# CHAPTER 5

## "Edge" effects in parity detection using triple correlation

### 5.1 Motivation

In chapter 2 we considered the SNR for detection of parity of a binary using the triple correlation. An idealized model of a speckle pattern was used there to simplify calculations. In particular statistical assumptions were made about the focal plane intensity correlations: this reduced the order of correlations involved by a factor of two. One could do with intensity correlations upto the sixth order rather than field correlations upto the twelfth order. The resulting simplification allowed us to calculate SNR for general light levels. The idealization of the PSF, by virtue of its simplicity, neglects certain effects which deserve justification on the basis of more detailed field calculations. For example, in chapter 2 the speckles were (somewhat artificially) confined to a pixel and speckles in different pixels were treated as statistically independent. Also neglected were the "edge" effects due to the finite size of the seeing disk.

In this chapter we estimate the effects neglected in chapter 2. The procedure followed is similar to that of chapter 4. We first obtain the focal plane triple correlation for a point source. This PSF triple correlation (PSFTC) contains features with different statistical origins and properties. Like the general PSF double correlation the PSF triple correlation is inversion symmetric about its center. The binary triple correlation is then written as a combination of PSF triple

correlation units with different strengths and locations. The asymmetric distribution of their strengths results in asymmetry in the triple correlation for the binary. As in the previous chapter we ask which of the PSFTC features yeilds better SNR for parity detection. The answer involves calculation of the SNR for a particular choise of weight function which emphasizes a particular feature. As already mentioned in chapter 4 "low flux" may mean different magnitudes for different weight functions. To decide what one means by low flux one has to calculate the variance for a given weight function. For a triple correlation this necessitates calculation of field correlations upto the twelfth order. So we take a shortcut here. First of all we let the binary shrink to a point. This is justified, for the noise considerations, for the following reason. For binaries with significant difference in the component brightness, the noise is basically due to the brighter star. In general letting the binary shrink to point neglects terms in the noise $b/6$ times smaller than that for a point source. A second simplifying factor is the use of the speckle model described in section 2.1. Note that both these simplifications apply only to the noise calculations and not to the signal calculation. The signal for any weight function is sixth order in the fields and is evaluated exactly for the weight functions considered. The reason for the difference in treatement for the signal and noise is that the PSF triple correlation is inversion symmetric about the flux center and therefore it is necessary to retain binary nature in computing the parity signal.

The effects considered in this chapter are analogous to the "edge" effects discussed in the previous chapter. In contrast to the double correlation method of parity detection, where the "edge" effects were entirely responsible for parity detection, the triple correlation method does not so crucially depend on such "edge" effects. The basic difference is again traced to the fact that the triple correlation method respects stationarity in the focal plane while the double correlation method (for parity) does not. The main result of this chapter is to show that the "edge" effects in the case of triple correlation method too have a small parameter $b/6$ in the SNR associated with them which makes them negligible in comparison to the leading spatially stationary effect considered in chapter 2. In reality , the multiscale turbulence in the earth's atmosphere results in multiscale correlations in the pupil plane fields. In this chapter, however, we deal only with single scale Gaussian correlations for analytical simplicity. Though the results are expected to be somewhat different from that for a Kolmogorov spectrum for the turbulence this simpler single scale model does contain the essential edge effects. Below we give qualitative (order of magnitude) scaling estimates of various contributions; details are left to the Appendices.

5.2 The triple correlation for a binary

The triple correlation, as used in the speckle masking method, for a point source

$$T_R(Y,Z) = \int d^2X \langle R(X)\,R(X+Y)\,R(X+Z)\rangle \qquad (5.1)$$

is defined on the four dimensional space $(Y_1, Y_2, Z_1, Z_2)$. We

emphasise that in this chapter we are dealing with focal plane triple correlation unlike in the chapter 1 where we considered the bispectrum. As the name suggests it correlates intensities at three points. We have shown this schematically in Fig 5.1 (page 5-32). In the $(Y_1, Z_1)$ plane ($(Y_2, Z_2)$ is similar) there are five features of this triple correlation. The feature A is a plateau region with extension of the order of the seeing disk. Then there are three ridges B, C and D with width of the order of speckle size and extension of the order of seeing disk. In the full four dimensions these are actually planar layers: for example the feature B is the plane $Y_1 \sim Z_1$ and $Y_2 \sim Z_2$, the layered nature comes because the equalities are not exact but allow for a tolerance of the order of the speckle size. The feature E is located at the center of the four-space and has a size of the order of the speckle size. In our Gaussian model all these features have Gaussian fall-offs with above mentioned length scales. At the origin of the four space the feature E is twice as strong as other features (which have equal strengths at the origin). The st'rength at the origin can be readily estimated. With our normalization $\pi N_0 R^2$ is the total photon count per exposure for a zeroth magnitude star. This flux is spread all over the seeing disk of size $\delta = f/k\ell$. Thus the average focal plane intensity density is $\langle R \rangle = \frac{N_0 k^2 \ell^2 R^2}{f^2}$. For far away points in the seeing disk the three intensities entering the triple correlation will be statistically independent so the scaling estimate for the triple correlation density for a typical triple correlation element within a seeing size is

$$\int d^2x \langle R \rangle^3 \sim (f/k\ell)^2 (N_0 k^2 \ell^2 R^2/f^2)^3 = N_0^3 k^4 \ell^4 R^6/f^4 \qquad (5.2)$$

This is precisely the term A arising from uncorrelated fluctuations in the three PSFs in the triple correlation. The ridges B, C and D appear because in this region of the triple correlation two (and only two) of the three points involved in the triple correlation are within a speckle size and thus have correlated intensity fluctuations. Thus in addition to the all pervading feature A there is enhancement of the triple correlation. Finally the .term E with weight two comes about because all the three points involved in triple correlation are within a speckle size. Note that the number of terms (counting E twice) is a reflection of the asymptotic Rayleigh distribution assumed in chapter 2. In those regions where some of these terms are equal (even though they might have fallen off considerably from their peak values at the center) we see that for general element only the A term, for ridges two terms, and near the center all six terms contribute. This is indicative of the Rayleigh asymptotic $\langle \mu \rangle = \langle \mu \rangle$ ; $\langle \mu^2 \rangle = 2\langle \mu \rangle^2$ ; $\langle \mu^3 \rangle = 6\langle \mu \rangle^3$ where $\mu$ is a Rayleigh variable with mean $\langle \mu \rangle$. This is the time to contemplate the Rayleigh approximation used in the second chapter. The deviations from the Rayleigh statistics show themselves in the following way: consider the ridges for example. At the center of the ridges the two terms (one is A and the other is either B,C, or D) are exactly equal giving the Rayleigh signature 2. But normal to the plane the features B, C, and D have much rapid fall off than the A feature. Similarly as one moves away from the center of the PSF triple correlation different features have different fall-off lengths.

However, the extensions of the correlation ridges are equal to that of the feature A. Deviations from the Rayleigh distribution used in chapter 2 are seen here to arise from different fall-off lengths of various features. The deviations, for binary separations smaller than the seeing disk, are negligible and can be absorbed in redefined paramters like the number of speckles etc. or alternatively by a numerical factor in the expression for the SNR. Our calculations for the double (previous chapter) and triple correlation (present chapter) show explicitly the origin and the limitations of the Rayleigh statistics assumed in the PSF model of section 2.1. The field calculations reported in this chapter give the same scalings of the SNR with $N_s$ and $\mathcal{N}$ as in the second chapter.

Let us call all the features (A,B,C,D,2E) collectively a unit of triple correlation (short for triple correlation transfer function). For a binary the intensity I(X) is given by

$$I(X) = \alpha_1 R(X) + \alpha_2 R(X-b) \qquad (5.3)$$

where b is the binary separation, $\alpha_1$ and $\alpha_2$ represent the relative strengths of the component stars. We align the focal plane axes such that the binary is along the first axis with subscript 1. The triple correlation $T_B$ for the binary can be written in terms of the same $T_R$ for the point source:

$$T_B(Y,Z) = (\alpha_1^3 + \alpha_2^3)\, T_R(Y,Z) + \alpha_1^2\alpha_2 [T_R(Y,Z-b) + T_R(Y-b,Z) + T_R(Y+b,Z+b)] + \alpha_1\alpha_2^2 [T_R(Y,Z+b) + T_R(Y+b,Z) + T_R(Y-b,Z-b)] \qquad (5.4)$$

It contains seven basic units with strengths and locations shown in Fig 5.2 (page 5.31). Note that this is just a convolution of

the basic triple correlation unit (Fig 5.1) and the triple correlation for the binary in the absence of atmospheric noise (Fig 2.6). Since the strengths of different units are asymmetric the triple correlation is also asymmetric and one has to extract the asymmetry in the best possible way. This leads to the interesting question of finding those regions of the four space which give better SNR for parity. Perhaps a more interesting question is to ask what is the relative importance of various features of the basic unit in determining parity. In chapter 2 all the emphasis was on the feature E i.e. the central region in the triple correlation transfer function. Since various features have negligible overlap volumewise, a superposition approximation (see section 5.6) for the SNR holds in that once we know what effect each kind of feature has one can figure out what weight function to use to get better SNR. So we consider following three cases: a) SNR due to the central E feature alone (this is the chapter 2 case), b) the SNR due to the correlation ridges B, C and D, c) the SNR due to the plateau feature A (long exposure case!).

