Weak Lensing Mass Mapping with Uncertainty Quantification

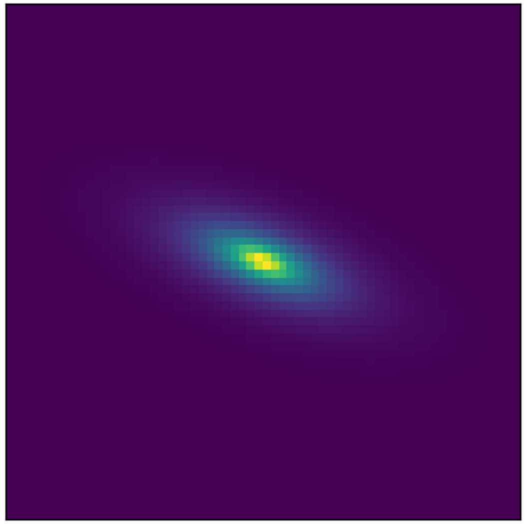
Hubert Leterme, Postdoc
Affiliated to GREYC CNRS-Ensicaen (Caen, France)
Co-supervised at CosmoStat, CEA DAp

Joint ARGOS-TITAN-TOSCA workshop, Heraklion

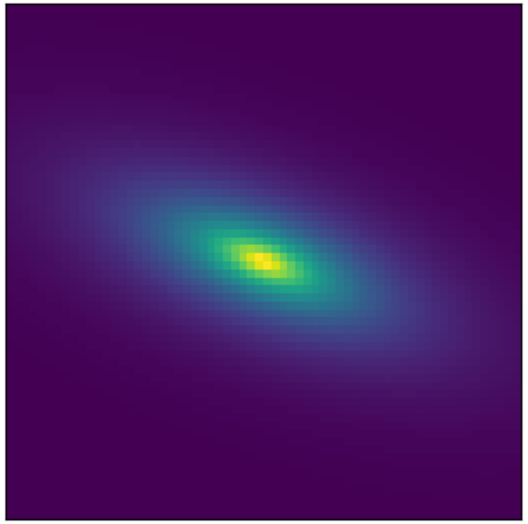
6th June 2024

Context and objectives

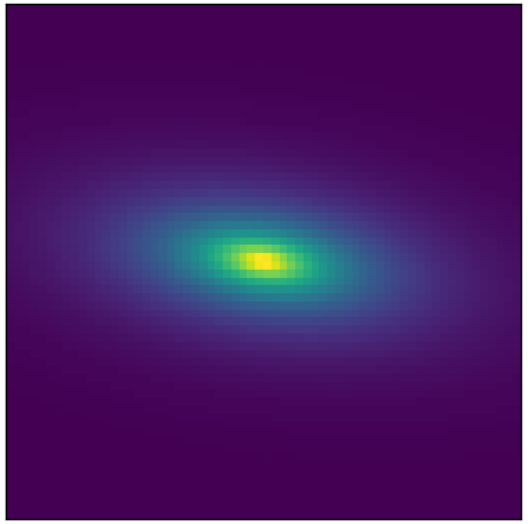
- Convergence map $\boldsymbol{\kappa} \in \mathbb{R}^K$: isotropic dilation of the galaxy image.
  - Proportional to the projected mass along the line of sight.
  - Used to constrain cosmological parameters ⇒ **variable of interest**.
  - However, $\boldsymbol{\kappa}$ cannot be directly measured.
- Shear map $\boldsymbol{\gamma} \in \mathbb{C}^K$: anisotropic stretching of the galaxy image.
- Relationship between shear and convergence maps: $\boldsymbol{\gamma} = \mathbf{A}\boldsymbol{\kappa}$, with $\mathbf{A} \in \mathbb{R}^{K \times K}$ (known).



Source galaxy, unlensed



Convergence only
$\kappa = 1$

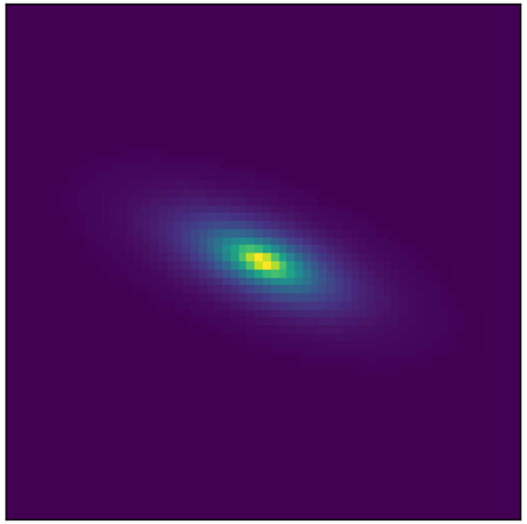


Convergence + shear
$\kappa = 1$ and $\gamma = (0.1 - 0.3\, i)$

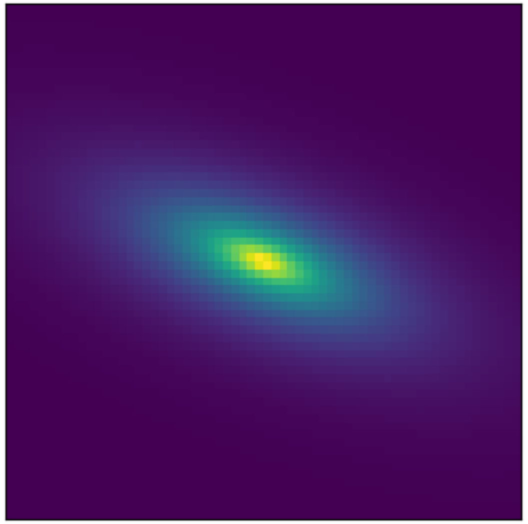
Context and objectives

- Convergence map $\boldsymbol{\kappa} \in \mathbb{R}^K$: isotropic dilation of the galaxy image.
  - Proportional to the projected mass along the line of sight.
  - Used to constrain cosmological parameters $\Rightarrow$ **variable of interest**.
  - However, $\boldsymbol{\kappa}$ cannot be directly measured.
- Shear map $\boldsymbol{\gamma} \in \mathbb{C}^K$: anisotropic stretching of the galaxy image.
- Relationship between shear and convergence maps: $\boldsymbol{\gamma} = \mathbf{A}\boldsymbol{\kappa}$, with $\mathbf{A} \in \mathbb{R}^{K \times K}$ (known).

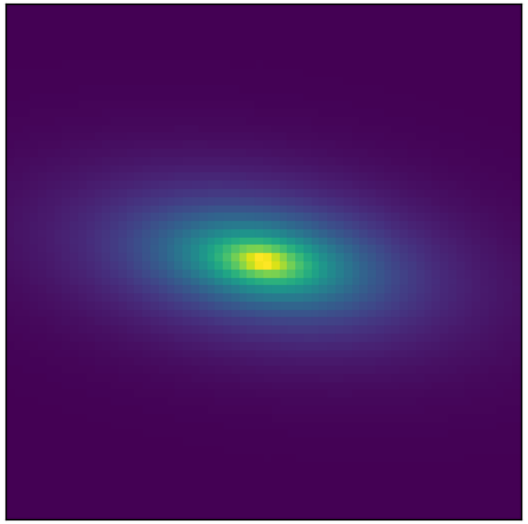
After mean-centering
(mass-sheet degeneracy)



Source galaxy, unlensed



Convergence only
$\kappa = 1$

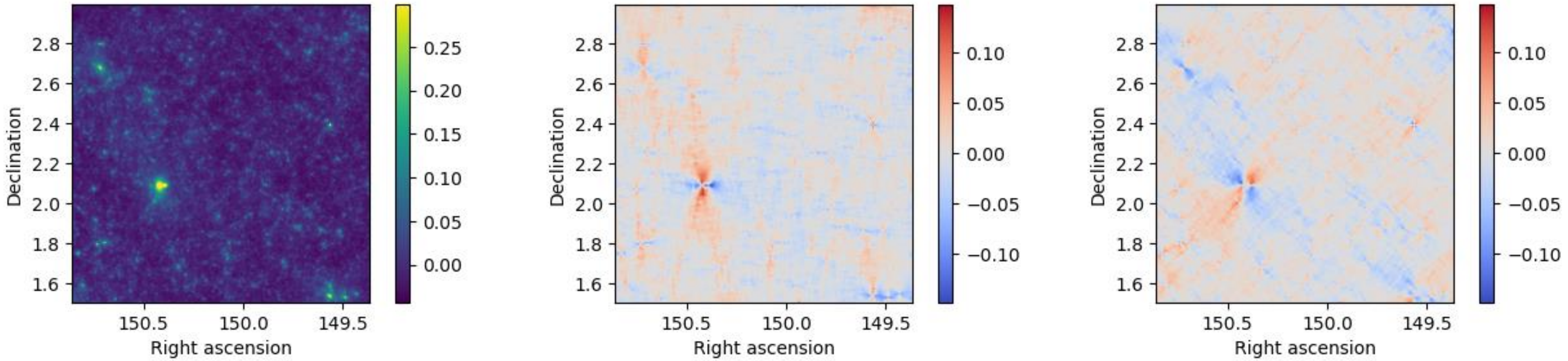


Convergence + shear
$\kappa = 1$ and $\gamma = (0.1 - 0.3\, i)$

Context and objectives

Example with the κTNG simulated dataset[1]

Simulated convergence map — Corresponding shear map (real and imaginary parts)



- As for the convergence map $\boldsymbol{\kappa}$, the true shear map $\boldsymbol{\gamma}$ cannot be directly measured.
- Unbiased estimator of $\boldsymbol{\gamma}$, obtained by measuring galaxy ellipticities: $\boldsymbol{\gamma} \longleftarrow \boldsymbol{\epsilon} - \langle \boldsymbol{\epsilon} \rangle$
- Relation between $\boldsymbol{\gamma}$ (observable) and $\boldsymbol{\kappa}$ (quantity of interest):

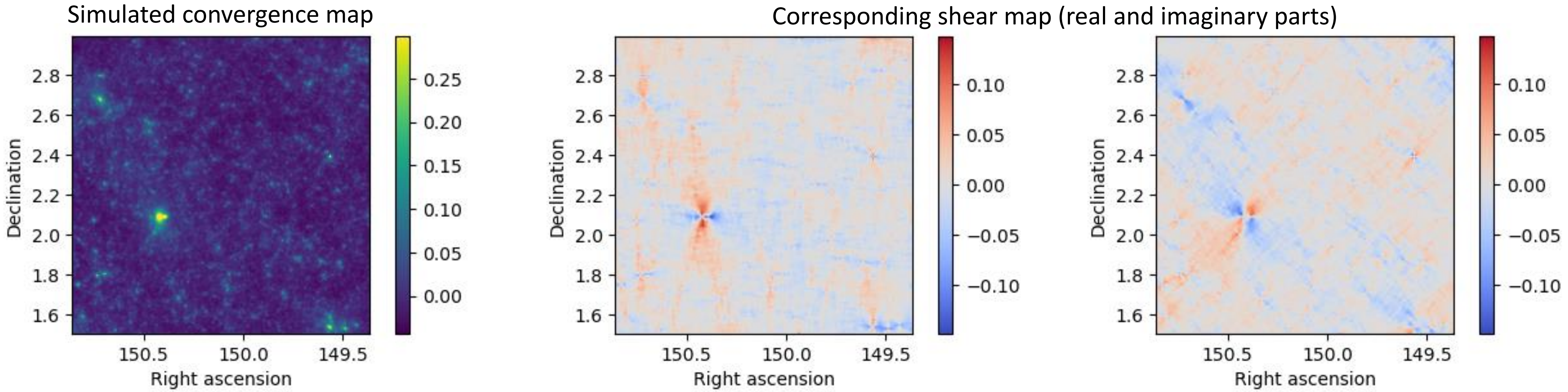
$$\boldsymbol{\gamma} = \mathbf{A}\boldsymbol{\kappa} + \boldsymbol{n},$$

  with noise $\boldsymbol{n}$ assumed Gaussian, zero-centered and with diagonal covariance matrix $\boldsymbol{\Sigma}$.
- Noise level (standard deviation per pixel): $\boldsymbol{\Sigma}[k,k] = \sigma / N_k$.

[1] K. Osato, J. Liu, and Z. Haiman, "κTNG: effect of baryonic processes on weak lensing with IllustrisTNG simulations," Monthly Notices of the Royal Astronomical Society, vol. 502, no. 4, pp. 5593–5602, Apr. 2021

Context and objectives

Example with the κTNG simulated dataset[1]



- As for the convergence map $\boldsymbol{\kappa}$, the true shear map $\boldsymbol{\gamma}$ cannot be directly measured.
- Unbiased estimator of $\boldsymbol{\gamma}$, obtained by measuring galaxy ellipticities: $\boldsymbol{\gamma} \longleftarrow \boldsymbol{\epsilon} - \langle \boldsymbol{\epsilon} \rangle$
- Relation between $\boldsymbol{\gamma}$ (observable) and $\boldsymbol{\kappa}$ (quantity of interest):

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

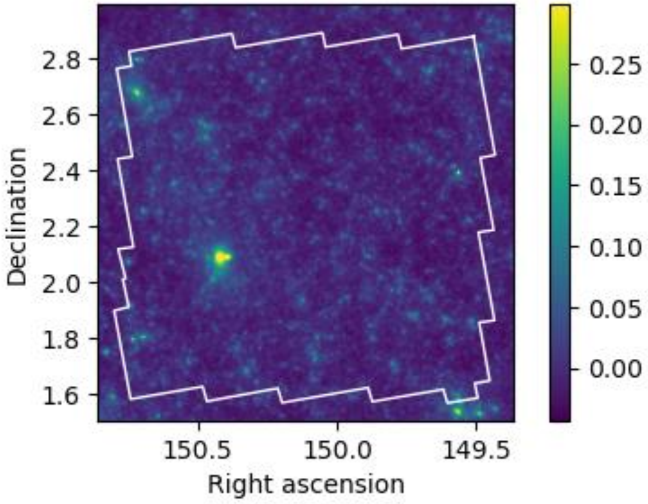
  with noise $\boldsymbol{n}$ assumed Gaussian, zero-centered and with diagonal covariance matrix $\boldsymbol{\Sigma}$.
- Noise level (standard deviation per pixel): $\boldsymbol{\Sigma}[k, k] = \sigma / N_k$. ← Nb measured galaxies; ← Intrinsic ellipticity (std)

[1] K. Osato, J. Liu, and Z. Haiman, "κTNG: effect of baryonic processes on weak lensing with IllustrisTNG simulations," Monthly Notices of the Royal Astronomical Society, vol. 502, no. 4, pp. 5593–5602, Apr. 2021

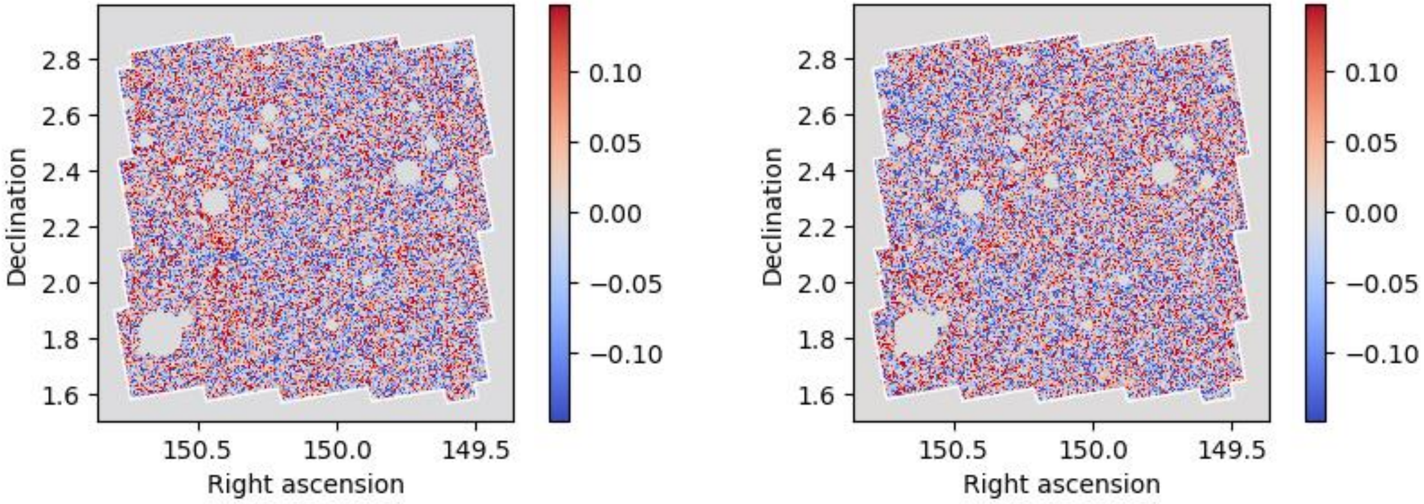
Context and objectives

Noisy shear maps (noise variance taken from the COSMOS shape catalog[1])

Simulated convergence map



Corresponding shear map with noise (real and imaginary parts)



Objective: given $\boldsymbol{\gamma}$, estimate $\widehat{\boldsymbol{\kappa}}^-$ and $\widehat{\boldsymbol{\kappa}}^+$ such that

$$\mathbb{E}[L(\boldsymbol{\kappa}, \widehat{\boldsymbol{\kappa}}^-, \widehat{\boldsymbol{\kappa}}^+)] \leq \alpha.$$

- Over which uncertainties the expected value is calculated?

