Phase calibration and water vapor radiometry for millimeter-wave arrays

O. P. Lay

Division of Physics, Mathematics and Astronomy, California Institute of Technology, Pasadena CA 91125, U.S.A.

Abstract. Correcting for the fluctuations in atmospheric path length caused by water vapor is a major challenge facing millimeter- and submillimeter-wave interferometers, and one that must be overcome to obtain routine sub-arcsecond resolution. Using the model for the power spectrum of phase fluctuations developed in Lay (1996), the existing technique of phase referencing to a bright calibrator object is analysed. It is shown that the phase errors after calibration have comparable contributions from both the target and calibrator measurements.

The technique of water vapor radiometry, where the amount of emission from water vapor in the beam of each antenna is used to estimate a path correction, is also examined. It is found that there are two levels on which a correction can be made. The simplest corrects just the fluctuations within each on-source period; the calibration requirements for the radiometers are modest, and this partial correction can give a substantial improvement in the resolution and coherence time of an interferometer. The atmospheric fluctuations on longer timescales remain uncorrected, however, and are significant. To remove these, a full correction is required, which measures the change in the path difference that occurs when moving between the calibrator and the target, in addition to the on-source fluctuations. Since there can be a large difference in airmass between the calibrator and the target, measuring this change requires that the radiometers have the same response to a given column of water vapor to within $\sim 0.1\%$. Two possible methods of achieving this very stringent limit are outlined.

For reasonable observing conditions at 230 GHz, it is predicted that the effective atmospheric 'seeing' (the apparent smearing of the sky brightness distribution due to the atmosphere) is improved from 0.6″ (phase referencing every 25 minutes) to 0.3″ (phase referencing and partial radiometric correction). A full radiometric correction would, in principle, restore perfect seeing.

Key words: Atmospheric effects – Instrumentation: interferometers – Site testing – Techniques: interferometric

1. Introduction

As radio interferometers push to higher frequencies and longer baselines, phase fluctuations resulting from the irregular distribution of water vapor in the troposphere become the dominant limitation on spatial resolution.

The technique of phase referencing, where observations of a calibrator object are interleaved with observations of the target, has been used to compensate for drifts in the instrumental response and also corrects for slow atmospheric variations. In its traditional 'slow' form, with calibrator observations every 20 minutes or so, phase referencing is insufficient for dealing with most atmospheric fluctuations, which are on shorter timescales. The possibility of very fast phase referencing, with a cycle time of order 10 seconds, is now being considered for new arrays (Holdaway 1992; Holdaway & Owen 1995), but sets stringent requirements on antenna agility and data aquisition rates.

Water vapor radiometry has long been proposed as a solution to the atmospheric phase problem (e.g. Westwater 1967; Schaper et al. 1970). As well as causing a propagation delay, water vapor along the line of sight emits radiation; the more water vapor present along the line of sight, the greater the emission, and the greater the propagation delay. By monitoring the emission from water vapor in an antenna beam as a function of time, it is possible to derive a phase correction that can be applied to the data. There are two basic approaches being developed to measure the water vapor emission: (1) monitoring fluctuations with the same receivers used for the astronomical observations (e.g. Welch 1994, Bremer et al. 1995), or (2) mounting a dedicated radiometer on each antenna, a technique being pursued at the Owens Valley Millimeter Array and the submillimeter interferometer comprising the Caltech Submillimeter Observatory and the James Clerk Maxwell Telescope on Mauna Kea. The algorithms for obtaining a path correction from an emission measure

are not considered here (see e.g. Sutton & Hueckstaedt 1996), although the accuracy needed is addressed.

The aim of this paper is to describe and quantify the impact of atmospheric phase fluctuations on astronomical observations, including phase referencing and water vapor radiometry. The analysis centers on the spectral density distribution of the phase fluctuations, which shows graphically the amount of phase fluctuation power on each timescale for a given baseline. The model is based on Kolmogorov turbulence 'frozen' in a layer of given thickness, and is presented with supporting data in Lay (1996), hereafter referred to as Paper I.
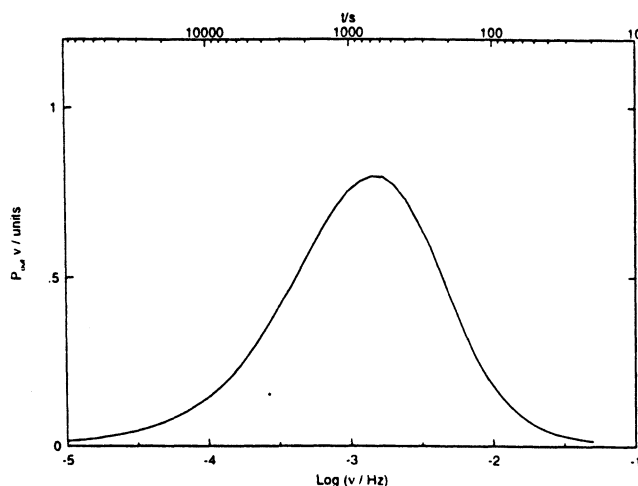

Fig. 1. Model for the frequency distribution of phase fluctuation power from a 1-km thick turbulent layer, measured on a 500-m baseline perpendicular to a 5 m s^{-1} wind. P_{out} is the spectral density of the phase variations at the output of the interferometer. The axes are chosen so that the area under the curve is proportional to the variance of the phase. For typical conditions at the Owens Valley Radio Observatory, the total area corresponds to a phase variance of $(110°)^2$ when observing at 230 GHz.

Figure 1 shows an example for a 500-m baseline, where the turbulent layer is 1-km thick and blows perpendicular to the baseline at 5 m s^{-1}. Most of the fluctuation power is on timescales of order 600 s (10 minutes). The distribution depends on a number of parameters, e.g.: longer baselines will show more power on longer timescales; the orientation of the baseline with respect to the wind direction affects the shape; a stronger wind shifts the curve to the right (doubling the windspeed halves the timescales); and the vertical scaling depends on the strength of the turbulence and is proportional to the square of the observing frequency. The effect of an outer scale to the turbulence has not been included; in this example, an outer scale size

of 10 km will reduce the phase power on timescales longer than 2000 s. The total area under the curve corresponds to a phase variance of $(110°)^2$ when observing at 230 GHz. Approximately one quarter of the area is accounted for by fluctuations with a period less than 400 s, so that these fluctuations alone would give an rms phase of 55°.

The next section examines the ability of phase referencing to correct for phase fluctuations. Section 3 shows how radiometric corrections can be incorporated into the phase referencing framework, and the demands this places on the precision of the correction. Section 4 illustrates how the different levels of phase correction improve the response to a point-like object. The summary is followed by an appendix that shows how only the fast fluctuations contribute to decorrelation in an integration.