## 5.3 Parity detection due to the feature E

In the triple correlation for the binary the feature E appears at seven places. The central one has a strength $\alpha_1^3 + \alpha_2^3$ which is symmetric in the fluxes of the component stars and thus does not contain any parity information. As discussed in chapter 2 only two of the remaining six locations have independent information, We choose the E-features of the first

and the fourth PSFTC to obtain an antisymmetric parity statistic. So our weight function of the first kind is shown in Fig 5.3. It has two ports where it takes nonzero values and these ports have four volume equal to the E feature. Note that though the weight function is mainly desinged to pick up the contribution due to the E feature it also gets a contribution from others. Label the seven PSFTCs as shown in Fig 5.3. We can then denote the contribution of a specific feature due to a specific unit: for example $D_3$ means the D feature of the third unit which passes through the port 2. The overlap of any of the features A-D with the weight function results in an equivalent E feature because the weight function has the four volume equal to the feature E. This is true as long as the binary separation is much smaller than the seeing. Otherwise the features A-D would have fallen off considerably at the location of the two ports and the overlap would be weaker than the E-features located at the ports. These equivalent E-features contributing to the ports 1 and 2 and their strengths are given in table 5.1.

For the parity statistic which is of the form

$$\int d^2y\, d^2z\; W(y,z)\, T_B(y,z) \qquad (5.5a)$$

the variance is given by (low flux)

$$2\int d^2y\, d^2z\; W^2(y,z)\, T_B(y,z) \qquad (5.5b)$$

This result follows from the general results on Poisson fluctuations derived in chapter 3 and the special nature of the weight function. An application of the diagramatic rule for the variance yields a total of 33 terms for the averaged squared modulus of a general triple correlation: a sixth order term, nine fifth order terms, eighteen fourth order terms and six third

order terms. Because of the symmetry of the weight function W and the triple correlation $T_B(Y,Z)$ the six lowest order terms can be paired. This explains the factor two in the above expression. For a general weight function there are two more terms in this lowest order in intensities. If we had made no use of the fact that only two of the peripheral six elements are independent and blindly used all the six asymmetric terms then these other two terms in the variance expression mentioned above would have made sure that the variance is nine times as large thus leaving SNR unaffected (see Appendix A5). To get the signal we weigh the port 1 by +1 and the port 2 by -1 and add. We must also multiply by the four volume of the order $(f/kR)^4$ ((speckle size)$^4$). Thus the parity signal due to the E-feature is

$$S_E \sim \underbrace{2\alpha_1\alpha_2(\alpha_1-\alpha_2)}_{\substack{\text{Port 1}\\ \text{minus}\\ \text{Port 2}}} \underbrace{(f/kR)^4}_{\substack{\text{four volume}\\ \text{of the ports}}} \underbrace{(N_0^3 k^4 \ell^4 R^6/f^4)}_{\substack{\text{triple correlation}\\ \text{density Eq 5.2}}} = 2\alpha_1\alpha_2(\alpha_1-\alpha_2)N_0^3\ell^4R^2$$

$$\sim 2\mathcal{N}_1\mathcal{N}_2(\mathcal{N}_1-\mathcal{N}_2)N_S \qquad (5.6)$$

where $\mathcal{N}=\frac{\alpha N_0 \ell^2}{16}$ is the count per speckle as in chapter 2 and $N_S$ is the average number of speckles, The variance is

$$V_E \sim \underbrace{2}_{\substack{\uparrow\\ \text{Eq 5.4}}} \underbrace{(4\alpha_1^3+14\alpha_1^2\alpha_2+14\alpha_1\alpha_2^2+4\alpha_2^3)}_{\text{Port 1 + Port 2}} \underbrace{(f/kR)^4}_{\substack{\text{four volume}\\ \text{of the ports}}} \underbrace{(N_0^3 k^4\ell^4R^6/f^4)}_{\substack{\text{triple correlation}\\ \text{density Eq 5.2}}}$$

$$= 4(2\alpha_1^3+7\alpha_1^2\alpha_2+7\alpha_1\alpha_2^2+2\alpha_2^3)N_0^3\ell^2R^2$$

The approximation is in using order of scaling estimates for the four volume of the E-feature and the triple correlation density. Thus the SNR for parity detection using E feature alone is

$$SNR_{TC(E)} = \frac{4}{3\sqrt{\pi}}\, q^{3/2}\, M^{1/2}\, N_s^{1/2}\, \frac{\mathcal{N}_1\,\mathcal{N}_2\,(\mathcal{N}_1-\mathcal{N}_2)}{\left[2\mathcal{N}_1^3+7\mathcal{N}_1^2\mathcal{N}_2+7\mathcal{N}_1\mathcal{N}_2^2+2\mathcal{N}_1^3\right]} \qquad (5.8)$$

The scaling is same as that derived in chapter 2 (Eq 2.15). The numerical coefficient is provided by the appendix A5 which evaluates quantities of interest exactly. This derivation gives us an explanation of the numerical coefficients as arising from that many terms. Also as the binary separation increases we can see that the contribution to the noise from the E feature remains constant but that due to all other terms falls off. The fall-off scale is, however, comparable to the seeing. We, therefore, conclude that the SNR due to E feature is weakly dependent on the binary separation.

## 5.4 SNR for parity detection due to correlation ridges

Fig 5.2 shows regions of the $(Y_i, Z_i)$ plane where correlation ridges exist. For clarity the ridges are shown slightly displaced, in realty ridges like $D_1$ and $D_6$, for example, coincide in the region common to both. We see nine strips in this plane where correlation ridges from two or three triple correlation units overlap. Because of the symmetry of the triple correlation only two (to be precise 1.5) of these nine strips contain independent information. To see this consider a strip with equation $Z = Y + C_o$. Then from the definition of the triple correlation the same strip appears at $Z=C_o$ and at $Y=C_o$:

$$\int d^2x\, I(x)I(x+Y)I(x+Y+C_o) = \int d^2x\, I(x-Y)I(x)I(x+C_o) \qquad (5.9)$$

i.e. if the coordinates on the first strip is $(Y, Z=Y+C_o)$ then the

same event is also written at $(-Y, Z=C_0)$. Fig 5.2 shows three strips $S_1$, $S_a$ and $S_3$ and their two copies each (denoted with a single and double primes). Note that the strips $S_1$, $S_2$ and $S_3$ have the same orientation as the B feature. In fact, for the PSFTC shown in Fig 5.1 only one of the features B, C and D contains independent information. Further reduction in the number of ridges is due to the symmetry of the triple correlation under the interchange of Y and Z. Because of this symmetry the strips 1 and 3 are identical for every realization. In the following we have used all the three strips instead of using any one of the strips 1 and 3 and half of the second strip split lengthwise. A Similar situation arose in the case of the double correlation analysis (section 4.3) where it was shown that both these schemes give the same SNR. The reasoning of section 4.3 is applicable here so we use all the three strips with equal relative weights. Although the weight function is designed to emphasise B type feature it will also pick up contribution from the feature A. The weight function will also pick up contribution due to the feature E. For the reasons given later in section 5.6 we need not consider it here. In Fig 5.4a we show strips 1, 2 and 3 in greater detail. The centers and the strengths of these features are shown in the figure. We are interested in the leading contribution in powers of binary separation b. Consider the strip 1 for example. The unit 1 at $(-b,-b)$ contributes to this strip via the feature $A_1$. Now since the binary separation is assumed small one can think of the contribution due to feature A on the strip to have effective center $B_1'$ shown in the figure. In fact this is the center of the overlap between the feature $A_1$ and the strip. The overlap is

equivalent to a feature **B**. In Fig **5.4a** the centers of such effective **B** features are shown by smaller dots and these are linked to the centers of the original A features by dotted lines. In Fig **5.4b** we see the centers and the strengths of such 'effective B' features. We are interested in the asymmetry in these three strips. Because of the Y↔Z symmetry of the triple correlation the lower half of the strip 1 is equivalent to the lower half of the strip 3 and so on. Thus asymmetry about the origin is equivalent to asymmetry of the individual strips about the $-45^{o}$ line passing through the center. For the central strip $S_2$ both these things mean the same. For the strips $S_1$ and $S_2$ the diagonally opposite parts, with respect to the origin, of the two strips can be replaced by corresponding part of the same strip. Fig **5.4c** shows the effect of the special weight function which is situated at a distance **x** away from the center of the effective feature. Only contribution from the central region between $\pm x$ survives. So for a weight function of this kind contribution depends linearly on the distance of the jump in the weight function from the center of the feature under consideration. It turns out that strips 1 and 3, which are identical have their centers shifted from the $-45^{o}$ line in the sense opposite to that of the middle strip which has its center shifted twice as much. This motivates our weight function shown in Fig **5.4d.** The asymmery in the triple correlation can be readily estimated when we note that the shift in the centers of the strips are of the order b. As described in appendix **B5** we get the signal due to the strips 1,2 and 3:

$$S_1 = S_3 = \tfrac{1}{2} S_2$$

$$S_2 = \frac{2}{\sqrt{3}\,\pi^{1.5}}\, N_S^{3/2}\, q^3\, \mathcal{N}_1 \mathcal{N}_2 (\mathcal{N}_1 - \mathcal{N}_2)\, \mathcal{B} \qquad (5.10)$$

For this weight function the variance calculation needs to be done to all orders to decide the low and high flux regimes. This is done in appendix B5 using the PSF model of section 2.1. The result is

$$V_B = 24 q^6 N_S^3 (\mathcal{N}_1 + \mathcal{N}_2)^6 + 4 q^5 N_S^3 (\mathcal{N}_1 + \mathcal{N}_2)^5 + \tfrac{4}{3} q^4 N_S^3 (\mathcal{N}_1 + \mathcal{N}_2)^4 + 4 q^3 N_S^2 (\mathcal{N}_1 + \mathcal{N}_2)^3 \qquad (5.11)$$

