Data Calibration

Robert LUCAS
IRAM mm school
November 2004

Introduction

- Each interferometer baseline provides a measurement of the source visibility at a given point in the u, v plane of spatial frequencies;
- The source brightness distribution can then be reconstructed by an appropriate Fourier Transform.
- In reality things are not so simple:
  - many electronic components will have variable gains both in amplitude and in phase
  - At millimeter wavelengths the atmospheric absorption and path length fluctuations will dominate the instrument imperfections in most cases.
  - As a consequence we have to calibrate visibilities before producing images.

Calibration Steps

- Radiometric Phase Correction
- Passband Calibration
- Phase Calibration
- Flux Calibration (Frédéric's lecture)
- Amplitude Calibration

Water vapor radiometry (1)

- A radiometry system is used to monitor the emission of water vapor at a suitable frequency in front of all antennas (dedicated instrument or astronomy receiver; water line at 22GHz or quasi continuum).
- The fluctuations in the path length can be of a few % of the total path length due to water vapor.

Water vapor radiometry (2)

- The fluctuations of the water emission are converted into fluctuations of path length by using an atmospheric model (see lecture by J.-M. Winters).
  In principle one could hope to correct for all phase fluctuations this way.
  However there are limitations due to:
  - radiometer stability and sensitivity,
  - uncertainty in the determination of the emission to path length conversion factor

  it is not yet possible to correct consistently for the variation of path length between the source and the calibrator.
- So far this method at IRAM/Plateau de Bure is used only to correct for on-source phase fluctuations.
- Its main effect is to remove the de-correlation effect due to short-term phase fluctuations, improving the precision of amplitude determination.

Definitions and formalism

- Observed (apparent) visibility: $\widetilde{V}_{ij}(\nu, t)$,
- True source visibility: $V_{ij}(\nu, t)$

$$\widetilde{V}_{ij}(t) = \mathcal{G}_{ij}(t) V_{ij}(t) + \epsilon_{ij}(t) + \eta_{ij}(t) \tag{1}$$

- $\mathcal{G}_{ij}(\nu, t)$: complex gains of each baseline
- $\eta_{ij}(t)$ is a noise term from thermal fluctuations in the receivers;
- $\epsilon_{ij}(t)$ is an offset term.
- Note that:
  - This assumes that the system is linear.
  - π phase switching applied on the first LO's is used for suppressing the offsets $\epsilon_{ij}(t)$; they are generally negligible and will not be considered any further.
  - This scalar description ignores polarization effects.

Antenna based gains

- Amplitude and phase distortions have different physical origins:

$$\mathcal{G}_{ij}(t) = g_i(t)g_j^*(t) = a_i(t)a_j(t)e^{i(\phi_i(t)-\phi_j(t))} \tag{2}$$

- We have split the gains into antenna based factors.
- The gains represent properties of the data acquisition chains which are in the analogue part of the system.
- The correlator itself is a digital machine and we assume it is perfectly working (including the clipping correction).
- Averaging in time or frequency may produce some baseline based effects.

Closure quantities

- Neglecting offsets and baseline-based factors, one sees that:

$$\tilde{\varphi}_{ij} + \tilde{\varphi}_{jk} - \tilde{\varphi}_{ik} = \varphi_{ij} + \varphi_{jk} - \varphi_{ik} \tag{3}$$

  The phase along a triangle is called a closure quantity; it is unaffected by the observing process if antenna-based.

- A similar relation for amplitudes is:

$$\frac{|\widetilde{V}_{ij}(t)||\widetilde{V}_{kl}(t)|}{|\widetilde{V}_{ik}(t)||\widetilde{V}_{jl}(t)|} = \frac{|V_{ij}(t)||V_{kl}(t)|}{|V_{ik}(t)||V_{jl}(t)|} \tag{4}$$

  There are $(N-1)(N-2)/2$ phase (and $N(N-3)/2$ amplitude) closure quantities. They may be used for imaging when phases are too unstable to be directly measured (VLBI, optics).

