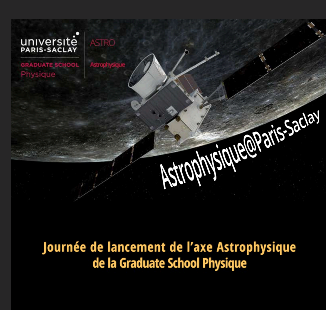
Rethinking data-driven point spread function modelling with a differentiable optical model

Tobias I. Liaudat

3rd year PhD student

tobias-liaudat.github.io

In collaboration with
Jean-Luc Starck, Martin Kilbinger and Pierre-Antoine Frugier



Journée de lancement de l'axe Astrophysique de la Graduate School Physique

16th November 2021

CosmoStat

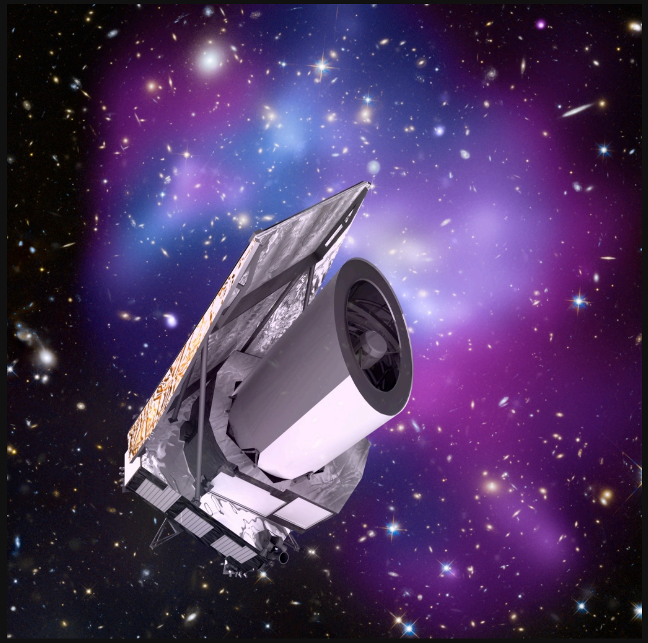
université PARIS-SACLAY

cea

Why do we care about the PSF?

Euclid Space mission

Launch: 2023



Optimised for **weak lensing**, a main science goal.

The PSF model is a crucial part of the mission.

Main source of systematic error.

Going to observe a vast part of the sky.

A lot of data to process → Processing time is an issue.

More complex diffraction-limited PSFs.

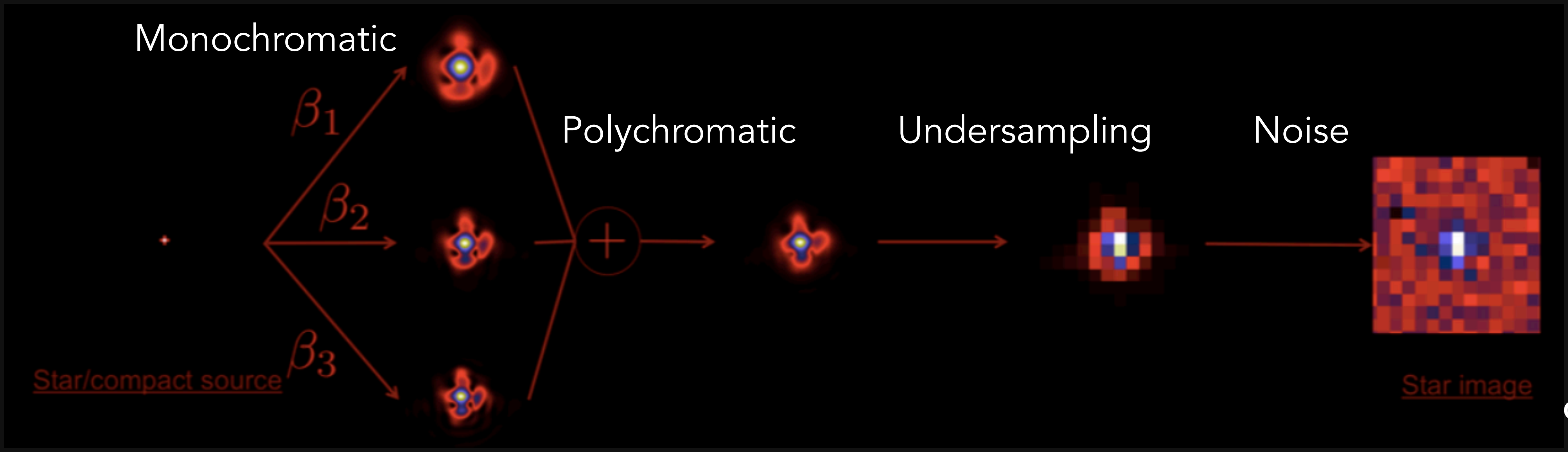
Observations are under-sampled.

Observations are in a single broad band [550,900]nm.

Necessary to model spectral variations.

Very low error requirements.

Several new challenges arising.



