General *theory* of data reduction

EuroSummer School

Observation and data reduction with the Very Large Telescope Interferometer

Goutelas, France
June 4-16, 2006

Guy Perrin
Observatoire de Paris
12 June 2006

Outline

1. The different types of beamcombiners
2. What’s the issue ?
3. The noises
4. Single-mode interferometers
5. Estimators
6. An example of *Fourier* estimator
7. Estimating error bars and calibrating visibilities

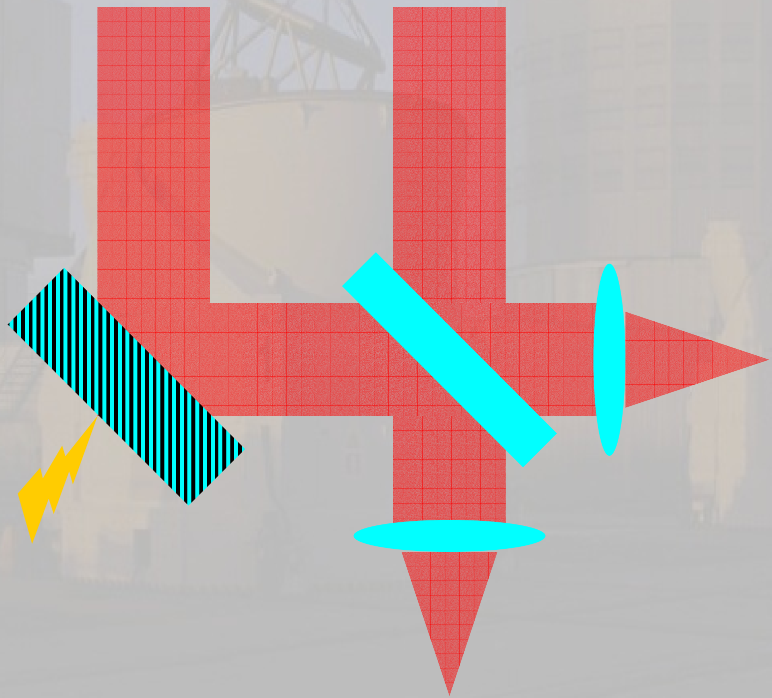
Outline

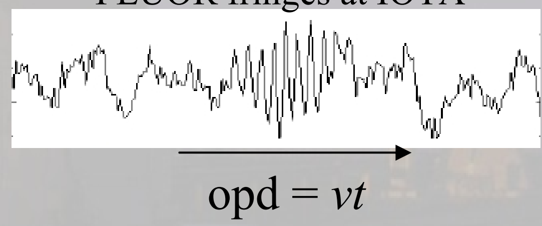
✓ The different types of beamcombiners

2. What's the issue ?
3. The noises
4. Single-mode interferometers
5. Estimators
6. An example of *Fourier* estimator
7. Estimating error bars and calibrating visibilities

The co-axial or ‘‘Michelson-type’’ beamcombiner

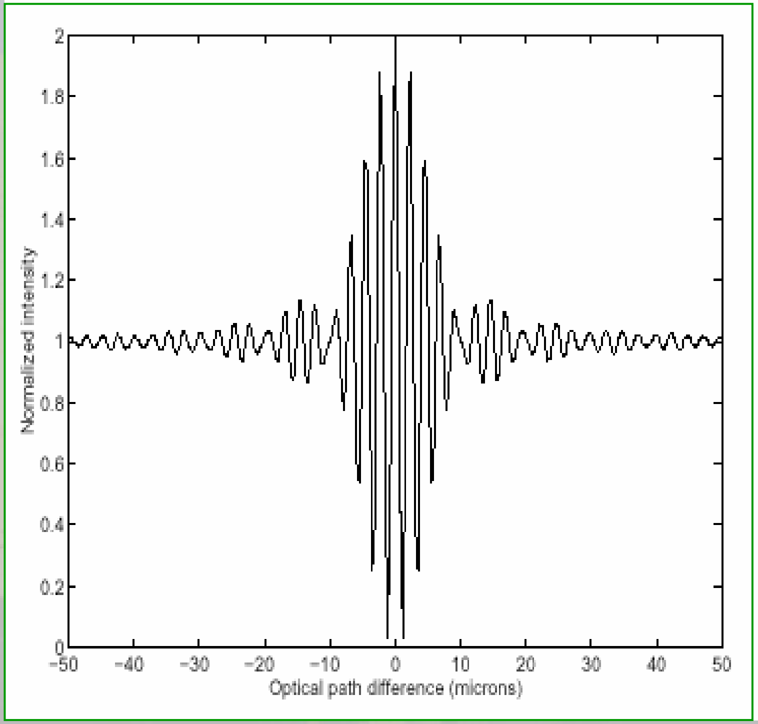
Temporal coding beam combiner:

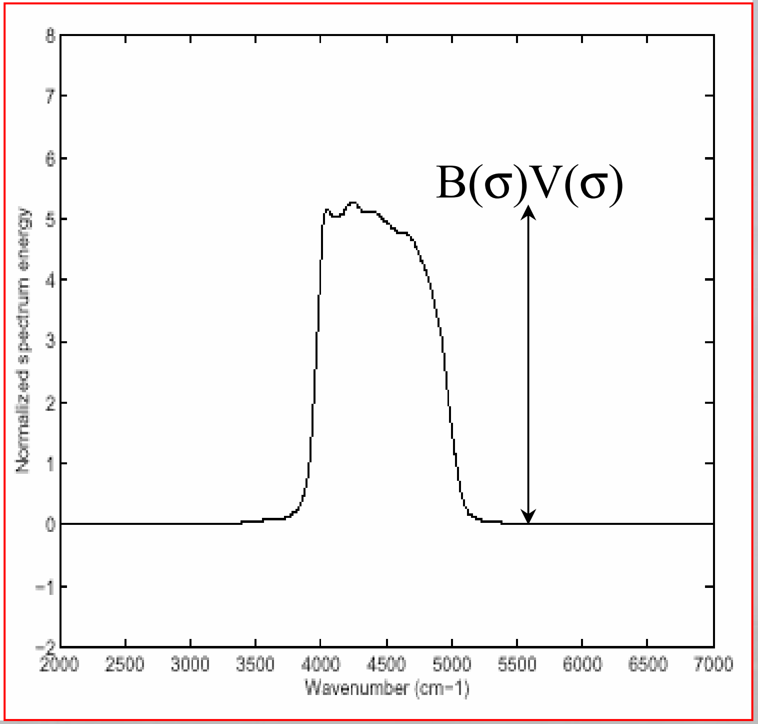




The opd is temporally modulated : $I(t) = I_0\left(1 \pm V\cos(2\pi\sigma v t)\right)$

Fourier transform





Scan

Intensity vs. x opd in μm

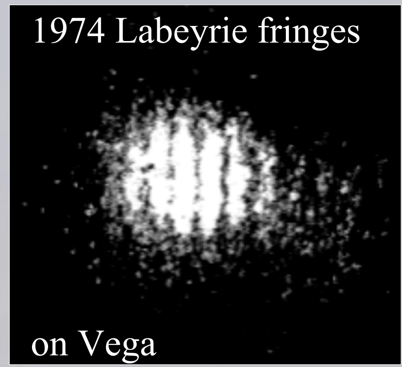
Spectrum B(σ)

$\sigma = 1/\lambda$ = wavenumber (cm^{-1})

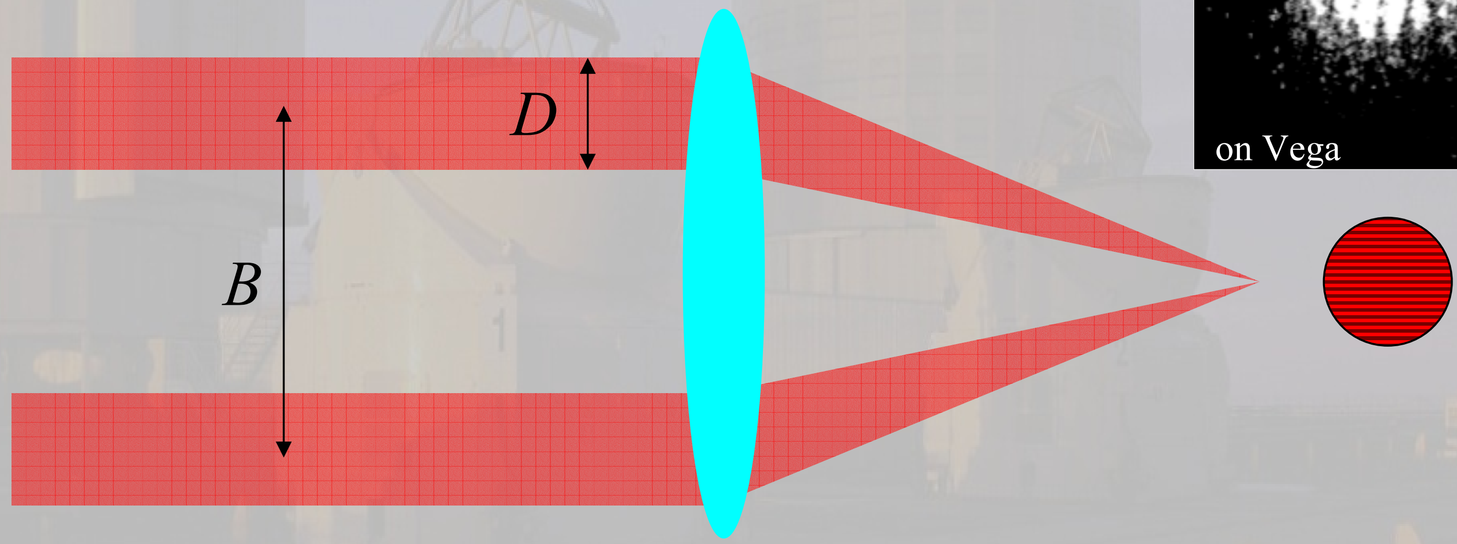
K band : 4000 - 5000 cm^{-1}

x and σ are conjugate variables through the Fourier Transform

The multi-axial or ‘‘Fizeau-type’’ beamcombiner



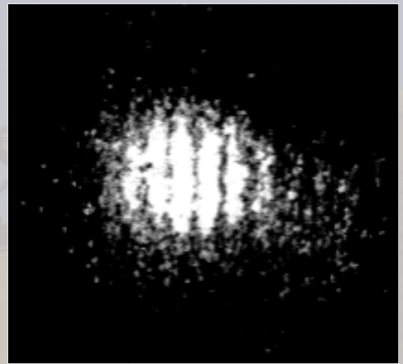
Spatial coding beam combiner:



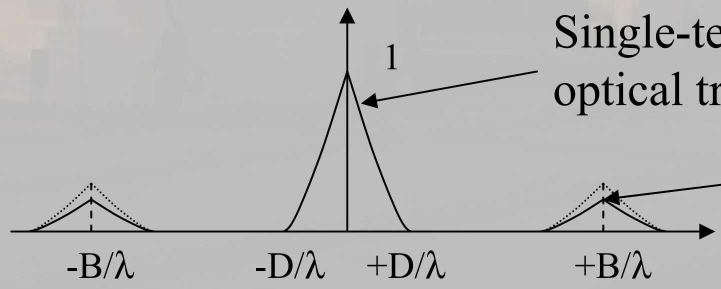
The opd is spatially modulated:

$$I(\theta) = \left(\frac{2J_1\left(\pi \frac{D}{\lambda}\theta\right)}{\pi \frac{D}{\lambda}\theta} \right)^2 \times \left(1 + V\cos\left(2\pi \frac{B}{\lambda}\theta\right)\right)$$

The multi-axial or ‘‘Fizeau-type’’ beamcombiner



Fourier transform



B/λ and θ are conjugate variables through the Fourier Transform

Type of beam combiner considered here

I will use the formalism of the co-axial beam combiner in the following.

