Zhiping Lin; Yau Wong; Ober, R.J., "Limit of the accuracy of parameter estimation for two molecules moving in close proximity," in *Circuits and Systems (ISCAS), 2015 IEEE International Symposium on* , vol., no., pp.441-444, 24-27 May 2015
doi: 10.1109/ISCAS.2015.7168665
keywords: {biomolecular electronics;matrix algebra;parameter estimation;Fisher information matrix;accuracy limit;biological process;biomolecular interactions;close proximity;closely spaced molecules;molecular motion;moving objects;nanometer scale distances;parameter estimation;photon emission intensity;planar trajectory;two molecules;Accuracy;Detectors;Genetic expression;Microscopy;Noise;Photonics;Trajectory},
URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7168665&isnumber=7168553

Limit of the Accuracy of Parameter Estimation for Two Molecules Moving in Close Proximity

Zhiping Lin
School of Electrical and Electronic Engineering
Nanyang Technological University
Singapore 639798
Email: ezplin@ntu.edu.sg

Yau Wong
School of Engineering
Nanyang Polytechnic
Singapore 569830
Email: wong_yau@nyp.edu.sg

Raimund J. Ober
Eric Jonsson School of Electrical Engineering and Computer Science
University of Texas at Dallas
Richardson, TX 75083-0688 USA
Email: ober@utdallas.edu

***Abstract*—Biomolecular interactions are central to biological processes and typically take place at nanometer scale distances. They often involve molecular motion which is known to affect the accuracy of the parameter estimates. Therefore, in this paper, we consider a case of two closely spaced molecules with planar trajectory and present a general expression of the Fisher information matrix in terms of their trajectory from which the benchmark for the accuracy of the parameter estimates is obtained. Through simulations, we show its application in the case of two moving objects and another case where only one of the two objects is moving. It is shown that the deterioration of the limit of the accuracy is not only dependent on the proximity of their starting position but also on their speed and direction of movement. The effect of differing photon emission intensities on the limit of the accuracy of parameter estimation is also investigated.**

I. Introduction

Biomolecular interactions, which typically take place at nanometer scale distances, are central to biological processes [1], [2]. With the development of highly sensitive detectors and the introduction of high quantum yield fluorescent proteins in recent years, there has been an increased use of single-molecule fluorescence microscopy in the study of biomolecular interactions. One such approach involves the acquisition of signals from closely spaced molecules simultaneously [3] from which the location, velocity, direction, etc., of the molecules of interest can be estimated using a specific estimation technique such as the maximum likelihood method. However, there is always the issue of estimation accuracy and thus it is important to have a benchmark, which is the primary focus of our paper, against which it can be measured.

In a recent paper by Ram *et al.* [4], they addressed the accuracy issue of two stationary point sources with the Cramér-Rao inequality. However, biomolecules are seldom stationary during interactions and their motions are known to affect the accuracy of the parameter estimates. Hence, in this paper we present expressions to determine the accuracy benchmark for both an ideal and practical case where movements of molecules in proximity are taken into account. To obtain these expressions, we adopt the theoretical framework laid out in [5] and extend it to a case where two molecules in close proximity are moving in a plane. We provide a general expression of the Fisher information matrix from which the benchmark or limit of the accuracy of the parameter estimates can be obtained. Even though this expression is in terms of the objects' trajectory, it can easily be adapted for use in cases where either one or both of the objects are stationary. When both objects are stationary, it reduces to that addressed by Ram *et al.* [4].

The paper is organized as follows. In Section II, we provide the general expression of the Fisher information matrix for the estimation problem that underlies the acquisition of images of two closely spaced molecules that move independently in a plane. In Section III, we provide some insights based on the simulation work done. Conclusions are presented in Section IV.

II. General Framework

In this paper, we consider a basic optical microscope setup where the emission and detection of photons from the fluorescent-labelled objects are inherently a random phenomenon. The acquired data comprise signals from both objects and are modelled as a spatio-temporal random process. It consists of the time points and the spatial coordinates of the detected photons. Since the measure of the accuracy of the parameter estimates is the standard deviation, the question of what is the best possible standard deviation of the parameter estimates for a given data set regardless of the estimation technique used arises. To this end, we consider the Cramér-Rao lower bound (CRLB) from which the limit of the accuracy of the parameter estimates or accuracy limit, in short, is obtained.

