

# Robust Beam Position Estimation with Photon Counting Detector Arrays in Free-Space Optical Communications

Muhammad Salman Bashir, *Senior Member, IEEE*, Ming-Cheng Tsai, *Student Member, IEEE*,  
and Mohamed-Slim Alouini, *Fellow, IEEE*

## Abstract

Optical beam center position on an array of detectors is an important parameter that is essential for estimating the angle-of-arrival of the incoming signal beam. In this paper, we have examined the beam position estimation problem for photon-counting detector arrays, and to this end, we have derived and analyzed the *Cramér-Rao lower bounds* on the mean-square error of the unbiased estimators of the beam position. Furthermore, we have also derived the Cramér-Rao lower bounds of other beam parameters such as peak intensity, and the intensity of background radiation on the array. In this sense, we have considered the problem of robust estimation of beam position in which none of the parameters are assumed to be known beforehand. Additionally, we have derived the Cramér-Rao lower bounds of beam parameters for observations based on both pilot and data symbols of a *pulse position modulation* (PPM) scheme. Finally, we have considered a two-step estimation problem in which the peak intensity and background radiation are estimated using a *method of moments* estimator, and the beam center position is estimated with the help of a *maximum likelihood* estimator.

## Index Terms

Angle-of-arrival, beam center position, Cramér-Rao lower bound, maximum likelihood estimator, method of moments estimator, photon-counting detector arrays, pulse position modulation.

## I. INTRODUCTION

Free-space optical (FSO) communications has typically been employed in deep-space communications due to the low divergence of optical beam that can help transmit data over much longer distances. However, the urgent need to provide connectivity to almost 4 billion people who currently do not have internet access has prompted the integration of FSO and millimeter wave/Terahertz communications in the backhaul of 6G wireless communication systems. Thus, FSO/millimeter/Terahertz wave links will be deployed in an integrated network of satellites, drones, high altitude platform and balloons that will form the backhaul of the next generation of wireless communications. In this regard, FSO is an emerging candidate for future communications due to its ability to support high data rates.

However, the problem of pointing, acquisition and tracking is significant in the FSO (as well as millimeter/Terahertz systems) domain because of the narrow beam widths associated with the optical signal. Acquisition is the process in which the two terminals acquire the initial location of each other before the actual data communication begins. However, after the acquisition is achieved, the system still needs to maintain the alignment between the transmitter and receiver assemblies due to physical factors such as random effects associated with atmospheric turbulence, the mechanical vibrations introduced in the transmitter/receiver assemblies due to random physical forces, or building sways due to wind vibrations. This misalignment leads to a loss of received signal energy at the receiver that may increase the outage

This work is supported by Office of Sponsored Research (OSR) at King Abdullah University of Science and Technology (KAUST).

M. S. Bashir, M. -C. Tsai and M. -S. Alouini are with the King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Kingdom of Saudi Arabia. e-mail: [salman.bashir@alumni.purdue.edu](mailto:salman.bashir@alumni.purdue.edu), [mingcheng.tsai@kaust.edu.sa](mailto:mingcheng.tsai@kaust.edu.sa), [slim.alouini@kaust.edu.sa](mailto:slim.alouini@kaust.edu.sa).

probability. Additionally, it goes without saying that the beam needs to be tracked/aligned for mobile platforms since the angle-of-arrival changes continuously due to relative motion.

In this paper, we consider the optical beam position estimation problem for a deep space optical communication system that employs multiple photon-counting detectors (array of detectors or a focal plane array) instead of one large (monolithic) detector at the receiver. Photon-counting detector arrays are typically used in deep space optical communications because of their ability to detect very low levels of received light (a few signal photons) [1]. Any changes in the angle-of-arrival of the beam on the receive aperture leads to a shift of the center of the focused beam on the focal plane array. Hence, the problem of estimating the angle-of-arrival is reduced to estimating the beam center deviation on the array. The deviation between the beam center and the array center (a measure of misalignment) is measured with a beam position estimation algorithm, and the misalignment may be corrected by orienting the receiver telescope in the direction of the incoming beam.

We further want to point out the advantage obtained with an array of detectors: It can be used for symbol detection as well as for estimating the beam position. Thus, a receiver based on an array of detectors is more efficient in terms of energy since no portion of the received energy has to be diverted to a separate tracking assembly<sup>1</sup>, and all the received energy can be used for detection of the transmitted symbol. However, a disadvantage associated with an array of detectors is the higher computational/hardware complexity as compared to a single detector. The complexity scales as  $\mathcal{O}(M)$  where  $M$  is the number of detectors in the array.

## II. LITERATURE REVIEW AND CONTRIBUTIONS/ORGANIZATION OF THIS PAPER

### A. Background Literature Review

There is a significant number of studies carried out on research in pointing, acquisition and tracking (PAT) systems in FSO that treat the tracking problem purely from a hardware point-of-view. In this respect, [2] provides a detailed overview of the current state-of-the-art hardware solutions for tracking the optical beam. Thus, we will cover the literature review from the theoretical/signal processing perspective since such a perspective is more relevant to our study in this paper.

The authors in [3] and [4] have discussed the performance of a proposed feedback (beam) tracking loop that acts on the error signal provided by a quadrant photodetector in the receiver assembly. The work in [4] actually builds on the study presented in [3] by carrying out the stability analysis of their proposed cooperative feedback loop. The authors in [5] present the performance analysis of centroid and maximum likelihood estimators of beam position for a “continuous”<sup>2</sup> array. Regarding the literature that covers communications with detector arrays in free-space optics, the authors in [6] propose beam position estimation algorithms and examine their mean-square error performance with simulations. The work in [7] extends the work in [6] by introducing Bayesian filtering algorithms, such as Kalman and particle filters, for tracking the time-varying beam position. The authors in [8] inspect the relationship between the probability of error and the estimation of beam position on the detector array, and using an argument based on Chernoff bounds, they show that precise estimation of beam center on the array is necessary in order to minimize the probability of error. Additionally, the author in [9] presents a mathematical argument to show that the probability of error decreases monotonically as the number of detectors in the array is increased. Furthermore, the authors in [10] analyze the acquisition performance of an FSO system that employs an array of detectors at the receiver. Finally, the authors in [11] consider time synchronization schemes based on an array of detectors.

Furthermore, we also briefly discuss the literature on pointing and tracking in FSO systems that examine the tracking problem from the perspective of a single detector. In this regard, [12] develops the pointing

<sup>1</sup>Typically, a quadrant photodetector is employed in the tracking assembly in order to track beam position.

<sup>2</sup>A continuous array is obtained if the number of detectors in the array goes to infinity for a fixed array area. In other words, the area per detector approaches zero for a continuous array. In this case, we have perfect information about the location of each photodetection in the continuous array. Therefore, continuous arrays lead to an optimal mean-square error performance.

error statistics for a circularly shaped detector and a Gaussian beam, and the outage capacity is optimized as a function of beam radius. The authors in [13] investigate a slightly different optimization problem concerning pointing: They have considered the maximization of link availability as a function of beam radius (for fixed signal power). Additionally, they also explore the minimization of transmitted power by tuning to the optimal beam radius under the constraint of a fixed link availability. In addition to these papers, the interested reader may be directed to [14]–[17] for a detailed study on the performance of FSO systems when the optical channel suffers degradation due to pointing errors for a single-detector receiver.

Readers who might be interested in deep space optical communications with photon-counting detector arrays are referred to [18]–[22].

### *B. Contributions of This Paper*

Even though the authors in [6] have proposed beam position estimation algorithms with an array of detectors, they have not considered the derivation of the Cramér-Rao lower bounds on the variance of unbiased estimators. We believe that an understanding of the Cramér-Rao bounds of beam position and other parameters is an important problem as these bounds will give us important insights into the behavior of estimators under different channel conditions. Additionally, beam parameters other than the beam position—such as peak intensity, beam radius and background radiation level<sup>3</sup>—are assumed by these authors as known quantities in their derivation of estimators. However, even though the beam radius on the focal plane array may be considered constant, the peak intensity and the background radiation intensity may change significantly with time, and they have to be estimated in real-time in order to improve the performance of beam position estimators. Additionally, it goes without saying that a knowledge of peak intensity and background noise is also important for allocating power in different channels of a multiple-input-multiple-output (MIMO), or a multiple-input-single-output (MISO), FSO communication system as these two quantities specify the signal-to-noise ratio of the channel [23].

In this study, we have derived and analyzed the Cramér-Rao lower bounds of beam position estimators for an array of detectors. The estimation problem discussed in this paper is robust since we also estimate the signal intensity as well as background radiation/noise power levels. In this regard, Cramér-Rao lower bounds are derived for beam position, peak intensity and noise intensity for various scenarios. Moreover, two types of observations are considered in this estimation problem: i) observations based on pilot symbols and ii) observations based on data symbols. Using energy of the signal based on data symbols for our estimation problem leads to a more bandwidth/energy efficient scheme. However, as we will see later in this study, the estimation performance (in terms of mean-square error) corresponding to data symbols suffers more at poorer signal-to-noise ratio as compared to pilot symbols.

For observations based on data symbols, we additionally consider two different observation intervals for estimation: i) a PPM symbol period, or, ii) PPM slot period containing the signal pulse. The estimation based on the slot period assumes that the signal pulse is present in the slot. If this assumption is true, the signal-to-noise ratio in the slot is much higher (as much as  $K$  times) than the signal-to-noise ratio in one  $K$ -PPM symbol period. This high signal-to-noise ratio leads to a better estimation performance. However, if the receiver makes an error, we use the “noise only” slot as our observation which results in a severe degradation in estimation performance. We discuss these ideas in more detail in Section V.

### *C. Model Assumptions*

One of the major assumptions in this study is that the array area is chosen to be large enough so that the beam footprint is much smaller than the footprint of the array. A practical system design requires that such an assumption should hold so that any outage in the received signal is precluded in case the beam wanders due to angle-of-arrival fluctuations. Furthermore, a large array area is also needed in order to

<sup>3</sup>Even though the background radiation is not strictly related to the beam, we loosely term it as another beam parameter.

| Symbols             | Definition                                                 |
|---------------------|------------------------------------------------------------|
| $I_0$               | peak beam intensity (average number of photons)            |
| $\lambda_n$         | noise intensity (average number of noise photons)          |
| $\rho$              | beam radius (mm)                                           |
| $(x_0, y_0)$        | location of the beam center on the array                   |
| $\mathcal{A}$       | region of the detector array                               |
| $ \mathcal{A} $     | area of the detector array (mm <sup>2</sup> )              |
| $M$                 | number of detectors in the array                           |
| $N$                 | number of pilot symbols                                    |
| $A_m$               | region of the $m$ th detector in the array                 |
| $A$                 | area of a given detector (mm <sup>2</sup> )                |
| $\ell(\mathcal{R})$ | length of one side of any square region $\mathcal{R}$ (mm) |

TABLE I: List of mathematical symbols

track the movement of the beam and align it—perhaps through a *fast steering mirror* (FSM) assembly—to the center of the array.

Secondly, all arrays are assumed to be of a square shape and each detector in the array is also assumed to be of a square shape as well. This is required in order to make the mathematical analysis more tractable.

Finally, the focus of this study is on non-Bayesian estimation techniques for beam position estimation. This is due to the fact that unless we are certain about the parameters of the prior random motion model of the beam on the array, we are likely going to incur a significant loss in performance if there is mismatch in our assumptions and the real world parameters<sup>4</sup> [24].

#### D. Organization of This Paper

This paper is organized as follows. Section III defines the beam profile and the Poisson model that governs the occurrence of photodetections<sup>5</sup> in the array of detectors. Section IV discusses the derivation of the Cramér-Rao lower bound of the beam parameter estimation problem when pilot symbols are used as an observation. In Section V, we derive the Cramér-Rao bounds for observations based on data symbols. Section VI considers the two-step estimation (method of moments and maximum likelihood estimators) algorithm to estimate beam parameters. The simulation results are explained in Section VII and Section VIII briefly discusses the complexity of the two estimators. The conclusions of this study are summarized in Section IX.

### III. SYSTEM MODEL

The received optical signal on the receiver aperture gives rise to photoelectrons or photodetections in each detector of the array due to *photoelectric effect*. The emission of these photoelectrons during the signal pulse interval helps us detect transmitted symbols. The photon count  $Z_m$  in the  $m$ th detector or cell of the array—during some specified observation interval—is modeled as a (Poisson) discrete random variable. Its probability mass function is characterized by the following expression:

$$P(\{Z_m = z_m\}) = \frac{\exp\left(-\iint_{A_m} [\lambda_s(x, y) + \lambda_n] dx dy\right) \left(\iint_{A_m} [\lambda_s(x, y) + \lambda_n] dx dy\right)^{z_m}}{z_m!}, \quad m = 1, \dots, M, \quad (1)$$

<sup>4</sup>This is especially true if the parameters themselves—such as the covariance matrices of the random motion model—are time-varying.

<sup>5</sup>We use the term photodetections and photons alternatively in this paper. A photodetection actually corresponds to an avalanche of electrons triggered by either an incoming (signal or noise) photon or a thermal noise electron.