Computing antenna based gains (1)

- The baseline based gains can be determined by observing a point source (strong quasar); then $V_{ij}(t)$ should all be equal to S

$$\mathcal{G}_{ij}(t) = g_i(t) g_j^*(t) = \frac{\widetilde{V}_{ij}(t)}{S} \tag{5}$$

- N complex unknowns and $N(N-1)/2$ equations.
- Real quantities: there are $N(N-1)$ measured values (amplitudes and phases)
- Only $2N-1$ unknowns since one may add a phase factor to all complex antenna gains
- When $N > 2$ the system is over determined and may be solved by a method of least squares.
- Using antenna based gains implicitly takes into account the closure relationships.

Computing antenna based gains (2)

- we note $\widetilde{V}_{ij} = \tilde{A}_{ij} e^{i\tilde{\varphi}_{ij}}$. if we impose the condition $\sum_{j=1}^{N} \phi_j = 0$, one may show that:

$$\phi_i = \frac{1}{N} \sum_{j \neq i} \tilde{\varphi}_{ij} \tag{6}$$

- For the amplitudes, if we define, in order to get a linear system:

$$\gamma_i = \log g_i, \quad \tilde{\alpha}_{ij} = \log \tilde{A}_{ij} \tag{7}$$

then:

$$\gamma_i = \frac{1}{N-1} \sum_{j \neq i} \alpha_{ij} - \frac{1}{(N-1)(N-2)} \sum_{j \neq i} \sum_{k \neq i, > j} \alpha_{jk} \tag{8}$$

at least three antennas are needed. For three antennas it reduces to

$$g_1^2 = \frac{\tilde{A}_{12} \tilde{A}_{13}}{\tilde{A}_{23}} \tag{9}$$

These formulae assume that the same weight is given to all baselines.

- The precision to which the antenna phases and amplitudes are determined is improved by a factor $\sqrt{N}$ over the precision of the measurement of the baseline amplitudes and phases.

Gain corrections

- The determination of antenna-based gains (amplitudes and phases) has an obvious advantage: the physical cause of the gain variations are truly antenna based.
- One may solve for the gains at the time of the observations, and correct the occurring problems to improve the quality of the data.
- One may re-point or re-focus the antennas to correct for an amplitude loss, correct for an instrumental delay (affecting the frequency dependence of the phases) . . .

Bandpass calibration

- We have considered a monochromatic system.
- Actually we have finite bandwidths: in principle the gain coefficients are functions of both frequency and time.

$$\widetilde{V}_{ij}(\nu,t) = \mathcal{G}_{ij}(\nu,t)V_{ij}(\nu,t) = g_i(\nu,t)g_j^*(\nu,t)V_{ij}(\nu,t) \tag{10}$$

- We assume that the passband shape does not change with time:

$$\mathcal{G}_{ij}(\nu,t) = \mathcal{G}_{\mathrm{B}ij}(\nu)\mathcal{G}_{\mathrm{C}ij}(t) \tag{11}$$

- Same decomposition for the antenna-based gains:

$$g_i(\nu,t) = g_{\mathrm{B}i}(\nu)g_{\mathrm{C}i}(t) \tag{12}$$

- $g_{\mathrm{B}i}(\nu)$ is the antenna complex passband shape
- normalized so that its integral over the observed bandpass is unity;
- $g_{\mathrm{C}i}(t)$ describes the time variation of the complex gains.

