DIFFRACTION-LIMITED IMAGING ON THE 200-INCH TELESCOPE

Thesis by

Tadashi Nakajima

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1989

(Submitted May 23, 1989)

Acknowledgements

I would like to add a history of the project seen from a student's perspective to these acknowledgements. One morning in July 1985, a few weeks after I arrived here, my adviser, Gerry Neugebauer suggested that I talk to Tony Readhead, since they were planning an experiment on *infrared interferometry* at Palomar in which I would be interested. Although that experiment and my entire thesis work turned out to be on optical interferometry, I now realize Gerry's intuition or instinct for finding a promising field. I am indebted to Gerry in many ways, but especially for his illuminating comments on drafts of my papers. Tony, the initiator of the project, spent a lot of time on my early education of VLBI imaging through 'blind tests,' in spite of his busy life as the then director of the Owens Valley Radio Observatory. Initial suggestions on simulations of atmospheric disturbances were given by Tony. Tim Pearson also helped me understand VLBI imaging and assisted with software works. Bev Oke often gave me words of technical advice, especially of optics. The project was accelerated and expanded by Shri Kulkarni who joined the project in 1986 and has been leading it since then. From Shri I received a lesson in how to get things going and get them done. He shared the hard experience with the water-cooling system of the Coude Camera with me. Peter Gorham and Tom Prince joined the project in 1987 to work on a speckle imaging technique which required an intensive computation on the NCUBE. Peter recently succeeded in producing images after a long struggle with instrumentally corrupted data. I will miss frank and direct discussions with Peter and Shri in the future. It was helpful to receive comments on the draft of my paper from Tom, who thoroughly checked the internal consistency. The number of graduate students doubled, to my relief, when Andrea Ghez joined the project in late 1987. She has been working mainly on laboratory experiments which we lacked before. I wish Andrea good luck in extracting scientific results from this rather tough project. Undergraduates Andrew Lee and Glenn Smith surprised me with their skill in programming on

peripherals of the observing system. Glenn, as a SURF student, also developed an excellent algorithm for speckle imaging. Our observing and data-acquisition systems were professionally built and programmed by Earl Emery, Fred Harris and Rick Nadler. Our early runs at Palomar often created chaos which was settled by their hard work and with the patience of the mountain staff. I acknowledge the financial support of the W.M. Keck Foundation and the Schlumberger Fellowship.

In my everyday life, I have been a lunch-time infrared and submillimeter astronomer. Conversations with Tom Soifer, Dave Sanders and Dave Carico on IRAS galaxies, kept my original interest in infrared astronomy. Discussions with James Graham, Keith Matthews and Dan Watson on astronomy or astronomical techniques have been a lot of pleasure. I wish John McGill were here and enjoyed infrared imaging. Lunches at Eddies Market and Friday beers with Thomas Büttgenbach, Jocelyn Keen, Edith Falgarone, James Graham, Eric Grossman, Tom Phillips, Colin Masson, Keith Matthews, Anneila Sargent, Nick Scoville, Gene Serabyn, Chris Walker, Dan Watson and Mike Wengler, kept my sanity. I am proud of having been a player of the Downs Syndrome, the anti-champion of the GSC league. I thank Pat Neil, Karen Waters and Carrah Wright for assistance with solving official matters.

I owe a high-school classmate, Shuichi Tasaki, and friends and advisers in Kyoto University for my basic education in physics and astronomy. Finally I would like to thank my parents, Hidekichi and Toshiko, and my sisters, Tomoko and Sachiko, for their continuous encouragement.

Abstract

We have used the technique of non-redundant masking at the Palomar 200-inch telescope and radio VLBI imaging software to make optical aperture synthesis maps of two binary stars, β Corona Borealis and σ Herculis. The dynamic range of the map of β CrB, a binary star with a separation of 230 milliarcseconds is 50:1. For σ Her, we find a separation of 70 milliarcseconds and the dynamic range of our image is 30:1. These demonstrate the potential of the non-redundant masking technique for diffraction-limited imaging of astronomical objects with high dynamic range.

We find that the optimal integration time for measuring the closure phase is longer than that for measuring the fringe amplitude. There is not a close relationship between amplitude errors and phase errors, as is found in radio interferometry. Amplitude self calibration is less effective at optical wavelengths than at radio wavelengths. Primary beam sensitivity correction made in radio aperture synthesis is not necessary in optical aperture synthesis.

The effects of atmospheric disturbances on optical aperture synthesis have been studied by Monte Carlo simulations based on the Kolmogorov theory of refractive-index fluctuations. For the non-redundant masking with r_c-sized apertures, the simulated fringe amplitude gives an upper bound of the observed fringe amplitude. A smooth transition is seen from the non-redundant masking regime to the speckle regime with increasing aperture size. The fractional reduction of the fringe amplitude according to the bandwidth is nearly independent of the aperture size. The limiting magnitude of optical aperture synthesis with r_c-sized apertures and that with apertures larger than r_c are derived.

Monte Carlo simulations are also made to study the sensitivity and resolution of the bispectral analysis of speckle interferometry. We present the bispectral modulation transfer function

and its signal-to-noise ratio at high light levels. The results confirm the validity of the heuristic interferometric view of image-forming process in the mid-spatial-frequency range. The signal-to-noise ratio of the bispectrum at arbitrary light levels is derived in the mid-spatial-frequency range.

The non-redundant masking technique is suitable for imaging bright objects with high resolution and high dynamic range, while the faintest limit will be better pursued by speckle imaging.

Table of Contents

Chapter I.

Introduction

"The new theory, which connects the wave function with probabilities for one photon, gets over the difficulty by making each photon go partly into each of the two components. Each photon then interferes only with itself. Interference between two different photons never occurs."

—— P.A.M. Dirac, *The Principle of Quantum Mechanics*

1. Purposes of Research

Progress and discovery in astronomy always take place along with the improvement in sensitivity and spatial resolution of an observing technique at any wavelength.

High-resolution interferometric observation from the ground is a technique to extract spatial structure of an astronomical object by a correlation analysis of an ensemble of instantaneous interference patterns made by transient partially coherent light. This overcomes the 'seeing' limit of the direct imaging of astronomical object and restores the potential resolution of a given telescope larger than 10 cm e.g., the diffraction limit of the 200-inch telescope is 25 milliarcseconds. Unlike the improvement in both sensitivity and spatial resolution expected from the Hubble Space Telescope through direct imaging from space, ground-based observations gain high angular resolution at the cost of sensitivity, because of the short life time of each interference pattern and the dependence on higher-order correlations. The resultant trade-off between sensitivity and resolution places high-resolution ground-based observations in a unique position in current astronomy as described later in this chapter. When one has to work at the very limit of a technique, astronomical goals and the technique are hardly separable.

In the development of high-spatial-resolution astronomy, two major issues have been 'imaging' and 'limiting magnitude.' Here imaging refers to the recovery of a two-dimensional source structure of the astronomical object, and limiting magnitude refers to the brightness limit of the astronomical object whose image is recoverable with a given signal-to-noise ratio. In this thesis, *Diffraction-Limited Imaging on the 200-inch Telescope*, these two issues are studied in two stages. The first stage is a demonstration of the potential of a high-resolution 'imaging' technique with ground-based optical telescopes, the non-redundant masking for making diffraction-limited images of astronomical objects with high dynamic range (Chapter II). The second stage is an estimate of

limitations of two high-resolution imaging techniques with ground-based optical telescopes, optical aperture synthesis including non-redundant masking (Chapter III) and the bispectral analysis of speckle interferometry (Chapter IV).

In this chapter, an introduction to current high-resolution astronomy is given in §2. In §3, one of the major goals of the high-resolution imaging effort at Caltech, high-resolution imaging of active galactic nuclei and quasars, is discussed in detail. High-resolution imaging techniques and some recent developments are introduced in §4. Future prospects of high-resolution imaging from the ground are discussed in §5.

2. Current High-Angular-Resolution Astronomy

In this section, high-resolution studies of astronomical objects at optical and infrared wavelengths are briefly reviewed.

Stellar Astronomy

High-resolution interferometric observation of astronomical objects was initiated as stellar interferometry by Michelson and Pease (1921). Among other fields of stellar interferometry, studies of binary stars have been systematically and extensively made. Binary stars are the fundamental source for the determination of stellar masses. Interferometric determinations of the magnitude differences and the orbits of double-lined spectroscopic binary star systems give the best estimates of masses. Systematic studies of binary stars are currently going on and the empirical mass-luminosity relation is still being improved (McAlister 1985, references therein). The measurements of spectroscopic binary stars also contribute to cosmology through independent distance determinations. Binary Cepheids and spectroscopic binaries in the Hyades cluster are

good targets of current high-resolution techniques. Some results on Hyades binaries have already been published (McAlister 1985).

Stellar diameters are the fundamental parameters in the study of stellar atmospheres. Until recently, however, the only systematic study of stellar diameters was made by the stellar intensity interferometer (Hanbury Brown 1974, references therein). Angular diameters of 32 stars brighter than B = 2.5 mag were obtained; the average measured angular diameter was 1.7 milliarcseconds and the spectral type of the stars ranged from O5f to F8. The measurement of about 30 more stars was made by speckle interferometry, amplitude interferometry and CERGA long-baseline interferometry in France with lower resolutions than the intensity interferometer. The status of interferometric measurements before 1984 has been summarized by Davis (1984). Very recently, the Mount Wilson Mark III interferometer (Shao et al. 1988) produced the preliminary results of the angular diameter measurement of 24 stars (Hutter et al. 1988) with a 12-m baseline. This fringe-tracking interferometer is promising in this field, taking into account the use of a 32-m baseline in the near future. Necessary resolution and sensitivity to cover various spectral types and luminosity classes are discussed by Davis (1978).

The massive end of the stellar mass distribution is ambiguous because of the high multiplicity found in very massive stars by high-resolution imaging. R136a, the central object of 30 Doradus with an estimated mass of 3000 $M_{\odot}$, has been resolved into a multiple stellar system by speckle imaging (Weigelt and Baier 1985). A systematic study of the duplicity and multiplicity of massive objects (e.g., Wolf-Rayet stars) should be made by high-resolution observations.

Until now, infrared study of dust-reflection nebulosities around both young stellar objects and evolved mass-loss stars has been made by one-dimensional speckle interferometry with a resolution of about 0.3 arcseconds (e.g., HL Tau; Beckwith et al. 1989). Major progress in this field is expected with the use of two-dimensional detectors with low read-out noise.

Solar System

Since Michelson measured the angular diameters of the Galilean satellites of Jupiter (1891), solar system objects have been the targets of high-resolution observations. Since the photographic discovery of a satellite of Pluto, Charon, by Christy and Harrington (1978), the orbit of the Pluto-Charon system has been traced by speckle interferometry (e.g., Beletic et al. 1989).

The solar system provides the brightest non-spherical solid bodies in the observable Universe, asteroids, for high-resolution 'imaging.' Since speckle interferometry was implemented, some observations of asteroids (e.g., Baier and Weigelt 1983) have been made. However, there has not been any systematic study of the three-dimensional geometry of an asteroid or a polarimetric study. Because of the transient nature of the objects, it is desirable to have a dedicated telescope which provides continuous observations.

Nuclei of Galaxies

Active galactic nuclei and quasars have often been referred to as major targets of high-resolution imaging. For example, the high-resolution imaging project at Caltech was initiated by the astronomers whose research interests are in these objects (e.g., Readhead et al. 1983; Neugebauer et al. 1987). However, at optical/IR wavelengths, there have not been many successful observations which extracted meaningful astrophysical parameters with a sub-arcsecond resolution from active galactic nuclei. A critical reexamination of the necessary dynamic range and resolution of images and observing strategy must be made, taking the future progress of the field made by the Hubble Space Telescope into account. Since a study of active galactic nuclei is also a major interest of the author, further discussion is given in the next section.

So far not much attention has been paid to normal galaxies. However, not only 'active' nuclei, but also nuclei of normal galaxies are interesting. By the study of stellar velocity dispersion around the nuclei of nearby galaxies, dynamically distinct cores have been found (e.g., Kormendy

1988). The stellar density profile of the core in M31 is likely to be obtained by high-resolution imaging.

By direct CCD imaging under good seeing conditions, stellar cusps were found in the nuclei of M87 and NGC6251 (Young et al. 1978, 1979). Resolving stellar cusps is a challenging project for speckle interferometry applied to faint objects.

3. Active Galactic Nuclei and Quasars

Since one of the original scientific motivations of the project *Diffraction-Limited Imaging on the 200-inch Telescope* was to study active galactic nuclei (AGN) and quasars, a discussion of future prospects of the field is given in this section.

Central power-law-continuum source

The physical dimension of the emitting region for black hole models will be order of 10 r_g, where r_g is the gravitational radius. For a black hole of 10^9 $M_\odot$, 10 r_g = 3×10^{10} km or 250 AU. At the distance of the Virgo cluster, the corresponding angular size is 10 μarcseconds. Mm-VLBI may achieve this resolution in the near feature. However, it is not likely that a central source can be resolved by optical/IR interferometry early in the next century.

Broad-Line Region (BLR)

At optical wavelengths, the environment of the AGN whose angular scale is smaller than a few arcseconds has been studied primarily by spectroscopy. The kinematics and the physical condition of the interstellar medium near the central power-law source are thought to be determined by its gravitational potential and radiation. The BLR is characterized by a high density (electron density $> 10^9$ cm^{-3}) and a high ionization state. There has not been any consensus on the

origin of high-velocity wings of emission lines (> 5000 km/s) in this region, though they may be partly explained by Keplerian motion around a massive object. By a simple scaling argument concerning the luminosity and the region size based on the photo-ionization model (Davidson and Netzer 1979, references therein), it is predicted that the BLRs of the brightest AGNs and QSOs (e.g., NGC4151, 3C273) at low redshifts have angular sizes of about 0.5 milliarcseconds. For instance, the physical dimension at the distance of NGC4151 is 0.05 pc. However the outer region may merge into the narrow-line region described below at a scale as large as 1 pc. The necessary angular resolution corresponds to a 60-m baseline interferometer in space at Lyα and a 300-m baseline at Hα. In order to extract the dynamics of the region it will be desirable to resolve spectrally into at least a few pixels with an appropriate fractional bandwidth of around 0.01. A breakthrough in ground-based long-baseline interferometry (§5) or a space interferometer in the next century will enable us to study the BLRs.

Narrow-Line Region (NLR)

The NLR lies between the outer envelope of the BLR and the host galaxy. The interstellar medium with a low density ($< 10^4$ cm^{-3}) and a small velocity dispersion (< 1000 km/s) compared to that in the BLR is ionized or heated by X-ray and UV radiation from the central object.

The scale of interest ranges from 0.1 pc to 1 kpc. In nearby AGNs, the NLRs are often spatially resolved by direct line imaging with a resolution of a few arcseconds. Recent studies show the evidence for spatial stratification of physical properties (velocity, density, ionization state) within the NLR (Wilson and Heckman 1985, references therein). Therefore any angular scale smaller than 1 arcsecond is of interest in nearby AGNs.

Using strong [OIII] doublets at 4959 Å and 5007 Å, Ebstein et al. (1989) have mapped the NLRs of NGC1068 and NGC4151 by speckle imaging with a resolution of 0.3 arcseconds. The spatial distribution of the [OIII] emission was found to be similar to that of the radio continuum

at 2 cm. The same tendency had been found previously at larger scales (Wilson and Heckman 1985).

Since the expected spatial structure of the emitting region is complicated, the true spatial resolution of a line image will depend on its dynamic range. High-dynamic-range imaging is a challenge for speckle techniques. The situation will be drastically improved by the Hubble Space Telescope which will have resolutions up to 20 milliarcseconds at near UV and 50 milliarcseconds at visual wavelengths with very high dynamic range. Beyond a resolution of tens of milliarcseconds, NLRs may be targets of the first-generation optical-aperture-synthesis array planned to be built in the last decade of the 20th century. However, it is not likely that the image quality of an aperture synthesis map of the NLR will be comparable with that of a direct image taken by the Hubble Space Telescope. Structures seen with a resolution of 50 milliarcseconds but with a high contrast may be lost at higher resolution because of the lower dynamic range of an image made by ground-based interferometry. Unless a breakthrough from the ground takes place (some candidates are discussed in §5), high-dynamic-range mapping from the ground may turn out to be impossible. The recognition of the absolute limitation of the ground-based techniques will become a force for steering the efforts of the high-resolution community to space interferometry.

Infrared continuum

So far the classification of distinct regions in AGNs has been made spectroscopically at optical wavelengths and the size estimates have been made mainly by the study of variability. The above discussions about the central continuum source, BLR and NLR, are simply based on the previously known facts. It is likely that an improvement in spatial resolution will be accompanied by a new discovery. At optical wavelengths, a true discovery will be made by the Hubble Space Telescope. Infrared continuum mapping of AGNs will be able to reveal at least three features in the environment of the nuclei: the stellar distribution, star-forming activity and

thermal re-radiation from hot dust heated by the central continuum source.

The stellar distribution in AGNs will be best studied in the near-infrared wavelengths because of extinction in the AGN. The diffraction limit of a 10-m telescope at 2 μm is comparable to the resolution of the Hubble Space Telescope at visual wavelengths. However, in the case of the heavily obscured nucleus found in NGC1068, the near-infrared continuum may be affected by scattered radiation of the obscured central source by dust. High-resolution imaging polarimetry will help determine the nature of near-infrared light in dusty objects. Circum-nuclear star-forming regions will be located by infrared continuum mapping beyond 3 μm and maybe best at 10 μm. Any angular scale smaller than a few arcseconds (kpc at the distance of Virgo) is of interest. These regions overlap with the optical NLR.

Dusty Seyfert galaxies, especially ones selected by IRAS observations, have most of their luminosity in the mid- to far-infrared wavelengths (Soifer et al. 1987, references therein). They show a wide variety of spectral energy distributions. For thermal re-radiation from dust heated by the central power-law source, it is possible to estimate the size of the emitting region at a given wavelength and observed flux density, by assuming the equilibrium temperature which peaks at that wavelength. In the case of the brightest Type-2 Seyfert, NGC1068, the size of the 10-μm emitting region will be between 10 milliarcseconds and 100 milliarcseconds depending on the optical depth. The angular scale of interest is considerably smaller at shorter wavelengths.

Infrared Line Mapping

A variety of hydrogen recombination lines in the Paschen and Brackett series, helium recombination lines, forbidden [FeII] lines and molecular hydrogen lines fall into near-infrared windows depending on the redshift of the galaxy.

Among others, molecular hydrogen lines having been found to be generally strong in AGN galaxies are of special interest (e.g., Kawara et al. 1987). From the study of runaway heating of

the interstellar medium (ISM) of the host galaxy by the X-ray from the central source, Begelman (1985) argues that even after the elimination of the cool component of the diffuse ISM ($T < 10^4$ K) within a scale of kpc from the nucleus, giant molecular clouds may survive, because of their small covering factors and large masses. Then the spatial distribution of the molecular hydrogen lines will be a key to understanding the circum-nuclear star formation and possibly the feeding mechanism of the central object. Any angular scale smaller than a few arcseconds has significant information.

Forbidden [FeII] at 1.644 μm has also been found to be strong in Seyfert galaxies (e.g., Moorwood and Olivia 1988). In dusty IRAS galaxies where optical narrow lines have significant extinction, [FeII] may be useful as a tracer of the NLR in addition to hydrogen recombination lines, provided that this line is excited by the radiation from the central power-law source (Graham et al. 1989). Angular scales of interest are the same for the optical NLR. Infrared spectroscopy of AGNs should precede infrared line imaging to provide a basic guide line.

4. High Resolution Imaging Techniques from the Ground

In this section, high-resolution imaging from the ground is introduced and the recent developments are described of two imaging techniques relevant to this thesis.

Two-Dimensional Photon-Counting Detection of an Astronomical Object

Turbulent mixing of the air causes local fluctuations of temperature and thus of density and refractive index. Multiple small-angle scattering by the local fluctuations of the refractive index turns an incoming plane wave from a star into a partially coherent wavefront. The partial coherence originates from the continuity of the wavefront and the finite phase-fluctuation power at small scales characterized by a coherence length which is about 10 cm at visual wavelengths.

If the partially coherent light is collected by an aperture larger than the coherence length (a large ground-based telescope), a random interference pattern or speckle pattern will be seen on the image plane. Each pair of coherent portions in the aperture plane forms an interferometer. If the baseline of the hypothetical interferometer is shorter than the diameter of the aperture, there will be multiple pairs which form the identical baseline. Then that baseline is redundant. Fringes made by the redundant baseline interfere destructively on the image plane and the modulation attenuates. The degree of attenuation depends on the redundancy of the baseline.

If a mask with a pair of holes whose size is much smaller than the coherence length (an ideal Michelson interferometer) is set over the aperture plane of the telescope, a fringe pattern with an amplitude of unity and a random phase will be seen on the image plane. Even if the number of holes is increased to three, a set of three superposed fringes with amplitudes of unity will be seen on the image plane, as long as the configuration of the holes does not form any redundant baseline. Although the individual fringes wander around according to their random phases, the 'closure phase', the sum of the three fringe phases, remains constant, setting a constraint on the triangular pattern.

The translation by the wind and the pattern change of the turbulent layer cause decorrelation of the speckle pattern or the fringe pattern in time. In order to preserve the modulation corresponding to a high spatial frequency, snap shots with an exposure time of 10 milliseconds must be taken. For an extended source, the modulation in a speckle or that in a fringe pattern is attenuated by incoherent addition (or convolution) of displaced point-source speckle patterns or that of displaced point-source fringe patterns in a static manner.

At high light levels, the signal-to-noise ratio of the modulation per speckle pattern is finite and the static source information has to be extracted from an ensemble of speckle patterns. On the other hand, a fringe pattern made by the ideal Michelson interferometer has an infinite signal-

to-noise ratio per snap shot.

At low light levels, the particle nature of the visual light sets further limits to the signal-to-noise ratio through photon noise. A wavefront disturbed by the atmosphere becomes a time-varying wave function of a single photon and a speckle pattern or a fringe pattern becomes a probability distribution for the detection of each photon. Instead of the ideal Michelson interferometer, masks with apertures whose sizes are about the lateral coherence length are used, in practice, to collect more photons at low light levels. We term this technique, 'non-redundant masking' (NRM), since the baseline configuration does not have redundancy. Only a limited number of discrete photon events are seen within a short-exposure frame. In practice, a photon-counting detector is used in our experiments and therefore the observables are the event location and the arrival time of each photon.

Image Reconstruction by Spatial Spectrum Analysis

An image of an extended object is in principle hidden in its instantaneous speckle pattern as the convolution function with the hypothetical simultaneous point-source speckle pattern. In the case of the ideal Michelson interferometer, the point-source fringe pattern is in principle known *a priori*, therefore the modulation in the object fringe pattern should immediately tell about the source structure. However, in practice, the quality of an instantaneous speckle pattern or fringe pattern obtained by NRM is severely limited both by the atmospheric disturbance and photon noise. Thus an image must be obtained from ensemble averaged observables. The atmospheric disturbance makes the absolute position of the speckle pattern and the absolute phase of the fringe pattern poor observables. Natural observables are correlation functions (autocorrelation, triple correlation ...) of the image or their counterparts in the Fourier domain (power spectrum, bispectrum ...), since these quantities extract structure of the patterns. The power spectrum gives the amplitude of the modulation leaving 180° ambiguity compared to a 'true image.' The

bispectrum gives phase information as the closure phase which is related to the asymmetry in the structure. When the amplitude and the closure phase are combined, a true image is obtained apart from the absolute position. The amplitude and the phase of the Fourier transform of the image are derived from the power spectrum and the bispectrum by supplying additional absolute position information. In practice, direct correlations of photon events are often taken at low light levels. Correlation functions whose orders are higher than the third are practically of no use, since they are much noisier than the autocorrelation and the triple correlation.

Speckle Imaging

The recognition of the use of full apertures of ground-based large telescopes by Labeyrie (1972) through speckle interferometry opened up a new era in high-resolution observations. A clue to speckle imaging using the phase difference was first proposed by Knox and Thompson (1974) and a practical implementation was made recently (e.g., Nisenson et al. 1985). A more powerful imaging technique, the bispectral analysis was developed by the Erlangen group (e.g., Lohmann et al. 1983, Wirnitzer 1985). A successful recovery of an image of a 11-mag multiple stellar system with a resolution of 0.1 arcsecond by Hoffman and Weigelt (1986) gathered the attention of radio astronomers who had been producing diffraction-limited maps under unstable conditions of phase through phase-closure imaging (Pearson and Readhead 1984, references therein). As was pointed out by Cornwell (1987), the bispectral analysis is the fully-filled-aperture version of phase-closure imaging. A comparison between the bispectral analysis and NRM in terms of the redundancy of the aperture configuration was discussed by Readhead et al. (1988).

In this thesis, the potential and the limitations of the bispectral analysis are studied by Monte Carlo simulations to give an estimate of the limiting magnitude and the resolution in Chapter IV.

Optical Aperture Synthesis Imaging

Diffraction-limited imaging of an astronomical object using an interferometric array under

unstable phase conditions was first achieved at radio wavelengths as phase-closure imaging of VLBI (Readhead and Wilkinson 1978). Although the application of phase-closure imaging to optical wavelengths had been considered to be straightforward (Readhead et al. 1980), measurements of the closure phase to test the feasibility of optical phase-closure imaging were made only recently by a group at Cambridge (Baldwin et al. 1986) and by one at Caltech (Readhead et al. 1988) through the NRM technique with ground-based optical telescopes. The successful measurement of the closure phase by the Cambridge group allowed the first optical aperture synthesis image of a binary star with a separation of a half arcsecond (Haniff et al. 1987), that by the Caltech group was demonstrated by two images of binary stars with separations of 70 and 230 milliarcseconds presented in Chapter II. The application of the radio imaging technique to optical aperture synthesis turned out not to be as straightforward as was expected. Fundamental differences that arise from the difference in interference-detection schemes at radio and optical wavelengths and associated problems are discussed in Chapter II.

The NRM technique can be regarded as a first step towards the next-generation optical aperture synthesis array under the constraint of passive optics. The effects of the atmospheric disturbance on the NRM and masking with apertures larger than the lateral coherence length are studied by Monte Carlo simulations and the sensitivity is obtained in Chapter III.

5. Future Prospects

Potential and Limitations of Ground-Based Imaging Techniques with Passive Optics

As estimated in Chapter III and Chapter IV respectively, the limiting visual magnitude of NRM is around 11 mag and that of speckle imaging is somewhere between 13 and 15 mag. The limiting magnitude of long-baseline interferometry is not better than those of single-telescope

techniques, since the atmospheric disturbance and the bandpass set more severe constraints on long-baseline interferometry. Here the limiting magnitude of 11 mag is taken as the best-case design. A magnitude-resolution diagram at optical wavelengths is shown in Fig. 1. In this diagram, the limiting magnitude of the power spectrum analysis (or Labeyrie's speckle interferometry) and the coverage of the Hubble Space Telescope are included.

Major progress is expected in the measurement of stellar diameters by long-baseline interferometry and in imaging of asteroids by single telescopes. The study of active galactic nuclei becomes difficult beyond the resolution of the Hubble Space Telescope.

Infrared high-resolution imaging is limited either by detector background or sky background depending on the type of detector. For an integration-type two-dimensional detector with read-out noise of 100 electrons per pixel, the limiting magnitude is around 7 mag, while for a low background detector such as a SSPM discussed in the following, the limiting magnitude will reach the sky-background limit of about 11 mag. The limiting magnitudes were obtained simply by equating the number of photon counts with that of background counts.

A magnitude-resolution diagram at 2 μm is also shown in Fig. 2. Reflection nebulosities around young stellar objects and T-Tauri stars are observable with a limiting magnitude of 6 mag or fainter with a resolution of one arcsecond. Nearby Seyfert nuclei are observable with a limiting magnitude of 8 mag or fainter. The brightnesses of a typical bipolar-outflow source, L1551-IRS5 and the brightest quasar, 3C273 are both around 9 mag. The study of narrow-line regions of active galactic nuclei will crucially depend on the availability of low background detectors.

Possible Breakthroughs

The ultimate solution of the problem of the atmospheric disturbance will be to set an observing system in space (space interferometer, telescope on the moon, etc.). However this is far beyond the scope of this thesis and probably a project for the mid 21st century. Here

possible breakthroughs whose success may drastically improve the sensitivity of the ground-based techniques in the near future are discussed.

Adaptive Optics + Artificial Guide Star

The fundamental limitation of imaging techniques using passive optics such as speckle interferometry or NRM is set by the finite coherent integration time and the necessity for higher-order correlations. The idea of adaptive optics such as wavefront compensation or fringe tracking is to make the integration time unlimited and obtain a direct image or true phase. Since the behavior of the wavefront must be known in real time, the adaptive optics requires many photons per coherent cell per coherence time. Therefore the limiting magnitude of 'pure' adaptive optics which activates a servomechanism solely by an astronomical object itself is worse than the passive methods and is limited at around 9 mag at visual wavelengths (Roddier 1988).

Although astronomers do not have any control over the atmospheric disturbance itself, they can possibly have a control over a light source in the sky which may illuminate the wavefront and activate the adaptive optics. An idea of creating an artificial guide star by laser pulses was first proposed by Foy and Labeyrie (1985). The first experiment using resonant scattering by the sodium layer in the mesopause (80-110km) was carried out by Thompson and Gardner (1987). Although the artificial star spread over one arcminute, the intensity of backscattered light was found to be nearly sufficient.

Because of the proximity to the ground and the finiteness of the scattering layer, a laser guide star is not likely to be useful for long-baseline interferometry. As an intermediate stage between a 'pure' ground-based interferometer which has the severe limitation in sensitivity and a space interferometer which is at least many decades ahead, a ground-based interferometer guided by an artificial star in space may be a realistic solution to the sensitivity problem. Since the satellite which carries a light source has to keep intervening the astronomical object and the ground-based

interferometer on the rotating earth, there may be a technical challenge in navigation. A laser is again the suitable light source and is especially ideal in the infrared where the elimination of the thermal background emission by a narrow bandpass is crucial.

Solid-State Photomultiplier

Although the detector technology both for long integrations (CCD) and for photon counting (photomultiplier) at optical wavelengths seems to have been nearly optimized, the detector technology in the infrared is still in a stage of rapid development.

The ideal detector for high-resolution imaging is characterized by a high quantum efficiency and low background noise. Even in the infrared, by reducing the thermal sky background by spatial filtering and/or spectral filtering, the advantage of a low background detector becomes obvious. The most promising and possibly the ultimate infrared detector in a low background situation is a solid-state photomultiplier (SSPM: Petroff et al. 1988). The SSPM is a solid-state photon-counting detector which utilizes electron multiplication by the impurity-band avalanche caused by impact ionization of one carrier. A quantum efficiency of 30% is observed at 20 μm and the linearity is good to 10^{10} counts per second. By choosing the dopant, the spectral response can be optimized to infrared windows of interest (Bharat et al. 1987). Although the first device was made in the form of a small linear array, there is no essential difficulty in making a two-dimensional detector. The limiting magnitude of continuum interferometry at 2 μm using a two-dimensional SSPM should be 11 mag per square arcsecond, the sky background limit. The SSPM will truly be ideal for line imaging.

6. Layout of Thesis

This thesis is composed of three separate papers.

Chapter II, *Diffraction-Limited Imaging II: Optical Aperture Synthesis Imaging of Two Binary Stars* (Nakajima, T., Kulkarni, S.R., Gorham, P.W., Ghez, A.M., Neugebauer, G., Oke, J.B., Prince, T.A., and Readhead, A.C.S. 1989. Astron. J. **97**, 1510), describes a demonstration of the potential of the NRM technique for making diffraction-limited images with high dynamic range. It also discusses fundamental differences between optical aperture synthesis mapping and its radio counterpart which arise from the difference in the interference-detection scheme.

Chapter III, *Atmospheric Disturbance on Optical Aperture Synthesis*, (to be submitted to J. Opt. Soc. Am. **A**), describes Monte Carlo simulations of the NRM technique and masking with apertures with larger than the coherence length. Fringe amplitudes are obtained as functions of aperture size, bandwidth and baseline length. By introducing a cut-off spatial frequency in the Kolmogorov spectrum, the effects of small-scale and large-scale disturbances are separately simulated and the relation of the results to the performance of a passive interferometer and active interferometer is discussed. Based on the simulated fringe amplitudes, limiting magnitudes of the NRM technique and masking with 50-cm apertures are derived.

Chapter IV, *Signal-to-Noise Ratio of the Bispectral Analysis of Speckle Interferometry*, (Nakajima, T. 1988. J. Opt. Soc. Am. **A5**, 1477), describes a study of the performance of a speckle imaging technique, the bispectral analysis, by Monte Carlo simulations and by a modeling of photo-detection process. The bispectral modulation transfer function is simulated and its signal-to-noise ratio at high light levels is obtained. A general expression of the signal-to-noise ratio of the bispectral analysis of photon-noise-affected images is derived. In the mid-spatial-frequency range, the signal-to-noise ratio is obtained as a function of the telescope transfer function, the number of speckles and the light level.

Chapter IV was published first, and then Chapter II and Chapter III were written in chronological order. Minor modifications in technical terms, notations and references were made to

original papers for the self-consistency as a thesis.

References

Baier, G. and Weigelt, G. 1983. Astron. Astrophys., **121**, 137.

Baldwin, J.E., Haniff, C.A., Mackay, C.D. and Warner, P.J. 1986. Nature, **320**, 595.

Beckwith, S.V.W., Sargent, A.I., Koresko, C.D., and Weintraub, D.A. 1989. Astrophys. J., In publication.

Begelman, M.C. 1985. Astrophys. J., **297**, 492.

Beletic, J.W., Goody, R.M. and Tholen, D.J. 1989. Ikaros, In publication.

Bharat, R., Petroff, M.D., and Stapelbroek, M.G. 1987. In *Infrared Astronomy with Arrays*, Wynn-Williams, C.G. and Becklin, E.E., eds. Univ. of Hawaii, Institute for Astronomy, Honolulu, Hawaii, p97.

Christy, J.W. and Harrington, R.S, 1978. Astron. J., **83**, 1005.

Cornwell, T.J. 1987. Astro. Astrophys., **180**, 269.

Davidson, K. and Netzer, H. 1979. Rev. Mod. Phys., **51**, 715.

Davis, J. 1979. In *High Angular Resolution Stellar Interferometry,* IAU Colloq. No. 50., Davis, J. and Tango, W.J., eds., Chatterton Astron. Dep., Univ. of Sydney.

Davis, J. 1984. In *Calibration of Fundamental Stellar Quantities, IAU Symp. No. 111,* Philip, A.G.D., Hayes, D.S., eds. Dordrecht: Reidel.

Ebstein, S.E., Carleton, N.P. and Papaliolis, C. 1989. Astrophys, J., **336**, 103.

Foy, R. and Labeyrie, A. 1985. Astro. Astrophys., **152**, L29.

Graham, J.R., Wright, G.S. and Longmore. 1989. Submitted to Astrophys. J.

Hanbury Brown, R. 1974. *The Intensity Interferometer*, London: Taylor & Francis.

Haniff, C.A., Mackay, C.D., Titterington, D.J., Sivia, J.E., Baldwin, J.E. and Warner, P.J. 1987. Nature, **328**, 694.

Hoffmann, K.-H. and Weigelt, G. 1986. Astron. Astrophys., **167**, L15.

Hutter, D.J., Johnston, K.J., Mozurkewich, D., Simon, R.S., Colavita, M.M., Pan, X.P., Shao, M., Hines, B.E., Staelin, D.H., Hershey, J.L., Hughes, J.A., and Kaplan, G.H. 1988. In *Proceeding of NOAO-ESO conference on High-Resolution Imaging by Interferometry*, Merkle, F., ed. (European Southern Observatory, Garching, 1988), p 855.

Kawara, K., Nishida, M. and Gregory, B. 1987. Astrophys. J., **321**, L35.

Knox, K.T. and Thompson, B.J. 1974. Astrophys. J., **193**, L45.

Kormendy, J. 1988. Astrophys. J., **325**, 128.

Labeyrie, A. 1970. Astro. Astrophys., **6**, 85.

Lohmann, A.W., Weigelt, G. and Wirnitzer, B. 1983. Appl. Opt., **22**, 4028.

McAlister, H.A. 1985. Ann. Rev. Astro. Astrophys., **23**, 59.

Michelson, A.A. 1891. Nature, **45**, 160.

Michelson, A.A. and Pease, F.G. 1921. Astrophys. J., **53**, 249.

Moorwood, A.F.M. and Olivia, E. 1988. Astro. Astrophys., **203**, 278.

Neugebauer, G., Elias, J., Matthews, K., McGill, J., Scoville, N. and Soifer, B.T. 1987. Astron. J., **93**, 1057.

Nisenson, P., Stachnik, V., Karovska, M. and Noyes, R. 1985. Astrophys. J., **297**, L17.

Pearson T.J. and Readhead, A.C.S. 1984. Annu. Rev. Astron. Astrophys., **22**, 97.

Petroff, M.D., Stapelbroek, M.G. and Kleinhans, W.A., *Solid State Photomultiplier*, United States Patent Number 4,586,068, granted April 29, 1986.

Readhead, A.C.S. and Wilkinson, P.N. 1978. Astrophys. J., **223**, 25.

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

Readhead, A.C.S., Hough, D.H., Ewing, M.S., Walker, R.C., and Romney, J.D. 1983. Astrophys. J., **265**, 107.

Readhead, A.C.S., Nakajima, T., Pearson, T.J., Neugebauer, G., Oke, J.B. and Sargent, W.L.W. 1988. Astron. J., **95**, 1278.

Roddier, F. 1988. In *Proceeding of NOAO-ESO conference on High-Resolution Imaging by Interferometry*, Merkle, F., ed. (European Southern Observatory, Garching, 1988), p 565.

Shao, M., Colavita, M.M.,Hines, B.E., Staelin, D.H., Hutter, D.J., Johnston, K.J., Mozurkewich, D., Simon, R.S., Hershey, J.L., Hughes, J.A., and Kaplan, G.H. 1988. Astro. Astrophys., **193**, 357.

Soifer, B.T., Houck, J.R. and Neugebauer G. 1987. Ann. Rev. Astro. Astrophys., **25**, 187.

Thompson, L.A. and Gardner, C.S. 1987. Nature, **328**, 229.

Weigelt, G. and Baier, G. 1985. Astro. Astrophys., **150**, L18.

Wilson, A.S. and Heckman, T.M. 1985. In *Astrophysics of Active Galaxies and Quasi-Stellar Objects*, Miller, J.S., ed. University Science Books.

Young, P.J., Westphal, J.A., Sargent, W.L.W., Kristian, J., Wilson, C.J., and Landauer, F.P. 1978. Astrophys. J., **221**, 721.

Young, P.J., Sargent, W.L.W., Kristian, J., and Westphal, J.A. 1979. Astrophys. J., **234**, 76.

Figure Captions

Fig. 1 - Magnitude-resolution diagram at visual wavelengths. A right arrow (e.g., of the intensity interferometer) indicates the limiting magnitude. An upward arrow (e.g., of the Hubble Space Telescope (HST)) indicates the the diffraction limited resolution of the observing system. The right arrow of the HST indicates a potentially higher resolution. Right and downward arrows of astronomical objects indicate necessary limiting magnitudes and resolutions respectively. The box of Pluto-Charon system indicates the ranges of the brightness and separation of the object. A left arrow (e.g., of 3C273) indicates the brightness and an unknown angular size of the object.

Fig. 2 - Magnitude-resolution diagram at 2 μm. The limiting magnitude of IR speckle imaging using an imager with a read-out noise of 100 electrons per pixel, is estimated to be 7.5. The limiting magnitude of IR speckle imaging with 2-μm continuum using a two dimensional SSPM is estimated to be 11. This corresponds to the sky brightness at Palomar of 11 mag per square arcsecond. The limiting magnitude of IR line imaging with a spectral resolution of 1000 is estimated to be 12.5. The spatial resolution of 2-μm direct imaging reaches 0.5 arcseconds. The meanings of arrows are the same as in Fig. 1.