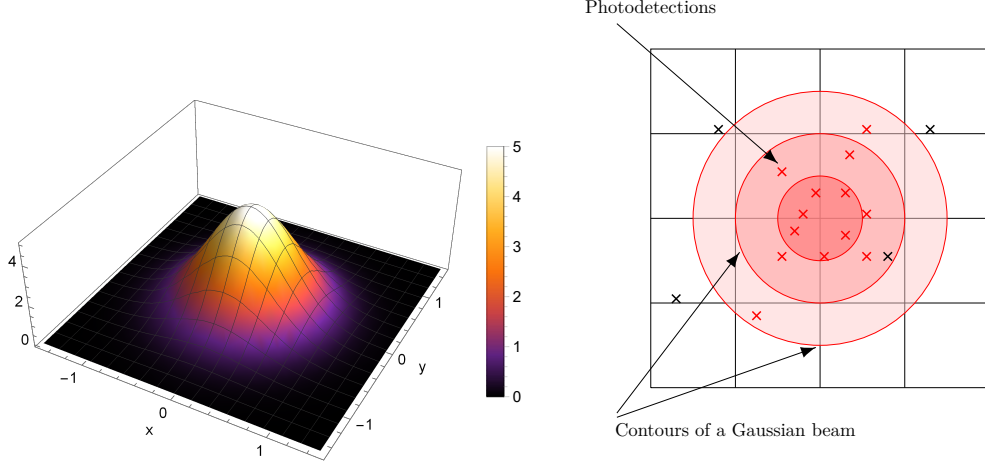


Fig. 1: The figure on left shows the profile of incident beam on an array of detectors. The figure on right depicts contours of incident light intensity and the resulting photodetections for a  $4 \times 4$  detector array. The red crosses represent the locations of signal photodetections, and the black ones correspond to noise.

where  $\lambda_s(x, y)$  is the scaled beam intensity<sup>6</sup> profile on the detector array,  $\lambda_n$  is the scaled noise intensity profile,  $A_m$  is the region of the  $m$ th detector on the detector array,  $Z_1, Z_2, \dots, Z_M$  are independent Poisson random variables and  $M$  is the total number of detectors in the array. As may have been discerned by the reader, the coordinate  $(x, y)$  stands for any point inside the region of the detector array. Moreover,  $\lambda_n$  is constant factor with respect to  $(x, y)$  that accounts for the background radiation and the thermal effects of the detector array [25].

We assume that the airy pattern of the beam on the focal plane array is well-approximated by a Gaussian function (see Fig. 1). The received (scaled) signal intensity at the detector array is given by the expression

$$\lambda_s(x, y) \triangleq \frac{I_0}{\rho^2} \exp\left(\frac{-(x - x_0)^2 - (y - y_0)^2}{2\rho^2}\right) \cdot \mathbf{1}_A(x, y), \quad (2)$$

where  $I_0/\rho^2$  is the peak intensity measured in terms of number of signal photons measured during an observation interval. Furthermore,  $\lambda_n$  is also measured in terms of number of noise photons generated during the same observation interval. The quantity  $\rho$  is known as the beam radius measured in millimeters, and  $(x_0, y_0)$  is the center of the Gaussian beam on the detector array. The function  $\mathbf{1}_A(\cdot)$  represents the *indicator* function over some (measurable) set  $A$ , and  $A$  is the region of the detector array.

Furthermore, it is a general assumption in the following sections that the center of the array has the coordinates  $(0, 0)$ . Additionally, the area of  $A_m$  is denoted by  $A$  since all detectors are assumed to have an equal area. The area of the detector array is denoted by  $|\mathcal{A}|$ . The length of one side of the array is denoted by  $\ell(\mathcal{A})$ . Please see Table I for a complete list of all mathematical symbols.

#### IV. CRAMÉR-RAO LOWER BOUNDS FOR BEAM PARAMETER ESTIMATION BASED ON PILOT SYMBOLS

In this section, we derive and analyze the Cramér-Rao lower bounds for the beam parameter estimation problem based on pilot symbols. The pilot symbol is transmitted as a known pulse position modulation symbol. For instance, we may transmit only the ‘0’ symbol (signal pulse only in the first slot) of a  $K$ -PPM scheme. The observation interval in this case is the first slot of every pilot symbol.

<sup>6</sup>The actual signal intensity,  $\lambda_{s_i}$ , and the actual noise intensity,  $\lambda_{n_i}$ , are both measured in terms of Joules/mm<sup>2</sup>/s. However, they are multiplied by the constant  $\frac{\eta T_p}{hc/\lambda}$  in order to obtain the intensity  $\lambda_s$  and  $\lambda_n$  for the photon generation model in (1). The constant  $h$  is known as the *Planck’s constant*, and its value is  $6.62607004 \times 10^{-34}$  m<sup>2</sup>kg/s. The constant  $c$  is the speed of light in vacuum which is about  $3 \times 10^8$  m/s,  $\lambda$  is the wavelength of light in meters,  $\eta$  stands for the photoconversion efficiency, and  $T_p$  represents signal pulse duration in seconds.

Let  $\boldsymbol{\theta} \triangleq [x_0 \ y_0 \ I_0 \ \lambda_n]^\top$ <sup>7</sup>. The likelihood function is given by

$$p(\mathbf{Z}|\boldsymbol{\theta}) = \prod_{m=1}^M e^{-\Lambda_m} \frac{\Lambda_m^{z_m}}{z_m!}, \quad (3)$$

where

$$\Lambda_m \triangleq \iint_{A_m} \left( \frac{I_0}{\rho^2} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} + \lambda_n \right) dx dy, \quad (4)$$

and the random vector  $\mathbf{Z} \triangleq [Z_1 \ Z_2 \ \cdots \ Z_M]^T$ . Let us define the total incident power on the array  $\Lambda_s \triangleq \sum_{m=1}^M \Lambda_m$ . Then,

$$\ln p(\mathbf{Z}|\boldsymbol{\theta}) = \sum_{m=1}^M z_m \ln \Lambda_m - \Lambda_m - \ln z_m! = \sum_{m=1}^M (z_m \ln \Lambda_m - \ln z_m!) - \Lambda_s. \quad (5)$$

#### A. Estimation Based on Pilot Symbols

In this section, we derive Cramér-Rao lower bounds based on the observations corresponding to pilot symbols.

##### 1) Cramér-Rao Lower Bound of $I_0$

As a first step in computing the Cramér-Rao lower bound for any unbiased estimator  $\hat{I}_0$ , we compute the first partial derivative of (5):

$$\frac{\partial \ln p(\mathbf{Z}|\boldsymbol{\theta})}{\partial I_0} = \sum_{m=1}^M \frac{Z_m}{\Lambda_m} \times \iint_{A_m} \frac{1}{\rho^2} \exp \left( -\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2} \right) dx dy, \quad (6)$$

and then,

$$\frac{\partial^2 \ln p(\mathbf{Z}|\boldsymbol{\theta})}{\partial I_0^2} = - \sum_{m=1}^M \frac{Z_m}{\Lambda_m^2} \times \left( \iint_{A_m} \frac{1}{\rho^2} \exp \left( -\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2} \right) dx dy \right)^2. \quad (7)$$

Now, taking the expectation with respect to  $Z_m$  and taking the negative of the resulting quantity, we have that

$$-\mathbb{E} \left[ \frac{\partial^2 \ln p(\mathbf{Z}|\boldsymbol{\theta})}{\partial I_0^2} \right] = \sum_{m=1}^M \frac{1}{\Lambda_m} \times \left( \iint_{A_m} \frac{1}{\rho^2} \exp \left( -\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2} \right) dx dy \right)^2. \quad (8)$$

##### 2) Cramér-Rao Lower Bound of $\lambda_n$

In this case, we assume that the background radiation is estimated at the receiver while the transmitter is turned off (no signal is present at the receiver). Therefore, in this case,  $\Lambda_m = \lambda_n A$ . Using the same line of argument as used in the derivation of (8), it can be easily shown that the Cramér-Rao lower bound on the variance of any unbiased estimator  $\hat{\lambda}_n$  for one noise-only slot is given by

$$-\mathbb{E} \left[ \frac{\partial^2 \ln p(\mathbf{Z}|\boldsymbol{\theta})}{\partial \lambda_n^2} \right] = \sum_{m=1}^M \frac{A^2}{\Lambda_m}. \quad (9)$$

<sup>7</sup>Here, we want to emphasize that the beam radius on the focal plane array is a known quantity that depends on the focal properties of the aperture lens, and hence does not need to be estimated as such.

### 3) Cramér-Rao Lower Bounds of $(x_0, y_0)$

The Cramér-Rao lower bound for  $\hat{x}_0$  and  $\hat{y}_0$  is derived in Section A of the appendix. The final expressions are produced here as follows:

$$\text{Var}[\hat{x}_0] \geq \frac{\sum_{m=1}^M \frac{1}{\Lambda_m} \left( \iint_{A_m} \frac{I_0}{\rho^4} (y - y_0) e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho^2}} dx dy \right)^2}{\Psi(x_0, y_0, I_0, \rho)}, \quad (10)$$

and

$$\text{Var}[\hat{y}_0] \geq \frac{\sum_{m=1}^M \frac{1}{\Lambda_m} \left( \iint_{A_m} \frac{I_0}{\rho^4} (x - x_0) e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho^2}} dx dy \right)^2}{\Psi(x_0, y_0, I_0, \rho)}, \quad (11)$$

where  $\Psi(x_0, y_0, I_0, \rho)$  is defined in (65).

### 4) Cramér-Rao Lower Bounds for Joint Estimation of $I_0$ and $(x_0, y_0)$

In this section, we state the Cramér-Rao lower bounds for the three-parameter estimation problem in which the three beam parameters are  $x_0, y_0$  and  $I_0$ <sup>8</sup>. We denote the  $3 \times 3$  Fisher Information Matrix by  $\mathcal{I}(x_0, y_0, I_0)$ . The Cramér-Rao lower bounds are given by

$$\begin{aligned} \text{Var}[\hat{x}_0] &\geq [\mathcal{I}^{-1}(x_0, y_0, I_0)]_{1,1} \\ &= \left[ \left( \sum_{m=1}^M \frac{1}{\Lambda_m \rho^4} \left( \iint_{A_m} e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho^2}} dx dy \right)^2 \right) \left( \sum_{m=1}^M \frac{I_0^2}{\Lambda_m \rho^8} \left( \iint_{A_m} e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho^2}} (y - y_0) dx dy \right)^2 \right) \right. \\ &\quad \left. - \left( \sum_{m=1}^M \frac{I_0}{\Lambda_m \rho^6} \iint_{A_m} e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho^2}} dx dy \iint_{A_m} e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho^2}} (y - y_0) dx dy \right) \right] |\mathcal{I}(x_0, y_0, I_0)|^{-1}. \quad (12) \end{aligned}$$

$$\begin{aligned} \text{Var}[\hat{y}_0] &\geq [\mathcal{I}^{-1}(x_0, y_0, I_0)]_{2,2} \\ &= \left[ \left( \sum_{m=1}^M \frac{1}{\Lambda_m \rho^4} \left( \iint_{A_m} e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho^2}} dx dy \right)^2 \right) \left( \sum_{m=1}^M \frac{I_0^2}{\Lambda_m \rho^8} \left( \iint_{A_m} e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho^2}} (x - x_0) dx dy \right)^2 \right) \right. \\ &\quad \left. - \left( \sum_{m=1}^M \frac{I_0}{\Lambda_m \rho^6} \iint_{A_m} e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho^2}} dx dy \iint_{A_m} e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho^2}} (x - x_0) dx dy \right) \right] |\mathcal{I}(x_0, y_0, I_0)|^{-1}. \quad (13) \end{aligned}$$

$$\begin{aligned} \text{Var}[\hat{I}_0] &\geq [\mathcal{I}^{-1}(x_0, y_0, I_0)]_{3,3} \\ &= \left[ \left( \sum_{m=1}^M \frac{I_0^2}{\Lambda_m \rho^8} \left( \iint_{A_m} e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho^2}} (x - x_0) dx dy \right)^2 \right) \left( \sum_{m=1}^M \frac{I_0^2}{\Lambda_m \rho^8} \left( \iint_{A_m} e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho^2}} (y - y_0) dx dy \right)^2 \right) \right. \\ &\quad \left. - \left( \sum_{m=1}^M \frac{I_0^2}{\Lambda_m \rho^8} \iint_{A_m} e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho^2}} (x - x_0) dx dy \iint_{A_m} e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho^2}} (y - y_0) dx dy \right) \right] |\mathcal{I}(x_0, y_0, I_0)|^{-1}. \quad (14) \end{aligned}$$

The determinant of the Fisher information matrix is given by

$$\begin{aligned} |\mathcal{I}(x_0, y_0, I_0)| &= - \sum_{m=1}^M \frac{I_0}{\rho^6 \Lambda_m} \psi_m^{(0)} \psi_m^{(1)} \left( \sum_{m=1}^M \frac{I_0^2}{\rho^8 \Lambda_m} (\psi_m^{(2)})^2 \sum_{m=1}^M \frac{I_0}{\rho^6 \Lambda_m} \psi_m^{(0)} \psi_m^{(1)} - \sum_{m=1}^M \frac{I_0}{\rho^6 \Lambda_m} \psi_m^{(0)} \psi_m^{(2)} \sum_{m=1}^M \frac{I_0^2}{\rho^8 \Lambda_m} \psi_m^{(2)} \psi_m^{(1)} \right) \end{aligned}$$

