A posteriori fringe sensing

Ferréol Soulez, Éric Thiébaut, Michel Tallon & Isabelle Tallon-Bosc

Équipe AiRi
Centre de Recherche Astronomique de Lyon
Université Claude Bernard Lyon I
École Normale Supérieure de Lyon

POLCA project funded by the French National Research Agency (ANR-10-BLAN-0511)

September 26, 2013

POLCA project

An opportunity to re-think signal processing from pixels to image in polychromatic interferometry

POLCA project

An opportunity to re-think signal processing from pixels to image in polychromatic interferometry

Image reconstruction point of view:

- simple if:
  - model as linear as possible,
  - noise statistics under control (*e.g.* no division of random variable. . .),
- computationally intensive,
  - no longer a constraint,

POLCA project

An opportunity to re-think signal processing from pixels to image in polychromatic interferometry

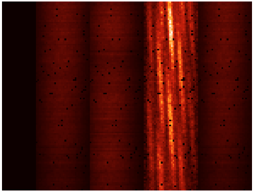
Image reconstruction point of view:

- simple if:
  - model as linear as possible,
  - noise statistics under control (*e.g.* no division of random variable. . .),
- computationally intensive,
  - no longer a constraint,

Image reconstruction directly using raw measurements? (no reduction)

Instantaneous brightness distribution in the image plane of an interferometric recombiner (*e.g.* AMBER):

$$i(\xi,t) = \sum_{j=1}^{J} a_j(\xi,\lambda)\, f_j(\lambda,t) + 2 \sum_{1\le j_1<j_2\le J} \sqrt{a_{j_1}(\xi,\lambda)\, f_{j_1}(\lambda,t)\, a_{j_2}(\xi,\lambda)\, f_{j_2}(\lambda,t)} \times \mathrm{Re}[v^{\mathrm{inst}}_{j_1,j_2}(\lambda)\, v^{\mathrm{obj}}_{j_1,j_2}(\lambda)\, \mathrm{e}^{\mathrm{i}\,[\omega_{j_1,j_2}(\lambda)\,\xi + \psi_{j_1,j_2}(\lambda,t)]}]$$



É. Tatulli *et al.* (2007)

Notations:

- ξ = position on detector;
- t = time;
- λ = wavelength;
- j = telescope number, (j_1, j_2) = pair of recombined telescopes (baseline), J = number of telescopes;
- $a_j(\xi, \lambda)$ = PSF of j-th output;
- $f_j(\lambda, t)$ = flux from j-th output;
- $v^{\mathrm{inst}}_{j_1,j_2}(\lambda)$ = instrumental visibility;
- $v^{\mathrm{obj}}_{j_1,j_2}(\lambda)$ = object complex visibility;
- $\omega_{j_1,j_2}(\lambda)$ = pulsation (for a given baseline);
- $\psi_{j_1,j_2}(\lambda, t)$ = phase term due to turbulence and fringe tracking errors;

Phase error term

Let $\delta_j(\lambda, t)$ be the optical path between the object and the j-th interferometric output. The phase of the fringes between j_1-th and j_2-th telescopes is that of the object complex visibility plus a perturbation phase term given by:

$$\begin{aligned}\psi_{j_1,j_2}(\lambda, t) &= \frac{2\,\pi}{\lambda}\,[\delta_{j_2}(\lambda, t) - \delta_{j_1}(\lambda, t)] \\ &\approx [\alpha_{j_2}(t) - \alpha_{j_1}(t)] + [\beta_{j_2}(t) - \beta_{j_1}(t)]/\lambda + \ldots\end{aligned}$$

Consequences:

- the perturbation $\psi_{j_1,j_2}(\lambda, t)$ prevents to integrate complex visibilities;
- only non-linear estimators (like powerspectrum and phase closures) which cancel $\psi_{j_1,j_2}(\lambda, t)$ can be integrated during an exposure;
  $\Longrightarrow$ loss of information;
  $\Longrightarrow$ lower SNR;
  $\Longrightarrow$ difficult to build an image given the non-linear measurements;

Can we estimate (perhaps up to some degeneracies) the unknown $\alpha_j(t)$ and $\beta_j(t)$ only from the actual data?

