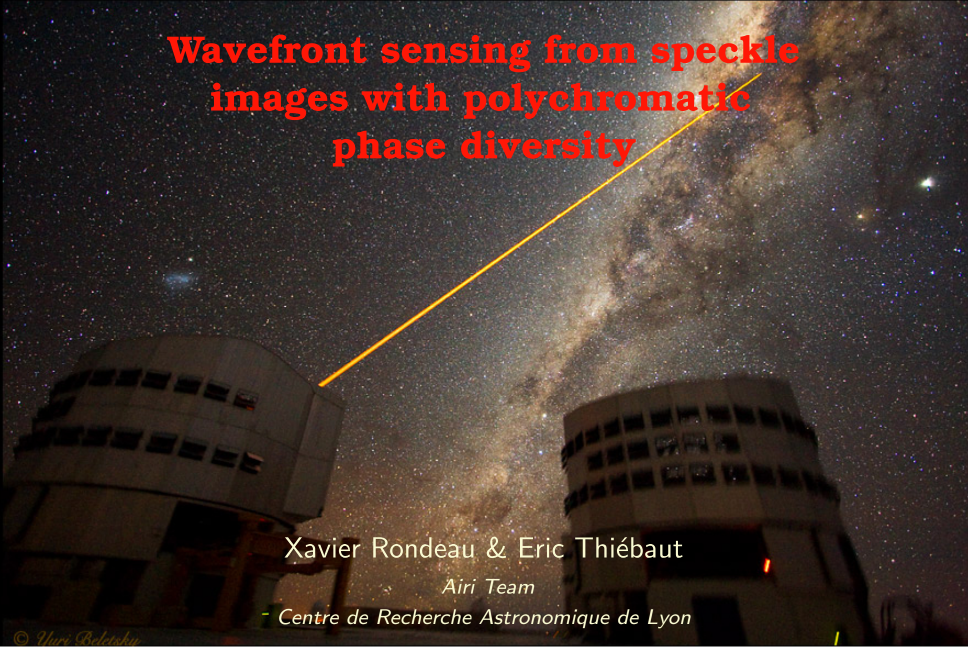
Wavefront sensing from speckle images with polychromatic phase diversity



Xavier Rondeau & Eric Thiébaut

Airi Team

Centre de Recherche Astronomique de Lyon

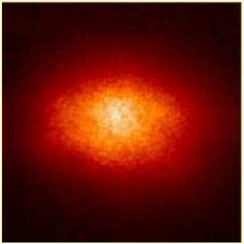
ADA-5th conference

Thursday, May 8$^{\mathrm{th}}$, 2008

High angular resolution in Astronomy

Angular resolution
- $\propto \lambda / D$ for an isolated telescope
- $\propto \lambda / r_0$ with atmospheric turbulence

$D/r_0(\lambda) \gg 1$



BOA-Onera

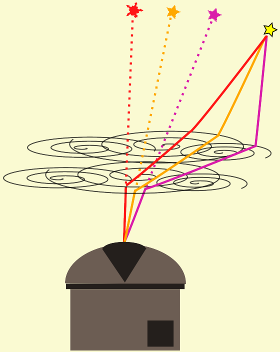
- $D \simeq 30\,,40\,\mathrm{m}$ (TMT,E-ELT), $\lambda = 500\,\mathrm{nm}$, $r_0(\lambda) = 15\,\mathrm{cm}$

$\Rightarrow$ **Resolution loss factor $D/r_0 \simeq 200$**

Polychromatic Laser Guide Star

R. Foy et al. 1995

AO limitations due to tip-tilt indetermination



⋆ Use of air refractive index chromaticity

Searched tip-tilt

Measured tip-tilts

$$\theta(\lambda) = \underbrace{\frac{n(\lambda) - 1}{n(\lambda_2) - n(\lambda_1)}}_{\zeta(\lambda_1, \lambda_2)} \times \overbrace{[\theta(\lambda_2) - \theta(\lambda_1)]}^{\text{Measured tip-tilts}}$$

