Processing of Polychromatic Interferometric Data for Astrophysics : the POLCA project

Michel Tallon & POLCA consortium

CRAL, Lyon, France
IPAG, Grenoble, France
Lagrange, Nice, France
LESIA, Paris, France

- Short presentation of POLCA project
- Polychromatic data
  - Analysis of (polychromatic) interferometric data
  - Visibilities for image reconstruction
    - → Ferréol Soulez's talk
- 3-D (x,y,λ) model fitting
- 3-D (x,y,λ) image reconstruction
  - specific approach → Jacques Kluska's talk
  - general approach → Ferréol Soulez's talk

Summary of POLCA / 1

- People (26) from 4 French labs
  - *CRAL (Lyon)*: Paul Berlioz-Arthaud, Ferréol Soulez, Michel Tallon, Isabelle Tallon-Bosc, Éric Thiébaut
  - *IPAG (Grenoble)*: Gilles Duvert, Jacques Kluska, Sylvain Lafrasse, Bernard Lazareff, Jean-Baptiste Lebouquin, Fabien Malbet, Guillaume Mella
  - *Lagrange (Nice)*: Philippe Bério, Olivier Chesneau, André Ferrari, David Mary, Florentin Millour, Denis Mourard, Romain Petrov, Antony Schutz, Céline Theys, Martin Vannier
  - *LESIA (Paris)*: Pierre Kervella, Sylvestre Lacour, Thibaut Paumard, Guy Perrin
- Ethnography :
  - Experienced observers in interferometry
  - Experts in signal processing
  - Experts on instruments (and their data)
    - AMBER, MIDI, VEGA, PIONIER, GRAVITY, MATISSE
- => Nobody understands everything
- Supported by ANR (French National Research Agency)
  - 4 years (2011 — 2015)
  - 400 k€
  - ~ 6 FTE during the 4 years

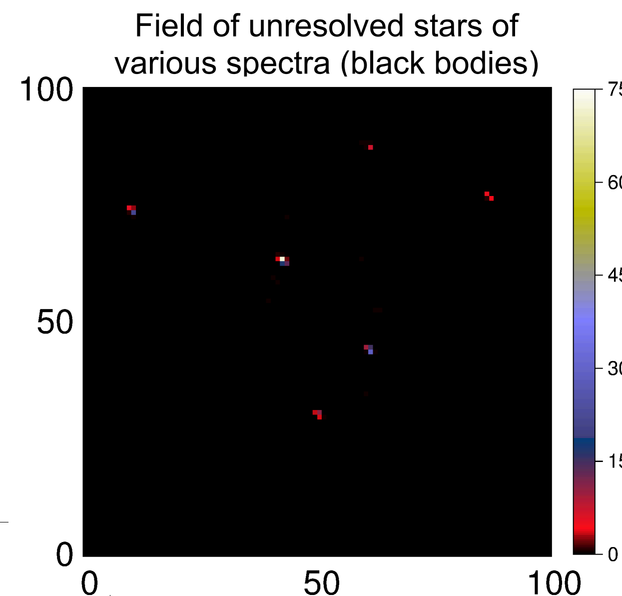
Summary of POLCA / 2

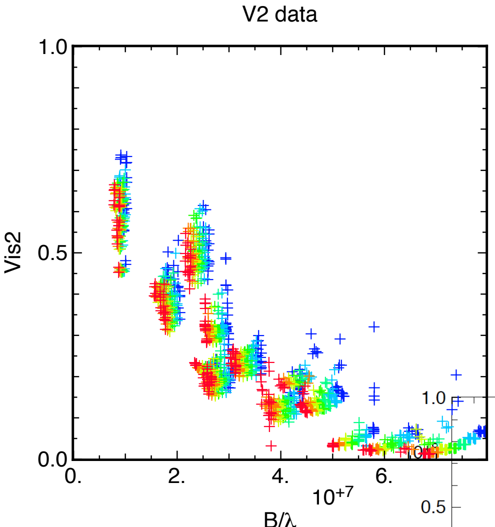
- Facts
  - Data now contain a lot of information along the wavelength (dispersed fringes + spectrum)
  - Needs of better tools to make the most of these data
    - Particularly, how to exploit differential visibilities ?
  - Need to make experts in interferometry and in signal processing work together
  - Need to work on real data

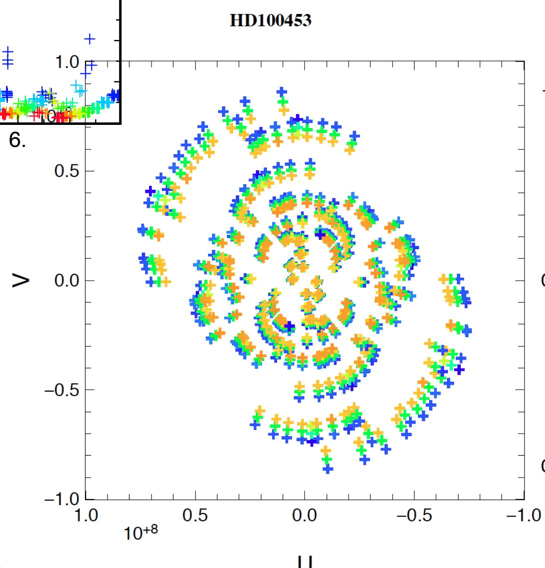
- Outputs
  - New methods and new algorithms for polychromatic data
    - 3-D image reconstruction
    - model fitting
    - image reconstruction + model-fitting
  - Better knowledge of data
    - requirements for better data
    - => better interferometers and instruments
  - Advance in astrophysics (as much as possible)
    - => work with real data

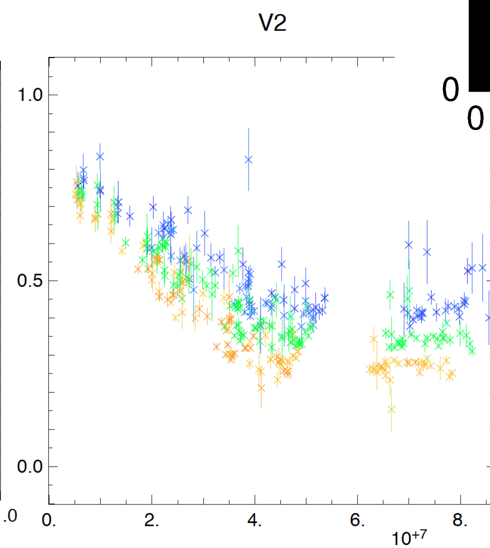
Selection of suitable data

- Selection of a limited set of data with chromatic features
  - in the objects
  - in the data
  - Simulated data in some cases (GRAVITY)