2. Traditional gain calibration

The usual observing procedure for millimeter arrays is to interleave observations of the target object with a point-like calibrator. This is known as phase referencing. An observing cycle usually consists of several integrations of duration T_{int} on the target, followed by a single integration on the calibrator. For example, there might be four 5-minute integrations on the target followed by a 5-minute integration on the calibrator, giving a cycle time T_{cyc} of 25 minutes. The amplitude and phase measured on each baseline for the calibrator change with time, as a result of instrumental drifts and atmospheric fluctuations. A complex gain curve is fitted to the calibrator observations as a function of time, and is then removed from the measurements of the target object.

The next two sections quantify how phase fluctuations get folded into this observing procedure. Some contribute to a decorrelation (i.e. reduction in the visibility amplitude) within each integration, some are effectively removed, while others result in phase errors in the complex visibility measurements.

2.1. Decorrelation within an integration

The complex visibility sampled by an interferometer can be written as $Ve^{i(\Phi+\phi(t))}$ where the $Ve^{i\Phi}$ is the visibility intrinsic to the brightness distribution on the sky, and $\phi(t)$ is the variation in the phase introduced by the atmosphere and instrumental drifts. Variations in the amplitude gain have been ignored. After integrating for time T_{int}, the averaged visibility is described by $fVe^{i(\Phi+\Delta\Phi)}$, where f ($0 \leq f \leq 1$) reflects a reduction in the amplitude through decorrelation, and $\Delta\Phi$ is a phase error.

Rapid fluctuations tend to get averaged out and contribute little to $\Delta\Phi$ but dominate in f; conversely, fluctuations with periods much longer than T_{int} have very little effect on f and a large impact on $\Delta\Phi$. It can be shown

(see Appendix) that the total phase power remaining after averaging the fluctuations over T_{int} is given by

$$\Phi_{rms}^2 = \int_0^\infty P_{out}(\nu) \, sinc^2 \, (\pi\nu T_{int}) \, d\nu, \qquad (1)$$

where P_{out} is the spectral density of the phase fluctuations at the output of the correlator (units: $rad^2 \, Hz^{-1}$). Figure 2 shows the $sinc^2$ function plotted on a logarithmic horizontal scale, and how it is well approximated by a step function that becomes zero for $\nu = (2.5T_{int})^{-1}$. The rest of the phase power goes into the decorrelation of the measured amplitude:

$$f = \left\langle \frac{1}{T_{int}} \int_{T_0}^{T_0+T_{int}} \cos\phi'(t)dt \right\rangle_{T_0} \qquad (2)$$

$$\simeq 1 - \frac{1}{2T_{int}} \left\langle \int_0^{T_{int}} (\phi'(t))^2 dt \right\rangle_{T_0}, \qquad (3)$$

where $\phi'(t)$ represents the phase fluctuations with period $\lesssim 2.5T_{int}$, the angled brackets denote an ensemble average over all values of T_0, and the approximation is valid for $\phi'_{rms} \ll 1$. The second integral gives the power of $\phi'(t)$, and can be expressed as an integral over the relevant frequencies of the power spectrum, so that

$$f \simeq 1 - \frac{1}{2T_{int}} \int_{(2.5T_{int})^{-1}}^\infty P_{out} \, d\nu. \qquad (4)$$

Changing the variable of integration from ν to $\log\nu$ gives

$$f \simeq 1 - \frac{1}{2} \int_{-\log(2.5T_{int})}^\infty \nu P_{out} \, d(\log\nu), \qquad (5)$$

which shows that the area of the distribution in Fig. 1 with $t < 2.5T_{int}$ is a measure of the decorrelation, $1 - f$.

The visibility amplitudes measured for the calibrator also exhibit decorrelation, which can be quantified as a function of time for each baseline and then used to correct the corresponding amplitudes on the target object. Although this restores the average amplitude to the correct level, the amount of decorrelation is somewhat random from one integration to the next, and the uncertainty in the corrected amplitude is increased (see Figs. 8a and b). Note also that decorrelation is fundamentally a baseline-based quantity and cannot be expressed as a combination of antenna-based contributions in the same way that is possible for phase errors.

2.2. Phase referencing

The phase measured on the target for a given baseline is referenced to the corresponding phase measured on the calibrator, a technique that eliminates most slow drifts in the instrumental response. The calibrator is chosen to be both bright enough to give sufficient signal-to-noise in a single integration, and as close to the target as possible to
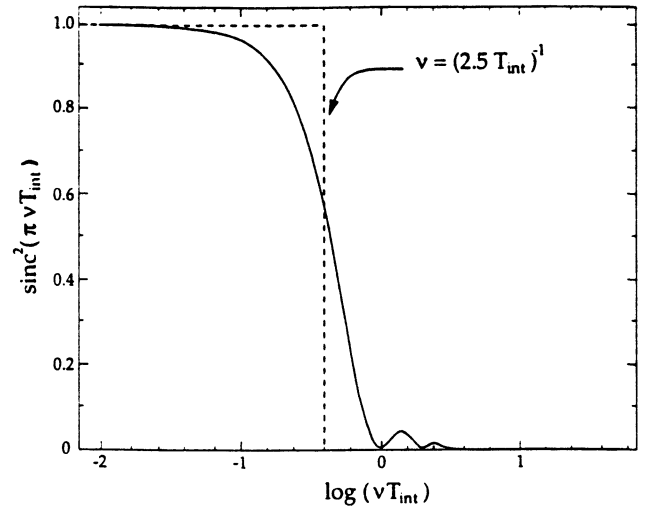

Fig. 2. $sinc^2$ $(\pi\nu T_{int})$ is well approximated by a step function with the transition at $\nu T_{int} = 2.5^{-1}$.

reduce position-dependent errors (e.g. from uncertainty in the relative locations of the antennas).