A. Fisher Information Matrix for the Ideal Case

We first consider an ideal case where images of two objects moving in close proximity on a plane are acquired simultaneously in the absence of extraneous noise with a non-pixelated detector of infinite size, i.e., $\mathcal{C} = \mathbb{R}^2$. This provides us with the best case scenario of what is theoretically possible sans the effect of pixelation and extraneous noise. For our work, we assume that the photon detection rate is independent of the parameter vector θ, which denotes the parameters to be estimated, such that

$$\Lambda(\tau) = \Lambda_1(\tau) + \Lambda_2(\tau), \quad \tau \geq t_0, \tag{1}$$

where $\Lambda_i(\tau)$, $i = 1, 2$, are the photon detection rates of the two objects, respectively. As for the photon distribution profile, it

978-1-4799-8391-9/15/$31.00 ©2015 IEEE

is the sum of the weighted photon distribution profiles of the respective objects and is expressed as

$$f_{\theta,\tau}(r) = \epsilon_1 f_{\theta,\tau,1}(r) + \epsilon_2 f_{\theta,\tau,2}(r), \\ r^2 = (x^2+y^2) \in \mathbb{R}^2, \tau \geq t_0,\ \theta \in \Theta, \quad (2)$$

where $\epsilon_i = \Lambda_i(\tau)/\Lambda(\tau)$ and $f_{\theta,\tau,i}(r)$, $i = 1,2$, are the weightage and photon distribution profiles of the two objects, respectively. The photon distribution profile can further be expressed as a scaled and shifted version of the image of the object [6] such that

$$f_{\theta,\tau,i}(r) = 1/M^2 q_i\left(x/M - x_{\theta,i}(\tau), y/M - y_{\theta,i}(\tau)\right), \\ i = 1,2, r^2 = (x^2+y^2) \in \mathbb{R}^2, \tau \geq t_0, \quad (3)$$

where q_i and $(x_{\theta,i}(\tau), y_{\theta,i}(\tau))$, $i = 1,2$, denote the image function and planar trajectory of the two objects, respectively, and $M > 0$ denotes the lateral magnification.

The image functions are assumed to be radially symmetric in order that the derivation is tractable. Moreover, it is unlikely that a correct model can be identified with ultimate certainty due to the inherent variability of biological samples. With this assumption and the following general expression of the Fisher information matrix $\mathbf{I}(\theta)$ for the image detection process

$$\mathbf{I}(\theta) = \int_{t_0}^{t}\int_{\mathcal{C}} \frac{\Lambda(\tau)}{f_{\theta,\tau}(r)} \left(\frac{\partial f_{\theta,\tau}(r)}{\partial\theta}\right)^T \left(\frac{\partial f_{\theta,\tau}(r)}{\partial\theta}\right) dr d\tau, \quad (4)$$

we obtain the below expression using the approach in [5].

$$\mathbf{I}(\theta) = \int_{t_0}^{t}\int_{\mathcal{C}} \frac{V_\theta^T(\tau)}{w_{2d}(\tau)} \begin{bmatrix} W_{11} & W_{12} \\ W_{12}^T & W_{22} \end{bmatrix} V_\theta(\tau) dx dy d\tau, \quad (5)$$

where

$$w_{2d}(\tau) = \sum_{i=1}^{2} \Lambda_i(\tau) q_i\left(x/M - x_{\theta,i}(\tau), y/M - y_{\theta,i}(\tau)\right),$$

$$V_\theta(\tau) = \left[\frac{-\partial x_{\theta,1}(\tau)}{\partial\theta} \quad \frac{-\partial y_{\theta,1}(\tau)}{\partial\theta} \quad \frac{-\partial x_{\theta,2}(\tau)}{\partial\theta} \quad \frac{-\partial y_{\theta,2}(\tau)}{\partial\theta}\right]^T,$$

$$W_{kl} = \Lambda_k(\tau)\Lambda_l(\tau) \\ \times diag\left(\frac{\partial q_k(\frac{x}{M} - x_{\theta,k}(\tau), \frac{y}{M} - y_{\theta,k}(\tau))}{\partial x} \frac{\partial q_l(\frac{x}{M} - x_{\theta,l}(\tau), \frac{y}{M} - y_{\theta,l}(\tau))}{\partial x}, \right. \\ \left. \frac{\partial q_k(\frac{x}{M} - x_{\theta,k}(\tau), \frac{y}{M} - y_{\theta,k}(\tau))}{\partial y} \frac{\partial q_l(\frac{x}{M} - x_{\theta,l}(\tau), \frac{y}{M} - y_{\theta,l}(\tau))}{\partial y}\right) \\ k,l = 1,2,\ \theta \in \Theta,\ \tau \geq t_0.$$

The above expression is also applicable to the case where the two moving objects are far apart since no assumption has been made with regard to the distance between them. Moreover, it can also be easily adapted to cases where either one or both of the objects are stationary by simply replacing the trajectory with the location of the objects concerned.

