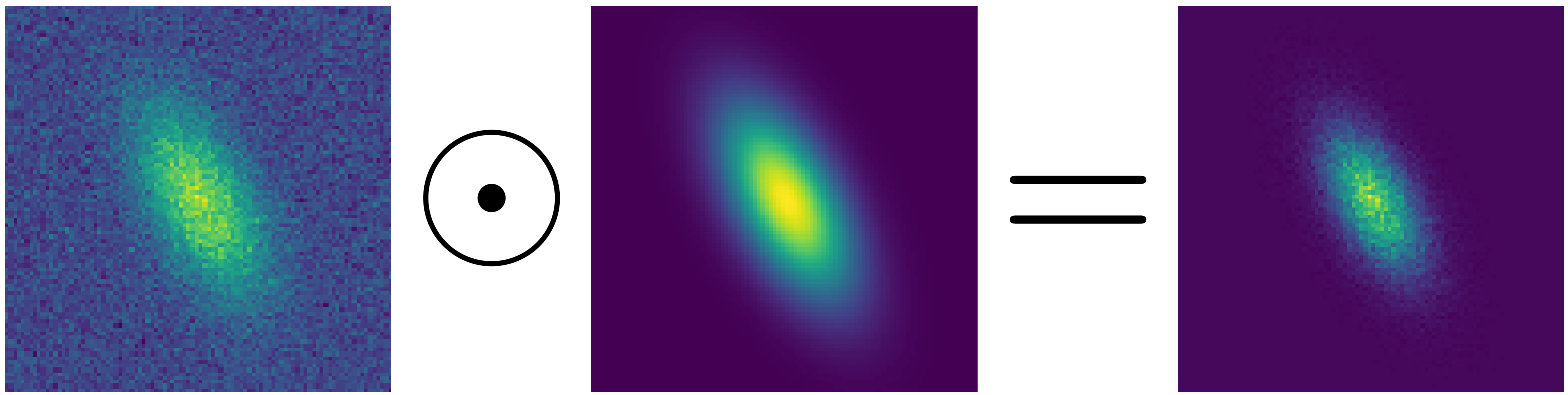
$$\mu_{s,t}(X) = \sum_{i,j} X[i,j](i - i_c)^s (j - j_c)^t$$

$$\gamma_1(X) = \frac{\mu_{2,0}(X) - \mu_{0,2}(X)}{\mu_{2,0}(X) + \mu_{0,2}(X)} \qquad \gamma_2(X) = \frac{2\mu_{1,1}(X)}{\mu_{2,0}(X) + \mu_{0,2}(X)}$$

$\gamma = \gamma_1 + i\gamma_2 = \dfrac{\mu_{2,0} - \mu_{0,2} + 2\mu_{1,1}}{\mu_{2,0} + \mu_{0,2}}$ is an unbiased estimator of “ellipticity” (Schneidez & Seitz 1994)

Moments Windowing

Does not work in presense of noise! Need to apply a window function.



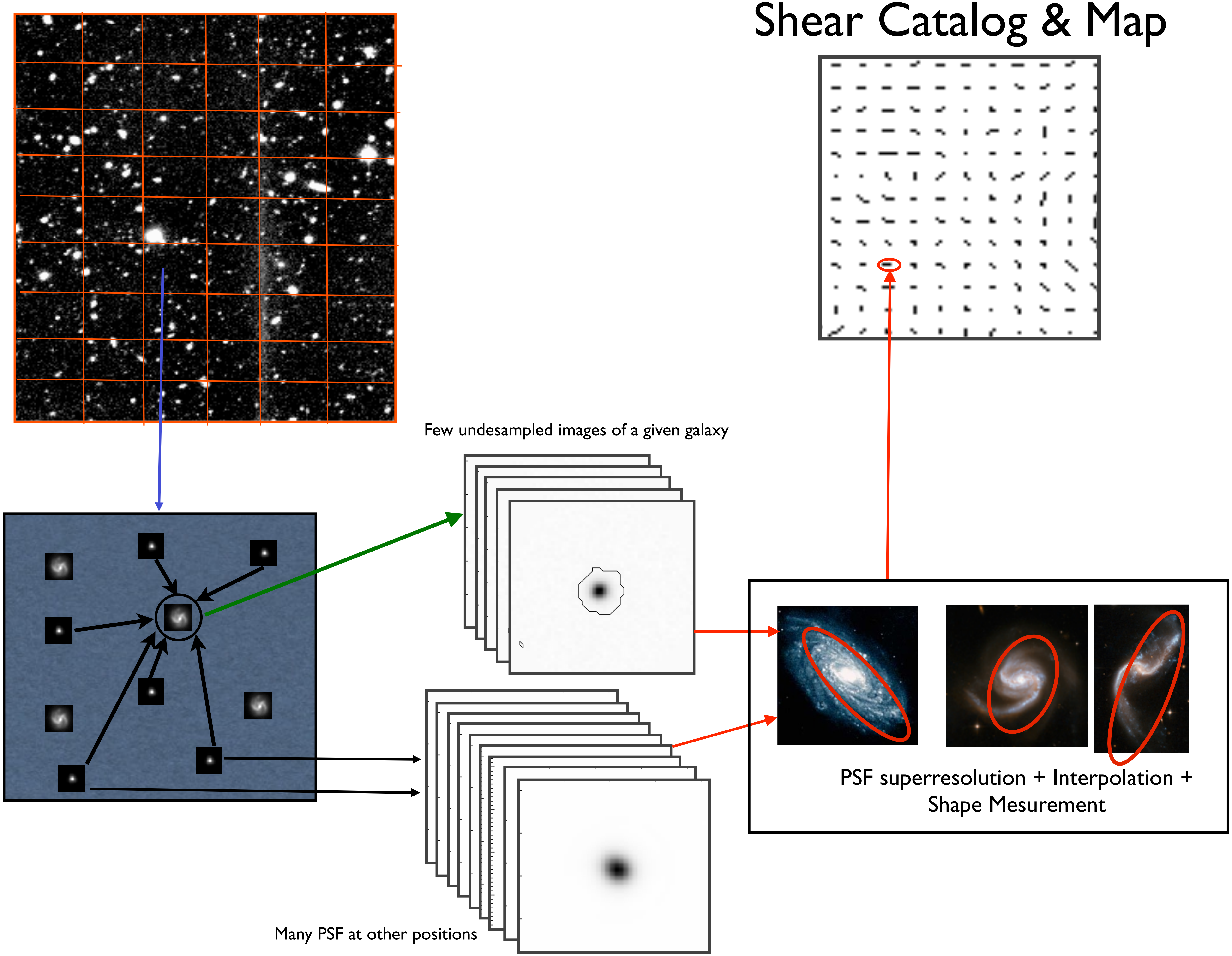
Observed galaxy Y

Matched Gaussian window W

Windowed galaxy $W \odot Y$

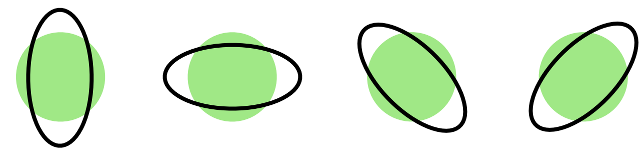
- "HSM" (or adaptive moments, Hirata & Seljak 2003; Mandelbaum et al 2004): match the gaussian window to the object
- Handling the PSF: Kaiser, Squires & Broadhurst 1995 (**KSB**).

Shear Catalog & Map

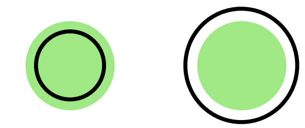


From Galaxies to Mass Maps

Euclid galaxy catalogue

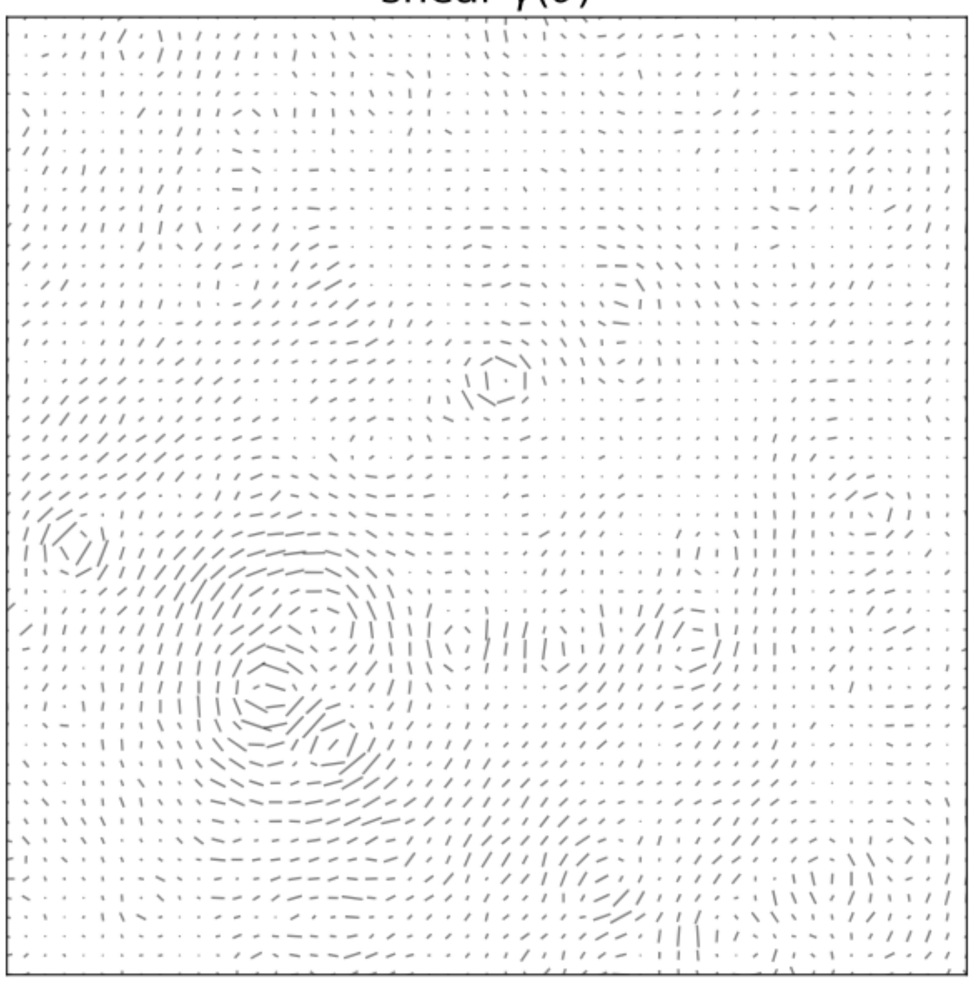


$$\gamma = \gamma_1 + \mathrm{i}\gamma_2$$

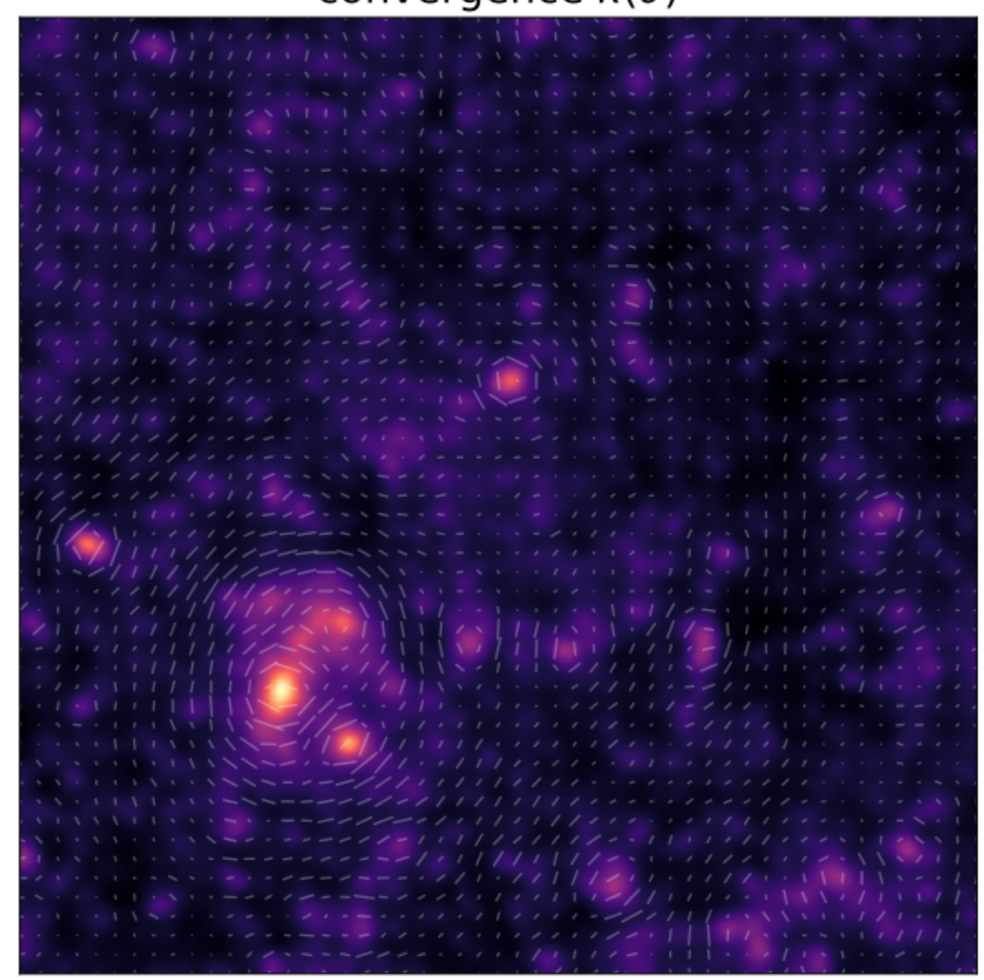


$$\kappa$$

shear $\gamma(\vec{\theta})$



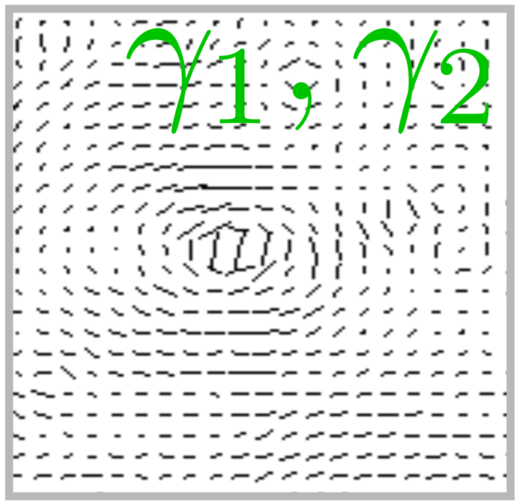
convergence $\kappa(\vec{\theta})$



Mass Mapping

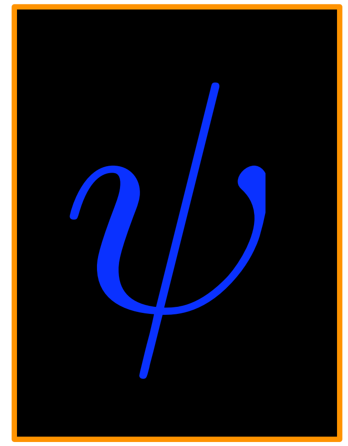
★ Starck, et al, A&A, Vol. 451, pp 1139-1150, 2006.

SIMULATED SHEAR MAP



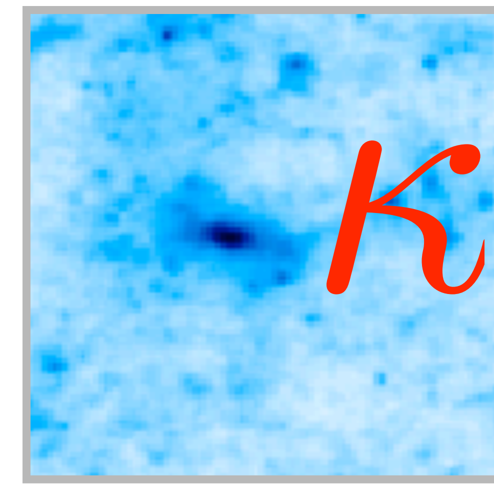
$$\gamma_1 = \frac{1}{2}(\partial_1^2 - \partial_2^2)\psi$$
$$\gamma_2 = \partial_1 \partial_2 \psi$$

LENSING POTENTIAL



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

SIMULATED MASS MAP
(Vale & White, 2003)



From mass to shear: $\gamma_i = \hat{P}_i \kappa$

From shear to mass: $\kappa = \hat{P}_1 \gamma_1 + \hat{P}_2 \gamma_2$

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

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

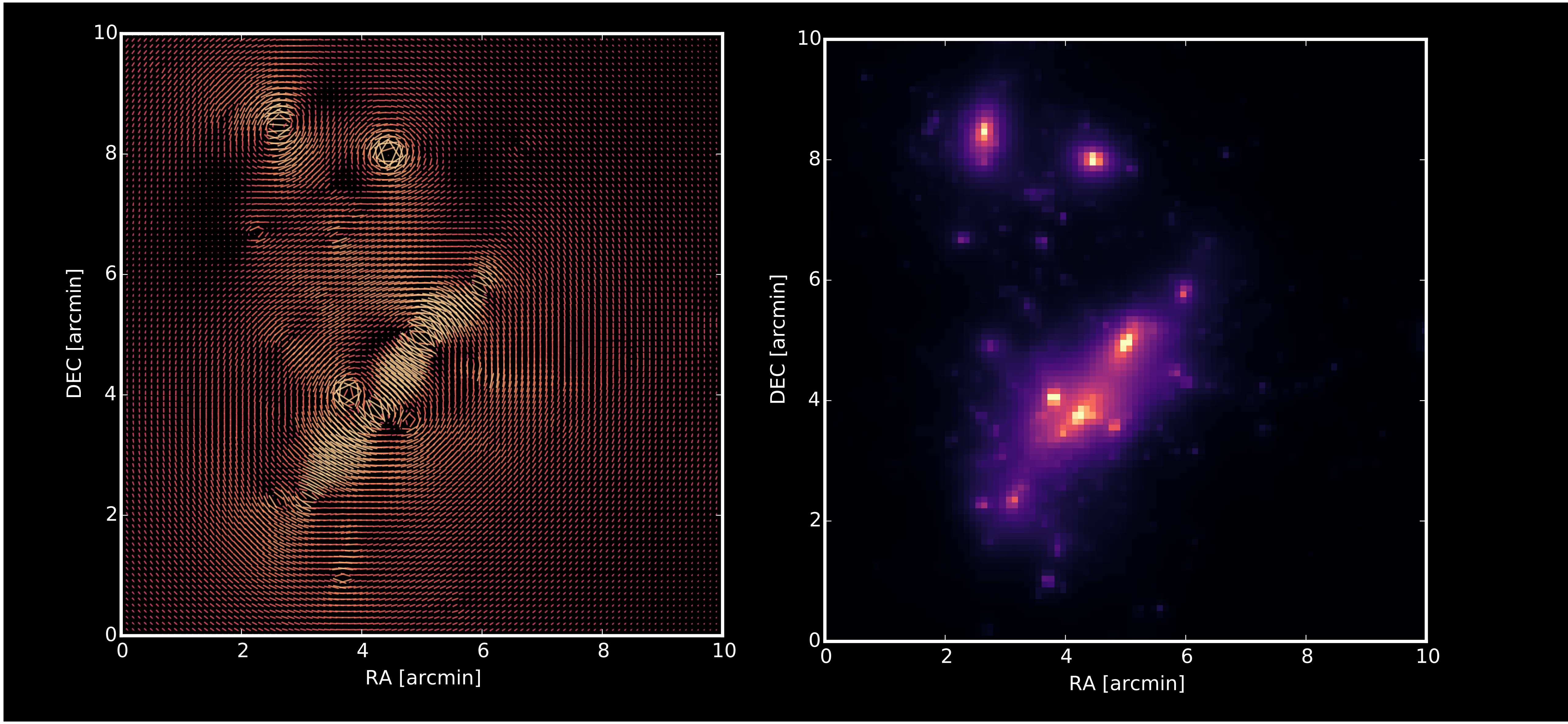
$$\begin{pmatrix} \hat{E}(\boldsymbol{k}) = \hat{\kappa}(\boldsymbol{k}) \\ \hat{B}(\boldsymbol{k}) \end{pmatrix} = \underbrace{\frac{1}{|\boldsymbol{k}|^2} \begin{pmatrix} k_1^2 - k_2^2 & 2k_1 k_2 \\ 2k_1 k_2 & -k_1^2 + k_2^2 \end{pmatrix}}_{A_{\boldsymbol{\kappa}}} \begin{pmatrix} \hat{\gamma}_1(\boldsymbol{k}) \\ \hat{\gamma}_2(\boldsymbol{k}) \end{pmatrix}$$