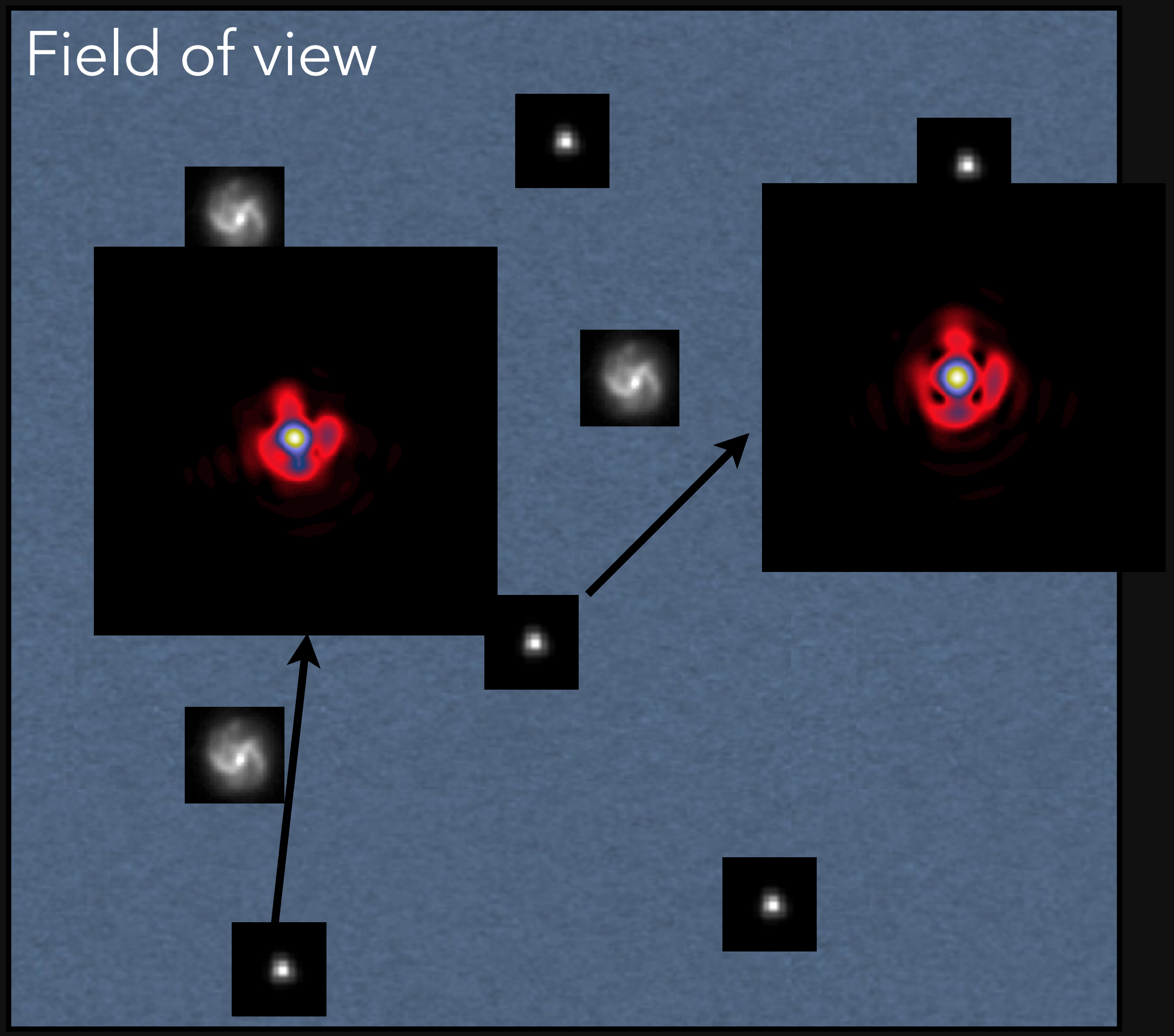
dat

PSF modelling

Need of a PSF model to correct the telescope's observations.

We can consider some star observations as samples of the PSF field.

Use them to build a model to infer the PSF at galaxy positions.



Inputs

PSFs at star positions

Train/fit →

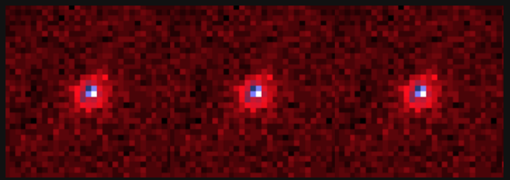
PSF model

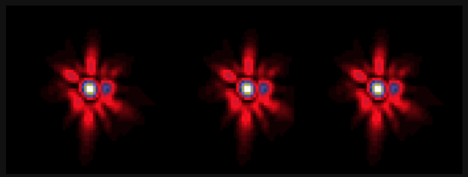
Infer →

Outputs

PSFs at galaxy positions

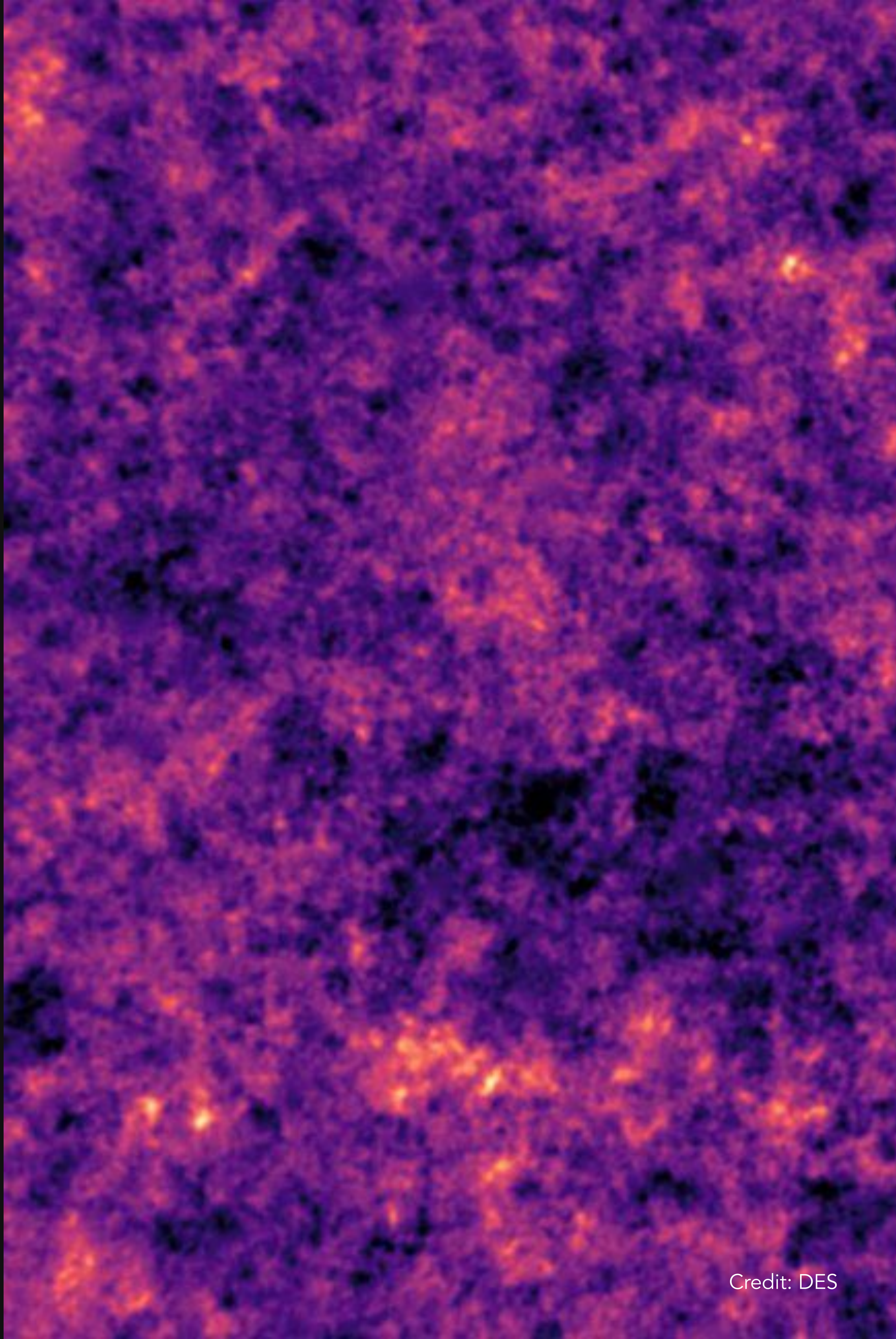
Stellar SEDs +





How do we model the PSF field?

euclid



Credit: DES

Modelling the PSF field

What variations we need to model?

