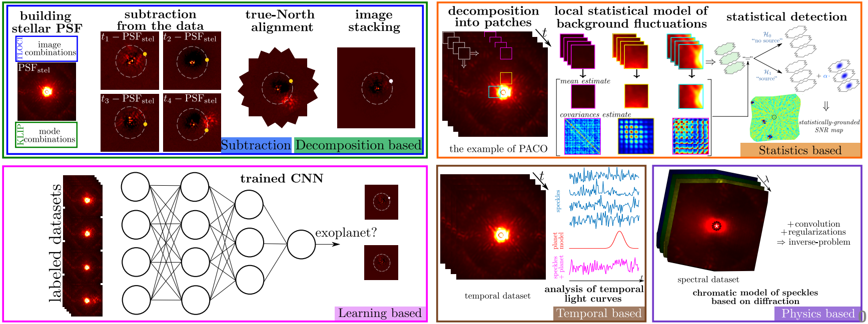
Post-processing algorithms for exoplanet detection and characterization at high contrast by angular (and spectral) differential imaging

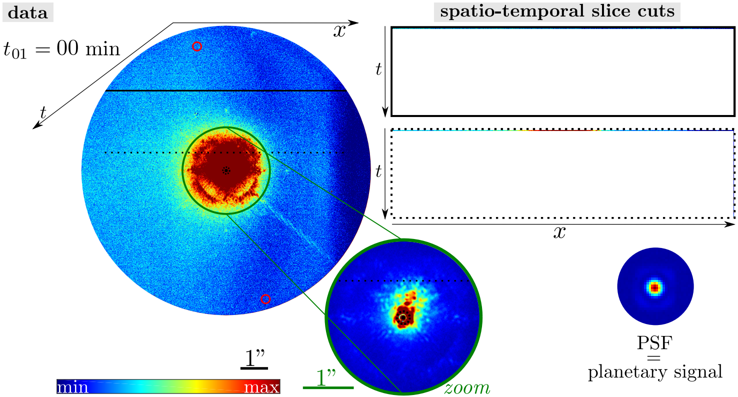
A focus on data-driven approaches

Olivier Flasseur



Typical dataset from VLT/SPHERE-IRDIS instrument

angular differential imaging (ADI) = temporal diversity



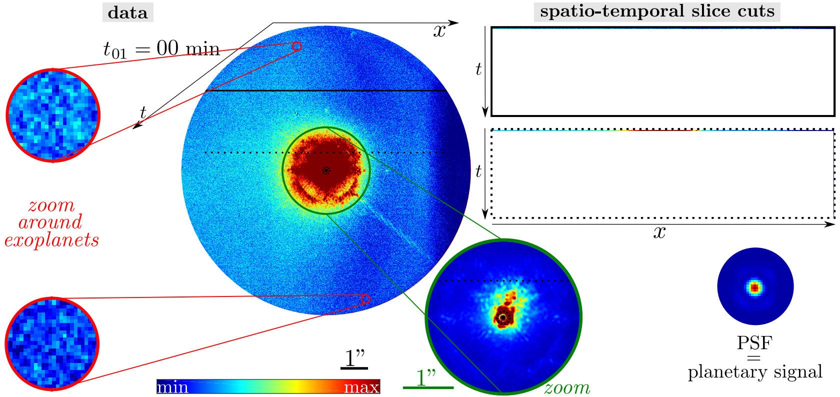
Peculiarities

- **Faint signal** from the exoplanets
- **Non-stationary** and **spatially correlated** strong background
- **Strong fluctuations** (stellar leakages)
- **Multi-spectral** data available

$\Rightarrow$ **Signal processing is mandatory**

Typical dataset from VLT/SPHERE-IRDIS instrument

angular differential imaging (ADI) = temporal diversity



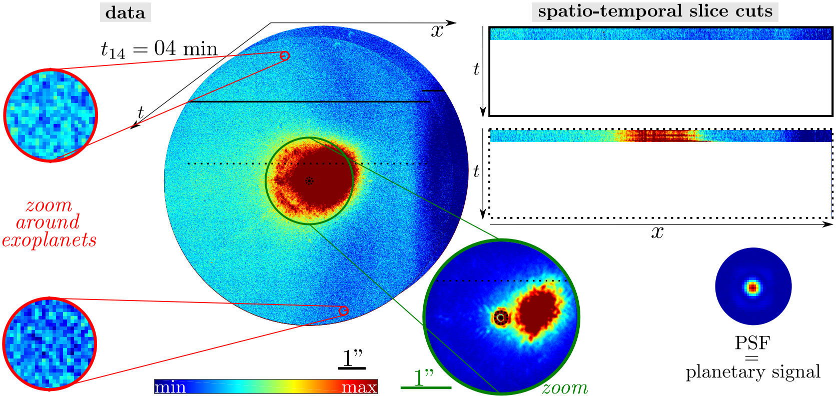
Peculiarities

- Faint signal from the exoplanets
- Non-stationary and spatially correlated strong background
- Strong fluctuations (stellar leakages)
- Multi-spectral data available

$\Rightarrow$ Signal processing is mandatory

Typical dataset from VLT/SPHERE-IRDIS instrument

angular differential imaging (ADI) = temporal diversity



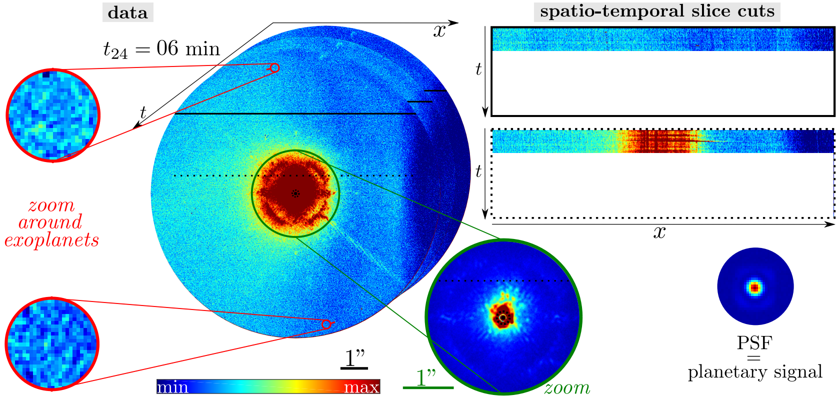
Peculiarities

- **Faint signal** from the exoplanets
- **Non-stationary** and **spatially correlated** strong background
- **Strong fluctuations** (stellar leakages)
- **Multi-spectral** data available

$\Rightarrow$ **Signal processing is mandatory**

Typical dataset from VLT/SPHERE-IRDIS instrument

angular differential imaging (ADI) = temporal diversity



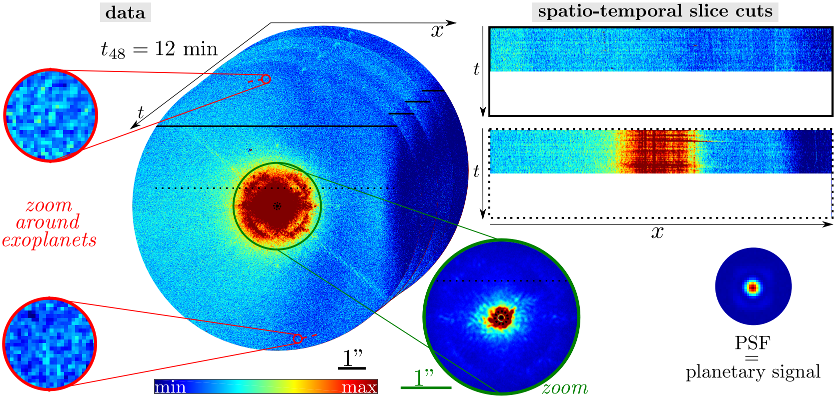
Peculiarities

- **Faint signal** from the exoplanets
- **Non-stationary** and **spatially correlated** strong background
- **Strong fluctuations** (stellar leakages)
- **Multi-spectral** data available

$\Rightarrow$ **Signal processing is mandatory**

Typical dataset from VLT/SPHERE-IRDIS instrument

angular differential imaging (ADI) = temporal diversity



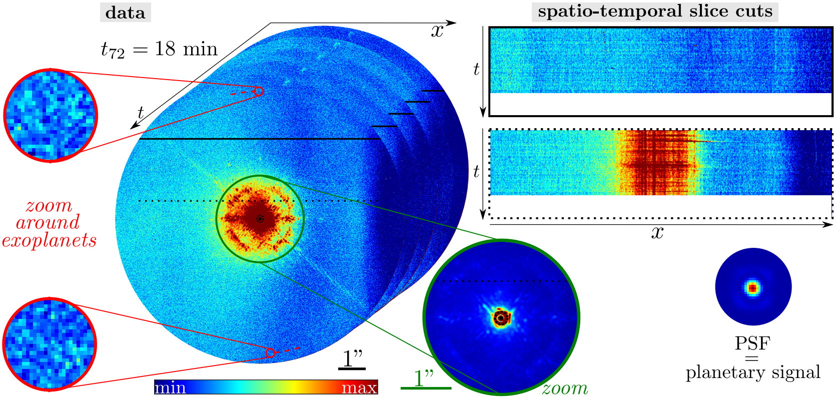
Peculiarities

- **Faint signal** from the exoplanets
- **Non-stationary** and **spatially correlated** strong background
- **Strong fluctuations** (stellar leakages)
- **Multi-spectral** data available

$\Rightarrow$ **Signal processing is mandatory**

Typical dataset from VLT/SPHERE-IRDIS instrument

angular differential imaging (ADI) = temporal diversity



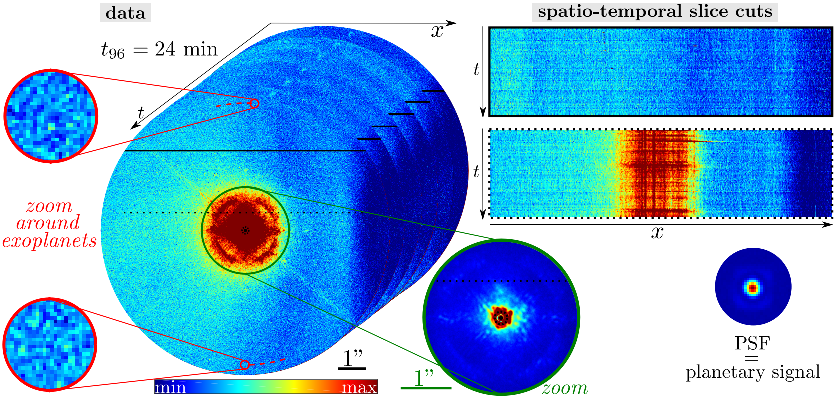
Peculiarities

- **Faint signal** from the exoplanets
- **Non-stationary** and **spatially correlated** strong background
- **Strong fluctuations** (stellar leakages)
- **Multi-spectral** data available