Regular shear sampling

$$\kappa = F^* P F \gamma$$

Shear γ → Convergence κ

complex field $\gamma(\theta) = \gamma_1 + i\gamma_2$

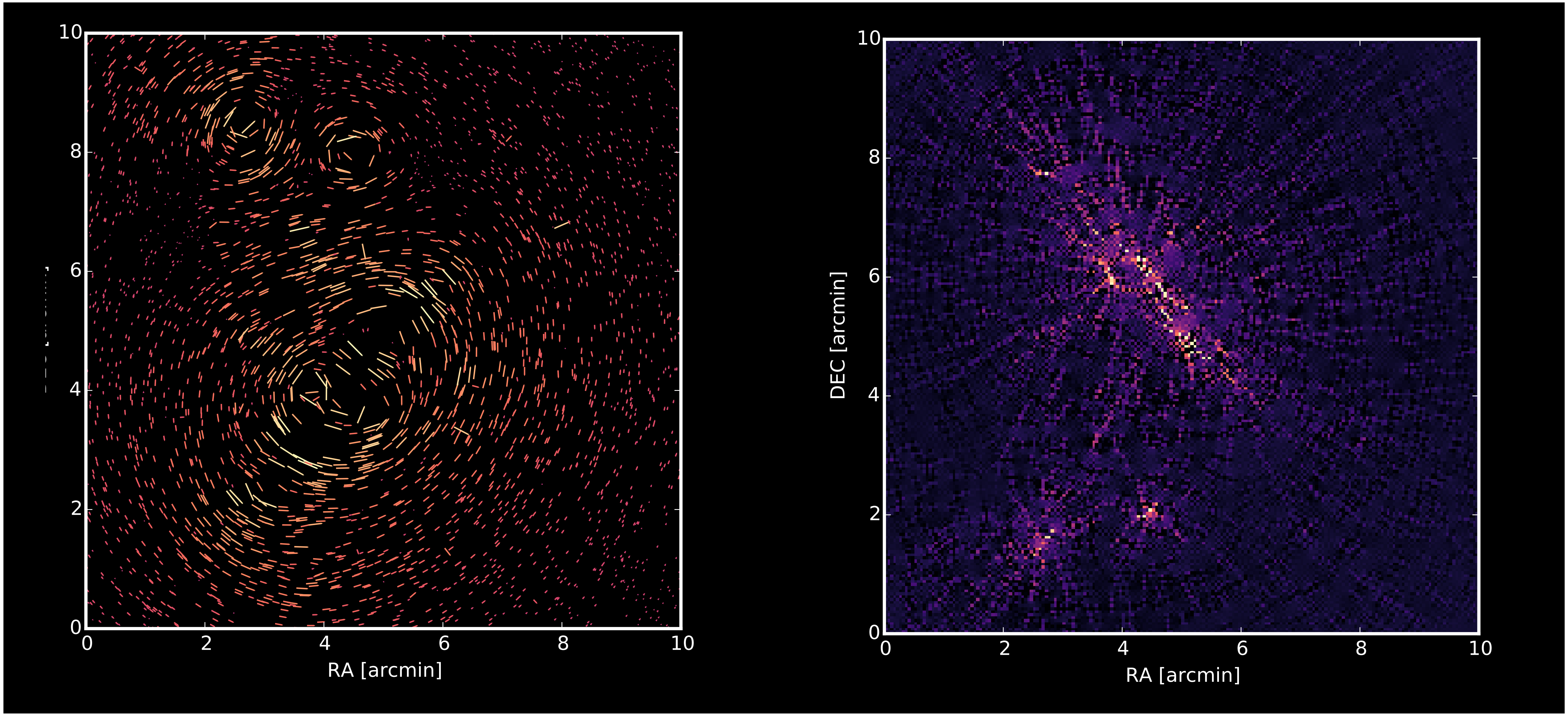


CosmoStat

Irregular shear sampling

Shear γ

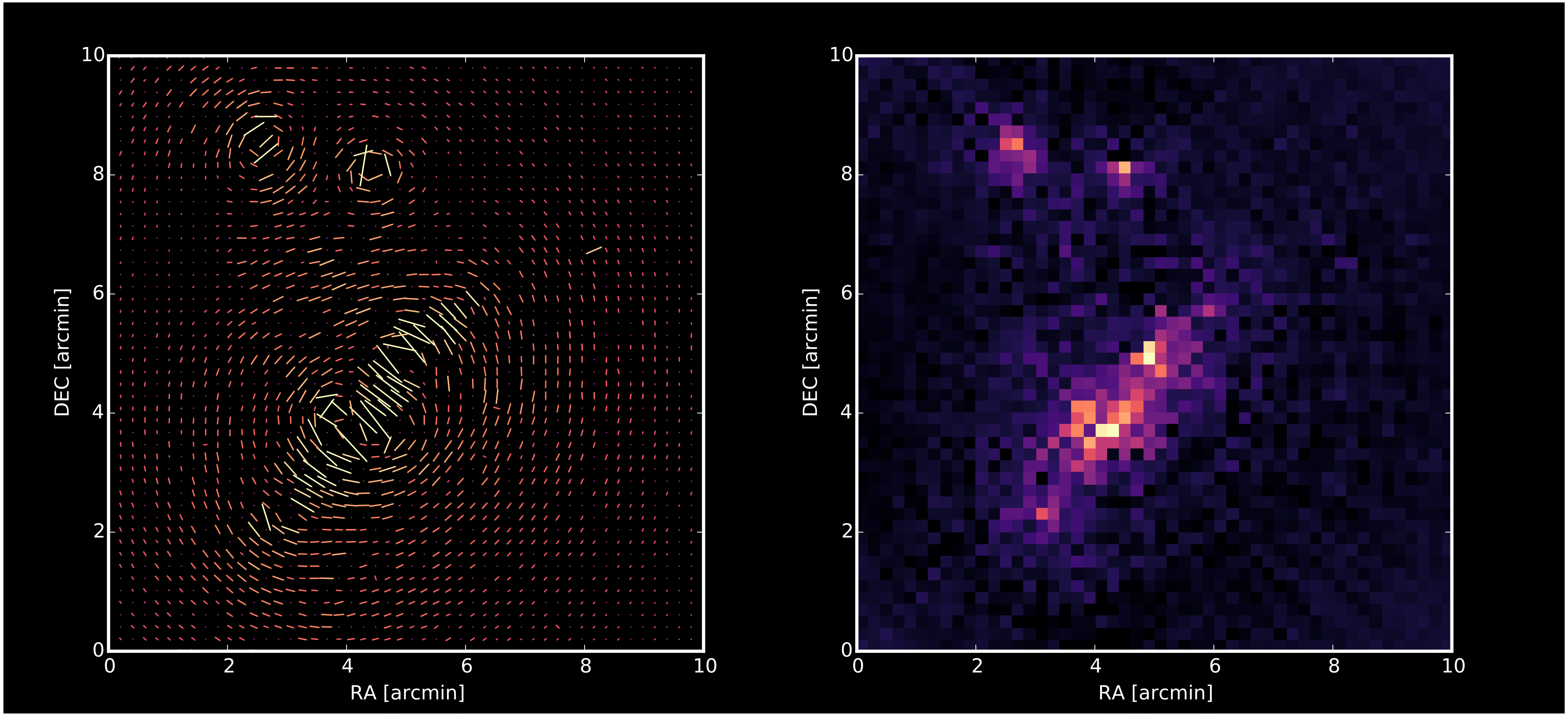
Convergence κ



$$\kappa = F^* P F M \gamma$$

Increasing the bin size (0.2 arcmin)

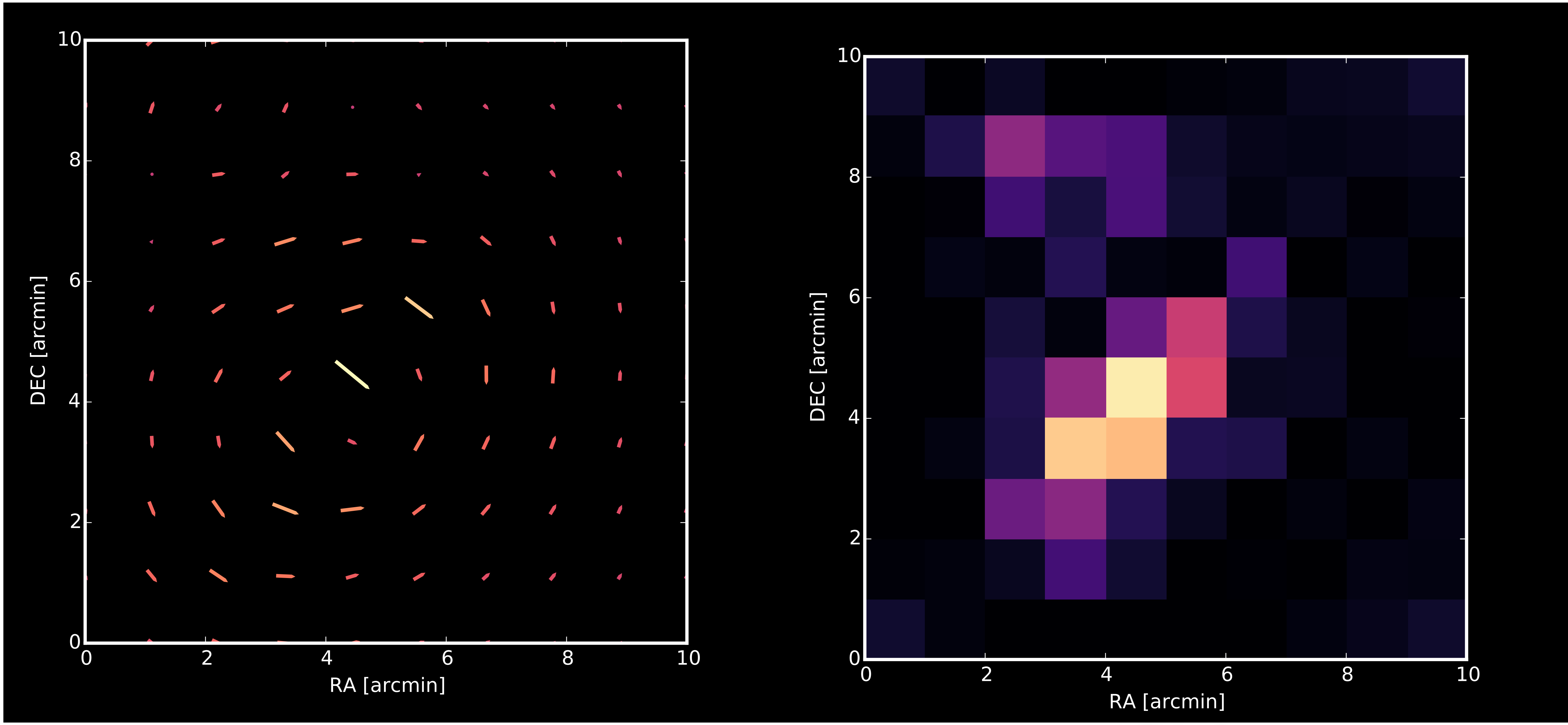
Shear γ Convergence κ



$$\kappa = F^* P F M \gamma$$

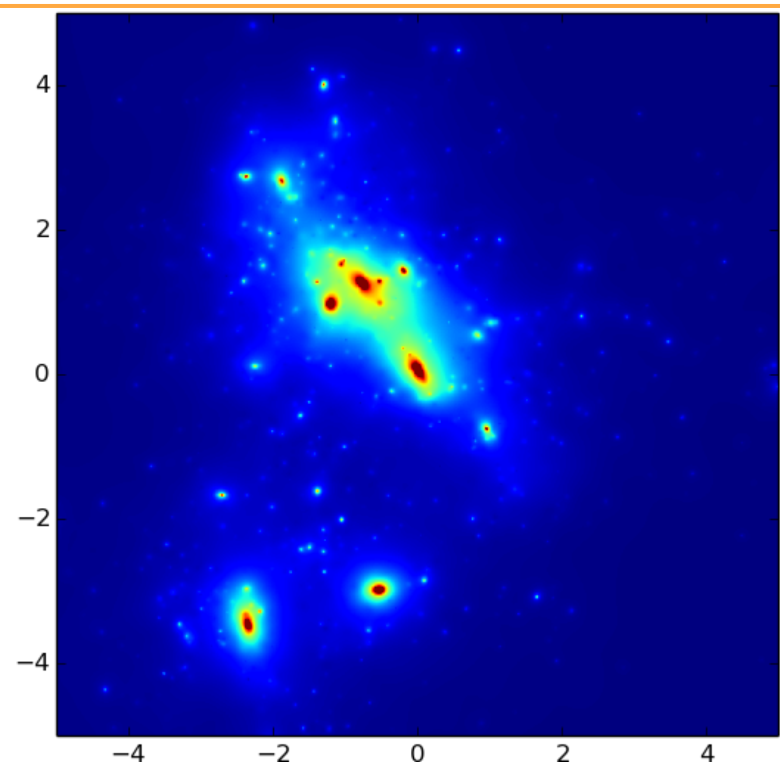
Increasing the bin size (1 arcmin)

Shear γ Convergence κ

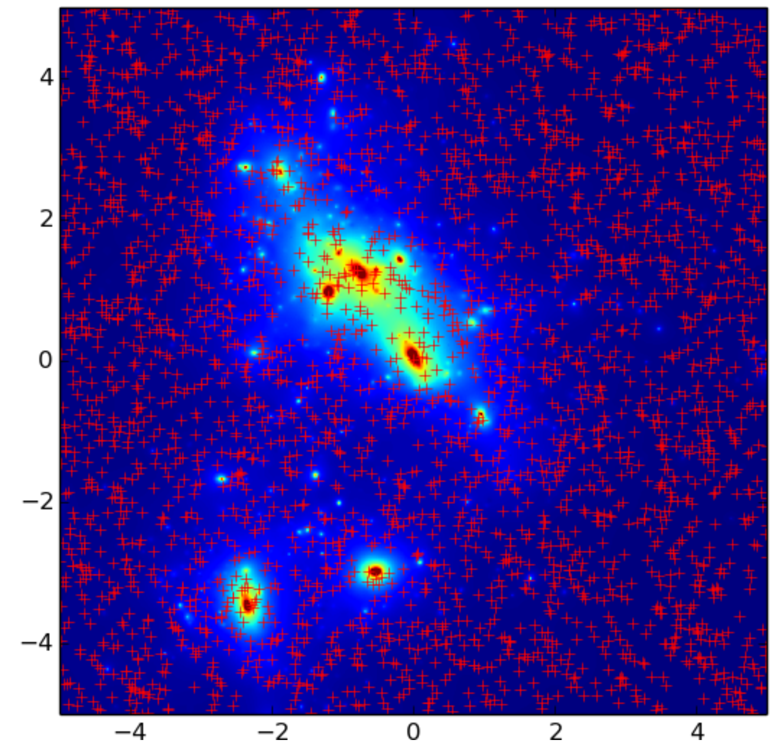


$$\kappa = F^* P F M \gamma$$

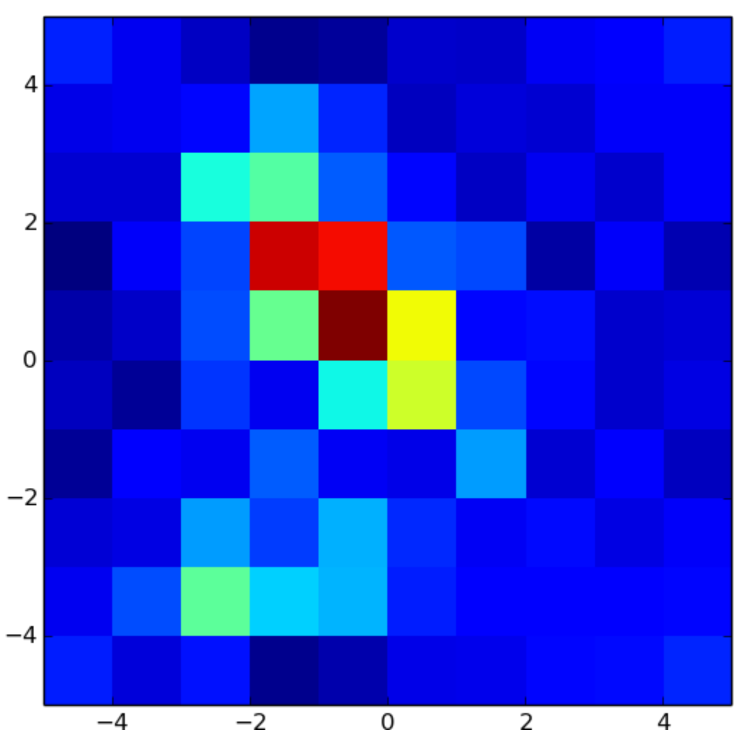
Handling Missing Data (no noise): Binning+Smoothing



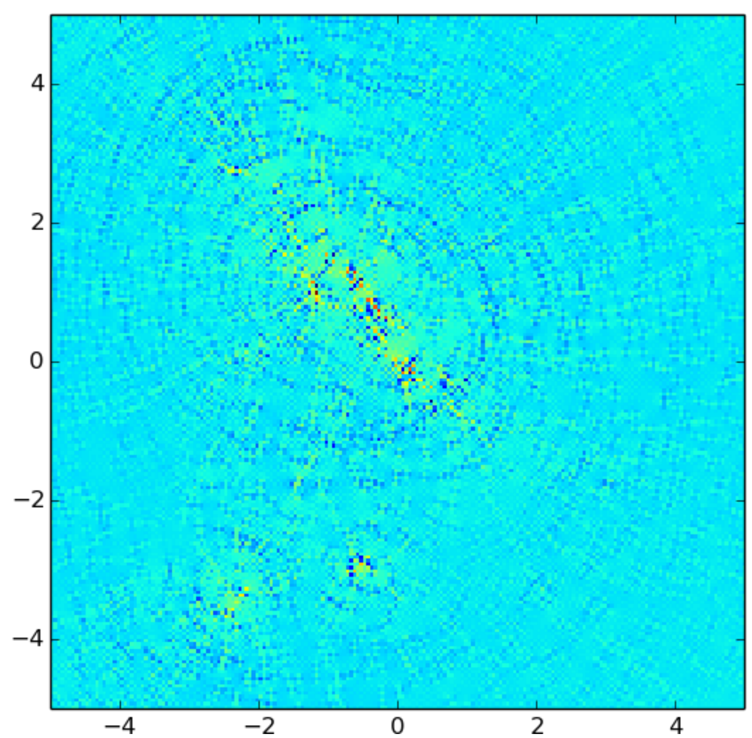
Input



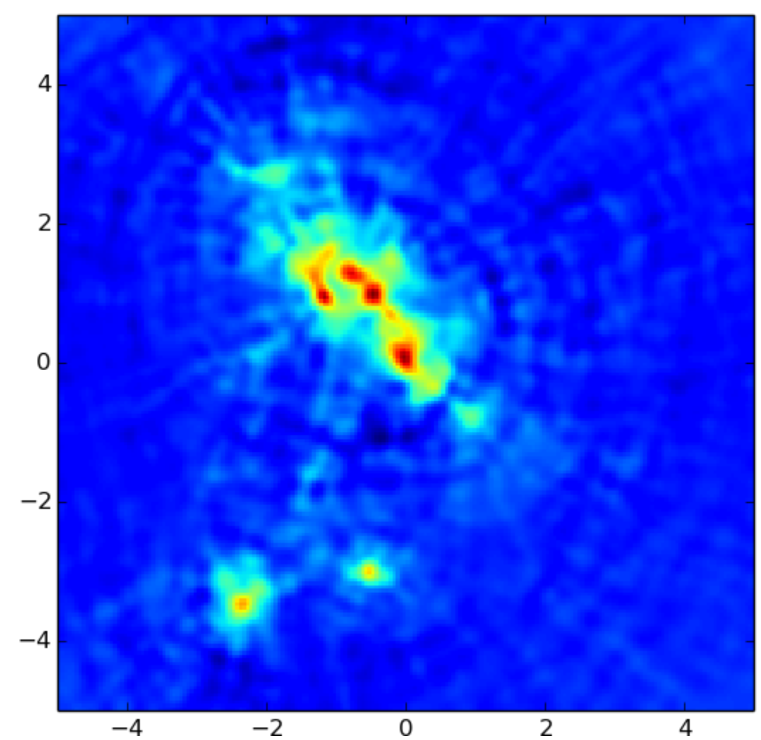
Galaxy catalogue with 30 gal/arcmin^2



Kaiser-Squires with 1' bins



Kaiser-Squires with 0.05' bins



KS with 0.05' bins + 0.1' smoothing

Mass Mapping

Advantages

- Simple linear operator – noise well understood
- Very easy to implement in Fourier space

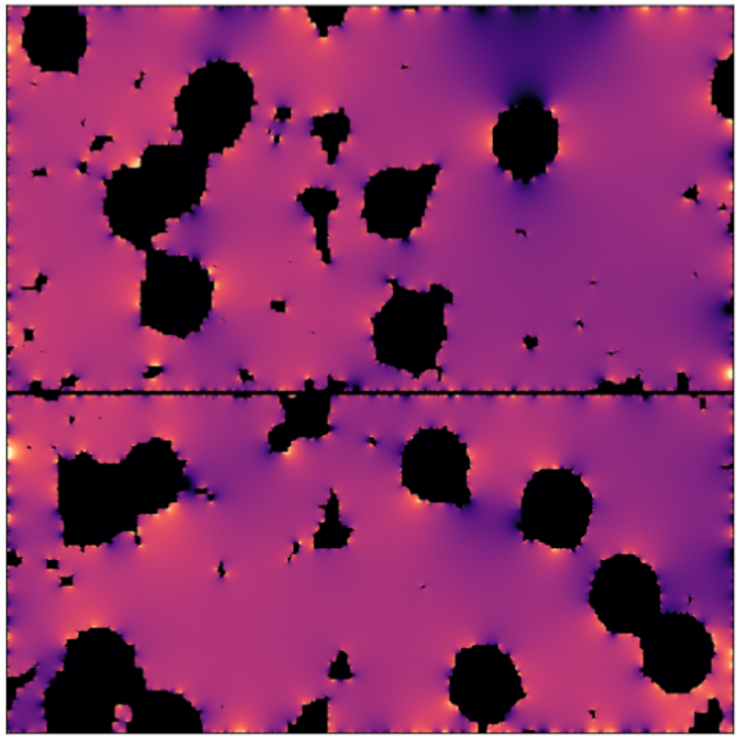
$$X_k = \sum_{n=0}^{N-1} x_n e^{-i2\pi kn/N}$$

$$x_n = \frac{1}{N} \sum_{k=0}^{N-1} X_k e^{i2\pi kn/N}$$

Practical difficulties

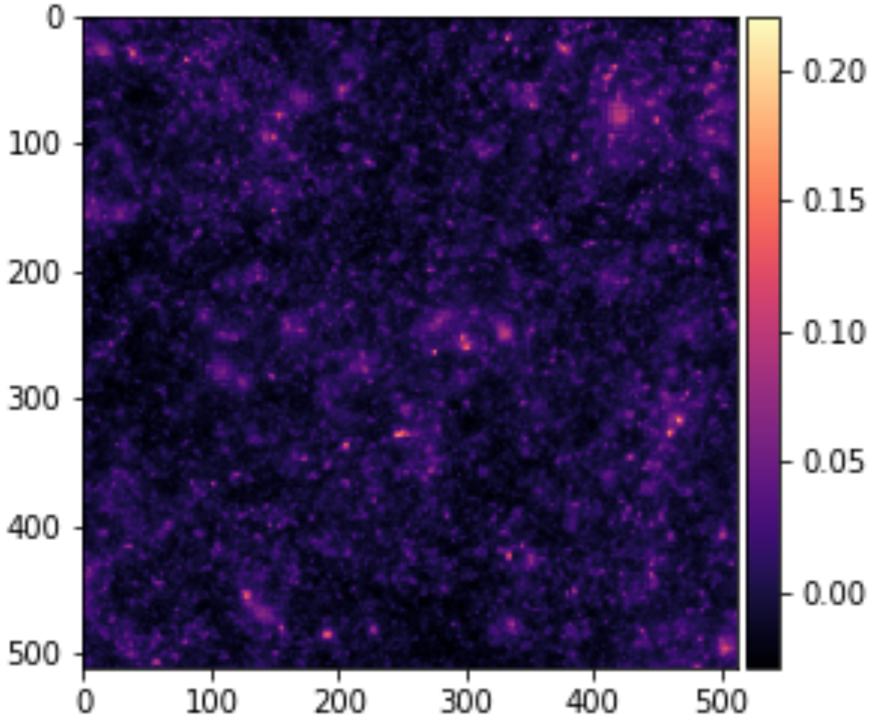
- Shear measurements are discrete, **noisy**, and **irregularly sampled**
- We actually measure reduced shear $g = \gamma/(1-\kappa)$
- Masks + integration over a subset of $\mathbb{R}^2$ lead to border errors => **missing data problem**
- Convergence recoverable up to a constant

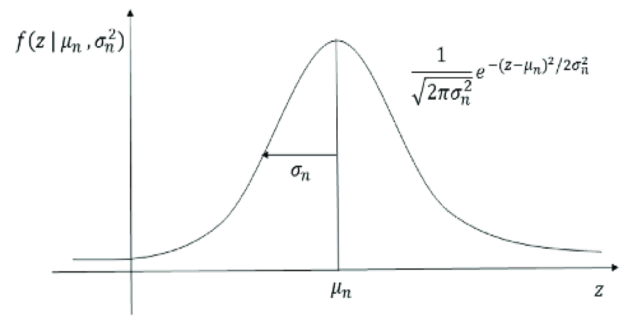
=> Mass-sheet degeneracy problem $\kappa^{'} = \lambda\kappa(1-\lambda)$



