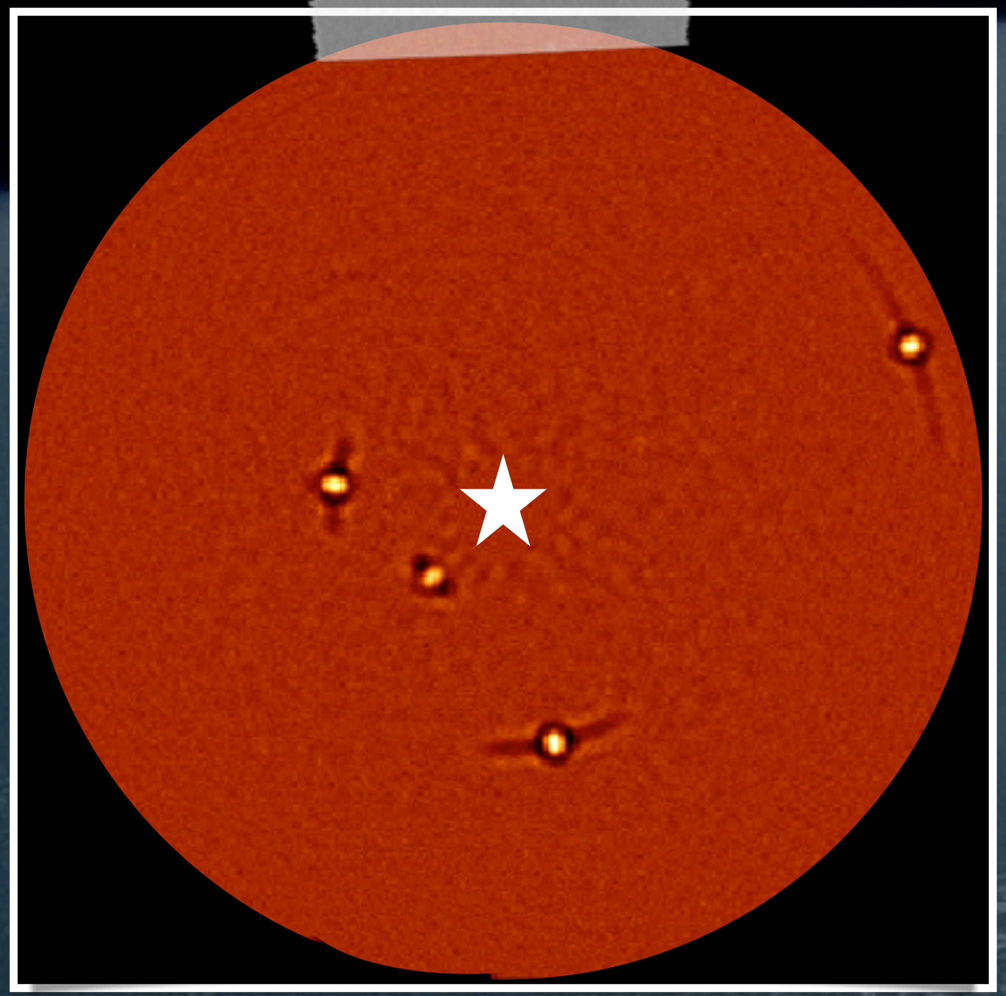
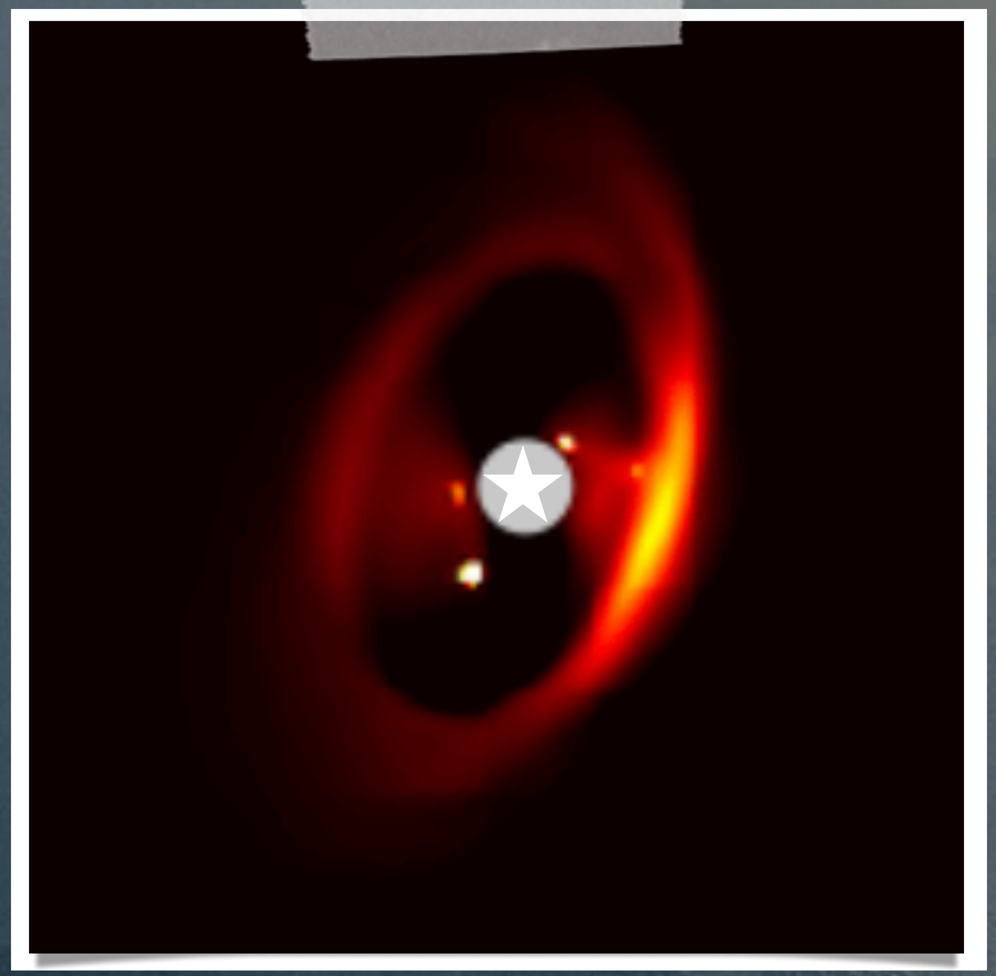
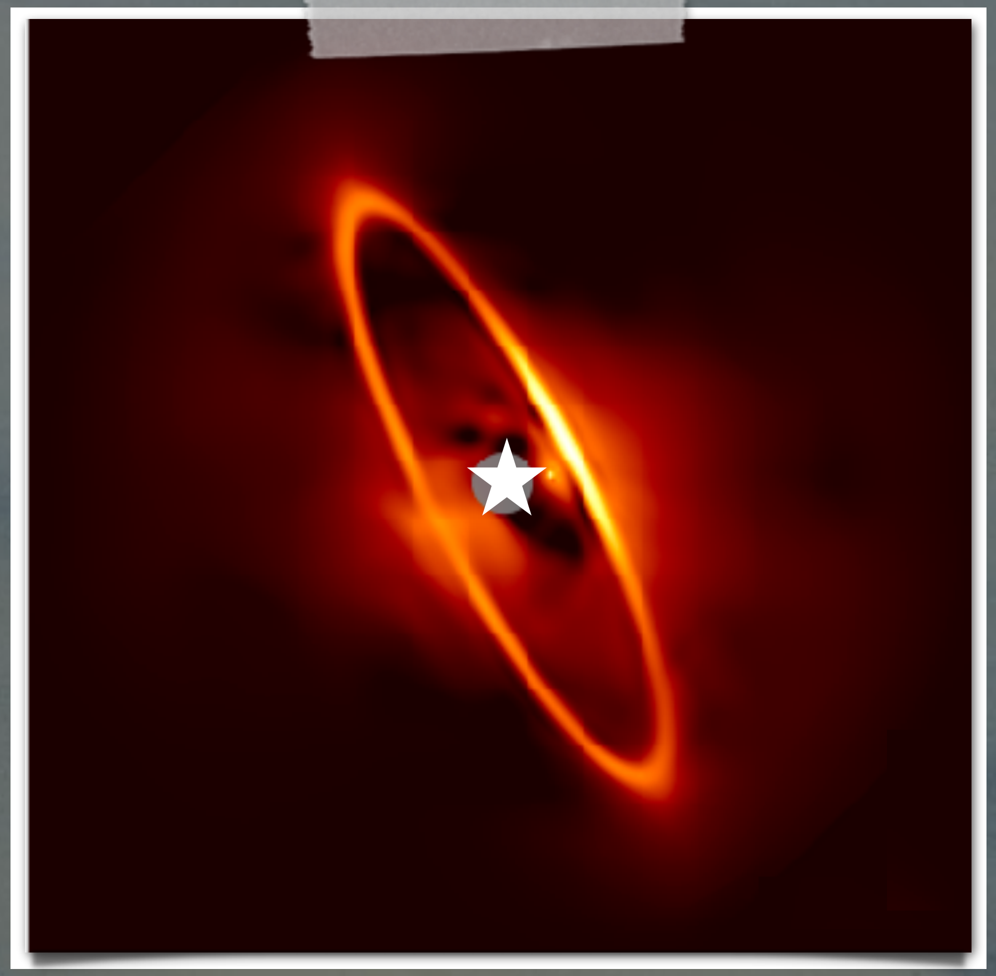
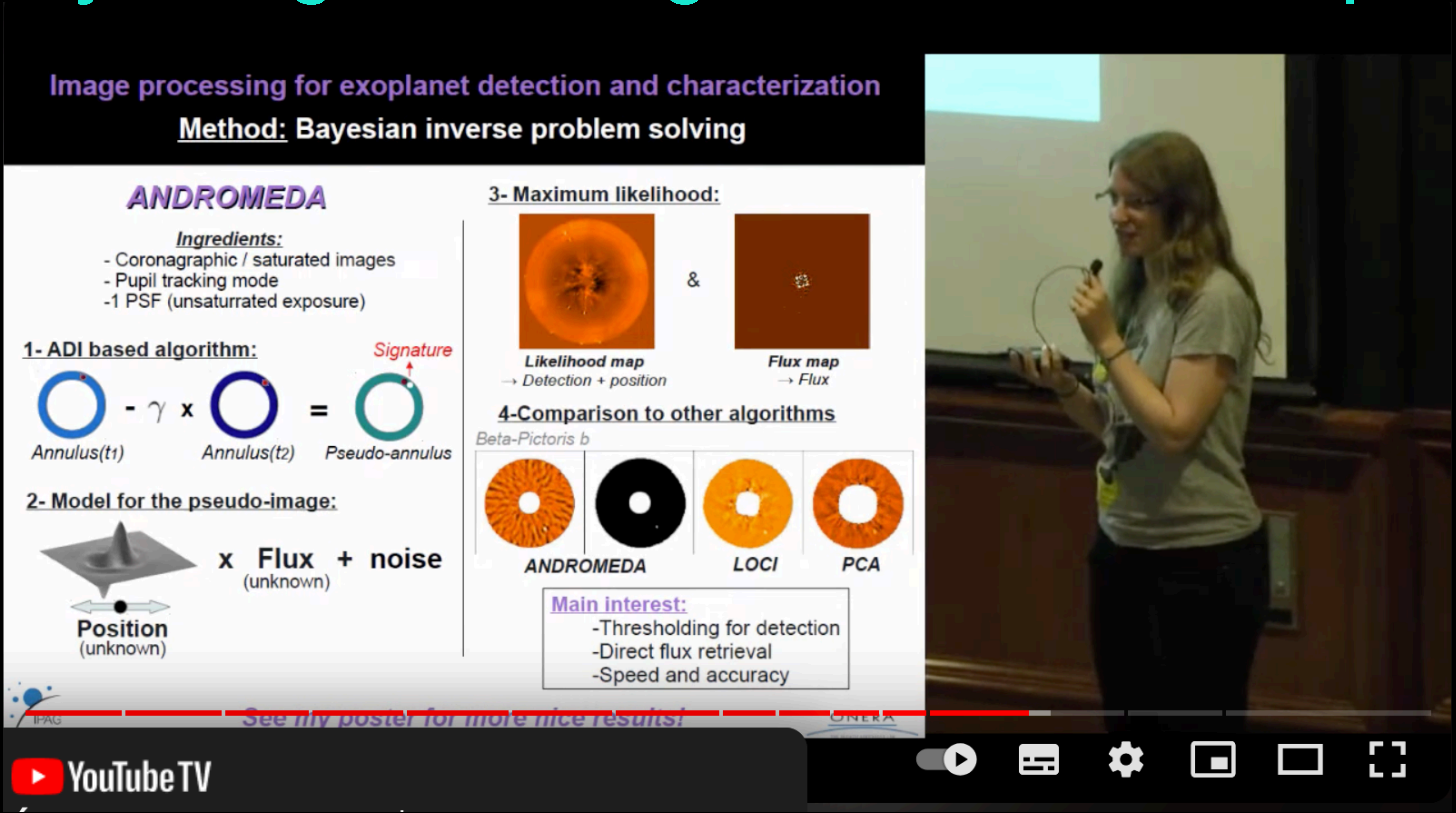
Post-processing for High-Contrast Imaging:
ground-based instruments

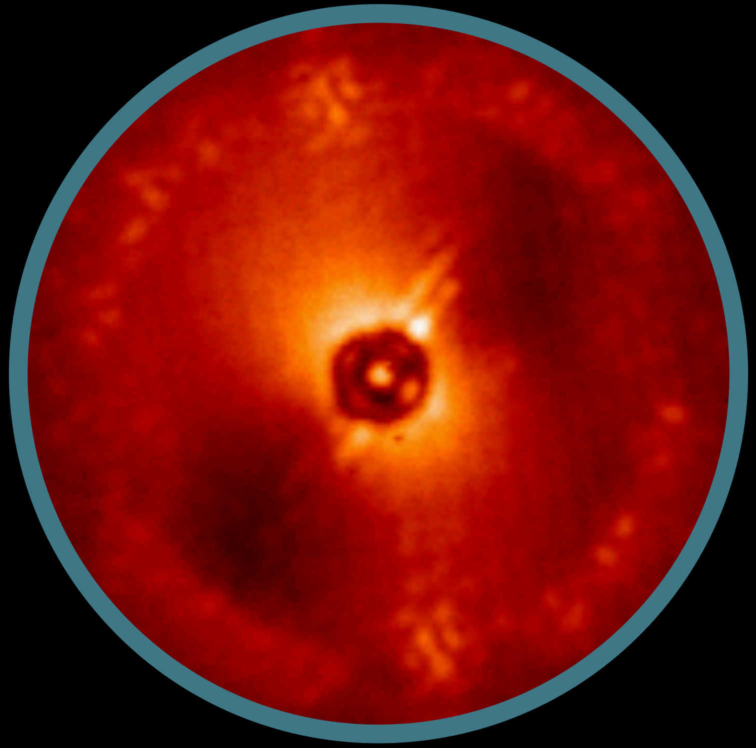




Faustine Cantalloube
Institut de Planétologie et d'Astrophysique de Grenoble
faustine.cantalloube@cnrs.fr

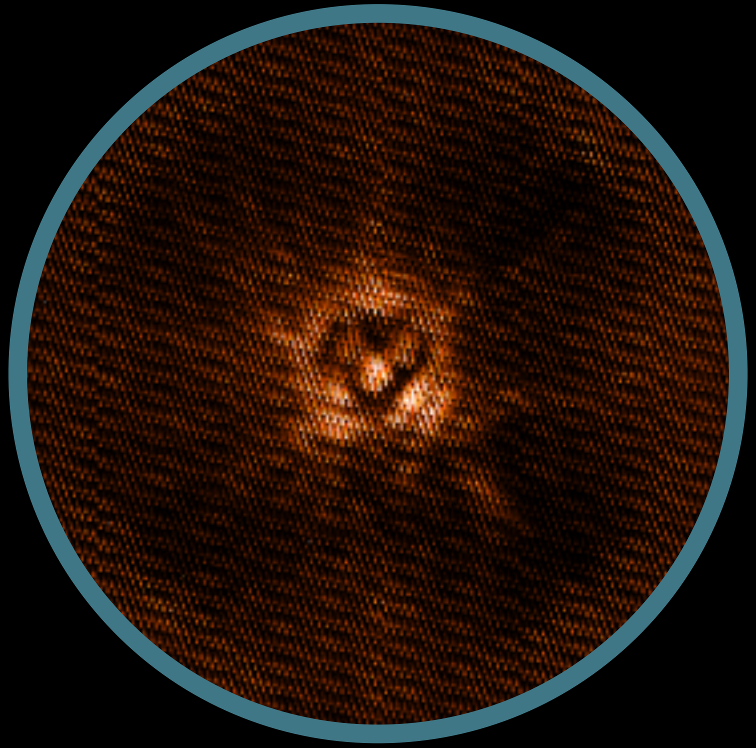
10-years ago at the Sagan Summer Workshop…



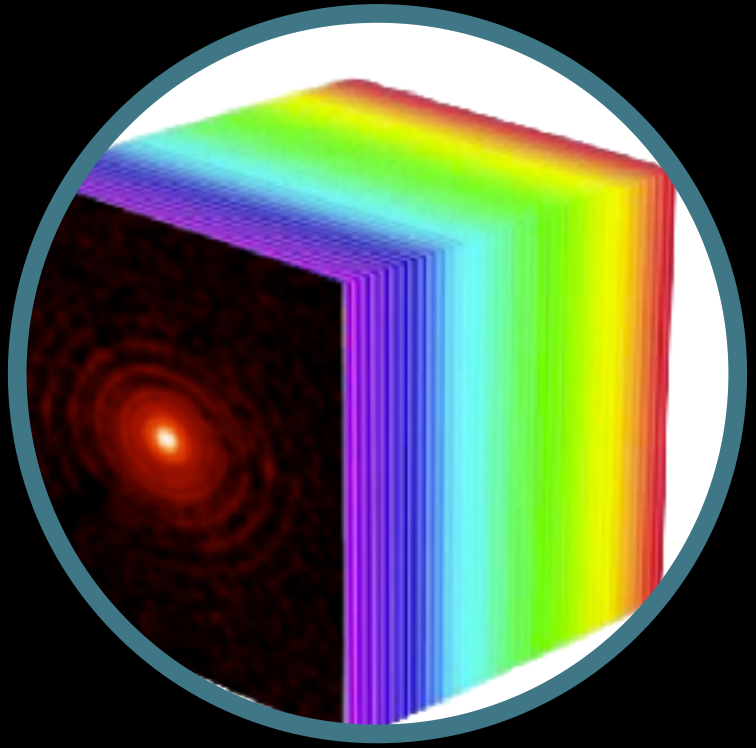
Direct imaging with ground-based telescopes



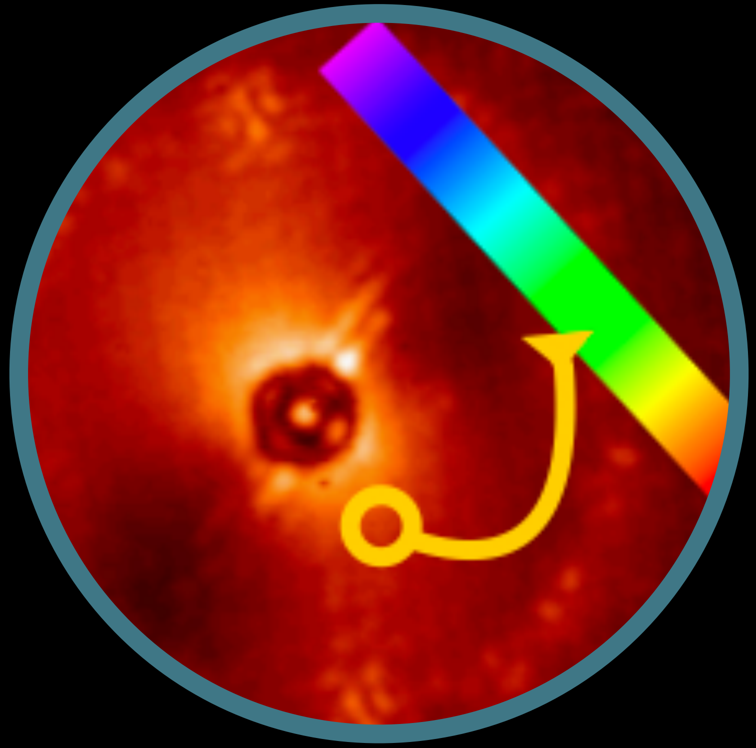
HCI



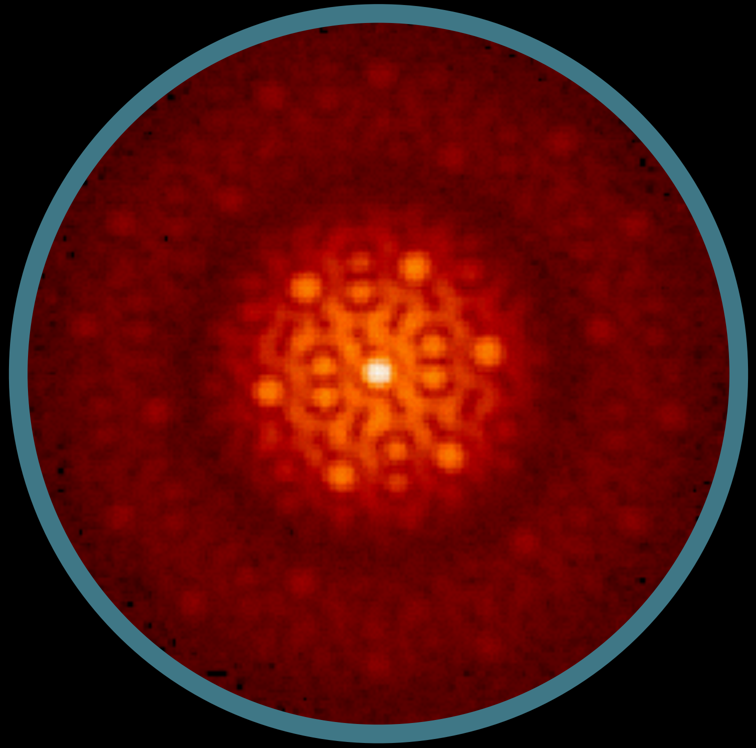
HCI-LRS



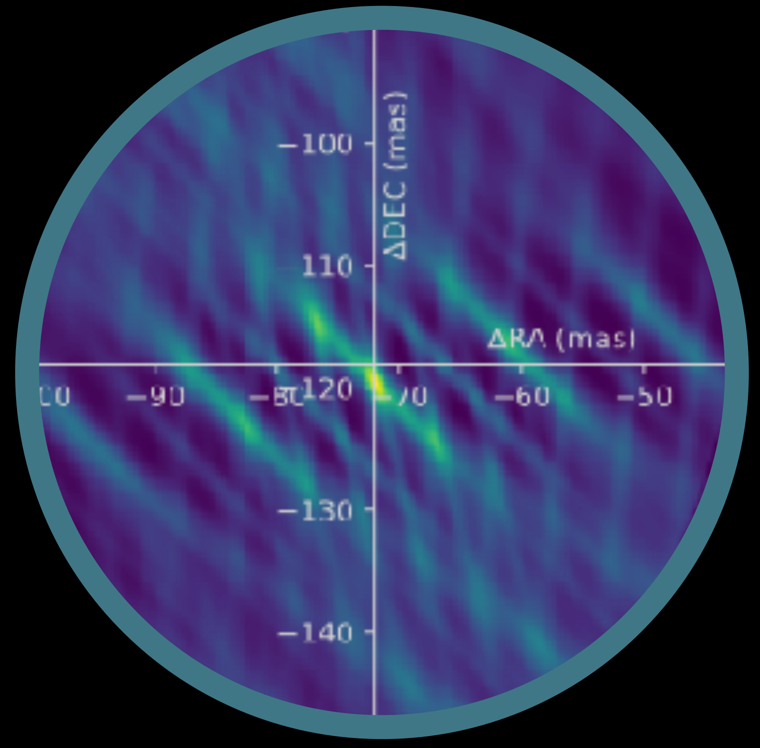
HRS



HDC



SAM



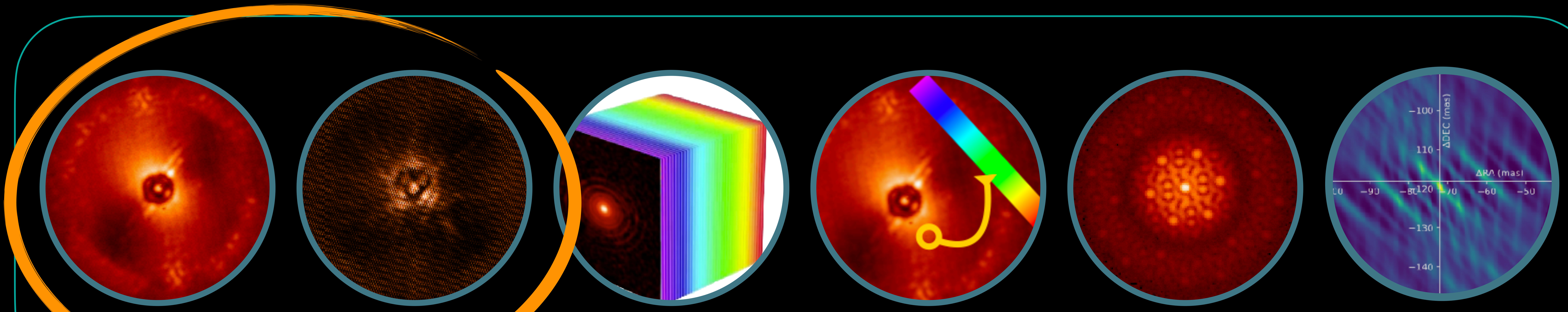
LBI

Wednesday 11.00am J.-B. Ruffio

Friday 11.00am Steph Sallum

Friday 8.30am Guillaume Bourdarot

Direct imaging with ground-based telescopes



HCI HCI-LRS HRS HDC SAM LBI

Faint objects

at close separation

to their bright host star

Sensitivity

High-resolution

High-dynamic

Ground based telescopes (8-m class)

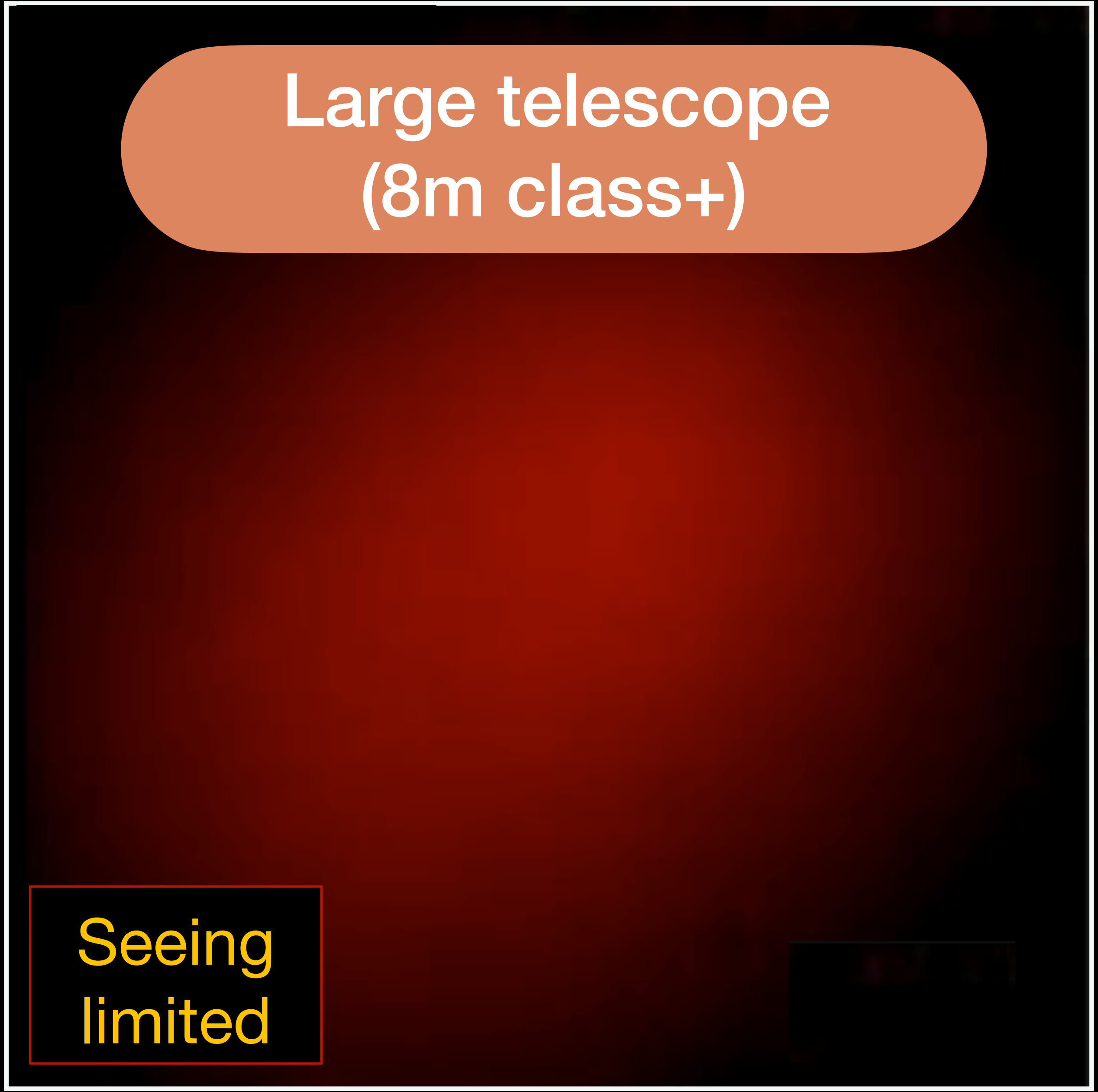
Near - thermal infrared

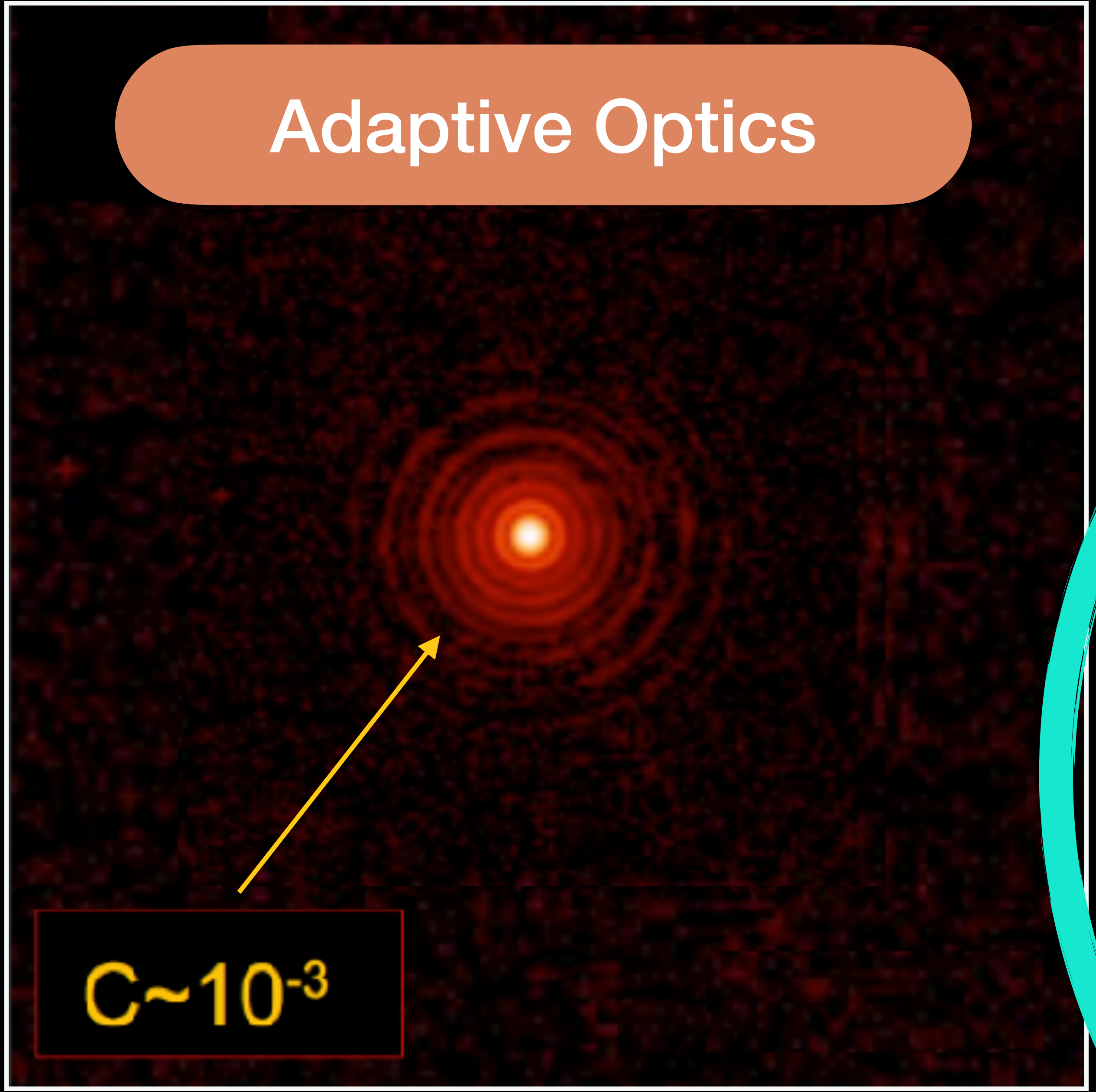
(H-band, 1.6μm to N-band, 10μm)

High-contrast imaging

Also called "coronagraphic imaging"