Spatial $PSF \sim f(x, y)$ Spectral $PSF \sim f(\lambda)$ Temporal $PSF \sim f(t)$

Parametric model

Completely model the telescope's optical system.

Express high variability with a small number of parameters.
Degenerate optimization problem.

The model is built in the wavefront-error space (WFE).

Need special calibration data.

Only done for the Hubble Space Telescope [Tiny Tim model].

Very sensible to any mismatch between the model and the observations.

Data-driven approaches led to better results [Hoffman & Anderson, 2017]

Modelling the PSF field

Data-driven model

What variations we need to model?

Spatial $PSF \sim f(x, y)$ Spectral $PSF \sim f(\lambda)$ Temporal $PSF \sim f(t)$

↓

Use observations at a given time
Limits the number of available stars

Tobias Liaudat

Modelling the PSF field

Data-driven model

What variations we need to model?

Spatial $PSF \sim f(x, y)$

↓

Use a constrained matrix factorisation scheme.

Different constraints and optimisation procedures define different models.

Model built in the pixel space → Linear combination of features.

Usually done independently on each CCD.

$$H = \mathcal{F}\{S\, A\}$$

- H ↑ Observations
- $\mathcal{F}$ ↑ Degradation operator
- S ↑ PSF features
- A ↖ Weights (spatial variations)

State-of-the-art examples:
- *MCCD*, Liaudat et al. A&A 646:A27
- *RCA*, Schmitz et al. A&A 636:A78 (Euclid consortium paper)
- *PSFEx*, Bertin. ASPC, 442, 435

Most of the used data-driven PSF models are built in a similar fashion.

Tobias Liaudat

Modelling the PSF field

Data-driven model

What variations we need to model?

Spectral $PSF \sim f(\lambda)$

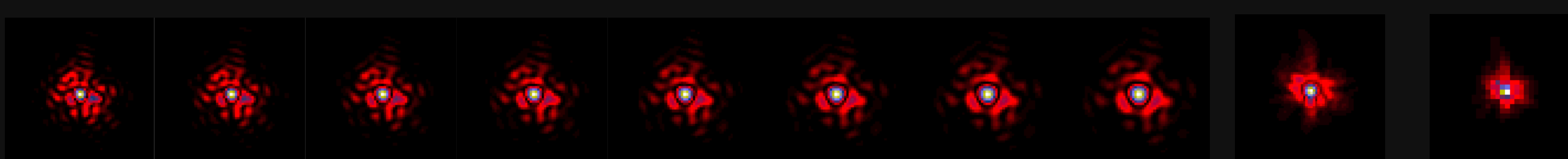
Very hard to model spectral variations with a data-driven approach.

Efforts in this directions included using Optimal Transport to interpolate between the two extreme wavelength PSFs. [Schmitz, PhD thesis, 2019]

Assumes spectral variations are smooth, but in real instruments this is not the case.

Better approach: **include the physics of the problem while remaining data-driven.**

