Optimal Power Allocation Between Beam Tracking and Symbol Detection Channels in a Free-Space Optical Communications Receiver

Muhammad Salman Bashir ¹ and Mohamed-Slim Alouini ¹

¹Affiliation not available

October 30, 2023

Abstract

Free-space optical (FSO) communications will play an important role in the backhaul of future generation of wireless networks in order to support high data rates. Because of narrow beamwidth inherent to an optical signal, acquisition and tracking form an important component of any FSO communication system. In this study, we have analyzed the optimization of received power allocation between tracking and data channels in an FSO receiver. The objective function that is optimized (minimized) are the probability of error and the probability of outage, and the optimization of power allocation is carried out as a function of parameters such as noise power, pointing error variance, pointing error correlation coefficient, and the threshold of outage. We have analyzed the optimization concerning the lognormal and exponentiated Weibull fading scenarios as well. We learn that the optimal power allocation is a function of the received signal-to-noise ratio: a lower signal-to-noise ratio dictates that a higher fraction of received power should be diverted to the tracking channel and vice versa.
Optimal Power Allocation Between Beam Tracking and Symbol Detection Channels in a Free-Space Optical Communications Receiver

Muhammad Salman Bashir, Senior Member, IEEE, and Mohamed-Slim Alouini, Fellow, IEEE

Abstract

Free-space optical (FSO) communications will play an important role in the backhaul of future generation of wireless networks in order to support high data rates. Because of narrow beamwidth inherent to an optical signal, acquisition and tracking form an important component of any FSO communication system. In this study, we have analyzed the optimization of received power allocation between tracking and data channels in an FSO receiver. The objective function that is optimized (minimized) are the probability of error and the probability of outage, and the optimization of power allocation is carried out as a function of parameters such as noise power, pointing error variance, pointing error correlation coefficient, and the threshold of outage. We have analyzed the optimization concerning the lognormal and exponentiated Weibull fading scenarios as well. We learn that the optimal power allocation is a function of the received signal-to-noise ratio: a lower signal-to-noise ratio dictates that a higher fraction of received power should be diverted to the tracking channel and vice versa.

Index Terms

Beam tracking, exponentiated Weibull fading, free-space optical communications, lognormal fading, optimal power allocation, pointing error, probability of error, probability of outage

I. INTRODUCTION

Free-space optical communications (FSO) is a promising candidate that will play a crucial role in supporting high data rates in the backhaul of future generation of wireless communications [1]. However,
due to narrow beamwidth of the optical signal, acquisition and tracking subsystems are required in order to acquire, and thereafter, maintain the link alignment (tracking) with narrow optical beams [2].

In most optical receivers employed in terrestrial FSO links, the received signal energy is divided into beam tracking and data/symbol detection paths or channels [3], [4]. In this regard, a beam splitter is used to split received energy between the two channels. In the tracking channel, an array of detectors is employed in order to track the beam, and a single detector is employed in the data or symbol detection channel. Even though the authors in [5] have proposed that a single array of detectors can be used for both signal detection and beam tracking (thereby doing away with the tracking channel altogether), the single detector scheme may still outperform low complexity combiners—such as an equal gain combiner—that are used to fuse the outputs of each detector in the array\(^1\). However, an array of detector—that employs either a selection combiner or a maximal ratio combiner—may outperform the single detector in terms of probability of error performance if the number of detectors in the array is large enough\(^2\). However, the computational cost of such fusion algorithms can be prohibitively high which, among other factors, depends on the number of detectors in the array [6]. Therefore, depending on the following factors, a single detector receiver based on separate beam tracking and data detection channels would be a more attractive choice as opposed to a detector array that is used for both beam tracking and data detection.

1) If the variance of the pointing error is not too large, a single-detector system can perform significantly better than an array of detectors (equal gain combiner), and can provide a satisfactory probability of error performance. We define the pointing error as the Euclidean distance between the beam center and the origin, where the origin is the center of the array in the tracking channel, or the center of the single-detector in the data channel.

2) The rate at which the detector is switched on and off depends on the junction capacitance of the detector which in turn is a function of the physical size of the detector: A larger detector implies a larger capacitance which in turn limits the rate at which the data can be sensed by the detector [4]. Thus, if an array of detectors is used for simultaneous beam tracking and symbol detection, then a significantly large number of smaller detectors may be required to achieve the appropriate surface area in order to satisfactorily track the pointing error of large variance. A large number of

\(^1\)This is because the equal gain combiner not only adds the signal power from each detector in the sufficient statistic, but also the noise power.

\(^2\)This is an intuitive argument since a maximal ratio combiner will optimally weight the output of each detector in the array. Hence, the output of each detector will be scaled proportionally to the SNR in the said detector before the fusion of all outputs takes place. In this context, a maximal ratio combiner may outperform a single-detector system.
detectors in the array would increase the computational complexity manifold with a maximal ratio [6] or a selection combiner. In this scenario, an array with a small number of detectors (large-sized detectors) may only be used in the tracking channel that does not have to have a large bandwidth (since it is not being used for symbol detection), and a single small detector may be employed in the data channel for a faster response.

A. Motivations of Current Study

In this context, the probability of error of a single-detector receiver is a function of two parameters: *pointing error variance* and the *signal-to-noise ratio* in the data detection channel. The variance of pointing error is a function of the mean-square error performance of the tracking system, which in turn is a function of the signal-to-noise ratio in the tracking channel. In order to maximize the performance of the receiver, it is important to optimally divide received beam power between beam tracking and data detection channels so that there is enough power in the two channels in order to minimize the probability of error. If the power going into the tracking channel is not enough, the tracking performance degrades and the resulting probability of error becomes worse due to a high variance of the resulting pointing error. On the other hand, if more than enough power is channeled into the tracking assembly but not enough in the data detection assembly, the probability of error suffers due to low signal-to-noise ratio in the data detection channel. Hence, the power allocation to each channel should be optimized\(^3\) in order to minimize the probability of error of the receiver.

II. Literature Review and Model Assumptions

References such as [2], [3], [4] and [7] discuss the division or split of energy between tracking and data detection channels in an FSO receiver without analyzing the optimization of the energy allocation between the two channels. The authors in [4] and [8] promote communications with an array of avalanche photodiodes. They claim that such arrays can be efficiently utilized for both beam tracking and symbol detection. The authors in [9], [10], [11], [12] discuss beam position tracking and symbol detection performance of photon-counting detector arrays for a photon-starved channel in a deep-space optical communication setting. Additionally, [13], [14], [15] consider the communications performance with an array of detectors for a deep-space communications channel. The authors in [16], [17], [18] analyze the performance and capacity of an FSO single-detector receiver that suffers from generalized pointing errors.

\(^3\)The power going into each of the two channels may be varied by the use of a variable beam splitter.
However, none of these studies describe or touch upon optimal techniques to split received beam energy or power between the tracking and data detection channels for a single-receiver system. For the purpose of the current study, we conducted a thorough search on optimal beam splitting in an optical receiver in the current FSO literature, and we were unable to find any study that deals with this issue conclusively.

Regarding the pointing error, the authors in [19] investigate the pointing error statistics for a single-detector receiver and the Gaussian beam, and the outage capacity of the system is optimized in terms of the received signal beam radius impinging on the aperture. The authors in [20] consider the maximization of link availability as a function of beam radius for a single-detector receiver.

In this paper, we have analyzed the optimal power allocation between beam tracking and data detection channels of a single-detector optical receiver. In this analysis, we have used a low complexity centroid algorithm for tracking the beam with an array of detectors in the tracking channel. In order to simplify the analysis, we have used the asymptotic approximation $M \to \infty$, where $M$ is the number of detectors in the array for a fixed array size or area. Such an asymptotic array is termed as the “continuous array” in this paper, as opposed to a “discrete array” whose $M$ is finite. For a fixed array area, $M \to \infty$ implies that the area of a single detector in the array is approaching zero. Fig. 1 illustrates a $4 \times 4$ or a 16-cell square array for which the area of each individual detector is the same.

As shown in Section VII, the performance of the actual (discrete) array is close to that of the continuous array if $M$ is large enough, and this closeness in performance justifies our choice of the continuous array in the analysis of this study. Analysis with a continuous array results in tractable mathematical expressions that are easier to analyze and simulate.

A. Model Assumptions

In this paper, we make the following assumptions about the signal model at the receiver:
1) The intensity profile of the spot on the detector/detector array is approximated by a Gaussian. The center \((x_0, y_0)\) of this Gaussian beam deviates about the origin \((0, 0)\). The beam moves about in a random fashion due to vibrations of the transmitter/receiver mechanical assemblies and the beam wander due to atmospheric turbulence. The deviation of the center of the beam from the origin is represented by random variable \(R\). The random variable \(R\) is also known as the pointing error in literature.

2) In the data detection channel, the detector radius \(a\) is much smaller than the radius \(\rho_d\) of the beam. This assumption is valid since the focus of current FSO research is on small sized detectors that are able to switch at fast enough rates due to their smaller junction capacitance [4]. A beam footprint (or radius) larger than the size of the detector is required in order to minimize the outage probability due to jitter or pointing error.

3) In the tracking channel, the dimensions of the array are much larger than the variance of the pointing error. Additionally, the dimensions of the array are also larger than the radius of the beam \(\rho_t\) (which does not have to be necessarily the same as the beam radius \(\rho_d\) in the data detection channel). These assumptions follows from the design practice in which arrays with large enough dimensions/area are chosen in order to account for the beam wander and minimize the outage probability.

4) Without any loss of generality, the shape of the array is chosen to be a square, and each detector in the array is also a square (please see Fig. 1).

5) We consider a high received signal energy scenario which usually holds for a ground-based free-space optical communications channel [21]. For the high energy model, the (discrete) photon counting model based on Poisson distribution converges to the (continuous) Gaussian distribution. Therefore, the sum of thermal and shot noise in a detector is modeled by a zero-mean Gaussian random variable with variance \(\sigma^2\).