Frequency dependence of the gains

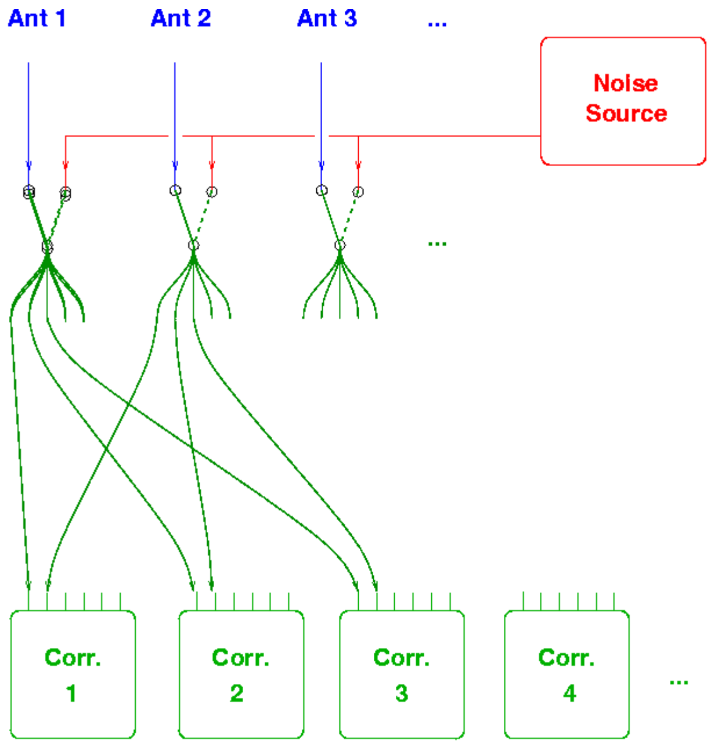
- In the correlator itself the anti-aliasing filters are very steep at the edges of each band; the phase slopes can be high there too.
- Any non-compensated delay offset in the IF can also be seen as a phase linearly dependent on frequency.
- The attenuation in the cables strongly depends on IF frequency, (this is normally compensated for, to first order)
- The receiver has a frequency dependent response both in amplitude and phase, due the IF amplifiers, and the frequency dependence on the mixer conversion loss.
- Antenna chromatism may also be important.
- The atmosphere itself may have some chromatic behavior, in the vicinity of a strong line (e.g. O_2 at 118 GHz) or if a weaker line (e.g. O_3) lies in the band.

Bandpass measurement

- Bandpass calibration usually relies on observing a very strong source for some time;
- the bandpass functions are obtained by normalizing the observed visibility spectra by their integral over frequency.
- It is *a priori* not necessary to observe a point source, as long as its visibility can be assumed to be, on all baselines, independent on frequency in the useful bandwidth.
- If there is some dependence on frequency, then one should take this into account.

IF passband calibration (1)

- the correlator can be split into independent sub-bands centered to different IF frequencies, observing different frequencies in the sky.
- they can be treated as different receivers (different anti-aliasing filters and different delay offsets)
- they need independent bandpass calibrations (can be done simultaneously on the same strong source).
- It is more practical to get a relative calibration of the sub-bands by switching the whole IF inputs to a noise source common to all antennas
- The switches are inserted before the IFs are split between sub-bands: the delay offsets of the sub-bands are also calibrated out.



IF Passband calibration setup

IF passband calibration (2)

- the SNR with the noise source is higher that for an astronomical source: it provides fully correlated signals to the correlator (delay tracking switched off)
- can be done quite often to suppress any gain drift due to thermal variations in the analogue part of the correlator.
- Since the sensitivity is high, this calibration is done by baseline, so that any closure errors are taken out.
- Only frequency dependent effects occurring in the signal path before the point where the noise source signal is inserted remain to be calibrated.
- Since at this point the signal is not yet split between sub-bands, the same passband functions are applicable to all correlator sub-bands.

At Plateau de Bure an 'IF passband calibration' is implemented. Precision in phase is $360/\sqrt{\Delta\nu\Delta t} = 0.5°$ at 100 kHz resolution.