We see that the transition to low flux goes through two, steps, The first transition occurs for objects fainter than $13^m$. The fourth order contribution takes over from the sixth order contribution which dominates at high light levels. The second transition occurs for objects fainter than $18^m$. The third order contribution dominates for lower light levels. The first occurs when a speckle receives less than one photon in an exposure: $18 q^2 \mathcal{N}^2 = 1$ and the second when all speckles together receive less than one phton per exposure: $q N_S \mathcal{N} = 3$. The SNR for parity detection using these correlation ridges is

$$SNR_B = \frac{\sqrt{2}}{3\pi\sqrt{\pi}}\, M^{1/2}\, \frac{\mathcal{N}_1 \mathcal{N}_2 (\mathcal{N}_1 - \mathcal{N}_2)}{(\mathcal{N}_1 + \mathcal{N}_2)^3}\, \mathcal{B} \qquad \text{high flux} \qquad (5.12a)$$

$$\downarrow 12^m$$

$$= \frac{2}{\pi\sqrt{\pi}}\, q M^{1/2}\, \frac{\mathcal{N}_1 \mathcal{N}_2 (\mathcal{N}_1 - \mathcal{N}_2)}{(\mathcal{N}_1 + \mathcal{N}_2)^2}\, \mathcal{B} \qquad \text{medium flux} \qquad (5.12b)$$

$$\downarrow 17.5^m$$

$$= \frac{2}{\pi\sqrt{3\pi}}\, q^{3/2} M^{1/2} N_S^{1/2}\, \frac{\mathcal{N}_1 \mathcal{N}_2 (\mathcal{N}_1 - \mathcal{N}_2)}{(\mathcal{N}_1 + \mathcal{N}_2)^{1.5}}\, \mathcal{B} \qquad \text{low flux} \qquad (5.12c)$$

We note that for the objects of present day interest (brighter than 18'th magnitude) and for close binaries the SNR due to the correlation ridges is never better than the SNR due to the main triple correlation feature E (the feature extensively discussed

in chapter 2). More specifically the following table 5.2 gives the binary separation, as a function of the magnitude of the brighter star, beyond which parity is better determined using the correlation ridges.

Table 5.2 Binary separation beyond which parity is better determined using the correlation ridges

| m | 13 | 14 | 15 | 16 | 17 |
|---|---|---|---|---|---|
| B | 27 | 12 | 7 | 5 | 3 |

We also note that in the medium flux regime this SNR (Eq 5.12b) has the same scaling as the SNR Eq 4.12 for parity detection using feature B of double correlation. Although more realistic models for the speckle phenomena and fine tuning of the weight functions need to be considered before one can compare numerical factors in these two expressions with a degree of confidence it appears that of these two schemes the double correlation method may have a slightly better SNR than the triple correlation method using correlation ridges.

## 5.5 SNR for parity detection using the feature A

For $+45^{o}$ strips (parallel to the B-feature) other than the correlation ridges the asymmetry starts in the cubic order in the separation. If we approximate the near center strength of the effective features by its peak value then for a general +45 strip contributions due to various triple correlation units cancel. One has to retain the quadratic near peak variation to get any asymmetry. Other reason for cubic asymmetry is that when the binary separation is large the features B and B' (overlap of feature A and a $45^{o}$ strip) differ in the second order in the

separation. In Fig 5.5 we show a representative weight function of the third kind designed to emphasis the asymmetry due the feature A. A straightforward but somewhat lengthy calculation gives the signal for this weight function

$$S_A = \frac{\pi^{2.5}}{3^{1.5}\,2^{5.5}}\, N_0^3 \alpha_1\alpha_2(\alpha_1-\alpha_2)\, k^3 l^3 b^3 R^6/f \qquad (5.13)$$

The variance,in principle, needs to be evaluated upto all orders. However, contribution to the variance is positive in every order so even if we consider only the third order contribtution to the variance:

$$V_A \geq \frac{3\pi^2}{2} N_0^3(\alpha_1+\alpha_2)^3 R^6 \qquad (5.14)$$

one gets an upper bound

$$SNR_A \leq \frac{16\, q^{3/2} m^{1/2} N_1 N_2 (N_1-N_2)}{9\,(N_1+N_2)^{3/2}} \qquad (5.15)$$

on SNR. We note that even this upper bound is no match to the SNR due to the feature E for small binary separations. For large separations parity is better determined by using long exposure image (appendix C5). Arguments, similar to those given in section 4.4, can be given to show the (noisy) equivalence of the feature A and long exposure image. We, therefore, conclude that although for large binary separations the parity may be better determined by the feature A it cannot be regarded as a genuine high resolution feature.

## 5.6 A discussion

A superposition approximation: So far we considered SNR for individual features like the central E-feature, the ridges and the feature A. We mentioned that a kind of superposition

approximation enables one to consider these features independently. The reason is the great difference in the four volumes of various features. To be specific, if we take the four volume of the E-feature (which is of the order $\rho^4$) as unity then the four volumes of the ridges and the feature A are $N_s$ and $N_s^2$ respectively. This is because the ridges have an extension of the order of the seeing disk (which is $N_s^{1/2}$ times the speckle size) in two dimensions while the feature A has the same extension in four dimensions. Now let us go back to SNR due to the feature E. The weight function used in this case had four volume equal to that of the E feature. The parity signal is purely due to the E feature and the noise comes from other features. For this reason one is justified in attributing the parity signal to the E feature. Next, consider the weight function of the second kind used to emphasise the feature B. In this case the signal not only comes from the feature B but also from the feature E. There are two reasons why we did not consider the contribution due to this feature in section 5.4. The first reason is that it's contribution to the signal and the noise separately is negligible when compared to that of the feature B. To see this consider the signal first. For concreteness we consider very low fluxes where signal and variance are proportional to the 4-volumes involved. The contribution to the signal due to the feature E is proportional to its four volume. The B feature has $N_s$ times larger volume and although it suffers from "edge" effects which, as described in the section 5.4, subdue its contribution by a factor $N_s^{1/2}$ its signal is still $N_s^{1/2}$ times larger than that due to the E feature. Unfortunately the variance does not suffer from

any "edge" effects and is larger by a factor $N_s$. In this case of very low fluxes (fainter than $18^m$) the two feature will have comparable SNR for close binaries. For flux levels of interest in speckle interferometry (between $13^m$ and $18^m$) the SNR due to the E feature is larger. Even in this case one can arrive at the superposition approximation by redifining the weight function so that it excludes all E features. For example, we may leave holes in the weight function at the locations of the E features. This wont affect the previous estimates because the size of such holes is very much smaller than the volume of the entire weight function. For similar reasons one can omit features E and B while considering parity due to the feature A.

Fine tuning the weight function: In this chapter (also in the previous one) we considered a rather simple weight function. The weight function was either 0 or $\pm 1$. It is certainly possible to improve the SNR by a smoother variation of the weight function. Intuitively we would expect the optimum weight function to follow the signal. The weight function should be large where more signal comes from. However, we also know that within the fall off scales the signal is more or less uniform. Therefore, we do not expect the fine tuning of the weight function to change the scalings with $N_s$ and $\lambda$ although the numerical coefficient may change a bit.

Effect of the Kolmogorov spectrum for the atmospheric inhomogeneities: The effect of the larger scales (comparable to or larger than the telescope pupil) in the turbulence is to tilt the entire wavefront. The speckle pattern will have its

instantaneous centroid wandering. However, unlike the double correlation method discussd in the previous chapter, the triple correlation does not share this focal plane motion of the speckle pattern. This is because the triple correlation defined in Eq 5.4 is invariant under shifts in the focal plane.

APPENDIX A5

Here we give the details of the SNR calculations described qualitatively in the text. We are not concerned with correlation effects due to secondary Airy rings in a real telescope so we use Gaussian apodization in the pupil plane. The focal plane field $\psi(x)$ is related to the pupil plane field $\varphi(\xi)$ by

$$\psi(x) = N_0^{1/2} \frac{k e^{i\frac{kx^2}{2f}}}{2\pi f i} \int d^2\xi \; e^{i\frac{kx\xi}{f}} \varphi(\xi) e^{-\xi^2/2R^2} \quad ; \quad R(x) = \psi\psi^*(x) \qquad (5.16)$$

The PSF triple correlation is given by

$$\langle R(x)R(x+y)R(x+z)\rangle = \underset{\text{(A)}}{\langle R(x)\rangle\langle R(x+y)\rangle\langle R(x+z)\rangle} + \underset{\text{(B)}}{\langle R(x)\rangle |\langle \psi(x+y)\psi^*(x+z)\rangle|^2}$$
$$+ \underset{\text{(C)}}{\langle R(x+y)\rangle |\langle \psi(x)\psi^*(x+z)\rangle|^2} + \underset{\text{(D)}}{\langle R(x+z)\rangle |\langle \psi(x)\psi^*(x+y)\rangle|^2}$$
$$+ \underset{2\text{(E)}}{2\langle \psi(x)\psi^*(x+y)\rangle\langle \psi(x+y)\psi^*(x+z)\rangle\langle \psi(x+z)\psi^*(x)\rangle} \qquad (5.17)$$

