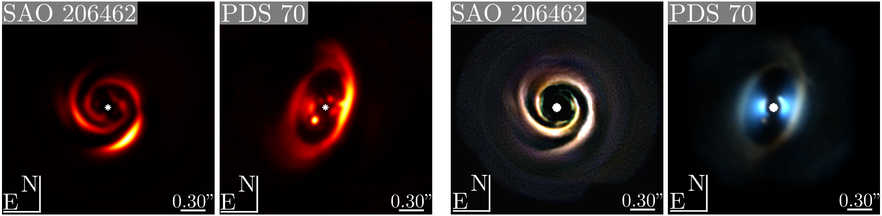
Post-processing algorithms for high contrast reconstruction of the circumstellar environment by angular (and spectral) differential imaging

A focus on inverse problem approaches

Olivier Flasseur

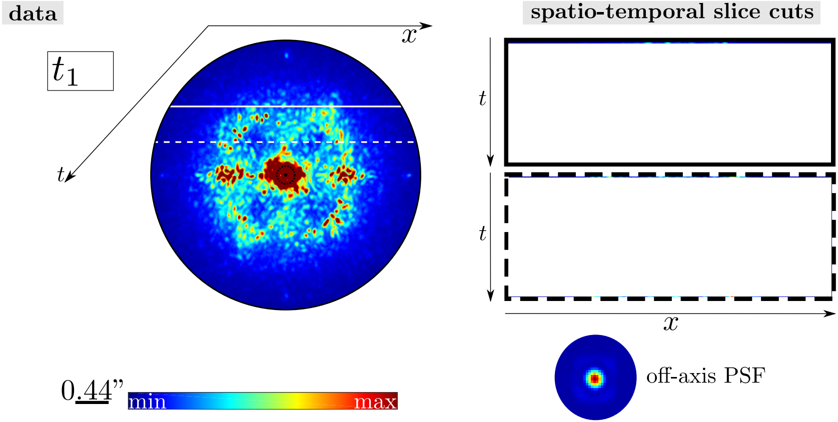
ADI (VLT/SPHERE-IRDIS) **ASDI (VLT/SPHERE-IFS)**



Workshop COBREX, 3rd October 2022

Context: typical dataset from VLT/SPHERE instrument

angular differential imaging (ADI) = temporal diversity



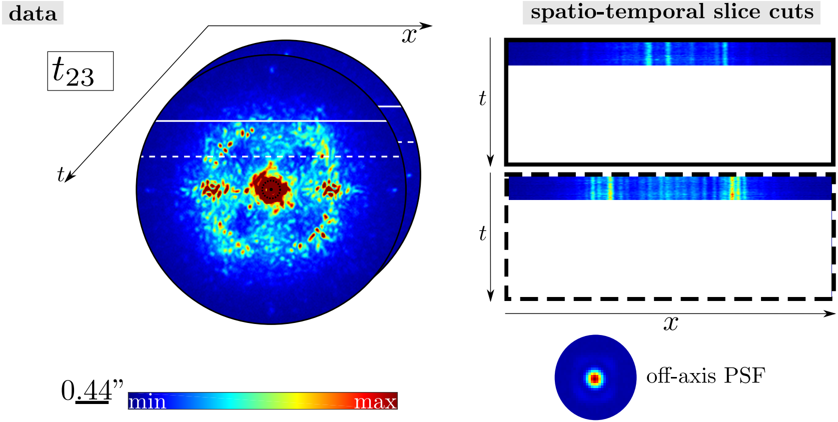
Specificities

- Disk (and exoplanet) **signal** stays **weak**
- **Non-stationary** and **multi-correlated** nuisance component

$\Rightarrow$ **Unmixing through signal processing is mandatory** $\Leftarrow$

Context: typical dataset from VLT/SPHERE instrument

angular differential imaging (ADI) = temporal diversity



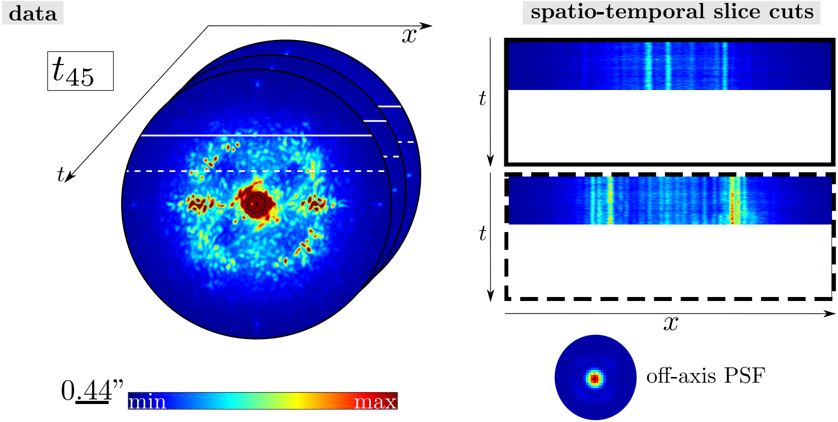
Specificities

- Disk (and exoplanet) **signal** stays **weak**
- **Non-stationary** and **multi-correlated** nuisance component

⇒ **Unmixing through signal processing is mandatory** ⇐

Context: typical dataset from VLT/SPHERE instrument

angular differential imaging (ADI) = temporal diversity



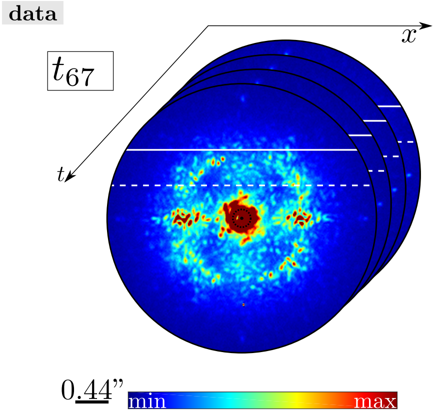
Specificities

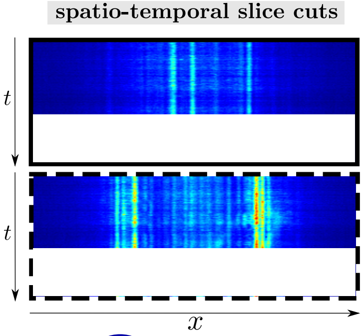
- Disk (and exoplanet) **signal** stays **weak**
- **Non-stationary** and **multi-correlated** nuisance component

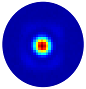
⇒ **Unmixing through signal processing is mandatory** ⇐

Context: typical dataset from VLT/SPHERE instrument

angular differential imaging (ADI) = temporal diversity







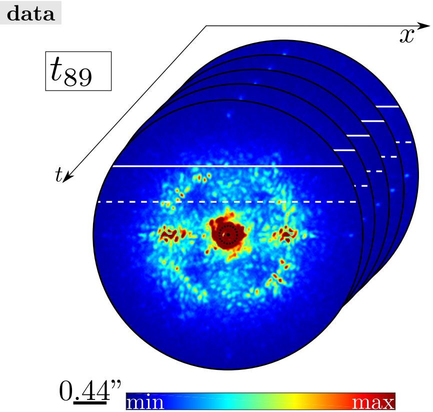
Specificities

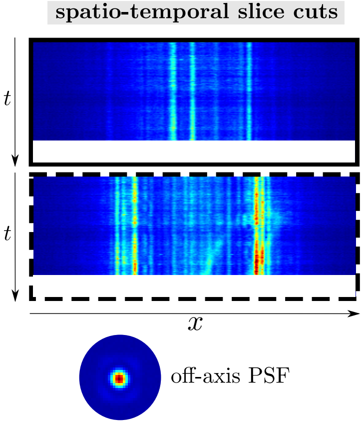
- Disk (and exoplanet) **signal** stays **weak**
- **Non-stationary** and **multi-correlated** nuisance component

$\Rightarrow$ **Unmixing through signal processing is mandatory** $\Leftarrow$

Context: typical dataset from VLT/SPHERE instrument

angular differential imaging (ADI) = temporal diversity





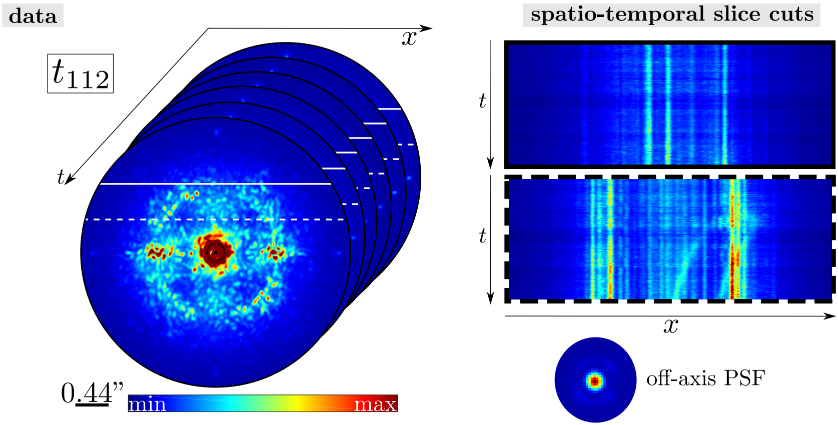
Specificities

- Disk (and exoplanet) **signal** stays **weak**
- **Non-stationary** and **multi-correlated** nuisance component

