Performance Analysis of a Noncoherently Combined Large Aperture Optical Heterodyne Receiver

C.-C. Chen
Communications Systems Research Section

The performance of a noncoherently combined, multiple-mirror heterodyne receiver is analyzed. It is shown that in the absence of atmospheric turbulence, the performance of the noncoherently combined receiver is inferior to that of a monolithic, diffraction-limited receiver with equivalent aperture area. However, when atmospheric turbulence is taken into consideration, the efficiency of a monolithic aperture heterodyne receiver is limited by the phase coherence length of the atmosphere and generally does not improve with increasing aperture size. In contrast, the performance of a noncoherently combined system improves with an increasing number of receivers. Consequently, given a fixed collecting area, the noncoherently combined system can offer superior performance. The performance of the noncoherently combined heterodyne receiver is studied by analyzing the combining loss of the receiver SNR. It is shown that, given the collecting area, the performance of the combined receiver is optimized when the diameter of each of the individual receivers is on the order of the phase coherence length r_0 of the atmospheric turbulence.

I. Introduction

Optical heterodyne reception [1] provides an alternative to direct-detection schemes for free-space optical communication applications. The ability to reject noncoherent background radiation has made the heterodyne system very attractive for applications where effective communication in the presence of strong background interference is required. Furthermore, the use of frequency and phase modulation schemes can remove the peak power constraint that currently limits the application of higher-order direct-detection pulse-position modulation (PPM) schemes.

The heterodyne receiver is more complicated than the direct-detection receiver. Accurate wavefront matching between the incoming signal and the local oscillator (LO) is needed to ensure effective heterodyne reception. Imperfect spatial mode matching can lead to destructive interference and, consequently, to degraded system performance. For systems employing small receiver apertures, the required wavefront alignment accuracy can be achieved relatively easily because the wavefront distortions due to the receiving optics and the atmosphere are negligible. For systems with large collecting apertures, such as those envisioned for deep-space reception, the difficulty in maintaining good optical surface quality across the input aperture can present a serious problem in achieving effective heterodyne reception. The presence of atmospheric turbulence further complicates the problem. Turbulence-induced beam wander and beam spreading can result in random fluctuations of the signal phase and amplitude that are difficult to compensate for. As a result, the performance of a large-aperture receiver generally does not improve with increasing collecting area.

Alternatively, a large effective aperture can be achieved by combining the output signals from an array of smaller receivers. These receivers can share a common support structure and tracking electronics to reduce construction cost. Because each receiver is now smaller than the roughness scale of the incoming signal wavefront, the local-oscillator output can be accurately matched to the signal to achieve effective heterodyne reception. Output signals from these receivers can then be combined electronically to improve the detection statistics. Ideally, output signals from different apertures should be combined coherently to optimize the overall receiver performance. In such a scheme, turbulence- and optics-induced phase shifts in the detected intermediate-frequency (IF) signals are first compensated electronically, and the outputs from these receivers are then added coherently. Since correcting the IF phase distortion is similar to correcting the incoming phase front, a coherent combining system can offer performance comparable to that of a monolithic diffraction-limited receiver. For systems with weak signal intensities, however, the low signal-to-noise ratio (SNR) and rapidly varying atmospheric conditions preclude the possibility of effective phase compensation. The use of adaptive optics [2] in conjunction with an artificial guide star [3] can compensate for the atmospherically induced wavefront distortion. However, such measures are costly and complicated. An alternative is to noncoherently combine the outputs of several subaperture receivers. In such a scheme, the detected IF signal of each individual heterodyne receiver is first noncoherently demodulated [4], and the outputs of these demodulators are then electronically combined.

This article presents a simple analysis of a noncoherently combined heterodyne receiver. A simple model for the output of the heterodyne receiver is first constructed. The performance of a noncoherently combined optical heterodyne receiver is then analyzed. It is shown that in the absence of atmospheric distortions, the performance of a noncoherently combined heterodyne receiver is worse than that of a single, monolithic-aperture heterodyne receiver of equivalent aperture size. This is because the noncoherent demodulation process is more susceptible to noise when the signal power is weak. Consequently, by subdividing the total collection area into a number of smaller receivers, the signal power collected by each receiver is smaller, thus leading to a degraded combiner performance. When the atmospheric turbulence effect is taken into consideration, however, the performance of the monolithic-aperture receiver is limited by the phase coherence length of the atmosphere. As a result, subdividing the aperture and noncoherently combining the subaperture outputs can actually lead to a significant improvement in performance. A simple analysis shows that the combiner SNR is optimized when the diameter of the individual subaperture is on the order of the phase-coherence length of the atmosphere.

II. Heterodyne Reception Technique

The structure of a simple dual-detector heterodyne receiver is shown in Fig. 1. The incoming signal is first spatially mixed with a strong local oscillator. The combined signal is then photodetected. The mixing of the signal and LO generates an intermediate frequency term at each detector output which oscillates at the beat frequency ω_{IF} between the signal and LO. The IF signal output of the dual-detector heterodyne receiver can be modeled as [1]:

$$s_{\mathrm{IF}} = \left(\frac{e\eta G}{h\nu}\right)\frac{1}{z_0}\int W_D(\mathbf{r})\vec{A}_S(\mathbf{r}) \times \vec{A}_{\mathrm{LO}}(\mathbf{r})\cos\left(\omega_{\mathrm{IF}}t + \phi_S(\mathbf{r}) - \phi_{\mathrm{LO}}(\mathbf{r})\right)d\mathbf{r} + n(t) \tag{1}$$

where $(e\eta/h\nu)$ is the responsivity of the detector, G is the detector gain, z_0 is the impedance of the photodetector, $\vec{A}_S(\mathbf{r})$ and $\vec{A}_{\mathrm{LO}}(\mathbf{r})$ are the amplitudes of the signal and LO electric fields, $\phi_S(\mathbf{r})$ and $\phi_{\mathrm{LO}}(\mathbf{r})$ are the signal and LO phases, ω_{IF} is the IF frequency, and $n(t)$ is the additive noise at the detector output. The function $W_D(\mathbf{r})$ in Eq. (1) defines the area of the receiving aperture, i.e.,

$$W_D(\mathbf{r}) = \begin{cases} 1 & \text{if } |\mathbf{r}| \leqslant \dfrac{D}{2} \\ 0 & \text{if } |\mathbf{r}| > \dfrac{D}{2} \end{cases} \tag{2}$$

In writing Eq. (1), it was assumed that the mixing process took place at the receiver input aperture. Alternatively, the heterodyne process can be described by the diffraction patterns of the signal and LO over the receiver focal plane. These two descriptions are completely equivalent for systems employing perfect receiving optics. For systems employing less-than-perfect optics, however, the description given by Eq. (1) must be modified to include the optics-induced perturbations in the signal.

