Data processing for Interferometry

Antoine Mérand

VLT(i) support astronomer
AMBER Paranal instrument scientists

VLTI school – Porquerolles 2010

This course is largely inspired by material provided by Guy Perrin from the 2006 school

Why this course?

For you to be critical:

- Understand the limitations of the method
- A DRS might gives you inconsistent results night to night...

For you to better understand the technique:

- Define better observations strategy
- Be able to interpret data in an astrophysical context

This course is not...

A recipe to write your own DRS

A how-to for current interferometers' DRS

An extensive overview

Basic ideas

- Understand the interferometric signal
- Break down the limitations
- See how instruments are designed to cope with the disturbances
- See what is left and how we can extract what we are interested in

Advanced techniques, to handle the biases, will be addressed in the second class.

What is the issue?

The fringes' signal has a simple form:

$$I(\delta) = 1 + \mathrm{Re}\left(Ve^{-2\pi i\delta/\lambda}\right)$$

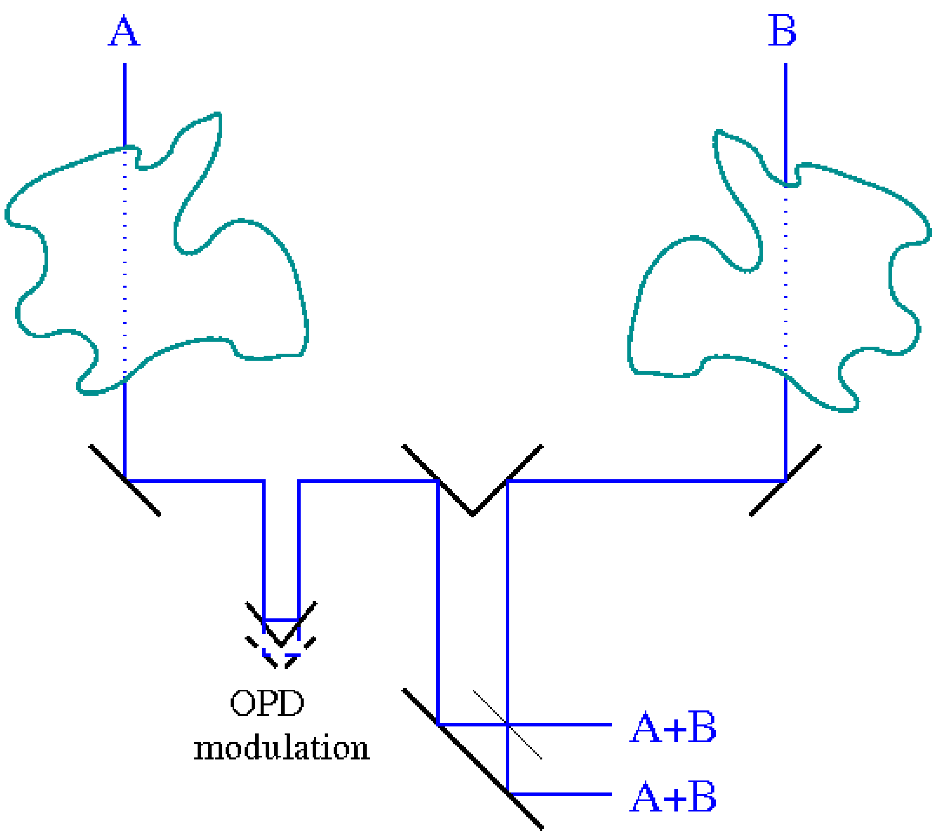
one can linearly estimate the visibility:

$$\mathrm{Re}(V) = I(0) - 1$$
$$\mathrm{Im}(V) = I(\lambda/4) - 1$$

So, there are no issues...

However...

- Each telescope sees a different atmospheric patch
- optical path = n(T,P) x L
- Temperature and Pressure have turbulent behaviors
- Optical Path Delay (OPD) jitter is very strong in the optical...



What is the issue

Actual fringes have a more complex expression

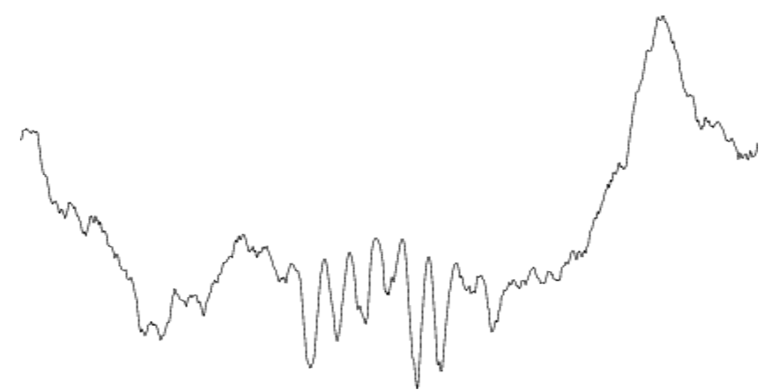
$$I(\delta(t)) = 1 + \mathrm{Re}\left(Ve^{-2\pi i\delta(t)/\lambda - 2\pi i \mathrm{OPD_{jitter}(t)}/\lambda}\right)$$

The linear approach fails... and the traditional approach is not robust to noise:

this works nicely on a sin wave:

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

real data look more like this:



Estimators robust to noise are necessary!

the real signal is noisy
Additive noises

Lets consider the ideal interferograms:

$$I(\delta) = 1 + \mathrm{Re}\left(Ve^{-2\pi i\delta/\lambda}\right)$$

There are many contributors:

$$I_n(\delta(t)) = I(\delta(t)) + n_{\mathrm{ph}} + n_{\mathrm{det}} + \mathrm{Back}(t) + n_{\mathrm{phot.Back}}$$

Photon shot — detector — background — background photon shot

with $var(n_{\mathrm{ph}}) = I(\delta(t))$

$$var(n_{\mathrm{phot.Back}}) = \mathrm{Back}(t)$$

Only “Back(t)” does not have 0-mean. If a dominant source of noise, can be removed by chopping (e.g. MIDI)

the real signal is noisy

Multiplicative noises

In case of unbalanced beams, the normalized interferogram becomes:

$$I(\delta) = \frac{P_\mathrm{A} + P_\mathrm{B} + 2\sqrt{P_\mathrm{A} P_\mathrm{B}} \times \mathrm{Re}\left(V e^{-2\pi i \delta/\lambda}\right)}{P_\mathrm{A} + P_\mathrm{B}}$$

In general, $P_A \neq P_B$ because of alignments, scintillation etc. The instantaneous contrast becomes biased by

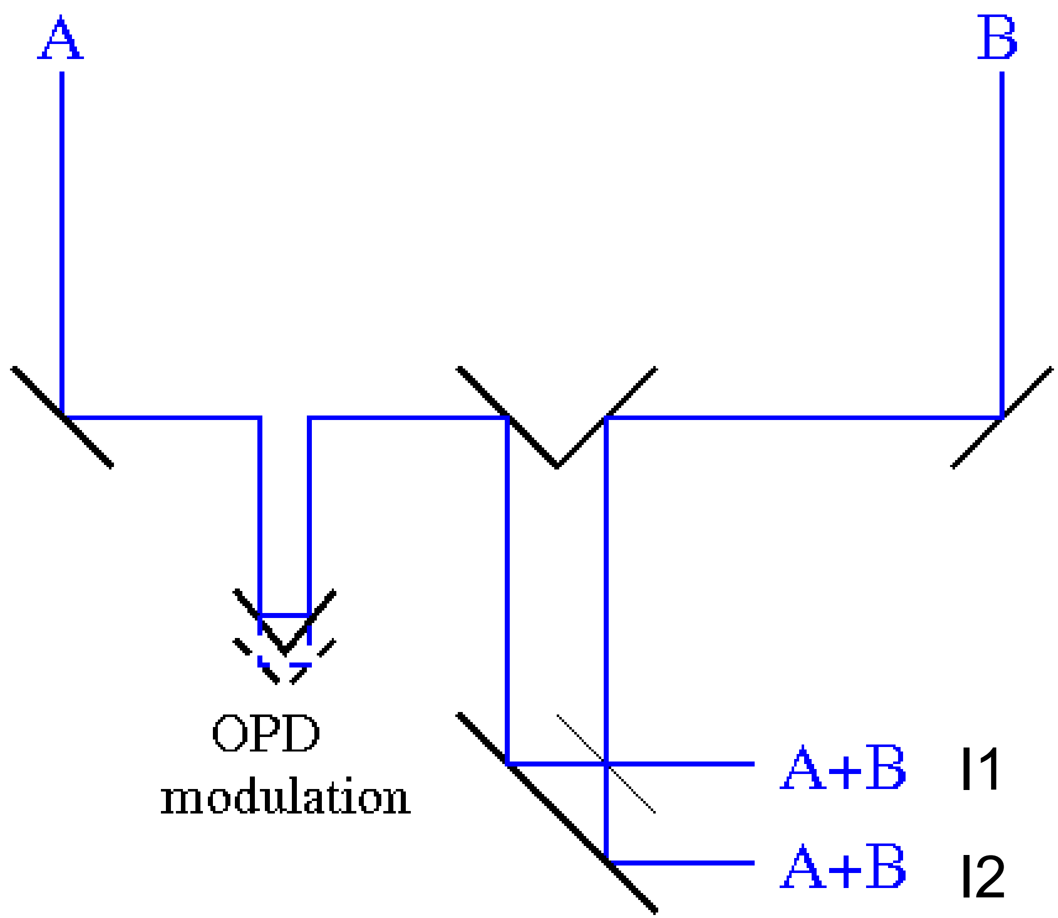
$$\frac{2\sqrt{P_\mathrm{A} P_\mathrm{B}}}{P_\mathrm{A} + P_\mathrm{B}} = 0.94 \text{ for } P_\mathrm{A} = 2P_\mathrm{B}$$

Photometric channels

Measure P_A/P_B using shutters, before and/or after taking fringes.

– or –

Simultaneously monitor the photometry using **dedicated channels.**

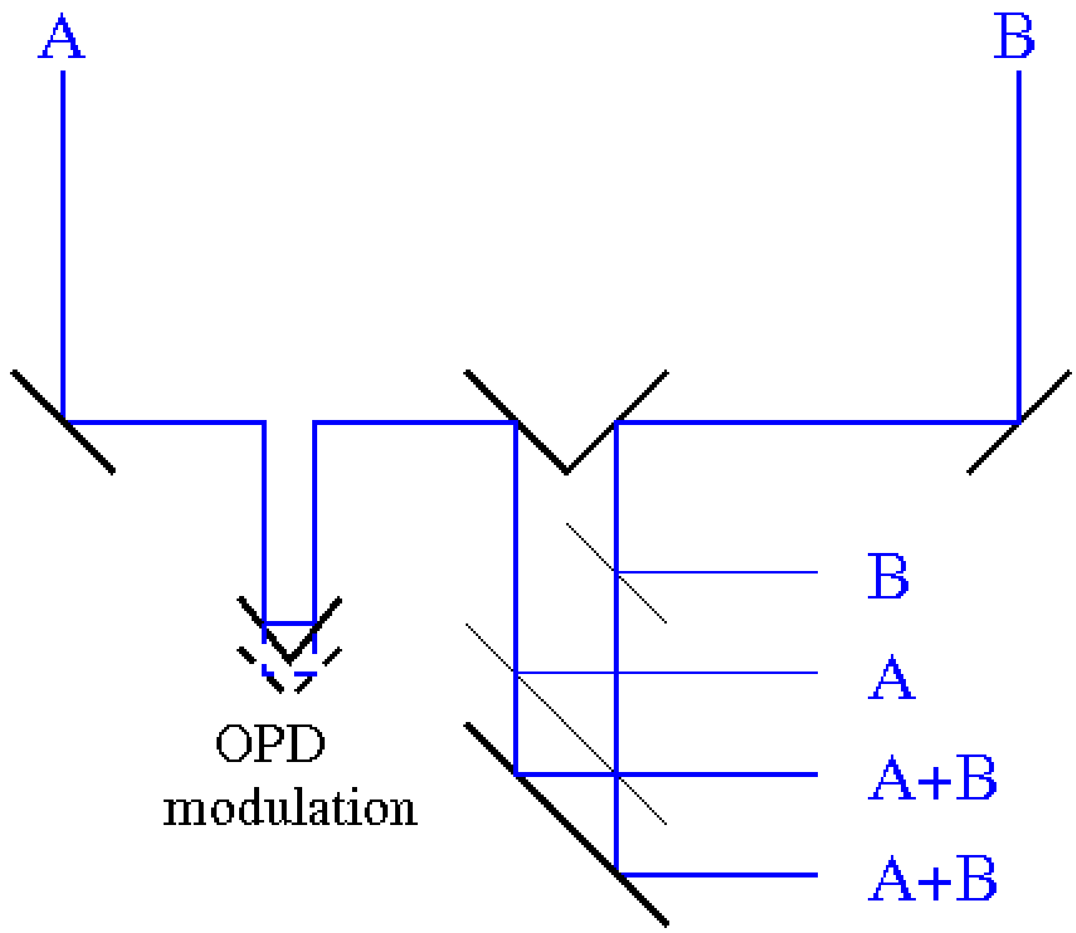


Photometric channels

Measure P_A/P_B using shutters, before and/or after taking fringes.

– or –

Simultaneously monitor the photometry using **dedicated channels.**



Other multiplicative noises...

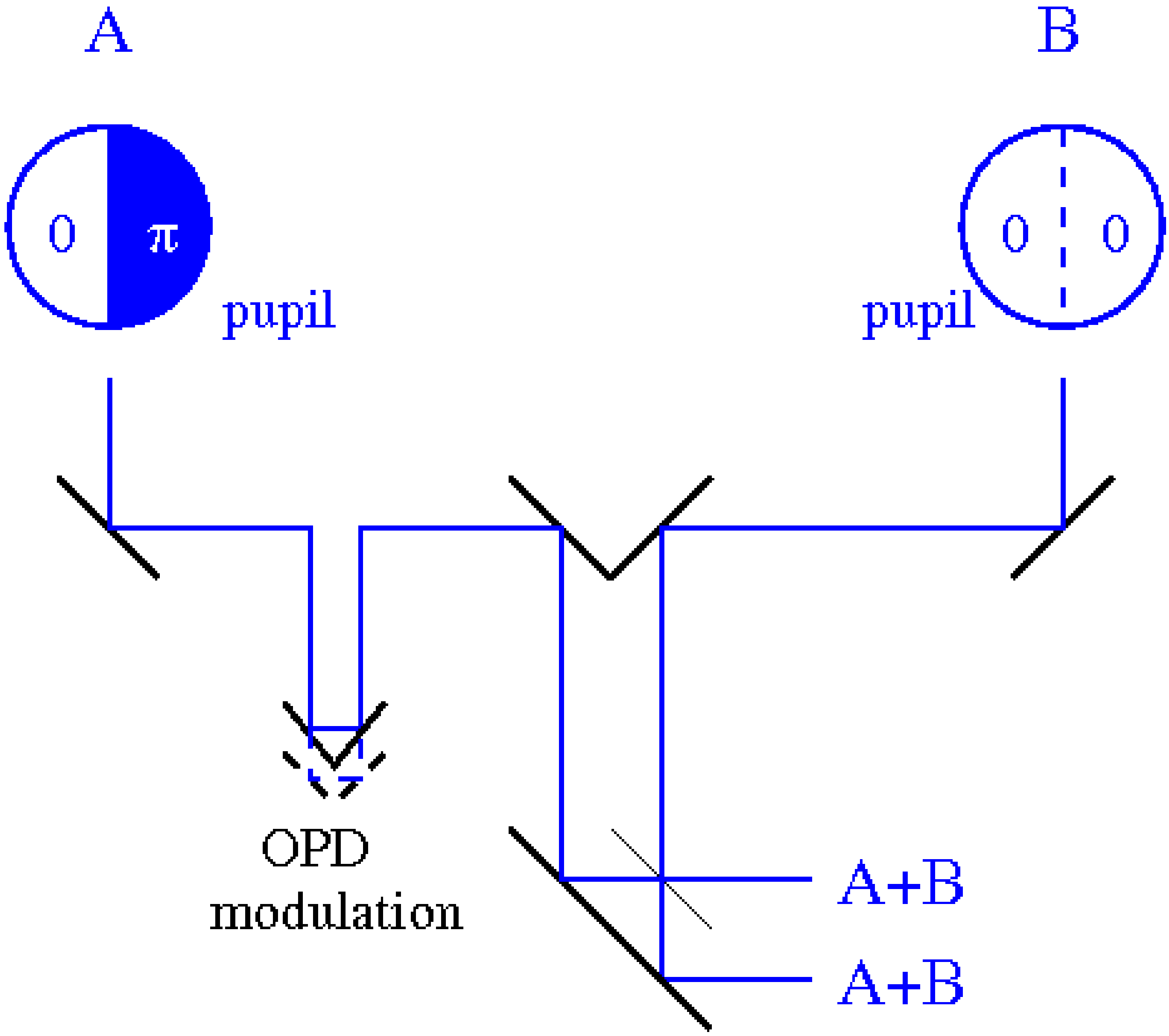
the really nasty one

The atmospheric turbulence have 2 effects:

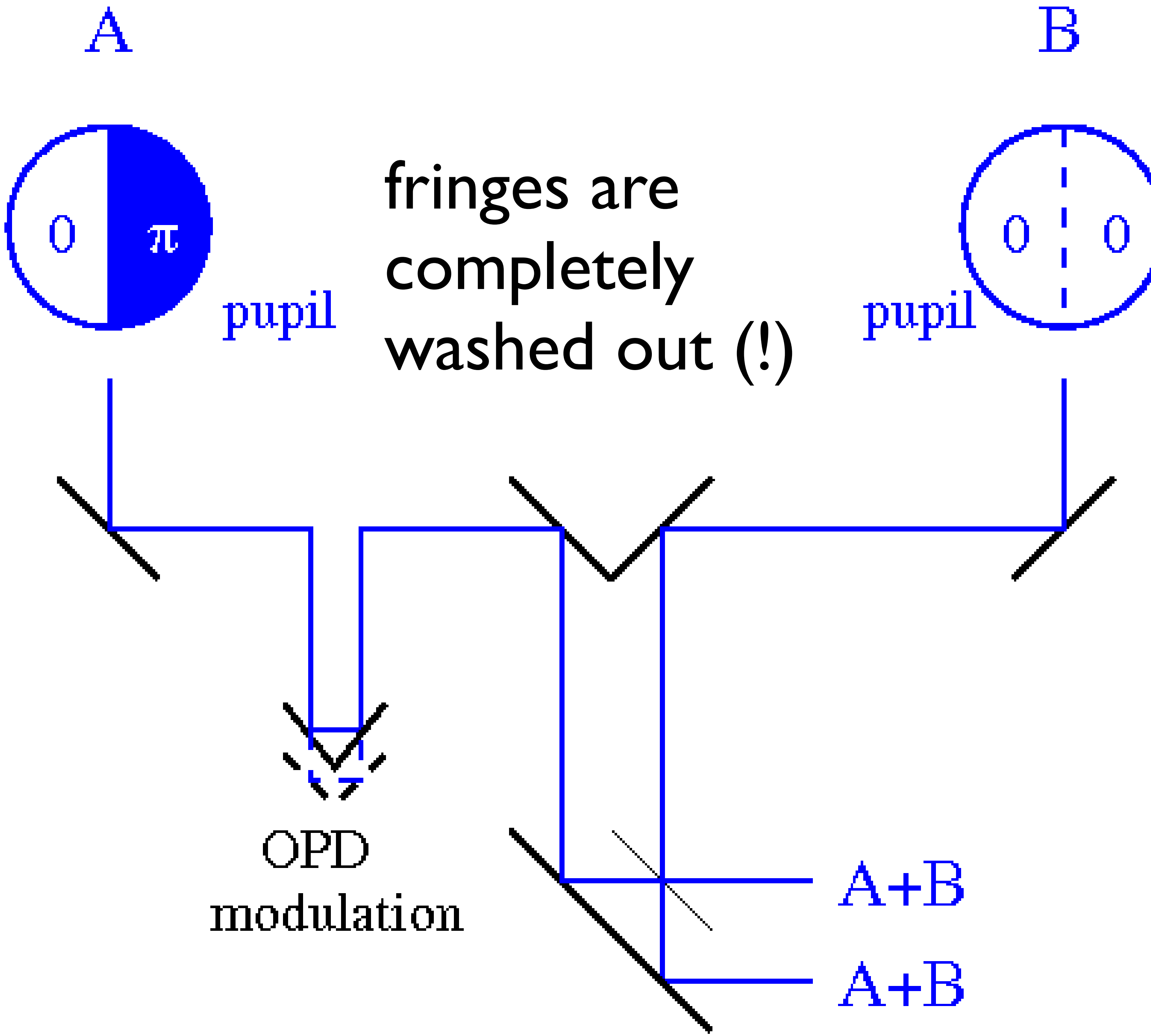
$$I[\delta(t)] = 1 + e^{-\sigma(t)^2} \mathrm{Re}\left[V e^{-2\pi i (\delta(t) + \mathrm{OPD_{jitter}(t)})/\lambda} \right]$$

We already saw the OPD jitter, but there is also the loss of coherence due to phase variance over the input pupils.

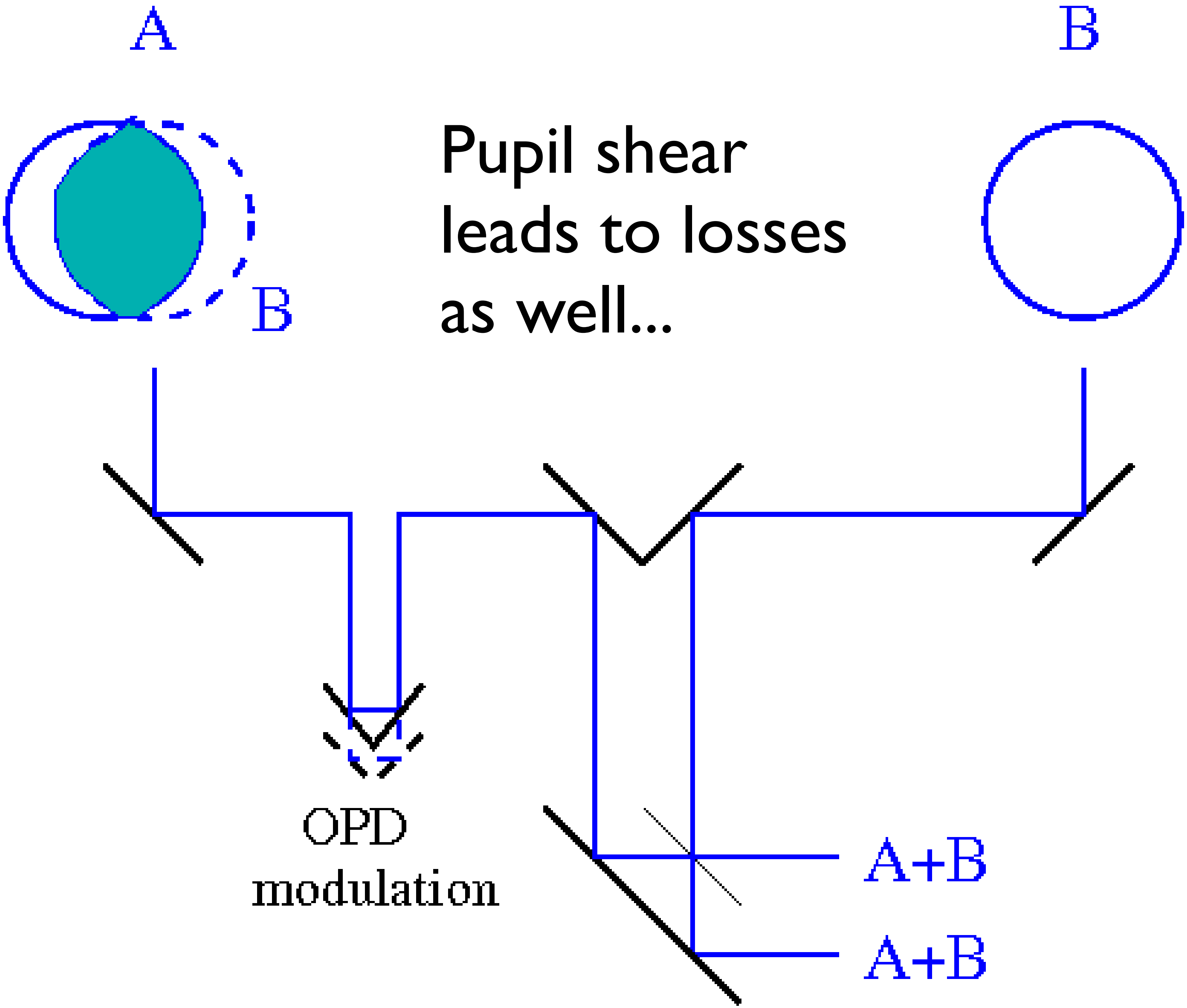
Basically, if the input pupils are larger than r_0, (atmospheric Fried Parameter) the loss is large and variable.


A
B
0
π
pupil
0
0
pupil
OPD
modulation
A+B
A+B

fringes are completely washed out (!)



Pupil shear leads to losses as well...



How to cope with the losses?

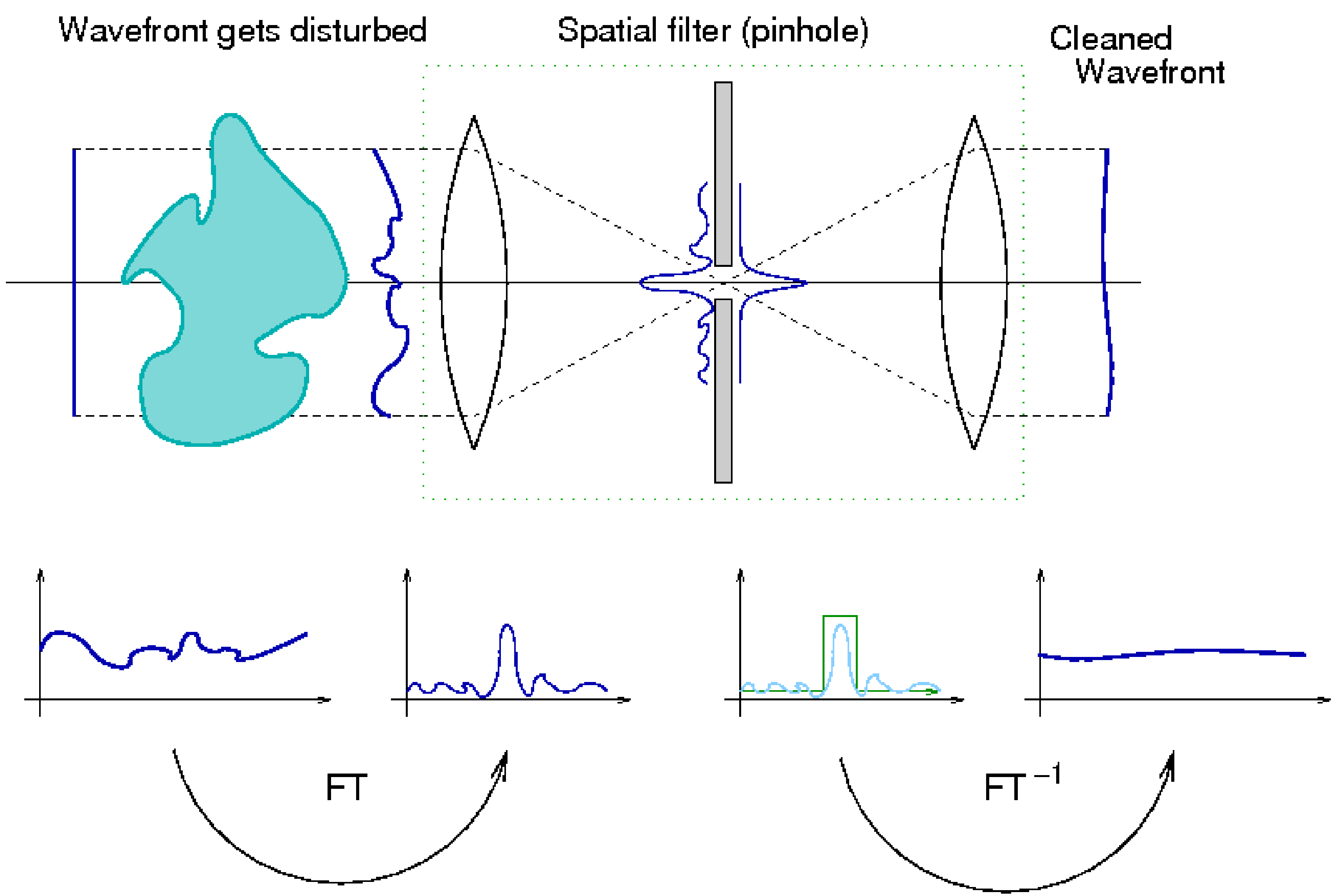
Deal with them:

- assume you can monitor the losses
- use stellar calibrators (with predictable visibilities)

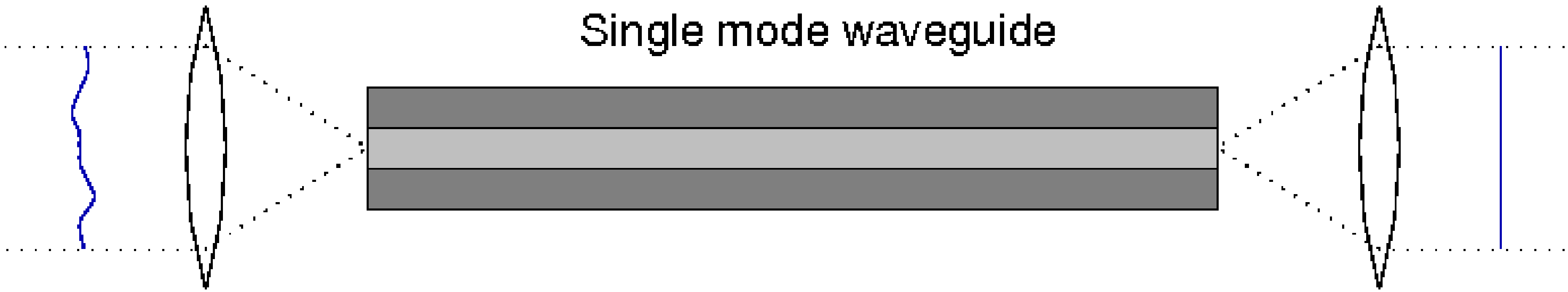
Reduce them:

- Stop your telescopes' aperture to r_0 (~10cm in the visible; <1m in the near infrared in the best sites)
- Use adaptive optics to correct the wavefront (AO known to have variable performances)

Alternative way: spatial filtering



A better way: modal filtering



- A single mode waveguide only propagates its fundamental mode (~gaussian)
- The other modes are lost (dissipated)
- The output wave front is almost perfectly flat and only has (important) intensity fluctuations

Spatial/Modal filtering

Advantages:

- instrumental contrast is close to 100%
- very stable instrument since mostly passive
- insensitive to pupil shear: decouple input/output alignments

Disadvantages

- not all the light gets through... but you keep the right photons
- requires (very) accurate alignment
- single mode fibers have chromatic dispersion issues
- important flux variations: requires simultaneous photometric monitoring

So we get the final signal

$$F(\delta, t, \lambda) = \sum \text{noises} + P_{\text{A}}(t) + P_{\text{B}}(t) +$$
$$2\sqrt{P_{\text{A}}(t) P_{\text{B}}(t)} \text{Re} \left[V e^{-2\pi i \delta/\lambda - 2\pi i \text{j}(t)/\lambda} \right]$$

- "exp(- ^{2})" bias is gone thanks to spatial/modal filtering
- P_A and P_B and monitored in real time so fringes can be corrected and normalized
- We have to extract the complex visibility V and
  - be robust to the noise
  - be robust to jitter [j(t) in the phase term]

Note about “Visibilities”

- We are talking about the fringes' contrast, not the actual visibility of the object.
- The measured visibility is not the visibility of the object: the instrumental response is not 100%. We use calibrator stars to correct from:
  - instrumental effects
  - atmospheric effects
- From now on, “visibility” means uncalibrated fringes' contrast...

Visibility estimators: |V| or V²?