Euclid Shape Noise

Noiseless convergence map

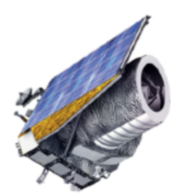




+

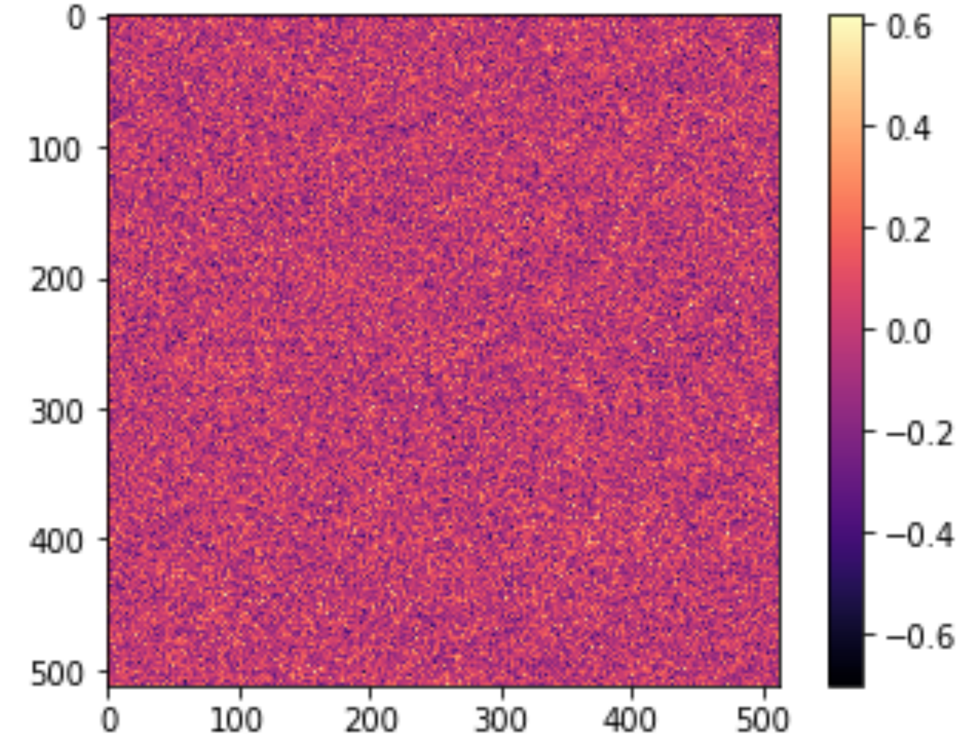
$$\sigma^2_{noise} = \frac{< \sigma_\lambda^2 >}{n_{gal} \Delta \boxtimes}$$

=



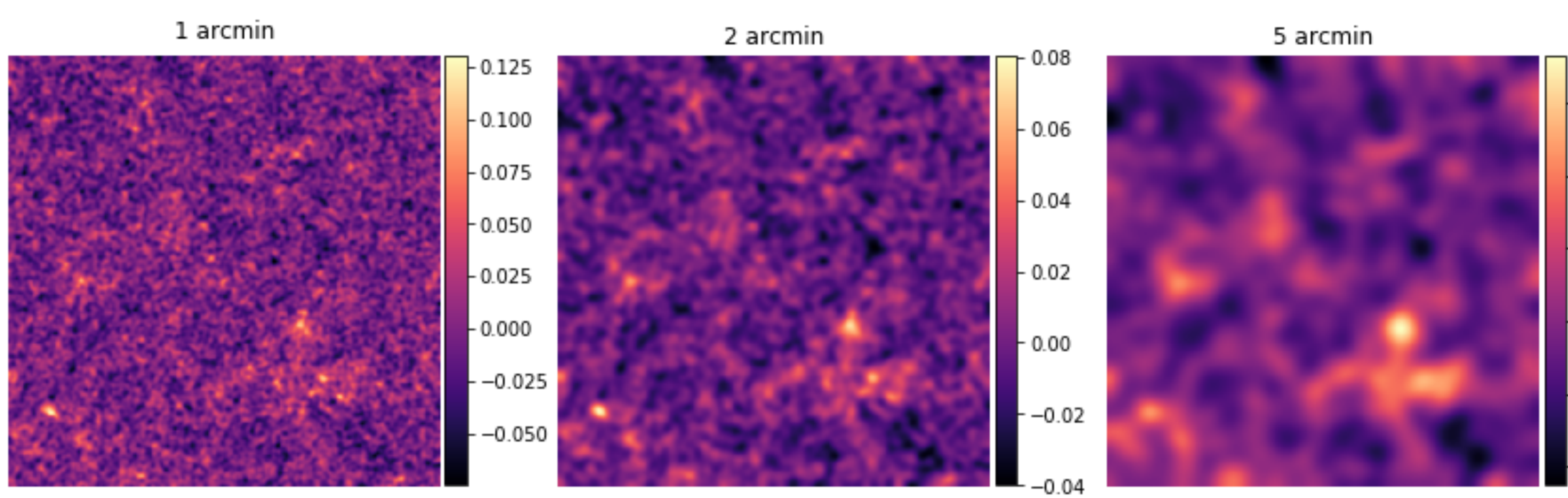
Euclid-like noise

Noisy convergence map



-Filtering noise

-Access cosmolgical information



Mass mapping as an inverse problem

Binned data: $\gamma = F^* P F \kappa$

Unbinned data: $\gamma = T^* P F \kappa$

T = Non Equispaced Discrete Fourier Transform (NDFT)

$$\tilde{\boldsymbol{\kappa}} = \min_{\boldsymbol{\kappa}} \|\boldsymbol{\gamma} - \mathbf{A}\boldsymbol{\kappa}\|^2_{\boldsymbol{\Sigma_n}}$$

with $\mathbf{A} = \boldsymbol{F}\mathbf{P}\boldsymbol{F}^*$

There is no unique and stable solution, it is an ill posed inverse problem.

CosmoStat

State of the Art

- **For clusters:**
  - Model fitting algorithms (Bartelmann et al, 1996; Bradac et al, 2005; Jullo and Kneib, 2009).
  - Aperture Mass (Seitz and Schneider, 1996; 2002).
- **For larger fields:**
  - Kaiser-Squires (1993) + Gaussian smoothing (used for HSC (2018), DES (2017))
  - **Wiener** (Jeffrey et al, 2020).
  - Sparsity: **Glimpse2D** (Lanusse et al, 2016), (Price et al., **2020**), used for DES (2018&2021)
  - Bayesian approches (Heavens et al, **2016**, Alsing et al , **2017**, Schneider et al, **2017**).
  - **DeepLearning**: (Jeffrey et al, **2020**).
  - **DeepLearning + Bayesian (**B. Remy et al, "Probabilistic Mapping of Dark Matter by Neural Score Matching", Machine Learning and the Physical Sciences Workshop, NeurIPS 2020.
- **3D Mass Mapping:**
  - 3D SVD Inversion (Simon et al, 2009) -> HSC (2018)
  - Bayesian approches (Bohm et al, **2017**).
  - **Glimpse3D** (Leonard et al, 2012, Leonard et al, **2014**).

Bayesian Reconstruction

$$\underbrace{p(\boldsymbol{\kappa}|\boldsymbol{\gamma}, \mathcal{M})}_{\textbf{Posterior}} \propto \underbrace{p(\boldsymbol{\gamma}|\boldsymbol{\kappa}, \mathcal{M})}_{\textbf{likelihood}} \; \underbrace{p(\boldsymbol{\kappa}|\mathcal{M})}_{\textbf{prior}}$$

$\mathcal{M}$ is the cosmological model

Proximal Wiener filtering

$$\boldsymbol{\kappa}_G = \min_{\boldsymbol{\kappa}} \|\gamma - \mathbf{A}\boldsymbol{\kappa}\|^2_{\boldsymbol{\Sigma}_{\boldsymbol{n}}} + \|\boldsymbol{\kappa}\|^2_{\boldsymbol{\Sigma}_{\boldsymbol{\kappa}}}$$

$$\boldsymbol{\kappa}_G = (\mathbf{A}\boldsymbol{\Sigma}_{\boldsymbol{\kappa}}\mathbf{A}^* + \boldsymbol{\Sigma}_{\boldsymbol{n}})^{-1}\mathbf{A}^*\boldsymbol{\Sigma}_{\boldsymbol{\kappa}}\gamma$$

Using the proximal optimisation, we get the following iterative Wiener filtering algorithm (Bobin et al, 2012):

Forward step: $$\mathbf{t} = \boldsymbol{\kappa}^n + 2\mu\mathbf{A}^*\boldsymbol{\Sigma}_n^{-1}(\gamma - \mathbf{A}\boldsymbol{\kappa}^n)$$

Backward step: $$\boldsymbol{\kappa}^{n+1} = \boldsymbol{F}^*\left(\mathbf{P}_{\boldsymbol{\kappa}}\left(\mathbf{P}_\eta + \mathbf{P}_{\boldsymbol{\kappa}}\right)^{-1}\right)\boldsymbol{F}\mathbf{t}$$

Mass mapping as an inverse problem

F. Lanusse, J.-L. Starck, A. Leonard, and S . Pires, High Resolution Weak Lensing Mass Mapping combining Shear and Flexion, A&A, 2016.

Binned data: $\gamma = F^* P F \kappa$

Unbinned data: $\gamma = T^* P F \kappa$

$$\mathbf{P} = T^* P F$$

T = Non Equispaced Discrete Fourier Transform (NDFT)

$$\min_{\kappa} \frac{1}{2} \parallel \gamma - \mathbf{P}\kappa \parallel_2^2 \quad + \quad \mathcal{C}(\kappa)$$

$$g = \frac{\gamma}{1-\kappa} \longrightarrow \min_{\kappa} \frac{1}{2} \parallel (1-\kappa) g - \mathbf{P}\kappa \parallel_2^2 \quad + \quad \mathcal{C}(\kappa)$$

CosmoStat

The 2D Glimpse Algorithm

$$\min_{\kappa} \frac{1}{2} F(\kappa) + \lambda \parallel \mathbf{\Phi}^t \kappa \parallel_1 \quad \text{with} \quad F(\kappa) = \frac{1}{2} \parallel (1-\kappa) g - \mathbf{P}\kappa \parallel_2^2$$

Primal-dual splitting:

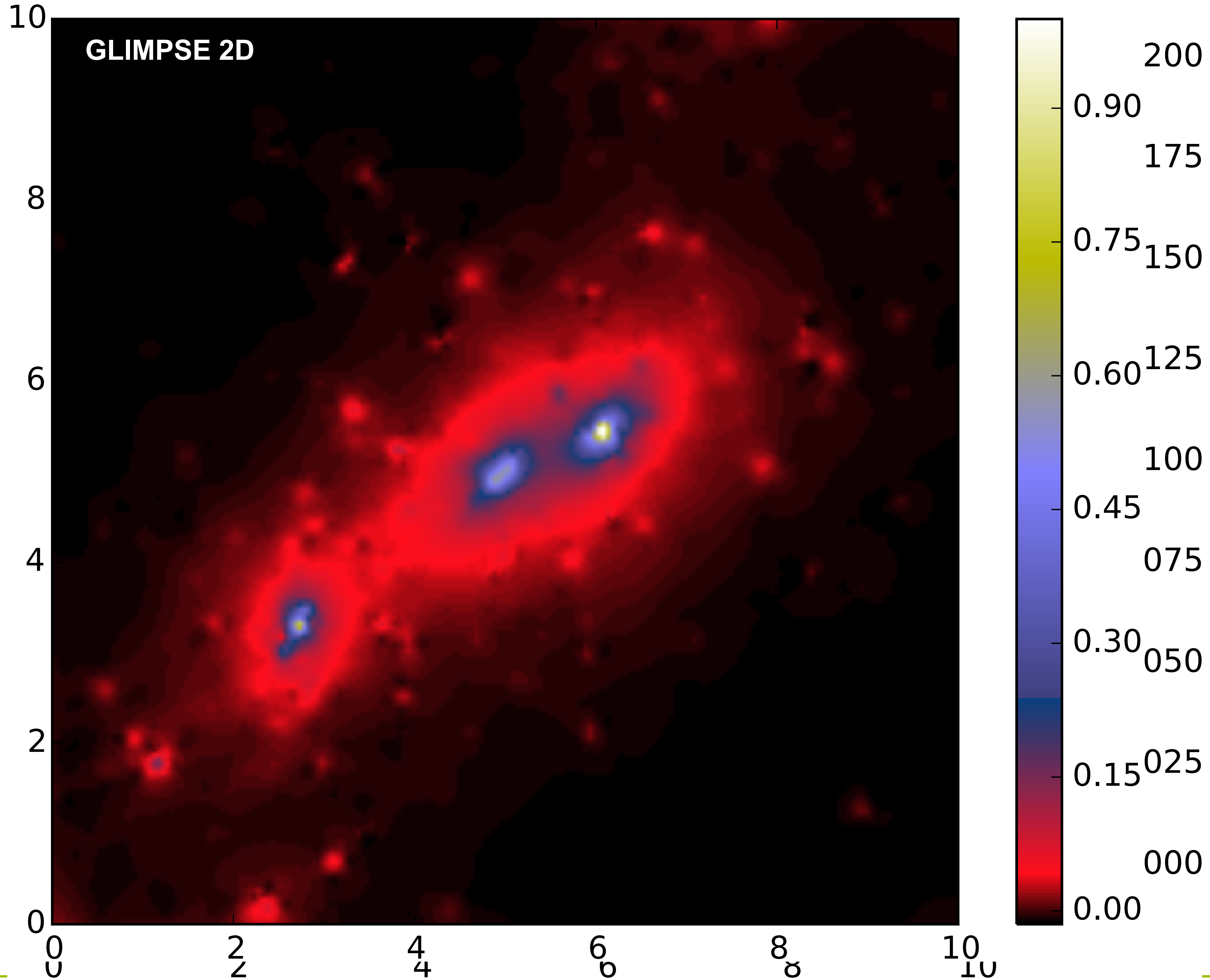
$$\begin{cases} \kappa^{(n+1)} & = & \kappa^{(n)} + \tau \left(\nabla F(\kappa^{(n)}) + \mathbf{\Phi} \alpha^{(n)} \right) \\ \alpha^{(n+1)} & = & (\mathrm{Id} - \mathrm{ST}_\lambda) \left(\alpha^{(n+1)} + \mathbf{\Phi}^t (2\kappa^{(n+1)} - \kappa^{(n)}) \right) \end{cases}$$

Condat-Vu algorithm, 2013

A few remarks:
- Recovers the convergence from the reduced shear
- **P** can be defined with and without binning the shear
- **P** can be ill-posed in case of missing data
- Sparse regularization of noise and missing data
  We use isotropic wavelets, well adapted to the recovery of clusters.

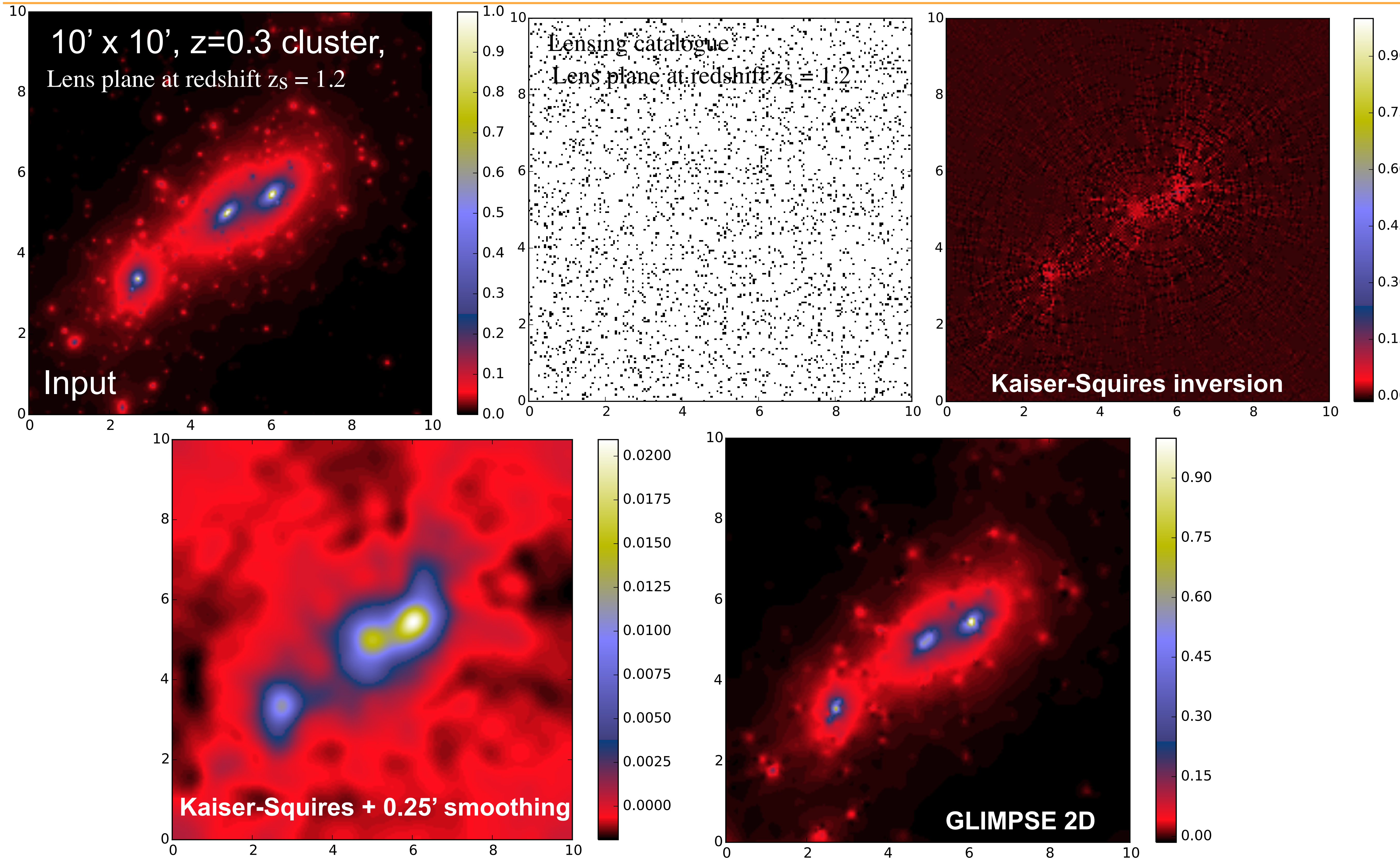
- Fast and flexible algorithm
- Sparsity constraint λ estimated locally by noise simulations $\Longrightarrow$ Accounts for **survey geometry**, **varying noise levels**

Example with 93 % of missing data



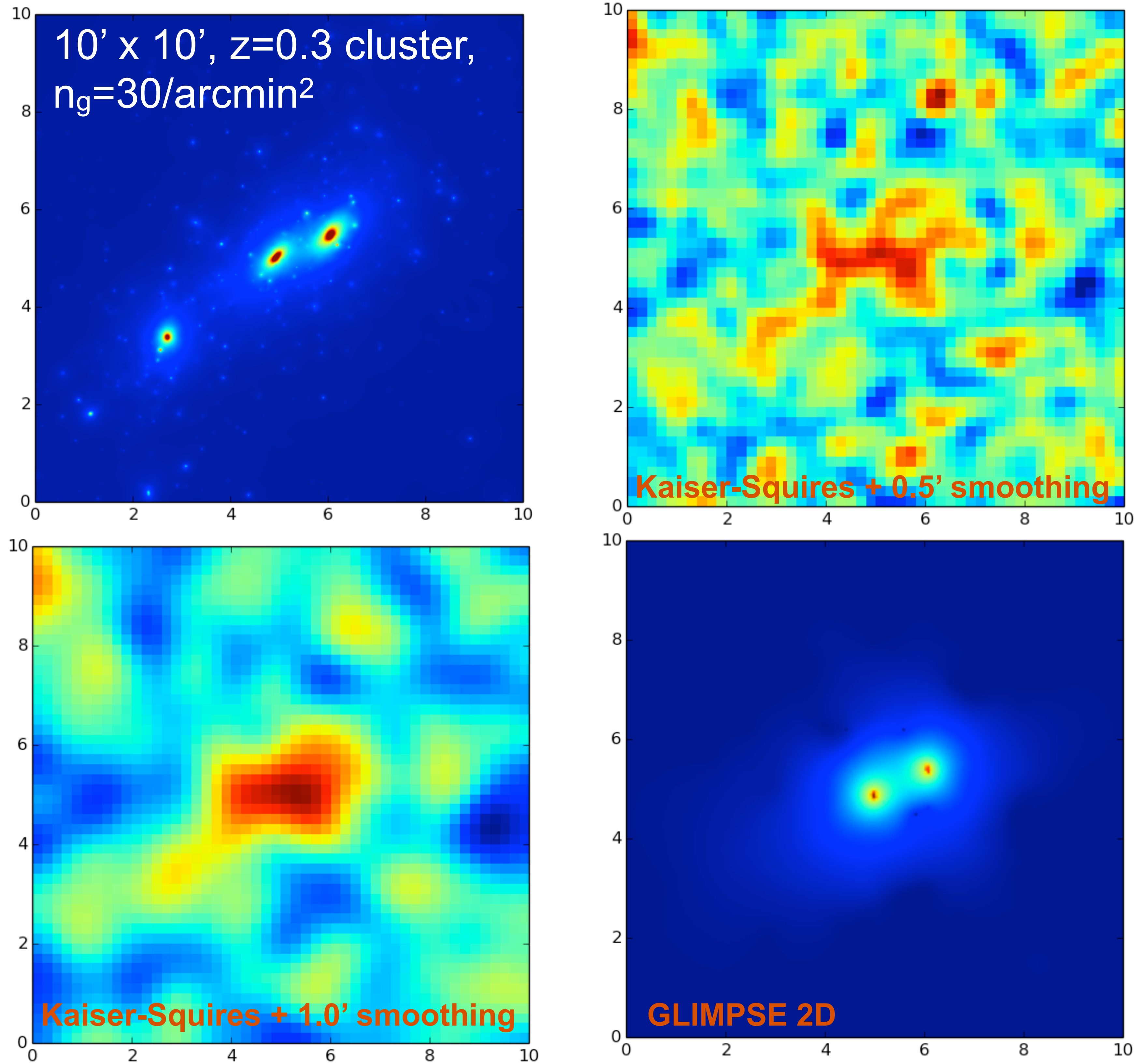
Galaxy distribution: **93% of missing pixels**, corresponding to 30 galaxies per square arcminute

Example with 93 % of missing data



Galaxy distribution: **93% of missing pixels**, corresponding to 30 galaxies per square arcminute

Missing Data + Noise



Mass Mapping & Cluster A520

A. Peel, F. Lanusse, J.-L. Starck, « SPARSE RECONSTRUCTION OF THE MERGING A520 CLUSTER SYSTEM », **ApJ**, 847, 1, id. 23, 2017.

A puzzling case

Abell 520 (the "cosmic train wreck") is a dynamically complex merging cluster system. Previous weak-lensing studies disagree about the presence of a mysterious dark mass peak—if real, it would challenge our current understanding of dark matter.

