Received 24 March 2021; revised 22 April 2021; accepted 22 April 2021. Date of publication 10 May 2021; date of current version 18 May 2021.

Digital Object Identifier 10.1109/OJCOMS.2021.3078091

Cramér–Rao Bounds for Beam Tracking With Photon Counting Detector Arrays in Free-Space Optical Communications

MUHAMMAD SALMAN BASHIR (Senior Member, IEEE), MING-CHENG TSAI (Graduate Student Member, IEEE), AND MOHAMED-SLIM ALOUINI (Fellow, IEEE)

Computer Electrical and Mathematical Sciences and Engineering Division, Department of Electrical Engineering, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia

CORRESPONDING AUTHOR: M.-C. TSAI (e-mail: mingcheng.tsai@kaust.edu.sa)

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

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 unbiased estimators of beam position. Furthermore, we have also derived the Cramér-Rao lower bounds of other system parameters such as signal peak intensity, and dark current noise power, on the array. In this sense, we have considered 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 and noise 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 signal peak and dark current noise intensities 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, Cramèr-Rao lower bound, dark current, maximum likelihood estimator, method of moments estimator, photon-counting detector arrays, pulse position modulation.

I. INTRODUCTION

FREE-SPACE optical (FSO) communications is an important technology that will help us support high data rates between satellites in deep-space communication systems [1]. However, the problem of pointing, acquisition and tracking is significant in FSO because of the narrow beam widths associated with the optical signal. Acquisition is the process whereby two terminals acquire each other's initial locations before the actual data communication begins. However, after the acquisition is achieved, the system still needs to maintain alignment between the transmitter and receiver assemblies. This requirement to maintain alignment arises due to the presence of physical factors such as atmospheric turbulence, or the mechanical vibrations introduced in transmitter/receiver assemblies by factors such as wind. In case there is a misalignment, the link would experience an outage due to a significant loss of received signal energy at the receiver. Similarly, the narrow optical beam needs to be tracked/aligned for mobile platforms in order to prevent outage at the receiver.

In this paper, we have considered the optical beam position estimation on a photon-counting detector array as part of "fine beam tracking" component in a deep space optical communication 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]. For these arrays, any fluctuations 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 this misalignment may be corrected by adjusting the receiver telescope, or/and with secondary optics (fast steering mirror (FSM) assemblies), in the receiver station.

In the current 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 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 that are 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 as compared to pilot symbols as the noise power becomes larger.

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 beam position estimation/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 optical beam tracking. In this study, we have chosen to cover the literature review from a signal processing perspective since such a perspective is more relevant to the analysis in this paper. In the following paragraph, we cover literature review on communications and beam tracking with an array of photon counting detectors in free-space optical communications.

The authors in [3] present the performance analysis of centroid and maximum likelihood estimators of beam position for a "continuous"[1] array. Building up on the work in [3], the authors in [4] propose beam position estimation algorithms for a "discrete" array of detectors since discrete arrays are the ones that are actually used in a practical setting. The work in [5] extends the work in [4] by recasting the random and time-varying beam position as a state space variable. They introduced Bayesian filtering algorithms—such as Kalman and particle filters—for tracking the beam position. Following up on the work [4] and [5], the authors in [6] 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 [7] presents a mathematical argument to show that for a fixed beam radius and signal power, the probability of error in symbol detection decreases monotonically as the number of detectors in the array is increased. Furthermore, the authors in [8] analyze the acquisition performance of an FSO system that employs an array of detectors at the receiver, and they demonstrated that the acquisition performance in terms of mean acquisition time can be significantly enhanced by employing a large number of detectors in the receiver array. In the same vein, the study in [9] considers time synchronization schemes based on an array of detectors, and this study showed that the synchronization performance improves significantly by using an array of detectors at the receiver.

For the sake of completeness, 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, the seminal work [10] develops the pointing 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 [11] investigate a slightly different optimization problem concerning pointing: They consider the maximization of link availability as a function of beam radius (for fixed signal power). In addition to these papers, the interested reader may be directed to [12]–[15] 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 [16]–[20].

B. 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 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 track the movement of the beam and align it—perhaps through a FSM assembly—to the center of the array.

Additionally, we have made two more assumptions in order to simplify the resulting analysis. First, we have assumed that the photon counting detectors (operated in the Geiger-mode) do not suffer from dead-time or blocking, and each incoming signal or noise photon is reported by the detectors. Secondly, we have assumed that the effect

1. A (hypothetical) 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.

of background radiation is minimal,[2] and the only source of noise is the dark current. This assumption is made in order to simplify the argument since the dark current noise power is assumed to be uniformly distributed over the array, but a more complex model is required in order to mimic the effect of background radiation (which is nonuniformly distributed on the array). In a future study, we would analyze the effect of both the dead-time and background radiation on the Crameŕ-Rao lower bounds of beam and noise parameter estimation.

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. The square shape is chosen in order to yield a more tractable mathematical analysis.

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 the prior model and real world parameters[3] [21].

C. 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[4] in the array of detectors. Section IV-A motivates the study of Cramér-Rao bounds of beam position, and Section IV-B argues the effect of the Cramér-Rao bound on the outage probability of the system. Section V 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 VI, we derive the Cramér-Rao bounds for observations based on data symbols. Section VII considers the two-step estimation (method of moments and maximum likelihood estimators) algorithm to estimate beam parameters. The simulation results are explained in Section VIII, and Section IX briefly discusses the complexity of the two estimators. The conclusions of this study are summarized in Section X.

III. SYSTEM MODEL

Photon-counting array receivers are typically used for long link distances [1] in deep space optics. In this case, the size or footprint of the received beam is much larger than the aperture area[5] since the beam diverges as it travels through space [8]. At the receiver aperture, only a small fraction of the total beam energy is captured and focused as a small spot or airy pattern on detector array that is located at the focal plane of the aperture lens.

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 mth 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:

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

where $\lambda_s(x,y)$ is the scaled beam intensity[6] profile on the detector array, λ_n is the scaled (dark current) noise intensity profile, A_m is the region of the mth detector on the detector array, $Z_1, Z_2, \ldots, Z_M$ 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, λ_n is constant factor with respect to (x,y) that accounts for the dark current noise in the detector array [22].

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

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

where I_0/ρ^2 is the peak intensity measured in terms of the average number of signal photons measured during an observation interval. Furthermore, λ_n is also measured in terms of the average number of noise photons generated during the same observation interval. The quantity ρ 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 $\mathcal{A}$ is the region of the detector array.

2. The assumption holds if the background radiation power can be minimized using effective optical filters.

3. This is especially true if the parameters themselves—such as the covariance matrices of the random motion model—are time-varying.

4. 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.

5. The beam radius ρ at a distance z meters from the transmitter is given by $\rho(z) = \rho_0\sqrt{1+(\frac{\lambda z}{\pi\rho_0^2})^2}$, where ρ_0 is the beam waist (in meters) and λ is the wavelength in meters.

6. The actual signal intensity, λ_{s_i}, and the actual noise intensity, λ_{n_i}, are both measured in terms of Joules/mm^2/s. However, they are multiplied by the constant $\frac{\eta T_s}{hc/\lambda}$ in order to obtain the intensity λ_s and λ_n for the photon generation model in (1). The constant h is known as the *Planck's constant*, and its value is $6.62607004\times10^{-34}\,\mathsf{m}^2\mathsf{kg/s}$. The constant c is the speed of light in vacuum which is about $3\times10^8\,\mathsf{m/s}$, λ is the wavelength of light in meters, η stands for the photoconversion efficiency, and T_s represents signal pulse duration in seconds.

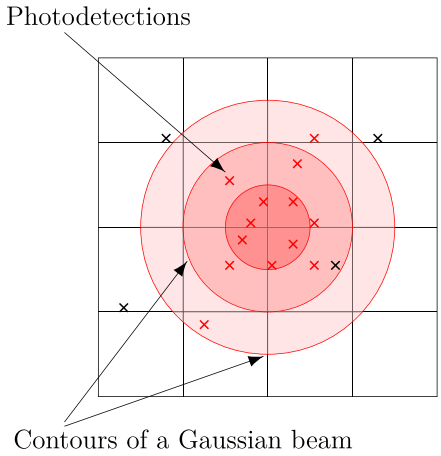

FIGURE 1. This figure depicts contours of incident light intensity (intensity is distributed as a Gaussian function) and the resulting photodetections for a 4 × 4 detector array. The red crosses represent the locations of signal photodetections, and the black ones correspond to noise.

TABLE 1. List of mathematical symbols.

Symbols	Definition
I_0	peak signal intensity (average number of photons)
λ_n	noise intensity (average number of noise photons)
ρ	beam radius (mm)
(x_0, y_0)	location of the beam center on the array
$\mathcal{A}$	region of the detector array
$\lvert\mathcal{A}\rvert$	area of the detector array (mm^2)
M	number of detectors in the array
N	number of pilot symbols
A_m	region of the mth detector in the array
A	area of a given detector (mm^2)
$\ell(\mathcal{R})$	length of one side of any square region $\mathcal{R}$ (mm)
Λ_m	total power (signal+noise) in the mth detector
$\lambda_m^{(s)}$	signal power in the mth detector
Λ_s	total power in the array (signal+noise)
$\mathbb{E}$	probabilistic expectation operator
Var	variance operator

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 1 for a complete list of all mathematical symbols.

IV. CRAMÉR-RAO LOWER BOUND

A. WHY COMPUTE THE CRAMÉR-RAO LOWER BOUND FOR BEAM POSITION ESTIMATORS?

As we know, the Cramér-Rao Lower Bound acts as a lower bound on the variance of any unbiased estimator. The motivation to use this bound in this paper stems from the fact that this bound help us obtain additional insights about the system performance that are helpful from a system design perspective.