→ Optical system



550nm → λ → 900nm

Polychromatic Super-resolved

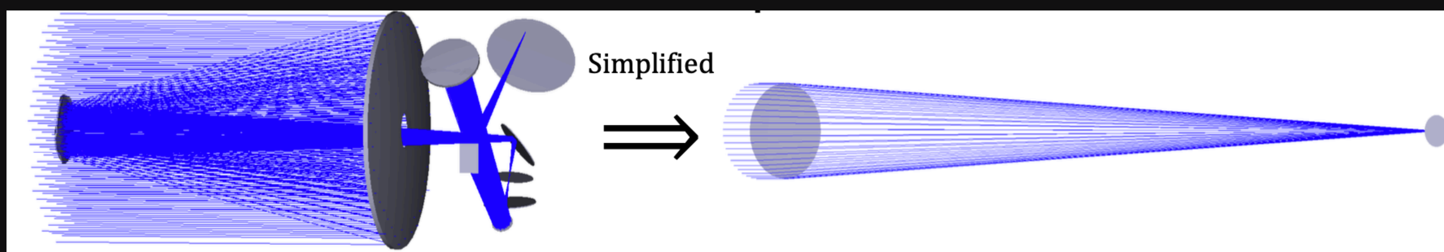
Polychromatic under-sampled

Tobias Liaudat

Changing the PSF model representation space

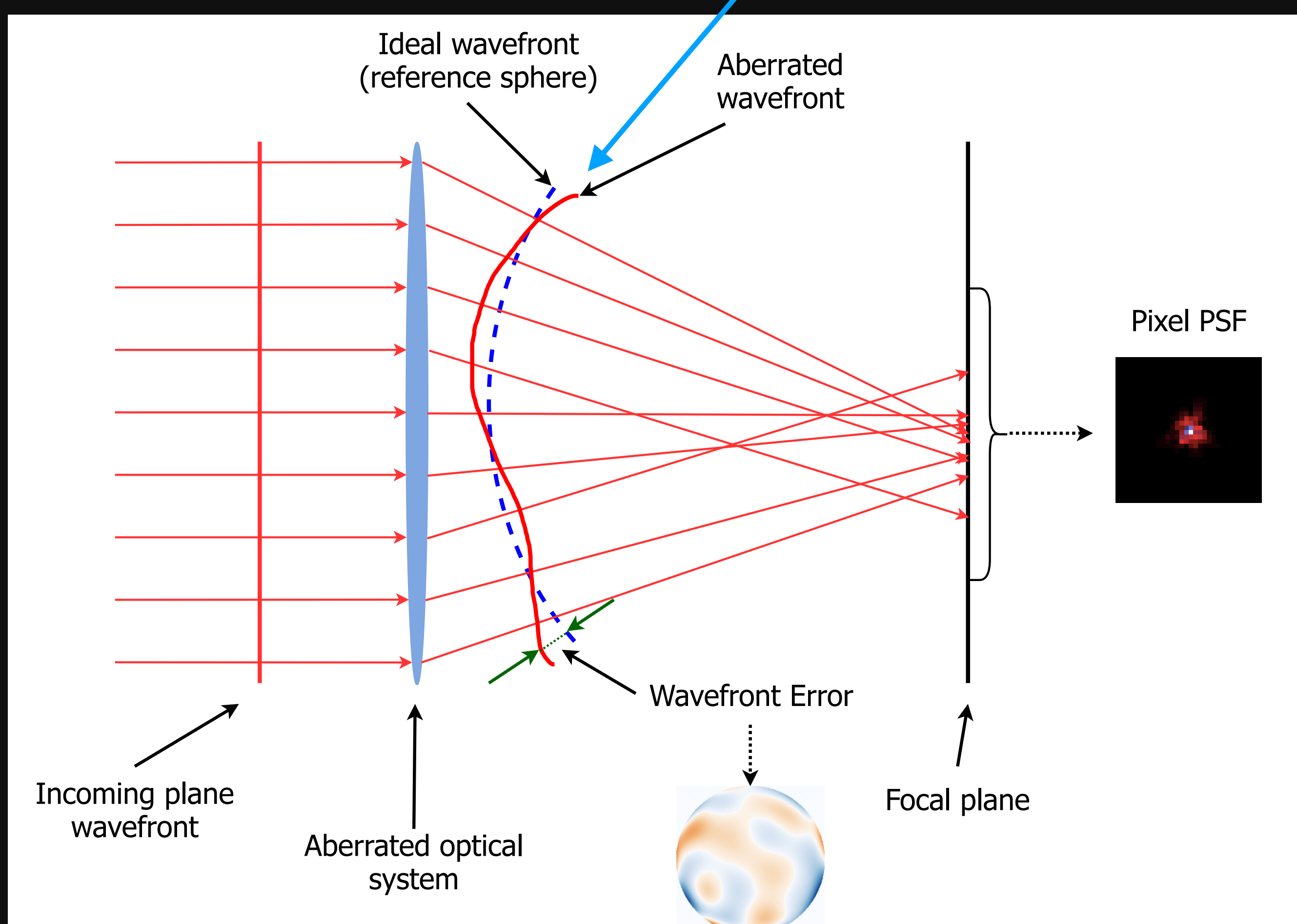
Euclid optic model

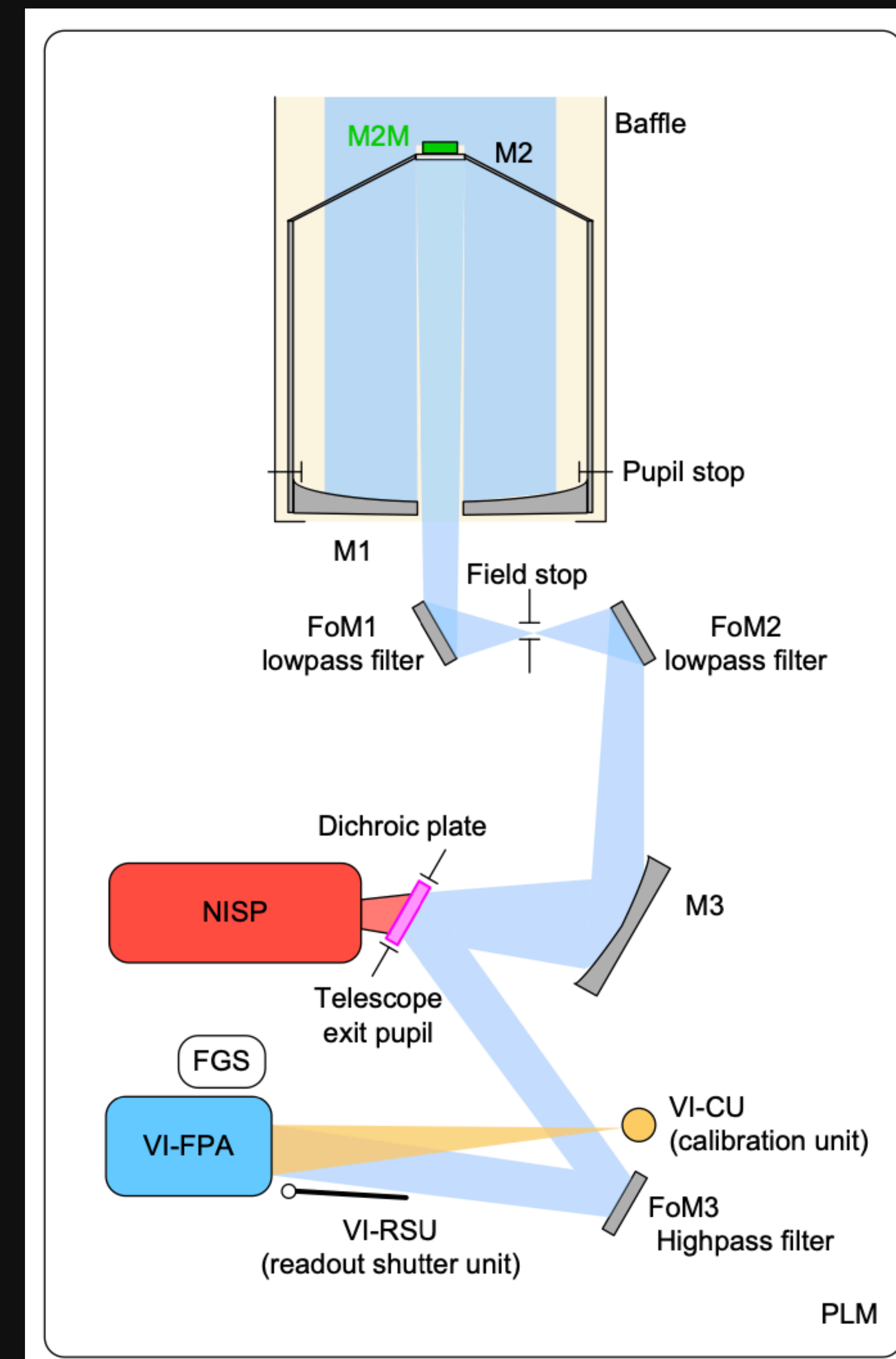
Simplified optic model



Credit: P.A. Frugier

PSF model in WFE space





Credit: G. Racca et al, 2016

Tobias Liaudat

Modelling the PSF in the WFE space

WFE = Parametric part + Non-Parametric part

The non-parametric (*data-driven*) part helps to correct the mismatch between the selected parametric model and the observations.

