

Chao, J.; Ward, E.S.; Ober, R.J., "Fisher information for EMCCD imaging with application to single molecule microscopy," in *Signals, Systems and Computers (ASILOMAR), 2010 Conference Record of the Forty Fourth Asilomar Conference on*, vol., no., pp.1085-1089, 7-10 Nov. 2010

doi: 10.1109/ACSSC.2010.5757570

keywords: {CCD image sensors;charge-coupled devices;electron multipliers;Cramer-Rao lower bound;EMCCD imaging;electron-multiplying charge-coupled device;single molecule microscopy;Accuracy;Charge coupled devices;Data models;Noise;Photonics;Pixel;Random variables;Branching process;Fisher information;electron multiplication;single molecule microscopy},

URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5757570&isnumber=5757211>

FISHER INFORMATION FOR EMCCD IMAGING WITH APPLICATION TO SINGLE MOLECULE MICROSCOPY

Jerry Chao^{1,2}, E. Sally Ward², and Raimund J. Ober^{1,2,}*

¹Department of Electrical Engineering, University of Texas at Dallas, Richardson, TX,

²Department of Immunology, University of Texas Southwestern Medical Center, Dallas, TX.

ABSTRACT

Owing to its high quantum efficiency, the charge-coupled device (CCD) is an important imaging tool employed in biological applications such as single molecule microscopy. Under extremely low light conditions, however, a CCD is generally unsuitable because its readout noise can easily overwhelm the weak signal. Instead, an electron-multiplying charge-coupled device (EMCCD), which stochastically amplifies the acquired signal to drown out the readout noise, can be used. We have previously proposed a framework for calculating the Fisher information, and hence the Cramer-Rao lower bound, for estimating parameters (e.g., single molecule location) from the images produced by an optical microscope. Here, we develop the theory that is needed for deriving, within this framework, performance measures pertaining to the estimation of parameters from an EMCCD image. Our results allow the comparison of a CCD and an EMCCD in terms of the best accuracy with which parameters can be estimated from their acquired images.

Index Terms— Branching process, electron multiplication, Fisher information, single molecule microscopy

1. INTRODUCTION

The charge-coupled device (CCD) is an important image detector that has found utility in diverse areas such as high sensitivity cellular microscopy and astronomy. The high quantum efficiency of a typical CCD enables it to detect a large fraction of the photons which impact its surface. However, a CCD is unsuitable for imaging under extremely low light conditions, since its measurement noise can overwhelm the signal when relatively few photons are detected from the imaged object. Measurement noise is added when the signal is read out from the CCD, and is commonly called the readout noise (e.g., [1]).

An intended solution for overcoming the readout noise under low light conditions is the electron-multiplying charge-coupled device (EMCCD). An EMCCD is similar to a CCD in that it accumulates electrons in proportion to the number

of photons it detects. Unlike a CCD, however, it can increase the number of electrons substantially via a multiplication process, thereby generating an amplified signal that can effectively drown out the readout noise. The multiplication occurs as electrons are transferred through a gain register consisting of typically several hundred stages. Specifically, each input electron to any given stage can, with certain probabilities, generate secondary electrons that are transferred, along with the input electron itself, to the next stage for further amplification. Given such a cascade, a large number of electrons can be produced at the output of the gain register per initial electron, even with the usually small probabilities for secondary electron generation (typically 0.01 to 0.02 [2] for one secondary electron per input electron per stage, and even smaller for more than one secondary electron per input electron).

From an image captured using a CCD or an EMCCD camera, parameters of interest can be estimated to obtain useful information about the imaged object. In single molecule microscopy (e.g., [3]), for example, an important problem has been the estimation of the location of a fluorescent molecule (e.g., [4, 5]). In this context, a general framework [6] has been proposed for calculating the Fisher information, and hence the Cramer-Rao lower bound [7], for the estimation of parameters from an image produced by a microscope. Using this framework, accuracy limits have been derived for estimating, for example, the location coordinates of a single molecule (e.g., [5]). These performance measures, however, have assumed the image to have been acquired with a CCD.