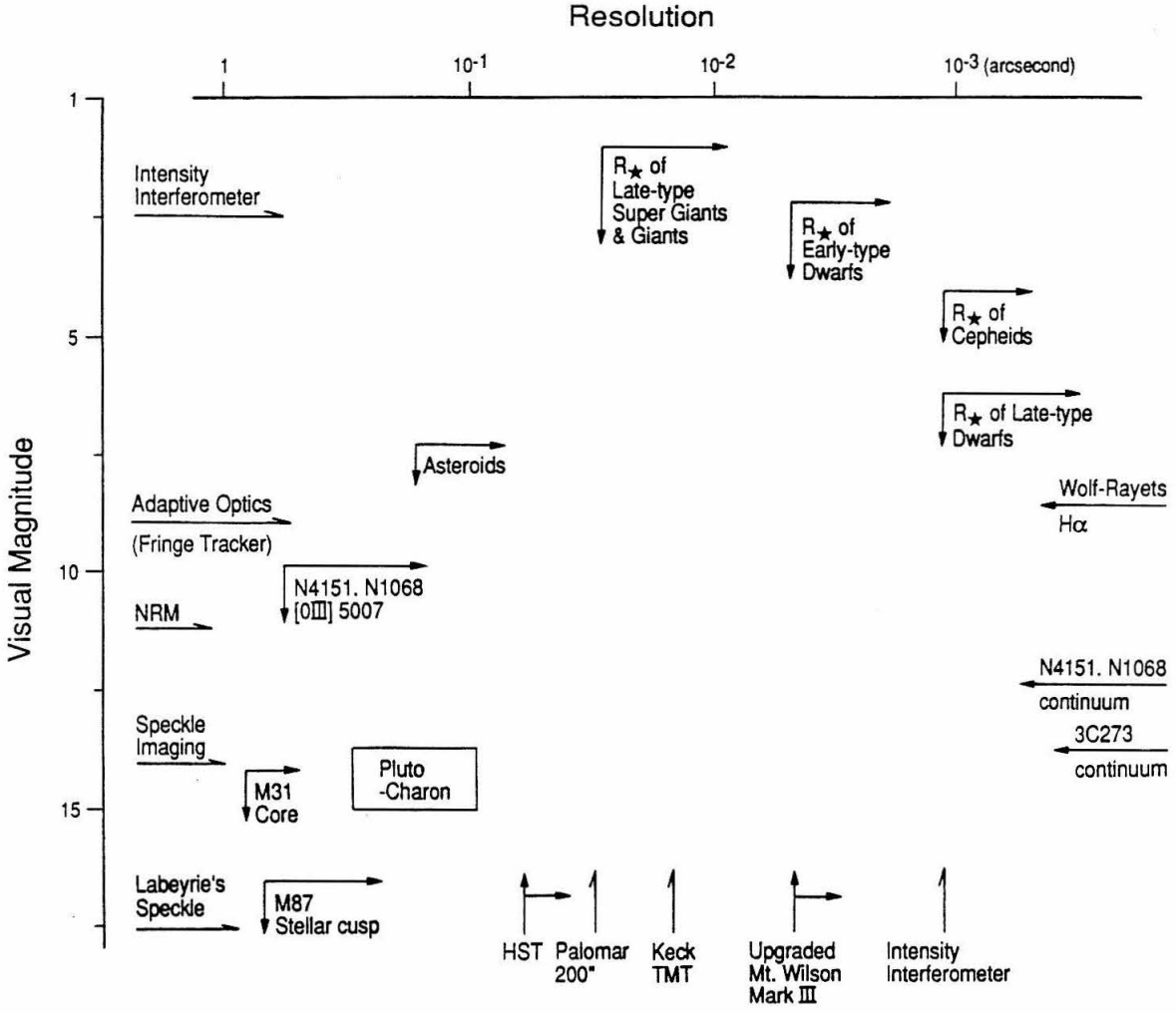
Resolution
1
10-1
10-2
10-3 (arcsecond)
Visual Magnitude
1
5
10
15
Intensity Interferometer
R★ of Late-type Super Giants & Giants
R★ of Early-type Dwarfs
R★ of Cepheids
R★ of Late-type Dwarfs
Asteroids
Wolf-Rayets
Hα
Adaptive Optics
(Fringe Tracker)
N4151. N1068
[OIII] 5007
NRM
N4151. N1068
continuum
3C273
continuum
Speckle Imaging
M31 Core
Pluto -Charon
Labeyrie's Speckle
M87 Stellar cusp
HST
Palomar 200"
Keck TMT
Upgraded Mt. Wilson Mark III
Intensity Interferometer

Figure 1

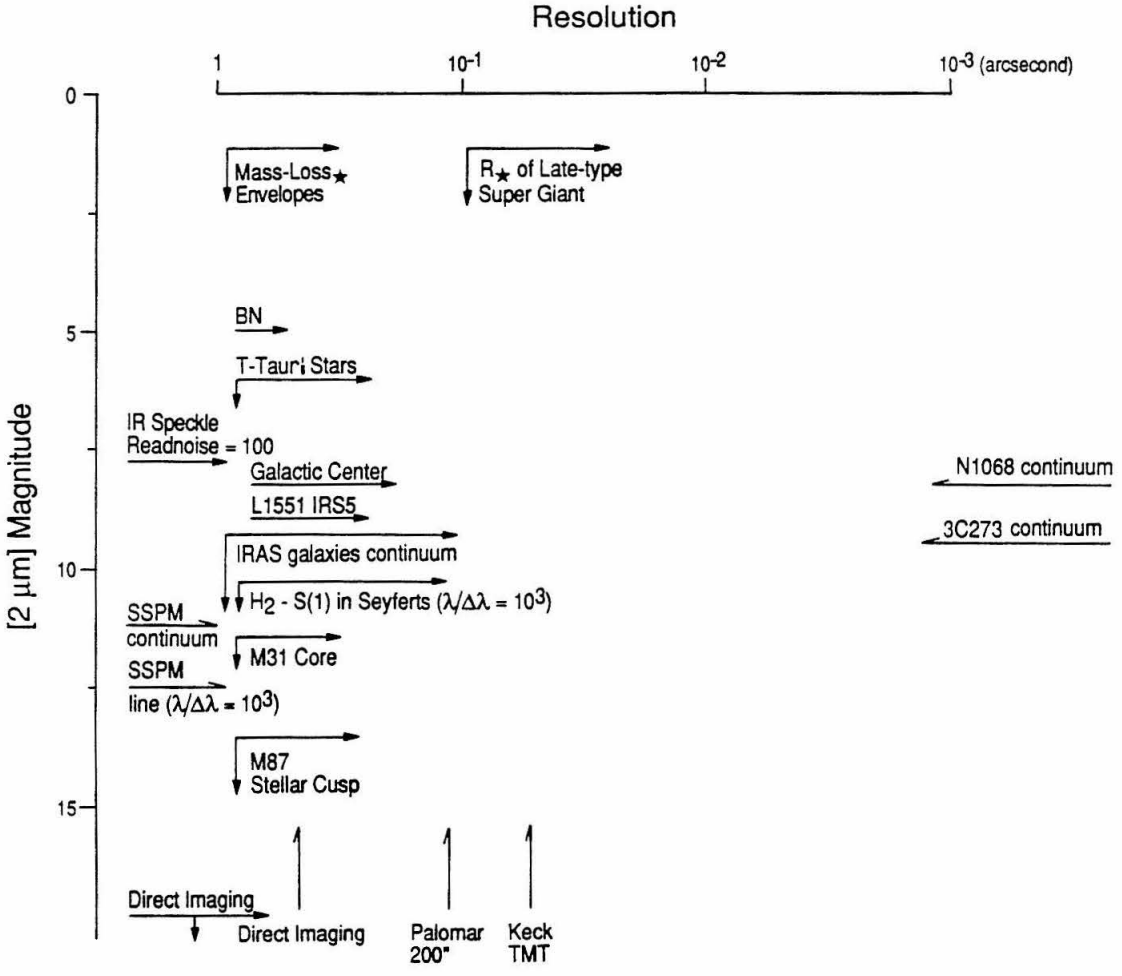
Resolution
1
10-1
10-2
10-3 (arcsecond)
0
5
10
15
[2 μm] Magnitude
Mass-Loss★ Envelopes
R★ of Late-type Super Giant
BN
T-Tauri Stars
IR Speckle Readnoise = 100
Galactic Center
L1551 IRS5
IRAS galaxies continuum
H2 - S(1) in Seyferts (λ/Δλ = 10^3)
SSPM continuum
M31 Core
SSPM line (λ/Δλ = 10^3)
M87 Stellar Cusp
N1068 continuum
3C273 continuum
Direct Imaging
Direct Imaging
Palomar 200"
Keck TMT

Figure 2

Chapter II.

Optical Aperture Synthesis Imaging of Two Binary Stars

1. Introduction

Three years ago we began a concerted program to obtain diffraction-limited images at optical wavelengths using the Hale 5-m telescope of the Palomar Observatory. The basic goal of this program is to understand quantitatively the strengths, limitations and sensitivities of different high-resolution imaging techniques. The background to this work and the demonstration of the feasibility of constructing diffraction-limited images on the 5-m telescope, using the techniques adapted from radio astronomy, have been described elsewhere (Readhead et al. 1988, hereafter Paper I). We have concentrated our efforts on understanding the trade-offs between the fully filled aperture (FFA) and the non-redundant masking (NRM) techniques. In this chapter we present the first diffraction-limited images obtained by optical aperture synthesis using non-redundant masks on the Hale 5-m telescope. In a subsequent paper we will present results obtained using the FFA (Gorham et al. 1989). In parallel with these observational efforts, we have embarked on a detailed analytical study of the problems associated with diffraction-limited imaging (Paper I; Chapter IV; Kulkarni 1989; Kulkarni and Nakajima 1989; Chapter III).

Other groups are engaged in similar studies aimed at exploiting the imaging potential of optical telescopes, and results have been published of a number of high-resolution observations using speckle interferometry (Labeyrie 1972), the Knox-Thompson algorithm (e.g., Nisenson et al. 1985), the speckle masking method (e.g., Hoffman and Weigelt 1986) and the NRM technique (Haniff et al. 1987).

Unlike speckle interferometry, which uses full *uv* coverage with attenuated complex amplitudes, the NRM technique uses limited *uv* coverage with high signal-to-noise ratio. A detailed study of the trade-offs between the speckle masking technique and the non-redundant masking technique shows that the NRM method produces higher dynamic range images at medium to high

light levels, while the speckle masking is better suited to pursuing the faint limit. (Kulkarni and Nakajima 1989). Finally, the NRM technique can be regarded as a first step towards an optical interferometric array and is important in establishing an optimal image reconstruction procedure and understanding the potential and limitations of the performance of an interferometric array under the constraint of passive optics.

For our first attempts, only bright binary stars were observed. The reconstructed images demonstrate intrinsically high resolution and potentially high dynamic range. The instrumentation and the observing procedure are described in §2 and §3. The data reduction and analysis are found in §4. Images of binary stars are presented in §5 along with the discussion of their dynamic range. In §6, the image reconstruction procedure and the estimation of the dynamic range are re-examined and distinctions between radio interferometry and optical interferometry are discussed in detail.

2. Instrumentation

A schematic drawing of the optics of the prime focus camera for the Hale 5-m telescope is shown in Fig. 1. An $f/3.3$ primary beam is collimated by a Nikon $f/2.8$ 85 mm camera lens. The lens forms an image of the primary mirror at a distance of about 85 mm (1 focal length) where a mask is placed on a stepper-motor-driven rotary stage controlled by a personal computer. One mm on the mask plane is equivalent to 19.8 cm on the primary mirror. Another camera lens identical to the first forms a second focus where the scale is $12''$ / mm. The image is expanded by a microscope objective with a power of 80. The final image size on the detector is $0.15''$ / mm. A 6300 Å interference filter with 30 Å bandpass was set between the microscope objective and the detector throughout the run.

The detector is a resistive anode position sensing photon counting detector (ITT #FM 4146M). It has a MA-2 photo cathode, essentially a red extended S-20, and five stages of microchannel plates (MCPs) arranged in a V-Z configuration. A net potential drop of about 5 kV is maintained from the cathode to the anode. Each primary photoelectron results in an avalanche of $10^7 - 10^8$ secondary electrons onto the resistive anode. Output currents from the four corners of the resistive anode are amplified and fed to an analog-to-digital position computer (Surface Science Lab, Mountain View, CA), which converts input currents into (x, y) coordinates of the avalanche. The position computer outputs 10 bits per axis as well as a strobe pulse; the position conversion takes approximately 10 μs and this is a source of one of the dead-times of the detector. Another source of dead time is the charge depletion resulting from the large gain of the tube. This has the effect that after an avalanche of secondary electrons, the pixel(s) corresponding to the location of the photoelectron are dead, for perhaps as long as several tens of milliseconds. Fortunately, for our case the fringes spread the light out and for the counting rates reported here ($< 1.5 \times 10^4$ s^{-1}) dead-time effects were negligible. A camera interface appended 12 bits of arrival time information to the photon coordinate data and passed the resulting "event" to a Digital Equipment Corporation (DEC) μVax workstation via a DEC DRV-11 parallel interface. At appropriate times, the interface unit also sends the absolute time (32 bits) to the computer, enabling a reconstruction of the absolute arrival time of each photon to better than 1 μs. The data were first recorded onto a hard disk of the μVax and later transferred to a magnetic tape for archival storage and processing.

In order to decrease the dark current, the detector was cooled to -20° C with liquid nitrogen. At this temperature the dark count rate was below 20 counts s^{-1} over the entire area of the detector. By comparing the actual count rate of stars of known magnitudes, the net throughput of the observing system was inferred to be 0.7%. This is determined primarily by the net detector efficiency of about 2% and the filter transmittance of 40%. The nominal spatial resolution of the

detector is 60 μm FWHM and the size of active area is 25 mm in diameter. The corresponding angular resolution and field of view are 9 milliarcseconds (mas) FWHM and 3.75″ in diameter respectively. The spatial linearity was tested by measuring an artificial fringe pattern in the laboratory and the non-linearity was found to be negligible.

3. Observing procedure

Observations were made over the nights of 3 and 4 April, 1988. Objects near the zenith were selected to minimize the path length in the atmospheric layer. The seeing condition varied between 1″ and 2.5″ during the run. Observations of the binary stars whose images were successfully reconstructed were made under 1.5″ visual seeing or better. A nearby point source calibrator was observed for each binary star with the same *uv* coverage. Observations of a program source and the corresponding calibrator were separated in time by less than 30 minutes.

Two kinds of masks were used depending on the necessary resolution and *uv* coverage. One was a five hole mask which had relatively short baselines and a uniform *uv* coverage at each position angle. The maximum baseline length of this mask was 220 cm. The other was a six hole mask which had three widely separated pairs of holes. Each pair formed a short baseline of about 40 cm. The pairs were separated by about 280 cm and the maximum baseline length was 310 cm. The equivalent size of each hole in both masks was 15 cm in diameter on the primary mirror. The observations were made at four or five different rotational angles. The *uv* coverage of these two masks are shown in Figs. 2 and 5 respectively. For each rotational angle, either $\sim 6 \times 10^5$ or $\sim 12 \times 10^5$ events were recorded.

4. Data reduction and analysis

The data reduction and analysis were carried out on a CONVEX C-1 minisupercomputer and a VAX/780 and consisted of several phases:

i. Determination of spatial frequencies.

ii. Estimation of the fringe amplitude and closure phase.

iii. Calibration of the fringe amplitude.

iv. Model fitting and hybrid mapping.

(i). *Determination of Spatial Frequencies.*

Initially 10 ms was adopted as a default integration time, τ, and power spectra of the "frames" were obtained; here a frame is an image obtained over integration time equal to τ. Given the small number of photons per frame, it was computationally advantageous to calculate and integrate the frame auto-correlation functions (ACF) followed by the Fourier transform of the integrated ACF to yield the power spectrum.

The purpose of calculating the power spectrum was to determine the spatial frequencies provided by the mask. In principle, the spatial frequencies could have been measured by measuring accurately the coordinates of the holes in the mask if the alignment of the optics had been perfect. Instead, however, we inferred the spatial frequencies from the integrated power spectrum of observed data.

Approximate spatial frequencies of fringes were obtained by centroiding power spectrum components which had finite extent in the spatial frequency domain. In a mask with n_h holes only $n_h - 1$ hole coordinates must be determined since the origin of the mask plane is arbitrary. Least squares fitting was therefore made to determine $n_h - 1$ unknown hole coordinates from $n_h(n_h - 1)/2$

power peaks. The n_h hole coordinates were pair-wise differenced to yield $n_h(n_h - 1)/2$ baseline vectors. This procedure ensured that all the triangles closed (see below).

From the knowledge of the focal lengths and magnification of the optics, the baseline vectors were converted from pixel units to spatial frequency units. The absolute errors in the measurement of the baseline vectors in the detector plane are determined by the signal-to-noise ratio of power spectrum, while the systematic error in the conversion of the baseline vectors to spatial frequencies depends on the precision of alignment of the optics. It was found that the systematic error of 3% in scaling dominated the errors in the spacial frequencies. The orientation of the detector on the sky was measured by a low-magnification image of the center of a globular cluster. The position angle towards north was estimated to be precise to 1°.

(ii). *Estimation of the Fringe Amplitude and Closure Phase.*

Once the spatial frequencies were determined, it was computationally advantageous to evaluate the discrete Fourier transform (DFT) of the raw data at those frequencies and to estimate the bispectrum or the triple product at only selected spatial frequency triplets.

The DFT was evaluated at the spatial frequencies derived earlier from the power spectrum data in each frame. The DFT component at a spatial frequency $\mathbf{u}$ of the j^{th} frame, $\tilde{D}_j(\mathbf{u})$ is given by,

$$\tilde{D}_j(\mathbf{u}) = \sum_{k=1}^{N_j} \exp(i\mathbf{u} \cdot \mathbf{x}_k), \tag{1}$$

where N_j is the number of photons in the j^{th} frame and $\mathbf{x}_k$ is the location of the k^{th} event. An unbiased estimator of the classical power spectrum for the j^{th} frame, is given by,

$$\tilde{Q}_j^{(2)}(\mathbf{u}) = |\tilde{D}_j(\mathbf{u})|^2 - N_j, \tag{2}$$

(Goodman and Belsher 1976, 1977; Dainty and Greenaway 1979). An unbiased estimator of the

classical bispectrum for the j^{th} frame, $\tilde{Q}_j^{(3)}(\mathbf{u_1}, \mathbf{u_2})$ is given by

$$\tilde{Q}_j^{(3)}(\mathbf{u_1}, \mathbf{u_2}) = \tilde{D}_j(\mathbf{u_1})\tilde{D}_j(\mathbf{u_2})\tilde{D}_j(-\mathbf{u_1} - \mathbf{u_2}) - \{|\tilde{D}_j(\mathbf{u_1})|^2 + |\tilde{D}_j(\mathbf{u_2})|^2 + |\tilde{D}_j(-\mathbf{u_1} - \mathbf{u_2})|^2 - 2\overline{N}_j\}, \tag{3}$$

(Wirnitzer 1985, Chapter IV).

The determination of spatial frequencies by centroiding power spectrum components is not necessarily free from error. This can result in a non-closing triplet of spatial frequencies which introduces a complex photon-noise bias in the bispectrum. Suppose three fringe frequencies, $\mathbf{u_1}, \mathbf{u_2}$ and $\mathbf{u_3}$ satisfy the closure relation only approximately as,

$$\mathbf{u_1} + \mathbf{u_2} + \mathbf{u_3} = \mathbf{w}, \tag{4}$$

where $\mathbf{w}$ is a small but non-zero residual spatial frequency and $|\mathbf{w} \cdot \mathbf{x_k}| \ll 1$. The triple product of the three DFT components is

$$\begin{aligned} &\tilde{D}_j(\mathbf{u_1})\tilde{D}_j(\mathbf{u_2})\tilde{D}_j(\mathbf{u_3}) \\ &= \tilde{D}_j(\mathbf{u_1})\tilde{D}_j(\mathbf{u_2})\tilde{D}_j(-\mathbf{u_1} - \mathbf{u_2}) + i\tilde{D}_j(\mathbf{u_1})\tilde{D}_j(\mathbf{u_2})\mathbf{w} \cdot \sum_{k=1}^{N_k} \mathbf{x}_k \exp\{i(-\mathbf{u_1} - \mathbf{u_2})\mathbf{x}_k\}, \end{aligned} \tag{5}$$

where $\exp(i\mathbf{w} \cdot \mathbf{x_k}) \approx 1 + i\mathbf{w} \cdot \mathbf{x_k}$ was used. The additional terms

$$\begin{aligned} &i\tilde{D}_j(\mathbf{u_1})\tilde{D}_j(\mathbf{u_2}) \sum_{k=1}^{N_k} \mathbf{w} \cdot \mathbf{x}_k \exp\{i(\mathbf{u_1} + \mathbf{u_2})\mathbf{x}_k\}, \\ &= \tilde{D}_j(\mathbf{u_1})\tilde{D}_j(\mathbf{u_2})\mathbf{w} \cdot \nabla_{\mathbf{u_1}+\mathbf{u_2}} \tilde{D}_j(-\mathbf{u_1} - \mathbf{u_2}), \end{aligned} \tag{6}$$

introduce a complex bias in the measured bispectrum. In principle, this bias can be taken into account provided $\mathbf{w}$ is known. The closure relation through the least-squares fitting was used to adjust the spatial frequencies and thus to ensure $\mathbf{w} = 0$, eliminating the need for the additional bias compensation. Squares of unbiased estimators were also integrated to determine variances. These are used in computation of the signal-to-noise ratio. After integration, mean fringe amplitudes

were obtained from the power spectrum normalized by the zero frequency component and mean closure phases were obtained from bispectral components along with their signal-to-noise ratios.

In order to obtain the maximum signal-to-noise ratio, DFTs were taken with different integration times ranging from 10 ms to 100 ms for a data set taken at one rotational angle of the mask for each object. Since the decorrelation caused by the atmospheric disturbance is continuous, there is no physical frame boundary. For this reason, frames were oversampled and each photon was processed twice. As a result of oversampling, the final signal-to-noise ratio increased by 30% to 40% compared to the case without oversampling and thus the procedure was justified. For the observations of the binary stars presented in this chapter, the optimal integration time was found to be around 20 ms. The behavior of the fringe amplitude, signal-to-noise ratios of the fringe amplitude and closure phase as functions of integration time is discussed in detail in §6.

(iii). *Calibration of the Fringe Amplitude.*

The method of choice for the amplitude calibration depends on the time scale of the calibration and the nature of dominant errors. As an analogy to radio interferometry, we first considered an application of amplitude self calibration. However, it was found to be difficult to carry out amplitude self calibration in optical interferometry and this difficulty is discussed in §6.

An attempted to correct long-term and static systematic errors was made by dividing fringe amplitudes of objects by those of corresponding point source calibrators. The extent to which this correction works depends on the stability of systematic errors. For one of the two binary stars whose images are presented here the correction was fairly successful and for the other binary star it was less successful (§5).

The calibration of object closure phase was not made, since closure phases of point source calibrators did not show a significant deviation from zero. The calibrated amplitudes and the

closure phases were transferred from the CONVEX to the VAX/780 on which the image reconstruction was made.

(iv). *Model Fitting and Hybrid Mapping.*

So far, photon coordinates with arrival times were reduced to calibrated fringe amplitudes and closure phases. The image construction was done using the Caltech VLBI hybrid mapping package (Pearson and Readhead, 1984) in two steps: (a) the best fit model was obtained by direct model fitting to the observables and (b) the final image was obtained by using the observables and the model as inputs to the hybrid mapping.

The input model for hybrid mapping was derived by a least-squares fitting of a model binary star with the calibrated fringe amplitudes and closure phases. A non-linear model-fitting program of the software package was used with some minor modifications. The modification consisted of changing the relative weights of the closure phase and amplitudes. This was done because in the VLBI software package the number of independent closure phases is assumed to be $(n_h - 1)(n_h - 2)/2$ which is true in a regime of high signal-to-noise ratio as in radio interferometry. This number is actually the number of mathematically independent phases in error free phase closure relations. However, data are always corrupted by finite errors and what is really relevant for model fitting is the number of statistically independent closure phases. As shown by Kulkarni (1988) for the Gaussian noise regime and Kulkarni and Nakajima (1989) for the Poisson regime, the number of statistically independent closure phases increases from $(n_h - 1)(n_h - 2)/2$ to $n_h(n_h - 1)(n_h - 2)/6$ as the signal-to-noise ratios of the closure phases decrease. In optical interferometry, it can safely be assumed that all the $n_h(n_h - 1)(n_h - 2)/6$ closure phases are independent, since their signal-to-noise ratios per short exposure frame are generally very low (< 0.1). The relative statistical weighting between the χ^2 of the fringe amplitude and the χ^2 of the closure phase was adjusted by increasing the weight of the χ^2 of the closure phase by

$\frac{n_h(n_h-1)(n_h-2)/6}{n_h(n_h-1)/2} = n_h/3.$

A hybrid mapping procedure was carried out in a single pass without iterations, because of the simplicity of the source structure. The major purposes of hybrid mapping were to obtain a true image with an estimate of dynamic range, to correct for any residual station-dependent amplitude errors and to look for any additional features of the object.

In order to synthesize the image from the *uv* data, closure phases must be reduced to fringe phases. Closure phase relations were algebraically solved with the aid of the model phases (Readhead and Wilkinson 1978). The amplitudes were then adjusted by self calibration using the method of Cornwell and Wilkinson (1981). However, as discussed in §6, baseline-dependent errors seemed to be significant at optical wavelengths and the effectiveness of amplitude self calibration was questionable.

Finally, the Fourier inversion was carried out. Based on the calibrated amplitudes and derived phases, a dirty map and a dirty beam were computed, and the final deconvolution was done using the CLEAN algorithm (Högbom 1974).

5. Results

We have fully reduced three binary stars, β Corona Borealis, σ Herculis and β Leonis Minoris. All of them have been studied by speckle interferometry (McAlister and Hartkopf 1984, references therein). We report the successful imaging of the first two and failure in the third.

β CrB. This is a spectroscopic binary with visual magnitude $\approx$ 3.7 mag (Hoffleit and Jaschek 1982). It has been continuously studied by speckle interferometry since 1973 and the separation ranged from 40 to 310 mas (McAlister and Hartkopf 1984).

Observations of this system were made on the night of 3 April, 1988, using the five hole mask at four different rotational angles. The seeing was about 1″. The *uv* coverage is shown in Fig. 2. 1.2×10^6 photons were recorded at a rate of 5.7 KHz for each rotational angle. For an integration time of 20 ms, there were 114 photons per frame and 2.1×10^5 frames in all, because of the oversampling by a factor of 2. Fringe amplitudes were calibrated by using observations of δ Bootis. The raw fringe amplitudes of the object and the calibrator and the calibrated fringe amplitude are plotted against the projected *uv* plane along the orientation of the binary in Fig. 3. The upper bound of the raw fringe amplitude of the calibrator is about 50%. This is consistent with Monte Carlo simulations of an atmospheric phase screen based on a Kolmogorov spectrum which give an average fringe amplitude of 60% under a condition of 1″ seeing (Chapter III). For the NRM with 15-cm holes, the effect of small-scale fluctuations which degrades wavefront across the aperture and the effect of large-scale fluctuations which causes wandering of the Airy disk appear to be equally significant in reducing the fringe amplitude according to the results of the simulations. A clear modulation in the calibrated fringe amplitude demonstrates the success of the empirical calibration procedure. There were 40 fringe amplitudes and 40 closure phases. Examples of closure phases and their signal-to-noise ratios are shown in Table 1. The average signal-to-noise ratios of the amplitudes and the closure phases were 25 and 8 respectively. We expect that the signal-to-noise ratio in the map will be limited by that of the closure phase and approximated by $8 \times \sqrt{40} = 50$.

The reconstructed image is shown in Fig. 4. A binary system with Δm = 1.47 mag, P.A.= 138.3° and $\Delta\theta$ = 231 mas was well resolved by a beam of 50 mas FWHM. As is mentioned earlier, the P.A. is good to 1° because of the systematic error in the determination of the orientation in the sky. The largest spurious component is minus two percent of the maximum. The dynamic range defined as the ratio between the maximum and the largest spurious component is about 50:1, which is consistent with the signal-to-noise ratio of the closure phases, as discussed above.

σ Her. This is a double line spectroscopic binary with a visual magnitude of 4.2 mag. It was first discovered to be a binary by speckle interferometry in 1972 and the separation ranged from <30 (unresolved) to 119 mas (McAlister and Hartkopf 1984).

The observation was made on 4 April 1988, using the six hole mask at five different rotational angles. The resultant *uv* coverage is shown in Fig. 5. 6.0×10^5 photons were recorded at a rate of 4.2 KHz for each rotational angle. The seeing was about 1.5″. For an integration time of 20 ms, there were 84 photons per frame and 1.4×10^5 frames in all. Fringe amplitudes were calibrated by a point source, *o* Herculis. The raw fringe amplitudes of the object and the calibrator and the calibrated fringe amplitude are plotted against the projected *uv* plane along the orientation of the binary in Fig. 6. The amplitude calibration of σ Her was not as successful as that of β CrB. There were 75 fringe amplitudes and 100 closure phases. The average signal-to-noise ratio of the amplitudes was 13, while that of closure phases was 5. Thus the expected signal-to-noise ratio in the map is dominated by the closure phase error and is $5 \times \sqrt{100} = 50$.

The reconstructed image is shown in Fig. 7. A binary system of a separation of 71 mas was resolved by a beam of 30 mas FWHM. Δm = 2.4 mag. P.A. = -174.4°. The largest spurious component is plus three percent of the maximum. The dynamic range is about 30:1, worse than expected from the signal-to-noise ratio of the closure phases or the amplitudes. We suspect that the decrease in dynamic range from the theoretically expected value is due to systematic errors in the calibrated amplitudes.

β LMi. This is a double line spectroscopic binary consisting of a G8III star with a visual magnitude of 4.2 mag and a F8V star with a visual magnitude of 6.1 mag (Hoffleit and Jascheck 1982). It has been observed by speckle interferometry since 1975, and the separation has been decreasing from 598 mas since 1977 (McAlister and Hartkopf 1984).

The observation was made on 3 April 1988, under a seeing condition of 2″. The optimal

integration time for this object was 5 ms and the mean fringe amplitude of the calibrator (ϵ Ursa Majoris) was below 20% for this integration time. The average signal-to-noise ratio of the amplitudes was 12 while that of the closure phases was 2.5. The low average signal-to-noise ratio of the closure phases is due to the poor seeing condition.

Although a map of this binary star has been presented as a preliminary result (Kulkarni 1988a), we were not able to confirm it by the reorganized image reconstruction procedure described in the previous section. A model fitting only with fringe amplitudes gave Δm = 1.67 mag and $\Delta\theta$ = 322 mas. P.A. = -120° or +60° with a systematic error of $+/-1°$. The quality of closure phases was too low to determine the orientation uniquely.

6. Discussion

In this section we first re-examine the image reconstruction procedure and the estimation of the dynamic range. Then we discuss fundamental differences between radio interferometry and optical NRM.

Potential of Hybrid Mapping for Imaging Unknown Sources.

Since the objects are known binary stars, we explicitly utilized this fact in the data analysis described in §4. In order to test the potential for imaging unknown sources, hybrid mapping was carried out for β CrB and σ Her using a point source as a starting input model instead of using the nature of the sources. It was found that hybrid mapping was effective in revealing companions of both of the binary stars and was successful in obtaining the identical images to the ones shown in §5.

In the case of β CrB, the companion appeared as the largest component among positive components and a negative spurious component was the largest in the CLEAN map after the first

iteration. As a common exercise of hybrid mapping, the CLEAN map was used as the input for model-fitting and the best-fit binary star was obtained. Therefore the image identical to that obtained by the procedure described in §4 was constructed in the second iteration. In the case of σ Her, the companion appeared as the largest component in the CLEAN map after the first iteration. Again model-fitting using the CLEAN map as the input gave the best-fit binary star and the second iteration was identical to the single path hybrid mapping described in §4. A point source is a better starting point for σ Her than for β CrB, since the magnitude difference of σ Her is larger.

The success of hybrid mapping using point sources as starting models implies that the NRM can be very sensitive to a faint secondary feature in the presence of a bright point source as a phase reference. For instance, a map of a 10 mag star with a dynamic range of 100 will contain meaningful features to 15 mag. Simulations of imaging more complicated objects are discussed in Paper I. However, a further investigation is necessary based on actual observing data.

Examination of Dynamic Range.

In §5, the dynamic range was defined crudely as the ratio of the maximum and the largest spurious component in the CLEAN map. Another way of estimating the dynamic range is to see the change of the goodness of fit according to the deviation of the model from the best-fit model. In Table 2, degrees of freedom and χ^2 are shown for different models of β CrB and σ Her: the best-fit binary star model, the best-fit point-source model, a binary star model with an additional flux of 2% to the second component to the best-fit one, a binary star model with an additional flux of 5% to the second component, a triple star model whose third component of 2% is additional to the best-fit binary star model and a triple star model whose third component of 5% is additional to the best-fit model. In the triple star model, three components geometrically form an equilateral right triangle whose right angle vertex is the primary component. Since

the goodness of fit is primarily determined by the systematic errors, fitting parameters do not behave as Gaussian random variables and therefore it is not possible to quantitatively estimate the likelihood of models based on the χ^2 distribution. Here we interpret the combination of the degrees of freedom and χ^2 semi-quantitatively.

In the case of β CrB, the best-fit point-source model is nowhere near the best-fit binary star model, while in the case of σ Her, the best-fit point-source model is not so far from the best-fit binary star model. This is because of the difference in the magnitude differences of the two binary stars. An additional component of 2% to the best-fit model of β CrB appreciably increases χ^2, while one to the best-fit model of σ Her affects χ^2 differently depending on where the additional component is. Apparently a change of flux by 2% is not significant in the case of σ Her. On the other hand, an additional component of 5% causes a significant change in χ^2 in both of the cases. These support the estimate given in §5 that the dynamic range of the map of β CrB is 50:1 and that of σ Her is 30:1.

Fundamental Differences between Radio Interferometry and Non-Redundant Masking.

We now discuss some fundamental differences between radio interferometry and optical NRM. These differences are primarily in the order of correlations of observables, the behavior of signal-to-noise ratios of observables according to the light level and interference-detection scheme of the light. Here the order of correlation of observables, n is defined mathematically as the number of complex fringe amplitudes needed to define a particular observable. For instance, the order of correlation of the power spectrum or fringe amplitude is two and that of the bispectrum or closure phase, three. The behavior of the signal-to-noise ratios of observables depends on the order of correlation and the light level. At high light levels, the signal-to-noise ratios of both the fringe amplitude and closure phase $\propto N^{\frac{1}{2}}$, where N is the number of photon events per integration time. At low light levels, the signal-to-noise ratio of the fringe amplitude $\propto N$, while that of the

closure phase $\propto N^{\frac{3}{2}}$. The general expressions of signal-to-noise ratios of the power spectrum and the bispectrum as functions of the number of photons are obtained by Dainty and Greenaway (1979) and Nakajima (Chapter IV) respectively. Henceforth the behavior of the signal-to-noise ratio of an observable $\propto N^{\alpha}$, is specified by the "SNR exponent", α.

In most cases of radio interferometry except for VLBI, both fringe amplitudes and phases or complex fringe amplitudes are observables. In the case of VLBI, observables are the fringe amplitude (second order correlation) and the closure phase (third order correlation) as in optical NRM. However there is a difference in the behavior of the signal-to-noise ratios of the observables between radio VLBI and optical NRM.

In radio interferometry, one heterodyne receiver is associated with each antenna and the characteristics of the antenna-receiver combinations cause station-dependent errors which dominate overall systematic errors. In NRM, the beams from all the holes (stations) are introduced to a two dimensional incoherent detector where the quantum mechanical interference of each photon takes place (e.g., Dirac 1958). The concept of station becomes vague, since it is impossible to measure the wave function at each hole. Within an integration time, there is only a sparse distribution of photon events which hardly looks like a fringe pattern. If it were not for the atmospheric disturbance, the fringe pattern or complex amplitude would become an observable from an ensemble of many photon events, since the classical fringe pattern would be a probability distribution for individual photons. Instead of the probability distribution, it is necessary because of the atmospheric disturbance, to estimate the second and third order correlations of the probability distribution from ensembles of photon event correlations. The necessity of higher order correlations leads to a very low signal-to-noise ratio per integration time and to the necessity of processing a large number of frames. The low signal-to-noise ratio per integration time sets a limitation in the time scale of amplitude calibration. Four differences found in the practice of

data reduction and analysis are discussed in the following.

Error in Closure Phase.

At radio wavelengths, the amplitude uncertainty (A/σ_A) is related to the phase error σ_ϕ as

$$\sigma_\phi = A/\sigma_A, \tag{7}$$

since α(fringe amplitude) = α(fringe phase) = $\frac{1}{2}$. The error in the closure phase can then be estimated from the propagation of phase errors:

$$\sigma^2_{jkl} = \sigma^2_{jk} + \sigma^2_{kl} + \sigma^2_{lj}, \tag{8}$$

since α(fringe phase) = α(closure phase) = $\frac{1}{2}$.

For an observation of a faint object at optical wavelengths, however, α(fringe amplitude) = 1, the fringe phase is not observable, and α(closure phase) = $\frac{3}{2}$. Therefore there is no simple relation between the amplitude error and the closure phase error.

Optimal Integration Times for Amplitude and Closure Phase.

Since α(fringe amplitude) $\neq$ α(closure phase), we do not necessarily expect the same optimal integration time ($\tau \propto N$) for these two quantities. One advantage of using a photon counting detector is in that the effect of the atmospheric disturbance can be traced as a function of integration time, τ. In Fig. 8, fringe amplitudes of β CrB at a position angle of the mask of 36° are plotted against τ. A slow decay of fringe amplitudes indicates good seeing conditions. Fringe amplitudes of the calibrator showed a very similar decay pattern which was necessary for a successful amplitude calibration. Signal-to-noise ratios of the power spectrum and the bispectrum are plotted as functions of τ in Fig. 9 and Fig. 10 respectively. The signal-to-noise ratio of the power spectrum (or fringe amplitude) peaks at an integration time of 20 ms, while that of bispectrum (or closure phase) has a broad peak between 20 ms and 60 ms. At this low photon

rate ($\approx$ 5 KHz), for a given data set (the total number of photons fixed), the signal-to-noise ratio of the power spectrum is expected to be proportional to

$$V_\tau{}^2 \times \sqrt{\tau}, \tag{9}$$

while that of bispectrum is expected to be proportional to

$$V_\tau{}^3 \times \tau, \tag{10}$$

where τ is the integration time and V_τ is the fringe amplitude as a function of τ (Appendix). Therefore a longer optimal integration time for the bispectrum is expected at a low photon rate. As the light level is lowered, the discrepancy of the optimal integration time for the power spectrum and that for the bispectrum becomes wider. Since the peak of the signal-to-noise ratio of the bispectrum was broad, both the power spectrum and bispectrum were calculated with an integration time of 20 ms for all the data sets taken at different rotational angles of the mask. However, the choice of separate integration times for the amplitude and the closure phase should be considered in observing fainter objects.

Absence of Primary Beam Correction.

Unlike the case in radio aperture synthesis, there is no primary beam correction in the optical NRM method. The 'primary beam correction' is a sensitivity correction made when the field of view of the synthesized map is comparable to the primary beam size of each radio antenna.

A variation of gain within a field of view occurs because each radio antenna is accompanied by a single heterodyne detector which is sensitive to only one angular mode (λ/D). In the case of optical NRM, all the beams from different holes (stations) on the mask are combined onto a two dimensional incoherent detector plane where no variation in sensitivity is seen in principle. Therefore there is no primary beam correction at optical wavelengths. In practice,

a finite bandwidth and physical dimension of the detector limits the field of view. The finite bandwidth effect on the field of view, θ_{FW} is estimated by the relation,

$$\theta_{FW} \sim \frac{\lambda}{\Delta\lambda} \cdot \frac{\lambda}{D}. \tag{11}$$

For the observations of binary stars presented in this chapter, $\theta_{FW} > 600$ mas and the effect is negligible.

Difficulty in Amplitude Self Calibration.

In radio VLBI, two methods of amplitude self calibration, 'closure amplitude' (Readhead et al., 1980) and the method of Cornwell and Wilkinson (1981) are successfully used. Although the two methods are different in approaches, both of them require certain signal-to-noise ratios of closure amplitudes or complex fringe amplitudes within each integration time and the absence of significant baseline-dependent errors. Self calibration is reviewed by Pearson and Readhead (1984, references therein).

At optical wavelengths, the signal-to-noise ratio of a fringe amplitude per integration time is typically very low (< 0.3), because of photon noise. A closure amplitude formed by four fringe amplitudes has much lower signal-to-noise ratio than each amplitude and is not practically measurable even after integrating over many frames. The method of Cornwell and Wilkinson is not applicable either within this short time scale. Realistically there is not any amplitude calibration within the time scale of integration.

At optical wavelengths, amplitude self calibration must be attempted in a time scale of many integrations in order to eliminate photon noise compared to other systematic errors. For the observations reported in this chapter, this time scale was order of several minutes. After the elimination of photon noise, time-averaged amplitudes may include both static or long-term systematic errors and short-term errors. A closure amplitude formed by four time-averaged fringe

amplitudes cancels out long-term station-dependent errors. The quality of the 'long-term closure amplitudes' of the point-source calibrator for β CrB, δ Boo were examined by comparing them with 'fake closure amplitudes' formed by non-closed combinations of two random amplitudes in numerator and two others in denominator. It was found to be poor, showing that the errors of the fringe amplitudes were dominated by short-term station-dependent errors and/or baseline-dependent errors.

A fairly good calibration of the amplitudes of β CrB by dividing those of δ Boo suggested that the atmospheric condition was rather stable during the observation of the pair of stars (§5). For this particular period the amplitude errors must have been dominated by static or long-term errors. Combined with the poor 'long-term closure amplitudes' of δ Boo, the successful division indicated the significance of static or long-term baseline-dependent errors. The presence of baseline-dependent errors reduces the effectiveness of the method of Cornwell and Wilkinson applied to a long time scale. It has been found by Monte Carlo simulations of the effect of the atmospheric disturbance on the NRM technique based on the Kolmogorov spectrum that the reduction of the fringe amplitude by the effect of large-scale disturbances is a major cause of long-term baseline-dependent errors (Chapter III). Astigmatism of the primary mirror is one of the suspected causes of static baseline-dependent errors.

7. Conclusions

We have carried out an experiment of optical aperture synthesis imaging using the non-redundant masking technique at the Hale 5-m telescope. Reduced observables, amplitudes and closure phases were processed by the Caltech VLBI hybrid mapping package.

Main conclusions are:

(1) Images of two bright binary stars were successfully constructed. β CrB, with a separation of 230 milliarcseconds, was resolved by a 50 milliarcsecond beam. The dynamic range of the image was 50:1. σ Her, with a separation of 70 milliarcseconds, was resolved by a 30 milliarcsecond beam. The dynamic range of the image was 30:1. The results demonstrate the potential of the technique for diffraction-limited imaging of astronomical objects with high dynamic range.

(2) In the data reduction, the optimal integration times for the fringe amplitude and the closure phase are not necessarily the same, since the signal-to-noise ratios of these quantities behave differently according the integration time. In general, the optimal integration time for the closure phase is longer than that for the fringe amplitude.

(3) The direct application of radio imaging technique worked well as a starting point for optical aperture synthesis. However there are some major differences between optical and radio regimes, which leave room for improvement of optical imaging.

At optical wavelengths, amplitude calibration is applied only to average fringe amplitudes over many integration times and is not effective to correct short term variations of the atmosphere. Unlike radio interferometry, significant baseline-dependent errors were found in our experiment. A close relation between amplitude errors and phase errors seen at radio wavelengths does not exist at optical wavelengths. As a result one important function of hybrid mapping procedure, amplitude self-calibration, does not have the basis for effectiveness at optical wavelengths. These facts are to be taken into account in future algorithms for optical aperture synthesis imaging.

(4) Primary beam correction of radio interferometry is not necessary in optical aperture synthesis, due to the incoherent detection scheme at optical wavelengths.

Appendix. Signal-to-Noise Ratio vs. Integration Time

The power spectrum and bispectrum analysis of photon-noise-affected images are well understood in terms of the behavior of the signal-to-noise ratio according to the light level (Goodman and Belsher 1976, 1977; Dainty and Greenaway 1979; Wirnitzer 1985; Chapter IV), including the covariance analysis (Kulkarni and Nakajima 1989). In this appendix, these results are applied to photon-noise-limited fringe patterns.

Here the dependence the signal-to-noise ratio of the power spectrum and that of the bispectrum on the integration time and the fringe amplitude, (9) and (10) are derived at low light levels. The discussion given here can easily be generalized to arbitrary light levels using the expression for the signal-to-noise ratio of the power spectrum per short-exposure frame (Dainty and Greenaway 1979) and that of the bispectrum (Chapter IV). It can also be generalized to a covariance analysis to obtain the number of statistically independent fringe amplitudes and that of statistically independent closure phases (Kulkarni and Nakajima 1989).

At low light levels, the signal-to-noise ratio per short-exposure frame of the power spectrum component at spatial frequency $\mathbf{u}$ is given by,

$$\overline{N}\langle|\tilde{i}(\mathbf{u})|^2\rangle, \qquad (A-1)$$

where $\overline{N}$ is the number of photons per frame, $\langle\rangle$ stands for an ensemble average, and $|\tilde{i}(\mathbf{u})|^2$ is the power spectrum component normalized by the zero spatial-frequency component (Dainty and Greenaway, 1979). The signal-to-noise ratio of the bispectrum component at $(\mathbf{u_1}, \mathbf{u_2})$ is given by

$$\overline{N}^{\frac{3}{2}}\langle|\tilde{b}^{(3)}(\mathbf{u_1}, \mathbf{u_2})|\rangle, \qquad (A-2)$$

where $\tilde{b}^{(3)}(\mathbf{u_1}, \mathbf{u_2})$ is the bispectrum component (Wirnitzer 1985; Chapter IV).

In the case of the NRM, each power spectrum component is related to the fringe amplitude V_τ as

$$\langle|\tilde{i}(\mathbf{u})|^2\rangle = \frac{{V_\tau}^2}{{n_h}^2}, \qquad (A-3)$$

where V_τ is the fringe amplitude as a function of the integration time and n_h is the number of holes. A bispectrum component is given by

$$\langle\tilde{b}^{(3)}(\mathbf{u_1},\mathbf{u_2})\rangle = \frac{\langle v_1 v_2 v_3\rangle}{{n_h}^3}, \qquad (A-4)$$

where $v_1 v_2 v_3$ is the triple product of three complex visibilities. In the case of a point source, or a binary star with a large magnitude difference ($\Delta m \geq 1$),

$$|\langle v_1 v_2 v_3\rangle| \approx {V_\tau}^3, \qquad (A-5)$$

since the closure phase is zero or small.