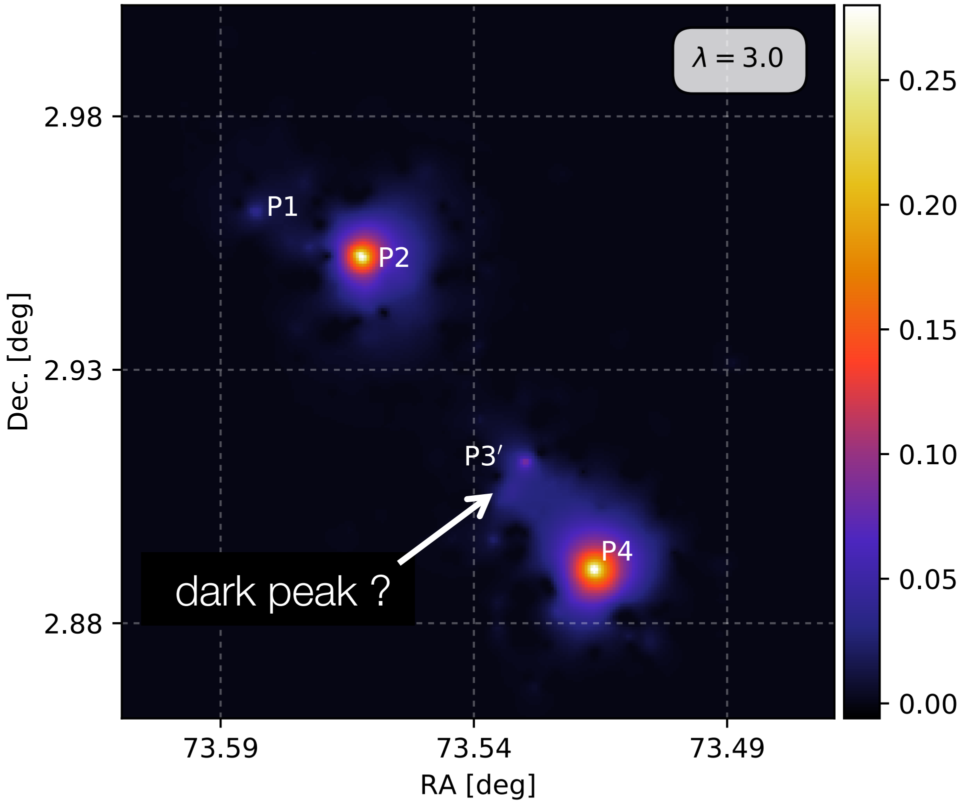
Sparsity-based mass mapping

We generated new mass maps of A520 using Glimpse2D*, a novel technique based on a sparsity prior.

Result

Based on a statistical noise analysis, we cannot confirm the existence of the dark peak.

Glimpse2D mass reconstruction



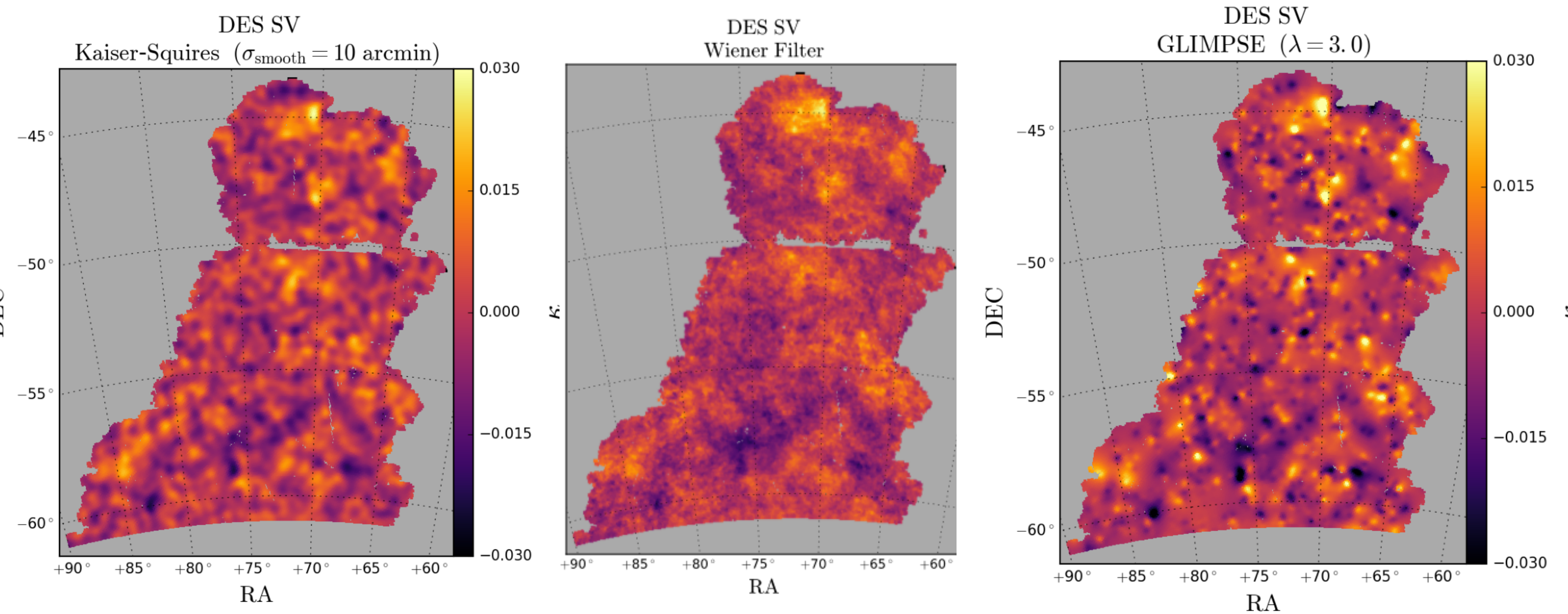
Upper limits on the significance of the P3' structure of **2.3σ and 1.0σ** for the J14 and C12 catalogs,

*http://www.cosmostat.org/software/glimpse

Glimpse2D on public DES SV data

Niall Jeffrey et al. 2018, MNRAS, 479, 3 arXiv:1801.08945

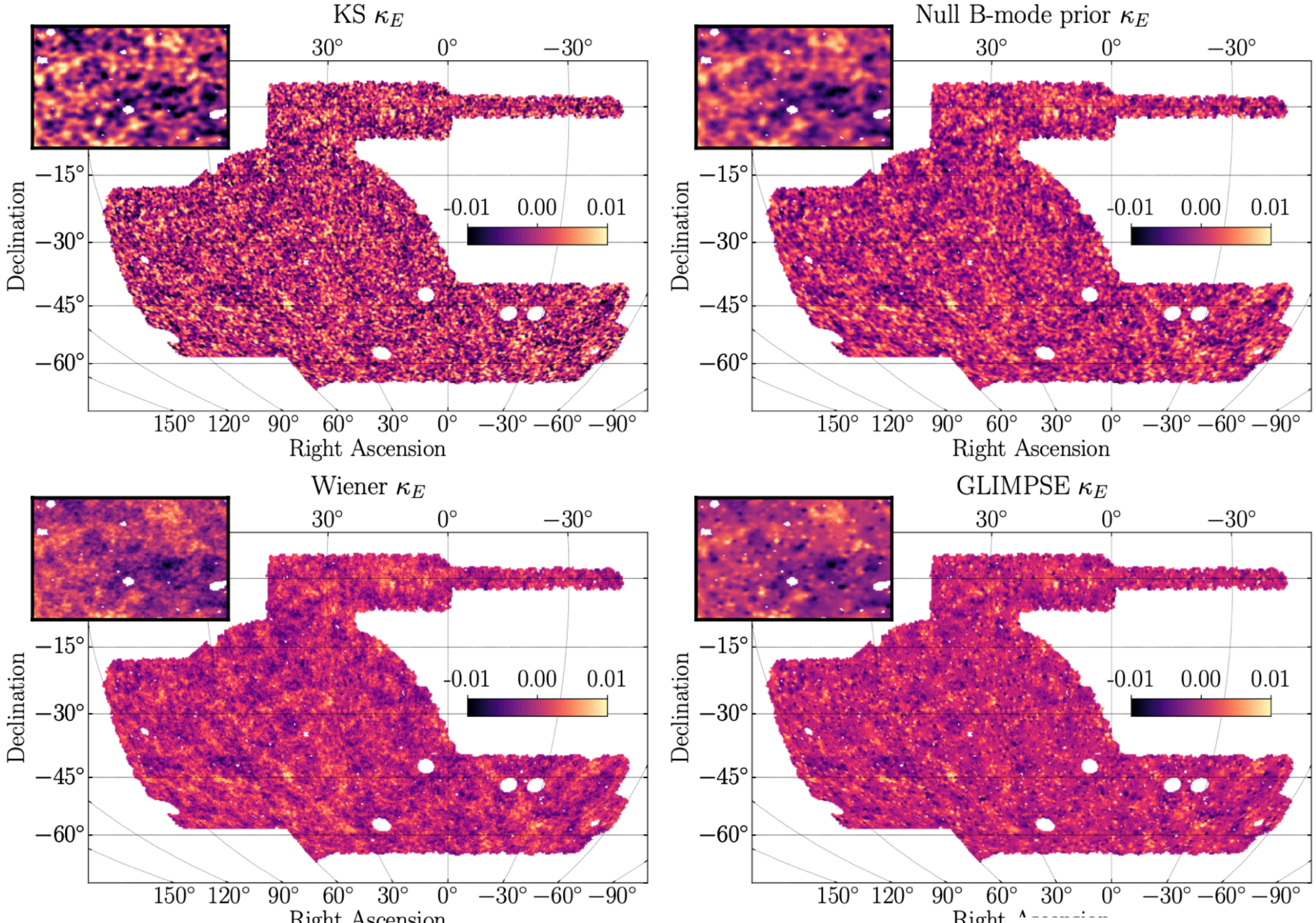
139 deg²



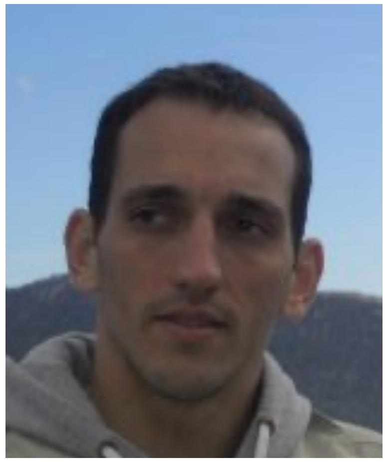
The maximum signal-to-noise value of **peak statistic increased by a factor of 9** using GLIMPSE.

DES Y3 Weak Lensing Mass Map

Jeffrey et al, Dark Energy Survey Year 3 results: curved-sky weak lensing mass map reconstruction, MNRS, 2021

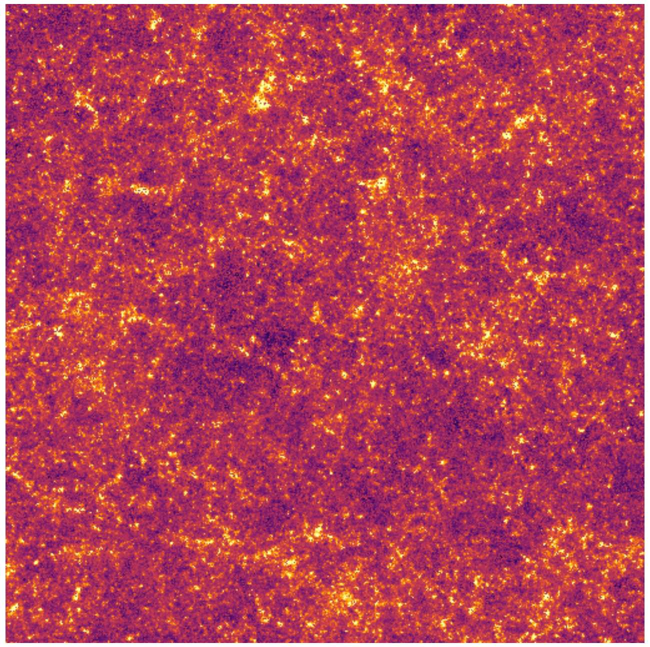


A New Model



$$\boldsymbol{\kappa} = \boldsymbol{\kappa}_{\mathrm{NG}} + \boldsymbol{\kappa}_{\mathrm{G}}.$$

Starck et al, A&A, 2021, https://arxiv.org/abs/2102.04127



$$\min_{\boldsymbol{\kappa}_G, \boldsymbol{\kappa}_{NG}} \|\boldsymbol{\gamma} - \mathbf{A}(\boldsymbol{\kappa}_G + \boldsymbol{\kappa}_{NG})\|^2_{\boldsymbol{\Sigma}_n} + \boldsymbol{C}_{\mathrm{G}}(\boldsymbol{\kappa}_G) + \boldsymbol{C}_{\mathrm{NG}}(\boldsymbol{\kappa}_{NG})$$

MCA (Morphological Component Analysis) performs an alternating minimization scheme:

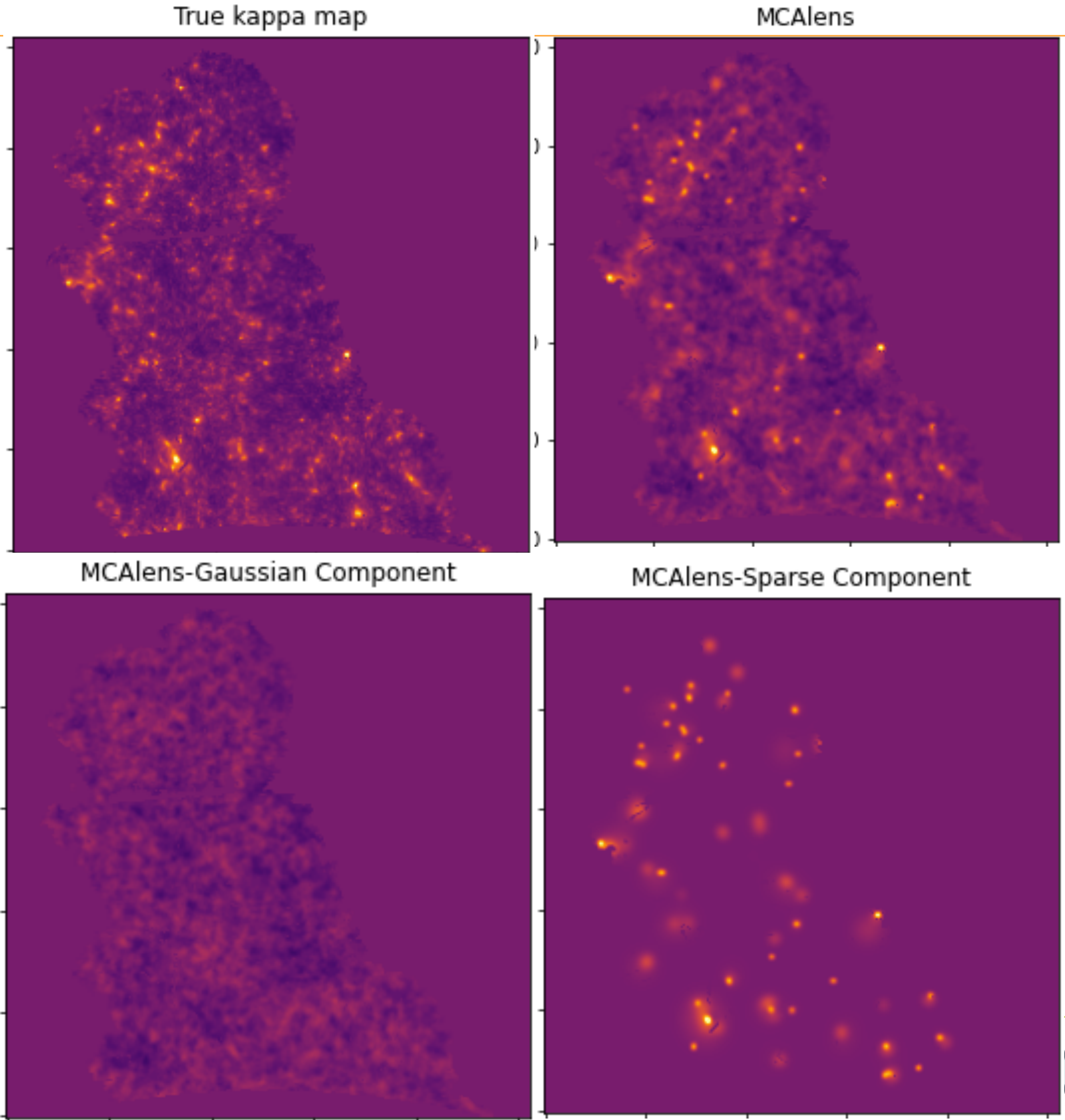
- Estimate $\boldsymbol{\kappa}_G$ assuming $\boldsymbol{\kappa}_{NG}$ is known:

$$\min_{\boldsymbol{\kappa}_G} \left\{ \|(\boldsymbol{\gamma} - \mathbf{A}\boldsymbol{\kappa}_{NG}) - A\boldsymbol{\kappa}_G)\|^2_{\boldsymbol{\Sigma}_n} + \boldsymbol{C}_{\mathrm{G}}(\boldsymbol{\kappa}_G) \right\}. \quad (1)$$

- Estimate $\boldsymbol{\kappa}_{NG}$ assuming $\boldsymbol{\kappa}_G$ is known:

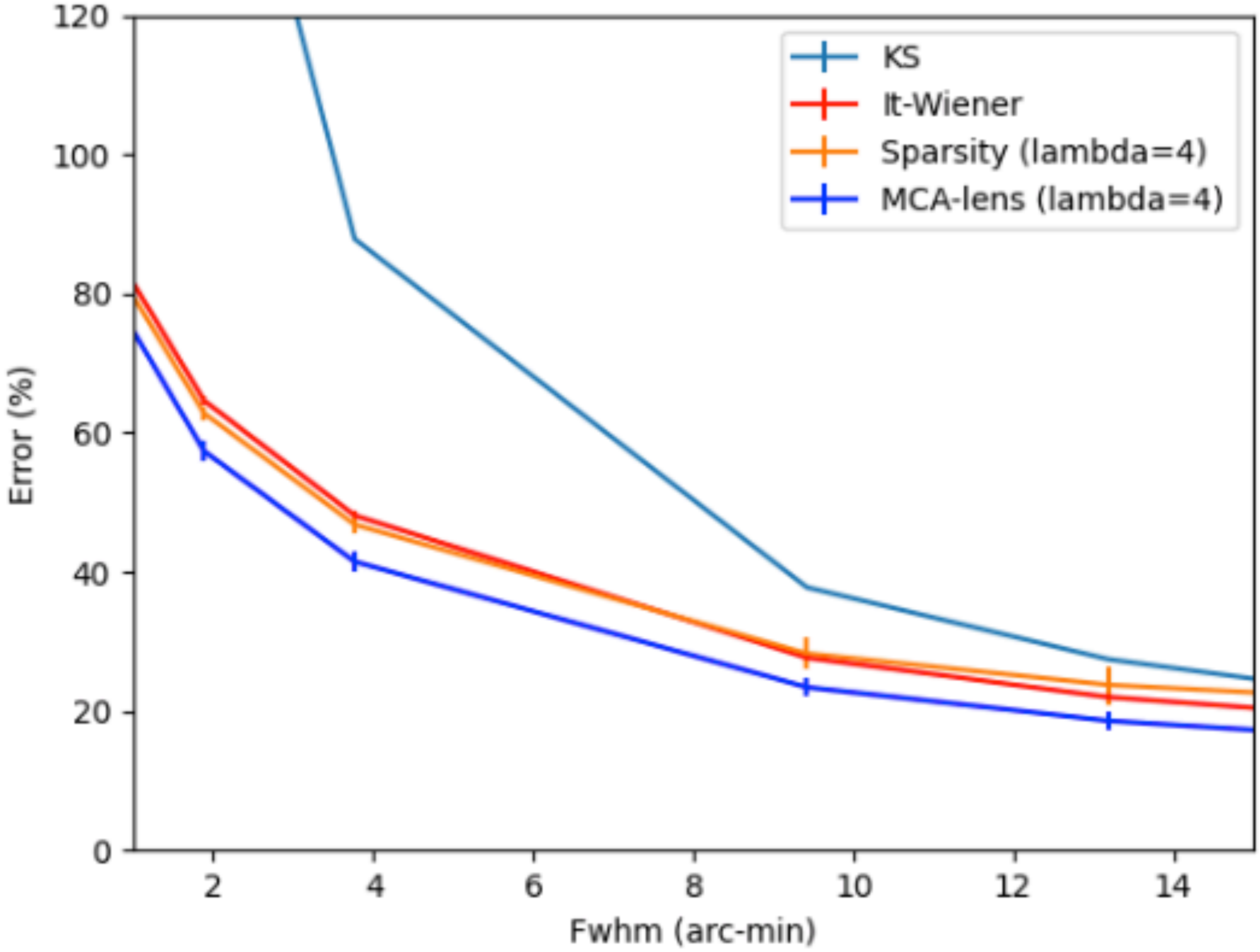
$$\min_{\boldsymbol{\kappa}_{NG}} \left\{ \|(\boldsymbol{\gamma} - \mathbf{A}\boldsymbol{\kappa}_G) - A\boldsymbol{\kappa}_{NG})\|^2_{\boldsymbol{\Sigma}_n} + \boldsymbol{C}_{\mathrm{NG}}(\boldsymbol{\kappa}_{NG}) \right\}. \quad (2)$$