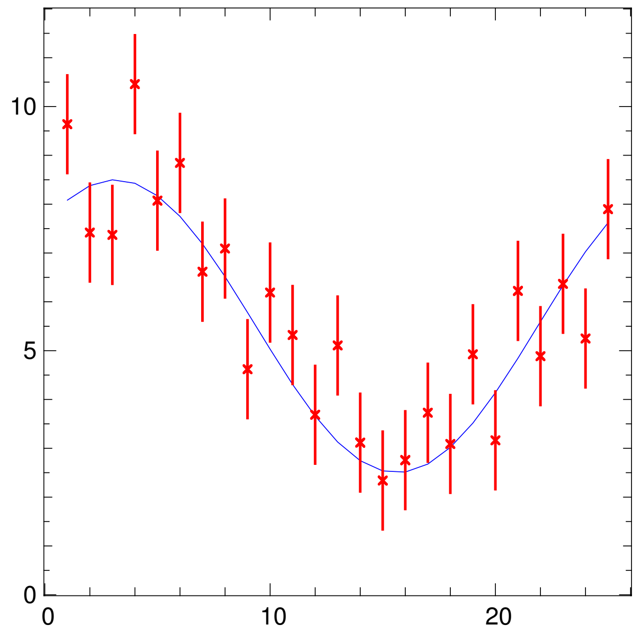


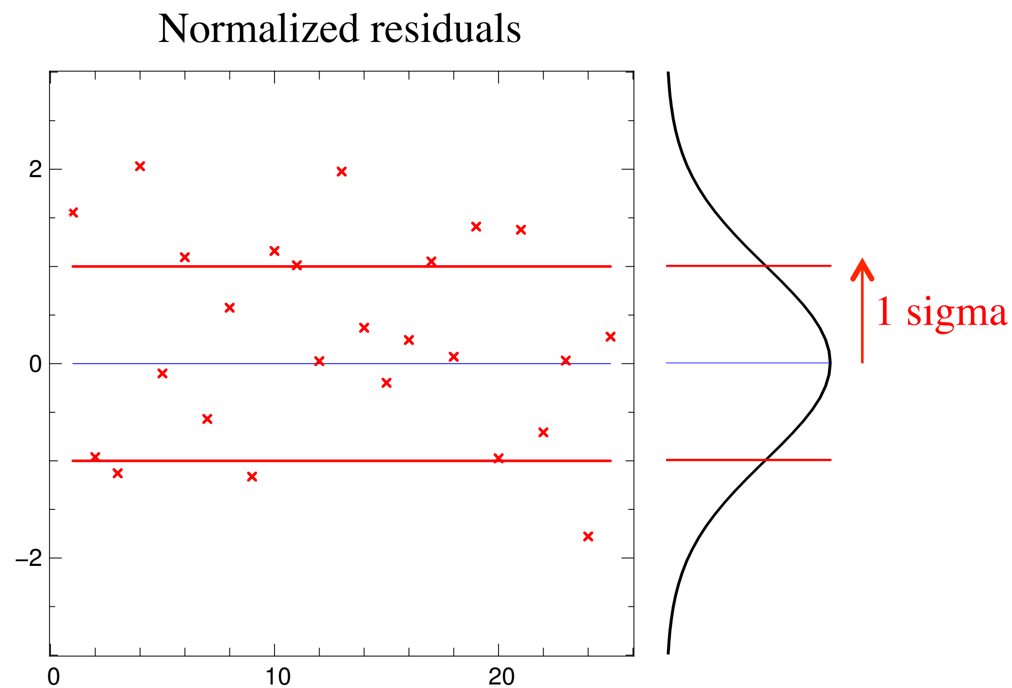
Statistics of polychromatic data

Current processings (image reconstruction, model fitting, …) assume uncorrelated gaussian data…

$$\text{Minimization of: } \chi^2(\boldsymbol{x}) = \sum_{i=1}^{N_{\text{data}}} \left(\frac{d_i - m_i(\boldsymbol{x})}{\sigma_i} \right)^2$$

Appearance of independence of data





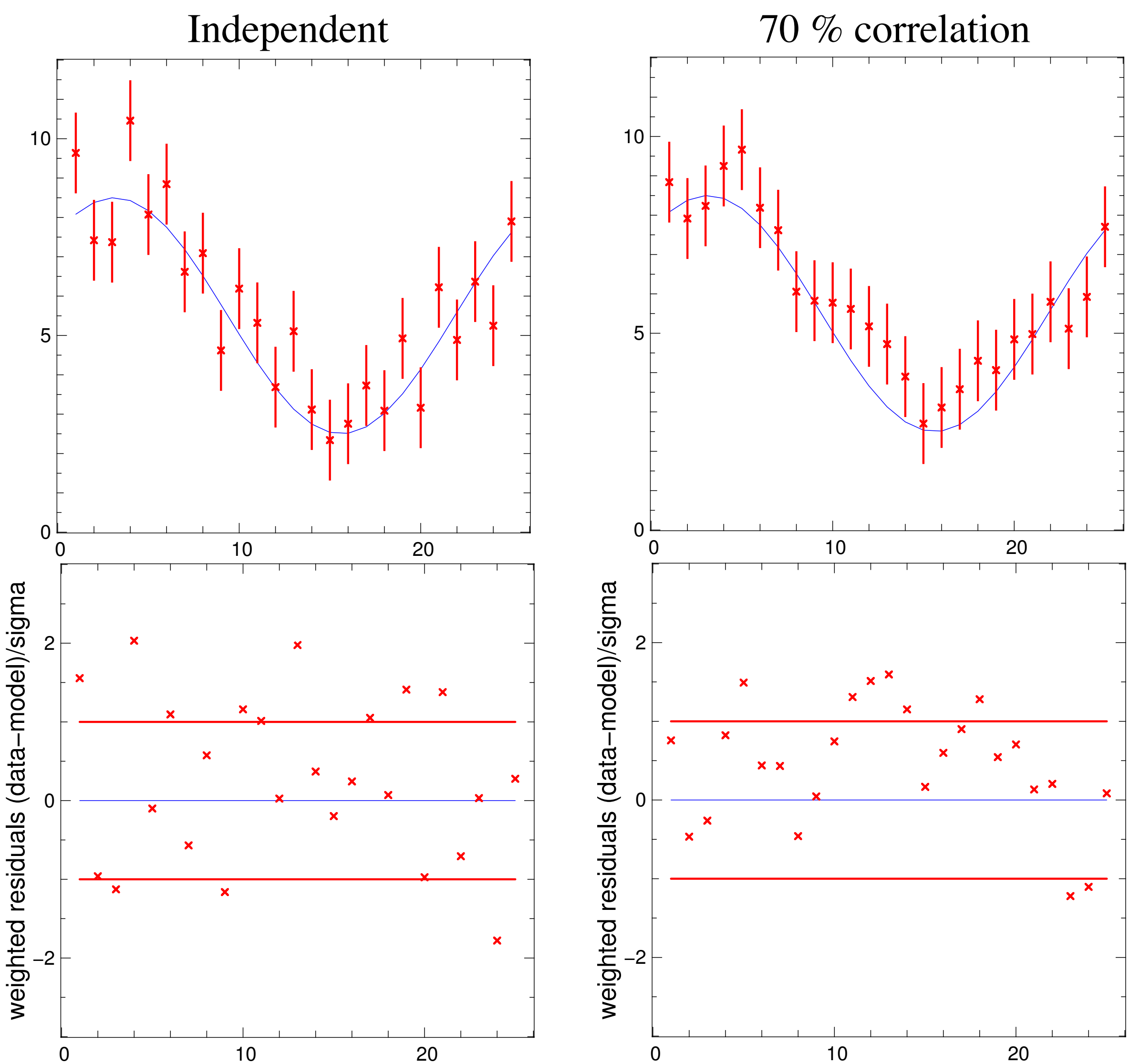
$$r_i = \frac{d_i - m_i}{\sigma_i}$$

- simulated data
- model is outside the error bars (1 sigma) for ~32% of the data

- ~32% of data out of the range [-σ, σ]

Beware : only one realization here !

