THE PACOME ALGORITHM

Optimal multi-epoch combination of direct imaging observations for exoplanet detection

J. Dallant[1], M. Langlois[1], É. Thiébaut[1], O. Flasseur[2]

[1]CRAL, UMR 5574, CNRS, Université de Lyon, Saint-Genis-Laval, France
[2]LESIA, Université de Paris, Meudon, France

COBREX week 2022

Outline

Table of Contents

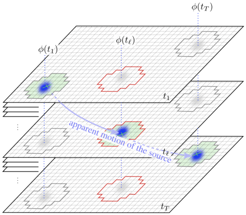
Introduction

Context

- direct imaging exoplanets detection → very challenging
  - very high star/companion contrast (10^4 to 10^8)
  - strong spatial & spectral correlations, stellar leakages + photon noise
- combining several observations → boost the exoplanet detection sensitivity

The PACO algorithm (Flasseur et al. 2018, 2020)

- cutting-edge ADI source detection algorithm
- learns the statistical model of the background from the data (in small local patches)
- achieves excellent detection performances
- sometimes not enough (for extremely faint signals)



The new approach

New PACO-based source detection method that optimally combines multi-epoch observations (assuming source's Keplerian motion)

→ The PACOME algorithm (PACO Multi-Epoch)

- detects exoplanets (with improved sensitivity)
- estimates simultaneously their orbital elements
- benefits from PACO's excellent sensitivity

THE PACOME ALGORITHM

(Dallant et al. in prep)

Table of Contents

Table of Contents

The PACOME algorithm

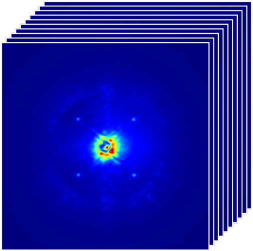
Mathematical formalism

Data $\boldsymbol{r}_{t,\ell,k}$ at time t, spectral channel ℓ and frame k:

$$\boldsymbol{r}_{t,\ell,k} = \alpha_{t,\ell}\, \boldsymbol{h}_{t,\ell,k}(\boldsymbol{\theta}_t(\boldsymbol{\mu})) + \boldsymbol{f}_{t,\ell,k}$$

data · flux · PSF · 2-D position · nuisance · orbital elements (to estimate)

$\{\boldsymbol{r}_{t,\ell,k}\}_{k=1:K}$



Nuisance term (stellar leakages & noise):

- considered Gaussian $\sim \mathcal{N}(\bar{\boldsymbol{f}}_{t,\ell}, \boldsymbol{\Sigma}_{t,\ell})$
- independent of the epoch of observation t and frame k

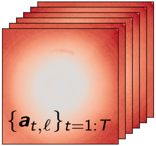
Total log-likelihood of the data given our model:

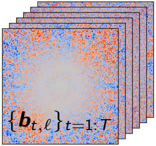
$$\mathcal{L}_\ell(\boldsymbol{\alpha}, \boldsymbol{\mu}) = \text{cste} - \frac{1}{2} \sum_{t,k} \left\| \boldsymbol{r}_{t,\ell,k} - \alpha_{t,\ell}\, \boldsymbol{h}_{t,\ell,k}(\boldsymbol{\theta}_t(\boldsymbol{\mu})) - \bar{\boldsymbol{f}}_{t,\ell} \right\|^2_{\boldsymbol{\Sigma}^{-1}_{t,\ell}}$$

$$= \text{cste} + \sum_t \left[\alpha_{t,\ell}\, \boldsymbol{b}_{t,\ell}(\boldsymbol{\theta}_t(\boldsymbol{\mu})) - \frac{1}{2}\, \alpha^2_{t,\ell}\, \boldsymbol{a}_{t,\ell}(\boldsymbol{\theta}_t(\boldsymbol{\mu})) \right]$$