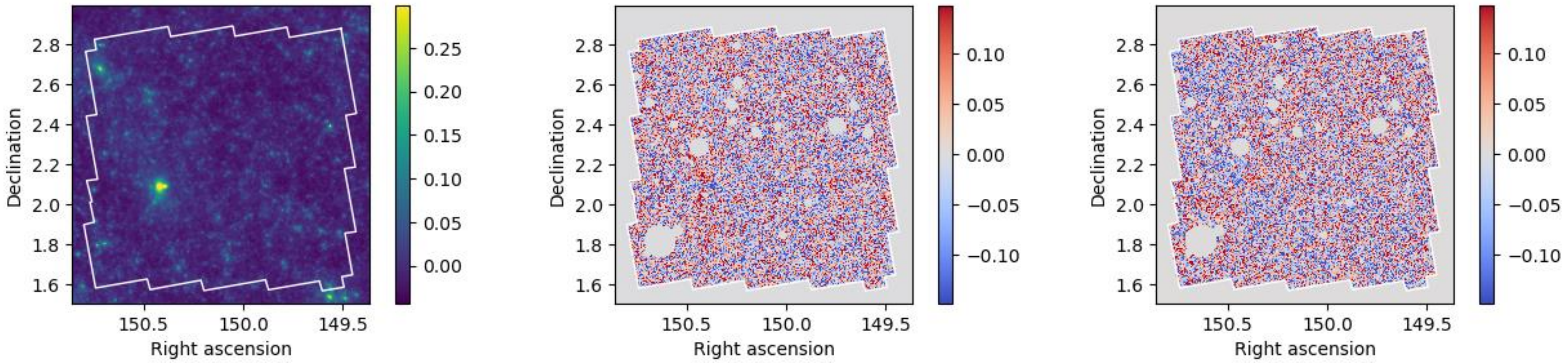
[1] https://astro.uni-bonn.de/en/m/schrabba/research

Context and objectives

Noisy shear maps (noise variance taken from the COSMOS shape catalog[1])

Simulated convergence map

Corresponding shear map with noise (real and imaginary parts)



Objective: given $\boldsymbol{\gamma}$, estimate $\widehat{\boldsymbol{\kappa}}^{-}$ and $\widehat{\boldsymbol{\kappa}}^{+}$ such that

Expected miscoverage rate
(% of pixels outside the bounds)

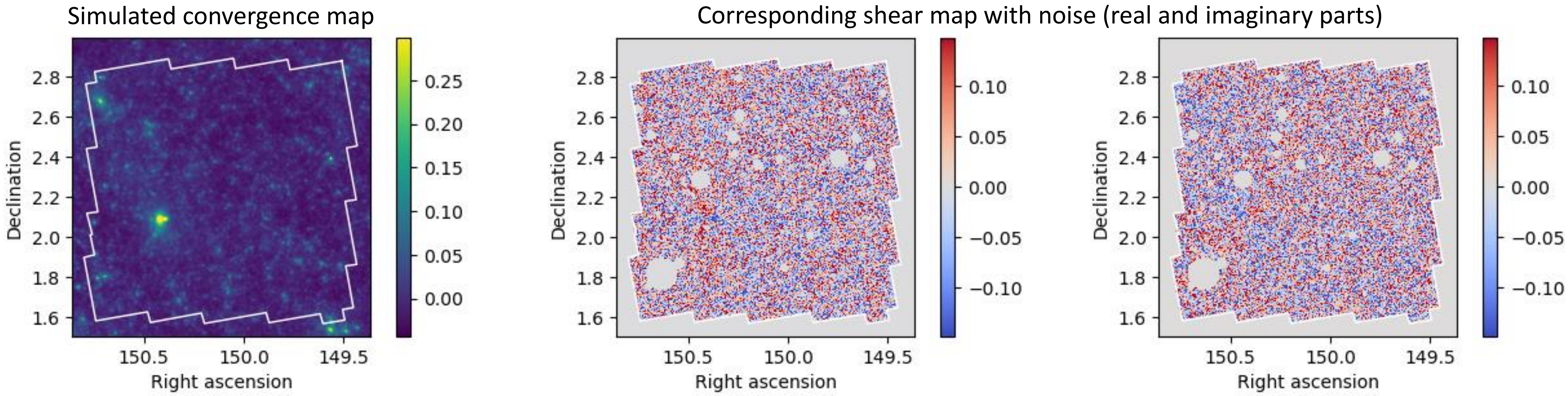
$$\mathbb{E}[L(\boldsymbol{\kappa}, \widehat{\boldsymbol{\kappa}}^{-}, \widehat{\boldsymbol{\kappa}}^{+})] \leq \alpha.$$

- Over which uncertainties the expected value is calculated?

[1] https://astro.uni-bonn.de/en/m/schrabba/research

Context and objectives

Noisy shear maps (noise variance taken from the COSMOS shape catalog[1])



Objective: given $\boldsymbol{\gamma}$, estimate $\widehat{\boldsymbol{\kappa}}^-$ and $\widehat{\boldsymbol{\kappa}}^+$ such that

$$\mathbb{E}[L(\boldsymbol{\kappa}, \widehat{\boldsymbol{\kappa}}^-, \widehat{\boldsymbol{\kappa}}^+)] \leq \alpha.$$

Confidence level $\in \,]0, 1[$

- Over which uncertainties the expected value is calculated?

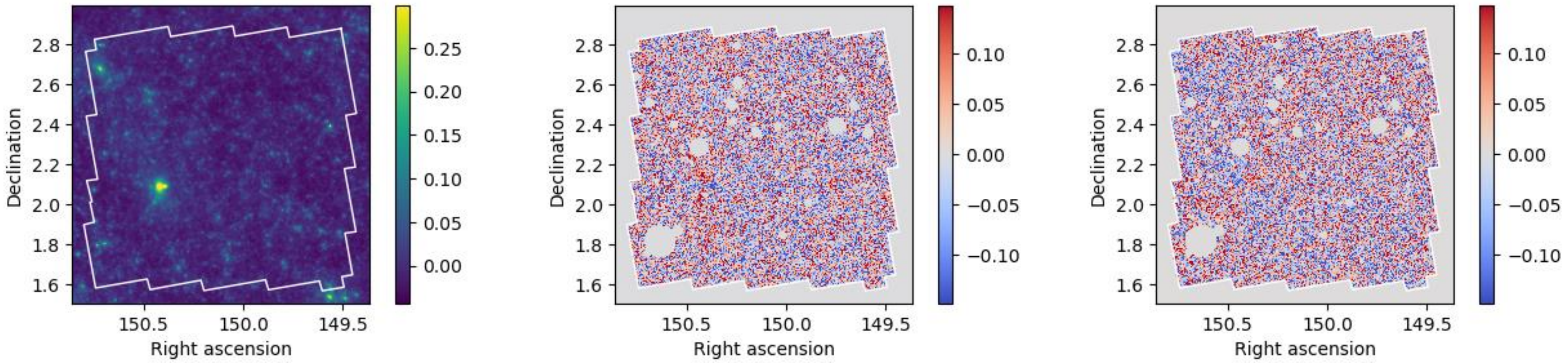
[1] https://astro.uni-bonn.de/en/m/schrabba/research

Context and objectives

Noisy shear maps (noise variance taken from the COSMOS shape catalog[1])

Simulated convergence map

Corresponding shear map with noise (real and imaginary parts)



Objective: given $\boldsymbol{\gamma}$, estimate $\widehat{\boldsymbol{\kappa}}^-$ and $\widehat{\boldsymbol{\kappa}}^+$ such that

$$\mathbb{E}[L(\boldsymbol{\kappa}, \widehat{\boldsymbol{\kappa}}^-, \widehat{\boldsymbol{\kappa}}^+)] \leq \alpha.$$

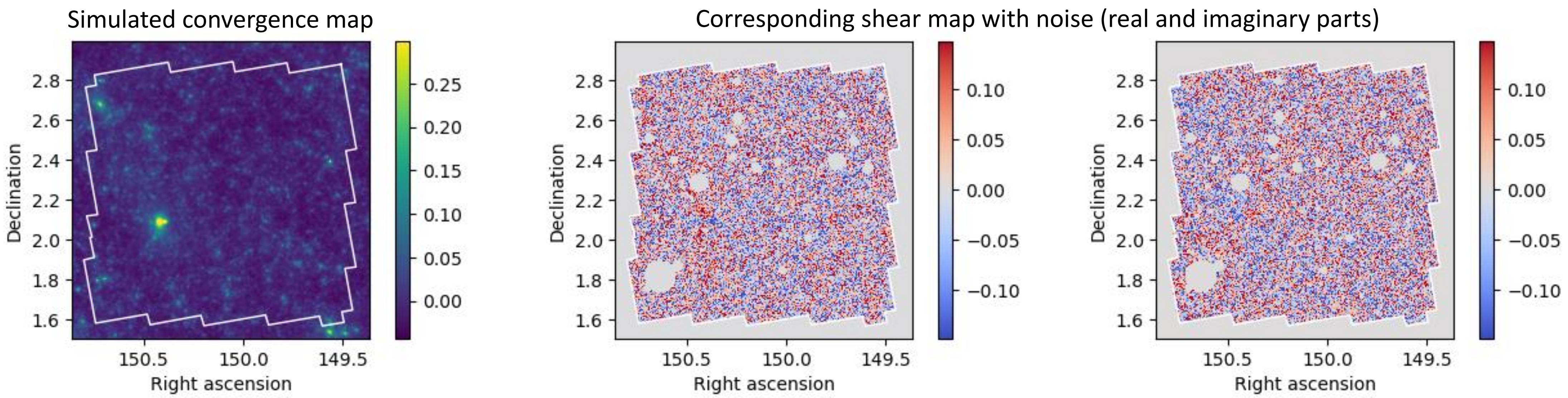
May be random

- Over which uncertainties the expected value is calculated?

[1] https://astro.uni-bonn.de/en/m/schrabba/research

Context and objectives

Noisy shear maps (noise variance taken from the COSMOS shape catalog[1])



Objective: given $\boldsymbol{\gamma}$, estimate $\widehat{\boldsymbol{\kappa}}^-$ and $\widehat{\boldsymbol{\kappa}}^+$ such that

$$\mathbb{E}[L(\boldsymbol{\kappa}, \widehat{\boldsymbol{\kappa}}^-, \widehat{\boldsymbol{\kappa}}^+)] \leq \alpha.$$

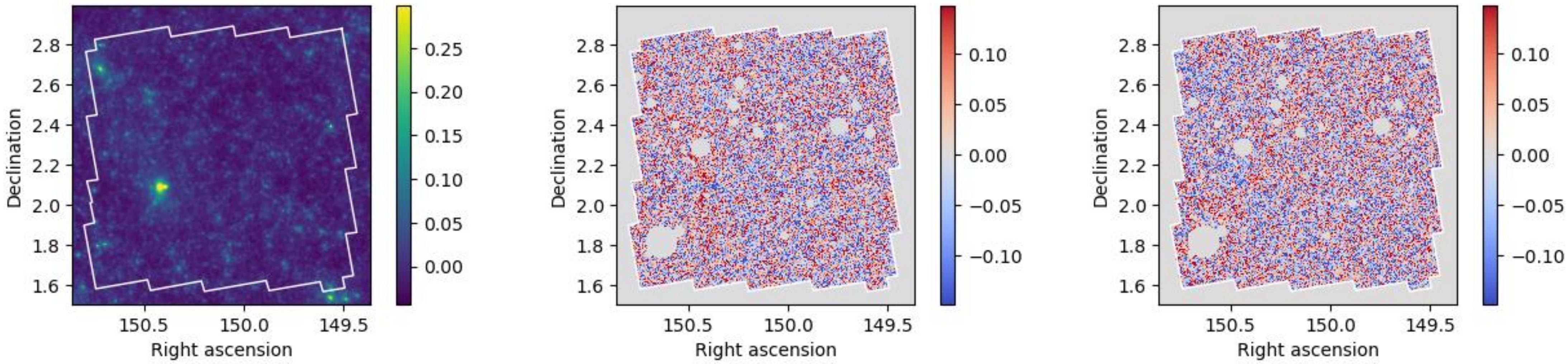
Depends on $\boldsymbol{\gamma} = \mathbf{A}\boldsymbol{\kappa} + \boldsymbol{n}$

- Over which uncertainties the expected value is calculated?

[1] https://astro.uni-bonn.de/en/m/schrabba/research

Context and objectives

Noisy shear maps (noise variance taken from the COSMOS shape catalog[1])



Objective: given $\boldsymbol{\gamma}$, estimate $\widehat{\boldsymbol{\kappa}}^{-}$ and $\widehat{\boldsymbol{\kappa}}^{+}$ such that

$$\mathbb{E}[L(\boldsymbol{\kappa}, \widehat{\boldsymbol{\kappa}}^{-}, \widehat{\boldsymbol{\kappa}}^{+})] \leq \alpha.$$

Two sources of randomness

Depends on $\boldsymbol{\gamma} = \mathbf{A}\boldsymbol{\kappa} + \boldsymbol{n}$

- Over which uncertainties the expected value is calculated?

[1] https://astro.uni-bonn.de/en/m/schrabba/research

Proposed approach

1. Compute a point estimate $\hat{\boldsymbol{\kappa}}$ and a residual $\hat{\boldsymbol{r}}$ using three mass mapping methods:
   a. Kaiser-Squires inversion;[1]
   b. iterative Wiener filtering;[2]
   c. MCALens.[3]
2. Set initial bounds:

$$\hat{\boldsymbol{\kappa}}^- := \hat{\boldsymbol{\kappa}} - \hat{\boldsymbol{r}} \quad \text{and} \quad \hat{\boldsymbol{\kappa}}^+ := \hat{\boldsymbol{\kappa}} + \hat{\boldsymbol{r}}$$

3. Post-processing: adjust residual $\hat{\boldsymbol{r}}$ using a **calibration set**.

→ Distribution-free UQ, does not assume any prior distribution on $\boldsymbol{\kappa}$.

→ Works for any blackbox prediction method, including deep learning.

[1] N. Kaiser and G. Squires, “Mapping the dark matter with weak gravitational lensing,” Astrophysical Journal, vol. 404, no. 2, pp. 441–450, 1993

[2] J. Bobin, J.-L. Starck, F. Sureau, and J. Fadili, “CMB Map Restoration,” Advances in Astronomy, vol. 2012, p. e703217, Apr. 2012

[3] J.-L. Starck, K. E. Themelis, N. Jeffrey, A. Peel, and F. Lanusse, “Weak-lensing mass reconstruction using sparsity and a Gaussian random field,” A&A, vol. 649, p. A99, May 2021

Uncertainty estimation before calibration

- Reminder: problem to solve: $\boldsymbol{\gamma} = \mathbf{A}\boldsymbol{\kappa} + \boldsymbol{n}$, with $\boldsymbol{n} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$.
- **Case 1:** linear operator: $\widehat{\boldsymbol{\kappa}} = \mathbf{B}\boldsymbol{\gamma}$.

$$\widehat{\boldsymbol{\kappa}} = \mathbf{B}\mathbf{A}\boldsymbol{\kappa} + \mathbf{B}\boldsymbol{n}$$
$$\widehat{\boldsymbol{\kappa}} \mid \boldsymbol{\kappa} \sim \mathcal{N}(\mathbf{B}\mathbf{A}\boldsymbol{\kappa}, \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^*)$$

- Hypothesis: $\widehat{\boldsymbol{\kappa}}$ unbiased estimator of $\boldsymbol{\kappa}$, i.e., $\mathbf{B}\mathbf{A}\boldsymbol{\kappa} = \boldsymbol{\kappa}$.
- Then, residual $\boldsymbol{r}$ obtained by considering 1D marginal distributions.

$$\mathbb{E}[L(\boldsymbol{\kappa}, \widehat{\boldsymbol{\kappa}}^-, \widehat{\boldsymbol{\kappa}}^+) \mid \boldsymbol{\kappa}] \leq \alpha$$

- What if hypothesis does not hold?
  - Kaiser-Squires: $\mathbf{B} = \mathbf{S}\mathbf{A}^\dagger$
  - Wiener: wrong if $p(\boldsymbol{\kappa})$ is small → assumes Gaussian prior
- **Proposed solution**: postprocessing step with calibration.

Uncertainty estimation before calibration

- Reminder: problem to solve: $\boldsymbol{\gamma} = \mathbf{A}\boldsymbol{\kappa} + \boldsymbol{n}$, with $\boldsymbol{n} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$.
- **Case 1:** linear operator: $\hat{\boldsymbol{\kappa}} = \mathbf{B}\boldsymbol{\gamma}$.

$$\hat{\boldsymbol{\kappa}} = \mathbf{B}\mathbf{A}\boldsymbol{\kappa} + \mathbf{B}\boldsymbol{n}$$

$$\hat{\boldsymbol{\kappa}} \mid \boldsymbol{\kappa} \sim \mathcal{N}(\mathbf{B}\mathbf{A}\boldsymbol{\kappa}, \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^*)$$

Diagonal elements only

- Hypothesis: $\hat{\boldsymbol{\kappa}}$ unbiased estimator of $\boldsymbol{\kappa}$, i.e., $\mathbf{B}\mathbf{A}\boldsymbol{\kappa} = \boldsymbol{\kappa}$.
- Then, residual $\boldsymbol{r}$ obtained by considering 1D marginal distributions.

$$\mathbb{E}[L(\boldsymbol{\kappa}, \hat{\boldsymbol{\kappa}}^-, \hat{\boldsymbol{\kappa}}^+) \mid \boldsymbol{\kappa}] \leq \alpha$$

- What if hypothesis does not hold?
  - Kaiser-Squires: $\mathbf{B} = \mathbf{S}\mathbf{A}^\dagger$
  - Wiener: wrong if $p(\boldsymbol{\kappa})$ is small → assumes Gaussian prior
- **Proposed solution**: postprocessing step with calibration.

Uncertainty estimation before calibration

- Reminder: problem to solve: $\boldsymbol{\gamma} = \mathbf{A}\boldsymbol{\kappa} + \boldsymbol{n}$, with $\boldsymbol{n} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$.
- **Case 1:** linear operator: $\widehat{\boldsymbol{\kappa}} = \mathbf{B}\boldsymbol{\gamma}$.

$$\widehat{\boldsymbol{\kappa}} = \mathbf{B}\mathbf{A}\boldsymbol{\kappa} + \mathbf{B}\boldsymbol{n}$$

$$\widehat{\boldsymbol{\kappa}} \mid \boldsymbol{\kappa} \sim \mathcal{N}(\mathbf{B}\mathbf{A}\boldsymbol{\kappa}, \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^*)$$

- Hypothesis: $\widehat{\boldsymbol{\kappa}}$ unbiased estimator of $\boldsymbol{\kappa}$, i.e., $\mathbf{B}\mathbf{A}\boldsymbol{\kappa} = \boldsymbol{\kappa}$.
- Then, residual $\boldsymbol{r}$ obtained by considering 1D marginal distributions.

$$\mathbb{E}[L(\boldsymbol{\kappa}, \widehat{\boldsymbol{\kappa}}^-, \widehat{\boldsymbol{\kappa}}^+) \mid \boldsymbol{\kappa}] \leq \alpha$$

- What if hypothesis does not hold?
  - Kaiser-Squires: $\mathbf{B} = \mathbf{S}\mathbf{A}^\dagger$
  - Wiener: wrong if $p(\boldsymbol{\kappa})$ is small → assumes Gaussian prior

Expected value conditionally to $\boldsymbol{\kappa}$

→ $\boldsymbol{n}$ only source of randomness

- **Proposed solution**: postprocessing step with calibration.