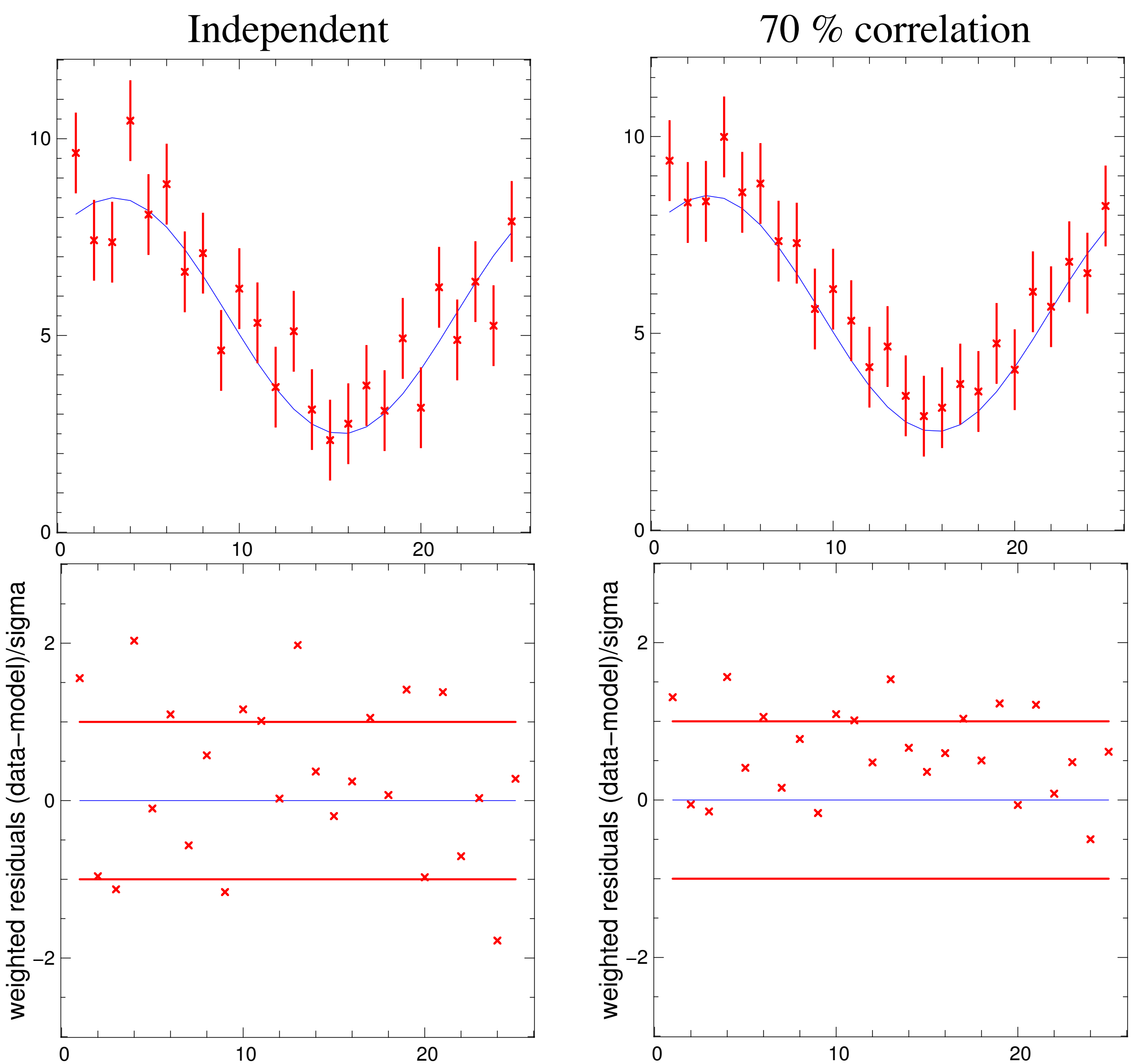
Data with local correlations: 70%



- correlation coefficient:
  - 70% with adjacent
  - 25% with next point
  - 0% farther away
- Similar effect as spectral correlations in real data
- alignments of successive points
- less dispersion of residuals

Beware : only one realization !

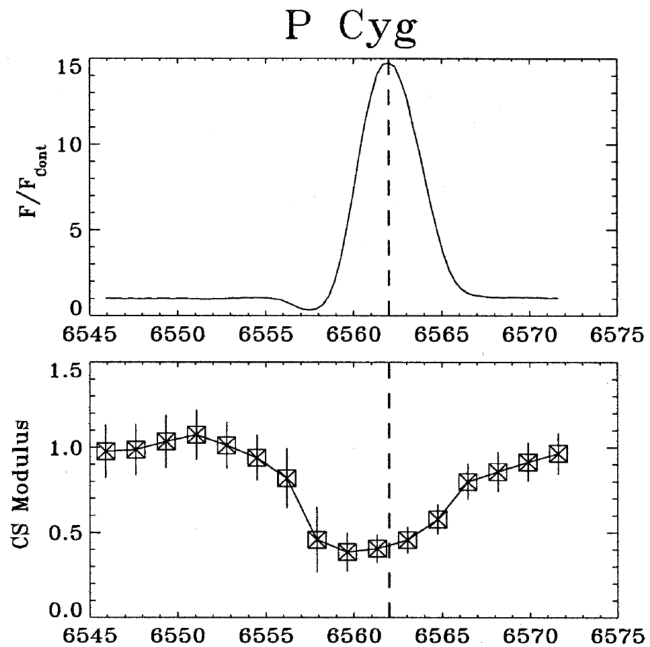
Data with global correlations: 70%



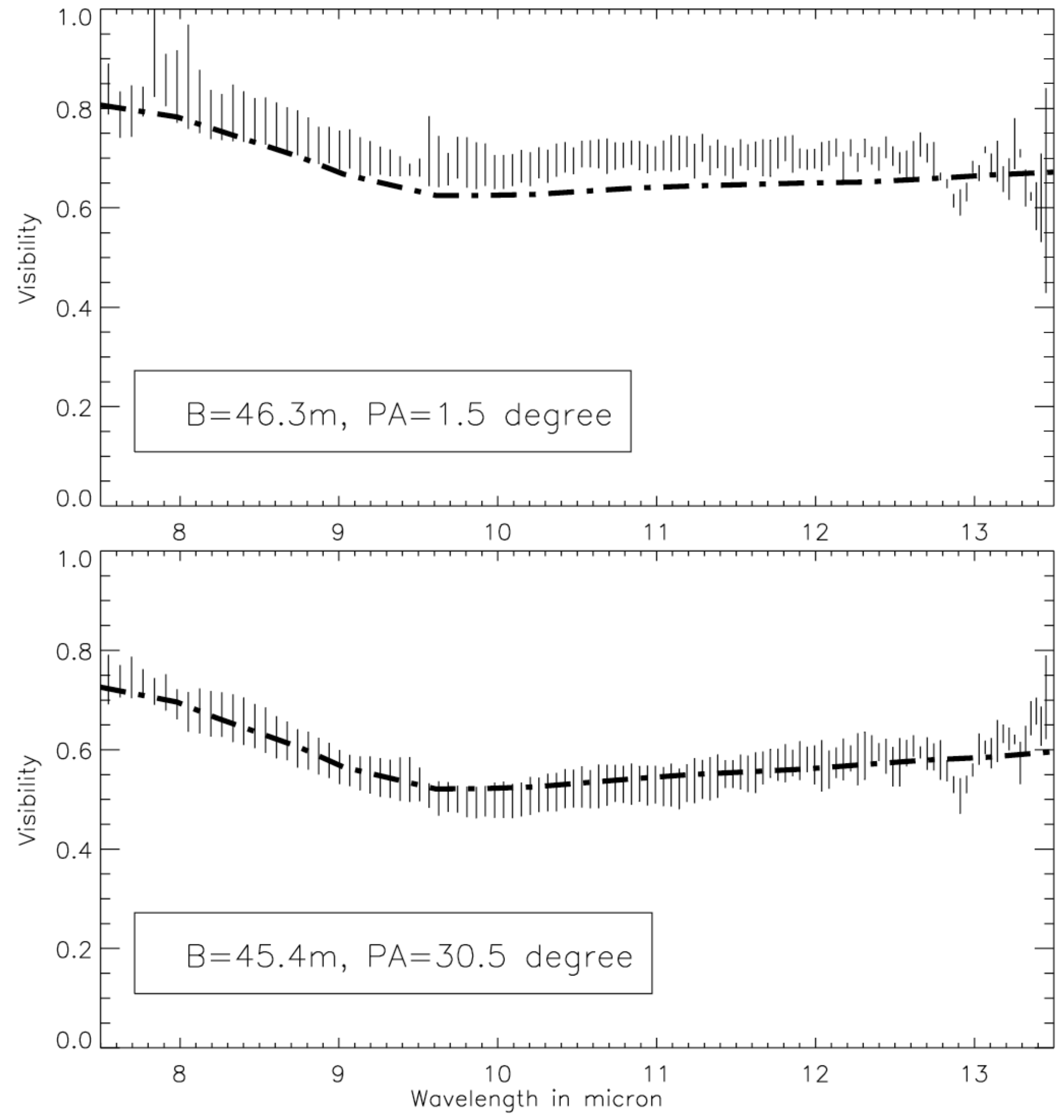
- 70% correlation between any points
- Similar effect as noise on normalization (incoherent flux, calibrator)
- less dispersion of residuals
- bias may appear

Beware : only one realization !