Tip-tilt estimation

- **Amplification of the tip-tilt error :**

$$\sigma_{\theta_\lambda} = \zeta(\lambda_1, \lambda_2) \times \sigma_{\Delta\theta_{\lambda_1,\lambda_2}} \quad \text{where} \quad \zeta(\lambda_1, \lambda_2) > 25$$

- **Degradation of Cramér-Rao lower bound :**

$$\tilde{\theta}_{\mathrm{ML}} = \max_{\theta} g(\mathbf{x}|\theta) \implies {\sigma_\theta^+}^2 = -\left(E\left[\frac{\partial^2 \ln(g(\mathbf{x}|\theta))}{\partial\theta^2}\right]\right)^{-1}_{\big|\theta=\theta^+}$$

Airy pdf : $\sigma_\theta^+ = \frac{\kappa}{\sqrt{N_{\mathrm{ph}}}} \times \frac{\lambda}{D}$

Average pdf of speckle images : $\sigma_\theta^+ \simeq \frac{0.3\,\kappa}{\sqrt{N_{\mathrm{ph}}}} \times \frac{\lambda}{r_0}$

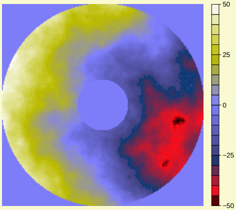
$\sim 0.3 \times D/r_0$ increase due to turbulence

Joint wavefront estimation from the images to attain the ultimate centering accuracy

The inverse problem

Forward model :

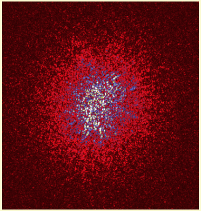
turbulent $\varphi(\mathbf{u})$



Telescope pupil

causes $\longrightarrow$

observed psf $h_{\lambda,t}(\mathbf{x})$



Telescope focal plane

$$\underbrace{\overset{\text{data}}{\mathbf{h}}}\; = \underbrace{\mathbf{m}}_{\text{model}}(\overset{\text{parameters}}{\varphi}) + \underbrace{\mathbf{n}}_{\text{noise}}$$

where $\mathbf{m}(\varphi) = \alpha \left| FT\left[\mathbf{P} \times \exp(j\varphi)\right]\right|^2$

The inverse approach :

Reconstruction of $\varphi(\mathbf{u})$ given data $h_{\lambda,t}(\mathbf{x})$, forward model $\mathbf{m}(\varphi)$ and statistics of noise $\mathbf{n}$

Bayesian formulation

$$\varphi^+ = \max_{\varphi}\{\Pr[\varphi|\mathbf{h}]\} = \max_{\varphi}\{\Pr[\mathbf{h}|\varphi] \times \Pr[\varphi]\}$$ **MAP criterion**

$$\rightarrow \textbf{optimization of } \mathbf{f}(\varphi) = \underbrace{\mathbf{f}_{\mathbf{exp}}(\varphi)}_{-\log(\Pr[\mathbf{h}|\varphi])} + \underbrace{\mu\, \mathbf{f}_{\mathbf{prior}}(\varphi)}_{-\log(\Pr[\varphi])}$$

Gaussian prior statistics

$$f_{\text{prior}}(\varphi) = \varphi^{\top}.\mathbf{C}_{\varphi}^{-1}.\varphi$$

where $\mathbf{C}_{\varphi} = \langle \varphi . \varphi^{\top} \rangle$ (Kolmogorov covariance)

Uncorrelated low count data

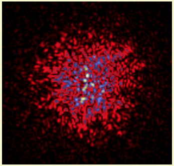
$$f_{\text{exp}}^{\text{Poisson}}(\varphi) = \sum_j [\mathbf{m}(\varphi) - \mathbf{h}\log(\mathbf{m}(\varphi))]_j$$