Uncertainty estimation before calibration

- Reminder: problem to solve: $\boldsymbol{\gamma} = \mathbf{A}\boldsymbol{\kappa} + \boldsymbol{n}$, with $\boldsymbol{n} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$.
- **Case 2**: nonlinear operator (MCALens): $\widehat{\boldsymbol{\kappa}} = \mathbf{B}(\boldsymbol{\gamma}) \times \boldsymbol{\gamma}$.
- Hypothesis: $\mathbf{B}(\boldsymbol{\gamma})$ stable to noise realizations $\boldsymbol{n}$:

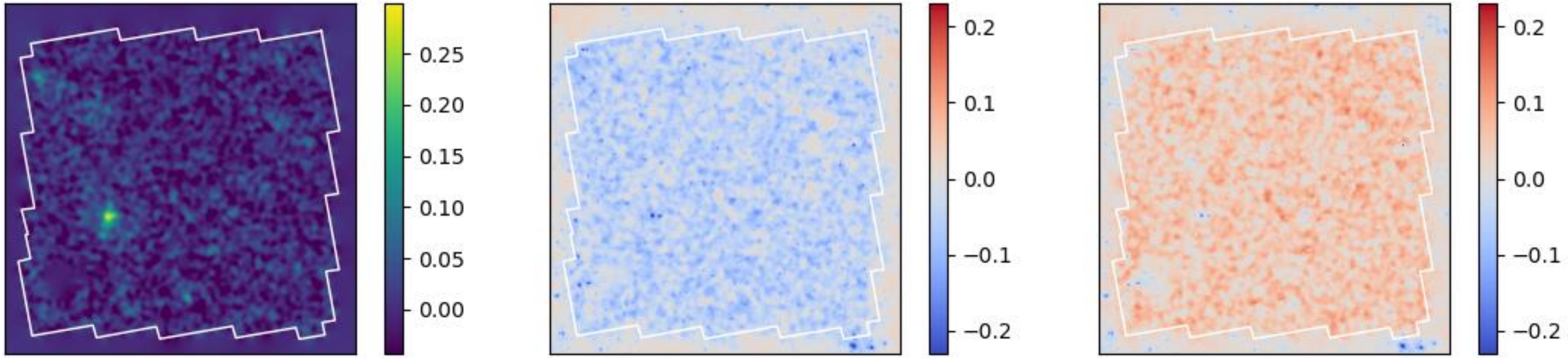
$$\mathbf{B}(\boldsymbol{\gamma}) \approx \mathbf{B}(\mathbf{A}\boldsymbol{\kappa})$$

→ Back to case 1 with $\mathbf{B} \leftarrow \mathbf{B}(\mathbf{A}\boldsymbol{\kappa})$ (linear operator if $\boldsymbol{\kappa}$ is fixed).

Point estimate and uncertainty bounds

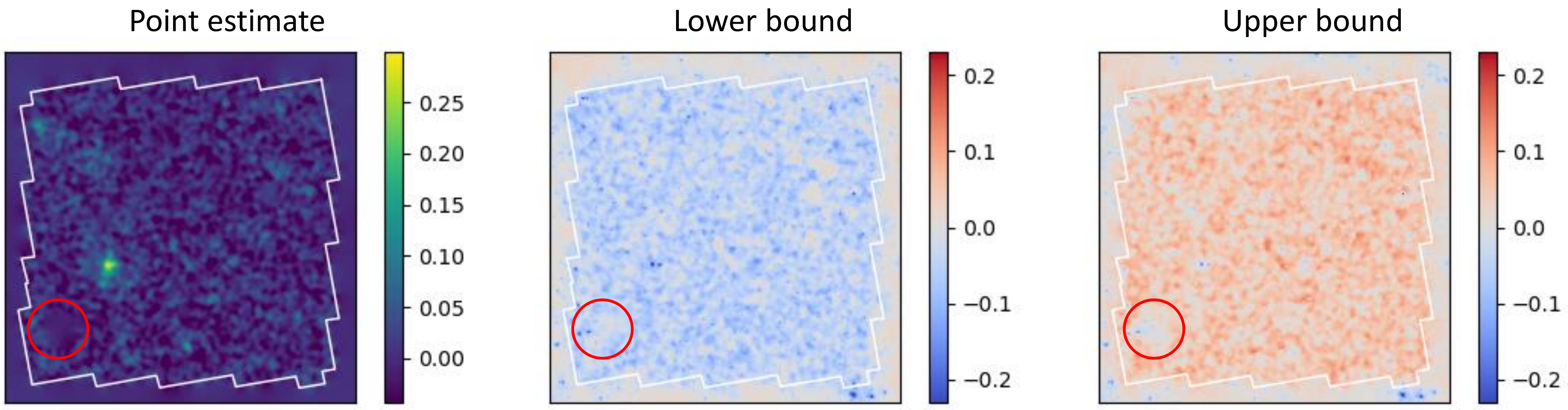
Kaiser-Squires

Point estimate | Lower bound | Upper bound



Point estimate and uncertainty bounds

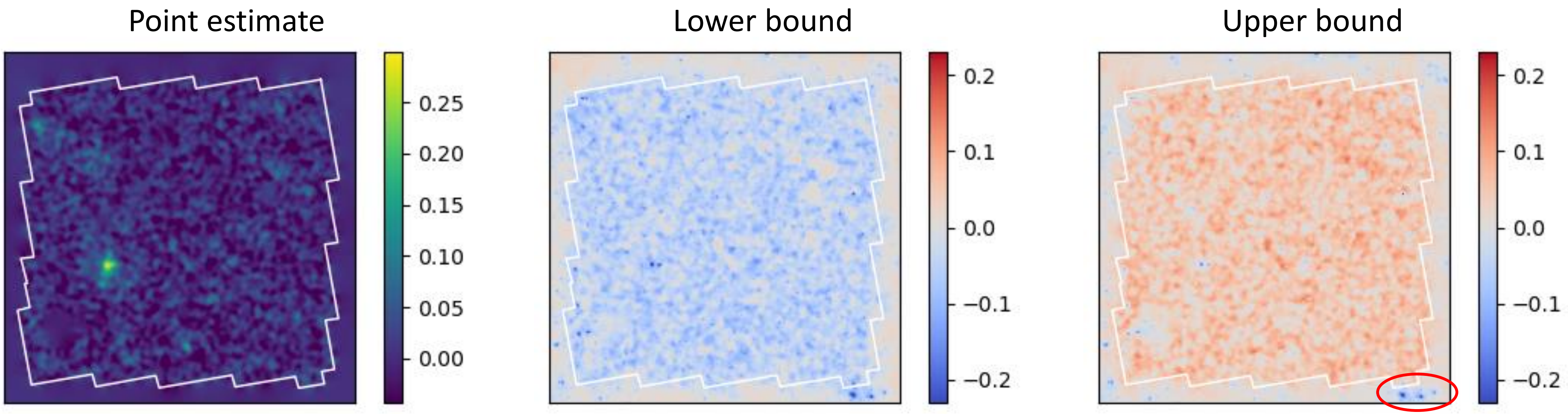
Kaiser-Squires



Mask not properly handled, excluded from results

Point estimate and uncertainty bounds

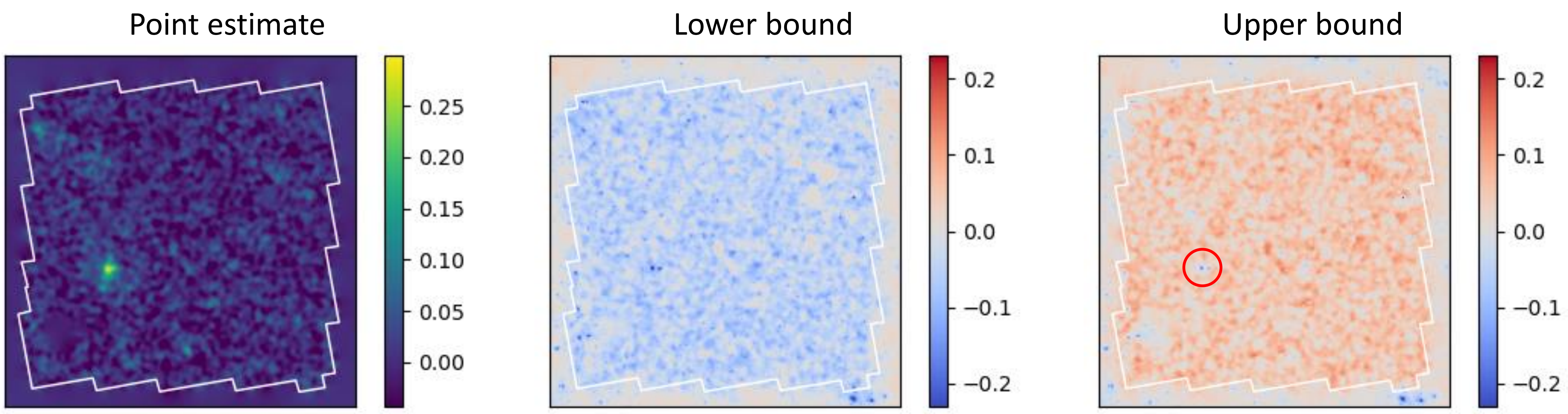
Kaiser-Squires



Undetected features outside survey boundaries

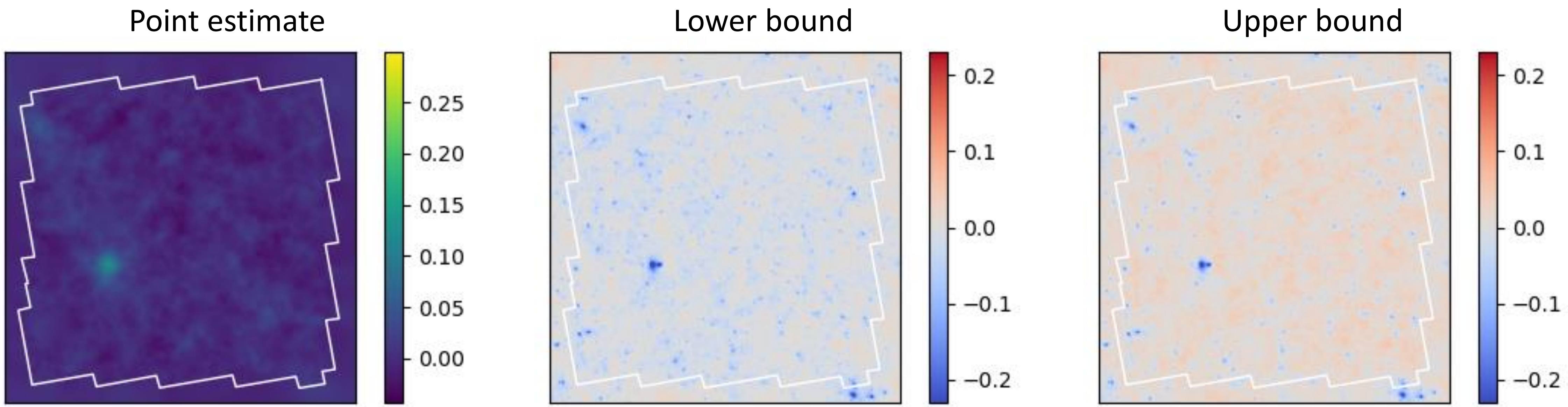
Point estimate and uncertainty bounds

Kaiser-Squires



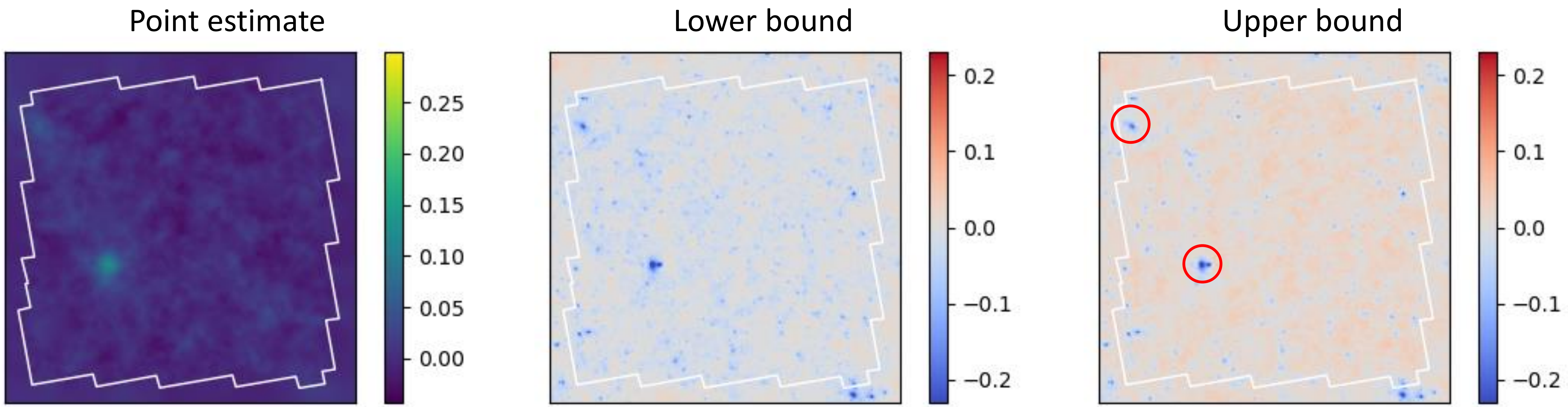
Miscoverage for high-density regions:
ground truth larger than upper bound

Point estimate and uncertainty bounds
Wiener



Point estimate and uncertainty bounds

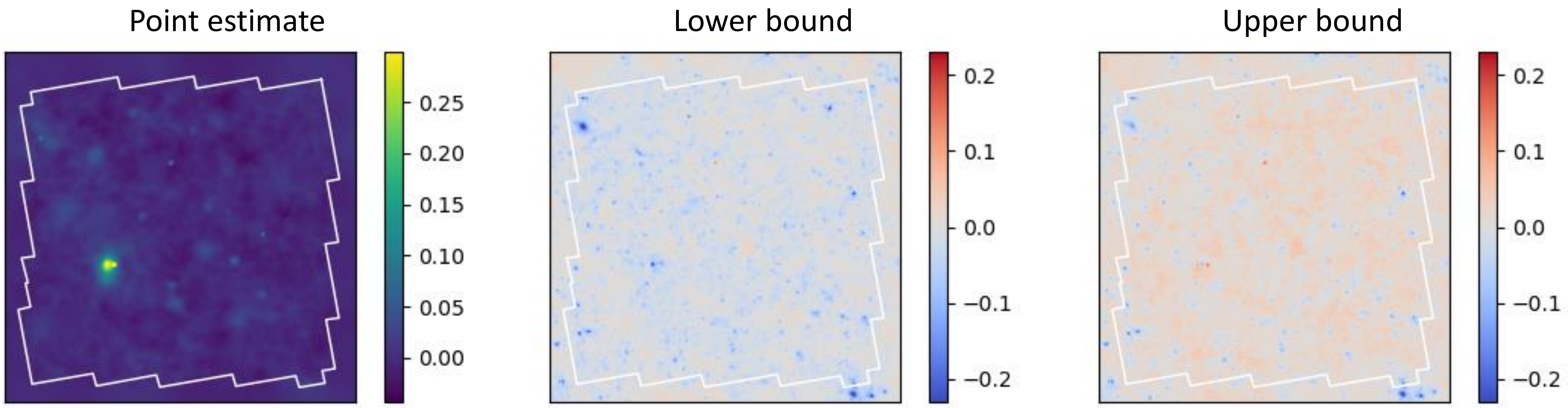
Wiener



Miscoverage for high-density regions:
ground truth larger than upper bound

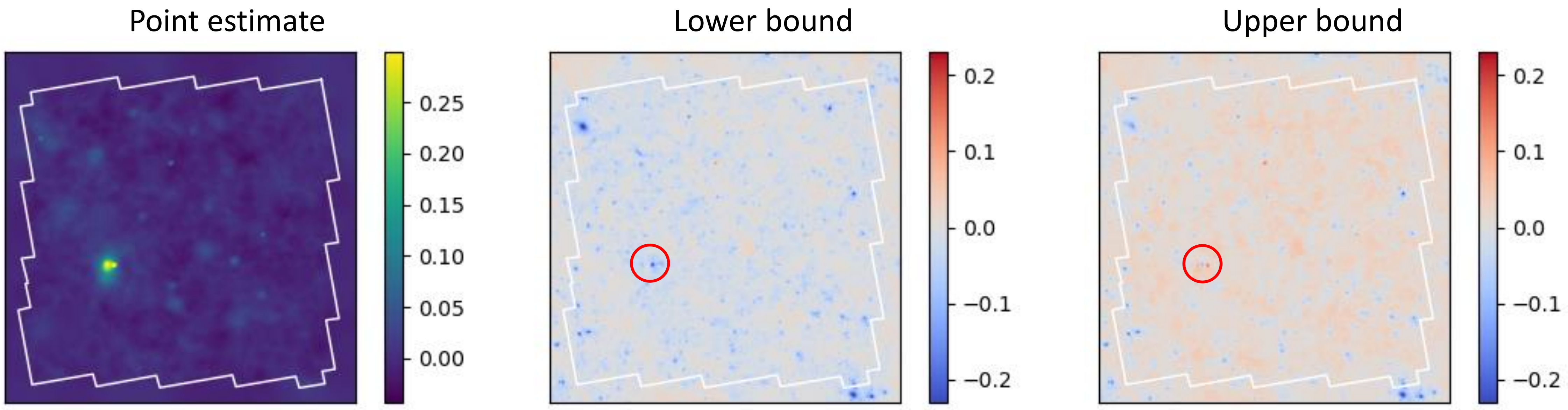
Point estimate and uncertainty bounds

MCALens



Point estimate and uncertainty bounds

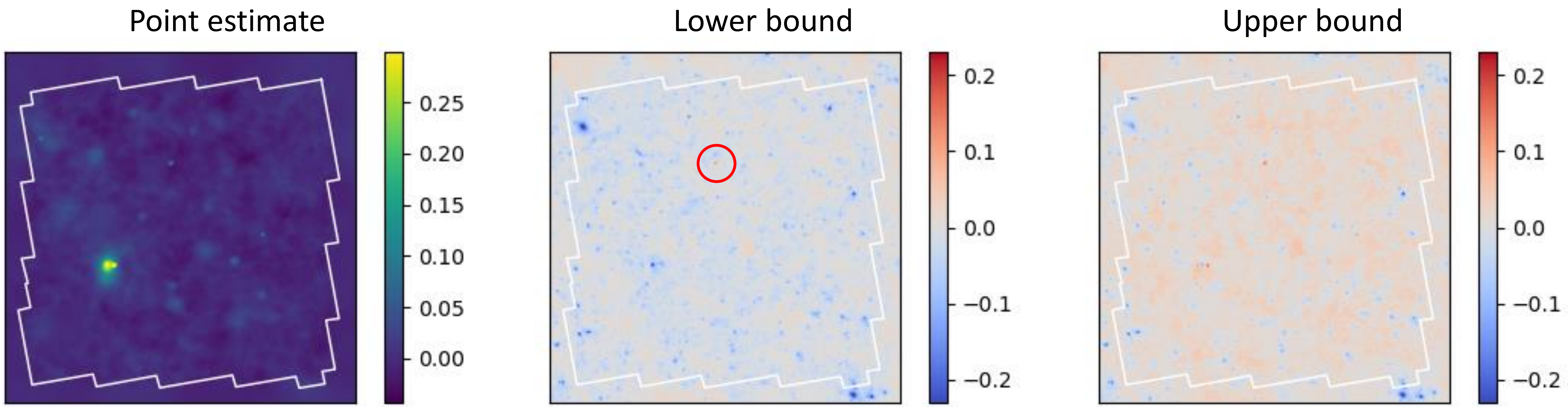
MCALens



Higher uncertainty near high-density regions

Point estimate and uncertainty bounds

MCALens



Ground truth smaller that lower bound. Hallucination?

