»We study how to uncover exoplanets in high-contrast imaging data by learning causal, pixel-wise noise models and including scientific domain knowledge.«

Physically constrained causal noise models for high-contrast imaging of exoplanets

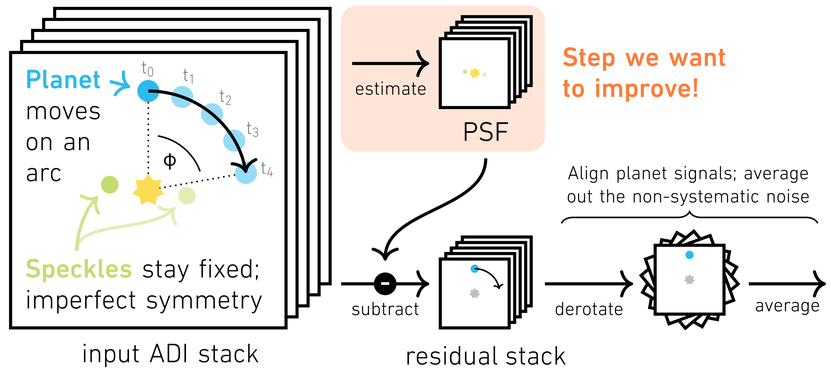
Timothy D. Gebhard[1,2], Markus J. Bonse[2], Sascha P. Quanz[2], Bernhard Schölkopf[1,3]

[1]Max Planck Institute for Intelligent Systems, Tübingen
[2]Institute for Particle Physics & Astrophysics, ETH Zurich
[3]Department of Computer Science, ETH Zurich

tgebhard@tue.mpg.de
tdgebhard

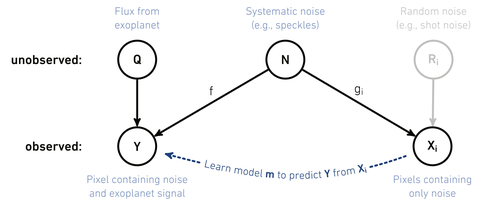
1. Background: High-contrast imaging (HCI)

- In HCI, we record a "video" of a star and then combine the frames to uncover any planets.
- Challenge: the star is much brighter than its planets. Additionally, there is systematic noise ("speckles") that often mimics a planet signal.
- Therefore, the most crucial part of any HCI post-processing pipeline is the denoising step.
- We try to improve this step by combining ML with existing domain knowledge (e.g., known symmetries in the data, temporal behavior, ...).



2. Half-sibling regression

- HSR is a denoising framework based on causality proposed by Schölkopf et al. (2016).
- The key idea is to exploit confounding effects to model systematic noise.
- Relevant: in physics, the causal structure of data-generating processes is often well-understood.

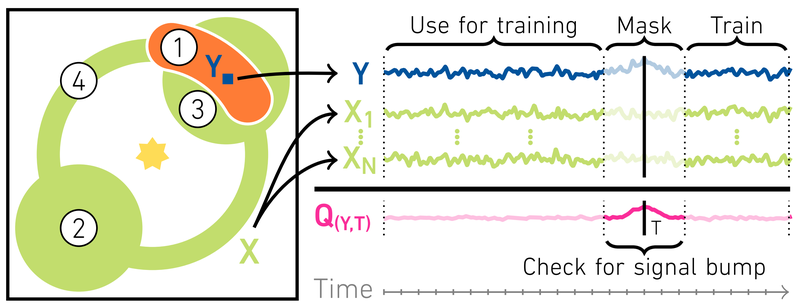


Assumption 1: $Y = f(N) + Q$ (additive noise)

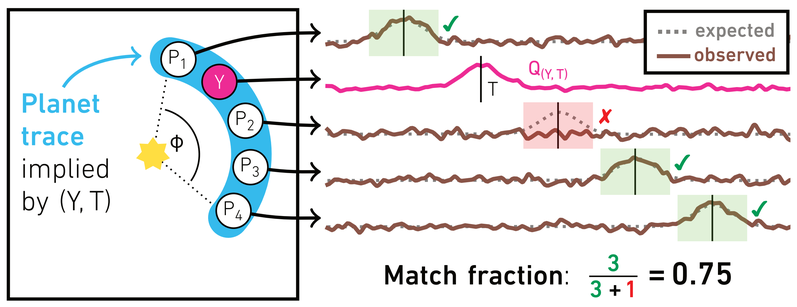
Assumption 2: $Q \perp\!\!\!\perp X_i, N$ (causal independence)

In this case: $m(X_i) \longrightarrow f(N)$ (given enough data)

Therefore: $Q \approx Y - m(X_i)$ (up to a constant offset)



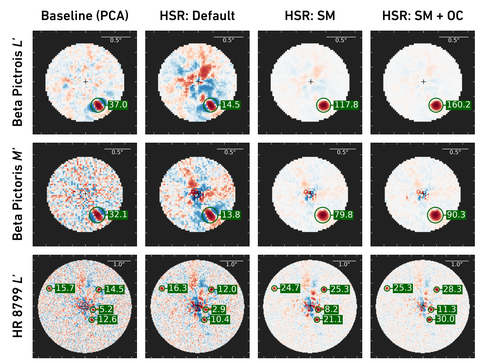
3. Method | For each pixel, we train several models (see paper for details).



4. Method | We run consistency checks to select the model for every pixel.

5. Results and outlook

- Very promising: HSR with signal masking (SM) looks better than PCA both visually and in SNR
- Observing conditions (OC) as additional predictors improves results further.
- Note: SNR alone is *not* yet sufficient to characterize HCIpp algorithms; further studies are necessary!



Preliminary, exemplary results for 3 real data sets that contain planets.

MAX PLANCK INSTITUTE FOR INTELLIGENT SYSTEMS
ETH zürich
arXiv:2010.05591
Proudly presented at the *Machine Learning and the Physical Sciences* workshop at NeurIPS 2020.
NEURAL INFORMATION PROCESSING SYSTEMS