RF bandpass calibration (1)

- To determine the functions $g_{\mathrm{B}i}(\nu)$ we observe a strong source, with a frequency-independent visibility.

$$\widetilde{V}_{ij}(\nu, t) = g_{\mathrm{B}i}(\nu) g_{\mathrm{B}j}^{*}(\nu) g_{\mathrm{C}i}(t) g_{\mathrm{C}i}^{*}(t) V_{ij}(t) \tag{13}$$

Then

$$g_{\mathrm{B}i}(\nu) g_{\mathrm{B}j}^{*}(\nu) = \frac{\widetilde{V}_{ij}(\nu, t)}{\int \widetilde{V}_{ij}(\nu, t) d\nu} \tag{14}$$

- One averages the measurements on a time long enough to get a decent signal-to-noise ratio.
- One solves for the antenna based coefficients in both amplitude and phase;
- then polynomial amplitude and phase passband curves are fitted to the data.

RF bandpass calibration accuracy (1)

- The passband calibrated visibility data will then be:

$$\widetilde{V}_{\mathrm{C}ij}(\nu, t) = \widetilde{V}_{ij}(\nu, t)/g_{\mathrm{B}i}(\nu)g_{\mathrm{B}j}^{*}(\nu) \quad (15)$$

  the amplitude and phase of which should be flat functions of frequency.

- Phase accuracy:
  It sets the uncertainty for relative positions of spectral features in the map. A rule of thumb is:

$$\Delta\theta/\theta_B = \Delta\phi/360 \quad (16)$$

  where

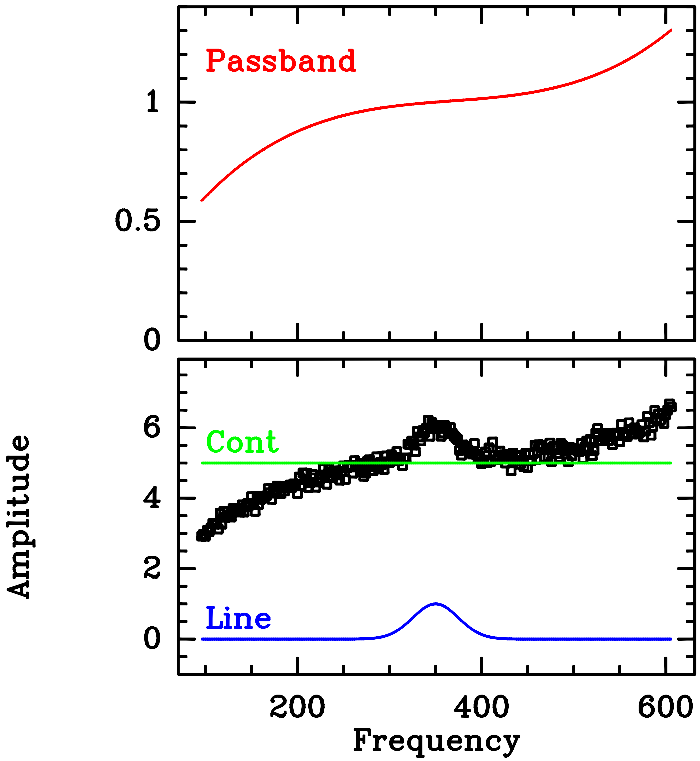
  – θ_B is the synthesized beam,
  – $\Delta\theta$ the relative position uncertainty.

  The SNR on the bandpass calibration should be better than the SNR of the spectral features observed; otherwise the relative positional accuracy will be limited by the accuracy of the passband calibration.

RF bandpass calibration accuracy (2)

Amplitude accuracy:

- The passband curve multiplies directly the detected signal
- if one measures a line (no continuum) the passband errors affect only the line shape
- if one wants to measure a weak line in front of a strong continuum, in particular for a broad line,
  one needs to measure the passband with an amplitude accuracy better than that is needed on source to get desired SNR.
  Example: we want to measure a line which is 10% of the continuum, with a SNR of 20 on the line strength : then the SNR on the continuum source should be 200, and the SNR on the passband calibration should be at least as good.