For instance, depending on the size of a single detector in the array, there is an optimal beam radius that will minimize the Cramér-Rao bound (see Fig. 9 in revised manuscript) for a given value of received signal-to-noise ratio. Additionally, the difference in performance of slot-period and symbol-period based observations of a data-aided tracking system (based on Pulse Position Modulation) is highlighted by the Cramér-Rao bounds (see Fig. 11, Fig. 12, Fig. 13 and Fig. 14 of revised manuscript). Thus, we are informed that for the observation period, we should focus on a slot—instead of a full symbol period—in order to minimize the estimation error (especially at high-to-moderate received signal-to-noise ratio). Last, but not the least, we also note from Fig. 12 that the difference in the Cramér-Rao bounds shrinks as we increase the number of detectors in the array. Thus, employing an array with too big a number of detectors will only lead to a more computationally complex system without a significant performance gain.

B. CRAMÉR-RAO LOWER BOUND AND OUTAGE PROBABILITY

Let us first define that the outage occurs as soon as the beam center coincides and moves past one of the edges of the array while moving away from the array center.[7] Let us assume that an efficient[8] estimator exists . Additionally, let us assume that the error of this estimator is Gaussian, and that the error in each of the two dimensions is independent and identically distributed with a certain variance σ_0^2. Let us denote the Cramér-Rao bound by $\mathcal{C}$. Then, the Cramér-Rao bound in this case is $\mathcal{C} \triangleq 2\sigma_0^2$. Also, it follows, that the deviation between the array center and the beam center is given by a two-dimensional Gaussian random vector $\boldsymbol{X}_0$ with covariance matrix $\sigma_0^2\mathbf{I}$, where $\mathbf{I}$ is a 2×2 identity matrix.

Before we analyze the outage, let us define how the tracking system works. The beam position is assumed to follow a random walk on the array in two dimensions due to physical affects such as turbulence and mechanical vibrations. We assume that the beam center's initial position lies at the origin, where the origin is defined to be the center of the array. We assume that the beam position evolves in time according to a Gauss-Markov process. The tracking block senses (or estimates) the beam position every T_s units of time, and aligns the beam center with the center of the array. We assume that the deviation of the beam center from the origin during time interval T_s is modeled by a two-dimensional Gaussian random vector $\boldsymbol{X}_1$ with a covariance matrix $\sigma_1^2\mathbf{I}$. We additionally assume that $\boldsymbol{X}_0$ and $\boldsymbol{X}_1$ are mutually independent Gaussian random vectors.

Due to the error in estimation, the beam is not perfectly aligned at the origin, but aligned with a two dimensional error $\boldsymbol{X}_0$. Thus the outage occurs when the quantity $R \triangleq \|\boldsymbol{X}_0 + \boldsymbol{X}_1\|_2$ exceeds a certain threshold. We note that R is a Rayleigh random variable with scale parameter $\sigma_R \triangleq$

7. Approximately, the outage occurs when the signal power falls below half of the maximum received power.

8. An estimator is said to be efficient if it is unbiased and achieves the Cramér-Rao bound.

$\sqrt{\sigma_0^2+\sigma_1^2}$. Then, an upper bound on the outage probability can be derived as

$$\mathbb{P}_{\text{out}} \leq \mathbb{P}\left(\left\{R \geq \frac{\sqrt{\mathcal{A}}}{2}\right\}\right) = \exp\left(-\frac{\mathcal{A}}{8\left(\sigma_1^2+\sigma_0^2\right)}\right) = \exp\left(-\frac{\mathcal{A}}{8\left(\sigma_1^2+\frac{\mathcal{C}}{2}\right)}\right), \tag{3}$$

where $\mathcal{A}$ is the area of the square array, and $\sqrt{\mathcal{A}}$ is the length of one side of the array. We note that a smaller value of Cramér-Rao bound $\mathcal{C}$ leads to a smaller (upper bound on) outage probability through the expression (3).

We remind the reader here that the upper bound in (3) represents the ideal scenario since it corresponds to an efficient estimator that achieves the bound. This is the best possible outage probability scenario for any unbiased estimator of beam position given the received SNR and other channel parameters.

V. CRAMÈR-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.

Let $\triangleq \begin{bmatrix} x_0 & y_0 & I_0 & \lambda_n \end{bmatrix}$.[9] The likelihood function is given by

$$p(\boldsymbol{Z}|) = \prod_{m=1}^{M} e^{-\Lambda_m} \frac{\Lambda_m^{z_m}}{z_m!}, \tag{4}$$

where

$$\Lambda_m \triangleq \iint_{A_m} \left(\frac{I_0}{\rho^2} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} + \lambda_n\right) dx\, dy, \tag{5}$$

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

$$\ln p(\boldsymbol{Z}|) = \sum_{m=1}^{M} (Z_m \ln \Lambda_m - \Lambda_m - \ln Z_m!) = \sum_{m=1}^{M} (Z_m \ln \Lambda_m - \ln Z_m!) - \Lambda_s. \tag{6}$$

9. 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.

A. ESTIMATION BASED ON PILOT SYMBOLS

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

a) *Cramèr-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 (6):

$$\frac{\partial \ln p(\boldsymbol{Z}|)}{\partial I_0} = \sum_{m=1}^{M} \frac{Z_m}{\Lambda_m} \iint_{A_m} \frac{1}{\rho^2} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx\, dy. \tag{7}$$

The second derivative of (6) with respect to I_0 is

$$\frac{\partial^2 \ln p(\boldsymbol{Z}|)}{\partial I_0^2} = -\sum_{m=1}^{M} \frac{Z_m}{\Lambda_m^2} \left(\iint_{A_m} \frac{1}{\rho^2} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx\, dy\right)^2. \tag{8}$$

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

$$-\mathbb{E}\left[\frac{\partial^2 \ln p(\boldsymbol{Z}|)}{\partial I_0^2}\right] = \sum_{m=1}^{M} \frac{1}{\Lambda_m} \left(\iint_{A_m} \frac{1}{\rho^2} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy\right)^2. \tag{9}$$

1) CRAMÈR-RAO LOWER BOUND OF λ_N

In this case, we assume that the dark current noise parameter λ_n 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 (9), it can be easily shown that the Cramèr-Rao lower bound on the variance of any unbiased estimator $\hat{\lambda}_n$ for one noise-only slot is given by

$$-\mathbb{E}\left[\frac{\partial^2 \ln p(\boldsymbol{Z}|)}{\partial \lambda_n^2}\right] = \sum_{m=1}^{M} \frac{A^2}{\Lambda_m}. \tag{10}$$

2) CRAMÈR-RAO LOWER BOUNDS OF (X_0, Y_0)

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

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

$$\text{Var}[\hat{y}_0] \geq \frac{\sum_{m=1}^{M} \frac{1}{\Lambda_m} \left(\iint_{A_m} \frac{I_0}{\rho^4} (x-x_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx dy\right)^2}{\Psi(x_0, y_0, I_0, \rho)}, \tag{12}$$

where $\Psi(x_0, y_0, I_0, \rho)$ is defined in (54).

3) CRAMÈR-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.[10] We denote the 3×3 Fisher Information Matrix by $\mathcal{I}(x_0, y_0, I_0)$.

For the sake of compactness, let us define the following quantities:

$$\psi_m^{(0)} \triangleq \iint_{A_m} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy,$$
$$\psi_m^{(1)} \triangleq \iint_{A_m} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} (y-y_0)\, dx\, dy,$$
$$\psi_m^{(2)} \triangleq \iint_{A_m} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} (x-x_0)\, dx\, dy.$$

Then,

$$\mathsf{Var}\big[\hat{x}_0\big] \geq \Big[\mathcal{I}^{-1}(x_0, y_0, I_0)\Big]_{1,1} = \left[\sum_{m=1}^{M} \frac{\left(\psi_m^{(0)}\right)^2}{\Lambda_m \rho^4} \sum_{m=1}^{M} \frac{I_0^2\left(\psi_m^{(1)}\right)^2}{\Lambda_m \rho^8} - \sum_{m=1}^{M} \frac{I_0 \psi_m^{(0)} \psi_m^{(1)}}{\Lambda_m \rho^6}\right] \times |\mathcal{I}(x_0, y_0, I_0)|^{-1}. \quad (13)$$

$$\mathsf{Var}\big[\hat{y}_0\big] \geq \Big[\mathcal{I}^{-1}(x_0, y_0, I_0)\Big]_{2,2} = \left[\sum_{m=1}^{M} \frac{\left(\psi_m^{(0)}\right)^2}{\Lambda_m \rho^4} \sum_{m=1}^{M} \frac{I_0^2\left(\psi_m^{(2)}\right)^2}{\Lambda_m \rho^8} - \sum_{m=1}^{M} \frac{I_0 \psi_m^{(0)} \psi_m^{(2)}}{\Lambda_m \rho^6}\right] \times |\mathcal{I}(x_0, y_0, I_0)|^{-1}. \quad (14)$$

$$\mathsf{Var}\Big[\hat{I}_0\Big] \geq \Big[\mathcal{I}^{-1}(x_0, y_0, I_0)\Big]_{3,3} = \left[\sum_{m=1}^{M} \frac{I_0^2\left(\psi_m^{(2)}\right)^2}{\Lambda_m \rho^8} \sum_{m=1}^{M} \frac{I_0^2\left(\psi_m^{(1)}\right)^2}{\Lambda_m \rho^8} - \sum_{m=1}^{M} \frac{I_0^2 \psi_m^{(1)} \psi_m^{(2)}}{\Lambda_m \rho^8}\right] \times |\mathcal{I}(x_0, y_0, I_0)|^{-1}. \quad (15)$$

The determinant of the Fisher information matrix is given by (16), As shown at the bottom of the page.

10. In order to lower the complexity of the estimation problem, we can estimate λ_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.

