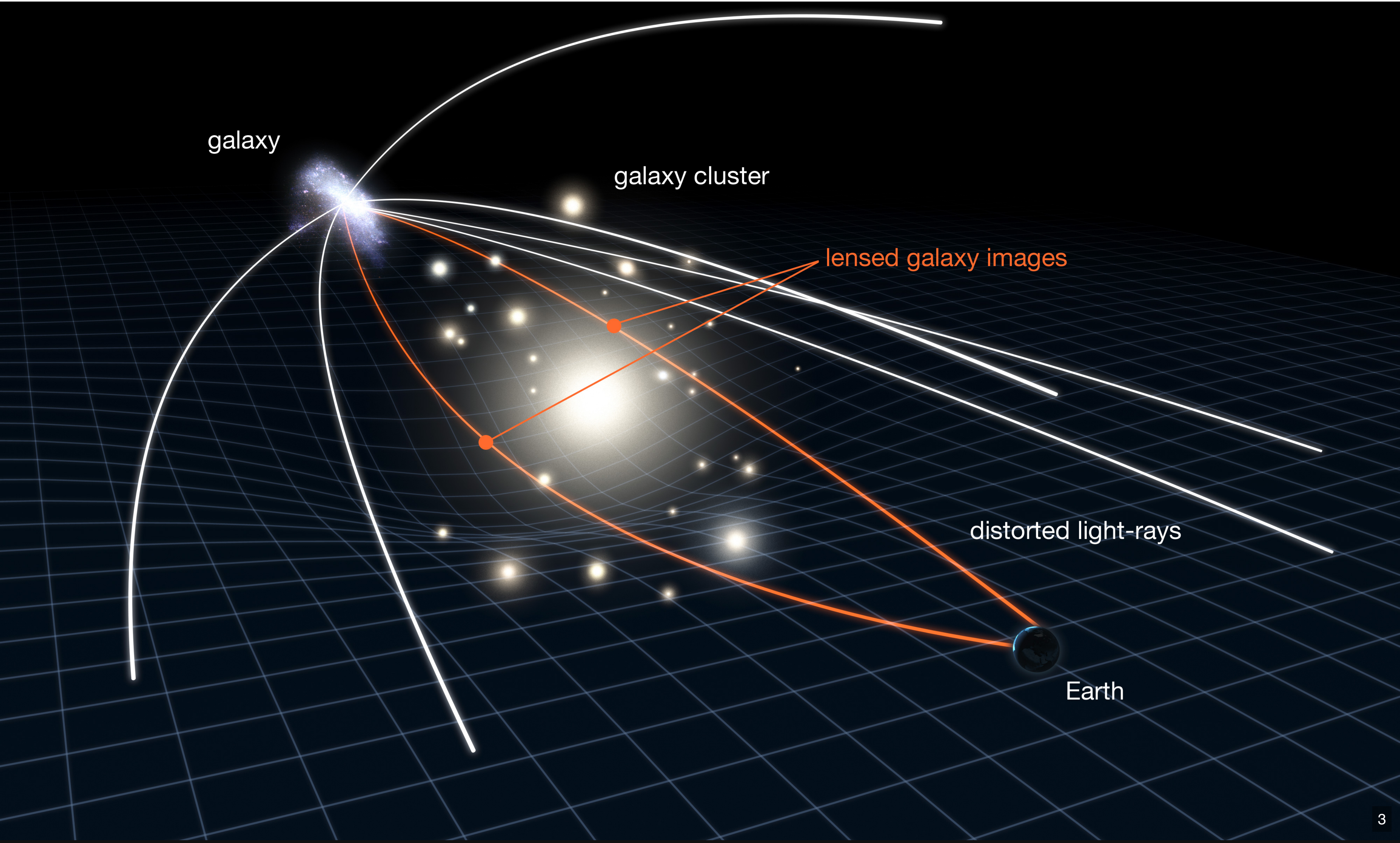
Impact of Weak Lensing Mass Mapping Algorithms on Cosmology Inference

June 11th, 2025

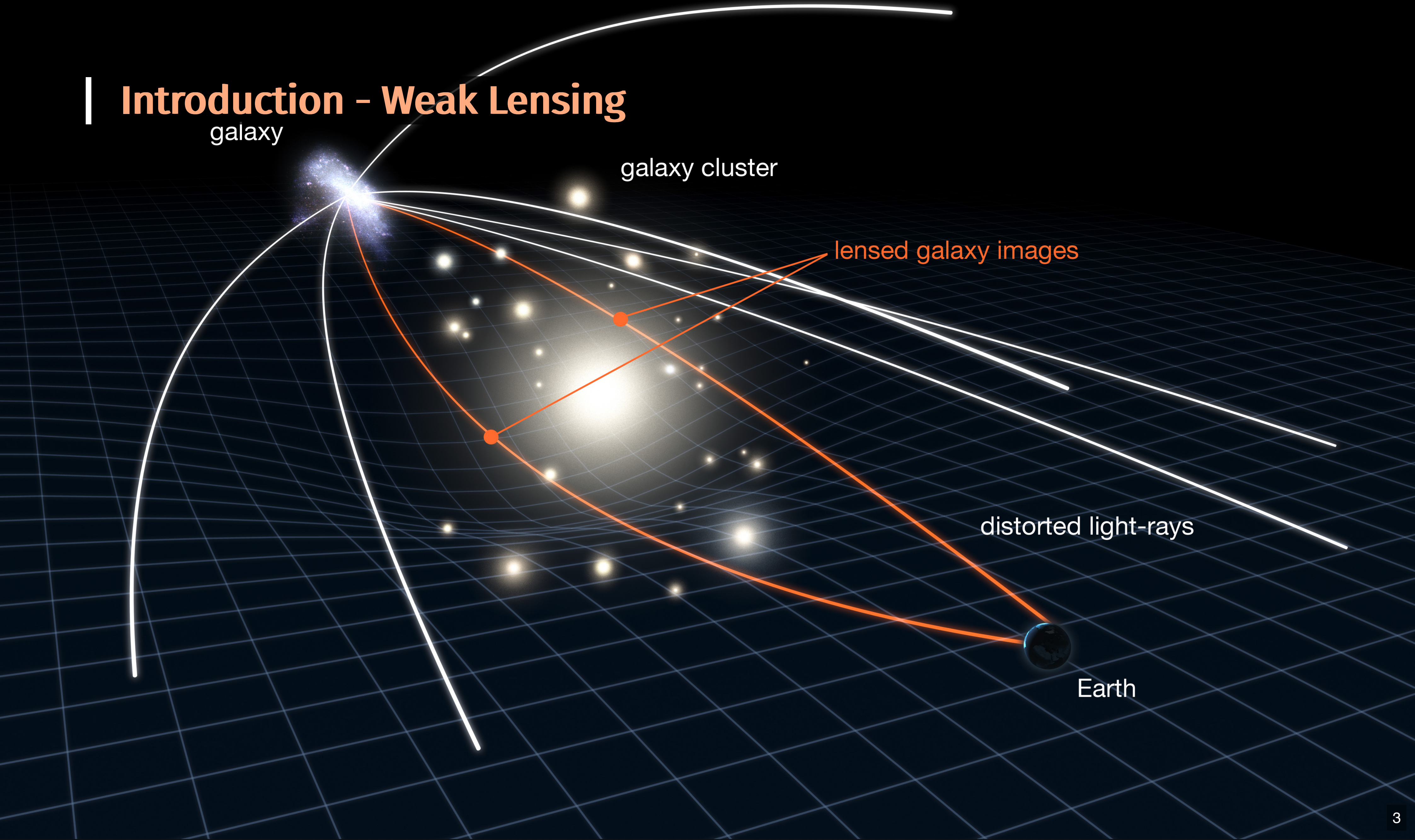
Andreas Tersenov

COLOURS workshop, Saclay

if you are seeing this in pdf, a nicer version of the slides is available at
andreastersenov.github.io/talks/COLOURS_Saclay_2025/