The terms in this expression are the five featrues A, B, C, D and (twice) E, in that order, and explicitly show their statistical origin. We have modeled the atmospheric distortions by a single scale Gaussian correlation function (the same as in the previous chapter):

$$\langle \varphi(0)\varphi^*(\xi)\rangle = \exp\{-4\xi^2/\ell^2\}$$

$$\langle \psi(x)\psi^*(x+y)\rangle = \frac{N_0 k^2 \ell^2 R^2}{16 f^2} e^{-\frac{k^2R^2}{4f^2}y^2} e^{-\frac{k^2\ell^2}{64f^2}(2x+y)^2} \qquad R \gg \ell \qquad (5.18)$$

for the pupil plane fields which are assumed to be Gaussian random fields i.e. higher order correlations have been broken down in terms of second order correlations with the help of the well known pairing theorem for Gaussian random variables. Since the random pathlength deviation for individual rays is belived to be hudreds of wavelengths in the pairing theorem every pair must have one field and one conjugate field (one starred and one without). In principle, pairs like $\langle \psi_1 \psi_2 \rangle$ exists but are exponentially small: of the order of $e^{-\langle \phi^2 \rangle}$ where $\langle \phi^2 \rangle$ is of the order of $\sim 10^2$. Thus for the assumed statistics of the fields we get the average PSF triple correlation

$$\langle R(x)R(x+y)R(x+z)\rangle = \frac{N_0^3 k^6 \ell^6 R^6}{2^{12} f^2} \Big\{ e^{-\frac{k^2\ell^2}{16f^2}[x^2+(x+y)^2+(x+z)^2]} + e^{-\frac{k^2\ell^2}{16f^2}x^2} e^{-\frac{k^2R^2}{2f^2}(y-z)^2} e^{-\frac{k^2\ell^2}{32f^2}(2x+y+z)^2} + e^{-\frac{k^2}{2f^2}\left[\frac{\ell^2}{8}(x+y)^2+R^2z^2+\frac{\ell^2}{16}(2x+z)^2\right]} + e^{-\frac{k^2}{2f^2}\left[\frac{\ell^2}{8}(x+z)^2+R^2y^2+\frac{\ell^2}{16}(2x+y)^2\right]} + 2\, e^{-\frac{k^2R^2}{4f^2}[y^2+(y-z)^2+z^2]} e^{-\frac{k^2\ell^2}{64f^2}[(2x+y)^2+(2x+y+z)^2+(2x+z)^2]} \Big\} \qquad (5.19)$$

The parity statistics is of the form

$$P = \int d^2y\, d^2z\, W(y,z) \int d^2x \langle I(x) I(x+y) I(x+z)\rangle \qquad (5.20)$$

where the weight fuction $W$ is chosen to emphasisevarious terms of the triple correlation. Note that the weight function does not depend on X and therefore on the orgin of the focal plane. Integration over X yields the PSFTC $T_R(y,z) = \int d^2x \langle R(x)R(x+y)R(x+z)\rangle$

$$T_R(y,z) = \frac{N_0^3 k^4 \ell^4 R^6}{3\, 2^8 f^4} \Big\{ e^{-\frac{k^2\ell^2}{24f^2}(y^2+z^2-yz)} + e^{-\frac{k^2R^2}{2f^2}(y-z)^2} e^{-\frac{k^2\ell^2}{96f^2}(y+z)^2} + e^{-\frac{k^2R^2}{2f^2}z^2} e^{-\frac{k^2\ell^2}{96f^2}(4y^2+z^2-4yz)} + e^{-\frac{k^2R^2}{2f^2}y^2} e^{-\frac{k^2\ell^2}{96f^2}(4z^2+y^2-4yz)}$$

$$+2\, e^{-\frac{k^2R^2}{2f^2}(y^2+z^2-yz)} \quad e^{-\frac{k^2\ell^2}{96f^2}(y^2+z^2-yz)} \Big\} \qquad (5.21)$$

Consider a binary

$$I(x) = \alpha_1 R(x) + \alpha_2 R(x-b) \qquad (5.3)$$

Then the averaged triple correlation for the binary is

$$T_B(y,z) = (\alpha_1^3+\alpha_2^3)\, T_R(y,z) + \alpha_1^2\alpha_2 [T_R(y,z-b) + T_R(y-b,z) + T_R(y+b,z+b)]$$

$$+ \alpha_1\alpha_2^2 [T_R(y,z+b) + T_R(y+b,z) + T_R(y-b,z-b)] \qquad (5.22)$$

We see that the triple correlation for the binary is made of seven PSFTCs whose strengths and locations are shown in Fig 5.2, The variance on the parity statistics Eq 5.5 is given by the diagramatic rule of the third chapter. The exact evaluation of all the orders is a tedious task and also unnecessary in many cases. We have considered three weight functions which emphasize three kinds of features in the triple correlation. The weight function of the first kind emphasizes the feature E which is the true triple correlation feature. For this weight function we have already worked out the variance in detail upto all orders. This calculation of the chapter 2, although based on simple model of PSF, tells us that below 13'th magnitude the dominent variance is in third order in the intensities. As mentioned several times before the neglect of edge effects is justified for variance. Thus the result of the second chapter can be used to restrict ourselves to the third order intensity calculations for feature E. The weight function of the second kind, which emphasizes all the correlation ridges, needs variance calculations upto all orders and this we do using the simpler PSF model of section 2.1.

The weight function of the third kind emphasize the plateau region A and is shown to be analogous to long exposure image considered in Appendix C5 which gives SNR for parity detection using long exposure images and compares it with double and triple correlation methods using the feature A alone. The result of these approximate calcultions for the weight functions of the second and the third kind result in SNR poorer than that due to the main feature E for binaries whose separation is near the diffraction limit of the telescope. For this reason it is not necessary to calculate the variance for these weight functions on the basis of detailed field calculations. In what follows we calculate the signal for all the weight functions on the basis of detailed field calculations which take into acount the edge effects. Also for the weight function of the first kind we calculate the variance as well but only in the lowest third order in the intensities which (from the results derived in chapter 2) is applicable to objects fainter than $13^{m}$ the regime where one is interested in comparing the Knox-Thompson and the triple correlation method.

a) SNR for parity detection due to the feature E

The feature E appears at seven places with strength and the locations shown in Fig 5.2. As mentioned in chapter 2 out of the six asymmetric elements it is enough to consider any two, as the others do not contain any independent information. This follows from the definition of the triple correlation Eq 5.1. If one had chosen all the six elements then the signal will be larger by a factor 3 while the variance will be larger by a factor 9 thus

giving the same SNR. This is because the three third order terms in the variance

$$2\int d^2Y\, d^2Z\; T_B(Y,Z)\, W(Y,Z)\{W(Y,Z)+W(-Y,Z-Y)+W(-Z,Y-Z)\} \qquad (5.24)$$

contribute equally and each one of them contributes three times what they would if only two elements were used. In general if the weight function is so chosen that there is no replication of the elements in the statistical sense then only one of the three terms survives. This is because in those cases the overlap of $W(Y,Z)$ with $W(-Y,Z-Y), W(-Z,Y-Z)$ is negligible. This will be the case for our choise of weight functions of the first and the second kind since we can beforehand figure out the independent elements and consider only them. We choose the weight function of the first kind shown in Fig 5.3 so that it emphasizes the feature E, in particular it has the same four-volume as the feature E but is unity within this volume. This weight function is chosen to be antisymmetric so that one get a antisymmetric parity statistics. Though this weight function is designed to emphasize the feature E it will get contributions from other features as well. For example the all pervading feature A due to all the PSF units will contribute. Calling the two elements of the weight function as port 1 and Port 2 we list the contributions to the two ports in the table 51 page 5-31. We see that the signal is just the total contribution at the port A minus the total contribution at the port B; while the variance is the total contribution at both the ports (there is a factor 2). One must multiply by the volume of the weight function and the triple correlation density. This product of the four volume and the

triple correlation density is $\frac{\pi^2 N_0^3 \ell^4 R^2}{9\, 2^4}$ . This gives us the signal and the noise for the weight function of the first kind given by Eq 5.8 .

b) SNR for parity detection due to the correlation ridges

As described in the text, of the nine ridges containing correlation ridges, only three are independent. So our choise for the weight function of the second kind is shown in Fig 5.4 . As the strips 1 and 3 are statistically equivalent in the average sense we can consider every strip individually. As mentioned in the text we are interested in the asymmetry in centers of these strips with respect to the $-45^0$ line passing through the flux center. When a feature like A is multiplied by the weight function we get an equivalent of a correlation ridge feature whose central strength is same but the fall off scale is different. Consider an equivalent feature situated distance away from the jump in the weight function. A B-feature centered at the origin has the following triple correlation density

$$B(Y,Z) = (N_0^3 k^4 \ell^4 R^6 / 3(2^8) f^4) \exp\{-\frac{k^2 R^2 (Y-Z)^2}{2f^2}\} \exp\{-\frac{k^2 \ell^2 (Y+Z)^2}{96 f^2}\} \qquad (5.25)$$

It is convenient to transform to coordinates $\xi$ and $\eta = \frac{1}{2}(Y \mp Z)$

$$B(Y,Z) \to B(\eta,\xi) = \frac{N_0^3 k^4 \ell^4 R^6}{3\, 2^8 f^4} \; e^{-\frac{2k^2R^2}{f^2}\xi^2} \; e^{-\frac{k^2\ell^2}{24 f^2}\eta^2} \qquad (5.26)$$

where a constant Jacobian of the transformation is not relevant as SNR cannot depend on a overall constant. The four volume of the B feature is

$$\int d^2\eta \, d^2\xi \, B(\eta,\xi) = 2^{-6} \pi^2 N_0^3 \ell^2 R^4 \qquad (5.27)$$

One can integrate the B feature density in $\eta_2, \xi_2$ plane and normal

to strip ( $\xi_1$ ) to get the projected line density of a B feature:

$$f(\eta_1) = \int d^2\xi \, d\eta_2 \; B = (\pi^{1.5} N_0^3 k \ell^3 R^4/\sqrt{3}\, 2^{7.5} f) \exp\{-k^2\ell^2\eta^2/24f^2\} \quad (5.28)$$

For a B feature with center $\eta_c$ away from the jump in the weight function and strength $\gamma$ the contribution to parity is given by

$$\int d\eta_1 \, W(\eta_1)\gamma f(\eta_1 - \eta_c) = \gamma 2\eta_c \, W(\eta_c)(\pi^{1.5} N_0^3 k\ell^3 R^4/\sqrt{3}\, 2^{7.5} f) \quad (5.29)$$

As shown in Fig only the region $\pm\eta_c$ about the center of the feature contributes. For small separations this can be approximated by the product of the width $2\eta_c$ and the peak strength $\frac{\pi^{1.5} N_0^3 k \ell^3 R^4}{\sqrt{3}\, 2^{7.5} f}$ . The contribution is +ve or -ve depending on the side to which the center lies from the jump. The effect of the weight function is a yield equal to the central part of the feature as the remaining two sides cancel each other. In Fig 5.4b we show the features contained in all the strips and their distances from the jump. We thus get the contribution to the signal

$$S_1 = S_3 = \tfrac{1}{2} S_2 \qquad S_2 = \frac{\pi^{1.5} N_0^3 k \ell^3 R^4 b}{\sqrt{3}\, 2^6 f}$$

$$S = \frac{4}{\sqrt{3}\, \pi^{1.5}} N_s^{3/2} \mathcal{N}_1 \mathcal{N}_2 (\mathcal{N}_1 - \mathcal{N}_2) \mathcal{B}$$

Appendix B5 gives the variance for this weight function in all orders based on the approximate PSF model of section 2.1. The SNR estimate for this weight function are given by Eq .