Toy Model

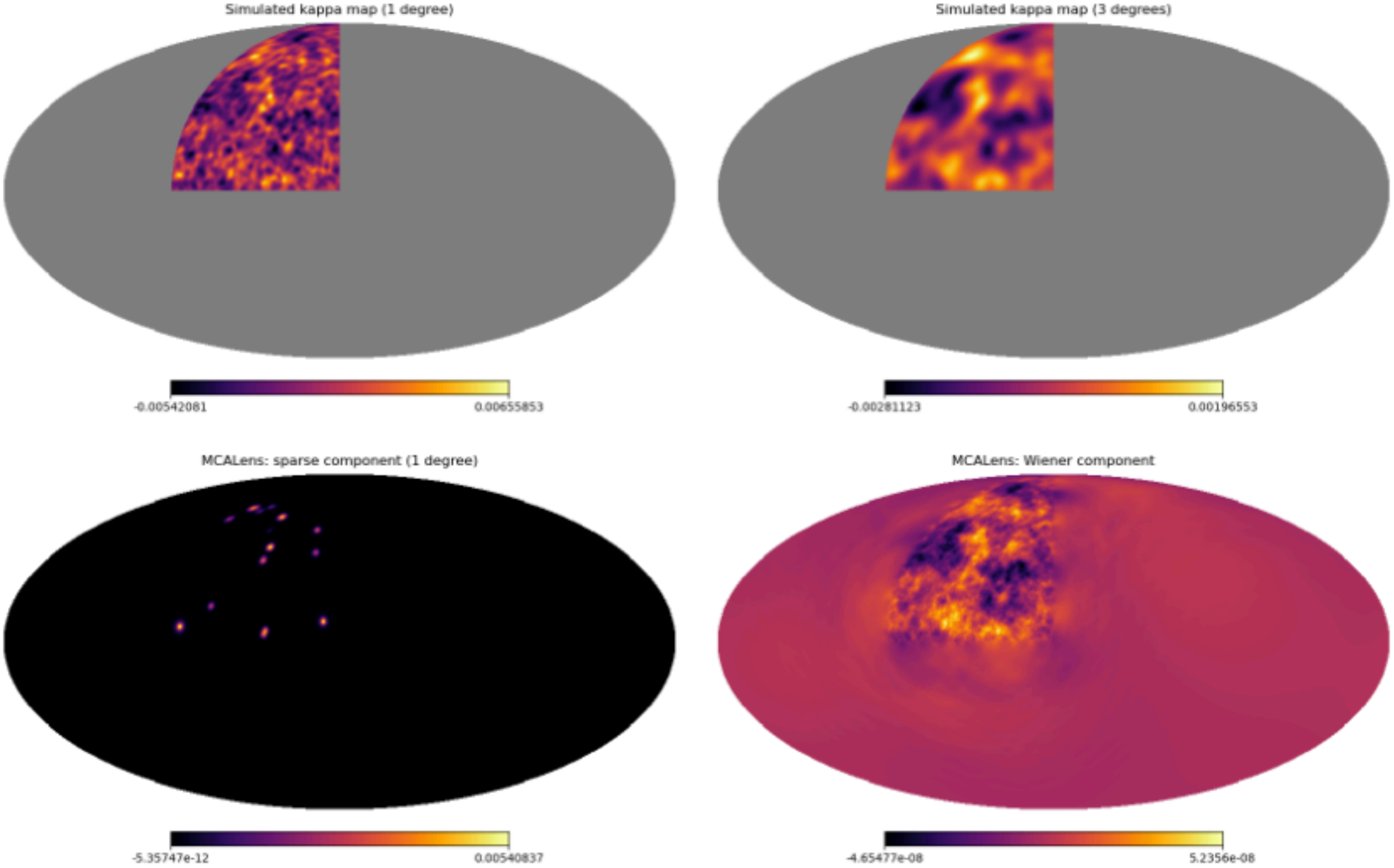


True kappa map
MCAlens
MCAlens-Gaussian Component
MCAlens-Sparse Component

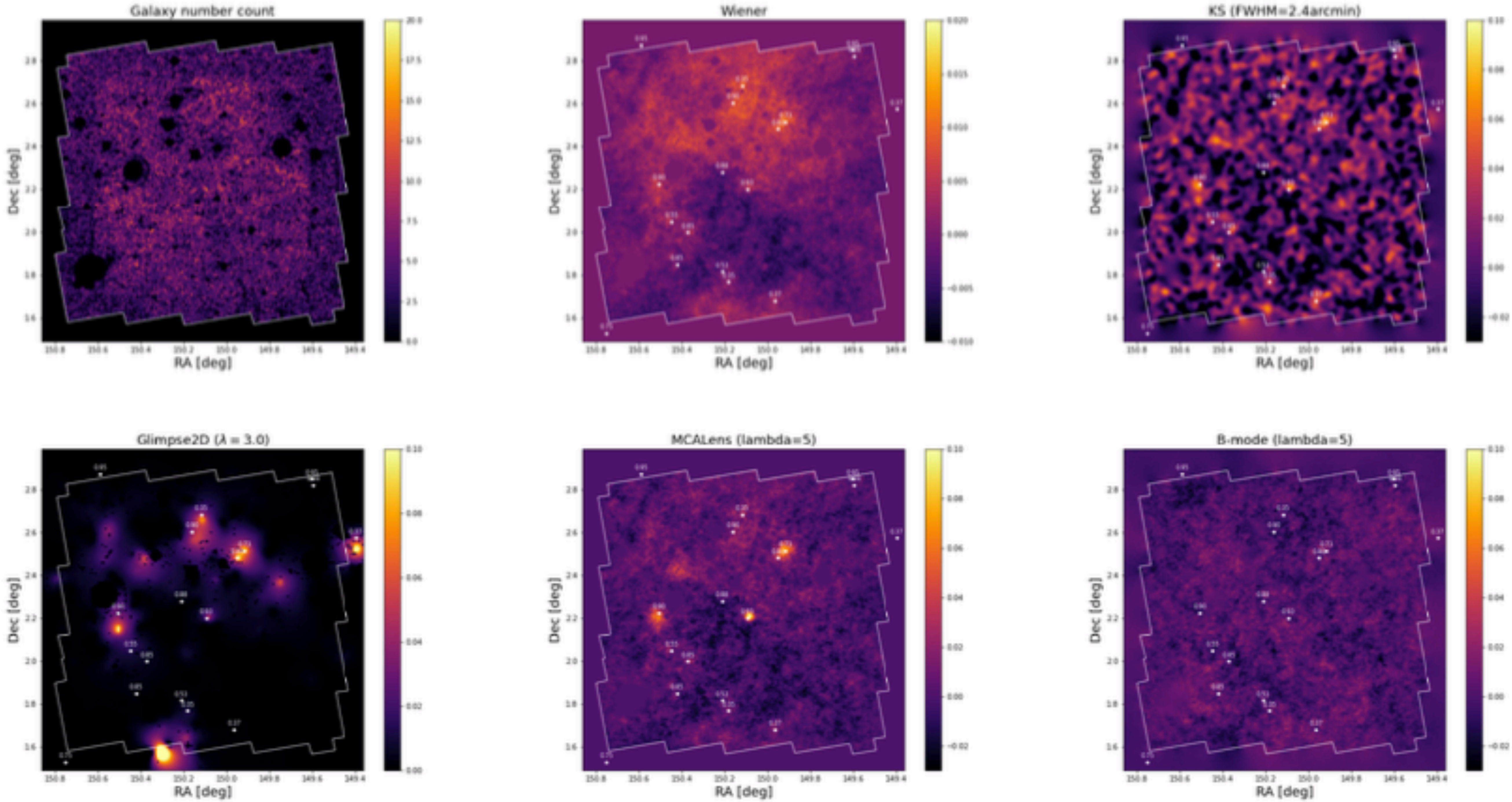
Error versus Scale



Data on the Sphere



COSMOS Data



COSMOS data: Top, galaxies count map, Wiener map, and Kaiser-Squires map smoothed with a Gaussian having a Full Width at Half Maximum of 2.4 arcmin. Bottom, Glimpse, MCAlens and MCAlens B-mode map.

DeepMass: First Deep Learning

N. Jeffrey, F. Lanusse, O. Lahav, J.-L. Starck, "Learning dark matter map reconstructions from DES SV weak lensing data", **Monthly Notices of the Royal Astronomical Society,** 492, 4, 2020.

N. Jeffrey

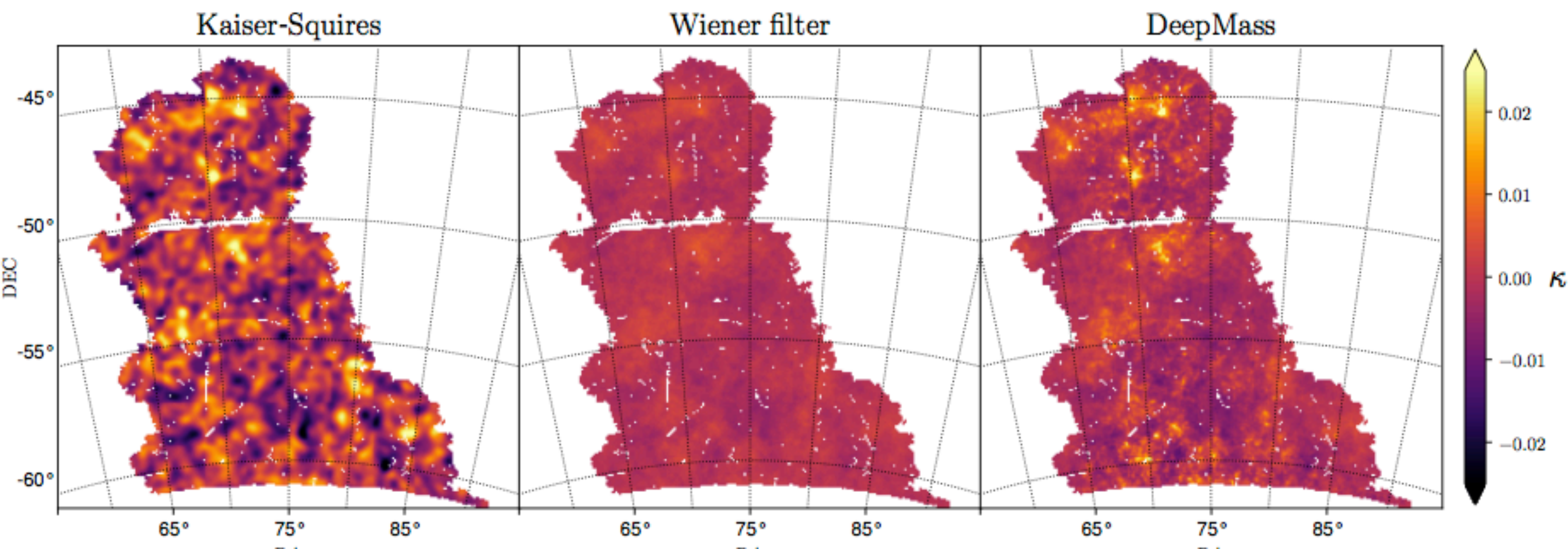
F. Lanusse

Approximate the function shear to convergence by a Convolutional Neural Networks $\mathcal{F}_\Theta(\cdot)$

Parameters Θ are learned, minimizing the Mean squared error:

$$J(\Theta) = \| \mathcal{F}_\Theta(\gamma) - \boldsymbol{\kappa}_{\text{true}} \|$$

Mean posterior estimate: $\hat{\boldsymbol{\kappa}} = \mathcal{F}_\Theta(\gamma) = \int \boldsymbol{\kappa} P(\boldsymbol{\kappa}|\boldsymbol{\gamma}) d\boldsymbol{\kappa}$



Mass Mapping in Deep Learning Era

B. Remy, F. Lanusse, Z. Ramzi, J. Liu, N. Jeffrey and J.-L. Starck, "Probabilistic Mapping of Dark Matter by Neural Score Matching", Machine Learning and the Physical Sciences Workshop, NeurIPS 2020.
arXiv: https://arxiv.org/abs/2011.08271, Code: https://github.com/CosmoStat/jax-lensing

Bayesian Deep Learning approach

==> Learn the prior from the data
==> Sampling with Annealed Hamiltonian Monte Carlo

Whether you are looking for the MAP or sampling with HMC, you only need access to the **score** of the posterior:

$$\frac{d \log p(x|y)}{dx}$$

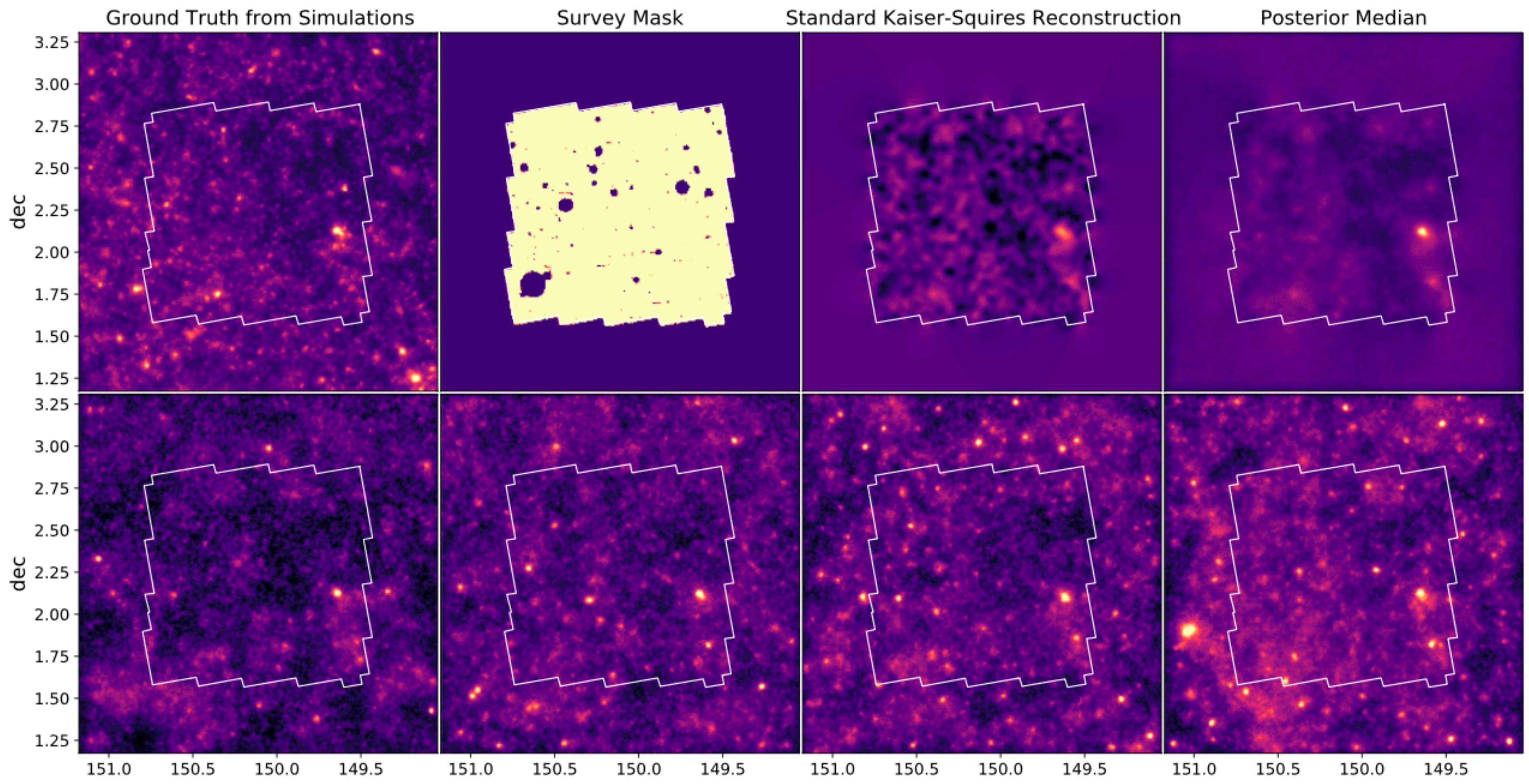
The score of the full posterior is:

$$\nabla_x \log p(x|y) = \underbrace{\nabla_x \log p(y|x)}_{\text{known}} + \underbrace{\nabla_x \log p(x)}_{\text{can be learned}}$$

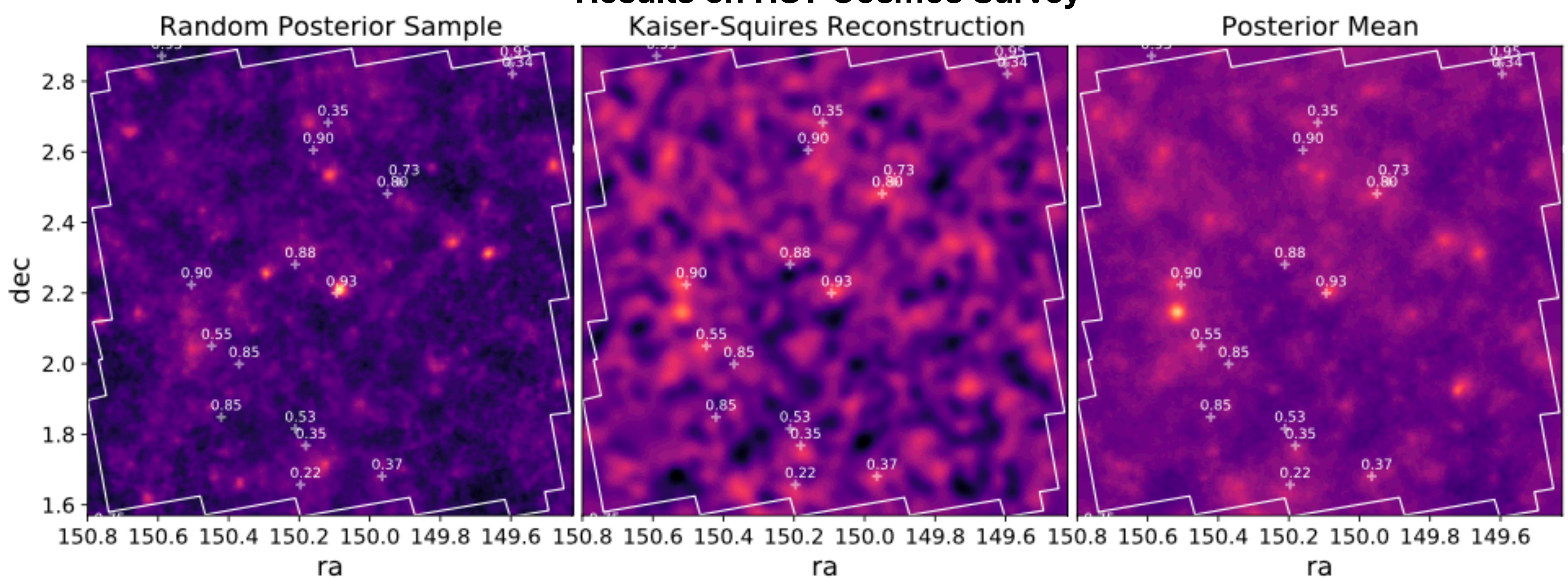
Mass Mapping in Deep Learning Era

Results on simulations

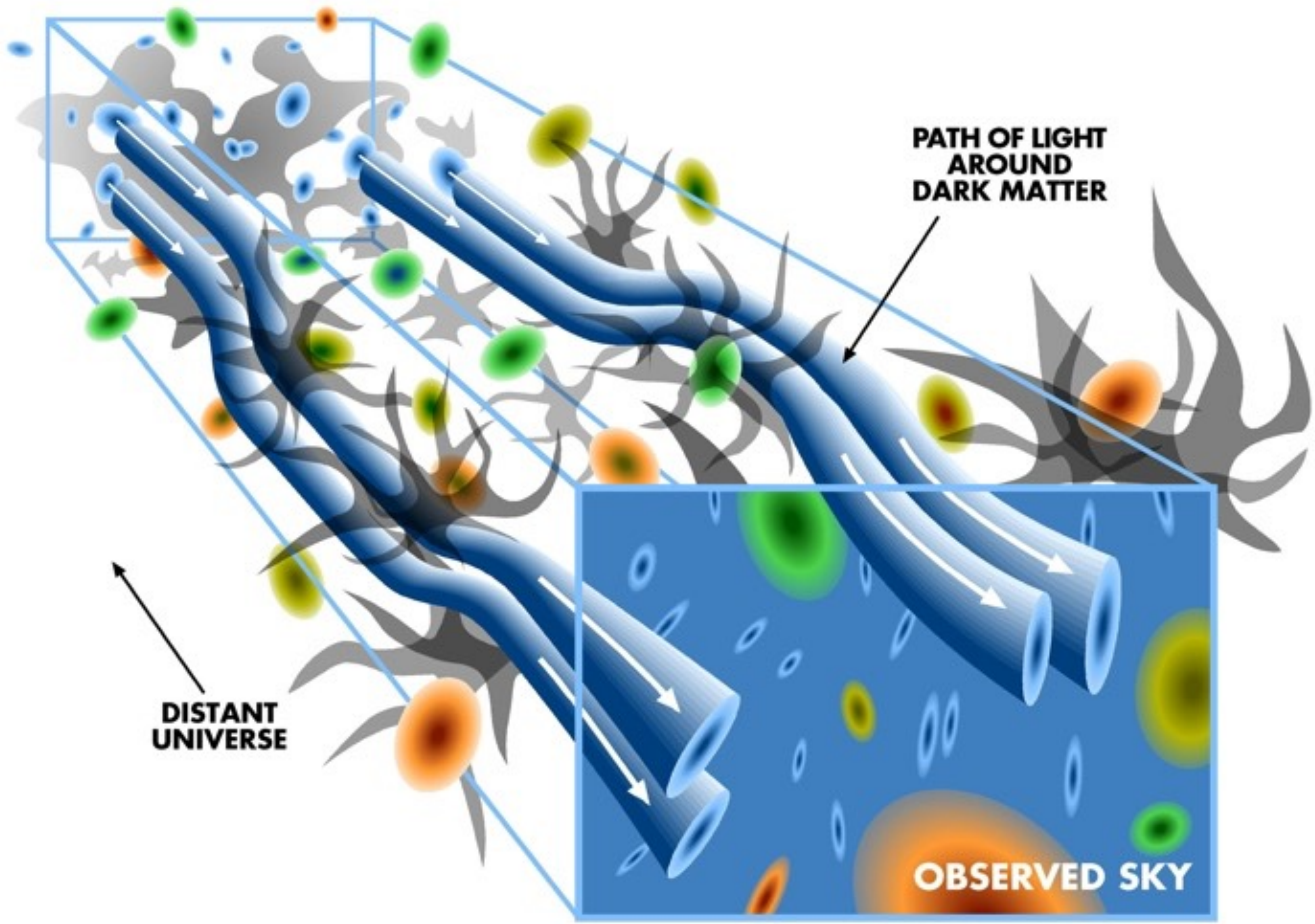
Posterior Samples

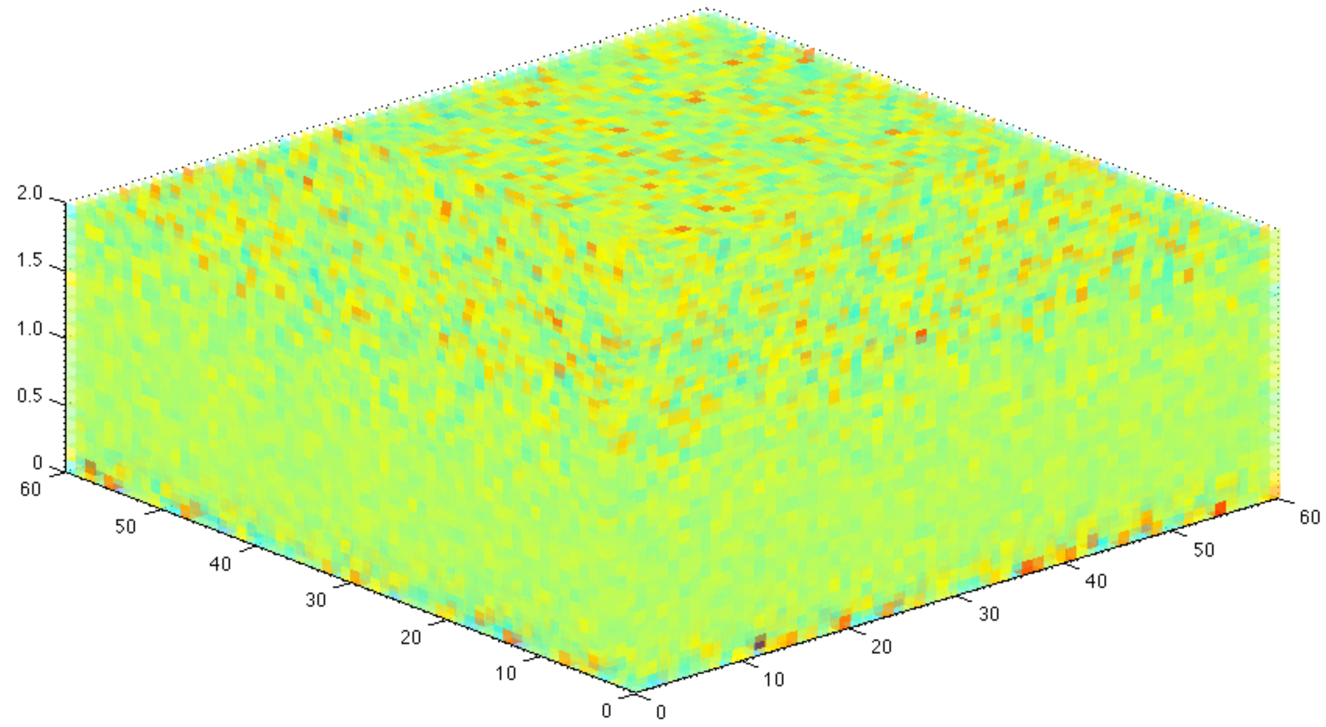


Results on HST Cosmos Survey



3D Weak Lensing




2.0
1.5
1.0
0.5
0
60
50
40
30
20
10
0
0
10
20
30
40
50
60

3D Mass Mapping

$$\gamma(\boldsymbol{\theta}) = \frac{1}{\pi} \int d^2\boldsymbol{\theta}' \mathcal{D}(\boldsymbol{\theta} - \boldsymbol{\theta}')\kappa(\boldsymbol{\theta}')$$