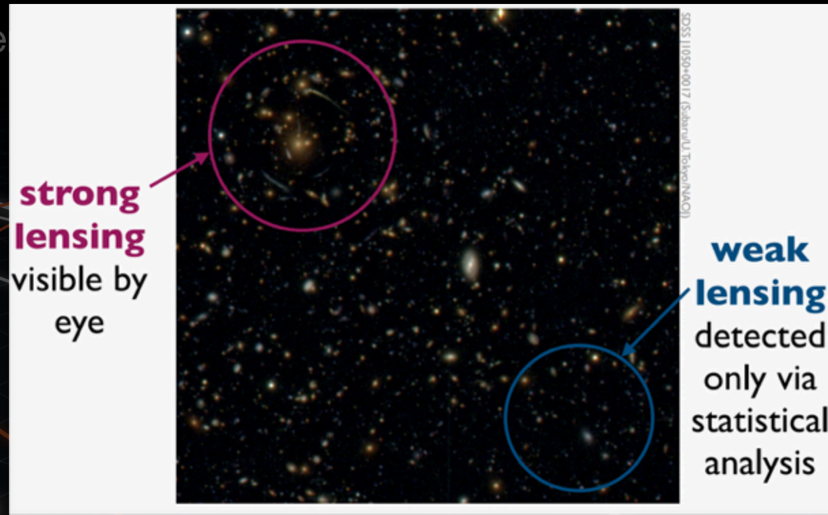

galaxy
galaxy cluster
lensed galaxy images
distorted light-rays
Earth

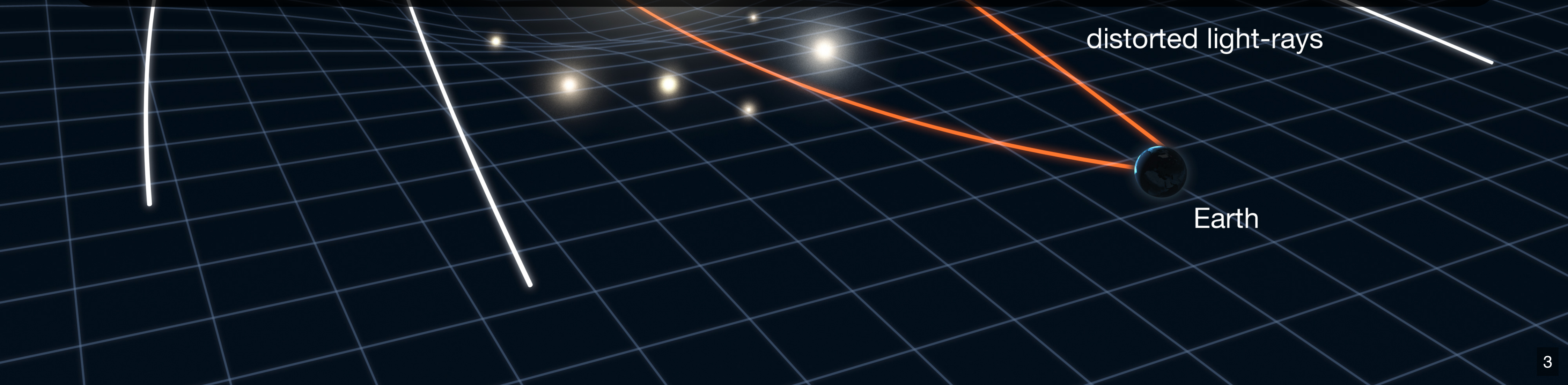
Introduction - Weak Lensing



Introduction – Weak Lensing

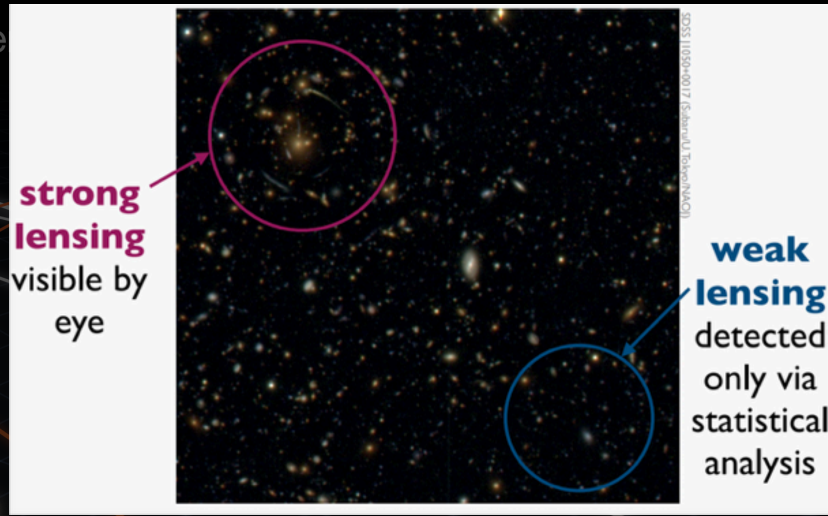
- **WL** = Observational technique in cosmology for studying the matter distribution in the universe
- **Principle:** deflection of light from distant galaxies by gravitational fields → causes image distortion
- **Weak** → subtle & coherent distortions of background galaxy shapes





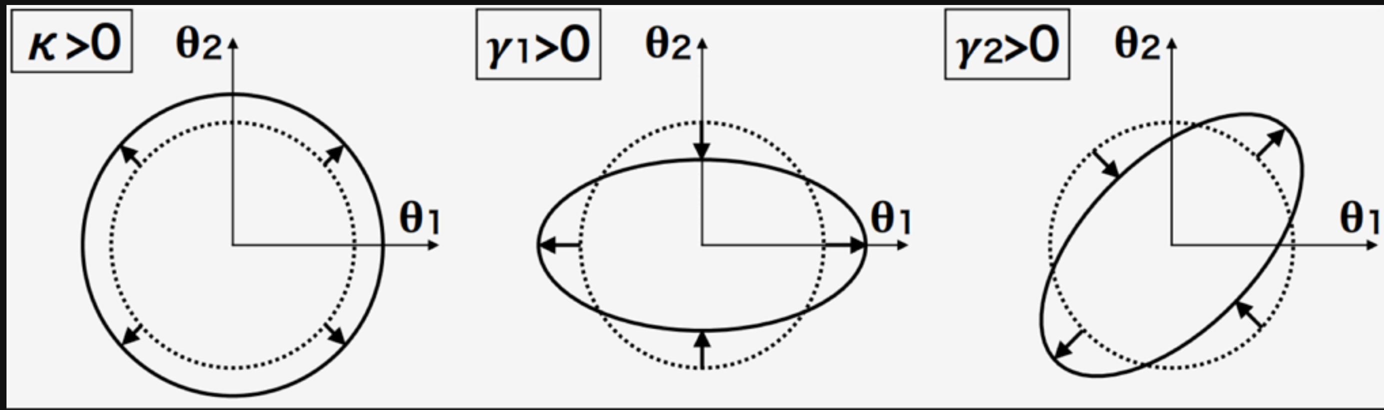
Introduction – Weak Lensing

- **WL** = Observational technique in cosmology for studying the matter distribution in the universe
- **Principle:** deflection of light from distant galaxies by gravitational fields → causes image distortion
- **Weak** → subtle & coherent distortions of background galaxy shapes

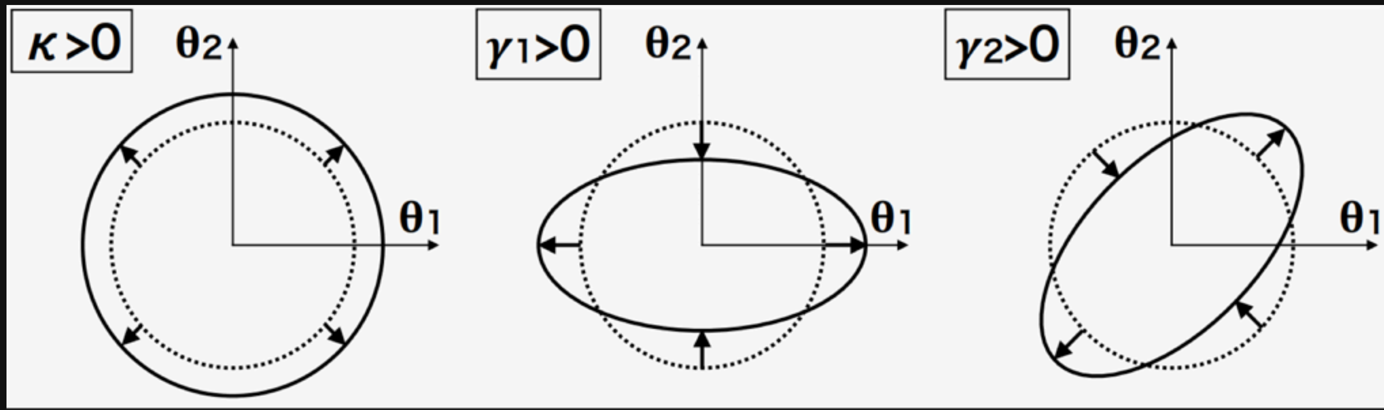


- WL provides a direct measurement of the gravitational distortion.
- WL enables us to probe the cosmic structure, investigate the nature of dark matter, and constrain cosmological parameters.