In this paper, we develop the theory that is necessary for deriving performance measures for estimating parameters from an EMCCD image. To arrive at the Fisher information for EMCCD data, an expression is needed for the probability distribution of the electron count that results from the multiplication process described above. To this end, we model the stochastic multiplication as a branching process (e.g., [8]), as others have done for EMCCDs (e.g., [9, 2]). However, as opposed to its typical description as a Bernoulli event, we describe the generation of secondary electrons using a geometric model of multiplication.

In addition to the derivation of a Fisher information expression for EMCCD data, we introduce the notion of a “noise coefficient” which enables the comparison of the Fisher infor-

This work was supported in part by the National Institutes of Health (R01 GM071048 and R01 GM085575). *Corresponding author, email: ober@utdallas.edu.

mation for different data models via a scalar quantity. Using the noise coefficient, we compare the Fisher information for CCD and EMCCD data as a function of the expected signal level. This is motivated by the fact that electron multiplication is a random process that introduces stochasticity of its own to the data, and hence should only be used when relatively low signal levels are expected. Comparison using the noise coefficient importantly allows a quantitative determination, based on the expected signal level, of the choice between the CCD and the EMCCD from the perspective of Fisher information.

The material presented in this paper represents an important subset of the content of [10]. In Section 2, we present a general result from which the Fisher information expressions for all data models considered in this paper can be derived. Based on this result, we also define the noise coefficient. In Section 3, a Fisher information expression is presented for data that can be described as the output of a branching process with a geometric model of multiplication, with added readout noise. By comparing its corresponding noise coefficient with the noise coefficient for CCD data, we examine the usefulness of multiplication as a function of the expected signal level. In Section 4, the theory developed in previous sections for a single signal is generalized for a collection of independent signals which comprise a CCD or an EMCCD image. An example is given which applies the generalized theory to the localization of a single molecule from an image.

2. THE NOISE COEFFICIENT

The first part of this paper deals with the analysis of the information content of a scalar random variable that models the data in a single pixel of an electron-multiplying image detector. Specifically, the signal that impinges upon the detector is modeled as a Poisson random variable, since photon emission (e.g., by a fluorescent molecule), and hence the detection of those photons by a camera, are typically modeled as Poisson processes. However, the data in the pixel is a readout noise-corrupted version of the signal that may have been amplified with the intention to drown out the added noise.

We are interested in calculating the Fisher information matrix pertaining to parameter estimation problems such as the localization of a single molecule from its image. These estimation problems take on a typical form. The probability distribution of the incident Poisson signal is parameterized by its mean ν . However, the mean ν itself is a function of the parameter vector θ that is of interest (e.g., the location coordinates of a single molecule). We first give an expression for the Fisher information of a random variable using this specific parameterization (see [10] for the proof).

Theorem 2.1 *Let Z_ν be a continuous (discrete) random variable with probability density (mass) function p_ν , where ν is a scalar parameter. Let $\nu = \nu_\theta$ be a reparameterization of p_ν through the possibly vector-valued parameter $\theta \in \Theta$, where*

Θ is the parameter space. We use the notation $Z_\theta(p_\theta)$ to denote $Z_{\nu_\theta}(p_{\nu_\theta})$, $\theta \in \Theta$. Then the Fisher information matrix $\mathbf{I}(\theta)$ of Z_θ with respect to θ is given by

$$\mathbf{I}(\theta) = \left(\frac{\partial \nu_\theta}{\partial \theta} \right)^T \frac{\partial \nu_\theta}{\partial \theta} \cdot E \left[\left(\frac{\partial}{\partial \nu_\theta} \ln(p_\theta(z)) \right)^2 \right]. \quad (1)$$

Note that the scalar expectation term in Eq. 1 is just the Fisher information of Z_θ with respect to ν_θ .