What happen if you average the visibility with additive noises:

$$\begin{aligned} V' &= V + n \\ |V'| &= |V + n| \\ \langle |V'| \rangle &= \langle |V + n| \rangle \end{aligned}$$

$$\begin{aligned} \left\langle |V'|^2 \right\rangle &= \left\langle |V + n|^2 \right\rangle \\ &= \left\langle |V|^2 \right\rangle + \langle 2\mathrm{Re}[Vn] \rangle + \left\langle |n|^2 \right\rangle \\ &= |V|^2 + 2\mathrm{Re}[V \langle n \rangle] + \left\langle |n|^2 \right\rangle \\ &= |V|^2 + \left\langle |n|^2 \right\rangle \end{aligned}$$

Averaging $|V'|^2$ instead of |V'| allows to correct from 0-mean additive noises

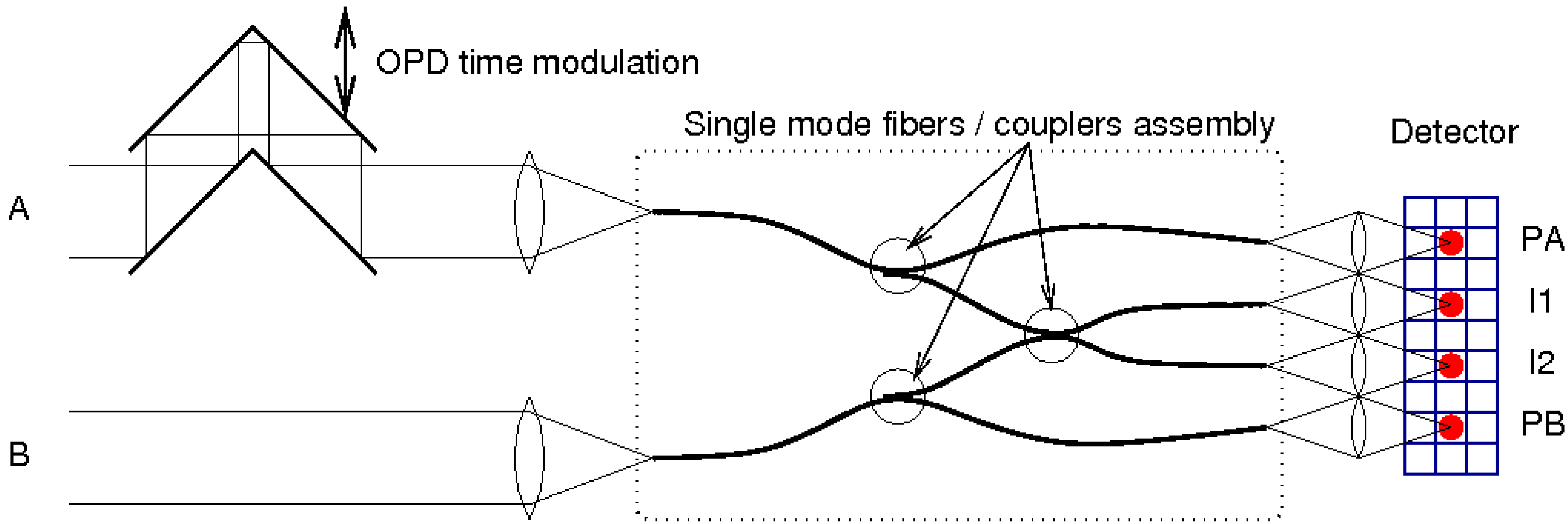
Estimation of the visibility's modulus

- Extract $|V'|^2$ (fringe contrast) for each frame
- Average it for all the batch: $\langle|V'|^2\rangle$
- Estimate the noise variance: $\langle|n|^2\rangle$
- Estimate the unbiased fringes' squared contrast:

$$\mu^2 = \left\langle |V'|^2 \right\rangle - \left\langle |n|^2 \right\rangle$$

- Measure μ^2 for a known target (calibrator) and predict its visibility from a model (V^2)
- Calibrate your object' visibility: $V^2_{\mathrm{sci}} = \mu^2_{\mathrm{sci}} \times \left[\frac{V^2_{\mathrm{cal}}}{\mu^2_{\mathrm{cal}}} \right]$

Example: FLUOR (or VINCI)



- 2 T in K band
- uses single mode fibers and couplers (~mixers)
- time modulation of the OPD ('Michelson' or co-axial)

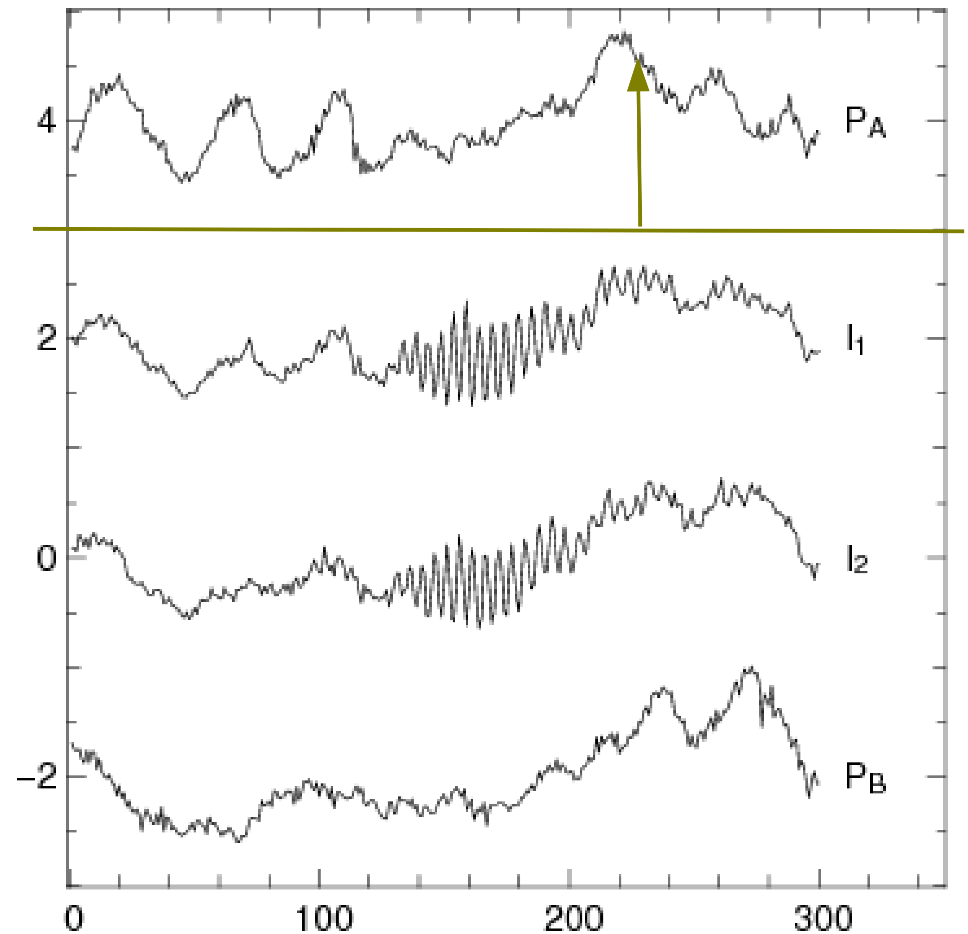
from raw signals to fringe signal

The steps described in the following for FLUOR are almost universal. You will find the same steps in AMBER and MIDI data reduction process.

raw signals

- photometric variations are important (and these are nice data...)
- 2 interferometric channels have opposite phases
- correction matrix:

$$\begin{aligned} I_1 &= \kappa_{1\mathrm{A}} P_\mathrm{A} + \kappa_{1\mathrm{B}} P_\mathrm{B} \\ I_2 &= \kappa_{2\mathrm{A}} P_\mathrm{A} + \kappa_{2\mathrm{B}} P_\mathrm{B} \end{aligned}$$

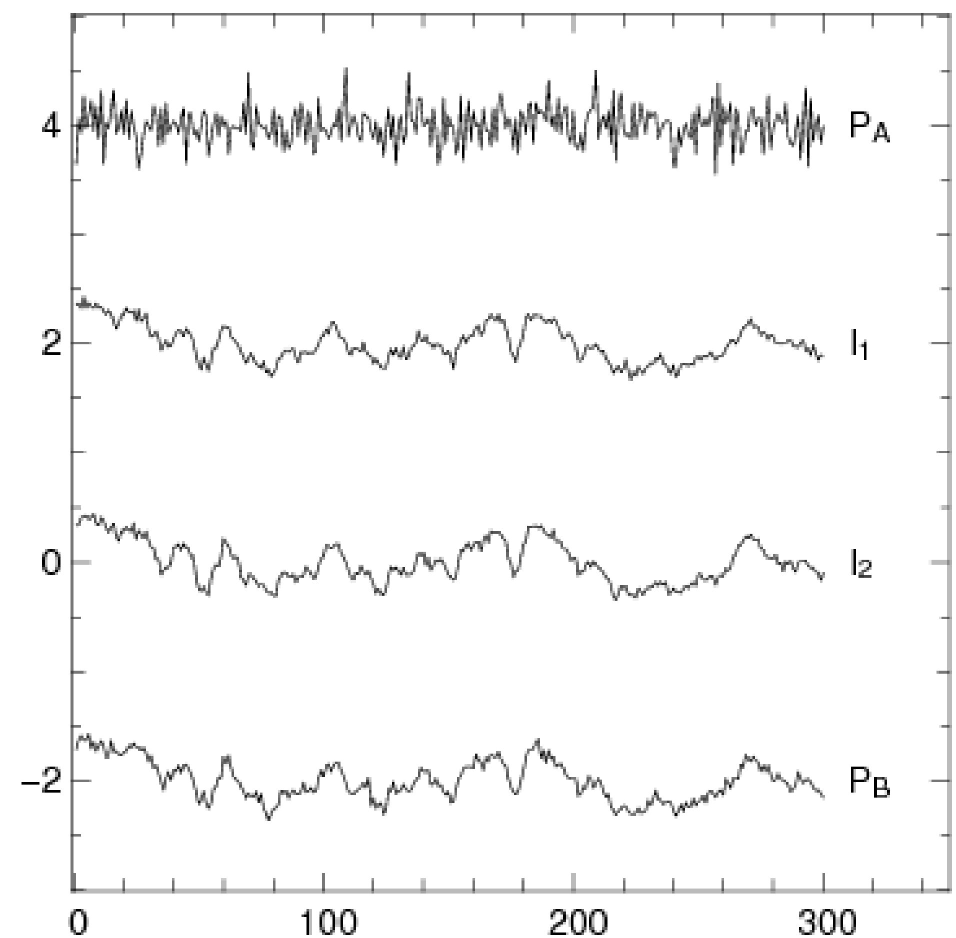


kappa matrix

Estimated by closing shutters in sequence

$$
\begin{aligned}
I_1 &= \kappa_{1\mathrm{A}} P_\mathrm{A} + \kappa_{1\mathrm{B}} P_\mathrm{B} \\
I_2 &= \kappa_{2\mathrm{A}} P_\mathrm{A} + \kappa_{2\mathrm{B}} P_\mathrm{B}
\end{aligned}
$$

estimated for each observation since FLUOR reads a single pixel per channel: output alignment is critical



Photometric correction

- Photometric correction
- normalization factor is smoothed and checked for 0-crossing
- final fringe signal is formed, to decrease correlated noises

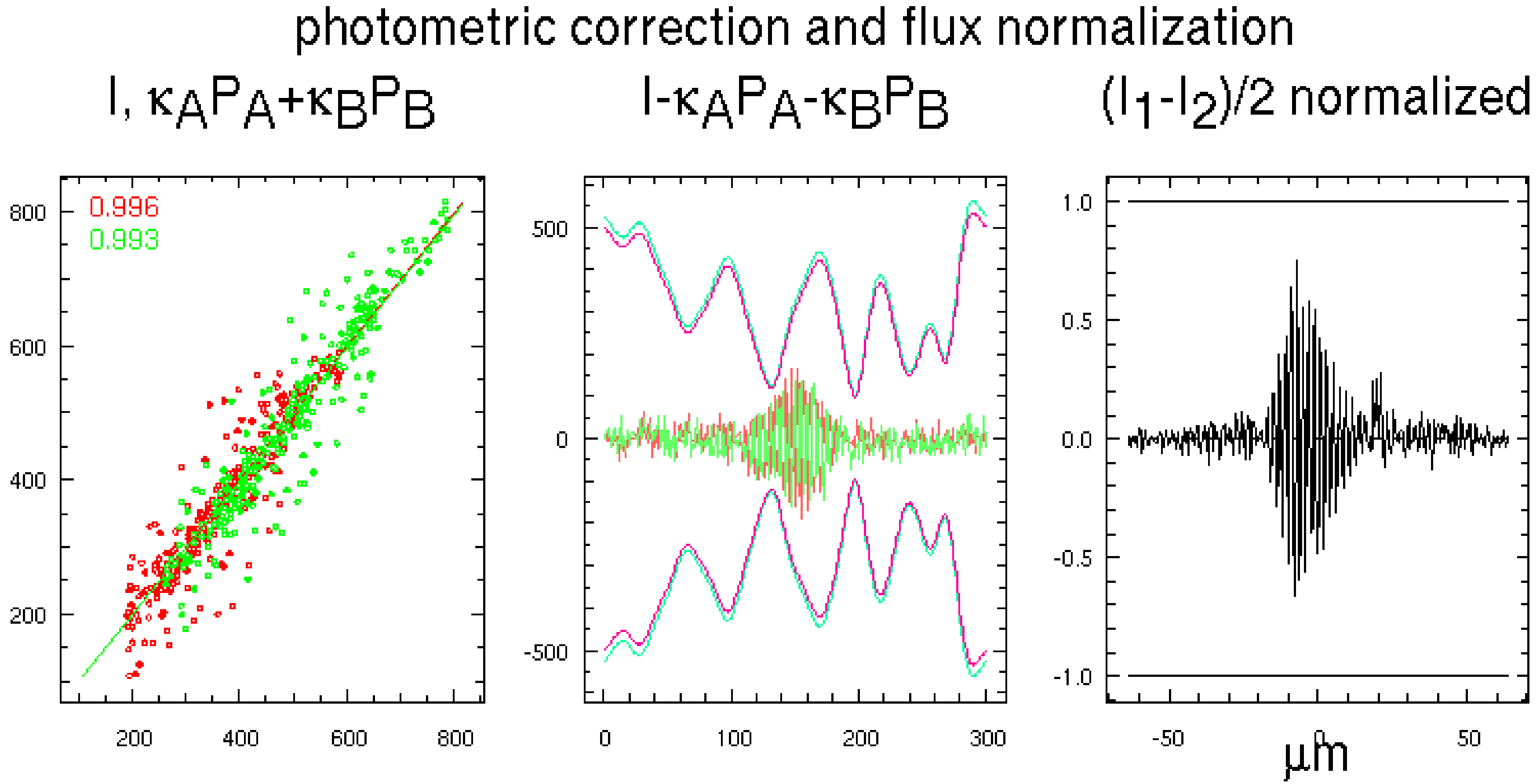
both I_1 and I_2 use PA, which has the same camera readout noise realisation

$$I_1 - \kappa_{1\mathrm{A}} P_\mathrm{A} - \kappa_{1\mathrm{B}} P_\mathrm{B}$$

$$I_1^c = \frac{I_1 - \kappa_{1\mathrm{A}} P_\mathrm{A} - \kappa_{1\mathrm{B}} P_\mathrm{B}}{2\left[\sqrt{\kappa_{1\mathrm{A}} P_\mathrm{A} \kappa_{1\mathrm{B}} P_\mathrm{B}}\right]_\mathrm{smoothed}}$$

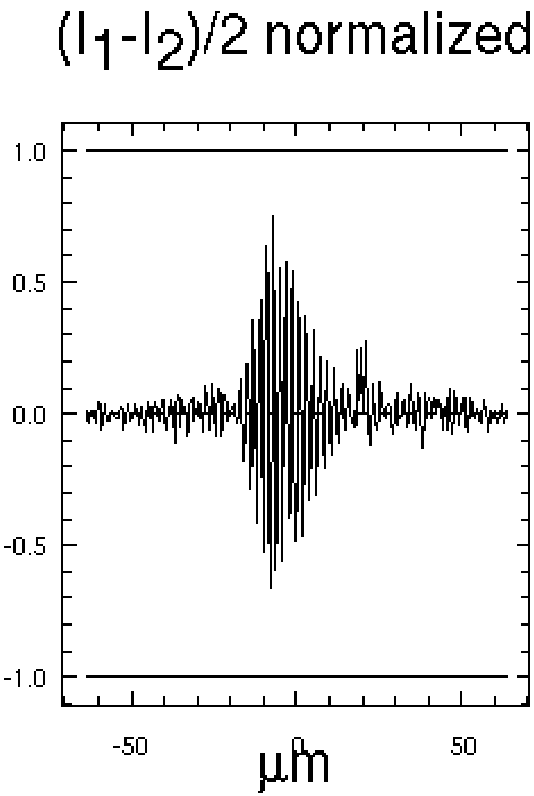
$$I^c = \frac{I_1^c - I_2^c}{2}$$

Photometric correction



FLUOR fringe 'packet'

- FLUOR fringe signal is localized in OPD
- this is because of the bandpass Δλ of the observing filter (the whole K band)



$$F(\delta, \lambda) = 1 + sin(2\pi\delta/\lambda)$$

$$\langle F(\delta) \rangle_{k=\frac{2\pi}{\lambda}} = 1 + \int_{k_{\min}}^{k_{\max}} T(k) \sin(k\delta) dk$$

Coherence length ~ fringe packet size:

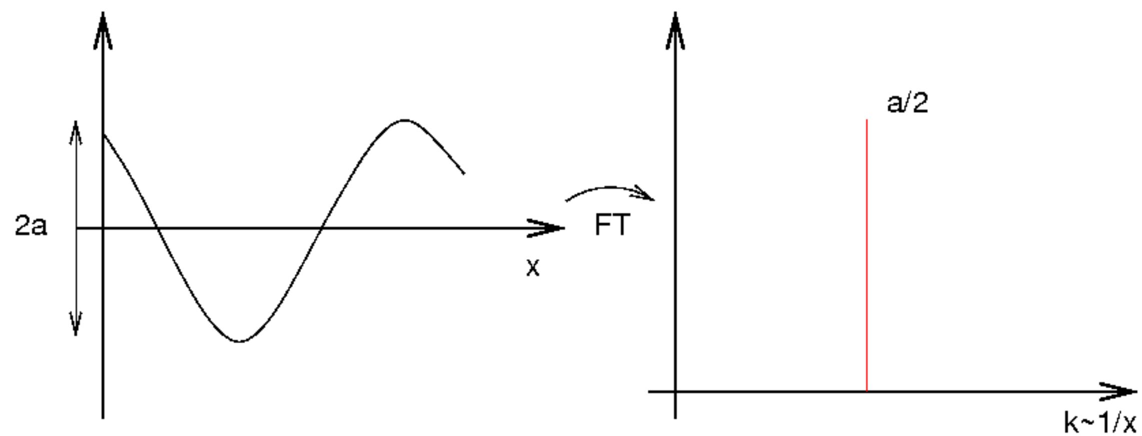
$$CL = 2\frac{\lambda^2}{\Delta\lambda} = 2\frac{2.2^2}{0.5^2} \approx 20\mu\text{m}$$

From fringe packets to Visibility

Again, this part is more specific to FLUOR, but retains some generality nonetheless

