THE ASTRONOMICAL JOURNAL VOLUME 109, NUMBER 3 MARCH 1995

FRINGE DETECTION METHODS FOR VERY LONG BASELINE ARRAYS

ALAN E. E. ROGERS AND SHEPERD S. DOELEMAN
Haystack Observatory, Westford, Massachusetts 01886
Electronic mail: arogers@wells.haystack.edu, dole@wells.haystack.edu

JAMES M. MORAN
Harvard–Smithsonian Center for Astrophysics, Cambridge, Massachusetts 02138
Electronic mail: moran@cfa.harvard.edu

Received 1994 August 17; revised 1994 November 18

ABSTRACT

In VLBI observations, atmospheric turbulence and local oscillator phase noise often limit the coherence time to a few seconds. When such coherence losses severely limit the integration time, detection thresholds rise and weak sources cannot be detected. We show here that under these conditions, the detection methods and measurements of visibility amplitude and phase can be reformulated in terms of incoherently averaged quantities. The theory is presented by examining the properties of averaged amplitudes and triple products. This paper reviews current detection methods, discusses new techniques, presents applications to recent 3 mm VLBI experiments, and shows that signal processing can significantly improve the probability of signal detection.

1. INTRODUCTION

VLBI at millimeter wavelengths is severely limited by atmospheric path length fluctuations. Under these conditions the fringe detection threshold can be improved by incoherent averaging (see Thompson *et al.* 1986; hereafter referred to as TMS) or by using the complex triple product. The triple product is also known as the bispectrum and its properties have been extensively studied by Kulkarni (1989) and Cornwell (1987). With an array, further improvements can be made by the simultaneous use of all elements in a global fringe search. A global fringe search algorithm has been given by Schwab & Cotton (1983) and is implemented in AIPS (1990). In this paper we give expressions for the detection threshold in a global fringe search, and explore related issues in a quantitative way.

2. CONVENTIONAL SINGLE BASELINE FRINGE SEARCH

The standard fringe detection method in VLBI is a single baseline search of the two-dimensional space of delay and delay rate for a maximum correlation amplitude. The signal-to-noise ratio (SNR) for a point source without coherence loss is

$$\mathrm{SNR}=L\rho\sqrt{(2BT)}, \tag{1}$$

where ρ is the correlation amplitude$=T_a/T_s$, (T_a/T_s) is the geometric mean of the ratio of source antenna temperature to system temperature on each baseline, B is the bandwidth (Hz), L is the digital loss factor ≤ 1 (1 for ideal analog processing), T is the coherent integration time (s). The probability of false detection or probability of error (PE) is given by

$$\mathrm{PE}=1-[1-e^{-R^2/2}]^n\approx ne^{-R^2/2}, \tag{2}$$

where R is the peak correlation in units of SNR, and n is the number of independent points in the two-dimensional search over delay and delay rate space.

The SNR is the ratio of the signal amplitude to the standard deviation of one component of the noise vector. With this definition the SNR becomes the inverse of the rms phase for large SNR. A SNR of 7 is sufficient to ensure a very small probability of false detection for a fringe search over 10^6 trial values of delay and delay rate (see TMS, p. 262). Alef & Porcas (1986) have demonstrated the use of antenna based residuals to narrow the search range on each baseline and lower the detection threshold by decreasing n in Eq. (2).

3. AVERAGING OF DATA SEGMENTS

Data averaging can be extended beyond the coherence time by incoherently averaging data segments. The field of view is usually small enough that the visibility phases change slowly and can be coherently averaged over the coherence interval or "segment." In this case an unbiased estimate A of the correlation amplitude can be made by averaging M segments of the correlation amplitude squared minus the expected value of the noise. A is given by

$$A=\left[\frac{1}{M}\sum_{i=1}^{M}(|a_i|^2-2)\right]^{1/2}, \tag{3}$$

where a_i is the complex correlation which results from coherent integration over time interval T in units of one standard deviation of one component of the noise vector. The estimate is unbiased because

$$\langle|a|^2\rangle=2+s^2, \tag{4}$$

where s is the signal in units of the rms noise, and hence

$$\langle A^2\rangle=s^2. \tag{5}$$

 0004-6256/95/109(3)/1391/11/$0.90

We note that the quantity within the square root of Eq. (3) can be negative, in which case we will take our estimate for A as zero. In the weak signal case, we define the SNR for A^2 as

$$\mathrm{SNR}_A = \frac{\langle A^2 \rangle}{\sqrt{\langle A^4 \rangle}}. \tag{6}$$

In the absence of a signal, the variance of A^2 is $4/M$ so that when $s \ll 1$, Eq. (6) becomes

$$\mathrm{SNR}_A = \frac{s^2}{2}\sqrt{M}. \tag{7}$$

The detection reliability can be estimated by comparing the measured value of A with the expected peak value of A due to the noise in the absence of a signal. For a large number of segments the probability distribution of A^2 becomes Gaussian by the central limit theorem so that the probability of false detection, PE_A, is determined by integrating the tail of a Gaussian probability distribution

$$\mathrm{PE}_A = 1 - \left[1 - \frac{1}{\sqrt{2\pi}} \int_{\mathrm{SNR}_A}^{\infty} e^{-x^2/2} dx\right]^n. \tag{8}$$

For the case where the integral in Eq. (8) is much less than unity, we use the usual asymptotic expansion for the error function and we obtain the approximation

$$\mathrm{PE}_A \approx \frac{n}{\sqrt{2\pi}\mathrm{SNR}_A} e^{-\mathrm{SNR}_A^2/2}, \tag{9}$$

which is valid for $\mathrm{PE}_A \ll 1$. With the SNR_A defined by Eq. (6), a SNR of 7 is sufficient to reduce the probability of false detection to 1.3×10^6 in a search of 10^6 points. We can convert the threshold for reliable detection corresponding to a particular value of SNR_A into an equivalent signal for a single segment by inverting Eq. (7),

$$s = \sqrt{2\mathrm{SNR}_A} M^{-1/4}. \tag{10}$$

If a value of SNR equal to SNR_T is required for reliable detection of unit flux density within a single segment, then the ratio of the flux density needed for detection with segmented data, s_{incoh}, compared to that of a single segment, s_{coh} is

$$\frac{s_{\mathrm{incoh}}}{s_{\mathrm{coh}}} = M^{-1/4} \sqrt{\frac{2\mathrm{SNR}_A}{\mathrm{SNR}_T^2 - 2}}. \tag{11}$$

The subtraction of 2 from SNR_T^2 is required to correct the coherent average for noise. Choosing values of $\mathrm{SNR}_T = 7$ and $\mathrm{SNR}_A = 6.6$, results in the same negligibly small probability of error (<0.01% in a search of 10^6 points) for both coherent [from Eq. (2)] and incoherent [from Eq. (9)] averaging. The above relation then becomes

$$\frac{s_{\mathrm{incoh}}}{s_{\mathrm{coh}}} = 0.53 M^{-1/4}. \tag{12}$$

The factor 0.53 is valid only for large M. For a smaller number of segments, the above analysis will not hold due to a breakdown of the Gaussian approximation introduced in Eq. (8). Flux density limits for small values of M are, however, accessible via two alternate routes. The first is to numerically integrate the exact expression for the probability distribution of A^2 (given in Appendix A). Since the probability of error is related to $p(A^2)$ by

$$\mathrm{PE} = 1 - \left[1 - \int_{A_0^2}^{\infty} p(A^2) d(A^2)\right]^n, \tag{13}$$

specification of PE and the number of search points n, yields the flux threshold A_0. The solid line in Fig. 1 shows this limit (normalized to the value for $M = 1$) as a function of segment number.