⇒ **Unmixing through signal processing is mandatory** ⇐

Context: typical dataset from VLT/SPHERE instrument

angular differential imaging (ADI) = temporal diversity



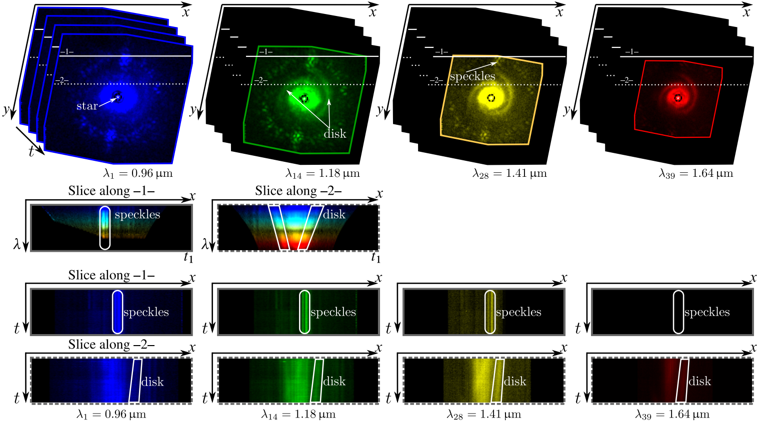
Specificities

- Disk (and exoplanet) **signal** stays **weak**
- **Non-stationary** and **multi-correlated** nuisance component

$\Rightarrow$ **Unmixing through signal processing is mandatory** $\Leftarrow$

Context: typical dataset from VLT/SPHERE instrument

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



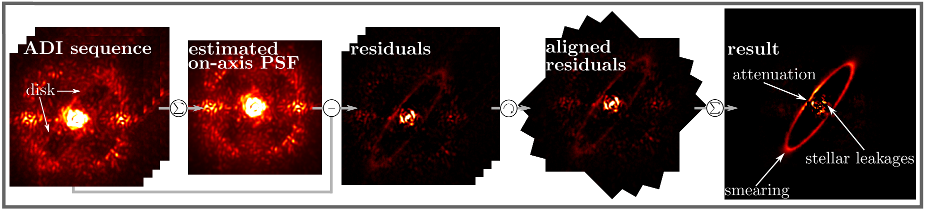
Specificities

- Disk (and exoplanet) **signal** stays **weak**
- **Non-stationary** and **multi-correlated** nuisance component

$\Rightarrow$ **Unmixing through signal processing is mandatory** $\Leftarrow$

Different categories of algorithms for disk reconstruction

The classical pipeline:



Key step: estimation of the on-axis PSF

- median or mean: cADI (Marois+, 2006), and many variants
- linear combination: {T, M, A}-LOCI (Marois+, 2014), (Wahhaj+, 2015)
- principal component analysis: KLIP (Soummer+, 2012), (Amara+, 2012)

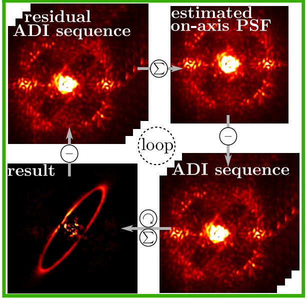
Limitations

- no explicit modeling of the nuisance component
  $\Rightarrow$ **high residual stellar leakages**
- no explicit modeling of the image formation process
  $\Rightarrow$ **high morphological and photometric distorsions**

Different categories of algorithms for disk reconstruction

More advanced algorithms:

artifacts mitigation without reference

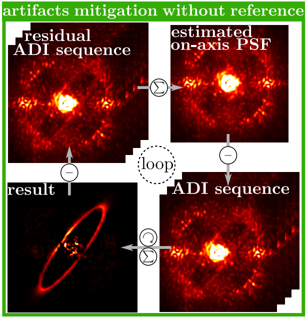


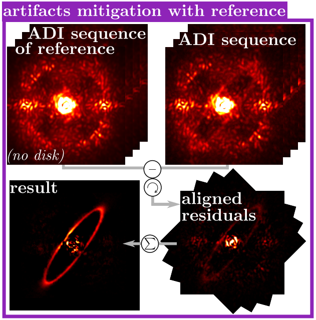
iterative PCA
(Pairet+, 2018)

data imputation strategy
(Ren+, 2020)

Different categories of algorithms for disk reconstruction

More advanced algorithms:





iterative PCA
(Pairet+, 2018)

data imputation strategy
(Ren+, 2020)

see Julien's & Sophia's focus

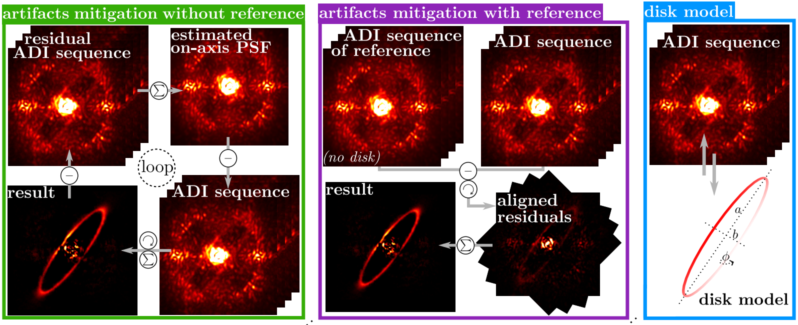
reference differential imaging (RDI)
= searching for similarities in images

RDI with a large library
(Gerard+, 2016) (Ren+, 2018)
(Xuan+, 2018) (Ruane+, 2019)

RDI with star hopping (Wahhaj+, 2021)

Different categories of algorithms for disk reconstruction

More advanced algorithms:



iterative PCA
(Pairet+, 2018)

data imputation strategy
(Ren+, 2020)

see Julien's & Sophia's focus

reference differential imaging (RDI)
= searching for similarities in images

RDI with a large library
(Gerard+, 2016) (Ren+, 2018)
(Xuan+, 2018) (Ruane+, 2019)

RDI with star hopping (Wahhaj+, 2021)

see Johan's focus

(physical) disk model
= parametric approaches

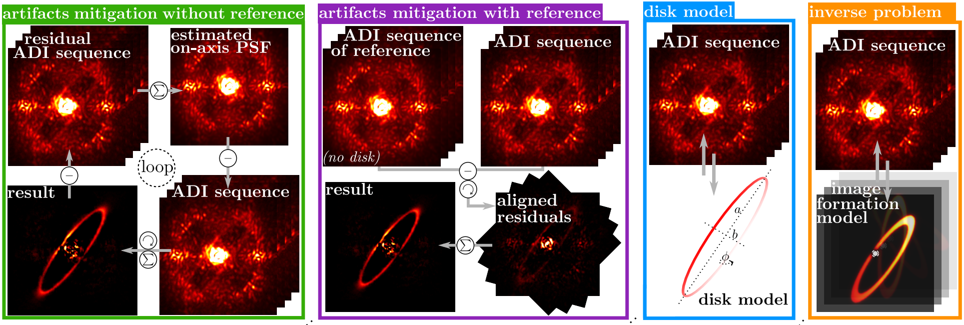
(Milli+, 2017)

(Esposito+, 2013)

DISKFM (Mazoyer+, 2020)

Different categories of algorithms for disk reconstruction

More advanced algorithms:



see Julien's & Sophia's focus *see Johan's focus* *focus of this presentation*

iterative PCA
(Pairet+, 2018)

data imputation strategy
(Ren+, 2020)

reference differential imaging (RDI)
= searching for similarities in images

RDI with a large library
(Gerard+, 2016) (Ren+, 2018)
(Xuan+, 2018) (Ruane+, 2019)

RDI with star hopping (Wahhaj+, 2021)

(physical) disk model
= parametric approaches

(Milli+, 2017)

(Esposito+, 2013)

DISKFM (Mazoyer+, 2020)

image formation model
= non-parametric approaches

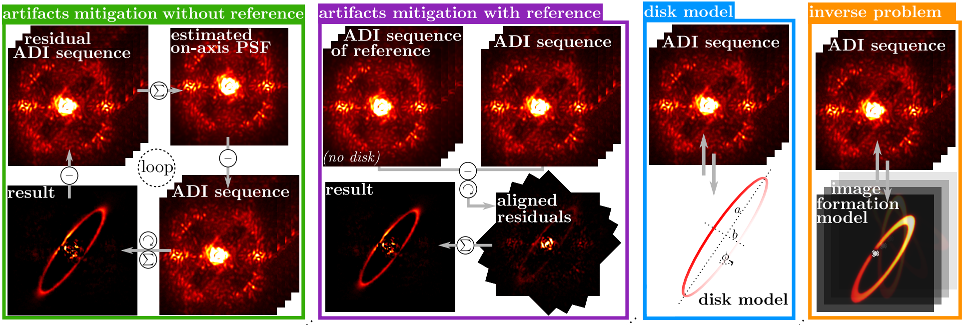
MAYO (Pairet+, 2021)

MUSTARD (Juillard+, 2022)

REXPACO (Flasseur+, 2021-22)

Different categories of algorithms for disk reconstruction

More advanced algorithms:



iterative PCA
(Pairet+, 2018)

data imputation strategy
(Ren+, 2020)

see Julien's & Sophia's focus

reference differential imaging (RDI)
= searching for similarities in images

RDI with a large library
(Gerard+, 2016) (Ren+, 2018)
(Xuan+, 2018) (Ruane+, 2019)

RDI with star hopping (Wahhaj+, 2021)

see Johan's focus

(physical) disk model
= parametric approaches

(Milli+, 2017)

(Esposito+, 2013)

DISKFM (Mazoyer+, 2020)

focus of this presentation

image formation model
= non-parametric approaches

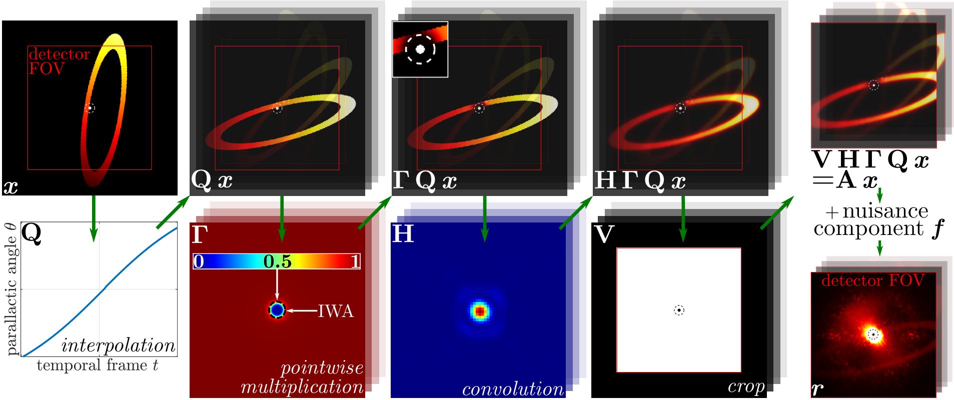
MAYO (Pairet+, 2021)