Examples on real data / 1

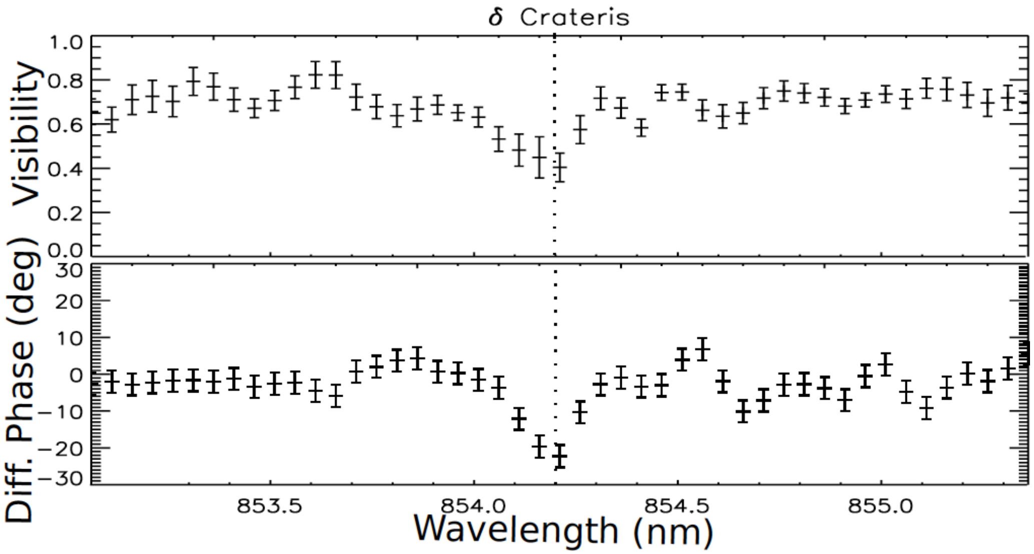


GI2T, Vakili et al. 1997



MIDI, Chesneau et al. 2007

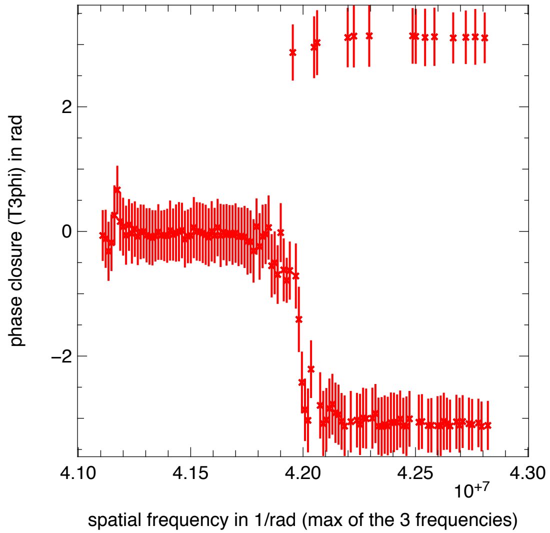
Examples on real data / 2

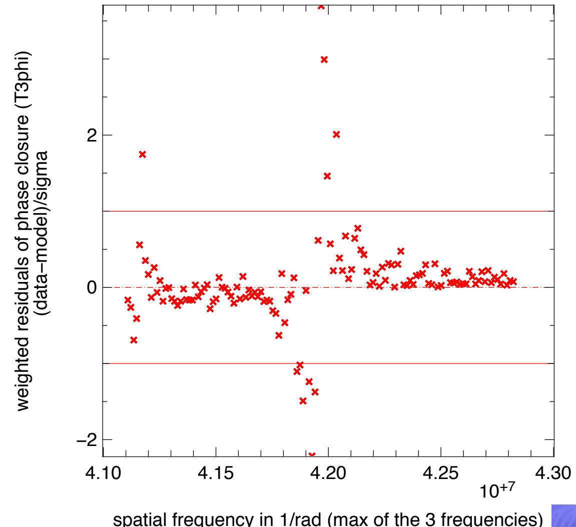


VEGA/CHARA, Bério et al 2011

- => Need for a study of statistics of real data.
- accurate error bars is necessary for reliable interferometry

AMBER data,
T3phi on Sirius
Duvert 2013





- On-going work done at Lagrange (Nice)
  - Antony Schutz, Martin Vannier, André Ferrari, David Mary, Florentin Millour, Romain Petrov
- Rationale
  - Analyze assumptions on visibility statistics (distribution, correlation)
  - Need to adapt data processing ?
  - Need to improve data reduction ?
  - Other
    - effective connection between experts from interferometry and signal processing
    - understand data reduction necessities and practices
- Data
  - Currently squared (differential) visibilities from AMBER
  - Work in progress for data from PIONIER and VEGA

Statistical analysis / dataset

Unit Telescopes + Medium Resolution + No Fringe Tracker

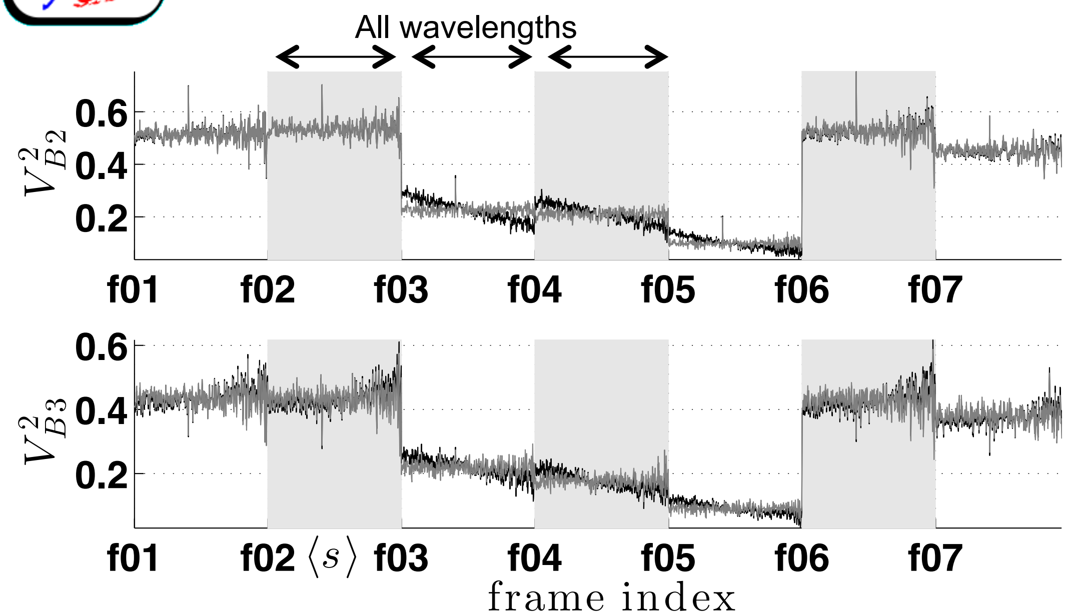
Auxiliary Telescopes + Medium Resolution + Fringe Tracker