Both are actually fully equivalent to within some details in the equations.

Outline

✓ The different types of beamcombiners

✓ What's the issue ?

3. The noises

4. Single-mode interferometers

5. Estimators

6. An example of *Fourier* estimator

2. Estimating error bars and calibrating visibilities

What’s the issue ?

A priori no issue.

We just need to measure a modulation.

Chris’s first lecture:

$$P(s_0,B,\delta) \;=\; I_{total}\{1+\mathrm{Re}\;[V\exp[-ik\delta]]\}$$

The visibility can easily be measured by sampling δ at 0 and $\lambda/4$:

$$\begin{cases} \mathrm{Re}[V]=\dfrac{P(s_0,B,0)}{I_{total}}-1 \\ \mathrm{Im}[V]=\dfrac{P(s_0,B,\lambda/4)}{I_{total}}-1 \end{cases}$$

What's the issue ?

Unfortunately atmospheric piston makes a very bad joke:

$$P(s_0,B,\delta) \;=\; I_{total}\left\{1+\mathrm{Re}\;\left[V\exp[-ik\delta-i\varphi_p]\right]\right\}$$

$$\begin{cases} \dfrac{P(s_0,B,0)}{I_{total}}-1=\mathrm{Re}\left[V\exp\left[-i\varphi_p\right]\right] \\ \dfrac{P(s_0,B,\lambda/4)}{I_{total}}-1=\mathrm{Im}\left[V\exp\left[-i\varphi_p\right]\right] \end{cases}$$

Such a linear estimator of the visibility unfortunately does not work (yet ?) as the phase is corrupted by piston

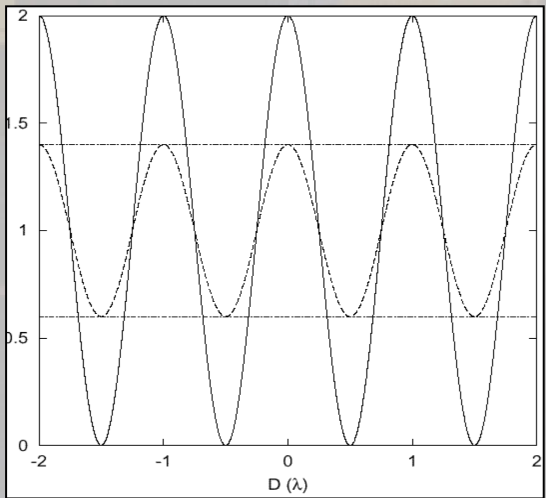
Non-linear estimators are needed to estimate the visibility modulus and phase separately

Classical visibility estimator

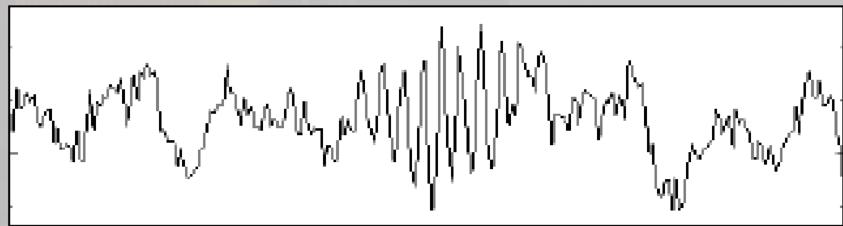
The Michelson estimator:

$$|V| = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}$$

Works well on:



But not so well on real data like:



Why ?

Estimators robust to noise are necessary

Outline

✓ The different types of beamcombiners

✓ What's the issue ?

✓ The noises

4. Single-mode interferometers

5. Estimators

6. An example of *Fourier* estimator

7. Estimating error bars and calibrating visibilities

From an ideal to a real noisy interferogram

1. Additive noise

Ideal interferogram:

$$P(s_0,B,\delta) \;=\; I_{total}\left\{1+\mathrm{Re}\,\left[V\exp[-ik\delta]\right]\right\}$$

With photon and detector noise:

$$\begin{cases} P_n(s_0,B,\delta) \;=\; P(s_0,B,\delta)+n_{\det}+n_{ph} \\ \mathrm{rms}(n_{ph})=\sqrt{P(s_0,B,\delta)} \end{cases}$$

With instrument and sky background noise:

$$\begin{cases} P_{n,b}(s_0,B,\delta) \;=\; P(s_0,B,\delta)+n_{\det}+n_{ph}+Back(t)+n_{Back(t)} \\ \mathrm{rms}(n_{Back(t)})=\sqrt{Back(t)} \end{cases}$$

From an ideal to a real noisy interferogram

1. Additive noise

$$P_{n,b}(s_0,B,\delta) = P(s_0,B,\delta) + n_{\det} + n_{ph} + Back(t) + n_{Back(t)}$$

Pure random noise
Only averages down to zero

Removable with the chopping technique (residuals will remain)
See L9 by O. Chesneau

From an ideal to a real noisy interferogram

2. Multiplicative noise

Ideal interferogram:

$$P(s_0,B,\delta) \;=\; I_{total}\{1+\mathrm{Re}\;[V\exp[-ik\delta]]\}$$

Interferogram with unbalanced beams:

$$P(s_0,B,\delta) \;=\; \mathrm{P_A}+\mathrm{P_B}+2\sqrt{\mathrm{P_A P_B}}\times\mathrm{Re}\;[V\exp[-ik\delta]]$$

Fringe contrast (phase unchanged):

$$C=\frac{2\sqrt{P_A P_B}}{P_A+P_B}\times|V|$$

$$P_A=2P_B \Rightarrow C=0.94\times V$$

From an ideal to a real noisy interferogram

2. Multiplicative noise

Interferogram with turbulence:

$$P(s_0,B,\delta) \;=\; \mathrm{P_A}+\mathrm{P_B}+2\sqrt{\mathrm{P_A P_B}}\times e^{-\sigma_\varphi^2}\times \mathrm{Re}\left[V\exp[-ik\delta-i\varphi_p(t)]\right]$$

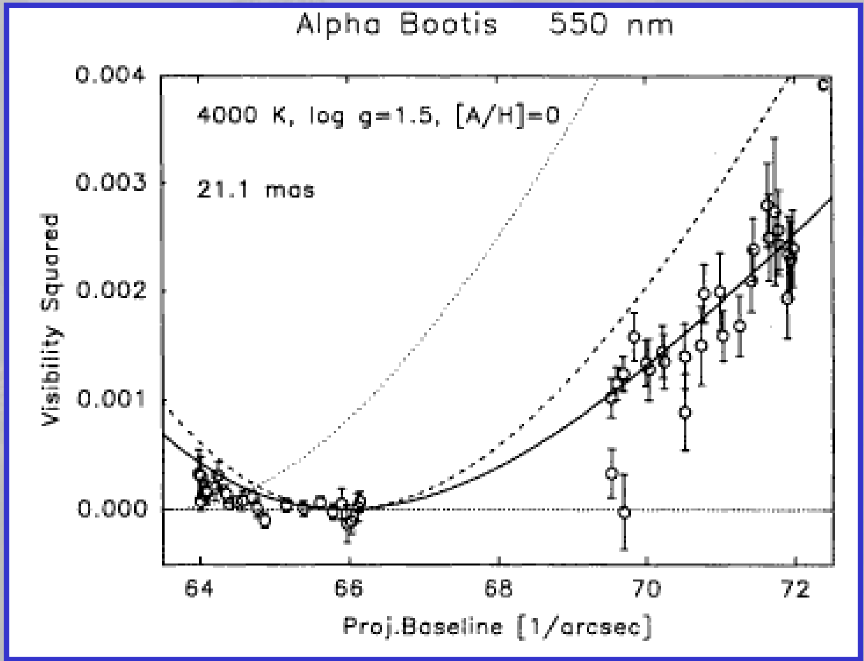
$e^{-\sigma_\varphi^2}$ is the coherent energy, σ_φ^2 is the phase variance over the pupil

Instantaneous fringe contrast:

$$C=\frac{2\sqrt{P_A P_B}}{P_A+P_B}\times e^{-\sigma_\varphi^2}\times|V|$$

This is a real catastrophy when the turbulence is not stable which unfortunately is the case in real life

The visibility bias decreases with the visibility



Quirrenbach et al. (1996)

Outline

✓ The different types of beamcombiners

✓ What’s the issue ?

✓ The noises

✓ Single-mode interferometers

5. Estimators

6. An example of *Fourier* estimator

2. Estimating error bars and calibrating visibilities

Different ways to handle the turbulence issue

1. Don’t do anything and assume that turbulence issues do not have an impact on the science case

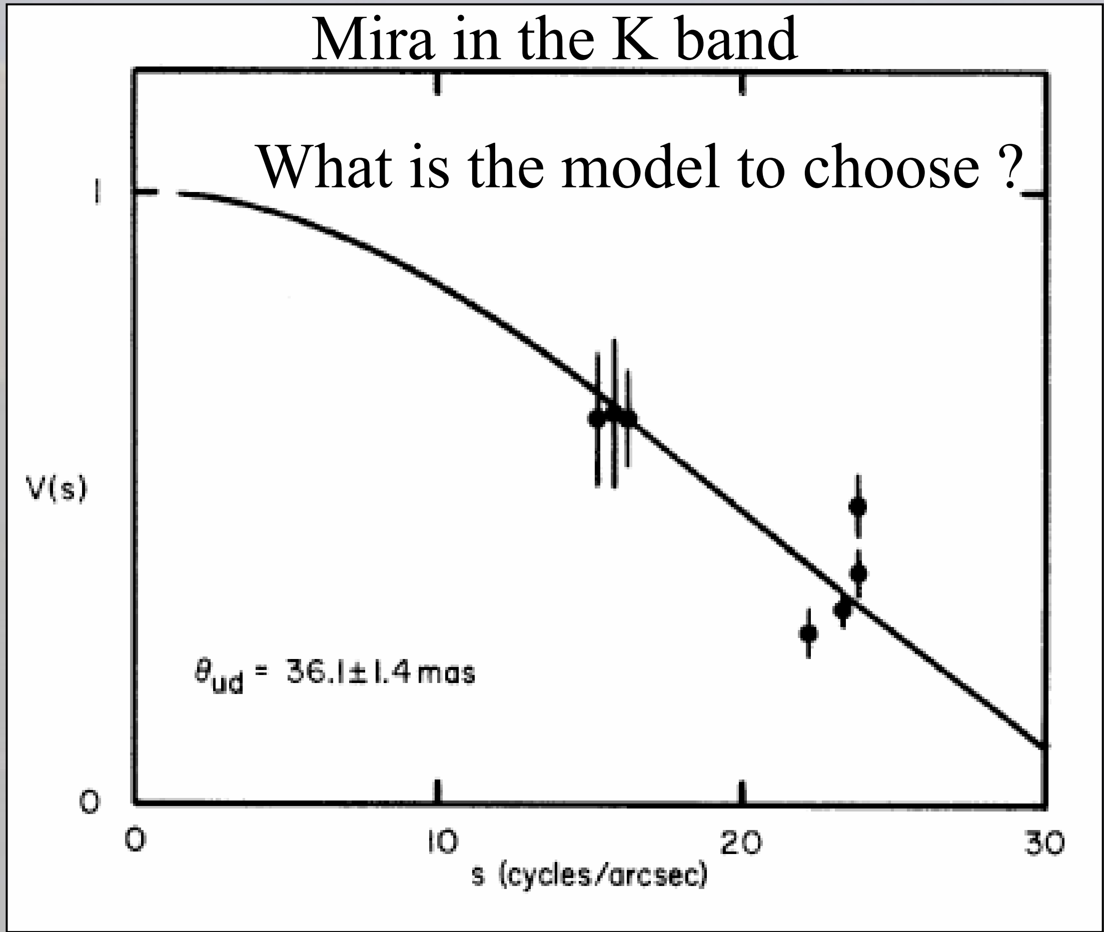
2. Stop the telescope pupils down to less than r_0