The noise current $n(t)$ at the receiver output consists of the signal shot noise and the detector thermal noise. In the limit of a strong LO, the shot noise term usually dominates the receiver thermal noise. For all practical purposes, this LO shot noise can be modeled as an additive white Gaussian noise (AWGN) with power spectral density

$$N_s = \frac{e^2 \eta G^2}{h\nu} P_{LO} = \left(\frac{e^2 \eta G^2}{h\nu}\right) \frac{1}{2z_0} \int W_D(\mathbf{r}) |\vec{A}_{LO}(\mathbf{r})|^2 \, d\mathbf{r} \tag{3}$$

where P_{LO} is the LO power. As a result, the optical heterodyne channel can often be modeled effectively as a Gaussian channel [1]. Demodulation techniques for signals in the presence of white Gaussian noise have been studied extensively [4]. In general, the IF signal can be demodulated either noncoherently or coherently. Noncoherent demodulation is often accomplished with the use of an envelope detector shown in Fig. 2. The output of the envelope detector can be written as

$$u = \sqrt{\frac{1}{2\sigma_N^2} \left[(A_c + n_c)^2 + (A_s + n_s)^2 \right]} \tag{4a}$$

where

$$A_c = \left(\frac{e\eta G}{h\nu z_0}\right) \int W_D(\mathbf{r}) \vec{A}_S(\mathbf{r}) \times \vec{A}_{LO}(\mathbf{r}) \cos\left(\phi_S(\mathbf{r}) - \phi_{LO}(\mathbf{r})\right) d\mathbf{r} \tag{4b}$$

$$A_s = \left(\frac{e\eta G}{h\nu z_0}\right) \int W_D(\mathbf{r}) \vec{A}_S(\mathbf{r}) \times \vec{A}_{LO}(\mathbf{r}) \sin\left(\phi_S(\mathbf{r}) - \phi_{LO}(\mathbf{r})\right) d\mathbf{r} \tag{4c}$$

and n_c, n_s are the Gaussian quadrature noise components with zero mean and variance equal to $\sigma_N^2 = 2N_s/T_s$. The probability distribution of the detected envelope is known to be Rician distributed with density function [4]

$$p_1(u) = 2u e^{-(\beta + u^2)} I_0\left(2u\sqrt{\beta}\right) \tag{5}$$

where

$$\beta = \frac{A_c^2 + A_s^2}{2\sigma_N^2}$$

$$= \left(\frac{e\eta P_S T_s}{h\nu}\right) \frac{1}{P_S P_{LO} \left(4z_0^2\right)} \iint W_D(\mathbf{r}) W_D(\mathbf{r}') \times [\vec{A}_S(\mathbf{r}) \times \vec{A}_{LO}(\mathbf{r})] [\vec{A}_S(\mathbf{r}') \times \vec{A}_{LO}(\mathbf{r}')]$$

$$\times \cos\left(\phi_S(\mathbf{r}) - \phi_{LO}(\mathbf{r}) - \phi_S(\mathbf{r}') + \phi_{LO}(\mathbf{r}')\right) d\mathbf{r} d\mathbf{r}'$$

$$\equiv \eta_{het} \left(\frac{e\eta P_S T_s}{h\nu}\right) \tag{6}$$

The factor η_{het} in Eq. (6) is known as the heterodyne efficiency of the receiver, and the quantity β is the signal-to-noise ratio of the envelope detector which can be interpreted as the number of signal photons incident on the receiver aperture multiplied by the heterodyne efficiency. For systems with perfect spatial mode matching, the heterodyne efficiency is equal to 1. When the spatial modes are not properly matched, the contribution to the IF signal from different parts of the receiver aperture can interfere destructively and result in reduced heterodyne efficiency.

In addition to the noncoherent demodulation scheme, the IF signal can also be coherently demodulated [4]. In this scheme, a local reference carrier is synchronized to the IF carrier. The received IF signal is then mixed with the reference carrier and the resulting baseband signal is matched-filter detected. In the absence of time-varying perturbations in the signal amplitude and phase, the coherent demodulation scheme can result in superior receiver performance. However, the coherent demodulator requires the generation of a local reference signal that is synchronized to the incoming IF carrier and is therefore more complicated. Furthermore, because of the rapidly varying atmospheric condition and the weak signal power expected over a deep-space link, carrier phase tracking can be very difficult to accomplish. The use of the noncoherent demodulation scheme eliminates the need to retrieve the IF signal phase and hence considerably simplifies the design of the receiver. Consequently, only the noncoherently demodulated heterodyne receiver will be included in the present analysis.

III. Noncoherently Combined Heterodyne Receiver

The performance of a single-aperture heterodyne receiver has been studied for free-space communication applications [5]. In some cases, achievement of an effective communication link requires the use of a large effective receiving aperture. However, the high cost of constructing a single monolithic-aperture heterodyne receiver with a large collecting area can be prohibitive because of the stringent demand on the quality of the optical surface. Furthermore, high-bandwidth adaptive optics must be used to compensate for the atmospherically induced wavefront distortion if reception inside Earth's atmosphere is desired. Such a measure can further increase the receiver cost. Alternatively, the performance of the communications link can be improved by electronically combining the output currents from several spatially separated heterodyne receivers. The block diagram of such a receiver is shown in Fig. 3. For the analysis, it will be assumed that an array of identical receivers is used to detect the incoming signal. These receivers can share a common support structure to reduce the construction cost.

Because the individual receivers are identical, the output signals from these receivers can be assumed to be identically distributed. Furthermore, the LO shot noise can be modeled as independently distributed. In the limit where the number of receivers is large, the Central Limit Theorem can be used to model the combiner output as Gaussian distributed. The error performance of this combined receiver can therefore be characterized by the combiner signal-to-noise ratio ρ_c which is given by

$$\rho_c = \frac{(E[s|\text{signal}] - E[s|\text{no signal}])^2}{\text{var}(s|\text{signal}) + \text{var}(s|\text{no signal})} \tag{7}$$

In Eq. (7), $s \equiv \Sigma u_i$ is the sum of individual demodulator outputs, and $E[s|X]$ and $\text{var}(s|X)$ denote the conditional mean and variance of the variable s, subjected to the condition X, respectively. For a binary Gaussian channel with SNR ρ_c, the bit error rate is [4]

$$BER = \frac{1}{2}\,\text{erfc}\left(\sqrt{\frac{\rho_c}{2}}\right) \tag{8}$$

Since the demodulator outputs $\{u_i\}$ are modeled as independent and identically distributed random variables, the mean and variance of s can be evaluated by summing the mean and variance of the individual demodulator outputs, and the combiner SNR can be simplified to

$$\rho_c = N \times \frac{(E[u|\text{signal}] - E[u|\text{no signal}])^2}{\text{var}(u|\text{signal}) + \text{var}(u|\text{no signal})} \tag{9}$$

where N is the number of combined subaperture receivers. Equation (9) shows that for systems employing an array of identical heterodyne receivers, the performance improves with increasing number of receivers.