B. Organization of this Paper

This paper is organized as follows. Section III lays out the probability of error expression for the single-detector receiver. Section IV briefly discusses the jitter model and Section V describes the statistics of the centroid error in detail for both the discrete and continuous arrays. Section VI discusses the justification for approximating the centroid error by a Gaussian distribution and Section VII carries out a bias/variance analysis of the centroid estimator of beam position for a continuous array. Section VIII investigates the
optimization of the error and outage probabilities for different channel conditions and fading channels. Section IX sums up the conclusions of this paper.

III. Probability of Error for the Single-Detector Receiver

For a Gaussian beam, let the total received signal energy in the data detection channel be given by

$$P_D \triangleq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \alpha \frac{\lambda_p}{2\pi \rho_d^2} \exp\left(-\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho_d^2}\right) dy \, dx,$$

(1)

where $\rho_d$ is the beam radius in the data detection channel, $\lambda_p$ is the peak intensity or the average power, $(x_0, y_0)$ is the center of the beam, and $0 < \alpha < 1$ is known as the power split factor or the power allocation factor. The power in the tracking channel is described by

$$P_T \triangleq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (1-\alpha) \frac{\lambda_p}{2\pi \rho_t^2} \exp\left(-\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho_t^2}\right) dy \, dx,$$

(2)

where $\rho_t$ is the beam radius in the tracking channel.

In the data channel, the signal energy received by a detector whose center lies at a distance $R$ from the center of the beam is,

$$P(R) \triangleq \int_{R-a}^{R+a} \int_{-\sqrt{a^2-(x-R)^2}}^{+\sqrt{a^2-(x-R)^2}} \alpha \frac{\lambda_p}{2\pi \rho_d^2} \exp\left(-\frac{(x^2+y^2)}{2\rho_d^2}\right) dy \, dx.$$

(3)
A. Probability of Error

Let the output of the data/symbol detector be given by

\[ Y = S + V, \tag{4} \]

where

\[ S = S(R) \overset{\Delta}{=} \int_0^{T_s} P(R) \, dt \tag{5} \]

for some observation interval \( T_s \) and \( V \sim \mathcal{N}(\mu_V, \sigma^2) \). The quantity \( V \) lumps the effects of noise power due to background radiation and thermal energy. Its mean \( \mu_V \overset{\Delta}{=} \lambda_n \pi a^2 \), where \( \lambda_n \) is the noise intensity in Joules/square millimeters/second. Let us transmit symbol \( 0 \leq j < M \) for an \( M \)-PPM scheme where PPM stands for pulse position modulation and \( M \) a positive integer that represents the order of a PPM symbol. For an equal gain combiner that operates on the maximum likelihood principle, we say that a symbol \( j \) is detected if \( Y_j > Y_i \) for all \( i \neq j, 0 \leq i < M \). Then, the conditional probability of the correct decision is

\[ P(c|j) = \left( P \left\{ Y_j - Y_i > 0 \right\} \right)^{M-1}, \tag{6} \]

where \( Y_j \sim \mathcal{N}(S(R) + \mu_V, \sigma^2) \) and \( Y_i \sim \mathcal{N}(\mu_V, \sigma^2) \). For an equiprobable symbol scheme, we have that \( P(c|j) = P(c) \). Therefore, the probability of error, \( P(e) \), is \( P(e) = 1 - P(c) \).

The quantity \( (Y_j - Y_i) \sim \mathcal{N}(S(R), 2\sigma^2) \). Therefore, \( P(e|R = r) = 1 - \left( Q \left( -\frac{S(r)}{\sqrt{2}\sigma} \right) \right)^{M-1} \), and

\[ P(e) = \int_0^\infty P(e|r)f_R(r) \, dr = \int_0^\infty \left( 1 - \left( Q \left( -\frac{S(r)}{\sqrt{2}\sigma} \right) \right)^{M-1} \right) f_R(r) \, dr, \tag{7} \]

where \( f_R(r) \) represents the probability density function of the pointing error, and the function \( Q(x) \overset{\Delta}{=} 1 - \Phi(x) \) where \( \Phi(x) \) is the cumulative distribution function of a standard normal random variable for any \( x \in \mathbb{R} \).

B. Pointing Error \( R \)

The pointing error \( R \)—at some discrete-time interval \( N \)—is the deviation of the beam center \((x_0, y_0)\) from the center of the detector (the center of the detector is also called the origin). Specifically, we say
that

\[ R \triangleq \sqrt{x_0^2 + y_0^2}. \]  \hspace{1cm} (8)

Additionally, \( R \) may be broken down as a sum of two random variables \( R_J \) and \( R_E \):

\[ R = R_J + R_E. \]  \hspace{1cm} (9)

The random variable \( R_J \) denotes the *jitter*. The jitter is defined as a deviation of the incoming beam’s center away from the detector during some time interval \( N \). The jitter occurs either because of atmospheric turbulence, or due to vibrations in the mechanical assemblies of the transmitter/receiver terminals. The job of the tracking system is to sense or estimate \( R_J \) every \( N \) time units, and send the error signal to a control block that activates a *fast steering mirror* (FSM) assembly in order to align the beam center back to the center of the array (please see Fig. 2). Therefore, the average value of \( R_J \) depends on how frequently the FSM is activated (the duration of interval \( N \)) in order to align the beam.

The random variable \( R_E \) is the Euclidean error between the actual and the estimated beam center locations, and depends on the mean-square error of the beam position tracker/estimator. Let us assume that the values of \( x_0 \) and \( y_0 \) are estimated every \( N \) time units with the help of an array of detectors in the tracking channel. Let us assume that the center of the array is the origin \((0, 0)\), and that the estimates of \( x_0 \) and \( y_0 \) are given by \( \hat{x}_0 \) and \( \hat{y}_0 \), respectively, at time \( N \). The control input at the \( N \)th time instant, \( U[N] \), is then equal to \( U[N] \triangleq \left[ -\hat{x}_0 \quad -\hat{y}_0 \right]^\top \), where \( \top \) denotes the transpose operator. Thus, the control input “tries” to bring the beam center back to the point \((0, 0)\) by subtracting \( \left[ \hat{x}_0 \quad \hat{y}_0 \right]^\top \) from \( \left[ x_0 \quad y_0 \right]^\top \).

Thus, when the beam position is adjusted by the control input every \( N \) time units, \((x_0, y_0)\) is not aligned exactly at the origin, but is aligned at \( \mathcal{E}^4 \), where

\[ \mathcal{E} = \begin{bmatrix} \mathcal{E}_x \\ \mathcal{E}_y \end{bmatrix} \triangleq \begin{bmatrix} x_0 - \hat{x}_0 \\ y_0 - \hat{y}_0 \end{bmatrix}, \] \hspace{1cm} (10)

is the error introduced by the estimator. It then follows that \( R_E \triangleq \|\mathcal{E}\|_2 \). The covariance matrix of this error is

\[ C_\mathcal{E} = \mathbb{E} [(\mathcal{E} - \mathbb{E}[\mathcal{E}])(\mathcal{E} - \mathbb{E}[\mathcal{E}])^\top]. \] \hspace{1cm} (11)

\(^4\text{We assume that the delay introduced by the FSM channel is negligible, i.e., the FSM assembly adjusts the beam position almost instantaneously.}\)
The deviation or error $R$ is a random variable with a certain variance $\sigma_R^2$. We will later show that for a continuous array, $R$ is Rayleigh distributed with scale parameter $\sigma_R$ which implies that $x_0 \sim \mathcal{N}(0, \sigma_R^2), y_0 \sim \mathcal{N}(0, \sigma_R^2)$. Additionally, for continuous arrays, we will also see that $R_J$ and $R_E$ are also Rayleigh distributed with scale parameters $\sigma_{R_J}$ and $\sigma_{R_E}$, respectively.

**IV. Distribution of $R_J$**

Let us assume that we estimate the beam center once every $N$ discrete-time units. Let $R_J = \|X[N]\|_2$, where

$$X[n] = \Phi X[n-1] + V[n], \quad (12)$$

where $X[n]$ is a 2×1 random vector at time $n$: $X[n] \triangleq \begin{bmatrix} X[n] & Y[n] \end{bmatrix}^T$, $\Phi$ is the 2×2 state transition matrix, and $V[n]$ is a 2×1 white Gaussian noise vector with covariance matrix $C_V = \sigma_V^2 I_2$ where $I_2$ is a 2×2 identity matrix. The quantity $(X[n], Y[n])$ is a discrete-time random process that represents the center of the beam at the discrete-time instant $n^5$. Specifically, $x_0 \triangleq X[N]$ and $y_0 \triangleq Y[N]$. It can be shown that

$$X[N] = \Phi^N X[0] + \sum_{\ell=1}^{N} \Phi^{N-\ell} V[\ell], \quad (13)$$

where we assume that $X[0] \sim \mathcal{N}(0, \sigma_0^2)$. The covariance matrix of $X[N]$ is

$$C_X = \sigma_0^2 \Phi^{2N} + \sigma_V^2 \sum_{\ell=1}^{N} \Phi^{2(N-\ell)}. \quad (14)$$

For the sake of simplification, we will assume that the state transition matrix is a diagonal matrix which means there is no coupling between the $x$ and $y$ dimensions of evolution noise. Moreover, we assume that the eigenvalues of $\Phi$ are equal. Additionally, since the eigenvalues of $\Phi$ are equal and the covariance matrix of $V[n]$ is diagonal with equal diagonal values, we have that $C_X$ is of the form $\sigma_X^2 I_2$ for some positive number $\sigma_X^2$.

Let us define $\sigma_{R_J}^2 \triangleq \sigma_X^2$. Now, we have that $R_J \triangleq \|X[N]\|_2$, where $X[N]$ is a zero-mean Gaussian vector with i.i.d $x$ and $y$ components. We note that $R_J$ is a Rayleigh random variable whose scale parameter $\sigma_{R_J}$ is dependent on $N$ through (14).