The cycling time between observations of the calibrator is T_{cyc}. This sampling rate can follow any component of $\Delta\Phi$ (the phase variation remaining after integration) with a period longer than the Nyquist limit of $2T_{cyc}$. These fluctuations constitute the true 'slow' component. A slow component *fitted* to the calibrator phases will not represent the true slow component, since there are errors introduced to each calibrator measurement by a 'fast' component, comprising those fluctuations with a frequency exceeding the Nyquist limit. This is illustrated in Fig. 3. The error caused by the fast component (shaded) looks like a slow component. This is the result of *aliasing*. The error component is white phase noise up to a high frequency cut-off of $(2T_{cyc})^{-1}$.
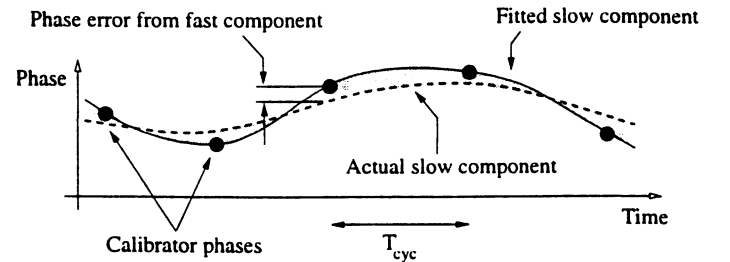

Fig. 3. Illustration of how phase fluctuations from the fast component ($t < 2T_{cyc}$) cause an error in the fitting of the slow component ($t > 2T_{cyc}$). The fast component is effectively aliased into a slow error component.
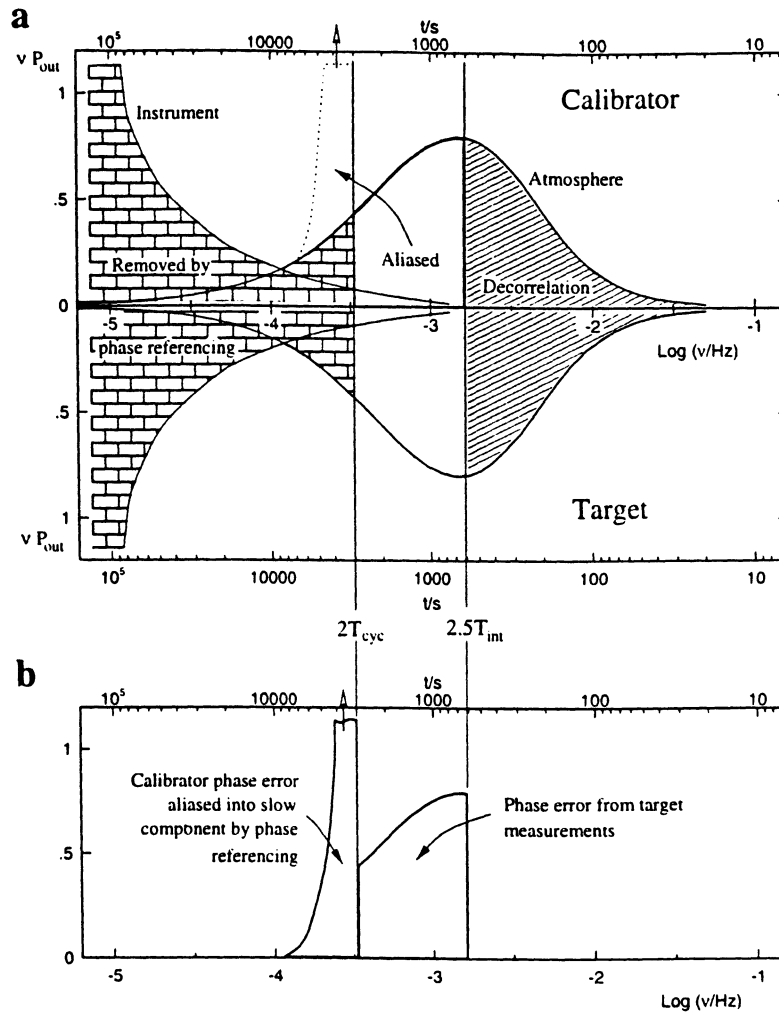
Fig. 4. a Phase power plot showing the contributions from the atmosphere and the instrument (schematic only). The upper and lower halves represent the calibrator and target measurements, respectively. The phase variance in the diagonally hatched region causes decorrelation of the amplitudes measured in time T_{int}. Phase fluctuations in the 'bricked' zone are correlated between the target and calibrator and are removed by phase referencing. The residual phase fluctuation power in the target observations after phase referencing comprises equal contributions from the target and the calibrator (blank regions in lower and upper halves, respectively). **b** Spectrum of the residual phase errors. The contribution from the calibrator is actually aliased to lower frequencies, so that it has a white noise spectrum up to a cut-off frequency of $(2T_{cyc})^{-1}$ (the sharp distribution is a result of the logarithmic horizontal scale).

The fitted slow component is subtracted from the phases measured on the target object as a function of time. This removes the effects of all phase fluctuations with a period longer than $2T_{cyc}$, but leaves the phase fluctuations from the target and calibrator for which $2.5T_{int} < t < 2T_{cyc}$, as illustrated in Fig. 4a. After phase referencing, there remain equal contributions to the phase error from the target and calibrator measurements. Since the calibrator contribution is aliased to lower frequencies, the spectrum of the phase errors has the appearance depicted in Fig. 4b.

The specific example shown in Fig. 4 has $T_{int} = 300$ s and $T_{cyc} = 1500$ s. For an observing frequency of 230 GHz, the phase variance from fluctuations with $t > 2.5T_{int}$ on each of the target and the calibrator is $\sim (70°)^2$, corresponding to $\sim 50\%$ decorrelation. The total phase error $\Delta\Phi_{rms}$ is $\sim 100°$, consisting of equal contributions from the calibrator (the aliased component) and target measurements.

One method for reducing the aliased component is to increase the shortest period $T_{fit,min}$ of the curve fitted to the calibrator phases (Fig. 3), so that it becomes over-sampled and no longer passes through all of the measurements. Only fluctuations with period $> 2T_{fit,min}$ are removed, so that the blank area in Fig. 4a corresponding to the uncorrected component is increased. Careful consideration of the effect of over-sampling, however, shows that the area of the aliased component on the calibrator is reduced by a factor $T_{fit,min}/T_{cyc}$, which can potentially more than offset the original increase.

The optimum choice of T_{int}, T_{cyc} and $T_{fit,min}$ depends on the phase power distribution (which depends in turn on the baseline length and orientation, windspeed, thickness of the turbulent layer, etc.) as well as instrumental constraints. Two cases are discussed briefly.

2.3. Regular phase referencing

When T_{cyc} is of order 1000 s, the spatial offset between the target and calibrator is unimportant for the cancellation of atmospheric fluctuations; the calibrator might as well be coincident with the target, as long as the elevation is not too low. This point is illustrated in Fig. 5. The wavelength of any fluctuation that is part of the slow component is much larger than the separation d_{XC} of the lines of sight to target and calibrator as they pass through the turbulent layer. It is then sufficient to subtract the fitted phase as a function of time on the calibrator from the measured target phases without worrying about the time offset d_{XC}/w, so that for the slow component of the phase error, $\phi_C(t) \simeq \phi_X(t)$.

Phase referencing becomes less effective at low elevations. There are two reasons for this: d_{XC} becomes larger, but, more significantly, the amplitude of large-scale phase fluctuations is proportional to the airmass $(1/\sin \epsilon)$. Although fluctuations with $t > 2T_{cyc}$ are correlated between target and calibrator, they are stronger for the object at lower elevation. The fractional error in the phase referencing correction is $\simeq \Delta\epsilon/\tan \epsilon \lesssim \Delta\epsilon/\epsilon$, where ϵ is the average of the two elevations and $\Delta\epsilon$ is the difference.