3. Use a perfect adaptive optics system and flatten the corrugated wavefront

4. Use a different technique to flatten the wavefront

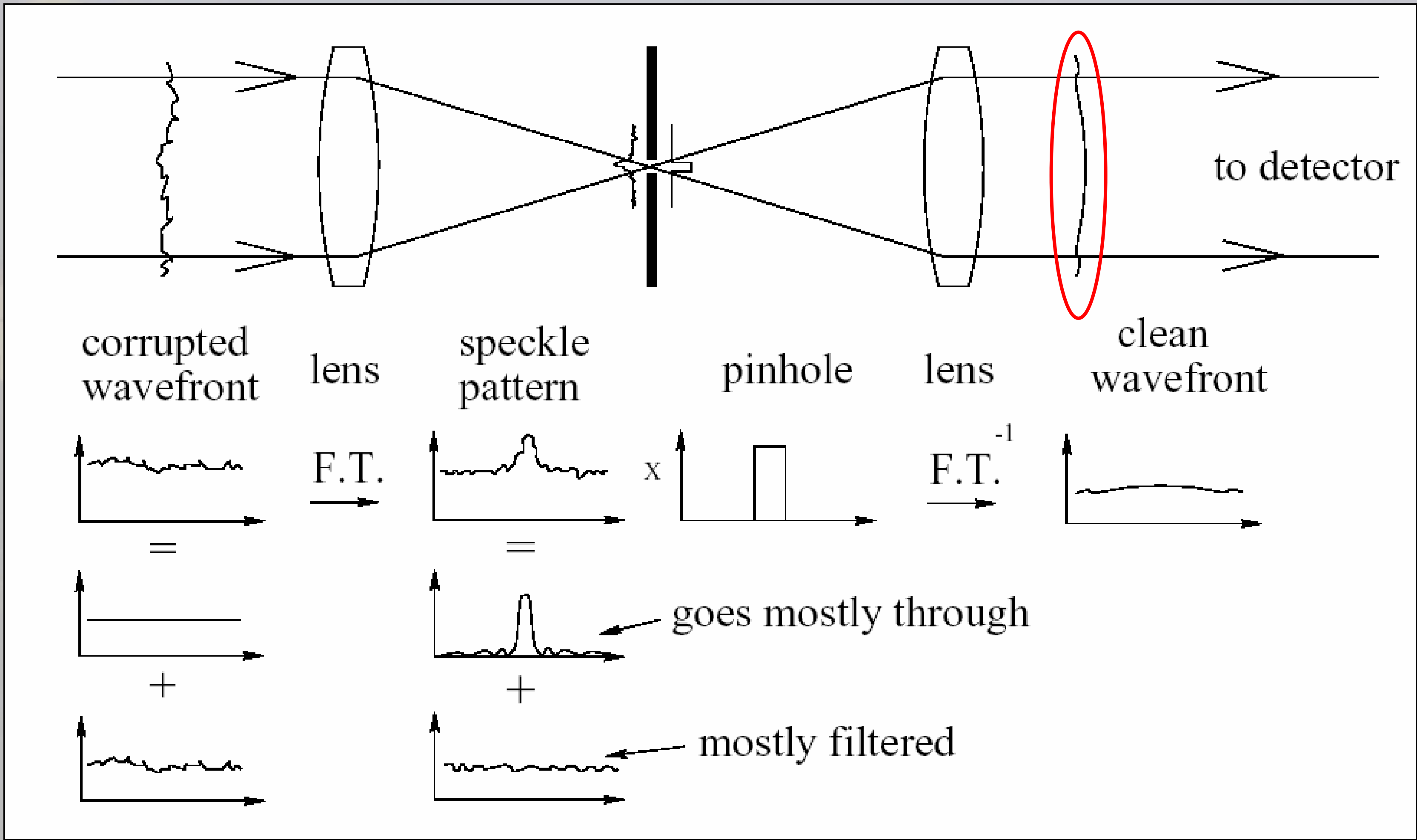
5. Mix of 3 and 4

The difficulty of interferometry without modal filtering

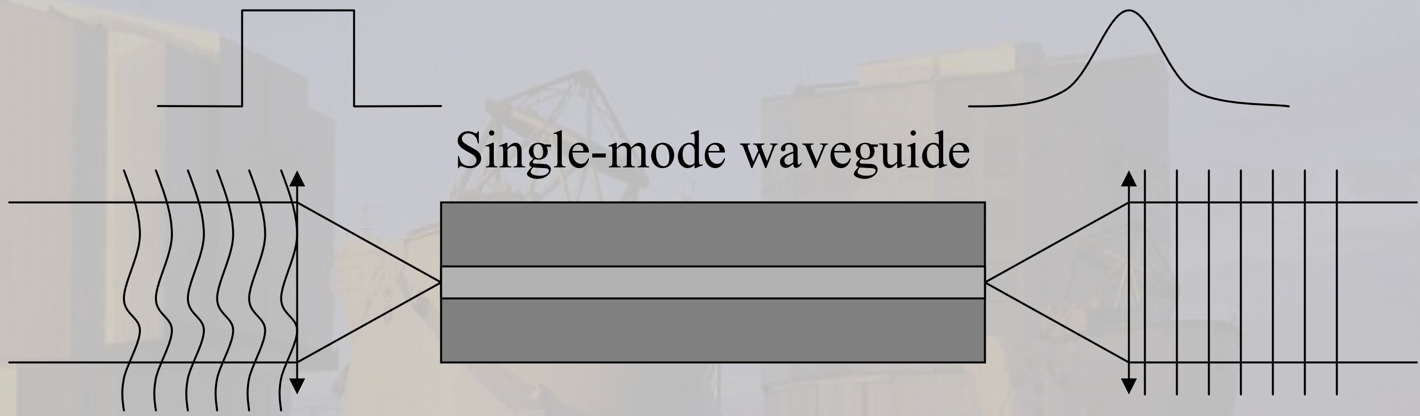


Ridgway et al. (1992)

Way out: spatial filtering



Better way out: *modal* filtering



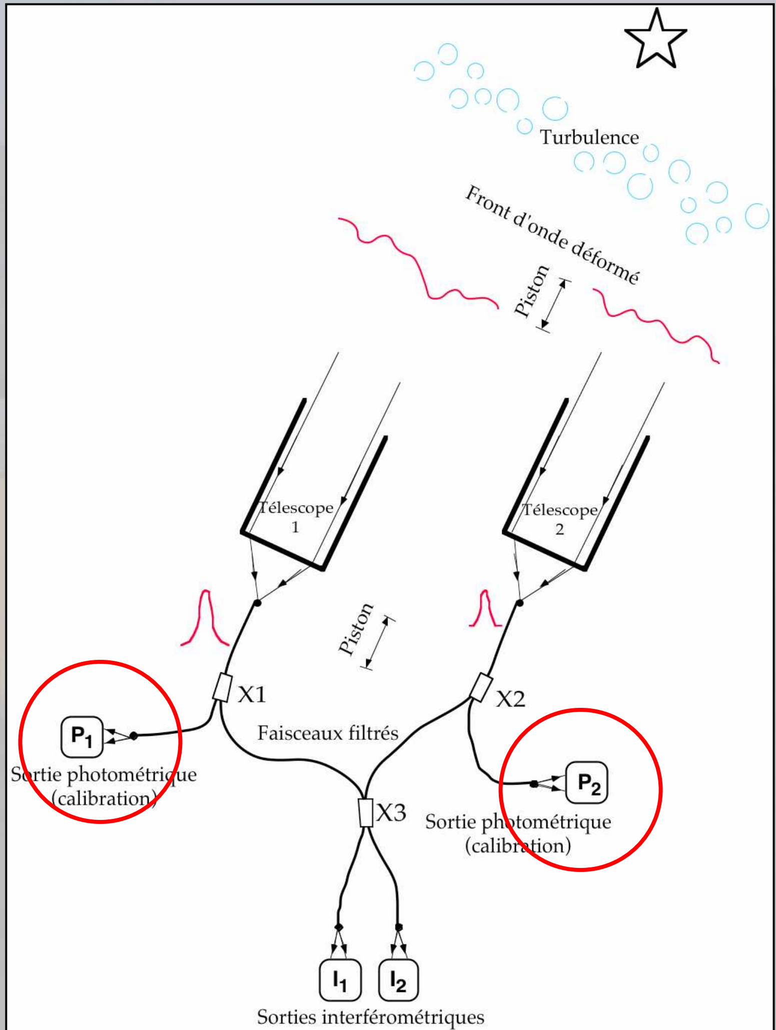
The spatial profile of the exit beam is the profile of the fiber fundamental mode (close to a gaussian).

The phase is flat in the waveguide. The spatial coherence is maximum.

The phase fluctuations of the input beam are traded against intensity fluctuations of the output. But these fluctuations can be measured.

Single-mode beam combiner
FLUOR (Fiber Linked Unit for Optical Recombination)

Precursor of the VINCI instrument on VLTI



Interferometric equation, revisited

$$P(s_0,B,\delta) \;=\; \mathrm{P_A} + \mathrm{P_B} + 2\sqrt{\mathrm{P_A}\mathrm{P_B}} \times e^{-\sigma_\varphi^2(t)} \times \mathrm{Re}\,\left[V \exp[-ik\delta - i\varphi_p(t)]\right]$$

With beams A and B modally filtered:

$$P(s_0,B,\delta) \;=\; \mathrm{P_A}(t) + \mathrm{P_B}(t) + 2\sqrt{\mathrm{P_A}(\mathrm{t})\mathrm{P_B}(t)} \times 1 \times \mathrm{Re}\,\left[V \exp[-ik\delta - i\varphi_p(t)]\right]$$

The maximum fringe contrast is the instantaneous atmospheric contrast:

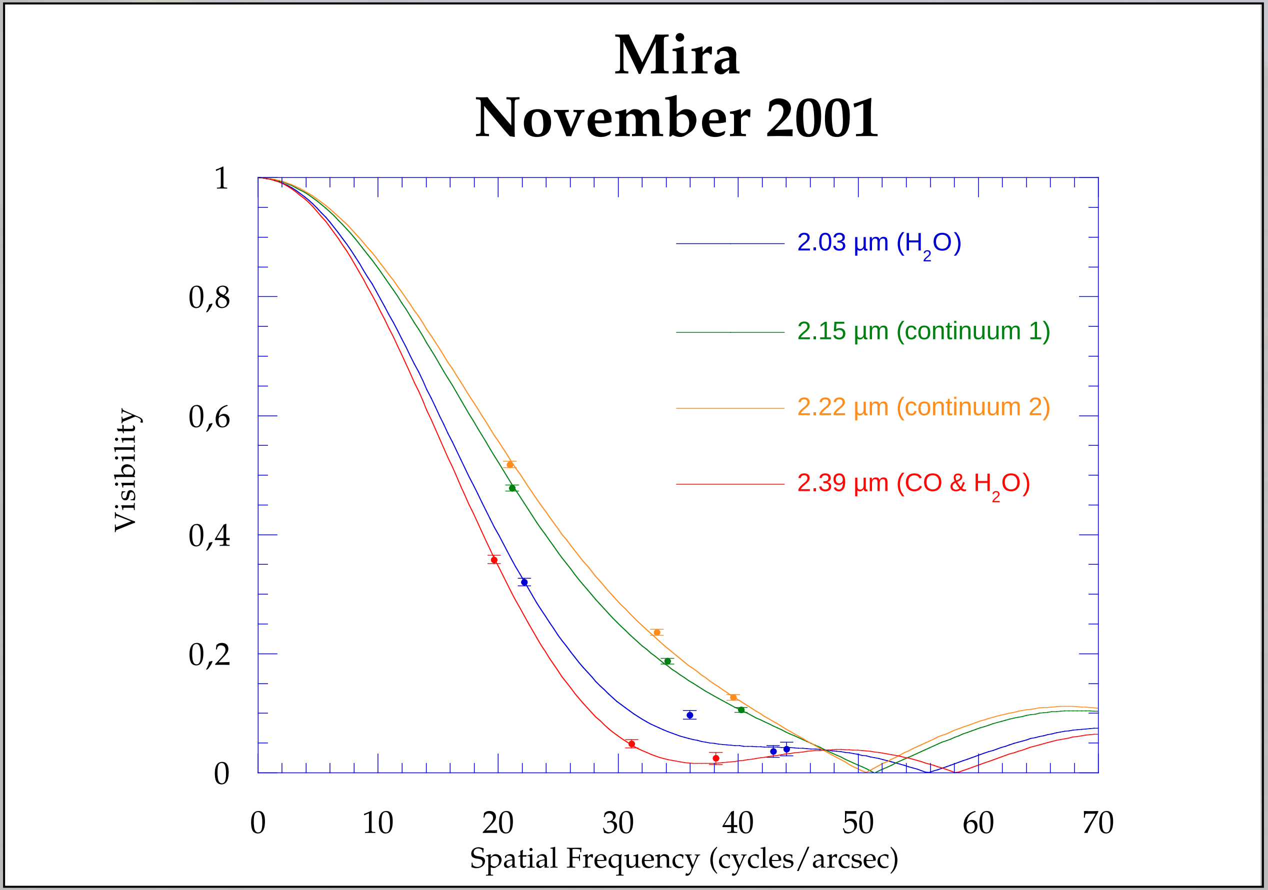
$$C_{atm}(t) = \frac{2\sqrt{P_A(t)P_B(t)}}{P_A(t) + P_B(t)}$$

Hence the visibility after photometric beams have been measured:

$$|V| = C(t)/C(t)_{atm}$$

The coherence loss due to turbulence has been removed.

Same object, with modal filtering

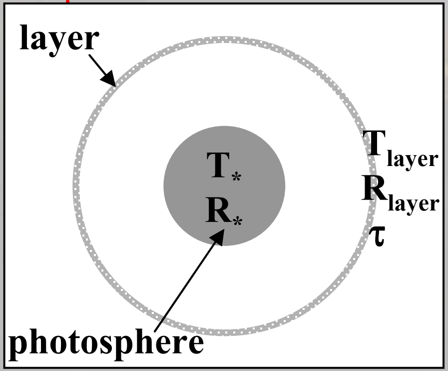


$\tau_{2.03\mu m} = 0.63 \pm 0.21$

$\tau_{2.15\mu m} = 0.19 \pm 0.05$

$\tau_{2.22\mu m} = 0.12 \pm 0.04$

$\tau_{2.39\mu m} = 0.76 \pm 0.50$



$R_* = 12.71 \pm 0.15$ mas **$R_{layer} = 24.95 \pm 0.10$ mas** **Phase: 0.19** **Type: M3-M4**

$T_* = 3600 \pm 67$ K **$T_{layer} = 1961 \pm 17$ K**

(Perrin et al. 2004)

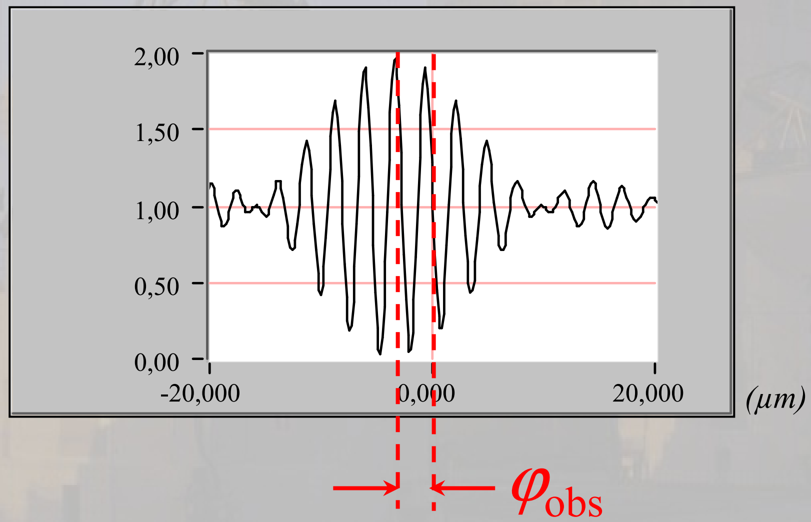
It is time for a 5 minute break

Outline

- ✓ The different types of beamcombiners
- ✓ What's the issue ?
- ✓ The noises
- ✓ Single-mode interferometers
- ✓ Estimators

6. An example of *Fourier* estimator
7. Estimating error bars and calibrating visibilities

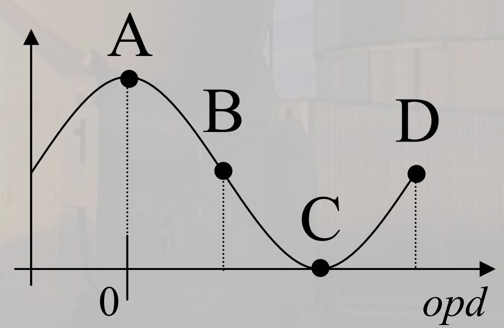
The phase of visibilities



Phase estimator

Measurement modulo 2π : *ABCD* method

The optical path difference is sampled at 4 points : 0, λ/4, λ/2 et 3λ/4

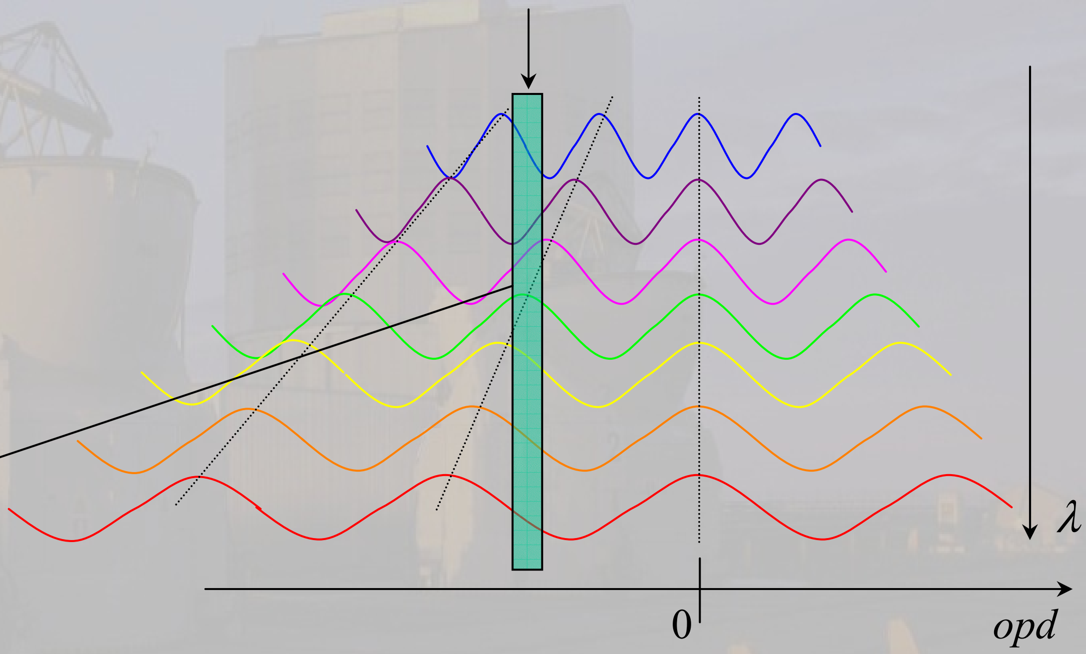


$$\varphi(2\pi) = \arctan\left(\frac{B-D}{A-C}\right)$$

Phase estimator

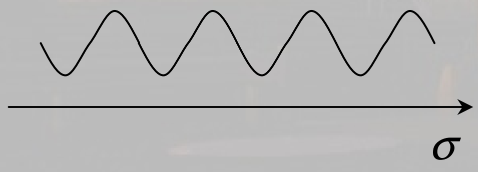
Multi-2π phase measurements.

Example 1: dispersed mode

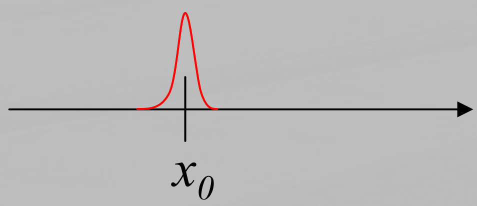


Group delay technique

$$I(\sigma) = \cos(2\pi\sigma x_0)$$



Fourier transform

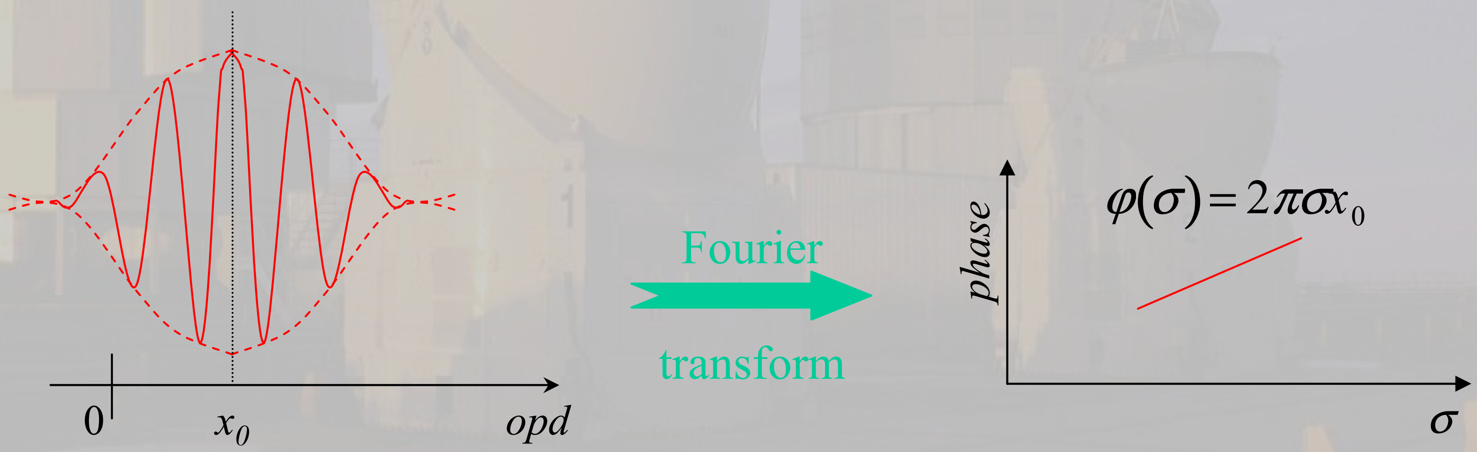


$$\varphi = 2\pi\sigma_0 x_0$$

Phase estimator

Multi-2π phase measurements.