Integration in k-th pixel, ℓ-th spectral channel and m-th time frame yields the expectation of the image of the fringes:

$$
\begin{aligned}
y_{k,\ell,m} &\overset{\text{def}}{=} \int_{\xi_k-\Delta\xi/2}^{\xi_k+\Delta\xi/2} \int_{t_m-\Delta t/2}^{t_m+\Delta t/2} i(\xi,t)\,\mathrm{d}t\,\mathrm{d}\xi \\
&= \sum_{j=1}^{J} A_{j,k,\ell}\, f_{j,\ell,m} \\
&\quad + \sum_{b=1}^{J\,(J-1)/2} \sqrt{f_{j_1,\ell,m}\, f_{j_2,\ell,m}}\, \mathcal{M}(\partial_t \psi_{b,\ell,m}\, \Delta t) \\
&\quad \times \mathrm{Re}[B_{b,k,\ell}\, v^{\text{obj}}_{b,\ell}\, \mathrm{e}^{\mathrm{i}\,\psi_{b,\ell,m}}]
\end{aligned}
$$

with:

$$
B_{b,k,\ell} = 2\sqrt{A_{j_1(b),k,\ell} A_{j_2(b),k,\ell}}\; v^{\text{inst}}_{b,\ell}\, \mathrm{e}^{\mathrm{i}\,\omega_{b,\ell}\,\xi_k}
$$

$$
\psi_{b,\ell,m} = [\alpha_{j_2(b),m} - \alpha_{j_1(b),m}] + [\beta_{j_2(b),m} - \beta_{j_1(b),m}]/\lambda_\ell
$$

Notations:

- ξ_k = position of k-th pixel;
- t_m = time in m-th frame;
- λ_ℓ = wavelength in ℓ-th channel;
- b = baseline index;
- $(j_1(b), j_2(b))$ = pair of recombined telescopes for b-th baseline;
- $A_{j,k,\ell}$ = PSF of j-th output;
- $f_{j,\ell,m}$ = flux from j-th output;
- $v^{\text{inst}}_{b,\ell}$ = instrumental visibility;
- $v^{\text{obj}}_{b,\ell}$ = object complex visibility;
- $\psi_{b,\ell,m}$ = random phase error term;
- $\alpha_{j,m}$, $\beta_{j,m}$ = optical path parameters;
- $\mathcal{M}(\ldots)$ accounts for loss of visibility due to fringe motion
- $\omega_{b,\ell}$ = pulsation;

In matrix form, the fringe pattern and the photometric measurements are:

$$\boldsymbol{y}^{\text{fringe}} = \mathbf{A} \cdot \boldsymbol{f} + \mathbf{B} \cdot \boldsymbol{v}^{\text{raw}} \,,$$
$$\boldsymbol{y}^{\text{phot}} = \mathbf{C} \cdot \boldsymbol{f} \,,$$

with $\boldsymbol{f}$ the fluxes and $\boldsymbol{v}^{\text{raw}}$ the *raw complex visibilities*:

$$v^{\text{raw}}_{b,\ell,m} \stackrel{\text{def}}{=} \sqrt{f_{j_1,\ell,m}\, f_{j_2,\ell,m}}\, \mathcal{M}(\partial_t \psi_{b,\ell,m}\, \Delta t)\, v^{\text{obj}}_{b,\ell}\, \mathrm{e}^{\mathrm{i}\,\psi_{b,\ell,m}} \,.$$

Putting all available data (fringe pattern and photometric channels) together:

$$\boldsymbol{y} = \mathbf{H} \cdot \boldsymbol{x}(\boldsymbol{\theta})$$

with:

$$\boldsymbol{y} = \begin{pmatrix} \boldsymbol{y}^{\text{phot}} \\ \boldsymbol{y}^{\text{fringe}} \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} \mathbf{C} & \mathbf{0} \\ \mathbf{A} & \mathbf{B} \end{pmatrix}, \quad \boldsymbol{x}(\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{f} \\ \boldsymbol{v}^{\text{raw}} \end{pmatrix},$$

and the sought parameters:

$$\boldsymbol{\theta} = \{\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{f}, \boldsymbol{v}^{\text{obj}}\} \,.$$

Nota bene: $\mathbf{H}$ is a generalization of the *visibility to pixel matrix* (V2PM).

Available data can be put in the form:

$$\tilde{\boldsymbol{y}} = \underbrace{\mathbf{H} \cdot \boldsymbol{x}(\boldsymbol{\theta})}_{\text{model}} + \underbrace{\tilde{\boldsymbol{n}}}_{\text{noise}}$$

with:

$$\boldsymbol{\theta} = \{\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{f}, \boldsymbol{v}^{\text{obj}}\} \quad \text{and} \quad \boldsymbol{x}(\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{f} \\ \boldsymbol{v}^{\text{raw}} \end{pmatrix},$$

where $\boldsymbol{f}$ are the fluxes, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are the optical path parameters, $\boldsymbol{v}^{\text{obj}}$ are the object complex visibilities and $\boldsymbol{v}^{\text{raw}}$ are the **raw complex visibilities**:

$$v^{\text{raw}}_{b,\ell,m} \stackrel{\text{def}}{=} \sqrt{f_{j_1,\ell,m}\, f_{j_2,\ell,m}}\, \mathcal{M}(\partial_t \psi_{b,\ell,m}\, \Delta t)\, v^{\text{obj}}_{b,\ell}\, \mathrm{e}^{\mathrm{i}\, \psi_{b,\ell,m}}$$

Our objectif: Fit the data $\tilde{\boldsymbol{y}}$ w.r.t. the sought parameters $\boldsymbol{\theta}$.