$\Rightarrow$ **Signal processing is mandatory**

Typical dataset from VLT/SPHERE-IRDIS instrument

angular differential imaging (ADI) = temporal diversity



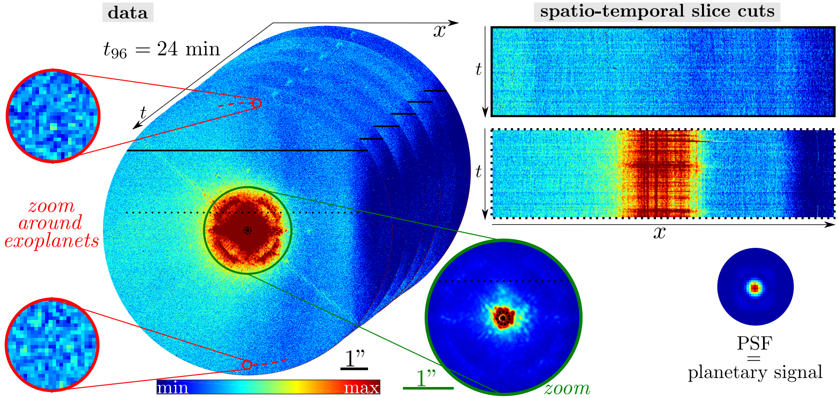
Peculiarities

- **Faint signal** from the exoplanets
- **Non-stationary** and **spatially correlated** strong background
- **Strong fluctuations** (stellar leakages)
- **Multi-spectral** data available

$\Rightarrow$ **Signal processing is mandatory**

Typical dataset from VLT/SPHERE-IRDIS instrument

angular differential imaging (ADI) = temporal diversity



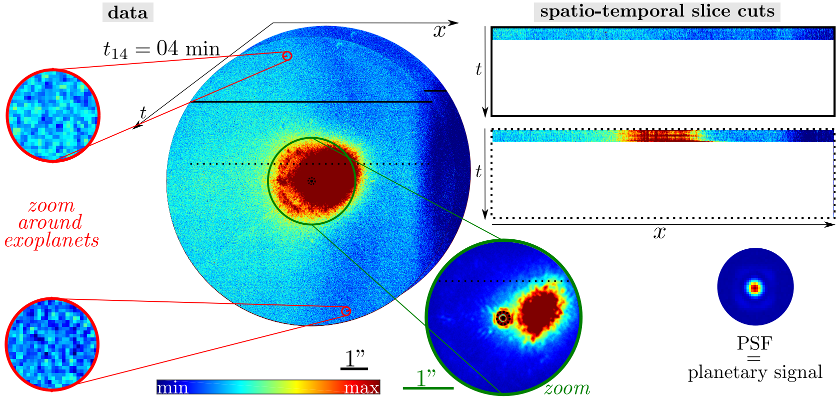
Peculiarities

- **Faint signal** from the exoplanets
- **Non-stationary** and **spatially correlated** strong background
- **Strong fluctuations** (stellar leakages)
- **Multi-spectral** data available

$\Rightarrow$ **Signal processing is mandatory**

Typical dataset from VLT/SPHERE-IRDIS instrument

angular differential imaging (ADI) = temporal diversity



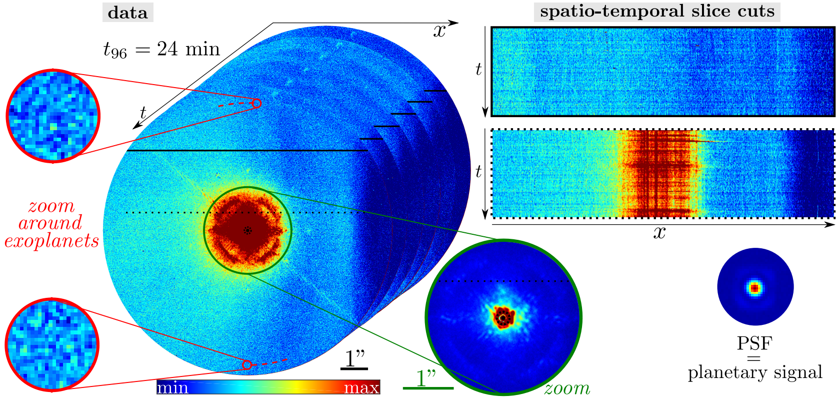
Peculiarities

- **Faint signal** from the exoplanets
- **Non-stationary** and **spatially correlated** strong background
- **Strong fluctuations** (stellar leakages)
- **Multi-spectral** data available

⇒ **Signal processing is mandatory**

Typical dataset from VLT/SPHERE-IRDIS instrument

angular differential imaging (ADI) = temporal diversity



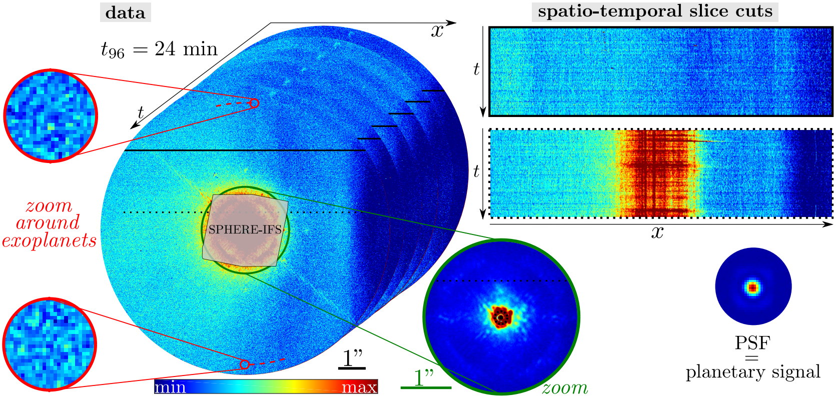
Peculiarities

- **Faint signal** from the exoplanets
- **Non-stationary** and **spatially correlated** strong background
- **Strong fluctuations** (stellar leakages)
- **Multi-spectral** data available

$\Rightarrow$ **Signal processing is mandatory**

Typical dataset from VLT/SPHERE-IRDIS instrument

angular differential imaging (ADI) = temporal diversity



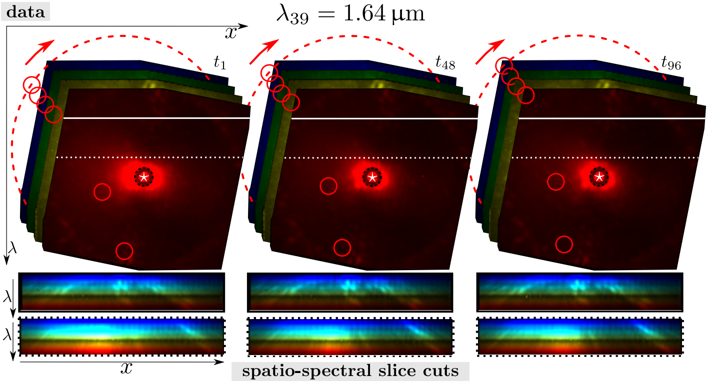
Peculiarities

- **Faint signal** from the exoplanets
- **Non-stationary** and **spatially correlated** strong background
- **Strong fluctuations** (stellar leakages)
- **Multi-spectral** data available

$\Rightarrow$ **Signal processing is mandatory** $\Leftarrow$

Typical dataset from VLT/SPHERE-IFS instrument

angular & spectral diff. im. (ASDI) = temporal & spectral diversity



Peculiarities

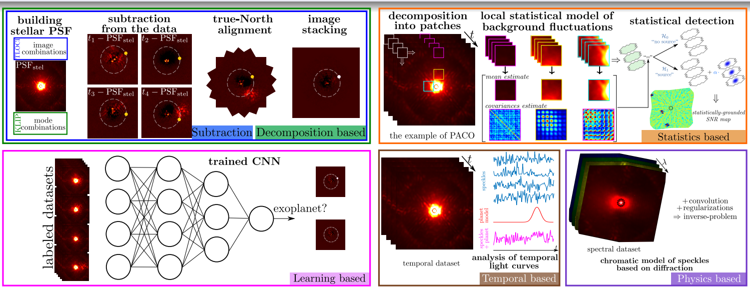
- **Faint signal** from the exoplanets
- **Non-stationary** and **spatially correlated** strong background
- **Strong fluctuations** (stellar leakages)
- **Multi-spectral** data available

$\Rightarrow$ **Signal processing is mandatory** $\Leftarrow$

State-of-the-art processing methods: *summary*

Existing methods

- **Subtraction** - **Decomposition**: KLIP [Soummer *et al.*, 2012]; TLOCI [Lafrénière *et al.*, 2014]
- **Statistics**: MOODS [Smith *et al.*, 2009]; ANDROMEDA [Cantalloube *et al.*, 2016]; PACO
- **Learning**: S4 [Fergus *et al.*, 2014] SODINN [Gonzalez *et al.*, 2017], CNN [Yip *et al.*, 2020]
- **Temporal**: RSM [Dahlqvist *et al.*, 2020], TRAP [Samland *et al.*, 2021]
- **Physics**: PeX [Devanay *et al.*, 2017], MEDUSAE [Cantalloube *et al.*, 2018]



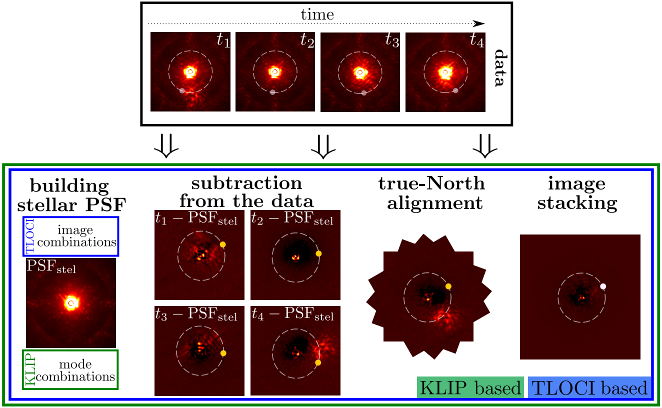
Main challenges

- dealing with the **high contrat**, **high-resolution**,
- accounting for the **non-stationarities** of the background,
- being **robust** against large **fluctuations & outliers**.