APPENDIX B5

Variance for the weight function of the second kind upto all orders: In the notation of chapter 2 the discretised version of the weight function of the second kind shown in Fig 5.4 for the central strip alone is

$$W_{ijk} = W_{i\,i+a\,i+c} = f_a\,\delta_{ac} = f_{j-i}\,\delta_{jk} \qquad f_a = sign\, a_1 \tag{5.30}$$

where the vector subscript represents the position of the i'th pixel which is of the speckle size. First we consider the classical sixth order variance. As mentioned in the text we let the binary shrink to a point to get the variance

$$(1+\alpha)^6 \sum_{ijklmn} W_{ijk}\, W_{lmn} \left[\langle \mu_i \mu_j \mu_k \mu_l \mu_m \mu_n \rangle - \langle \mu_i \mu_j \mu_k \rangle \langle \mu_l \mu_m \mu_n \rangle\right] \tag{5.31}$$

in terms of the speckle intensity $\mu_i$ on the i'th pixel. Using the form Eq5.30 for the weight function we get for the classical variance

$$(1+\alpha)^6 \sum_{ijkl} f_{j-i}\, f_{l-k} \left[\langle \mu_i \mu_j^2 \mu_k \mu_l^2 \rangle - \langle \mu_i \mu_j^2 \rangle \langle \mu_k \mu_l^2 \rangle\right] \tag{5.32}$$

The following results are needed

$$\begin{aligned}\langle \mu_i \mu_j^2 \mu_k \mu_l^2 \rangle = {} & 4\langle\mu_i\rangle\langle\mu_j\rangle^2\langle\mu_k\rangle\langle\mu_l\rangle^2 + 8\,\delta_{ij}\langle\mu_i\rangle^3\langle\mu_k\rangle\langle\mu_l\rangle^2 + 4\,\delta_{ik}\langle\mu_i\rangle^2\langle\mu_j\rangle^2\langle\mu_l\rangle^2 \\ & + 8\delta_{il}\langle\mu_i\rangle^3\langle\mu_j\rangle^2\langle\mu_k\rangle + 8\,\delta_{jk}\langle\mu_i\rangle\langle\mu_j\rangle^3\langle\mu_l\rangle^2 + 20\,\delta_{jl}\langle\mu_i\rangle\langle\mu_j\rangle^4\langle\mu_k\rangle \\ & + 8\,\delta_{kl}\langle\mu_i\rangle\langle\mu_j\rangle^2\langle\mu_k\rangle^3 + 16\,\delta_{ij}\delta_{kl}\langle\mu_i\rangle^3\langle\mu_k\rangle^3 + 20\,\delta_{ik}\delta_{jl}\langle\mu_i\rangle^2\langle\mu_j\rangle^4 \\ & + 16\delta_{il}\,\delta_{jk}\langle\mu_i\rangle^3\langle\mu_j\rangle^3 + 24\,\delta_{ijk}\langle\mu_i\rangle^4\langle\mu_l\rangle^2 + 80\,\delta_{jkl}\langle\mu_i\rangle\langle\mu_j\rangle^5 \\ & + 80\,\delta_{ijl}\langle\mu_i\rangle^5\langle\mu_k\rangle + 24\,\delta_{ikl}\langle\mu_i\rangle^4\langle\mu_j\rangle^2 + 400\,\delta_{ijkl}\langle\mu_i\rangle^6\end{aligned} \tag{5.33}$$

$$\begin{aligned}\langle \mu_i \mu_j^2 \rangle \langle \mu_k \mu_l^2 \rangle = {} & 4\langle\mu_i\rangle\langle\mu_j\rangle^2\langle\mu_k\rangle\langle\mu_l\rangle^2 + 8\,\delta_{ij}\langle\mu_i\rangle^3\langle\mu_k\rangle\langle\mu_l\rangle^2 \\ & + 8\,\delta_{kl}\langle\mu_i\rangle\langle\mu_j\rangle^2\langle\mu_k\rangle^3 + 16\,\delta_{ij}\,\delta_{kl}\langle\mu_i\rangle^3\langle\mu_k\rangle^3\end{aligned} \tag{5.34}$$

This gives us the variance in the sixth order

$$\begin{aligned}(1+\alpha)^6 \sum_{ijkl} & f_{j-i}\, f_{l-k} \{ 4\,\delta_{ik}\langle\mu_i\rangle^3\langle\mu_j\rangle^2\langle\mu_l\rangle^2 + 8\langle\mu_i\rangle^3\langle\mu_j\rangle^2\langle\mu_k\rangle\,\delta_{il} + 8\delta_{jk}\langle\mu_i\rangle\langle\mu_j\rangle^3\langle\mu_l\rangle^2 \\ & + 20\,\delta_{jl}\langle\mu_i\rangle\langle\mu_j\rangle^4\langle\mu_k\rangle + 20\,\delta_{ik}\,\delta_{jl}\langle\mu_i\rangle^2\langle\mu_j\rangle^4 + 16\,\delta_{il}\,\delta_{jk}\langle\mu_i\rangle^3\langle\mu_j\rangle^3 \\ & + 24\,\delta_{ijk}\langle\mu_i\rangle^4\langle\mu_l\rangle^2 + 80\,\delta_{ijl}\langle\mu_i\rangle^5\langle\mu_k\rangle + 24\,\delta_{ikl}\langle\mu_i\rangle^4\langle\mu_j\rangle^2 + 8\,\delta_{jkl}\langle\mu_i\rangle\langle\mu_j\rangle^5 \\ & + 400\,\delta_{ijkl}\langle\mu_i\rangle^6 \}\end{aligned} \tag{5.35}$$

Other terms vanish because of the factors like $f_0 = 0$.

For our weight function it can be seen from Fig 5.5 that

$$\sum_{j_1} f_{j_1 - i_1} = -2 i_1 \tag{5.36}$$

Following the spirit of the chapter 2 we consider the statistics of the $\mu$'s to be uniform within the seeing disk, so $i_1, i_2 \cdots$ run from $-\frac{N_S^{1/2}}{2}$ to $\frac{N_S^{1/2}}{2}$ so that $i_1$ takes $N_S^{1/2}$ values. The total number of pixels within the seeing disk is $\sum_{i_1, i_2} = N_S$ . This range of variation of $i_1$ actually comes from factors like $\langle \mu_i \rangle^4$ etc which are nonzero only within the seeing disk. Thus it follows that

$$\sum_{ijk} f_{j-i}\, f_{k-i} = N_S^{3/2} \sum_{i_1 j_1 k_1} f_{j_1 - i_1} f_{k_1 - i_1} = 4 N_S^{3/2} \sum_{i_1} i_1^2 = \frac{1}{3} N_S^3 \tag{5.37}$$

This gives us the sixth order variance for the weight function of the second kind

$$\frac{8}{3} N_S^3 (\mathcal{N}_1 + \mathcal{N}_2)^6 + 4 N_S^2 (\mathcal{N}_1 + \mathcal{N}_2)^6 \sim \frac{8}{3} N_S^3 q^6 (\mathcal{N}_1 + \mathcal{N}_2)^6 \tag{5.38}$$

From the chapter 3 the fifth order terms in the variance are

$$(1+\alpha)^5 \sum_{ijklm} \langle \mu_i \mu_j \mu_k \mu_l \mu_m \rangle W_{ijk} \{ W_{ilm} + W_{lim} + W_{lmi} + W_{jlm} + W_{ljm} + W_{lmj} + W_{klm} + W_{lkm} + W_{lmk} \} \tag{5.39}$$

which reduce to

$$(1+\alpha)^5 \sum_{ijklm} \langle \mu_i \mu_j \mu_k \mu_l \mu_m \rangle W_{ijk} \{ W_{ilm} + 4 W_{lim} + 4 W_{ljm} \} \tag{5.40}$$

on using the fact that two $W$'s appear or the symmetry

$$W_{ijk} = W_{ikj} \tag{5.41}$$

of the weight function. These terms can be shown to be

$$\sum \langle \mu_i \mu_j \mu_k \mu_l \mu_m \rangle W_{ijk} W_{ilm} = \frac{4}{3} N_S^3 (N_1+N_2)^5 + 20 N_S^2 (N_1+N_2)^5$$

$$4 \sum \langle \text{———}\,,,\,\text{———} \rangle W_{ijk} W_{lim} = -\frac{16}{3} N_S^3 (N_1+N_2)^5 - 32 N_S^2 (N_1+N_2)^5$$

$$4 \sum \langle \text{———}\,,,\,\text{———} \rangle W_{ijk} W_{ljm} = 8 N_S^3 (N_1+N_2)^5 + 24 N_S^2 (N_1+N_2)^5 \qquad (5.42)$$