It should be noted that the size of the Fisher information matrix is dependent on that of the weighting function $V_\theta(\tau)$ which in turn is dependent on the number of parameters to be estimated. Thus, in the absence of prior knowledge for the parameters, we can expect the size of the matrix to be the largest when both objects are moving and the smallest when both are stationary [7]. To obtain the accuracy limit of the parameter estimates or simply the fundamental limit, we take the square root of the diagonal elements of the inverse Fisher information matrix.

B. Fisher Information Matrix for the Practical Case

In practice, images are often acquired by a pixelated detector of finite size, e.g. charge coupled device (CCD) detector. Moreover, the photon emission by an object and its detection by a detector are typically assumed to follow a Poisson process. In this section, we consider two cases; one where extraneous noise is present and another where it is absent. The latter hypothetical case is important because it allows the experimentalists to compare results obtained from any practical detector against it to determine the extent to which detector noise deteriorates the obtainable parameter estimation accuracy. Moreover, it enables experimentalists to evaluate the effect of pixelation when compared to the ideal case. We refer to the accuracy limit obtained in this section as the practical limit.

In the case of the hypothetical noiseless detector, the image is assumed to contain just a Poisson-distributed number of photons in each of its pixels, uncorrupted by detector noise. Hence, the Fisher information matrix comprising N_p pixels and acquired over the exposure interval $[t_0, t]$ is given by [8]

$$\mathbf{I}(\theta) = \sum_{k=1}^{N_p} \frac{1}{\mu_{\theta,k}} \left(\frac{\partial\mu_{\theta,k}}{\partial\theta}\right)^T \left(\frac{\partial\mu_{\theta,k}}{\partial\theta}\right), \quad \theta \in \Theta, \quad (6)$$

where $\mu_{\theta,k}$ is the mean of the number of detected photons at the k^{th} pixel due to the objects of interest for the exposure interval $[t_0, t]$. It is expressed as [4]

$$\mu_{\theta,k} = \int_{t_0}^{t}\int_{C_k} \Lambda(\tau) f_{\theta,\tau}(x,y) dx dy d\tau, \\ k = 1,2,\ldots,N_p, \tau \geq t_0,\ \theta \in \Theta. \quad (7)$$

In the case of the pixelated detector, the image is assumed to contain, in each of its pixels, Poisson-distributed photon signals from both the objects of interest and background components. These signals are further corrupted by the detector's additive readout noise. The readout noise is modelled as a Gaussian random variable and we let the readout noise in the k^{th} pixel be Gaussian-distributed with mean η_k and variance $\sigma_{w,k}^2$, $k = 1,2,\ldots,N_p$. For this data model, the Fisher information matrix comprising N_p pixels and acquired over the exposure interval $[t_0, t]$ is given by [5]

$$\mathbf{I}(\theta) = \sum_{k=1}^{N_p} \left(\frac{\partial\mu_{\theta,k}}{\partial\theta}\right)^T \left(\frac{\partial\mu_{\theta,k}}{\partial\theta}\right) \\ \times \left(\int_R \frac{\left(\sum_{l=1}^{\infty} \frac{[\upsilon_{\theta,k}]^{l-1} e^{-\upsilon_{\theta,k}}}{(l-1)!} \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{1}{2}\left(\frac{z-l-\eta_k}{\sigma_{w,k}}\right)^2}\right)^2}{p_{\theta,k(z)}} dz - 1 \right), \quad (8)$$

where $\upsilon_{\theta,k} = \mu_{\theta,k} + \beta_k$ and β_k, which is the mean of the number of detected photons at the k^{th} pixel due to the background components for the exposure interval $[t_0, t]$, is expressed as

$$\beta_k = \int_{t_0}^{t}\int_{C_k} \Lambda_b(\tau) f_b(x,y) dx dy d\tau, \quad k = 1,2,\ldots,N_p. \quad (9)$$

where $\Lambda_b(\tau)$ denotes the photon detection rate from the background components and f_b denotes its photon distribution

profile which is assumed to be independent of the time point τ. The Poisson-Gaussian mixture probability density function $p_{\theta,k(z)}$ is given by

$$p_{\theta,k(z)} = \frac{1}{\sqrt{2\pi}\sigma_k} \sum_{l=0}^{\infty} \frac{[\upsilon_{\theta,k}]^l \, e^{-\upsilon_{\theta,k}}}{l!} e^{-\frac{1}{2}\left(\frac{z-l-\eta_k}{\sigma_{w,k}}\right)^2}, \quad z \in \mathbb{R},\ k = 1, 2, \ldots, N_p.$$