---

5This process as known as the Gauss-Markov process in the control literature.
V. Derivation of Distribution Function of 
\((\hat{x}_0, \hat{y}_0)\) for Centroid Estimator

Since \(R\) depend on \((\hat{x}_0, \hat{y}_0)\), this section is dedicated to the derivation of the statistics of estimate \((\hat{x}_0, \hat{y}_0)\) that is computed by a centroid estimator. For the purpose of the following analysis, we assume that \((x_0, y_0)\) is fixed on the array in the tracking assembly (a deterministic but hidden quantity). The region of the square detector array is defined by the set \(A\). The center of the array lies at the origin \((0, 0)\), and the array extends from \(-\frac{\ell(A)}{2}\) to \(\frac{\ell(A)}{2}\) along each axis. Thus, the length of one side of the array is \(\ell(A)\).

The amount of power allocated for the tracking channel is \(P_T\) given by (2).

A. Centroid is the Weighted Least Squares Estimate of \((x_0, y_0)\)

A weighted least squares estimate of the beam position on the array is defined as

\[
\hat{x}_0 \triangleq \arg \min_{x_0} (x_0 - x) \mathbf{P}(x_0 - x), \quad \hat{y}_0 \triangleq \arg \min_{y_0} (y_0 - y) \mathbf{P}(y_0 - y),
\]

(15)

where \(x \triangleq [x_1 \ x_2 \ \cdots \ x_M]^\top\), \(y \triangleq [y_1 \ y_2 \ \cdots \ y_M]^\top\), and \(\mathbf{P} \triangleq \begin{bmatrix} P_1 & 0 & 0 & \cdots & 0 \\
0 & P_2 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & P_M \end{bmatrix}\). Here, \((x_m, y_m)\) denote the center of the \(m\)th cell, and \(P_m\) is the total power (or energy) in the \(m\)th detector. The integer \(M\) is the total number of detectors in the array.

Equation (15) can be rewritten as

\[
\hat{x}_0 \triangleq \arg \min_{x_0} \sum_{m=1}^{M} P_m (x_0 - x_m)^2, \quad \hat{y}_0 \triangleq \arg \min_{y_0} \sum_{m=1}^{M} P_m (y_0 - y_m)^2,
\]

(16)

where the output \(P_m\) of the \(m\)th detector is

\[
P_m \triangleq \int_A \int_{A_m} \frac{(1 - \alpha) \lambda_p}{2\pi \rho_i^2} \exp\left(-\frac{(x - x_0)^2 + (y - y_0)^2}{2\rho_i^2}\right) dy \, dx + W_m,
\]

(17)

where \(W_m\) is the noise contribution of the \(m\)th detector. The quantity \(W_m\) is a discrete-time white noise process with \(E[W_m] = 0\) and \(E[W_m^2] = \sigma_m^2\) for all \(m = 1, 2, \ldots, M\).
Now, from (16), we have that

\[ \frac{\partial}{\partial x_0} \sum_{m=1}^{M} P_m (x_0 - x_m)^2 = 0 \implies \sum_{m=1}^{M} P_m (x_0 - x_m) = 0 \]

\[ \implies x_0 P_0 - \sum_{m=1}^{M} P_m x_m = 0 \implies \hat{x}_0 = \frac{1}{P_0} \sum_{m=1}^{M} P_m x_m, \tag{18} \]

where \( P_0 \triangleq \sum_{m=1}^{M} P_m \) is the total signal power in the tracking channel. Expression (18) is the well known centroid estimator of \( x_0 \). Similarly, it can be shown that the linear least squares estimate of \( y_0 \) is \( \hat{y}_0 = \frac{1}{P_0} \sum_{m=1}^{M} P_m y_m \).

### B. Finite M Case

In this regard, let the noise power from a single detector of the array \( \sigma^2_M \triangleq N_0 \frac{|A|}{M} \), where \( N_0 \) is noise power per unit area (Watts/mm²). Moreover, let

\[ \Lambda_m \triangleq \int_{A_m} \frac{(1-\alpha) \lambda_p}{2\pi \rho_t^2} \exp \left( -\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho_t^2} \right) \, dy \, dx, \tag{19} \]

which is the signal component in the \( m \)th detector. Then, from (18), the value of \( \hat{x}_0 \) is given by

\[ \hat{x}_0 = \frac{\sum_{m=1}^{M} x_m \Lambda_m + \sum_{m=1}^{M} x_m W_m}{\int_{A} (1-\alpha) \lambda_p \exp \left( -\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho_t^2} \right) \, dy \, dx + \sum_{m=1}^{M} W_m} \approx \frac{\sum_{m=1}^{M} x_m \Lambda_m + \sum_{m=1}^{M} x_m W_m}{(1-\alpha) \lambda_p + \sum_{m=1}^{M} W_m}, \tag{20} \]

where we have used the fact that \( \int_{A} (1-\alpha) \lambda_p \exp \left( -\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho_t^2} \right) \, dy \, dx = 1 \) since we assume that \( \pi \rho_t^2 << |A| \).

Let us denote \( U_1 \triangleq \sum_{m=1}^{M} x_m \Lambda_m + \sum_{m=1}^{M} x_m W_m \), and \( U_0 \triangleq (1-\alpha) \lambda_p + \sum_{m=1}^{M} W_m \). Then, \( \hat{x}_0 = \frac{U_1}{U_0} \).

Additionally, both \( U_1 \) and \( U_0 \) are Gaussian, and \( E[U_1] = \sum_{m=1}^{M} x_m \Lambda_m \), \( \text{VAR}[U_1] = \sigma^2_M \sum_{m=1}^{M} x_m^2 \), \( E[U_0] = (1-\alpha) \lambda_p \), and \( \text{VAR}[U_0] = M \sigma^2_M = N_0 |A| \). Additionally,

\[ E[U_1 U_0] = E \left[ \left( \sum_{m=1}^{M} x_m \Lambda_m + \sum_{m=1}^{M} x_m W_m \right) \left( (1-\alpha) \lambda_p + \sum_{m=1}^{M} W_m \right) \right] \tag{21} \]

\[ = (1-\alpha) \lambda_p \sum_{m=1}^{M} x_m \Lambda_m + \sum_{m=1}^{M} x_m \sum_{\ell=1}^{M} \text{VAR}[W_\ell W_\ell] = (1-\alpha) \lambda_p \sum_{m=1}^{M} x_m \Lambda_m + \sigma^2_M \sum_{m=1}^{M} x_m \]

\[ = (1-\alpha) \lambda_p \sum_{m=1}^{M} x_m \Lambda_m = E[U_1]E[U_0]. \tag{22} \]

We thus conclude that \( U_1 \perp U_0 \).
In a similar fashion, \( \hat{y}_0 = \frac{\sum_{m=1}^{M} y_m \Lambda_m}{(1-\alpha) \lambda_p + \sum_{m=1}^{M} y_m W_m} \), and denoting \( V_1 \triangleq \sum_{m=1}^{M} y_m \Lambda_m + \sum_{m=1}^{M} y_m W_m \), we have that \( \hat{y}_0 = \frac{V_1}{U_0} \), where \( V_1 \) is also a Gaussian random variable: \( \mathbb{E}[V_1] = \sum_{m=1}^{M} y_m \Lambda_m \), and \( \text{VAR}[V_1] = \sigma^2_M \sum_{m=1}^{M} y_m^2 \). Additionally, we can easily show that \( V_1 \perp U_0 \).

1) Approximate Cumulative Distribution Functions of \( \hat{x}_0 \) and \( \hat{y}_0 \)

Under conditions of high signal-to-noise ratio, the random variables \( \hat{x}_0 \) and \( \hat{y}_0 \) are approximately distributed as quasiGaussian random variables [22]. Thus,

\[
F_{\hat{x}_0}(x) \rightarrow \Phi \left( \frac{x - \mu_{u_1}/\mu_{u_0}}{\sigma_{u_1} \sigma_{u_0} a(x)/\mu_{u_0}} \right),
\]

as \( \lambda_p/N_0 \rightarrow \infty \), and

\[
a(x) \triangleq \left( \frac{x^2}{\sigma^2_{u_1}} + \frac{1}{\sigma^2_{u_0}} \right)^{\frac{1}{2}},
\]

\( \mu_{u_1} \triangleq \mathbb{E}[U_1], \mu_{u_0} \triangleq \mathbb{E}[U_0], \sigma^2_{u_1} \triangleq \text{VAR}[U_1], \) and \( \sigma^2_{u_0} \triangleq \text{VAR}[U_0] \). The function \( \Phi(\cdot) \) in (23) is the cumulative distribution function of a standard normal random variable. The distribution \( F_{\hat{x}_0}(x) \) is not a Gaussian distribution since the variance of this distribution is a function of the argument \( x \) through \( a(x) \). However, if \( a(x) \) were a constant, then (23) would be a Gaussian distribution. We will later argue in Section VI that for \( a(x) \triangleq a(\mu_{u_1}/\mu_{u_0}) \), the resulting (Gaussian) distribution approximates the quasiGaussian distribution close enough for a large enough signal-to-noise ratio.

Similarly,

\[
F_{\hat{y}_0}(y) \rightarrow \Phi \left( \frac{y - \mu_{v_1}/\mu_{u_0}}{\sigma_{v_1} \sigma_{u_0} a(y)/\mu_{u_0}} \right),
\]

as signal-to-noise ratio \( \rightarrow \infty \). Additionally, \( \mu_{v_1} \triangleq \mathbb{E}[V_1] \) and

\[
a(y) \triangleq \left( \frac{y^2}{\sigma^2_{v_1}} + \frac{1}{\sigma^2_{u_0}} \right)^{\frac{1}{2}}.
\]

and in order to approximate (25) by the Gaussian distribution, we use \( a(y) \approx a(\mu_{v_1}/\mu_{u_0}) \).