The PACOME algorithm

Mathematical formalism





$\boldsymbol{a}_{t,\ell}$ and $\boldsymbol{b}_{t,\ell}$ are pre-calculated by the PACO algorithm as:

$$\boldsymbol{a}_{t,\ell}(\boldsymbol{\theta}_t(\boldsymbol{\mu})) = \sum_k \boldsymbol{h}_{t,\ell,k}(\boldsymbol{\theta}_t(\boldsymbol{\mu}))^{\mathsf{T}}\, \boldsymbol{\Sigma}_{t,\ell}^{-1}\, h_{t,\ell,k}(\boldsymbol{\theta}_t(\boldsymbol{\mu}))$$

$$\boldsymbol{b}_{t,\ell}(\boldsymbol{\theta}_t(\boldsymbol{\mu})) = \sum_k \boldsymbol{h}_{t,\ell,k}(\boldsymbol{\theta}_t(\boldsymbol{\mu}))^{\mathsf{T}}\, \boldsymbol{\Sigma}_{t,\ell}^{-1}\, (\boldsymbol{r}_{t,\ell,k} - \bar{\boldsymbol{f}}_{t,\ell})$$

Deriving $\mathcal{L}_\ell(\boldsymbol{\alpha}, \boldsymbol{\mu})$ → **source flux estimator** for each epoch

$$\widehat{\alpha}_{t,\ell} = \underset{\alpha_{t,\ell}\geq 0}{\arg\max}\ \mathcal{L}_\ell(\boldsymbol{\alpha}, \boldsymbol{\mu}) = \frac{\left[\boldsymbol{b}_{t,\ell}(\boldsymbol{\theta}_t(\boldsymbol{\mu}))\right]_+}{\boldsymbol{a}_{t,\ell}(\boldsymbol{\theta}_t(\boldsymbol{\mu}))}$$

Injecting $\widehat{\alpha}_{t,\ell}$ in $\mathcal{L}_\ell(\boldsymbol{\alpha}, \boldsymbol{\mu})$ and re-deriving → **orbital elements estimator**

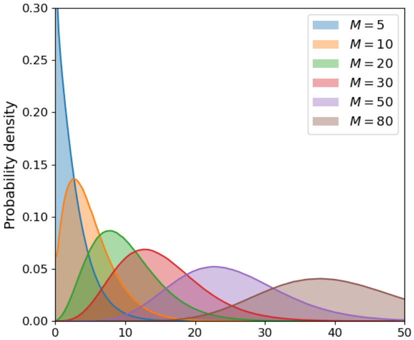
$$\widehat{\boldsymbol{\mu}}_\ell = \underset{\boldsymbol{\mu}}{\arg\max} \sum_t \frac{\left[\boldsymbol{b}_{t,\ell}(\boldsymbol{\theta}_t(\boldsymbol{\mu}))\right]_+^2}{\boldsymbol{a}_{t,\ell}(\boldsymbol{\theta}_t(\boldsymbol{\mu}))}$$

$$\mathcal{C}_\ell(\boldsymbol{\mu}) = \sum_t \frac{\left[\boldsymbol{b}_{t,\ell}(\boldsymbol{\theta}_t(\boldsymbol{\mu}))\right]_+^2}{\boldsymbol{a}_{t,\ell}(\boldsymbol{\theta}_t(\boldsymbol{\mu}))} \Leftrightarrow \mathrm{SNR}_\ell(\boldsymbol{\mu}) = \sqrt{\sum_t \frac{\left[\boldsymbol{b}_{t,\ell}(\boldsymbol{\theta}_t(\boldsymbol{\mu}))\right]_+^2}{\boldsymbol{a}_{t,\ell}(\boldsymbol{\theta}_t(\boldsymbol{\mu}))}}$$

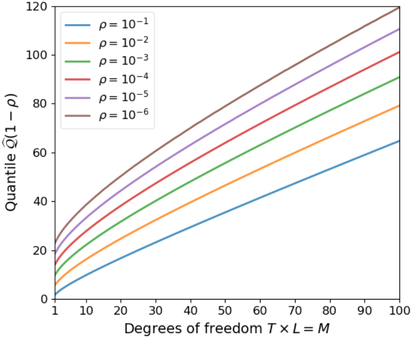
The PACOME algorithm

Statistical guarantees

- Theoretical distribution of the criterion $\mathcal{C}_\ell \equiv \sum$ (rectified Normal distrib.)^2
- Possible to assess the statistical relevance of multi-epoch detection



PDF of the criterion given different degrees of freedom M.