<sup>8</sup>In order to lower the complexity of the estimation problem, we can estimate  $\lambda_n$  independently of  $x_0, y_0$  and  $I_0$ . In this case, all we need to do is to estimate the average number of noise photons by occasionally turning the transmitter off.

$$\begin{aligned}
& + \sum_{m=1}^M \frac{I_0}{\rho^6 \Lambda_m} \psi_m^{(0)} \psi_m^{(2)} \left( \sum_{m=1}^M \frac{I_0}{\rho^6 \Lambda_m} \psi_m^{(0)} \psi_m^{(1)} \sum_{m=1}^M \frac{I_0^2}{\rho^8 \Lambda_m} \psi_m^{(2)} \psi_m^{(1)} - \sum_{m=1}^M \frac{I_0}{\rho^6 \Lambda_m} \psi_m^{(0)} \psi_m^{(2)} \sum_{m=1}^M \frac{I_0^2}{\rho^8 \Lambda_m} (\psi_m^{(1)})^2 \right) \\
& + \sum_{m=1}^M \frac{1}{\rho^4 \Lambda_m} (\psi_m^{(0)})^2 \left( \sum_{m=1}^M \frac{I_0^2}{\rho^8 \Lambda_m} (\psi_m^{(2)})^2 \sum_{m=1}^M \frac{I_0^2}{\rho^8 \Lambda_m} (\psi_m^{(1)})^2 - \left( \sum_{m=1}^M \frac{I_0^2}{\rho^8 \Lambda_m} \psi_m^{(2)} \psi_m^{(1)} \right)^2 \right). \quad (15)
\end{aligned}$$

where

$$\psi_m^{(0)} \triangleq \iint_{A_m} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy, \quad (16)$$

$$\psi_m^{(1)} \triangleq \iint_{A_m} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} (y-y_0) dx dy, \quad (17)$$

$$\psi_m^{(2)} \triangleq \iint_{A_m} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} (x-x_0) dx dy. \quad (18)$$

### B. Cramér-Rao Lower Bounds of $(x_0, y_0)$ : Asymptotic Case ( $M \rightarrow \infty$ )

We know that each detector in the array counts or reports the photodetections that occur inside its region in a given observation interval for the purpose of beam position estimation. However, the detector does not specify the exact location of the photodetection inside its region. In the ideal case when  $M \rightarrow \infty$  for fixed array area, the true location of each photodetection can be reported by the infinitesimally small detector. When  $M \rightarrow \infty$ , we call this limiting array a “continuous” array. This asymptotic case is of interest since the probability of error/tracking performance of a practical array can be reasonably approximated with the continuous array when the number of detectors is large enough ( $M \geq 64$ ) [6], [9]. Therefore, in this section, we look at the Cramér-Rao lower bound of  $(x_0, y_0)$  for the  $M \rightarrow \infty$  case for the low and high signal-to-noise-ratio regimes, and the convergence rates of the Cramér-Rao lower bounds are derived in terms of beam radius  $\rho$ .

In the following analysis, let us analyze the Cramér-Rao lower bound of  $\hat{x}_0$  only since the same analysis will hold for  $\hat{y}_0$  due to the symmetric nature of the Gaussian beam.

#### 1) Estimation of $x_0$ (High Signal-to-Noise Ratio)

For high signal-to-noise ratio,  $\lambda_n A \ll \iint_{A_m} \frac{I_0}{\rho^2} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy$ . Then,  $\Lambda_m \approx \iint_{A_m} \frac{I_0}{\rho^2} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy$ .

When  $M$  is large,  $\Lambda_m \approx \frac{I_0}{\rho^2} e^{-\frac{(x_m-x_0)^2+(y_m-y_0)^2}{2\rho^2}} \Delta_M$ , where  $(x_m, y_m)$  is the center of the  $m$ th small detector, and  $\Delta_M$  is its infinitesimal area. Then, the numerator of (10) simplifies as

$$\begin{aligned}
& \sum_{m=1}^M \frac{1}{\Lambda_m} \left( \iint_{A_m} \frac{I_0}{\rho^4} (y-y_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy \right)^2 \approx \sum_{m=1}^M \frac{\left( \frac{I_0}{\rho^4} (y_m-y_0) e^{-\frac{(x_m-x_0)^2+(y_m-y_0)^2}{2\rho^2}} \Delta_M \right)^2}{\frac{I_0}{\rho^2} e^{-\frac{(x_m-x_0)^2+(y_m-y_0)^2}{2\rho^2}} \Delta_M} \\
& = \sum_{m=1}^M \frac{I_0}{\rho^6} (y_m-y_0)^2 e^{-\frac{(x_m-x_0)^2+(y_m-y_0)^2}{2\rho^2}} \Delta_M \approx \frac{I_0 2\pi}{\rho^4} \sum_{m=1}^M \iint_{A_m} \frac{1}{2\pi\rho^2} (y-y_0)^2 e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy \\
& = \frac{I_0 2\pi}{\rho^4} \iint_{\mathcal{A}} \frac{1}{2\pi\rho^2} (y-y_0)^2 e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy \approx \frac{I_0 2\pi}{\rho^4} \rho^2 = \frac{I_0 2\pi}{\rho^2}, \quad (19)
\end{aligned}$$

where, in the last approximation of (19) we have used the fact that

$\iint_{\mathcal{A}} \frac{1}{2\pi\rho^2} (y-y_0)^2 e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy \approx \iint_{-\infty}^{\infty} \frac{1}{2\pi\rho^2} (y-y_0)^2 e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy = \rho^2$  since  $\rho \ll \ell(|\mathcal{A}|)$ . The positive term in the denominator (see (65)) can be simplified in a similar fashion. The square root of the term with minus sign can be simplified as

$$\sum_{m=1}^M \frac{1}{\Lambda_m} \iint_{A_m} \frac{I_0}{\rho^4} (y-y_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy \iint_{A_m} \frac{I_0}{\rho^4} (x-x_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy$$

$$\begin{aligned}
&\approx \sum_{m=1}^M \frac{\frac{I_0}{\rho^4}(y_m - y_0)e^{-\frac{(x_m - x_0)^2 + (y_m - y_0)^2}{2\rho^2}} \Delta_M}{\frac{I_0}{\rho^2}e^{-\frac{(x_m - x_0)^2 + (y_m - y_0)^2}{2\rho^2}} \Delta_M} \times \frac{I_0}{\rho^4}(x_m - x_0)e^{-\frac{(x_m - x_0)^2 + (y_m - y_0)^2}{2\rho^2}} \Delta_M \\
&\approx \frac{I_0 2\pi}{\rho^4} \sum_{m=1}^M \iint_{A_m} \frac{1}{2\pi\rho^2} (y - y_0)(x - x_0) e^{-\frac{(x - x_0)^2 + (y - y_0)^2}{2\rho^2}} dx dy \\
&= \frac{I_0 2\pi}{\rho^4} \iint_{\mathcal{A}} \frac{1}{2\pi\rho^2} (y - y_0)(x - x_0) e^{-\frac{(x - x_0)^2 + (y - y_0)^2}{2\rho^2}} dx dy \approx 0,
\end{aligned} \tag{20}$$

where in the last approximation of (20), we have used the fact that  $\iint_{\mathcal{A}} \frac{1}{2\pi\rho^2} (y - y_0)(x - x_0) e^{-\frac{(x - x_0)^2 + (y - y_0)^2}{2\rho^2}} dx dy \approx \iint_{-\infty}^{\infty} \frac{1}{2\pi\rho^2} (y - y_0)(x - x_0) e^{-\frac{(x - x_0)^2 + (y - y_0)^2}{2\rho^2}} dx dy = \mathbb{E}[X - \mathbb{E}[X]]\mathbb{E}[Y - \mathbb{E}[Y]] = 0$  where  $X$  and  $Y$  are independent Gaussian random variables with the same variance  $\rho^2$ , but with different means:  $\mathbb{E}[X] = x_0, \mathbb{E}[Y] = y_0$ . Therefore,

$$\text{Var}[\hat{x}_0] \geq \frac{\frac{I_0 2\pi}{\rho^2}}{\frac{I_0 2\pi}{\rho^2} \times \frac{I_0 2\pi}{\rho^2}} = \frac{\rho^2}{I_0 2\pi}. \tag{21}$$

We note that the Cramér-Rao lower bound is minimized by minimizing  $\rho$  (a more focused beam) for fixed signal power. The Cramér-Rao lower bounds goes to zero at the rate  $\mathcal{O}(\rho^2)$  as  $\rho \rightarrow 0$ , where  $\mathcal{O}$  represent the “big O” notation. Moreover, the Cramér-Rao lower bounds goes to zero in terms of  $I_0$  at the rate  $\mathcal{O}(I_0^{-1})$ .

## 2) Estimation of $x_0$ (Low Signal-to-Noise Ratio)

In this case, let us assume that  $\lambda_n A \gg \iint_{A_m} \frac{I_0}{\rho^2} e^{-\frac{(x - x_0)^2 + (y - y_0)^2}{2\rho^2}} dx dy$ . Then,  $\Lambda_m \approx \lambda_n A$ . In this case, the square root of the term with the minus sign in the denominator (the denominator is given by (65)) is

$$\frac{1}{\lambda_n A} \sum_{m=1}^M \iint_{A_m} \frac{I_0}{\rho^4} (y - y_0) e^{-\frac{(x - x_0)^2 + (y - y_0)^2}{2\rho^2}} dx dy \times \iint_{A_m} \frac{I_0}{\rho^4} (x - x_0) e^{-\frac{(x - x_0)^2 + (y - y_0)^2}{2\rho^2}} dx dy, \tag{22}$$

which is zero due to the symmetric nature of the Gaussian beam. Therefore, by further simplification,

$$\text{Var}[\hat{x}_0] \geq \frac{\lambda_n \rho^8}{I_0^2} \times \frac{A}{\sum_{m=1}^M \left( \iint_{A_m} (x - x_0) e^{-\frac{(x - x_0)^2 + (y - y_0)^2}{2\rho^2}} dx dy \right)^2} \tag{23}$$

which goes to

$$\text{Var}[\hat{x}_0] \geq \frac{\frac{\lambda_n \rho^8}{I_0^2}}{\iint_{\mathcal{A}} (x - x_0)^2 e^{-\frac{(x - x_0)^2 + (y - y_0)^2}{\rho^2}} dx dy} = \frac{2\rho^4}{\pi \left( \frac{I_0^2}{\lambda_n} \right)} \tag{24}$$

as  $M \rightarrow \infty$ . In this case, the Cramér-Rao lower bound goes to zero at a rate  $\mathcal{O}(\rho^4)$  as  $\rho \rightarrow 0$ . This is a faster rate of convergence than  $\mathcal{O}(\rho^2)$  for the high signal-to-noise ratio case. Additionally, Cramér-Rao lower bound converges to zero at the rate  $\mathcal{O}(I_0^{-2})$  in the low signal-to-noise ratio regime.

## V. CRAMÉR-RAO LOWER BOUNDS FOR BEAM PARAMETER ESTIMATION BASED ON DATA SYMBOLS

Since pilot symbols incur a loss in energy and bandwidth, there is a motivation to use data symbols for the estimation of beam parameters even though the use of data symbols may result in some loss in estimation performance. In this section, we derive the Cramér-Rao lower bounds of beam parameters based on PPM data symbols. As discussed in Section II-B, we can either use either one PPM symbol period or a PPM slot (slot containing the pulse) period as our observation interval. We first look at the Cramér-Rao lower bounds related to the symbol period based observation in the next section.

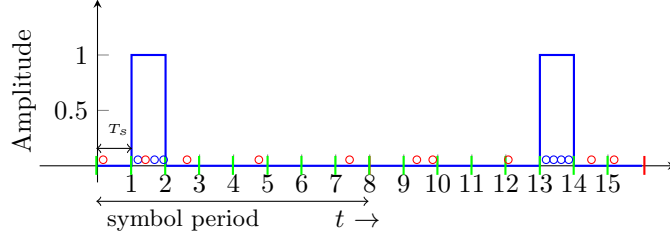


Fig. 2: This figure shows an 8-PPM scheme. The blue circles indicate signal photons and the red circles indicate noise photons. The observation based on the (signal) slot period contains a higher signal-to-noise ratio (eight times higher) as compared to the observation based on one symbol period.

#### A. Observations Based on Symbol Period

In this case, the noise power goes up  $K$  times where  $K$  is the number of slots in PPM. Thus, the new  $\lambda'_n \triangleq K\lambda_n$ , and  $Z_m \sim \text{Poisson}(\Lambda'_m)$ , where

$$\Lambda'_m \triangleq \iint_{A_m} \left( \frac{I_0}{\rho^2} \exp \left( -\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho^2} \right) + \lambda'_n \right) dx dy. \quad (25)$$

Thus, in this case,

$$p(\mathbf{Z}|\boldsymbol{\theta}) \triangleq \prod_{m=1}^M \frac{e^{-\Lambda'_m} (\Lambda'_m)^{Z_m}}{Z_m!}. \quad (26)$$

#### B. Observations Based on Slot Period: A Decision-Aided Scheme

The motivation behind choosing the slot period is to maximize the signal-to-noise ratio in the sufficient statistic. If the slot containing the signal is chosen, the resulting signal-to-noise ratio is  $K$  times bigger than the signal-to-noise ratio available in a symbol period. However, for the slot period case, we depend on the correct decision of the receiver to choose the “right” slot that contains the signal. If the receiver makes a mistake, we end up choosing a “noise-only” slot and the resulting noise photons do not give us any information about the beam parameters. Therefore, if the receiver starts making too many errors, the estimation performance will take a significant hit. Thus, in the slot period case, the correct symbol decision is the key to a good estimation performance, and we term the estimation based on slot period alternatively as *decision-aided estimation* of beam parameters.