Thus, for high enough signal-to-noise ratio, \( \hat{x}_0 \sim \mathcal{N}(\mu_{x_0}, \sigma^2_{x_0}) \), and \( \hat{y}_0 \sim \mathcal{N}(\mu_{y_0}, \sigma^2_{y_0}) \), where

\[
\mu_{\hat{x}_0} \triangleq \sum_{m=1}^{M} x_m \int_{A_m} \frac{1}{2\pi \rho^2} \exp \left( -\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho^2} \right) dy \, dx,
\]

\[
\mu_{\hat{y}_0} \triangleq \sum_{m=1}^{M} y_m \int_{A_m} \frac{1}{2\pi \rho^2} \exp \left( -\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho^2} \right) dy \, dx.
\]
We note that the bias vector \( \begin{bmatrix} x_0 - \mu_{x_0} & y_0 - \mu_{y_0} \end{bmatrix}^T \) is not necessarily zero. Moreover, from (23) and (25),

\[
\sigma_{\hat{x}_0} = \frac{\sigma_{u_1} \sigma_{u_0} a(\mu_{u_1} / \mu_{u_0})}{\mu_{u_0}}, \quad \sigma_{\hat{y}_0} = \frac{\sigma_{v_1} \sigma_{u_0} a(\mu_{v_1} / \mu_{u_0})}{\mu_{u_0}},
\]

(29)

where \( \sigma_{u_1}^2 = \sigma_{v_1}^2 = \sigma_M^2 \sum_{m=1}^M x_m^2 \) and \( \sigma_{u_0}^2 = M \sigma_M^2 = N_0 |A| \). It can then be shown that

\[
\sigma_{\hat{x}_0}^2(x_0, y_0) = \left( \frac{N_0^2 |A|^2}{M(1 - \alpha)^2 \lambda_p^2} \sum_{m=1}^M x_m^2 \right) \left( \frac{M \left( \sum_{m=1}^M x_m \lambda_m \right)^2}{((1 - \alpha)e_p^2) \left( N_0 |A| \sum_{m=1}^M x_m^2 \right) + 1} \right),
\]

(30)

\[
\sigma_{\hat{y}_0}^2(x_0, y_0) = \left( \frac{N_0^2 |A|^2}{M(1 - \alpha)^2 \lambda_p^2} \sum_{m=1}^M y_m^2 \right) \left( \frac{M \left( \sum_{m=1}^M y_m \lambda_m \right)^2}{((1 - \alpha)e_p^2) \left( N_0 |A| \sum_{m=1}^M y_m^2 \right) + 1} \right),
\]

(31)

where \( \sigma_{\hat{x}_0}^2 \) and \( \sigma_{\hat{y}_0}^2 \) are functions of \( x_0 \) and \( y_0 \) since \( \Lambda_m \) is a function of \((x_0, y_0)\).

The correlation between \( \hat{x}_0 \) and \( \hat{y}_0 \) is \( \mathbb{E}[\hat{x}_0 \hat{y}_0] = \mathbb{E} \left[ \frac{\hat{U}_1 \hat{V}_1}{\hat{U}_0^2} \right] \), which implies

\[
\mathbb{E}[\hat{x}_0 \hat{y}_0] = \left( \sum_{m=1}^M \sum_{\ell=1}^M x_m y_\ell \Lambda_m \Lambda_\ell + \sigma_M^2 \sum_{m=1}^M x_m y_m \right) \mathbb{E} \left[ \frac{1}{\hat{U}_0^2} \right],
\]

(32)

and, since \( \sum_{m=1}^M x_m y_m = 0 \),

\[
\mathbb{E}[\hat{x}_0 \hat{y}_0] = \left( \sum_{m=1}^M \sum_{\ell=1}^M x_m y_\ell \Lambda_m \Lambda_\ell \right) \int_{-\infty}^{\infty} \frac{1}{u^2} f_{U_0}(u) \, du.
\]

(33)

where \( U_0 \sim \mathcal{N}(0, (1 - \alpha) e_p, N_0 |A|) \).

a) **High SNR Case:** For high signal-to-noise ratio, it is argued in Section V-C3 that the integral \( \int_{-\infty}^{\infty} \frac{1}{u^2} f_{U_0}(u) \, du \) can be approximated by \( (1 - \alpha)^2 e_p^2 \). Thus, for high signal-to-noise ratio,

\[
\mathbb{E}[\hat{x}_0 \hat{y}_0] = \frac{\sum_{m=1}^M \sum_{\ell=1}^M x_m y_\ell \Lambda_m \Lambda_\ell}{(1 - \alpha)^2 e_p^2}.
\]

(34)

C. **\( M \rightarrow \infty \) Case**

In this case, the area of each detector, \(|A_m|\), goes to zero, and we achieve a “continuum” on the array in the limit. Since the noise from each infinitesimal point detector of this continuous array is independent, the noise over the region of the array is modeled by a white noise process \( W(x, y) \), which is a zero-mean wide-sense stationary Gaussian process in two dimensions. In this case, we assume that the autocorrelation function \( R_W(\tau, \varsigma) = N_0 \delta(\tau, \varsigma) \), where \( N_0 > 0 \), and \( \delta(x, y) \) is the Dirac delta function in
two dimensions. Specifically, \( \delta(x, y) = \infty \) if \( x = 0 \) and \( y = 0 \), \( \delta(x, y) = 0 \) if \( x \neq 0 \) or \( y \neq 0 \), \( \delta(x)\delta(y) \), and \( \int_{-\epsilon}^{\epsilon} \int_{-\epsilon}^{\epsilon} \delta(x, y) \, dy \, dx = 1 \) for any \( \epsilon > 0 \). It then follows that

\[
E[W(x, y)W(s, t)] = N_0 \delta(x - s, y - t) = N_0 \delta(x - s)\delta(y - t).
\]  

(35)

Let us assume that we divide the continuous array into a number of very small square regions \( A_m \). Let the total number of regions be \( M \), where \( M \) is a large number. The power in each small region \( A_m \) is

\[
P_m \approx \frac{(1-\alpha)\lambda_p}{2\pi \rho_t^2} \exp \left(-\frac{(x_m-x_0)^2+(y_m-y_0)^2}{2\rho_t^2}\right) \Delta_M + \iint_{A_m} W(x, y) \, dy \, dx,
\]

and the centroid estimator of \( x_0 \) is

\[
\hat{x}_0 = \frac{\sum_{m=1}^{M} x_m \iint_{A_m} \left(\frac{(1-\alpha)\lambda_p}{2\pi \rho_t^2} \exp \left(-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho_t^2}\right) + W(x, y)\right) \, dy \, dx}{(1-\alpha)\lambda_p + W_0},
\]

where \( W_0 \triangleq \iint_{A} W(x, y) \, dy \, dx \). In the limit as \( M \to \infty \),

\[
\lim_{M \to \infty} \hat{x}_0 = \frac{\sum_{m=1}^{M} x_m \frac{(1-\alpha)\lambda_p}{2\pi \rho_t^2} \exp \left(-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho_t^2}\right) \Delta_M + \sum_{m=1}^{M} x_m \iint_{A_m} W(x, y) \, dx \, dy}{(1-\alpha)\lambda_p + W_0}
\]

\[
= \frac{\iint_{A} x \frac{(1-\alpha)\lambda_p}{2\pi \rho_t^2} \exp \left(-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho_t^2}\right) \, dy \, dx + W_1}{(1-\alpha)\lambda_p + W_0} = \frac{(1-\alpha)\lambda_p x_0 + W_1}{(1-\alpha)\lambda_p + W_0},
\]

(37)

where \( W_1 \triangleq \iint_{A} x W(x, y) \, dy \, dx \). In a similar fashion,

\[
\hat{y}_0 = \frac{(1-\alpha)\lambda_p y_0 + W_1}{(1-\alpha)\lambda_p + W_0}.
\]

(38)

**Note:** The centroid estimator’s performance is independent of beam radius \( \rho_t \) in the tracking channel (please see (37) and (38)).
1) Variance of $W_0$ and $W_1$

It can be easily shown that both $W_0$ and $W_1$ are zero-mean Gaussian random variables. The variance of $W_0$ is

$$\mathbb{E}[W_0^2] = \mathbb{E} \left[ \iint_A W(x, y) \, dy \, dx \iint_A W(s, t) \, ds \, dt \right] = \iint_A \iint_A \mathbb{E} [W(x, y)W(s, t)] \, dt \, ds \, dy \, dx$$

$$= N_0 \iint_A \iint_A \delta(x - s)\delta(y - t) \, dt \, ds \, dy \, dx = \iint_A \iint_A \frac{\ell(A)}{2}^2 \delta(x - s)\delta(y - t) \, dt \, ds \, dy \, dx$$

$$= N_0 \ell^2(A) = N_0 |A|. \quad (39)$$

The variance of $W_1$ is

$$\mathbb{E}[W_1^2] = \mathbb{E} \left[ \iint_A xW(x, y) \, dy \, dx \iint_A sW(s, t) \, ds \, dt \right] = \iint_A \iint_A x \mathbb{E}[W(x, y)W(s, t)] \, dt \, ds \, dy \, dx$$

$$= N_0 \iint_A \iint_A x s \delta(x - s)\delta(y - t) \, dt \, ds \, dy \, dx = N_0 \iint_A \iint_A x s \delta(x - s) \, ds \, dx \iint_A \iint_A \frac{\ell(A)}{2}^2 \delta(y - t) \, dt \, dy$$

$$= N_0 \frac{|A|^2}{12}. \quad (40)$$

2) Uncorrelatedness of $W_0$ and $W_1$

$$\mathbb{E}[W_1W_0] = \mathbb{E} \left[ \iint_A xW(x, y) \, dy \, dx \iint_A W(s, t) \, dt \, ds \right] = \iint_A \iint_A x \mathbb{E}[W(x, y)W(s, t)] \, dt \, ds \, dy \, dx$$

$$= N_0 \iint_A \iint_A x \delta(x - s) \, ds \, dx \iint_A \iint_A \frac{\ell(A)}{2}^2 \delta(y - t) \, dt \, dy$$

$$= N_0 \iint_A \frac{\ell(A)}{2} \, dx \iint_A \frac{\ell(A)}{2} \, dy = 0 \quad (41)$$

which is equal to $\mathbb{E}[W_1]\mathbb{E}[W_0]$. Therefore, $W_0 \perp W_1$.