Fourier Analysis

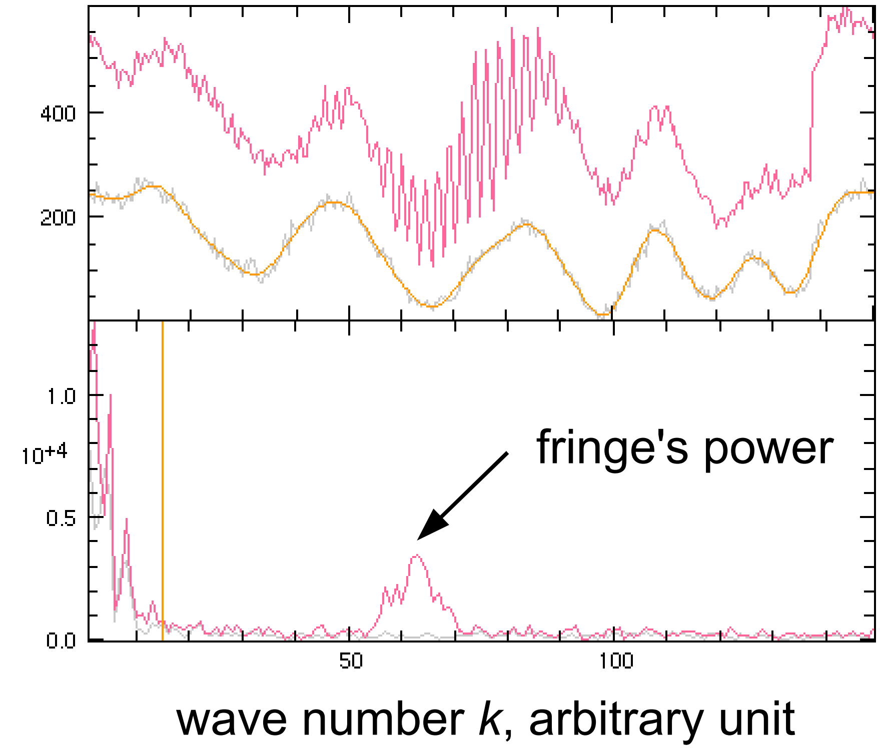
- Fringes are by definition sin waves
- The Fourier transform projects a sin wave in a single point in the reciprocal variable



It is natural to use Fourier Analysis to extract V^2

FFT of FLUOR Fringes

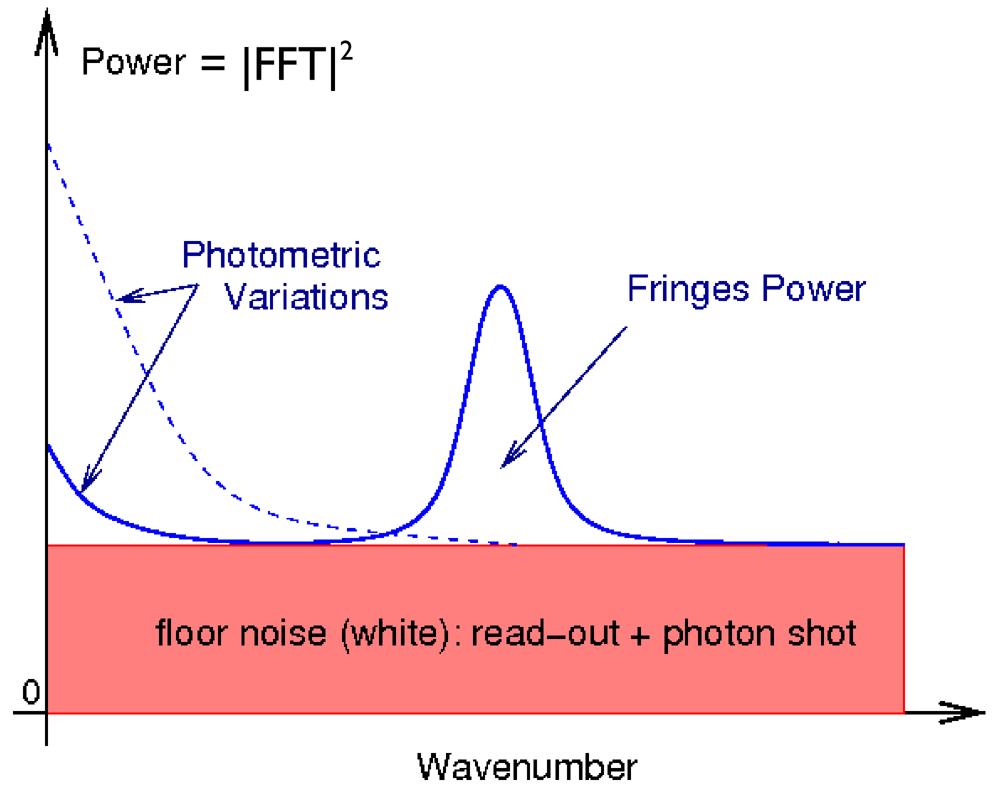
- Fringes frequency is chosen to be >> typical photometric variations
- Fringe sampling is 5samples / fringe
- the floor noise contains the camera readout-noise and the photon shot noise



$$
\begin{aligned}
F(\delta,\lambda) &= 1 + sin(2\pi\delta/\lambda) \\
\langle F(\delta)\rangle_{k=\frac{2\pi}{\lambda}} &= 1 + \int_{k_{\min}}^{k_{\max}} T(k)\sin(k\delta)dk
\end{aligned}
$$

Fourier Estimator

- Compute the integral of the fringes' power: $|V'|^2$
- bias is the floor noise, read-out and photon noise are white noises: $|n|^2$
- photometric variations are reduced thanks to the photometric correction: do not contaminate fringes



$$\mu^2 = \langle |V'|^2 \rangle - \langle |n|^2 \rangle$$

Fourier Estimator

The actual power integral is (in wavenumber k):

spectrum transmission Visibility (uncalibrated)

$$\mu^2 = \frac{\int_0^\infty B(k)^2 T(k)^2 V(k)^2 dk}{\left(\int_0^\infty B(k)T(k)dk\right)^2}$$

Photometric normalization

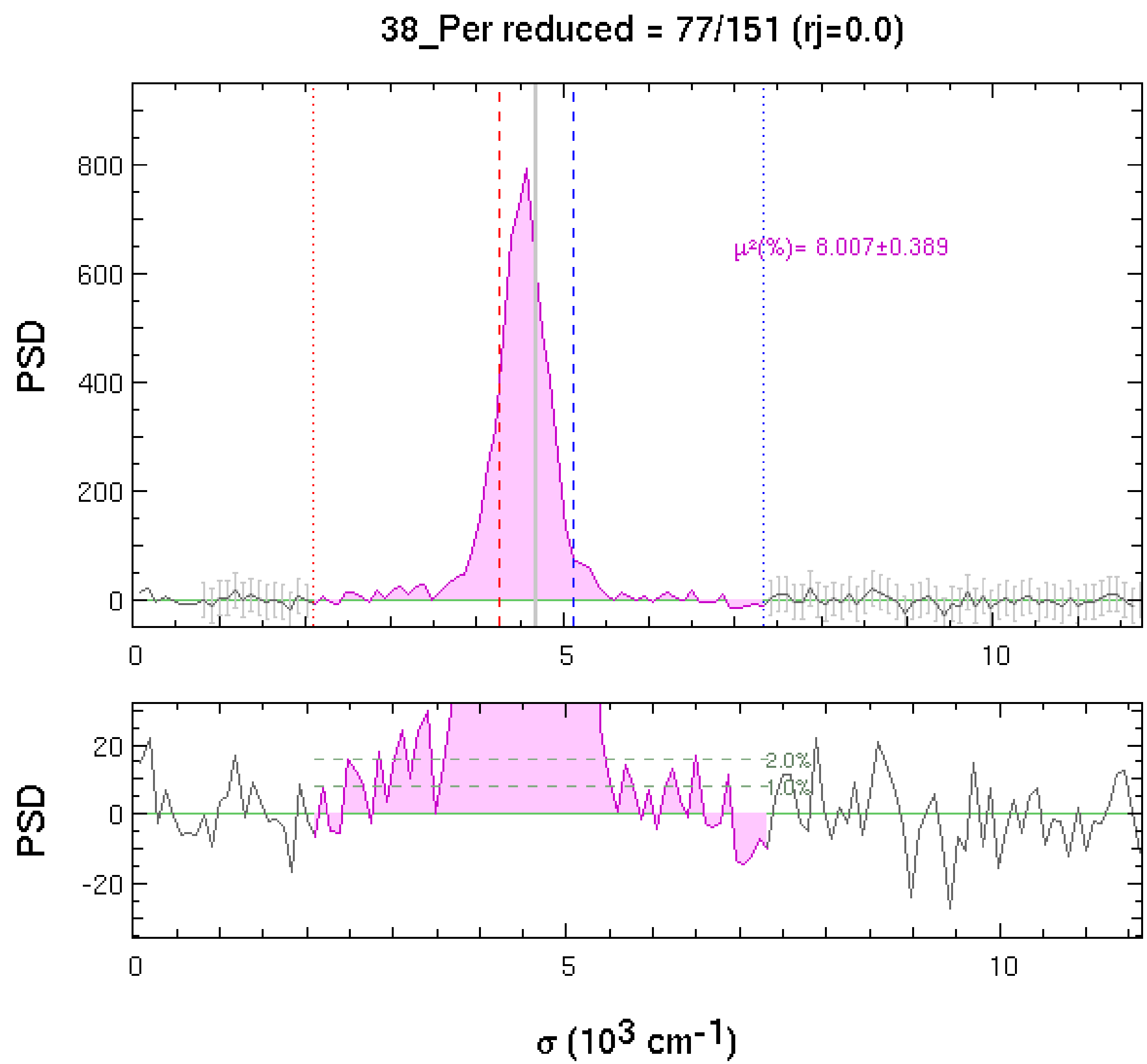
It turns out, because the bandpass is relatively small we have:

$$\mu^2 \quad \tilde{\propto} \quad \frac{\int_0^\infty B(k)^2 T(k)^2 V(k)^2 dk}{\int_0^\infty B(k)^2 T(k)^2 dk}$$

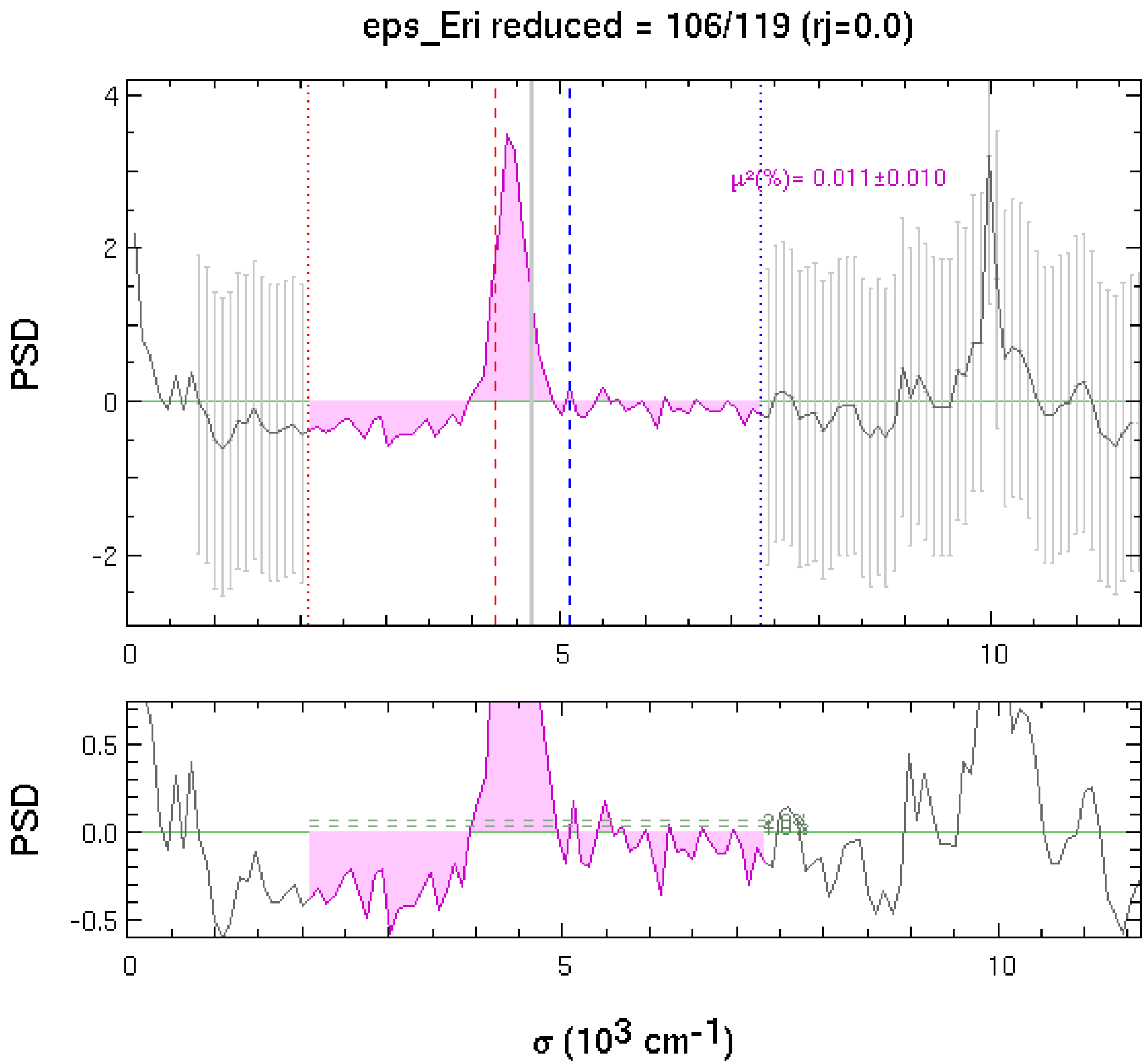
$$\mu^2 \quad \tilde{\propto} \quad \frac{\int_0^\infty B(k)^2 T(k)^2 V(k)^2 dk}{\int_0^\infty B(k)^2 T(k)^2 V(B=0,k)^2 dk}$$

FLUOR visibilities

- excellent photometric correction
- floor noise estimated outside the peak and substracted
- noise correction is very good (flat)



What can go wrong?

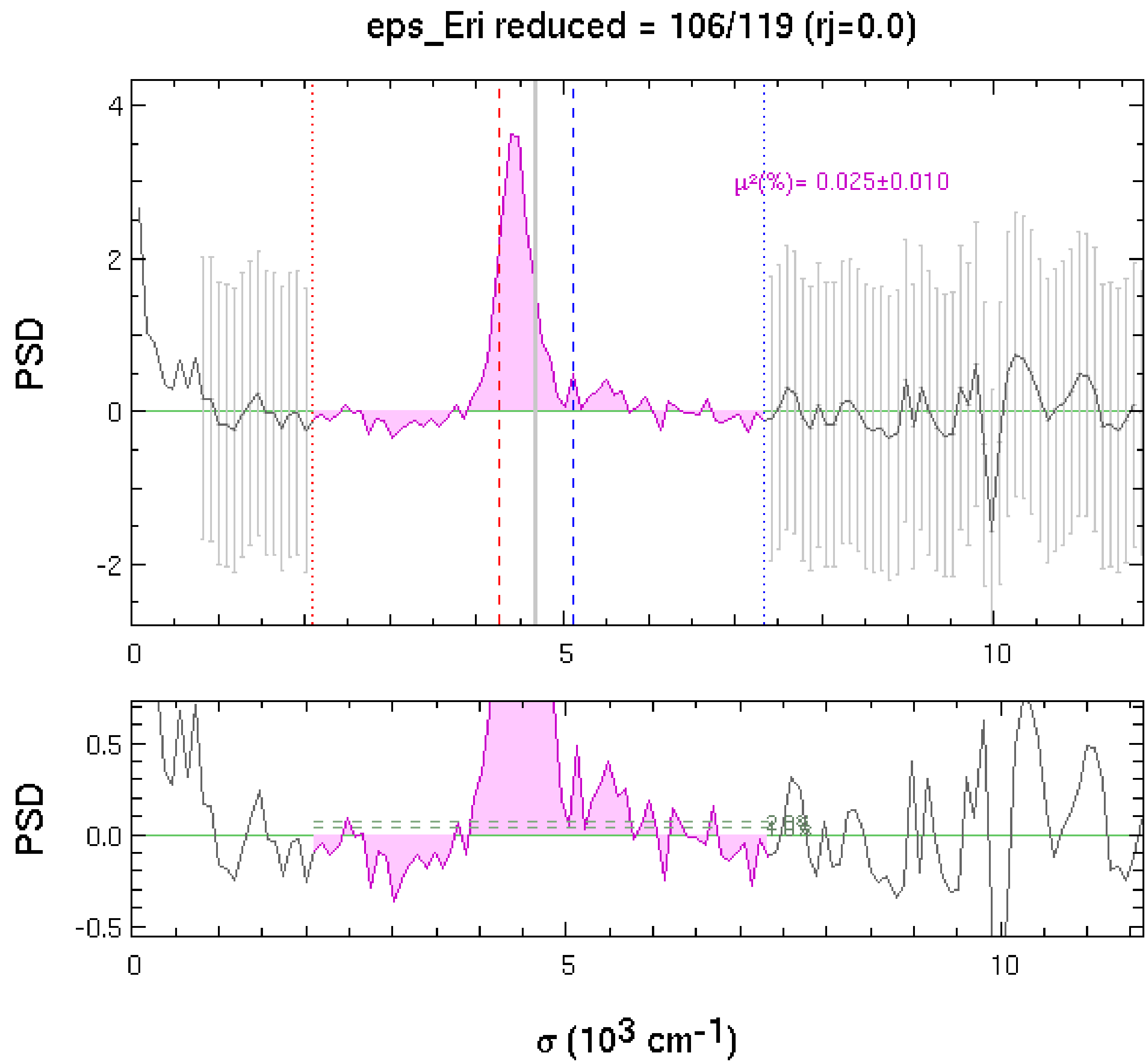


Readout noise is not white...

