



CHARA TECHNICAL REPORT

No. 69 10 JULY 1998

Optical Path Stability and Fringe Tracking in Optical Stellar Interferometry²

R. A. FRAZIN³

DEPARTMENTS OF ASTRONOMY AND ELECTRICAL AND COMPUTER ENGINEERING,
UNIVERSITY OF ILLINOIS, URBANA, IL 61801

ABSTRACT: In this paper we analyze the problems of fringe pattern estimation in optical stellar interferometry that are associated with optical path errors. Our first finding is that an optimal fringe tracking system must track with path length errors of less than about one-eighth of the coherence length. The second is that if phase closure estimation is to be achieved at the lowest possible light levels, the time scale of fringe wander must not be less than about 10% of the total observation time.

1. INTRODUCTION

A Michelson stellar interferometer operating in a phase-coherent mode uses adaptive systems to compensate for atmospheric propagation effects (Tango & Twiss 1980, Shao & Colavita 1992). One of these systems is called a *fringe tracker* which tries to compensate for the differential time delay that it takes the light to reach the different telescopes. After the beams have been corrected for the differential time delays (and wavefront distortion effects, such as tip-tilt), they are delivered to a beam combination system in which they make an interference pattern. This interference pattern gives phase and amplitude information which allows for image construction. General introductions to optical astronomical interferometry are to be found in the books by Goodman (1985) and Thompson, Moran & Swenson (1986), and the reviews by Tango & Twiss (1980) and Shao & Colavita (1992). Shao *et al.* (1988) describe a fringe tracker in their discussion on the Mark III stellar interferometer. Others have analyzed problems involved with fringe tracking using statistical methods (Lawson 1995, ten Brummelaar 1997).

In practice, the fringe tracker does not do its job perfectly, and the remaining differential time delay is called the *path length compensation error (PLCE)*. The size of the PLCE's may be many wavelengths (ten Brummelaar 1994). The PLCE is a function of time and this temporal variation is called *fringe wander*. The typical length of time for the fringes to move enough to cause significant smearing is called the *coherence time*. The term "coherence

¹Center for High Angular Resolution Astronomy, Georgia State University, Atlanta GA 30303-3083
Tel: (404) 651-2932, FAX: (404) 651-1389, Anonymous ftp: chara.gsu.edu, WWW: <http://chara.gsu.edu>

²to be published in Optics Communications, **153**, 323-330, 1998

³now at Smithsonian Astrophysical Observatory, 60 Garden St., Cambridge, MA, 02138

time” also refers to a spectral property of the light, but we will not use it in that context here. The problem of determining the optimal exposure time for the individual frames has been discussed in the contexts of speckle interferometry (O’Donnell 1980), and optical stellar interferometry (Buscher 1988, ten Brummelaar 1997). The general consensus is that the optimal integration is on the order of the coherence time. The aim of this paper is to evaluate the information content of three beams in an image-plane combination system as a function of their PLCE’s. We address two questions:

1. How much information is lost when the PLCE’s are larger than a few wavelengths?
2. How much information is lost when the entire observation period is much greater than the coherence time, making it necessary to average incoherently over many exposures?

We will answer the first question by considering the amount of information contained in a single exposure taken when the the PLCE’s are of the same order as the *coherence length* (which is the width of the the Fourier transform of the bandpass function). In order to answer the second question and not confuse the discussion with issues involved in the first question, we will consider PLCE’s which are smaller than the coherence length, but random within 0 to 2π radians. In both discussions we will assume the beams have passed through a separate fringe tracking system, which, as far as this analysis is concerned, is a “black box” that gives the beams their differential time delays.

2. THE MODEL

We will assume that the corrected light leaving the fringe trackers goes into an image-plane beam combination system, as follows: Consider a variant of Young’s double slit experiment, which has three collinear pinholes spaced at a ratio of 1:2. The light exiting the pinholes is from the fringe trackers, so it is partially coherent. The intensity pattern on the detector is an Airy pattern (the diffraction pattern of a single pinhole) multiplied by a sum of sinusoids which have spatial frequencies of ratios 1 : 2 : 3, with the highest frequency being due to the most distantly separated pair. Many optics texts have a picture of such a pattern made with two beams in connection a discussion of partial coherence, two examples are the texts by Goodman (1985) and Born & Wolf (1965). This set-up is a simple example of a non-redundant beam combining system. Non-redundant beam combination has been discussed in the contexts of aperture masking and optical stellar interferometry by at least several authors (Haniff 1989, Buscher 1988, ten Brumelaar 1994).

Fringe wander causes the interference pattern to move across the detector. In this paper we will take the point of view that the total amount of time for all of the measurements is fixed at T . This time is divided up into m equal subintervals of duration $t = T/m$, which are of the order of the coherence time. We will make the simplifying assumption that there is no fringe wander in each of the m subintervals, but allow each subinterval to have its own independent PLCE’s. We will assume that the exposures are statistically identical.

We will assume that a two-dimensional detector array is centered on the diffraction envelope (the Airy pattern due to the pinholes), which is given by the real, positive function $Q'(x, y)$, and normalized to unity. Thus, $\int Q'(x, y) dx dy = 1$. If the light and dark bands are parallel to the y -axis, we may integrate all of the photon counts over y and define a new diffraction envelope function $Q(x) = \int Q'(x, y) dy$.