3) Approximate Cumulative Distribution Functions of $\hat{x}_0$ and $\hat{y}_0$

Let us assume that $U_1 \triangleq (1 - \alpha)\lambda_p + W_1$, and $U_0 \triangleq (1 - \alpha)\lambda_p + W_0$. Additionally, let $V_1 \triangleq (1 - \alpha)\lambda_p + W_1$. Then, $\hat{x}_0 = \frac{V_1}{U_0}$, and $\hat{y}_0 = \frac{V_1}{U_0}$. Since $U_1 \perp U_0$ and $V_1 \perp U_0$, the random variables $\hat{x}_0$ and $\hat{y}_0$ are approximately distributed as Gaussian random variables for high signal-to-noise ratio as discussed in V-B1. The expressions for $F_{\hat{x}_0}(x), a(x), F_{\hat{y}_0}(y)$ and $a(y)$ are the same as given in (23), (24), (25) and (26), respectively.
Thus, \( \hat{x}_0 \sim \mathcal{N}(\mu_{x_0}, \sigma_{x_0}^2) \), and \( \hat{y}_0 \sim \mathcal{N}(\mu_{y_0}, \sigma_{y_0}^2) \), where

\[
\mu_{\hat{x}_0} \triangleq \frac{\mu_{u_1}}{\mu_{u_0}} = \frac{(1 - \alpha)\lambda_p x_0}{(1 - \alpha)\lambda_p} = x_0, \quad \mu_{\hat{y}_0} \triangleq \frac{\mu_{v_1}}{\mu_{v_0}} = \frac{(1 - \alpha)\lambda_p y_0}{(1 - \alpha)\lambda_p} = y_0, \tag{42}
\]

and we note that \( \hat{x}_0 \) and \( \hat{y}_0 \) are unbiased estimators of \( x_0 \) and \( y_0 \), respectively. Additionally,

\[
\sigma_{\hat{x}_0}^2 = \frac{\sigma_{u_1}^2}{\mu_{u_0}} = \frac{\sigma_{u_0}^2}{\mu_{u_0}}, \quad \sigma_{\hat{y}_0}^2 = \frac{\sigma_{v_1}^2}{\mu_{v_0}} = \frac{\sigma_{v_0}^2}{\mu_{v_0}}, \tag{43}
\]

where \( \sigma_{u_1}^2 = \sigma_{v_1}^2 = \frac{N_0}{12} |A|^2 \) and \( \sigma_{u_0}^2 = \sigma_{v_0}^2 = N_0 |A| \). After a few simplifications, we obtain

\[
\sigma_{\hat{x}_0}^2(x_0) = \frac{N_0 |A|^2}{12(1 - \alpha)^2 \lambda_p^2} \left( 1 + \frac{12x_0^2}{|A|} \right), \quad \sigma_{\hat{y}_0}^2(y_0) = \frac{N_0 |A|^2}{12(1 - \alpha)^2 \lambda_p^2} \left( 1 + \frac{12y_0^2}{|A|} \right). \tag{44}
\]

The correlation between \( \hat{x}_0 \) and \( \hat{y}_0 \) is

\[
\mathbb{E} [\hat{x}_0 \hat{y}_0] = \mathbb{E} \left[ \frac{U_1 V_1}{U_0^2} \right] = \mathbb{E} [U_1 V_1] \mathbb{E} \left[ \frac{1}{U_0^2} \right] = \mathbb{E} \left[ \frac{(1 - \alpha)^2 \lambda_p^2 x_0 y_0 + (1 - \alpha)\lambda_p (x_0 + y_0) W_1 + W_1^2}{U_0^2} \right], \tag{45}
\]

\[
= \left( 1 - \alpha \right)^2 \lambda_p^2 x_0 y_0 + N_0 \frac{|A|^2}{12} \mathbb{E} \left[ \frac{1}{U_0^2} \right], \tag{46}
\]

which implies that

\[
\mathbb{E} [\hat{x}_0 \hat{y}_0] = \left( 1 - \alpha \right)^2 \lambda_p^2 x_0 y_0 + N_0 \frac{|A|^2}{12} \int_{-\infty}^{\infty} \frac{1}{u^2} f_{U_0}(u) \, du, \tag{47}
\]

where \( U_0 \sim \mathcal{N}((1 - \alpha)\lambda_p, N_0 |A|) \). Therefore,

\[
\mathbb{E} [\hat{x}_0 \hat{y}_0] = \left( 1 - \alpha \right)^2 \lambda_p^2 x_0 y_0 + N_0 \frac{|A|^2}{12} \int_{-\infty}^{\infty} \frac{1}{u^2} \frac{1}{\sqrt{2\pi N_0 |A|}} \exp \left( -\frac{(u - (1 - \alpha)\lambda_p)^2}{2N_0 |A|} \right) \, du, \tag{48}
\]

where (48) shows the dependence of the cross-correlation \( \mathbb{E} [\hat{x}_0 \hat{y}_0] \) on \( \alpha \). Even though the integrand in (48) is not integrable, we can get an approximate result for the high signal-to-noise ratio case. For large signal-to-noise ratio \( (\lambda_p \gg N_0) \), the Gaussian distribution in (48) is approximated by a Dirac delta function. Then, for \( 0 < \alpha < 0.9 \), (48) can be rewritten as

\[
\mathbb{E} [\hat{x}_0 \hat{y}_0] \approx (1 - \alpha)^2 \lambda_p^2 x_0 y_0 \int_{-\infty}^{\infty} \frac{1}{u^2} \delta(u - (1 - \alpha)\lambda_p) \, du = (1 - \alpha)^2 \lambda_p^2 x_0 y_0 \frac{1}{(1 - \alpha)^2 \lambda_p^2} \tag{49}
\]

\[
= x_0 y_0, \tag{50}
\]

where \( \delta(\cdot) \) is the Dirac delta function. Hence, we see that for the high signal-to-noise ratio case, \( \hat{x}_0 \) and \( \hat{y}_0 \) are approximately uncorrelated for the continuous array.
The centroid estimator for a continuous has a number of features that make it attractive in terms of optimization. For one, the centroid is unbiased with uncorrelated components (for high signal-to-noise ratio). This implies that the radial distance of the centroid error is zero-mean. In the following sections, we will show that the (average) variance of the centroid error in each dimension is equal, and that the centroid error is approximately Gaussian in each dimension. This implies that the radial error \( R_E \) of the centroid is a Rayleigh distributed random variable. Finally, the expression of the centroid for the continuous array is straightforward (see (37) and (38)) as opposed to the centroid for a discrete array (see (20)) which is a function of number of detectors \( M \).

VI. Gaussian Approximation of Centroid Error for Continuous Arrays

In this section, we argue that the centroid error for a continuous array can be approximated by a Gaussian distribution for a large signal-to-noise ratio. We know that \( \hat{x}_0 = \frac{U_1}{U_0} \) and \( \hat{y}_0 = \frac{V_1}{U_0} \) where \( U_1, U_0 \) and \( V_1 \) are defined in Section V-C3. The quasiGaussian approximation for \( \hat{x}_0 \) and \( \hat{y}_0 \) is furnished by (23) and (25), respectively.

As one notes from (23) and (25), the variance of the quasiGaussian distribution increases monotonically with the absolute value of the argument (\( x \) or \( y \)) since the parameter \( a(x) \) or \( a(y) \) increase monotonically with \( x \) and \( y \). Let us look at the \( x_0 \) case only since the results obtained regarding the analysis of \( x_0 \) will also hold for \( y_0 \). This quasiGaussian distribution is plotted as a function of \( x \) in Fig. 3 by using the following expression:

\[
f_{\hat{x}_0}(x) = \frac{1}{\sqrt{2\pi \sigma_{u_1} \sigma_{u_0} a(x)/\mu_{u_0}}} \exp \left( -\frac{(x - \mu_{u_1}/\mu_{u_0})^2}{2 \left( \sigma_{u_1} \sigma_{u_0} a(x)/\mu_{u_0} \right)^2} \right)
\]  

(51)

Let us denote the Gaussian approximation of (51) by \( f'_{\hat{x}_0}(x) \):

\[
f'_{\hat{x}_0}(x) \triangleq \frac{1}{\sqrt{2\pi \sigma_{u_1} \sigma_{u_0} a(\mu_{u_1}/\mu_{u_0})/\mu_{u_0}}} \exp \left( -\frac{(x - \mu_{u_1}/\mu_{u_0})^2}{2 \left( \sigma_{u_1} \sigma_{u_0} a(\mu_{u_1}/\mu_{u_0})/\mu_{u_0} \right)^2} \right).
\]  

(52)

Additionally, let us denote the actual probability density function of \( \frac{U_1}{U_0} \) by the ratio distribution \( f_U(u) \). The ratio distribution \( f_U(u) \) is plotted by the (normalized) histogram by randomly sampling \( U_1 \) and \( U_0 \) and then computing their ratio. Fig. 3 indicates the histogram of the actual distribution of \( \hat{x}_0 \), the quasiGaussian approximation, and the Gaussian approximation for different values of signal-to-noise ratio. From this figure, we note that the Gaussian and quasiGaussian distributions converge to the histogram as the signal-to-noise ratio grows large.
In this section, we test the normality of the centroid error by an alternative technique. Here, we look at two commonly used statistical tests, namely the Anderson-Darling test and the Shapiro test, in order to check the normality of $\hat{x}_0$ and $\hat{y}_0$. The interested readers may consult [23] and [24] for more details on these tests.

Let $H_0$ be the hypothesis that the sample comes from a normal distribution. In the case of Anderson-Darling test, we have five possible values of significance level and the critical value. We conduct the Anderson-Darling test against this set of five values. The results are tabulated in Table I.