Advantages (compared to phase closure):

- no loss of frames;
- no loss of Fourier phase information;
- **maximum likelihood** $\Rightarrow$ **optimal** and **unbiased** estimators.

The sought parameters:

$$\boldsymbol{\theta} = \{\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{f}, \boldsymbol{v}^{\mathsf{obj}}\}$$

(fluxes $\boldsymbol{f}$, optical path parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ and object complex visibilities $\boldsymbol{v}^{\mathsf{obj}}$) are estimated by maximum likelihood (non-stationary Gaussian noise):

$$\boldsymbol{\theta}^{\star} = \arg\min_{\boldsymbol{\theta}} \Big\{ L(\boldsymbol{\theta}) = (\mathbf{H} \cdot \boldsymbol{x}(\boldsymbol{\theta}) - \tilde{\boldsymbol{y}})^{\top} \cdot \mathbf{W} \cdot (\mathbf{H} \cdot \boldsymbol{x}(\boldsymbol{\theta}) - \tilde{\boldsymbol{y}}) \Big\},$$

with $\mathbf{W} = \mathrm{Cov}(\tilde{\boldsymbol{y}})$. The co-log-likelihood $L(\boldsymbol{\theta})$ can be rewritten as:

$$L(\boldsymbol{\theta}) = \text{const.} + (\boldsymbol{x}(\boldsymbol{\theta}) - \tilde{\boldsymbol{x}})^{\top} \cdot \mathbf{Q} \cdot (\boldsymbol{x}(\boldsymbol{\theta}) - \tilde{\boldsymbol{x}})$$

with the **reduced data** $\tilde{\boldsymbol{x}}$ given by solving the *normal equations*:

$$\underbrace{(\mathbf{H}^{\top} \cdot \mathbf{W} \cdot \mathbf{H})}_{\mathbf{Q}} \cdot \tilde{\boldsymbol{x}} = \mathbf{H}^{\top} \cdot \mathbf{W} \cdot \tilde{\boldsymbol{y}} \quad \Longrightarrow \quad \tilde{\boldsymbol{x}} = \underbrace{(\mathbf{H}^{\top} \cdot \mathbf{W} \cdot \mathbf{H})^{\dagger} \cdot \mathbf{H}^{\top} \cdot \mathbf{W}}_{\mathbf{R}} \cdot \tilde{\boldsymbol{y}}$$

Nota bene: $\mathbf{R}$ is a generalization of the *pixel to visibility matrix* (P2VM).

Important

Using the reduced data $\tilde{\boldsymbol{x}}$ with weights $\mathbf{Q}$ is equivalent (no loss of information) to using the pixel data $\tilde{\boldsymbol{y}}$.

We need the frame data $\tilde{\boldsymbol{y}}$ and a good estimate of their statistical weights $\mathbf{W}$.

- Poisson (photon noise) + Gaussian (detector noise) distribution approximated by non-stationary independent Gaussian distribution.
- Signal-dependent variance of the measured pixels:

$$\mathrm{Var}(\tilde{y}_{k,\ell,m}) = \eta\, y_{k,\ell,m} + \zeta$$

  where $y_{k,\ell,m} = \mathbb{E}(\tilde{y}_{k,\ell,m})$, $\eta = 1/\gamma$ and $\zeta = (\sigma/\gamma)^2 + 1/12$, with γ the detector gain (in e$^-$ per digital level) and σ the standard deviation of the detector noise (in e$^-$ per pixel per frame);
- Since $y_{k,\ell,m}$ unknown in practice, we considered different approximations of $y_{k,\ell,m} = \mathbb{E}(\tilde{y}_{k,\ell,m})$:
  - Mistral (Mugnier et al., 2004):

$$y_{k,\ell,m} \approx y^{\text{Mistral}}_{k,\ell,m} = \max\{0, \tilde{y}_{k,\ell,m}\}$$

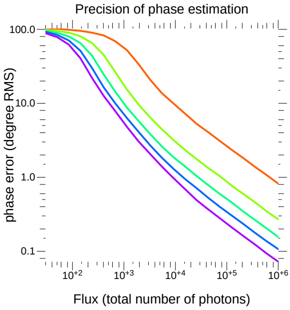
  - uniform level:

$$y_{k,\ell,m} \approx y^{\text{avg}}_{\ell,m} = \langle \tilde{y}_{k,\ell,m} \rangle_k$$

  - maximum likelihood:

$$y_{k,\ell,m} \approx y^{\text{ML}}_{k,\ell,m} = \arg\max_{y \geq 0} \left\{ \frac{(\tilde{y}_{k,\ell,m} - y)^2}{\eta\, y + \zeta} + \log(\eta\, y + \zeta) \right\}$$

Precision on the phase of the raw visibilities



Precision of the phase estimation of the raw complex visibilities for different fluxes (horizontal axis) and different fringe constrats (from 0.1 in orange to 0.9 in purple). Settings similar to AMBER (detector gain: 4.18 e^-/ADU, detector noise: 9.0 e^-/pixel RMS).