This gives the fifth order variance

$$4(N_S^3+3N_S^2)(N_1+N_2)^5 \sim 4 N_S^3 q^5 (N_1+N_2)^5 \qquad (5.43)$$

The fourth order contribution to the variance is (chapter 3)

$$(1+\alpha)^4 \sum_{ijkl} \langle \mu_i \mu_j \mu_k \mu_l \rangle W_{ijk} \{ W_{ijl} + W_{ilj} + W_{lij} + W_{jil} + W_{jli} + W_{lji} + W_{ikl} + W_{kil} + W_{ilk} + W_{kli} + W_{lki} + W_{lik} + W_{jkl} + W_{kjl} + W_{jlk} + W_{klj} + W_{ljk} + W_{lkj} \} \qquad (5.44)$$

This follows from chapter 2. Again using the symmetry Eq 5.41 and the fact that two W's appear we can reduce this to

$$2(1+\alpha)^4 \sum_{ijkl} \langle \mu_i \mu_j \mu_k \mu_l \rangle W_{ijk} \{ 2 W_{ijl} + 4 W_{lij} + 2 W_{jli} + W_{ljk} \} \qquad (5.45)$$

The following averages are needed

$$4 \sum \langle \mu_i \mu_j \mu_k \mu_l \rangle W_{ijk} W_{ijl} = (96 N_S^2 + 384 N_S)(N_1+N_2)^4$$

$$8 \quad \langle \text{———}\,,,\,\text{———} \rangle W_{ij} \; W_{lij} = 0$$

$$4 \sum \langle \text{———}\,,,\,\text{———} \rangle W_{ijk} W_{jli} = (-16 N_S^2 - 80 N_S)(N_1+N_2)^4$$

$$\mathbf{2} \sum \langle \text{———}\,,,\,\text{———} \rangle W_{ijk} W_{ljk} = (\tfrac{4}{3} N_S^3 + 4 N_S^2)(N_1+N_2)^4 \qquad (5.46)$$

This gives the fourth order variance

$$(\tfrac{4}{3} N_S^3 + 84 N_S^2 + 304 N_S)(N_1+N_2)^4 \sim \tfrac{4}{3} N_S^3 q^4 (N_1+N_2)^4 \qquad (5.47)$$

The third order contribtuion contains six terms (chapter 3)

$$(1+\alpha)^3 \sum_{ijk} \langle \mu_i \mu_j \mu_k \rangle W_{ijk} \{ W_{ijk} + W_{ikj} + W_{jik} + W_{jki} + W_{kij} + W_{kji} \} \qquad (5.48)$$

which on using the symmetry of W reduces to

$$2(1+\alpha)^3 \sum \langle \mu_i \mu_j \mu_k \rangle W_{ijk} \{ W_{ijk} + W_{jik} + W_{kij} \} \qquad (5.49)$$

Since

$$2 \sum \langle \mu_i \mu_j \mu_k \rangle W_{ijk}^2 = 4(N_S^2+2N_S)(N_1+N_2)^3 \qquad (5.50)$$

$$\Sigma\langle\mu_i\mu_j\mu_k\rangle W_{ijk}\,W_{jik}=\Sigma\langle\mu_i\mu_j\mu_k\rangle W_{ijk}\,W_{kij}=0$$

we get the third order contribution to be

$$4N_S^2+2N_S)q^3(N_1+N_2)^3\sim 4N_S^2q^3(N_1+N_2)^3 \qquad (5.51)$$

Since $N_S$ is much larger than unity one needs to retain only the leading terms in the every order in $N$. We have considered only one strip so we must multiply by 3 to get variance due to all the strips. This gives us the variance upto all orders

$$72N_S^3q^6N^6+12N_S^3q^5N^5+4N_S^3q^4N^4+12N_S^2q^3N^3 \quad ; N=N_1+N_2 \qquad (5.52)$$

We see that there are two transitions. The first is when the fourth order variance takes over from the classical sixth order contribution. This happens when individual speckles recieve less than one photon in an exposure. The second transition occurs when the third order contribution dominates the fourth order. This happens when the entire photon count becomes less than unity.

APPENDIX C5

First order parity statistics: SNR

In chapter 4 and 5 we encountered the feature A which originated in uncorrelated regions of double and triple correlations respectively. We argued in chapter 4 that this is equivalent to a long exposure image with a noisy weight function. It is therefore natural to consider parity determined from the long exposure image as well, although, everyone will admit without a moments paouse that this is not a genuine high resolution method. Anyway we consider a first order statistics of the form

$$S=\sum \left(i_1-\frac{\alpha_2 b}{\alpha_1+\alpha_2}\right)^3(\mu_i+\alpha\mu_{i-b}) \quad \because \bar{n}_i=\mu_i+\alpha\mu_{i-b} \tag{5.53}$$

where we have used the discrete pixel model of section 2.1. We have centered the weight function on the flux center $\frac{\alpha_2 b}{\alpha_1+\alpha_2}$ The cubic nature of the weight function follows because one is interested in asymmetry in the image so the weight function should contain odd powers; the first power gives zero by the definition of the flux center. One can expand the cube and retain only the even powers of i. The odd powers vanish on account of symmetry of $\mu_i$ about the origin which is the location of star 1 whose PSF is the $\mu_i$'s . In Eq 5.53 b is the binary separation in units of speckle size (in the notation of section 2.1) which is the same as of chapters 4 and 5.

$$S=\left[\alpha\frac{\alpha_1 b}{\alpha_1+\alpha_2}-\frac{\alpha_2 b}{\alpha_1+\alpha_2}\right]\sum i_1^3\langle\mu_i\rangle+\left[\alpha\left(\frac{\alpha_1 b}{\alpha_1+\alpha_2}\right)^3-\left(\frac{\alpha_2 b}{\alpha_1+\alpha_2}\right)^3\right]\sum\langle\mu_i\rangle \qquad \alpha=\frac{\alpha_2}{\alpha_1}$$

$$=N_S b^3 N_1 N_2(N_1^2-N_2^2)/(N_1+N_2)^3 \tag{5.54}$$

For variance calculation we let the binary shrink to a point. The variance for the first order parity statistics considered here cosists of a classical second order and a first order photonic contribution

$$(1+\alpha)^2\sum i_1^3 j_1^3\left[\langle\mu_i\mu_j\rangle-\langle\mu_i\rangle\langle\mu_j\rangle\right]+\sum i_1^6\langle\mu_i\rangle(1+\alpha) \tag{5.55}$$

A straightforward calculation gives

$$V=\frac{q^2 N_S^4(N_1+N_2)^2}{7\quad 2^6}+\frac{qN_S^4(N_1+N_2)}{7\quad 2^6} \tag{5.56}$$

The SNR is

$$SNR_{LE}=8\sqrt{7}\,m^{1/2}\frac{IB^3}{N_S}\frac{N_1N_2(N_1^2-N_2^2)}{(N_1+N_2)^4} \tag{5.57a}$$

$$=8\sqrt{7}\,q^{1/2}m^{1/2}\frac{IB^3}{N_S}\frac{N_1N_2(N_1^2-N_2^2)}{(N_1+N_2)^{3.5}} \tag{5.57b}$$

Note the characteristic $B^3$ dependence which we got in all SNRs (double or triple correlation) using the feature A. The scaling with $N_S$ and $N$ is, however, different and reflects the order of the statistics. For this weight function the transition to low flux occurs when the first order contribtuion dominates the classical second order contribution. This means $qN=1$ i.e $12^m$. The table 5.3 below gives the limiting faintness of the fainter component for various binary separations.

Table 5.3 Limiting faintness $m_2$ for parity determination using long exposure images

| $m_1$ | Binary separation in units of speckle size | | | | | | |
|---|---|---|---|---|---|---|---|
| | 1 | 4 | 7 | 10 | 13 | 16 | 19 |
| 7 | X | 12 | 13.5 | 15 | 15.5 | 16.5 | 17 |
| 8 | X | 13 | 14.5 | 15.7 | 16.5 | 17.5 | 18 |
| 9 | X | 14 | 15.5 | 16.7 | 17.5 | 18.2 | 19 |
| 10 | X | 14.5 | 16.5 | 17.5 | 18.5 | 19 | 19.5 |
| 11 | X | 15 | 17 | 18.5 | 19 | 20 | 20 |
| 12 | X | 16 | 17.7 | 19 | 20 | 20 | 20 |
| 13 | X | 16.5 | 18.5 | 19.5 | 20 | 20 | 20 |
| 14 | X | 16.7 | 19 | 20 | 20 | ↓ | ↓ |
| 15 | X | 16.7 | 19.5 | 20 | ↓ | ↓ | ↓ |
| 16 | X | X | 20 | ↓ | ↓ | ↓ | ↓ |
| upto 19 | X | X | 20 | ↓ | ↓ | ↓ | ↓ |

$m_1$: magnitude of brighter component, $m_2$: fainter component

In the above table we limit ourselves to magnitudes brighter than $20^m$ keeping in mind the sky background. We note that except for close binaries the limiting faintness is indeed quite good. It is necessary to emphasise that this detection of parity by first order statistics has been discussed for the sake of completeness and comparison with higher order methods. It is not being suggested as a practical method for the following reason. A typical speckle image even of a single star has gross asymmetries produced by large scale tilts which are described by the

Kolmogorov model of phase variations on the pupil. In the present model these asymmetries do not occur. They will adversely affect the first order statistics defined here.