Another important point is that the impact of phase fluctuations does not increase indefinitely as the baseline gets longer. This is discussed further in Sect. 4.
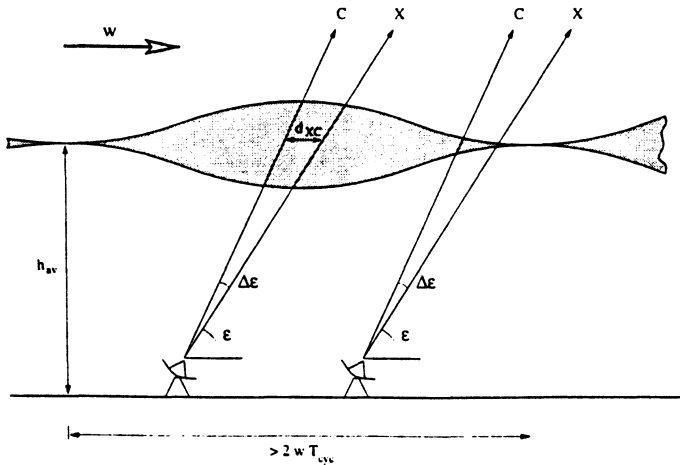

Fig. 5. A water vapor fluctuation (part of the slow component) at average height h_{av} with wavelength $\lambda > 2wT_{\text{cyc}}$. The distance between the lines of sight to the target (X) and calibrator (C) at height h_{av} is $d_{\text{XC}} \simeq h_{\text{av}}\Delta\epsilon/\sin \epsilon$. For $w = 5$ m s^{-1}, $T_{\text{cyc}} = 1500$ s, $\Delta\epsilon = 20°$, $\epsilon = 45°$ and $h_{\text{av}} = 2$ km, values of $d_{\text{XC}} \simeq 0.5$ km and $\lambda > 15$ km are obtained.

2.4. Fast phase referencing

In principle, T_{cyc} (and therefore also $T_{\text{fit.min}}$ and T_{int}) can be reduced to timescales of order 10 s, below which there is very little phase fluctuation power. In this regime, the separation of the lines of sight to the target and calibrator, d_{XC}, is comparable to the wavelength of the smallest fluctuation that can be sampled, $2wT_{\text{cyc}}$. For such fluctuations, it is no longer a good approximation to say that $\phi_C(t) = \phi_X(t)$; instead, a time lag needs to be introduced: $\phi_C(t) = \phi_X(t - T_{\text{lag}})$, where in the one-dimensional case $T_{\text{lag}} = d_{\text{XC}}/w$. The two-dimensional reality requires that d_{XC} and w are treated as vectors, with $T_{\text{lag}} = \mathbf{d}_{\text{XC}} \cdot \mathbf{w}/w^2$. There is then a residual, uncorrected phase error associated with the separation of the target and calibrator perpendicular to the wind direction.

The time overhead associated with rapid switching between sources is generally prohibitive for existing arrays – calibrators can be over 10° from the target and the antennas are not designed to be agile – but fast phase calibration is being seriously considered for future arrays (e.g. Holdaway 1992; Holdaway & Owen 1995).

3. Water Vapor Radiometry

Any radiometry system, whether it measures fluctuations in the noise level of existing receivers, or consists of separate, dedicated instruments, measures a signal that is related to the amount of emission E from the water vapor. This is converted into the path excess S through a gain factor g, i.e. $S = gE$. The fluctuations in the amount of water vapor along a line of sight can be typically of order a few percent of the total water vapor column (Paper I), so that the fluctuating part of S is small compared to its absolute value. The latter scales linearly with the airmass and in the Owens Valley, site of the millimeter array run by Caltech, ranges from ~ 10 to 100 mm in the zenith direction.

3.1. The absolute water vapor column

The path excess derived from a single radiometer is given by $S_1 = g_1 E_1$, where g_1 is an estimate of the true gain factor: $g_1 = g + \Delta g_1$. The error Δg_1 includes uncertainties associated with the instrument (gain, bandpass, spillover, etc.), and also uncertainties in the atmospheric model used to convert emission into path delay (e.g. temperature profile, altitude of the turbulence). Ideally, the path excess should be measured to within $\lambda/20$; when observing at 230 GHz ($\lambda = 1.3$ mm), the required accuracy is therefore $\sim 50 \mu$m. Since S_1 can be ~ 150 mm (50 mm zenith delay at 20° elevation), it is necessary that $\Delta g_1/g < 3 \times 10^{-4}$ for this absolute measurement.
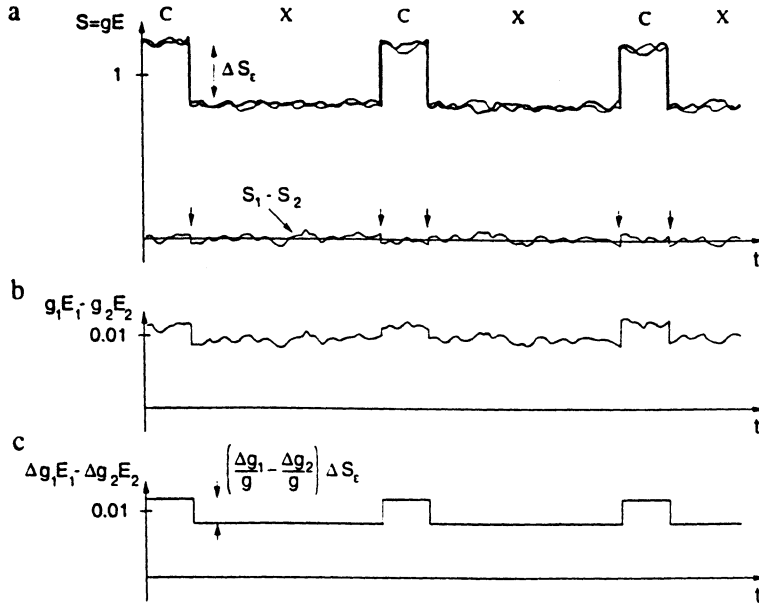
Fig. 6. a-c Plots showing how small gain differences in the radiometers can lead to a bad path correction. **a** Propagation delays (in arbitrary units) to two antennas as a function of time, alternating between the calibrator (C) and target object (X). The abrupt changes are the result of a difference in elevation between C and X (C has lower elevation in this case). Also shown is the difference between these quantities, which is the required path correction. This also has small discontinuities (arrowed), since the path correction is different for the lines of sight to X and C. **b** The derived path correction if $\Delta g_1/g - \Delta g_2/g = 10^{-2}$. **c** The resulting error in the path correction. The average error level is removed by phase referencing, but the steps remain.