MUSTARD (Juillard+, 2022)

REXPACO (Flasseur+, 2021-22)

+ specific algorithms for polarization data *see Maud's focus*
e.g., inverse problem approach: RAPSHODIE (Denneulin+, 2021)

The common ingredient: the image formation model

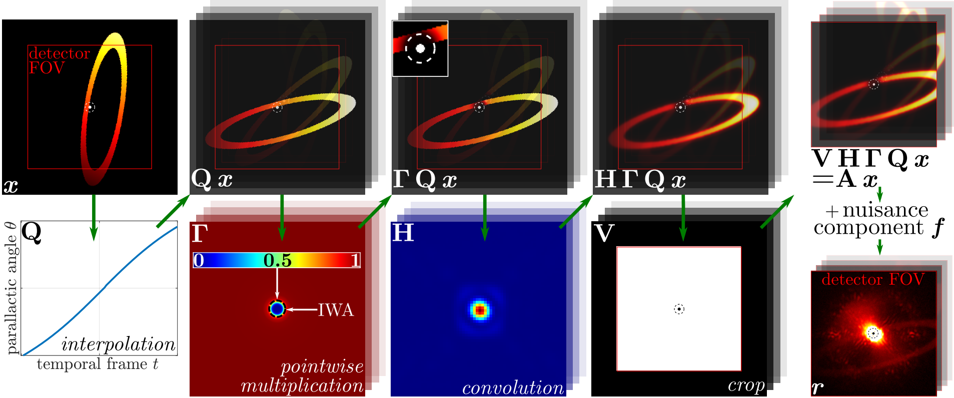


Operators / *implementation*:

- **Q**: **rotation** / *(sparse) interpolation matrix*
- **Γ**: **attenuation** / *diagonal matrix*
- **H**: **blur** / *bi-dimensional discrete convolution*
- **V**: **truncation** / *sparse matrix*

Subject to small variations depending on the algorithm.

The example of the REXPACO-based algorithms



Specificities of REXPACO-based algorithms:

⇒ accounting for the statistics Ω of the nuisance f ⇐

- REXPACO (Flasseur+, 2021): *for ADI observations*
- robust REXPACO (Flasseur+, 2022): *temporal robustness*
- REXPACO ASDI (Flasseur+, sub., ArXiv): *for ASDI observations*

Regularized reconstruction: *framework*

Model of the observed intensity

$$\boldsymbol{r} = \mathbf{A}\,\boldsymbol{x} + \boldsymbol{f}\,,$$

- $\boldsymbol{r}$ $(\mathbb{R}^{N\times T})$: total intensity in ADI stack of T frames with N pixels,
- $\boldsymbol{x}$ $\left((\mathbb{R}^{+})^{M}\right)$: unknown object flux,
- $\mathbf{A}$ $(\mathbb{R}^{M} \rightarrow \mathbb{R}^{N\times T})$: linear operator describing the image formation,
- $\boldsymbol{f}$ $(\mathbb{R}^{N\times T})$: noise; $\boldsymbol{f} \gg \mathbf{A}\,\boldsymbol{x}$, nonstationary, fluctuates over time.

Regularized reconstruction: *framework*

Model of the observed intensity

$$r = \mathbf{A}\, x + f\,,$$

- r $(\mathbb{R}^{N\times T})$: total intensity in ADI stack of T frames with N pixels,
- x $\left((\mathbb{R}^{+})^{M}\right)$: unknown object flux,
- $\mathbf{A}$ $(\mathbb{R}^{M} \to \mathbb{R}^{N\times T})$: linear operator describing the image formation,
- f $(\mathbb{R}^{N\times T})$: noise; $f \gg \mathbf{A}\, x$, nonstationary, fluctuates over time.

Regularized reconstruction of the object flux

Resolution of an inverse problem:

$$\hat{x} = \underset{x>0}{\arg\min}\{\mathscr{C}(r, x, \mathbf{A}, \mathbf{\Omega}, \mu) = \mathscr{D}(r, \mathbf{A}\, x, \mathbf{\Omega}) + \mathscr{R}(x, \mu)\}\,,$$

- $\mathscr{D}(r, \mathbf{A}\, x, \mathbf{\Omega})$: data-fidelity term, depends on $\mathbf{\Omega}$ statistics of f,
- $\mathscr{R}(x, \mu)$: regularization term, depends on hyperparameters μ.

Modeling of the nuisance component

Statistical model

Multi-variate Gaussian ($\boldsymbol{\Omega} = \{\boldsymbol{m}, \mathbf{C}\}$)

$\Rightarrow \boldsymbol{f} = \boldsymbol{m} + \boldsymbol{u}$ where $\boldsymbol{u} \sim \mathcal{N}(\mathbf{0}, \mathbf{C})$

Co-log-likelihood:

$$\mathscr{D}(\boldsymbol{r}, \mathbf{A}\,\boldsymbol{x}, \boldsymbol{\Omega}) = \frac{T}{2} \log \det \mathbf{C} + \frac{1}{2} \sum_{t=1}^{T} \| \boldsymbol{r}_t - \boldsymbol{m} - [\mathbf{A}\,\boldsymbol{x}]_t \|^2_{\mathbf{C}^{-1}} \,.$$

Modeling of the nuisance component

Statistical model

Multi-variate Gaussian ($\boldsymbol{\Omega} = \{\boldsymbol{m}, \mathbf{C}\}$)

$$\Rightarrow \boldsymbol{f} = \boldsymbol{m} + \boldsymbol{u} \quad \text{where} \quad \boldsymbol{u} \sim \mathcal{N}(\mathbf{0}, \mathbf{C})$$

Co-log-likelihood:

$$\mathscr{D}(\boldsymbol{r}, \mathbf{A}\,\boldsymbol{x}, \boldsymbol{\Omega}) = \frac{T}{2} \log \det \mathbf{C} + \frac{1}{2} \sum_{t=1}^{T} \|\boldsymbol{r}_t - \boldsymbol{m} - [\mathbf{A}\,\boldsymbol{x}]_t\|^2_{\mathbf{C}^{-1}} .$$

Statistical learning

Estimators from the **maximum likelihood**:

- $\widehat{\boldsymbol{m}} = \frac{1}{T} \sum_{t=1}^{T} (\boldsymbol{r}_t - [\mathbf{A}\,\boldsymbol{x}]_t)$,
- $\widehat{\mathbf{C}} = \frac{1}{T} \sum_{t=1}^{T} (\boldsymbol{r}_t - \boldsymbol{m} - [\mathbf{A}\,\boldsymbol{x}]_t)(\boldsymbol{r}_t - \boldsymbol{m} - [\mathbf{A}\,\boldsymbol{x}]_t)^\top$.

Modeling of the nuisance component

Statistical model

Multi-variate Gaussian ($\boldsymbol{\Omega} = \{\boldsymbol{m}, \mathbf{C}\}$)

$$\Rightarrow \boldsymbol{f} = \boldsymbol{m} + \boldsymbol{u} \quad \text{where} \quad \boldsymbol{u} \sim \mathcal{N}(\mathbf{0}, \mathbf{C})$$

Co-log-likelihood:

$$\mathscr{D}(\boldsymbol{r}, \mathbf{A}\,\boldsymbol{x}, \boldsymbol{\Omega}) = \frac{T}{2} \log\det \mathbf{C} + \frac{1}{2} \sum_{t=1}^{T} \| \boldsymbol{r}_t - \boldsymbol{m} - [\mathbf{A}\,\boldsymbol{x}]_t \|^2_{\mathbf{C}^{-1}} \,.$$

Statistical learning

Estimators from the **maximum likelihood**:

- $\widehat{\boldsymbol{m}} = \frac{1}{T} \sum_{t=1}^{T} (\boldsymbol{r}_t - [\mathbf{A}\,\boldsymbol{x}]_t)\,,$
- $\widehat{\mathbf{C}} = \frac{1}{T} \sum_{t=1}^{T} (\boldsymbol{r}_t - \boldsymbol{m} - [\mathbf{A}\,\boldsymbol{x}]_t)(\boldsymbol{r}_t - \boldsymbol{m} - [\mathbf{A}\,\boldsymbol{x}]_t)^\top \,.$

? Limited number T of samples to estimate $\widehat{\mathbf{C}}$

? The estimators $\widehat{\boldsymbol{m}}$ and $\widehat{\mathbf{C}}$ depend on the unknown object flux $\boldsymbol{x}$

Modeling of the nuisance component

Statistical model

Multi-variate Gaussian $(\boldsymbol{\Omega} = \{\boldsymbol{m}, \mathbf{C}\})$

$$\Rightarrow \boldsymbol{f} = \boldsymbol{m} + \boldsymbol{u} \text{ where } \boldsymbol{u} \sim \mathcal{N}(\mathbf{0}, \mathbf{C})$$

Co-log-likelihood:

$$\mathscr{D}(\boldsymbol{r}, \mathbf{A}\,\boldsymbol{x}, \boldsymbol{\Omega}) = \frac{T}{2} \log\det \mathbf{C} + \frac{1}{2}\sum_{t=1}^{T} \|\boldsymbol{r}_t - \boldsymbol{m} - [\mathbf{A}\,\boldsymbol{x}]_t\|^2_{\mathbf{C}^{-1}}\,.$$

Statistical learning

Estimators from the **maximum likelihood**:

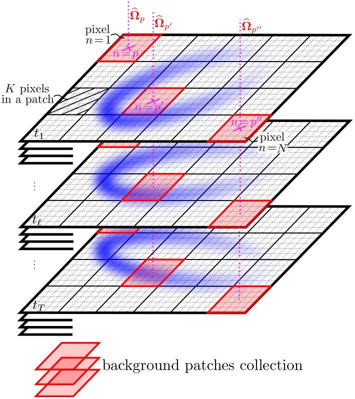
- $\widehat{\boldsymbol{m}} = \frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{r}_t - [\mathbf{A}\,\boldsymbol{x}]_t)\,,$
- $\widehat{\mathbf{C}} = \frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{r}_t - \boldsymbol{m} - [\mathbf{A}\,\boldsymbol{x}]_t)(\boldsymbol{r}_t - \boldsymbol{m} - [\mathbf{A}\,\boldsymbol{x}]_t)^\top\,.$