1) Significance Level = 15, Critical Value = 0.538
2) Significance Level = 10, Critical Value = 0.613
3) Significance Level = 5, Critical Value = 0.736
4) Significance Level = 2.5, Critical Value = 0.858
5) Significance Level = 1, Critical Value = 1.021

In case of the Shapiro test, the critical value is set equal to 0.05, and the p-Value is compared against this critical value in order to decide between i) reject $H_0$, or ii) failure to reject $H_0$. As seen in the outcome of the Anderson-Darling test shown in Table I, the test fails to reject the null hypothesis $H_0$.

| Noise Power | Statistic | Conclusion |
|-------------|-----------|------------|
| $N_0 = 0.001$ | 0.415 | Fail to reject $H_0$ for all five cases |
| $N_0 = 0.003$ | 0.487 | Fail to reject $H_0$ for all five cases |
| $N_0 = 0.005$ | 0.554 | Fail to reject $H_0$ for 2, 3, 4 and 5 |
| $N_0 = 0.007$ | 0.613 | Fail to reject $H_0$ for 2, 3, 4, and 5 |
| $N_0 = 0.009$ | 0.663 | Fail to reject $H_0$ for 3, 4, and 5 |
| $N_0 = 0.01$  | 0.687 | Fail to reject $H_0$ for 3, 4, and 5 |

TABLE I: Anderson-Darling Test
| Noise Power | Statistic | p-Value  | Conclusion       |
|-------------|-----------|----------|------------------|
| $N_0 = 0.001$ | 0.977     | 0.418    | Fail to reject $H_0$ |
| $N_0 = 0.003$ | 0.974     | 0.320    | Fail to reject $H_0$ |
| $N_0 = 0.005$ | 0.970     | 0.242    | Fail to reject $H_0$ |
| $N_0 = 0.007$ | 0.968     | 0.191    | Fail to reject $H_0$ |
| $N_0 = 0.009$ | 0.966     | 0.153    | Fail to reject $H_0$ |
| $N_0 = 0.01$  | 0.965     | 0.137    | Fail to reject $H_0$ |

TABLE II: Shapiro Test

that the samples come from a Gaussian distribution given by (52) for most significance level and critical values for the high signal-to-noise ratio scenarios. Moreover, the p-values for the Shapiro test also remain higher than the predefined critical value of 0.05, and this test fails to reject $H_0$ for a range of $N_0$ values as shown in Table II. From the tests conducted in this section, we conclude that the centroid error is roughly distributed as a Gaussian random variable for $\lambda_p = 1$ Watts and $N_0 \leq 0.01$ Watts/mm$^2$.

VII. Bias/Variance Analysis of Centroid Estimator

The mean-square error for the estimation of beam center $(x_0, y_0)$ is given by [25]

$$\text{MSE}(\theta) \triangleq \text{Tr}(C(\theta)) + b^T(\theta)b(\theta),$$

(53)

where $\theta \triangleq \begin{bmatrix} x_0 & y_0 \end{bmatrix}^T$, $C(\theta)$ is the $2 \times 2$ covariance matrix of $\theta$, $b(\theta)$ is the bias vector: $b(\theta) \triangleq \begin{bmatrix} b_{x_0}(x_0) & b_{y_0}(y_0) \end{bmatrix}^T$. The mean-square error is defined as $\text{MSE}(\theta) \triangleq \mathbb{E}\left( (\theta - \hat{\theta})^T(\theta - \hat{\theta}) \right)$.

A. Discrete Array

From (27) and (28), we know that the bias function of a centroid estimator for a discrete array is given by

$$b_{x_0}(x_0) \triangleq \sum_{m=1}^{M} x_m \frac{\Lambda_m}{(1 - \alpha)\lambda_p} - x_0$$

(54)

$$b_{y_0}(y_0) \triangleq \sum_{m=1}^{M} y_m \frac{\Lambda_m}{(1 - \alpha)\lambda_p} - y_0.$$  

(55)

The variance of the centroid estimator for discrete case is given by (30) and (31). Thus, by (53), we have that

$$\text{MSE}(\theta) = \sigma^2_{x_0}(x_0, y_0) + \sigma^2_{y_0}(x_0, y_0) + b^2_{x_0}(x_0) + b^2_{y_0}(y_0).$$

(56)
B. Continuous Array

Since the centroid estimate is unbiased for a continuous array, the error in (10) is zero-mean. In addition to this, \( \hat{x}_0 \) and \( \hat{y}_0 \) are Gaussian and uncorrelated, and therefore, independent\(^6\) for a continuous array. Thus, we have that \( R_0 \) is distributed as a Rayleigh random variable with scale parameter \( \sigma_{R_0} = \sigma_{\hat{x}_0} = \sigma_{\hat{y}_0} \).

We note from (44) that \( \sigma^2_{\hat{x}_0}(x_0) \) and \( \sigma^2_{\hat{y}_0}(y_0) \) are functions of \( x_0 \) and \( y_0 \), respectively. In order to obtain the “average” variance (which we now denote by \( \sigma^2_{\hat{x}_0} \) or \( \sigma^2_{\hat{y}_0} \) to indicate no explicit dependence on \( x_0 \) and \( y_0 \)), we average over the pdf of \( x_0 \) and \( y_0 \). For the sake of simplicity, let us look at only the variance of \( x_0 \) (since the variance of \( y_0 \) will follow in the similar manner). The pdf of \( x_0 \) is a Gaussian density function: \( x_0 \sim N(0, \sigma^2_{\hat{x}_0}) \). Therefore, we have that,

\[
\sigma^2_{\hat{x}_0} \triangleq \int_{-\infty}^{\infty} \sigma^2_{\hat{x}_0}(x_0) f(x_0) dx_0 = \int_{-\infty}^{\infty} \frac{N_0 |A|^2}{12(1 - \alpha)^2 \lambda_p^2} \left( 1 + \frac{12 \sigma^2_{\hat{x}_0}}{|A|} \right) f(x_0) dx_0
\]

\[
= \frac{N_0 |A|^2}{12(1 - \alpha)^2 \lambda_p^2} \left( 1 + \frac{12 \sigma^2_{\hat{x}_0}}{|A|} \right) = \frac{N_0 |A|^2}{12(1 - \alpha)^2 \lambda_p^2} \left( 1 + \frac{12 (\sigma^2_{R_J} + \sigma^2_{R_0})}{|A|} \right)
\]

(57)

\[
= \frac{N_0 |A|^2}{12(1 - \alpha)^2 \lambda_p^2} \left( 1 + \frac{12 (\sigma^2_{R_J} + \sigma^2_{\hat{x}_0})}{|A|} \right),
\]

(58)

where the second equality in (57) comes from the independence of \( R_J \) and \( R_0 \). The independence of these two random variables can be proved quickly. Let us assume that \( f(r_j, r_e) \) is the joint distribution of \( R_J \) and \( R_0 \), and \( f(r_e|r_j) \) is the conditional distribution of \( R_0 \) given \( R_J = r_j \). Additionally, we also know that \( \mathbb{E}[R_J] = \mathbb{E}[R_0] = 0 \). Therefore,

\[
\mathbb{E}[R_J R_0] = \int_{0}^{\infty} \int_{0}^{\infty} r_j r_e f(r_j, r_e) dr_j dr_e = \int_{0}^{\infty} r_j \int_{0}^{\infty} r_e f(r_e|r_j) dr_e f(r_j) dr_j,
\]

(59)

where \( \int_{0}^{\infty} r_e f(r_e|r_j) dr_e = 0 \) since the centroid error is zero mean for a continuous array for any \( r_j \). Thus, it follows that \( \mathbb{E}[R_0 R_J] = 0 = \mathbb{E}[R_J] \mathbb{E}[R_0] \), which implies that \( R_0 \perp R_J \) since both \( R_J \) and \( R_0 \) are Gaussian random variables.

After a few easy manipulations of (58), we have that

\[
\sigma^2_{\hat{x}_0} = \frac{N_0 |A| (|A| + 12 \sigma^2_{R_J})}{12 (|A| + 12 \sigma^2_{R_J} - N_0 |A|)},
\]

(60)

and similarly, it follows that

\[
\sigma^2_{\hat{y}_0} = \frac{N_0 |A| (|A| + 12 \sigma^2_{R_J})}{12 (|A| + 12 \sigma^2_{R_J} - N_0 |A|)}.
\]

\(^6\)The independence follows from the (approximate) Gaussianity of \( \hat{x}_0 \) and \( \hat{y}_0 \).
Finally, because of the unbiasedness of the centroid for a continuous array, we have that

$$\text{MSE}(\theta) = \text{Tr}(C_e) = \text{Tr} \begin{pmatrix} \sigma_{x_0}^2 & 0 \\ 0 & \sigma_{y_0}^2 \end{pmatrix} = \sigma_{x_0}^2 + \sigma_{y_0}^2,$$

(61)

Fig. 4 compares the mean-square error and the bias performance of the centroid estimator as a function of noise parameter $N_0$ and $\rho_t$. The figure indicates the the performance of a discrete array approaches the continuous array as $M \to \infty$. Roughly, the performance is comparable to a continuous array when the number of detectors $M$ is equal to or greater than 16.

In the following section, we will carry out the optimization of error and outage probabilities assuming that the array is continuous. A continuous array leads to an optimization problem formulation that is easier to analyze as compared to discrete arrays. We expect that the optimization results obtained with a continuous array would be roughly the same for a discrete array whose number of detectors $M \geq 16$.

VIII. Optimization Problems

In order to simplify the analysis for the purpose of optimization, we use the “point-detector” approximation for (3) in this paper defined as follows. When $\rho/a \gg 1$, (3) can be approximated by

$$P(R) \approx \int_{R-a}^{R+a} \frac{\lambda_p}{2\pi \rho_d^2} \exp \left( -\frac{(x^2 + y^2)}{2\rho_d^2} \right) \, dy \, dx \approx \alpha \frac{\lambda_p}{2\pi \rho_d^2} \exp \left( -\frac{R^2}{2\rho_d^2} \right) \pi a^2. \quad (62)$$

This approximation gets rid of the double integral in (3) and saves us the extra computational complexity in carrying out the optimization.
Additionally, as discussed towards the end of Section VII, the optimization problem is solved under the assumption that the array is continuous. The area of the array is fixed at 4 mm$^2$, and the single-detector radius $a = 0.05$ mm for all the experiments. The value of $\lambda_p = 1$ Watts and the order of PPM is $M = 8$. Additionally, the value of noise parameter $N_0$ ranges from 0.001 to 0.01 Watts/mm$^2$. This range of values is chosen based on the data available in [19], where the authors have set the noise power equal to $10^{-14}$ Watts/Hz. Let us assume that the bandwidth of the filter is 1 MHz (assuming that the data rate is approximately 1 Megasymbols per second). Then, the resulting noise power is approximately $10^{-8}$ Watts. In our case, the total noise power turns out to be $N_0 \pi a^2$ where the detector radius $a = 0.05$ mm. Hence, the total noise power in our study ranges from approximately $8 \times 10^{-6}$ to $8 \times 10^{-5}$ Watts.