By taking the square root of the diagonal element of the inverse of (6) and (8), we can obtain the practical limit of the hypothetical detector and that where extraneous noise is present, respectively

III. Simulation Results

To gain insights into how the movement of molecule or molecules affect the accuracy limit of the parameter estimates of molecules in close proximity, we consider three scenarios where there are two molecules in close proximity. Since the molecules are usually much smaller than the optical resolution of the optical microscope used in acquiring their images, they can reasonably be assumed to be point sources. Thus, the acquired image consists of diffraction-limited spots from both point sources. In the first scenario, we consider the two molecules to be moving towards each other while in the second, only one is moving. The imaging condition for both scenarios is considered to be analogous to that of the Rayleigh's criterion, i.e., identical, self-luminous in-focus point sources emitting unpolarized incoherent light. The third scenario is similar to the second except that their photon detection rates are different. In the simulations, the photon detection rates are assumed to be constant, i.e., $\Lambda_1(\tau) = \Lambda_1$, $\Lambda_2(\tau) = \Lambda_2 \in \mathbb{R}^+$, and their image functions are described by a Gaussian profile since it provides an excellent approximation to the intensity distribution in the central region of the diffraction limited image [9]. The background component is assumed to be a constant, i.e., $\Lambda_b(\tau) = \Lambda_b \in \mathbb{R}^+$, and uniformly distributed.

In the first scenario, we denote the starting points of the two linearly moving objects by (x_1, y_1) and (x_2, y_2). They move towards each other at the same speed, i.e., $\upsilon_1 = \upsilon_2$, and their directions of movement are denoted by ϕ_1 and ϕ_2. The parametric expressions of their trajectory for the time interval $[t_0, t]$ are defined as $x_{\theta,i}(\tau) = x_i + \upsilon_i(\tau - t_0)\cos\phi_i$, $y_{\theta,i}(\tau) = y_i + \upsilon_i(\tau - t_0)\sin\phi_i$, $i = 1, 2$, $t_0 \leq \tau \leq t$ and the unknown parameter vector $\theta = (x_1, y_1, \phi_1, \upsilon_1, x_2, y_2, \phi_2, \upsilon_2)$.

From Fig. 1, it can be seen that the fundamental and practical limit deteriorate with decreasing distance between their starting points. For a given speed and time interval, the objects move an equal distance towards each other. Therefore when their starting points are closer, the distance travelled by them increases the amount of image overlap. This results in the deterioration of the accuracy limit because more photons become indistinguishable. The results obtained differ from that of two stationary objects whose image overlap is due just to their distance apart [4]. In the case of two moving objects, it is dependent not only on how far apart is their starting points but also on the speed and direction of movement. We can therefore expect the accuracy limit to be at its poorest when the two objects are moving diametrically towards one other because maximum overlapping of their images occurs under these circumstances. An interesting point to note here is that for two identical objects moving at equal speed towards one another,

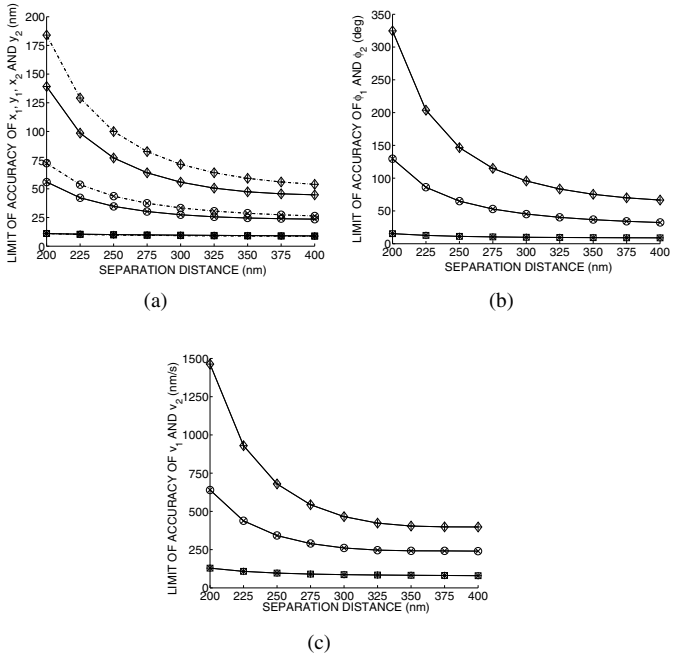
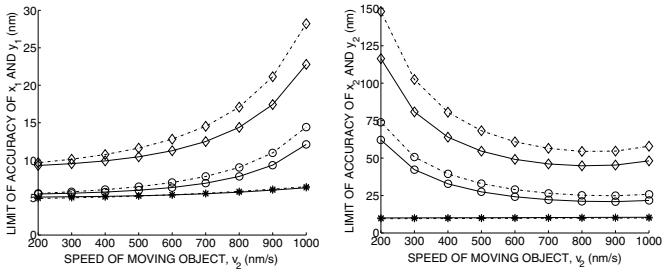