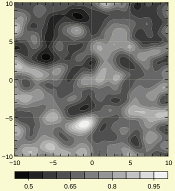
Higher count or additive gaussian noise

$$f_{\text{exp}}^{\text{Gauss}}(\varphi) = [\mathbf{h} - \mathbf{m}(\varphi)]^{\top}.\mathbf{C}_{\mathbf{n}}^{-1}.[\mathbf{h} - \mathbf{m}(\varphi)]$$

Ill-posed inverse problem

- Intrinsic degeneracies of $\mathbf{m}(\varphi)$
  - $\rightarrow$ sign of all even phase modes : $m\,[\varphi(\mathbf{r})] = m\,[-\varphi(-\mathbf{r})]$
  - $\rightarrow$ modulo $2\,\pi$ periodicity : $m\,[\varphi] = m\,[\varphi + 2\,\mathbf{k}(\mathbf{r})\,\pi] \quad \forall\,\varphi_j$
- Strong local minima of $f_{\exp}(\varphi)$ (non-convex criterion)
  - $\rightarrow$ Translation ambiguity (speckle overlapping between $\mathbf{m}$ and $\mathbf{h}$)





  - $\rightarrow$ Loose sign ambiguity of individual even phase modes
  - $\rightarrow$ Alteration of the speckle pattern (new speckles) due to the noise

Need of a global optimization strategy

Large-scale non-linear optimization

Though the criterion is non-convex, faster to make the most of the good local continuity — given size of the pb — than Monte-Carlo

Local quadratic approximation :

$$f(\varphi + \delta\varphi) - f(\varphi) = \mathbf{g}_\varphi{}^\top.\delta\varphi + \tfrac{1}{2}\delta\varphi^\top.\mathbf{B}_\varphi.\delta\varphi + o(\|\delta\varphi\|^2)$$

$\Rightarrow$ Newton step $\boxed{\delta\varphi^+ = -\mathbf{B}_\varphi{}^{-1}.\mathbf{g}_\varphi}$ with

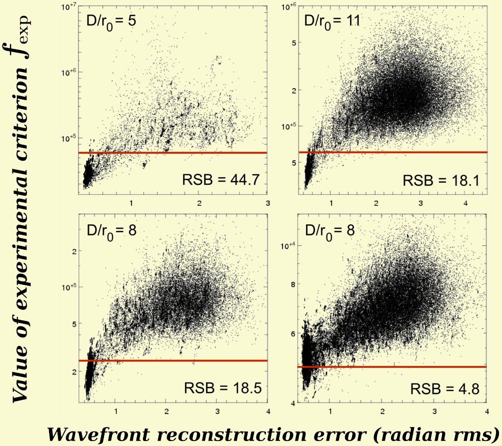
- $\mathbf{g} = \nabla f(\varphi)$ (gradient)
- $\mathbf{B} = \nabla^2 f(\varphi)$ (hessian)
- $o(\|\delta\varphi\|^2) \simeq 0$

- $\mathbf{B}_\varphi$ too expensive to compute owing to the framerate (~ 50 image/s)
- $\mathbf{B}_\varphi$ too big for inversion ($N_\varphi \times N_{\mathrm{data}} \sim 10^8$ scalars)

Need of fast and very efficient limited-memory local optimization

Correlation between values of criterion f_{exp} and wavefront reconstruction error

observable



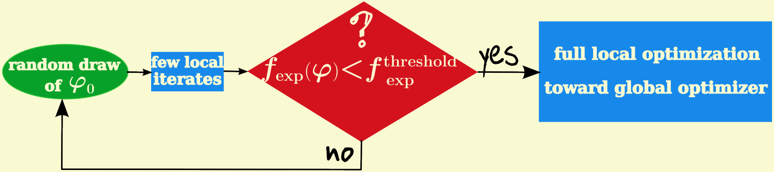
unknown

Global optimization strategy

➢ Uncorrelated wavefronts

No proper initialization available for local optimisation but:

- Global convergence guaranteed below some threshold $f_{\mathrm{exp}}^{\mathrm{threshold}}$
- Decision can be made in very few iterations because of faster convergence rate in the first iterations



➢ Time-correlated wavefronts

Only local optimization is needed by taking last reconstruction as an initial guess for next wavefront reconstruction