B. CRAMÉR-RAO LOWER BOUNDS OF (X_0, Y_0): ASYMPTOTIC CASE ($M \to \infty$)

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 $M \to \infty$ for fixed array area, the true location of each photodetection can be reported by the infinitesimally small detector. When $M \to \infty$, 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 ($M \geq 64$) [4], [7]. Therefore, in this section, we look at the Cramér-Rao lower bound of (x_0, y_0) for the $M \to \infty$ 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 ρ.

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,

$$\lambda_n A << \iint_{A_m} \frac{I_0}{\rho^2} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy.$$

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 mth small detector, and Δ_M is its infinitesimal area. Then, the numerator of (11) simplifies as

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

$$|\mathcal{I}(x_0, y_0, I_0)| = -\sum_{m=1}^{M} \frac{I_0}{\rho^6 \Lambda_m} \psi_m^{(0)} \psi_m^{(1)} \left(\sum_{m=1}^{M} \frac{I_0^2}{\rho^8 \Lambda_m} \left(\psi_m^{(2)}\right)^2 \sum_{m=1}^{M} \frac{I_0}{\rho^6 \Lambda_m} \psi_m^{(0)} \psi_m^{(1)} - \sum_{m=1}^{M} \frac{I_0}{\rho^6 \Lambda_m} \psi_m^{(0)} \psi_m^{(2)} \sum_{m=1}^{M} \frac{I_0^2}{\rho^8 \Lambda_m} \psi_m^{(2)} \psi_m^{(1)}\right)$$
$$+ \sum_{m=1}^{M} \frac{I_0}{\rho^6 \Lambda_m} \psi_m^{(0)} \psi_m^{(2)} \left(\sum_{m=1}^{M} \frac{I_0}{\rho^6 \Lambda_m} \psi_m^{(0)} \psi_m^{(1)} \sum_{m=1}^{M} \frac{I_0^2}{\rho^8 \Lambda_m} \psi_m^{(2)} \psi_m^{(1)} - \sum_{m=1}^{M} \frac{I_0}{\rho^6 \Lambda_m} \psi_m^{(0)} \psi_m^{(2)} \sum_{m=1}^{M} \frac{I_0^2}{\rho^8 \Lambda_m} \left(\psi_m^{(1)}\right)^2\right)$$
$$+ \sum_{m=1}^{M} \frac{1}{\rho^4 \Lambda_m} \left(\psi_m^{(0)}\right)^2 \left(\sum_{m=1}^{M} \frac{I_0^2}{\rho^8 \Lambda_m} \left(\psi_m^{(2)}\right)^2 \sum_{m=1}^{M} \frac{I_0^2}{\rho^8 \Lambda_m} \left(\psi_m^{(1)}\right)^2 - \left(\sum_{m=1}^{M} \frac{I_0^2}{\rho^8 \Lambda_m} \psi_m^{(2)} \psi_m^{(1)}\right)^2\right) \quad (16)$$

$$= \sum_{m=1}^{M} \frac{I_0}{\rho^6}(y_m - y_0)^2 e^{-\frac{(x_m-x_0)^2+(y_m-y_0)^2}{2\rho^2}} \Delta_M$$
$$\approx \frac{I_0 2\pi}{\rho^4} \sum_{m=1}^{M} \iint_{A_m} \frac{1}{2\pi\rho^2}(y - y_0)^2 e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx\, dy$$
$$= \frac{I_0 2\pi}{\rho^4} \iint_{\mathcal{A}} \frac{1}{2\pi\rho^2}(y - y_0)^2 e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx\, dy$$
$$\approx \frac{I_0 2\pi}{\rho^4}\rho^2 = \frac{I_0 2\pi}{\rho^2}, \tag{17}$$

where, in the last approximation of (17) 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 $\rho << \ell(|\mathcal{A}|)$.

The positive term in the denominator (see (54)) can be simplified in a similar fashion. In the denominator, the square root of the term with minus sign can be simplified as

$$\sum_{m=1}^{M}\left(\frac{1}{\Lambda_m}\iint_{A_m}\frac{I_0}{\rho^4}(y - y_0)e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx\, dy \times \iint_{A_m}\frac{I_0}{\rho^4}(x - x_0)e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx\, dy\right)$$
$$\approx \sum_{m=1}^{M}\left(\frac{\frac{I_0}{\rho^4}(y_m - y_0)e^{-\frac{(x_m-x_0)^2+(y_m-y_0)^2}{2\rho^2}}\Delta_M}{\frac{I_0}{\rho^2}e^{-\frac{(x_m-x_0)^2+(y_m-y_0)^2}{2\rho^2}}\Delta_M} \times \frac{I_0}{\rho^4}(x_m - x_0)e^{-\frac{(x_m-x_0)^2+(y_m-y_0)^2}{2\rho^2}}\Delta_M\right)$$
$$\approx \frac{I_0 2\pi}{\rho^4}\sum_{m=1}^{M}\iint_{A_m}\frac{1}{2\pi\rho^2}(y - y_0)(x - x_0)e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dxdy$$
$$= \frac{I_0 2\pi}{\rho^4}\iint_{\mathcal{A}}\frac{1}{2\pi\rho^2}(y - y_0)(x - x_0)e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx\, dy$$
$$\approx 0, \tag{18}$$

where in the last approximation of (18), 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 ρ^2, but with different means: $\mathbb{E}[X] = x_0$, $\mathbb{E}[Y] = y_0$. Therefore,

$$\mathsf{Var}[\hat{x}_0] \geq \frac{\frac{I_0 2\pi}{\rho^2}}{\frac{I_0 2\pi}{\rho^2}\times\frac{I_0 2\pi}{\rho^2}} = \frac{\rho^2}{I_0 2\pi}. \tag{19}$$

We note that the Cramér-Rao lower bound is minimized by minimizing ρ (a more focused beam) for fixed signal power. The Cramér-Rao lower bounds goes to zero at the rate $\mathcal{O}(\rho^2)$ as $\rho \to 0$, 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 $\boldsymbol{X_0}$: LOW SIGNAL-TO-NOISE RATIO

In this case, let us assume that

$$\lambda_n A >> \iint_{A_m}\frac{I_0}{\rho^2}e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx\, dy.$$

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 (54)) is

$$\frac{1}{\lambda_n A}\sum_{m=1}^{M}\left(\iint_{A_m}\frac{I_0}{\rho^4}(y - y_0)e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx\, dy \times \iint_{A_m}\frac{I_0}{\rho^4}(x - x_0)e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx\, dy\right)$$
$$\xrightarrow{M \text{ large}} \frac{1}{\lambda_n \Delta_M}\sum_{m=1}^{M}\frac{I_0^2}{\rho^8}(y_m - y_0)(x_m - x_0)e^{-\frac{(x_m-x_0)^2+(y_m-y_0)^2}{\rho^2}} \times \Delta_M^2$$
$$\approx \frac{\pi\rho^2 I_0^2}{\rho^8\lambda_n}\iint_{\mathcal{A}}\frac{1}{2\pi\frac{\rho^2}{2}}(y - y_0)(x - x_0)e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2/2}} dx\, dy$$
$$\approx 0. \tag{20}$$

Therefore, by further simplification,

$$\mathsf{Var}[\hat{x}_0] \geq \frac{\frac{\lambda_n\rho^8}{I_0^2}A}{\sum_{m=1}^{M}\left(\iint_{A_m}(x - x_0)e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} dx\, dy\right)^2}$$
$$\xrightarrow{M\to\infty} \frac{\frac{\lambda_n\rho^8}{I_0^2}}{\iint_{\mathcal{A}}(x - x_0)^2 e^{-\frac{(x-x_0)^2+(y-y_0)^2}{\rho^2}} dx\, dy} = \frac{2\rho^4}{\pi\left(\frac{I_0^2}{\lambda_n}\right)}. \tag{21}$$

In this case, the Cramér-Rao lower bound goes to zero at a rate $\mathcal{O}(\rho^4)$ as $\rho \to 0$. 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.

VI. CRAMÈR-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

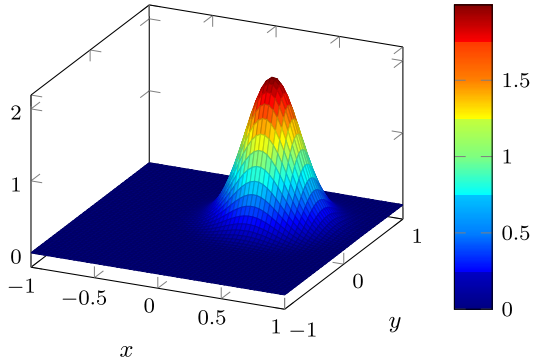

FIGURE 2. The airy pattern's energy distribution on the array is modeled by a circularly symmetric Gaussian in two dimensions.

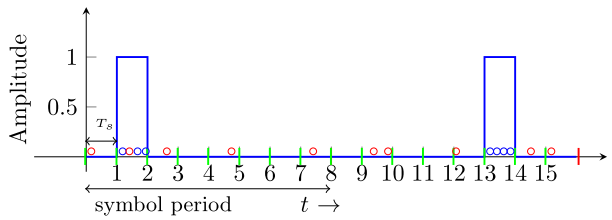

FIGURE 3. This figure shows an 8-PPM scheme. The blue circles indicate signal photons and the red circles indicate noise photons.

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. In this regard, we may either use 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.

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 $Z_m \sim Poisson(\Lambda_m')$, where

$$\Lambda_m' \triangleq \iint_{A_m} \left(\frac{I_0}{\rho^2} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} + \lambda_n' \right) dx\, dy. \tag{22}$$

Thus, in this case,

$$p(\mathbf{Z}|) \triangleq \prod_{m=1}^{M} \frac{e^{-\Lambda_m'} \left(\Lambda_m'\right)^{Z_m}}{Z_m!}. \tag{23}$$

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. 5 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 jth slot as its observation interval.