For a data set with the total number of photons, N_{tot}, the signal-to-noise ratio of the integrated power spectrum, S_P is given by,

$$\begin{aligned} S_P &= \overline{N}\frac{{V_\tau}^2}{{n_h}^2}\sqrt{\frac{N_{tot}}{\overline{N}}} \\ &= \frac{\sqrt{RN_{tot}}}{{n_h}^2}{V_\tau}^2\sqrt{\tau}, \end{aligned} \qquad (A-6)$$

where R is the photon rate and thus $\overline{N} = R\tau$. The signal-to-noise ratio of the integrated bispectrum, S_B is given by,

$$\begin{aligned} S_B &= \overline{N}^{\frac{3}{2}}\frac{{V_\tau}^3}{{n_h}^3}\sqrt{\frac{N_{tot}}{\overline{N}}} \\ &= \frac{R\sqrt{N_{tot}}}{{n_h}^3}{V_\tau}^3\tau. \end{aligned} \qquad (A-7)$$

It is intuitively clear that the S_P and S_B peak at different τ and that the optimal integration time for the power spectrum (or fringe amplitude) is shorter. A brief confirmation is given here.

$$\frac{d\log(S_P)}{d\tau} = 2\frac{d\log V_\tau}{d\tau} + \frac{1}{2}\frac{1}{\tau}, \qquad (A-8)$$

while

$$\frac{d\log(S_B)}{d\tau} = 3\frac{d\log V_\tau}{d\tau} + \frac{1}{\tau}. \qquad (A-8)$$

S_B peaks when,

$$2\frac{d\log V_\tau}{d\tau}^* + \frac{1}{2}\frac{1}{\tau^*} = 0, \qquad (A-9)$$

and

$$3\frac{d\log V_\tau}{d\tau}^* + \frac{1}{\tau^*} = -\frac{d\log V_\tau}{d\tau}^*, \qquad (A-10)$$

where * stands for a value at the peak of S_P. Since V_τ is a monotonically decreasing function of τ, $-\frac{d\log V_\tau}{d\tau}^*$ is positive. Therefore when S_P peaks at a given τ^*, S_B is still increasing.

References

Cornwell, T.J. and Wilkinson, P.N. 1981. Mon. Not. R. Astron. Soc. **196**, 106.

Dainty, J.C. and Greenaway, A.H. 1979. J. Opt. Soc. **69**, 786.

Dirac, P.A.M. 1958. *The Principles of Quantum Mechanics* (4th ed.), Oxford University Press.

Goodman, J.W. and Belsher, J.F. 1976, 1977. Tech. Rep. RADC-TR-76-50, RADC-TR-76-382 (Rome Air Development Center, New York), 1977, RADC-TR-77-175 (Rome Air Development Center, New York).

Gorham, P.W., Ghez, A.M., Kulkarni, S.R., Nakajima, T., Neugebauer, G., Oke, J.B., and Prince, T.A. 1989. Submitted to Astro. J..

Haniff, C.A., Mackay, C.D, Titterington, D.J., Sivia, D., Baldwin, J.E., and Warner, P.J. 1987. Nature **328**, 694.

Hoffleit, D. and Jaschek, C. 1982. *The Bright Star Catalog* (4th ed.), Yale University Observatory.

Hofmann, K.-H. and Weigelt, G. 1986. Astron. Astrophys. **167**, L15.

Högbom, J.A. 1974. Astron. Astrophys. Suppl. **15**, 417.

Knox, K.T. and Thompson, B.J. 1974. Astrophys. J. **193**, L45.

Kulkarni, S.R. 1988. In *Proceedings of the Joint ESO/NOAO conference on High-Resolution Imaging by Interferometry*, Merkle, F., ed., (European Southern Observatory, Garching), p 595.

Kulkarni, S.R. 1989. Astron. J.. In publication.

Kulkarni, S.R. and Nakajima, T., 1989. To be submitted to J. Opt. Soc. Am. A.

Labeyrie, A. 1972. Astron. Astrophys. **6**, 85.

McAlister, H.A. and Hartkopf, W.I. 1984. *Catalog of Interferometric Measurements of Binary Stars*, Center for High Angular Resolution Astronomy, Georgia State University.

Nisenson, P., Stachnik, V., Karovska, M., and Noyes, R. 1985. Astrophys. J. **297**, L17.

Pearson, T.J. and Readhead, A.C.S. 1984. Ann. Rev. Astron. Astrophys. **22**, 97.

Readhead, A.C.S and Wilkinson, P.N. 1978. Astrophys. J, **223**, 25.

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

Readhead, A.C.S., Nakajima, T., Pearson, T.J., Neugebauer, G., Oke, J.B. and Sargent, W.L.W. 1988. Astron. J. **95**, 1278. *Paper I.*

Wirnitzer, B. 1985. J. Opt. Soc. Am. **A2**, 14.

hole number	x-coordinate(cm)	y-coordinate(cm)
1	-125.2	27.8
2	-132.9	175.4
3	4.2	130.6
4	21.8	192.9
5	-36.0	200.5
hole combination	fringe amplitude	SNR
12	0.687	27.5
13	0.507	14.2
14	0.818	20.2
15	0.741	25.3
23	0.695	24.4
24	0.830	29.2
25	0.599	24.8
34	0.715	28.9
35	0.575	24.6
45	0.647	27.2
hole combination	closure phase	SNR
123	10.2	6.6
124	-0.1	7.4
125	7.0	9.2
134	0.3	6.1
135	8.2	6.8
145	4.8	8.4
234	-3.1	9.7
235	14.5	6.9
245	-7.0	9.2
345	-24.1	11.0

Table 1 - An example of calibrated data. Calibrated fringe amplitudes and closure phases of β CrB at the position angle of the mask, 144° are shown along with their signal-to-noise ratios and hole coordinates.

β CrB	best fit binary	best fit point	2nd+2%	2nd+5%	triple(3rd=2%)	triple(3rd=5%)
degrees of freedom	76	79	76			
χ^2	234	323	236	268	237	269

σ Her	best fit binary	best fit point	2nd+2%	2nd+5%	triple(3rd=2%)	triple(3rd=5%)
degrees of freedom	171	174	171			
χ^2	275	292	281	297	275	284

Table 2 - χ^2 for different models of β CrB and σ Her. The first row has the star name and the type of models. The second row and third row are degrees of freedom and χ^2 respectively. From the second column to the seventh column are the best-fit binary star model, the best-fit point-source model, a binary star model whose second component is increased by 2% of the maximum, a binary star model whose second component is increased by 5% of the maximum, a triple star model made from the best-fit binary-star model by adding the third component of 2% of the maximum, and a triple-star model made from the best-fit binary star model by adding the third component of 5% of the maximum. In the triple star model, the three components form an equilateral right triangle whose right angle vertex is the primary component.

Figure Captions

Fig. 1 - Schematic of the optics of the prime focus camera. A mask is placed at the image of the primary mirror made by the collimator lens.

Fig. 2 - *Uv* coverage for the observation of β CrB. The unit is not wavelength as usual, but baseline length in cm.

Fig. 3 - Raw fringe amplitudes of β CrB (+), raw fringe amplitudes of δ Boo (o) and calibrated amplitudes (x), plotted against the projected *uv* plane along the orientation of the binary.

Fig. 4 - Reconstructed image of β CrB. Contour levels are -2, -1, 1, 2, 3, 5, 10, 20, 30, 40, 50 and 60% of the maximum. The top to the north and the left to the east.

Fig. 5 - *Uv* coverage for the observation of σ Her. The unit is baseline length in cm.

Fig. 6 - Raw fringe amplitudes of σ Her (+), raw fringe amplitudes of o Her (o) and calibrated amplitudes (x), plotted against the projected *uv* plane along the orientation of the binary.

Fig. 7 - Reconstructed image of σ Her. Contour levels are -2, -1, 1, 2, 3, 5, 10, 20, 30, 40, 50 and 60% of the maximum. The top to the north and the left to the east.

Fig. 8 - Raw fringe amplitudes of β CrB for the mask position angle of 36° plotted against integration time. Ten different symbols denote fringe amplitudes of different baselines.

Fig. 9 - Signal-to-noise ratios of power spectrum components of β CrB for the mask position angle of 36° plotted against integration time. The observed photon rate was 5.7 KHz, the total number of photons was about 1.2×10^5.

Fig. 10 - Signal-to noise-ratios of bispectrum components of β CrB for the mask position angle of 36° plotted against integration time. Different symbols denote different triangles. The observed photon rate was 5.7 KHz, the total number of photons was about 1.2×10^5.

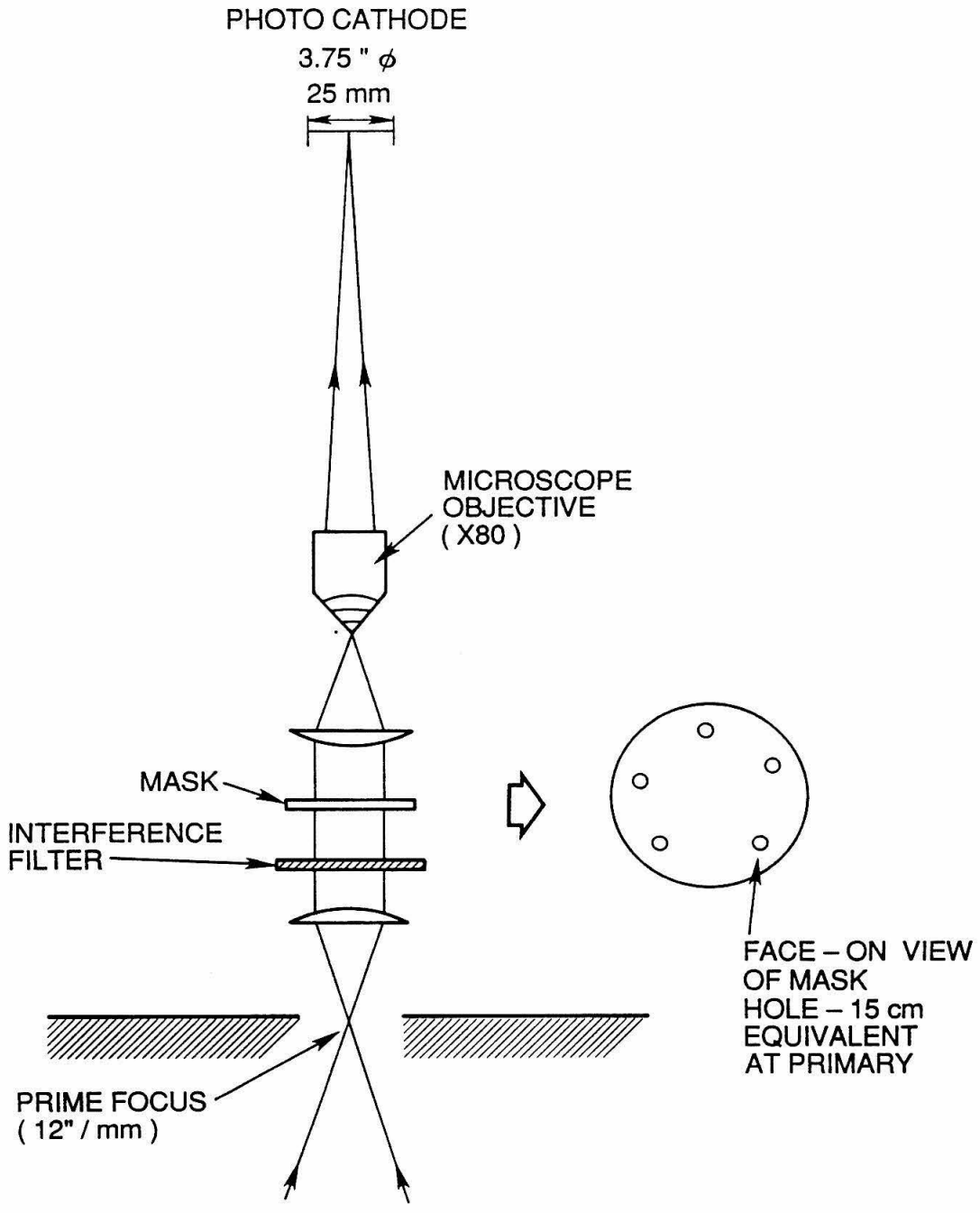
PHOTO CATHODE
3.75 " ϕ
25 mm
MICROSCOPE
OBJECTIVE
(X80)
MASK
INTERFERENCE
FILTER
PRIME FOCUS
(12" / mm)
FACE – ON VIEW
OF MASK
HOLE – 15 cm
EQUIVALENT
AT PRIMARY

Figure 1

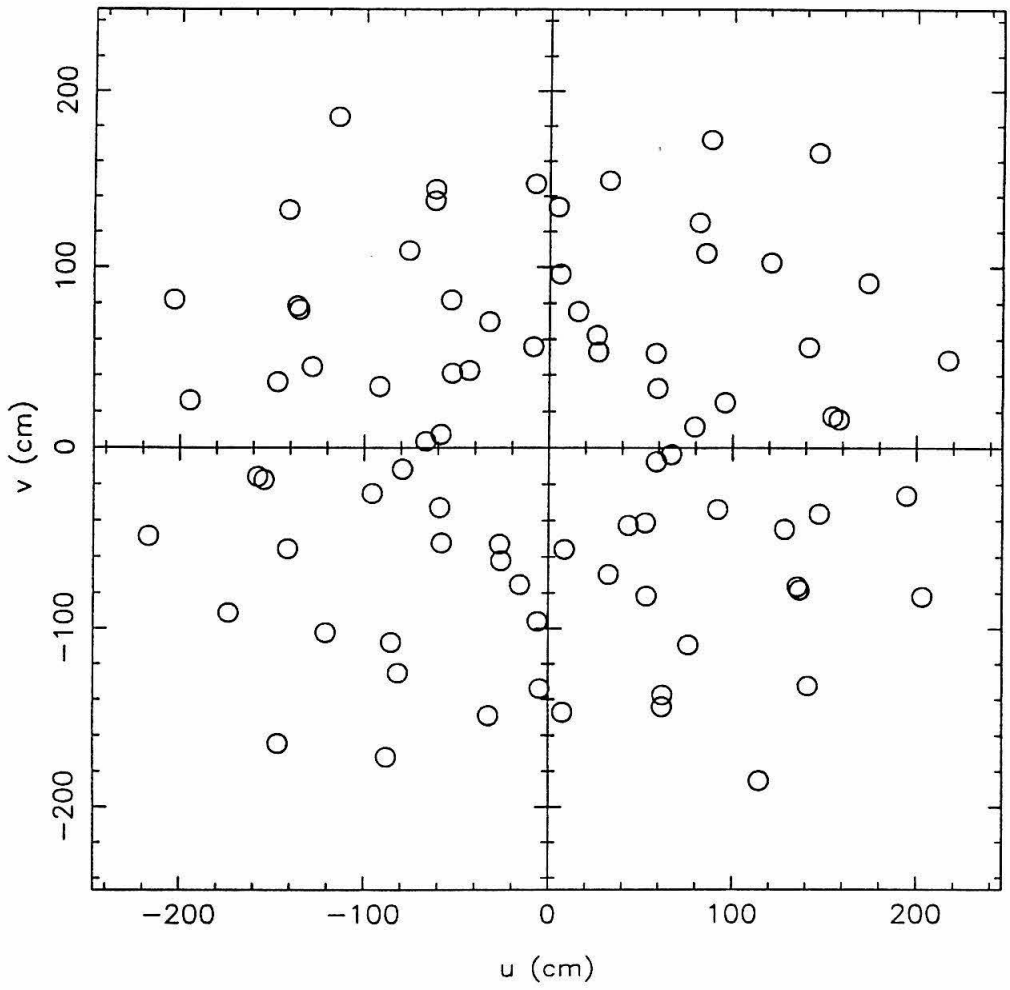

Figure 2

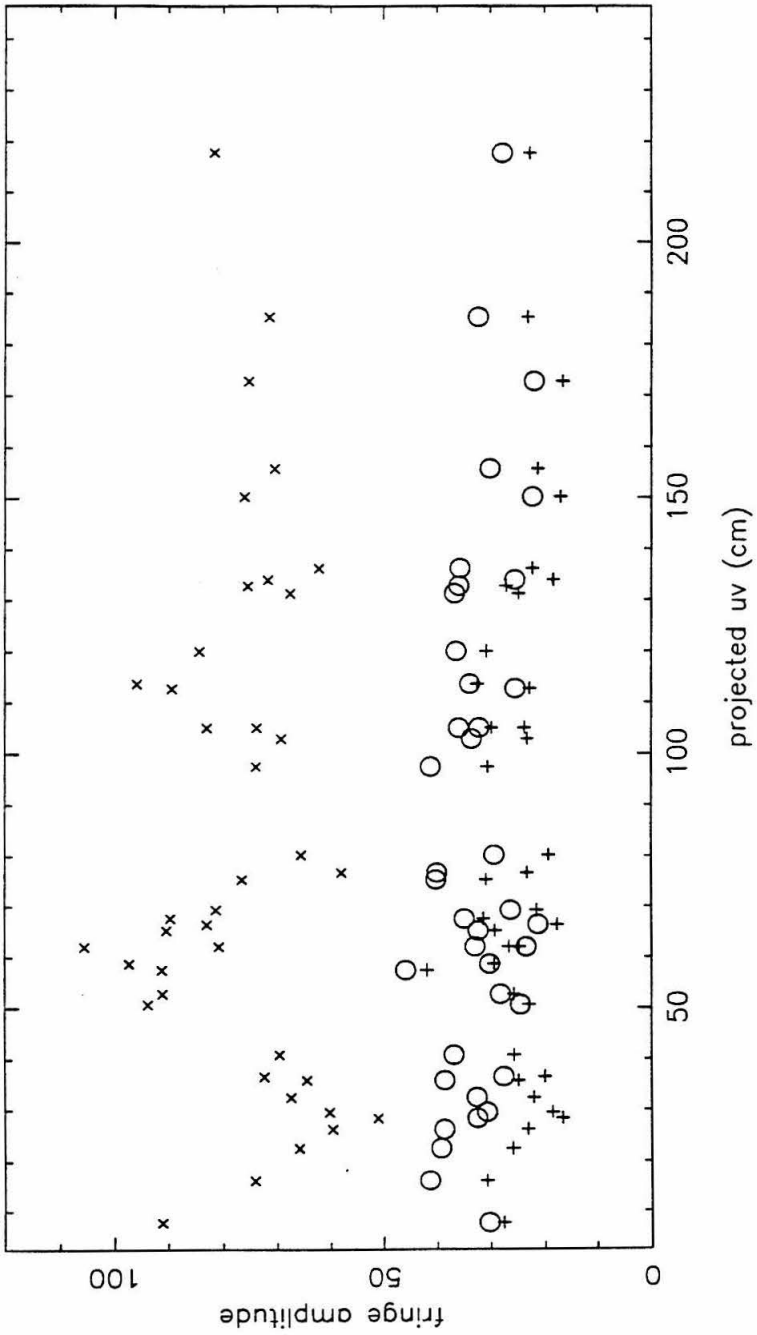

Figure 3

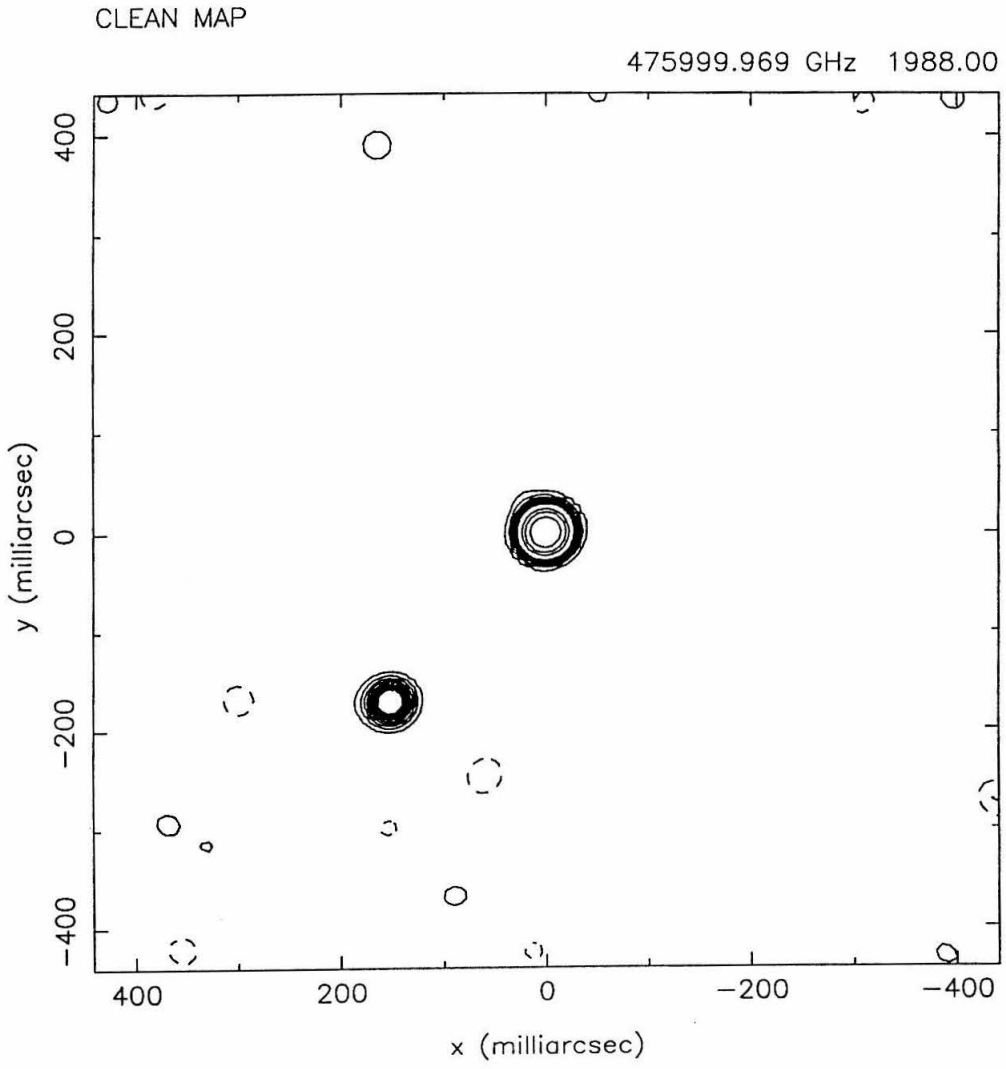
CLEAN MAP
475999.969 GHz 1988.00
400
200
0
−200
−400
y (milliarcsec)
400
200
0
−200
−400
x (milliarcsec)

Figure 4

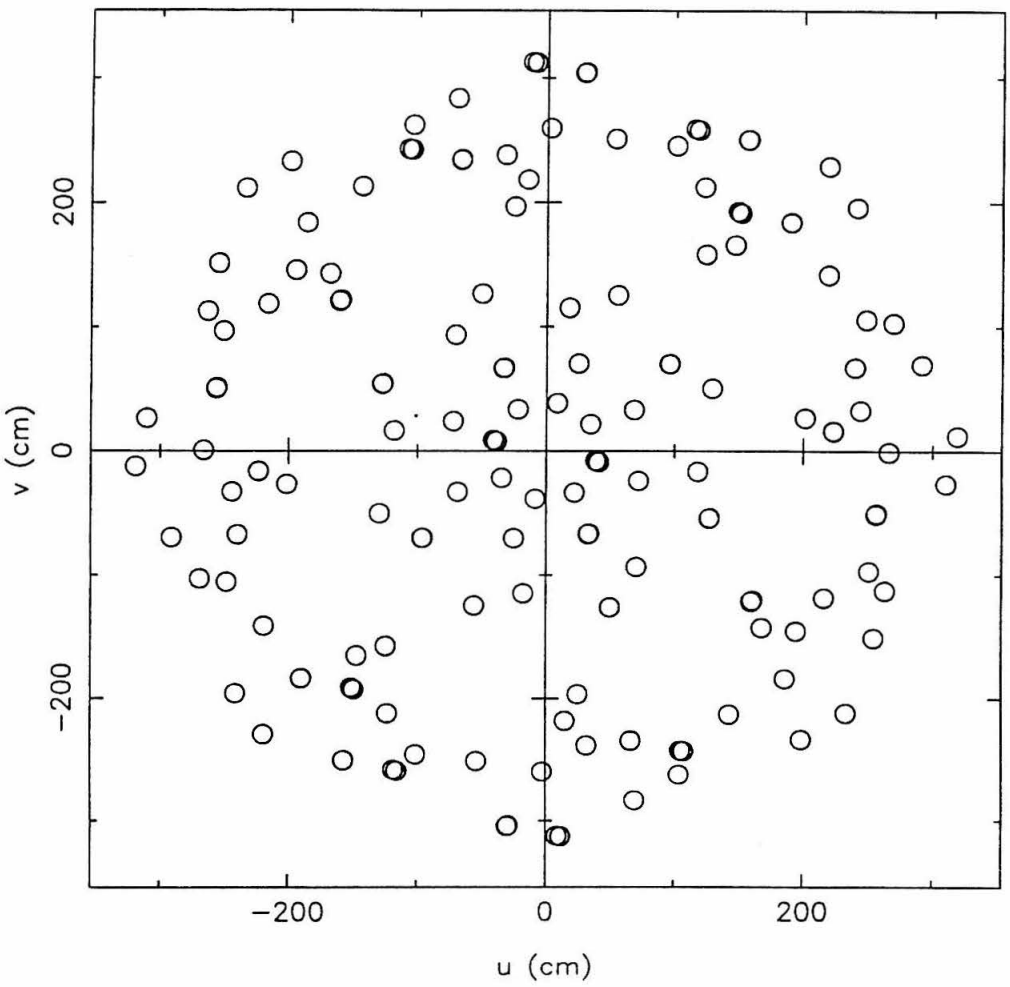

Figure 5

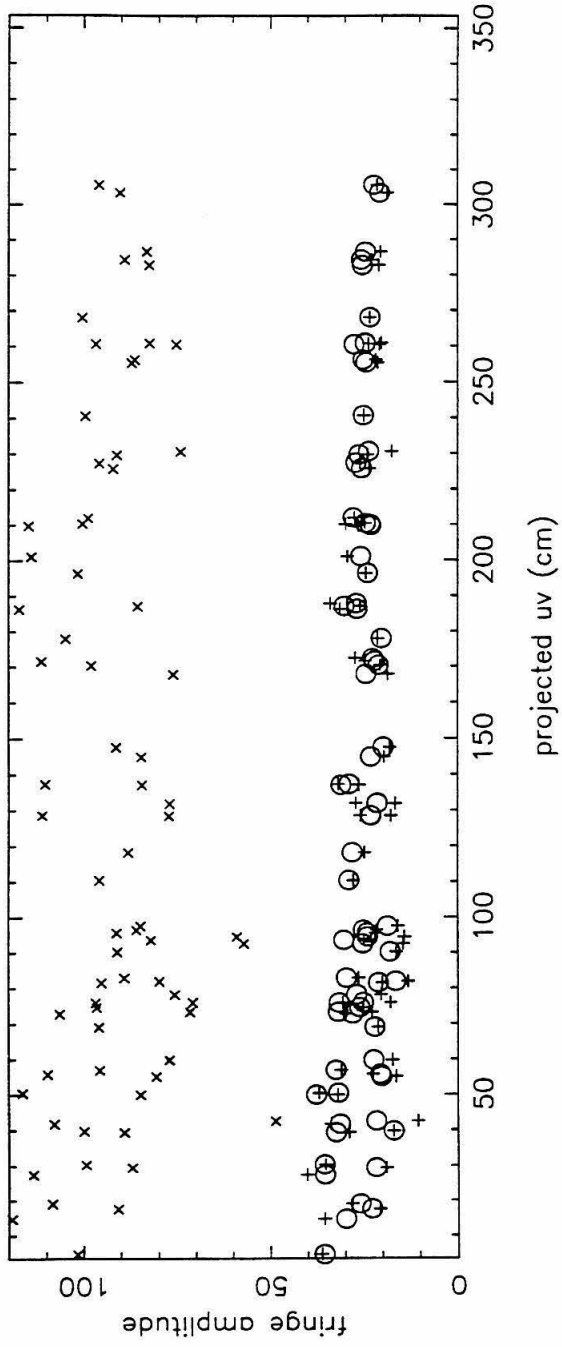
fringe amplitude
0
50
100
projected uv (cm)
50
100
150
200
250
300
350

Figure 6

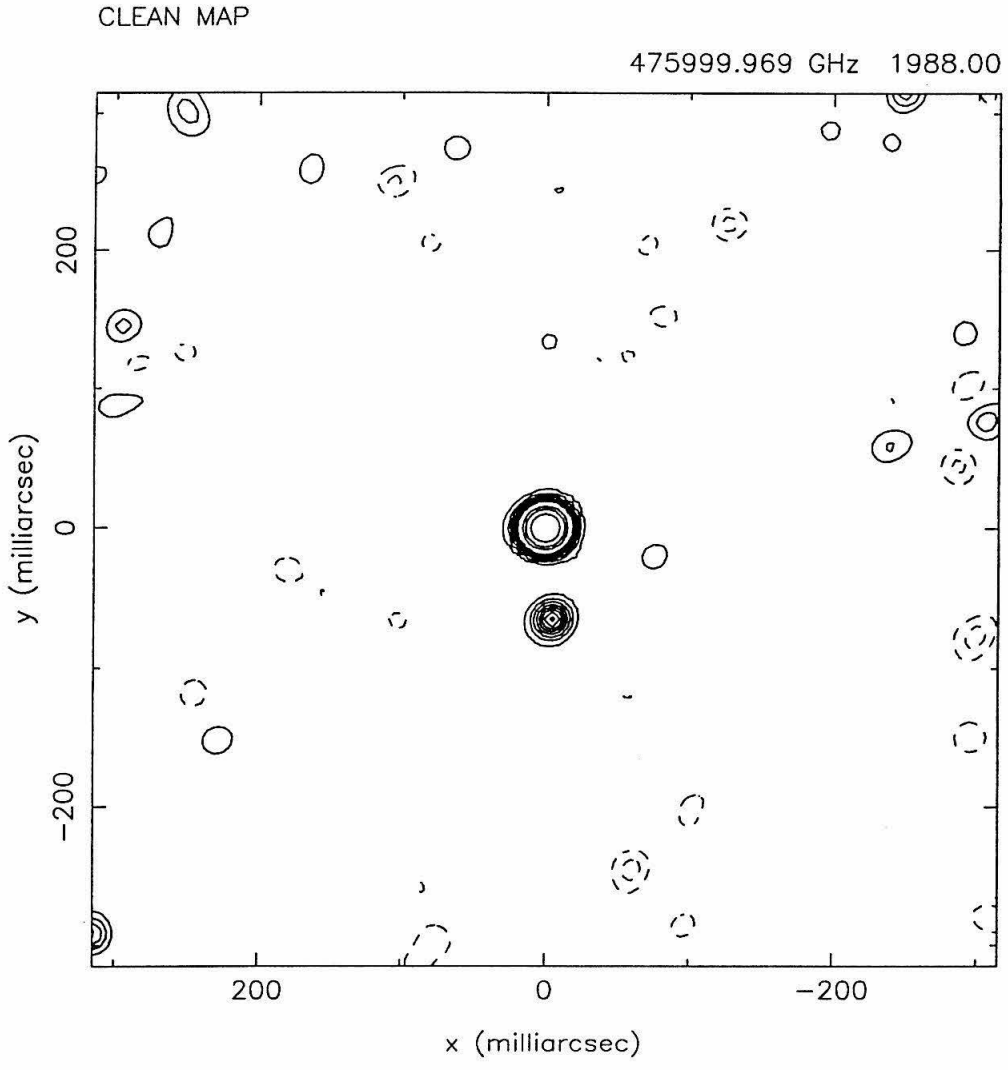
CLEAN MAP
475999.969 GHz 1988.00
200
0
−200
y (milliarcsec)
200
0
−200
x (milliarcsec)

Figure 7

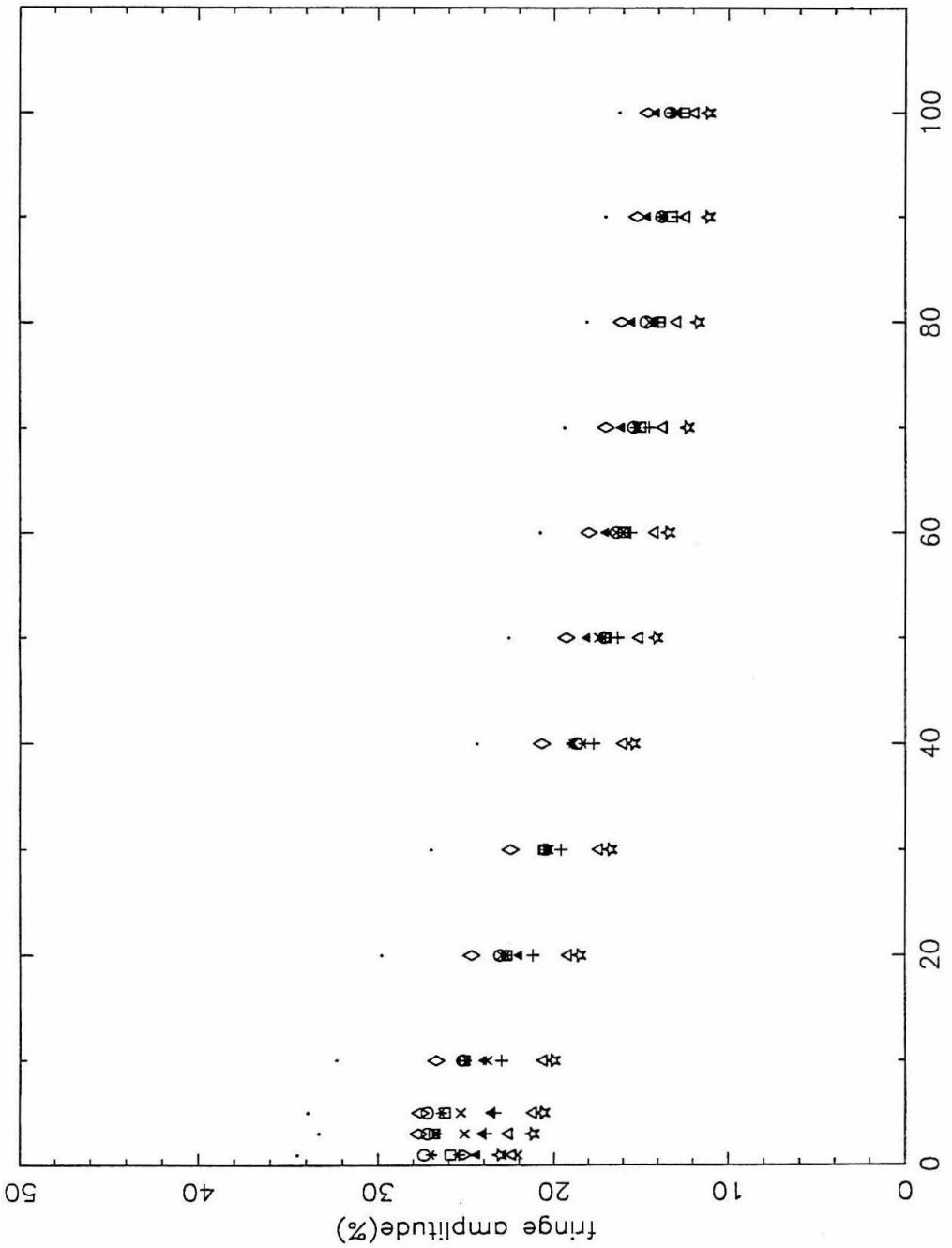
fringe amplitude(%)
0
10
20
30
40
50
0
20
40
60
80
100
integration time (ms)

Figure 8

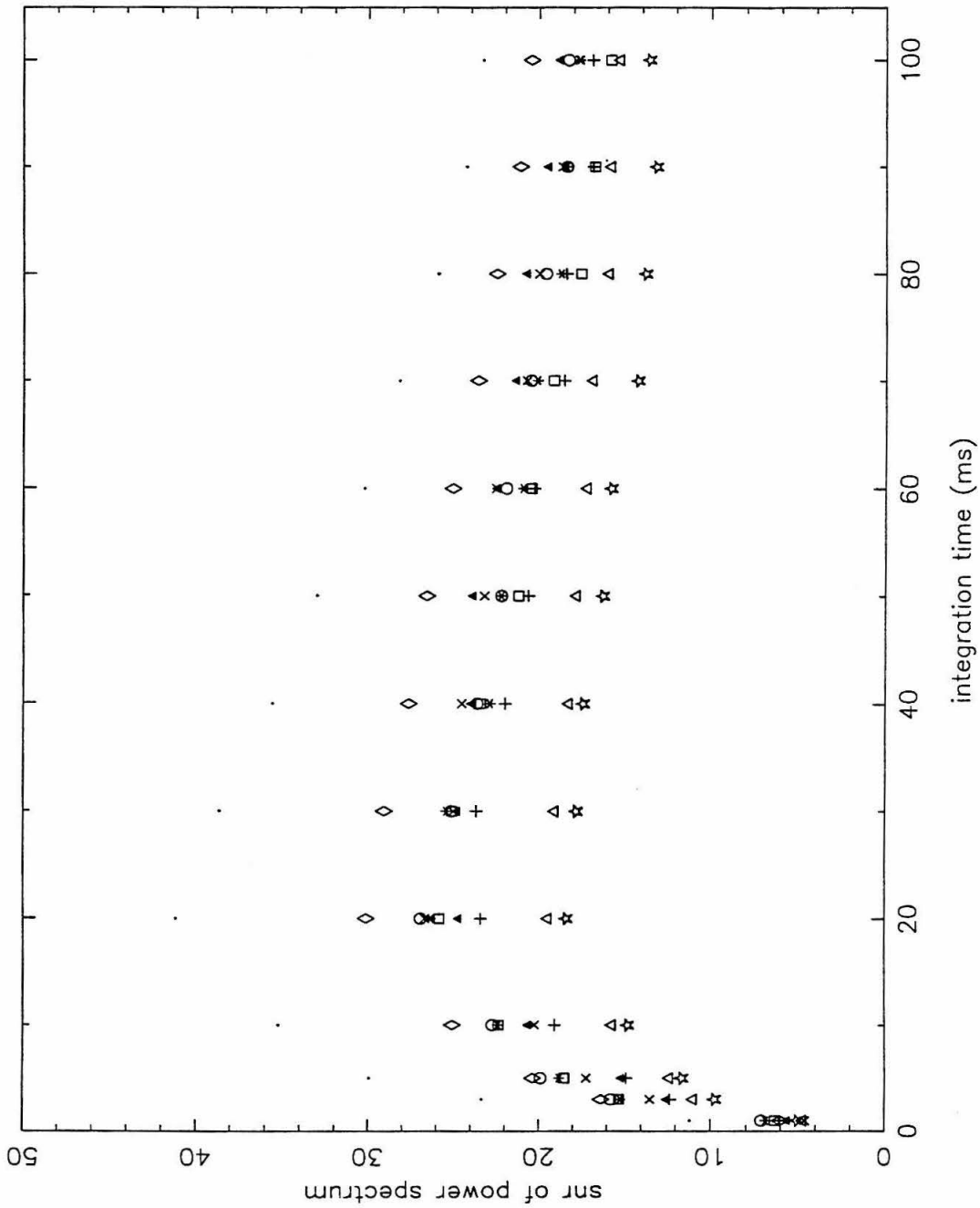
snr of power spectrum
0
10
20
30
40
50
0
20
40
60
80
100
integration time (ms)

Figure 9

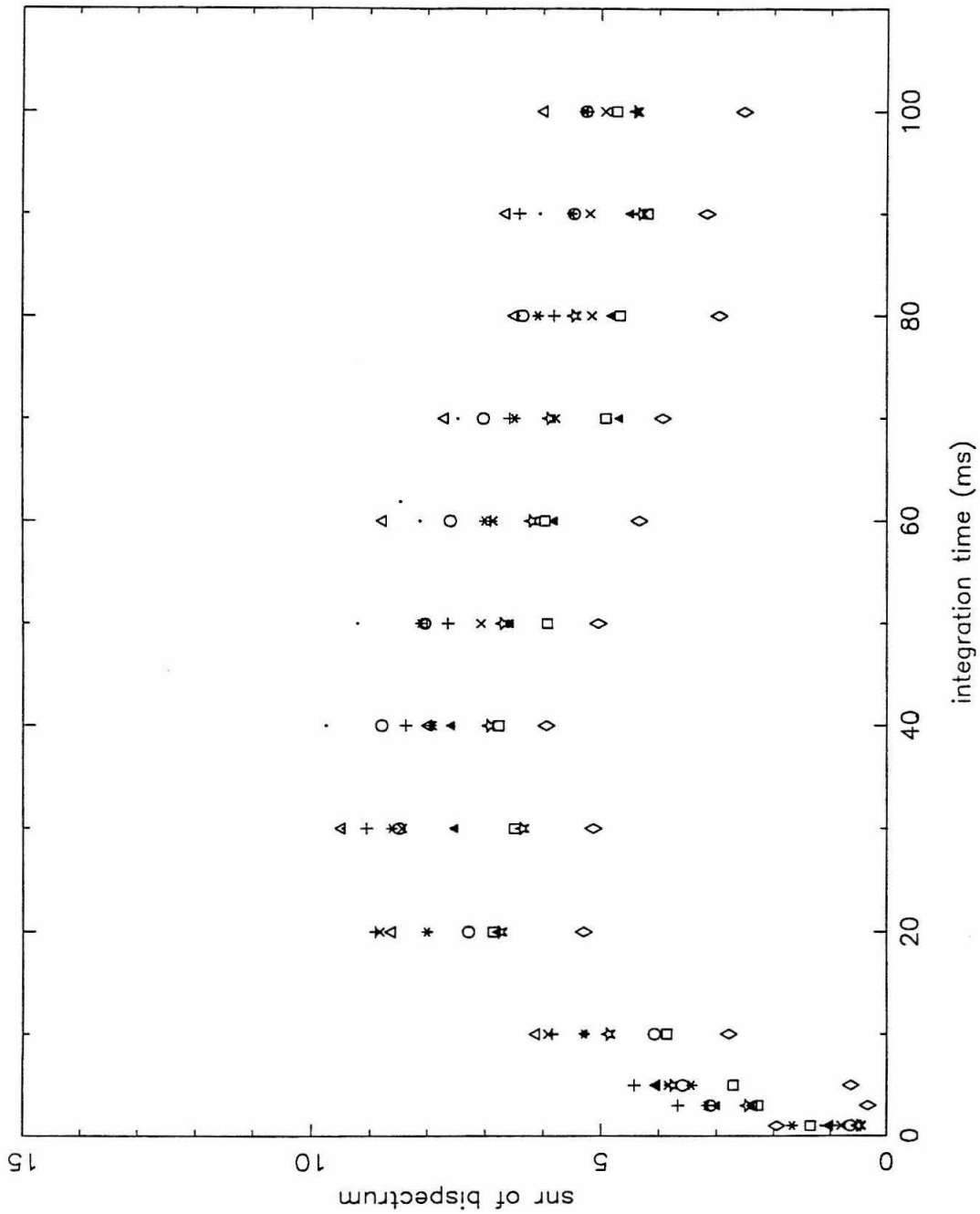
snr of bispectrum
0
5
10
15
0
20
40
60
80
100
integration time (ms)

Figure 10

Chapter III.

Atmospheric Disturbance on Optical Aperture Synthesis

1. Introduction

The method of non-redundant masking (NRM) is a promising technique for high-dynamic-range diffraction-limited imaging with ground-based telescopes. The potential of this technique has been demonstrated by two groups, one at Cambridge (Haniff et al. 1987) and the other at Caltech (Chapter II). In the course of the demonstrations, two major problems were encountered; these were low fringe amplitudes and baseline-dependent amplitude errors. Since the NRM technique is the first step towards optical aperture synthesis with a discrete array of telescopes, it is important to understand the origin of the problems and to find suitable solutions.