Table 5.1 Contributions to Port 1 and 2.

| | Strength | Contributing features | contribution |
|---|---|---|---|
| Port 1 | $\alpha_1^2\alpha_2$ | $A_1, B_1, C_1, D_1, 2E_1, A_3, A_5$ | $8\alpha_1^2\alpha_2$ |
| | $\alpha_1^3+\alpha_2^3$ | $A_7, B_7$ | $2(\alpha_1^3+\alpha_2^3)$ |
| | $\alpha_1\alpha_2^2$ | $D_6, C_2, A_2, A_6, A_4, B_4$ | $\frac{6\alpha_1\alpha_2^2}{8\alpha_1^2\alpha_2+2(\alpha_1^3+\alpha_2^3)+6\alpha_1\alpha_2^2}$ |
| Port 2 | $\alpha_1^2\alpha_2$ | $A_3, D_3, A_1, B_1, A_5, C_5$ | $6\alpha_1^2\alpha_2$ |
| | $\alpha_1^3+\alpha_2^3$ | $A_7, B_7$ | $2(\alpha_1^3+\alpha_2^3)$ |
| | $\alpha_1\alpha_2^2$ | $A_4, B_4, C_4, D_4, 2E_4, A_2, A_6$ | $\frac{8\alpha_1\alpha_2^2}{6\alpha_1^2\alpha_2+2(\alpha_1^3+\alpha_2^3)+8\alpha_1\alpha_2^2}$ |

Fig 5.1

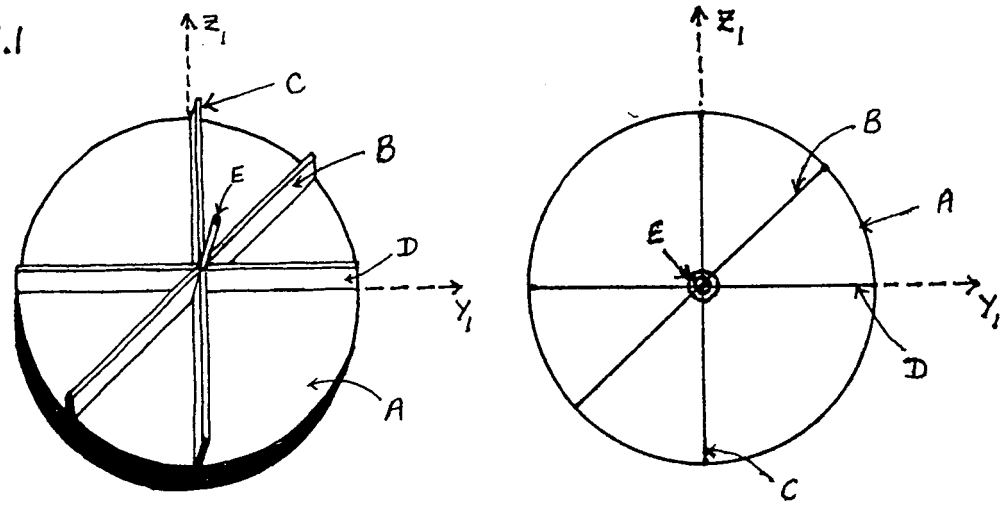


Fig 5.1 PSF triple correlation: left) schematic; right) Symbolic

Fig 5.2

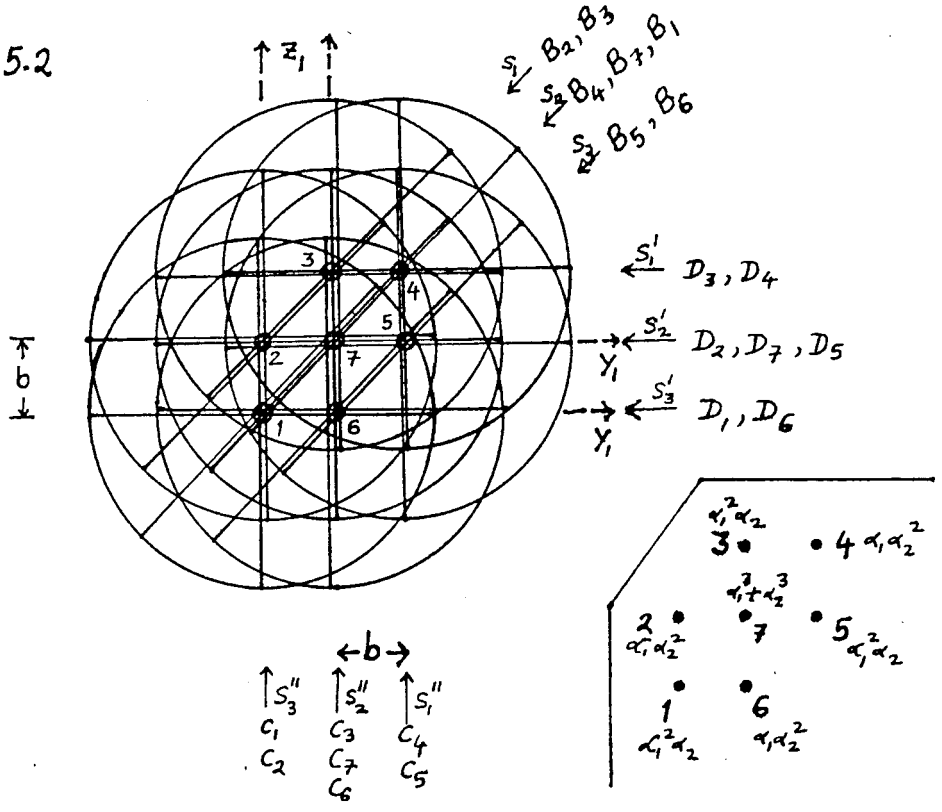


Fig 5.2. Triple correlation for a binary: ridges slightly displaced for clarity. Inset: seven PSF units comprising the binary T.C.

Fig 5.3

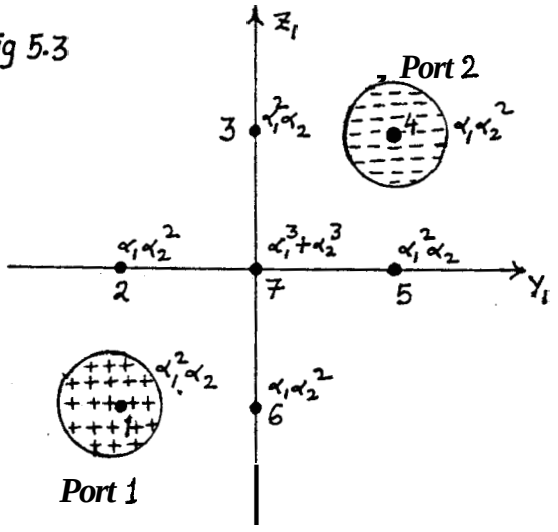


Fig 5.3. Weight function used to emphasise the E-feature

Fig 5.4 a)

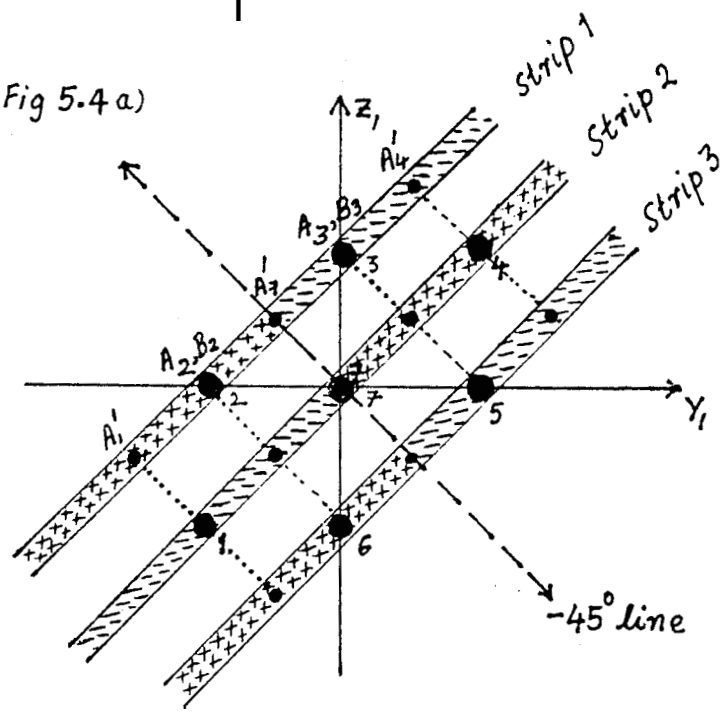


Fig 5.4 a)

Weight function used to emphasise the correlation ridges.

Fig 5.4 b)

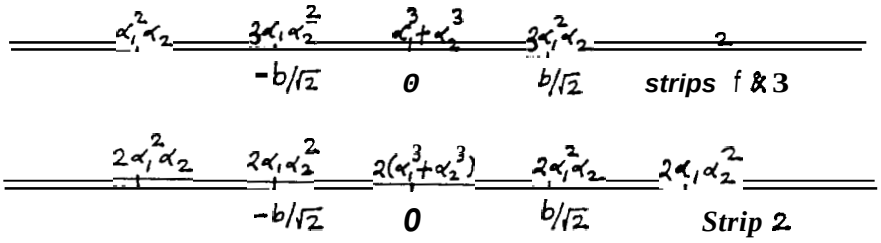


Fig 5.4b) Centers and strengths of 'effective' B-features on ships.