By integrating with respect to the probability density function given in Eq. (5), the first two moments of the noncoherent demodulator output can be written as

$$E[u] = \frac{\sqrt{\pi}}{2} e^{-\beta} {}_1F_1(3/2,1,\beta) \tag{10}$$

$$E[u^2] = (1+\beta) \tag{11}$$

The function ${}_1F_1(a,b,c)$ is the degenerate hypergeometric function. The combiner SNR can then be evaluated by substituting the demodulator mean and variance calculated from Eqs. (10) and (11) into Eq. (9). For $\beta \ll 1$, the degenerate hypergeometric function in Eq. (10) can be expanded in a power series:

$$E[u] \approx \frac{\sqrt{\pi}}{2}\left[1 + \frac{\beta}{2} - \frac{\beta^2}{16}\right], \qquad \beta \ll 1 \tag{12a}$$

and the resulting combiner SNR is given by

$$\rho_c = N \times \frac{\frac{\pi}{4}\left[\frac{\beta}{2} - \frac{\beta^2}{16}\right]^2}{\left(1 - \frac{\pi}{4}\right)(2+\beta) - \frac{3\pi}{64}\beta^2}, \qquad \beta \ll 1 \tag{12b}$$

When $\beta \gg 1$, on the other hand, an approximation to the mean demodulator output can be found by Taylor expanding the square root in Eq. (4a) and taking the expectation:

$$E[u] \approx \sqrt{(1+\beta)}\left[1 - \frac{1+2\beta}{8(1+\beta)^2}\right], \qquad \beta \gg 1 \tag{13a}$$

The resulting combiner SNR is therefore

$$\rho_c \approx N \times \frac{\left[\sqrt{1+\beta}\left(1 - \frac{(1+2\beta)}{8(1+\beta)^2}\right) - \frac{\sqrt{\pi}}{2}\right]^2}{1 - \frac{\pi}{4} + \frac{1+2\beta}{4(1+\beta)}}, \qquad \beta \gg 1 \tag{13b}$$

Equations (12b) and (13b) show that for a system consisting of a fixed number of receivers, the performance can be improved by increasing the SNR of the individual receivers. This can be accomplished by increasing the aperture of the individual receivers or by increasing the signal power. For a fixed size of individual receiver apertures and incoming signal intensity, the combiner SNR increases linearly with the number of receivers N. However, for a system with a fixed overall collecting area, subdividing the receiving area into several smaller receivers can result in a corresponding decrease of the individual receiver SNR. Therefore, even though the combiner SNR depends explicitly on the number of receivers N, the reduction in β with increasing N will actually lead to a reduction in the combiner SNR. This fact is demonstrated in Fig. 4 where the combining loss, which is the ratio of the combiner SNR to the SNR of a single monolithic aperture of equivalent size, is plotted against the number of receivers. The figure was generated with no turbulence-induced degradation taken into account. The parameter ρ_0 is the SNR of a monolithic receiver.

Note that at small N, the combiner SNR decreases slowly with an increasing number of aperture segments. When the amount of signal power received over each aperture decreases, however, the efficiency of the noncoherent demodulator decreases and the combining loss increases rapidly. This is because of the nonlinear nature of the noncoherent demodulation process. At very low signal powers, the noise contribution to the demodulator output is more significant and, as a result, the efficiency of the demodulator decreases with decreasing signal power. In other words, given the overall receiving area, there is a penalty for subdividing the aperture and then noncoherently combining the demodulator outputs. For a system with ρ_0 = 20 dB, the noncoherent combining loss when the aperture is subdivided into 1000 receivers is approximately 15 dB. This combining loss decreases with increasing signal power. It should be noted that this combining loss is due to the noncoherent demodulation process. For systems employing coherent combining schemes, the combiner performance will depend only on the total collecting area, and thus will not degrade with an increasing number of receivers.

IV. Atmospheric Turbulence Effect

The above analysis showed that, in the absence of atmospheric turbulence, the performance of a noncoherently combined optical heterodyne receiver is inferior to that of a monolithic-aperture system. In the presence of atmospheric turbulence, both the amplitude and phase of the incoming signal wavefront will be distorted. Under the condition of weak turbulence, the log amplitude $\ell(\mathbf{r})$ and the phase $\phi(\mathbf{r})$ of the optical signal, after propagating through the atmosphere, can be modeled as stationary, Gaussian random processes [5]. The autocorrelation of the amplitude and phase can be characterized by the atmospheric structure function $\mathscr{D}(\mathbf{r},\mathbf{r}')$ which can be written as

$$\begin{aligned}\mathscr{D}(\mathbf{r},\mathbf{r}') &= \mathscr{D}_\ell(\mathbf{r},\mathbf{r}') + \mathscr{D}_\phi(\mathbf{r},\mathbf{r}') \\ &= \langle |\ell(\mathbf{r}) - \ell(\mathbf{r}')|^2 \rangle + \langle |\phi(\mathbf{r}) - \phi(\mathbf{r}')|^2 \rangle\end{aligned} \tag{14}$$

The angle brackets in Eq. (14) denote the ensemble average. Tatarski [6] showed that, under the condition of weak turbulence, the structural function of the atmosphere can be characterized by its phase coherence length r_0 as

$$\mathscr{D}(\mathbf{r},\mathbf{r}') = 6.88(|\mathbf{r}-\mathbf{r}'|/r_0)^{5/3} \tag{15}$$

Under normal viewing conditions, r_0 is typically between 5 and 30 cm, and at a few outstanding sites, such as Mauna Kea, an r_0 in excess of 40 cm can occasionally be observed.

The effect of turbulence on the performance of a heterodyne receiver is to reduce the heterodyne efficiency. Since the atmospheric condition varies dynamically, the instantaneous IF signal amplitude and phase vary continuously. If the integration time is much shorter than the characteristic time in which the atmospheric properties change significantly, however, it is reasonable to model the atmospheric turbulence as "frozen" over the integration period. In this case the time dependence of the signal phase and amplitude can be ignored, and the receiver SNR can be described adequately using Eq. (6). The expression for the SNR can be further simplified by noting that, under the condition of weak turbulence, the structural function $\mathscr{D}(\mathbf{r},\mathbf{r}')$ is dominated by the phase structural function $\mathscr{D}_\phi(\mathbf{r},\mathbf{r}')$. Over a sufficiently large aperture, the fluctuation in signal amplitude (scintillation) will be averaged so that its effect on the receiver SNR can be ignored. Under these assumptions, the SNR of the noncoherent demodulator in Eq. (6) can be written as

$$\beta = (\beta_0 A_R)\frac{1}{A_R^2}\iint W_D(\mathbf{r})W_D(\mathbf{r}')e^{i(\phi_s(\mathbf{r})-\phi_s(\mathbf{r}'))}d\mathbf{r}d\mathbf{r}' \tag{16}$$

where $\beta_0 = (\eta P_S T_s/h\nu A_R)$ is the number of signal photons received per unit area or, equivalently, the receiver SNR per unit area. In writing Eq. (16), it has been assumed that the LO can be approximated by a plane wave and that the amplitude fluctuation of the signal can be ignored.

The statistical properties of the heterodyne SNR inside the atmosphere were first studied by Fried [7],[8]. By taking the expectation of Eq. (16) with respect to the signal phase and using the approximation that the signal phase is a zero-mean, Gaussian random process with structure function $\mathscr{D}_\phi(\mathbf{r},\mathbf{r}')$, Fried successfully calculated the average receiver SNR $\bar{\beta}$ in the presence of turbulence as [7]

$$\bar{\beta} = \beta_0\frac{\pi r_0^2}{4}\psi(D/r_0) \tag{17}$$

where

$$\psi(x) = \frac{32x^2}{\pi} \int_0^1 \frac{u}{2}[\cos^{-1} u - u(1-u^2)^{1/2}]\, e^{-3.44x^{5/3}u^{5/3}}\, du \tag{18}$$

Figure 5 is a plot of the function $\psi(D/r_0)$ versus the diameter of the receiver aperture D/r_0. For $D/r_0 \ll 1$, the function $\psi(D/r_0)$ increases as the square of the aperture diameter. For $D/r_0 \gg 1$, the value of $\psi(D/r_0)$ approaches 1 asymptotically. Note that when $D = r_0$, the SNR is only 3 dB below what can be achieved with $D \rightarrow \infty$. As a result, a further increase in the aperture size will only result in a marginal increase in the receiver SNR, and very little can be gained by increasing the diameter of the receiver aperture beyond r_0.