Fig. 1: Accuracy limits of the parameter estimates as a function of the distance between the starting points of the two objects that moved towards one another. Panel (a) shows the limit of the accuracy of x_i (—-) and y_i $(- \cdot -)$, $i = 1, 2$. Panel (b) shows the limit of the accuracy of ϕ_1 and ϕ_2 and Panel (c) of υ_1 and υ_2. The markers $(\circ)$ and $(\times)$ correspond to object 1 and 2, respectively, in the absence of extraneous noise whereas $(\diamond)$ and $(+)$ correspond to object 1 and 2, respectively, in the presence of extraneous noise that consists of Poisson noise (Λ_b) of 2 photons/pixel/s and Gaussian noise (σ_k) of 4 e^-/pixel. The markers $(*)$ and $(\square)$ correspond to the fundamental limit of object 1 and 2, respectively. For all plots, the starting points of the moving objects are located in the object space at $(-87, -87)$ nm and $(86.2, 13)$ nm with respect to the origin of the $x-y$ axes that passes through the center of the detector. The objects move at $\upsilon_1 = \upsilon_2 = 500$ nm/s. The direction of movement $\phi_1 = 30^0$ and $\phi_2 = 210^0$. The photon detection rate $\Lambda_1 = \Lambda_2 = 2000$ photons/s, $\sigma = 84$ nm, the magnification $M = 100$, the acquisition time is 0.2 s, the pixel size is 6.8 μm $\times$ 6.8 μm and the array size is 21 $\times$ 21 pixels.

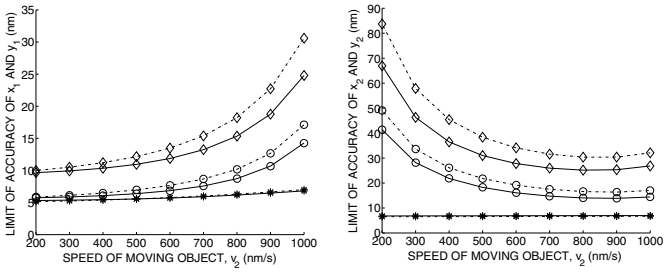
their accuracy limits are equal even though their starting points are different. Thus the accuracy limit is independent of the starting points. We can also infer that their accuracy limits will be equal when both objects move diametrically away from one another at equal speed.

In the second scenario, the stationary object is arbitrarily located at (x_1, y_1) while the starting point of the moving object (x_2, y_2) is located 200 nm to its right at a relative angle of 30^o and thus the unknown parameter vector $\theta = (x_1, y_1, x_2, y_2, \phi_2, \upsilon_2)$. From Fig. 2, the detrimental effect of motion on the localization limit becomes apparent from the fact that the localization limit of the moving object is significantly poorer than that of the stationary object. It also shows how the movement of object affects the accuracy limit of those in its vicinity. For the stationary object, its localization limit is expected to be constant regardless of the speed of the moving object if they are distinctly apart. However, due to the overlapping of their images as the moving object approaches,



(a) (b)

Fig. 2: Localization limits as a function of the speed of an object moving linearly towards a stationary object. Panel (a) shows the limit of the accuracy of x_1 (—) and y_1 $(-\cdot-)$ and Panel (b) of x_2 (—) and y_2 $(-\cdot-)$. The markers $(*)$, $(\circ)$ and $(\diamond)$ correspond to the ideal case, the practical case without extraneous noise and where extraneous noise is present, respectively. All other settings are as given in Fig. 1.