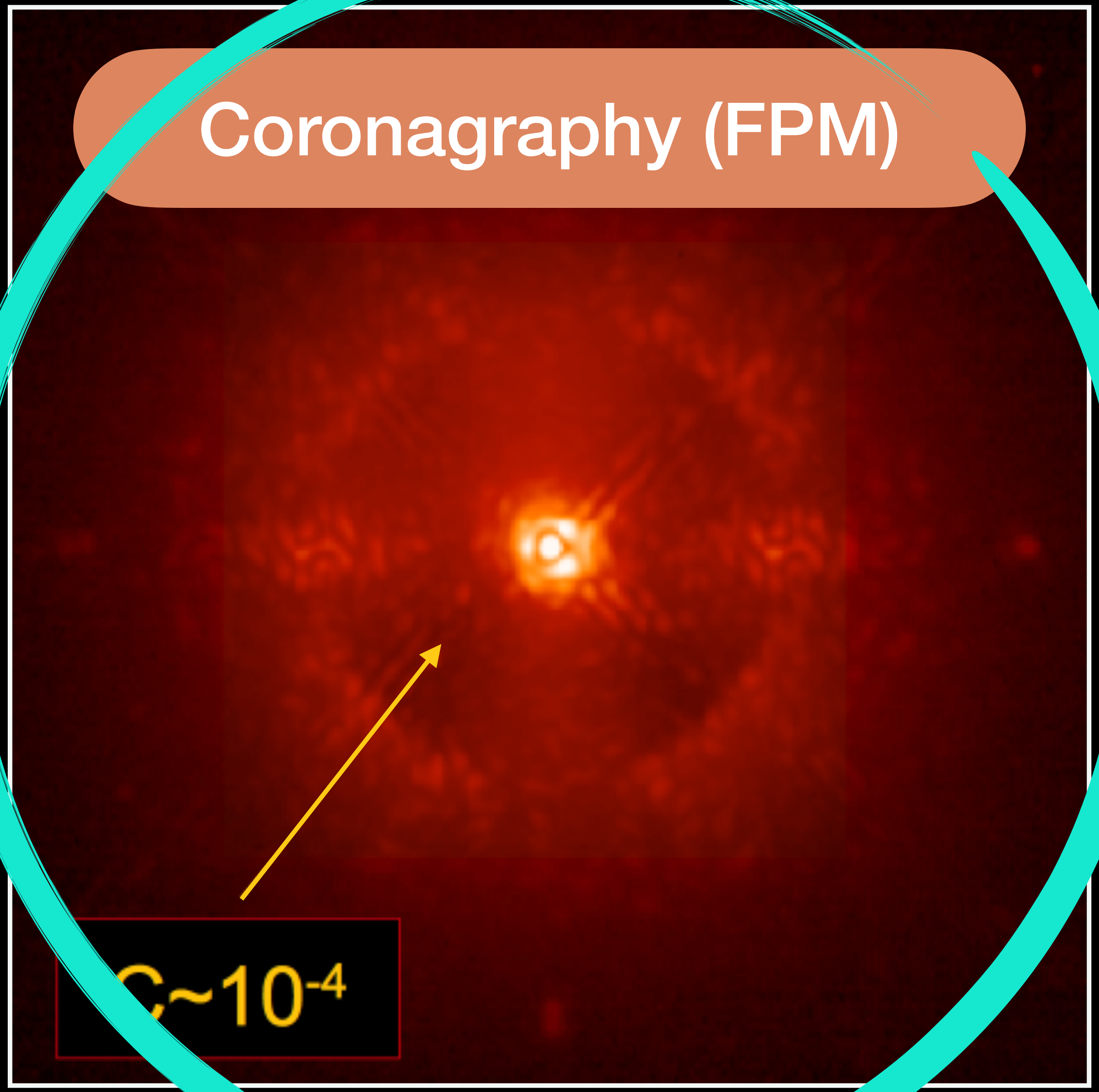
High Angular resolution few milliarcseconds & **High Contrast** more than 10^{-6}

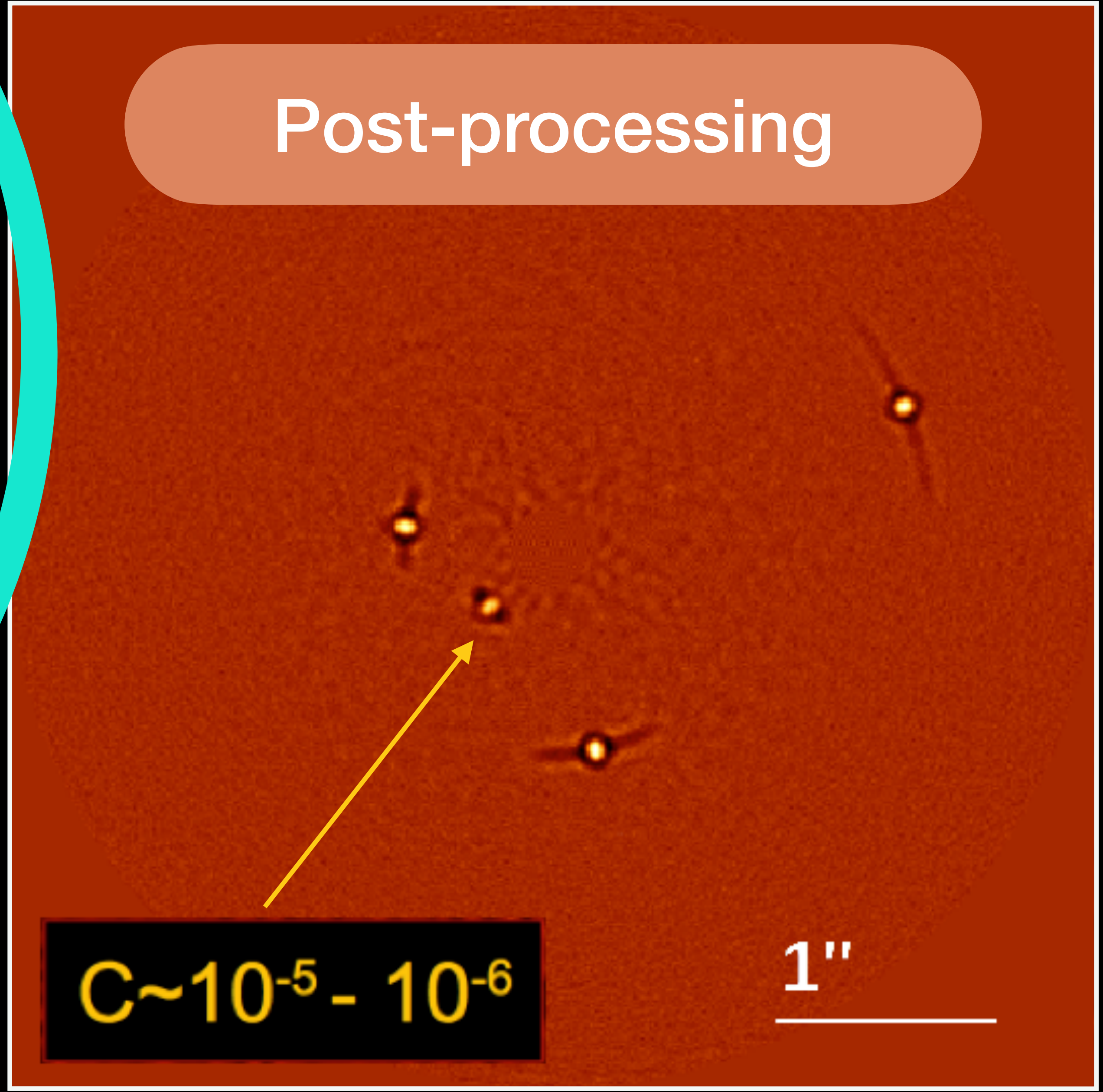




Bright starlight residuals !

Often called "stellar halo"
or "speckle field", …
"stellar glare"



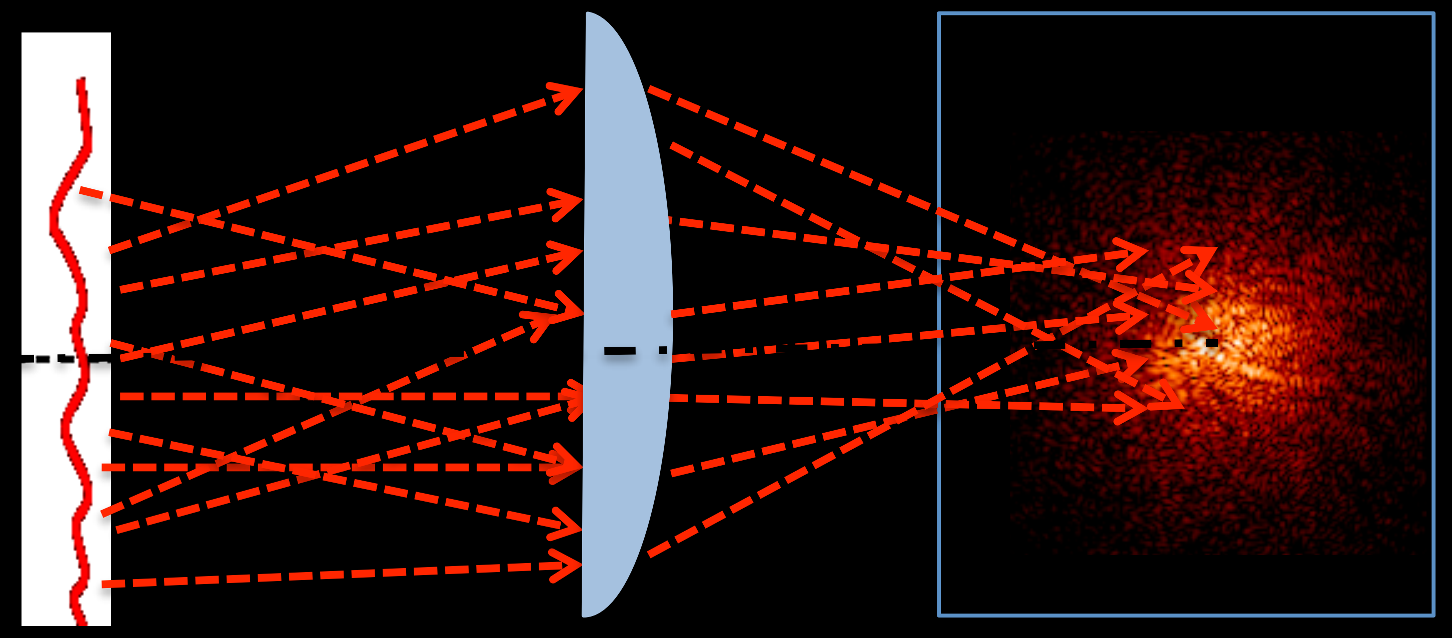


VLT/SPHERE-IR

High-contrast imaging

Bright starlight residuals !

➡Incoming wavefront is **not perfectly flat**



➡Light-rays **interfere** in the focal plane

High-contrast imaging

Wednesday 10.30am Rob De Rosa

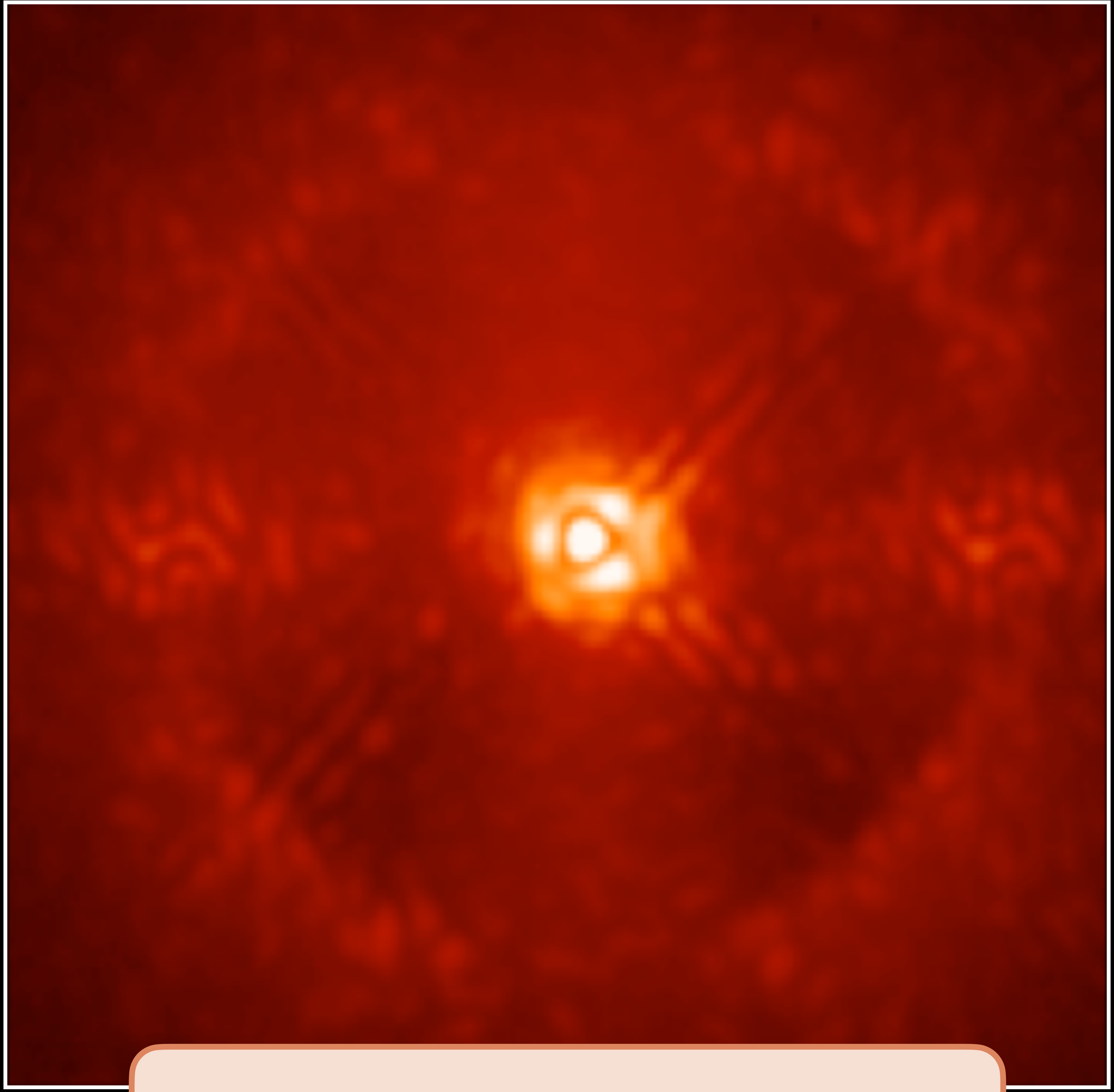
There is a **pre-processing** step here

Also called "cosmetic", "calibration", "data reduction"…

- **#1- Adaptive optics residuals:** usually halos (long exposure time ~1sec)
- **#2- Low order aberrations:** starlight leakage (vibration/jitter, low wind effect …)
- **#3- Non-common path aberrations:** high order spatial frequencies
- **#4- Detection noise:** detector noise, background, readout noise, bad pixels…

Summary of our task

Saying that at a given radius, mean and variance is the same is usually **wrong**



Thursday 11am Julien Milli

What do we want / need:

- **Detection**: discriminate H_0 from H_1 + confidence level
- **Characterisation**: astrometry & photometry + uncertainties
- **Detection limit**: algorithm performance (astro-centered)
- **Comparison**: apply different algorithms (algo-centered)

Maximizing True positives, minimizing False negatives

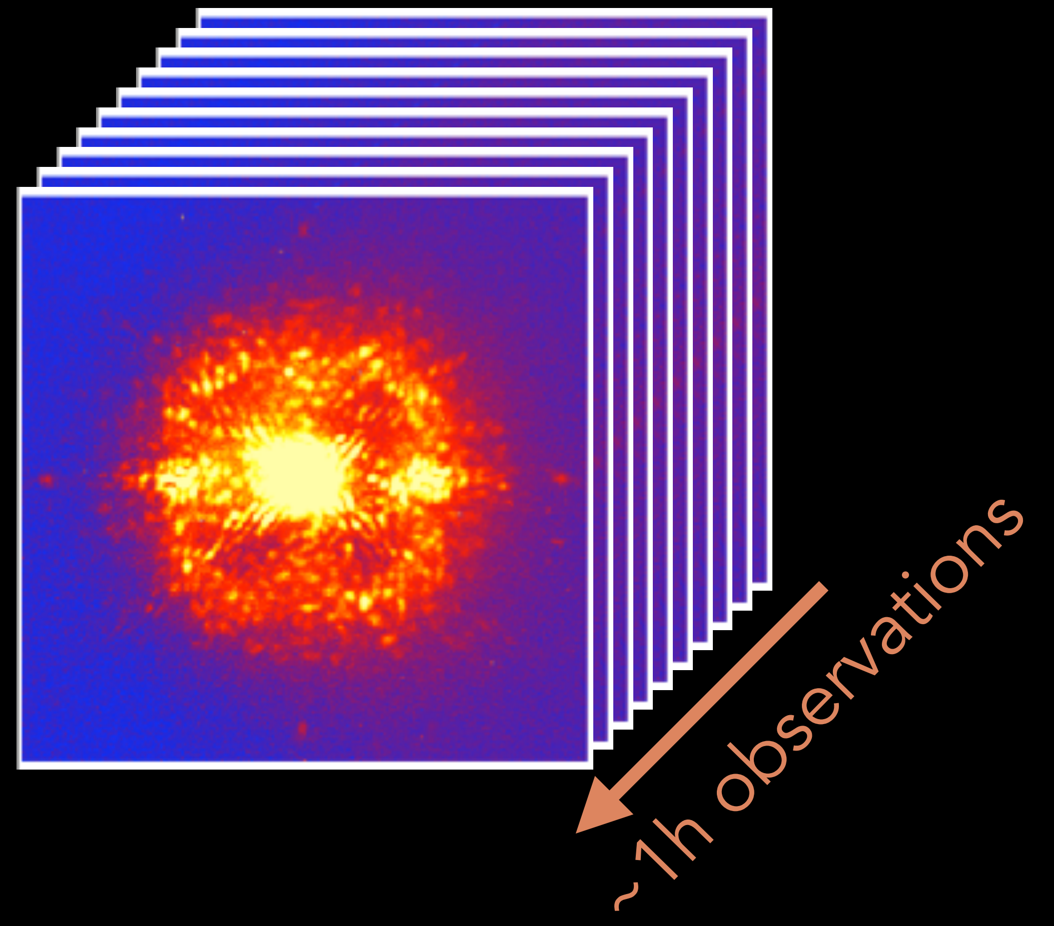
For point source (substellar companions) **&** extended source (circumstellar disks)

Tailored image processing techniques to carve out the residual starlight —> **10^{-6} contrast**

Differential imaging 101

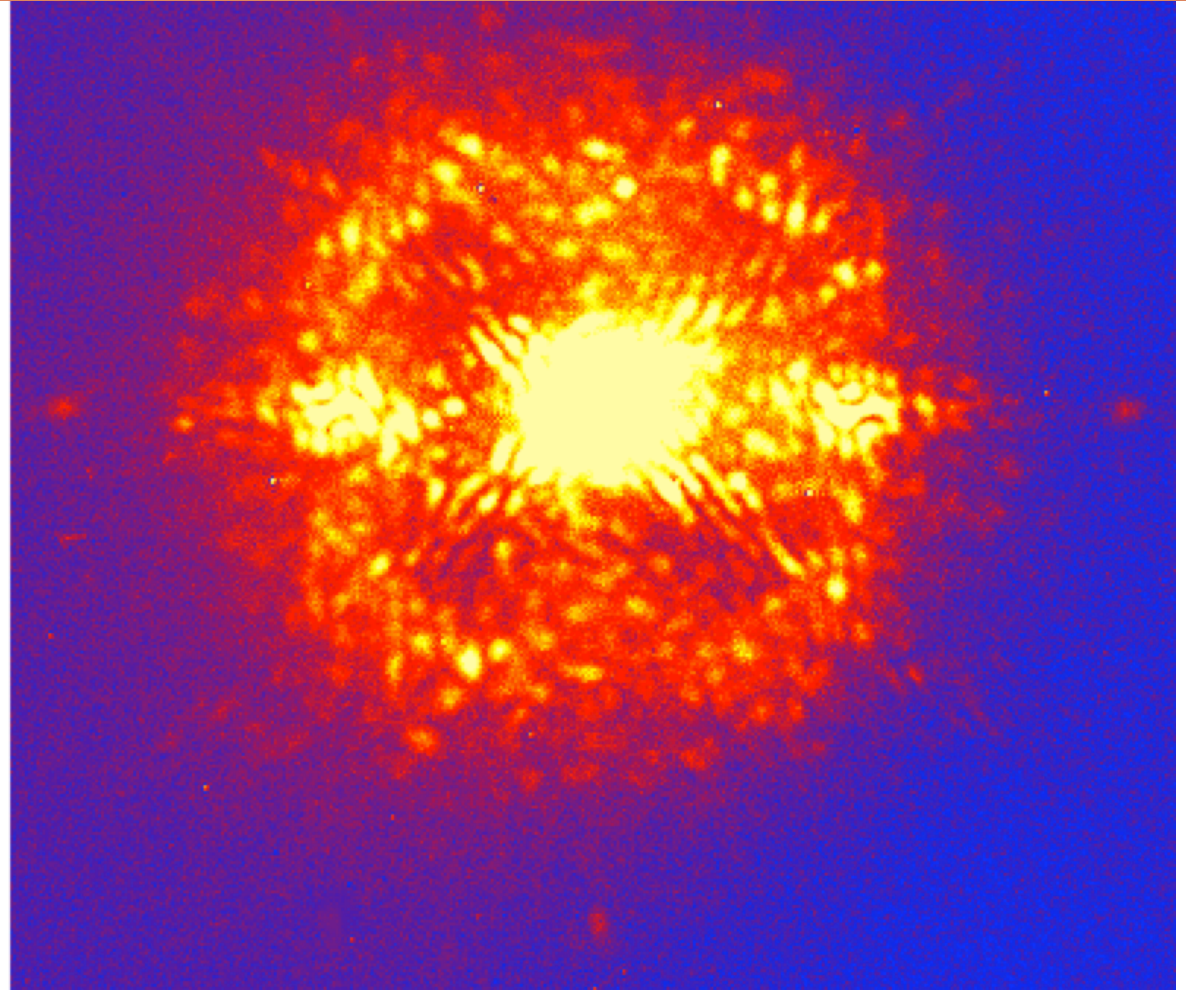
Find a **different behaviour** between **(1)** the astrophysical signals
and **(2)** the starlight residuals

➡ Exploit this **diversity** to recover the signal !



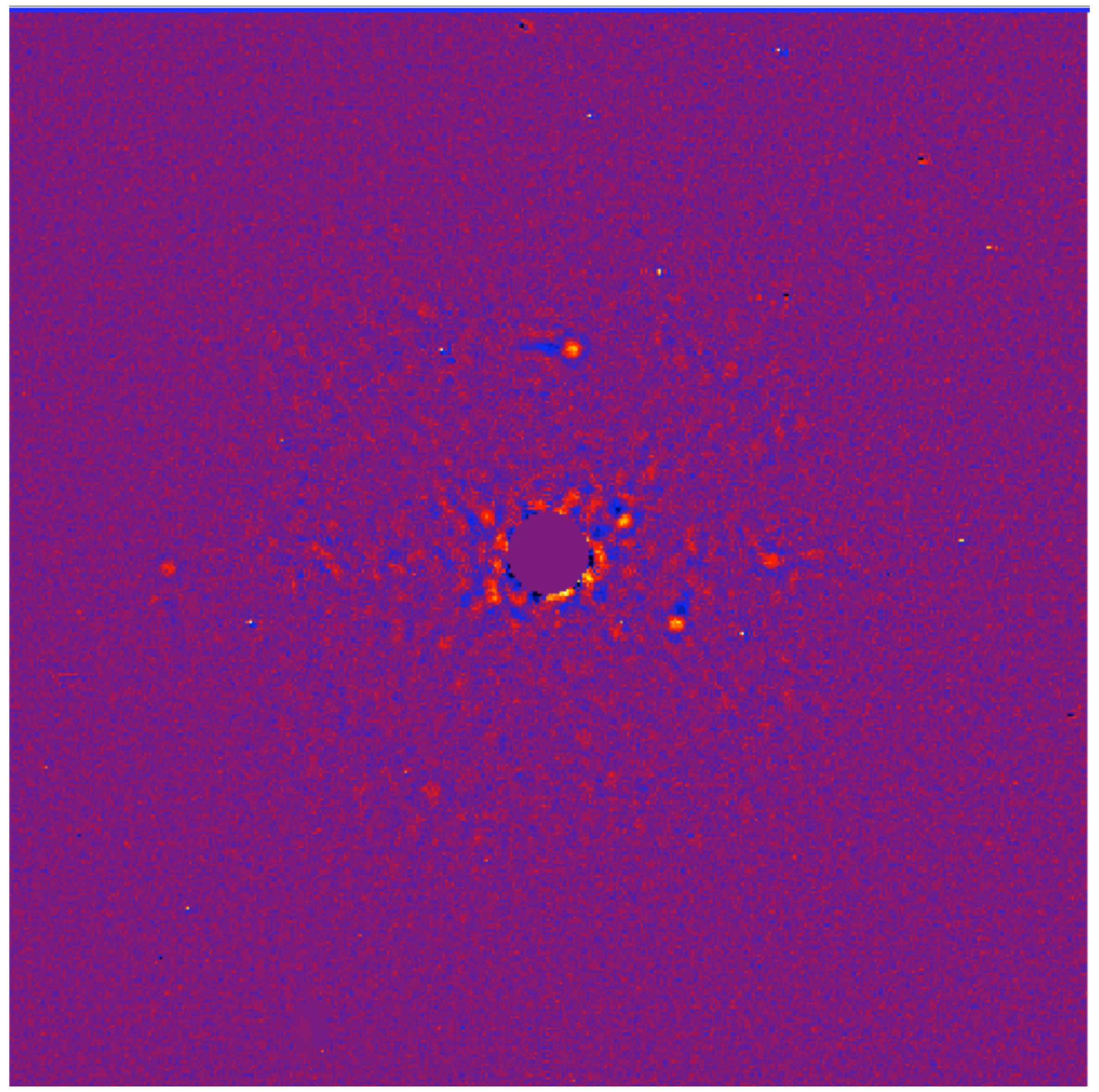
1. Estimate the star image

Often called "reference" PSF



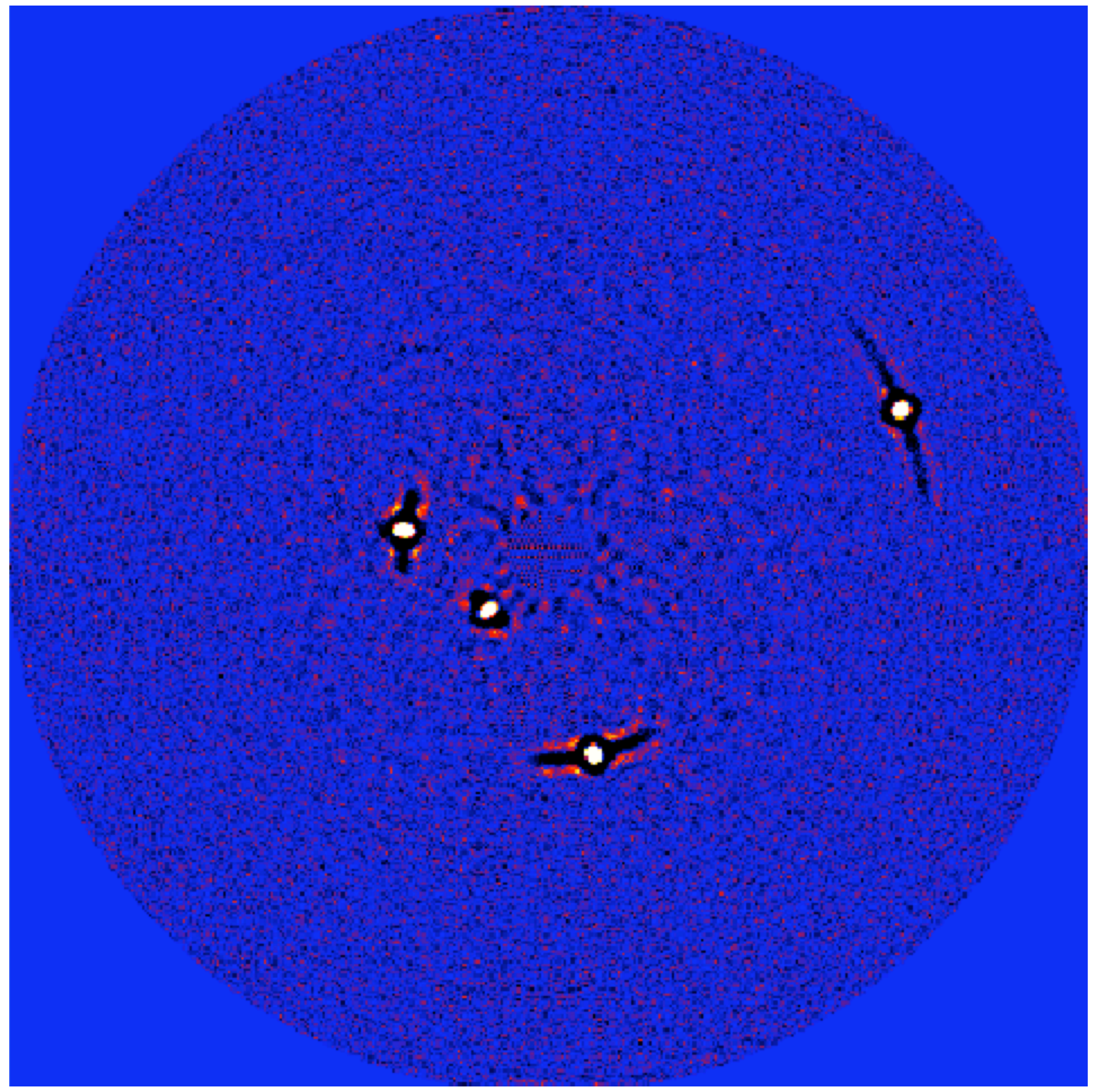
This step is **critical !!!**

2. Subtract it to the image



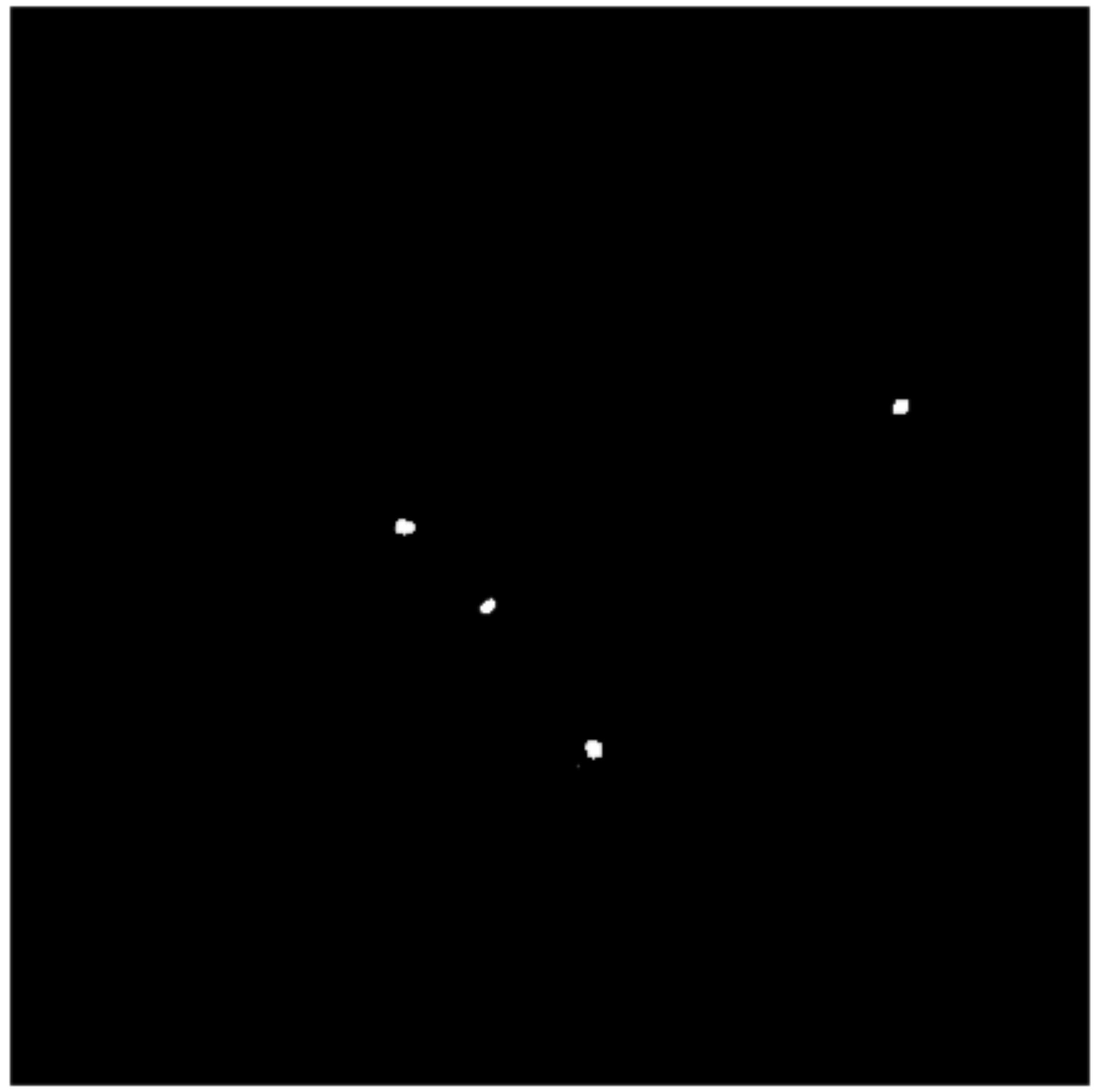
Whitens the residuals !