3.2. The differential water vapor column

For an interferometer with radiometers on antennas 1 and 2, the differential path correction $S_1 - S_2$ is estimated by $g_1 E_1 - g_2 E_2$. The calibration requirement in this case is that $|\Delta g_1/g - \Delta g_2/g| < 3 \times 10^{-4}$, i.e. the gain factors have to *agree with each other* to within 0.03%. The errors in the gain factors, Δg_1 and Δg_2, can be substantial, so long as they are the same for all radiometers. It follows that the uncertainty in the atmospheric model is largely unimportant for this differential measurement. However, a small quantity (the difference in the effective path lengths) is being estimated by subtracting two large numbers (S_1 and S_2) which are measured independently, and the level of calibration needed is still daunting. These ideas are illustrated schematically in Fig. 6. The difference signal in Fig. 6a is small compared to the individual measurements, and a small error in the relative calibration of the radiometer gains generates a path correction (Fig. 6b) that is in error by the amount shown in (Fig. 6c).

In practice, this stringent calibration requirement can be relaxed slightly, since it is only necessary to reference the phase on the target source X to the phase on the calibrator C, i.e. it is sufficient to measure the quantity $g_1(E_{1X} - E_{1C}) - g_2(E_{2X} - E_{2C})$ as an estimate of the phase change that water vapor introduces between observations of the target and calibrator. Phase referencing then removes the error that is common to both target and calibrator measurements, and only the steps in Fig. 6c remain.

The difference in the water vapor column between the lines of sight to the target and the calibrator can still be substantial, however. The difference in path excess, ΔS_ϵ,

to an antenna for sources at elevation ϵ but separated by a small elevation offset $\Delta\epsilon$ is given by

$$\Delta S_\epsilon \simeq -S_z \frac{\cos\epsilon}{\sin^2\epsilon}\Delta\epsilon, \qquad (6)$$

where S_z is the path excess in the zenith direction. The worst case is for low ϵ and high $\Delta\epsilon$. For example, for $\epsilon = 25°$, $\Delta\epsilon = 10°$ and $S_z = 50$ mm, $\Delta S_\epsilon = 42$ mm. The requirement is now that the radiometers must be calibrated to the level of $|\Delta g_1/g - \Delta g_2/g| < 10^{-3}$ with respect to each other, i.e. all radiometers in the system must give the same value for the path delay from a given column of water vapor to within 0.1%. Two methods that attempt to achieve this are outlined in Sect. 3.4.

The value of 10° chosen for $\Delta\epsilon$ is typical (in many cases optimistic) for existing millimeter arrays. Arrays proposed for the future (e.g. the National Radio Astronomy Observatory Millimeter Array and the Japanese Large Millimeter and Submillimeter Array) have substantially higher sensitivity, and therefore have access to a much larger number of calibrator sources. If $\Delta\epsilon = 0.5°$ is used in the example above, then only 2% precision is required in the agreement between the radiometers.

3.3. Correcting on-source fluctuations only

The method just described assumes that radiometry is used to correct the phase changes due to water vapor introduced in moving from the target to the calibrator and back (the small discontinuities in $S_1 - S_2$ in Fig. 6a), in addition to correcting the fluctuations during each on-source period. It is the former that demands tight calibration of the radiometers, and so it is instructive to investigate the level of phase errors remaining if only the on-source fluctuations are removed. In this case, the average phase correction for each on-source period is made equal to zero.

Figure 7 shows the result of doing this, for a case where the time T_X spent on the target is four times longer than the time T_C spent on the calibrator. The plot is similar to Fig. 4, except that the radiometry corrects all fluctuations with $t < 2.5T_X$ for the target and $t < 2.5T_C$ for the calibrator. The residual rms phase for this specific example can be estimated from the diagram: the uncorrected (aliased) area is approximately 40% of the total for the calibrator, corresponding to an rms phase error of 70° at 230 GHz.

It is also possible to estimate the level of calibration needed for the radiometers in this case. The rms phase with $t < 2.5T_X$ that is being corrected by the radiometry while observing the target is $\sim 100°$, i.e. an rms path correction of $\sim 400\mu$m. Path corrections of ~ 1 mm would not be uncommon, so that an accuracy of 50 μm requires that $\Delta g/g \lesssim 5\%$.

By reducing decorrelation and increasing the coherence time of the interferometer, this partial radiometric correction offers a substantial improvement over phase referencing alone. Using radiometry to its full potential, however, requires a method of achieving the much more stringent calibration requirements outlined in 3.2.
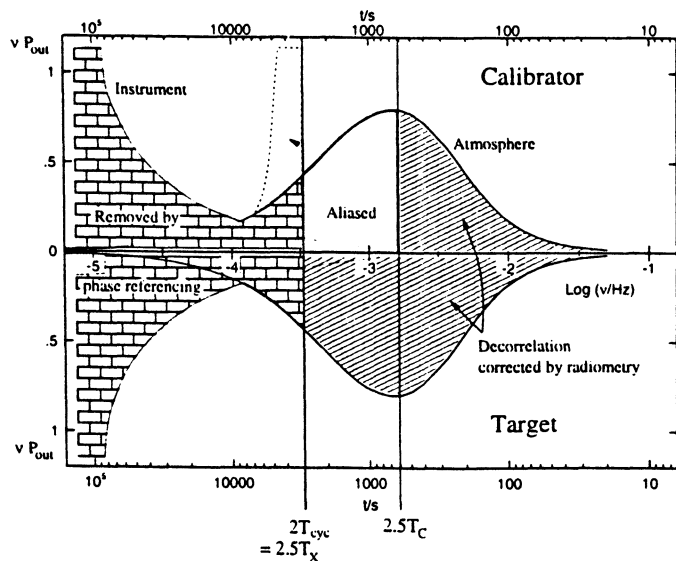

Fig. 7. Phase power plots showing the result of correcting on-source fluctuations only. This example has $T_X = 1200$ s and $T_C = 300$ s. The residual phase error is dominated by the aliased contribution from the calibrator.

3.4. Calibrating the Radiometers

To ensure that the full radiometric phase correction does not do more harm than good, all radiometers need to be calibrated to have the same sensitivity to water vapor to within 1 part in 10^3 (Sect. 3.2), i.e. if all the radiometers were to look at the same column of water vapor (difficult to arrange in practice) then they must give the same reading to within 0.1%. Although the response of radiometers to hot and cold loads can be determined to this level of precision, small differences in the bandpass shapes, spillover, scattering, etc., will inevitably introduce systematic uncertainties in the response to water vapor. The errors introduced will be a function of elevation and the prevailing atmospheric conditions (e.g. temperature profile, altitude of turbulence), both of which vary with time. There may also be a residual gain variation of the radiometers themselves that is not removed by a hot and cold load calibration. The relative calibration of the radiometers should therefore be made *against a column of water vapor on the sky*, frequently enough to follow variations in the response. Two possible calibration schemes have been identified.