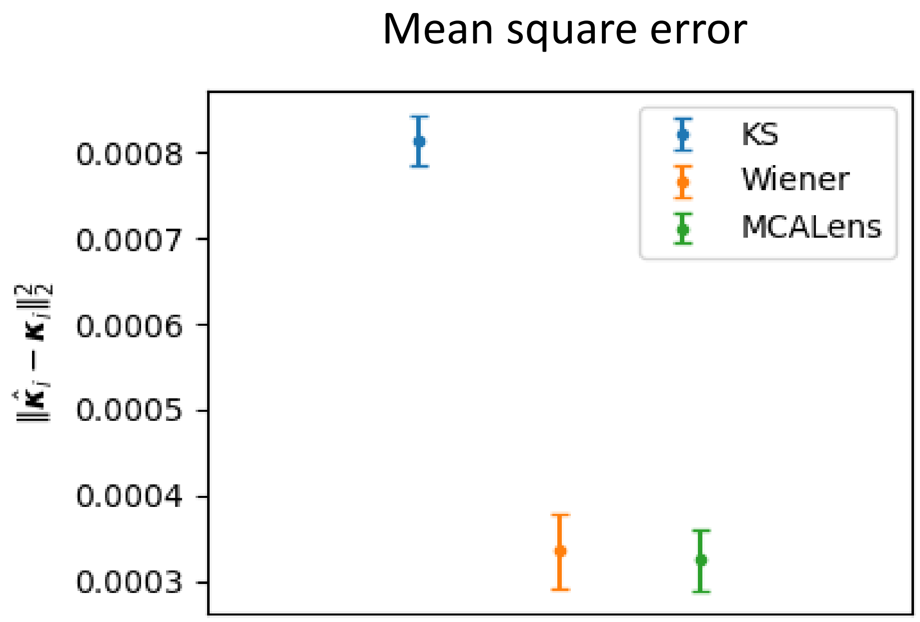
Results before calibration

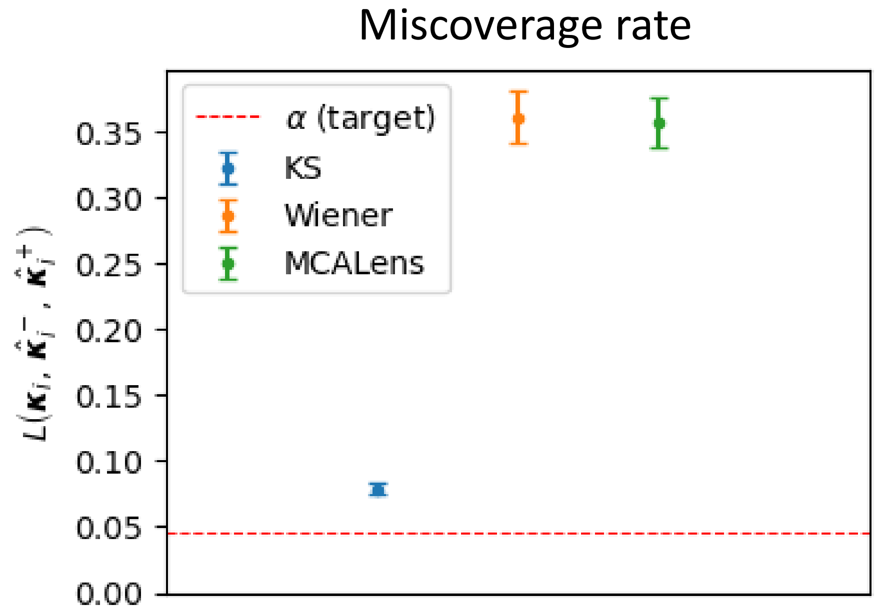
- Target: 2σ-confidence ($\alpha \approx 4.6\%$).
- MSE and rate of ill-predicted pixels computed on a test set of 125 images.

Predictions are way above target!

→ **Undercoverage**

→ **Calibration needed**





Calibration methods

Objective (reminder): given $\boldsymbol{\gamma}$, estimate $\widehat{\boldsymbol{\kappa}}^-$ and $\widehat{\boldsymbol{\kappa}}^+$ such that

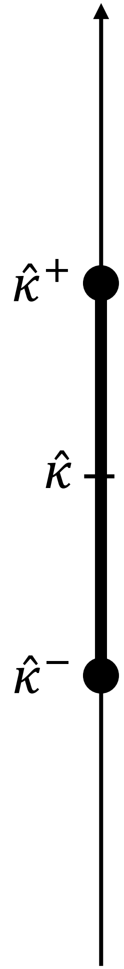
$$\mathbb{E}[L(\boldsymbol{\kappa}, \widehat{\boldsymbol{\kappa}}^-, \widehat{\boldsymbol{\kappa}}^+)] \leq \alpha.$$

Two postprocessing **calibration methods**:

- Conformalized quantile regression (CQR);[1]
- Risk-controlling prediction sets (RCPS).[2]

General principles: consider a calibration set $(\boldsymbol{\gamma}_i, \boldsymbol{\kappa}_i)_{i=1}^n$.

1. Compute point estimates $\widehat{\boldsymbol{\kappa}}_i$ and residuals $\widehat{\boldsymbol{r}}_i$ for each input;
2. Compute a calibration parameter λ from $(\widehat{\boldsymbol{\kappa}}_i, \widehat{\boldsymbol{r}}_i, \boldsymbol{\kappa}_i)_{i=1}^n$ and α;
3. Adjust the residual $\widehat{\boldsymbol{r}}$, using a calibration function g_λ.



[1] Y. Romano, E. Patterson, and E. Candès, “Conformalized Quantile Regression,” NeurIPS, 2019

[2] A. N. Angelopoulos et al., “Image-to-Image Regression with Distribution-Free UQ and Applications in Imaging,” ICML, 2022

Calibration methods

Objective (reminder): given $\boldsymbol{\gamma}$, estimate $\widehat{\boldsymbol{\kappa}}^-$ and $\widehat{\boldsymbol{\kappa}}^+$ such that

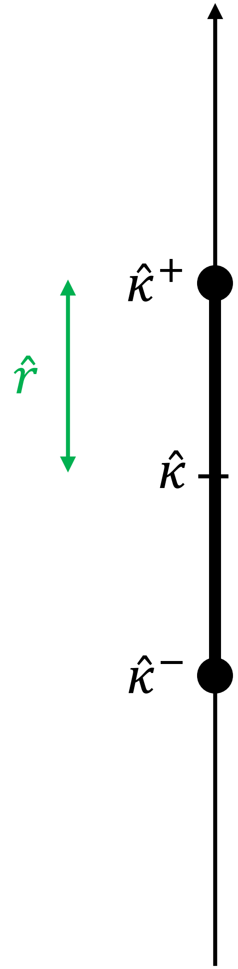
$$\mathbb{E}[L(\boldsymbol{\kappa}, \widehat{\boldsymbol{\kappa}}^-, \widehat{\boldsymbol{\kappa}}^+)] \leq \alpha.$$

Two postprocessing **calibration methods**:

- Conformalized quantile regression (CQR);[1]
- Risk-controlling prediction sets (RCPS).[2]

General principles: consider a calibration set $(\boldsymbol{\gamma}_i, \boldsymbol{\kappa}_i)_{i=1}^n$.

1. Compute point estimates $\widehat{\boldsymbol{\kappa}}_i$ and residuals $\widehat{\boldsymbol{r}}_i$ for each input;
2. Compute a calibration parameter λ from $(\widehat{\boldsymbol{\kappa}}_i, \widehat{\boldsymbol{r}}_i, \boldsymbol{\kappa}_i)_{i=1}^n$ and α;
3. Adjust the residual $\widehat{\boldsymbol{r}}$, using a calibration function g_λ.



[1] Y. Romano, E. Patterson, and E. Candès, "Conformalized Quantile Regression," NeurIPS, 2019

[2] A. N. Angelopoulos et al., "Image-to-Image Regression with Distribution-Free UQ and Applications in Imaging," ICML, 2022

Calibration methods

Objective (reminder): given $\boldsymbol{\gamma}$, estimate $\hat{\boldsymbol{\kappa}}^-$ and $\hat{\boldsymbol{\kappa}}^+$ such that

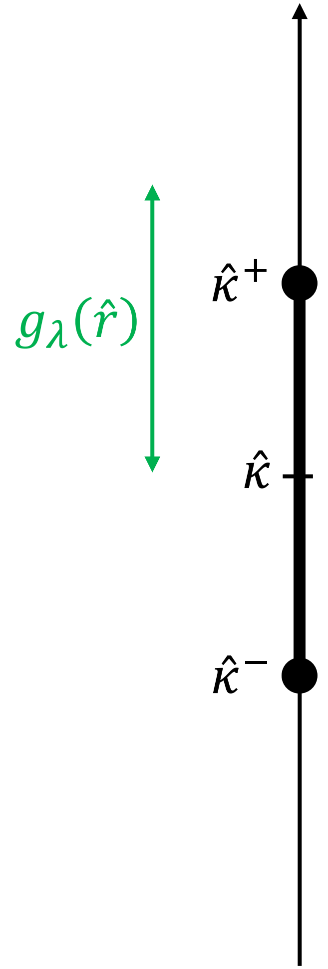
$$\mathbb{E}[L(\boldsymbol{\kappa}, \hat{\boldsymbol{\kappa}}^-, \hat{\boldsymbol{\kappa}}^+)] \leq \alpha.$$

Two postprocessing **calibration methods**:

- Conformalized quantile regression (CQR);[1]
- Risk-controlling prediction sets (RCPS).[2]

General principles: consider a calibration set $(\boldsymbol{\gamma}_i, \boldsymbol{\kappa}_i)_{i=1}^n$.

1. Compute point estimates $\hat{\boldsymbol{\kappa}}_i$ and residuals $\hat{\boldsymbol{r}}_i$ for each input;
2. Compute a calibration parameter λ from $(\hat{\boldsymbol{\kappa}}_i, \hat{\boldsymbol{r}}_i, \boldsymbol{\kappa}_i)_{i=1}^n$ and α;
3. Adjust the residual $\hat{\boldsymbol{r}}$, using a calibration function g_λ.



[1] Y. Romano, E. Patterson, and E. Candès, “Conformalized Quantile Regression,” NeurIPS, 2019

[2] A. N. Angelopoulos et al., “Image-to-Image Regression with Distribution-Free UQ and Applications in Imaging,” ICML, 2022

Calibration methods

Objective (reminder): given $\boldsymbol{\gamma}$, estimate $\widehat{\boldsymbol{\kappa}}^-$ and $\widehat{\boldsymbol{\kappa}}^+$ such that

$$\mathbb{E}[L(\boldsymbol{\kappa}, \widehat{\boldsymbol{\kappa}}^-, \widehat{\boldsymbol{\kappa}}^+)] \leq \alpha.$$

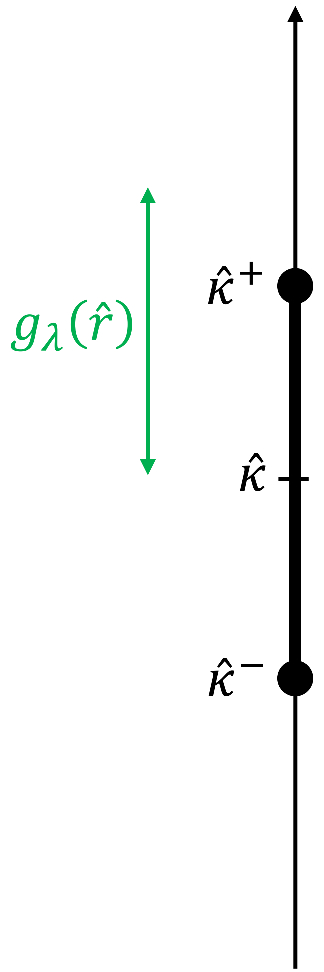
Two postprocessing **calibration methods**:

- Conformalized quantile regression (CQR);[1]
- Risk-controlling prediction sets (RCPS).[2]

General principles: consider a calibration set $(\boldsymbol{\gamma}_i, \boldsymbol{\kappa}_i)_{i=1}^n$.

1. Compute point estimates $\widehat{\boldsymbol{\kappa}}_i$ and residuals $\widehat{\boldsymbol{r}}_i$ for each input;
2. Compute a calibration parameter λ from $(\widehat{\boldsymbol{\kappa}}_i, \widehat{\boldsymbol{r}}_i, \boldsymbol{\kappa}_i)_{i=1}^n$ and α;
3. Adjust the residual $\widehat{\boldsymbol{r}}$, using a calibration function g_λ.

E.g., $g_\lambda(\hat{r}) = \hat{r} + \lambda$



[1] Y. Romano, E. Patterson, and E. Candès, "Conformalized Quantile Regression," NeurIPS, 2019

[2] A. N. Angelopoulos et al., "Image-to-Image Regression with Distribution-Free UQ and Applications in Imaging," ICML, 2022

Calibration methods

Objective (reminder): given $\boldsymbol{\gamma}$, estimate $\widehat{\boldsymbol{\kappa}}^-$ and $\widehat{\boldsymbol{\kappa}}^+$ such that

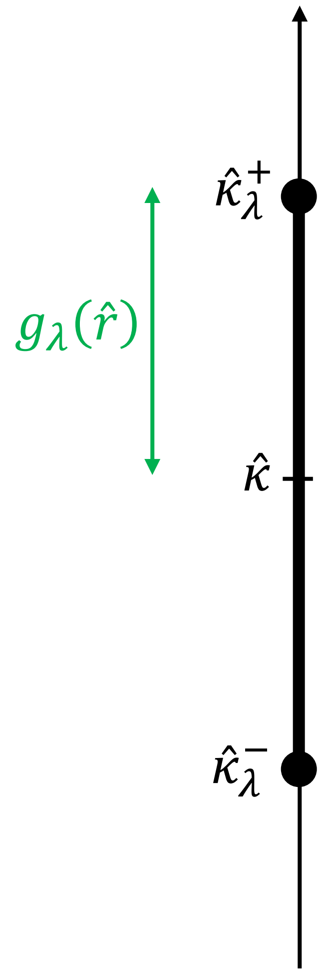
$$\mathbb{E}[L(\boldsymbol{\kappa}, \widehat{\boldsymbol{\kappa}}^-, \widehat{\boldsymbol{\kappa}}^+)] \leq \alpha.$$

Two postprocessing **calibration methods**:

- Conformalized quantile regression (CQR);[1]
- Risk-controlling prediction sets (RCPS).[2]

General principles: consider a calibration set $(\boldsymbol{\gamma}_i, \boldsymbol{\kappa}_i)_{i=1}^n$.

1. Compute point estimates $\widehat{\boldsymbol{\kappa}}_i$ and residuals $\widehat{\boldsymbol{r}}_i$ for each input;
2. Compute a calibration parameter λ from $(\widehat{\boldsymbol{\kappa}}_i, \widehat{\boldsymbol{r}}_i, \boldsymbol{\kappa}_i)_{i=1}^n$ and α;
3. Adjust the residual $\widehat{\boldsymbol{r}}$, using a calibration function g_λ.



[1] Y. Romano, E. Patterson, and E. Candès, "Conformalized Quantile Regression," NeurIPS, 2019

[2] A. N. Angelopoulos et al., "Image-to-Image Regression with Distribution-Free UQ and Applications in Imaging," ICML, 2022

Calibration methods

Objective (reminder): given $\boldsymbol{\gamma}$, estimate $\hat{\boldsymbol{\kappa}}^-$ and $\hat{\boldsymbol{\kappa}}^+$ such that

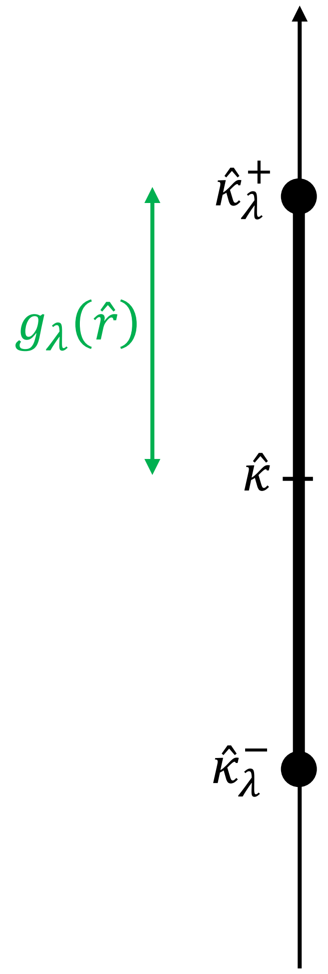
$$\mathbb{E}[L(\boldsymbol{\kappa}, \hat{\boldsymbol{\kappa}}^-, \hat{\boldsymbol{\kappa}}^+)] \leq \alpha.$$

Two postprocessing **calibration methods**:

- Conformalized quantile regression (CQR);[1]
- Risk-controlling prediction sets (RCPS).[2]

General principles: consider a calibration set $(\boldsymbol{\gamma}_i, \boldsymbol{\kappa}_i)_{i=1}^n$.

1. Compute point estimates $\hat{\boldsymbol{\kappa}}_i$ and residuals $\hat{\boldsymbol{r}}_i$ for each input;
2. Compute a calibration parameter λ from $(\hat{\boldsymbol{\kappa}}_i, \hat{\boldsymbol{r}}_i, \boldsymbol{\kappa}_i)_{i=1}^n$ and α;
3. Adjust the residual $\hat{\boldsymbol{r}}$, using a calibration function g_λ.