3. Combine all the images



This can be **optimized**

4. Build a detection map



Detection by thresholding

Differential imaging 101

Need a specific **observing strategy** & **calibration** procedure

➡ Provide with various **diversities**

Binary Differential Imaging: *Rodigas et al., 2015*

Reference Differential Imaging: *Mawet et al., 2009, Rameau et al. 2012*

multi-Reference Differential Imaging: *Xuan et al. 2018, Bohn et al. 2019, Ruane et al. 2019*

Polarimetric Differential Imaging: *Kuhn et al. 2001*

Spectral Differential Imaging: *Racine et al. 1999, Sparks and Ford 2002*

Coherence Differential Imaging: *Baudoz et al. 2005*

Angular Differential Imaging: *Marois 2006, Davies 1980*

...

Focus on ADI-based techniques !

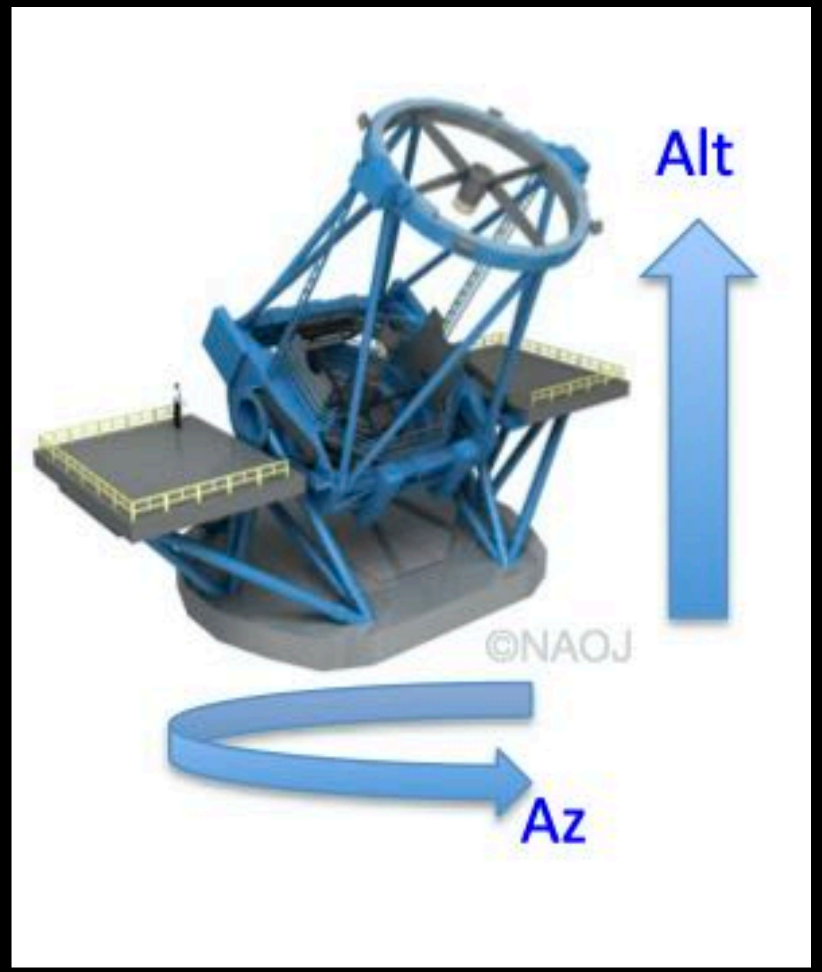
Angular Differential Imaging

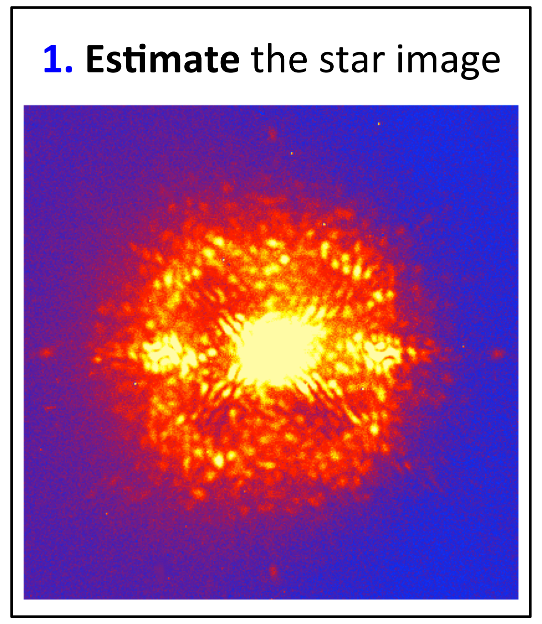
Pupil tracking mode:

For an alt-az mount telescope
Disable the field derotator

It brings **angular** diversity

ADI is a technique, not an observing mode or data set type



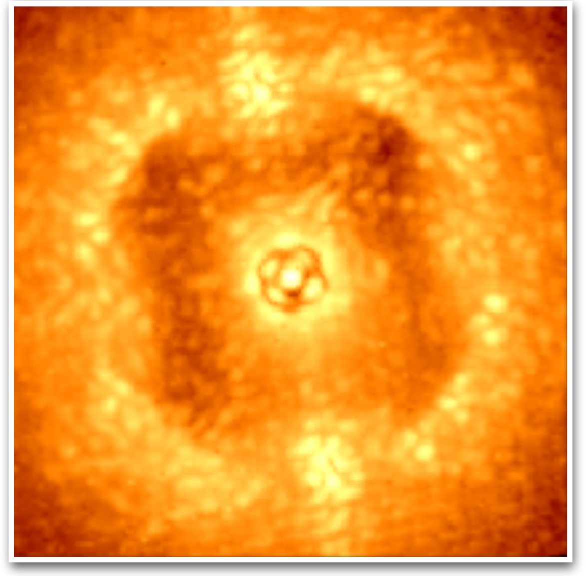


1.Estimating the star image

The basic approach

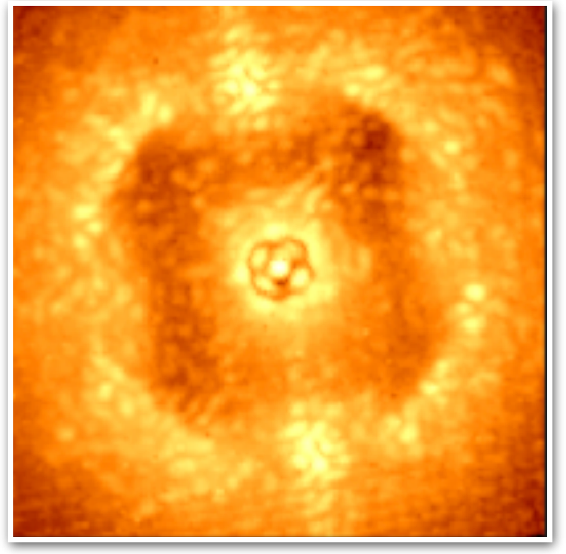
Pairwise subtraction:

The closest frames in time are the most correlated
while not self-subtracting the signal



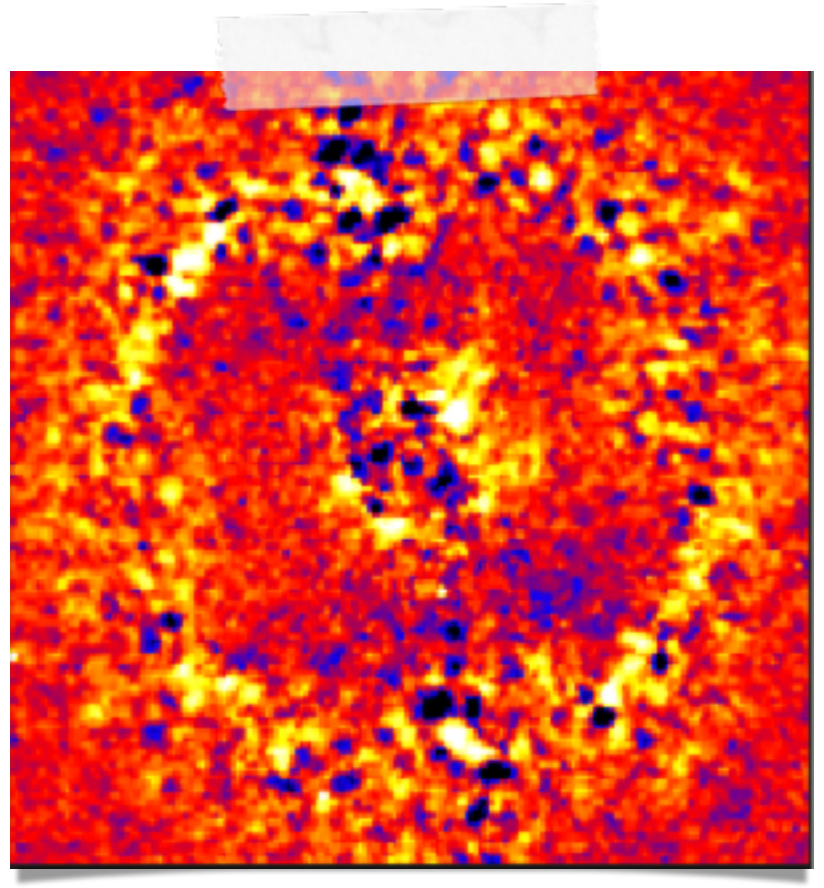
frame (t1)

\-



frame (t2)

=

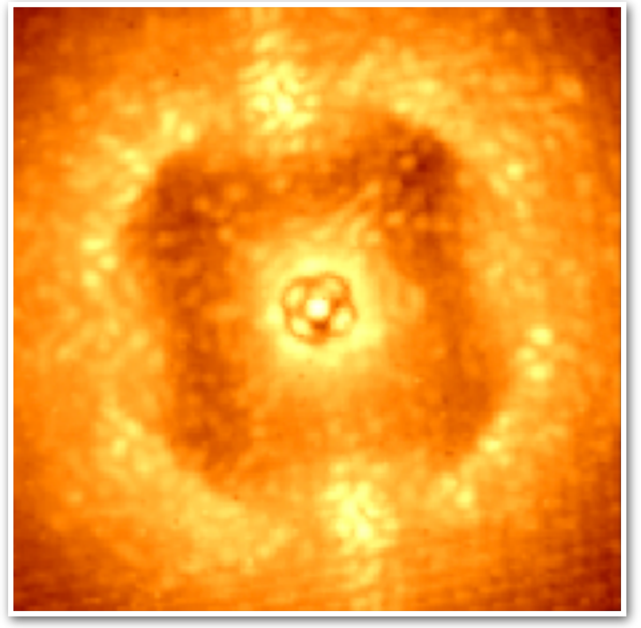


The optimal Δα is **0.5 lambda/D**

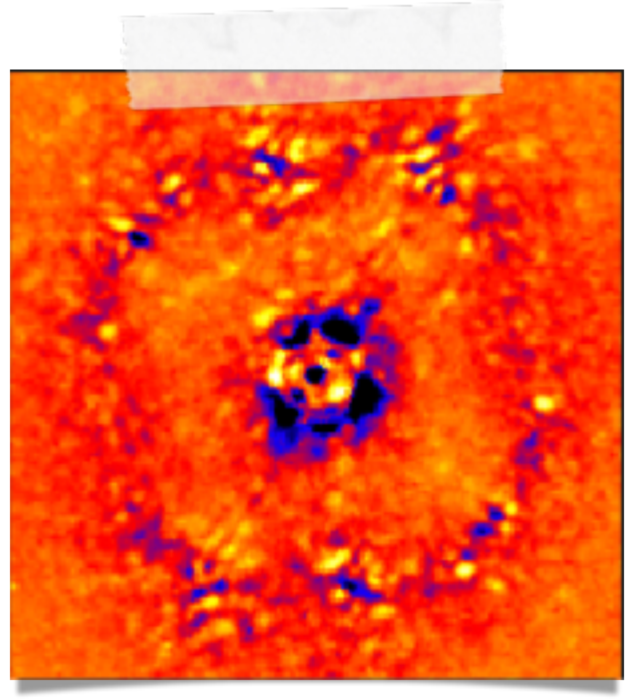
Temporal median:

The median represents the 'typical' image
while the moving signal is not taken into account

Marois et al., 2006

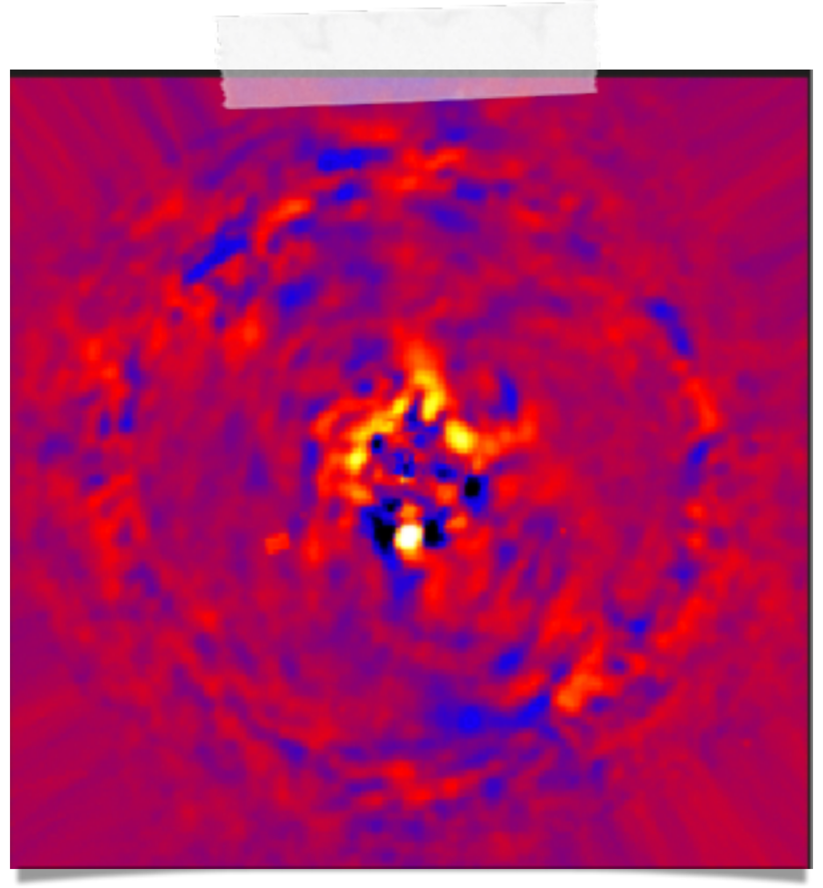


Temporal median



frame (t1) -median

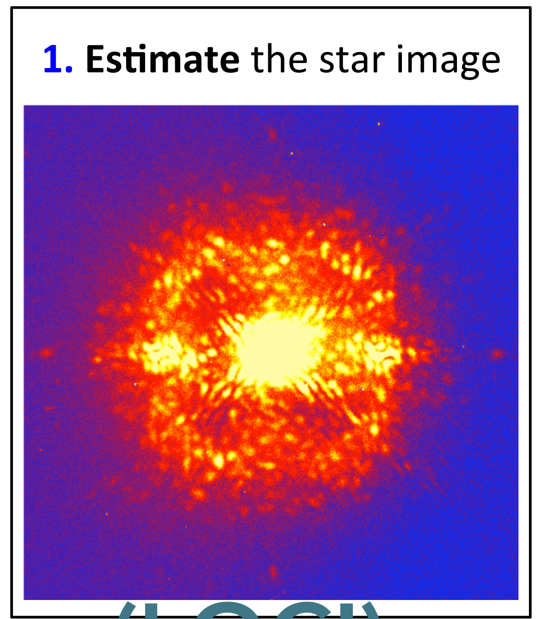
…



It comes with various flavours: smart-ADI, annular-ADI
And variations: Image Rotation Subtraction *Dou et al., 2015*

1. Estimating the star image

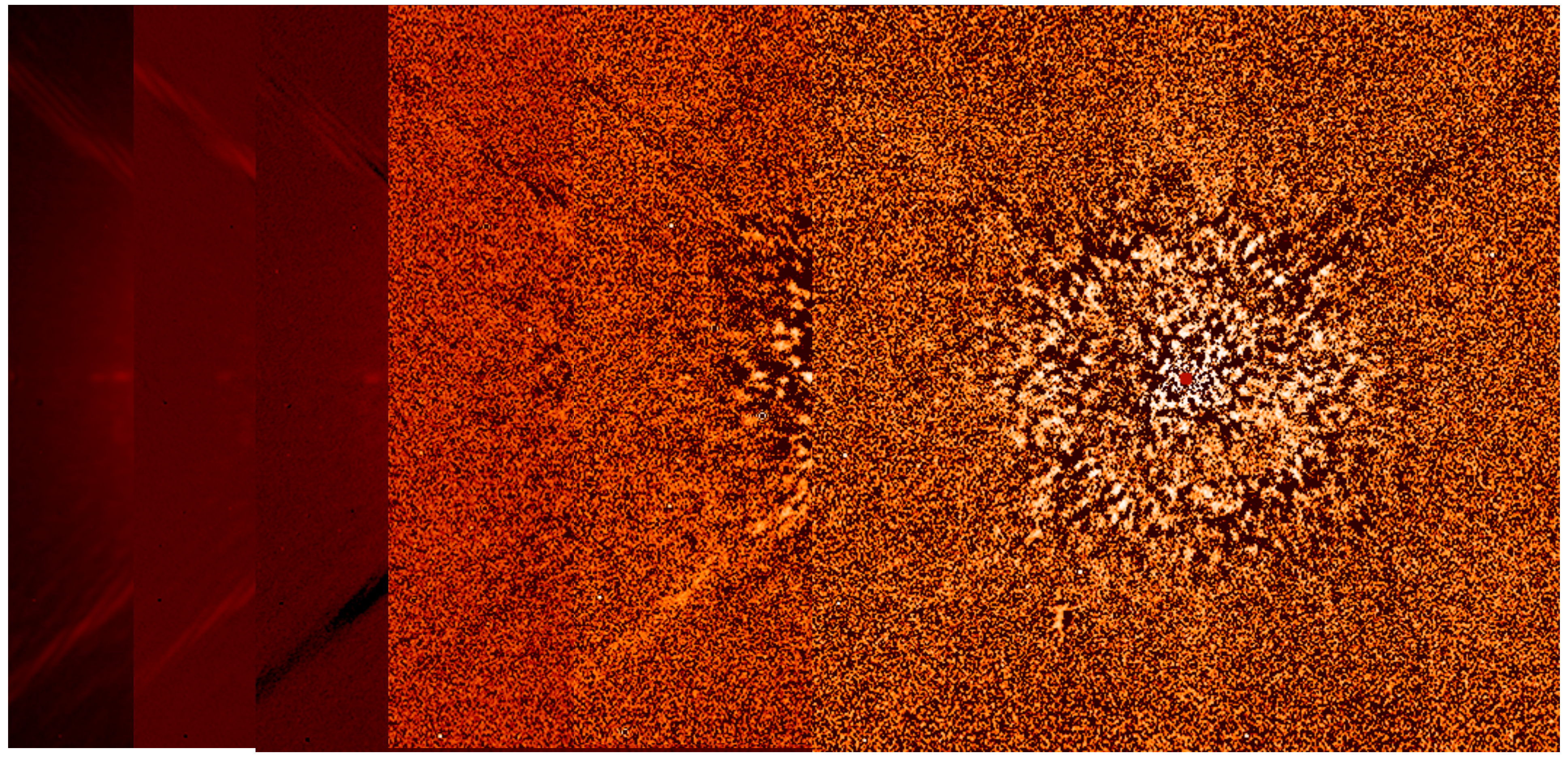
The classic approach



Principal Component Analysis (PCA):

Linear combination of the images of the cube decomposed over orthogonal basis (eigen-images)

Soummer et al., 2012 | *Amara & Quanz 2012*



PC:
#1
#2
#3
#25
#50
#197
(last)

And also…
smart-PCA, *Absil et al., 2013*
LLGS, *Gonzalez et al., 2016*
AMAT, *Daglaya et al., subm.*
NMF, *Ren et al., 2018*
Space-Time KLIP, *Lewis et al., 2023*

Locally Optimized Combination of Images (LOCI):

Find the linear combination that minimises the residuals while the moving signal is not taken into account

Lafrenière et al., 2007

Model !

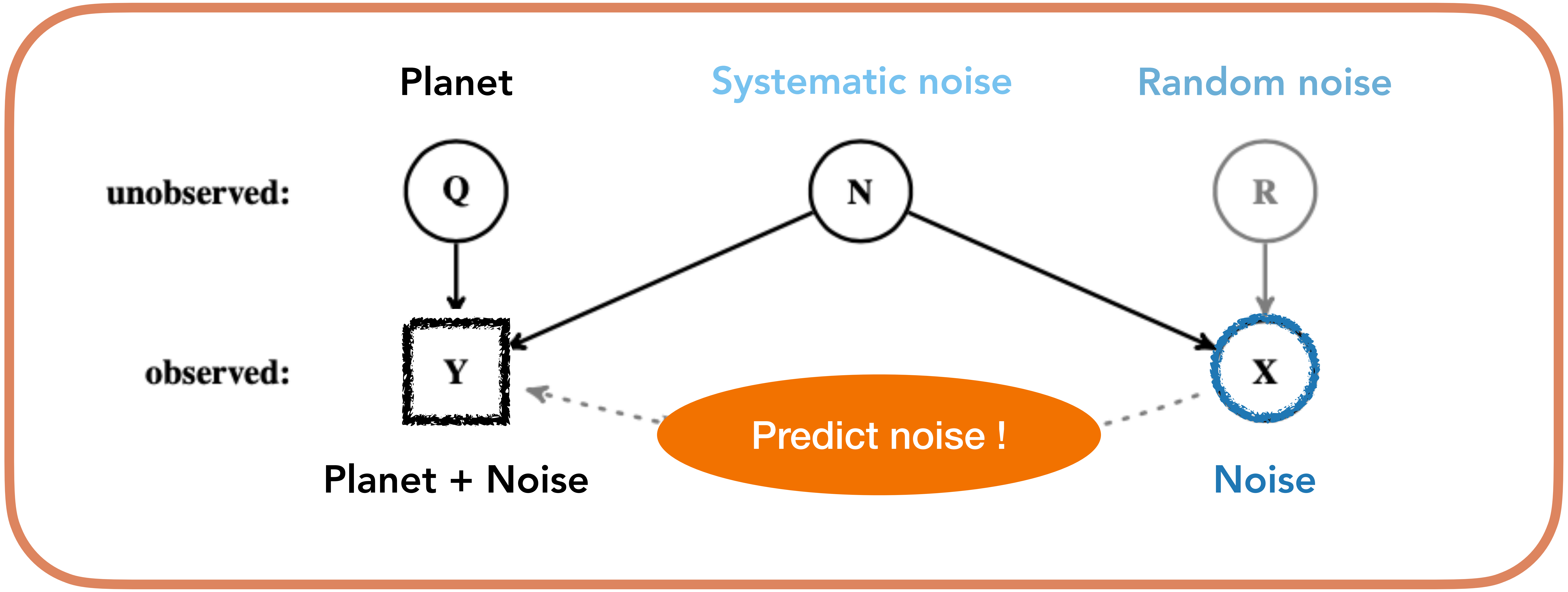
$$\mathrm{Min}(\ \sigma_{res}^{2}) = \mathrm{Min}\sum_i m_i \left(O_i^T - \sum_k c^k O_i^k\right)^2;$$

All the images (t)
Binary mask
Images
Coefficients
Subset of images

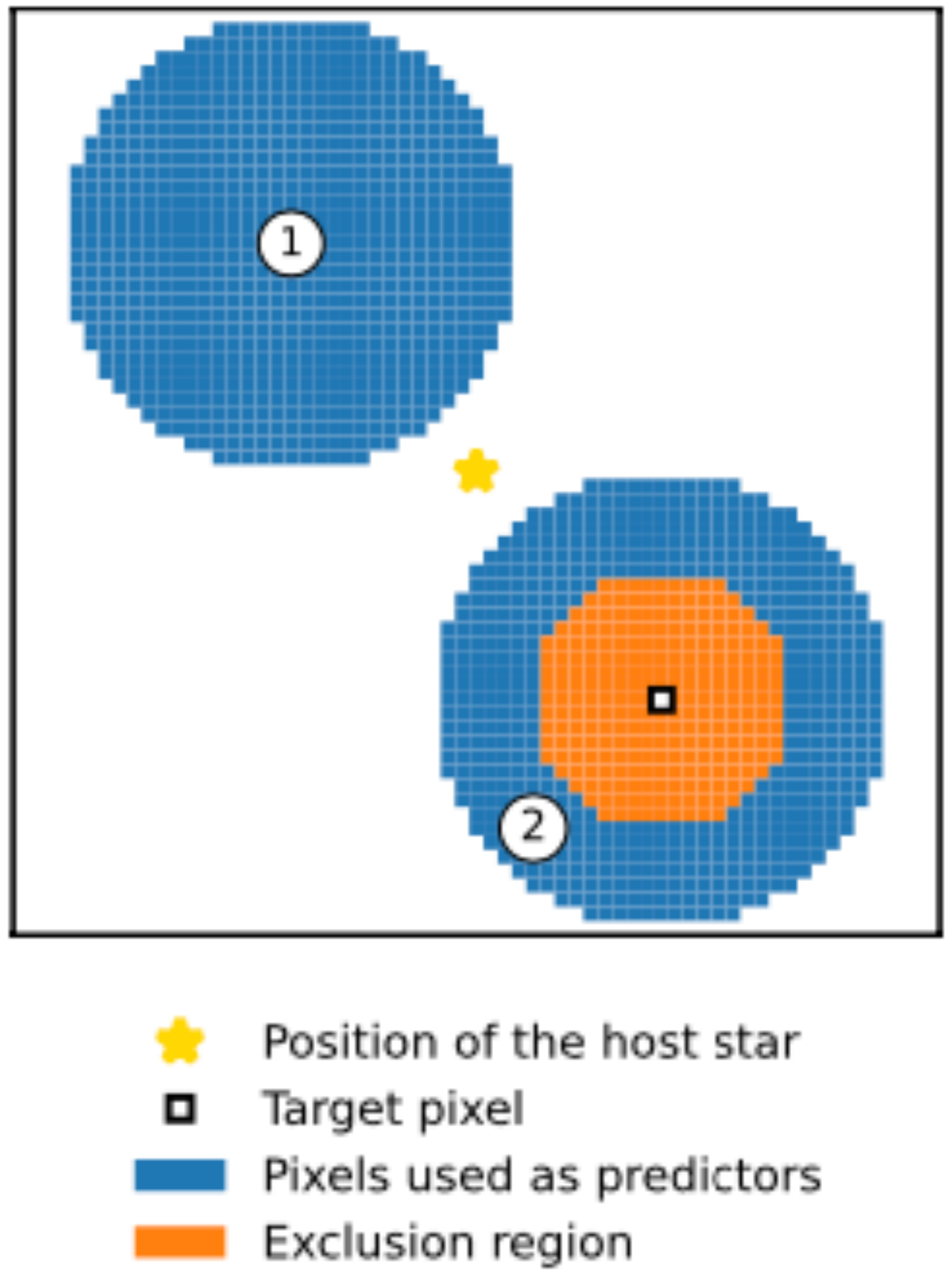
And also…
Template-LOCI, *Marois et al., 2014*
Adaptive-LOCI, *Currie et al., 2012*
Matched-LOCI, *Wahhaj et al., 2015*
Damped-LOCI, *Pueyo et al., 2016*

1. Estimating the star image

Half Sibling Regression



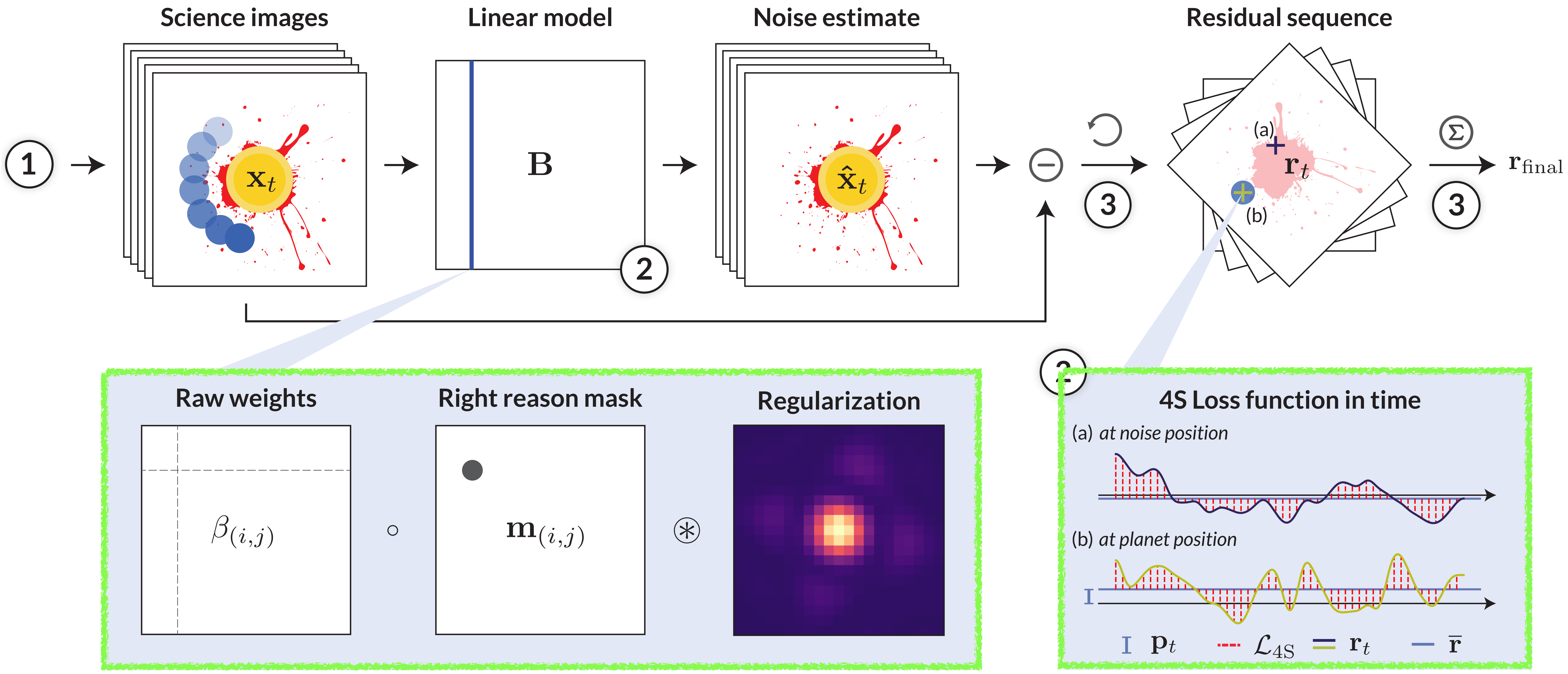
A causal model to predict the systematic noise



Gebhard et al., 2022

1. Estimating the star image

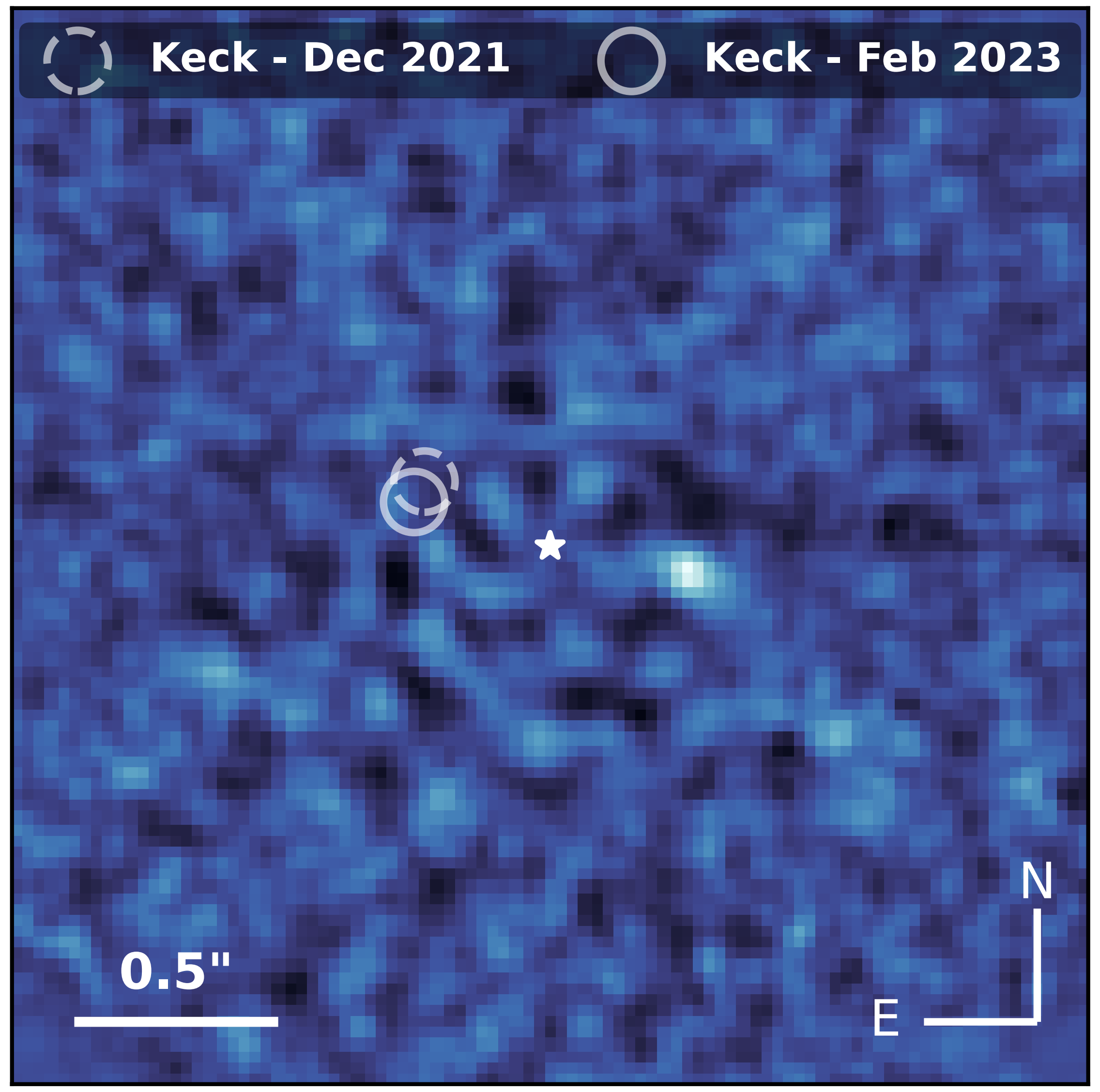
Signal Safe Speckle Subtraction (4S)