Two corollaries follow from Theorem 2.1 which pertain to data acquired under two important scenarios. The first corollary (see [10] for the proof) gives the Fisher information matrix for the ideal scenario where a Poisson-distributed number of electrons are read out from a device without being corrupted by readout noise. This scenario represents the best case wherein a CCD is able to read out a signal without introducing readout noise, and will serve as the benchmark against which practical scenarios are compared. From here onwards, the function ν_θ will represent the mean of the Poisson signal.

Corollary 2.1 *Let Z_θ be a Poisson random variable with mean $\nu_\theta > 0$, such that its probability mass function is $p_{\theta,P}(z) = \frac{e^{-\nu_\theta} \nu_\theta^z}{z!}$, $z = 0, 1, \dots$. The Fisher information matrix $\mathbf{I}_P(\theta)$ of Z_θ is given by*

$$\mathbf{I}_P(\theta) = \left(\frac{\partial \nu_\theta}{\partial \theta} \right)^T \frac{\partial \nu_\theta}{\partial \theta} \cdot \frac{1}{\nu_\theta}. \quad (2)$$

The second corollary (see [10] for the proof) gives the Fisher information matrix for the practical scenario where readout noise is added to a Poisson-distributed number of electrons when they are read out from a device. Readout noise is typically modeled as a Gaussian random variable, and is assumed to be such by this corollary. This scenario corresponds to the practical operation of a CCD.

Corollary 2.2 *Let $Z_\theta = V_\theta + W$, where V_θ is a Poisson random variable with mean $\nu_\theta > 0$, and W is a Gaussian random variable with mean η_w and variance σ_w^2 . Let V_θ and W be stochastically independent of each other, and let W be not dependent on θ . The probability density function of Z_θ is then the convolution of the Poisson probability mass function with mean ν_θ and the Gaussian probability density function with mean η_w and variance σ_w^2 , given by $p_{\theta,R}(z) = \frac{1}{\sqrt{2\pi}\sigma_w} \sum_{j=0}^{\infty} \frac{e^{-\nu_\theta} \nu_\theta^j}{j!} e^{-\frac{1}{2} \left(\frac{z-j-\eta_w}{\sigma_w} \right)^2}$, $z \in \mathbb{R}$. The Fisher information matrix $\mathbf{I}_R(\theta)$ of Z_θ is given by*

$$\mathbf{I}_R(\theta) = \left(\frac{\partial \nu_\theta}{\partial \theta} \right)^T \frac{\partial \nu_\theta}{\partial \theta} \cdot \left(\int_{-\infty}^{\infty} \frac{1}{p_{\theta,R}(z)} \left(\frac{e^{-\nu_\theta}}{\sqrt{2\pi}\sigma_w} \sum_{j=1}^{\infty} \frac{\nu_\theta^{j-1} e^{-\frac{1}{2} \left(\frac{z-j-\eta_w}{\sigma_w} \right)^2}}{(j-1)!} \right)^2 dz - 1 \right). \quad (3)$$

Note that $p_{\theta,R}$ in the above corollary can be found in [1].

By Theorem 2.1 and demonstrated by Corollaries 2.1 and 2.2, different distributions of the random variable Z_θ will yield Fisher information matrices that differ from one another via the Fisher information of Z_θ with respect to ν_θ (i.e., via

the scalar expectation term of Eq. 1). Hence, we introduce, for the purpose of comparing the Fisher information of different data models, a “noise coefficient” based on this quantity. Since Eq. 2 represents the best case scenario (i.e., a Poisson signal that is not corrupted by readout noise), we take its Fisher information with respect to ν_θ (i.e., the quantity $\frac{1}{\nu_\theta}$) as the reference, and define the noise coefficient as follows.