Kappa (or convergence) is a dimensionless surface mass density of the lens

$$\kappa(\boldsymbol{\theta}, w) = \frac{3H_0^2\Omega_M}{2c^2} \int_0^w dw' \frac{f_K(w')f_K(w-w')}{f_K(w)} \frac{\delta[f_K(w')\boldsymbol{\theta}, w']}{a(w')} ,$$

f_K is the angular diameter distance, which is a function of the comoving radial distance r and the curvature K.

$$\boldsymbol{\gamma} = \mathbf{P}_{\gamma\kappa}\boldsymbol{\kappa} + \boldsymbol{n}_\gamma ,$$

$$\kappa = Q\delta + n$$

$$\gamma = \mathbf{R}\delta + n$$

- Galaxies are not intrinsically circular: intrinsic ellipticity ~ 0.2-0.3; gravitational shear ~ 0.02
- Reconstructions require knowledge of distances to galaxies

Linear Methods

✧Assume uncorrelated Gaussian noise*

✧Linear methods

✧Wiener/inverse variance filter (Simon et al., 2009)

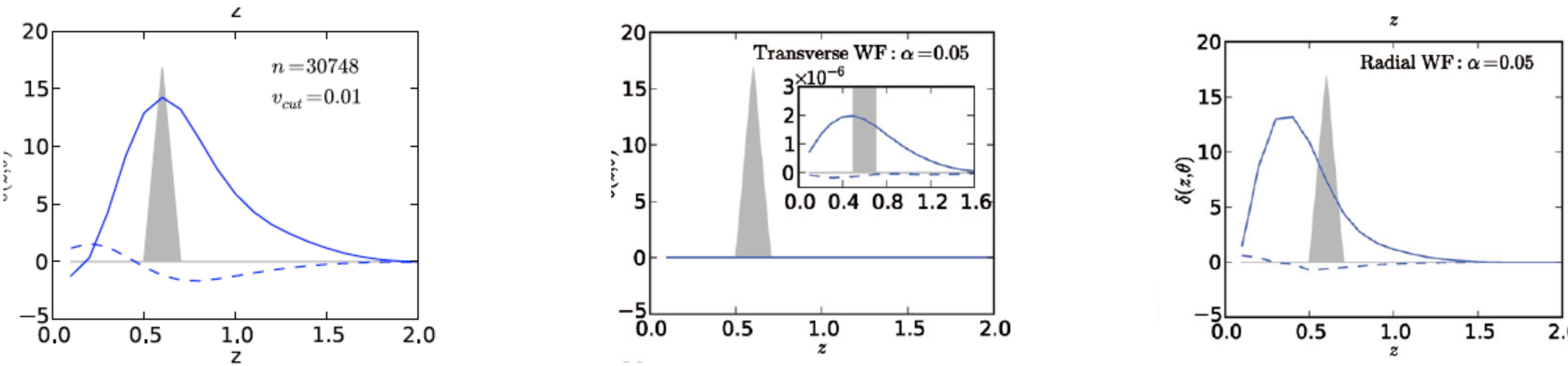
✧ $$\hat{s}_{MV} = [\alpha\mathbf{1} + \mathbf{S}\mathbf{R}^\dagger\mathbf{\Sigma}^{-1}\mathbf{R}]^{-1}\mathbf{S}\mathbf{R}^\dagger\mathbf{\Sigma}^{-1}d \ .$$

✧SVD decomposition & thresholding (VanderPlas et al., 2011)

$$\hat{s}_{IV} = \mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{U}^\dagger\mathbf{\Sigma}^{-1/2}d \ ,$$

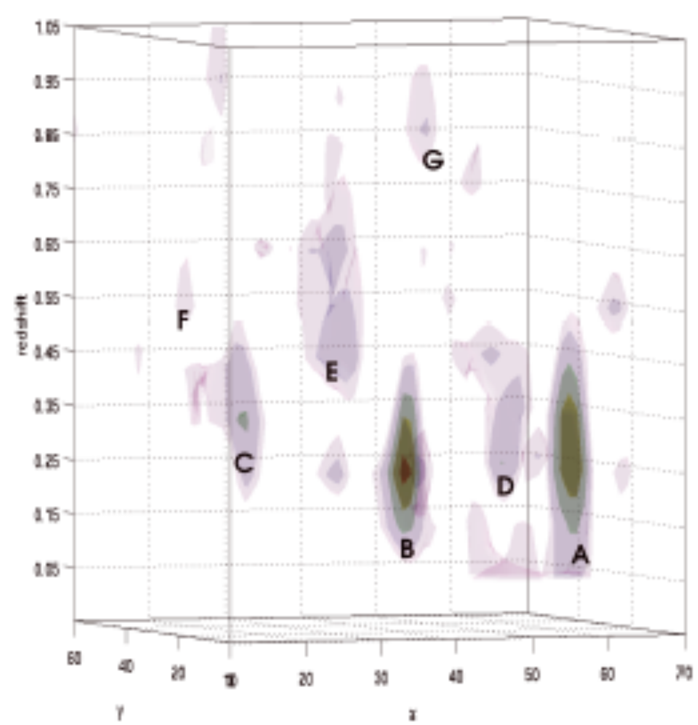
Reconstruction resolution limited by resolution of data

Linear Methods

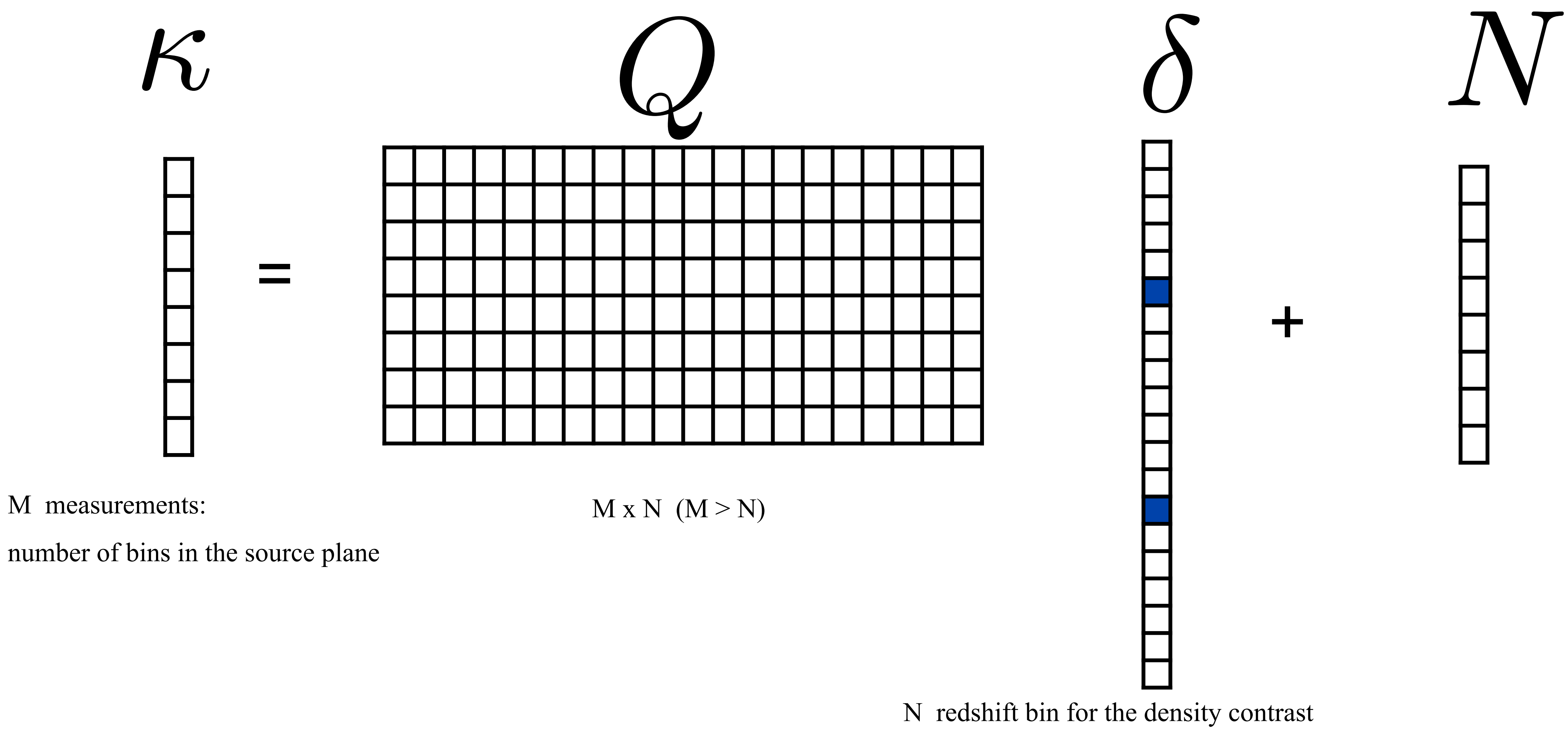


Target Areas for Improvement

- ✧Redshift bias in location of detected peaks
- ✧Smearing along the line of sight
- ✧Damping of the reconstruction
- ✧Sensitivity at high redshift
- ✧Improving resolution in reconstructions



3D Mass Mapping



δ is sparse.

Q spreads out the information in δ along κ bins.

More unkown than measurements

Weak Lensing & 3D Matter Distribution

A. Leonard, F.X. Dupe, and J.-L. Starck, "A Compressed Sensing Approach to 3D Weak Lensing", **Astronomy and Astrophysics**, 539, A85, 2012.

A. Leonard, F. Lanusse, J-L. Starck, GLIMPSE: Accurate 3D weak lensing reconstruction usiing sparsity, Astronomy and Astrophysics, 2014

$$\gamma = P\kappa$$
$$\kappa = Q\delta$$

$$\delta = \Phi\alpha \quad \rightarrow \quad \gamma = PQ\Phi\alpha = R\Phi\alpha$$

$$R = PQ \qquad \alpha$$

$$\gamma$$

Shear Measurements
M measurements: number of bins in the source plan x number of pixels at for given bin

Density constrast wavelet coefficients

Related to Compressed Sensing theorem
==> Use Sparse recovery and Proximal optimization theory

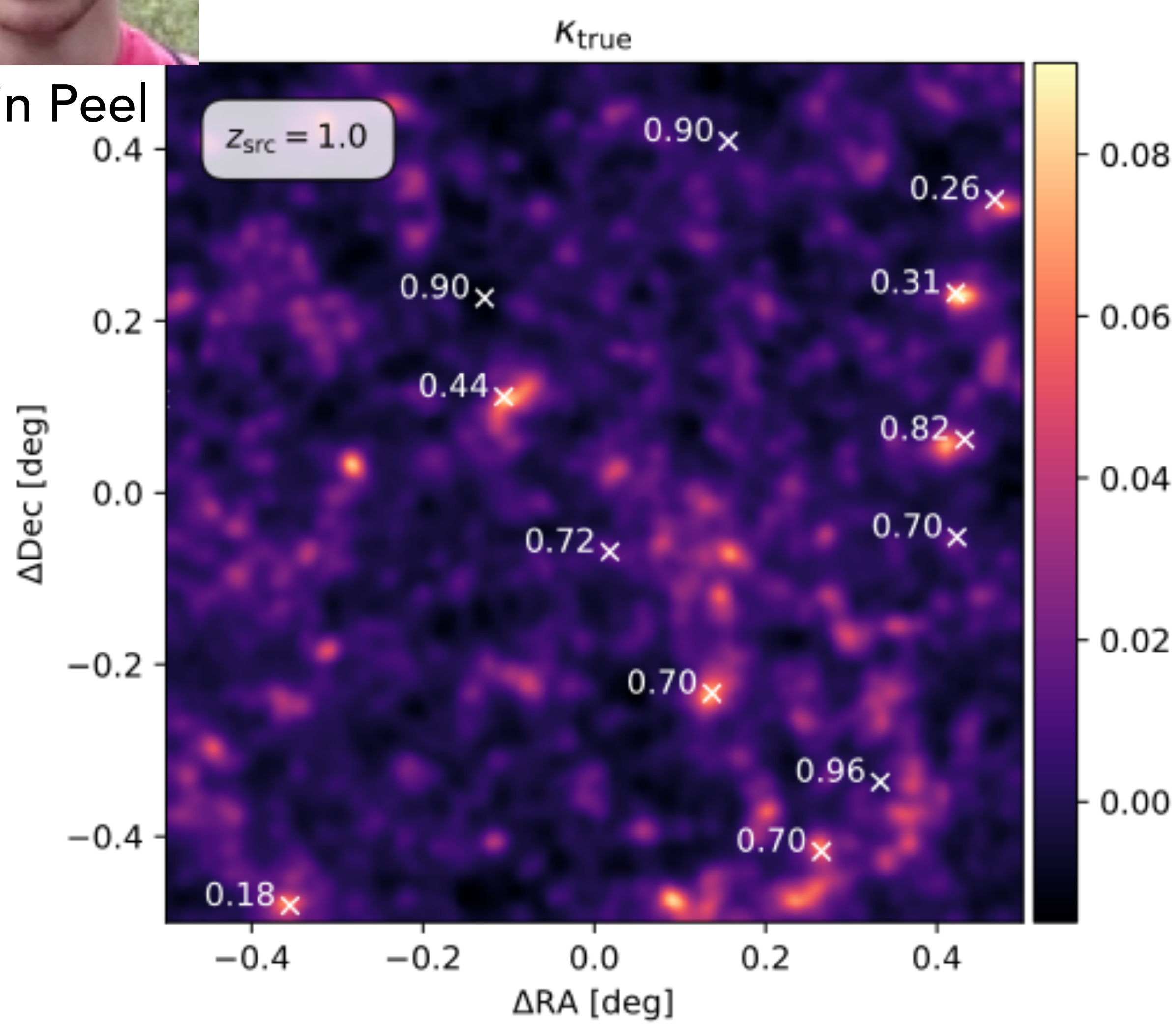
$$\min_{\alpha} \parallel \alpha \parallel_1 \quad s.t. \ \frac{1}{2} \parallel \gamma - R\Phi\alpha \parallel^2_{\mathbf{\Sigma}^{-1}} \leq \epsilon$$

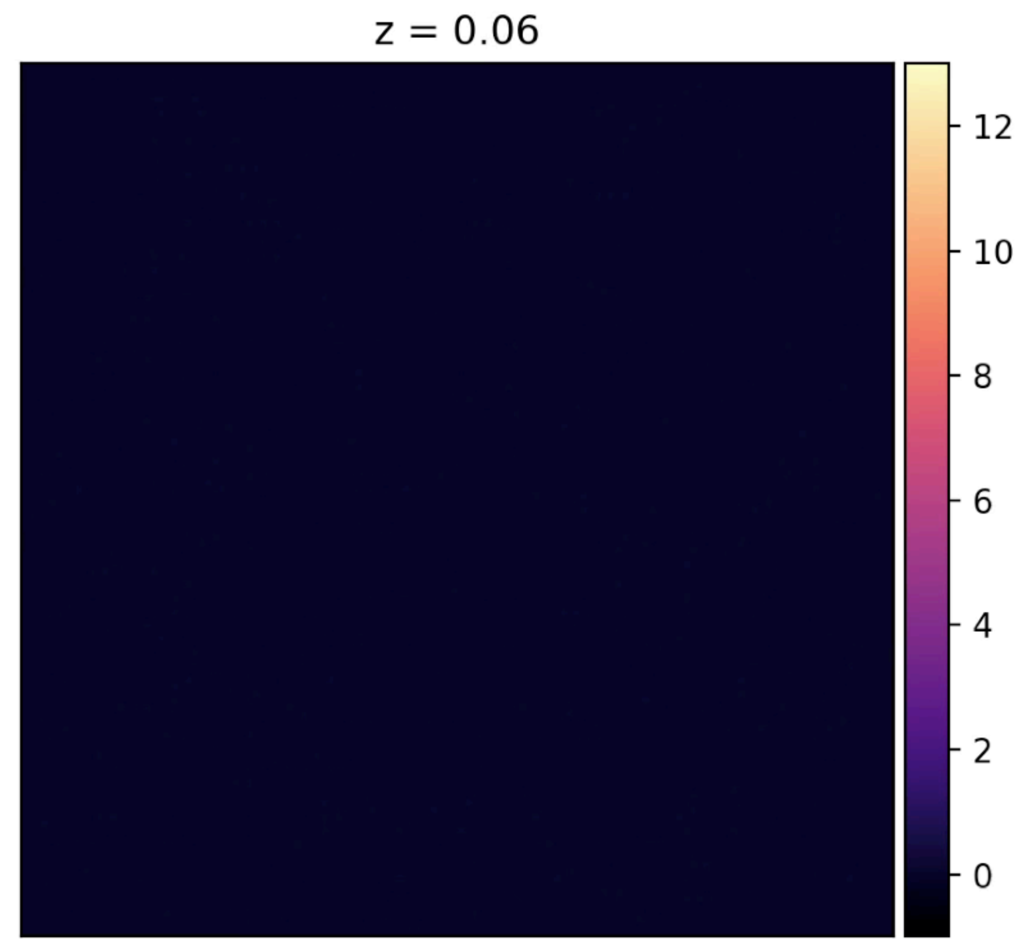
$\delta = \Phi\alpha$ $\qquad \Phi$ = 2D Wavelet Transform on each redshift bin

Glimpse3D on Euclid calibration mock

Austin Peel

1 deg² example tile





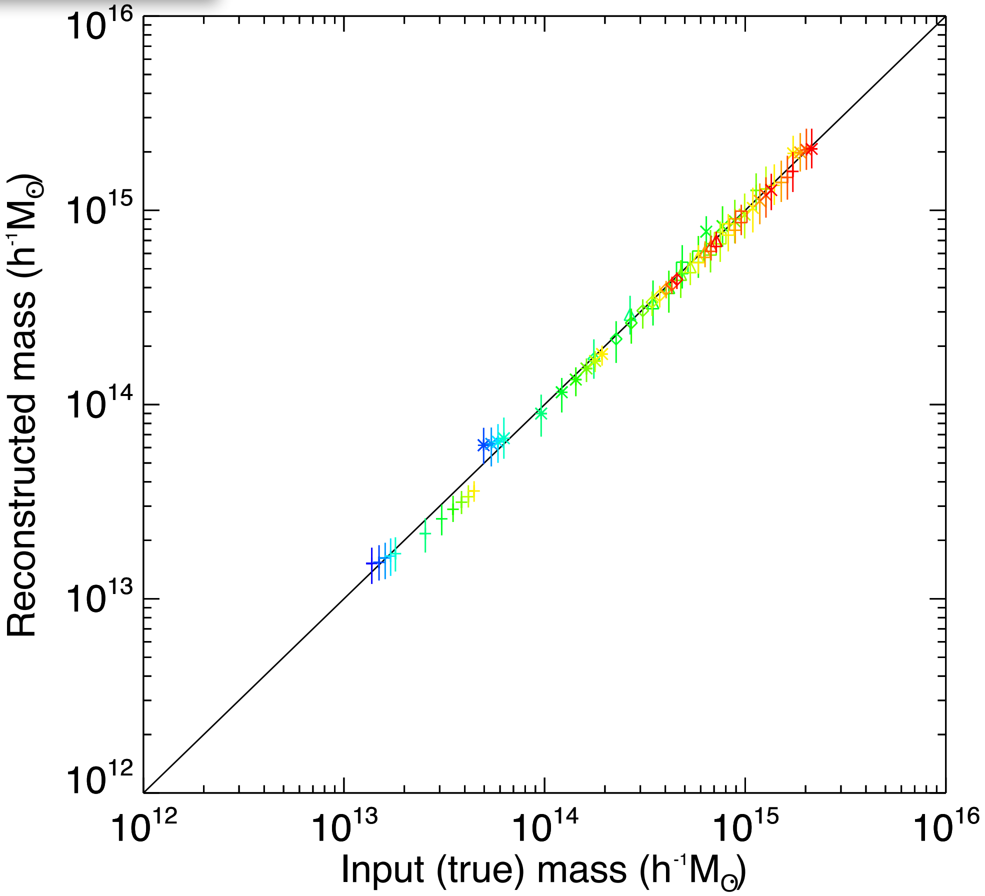
1 pix = **0.23** arcmin

no Glimpse3D smoothing

density reconstruction on redshift slices

Mass Estimation

Cluster Masses from 3D Weak

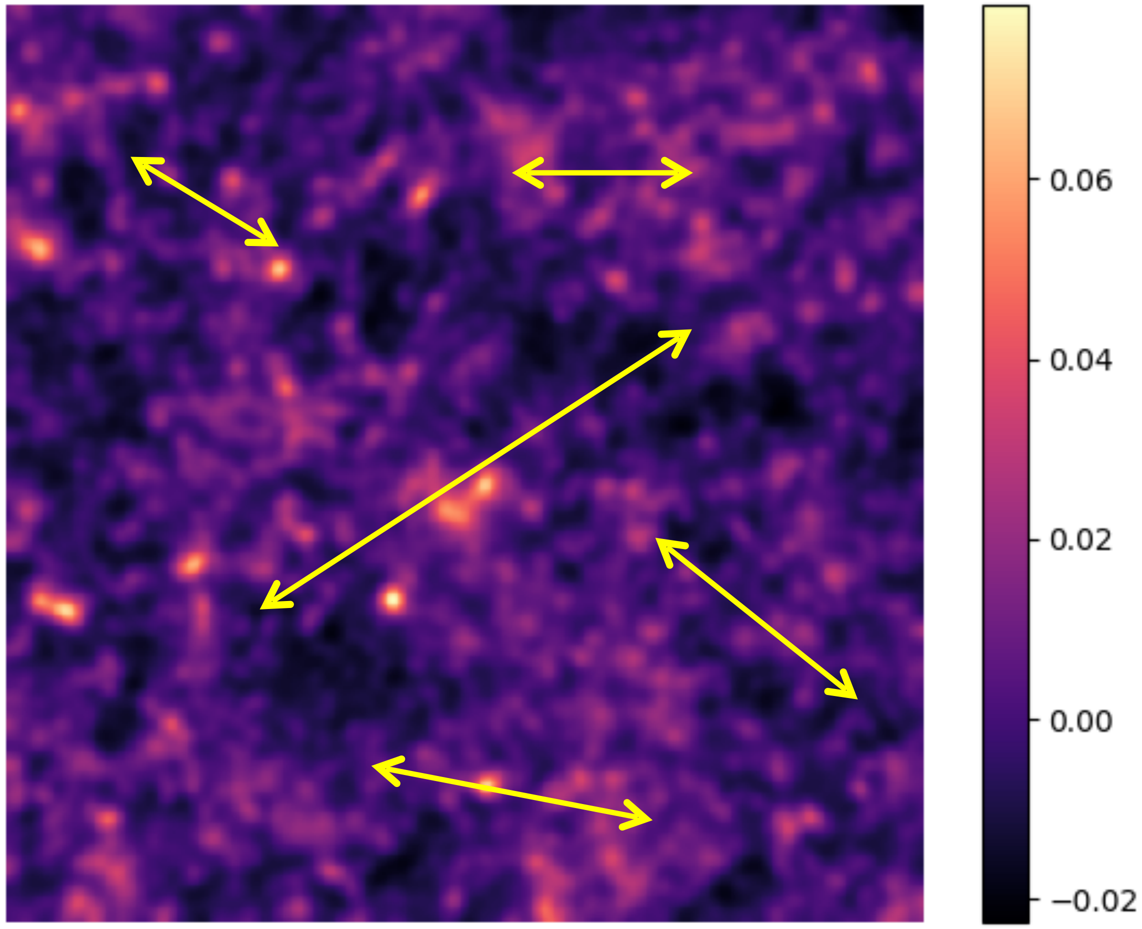


- GLIMPSE 3D reconstructions provide a direct, unbiased & nonparametric estimate of the cluster mass (Leonard, Lanusse & Starck 2014, MNRAS, 440, 1281)
- Masses estimated integrating the density in the central 4 x 4 arcmin
- Error bars reflect the standard deviation in mass estimates 1000 Monte Carlo simulations of each cluster
- Cluster masses $2 \times 10^{13} h^{-1} M_{\odot} \leq M_{vir} \leq 10^{15} h^{-1} M_{\odot}$
- Cluster redshifts $0.05 \leq z \leq 0.75$

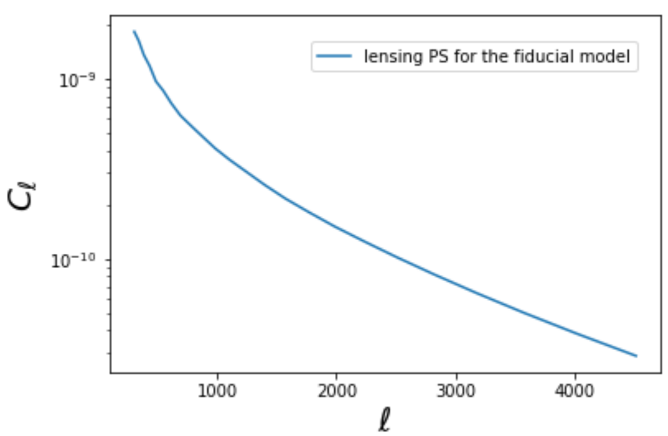
•A. Leonard, F. Lanusse, & J.-L. Starck, "Weak lensing reconstructions in 2D & 3D: implications for cluster studies", MNRAS, 449, 1146–1157, 2015.
•A. Leonard, F. Lanusse & J.-L. Starck, A&A, “GLIMPSE: Accurate 3D weak lensing reconstructions using sparsity”, 2014.