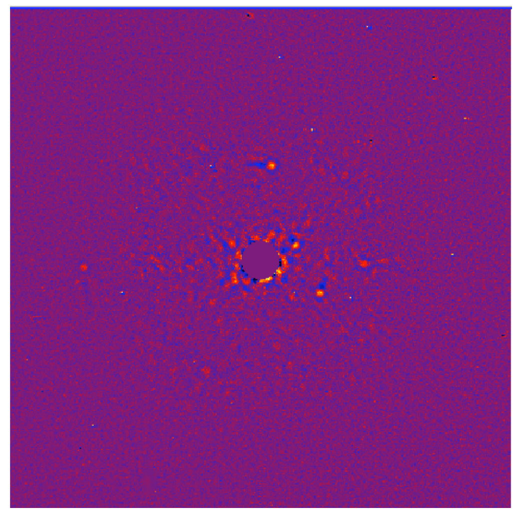
Explainable Machine Learning

Auto-Grad against signal loss

AF Lep b (2011)



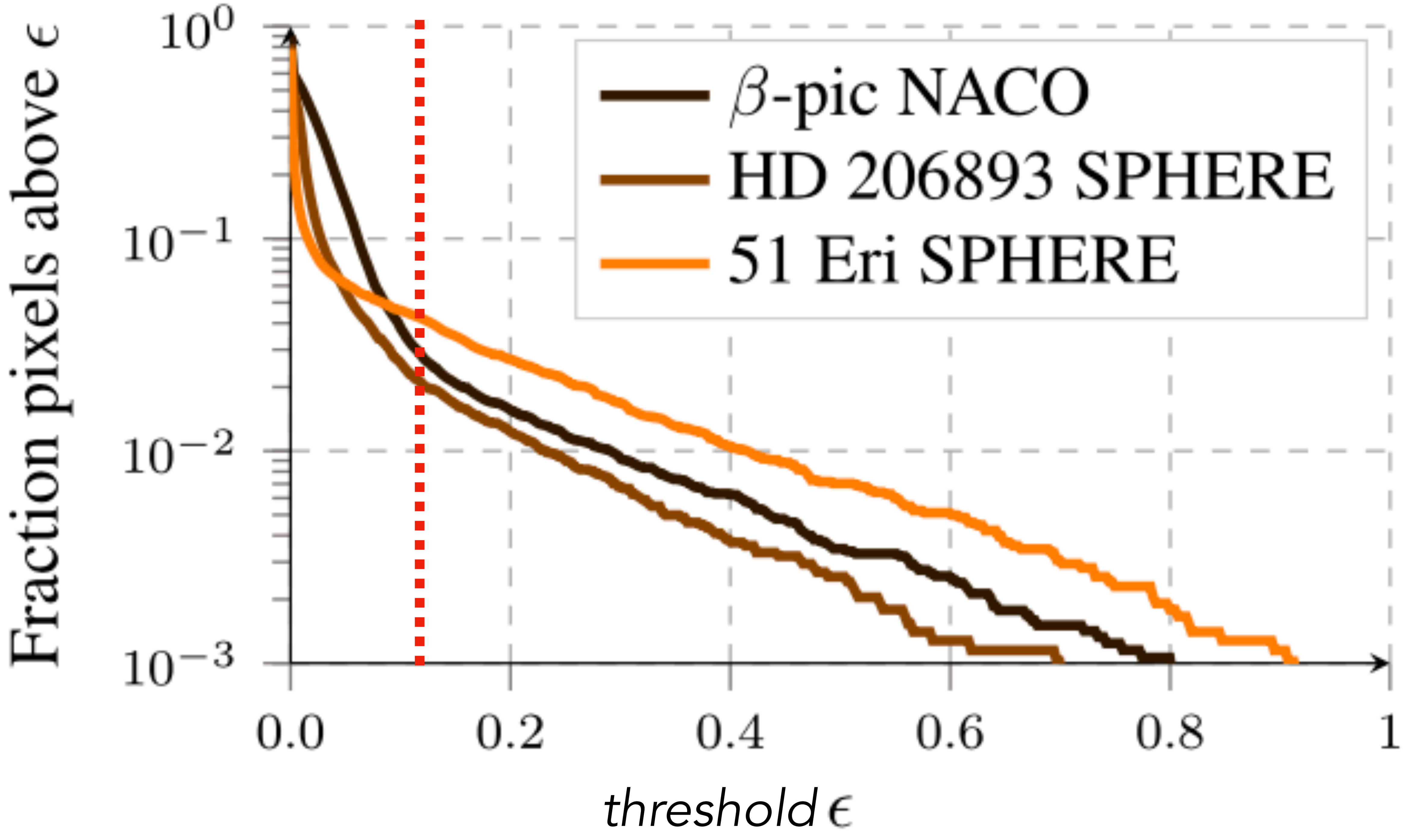
Bonse et al., 2024

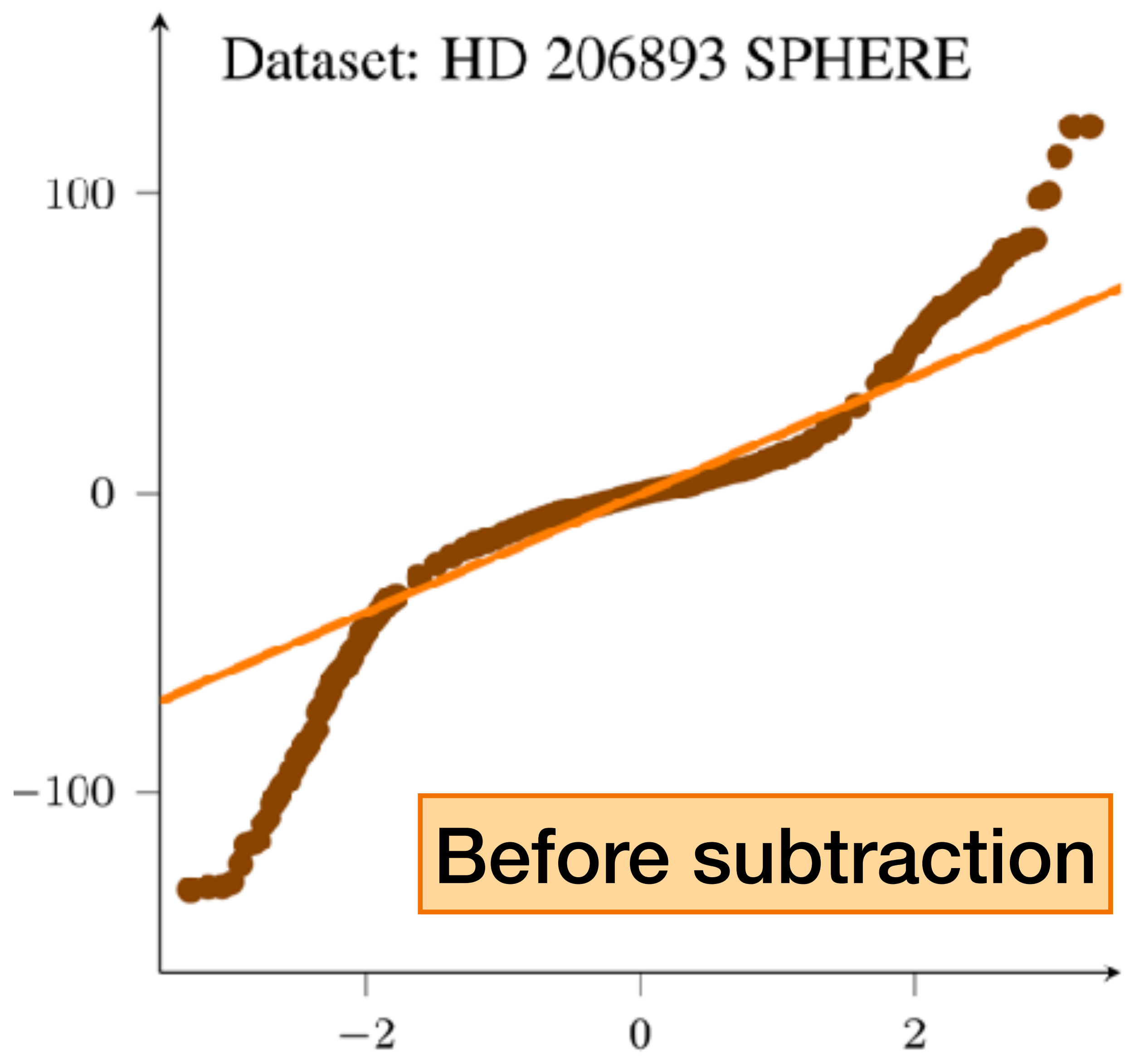


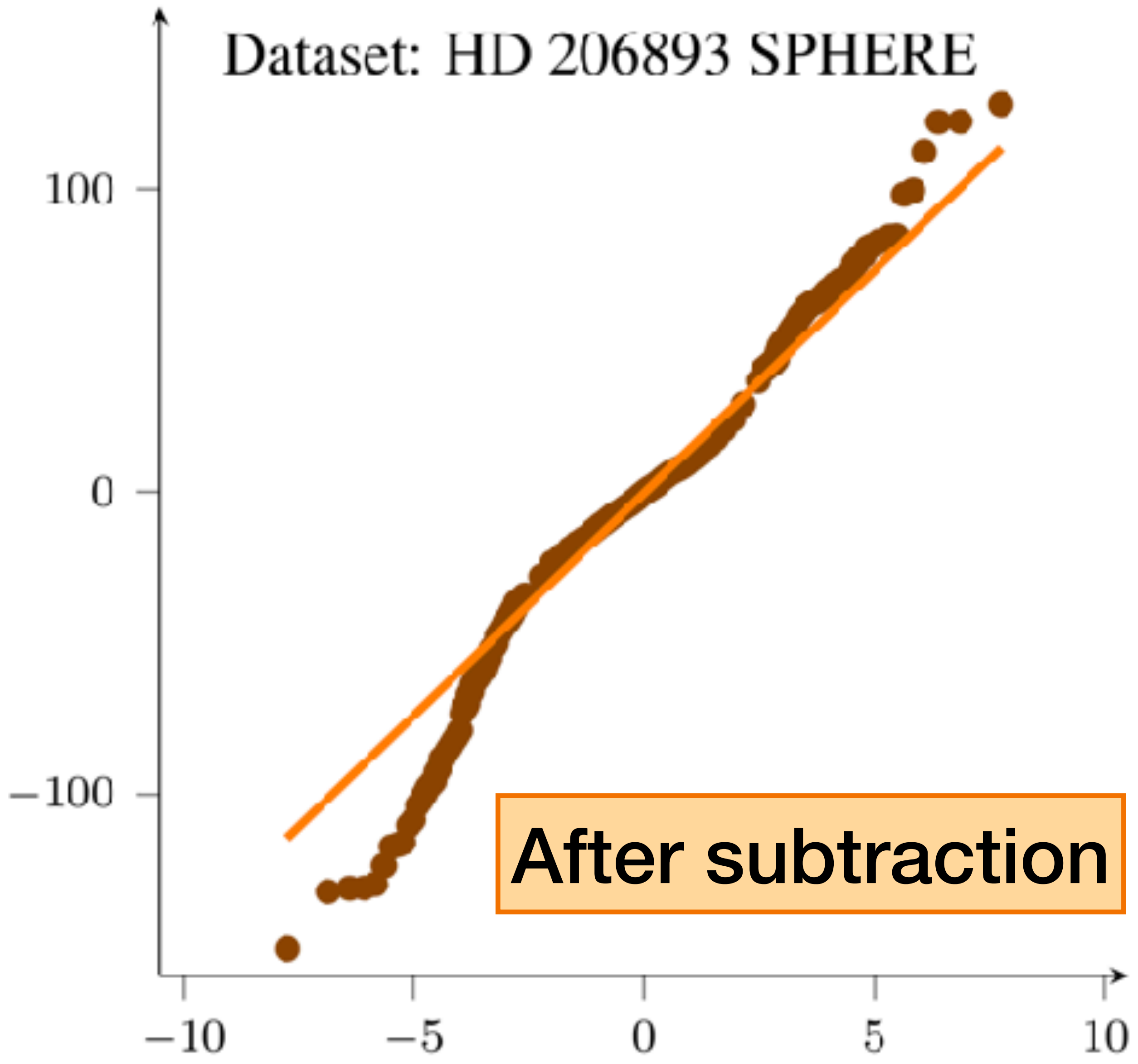
2. Residuals after subtraction

Also called 'subtraction residuals', 'differential imaging' residuals, 'post-processing residuals'

The noise distribution of the residuals is sub-exponential (and <u>not</u> Gaussian)
> hence the high number of false positives !







It depends:
- On the data itself (instruments and conditions)
- On the distance to the star

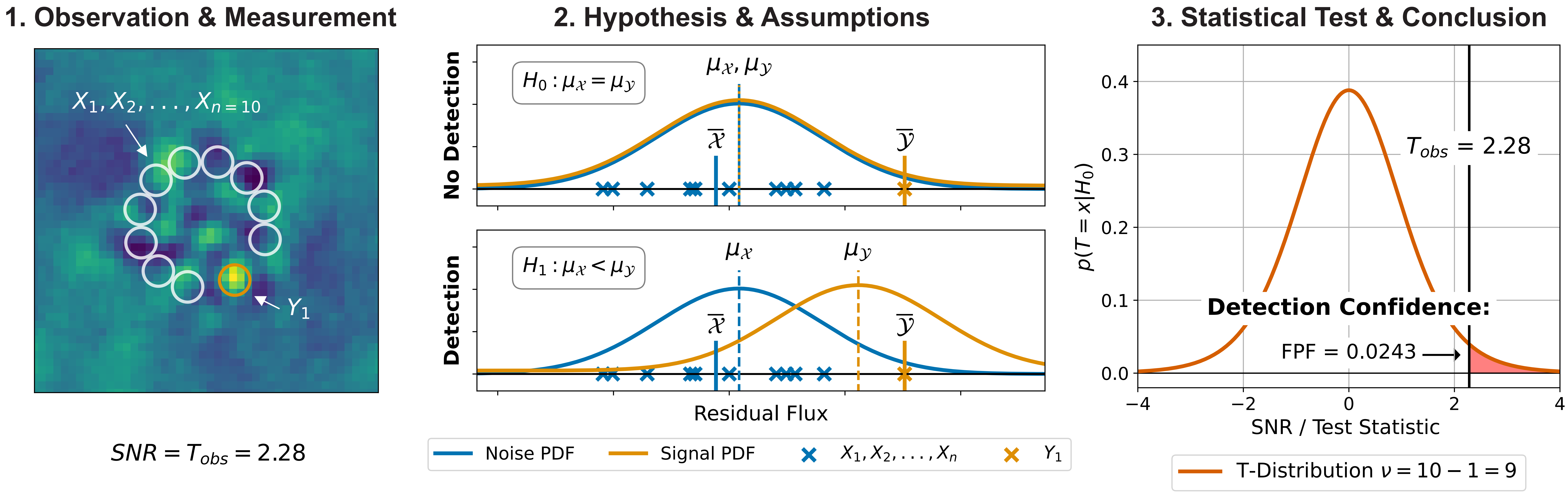
Pairet et al., 2019

2. Residuals after subtraction

Beyond the 5-σ contrast curve for non-Gaussian noise

Jensen-Clem et al., 2017

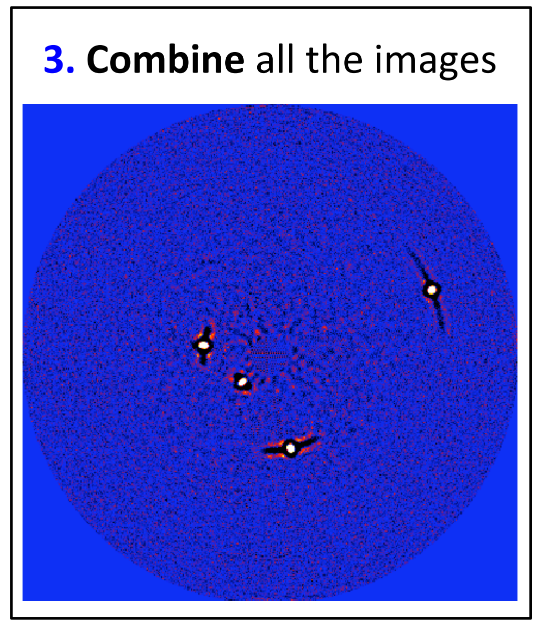
Bonse et al., 2022



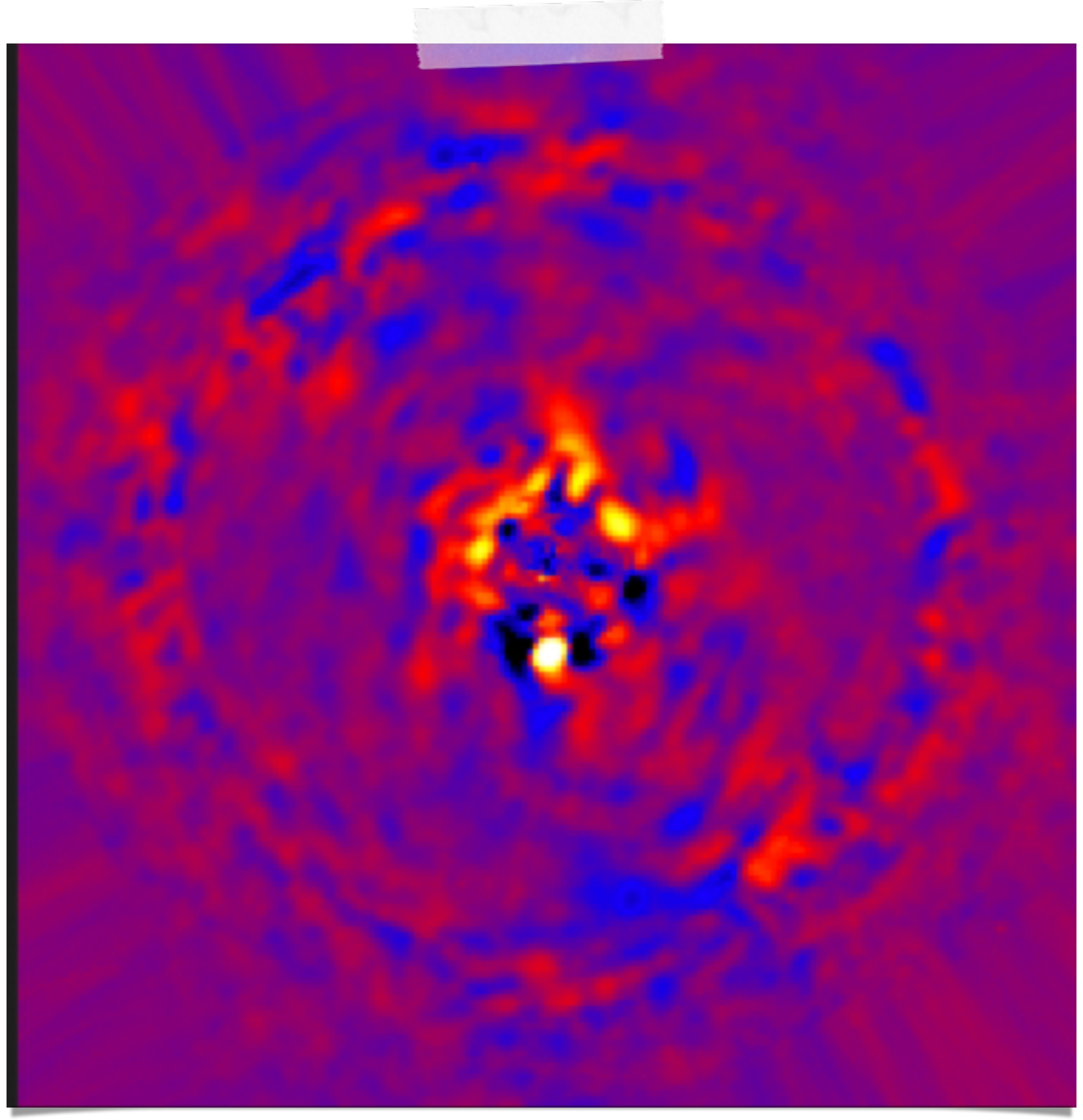
It is essential to have a realistic estimate of the distribution of the residual noise

Confidence level

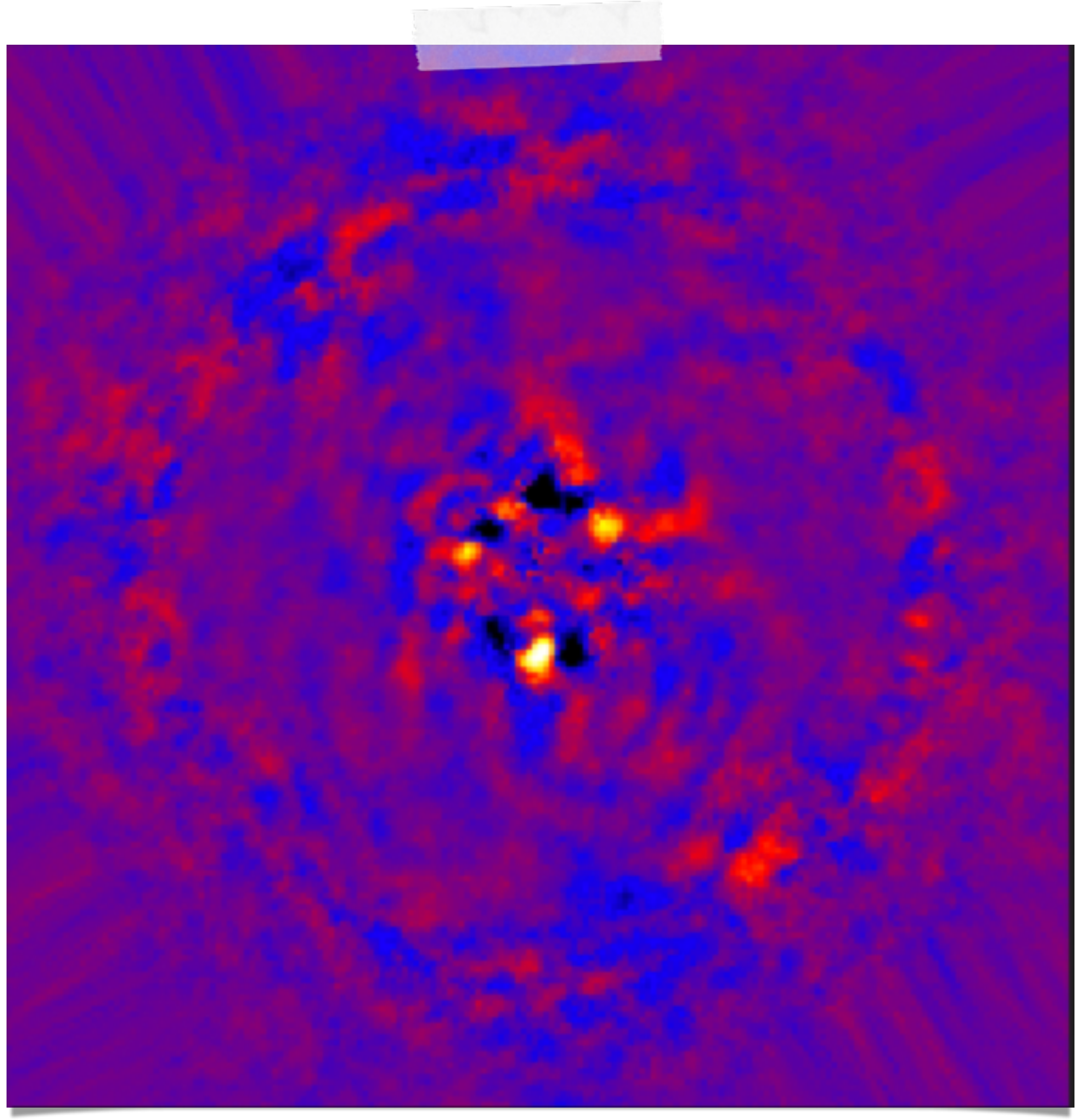
3. Combining the images



Mean combination

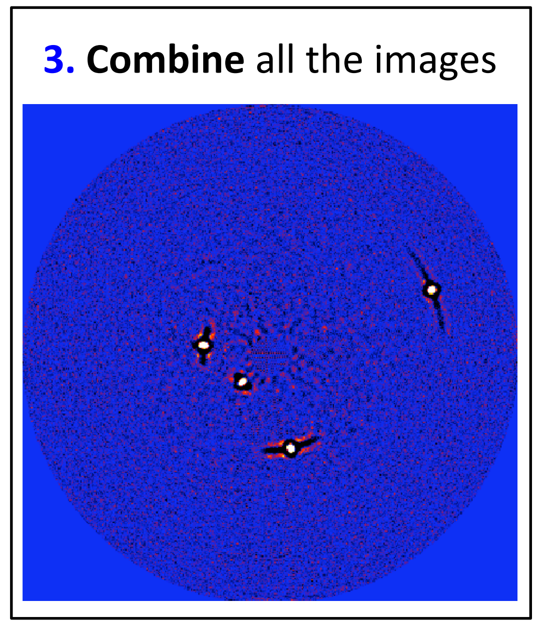


Median combination



There is no formal proof that one is better than the other…

3. Combining the images



Noise weighted combination

Bottom et al., 2017

Optimal weight $F_{opt} = \frac{1}{\sum_i \frac{1}{\sigma_i^2}} \sum_i \frac{F_i}{\sigma_i^2}$

STIM map

Pairet et al., 2019

$$STIM = \frac{\hat{\mu}_g}{\hat{\sigma}_g}$$

Normalization
factors to optimize SNR

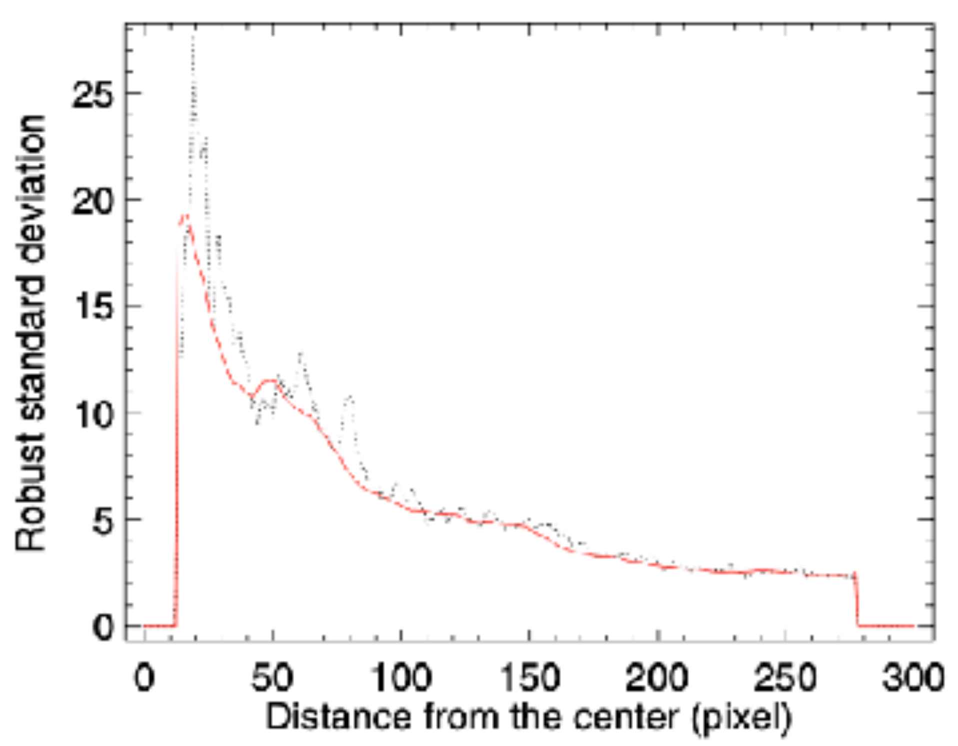
Statistical testing

Mawet et al., 2014

Accounting for small-sample statistics

Empirical normalisation

Cantalloube et al., 2015



Similar to adapting the threshold

Multinest approach

Golomb et al., 2020

Nested sampling to compute the evidence for H0 and H1.