Easier to model chromatic variations and still be data-driven.

Harder to constraint from star observations, it's a degenerate optimization problem.

We aim to avoid the use of special calibration data.

Build a forward model based on the telescope's optics, WFE → pixels

Includes diffraction phenomena (Fraunhofer approx.), obscuration, downsampling, etc..

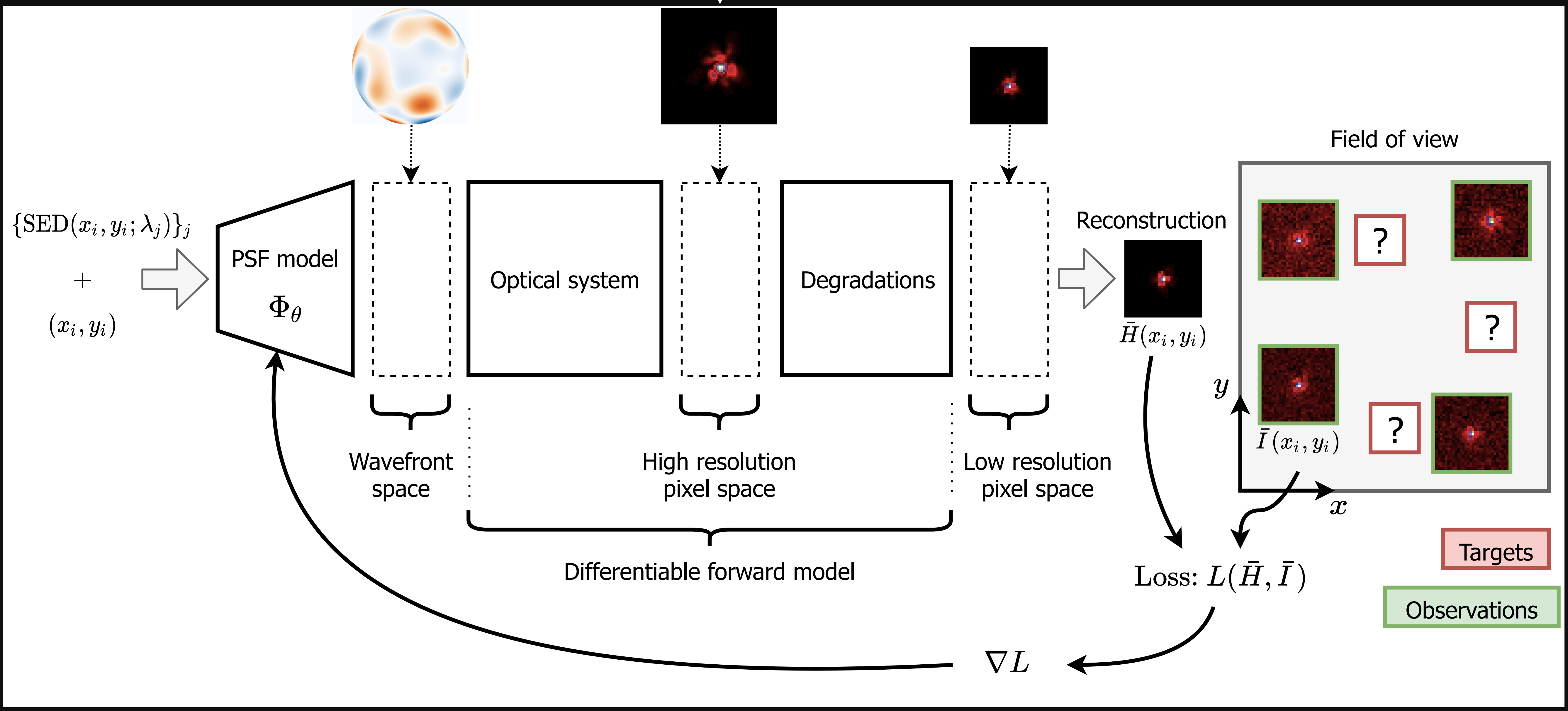
End-to-end differentiable!

Based on an automatic differentiation framework → TensorFlow.

Fast computations on GPU.

Overview of the proposed approach

Super-resolved
monochromatic
PSF reconstruction



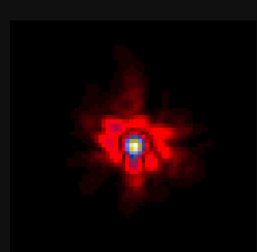
Tobias Liaudat

Modelling the PSF in the WFE space

Differentiable optical forward model

$$\hat{H}(x, y; \lambda) \propto \mathscr{D}\left\{ P \odot \exp\left[2\pi \mathrm{i} \left(\frac{\Phi^{Z}(x, y)}{\lambda} + \Phi^{DD}(x, y; \lambda) \right) \right] \right\}$$

Super-resolved monochromatic PSF reconstruction



Diffraction operator

Telescope's obscurations

WFE parametric part

WFE non-parametric part

Wavefront PSF model
$\Phi_{\theta}(x, y; \lambda)$

$\Phi^{Z}\,,\Phi^{DD}\,,P \in \mathbb{R}^{n \times n}$

$\hat{H}(x, y; \lambda) \in \mathbb{R}^{p \times p}$

$x, y, \lambda \in \mathbb{R}$

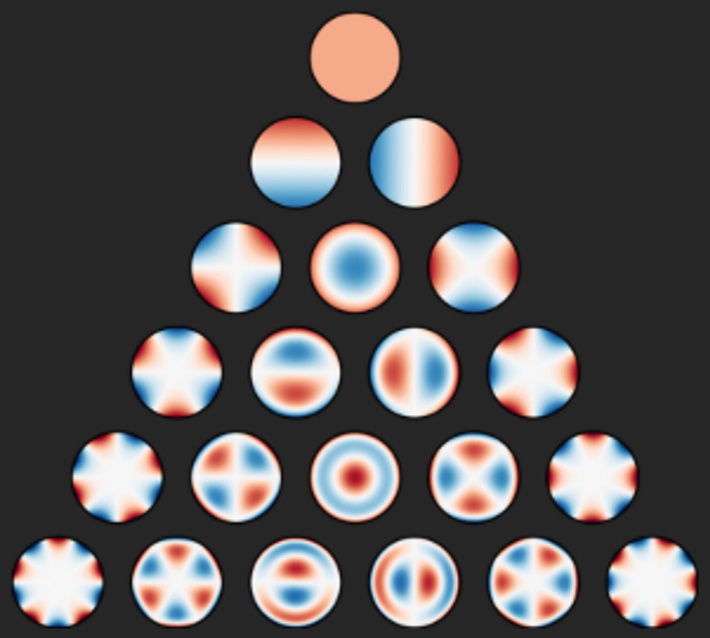
Tobias Liaudat

WFE parametric part

$$\Phi^Z(x, y) = \sum_{j=1}^{N^Z} f_j^Z(x, y)\, S_j^Z$$

e.g. $f_j^Z(x, y) = c_0^j + c_1^j\, x + c_2^j\, y$

Zernike Polynomials S_j^Z



Based on Zernike polynomials up to mode N^Z.