Finally, the optimization is performed by a gradient descent algorithm with a step size of 0.05.

A. Optimization of Probability of Error

1) Uncorrelated Pointing Error

For a high data-rate scheme for which the symbol period $T_s$ is very small, we can assume that $S(R)$ in (5) can be approximated as $S(R) \approx P(R)T_s$. Without loss of generality, we set $T_s = 1$ time unit in order to simplify the analysis. Therefore, we have that $S(R) = P(R)$, and (7) is rewritten as

$$P(e) = \int_0^\infty \left(1 - \left(Q\left(-\frac{\sqrt{2}\sigma}{\sigma_R (\alpha)}\right)\right)^{M-1}\right) f_R(r) \, dr \quad (63)$$

$$= \int_0^\infty \left(1 - \left(Q\left(-\frac{\alpha \lambda_p}{2\pi^2 \sigma_R^2} \exp\left(-\frac{r^2}{2\sigma^2 R(\alpha)}\right)\right)\right)^{M-1}\right) r \sigma_R^2 (\alpha) \exp\left(-\frac{r^2}{2\sigma^2 R(\alpha)}\right) \, dr. \quad (64)$$

In (64), the distribution of $R$ is characterized by a Rayleigh random variable whose scale parameter squared is given by

$$\sigma_R^2 (\alpha) = \sigma_{R_j}^2 + \sigma_{R_e}^2 = \sigma_{R_j}^2 + \frac{N_0 |A| (|A| + 12\sigma_{R_j}^2)}{12 ((1 - \alpha)^2 \lambda_p^2 - N_0 |A|)} \quad (65)$$

since $\sigma_{R_e}^2 = \sigma_{y_0}^2 = \sigma_{y_0}^2$, where $\sigma_{y_0}^2$ is given by (60).

The objective function we want to minimize is the probability of error given by (64).

a) Optimization of $\alpha$:

$$\text{minimize}_{\alpha} \quad P(e)$$

subject to $0 < \alpha < 1$. 

b) Joint Optimization of $\alpha$ and $\rho_d$:

\[
\begin{align*}
\text{minimize} & \quad P(e) \\
\text{subject to} & \quad i) \ 0 < \alpha < 1, \\
& \quad ii) \ \rho_{d_0} < \rho_d < \rho_{d_1}.
\end{align*}
\]

The value of $\rho_{d_0} = 0.1$ mm and $\rho_{d_1} = 0.5$ mm for all the experiments in this paper. The beam radius $\rho_d$ projected on the detector depends on the characteristics of the lens used to focus the beam.

Fig. 5 indicates the probability of error as a function of beam split factor $\alpha$ and beam radius (detection channel) $\rho_d$ for three different values of noise parameter $N_0$. Fig. 6 depicts the probability of error as a function of $\alpha$ for different values of jitter standard deviation $\sigma_R$. We note from these figures that there is a particular value of $\alpha$ and $\rho_d$ that will minimize the probability of error.

Fig. 7 illustrate the variation in optimal values $\alpha^*$ and $\rho_d^*$ as a function of noise parameter $N_0$ for different values of jitter standard deviation $\sigma_R$. We conclude from this figure that the smaller the signal-to-noise ratio, the larger the fraction of energy, $(1 - \alpha)$, that goes into the tracking channel (or the smaller the amount of energy in the data detection channel) in order to minimize the probability of error. Additionally, it follows that a larger jitter corresponds to a smaller signal-to-noise ratio, which again leads to a larger fraction of energy going into the tracking channel for the optimization of error probability. Moreover, in order to minimize the probability of error, a poorer signal-to-noise ratio also leads to the requirement of a more focused beam (a smaller beam radius $\rho_d$) in the data detection channel.
24

Fig. 6. This figure shows probability of error as a function of $\alpha$ for different values of jitter standard deviation $\sigma_{RJ}$. The value of other parameters are as follows: noise power is $N_0 = 0.001$ Watts/mm$^2$, beam radius $\rho_d = 0.2$ mm, and $M = 8$.

Fig. 7. This figure shows the plots of jointly optimized parameters, $\alpha^*$ (left) and $\rho_d^*$ (center) as a function of noise power $N_0$ for different values of jitter standard deviation $\sigma_{RJ}$. The figure on the right shows the minimum probability of error as a function of noise power for optimized $\alpha^*$ and $\rho_d^*$.

2) Correlated Pointing Error

In this case, we assume that the $x$ and $y$ components of the jitter are correlated. The more general expression for the probability of error in this case is

$$P(e) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(1 - Q \left( -\frac{\alpha \frac{\lambda_0}{2\pi \rho_d} \exp \left( -\frac{(x^2+y^2)}{2\rho_d^2} \right) \pi \sigma^2 \right) \right)^{M-1} f_{X_0Y_0}(x,y) \, dx \, dy.$$

where

$$f_{X_0Y_0}(x,y) \triangleq \frac{1}{2\pi \sigma^2_R(\alpha) \sqrt{1-\gamma^2}} \exp \left( -\frac{(x^2+y^2-2\gamma xy)}{2(1-\gamma^2)\sigma^2_R(\alpha)} \right).$$

where $\gamma$ is the correlation coefficient and $\sigma^2_R(\alpha)$ is furnished by (65). In the probability density function represented by (67), we have assumed equal variance in each of the two dimensions: $\sigma^2_X = \sigma^2_Y = \sigma^2_R(\alpha)$.

Fig. 8 shows the effect of correlation coefficient of the pointing error on the minimization of error.
Fig. 8. This figure shows the probability of error as a function of $\alpha$ and $\rho_d$ for the continuous array for different values of correlation coefficient between $x_0$ and $y_0$. The value of noise parameter $N_0 = 0.001$ Watts/mm$^2$, and the jitter variance $\sigma^2_{R_j} = 0.01$ mm. For the plot on left, the beam radius $\rho_d = 0.2$ mm, and for the plot on right, $\alpha = 0.5$.

We see that a larger value of correlation coefficient implies a higher probability of error.

B. Optimization of Probability of Outage

We say that the optical receiver experiences an “outage” if the received signal-to-noise ratio goes below a certain threshold $\Upsilon_{th}$. The signal-to-noise ratio for the point-detector case is

$$\text{SNR} = \frac{P(R)}{N_0 \pi a^2} = \frac{\alpha \lambda}{2 \pi \rho_d^2} \exp \left( \frac{R^2}{2 \rho_d^2} \right),$$

and the outage probability, $P_{out}$, is defined as

$$P_{out} = P \left( \left\{ \frac{\alpha \lambda}{2 \pi \rho_d^2} \exp \left( \frac{-R^2}{2 \rho_d^2} \right) < \Upsilon_{th} \right\} \right) = P \left( \left\{ \exp \left( \frac{-R^2}{2 \rho_d^2} \right) < \frac{\Upsilon_{th} N_0 2 \pi \rho_d^2}{\alpha \lambda_p} \right\} \right)$$

$$= P \left( \left\{ R^2 > 2 \rho_d^2 \left( -\ln \left( \frac{\Upsilon_{th} N_0 2 \pi \rho_d^2}{\alpha \lambda_p} \right) \right) \right\} \right).$$

The quantity $R^2$ is an exponential random variable with mean $2\sigma^2_R$. Thus, $P_{out}$ in this case is the complementary cumulative distribution function of this exponential random variable. Therefore,

$$P_{out} = \exp \left( -\frac{2 \rho_d^2 \left( -\ln \left( \frac{\Upsilon_{th} N_0 2 \pi \rho_d^2}{\alpha \lambda_p} \right) \right)}{2 \sigma^2_R(\alpha)} \right) = \exp \left( \frac{\rho_d^2 \ln \left( \frac{\Upsilon_{th} N_0 2 \pi \rho_d^2}{\alpha \lambda_p} \right)}{\sigma^2_R(\alpha)} \right)$$

where $\sigma^2_R(\alpha)$ is supplied by (65).

Fig. 9 shows the probability of outage as a function of $\alpha$ for different values of noise parameter $N_0$ and threshold $\Upsilon_{th}$. We note from Fig. 10 a similar trend in the optimized values of $\alpha$ and $\rho_d$ as the probability
### C. Optimal Power Allocation for a Fading Channel

In this section, we optimize the power split factor \( \alpha \) for a fading channel. We have considered two types of common fading techniques, the exponentiated Weibull fading and the lognormal fading, that simulate both the strong and weak turbulent regimes via selection of appropriate parameters.

1) **Exponentiated Weibull Fading**

The exponentiated Weibull (EW) distribution has been shown to provide a better fit to the experimental data under weak and moderate turbulence and all aperture averaging conditions when compared to lognormal and Gamma-Gamma distributions [26]. Another attractive property of EW distribution is the
tractable and closed-form expression of its probability density function which is given by

\[ f(z) = \frac{ab}{\eta} \left( \frac{z}{\eta} \right)^{b-1} \exp \left( - \left( \frac{z}{\eta} \right)^b \right) \left( 1 - \exp \left( - \left( \frac{z}{\eta} \right)^b \right) \right)^{a-1} \cdot 1_{[0,\infty)}(z). \] (71)

In (71), \( b \) is a shape parameter that is related to the scintillation index, \( a \) is another shape parameter that depends on the receiver aperture size, and \( \eta \) is the scale parameter. The function \( 1_A(x) \) is the indicator function over some measurable set \( A \).