Equation (17) shows that the limiting performance of a single monolithic-aperture heterodyne receiver inside the atmosphere is equivalent to that of a single aperture with diameter r_0 in the absence of turbulence. In addition to limiting the average SNR, the presence of atmospheric turbulence also induces a severe fluctuation in the detected SNR. By squaring the expression for β in Eq. (16) and taking the expectation with respect to the signal phase, the mean-square receiver SNR can be written as

$$\langle \beta^2 \rangle = \beta_0^2 \frac{1}{A_R^2} \int W_D(\mathbf{r}_1) W_D(\mathbf{r}_2) W_D(\mathbf{r}_3) W_D(\mathbf{r}_4) \times \langle e^{i\left(\phi_s(\mathbf{r}_1)-\phi_s(\mathbf{r}_2)+\phi_s(\mathbf{r}_3)-\phi_s(\mathbf{r}_4)\right)} \rangle \times d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 d\mathbf{r}_4 \tag{19}$$

Equation (19) is very difficult to evaluate in a closed form. However, by applying Fried's approximation, a simple upper bound on the mean square SNR can be given by (Appendix A)

$$\langle \beta^2 \rangle \leqslant e^{0.461(D/r_0)^{5/3}} \beta_0^2 \left(\frac{\pi r_0}{4}\right)^2 \left(\frac{512}{\pi^3}\right) \times \left[\left(\frac{D}{r_0}\right)^4 \int K(w) w e^{-(D/r_0)^{5/3} w^2}\, dw\right] \tag{20a}$$

where

$$K(w) = \int_0^P dp \int_0^Q dq f\left(p + \tfrac{1}{2} w/D, q\right) f\left(p - \tfrac{1}{2} w/D, q\right)$$

$$P = 1 - w/2D$$

$$Q = [1 - (w/2D)^2]^{1/2}$$

$$f(x,y) = \cos^{-1}[(x^2+y^2)^{1/2}] - (x^2+y^2)^{1/2}[1-(x^2+y^2)]^{1/2} \tag{20b}$$

Equation (20a) can be evaluated numerically. The result, which is shown in Fig. 6, provides an upper bound for the mean square SNR fluctuation. Note that the upper bound diverges exponentially for $D/r_0 \geqslant 1$. Also shown in Fig. 6 are the results of a Monte Carlo simulation of the mean square SNR. In contrast to the diverging upper bound, the simulation data show that the mean square SNR actually converges at $D/r_0 \gg 1$. This convergence can be argued as follows: for $|\mathbf{r}_i - \mathbf{r}_j| \gg r_0$, the phase of the incoming signal at $\mathbf{r}_i$ and $\mathbf{r}_j$ will be completely uncorrelated so that the expectation of the exponent in Eq. (19) is negligible. By applying a simple geometrical argument, it can be reasoned that the integrand in Eq. (19) is negligible except when $\mathbf{r}_1 \approx \mathbf{r}_2$ and $\mathbf{r}_3 \approx \mathbf{r}_4$, or when $\mathbf{r}_1 \approx \mathbf{r}_3$ and $\mathbf{r}_2 \approx \mathbf{r}_4$. Consequently, only two of the four spatial integrals increase with increasing D, and the integral in Eq. (19) increases only as D^4. As a result, the mean square SNR converges as $D \rightarrow \infty$.

When the effect of turbulence is taken into account, the performance of the noncoherently combined heterodyne receiver depends on the actual probability distribution of β. Given $P_\beta(\beta)$, the probability distribution of β, the first two moments of the noncoherent demodulator can be calculated by averaging Eqs. (9) and (10) over the distribution of β as

$$E[u] = \int_0^\infty \frac{\sqrt{\pi}}{2} e^{-\beta} {}_1F_1(3/2,1,\beta) P_\beta(\beta)\, d\beta \tag{21a}$$

$$E[u^2] = \int_0^\infty (1+\beta) P_\beta(\beta)\, d\beta = 1 + \bar{\beta} \tag{21b}$$

A detailed description of the distribution of β is needed to evaluate these integrals. Unfortunately, due to the complexity of the expression shown in Eq. (16), the statistical properties for β are very difficult to characterize. However, when the variance of β is small, a simple approximation for the mean demodulator output can be found by expanding the conditional expectation of the demodulator output in a Taylor series (Appendix B)

$$E[u] \equiv \int_0^\infty \frac{\sqrt{\pi}}{2} e^{-\beta} {}_1F_1(3/2,1,\beta) P_\beta(\beta)\, d\beta$$

$$\approx \frac{\sqrt{\pi}}{2} e^{-\bar{\beta}} {}_1F_1(3/2,1,\bar{\beta}) - \frac{\sqrt{\pi}}{16} e^{-\bar{\beta}} {}_1F_1(3/2,3,\bar{\beta})\, \sigma_\beta^2 \tag{22a}$$

Because the second derivative of the conditional average of the demodulator output is always negative, it can also be shown that the average demodulator output is bounded from above by (Appendix B):

$$E[u] < \frac{\sqrt{\pi}}{2} e^{-\bar{\beta}} {}_1F_1(3/2,1,\bar{\beta}) \tag{22b}$$

For the analysis of noncoherently combined heterodyne receivers inside the atmosphere, it will be assumed that the individual receivers are spatially separated by a distance much greater than the phase coherence length r_0 of the atmosphere so that the output of these receivers can be modeled as independent and identically distributed. Therefore, by substituting the expressions for the first two moments of the demodulator output from Eqs. (21b) and (22b) into Eq. (7), an estimate of the combiner SNR can be found for a noncoherently combined heterodyne receiver inside atmosphere. Since Eq. (22b) presents an upper bound for the mean demodulator output, it follows from Eq. (7) that the combiner SNR derived is an upper bound for the actual combiner SNR. Similarly, by substituting Eqs. (21a) and (22a) into Eq. (7), an approximation to the combiner SNR can be derived. If we further substitute in the upper bound of the variance of β given in Eq. (20), the resulting combiner SNR will be an approximate lower bound of the actual combiner performance.

Given the expression for the combiner SNR, the performance of the combined receiver can be investigated. Figure 7 is a plot of the combiner SNR versus the total receiving area for a noncoherently combined heterodyne receiver with several values of N, the number of receivers. The incoming signal intensity β_0 and the atmospheric coherence length r_0 are fixed. Also plotted in the figure is the SNR for a coherently combined heterodyne receiver with no turbulence. This curve represents the best achievable performance for a heterodyne receiver at a given receiving area. Note that when the total collecting area is small, the performance of the combined system degrades with an increase in the number of receivers. This is because the efficiency of the noncoherent demodulation process degrades rapidly with decreasing signal power. However, as the diameter of the receiver aperture increases beyond r_0, the SNR of the single-aperture receiver begins to saturate because of the atmospheric turbulence. On the other hand, for systems employing a number of receivers, the combiner SNR continues to improve until the diameter of each receiver is greater than r_0. Therefore, for a system which requires a large collecting area, the performance can indeed be improved by combining the outputs from several receivers. It should be noted, however, that despite the improvement in combiner performance with an increasing number of receivers, the performance of the noncoherently combined receiver is still far inferior to that of a coherently combined system. In fact, for the given signal intensity assumed in Fig. 7, the noncoherently combined receiver suffers more than 20 dB loss in the receiver SNR.