Orthogonal in the unit disk.

Widely used in optics.

$\Phi^Z \in \mathbb{R}^{n \times n}$

$S_j^Z \in \mathbb{R}^{n \times n}$

$x, y \in \mathbb{R}$

FoV variations based on FoV position polynomials of Zernike coefficients.

Chromatic variations follow the $1/\lambda$ dependence of diffraction.

Small number of parameters to represent all the variability.

WFE non-parametric part

Based on a matrix factorisation scheme

Diffraction-based wavelength dependence

Number of PSF features

$\Phi^{DD} \in \mathbb{R}^{n \times n}$

$S_l^{DD} \in \mathbb{R}^{n \times n}$

$$\Phi^{DD}(x, y; \lambda) = \frac{1}{\lambda} \sum_{l=1}^{N^{DD}} f_l^{DD}(x, y) \, S_l^{DD}$$

PSF features weights

Data-driven WFE PSF features

PSF wavefront features are completely data-driven.

Different choices for the PSF feature weights will define different flavours of the model.

We present one using polynomials of FoV positions.

We could use the graph constraint for localised variations from the RCA method.

For the moment, we use a diffraction-based wavelength dependence.

We could easily add more sophisticated chromatic functions (e.g. refractive elements).

Tobias Liaudat

Optimisation and inference

Optimisation

Using weighted Mean Squared Error loss function over the star observations.

Add a regulariser of the model's parameters R_θ depending on the model's flavour.

Use a noise std dev estimator for the weights.

$$L(\theta) \propto \sum_i \frac{1}{\hat{\sigma}_i} \| \bar{I}_i - \bar{H}_i(\Phi_\theta) \|_F^2 + R(\Phi_\theta)$$

Optimising with Rectified Adam (advanced stochastic gradient descend method).

Allowed by the automatic differentiation framework.

Inference - PSF recovery

Straightforward and fast

Evaluate $\Phi_\theta(x, y)$ on the new position and propagate through the forward model.

Tobias Liaudat

Numerical experiments

Credit: NASA/ESA

Experiment set-up

Simulating one FoV with 2000 stars for training and 400 stars for testing.

Simulations at Euclid resolution (under-sampled), images 32 x 32.

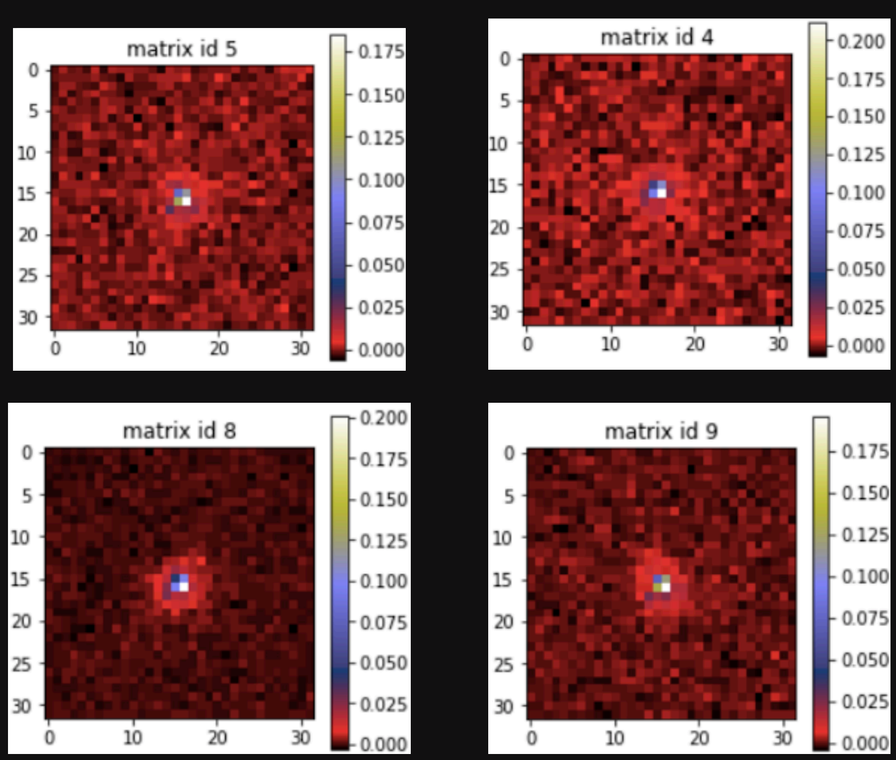
Using real stellar SEDs for star observations.

Added Gaussian noise to achieve flat distribution of random SNRs.

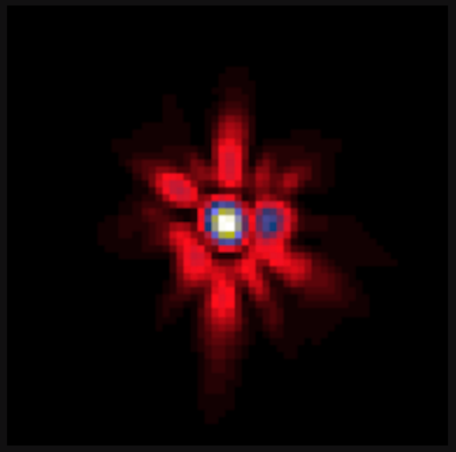
PSF field simulated using parametric part:

2D position polynomials of degree 2 and 45 Zernike modes.

