»We explore a possible alternative to PCA-based PSF estimation for **direct imaging** of **exoplanets** which works by **learning** causal pixel-wise **noise models**.«

Half-Sibling Regression meets Direct Imaging: A Causal Approach for Uncovering Exoplanets

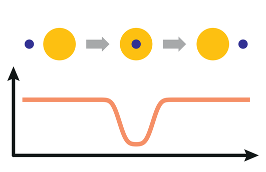
Timothy D. Gebhard[1,2], *Sascha P. Quanz*[3], *Bernhard Schölkopf*[1]

[1]Max Planck Institute for Intelligent Systems | [2]Max Planck ETH Center for Learning Systems | [3]Institute for Particle Physics & Astrophysics (ETHZ)

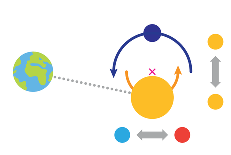
Corresponding author:
tgebhard@tue.mpg.de

Detecting Exoplanets

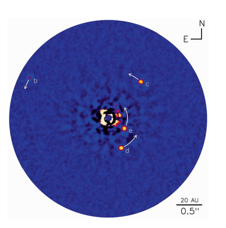
1 Transit Photometry
Identify periodic "dips" in the light curves of stars (→ *e.g.*, Kepler)



2 Astrometry / Radial Velocity
Observe periodic changes in star position / redshift (→ *e.g.*, Gaia)

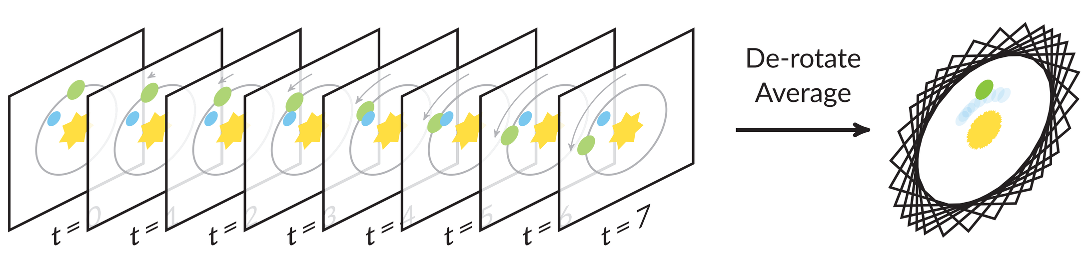


3 Direct Imaging
Capture photons from planet, mostly in the infrared (→ *e.g.*, NACO/VLT)



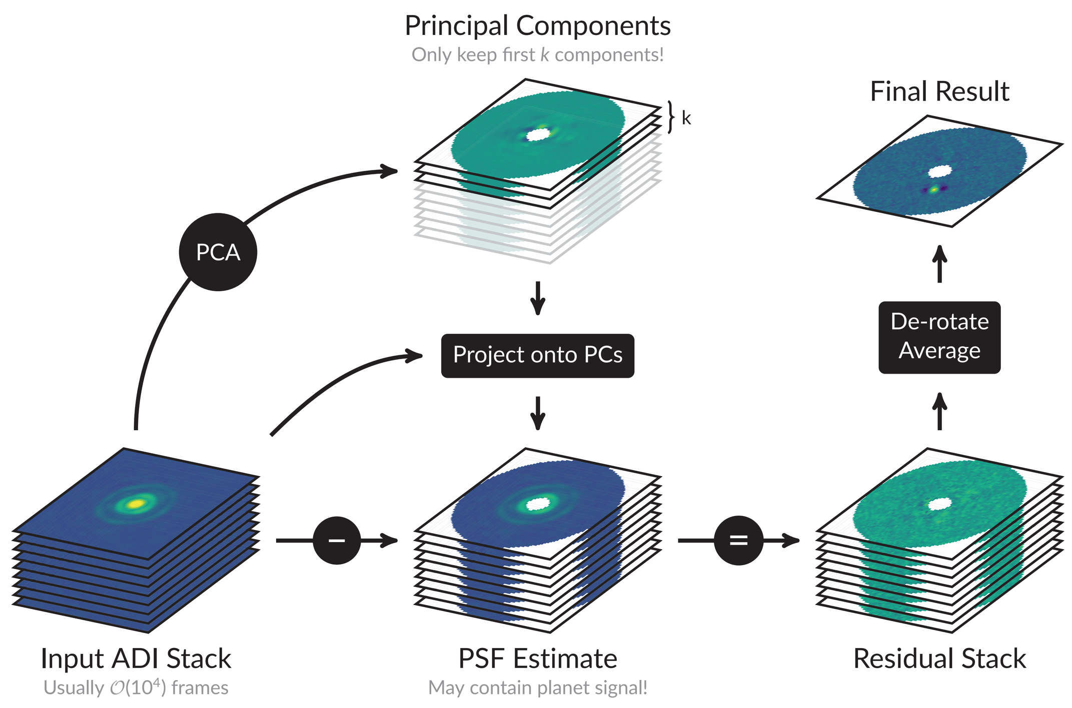
Angular Differential Imaging

Angular Differential Imaging (ADI) is an **observation technique** first introduced by Marois *et al.* (2005):



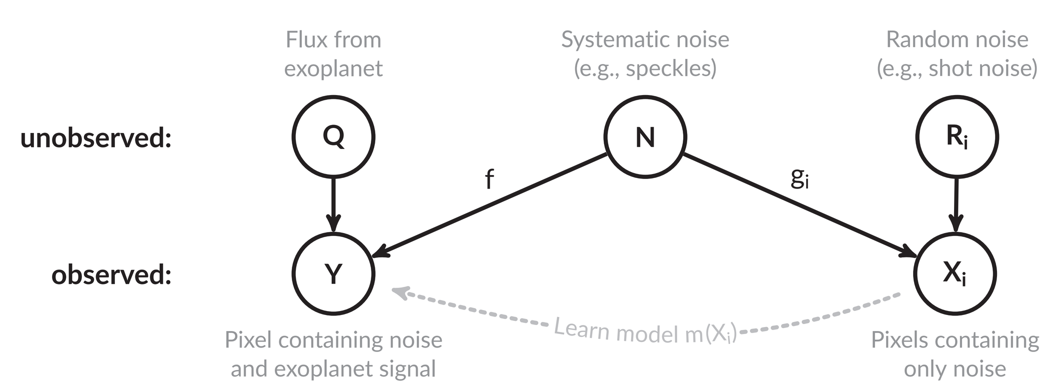