Auxiliary Telescopes + Low Resolution + Fringe Tracker

Dataset name	UT-MR-NoFT	AT-MR-FT	AT-LR-FT
Date	2011/01/16	2012/02/22	2012/10/09
Nb of files	**2**	**40**	**37**
Time Span	0.2 h	5.7 h	3.5 h
Nb frames per file	**970**	**37**	**970**
Nb λ per frame	**400**	**400**	**13**
integration time (s)	**0.19**	**2.0**	**0.026**
Spectral Res.	1500	1500	35
Spectral Band	K	K	K
Fringe Tracker	OFF	ON	ON
Seeing (arcsec) (min., avg., max.)	0.9, 1.0, 1.0	0.6, 0.9, 1.4	0.5,0.6,0.8,

- AMBER data
- calibrators
- with and without fringe tracker
- UT / AT
- various time exposure
- # frames / # spectral channels

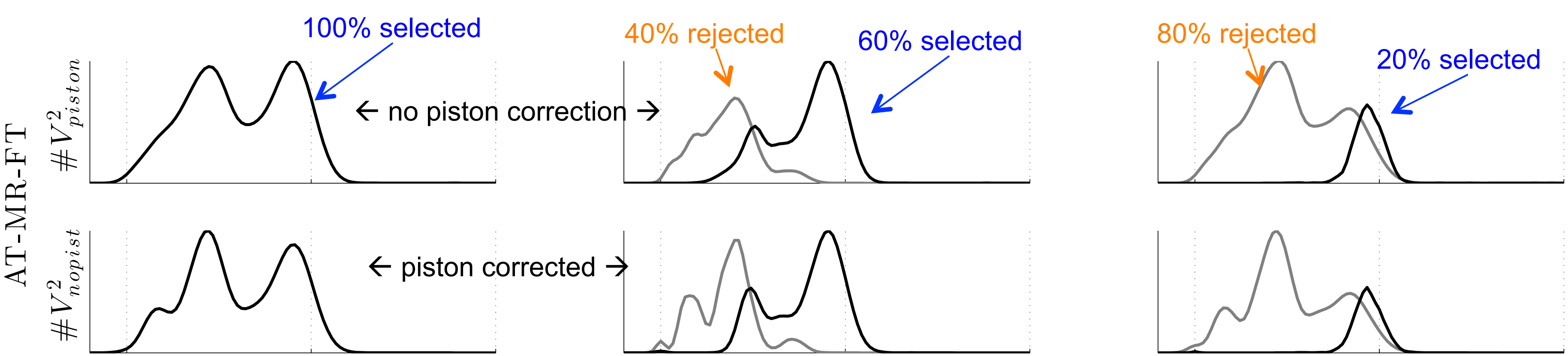
Statistical analysis / squared visibility



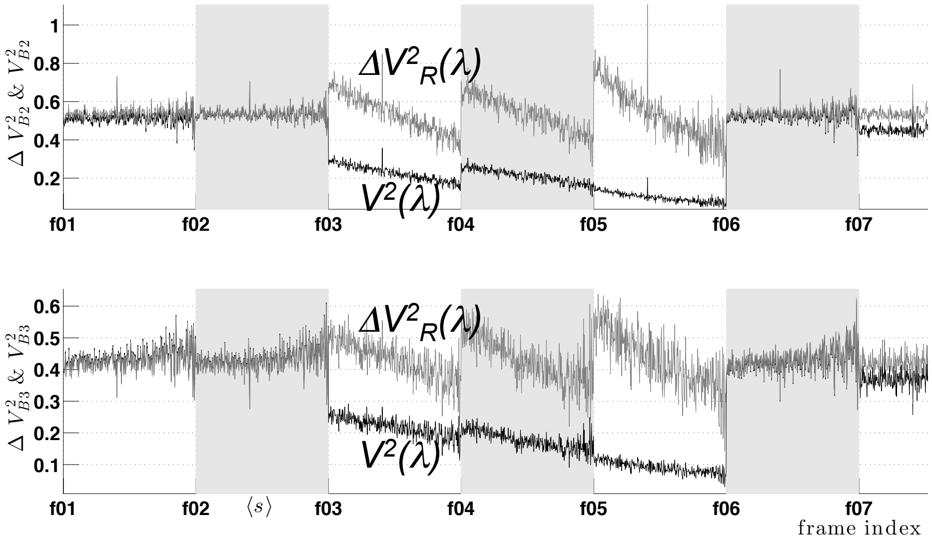
- Sudden visibility losses
- with (gray) / without (black) piston correction
- visibility loss and piston introduce correlations *between wavelengths*

Example of histograms for AT-MR-FT dataset

- Selection based on fringe SNR
- Multi-modal distributions can appear



Statistical analysis / differential visibility



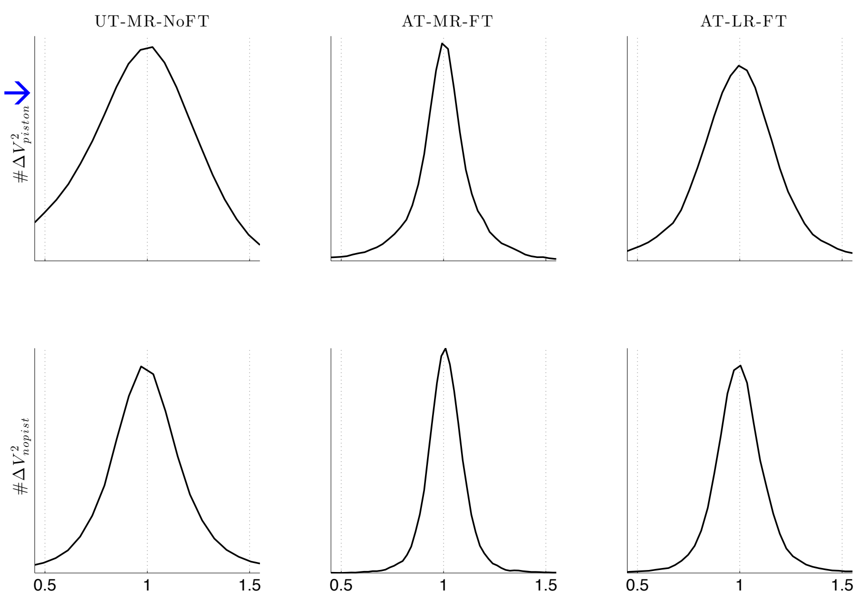
AMBER definition :

$$\Delta V^2(\lambda_i) = \frac{V^2(\lambda_i)}{V^2_{\mathrm{Ref}}(\lambda_i)}, \text{ with } V^2_{\mathrm{Ref}}(\lambda_i) = \frac{1}{n_\lambda - 1} \sum_{i=1, j\neq i}^{n_\lambda} V^2(\lambda_j)$$

- ΔV^2 corrected from visibility loss
- Rescaling: $\Delta V^2 \rightarrow \Delta V^2_R$
  - rescaled to the mean value of visibility with 20% selection
- Variance increased