It is interesting to evaluate the combining loss versus the number of receivers. Plotted in Fig. 8 is the combining loss of a noncoherently combined heterodyne receiver versus the number of receivers. The total collecting aperture is fixed so that an increase in the number of receivers corresponds to a decrease in the diameter of the individual receiver apertures. For systems employing a large number of small-aperture receivers, the loss due to noncoherent demodulation dominates so that the combining loss increases with an increasing number of receivers. For systems using a small number of receivers, on the other hand, the SNR of each receiver is limited by the coherence length of the atmosphere, and hence the receiver performance improves with an increasing number of receivers. Consequently, given the overall collecting area and the phase coherence length r_0, an optimal number of receivers can be found which maximizes the combiner SNR. It should be noted, however, that Fig. 8 was generated by fixing the total collection area of the combined receiver. When the size of the individual receiver is fixed, it follows directly from Eq. (7) that the performance of the combined receiver improves linearly with N. It is also interesting to investigate the dependence of combining loss with different values of r_0, the phase coherence length. Figure 9 is a plot of the combining loss versus the diameter of the individual aperture at different values of r_0. The total collecting area of the combined receiver is again assumed to be equivalent to that of a monolithic receiver with a 10-meter diameter. It is seen that the performance of the combined receiver is optimized when the diameter of the individual receiver D is approximately equal to the phase coherence length r_0. The combining loss increases rapidly when the size of the individual aperture becomes much larger than r_0.

V. Conclusions

Combining outputs from a number of small-aperture heterodyne receivers can significantly reduce the cost of constructing a receiver with a large collecting area. In principle, output signals from these small-aperture receivers can be combined to improve the detection statistics. Compared to systems employing a single, diffraction-limited aperture, however, the noncoherently combined receiver suffers from the combining loss due to the noncoherent demodulation process. This combining loss increases rapidly with an increasing number of aperture segments. The loss due to noncoherent combining is worse for systems with weak incoming signals. For systems with a large number of segments ($N \approx 100$), the combining loss is on the order of 15–25 dB for systems with an effective single-aperture SNR of 10–20 dB. The combining loss is smaller for systems with higher signal powers.

When the atmospheric turbulence is taken into consideration, however, the performance of a monolithic-aperture heterodyne receiver is limited by the size of the phase-coherence cell of the atmosphere. As a result, increasing the size of a monolithic aperture will not result in a corresponding improvement in system performance. For such a system, the noncoherently combined receiver can provide a significant improvement in system performance. The analysis showed that, given the size of the overall collecting aperture, the performance of the receiver is optimized when the diameter of the individual subaperture is equal to the phase coherence length of the atmosphere r_0. Increasing the size of the individual aperture beyond r_0 can result in a saturation of individual SNR. On the other hand, reducing the subaperture to a size smaller than r_0 will increase the combining loss and hence degrade the system performance.

It should be noted that for large N, the combining loss of the noncoherently combined heterodyne receiver is due primarily to the inefficiency in the noncoherent demodulation process. It was assumed throughout this article that the coherent combining is ineffective due to the weak signal power and the rapidly varying atmospheric condition. For systems in which this assumption can be relaxed, the coherent combining scheme can be used to improve receiver performance significantly. One such system is the large-aperture heterodyne receiver in space. Although the alignment errors between individual receivers can induce random phase shifts between the detector outputs, these phase errors are fixed and can therefore be accurately measured and properly compensated in the absence of rapidly varying atmospheric conditions. Another scheme is to use the adaptive optic technique to correct for the atmospherically induced wavefront distortion. The required high signal power can be provided through the use of natural or artificial guide stars [3]. Receiver performance comparable to that of a monolithic, diffraction-limited receiver can be obtained, at least in principle, with the use of a coherent signal combining scheme.

Acknowledgments

The author acknowledges Dr. James R. Lesh for his helpful discussions, and Mr. Moe Zaw Win for carrying out the computer simulation.

References

[1] R. Gagliardi and S. Karp, *Optical Communications*, New York: Wiley, 1975.

[2] J. W. Hardy, "Active Optics: A New Technology for the Control of Light," *Proc. IEEE*, vol. 66, no. 6, pp. 651–697, June 1978.

[3] L. A. Thompson and C. S. Gardner, "Experiment on Laser Guide Star at Mauna Kea Observatory for Adaptive Imaging in Astronomy," *Nature*, vol. 328, no. 6127, pp. 229–231, July 16, 1987.

[4] W. Lindsey and M. K. Simon, *Telecommunication Systems Engineering*, Englewood Cliffs, New Jersey: Prentice-Hall, 1973.

[5] J. E. Kaufmann and L. L. Jeromin, "Optical Heterodyne Intersatellite Links Using Semiconductor Lasers," presented at the IEEE GLOBECOM'84, Atlanta, Georgia, November 27–29, 1984.

[6] V. I. Tatarski, *Wave Propagation in a Turbulent Medium*, New York: McGraw-Hill, 1961.

[7] D. L. Fried, "Optical Heterodyne Detection of an Atmospherically Distorted Signal Wave Front," *Proc. IEEE*, vol. 55, no. 1, pp. 57–67, January 1967.

[8] D. L. Fried, "Atmospheric Modulation Noise in an Optical Heterodyne Receiver," *IEEE J. Quantum Electron.*, vol. QE-3, no. 6, pp. 213–221, 1967.