$\delta L = C / R$

R is the signal to noise ratio on the passband

$L / \delta L = (L / C) . R$

Side band calibration (1)

- A millimeter-wave interferometer can be used to record separately the signal in both sidebands of the first LO (see previous lecture).
- If the first mixer does not attenuate the image side band, it is useful to average both side bands (more continuum sensitivity, both for detection and calibration)
- The relative phases of the two sidebands can be arbitrary (Plateau de Bure where the IF frequency is variable)
- Calibration of the relative phases is done by measuring the phases of the upper and lower sidebands on the passband calibrator observation.

$$e^{i\phi_U} = \frac{\int \widetilde{V}_{ij,USB}(\nu,t)d\nu}{|\int \widetilde{V}_{ij,USB}(\nu,t)d\nu|} \quad (17)$$

$$e^{i\phi_L} = \frac{\int \widetilde{V}_{ij,LSB}(\nu,t)d\nu}{|\int \widetilde{V}_{ij,LSB}(\nu,t)d\nu|} \quad (18)$$

Side band calibration (2)

- These values can be used later to correct each side band phase to compensate for their phase difference.
  Double side band visibility is:

$$\widetilde{V}_{ij,DSB}(\nu,t) = e^{-i\phi_U}\widetilde{V}_{ij,USB}(\nu,t) + e^{-i\phi_L}\widetilde{V}_{ij,LSB}(\nu,t) \qquad (19)$$

  the two terms at right have zero phase at the time of the pass band calibration and they keep the same phase during the whole observing session.
- At observing time, offsets on the first and second LOs can be introduced so that both ϕ_U and ϕ_L are very close to zero when a project is done (actually done at Plateau de Bure, at the same time when the side band gain ratio is measured). This allows to plot DSB raw data without previous side band calibration.

Phase calibration: geometry

- Correct for the time dependence of the complex gains, contained in the functions $g_{\mathrm{C}i}(t)$.
- Variations occur in both amplitude and phase.
- Typical effects to be calibrated:
  - Interferometer geometry: Small errors in the baseline determinations cause slow phase drifts (period 24 hours);
  - these errors are dependent on the source direction so they cannot be properly calibrated out by phase referencing on a calibrator, only reduced by a factor of the order of the source to calibrator distance expressed in radians. (cf lecture on astrometry?)

Phase calibration: atmosphere

- The atmosphere introduces phase fluctuations on time scales 1s to a few hours, depending on baseline length and atmospheric conditions (see previous lectures).
- Fluctuations on short time scales cause loss of amplitude by de-correlation,
- On the long term the phase fluctuations can be mistaken for structure in the source itself.
- The critical time there is the time it takes for the projected baseline vector to move by half the diameter of one antenna:

$$\Delta t_C = \frac{D}{2} \frac{86400}{2\pi b} = 224 \frac{D}{15} \frac{450}{b} \tag{20}$$

  This time ranges from 4 minutes to 1 hour for Plateau de Bure, depending on the baseline length.
- The phase fluctuations on short time scales may be corrected by applying the radiometric phase correction method, if it is available.

Phase calibration: antennas

- The antennas may cause variations in amplitude gains due to degradation in pointing and in focus due to thermal deformations.
- These can be corrected to first order by amplitude calibration
- it is much better to keep the errors low by proper frequent monitoring of the pointing and focus, since these errors affect differently sources at the center and at the edges of the field.
- Focus errors also cause strong phase errors due to the additional path length which is twice the sub-reflector motion in the axial direction.
- (these phase errors are corrected in real time)

Phase calibration: electronics

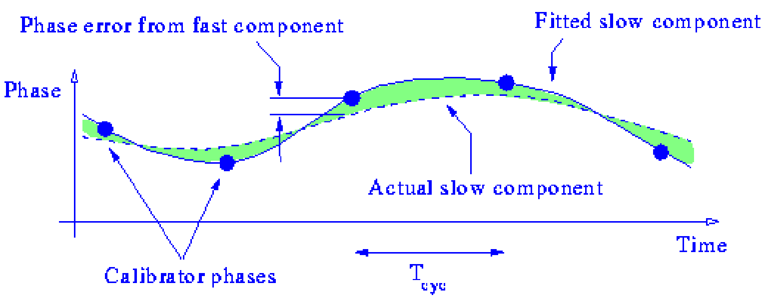
- Phase and amplitude drifts in the electronics are kept low by efficient design.
- The electronics phase drifts are generally slow and of low amplitude, hardware problems excepted.

