Invited Paper

Single emitter localization analysis in the presence of background

S. Stallinga

Department of Imaging Physics, Technische Univ. Delft, Lorentzweg 1, 2628 CJ,

Delft, The Netherlands

ABSTRACT

Localization microscopy for imaging at the nano-scale relies on the quality of fitting the emitter positions from the measured light spots. The type and magnitude of the noise in the detection process, the background light level and the Point Spread Function model that is used in the fit are of paramount importance for the precision and accuracy of the fit. We present several developments on the computational methods and performance limits of single emitter localization, targeting specifically these three aspects.

Keywords: Localization microscopy, single molecule, maximum likelihood estimation, Cramér Rao lower bound, background

1. INTRODUCTION

Since the advent of PALM,[1] STORM[2] and fPALM[3] in 2006 localization microscopy based on stochastically and sparsely activated fluorescent emitters has emerged as a powerful tool to study molecular building blocks of life on the nano-scale.[4] One of the cornerstones of localization microscopy is the determination of the positions of the sparse emitters by fitting a model Point Spread Function (PSF) to the actually observed spot shape over a small Region Of Interest (ROI).[5,6] These ROIs are selected from the raw frames of the entire sequence of acquired images by an initial spot detection algorithm.

In this paper we present formal developments concerning three important aspects of single emitter localization. The first concerns the noise model used in localization. The emitter location is found via a process called Maximum Likelihood Estimation (MLE). An equation for the likelihood of the observed spot shape given a model for the PSF centered at the position of the emitter is set up based on the noise sources in the photon detection process. Noise sources are shot noise originating from the quantum character of light, the readout noise of the camera, and possibly, excess noise originating from electron multipication when Electron Multiplying (EM) CCDs are used.[7] We give an outline of MLE that interpolates between the shot noise only and the readout noise only limits. The second aspect concerns the background that usually obscures the single molecule signals. The signal-to-background level appears to be of eminent importance for the precision of the single molecule localization algorithm.[8] Possible non-unformities in the background have a negative impact on the accuracy of the localization.[9] We present an extension of the previously presented formalism[8] for the minimum variance in the fit parameters of the MLE procedure, the so-called Cramér Rao Lower Bound (CRLB). In our previous study the background was assumed to be a fixed parameter, in the current treatment we take the background as an additional statistical parameter. The third aspect concerns the model assumed for the PSF. The conventional choice is a Gaussian model, but this choice is based entirely on convenience and convention, not on physical principle. It has been shown that the chocice of a Gaussian PSF model leads to satisfactory results provided the actual PSF is symmetric[10] but there is no reason to exclude other heuristic PSF models that could work equally well as or perhaps better than the Gaussian model. In particular we will focus on the Lorentzian PSF model and will compare the anticipated localization precision to that of the Gaussian PSF model.

E-mail: s.stallinga@tudelft.nl

Optical Systems Design 2015: Computational Optics, edited by Daniel G. Smith, Frank Wyrowski, Andreas Erdmann, Proc. of SPIE Vol. 9630, 96300V · © 2015 SPIE
CCC code: 0277-786X/15/$18 · doi: 10.1117/12.2192043

2. MAXIMUM LIKELIHOOD ESTIMATION AND NOISE MODEL

The state-of-the-art in localization algorithms is described in Ref. 11, which describes two algorithmical innovations targeting sCMOS-cameras. The first is related to non-uniformities in offset, gain and readout noise variance. We will neglect that here. The second is a relatively simple way to incorporate the effects of both shot noise and readout noise in spot fitting. Starting point is the expression for the log-likelihood as a summation over all pixels k in the selected ROI:

$$W = \sum_k \left[\left(n_k + \sigma_r^2\right) \log\left(\mu_k + \sigma_r^2\right) - \left(\mu_k + \sigma_r^2\right) - \log\Gamma\left(n_k + \sigma_r^2 + 1\right) \right], \tag{1}$$

where n_k is the observed photon count, $\Gamma(x) = \int_0^\infty ds\, s^{x-1} \exp(-x)$ is the Gamma-function, σ_r the rms readout noise, and the expected photon count at pixel k is given by the integration of the PSF over the pixel area $\mathcal{A}_k$:

$$\mu_k = \int_{\mathcal{A}_k} dxdy\, H\left(x - x_0, y - y_0\right), \tag{2}$$

where the PSF is conventionally approximated by a Gaussian:

$$H\left(x, y\right) = \frac{N}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} + \frac{b}{a^2}. \tag{3}$$