No piston correction →

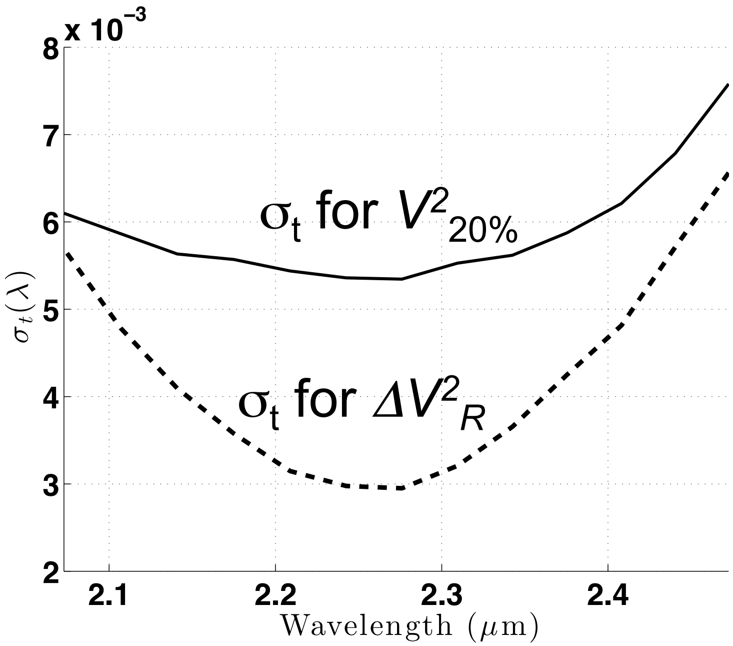
Piston corrected →



Histogram of ΔV^2 for the 3 datasets

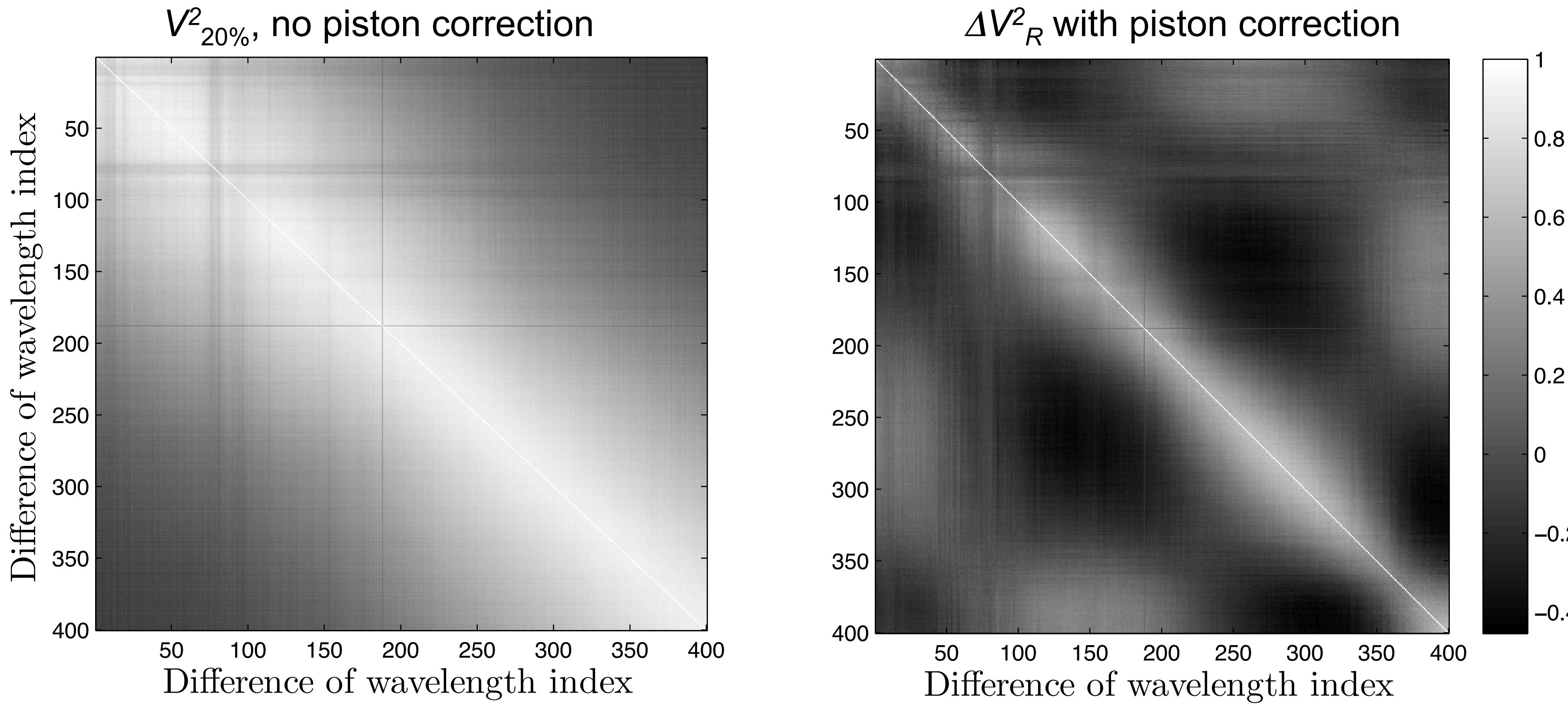
- 100% selected
- Unimodal
- Mean close to one

Statistical analysis / comparison of standard deviations



- Example of standard deviation per exposure file for visibility and differential visibility
  - results averaged over all files and bases
  - divided by the squared root of the number of frames

Statistical analysis / comparison of correlations



- Examples of matrices of the temporal correlation coefficients between spectral channels for dataset UT-MT-NoFT, averaged over observation files and over the three baselines.

Statistical analysis / probability distribution of data ?

- Compatibility of empirical distribution of data with standard distributions ?
  - χ^2 goodness-to-fit test with 5% probability of false alarm
- Tested standard distributions
  - Normal (generally assumed)
  - Cauchy (ratio of 2 normal random variable)
  - Student (able to fit both Normal and Cauchy and between)
  - Log Normal (used in previous work)

Rejection rate in % for AT-MR-FT dataset

Normal **Student** Log Normal Cauchy

	AT-MR-FT/piston	Spectral Analysis				Temporal analysis			
		$\mathcal{N}$	t	Log $\mathcal{N}$	$\mathcal{C}$	$\mathcal{N}$	t	Log $\mathcal{N}$	$\mathcal{C}$
100% selected →	V^2_{100}	59	**29**	53	100	34	**11**	74	12
20% selected →	V^2_{20}	59	**22**	49	100	1	**0**	1	4
differential Vis. rescaled→	ΔV^2_{R}	59	**19**	53	99	22	**1**	29	5