Important result:

The error on the estimation of the phase of the raw complex visibilities is almost the same as using the exact (but unknown) variance whatever our approximation of the variance.

Remind that the raw visibilities are part of the reduced data given by:

$$\tilde{\boldsymbol{x}} = (\mathbf{H}^\top \cdot \mathbf{W} \cdot \mathbf{H})^\dagger \cdot \mathbf{H}^\top \cdot \mathbf{W} \cdot \tilde{\boldsymbol{y}}$$

with $\mathbf{H}$ the generalized *visibility to pixel matrix* and the weighting matrix $\mathbf{W} = \mathrm{diag}(1/\,\mathrm{Var}(\tilde{\boldsymbol{y}}))$ depending on the approximation of the variance.

- Being linear combinations of the pixel values, measured raw complex visibilities $\tilde{v}_{b,\ell,m}^{\mathsf{raw}}$, and photometric fluxes $f_{j,\ell,m}$ are approximately **Gaussian random variables**.
- Based on mathematical approximations and simulations:

$$\left\{ \begin{array}{l} \mathrm{Var}\{\mathrm{Re}(\tilde{v}_{b,\ell,m}^{\mathsf{raw}})\} \approx \mathrm{Var}\{\mathrm{Im}(\tilde{v}_{b,\ell,m}^{\mathsf{raw}})\} \\ \mathrm{Cov}\{\mathrm{Re}(\tilde{v}_{b,\ell,m}^{\mathsf{raw}}), \mathrm{Im}(\tilde{v}_{b,\ell,m}^{\mathsf{raw}})\} \approx 0 \end{array} \right.$$

$$\Longrightarrow \quad \mathrm{Cov}(\tilde{v}_{b,\ell,m}^{\mathsf{raw}}) \approx \sigma_{b,\ell,m}^{2}\,\mathbf{I}$$

  thus Goodman (1985) model applies.
- Not yet checked:
  - independant random variables — for different (b, ℓ, m);
  - still true for multiplexed fringes (*e.g.* AMBER)

A critical issue to allow for integration of complex visibilities is the ability to align the fringes of all the frames. That is to compensate for the phase error terms $\psi_{b,\ell,m}$.

For a given baseline b, we propose to align the fringes of frame m_1 and frame m_2 for all spectral channels by minimizing the co-log-likelihood:

$$L = \sum_{\ell} \left\| \rho_{b,\ell,m_2}\, \mathbf{R}(\Delta\psi_{b,\ell,m_1,m_2}) \cdot \tilde{\boldsymbol{v}}^{\text{raw}}_{b,\ell,m_1} - \rho_{b,\ell,m_1}\, \tilde{\boldsymbol{v}}^{\text{raw}}_{b,\ell,m_2} \right\|^2_{\mathbf{W}_{b,\ell,m_1,m_2}}$$

with respect to $\Delta\boldsymbol{\alpha}$ and $\Delta\boldsymbol{\beta}$ such that:

$$\Delta\psi_{b,\ell,m_1,m_2} = \psi_{b,\ell,m_2} - \psi_{b,\ell,m_1} = \Delta\alpha_{b,\ell,m_1,m_2} + \Delta\beta_{b,\ell,m_1,m_2}/\lambda_\ell\,.$$

Nota bene:

- $\mathbf{R}(\Delta\psi)$ is a 2-D rotation matrix (same as multiplying the raw complex visibility by a phasor of argument $\Delta\psi$) and allows to align the phases;
- the non-negative factors $\rho_{b,\ell,m}$ account for contrast change;
- the statistical weights are:

$$\begin{aligned} \mathbf{W}_{b,\ell,m_1,m_2} &= \mathrm{Cov}(\rho_{b,\ell,m_2}\, \mathbf{R} \cdot \tilde{\boldsymbol{v}}^{\text{raw}}_{b,\ell,m_1} - \rho_{b,\ell,m_1}\, \tilde{\boldsymbol{v}}^{\text{raw}}_{b,\ell,m_2})^{-1} \\ &= \left[\rho^2_{b,\ell,m_2}\, \mathbf{R} \cdot \mathrm{Cov}(\tilde{\boldsymbol{v}}^{\text{raw}}_{b,\ell,m_1}) \cdot \mathbf{R}^\top + \rho^2_{b,\ell,m_1} \cdot \mathrm{Cov}(\tilde{\boldsymbol{v}}^{\text{raw}}_{b,\ell,m_2})\right]^{-1} \\ &\approx (\rho^2_{b,\ell,m_2}\, \sigma^2_{b,\ell,m_1} + \rho^2_{b,\ell,m_1}\, \sigma^2_{b,\ell,m_2})^{-1}\, \mathbf{I} \end{aligned}$$