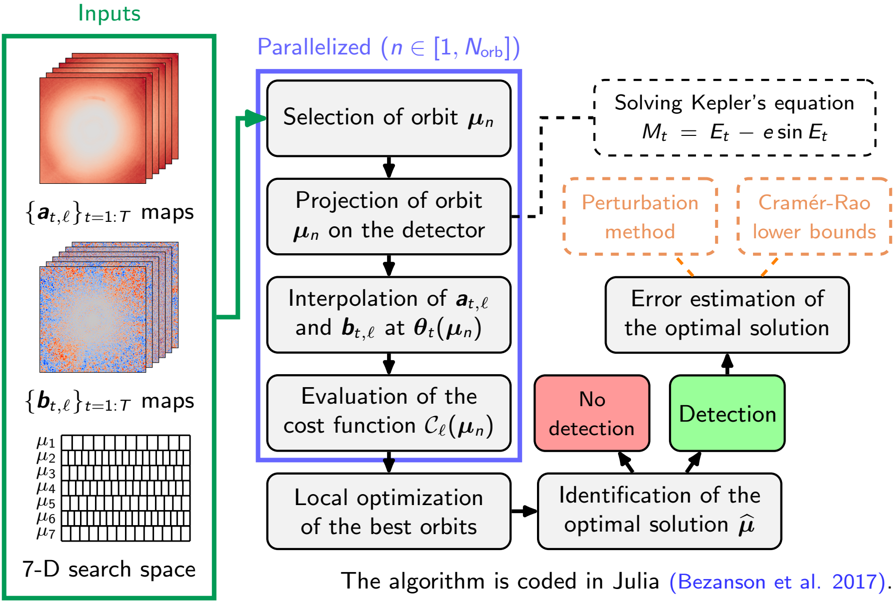


Empirical quantile upper bound for different degrees of freedom M and confidence levels $1 - \rho$.

Table of Contents

The PACOME algorithm

Schematic diagram



The algorithm is coded in Julia (Bezanson et al. 2017).

Table of Contents

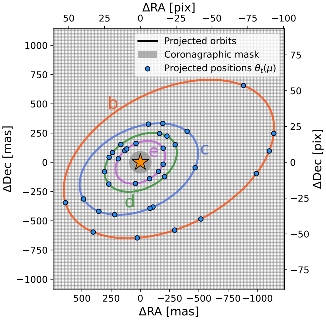
Results on injected sources

Fake orbits injections

Injections of 4 sources at 9 $\neq$ epochs

Data	ADI
Instru.	SPHERE/IRDIS (Beuzit et al. 2019)
Ref.	HR 8799 @ 2016-11-18
Band	DB H23
N_{sources}	4
N_{epochs}	9
$t_{\text{exp-time}}$	$\sim$ 5h

Injection data



Injected orbits

Injected fluxes $< 5\sigma$ limit on average

Source	$\overline{\textbf{SNR}}_{t,\ell}$	$\overline{\alpha}_{t,\ell}$
b	4.6	1.8×10^{-6}
c	2.5	1.7×10^{-6}
d	2.4	2.9×10^{-6}
e	2.17	8.8×10^{-6}

Average injected SNR & flux per epoch

Search strategy:
Find Brighter, Mask & Restart

Interp.	Catmull-Rom
ℓ	1 (H2)
N_{orb}	$\sim 42 \times 10^9 (\times 4)$
N_{threads}	12
t_{exec}	11h$\times$4 (CPU)

Settings of the PACOME run

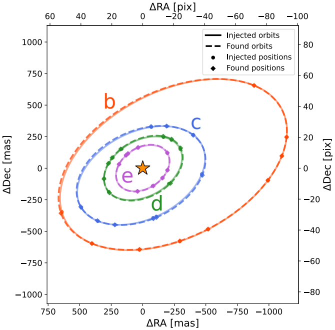
Results on injected sources

Optimal orbits found with PACOME

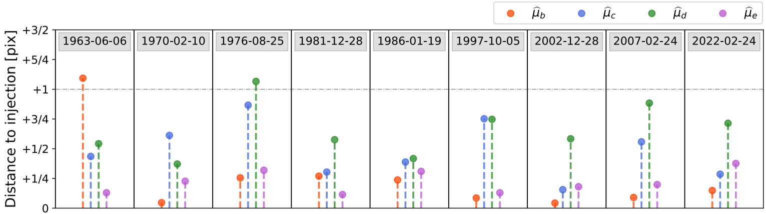
All optimal orbits $\widehat{\boldsymbol{\mu}}$:
- < 1 pix from the injected orbits (**RMSD**)
- $\mathcal{C} > 1 - 10^{-6}$ confidence level