For observation based on one slot, $Z_m \sim Poisson(\Lambda_m)$, with probability P_c and $Z_m \sim Poisson(\lambda_n A)$ 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:

$$p(\xi) \;= P_c \delta(\xi - \Lambda_m) + (1 - P_c)\delta(\xi - \lambda_n A), \tag{24}$$

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

$$p(\mathbf{Z}|) \triangleq P_c \prod_{m=1}^{M} \frac{e^{-\Lambda_m} \Lambda_m^{Z_m}}{Z_m!} + (1 - P_c) \prod_{m=1}^{M} \frac{e^{-\lambda_n A} (\lambda_n A)^{Z_m}}{Z_m!}. \tag{25}$$

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

$$P_c = (P(\{Z_s > Z_n\}))^{K-1} = (P(\{(Z_s - Z_n) > 0\}))^{K-1}. \tag{26}$$

In (26), $Z_s \sim Poisson(\Lambda_s)$ and $Z_n \sim Poisson(\lambda_n|\mathcal{A}|)$. The random variable $Z \triangleq Z_s - Z_n$ is a (discrete) *Skellam* random variable whose distribution is

$$P(\{Z = z\}) = e^{-(\Lambda_s + \lambda_n|\mathcal{A}|)} \left(\frac{\Lambda_s}{\lambda_n|\mathcal{A}|} \right)^{z/2} I_z\left(2\sqrt{\Lambda_s \lambda_n|\mathcal{A}|}\right), \tag{27}$$

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

$$\begin{aligned} P(\{Z > 0\}) &= \sum_{z=1}^{\infty} P(\{Z = z\}) \\ &= \sum_{z=1}^{\infty} e^{-(\Lambda_s + \lambda_n|\mathcal{A}|)} \left(\frac{\Lambda_s}{\lambda_n|\mathcal{A}|} \right)^{z/2} I_z\left(2\sqrt{\Lambda_s \lambda_n|\mathcal{A}|}\right). \end{aligned} \tag{28}$$

Fig. 4 shows the probability of correct decision P_c for different values of beam radius ρ. 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.

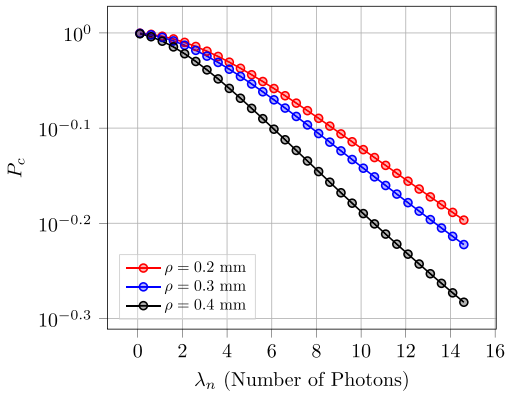

FIGURE 4. This figure shows the probability of correct decision of the equal gain combiner as a function of noise parameter λ_n for different values of beam radius ρ for a 4 × 4 detector array. The modulation scheme considered in this case is 8-PPM. The received signal intensity is measured in terms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

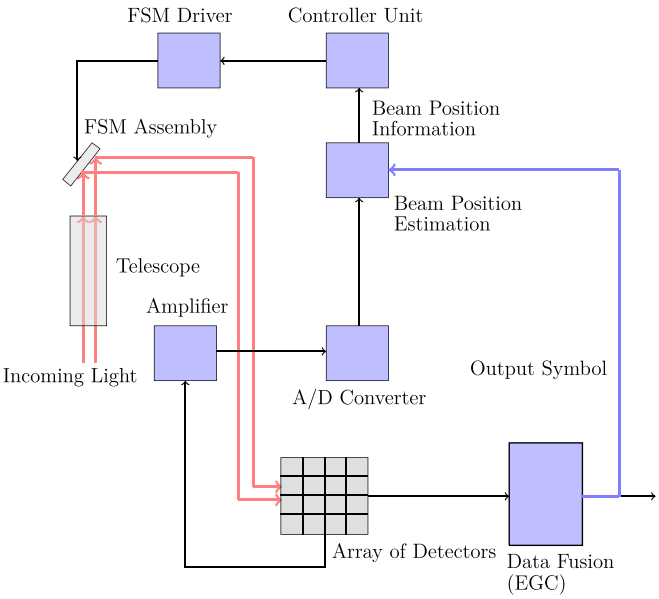

FIGURE 5. This figure shows the block diagram of the decision-aided beam position estimation system. The output of the equal gain combiner is fed into the beam position estimation block so that the "right" slot is chosen for estimation purpose.

Finally, since $P_e = 1 - P_c$, we have that

$$P_e = 1 - \left(\sum_{z=1}^{\infty} e^{-(\Lambda_s + \lambda_n |\mathcal{A}|)} \left(\frac{\Lambda_s}{\lambda_n |\mathcal{A}|} \right)^{z/2} I_z \left(2\sqrt{\Lambda_s \lambda_n |\mathcal{A}|} \right) \right)^{K-1}. \tag{29}$$

1) MONTE CARLO EXPECTATION

It is not straightforward to compute the probabilistic expectations $\mathbb{E}[\frac{\partial^2 \ln p(\mathbf{Z}|)}{\partial I_0^2}]$, $\mathbb{E}[\frac{\partial^2 \ln p(\mathbf{Z}|)}{\partial I_0 \partial x_0}]$, $\mathbb{E}[\frac{\partial^2 \ln p(\mathbf{Z}|)}{\partial I_0 \partial y_0}]$, $\mathbb{E}[\frac{\partial^2 \ln p(\mathbf{Z}|)}{\partial x_0^2}]$, $\mathbb{E}[\frac{\partial^2 \ln p(\mathbf{Z}|)}{\partial y_0^2}]$ and $\mathbb{E}[\frac{\partial^2 \ln p(\mathbf{Z}|)}{\partial x_0 \partial y_0}]$ for the likelihood function in (25). Thus, we resort to the 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 $Poisson(\Lambda_2), \ldots, Z_M$ from $Poisson(\Lambda_M)$. Else, sample Z_1 from $Poisson(\lambda_n A)$, Z_2 from $Poisson(\lambda_n A), \ldots, Z_M$ from $Poisson(\lambda_n A)$.
3) Substitute the Z_m's obtained from Step 2 into each of the second order partial derivatives: $\frac{\partial^2 \ln p(\mathbf{Z}|)}{\partial I_0^2}$, $\frac{\partial^2 \ln p(\mathbf{Z}|)}{\partial I_0 \partial x_0}$, $\frac{\partial^2 \ln p(\mathbf{Z}|)}{\partial I_0 \partial y_0}$, $\frac{\partial^2 \ln p(\mathbf{Z}|)}{\partial x_0^2}$, $\frac{\partial^2 \ln p(\mathbf{Z}|)}{\partial y_0^2}$ and $\frac{\partial^2 \ln p(\mathbf{Z}|)}{\partial x_0 \partial y_0}$, 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.

VII. 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 dark current noise parameter λ_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 $\ln p(\mathbf{Z}|)$, 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 λ_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 $(\hat{x}_0, \hat{y}_0)$.

Here, we want to point out that even though the centroid estimator can also be used to estimate the beam position on the array (and that the complexity of centroid is significantly smaller), we have deliberately chosen to focus on the maximum likelihood estimator in this study. This is because of the fact that the centroid estimator is heavily biased under conditions of moderate to low signal-to-noise ratio [4], [23], and the bias in maximum likelihood estimator is negligible even for moderate signal-to-noise ratio [23]. Therefore, the comparison of the maximum likelihood estimator's means-square error with the Crameŕ-Rao lower bound is more meaningful since the Crameŕ-Rao lower bound is a lower bound on the variance (or mean-square error) of an unbiased estimator.

A. METHOD OF MOMENTS ESTIMATOR OF I_0 AND λ_N

1) PILOT SYMBOL CASE

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

$$\hat{I}_0 = \frac{1}{2\pi N}\left(\sum_{i=1}^{N}\sum_{m=1}^{M} Z_{i,m}^{(s)}\right) - \frac{\hat{\lambda}_n|\mathcal{A}|}{2\pi}, \tag{30}$$

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

$$\hat{\lambda}_n = \frac{1}{|\mathcal{A}|N}\sum_{i=1}^{N}\sum_{m=1}^{M} Z_{i,m}^{(n)} \tag{31}$$

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

$$\begin{aligned}\mathbb{E}\Big[\hat{I}_0\Big] &= \frac{1}{2\pi N}\left(\sum_{i=1}^{N}\sum_{m=1}^{M}\mathbb{E}\Big[Z_{i,m}^{(s)}\Big]\right) - \mathbb{E}\Big[\hat{\lambda}_n\Big]\frac{|\mathcal{A}|}{2\pi}\\ &= \frac{1}{2\pi}\sum_{m=1}^{M}\iint_{A_m}\frac{I_0}{\rho^2}e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\,dx\,dy + \lambda_n A - \lambda_n\frac{|\mathcal{A}|}{2\pi}\\ &= \frac{1}{2\pi}(I_0 2\pi + \lambda_n|\mathcal{A}|) - \frac{\lambda_n|\mathcal{A}|}{2\pi} = I_0.\end{aligned} \tag{32}$$

Thus, both $\hat{I}_0$ and $\hat{\lambda}_n$ are unbiased estimators of I_0 and λ_n, respectively. Additionally, it is straightforward to verify that

$$\mathbb{E}\left[\Big(\hat{\lambda}_n - \lambda_n\Big)^2\right] = \mathbb{E}\Big[\hat{\lambda}_n^2\Big] - \lambda_n^2 = \frac{\lambda_n}{NMA} = \frac{\lambda_n}{N|\mathcal{A}|}. \tag{33}$$