Works for any blackbox predictor!

[1] Y. Romano, E. Patterson, and E. Candès, “Conformalized Quantile Regression,” NeurIPS, 2019

[2] A. N. Angelopoulos et al., “Image-to-Image Regression with Distribution-Free UQ and Applications in Imaging,” ICML, 2022

Calibration methods

Comparative table

	CQR	**RCPS**
Calibration parameter λ	Different for each pixel	Shared over the whole image
Calculated using	The $(1-\alpha)(1+1/n)$-th quantile of a conformity score	Hoeffding's upper confidence bound
Depends on	α, n	α, δ, n
Theoretical guarantees	$\alpha - 1/n \leq \mathbb{E}[L(\boldsymbol{\kappa}, \widehat{\boldsymbol{\kappa}}^-, \widehat{\boldsymbol{\kappa}}^+)] \leq \alpha$	$\mathbb{P}\{\mathbb{E}[L(\boldsymbol{\kappa}, \widehat{\boldsymbol{\kappa}}^-, \widehat{\boldsymbol{\kappa}}^+) \mid (\boldsymbol{\gamma}_i, \boldsymbol{\kappa}_i)_{i=1}^n] \leq \alpha\} \geq 1-\delta$

Calibration methods

Comparative table

	CQR	RCPS
Calibration parameter λ	Different for each pixel	Shared over the whole image
Calculated using	The $(1-\alpha)(1+1/n)$-th quantile of a conformity score	Hoeffding's upper confidence bound Additional parameter
Depends on	α, n	α, δ, n
Theoretical guarantees	$\alpha - 1/n \leq \mathbb{E}[L(\boldsymbol{\kappa}, \widehat{\boldsymbol{\kappa}}^-, \widehat{\boldsymbol{\kappa}}^+)] \leq \alpha$	$\mathbb{P}\{\mathbb{E}[L(\boldsymbol{\kappa}, \widehat{\boldsymbol{\kappa}}^-, \widehat{\boldsymbol{\kappa}}^+) \mid (\boldsymbol{\gamma}_i, \boldsymbol{\kappa}_i)_{i=1}^n] \leq \alpha\} \geq 1 - \delta$

Calibration methods

Comparative table

	CQR	RCPS
Calibration parameter λ	Different for each pixel	Shared over the whole image
Calculated using	The $(1-\alpha)(1+1/n)$-th quantile of a conformity score	Hoeffding's upper confidence bound
Depends on	α, n	α, δ, n
Theoretical guarantees	$\alpha - 1/n \leq \mathbb{E}[L(\boldsymbol{\kappa}, \widehat{\boldsymbol{\kappa}}^-, \widehat{\boldsymbol{\kappa}}^+)] \leq \alpha$	$\mathbb{P}\{\mathbb{E}[L(\boldsymbol{\kappa}, \widehat{\boldsymbol{\kappa}}^-, \widehat{\boldsymbol{\kappa}}^+) \mid (\boldsymbol{\gamma}_i, \boldsymbol{\kappa}_i)_{i=1}^n] \leq \alpha\} \geq 1 - \delta$

Upper bound: coverage guarantee

Calibration methods

Comparative table

	CQR	RCPS
Calibration parameter λ	Different for each pixel	Shared over the whole image
Calculated using	The $(1-\alpha)(1+1/n)$-th quantile of a conformity score	Hoeffding's upper confidence bound
Depends on	α, n	α, δ, n
Theoretical guarantees	$\alpha - 1/n \leq \mathbb{E}[L(\boldsymbol{\kappa}, \hat{\boldsymbol{\kappa}}^-, \hat{\boldsymbol{\kappa}}^+)] \leq \alpha$	$\mathbb{P}\{\mathbb{E}[L(\boldsymbol{\kappa}, \hat{\boldsymbol{\kappa}}^-, \hat{\boldsymbol{\kappa}}^+) \mid (\boldsymbol{\gamma}_i, \boldsymbol{\kappa}_i)_{i=1}^n] \leq \alpha\} \geq 1-\delta$

Lower bound: prevents overconservative prediction bounds. Only for CQR!

Calibration methods

Comparative table

	CQR	**RCPS**
Calibration parameter λ	Different for each pixel	Shared over the whole image
Calculated using	The $(1-\alpha)(1+1/n)$-th quantile of a conformity score	Hoeffding's upper confidence bound
Depends on	α, n	α, δ, n
Theoretical guarantees	$\alpha - 1/n \leq \mathbb{E}[L(\boldsymbol{\kappa}, \hat{\boldsymbol{\kappa}}^-, \hat{\boldsymbol{\kappa}}^+)] \leq \alpha$	$\mathbb{P}\{\mathbb{E}[L(\boldsymbol{\kappa}, \hat{\boldsymbol{\kappa}}^-, \hat{\boldsymbol{\kappa}}^+) \mid (\boldsymbol{\gamma}_i, \boldsymbol{\kappa}_i)_{i=1}^n] \leq \alpha\} \geq 1 - \delta$

Three sources of randomness:

- ground-truth convergence maps $\boldsymbol{\kappa}$;
- noise $\boldsymbol{n}$, since $\boldsymbol{\gamma} = \mathbf{A}\boldsymbol{\kappa} + \boldsymbol{n}$;
- calibration set $(\boldsymbol{\gamma}_i, \boldsymbol{\kappa}_i)_{i=1}^n$.

Calibration methods

Comparative table

	CQR	**RCPS**
Calibration parameter λ	Different for each pixel	Shared over the whole image
Calculated using	The $(1-\alpha)(1+1/n)$-th quantile of a conformity score	Hoeffding's upper confidence bound
Depends on	α, n	α, δ, n
Theoretical guarantees	$\alpha - 1/n \leq \mathbb{E}[L(\boldsymbol{\kappa}, \widehat{\boldsymbol{\kappa}}^-, \widehat{\boldsymbol{\kappa}}^+)] \leq \alpha$	$\mathbb{P}\{\mathbb{E}[L(\boldsymbol{\kappa}, \widehat{\boldsymbol{\kappa}}^-, \widehat{\boldsymbol{\kappa}}^+) \mid (\boldsymbol{\gamma}_i, \boldsymbol{\kappa}_i)_{i=1}^n] \leq \alpha\} \geq 1-\delta$

CQR: expected value computed over:

- ground-truth convergence maps $\boldsymbol{\kappa}$;
- noise $\boldsymbol{n}$, since $\boldsymbol{\gamma} = \mathbf{A}\boldsymbol{\kappa} + \boldsymbol{n}$;
- calibration set $(\boldsymbol{\gamma}_i, \boldsymbol{\kappa}_i)_{i=1}^n$.

Calibration methods

Comparative table

	CQR	RCPS
Calibration parameter λ	Different for each pixel	Shared over the whole image
Calculated using	The $(1-\alpha)(1+1/n)$-th quantile of a conformity score	Hoeffding's upper confidence bound
Depends on	α, n	α, δ, n
Theoretical guarantees	$\alpha - 1/n \leq \mathbb{E}[L(\boldsymbol{\kappa}, \hat{\boldsymbol{\kappa}}^{-}, \hat{\boldsymbol{\kappa}}^{+})] \leq \alpha$	$\mathbb{P}\{\mathbb{E}[L(\boldsymbol{\kappa}, \hat{\boldsymbol{\kappa}}^{-}, \hat{\boldsymbol{\kappa}}^{+}) \mid (\boldsymbol{\gamma}_i, \boldsymbol{\kappa}_i)_{i=1}^{n}] \leq \alpha\} \geq 1 - \delta$

Fixed calibration set

RCPS: expected value computed over:

- ground-truth convergence maps $\boldsymbol{\kappa}$;
- noise $\boldsymbol{n}$, since $\boldsymbol{\gamma} = \mathbf{A}\boldsymbol{\kappa} + \boldsymbol{n}$;
- calibration set $(\boldsymbol{\gamma}_i, \boldsymbol{\kappa}_i)_{i=1}^{n}$.

Calibration methods

Comparative table

	CQR	RCPS
Calibration parameter λ	Different for each pixel	Shared over the whole image
Calculated using	The $(1-\alpha)(1+1/n)$-th quantile of a conformity score	Hoeffding's upper confidence bound
Depends on	α, n	α, δ, n
Theoretical guarantees	$\alpha - 1/n \leq \mathbb{E}[L(\boldsymbol{\kappa}, \hat{\boldsymbol{\kappa}}^-, \hat{\boldsymbol{\kappa}}^+)] \leq \alpha$	$\mathbb{P}\{\mathbb{E}[L(\boldsymbol{\kappa}, \hat{\boldsymbol{\kappa}}^-, \hat{\boldsymbol{\kappa}}^+) \mid (\boldsymbol{\gamma}_i, \boldsymbol{\kappa}_i)_{i=1}^n] \leq \alpha\} \geq 1 - \delta$

Controls the risk of selecting a statistically deviant calibration set

RCPS: expected value computed over:

- ground-truth convergence maps $\boldsymbol{\kappa}$;
- noise $\boldsymbol{n}$, since $\boldsymbol{\gamma} = \mathbf{A}\boldsymbol{\kappa} + \boldsymbol{n}$;
- calibration set $(\boldsymbol{\gamma}_i, \boldsymbol{\kappa}_i)_{i=1}^n$.

Calibration methods

Comparative table

	CQR	RCPS
Calibration parameter λ	Different for each pixel	Shared over the whole image
Calculated using	The $(1-\alpha)(1+1/n)$-th quantile of a conformity score	Hoeffding's upper confidence bound
Depends on	α, n	α, δ, n
Theoretical guarantees	$\alpha - 1/n \leq \mathbb{E}[L(\boldsymbol{\kappa}, \widehat{\boldsymbol{\kappa}}^-, \widehat{\boldsymbol{\kappa}}^+)] \leq \alpha$	$\mathbb{P}\{\mathbb{E}[L(\boldsymbol{\kappa}, \widehat{\boldsymbol{\kappa}}^-, \widehat{\boldsymbol{\kappa}}^+) \mid (\boldsymbol{\gamma}_i, \boldsymbol{\kappa}_i)_{i=1}^n] \leq \alpha\} \geq 1-\delta$

Controls the risk of selecting a statistically deviant calibration set

RCPS: expected value computed over:

- ground-truth convergence maps $\boldsymbol{\kappa}$;
- noise $\boldsymbol{n}$, since $\boldsymbol{\gamma} = \mathbf{A}\boldsymbol{\kappa} + \boldsymbol{n}$;
- calibration set $(\boldsymbol{\gamma}_i, \boldsymbol{\kappa}_i)_{i=1}^n$.

Condition for theoretical guarantees: **exchangeability** of calibration and test data.

Results

Miscoverage rate

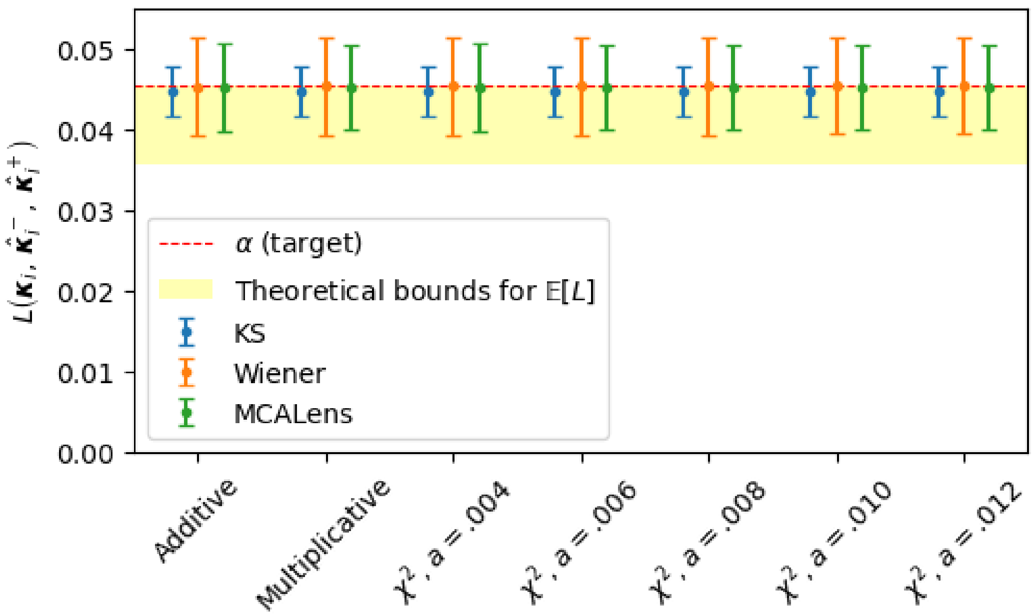
- Calibration set of 100 images from κTNG simulations
- Test set of 125 images from κTNG simulations
- Target: $\alpha \approx 4{,}6\%$ (2σ-confidence)
- CQR: the minimal size depends on the desired confidence level:

  2σ-confidence → $n_{\min} = 21$
  3σ-confidence → $n_{\min} = 370$
  4σ-confidence → $n_{\min} = 15\,787$

Results

Miscoverage rate

CQR



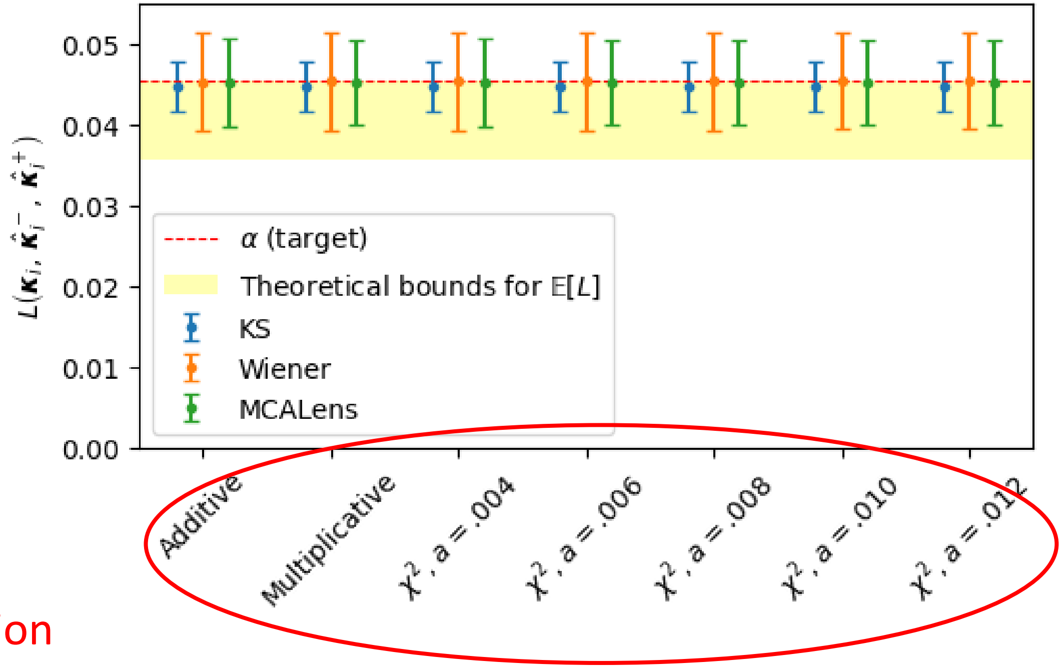
- Calibration set of 100 images from κTNG simulations
- Test set of 125 images from κTNG simulations
- Target: $\alpha \approx 4{,}6\%$ (2σ-confidence)
- CQR: the minimal size depends on the desired confidence level:

  2σ-confidence → $n_{\min} = 21$
  3σ-confidence → $n_{\min} = 370$
  4σ-confidence → $n_{\min} = 15\,787$

Results

Miscoverage rate

CQR



- Calibration set of 100 images from κTNG simulations
- Test set of 125 images from κTNG simulations
- Target: $\alpha \approx 4{,}6\%$ (2σ-confidence)
- CQR: the minimal size depends on the desired confidence level:

  2σ-confidence → $n_{\min} = 21$
  3σ-confidence → $n_{\min} = 370$
  4σ-confidence → $n_{\min} = 15\,787$

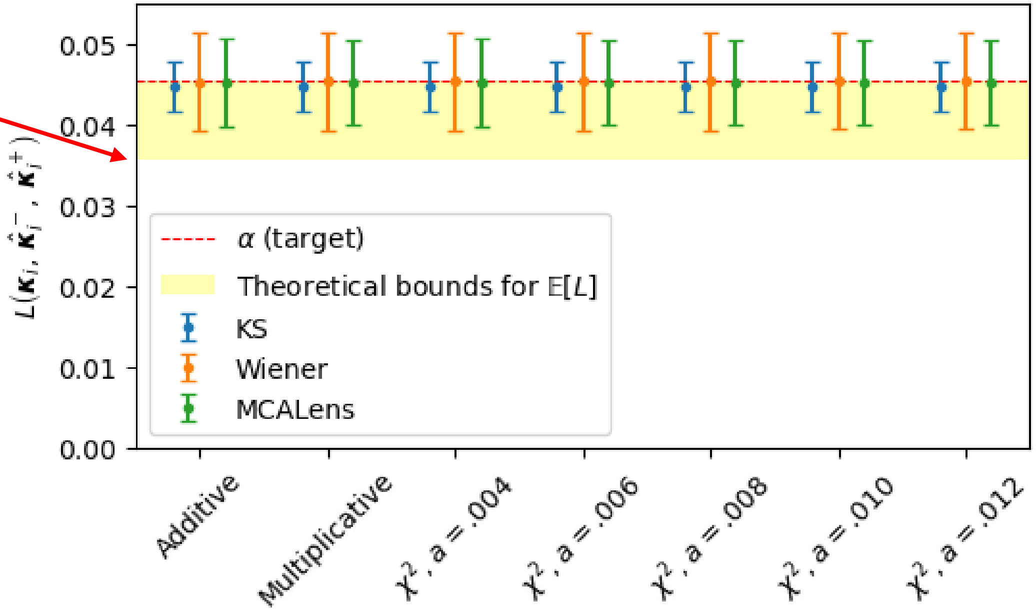
Various calibration functions g_λ

Results

Miscoverage rate

Lower bound $\alpha - 1/n$

CQR



- Calibration set of 100 images from κTNG simulations
- Test set of 125 images from κTNG simulations
- Target: $\alpha \approx 4{,}6\%$ (2σ-confidence)
- CQR: the minimal size depends on the desired confidence level:

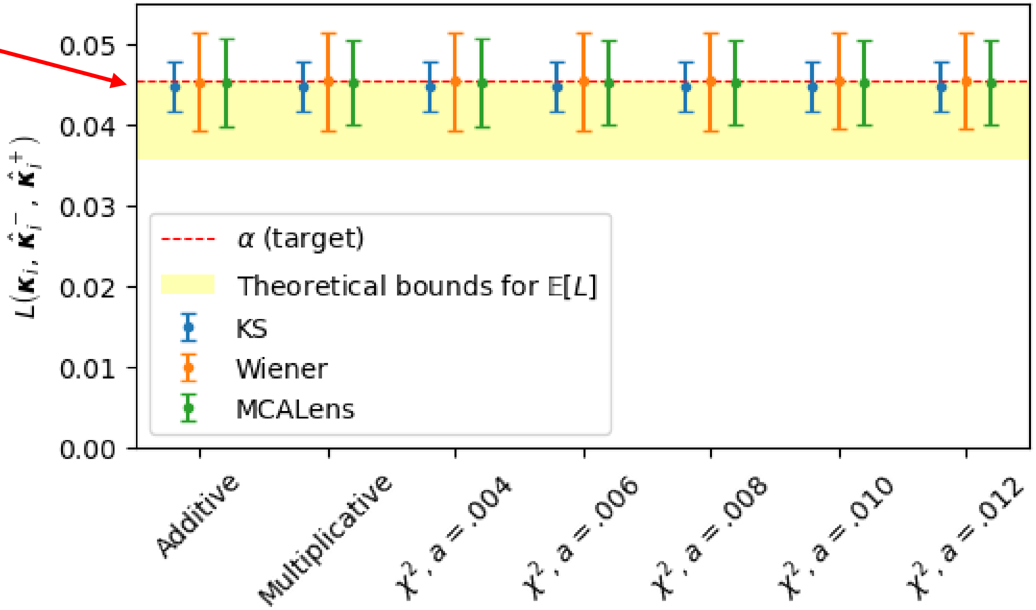
  2σ-confidence → $n_{\min} = 21$
  3σ-confidence → $n_{\min} = 370$
  4σ-confidence → $n_{\min} = 15\,787$

Results

Miscoverage rate

Upper bound α

CQR



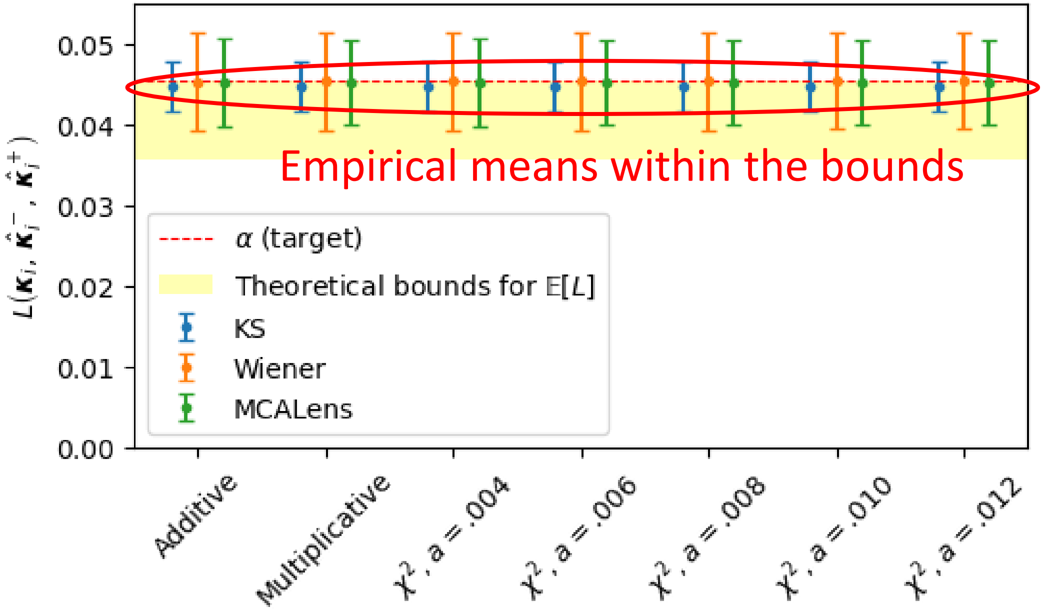
- Calibration set of 100 images from κTNG simulations
- Test set of 125 images from κTNG simulations
- Target: $\alpha \approx 4{,}6\%$ (2σ-confidence)
- CQR: the minimal size depends on the desired confidence level:

  2σ-confidence → $n_{\min} = 21$
  3σ-confidence → $n_{\min} = 370$
  4σ-confidence → $n_{\min} = 15\ 787$

Results

Miscoverage rate

CQR



- Calibration set of 100 images from κTNG simulations
- Test set of 125 images from κTNG simulations
- Target: $\alpha \approx 4{,}6\%$ (2σ-confidence)
- CQR: the minimal size depends on the desired confidence level:

  2σ-confidence → $n_{\min} = 21$
  3σ-confidence → $n_{\min} = 370$
  4σ-confidence → $n_{\min} = 15\,787$

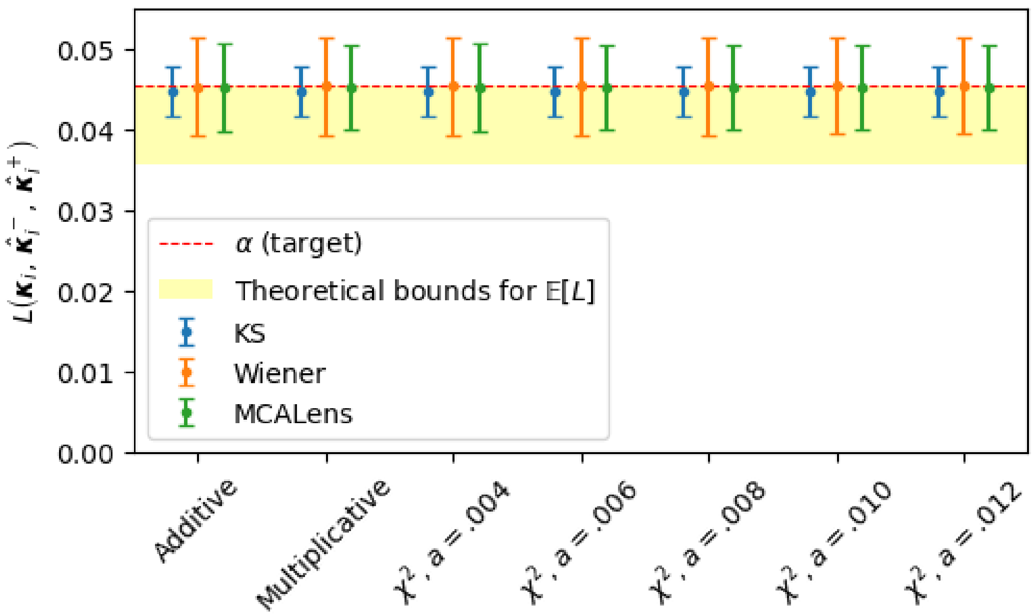
Results

Miscoverage rate

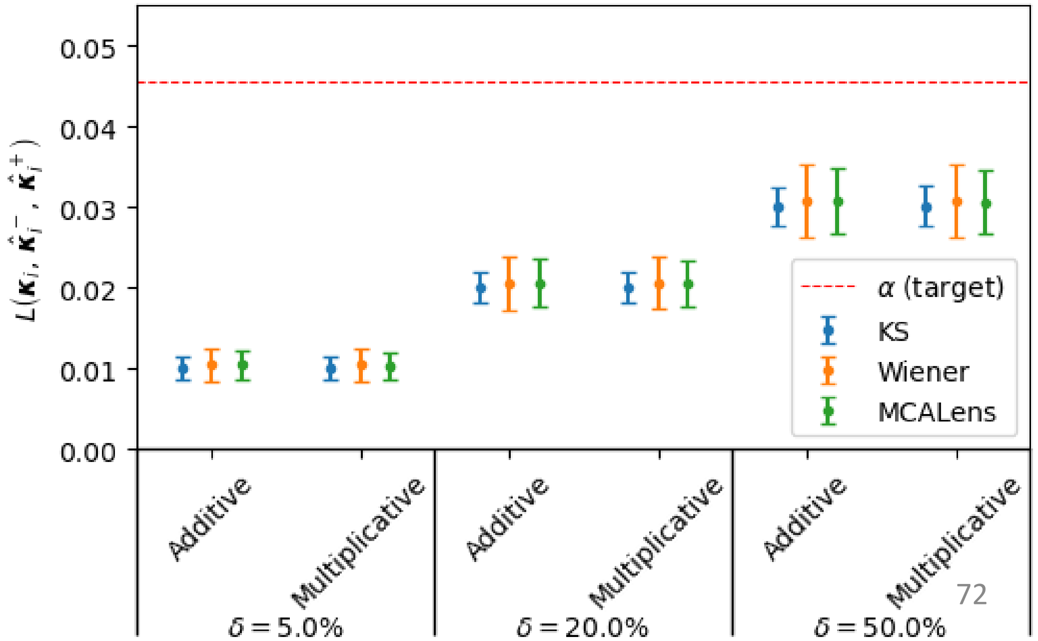
- Calibration set of 100 images from κTNG simulations
- Test set of 125 images from κTNG simulations
- Target: $\alpha \approx 4{,}6\%$ (2σ-confidence)
- CQR: the minimal size depends on the desired confidence level:

  2σ-confidence → $n_{\min} = 21$
  3σ-confidence → $n_{\min} = 370$
  4σ-confidence → $n_{\min} = 15\,787$

CQR



RCPS



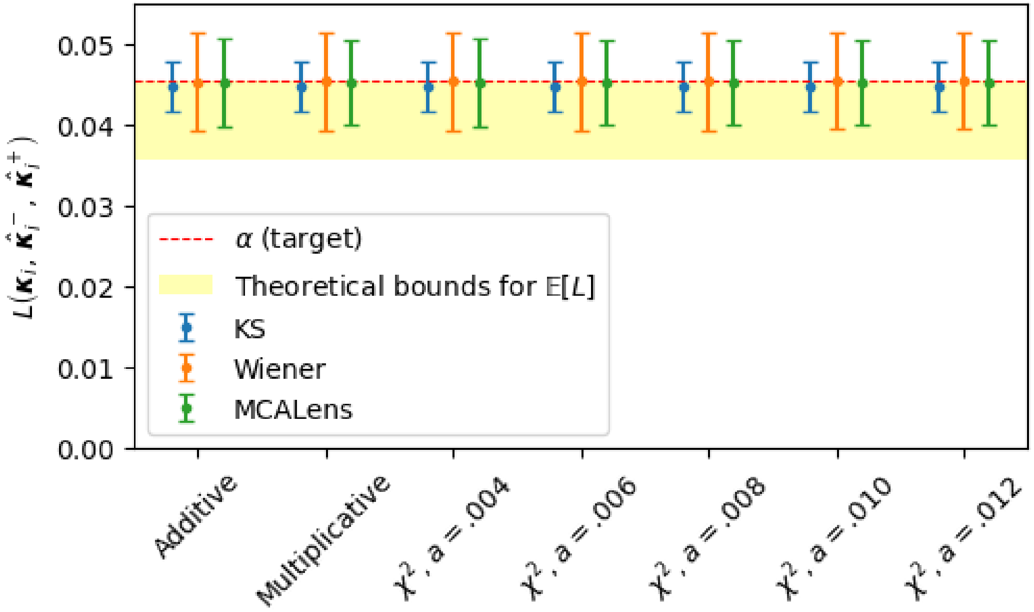
Results

Miscoverage rate

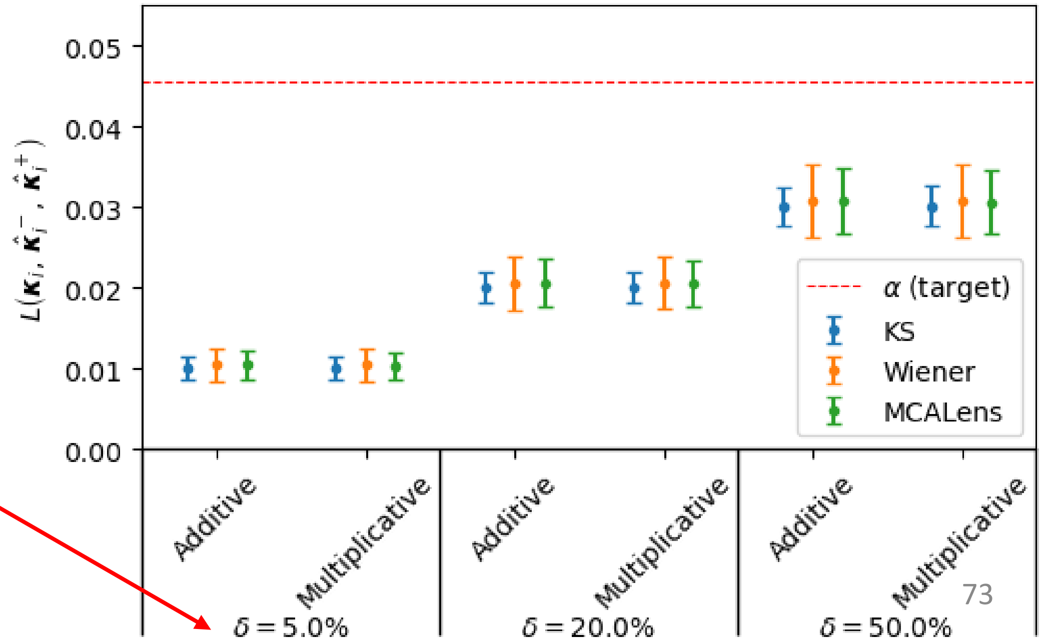
- Calibration set of 100 images from κTNG simulations
- Test set of 125 images from κTNG simulations
- Target: $\alpha \approx 4{,}6\%$ (2σ-confidence)
- CQR: the minimal size depends on the desired confidence level:

  2σ-confidence → $n_{\min} = 21$
  3σ-confidence → $n_{\min} = 370$
  4σ-confidence → $n_{\min} = 15\ 787$

CQR



RCPS



Various values of δ

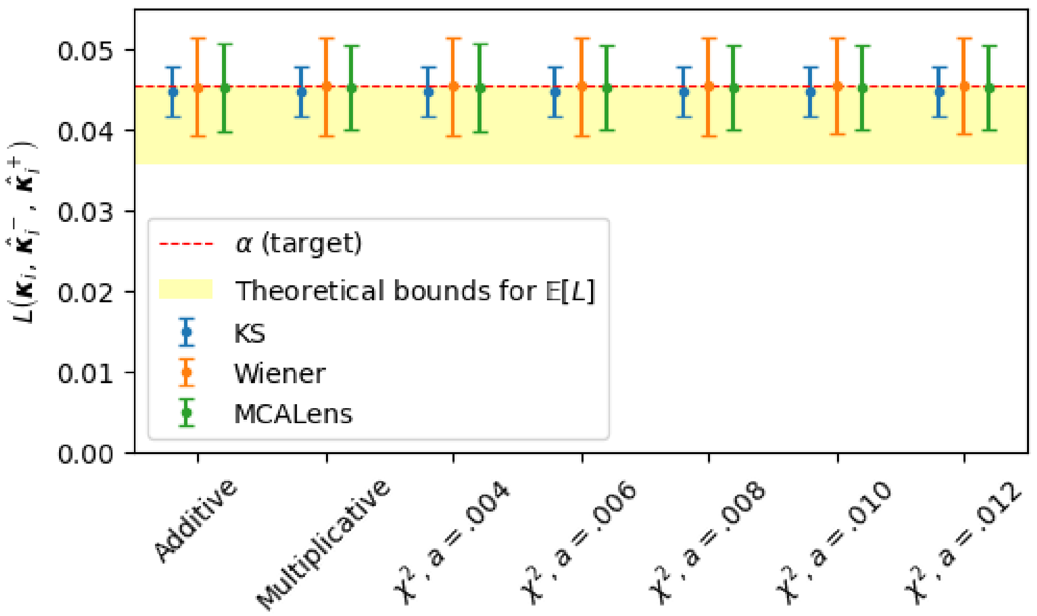
Results

Miscoverage rate

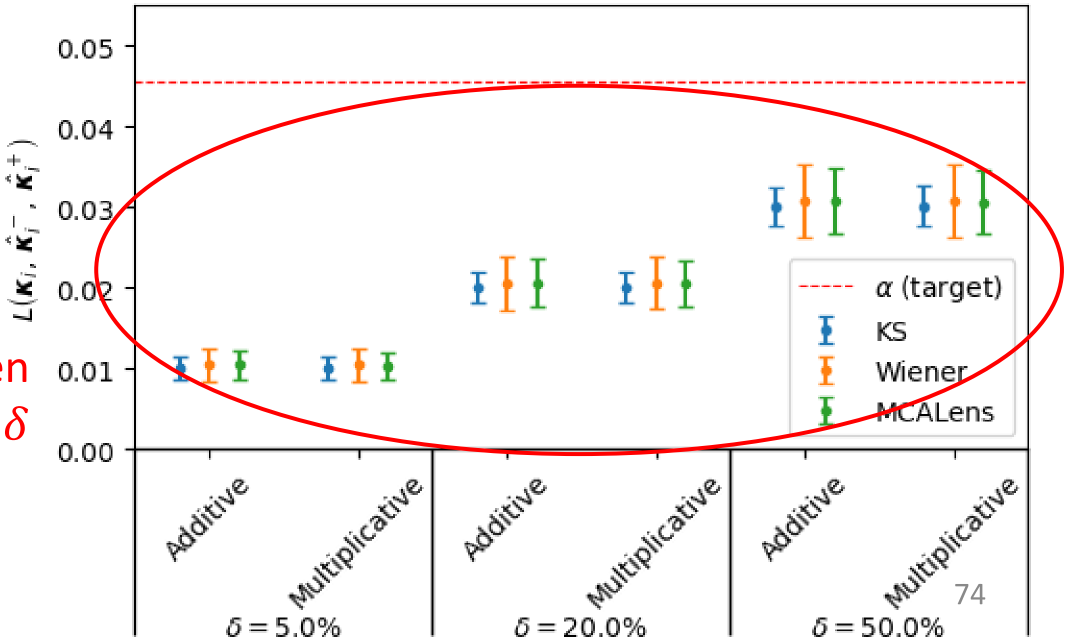
- Calibration set of 100 images from κTNG simulations
- Test set of 125 images from κTNG simulations
- Target: $\alpha \approx 4{,}6\%$ (2σ-confidence)
- CQR: the minimal size depends on the desired confidence level:

  2σ-confidence → $n_{\min} = 21$
  3σ-confidence → $n_{\min} = 370$
  4σ-confidence → $n_{\min} = 15\,787$