Choosing $\rho_{b,\ell,m} = |\tilde{v}^{\mathsf{raw}}_{b,\ell,m}|$ yields:

$$\boxed{L = \sum_{\ell} w_{b,\ell,m_1,m_2} \left| \mathrm{e}^{\mathrm{i}\,\widetilde{\Delta\psi}_{b,\ell,m_1,m_2}} - \mathrm{e}^{\mathrm{i}\,\Delta\psi_{b,\ell,m_1,m_2}} \right|^2}$$

with weights:

$$w_{b,\ell,m_1,m_2} = \frac{\rho^2_{b,\ell,m_1}\,\rho^2_{b,\ell,m_2}}{\rho^2_{b,\ell,m_2}\,\sigma^2_{b,\ell,m_1} + \rho^2_{b,\ell,m_1}\,\sigma^2_{b,\ell,m_2}}$$

and phase difference data:

$$\widetilde{\Delta\psi}_{b,\ell,m_1,m_2} = \arg(\tilde{v}^{\mathsf{raw}}_{b,\ell,m_2}) - \arg(\tilde{v}^{\mathsf{raw}}_{b,\ell,m_1})$$

modelled by:

$$\Delta\psi_{b,\ell,m_1,m_2} = \Delta\alpha_{b,\ell,m_1,m_2} + \Delta\beta_{b,\ell,m_1,m_2}/\lambda_\ell$$

with

$$\begin{aligned}
\Delta\alpha_{b,\ell,m_1,m_2} &= (\alpha_{j_2(b),m_2} - \alpha_{j_1(b),m_2}) - (\alpha_{j_2(b),m_1} - \alpha_{j_1(b),m_1}) \\
\Delta\beta_{b,\ell,m_1,m_2} &= (\beta_{j_2(b),m_2} - \beta_{j_1(b),m_2}) - (\beta_{j_2(b),m_1} - \beta_{j_1(b),m_1})
\end{aligned}$$

We now have separable 2-D problems (for given baseline b and frame numbers m_1 and m_2) of minimizing:

$$L(\Delta\alpha, \Delta\beta) = \sum_{\ell} w_\ell \left| \mathrm{e}^{\mathrm{i}\,\widetilde{\Delta\psi}_{b,\ell,m_1,m_2}} - \mathrm{e}^{\mathrm{i}\,(\Delta\alpha + \Delta\beta/\lambda_\ell)} \right|^2$$

with respect to $\Delta\alpha$ and $\Delta\beta$. Require **global optimization**. However analytical solution for:

$$\Delta\alpha^+ \overset{\text{def}}{=} \arg\min_{\Delta\alpha} L(\Delta\alpha, \Delta\beta)$$

exists:

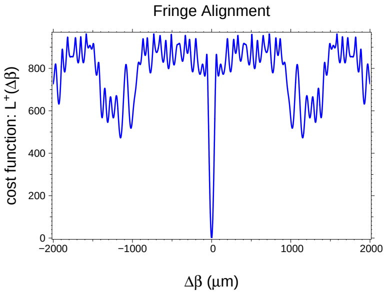
$$\mathrm{e}^{\mathrm{i}\,\Delta\alpha^+} = \frac{u(\Delta\beta)}{|u(\Delta\beta)|} \quad \text{with:} \quad \boxed{u(\Delta\beta) = \sum_{\ell} w_\ell\, \mathrm{e}^{\mathrm{i}\,\left(\widetilde{\Delta\psi}_{b,\ell,m_1,m_2} - \Delta\beta/\lambda_\ell\right)}}$$

replacing $\Delta\alpha$ by $\Delta\alpha^+$ yields a criterion which only depends on $\Delta\beta$:

$$L^+(\Delta\beta) \overset{\text{def}}{=} L(\Delta\alpha^+, \Delta\beta) = \sum_{\ell} w_\ell \left| \mathrm{e}^{\mathrm{i}\,\widetilde{\Delta\psi}_{b,\ell,m_1,m_2}} - \frac{u(\Delta\beta)}{|u(\Delta\beta)|}\, \mathrm{e}^{\mathrm{i}\,\Delta\beta/\lambda_\ell} \right|^2$$

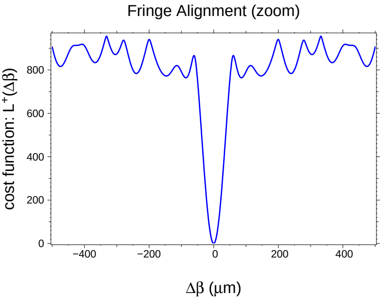
which requires only 1D global optimization (easy).