Subtraction-based: *general principle*

Subtraction: cADI [Marois+, 2006]; KLIP/PCA [Soummer+, 2012]; TLOCI [Lafrénière+, 2014]

and many variants...



+ possibility to add a forward model of planet signature (KLIP-FM: [Pueyo+, 2016], FMMF: [Ruffio+, 2017])

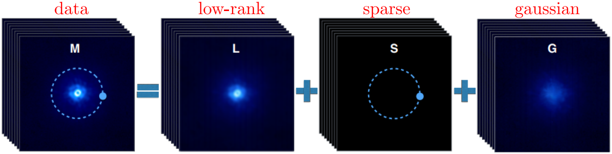
$\Rightarrow$ **Used routinely in direct imaging...**

but limited sensibility & no control false alarm / detection probabilities

Subtraction-based $\Rightarrow$ decomposition-based: *LLSG*

local data decomposition: low rank + sparse + gaussian

Model of the observations: $\mathbf{M} = \mathbf{L} + \mathbf{S} + \mathbf{N}$



[Gomez Gonzalez *et al.*, 2016]

$\Rightarrow$ **explicit unmixing of the planet signal $\neq$ PCA, (T)LOCI**

- **Inverse problem formulation**:

$$\Rightarrow \{\widehat{\mathbf{L}}, \widehat{\mathbf{S}}\} = \text{arg min}_{\mathbf{L},\mathbf{S}} \, \tfrac{1}{2}||\mathbf{M} - \mathbf{L} - \mathbf{S}||_2^2 \quad \text{s.t.} \quad \text{rank}(\mathbf{L}) \leq r \, , ||\mathbf{S}||_0 \leq s$$

$$\Rightarrow \{\widehat{\mathbf{L}}, \widehat{\mathbf{S}}\} = \text{arg min}_{\mathbf{L},\mathbf{S}} \, \tfrac{1}{2}||\mathbf{M} - \mathbf{L} - \mathbf{S}||_2^2 \quad \text{s.t.} \quad ||\mathbf{L}||_* \leq \tau_* \, , ||\mathbf{S}||_1 \leq \tau_1$$

- **Alternate low-rank plus sparse separation**:

$$\widehat{\mathbf{L}}_i = \text{arg min}_{\mathbf{L}} \, ||\mathbf{M} - \mathbf{L} - \widehat{\mathbf{S}}_{i-1}||_F^2 \quad \text{s.t.} \quad \text{rank}(\mathbf{L}) \leq r$$

$$\widehat{\mathbf{S}}_i = \text{arg min}_{\mathbf{S}} \, ||\mathbf{M} - \widehat{\mathbf{L}}_i - \widehat{\mathbf{S}}||_F^2 \quad \text{s.t.} \quad ||\mathbf{L}||_0 \leq s$$

- **Subproblems $\simeq$ solved with a greedy approach of truncated SVD**:

$$\widehat{\mathbf{L}}_i = \mathscr{H}_k^{\text{SVD}}(\mathbf{M} - \widehat{\mathbf{S}}_{i-1}) \quad \text{and} \quad \widehat{\mathbf{S}}_i = \mathscr{S}_\lambda(\mathbf{M} - \widehat{\mathbf{L}}_i)$$

An approach to combine residuals from $\neq$ algorithms

Regime Switching Model: state of a system / a time series

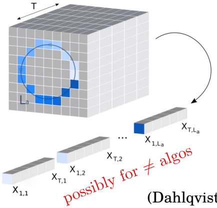
target model: no planet/planet ; *background model*: Gauss/Lap.

- **Set of linear equations describing the RSM model**:

$$\mathbf{X}_{i_a} = \mu + \beta R_{i_a}\mathbf{P} + \epsilon_{s,i_a} = \begin{cases} \mu + \epsilon_{0,i_a} & \text{if } S_{i_a} = 0 \\ \mu + \beta\mathbf{P} + \epsilon_{1,i_a} & \text{if } S_{i_a} = 1 \end{cases}$$

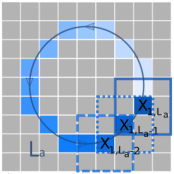
- **Probability ξ_{s,i_a} of $\mathbf{X}_{i_a}$ being in a state $S_{i_a} = s$ at step i_a is**:

$$\xi_{s,i_a} = \mathrm{P}(S_{i_a} | \{\mathbf{X}_{i_a}, \mathbf{X}_{i_a-1}\}, \mathbf{P}, \mu, \beta, \sigma) = \sum_{q=0}^{1} \frac{\eta_{s,i_a} p_{q,s} \xi_{q,i_a-1}}{\sum_{q'=0}^{1}\sum_{s'=0}^{1} \eta_{s',i_a} p_{q',s'} \xi_{q',i_a-1}}$$



(Dahlqvist+, 2020)

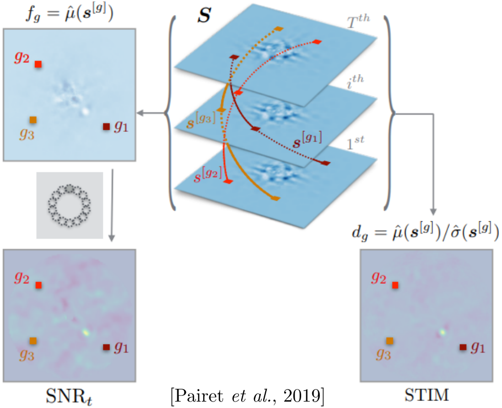
annulus-wise processing



- Iterative inference algorithm to estimate the model parameters
- The probability of being in a state depends on the previous state and on the transition probability (preset)

Different ways to compute SNR from residuals

SNR$_t$: t-test based on small sample statistics (Mawet+, 2014)
versus
STIM: Standardized Trajectory Intensity Mean (Pairet+, 2019)



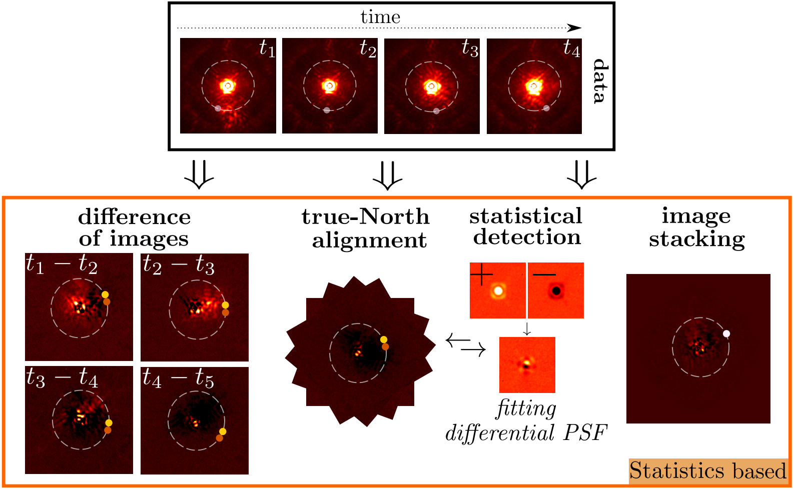
[Pairet *et al.*, 2019]

The absence of explicit computation of SNR map for approaches based on image subtraction/decomposition remains a problem

Statistics-based: *general principle*

Statistics-based: MOODS [Smith+, 2009]; ANDROMEDA [Cantalloube+, 2016]

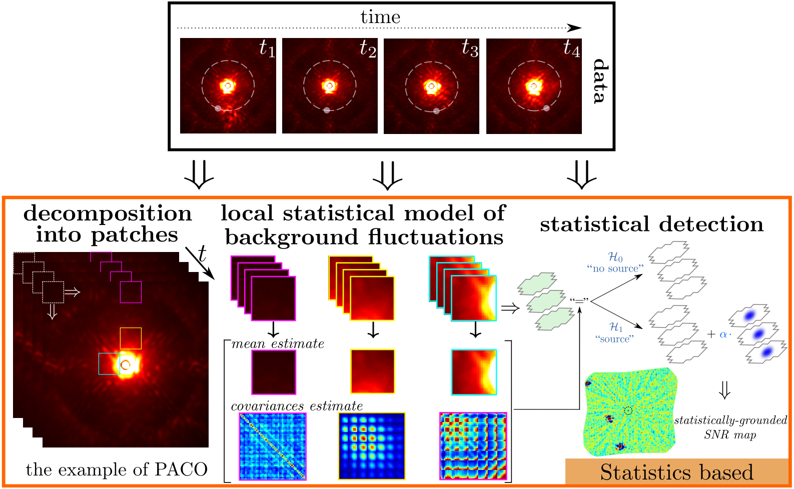
ANDROMEDA general principle:



Statistics-based: *general principle*

Statistics-based: `PACO`, `robust PACO`, `PACO ASDI` [Flasseur+, 2018, 2019, 2020]

`PACO` general principle:

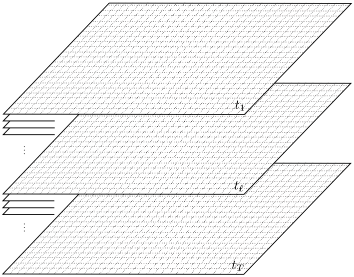


PACO: data-driven exoplanet detection & characterization

PACO (**PA**tch **CO**variances): local learning of the background covariances

PACO principle

- Accounts for background fluctuations
- Local modeling: $\simeq$ 50 pixels/patch
  $\Rightarrow$ **Local adaptivity**
- **Detection**: binary hypothesis test
- **Characterization**: max. likelihood
  - Unbiased astrometry
  - Unbiased photometry



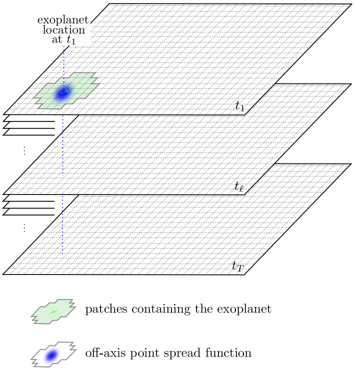
$\Rightarrow$ **Parameter-free algorithm**

PACO: data-driven exoplanet detection & characterization

PACO (PAtch COvariances): local learning of the background covariances

PACO principle

- Accounts for background fluctuations
- Local modeling: $\simeq$ 50 pixels/patch
  $\Rightarrow$ **Local adaptivity**
- **Detection**: binary hypothesis test
- **Characterization**: max. likelihood
  - Unbiased astrometry
  - Unbiased photometry