Shear & Convergence



Shear & Convergence



Convergence κ

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

$\longrightarrow$ difficult to measure

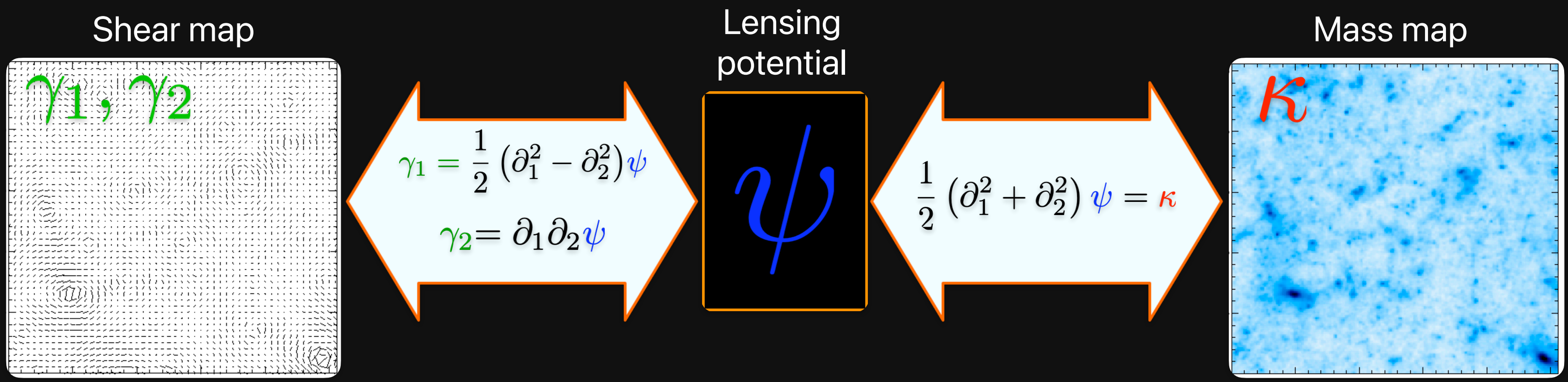
Shear γ

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

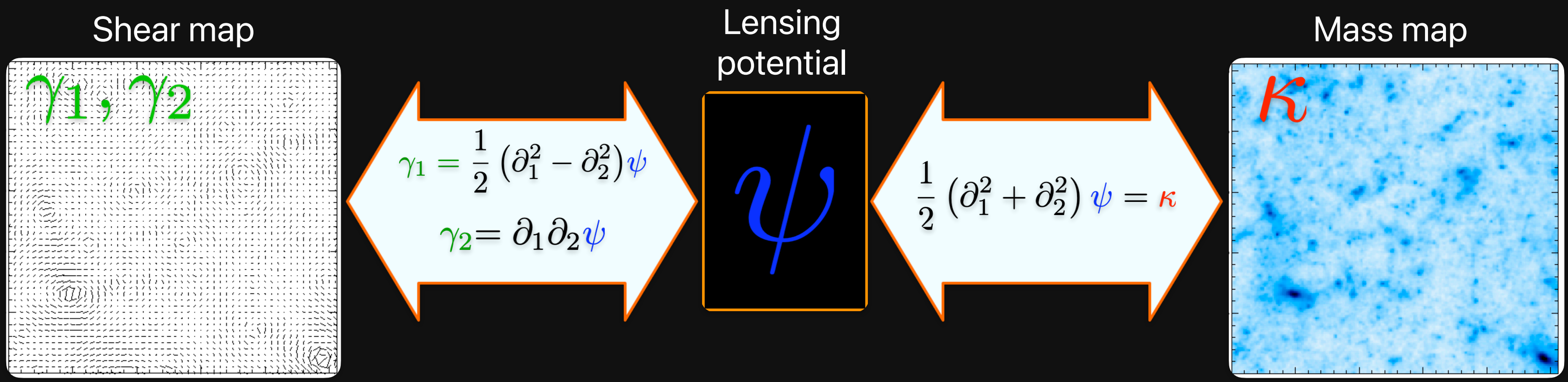
$\longrightarrow$ can be measured by statistical analysis of galaxy shapes

Relation between κ and γ

Relation between κ and γ



Relation between κ and γ



- From **convergence** to **shear**: $\gamma_i = \hat{P}_i\kappa$
- From **shear** to **convergence**: $\kappa = \hat{P}_1\gamma_1 + \hat{P}_2\gamma_2$

$$\hat{P}_1(k) = \frac{k_1^2 - k_2^2}{k^2}, \quad \hat{P}_2(k) = \frac{2k_1k_2}{k^2}$$

Kaiser-Squires inversion

Kaiser-Squires inversion

Advantages:

- Simple ***linear*** operator
- Very easy to implement in Fourier space
- Optimal, in theory

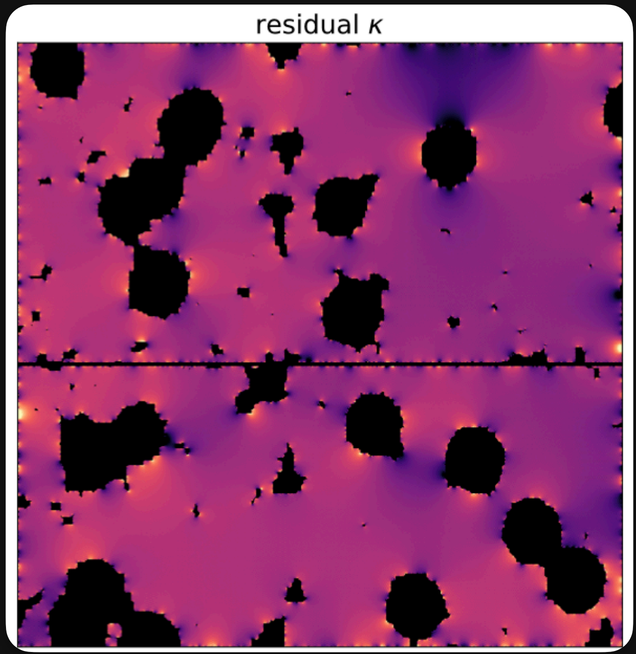
Kaiser-Squires inversion

Advantages:

- Simple *linear* operator
- Very easy to implement in Fourier space
- Optimal, in theory

Practical difficulties:

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



Bayesian reconstruction

Bayesian reconstruction

- Mass mapping problem → statistical inference problem
- **Goal**: infer most probable value of κ-field given observed shear data

$$\underbrace{p(\boldsymbol{\kappa} \mid \boldsymbol{\gamma}, \mathrm{M})}_{\text{Posterior}} \propto \underbrace{p(\gamma \mid \boldsymbol{\kappa}, \mathrm{M})}_{\text{likelihood}} \underbrace{p(\boldsymbol{\kappa} \mid \mathrm{M})}_{\text{prior}}$$

M: cosmological model

Bayesian reconstruction

- Mass mapping problem → statistical inference problem
- **Goal**: infer most probable value of κ-field given observed shear data

$$p(\underbrace{\boldsymbol{\kappa} \mid \gamma}_{\text{Posterior}}, \mathrm{M}) \propto p(\underbrace{\gamma \mid \boldsymbol{\kappa}}_{\text{likelihood}}, \mathrm{M})\, p(\underbrace{\boldsymbol{\kappa} \mid}_{\text{prior}} \mathrm{M})$$

M: cosmological model

- **Maximum A Posteriori solution:** $\hat{x} = \text{argmax}$

Bayesian reconstruction

- Mass mapping problem → statistical inference problem
- **Goal**: infer most probable value of κ-field given observed shear data

$$\underbrace{p(\boldsymbol{\kappa} \mid \boldsymbol{\gamma}, \mathrm{M})}_{\text{Posterior}} \propto \underbrace{p(\gamma \mid \boldsymbol{\kappa}, \mathrm{M})}_{\text{likelihood}} \underbrace{p(\boldsymbol{\kappa} \mid \mathrm{M})}_{\text{prior}}$$

M: cosmological model

- **Maximum A Posteriori solution:** $\hat{x} = \mathrm{argmax}$

- **In practice:** this is usually solved iteratively by alternating two steps:
  - Move toward better data fit (gradient of likelihood)
  - Enforce the prior using a **proximal operator**

Bayesian reconstruction

- Mass mapping problem → statistical inference problem
- **Goal**: infer most probable value of κ-field given observed shear data

$$\underbrace{p(\boldsymbol{\kappa} \mid \gamma, \mathrm{M})}_{\text{Posterior}} \propto \underbrace{p(\gamma \mid \boldsymbol{\kappa}, \mathrm{M})}_{\text{likelihood}} \underbrace{p(\boldsymbol{\kappa} \mid \mathrm{M})}_{\text{prior}}$$

M: cosmological model

- **Maximum A Posteriori solution:** $\hat{x} = \operatorname{argmax}$

- **In practice:** this is usually solved iteratively by alternating two steps:
  - Move toward better data fit (gradient of likelihood)
  - Enforce the prior using a **proximal operator**

Proximal operator

- Acts as a "smart denoiser" by finding the closest solution that satisfies the prior.
- *Example*: sparsity prior → proximal operator performs thresholding to enforce sparsity in the solution.