CQR



RCPS



RCPS overconservative, even for large values of δ

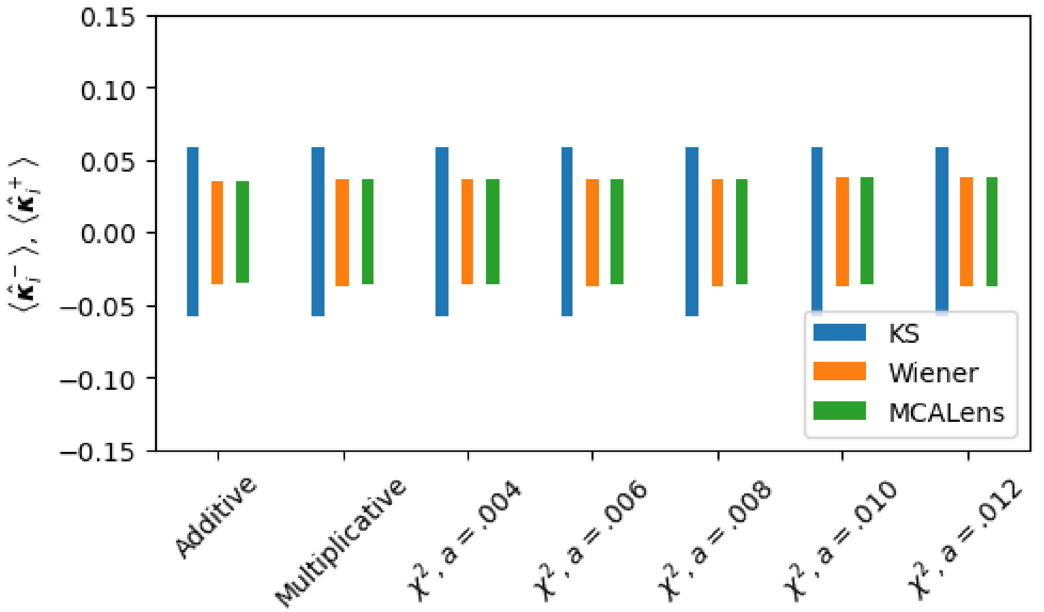
Results

Size of prediction intervals

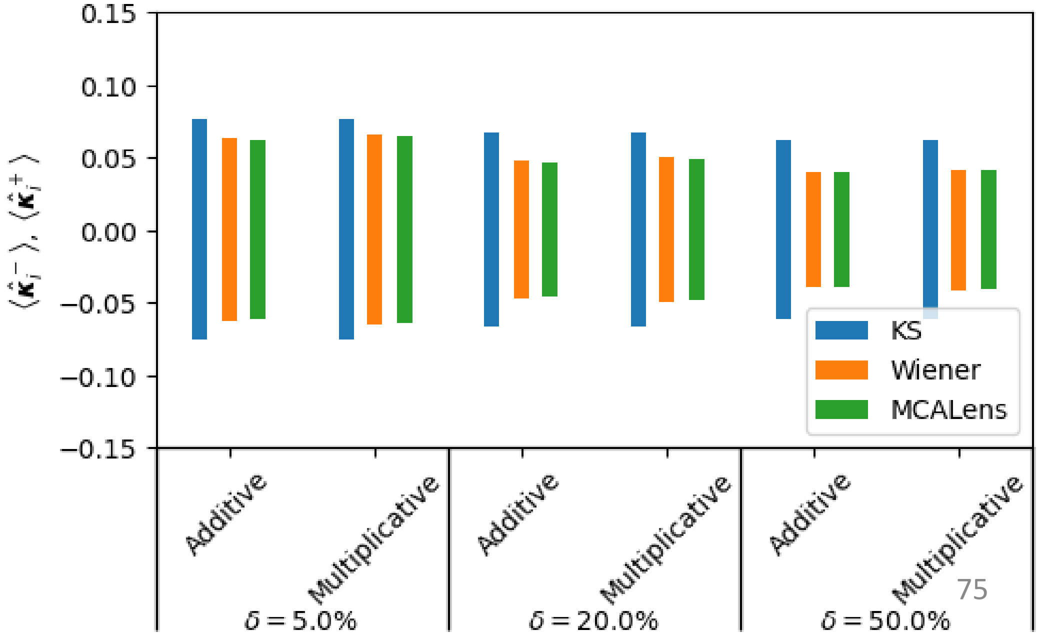
- Calibration set of 100 images from κTNG simulations
- Test set of 125 images from κTNG simulations
- Target: $\alpha \approx 4{,}6\%$ (2σ-confidence)
- CQR: the minimal size depends on the desired confidence level:

  2σ-confidence → $n_{\text{min}} = 21$
  3σ-confidence → $n_{\text{min}} = 370$
  4σ-confidence → $n_{\text{min}} = 15\,787$

CQR



RCPS



Results

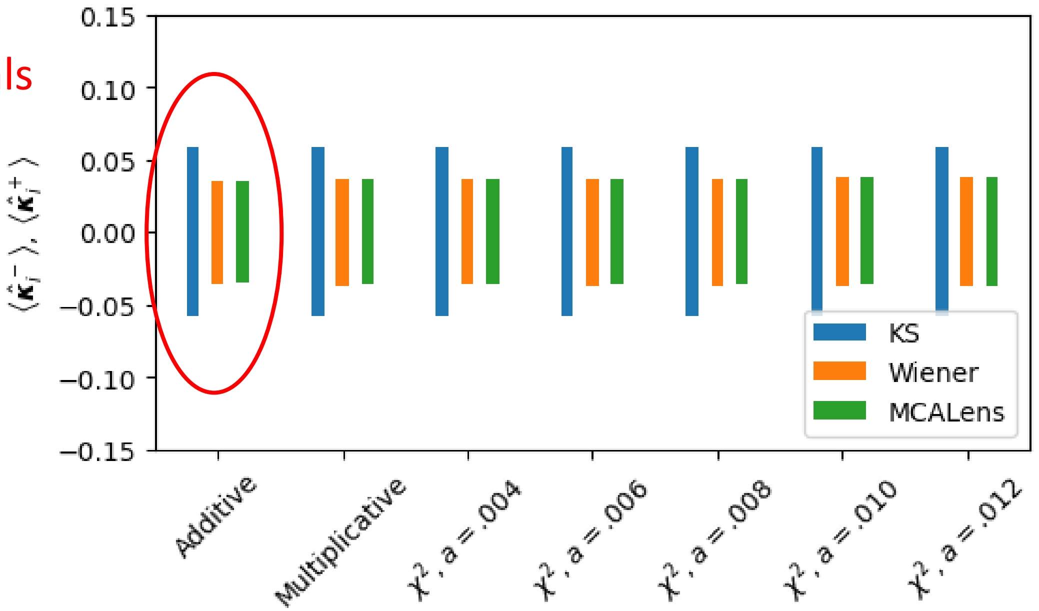
Size of prediction intervals

Smallest confidence intervals

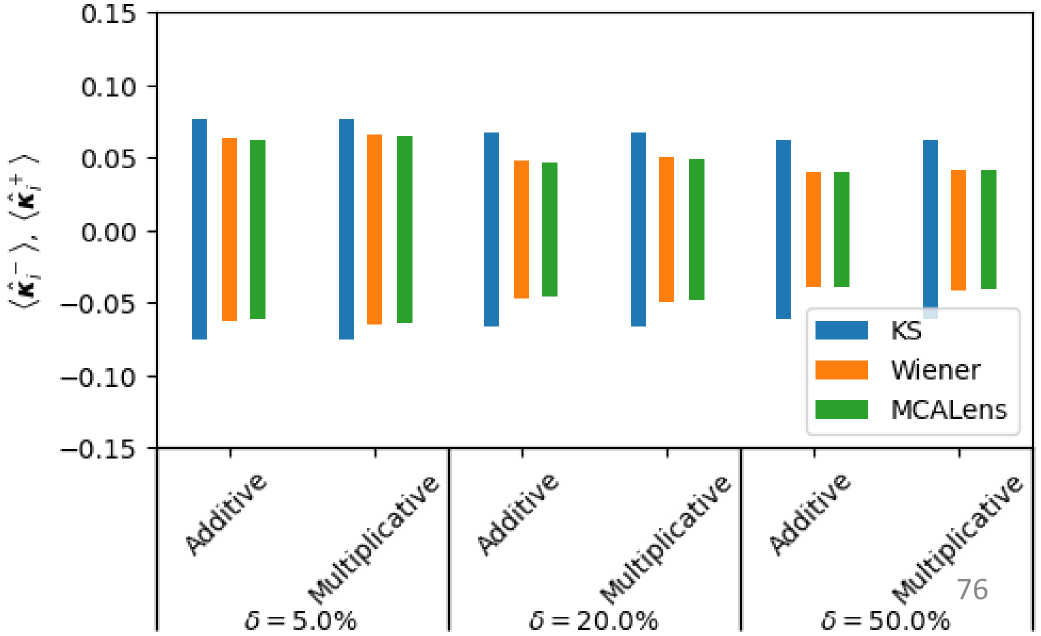
- Calibration set of 100 images from κTNG simulations
- Test set of 125 images from κTNG simulations
- Target: $\alpha \approx 4{,}6\%$ (2σ-confidence)
- CQR: the minimal size depends on the desired confidence level:

  2σ-confidence → $n_{\min} = 21$
  3σ-confidence → $n_{\min} = 370$
  4σ-confidence → $n_{\min} = 15\,787$

CQR

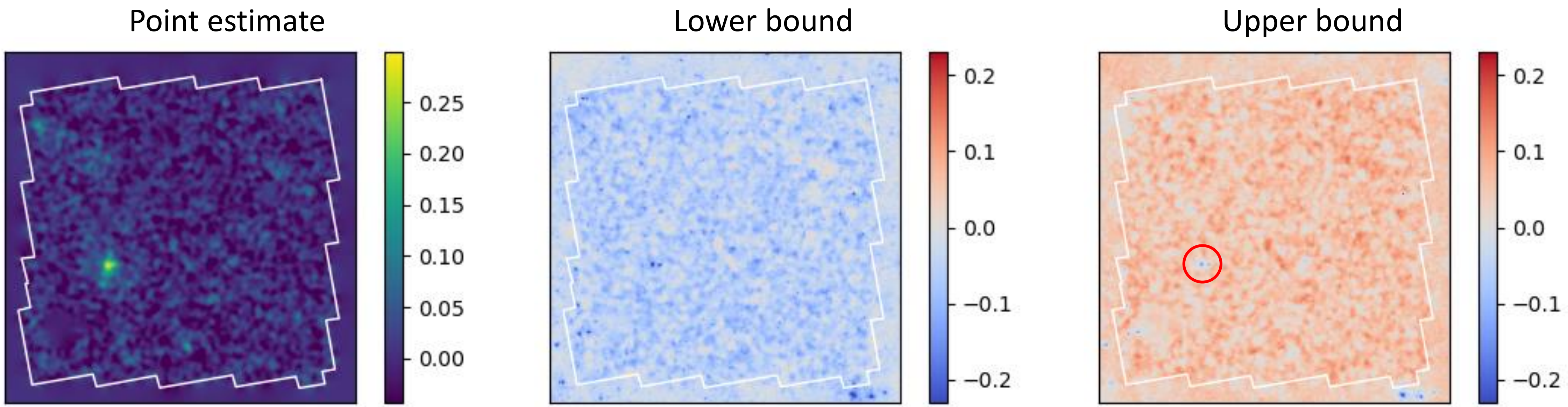


RCPS



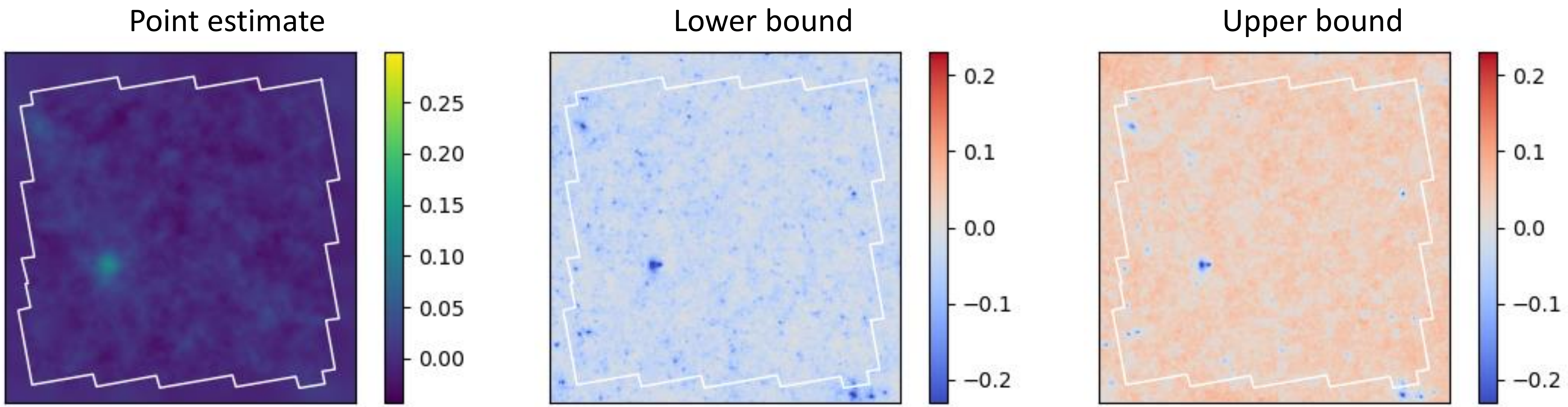
Uncertainty bounds after CQR

Kaiser-Squires

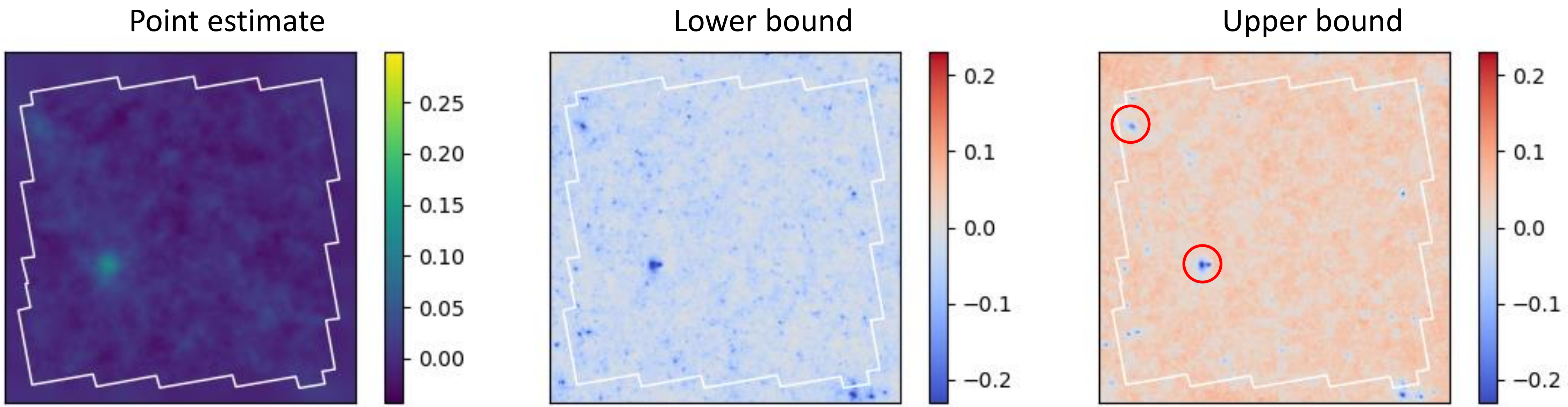


Miscoverage for high-density regions: ground truth larger than upper bound, even after calibration

Uncertainty bounds after CQR
Wiener

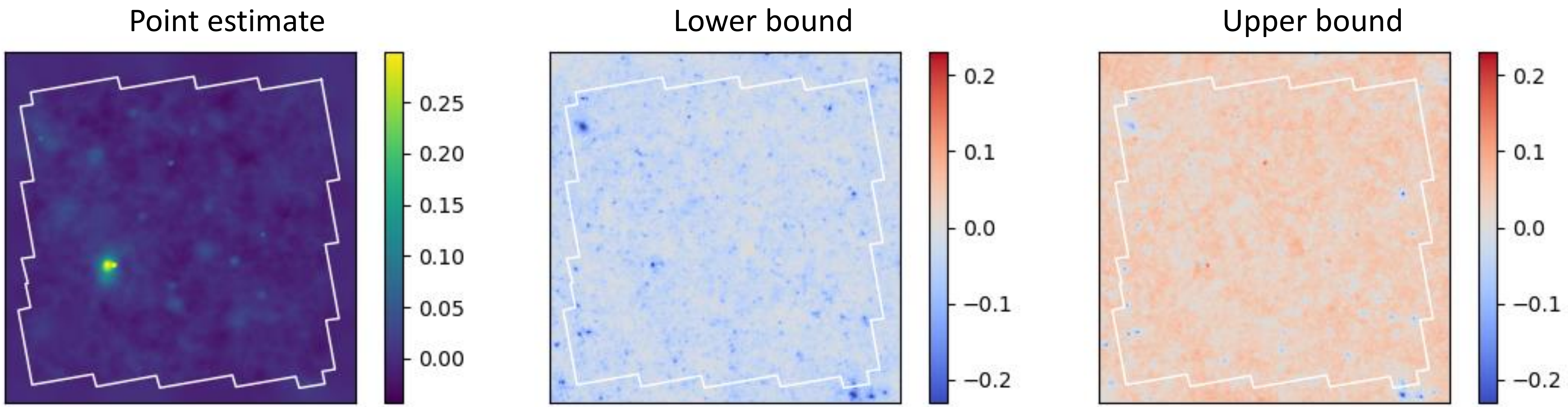


Uncertainty bounds after CQR
Wiener



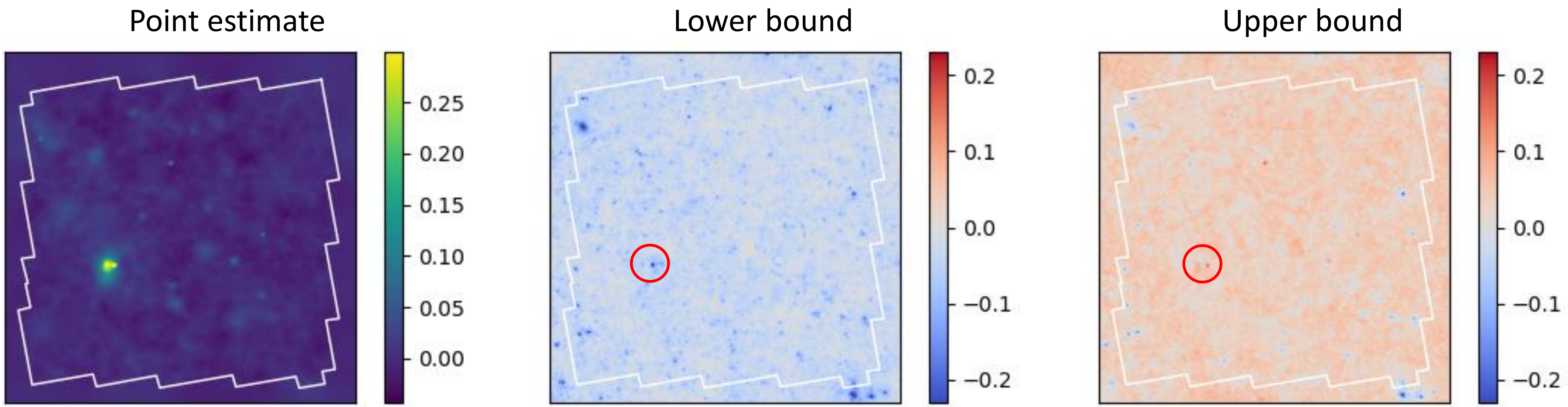
Miscoverage for high-density regions: ground truth larger than upper bound, even after calibration