Example 2: co-axial and non-dispersed fringes in wide band



$$\varphi = 2\pi\sigma_0 x_0$$

Visibility estimator

Some vocabulary first

The measured fringe contrast needs to be calibrated to be useful otherwise it is dramatically biased by coherence losses (polarization, dispersion, atmosphere, …)

<u>Fringe contrast before calibration:</u>

μ ***Coherence Factor*** or ***Uncalibrated Visibility***

<u>Fringe contrast after calibration (the observable linked to the object):</u>

V ***Visibility*** or ***Calibrated Visibility***

Visibility estimator

V or V^2 ?

By V, interferometrists indeed mean $|V|$

The data reduction process provides a noisy estimate of $|V|$

The noise applies to the complex quantity V and therefore the estimated fringe contrast for each individual scan is corrupted by noise so that:

$$\tilde{V} = V + n \qquad and \qquad |\tilde{V}| = |V + n|$$

When averaging individual visibility estimates to improve the SNR:

$$\langle |\tilde{V}| \rangle = \langle |V + n| \rangle \neq |V|$$

Visibility estimator

V or V^2 ?

The $|V|$ estimator is therefore *biased* by additive noise.

The bias is all the larger as the visibility is low or as the source is faint

The solution is to average $|V|^2$ instead of $|V|$

$$\left\langle |\tilde{V}|^2 \right\rangle = \left\langle |V+n|^2 \right\rangle = \left\langle |V|^2 + 2\times \mathrm{Re}\{Vn\} + |n|^2 \right\rangle$$

$$= \left\langle |V|^2 \right\rangle + \left\langle 2\times \mathrm{Re}\{Vn\} \right\rangle + \left\langle |n|^2 \right\rangle$$

$$= |V|^2 + 2\times \mathrm{Re}\{V \cancel{\langle n \rangle}\} + \left\langle |n|^2 \right\rangle$$

$$= |V|^2 + \left\langle |n|^2 \right\rangle$$

Visibility estimator

V or V^2 ?

The $|V|^2$ estimator is biased but can easily be unbiased

The real $|V|^2$ estimator is:

$$\tilde{V}^2 = \left\langle |V+n|^2 \right\rangle - \left\langle |n|^2 \right\rangle$$

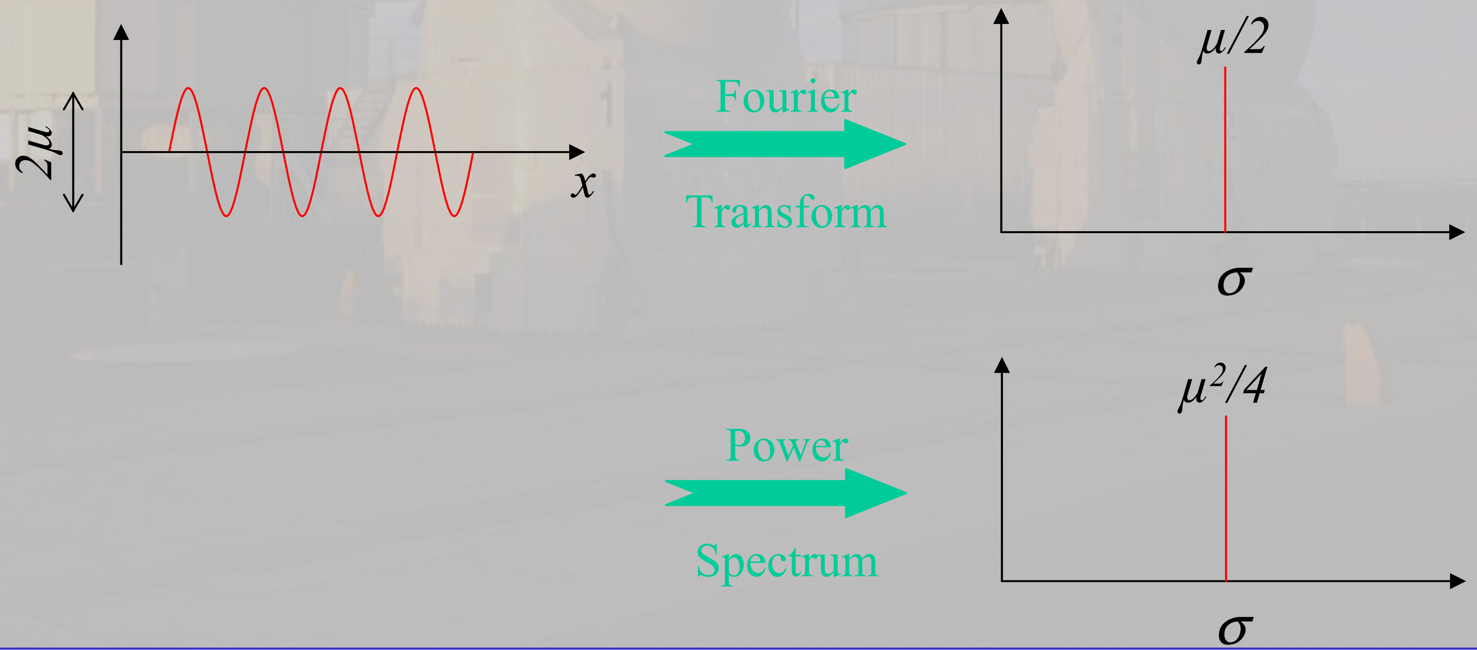
An unbiased estimator of the squared visibility is therefore obtained by subtracting the variance of the additive noise

This is why interferometrists talk about V^2 instead of V

Visibility estimator

How is μ^2 measured ?

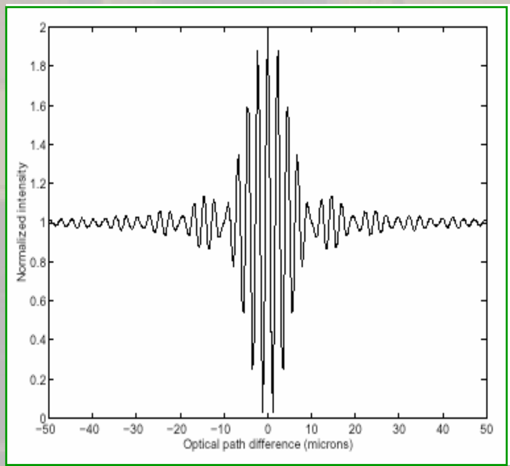
Very traditionally, a Fourier analysis of the fringe pattern is performed



Visibility estimator

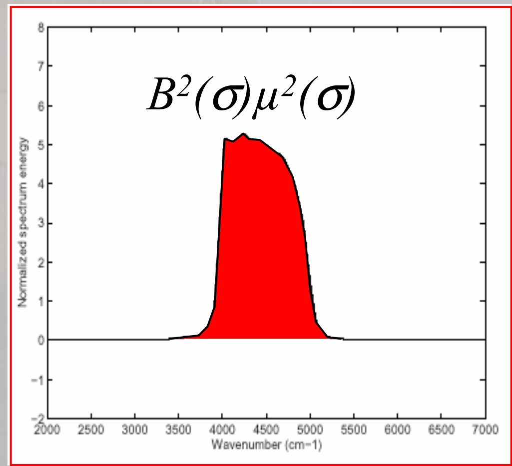
How is μ^2 measured ?

With a wide spectral band:



Power
Spectrum

B(σ) is the spectrum of the source



The integral of the PSD is proportional to $\left\langle \mu^2(\sigma) \right\rangle_{band}$

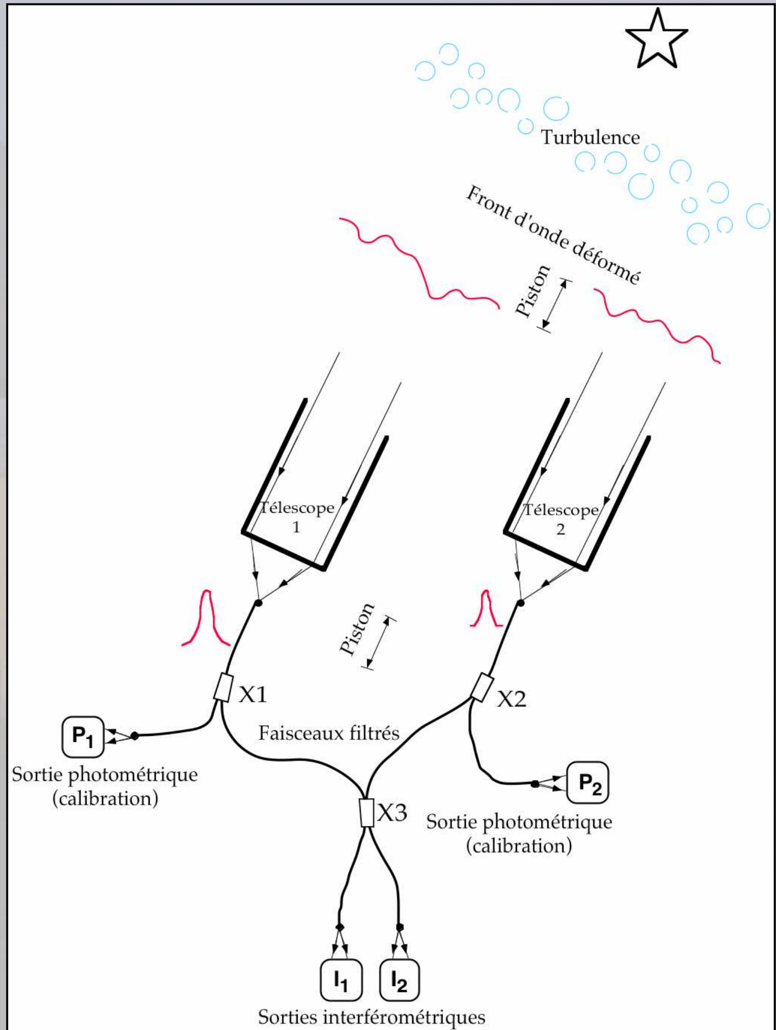
Outline

- ✓ The different types of beamcombiners
- ✓ What's the issue ?
- ✓ The noises
- ✓ Single-mode interferometers
- ✓ Estimators
- ✓ An example of *Fourier* estimator

7. Estimating error bars and calibrating visibilities

Single-mode beam combiner

FLUOR (Fiber Linked Unit for Optical Recombination)



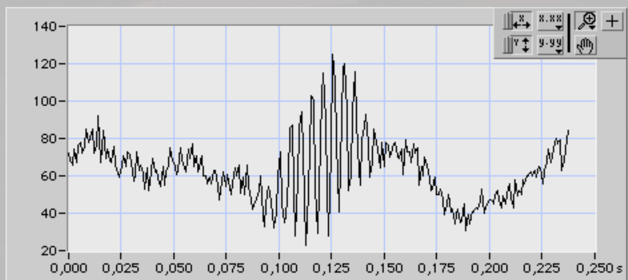
What do the signals look like ?