Definition 2.1 Let Z_θ be a continuous (discrete) random variable with probability density (mass) function p_θ . Let p_θ be parameterized by θ through the mean $\nu_\theta > 0$ of a Poisson-distributed random variable. Then the noise coefficient (with respect to ν_θ) of Z_θ , denoted by α , is given by

$$\alpha = \nu_\theta \cdot E \left[\left(\frac{\partial}{\partial \nu_\theta} \ln(p_\theta(z)) \right)^2 \right]. \quad (4)$$

The noise coefficient of Eq. 4 is just the ratio of the Fisher information of Z_θ to that of the ideal, uncorrupted Poisson signal, both with respect to ν_θ . Using this quantity, the Fisher information matrix $\mathbf{I}(\theta)$ of a random variable which satisfies the conditions of Definition 2.1 can be expressed as $\mathbf{I}(\theta) = \alpha \cdot \mathbf{I}_P(\theta)$, where $\mathbf{I}_P(\theta)$ is the matrix of Eq. 2.

For the ideal scenario of Corollary 2.1 where the data is Poisson-distributed, the noise coefficient is trivially $\alpha_P = 1$. For the practical scenario of Corollary 2.2 where Gaussian-distributed readout noise is added to the Poisson signal, the noise coefficient α_R is just

$$\alpha_R = \left(\int_{-\infty}^{\infty} \frac{\nu_\theta}{p_{\theta,R}(z)} \left(\frac{e^{-\nu_\theta}}{\sqrt{2\pi}\sigma_w} \sum_{j=1}^{\infty} \frac{\nu_\theta^{j-1} e^{-\frac{1}{2} \left(\frac{z-j-\eta_w}{\sigma_w} \right)^2}}{(j-1)!} \right)^2 dz - \nu_\theta \right) \quad (5)$$

and the Fisher information matrix $\mathbf{I}_R(\theta)$ can be expressed as $\mathbf{I}_R(\theta) = \alpha_R \cdot \mathbf{I}_P(\theta)$.

The results presented so far, as well as the result to be presented in Section 3, involve data that can be described as a Poisson signal with mean ν_θ that may have been stochastically multiplied (i.e., amplified) by some random function M before potentially being corrupted by some additive readout noise W . Since neither the stochasticity introduced by the multiplication nor the readout noise is dependent on θ , they contribute no additional information about θ . Therefore, the noise coefficient α for these data models can be expected to be at most 1 (i.e., at most α_P). We state this result formally in the following theorem (see [10] for the proof).

Theorem 2.2 Let Θ be a parameter space and let $Z_\theta = M(V_\theta) + W$, $\theta \in \Theta$, where V_θ is a Poisson random variable with mean $\nu_\theta > 0$, M is a random function, and W is a scalar-valued random variable. We assume that V_θ , M , and W are stochastically independent. Then, for the noise coefficient α of Z_θ , $0 \leq \alpha \leq 1$.

3. GEOMETRIC SIGNAL MULTIPLICATION

As mentioned in Section 1, we model electron multiplication as a branching process (e.g., [8]) that is geometrically multiplied. Specifically, an initial Poisson-distributed number of

electrons are fed into a series of stages where, in each stage, an input electron can generate secondary electrons, such that the probability of obtaining k electrons per input electron per stage (including the input electron itself) is given by the zero modified geometric distribution [11] defined as follows.

Definition 3.1 A zero modified geometric distribution is a probability distribution $(p_k)_{k=0,1,\dots}$ given by

$$p_k = \begin{cases} a, & k = 0, \\ (1-a)(1-b)b^{k-1}, & k = 1, 2, \dots, \end{cases} \quad (6)$$

where $0 \leq a < 1$ and $0 \leq b < 1$.

The zero modified geometric distribution has mean $m = \frac{1-a}{1-b}$ and variance $\sigma^2 = \frac{(1-a)(b+a)}{(1-b)^2}$, and has a probability generating function of linear fractional form. Due to this special property, the probability distribution of the number of electrons $X_{N,\theta}$ at the output of an N -stage branching process with a zero modified geometric model of multiplication can be expressed explicitly without recursion (e.g., [8]). The zero modified geometric distribution also importantly allows the possibility of having zero or more electrons per input electron per stage (including the input electron itself). This makes it suitable for modeling electron multiplication in an EMCCD, where more than one secondary electron can be generated per input electron per stage [9], and where electron loss mechanisms may exist. Note that by setting the parameter $a = 0$, the zero modified geometric distribution of Eq. 6 reduces to the standard geometric distribution $p_k = (1-b)b^{k-1}$, $k = 1, 2, \dots$, with mean $m = \frac{1}{1-b}$ and variance $\sigma^2 = \frac{b}{(1-b)^2}$. While the more general zero modified geometric distribution will be used in the theorem that follows, the standard geometric distribution will be used for all subsequent illustrations.