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

$$\begin{aligned}\mathbb{E}\left[\Big(\hat{I}_0 - I_0\Big)^2\right] &= \mathsf{Var}\Big(\hat{I}_0 - I_0\Big) = \mathsf{Var}\Big(\hat{I}_0\Big)\\ &= \frac{1}{(2\pi N)^2}\sum_{i=1}^{N}\sum_{m=1}^{M}\mathsf{Var}\Big(Z_{i,m}^{(s)}\Big) + \left(\frac{AM}{2\pi}\right)^2\mathsf{Var}\Big(\hat{\lambda}_n\Big)\\ &= \frac{I_0}{2\pi N} + \frac{\lambda_n AM}{2\pi^2 N} = \frac{1}{2\pi N}\left(I_0 + \frac{\lambda_n AM}{\pi}\right).\end{aligned} \tag{34}$$

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,

$$\mathbb{E}\Big[\hat{I}_0\Big] = \frac{1}{2\pi N}\sum_{i=1}^{N}\sum_{m=1}^{M}\mathbb{E}[Z_m] - \mathbb{E}\Big[\hat{\lambda}_n\Big]\frac{AM}{2\pi}, \tag{35}$$

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

$$\mathbb{E}[Z_m] = \Lambda_m^{(s)}P_c + \lambda_n A. \tag{36}$$

Therefore,

$$\begin{aligned}\mathbb{E}\Big[\hat{I}_0\Big] &= \frac{1}{2\pi N}\sum_{i=1}^{N}\sum_{m=1}^{M}\Big(\Lambda_m^{(s)}P_c + \lambda_n A\Big) - \frac{\lambda_n AM}{2\pi}\\ &= P_c I_0.\end{aligned} \tag{37}$$

Moreover,

$$\begin{aligned}\mathsf{Var}\Big[\hat{I}_0\Big] &= \frac{1}{(2\pi N)^2}\sum_{i=1}^{N}\sum_{m=1}^{M}\mathsf{Var}[Z_m] + \mathsf{Var}\Big[\hat{\lambda}_n\Big]\left(\frac{AM}{2\pi}\right)^2\\ &= \frac{1}{(2\pi N)^2}\sum_{i=1}^{N}\sum_{m=1}^{M}\Lambda_m^{(s)}(P_c + \lambda_n A) + \frac{\lambda_n AM}{N(2\pi)^2}\\ &= \frac{P_c I_0}{2\pi N} + \frac{\lambda_n AM}{2\pi^2 N} = \frac{1}{2\pi N}\left(P_c I_0 + \frac{\lambda_n AM}{\pi}\right).\end{aligned} \tag{38}$$

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

$$\begin{aligned}\mathbb{E}\left[\Big(I_0 - \hat{I}_0\Big)^2\right] &= \mathsf{Var}\Big[I_0 - \hat{I}_0\Big] + \Big(\mathbb{E}\Big[I_0 - \hat{I}_0\Big]\Big)^2\\ &= \mathsf{Var}\Big(\hat{I}_0\Big) + (I_0 - I_0P_c)^2\\ &= \frac{1}{2\pi N}\left(P_c I_0 + \frac{\lambda_n AM}{\pi}\right) + I_0^2(1-P_c)^2.\end{aligned} \tag{39}$$

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 [4]:

$$\begin{aligned}(\hat{x}_0, \hat{y}_0) &\triangleq \arg\max_{x_0,y_0} \ln p(Z_1, Z_2, \ldots, Z_M|x_0, y_0)\\ &= \arg\max_{x_0,y_0}\sum_{m=1}^{M}\left(Z_m \ln\left(\hat{I}_0 2\pi\left[\Phi\left(\frac{y_{m_2}-y_0}{\rho}\right) - \Phi\left(\frac{y_{m_1}-y_0}{\rho}\right)\right]\right.\right.\\ &\qquad\left.\left.\times\left[\Phi\left(\frac{x_{m_2}-x_0}{\rho}\right) - \Phi\left(\frac{x_{m_1}-x_0}{\rho}\right)\right] + \hat{\lambda}_n A\right)\right)\end{aligned} \tag{40}$$

$$\begin{aligned}&- \left(\hat{I}_0 2\pi\left[\Phi\left(\frac{\frac{\ell(\mathcal{A})}{2}-y_0}{\rho}\right) - \Phi\left(\frac{-\frac{\ell(\mathcal{A})}{2}-y_0}{\rho}\right)\right]\right.\\ &\qquad\left.\times\left[\Phi\left(\frac{\frac{\ell(\mathcal{A})}{2}-x_0}{\rho}\right) - \Phi\left(\frac{-\frac{\ell(\mathcal{A})}{2}-x_0}{\rho}\right)\right] + \hat{\lambda}_n|\mathcal{A}|\right),\end{aligned} \tag{41}$$

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 λ_n, respectively. The quantity (x_{m_2}, y_{m_2}) is the location of the north east corner of the mth detector, and (x_{m_1}, y_{m_1}) is the position of south west corner.

For estimation based on data symbols, the maximum likelihood estimate is obtained by maximizing (23) (symbol period) or by maximizing (25) (slot period). Regarding the maximization of all loglikelihood functions, we utilize a genetic algorithm to search 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.

VIII. 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, and $(x_0, y_0) = (0.4, 0.4)$ for all figures except for Fig. 10. The low photon rate channels are of interest in deep space communications where the received signal energy can be so low that we are only able to detect a few signal photons per slot of a PPM symbol [1], [4]. 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. 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 is "blocked" for a few microseconds before it is ready to detect the next incoming photon.

For all experiments, the area of the detector array $|\mathcal{A}| = 4$ 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 detectors 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 CRLB(x_0, y_0) denotes the sum of individual Cramér-Rao lower bounds: CRLB(x_0) and 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. 6 indicates the Cramér-Rao lower bound plots as a three-parameter estimation problem as defined in (13), (14) and (15), whereas Fig. 7 indicates the decay of the Cramér-Rao bound as a function of the average number of signal photons when the noise power is fixed. Fig. 8 depicts the Cramér-Rao lower bound curves as a function of number of pilot symbols used as observations in the estimation of parameters.

Fig. 9 shows the Cramér-Rao lower bound plots as a function of beam radius ρ (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 ρ. However, for the CRLB of $(\hat{x}_0, \hat{y}_0)$, we see that there is an optimum value of ρ (lets call it ρ^*_M for the M-cell array) 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 to estimate the beam position more accurately as opposed to a more "spread out" beam.

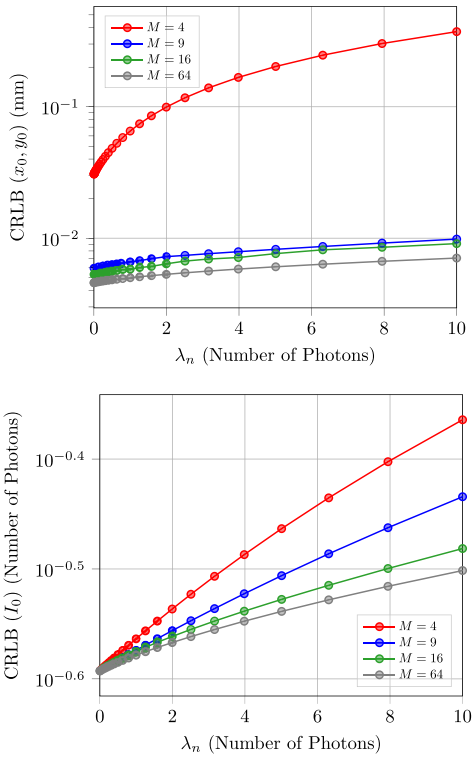

FIGURE 6. This figure shows the Cramèr-Rao lower bounds for different values of detectors M in the array as a function of noise parameter λ_n for one pilot symbol. The beam radius is 0.2 mm.

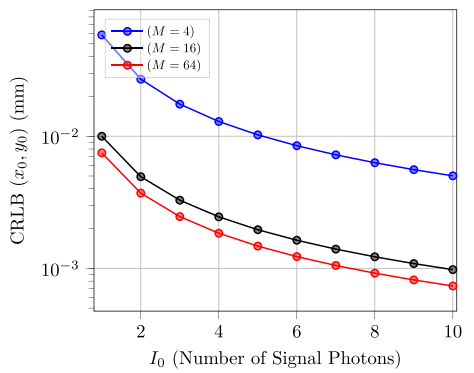

FIGURE 7. This figure shows the effect of the number of signal intensity (measure in terms of an average number of signal photons) on the Cramér-Rao bound of the beam position estimator. The value of noise parameter is $\lambda_n = 0.1$ and the beam radius is 0.2 millimeters.

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 all 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 ρ diminishes beyond a certain (optimum) value.

Fig. 10 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)

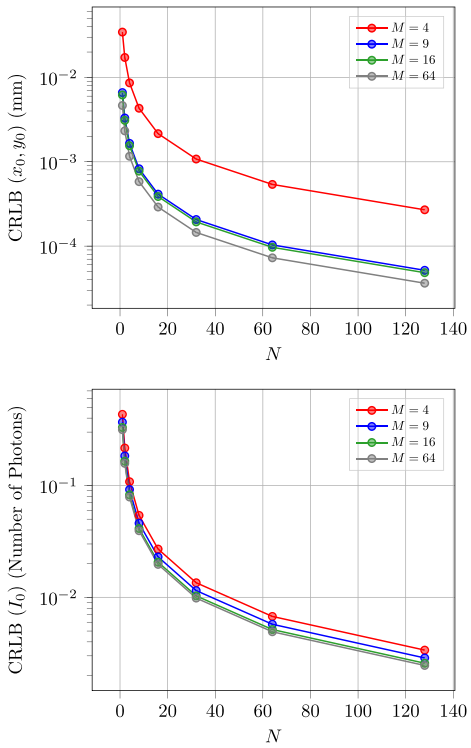

FIGURE 8. This figure shows the Cramèr-Rao lower bounds as a function of the number of pilot symbols N. The average noise power is $\lambda_n = 1$ photons/slot period for the figure on top, and $\lambda_n = 10$ photons/slot period for the figure at the bottom. The beam radius is 0.2 mm.