? Limited number T of samples to estimate $\widehat{\mathbf{C}}$

$\Rightarrow$ **Local modeling of PAtch COvariances**

Local learning of PAtch COvariances

REXPACO: Reconstruction of Extended features by learning of PAtch COvariances



REXPACO principle

Accounts for background fluctuations

$$\widehat{\Omega}_n = \{\widehat{m}_n, \widehat{\mathbf{C}}_n\}$$

- **Local modeling**: $K \simeq 80$ pix/patch

  $\Rightarrow$ **local adaptivity** $\Leftarrow$

- **Reconstruction**: all patches

Local learning of PAtch COvariances

REXPACO: Reconstruction of Extended features by learning of PAtch COvariances

REXPACO principle

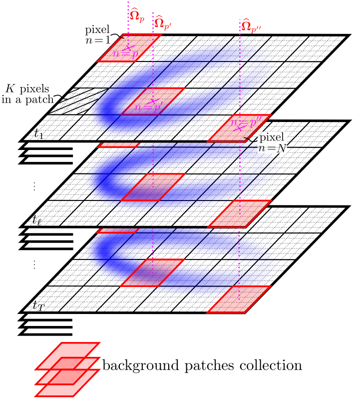
Accounts for background fluctuations

$$\widehat{\boldsymbol{\Omega}}_n = \{\widehat{\boldsymbol{m}}_n, \widehat{\mathbf{C}}_n\}$$

- **Local modeling**: $K \simeq 80$ pix/patch

  $\Rightarrow$ **local adaptivity** $\Leftarrow$

- **Reconstruction**: all patches



? In spite of local modeling, $K \approx T$

$\Rightarrow$ **A form of regularization on covariances should be enforced**

Local learning of PAtch COvariances – *shrinkage*

Issue and proposed approach

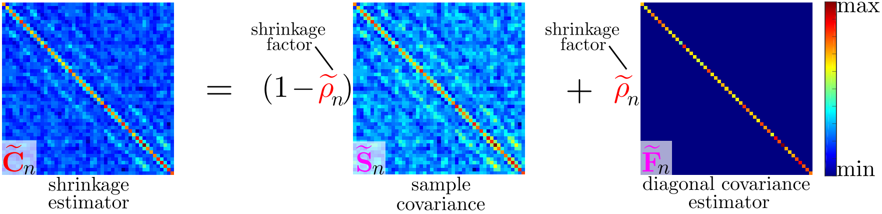
- Limited number of samples ($T \approx K$) to estimate $\mathbf{C}_n$ ($K \times K$)

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

*A form of **regularization** should be enforced.*

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

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



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

Local learning of PAtch COvariances – *shrinkage*

Issue and proposed approach

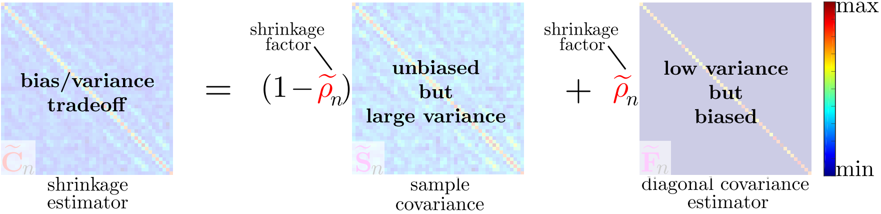
- Limited number of samples ($T \approx K$) to estimate $\mathbf{C}_n$ ($K \times K$)

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

  A form of ***regularization*** *should be enforced.*

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

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



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

Modeling of the nuisance component

Statistical model

Multi-variate Gaussian ($\boldsymbol{\Omega}_n = \{\boldsymbol{m}_n, \mathbf{C}_n\}$)

$$\Rightarrow \boldsymbol{f}_n = \boldsymbol{m}_n + \boldsymbol{u}_n \quad \text{where} \quad \boldsymbol{u}_n \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_n)$$

Co-log-likelihood:

$$\mathscr{D}(\boldsymbol{r}, \mathbf{A}\,\boldsymbol{x}, \boldsymbol{\Omega}) = \frac{T}{2} \sum_{n=1:K}^{N} \log\det \widetilde{\mathbf{C}}_n + \frac{1}{2} \sum_{n=1:K}^{N} \sum_{t=1}^{T} \| \boxed{\mathbf{P}_n} (\boldsymbol{r}_t - \widehat{\boldsymbol{m}} - [\mathbf{A}\,\boldsymbol{x}]_t) \|^2_{\widetilde{\mathbf{C}}_n^{-1}} \,.$$

$\boxed{\mathbf{P}_n}$: patch-extractor operator around pixel n

Statistical learning

- $\widehat{\boldsymbol{m}} = \frac{1}{T} \sum_{t=1}^{T} (\boldsymbol{r}_t - [\mathbf{A}\,\boldsymbol{x}]_t)$,
- $\widetilde{\mathbf{S}}_n = \frac{1}{T} \sum_{t=1}^{T} \left(\boxed{\mathbf{P}_n} (\boldsymbol{r}_t - \boldsymbol{m} - [\mathbf{A}\,\boldsymbol{x}]_t) \right) \left(\boxed{\mathbf{P}_n} (\boldsymbol{r}_t - \boldsymbol{m} - [\mathbf{A}\,\boldsymbol{x}]_t) \right)^{\top}$,
- $\widetilde{\mathbf{C}}_n = (1 - \widetilde{\rho}_n)\, \widetilde{\mathbf{S}}_n + \widetilde{\rho}_n\, \widetilde{\mathbf{F}}_n$.

Modeling of the nuisance component

Statistical model

Multi-variate Gaussian ($\boldsymbol{\Omega}_n = \{\boldsymbol{m}_n, \mathbf{C}_n\}$)

$$\Rightarrow \boldsymbol{f}_n = \boldsymbol{m}_n + \boldsymbol{u}_n \quad \text{where} \quad \boldsymbol{u}_n \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_n)$$

Co-log-likelihood:

$$\mathscr{D}(\boldsymbol{r}, \mathbf{A}\,\boldsymbol{x}, \boldsymbol{\Omega}) = \frac{T}{2} \sum_{n=1:K}^{N} \log\det \widetilde{\mathbf{C}}_n + \frac{1}{2} \sum_{n=1:K}^{N} \sum_{t=1}^{T} \| \boxed{\mathbf{P}_n} (\boldsymbol{r}_t - \widehat{\boldsymbol{m}} - [\mathbf{A}\,\boldsymbol{x}]_t) \|^2_{\widetilde{\mathbf{C}}_n^{-1}} \,.$$

$\boxed{\mathbf{P}_n}$: patch-extractor operator around pixel n

Statistical learning

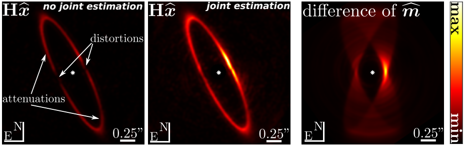
- $\widehat{\boldsymbol{m}} = \frac{1}{T} \sum_{t=1}^{T} (\boldsymbol{r}_t - [\mathbf{A}\,\boldsymbol{x}]_t)$,
- $\widehat{\mathbf{S}}_n = \frac{1}{T} \sum_{t=1}^{T} \left(\boxed{\mathbf{P}_n} (\boldsymbol{r}_t - \boldsymbol{m} - [\mathbf{A}\,\boldsymbol{x}]_t) \right) \left(\boxed{\mathbf{P}_n} (\boldsymbol{r}_t - \boldsymbol{m} - [\mathbf{A}\,\boldsymbol{x}]_t) \right)^{\top}$,
- $\widetilde{\mathbf{C}}_n = (1 - \widetilde{\rho}_n)\, \widehat{\mathbf{S}}_n + \widetilde{\rho}_n\, \widehat{\mathbf{F}}_n$.

? The estimators $\widehat{\boldsymbol{m}}$ and $\widetilde{\mathbf{C}}$ depend on the unknown object flux $\boldsymbol{x}$

Unbiased estimation of the nuisance statistics

Alternate/joint strategy

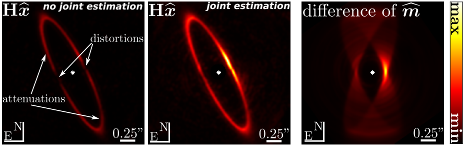
- Statistics biased by the object
  $\Rightarrow$ Alternate/joint estimation of $\widehat{\boldsymbol{\Omega}}$ and $\widehat{\boldsymbol{x}}$

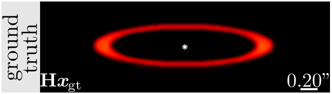


Unbiased estimation of the nuisance statistics

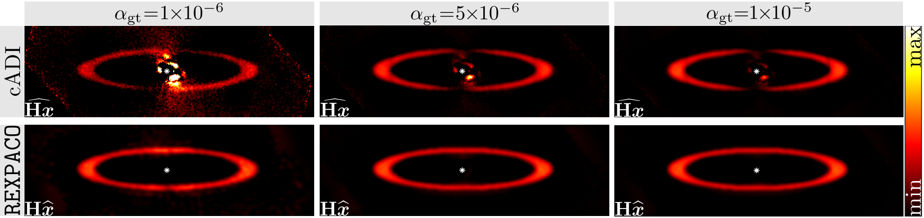
Alternate/joint strategy

- Statistics biased by the object
  $\Rightarrow$ Alternate/joint estimation of $\widehat{\Omega}$ and $\widehat{x}$





– a single reconstruction :