In [10], a probability mass function was derived, using the theory of probability generating functions, for the number of electrons $X_{N,\theta}$ at the output of an N -stage branching process with an initial Poisson-distributed number of electrons and a zero modified geometric model of multiplication. Since the data Z_θ in a given pixel of an EMCCD is modeled as the sum of $X_{N,\theta}$ and a Gaussian random variable W representing the readout noise, the probability density function of Z_θ is just the convolution of the mass function of $X_{N,\theta}$ and a Gaussian density function. The following Theorem gives this density function and the corresponding Fisher information.

Theorem 3.1 Let $Z_\theta = X_{N,\theta} + W$, where $X_{N,\theta}$, $N \in \{0, 1, \dots\}$, is the number of particles at the output of an N -stage branching process with an initial Poisson-distributed particle count with mean $\nu_\theta > 0$ and a zero modified geometric model of multiplication, and W is a Gaussian random variable with mean η_w and variance σ_w^2 . Let $X_{N,\theta}$ and W be stochastically independent, and let W be not dependent on θ . 1. The probability density function of Z_θ is, for $z \in \mathbb{R}$,

$$p_{\theta,GeomR}(z) = \frac{e^{-\nu_\theta} \frac{A}{B}}{\sqrt{2\pi}\sigma_w}.$$

$$\left[e^{-\left(\frac{z-\eta_w}{\sqrt{2}\sigma_w}\right)^2} + \sum_{l=1}^{\infty} e^{-\left(\frac{z-l-\eta_w}{\sqrt{2}\sigma_w}\right)^2} \sum_{j=0}^{l-1} \frac{\binom{l-1}{j} C^{l-1-j} (D\nu_\theta)^{j+1}}{(j+1)! B^{j+l+1}} \right], \quad (7)$$

where $A = (1-a)(m-1)m^N$, $B = b(m^N-1)m + (1-a)(m-1)$, $C = b(m^N-1)m$, $D = m^N(1-a)^2(m-1)^2$, $m = \frac{1-a}{1-b} \neq 1$, and $\binom{l-1}{j}$ denotes “ $l-1$ choose j ”.

2. The noise coefficient corresponding to $p_{\theta, \text{GeomR}}$ is

$$\alpha_{\text{GeomR}} = -\nu_\theta \frac{A^2}{B^2} + \int_{-\infty}^{\infty} \frac{\nu_\theta D^2 e^{-2\nu_\theta \frac{A}{B}}}{p_{\theta, \text{GeomR}}(z)} \left(\sum_{l=1}^{\infty} \frac{e^{-\left(\frac{z-l-\eta_w}{\sqrt{2}\sigma_w}\right)^2}}{\sqrt{2\pi}\sigma_w} \sum_{j=0}^{l-1} \frac{\binom{l-1}{j} C^{l-1-j} (D\nu_\theta)^{-j}}{j! B^{j+l+1}} \right)^2 dz. \quad (8)$$

The Fisher information matrix of Z_θ is $\mathbf{I}_{\text{GeomR}}(\theta) = \alpha_{\text{GeomR}} \cdot \mathbf{I}_P(\theta)$, with $\mathbf{I}_P(\theta)$ as given in Eq. 2.

The term m^N in Eqs. 7 and 8 is called the *mean gain*, and is the average number of particles (i.e., electrons) at the output of the multiplication process given a single initial particle.