Maximum likelihood fringe alignment (results)



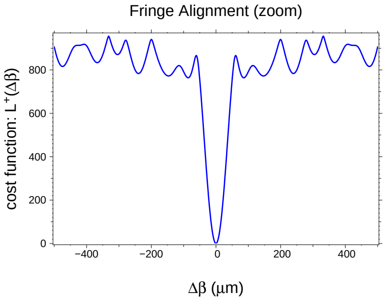
- Maximum likelihood estimate of β is done by 1D global optimization of the criterion (see curve above):

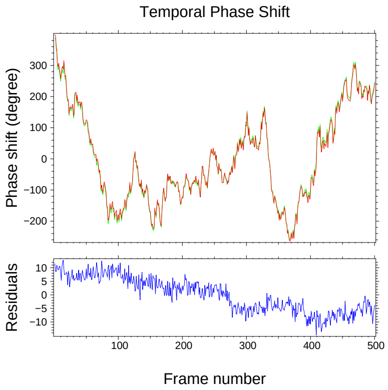
$$L^{+}(\Delta\beta) \stackrel{\text{def}}{=} L(\Delta\alpha^{+}, \Delta\beta)$$



- Maximum likelihood estimate of β is done by 1D global optimization of the criterion (see curve above):

$$L^{+}(\Delta\beta) \stackrel{\text{def}}{=} L(\Delta\alpha^{+}, \Delta\beta)$$

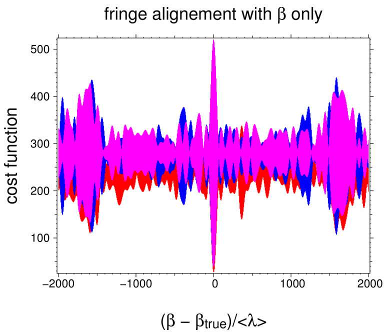


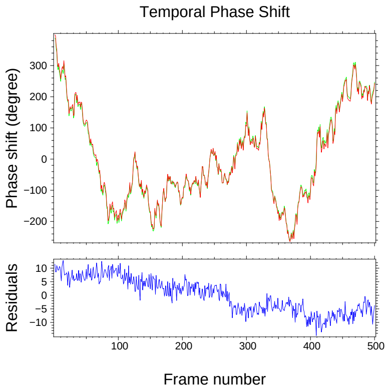


In green: the true phase shift at 2.2 μm. In red: the estimated phase shift. In blue: the residuals. Data: 20 spectral channels in K band.

- Maximum likelihood estimate of β is done by 1D global optimization of the criterion (see curve above):

$$L^{+}(\Delta\beta) \stackrel{\text{def}}{=} L(\Delta\alpha^{+}, \Delta\beta)$$





In green: the true phase shift at 2.2 µm. In red: the estimated phase shift. In blue: the residuals.
Data: 20 spectral channels in K band.

- Maximum likelihood estimate of β is done by 1D global optimization of the criterion (see curve above): Sans paramètre $\Delta\alpha$

$$L^+(\Delta\beta) \stackrel{\text{def}}{=} L(\Delta\beta)$$

Maximum likelihood fringe alignment (summary)

Our alignment method:

1. is **insentive to the phase of the object complex visibility**;
2. is **robust with respect to constrast changes and insensitive to amplitude of object complex visibility**;
3. can be solved easily (separable 1D global optimization problems).

Nota bene: points 1 and 2 are two major issues with fringe sensors.

Next (and final) stage is to fit the other parameters...

Our objective is to find the global minimum of the co-log-likelihood criterion:

$$L(\boldsymbol{\theta}) = \sum_{b,\ell,m} \frac{\left| \sqrt{f_{j_1(b),\ell,m}\, f_{j_2(b),\ell,m}}\, \mathcal{M}(\partial_t \psi_{b,\ell,m}\, \Delta m)\, v^{\mathsf{obj}}_{b,\ell}\, \mathrm{e}^{\mathrm{i}\,\psi_{b,\ell,m}} - \tilde{v}^{\mathsf{raw}}_{b,\ell,m} \right|^2}{\sigma^2_{b,\ell,m}} + \sum_{j,\ell,m} \frac{[f_{j,\ell,m} - \tilde{f}_{j,\ell,m}]^2}{\mathrm{Var}(\tilde{f}_{j,\ell,m})}$$

with:

$$\psi_{b,\ell,m} = [\alpha_{j_2(b),m} - \alpha_{j_1(b),m}] + [\beta_{j_2(b),m} - \beta_{j_1(b),m}]/\lambda_\ell$$

We assume that the proposed alignment method provides good initial values for the optical path parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ to devise our optimization strategy...