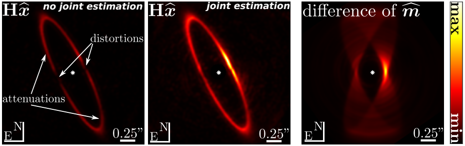


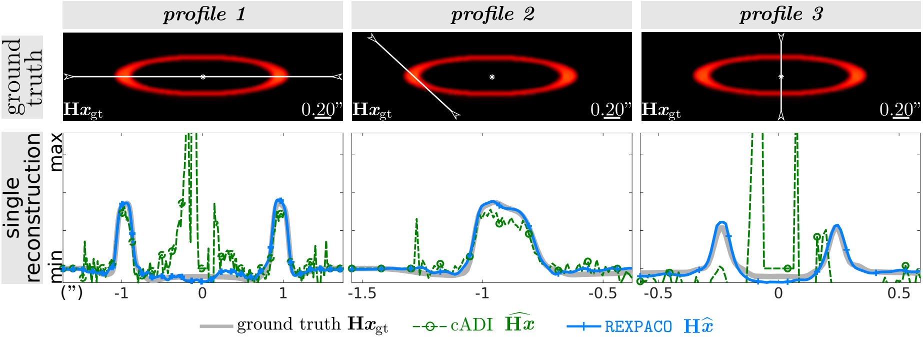
$\Rightarrow$ **The photometry is (mostly) preserved by the method.**

Unbiased estimation of the nuisance statistics

Alternate/joint strategy

- Statistics biased by the object
  $\Rightarrow$ Alternate/joint estimation of $\widehat{\Omega}$ and $\widehat{x}$



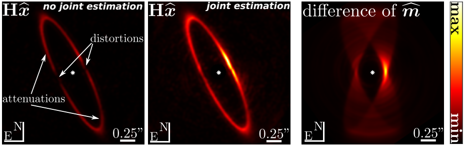


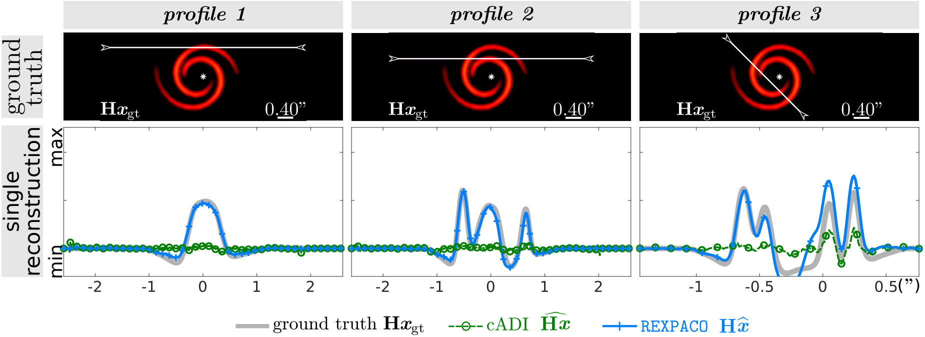
$\Rightarrow$ **The photometry is (mostly) preserved by the method.**

Unbiased estimation of the nuisance statistics

Alternate/joint strategy

- Statistics biased by the object
  $\Rightarrow$ Alternate/joint estimation of $\widehat{\Omega}$ and $\widehat{x}$





$\Rightarrow$ **The photometry is (mostly) preserved by the method.**

Unsupervised regularization & optimization

Unsupervised estimation of μ with SURE

$$\mathscr{R}(\boldsymbol{x}, \boxed{\boldsymbol{\mu}}) = \boxed{\mu_{\ell_1}} \sum_{n=1}^{N} |x_n| + \boxed{\mu_{\text{smooth}}} \sum_{n=1}^{N} \sqrt{||\boldsymbol{\Delta}_n \boldsymbol{x}||_2^2 + \epsilon^2}\,.$$

- SURE; unbiased estimator of MSE [Stein (1981)]

$\boxed{\Rightarrow \textbf{accounting for the local statistics } \boldsymbol{\Omega} \textbf{ of } \boldsymbol{f}}$:

$$\text{SURE}(\boldsymbol{\mu}) = \sum_{n \in \mathbb{P}} \sum_{t} ||\boldsymbol{r}_{n,t} - \widehat{\boldsymbol{m}}_n - [\mathbf{A}\, \mathbf{v}_{\boldsymbol{\mu}}(\boldsymbol{r})]_{n,t}||^2_{\widehat{\boldsymbol{\sigma}}^{-2}_{n,t} \widehat{\mathbf{C}}^{-1}_n} + 2\,\mathrm{tr}\left(\mathbf{A}\, \mathbf{J}_{\mathbf{v}_{\boldsymbol{\mu}}}(\boldsymbol{r})\right) - N\,,$$

...BUT no closed-form expression of $\mathbf{J}_{\mathbf{v}_{\boldsymbol{\mu}}}(\boldsymbol{r})$, the Jacobian of $\mathbf{v}_{\boldsymbol{\mu}}$ w.r.t $\boldsymbol{r}$.

- Evaluation of $\mathrm{tr}\left(\mathbf{A}\, \mathbf{J}_{\mathbf{v}_{\boldsymbol{\mu}}}(\boldsymbol{r})\right)$ with a *black-box approach* [Ramani (2012)]:

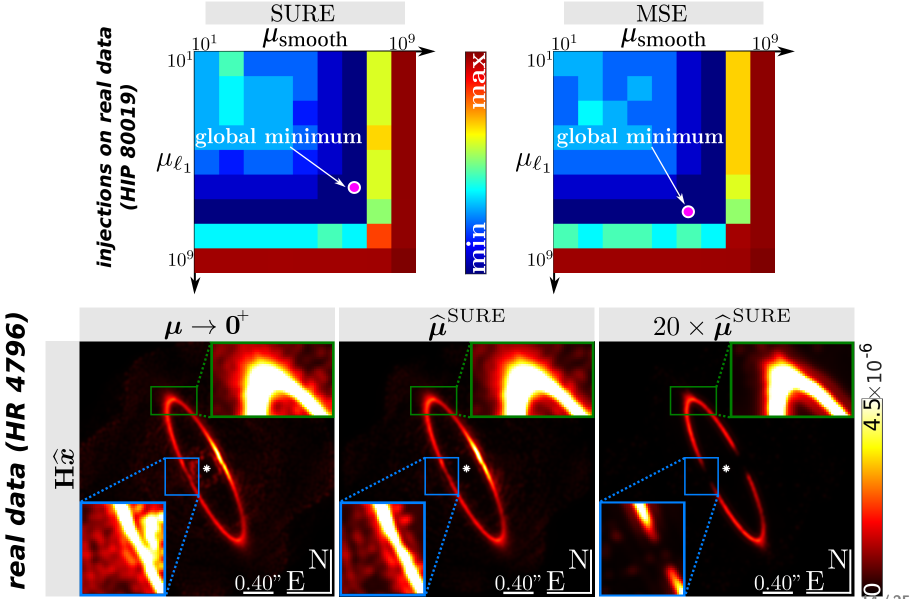
$$\mathrm{tr}\left(\mathbf{A}\, \mathbf{J}_{\mathbf{v}_{\boldsymbol{\mu}}}(\boldsymbol{r})\right) \approx \xi^{-1} \boldsymbol{b}^\top \mathbf{A}\, [\mathbf{v}_{\boldsymbol{\mu}}(\boldsymbol{r} + \xi \boldsymbol{b}) - \mathbf{v}_{\boldsymbol{\mu}}(\boldsymbol{r})]\,,$$

Optimization

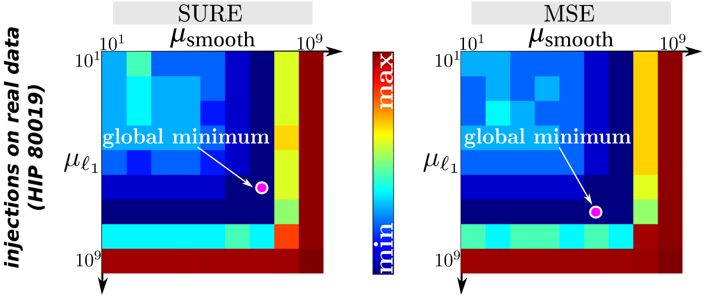
- bound constraints: $\boldsymbol{x} > \mathbf{0}$
- differentiable objective function

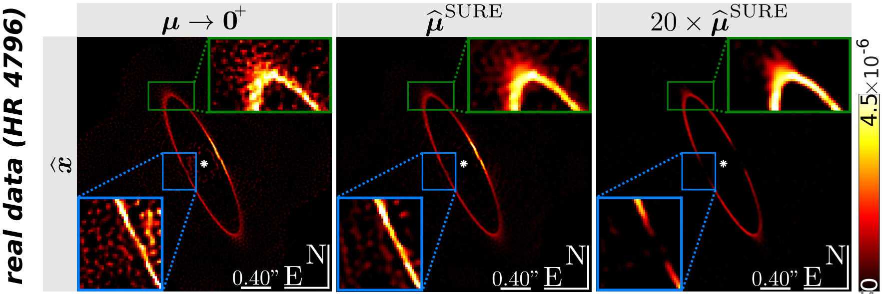
$\Rightarrow$ solved with VMLMB [Thiébaut (2002)]