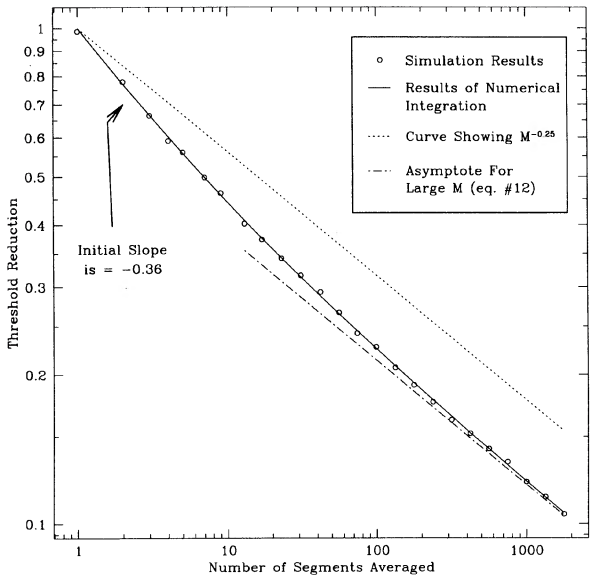

FIG. 1. Reduction in the flux detection threshold obtained by searching over an incoherent average of many segments. The open circles show results of a simulation in which a maximum in segmented and averaged amplitude was found in a search of 100 points in (delay–delay rate) space. Two thousand of these searches were made and the threshold was defined as the dividing line between the top 1% and lower 99%. The solid line is calculated by a numerical integration of the probability density for A_0^2 from ∞ [a large enough value to make $p(A_0^2)$ extremely small] down towards 0 until a value of 10^{-4} is reached. Both methods are equivalent to a combined PE=0.01 and $n = 100$ for which $\mathrm{SNR}_T = 4.3$ and $\mathrm{SNR}_A = 3.8$. Normalization is to the $M = 1$ point on the solid curve.

The second method uses simulations to statistically estimate flux thresholds. A single simulation represents a search for a maximum over n samples of A *with no signal.* After making a large number of searches, we separate the resulting distribution of noise maxima so that a fixed percentage are above our threshold value; this percentage is PE from Eq. (13). In Fig. 1 the simulation data corresponds to the same PE and n as the numerical integration of $p(A^2)$. The asymptotic limit described by Eq. (11) is also shown and agrees well with both the simulation and the exact determinations of flux limits for large values of M. Earlier work (see TMS, p. 269) has shown that incoherent averaging improves the detection threshold by the factor $M^{1/4}$. For small values of M, we find the sensitivity improves at a faster rate of $\sim M^{0.36}$. Because of this faster than expected reduction in detection threshold, the potential of incoherent averaging is more powerful than previously recognized. It should be noted that the

difference between the asymptote given by Eq. (11) and the curve for a simple $M^{1/4}$ rate improvement increases for larger searches and more stringent detection criteria (i.e., smaller PE).

The SNR given in Eq. (6) needs to be modified, following TMS, to include the presence of a signal when used to estimate the standard deviation of the correlation amplitude. In this case

$$\mathrm{SNR}'_A = \frac{\langle A^2 \rangle}{\sqrt{\langle A^4 \rangle - \langle A^2 \rangle^2}}. \tag{14}$$

Since

$$\langle |a|^4 \rangle = s^4 + 8s^2 + 8, \tag{15}$$

then

$$\mathrm{SNR}'_A = \frac{s^2}{2} \sqrt{\frac{M}{s^2+1}}. \tag{16}$$

In the limit when $s \gg 1$,

$$\mathrm{SNR}'_A \approx \frac{s}{2} \sqrt{M}. \tag{17}$$

The fractional error ϵ in the amplitude estimate, calculated by noting that SNR'_A is the signal to noise of the amplitude squared, is given by

$$\epsilon \approx \sqrt{1 + \frac{1}{\mathrm{SNR}'_A}} - 1, \tag{18}$$

or for $s \gg 1$ [by the use of Eq. (17)],

$$\epsilon \approx \frac{1}{s\sqrt{M}}, \tag{19}$$

which is the same as would be obtained for a coherent average over the entire data duration, with the assumption of no coherence loss.

Incoherent averaging of segments of duration T is equivalent to incoherent averaging of amplitudes in the fringe rate power spectrum with a sinc^2 windowing function of width $(1/T)$ Hz. This equivalence has been shown by TMS using Parseval's theorem. Incoherent averaging in the spectral domain allows a windowing function to be chosen that matches the spectrum of atmospheric phase fluctuations better than the sinc^2 function. Clark (1968) has shown that in the limiting case when the coherent integration is shortened to the inverse bandwidth the detection scheme becomes equivalent to the intensity interferometer described by Hanbury Brown (1974).

In segmenting complex amplitude data it has been the practice to average coherently nonoverlapping segments rather than to compute a running average for each input data point and the SNR theory has been given for M independent nonoverlapping segments. Two sets of segments can be derived from a common dataset by offsetting the segment boundaries so that the boundaries of one set are midway between the boundaries of the other set. In this case it can be shown that in the absence of a signal and in the limit of a large number of segments the incoherent average A^2 from one set is 50% correlated with the other set. Thus incoherently averaging the combined overlapping set results in a further small reduction of the flux density threshold by a factor of $(3/4)^{1/4}$ which corresponds to an improvement of about 7%. To realize the greatest threshold reduction, a running mean can be used to generate the segmented amplitudes. In this case the data which is made up of correlation accumulation periods (APs), is segmented such that each segment is offset from the previous one by a single AP. Incoherently averaging the entire set of segments lowers the threshold by a factor of $(2/3)^{1/4}$ or makes an improvement of about 10% in the limits of large M and a large number of APs per segment.