It has been observed that the fringe amplitude of a point source is much lower than 100% (Haniff et al. 1987, Chapter II) even under seeing conditions as good as 1″. The maximum fringe amplitude is around 50% for baselines shorter than 1 m, and the fringe amplitude gradually decreases with increasing baseline length. Both disturbances in the atmosphere and imperfections in the telescope optics can be responsible for the reduction of the fringe amplitude. In this chapter, the reduction of the fringe amplitude due to atmospheric disturbances is estimated by simulations based on a Kolmogorov spectrum. Assuming a fringe amplitude of unity for a point source, Kulkarni and Nakajima (1989) obtained a lower bound to the limiting magnitude of the NRM technique, taking into account the covariance analysis of the amplitude and closure phase and thus the signal-to-noise ratio of an ideally constructed map. In this paper, more realistic estimates of the signal-to-noise ratio will be given taking atmospheric disturbances into account.

A fringe pattern formed on the image plane suffers from atmospheric disturbances in two ways:

Small-scale. High frequency fluctuations across each aperture degrade each Airy disk and thus result in station-dependent amplitude reduction.

Large-scale. Low frequency fluctuations prevent different Airy disks from overlapping perfectly on the image plane and thus result in baseline-dependent amplitude reduction.

An advantage of simulations over real observations is that it is possible to simulate the effects of small-scale and large-scale disturbance separately and to study the relative significance by introducing a cut-off frequency in the Kolmogorov spectrum. Separating the small-scale effect and the large-scale effect is also important in evaluating the performance of an active interferometer which in principle can compensate for the path difference and the large-scale effects in real time.

An obvious disadvantage of the NRM technique is the waste of photons. By increasing the aperture size, the fringe amplitude is reduced due to redundancy in baselines, but there is a gain in the number of photons and uv coverage. A transition from the NRM regime with r_c-sized apertures to the speckle regime and trade-offs are studied from the behavior of the fringe amplitude as a function of the aperture size. Throughout this chapter, the size of an aperture or the aperture size refers to the diameter of a circular aperture unless specified otherwise.

The algorithm of the simulations and the computation are described in §2 and §3 respectively. The results of the simulations are presented in §4. A discussion on the simulated results is given in §5. Finally the limiting magnitudes of the NRM technique and masking with 50-cm apertures are derived in §6.

2. Algorithm

The simulations are based on the Kolmogorov theory of turbulence and refractive-index fluctuations (Kolmogorov 1961; Tatarskii 1961, 1967) and on recent observations of the altitude dependence of the refractive-index structure constant ${C_n}^2$ (e.g., La Silla Seeing Campaign: Vernin 1986; Roddier and Roddier 1986). The Kolmogorov theory provides mathematical expressions

of the atmospheric disturbance on a light-wave propagation as two-dimensional spectral densities of phase fluctuation, amplitude fluctuation and the cross spectral density of the two. On the other hand, recent observations indicate that most of the turbulence is produced at the boundary layer and that the high altitude turbulence contributes a relatively small fraction of the overall seeing degradation. We therefore assume that the major disturbance appears as phase fluctuations in the near-field limit and that amplitude fluctuations (scintillation) and the cross correlation of the phase and the amplitude fluctuations are negligible to first order.

In the near-field limit, the spectral density of the phase fluctuation at the aperture plane of a telescope is given by,

$$F_S(\kappa_r) = 0.033 \cdot 2\pi k^2 \Big[\int {C_n}^2(L) dL \Big] {\kappa_r}^{-11/3}, \tag{1}$$

where $\kappa_r = \sqrt{{\kappa_x}^2 + {\kappa_y}^2}$ is the two-dimensional radial spatial frequency, k is the wave number, and $\int {C_n}^2(L) dL$ is the integrated structure constant of the refractive-index fluctuation over the optical path through a turbulent medium.

Since the spectral density specifies only the frequency-dependent variance of the fluctuation but not the probability distribution, we further assume a Gaussian probability distribution with zero mean. In what follows, $\Phi(\mathbf{x})$ denotes the phase at the aperture plane of a telescope and is a real function of $\mathbf{x}$, and $\tilde{\Phi}(\kappa)$ denotes its Fourier transform, which is a conjugate-symmetric-complex function of κ. In the simulations, $\tilde{\Phi}(\kappa)$ is generated at each point over one half of the κ space; $\tilde{\Phi}(\kappa)$ is a complex random number whose modulus is a Gaussian random number with a variance of $F_S(\kappa_r)\Delta\kappa$ and whose phase is a uniform random number between 0 and 2π. $\Delta\kappa$ denotes an area in κ space equaling $(\frac{2\pi}{l_{max}})^2$, where l_{max} is the size of the square phase screen. A conjugate-symmetric-to-real Fourier transform from κ to $\mathbf{x}$ space then creates a monochromatic phase screen $\Phi(\mathbf{x})$ at $\lambda = \frac{2\pi}{k}$. The same algorithm is used in the simulations of speckle interferometry in Chapter IV.

Idealized masking or aperture synthesis is simulated simply by calculating the wave function at multiple circular apertures on the phase screen and its propagation to the image plane. At each point of the aperture plane, the complex amplitude $\Psi(\mathbf{x}) = \exp(i\Phi(\mathbf{x}))$ is calculated. Another Fourier transform simulates the light-wave propagation from the aperture plane to the image plane, and from the squared modulus of the Fourier transform of $\Psi(\mathbf{x})$, a monochromatic fringe pattern, $|\tilde{\Psi}(\mathbf{s})|^2$ is obtained; $\mathbf{s}$ denotes the coordinate on the image plane. Since the phase fluctuations are proportional to the wave number k, finite bandwidth effects are taken into account by averaging over monochromatic fringe patterns at equally spaced wave numbers covering the bandpass. Thus from one evaluation of $\Psi(\mathbf{x})$, multiple $|\tilde{\Psi}(\mathbf{s})|^2$ are generated and averaged to produce one interferogram $I(\mathbf{s})$. If Taylor's hypothesis of frozen-in turbulence (Taylor 1961) and a uniform translation by a constant wind velocity are assumed, a continuous observation can be simulated by considering a series of masks displaced by a distance that is typically the coherence length, r_c.

In the data reduction, the fringe amplitude is obtained statistically to a required accuracy by combining many frames. The Fourier transform $\tilde{I}_j(\mathbf{u})$ of the j^{th} fringe pattern, $I_j(\mathbf{s})$, is taken to form the power spectrum,

$$\tilde{P}_j(\mathbf{u}) = \tilde{P}_j(-\mathbf{u}) = \tilde{I}_j(\mathbf{u})\tilde{I}_j(-\mathbf{u}), \tag{2}$$

where $\mathbf{u}$ denotes the spatial frequency on the image plane. For n speckle interferograms, the sum of power spectra,

$$\tilde{P}(\mathbf{u}) = \sum_{j=1}^{n} \tilde{P}_j(\mathbf{u}), \tag{3}$$

is calculated. Then at each spatial frequency $\mathbf{u}$, corresponding to each mean baseline vector, the fringe amplitude (or visibility) $V(\mathbf{u})$ is calculated as,

$$V(\mathbf{u}) = n_h \sqrt{\frac{\tilde{P}(\mathbf{u})}{\tilde{P}(\mathbf{0})}}, \tag{4}$$

where n_h is the number of holes in the mask. In some cases, the signal-to-noise ratio of $\tilde{P}(\mathbf{u})$ is

also calculated. Since the purpose of this paper is to study the effect of atmospheric disturbances, we study the simple case of a point source.

3. Computation

The simulations were made at a central wavelength of 6300 Å. The aperture configuration was the same as one of the masks used in the experiment discussed in Chapter II and the resultant uv coverage is shown in Fig. 1. The simulations were made with six different aperture sizes of 10, 15, 20, 30, 40 and 50 cm and five different bandwidths of 0, 30, 100, 300 and 1000 Å.

The integrated structure constant of the refractive-index fluctuation,

$$\int C_n^2(L)dL = 5 \times 10^{-13}\ \mathrm{m}^{\frac{1}{3}} \tag{5}$$

was adopted corresponding to 1″ seeing. This is approximately Roddier's value (Roddier 1981). As discussed later in §5, the coherence length r_c, defined as a scale length within which the rms phase fluctuations are smaller than one radian, was 14 cm under these conditions. It was experimentally found that five monochromatic fringe patterns produced at equally spaced wave numbers within the bandpass were enough to obtain a reasonable averaged interferogram.

In order to include wavefront degradation at small scales, the sampling interval on the phase screen must be significantly smaller than the coherence length r_c. On the other hand, the linear dimension of the phase screen has to be significantly larger than the longest baseline so that the large-scale disturbance is simulated properly. Both the sampling interval, Δl, and the size of the phase screen, l_{max}, were determined empirically. It was found that, for sampling intervals, $\Delta l <$ 2 cm, the simulated amplitudes were constant to 1% even for a mask with 10-cm apertures which was most affected by the small-scale disturbance. Hence, Δl was set to 2 cm. Masks with 15-cm

and 50-cm apertures with 300-Å bandpass were simulated with l_{max} = 10 m and 20 m and the results were compared. In the case of the mask with 15-cm apertures, and baselines of 1.5 m or shorter, the resultant fringe amplitudes did not show any difference larger than 1%. Therefore the fringe amplitudes of baselines shorter than 1.5 m were fully determined by the fluctuations whose scale lengths were smaller than 10 m. Fluctuations whose scale lengths were between 10 m and 20 m only caused overall wandering of the superposed fringe pattern, but did not cause the dislocation of the Airy disks. For the baselines longer than 1.5 m, the fringe amplitudes simulated with l_{max} = 20 m were lower than their counterparts with l_{max} = 10 m by at most 2% at the longest baseline. In the simulation of a mask with 50-cm apertures, the resultant amplitudes did not show any difference at all baselines for different l_{max}. l_{max} was set to 10.24 m for the convenience of computation. The array size on the computer was thus $(\frac{l_{max}}{\Delta l})^2 = 512^2$. In producing fringe patterns, 512^2-sized fast Fourier transforms were made to satisfy the Nyquist sampling requirement.

Throughout this chapter, any effect smaller than 2% in the difference of fringe amplitude is regarded as negligible. A comparison of different sizes of the phase screen indicated that for a baseline of length B, any fluctuation of scale length $L > 6B$ was fully negligible. For a sinusoidal fluctuation of wavelength $6B$, an area of the wavefront of the size $B \times B$ is often tilted, but is always flat. Intuitively, sinusoidal fluctuations with scale lengths smaller than $2d$ degrade and spread the Airy disk made by a d-sized aperture, but do not cause wandering of the Airy disk. This disturbance is fully 'small-scale' and the reduction of the fringe amplitude is station (or hole) dependent. For the same reason, fluctuations with scale lengths larger than $6d$ cause wandering of the Airy disk made by a d-sized aperture, but do not cause any significant degradation of the Airy disk. For a multi-aperture mask, the dislocation of the Airy disks reduces the fringe amplitude. This disturbance is fully 'large-scale' and the reduction of the fringe amplitude is baseline dependent. Fluctuations with scale lengths between $2d$ and $6d$ contribute to both types

of disturbances. In this chapter, for a d-sized aperture, atmospheric disturbances by fluctuations with scale lengths smaller than $4d$ are *defined* as 'small-scale disturbances (SSD)' and those with scale lengths larger than $4d$ are defined as 'large-scale disturbances (LSD).' Based on these definitions, the cutoff spatial frequency κ_c of the Kolmogorov spectrum was introduced as

$$\kappa_c = \frac{2\pi}{4d}. \tag{6}$$

In practice a smooth cutoff was adopted by attenuating the Kolmogorov spectrum below or above κ_c by a unit-Gaussian attenuation factor which peaked at κ_c with a width of 10 cm. Hereafter, the Kolmogorov spectra with the lower and upper cutoffs are called the spectrum of small-scale disturbances (SSSD) and the spectrum of large-scale disturbances (SLSD) respectively. The simulations were carried out on a Convex C-1 minisupercomputer.

4. Results

Aperture-Size Dependence of Fringe Amplitudes with Monochromatic Light

Monochromatic fringe amplitudes are plotted as functions of the baseline length for masks with different aperture sizes in Figs. 2, 3, and 4 respectively for the Kolmogorov spectrum, SSSD, and SLSD. As expected, there is no baseline-length dependence of the fringe amplitude for SSSD, but LSD cause a baseline-dependent amplitude reduction.

The aperture dependence of the relative importance between SSD and LSD shows an interesting behavior. By averaging fringe amplitudes over different baselines, the mean amplitude for each mask is obtained. In Fig. 5, the mean amplitudes are plotted as functions of aperture size d. For small apertures, SSD and LSD are equally important, while for large apertures SSD are dominant. For small apertures, the quality of Airy disks is so high that small displacements

of the Airy disks by LSD are equally important as SSD. For large apertures, intrinsically smaller Airy disks are already degraded and spread to the size of a seeing disk by SSD and the effect of LSD is less appreciable.

The fluctuation spectrum (spectral density) of the phase gradient is an indicator of LSD. As shown in Appendix 1, the spectral density of phase gradient follows a $-\frac{5}{3}$ power law. The negative power law exponent implies that the effect of LSD will dominate the disturbance in long-baseline interferometry with small apertures.

Finite-Bandpass Effects

The finite-bandpass causes further decorrelation in fringes. Table 1 summarizes the mean amplitudes of all the simulated combinations of the mask and bandpass for the Kolmogorov spectrum. It is again emphasized that simulated fringe amplitudes are accurate to 2%. Figs. 6, 7 and 8 are the Kolmogorov spectrum, SSSD, and SLSD respectively for bandwidths of 0, 30, 100, 300 and 1000 Å. The effect of finite bandpass on SSD is negligible. As seen from Figs. 6 and 8, finite bandwidth combined with LSD causes baseline-dependent amplitude reduction. For $\Delta\lambda$ = 0, 30 and 100 Å, there is no appreciable effect, but the reduction of the amplitude toward longer baselines is seen for $\Delta\lambda$ = 300 and 1000 Å.

The effects of finite bandpass and aperture size are empirically found to be nearly independent of each other. In Fig. 9, mean fringe amplitudes for six different aperture sizes are plotted as functions of bandpass. The finite bandpass affects the fringe amplitude in the same manner for any aperture size. Therefore the mean amplitude $\overline{V}(d, \Delta\lambda)$ as a function of the aperture size d and bandpass $\Delta\lambda$ is factored as

$$\overline{V}(d, \Delta\lambda) = \overline{V}_0(d) R(\Delta\lambda), \tag{7}$$

where $\overline{V}_0(d)$ is the mean fringe amplitude for monochromatic light and $R(\Delta\lambda)$ is the amplitude-reduction factor dependent only on the bandpass. $R(\Delta\lambda)$ is used in estimating the signal-to-noise

ratio of optical aperture synthesis images in §6.

5. Discussion

Comparison between Simulations and Observations

One of the observations described in Chapter II was made using a five-hole mask with 15-cm apertures and 30-Å bandpass. At 6300 Å, 15 cm corresponds nearly to the lateral coherence length for 1″ seeing. For a point source calibrator, the observed fringe amplitude ranged from 45% to 25% under a condition of 1″. On average, the fringe amplitude decreased from 35% on 60-cm baselines to 30% on 2-m baselines. However a much larger anisotropy in the uv plane ($\approx$ 10%) than the radial dependence was observed on baselines with nearly the same lengths.

The simulated fringe amplitudes with 30-Å bandpass for the Kolmogorov spectrum, SSSD (fluctuation scales shorter than 60 cm) and SLSD (fluctuation scales larger than 60 cm) are plotted in Fig. 10. In all the simulations, the rms error in fringe amplitudes is smaller than 2%. The simulated fringe amplitude with the Kolmogorov spectrum monotonically decreased from 63% at shorter baselines to 60% at longer baselines. With SSSD, fringe amplitudes of 73% were obtained, independently of baseline length. With SLSD, the simulated fringe amplitude showed a weak decreasing tendency from 78% at 60-cm baselines to 73% at 2-m baselines. Both the effects of SSD and LSD were equally significant in the reduction of the fringe amplitude. However, observed fringe amplitudes were typically only half of the simulated fringe amplitudes. Because of the uncertainty in the estimate of the seeing condition in the observations, precise comparison between the results of the simulations and the observations is impossible. However, the simulated fringe amplitudes seem to give an upper bound of the observed fringe amplitudes.

The most striking qualitative discrepancy between the simulations and the observations is the large anisotropy of the observed fringe amplitude in the uv plane. This anisotropy must

originate from either imperfections in the telescope optics or from incorrect assumptions of the statistical isotropy in the Kolmogorov theory. One clue to distinguish the telescope imperfections from the atmospheric disturbance may be given by the time variability of the anisotropy. As discussed in Chapter II, a fairly successful baseline-dependent calibration of the amplitude of the object by that of the calibrator indicated that the anisotropy was almost static over a period of 20 minutes, and thus we conclude that the anisotropy was caused by the telescope imperfections.

Transition from NRM Regime to Speckle Regime

By a linear fitting of a logarithmic plot (Fig. 5), the aperture dependence of the fringe amplitude with SSSD (o), V_D, for apertures larger than 30 cm, is empirically given by

$$V_D = \left(\frac{14}{d}\right)^{1.0}. \tag{8}$$

From the heuristic interferometric view of the image-forming process that is applicable to the effect of SSD (Rogers 1963), the fringe amplitude is expected to be the reciprocal of the square root of the redundancy of the baseline $\left(\frac{d}{r_c}\right)^2$ (Korff et al. 1972, Appendix 2 of Chapter IV) and thus $V = \frac{r_c}{d}$. Therefore the coherence length, r_c, estimated from the interferometric view is 14 cm at 6300 Å with the integrated structure constant of 5×10^{-13} $m^{-\frac{1}{3}}$. For aperture sizes of 50 cm or larger, SSD are the dominant cause of the amplitude reduction and aperture masking is essentially in the speckle regime. A smooth transition from the NRM regime to the speckle regime is seen in Fig. 5.

Origin of Finite-Bandpass Effects

In the uv plane, the diameter of a statistically independent spatial frequency circle at $\frac{B}{\lambda}$ determined by atmospheric disturbances is $\frac{r_c}{\lambda}$ for monochromatic light. By widening the band-width, uncorrelated phasors of baselines adjacent to B at wavelengths adjacent to λ, incoherently blend into the phasor at the spatial frequency, $\frac{B}{\lambda}$. If it were not for the effect of LSD, no blending

should occur for a r_c-sized aperture, because of the geometrical absence of the adjacent baselines. Therefore no amplitude reduction caused by the finite bandpass is seen in the simulation of NRM with SSSD. On the other hand, obvious reduction with a wide bandpass is seen for 50-cm apertures with SSSD because of the geometrical presence of neighboring baselines (Fig. 11). The finite-bandpass effect becomes significant when

$$\frac{\Delta\lambda}{\lambda} \approx \frac{r_c}{B}. \tag{9}$$

For B = 2 m, $\Delta\lambda \approx 300$ Å, the estimate is therefore consistent with the simulations.

For a long-baseline interferometer with r_c-sized apertures, the finite-bandpass effect induced by LSD will dominate the amplitude reduction. By a real-time compensation of both the path difference and the effect of LSD, high quality fringes affected only by SSD can in principle be obtained.

6. Signal-to-Noise Ratio of Aperture Masking

The limiting magnitudes of the NRM technique and that of masking with apertures larger than r_c are estimated at low light levels. These estimates are based on the results of simulations, on the theory discussed by Kulkarni and Nakajima (1989), and on the experience with real observations described in Chapter II.

Light Level Dependence

Before proceeding to the discussion of limiting magnitudes, the light level dependence of the signal-to-noise ratio (SNR) of the power spectrum ($|\text{fringe amplitude}|^2$) of NRM is considered in comparison with that of an ideal Michelson interferometer. The SNR of a power spectrum

component per short-exposure frame is given by

$$SNR = \frac{(\frac{V}{n_h})^2 \overline{N}^2}{\sqrt{\overline{N}^2 + 2\overline{N}^3 (\frac{V}{n_h})^2 + \overline{N}^4 {\sigma_A}^2}} = \frac{V^2 \overline{n}}{\sqrt{{n_h}^2 + 2 n_h V^2 \overline{n} + {n_h}^4 \overline{n}^2 {\sigma_A}^2}}, \quad (10)$$

where V, n_h, $\overline{N}$, $\overline{n} = \frac{\overline{N}}{n_h}$ and ${\sigma_A}^2$ are the fringe amplitude, the number of holes, the mean photon counts per short-exposure frame, the mean photon counts per hole per short-exposure frame and the variance of the power spectrum component purely due to atmospheric disturbances respectively (e.g., Kulkarni and Nakajima 1989). Since the object under consideration is a point source, the fringe amplitude V represents the atmospheric disturbance. We consider the case of a five-hole mask with 15-cm apertures and 100-Å bandpass. For ideal Michelson interferometry free from atmospheric disturbances, $V = 1.0$ and $\sigma_A = 0.0$. The simulation of the NRM including atmospheric disturbances gives $V = 0.6$ and $\sigma_A = 0.01$. In Fig. 12, the SNRs of the ideal Michelson interferometer and the simulated NRM are plotted as functions of $\overline{n}$. At low light levels ($\overline{n} < 10$), the SNRs of both the ideal Michelson interferometer and the simulated NRM are proportional to $\overline{n}$. The SNR of the simulated NRM is lower than that of the ideal Michelson by a factor of $V^2 = 0.36$ for a given $\overline{n}$. At high light levels ($\overline{n} > 10$), the SNR of the ideal Michelson interferometer is proportional to $\overline{n}^{\frac{1}{2}}$, while that of the simulated NRM saturates to 1.4. For an observing system with a net throughput of η%, the magnitude at the transition light level $\overline{n} = 10$ is

$$M_{tr} = 6.1 + 2.5 \log \eta \quad \text{mag.} \quad (11)$$

Low Light Levels

In Table 2, expressions of the SNR of a power spectrum component ($\approx$ SNR of fringe amplitude), a bispectrum component ($\approx$ SNR of closure phase) and the total SNRs in the power spectrum and bispectrum are given in terms of the fringe amplitude, V, the number of holes, n_h, the mean photon counts per integration time, $\overline{n}$ and the number of short-exposure frames, M

(Appendix 2.). Here the total SNR in the power spectrum is defined by

$$\text{SNR of a power spectrum component} \times \sqrt{\text{number of statistically independent components}},$$

and the total SNR in the bispectrum is defined by

$$\text{SNR of a bispectrum component} \times \sqrt{\text{number of statistically independent components}}.$$

The total SNRs in the power spectrum and the bispectrum can also be interpreted as the SNRs of an ideally constructed autocorrelation function (ACF) of an object and an ideally constructed map of the object respectively. Since the total SNR in the bispectrum (third order quantity) is always lower than that in the power spectrum (second order quantity), the SNR of a map is determined by the total SNR of the bispectrum which supplies phase information. A detailed discussion of this subject is found in Kulkarni (1989).

The expression 'ideally' is used since some practical complications were not included in these estimates of the SNRs of the ACF and the map. Individual short-exposure frames are limited by photon noise. The same expressions for the SNRs would be obtained for an error-free fringe amplitude instead of the mean fringe amplitude affected by the atmospheric disturbance. However, after eliminating the predominance of photon noise by an ensemble average of $M(\gg 1)$ frames, the quality of a data set is limited by long-term systematic errors caused by the atmospheric disturbance (Chapter II) which do not show up in the given expressions. Any calibration problem originating from these hidden systematic errors limits the total SNR, especially of the power spectrum, and hence the SNRs of both the ACF and the map. One obvious way to avoid this calibration problem is to keep the mean fringe amplitude high and thus minimize the effects of atmospheric disturbances. If this is the only solution for the calibration problem, high-dynamic-range images can be obtained only by interferometry with r_c-sized apertures.

Limiting Magnitude of Non-Redundant Masking

The limiting magnitude of the NRM technique is estimated under rather strict constraints for high-dynamic-range imaging. The SNR of a bispectrum component is maximized by using a three-hole mask. However, the *uv* coverage is so poor that mapping with a three-hole mask is not very practical. By using a mask with more than seven holes the SNR of the map does not increase, but the SNR of each bispectrum component decreases. Therefore a seven-hole mask is nearly optimal for a non-redundant configuration (Kulkarni and Nakajima 1989). If a seven-hole mask with 15-cm apertures is used at $\lambda = 6300$ Å, a constant fringe amplitude of $V = 0.6$ is expected for up to 2-m baselines with $\Delta\lambda \leq 100$ Å. Two conservative constrains are imposed based on the experience of data reduction and analysis described in Chapter II. First, a minimum SNR of three is required for each bispectrum component. Second, the maximum observing time per aperture configuration (e.g., rotational angle of the mask) required for the purpose of a successful amplitude calibration (between point sources, a cross calibration) is assumed to be on the order of ten minutes. An integration time of 20 milliseconds is adopted as in the observation described in Chapter II. Henceforth the limiting magnitude, $M_{lim}(3\sigma; 10; 20)$, is obtained with the assumptions of a ten-minute observing time and 20-millisecond integration times. The corresponding total number of frames, M, is about 5×10^4 including the over-sampling factor of two. The inferred number of photons per hole is $\overline{n} = 1$. For an observing system with a net throughput of $\eta\%$, the resultant limiting magnitude is

$$M_{lim}(3\sigma; 10; 20) = 8.6 + 2.5 \log \eta \quad \text{mag.} \tag{12}$$

With the current technology of photo-multipliers the best-case design of the throughput η will be 10%. The limiting magnitude of the NRM technique for high-dynamic-range imaging is then around 11 mag. For each configuration, the SNR of each power spectrum component is 12. The SNRs of the ACF and the map will be 50 and 18 respectively. It should again be emphasized that the 3σ limit is imposed on a bispectrum component, while the constraint on the observing time is imposed by amplitude calibration. Under conditions that T-minute observations and integration

times of τ milliseconds are allowed, the limiting magnitude, $M_{lim}(3\sigma; T; \tau)$, will be given by

$$M_{lim}(3\sigma; T; \tau) = M_{lim}(3\sigma; 10; 20) + 1.25 \log(T/10) + 2.5 \log(\tau/20) \quad \text{mag.} \qquad (13)$$

With several different aperture configurations, a map with a dynamic range over 100 will be obtained. In the presence of a point source reference of 11 mag or brighter, features as faint as 16 mag can be imaged. If a laser guiding star brighter than 11 mag is successfully created (Foy and Labeyrie 1985, Thompson and Gardner 1987), but not bright enough to activate adaptive optics, the NRM technique may well be the method of choice for high-dynamic-range imaging.

Masking in Speckle Regime

As discussed in §5, a smooth transition from the NRM regime to the speckle regime is seen when the aperture size is increased from 10 cm to 50 cm. Beyond d = 30 cm, the amplitude reduction is dominated by the effect of SSD. With 50-cm apertures, aperture masking is practically in the speckle regime. The limiting magnitude of aperture masking in the speckle regime is of practical interest, since many of the currently planned optical-aperture-synthesis arrays have aperture sizes around 50 cm (e.g., COAST, Mackay and Baldwin 1988). Therefore the sensitivity limit of the masking with 50-cm apertures sets the upper bound of those of long-baseline interferometers since more severe atmospheric disturbances are expected on longer baselines.

For $d \geq 50$ cm, the fringe amplitude is expected to behave as $V = \frac{r_0}{d}$ for $\Delta\lambda \leq 100$ Å. Table 3 shows the SNRs of a power spectrum component and bispectrum component, and the total SNRs. Statistically independent spatial frequency components are extended in the uv plane. The mean amplitude, V is constant and equals $\frac{r_0}{d}\frac{1}{n_h}$ over a circle of diameter $\frac{d}{\lambda}$ in the uv plane. The SNR of the fringe amplitude per statistically independent spatial component does not depend on the aperture size (see Table 3). The reduction in amplitude and the increase in the number of photons are balanced in the photon noise limit. Therefore if the systematic errors caused by the

atmospheric disturbance are negligible, the SNR in the ACF will be gained simply by better uv coverage of larger apertures.

In the following, the same observing parameters as used to test the NRM technique are first assumed aiming at high-dynamic-range mapping. If a seven-hole mask is used with a bandpass of 100 Å, the 3σ limit in the bispectrum per configuration for a ten-minute observation corresponds to $\overline{n}_c = 0.9$, where $\overline{n}_c$ is the number of photon events per r_c-sized coherent cell. $\overline{n}_c = 0.9$ is nearly the same object brightness as $\overline{n} = 1.0$ for the NRM with 15-cm apertures. At this flux level, the SNR in amplitude is 29, the SNRs of ACF and map are 466 and 218 respectively. In reality, such a high dynamic range has never been seen in images produced by speckle interferometry. The hidden atmospheric effect must be well understood especially in this speckle regime. Is the estimated SNR of the map wrong or is there any calibration procedure which overcomes this problem? The limiting magnitude for potential high-dynamic-range imaging is the same as that for the NRM technique. For objects brighter than $\overline{n}_c \approx 1$, the NRM technique is advantageous over masking with larger apertures unless a good calibration procedure exists in the speckle regime.

Now we relax the constraints on the bandpass, the number of holes and the limiting SNR per closure phase, in order to pursue the faintest limit. As discussed in §5, the amplitude-reduction factor due to the finite bandpass, $R(\Delta\lambda)$, is independent of aperture size d. Then, $G(\Delta\lambda) = R(\Delta\lambda)\sqrt{\frac{\Delta\lambda}{100}}$ represents a gain in the apparent SNR compared to 100 Å bandpass because of the increased number of photons. G^2 and G^3 correspond to the gain factors in power spectrum and bispectrum respectively. From Fig. 11, $R(\Delta\lambda)$ is obtained at a given baseline. For a 2-m baseline, $G(300) = 1.55$ and $G(1000) = 2.0$. This gain factor seems to saturate beyond 1000 Å, judging from the change of the slope. $\Delta\lambda = 1000$ Å will be the practical maximum bandpass and the maximum gain in the SNR of the bispectrum component is 8. Since there are many more statistically independent bispectrum components, $n_h = 7$ is no longer a constraint. Here

the minimal value, n_h = 3 is adopted. Compared to NRM which has fully discrete triangles, masking with large apertures has a contiguous distribution of triangles. Since contiguous data points are easier to handle (e.g., by smoothing) than fully discrete data points, the requirement for the minimum SNR per bispectrum component could be reduced to one. Taking all these gain factors into account, the limiting magnitude for 1σ detection with a three-hole mask with 50-cm apertures and 1000-Å bandpass is

$$M_{lim}(1\sigma; 10; 20) = 11.8 + 2.5 \log \eta. \tag{14}$$

At this limit, the SNR of the map is 12.

7. Conclusions

In this chapter the effects of the atmospheric disturbance on the method of NRM and masking with apertures larger than the coherence length are studied by Monte Carlo simulations based on the Kolmogorov theory of refractive-index fluctuations. The simulations are made at 6300 Å with different bandwidths ranging from 0 to 1000 Å under a seeing condition of 1″. The baseline length ranges from 60 cm to 2.2 m. The limiting magnitudes of the NRM technique and masking with 50-cm apertures are estimated, based on the results of the simulations, the theory by Kulkarni and Nakajima (1989) and the practical experience described in Chapter II.

Major conclusions are as follows:

(1) The mean fringe amplitude simulated for the NRM technique with r_c-sized apertures is about 60%. This value gives an upper bound for the observed fringe amplitudes which are also affected by telescope imperfections. The effects of small-scale disturbances (SSD) and large-scale disturbances (LSD) are equally significant in the reduction of the fringe amplitude.

(2) A smooth transition from the NRM regime to the speckle regime is simulated by increasing the aperture size from 10 cm to 50 cm. In the speckle regime, the monochromatic fringe amplitude is consistent with the value predicted by the heuristic interferometric view of image-forming process. A real-time compensation of the effect of LSD does not improve the mean fringe amplitude for apertures larger than $2r_c$.

(3) A criterion for a significant amplitude reduction for a baseline B is given by $\frac{\Delta\lambda}{\lambda} \geq \frac{r_c}{B}$. For the NRM technique, the effect of LSD combined with the finite bandwidth causes baseline-dependent amplitude reduction. For larger apertures, the finite-bandwidth and the effect of SSD cause the mixing of uncorrelated baseline phasors at a given spatial frequency because of the geometrical presence of multiple baselines corresponding to that particular spatial frequency. The fractional reduction of fringe amplitudes according to the bandwidth is nearly independent of the aperture size.

(4) The limiting magnitudes for high-dynamic-range imaging are obtained with the following assumptions; the 3σ limit in each bispectrum component, an observing time of ten minutes for a successful amplitude calibration, and an integration time of 20 milliseconds. For an observing system with a net throughput of $\eta\%$, the estimated limiting magnitudes, $M_{lim}(3\sigma; 10; 20)$ of the NRM technique and the masking with 50-cm apertures and 100-Å bandwidth were the same, and are $8.6 + 2.5 \log \eta$. With the best system currently available, $M_{lim}(3\sigma; 10; 20) \approx 11$ mag. For objects brighter than the limiting magnitude given above, NRM will be the method of choice for the purpose of high-dynamic-range imaging unless a good calibration scheme is found in the speckle regime. The faintest limit for the phase-closure imaging will be pursued by relaxing the requirements for the detection limit of each bispectrum component to 1σ and the bandpass to 1000 Å. For the masking with 50-cm apertures, the limiting magnitude, $M_{lim}(1\sigma; 10; 20)$ is $11.8 + 2.5 \log \eta$ mag. With the best system currently available, $M_{lim}(3\sigma; 10; 20) \approx 14$ mag.

Under conditions that T-minute observing times and τ-millisecond integrations are allowed, the limiting magnitude will be given by, $M_{lim}(T;\tau) = M_{lim}(10;20) + 1.25\log(T/10) + 2.5\log(\tau/20)$ mag.

Appendix 1. Kolmogorov Spectrum of Phase Gradient

In order to characterize the effect of LSD quantitatively, it would be desirable to obtain the correlation function of the angle θ of the wavefront normal, $D_\theta(r)$. However, θ is a non-linear function of the phase S and it is next to impossible to calculate $D_\theta(r)$ analytically. Instead, the Kolmogorov spectrum of fluctuations of phase gradient, $F_{|\nabla S|}$, is derived from the Kolmogorov spectrum of phase fluctuations, $F_S(\kappa_r)$, as an indicator of the effect of LSD.

First the power spectrum of $\frac{\partial S}{\partial x}$ is derived as

$$\begin{aligned} F_{\frac{\partial S}{\partial x}}(\kappa_x, \kappa_y) &= F_{\delta'(x)}(\kappa_x, \kappa_y) \times F_S(\kappa_r) \\ &= |i\frac{\kappa_x}{2\pi}|^2 \times F_S(\kappa_r). \end{aligned} \qquad (A-1)$$

Similarly,

$$F_{\frac{\partial S}{\partial y}}(\kappa_x, \kappa_y) = |i\frac{\kappa_y}{2\pi}|^2 \times F_S(\kappa_r). \qquad (A-2)$$

Because of statistical isotropy in the turbulence,

$$\begin{aligned} F_{|\nabla S|}(\kappa_r) &= F_{\frac{\partial S}{\partial x}} + F_{\frac{\partial S}{\partial y}} \\ &= (\frac{\kappa_r}{2\pi})^2 F_S(\kappa_r) \\ &\propto \kappa_r^{-\frac{5}{3}}. \end{aligned} \qquad (A-3)$$

Appendix 2. Signal-to-Noise Ratios of Power Spectrum and Bispectrum

The power spectrum and bispectrum analysis of of photon noise affected images are well understood in terms of the behavior of the signal-to-noise ratios (SNR) according to the light levels. (Goodman and Belsher 1976, 1977; Dainty and Greenaway 1979; Wirnitzer 1985; Chapter IV). These results are applied to photon-noise-limited fringe patterns to derive the formulae of SNRs given in Tables 2 and 3.

By extending the variance analysis to the covariance analysis, the number of statistically independent fringe amplitudes and that of statistically independent closure phases have been obtained (Kulkarni and Nakajima 1989). By using the results of the covariance analysis, SNR of a point source in a synthesized map are given in Tables 2 and 3.

At low light levels, the SNR per short-exposure frame of the power spectrum component at spatial frequency $\mathbf{u}$ is given by

$$\overline{N}\langle|\tilde{i}(\mathbf{u})|^2\rangle, \qquad (A-4)$$

where $\overline{N}$ is the number of photons per frame, $\langle\rangle$ stands for an ensemble average, and $|\tilde{i}(\mathbf{u})|^2$ is the power spectrum component normalized by the zero spatial frequency component (Dainty and Greenaway 1979). The SNR of the bispectrum component at $(\mathbf{u_1}, \mathbf{u_2})$ is given by

$$\overline{N}^{\frac{3}{2}}\langle|\tilde{b}^{(3)}(\mathbf{u_1}, \mathbf{u_2})|\rangle, \qquad (A-5)$$

where $\tilde{b}^{(3)}(\mathbf{u_1}, \mathbf{u_2})$ is the bispectrum component (Wirnitzer 1985).

In the case of NRM, each power spectrum component is related to the fringe amplitude V as

$$\langle|\tilde{i}(\mathbf{u})|^2\rangle = \frac{V^2}{{n_h}^2}, \qquad (A-6)$$

where V is the visibility amplitude and n_h is the number of holes. A bispectrum component is given by

$$\langle \tilde{b}^{(3)}(\mathbf{u_1}, \mathbf{u_2}) \rangle = \frac{\langle v_1 v_2 v_3 \rangle}{{n_h}^3}, \qquad (A-7)$$

where $v_1 v_2 v_3$ is the triple product of three complex visibilities. In the case of a point source,

$$|\langle v_1 v_2 v_3 \rangle| \approx V^3, \qquad (A-8)$$

since the mean closure phase is zero. This approximation is valid only if V is not very noisy or not too small compared to unity (> 0.3).

By introducing the number of photons per hole, $\overline{n}_h$, $\overline{N} = n_h \overline{n}$, for an ensemble of M frames, the SNR of the amplitude (power spectrum) is,

$$\overline{N} \langle |\tilde{i}(\mathbf{u})|^2 \rangle \sqrt{M} = V^2 \frac{\overline{n}}{n_h} \sqrt{M}. \qquad (A-9)$$

Similarly, the SNR of the closure phase (bispectrum) is,

$$\overline{N}^{\frac{3}{2}} \langle |\tilde{b}^{(3)}(\mathbf{u_1}, \mathbf{u_2})| \rangle \sqrt{M} = V^3 (\frac{\overline{n}}{n_h})^{\frac{3}{2}} \sqrt{M}. \qquad (A-10)$$

In the case of aperture masking in the speckle regime, each power spectrum component is related to the fringe amplitude V as

$$\langle |\tilde{i}(\mathbf{u})|^2 \rangle = \frac{1}{{n_h}^2} (\frac{r_c}{d})^2, \qquad (A-11)$$

where r_c is the coherence length and d is the aperture size (Korff et al. 1972, Appendix 1 of Chapter IV).

A bispectrum component is given by

$$\langle \tilde{b}^{(3)}(\mathbf{u_1}, \mathbf{u_2}) \rangle = \frac{1}{{n_h}^3} (\frac{r_c}{d})^4 \qquad (A-12)$$

(Readhead et al. 1988, Appendix 1 of Chapter IV).

By introducing the number of photons per coherent cell, $\overline{n}_c$, $\overline{N} = n_h \overline{n}_c (\frac{d}{r_c})^2$. The SNR of amplitude is given as

$$\overline{N} \langle |\tilde{i}(\mathbf{u})|^2 \rangle \sqrt{M} = \frac{\overline{n}_c}{n_h} \sqrt{M}. \qquad (A-13)$$

The SNR of closure phase is given as

$$\overline{N}^{\frac{3}{2}} \langle |\tilde{b}^{(3)}(\mathbf{u_1}, \mathbf{u_2})| \rangle \sqrt{M} = (\frac{r_c}{d})(\frac{\overline{n}_c}{n_h})^{\frac{3}{2}} \sqrt{M}. \qquad (A-14)$$

At low light levels, all the fringe amplitudes and the closure phases are effectively statistically independent (Kulkarni and Nakajima 1989). In the case of NRM, there are $n_h(n_h - 1)/2$ amplitudes and $n_h(n_h - 1)(n_h - 2)/6$ closure phases. In the case of the masking in the speckle regime, there are $n_h(n_h - 1)/2 \times (\frac{d}{r_c})^2$ amplitudes and $n_h(n_h - 1)(n_h - 2)/6 \times (\frac{d}{r_c})^4$ closure phases. Eq. (A-9), multiplied by $\sqrt{n_h(n_h - 1)/2}$ gives the SNR of an ideally constructed amplitude map. Similarly, the SNR of a phase map of NRM and those of an amplitude map and a phase of the masking in the speckle regime are obtained as given in Tables 2 and 3.

References

Dainty, J.C. and Greenaway A.H. 1979. J. Opt. Soc. Am. **69**, 786.

Foy, R. and Labeyrie, A. 1985. Astro. Astrophys., **152**, L29.

Goodman, J.W. and Belsher, J.F. 1976, 1977. ARPA Order No. 2646, Technical Report RADC-TR-76-50. (Rome Air Development Center, NY, 1976). ARPA Order No. 2646, Technical Report RADC-TR-76-382, (Rome Air Development Center, NY, 1976). Technical Report RADC-TR-77-175 (Rome Air Development Center, NY, 1977).

Haniff, C.A., Mackay, C.D., Titterington, D.J., Sivia, D., Baldwin, J.E. and Warner, P.J. 1987. Nature 328, 694.

Kolmogorov, A.N. 1961. In *Turbulence, Classical Papers on Statistical Theory*, Friedlander, S.K. and Topper, L., eds. (Wiley-Interscience, New York).

Korff, D., Dryden, G. and Miller, M.G. 1972. Opt. Comm., **5**, 187.

Kulkarni, S.R. 1989. Astron. J. In publication.

Kulkarni, S.R. and Nakajima, T. 1989. To be submitted to J. Opt. Soc. Am. **A**.

Mackay, C.D. and Baldwin, J.E. 1988. In *Proceedings of NOAO-ESO conference on High-Resolution Imaging by Interferometry*, Merkle, F., ed. (European Southern Observatory, Garching), p 565.

Readhead, A.C.S., Nakajima, T., Pearson, T.J., Neugebauer, G., Oke, J.B. and Sargent, W.L.W. 1988. Astron. J, **95**, 1278.

Roddier, C. and Roddier, F. 1986. In *Proceedings of the Second Workshop on ESO's Very Large Telescope*, D'Odorico, S. and Swings, J.P., eds. (European Southern Observatory, Garching), p 269.

Roddier, F. 1981. In *Proceeding of the ESO conference on The Scientific Importance of High Angular Resolution at Infrared and Optical Wavelength*, Ulrich, M.H. and Kjar, K., eds. (European Southern Observatory, Garching), p 5.

Rogers, G.L. 1963. Proc. Phys. Soc. **81**, 323.

Tatarskii, V.I. 1961. *Wave Propagation in a Turbulent Medium*, McGraw-Hill, New York. (Translation by R.S. Silverman).

Tatarskii, V.I. 1967. *The Effects of the Turbulent Atmosphere on Wave Propagation.* Moscow, Nauka. (translation by Israel Program for Scientific Translations, Keter Press, Jerusalem, 1971).

Thompson, L.A. and Gardner, C.S. 1987. Nature, **328**, 229.

Vernin J. 1986. in *Proceedings of Second Workshop on ESO's Very Large Telescope*, D'Odorico, S. and Swings, J. P., eds. (European Southern Observatory, Garching, 1986), p 279.

Wirnitzer, B. 1985. J. Opt. Soc. Am. **A2**, 14.

Amplitudes averaged over baselines

$d \setminus \Delta\lambda$	0 Å	30 Å	100 Å	300 Å	1000 Å
10 cm	0.78	0.76	0.76	0.64	0.43
15 cm	0.62	0.61	0.63	0.54	0.35
20 cm	0.51	0.50	0.51	0.43	0.29
30 cm	0.38	0.38	0.38	0.33	0.24
40 cm	0.30	0.30	0.29	0.26	0.17
50 cm	0.25	0.25	0.24	0.21	0.15

Table 1.