Unsupervised regularization & optimization

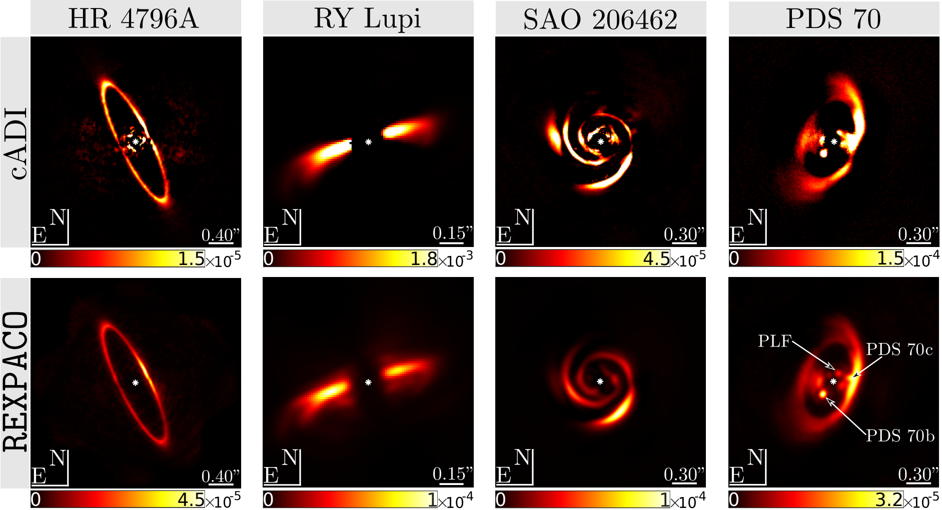


Unsupervised regularization & optimization



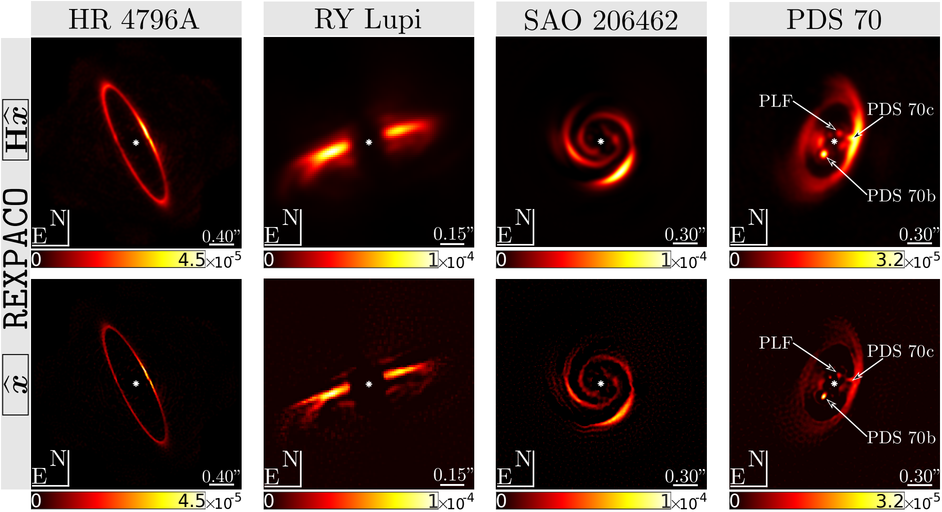


Comparison with cADI/PCA on VLT/SPHERE IRDIS data



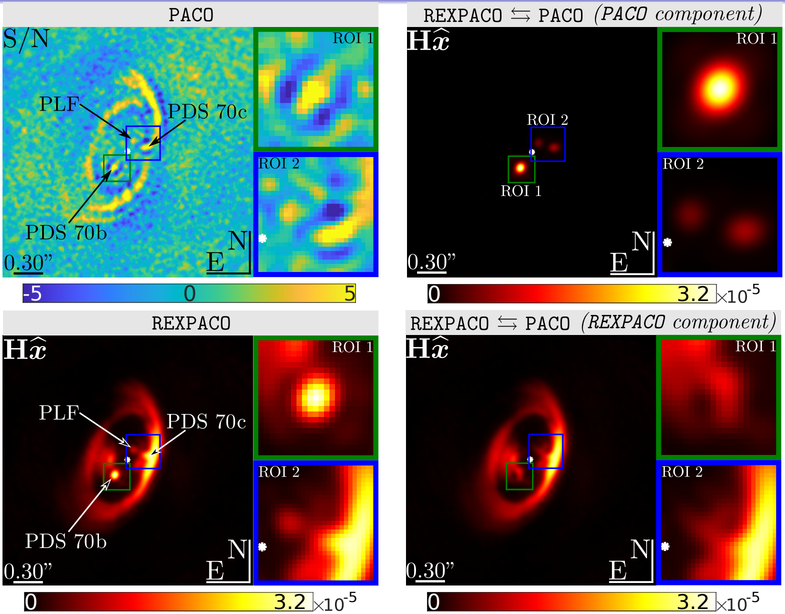
statistical model ⇒ **residual stellar leakages are reduced**
image formation model ⇒ **non-physical artefacts are reduced**

Comparison with cADI/PCA on VLT/SPHERE IRDIS data

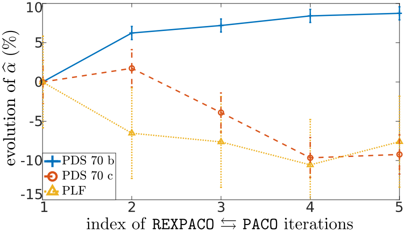


statistical model **⇒ residual stellar leakages are reduced**
image formation model **⇒ non-physical artefacts are reduced**
image formation model **⇒ angular resolution is improved**

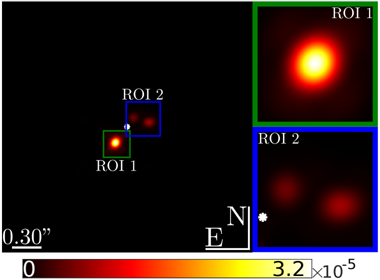
Unmixing point-like and extended features

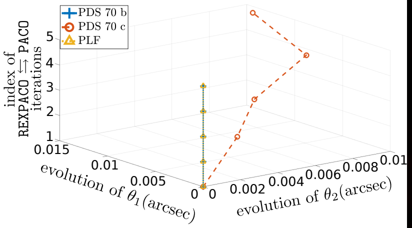


Unmixing point-like and extended features

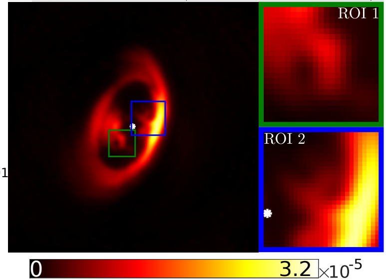


REXPACO ⇆ PACO *(PACO component)*



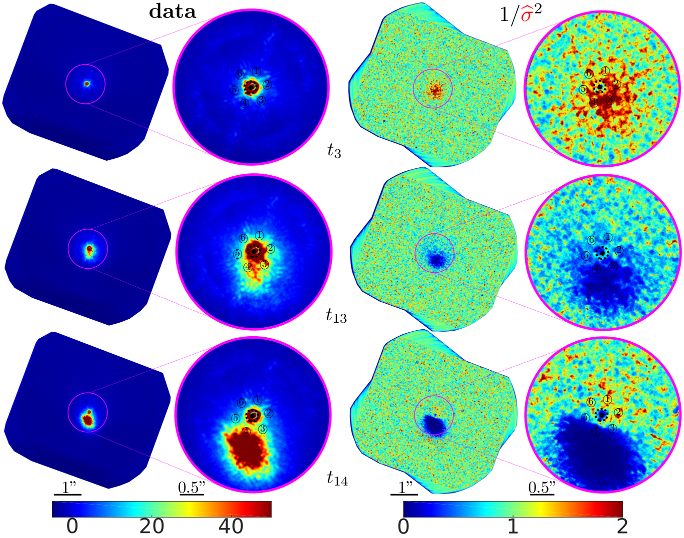


REXPACO ⇆ PACO *(REXPACO component)*



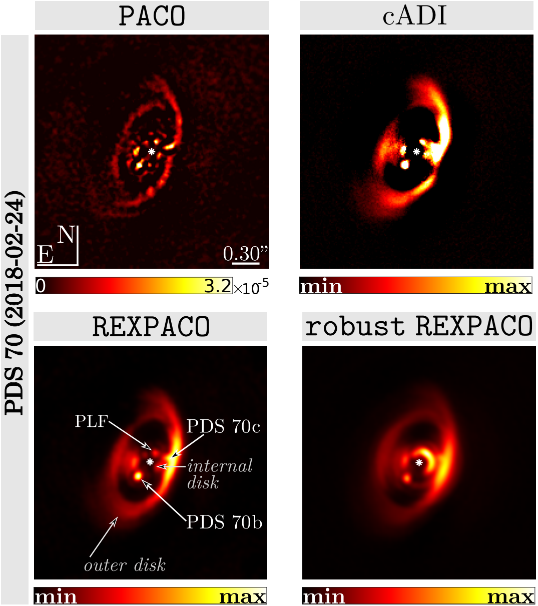
Improving the robustness by temporal weighting

local + data-driven identification and neutralization of outliers



⇒ **impact of large fluctuations is decreased, robustness is improved**

Improving the robustness by temporal weighting



robustness benefits:

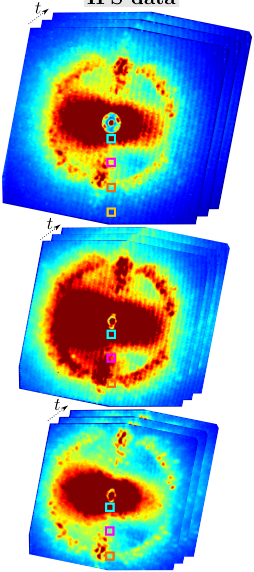
statistical model ⇒ **better rejection of nuisance comp.**

statistical model ⇒ **better reconstruct. of fine structures at short separations**