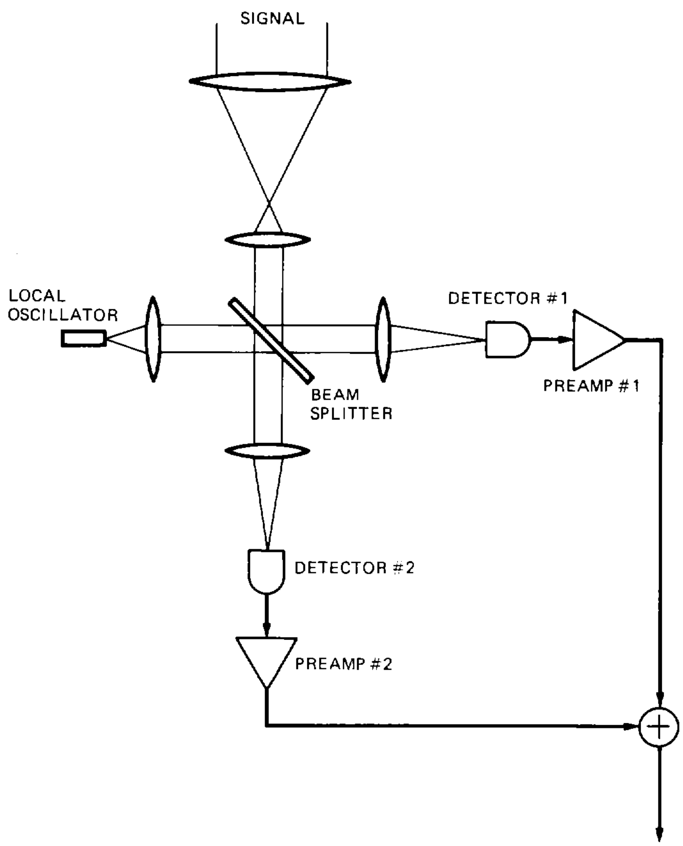

Fig. 1. Block diagram of a dual-detector heterodyne receiver

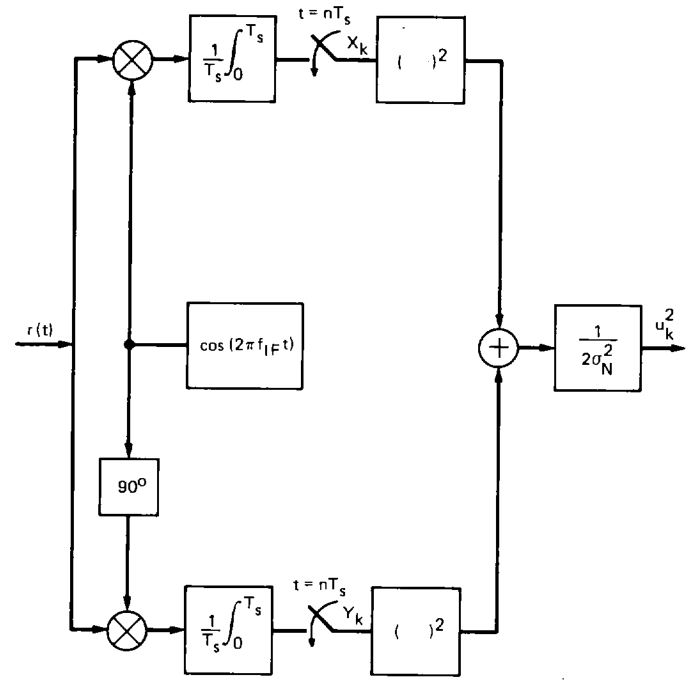

Fig. 2. Structure of a noncoherent envelope demodulator

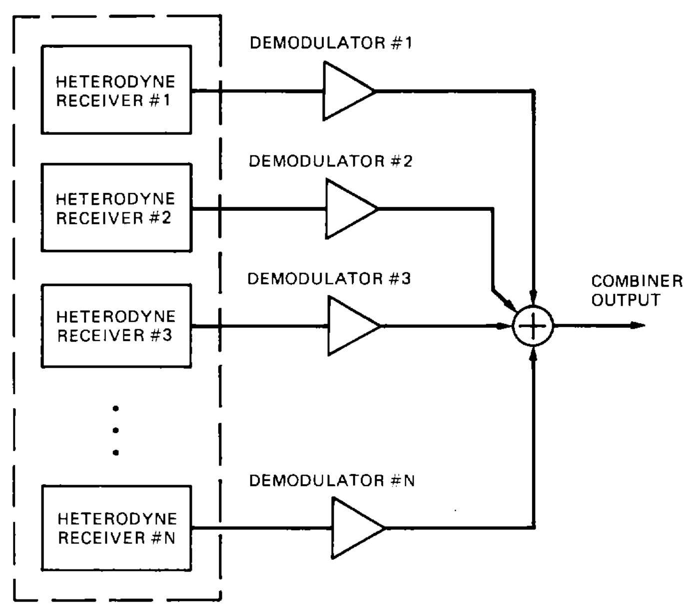

Fig. 3. Block diagram of a noncoherently combined heterodyne receiver

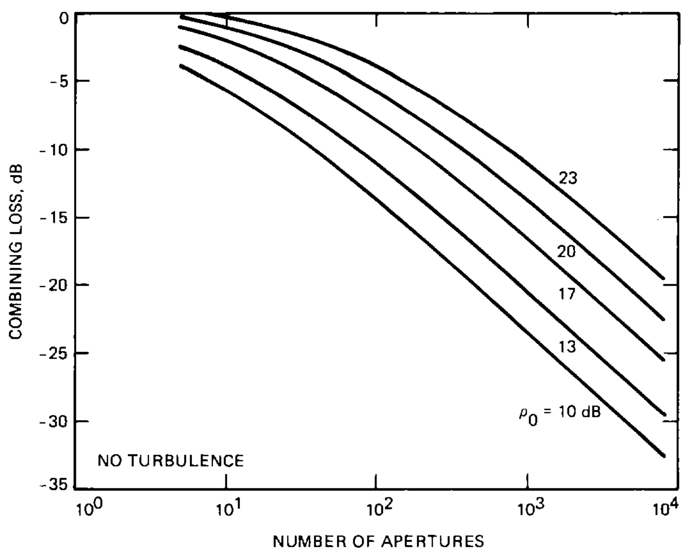

Fig. 4. Combining loss of a noncoherently combined heterodyne receiver with a fixed overall aperture size

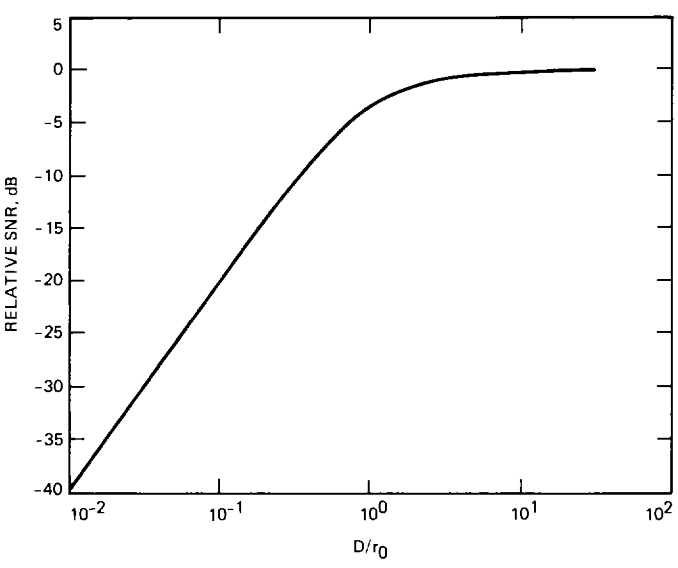

Fig. 5. The value of $\psi(D/r_0)$ versus the diameter of the aperture D/r_0

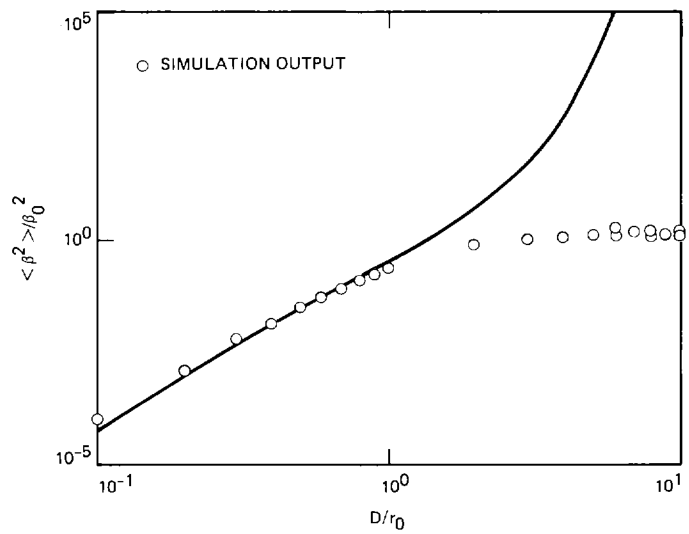

Fig. 6. Mean square value of the demodulator SNR $<\beta^2>$ versus the diameter of the receiver aperture. Also plotted are the mean square values calculated from the Monte Carlo simulations.