The two complementary interferometric channels:

$$I(t) = \kappa_A^{\pm} P_A(t) + \kappa_B^{\pm} P_B(t) \pm 2\sqrt{\kappa_A^{\pm}\kappa_B^{\pm}}\sqrt{P_A(t)P_B(t)}\int_{\sigma} C_{loss}(\sigma)\times V(\sigma)\times e^{-2i\pi\sigma v t - i\varphi_p(t)}\,d\sigma$$

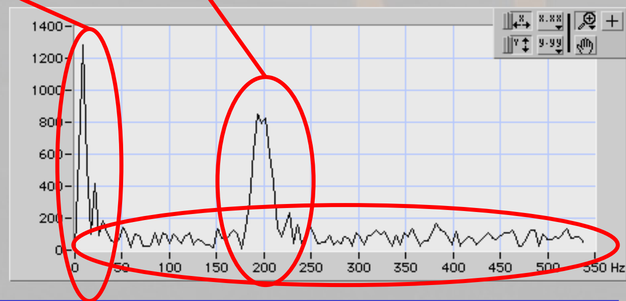
$$= \kappa_A^{\pm} P_A(t) + \kappa_B^{\pm} P_B(t) \pm 2\sqrt{\kappa_A^{\pm}\kappa_B^{\pm}}\sqrt{P_A(t)P_B(t)}\int_{\sigma} \mu(\sigma)\times e^{-2i\pi\sigma v t - i\varphi_p(t)}\,d\sigma$$

With $P_A(t)$ and $P_B(t)$ the photometric signals measured by the photometric channels and $\kappa_A^{\pm}$ and $\kappa_B^{\pm}$ the gains of the interferometric channels relative to the photometric channels

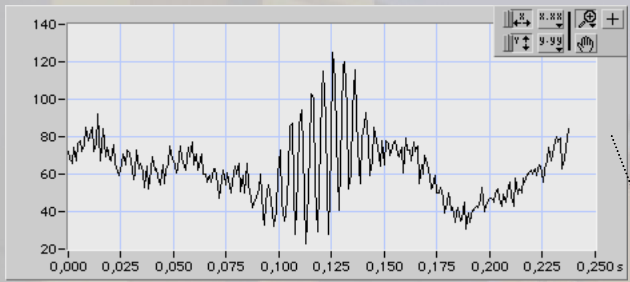


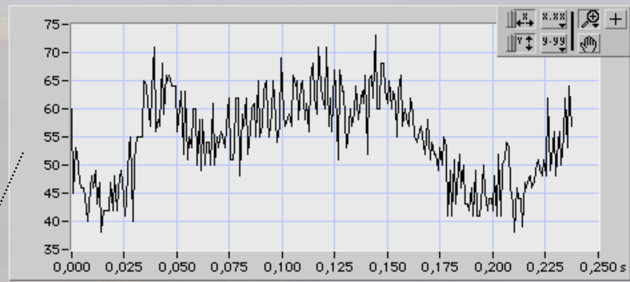
Fourier

Transform

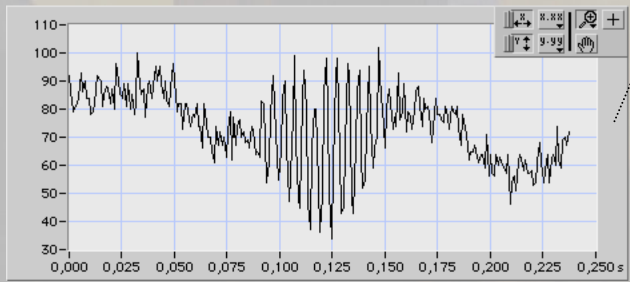


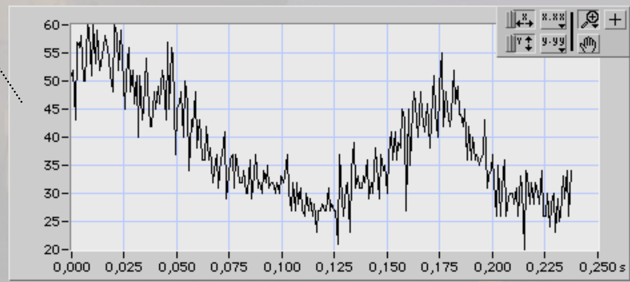
First step: measuring the gains





$$\begin{pmatrix} I^+ \\ I^- \end{pmatrix}_{LF} = \begin{bmatrix} \kappa_A^+ & \kappa_B^+ \\ \kappa_A^- & \kappa_B^- \end{bmatrix} \begin{pmatrix} P_A \\ P_B \end{pmatrix}$$

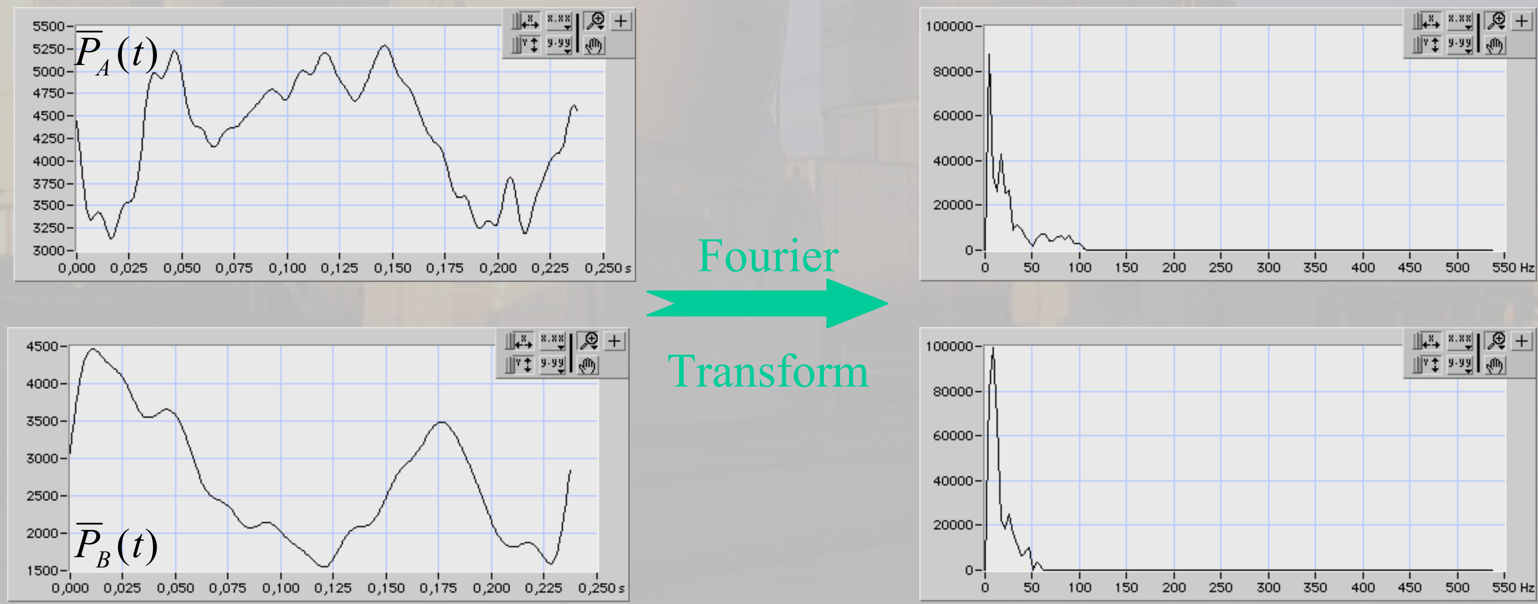




Gains are measured by alternatively blocking beams A and B and by fitting the interferometric signals with the photometric signals (the parameters are the κs)

Second step: estimating the photometric signals

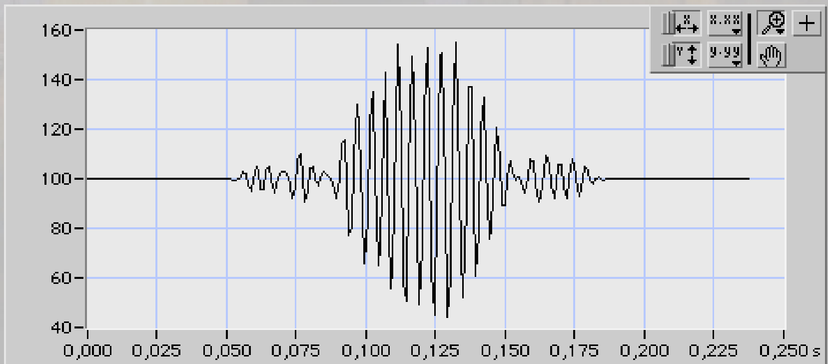
Photometric signals are filtered by an optimum filter to minimize rms fluctuations but keep turbulent fluctuations



Third step: normalizing the interferogram

Normalization of the interferogram (correction of the effects of turbulence, except for piston):

$$I^{\pm}_{norm}(t) = \frac{I^{\pm}(t) - \kappa^{\pm}_A \overline{P}_A(t) - \kappa^{\pm}_B \overline{P}_B(t)}{2\sqrt{\kappa^{\pm}_A \kappa^{\pm}_B \overline{P}_A(t) \overline{P}_B(t)}}$$

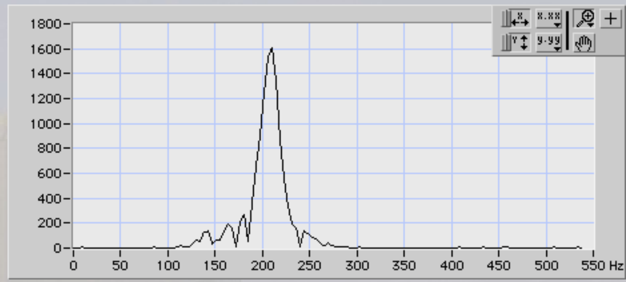


The same process is applied to the noise sequences that have been recorded

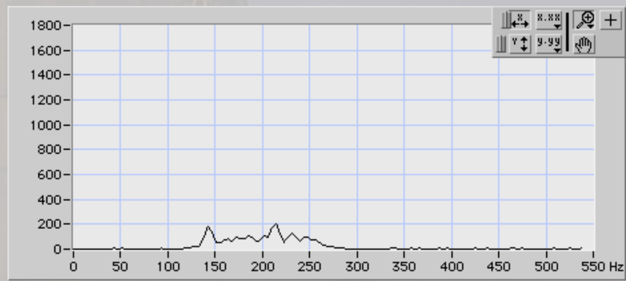
Why ?

Fourth step: computing coherence factors

Integrate the corrected interferogram PSD



Subtract the integrated ***processed noise PSD***



One noise PSD has not been subracted though. Which one ?