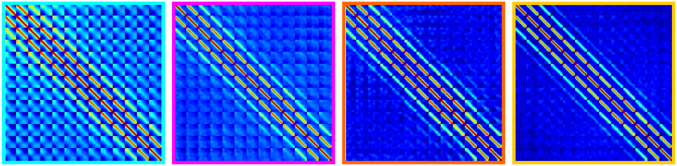
see Maud's focus for more results

Joint multi-spectral processing: *general principle*

IFS data

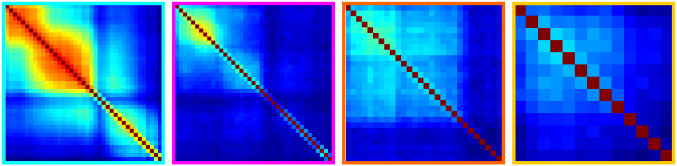


estimated **spatial** correlations



distance to the star

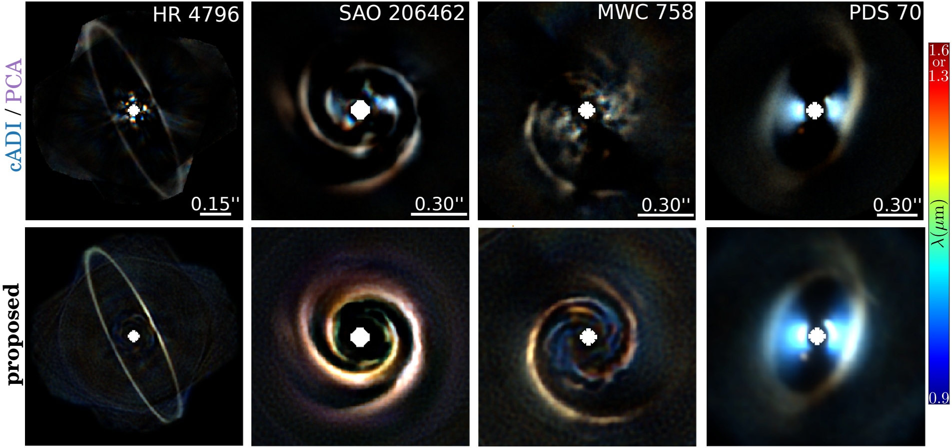
estimated **spectral** correlations



distance to the star

accounting for nonstationary spatial & spectral correlations

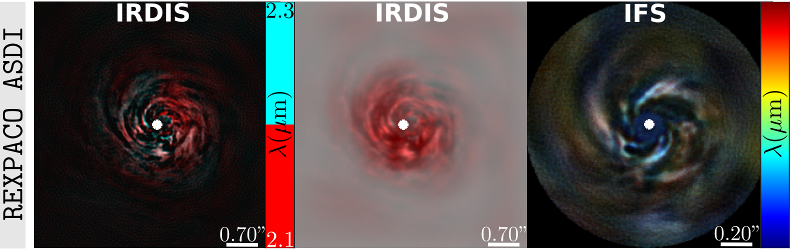
Comparison with cADI/PCA on VLT/SPHERE IFS data



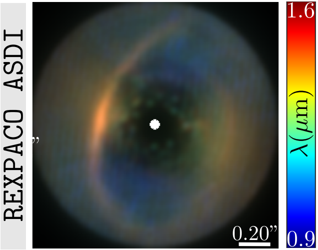
statistical model **⇒ residual stellar leakages are reduced**
image formation model **⇒ non-physical artefacts are reduced**
image formation model **⇒ angular resolution is improved**
spectral diversity **⇒ the key for disks with a circular symmetry**

VLT/SPHERE IFS reconstructions - *other targets*

AB Aurigae



HD 163296



A focus on MAYONNAISE, MUSTARD algorithms

MAYONNAISE (Pairet 2021+)

inverse problem approach, with specific regularization terms, no statistical modeling of the nuisance component

Model of the observed intensity

$$\boldsymbol{r} = \mathbf{A}\,(\boldsymbol{x}_d + \boldsymbol{x}_p) + \boldsymbol{f}\,,$$

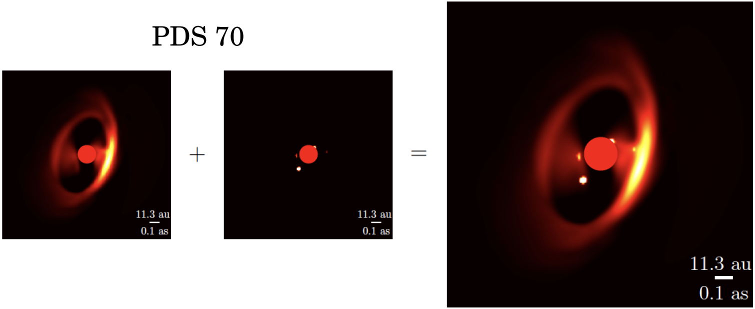
- $\boldsymbol{r}\;(\mathbb{R}^{N\times T})$: total intensity in ADI stack of T frames with N pixels,
- $\boldsymbol{x} = \boldsymbol{x}_d + \boldsymbol{x}_p\;\left((\mathbb{R}^{+})^{M}\right)$: unknown object flux (disk + planets),
- $\mathbf{A}\;(\mathbb{R}^{M} \to \mathbb{R}^{N\times T})$: image formation model (rotation + blur),
- $\boldsymbol{f}\;(\mathbb{R}^{N\times T})$: noise; $\boldsymbol{f} \gg \mathbf{A}\,\boldsymbol{x}$, nonstationary, fluctuates over time.

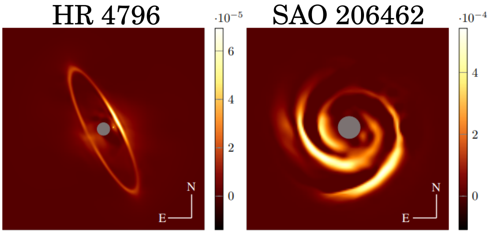
Regularized reconstruction

$$\{\widehat{\boldsymbol{x}}_d, \widehat{\boldsymbol{x}}_p, \widehat{\boldsymbol{f}}\} = \underset{\boldsymbol{x}_d,\,\boldsymbol{x}_p,\,\boldsymbol{f}}{\arg\min}\{\mathscr{L}\,(\boldsymbol{r} - \mathbf{A}\,(\boldsymbol{x}_d + \boldsymbol{x}_p) - \boldsymbol{f}) + \mathscr{R}(\boldsymbol{x}_d, \boldsymbol{x}_p)\}\,,$$

$\mathscr{L}$:= Huber loss function ; $\mathscr{R}$:= regularization term ($\boldsymbol{f}$ is low rank, $\boldsymbol{x}_p$ is sparse in space domain, $\boldsymbol{x}_d$ is sparse in transformed domain).

A focus on MAYONNAISE, MUSTARD algorithms



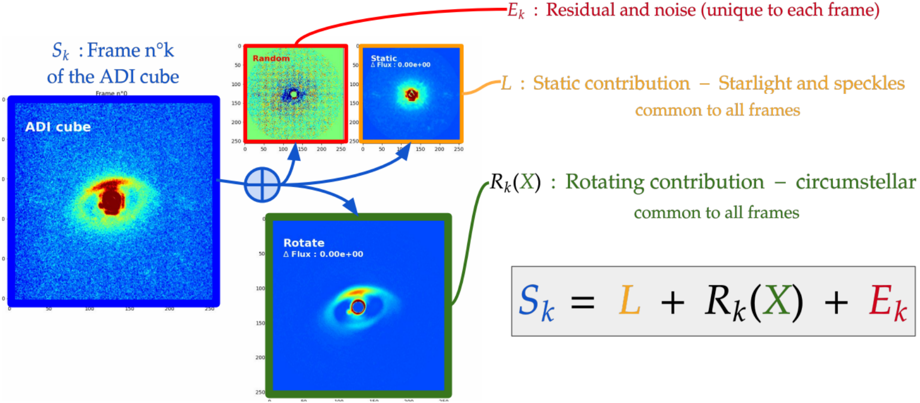


Courtesy: extracted from (Pairet 2021+)

A focus on MAYONNAISE, MUSTARD algorithms

MUSTARD (Juillard+, 2022, in prep.)

MAYO with decomposition of nuisance component in two terms



$$S_k = L + R_k(X) + E_k$$

Coutesy: S. Juillard, extracted from a presentation available at: https://orbi.uliege.be/bitstream/2268/291212/1/PDS70-%20resume.pdf

Conclusions

Different classes of post-processing algorithms for disk imaging:

- *subtraction (cADI, PCA, TLOCI),*
- *artifacts mitigation (iterative PCA, data imputation strategy)*
- *reference differential imaging,*
- *parametric approaches with a disk model,*
- *non-parametric approaches with an image formation model.*

Advanced algorithms allows:

- *detection at better contrasts,*
- *better preservation of the disk morphology and photometry*
  - *reduce classical artifacts (e.g., self-subtraction),*
  - *reduce stellar leakages,*
- unmixing of point-like and extended sources.

Specificities of `REXPACO`-based algorithms:

- *encompass a statistical modeling of the nuisance component,*
- *spectral diversity is the key for circulo-symmetric disks.*

Classical algorithms
Marois+ 2006, "Angular differential imaging: a powerful high-contrast imaging technique", APJ, 641(1), 556 (cADI)
Marois+ 2014, "GPI PSF subtraction with TLOCI: the next evolution in exoplanet/disk high-contrast imaging", SPIE Adaptive Optics Systems, 9148 (TLOCI)
Soummer+ 2012, "Detection and characterization of exoplanets and disks using projections on Karhunen–Loève eigenimages", APJ Letters, 755(2), L28 (KLIP/PCA)

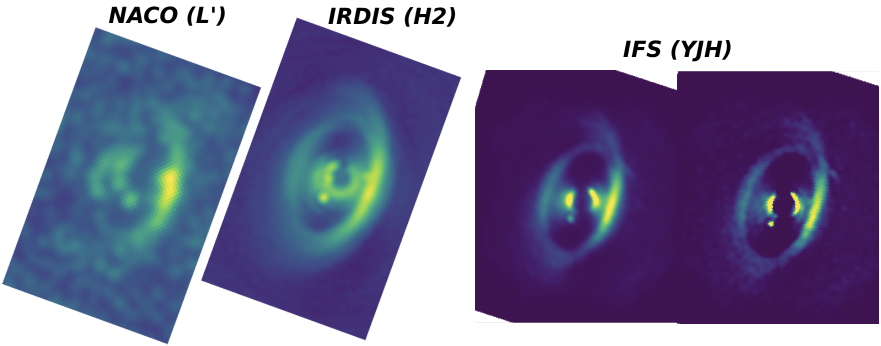
Artifacts mitigation without reference
Pairet+ 2018, "Reference-less algorithm for circumstellar disks imaging", ArXiv (iterative PCA)
Ren+ 2020, "Using data imputation for signal separation in high-contrast imaging", APJ, 892(2), 74 (data imputation)