OPTICAL PATH STABILITY

If the light beams enter the beam combiner through identical pinholes, the diffraction envelope will be given by the formula for Fraunhofer diffraction at a circular aperture (Born & Wolf 1965). Thus, $Q'(x, y) \propto J_1(\lambda\sqrt{x^2 + y^2})/\lambda\sqrt{x^2 + y^2}$, where J_1 is the Bessel function of the first kind, x and y are unitless and λ is a unitless scale factor. Fortunately, the side-lobes of this function are small and a two-dimensional Gaussian fits the main lobe reasonably well. We will approximate the one-dimensional diffraction envelope by

$$Q(x) = \frac{4}{\sqrt{\pi}} \exp[-16x^2]. \quad (1)$$

Equation 1 is arranged so that $Q(1/2) = .0183Q(0)$. For convenience, we will take the length of the detector to be 1. In our discussion of Question 2 we will use assume that $Q(x) = 1$.

We will allow the light be polychromatic, but confined to a band $\Delta\nu$ about a frequency ν_0 . For simplicity, we will assume the mutual coherence functions are not functions of wavelength. When the image is a very strong function of wavelength, or when the bandpass and baselines are large enough so that the spatial frequencies vary significantly over the bandpass, a more complicated analysis is required. The *complex visibilities* are the normalized mutual coherence functions (Goodman 1985), and they have magnitudes V_{12} , V_{23} and V_{13} and phases ϕ_{12} , ϕ_{23} and ϕ_{13} , where $\phi_{ij} = -\phi_{ji}$. Since only the differences in compensation errors are important, we only need two parameters to describe them in a three-beam system. The PLCE's in the k^{th} time subinterval are $\delta_{12,k}$ and $\delta_{13,k}$, and they are measured relative to the first beam as the notation indicates. The expected number of photons from each beam in the entire observation period is N_0 . We will assume that this number can be measured accurately, and we will take it to be a known quantity. The column vector of unknown parameters is

$$\boldsymbol{\theta} \equiv [V_{12}, V_{23}, V_{13}, \phi_{12}, \phi_{23}, \phi_{31}, \delta_{12,1}, \dots, \delta_{12,m}, \delta_{13,1}, \dots, \delta_{13,m}]^T, \quad (2)$$

where T denotes the transpose operator. $\boldsymbol{\theta}$ represents the minimal set of parameters which describe the intensity pattern.

P_{12} , P_{23} and P_{13} are the number of periods of each sinusoid that fit on the detector array at the central frequency; they are proportional to the spatial frequencies of the fringes. The ratio $P_{12} : P_{23} : P_{13}$ is 1:2:3 (i.e., $P_{12} = \frac{1}{3}P_{13}$, and $P_{23} = \frac{2}{3}P_{13}$), which is the same as the distance ratios of the three pairs of pinholes. We will assume a square bandpass function, $\frac{1}{\Delta\nu} \Pi(\frac{\nu-\nu_0}{\Delta\nu})$, where Π is the rectangle function as described by Bracewell (1978). The expected value of the intensity pattern on the detector may be derived from Walkup and Goodman's quasi-monochromatic expression by multiplying the interference terms by a phase shift that is proportional to the temporal frequency and integrating (Goodman 1985). The expected value of the number of photons in the pixel centered at x_j in the k^{th} time subinterval turns out to be

$$\begin{aligned} \langle N_{jk} \rangle &\equiv \mathcal{E}[N_k(x_j)] \\ &= Q(x) \left[\begin{aligned} &3 \frac{N_0}{mn} + 2 \frac{N_0}{mn} V_{12} \cos[2\pi(P_{12}x_j + \delta_{12,k}) + \phi_{12}] \text{sinc}[\frac{\Delta\nu}{\nu_0}(P_{12}x_j + \delta_{12,k})] \\ &+ 2 \frac{N_0}{mn} V_{23} \cos[2\pi(P_{23}x_j + \delta_{13,k} - \delta_{12,k}) + \phi_{23}] \text{sinc}[\frac{\Delta\nu}{\nu_0}(P_{23}x_j + \delta_{13,k} - \delta_{12,k})] \\ &+ 2 \frac{N_0}{mn} V_{13} \cos[2\pi(P_{13}x_j - \delta_{13,k}) + \phi_{31}] \text{sinc}[\frac{\Delta\nu}{\nu_0}(P_{13}x_j - \delta_{13,k})] \end{aligned} \right] \quad (3) \end{aligned}$$

In Equation 3 n is the number of pixels, and m is the number of time subintervals in the observation period. Thus, $1 \leq j \leq n$, $1 \leq k \leq m$, and $-1/2 \leq x_j \leq 1/2$. $\mathcal{E}[\]$ is the expectation operator, which amounts to averaging over a (possibly multi-dimensional) Poisson distribution in this paper.

3. PHASE CLOSURE

The concept of *phase closure* has played an important role in astronomical aperture synthesis. A closure phase is a linear combination of phases that is independent of the PLCE's. Phase closure is discussed in considerable detail by Thompson, Moran, & Swenson (1986), and the problem of estimating closure phases in the context of optical aperture synthesis has received attention from a number of authors (Buscher 1988, Kulkarni 1989, Haniff 1989, Roggemann 1996). The best way to illustrate the idea is by example. Let a_{12} , a_{23} and a_{31} stand for the phases of the cosines in Equation 3, e.g., $a_{23} = 2\pi(\delta_{13,k} - \delta_{12,k}) + \phi_{23}$. Evidently,

$$a_{12} + a_{23} + a_{31} = \phi_{12} + \phi_{23} + \phi_{31} \equiv \psi. \quad (4)$$

ψ is independent of the compensation errors and is called a *closure phase*. We should note that one must take care to make sure that the measured value of ψ is not incorrect by a multiple of 2π . One way around this is to calculate the value of ψ by multiplying the observed complex visibilities (Kulkarni 1989).