Cosmology with Maps

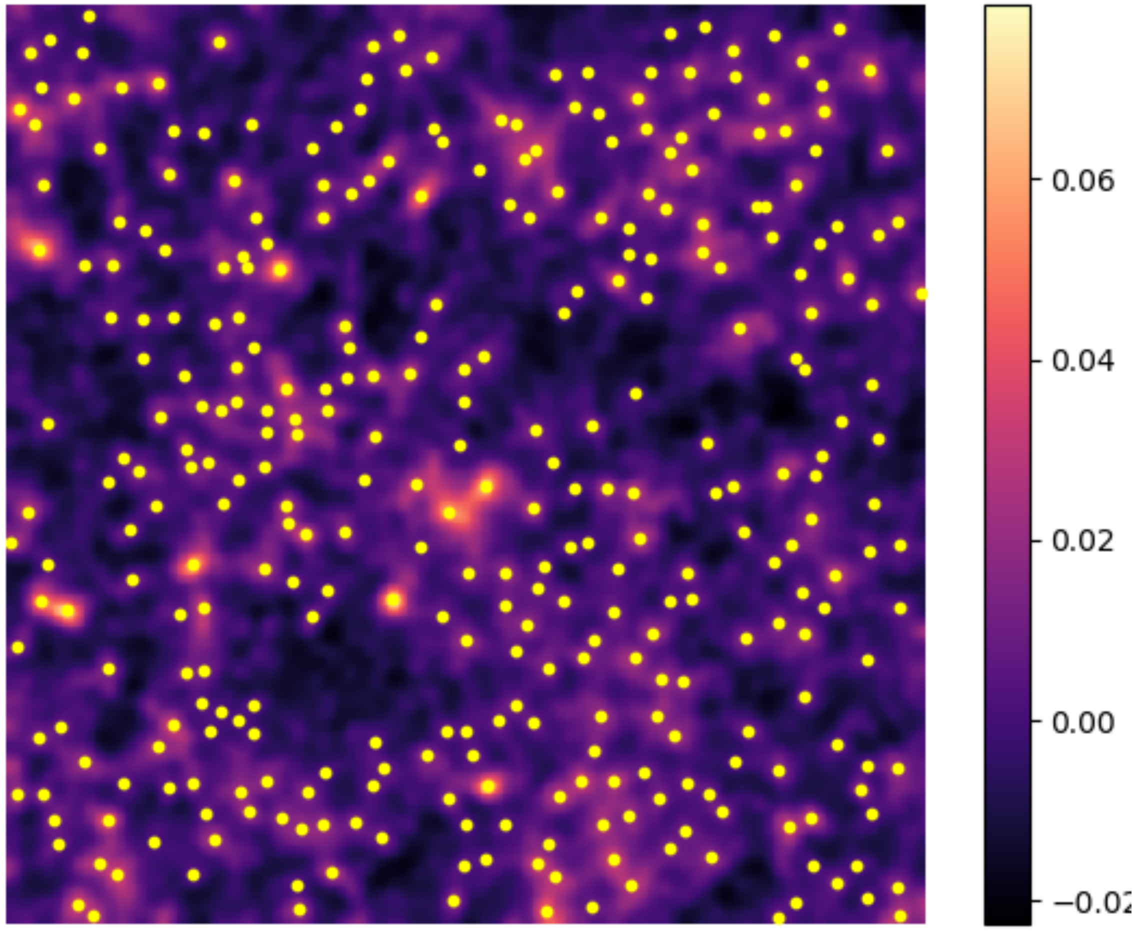
Second Order



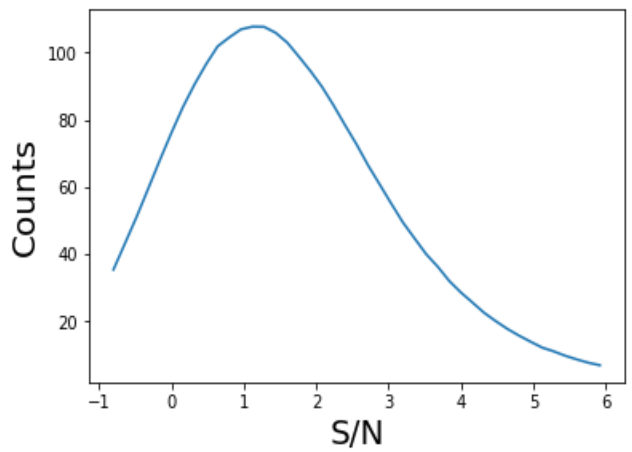
Power spectrum



Higher Order

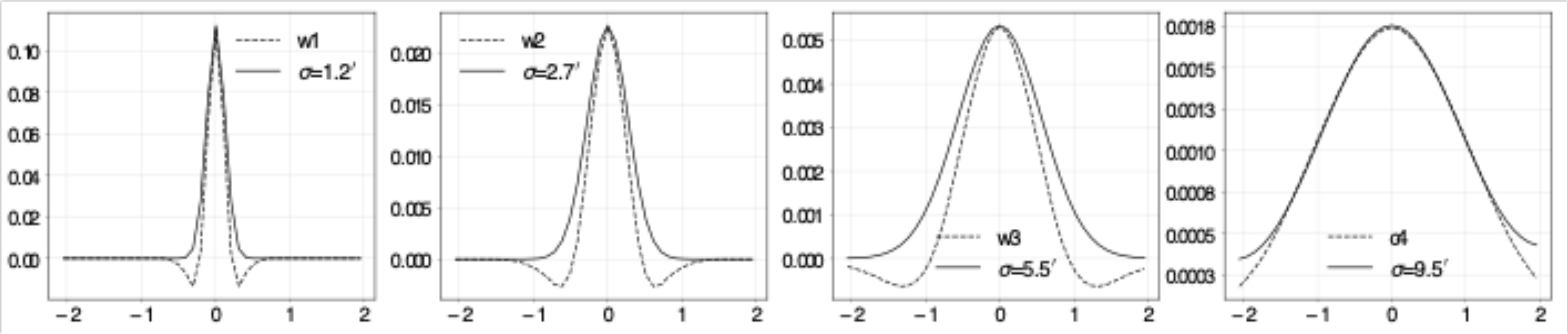


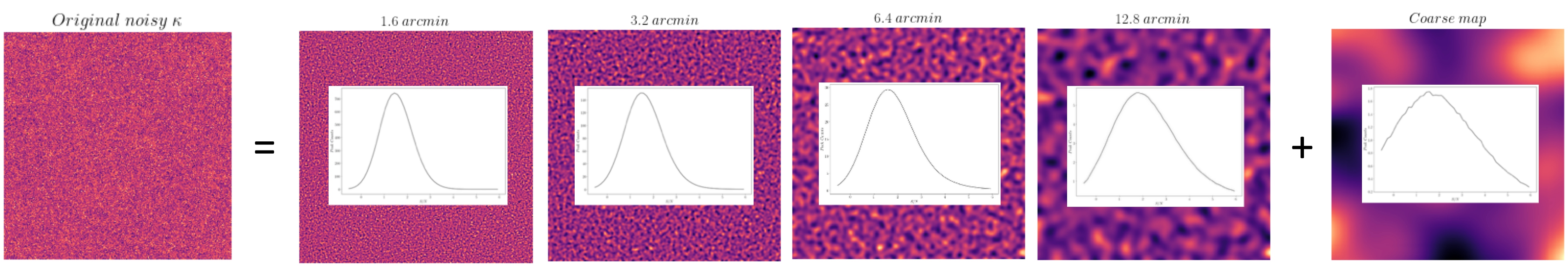
Peak counts



Wavelet Peaks

https://arxiv.org/abs/2001.10993 V. Ajani, A. Peel, V. Pettorino, J-L. Starck, Z. Li, J. Liu, Phys. Rev. D 102, 103531, (2020)





Wavelet Peaks

V. Ajani

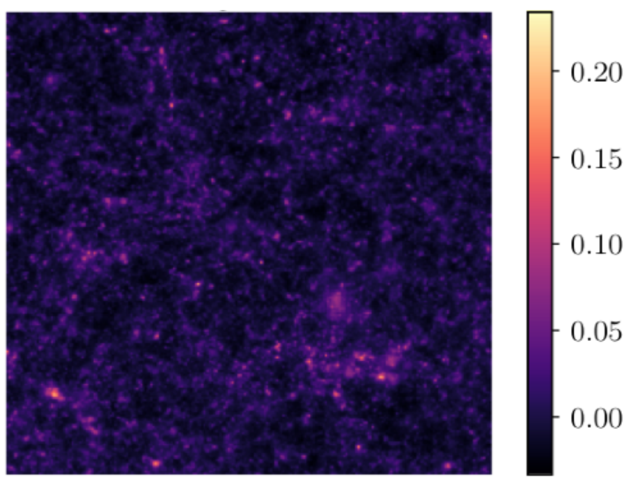
Convergence Map from **MassiveNus** simulations

(http://columbialensing.org)

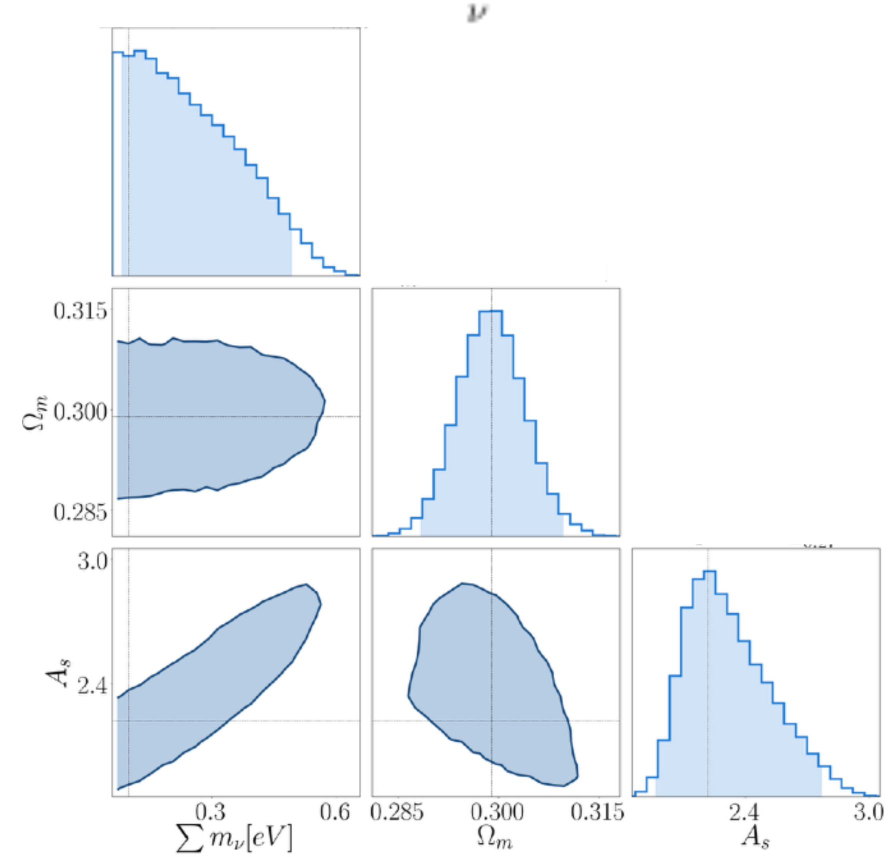
100 cosmological models
10000 realisations

Model $\sum_\nu m_\nu, \quad \Omega_m, \quad A_s$

Mock data



$$\log \mathcal{L}(\theta) = -\frac{1}{2}(d - \mu(\theta))^T C^{-1}(d - \mu(\theta))$$

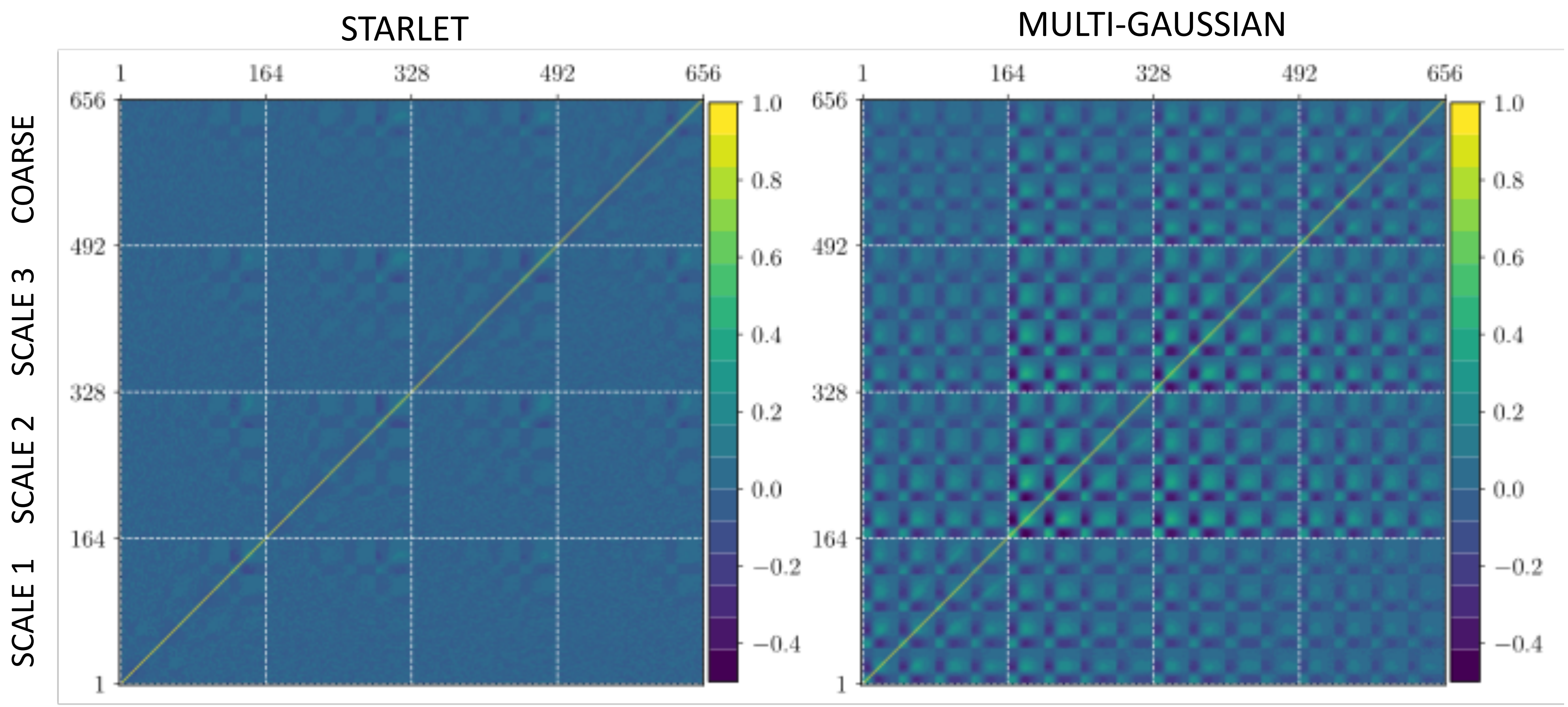


Constraints on parameters

*Convergence Map from **MassiveNus** simulations (http://columbialensing.org)

Wavelet Peaks

V. Ajani

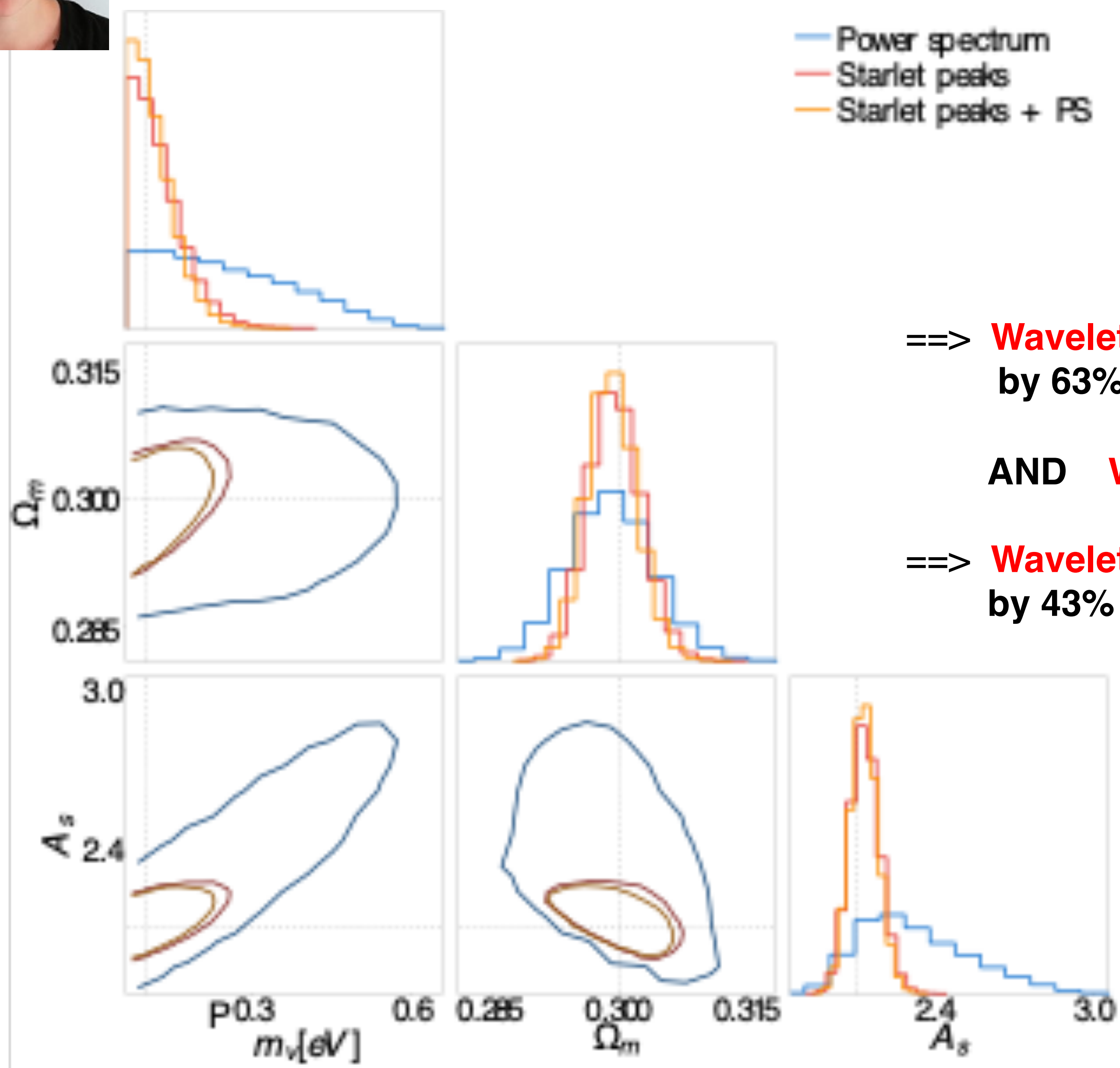


Starlet filter tends to make the covariance matrix more diagonal

https://arxiv.org/abs/2001.10993 Ajani et al, Phys. Rev. D 102, 103531, (2020)

Wavelet Peaks

V. Ajani, A. Peel, V. Pettorino, J.-L. Starck, Z. Li, J. Liu, "**Constraining neutrino masses with weak-lensing starlet peak counts**", Physical Review D, 102, 103531, 2020, DOI: 10.1103/PhysRevD.102.103531, [arXiv:2001.10993].