(better guess by modeling the wavefront evolution according to Taylor hypothesis)

Tuning of the regularization weight

Weight of convex f_{prior} is enforced in first iterations by tuning

$$\mu \quad \text{so that} \quad \delta f_{\text{exp}}[\varphi, \delta\boldsymbol{\varphi}(\mu)] \simeq \underset{\in\,[0,1]}{\underset{\uparrow}{\epsilon}} \times \delta f_{\text{exp}}[\varphi, \delta\boldsymbol{\varphi}(\mu = 0)]$$

where $\delta\boldsymbol{\varphi}(\mu) = -(\mathbf{B}_{\text{exp}} + \mu\,\mathbf{B}_{\text{prior}})^{-1} \cdot (\mathbf{g}_{\text{exp}} + \mu\,\mathbf{g}_{\text{prior}})$

★ **constrain the reconstruction to consistent φ by :**

- disentangling $m(\varphi)$ degeneracies
- smoothing out some local minima

★ **automatic adaptation to noise and turbulence strength**

Local optimization algorithm

Subspace of priviledged search directions: $\delta\varphi = \sum_{i=1}^{N_{\mathrm{dir}} \ll N_{\varphi}} \beta_i \mathbf{s}_i$

$$\boxed{\beta^+ = -\mathbf{B}'^{-1} . \mathbf{g}'} \quad \Longleftarrow \quad \delta f^{\mathrm{quad}}(\beta) = \mathbf{g}'^{\top} . \beta + \frac{1}{2} \beta^{\top} . \mathbf{B}' . \beta$$

where $\mathbf{B}' = \mathbf{S}^{\top} . \mathbf{B} . \mathbf{S}$ and $\mathbf{g}' = \mathbf{S}^{\top} . \mathbf{g}$

Local optimization algorithm

Subspace of priviledged search directions: $\delta\varphi = \sum_{i=1}^{N_{\rm dir} \ll N_\varphi} \beta_i \mathbf{s}_i$

$$\boxed{\beta^+ = -\mathbf{B}'^{-1}.\mathbf{g}'} \quad \Longleftarrow \quad \delta f^{\rm quad}(\beta) = \mathbf{g}'^\top.\beta + \frac{1}{2}\,\beta^\top.\mathbf{B}'.\beta$$

where $\mathbf{B}' = \mathbf{S}^\top.\mathbf{B}.\mathbf{S}$ and $\mathbf{g}' = \mathbf{S}^\top.\mathbf{g}$

- *Directions built according to Taylor expansion, using the prior to ensure valid metric with fast fractal implementation* (*FRIM Thiébaut & Tallon* 2008)

- $\mathbf{s}_1^{(k)} = \delta\varphi^{(k-1)}$
- $\mathbf{s}_2 = -\mathbf{B}_{\rm prior}^{-1}.\mathbf{g}_{\rm prior}$
- $\mathbf{s}_3 = -\mathbf{B}_{\rm prior}^{-1}.\mathbf{g}_{\rm exp}$
- $\mathbf{s}_{i+1} = -\mathbf{B}_{\rm prior}^{-1}.\mathbf{B}_{\rm exp}.\mathbf{s}_i$ $\quad$ ($e.g.$ $i = 4, 5$)

Local optimization algorithm

Subspace of priviledged search directions: $\delta\varphi = \sum_{i=1}^{N_{\rm dir} \ll N_\varphi} \beta_i \mathbf{s}_i$

$$\boxed{\beta^+ = -\mathbf{B}'^{-1}.\mathbf{g}'} \quad \Longleftarrow \quad \delta f^{\rm quad}(\beta) = \mathbf{g}'^\top.\beta + \frac{1}{2}\,\beta^\top.\mathbf{B}'.\beta$$

where $\mathbf{B}' = \mathbf{S}^\top.\mathbf{B}.\mathbf{S}$ and $\mathbf{g}' = \mathbf{S}^\top.\mathbf{g}$