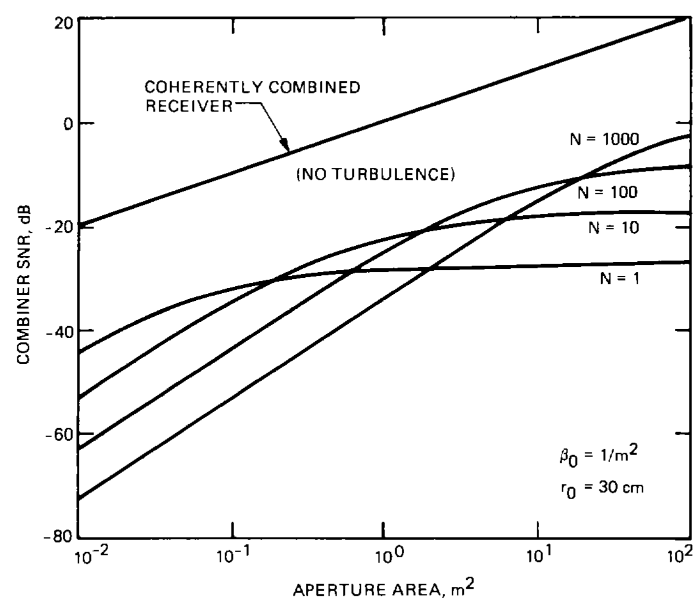

Fig. 7. Combiner SNR versus the total area of the collecting aperture for several values of N, the number of receivers. Also plotted is the SNR of a coherently combined receiver of equivalent aperture area.

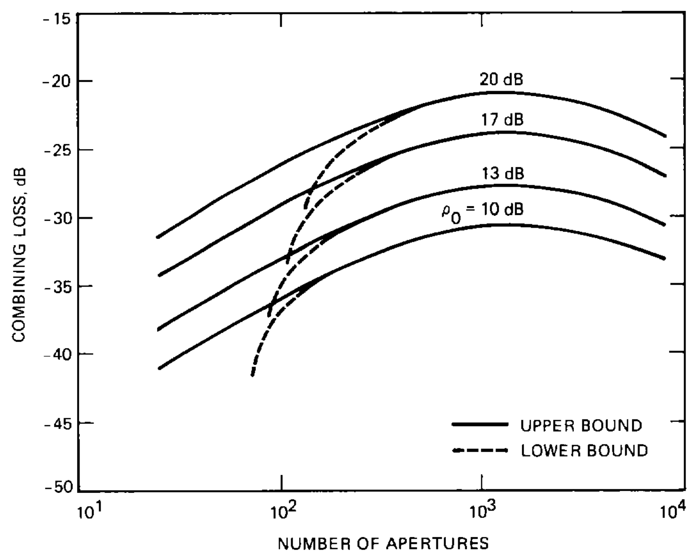

Fig. 8. Combining loss of a noncoherently combined heterodyne receiver with the saturation effect of the atmosphere taken into account. The total collecting area of the receiver is equivalent to a 10-m-diameter monolithic aperture, and the atmospheric coherence length r_0 is 30 cm.

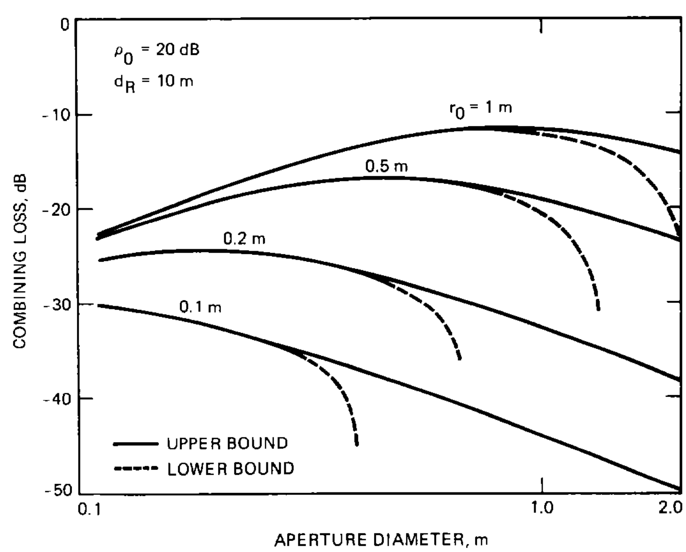

Fig. 9. Combining loss versus the diameter of the individual aperture for a system with an equivalent aperture diameter of 10 meters

Appendix A
Derivation of the Mean Square Receiver SNR

Given the expression of β in Eq. (16), the mean square SNR, $\langle \beta^2 \rangle$, can be written as

$$\langle \beta^2 \rangle = \beta_0^2 \frac{1}{A_R^2} \int W_D(\mathbf{r}_1) \ldots W_D(\mathbf{r}_4)$$

$$\times \langle e^{i(\phi_S(\mathbf{r}_1)-\phi_S(\mathbf{r}_2)+\phi_S(\mathbf{r}_3)-\phi_S(\mathbf{r}_4))} \rangle$$

$$\times\, d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 d\mathbf{r}_4 \qquad \text{(A-1)}$$

The expectation over $\phi_S(\mathbf{r}_1)$ can be evaluated by using the assumption that $\Phi_S(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3,\mathbf{r}_4) = \phi_S(\mathbf{r}_1) - \phi_S(\mathbf{r}_2) + \phi_S(\mathbf{r}_3) - \phi_S(\mathbf{r}_4)$ is a zero-mean Gaussian random variable. The resulting expectation value of the exponential is given by

$$\langle e^{i\Phi_S(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3,\mathbf{r}_4)} \rangle = e^{-\sigma_\Phi^2/2} \qquad \text{(A-2)}$$

where σ_Φ^2 is the variance of $\Phi_S(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3,\mathbf{r}_4)$. Equation (A-2) can be further simplified by realizing that

$$(a-b+c-d)^2 = (a-b)^2 + (a-d)^2 + (b-c)^2 + (c-d)^2$$

$$- (a-c)^2 - (b-d)^2 \qquad \text{(A-3)}$$

and that the expectation of the phase-difference square is simply the phase structural function, i.e., $\mathscr{D}_\phi(\mathbf{r}_1 - \mathbf{r}_2) = \langle |\phi_S(\mathbf{r}_1) - \phi_S(\mathbf{r}_2)|^2 \rangle$. By substituting Eq. (A-2) into Eq. (A-1), the mean-square SNR can be written as

$$\langle \beta^2 \rangle = \beta_0^2 \frac{1}{A_R^2} \int W_D(\mathbf{r}_1) W_D(\mathbf{r}_2) W_D(\mathbf{r}_3) W_D(\mathbf{r}_4)\, e^{\left\{1/2[-\mathscr{D}_\phi(|\mathbf{r}_1-\mathbf{r}_2|)-\mathscr{D}_\phi(|\mathbf{r}_3-\mathbf{r}_4|)]\right\}}$$