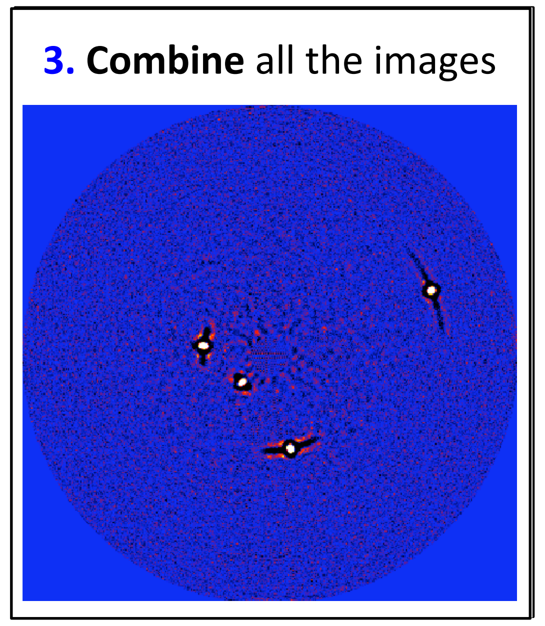
SNAP approach

Thompson & Marois, 2021

Optimisation of the S/N ratio

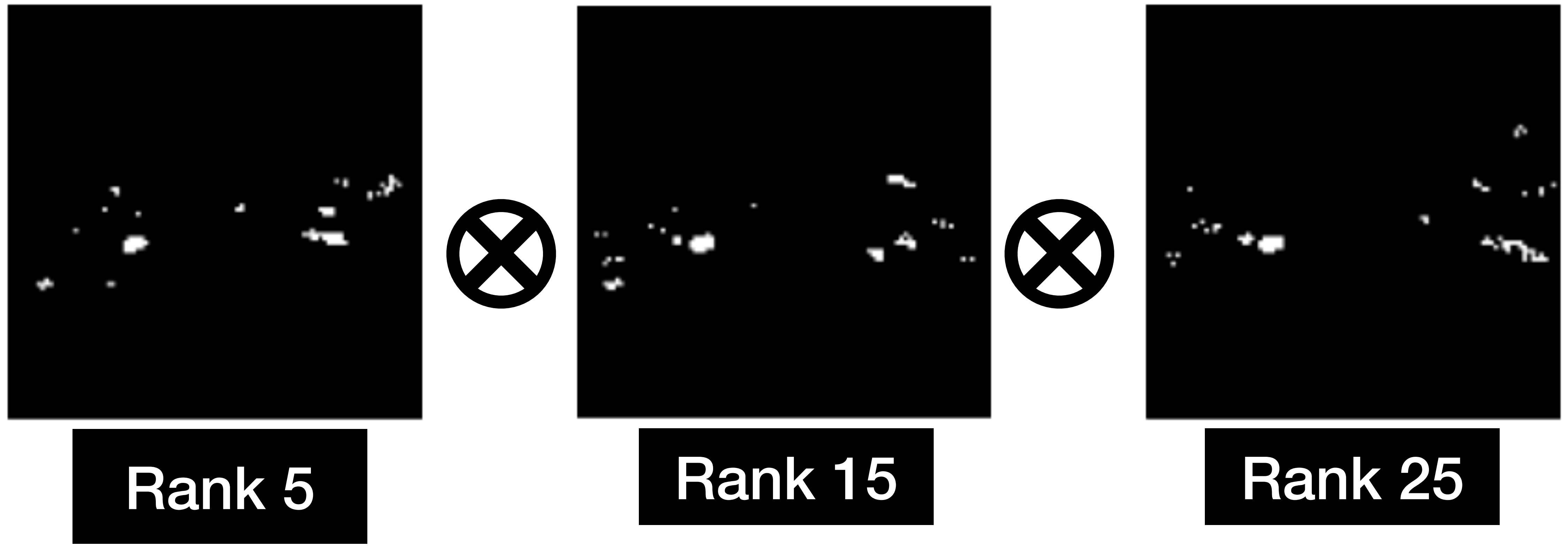
Balance noise reduction
vs. self-subtraction

3. Combining the images



STIM Largest Intensity Mask (SLIMask):

Pairet et al., 2021 (PhD thesis)



Mask to apply on STIM-maps:
Average location of the largest entries for a range of ranks

Regime Switching Model (RSM):

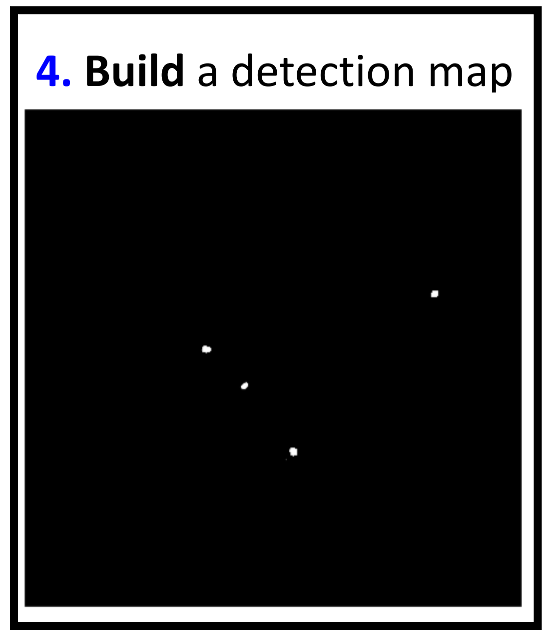
Dahlqvist et al., 2020, 2021ab, 2022

Build a time series in residual cubes:
Probability of H_1 at t, knowing state at t-1

Poster #39 Mariam Saballal

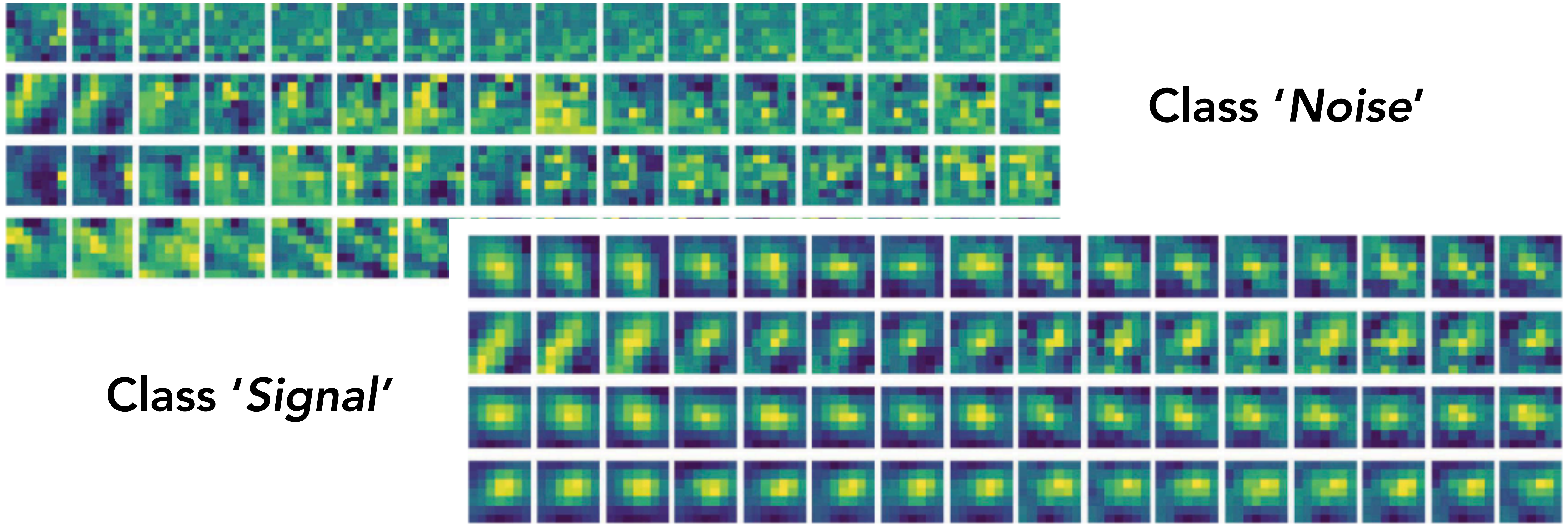
4. Detection map

Supervised binary classification



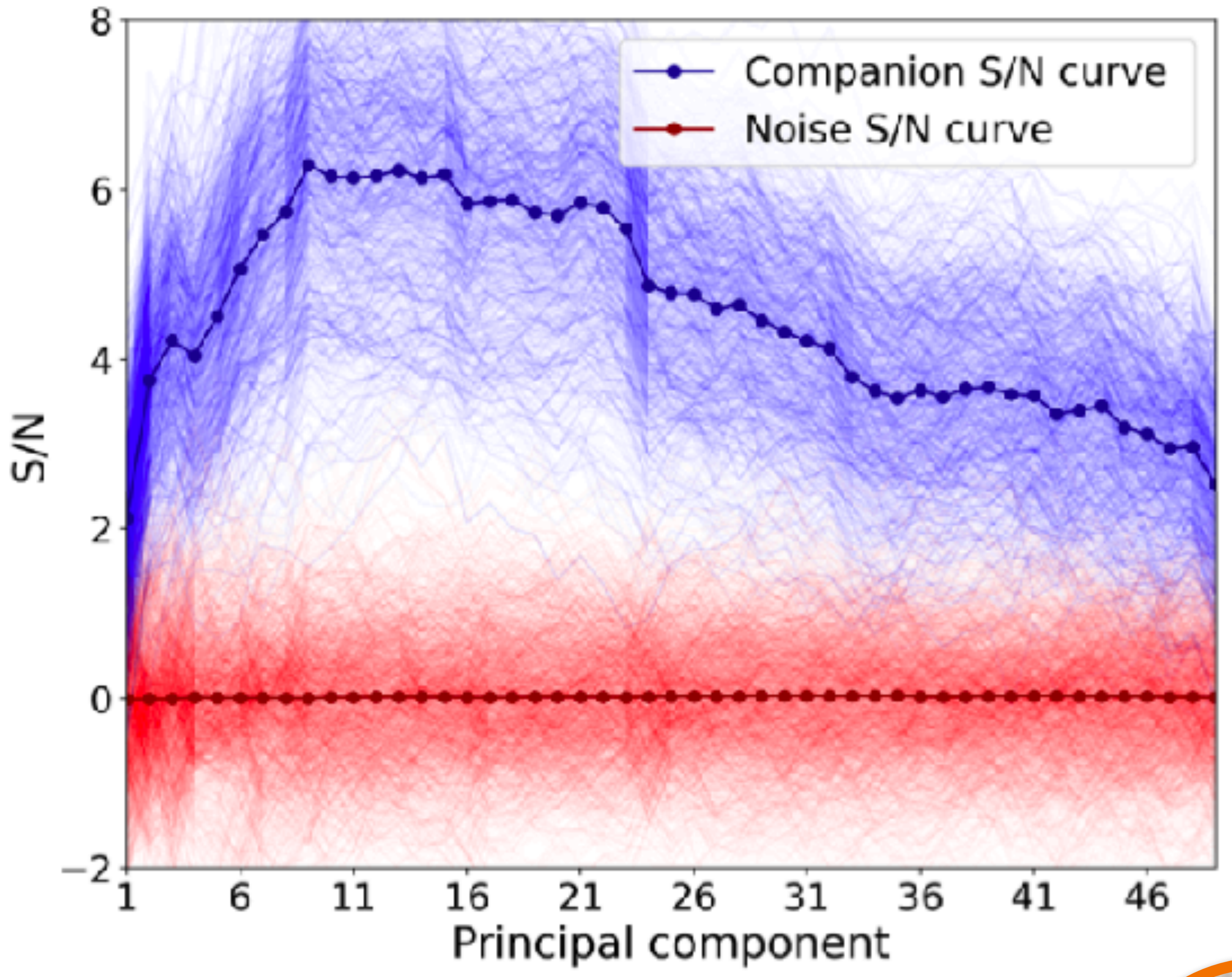
Supervised exOplanet detection via Direct Imaging with deep Neural Networks (SODINN)

Binary Classifier after a PCA subtraction



Noise Adaptative SODDIN (NA-SODINN)

Add SNR curves to support the training process



Gomez Gonzalez et al., 2018

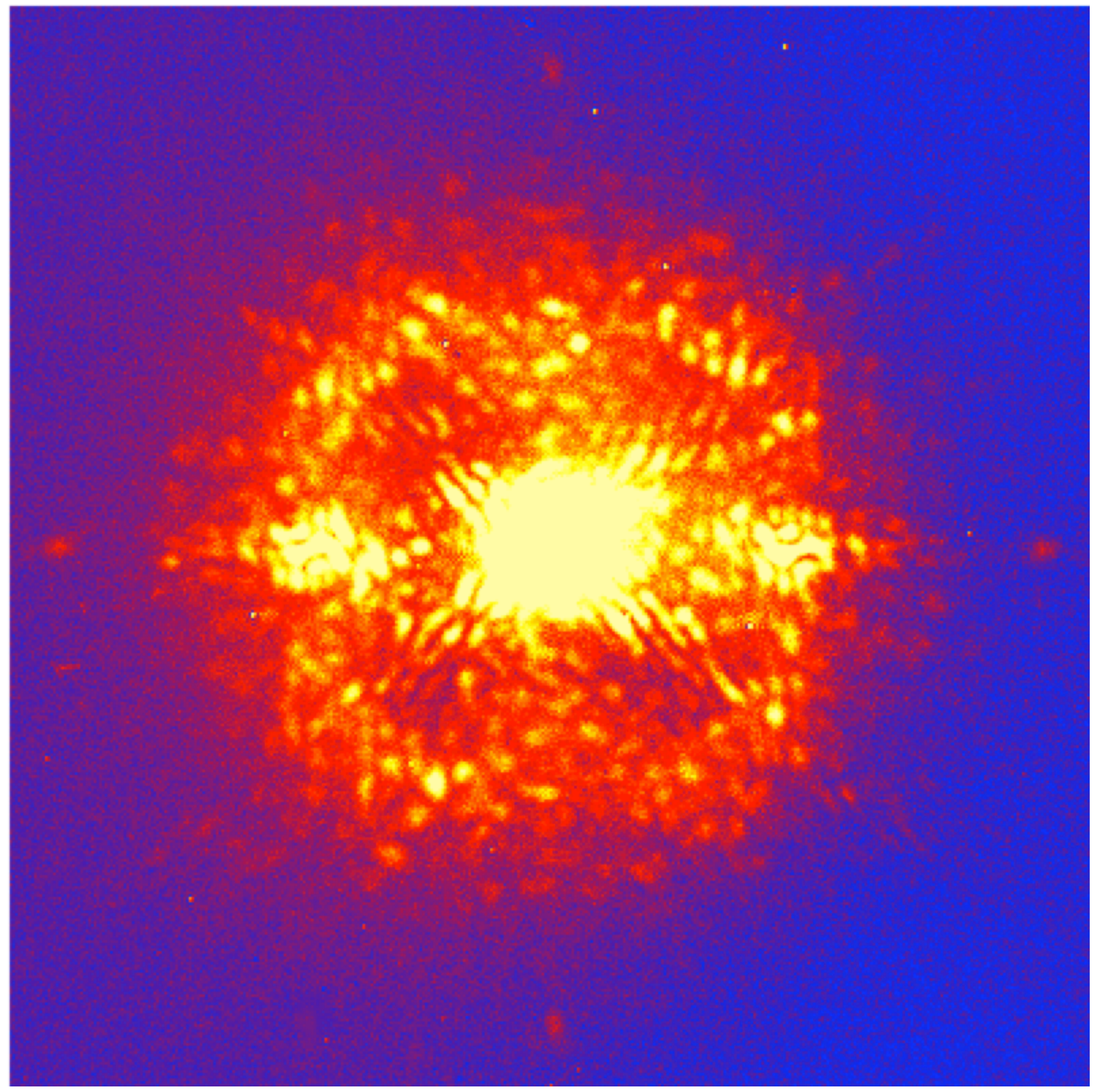
Cantero Mitijans et al., 2023

Characterization of the point-sources

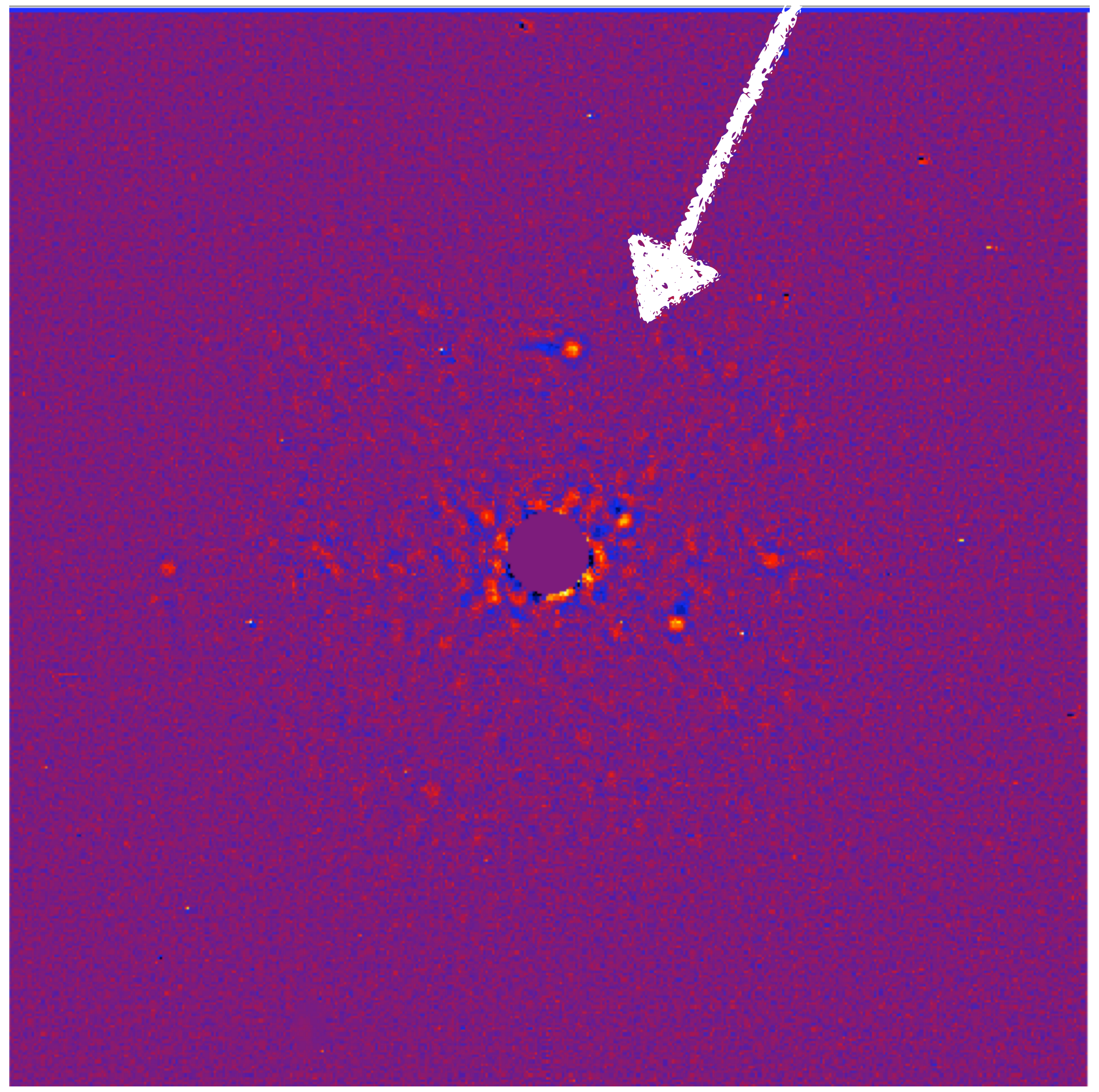
Forward Modeling
(inclusing assumption on noise distribution)

Negative Fake Companion
injection (NEGFC) + minimization

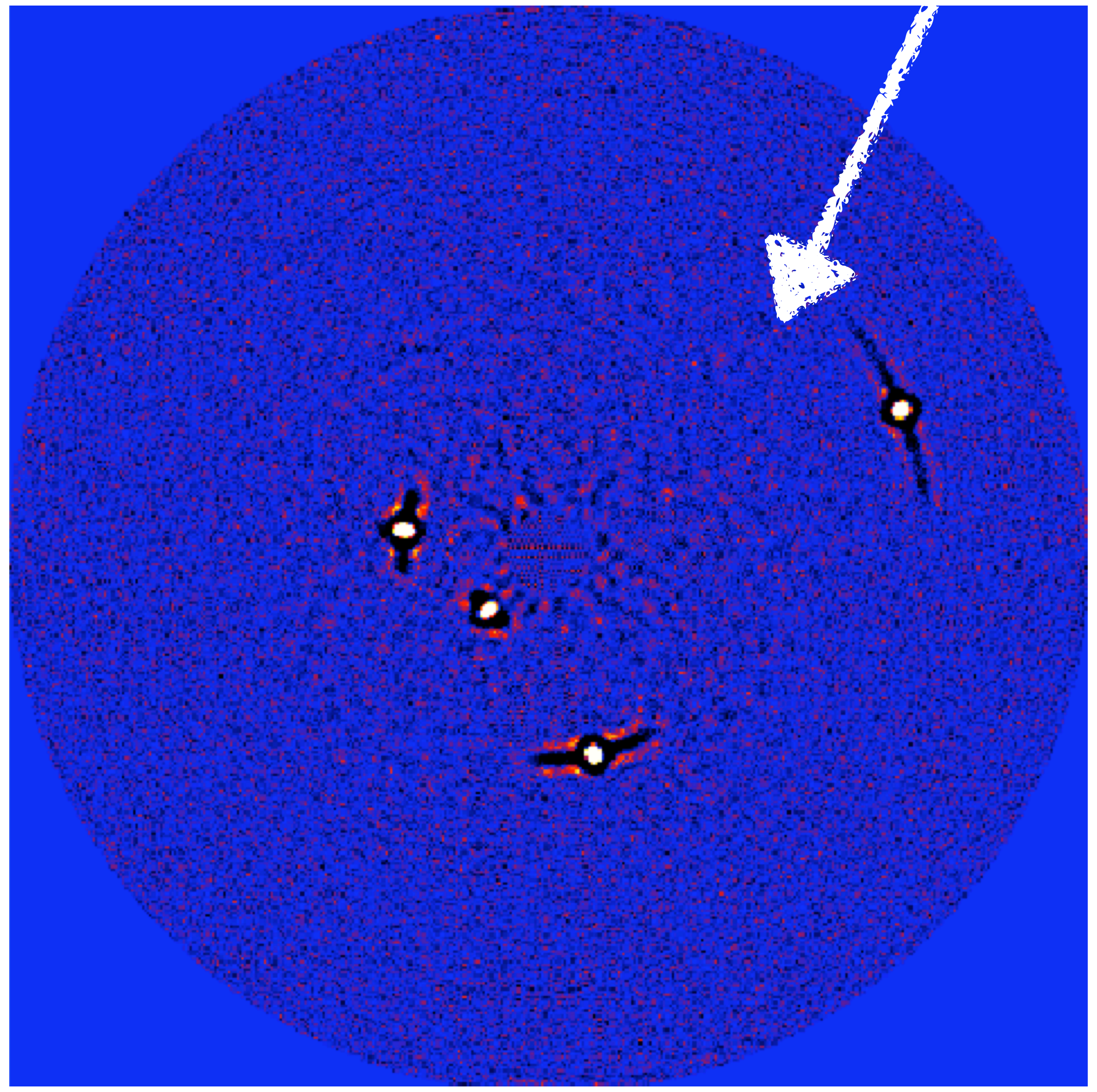
1. Estimate the star image



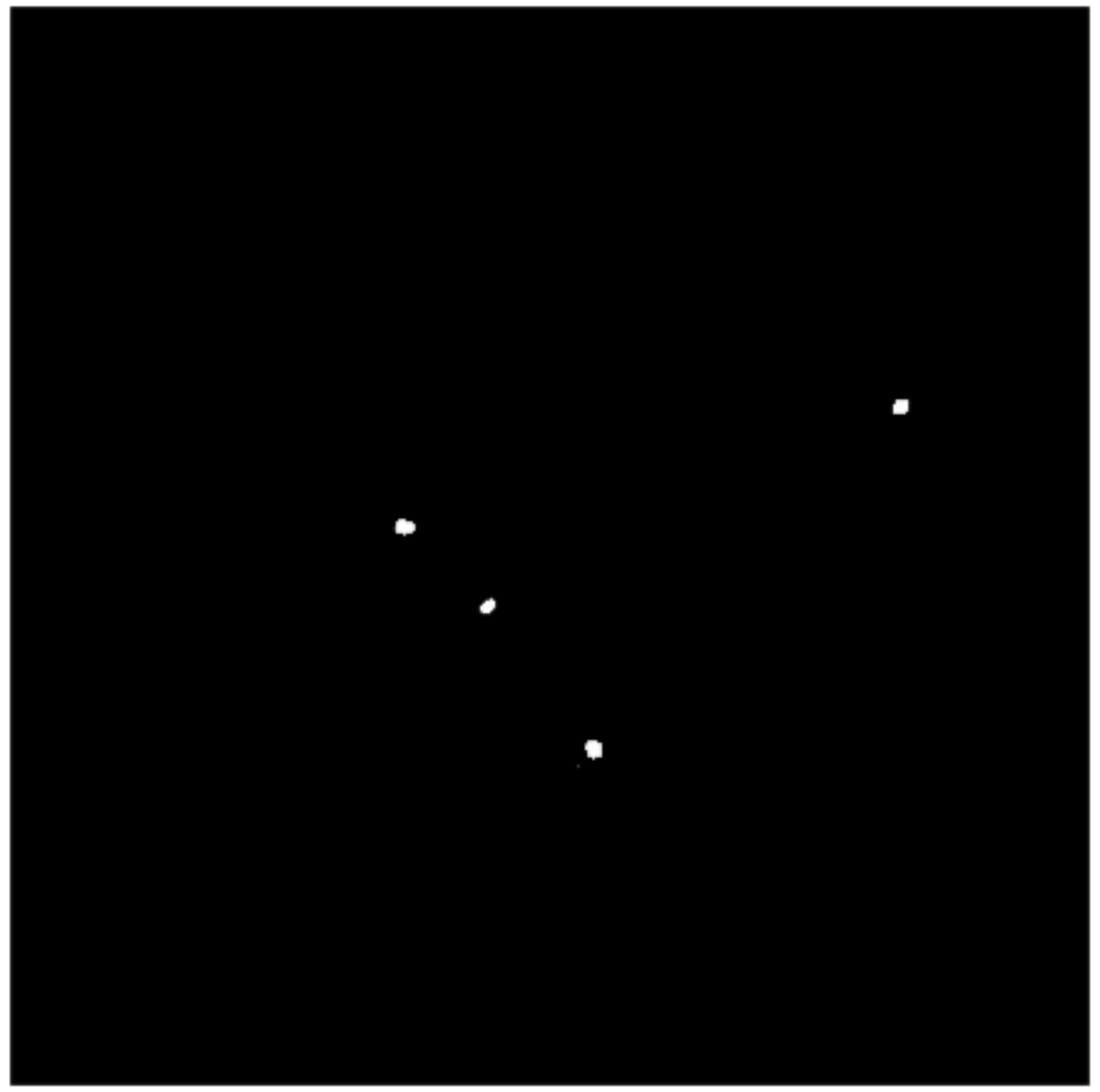
2. Subtract it to the image



3. Combine all the images



4. Build a detection map



direct SNR map estimate

$\mathbf{SNR = f_p/\sigma(f_p)}$

Forward Modeling of the planetary signal

Basic concept

Also called "inverse problem" or "Match Filtering" approach

ANDROMEDA

Mugnier et al., 2009

Cantalloube et al., 2015

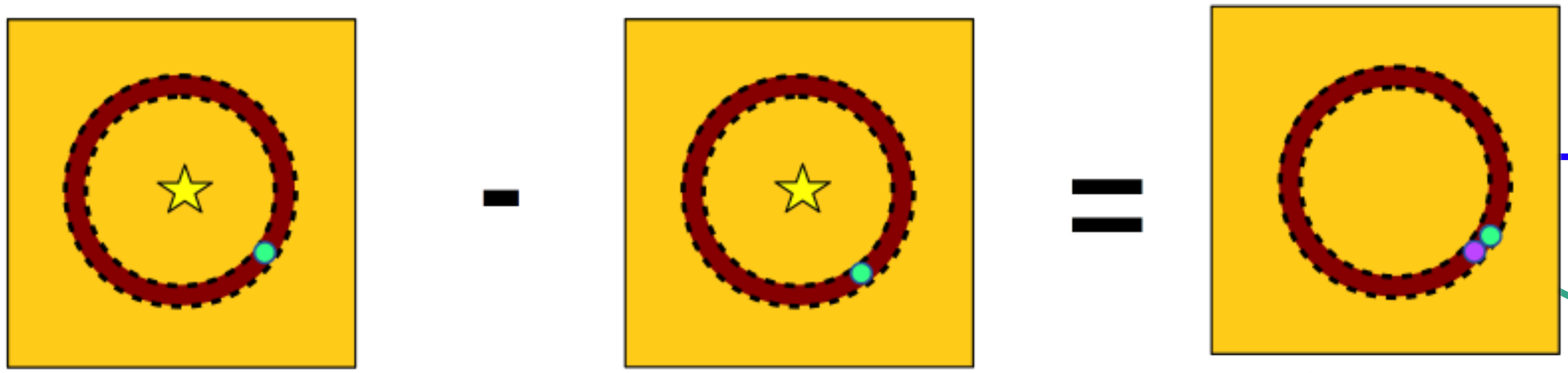
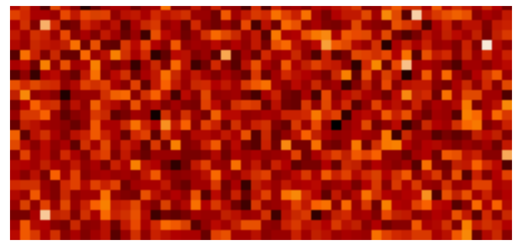


Image at t_1 - **Image at t_2** = **Differential image Δ**

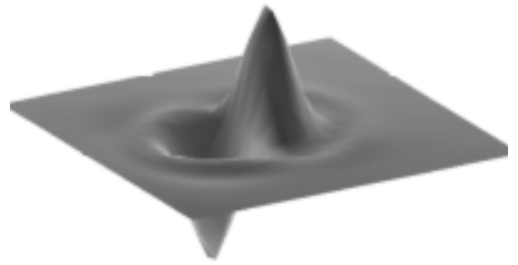
→ *smallest time interval (t_1-t_2)*

→ *sufficient signal distance (θ_1-θ_2)*

H_0: model of the residuals
White, Gaussian



H_1: model planet signature
Subtraction of 2 off-axis PSF



Maximum likelihood estimation:

$$L(r_0, a) \propto exp\left(-\frac{1}{2}\left\|\frac{\Delta(r,k) - a\, p(r,k;r_0)}{\sigma_\Delta(r)}\right\|_2^2\right)$$

Maximisation over:

- planet flux, a
- planet position, r_0

There is an analytical solution !

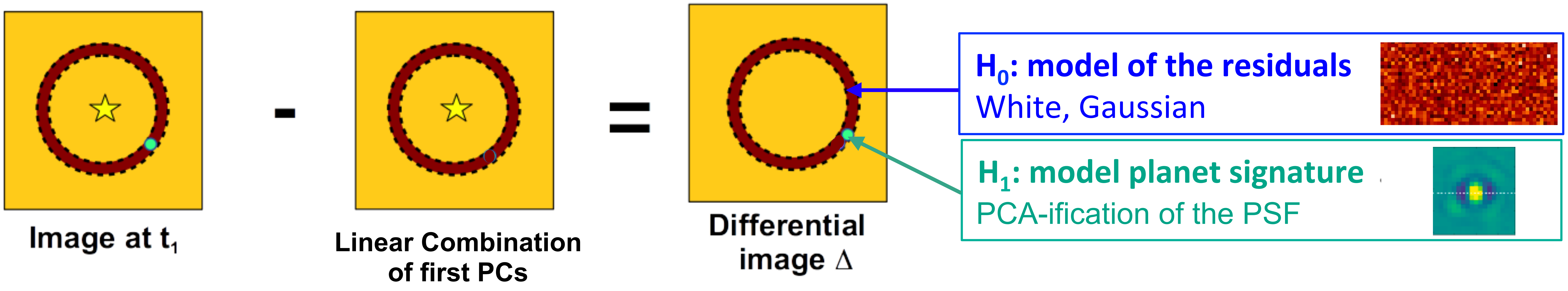
Forward Modeling of the planetary signal

Better subtraction ?

FMMF

Pueyo et al., 2016

Ruffio et al., 2017



Maximum likelihood estimation:

$$L(r_0, a) \propto exp\left(-\frac{1}{2}\left\|\frac{\Delta(r,k) - a\,p(r,k;r_0)}{\sigma_\Delta(r)}\right\|_2^2\right)$$

Maximisation over:

- planet flux, a
- planet position, r_0

There is an analytical solution !

Forward Modeling of the planetary signal

Better differential residuals model ?

r-ANDROMEDA

Cantalloube et al., subm. in 2019

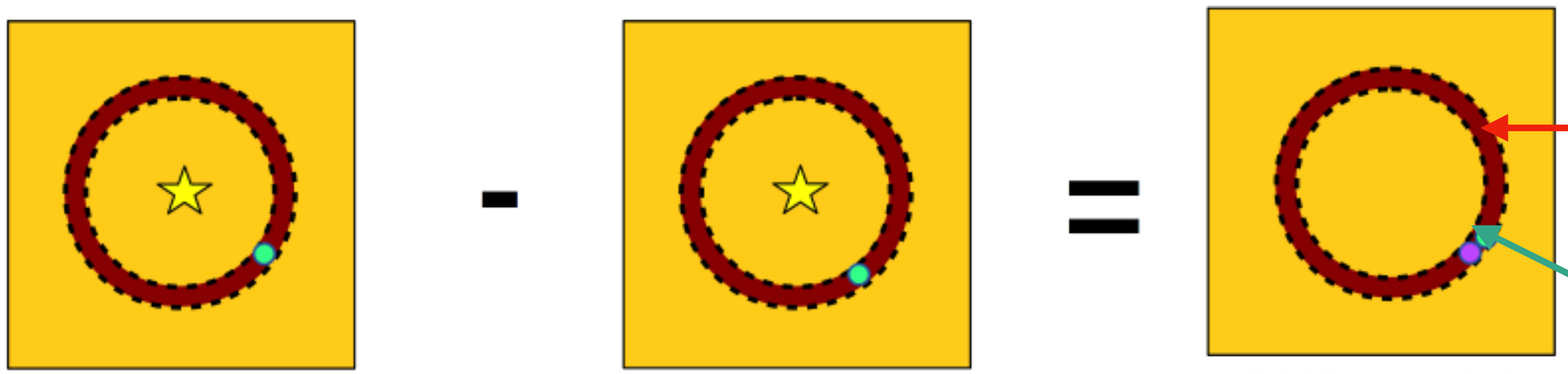
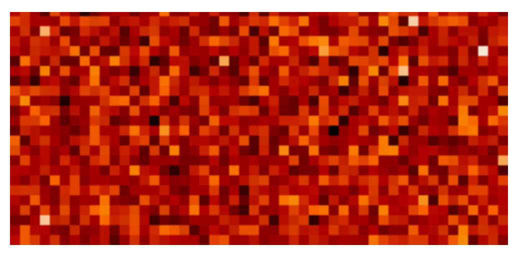
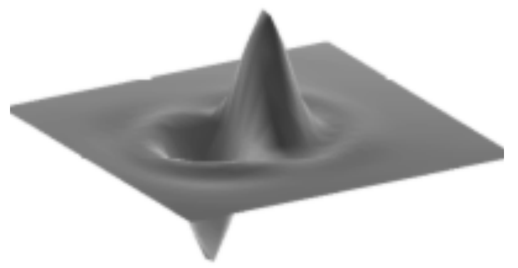


Image at t_1 - Image at t_2 = Differential image Δ

H_0: model of the residuals
White, Laplacian



H_1: model planet signature
Subtraction of 2 off-axis PSF



Maximum likelihood estimation:

$$L(r_0, a) \propto exp\left(-\left\|\frac{\Delta(r,k) - a\,p(r,k;r_0)}{b_\Delta(r)}\right\|_1\right)$$

Maximisation over:

- planet flux, a
- planet position, r_0

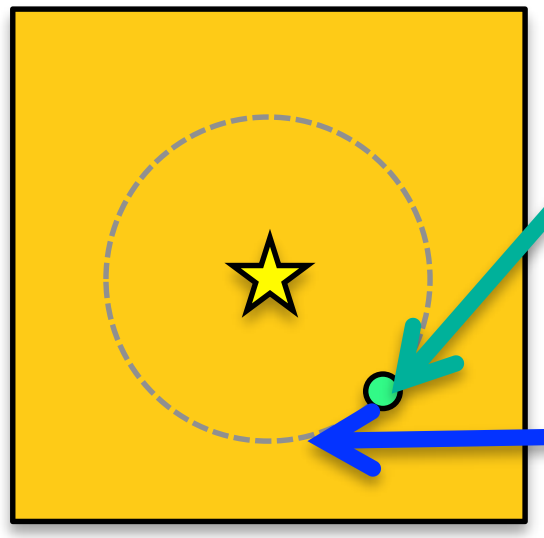
There is <u>no</u> analytical solution

Forward Modeling of the planetary signal

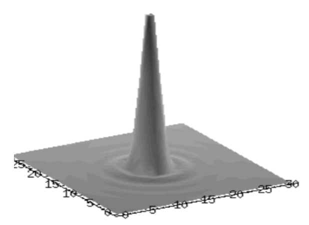
Not even subtraction ?