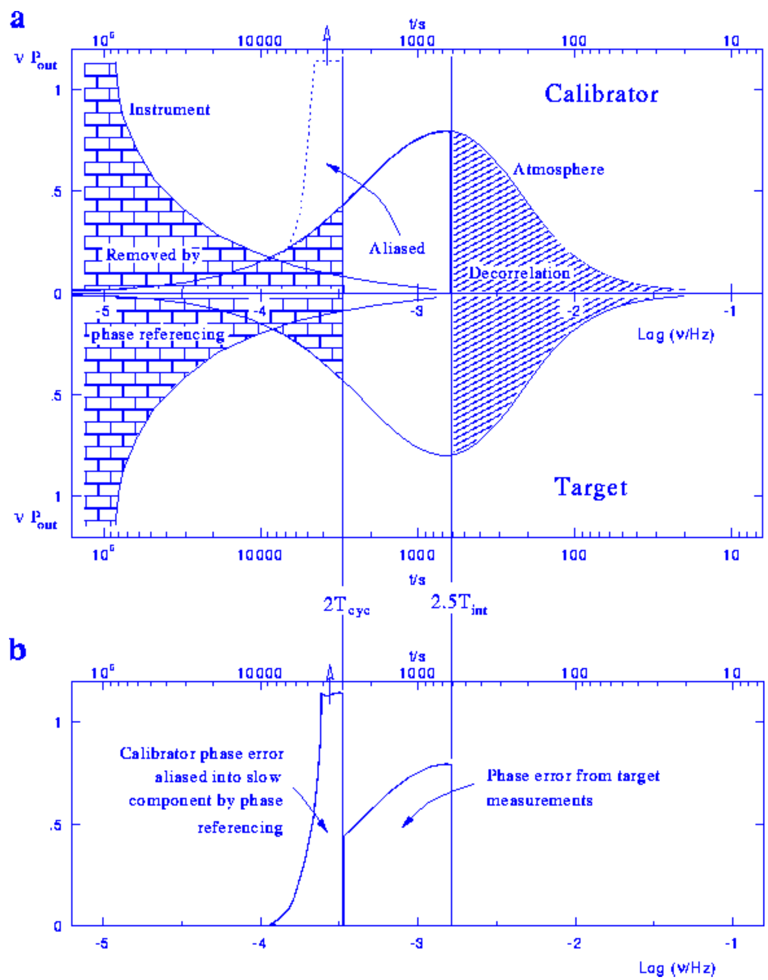
Phase referencing by a point source (1)

- This is the standard, traditional way to calibrate the phases with current interferometers.
- A point source calibrator is typically observed for T_{I} (a few minutes) every T_{C} (20-30 minutes).
- One fits a gain curve to the data observed on the calibrator, this gain curve is an estimate of the actual gain curve $g_{\mathrm{C}i}(t)$.
- This enables removing most long-term phase drifts from the observation of the target source.
- However this does not remove the bulk of atmospheric fluctuations.

Phase referencing by a point source (2)

The slow component of $g_{\mathrm{C}i}(t)$ to be calibrated out is sampled at intervals T_{C} so that only variations with periods longer than $2T_{\mathrm{C}}$ can be followed.
However one fits a slow component into the data points so one is sensitive to errors due to the presence of the fast component: the slow component is *aliased* into a slow component.
It is then recommended then to fit a curve that does not go through the points.





from O. Lay, 1997A&AS..122..547L

Phase referencing using another frequency (1)

- It is generally easier to measure the path length fluctuations at a lower frequency (even though the rms phase scales like frequency),
  due to both better receiver sensitivity and larger flux of the referencing source.
- In marginal weather conditions, if the rms phase fluctuations at 90 GHz is $\sim 40^\circ$, then at 230 GHz they are of $\sim 100^\circ$, and the phase becomes impossible to track directly due to 2π ambiguities.