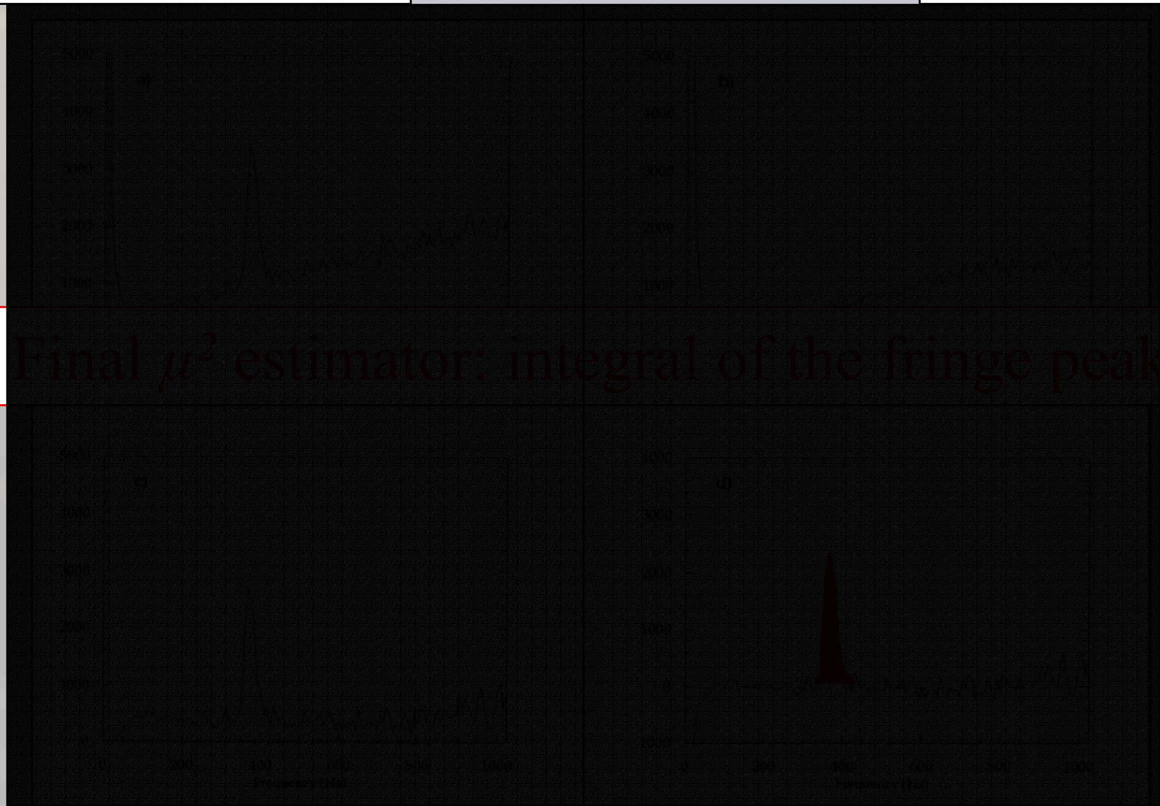
μ^2 processed source photon noise bias:

$$b_{ph}^{\pm} = \int_{-\infty}^{+\infty} \left[g^{\pm}(t)\right]^2 I^{\pm}(t)\,dt \qquad \text{with} \qquad g^{\pm}(t) = \frac{1}{2\sqrt{\kappa_A^{\pm}\kappa_B^{\pm}\overline{P}_A(t)\overline{P}_B(t)}}$$

An illustrated example

PSD of the normalized interferogram

PSD of the normalized noise



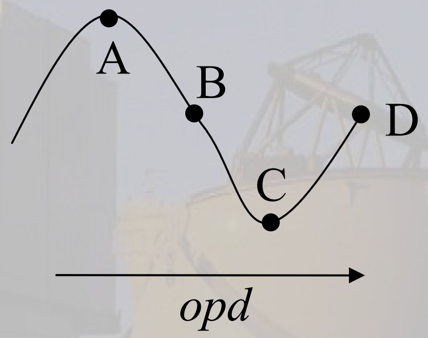
Final μ^2 estimator: integral of the fringe peak

PSD interferogram - PSD noise

PSD interferogram - PSD noise - PSD photon noise

Other co-axial V^2 estimator

ABCD:

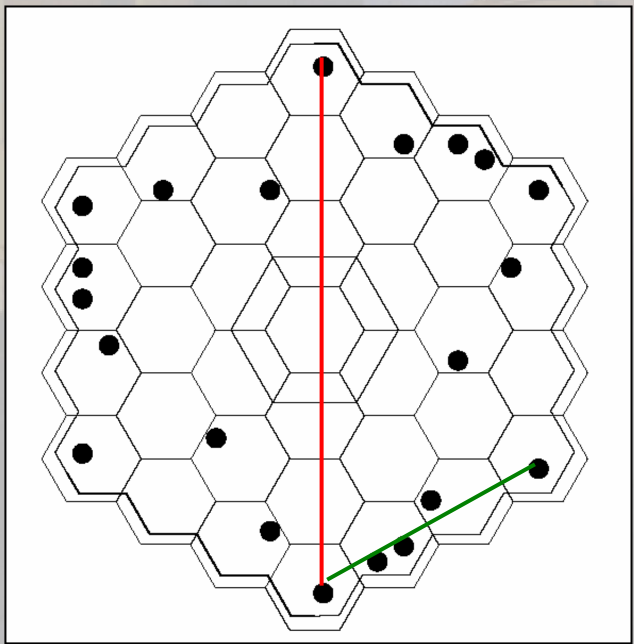


$$V^2 = 4\frac{(A-C)^2+(B-D)^2}{(A+B+C+D)^2}$$

It is indeed a 1-fringe Fourier estimator

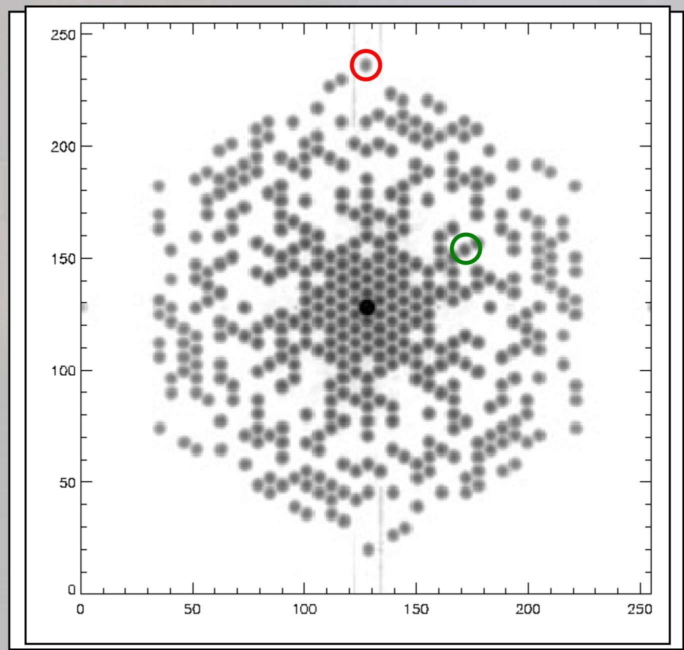
Quick example with a multi-axial beamcombiner

Pupil mask (Keck experiment)



Frequency peaks

Image plane



Outline

- ✓ The different types of beamcombiners
- ✓ What's the issue ?
- ✓ The noises
- ✓ Single-mode interferometers
- ✓ Estimators
- ✓ An example of *Fourier* estimator
- ✓ Estimating error bars and calibrating visibilities

Importance of error bars assessment

Not such a huge issue in radio interferometry (this is a personal opinion) as the amount of *uv* points is usually large (except probably for VLBI) and statistics can be directly estimated from calibrated visibilities.

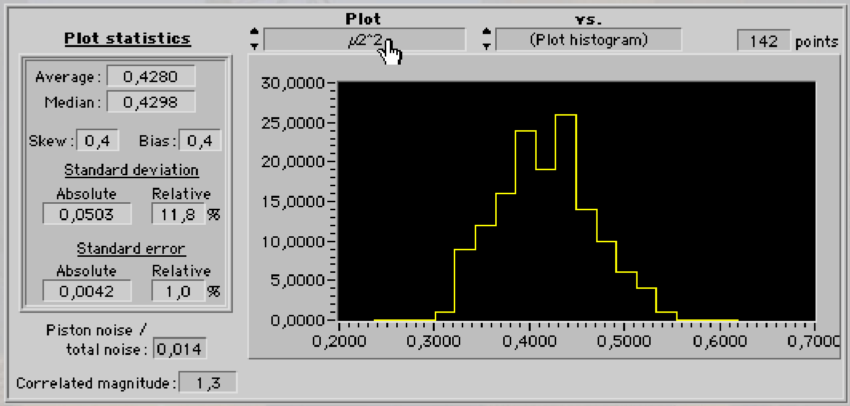
At optical wavelengths, *uv* point sampling is poorer (I hope this will change !).

Each visibility point therefore has a large relative weight in the data set.

Each point should therefore be well calibrated and the error be well estimated.

Final μ^2 estimate and error bar

Squared coherence factors are computed for each scan in each interferometric channel

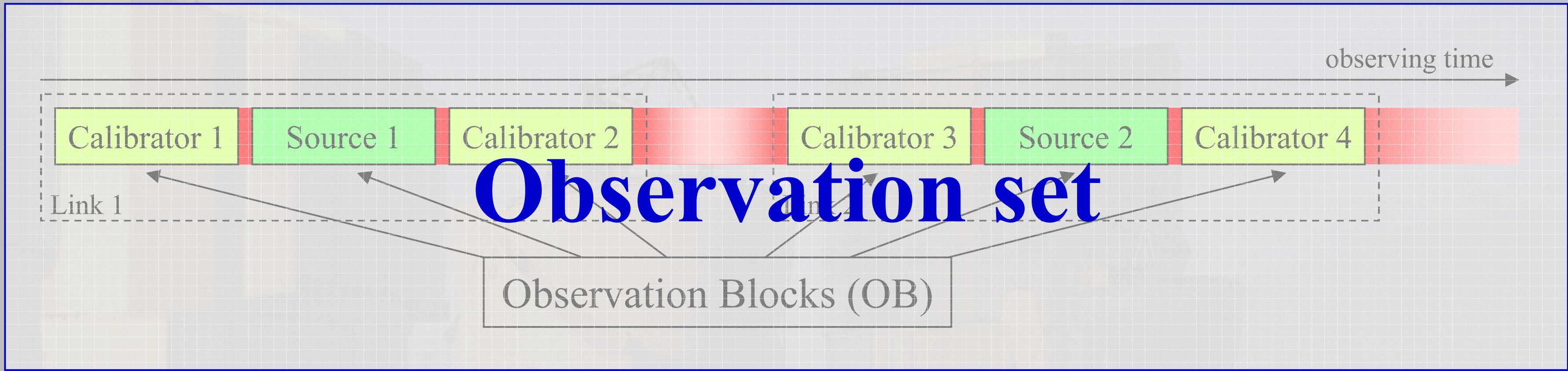


They define a statistics (histogram) from wich a standard deviation is derived

Eventually one gets:

$$\left.\begin{array}{l}\mu_+^2 \pm \sigma(\mu_+^2) \\ \mu_-^2 \pm \sigma(\mu_-^2)\end{array}\right\} \Rightarrow \boxed{\mu^2 \pm \sigma(\mu^2)}$$

Principle of calibration



- 1 Observation set = 1 set-up
  - same night
  - same detector parameters (frame rate, number of frames, ...)
  - same filter...