MCALens

- Models κ-field as a sum of a **Gaussian** and **non-Gaussian** component

$$\kappa = \underbrace{\kappa_{\mathrm{NG}}}_{\text{Standard Wiener filter approach}} + \underbrace{\kappa_{\mathrm{G}}}_{\text{Modified wavelet approach}}$$

$$\min_{\kappa_G, \kappa_{NG}} \|\gamma - \mathbf{A}\left(\kappa_G + \kappa_{NG}\right)\|^2_{\Sigma_n} + C_{\mathrm{G}}\left(\kappa_G\right) + C_{\mathrm{NG}}\left(\kappa_{NG}\right)$$

- **MCA** (morphological Component Analysis) performs an alternating minimization scheme:
  - Estimate κ_{G} assuming κ_{NG} is known:

$$\min_{\kappa_G} \|(\gamma - \mathbf{A}\kappa_{NG}) - \mathbf{A}\kappa_G\|^2_{\Sigma_n} + C_{\mathrm{G}}\left(\kappa_G\right)$$

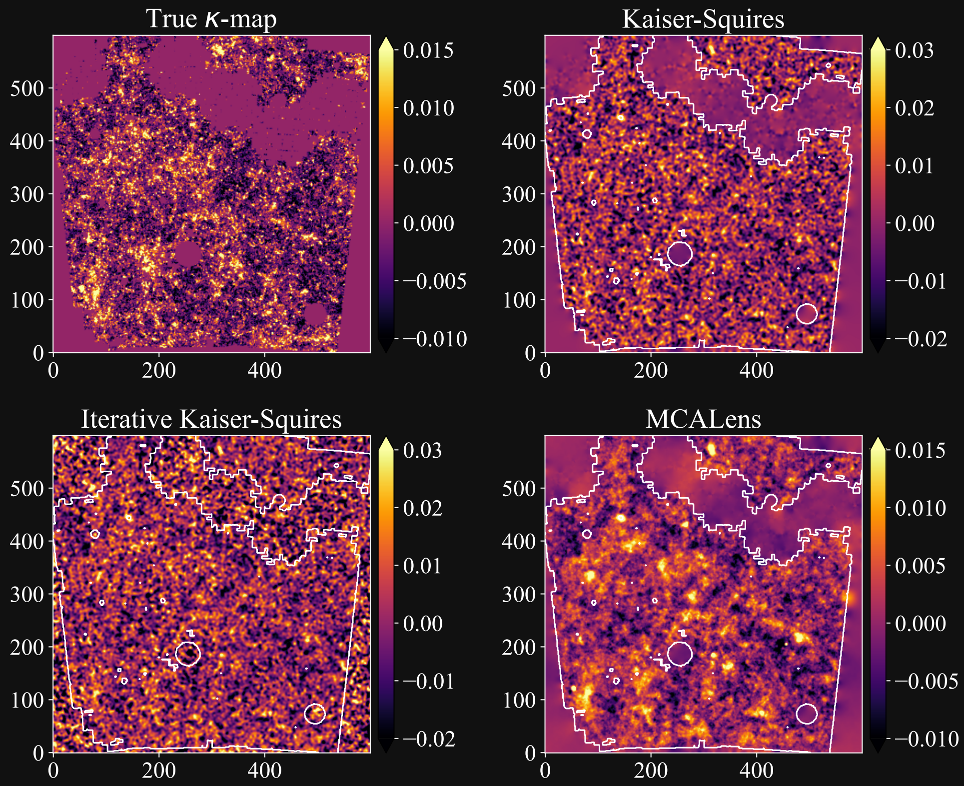
  - Estimate κ_{NG} assuming κ_{G} is known:

$$\min_{\kappa_{NG}} \|(\gamma - \mathbf{A}\kappa_G) - \mathbf{A}\kappa_{NG}\|^2_{\Sigma_n} + C_{\mathrm{NG}}\left(\kappa_{NG}\right)$$

Does the choice of mass-mapping method matter for cosmology?

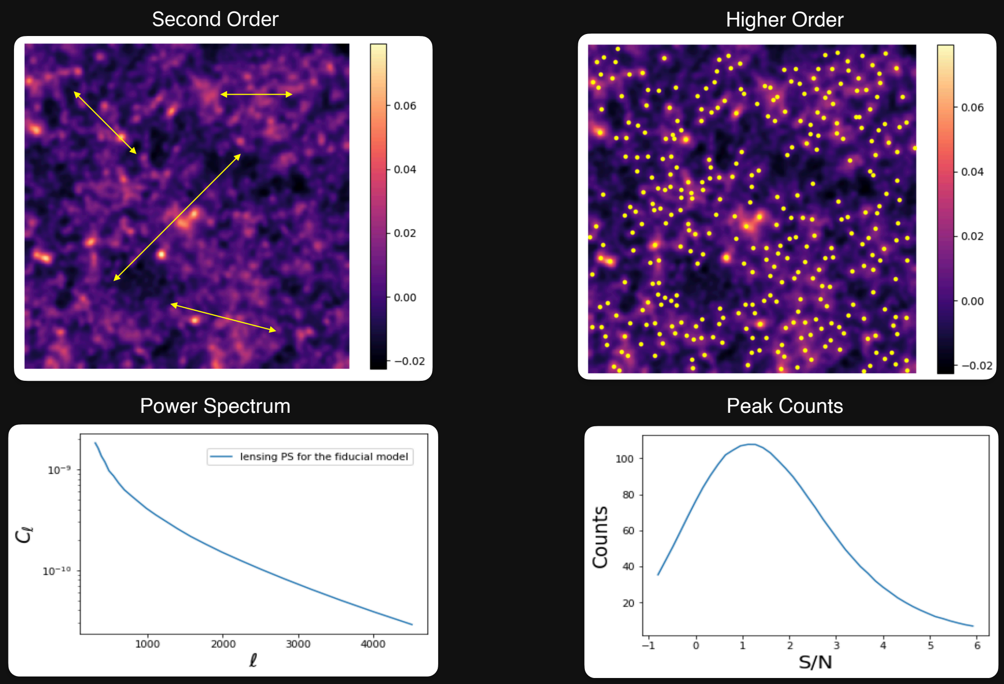
A. Tersenov, L. Baumont, J.L. Starck, M. Kilbinger, doi.org/10.1051/0004-6361/202553707

Mass mapping methods:



Method	RMSE↓
KS	1.1×10^{-2}
iKS	1.1×10^{-2}
MCALens	9.8×10^{-3}

Higher Order Statistics: Peak Counts



- **Peaks**: local maxima of the SNR field $\nu = \dfrac{(\mathcal{W} * \kappa)(\theta_{\text{ker}})}{\sigma_n^{\text{filt}}}$
- Peaks trace regions where the value of κ **is high** → they are associated to **massive structures**

Wavelet peaks

Wavelet peaks

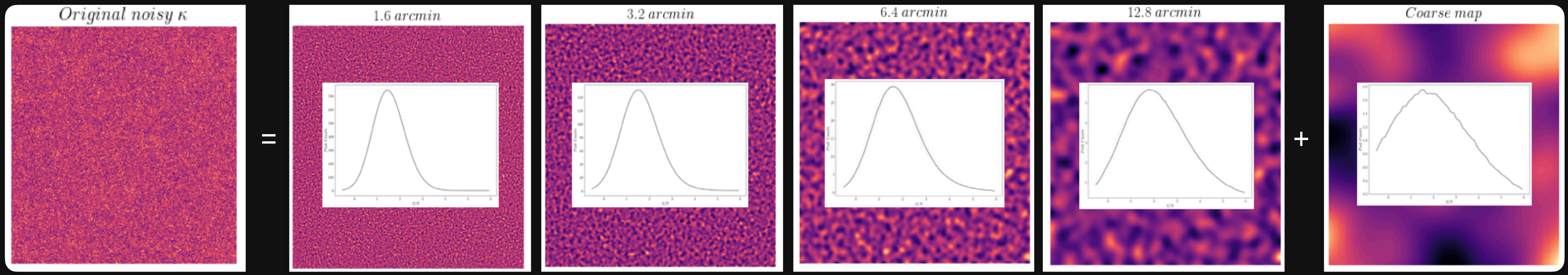
- We consider a **multi-scale analysis** compared to a single-scale analysis

Wavelet peaks

- We consider a **multi-scale analysis** compared to a single-scale analysis
- Apply (instead of Gaussian filter) a **starlet transform** → allows us to represent an image I as a sum of wavelet coefficient images and a coarse resolution image

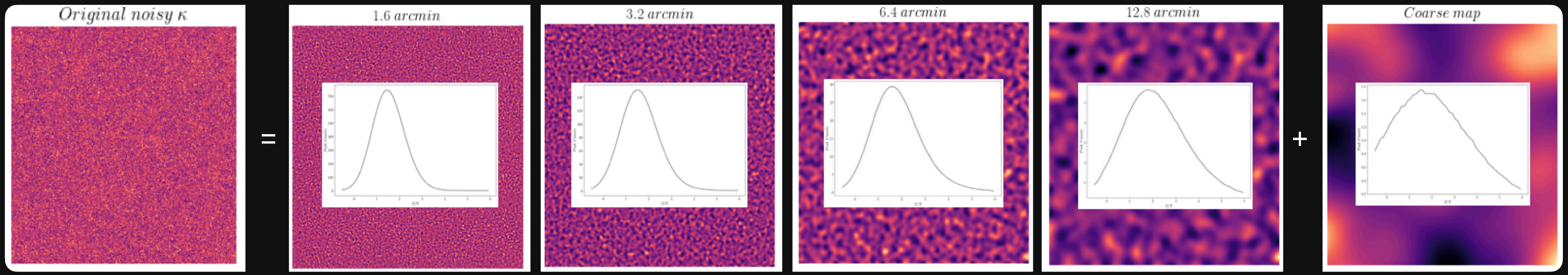
Wavelet peaks

- We consider a **multi-scale analysis** compared to a single-scale analysis
- Apply (instead of Gaussian filter) a **starlet transform** → allows us to represent an image I as a sum of wavelet coefficient images and a coarse resolution image