4. THE ESTIMATION PROBLEM

4.1. Cramer-Rao bounds

We will address Question 1 of the introduction with a technique from estimation theory. The general results in the section, i.e. those that are not specific to our problem, can be found in books on estimation and detection theory, for example, the texts by Van Trees (1968) and Helstrom (1995).

The Cramer-Rao inequality allows us to put a lower bound on the covariance matrix of *any unbiased estimator*. This lower bound is called the Cramer-Rao bound (CRB), and we will use it to estimate the variance of the measurement of the closure phase ψ , which is a function of the parameter vector $\boldsymbol{\theta}$. Under fairly general circumstances, in the limit of high signal-to-noise, this bound is tight and the maximum likelihood estimator (MLE) achieves it (the MLE is asymptotically unbiased; see Van Trees 1968, Helstrom 1995). Other bounding procedures are available, many of which have important advantages over the CRB in certain regimes. However, all of them are more difficult to calculate than the CRB. The paper by Bell *et al.* has a discussion of such procedures in the introduction (Bell et al. 1997). The Cramer-Rao inequality states that for any suitably dimensioned column vector \mathbf{a} ,

$$\mathbf{a}^T \mathcal{E}[(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T] \mathbf{a} \geq \mathbf{a}^T \mathbf{F}^{-1} \mathbf{a}, \quad (5)$$

where $\hat{\boldsymbol{\theta}}$ is the estimator (e.g., the discrete Fourier transform (DFT)), $\mathcal{E}[(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T]$ is the covariance matrix and \mathbf{F} is *Fisher's information matrix*. The entries in Fisher's information matrix are given by

$$F_{ij} = -\mathcal{E}\left[\frac{\partial^2 \ln P(\mathbf{v} | \boldsymbol{\theta})}{\partial \theta_i \partial \theta_j}\right], \quad (6)$$

where $P(\mathbf{v} | \boldsymbol{\theta})$ is the conditional probability of measuring the data vector \mathbf{v} given the parameter vector $\boldsymbol{\theta}$, of which θ_i is the i^{th} component.

In our problem the data vector \mathbf{v} corresponds to the vector of measured photo-counts with components N_{jk} , and $\boldsymbol{\theta}$ was defined in Equation 2. Since determining the values of the

PLCE's is not the ultimate goal of most stellar interferometry experiments we will call $\delta_{12,1}, \dots, \delta_{12,m}, \delta_{13,1}, \dots, \delta_{13,m}$ *nuisance parameters*. In the presence of only photon noise, the conditional probability in Equation 6 is given by Poisson counting statistics (Goodman 1985). Since each of the observations N_{jk} is statistically independent, their probability density functions $P_{jk}(N_{jk})$ multiply, and we have

$$\ln P(\mathbf{v} | \boldsymbol{\theta}) = \sum_{k=1}^m \sum_{j=1}^n [N_{jk} \ln(\langle N_{jk}(\boldsymbol{\theta}) \rangle) - \langle N_{jk}(\boldsymbol{\theta}) \rangle - \ln(N_{jk}!)], \quad (7)$$

where $\langle N_{jk} \rangle$ was given by Equation 3.

To calculate the bounds on a quantity which depends on the parameter vector $\boldsymbol{\theta}$, such as the closure phase ψ , we must transform the inverse of the Fisher information matrix. From the definition of ψ in Equation 4, we can write $\psi = \psi(\boldsymbol{\theta})$. The bound is given by (Walkup & Goodman 1973)

$$\mathcal{E}[(\hat{\psi} - \psi)^2] = \sum_i \sum_j \frac{\partial \psi}{\partial \theta_i} (\mathbf{F}^{-1})_{ij} \frac{\partial \psi}{\partial \theta_j}. \quad (8)$$

In this particular problem, the CRB's on the estimators of the non-nuisance parameters are independent of the number of subintervals m , and we demonstrate this fact in the Appendix. This implies that we cannot use the CRB to calculate the dependence of the error covariance matrix on m . Clearly, the errors must ultimately increase with m because when $m = 3N_0$ there is only about one photon per interval, a situation likely to lead to less than satisfactory results. The CRB is significantly less than the greatest lower bound when the signal-to-noise ratio falls below a certain value. This value is called the *threshold* (Seidman 1970). We must use another procedure to analyze the performance of the system below the threshold.

4.2. The DFT closure phase estimator

In our problem, the CRB's on the non-nuisance parameters is independent of the number of subintervals m (see Appendix), so we must use another method to address Question 2 in the Introduction. We may gain some insight into the problem by assuming a constant diffraction envelope and no leakage (which requires that an integer number of fringe periods fit across the detector; see Goodman 1985, Nakadate 1988), setting the bandwidth to zero, and using Roggemann's analytical results. Roggemann's method is an implementation of the ideas presented by Walkup (1973), Goodman (1985), and Kulkarni (1989). The probability distribution of the discrete Fourier transform (DFT) estimate of the closure phase is a function of the parameter κ , which itself is a function of the visibility magnitudes and light level. For simplicity we will assume that all of the visibility magnitudes are the same. Using our notation and making some obvious modifications, the variance of the DFT closure phase estimator (setting the reference phase to zero) is given by Roggemann (1996)⁴.