4. CLOSURE PHASE

The closure phase (see Jennison 1958; Rogers *et al.* 1974) is the sum of interferometer phases around a triangle of baselines. This phase is independent of instrumental phases and atmospheric path delays and depends only on the source structure. Neglecting noise, the closure phase for a point source is zero. In the following analysis in this section we assume that the source produces zero closure phase (i.e., it is either unresolved or has symmetric structure). To estimate the noise in the closure phase it is convenient to define a phase noise loss function $L(s)$ (see Rogers *et al.* 1984), which is defined for a random phase θ by the relation

$$L(s) = \langle \cos\, \theta \rangle. \tag{20}$$

Since the probability density distribution of the phase is given by

$$p(\theta) = \frac{e^{-s^2/2}}{2\pi} \int_0^\infty r e^{-r^2/2} e^{sr \cos\, \theta} dr, \tag{21}$$

(which reduces to a uniform probability distribution for $s=0$) one can show that

$$L(s) = \sqrt{\frac{\pi}{8}}\, s e^{-s^2/4} \left[I_0\left(\frac{s^2}{4}\right) + I_1\left(\frac{s^2}{4}\right) \right], \tag{22}$$

where I_0 and I_1 are hyperbolic Bessel functions. For low SNR, Eq. (22) reduces to

$$L(s) \approx \sqrt{\frac{\pi}{8}}\, s. \tag{23}$$

For the closure phase, θ_c, given by

$$\theta_c = \theta_1 + \theta_2 + \theta_3, \tag{24}$$

the associated loss function is

$$L(s_c) = \langle \cos(\theta_1 + \theta_2 + \theta_3) \rangle. \tag{25}$$

Since θ_1, θ_2, and θ_3 are independent and $\langle \sin(\theta_i) \rangle = 0$, we can expand Eq. (25) to obtain the result

$$L(s_c) = L(s_1) L(s_2) L(s_3). \tag{26}$$

If we define the SNR of the closure phase as

$$\mathrm{SNR}_c = L^{-1}[L(s_c)] = L^{-1}[L(s_1) L(s_2) L(s_3)], \tag{27}$$

hence, for low SNR [see Eq. (23)], we obtain

$$L^{-1}(s)\approx\sqrt{\frac{8}{\pi}}\,s, \tag{28}$$

so that

$$\mathrm{SNR}_c\approx\frac{\pi}{8}\,s_1s_2s_3, \tag{29}$$

which is valid when s_1, s_2, $s_3\ll 1$. In the high SNR limit,

$$L(s)\approx 1-\frac{1}{2s^2}, \tag{30}$$

so that

$$\mathrm{SNR}_c\approx\frac{s_1s_2s_3}{\sqrt{s_1^2s_2^2+s_1^2s_3^2+s_2^2s_3^2}}. \tag{31}$$

When $s_1=s_2=s_3=s$,

$$\mathrm{SNR}_c\approx\frac{s}{\sqrt{3}}. \tag{32}$$

5. BISPECTRUM OR TRIPLE PRODUCT

The bispectrum is the product of the complex interferometric amplitudes on the three baselines of a triangle and is given by

$$\mathbf{b}=|a_1\|a_2\|a_3|e^{i(\theta_1+\theta_2+\theta_3)}=|a_1\|a_2\|a_3|e^{i\theta_c}, \tag{33}$$

where a_i and θ_i are the amplitudes and phases of the visibility on the three baselines. The expectation of **b** is given by

$$\langle\mathbf{b}\rangle=\langle(\mathbf{s}_1+\mathbf{n}_1)(\mathbf{s}_2+\mathbf{n}_2)(\mathbf{s}_3+\mathbf{n}_3)\rangle=\mathbf{s}_1\mathbf{s}_2\mathbf{s}_3, \tag{34}$$

where $(\mathbf{s}_1+\mathbf{n}_1)$, etc., are the complex sum of signal and noise phasors. Each noise term component has unit variance. For low SNR on each baseline,

$$\langle|\mathbf{b}|^2\rangle=\langle|\mathbf{n}_1|^2|\mathbf{n}_2|^2|\mathbf{n}_3|^2\rangle=8. \tag{35}$$

If we use the ratio of the signal (which will appear in the real part of **b** for a point source) to one component of the noise vector as a measure of SNR then

$$\mathrm{SNR}_b=\frac{2\ \mathrm{Re}\ \mathbf{b}}{\sqrt{\langle|\mathbf{b}|^2\rangle}}, \tag{36}$$

or, since $\langle(\mathrm{Re}\,\mathbf{b})^2\rangle=\langle|\mathbf{b}|^2\rangle/2$,

$$\mathrm{SNR}_b\approx\frac{s^3}{2}, \tag{37}$$

where $s_i=|\mathbf{s}_i|$ and $s=s_1=s_2=s_3\ll 1$.

We can write an expression for the bispectral SNR that is valid for a wider range of SNR by including the signal terms previously neglected in Eq. (35) and noting that $\langle\mathbf{s}_i\mathbf{n}_i\rangle=0$. In this case,

$$\mathrm{SNR}_b\approx\frac{s_1s_2s_3}{\sqrt{4+(s_1^2s_2^2+s_1^2s_3^2+s_2^2s_3^2)+2(s_1^2+s_2^2+s_3^2)}}, \tag{38}$$

where $s_1^2=\mathbf{s}_1\mathbf{s}_1^*$, etc. Kulkarni (1989) gives an even more general expression for the bispectrum SNR. Note however, that he defined the SNR as the ratio of the signal to the total noise.

6. AVERAGE OF CLOSURE PHASORS AND TRIPLE PRODUCTS

If the bispectrum or triple product for coherent integration time T is averaged for M equal segments the SNR will improve. First, consider the closure phasor average

$$C=\frac{1}{M}\sum e^{i\theta_c}. \tag{39}$$

For weak signals, we use Eqs. (23), (25), and (27) to obtain

$$\langle\mathrm{Re}\ \mathbf{C}\rangle=\langle\cos\ \theta_c\rangle=\sqrt{\left(\frac{\pi}{8}\right)^3}\,s^3, \tag{40}$$

and in the absence of a signal (where θ_c is a uniformly distributed random variable)

$$\langle(\mathrm{Re}\ \mathbf{C})^2\rangle=\frac{1}{M}\langle\cos^2\ \theta_c\rangle=\frac{1}{2M}. \tag{41}$$

If we use $\langle\mathrm{Re}\ \mathbf{C}\rangle/\sqrt{\langle(\mathrm{Re}\ \mathbf{C})^2\rangle}$ as a measure of the SNR, then

$$\mathrm{SNR}_C=\sqrt{\left(\frac{\pi}{8}\right)^3}\sqrt{2M}s^3, \tag{42}$$

and the probability of error, based on the analysis leading to Eq. (9) is given by

$$\mathrm{PE}\approx\frac{n}{\sqrt{2\pi}\mathrm{SNR}_C}\,e^{-\mathrm{SNR}_C^2/2}. \tag{43}$$

We follow a similar method for the bispectrum analysis. The bispectrum averaged over M segments is given by

$$\mathbf{B}=\frac{1}{M}\sum|a_1||a_2||a_3|e^{i\theta_c}. \tag{44}$$

For equal signal amplitudes the expectation of the real part of **B** is

$$\langle\mathrm{Re}\ \mathbf{B}\rangle=s^3, \tag{45}$$

and its second moment is

$$\langle(\mathrm{Re}\ \mathbf{B})^2\rangle=\frac{1}{M}\langle|\mathbf{b}|^2\rangle\langle\cos^2\ \theta_c\rangle. \tag{46}$$

Hence for low SNR,

$$\langle(\mathrm{Re}\ \mathbf{B})^2\rangle=\frac{4}{M}, \tag{47}$$

and with the definition, $\mathrm{SNR}_B=\langle\mathrm{Re}\ \mathbf{B}\rangle/\sqrt{\langle(\mathrm{Re}\ \mathbf{B})^2\rangle}$, we obtain

$$\mathrm{SNR}_B=\tfrac{1}{2}\sqrt{M}s^3. \tag{48}$$

Following Kulkarni (1989), for a wider range of SNR the more general expression for SNR_B is

$$\mathrm{SNR}_B\approx\frac{s_1s_2s_3\sqrt{M}}{\sqrt{4+(s_1^2s_2^2+s_1^2s_3^2+s_2^2s_3^2)+2(s_1^2+s_2^2+s_3^2)}}. \tag{49}$$