Wavelet peaks

- We consider a **multi-scale analysis** compared to a single-scale analysis
- Apply (instead of Gaussian filter) a **starlet transform** → allows us to represent an image I as a sum of wavelet coefficient images and a coarse resolution image



- Allows for the **simultaneous** processing of data at different scales → **efficiency**
- Each wavelet band covers a different frequency range, which leads to an almost **diagonal peak count covariance matrix**

Inference with HOS

Inference with HOS

- Use \texttt{cosmoSLICS} simulations: suite designed for the analysis of WL data beyond the standard 2pt cosmic shear

Inference with HOS

- Use \texttt{cosmoSLICS} simulations: suite designed for the analysis of WL data beyond the standard 2pt cosmic shear
- \texttt{cosmoSLICS} cover a wide parameter space in \left[\Omega_m, \sigma_8, w_0, h \right].

Inference with HOS

- Use \texttt{cosmoSLICS} simulations: suite designed for the analysis of WL data beyond the standard 2pt cosmic shear
- \texttt{cosmoSLICS} cover a wide parameter space in \left[\Omega_m, \sigma_8, w_0, h \right].
- For **Bayesian inference** → use a **Gaussian likelihood** for a cosmology independent covariance, and a **flat prior**.

Inference with HOS

- Use \texttt{cosmoSLICS} simulations: suite designed for the analysis of WL data beyond the standard 2pt cosmic shear

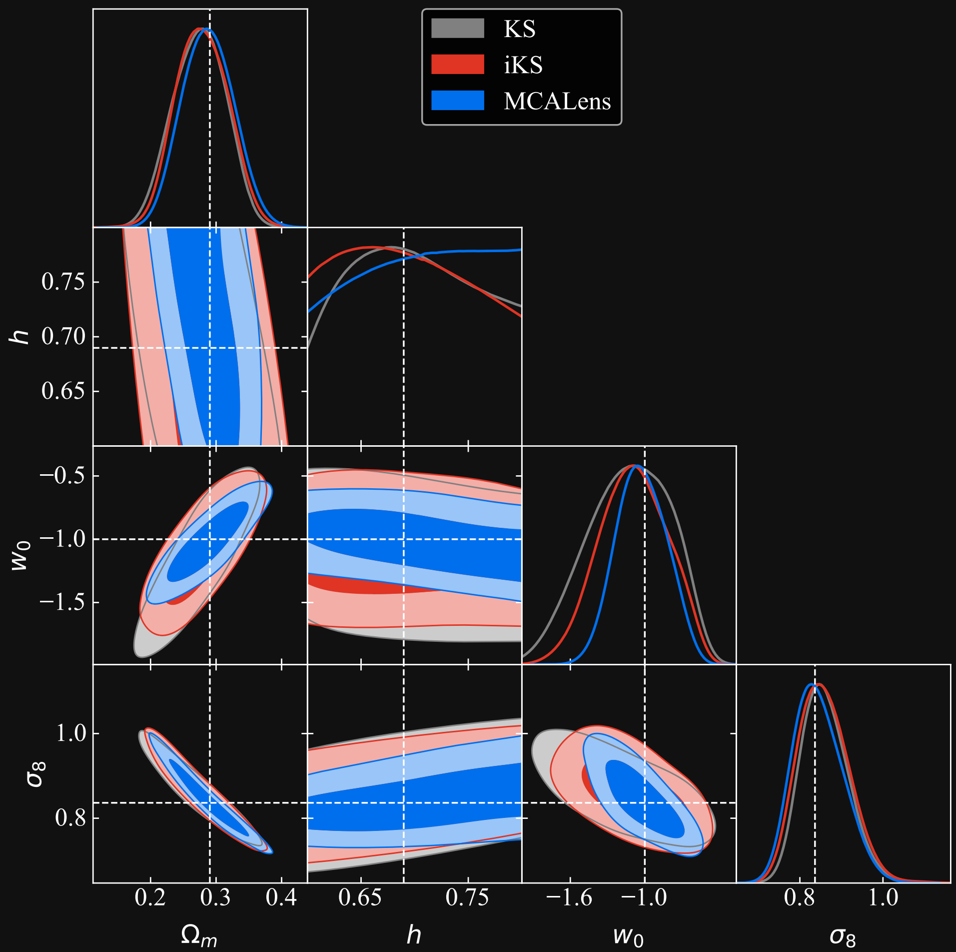
- \texttt{cosmoSLICS} cover a wide parameter space in \left[\Omega_m, \sigma_8, w_0, h \right].

- For **Bayesian inference** → use a **Gaussian likelihood** for a cosmology independent covariance, and a **flat prior**.

- To have a prediction of each HOS **given a new set of parameters**→ employ an interpolation with **Gaussian Process Regressor** (GPR)

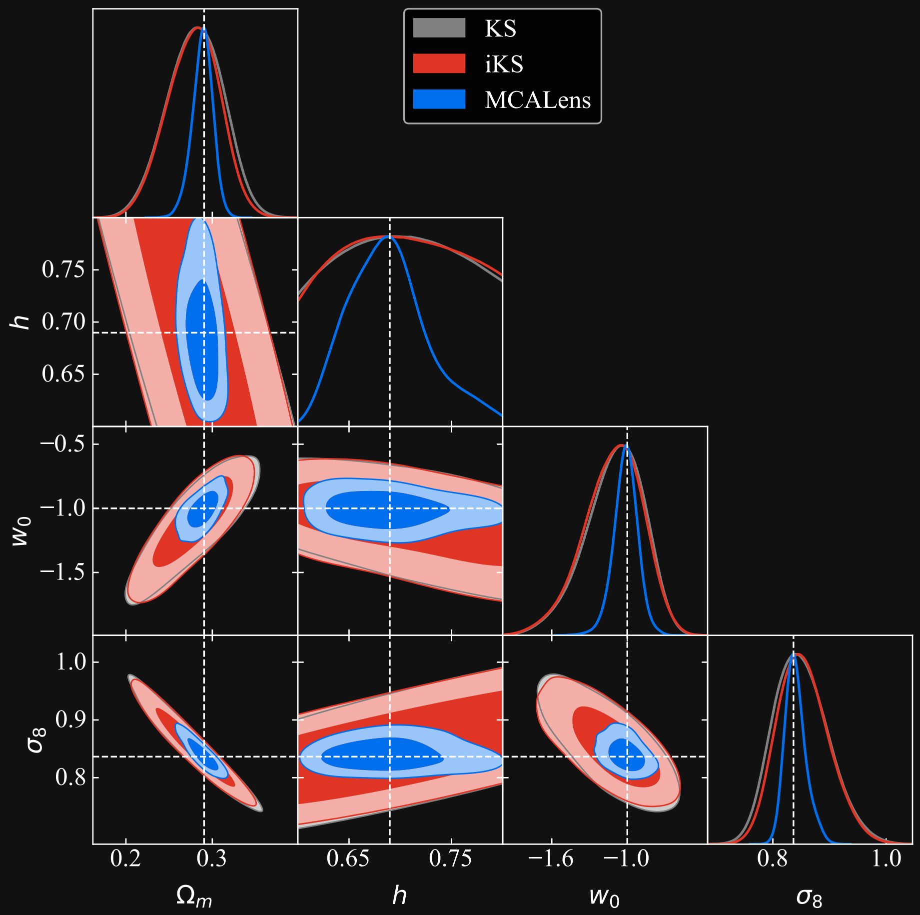
So does the choice of the mass mapping algorithm matter for the final constraints?