To demonstrate a comparison of the Fisher information of different data models using the noise coefficient, Fig. 1 plots α_{GeomR} of Eq. 8 (with $a = 0$ for standard geometric multiplication) for different mean gain values, and α_R of Eq. 5, as a function of the mean ν_θ of the initial electron count. As per Theorem 2.2, the plot shows that α_{GeomR} and α_R have values between 0 and 1 regardless of the value of ν_θ . The Fisher information matrices $\mathbf{I}_{\text{GeomR}}(\theta)$ and $\mathbf{I}_R(\theta)$ are hence no greater than $\mathbf{I}_P(\theta)$ of the ideal scenario of Corollary 2.1.

Fig. 1 further shows, for the settings specified therein, that α_{GeomR} is greater than α_R for ν_θ values of up to roughly 60 electrons. In this range of ν_θ values, a higher mean gain generally yields a larger α_{GeomR} . Beyond roughly $\nu_\theta = 60$ electrons, however, α_{GeomR} starts to drop below α_R in order of decreasing mean gain. By roughly $\nu_\theta = 130$ electrons, multiplication with any of the given mean gain values produces an α_{GeomR} that is less than α_R . Fig. 1 thus demonstrates that multiplication is beneficial only when the expected signal level is relatively small (or equivalently, when the readout noise level is relatively significant). Importantly, this is demonstrated quantitatively as a function of the expected signal level, and from the perspective of Fisher information.

4. GENERALIZATION TO AN IMAGE

The theory of the previous sections applies to a single pixel of a CCD-based detector. However, by assuming the data in different pixels of an image to be independent measurements, the Fisher information matrix for an image is just the sum of the Fisher information matrices for its pixels. For an image of K pixels, its Fisher information matrix can thus be written as $\mathbf{I}_{im}(\theta) = \sum_{k=1}^K \mathbf{I}_k(\theta) = \sum_{k=1}^K \alpha_k \cdot \mathbf{I}_{P,k}(\theta)$, where the notation is as before and the subscript k denotes quantity for the k^{th} pixel. It follows that for an ideal K -pixel image of uncorrupted Poisson signals, (i.e., $\alpha_k = 1$ for $k = 1, \dots, K$), its Fisher information matrix is just $\mathbf{I}_{im,P}(\theta) = \sum_{k=1}^K \mathbf{I}_{P,k}(\theta)$.

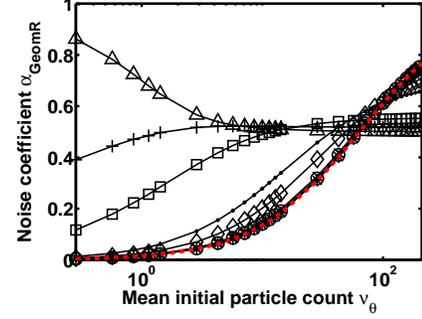


Fig. 1. Noise coefficient α_{GeomR} (Eq. 8 with $a = 0$), for the scenario of a Poisson-distributed signal that is amplified by multiplication and subsequently corrupted by readout noise. The noise coefficient is shown as a function of the mean ν_θ of the signal, which ranges in value from 0.29 to 199.70. The signal is amplified through $N = 536$ stages (as in the gain register of a CCD97 chip, E2V Technologies, Chelmsford, UK) of standard geometric multiplication, and the different curves correspond to mean gain values of $m^N = 1.01$ (*), 1.03 (o), 1.06 (x), 1.31 (◇), 1.71 (-), 4.98 (□), 14.49 (+), and 1015.46 (△). The readout noise is Gaussian with mean $\eta_w = 0$ and standard deviation $\sigma_w = 8$, and as a reference, the red curve shows the noise coefficient α_R (Eq. 5) for the scenario where there is no signal amplification.

Using these expressions, we next give an inequality (see [10] for the proof) which relates $\mathbf{I}_{im}(\theta)$ for a practical image to $\mathbf{I}_{im,P}(\theta)$ for its corresponding ideal image.