$$\text{Var}[\hat{\psi}] = \int_{-\pi}^{\pi} \hat{\psi}^2 P_{\Psi}(\hat{\psi}; \kappa) d\hat{\psi}, \quad (9)$$

⁴Roggemann's Equation 22, the probability density function of the DFT phase closure estimator, has an error. The argument of the error function needs to be divided by $\sqrt{2}$.

where $P_\Psi(\hat{\psi}; \kappa)$ is the probability of measuring $\hat{\psi}$, given the parameter κ , and κ is given by (Roggemann 1996)

$$\frac{m}{\kappa^2} = \frac{3m}{V_r^2 N_{0e}} + \frac{6m^2}{V_r^4 N_{0e}^2} + \frac{4m^3}{V_r^6 N_{0e}^3}. \quad (10)$$

Since the Roggemann's formula for κ was derived for a two-beam system, it needs to be modified so that the two beam system has the same number of photons and the same fringe contrast as our three-beam system. V_r is the reduced fringe contrast, which takes into account the fact that when three or more beams form fringes on a detector, the effective contrast of the fringes decreases. N_{0e} is the enhanced number of photons, it makes it so that the total photon count is the same. In our case there are three beams, so $V_r = 2V/3$ and $N_{0e} = 3N_0/2$. It is clear from Equation 10 that when N_0/m is large, κ is independent of m . This behavior is consistent with the results of our CRB calculations discussed above, and numerical experiments show quantitative agreement in the low visibility, high light level regime. In fact, we can say that when either the second or third terms in Equation 10 are comparable to the first, the system is operating below threshold (see Section 4.1).

5. RESULTS OF NUMERICAL CALCULATIONS

5.1. Mode coupling

Our numerical calculations show that the CRB on the variance of the estimate for any one parameter (e.g., V_{12}) depends on the values of the phases and amplitudes of the other fringes. We will call this effect *mode coupling* for its analogy to problems in nonlinear dynamics. The existence of mode coupling may be bit of a surprise, and it seems contrary to the result by Walkup & Goodman. However, it is important to note that the maximum likelihood estimate (which has a relationship to the CRB as described above) and the DFT are only equivalent in the low-visibility regime (Walkup & Goodman 1973). Mode coupling probably is the result of the multiplicative nature of photon noise.

Our numerical experiments indicate that the effect is no greater than about 5% when the visibilities are near 0.1, and it insignificant for visibility values near 0.01. However, for visibilities near 0.5, the CRB's change by as much as a factor two as the relative phases vary. Since N_0 factors out of the Fisher information matrix (see Appendix), our numerical experiments could not determine the extent to which mode coupling depends on the number of photons. Studying mode coupling would require a more sophisticated study which compares the DFT and ML estimators at moderate visibility ranges with varying light levels.

5.2. CRB results

Before describing Figures 1 and 2 in detail, a few words are in order. Since the expected number of photons per beam N_0 factors out of Fisher's information matrix (see Appendix), we set $N_0 = 1$. In order to remove the effects of mode coupling, all of the calculations that follow in this section have the visibilities all set to 0.01. The reader may use a simple, approximate procedure to "move" these figures into the visibility and intensity regime of interest. All that is required is to multiply the variances by $N_0^{-1}(.01/V)^2$, where N_0 , and V are the desired number of photons and visibility, respectively. Division by the visibility ratio squared is reasonable because the variance DFT estimator scales as V^{-2} , and the CRB and DFT variance are closely related as discussed above. This approximate adjustment breaks down in the low intensity and high visibility regimes, for reasons discussed above. The

reader may also wonder why the CRB's are greater than the maximum possible variance of the parameters since the variance of any reasonable estimator cannot exceed $\pi^2/3$. This is because the CRB is only a measure of the local curvature of the probability distribution, which is insensitive to the boundaries of the parameter support in the high SNR regime. By the words "period number" we mean P_{13} and the reader will recall that $P_{12} = P_{13}/3$ and $P_{23} = 2P_{13}/3$.

Figures 1 and 2 show how the CRB on the variance of the closure phase estimator changes as a function of period number for different PLCE values. For simplicity we have set $\delta_{12} = \delta_{13} \equiv \text{PLCE}$. The curves in Figures 1 and 2 are calculated for 1% and 10% bandpasses, respectively. We will take the number of oscillations in the central lobe to be the definition of the coherence length. For example, the central lobe of the fringe envelope contains 200 oscillations when the bandpass is 1% (assuming a sinc function, as we have used here), so the coherence length would be equal to 200 wavelengths. Figure 1 clearly demonstrates that the variance doubles its minimum value and rapidly increases the PLCE increases past 1/8 of the coherence length, or 25 wavelengths, unless the period number is large. Figure 2 leads us to the same conclusion about what happens to the information content as the PLCE increases beyond 1/8 of the coherence length.

The reason that the variances in Figures 1 and 2 increase as a function of period number for PLCE's of zero is that more of the fringe envelope fits into the diffraction envelope while the total number of photons is held constant. This increases the variance because the fringe contrast is small in the outer regions of the fringe envelope. The fact that the amount of the fringe envelope that fits into the diffraction envelope increases with period number also explains why the variance decreases as a function of period number when the PLCE's are large: When the period number is small only the low contrast part of the fringe envelope fits into the diffraction envelope, but the higher contrasts regions of the fringe envelope "move" into the diffraction envelope as the period number increases.