The (standard) mono-scale peak counts



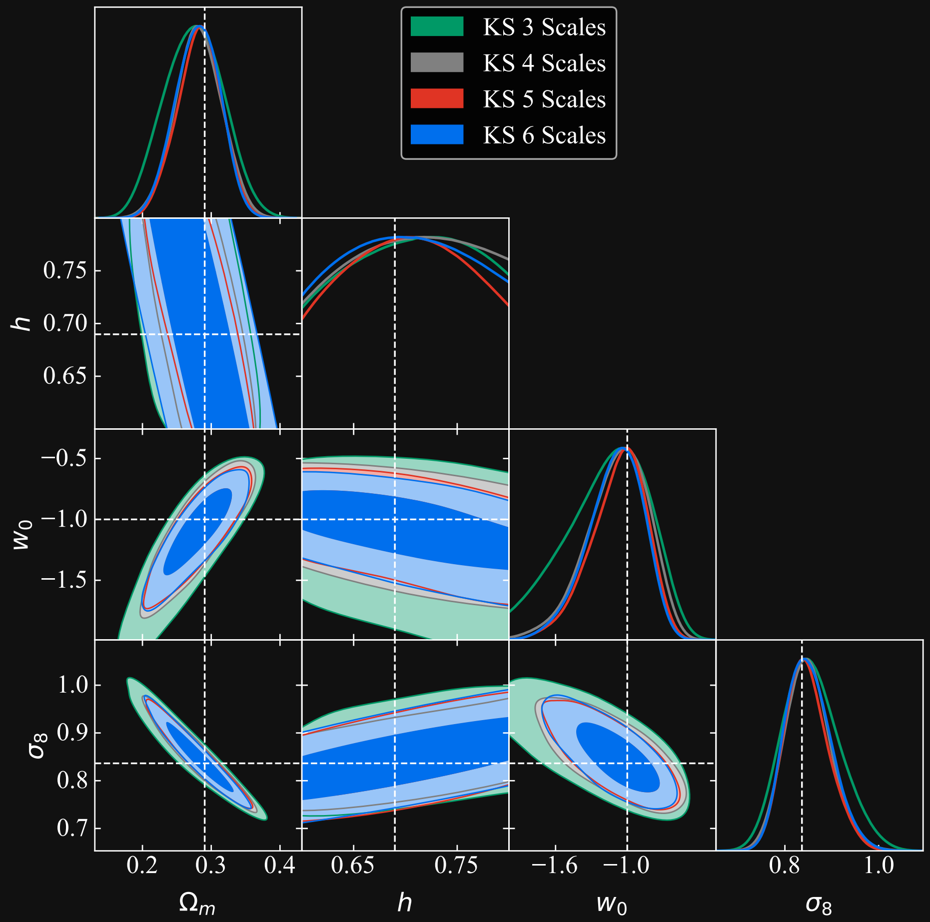
FoM	KS	iKS	MCALens
(Ω_m, h)	476	453	450
(Ω_m, w_0)	152	141	233
(Ω_m, σ_8)	1323	1285	1740
(h, w_0)	55	63	87
(h, σ_8)	336	292	293
(w_0, σ_8)	75	72	124
$(\Omega_m, h, w_0, \sigma_8)$	492	444	578

Wavelet multi-scale peak counts

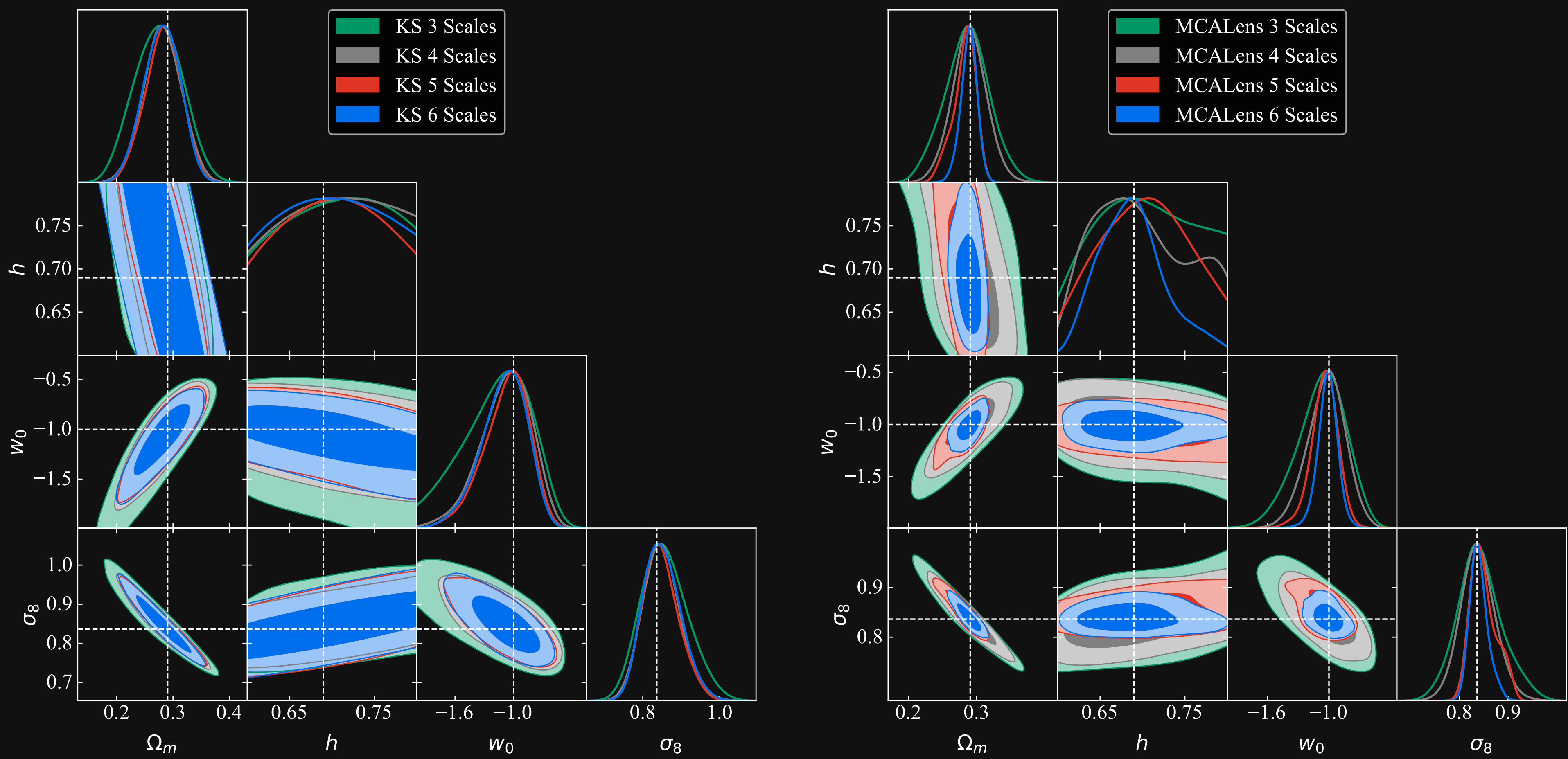


FoM	KS	iKS	MCALens
(Ω_m, h)	670	702	2159
(Ω_m, w_0)	247	244	1051
(Ω_m, σ_8)	2414	2517	9039
(h, w_0)	82	80	259
(h, σ_8)	411	433	1335
(w_0, σ_8)	131	129	577
$(\Omega_m, h, w_0, \sigma_8)$	758	755	1947

Where does this improvement come from?



Where does this improvement come from?



Part 2

A plug-and-play approach with fast uncertainty quantification for weak lensing mass mapping

H. Leterme, A. Tersenov, J. Fadili, and J.-L. Starck (in prep.)

Wait... what about Deep Learning for Mass Mapping?

Wait... what about Deep Learning for Mass Mapping?

Yep, people have tried it! ...And it works!

Wait... what about Deep Learning for Mass Mapping?

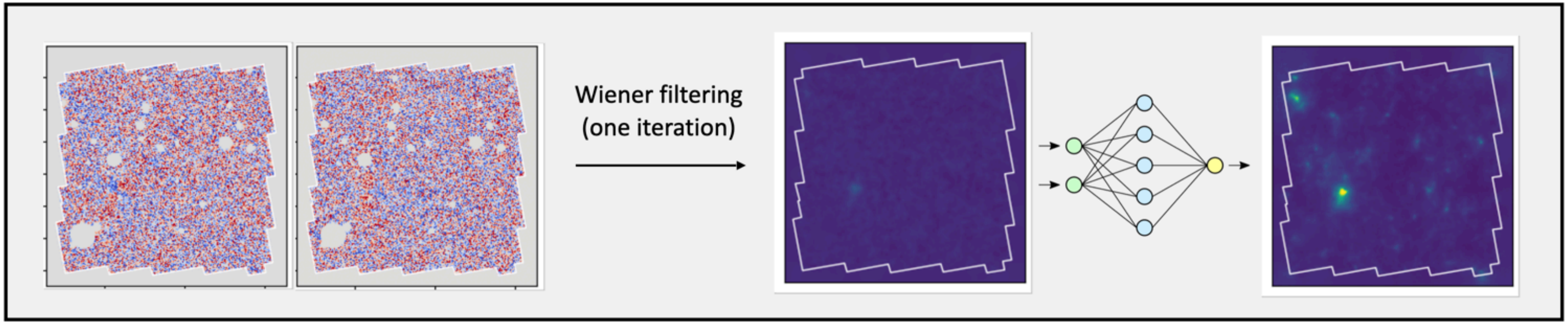
Yep, people have tried it! ...And it works!

Example: DeepMass

Wait... what about Deep Learning for Mass Mapping?

Yep, people have tried it! ...And it works!