- $\mathbf{A}_{\rm exp}$ *is an approximation of the too expensive hessian* $\mathbf{B}_{\rm exp}$*:*

Linearizing $\mathbf{m}(\varphi)$ yields

$$\mathbf{A}_{\rm exp}^{\rm gauss} = \mathbf{J}^\top.\mathbf{C}_\mathbf{n}^{-1}.\mathbf{J}$$

$$\mathbf{A}_{\rm exp}^{\rm poisson} = \mathbf{J}^\top.\mathbf{Diag}\left[\frac{\mathbf{h}}{\mathbf{m}^2}\right].\mathbf{J}$$

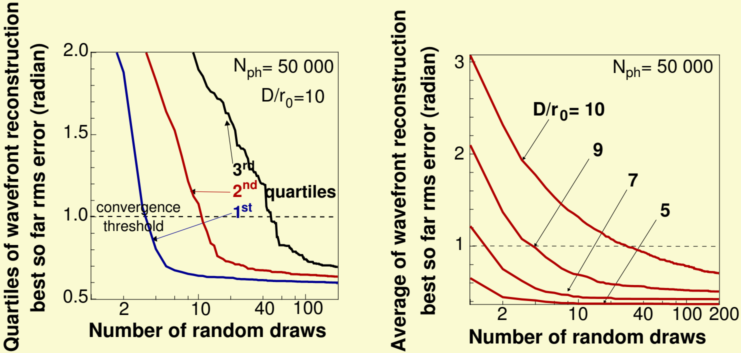
where jacobian $\mathbf{J}_{k,l} = \frac{\partial m_k(\varphi)}{\partial \varphi_l}$ need not be formed and is implemented for fast vector product

- *Step-length controled via trust-region to valid quadradic approximation*

$$\min_{\delta\varphi} \delta f_\varphi^{\rm quad}(\delta\varphi) \quad \text{so that} \quad \|\delta\varphi\|_\mathbf{M}^2 \leqslant \Delta$$

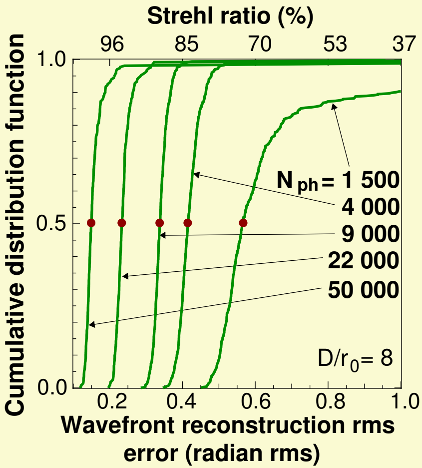
Cost with respect to the turbulence strength

Number of random wavefront draws (with prior covariance) needed to attain a given reconstruction error after very few local iterates is investigated through quartiles of a large sampling of wavefront's statistics



High increase of local minima with D/r_0 (bad tip-tilt estimation by image centroïd)

Final convergence



In agreement with SNR degradation. Could be compared to a linearized estimator of the covariance error assuming local convexity

Polychromatic phase diversity

Use of images at different wavelengths with same optical path

Polychromatic model: $\varphi_{\lambda_i}(\mathbf{r}) = \frac{\lambda_{\mathrm{ref}}}{\lambda_i} \frac{(n_{\lambda_i} - 1)}{(n_{\lambda_{\mathrm{ref}}} - 1)} \times \varphi_{\lambda_{\mathrm{ref}}}(\mathbf{r})$

MAP criterion: $f(\varphi_{\lambda_{\mathrm{ref}}}) = \sum_k \mu_k \, f_{\mathrm{exp},\lambda_k}(\varphi_{\lambda_{\mathrm{ref}}}) + f_{\mathrm{prior}}(\varphi_{\lambda_{\mathrm{ref}}})$

★ $f_{\mathrm{exp},\lambda}$ at long λ used as a smooth guide for shorter wavelength: (local minima and degeneracies $\nearrow$ when $\lambda \searrow$)