3.4.1. Monitoring the difference between the derived phase corrections

For a given baseline, windspeed and thickness of the turbulent layer, there is a timescale T_{max} beyond which there is very little fluctuation power. The data presented in Paper I indicate that for a 100-m baseline, during typical conditions in the Owens Valley, there is very little power for $T_{max} > 1$ hour. This timescale is longer for the 500-m model shown in Fig. 1, but this does not include the effects of an outer scale to the turbulence, which will reduce the power on long timescales.

The radiometer gains g_1, g_2, etc., should therefore be scaled to ensure that the phase corrections derived for each baseline (Fig. 6b) average to zero, over periods exceeding $T_{max}/2.5$ (Fig. 2 for factor of 2.5). This effectively calibrates out all variations in g_1, g_2, etc. with period greater than T_{max}. By also monitoring over a long period how the difference between radiometer measurements on a baseline vary as a function of elevation and azimuth, the effects of different spillover patterns can also be reduced.

This method assumes that the antennas are at the same altitude and that the atmosphere can be considered planar over the area of the array. If these conditions are not satisfied, the long term average of the atmospheric phase fluctuations will be non-zero, but it may still be possible to estimate the offsets to the required level of precision. The long averaging times needed for long baselines means that this technique is likely to be more practical for short baselines.

3.4.2. Two phase calibrators

Another possibility is to observe two bright calibrators whose positions are well known and well separated in elevation (thereby emphasizing the steps in Fig. 6). In the absence of atmospheric fluctuations the phase measured

by an interferometer should be the same for each source, assuming that each is at the phase center and that baseline errors are negligible. This should also be the case after radiometry has been used to correct for atmospheric fluctuations, and the radiometer gain factors can be scaled to ensure this. Because of the 2π ambiguity in phase, there are many scalings that satisfy this condition. The correct relative scaling can be found by measuring several bright calibrators at different elevations initially, and then subsequently correcting for drifts by using just two calibrators.

An observing cycle that uses two calibrators (in addition to the target source) has the advantage that while one is used as a phase reference to remove instrumental drifts, the other one can be mapped in parallel with the target, as a useful check on the phase correction and imaging procedure.

This method can be used to remove variations in g_1, g_2, etc., with periods exceeding twice the observing cycle time. The disadvantages are that time must be spent observing an object that is not of direct interest, and the method will not correct errors introduced by different spillover patterns.

3.5. Alternative schemes

There are several other possibilities that could improve phase correction without the stringent radiometer calibration described above.

3.5.1. Self-calibration:

The technique of phase self-calibration (e.g. Readhead et al. 1980) can be used to remove phase errors that are not measured directly. The use of radiometry for correcting fluctuations during each on-source period means that self-calibration can be applied to periods of data up to the on-source time T_X. If the instrumental drift is slow, then T_X can be made large, and it will be possible to self-calibrate on much fainter objects than was possible without radiometry; radiometry effectively increases the coherence time of the interferometer.

3.5.2. Filter out the error signal:

Another possibility is to filter out the frequencies in the derived phase correction that result from the error pattern shown in Fig. 6c, before the correction is applied to the data in post processing. The power spectrum of this error signal consists of peaks with frequencies of n/T_{cyc} (where n is a positive integer), widths that depend on the variation in $\Delta g_1 - \Delta g_2$ over time, and strengths proportional to $\mathrm{sinc}^2(\pi n T_C/T_{cyc})$. By removing those frequencies where the error phase power exceeds the atmospheric phase power, the phase error can be reduced substantially. This technique will be more successful for shorter baselines where there is less power on long timescales, and when the gain factors of the radiometers vary only slowly with time,

giving narrower peaks in the error power spectrum. It is also important that the observations of the target and calibrator are as regular as possible.

3.5.3. Increase sensitivity and move to drier site:

An increase in the array sensitivity through larger antennas, a greater number of antennas, or more sensitive receivers, for example, allows fainter calibrators to be used that can be found much closer to the target. This reduces the size of the steps in S_1 and S_2 in Fig. 6a and therefore reduces the error introduced by differences in the radiometer gains. The same is true for a site that has a lower column of water vapor above it. Phase correction using radiometry will therefore be much easier for the future generation of millimeter arrays, with many antennas at a drier, high altitude site, than for the existing arrays.

3.6. How often should a radiometric correction be applied?

If an average radiometric path correction is derived and applied to the incoming signals for contiguous time intervals of duration T_{rad}, then all fluctuations with $t \gtrsim 2T_{rad}$ can be corrected. In the ideal system, the correction would be made in real-time, either by applying appropriate phase offsets to the local oscillators at each antenna, or by applying the offsets in the correlator. The limiting value of T_{rad} is then set by the sensitivity of the radiometry system, and could easily be less than 1 second. This form of correction is irreversible, however, so that a bad radiometer measurement could do more harm than good to the astronomical data. Until water vapor radiometry becomes a reliable technique, it would be prudent to apply as much of the correction as possible offline. Offline corrections will also be needed to account for gain changes in the radiometers, as might be determined from the two calibration procedures described above.

At the Owens Valley Millimeter Array, continuum data are recorded every 10 s, so that with a purely offline correction, fluctuations with $t \lesssim 20$ s go uncorrected. Reference to Fig. 8 of Paper I indicates that there is generally little phase power on these timescales in the Owens Valley, so that an offline correction every 10 s would be adequate in most cases for the continuum data. The much larger quantity of spectral line data is typically recorded at intervals of 1 to 5 minutes, however. A purely offline correction is much less effective in this case, unless the sampling time can be reduced substantially.