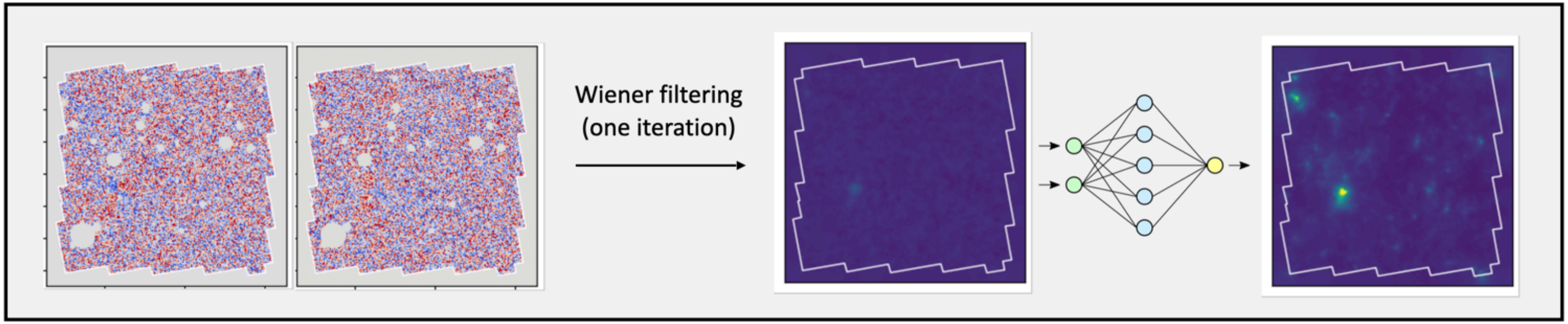
Example: DeepMass



Wait... what about Deep Learning for Mass Mapping?

Yep, people have tried it! ...And it works!

Example: DeepMass



So what's the problem?

Deep Learning for Mass Mapping?

Mass mapping method	Type	Accurate	Flexible	Fast rec.	Fast UQ
Iterative Wiener	Model-driven (Gaus. prior)	✗	✓	✓	✗
MCALens	Model-driven (Gaus. + sparse)	≈	✓	✗	✗
DeepMass	Data-driven (UNet)	✓	✗*	✓	✓
DeepPosterior	Data-driven (UNet + MCMC)	✓	✓	✗	✗
MMGAN	Data-driven (GAN)	✓	✗*	≈	≈

Deep Learning for Mass Mapping?

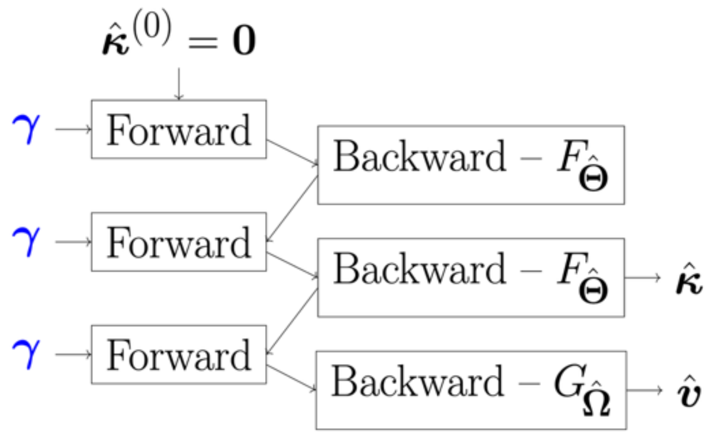
Mass mapping method	Type	Accurate	Flexible	Fast rec.	Fast UQ
Iterative Wiener	Model-driven (Gaus. prior)	✗	✓	✓	✗
MCALens	Model-driven (Gaus. + sparse)	≈	✓	✗	✗
DeepMass	Data-driven (UNet)	✓	✗*	✓	✓
DeepPosterior	Data-driven (UNet + MCMC)	✓	✓	✗	✗
MMGAN	Data-driven (GAN)	✓	✗*	≈	≈
What we'd like	**Data-driven**	✓	✓	✓	✓

Plug-and-Play Mass Mapping

Plug-and-Play Mass Mapping

Main idea:

- *Use PnP framework:* replace **prox** by an on-the-shelf **deep denoiser** trained on simulations $\kappa^{n+1} = \mathrm{prox}_{\tau g} \left(\kappa^n - \tau \nabla f(\kappa^n) \right) \kappa^{n+1} = F_{\Theta} \left(\kappa^n - \tau \mathbf{B} \left(\mathbf{A} \kappa^n - \gamma \right) \right) \kappa^{n+1} = \mathrm{Denoiser} \left(\kappa^n + \mathrm{Data \, residual} \right)$
- Series **converges** towards a **fixed point** $\hat{\kappa}$
- If we choose $\mathbf{B} = \mathbf{A}^T \Sigma^{-1}$ → training phase independent of the noise covariance matrix



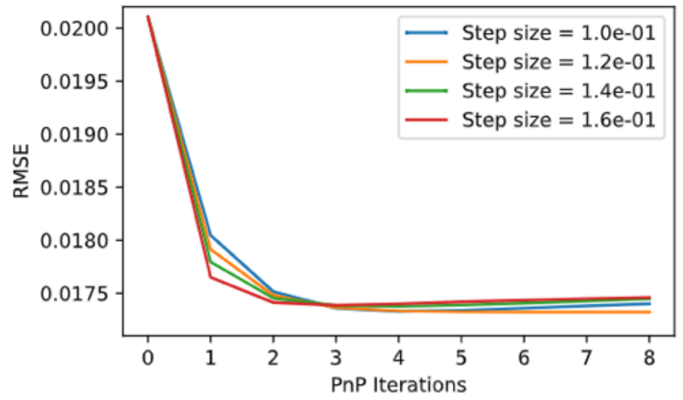
Instead of explicitly writing down a prior, we learn what a "likely" \kappa looks like from simulations and enforce it through denoising.

Implementation

Implementation

Training

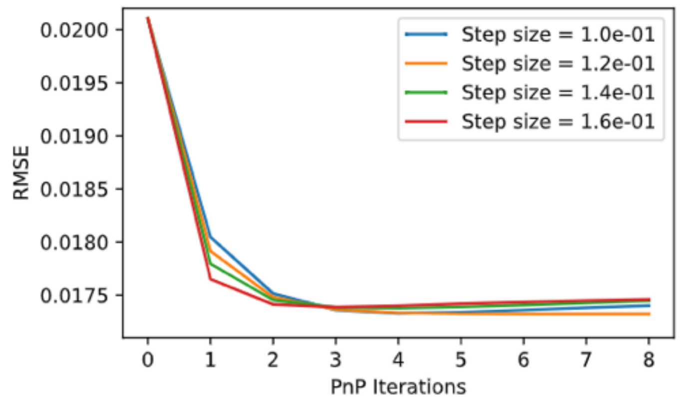
- Denoiser models implemented: **DRUNet** & **SUNet**
- Trained on **κTNG** and **cosmoSLICS** simulations, using pairs of (\kappa_{\rm true}, \gamma_{\rm obs}) as training data

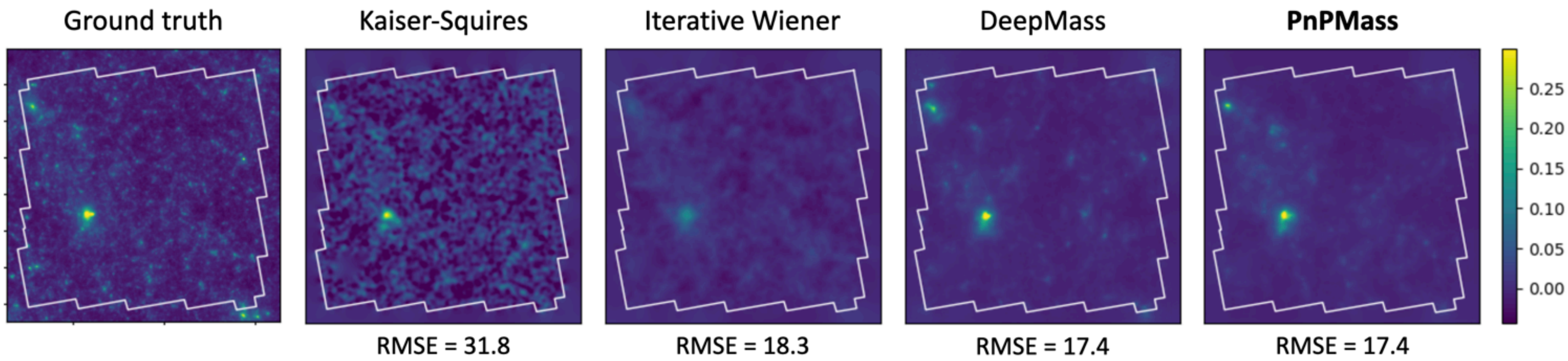


Implementation

Training

- Denoiser models implemented: **DRUNet** & **SUNet**
- Trained on **κTNG** and **cosmoSLICS** simulations, using pairs of (\kappa_{\rm true}, \gamma_{\rm obs}) as training data





How do we estimate uncertainties?