2) Lognormal Fading

The expression for lognormal fading is given by

\[ f(z) = \frac{1}{\sqrt{2\pi}\sigma z} \exp \left( -\frac{(\ln z + \sigma^2/2)^2}{2\sigma^2} \right) \cdot 1_{[0,\infty)}(z), \] (72)

where \( \sigma \) is known as the scintillation level, and its value typically ranges between 0.1 and 0.5, but will never exceed 0.75 [27].

3) Error and Outage Probabilities

The probability of error for the fading channel is given by

\[ P(e) = \int_0^\infty \int_0^\infty \left( 1 - Q \left( -\frac{\alpha z^2}{2\pi} \exp \left( -\frac{r^2}{2\rho^2} \pi a^2 \right) \right) \right)^{M-1} \frac{r}{\sigma_R^2(\alpha, z)} \exp \left( -\frac{r^2}{2\sigma_R^2(\alpha, z)} \right) f(z) \, dr \, dz \] (73)

where

\[ \sigma_R^2(\alpha, z) = \sigma^2_{RJ} + \frac{N_0 |A| (|A| + 12\sigma^2_{RJ})}{12 ((1 - \alpha)^2z^2 - N_0 |A|)}, \] (74)

and \( f(z) \) is the fading distribution.

The outage probability is given by

\[ P_{\text{out}} = \int_0^\infty \exp \left( \frac{\rho^2}{\sigma^2_R(\alpha, z)} \ln \left( \frac{\chi_{\nu, N_0 2\pi \rho^2}}{\alpha z} \right) \right) f(z) \, dz. \] (75)

Fig. 11 shows the curves for the pdf’s of exponentiated Weibull and lognormal fading for different parameter values. These parameter values are taken from [26] for different fading and aperture averaging conditions. Fig. 12 depicts the probability of error as a function of \( \alpha \) for the two fading distributions. We note from these two figures that harsher fading conditions lead to a smaller optimum \( \alpha \) (more power in
the tracking channel) that will minimize the error probability.

IX. Conclusions

In this paper, we have analyzed the optimal division or allocation of power between beam tracking and data detection channels in a free-space optical receiver. In this regard, we analyzed the error and outage probability optimization under different scenarios of pointing error variance, signal-to-noise ratio values, threshold values (for outage probability) and different fading conditions. The main conclusion from this study is that for large signal-to-noise ratio, most of the power should be diverted to the data detection channel for optimal error and outage probability performance. However, as the signal-to-noise ratio degrades, a higher fraction of received energy should be channeled into the tracking system in order to minimize the error and outage probabilities.
ACKNOWLEDGMENT

The authors would like to thank Heyou Liu for his contribution in the simulations and plots related to Fig. 5 and Fig. 9 in this paper.

REFERENCES

[1] A. Trichili, M. A. Cox, B. S. Ooi, and M.-S. Alouini, “Roadmap to free space optics,” J. Opt. Soc. Am. B, vol. 37, no. 11, pp. A184–A201, Nov 2020. [Online]. Available: http://josab.osa.org/abstract.cfm?URI=josab-37-11-A184

[2] Y. Kaymak, R. Rojas-Cessa, J. Feng, N. Ansari, M. Zhou, and T. Zhang, “A survey on acquisition, tracking and pointing mechanisms for mobile free-space optical communications,” IEEE Communications Surveys & Tutorials, vol. 20, no. 2, Second Quarter 2018.

[3] Q. Li, S. Xu, J. Yu, L. Yan, and Y. Huang, “An improved method for the position detection of a quadrant detector for free space optical communication,” Sensors, vol. 19, no. 1, p. 175, Jan 2019. [Online]. Available: http://dx.doi.org/10.3390/s19010175

[4] M. S. Ferraro, W. R. Clark, W. S. Rabinovich, R. Mahon, J. L. Murphy, P. G. Goetz, L. M. Thomas, H. R. Burris, C. I. Moore, W. D. Waters, K. Vaccaro, and B. D. Krejca, “InAlAs/InGaAs avalanche photodiode arrays for free space optical communication,” Appl. Opt., vol. 54, no. 31, pp. F182-F188, Nov 2015. [Online]. Available: http://ao.osa.org/abstract.cfm?URI=ao-54-31-F182

[5] M. S. Bashir, M.-C. Tsai, and M.-S. Alouini, “Robust beam position estimation with photon counting detector arrays in free-space optical communications,” Journal of Lightwave Technology, November 2020, submitted for publication (available on engrXiv at https://engrxiv.org/nqf3y/).

[6] M. S. Bashir and M.-S. Alouini, “Free-space optical MISO communications with an array of detectors,” IEEE Open Journal of the Communications Society, pp. 1–1, 2020.

[7] J. Zhang, W. Qian, G. Gu, K. Ren, Q. Chen, C. Mao, G. Cai, Z. Liu, and L. Xu, “Quadrant response model and error analysis of four-quadrant detectors related to the non-uniform spot and blind area,” Appl. Opt., vol. 57, no. 24, pp. 6898–6905, Aug 2018. [Online]. Available: http://ao.osa.org/abstract.cfm?URI=ao-57-24-6898

[8] M. S. Ferraro, R. Mahon, W. S. Rabinovich, W. T. Freeman, J. L. Murphy, P. G. Goetz, H. R. Burris, C. I. Moore, L. M. Thomas, W. R. Clark, and W. D. Waters, “InGaAs avalanche photodiode arrays for simultaneous communications and tracking,” in Free-Space and Atmospheric Laser Communications XI, A. K. Majumdar and C. C. Davis, Eds., vol. 8162, International Society for Optics and Photonics. SPIE, 2011, pp. 101 – 108. [Online]. Available: https://doi.org/10.1117/12.895116

[9] M. S. Bashir and M. R. Bell, “Optical beam position estimation in free-space optical communication,” IEEE Transactions on Aerospace and Electronic Systems, vol. 52, no. 6, December 2016.

[10] ——, “Optical beam position tracking in free-space optical communication systems,” IEEE Transactions on Aerospace and Electronic Systems, vol. 20, no. 2, April 2018.

[11] M. S. Bashir and M. R. Bell, “The impact of optical beam position estimation on the probability of error in free-space optical communications,” IEEE Transactions on Aerospace and Electronic Systems, vol. 55, no. 3, pp. 1319–1333, June 2019.

[12] M. S. Bashir, “Free-space optical communications with detector arrays: A mathematical analysis,” IEEE Transactions on Aerospace and Electronic Systems, vol. 56, no. 2, pp. 1420–1429, 2020.

[13] M. Srinivasan, K. S. Andrews, W. H. Farr, and A. Wong, “Photon counting detector array algorithms for deep space optical communications,” in Free-Space Laser Communication and Atmospheric Propagation XXVIII, H. Hemmati and D. M. Boroson, Eds., vol. 9739, International Society for Optics and Photonics. SPIE, 2016, pp. 267 – 282. [Online]. Available: https://doi.org/10.1117/12.2217971
[14] V. A. Vilnrotter and M. Srinivasan, “Adaptive detector arrays for optical communications receivers,” IEEE Transactions on Communications, vol. 50, no. 7, pp. 1091–1097, July 2002.

[15] V. Vilnrotter, C. . Lau, M. Srinivasan, K. Andrews, and R. Mukai, “Optical array receiver for communication through atmospheric turbulence,” Journal of Lightwave Technology, vol. 23, no. 4, pp. 1664–1675, April 2005.

[16] I. Ansari, F. Yilmaz, and M. -S. Alouini, “Performance analysis of free-space optical links over Malaga (M) turbulence channels with pointing errors,” IEEE Transactions on Wireless Communications, vol. 15, no. 1, pp. 91–102, January 2016.

[17] E. Zedini, H. Soury, and M. -S. Alouini, “Dual-hop FSO transmission systems over gamma-gamma turbulence with pointing errors,” IEEE Transactions on Wireless Communications, vol. 16, no. 2, pp. 784–796, February 2017.

[18] H. Al-Quwaiee, H. C. Yang, and M. -S. Alouini, “On the asymptotic capacity of dual-aperture FSO systems with a generalized pointing error model,” IEEE Transactions on Wireless Communications, vol. 15, no. 9, pp. 6502–6512, September 2016.

[19] A. A. Farid and S. Hranilovic, “Outage capacity optimization for free-space optical links with pointing errors,” Journal of Lightwave Technology, vol. 25, no. 7, July 2007.

[20] V. V. Mai and H. Kim, “Adaptive beam control techniques for airborne free-space optical communication systems,” Applied Optics, vol. 57, no. 26, September 2018.

[21] M. K. Simon and V. A. Vilnrotter, “Alamouti-type space-time coding for free-space optical communication with direct detection,” IEEE Transactions on Wireless Communications, vol. 4, no. 1, pp. 35–39, 2005.

[22] D. V. Hinkley, “On the ratio of two correlated normal random variables,” Biometrika, vol. 56, no. 3, pp. 635–639, 1969. [Online]. Available: http://www.jstor.org/stable/2334671

[23] T. W. Anderson and D. A. Darling, “A test of goodness of fit,” Journal of the American Statistical Association, vol. 49, no. 268, pp. 765–769, 1954. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1080/01621459.1954.10501232

[24] S. S. Shapiro and M. B. Wilk, “An analysis of variance test for normality (complete samples),” Biometrika, vol. 52, no. 3/4, pp. 591–611, 1965. [Online]. Available: http://www.jstor.org/stable/2333709

[25] S. M. Kay, Fundamentals of statistical signal processing, ser. Prentice Hall signal processing series. Upper Saddle River, NJ: Prentice Hall PTR, 1993. [Online]. Available: https://cds.cern.ch/record/2012069

[26] R. Barrios and F. Dios, “Exponentiated weibull distribution family under aperture averaging for gaussian beam waves,” Opt. Express, vol. 20, no. 12, pp. 13055–13064, Jun 2012. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-20-12-13055

[27] L. Yang, X. Song, J. Cheng, and J. F. Holzman, “Free-space optical communications over lognormal fading channels using OOK with finite extinction ratios,” IEEE Access, vol. 4, pp. 574–584, 2016.