5.3. DFT results

Figure 3 shows how the variance of the DFT closure phase estimator depends on the number of subintervals m and the number of photons per beam N_0 . The y-axis is the variance of the DFT closure phase estimator and the x-axis is m . The numbers next to the curves correspond to N_0 . This calculation was done with all three visibility magnitudes set to 0.1. This is reasonable, because any imaging interferometer must be able to perform in the moderate visibility regime. The CRB's predict that the phase closure variance of an efficient estimator does not depend on m , and Figure 3 shows this to be the case when N_0/m is large. When looking at Figure 3, the reader should keep in mind that variance of 0.1 rad² implies an RMS phase error of about 18°, which is on the order of the maximum tolerable error. It is clear from this graph that if N_0 is such that it produces a variance of about 0.1 when $m = 1$, the variance does not double its minimum value until m is over 10.

6. CONCLUSION

This paper helps to place performance standards on fringe tracking systems which may be used to optimize ground-based stellar interferometers. As we have stated several times, one of our results is that the PLCE must be less 1/8 of the coherence length or the information losses will be severe. Figures 1 and 2 show that, in principle, it is possible to make up

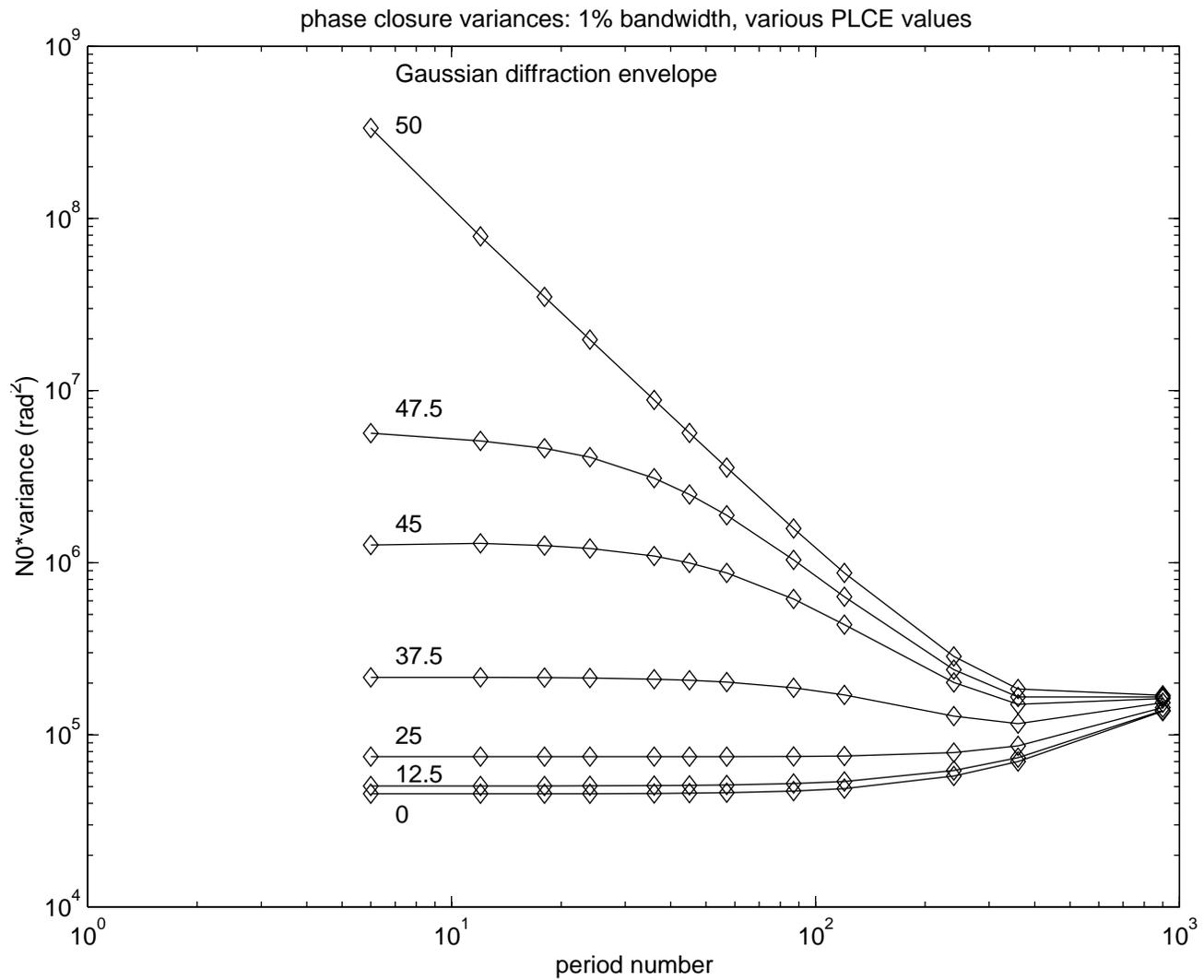


FIGURE 1. CRB of phase closure variance vs. period number. Each curve is for a different PLCE, according to its label. All three visibilities are set to 0.01. The bandpass is 1%, and a Gaussian diffraction envelope is assumed.