The fit parameters are the coordinates of the emitter x_0 and y_0, the spot width σ, the emitter photon count N and the number of background photons per pixel b, and are determined by maximizing the log-likelihood Eq. (1). The observed photon count n_k satisfies:

$$\langle n_k \rangle = \mu_k, \tag{4}$$

$$\langle n_k^2 \rangle - \langle n_k \rangle^2 = \mu_k + \sigma_r^2. \tag{5}$$

In the limit of zero readout noise $\sigma_r \downarrow 0$ Eq. (1) reduces to the classical shot-noise only log-likelihood:

$$W = \sum_k \left[n_k \log \mu_k - \mu_k - \log\left(n_k!\right) \right], \tag{6}$$

and in the limit of large readout noise $\sigma_r \gg n_k$ for all pixels k we retrieve the least-squares log-likelihood:

$$W = -\sum_k \frac{\left(n_k - \mu_k\right)^2}{2\sigma_r^2}, \tag{7}$$

So, the log-likelihood Eq. (1) interpolates between least-squares fitting and shot noise only maximum likelihood fitting.

In optimization routines as e.g. the Levenberg-Marquardt algorithm the derivatives of the log-likelihood w.r.t. the fit parameters $\theta = \{x_0, y_0, \sigma, N, b\}$ are needed. Using the chain rule we find:

$$\frac{\partial W}{\partial \theta_j} = \sum_k \frac{n_k - \mu_k}{\mu_k + \sigma_r^2} \frac{\partial \mu_k}{\partial \theta_j}, \tag{8}$$

$$\frac{\partial^2 W}{\partial \theta_j \partial \theta_l} = -\sum_k \frac{n_k + \sigma_r^2}{\left(\mu_k + \sigma_r^2\right)^2} \frac{\partial \mu_k}{\partial \theta_j} \frac{\partial \mu_k}{\partial \theta_l} + \sum_k \frac{n_k - \mu_k}{\mu_k + \sigma_r^2} \frac{\partial^2 \mu_k}{\partial \theta_j \partial \theta_l}. \tag{9}$$

It is mentioned that the second term in the second order derivatives is usually discarded as it is generally small close to the optimum of W, and it does not affect the set of parameters for which W is maximal. These derivatives also give rise to a Fisher matrix:

$$F_{jl} = \left\langle \frac{\partial W}{\partial \theta_j} \frac{\partial W}{\partial \theta_l} \right\rangle = -\left\langle \frac{\partial^2 W}{\partial \theta_j \partial \theta_l} \right\rangle = \sum_k \frac{1}{\mu_k + \sigma_r^2} \frac{\partial \mu_k}{\partial \theta_j} \frac{\partial \mu_k}{\partial \theta_l}. \tag{10}$$

The CRLB is found from the diagonal elements of the inverse of the Fisher matrix. From the derived expression for the Fisher-matrix (the term $\mu_k + \sigma_r^2$ in the denominator) we may infer that the effect of readout noise on the best possible emitter position estimation is simply to add the readout noise variance to the background photon count giving an effective background:

$$b_{\text{eff}} = b + \sigma_r^2. \tag{11}$$

Therefore, without loss of generality, we will set $\sigma_r = 0$ in the following. The corresponding equations for the case of non-zero readout noise can then be found by subtituting the background b with the effective background b_{eff}.

3. CRAMÉR-RAO LOWER BOUND OF GAUSSIAN SPOT FITTING

We expand on the treatment of Ref. 8 by adding the background as an additional statistical parameter that is estimated by the MLE procedure. Appearing summations over the N_{pix} pixels in the ROI are at some point approximated by an integration over the entire detector surface. The derivatives of the PSF w.r.t. the fit parameters $\theta = \{x_0, y_0, \sigma, N, b\}$ are:

$$\frac{\partial H}{\partial x_0} = \frac{N\,(x - x_0)}{2\pi\sigma^4} e^{-t}, \tag{12}$$

$$\frac{\partial H}{\partial y_0} = \frac{N\,(y - y_0)}{2\pi\sigma^4} e^{-t}, \tag{13}$$