Fig. 4 shows the block diagram of the decision-aided beam position estimation scheme in which the output of the *equal gain combiner* is fed into the beam position estimation block. If the equal gain combiner declares some symbol  $j$  as the transmitted ( $K$ -PPM) symbol for  $0 \leq j < K$ , the beam position estimation block chooses the  $j$ th slot as its observation interval.

For observation based on one slot,  $Z_m \sim \text{Poisson}(\Lambda_m)$ , with probability  $P_c$  and  $Z_m \sim \text{Poisson}(\lambda_n A)$  with probability  $(1 - P_c)$ . Thus,  $Z_m$  is a *doubly stochastic Poisson process* or a *Cox process* whose intensity varies randomly according to the Bernoulli distribution as follows:

$$p(\xi) = P_c \delta(\xi - \Lambda_m) + (1 - P_c) \delta(\xi - \lambda_n A), \quad (27)$$

where  $\delta(\cdot)$  is the *Dirac delta function*. Therefore, the likelihood function becomes

$$p(\mathbf{Z}|\boldsymbol{\theta}) \triangleq P_c \prod_{m=1}^M \frac{e^{-\Lambda_m} (\Lambda_m)^{Z_m}}{Z_m!} + (1 - P_c) \prod_{m=1}^M \frac{e^{-\lambda_n A} (\lambda_n A)^{Z_m}}{Z_m!}. \quad (28)$$

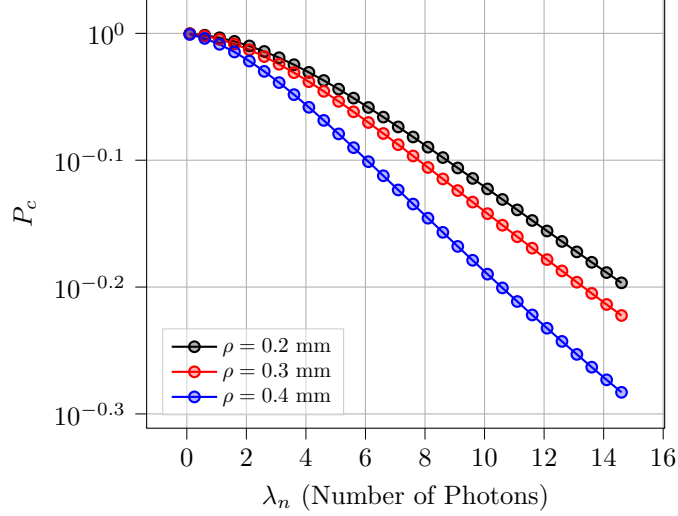


Fig. 3: This figure shows the probability of correct decision of the equal gain combiner receiver as a function of noise parameter  $\lambda_n$  for different values of beam radius  $\rho$  for the  $4 \times 4$  detector array. The modulation scheme considered in this case is 8-PPM. The received signal intensity is measured in terms of 10 signal photons received on average, and the area of the array  $|\mathcal{A}| = 4 \text{ mm}^2$ . The value of  $(x_0, y_0) = (0.4, 0.4)$ .

The quantity  $P_c$  is the probability of a correct decision of the equal gain combiner. It can be shown that for a *maximum a posteriori probability* detector that operates on a  $K$ -PPM symbol, we have that

$$P_c = (P(\{Z_s > Z_n\}))^{K-1} = (P(\{(Z_s - Z_n) > 0\}))^{K-1}. \quad (29)$$

In (29),  $Z_s \sim \text{Poisson}(\Lambda_s)$  and  $Z_n \sim \text{Poisson}(\lambda_n |\mathcal{A}|)$ . The random variable  $Z \triangleq Z_s - Z_n$  is a (discrete) *Skellam* random variable whose distribution is

$$P(\{Z = z\}) = e^{-(\Lambda_s + \lambda_n |\mathcal{A}|)} \left( \frac{\Lambda_s}{\lambda_n |\mathcal{A}|} \right)^{z/2} I_z \left( 2\sqrt{\Lambda_s \lambda_n |\mathcal{A}|} \right), \quad (30)$$

where  $I_z(\cdot)$  is the *modified Bessel function* of the first kind (not to be confused with peak intensity  $I_0$ ). Thus,

$$P(\{Z > 0\}) = \sum_{z=1}^{\infty} P(\{Z = z\}) = \sum_{z=1}^{\infty} e^{-(\Lambda_s + \lambda_n |\mathcal{A}|)} \left( \frac{\Lambda_s}{\lambda_n |\mathcal{A}|} \right)^{z/2} I_z \left( 2\sqrt{\Lambda_s \lambda_n |\mathcal{A}|} \right).$$

Fig. 3 shows the probability of correct decision  $P_c$  for different values of beam radius  $\rho$ . A large beam radius results in some loss of energy since some of the beam energy falls off the edge of the array. This leads to a lower probability of correct decision for larger beam radii. Finally, since  $P_e = 1 - P_c$ , we have that

$$P_e = 1 - \left( \sum_{z=1}^{\infty} e^{-(\Lambda_s + \lambda_n |\mathcal{A}|)} \left( \frac{\Lambda_s}{\lambda_n |\mathcal{A}|} \right)^{z/2} I_z \left( 2\sqrt{\Lambda_s \lambda_n |\mathcal{A}|} \right) \right)^{K-1}. \quad (31)$$

#### 1) Monte Carlo Expectation

It is not straightforward to compute the probabilistic expectations  $E \left[ \frac{\partial^2 \ln p(\mathbf{Z}|\boldsymbol{\theta})}{\partial I_0^2} \right]$ ,  $E \left[ \frac{\partial^2 \ln p(\mathbf{Z}|\boldsymbol{\theta})}{\partial I_0 \partial x_0} \right]$ ,  $E \left[ \frac{\partial^2 \ln p(\mathbf{Z}|\boldsymbol{\theta})}{\partial I_0 \partial y_0} \right]$ ,  $E \left[ \frac{\partial^2 \ln p(\mathbf{Z}|\boldsymbol{\theta})}{\partial x_0^2} \right]$ ,  $E \left[ \frac{\partial^2 \ln p(\mathbf{Z}|\boldsymbol{\theta})}{\partial y_0^2} \right]$  and  $E \left[ \frac{\partial^2 \ln p(\mathbf{Z}|\boldsymbol{\theta})}{\partial x_0 \partial y_0} \right]$  for the likelihood function in (28). Thus, we resort to the

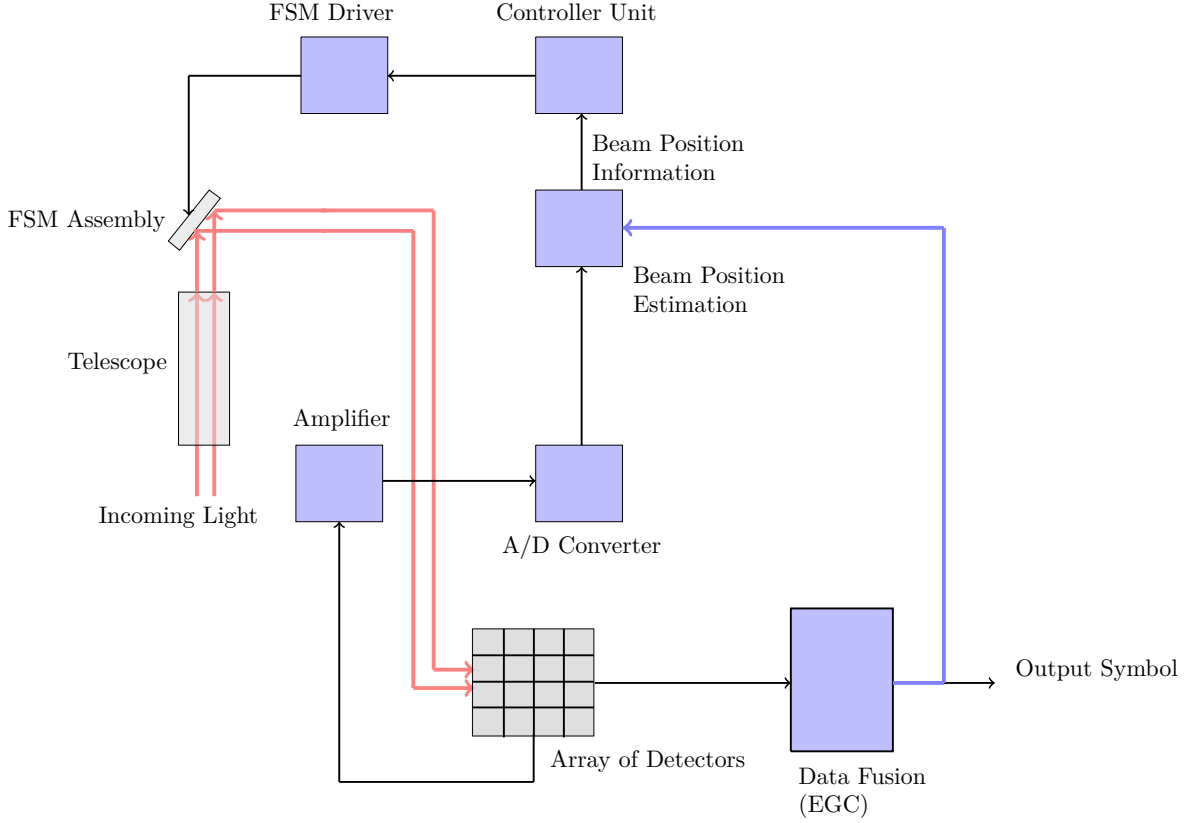


Fig. 4: This figure shows the block diagram of the decision-aided beam position estimation system. The output of the equal gain combiner is fed into the beam position estimation block so that the “right” slot is chosen for estimation purpose.

Monte Carlo simulations to compute these expectations. The simulations are carried out as follows:

- 1) Sample 1 with probability  $P_c$  and 0 with probability  $1 - P_c$ .
- 2) If 1, then sample  $Z_1$  from  $Poisson(\Lambda_1)$ ,  $Z_2$  from  $Poisson(\Lambda_2)$ ,  $\dots$ ,  $Z_M$  from  $Poisson(\Lambda_M)$ . Else, sample sample  $Z_1$  from  $Poisson(\lambda_n A)$ ,  $Z_2$  from  $Poisson(\lambda_n A)$ ,  $\dots$ ,  $Z_M$  from  $Poisson(\lambda_n A)$ .
- 3) Substitute the  $Z_m$ 's obtained from Step 2 into each of the second order partial derivatives:  $\frac{\partial^2 \ln p(\mathbf{Z}|\boldsymbol{\theta})}{\partial I_0^2}$ ,  $\frac{\partial^2 \ln p(\mathbf{Z}|\boldsymbol{\theta})}{\partial I_0 \partial x_0}$ ,  $\frac{\partial^2 \ln p(\mathbf{Z}|\boldsymbol{\theta})}{\partial I_0 \partial y_0}$ ,  $\frac{\partial^2 \ln p(\mathbf{Z}|\boldsymbol{\theta})}{\partial x_0^2}$ ,  $\frac{\partial^2 \ln p(\mathbf{Z}|\boldsymbol{\theta})}{\partial y_0^2}$  and  $\frac{\partial^2 \ln p(\mathbf{Z}|\boldsymbol{\theta})}{\partial x_0 \partial y_0}$ , and store the resulting values.
- 4) Repeat Step 1, Step 2 and Step 3 until the required number of simulations is reached.
- 5) Compute the sample mean of the values obtained in Step 3.

## VI. TWO-STEP ESTIMATION OF BEAM PARAMETERS

In this section, we will look at a two-step estimation algorithm that is used for estimating the beam parameters. The two-step estimation algorithm is defined as follows:

- 1) In the first step, the peak intensity  $I_0$  and background radiation  $\lambda_n$  are estimated using a method of moments estimator.
- 2) The estimates  $\hat{I}_0$  and  $\hat{\lambda}_n$  obtained from Step 1 are substituted into the loglikelihood function  $\ln p(\mathbf{Z}|\boldsymbol{\theta})$  and the estimate of  $(x_0, y_0)$  is obtained by maximizing the loglikelihood function (maximum likelihood estimation).

Alternatively, all the four parameters  $(x_0, y_0, I_0, \lambda_n)$  can be estimated via the maximum likelihood estimator. However, since no closed-form expressions for the maximum likelihood estimator are available, we have to resort to numerical optimization techniques (such as a genetic algorithm) in order to find the peak of

the loglikelihood function. This incurs a much higher computational complexity if all the four parameters are estimated with the maximum likelihood estimator. The two-step estimation algorithm reduces the complexity since two of the four parameters ( $I_0$  and  $\lambda_n$ ) can be estimated via the computationally efficient method of moments estimator without any knowledge of  $(x_0, y_0)$ , and the numerical search for the maximum of loglikelihood function is limited to just two dimensions in order to find  $(\hat{x}_0, \hat{y}_0)$ .