Artifacts mitigation with reference
Gerard+ 2016, "Planet detection down to a few λ/D: an RSDI/TLOCI approach to PSF subtraction", SPIE Adaptive Optics (RSDI/TLOCI)
Ren+ 2018, "Non-negative matrix factorization: robust extraction of extended structures", APJ, 852(2), 104 (NMF)
Xuan+ 2018, "Characterizing the performance of the NIRC2 vortex coronagraph at WM Keck Observatory", APJ, 156(4), 156 (RDI ADI on KECK/NIRC2 data)
Ruane+ 2019, "Reference star differential imaging of close-in companions and circumstellar disks with the NIRC2 vortex coronagraph at the WM Keck Observatory", APJ, 157(3), 118 (RDI ADI on KECK/NIRC2 data)
Wahhaj+ 2021, "A search for a fifth planet around HR 8799 using the star-hopping RDI technique at VLT/SPHERE", A&A, 648, A26 (star-hopping RDI on VLT/SPHERE data)

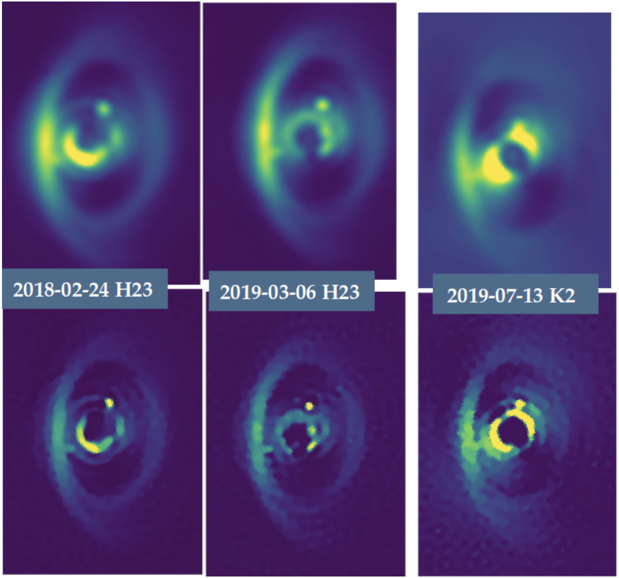
Disk models
Milli+ 2017, "Near-infrared scattered light properties of the HR 4796 A dust ring - A measured scattering phase function from 13.6° to 166.6°", A&A, 599, A108 (disk model fitting on HR 4796 data)
Esposito+ 2013, "Modeling self-subtraction in angular differential imaging: Application to the HD 32297 debris disk", APJ, 780(1), 25
Mazoyer+ 2020, "A forward modeling tool for disk analysis with coronagraphic instruments", SPIE Ground-based and Airborne Instrumentation for Astronomy, 11447 (DiskFM: forward-backward modeling for disk)

Inverse problems
Pairet+ 2021, "MAYONNAISE: a morphological components analysis pipeline for circumstellar discs and exoplanets imaging in the near-infrared", MNRAS, 503(3) (MAYONNAISE)
Julliard+ 2022, "A spiral arm in the protoplanety disk PDS70?" (presentation) (MUSTARD)
Flasseur+ 2021, "REXPACO: An algorithm for high contrast reconstruction of the circumstellar environment by angular differential imaging", A&A, 651, A62 (`REXPACO`)
Flasseur+ 2022, "Multispectral image reconstruction of faint circumstellar environments from high contrast angular spectral differential imaging (ASDI) data", SPIE Adaptive Optics Systems, 12185 (`robust REXPACO`)
Flasseur+ (sub), "Joint unmixing and deconvolution for angular and spectral differential imaging", ArXiv (`REXPACO ASDI`)

Multi-instruments



Multi-epochs



Reconstruction framework – *data fidelity*

Data fidelity term

Gaussian Scale Mixture ($\boldsymbol{\Omega}_{n,t} = \{\boldsymbol{m}_n, \boldsymbol{\sigma}_{n,t}, \mathbf{C}_n\}$)

$$\Rightarrow \boldsymbol{f}_{n,t} = \boldsymbol{m}_n + \boldsymbol{\sigma}_{n,t}\,\boldsymbol{u}_n \quad \text{where} \quad \boldsymbol{u}_n \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_n)$$

Co-log-likelihood:

$$\mathscr{D}(\boldsymbol{r}, \mathbf{A}\,\boldsymbol{x}, \boldsymbol{\Omega}) = \frac{1}{2}\sum_{n\in\mathbb{P}}\sum_{t} \log\det \widehat{\boldsymbol{\sigma}}^2_{n,t}\,\widehat{\mathbf{C}}_n + \frac{1}{2}\sum_{n\in\mathbb{P}}\sum_{t} \|\widehat{\boldsymbol{v}}_{n,t}\|^2_{\widehat{\boldsymbol{\sigma}}^{-2}_{n,t}\widehat{\mathbf{C}}^{-1}_n},$$

$\widehat{\boldsymbol{v}}_{n,t} = \boldsymbol{r}_{n,t} - \widehat{\boldsymbol{m}}_n - [\mathbf{A}\,\boldsymbol{x}]_{n,t}$: residual intensity patch around pixel n.

Statistical background modeling

- Scaling factor: $\widehat{\boldsymbol{\sigma}}^2_{n,t} = (1/K)\,\widehat{\boldsymbol{v}}_{n,t}\,\widehat{\mathbf{C}}^{-1}_n\,\widehat{\boldsymbol{v}}^{\top}_{n,t}$
- Sample mean: $\widehat{\boldsymbol{m}}_n = \frac{1}{T}\sum_{t=1}^{T} \widehat{\boldsymbol{\sigma}}^{-2}_{n,t}\left(\boldsymbol{r}_{n,t} - [\mathbf{A}\,\boldsymbol{x}]_{n,t}\right),$
- Sample covariance: $\widehat{\mathbf{S}}_n = \frac{1}{T}\sum_{t=1}^{T} \widehat{\boldsymbol{\sigma}}^{2}_{n,t}\,\widehat{\boldsymbol{v}}_{n,t}\,\widehat{\boldsymbol{v}}^{\top}_{n,t},$
- *Shrunk* covariance: $\widehat{\mathbf{C}}_n = (1-\widehat{\rho}_n)\,\widehat{\mathbf{S}}_n + \widehat{\rho}_n\,\widehat{\mathbf{F}}_n = \widehat{\mathbf{W}}_n \odot \widehat{\mathbf{S}}_n\,.$

? The statistics $\boldsymbol{\Omega}$ depends on the sought object $\boldsymbol{x}$

$\Rightarrow$ **alternate or hierarchical estimation of $\boldsymbol{\Omega}$ and $\boldsymbol{x}$ is mandatory**

Reconstruction framework – *data fidelity*

Data fidelity term

$$\text{Gaussian Scale Mixture } (\boldsymbol{\Omega}_{n,t} = \{\boldsymbol{m}_n, \boldsymbol{\sigma}_{n,t}, \mathbf{C}_n\})$$

$$\Rightarrow \boldsymbol{f}_{n,t} = \boldsymbol{m}_n + \boldsymbol{\sigma}_{n,t}\,\boldsymbol{u}_n \quad \text{where} \quad \boldsymbol{u}_n \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_n)$$

Co-log-likelihood:

$$\mathscr{D}_{\text{joint}}(\boldsymbol{r}, \mathbf{A}\,\boldsymbol{x}, \boldsymbol{\Omega}) = \frac{1}{2}\sum_{n\in\mathbb{P}}\sum_{t} \log\det \widehat{\boldsymbol{\sigma}}^2_{n,t}(\boldsymbol{x})\,\widehat{\mathbf{C}}_n(\boldsymbol{x}) + \frac{1}{2}\sum_{n\in\mathbb{P}} \operatorname{tr}\left[\widehat{\mathbf{C}}_n^{-1}(\boldsymbol{x})\left(\widehat{\mathbf{W}}_n \odot \sum_t \widehat{\boldsymbol{\sigma}}^{-2}_{n,t}(\boldsymbol{x})\,\widehat{\boldsymbol{v}}_{n,t}(\boldsymbol{x})\,\widehat{\boldsymbol{v}}_{n,t}(\boldsymbol{x})^\top\right)\right],$$

$\widehat{\boldsymbol{v}}_{n,t}(\boldsymbol{x}) = \boldsymbol{r}_{n,t} - \widehat{\boldsymbol{m}}_n(\boldsymbol{x}) - [\mathbf{A}\,\boldsymbol{x}]_{n,t}$: residual intensity patch around pixel n.

Statistical background modeling

- Scaling factor: $\widehat{\boldsymbol{\sigma}}^2_{n,t}(\boldsymbol{x}) = (1/K)\,\widehat{\boldsymbol{v}}_{n,t}\left(\widehat{\mathbf{W}}_n \odot \widehat{\mathbf{C}}_n^{-1}\right)\widehat{\boldsymbol{v}}_{n,t}^\top$
- Sample mean: $\widehat{\boldsymbol{m}}_n(\boldsymbol{x}) = \frac{1}{T}\sum_{t=1}^{T} \widehat{\boldsymbol{\sigma}}^{-2}_{n,t}\left(\boldsymbol{r}_{n,t} - [\mathbf{A}\,\boldsymbol{x}]_{n,t}\right)$,
- Sample covariance: $\widehat{\mathbf{S}}_n(\boldsymbol{x}) = \frac{1}{T}\sum_{t=1}^{T} \widehat{\boldsymbol{\sigma}}^2_{n,t}\,\widehat{\boldsymbol{v}}_{n,t}\,\widehat{\boldsymbol{v}}_{n,t}^\top$,
- *Shrunk* covariance: $\widehat{\mathbf{C}}_n(\boldsymbol{x}) = (1-\widehat{\rho}_n)\,\widehat{\mathbf{S}}_n + \widehat{\rho}_n\,\widehat{\mathbf{F}}_n = \widehat{\mathbf{W}}_n \odot \widehat{\mathbf{S}}_n$.