The probability of error, from Eq. (9), is

$$\mathrm{PE}\approx\frac{n}{\sqrt{2\pi}\mathrm{SNR}_B}\,e^{-\mathrm{SNR}_B^2/2}. \tag{50}$$

The detection threshold for the bispectrum average is about 30% lower in units of SNR or about 10% lower in units of s, or flux density, than the average of closure phasors. Thus, it is advantageous to average the triple product instead of the closure phasor. In either case the s^3 dependence of the SNR imposes a sharp flux density threshold. In the case of low SNR on one or two baselines of the triangle the bispectrum has a SNR advantage of about 10% and 20%, respectively, compared with an average of closure phasors. As in the case of incoherent averaging, a further small improvement can be made by averaging triple products from 50% overlapping segments. In this case the overlapping segments are 25% correlated and hence the threshold is reduced by $(5/8)^{1/6}$ or about 7%. In the running mean limit the threshold is improved by $(1/2)^{1/6}$ or about 11%.

If no assumption is made about source structure the closure phase can have an arbitrary value and we must examine the magnitude of the average bispectrum. In this case the SNR is reduced by $\sqrt{2}$ and becomes

$$\mathrm{SNR}'_{\mathbf{B}}=\frac{\langle|\mathbf{B}|\rangle}{\sqrt{\langle|\mathbf{B}|^2\rangle}}\approx\sqrt{\frac{M}{8}}s^3, \tag{51}$$

since $\langle|\mathbf{B}|\rangle\approx|\langle\mathbf{B}\rangle|$ and $\langle|\mathbf{B}|^2\rangle=2\langle(\mathrm{Re}\,\mathbf{B})^2\rangle$. The corresponding approximation for the probability of false detection is

$$\mathrm{PE}\approx\frac{n}{\sqrt{2\pi}\mathrm{SNR}'_B}e^{-\mathrm{SNR}'^2_{\mathbf{B}}/2}. \tag{52}$$

7. FRINGE SEARCHING WITH AN ARRAY

For weak signals it may not be possible to detect fringes on any individual baseline so that a global search is required. While the sensitivity of a coherent array of N equal elements improves by the square root of the number of baselines,

$$\sqrt{\frac{N(N-1)}{2}}, \tag{53}$$

the detection threshold for an array whose elements have not been "phased up" is improved by only about

$$\sqrt{\frac{N}{2}} \tag{54}$$

over that of a single baseline (see Rogers 1991). The failure of the threshold to improve by the square root of the number of baselines is the result of the vast increase in search parameter space. When a search is made over $n^{(N-1)}$ independent points, where n is the total number of independent parameters of delay, rate, and phase in the search to each element (beyond the first which acts as a reference), the detection threshold is increased by $\sqrt{(N-1)}$ as shown in Appendix B. For an array we consider the n parameters of each element as one dimension. Note that n can be very large, e.g., 100 rates, 100 delays and 4 phases give $n=40{,}000$. Because the search points may not be completely independent, it is shown in Appendix C that the expected noise level for an $(N-1)$ dimensional search is increased by at least $\sqrt{8(N-1)/9}$, for one specific prescription of a fringe search, which thereby sets a lower bound on the strength of a noise spike. Also, a least-squares estimate of station phases from $[N(N-1)/2]$ independent baseline phases has a standard deviation of $\sqrt{2/N}$, which shows that in the high SNR regime the added degrees of freedom needed to phase up an array result in an SNR loss of $\sqrt{(N-1)}$. This result comes from the covariance matrix of estimators derived from the normal equations as shown in Appendix D. The separation of the bounds placed by Appendix B and C is always less than 6% for any value of N, thereby ensuring that the factor $\sqrt{N/2}$ is accurate to within 6%. To summarize, the relative flux density needed for signal detection with an array, s_{array}, compared with that needed for detection with a single baseline, s_{base}, is given by

$$0.94\sqrt{\frac{2}{N}}\leq\frac{s_{\mathrm{array}}}{s_{\mathrm{base}}}\leq\sqrt{\frac{2}{N}}. \tag{55}$$

In the presence of atmospheric phase variations a global fringe search must be performed on segmented data. Data within a segment can be coherently added while the M segments are incoherently combined. The variable to be maximized is

$$G=\left[\frac{1}{M}\sum\left(\frac{2}{N(N-1)}\,\mathrm{Re}\sum a_{ij}e^{-i\theta_{ij}}\right)^2-\mathrm{bias}\right]^{1/2}, \tag{56}$$

where the inner summation is the coherent sum of complex visibilities a_{ij} over all baselines, frequencies, and samples within the coherence interval. θ_{ij} are the phases needed to counter rotate each data sample,

$$\theta_{ij}=\phi_i-\phi_j+\omega(\tau_i-\tau_j)+(\omega_{r_i}-\omega_{r_j})t, \tag{57}$$

where $\phi_1,\ldots,\phi_N$ are the instrumental phases, $\tau_1,\ldots,\tau_N$ are the instrumental delays, and $\omega_{r_1},\ldots,\omega_{r_N}$ are the instrumental rates (these parameters describe the search space), with the boundary condition that $\phi_1=\tau_1=\omega_{r_1}=0$.

The search algorithm given by Eq. (56) is based on the assumption that the source is unresolved. Since the visibility phases in this case are zero, the signal will appear in the real part of the inner summation. Consider only the contribution of the signal components to $\langle G^2\rangle$. In this case

$$\langle G^2\rangle=\left[\frac{2}{N(N-1)}\right]^2\sum_l\sum_m s^2\langle\cos\Phi_l\cos\Phi_m\rangle, \tag{58}$$

where Φ_l and Φ_m are the phase residuals on the l and m baselines [i.e., the visibility phases in Eq. (56) counter rotated by the instrumental phases]. For a large number of stations the terms for which $l\neq m$ are dominant. Also the residuals on different baselines should be uncorrelated so that

$$\langle G^2\rangle\approx s^2\langle\cos\Phi\rangle^2. \tag{59}$$

Since the global search results in only an estimate of the instrumental phases $\langle\cos\Phi\rangle$ can be viewed as a loss factor that results from the phase noise in the estimate. The reduction in signal is approximately given by the loss factor function of Eq. (22), whose argument is scaled by a factor f, so that

$$\langle G^2\rangle\approx s^2L^2(sf). \tag{60}$$

In the global search to maximize G there are $(N-1)$ unknown phases. With $(N-1)$ unknown phases to be determined from $N(N-1)/2$ baseline phases, the value of s, the SNR of a single baseline, is scaled by