$$\frac{\partial H}{\partial \sigma} = \frac{N}{\pi\sigma^3}\,(t-1)\,e^{-t}, \tag{14}$$

$$\frac{\partial H}{\partial N} = \frac{1}{2\pi\sigma^2} e^{-t}, \tag{15}$$

$$\frac{\partial H}{\partial b} = \frac{1}{a^2}, \tag{16}$$

with:

$$t = \frac{(x - x_0)^2 + (y - y_0)^2}{2\sigma^2}. \tag{17}$$

The Fisher-matrix elements for the set of parameters $\theta = \{x_0, y_0, \sigma, N, b\}$ then follow from:

$$F_{jl} = \sum_{k=1}^{N_{\text{pix}}} \frac{1}{Ha^2} \frac{\partial Ha^2}{\partial \theta_j} \frac{\partial Ha^2}{\partial \theta_l}, \tag{18}$$

where the derivatives are evaluated at the pixel centers (x_k, y_k) and where we have neglected the variation of the PSF over the area of single pixels. The Fisher matrix elements involving only the parameters $\{x_0, y_0, \sigma, N\}$ may be approximated by the integrals:

$$F_{jl} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dxdy\, \frac{1}{H} \frac{\partial H}{\partial \theta_j} \frac{\partial H}{\partial \theta_l}, \tag{19}$$

as the integrands are rapidly decaying functions of position w.r.t. the spot center. This results in the Fisher matrix elements:[8]

$$F_{x_0 x_0} = F_{y_0 y_0} = \frac{N}{\sigma^2}\,(1 - I_1)\,, \tag{20}$$

$$F_{NN} = \frac{1}{N}\,(1 - I_0)\,, \tag{21}$$

$$F_{\sigma\sigma} = \frac{4N}{\sigma^2}\,(1 - 2I_2 + 2I_1 - I_0)\,, \tag{22}$$

$$F_{\sigma N} = \frac{2}{\sigma}\,(I_0 - I_1)\,, \tag{23}$$

where the parameters:

$$I_n = \frac{\tau}{n!}\int_0^\infty dt\,\frac{t^n}{1+\tau e^t}, \tag{24}$$

for integer n depend on the dimensionless background parameter $\tau = 2\pi b\sigma^2/Na^2$. The Fisher-matrix elements involving the background b can be evaluated in a similar way. For example:

$$\begin{aligned} F_{bb} &= \sum_{l=1}^{N_{\rm pix}} \frac{1}{Ha^2} = \frac{N_{\rm pix}}{b} + \sum_{l=1}^{N_{\rm pix}}\left[\frac{1}{Ha^2} - \frac{1}{b}\right] \\ &\approx \frac{N_{\rm pix}}{b} + \frac{1}{a^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dxdy\,\left[\frac{1}{b + (Na^2/2\pi\sigma^2)\,e^{-(x^2+y^2)/2\sigma^2}} - \frac{1}{b}\right] \\ &= \frac{N_{\rm pix}}{b} - \frac{2\pi\sigma^2}{ba^2}\int_0^\infty dt\,\frac{1}{1+\tau e^t} \\ &= \frac{N_{\rm pix}}{b} - \frac{N\tau I_0}{b^2}. \end{aligned} \tag{25}$$

The other non-zero Fisher matrix elements involving the background b are derived in the same spirit:

$$\begin{aligned} F_{bN} &= \sum_{l=1}^{N_{\rm pix}} \frac{1}{(b + (Na^2/2\pi\sigma^2)\,e^{-t})}\frac{a^2}{2\pi\sigma^2}e^{-t} \\ &\approx \frac{\tau I_0}{b} \end{aligned} \tag{26}$$

$$\begin{aligned} F_{b\sigma} &= \sum_{l=1}^{N_{\rm pix}} \frac{1}{(b + (Na^2/2\pi\sigma^2)\,e^{-t})}\frac{Na^2}{\pi\sigma^3}\,(t-1)\,e^{-t} \\ &\approx \frac{2N\tau}{\sigma b}\,(I_1 - I_0)\,. \end{aligned} \tag{27}$$