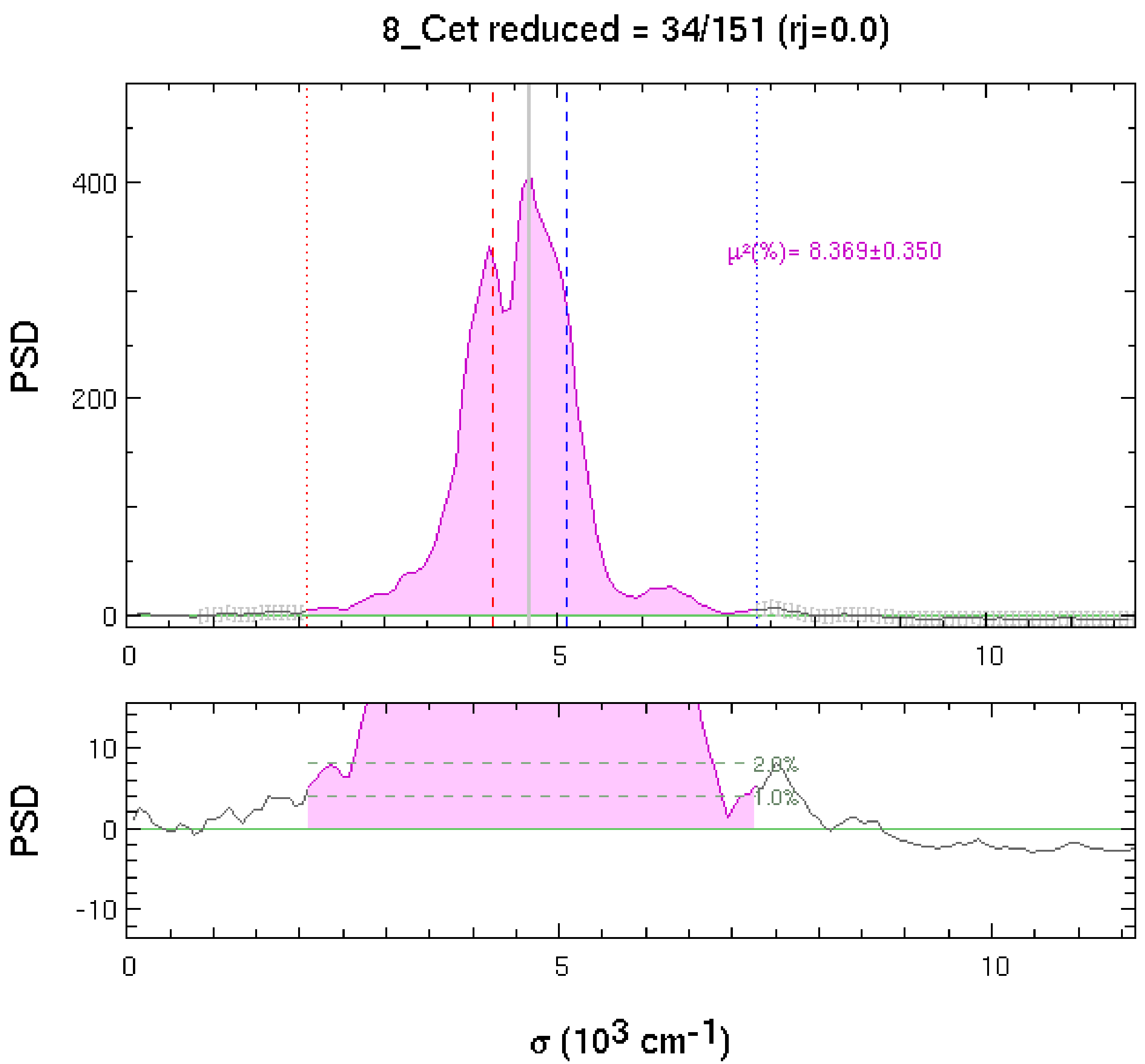
Correction introduced a bias

Solution?

- Use the darks taken during data acquisition
- Use the same algorithm (phot correction)
- Remove it from the PSD
- Looks nicer...



What else can go wrong?



OPD jitter... (piston)

Remember the jitter?

$$I(\delta(t)) = 1 + \mathrm{Re}\left(Ve^{-2\pi i\delta(t)/\lambda - 2\pi i \mathrm{OPD}_{\mathrm{jitter}}(\mathrm{t})/\lambda}\right)$$

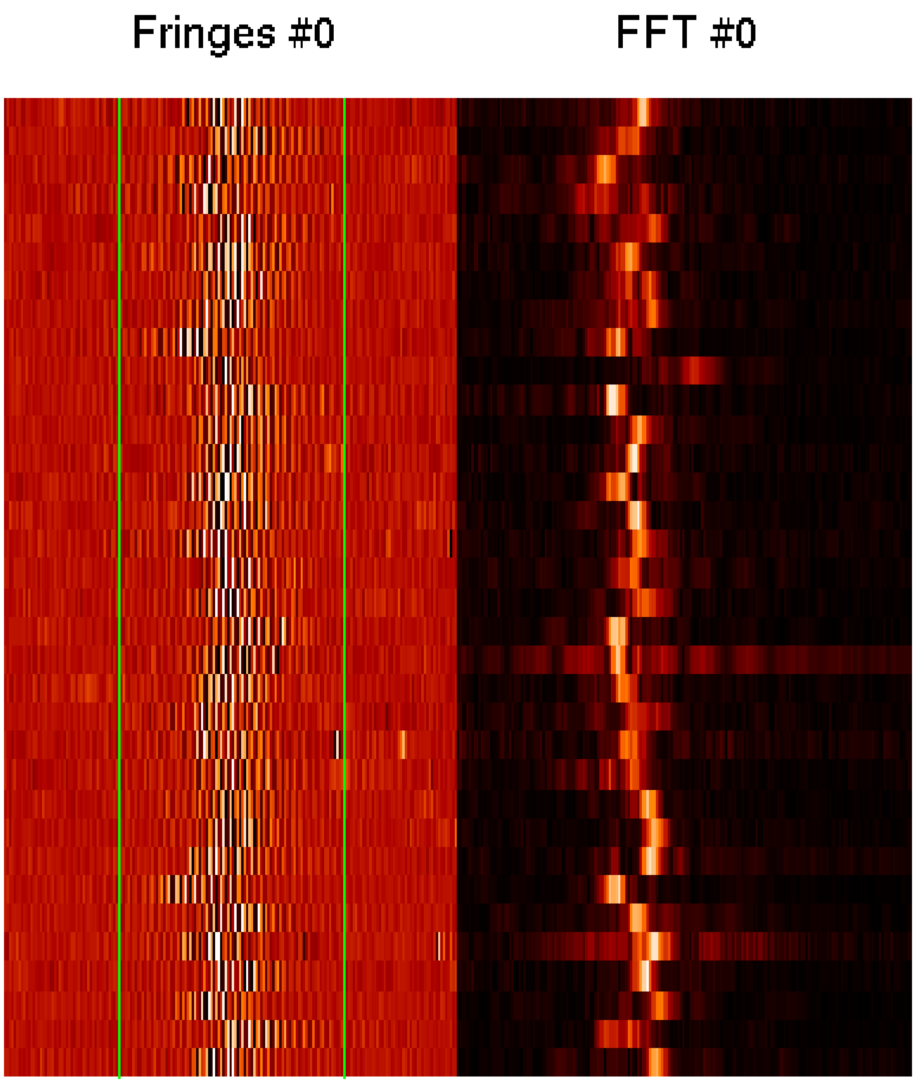
$$\delta(t) = tv_{\mathrm{mod}} \text{ (temporal modulation)}$$

$$I(\delta(t)) = 1 + \mathrm{Re}\left(Ve^{-2\pi i[tv_{\mathrm{mod}} - (p_0 + tp_1 + t^2 p_2 \ldots]/\lambda}\right)$$

- 0-order: the jitter introduces a phase. No phase measurement with 2 telescopes...
- 1-order: the jitter shifts the fringes frequency (time domain)

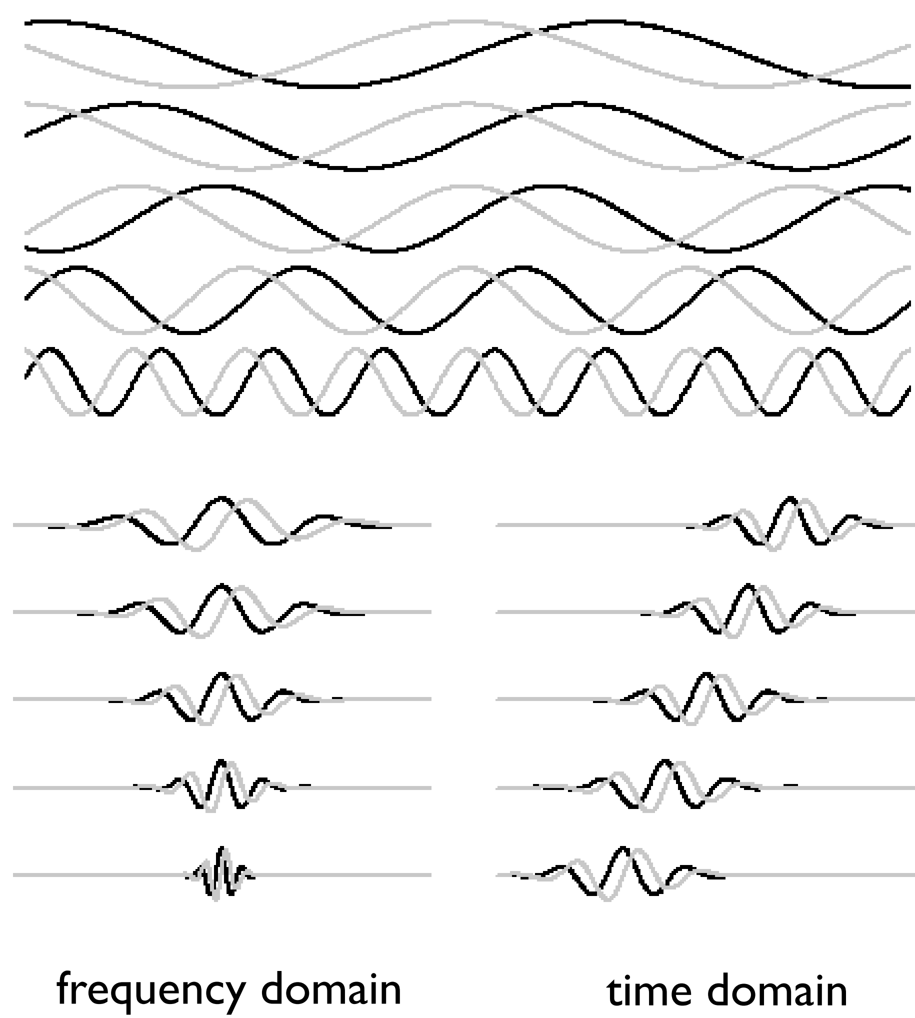
OPD jitter

- The fringes have been re-centered during the observations (coherencing)
- The jitter is clearly visible in the Fourier domain...
- Integration range should be larger



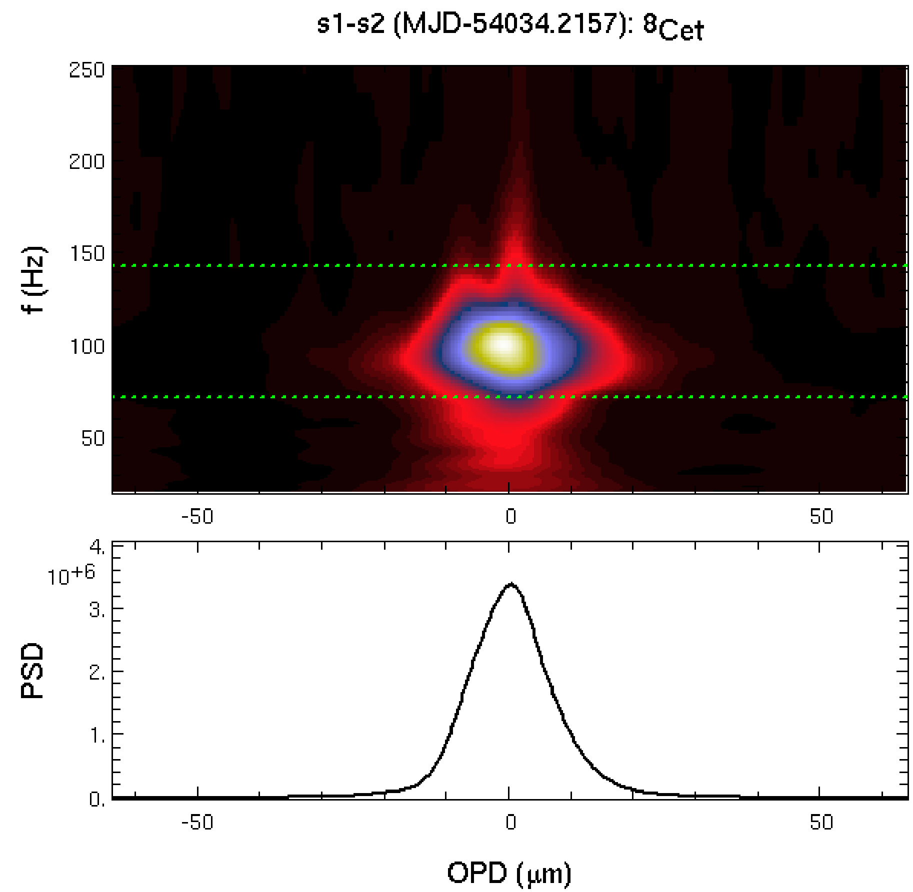
Wavelet Transform (WT)

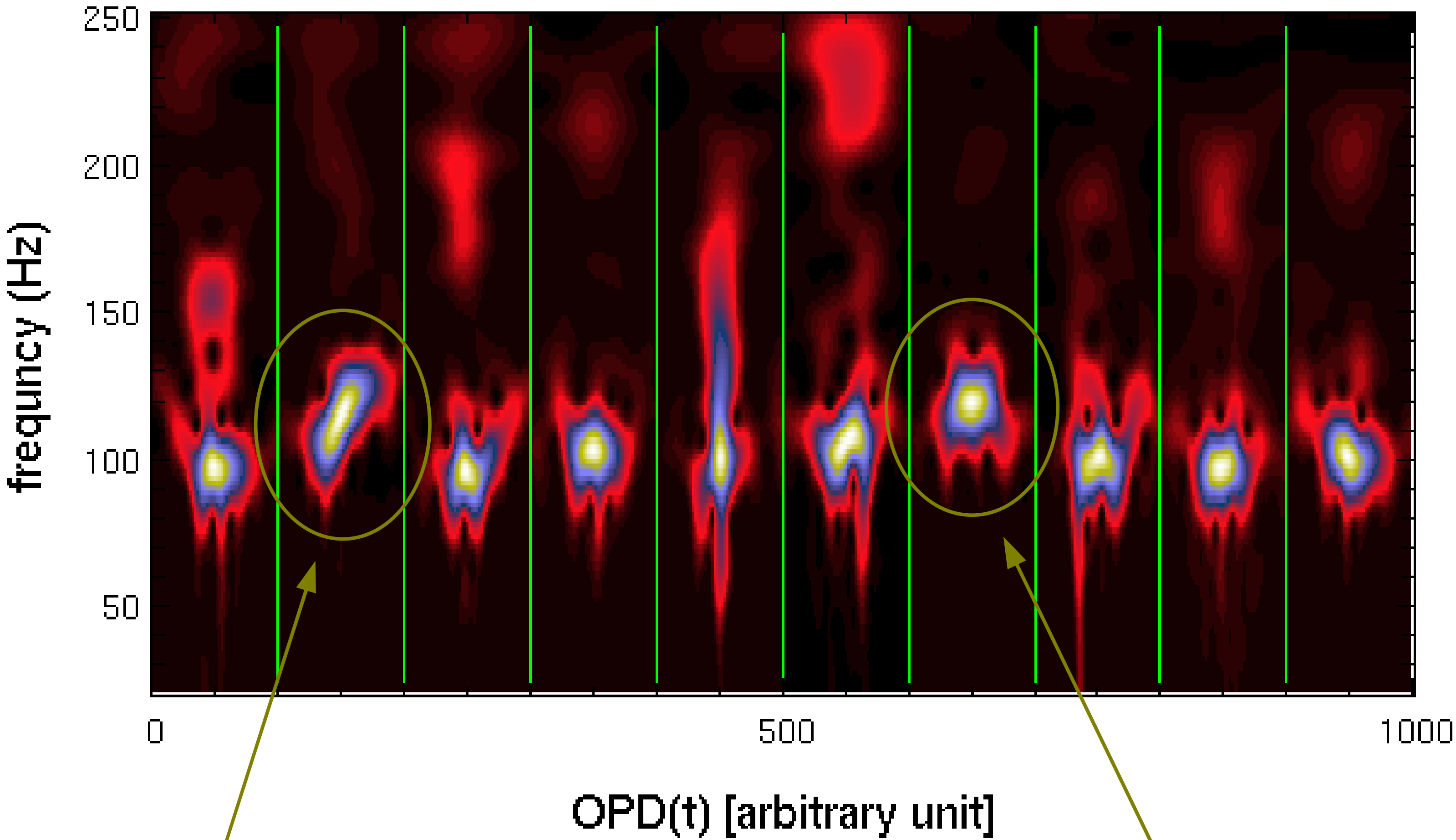
- Fourier transform uses a base of sin/cos waves of various **frequencies**: 1D
- Wavelet transform uses a base of **localized** waves with various **frequencies**: 2D