Key idea: **turn off the field derotator** of the telescope. **De-rotating** every frame by its parallactic angle and **averaging** will then **reduce quasi-static speckle noise**.

PCA-based PSF subtraction



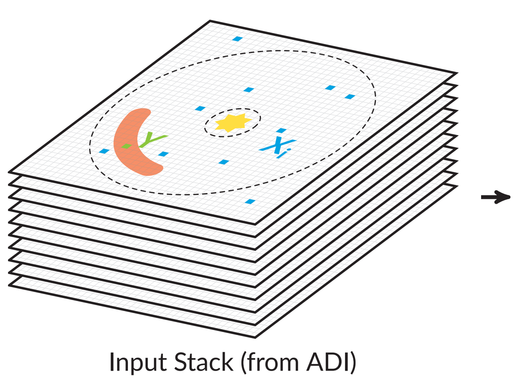
Half-Sibling Regression

Assume we are interested in the latent quantity $\mathbf{Q} \perp\!\!\!\perp \mathbf{X}_i$:



Schölkopf *et al.* (2016) have shown that by regressing **Y** onto $\mathbf{X}_i$, one can learn an estimate for **f(N)**, allowing us to reconstruct **Q** (up to a constant) as $\hat{\mathbf{Q}} = \mathbf{Y} - \mathbf{f(N)}$.

Application to Direct Imaging



For every pixel **Y** in the ROI, we learn a model **m**:

$$\mathbf{Y} = \mathbf{m}(\mathbf{X}_i) + \varepsilon$$

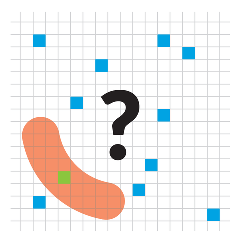
Idea: Use $\mathbf{m}(\mathbf{X}_i)$ as an estimate for the systematics.

→ ε is the estimated planet signal in pixel Y!

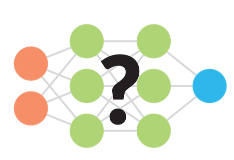
Advantage: Lots of freedom in choosing **m** (and $\mathbf{X}_i$)!
→ Reduce self-subtraction, include time domain, ... ?

Current Questions

Predictor Selection
- Which pixels should we use?
- How to define the exclusion region?



Machine Learning
- Which model class should we use?
- How to split for training / test?



Evaluation
- Physical metrics vs. ML metrics?
- How to quantify self-subtraction?