The diagonal elements of the inverse of the Fisher matrix results in the CRLB of the fit parameters:

$$(\Delta x)^2 = (\Delta y)^2 = \frac{\sigma^2}{N}\,\frac{1}{1-I_1}, \tag{28}$$

$$(\Delta\sigma)^2 = \frac{\sigma^2}{4N}\,\frac{1 - I_0 - \left(2\pi\sigma^2/N_{\rm pix}a^2\right)I_0 A}{B - C - (2\pi\sigma^2/N_{\rm pix}a^2)\,AC}, \tag{29}$$

$$(\Delta N)^2 = N\,\frac{B - \left(2\pi\sigma^2/N_{\rm pix}a^2\right)C}{B - C - (2\pi\sigma^2/N_{\rm pix}a^2)\,AC}, \tag{30}$$

$$(\Delta b)^2 = \frac{b}{N_{\rm pix}}\,\frac{B - C}{B - C - (2\pi\sigma^2/N_{\rm pix}a^2)\,AC}, \tag{31}$$

with:

$$A = 1 - I_0 + \tau I_0, \tag{32}$$

$$B = 1 - 2I_2 + 2I_1 - I_0, \tag{33}$$

$$C = I_0 - 2I_0 I_2 + I_1^2. \tag{34}$$

Compared to the results obtained in Ref. 8 we see the ratio of the spot area and the ROI area $2\pi\sigma^2/N_{\rm pix}a^2$ appearing. When this ratio is much smaller than one we retrieve the CRLB for σ and N of Ref. 8. A typical ROI is $(6\sigma + a) \times (6\sigma + a)$ wide.[12] With $\sigma \approx a$ this leads to $2\pi\sigma^2/N_{\rm pix}a^2 \approx 0.13$ implying that the corrections related to the finite ROI size over the results of Ref. 8 are relatively small.

For $\tau \to 0$ (low background limit) we find the relatively simple expressions:

$$(\Delta x)^2 = \frac{\sigma^2}{N}, \tag{35}$$

$$(\Delta\sigma)^2 = \frac{\sigma^2}{4N}, \tag{36}$$

$$(\Delta N)^2 = N, \tag{37}$$

$$(\Delta b)^2 = \frac{b}{N_{\rm pix}}. \tag{38}$$

A good approximation to the CRLB for the emitter coordinates has been proposed in Refs. 8, 13:

$$(\Delta x)^2 = \frac{\sigma^2 + a^2/12}{N}\left(1 + 4\tau + \sqrt{\frac{2\tau}{1+4\tau}}\right), \tag{39}$$

which also takes into account the effect of non-zero pixel size, and which is more convenient than the exact result of Mortensen et al.[14] and wich improves over the often quoted result of Thompson, Larson and Webb.[15]

4. ALTERNATIVES TO GAUSSIAN PSF MODEL

The Gaussian function has a width that is characterized by its Full Width Half Maximum (FWHM):

$$\mathrm{FWHM}_{\rm Gauss} = 2\sqrt{2\ln(2)}\sigma, \tag{40}$$

leading to a localization precision at zero background:

$$\Delta x = \frac{\sigma}{\sqrt{N}} = \frac{\mathrm{FWHM}_{\rm Gauss}}{2\sqrt{2\log(2)}\sqrt{N}} \approx 0.425\frac{\mathrm{FWHM}_{\rm Gauss}}{\sqrt{N}} \tag{41}$$

An alternative to the Gaussian might be the Lorentzian, which is also a bell-shaped function:

$$H(x,y) = \frac{N}{\pi\omega^2\left(1 + \left(x^2+y^2\right)/\omega^2\right)^2}, \tag{42}$$

The FWHM of this distribution is:

$$\mathrm{FWHM}_{\rm Lorentz} = 2\sqrt{\sqrt{2}-1}\,\omega. \tag{43}$$

The question now is what the expected localization precision is for fitting with a Lorentzian PSF model. To that end we must evaluate the Fisher-matrix element:

$$\begin{aligned} F_{x_0x_0} &= \int dxdy\,\frac{1}{H(x,y)}\frac{\partial H(x,y)}{\partial x_0}^2 \\ &= \frac{16N}{\pi\omega^6}\int_0^\infty dr\int_0^{2\pi} d\phi\,\frac{r^3\cos^2\phi}{\left(1 + r^2/\omega^2\right)^4} \\ &= \frac{4N}{3\omega^2}. \end{aligned} \tag{44}$$

This results in a localization precision:

$$\Delta x = \frac{\sqrt{3}\,\omega}{4\sqrt{N}} = \frac{\sqrt{3}\mathrm{FWHM}_{\rm Lorentz}}{4\sqrt{\sqrt{2}-1}\sqrt{N}} \approx 0.673\frac{\mathrm{FWHM}_{\rm Lorentz}}{\sqrt{N}}. \tag{45}$$

Given the same FWHM this seems to be a factor 1.6 worse than the Gaussian PSF model. However, in the MLE procedure σ (for the Gaussian PSF model) and ω (for the Lorentzian PSF model) will be adjusted to give the best overall fit of the actual spot shape. This will result in a FWHM of the fitted model that is smaller for the Lorentzian than for the Gaussian, giving a localization precision that differs less than the factor 1.6 for the equal FWHM case.