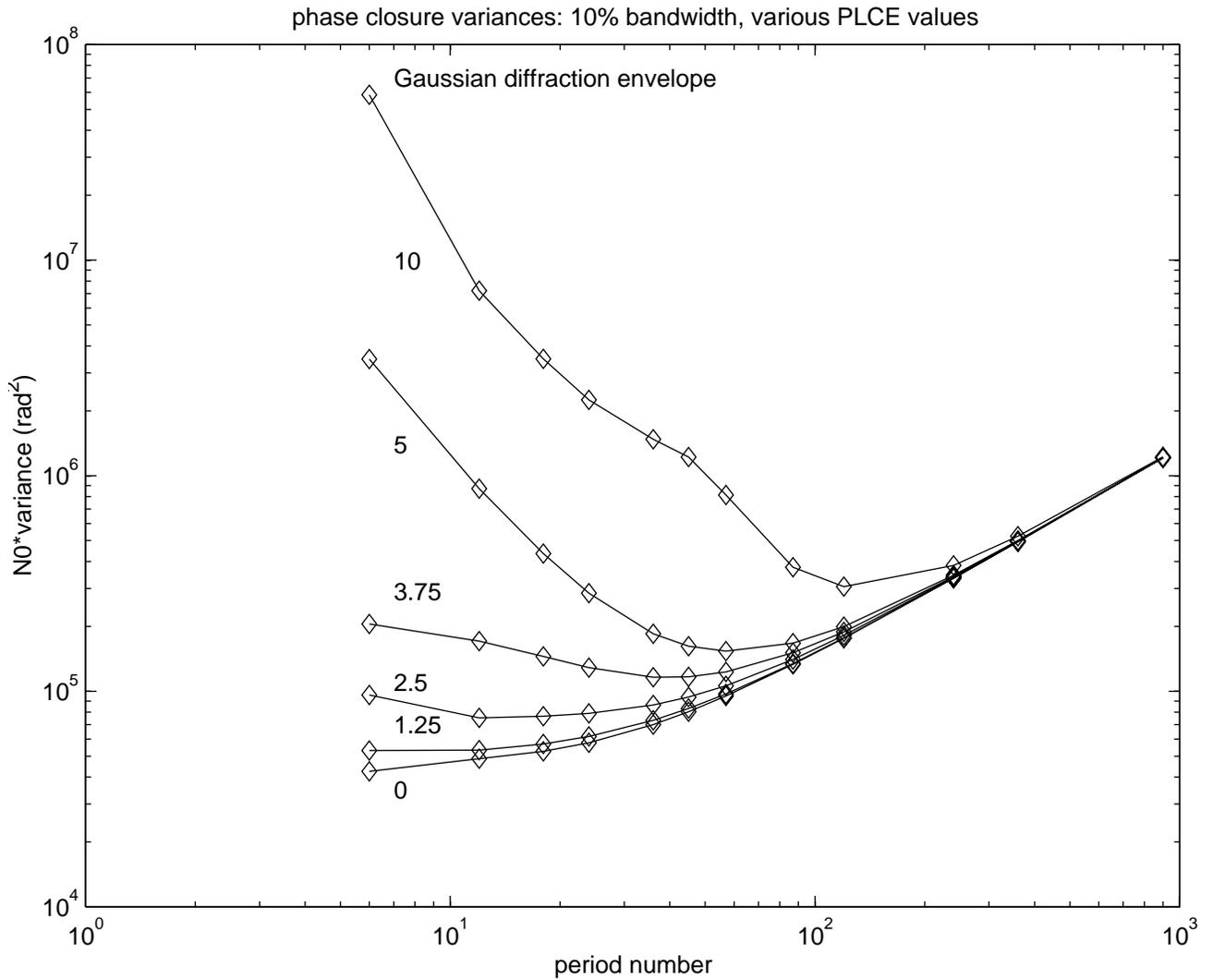


FIGURE 2. CRB of phase closure variance vs. period number. Each curve is for a different PLCE, according to its label. All three visibilities are set to 0.01. The bandpass is 10%, and a Gaussian diffraction envelope is assumed.