		$\widehat{\boldsymbol{\mu}}_b$	$\widehat{\boldsymbol{\mu}}_c$	$\widehat{\boldsymbol{\mu}}_d$	$\widehat{\boldsymbol{\mu}}_e$
$\overline{\text{SNR}_t}$	-	4.63	2.49	2.40	2.20
$\mathcal{C}$	-	222.67	59.53	57.49	47.75
SNR	-	14.92	7.72	7.58	6.91
RMSD	**pix**	**0.40**	**0.53**	**0.69**	**0.24**
$\widehat{\mathcal{Q}}_{\mathcal{C}}(1-10^{-6})$		37.13 ± 0.25			

Optimal solution $\widehat{\mu}$ found for each planet



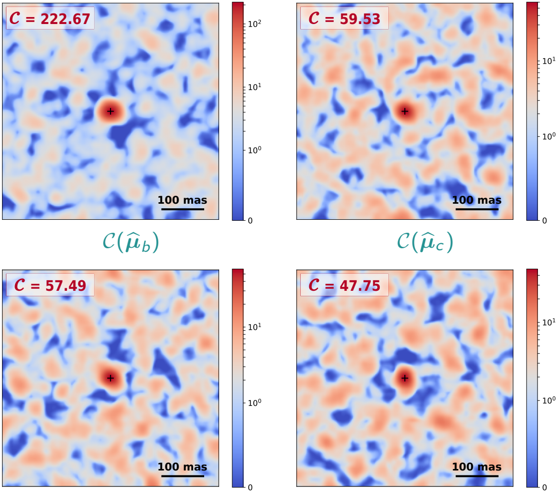
Projections of the injected orbits and PACOME solutions



Distance (per epoch) between the injections and the best solutions found by PACOME

Maps of the criterion

Cost function map around the best solutions (9 epochs combined) :



$\mathcal{C}(\widehat{\boldsymbol{\mu}}_b)$ $\mathcal{C}(\widehat{\boldsymbol{\mu}}_c)$

$\mathcal{C}(\widehat{\boldsymbol{\mu}}_d)$ $\mathcal{C}(\widehat{\boldsymbol{\mu}}_e)$

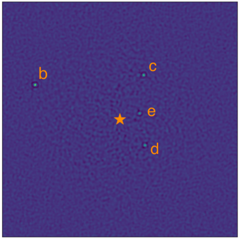
→ **We detect planets that were not detectable on any individual epoch**

Table of Contents

Results on the HR 8799 system

Setup of the PACOME run

Aim: Find all 4 known planets in one large blind search covering all the field



The HR 8799 system
(2017-06-14)

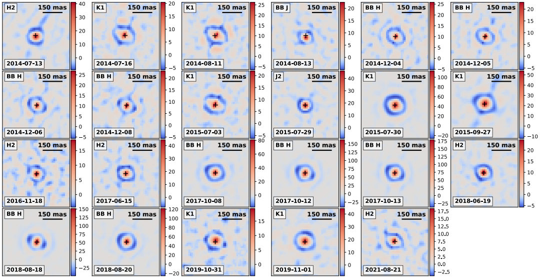
Interp.	Lanczos
$\boldsymbol{\ell}$	1
$\boldsymbol{N}_{\mathrm{orb}}$	$\sim 21 \times 10^{9}$
$\boldsymbol{N}_{\mathrm{threads}}$	12
$\boldsymbol{t}_{\mathrm{exec}}$	11h (CPU)

Settings of the run

Data	ADI
Instru.	SPHERE/IRDIS
Bands	H23, K12, BB J, BB H
$\boldsymbol{N}_{\mathrm{epochs}}$	23
$\boldsymbol{t}_{\mathrm{exp\text{-}time}}$	$\sim$ 37h

HR 8799 data

Best on-grid solution → **Planet b**

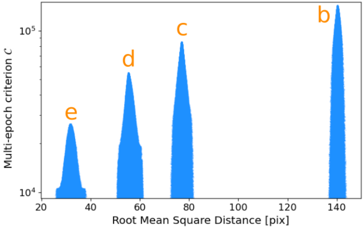


Individual SNR_t maps centered on $\widehat{\boldsymbol{\mu}}$

$\min(\mathrm{SNR}_t) = 12.5$ / $\overline{\mathrm{SNR}_t} = 57.0$ / $\max(\mathrm{SNR}_t) = 180.7$