$$f=\sqrt{\frac{N(N-1)}{2(N-1)}}=\sqrt{\frac{N}{2}}. \tag{61}$$

For a single baseline $\cos\Phi=1$ and no signal is lost by the presence of an error in the station phases since the search will remove the phase error in this case. In this case Eq. (58) becomes the same as Eq. (5). In practice, for signals around and above the detection threshold, $s\geq 1$, $L^2(sf)$ is close to unity and the approximation for L in Eq. (30) is valid. The bias term in G is given by

$$\text{bias}\approx\frac{0.9\times 2\times(N-1)}{N(N-1)/2}\approx\frac{3.6}{N}, \tag{62}$$

where the first factor of $(N-1)$ is the result of the $(N-1)$ dimensional search and the factor of 2 is the normal bias without searching. Note that there is a weak dependence of the bias on the search parameter n, which we have ignored here because it enters only as $\sqrt{2\log_e n}$. The numerical factor of 0.9, which results from the use of the real part rather than the magnitude in Eq. (56), was determined by numerical simulation. It is approximately the ratio of a maximum found using a search for the real part to that maximum found using a search for the magnitude. For a source of arbitrary structure the real part in the inner coherent sum should be replaced by the magnitude. In this case the numerical factor in the bias becomes approximately unity. In the absence of a signal

$$\langle G^4\rangle=\frac{4}{M}\frac{(N-1)^2}{[N(N-1)/2]^2}, \tag{63}$$

where the factor $(N-1)^2$ accounts for the change in probability distribution that results from choosing the maximum from an $(N-1)$ dimensional search for each segment. If we define

$$\text{SNR}_G=\frac{\langle G^2\rangle}{\sqrt{\langle G^4\rangle}}\approx\frac{s^2}{2}\frac{N}{2}\sqrt{M}L^2(sf), \tag{64}$$

we obtain an expression for the probability of error, with a large number of segments of [see Eqs. (8) and (9)]

$$\text{PE}\approx\frac{n}{\sqrt{2\pi}\text{SNR}'_G}e^{-\text{SNR}'^2_G/2}. \tag{65}$$

Unfortunately, the bias in Eq. (56) is somewhat dependent on the extent of the search, but may be estimated from the data in a region of search space in which it is known that there is no signal. Numerical simulations show that the use of the magnitude in place of the real part in Eq. (56), as is required when searching for a source of arbitrary structure, makes little difference to the detection threshold for $N>3$. In the case of a single baseline the process of taking the real part and searching through values of the second station phase for a maximum gives the same value for G as taking the magnitude without a search. When the signals are weak the loss factor is significant so that it may be better to incoherently average the individual baselines and not attempt to determine the station phases for the individual segments. In this case the individual baselines should be incoherently averaged in the same way as individual data segments. Using the theory given in Sec. 3 [see Eq. (10)] we get

$$\text{SNR}'_G=\frac{s^2}{2}\sqrt{M}\sqrt{\frac{N(N-1)}{2}}. \tag{66}$$

Equation (66) is only valid when the product of the number of segments and the number of baselines is large and the SNR is low, but otherwise has no other approximations. Note that for $N>2$, Eq. (66) gives a higher value for the SNR than Eq. (64). However we have not accounted for the extent of the search. Whereas the coherent search always requires an $(N-1)$ dimensional search, because the station phases are unknown, the incoherent search is far less demanding. For example, when the rates and delays are already determined on a strong calibrator it may be sufficient to search in only two dimensions for a right ascension and declination offset. The incoherent global search for the coordinates of the source requires relatively little computation, compared with a coherent global search. Hence the incoherent search is a very useful technique since, in many cases, it is just as sensitive as a fully coherent segmented search. Once a significant peak is found for the source's right ascension and declination, the visibility information is then best extracted by computing the bispectrum on all the available triangles for the phase information and computing incoherent averages on each baseline for the amplitude information. In order to include the phase information in existing image processing packages, the closure phases from the bispectral averages can be converted back to baseline phases by linear least squares using weights from the SNRs of the bispectral averages. Note that all the bispectral components should be used in the analysis since they are always at least partially independent [see also Kulkarni (1989), *Sec.* VI d.]

It is important to recognize that, as with fringe searching on a single baseline, there is a bias in global fitting with an array. If there is no source present then global fitting, without the bias correction term, will find a source of approximately $\sqrt{N-1}$ times the 1σ noise level for the phased array. Also, without bias correction the global fitting will increase the apparent flux of a detected source by approximately $\sqrt{(N-1)}\,\sigma$ when the source is so weak that the SNR on each baseline is unity or less. (Using global fitting without bias correction at the VLA is known to result in biased flux measurements.)

If the elements of the array have unequal sensitivities, the baselines in the summation of Eq. (56) should be weighted in proportion to the expected SNR for each baseline to maximize the SNR of the sum. For example, consider adding a baseline with SNR$=s$ to one with unit SNR using weight w. In this case the SNR of the sum is

$$\frac{(1+ws)}{\sqrt{(1+w^2)}}, \tag{67}$$

which is maximized to a SNR of $\sqrt{(1+s^2)}$ when $w=s$. This result can be readily extended to an array [e.g., Vilnrotter *et al.* (1992)]. For an array whose elements have their sensi-

tivity given by their system equivalent flux densities (SEFDs) the overall coherent sensitivity for a phased-up array is

$$\mathrm{SEFD}_c=\left[\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{1}{\mathrm{SEFD}_i\mathrm{SEFD}_j}\right]^{-1/2}, \tag{68}$$

where the summation is carried out over all baselines. For a single baseline the SEFD is the geometric mean of the SEFDs of the two elements. As with an array of equal elements the detection threshold is degraded by the vast increase in search parameter space. A lower bound on this degradation can be estimated using the method of Appendix C, by including the appropriate optimal weighting factors, to estimate the noise peak present in a search. The SEFD is effectively increased by the factor F given by

$$F=\frac{\Sigma_{j=2}^{N}\sqrt{\Sigma_{i=1}^{j-1}\frac{1}{\mathrm{SEFD}_i\mathrm{SEFD}_j}}}{\sqrt{\Sigma_{i=1}^{N-1}\Sigma_{j=i+1}^{N}\frac{1}{\mathrm{SEFD}_i\mathrm{SEFD}_j}}}. \tag{69}$$

Unless the SEFDs are very different (more than a factor of 10),

$$F\approx\sqrt{N-1}, \tag{70}$$

to an accuracy of better than about 10%. Another measure of the array sensitivity, useful for mapping, is given by the equivalent SEFD for fringes from the kth element to all other elements phased up and optimally weighted. In this case, for equal elements the sensitivity is improved by a factor of $\sqrt{N-1}$ relative to a single baseline, compared with the improvement factor of $\sqrt{N/2}$ for a global search on all elements of the array.