Fig 5.4 c)

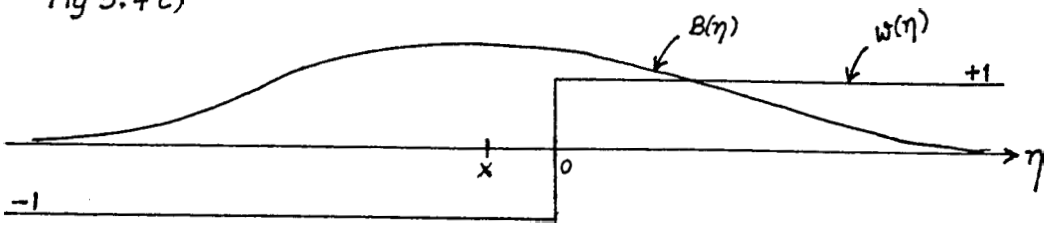


Fig 5.4 c) An effective B-feature $B(\eta)$ with center a distance $X$ away from the jump in the weight function $w(\eta)$

Fig 5.5

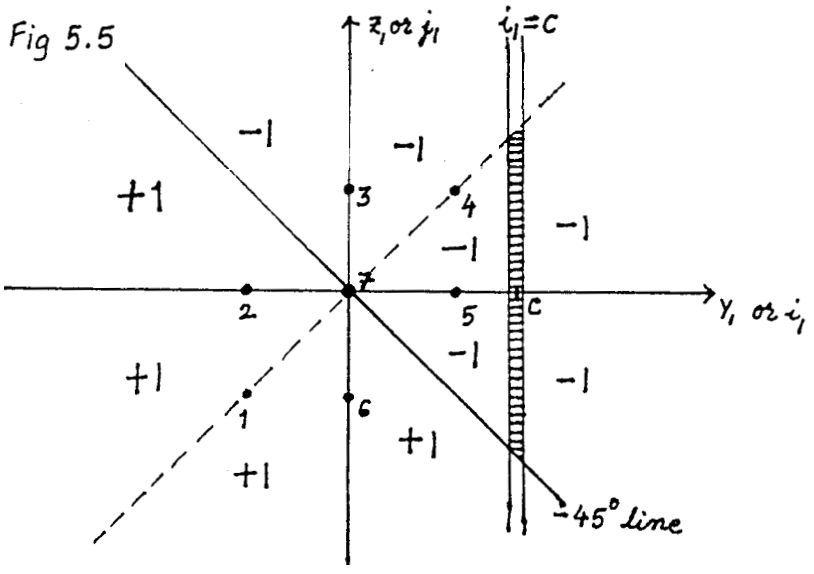


Fig 5.5. Weight function used to emphasise the feature A

The weight function is +1 to the left of the $-45^0$ line and -1 to the right of this line.

## CONCLUSIONS

In this thesis we have estimated the SNR for parity determination using the focal plane intensity correlations of the first (long exposure), second (Knox-Thompson) and third (Weigelt) order. The PSFs for these correlations have one, two and three kinds of features respectively. Different kinds of features yield SNR with different dependence on $N_s$, $\mathcal{N}$ and $b$. The following table CI lists these scalings.

Table CI SNR SCALINGS

| Method (Statistics) | High flux brighter than $13^m$ | medium flux $13^m$-$18^m$ | Low flux fainter than $18^m$ |
|---|---|---|---|
| TOTAL FLUX * | $N_s^{1/2}$ | $N_s^{1/2}\mathcal{N}^{1/2}$ | $N_s^{1/2}\mathcal{N}^{1/2}$ |
| AUTOCORRELATION* (Labeyrie) | $N_s^{1/2}$ | $N_s^{1/2}\mathcal{N}$ | $N_s^{1/2}\mathcal{N}$ |
| PARITY | | | |
| 1) Long exposure | $\mathcal{B}^3/N_s$ | $\mathcal{N}^{1/2}\mathcal{B}^3/N_s$ | $\mathcal{N}^{1/2}\mathcal{B}^3/N_s$ |
| 2) Double correlation (Knox-Thompson) | | | |
| 1.1) Ridges | $1\mathcal{B}$ | $\mathcal{N}\mathcal{B}$ | $\mathcal{N}\mathcal{B}$ |
| 1.2) Plateau | $\mathcal{B}^3/N_s$ | $\mathcal{N}^{1/2}\mathcal{B}^3/N_s$ | $\mathcal{N}\mathcal{B}^3/N_s^{1/2}$ |
| 3) Triple correlation (Weigelt) | | | |
| 2.1) Peak | $N_s^{1/2}$ | $N_s^{1/2}\mathcal{N}^{3/2}$ | $N_s^{1/2}\mathcal{N}^{3/2}$ |
| 2.2) Ridges | $1\mathcal{B}$ | $\mathcal{N}\mathcal{B}$ | $N_s^{1/2}\mathcal{B}\mathcal{N}^{3/2}$ |

Observational parameters: 4m telescope, 100A bandwidth, exposure time 10ms, $\mathcal{B}$ is binary separation in units of the diffraction limit of the telescope, $\mathcal{N}$ is photon count in an exposure per speckle, *: results known in the literature.

In evaluating the SNR for some of the features an implicit assumption was that the binary nature of source was known

beforehand. Since the limiting faintness for parity determination turns out to be always ~~fainter~~ brighter than that for autocorrelation (table 2.8 ) this assumption is justified a posteriori. From table 5.3 which gives the limiting faintness for parity detection using the long exposure image (observation time equivalent to the above parameters) would appear to have as good a SNR as the feature E of triple correlation for binary separation five time the resolution limit (see however remarks at the end of Appendix C5 page 5-30). So for the specific case of wide binaries (still within the seeing disk) it it may well be better to use much simpler "monocorrelation". As mentioned earlier, we treat binaries as guinea-pigs on which to test various schemes of phase reconstruction. It is therefore necessary to supplement these calculations with a scrutiny of various features with high resolution applications in mind. For example, although the long exposure method is better than any other for separations significantly greater than the resolution limit it can not be considered as a high resolution feature because all it detects is asymmetry in the source. Every binary (unequal components) is asymmetric but not every asymmetric source is a binary. For complex sources only genuine high resolution methods are relevant. There are only three such feature: the feature B of double correlation and the features B and E of triple correlation. From the above table we see that the correlation ridges (features B) have the same scalings in the medium range light levels for both double and triple correlation. From Eq 4.12 and Eq 5.12b we conclude that the Knox-Thompson

method is slightly better than the feature B of triple correlation. In the following table CII we give the limiting faintness for the remining two features. For comparison we have included the long exposure method.

Table CII Limiting faintness $m_2$ for parity detection

| | Binary separation | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 7 | | 10 | | all |
| $m_1$ | LE | KT-B | LE | KT-B | LE | KT-B | LE | KT-B | TC-E |
| 13 | X | 16.2 | 16.5 | 17.7 | 18.5 | 18.2 | 19.5 | 18.7 | 17.2 |
| 14 | X | 15.7 | 16.5 | 17.7 | 19 | 18.2 | 20 | 18.7 | 17 |
| 15 | X | X | 16.5 | 17.7 | 19.5 | 18 | 20 | 18.7 | 16 |
| 16 | X | X | X | X | 19.7 | 18 | 20 | 18.7 | X |
| 17 | X | X | X | X | 20 | X | 20 | X | X |
| 18 | X | X | X | X | 20 | X | 20 | X | X |

$m_1$: magnitude of brighter star, $m_2$: fainter star, X: none with SNR > 3

We conclude that for close binaries the triple correlation method is better while for wider separations the Knox-Thompson is better.

In summary the new results are i) Systematic study of the $N_s$ and $N$ dependence of the SNR for parity in different schemes. ii) The binary separation dependence of the SNR for phase reconstruction. iii) Diagramatic rules for computing the Poisson fluctuations in general image intensity correlations.

## REFERENCES

### REVIEWS


a. Bates, R.H.T. Astronomical speckle imaging, Phy. Rep. 90 203 (1982)

b. Brown R-H, The intensity interferometer & its applications to astronomy, Taylor & Francis Ltd. ISBN 0 85066 072 6

c. Coulman C.E., Annu.Rev.Astron.Astrophy. 23,19 (1985)

d. Labeyrie A., High resolutiontechniques in optical astronomy Progress in optics, XIV, 47 (1977)

e. Labeyrie A., Stellar interferometric methods, Annu.Rev.Astron.Astrophy. 16,77 (1978)

f. McAlister H.A., High angular resolution measurements of stellar properties, Annu.Rev.Astron.Astrophy.23,59(1985)

g. Minnaert M., Light and colour in the open air, DOVER(1940)

h. Roddier F., The effects of atmospheric turbulence in optical astronomy, Progress in optics XIX 281 (1981)

i. Roddier F., Interferometric imaging in optical astronomy Phys. Rep. 170, 99 (1988)

j. Tango W.J. & Twiss R.Q., Michelson stellar interferometry Progress in optics

k. Woolf N.J., High resolution astronomy from the ground Annu.Rev.Astron.Astrophy. 20,367 (1982)


### PAPERS


1. Ayers et al JOSA A5, 963 (1988)

2. Bartlet et al Appl. Opt. 23, 3121 (1984)

3. Dainty J.C. MNRAS 169, 631 (1974)

4. Dainty et al JOSA 69, 786 (1979)

5. Gezari et al AJ 173 L1 (1972)

6. Karbelkar et al Jou. Astron. Astrophys. 3,271(1987)

7. Knox et al Ap.J. 193 L45 (1974)

8. Labeyrie A., A & A 6, 85 (1970)

9. Lohmann et al Appl.Opt. 22, 4028 (1983)

10. McAlister et al AJ 96, 1431 (1988)

11. Proceedings of NOAO-ESO conference on 'High resolution imaging by interferometry' Garching 1988, Ed. Merckle F

12. Michelson A.A., Phil. Mag. 30, 1 (1890)

13. Michelson et al Ap.J 53, 249 (1921)

14. Nakajima T., JOSA A5, 1477 (1988)

15. Nisenson et al Opt.Commun. 47, 91 (1983)

16. Ramchandran G.N., Proc. Indian Academy of Sciences A18 (1943)

17. Risken H., Progress in optics VIII 239 (1970)

18. Weigelt G., Opt. Commun. 21, 55 (1977)

19. Wirnitzer B., JOSA A2 14 (1985)