Results on the HR 8799 system

Unravelling all four exoplanets



Cost function of all explored orbits vs. their RMS distance to the stellar center

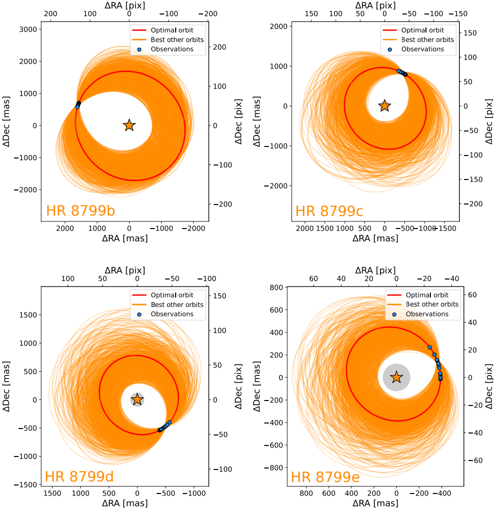
Adding **constraints** of:

- **coplanarity** (i, Ω)
- near **1:2:4:8 resonance***
- same **stellar mass** (K)**

*Gozdziewski et Migaszewski 2020

**Sepulveda et Bowler 2022

Optimal orbits of each peak & the next 1000 best orbits with inclination $20° < i < 30°$
(agree with literature (Zurlo et al. 2016, Wang et al. 2018))



→ We keep the quadruplet of orbits maximising $\mathcal{C}$ and satisfying the constraints (–)

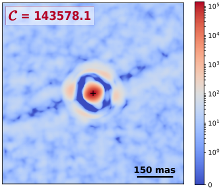
Results on the HR 8799 system

Maps of the criterion

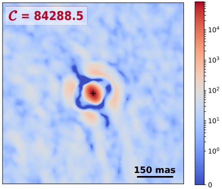
	$\widehat{\boldsymbol{\mu}}_b$	$\widehat{\boldsymbol{\mu}}_c$	$\widehat{\boldsymbol{\mu}}_d$	$\widehat{\boldsymbol{\mu}}_e$
$\mathcal{C}$	14.4×10^4	8.4×10^4	5.5×10^4	2.6×10^4
SNR	379.5	289.8	234.5	161.2
$\widehat{\mathcal{Q}}_{\mathcal{C}}(1 - 10^{-6})$	53.5 ± 0.3			

$\mathcal{C}$ values of the optimal solutions found with PACOME

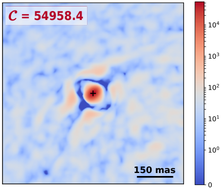
Cost function map around the best solutions:



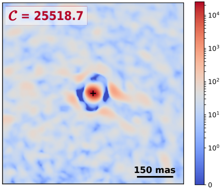
$\mathcal{C}(\widehat{\boldsymbol{\mu}}_b)$



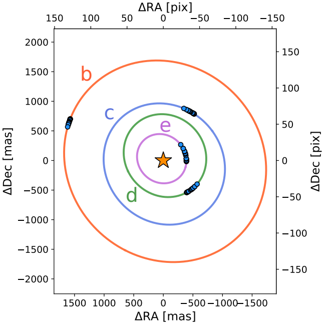
$\mathcal{C}(\widehat{\boldsymbol{\mu}}_c)$



$\mathcal{C}(\widehat{\boldsymbol{\mu}}_d)$



$\mathcal{C}(\widehat{\boldsymbol{\mu}}_e)$



Optimal orbits found with PACOME

Table of Contents

Conclusion

The algorithm...

- combines multi-epoch direct-imaging observations
- **increases the detection sensitivity** of keplerian sources ($\sim \sqrt{N_{\text{epochs}}}$)
- **estimates** simultaneously their **orbital elements** (best scheduling of future obs.)
- is **optimal** in the **maximum likelihood** sense
- is very **fast**

Future prospects

Enhancing the algorithm

- smarter sampling strategy (on-search grid refinement)
- data pre-processing & data flux calibration (waffles)

Switching to IRDIS/ASDI and IFS data

- account for the spectral correlations and improve the detection further

Applying PACOME on other systems

- Re-exploring HR 8799 to search for a 5^{th} planet (Wahhaj et al. 2021)
- Exploring HD 95086 & β Pictoris

Questions ? ☺