If signals from several antannea are optimally combined to form an element in a larger array then the SEFD of the arrayed element is given by

$$\mathrm{SEFD}_{\mathrm{array}}=\left[\sum_{i=1}^{N}\frac{1}{\mathrm{SEFD}_i}\right]^{-1}. \tag{71}$$

Equation (71) is the well-known result that the equivalent aperture of an array is the sum of the individual apertures. Note that Eq. (71) requires that the signals from the antennas be optimally weighted (see Dewey 1994) (which is not done at the VLA) and that there be no loss due to poor phasing. The SEFD of the baseline formed by this arrayed element and the kth element of the array is

$$\mathrm{SEFD}_{k-\mathrm{array}}=\sqrt{\frac{\mathrm{SEFD}_k}{\Sigma_{i\neq k}\frac{1}{\mathrm{SEFD}_i}}}.$$

If baselines are added incoherently, and we seek to optimize the detection of a weak source, the correlations should be weighted in proportion to the expected baseline SNR^2 so that

$$\mathrm{SEFD}_{\mathrm{incoh}}=\left[\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{1}{(\mathrm{SEFD}_i\mathrm{SEFD}_j)^2}\right]^{-1/4}. \tag{73}$$

Hence, the better systems become even more dominant in the overall performance of the array for source detection than in the case of a coherent search.

8. EXAMPLES

8.1 Closure Phase from Bispectrum

Figure 2 shows the results of determining the closure phase from an average of M segments of data. The SNR was computed from the data by taking the ratio of the signal voltage to the square root of the sum of noise power components in a direction normal to the signal vector. That is,

$$\mathrm{SNR}=\frac{\Sigma_{i=1}^{M}\mathrm{amp}_i\cos(\theta_{c_i}-\theta_c)}{\sqrt{\Sigma_{i=1}^{M}\mathrm{amp}_i^2\sin^2(\theta_{c_i}-\theta_c)}}, \tag{74}$$

where amp_i is the triple product of the amplitudes (s_1, s_2, and s_3), and θ_{c_i} is the closure phase for each segment, re-

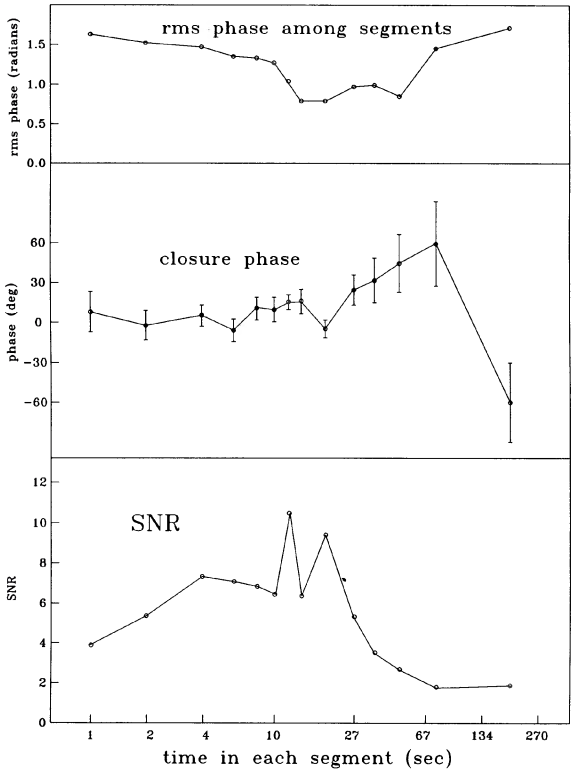

FIG. 2. Closure phase from the triple product for an observation of 380 s duration as a function of the segment duration. The data for a triangle of baselines formed by antennas at Haystack(K), Kitt Peak(P), and Owens Valley(O), observing the quasar NRAO530 at 86 GHz on 1994 April 4 at 0822 UT. The coherent SNRs for the full length of the observation are 15, 5, and 10, while the SNRs for 1 s segments are 1.5, 0.6, and 0.8 on the PO, KO, and KP baselines respectively. The SNR for the bispectrum shown is calculated by the method described in the text [see Eq. (72)]. Also shown is the rms variation of closure phases among segments and the approximate $\pm1\sigma$ error bars on the closure phase average from the inverse of the SNR. This approximation for the closure phase error is valid for SNR>1. In the limit of no signal or very short segments, the 1σ phase error approaches $\pi/\sqrt{3}$ radians. In this plot the minimum segment length is the coherent accumulation period of the hardware correlator (1 s) and the maximum length is that which produces two segments.

spectively, and θ_c is the closure phase for the sum. s_1, s_2, s_3 are estimated from the data on each baseline incoherently averaged with the same segment length. The SNR computed this way agrees with the SNR calculated from Eq. (49). As with data on a single baseline, the 1σ error in a measurement of phase is 1/SNR for SNR$\gg$1.

From Fig. 2 it is clear that there is an optimum coherent integration time that maximizes the SNR and minimizes the error in the determination of the closure phase. Too long a coherent integration degrades the SNR because of coherence loss and results in a corrupted value of closure phase. On the other hand, a very short coherent integration results in a low SNR. This happens when the SNR of each segment drops to unity, and the signal drops very rapidly owing to the s^3 term in the bispectrum. In the case shown here, two baselines are weak and the third is stronger. If only one baseline were weak, the decline of SNR would not occur until the segments are so short that the SNR of each segment on the stronger baselines approach unity.

8.2 Fringe Search on Incoherent Average of Segments

Figure 3 shows the result of a single baseline search in delay and delay rate using data from seven observations divided into 10 s segments. In this case, a single observation of 270 s duration does not result in a significant detection whereas the incoherent addition of 200 10 s duration segments results in the detection of the radio source with very small probability of false detection.

8.3 Global Fringe Search on Segmented Data

Figure 4 shows a global search for fringes on three baselines using a single observation of 27 10 s segments. In this case, although the individual baselines only show marginal fringes on two of the three baselines, the search clearly establishes the station clocks.

Figure 5 shows a global search on the same triangle of baselines. In this case, the sensitivity of the elements have been reduced by a factor of about 2 owing to added atmospheric attenuation as the source set. A search on each baseline failed to detect fringes. With narrow search windows, fringes were detected only on the baseline PO. The global search detected fringes just above the threshold with SNR$_G \approx 6$, and gave values for the clocks [K$=-5$ ns, $P=-60$ ns, O$=0$ ns (reference)] which agree with the values from the earlier observations, when corrected for the clock rates. Figure 6 shows the SNR, for the same observation, as a function of delay for each baseline separately, clearly

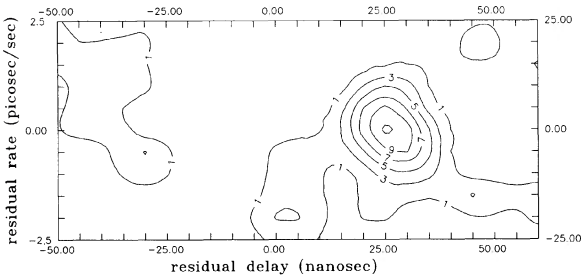

FIG. 3. Contour plot of SNR$_A$ for a two-dimensional search for fringes using 200 10 s segments of the PO baseline taken on the Galactic Center (Sgr A*) at 86 GHz on 1993 November 13 from 2030 to 2345 UT. The peak at 25 ns and zero fringe rate is a significant detection. The coherent SNRs for the individual 270 s duration observations taken individually do not provide a significant detection. The residual delay is shown for a range of 125 ns which corresponds to the multiband delay ambiguity formed by the synthesis of a 56 MHz band using seven 8 MHz single-band channels spaced 8 MHz apart.