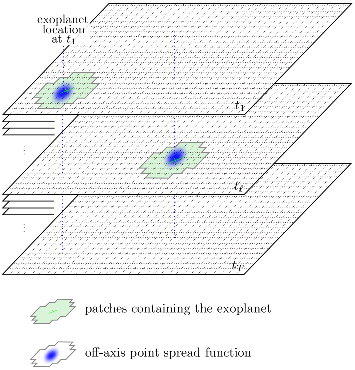
$\Rightarrow$ **Parameter-free algorithm**

PACO: data-driven exoplanet detection & characterization

PACO (**PA**tch **CO**variances): local learning of the background covariances

PACO principle

- Accounts for background fluctuations
- Local modeling: $\simeq$ 50 pixels/patch
  $\Rightarrow$ **Local adaptivity**
- **Detection**: binary hypothesis test
- **Characterization**: max. likelihood
  - Unbiased astrometry
  - Unbiased photometry



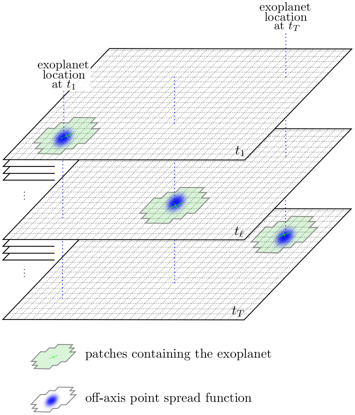
$\Rightarrow$ **Parameter-free algorithm**

PACO: data-driven exoplanet detection & characterization

PACO (**PA**tch **CO**variances): local learning of the background covariances

PACO principle

- Accounts for background fluctuations
- Local modeling: $\simeq$ 50 pixels/patch
  $\Rightarrow$ **Local adaptivity**
- **Detection**: binary hypothesis test
- **Characterization**: max. likelihood
  - Unbiased astrometry
  - Unbiased photometry



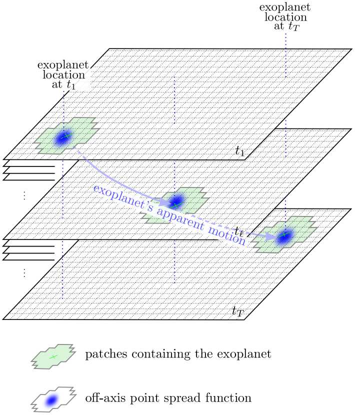
$\Rightarrow$ **Parameter-free algorithm**

PACO: data-driven exoplanet detection & characterization

PACO (**PA**tch **CO**variances): local learning of the background covariances

PACO principle

- Accounts for background fluctuations
- Local modeling: $\simeq$ 50 pixels/patch
  $\Rightarrow$ **Local adaptivity**
- **Detection**: binary hypothesis test
- **Characterization**: max. likelihood
  - Unbiased astrometry
  - Unbiased photometry



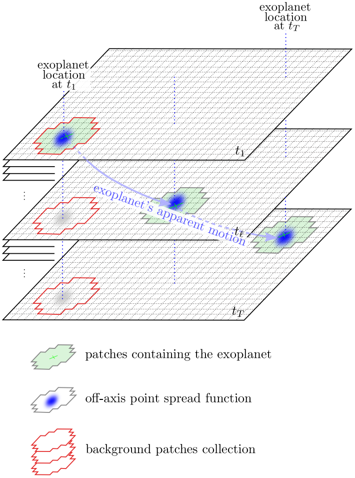
$\boxed{\Rightarrow \textbf{Parameter-free algorithm}}$

PACO: data-driven exoplanet detection & characterization

PACO (**PA**tch **CO**variances): local learning of the background covariances

PACO principle

- Accounts for background fluctuations
- Local modeling: $\simeq$ 50 pixels/patch
  $\Rightarrow$ **Local adaptivity**
- **Detection**: binary hypothesis test
- **Characterization**: max. likelihood
  - Unbiased astrometry
  - Unbiased photometry



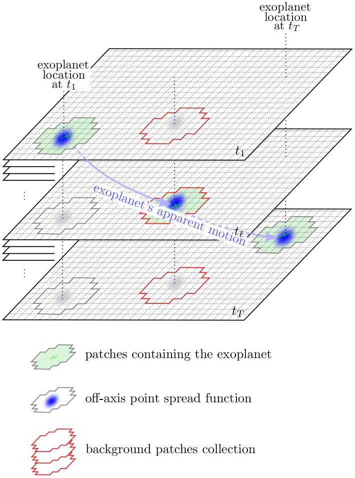
$\Rightarrow$ **Parameter-free algorithm**

PACO: data-driven exoplanet detection & characterization

PACO (**PA**tch **CO**variances): local learning of the background covariances

PACO principle

- Accounts for background fluctuations
- Local modeling: $\simeq$ 50 pixels/patch
  $\Rightarrow$ **Local adaptivity**
- **Detection**: binary hypothesis test
- **Characterization**: max. likelihood
  - Unbiased astrometry
  - Unbiased photometry



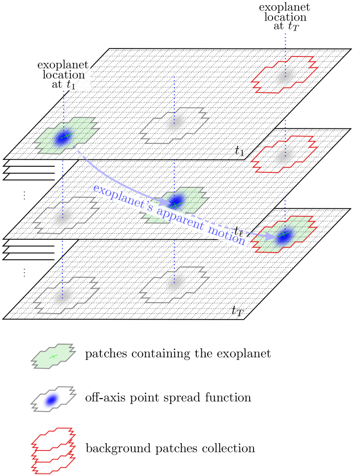
$\Rightarrow$ **Parameter-free algorithm**

PACO: data-driven exoplanet detection & characterization

PACO (**PA**tch **CO**variances): local learning of the background covariances

PACO principle

- Accounts for background fluctuations
- Local modeling: $\simeq$ 50 pixels/patch
  $\Rightarrow$ **Local adaptivity**
- **Detection**: binary hypothesis test
- **Characterization**: max. likelihood
  - Unbiased astrometry
  - Unbiased photometry



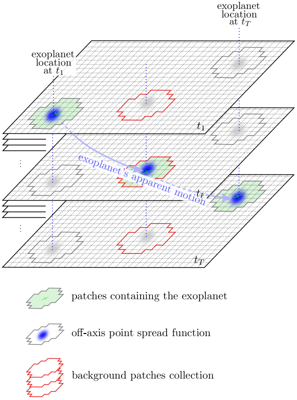
$\Rightarrow$ **Parameter-free algorithm**

PACO: statistical framework

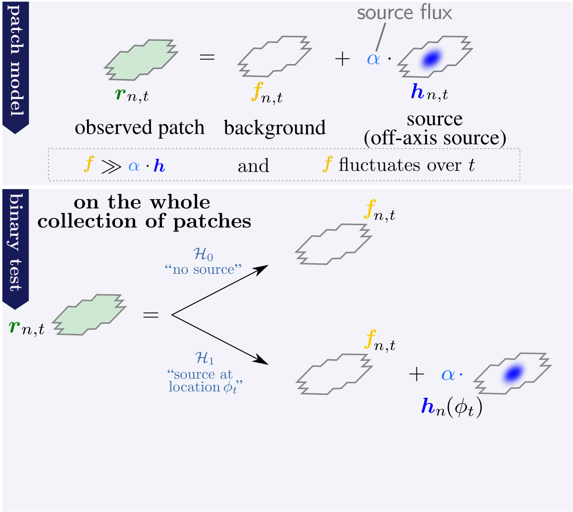
patch model

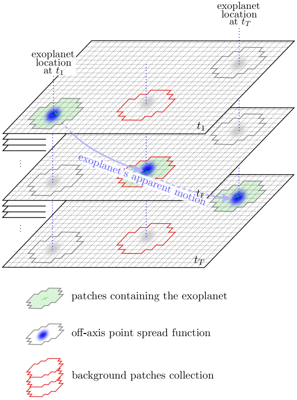
$$\underbrace{\boldsymbol{r}_{n,t}}_{\text{observed patch}} = \underbrace{\boldsymbol{f}_{n,t}}_{\text{background}} + \overbrace{\alpha}^{\text{source flux}} \cdot \underbrace{\boldsymbol{h}_{n,t}}_{\text{source (off-axis source)}}$$

$\boldsymbol{f} \gg \alpha \cdot \boldsymbol{h}$ and $\boldsymbol{f}$ fluctuates over t



PACO: statistical framework





PACO: statistical framework

patch model

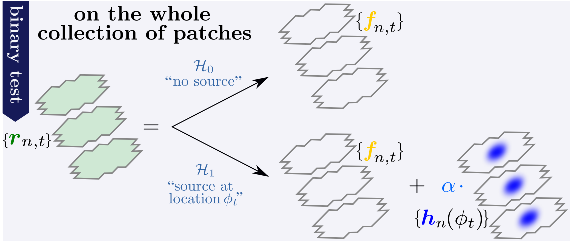
$r_{n,t}$ = $f_{n,t}$ + $\alpha \cdot h_{n,t}$

observed patch — background — source (off-axis source)

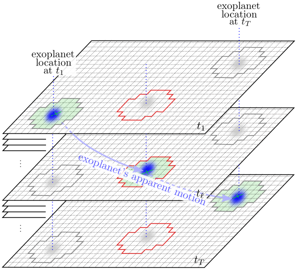
$f \gg \alpha \cdot h$ and f fluctuates over t

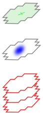
binary test

on the whole collection of patches



GLRT: $$\log \frac{\max\limits_{\alpha \geq 0} \mathscr{L}(\ |\alpha)}{\mathscr{L}(\ |\alpha = 0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta$$





f? ⇒**stat. model**
- learn correlations
- be robust

PACO: modeling the fluctuations of the nuisance component

Statistical model

Gaussian Scale Mixture (GSM) to model a **patch**

patch $\boldsymbol{f}_{n,t}$ at pixel n and time t: $\boldsymbol{f}_{n,t} = \boldsymbol{m}_n + \boxed{\sigma_{n,t}}\boldsymbol{u}_{n,t}$ where $\boldsymbol{u}_{n,t} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_n)$

$$\mathrm{p}_{\boldsymbol{f}}(\{\boldsymbol{f}_{n,t}\}_{t=1:T}) = \prod_{t=1}^{T} \mathcal{N}\left(\boldsymbol{f}_{n,t} \,\middle|\, \boldsymbol{m}_n, \boxed{\{\sigma_{n,t'}\}_{t'=1:T}}, \mathbf{C}_n\right) \quad \text{where} \quad n = \lfloor \phi_t \rceil$$