Uncertainty bounds after CQR
MCALens



Uncertainty bounds after CQR

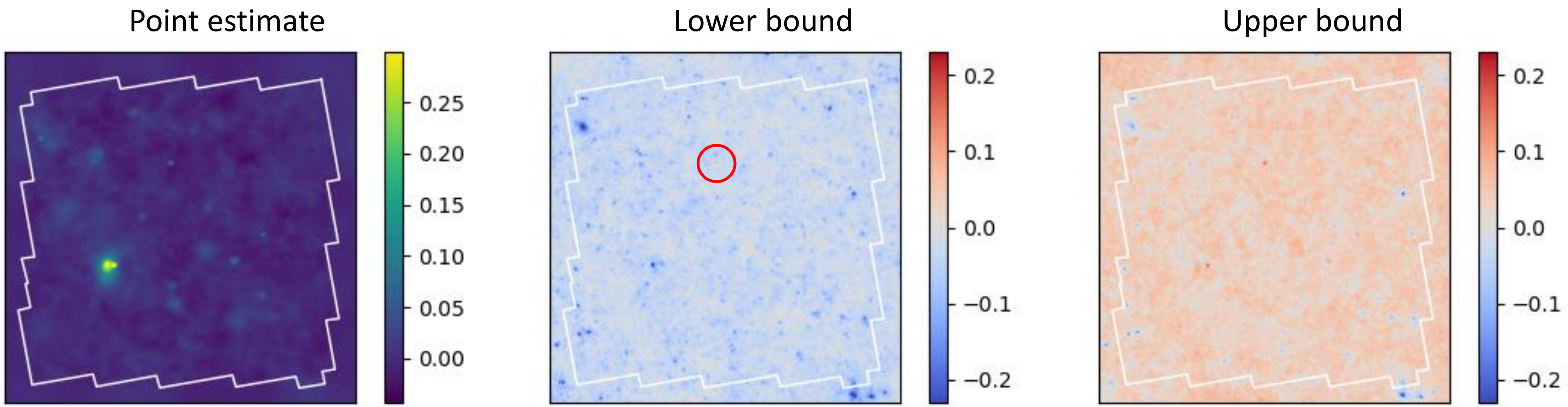
MCALens



Higher uncertainty near high-density regions

Uncertainty bounds after CQR

MCALens

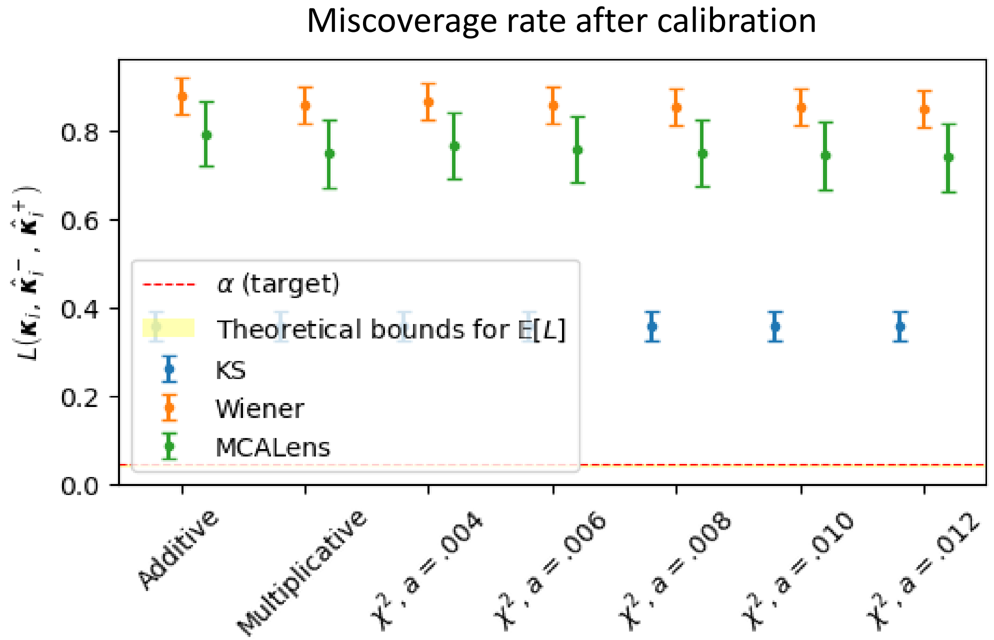


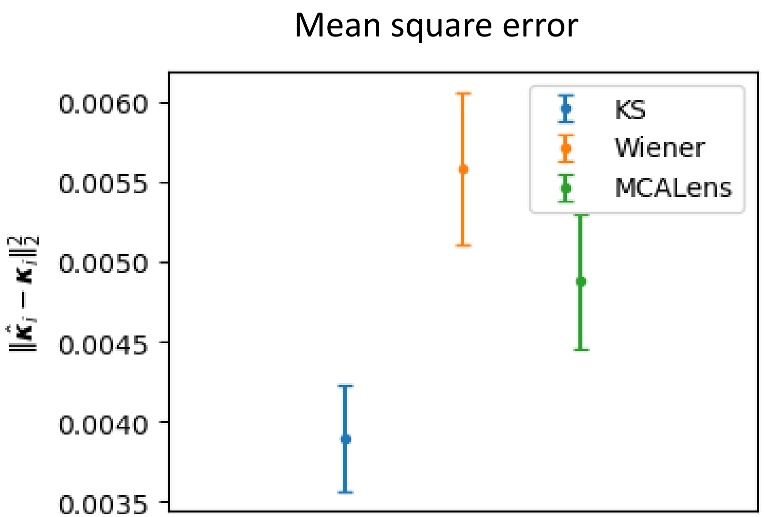
Still hallucinating

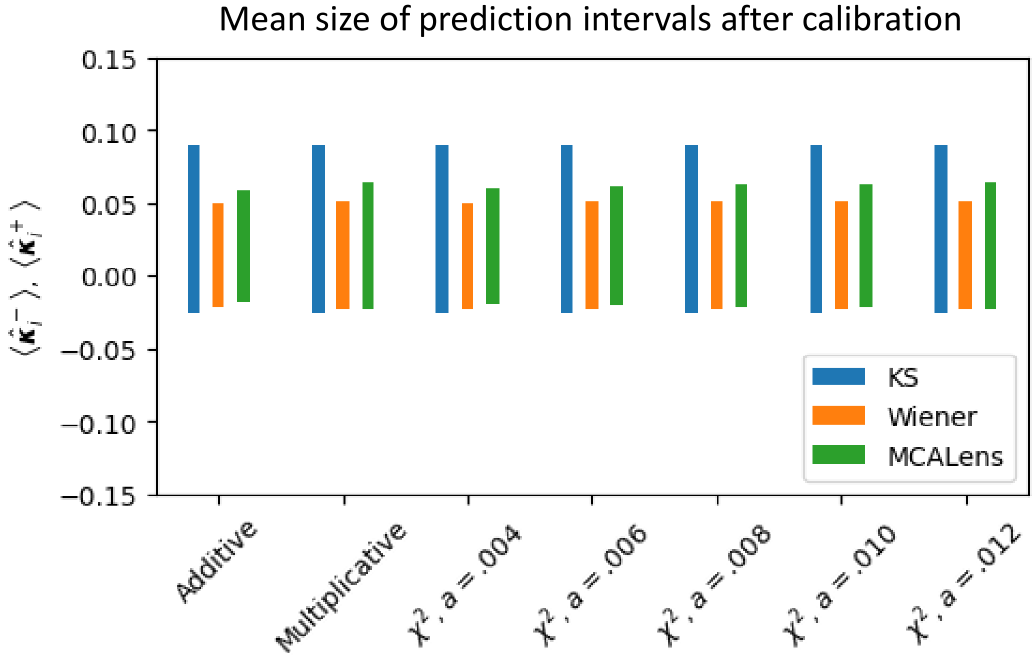
Discussion

Focus on high-density regions

- Theoretical guarantees apply on images as a whole. What happens if we focus on high density regions?
- Results for pixels where the ground truth convergence exceeds 0.1 (2.6% of total pixels).







Discussion

Bayesian vs non-parametric UQ

- Bayesian UQ (e.g., Remy et al. 2023,[1] Liaudat et al. 2023[2]): aims at getting coverage guarantees **assuming $\boldsymbol{\kappa} \sim \mu$ for a certain prior distribution** μ:

$$\mathbb{E}_{\boldsymbol{\kappa}\sim\mu}[L(\boldsymbol{\kappa}, \widehat{\boldsymbol{\kappa}}^{-}, \widehat{\boldsymbol{\kappa}}^{+})] \leq \alpha.$$

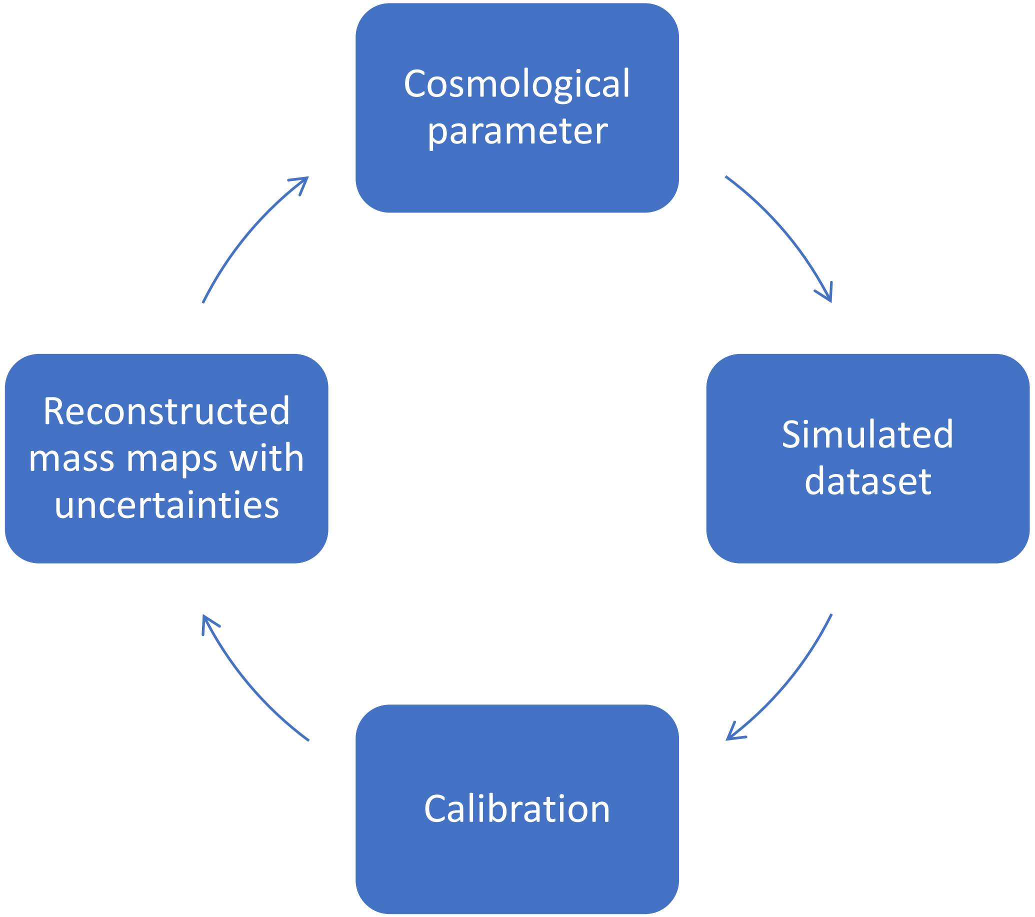
- In contrast, distribution-free approaches do not assume anything on the distribution of $\boldsymbol{\kappa}$. We only assume that the simulated convergence maps $\boldsymbol{\kappa}_i$ from the calibration set follow the same oracle distribution μ^* as $\boldsymbol{\kappa}$ (with exchangeability).
- Bayesian uncertainty bounds could therefore benefit from being calibrated using CQR.

[1] B. Remy et al., "Probabilistic mass-mapping with neural score estimation," A&A, vol. 672, p. A51, Apr. 2023

[2] T. I. Liaudat et al., "Scalable Bayesian uncertainty quantification with data-driven priors for radio interferometric imaging." arXiv, Nov. 2023

Discussion

Inferring cosmological parameters



- Using simulations from a given cosmology may create biases when inferring cosmological parameters.
- Idea: each simulated convergence map uses its own set of cosmological parameters, randomly drawn according to a predefined distribution.
- Re-introduces distribution assumptions, this time over the cosmological parameters.

Conclusion

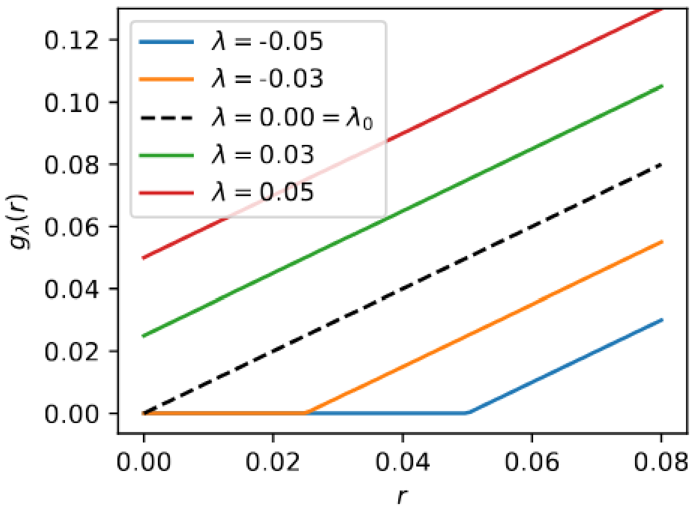
- Distribution-free UQ for mass mapping: provides coverage guarantees with a limited number of calibration examples.
- Works for any mass mapping method, including deep learning → can be adapted to more advanced approaches.
- Does not prevent hallucinations, nor undercoverage near high-density regions. Possible improvement using conformal prediction masks?[1]
- Does not assume any prior distribution on the convergence maps, but a specific attention is required for the calibration set, especially regarding the choice of cosmology from which it is simulated.
- Possible extension: exploit correlation between pixels to get tighter confidence regions.[2]

[1] G. Kutiel, R. Cohen, M. Elad, D. Freedman, and E. Rivlin, "Conformal Prediction Masks: Visualizing Uncertainty in Medical Imaging," presented at the ICLR 2023 Workshop on Trustworthy Machine Learning for Healthcare, Apr. 2023

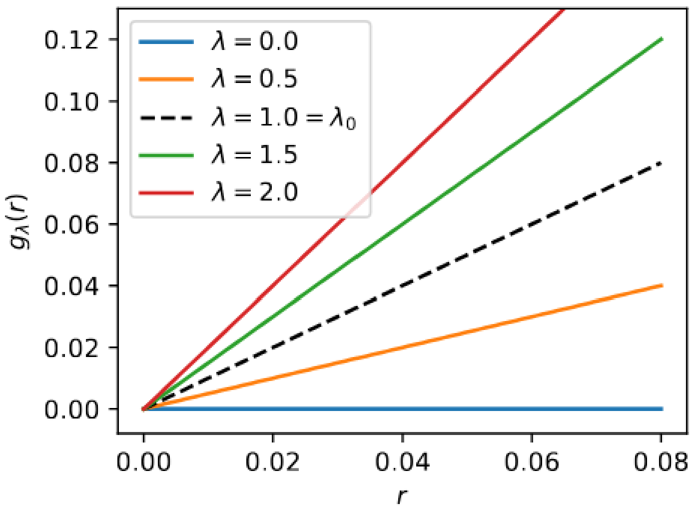
[2] O. Belhasin, Y. Romano, D. Freedman, E. Rivlin, and M. Elad, "Principal Uncertainty Quantification with Spatial Correlation for Image Restoration Problems." arXiv, May 17, 2023.

Ευχαριστώ!

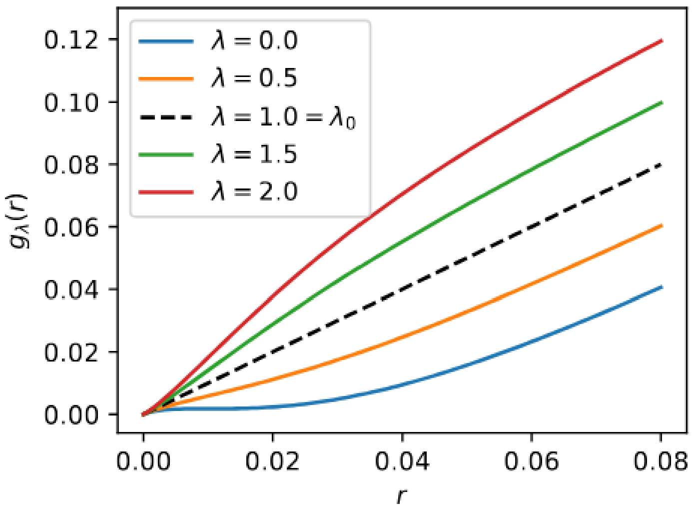
Examples of calibration functions



(a) $g_\lambda : r \mapsto \max(r + \lambda, 0)$



(b) $g_\lambda : r \mapsto \lambda r$



(c) $g_\lambda : r \mapsto r + bF_{\chi^2(k)}(r/a)(\lambda - 1)$