WT of a fringe packet

- The power is localized in frequency and OPD
- Formalism of estimator is equivalent to Fourier
- Advantages: allow to estimate the floor noise at the same frequency as the fringes!





second order effect:
the fringes slide during
the acquisition

first order effect:
the fringes peak shifts
in frequency

$$\delta(t) = tv_{\text{mod}} \text{ (temporal modulation)}$$

$$I(\delta(t)) = 1 + \text{Re}\left(Ve^{-2\pi i[tv_{\text{mod}} - (p_0 + tp_1 + t^2 p_2 \ldots]/\lambda}\right)$$

Jitter first order correction

The jitter introduces, at the first order, a multiplicative bias:

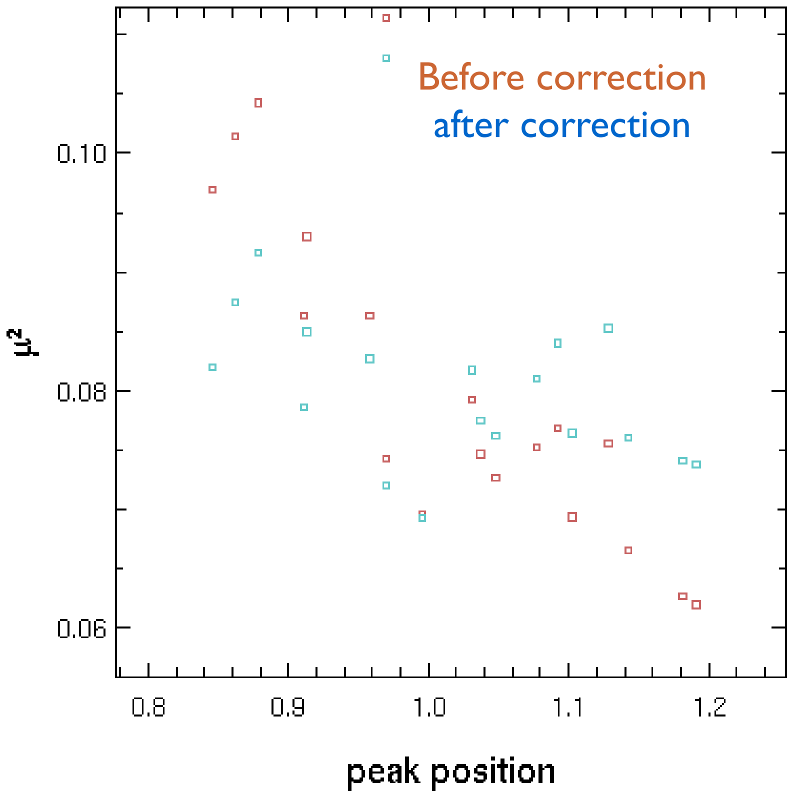
$$\mu'^2 \sim \int_0^\infty B(k)^2 T(k)^2 V(k)^2 dk'$$

$$k' = ak$$

$$\mu'^2 \sim \int_0^\infty B(k)^2 T(k)^2 V(k)^2 a dk$$

$$\mu'^2 \sim a \int_0^\infty B(k)^2 T(k)^2 V(k)^2 dk$$

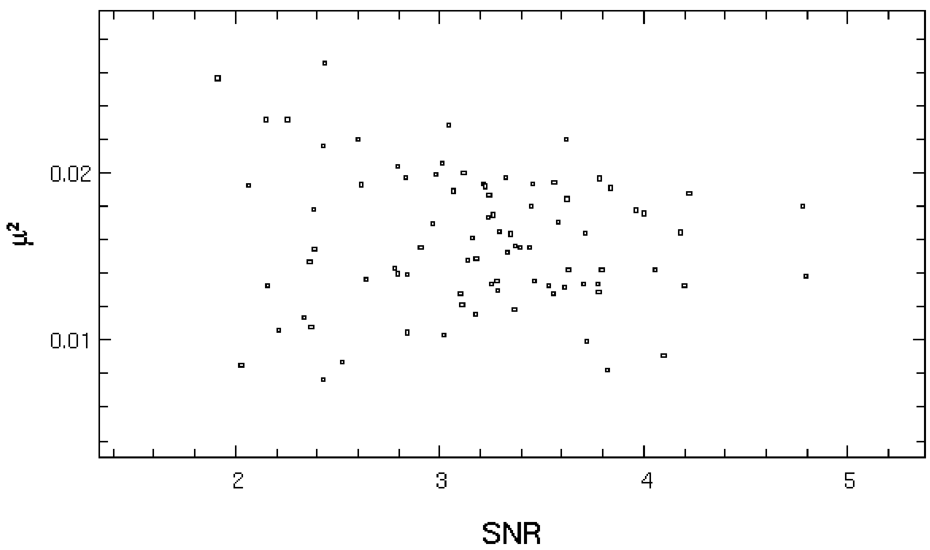
$$\mu'^2 \sim a\mu^2$$

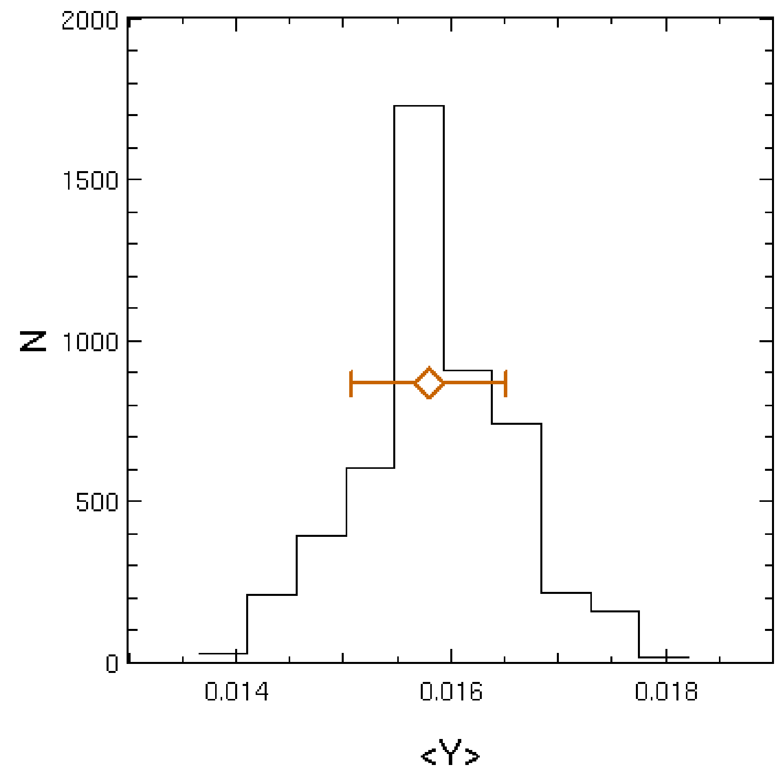


Can be linearly corrected using the PSD peak position

Final estimation

- A lot of interferograms are recorded
- Each interferogram is reduced and gives a μ^2
- The distribution as function photometric SNR is 'trumpet' like
- final values (and error bar) estimated using boot-strapping





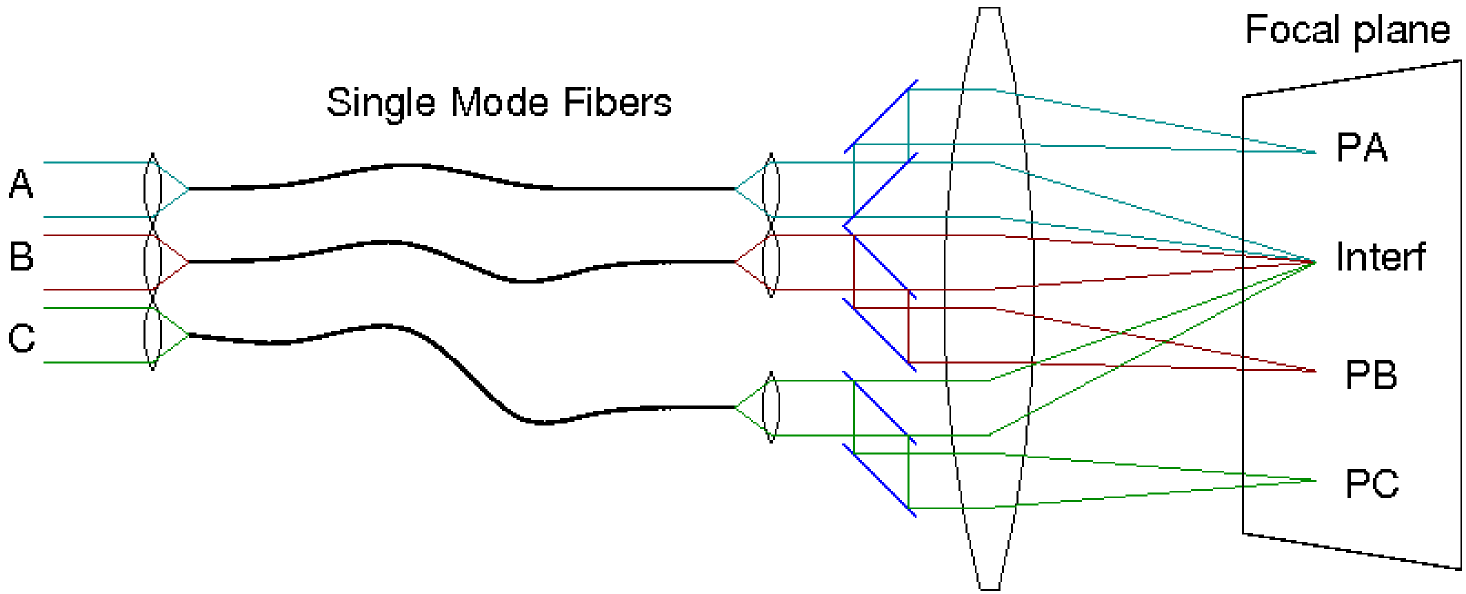
FLUOR 'take home' message

FLUOR observes a bandwidth-smeared squared visibility

What limits the precision?

- FLUOR scans fringes very quickly (1 fringe in 0.01s)
- under most conditions the jitter is 'frozen' so the first order effects dominate
- higher orders limit the ultimate precision to about a few 0.1% in high SNR conditions (typical is 1%).

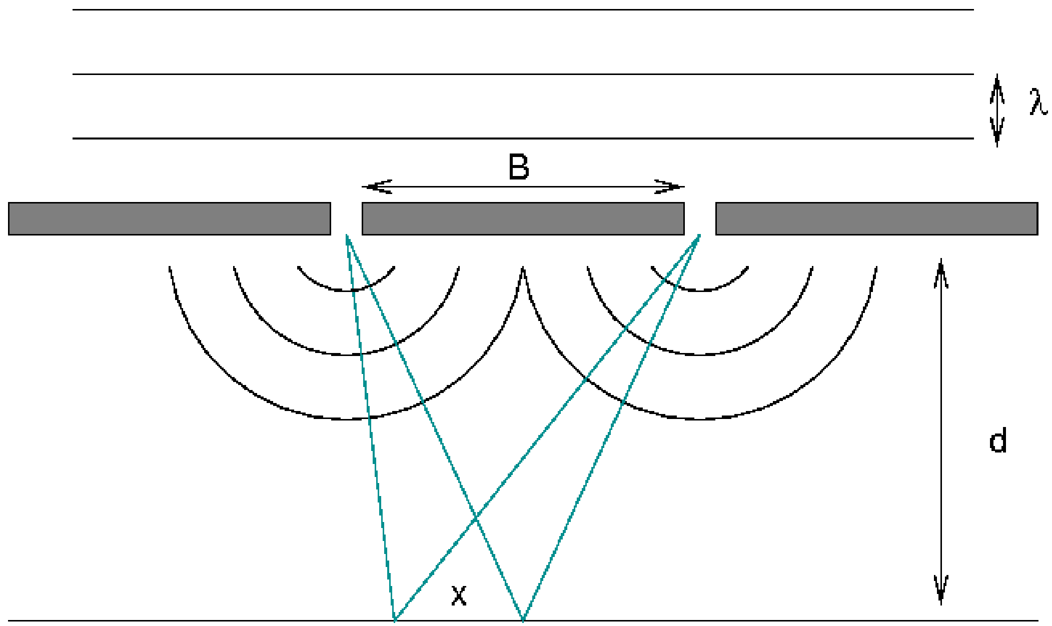
A little more complex: AMBER



- 3 Telescopes in JHK (1 - 2.5 microns)
- Fizeau (multi-axial) recombination of the 3 baselines
- Modal filtering and simultaneous photometric channels
- Spectral dispersion

Fringe coding

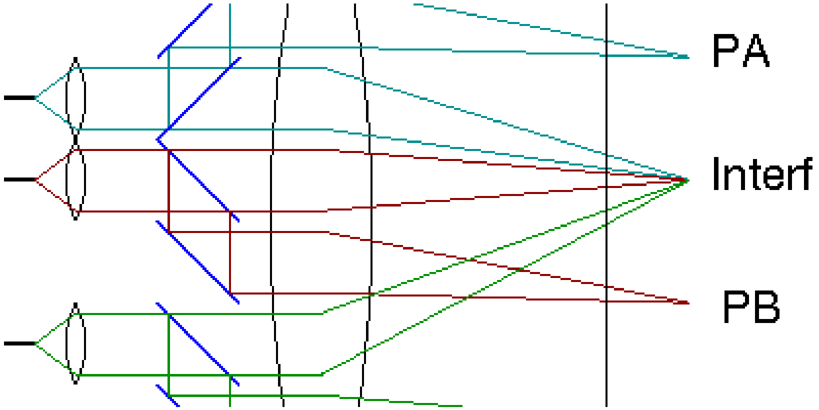
Similar to Young's slit interferometer:

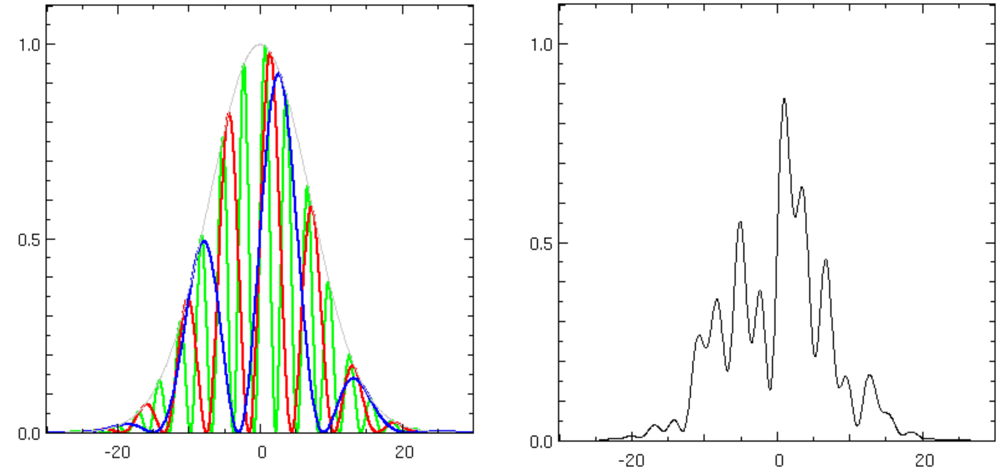


phase between the 2 points is $\Delta\phi \sim 2\pi \frac{xd}{B\lambda}$

AMBER fringe coding

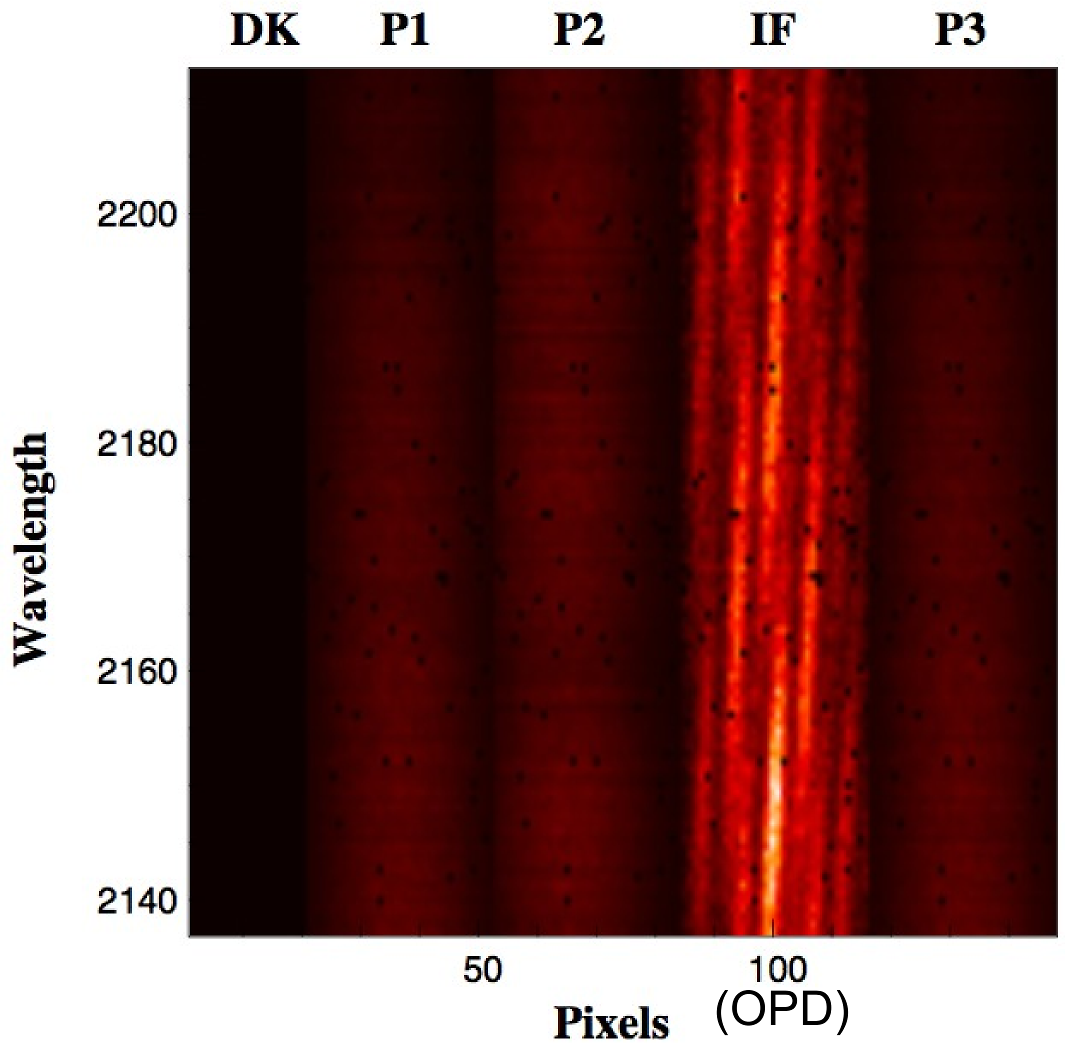
- The output fibers are arranged to have non redundant separations
- Each pair (baseline) has a unique fiber separation
- The envelope of the fringes is the diffraction pattern of the fiber