How do we estimate uncertainties?

Step 1

- We train a second neural network to estimate the **posterior variance** of \kappa: \arg\min_{\Omega} \mathbb{E} \left[\left\| G_{\Omega}(\gamma) - \big(\kappa - F_{\Theta}(\gamma) \big)^2 \right\|^2_2 \right]
- Trained on simulated pairs (\kappa, \gamma), just like the denoiser — but now focused on uncertainty.
- This gives fast, pixel-wise error bars
- Uncertainty estimation is fast — just one extra iteration after $\hat{\kappa}$ is computed → adds almost no overhead

How do we estimate uncertainties?

Step 1

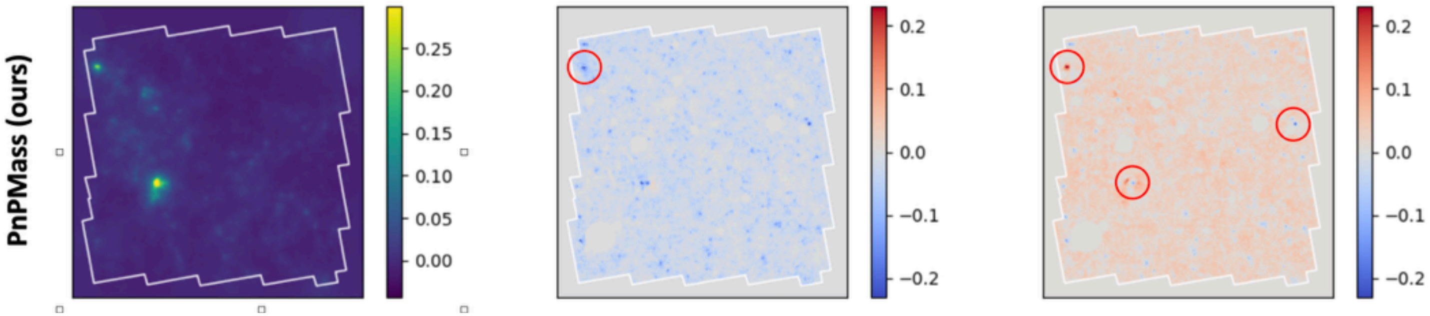
- We train a second neural network to estimate the **posterior variance** of κ: $\arg\min_{\Omega} \mathbb{E} \left[\left\| G_{\Omega}(\gamma) - \big(\kappa - F_{\Theta}(\gamma) \big)^2 \right\|^2_2 \right]$
- Trained on simulated pairs (κ, γ), just like the denoiser — but now focused on uncertainty.
- This gives fast, pixel-wise error bars
- Uncertainty estimation is fast — just one extra iteration after $\hat{\kappa}$ is computed → adds almost no overhead

Step 2

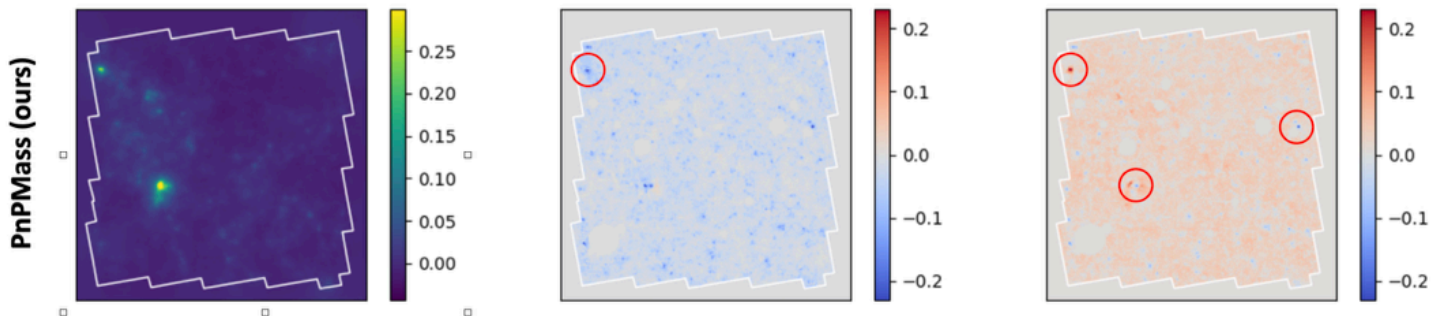
- Neural networks tend to produce miscalibrated uncertainties.
- We apply conformal quantile regression (CQR) to adjust uncertainty intervals so they have guaranteed statistical coverage.
- CQR uses a held-out calibration set to compute a correction factor for each pixel.
- Result: Reliable, data-driven uncertainty maps — with built-in coverage guarantees.

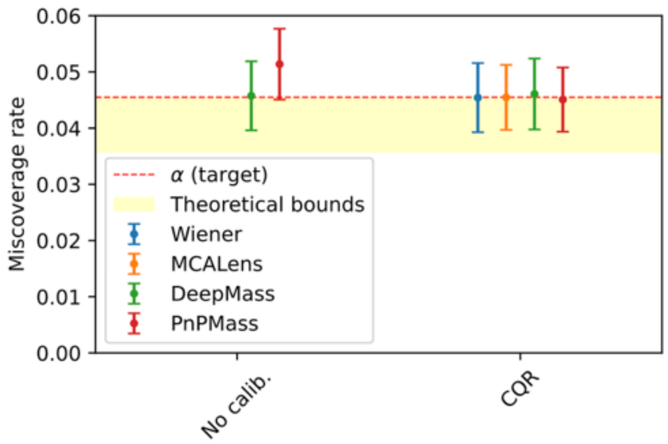
Uncertainty bounds

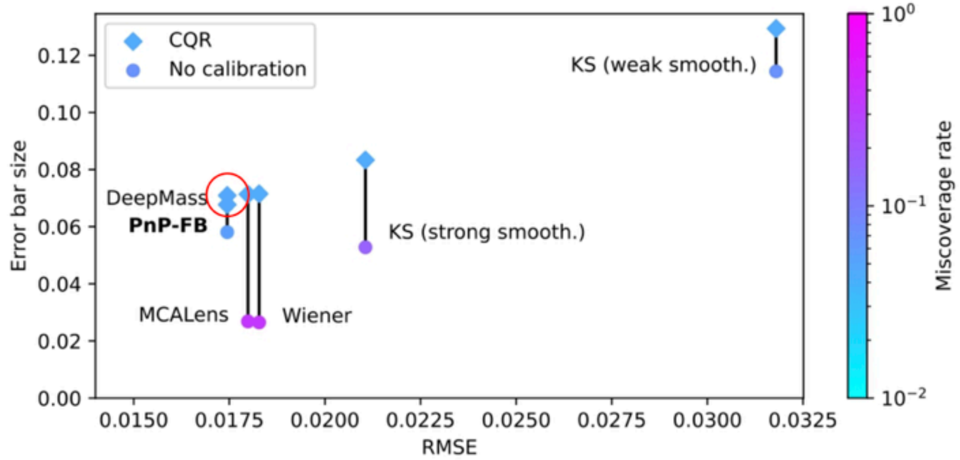
Uncertainty bounds



Uncertainty bounds







Conclusions

Conclusions

- Investigated how different **mass mapping algorithms** affect **cosmological inference** using HOS from WL data.
- Constructed **new mass mapping algorithm** based on the **PnP** formalism.

Conclusions

- Investigated how different **mass mapping algorithms** affect **cosmological inference** using HOS from WL data.
- Constructed **new mass mapping algorithm** based on the **PnP** formalism.

Results

- With a **state-of-the-art** mass-mapping method (MCALens) we managed to get \sim 157\% improvement in FoM over KS.
- Increase in constraining power comes from the more accurate recovery of the **smaller scales**.
- **Wavelet Peak Counts**: Provide tighter constraints than single-scale peak counts.
- **PnP Mass Mapping**:
  - Provides a **fast**, **flexible**, and **accurate** mass mapping algorithm that can be used with any denoiser.
  - Provides **fast** and **reliable** uncertainties using **moment networks** and **conformal quantile regression**.

Conclusions

- Investigated how different **mass mapping algorithms** affect **cosmological inference** using HOS from WL data.
- Constructed **new mass mapping algorithm** based on the **PnP** formalism.

Results

- With a **state-of-the-art** mass-mapping method (MCALens) we managed to get \sim 157\% improvement in FoM over KS.
- Increase in constraining power comes from the more accurate recovery of the **smaller scales**.
- **Wavelet Peak Counts**: Provide tighter constraints than single-scale peak counts.
- **PnP Mass Mapping**:
  - Provides a **fast**, **flexible**, and **accurate** mass mapping algorithm that can be used with any denoiser.
  - Provides **fast** and **reliable** uncertainties using **moment networks** and **conformal quantile regression**.

Takeaway

- **Mass-mapping Matters**: Choosing an advanced mass mapping method significantly enhances constraints on cosmological parameters from HOS.