Theorem 4.1 Let $\mathbf{I}_{im}(\theta) = \sum_{k=1}^K \alpha_k \cdot \mathbf{I}_{P,k}(\theta)$, and let $\mathbf{I}_{im,P}(\theta) = \sum_{k=1}^K \mathbf{I}_{P,k}(\theta)$. Let α_{min} and α_{max} denote, respectively, the smallest and the largest elements in the sequence $(\alpha_k)_{k=1, \dots, K}$. Then we have

$$\alpha_{min} \cdot \mathbf{I}_{im,P}(\theta) \leq \mathbf{I}_{im}(\theta) \leq \alpha_{max} \cdot \mathbf{I}_{im,P}(\theta).$$

Theorem 4.1 can be used to assess, in terms of the Fisher information, how close a practical image is to its corresponding ideal image. It will be used in the example that follows.

To conclude this paper, we apply our theory to the localization of a fluorescent molecule. We consider the estimation of the location of an in-focus point source (i.e., single molecule) from its image as observed through a fluorescence microscope and detected by a CCD or an EMCCD camera. For this problem, the mean of the Poisson-distributed electron count at the k^{th} pixel of the device due to the photons detected from the point source can be shown to be [6]

$$\nu_{\theta,k} = \frac{N_{photon}}{M^2} \int_{C_k} q(x/M - x_0, y/M - y_0) dx dy, \quad (9)$$

where N_{photon} is the expected number of photons detected from the point source, M is the magnification of the microscope, C_k is the region in the xy -plane occupied by the pixel, x_0 and y_0 are the x and y coordinates of the point source in the object space where it resides, and q is the classical Airy point spread function (see [10] for the definition) which describes the image formed from the detected photons.

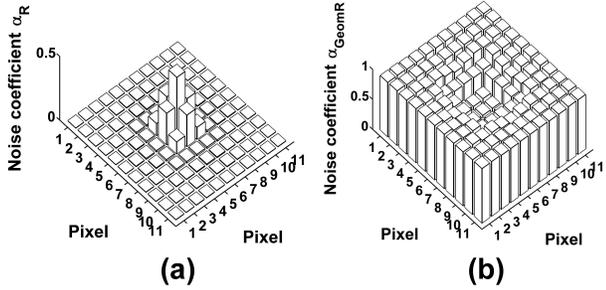


Fig. 2. Noise coefficient (α_R for (a), α_{GeomR} for (b)) profile for (a) a CCD image and (b) an EMCCD image of an in-focus point source. The point source is assumed to emit photons of wavelength $\lambda = 680$ nm, which are collected by an objective lens with magnification $M = 100$ and numerical aperture $n_a = 1.4$. The image of the point source is given by the Airy point spread function, and is centered on an 11-by-11 array of $16 \mu\text{m}$ by $16 \mu\text{m}$ pixels (i.e., $x_0 = y_0 = 880$ nm, assuming the upper left corner of the pixel array is $(0, 0)$). The expected number of detected photons is set to $N_{photon} = 200$. In (a), readout noise with mean $\eta_w = 0 e^-$ and standard deviation $\sigma_w = 8 e^-$ is assumed for every pixel. In (b), the standard deviation is higher at $\sigma_w = 24 e^-$, and standard geometric multiplication with a mean gain of $m^{536} = 1015.46$ is assumed.

Noise coefficients are computed for an 11-by-11 pixel array (i.e., image), with the mean initial electron count $\nu_{\theta,k}$ (Eq. 9) at the k^{th} pixel calculated with the point source attributes and imaging parameters given in Fig. 2. For the CCD scenario, Fig. 2(a) shows that α_R of Eq. 5 is relatively small for every pixel, with the center pixel having the maximum α_R of only 0.471. By Theorem 4.1, this implies that the Fisher information for this scenario is less than half of that for the ideal scenario. For the EMCCD scenario, Fig. 2(b) shows that α_{GeomR} of Eq. 8 is at least 0.5 in every pixel, demonstrating a significant increase in the information content of every pixel due to the high mean gain multiplication that is assumed. The minimum α_{GeomR} value is 0.502, which implies, by Theorem 4.1, that the Fisher information for this scenario is greater than half of that for the ideal scenario.