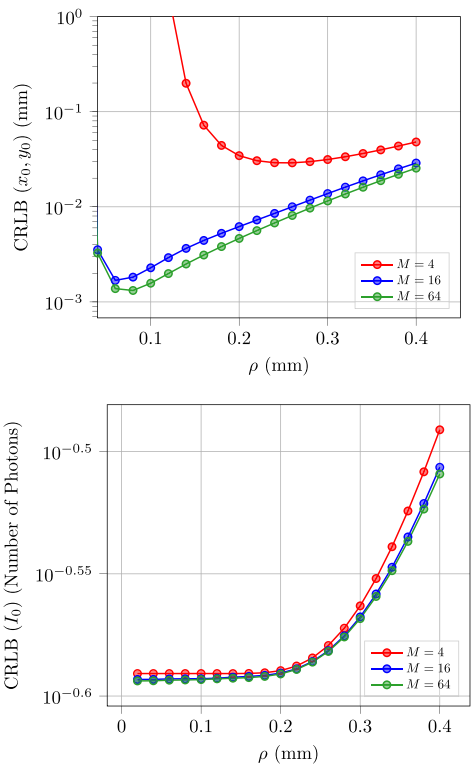

FIGURE 9. This figure shows the Cramèr-Rao lower bounds as a function of the beam radius ρ. The average number of noise photons $\lambda_n = 1$ photon/slot period for both the figures.

for 4×4 and 8×8 detector arrays 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 ρ, 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 $\rho = 0.14$ mm, we note that the diameter of the beam 2ρ is much smaller than the length of a single detector in this case (the length of a single detector for a 4×4 array is $\ell(A_m) = 0.5$ mm). Thus. $2\rho << \ell(A_m)$. 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 minimum values on the edges. This pattern is explained by our earlier understanding (as argued during elaboration of Fig. 9) that when the beam is very thin and its center resides at 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 $2\rho \approx \ell(A_m)$, 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 a change in the detected energy sensed 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 goes up as compared to the smaller beam radius scenario. This is because for a larger beam radius, the signal energy is spread out to a larger number of detectors which leads to a smaller signal-to-noise ratio per detector. 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 $\rho \to 0$ for the continuous array (see (19), and (21)).

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 4×4 detector and the 8×8 case is not significant.

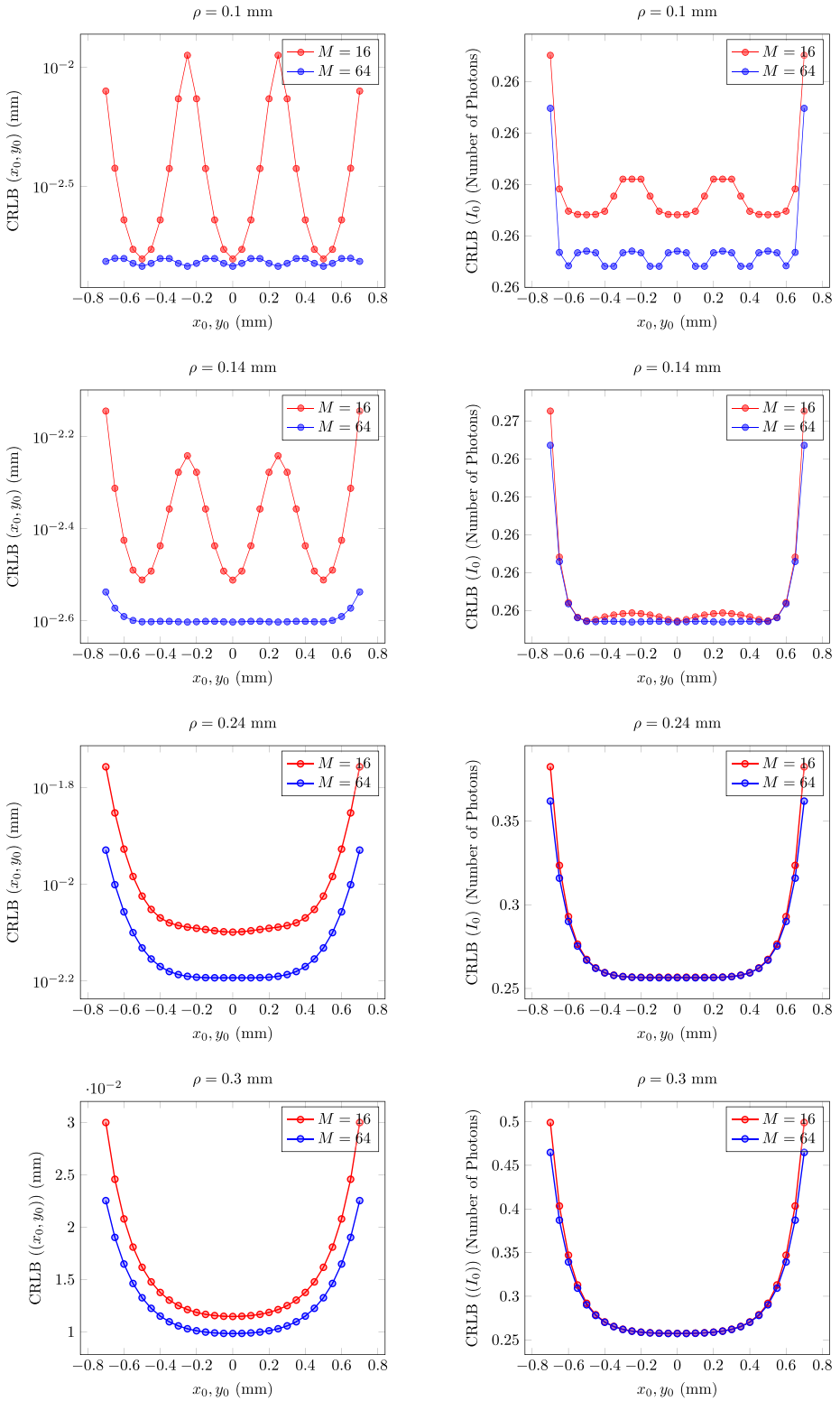

FIGURE 10. This figure shows the effect of beam radius on the Cramér-Rao lower bounds for 4 × 4 and 8 × 8 detector arrays. The value of noise parameter λ_n is 1 photon/slot period.

Fig. 11 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

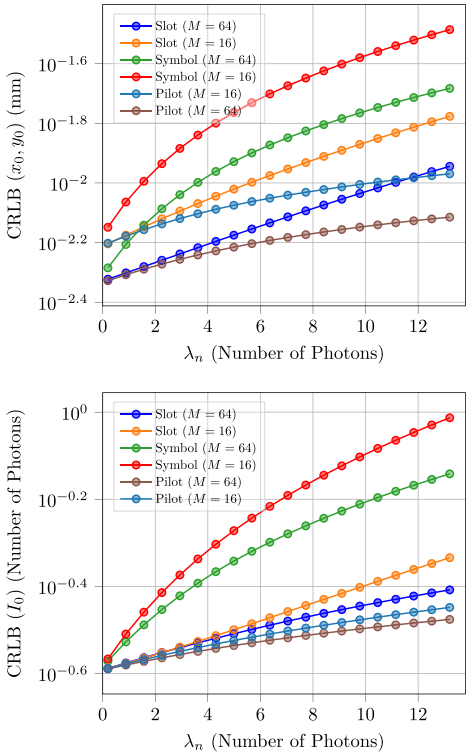

FIGURE 11. This figure shows the Cramèr-Rao lower bounds for three types of observations: pilot symbol, symbol period, and (signal) slot period. The beam radius is 0.2 mm.

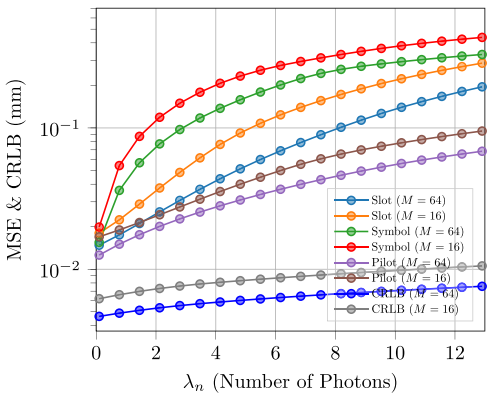

FIGURE 12. This figure shows the performance of the maximum likelihood estimator of (x_0, y_0) which is carried out for three different observations: pilot symbol, symbol period, and (signal) slot period. The beam radius is 0.2 mm. The performance of the maximum likelihood estimator is compared to the Cramér-Rao lower bound for the pilot symbol observations.

bound based on slot period performs better than the Cramér-Rao lower bound related to symbol period. Fig. 12 shows 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 Cramér-Rao lower bound curves in Fig. 11.

Fig. 13 and Fig. 14 depict the performance of the method of moment estimator of λ_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. 14 an interesting case where the mean-square error

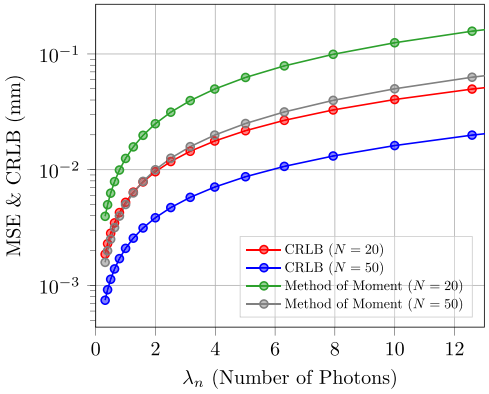

FIGURE 13. This figure shows the performance of the method of moments estimator of the noise parameter λ_n for a 4 × 4 array.