- Principle : follow slow coherence loss fluctuations

Steps

1. Derive the ***expected visibility of the calibrator***

   usually a uniform disk diameter is used to predict visibility at the spatial frequency *S*

$$V_{\exp}(S) = \left| \frac{2J_1(\pi\theta_{UD}S)}{\pi\theta_{UD}S} \right|$$

2. Derive the ***instantaneous transfer function*** for each channel

$$T_i^2(t_1) = \frac{\mu_i^2}{V_{\exp}^2(S)}$$

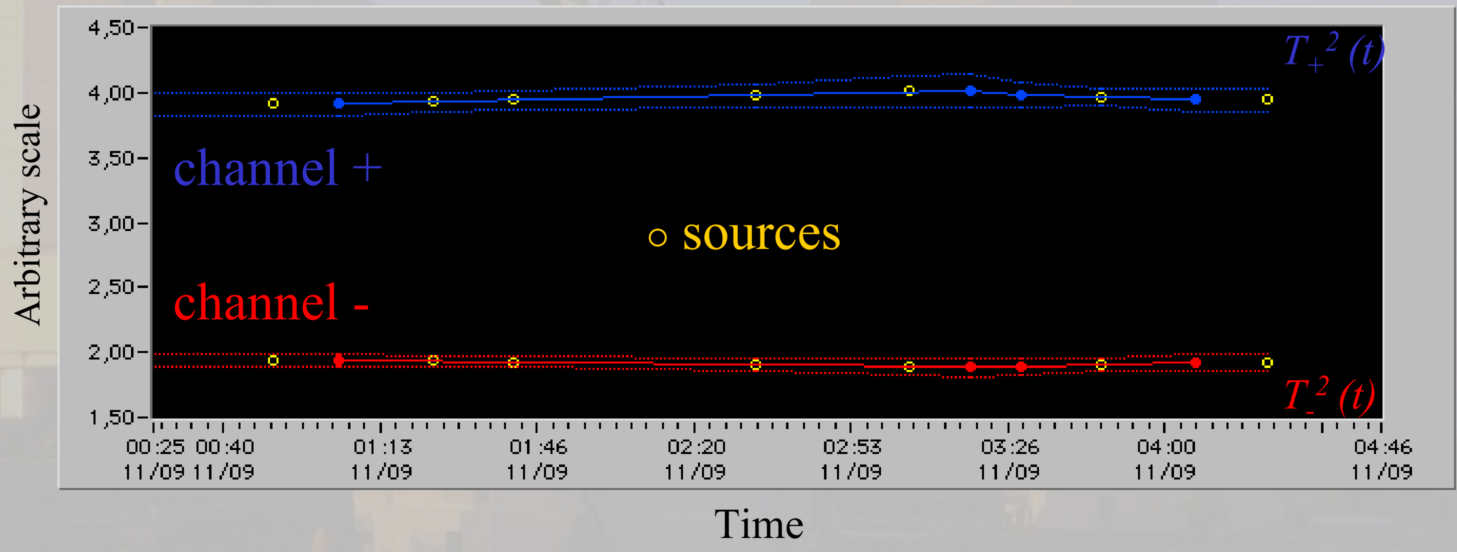
3. Estimate the transfer function ***at the time when the science target was observed***

Different methods may be used

4. ***Calibrate the visibility of the science target***

$$V^2 = \frac{\mu^2}{T^2(\tau)}$$

Example of transfer function



Propagation of errors

<u>Sources of errors (1σ error bars):</u>

- errors on coherence factors (detector noise, photon noise, piston noise)
- errors on the diameter of calibrators

<u>Propagation of errors:</u>

- The final estimate of the squared visibility is the product and ratio of hopefully gaussian random variables.

- The ratio of two centered gaussian random variables is a Cauchy distribution of non-defined mean and variance ! This is potentially dangerous when the SNR is low.

Propagation of errors

1st method to propagate errors:

$$V^2 = \frac{\mu^2}{\mu_c^2} \times V_c^2$$

- make an expansion of the V^2 estimator if error bars are small

$$dV^2 = \frac{V^2}{\mu^2} \times d\mu^2 + \frac{V^2}{V_c^2} \times dV_c^2 - \frac{V^2}{\mu_c^2} \times d\mu_c^2$$

- and sum the weighted variances of the errors

$$\sigma^2(V^2) \approx \left(\frac{V^2}{\mu^2}\right)^2 \times \sigma^2(\mu^2) + \left(\frac{V^2}{V_c^2}\right)^2 \times \sigma^2(V_c^2) + \left(\frac{V^2}{\mu_c^2}\right)^2 \times \sigma^2(\mu_c^2)$$

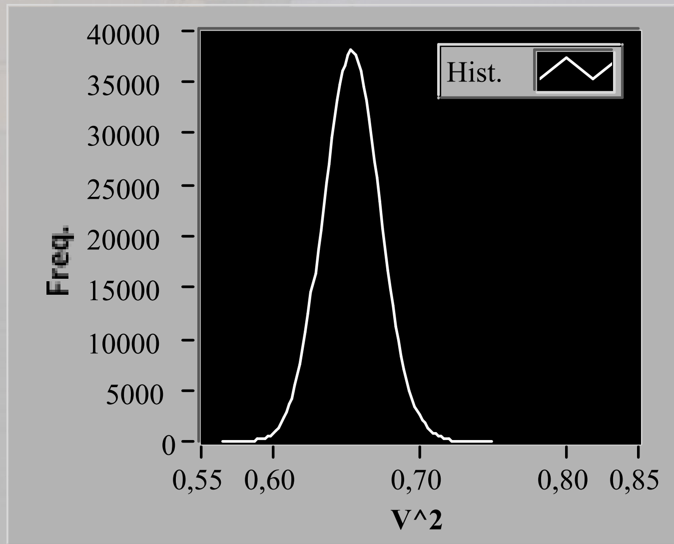
only valid if errors are small

Propagation of errors

2nd method to propagate errors:

$$V^2 = \frac{\mu^2}{\mu_c^2} \times V_c^2$$

- simulate the random variable and compute the variance of the simulated statistical distribution



$$\left.\begin{aligned} \mu^2 &= 0.400 \pm 0.010 \\ \mu_c^2 &= 0.600 \pm 0.010 \\ V_c^2 &= 0.980 \pm 0.001 \end{aligned}\right\} \Rightarrow \quad V^2 = 0.654 \pm 0.020$$

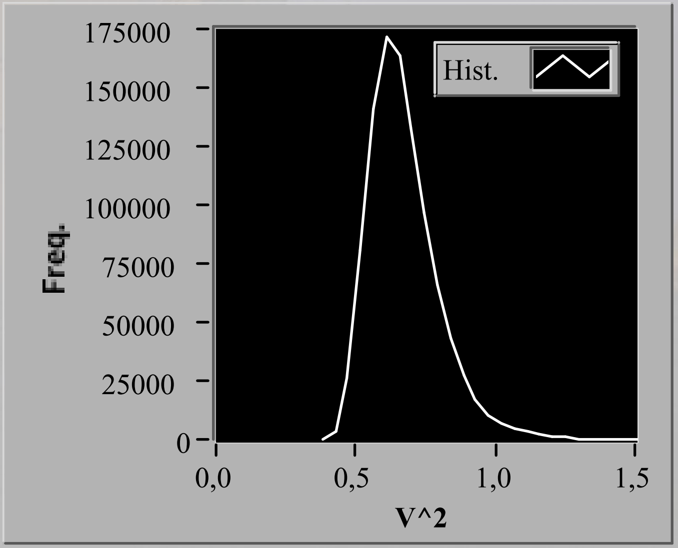
Analytical method $\Rightarrow \quad V^2 = 0.653 \pm 0.020$

Propagation of errors

2nd method to propagate errors:

$$V^2 = \frac{\mu^2}{\mu_c^2} \times V_c^2$$

- simulate the random variable and compute the variance of the simulated statistical distribution



$$\left.\begin{aligned} \mu^2 &= 0.400 \pm 0.010 \\ \mu_c^2 &= 0.600 \pm 0.100 \\ V_c^2 &= 0.980 \pm 0.001 \end{aligned}\right\} \Rightarrow V^2 = 0.670 \pm 0.126$$

Analytical method $\Rightarrow V^2 = 0.653 \pm 0.110$

- this method is more robust as it also works with large error bars

Another issue, often overlooked

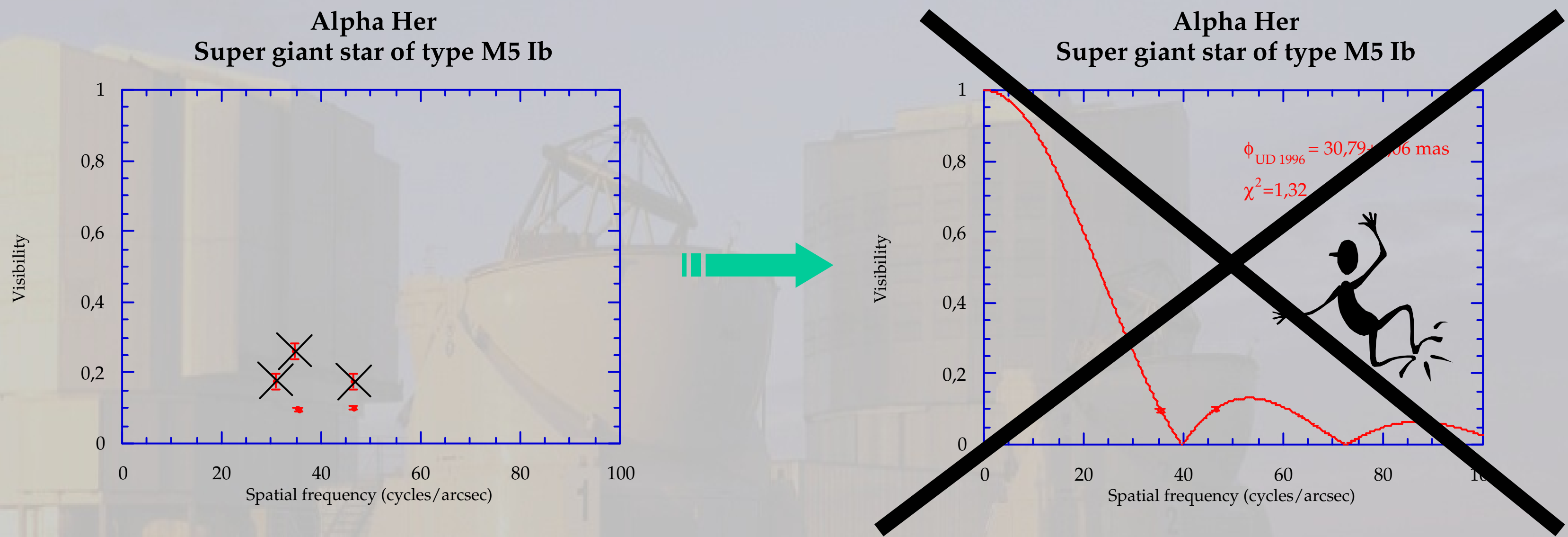
Correlations between visibilities recorded at different times, with a different baseline,..., also have to be taken into account for model fitting.

In this case the correlation is due to the use of common calibrators (the expected visibilities are then correlated)

It is therefore necessary that the data reduction program outputs numbers to compute the correlation *a posteriori* …

Examples will be shown in L11

Rejecting bad data

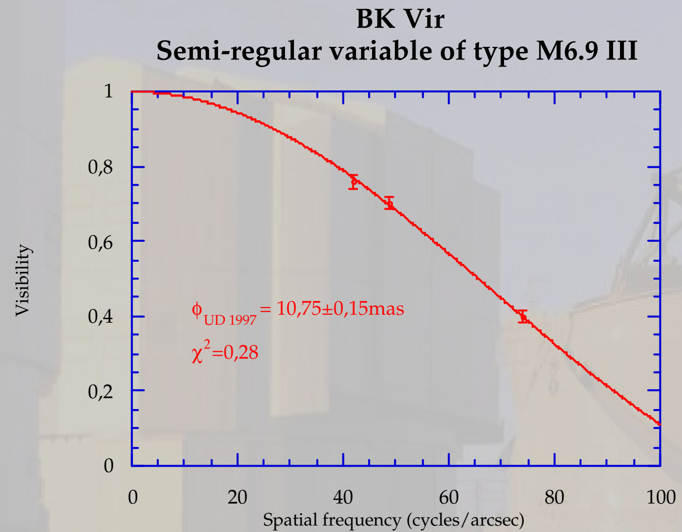


<u>Examples of selection criteria:</u>

- reject data for which the instrument was not stable (varying transfer function)
- (probably) reject data for which statistical distributions of μ^2 are not gaussian

Examples will be shown in L11

Have visibilities been well estimated?



Best accuracies with single-mode interferometers: a few 0.1%