0. pre-align frames as explained, yields $\alpha_{j,m}^{(0)}$ and $\beta_{j,m}^{(0)}$, and choose initial fluxes, for instance:
$$f_{j,\ell,m}^{(0)} = \tilde{f}_{j,\ell,m}\,;$$
set $n = 0$ (iteration number); and repeat steps 1–4 until convergence:
1. fit object complex visibilities (closed solution):
$$\boldsymbol{v}^{\mathsf{obj}\,(n+1)} = \operatorname*{arg\,min}_{\boldsymbol{v}^{\mathsf{obj}}} L(\boldsymbol{\alpha}^{(n)}, \boldsymbol{\beta}^{(n)}, \boldsymbol{f}^{(n)}, \boldsymbol{v}^{\mathsf{obj}})$$
2. fit piston parameters:
$$(\boldsymbol{\alpha}^{(n+1)}, \boldsymbol{\beta}^{(n+1)}) = \operatorname*{arg\,min}_{\boldsymbol{\alpha},\boldsymbol{\beta}} L(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{f}^{(n)}, \boldsymbol{v}^{\mathsf{obj}\,(n+1)})$$
done by local optimization, starting at the previous solution;
3. fit fluxes ($L \times T$ separable global optimisation problems of size J, the number of telescopes):
$$f_{j,\ell,m}^{(n+1)} = \operatorname*{arg\,min}_{f_{j,\ell,m}} L(\boldsymbol{\alpha}^{(n+1)}, \boldsymbol{\beta}^{(n+1)}, \boldsymbol{f}, \boldsymbol{v}^{\mathsf{obj}\,(n+1)})$$
4. cancel average piston;

Frame alignment is a post-processing coherent integration of visibilities:

- preserves most of the information,
- the errors are still Gaussian,
- all telescopes do not need to be cophased at the same time (unlike phase closures),
- gives object complex visibilies $\boldsymbol{v}^{\mathrm{obj}}$ up to a chromatic piston $\longrightarrow$ **self calibration**

After the proposed pre-processing, the data are complex visibilities known up to a phase shift:

$$\tilde{y}_{b,\ell,p} = t_{b,\ell,p}\,(\mathbf{F}\cdot\boldsymbol{x})_{b,\ell,p}\,\mathrm{e}^{\mathrm{i}\,[\Delta\alpha_{b,p}+\Delta\beta_{b,p}/\lambda_\ell]} + \underbrace{\tilde{n}_{b,\ell,p}}_{\text{noise}}$$

in matrix notation:

$$\tilde{\boldsymbol{y}} = \boldsymbol{t}\,\mathbf{R}(\boldsymbol{\alpha},\boldsymbol{\beta})\cdot\mathbf{F}\cdot\boldsymbol{x} + \tilde{\boldsymbol{n}}$$

with $\mathbf{R}(\boldsymbol{\alpha},\boldsymbol{\beta})$ a block diagonal matrix with 2×2 blocks corresponding to rotations by angle $\Delta\alpha_{b,p} + \Delta\beta_{b,p}/\lambda_\ell$.

Unknowns:

- $\boldsymbol{x}$ = 3-D (spatio-spectral) object image;
- $(\boldsymbol{\alpha},\boldsymbol{\beta})$ = phase error parameters;
- $\boldsymbol{t}$ = instruments tranfert function.

Notations:

- $\tilde{\boldsymbol{y}}$ = complex visibility data;
- $\boldsymbol{x}$ = 3-D (spatio-spectral) object image;
- unknown phase errors:

$$\Delta\alpha_{b,p} = \alpha_{j_2(b),p} - \alpha_{j_1(b),p}$$
$$\Delta\beta_{b,p} = \beta_{j_2(b),p} - \beta_{j_1(b),p}$$

- $\boldsymbol{t}$ = Instrument transfert function;
- $\tilde{\boldsymbol{n}}$ = noise;
- $\mathbf{F}$ = non-uniform Fourier transform;
- p = exposure index;
- ℓ = wavelength index;
- b = baseline index;
- $j_1(b)$ and $j_2(b)$ = telescopes involved in b-th baseline;