PAtch COvariance (PACO)

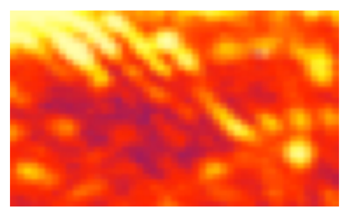
Flasseur et al., 2018



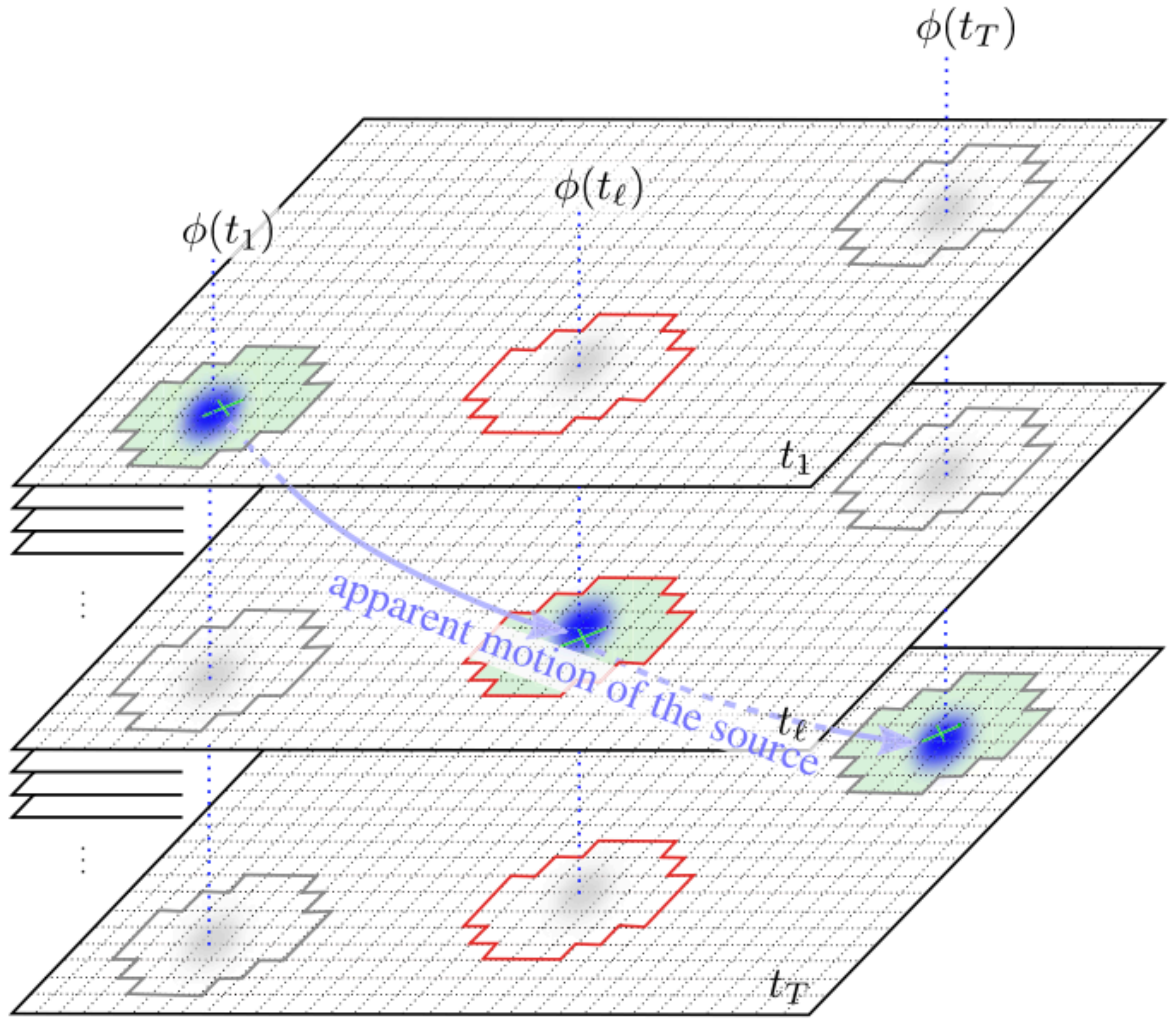
H_1: model planet signature h
off-axis PSF



H_0: model of the background f
Multivariate Gaussian



Data = α.h + f



Empirical mean and covariance on temporal patches

Maximum likelihood estimation with multivariate Gaussian noise:

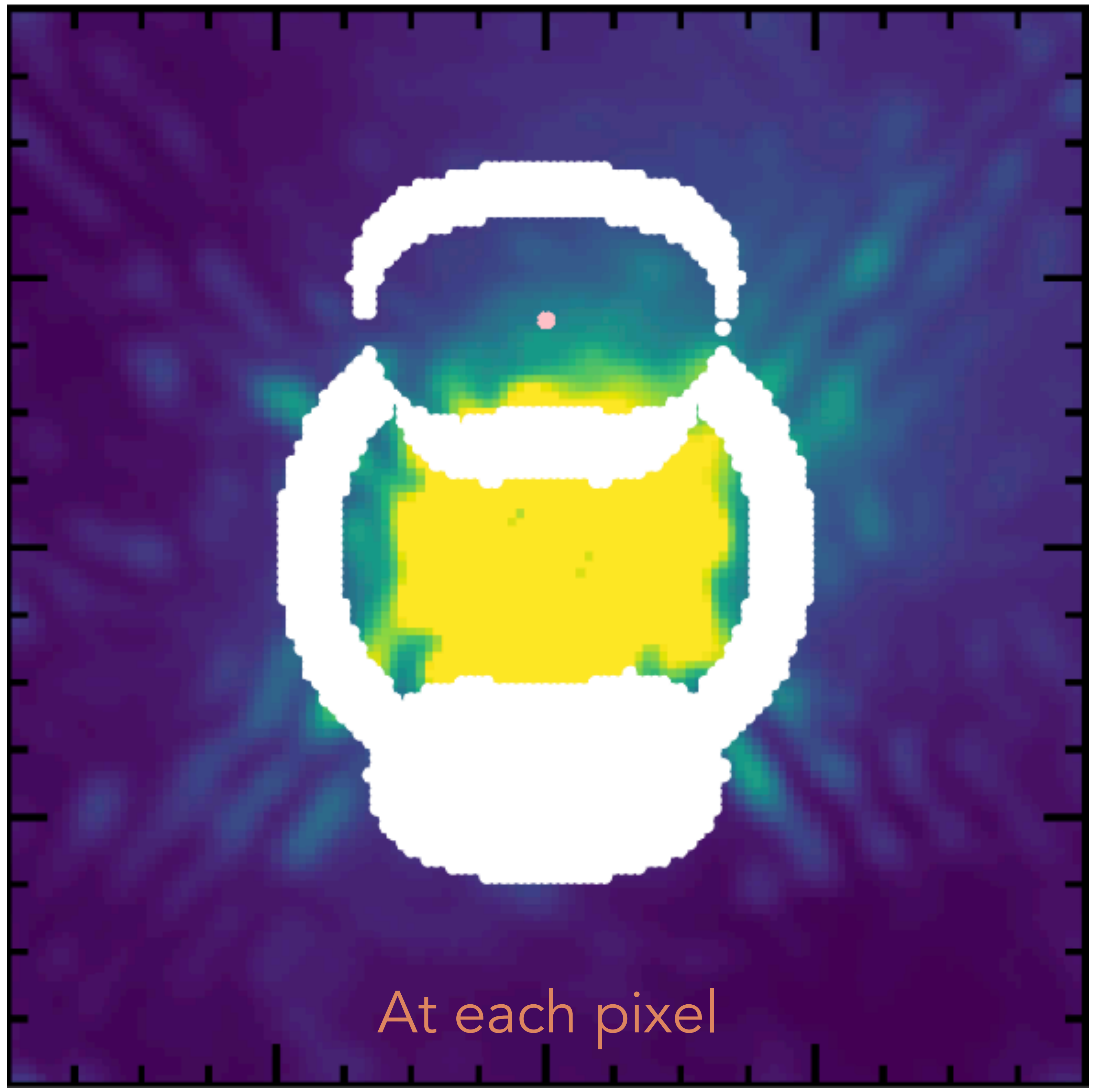
$$\mathrm{p}_f(\{\boldsymbol{f}_{\lfloor\phi_t\rceil,t}\}_{t=1:T}) = \prod_{t=1}^{T} \mathcal{N}\left(\boldsymbol{f}_{\lfloor\phi_t\rceil,t} \,\middle|\, \boldsymbol{m}_{\lfloor\phi_t\rceil}, \mathbf{C}_{\lfloor\phi_t\rceil}\right)$$

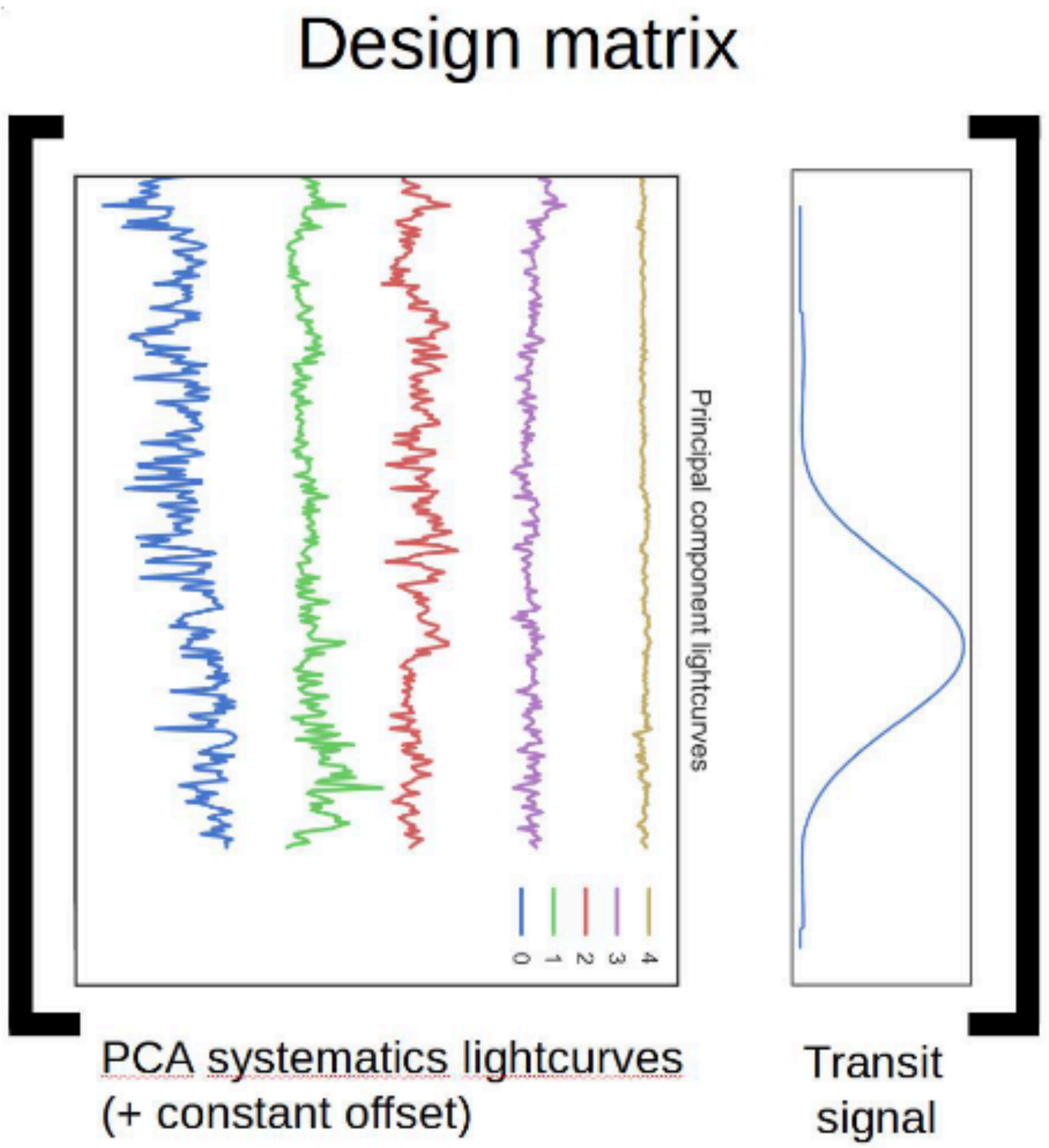
Forward Modeling

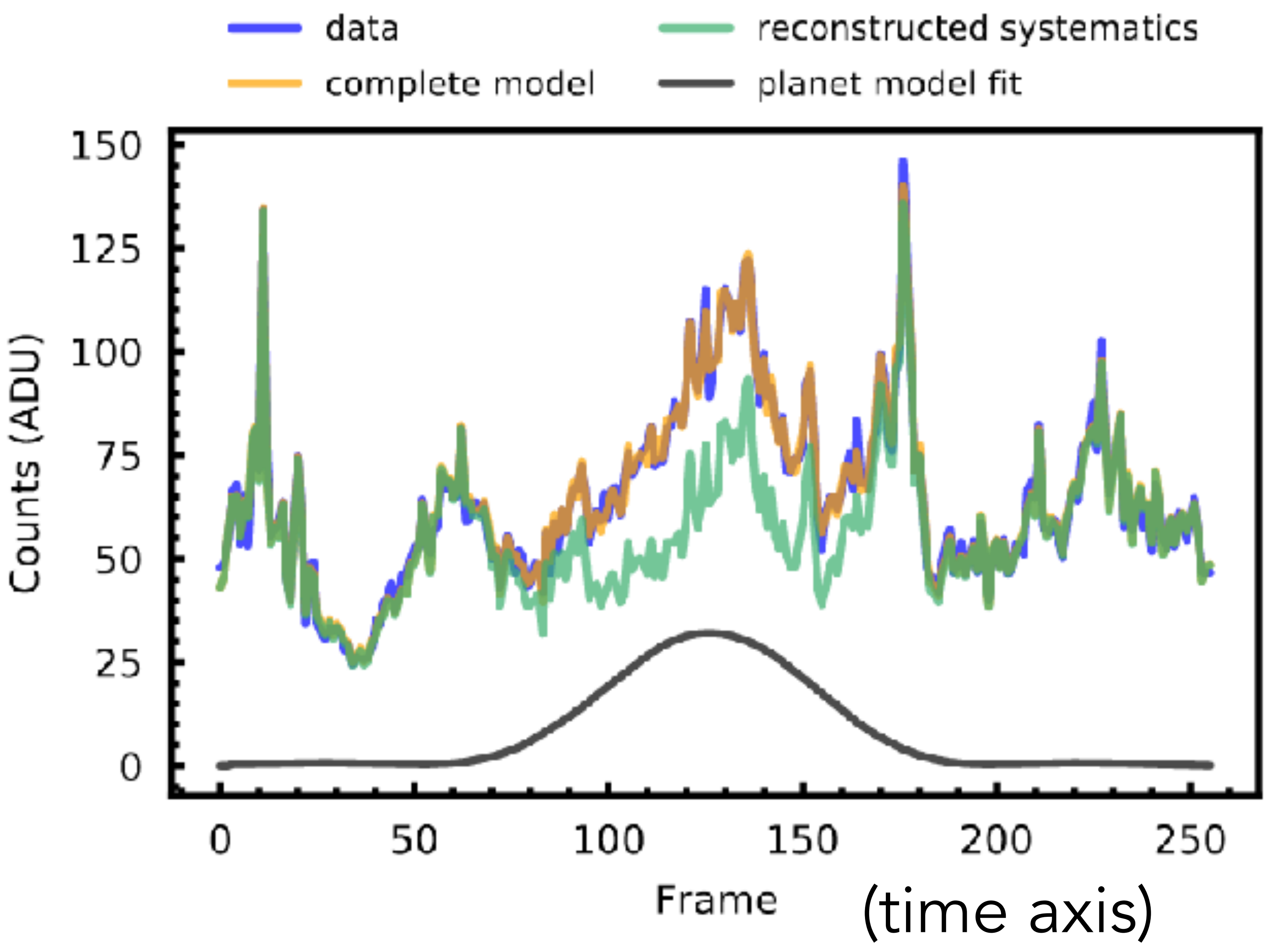
Samland et al., 2021

Temporal Reference Analysis of Planets (TRAP)

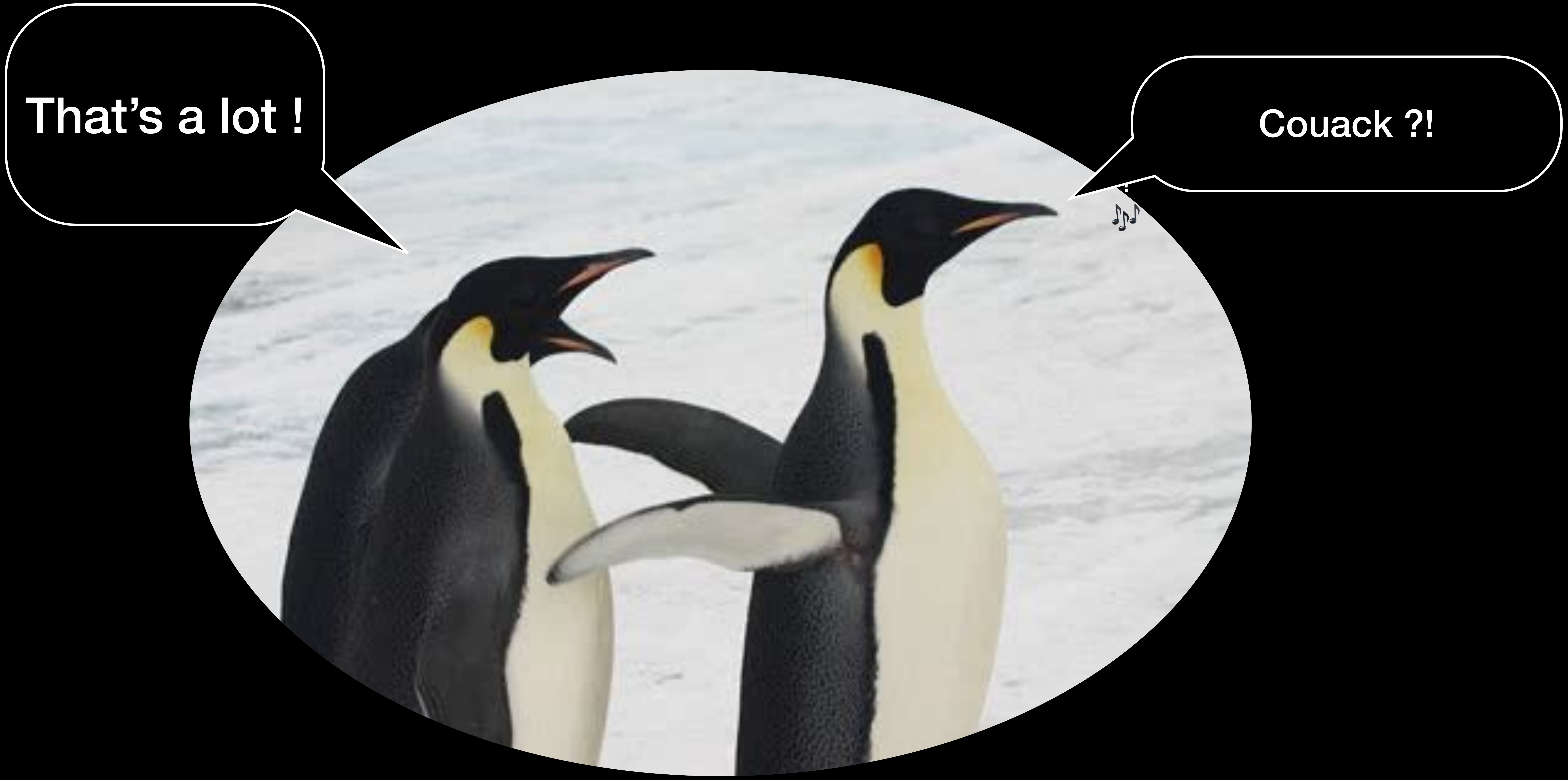
Temporal PCA model of the starlight
+ Forward Modeling







Penguin interlude



ADI-based post-processing techniques

Ground-based, point-source detection

classical ADI

Marois et al., 2006

Dou et al., 2015

Golomb et al., 2020

Mawet et al., 2014

LOCI

Lafrenière et al., 2007

Template-LOCI, *Marois et al., 2014*
Adaptive-LOCI, *Currie et al., 2012*
Matched-LOCI, *Wahhaj et al., 2015*
Damped-LOCI, *Pueyo et al., 2016*

PCA-based DI

Soummer et al., 2012

Amara & Quanz 2012

smart-PCA, *Absil et al., 2013*
LLGS, *Gonzalez et al., 2016*
NMF, *Ren et al., 2018*
Slim-mask, Pairet PhD 2020

Bottom et al., 2017

Thompson & Marois, 2021

Gomez Gonzalez et al., 2018

Dahlqvist et al., 2020, 2021ab, 2022

Gebhard et al., 2022

Stolker et al., pynpoint

Wang et al., pyKLIP

Christiaens et al., VIP

Pueyo et al., 2016

Ruffio et al., 2017

Samland et al., 2021

Pairet et al., 2019

Mugnier et al., 2009

Cantalloube et al., 2015

Bonse et al., 2024

Daglayan Sevim et al., 2022

Cantero-Mitijans et al., 2023

Flasseur et al., 2018

Le Coroller et al., 2015

Flasseur et al., 2019, 2020,2021, 2022, 2023, 2024

Exoplanet Imaging Data Challenge

a community-wide effort



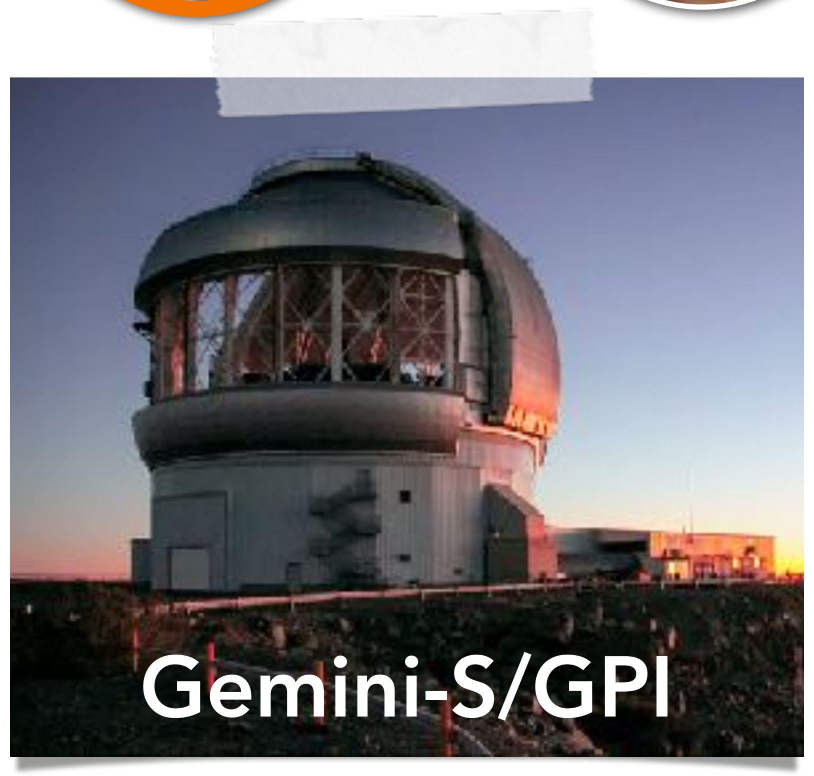


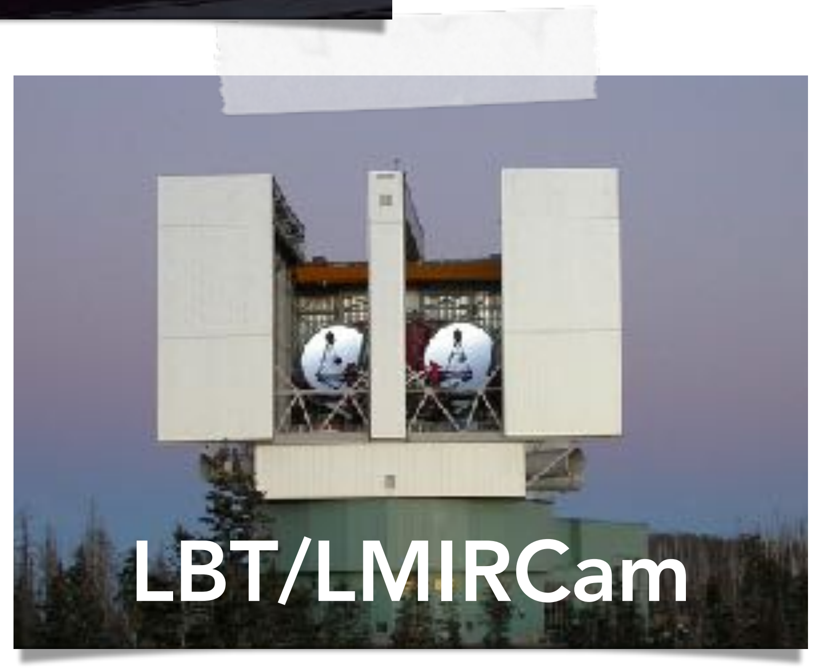







- Started in 2019 !
- First phase launched in Sept. 2019
- Workshop HCI post-processing, Berlin, Germany
- First phase closed in Oct. 2020
- Publication SPIE 2020
- Second phase (**characterization**) launched Apr. 2022
- Third phase (**disk imaging**) for ~2025
- Fourth phase (**high resolution spectroscopy**) for ~2026





https://exoplanet-imaging-challenge.github.io/

Exoplanet Imaging Data Challenge

Ressources

Website+

Data repository

Competition platform

EIDC website

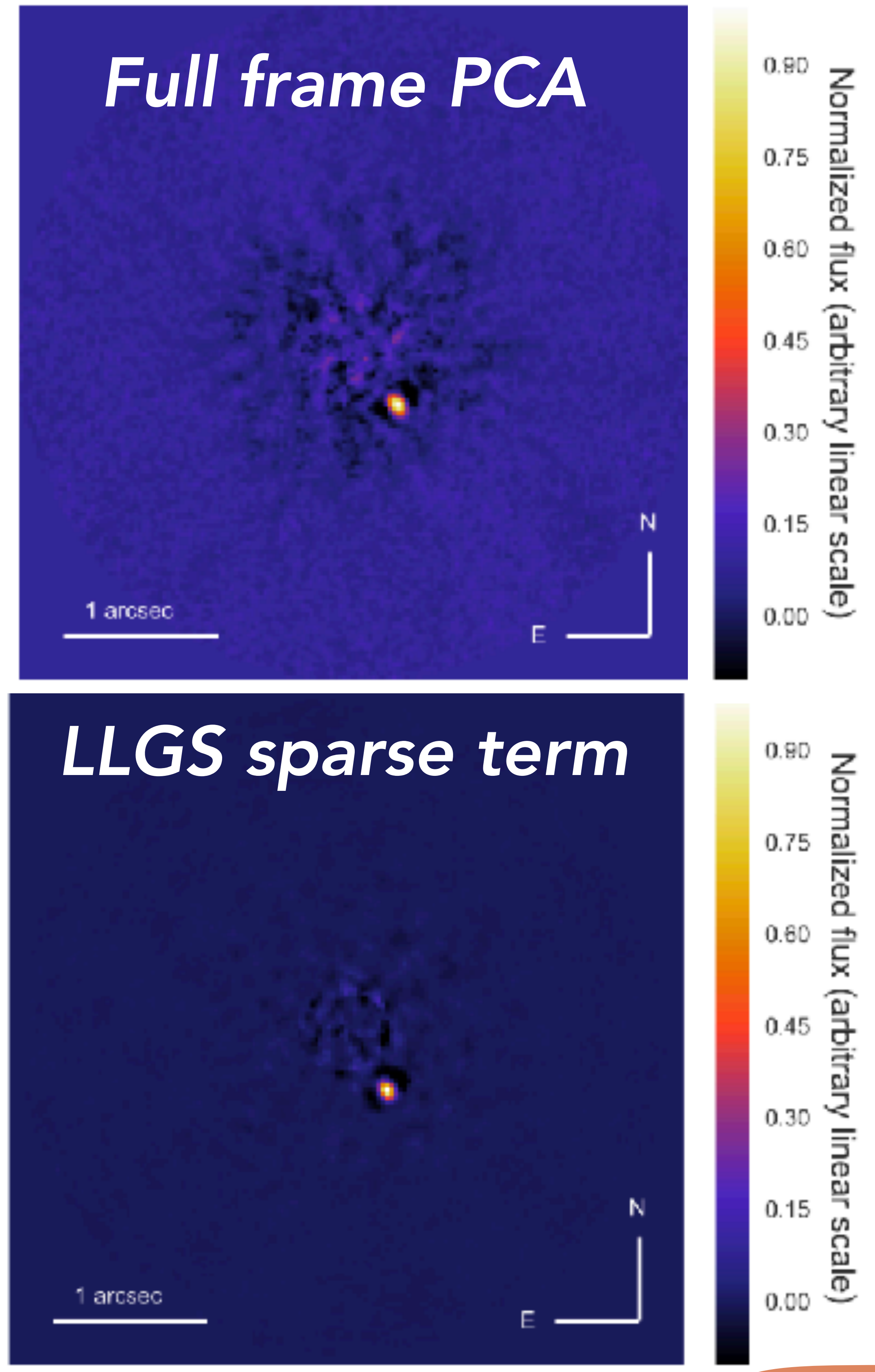
Cantalloube et al., SPIE 2020

Cantalloube et al., SPIE 2022

*Cantalloube et al., SPIE **2024***

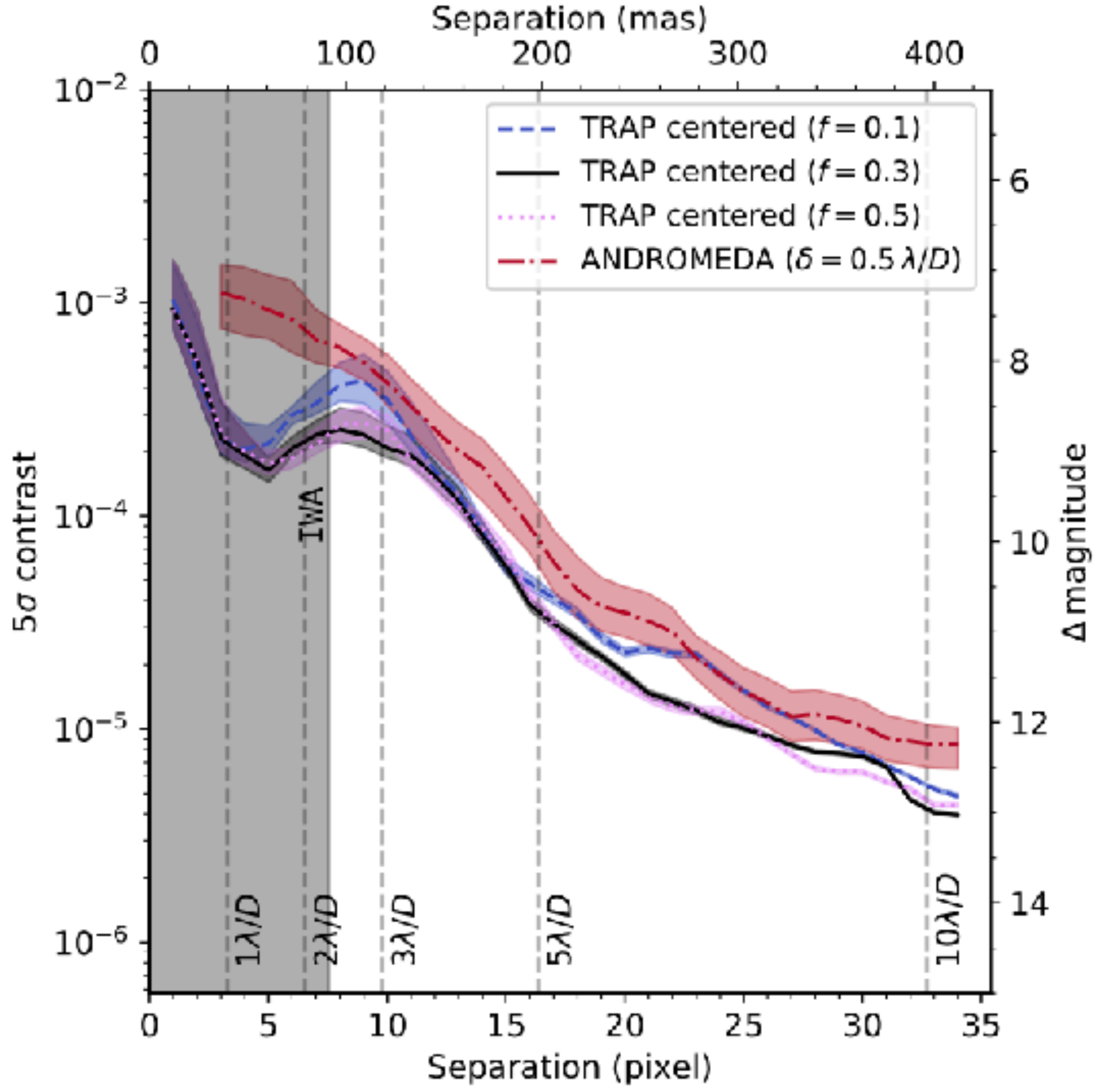
Post-processing techniques performance assessement

Images comparison
(Residual maps)