AMBER focal plane

- 2 dimensions are used
- horizontal dimension is the OPD
- vertical dimension is the spectral domain



Side note: ABCD

- Imaginge we observe only one fringe on 4 pixels
- V2 and phase estimate from linear products (X,Y,N)
- of course, it is biased but can be de-biased (Colavita 1999)

$$
\begin{aligned}
A, B, C, D &= 1 + \mathrm{Re}\left(V e^{2i\pi[0,1,2,3]/4 + i\phi}\right) \\
X &= A - C \\
Y &= B - D \\
N &= A + B + C + D \\
& \cos(a - \pi/4) - \cos(a + \pi/4) = 2\sin(\pi/4)\sin(a) \\
X, Y &= 2V \sin(\pi/4)\sin(\phi[+,-]\pi/4) \\
V^2 &= \frac{\pi^2}{2}\frac{X^2 + Y^2}{N^2} \\
\phi &= \tan^{-1}\frac{Y}{X}
\end{aligned}
$$

ABCD

- The ABCD is actually a FFT (Fast Fourier Transform)
- It is highly optimized
  - uses the minimum amount of information
  - it returns all the observables (V^2 and phase)
- However
  - it requires to stay on the central fringe (fringe packet envelope will bias the visibility)
  - it requires the ABCD to be $\lambda/4$ apart

ABCD vs. Fourier

Fourier

- has no a priori
- requires extra data to be able to integrate past the peak
- is not optimized: a fringe packet is not a sin wave

Robust

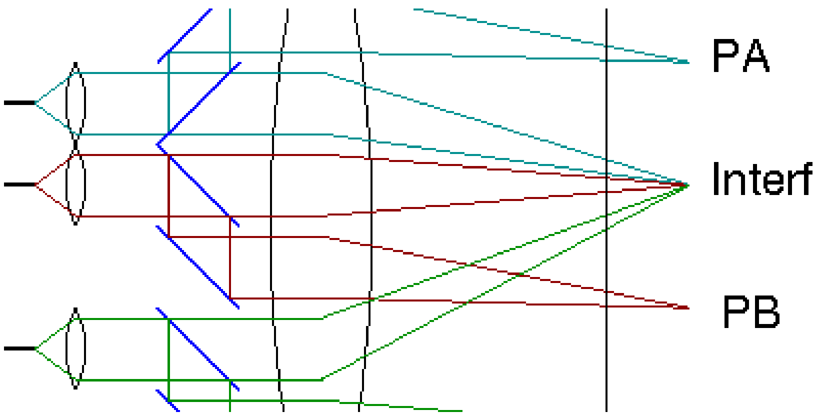
ABCD

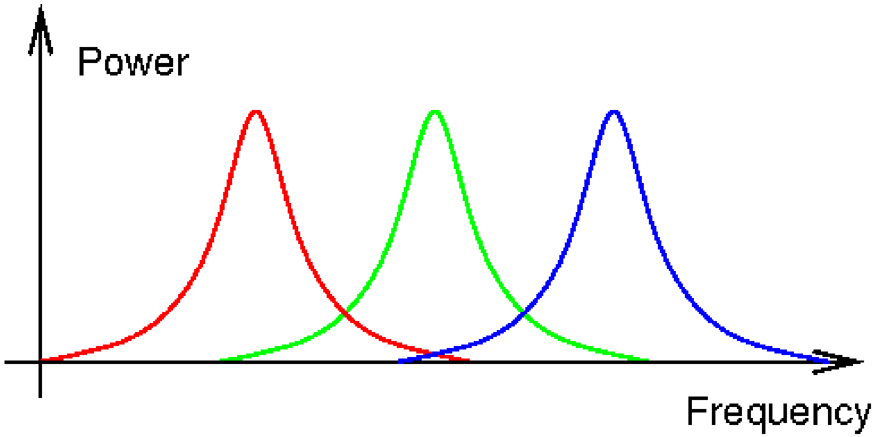
- biased by envelope effects
- require an a priori on the fringe signal
- optimized: a fringe is a sin wave

Optimized

AMBER: robust or optimized?

- In the Fourier space, the 3 peaks are superimposed...
- If Fourier is used, there will be some cross talk
- AMBER Fringe spacing is set by fibers' separation
- But the jitter will NOT affect the fringe peak





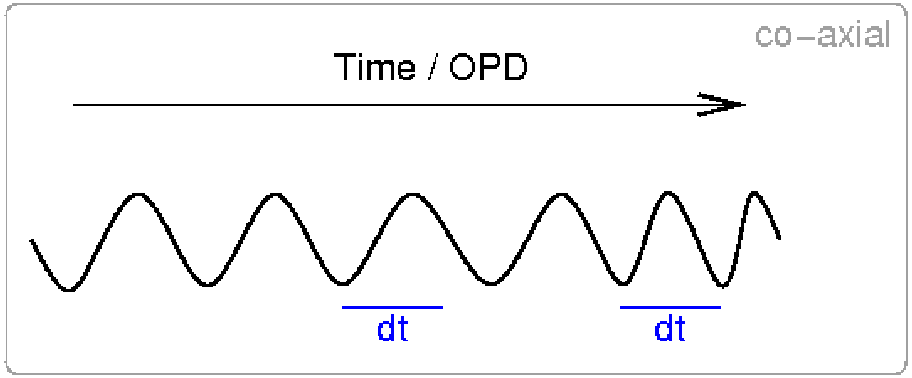
Co-axial / Multi-axial

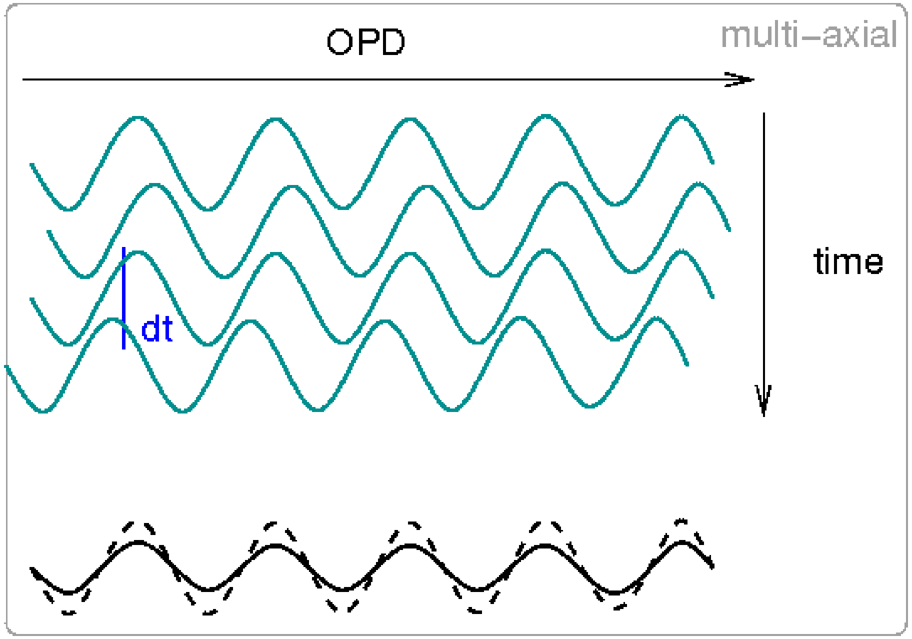
in a co-axial interferometer, time and OPD are mixed

> first orders of jitter acts like a accordion

in a multi axial interferometer, time and OPD are separate

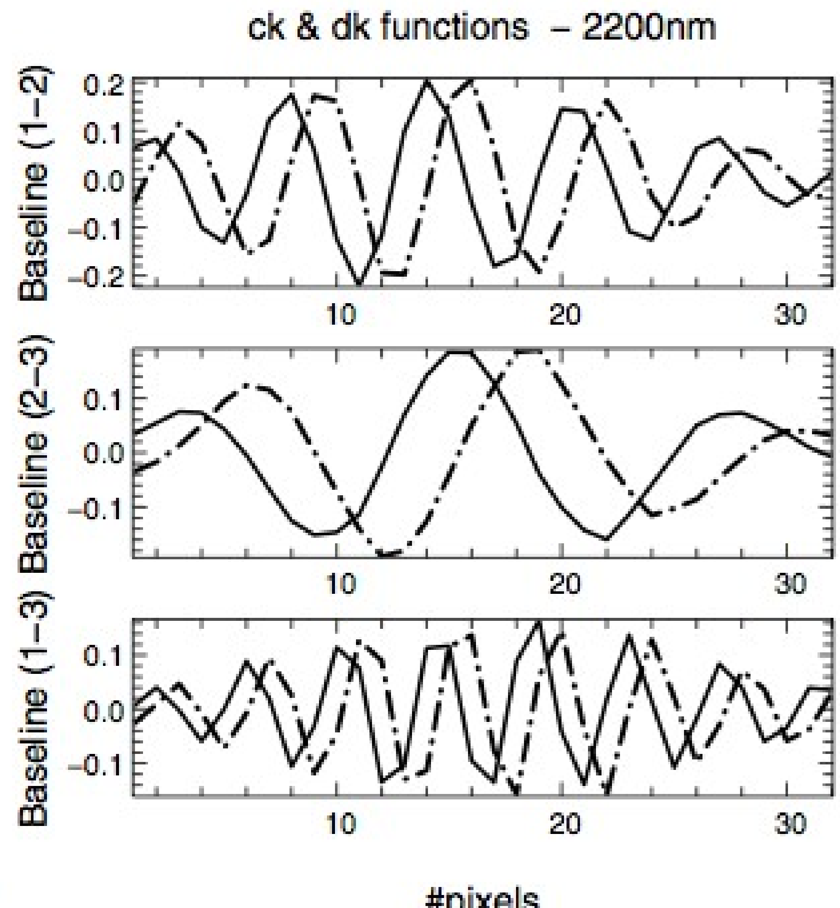
> first orders of jitter reduces the contrast





Processing for AMBER

- use a generalization of the ABCD algo
- uses more than 4 pix...
- 'X', 'Y' and 'N' like variables are obtained linearly using a matrix **(Pixel to Visibility Matrix)**
- knowledge of the interferogram is required



lamp fringes obtain before each series of observations for AMBER (0 and π/4 phases)

AMBER Pixel to Visibility Matrix

The P2VM varies

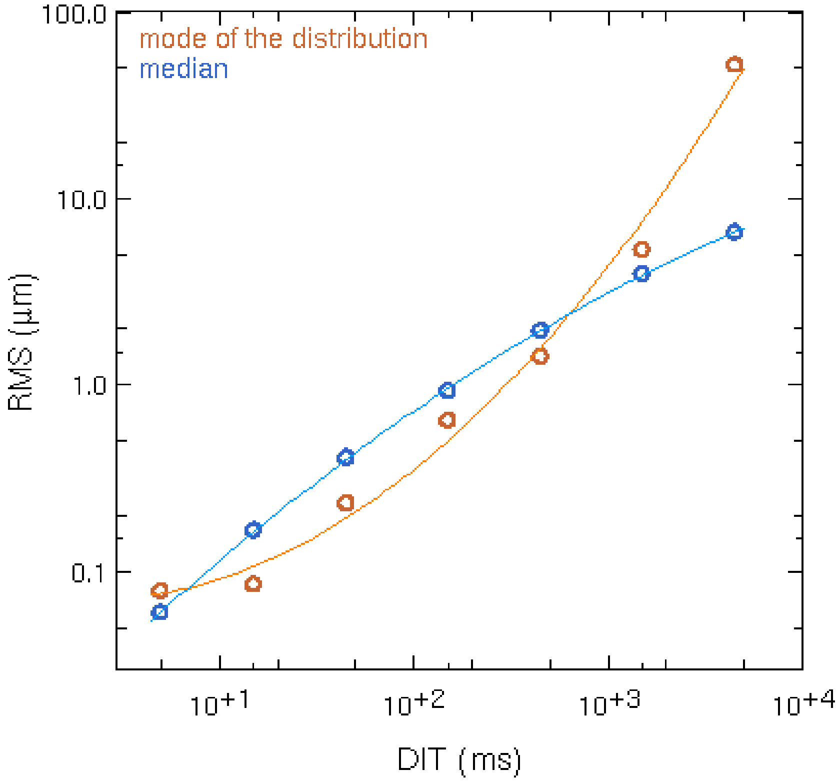
- position of the grism (not repeatable)
- alignment of the fibers
  - AMBER has 3x3 fibers: 3 bands treated separately
  - lots of optics between the fibers and the spectro...

a P2VM has to be taken for each change of setup

- a set of data can only be reduced with the P2VM taken right before
- P2VM is valid for ~1/2 night

AMBER Frame selection

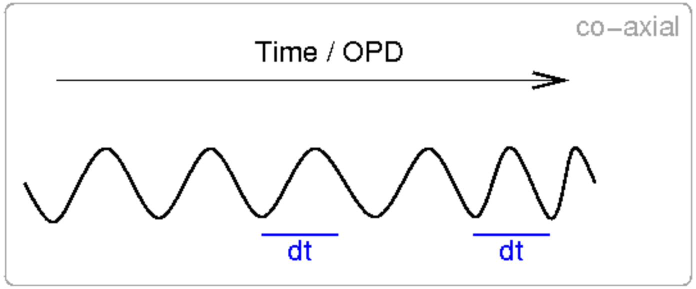
- The exposure times are small, to freeze the atmospheric jitter
- Each short exposure 'frame' leads to visibilities / phases.
- AMBER is strongly affected if the jitter is not frozen

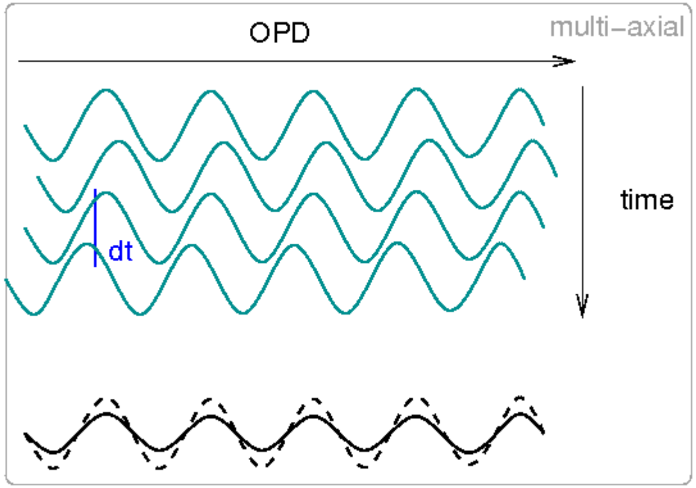


Jitter average RMS for different DITs, estimated from FINITO phase measurements *(not particularly a good night)*

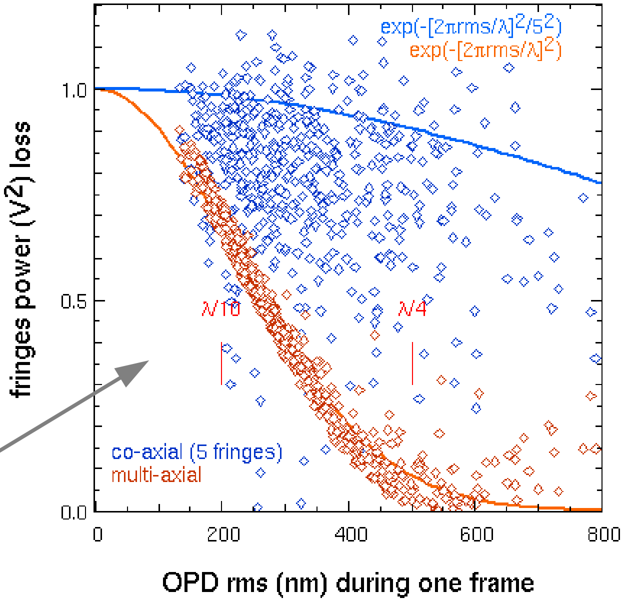
Effect of the jitter on AMBER data

- Multi-axial is less robust to jitter...
- AMBER was designed to sit and integrate behind a perfect fringe tracker