Image reconstruction (inverse problem)

$$\begin{pmatrix} \boldsymbol{x}^\star \\ \boldsymbol{\alpha}^\star \\ \boldsymbol{\beta}^\star \\ \boldsymbol{t}^\star \end{pmatrix} = \underset{\boldsymbol{x},\boldsymbol{\alpha},\boldsymbol{\beta}}{\arg\min} \Big\{ \underbrace{\mathcal{J}_{\text{data}}(\boldsymbol{t}\,\mathbf{R}(\boldsymbol{\alpha},\boldsymbol{\beta})\cdot\mathbf{F}\cdot\boldsymbol{x})}_{\text{data fidelity}} + \underbrace{\mu\,\mathcal{J}_{\text{prior}}(\boldsymbol{x})}_{\text{regularization}} \Big\} \text{s.t.} \left\{ \begin{array}{l} \boldsymbol{x} \geq \mathbf{0} \\ \underbrace{\|\boldsymbol{t}\,\boldsymbol{c} - \tilde{\boldsymbol{c}}\|_2^2}_{\text{Calibration}} \leq \eta \end{array} \right.$$

Alternate minimization algorithm:

0. Choose initial phase shift parameters $\boldsymbol{\alpha}^{(0)}$ and $\boldsymbol{\beta}^{(0)}$. Set $\boldsymbol{t} = \tilde{\boldsymbol{c}}/\boldsymbol{c}$, $k = 0$.
   Repeat until convergence:
1. image reconstruction step:

$$\boldsymbol{x}^{(k+1)} = \underset{\boldsymbol{x}\geq\mathbf{0}}{\arg\min} \Big\{ \mathcal{J}_{\text{data}}(\mathbf{R}(\boldsymbol{\alpha}^{(k)},\boldsymbol{\beta}^{(k)})\cdot\mathbf{F}\cdot\boldsymbol{x}) + \mu\,\mathcal{J}_{\text{prior}}(\boldsymbol{x}) \Big\};$$

2. piston self-calibration step:

$$(\boldsymbol{\alpha}^{(k+1)},\boldsymbol{\beta}^{(k+1)}) = \underset{\boldsymbol{\alpha},\boldsymbol{\beta}}{\arg\min}\, \mathcal{J}_{\text{data}}(\mathbf{R}(\boldsymbol{\alpha},\boldsymbol{\beta})\cdot\mathbf{F}\cdot\boldsymbol{x}^{(k+1)});$$

3. transfert function calibration step
4. let $k = k + 1$;

We propose a method for **frame alignment** which:

- is robust w.r.t. noise and insensitive to object complex visibility (unlike conventional fringe tracking);
- exploits multiple wavelengths;
- allows for **coherent integration** of visibilities during an exposure;
- avoids loss of frames;
- avoids loss Fourier phase information (compared to phase closures);
- could account for atmospheric refraction (higher order expansion of the phase error terms);
- provides object complex visibilities up to unknown phase bias which depends on few parameters
  $\Longrightarrow$ image reconstruction requires self-calibration (much easier than working with phase closures and powerspectrum);

Complex variables are also 2D real vectors: $\mathbb{C} \sim \mathbb{R}^2$

Convention:

$$\forall u \in \mathbb{C}, \quad u \sim \boldsymbol{u} = (u_1, u_2)^\top \in \mathbb{R}^2 \quad \text{with } u_1 = \mathrm{Re}(u) \text{ and } u_2 = \mathrm{Im}(u).$$

With this convention, any complex random variable $\tilde{u} \in \mathbb{C}$ has an expected value:

$$\mathbb{E}(\tilde{u}) \sim \mathbb{E}(\tilde{\boldsymbol{u}}) = \left[\begin{array}{c} \mathrm{Re}\big(\mathbb{E}(\tilde{u})\big) \\ \mathrm{Im}\big(\mathbb{E}(\tilde{u})\big) \end{array} \right],$$

(which directly follows from linearity) and has an associated 2×2 covariance matrix defined by:

$$\mathrm{Cov}(\tilde{u}) \sim \mathrm{Cov}(\tilde{\boldsymbol{u}}) = \begin{pmatrix} C_1 & C_3 \\ C_3 & C_2 \end{pmatrix}, \quad \text{with: } \begin{array}{l} C_1 = \mathrm{Var}\{\mathrm{Re}(\tilde{u})\}, \\ C_2 = \mathrm{Var}\{\mathrm{Im}(\tilde{u})\}, \\ C_3 = \mathrm{Cov}\{\mathrm{Re}(\tilde{u}), \mathrm{Im}(\tilde{u})\}. \end{array}$$

The corresponding weighting matrix is:

$$\mathbf{W} = \mathrm{Cov}(\tilde{\boldsymbol{u}})^{-1} = \frac{1}{C_1\, C_2 - C_3^2} \begin{pmatrix} C_2 & -C_3 \\ -C_3 & C_2 \end{pmatrix}.$$