SNR of Non-Redundant Masking

NRM	SNR	Total SNR
Power spectrum	$V^2 \frac{\overline{n}}{n_h} \sqrt{M}$	$V^2 \overline{n} \sqrt{\frac{n_h - 1}{2 n_h}} \sqrt{M}$
Bispectrum	$V^3 (\frac{\overline{n}}{n_h})^{\frac{3}{2}} \sqrt{M}$	$V^3 \overline{n}^{\frac{3}{2}} \sqrt{\frac{(n_h - 1)(n_h - 2)}{6 {n_h}^2}} \sqrt{M}$

Table 2.

SNR in Speckle Regime

Speckle Regime	SNR	Total SNR
Power spectrum	$\frac{\overline{n}_c}{n_h} \sqrt{M}$	$(\frac{d}{r_c}) \overline{n}_c \sqrt{\frac{n_h - 1}{2 n_h}} \sqrt{M}$
Bispectrum	$(\frac{r_c}{d})(\frac{\overline{n}_c}{n_h})^{\frac{3}{2}} \sqrt{M}$	$(\frac{d}{r_c}) \overline{n}_c^{\frac{3}{2}} \sqrt{\frac{(n_h - 1)(n_h - 2)}{6 {n_h}^2}} \sqrt{M}$

Table 3.

Figure Captions

Fig. 1 - UV coverage of the simulated five-hole mask. The unit is in meters.

Fig. 2 - Aperture-size dependence of the fringe amplitude for the Kolmogorov spectrum. $\Delta\lambda$ = 0 Å. Different symbols denote different aperture sizes. d = 10, 15, 20, 30, 40 and 50 cm from the top to the bottom.

Fig. 3 - Aperture-size dependence of the fringe amplitude for SSSD. $\Delta\lambda$ = 0 Å. Different symbols denote different aperture sizes. d = 10, 15, 20, 30, 40 and 50 cm from the top to the bottom.

Fig. 4 - Aperture-size dependence of the fringe amplitude for SLSD. $\Delta\lambda$ = 0 Å. Different symbols denote different aperture sizes. d = 10, 15, 20, 30, 40 and 50 cm from the top to the bottom.

Fig. 5 - Mean fringe amplitudes plotted as functions of the aperture size in logarithmic scale. $\Delta\lambda$ = 0 Å. From the top to the bottom, three different symbols denote SLSD (x), SSSD (o) and the Kolmogorov spectrum (*).

Fig. 6 - Finite-bandwidth effects on NRM with the Kolmogorov spectrum. d = 15 cm. Different symbols denote different bandwidths. $\Delta\lambda$ = 0, 30, 100, 300 and 1000 Å from the top to the bottom.

Fig. 7 - Finite-bandwidth effects on NRM with SSSD. d = 15 cm. $\Delta\lambda$ = 0, 30, 100, 300 and 1000 Å.

Fig. 8 - Finite-bandwidth effects on NRM with SLSD. d = 15 cm. Different symbols denote different bandwidths. $\Delta\lambda$ = 0, 30, 100, 300 and 1000 Å from the top to the bottom.

Fig. 9 - Finite-bandwidth and aperture size. The mean fringe amplitudes are plotted as functions of the bandpass in logarithmic scale. $\Delta\lambda$ = 30, 100, 300, 1000 Å. Different symbols denote different aperture sizes. d = 10, 15, 20, 30, 40 and 50 from the top to the bottom. The

fractional amplitude reduction caused by the finite-bandwidth effect is nearly independent of the aperture size.

Fig. 10 - NRM with monochromatic light. d = 15 cm, $\Delta\lambda$ = 0 Å. Circles, crosses and square boxes are the fringe amplitudes for the Kolmogorov spectrum, SSSD, and SLSD respectively.

Fig. 11 - Atmospheric disturbance on masking in the speckle regime. d = 50 cm. Different symbols denote different bandwidths. From the top to the bottom, $\Delta\lambda$ = 0, 30, 100, 300 and 1000 Å.

Fig. 12 - SNR of an ideal Michelson interferometer with 15-cm apertures (*) and that of NRM with 15-cm apertures (solid line). SNR of NRM saturates at high light levels due to the atmospheric disturbance.

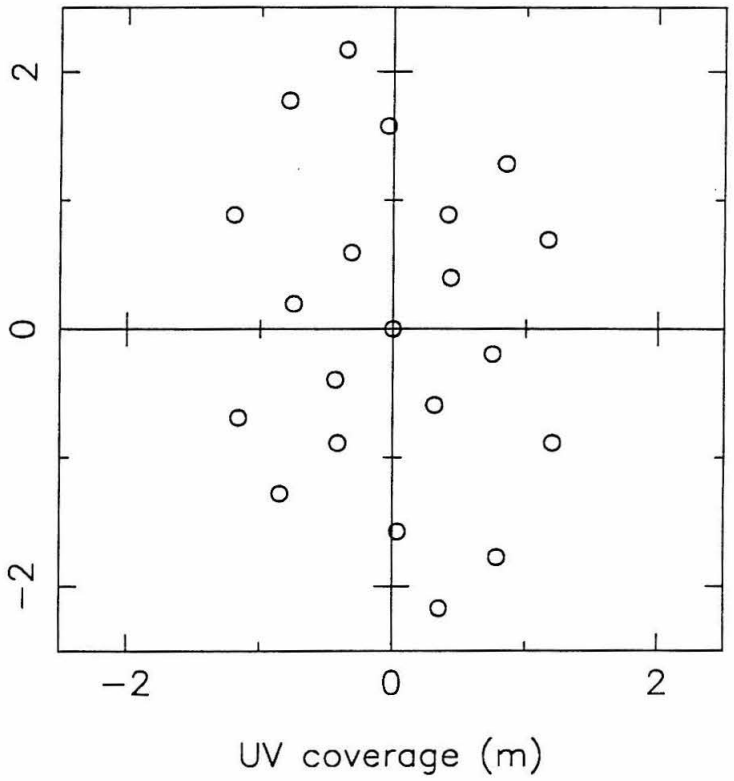

Figure 1

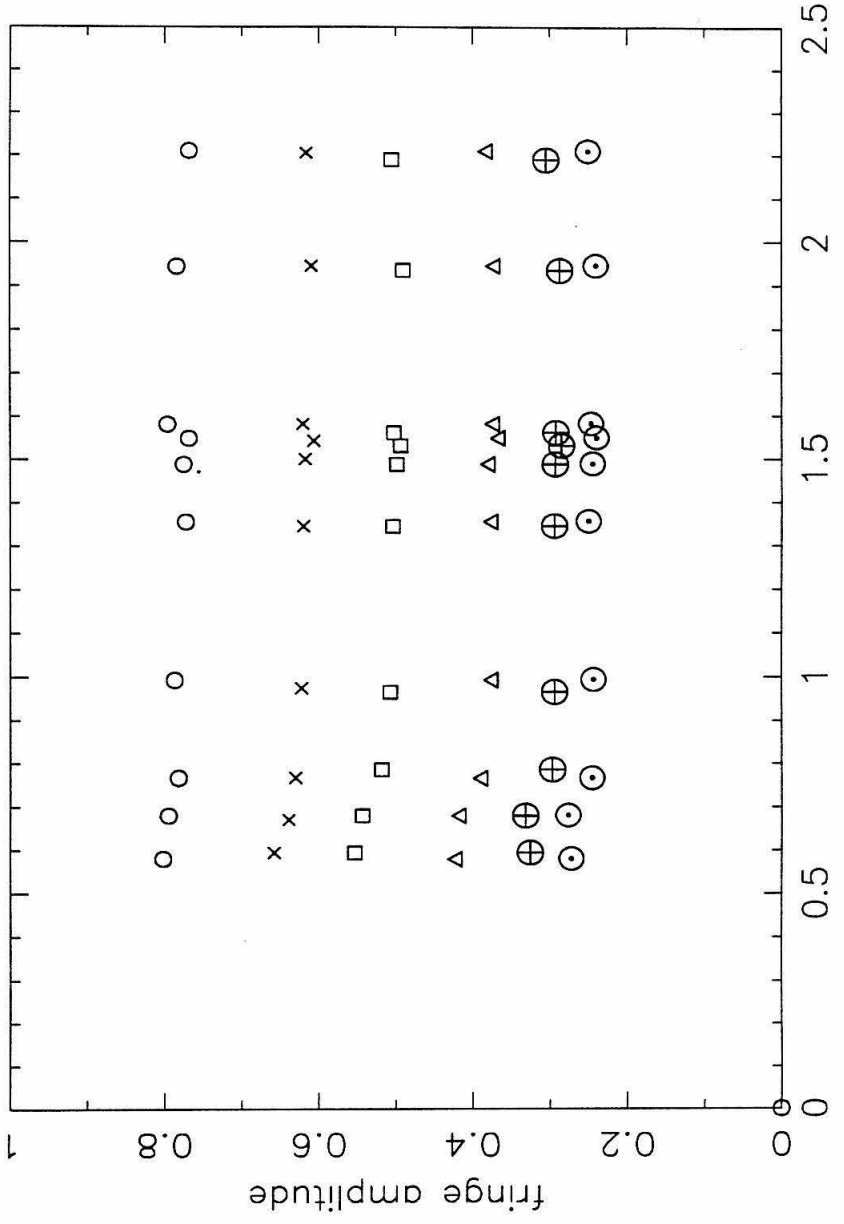

Figure 2

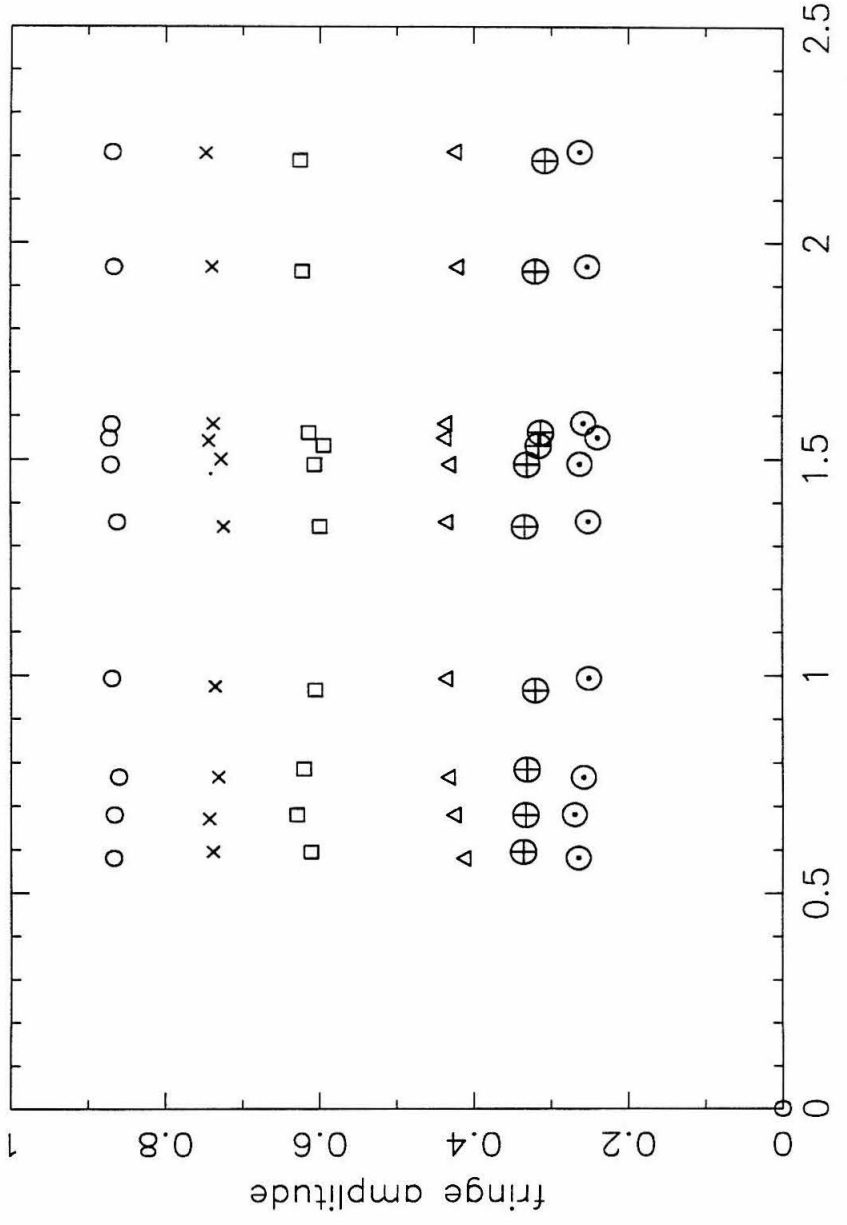

Figure 3

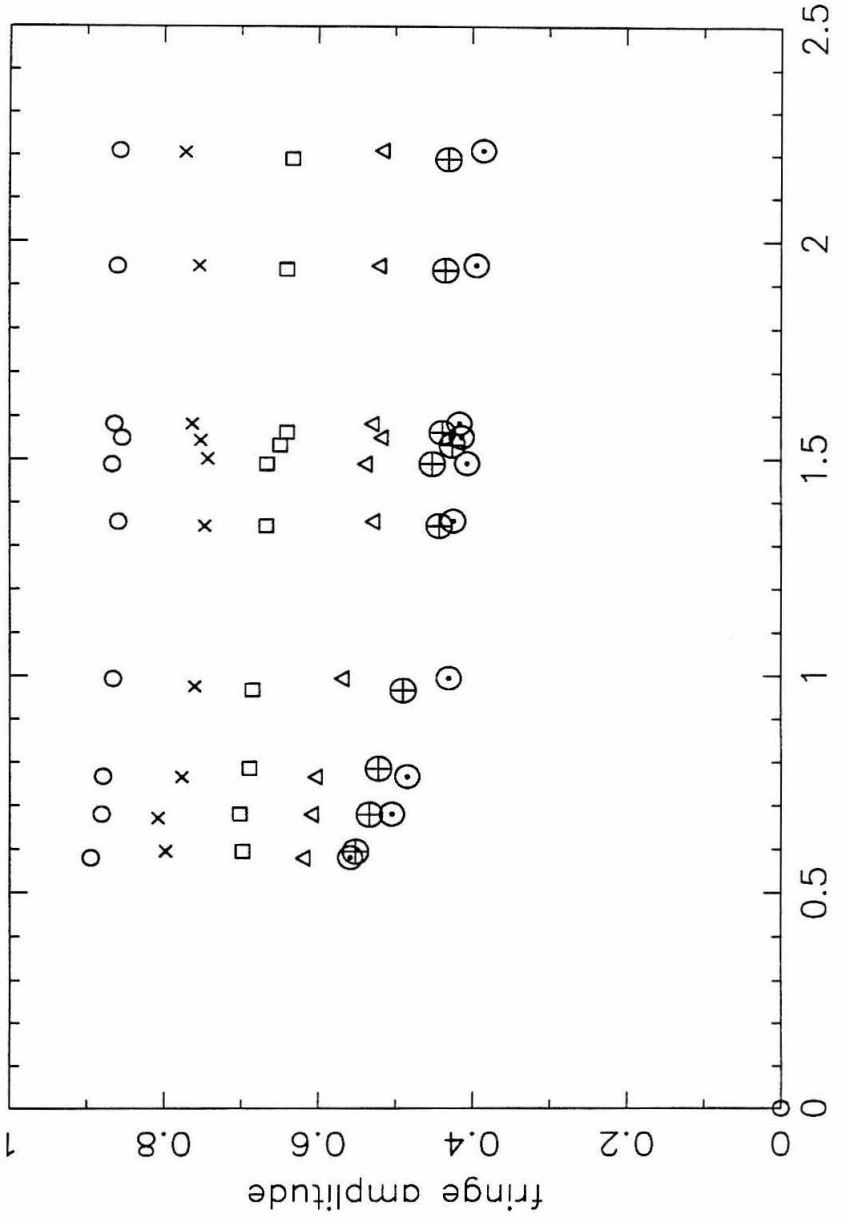

Figure 4

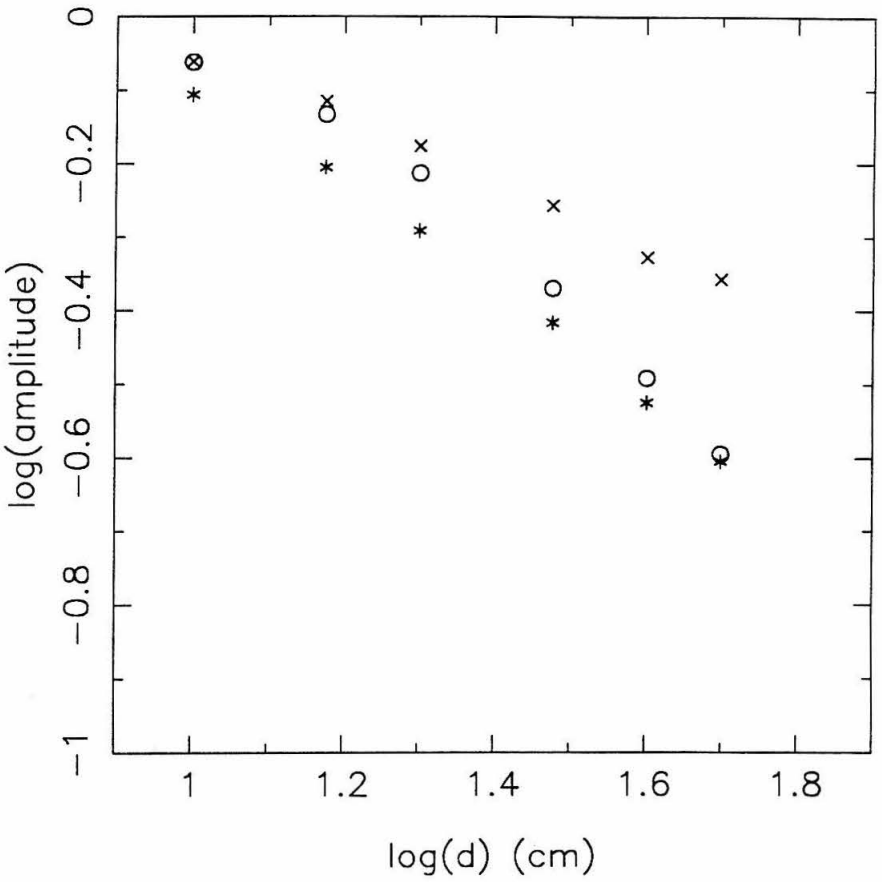

Figure 5

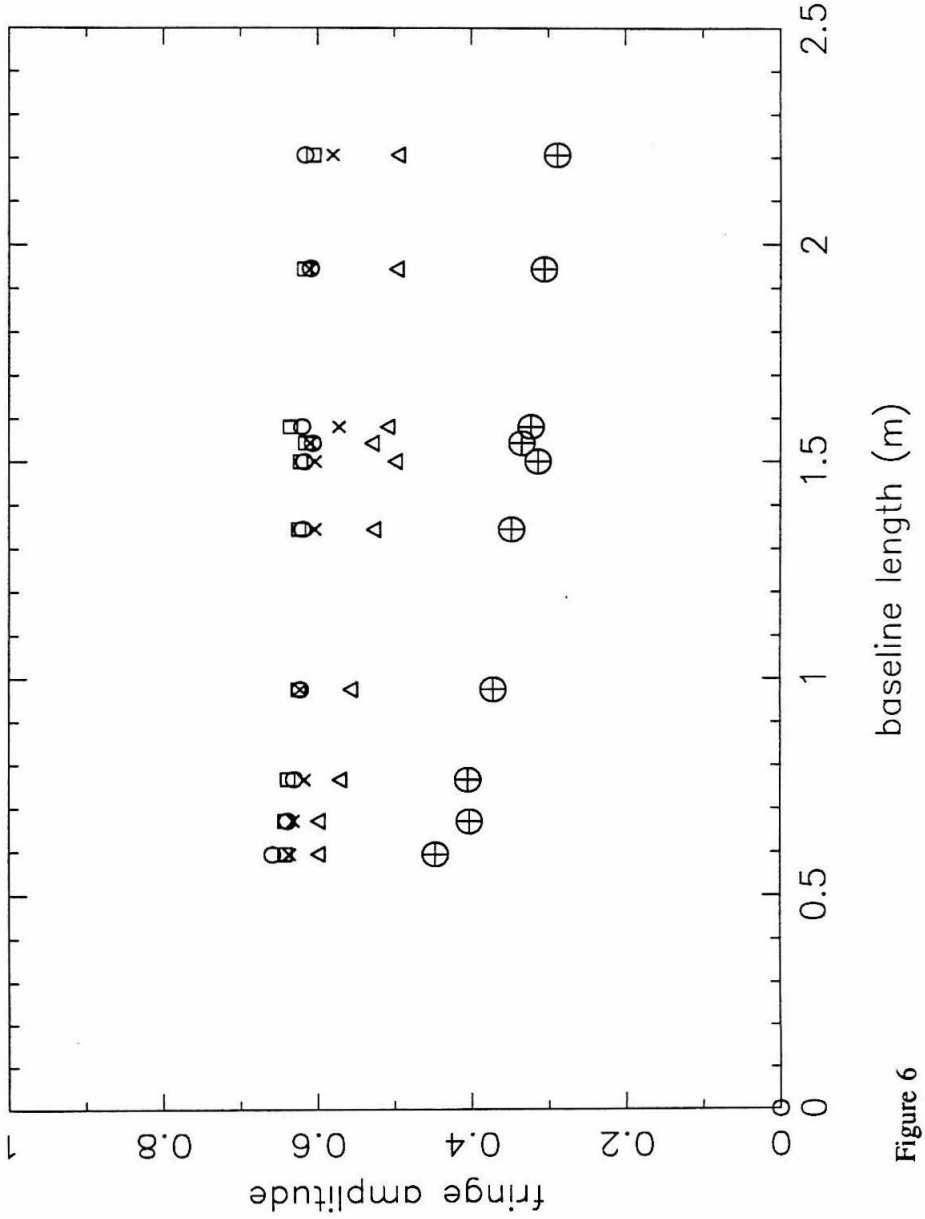

Figure 6

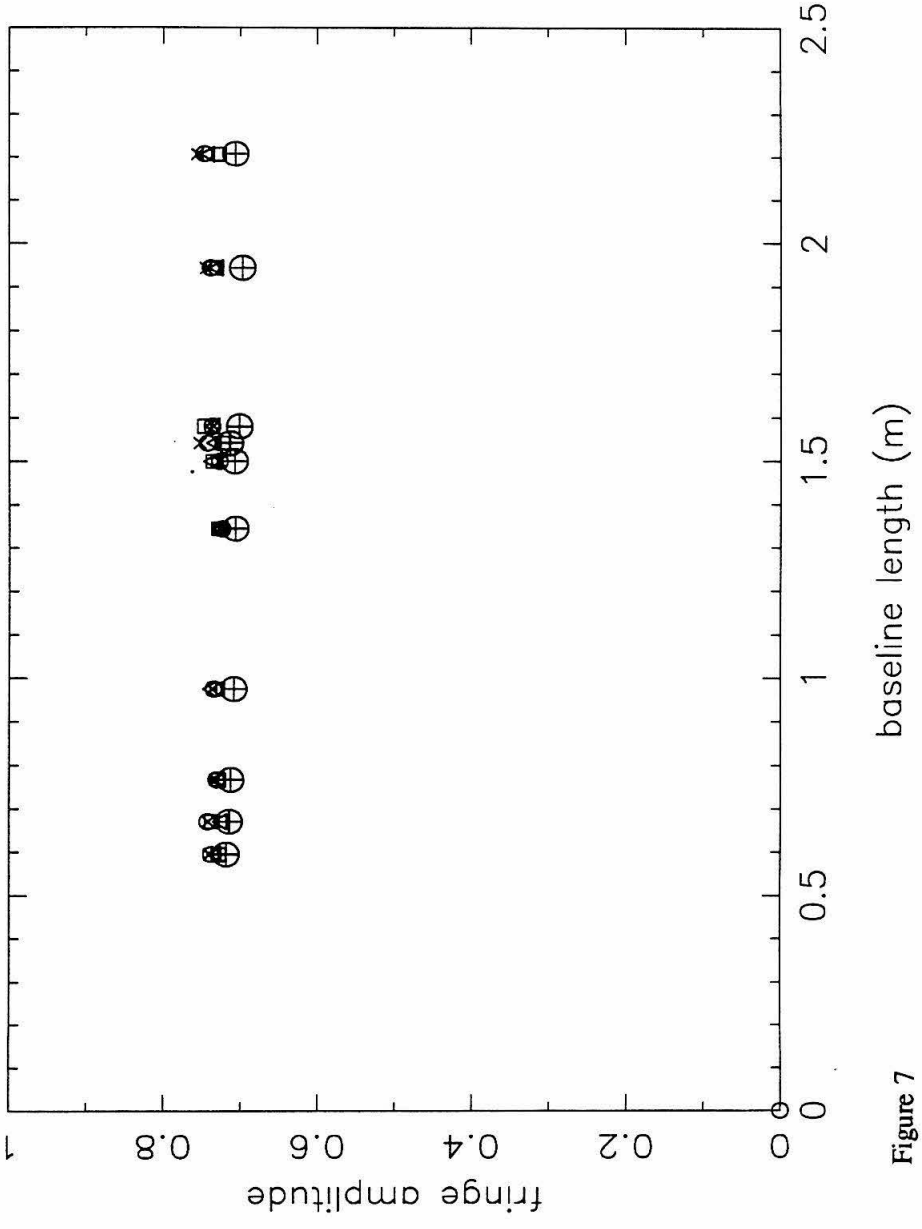

Figure 7

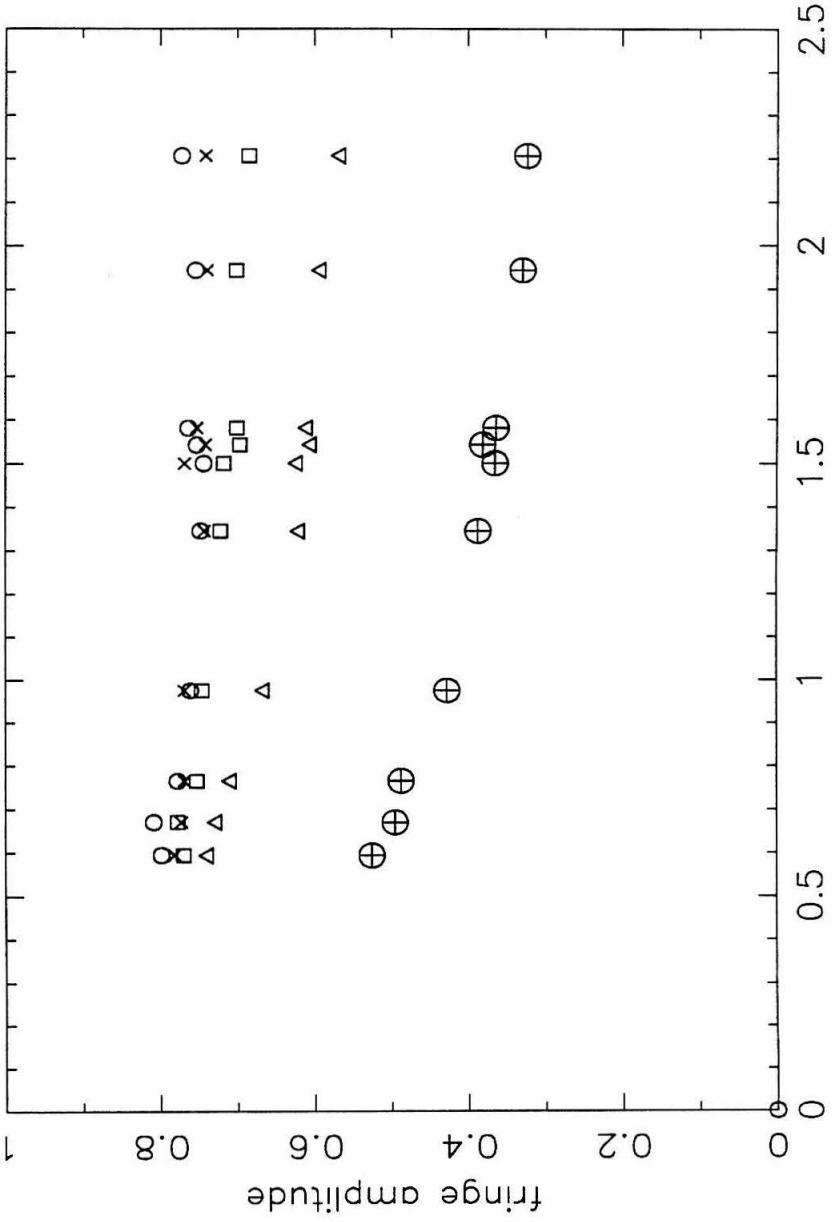

Figure 8

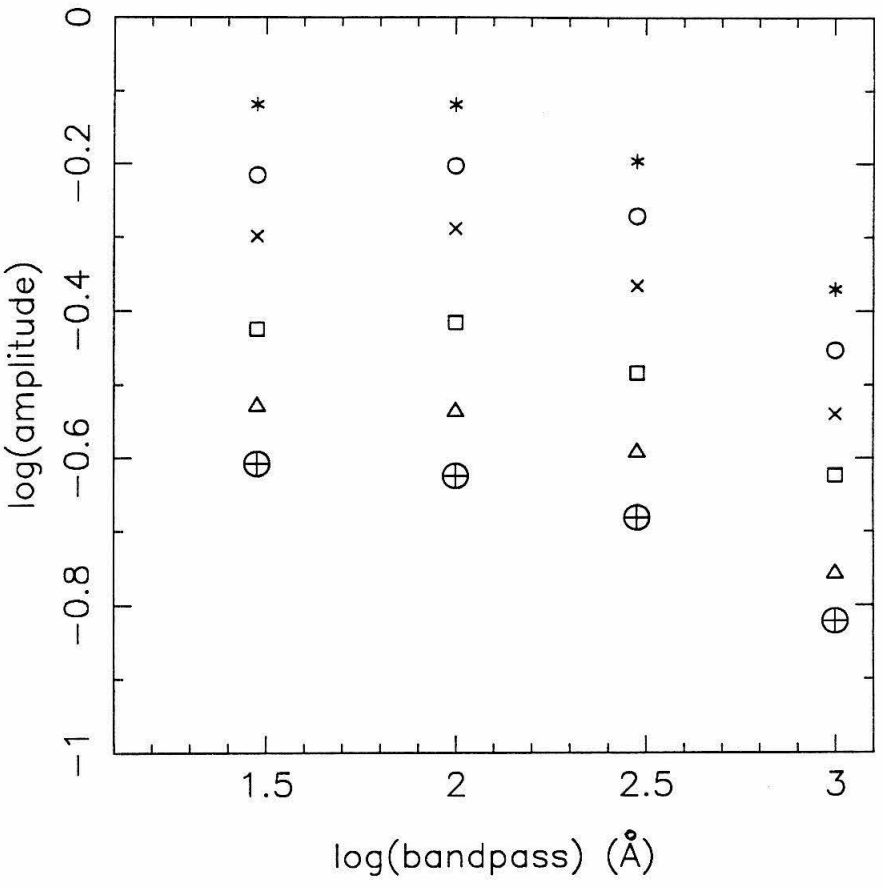

Figure 9

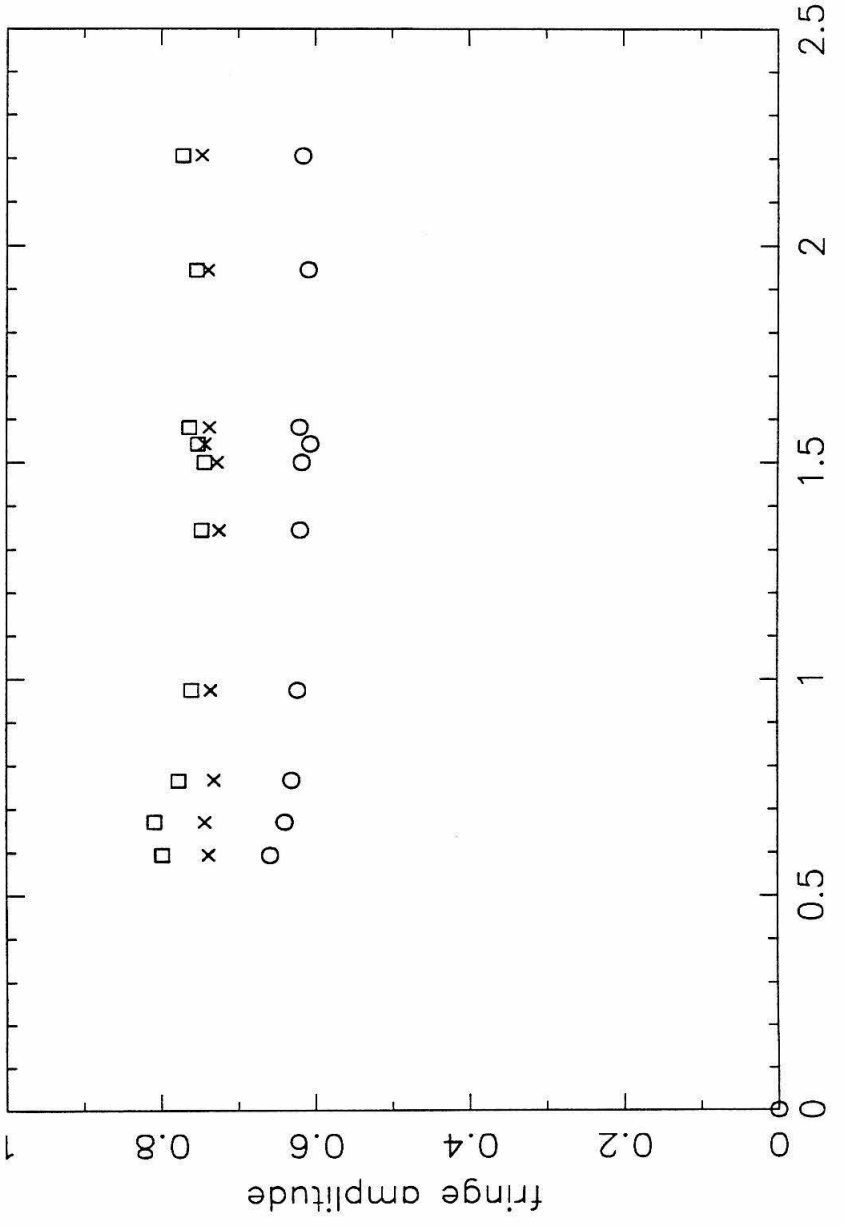

Figure 10

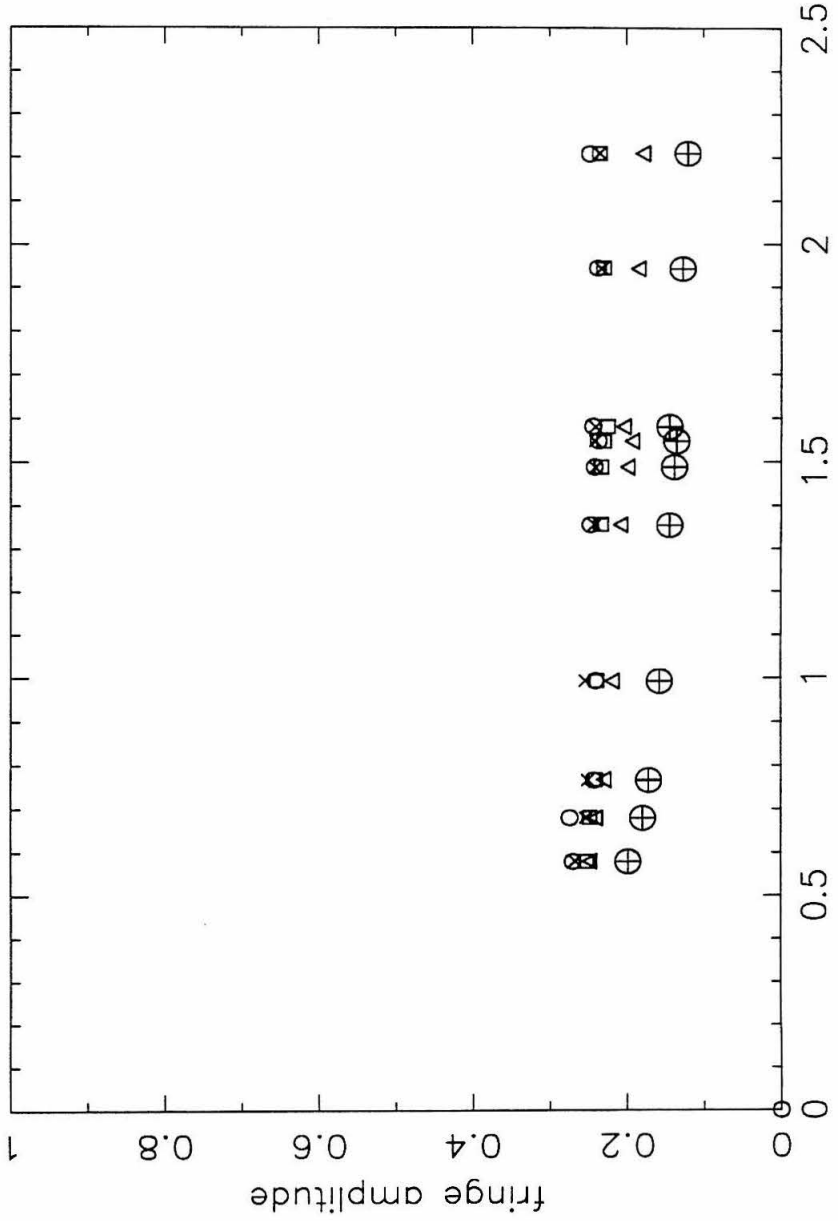

Figure 11

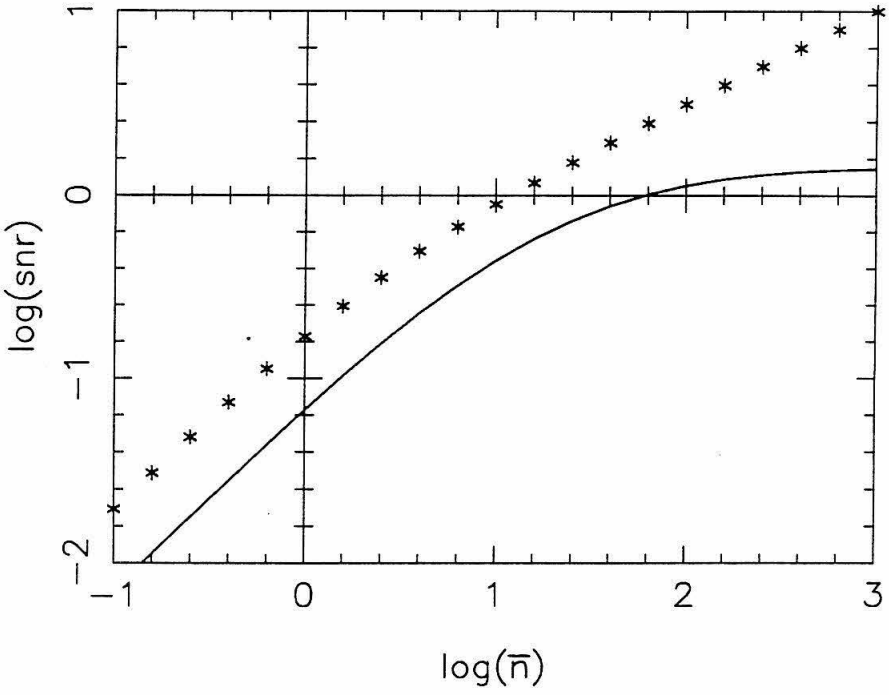

Figure 12

Chapter IV.

Signal-to-Noise Ratio of the Bispectral Analysis of Speckle Interferometry

1. Introduction

Speckle interferometry, (Labeyrie 1970) was first extended to full imaging with phase by Knox and Thompson (1974). A more powerful imaging technique, based on the use of closure phase, was developed in radio astronomy (Readhead and Wilkinson 1978). Independently for optical wavelengths, a method to extract closure-phase information from speckle observations by means of the bispectrum was developed by the Erlangen group (Lohmann et al. 1983, references therein). So far, the method called bispectral analysis has been successful in recovering a 10th-magnitude multiple stellar system (Hoffman and Weigelt 1986). However the potential and the limitations of the method have not yet been investigated fully both in sensitivity and resolution. It is important to quantify the behavior of the signal-to-noise ratio (SNR) of the bispectrum, which depends on both the spatial frequency and the light level.

The analysis of the signal-to-noise ratio of the bispectral analysis is parallel to that of the power spectrum analysis and comprises two stages. First the modulation transfer functions (MTFs) which describe the combined effect of the telescope and the atmospheric disturbance are obtained by treating an incoming light as a wave and a speckle interferogram as an intensity distribution. By taking the influence of photon noise into account, the SNR of the bispectrum at arbitrary light levels is determined as a function of the classical MTFs and the mean photon counts.

In most of the literature of speckle interferometry where the SNR is discussed, MTFs are derived in the mid-spatial-frequency range based on the heuristic interferometric view (HIV) of the image-forming process (Rogers 1963). From this point of view, a speckle pattern is regarded as a random interference pattern produced by a partially coherent incident wave. The validity of this heuristic treatment is known empirically in the case of the power spectrum analysis. The effect of the atmospheric disturbance is included in only one parameter, the coherence length.

The existence of the steep Kolmogorov spectrum in phase fluctuations suggests that there may be some important effects that are not predicted by this simple approach. A more thorough derivation of the power spectrum MTF, based on the phase structure function of the Kolmogorov theory, was derived by Korff (1973), who used a semi-analytical approach that took the atmospheric turbulence properly into account. The derivation of the power spectrum MTF is close to the limit of what can be done analytically. In order to obtain higher-order MTFs, such as the bispectral MTF, that take the Kolmogorov theory into account, it is necessary to resort to Monte Carlo simulations. This method enables us to test the predictions of the HIV of the bispectral analysis (Roddier 1986, Karbelkar and Nityananda 1987, Readhead et al. 1988) and it is shown here that they can be used as a guide to the correct first-order results. Predictions of the HIV are summarized in Appendix 1 of this chapter.

The modeling of photodetection process, based on the rules of conditional statistics, and its application to the power spectrum analysis, were given by Goodman and Belsher (1976, 1977), who formulated an unbiased estimator of the classical power spectrum from an ensemble of photon noise limited images. They also obtained an expression for the signal-to-noise ratio of the power spectrum in terms of the classical modulation transfer function and the mean photon counts. Their analysis is applicable to non-photon-counting detection. In other words, they treated a case in which average photon counts per image were measurable, but neither positions of individual photons nor the total photon counts of individual images were known. Dainty and Greenaway (1979) applied the approach of Goodman and Belsher to photon-counting detection and pointed out that an unbiased estimator of the power spectrum is given in the same manner as in the case of non-photon-counting detection but the expression for the SNR is different. Since the photon noise bias can be removed in each frame, the variance of the power spectrum does not include terms originating from the fluctuations of the bias.

Wirnitzer (1985) gave an unbiased estimator of the classical bispectrum, applying the method of Goodman and Belsher to the bispectral analysis for photon-counting detection. Wirnitzer also obtained the SNR in the high and low light limits, by evaluating the corresponding leading terms in the power of photon counts. This was the first realistic attempt to estimate the limiting magnitude of the bispectral analysis. However, as was first pointed out by Karbelkar and Nityananda, the classical bispectral MTF adopted by Wirnitzer and the SNR in the high-light limit do not agree with those derived from the treatment based on the HIV.