#### A. Method of Moments Estimator of $I_0$ and $\lambda_n$

##### 1) Pilot Symbol Case

The method of moments estimator of  $I_0$  for the pilot symbol case is

$$\hat{I}_0 = \frac{1}{2\pi N} \left( \sum_{i=1}^N \sum_{m=1}^M Z_{i,m}^{(s)} \right) - \frac{\hat{\lambda}_n |\mathcal{A}|}{2\pi}, \quad (32)$$

where  $Z_{i,m}^{(s)}$  is a Poisson random variable with mean  $\Lambda_m$ . The method of moments estimator of  $\lambda_n$  is

$$\hat{\lambda}_n = \frac{1}{|\mathcal{A}|N} \sum_{i=1}^N \sum_{m=1}^M Z_{i,m}^{(n)} \quad (33)$$

where  $Z_{i,m}^{(n)}$  is a Poisson random variable with mean  $\lambda_n A$ . It can be easily shown that  $\mathbb{E}[\hat{\lambda}_n] = \lambda_n$ , and

$$\mathbb{E}[\hat{I}_0] = \frac{1}{2\pi N} \left( \sum_{i=1}^N \sum_{m=1}^M \mathbb{E}[Z_{i,m}^{(s)}] \right) - \mathbb{E}[\hat{\lambda}_n] \frac{|\mathcal{A}|}{2\pi} \quad (34)$$

$$= \frac{1}{2\pi} \sum_{m=1}^M \iint_{A_m} \frac{I_0}{\rho^2} e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho^2}} dx dy + \lambda_n A - \lambda_n \frac{|\mathcal{A}|}{2\pi} \quad (35)$$

$$= \frac{1}{2\pi} (I_0 2\pi + \lambda_n |\mathcal{A}|) - \frac{\lambda_n |\mathcal{A}|}{2\pi} = I_0. \quad (36)$$

Thus, both  $\hat{I}_0$  and  $\hat{\lambda}_n$  are unbiased estimators of  $I_0$  and  $\lambda_n$ , respectively.

It is straightforward to verify that

$$\mathbb{E}[(\hat{\lambda}_n - \lambda_n)^2] = \mathbb{E}[\hat{\lambda}_n^2] - \lambda_n^2 = \frac{\lambda_n}{NMA} = \frac{\lambda_n}{N|\mathcal{A}|}. \quad (37)$$

The mean-square error between  $\hat{I}_0$  and  $I_0$  is

$$\mathbb{E}[(\hat{I}_0 - I_0)^2] = \text{Var}(\hat{I}_0 - I_0) = \text{Var}(\hat{I}_0) \quad (38)$$

$$= \frac{1}{(2\pi N)^2} \sum_{i=1}^N \sum_{m=1}^M \text{Var}(Z_{i,m}^{(s)}) + \left( \frac{AM}{2\pi} \right)^2 \text{Var}(\hat{\lambda}_n) \quad (39)$$

$$= \frac{1}{(2\pi N)^2} \sum_{i=1}^N \sum_{m=1}^M (\lambda_m^{(s)} + \lambda_n A) + \left( \frac{AM}{2\pi} \right)^2 \frac{\lambda_n}{NMA} \quad (40)$$

$$= \frac{I_0}{2\pi N} + \frac{\lambda_n AM}{2\pi^2 N} = \frac{1}{2\pi N} \left( I_0 + \frac{\lambda_n AM}{\pi} \right). \quad (41)$$

##### 2) Decision-Aided Estimation (Observations Based on Slot Period)

In this case, the generation of photon counts are governed by a doubly stochastic Poisson process. Thus,

$$\mathbb{E}[\hat{I}_0] = \frac{1}{2\pi N} \sum_{i=1}^N \sum_{m=1}^M \mathbb{E}[Z_m] - \mathbb{E}[\hat{\lambda}_n] \frac{AM}{2\pi}, \quad (42)$$

where  $\mathbb{E}[Z_m|c] = \Lambda_m = \Lambda_m^{(s)} + \lambda_n A$  and  $\mathbb{E}[Z_m|e] = \lambda_n A$ . Therefore,

$$\mathbb{E}[Z_m] = \Lambda_m^{(s)} P_c + \lambda_n A. \quad (43)$$

Therefore,

$$\mathbb{E}[\hat{I}_0] = \frac{1}{2\pi N} \sum_{i=1}^N \sum_{m=1}^M (\Lambda_m^{(s)} P_c + \lambda_n A) - \frac{\lambda_n AM}{2\pi} \quad (44)$$

$$= P_c I_0. \quad (45)$$

Moreover,

$$\text{Var}[\hat{I}_0] = \frac{1}{(2\pi N)^2} \sum_{i=1}^N \sum_{m=1}^M \text{Var}[Z_m] + \text{Var}[\hat{\lambda}_n] \left(\frac{AM}{2\pi}\right)^2 \quad (46)$$

$$= \frac{1}{(2\pi N)^2} \sum_{i=1}^N \sum_{m=1}^M \Lambda_m^{(s)} (P_c + \lambda_n A) + \frac{\lambda_n AM}{N(2\pi)^2} \quad (47)$$

$$= \frac{1}{(2\pi N)^2} [P_c I_0 2\pi N + \lambda_n AMN] + \frac{\lambda_n AM}{N(2\pi)^2} = \frac{P_c I_0}{2\pi N} + \frac{\lambda_n AM}{2\pi^2 N} \quad (48)$$

$$= \frac{1}{2\pi N} \left( P_c I_0 + \frac{\lambda_n AM}{\pi} \right). \quad (49)$$

Finally, since  $\text{Var}(\hat{I}_0 - I_0) = \text{Var}(\hat{I}_0)$ , we have that

$$\mathbb{E}[(I_0 - \hat{I}_0)^2] = \text{Var}[I_0 - \hat{I}_0] + (\mathbb{E}[I_0 - \hat{I}_0])^2 = \text{Var}(\hat{I}_0) + (I_0 - I_0 P_c)^2 \quad (50)$$

$$= \frac{1}{2\pi N} \left( P_c I_0 + \frac{\lambda_n AM}{\pi} \right) + I_0^2 (1 - P_c)^2. \quad (51)$$

### B. Maximum Likelihood Estimation of $(x_0, y_0)$

For the pilot symbol scheme, the maximum likelihood estimator of beam position  $(x_0, y_0)$  on the array is given by [6]:

$$\begin{aligned} (\hat{x}_0, \hat{y}_0) &\triangleq \arg \max_{x_0, y_0} \ln p(Z_1, Z_2, \dots, Z_M | x_0, y_0) \\ &= \arg \max_{x_0, y_0} \sum_{m=1}^M Z_m \ln \left( \hat{I}_0 2\pi \left[ \Phi\left(\frac{y_{m2} - y_0}{\rho}\right) - \Phi\left(\frac{y_{m1} - y_0}{\rho}\right) \right] \left[ \Phi\left(\frac{x_{m2} - x_0}{\rho}\right) - \Phi\left(\frac{x_{m1} - x_0}{\rho}\right) \right] + \hat{\lambda}_n A \right) \\ &\quad - \left( \hat{I}_0 2\pi \left[ \Phi\left(\frac{\frac{\ell(\mathcal{A})}{2} - y_0}{\rho}\right) - \Phi\left(\frac{-\frac{\ell(\mathcal{A})}{2} - y_0}{\rho}\right) \right] \left[ \Phi\left(\frac{\frac{\ell(\mathcal{A})}{2} - x_0}{\rho}\right) - \Phi\left(\frac{-\frac{\ell(\mathcal{A})}{2} - x_0}{\rho}\right) \right] + \hat{\lambda}_n |\mathcal{A}| \right), \quad (52) \end{aligned}$$

where  $\Phi(\cdot)$  is the cumulative distribution function of a standard normal random variable, and  $\hat{I}_0$  and  $\hat{\lambda}_n$  are the method of moments estimates of  $I_0$  and  $\lambda_n$ , respectively. The quantity  $(x_{m2}, y_{m2})$  is the location of the north east corner of the  $m$ th detector, and  $(x_{m1}, y_{m1})$  is the position of south west corner.

For estimation based on data symbols, the maximum likelihood estimate is obtained by maximizing (26) (symbol period) or by maximizing (28) (slot period).

Regarding the maximization of the loglikelihood functions (3), (26) and (28), we have to resort to a genetic algorithm to look for the global maximum.

The mean-square error of the maximum likelihood estimator is computed via Monte Carlo simulations. The average of the squared errors is computed by repeating the experiment many times and then computing the sample average of the squared errors.

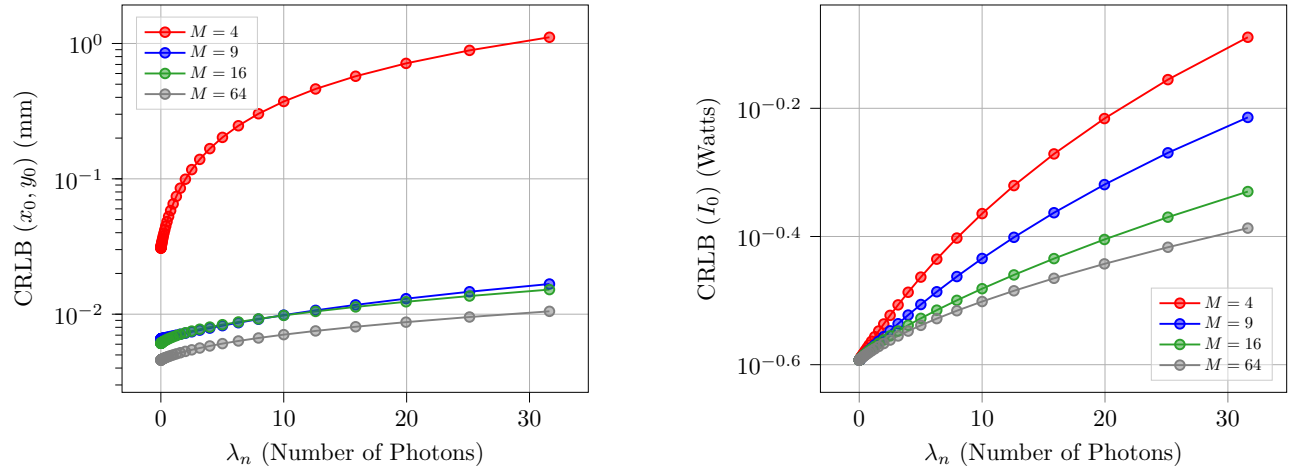


Fig. 5: This figure shows the Cramér-Rao lower bound on the variance of unbiased estimators of beam center position  $(x_0, y_0)$  (left) and the peak intensity  $I_0$  (right) as a function of noise parameter  $\lambda_n$  for one pilot symbol. These curves are plotted for different values of number of detectors  $M$  in the array. The value of  $I_0 = 1.6$  which gives an average signal photon rate of 10 photons/slot period. The value of  $(x_0, y_0) = (0.4, 0.4)$ .

## VII. SIMULATION RESULTS AND DISCUSSION

In this section, we interpret the simulations results that we have obtained in this study. In these simulations, we have considered the low photon rate regime. In this regard, we have considered—on average—10 signal photons for the entire array during the observation interval. These low photon rate channels are of interest in deep space communications where the received signal energy is so low that we are only able to detect a few signal photons per slot of a PPM symbol [1], [6]. Additionally, the low rate of photons has also to do with the “blocking” phenomenon of avalanche photodetectors that are operated in Geiger mode as photon counters. The blocking occurs because the detection of the first signal photon causes an avalanche of electrons, and this avalanche has to be *quenched* by an avalanche recovery circuit, and the bias has to be restored, before the next photon can be detected. Thus, the detector “sleeps” or gets “blocked” for a few microseconds before it is ready to detect the next incoming photon.

In the discussion that follows,  $\lambda_n$  is measured in terms of average number of noise photons that occur during an observation interval.

For all the experiments, the area of the detector array  $|\mathcal{A}| = 4 \text{ mm}^2$ . This detector array extends from -1 mm to 1 mm in each of the two dimensions, and the center of the array coincides with the origin. Additionally, we want to emphasize that the area of the array  $|\mathcal{A}|$  is fixed irrespective of the number of detector  $M$  in the array. Thus, a larger  $M$  implies a smaller area per detector.

In terms of notation, we want to point out that the expression  $\text{CRLB}(x_0, y_0)$  denotes the sum of individual Cramér-Rao lower bounds— $\text{CRLB}(x_0)$  and  $\text{CRLB}(y_0)$ . This is true since these two parameters can be treated independently of each other due to circularly symmetric nature of Gaussian beam.

Fig. 5 indicates the Cramér-Rao lower bound plots for  $(\hat{x}_0, \hat{y}_0, \hat{I}_0)$  as a three-parameter estimation problem as defined in (12), (13) and (14). These curves are plotted as a function of noise parameter  $\lambda_n$ . Fig. 6 depicts the Cramér-Rao lower bound curves as a function of number of pilot symbols used as observations in the estimation of parameters.