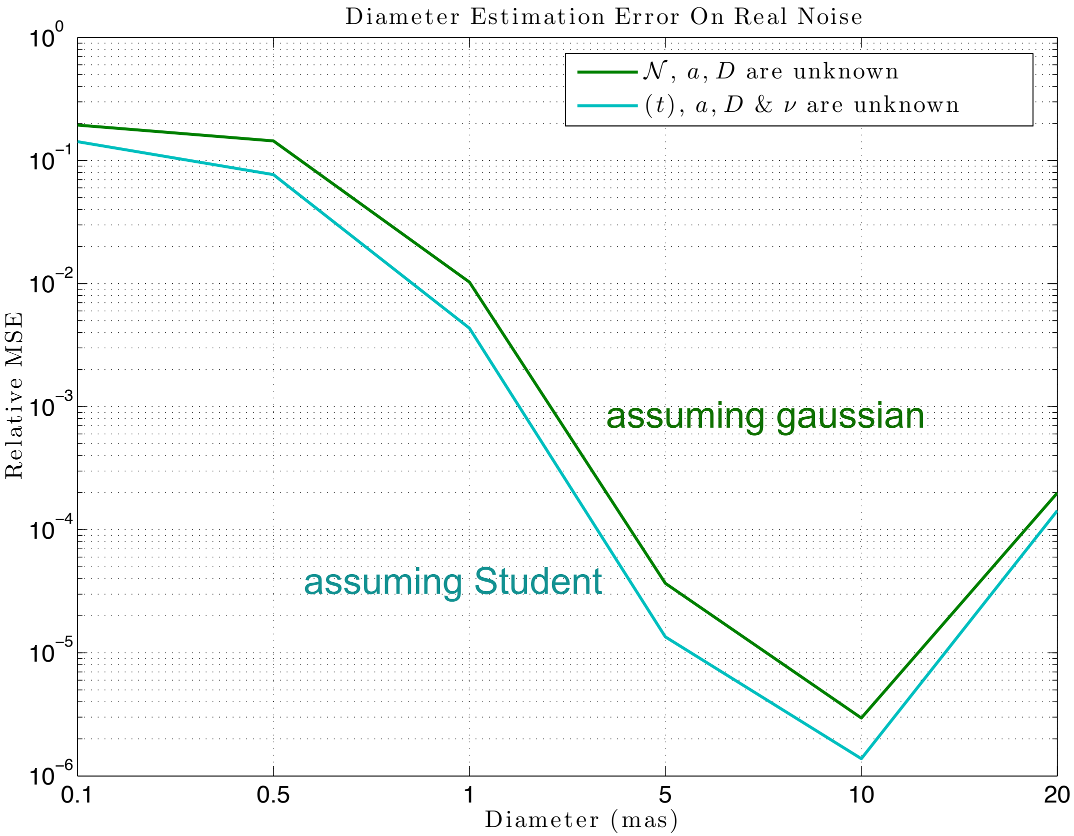
AT-MR-FT/no piston	$\mathcal{N}$	t	Log $\mathcal{N}$	$\mathcal{C}$	$\mathcal{N}$	t	Log $\mathcal{N}$	$\mathcal{C}$
V^2_{100}	88	**12**	51	99	34	**11**	70	18
V^2_{20}	83	**5**	52	100	11	**0**	12	10
ΔV^2_{R}	88	**12**	51	99	15	**1**	16	16

Statistical analysis / summary

- Student distribution is more suitable than Normal, Cauchy and Log Normal distributions
  - True for visibility and differential visibility

- Visibility versus differential visibility
  - Visibility:
    - Visibility loss and piston effect depend on the selection threshold (empirical tuned)
    - Can be multi-modal (depends on selection threshold)
    - Correlated (depends on selection threshold)
  - Rescaled differential visibility: alternative estimator of visibility ?
    - No selection (nothing to tune)
    - Larger number of frames (here, x 5), since no selection
    - Lower standard deviation, since larger number of frames
    - Less correlated
    - Best distribution fitting score (Student)
- Next
  - data from VEGA, PIONIER
  - differential phase, closure phase

Example of possible effect for model fitting



- Simulated data for uniform disks (6 different diameters) with ***real noise***.
  - real noise is closer to Student law
- Fit of this data assuming either gaussian or Student statistics
  - better accuracy with Student statistics : relative MSE twice smaller.

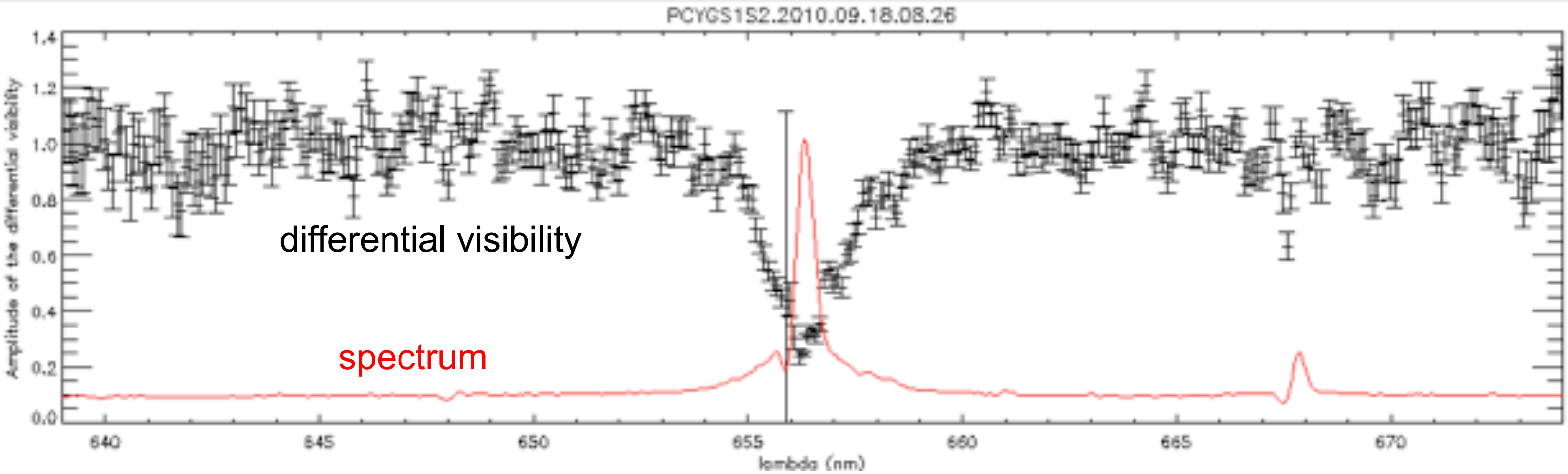
Data more suitable to image reconstruction ?

- Current polychromatic data are not convenient
  - closure phase and powerspectrum not linear (lower SNR, loss of information)
  - differential phase very helpful : how to use it in a global approach ?
- Image reconstruction = inverse problem
  - Compute (synthetic) data *knowing* the object (direct model) and priors
  - Compare with real data
  - Find the model of the object that gives the best match
- Attempt: allow other unknowns that can help in the inverse problem
  - We want to keep linearity => complex visibilities
  - We want to accumulate frames => fringe tracking in post-processing (multiple wavelengths)
  - We don't want to spoil the statistics => account for statistics (and keep all frames)
  - Reject optimally (i.e. robustly) the information we don't need.
  - Keep remaining unknowns that cannot be rejected => ~"self-calibration"
- → Ferréol Soulez's talk

3-D (x,y,λ) model fitting

One example of polychromatic "difficulties"