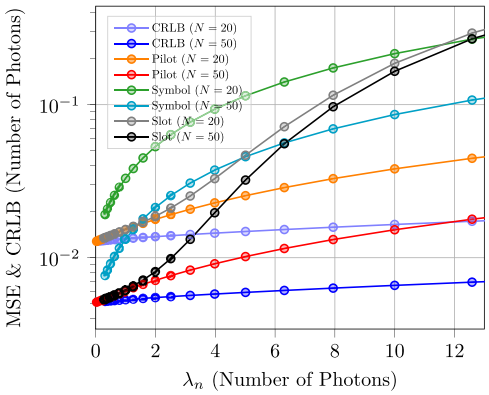

FIGURE 14. This figure shows the performance of the method of moments estimator of I_0 for the three different observations: pilot symbol, symbol period, and (signal) slot period. The beam radius is 0.2 mm. The performance of the method of moments estimator is compared with the Cramér-Rao lower bound of pilot symbol scheme as well. The configuration of the array is 4 × 4.

of the method of moments estimator based on slot period exceeds the mean-square error based on symbol period for large λ_n.

IX. A COMPLEXITY ANALYSIS OF ESTIMATORS

It is easy to see from (30) and (31), 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, firstly we need $\mathcal{O}(M)$ real additions and $\mathcal{O}(M)$ real multiplies (see (41)) 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 [4]. We note from [4] that the complexity of the genetic algorithm is a function of number of chromosomes, N_c, and the number of generations,[11] 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

11. 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.

Algorithm	Complexity
MLE for x_0, y_0, Method of Moments for I_0 and λ_n	$\mathcal{O}(\mathcal{A}^2M) + \mathcal{O}(\mathcal{A}^3)$ $+\mathcal{O}(M)$
MLE for x_0, y_0, I_0, Method of Moments for λ_n	$\mathcal{O}(\mathcal{A}^3M) + \mathcal{O}\left(\mathcal{A}^{\frac{9}{2}}\right)$ $+\mathcal{O}(M)$
MLE for x_0, y_0, I_0, λ_n	$\mathcal{O}(\mathcal{A}^4M) + \mathcal{O}\left(\mathcal{A}^6\right)$

FIGURE 15. Table comparing complexities for different scenarios.

convergence to 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. Therefore, we assume that $N_c = \mathcal{O}(\mathcal{A}^{\frac{D}{2}})$ and $N_g = \mathcal{O}(\mathcal{A}^{\frac{D}{2}})$, where D is the number of parameters (or dimensions) that we want to estimate with the maximum likelihood estimator.

Since the value of the loglikelihood function is computed for each chromosome, the total complexity of computing the likelihood function is approximately $N_c \times N_g \times \mathcal{O}(M)$ real multiplications and real additions. In addition to this, we have the complexity required for the comparison of fitness of the chromosomes during each iteration of the algorithm. Assuming that we use bubble sort, the we have to compute approximately N_c^2 comparisons during each generation in order to retain the chromosomes with the highest fitness. Thus, a total complexity of $N_c^2 N_g$ is also required in the computation of the maximum likelihood estimator.

Therefore, the total complexity of the maximum likelihood estimator turns out to be roughly $N_c \times N_g \times \mathcal{O}(M) + N_c^2 N_g$, which is $\mathcal{O}(\mathcal{A}^{\frac{D}{2}})\mathcal{O}(\mathcal{A}^{\frac{D}{2}})\mathcal{O}(M) + \mathcal{O}(\mathcal{A}^D)\mathcal{O}(\mathcal{A}^{\frac{D}{2}})$, which is equal to $\mathcal{O}(\mathcal{A}^D M) + \mathcal{O}(\mathcal{A}^{\frac{3D}{2}})$. Thus, for instance, if the maximum likelihood estimator is used in order to estimate only the beam position, then $D = 2$, and the computational complexity is roughly $\mathcal{O}(\mathcal{A}^2 M) + \mathcal{O}(\mathcal{A}^3)$.

Fig. 15 shows the complexity for three scenarios: i) when the maximum likelihood estimator is used to estimate (x_0, y_0) and method of moments estimator is used to estimate I_0 and λ_n (this is the two-step estimator); ii) when the maximum likelihood estimator is used to estimate (x_0, y_0, I_0) and method of moments estimator is used for λ_n; iii) and the final scenario in which the maximum likelihood estimator is used for the estimation of all four parameters $(x_0, y_0, I_0, \lambda_n)$.

In our simulations, we have set $N_c = 50$, and $N_g = 400$. The interested reader is referred to [24] for more details on genetic algorithms.

X. CONCLUSION AND FUTURE WORK

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 of the pulse position modulation scheme. As expected, estimation based on pilot symbols provide a superior mean-square error performance. However, estimation with data symbols does not require extra bandwidth and energy expenditure in terms of pilot symbols. 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.

In future, we will analyze the Cramer-Rao lower bounds for photon-counting detectors that suffer from dead-time/blocking, and a more comprehensive noise model that deals with location dependent background radiation as well.

APPENDIX

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

$$\begin{aligned}
&\frac{\partial \ln p(\mathbf{Z}|x_0, y_0)}{\partial x_0} \\
&= \sum_{m=1}^{M} \left(\frac{Z_m}{\Lambda_m} \iint_{A_m} \frac{I_0}{\rho^2} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} \frac{(x-x_0)}{\rho^2}\, dx\, dy \right. \\
&\qquad \left. - \iint_{\mathcal{A}} \frac{I_0}{\rho^2} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} \frac{(x-x_0)}{\rho^2}\, dx\, dy \right) \\
&\approx \sum_{m=1}^{M} \frac{Z_m}{\Lambda_m} \iint_{A_m} \frac{I_0}{\rho^4} (x-x_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy \qquad (42)
\end{aligned}$$

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

$$\iint_{\mathcal{A}} \frac{I_0}{\rho^2} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}} \frac{(x-x_0)}{\rho^2}\, dx\, dy \approx \kappa_0 (\mathbb{E}[X] - x_0) \qquad (43)$$

where κ_0 is a constant, and $\mathbb{E}[X] = x_0$. This approximation is valid when $\ell(\mathcal{A}) >> \rho$ and the beam resides well within the boundaries of the array. Additionally,

$$\begin{aligned}
&\frac{\partial^2 \ln p(\mathbf{Z}|x_0, y_0)}{\partial x_0^2} \\
&\approx \sum_{m=1}^{M} -\frac{Z_m}{\Lambda_m^2} \left(\iint_{A_m} \frac{I_0}{\rho^4} (x-x_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy \right)^2 \\
&\quad - \sum_{m=1}^{M} \frac{Z_m}{\Lambda_m} \iint_{A_m} \frac{I_0}{\rho^4} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy \\
&\quad + \sum_{m=1}^{M} \frac{Z_m}{\Lambda_m} \iint_{A_m} \frac{I_0}{\rho^6} (x-x_0)^2 e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy. \qquad (44)
\end{aligned}$$

Now, the expectation is taken with respect to Z_m:

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

$$
\begin{aligned}
&- \sum_{m=1}^{M} \iint_{A_m} \frac{I_0}{\rho^6}(x-x_0)^2 e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy \\
&= \sum_{m=1}^{M} \frac{1}{\Lambda_m}\left(\iint_{A_m} \frac{I_0}{\rho^4}(x-x_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy\right)^2 \\
&\quad + \iint_{\mathcal{A}} \frac{I_0}{\rho^4} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy \\
&\quad - \iint_{\mathcal{A}} \frac{I_0}{\rho^6}(x-x_0)^2 e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy \\
&= \sum_{m=1}^{M} \frac{1}{\Lambda_m}\left(\iint_{A_m} \frac{I_0}{\rho^4}(x-x_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy\right)^2 \\
&\quad + \frac{2\pi I_0}{\rho^2} - \frac{2\pi I_0}{\rho^2} \\
&= \sum_{m=1}^{M} \frac{1}{\Lambda_m}\left(\iint_{A_m} \frac{I_0}{\rho^4}(x-x_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy\right)^2.
\end{aligned} \tag{45}
$$

Similarly, it can be shown that

$$
\begin{aligned}
&-\mathbb{E}\left[\frac{\partial^2 \ln p(\boldsymbol{Z}|x_0, y_0)}{\partial y_0^2}\right] \\
&= \sum_{m=1}^{M} \frac{1}{\Lambda_m}\left(\iint_{A_m} \frac{I_0}{\rho^4}(y-y_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy\right)^2.
\end{aligned} \tag{46}
$$

Furthermore,

$$
\begin{aligned}
&\frac{\partial^2 \ln p(\boldsymbol{Z}|x_0, y_0)}{\partial x_0 \partial y_0} \\
&= \sum_{m=1}^{M}\left(-\frac{Z_m}{\Lambda_m^2}\iint_{A_m} \frac{I_0}{\rho^4}(y-y_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy \right. \\
&\qquad \left. \times \iint_{A_m} \frac{I_0}{\rho^4}(x-x_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy\right) \\
&\quad + \sum_{m=1}^{M} \frac{Z_m}{\Lambda_m} \iint_{A_m} \frac{I_0}{\rho^6}(x-x_0)(y-y_0) \\
&\quad \times e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy.
\end{aligned} \tag{47}
$$

Additionally,

$$
\begin{aligned}
&-\mathbb{E}\left[\frac{\partial^2 \ln p(\boldsymbol{Z}|x_0, y_0)}{\partial x_0 \partial y_0}\right] \\
&= \sum_{m=1}^{M}\left(\frac{1}{\Lambda_m}\iint_{A_m} \frac{I_0}{\rho^4}(y-y_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy \right. \\
&\qquad \left. \times \iint_{A_m} \frac{I_0}{\rho^4}(x-x_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy\right) \\
&\quad - \iint_{\mathcal{A}} \frac{I_0}{\rho^6}(x-x_0)(y-y_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy
\end{aligned} \tag{48}
$$