Monte Carlo simulations of an atmospheric phase screen, based on the Kolmogorov spectrum (Kolmogorov 1961, Tatarskii 1961) were made in order to study statistical properties of the bispectral MTF at high light levels. The algorithm and the computation are described in §2 and §3, respectively. The results of the simulations for a 2-m telescope are presented in §4. The results are compared with those obtained using the interferometric view in §5. The MTF obtained by the simulations is re-interpreted to yield the SNR at low light levels in §6. In §7, the discussion of SNR is generalized to arbitrary light levels and arbitrary telescope sizes by modeling of the photo-detection process and the approximate modulation MTFs in the mid-frequency range. Finally in §8, the SNR in the recovered map is considered. An estimate of the practical limiting magnitude is discussed along with the limitation in resolution.

2. Algorithm

The algorithm of creating a phase screen at the aperture plane is identical to that described in Chapter III. The simulations are based on the Kolmogorov theory of turbulence and refractive-index fluctuation (Kolmogorov 1961; Tatarskii 1961, 1967) and recent observations of the altitude dependence of the refractive-index structure constant $C_n{}^2$ (e.g the La *Silla Seeing* Campaign :

Vernin 1986, Roddier and Roddier 1986). The Kolmogorov theory provides mathematical expressions of the atmospheric disturbance on a light-wave propagation as two dimensional spectral densities of phase fluctuation, amplitude fluctuation and the cross spectral density of the two. On the other hand, recent observations indicate that most of the turbulence is produced at the boundary layer and that the high-altitude turbulence contributes a relatively small fraction of the overall seeing degradation. We therefore assume that the major disturbance appears as phase fluctuations in the near-field limit and that the amplitude fluctuation (scintillation) and the cross correlation of the phase and the amplitude fluctuations are negligible to first order.

In the near-field limit, the spectral density of the phase fluctuation at the aperture plane of a telescope is given by

$$F_S(\kappa_r) = 0.033 \cdot 2\pi k^2 \Big[\int C_n^2(L) dL \Big] \kappa_r^{-11/3}, \tag{1}$$

where $\kappa_r = \sqrt{\kappa_x^2 + \kappa_y^2}$ is the two-dimensional radial spatial frequency, k is the wave number, and $\int C_n^2(L)dL$ is the integrated structure constant of the refractive-index fluctuation over the optical path through a turbulent medium.

Since the spectral density constrains the frequency-dependent variance of the fluctuation but not the probability distribution, we further assume a Gaussian probability distribution with zero mean. In the following, $\Phi(\mathbf{x})$ denotes the phase at the aperture plane of a telescope, which is a real function of $\mathbf{x}$ and $\tilde{\Phi}(\kappa)$ denotes its Fourier transform which is a conjugate-symmetric-complex function of κ. At each point over one half of the κ space, $\tilde{\Phi}(\kappa)$, a complex random number whose modulus is a Gaussian random number with a variance of $F_S(\kappa_r)\Delta\kappa$ and whose phase is a uniform random number between 0 and 2π, is generated. $\Delta\kappa$ denotes an area in κ space equalling $(\frac{2\pi}{l_{max}})^2$, where l_{max} is the size of the square phase screen. A conjugate-symmetric-to-real Fourier transform from κ to $\mathbf{x}$ space then creates a monochromatic phase screen $\Phi(\mathbf{x})$ at $\lambda = \frac{2\pi}{k}$.

An idealized telescope is simulated simply by a circular aperture on the phase screen. At each point of the aperture plane, the complex amplitude $\Psi(\mathbf{x}) = \exp(i\Phi(\mathbf{x}))$ is calculated. Another Fourier transform simulates the light-wave propagation from the aperture plane to the image plane and from the squared modulus of the Fourier transform of $\Psi(\mathbf{x})$, a monochromatic speckle pattern, $|\tilde{\Psi}(\mathbf{s})|^2$ is obtained, where $\mathbf{s}$ denotes the coordinate on the image plane. Since the phase fluctuations are simply proportional to the wave number k, the finite-bandwidth effect is taken into account by averaging over monochromatic speckle patterns at equally spaced wave numbers covering the bandpass. Thus, from one evaluation of $\Psi(\mathbf{x})$, multiple $|\tilde{\Psi}(\mathbf{s})|^2$s are generated and averaged to produce one speckle interferogram $I(\mathbf{s})$. If Taylor's hypothesis of frozen-in turbulence (Taylor 1961) and a uniform translation by a constant wind velocity are assumed, a continuous observation can be simulated by considering a series of apertures displaced by a distance which is typically the coherence length, r_c.

In the data reduction, many short-exposure frames are processed to derived statistics. The Fourier transform $\tilde{I}_j(\mathbf{u})$ of the speckle interferogram of the j th frame $I_j(\mathbf{s})$ is taken to form the bispectrum

$$\tilde{B}_j^{(3)}(\mathbf{u_1}, \mathbf{u_2}) = \tilde{I}_j(\mathbf{u_1})\tilde{I}_j(\mathbf{u_2})\tilde{I}_j(-\mathbf{u_1} - \mathbf{u_2}), \tag{2}$$

where $\mathbf{u}$ denotes the spatial frequency on the image plane. For n frames, both the sum of bispectra

$$\sum_{j=1}^{n} \tilde{B}_j^{(3)}(\mathbf{u_1}, \mathbf{u_2}) \tag{3}$$

and the sum of square moduli

$$\sum_{j=1}^{n} |\tilde{B}_j^{(3)}(\mathbf{u_1}, \mathbf{u_2})|^2, \tag{4}$$

are calculated. The unbiased estimator of the ensemble average bispectrum is

$$\langle \tilde{B}^{(3)}(\mathbf{u_1}, \mathbf{u_2}) \rangle = \frac{\sum_{j=1}^{n} \tilde{B}_j^{(3)}(\mathbf{u_1}, \mathbf{u_2})}{n}, \tag{5}$$

where $\langle\rangle$ indicates an ensemble average. Likewise the unbiased estimator of the variance of the bispectrum per frame is

$$\sigma^2(\tilde{B}^{(3)}(\mathbf{u_1},\mathbf{u_2})) = \frac{\sum_{j=1}^{n} |\tilde{B}_j^{(3)}(\mathbf{u_1},\mathbf{u_2})|^2 - n|\langle\tilde{B}^{(3)}(\mathbf{u_1},\mathbf{u_2})\rangle|^2}{n-1}. \tag{6}$$

In the case of the bispectral MTF, the mean value is real (Lohmann et al. 1983), since the atmospheric disturbance is statistically isotropic and the ideal telescope is static and symmetric. After averaging over enough samples, the SNR of the MTF per frame is defined as

$$SNR(\tilde{B}^{(3)}(\mathbf{u_1},\mathbf{u_2})) = \frac{Real(\langle\tilde{B}^{(3)}(\mathbf{u_1},\mathbf{u_2})\rangle)}{\sqrt{\sigma^2(\tilde{B}^{(3)}(\mathbf{u_1},\mathbf{u_2}))}}. \tag{7}$$

Henceforth a SNR means a SNR per frame, unless specified otherwise. The SNR in the recovered map will be discussed in §8.

3. Computation

The simulations were made at a wavelength λ = 0.55 μm with a fractional bandwidth of 0.1. At this wavelength, the integrated structure constant of the refractive-index fluctuation,

$$\int C_n^2(L)dL = 5 \times 10^{-13}\ \mathrm{m}^{\frac{1}{3}}, \tag{8}$$

was adopted corresponding to 1-arcsec seeing. This is approximately Roddier's value (Roddier 1981). It was found experimentally that five monochromatic speckle interferograms produced at equally spaced wave numbers within the bandpass were enough to obtain a reasonable averaged speckle interferogram.

In order to include wave front degradation at small scales, the sampling interval on the phase screen must be significantly smaller than the coherence length r_c. On the other hand, the

linear size of the phase screen (l_{max}) has to be significantly larger than the primary mirror of the telescope so that the large-scale disturbance is simulated properly. Both the sampling interval and the ratio between the size of the phase screen and the diameter of the primary mirror $\frac{l_{max}}{D}$ were determined empirically. The sampling interval Δl was chosen as 2 cm, so that 49 phase data were obtained within a square area of ${r_c}^2$ for r_c = 14 cm. It was found that fluctuations with correlation scales larger than 4 times the telescope diameter mainly caused image wandering, but did not affect the power spectrum or the bispectrum of a speckle interferogram. The ratio $\frac{l_{max}}{D}$ must be at least 4.

The maximum practical array size on the computer used, a VAX/750, is 512^2 when the memory access time and the CPU-time for the Fast Fourier Transform(FFT) are taken into account. When the array size is combined with the sampling interval Δl, the size of phase screen is $(512\Delta l)^2 \approx$ 10 m $\times$ 10 m. Since the size of the phase screen is still smaller than the typical outer scale of the atmospheric turbulence, the Kolmogorov theory can safely be assumed. The maximum primary mirror size of the resultant simulated telescope is about 2.5 m or 128 pixels in diameter. In practice, we concentrated on a 2-m telescope, but also made calculations for a 1-m telescope for comparison. In producing images, 256^2-sized fast Fourier transforms were made to satisfy the Nyquist sampling requirement. A cross section of a phase screen is shown in Fig. 1.

Unfortunately the entire four-dimensional bispectrum is beyond the capacity of the computer used. It is however, possible to get a good insight into the MTF from a two-dimensional cross section of the bispectrum, since the atmospheric disturbance was already assumed to be locally isotropic in the Kolmogorov theory. The cross section of the bispectral MTF was chosen so that $\mathbf{u_1}$ and $\mathbf{u_2}$ are perpendicular to each other because of the convenience in drawing two-dimensional contour maps (Fig. 2). Therefore in a four-dimensional expression, the cross section is

$$\tilde{B}^{(3)}(\zeta, 0, 0, \eta) = \tilde{I}(\zeta, 0)\tilde{I}(0, \eta)\tilde{I}(-\zeta, -\eta). \tag{9}$$

It should be noted that the third spatial frequency $(-\mathbf{u_1} - \mathbf{u_2})$ has the largest modulus $(\sqrt{\zeta^2 + \eta^2})$ among the three and the circle,

$$\zeta^2 + \eta^2 = (\frac{D}{\lambda})^2 \tag{10}$$

forms the boundary of the bispectrum, where $\frac{D}{\lambda}$ is the telescope cutoff frequency.

4. Results of the Simulations

The results of the computations are the normalized bispectral MTFs defined as

$$\langle b^{(3)}(\mathbf{u_1}, \mathbf{u_2})\rangle = \frac{\langle B^{(3)}(\mathbf{u_1}, \mathbf{u_2})\rangle}{\langle B^{(3)}(\mathbf{0}, \mathbf{0})\rangle}. \tag{11}$$

Since fluctuations in the total intensity $\tilde{I}(\mathbf{0}) = \int I(\mathbf{x})d\mathbf{x}$ are not considered,

$$\langle B^{(3)}(\mathbf{0}, \mathbf{0})\rangle = \tilde{I}(\mathbf{0})^3 = const., \tag{12}$$

and then

$$\langle b^{(3)}(\mathbf{u_1}, \mathbf{u_2})\rangle = \langle \tilde{i}(\mathbf{u_1})\tilde{i}(\mathbf{u_2})\tilde{i}(-\mathbf{u_1} - \mathbf{u_2})\rangle, \tag{13}$$

where $\tilde{i}(\mathbf{u}) = \frac{\tilde{I}(\mathbf{u})}{\tilde{I}(\mathbf{0})}$. For the same reason, the SNR in the normalized bispectral MTF is

$$SNR(b^{(3)}(\mathbf{u_1}, \mathbf{u_2})) = SNR(B^{(3)}(\mathbf{u_1}, \mathbf{u_2})). \tag{14}$$

The ζ and η axis on the ζ-η plane correspond to the normalized power spectrum MTF:

$$\begin{aligned}\langle b^{(3)}(\zeta, 0, 0, 0)\rangle &= \langle \tilde{i}(\zeta, 0)\tilde{i}(0, 0)\tilde{i}(-\zeta, 0)\rangle \\ &= \langle \tilde{i}(\zeta, 0)1\tilde{i}(\zeta, 0)^*\rangle \\ &= \langle |\tilde{i}(\zeta, 0)|^2\rangle,\end{aligned} \tag{15}$$

and similarly,

$$\langle b^{(3)}(0, 0, 0, \eta)\rangle = \langle |\tilde{i}(0, \eta)|^2\rangle. \tag{16}$$

The normalized MTF of a 2-m telescope is shown in Fig. 3. The statistics are derived from 500 interferograms sampled from independent portions of 50 different phase screens to assure the statistical independence of instantaneous bispectra. One pixel in spatial frequency corresponds to $0.173(\text{arcsec})^{-1} = \frac{1.95cm}{0.55\mu} = 0.01\frac{D}{\lambda}$ and the 102nd pixel corresponds to the telescope cutoff frequency, $\frac{D}{\lambda}$.

In understanding the physics of the results, it is convenient to introduce a concept of the attenuation of the bispectral MTF, since the atmospheric disturbance is regarded as a low-pass filter of the spatial-frequency information. The attenuation factor (ATF) of the bispectral MTF, $\tilde{a}^{(3)}(\zeta,0,0,\eta)$ is defined along with the the optical transfer function (OTF) of the telescope, $\tilde{t}^{(3)}(\zeta,0,0,\eta)$, as

$$\langle\tilde{a}^{(3)}(\zeta,0,0,\eta)\rangle = \frac{\langle\tilde{b}^{(3)}(\zeta,0,0,\eta)\rangle}{\tilde{t}^{(3)}(\zeta,0,0,\eta)}, \tag{17}$$

where $\tilde{t}^{(3)}(\zeta,0,0,\eta)$ is static and real. The OTF is the normalized bispectrum of the Airy pattern and thus is the normalized MTF under the coherent illumination. It could also be interpreted as the relative weight of the frequency components or the relative redundancy of the triangular baselines (both closed and nonclosed) on the primary mirror. To avoid confusion, it should be noted that the attenuation is the combined effect of the atmosphere and the optics and thus that the ATF depends on the OTF even for a given atmospheric condition. The OTF and the ATF are shown in Figs. 4 and 5, respectively. The OTF is a monotonically decreasing function of ζ and η. It is fairly flat at $\sqrt{\zeta^2+\eta^2} \leq 0.5\frac{D}{\lambda}$ and then falls more steeply at higher frequencies.

The ATF, and thus the MTF, behaves in a more complicated manner. The SNR of the classical MTF or the saturated SNR at high light levels is shown in Fig. 6. The behavior of the SNR is quite similar to that of the ATF. Semi-quantitatively, the contour maps of the ATF and SNR can be classified into five distinct regions in spatial frequency as follows.

i) Low-Frequency Region ($\sqrt{\zeta^2+\eta^2} \leq 0.1\frac{D}{\lambda}$)

The ATF is larger than 0.01 and the SNR is larger than unity. The information of this region originates from the envelopes of instantaneous interferograms. Even this low-frequency region has better information than a seeing disk obtained by a long time exposure, since the effect of image wandering is removed.

ii) On-Axis Region ($\zeta = 0$, $\eta = 0$)

The power spectrum ATF is plotted as a function of radial frequency in Fig. 7, since it is statistically isotropic. The ATF falls off steeply at low frequencies ($\leq 0.1\frac{D}{\lambda}$) and then levels between $0.1\frac{D}{\lambda}$ and $0.8\frac{D}{\lambda}$ at around 4×10^{-3}. Above $0.8\frac{D}{\lambda}$, the ATF slowly rises up to 10^{-1} at $\frac{D}{\lambda}$. This increase of the ATF at high spatial frequency is not fast enough to compensate the steep fall of the OTF and the power spectrum MTF monotonically falls off as is shown on the axes in Fig. 3. The SNR is larger than 0.8 up to $\frac{D}{\lambda}$.

iii) Near-Axis Region ($\zeta \leq 0.1\frac{D}{\lambda}$, or $\eta \leq 0.1\frac{D}{\lambda}$)

The ATF falls off steeply to 10^{-4} and the SNR decreases to 0.3 as the plot moves vertically away from each axis. The bispectral components are the combination of the low-frequency Fourier components and the power spectrum. Phases of the bispectral components in this region are effectively local phase differences of nearby Fourier components which are used in the Knox-Thompson method (Knox and Thompson 1974).

iv) Mid-Frequency Region ($\zeta, \eta \geq 0.1\frac{D}{\lambda}$ and $\sqrt{\zeta^2 + \eta^2} \leq 0.8\frac{D}{\lambda}$)

A large triangular plateau of the ATF with a mean value of 3×10^{-5} is evident in Fig. 5. In this region, the behavior of the MTF is determined mainly by that of the OTF. The MTF falls from 10^{-4} to 10^{-6}. The SNR is between 0.1 and 0.2.

v) High frequency region ($\zeta^2 + \eta^2 \approx (\frac{D}{\lambda})^2$)

The diffraction limited information lies in this region. Because of the steep fall of the OTF, the MTF is very small ($\leq 10^{-7}$). The SNR is smaller than 0.1.

5. Comparison with the Heuristic Interferometric View

It is interesting to compare the above results with the predictions by the heuristic interferometric view (HIV) of the image-forming process. A brief description of this view is given in Appendix 1, and a detailed discussion is found in Readhead et al. (1988). The HIV predicts that the power spectrum ATF is approximated by $n_s^{-1} = (\frac{r_c}{D})^2$ where n_s is the number of speckles, and that the SNR of the power spectrum is unity at the 'mid frequency range'. It also estimates that the bispectral ATF in the mid-frequency range is about $n_s^{-2} = (\frac{r_c}{D})^4$ and the SNR is given by $n_s^{-\frac{1}{2}} = \frac{r_c}{D}$. The bispectral ATF and its SNR are therefore related by

$$\langle \tilde{a}^{(3)}(\mathbf{u_1}, \mathbf{u_2}) \rangle \approx \{ SNR(\tilde{b}^{(3)}(\mathbf{u_1}, \mathbf{u_2})) \}^4. \tag{18}$$

Thus the similarity of the contour maps of the ATF and the SNR, which are plotted in logarithmic scale in Figs. 5 and 6, is naturally explained by the HIV. For r_c = 14 cm and D = 2 m, n_s = 204. The flat portion of the simulated power spectrum ATF between $0.2\frac{D}{\lambda}$ and $0.8\frac{D}{\lambda}$ is 4×10^{-3} on average, while the value predicted by the HIV is 5×10^{-3}. The SNR of the power spectrum lies between 0.6 and 0.8 and is approximately unity. At the mid-frequency region, the simulated bispectral ATF has an average value of 3×10^{-5}, while the value predicted by the HIV is 2×10^{-5}. We consider this agreement good.

The simulations were also made for a 1-m telescope. For a telescope of this size, the mid-frequency of the power spectrum ranges from $0.3\frac{D}{\lambda}$ to $0.6\frac{D}{\lambda}$ and the simulated power-spectrum ATF is 3×10^{-2} while the predicted value is 2×10^{-2}. The SNR of the power spectrum is between 0.7 and 1.0. The simulated bispectral ATF at the mid frequency is 2×10^{-3} in average,

which is somewhat larger than the predicted value, 4×10^{-4}. The agreement is not as good as for a 2-m telescope.

The rise of the power spectrum ATF at the high-frequency region ($\geq 0.8\frac{D}{\lambda}$ for $D = 2$ m) can qualitatively be interpreted by the HIV. At the high-frequency region, the redundancy (or OTF) of the baselines is so small and the identical baselines are so localized on the primary mirror that the phasors of those baselines are correlated and increase the ATF. The approximate validity of the heuristic interferometric view of the image-forming process is confirmed by the simulations. The simulations also clarified the boundaries of the mid frequency for given apertures. Because of the higher redundancy and the wider mid-frequency range, the predictions by the HIV work better for larger telescopes.

6. Signal-to-Noise Ratio at Low Light Levels

At low light levels, the signal-to-noise ratio per frame of an unbiased estimator of the classical bispectral modulation transfer function is approximated by

$$\langle \tilde{b}^{(3)}(\mathbf{u_1}, \mathbf{u_2}) \rangle \overline{N}^{\frac{3}{2}}, \tag{19}$$

where $\overline{N}$ is the average photon counts per frame (Wirnitzer 1985). Therefore the contour map of $\langle \tilde{b}^{(3)}(\mathbf{u_1}, \mathbf{u_2}) \rangle$ can be immediately converted to that of the SNR.

Fig. 8 shows the signal-to-noise ratio of a V = 12.3 mag star with a 2-m telescope after integrating 10^4 frames; this magnitude corresponds to one photon per speckle in a 10% fractional bandwidth, with 10% efficiency of the observing system and 10 ms integration time. The SNR = 3 contour reaches the diffraction limit on the axes, but stays near the axis as ζ or η increases. The slope of contours is the steepest diagonally. Fig. 9 shows SNR = 3 contours according to the

brightness of sources. At 9.0 mag, the bispectral analysis is diffraction limited in the sense of a three-sigma detection, whereas at 13.9 mag, even the power-spectrum analysis is not necessarily diffraction limited, and the region of high SNR is strictly near the axes.

The contour maps immediately show that the power spectrum in general has better SNR than the bispectrum. The closure-phase information obtained by near-axis bispectral components is effectively local phase differences of neighboring Fourier components. For a simple source such as a multiple stellar system, the autocorrelation function contains most of the source structure. The behavior of the phase in Fourier space is fairly regular and thus the local phase differences are enough to recover a full image. Hofmann and Weigelt (1986) used only the 5% of the bispectrum near the axes with the highest SNR for their image recovery. The result of the simulations is consistent with their observations. The wide mid-frequency range contains global closure-phase information with lower quality. For a complicated source, mid-frequency components may be crucial in recovering a full image. Intensive computations are required for utilization of the full bispectrum.

7. Signal-to-Noise Ratio at Arbitrary Light Levels

In the simulations, the incoming light has been treated as a wave. Thus from the point of view of photon detection, a limiting case with an infinite number of photons is considered. In this section, by using the modeling of the photodetection process by Goodman and Belsher (1976, 1977; Goodman 1985) and following the treatment of the influence of photon noise on the bispectral analysis by Wirnitzer (1985), the derivation of an unbiased estimator of the classical bispectrum is reviewed and then an expression for the SNR of the bispectral MTF is obtained as a function of the mean photon count, OTF's of telescopes and the number of speckles.

We consider a speckle observation using a photon-counting detector which records positions of individual photons detected on the image plane. The raw intensity of the jth frame is given as

$$D_j(\mathbf{x}) = \sum_{k=1}^{N_j} \delta(\mathbf{x} - \mathbf{x}_k), \tag{20}$$

where x_k is the position of the kth photon and N_j is the total number of photons. The Fourier transform of Eq. (20) is

$$\begin{aligned}\tilde{D}_j(\mathbf{u}) &= \int \sum_{k=1}^{N_j} \delta(\mathbf{x} - \mathbf{x}_k)\exp(i\mathbf{u}\mathbf{x})d\mathbf{x} \\ &= \sum_{k=1}^{N_j} \exp(i\mathbf{u}\mathbf{x}_k).\end{aligned} \tag{21}$$

The bispectrum of the raw data is given as

$$\begin{aligned}\tilde{D}_j^{(3)}(\mathbf{u_1}, \mathbf{u_2}) &= \tilde{D}_j(\mathbf{u_1})\tilde{D}_j(\mathbf{u_2})\tilde{D}_j(-\mathbf{u_1} - \mathbf{u_2}) \\ &= \sum_{k=1}^{N_j}\sum_{l=1}^{N_j}\sum_{m=1}^{N_j} \exp(i\{\mathbf{u_1}(\mathbf{x}_k - \mathbf{x}_m) + \mathbf{u_2}(\mathbf{x}_l - \mathbf{x}_m)\}).\end{aligned} \tag{22}$$

The expected value of $\tilde{D}_j^{(3)}(\mathbf{u_1}, \mathbf{u_2})$ is evaluated over the conditional statistics of $\mathbf{x}_k$s, N_j and the rate function $\lambda_j(\mathbf{x})$, which is proportional to the classical intensity, $I_j(\mathbf{x})$. For a given N_j and $\lambda_j(\mathbf{x})$, the event locations $\mathbf{x}$ are independent random variables, with common probability-density function,

$$p_j(\mathbf{x}) = \frac{\lambda_j(\mathbf{x})}{\int \lambda_j(\mathbf{x})d\mathbf{x}} = \frac{I_j(\mathbf{x})}{\int I_j(\mathbf{x})d\mathbf{x}}. \tag{23}$$

The characteristic function of $p_j(\mathbf{x})$ equals the normalized Fourier transform of the classical intensity distribution $\tilde{i}_j(\mathbf{u})$:

$$\begin{aligned}\tilde{p}_j(\mathbf{u}) &= \int p_j(\mathbf{x})\exp(i\mathbf{u}\mathbf{x})d\mathbf{x} \\ &= \frac{\int I_j(\mathbf{x})\exp(i\mathbf{u}\mathbf{x})d\mathbf{x}}{\int I_j(\mathbf{x})d\mathbf{x}} \\ &= \frac{\tilde{I}_j(\mathbf{u})}{\tilde{I}_j(\mathbf{0})} \\ &= \tilde{i}_j(\mathbf{u}).\end{aligned} \tag{24}$$

First $\tilde{D}_j^{(3)}(\mathbf{u_1}, \mathbf{u_2})$ is averaged over the conditional statistics of $\mathbf{x}_k$, $\mathbf{x}_l$ and $\mathbf{x}_m$ and then averaged over N_j and $\lambda_j(\mathbf{x})$. The starting point is the evaluation of

$$E_{klm}[\tilde{D}_j^{(3)}(\mathbf{u_1}, \mathbf{u_2})] = E_{klm}[\sum_{k=1}^{N_j}\sum_{l=1}^{N_j}\sum_{m=1}^{N_j} \exp(i\{\mathbf{u_1}(\mathbf{x}_k - \mathbf{x}_m) + \mathbf{u_2}(\mathbf{x}_l - \mathbf{x}_m)\})]$$

$$= \sum_{k=1}^{N_j}\sum_{l=1}^{N_j}\sum_{m=1}^{N_j} E_{klm}[\exp(i\{\mathbf{u_1}(\mathbf{x}_k - \mathbf{x}_m) + \mathbf{u_2}(\mathbf{x}_l - \mathbf{x}_m)\})], \tag{25}$$

where E_{klm} stands for an average over $\mathbf{x}_k$, $\mathbf{x}_l$ and $\mathbf{x}_m$. The ${N_j}^3$ terms are classified as follows:

(1) $k = l = m$ $\quad$ N_j terms

$$E_{klm}[1] = \int p_j(\mathbf{x}_k)d\mathbf{x}_k = 1 \tag{26}$$

(2) $k \neq l = m$ $\quad$ $N_j(N_j - 1)$ terms

$$E_{klm}[\exp(i\mathbf{u_1}(\mathbf{x}_k - \mathbf{x}_l))] = \int\int \exp(i\mathbf{u_1}(\mathbf{x}_k - \mathbf{x}_l))p_j(\mathbf{x}_k)p_j(\mathbf{x}_l)d\mathbf{x}_k d\mathbf{x}_l$$

$$= (\int p_j(\mathbf{x}_k)\exp(i\mathbf{u_1}\mathbf{x}_k)d\mathbf{x}_k)(\int p_j(\mathbf{x}_l)\exp(-i\mathbf{u_1}\mathbf{x}_l)d\mathbf{x}_l)$$

$$= \tilde{i}_j(\mathbf{u_1})\tilde{i}_j(-\mathbf{u_1})$$

$$= |\tilde{i}_j(\mathbf{u_1})|^2, \tag{27}$$

where $|\tilde{i}_j(\mathbf{u_1})|^2$ is the normalized power spectrum.

(3) $k = m \neq l$ $\quad$ $N_j(N_j - 1)$ terms

$$E_{klm}[\exp(i\mathbf{u_1}(\mathbf{x}_l - \mathbf{x}_m))] = |\tilde{i}_j(\mathbf{u_2})|^2 \tag{28}$$

(4) $k = l \neq m$ $\quad$ $N_j(N_j - 1)$ terms

$$E_{klm}[\exp(i(-\mathbf{u_1} - \mathbf{u_2})(\mathbf{x}_l - \mathbf{x}_m))] = |\tilde{i}_j(-\mathbf{u_1} - \mathbf{u_2})|^2 \tag{29}$$

(5) $k \neq l \neq m$ $N_j(N_j-1)(N_j-2)$ terms

$$
\begin{aligned}
&E_{klm}[\exp(i\{\mathbf{u_1}(\mathbf{x}_k-\mathbf{x}_m)+\mathbf{u_2}(\mathbf{x}_l-\mathbf{x}_m)\})] \\
&=\Big(\int p_j(\mathbf{x}_k)\exp(i\mathbf{u_1}\mathbf{x}_k)d\mathbf{x}_k\Big)\Big(\int p_j(\mathbf{x}_l)\exp(i\mathbf{u_2}\mathbf{x}_l)d\mathbf{x}_l\Big)\Big(\int p_j(\mathbf{x}_m)\exp(i(-\mathbf{u_1}-\mathbf{u_2})\mathbf{x}_m)d\mathbf{x}_m\Big) \\
&=\tilde{i}_j(\mathbf{u_1})\tilde{i}_j(\mathbf{u_2})\tilde{i}_j(-\mathbf{u_1}-\mathbf{u_2}) \\
&=\tilde{b}_j^{(3)}(\mathbf{u_1},\mathbf{u_2}),
\end{aligned}
\tag{30}
$$

where $\tilde{b}_j^{(3)}(\mathbf{u_1},\mathbf{u_2})$ is the normalized bispectrum. Thus the average of $\tilde{D}_j^{(3)}(\mathbf{u_1},\mathbf{u_2})$ over the statistics of $\mathbf{x}_k$, $\mathbf{x}_k$ and $\mathbf{x}_k$ is

$$
\begin{aligned}
E_{klm}[\tilde{D}_j^{(3)}(\mathbf{u_1},\mathbf{u_2})] &= N_j + N_j(N_j-1)\{|\tilde{i}_j(\mathbf{u_1})|^2 + |\tilde{i}_j(\mathbf{u_2})|^2 + |\tilde{i}_j(-\mathbf{u_1}-\mathbf{u_2})|^2\} \\
&\quad + N_j(N_j-1)(N_j-2)\tilde{b}_j^{(3)}(\mathbf{u_1},\mathbf{u_2}).
\end{aligned}
\tag{31}
$$

Next $E_{klm}[\tilde{D}_j^{(3)}(\mathbf{u_1},\mathbf{u_2})]$ is averaged over the Poisson statistics of N_j. For Poisson statistics,

$$
E[N_j(N_j-1)\cdots(N_j-r+1)] = \overline{N}_j{}^r, \tag{32}
$$

where $\overline{N}_j$ denotes the Poisson mean of N_j. For a given rate function $\lambda_j(\mathbf{x})$

$$
\begin{aligned}
E_{klm,N_j}[\tilde{D}_j^{(3)}(\mathbf{u_1},\mathbf{u_2})] &= \overline{N}_j + \overline{N}_j{}^2\{|\tilde{i}_j(\mathbf{u_1})|^2 + |\tilde{i}_j(\mathbf{u_2})|^2 + |\tilde{i}_j(-\mathbf{u_1}-\mathbf{u_2})|^2\} \\
&\quad + \overline{N}_j{}^3\tilde{b}_j^{(3)}(\mathbf{u_1},\mathbf{u_2}).
\end{aligned}
\tag{33}
$$

Finally, averaging over the ensemble of $\lambda_j(\mathbf{x})$ or $p_j(\mathbf{x})$, yields

$$
\begin{aligned}
E[\tilde{D}^{(3)}(\mathbf{u_1},\mathbf{u_2})] &= \overline{N}_j + \overline{N}_j{}^2\{\langle|\tilde{i}(\mathbf{u_1})|^2\rangle + \langle|\tilde{i}(\mathbf{u_2})|^2\rangle + \langle|\tilde{i}(-\mathbf{u_1}-\mathbf{u_2})|^2\rangle\} \\
&\quad + \overline{N}_j{}^3\langle\tilde{b}^{(3)}(\mathbf{u_1},\mathbf{u_2})\rangle.
\end{aligned}
\tag{34}
$$

If $\overline{N}_j$ does not fluctuate from frame to frame, *i.e.*

$$
E[\overline{N}_j{}^r] = E[N_j]^r = \overline{N}^r, \tag{35}
$$

for an arbitrary r,

$$E[\tilde{D}^{(3)}(\mathbf{u_1},\mathbf{u_2})] = \overline{N} + \overline{N}^2\{\langle|\tilde{i}(\mathbf{u_1})|^2\rangle + \langle|\tilde{i}(\mathbf{u_2})|^2\rangle + \langle|\tilde{i}(-\mathbf{u_1}-\mathbf{u_2})|^2\rangle\} + \overline{N}^3\langle\tilde{b}^{(3)}(\mathbf{u_1},\mathbf{u_2})\rangle. \tag{36}$$

In order to express an unbiased estimator of $\langle\tilde{b}^{(3)}(\mathbf{u_1},\mathbf{u_2})\rangle$ using the quantities observed directly,

$$E[|\tilde{D}(\mathbf{u})|^2] = \overline{N} + \overline{N}^2\langle|\tilde{i}(\mathbf{u})|^2\rangle, \tag{37}$$

is useful. This relation was first obtained by Goodman and Belsher for non-photon-counting detection and and was also derived by Dainty and Greenaway (1979) for photon-counting detection. From it we obtain

$$\overline{N}^3\langle\tilde{b}^{(3)}(\mathbf{u_1},\mathbf{u_2})\rangle = E[\tilde{D}^{(3)}(\mathbf{u_1},\mathbf{u_2}) - \{|\tilde{D}(\mathbf{u_1})|^2 + |\tilde{D}(\mathbf{u_2})|^2 + |\tilde{D}(-\mathbf{u_1}-\mathbf{u_2})|^2 - 2\overline{N}\}]. \tag{38}$$

Thus an unbiased estimator of the bispectrum for the jth frame becomes

$$\tilde{Q}_j^{(3)}(\mathbf{u_1},\mathbf{u_2}) = \tilde{D}_j^{(3)}(\mathbf{u_1},\mathbf{u_2}) - \{|\tilde{D}_j(\mathbf{u_1})|^2 + |\tilde{D}_j(\mathbf{u_2})|^2 + |\tilde{D}_j(-\mathbf{u_1}-\mathbf{u_2})|^2 - 2\overline{N}_j\}, \tag{39}$$

where the terms in $\{\}$ are the photon noise bias. Equation (39) was first obtained by Wirnitzer (1985). This estimator can be rewritten as

$$\tilde{Q}_j^{(3)}(\mathbf{u_1},\mathbf{u_2}) = \sum_{k\neq l\neq m} \exp(i\{\mathbf{u_1}(\mathbf{x}_k - \mathbf{x}_m) + \mathbf{u_2}(\mathbf{x}_l - \mathbf{x}_m)\}). \tag{40}$$

Since the observables are the positions of individual photons, it is also possible to directly calculate the $N_j(N_j-1)(N_j-2)$ exponential terms through Eq. (40). The absolute minimum number of photons per frame is three, since triple cross correlations of different photon events contribute to the unbiased estimator of the classical bispectrum.

The next goal is to find the variance of the unbiased estimator $\tilde{Q}_j^{(3)}(\mathbf{u_1},\mathbf{u_2})$ and its SNR per frame. In evaluating the variance, $\sigma^2(\tilde{Q}_j^{(3)}(\mathbf{u_1},\mathbf{u_2}))$, it is necessary to calculate $E[|\tilde{Q}_j^{(3)}(\mathbf{u_1},\mathbf{u_2})|^2]$.

The derivation is systematic but lengthy and is given in Appendix 2. Here only the resultant expression is presented:

$$\begin{aligned}
\sigma^2(\tilde{Q}_j^{(3)}(\mathbf{u}_1,\mathbf{u}_2)) = \; & \overline{N}^3[1 + \langle|\tilde{i}(\mathbf{u}_1-\mathbf{u}_2)|^2\rangle + \langle|\tilde{i}(2\mathbf{u}_1-\mathbf{u}_2)|^2\rangle + \langle|\tilde{i}(\mathbf{u}_1+2\mathbf{u}_2)|^2\rangle] \\
& + \overline{N}^4[\langle|\tilde{i}(\mathbf{u}_1)|^2\rangle + \langle|\tilde{i}(\mathbf{u}_2)|^2\rangle + \langle|\tilde{i}(-\mathbf{u}_1-\mathbf{u}_2)|^2\rangle \\
& \quad + \langle\tilde{b}^{(3)}(\mathbf{u}_1,-\mathbf{u}_2)\rangle + c.c. \\
& \quad + \langle\tilde{b}^{(3)}(\mathbf{u}_1+\mathbf{u}_2,\mathbf{u}_1)\rangle + c.c. \\
& \quad + \langle\tilde{b}^{(3)}(\mathbf{u}_2,\mathbf{u}_1+\mathbf{u}_2)\rangle + c.c. \\
& \quad + \langle|\tilde{i}(\mathbf{u}_1)|^2|\tilde{i}(\mathbf{u}_1+2\mathbf{u}_2)|^2\rangle \\
& \quad + \langle|\tilde{i}(\mathbf{u}_2)|^2|\tilde{i}(2\mathbf{u}_1+\mathbf{u}_2)|^2\rangle \\
& \quad + \langle|\tilde{i}(\mathbf{u}_1+\mathbf{u}_2)|^2|\tilde{i}(\mathbf{u}_1-\mathbf{u}_2)|^2\rangle \\
& \quad + \langle\tilde{f}^{(4)}(\mathbf{u}_1-\mathbf{u}_2,\mathbf{u}_1+2\mathbf{u}_2,-\mathbf{u}_1-\mathbf{u}_2)\rangle + c.c. \\
& \quad + \langle\tilde{f}^{(4)}(\mathbf{u}_1,\mathbf{u}_1+2\mathbf{u}_2,-2\mathbf{u}_1-\mathbf{u}_2)\rangle + c.c. \\
& \quad + \langle\tilde{f}^{(4)}(2\mathbf{u}_1+\mathbf{u}_2,-\mathbf{u}_1+\mathbf{u}_2,-\mathbf{u}_1-\mathbf{u}_2)\rangle + c.c.] \\
& + \overline{N}^5[\langle|\tilde{i}(\mathbf{u}_1)|^2|\tilde{i}(-\mathbf{u}_1-\mathbf{u}_2)|^2\rangle + \langle|\tilde{i}(\mathbf{u}_2)|^2|\tilde{i}(-\mathbf{u}_1-\mathbf{u}_2)|^2\rangle + \langle|\tilde{i}(\mathbf{u}_1)|^2|\tilde{i}(\mathbf{u}_2)|^2\rangle \\
& \quad + \langle|\tilde{i}(\mathbf{u}_1)|^2(\tilde{b}^{(3)}(\mathbf{u}_2,\mathbf{u}_1+\mathbf{u}_2) + c.c.)\rangle \\
& \quad + \langle|\tilde{i}(\mathbf{u}_2)|^2(\tilde{b}^{(3)}(\mathbf{u}_1,\mathbf{u}_1+\mathbf{u}_2) + c.c.)\rangle \\
& \quad + \langle|\tilde{i}(-\mathbf{u}_1-\mathbf{u}_2)|^2(\tilde{b}^{(3)}(\mathbf{u}_1,-\mathbf{u}_2) + c.c.)\rangle] \\
& + \overline{N}^6[\langle|\tilde{b}^{(3)}(\mathbf{u}_1,\mathbf{u}_2)|^2\rangle - |\langle\tilde{b}^{(3)}(\mathbf{u}_1,\mathbf{u}_2)\rangle|^2], \qquad (41)
\end{aligned}$$

where

$$\tilde{f}^{(4)}(\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3) = \tilde{i}(\mathbf{u}_1)\tilde{i}(\mathbf{u}_2)\tilde{i}(\mathbf{u}_3)\tilde{i}(-\mathbf{u}_1-\mathbf{u}_2-\mathbf{u}_3), \qquad (42)$$

is the normalized fourth order spectrum and *c.c.* denotes complex conjugate. The SNR of the

unbiased estimator of the bispectrum is given by

$$SNR(\tilde{Q}^{(3)}(\mathbf{u_1},\mathbf{u_2})) = \frac{\overline{N}^3\langle\tilde{b}^{(3)}(\mathbf{u_1},\mathbf{u_2})\rangle}{\sqrt{\sigma^2(\tilde{Q}^{(3)}(\mathbf{u_1},\mathbf{u_2}))}}. \tag{43}$$

In order to obtain a more useful expression, it is convenient to use approximate attenuation factors of MTFs based on the HIV of the image-forming process (Appendix 1). In the mid-frequency range, MTFs are expressed by the number of speckles n_s and normalized OTF's of the telescope as follows:

$$\langle|\tilde{i}(\mathbf{u})|^2\rangle = n_s^{-1}|\tilde{t}(\mathbf{u})|^2 \tag{44}$$

$$\langle\tilde{b}^{(3)}(\mathbf{u_1},\mathbf{u_2})\rangle = n_s^{-2}\tilde{t}^{(3)}(\mathbf{u_1},\mathbf{u_2}) \tag{45}$$

$$\langle\tilde{f}^{(4)}(\mathbf{u_1},\mathbf{u_2},\mathbf{u_3})\rangle = n_s^{-3}\tilde{t}^{(4)}(\mathbf{u_1},\mathbf{u_2},\mathbf{u_3}) \tag{46}$$

$$\langle|\tilde{i}(\mathbf{u_1})|^2|\tilde{i}(\mathbf{u_2})|^2\rangle = n_s^{-2}|\tilde{t}(\mathbf{u_1})|^2|\tilde{t}(\mathbf{u_2})|^2 \tag{47}$$

$$\langle|\tilde{i}(\mathbf{u_1})|^2\tilde{b}^{(3)}(\mathbf{u_2},\mathbf{u_3})\rangle = n_s^{-3}|\tilde{t}(\mathbf{u_1})|^2\tilde{t}^{(3)}(\mathbf{u_2},\mathbf{u_3}) \tag{48}$$

$$\langle|\tilde{b}^{(3)}(\mathbf{u_1},\mathbf{u_2})|^2\rangle = n_s^{-3}|\tilde{t}^{(3)}(\mathbf{u_1},\mathbf{u_2})|^2 \tag{49}$$

$$|\langle\tilde{b}^{(3)}(\mathbf{u_1},\mathbf{u_2})\rangle|^2 = n_s^{-4}|\tilde{t}^{(3)}(\mathbf{u_1},\mathbf{u_2})|^2. \tag{50}$$

Typically $n_s = (\frac{D}{r_c})^2 \geq 10^2$, and $|\tilde{t}(\mathbf{u})|^2 \approx 10^{-1}$ in the mid frequency range. By selecting leading terms of each order of $\overline{N}$ in Eq. (41), the variance of the unbiased estimator of the classical bispectral MTF, $\sigma^2(\tilde{Q}^{(3)}(\mathbf{u_1},\mathbf{u_2}))$ is found to be,

$$\begin{aligned}\sigma^2(\tilde{Q}^{(3)}(\mathbf{u_1},\mathbf{u_2})) = \; & \overline{N}^3 + \overline{N}^4 n_s^{-1}[|\tilde{t}(\mathbf{u_1})|^2 + |\tilde{t}(\mathbf{u_2})|^2 + |\tilde{t}(-\mathbf{u_1}-\mathbf{u_2})|^2] \\ & + \overline{N}^5 n_s^{-2}[|\tilde{t}(\mathbf{u_1})|^2|\tilde{t}(-\mathbf{u_1}-\mathbf{u_2})|^2 + |\tilde{t}(\mathbf{u_2})|^2|\tilde{t}(-\mathbf{u_1}-\mathbf{u_2})|^2 + |\tilde{t}(\mathbf{u_1})|^2|\tilde{t}(\mathbf{u_2})|^2] \\ & + \overline{N}^6 n_s^{-3}\tilde{t}^{(3)}(\mathbf{u_1},\mathbf{u_2}).\end{aligned} \tag{51}$$

The SNR of the unbiased estimator is then given by

$$SNR(\tilde{Q}^{(3)}(\mathbf{u_1},\mathbf{u_2})) = \frac{\overline{N}^3 n_s^{-2}\tilde{t}^{(3)}(\mathbf{u_1},\mathbf{u_2})}{\sqrt{\sigma^2(\tilde{Q}^{(3)}(\mathbf{u_1},\mathbf{u_2}))}}. \tag{52}$$

If the mean number of photons per speckle, $\overline{n}$ is defined as $\overline{n} = \frac{\overline{N}}{n_s}$, the SNR is alternatively expressed as

$$SNR(\tilde{Q}^{(3)}(\mathbf{u_1}, \mathbf{u_2})) = \frac{n_s^{-\frac{1}{2}} \overline{n}^{\frac{3}{2}} \tilde{t}^{(3)}(\mathbf{u_1}, \mathbf{u_2})}{\sqrt{A}},$$

where

$$\begin{aligned} A = \; & 1 + \overline{n}(|\tilde{t}(\mathbf{u_1})|^2 + |\tilde{t}(\mathbf{u_2})|^2 + |\tilde{t}(-\mathbf{u_1} - \mathbf{u_2})|^2) \\ & + \overline{n}^2(|\tilde{t}(\mathbf{u_1})|^2|\tilde{t}(\mathbf{u_2})|^2 + |\tilde{t}(\mathbf{u_1})|^2|\tilde{t}(-\mathbf{u_1} - \mathbf{u_2})|^2 + |\tilde{t}(\mathbf{u_2})|^2|\tilde{t}(-\mathbf{u_1} - \mathbf{u_2})|^2) \\ & + \overline{n}^3|\tilde{t}^{(3)}(\mathbf{u_1}, \mathbf{u_2})|^2, \end{aligned} \tag{53}$$

and $\sqrt{n_s} SNR(\tilde{Q}^{(3)}(\mathbf{u_1}, \mathbf{u_2}))$ is independent of $n_s = (\frac{D}{r_c})^2$. Ayers et al. (1988) obtained, by means of a different logical path, an expression identical to Eq. (53). In Fig. 10, $\sqrt{n_s} \times SNR(\tilde{Q}^{(3)}(x\frac{D}{\lambda}, 0, 0, x\frac{D}{\lambda}))$ at x = (0.2, 0.3, 0.4, 0.5, 0.6) is plotted as a function of $\overline{n}$ in logarithmic scale. An estimate of the SNR for a telescope of diameter D meters can be obtained by lowering the value on the plot by $\log(\frac{D}{r_c})$. It should again be emphasized that the above approximate estimations are valid only in the mid-frequency range.