$$\times\, e^{\left\{1/2[-\mathscr{D}_\phi(|\mathbf{r}_1-\mathbf{r}_4|)-\mathscr{D}_\phi(|\mathbf{r}_2-\mathbf{r}_3|)+\mathscr{D}_\phi(|\mathbf{r}_1-\mathbf{r}_3|)+\mathscr{D}_\phi(|\mathbf{r}_2-\mathbf{r}_4|)]\right\}}\, d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 d\mathbf{r}_4$$

$$= \beta_0^2 \left(\frac{16D^4}{\pi^2}\right) \int W_1(\mathbf{u}_1) W_1(\mathbf{u}_2) W_1(\mathbf{u}_3) W_1(\mathbf{u}_4)\, e^{3.44(D/r_0)^{5/3}[-|\mathbf{u}_1-\mathbf{u}_2|^{5/3}-|\mathbf{u}_3-\mathbf{u}_4|^{5/3}]}$$

$$\times\, e^{3.44(D/r_0)^{5/3}[-|\mathbf{u}_1-\mathbf{u}_4|^{5/3}-|\mathbf{u}_2-\mathbf{u}_3|^{5/3}+|\mathbf{u}_1-\mathbf{u}_3|^{5/3}+|\mathbf{u}_2-\mathbf{u}_4|^{5/3}]}\, d\mathbf{u}_1 d\mathbf{u}_2 d\mathbf{u}_3 d\mathbf{u}_4$$

(A-4)

where we have carried out the substitution of variables, $\mathbf{r}_i = D\mathbf{u}_i$, and applied the 5/3 power law of the phase structural function to factor out the term $(D/r_0)^{5/3}$.

The mean square SNR given in Eq. (A-4) is very difficult to evaluate numerically. However, a simple upper bound of the integral can be obtained by using the fact that since $|\mathbf{u}_i| < 1/2$, $|\mathbf{u}_i - \mathbf{u}_j| = u_{ij} \leqslant 1$, and the 5/3 power of u_{ij} can be bounded from above and below by

$$u_{ij}^2 \leqslant u_{ij}^{5/3} \leqslant u_{ij}^2 + \left[\left(\frac{5}{6}\right)^5 - \left(\frac{5}{6}\right)^6\right] = u_{ij}^2 + 0.06698$$

$$0 \leqslant u \leqslant 1 \qquad \text{(A-5)}$$

The bounds in Eq. (A-5) are derived by solving for the maximum and minimum of the function $y(x) = x^{5/3} - x^2$. By substituting the upper bound into Eq. (A-4) for terms with positive exponents and the lower bound for terms with the negative exponent, a simple upper bound for the mean square SNR can be easily derived,

$$\langle \beta^2 \rangle \leqslant e^{0.461(D/r_0)^{5/3}} \beta_0^2 \left(\frac{\pi r_0^2}{4}\right)^2 \phi(D/r_0) \quad \text{(A-6)}$$

where

$$\phi(D/r_0) = \frac{256}{\pi^4}\left(\frac{D}{r_0}\right)^4 \int W_1(\mathbf{u}_1)\ldots W_1(\mathbf{u}_4)$$

$$\times\, e^{[-(D/r_0)^{5/3}|\mathbf{u}_1-\mathbf{u}_2+\mathbf{u}_3-\mathbf{u}_4|^2]}\, d\mathbf{u}_1 d\mathbf{u}_2 d\mathbf{u}_3 d\mathbf{u}_4 \quad \text{(A-7)}$$

Since the integral involves only a single variable of the form $\mathbf{w} = |\mathbf{u}_1 - \mathbf{u}_2 + \mathbf{u}_3 - \mathbf{u}_4|^2$, the 4-fold integral in Eq. (A-7) can be collapsed into a single one-dimensional integral of the form [7]:

$$\int W_1(\mathbf{u}_1)\ldots W_1(\mathbf{u}_4) e^{-t^{5/3}|\mathbf{u}_1-\mathbf{u}_2+\mathbf{u}_3-\mathbf{u}_4|^2}\, d\mathbf{u}_1 \ldots d\mathbf{u}_4 =$$

$$2\pi \int K(w) w e^{-t^{5/3} w^2}\, dw \quad \text{(A-8)}$$

where

$$K(w) = \int_0^P dp \int_0^Q dq f\left(p + \frac{1}{2}w/D, q\right) f\left(p - \frac{1}{2}w/D, q\right)$$

$$P = 1 - w/2D$$

$$Q = [1 - (w/2D)^2]^{1/2}$$

$$f(x,y) = \cos^{-1}[(x^2+y^2)^{1/2}] - (x^2+y^2)^{1/2}[1-(x^2+y^2)]^{1/2} \quad \text{(A-9)}$$

Appendix B
Derivation of Eq. (22)

Given the probability distribution of β, $P_\beta(\beta)$, the expectation value of the noncoherent demodulator output is given by

$$E[u] \equiv \int_0^\infty f(\beta) P_\beta(\beta)\, d\beta \tag{B-1}$$

where $f(\beta) = E[u|\beta] = (\sqrt{\pi}/2)e^{-\beta}{}_1F_1(3/2,1,\beta)$ is the conditional expectation of u. The integral in Eq. (B-1) depends on the distribution of β and hence is in general difficult to evaluate. However, when the distribution of β is sufficiently narrow, we can expand $f(\beta)$ around the mean SNR $\bar{\beta}$ as

$$f(\beta) \approx f(\bar{\beta}) + f'(\bar{\beta})(\beta - \bar{\beta}) + \frac{f''(\bar{\beta})}{2}(\beta - \bar{\beta})^2 + \ldots \tag{B-2}$$

The derivatives of the function $f(\beta)$ can be evaluated by taking the derivatives of the degenerated hypergeometric function, and applying the recurrence relations:

$$\frac{df(\beta)}{d\beta} = \frac{d}{d\beta}\left(\frac{\sqrt{\pi}}{2}\, e^{-\beta}{}_1F_1(3/2,1,\beta)\right)$$

$$= \frac{\sqrt{\pi}}{4}\, e^{-\beta}{}_1F_1(3/2,2,\beta) \geqslant 0 \tag{B-3}$$

$$\frac{d^2f(\beta)}{d\beta^2} = -\frac{\sqrt{\pi}}{8}\, e^{-\beta}{}_1F_1(3/2,3,\beta) \leqslant 0 \tag{B-4}$$

By substituting the Taylor expansion in Eq. (B-2) into Eq. (B-1), the expectation value can be written as

$$E[u] \approx \int_0^\infty \left(f(\bar{\beta}) + f'(\bar{\beta})(\beta - \bar{\beta}) + \frac{f''(\bar{\beta})}{2}(\beta - \bar{\beta})^2\right) P_\beta(\beta)\, d\beta$$

$$= f(\bar{\beta}) + \frac{f''(\bar{\beta})}{2}\sigma_\beta^2 \tag{B-5}$$

where the first order term vanishes because of the fact that

$$\beta = \int_0^\infty \beta P_\beta(\beta)\, d\beta \tag{B-6}$$

To show that $f(\bar{\beta})$ is an upper bound of Eq. (B-1), note that since the second derivative of $f(\beta)$ is less than zero for all β, the function $f(\beta)$ is always less than the linear term $g(\beta) = f(\bar{\beta}) + f'(\bar{\beta})(\beta - \bar{\beta})$. Consequently, $f(\bar{\beta})$, which is derived by substituting $g(\beta)$ for $f(\beta)$ in Eq. (B-1), is always greater than the actual expectation value.