Fig. 7 shows the Cramér-Rao lower bound plots as a function of beam radius  $\rho$  at the point  $(x_0, y_0) = (0.4, 0.4)$ . We note that the Cramér-Rao lower bound of  $\hat{I}_0$  increases monotonically with  $\rho$ . However, for the CRLB of  $(\hat{x}_0, \hat{y}_0)$ , we see that there is an optimum value of  $\rho$  (lets call it  $\rho_M^*$  for the  $M$ -cell array)

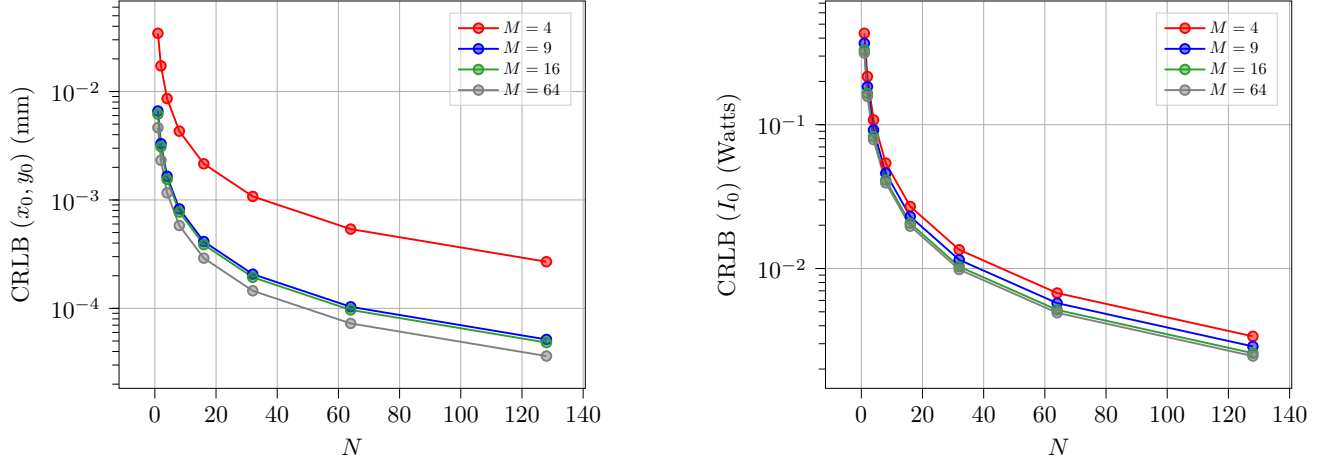


Fig. 6: This figure shows the Cramér-Rao lower bound on the variance of unbiased estimators of beam center position  $(x_0, y_0)$  (left) and the peak intensity  $I_0$  (right) as a function of the number of pilot symbols  $N$  needed in order to estimate  $(x_0, y_0)$ . The average signal photon rate is 10 photons/slot period. The average noise power for both cases is  $\lambda_n = 1$  photons/slot period for the figure on top, and  $\lambda_n = 10$  photons/slot period for the figure at the bottom.

at which the Cramér-Rao lower bound is minimized. Additionally,  $\rho_N^* < \rho_M^*$  for  $N > M$ . Intuitively, these observations are straightforward to explain. For a fixed signal-to-noise ratio, if the beam footprint is small, but at least covers one detector completely, then such a small beam footprint will minimize the mean-square error. This is true since all the power is focused into a small region on the array where the number of noise photons (on average) is relatively small, and this fact will help the estimator in order to estimate the beam position more accurately as opposed to a more “spread out” beam.

However, if the beam radius is much smaller than the dimensions of a single detector, then the beam will only give rise to photons in the detector in which it is located, and the neighboring detectors will not register any signal photons. Since we round off the locations of the photons—that occur inside a given detector—to the center of that detector, any movement of the “super thin” beam inside the given detector cannot be tracked. Therefore, the Cramér-Rao lower bound rises if  $\rho$  diminishes beyond a certain (optimum) value.

Fig. 8 and Fig. 9 show the effect of beam radius on the Cramér-Rao lower bounds of  $(\hat{x}_0, \hat{y}_0)$  and  $\hat{I}_0$  as a function of  $(x_0, y_0)$  for a  $4 \times 4$  detector array when pilot symbols are used for parameter estimation. In these figures, the beam center  $(x_0, y_0)$  varies along a diagonal on the array, i.e. from point  $(-0.8, -0.8)$  to point  $(0.8, 0.8)$ . For  $(\hat{x}_0, \hat{y}_0)$ , we note that for small  $\rho$ , the Cramér-Rao lower bound is highly sensitive to the location of the beam center  $(x_0, y_0)$ . For example, if we consider the case for  $\rho = 0.12$  mm, we note that the diameter of the beam  $2\rho$  is much smaller than the breadth of a single detector in this case (the breadth of a single detector for a  $4 \times 4$  array is  $\ell(A_m) = 0.5$  mm). Thus,  $2\rho \ll \ell(A_m)$ . First, we note that the points  $(-0.75, -0.75)$ ,  $(-0.25, -0.25)$ ,  $(0.25, 0.25)$ ,  $(0.75, 0.75)$  correspond to the centers of the detectors on the diagonal, and the points  $(-0.5, -0.5)$ ,  $(0, 0)$ ,  $(0.5, 0.5)$  corresponds to the edges of detectors. We additionally note that the Cramér-Rao lower bound attains its peak value at the centers of detectors and its minimum values on the edges. This pattern is explained by our earlier understanding (as argued during elaboration of Fig. 7) that when the beam is very thin and its center resides on the center of a particular detector, then all the energy of the beam resides in that particular detector, and a small movement of the beam cannot be detected by the array. However, if the beam center of such a thin beam lies on the edge of a detector, its slightest movement can be tracked by detecting the change in the energy difference of the two detectors that share the edge. For the case of  $M = 64$ ,  $\ell(A_m) = 0.25$  mm, and since

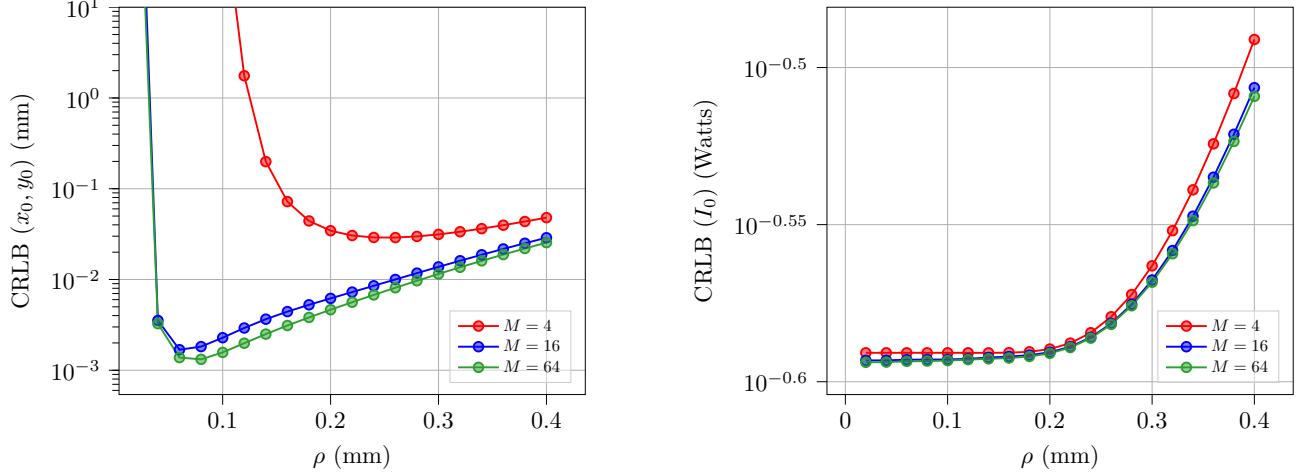


Fig. 7: This figure shows the Cramér-Rao lower bound on the variance of unbiased estimators of beam center position  $(x_0, y_0)$  (left) and the peak intensity  $I_0$  (right) as a function of the beam radius  $\rho$ . The value of  $x_0$  and  $y_0$  are both set to 0.4 mm on the array. The average signal photon rate is 10 photons/slot period, and the average number of noise photons  $\lambda_n = 1$  photon/slot period for both the figures.

$2\rho \approx \ell(A_m)$ , we observe that the fluctuation of the Cramér-Rao lower bound as a function of  $(x_0, y_0)$  is almost negligible.

When the beam radius is larger such that  $2\rho > \ell(A_m)$ , then the energy of the beam is not confined to a single detector regardless of where the beam center resides on the array. In this case, any movement of the beam will be registered because of change in the detected energy reported by the detectors. However, whether the beam diameter is large or small, once the beam center gets too close to the edge of the array, part of the beam energy will fall off the edge of the array and the detector array will experience a net loss in received signal energy. This leads to a higher Cramér-Rao lower bound at the edges of the detector array. Additionally, even though the fluctuation in Cramér-Rao lower bound is minimized for a higher beam radius, the minimum value of the Cramér-Rao lower bound suffers significantly (the minimum value of Cramér-Rao lower bound goes up) as compared to the smaller beam radius scenario. This is because with a larger beam radius, the signal energy is spread out to a larger number of detectors, which results in a smaller signal-to-noise ratio per detector. With a smaller beam radius, the beam energy is confined to a smaller number of detectors, and the average signal-to-noise ratio in these detectors remains significantly higher compared to the large beam radius case. This observation corroborates our previous assertion that the Cramér-Rao lower bound decays as  $\mathcal{O}(\rho^2)$  and  $\mathcal{O}(\rho^4)$  for the high and low signal-to-noise ratio, respectively, as  $\rho \rightarrow 0$  for the continuous array (see (21), and (24)).

Regarding the Cramér-Rao lower bound of  $\hat{I}_0$ , we observe that the Cramér-Rao lower bound is not very sensitive to the beam radius and the performance does not change significantly over the chosen range of beam radii. Additionally, the Cramér-Rao lower bound performance between the  $4 \times 4$  detector and the  $8 \times 8$  case is not significant.

Fig. 10 illustrates the Cramér-Rao lower bound comparisons for the observations based on pilot symbols versus data symbols. We note that—as expected—the Cramér-Rao lower bound based on pilot symbols outperforms the Cramér-Rao lower bound based on data symbols (both for slot period and symbol period). We also note that the Cramér-Rao lower bound based on slot period performs better than the Cramér-Rao lower bound related to symbol period for large SNR, and the gap between the performance of the two schemes closes as  $\lambda_n$  grows large. Fig. 11 show the performance comparison of the maximum likelihood estimator of  $(x_0, y_0)$  for the pilot and data symbol based observations, and we see a similar trend as the

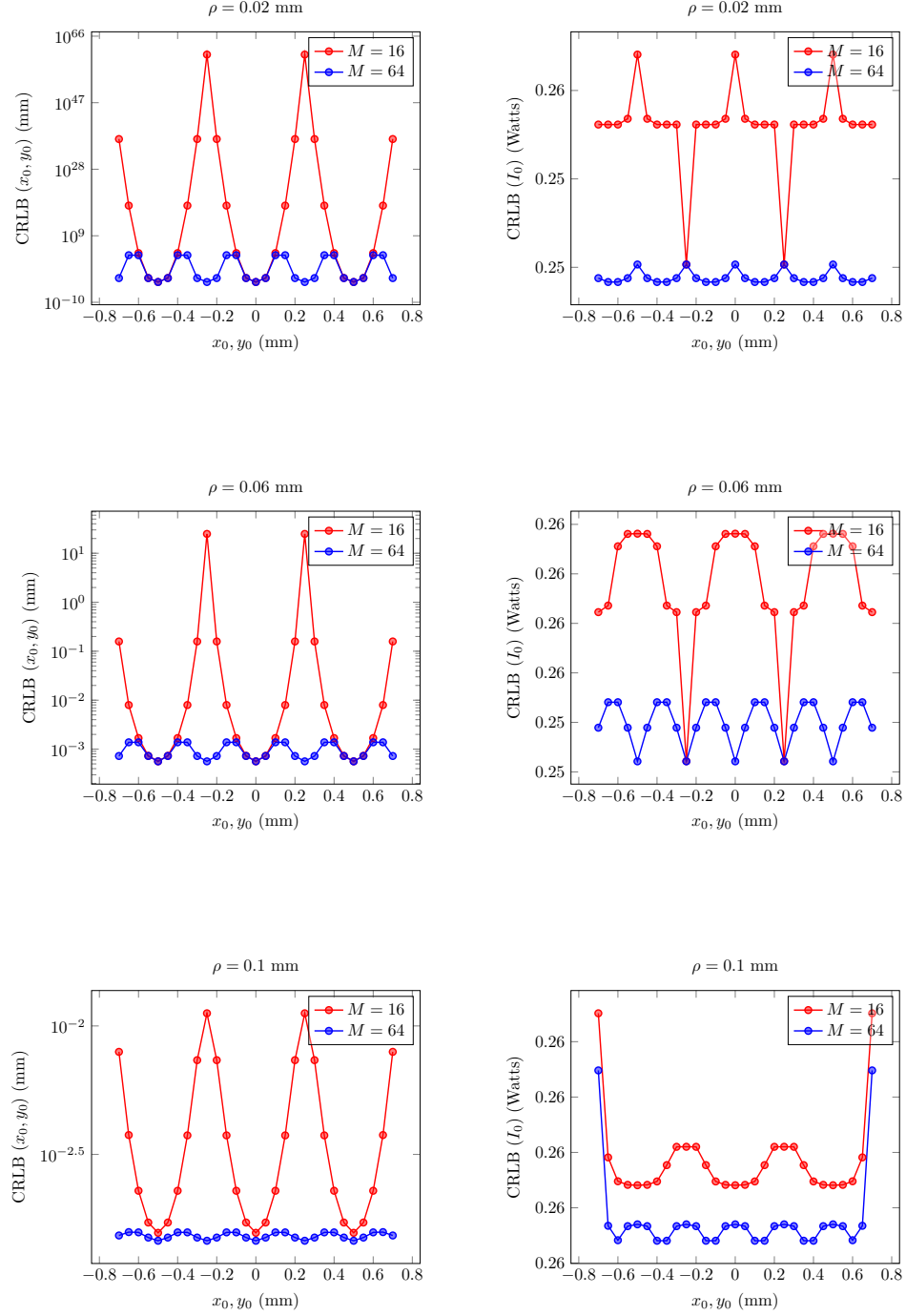


Fig. 8: This figure shows the effect of beam radius  $\rho$  on the Cramér-Rao lower bound on the variance of unbiased estimators of  $(x_0, y_0)$  and  $I_0$  for the  $4 \times 4$  and  $8 \times 8$  detector arrays. The value of noise parameter  $\lambda_n$  is 1 photon/slot period. The average signal photon rate is 10 photons/slot period.