4. Phase correction and resolving power

It has been shown that it is possible to predict the residual uncorrected phase variance after various calibration schemes have been applied. But what impact does this phase variance really have? Perhaps the most direct approach is to consider the response to a point-like object. What upper limit can be set on the size of the object?
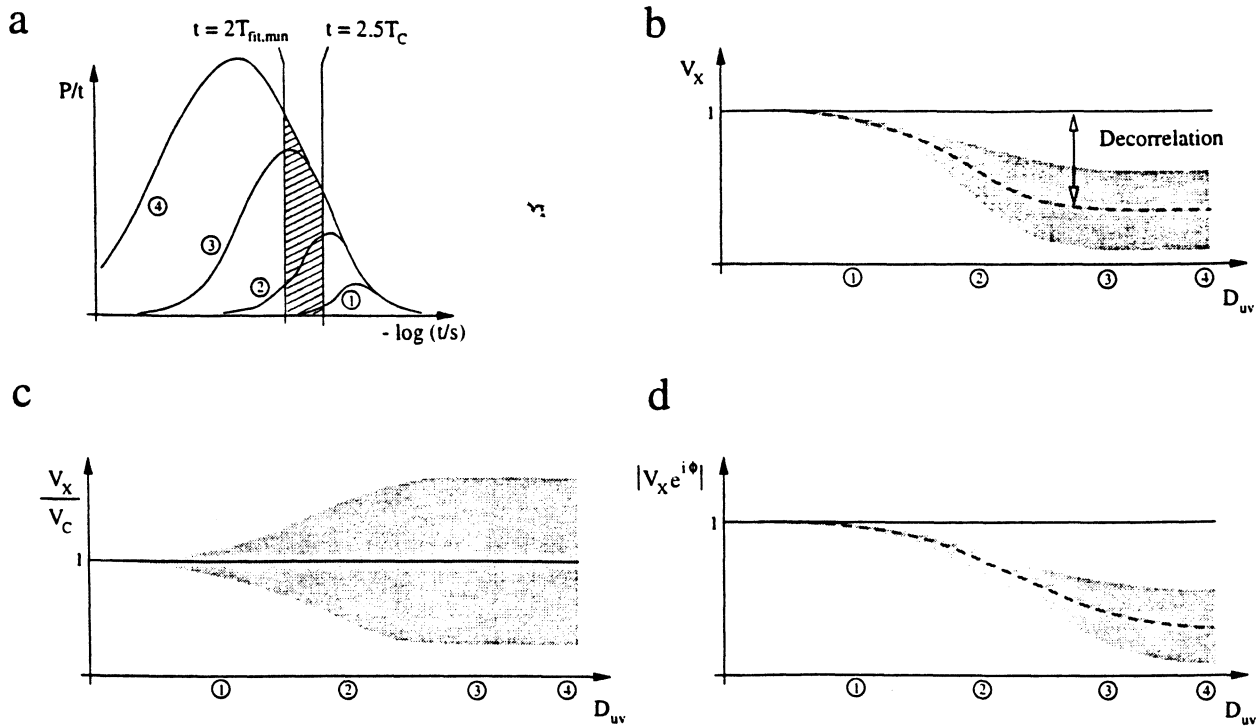
a

$t = 2T_{\text{fit,min}}$ $t = 2.5T_C$

P/t

④
③
②   ①

- log (t/s)

b

V_X

1

Decorrelation

①   ②   ③   ④

D_{uv}

c

$\dfrac{V_X}{V_C}$

1

①   ②   ③   ④

D_{uv}

d

$|V_X e^{i\phi}|$

1

①   ②   ③   ④

D_{uv}

Fig. 8. a-d The effect of residual phase errors on the ability to resolve a small object. **a** Phase power distributions as a function of fluctuation period t for four different baseline lengths. The hatched region indicates the fluctuations contributing to the phase error for each integration on the calibrator. **b** Visibility amplitude V_X from one integration as a function of baseline length for an unresolved point source of unit flux, with (dashed line) and without (solid line) phase fluctuations. The shaded region shows the distribution of amplitudes that might be expected, scattered about the average value. **c** Amplitude for one integration, corrected for decorrelation by dividing by the corresponding calibrator amplitude. **d** The amplitude expected after vector averaging over many integrations, with radiometry used to correct the decorrelation in each integration.

The problem is best addressed in the visibility domain, illustrated schematically in Fig. 8. Figure 8a shows how the phase power distribution changes as the length of the baseline is increased. The phase power on short timescales is independent of the baseline length. When the peak of the distribution is on a timescale slower than $2.5T_{\text{int}}$, the decorrelation in an integration becomes independent of baseline length (Fig. 8b); when it is slower than $2T_{\text{cyc}}$ the phase error is also independent of baseline length. The timescale for the peak ranges from 5 to 15 times d_{proj}/w, for wind w blowing along and perpendicular to a projected baseline length d_{proj}, respectively (Paper I). For $T_{\text{cyc}} = 1500$ s and $w = 5$ m s^{-1} the residual phase power becomes constant for baselines exceeding 3.7 km if the wind blows along the baseline, and 1 km if the wind blows perpendicular. These can be reduced through decreasing T_{cyc}.

The scatter in the decorrelation also increases with baseline. The lower amplitudes measured on longer baselines can be corrected by dividing by the corresponding amplitude measured on the calibrator, but the uncertainty in the amplitude is increased, as shown in Fig. 8c, and there is no fundamental gain in signal-to-noise. If radiom-

etry is used to correct for the on-source fluctuations only, then the amplitude measured in each integration is restored to its correct value (Fig. 8b) but the uncorrected phase errors reduce the amplitude obtained when a number of complex visibilities are averaged together (Fig. 8d)

Fig. 9 shows a specific calculated example of the average coherence obtained over many integrations as a function of baseline length. This example is for a 5 m s^{-1} wind blowing a layer of turbulence 1 km thick in a direction perpendicular to the baseline. The strength of the turbulence corresponds to reasonable conditions in the Owens Valley (38° rms phase at 230 GHz measured on a 100 m baseline over a 5 minute interval) and the observing frequency is 230 GHz. The integration time T_{int} is 300 s and the cycle time T_{cyc} is 1500 s, the same as the values used for Fig. 4 and Fig. 7.

The total decorrelation with no radiometric correction, indicated by the crosses in Fig. 9, is the result of phase fluctuations on the target with $t < 2T_{\text{cyc}}$ and those on the calibrator with $2.5T_{\text{int}} < t < 2T_{\text{cyc}}$ that are introduced by phase referencing (see Fig. 4). The solid Gaussian curve in Fig. 9 is the visibility curve that would be obtained under perfect conditions for a circular Gaussian source on
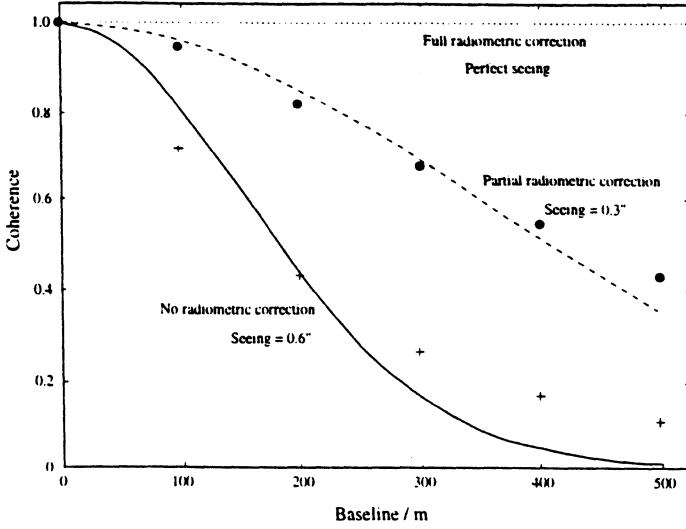
Fig. 9. Model prediction for the overall coherence as a function of baseline length, for a frequency of 230 GHz under reasonable conditions (see main text for details). The crosses are the expected values for phase referencing only, and circles also include a radiometric correction for on-source fluctuations. The curves are Gaussians, and represent the visiblity curves that would be obtained under perfect conditions from circular Gaussian sources on the sky with FWHM of 0.6″ and 0.3″. These are a measure of the effective seeing. A full radiometric correction restores perfect seeing.