PACO: modeling the fluctuations of the nuisance component

Statistical model

Gaussian Scale Mixture (GSM) to model a **patch**

patch $\boldsymbol{f}_{n,t}$ at pixel n and time t: $\boldsymbol{f}_{n,t} = \boldsymbol{m}_n + \boxed{\sigma_{n,t}}\boldsymbol{u}_{n,t}$ where $\boldsymbol{u}_{n,t} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_n)$

$$\mathrm{p}_{\boldsymbol{f}}(\{\boldsymbol{f}_{n,t}\}_{t=1:T}) = \prod_{t=1}^{T} \mathcal{N}\left(\boldsymbol{f}_{n,t} \,\middle|\, \boldsymbol{m}_n, \boxed{\{\sigma_{n,t'}\}_{t'=1:T}}, \mathbf{C}_n\right) \quad \text{where} \quad n = \lfloor \phi_t \rceil$$

Statistical learning

Estimated through **fixed-point iterations**:

- Scaling factor: $\boxed{\widehat{\sigma}^2_{n,t}} = (1/K)\,(\boldsymbol{r}_{n,t} - \boldsymbol{m}_{n,t})\mathbf{C}_n^{-1}(\boldsymbol{r}_{n,t} - \boldsymbol{m}_{n,t})^{\mathrm{t}}\,.$
- Sample mean: $\widehat{\boldsymbol{m}}_n = \left(\sum_{t=1}^{T} \boxed{\dfrac{1}{\sigma^2_{n,t}}} \boldsymbol{r}_{n,t}\right) \Big/ \left(\sum_{t=1}^{T} \boxed{\dfrac{1}{\sigma^2_{n,t}}}\right)\,.$
- Sample cov.: $\widehat{\mathbf{S}}_n = \frac{1}{T} \sum_{t=1}^{T} \boxed{\dfrac{1}{\sigma^2_{n,t}}} (\boldsymbol{r}_{n,t} - \boldsymbol{m}_{n,t})(\boldsymbol{r}_{n,t} - \boldsymbol{m}_{n,t})^{\mathrm{t}}\,.$
- *Shrunk* cov. [Ledoit&Wolf, 2004]; [Chen *et al.*, 2010]: $\widehat{\mathbf{C}}_n = (1 - \widehat{\rho}_n)\widehat{\mathbf{S}}_n + \widehat{\rho}_n \widehat{\mathbf{F}}_n\,.$

PACO: modeling the fluctuations of the nuisance component

Statistical model

Gaussian Scale Mixture (GSM) to model a **patch**

patch $\boldsymbol{f}_{n,t}$ at pixel n and time t: $\boldsymbol{f}_{n,t} = \boldsymbol{m}_n + \boxed{\sigma_{n,t}}\,\boldsymbol{u}_{n,t}$ where $\boldsymbol{u}_{n,t} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_n)$

$$\mathrm{p}_f(\{\boldsymbol{f}_{n,t}\}_{t=1:T}) = \prod_{t=1}^{T} \mathcal{N}\left(\boldsymbol{f}_{n,t} \,\middle|\, \boldsymbol{m}_n, \boxed{\{\sigma_{n,t'}\}_{t'=1:T}}, \mathbf{C}_n\right) \quad \text{where} \quad n = \lfloor \phi_t \rceil$$

Statistical learning

Estimated through **fixed-point iterations**:

- Scaling factor: $\boxed{\widehat{\sigma}^2_{n,t}} = (1/K)\,(\boldsymbol{r}_{n,t} - \boldsymbol{m}_{n,t})\mathbf{C}_n^{-1}(\boldsymbol{r}_{n,t} - \boldsymbol{m}_{n,t})^{\mathrm{t}}$.
- Sample mean: $\widehat{\boldsymbol{m}}_n = \left(\sum_{t=1}^{T} \boxed{\dfrac{1}{\sigma^2_{n,t}}}\boldsymbol{r}_{n,t}\right) \Big/ \left(\sum_{t=1}^{T} \boxed{\dfrac{1}{\sigma^2_{n,t}}}\right)$.
- Sample cov.: $\widehat{\mathbf{S}}_n = \frac{1}{T}\sum_{t=1}^{T} \boxed{\dfrac{1}{\sigma^2_{n,t}}}(\boldsymbol{r}_{n,t} - \boldsymbol{m}_{n,t})(\boldsymbol{r}_{n,t} - \boldsymbol{m}_{n,t})^{\mathrm{t}}$.
- *Shrunk* cov. [Ledoit&Wolf, 2004]; [Chen *et al.*, 2010]: $\widehat{\mathbf{C}}_n = (1 - \widehat{\rho}_n)\widehat{\mathbf{S}}_n + \widehat{\rho}_n \widehat{\mathbf{F}}_n$.

PACO: *shrinkage* estimation of covariances

Issue and proposed approach

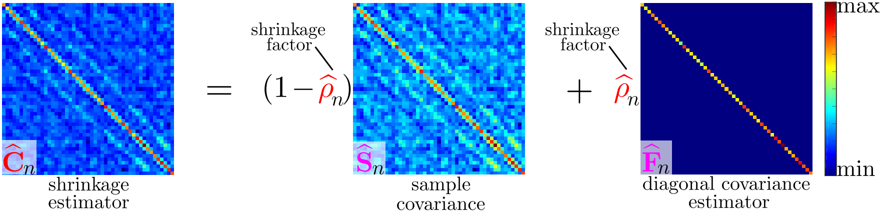
- Limited number of samples $(T \approx K)$ to estimate $\mathbf{C}_n$ $(K \times K)$

  $\Rightarrow \widehat{\mathbf{C}}_n$ is **very noisy** and **rank deficient**.

  A form of ***regularization*** *has to be enforced.*

- *Shrinkage* approach [Ledoit & Wolf, (2004)]; [Chen *et al.*, 2010]

  $\Rightarrow$ **A bias/variance tradeoff: automatic and locally adaptive.**



with $\quad \widehat{\rho}_n = \dfrac{\mathrm{tr}(\widehat{\mathbf{S}}_n^2) + \mathrm{tr}^2(\widehat{\mathbf{S}}_n) - 2\sum_{k=1}^{K} [\widehat{\mathbf{S}}_n]_{kk}^2}{(T+1)(\mathrm{tr}(\widehat{\mathbf{S}}_n^2) - \sum_{k=1}^{K} [\widehat{\mathbf{S}}_n]_{kk}^2)}$

PACO: *shrinkage* estimation of covariances

Issue and proposed approach

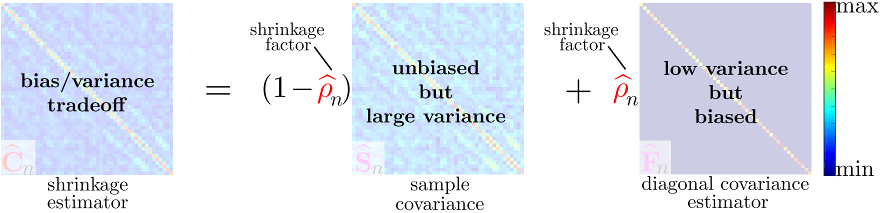
- Limited number of samples ($T \approx K$) to estimate $\mathbf{C}_n$ ($K \times K$)

  $\Rightarrow \widehat{\mathbf{C}}_n$ is **very noisy** and **rank deficient**.

  *A form of **regularization** has to be enforced.*

- *Shrinkage* approach [Ledoit & Wolf, (2004)]; [Chen *et al.*, 2010]

  $\Rightarrow$ **A bias/variance tradeoff: automatic and locally adaptive.**



with $$\widehat{\rho}_n = \frac{\mathrm{tr}(\widehat{\mathbf{S}}_n^2) + \mathrm{tr}^2(\widehat{\mathbf{S}}_n) - 2\sum_{k=1}^{K}[\widehat{\mathbf{S}}_n]_{kk}^2}{(T+1)(\mathrm{tr}(\widehat{\mathbf{S}}_n^2) - \sum_{k=1}^{K}[\widehat{\mathbf{S}}_n]_{kk}^2)}$$

PACO: modeling the fluctuations of the nuisance component

Statistical model

Gaussian Scale Mixture (GSM) to model a **patch**

patch $\boldsymbol{f}_{n,t}$ at pixel n and time t: $\boldsymbol{f}_{n,t} = \boldsymbol{m}_n + \boxed{\sigma_{n,t}}\,\boldsymbol{u}_{n,t}$ where $\boldsymbol{u}_{n,t} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_n)$

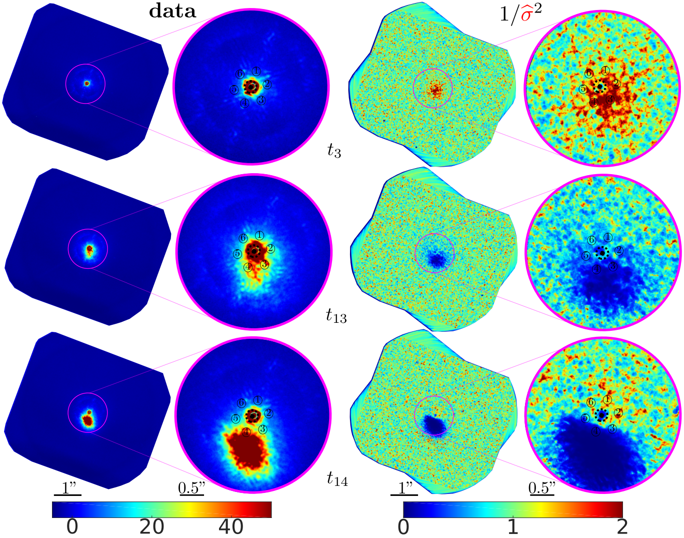
$$\mathrm{p}_f(\{\boldsymbol{f}_{n,t}\}_{t=1:T}) = \prod_{t=1}^{T} \mathcal{N}\left(\boldsymbol{f}_{n,t} \,\middle|\, \boldsymbol{m}_n, \boxed{\{\sigma_{n,t'}\}_{t'=1:T}}, \mathbf{C}_n\right) \quad \text{where} \quad n = \lfloor \phi_t \rceil$$

Statistical learning

Estimated through **fixed-point iterations**:

- Scaling factor: $\boxed{\widehat{\sigma}^2_{n,t}} = (1/K)\,(\boldsymbol{r}_{n,t} - \boldsymbol{m}_{n,t})\mathbf{C}_n^{-1}(\boldsymbol{r}_{n,t} - \boldsymbol{m}_{n,t})^{\mathrm{t}}\,.$
- Sample mean: $\widehat{\boldsymbol{m}}_n = \left(\sum_{t=1}^{T} \boxed{\frac{1}{\sigma^2_{n,t}}} \boldsymbol{r}_{n,t}\right) \Big/ \left(\sum_{t=1}^{T} \boxed{\frac{1}{\sigma^2_{n,t}}}\right)\,.$
- Sample cov.: $\widehat{\mathbf{S}}_n = \frac{1}{T}\sum_{t=1}^{T} \boxed{\frac{1}{\sigma^2_{n,t}}} (\boldsymbol{r}_{n,t} - \boldsymbol{m}_{n,t})(\boldsymbol{r}_{n,t} - \boldsymbol{m}_{n,t})^{\mathrm{t}}\,.$
- *Shrunk* cov. [Ledoit&Wolf, 2004]; [Chen *et al.*, 2010]: $\widehat{\mathbf{C}}_n = (1 - \widehat{\rho}_n)\widehat{\mathbf{S}}_n + \widehat{\rho}_n \widehat{\mathbf{F}}_n\,.$

Weighting maps $1/\widehat{\sigma}^2$



$\Rightarrow$ **impact of large fluctuations is decreased**
$\Rightarrow$ **robustness is improved**

PACO: modeling the fluctuations of the nuisance component

Statistical model

Gaussian Scale Mixture (GSM) to model a **patch**

patch $\boldsymbol{f}_{n,t}$ at pixel n and time t: $\boldsymbol{f}_{n,t} = \boldsymbol{m}_n + \boxed{\sigma_{n,t}}\boldsymbol{u}_{n,t}$ where $\boldsymbol{u}_{n,t} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_n)$

$$\mathrm{p}_{f}(\{\boldsymbol{f}_{n,t}\}_{t=1:T}) = \prod_{t=1}^{T} \mathcal{N}\left(\boldsymbol{f}_{n,t} \,\middle|\, \boldsymbol{m}_n, \boxed{\{\sigma_{n,t'}\}_{t'=1:T}}, \mathbf{C}_n\right) \quad \text{where} \quad n = \lfloor \phi_t \rceil$$

Statistical learning

Estimated through **fixed-point iterations**:

- Scaling factor: $\boxed{\widehat{\sigma}^2_{n,t}} = (1/K)\,(\boldsymbol{r}_{n,t} - \boldsymbol{m}_{n,t})\mathbf{C}_n^{-1}(\boldsymbol{r}_{n,t} - \boldsymbol{m}_{n,t})^{\mathrm{t}}$.
- Sample mean: $\widehat{\boldsymbol{m}}_n = \left(\sum_{t=1}^{T} \boxed{\dfrac{1}{\sigma^2_{n,t}}} \boldsymbol{r}_{n,t}\right) / \left(\sum_{t=1}^{T} \boxed{\dfrac{1}{\sigma^2_{n,t}}}\right)$.
- Sample cov.: $\widehat{\mathbf{S}}_n = \frac{1}{T}\sum_{t=1}^{T} \boxed{\dfrac{1}{\sigma^2_{n,t}}} (\boldsymbol{r}_{n,t} - \boldsymbol{m}_{n,t})(\boldsymbol{r}_{n,t} - \boldsymbol{m}_{n,t})^{\mathrm{t}}$.
- *Shrunk* cov. [Ledoit&Wolf, 2004]; [Chen *et al.*, 2010]: $\widehat{\mathbf{C}}_n = (1 - \widehat{\rho}_n)\widehat{\mathbf{S}}_n + \widehat{\rho}_n \widehat{\mathbf{F}}_n$.

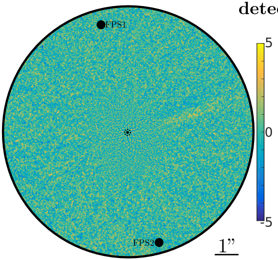
$\Rightarrow$ **Is this model relevant?**

PACO: statistically grounded detection criterion

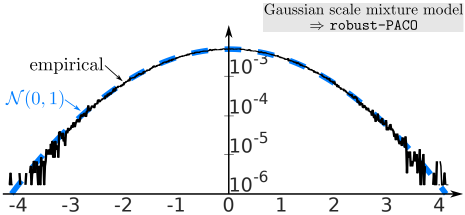
Derivation of the Generalized Likelihood Ratio Test (GLRT)

"background at $n, t \sim \mathcal{N}(\widehat{\boldsymbol{m}}_n, \widehat{\sigma}_{n,t}\widehat{\mathbf{C}}_n)$"

- criterion: $\mathrm{S/N} = \dfrac{\widehat{\alpha}}{\widehat{\sigma}_\alpha} = \dfrac{\sum_{t=1}^{T} \frac{1}{\hat{\sigma}^2{}_{n,t}} \boldsymbol{h}_n(\phi_t)^{\mathrm{t}} \cdot \widehat{\mathbf{C}}_n^{-1} \cdot \left(\boldsymbol{r}_{n,t} - \widehat{\boldsymbol{m}}_n\right)}{\sqrt{\sum_{t=1}^{T} \frac{1}{\hat{\sigma}^2{}_{n,t}} \boldsymbol{h}_n(\phi_t)^{\mathrm{t}} \cdot \widehat{\mathbf{C}}_n^{-1} \cdot \boldsymbol{h}_n(\phi_t)}} \underset{\mathscr{H}_0}{\overset{\mathscr{H}_1}{\gtrless}} \tau$

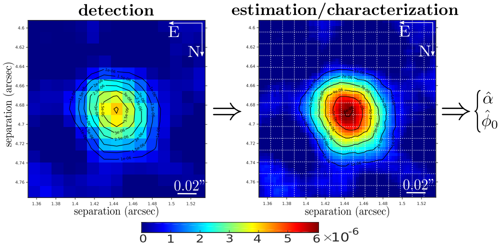


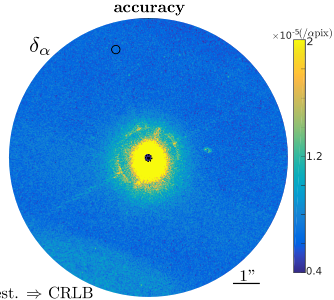
detection criterion in absence of source



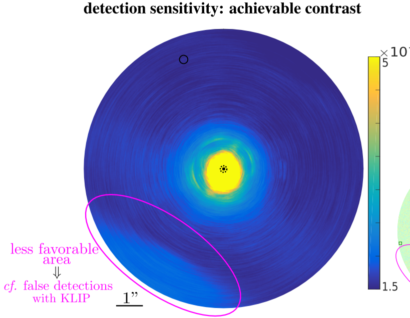
$\boxed{\textbf{S/N follows } \mathcal{N}(0,1) \textbf{ if no source} \Rightarrow \textbf{controlled PFA or FDR}}$

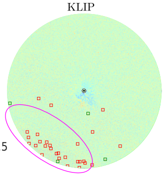
PACO: statistically grounded astro-photometry

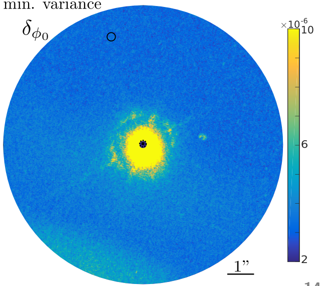




unbiased est. $\Rightarrow$ CRLB
$=$ theoretical min. variance



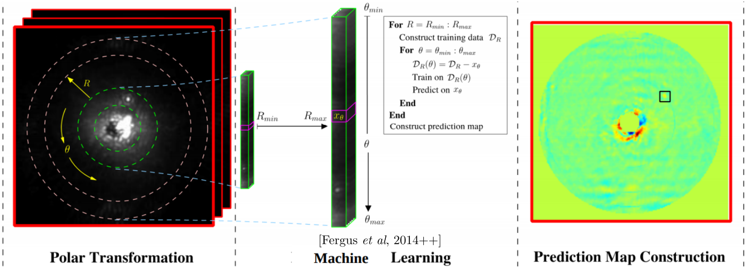




$S4$: a discriminative model based on SVM

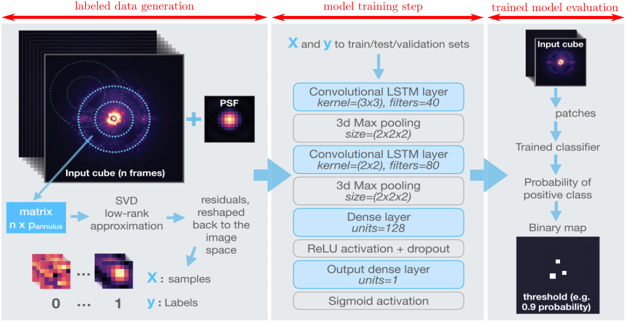
local Spatial-Spectral model for Speckle Suppression (Fergus+, 2014)

- **data representation**:
  - *exploiting radial motion of speckles wrt. wavelength*
  - *patches in polar coordinates (samples: angles × exposures)*
- **model**:
  - *discriminative: SVM-based, + combined with injections*
- **learning & testing**
  - *separating slices within annulus into train/test*
  - *train new model for each location*



SODINN: a discriminative model based on CNN

- **generating labeled data**: injections of fake faint sources
  - *applying a truncated SVD for various ranks k*
  - *forming the labeled groundtruths with the residual patches*
- **training step:** a (deep) discriminative model
  - *a random forest classifier or a CNN*
- **testing step:** testing each location of the FOV

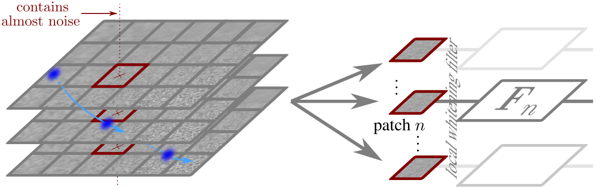


[Gomez Gonzalez et al., 2017]

Initial approach: SODINN/SODIRF (Gonzalez+ 2017). Improvements in progress (Cantero+ in prep.)

deep PACO: a discriminative model based on stats & CNN

Preprocessing: centering and local whitening with PACO model → *to improve the SNR and the stationarity*



- local adaptivity
- parameter-free
- tradeoff between model complexity / data fidelity

Learning: semantic segmentation and regression

→ *supervised training with simulated exoplanets*

samples:

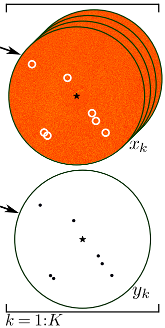
preprocessed images
+
shuffling
+
injected fake exoplanets

data augmentation to deal with lack of groundtruth

groundtruths:

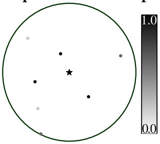
location
+
intensity

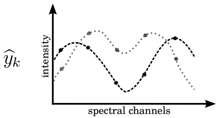
- more flexible model



$\Rightarrow$

	semantic segmentation	regression
	CNN (full frames)	CNN (patches)
	⇓ **output: detection map**	⇓ **output: estimated intensity**

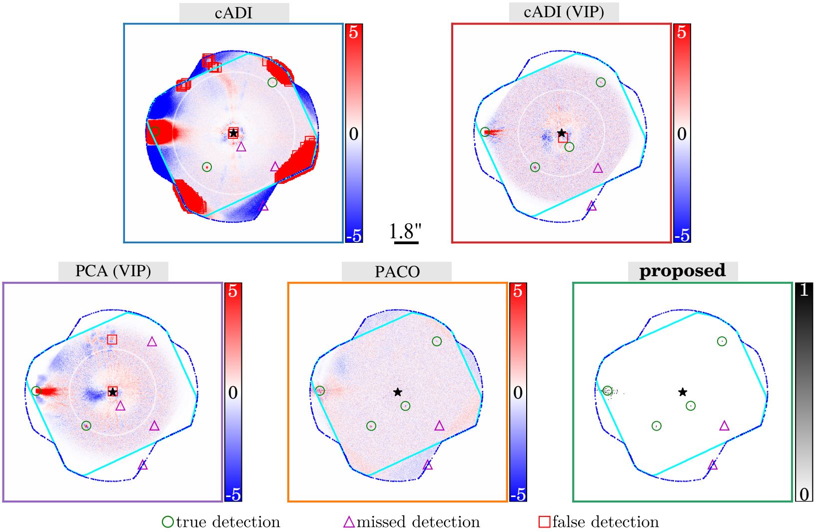




	semantic segmentation	regression
loss:	Dice score (overlap measure)	MSE
validation metric:	F1R score (tradeoff precision/recall)	MSE
architecture:	U-Net (backbone Res-Net18) *trained from scratch*	VGG-like

deep PACO: a discriminative model based on stats & CNN

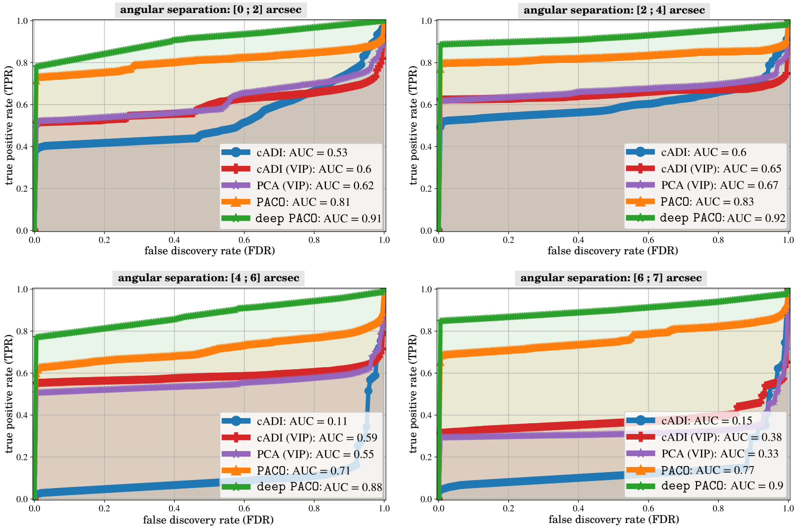
Example of detection maps (HIP 88399, 2015-05-10)



(Flasseur, Bodrito+, 2022)

deep PACO: a discriminative model based on stats & CNN

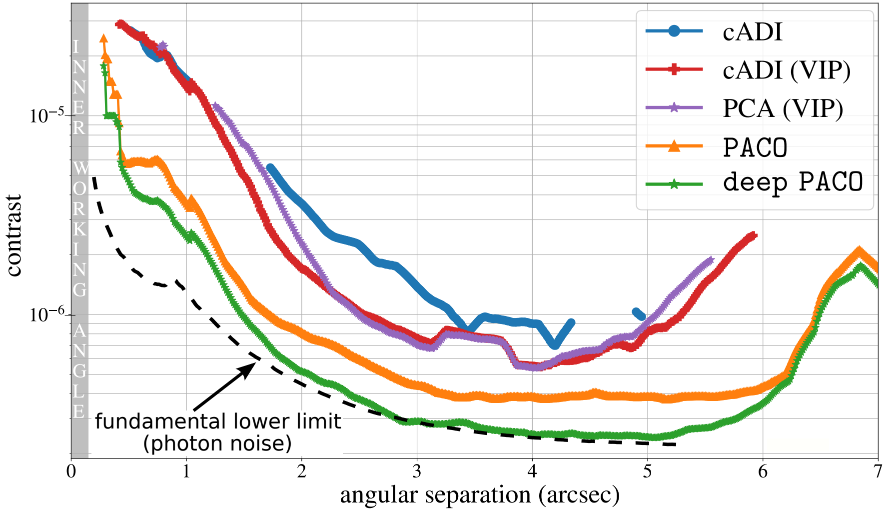
Example of ROCs curves (HIP 88399, 2015-05-10)



(Flasseur, Bodrito+, 2022)

deep PACO: a discriminative model based on stats & CNN

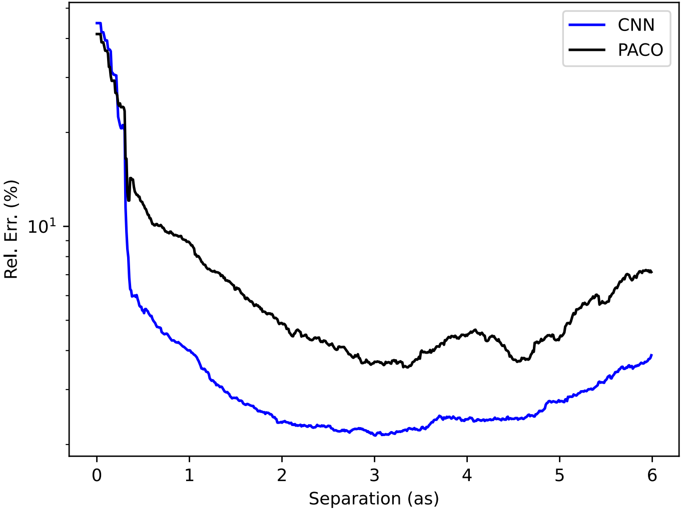
Example of contrast curves (HIP 88399, 2015-05-10)



(Flasseur, Bodrito+, 2022)

deep PACO: a discriminative model based on stats & CNN

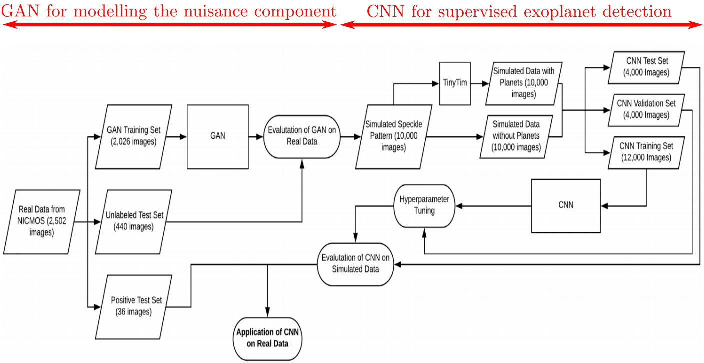
Example of photometric errors (HIP 88399, 2015-05-10)



(Flasseur, Bodrito+, 2022)

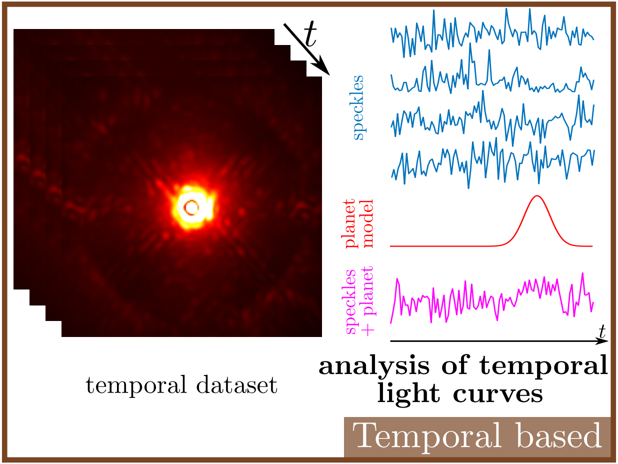
A generative and discriminative approach

- **generating labeled data**: a deep generative model
  - *training a GAN to produce "pure speckles" (prominent component)*
  - *injecting fake faint sources into the generated examples*
- **training step**: a deep discriminative model
  - *training a CNN based on the generated examples*
- **testing step**: testing from an complete image directly



[Yip et al., 2019]

Temporal-based: *general principle*

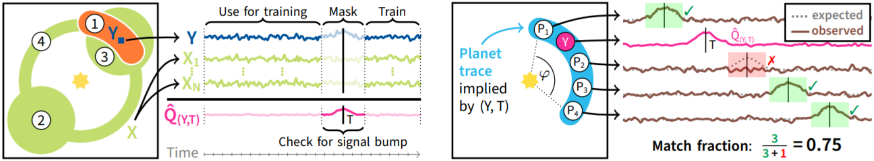


⇒ **Different paradigm: samples are temporal not spatial**

HSR: a temporal approach exploiting metadata

A temporal approach = denoise a time series based on predictors

- **choice of the predictors**: similar pure background time series
  - *excluding the predicted trajectory*
  - *encompassing the close + opposite + annular areas*
  - *exploiting metadata containing information about the systematics*
- **model**: ridge regression
  - *excluding the time samples affected by a putative exoplanet*
- **decision**: (for a single location)
  - *forming the residuals "current time series - prediction"*
  - *checking for signal bump at all time → candidates*
  - *applying a consistent test for each candidate*



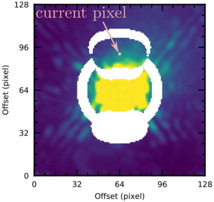
(Gebhard+, 2021, 2022)

TRAP: a closely related approach based on inverse problem

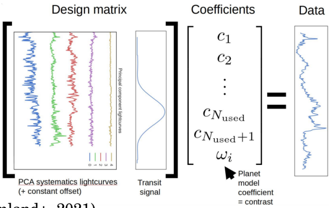
A temporal approach = denoise a time series based on predictors