8. Estimates of the Limiting Magnitude and Resolution

In order to estimate the limiting magnitude, the statistically independent volume of the bispectrum must be obtained first. The heuristic interferometric view suggests that the Fourier components in the mid-frequency range are statistically independent. The volume is proportional to n_s^2 which must be multiplied by a factor related to the symmetry and the boundary of the bispectrum. Wirnitzer (1985) estimated the bispectral volume as $\frac{n_s^2}{4}$. In Readhead et al. (1988), $\frac{3{n_s}^2}{4}$ for a square aperture, and $\frac{\pi^2}{32}{n_s}^2$ for a circular aperture, were obtained, assuming statistical independence of all the triangles. Since only the mid-frequency components are statistically

independent, these values give upper limits. An estimate of the SNR of an ideally recovered map from the bispectral MTF is given as

$$SNR(map) = \sqrt{\frac{\pi^2}{32} n_s^2 \overline{SNR(\tilde{Q}^{(3)}(\mathbf{u_1}, \mathbf{u_2}))} \sqrt{M}}, \tag{54}$$

where $\overline{SNR(\tilde{Q}^{(3)}(\mathbf{u_1}, \mathbf{u_2}))}$ is the average SNR over the mid-frequency range and M is the number of frames. As is estimated from Eq. (53) and is shown in Fig. 10, the SNR at the mid-frequency for $\overline{n} \leq 1$ is approximated well by

$$SNR(\tilde{Q}^{(3)}(\mathbf{u_1}, \mathbf{u_2})) \approx n_s^{-\frac{1}{2}} \overline{n}^{\frac{3}{2}} \tilde{t}^{(3)}(\mathbf{u_1}, \mathbf{u_2}), \tag{55}$$

since $A \approx 1$ in Eq. (53). Equation (55) can then be rewritten as

$$\begin{aligned} SNR(map) &= \sqrt{\frac{\pi^2}{32}} (n_s \overline{n}^3)^{\frac{1}{2}} \overline{\tilde{t}^{(3)}(\mathbf{u_1}, \mathbf{u_2})} \sqrt{M} \\ &\approx 0.027 (n_s \overline{n}^3)^{\frac{1}{2}} \sqrt{M}, \end{aligned} \tag{56}$$

where $\overline{\tilde{t}^{(3)}(\mathbf{u_1}, \mathbf{u_2})}$ is the average OTF over the mid-frequency and about 5×10^{-2}. $\overline{n}$ can be expressed as a function of the magnitude of the object m, fractional bandwidth $\Delta\lambda/\lambda$, efficiency of the detection system η, integration time $\Delta\tau$, and the coherence length r_c. Then the limiting magnitude at $\lambda = 0.55\mu$m is given as

$$\begin{aligned} m_{lim} = 13.3 + 2.5[&\log(\frac{\Delta\lambda/\lambda}{0.1}) + \log(\frac{\eta}{0.1}) + \log(\frac{\Delta\tau}{10ms}) \\ &+ \frac{4}{3}\log(\frac{r_c}{14cm}) + \frac{2}{3}\log(\frac{D}{1m}) + \frac{1}{3}\log(\frac{M}{10^4}) - \frac{2}{3}\log(\frac{SNR(map)}{10})]. \end{aligned} \tag{57}$$

With a resolution about a factor two lower than the diffraction limit, a SNR of 10 will be obtained from the attainable number of frames of 10^4 for a point source of 13.3 mag with a 1-m telescope and of 14.5 mag with a 5-m telescope. For a good observing condition, $\Delta\tau$ may be somewhat longer and the limiting magnitude may reach 15 mag. As can be seen in Fig. 10, the SNR at frequencies above $0.5\frac{D}{\lambda}$ decreases drastically according to the behavior of the OTF.

For the high frequency region, the non-redundant-masking method (Baldwin et al. 1986, Haniff et al. 1987, Readhead et al. 1988, Chapter II) is more promising than the fully-filled-aperture method of the conventional speckle. From the interferometric view, for a certain Fourier component, other Fourier components behave as backgrounds. In the presence of overwhelming lower frequency components, high frequency components are strongly suppressed because of the low redundancy of long baselines. However before we proceed to a quantitative comparison between the fully-filled aperture and the non-redundant masking, there are remaining problems to be solved such as the estimation of the independent bispectral volume for the non-redundant masking (Kulkarni and Nakajima 1989).

9. Conclusions

In this chapter the behavior of the signal-to-noise ratio of the bispectral analysis of speckle interferometry has been studied in two stages. At the high light limit, the Monte Carlo simulations of an atmospheric phase screen based on the Kolmogorov theory and recent observations of the atmospheric disturbance are used to derive statistical properties of the classical bispectral modulation transfer function. The influence of photon noise is taken into account, by modeling the photodetection process.

A general expression of the signal-to-noise ratio of the bispectrum at arbitrary light levels is obtained in terms of the classical modulation transfer functions and the mean photon counts. In the mid-frequency range, a practical expression is obtained for the signal-to-noise ratio as a function of the optical transfer functions of the telescope optics, the number of speckles and the mean photon counts was obtained.

Major conclusions are:

(1) The overall behavior of the atmospheric transfer function is qualitatively consistent with the heuristic interferometric view of the image-forming process, and, especially in the mid-frequency range, the quantitative predictions of the heuristic interferometric view agree approximately with the simulated results. At the mid frequencies, the attenuation of the bispectral modulation transfer function and the signal-to-noise ratio are approximated by the predicted values n_s^{-2} and $n_s^{-\frac{1}{2}}$ respectively.

(2) At low light levels, only bispectral components near the axes have a high signal-to-noise ratio. Closure phases near the axes are effectively local phase differences. For simple sources, the behavior of the phase in Fourier space is so regular that local phase differences are enough for a full image recovery. In recovering complicated sources, global closure phases contained in the mid-frequency range may be crucial for the reconstruction of images of complicated sources. However, the signal-to-noise ratio at the mid frequency falls off so drastically at low light levels that the effective limiting magnitudes are much lower than those of simple sources.

(3) As estimated from the signal-to-noise ratio in the mid-frequency range, the practical limiting magnitude of the bispectral analysis at a visual wavelength will be between 13 and 15 mag depending on the size of the telescope and observing conditions. This limit is achieved with a resolution half of the diffraction limit of a given telescope.

Appendix 1. Predictions Obtained by the Treatment Based on the Heuristic Interferometric View of the Image-Forming Process

A detailed treatment based on heuristic interferometric view of the image-forming process was discussed in Readhead et al.. In this appendix the major predictions are reviewed briefly.

A speckle pattern is regarded as an instantaneous interference pattern formed by a number of

elementary coherent areas on the aperture plane, whose linear sizes are about r_c. The discussion has to be restricted to mid spatial frequency, where a certain Fourier component on the image plane is given as a sum of random phasors originating from identical baselines on the aperture plane. A mid-frequency component satisfies the following two conditions. First, the corresponding baselines to a mid-frequency component are so sufficiently longer than r_c that the rms phase(σ_Φ) of the baselines is significantly larger than 2π. In other words, the rms phase correlation function of a pair of elementary areas is significantly larger than 2π. Then the unit phasor $\exp(i\Phi)$ of a baseline becomes a uniform random number on the unit circle on the complex plane. Effectively a Φ can be regarded as a uniform random number between $-\pi$ and π. Second, the redundancy of the baseline must be high so that the number of random phasors is large enough for an incoherent average to be performed. Although individual phasors have uniformly random phases, phases of neighboring baselines are correlated. In order to average well over the random phasors, the number of phasors must be significantly larger than 2π (Recall that the translation by r_c on the aperture plane causes rms phase change of 1 radian). At the highest-frequency region of a circular aperture, this condition is not satisfied, thus we restrict our discussion to the mid-frequency range.

The dependences of the power spectrum ATF, the bispectral ATF and their SNRs on the number of speckles ($n_s = (\frac{D}{r_c})^2$) are determined below. For simplicity, we neglect scintillation and assume unit phasors originating from individual baselines.

$\mathbf{u_1}$, $\tilde{I}(\mathbf{u_1})$ and $N(\mathbf{u_1})$ denote a mid spatial frequency, the Fourier component and the number of phasors or redundancy of the baseline corresponding to $\mathbf{u_1}$, respectively. The Fourier component is given as

$$\tilde{I}(\mathbf{u_1}) = \sum_{k=1}^{N(\mathbf{u_1})} \exp(i\Phi_k),$$

where Φ_k is the phase of the kth phasor. And the average power spectrum MTF over an ensemble

of speckle patterns is

$$\langle |\tilde{I}(\mathbf{u_1})|^2 \rangle = \langle \sum_{k=1}^{N(\mathbf{u_1})} \sum_{l=1}^{N(\mathbf{u_1})} \exp(i(\Phi_k - \Phi_l)) \rangle$$

$$= \sum_{k=1}^{N(\mathbf{u_1})} \sum_{l=1}^{N(\mathbf{u_1})} \langle \exp(i(\Phi_k - \Phi_l)) \rangle.$$

Since Φ_k and Φ_l are not correlated unless $k = l$,

$$\langle \exp(i(\Phi_k - \Phi_l)) \rangle = 1 \quad k = l \quad N(\mathbf{u_1}) \quad \text{terms}$$

$$= 0 \quad k \neq l \quad N(\mathbf{u_1})(N(\mathbf{u_1}) - 1) \quad \text{terms;}$$

then

$$\langle |\tilde{I}(\mathbf{u_1})|^2 \rangle = N(\mathbf{u_1}).$$

In the absence of atmospheric disturbance,

$$|\tilde{I}(\mathbf{u_1})|^2 = N(\mathbf{u_1})^2.$$

Thus the ATF or the power spectrum ATF is $N(\mathbf{u_1})^{-1}$. In addition to the ATF due to the atmosphere, there is an atmospheric noise factor $\sim N(\mathbf{u_1})$. Thus the SNR of $\langle |\tilde{I}(\mathbf{u_1})|^2 \rangle$ is 1. At the mid-frequency region, the redundancy, $N(\mathbf{u})$ is proportional to and of the order of the number of speckles, $n_s = (\frac{D}{r_c})^2$. The ATF is about $n_s{}^{-1}$.

The bispectral ATF is obtained in the same manner. The ensemble average of bispectral component at $(\mathbf{u_1}, \mathbf{u_2})$ is

$$\langle \tilde{I}^{(3)}(\mathbf{u_1}, \mathbf{u_2}) \rangle = \langle \sum_{k=1}^{N(\mathbf{u_1})} \exp(i\Phi_{12,k}) \sum_{l=1}^{N(\mathbf{u_2})} \exp(i\Phi_{23,l}) \sum_{m=1}^{N(-\mathbf{u_1}-\mathbf{u_2})} \exp(i\Phi_{31,m}) \rangle$$

$$= \sum_{k=1}^{N(\mathbf{u_1})N(\mathbf{u_2})N(-\mathbf{u_1}-\mathbf{u_2})} \sum_{l=1} \sum_{m=1} \langle \exp(i(\Phi_{12,k} + \Phi_{23,l} + \Phi_{31,m})) \rangle,$$

where 12, 23 and 31 stand for baselines corresponding to the frequencies $\mathbf{u_1}$, $\mathbf{u_2}$ and $-\mathbf{u_1} - \mathbf{u_2}$ respectively. The only terms with $k = l = m$ have have finite contributions, and other terms have

zero mean. In the ideal case that the closure-phase cancellation is perfect,

$$\langle \tilde{I}^{(3)}(\mathbf{u}_1, \mathbf{u}_2) \rangle = Min(N(\mathbf{u}_1), N(\mathbf{u}_2), N(-\mathbf{u}_1 - \mathbf{u}_2)),$$

where $Min(N(\mathbf{u}_1), N(\mathbf{u}_2), N(-\mathbf{u}_1 - \mathbf{u}_2))$ is the minimum among the three redundancies and the number of closed triangles. In the absence of atmospheric disturbance,

$$\tilde{I}^{(3)}(\mathbf{u}_1, \mathbf{u}_2) = N(\mathbf{u}_1)N(\mathbf{u}_2)N(-\mathbf{u}_1 - \mathbf{u}_2),$$

then the bispectral ATF at $(\mathbf{u}_1, \mathbf{u}_2)$ is

$$\frac{Min(N(\mathbf{u}_1), N(\mathbf{u}_2), N(-\mathbf{u}_1 - \mathbf{u}_2))}{N(\mathbf{u}_1)N(\mathbf{u}_2)N(-\mathbf{u}_1 - \mathbf{u}_2)}.$$

After $N(\mathbf{u}_1)N(\mathbf{u}_2)N(-\mathbf{u}_1 - \mathbf{u}_2)$ terms are added, the average becomes $Min(N(\mathbf{u}_1), N(\mathbf{u}_2), N(-\mathbf{u}_1 - \mathbf{u}_2))$. Therefore the SNR is

$$\frac{Min(N(\mathbf{u}_1), N(\mathbf{u}_2), N(-\mathbf{u}_1 - \mathbf{u}_2))}{\sqrt{N(\mathbf{u}_1)N(\mathbf{u}_2)N(-\mathbf{u}_1 - \mathbf{u}_2)}}.$$

Estimates of the bispectral ATF and the SNR are given as ${n_s}^{-2}$ and ${n_s}^{-\frac{1}{2}}$ respectively.

In what follows, higher-order ATF's used to derive Eqs.(46)-(50) are calculated.

$$\begin{aligned}
\langle \tilde{a}^{(4)}(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) \rangle &= \frac{\sum_{k=1}^{N(\mathbf{u}_1)} \sum_{l=1}^{N(\mathbf{u}_2)} \sum_{m=1}^{N(\mathbf{u}_3)} \sum_{n=1}^{N(-\mathbf{u}_1-\mathbf{u}_2-\mathbf{u}_3)} \langle \exp(i(\Phi_{12,k} + \Phi_{23,l} + \Phi_{34,m} + \Phi_{41,n})) \rangle}{N(\mathbf{u}_1)N(\mathbf{u}_2)N(\mathbf{u}_3)N(-\mathbf{u}_1 - \mathbf{u}_2 - \mathbf{u}_3)} \\
&= \frac{Min(N(\mathbf{u}_1), N(\mathbf{u}_2), N(\mathbf{u}_3), N(-\mathbf{u}_1 - \mathbf{u}_2 - \mathbf{u}_3))}{N(\mathbf{u}_1)N(\mathbf{u}_2)N(\mathbf{u}_3)N(-\mathbf{u}_1 - \mathbf{u}_2 - \mathbf{u}_3)} \\
&\approx n_s^{-3},
\end{aligned}$$

since the following closure-phase cancellations work only the terms with $k = l = m = n$,

$$\begin{aligned}
\Phi_{12,k} + \Phi_{23,k} + \Phi_{34,k} + \Phi_{41,k} &= (\Phi_{12,k} + \Phi_{23,k} + \Phi_{31,k}) + (\Phi_{13,k} + \Phi_{34,k} + \Phi_{41,k}) \\
&= 0.
\end{aligned}$$

Thus we have

$$\begin{aligned}
\langle |\tilde{a}(\mathbf{u}_1)|^2 |\tilde{a}(\mathbf{u}_2)|^2 \rangle &= \frac{\sum_{k=1}^{N(\mathbf{u}_1)} \sum_{l=1}^{N(\mathbf{u}_1)} \sum_{m=1}^{N(\mathbf{u}_2)} \sum_{n=1}^{N(\mathbf{u}_2)} \langle \exp(i(\Phi_{12,k} - \Phi_{12,l} + \Phi_{34,m} - \Phi_{34,n})) \rangle}{N(\mathbf{u}_1)^2 N(\mathbf{u}_2)^2} \\
&= \frac{N(\mathbf{u}_1)N(\mathbf{u}_2)}{N(\mathbf{u}_1)^2 N(\mathbf{u}_2)^2} \\
&\approx n_s^{-2},
\end{aligned}$$

since only for the terms with $k = l$, $m = n$,

$$\langle \exp(i(\Phi_{12,k} - \Phi_{12,l} + \Phi_{34,m} - \Phi_{34,n}))\rangle = 1;$$

$$\begin{aligned}\langle |\tilde{a}(\mathbf{u_1})|^2 \tilde{a}^{(3)}(\mathbf{u_2},\mathbf{u_3})\rangle &= \sum_{k=1}^{N(\mathbf{u_1})}\sum_{l=1}^{N(\mathbf{u_1})}\sum_{m=1}^{N(\mathbf{u_2})}\sum_{n=1}^{N(\mathbf{u_3})}\sum_{o=1}^{N(-\mathbf{u_2}-\mathbf{u_3})} \langle \exp(i(\Phi_{12,k} - \Phi_{12,l} + \Phi_{34,m} + \Phi_{45,n} + \Phi_{53,o}))\rangle \\ &/N(\mathbf{u_1})^2 N(\mathbf{u_2}) N(\mathbf{u_3}) N(-\mathbf{u_2}-\mathbf{u_3}) \\ &= \frac{N(\mathbf{u_1}) \times Min(N(\mathbf{u_2}), N(\mathbf{u_3}), N(-\mathbf{u_2}-\mathbf{u_3}))}{N(\mathbf{u_1})^2 N(\mathbf{u_2}) N(\mathbf{u_3}) N(-\mathbf{u_2}-\mathbf{u_3})} \\ &\approx n_s^{-3},\end{aligned}$$

since only for the terms with $k = l$, $m = n = o$,

$$\langle \exp(i(\Phi_{12,k} - \Phi_{12,l} + \Phi_{34,m} + \Phi_{45,n} + \Phi_{53,o}))\rangle = 1,$$

$$\begin{aligned}\langle |\tilde{a}^{(3)}(\mathbf{u_1},\mathbf{u_2})|^2\rangle &= \sum_{k=1}^{N(\mathbf{u_1})}\sum_{l=1}^{N(\mathbf{u_2})}\sum_{m=1}^{N(-\mathbf{u_1}-\mathbf{u_2})}\sum_{n=1}^{N(\mathbf{u_1})}\sum_{o=1}^{N(\mathbf{u_2})}\sum_{p=1}^{N(-\mathbf{u_1}-\mathbf{u_2})} \\ &\langle \exp(i(\Phi_{12,k} + \Phi_{23,l} + \Phi_{31,m} - \Phi_{12,n} - \Phi_{23,o} - \Phi_{31,p}))\rangle \\ &/N(\mathbf{u_1})^2 N(\mathbf{u_2})^2 N(-\mathbf{u_1}-\mathbf{u_2})^2 \\ &= \frac{N(\mathbf{u_1})N(\mathbf{u_2})N(-\mathbf{u_1}-\mathbf{u_2}) + Min(N(\mathbf{u_2}), N(\mathbf{u_3}), N(-\mathbf{u_2}-\mathbf{u_3}))^2}{N(\mathbf{u_1})^2 N(\mathbf{u_3})^2 N(-\mathbf{u_1}-\mathbf{u_2})^2} \\ &\approx n_s^{-3};\end{aligned}$$

since only for the terms with $k = n, l = o, m = p$ or $k = l = m, o = p = q$,

$$\langle \exp(i(\Phi_{12,k} + \Phi_{23,l} + \Phi_{31,m} - \Phi_{12,n} - \Phi_{23,o} - \Phi_{31,p}))\rangle = 1.$$

Appendix 2. Calculations of the Variance of the Bispectrum

In the estimation of the expected value of the variance, the starting point is the modulus squared of Eq. (40):

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\sum_{\alpha\neq\beta\neq\gamma}\sum_{\delta\neq\epsilon\neq\zeta} \exp(i\{\mathbf{u_1}(\mathbf{x}_\alpha - \mathbf{x}_\gamma - \mathbf{x}_\delta + \mathbf{x}_\zeta) + \mathbf{u_2}(\mathbf{x}_\beta - \mathbf{x}_\gamma - \mathbf{x}_\epsilon + \mathbf{x}_\zeta)\})]$$

$$= \sum_{\alpha\neq\beta\neq\gamma}\sum_{\delta\neq\epsilon\neq\zeta} E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{\mathbf{u_1}(\mathbf{x}_\alpha - \mathbf{x}_\gamma - \mathbf{x}_\delta + \mathbf{x}_\zeta) + \mathbf{u_2}(\mathbf{x}_\beta - \mathbf{x}_\gamma - \mathbf{x}_\epsilon + \mathbf{x}_\zeta)\})],$$

where $E_{\alpha\beta\gamma\delta\epsilon\zeta}$ stands for an average over $\mathbf{x}_\alpha$, $\mathbf{x}_\beta$, $\mathbf{x}_\gamma$, $\mathbf{x}_\delta$, $\mathbf{x}_\epsilon$ and $\mathbf{x}_\zeta$. The $\{N_j(N_j-1)(N_j-2)\}^2$ terms are classified as follows:

(1) $\alpha=\delta, \beta=\epsilon, \gamma=\zeta$ $\quad N_j(N_j-1)(N_j-2)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[1]=1$$

(2) $\alpha=\epsilon, \beta=\delta, \gamma=\zeta$ $\quad N_j(N_j-1)(N_j-2)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i(\mathbf{u_1}-\mathbf{u_2})(\mathbf{x}_\alpha-\mathbf{x}_\beta))]=|\tilde{i}_j(\mathbf{u_1}-\mathbf{u_2})|^2$$

(3) $\alpha=\zeta, \beta=\epsilon, \gamma=\delta$ $\quad N_j(N_j-1)(N_j-2)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i(2\mathbf{u_1}+\mathbf{u_2})(\mathbf{x}_\alpha-\mathbf{x}_\gamma))]=|\tilde{i}_j(2\mathbf{u_1}+\mathbf{u_2})|^2$$

(4) $\alpha=\delta, \beta=\zeta, \gamma=\epsilon$ $\quad N_j(N_j-1)(N_j-2)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i(\mathbf{u_1}+2\mathbf{u_2})(\mathbf{x}_\beta-\mathbf{x}_\gamma))]=|\tilde{i}_j(\mathbf{u_1}+2\mathbf{u_2})|^2$$

(5) $\alpha=\epsilon, \beta=\zeta, \gamma=\delta$ $\quad N_j(N_j-1)(N_j-2)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{(\mathbf{u_1}-\mathbf{u_2})\mathbf{x}_\alpha+(\mathbf{u_1}+2\mathbf{u_2})\mathbf{x}_\beta+(-2\mathbf{u_1}-\mathbf{u_2})\mathbf{x}_\gamma\})]=\tilde{b}_j^{(3)}(\mathbf{u_1}-\mathbf{u_2},\mathbf{u_1}+2\mathbf{u_2})$$

(6) $\alpha=\zeta, \beta=\delta, \gamma=\epsilon$ $\quad N_j(N_j-1)(N_j-2)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{(2\mathbf{u_1}+\mathbf{u_2})\mathbf{x}_\alpha+(-\mathbf{u_1}+\mathbf{u_2})\mathbf{x}_\beta+(-\mathbf{u_1}-2\mathbf{u_2})\mathbf{x}_\gamma\})]=\tilde{b}_j^{(3)}(2\mathbf{u_1}+\mathbf{u_2},-\mathbf{u_1}+\mathbf{u_2})$$

(7) $\alpha\neq\delta, \beta=\epsilon, \gamma=\zeta$ $\quad N_j(N_j-1)(N_j-2)(N_j-3)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\mathbf{u_1}(\mathbf{x}_\alpha-\mathbf{x}_\delta))]=|\tilde{i}_j(\mathbf{u_1})|^2$$

(8) $\alpha = \delta, \beta \neq \epsilon, \gamma = \zeta$ $\quad N_j(N_j-1)(N_j-2)(N_j-3)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\mathbf{u_2}(\mathbf{x}_\beta - \mathbf{x}_\epsilon))] = |\tilde{i}_j(\mathbf{u_2})|^2$$

(9) $\alpha = \delta, \beta = \epsilon, \gamma \neq \zeta$ $\quad N_j(N_j-1)(N_j-2)(N_j-3)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i(-\mathbf{u_1}-\mathbf{u_2})(\mathbf{x}_\gamma - \mathbf{x}_\zeta))] = |\tilde{i}_j(-\mathbf{u_1}-\mathbf{u_2})|^2$$

(10) $\alpha \neq \epsilon, \beta = \delta, \gamma = \zeta$ $\quad N_j(N_j-1)(N_j-2)(N_j-3)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{\mathbf{u_1}\mathbf{x}_\alpha + (-\mathbf{u_1}+\mathbf{u_2})\mathbf{x}_\beta - \mathbf{u_2}\mathbf{x}_\epsilon\})] = \tilde{b}_j^{(3)}(\mathbf{u_1}, -\mathbf{u_2})$$

(11) $\alpha = \epsilon, \beta \neq \delta, \gamma = \zeta$ $\quad N_j(N_j-1)(N_j-2)(N_j-3)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{(\mathbf{u_1}-\mathbf{u_2})\mathbf{x}_\alpha + \mathbf{u_2}\mathbf{x}_\beta - \mathbf{u_1}\mathbf{x}_\delta\})] = \tilde{b}_j^{(3)}(-\mathbf{u_1}, \mathbf{u_2})$$
$$= \tilde{b}_j^{(3)}(\mathbf{u_1}, -\mathbf{u_2})^*$$

(12) $\alpha = \delta, \beta \neq \zeta, \gamma = \epsilon$ $\quad N_j(N_j-1)(N_j-2)(N_j-3)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{\mathbf{u_2}\mathbf{x}_\beta - (\mathbf{u_1}+2\mathbf{u_2})\mathbf{x}_\gamma + (\mathbf{u_1}+\mathbf{u_2})\mathbf{x}_\zeta\})] = \tilde{b}_j^{(3)}(\mathbf{u_1}+\mathbf{u_2}, \mathbf{u_2})$$

(13) $\alpha = \delta, \beta = \zeta, \gamma \neq \epsilon$ $\quad N_j(N_j-1)(N_j-2)(N_j-3)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{(\mathbf{u_1}+2\mathbf{u_2})\mathbf{x}_\beta - (\mathbf{u_1}+\mathbf{u_2})\mathbf{x}_\gamma + \mathbf{u_2}\mathbf{x}_\epsilon\})] = \tilde{b}_j^{(3)}(-\mathbf{u_1}-\mathbf{u_2}, \mathbf{u_2})$$
$$= \tilde{b}_j^{(3)}(\mathbf{u_1}+\mathbf{u_2}, \mathbf{u_2})^*$$

(14) $\alpha \neq \zeta, \beta = \epsilon, \gamma = \delta$ $\quad N_j(N_j-1)(N_j-2)(N_j-3)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{\mathbf{u_1}\mathbf{x}_\alpha - (2\mathbf{u_1}+\mathbf{u_2})\mathbf{x}_\gamma + (\mathbf{u_1}+\mathbf{u_2})\mathbf{x}_\zeta\})] = \tilde{b}_j^{(3)}(\mathbf{u_1}+\mathbf{u_2}, \mathbf{u_1})$$

(15) $\alpha = \zeta, \beta = \epsilon, \gamma \neq \delta$ $\quad N_j(N_j-1)(N_j-2)(N_j-3)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{(2\mathbf{u_1}+\mathbf{u_2})\mathbf{x}_\alpha - (\mathbf{u_1}+\mathbf{u_2})\mathbf{x}_\gamma - \mathbf{u_1}\mathbf{x}_\delta\})] = \tilde{b}_j^{(3)}(-\mathbf{u_1}-\mathbf{u_2}, -\mathbf{u_1})$$
$$= \tilde{b}_j^{(3)}(\mathbf{u_1}+\mathbf{u_2}, \mathbf{u_1})^*$$

(16) $\alpha \neq \delta, \beta = \zeta, \gamma = \epsilon$ $N_j(N_j-1)(N_j-2)(N_j-3)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{\mathbf{u_1}(\mathbf{x}_\alpha - \mathbf{x}_\delta) + (\mathbf{u_1}+2\mathbf{u_2})(\mathbf{x}_\beta - \mathbf{x}_\gamma)\})] = |\tilde{i}_j(\mathbf{u_1})|^2 |\tilde{i}_j(\mathbf{u_1}+2\mathbf{u_2})|^2$$

(17) $\alpha = \zeta, \beta \neq \epsilon, \gamma = \delta$ $N_j(N_j-1)(N_j-2)(N_j-3)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{(\mathbf{u_2}(\mathbf{x}_\beta - \mathbf{x}_\epsilon) + (2\mathbf{u_1}+\mathbf{u_2})(\mathbf{x}_\alpha - \mathbf{x}_\gamma)\})] = |\tilde{i}_j(\mathbf{u_2})|^2 |\tilde{i}(2\mathbf{u_1}+\mathbf{u_2})|^2$$

(18) $\alpha = \epsilon, \beta = \delta, \gamma \neq \zeta$ $N_j(N_j-1)(N_j-2)(N_j-3)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{(\mathbf{u_1}-\mathbf{u_2})(\mathbf{x}_\alpha - \mathbf{x}_\beta) - (\mathbf{u_1}+\mathbf{u_2})(\mathbf{x}_\gamma - \mathbf{x}_\zeta)\})] = |\tilde{i}_j(\mathbf{u_1}-\mathbf{u_2})|^2 |\tilde{i}_j(\mathbf{u_1}+\mathbf{u_2})|^2$$

(19) $\alpha = \epsilon, \beta = \zeta, \gamma \neq \delta$ $N_j(N_j-1)(N_j-2)(N_j-3)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{(\mathbf{u_1}-\mathbf{u_2})\mathbf{x}_\alpha + (\mathbf{u_1}+2\mathbf{u_2})\mathbf{x}_\beta - (\mathbf{u_1}+\mathbf{u_2})\mathbf{x}_\gamma - \mathbf{u_1}\mathbf{x}_\delta\})]$$

$$= \tilde{i}_j(\mathbf{u_1}-\mathbf{u_2})\tilde{i}_j(\mathbf{u_1}+2\mathbf{u_2})\tilde{i}_j(-\mathbf{u_1}-\mathbf{u_2})\tilde{i}_j(-\mathbf{u_1})$$

$$= \tilde{f}_j^{(4)}(\mathbf{u_1}-\mathbf{u_2}, \mathbf{u_1}+2\mathbf{u_2}, -\mathbf{u_1}-\mathbf{u_2})$$

(20) $\alpha \neq \zeta, \beta = \delta, \gamma = \epsilon$ $N_j(N_j-1)(N_j-2)(N_j-3)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{\mathbf{u_1}\mathbf{x}_\alpha + (-\mathbf{u_1}+\mathbf{u_2})\mathbf{x}_\beta - (\mathbf{u_1}+2\mathbf{u_2})\mathbf{x}_\gamma + (\mathbf{u_1}+\mathbf{u_2})\mathbf{x}_\delta\})]$$

$$= \tilde{i}_j(\mathbf{u_1})\tilde{i}_j(-\mathbf{u_1}+\mathbf{u_2})\tilde{i}_j(-\mathbf{u_1}-2\mathbf{u_2})\tilde{i}_j(\mathbf{u_1}+\mathbf{u_2})$$

$$= \tilde{f}_j^{(4)}(\mathbf{u_1}-\mathbf{u_2}, \mathbf{u_1}+2\mathbf{u_2}, -\mathbf{u_1}-\mathbf{u_2})^*$$

(21) $\alpha \neq \epsilon, \beta = \zeta, \gamma = \delta$ $N_j(N_j-1)(N_j-2)(N_j-3)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{\mathbf{u_1}\mathbf{x}_\alpha + (\mathbf{u_1}+2\mathbf{u_2})\mathbf{x}_\beta - (2\mathbf{u_1}+\mathbf{u_2})\mathbf{x}_\gamma - \mathbf{u_2}\mathbf{x}_\epsilon\})] = \tilde{f}_j^{(4)}(\mathbf{u_1}, \mathbf{u_1}+2\mathbf{u_2}, -2\mathbf{u_1}-\mathbf{u_2})$$

(22) $\alpha = \zeta, \beta \neq \delta, \gamma = \epsilon$ $N_j(N_j-1)(N_j-2)(N_j-3)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{(2\mathbf{u_1}+\mathbf{u_2})\mathbf{x}_\alpha + \mathbf{u_2}\mathbf{x}_\beta - (\mathbf{u_1}+2\mathbf{u_2})\mathbf{x}_\gamma - \mathbf{u_1}\mathbf{x}_\delta\})] = \tilde{f}_j^{(4)}(\mathbf{u_1}, \mathbf{u_1}+2\mathbf{u_2}, -2\mathbf{u_1}-\mathbf{u_2})^*$$

(23) $\alpha = \zeta, \beta = \delta, \gamma \neq \epsilon \qquad N_j(N_j-1)(N_j-2)(N_j-3)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{(2\mathbf{u_1}+\mathbf{u_2})\mathbf{x}_\alpha + (-\mathbf{u_1}+\mathbf{u_2})\mathbf{x}_\beta - (\mathbf{u_1}+\mathbf{u_2})\mathbf{x}_\gamma - \mathbf{u_2}\mathbf{x}_\epsilon\})]$$

$$= \tilde{f}_j^{(4)}(2\mathbf{u_1}+\mathbf{u_2}, -\mathbf{u_1}+\mathbf{u_2}, -\mathbf{u_1}-\mathbf{u_2})$$

(24) $\alpha = \epsilon, \beta \neq \zeta, \gamma = \delta \qquad N_j(N_j-1)(N_j-2)(N_j-3)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{(\mathbf{u_1}-\mathbf{u_2})\mathbf{x}_\alpha + \mathbf{u_2}\mathbf{x}_\beta + (-2\mathbf{u_1}+\mathbf{u_2})\mathbf{x}_\gamma + (\mathbf{u_1}+\mathbf{u_2})\mathbf{x}_\epsilon\})]$$

$$= \tilde{f}_j^{(4)}(2\mathbf{u_1}+\mathbf{u_2}, -\mathbf{u_1}+\mathbf{u_2}, -\mathbf{u_1}-\mathbf{u_2})^*$$

(25) $\alpha \neq \delta, \beta \neq \epsilon, \gamma = \zeta \qquad N_j(N_j-1)(N_j-2)(N_j-3)(N_j-4)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{\mathbf{u_1}(\mathbf{x}_\alpha - \mathbf{x}_\delta) + \mathbf{u_2}(\mathbf{x}_\beta - \mathbf{x}_\epsilon)\})] = |\tilde{i}_j(\mathbf{u_1})|^2 |\tilde{i}_j(\mathbf{u_2})|^2$$

(26) $\alpha \neq \delta, \beta = \epsilon, \gamma \neq \zeta \qquad N_j(N_j-1)(N_j-2)(N_j-3)(N_j-4)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{\mathbf{u_1}(\mathbf{x}_\alpha - \mathbf{x}_\delta) - (\mathbf{u_1}+\mathbf{u_2})(\mathbf{x}_\gamma - \mathbf{x}_\zeta)\})] = |\tilde{i}_j(\mathbf{u_1})|^2 |\tilde{i}_j(-\mathbf{u_1}-\mathbf{u_2})|^2$$

(27) $\alpha = \delta, \beta \neq \epsilon, \gamma \neq \zeta \qquad N_j(N_j-1)(N_j-2)(N_j-3)(N_j-4)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{\mathbf{u_2}(\mathbf{x}_\beta - \mathbf{x}_\epsilon) - (\mathbf{u_1}+\mathbf{u_2})(\mathbf{x}_\gamma - \mathbf{x}_\zeta)\})] = |\tilde{i}_j(\mathbf{u_2})|^2 |\tilde{i}_j(-\mathbf{u_1}-\mathbf{u_2})|^2$$

(28) $\alpha \neq \beta \neq \delta \neq \zeta, \gamma = \epsilon \qquad N_j(N_j-1)(N_j-2)(N_j-3)(N_j-4)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{\mathbf{u_1}(\mathbf{x}_\alpha - \mathbf{x}_\delta) + \mathbf{u_2}(\mathbf{x}_\beta - \mathbf{x}_\gamma) + (\mathbf{u_1}+\mathbf{u_2})(\mathbf{x}_\zeta - \mathbf{x}_\gamma)\})] = |\tilde{i}_j(\mathbf{u_1})|^2 \tilde{b}_j^{(3)}(\mathbf{u_2}, \mathbf{u_1}+\mathbf{u_2})$$

(29) $\alpha \neq \gamma \neq \delta \neq \epsilon, \beta = \zeta \qquad N_j(N_j-1)(N_j-2)(N_j-3)(N_j-4)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{\mathbf{u_1}(\mathbf{x}_\alpha - \mathbf{x}_\delta) - \mathbf{u_2}(\mathbf{x}_\epsilon - \mathbf{x}_\beta) - (\mathbf{u_1}+\mathbf{u_2})(\mathbf{x}_\gamma - \mathbf{x}_\beta)\})] = |\tilde{i}_j(\mathbf{u_1})|^2 \tilde{b}_j^{(3)}(\mathbf{u_2}, \mathbf{u_1}+\mathbf{u_2})^*$$

(30) $\alpha \neq \beta \neq \epsilon \neq \zeta, \gamma = \delta \qquad N_j(N_j-1)(N_j-2)(N_j-3)(N_j-4)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{\mathbf{u_2}(\mathbf{x}_\beta - \mathbf{x}_\epsilon) + \mathbf{u_1}(\mathbf{x}_\alpha - \mathbf{x}_\gamma) + (\mathbf{u_1}+\mathbf{u_2})(\mathbf{x}_\zeta - \mathbf{x}_\gamma)\})] = |\tilde{i}_j(\mathbf{u_2})|^2 \tilde{b}_j^{(3)}(\mathbf{u_1}, \mathbf{u_1}+\mathbf{u_2})$$

(31) $\beta \neq \gamma \neq \delta \neq \epsilon, \alpha = \zeta$ $N_j(N_j-1)(N_j-2)(N_j-3)(N_j-4)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{\mathbf{u_2}(\mathbf{x}_\beta - \mathbf{x}_\epsilon) - \mathbf{u_1}(\mathbf{x}_\delta - \mathbf{x}_\alpha) - (\mathbf{u_1}+\mathbf{u_2})(\mathbf{x}_\gamma - \mathbf{x}_\alpha)\})] = |\tilde{i}_j(\mathbf{u_2})|^2 \tilde{b}_j^{(3)}(\mathbf{u_1}, \mathbf{u_1}+\mathbf{u_2})^*$$

(32) $\alpha \neq \gamma \neq \epsilon \neq \zeta, \beta = \delta$ $N_j(N_j-1)(N_j-2)(N_j-3)(N_j-4)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{-(\mathbf{u_1}+\mathbf{u_2})(\mathbf{x}_\gamma - \mathbf{x}_\zeta) + \mathbf{u_1}(\mathbf{x}_\alpha - \mathbf{x}_\beta) - \mathbf{u_2}(\mathbf{x}_\epsilon - \mathbf{x}_\beta)\})] = |\tilde{i}_j(-\mathbf{u_1}-\mathbf{u_2})|^2 \tilde{b}_j^{(3)}(\mathbf{u_1}, -\mathbf{u_2})$$

(33) $\beta \neq \gamma \neq \delta \neq \zeta, \alpha = \epsilon$ $N_j(N_j-1)(N_j-2)(N_j-3)(N_j-4)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{-(\mathbf{u_1}+\mathbf{u_2})(\mathbf{x}_\gamma - \mathbf{x}_\zeta) - \mathbf{u_1}(\mathbf{x}_\delta - \mathbf{x}_\alpha) + \mathbf{u_2}(\mathbf{x}_\beta - \mathbf{x}_\alpha)\})] = |\tilde{i}_j(-\mathbf{u_1}-\mathbf{u_2})|^2 \tilde{b}_j^{(3)}(\mathbf{u_1}, -\mathbf{u_2})^*$$

(34) $\alpha \neq \beta \neq \gamma \neq \delta \neq \epsilon \neq \zeta$ $N_j(N_j-1)(N_j-2)(N_j-3)(N_j-4)(N_j-5)$ terms

$$E_{\alpha\beta\gamma\delta\epsilon\zeta}[\exp(i\{\mathbf{u_1}(\mathbf{x}_\alpha - \mathbf{x}_\delta) + \mathbf{u_2})(\mathbf{x}_\beta - \mathbf{x}_\epsilon) - (\mathbf{u_1}+\mathbf{u_2})(\mathbf{x}_\gamma - \mathbf{x}_\zeta)\})] = |\tilde{i}_j(\mathbf{u_1})|^2 |\tilde{i}_j(\mathbf{u_2})|^2 |\tilde{i}_j(-\mathbf{u_1}-\mathbf{u_2})|^2$$
$$= |\tilde{b}_j^{(3)}(\mathbf{u_1}, \mathbf{u_2})|^2$$

Averaging over the statistics of N_j and $\lambda_j(\mathbf{x})$ and assuming that Eq. (35), we obtain Eq. (41).