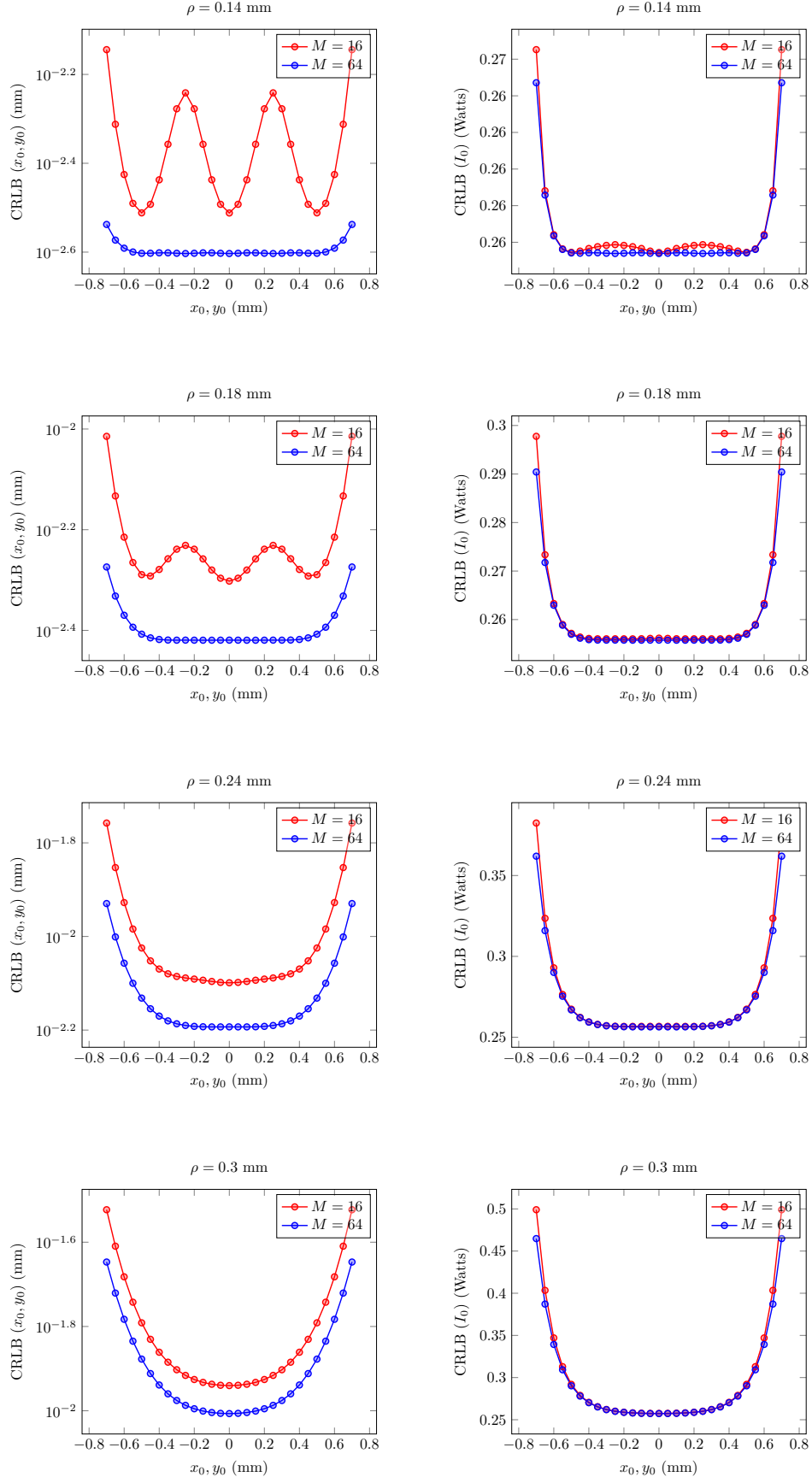


Fig. 9: This figure shows the effect of beam radius on the Cramér-Rao lower bound on the variance of unbiased estimators of  $(x_0, y_0)$  and  $I_0$  for the  $4 \times 4$  and  $8 \times 8$  detector arrays. The number of average received signal photons/symbol is 11. The value of noise parameter  $\lambda_n$  is 1 photon/slot period. The average signal photon rate is 10 photons/slot period.

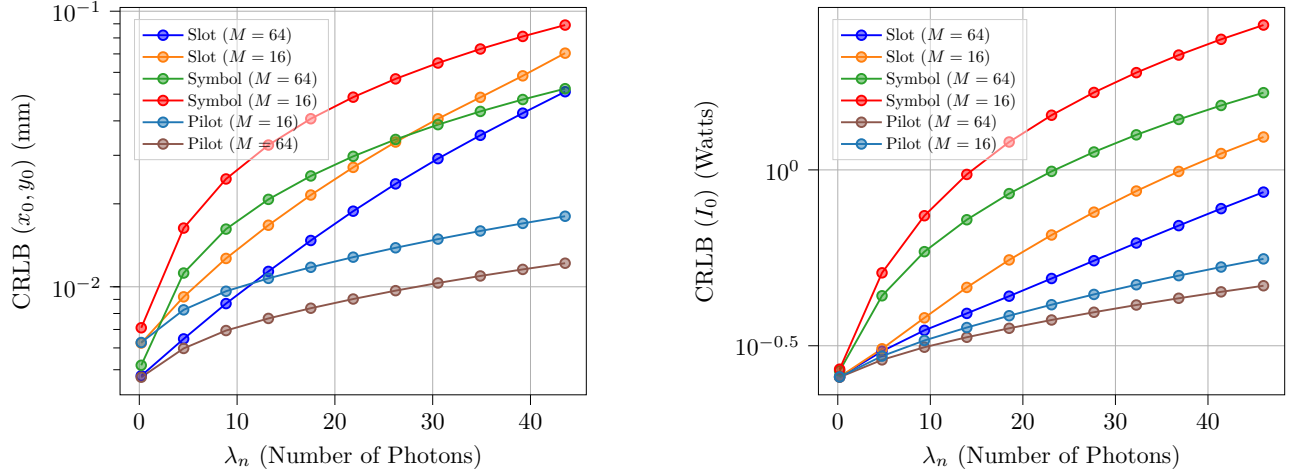


Fig. 10: This figure shows the Cramér-Rao lower bound on the variance of unbiased estimators of beam center position  $(x_0, y_0)$  (left) and the peak intensity  $I_0$  (right) as a function of noise parameter  $\lambda_n$  for three types of observations: pilot symbol, symbol period, and (signal) slot period. The average signal photon rate is 10 photons/slot period and the beam radius is 0.2 mm.

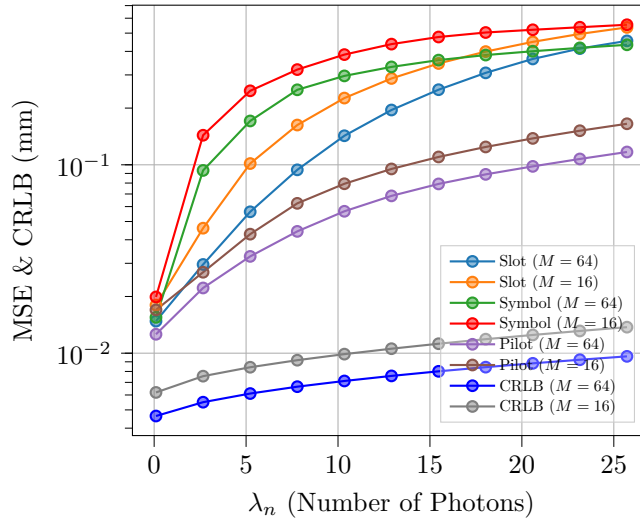


Fig. 11: This figure shows the performance of the the maximum likelihood estimator of  $(x_0, y_0)$  which is carried out for three different observations: pilot symbol, symbol period, and (signal) slot period. The average signal photon rate is 10 photons/slot period, and the beam radius is 0.2 mm. The performance of the maximum likelihood estimator is compared to the Cramér-Rao lower bound for the pilot symbol observations.

Cramér-Rao lower bound curves in Fig. 10.

Fig. 12 and Fig. 13 depict the performance of the method of moment estimator of  $\lambda_n$  and  $I_0$ , respectively. The performance of the method of moments estimator is compared with the Cramér-Rao lower bound as well. We note in Fig. 13 an interesting case where the mean-square error of the method of moments estimator based on slot period exceeds the mean-square error based on symbol period for large  $\lambda_n$ .

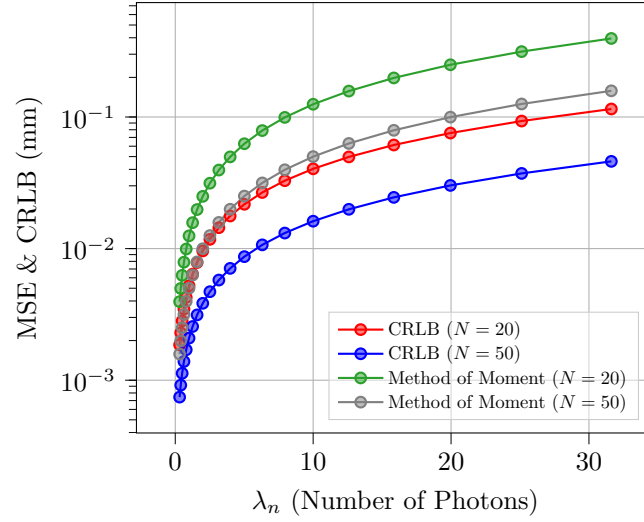


Fig. 12: This figure shows the performance of the method of moments estimator of the noise parameter  $\lambda_n$  for a  $4 \times 4$  array. The noise parameter is estimated by simply turning the transmitter off during regular intervals of time.

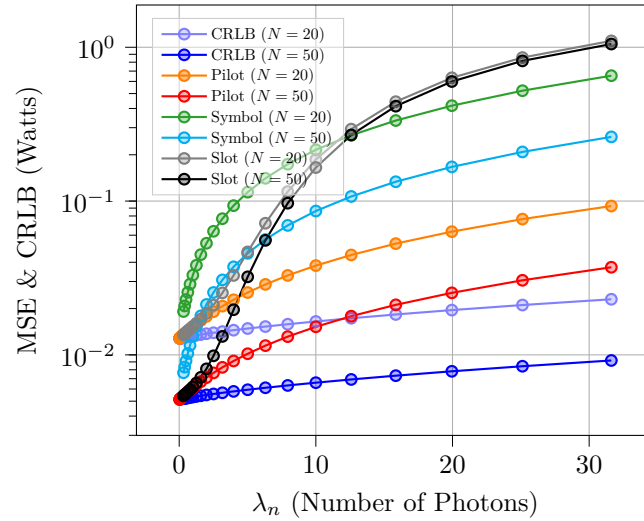


Fig. 13: This figure shows the performance of the method of moments estimator of  $I_0$  for the three different observations: pilot symbol, symbol period, and (signal) slot period. The average signal photon rate is 10 photons/slot period, and the beam radius is 0.2 mm. The performance of the method of moments estimator is compared with the Cramér-Rao lower bound of pilot symbol scheme as well. The number of detectors in the array is  $M = 16$  (a  $4 \times 4$  array).

### VIII. A BRIEF COMPLEXITY ANALYSIS OF ESTIMATORS

It is easy to see from (32) and (33), that the computational complexity of the method of moments estimator is  $\mathcal{O}(M)$  real additions if just one symbol is used for the purpose of estimation.

Regarding the maximum likelihood estimator, first we need  $\mathcal{O}(M)$  real additions and  $\mathcal{O}(M)$  real multiplies (see (52)) in order to compute the loglikelihood function. Thereafter, a *genetic algorithm* is employed in order to find the global maximum of the loglikelihood function. The complexity of the real number genetic algorithm is discussed in detail [6]. We note from [6] that the complexity of the genetic algorithm is a function of number of chromosomes,  $N_c$ , and the number of generations<sup>9</sup>,  $N_g$ . The values of  $N_c$  and  $N_g$  have to be chosen according to the nature of the objective function—a “spikier” function requires relatively large  $N_g$  and  $N_c$  for convergence to the true maximum. Additionally, the larger the number of dimensions of the objective function, the larger the values of  $N_c$  and  $N_g$  in order to speed up convergence. In our simulations, we set  $N_c = 50$ , and  $N_g = 400$ . Since the value of the loglikelihood function is computed for each chromosome, the total complexity of the maximum likelihood estimator is approximately  $N_c \times N_g \times \mathcal{O}(M)$  real multiplications and real additions<sup>10</sup>.