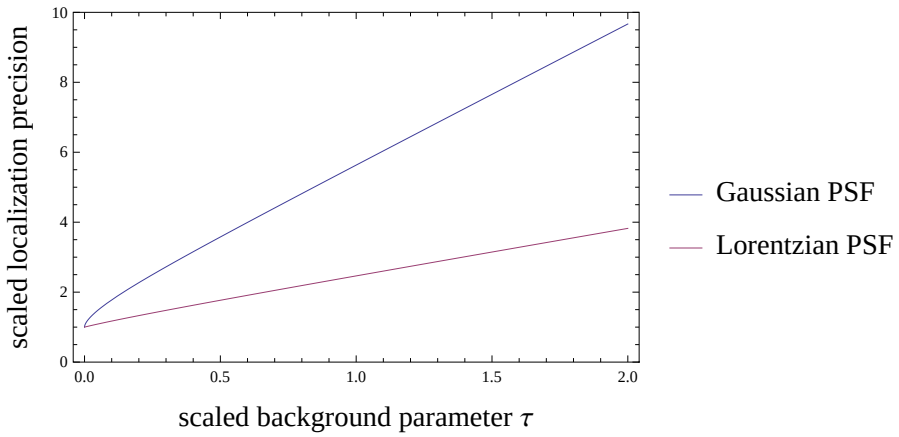

Figure 1. Localization precision scaled by the precision at zero background as a function of the scaled background parameter τ for the Gaussian and Lorentzian PSF model, assuming constant fitted widths or all background levels.

The localization precision for the Lorentz PSF model at non-zero background can be computed in a similar way. Now the PSF is given by:

$$H(x,y) = \frac{N}{\pi\omega^2\left(1+\left(x^2+y^2\right)/\omega^2\right)^2} + \frac{b}{a^2}, \tag{46}$$

and the relevant Fisher-matrix element is:

$$\begin{aligned} F_{x_0x_0} &= \int dxdy\, \frac{1}{H(x,y)} \frac{\partial H(x,y)^2}{\partial x_0} \\ &= \frac{16N}{\pi\omega^6} \int_0^\infty dr \int_0^{2\pi} d\phi\, \frac{r^3\cos^2\phi}{\left(1+r^2/\omega^2\right)^4\left(1+\kappa\left(1+r^2/\omega^2\right)\right)} \\ &= \frac{4N}{3\omega^2}\left[1-3\kappa\left(1+2\kappa\right)+6\kappa^2\left(1+\kappa\right)\log\left(\frac{1+\kappa}{\kappa}\right)\right], \end{aligned} \tag{47}$$

with the dimensionless background parameter $\kappa = \pi b\omega^2/Na^2$, leading to a localization precision:

$$\Delta x = \frac{\sqrt{3}\,\omega}{4\sqrt{N}\sqrt{1-3\kappa\left(1+2\kappa\right)+6\kappa^2\left(1+\kappa\right)\log\left(\left(1+\kappa\right)/\kappa\right)}}. \tag{48}$$

Fig. 1 shows the localization precision for the Gaussian and the Lorentzian PSF models as a function of the background parameter τ of the Gaussian PSF model for $\omega = 2\sigma/\sqrt{3}$, which leads to the same precision at zero background and to $\kappa = 2\tau/3$. We see that the Lorentzian PSF model has a favourable dependence on background in this graph. However, it assumes that the MLE routine will fit Lorentzian spots with the same width $\omega = 2\sigma/\sqrt{3}$ for al levels of background, which may not be the case after all. For an elevated background the tail of the Lorentzian spot shape may be interpreted by the fitting routine as part of the background, leaving only the central part for fitting the actual spot, which will generally lead to a larger value for ω and to a worse precision.