the sky that has a Full-Width-to-Half-Maximum of 0.6″. This figure is therefore a measure of the effective 'seeing' – the amount of smearing of the sky brightness distribution due to the atmosphere – when there is no radiometric correction.

The circles show the behavior to be expected if radiometry is used to correct the fluctuations during each on-source period only, with the reduced coherence resulting from just the aliased calibrator component of phase fluctuations, as shown in Fig. 7. The dashed curve represents the visibility obtained for a Gaussian source on the sky with FWHM of 0.3″ (note that a wide curve in this visibility plot corresponds to a narrow distribution on the sky).

The effective seeing is clearly improved by this limited radiometric correction, but obtaining the highest possible spatial resolution on faint objects requires the full correction (dotted line in Fig. 9), which in turn requires very precise calibration of the radiometers.

5. Summary

1. The turbulent distribution of water vapor causes fluctuations in the path length through the atmosphere, with periods ranging from seconds to ~ 1 hour. These fluctuations interact with the observing process for in-

terferometers in different ways, depending on their period t.

2. Fluctuations with $t \lesssim 2.5 T_{\rm int}$, where $T_{\rm int}$ is the integration time, reduce the amplitude measured in each integration; those with $t \gtrsim 2.5 T_{\rm int}$ produce an error in the phase measured for the integration.

3. Phase referencing to a nearby bright calibrator source removes fluctuations with $t > 2 T_{\rm cyc}$, but adds the fluctuations on the calibrator with $2.5 T_{\rm int} < t < 2 T_{\rm cyc}$, where $T_{\rm cyc}$ is the period of the observing cycle. This component is aliased to longer timescales.

4. Water vapor radiometry can be used to correct for the residual fluctuations. There are two levels on which the correction can be applied. Correcting the fluctuations during each on-source period constitues a partial correction, since the aliased component from the calibrator is left intact. The coherence time of the interferometer is increased, and the calibration requirements for the radiometers are modest ($\sim 5\%$).

5. A full correction also measures the change in the path correction between observations of the target and calibrator. Since there can be a large difference in airmass between the calibrator and the target (at least for existing arrays), measuring this change requires that the radiometers have the same response to a given column of water vapor to within $\sim 0.1\%$. This stringent calibration may be possible to achieve by monitoring the average derived correction on timescales longer than the atmospheric fluctuations, or by observing more than one calibrator.

6. For reasonable observing conditions at 230 GHz, the prediction for effective seeing is improved from 0.6″ (phase referencing every 25 minutes) to 0.3″ (phase referencing and partial radiometric correction). The full radiometric correction would, in principle, restore perfect seeing.

Appendix: frequencies contributing to decorrelation and phase error

What is the average phase remaining after averaging a sinusoidal phase fluctuation of frequency ν over a time interval $T_{\rm int}$? The random fluctuation from a single spectral component at frequency ν can be represented by

$$\phi(t, \nu) = A(\nu)\{g_{\rm c} \cos(2\pi\nu t) + g_{\rm s} \sin(2\pi\nu t)\}, \qquad (7)$$

where $A(\nu)$ is the rms amplitude of the fluctuation, and $g_{\rm c}$ and $g_{\rm s}$ are Gaussian random variables with a variance of unity. The averaged phase is given by

$$\phi_{\rm av}(\nu) = \frac{1}{T_{\rm int}} \int_{-T_{\rm int}/2}^{+T_{\rm int}/2} A(\nu)\{g_{\rm c} \cos(2\pi\nu t)$$
$$+ g_{\rm s} \sin(2\pi\nu t)\} \, dt \qquad (8)$$
$$= \frac{1}{T_{\rm int}} \int_{-T_{\rm int}/2}^{+T_{\rm int}/2} A(\nu) g_{\rm c} \cos(2\pi\nu t) \, dt \qquad (9)$$

$$= A(\nu) g_c \, \text{sinc} \, (\pi \nu T_{\text{int}}). \qquad (10)$$

The variance of ϕ_{av} for a given value of ν is

$$\nu)\rangle = A^2(\nu) \, \text{sinc}^2 \, (\pi \nu T_{\text{int}}) \qquad (11)$$

$$= P_{\text{out}}(\nu) \, \text{sinc}^2 \, (\pi \nu T_{\text{int}}). \qquad (12)$$

The total phase variance after averaging is obtained by summing all of the independent random contributions. If $\phi(t)$ is an infinite time series, then

$$\langle \phi_{\text{av}}^2 \rangle = \int_0^\infty P_{\text{out}}(\nu) \, \text{sinc}^2 \, (\pi \nu T_{\text{int}}) \, d\nu, \qquad (13)$$

as given in Eq. 1.

References

Bremer, M., Guilloteau, S., Lucas, R., 1995, The Phase Project: Observations of Quasars, IRAM Working Report No. 238

Holdaway, M. A., 1992, Possible phase calibration schemes for the MMA, MMA Memo 84

Holdaway, M. A., Owen, F. N., 1995, A test of fast switching phase calibration with the VLA at 22 GHz, MMA Memo ˙26

, O. P. 1996, A&A, this volume (Paper I)

dhead, A. C. S., Walker, R. C., Pearson, T. J., Cohen, M. H., 1980, Nature, 285, 137

Schaper, L. W., Staelin, D. H., Waters, J. W., 1970, Proc IEEE, 58, 272

Sutton, E. C., Hueckstaedt, R. M., 1996, A&A, in press

Welch, Wm. J., 1994, The Berkeley–Illinois–Maryland–Association Array. In: Ishiguro, M. & Welch, W. J. (eds.) Astronomy with Millimeter and Submillimeter Wave Interferometry, ASP Conference Series Vol. 59, p. 74

Westwater, E. R. 1967, An analysis of the correction of range errors due to atmospheric refraction by microwave radiometric techniques, ESSA Tech. Rpt. IER 30-ITSA 30, Institute for Telecommunication Sciences, Boulder, Colorado