- μ_1^+ so that $\delta f_{\mathrm{exp},\lambda_1}[\delta\varphi(\mu_1, \mu_2{=}0)] \simeq \underset{\in [0,1]}{\underset{\downarrow}{\epsilon_1}} \times \delta f_{\mathrm{exp},\lambda_1}[\delta\varphi(\mu_1 = \infty, \mu_2{=}0)]$

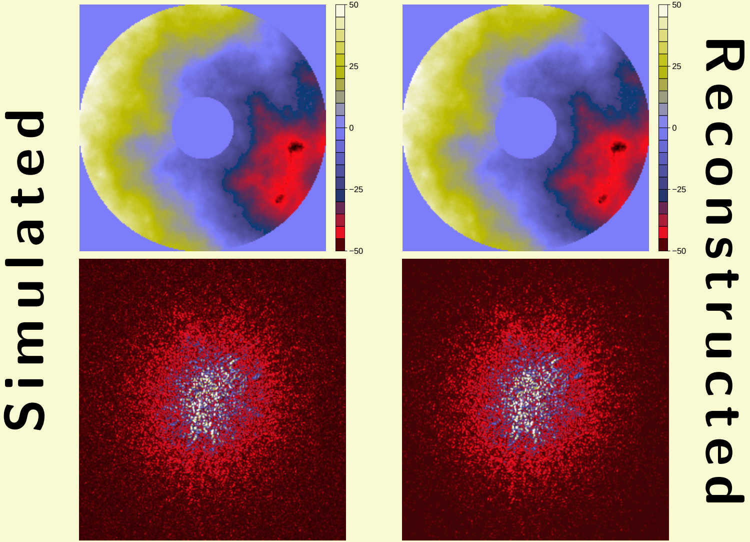
then

- μ_2^+ so that $\delta f_{\mathrm{exp},\lambda_2}[\delta\varphi(\mu_1^+, \mu_2)] \simeq \underset{\in [0,1]}{\underset{\downarrow}{\epsilon_2}} \times \delta f_{\mathrm{exp},\lambda_2}[\delta\varphi(\mu_1^+, \mu_2 = \infty)]$

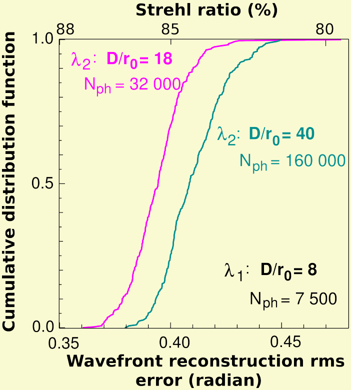
Polychromatic level-arm enables to deal with much higher D/r_0 (~80 i.e. $N_\varphi \geqslant 7000$)

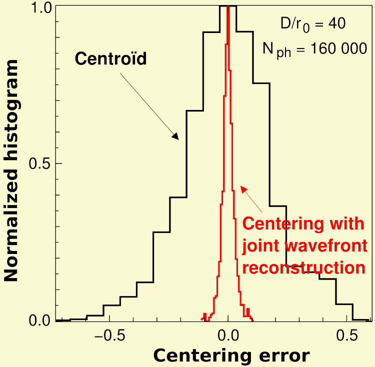
Results

$D/r_0 = 58$ et $N_{ph} = 500\,000$



Results





Effective ultimate centering accuracy thanks to joint wavefront sensing

Conclusion and perpectives

Rondeau et al. 2007

★ *Global optimisation for the wavefront reconstruction pb*

- **D/r_0 from 4** (previous works) **to 11** with a single wavelength
- **$D/r_0 \simeq 70$** using **polychromatic diversity** with 2 wavelengths
- **Automatic strategy** with respect to noise and turbulence fluctuations
- **Fast** limited-memory **local optimization and global strategy**

☞ Application to myopic deconvolution or adaptive optics

★ *Benefit for tip-tilt estimation* $\simeq \mathbf{0.3\, D/r_0}$

☞ Blind deconvolution with an interferometric P-LGS