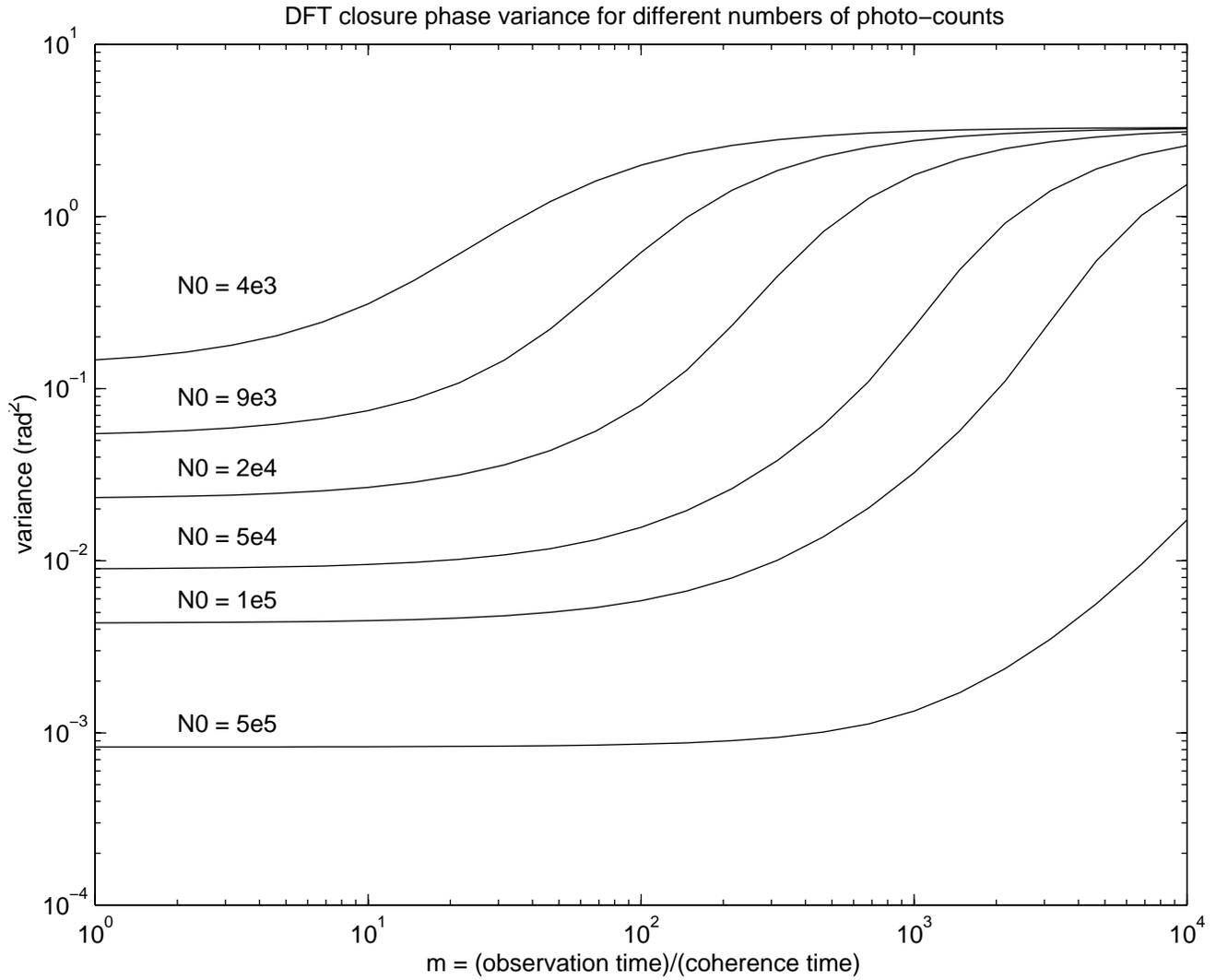


FIGURE 3. DFT phase closure variance vs. $m = \frac{\text{observation time}}{\text{coherence time}}$. All three visibilities are set to 0.1. Each curve is for a different number of photons, N_0 , according to its label. The bandpass is infinitesimal and the diffraction envelope is taken to be a constant.

for large PLCE's by increasing the period number. However, the required period numbers are large and increase with decreasing bandwidth. These figures also show that the phase closure variance should not be a strong function of the period number as long as the PLCE's are smaller than about one-quarter of the coherence length. Figure 3 shows that one can obtain useful phase closure measurements with the minimum possible number of photons when the coherence time is greater than about 10% of the observation time. There are two applications of the work presented here:

- A comparison of Figures 1 and 2 suggests that dispersing the light spectrally within beam combination system (which would correspond to separating the wavelengths in the y - direction in this example) will help to reduce the information loss associated with large PLCE's. This work also suggests that a quantitative information theoretical model, more elaborate than that given here, might prove beneficial when designing a beam combination system.
- Comparing Figures 1, 2 and 3 to "light level vs. rms PLCE" and "light level vs. coherence time" curves of a fringe tracking system can help determine the optimal way to divide the available light between the beam combination and the fringe tracking systems.

The analysis in this paper leads to two questions whose answers will determine the future of ground based optical stellar interferometry. The first is, "What are the signal requirements (i.e., light level and degree of wavefront correction) for fringe tracking system to perform as outlined above?" The next question is a logical extension of first, "How does one make an adaptive optics system which meets the requirements of such a fringe tracker?" ten Brummelaar *et al.* have done some work which goes toward answering the first question (ten Brummelaar et al. 1995).

7. ACKNOWLEDGMENTS

The author would like to thank R. E. Blahut, L. Roberts, E. C. Sutton, T. A. ten Brummelaar, and W. A. Traub for their useful comments. This work would not have been possible without George Swenson's support and patience. I am deeply indebted to Yoram Bresler for his scientific guidance. U.S. Army Construction Engineering Research Laboratories contract #DACA-88-95-C-0100 provided the funds for this research.