==> **Wavelet peak count > power spectrum,** **by 63% on M_ν, 40% on Ω_m, 72% on A_s.**

AND Wavelet peak count + power spectrum = Wavelet peak count

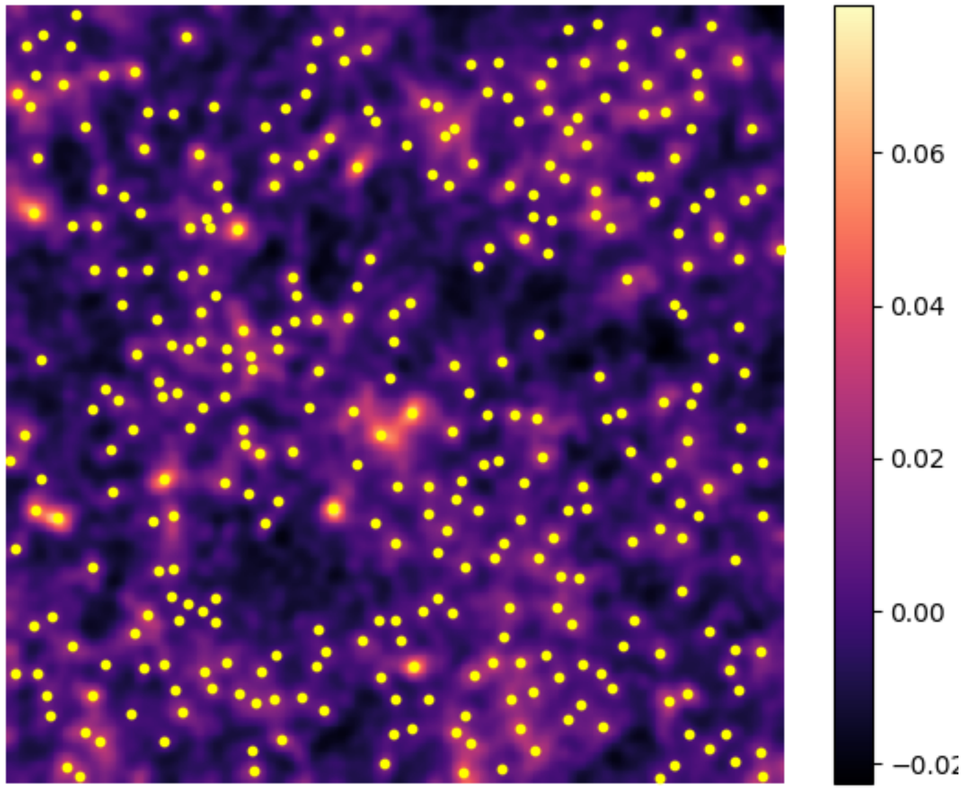
==> **Wavelet peak count > mono-scale peaks,** **by 43% on M_ν, 25% on Ω_m, 34% on A_s.**

Multi-scale peaks significantly outperfom single-scale peaks and power spectrum

Multi-scale peaks alone perform as well as multi-scale peaks + power spectrum

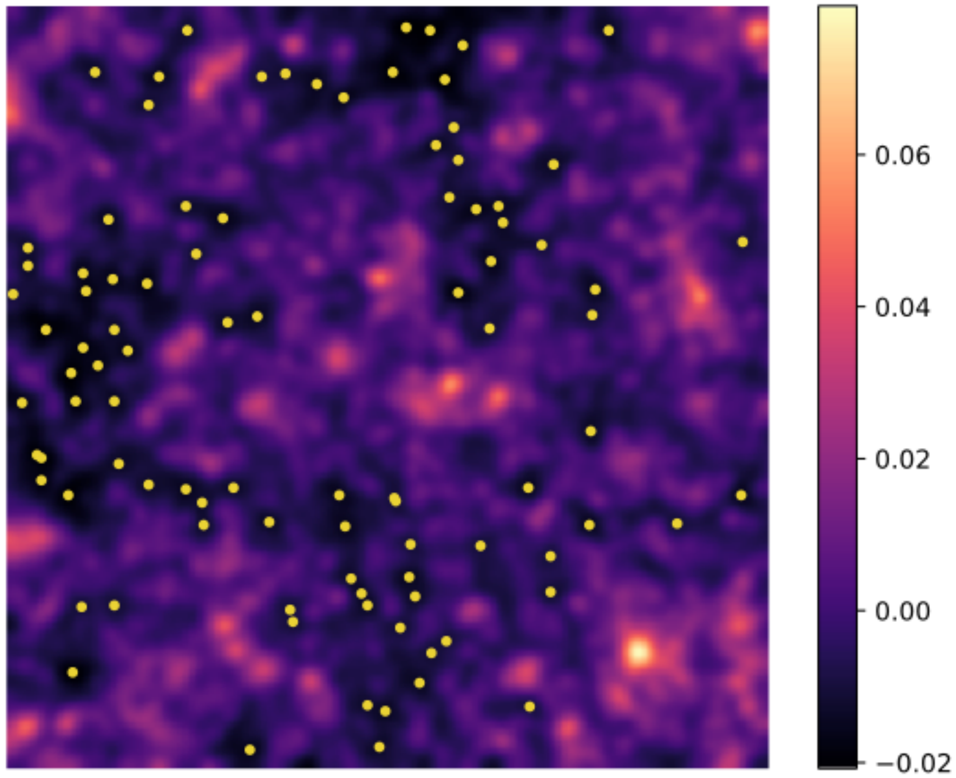
Wavelet Peaks and Voids

Peak counts



Local maxima in the map

Void counts



Local minima in the map

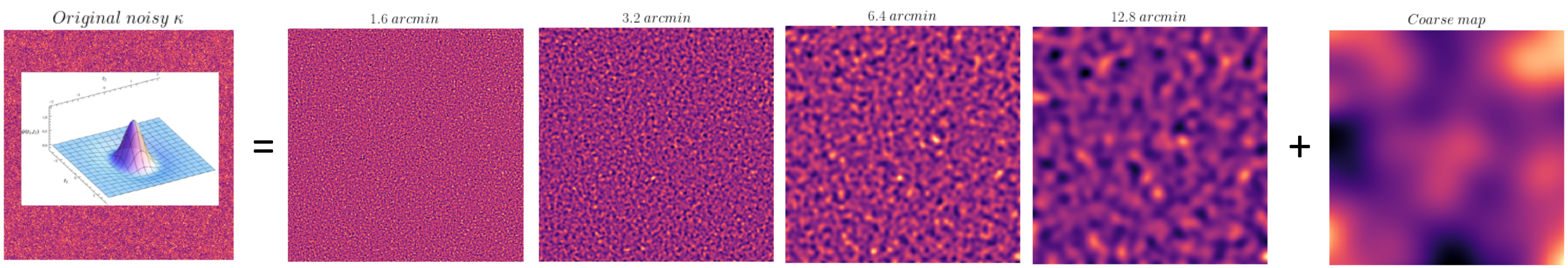
Depending on the computed statistic information gained, information lost

What happens if we consider **all pixels** instead of selecting **multi-scale** minima and maxima?

Starlet ℓ_1-NORM

V. Ajani

V. Ajani, J-L. Starck, V. Pettorino, 2021 A&A Letters, [arXiv:2101.01542](https://arxiv.org/abs/2101.01542)



$$l_1^{j,i} = \sum_{u=1}^{\#coef(S_{j,i})} | S_{j,i}[u] | = \|S_{j,i}\|_1$$

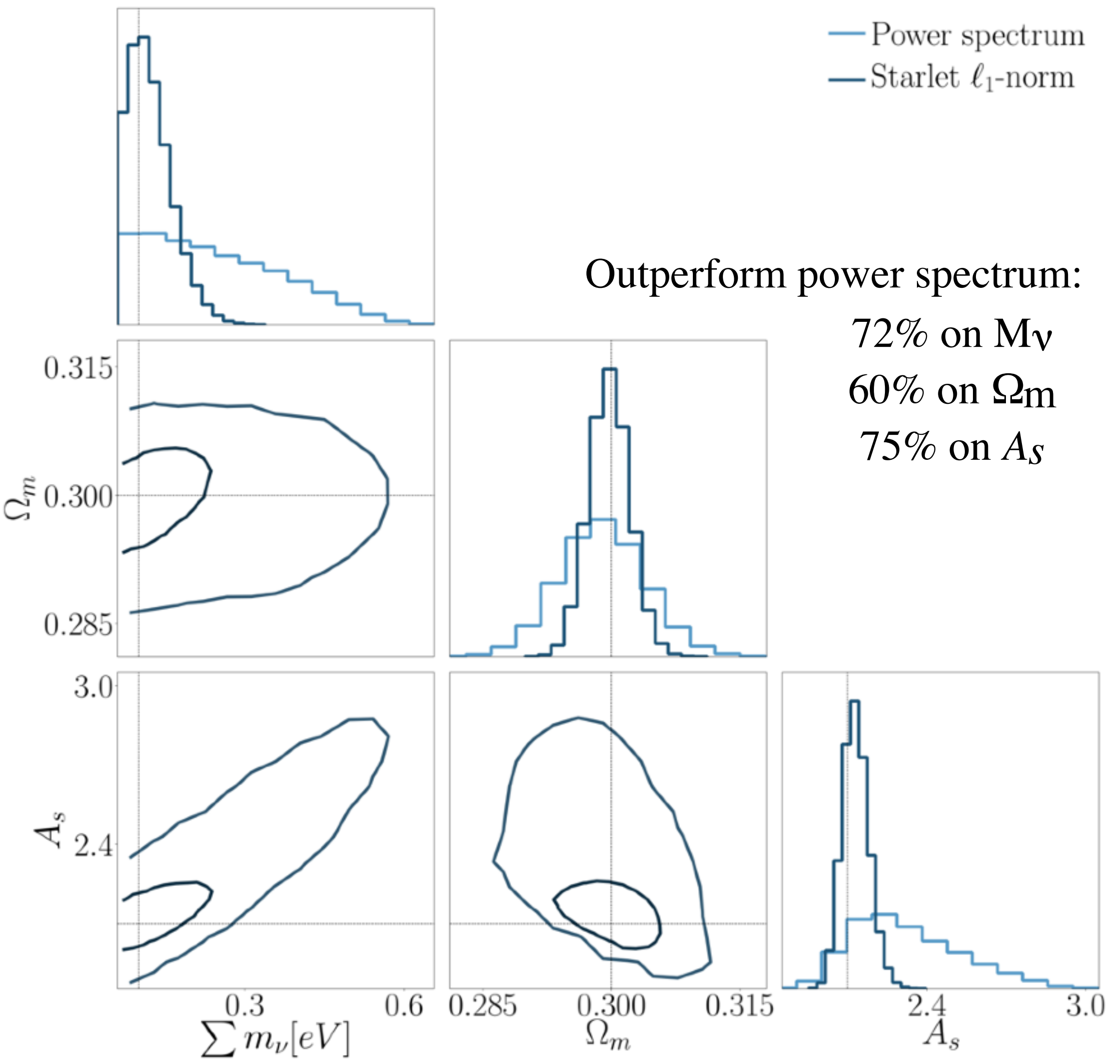
$$S_{j,i} = \{w_{j,k} / B_i < w_{j,k} < B_{i+1}\}$$

wavelet coefficient

- **information** encoded in **all pixels**
- automatically includes **peaks and voids**
- multi-scale approach
- **avoids the problem of defining peaks and voids**

Starlet ℓ_1-NORM

V. Ajani, J.-L. Starck, V. Pettorino, J. Liu, **"Starlet ℓ_1- norm for weak lensing cosmology", A&A letters,** [arXiv:2101.01542](https://arxiv.org/abs/2101.01542)

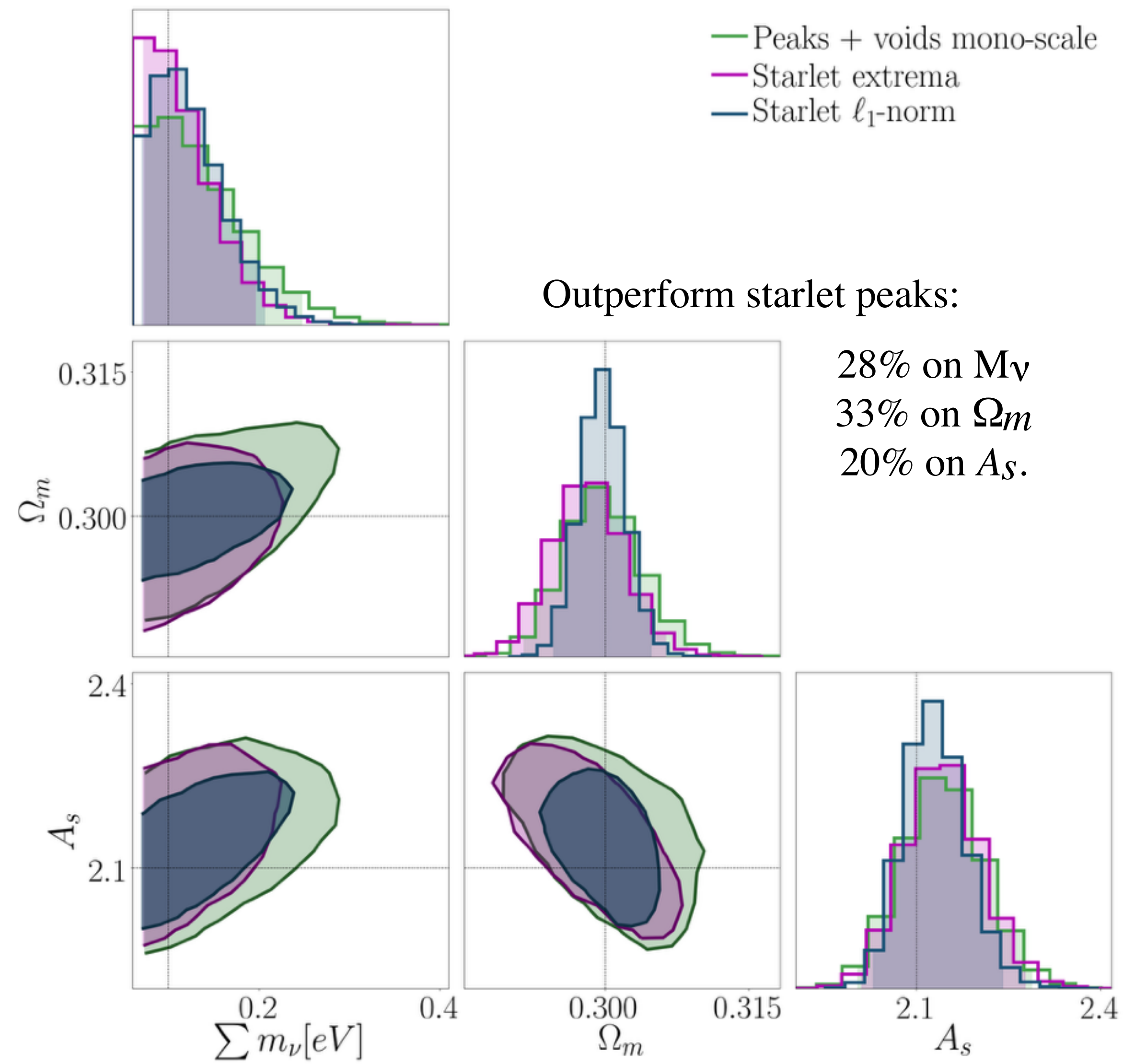


Outperform power spectrum:

72% on M_ν

60% on Ω_m

75% on A_s



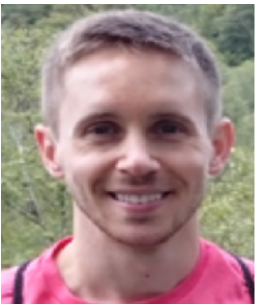
Outperform starlet peaks:

28% on M_ν

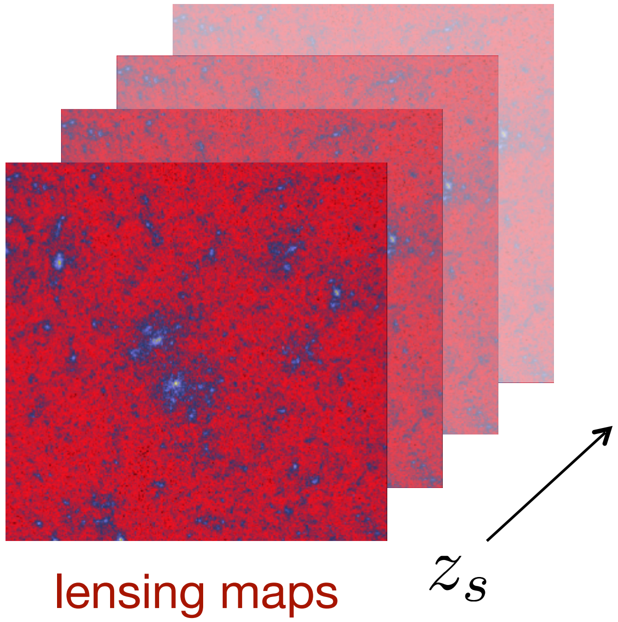
33% on Ω_m

20% on A_s.

==> unified framework to simultaneously account for peaks+voids , and **outperforms power spectrum and state of the art peaks and void statistics**



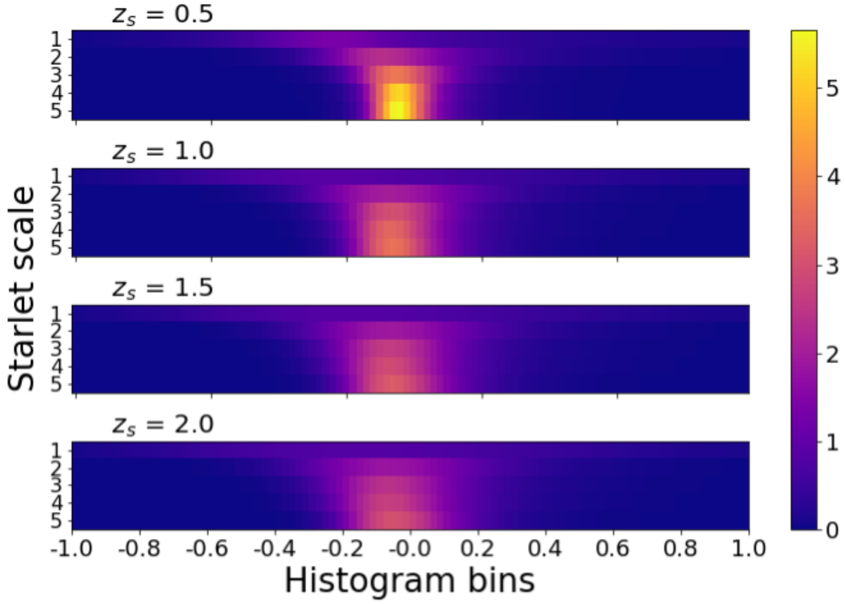
Distinguishing degenerate cosmological models with Machine Learning



lensing maps

dimensionality reduction

with wavelets



Modified grav. + massive neutrinos can **mimic ΛCDM** in terms of the weak-lensing signal.

A **CNN** trained on a wavelet-based representation of the data can discriminate well between models and is **more robust** to noise than conventional statistics like higher-order moments and peak counts.

Prediction (columns) vs. Truth (rows):

$\sigma_{\rm noise} = 0$	ΛCDM	$f_5(R)$ $M_\nu = 0$ eV	$f_5(R)$ $M_\nu = 0.1$ eV	$f_5(R)$ $M_\nu = 0.15$ eV
ΛCDM	1.00	0.00	0.00	0.00
$f_5(R)$ $M_\nu = 0$ eV	0.00	0.88	0.12	0.00
$f_5(R)$ $M_\nu = 0.1$ eV	0.00	0.13	0.83	0.04
$f_5(R)$ $M_\nu = 0.15$ eV	0.00	0.00	0.04	0.96

A. Peel, F. Lalande, J.-L. Starck, V. Pettorino, et al., A&A, 619, id.A38, 2019, arXiv:1810.11030
J. Merten et al, MNRAS, 961 , 2019. arXiv:1810.11027

Forward Modeling

D. Lanzieri

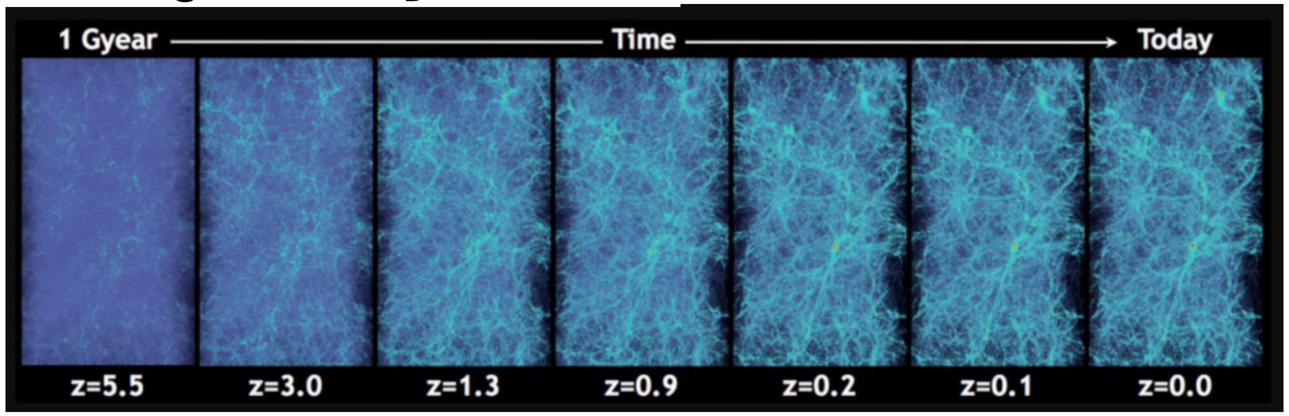
F. Lanusse

Simulating the Universe in a fast and differentiable way using Automatic Differentiation and Gradients in TensorFlow

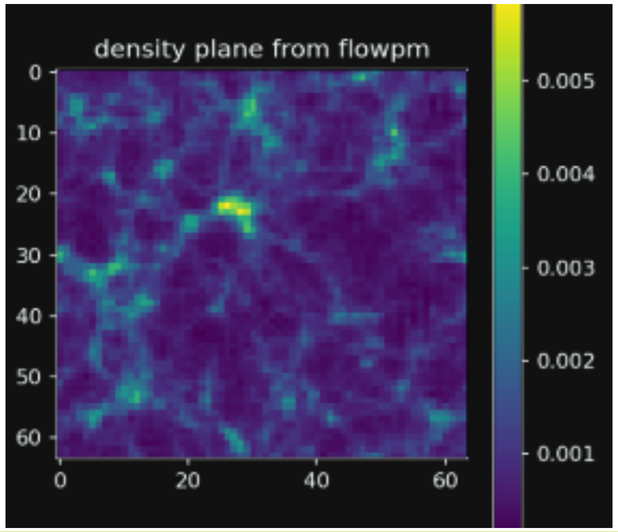
TensorFlow compute automatically derivatives of arbitrary order by applying the chain rule repeatedly to elementary arithmetic operations

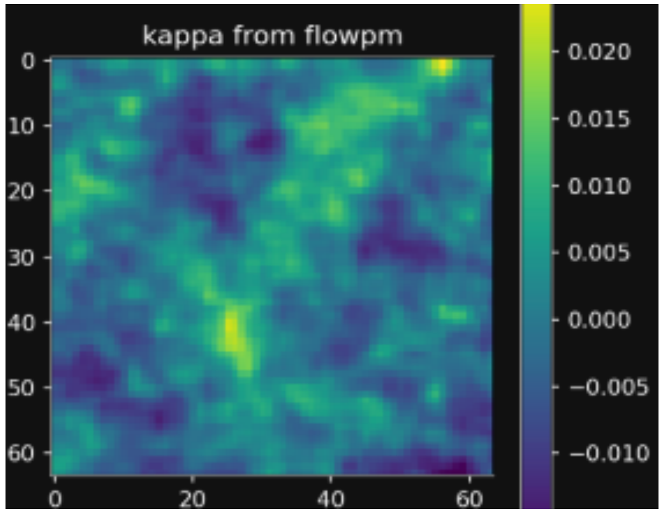
To differentiate automatically, TensorFlow remember what operations happen in what order during the forward pass and traverses this list of operations in reverse order to compute gradients.

Cosmological N-Body Simulations



Lensing lightcones implementing gravitational lensing ray-tracing in FlowPM framework (Born approximation)





Conclusions

✓ Sparsity very efficient for inverse problems

https://github.com/CosmoStat

✓ New Deep Learning Techniques very Promising

➡ Impact of the cosmology used for the training data set ?

➡ Generalisation problem for the PSF recovery.

✓ Wavelets based statistics very efficient.

➡ To be replaced by deep learning as well ?

✓ Forward Modelling very interesting and promising.