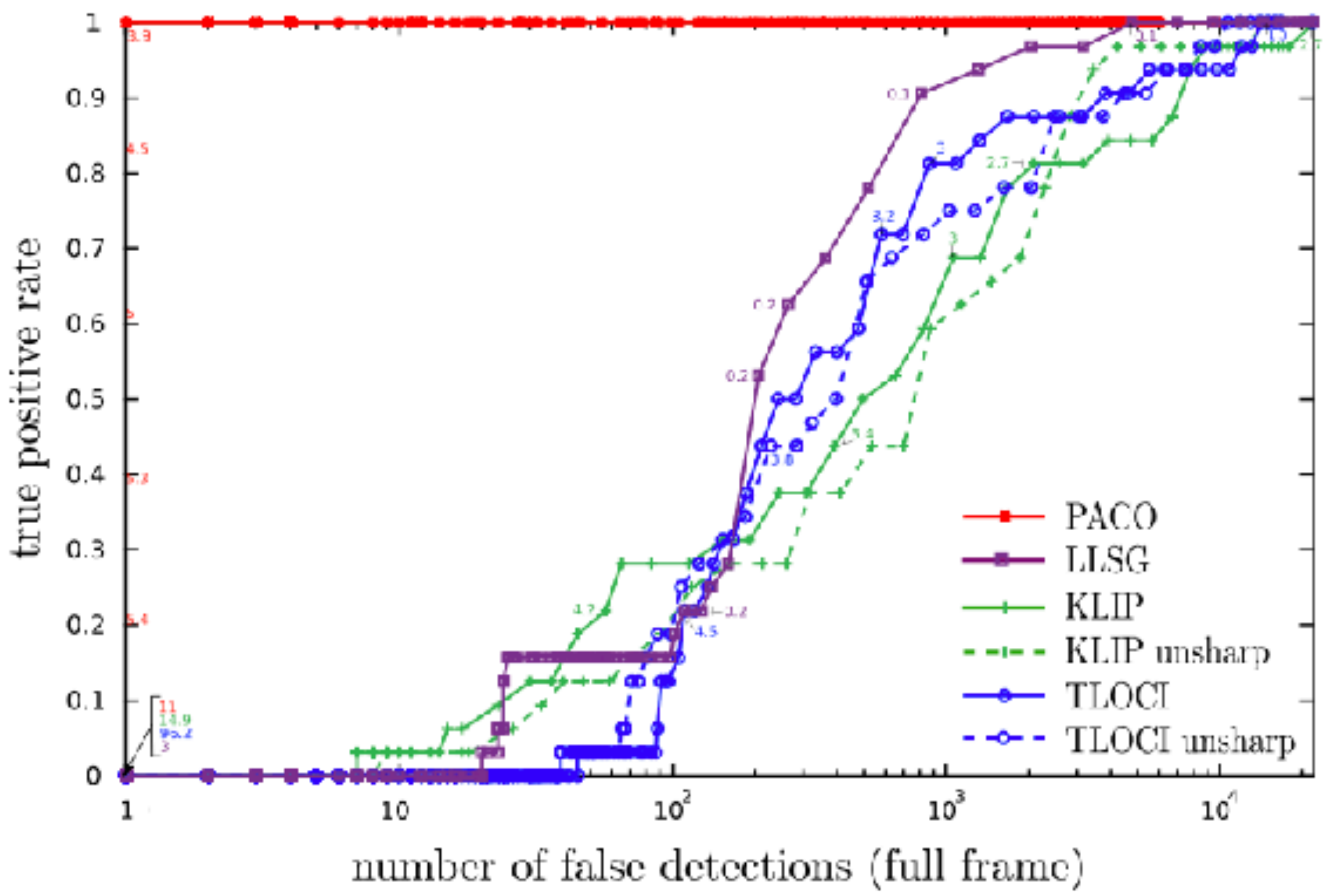
Gomez Gonzalez et al., 2016

Detection limits
(5-σ Contrast curves)



Samland et al., 2021

ROC curves
(Receiver Operating C)



Flasseur et al., 2018

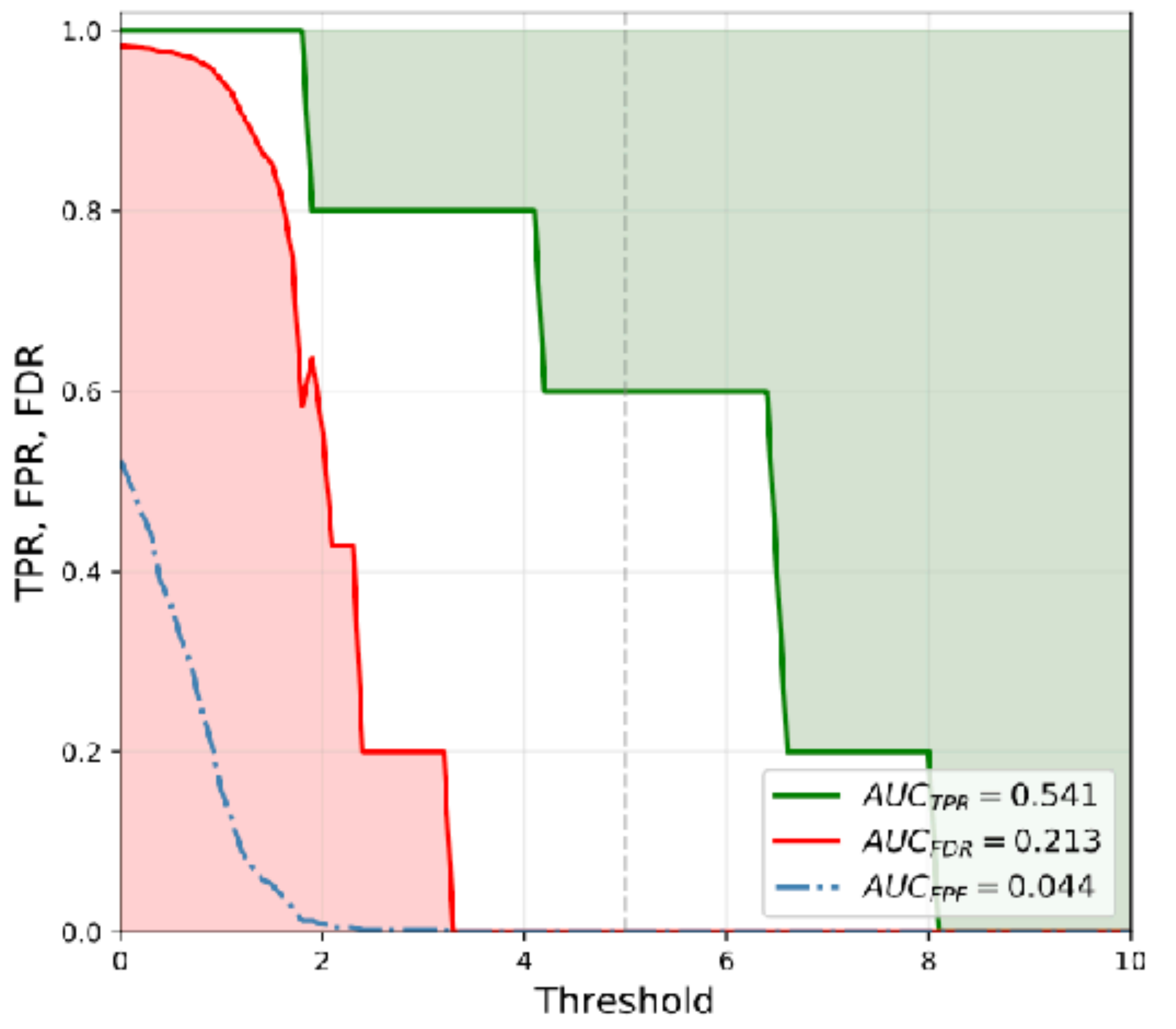
Detection map + threshold + posterior (spectro)-photometry

- Counting True and False positives

SADI-based detection

Metrics, SNR map, ranking

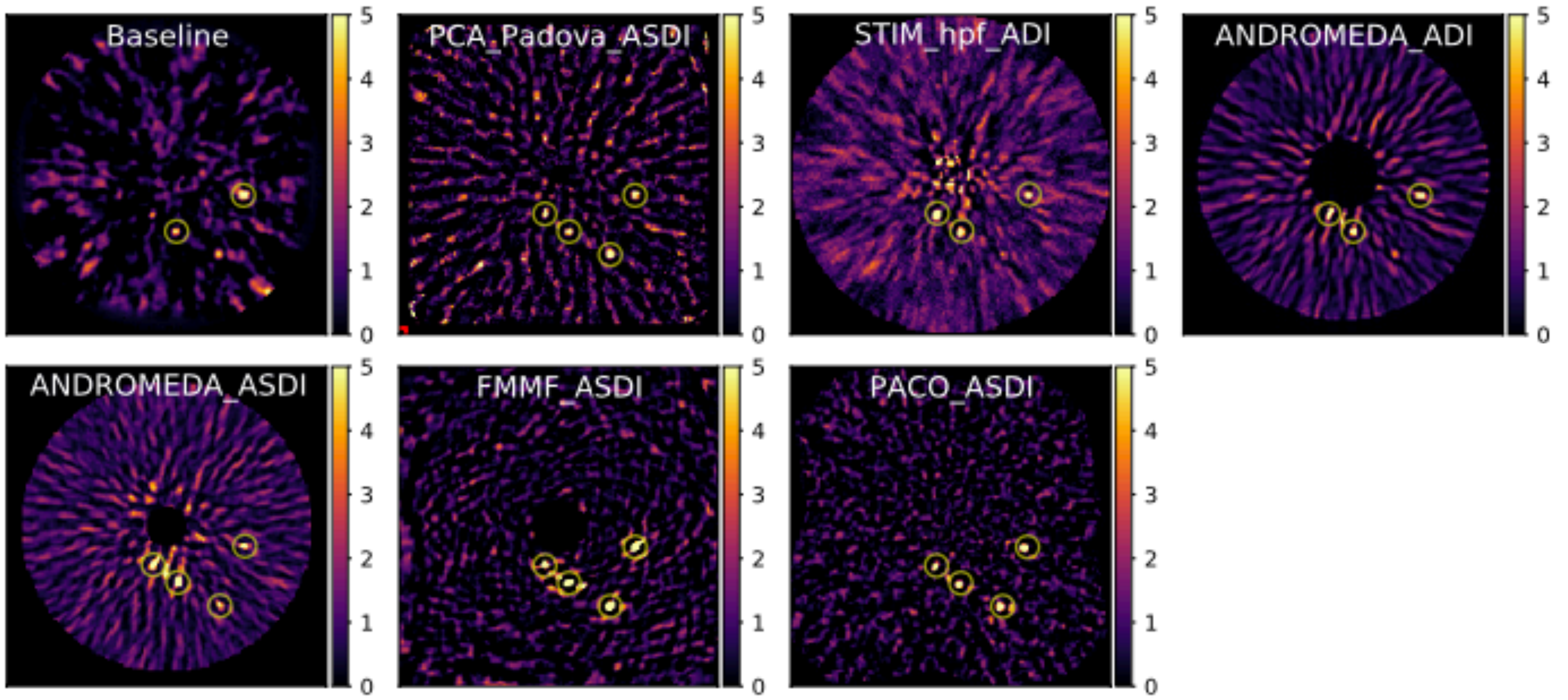
Metrics

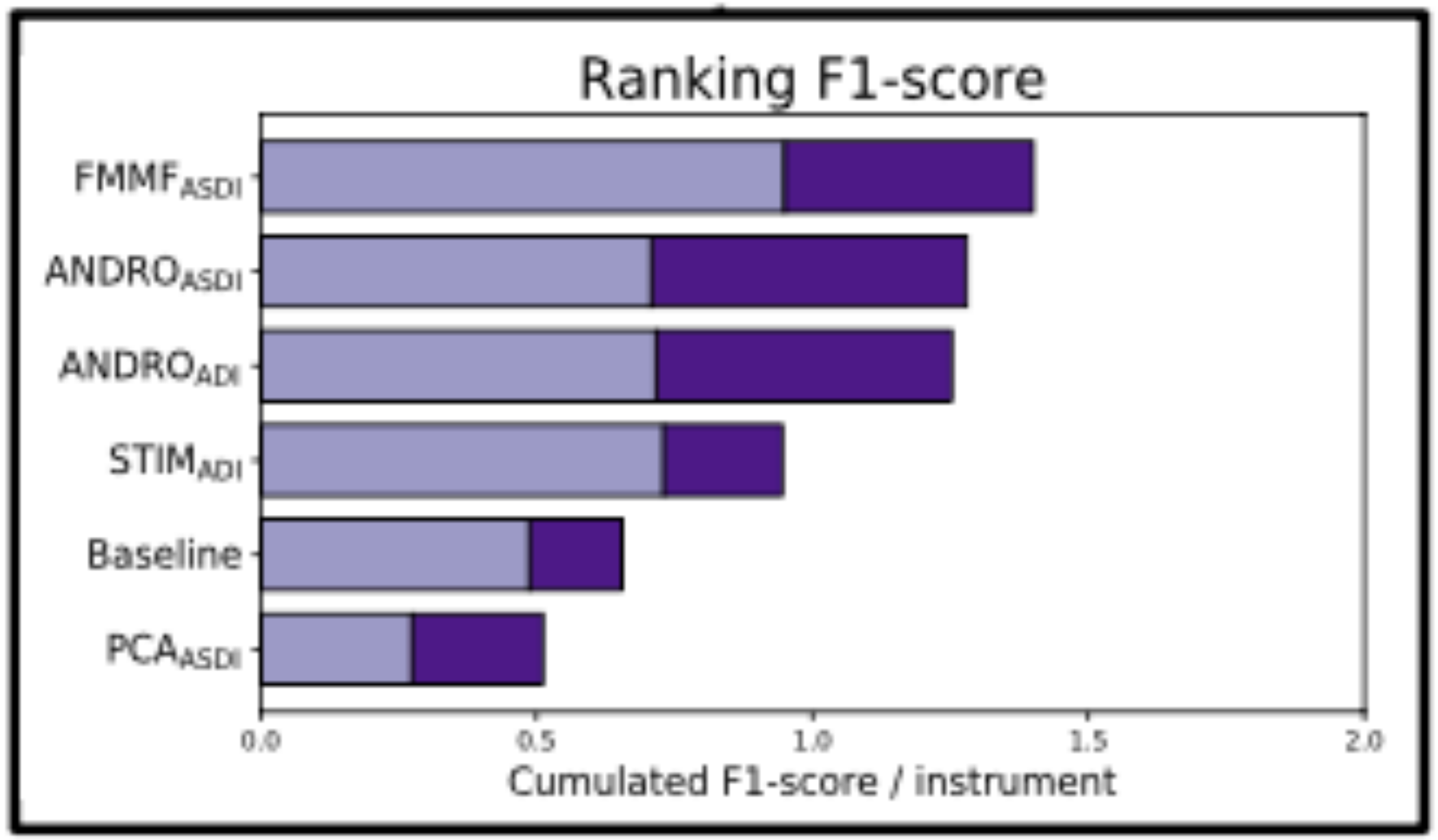


- True positive rate: TPR = TP/(TP+FN)
- False discovery rate: FDR = FP/(FP+TP)
- False positive rate: FPR = FP/(FP+TN)

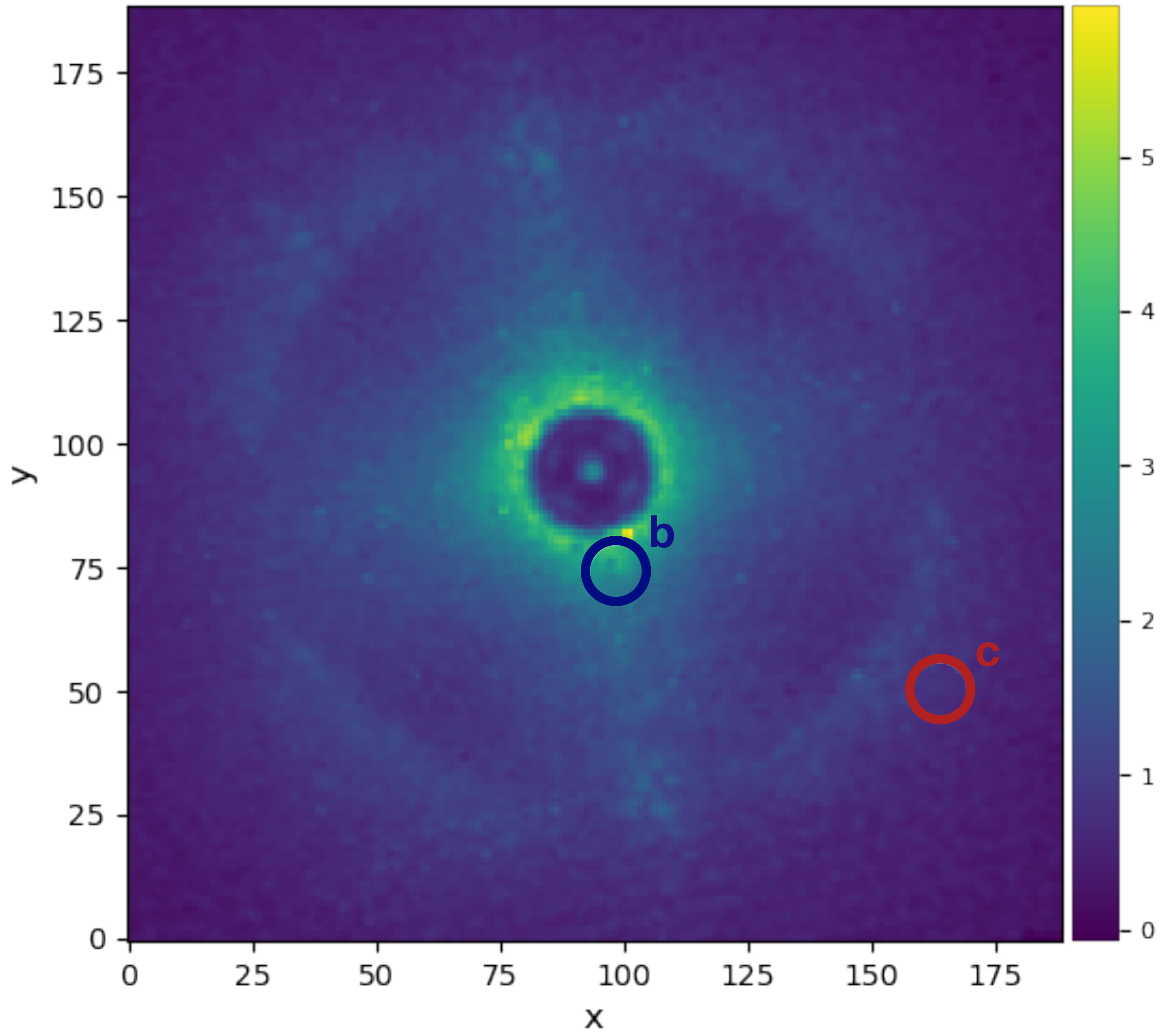
At the submitted threshold, we compute:

- F1-score = 2 TP / (2 TP+FP+FN)

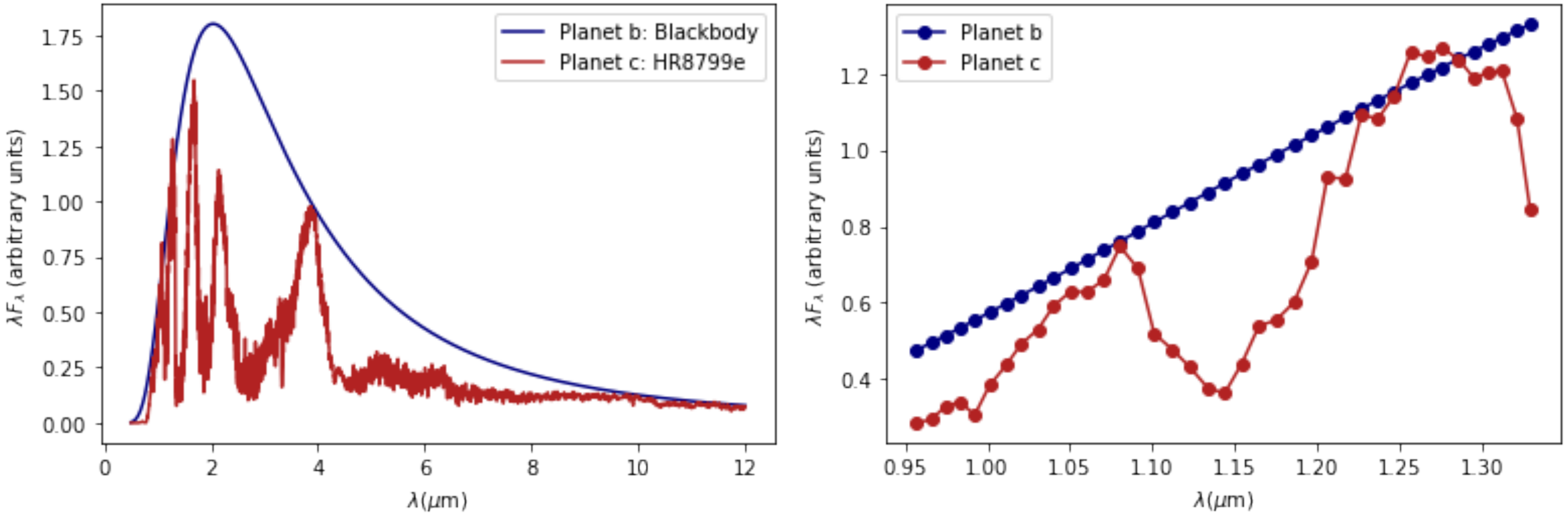




SADI-based characterization: Data type



Multispectral image cube



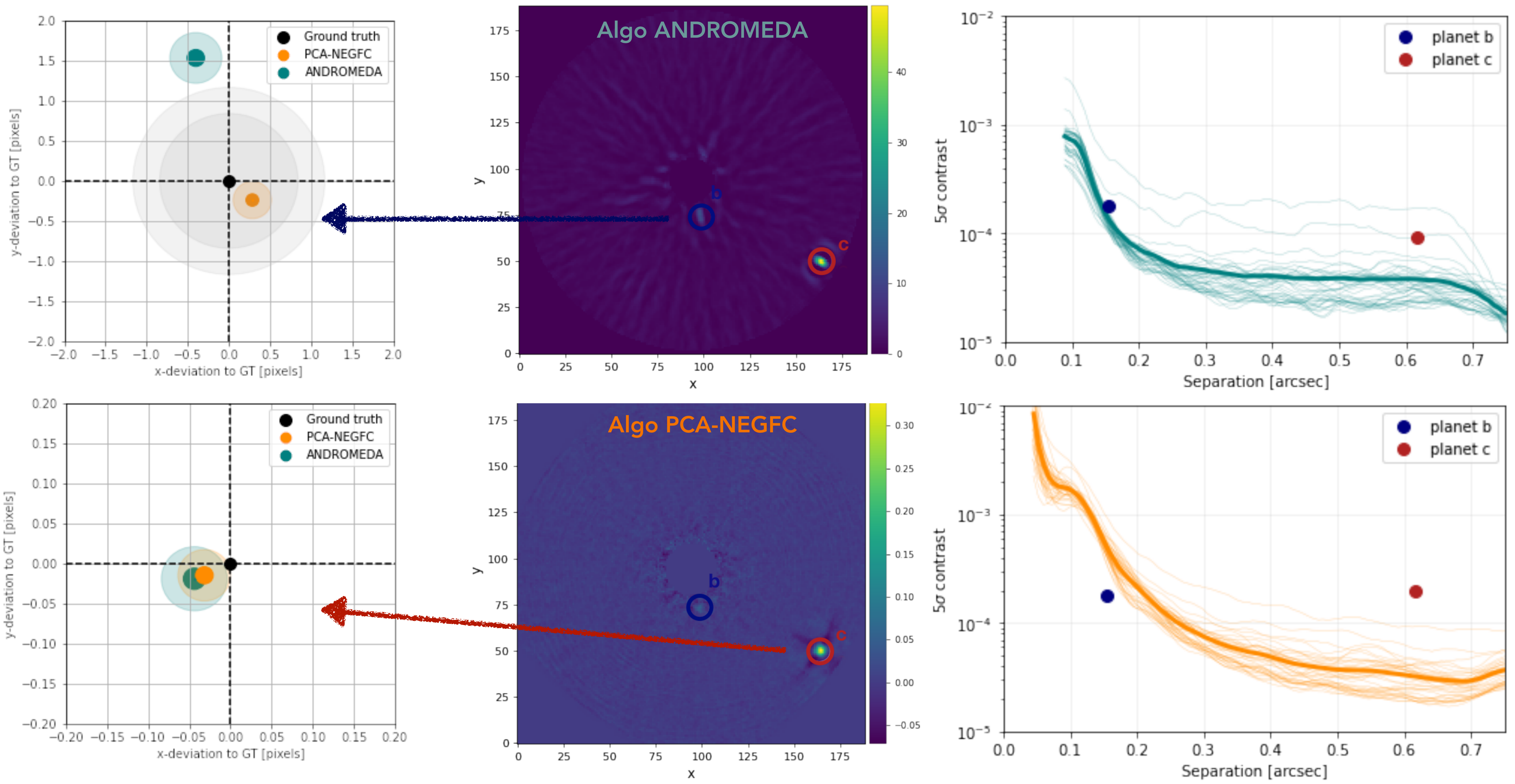
Injected planetary companions spectra

Estimates [+ uncertainties + posterior distribution]

- (spectro)-photometry
- Astrometry

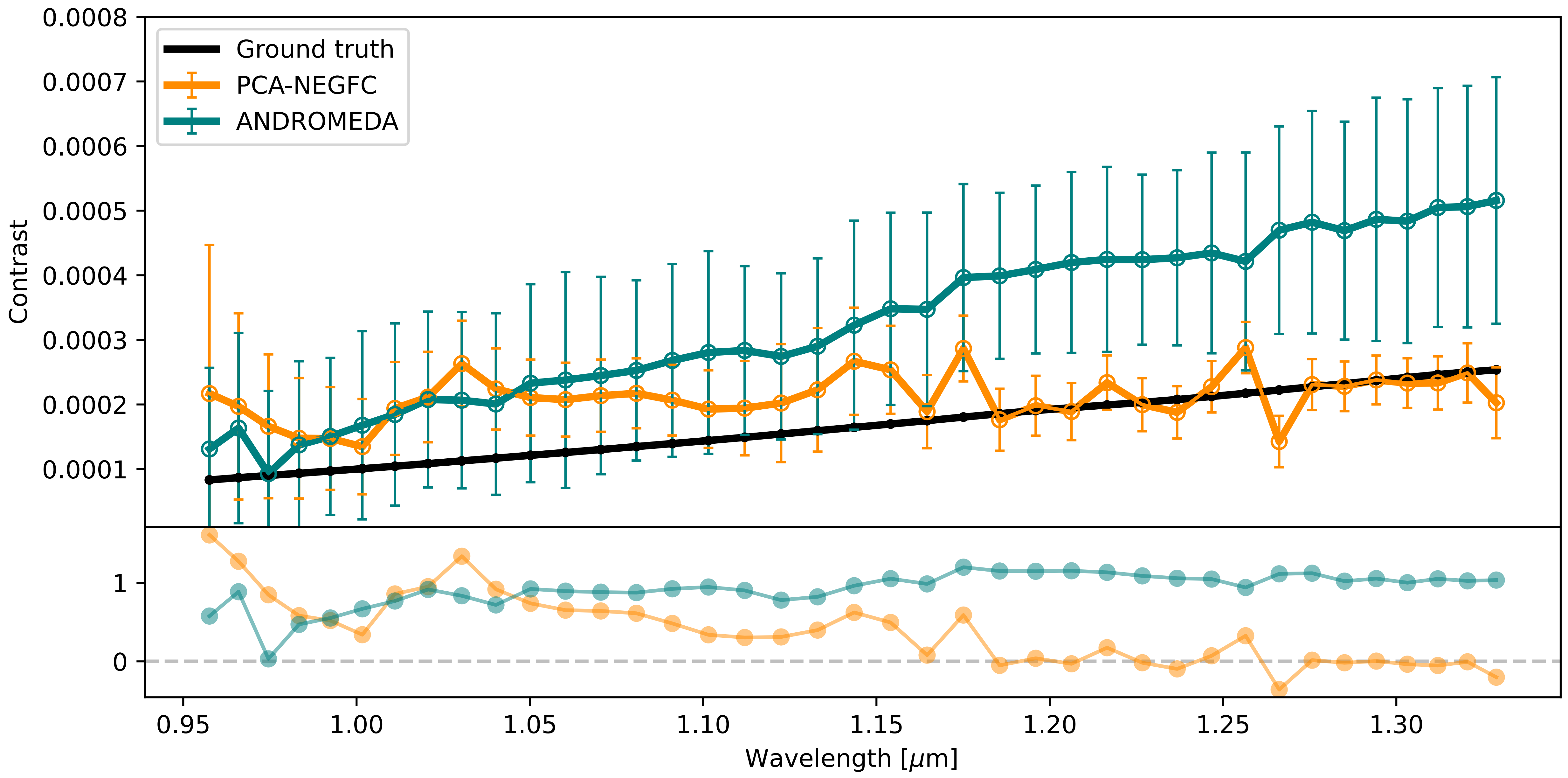
1. Astrometry of point sources

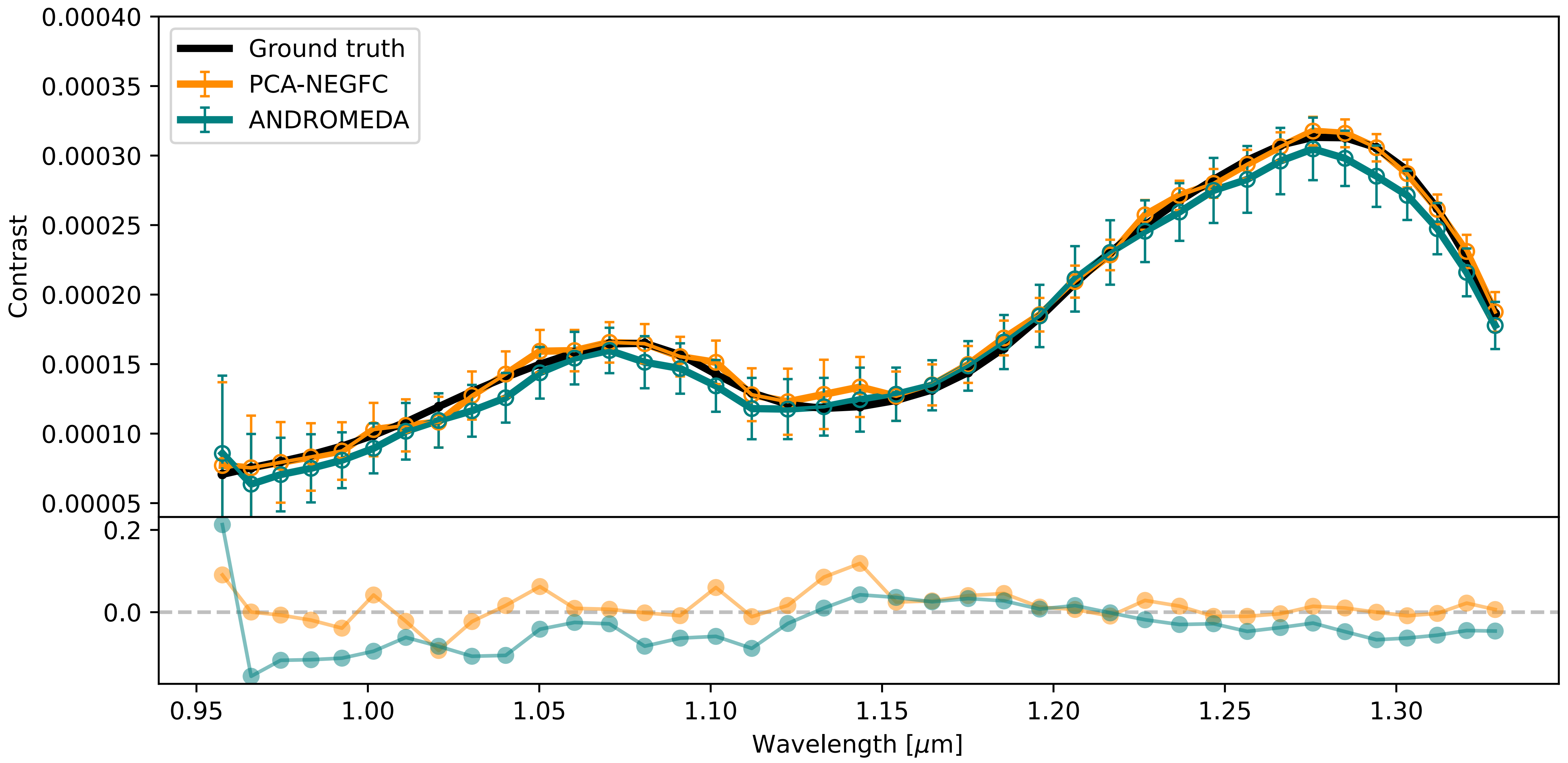
Thursday 9.15am Sarah Blunt



2. Spectrophotometry of point sources

Thursday 8.30am Eileen Gonzales





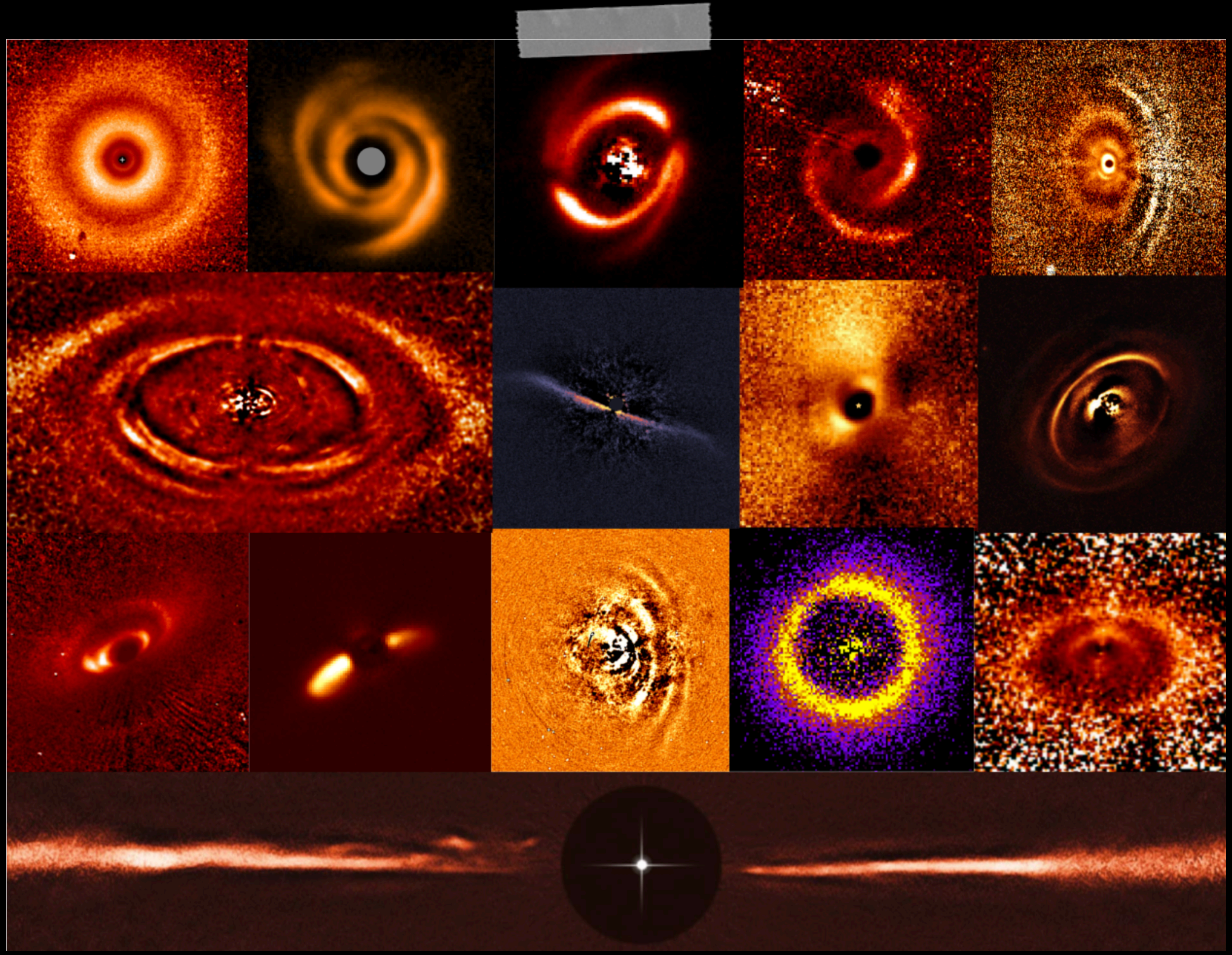
Final "ranking"