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

$$
\begin{aligned}
&\iint_{\mathcal{A}} \frac{I_0}{\rho^6}(x-x_0)(y-y_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy \\
&\quad \approx \kappa_1\big[(\mathbb{E}[X]-x_0)(\mathbb{E}[Y]-y_0)\big]
\end{aligned} \tag{49}
$$

where $\mathbb{E}[X] = x_0$ and $\mathbb{E}[Y] = y_0$ for some constant κ_1. This approximation is valid if $\ell(\mathcal{A}) >> \rho$ and the beam resides well within the boundaries of the array. Therefore,

$$
\begin{aligned}
&-\mathbb{E}\left[\frac{\partial^2 \ln p(\boldsymbol{Z}|x_0, y_0)}{\partial x_0 \partial y_0}\right] \\
&= \sum_{m=1}^{M}\left(\frac{1}{\Lambda_m}\iint_{A_m} \frac{I_0}{\rho^4}(y-y_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy \right. \\
&\qquad \left. \times \iint_{A_m} \frac{I_0}{\rho^4}(x-x_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy\right) \\
&= -\mathbb{E}\left[\frac{\partial^2 \ln p(\boldsymbol{Z}|x_0, y_0)}{\partial y_0 \partial x_0}\right].
\end{aligned} \tag{50}
$$

Moreover, the *Fisher Information Matrix* is

$$
\mathcal{I}(x_0, y_0) = \begin{bmatrix} -\mathbb{E}\left[\frac{\partial^2 \ln p(\boldsymbol{Z}|x_0,y_0)}{\partial x_0^2}\right] & -\mathbb{E}\left[\frac{\partial^2 \ln p(\boldsymbol{Z}|x_0,y_0)}{\partial x_0 \partial y_0}\right] \\ -\mathbb{E}\left[\frac{\partial^2 \ln p(\boldsymbol{Z}|x_0,y_0)}{\partial y_0 \partial x_0}\right] & -\mathbb{E}\left[\frac{\partial^2 \ln p(\boldsymbol{Z}|x_0,y_0)}{\partial y_0^2}\right] \end{bmatrix}, \tag{51}
$$

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

$$
\mathsf{Var}[\hat{x}_0] \geq \frac{\sum_{m=1}^{M} \frac{1}{\Lambda_m}\left(\iint_{A_m} \frac{I_0}{\rho^4}(y-y_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy\right)^2}{\Psi(x_0, y_0, I_0, \rho)} \tag{52}
$$

$$
\mathsf{Var}[\hat{y}_0] \geq \frac{\sum_{m=1}^{M} \frac{1}{\Lambda_m}\left(\iint_{A_m} \frac{I_0}{\rho^4}(x-x_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy\right)^2}{\Psi(x_0, y_0, I_0, \rho)}, \tag{53}
$$

where

$$
\begin{aligned}
&\Psi(x_0, y_0, I_0, \rho) \\
&\triangleq \sum_{m=1}^{M} \frac{1}{\Lambda_m}\left(\iint_{A_m} \frac{I_0}{\rho^4}(x-x_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy\right)^2 \\
&\quad \times \sum_{m=1}^{M} \frac{1}{\Lambda_m}\left(\iint_{A_m} \frac{I_0}{\rho^4}(y-y_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy\right)^2 \\
&\quad - \left(\sum_{m=1}^{M} \frac{1}{\Lambda_m}\iint_{A_m} \frac{I_0}{\rho^4}(y-y_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy \right. \\
&\qquad \left. \times \iint_{A_m} \frac{I_0}{\rho^4}(x-x_0) e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\rho^2}}\, dx\, dy\right)^2.
\end{aligned} \tag{54}
$$

REFERENCES

[1] H. Hemmati, A. Biswas, and I. B. Djordjevic, "Deep-space optical communications: Future perspectives and applications," *Proc. IEEE*, vol. 99, no. 11, pp. 2020–2039, Nov. 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 Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 1104–1123, 2nd Quart., 2018.

[3] B. J. Slocumb and D. L. Snyder. (1990). *Maximum Likelihood Estimation Applied to Quantum-Limited Optical Position-Sensing*. [Online]. Available: https://doi.org/10.1117/12.2322208

[4] M. S. Bashir and M. R. Bell, "Optical beam position estimation in free-space optical communication," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 6, pp. 2896–2905, Dec. 2016.

[5] M. S. Bashir and M. R. Bell, "Optical beam position tracking in free-space optical communication systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 2, pp. 520–536, Apr. 2018.

[6] 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 Trans. Aerosp. Electron. Syst.*, vol. 55, no. 3, pp. 1319–1333, Jun. 2019.

[7] M. S. Bashir, "Free-space optical communications with detector arrays: A mathematical analysis," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 2, pp. 1420–1429, Apr. 2020.

[8] M. S. Bashir and M.-S. Alouini, "Signal acquisition with photon-counting detector arrays in free-space optical communications," *IEEE Trans. Wireless Commun.*, vol. 19, no. 4, pp. 2181–2195, Apr. 2020.

[9] M. S. Bashir and S. S. Muhammad, "Time synchronization in photon-limited deep space optical communications," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 1, pp. 30–40, Feb. 2020.

[10] A. A. Farid and S. Hranilovic, "Outage capacity optimization for free-space optical links with pointing errors," *J. Lightw. Technol.*, vol. 25, no. 7, pp. 1702–1710, Jul. 2007.

[11] V. V. Mai and H. Kim, "Adaptive beam control techniques for airborne free-space optical communication systems," *Appl. Opt.*, vol. 57, no. 26, pp. 7462–7471, Sep. 2018.

[12] I. Ansari, F. Yilmaz, and M.-S. Alouini, "Performance analysis of free-space optical links over Malaga (M) turbulence channels with pointing errors," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 91–102, Jan. 2016.

[13] E. Zedini, H. Soury, and M.-S. Alouini, "Dual-hop FSO transmission systems over gamma-gamma turbulence with pointing errors," *IEEE Trans. Wireless Commun.*, vol. 16, no. 2, pp. 784–796, Feb. 2017.

[14] 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 Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6502–6512, Sep. 2016.

[15] C. B. Issaid, K.-H. 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 Trans. Wireless Commun.*, vol. 16, no. 10, pp. 6822–6837, Oct. 2017.

[16] V. A. Vilnrotter and M. Srinivasan, "Adaptive detector arrays for optical communications receivers," *IEEE Trans. Commun.*, vol. 50, no. 7, pp. 1091–1097, Jul. 2002.

[17] V. Vilnrotter, C. Lau, M. Srinivasan, K. Andrews, and R. Mukai, "Optical array receiver for communication through atmospheric turbulence," *J. Lightw. Technol.*, vol. 23, no. 4, pp. 1664–1675, Apr. 2005.

[18] 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*, vol. 9739, H. Hemmati and D. M. Boroson, Eds. Bellingham, WA, USA: Int. Soc. Opt. Photon., 2016, pp. 267–282. [Online]. Available: https://doi.org/10.1117/12.2217971

[19] 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*, vol. 10524, H. Hemmati and D. M. Boroson, Eds., Bellingham, WA, USA: Int. Soc. Opt. Photon., 2018, pp. 47–59. [Online]. Available: https://doi.org/10.1117/12.2296740

[20] *APD Arrays: Geiger-Mode APD Arrays Detect Low Light*. Accessed: Apr. 6, 2020. [Online]. Available: https://www.laserfocusworld.com/detectors-imaging/article/16555170/apd-arrays-geigermode-apd-arrays-detect-low-light

[21] 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 Syst. Mag.*, vol. 37, no. 5, pp. 70–89, Oct. 2017.

[22] D. L. Snyder and M. I. Miller, *Random Point Processes in Time and Space*. New York, NY, USA: Springer, 1991.

[23] M. S. Bashir and M.-S. Alouini, "Beam tracking with photon-counting detector arrays in free-space optical communications," 2020. [Online]. Available: https://arxiv.org/pdf/2001.04007.pdf

[24] S. S. Rao, *Engineering Optimization Theory and Practice*. Hoboken, NJ, USA: Wiley, 2009.

MUHAMMAD SALMAN BASHIR (Senior Member, IEEE) received the M.S. and Ph.D. degrees in electrical and computer engineering from Purdue University, West Lafayette, IN, USA, in 2014 and 2017, respectively. He served as a Faculty Member with the National University of Computer and Emerging Sciences Lahore, before joining the Communication Theory Lab, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia, as a Postdoctoral Fellow in 2019. His current research interests are in the area of statistical signal processing for wireless communications. He was the recipient of the International Fulbright Science and Technology Award for his graduate studies at Purdue University.

MING-CHENG TSAI (Graduate Student Member, IEEE) was born in Fujian, China. He received the B.E. degree in electrical engineering from the National Taipei University of Technology, Taipei, Taiwan, in 2015. From 2015 to 2018, he was a student of Communication Engineering with National Tsinghua University. He is currently pursuing the M.S./Ph.D. degree in electrical engineering with the King Abdullah University of Science and Technology. His research interests include device-to-device communication, digital communication, and error correcting codes.

MOHAMED-SLIM ALOUINI (Fellow, IEEE) was born in Tunis, Tunisia. He received the Ph.D. degree in electrical engineering from the California Institute of Technology in 1998. He served as a Faculty Member with the University of Minnesota then with Texas A&M University at Qatar before joining in 2009, the King Abdullah University of Science and Technology, where he is currently a Distinguished Professor of Electrical and Computer Engineering. He is currently particularly interested in addressing the technical challenges associated with the uneven distribution, access to, and use of information and communication technologies in far-flung, rural, low-density populations, low-income, and/or hard-to-reach areas. He is a Fellow of OSA.