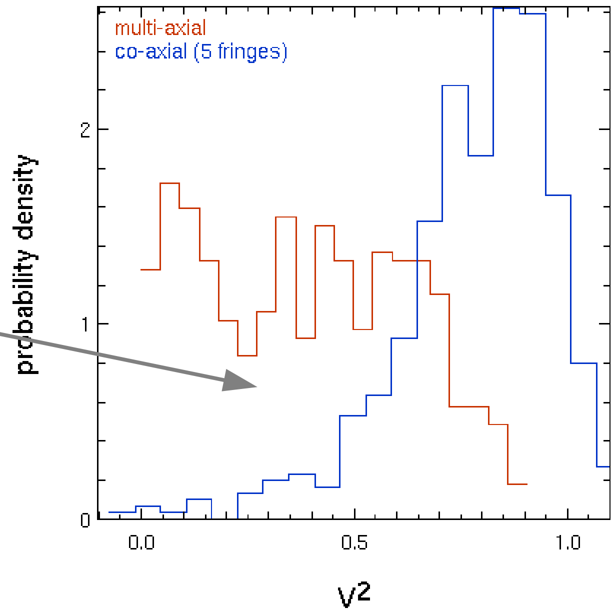




The jitter produces a drastic contrast loss in AMBER (muti-axial)

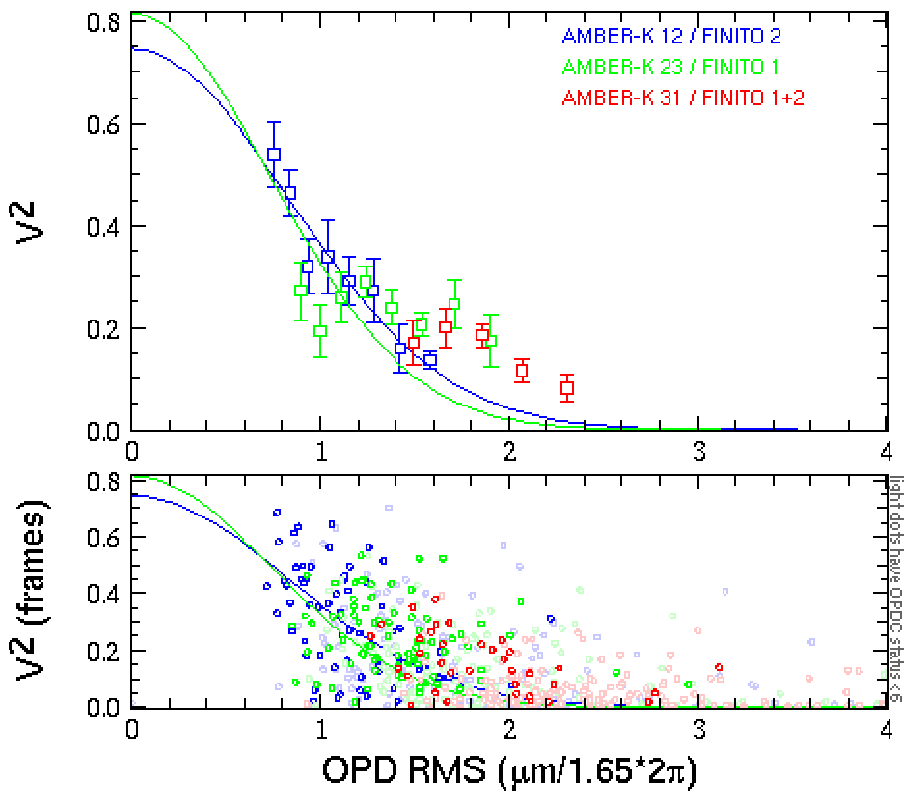


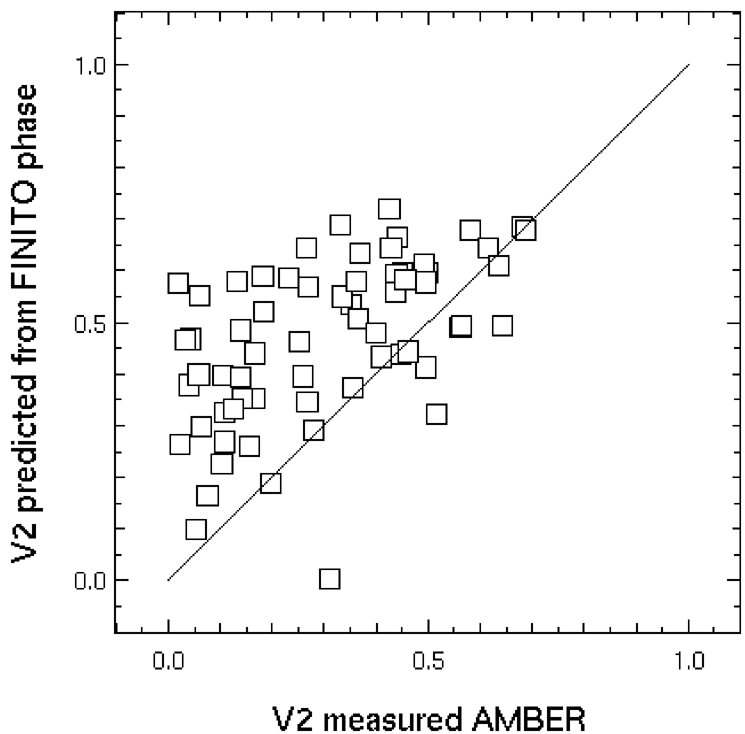
Because the distribution has no mode, averaging all the data does not work very well



What we should be doing

- If one knew the jitter durning the AMBER exposure, one could debias the visibility
- Using phases measured by FINITO (every 1ms), we can do that
- FINITO data recorder is a mode under commissioning...





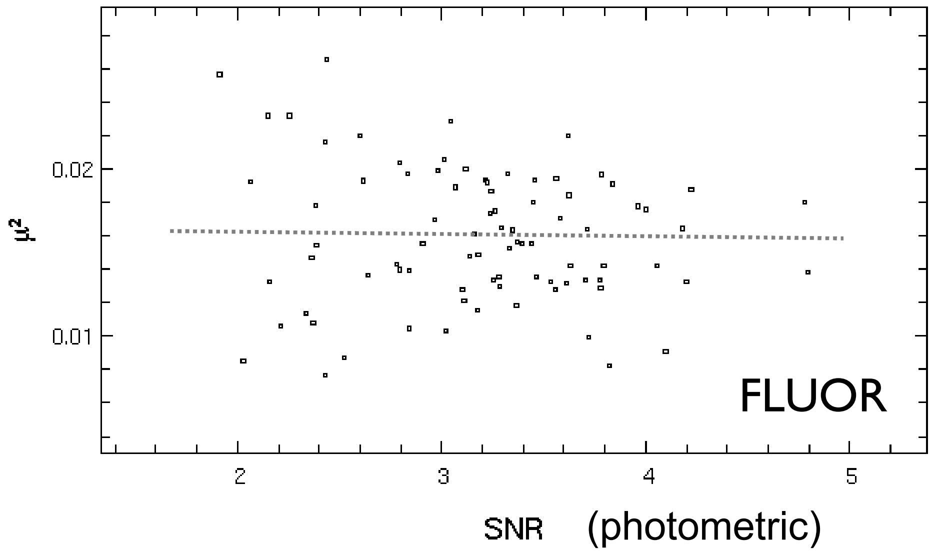
What people have been doing

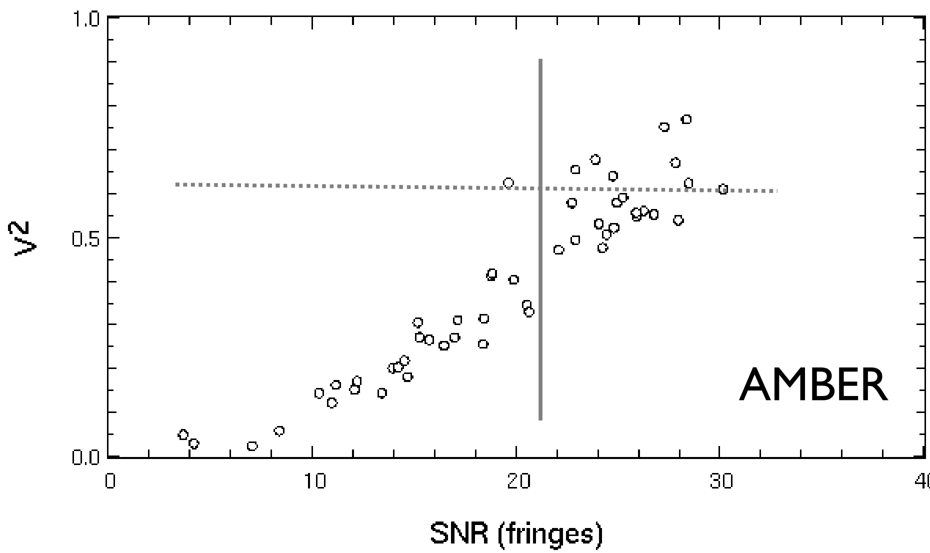
- **Select frames** using signal to noise ratio (SNR)
- WARNING: SNR defined in AMBER is the **fringes SNR = (FV)²**

  dominant noise is jitter

- 'Trumpet' diagram for FLUOR are ploted wrt **photometric SNR =F²**

  dominant noise is photon and readout





Why it 'kind of' works

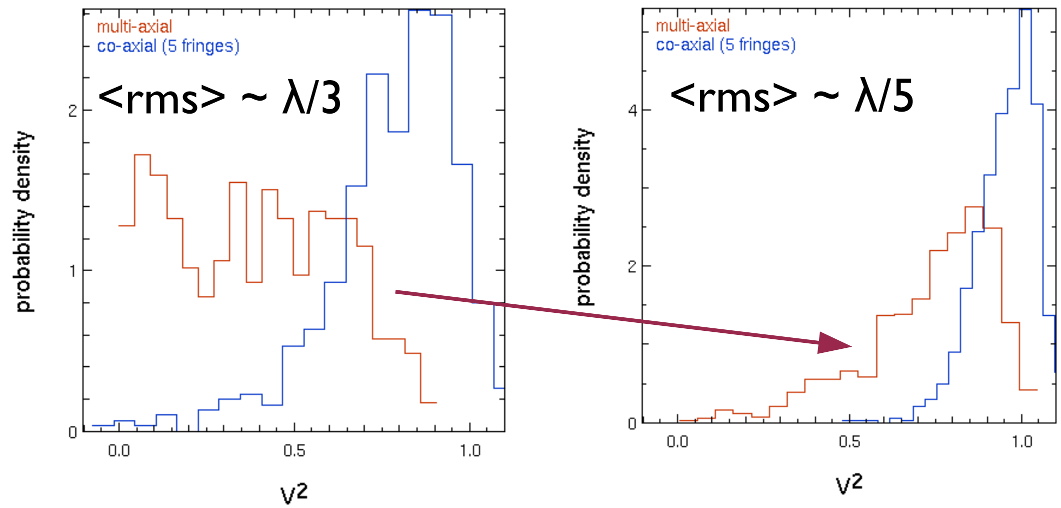
- 2 regimes in the V2 bias function

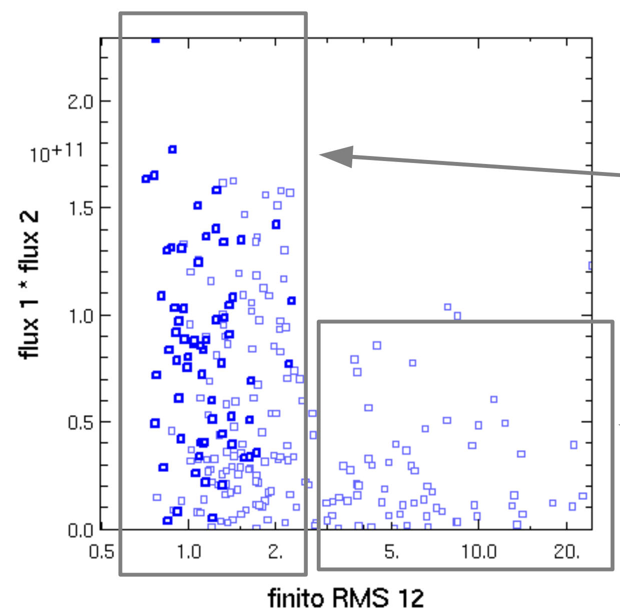
  small jitter leads to a peaked distribution

- seeing and jitter are somewhat correlated

  bad seeing decreases coupling efficiency in the fibers

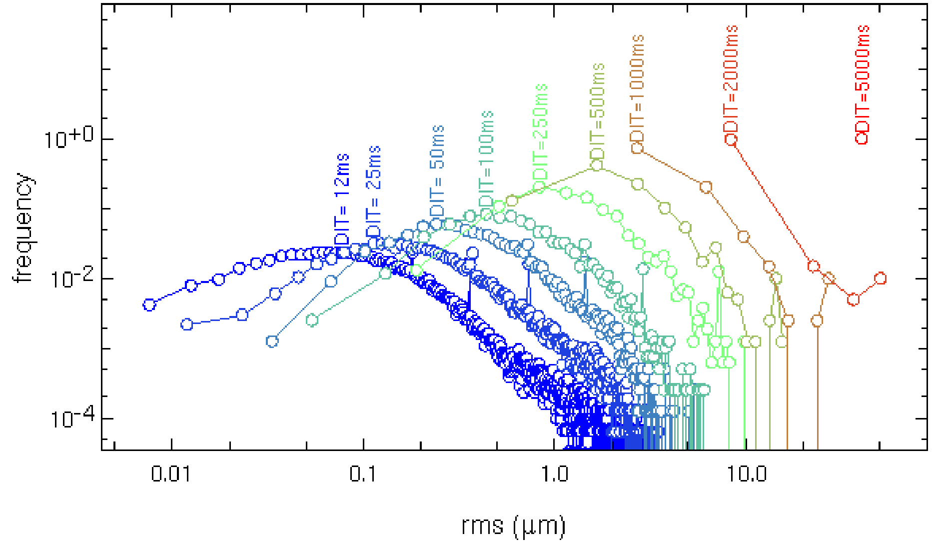
idea: **frames with higher fringes SNR had better conditions.** not perfect, but better than nothing...

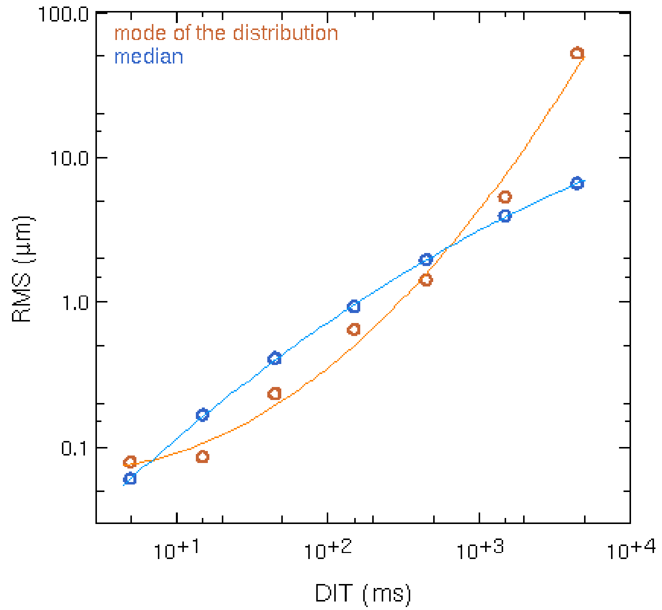




Interferometric coherence time

- AMBER gives exploitable fringes for a jitter < λ/5
- **jitter coherence time** ~ ½ frames have a RMS< λ/5
- For this particular night (not good), it is 40ms in K Band, 30ms in H Band





the jitter is here the jitter from the atmosphere

AMBER 'Take Home' message

- AMBER is multi axial: has some limitations
- P2VM: main element of the calibration plan
- visibilities unbiased by the jitter are very hard to measure: minimum DIT is 25ms ~ coherence time of the atmosphere.
- FINITO helps to stabilize the jitter
- we did not talk about:
  - differential visibilities
  - differential phases and phase closures

Will be addressed in the second lecture on the biases

Conclusions

- Interferometric data reduction is heavy signal processing
- intrinsically biased methods...
  - use of unbiased estimators
  - use of stellar calibrators to measure the instrumental / atmospheric losses
- It is easy to forget to be critical after battling to obtain visibilities
- Use statistical tools and calibrators to check the self-consistency

Bibliography (selective)

ABCD (PTI):

- Colavita (1998)

FLUOR/VINCI:

- Foresto et al. (1997), Kervella et al. (2004), Mérand et al. (2006)

AMBER:

- Tatulli et al. (2007), Chelli et al. (2009)

MIDI:

- Jaffe (2004), Meisner (several paper)