The interested reader is referred to [26] for more details on genetic algorithms.

### IX. CONCLUSION

In this paper, we have analyzed the Cramér-Rao lower bounds for the robust beam position estimation problem for a deep space optical communication system that uses an array of photon counting detectors at the receiver. In this regard, we derived the Cramér-Rao lower bounds for observations based on pilot symbols as well as data symbols (symbol period and slot period) of the pulse position modulation scheme. Even though the pilot symbols provide a superior performance in terms of estimation of beam parameters, we pay an extra price in terms of a larger bandwidth and a higher energy expenditure for transmission of dedicated pilot symbols. With data symbols, no extra bandwidth and energy expenditure is required. Additionally, for estimation of beam parameters with data symbols, we discovered that the performance of estimators based on slot period outperforms the estimators based on a symbol period, especially for the high signal-to-noise ratio case.

### APPENDIX

Taking the first partial derivative of the loglikelihood function, we have that

$$\begin{aligned} \frac{\partial \ln p(\mathbf{Z}|x_0, y_0)}{\partial x_0} &= \sum_{m=1}^M \frac{Z_m}{\Lambda_m} \iint_{A_m} \frac{I_0}{\rho^2} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} \frac{(x-x_0)}{\rho^2} dx dy - \iint_A \frac{I_0}{\rho^2} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} \frac{(x-x_0)}{\rho^2} dx dy \\ &\approx \sum_{m=1}^M \frac{Z_m}{\Lambda_m} \iint_{A_m} \frac{I_0}{\rho^4} (x-x_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy \end{aligned} \quad (53)$$

since  $\iint_A \frac{I_0}{\rho^2} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} \frac{(x-x_0)}{\rho^2} dx dy$  is approximately zero because it can be approximated as

$$\iint_A \frac{I_0}{\rho^2} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} \frac{(x-x_0)}{\rho^2} dx dy \approx \kappa_0 (\mathbb{E}[X] - x_0) \quad (54)$$

where  $\kappa_0$  is a constant, and  $\mathbb{E}[X] = x_0$ . This approximation is valid when  $\ell(\mathcal{A}) \gg \rho$  and the beam resides well within the boundaries of the array. Additionally,

$$\frac{\partial^2 \ln p(\mathbf{Z}|x_0, y_0)}{\partial x_0^2} \approx \sum_{m=1}^M -\frac{Z_m}{\Lambda_m^2} \left( \iint_{A_m} \frac{I_0}{\rho^4} (x-x_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy \right)^2$$

<sup>9</sup>The number of generations can be regarded as the number of iterations required in order to converge to the true maximum/minimum of the objective function.

<sup>10</sup>This does not include the complexity involved in comparing the fitness of the chromosomes during each iteration of the algorithm.

$$- \sum_{m=1}^M \frac{Z_m}{\Lambda_m} \iint_{A_m} \frac{I_0}{\rho^4} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy + \sum_{m=1}^M \frac{Z_m}{\Lambda_m} \iint_{A_m} \frac{I_0}{\rho^6} (x-x_0)^2 e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy. \quad (55)$$

Now, the expectation is taken with respect to  $Z_m$ :

$$\begin{aligned} -\mathbb{E} \left[ \frac{\partial^2 \ln p(\mathbf{Z}|x_0, y_0)}{\partial x_0^2} \right] &= \sum_{m=1}^M \frac{1}{\Lambda_m} \left( \iint_{A_m} \frac{I_0}{\rho^4} (x-x_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy \right)^2 \\ &+ \sum_{m=1}^M \iint_{A_m} \frac{I_0}{\rho^4} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy - \sum_{m=1}^M \iint_{A_m} \frac{I_0}{\rho^6} (x-x_0)^2 e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy \\ &= \sum_{m=1}^M \frac{1}{\Lambda_m} \left( \iint_{A_m} \frac{I_0}{\rho^4} (x-x_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy \right)^2 \\ &+ \iint_{\mathcal{A}} \frac{I_0}{\rho^4} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy - \iint_{\mathcal{A}} \frac{I_0}{\rho^6} (x-x_0)^2 e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy \\ &= \sum_{m=1}^M \frac{1}{\Lambda_m} \left( \iint_{A_m} \frac{I_0}{\rho^4} (x-x_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy \right)^2 + \frac{2\pi I_0}{\rho^2} - \frac{2\pi I_0}{\rho^2} \\ &= \sum_{m=1}^M \frac{1}{\Lambda_m} \left( \iint_{A_m} \frac{I_0}{\rho^4} (x-x_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy \right)^2. \end{aligned} \quad (56)$$

Similarly, it can be shown that

$$-\mathbb{E} \left[ \frac{\partial^2 \ln p(\mathbf{Z}|x_0, y_0)}{\partial y_0^2} \right] = \sum_{m=1}^M \frac{1}{\Lambda_m} \left( \iint_{A_m} \frac{I_0}{\rho^4} (y-y_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy \right)^2. \quad (57)$$

Furthermore,

$$\begin{aligned} \frac{\partial^2 \ln p(\mathbf{Z}|x_0, y_0)}{\partial x_0 \partial y_0} &= \sum_{m=1}^M -\frac{Z_m}{\Lambda_m^2} \iint_{A_m} \frac{I_0}{\rho^4} (y-y_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy \times \iint_{A_m} \frac{I_0}{\rho^4} (x-x_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy \\ &+ \sum_{m=1}^M \frac{Z_m}{\Lambda_m} \iint_{A_m} \frac{I_0}{\rho^6} (x-x_0)(y-y_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy \end{aligned} \quad (58)$$

$$\begin{aligned} -\mathbb{E} \left[ \frac{\partial^2 \ln p(\mathbf{Z}|x_0, y_0)}{\partial x_0 \partial y_0} \right] &= \sum_{m=1}^M \frac{1}{\Lambda_m} \iint_{A_m} \frac{I_0}{\rho^4} (y-y_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy \iint_{A_m} \frac{I_0}{\rho^4} (x-x_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy \\ &- \iint_{\mathcal{A}} \frac{I_0}{\rho^6} (x-x_0)(y-y_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy \end{aligned} \quad (59)$$

where,  $\iint_{\mathcal{A}} \frac{I_0}{\rho^6} (x-x_0)(y-y_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy \approx 0$  since it can be approximated as

$$\iint_{\mathcal{A}} \frac{I_0}{\rho^6} (x-x_0)(y-y_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy \approx \kappa_1 [(\mathbb{E}[X] - x_0)(\mathbb{E}[Y] - y_0)] \quad (60)$$

where  $\mathbb{E}[X] = x_0$  and  $\mathbb{E}[Y] = y_0$  for some constant  $\kappa_1$ . This approximation is valid if  $\ell(\mathcal{A}) \gg \rho$  and the beam resides well within the boundaries of the array. Therefore,

$$-\mathbb{E} \left[ \frac{\partial^2 \ln p(\mathbf{Z}|x_0, y_0)}{\partial x_0 \partial y_0} \right] = \sum_{m=1}^M \frac{1}{\Lambda_m} \iint_{A_m} \frac{I_0}{\rho^4} (y-y_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy \iint_{A_m} \frac{I_0}{\rho^4} (x-x_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy$$

$$= -\mathbb{E} \left[ \frac{\partial^2 \ln p(\mathbf{Z}|x_0, y_0)}{\partial y_0 \partial x_0} \right]. \quad (61)$$

Moreover, the *Fisher Information Matrix* is

$$\mathcal{I}(x_0, y_0) = \begin{bmatrix} -\mathbb{E} \left[ \frac{\partial^2 \ln p(\mathbf{Z}|x_0, y_0)}{\partial x_0^2} \right] & -\mathbb{E} \left[ \frac{\partial^2 \ln p(\mathbf{Z}|x_0, y_0)}{\partial x_0 \partial y_0} \right] \\ -\mathbb{E} \left[ \frac{\partial^2 \ln p(\mathbf{Z}|x_0, y_0)}{\partial y_0 \partial x_0} \right] & -\mathbb{E} \left[ \frac{\partial^2 \ln p(\mathbf{Z}|x_0, y_0)}{\partial y_0^2} \right] \end{bmatrix}, \quad (62)$$

and  $\text{Var}[\hat{x}_0] \geq [\mathcal{I}^{-1}(x_0, y_0)]_{1,1}$ , and  $\text{Var}[\hat{y}_0] \geq [\mathcal{I}^{-1}(x_0, y_0)]_{2,2}$ . Finally,

$$\text{Var}[\hat{x}_0] \geq \frac{\sum_{m=1}^M \frac{1}{\Lambda_m} \left( \iint_{A_m} \frac{I_0}{\rho^4} (y - y_0) e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho^2}} dx dy \right)^2}{\Psi(x_0, y_0, I_0, \rho)}, \quad (63)$$

and

$$\text{Var}[\hat{y}_0] \geq \frac{\sum_{m=1}^M \frac{1}{\Lambda_m} \left( \iint_{A_m} \frac{I_0}{\rho^4} (x - x_0) e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho^2}} dx dy \right)^2}{\Psi(x_0, y_0, I_0, \rho)}, \quad (64)$$

where

$$\begin{aligned} & \Psi(x_0, y_0, I_0, \rho) \\ & \triangleq \left[ \sum_{m=1}^M \frac{1}{\Lambda_m} \left( \iint_{A_m} \frac{I_0}{\rho^4} (x - x_0) e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho^2}} dx dy \right)^2 \times \sum_{m=1}^M \frac{1}{\Lambda_m} \left( \iint_{A_m} \frac{I_0}{\rho^4} (y - y_0) e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho^2}} dx dy \right)^2 \right. \\ & \left. - \left( \sum_{m=1}^M \frac{1}{\Lambda_m} \iint_{A_m} \frac{I_0}{\rho^4} (y - y_0) e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho^2}} dx dy \iint_{A_m} \frac{I_0}{\rho^4} (x - x_0) e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\rho^2}} dx dy \right)^2 \right]. \quad (65) \end{aligned}$$

## REFERENCES

- [1] H. Hemmati, A. Biswas, and I. B. Djordjevic, "Deep-space optical communications: Future perspectives and applications," *Proceedings of the IEEE*, vol. 99, no. 11, pp. 2020–2039, 2011.
- [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] R. M. Gagliardi and M. Sheikh, "Pointing error statistics in optical beam tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-16, no. 5, pp. 674–682, Sep. 1980.
- [4] G. Marola, D. Santerini, and G. Prati, "Stability analysis of direct-detection cooperative optical beam tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 25, no. 3, pp. 325–334, May 1989.
- [5] B. J. Slocumb and D. L. Snyder, "Maximum likelihood estimation applied to quantum-limited optical position-sensing," 1990. [Online]. Available: <https://doi.org/10.1117/12.2322208>
- [6] 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.
- [7] —, "Optical beam position tracking in free-space optical communication systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 20, no. 2, April 2018.
- [8] 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.
- [9] 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.
- [10] M. S. Bashir and M.-S. Alouini, "Signal acquisition with photon-counting detector arrays in free-space optical communications," *IEEE Transactions on Wireless Communications*, vol. 19, no. 4, pp. 2181–2195, 2020.
- [11] M. S. Bashir and S. S. Muhammad, "Time synchronization in photon-limited deep space optical communications," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 1, pp. 30–40, 2020.
- [12] 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.
- [13] 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.

- [14] 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.
- [15] 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.
- [16] 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.
- [17] C. B. Issaid, K. Park, and M. -S. Alouini, "A generic simulation approach for the fast and accurate estimation of the outage probability of single hop and multihop FSO links subject to generalized pointing errors," *IEEE Transactions on Wireless Communications*, vol. 16, no. 10, pp. 6822–6837, October 2017.
- [18] 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.
- [19] 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.
- [20] 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>
- [21] E. Alerstam, K. Andrews, M. Srinivasan, and A. Wong, "The effect of photon counting detector blocking on centroiding for deep space optical communications," in *Free-Space Laser Communication and Atmospheric Propagation XXX*, H. Hemmati and D. M. Boroson, Eds., vol. 10524, International Society for Optics and Photonics. SPIE, 2018, pp. 47 – 59. [Online]. Available: <https://doi.org/10.1117/12.2296740>
- [22] "APD arrays: Geiger-mode APD arrays detect low light," <https://www.laserfocusworld.com/detectors-imaging/article/16555170/apd-arrays-geigermode-apd-arrays-detect-low-light>, accessed: 06/04/2020.
- [23] M. S. Bashir and M. -S. Alouini, "Free-space optical MISO communications with an array of detectors," *IEEE Open Journal of the Communications Society*, July 2020, submitted for publication (available on arXiv at <https://arxiv.org/pdf/2009.00380.pdf>).
- [24] Y. S. Shmaliy, S. Zhao, and C. K. Ahn, "Unbiased finite impulse response filtering: An iterative alternative to Kalman filtering ignoring noise and initial conditions," *IEEE Control Systems Magazine*, vol. 37, no. 5, October 2017.
- [25] D. L. Snyder and M. I. Miller, *Random Point Processes in Time and Space*. New York, NY: Springer-Verlag, 1991.
- [26] S. S. Rao, *Engineering Optimization Theory and Practice*. Hoboken, NJ: John Wiley & Sons, Inc., 2009.