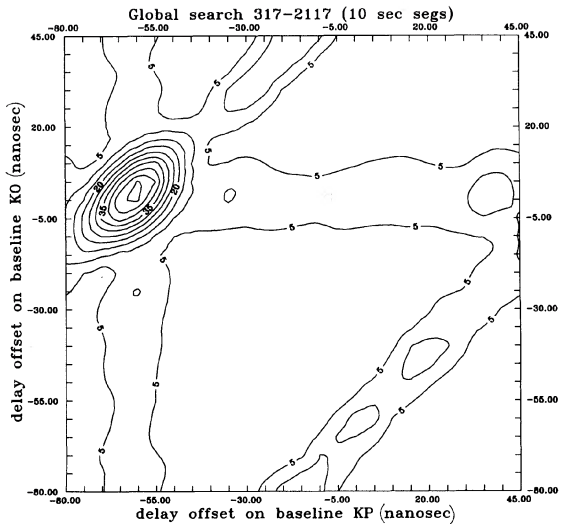

FIG. 4. A global fringe search on the Kitt Peak(P), OVRO(O), Haystack(K) triangle for a single 270 s observation at 86 GHz on the source NRAO530. The data were segmented into 10 s segments and contours of SNR$_G$ were calculated. As in the other figures, the delay is shown over a range of one ambiguity.

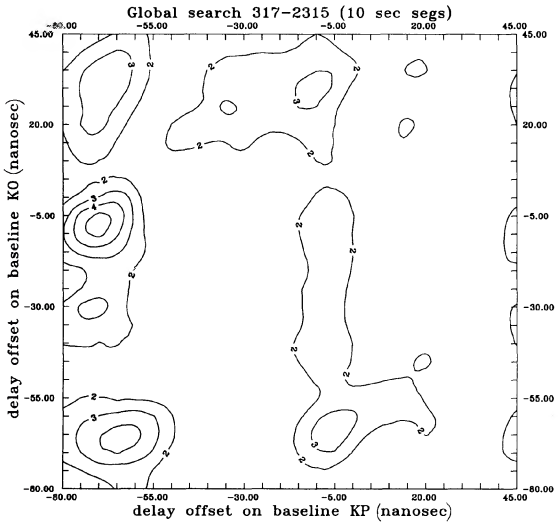

FIG. 5. A later observation on NRAO530. The value of SNR$_A$ for the individual baselines KP, KO, and PO using clocks from the global search are equal to 2.6, 0.1, and 11 respectively. The global search produces a value of SNR$_G$ equal to 5.8 at the peak, and corresponding delay offsets close to those of the earlier observation shown in Fig. 4.

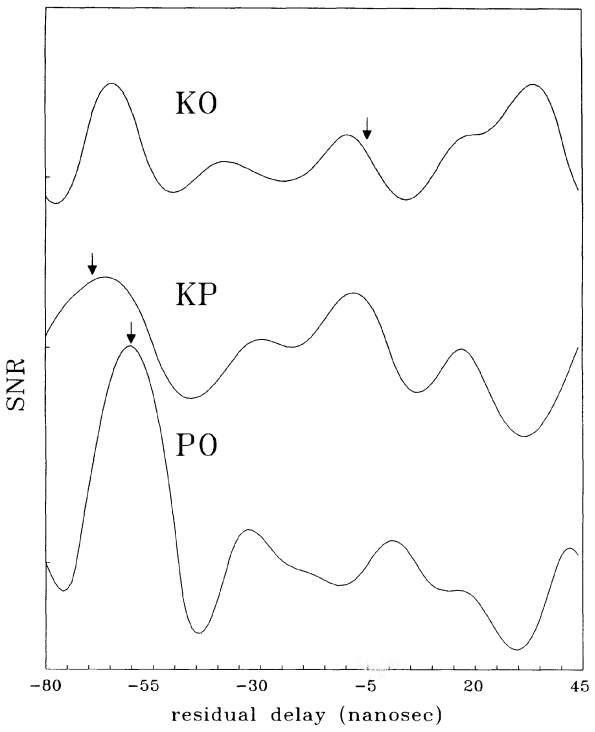

FIG. 6. The SNR of the individual baselines for the observation of Fig. 5. The delay range of 125 ns is the full range of one multiband delay ambiguity. The arrows mark the location of signal, in delay for each baseline, found in the global search.

showing that fringes cannot be correctly identified on the KO and KP baselines without the global search.

9. SUMMARY

Table 1 summarizes the detection thresholds for various detection methods relative to the conventional coherent single baseline search of a single segment. An array improves as $\sqrt{N/2}$ for a coherent search with an added improvement of better than $M^{1/4}$ for the incoherent averaging of many segments. The detection thresholds in Table 1, in all cases except the last one, result in a probability of false detection of less than 0.01% in a search of 10^6 independent values of rate and delay for each element of the array (excluding the reference element). In the last case, for incoherent averaging over time segments and baselines, the search is assumed to span only the two dimensions of right ascension and declination. The thresholds for segmented data are valid only for large M, as discussed in Sec. 3. If the segmented data thresholds are to be compared with those that could be obtained from a coherent integration of the entire dataset, without coherence loss, then these entries should be multiplied by $\sqrt{M}$.

The data used as examples in this paper were taken as part of the millimeter VLBI project to study the Galactic Center headed by Don Backer and Melvyn Wright of the University of California at Berkeley. Millimeter VLBI at the Haystack Observatory is supported by the National Science Foundation.

APPENDIX A: PROBABILITY DISTRIBUTION AND CUMMULATIVE PROBABILITY OF A^2

The probability density of A^2 can be derived from a successive convolution of Rayleigh probability densities normalized by M and with an added offset of 2. The result is the relatively simple expression

$$p(A^2)=\left(\frac{M}{2}\right)^M \frac{(A^2+2)^{M-1}}{(M-1)!}\, e^{-M(A^2+2)/2}. \tag{A1}$$

Figure 7 shows this distribution for various values of M, as well as one of the Gaussian approximations used for large M. In order to fix a detection threshold we integrate the tail of Eq. (A1), finding a value of A_0^2 that satisfies Eq. (8). This can either be done by straightforward numerical integration or by using an analytic expression for the cumulative probability which we show here:

$$p(A_0^2)=\int_{A_0^2}^{\infty} p(A^2)\,d(A^2)$$

$$=e^{-M(A_0^2+2)/2}\sum_{k=0}^{M-1}\frac{\left[\frac{M}{2}(A_0^2+2)\right]^k}{k!}. \tag{A2}$$

TABLE 1. Thresholds for various detection methods.