(a) (b)

Fig. 3: Localization limits as a function of the speed of an object moving linearly towards a stationary object. Panel (a) shows the limit of the accuracy of x_1 (—) and y_1 $(-\cdot-)$ and Panel (b) of x_2 (—) and y_2 $(-\cdot-)$. Other than the photon detection rates which are $\Lambda_1 = 2000$ photons/s and $\Lambda_2 = 4000$ photons/s, all markers and settings remain the same as Fig. 2.

the localization limit of the stationary object deteriorates when the detected photons become indistinguishable. In the presence of extraneous noise, the deterioration of the practical localization limit exacerbates. As for the moving object, its poor practical localization limit at low speed is due primarily to a lack of information because of the short distance travelled. This practical localization limit improves with increasing speed but subsequently deteriorates because of the influence from the stationary object.

Thus far the imaging conditions used are analogous to those of the Rayleigh's criterion. However, in practice, the intensity of the fluorescent-labelled molecules may differ [2]. Therefore, in this scenario we assume that the photon detection rate of the moving object is twice that of the stationary object, i.e., $\Lambda_2 = 2\Lambda_1$. Under these circumstances, the localization limit as shown in Fig. 3 has similar trend to that in Fig. 2 except for their values. Contrast to that in Fig. 2, the localization limit of x_2 and y_2 are better since now there are more detected photons from object 2, i.e., stronger signal. Since the signal from the moving object is stronger, it therefore exerts a stronger influence on the stationary object when their images overlap at higher speed. Thus, it results in a poorer localization limit for x_1 and y_1 when compared to the corresponding results in Fig. 2.

IV. Conclusions

In this paper, we have presented a general expression of the Fisher information matrix from which the accuracy limit of the parameter estimates can be obtained for the case where two objects are moving in a plane. It can also be easily adapted to cases where one or both objects are stationary. Through simulations, we have shown that the deterioration of the accuracy limit is due to the amount of image overlap. In the case of moving objects, the amount of image overlap is not only dependent on the proximity of their starting point but also on their speed and direction of movement. Moreover, for identical objects moving diametrically towards each other at equal speed, their accuracy limits are equal and independent of their starting points. For objects with different intensity, the localization limit of the object with the stronger signal improves at the expense of the weaker one when they are in close proximity.

Acknowledgment

This work is supported in part by NIH grant RO1 GM085575.

References

[1] J. Yu, J. Xiao, X. Ren, K. Lao and X.S. Xie, "Probing gene expression in live cells, one protein molecule at a time," *Science*, vol. 311, pp. 1600-1603, Mar. 2006.

[2] Y. Sako, S. Minoguchi and T. Yanagida, "Single-molecule imaging of EGFR signalling on the surface of living cells," *Nature Cell Biology*, vol. 2, pp. 168-172, Feb. 2000.

[3] W.E. Moerner and D.P. Fromm, "Methods of single-molecule fluorescence spectroscopy and microscopy," *Review of Scientific Instruments*, vol. 74, pp. 3597-3619, Aug. 2003.

[4] S. Ram, E.S. Ward and R.J. Ober, "A stochastic analysis of distance estimation approaches in single molecule microscopy-quantifying the resolution limits of photon-limited imaging systems," *Multidimensional Systems and Signal Processing*, vol. 24, no. 3, pp. 503-542, Sep. 2013.

[5] Y. Wong, Z. Lin and R.J. Ober, "Limit of the accuracy of parameter estimation for moving single molecules imaged by fluorescence microscopy," *IEEE Transactions of Signal Processing*, vol. 59, no. 3, pp. 895-911, Mar. 2011.

[6] I.T. Young, "Quantitative microsocpy," *IEEE Engineering in Medicine and Biology*, vol. 15, pp. 59-66, Jan. 1996.

[7] Z. Lin, Y. Wong and R.J. Ober, "Influence of prior knowledge on the accuracy limit of parameter estimation in single-molecule fluorescence microscopy ," *in Proc. IEEE ISCAS* 2013, *Beijing*, *China*, May 2013.

[8] D.L. Snyder and M.I. Miller, *Random Point Processes in Time and Space (2nd edition).* New York: Springer Verlag, 1999.

[9] K.I. Mortensen, L.S. Churchman, J.A. Spudich and H. Flyvbjerg, "Optimized localization analysis for single-molecule tracking and super-resolution microscopy," *Nature Methods*, pp. 377-381, Apr. 2010.