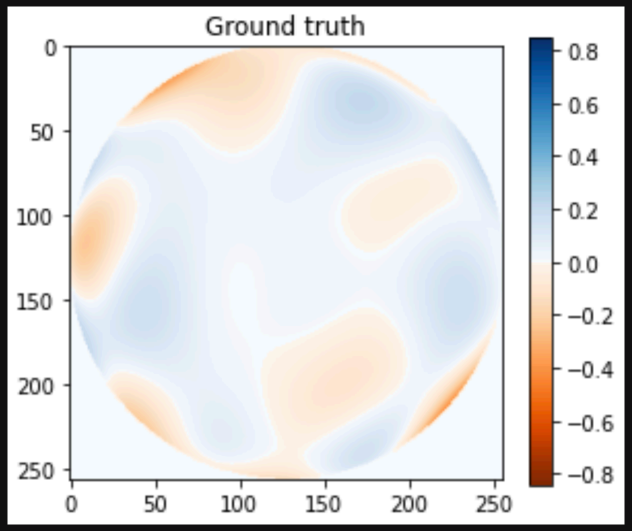
Observations



Ground truth PSF at 3x observation resolution



Ground truth $\Phi(x_i, y_i)$



Tobias Liaudat

Model comparison

- Slightly imperfect parametric PSF model

Almost the same model as the simulation.

Using 40 Zernike modes.

- Imperfect parametric PSF model

Only 15 Zernike modes used.

(Instead of 45)

- Resolved Component Analysis (RCA)

State-of-the-art, designed for Euclid

RCA, Schmitz et al. A&A 636:A78

- Proposed data-driven WFE PSF model

Badly specified parametric part

Only 15 Zernike modes used.

Non-parametric part with 21 PSF features.

Position polynomial of degree 5.

- PSFEx

Widely used state-of-the-art model

PSFEx, Bertin. ASPC, 442, 435

All of them are given the true stellar SEDs as input.

No calibration data used.

Pixel reconstruction results

Reconstruction of test stars at x1 and x3 observation resolution.

Super-Resolution (SR) task

Performance gap in SR between models with forward model and previous SOTA.

	RMSE [$\times 10^{-5}$] (relative)	
PSF model	Resolution x1	Resolution x3
i) Zernike 15	72.3 (10.0%)	18.3 (12.4%)
ii) Zernike 40	22.2 (3.0%)	5.75 (3.9%)
iii) Zernike 15 + DD	**8.34** (**1.1**%)	**4.47** (**3.0**%)
iv) PSFEx	69.2 (9.5%)	66.3 (43.0%)
v) RCA	39.6 (5.4%)	85.3 (55.5%)

Even if the parametric part lacks complexity the Z15+DD is the best performing.

Importance of the **data-driven** part in the model.

Breakthrough in performance w.r.t. current SOTA models.