Appendix 3. Variance for Non-Photon-Counting Detection

For an observation with a non-photon-counting detector, it is impossible to remove bias terms frame by frame. An unbiased estimator of classical bispectrum is again given by Eq. (39). However, fluctuations of the bias terms cause additional terms in the variance of the unbiased

estimator of the classical bispectrum, Eq. (40). Those terms are evaluated as

$$\overline{N} + \overline{N}^2 + 2(\overline{N}^3 + 2\overline{N}^2)[\langle|\tilde{i}(\mathbf{u_1})|^2\rangle + \langle|\tilde{i}(\mathbf{u_2})|^2\rangle + \langle|\tilde{i}(-\mathbf{u_1}-\mathbf{u_2})|^2\rangle]$$

$$+ (\overline{N}^4 + 3\overline{N}^3)[\langle\tilde{b}^{(3)}(\mathbf{u_1},\mathbf{u_2})\rangle + c.c.]$$

$$+ (\overline{N}^4 + 4\overline{N}^3 + 2\overline{N}^2)[\langle|\tilde{i}(\mathbf{u_1})|^4\rangle + \langle|\tilde{i}(\mathbf{u_2})|^4\rangle + \langle|\tilde{i}(-\mathbf{u_1}-\mathbf{u_2})|^4\rangle$$

$$+ 2(\langle|\tilde{i}(\mathbf{u_1})|^2|\tilde{i}(\mathbf{u_2})|^2\rangle + \langle|\tilde{i}(\mathbf{u_2})|^2|\tilde{i}(-\mathbf{u_1}-\mathbf{u_2})|^2\rangle + \langle|\tilde{i}(-\mathbf{u_1}-\mathbf{u_2})|^2|\tilde{i}(\mathbf{u_1})|^2\rangle)]$$

$$+ (\overline{N}^5 + 6\overline{N}^4 + 6\overline{N}^3)\langle[\tilde{b}^{(3)}(\mathbf{u_1},\mathbf{u_2})\rangle + c.c.][|\tilde{i}(\mathbf{u_1})|^2 + |\tilde{i}(\mathbf{u_2})|^2 + |\tilde{i}(-\mathbf{u_1}-\mathbf{u_2})|^2]\rangle$$

References

Ayers, G.R., Northcott, M.J., and Dainty, J.C. 1988 J. Opt. Soc. Am. A5, 963.

Baldwin, J.E., Haniff, C.A., Mackay, C.D., and Warner P.J. 1986. Nature **320**, 595.

Dainty, J.C. and A.H. Greenaway, A.H. 1979. J. Opt. Soc. Am. **64**, 786.

Goodman, J.W. 1985. *Statistical Optics*, John Wiley & Sons.

Goodman, J.W. and Belsher, J.F. 1976. ARPA Order No. 2646, Technical Report RADC-TR-76-50. (Rome Air Development Center, NY, 1976). ARPA Order No. 2646, Technical Report RADC-TR-76-382, (Rome Air Development Center, NY, 1976).

Goodman, J.W. and Belsher, J.F. 1977. Technical Report RADC-TR-77-175 (Rome Air Development Center, NY, 1977).

Haniff, C.A., Mackay, C.D., Titterington, D.J., Sivia, D., Baldwin, J.E., and Warner, P.J. 1987. Nature **328**, 694.

Hofmann, K.-H. and Weigelt, G. 1986. Astron. Astrophys. **167**, L15.

Karbelkar, S.N. and Nityananda, R. 1987. J. Astrophys. Astr. **8**, 271.

Knox, K.T. and Thompson, B.J. 1974. Astrophys. J. **193**, L45.

Korff, D.J. 1973. J. Opt. Soc. Am., **63**, 971.

Kolmogorov, A.N. 1961. In *Turbulence, Classical Papers on Statistical Theory*, eds, S.K. Friedlander and L. Topper, Wiley-Interscience, New York.

Kulkarni, S.R. and Nakajima, T. 1989. To be submitted to J. Opt. Soc. Am. **A**.

Labeyrie, A. 1970. Astron. Astrophys. **6**, 85.

Lohmann, A.W., Weigelt, G.P., and Wirnitzer, B. 1983. Appl. Opt. **22**, 4028.

Readhead, A.C.S. and Wilkinson, P.N. 1978. Astrophys. J. **223**, 25.

Readhead, A.C.S., Nakajima, T., Pearson, T.J., Neugebauer, G., Oke, J.B. and Sargent, W.L.W. 1988. Astron. J, **95**, 1278.

Roddier, C. and Roddier, F. 1986. In *Proceedings of second workshop on ESO's Very Large Telescope*, D'Odorico, S. and Swings, J.P., eds. (European Southern Observatory, Garching), p 269.

Roddier, F. 1981. In *Proceedings of the ESO conference on The Scientific Importance of High Angular Resolution at Infrared and Optical Wavelengths*, Ulrich, M.H. and Kjar, K., eds. (European Southern Observatory, Garching), p 5.

F. Roddier, 1986. Optics. Comm., **60**, 145.

Rogers, G.L. 1963. Proc. Phys, Soc., **81**, 323.

Tatarskii, V.I. 1961. *Wave Propagation in a Turbulent Medium*, New York: McGraw-Hill.

Tatarskii, V.I. 1967. *The Effects of the Turbulent Atmosphere on Wave Propagation*, Moscow, Nauka. (1967) (transl. by Israel Program for Scientific Translations, Keter Press, Jersalem, 1971).

Taylor, G.I. 1961. In *Turbulence, Classical Papers on Statistical Theory* Friedlander, S.K. and Topper, L., eds., Wiley-Interscience, New York.

Vernin, J. 1986. In *Proceedings of second workshop on ESO's Very Large Telescope*, D'Odorico, S. and Swings, J.P., eds. (European Southern Observatory, Garching), p 279.

Wirnitzer, B. 1985, J. Opt. Soc. Am. **A2**, 14.

Figure Captions

Fig. 1 - A cross section of a simulated phase screen, plotted as a function of x coordinate on the aperture plane.

Fig. 2 - Two-dimensional coordinate system (ζ, η) adopted to represent a two-dimensional cross section of the four-dimensional bispectrum $\tilde{b}^{(3)}(\zeta,0,0,\eta)$ and that of the signal-to-noise ratio. $\mathbf{u_1} = (\zeta,0)$ and $\mathbf{u_2} = (0,\eta)$ are perpendicular to each other and the third spatial frequency $-\mathbf{u_1} - \mathbf{u_2} = (-\zeta,-\eta)$ has the largest modulus $\sqrt{\zeta^2+\eta^2}$. Therefore the circle $\zeta^2+\eta^2 = (\frac{D}{\lambda})^2$ forms the boundary, where $\frac{D}{\lambda}$ is the telescope cutoff frequency.

Fig. 3 - Contour map of the normalized bispectral modulation transfer function $\tilde{b}^{(3)}(\zeta,0,0,\eta)$ for a 2-m telescope, drawn on the ζ-η plane in a logarithmic scale. The numbers labeling contours indicate powers of 10.

Fig. 4 - Contour map of the normalized optical transfer function of the telescope, $\tilde{t}^{(3)}(\zeta,0,0,\eta)$, drawn on the ζ-η plane in a logarithmic scale. The numbers labeling contours indicate powers of 10.

Fig. 5 - Contour map of the attenuation factor, $\tilde{a}^{(3)}(\zeta,0,0,\eta)$ for a 2-m telescope, drawn on the ζ-η plane in a logarithmic scale. The number labeling contours indicate powers of 10.

Fig. 6 - Contour map of the signal-to-noise ratio of the bispectral modulation transfer function $SNR(\tilde{b}^{(3)}(\zeta,0,0,\eta))$ for a single frame obtained with a 2-m telescope, drawn on the ζ-η plane. This map shows the saturated signal-to-noise ratio at high light levels.

Fig. 7 - Power spectrum attenuation factor $|\tilde{a}(\mathbf{u})|^2$, plotted as a function of the modulus of the radial spatial frequency $|\mathbf{u}|$.

Fig. 8 - Contour map of the signal-to-noise ratio for a V=12.3 mag star obtained with a 2-m telescope after integrating 10^4 frames, assuming 10% observing efficiency, 10% fractional bandwidth and 10-ms integration time. These brightness and observing efficiency correspond to one photon per speckle.

Fig. 9 - Behavior of 3 σ contours, plotted according to the light levels. Magnitudes are calculated assuming the same conditions as for Fig. 8.

Fig. 10 - Light level and spatial-frequency dependence of the signal-to-noise ratio, $\sqrt{n_s} \cdot SNR(\tilde{Q}^{(3)}(x \cdot \frac{D}{\lambda}, 0, 0, x \cdot \frac{D}{\lambda}))$, which is independent of the size of the telescope, plotted as a function of the mean photon counts per speckle $\overline{n}$ and the normalized spatial frequency x. Since $\sqrt{n_s} = \frac{D}{r_c}$, the signal-to-noise ratio for a telescope of diameter of D meters can be obtained by lowering the whole plot by $\log(\frac{D}{r_c})$ as indicated by the arrow.

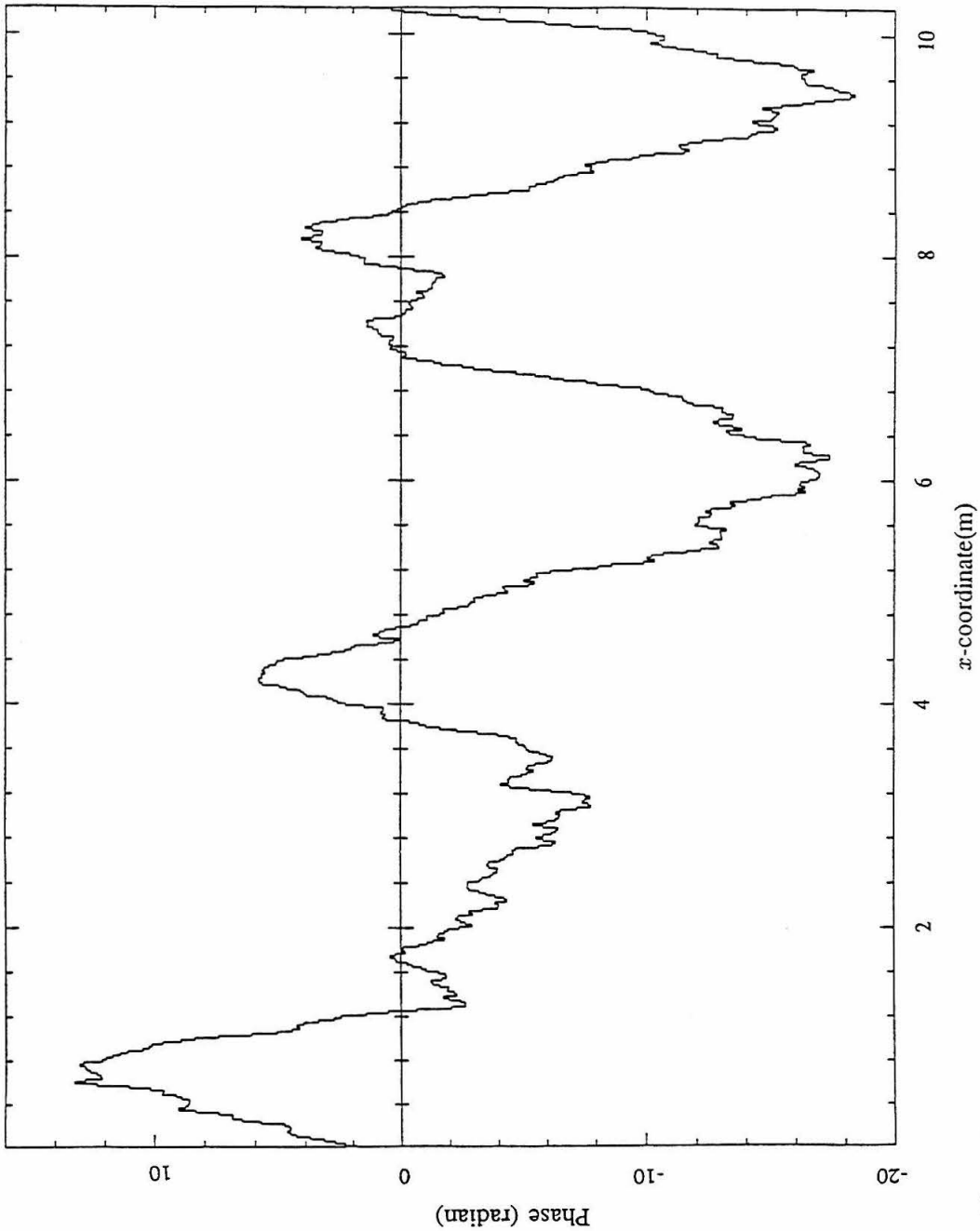

Figure 1

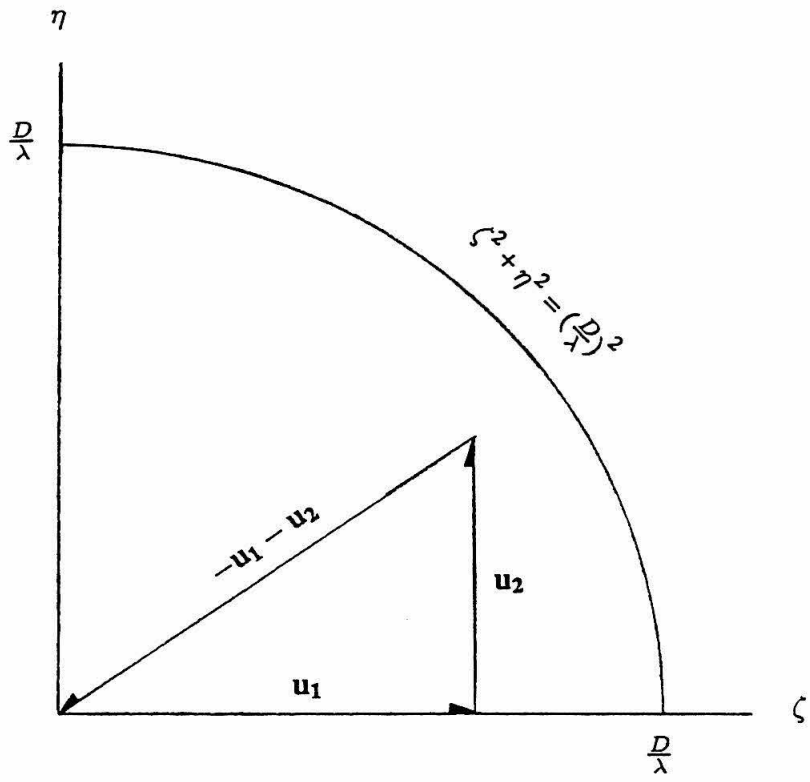

Figure 2

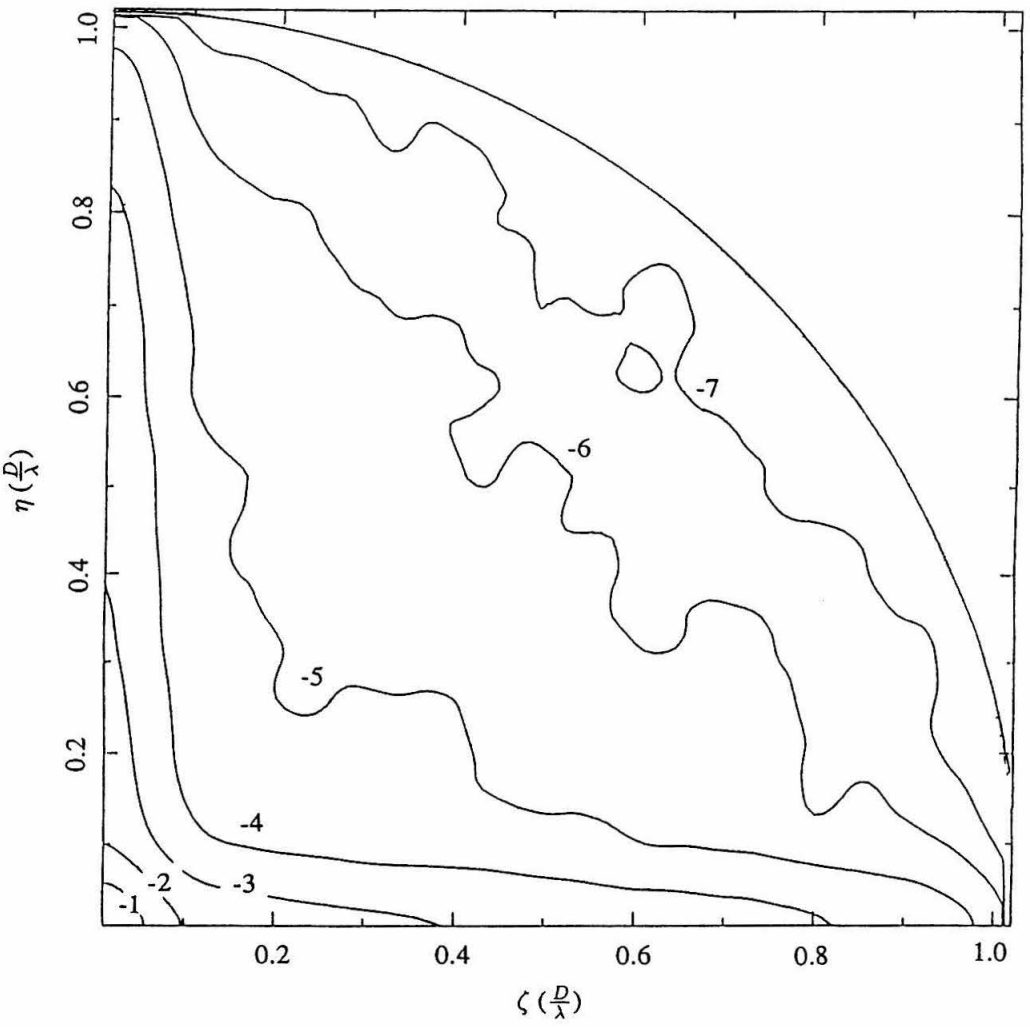
1.0
0.8
0.6
0.4
0.2
$\eta(\frac{D}{\lambda})$
-7
-6
-5
-4
-3
-2
-1
0.2
0.4
0.6
0.8
1.0
$\zeta(\frac{D}{\lambda})$

Figure 3

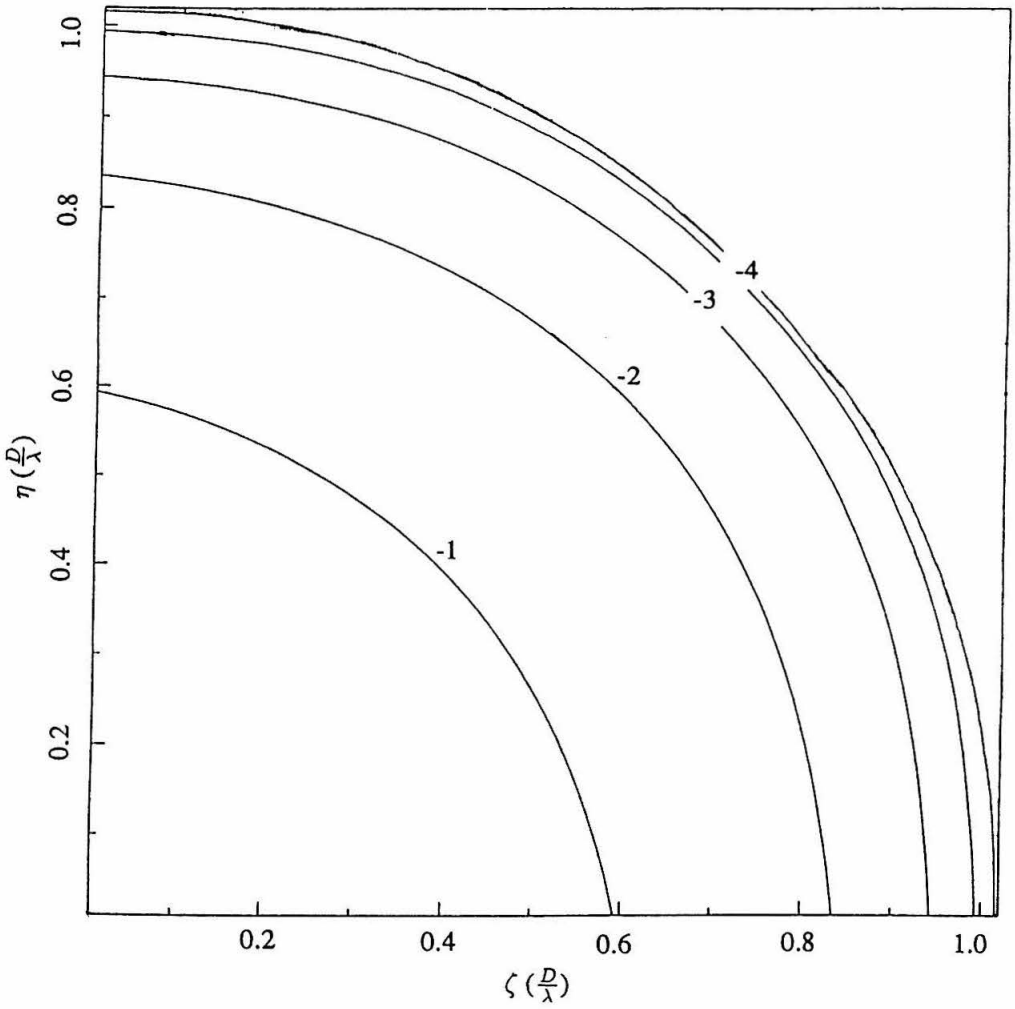
1.0
0.8
0.6
0.4
0.2
$\eta\,(\frac{D}{\lambda})$
-4
-3
-2
-1
0.2
0.4
0.6
0.8
1.0
$\zeta\,(\frac{D}{\lambda})$

Figure 4

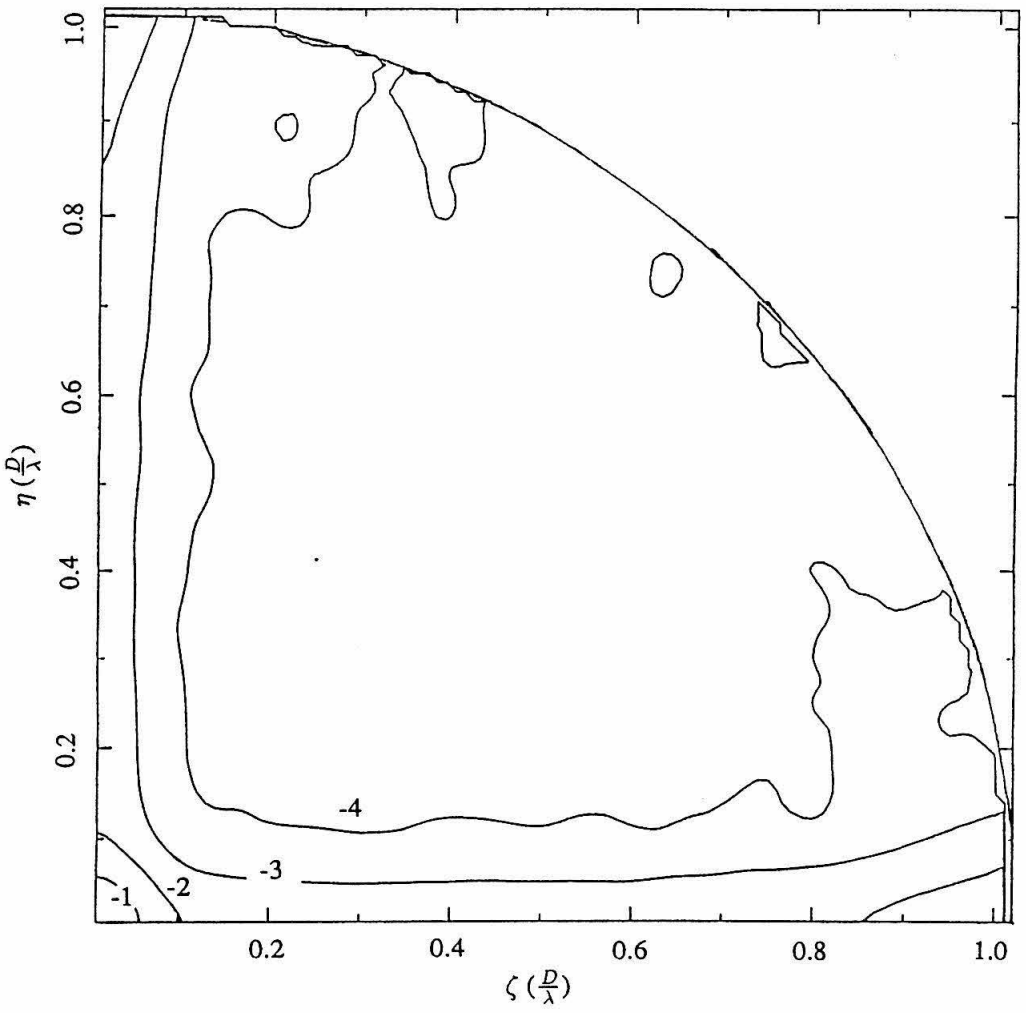
1.0
0.8
0.6
0.4
0.2
$\eta\,(\frac{D}{\lambda})$
-4
-3
-2
-1
0.2
0.4
0.6
0.8
1.0
$\zeta\,(\frac{D}{\lambda})$

Figure 5

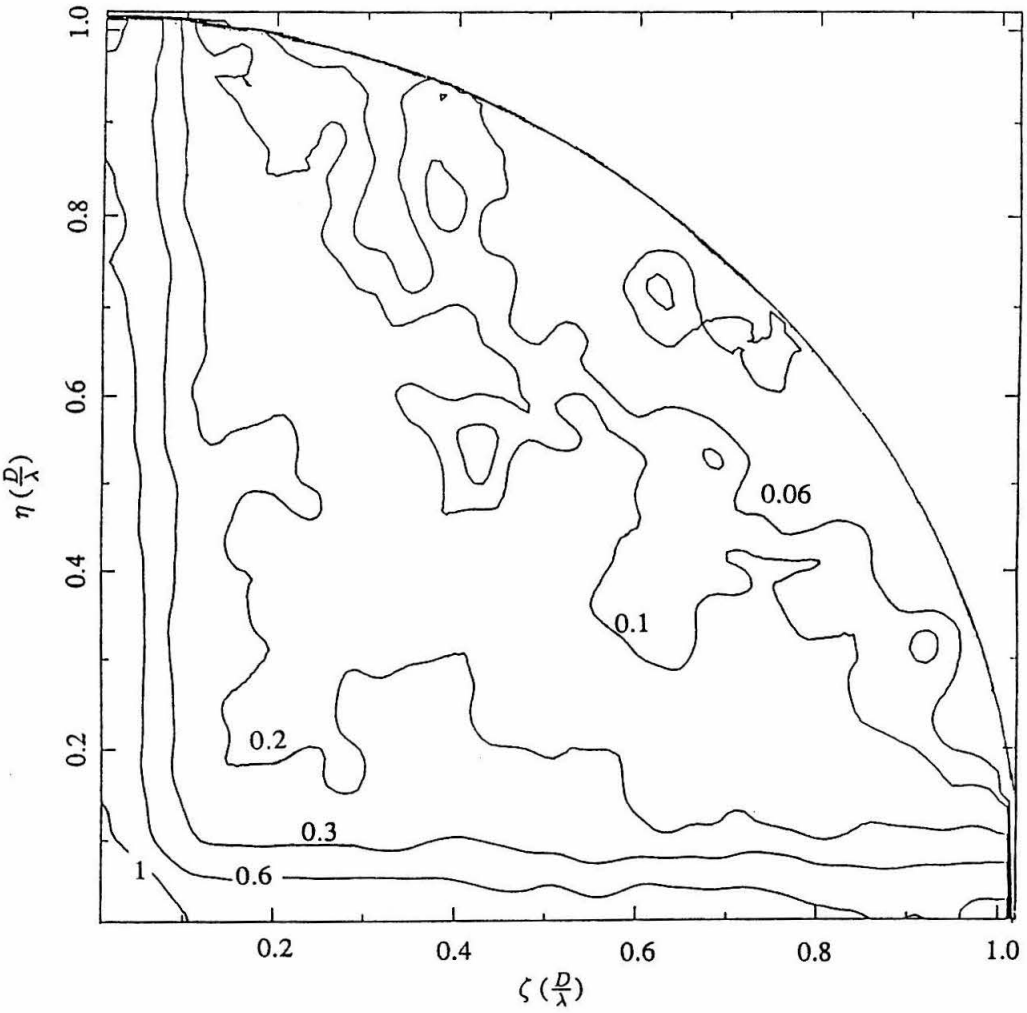
1.0
0.8
0.6
0.4
0.2
$\eta(\frac{D}{\lambda})$
0.06
0.1
0.2
0.3
0.6
1
0.2
0.4
0.6
0.8
1.0
$\zeta(\frac{D}{\lambda})$

Figure 6

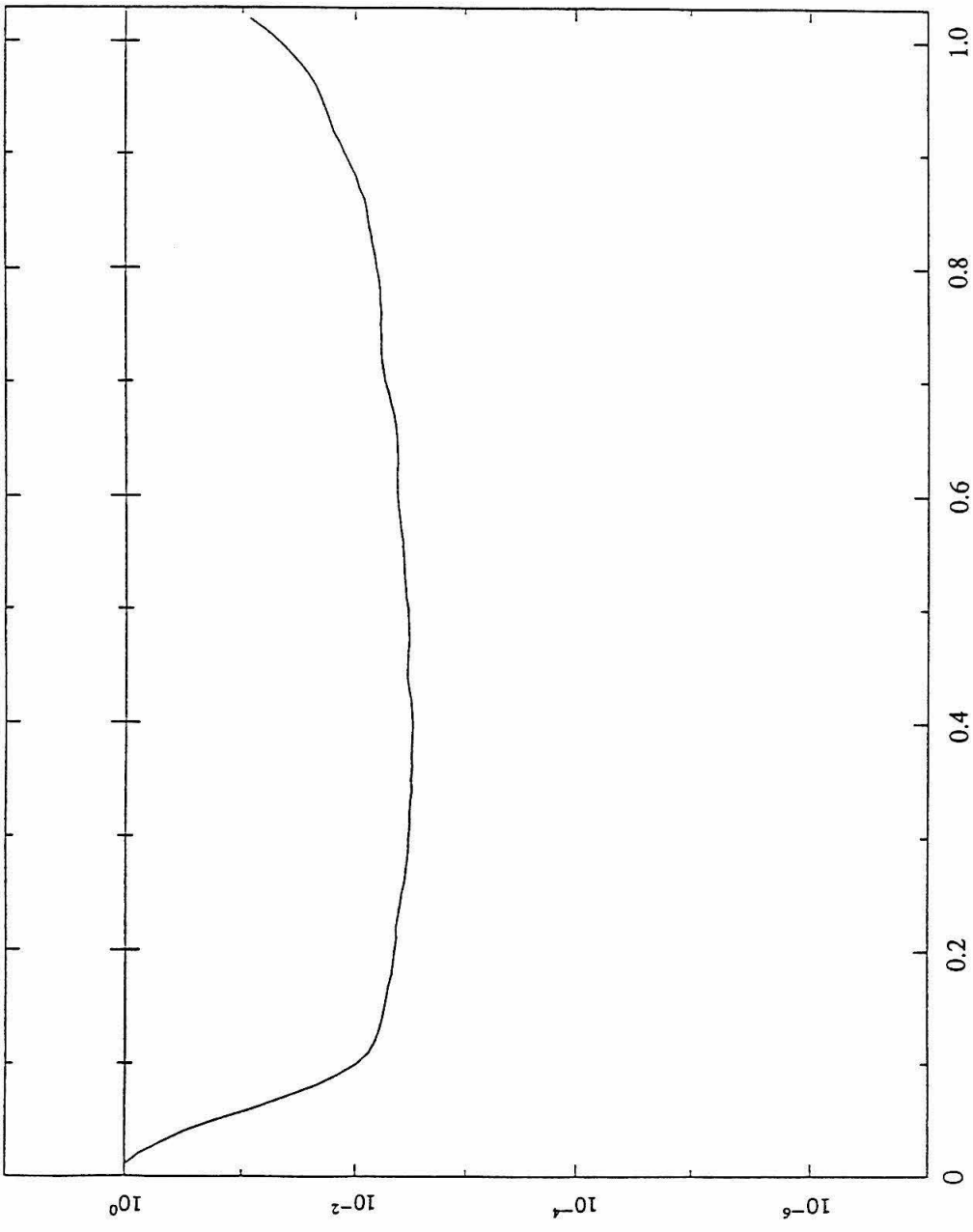
10^0
10^{-2}
10^{-4}
10^{-6}
0
0.2
0.4
0.6
0.8
1.0
Radial Spatial Frequency $|\mathbf{u}|$ $(\frac{D}{\lambda})$

Figure 7

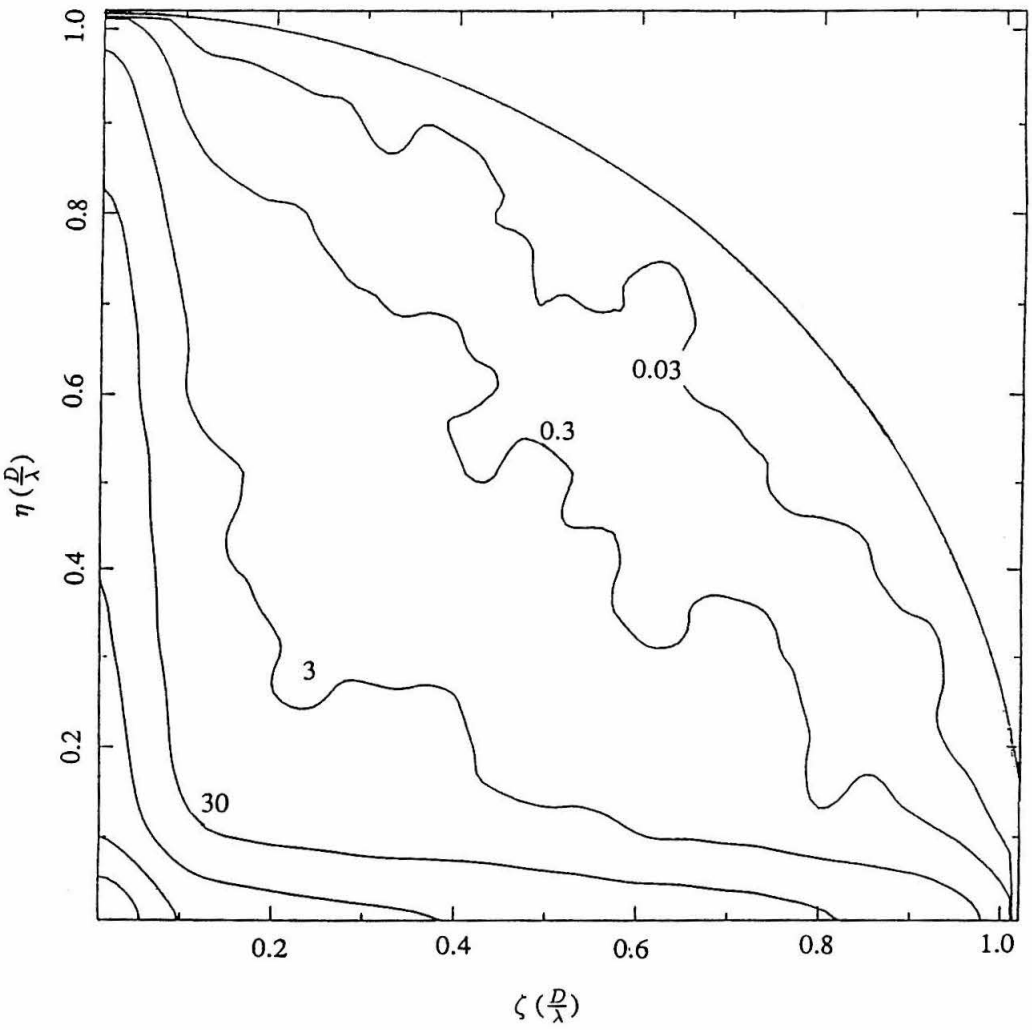

Figure 8

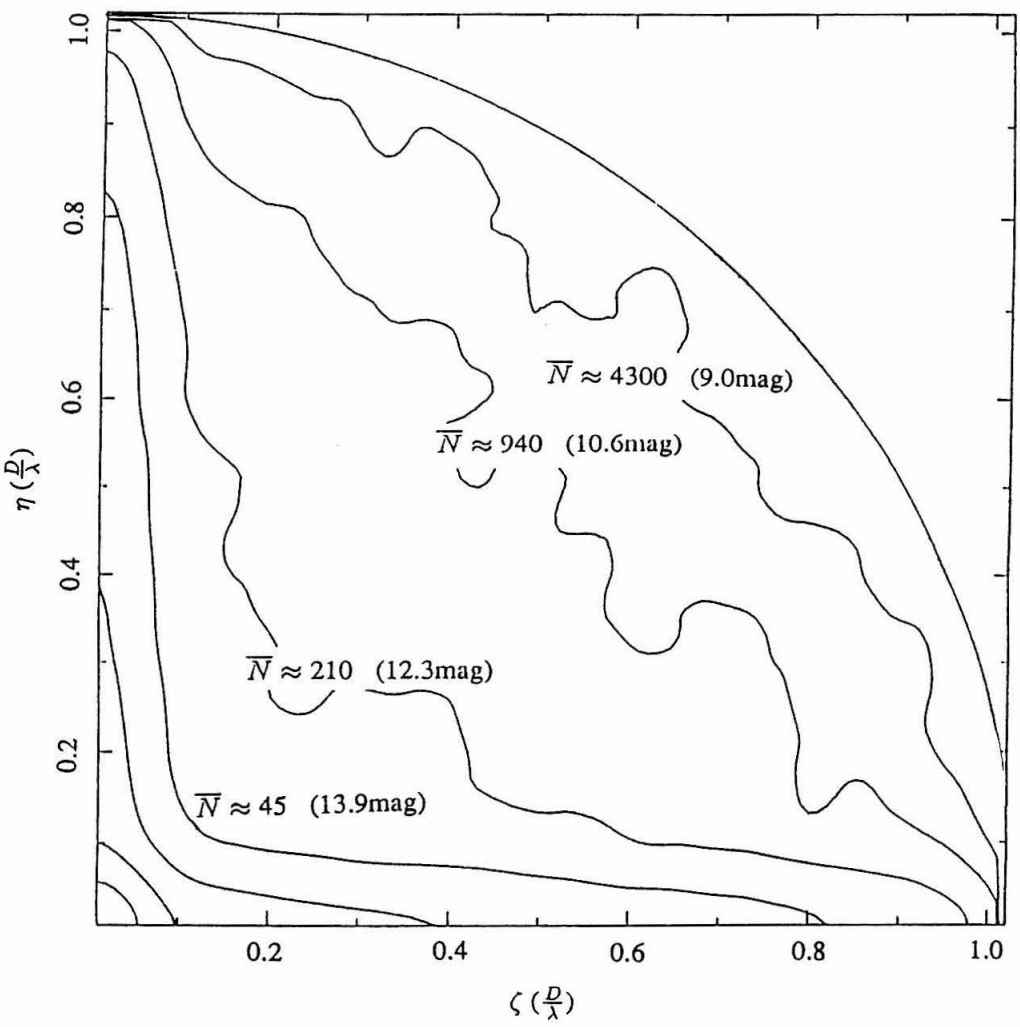
$\overline{N} \approx 4300$ (9.0mag)
$\overline{N} \approx 940$ (10.6mag)
$\overline{N} \approx 210$ (12.3mag)
$\overline{N} \approx 45$ (13.9mag)
$\eta\,(\frac{D}{\lambda})$
$\zeta\,(\frac{D}{\lambda})$
0.2
0.4
0.6
0.8
1.0

Figure 9

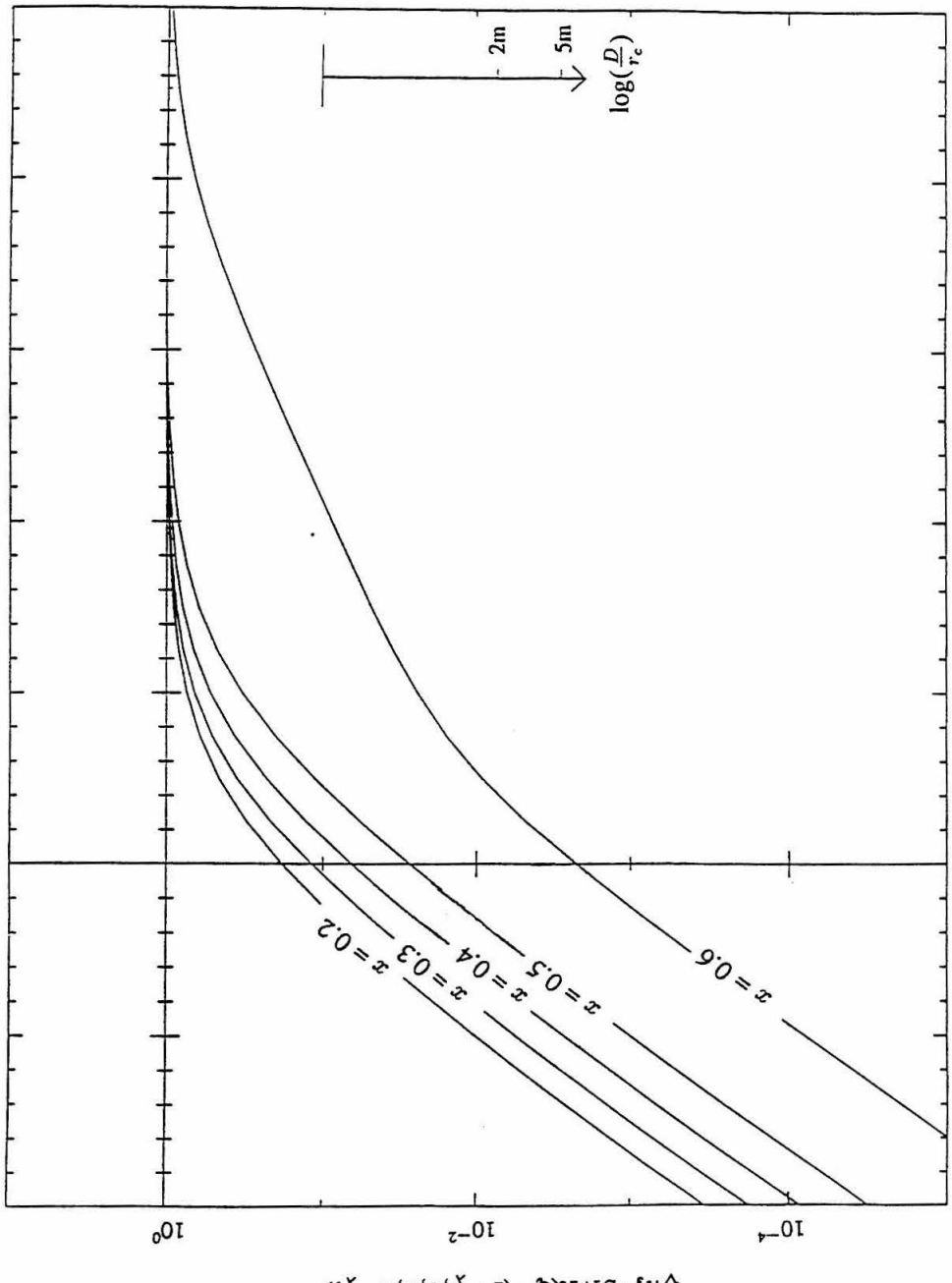
$\sqrt{n_s} \cdot SNR(\hat{Q}^{(2)}(x \cdot \frac{D}{\lambda}, 0, 0, x \cdot \frac{D}{\lambda}))$
10^0
10^{-2}
10^{-4}
10^{-2}
10^{-1}
10^0
10^1
10^2
10^3
10^4
10^5
Mean Photon Counts per Speckle
$\log(\frac{D}{r_c})$
2m
5m
$x = 0.6$
$x = 0.5$
$x = 0.4$
$x = 0.3$
$x = 0.2$

Figure 10