- **choice of the predictors**: $\simeq$ same criteria
  - *same collections for all set of tested pixels (no masking)*
- **model**: causal regression model
  - *dimensionality reduction: truncated SVD on the set of predictors*
  - *simultaneous fit of the transiting planet + speckles patterns*
- **decision**:
  - *linear system solved analytically (flux + variance)*
  - *heuristic to find the exoplanet flux (weighted mean)*
  - *computation of SNR map*

choice of the predictors



regression



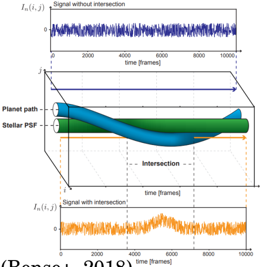
(Samland+, 2021)

Wavelet-based denoising as a pre-conditionner

A temporal approach: frequencies of speckles variations + time dependence

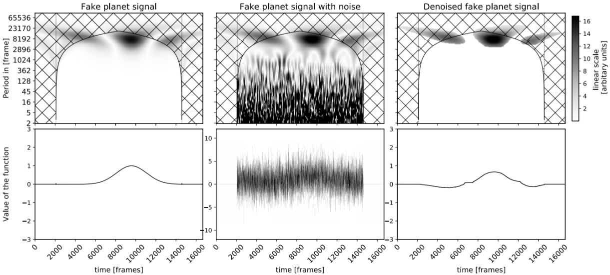
- **multi-level/resolution analysis** (time/frequency)
  - *applying a wavelet transform to each (derotated) time series*
  - *denoising by soft-thresholding (MAD on the 1st wavelet layer)*
  - *applying inverse transform $\rightarrow$ preconditioned data*
  - *applying a space-based detection algorithm*

data model



(Bonse+, 2018)

wavelet denoising = suppression of short-term temporal variations



Conclusions

Different classes of post-processing algorithms for exoplanet detection and characterization:

- *subtraction / decomposition-based*
- $\boxed{\textit{statistics-based}}$
- $\boxed{\textit{learning-based}}$
- $\boxed{\textit{temporal-based}}$

Statistics & temporal approaches

- *currently best tradeoff between sensitivity and confidence*

Learning-based approaches

- *very interesting detection sensitivity...*
- *...but stay "black-box approaches"*

$\Rightarrow$ see brainstorming sessions this afternoon about control of the uncertainties, model-based approaches, and integration of the metadata.

References (1/2)

Subtraction-based:

- R. Soummer et al., "Detection and characterization of exoplanets and disks using projections on Karhunen-Loève eigenimages", The Astrophysical Journal Letters, 755(2), L28, 2012. *[KLIP]*
- A. Amara et al., "PYNPOINT: an image processing package for finding exoplanets", Monthly Notices of the Royal Astronomical Society, 427(2), 2012. *[KLIP]*
- D. Lafrénière et al., "A new algorithm for point-spread function subtraction in high-contrast imaging: a demonstration with angular differential imaging", The Astrophysical Journal, 660(1), 2007. *[LOCI]*
- C. Marois et al., "Exoplanet imaging with LOCI processing: photometry and astrometry with the new SOSIE pipeline", In Adaptive Optics Systems II, SPIE, 7736, 2010. *[LOCI]*
- C. Marois et al., "TLOCI: A fully loaded speckle killing machine", In proceedings of the International Astronomical Union, 8(S299), 2013. *[TLOCI]*
- C. Marois et al., "GPI PSF subtraction with TLOCI: the next evolution in exoplanet/disk high-contrast imaging", In Adaptive Optics Systems IV, SPIE, 9148, 2014. *[TLOCI]*
- L. Pueyo et al., "Detection and characterization of exoplanets using projections on karhunen–loeve eigenimages: Forward modeling", The Astrophysical Journal, 824(2), 2016. *[KLIP-FM]*
- J.-B. Ruffio et al., "Improving and assessing planet sensitivity of the GPI exoplanet survey with a forward model matched filter", The Astrophysical Journal, 842(1), 2017. *[FMMF]*
- B. Pairet et al., "STIM map: detection map for exoplanets imaging beyond asymptotic Gaussian residual speckle noise", Monthly Notices of the Royal Astronomical Society, 487(2), 2019. *[STIM]*

Decomposition-based:

- C. G. Gonzalez et al., "Low-rank plus sparse decomposition for exoplanet detection in direct-imaging ADI sequences - The LLSG algorithm", Astronomy & Astrophysics, 589, A54, 2016. *[LLSG]*
- B. Ren et al., "Non-negative matrix factorization: robust extraction of extended structures", The Astrophysical Journal, 852(2), 104, 2018. *[NMF]*

Statistics-based:

- I. Smith et al., "Detection of a moving source in speckle noise. Application to exoplanet detection", IEEE Transactions on Signal Processing, 57(3), 2009. *[MOODS]*
- F. Cantalloube et al., "Direct exoplanet detection and characterization using the ANDROMEDA method: Performance on VLT/NaCo data", Astronomy & Astrophysics, 582, A89, 2015. *[ANDROMEDA]*
- O. Flasseur et al., "Exoplanet detection in angular differential imaging by statistical learning of the nonstationary patch covariances-The PACO algorithm", Astronomy & Astrophysics, 618, A138, 2018. *[PACO]*
- O. Flasseur et al., "Robustness to bad frames in angular differential imaging: a local weighting approach", Astronomy & Astrophysics, 634, A2, 2019. *[robust PACO]*

References (2/2)

Statistics-based:

- O. Flasseur et al., "PACO ASDI: an algorithm for exoplanet detection and characterization in direct imaging with integral field spectrographs", Astronomy & Astrophysics, 637, A9, 2020. *[PACO ASDI]*
- F. Cantalloube et al., "Direct exoplanet detection and characterization using the ANDROMEDA method: Performance on VLT/NaCo data", Astronomy & Astrophysics, 582, A89, 2015. *[ANDROMEDA]*

Leaning-based:

- R. Fergus et al., "S4: A Spatial-spectral model for Speckle Suppression. The Astrophysical Journal, 794(2), 161, 2014. *[S4]*
- C. G. Gonzalez *et al.*, "Supervised detection of exoplanets in high-contrast imaging sequences", Astronomy & Astrophysics, 613, A71, 2018. *[SODIRF, SODINN]*
- K. H. Yip et al., "Pushing the limits of exoplanet discovery via direct imaging with deep learning", In Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Springer, 2019. **[CNN]**
- T. D. Gebhard *et al.*, "Physically constrained causal noise models for high-contrast imaging of exoplanets", arXiv:2010.05591, 2021. *[HSR]*
- T. D. Gebhard *et al.*, "Half-sibling regression meets exoplanet imaging: PSF modeling and subtraction using a flexible, domain knowledge-driven, causal framework", A&A, 666, A9, 2022'. *[HSR]*
- O. Flasseur *et al.*, "Exoplanet detection in angular differential imaging: combining a statistics-based learning with a deep-based learning for improved detections", SPIE Adaptive Optics Systems, 12185, 2022. *[deep PACO]*

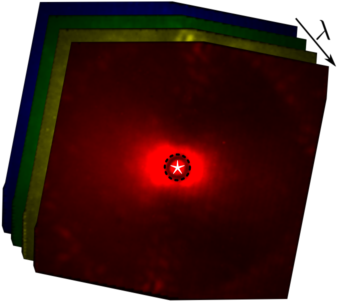
Temporal-based:

- M. Samland et al., "TRAP: A temporal systematics model for improved direct detection of exoplanets at small angular separations", Astronomy & Astrophysics, 646, A24, 2021. *[TRAP]*
- M. J. Bonse et al., "Wavelet based speckle suppression for exoplanet imaging-Application of a de-noising technique in the time domain", arXiv:1804.05063, 2018. *[wavelet]*
- C. H. Dahlqvist et al., "Regime-switching model detection map for direct exoplanet detection in ADI sequences", Astronomy & Astrophysics, 633, A95, 2020. *[RSM]*

Physics-based:

- N. Devaney et al., "PeX 1. Multispectral expansion of residual speckles for planet detection", Monthly Notices of the Royal Astronomical Society, 472(3), 2017. *[PeX]*
- F. Cantalloube et al., "Status of the MEDUSAE post-processing method to detect circumstellar objects in high-contrast multispectral images", arXiv:1812.04312, 2018. *[MEDUSAE]*

Physics-based: *general principle*



spectral dataset

$+$convolution
$+$regularizations
$\Rightarrow$ inverse-problem

chromatic model of speckles based on diffraction

Physics based

$\Rightarrow$ **Different paradigm: samples are spectral not spatial/temporal**

Decomposition-based: *Non-Negative Matrix factorization*

- **constructing basis of components with reference $\mathbf{R}$:**

$$\{\widehat{\mathbf{W}}, \widehat{\mathbf{H}}\} = \text{arg min}_{\mathbf{W},\mathbf{H}} \frac{1}{2}||\mathbf{R} - \mathbf{W}\mathbf{H}||_F^2 \quad \text{s.t.} \quad \text{rank}(\mathbf{W}\mathbf{H}) < \text{rank}(\mathbf{R})$$

$$\widehat{\mathbf{W}}^{(k+1)} = \widehat{\mathbf{W}}^{(k)}.* \left[\mathbf{R}\widehat{\mathbf{H}}^{(k)^{\mathrm{t}}}\right] ./ \left[\widehat{\mathbf{W}}^{(k)}\widehat{\mathbf{H}}^{(k)}\widehat{\mathbf{H}}^{(k)^{\mathrm{t}}}\right]$$

$$\widehat{\mathbf{H}}^{(k+1)} = \widehat{\mathbf{H}}^{(k)}.* \left[\widehat{\mathbf{W}}^{(k)^{\mathrm{t}}}\mathbf{R}\right] ./ \left[\widehat{\mathbf{W}}^{(k)^{\mathrm{t}}}\widehat{\mathbf{W}}^{(k)}\widehat{\mathbf{H}}^{(k)}\right]$$

- **modeling any target $\mathbf{M}$ with the component basis $\widehat{\mathbf{H}}$:**

$$\text{scaling: } \widehat{w}^{(k+1)} = \widehat{w}^{(k)}.* \left[\mathbf{M}\widehat{\mathbf{H}}^{\mathrm{t}}\right] ./ \left[\widehat{w}^{(k)}\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathrm{t}}\right]$$

$$\text{projection: } \mathbf{M}_{\text{NMF}} = \widehat{w}\widehat{\mathbf{H}}$$

- **NMF $\neq$ KLIP, LLSG:**

NMF does not remove the mean of every image + entries non-negative

$\Rightarrow$ non-orthogonal component basis

projection is iterative

$\Rightarrow$ finding a non-neg combination of components to model $\mathbf{M}$