Method	PCA-NEGFC			ANDROMEDA		
	planet b	planet c	all	planet b	planet c	all
Astrometry	0.37	0.03	0.20	1.95	0.06	1.01
Spectro-photometry	16.98	1.03	9.00	35.67	2.29	18.99

EIDC website

Penguin interlude #2

Circumstellar extended structures:
Protoplanetary disks and debris disks



Garuffi et al., 2017

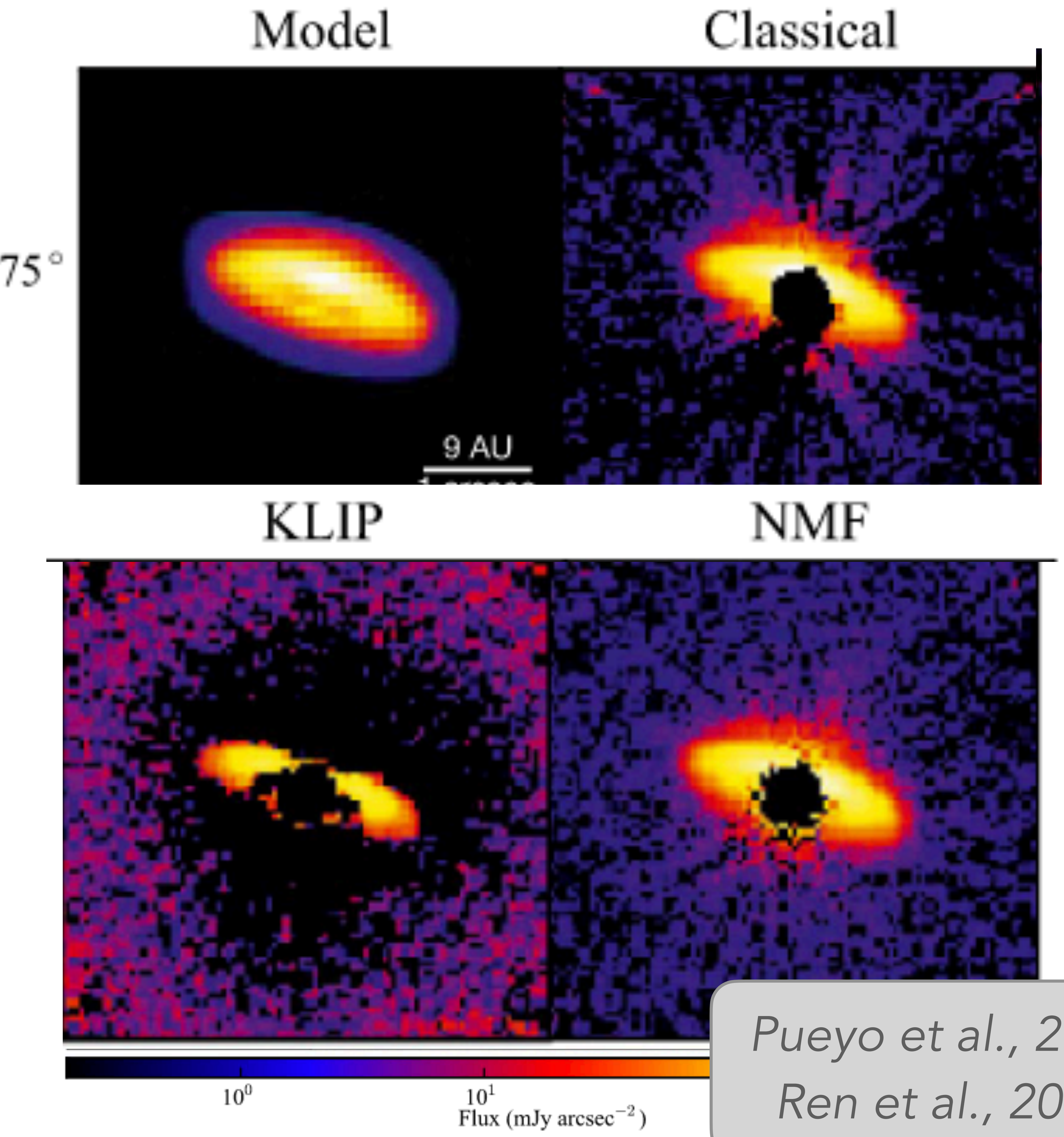
Total intensity image in IR:

- Face-on circular disks
- Structures on edge-on disks
- Spiral structure
- Shadows, dips, gaps…
- Disentangling planets-disks

Towards EIDC Phase 3 !

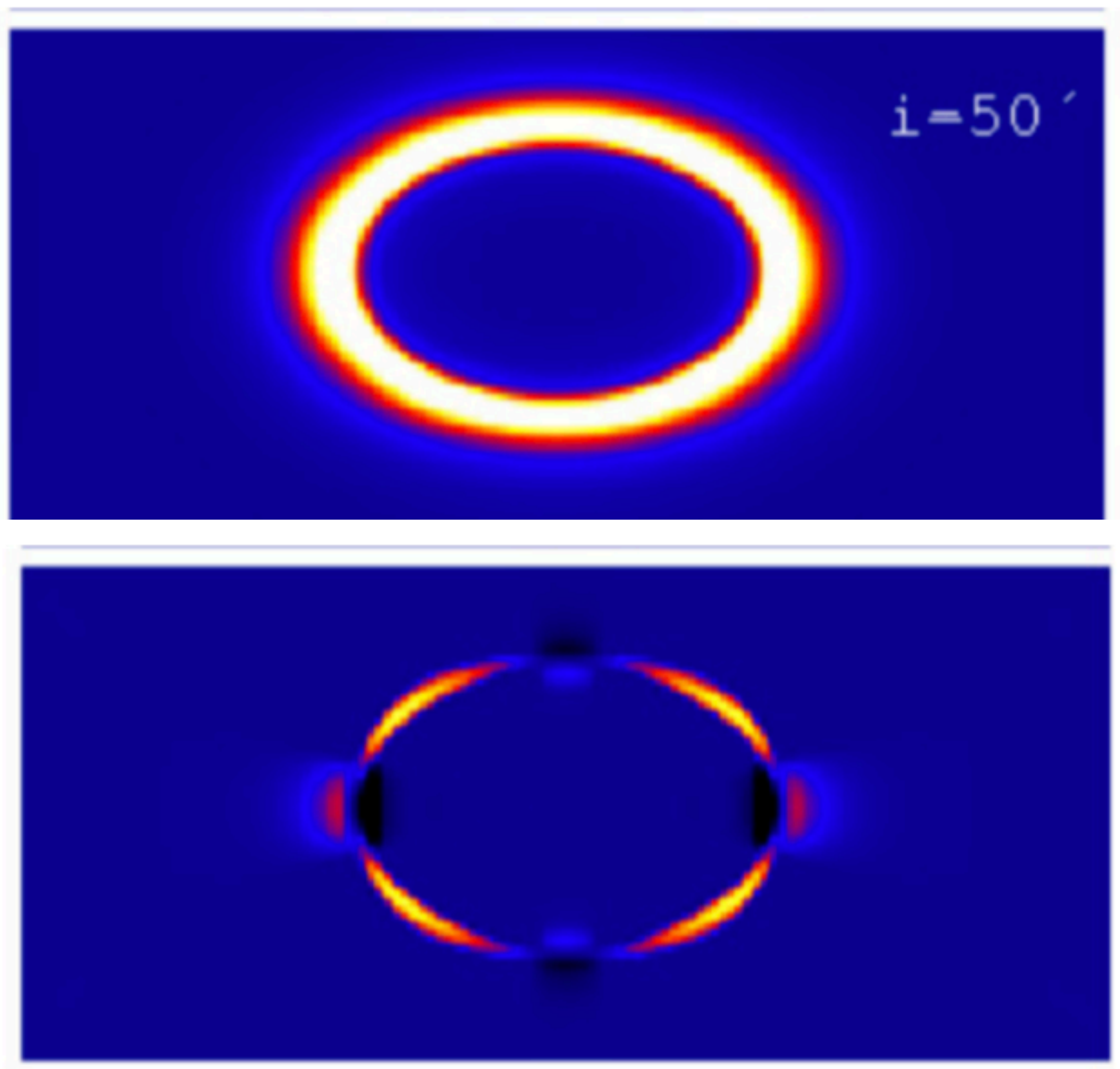
Classic approaches

Breaking down optimisation regions enforcing positivity and sparcity



Pueyo et al., 2016
Ren et al., 2018

Mask the signal
Analyse the ADI-made distortion



Milli et al., 2012
Ren et al., 2020a

Stapper & Ginsky, 2022

Iterate the ADI subtraction to minimize self-subtraction

Pairet et al., 2018

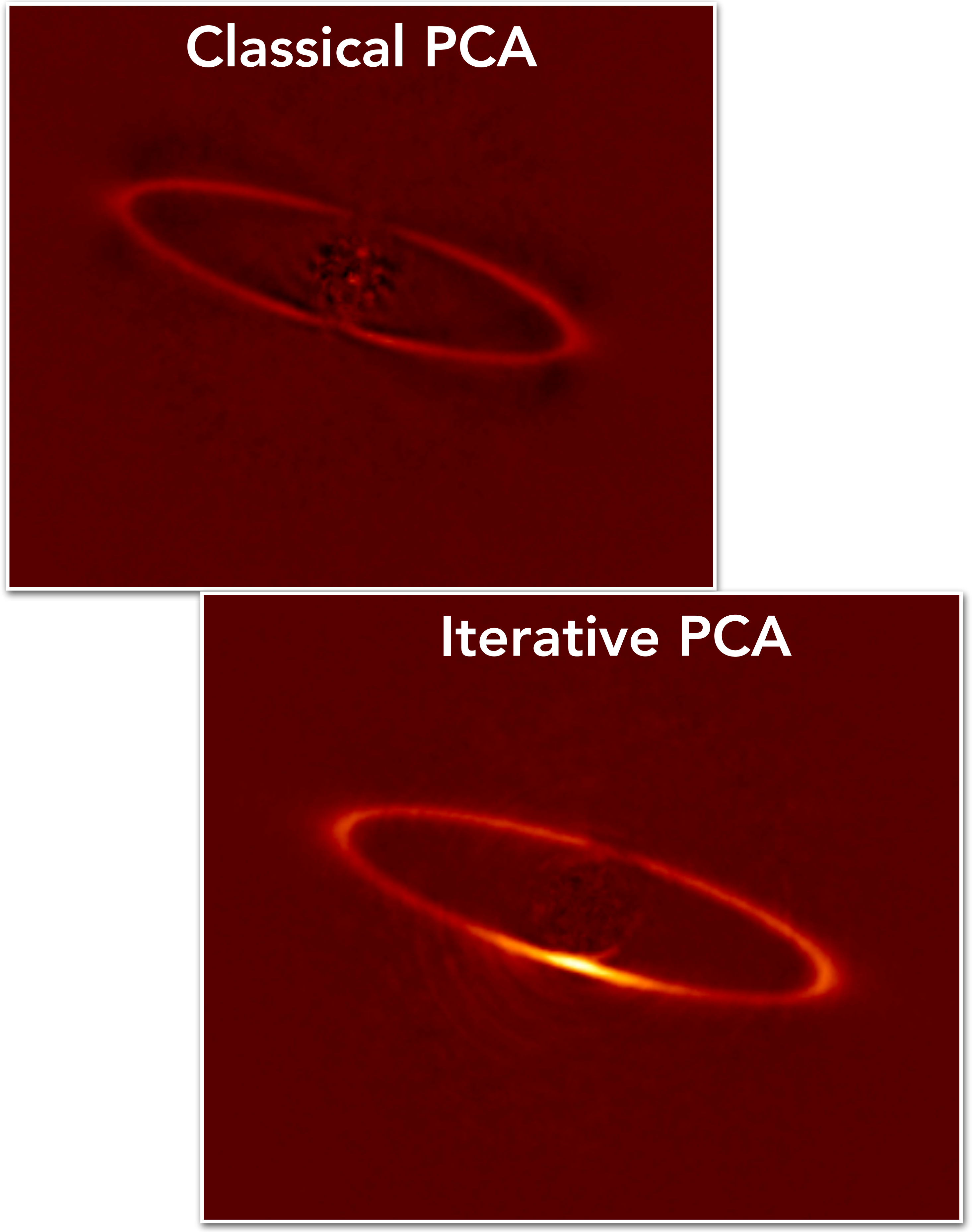
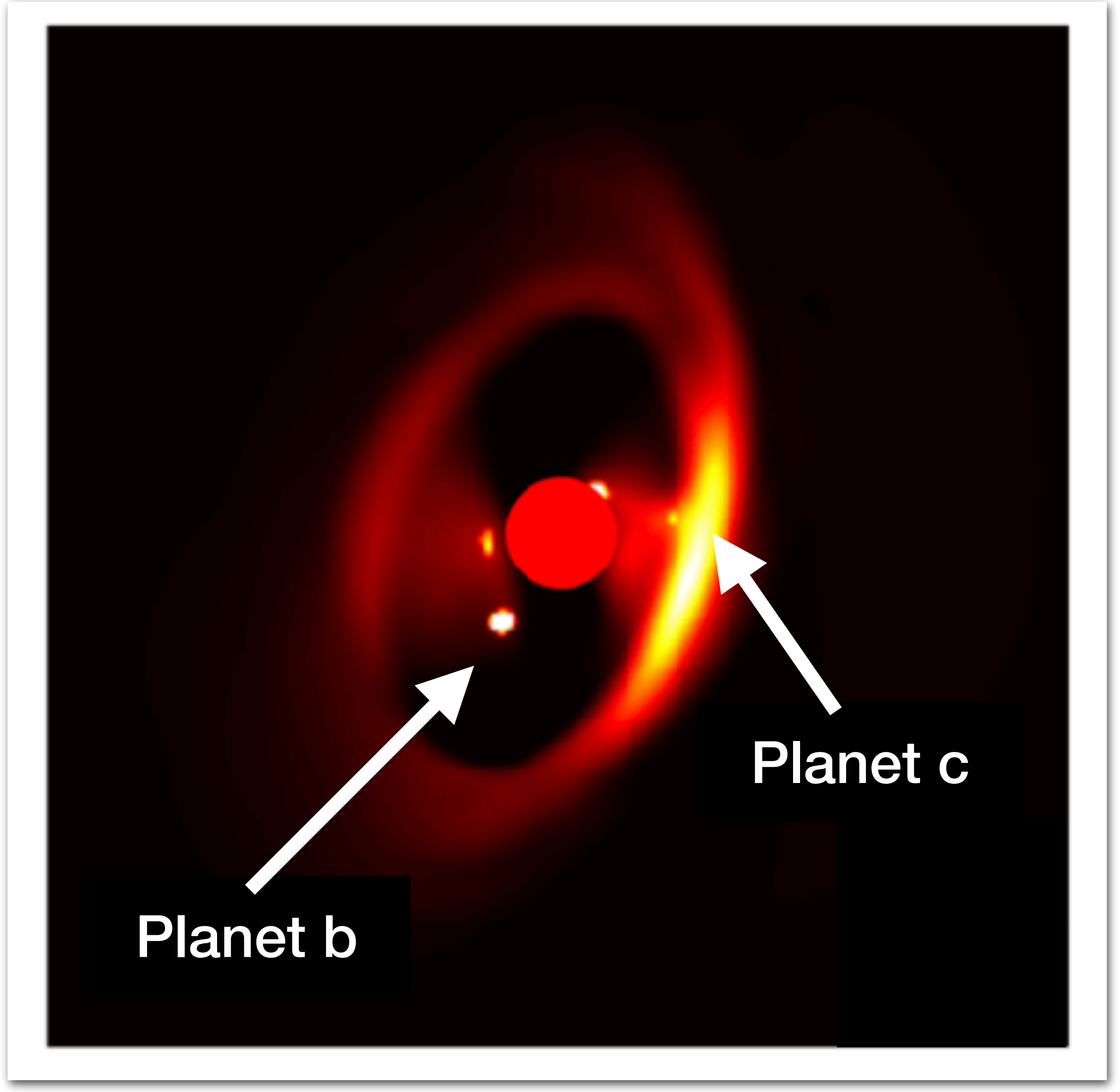


Image reconstruction approaches

MAYONNAISE

Pairet et al., 2020

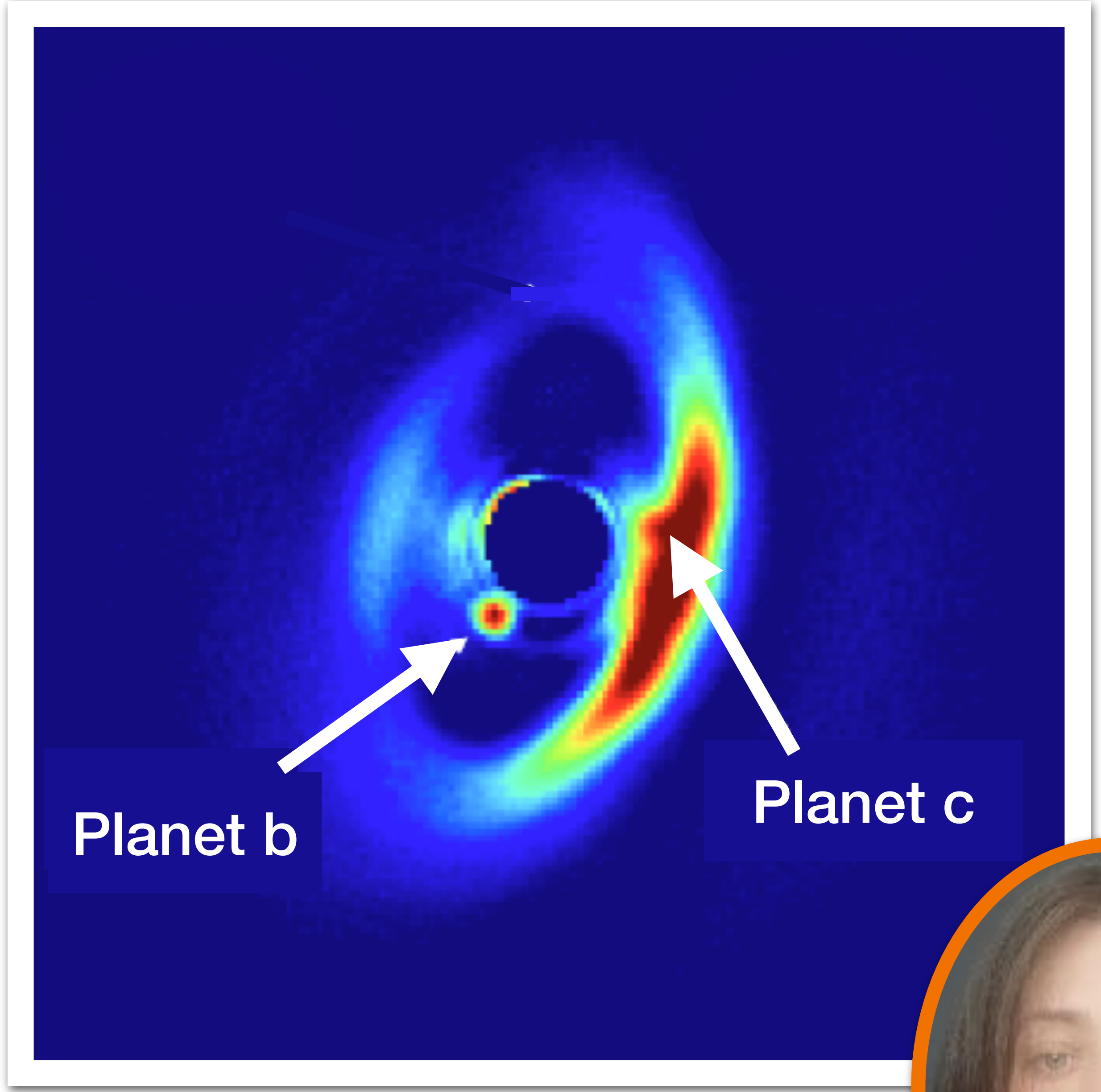
Morphological Component Analysis:
-Disk in shearlet space
-Planet in direct space



MUSTARD

Juillard et al., 2023

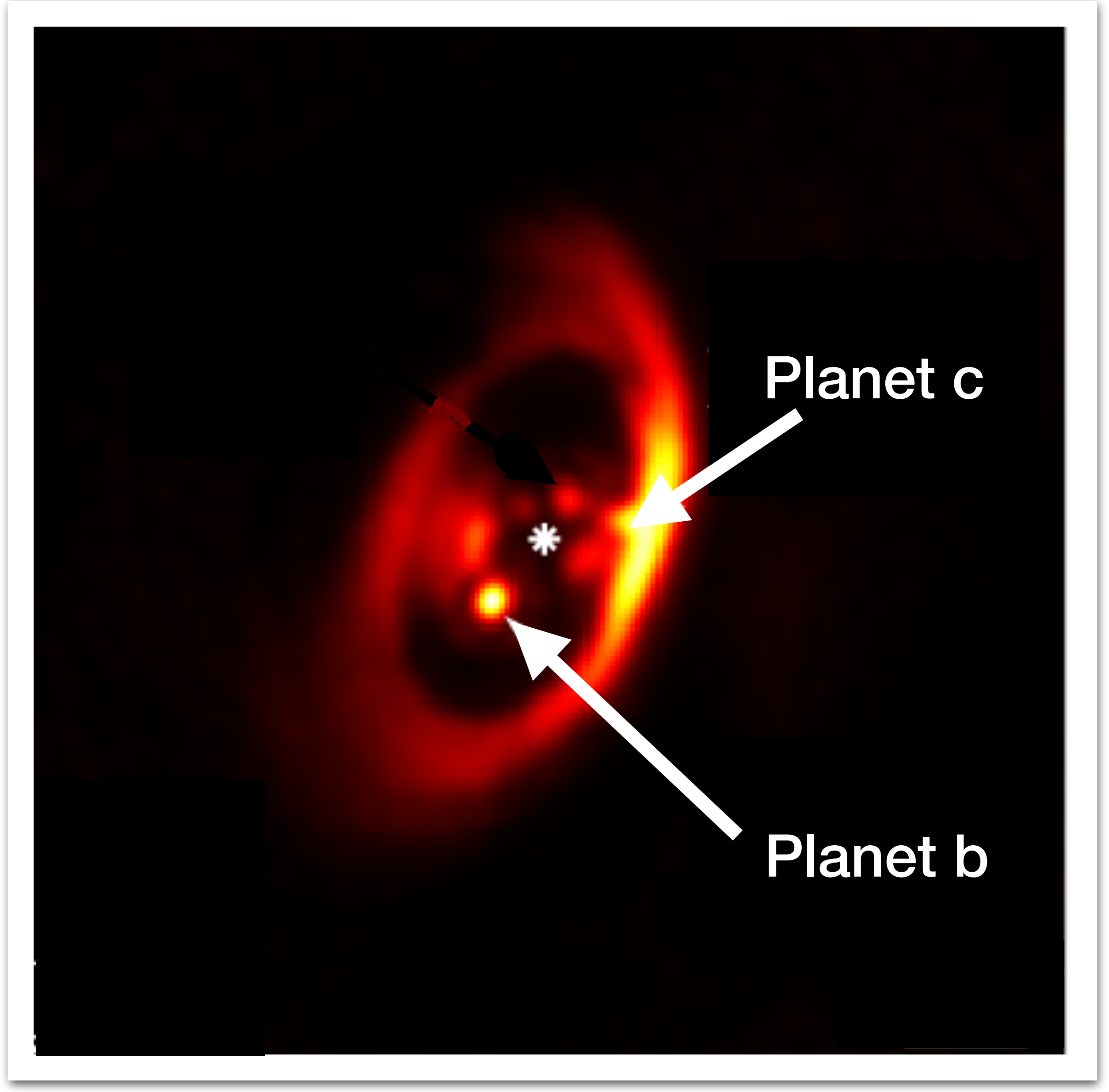
Mask the ambiguous region
due to rotation (not known)



REXPACO

Flasseur et al., 2021

PACO framework to estimate noise
Iterates on the disk estimation



To go further !

Building the reference PSF

multi-Reference RDI:
Using a library of images as a database

Using Data Imputation with NMF — *Ren et al., 2018*

Using Structure Similarity Index — *Ruane et al., 2019*

ConStruct (Auto-encoders based) — *Wolf et al., 2023*

Data Imputation with semi-supervised CNN

Poster #27 Sandrine Juillard

Juillard et al., 2024

Poster #8 Cao Fangyi

IPCA: Combine ADI + RDI

Observing strategies for RDI:
Using another reference star

Snapshot of similar targets — *Bohn et al. 2020a*

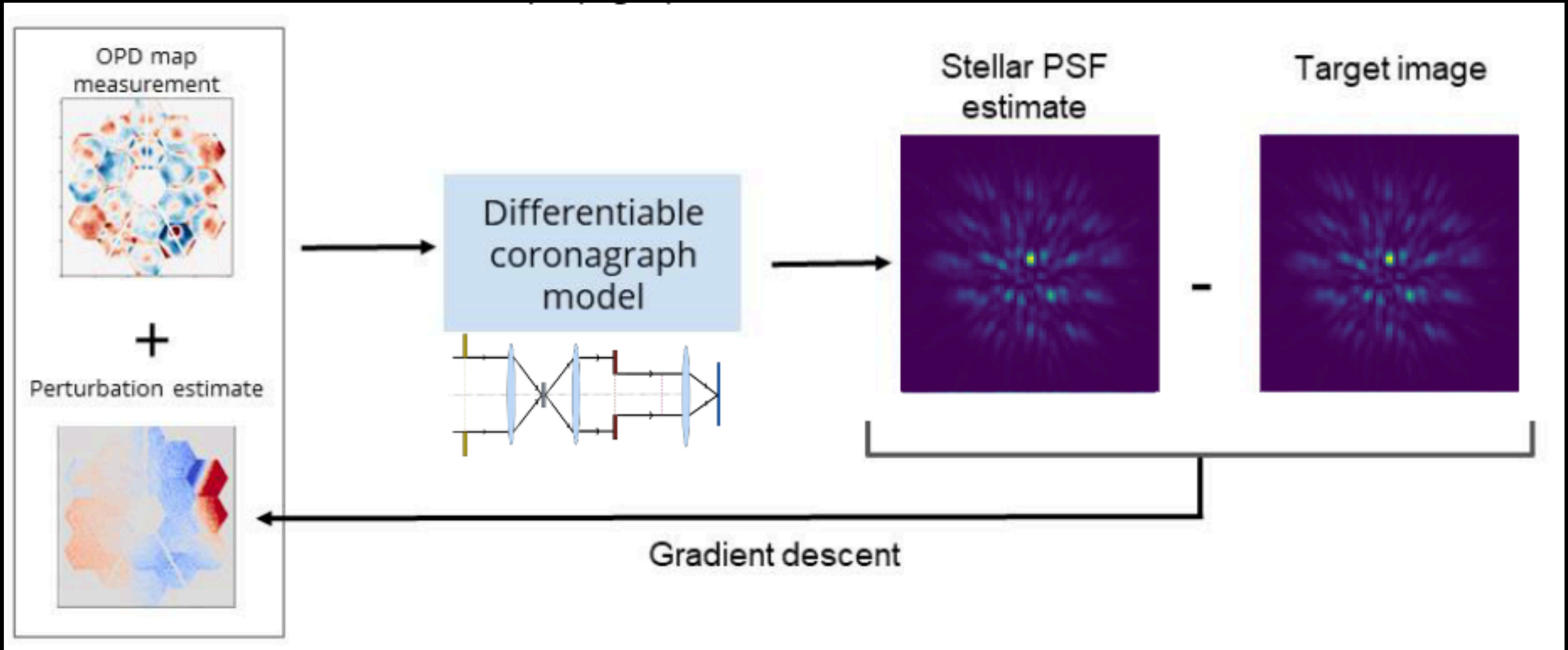
Star-hopping observations — *Wahhaj et al. 2021*

Poster #32 Pengyu Liu

Poster #47 Richelle Cvan Capelleveen

To go further !

Building the reference PSF



Estimation with an instrumental model

Poster #18 Rodrigo Ferrer-Chavez

Pipelines

- Preprocessing tools
- Library of post-processing algo
PCA, LOCI, ANDROMEDA, PACO
- Characterization tools
NEGFC, MCMC…

Library of algorithms !

https://vip.readthedocs.io

PyKLIP

- Post-processing algo + detection maps
KLIP, mRDI + *planetevidence* (multinest)
- Ready-made configuration
GPI, CHARIS, SPHERE, NIRC2, VisAO
- Characterization FM (disk incl.)

Complete toolbox FM-based

https://pyklip.readthedocs.io

PynPoint

- Preprocessing tools
w/ pre-configuration files SPHERE, NaCo
- Post-processing algo
PCA mainly, in-house PACO…
- Characterization tools
NEGFC, MCMC…

Large data sample management !

https://pynpoint.readthedocs.io

And also, CHARIS, GRAPHICS, SPHERE-DC, Data Cruncher etc.

You can use all these beautiful tools !

Summary of key points:

post-processing is essential to gain > 1mag

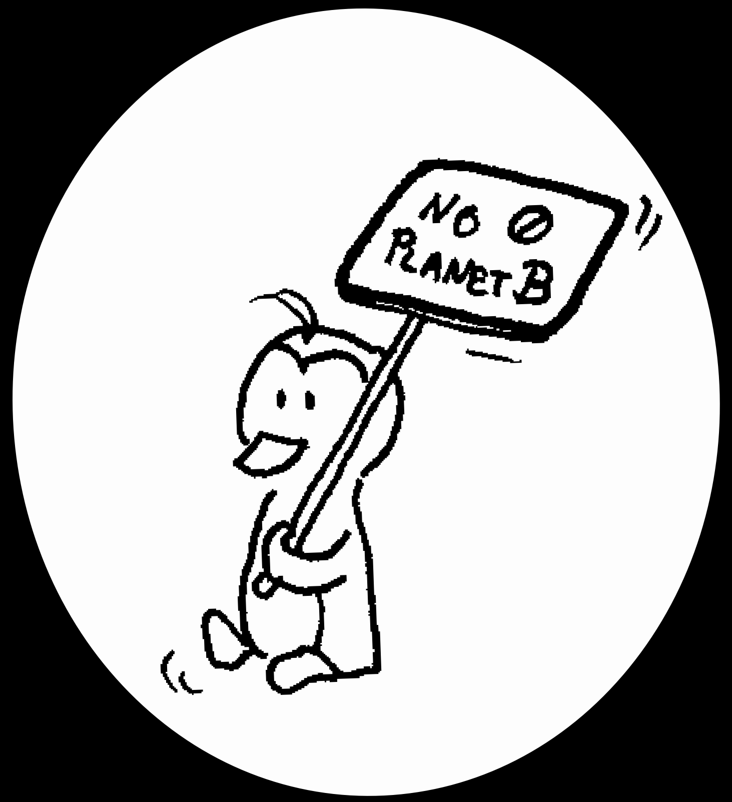
- Understanding the limitations of HCI: temporal stability is key
- Relies on specific observing strategies and calibration
- Characterising the starlight residuals and differential residuals distribution

- All algorithm provide different outputs requiring different interpretation
- Assessing the performance is <u>not</u> obvious at all
- Data challenges are a great tool for homogeneous comparison

EIDC website

Advanced post-processing techniques are **available** and documented !

Use several concept to achieve better astrophysical input



faustine.cantalloube@cnrs.fr