The observations made about the Fisher information are reflected in the limits of the localization accuracy calculated for the ideal, the CCD, and the EMCCD data models. Shown in Table 1, the ideal scenario of Poisson data has the best accuracy limit of 8.18 nm. In comparison, the CCD scenario has a significantly worse accuracy limit of 20.18 nm due to the addition of readout noise. In contrast, with the use of high mean gain multiplication to drown out the readout noise, the EMCCD scenario has a much improved accuracy limit of 11.17 nm. By defining the estimated parameters as the coordinates of the point source, i.e., $\theta = (x_0, y_0)$, these limits of accuracy were obtained as the square root of the Cramer-Rao lower bound on the variance of the estimates of x_0 .

Table 1 also shows, for each data model, the mean and standard deviation of the estimates of the x_0 coordinate from

Table 1. Limits of the localization accuracy and results of maximum likelihood estimations using simulated images

Data model	No. of x_0 estimates	True x_0 (nm)	Mean of x_0 estimates (nm)	Limit of the localization accuracy (nm)	Standard deviation of x_0 estimates (nm)
Ideal	1000	880	879.81	8.18	8.31
CCD	1000	880	880.20	20.18	19.91
EMCCD	1000	880	879.92	11.17	11.42

maximum likelihood estimations (see [10] for details) carried out on 1000 simulated images of the point source. For each data scenario, the mean of the estimates recovers reasonably closely the true value of x_0 , while the standard deviation of the estimates comes reasonably close to the corresponding limit of the localization accuracy. These results suggest that the maximum likelihood estimator is capable of attaining the Cramer-Rao lower bound.

5. REFERENCES

- [1] D. L. Snyder, C. W. Helstrom, A. D. Lanterman, M. Faisal, and R. L. White, "Compensation for readout noise in CCD images," *J. Opt. Soc. Am. A*, vol. 12, pp. 272–283, 1995.
- [2] A. G. Basden, C. A. Haniff, and C. D. Mackay, "Photon counting strategies with low-light-level CCDs," *Mon. Not. R. Astron. Soc.*, vol. 345, pp. 985–991, 2003.
- [3] N. G. Walter, C. Huang, A. J. Manzo, and M. A. Sobhy, "Do-it-yourself guide: how to use the modern single-molecule toolkit," *Nat. Methods*, vol. 5, pp. 475–489, 2008.
- [4] R. E. Thompson, D. R. Larson, and W. W. Webb, "Precise nanometer localization analysis for individual fluorescent probes," *Biophys. J.*, vol. 82, pp. 2775–2783, 2002.
- [5] R. J. Ober, S. Ram, and E. S. Ward, "Localization accuracy in single molecule microscopy," *Biophys. J.*, vol. 86, pp. 1185–1200, 2004.
- [6] S. Ram, E. S. Ward, and R. J. Ober, "A stochastic analysis of performance limits for optical microscopes," *Multidim. Syst. Sig. Process.*, vol. 17, pp. 27–57, 2006.
- [7] C. R. Rao, *Linear statistical inference and its applications*, Wiley, New York, USA, 1965.
- [8] T. E. Harris, *The theory of branching processes*, Prentice-Hall, USA, 1963.
- [9] J. Hyneczek and T. Nishiwaki, "Excess noise and other important characteristics of low light level imaging using charge multiplying CCDs," *IEEE Trans. Electron Devices*, vol. 50, pp. 239–245, 2003.
- [10] J. Chao, E. S. Ward, and R. J. Ober, "Fisher information matrix for branching processes with application to electron-multiplying charge-coupled devices," *under review*.
- [11] P. Haccou, P. Jagers, and V. A. Vatutin, *Branching processes: variation, growth, and extinction of populations*, Cambridge University Press, Cambridge, UK, 2005.