Need of a model of the spectrograph

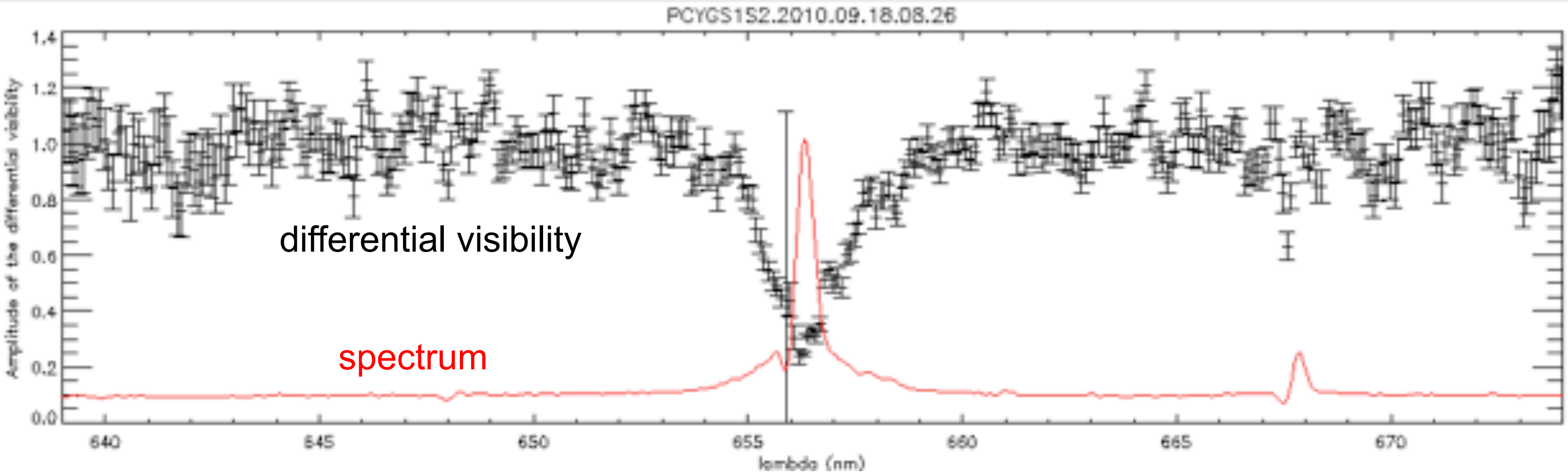


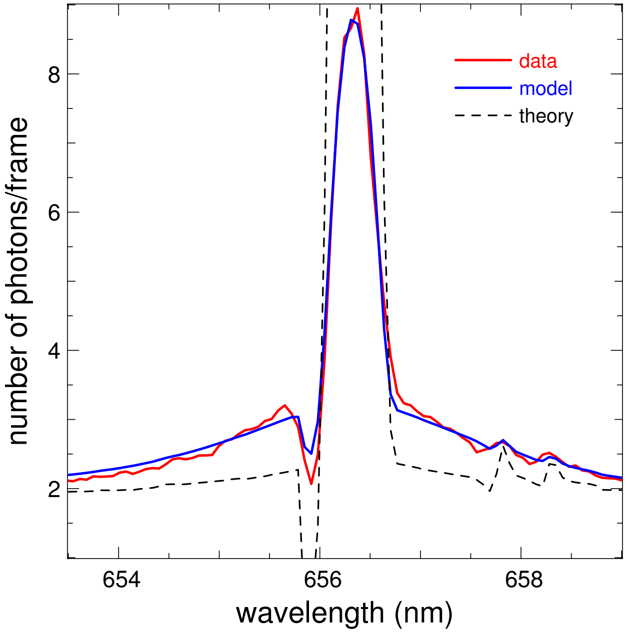
- Data on P Cygni, from VEGA
  - Width of differential visibility difficult to explain in H_α line
  - Spectrum is affected by detector artifacts (saturation, …)

- P Cyg spectrum is quite well modeled by radiative transfer codes

  ⇒ make use of the theoretical spectrum for modeling the object "as seen by VEGA"

Need of a model of the spectrograph





- Spectrograph model includes
  - PSF
  - wide hallo
  - detector saturation, non-uniformity
- To be included in LITpro at the end of the project

Polychromatic image reconstruction

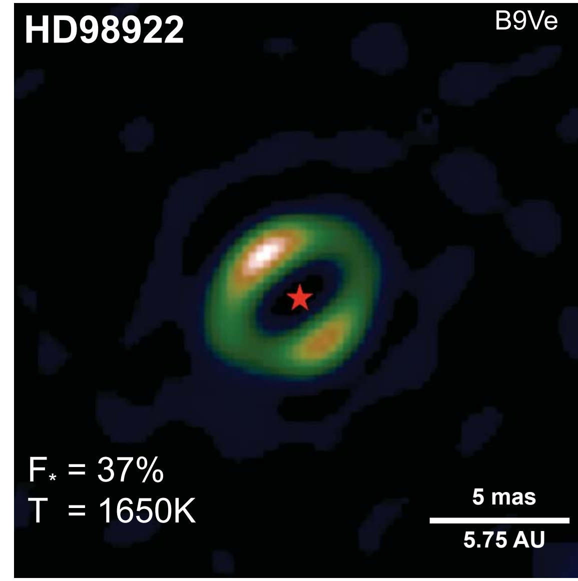
- Specific approach → Jacques Kluska's talk
- General approach → Ferréol Soulez's talk
- Other work done in Nice using sparse representations

Polychromatic image reconstruction : specific approach

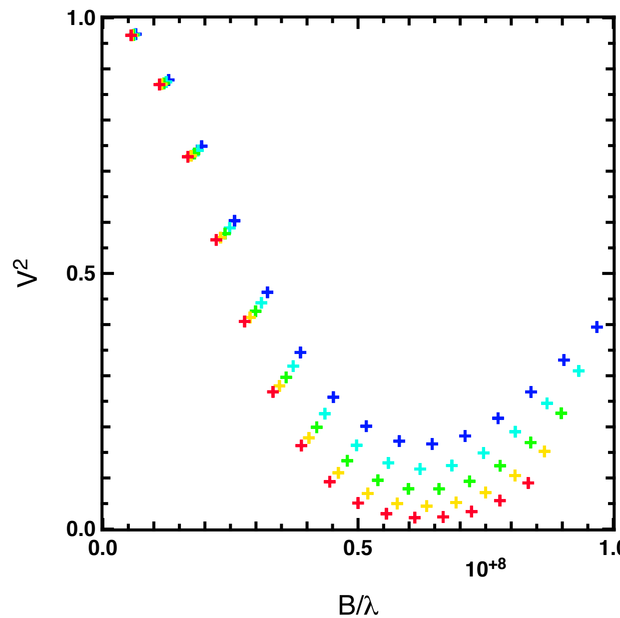
- The image is the addition of components with weights that depend on the wavelength:

$$i(\boldsymbol{x}, \lambda) = \sum_{k=1}^{n} w_k(\lambda)\, i_k(\boldsymbol{x})$$

  - $i_k(\boldsymbol{x})$ obtained with usual image reconstruction or model fitting



- → Jacques Kluska's talk

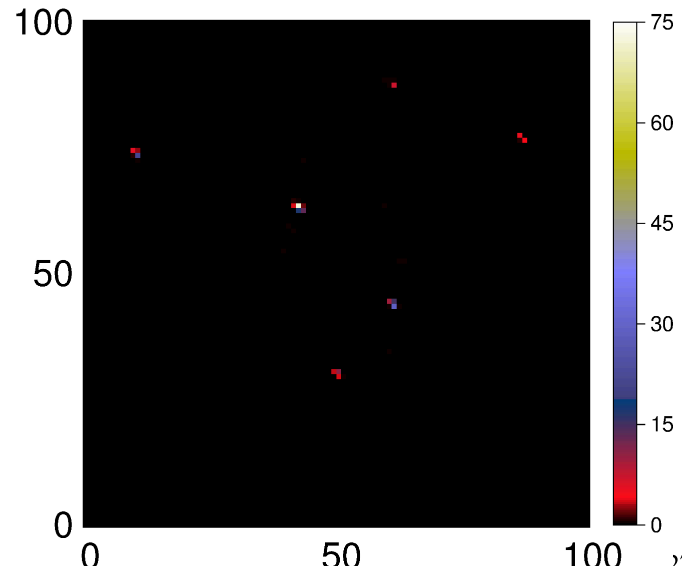


Polychromatic image reconstruction : general approach

- Make use of the strong relationship between the wavelengths
  - Reconstruction of a 3D cube (x,y,λ)
  - The object shape depends continuously on the wavelength (with exceptions)

- Needs
  - A 3D cube ! => need to be fast enough (i.e. converge easily)
  - Account for the spectrum complexity (lines, dynamics, …)

- → Ferréol Soulez's talk



Conclusions

- POLCA = Processing of Polychromatic Interferometric Data for Astrophysics
  - Focus on real data (AMBER, MIDI, VEGA, PIONIER)
    - simulated data if not possible (GRAVITY)
  - Improve model of data statistics
  - Formalize measurements by current interferometers
  - 3-D (x,y,λ) image reconstruction
  - 3-D (x,y,λ) model fitting (and coupling with image reconstruction)

- Evolutions of the project
  - initially starting from reduced data, but now look for possible more appropriate estimate using wavelength dependency (phase difference, …)

- Improving the performances of interferometers ?
  - Do not forget to improve signal processing