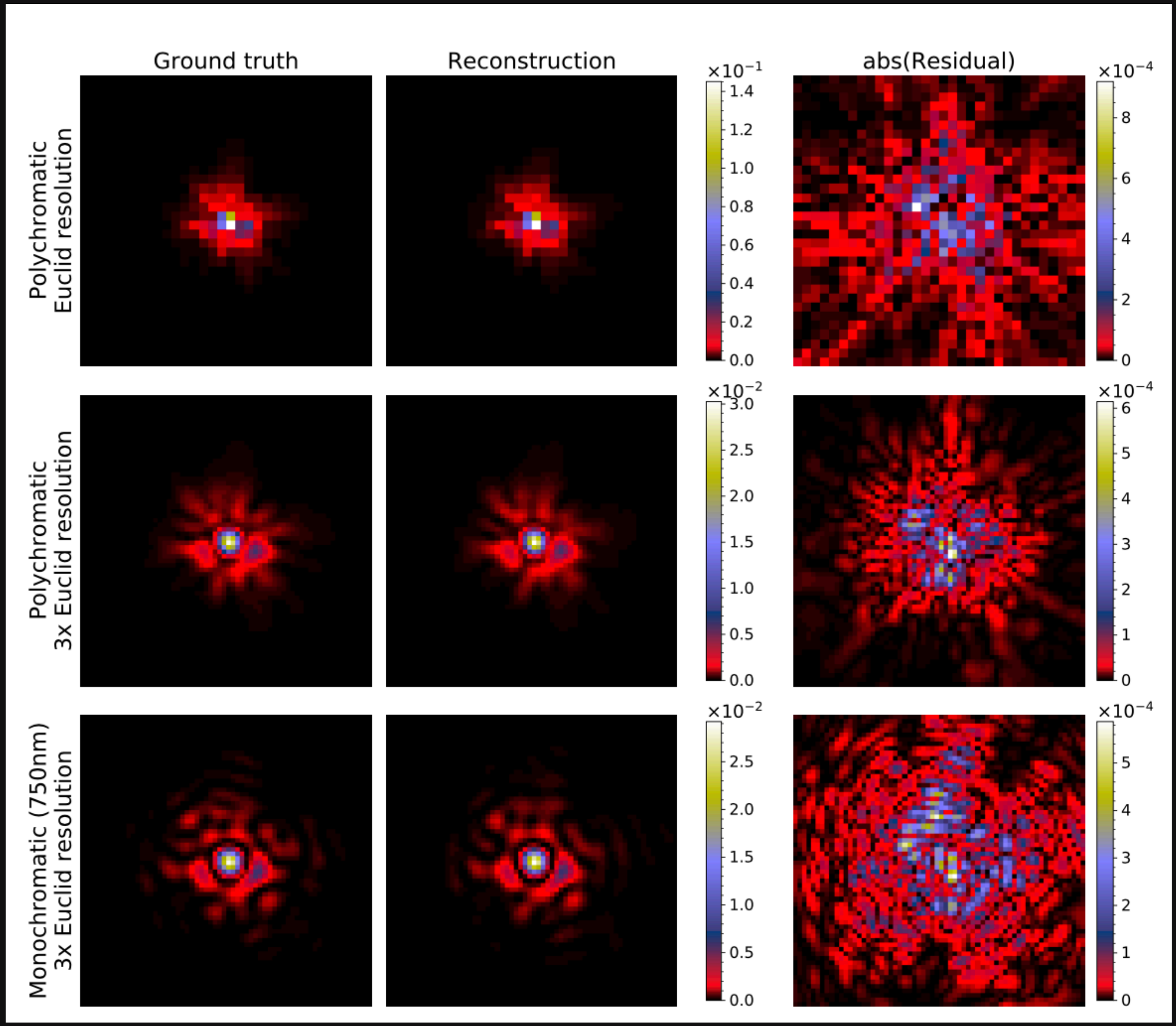
Tobias Liaudat

Pixel reconstruction results

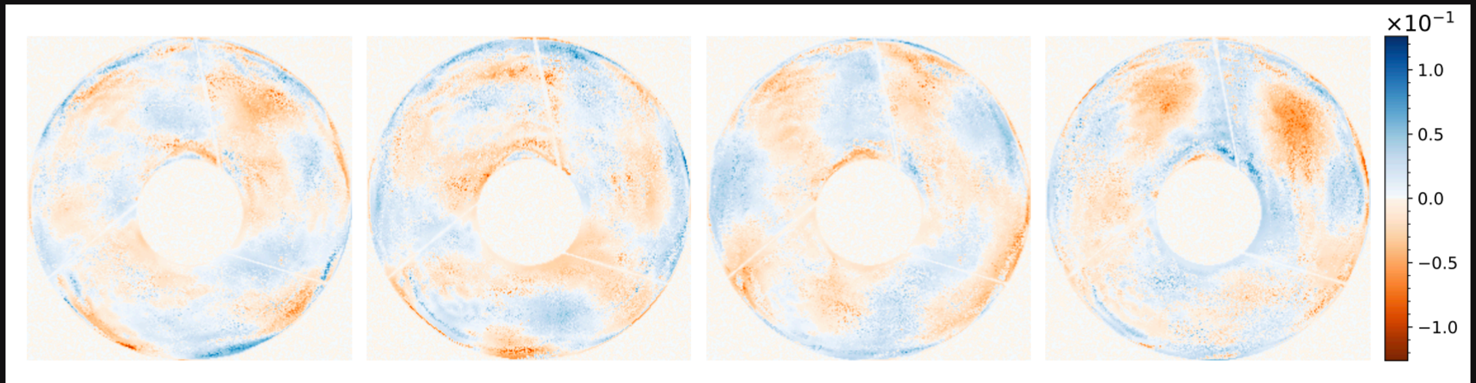
Zernike15+DD PSF model reconstruction examples:

Very good pixel reconstruction.

@750nm



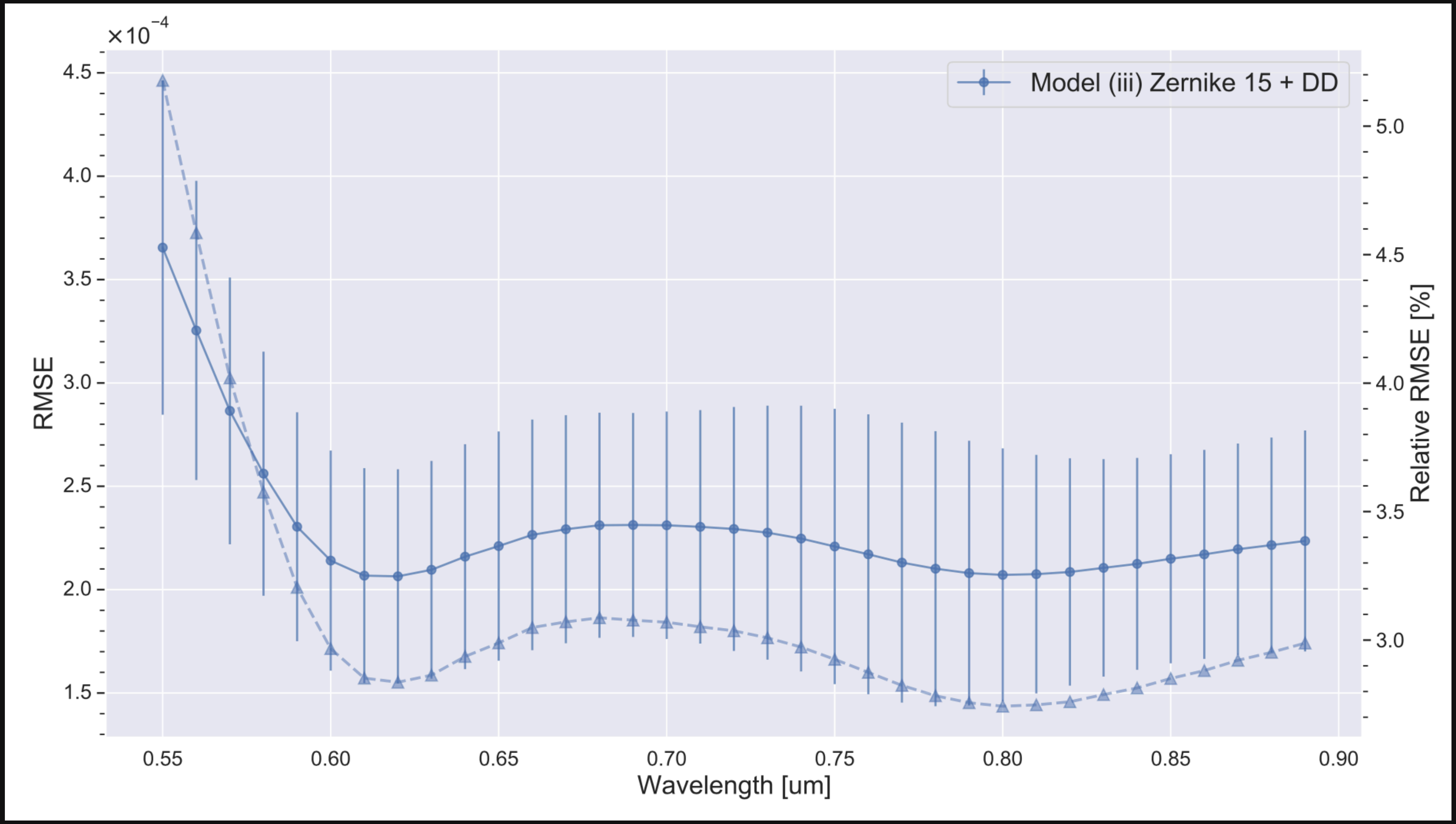
Examples of learned WFE data-driven features:



Tobias Liaudat

Monochromatic pixel reconstruction

Test star reconstruction as a function of wavelength at 3x observation resolution.



First data-driven model to effectively model chromatic variations!

The estimated model is not degenerating w.r.t. wavelength.

Tobias Liaudat

Conclusions

First data-driven model built in the WFE space up to my knowledge.

- Able to model spectral variations!

Obtained a very low pixel error even with an incomplete parametric part.

- The non-parametric part is effective in capturing the mismatches of the parametric part.
- Better results than the slightly imperfect parametric model.

Does not require special calibration data.

Good results on a realistic dataset.

Built over the Tensorflow framework → end-to-end differentiable.

- Allows for fast GPU calculations.
- Could be easily used to introduce physics into Neural Networks.

Promising approach for the Euclid mission!

Thank you!