Phase referencing using another frequency (2)

- If two receivers are available simultaneously, one may subtract to the high frequency phase the phase measured at the low frequency, corrected by the ratio of frequencies.
- The atmospheric fluctuations are canceled and only a slow instrumental drift remains.
- The gain curve at the high frequency is then determined as the sum of two terms:
  - the low frequency gain curve (including the slow atmospheric terms), scaled by frequency
  - plus the slow instrumental drift (which represents any phase fluctuation affecting one of the signal paths of the two receivers).

Fast phase referencing

- One may reduce $T_{\rm C}$ as much as possible to remove a larger part of the atmospheric fluctuation spectrum.
- Time scales of the order of 10s may be used, at the expense of:
  - the time efficiency is decreased (relatively more time is spent on the calibrator resulting in a larger overhead)
  - caution must be taken than the time $T_{\rm C}$ may become comparable to the time it takes to water eddies to drift along the apparent distance between source and calibrator.

  This is most useful for long baselines, for which there are strong atmospheric phase fluctuations at low frequencies.
- Very 'agile' antennas are required
- Planned for ALMA...

Phase referencing by a point source in the primary beam

- Simple case where the field contains a strong point source: it can be a continuum source (quasar) or a line source (maser).
- In that case all phase fluctuations with period longer than $\sim 2T_{\rm I}$ are removed, where $T_{\rm I}$ is the integration time.
- However statistical errors may be mistaken for true atmospheric phase fluctuations, causing additional de-correlation.
- This method gives good results, but for very specific projects which can be observed in very poor atmospheric conditions (e.g. observation of radio emitting quasars, of stars with strong maser lines).

Amplitude Calibration

- One has determined a temperature scale that takes into account:
  - variations in electronic gain
  - variations of atmospheric absorption
- However, at least two origins of amplitude errors remain
  - Variations of antenna gain (pointing, focus)
  - Decorrelation due to phase fluctuations
- This is why we do amplitude referencing using a strong, nearby source.
- If the source is too weak, as amplitude noise is biased the result may be wrong.

Decorrelation

- It can be shown that the decorrelation factor for a given baseline is approximated by:

$$f \sim 1 - \frac{1}{2} \int_{-log2.5T_{\mathrm{I}}}^{\infty} \nu P_{\phi}(\nu) d(\log \nu) \quad (21)$$

  ($P_{\phi}(\nu)$ is the power spectrum of the phase error function).
- The decorrelation is fundamentally a baseline based quantity: It cannot generally be expressed as a product of antenna-based factors.
- Both the target source and the calibrator are affected, so amplitude referencing will correct for decorrelation.
- However, the amount of de-correlation will vary from an integration to the next, so that the amplitude uncertainty is increased.

Amplitude referencing

- The amplitude gain is measured on the calibrators by dividing the baseline visibility amplitudes by the source fluxes (valid only for point sources)
- These gains are interpolated in time (using cubic splines) and will be applied to the target source(s).
- The calibrator fluxes should be known, this is tricky as they are variable (more to come !).

Summary, Concluding remarks

- Phase:
  - Only long-term atmospheric phase calibrated out (≥ 30min.)
  - Short-term atmospheric phase canceled by phase correction.
  - Future system will not have simultaneous multi-frequency observations (no real impact)
- Amplitude:
  - Decorrelation can be suppressed by phase correction.
  - Precision amplitude calibration relies on flux measurements.
- Bandpass:
  - Can be the limiting factor for some experiments.
- Limitations motivated by current Plateau de Bure system:
  - Only scalar treatment, polarization ignored (so far)
  - Self-calibration not discussed (few strong sources)