5. CONCLUSION

We have presented several advances concerning the method of single emitter localization. An integrated treatment of MLE including effects of both readout noise and shot noise has been given, leading to the conclusion that the readout noise variance effectively adds to the background. A CRLB analysis of the estimation of background, signal photon count and spot width leads to new analytical expressions for the best possible variance of these

parameters. These expressions provide a correction to previously obtained expressions by taking the effect of a finite ROI into account. In particular, it is found for a high signal-to-background that the background itself can be estimated with a precision that scales inversely proportional to the square root of the number of pixels in the ROI. Finally, an alternative heuristic PSF model is proposed, namely the Lorentzian spot model. This may work equally well as the almost universally used Gaussian PSF model. The relative performance of the Lorentzian PSF model compared to the Gaussian PSF model depends on the level of background and the fitted spot widths for the two cases. A next step would be to compare the fit results for a fully vectorial ground truth, similar to the study of Ref. 10.

5.1 Acknowledgments

Bernd Rieger is thanked for feedbcak on the manuscript.

REFERENCES

[1] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, "Imaging Intracellular Fluorescent Proteins at Nanometer Resolution," *Science* **313**, pp. 1643–1645, 2006.

[2] M. J. Rust, M. Bates, and X. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)," *Nat. Methods* **3**, pp. 793–795, 2006.

[3] S. T. Hess, T. P. Girirajan, and M. D. Mason, "Ultra-high resolution imaging by fluorescence photoactivation localization microscopy", *Biophys J.* **91**, pp. 4258–4272, 2006.

[4] H. Deschout, F. C. Zanacchi, M. Mlodzianoski, A, Diaspro, J. Bewersdorf, S. T. Hess, and K. Braeckmans, "Precisely and accurately localizing single emitters in fluorescence microscopy," *Nat. Methods* **11**, pp. 253–266, 2014.

[5] A. Small, and S. Stahlheber, "Fluorophore localization algorithms for super-resolution microscopy," *Nat. Methods* **11**, pp. 267–279, 2014.

[6] D. Sage, H. Kirshner, T. Pengo, N. Stuurman, J. Min, S. Manley, and M. Unser, "Quantitative evaluation of software packages for single-molecule localization microscopy," *Nat. Methods*, doi:10.1038/nmeth.3442, 2015.

[7] J. Chao, E. S. Ward, and R. J. Ober, "Fisher information matrix for branching processes with application to electron-multiplying charge-coupled devices," *Multidim. Syst. Sign. Process.* **23**, pp. 349–379, 2012.

[8] B. Rieger, and S. Stallinga, "The lateral and axial localization uncertainty in super-resolution light microscopy," *ChemPhysChem* **15**, pp. 664–670, 2014.

[9] E. Hoogendoorn, K. C. Crosby, D. Leyton-Puig, R. M. P. Breedijk, K. Jalink, T. W. J. Gadella, and M. Postma, "The fidelity of stochastic single-molecule super-resolution reconstructions critically depends upon robust background estimation," *Sci. Reports* **4**, 3854, 2014.

[10] S. Stallinga, and B. Rieger, "Accuracy of the Gaussian Point Spread Function model in 2D localization microscopy," *Opt. Express* **18**, pp. 24461–24476, 2010.

[11] F. Huang et al., "Video-rate nanoscopy using sCMOS camera-specific single-molecule localization algorithms," *Nat. Methods* **10**, pp. 653–658, 2013.

[12] C. S. Smith, N. Joseph, B. Rieger, and K. A. Lidke, "Fast, single-molecule localization that achieves theoretically minimum uncertainty," *Nat. Methods* **7**, pp. 373–375, 2010.

[13] S. Stallinga, and B. Rieger, "The effect of background on the localization uncertainty in single emitter imaging," *, IEEE Int. Symp. Biomedical Imaging*, pp. 988–991, 2012.

[14] K. I. Mortensen, L. Stirling Churchman, J. A. Spudich, and H. Flyvbjerg, "Optimized localization analysis for single-molecule tracking and super-resolution microscopy," *Nat. Methods* **7**, pp. 377–381, 2010.

[15] R. E. Thompson, D. R. Larson, and W. W. Webb, "Precise nanometer localization analysis for individual fluorescent probes," *Biophys. J.* **82**, pp. 2775–2783, 2002.