Method	Asymptotic Flux Density Threshold	
1-Baseline Coherent	1	1
1-Baseline Incoherent Averaging	$\frac{0.53}{M^{1/4}}$	0.14 (M=200)
3-Baseline Bispectrum	$\left(\frac{4}{M}\right)^{1/6}$	0.52 (M=200)
N-Element Array Coherent-Global Search	$\left(\frac{2}{N}\right)^{1/2}$	0.45 (N=10)
Global Search with Incoherent Averaging	$0.53\left(\frac{N^2}{4M}\right)^{1/4}$	0.05 (M=200, N=10)
Incoherent Averaging over Segements & Baselines	$0.53\left(\frac{2}{MN(N-1)}\right)^{1/4}$	0.05 (M=200, N=10)

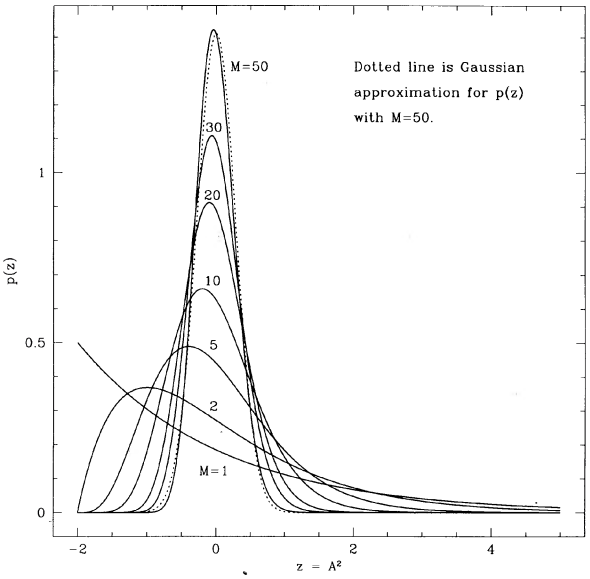

FIG. 7. Probability density functions of A^2 for incoherent averages with an increasing number of segments. For one segment ($M=1$) the density is an exponential probability density with an offset of 2. As M increases, the density approaches a Gaussian function as expected from the central limit theorem.

The value of A_0 found this way corresponds to a 1% probability of false detection in a search of 100 independent trials and agrees well with the values plotted in Fig. 1.

APPENDIX B: SEARCH OF A LARGE NUMBER OF RAYLEIGH DISTRIBUTED RANDOM VARIABLES

When a search is made over a large number of independent Rayleigh distributed random variables (see TMS, p. 264),

$$p(Z_m)=nZ_m e^{-z_m^2/2}[1-e^{-z_m^2/2}]^{n-1}, \tag{B1}$$

where $p(Z_m)$ is probability distribution of the maximum of n variables of unit variance. For large n, $\langle Z_m\rangle \approx \sqrt{2\log_e n}$ so that the expected value for the noise peak in a search on a single baseline is increased by $\sqrt{2\log_e n}$ as the result of a fringe search over n independent rates and delays. For a search over $(N-1)$ dimensions or $n^{(N-1)}$ points

$$\langle Z_m\rangle=\sqrt{2\log_e n^{N-1}}=\sqrt{2(N-1)\log_e n} \tag{B2}$$

APPENDIX C: LOWER BOUND ON THE STRENGTH OF A NOISE SPIKE

Consider the following search procedure for the purpose of evaluating the noise. First, search for a maximum in the magnitude on a single baseline from station 2 to station 1. Adjust the phase of station 2 to make the baseline phase zero. Then form a partial sum of the counter rotated cross spectra on just the baselines from station 3 to stations 1 and 2. Find a peak in the magnitude by adjusting the delay and rate of station 3 and then adjust the phase of station 3 (which is in common to both baselines) to zero the phase. Continue this process up to N stations. The noise peak for the partial sum to the kth station is

$$\frac{2\sqrt{k-1}}{N(N-1)}\sqrt{2\log_e n}, \tag{C1}$$

where the first factor is the noise caused by adding $(k-1)$ baselines and dividing by the total number of baselines. The values of delay rate and phase found using this procedure will give a sum over all baselines which is a real quantity because the phases were adjusted to make the partial sums real and will have a value equal to the sum of the magnitudes of each partial sum for the search to each added station. Thus the noise peak on the average of all baselines is

$$\frac{2}{N(N-1)}\sqrt{2\log_e n}[1+\sqrt{2}+\sqrt{3}\cdots+\sqrt{(N-1)}], \tag{C2}$$

which becomes [e.g., Gradshteyn & Ryzhik (1980)]

$$\geqslant\sqrt{2\log_e n}\,\frac{2^{3/2}}{3}\sqrt{\frac{2}{N}}, \tag{C3}$$

where the equality is approached for large N. The noise peak is increased by more than

$$\frac{2^{3/2}}{3}\sqrt{\frac{2}{N}}\sqrt{\frac{N(N-1)}{2}}=\sqrt{\frac{8(N-1)}{9}} \tag{C4}$$

as a result of the $(N-1)$ dimensional search. Since this represents only one possible prescription for finding a maximum in the average of cross-spectral functions over all baselines, it is an upper bound on the sensitivity improvement over a single baseline.

APPENDIX D: COVARIANCE MATRIX OF ESTIMATED STATION PHASES

If we use the method of least squares to estimate $N-1$ station phases represented by column vector $\mathbf{X}$ from the observed baseline phases represented by the column vector $\mathbf{Y}$

$$\mathbf{Y}=\mathbf{A}\cdot\mathbf{X}+\boldsymbol{\epsilon}, \tag{D1}$$

where the column vector $\boldsymbol{\epsilon}$ are the measurement errors on each baseline. The normal equations are

$$\mathbf{A}^{\mathbf{T}}\cdot\mathbf{X}=\mathbf{A}^{\mathbf{T}}\cdot\mathbf{Y} \tag{D2}$$

and $\mathbf{A}$ is the matrix of points for $N-1$ phases to be determined from $N(N-1)/2$ baselines. The matrix $\mathbf{A}^T\mathbf{A}$ has diagonal elements of value $N-1$ and all other elements have a value of -1. It can be shown be inspection that the inverse of $\mathbf{A}^T\mathbf{A}$, which is the covariance matrix, has diagonal elements of value $2/N$ and all other elements of value $1/N$.

REFERENCES

AIPS 1990, Astronomical Image Processing Users Manual (National Radio Astronomy Observatory, Charlottesville)
Alef, W., & Porcas, R. W. 1986, A&A, 168, 365
Clark, B. G. 1968, IEEE Transactions on Antennas and Propagation, Ap-16, 143
Cornwell, T. J. 1987, A&A, 180, 269
Dewey, R. J. 1994, AJ, 108, 337
Gradshteyn, I. S., & Ryzhik, I. M. 1980, Table of Integrals, Series, and Products, (Academic, New York), p. 1
Hanbury Brown, R. 1974, The Intensity Interferometer (Taylor and Francis, London)
Jennison, R. C. 1958, MNRAS, 118, 276
Kulkarni, S. R. 1989, AJ, 98, 1112
Rogers, A. E. E. 1991, in Frontiers of VLBI (Universal Academy Press, Tokyo), 341
Rogers, A. E. E., *et al.* 1974, ApJ, 193, 293
Rogers, A. E. E., Moffet, A. T., Backer, D. C., & Moran, J. M. 1984, Radio Science, 19, 1552
Schwab, F. R., & Cotton, W. D. 1983, AJ, 88, 688
Thompson, A. R., Moran, J. M., & Swenson, G. W. 1986, Interferometry and Synthesis in Radio Astronomy (Wiley, New York) (TMS)
Vilnrotter, V. A., Rodemich, E. R., & Dolinar, S. J. 1992, IEEE Transactions on Communications, 40, 604