8. REFERENCES

- Bell, K. L., Steinberg, Y., Ephraim, Y., & Van Trees, H. L., "Extended Ziv-Zakai lower bound for vector parameter estimation," *IEEE Trans. Inform. Theory*, IT-43, 624-637 (1997).
- Born, M., & Wolf, E., *Principles of Optics*, 3rd ed. (Pergamon Press, New York 1965).
- Bracewell, R. N., *The Fourier Transform and Its Applications*, 2nd ed. (McGraw-Hill, New York 1978).
- ten Brummelaar, T. A., "Visible light imaging," Appendix K in *The CHARA Array: Final Report to NSF: Cooperative Agreement 90-08941*, Atlanta (1994). This and other CHARA documents are available on the World Wide Web at <http://www.chara.gsu.edu/>
- ten Brummelaar, T. A., "Correlation measurement and group delay tracking in optical stellar interferometry with a noisy detector," *Mon. Not. R. Astr. Soc.*, 285, 135-150 (1997).
- ten Brummelaar, T. A., Bagnuolo, W. J., & Ridgway, S. T., "Strehl ratio and visibility in long-baseline stellar interferometry," *Opt. Lett.*, 20, 521-523 (1995).
- Buscher, D. F., "Optimizing a ground-based optical interferometer for sensitivity at low light levels," *Mon. Not. R. Astr. Soc.*, 235, 1203-1226 (1988)
- Goodman, J. W., *Statistical Optics*. (John Wiley & Sons, Inc., New York, 1985).
- Haniff, C. A., "Phase closure imaging - theory and practice," in *Diffraction Limited Imaging with Very Large Ground Based Telescopes* (Kluwer Academic Publishers, Boston, 1989)
- Helstrom, C. W., *Elements of Signal Detection and Estimation*. (Prentice-Hall, Englewood Cliffs, New Jersey, 1995).
- Kulkarni, S. R., "Self-noise in interferometers: radio and infrared," *Astron. J.*, 98, 1112-1130 (1989).
- Lawson, P. R., "Group-delay tracking in optical stellar interferometry with the fast Fourier transform," *J. Opt. Soc. Am. A*, 12, 366-374 (1995).
- Nakadate, S., "Phase detection of equidistant fringes for highly sensitive optical sensing. I. Principle and error analyses," *J. Opt. Soc. Am. A*, 5, 1258-1264 (1988).
- O'Donnell, K. A., & Dainty, J. C., "Space-time analysis of photon-limited stellar speckle interferometry," *J. Opt. Soc. Am.*, 70, 1354-1361 (1980).
- Roggemann, M. C., "Photon-noise limits to the detection of the closure phase in optical interferometry," *Appl. Opt.*, 35, 1809-1814 (1996).
- Seidman, L. P., "Performance limitations and error calculations for parameter estimation," *Proc. IEEE*, 58, (1970).
- Shao, M., & Colavita, M. M., "Long-baseline optical and infrared stellar interferometry," *Ann. Rev. Astron. Astrophys.*, 30, 457-498 (1992)
- Shao, M., *et al.*, "The Mark III stellar interferometer," *Astronomy and Astrophysics*, 193, 357-371, (1988).
- Tango, W. J., & Twiss, R. Q., "The modern Michelson stellar interferometer," in *Progress in Optics vol. XVII*, pp. 239-277 (1980).
- Thompson, A. R., Moran, J. M., & Swenson, G. W., *Interferometry and Synthesis in Radio Astronomy*. (Wiley, New York, 1986).
- Van Trees, H. L., *Detection, Estimation, and Modulation Theory, vol. I*. (Wiley, New York, 1968).
- Walkup, J. F., & Goodman, J. W., "Limitations of fringe-parameter estimation at low light levels," *J. Opt. Soc. Am.*, 63, 399-407 (1973).

Appendix: Demonstration that the Cramer-Rao bound is independent of the number of subintervals

In the analysis to follow, the three-beam experiment is unnecessarily complicated and considering a two-beam experiment is sufficient. The equation that follows is a simplified version of Equation 3. There is no need for the subscripts that contain the beam indices, so we drop them, e.g., $V_{12} \rightarrow V$. The result is

$$\langle N_{jk} \rangle = Q(x) \left\{ 2 \frac{N_0}{mn} + 2 \frac{N_0}{mn} V \cos[2\pi(Px_j + \delta_k) + \phi] \operatorname{sinc} \left[\frac{\Delta\nu}{\nu_0} P(x_j + \delta_k) \right] \right\}. \quad (11)$$

For the sake of further algebraic simplicity, we will also assume that the parameter V is known ahead of time, which reduces the number of parameters that we must determine. We used numerical experiments to verify its extension to the full three-beam case. This example also serves to illustrate the procedure for calculating CRB's.

Since we assume that we already know V and need to determine bounds on ϕ and $\{\delta_1, \dots, \delta_m\}$, the vector of unknowns is

$$\boldsymbol{\theta} = [\phi, \delta_1, \dots, \delta_m]^T. \quad (12)$$

Using Equations 7 and 6, we may write the following expression for the elements of the Fisher information matrix

$$f_{il} \equiv (\mathbf{F})_{il} = -\mathcal{E} \left[\frac{\partial^2 \ln P(\mathbf{v} | \boldsymbol{\theta})}{\partial \theta_i \partial \theta_l} \right] = \sum_{k=1}^m \sum_{j=1}^n \frac{1}{\langle N_{jk} \rangle} \frac{\partial \langle N_{jk} \rangle}{\partial \theta_i} \frac{\partial \langle N_{jk} \rangle}{\partial \theta_l}. \quad (13)$$

The matrix described by Equation 13 has a very simple structure. We assumed that the observations in each time subinterval are independent, which makes the δ_{k1} - δ_{k2} terms zero unless $k1 = k2$. If the δ 's are not large enough to shift the intensity pattern through a significant fraction of its fringe envelope, each of these diagonal terms will be the same (each term is a sum over all of the pixels). Furthermore, $\langle N_{jk} \rangle$ and its first derivatives are all proportional to N_0/m . A glance at Equation 13 shows that each of the diagonal elements, except the first, can be written as cN_0/m , where c is a constant. Similar reasoning shows that the ϕ - δ matrix elements are all equal and have the form bN_0/m , where b is a constant. The ϕ - ϕ element is a sum with mn terms, each of which is proportional to N_0/m . Therefore, it does not depend on m and can be written in the form N_0a , where a is another constant. The $(m+1) \times (m+1)$ Fisher information matrix takes the form

$$\mathbf{F} = N_0 \begin{pmatrix} a & b/m & \dots & \dots & b/m \\ b/m & c/m & 0 & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ b/m & 0 & \dots & 0 & c/m \end{pmatrix}. \quad (14)$$

At this point it is also clear that \mathbf{F}^{-1} scales as $1/N_0$. The final step is to show that the 1-1 element of \mathbf{F}^{-1} is independent of m . This involves almost no more effort than solving a 2×2 system of equations.