

A Mixture Model for Quantum Dot Images of Kinesin Motor Assays

John Hughes* and John Fricks**

Department of Statistics, The Pennsylvania State University, University Park, Pennsylvania 16802, U.S.A.

**email:* jph264@psu.edu

***email:* fricks@stat.psu.edu

SUMMARY. We introduce a nearly automatic procedure to locate and count the quantum dots in images of kinesin motor assays. Our procedure employs an approximate likelihood estimator based on a two-component mixture model for the image data; the first component has a normal distribution, and the other component is distributed as a normal random variable plus an exponential random variable. The normal component has an unknown variance, which we model as a function of the mean. We use B-splines to estimate the variance function during a training run on a suitable image, and the estimate is used to process subsequent images. Parameter estimates are generated for each image along with estimates of standard errors, and the number of dots in the image is determined using an information criterion and likelihood ratio tests. Realistic simulations show that our procedure is robust and that it leads to accurate estimates, both of parameters and of standard errors.

KEY WORDS: Bio-imaging; Fluorescence microscopy; Kinesin motor protein; Maximum likelihood; Mixture model; Quantum dot; Spline; Variance-function estimation.

1. Introduction

Precise tracking of microscopic biological specimens can provide insights into the function and dynamics of those specimens (Yildiz and Selvin, 2005). This tracking can be accomplished by analyzing digital images (usually TIFF) produced by a CCD (charge-coupled device) camera interfaced with a fluorescence microscope. In fluorescence microscopy, a specimen of interest is marked with a fluorescent particle, or fluorophore, before being irradiated with light at the excitation wavelength of the fluorophore. When the excited electrons revert to their ground state, they emit photons at the fluorophore's emission wavelength. A filter separates the emitted light from the excitation light so that only the light from the fluorophore can pass through the microscope to the camera system (Rost, 1992).

Traditional organic fluorophores, for example, green fluorescent proteins (GFPs), present several limitations including dimness, photo-instability, and narrow excitation bandwidths. Consequently, quantum dots, which are composed of semiconductor nanocrystals, have emerged as an attractive alternative for bio-imaging applications. Quantum dots are bright and photostable, and have broad excitation spectra and narrow emission spectra, owing to the so-called quantum confinement effect exhibited by semiconductor nanocrystals (Arya et al., 2005).

In conjunction with the Hancock Lab at Penn State University, we use fluorescence microscopy to study motor proteins, in particular kinesin motor proteins, which are important agents of intracellular transport (Vale, 2007). Our images are unique in some respects. Each fluorophore is attached to a head (roughly 5 nm in diameter) or near the neck linker of a kinesin motor. In order to infer the dynamics of the kinesin head, it is important to determine the location

of the fluorophore to within a few nanometers, a task that is complicated by a pixel width on the order of several tens of nanometers. Luckily, the true location of the fluorophore is at the center of a point spread function that contains a great deal of information about the center. By extracting this information, we can overcome the limited resolution of the imaging apparatus. Additionally, the motor assays contain a relatively small number of fluorophores, and so the methods presented here may not be efficient enough computationally to be used in a more general setting.

Until recently our kinesin assays employed only GFPs, and so, in 2009, we presented a full likelihood-based method for counting and locating the fluorophores in GFP images (Hughes, Fricks, and Hancock, 2009). Our unified parametric approach makes it possible to efficiently and automatically locate and count the fluorophores in an image so that experimentalists can process many images (hundreds or thousands) in a short time and without human intervention. Moreover, the statistical efficiency of maximum likelihood estimators allows for precise localization of fluorophores.

When the Hancock Lab subsequently began using quantum dots, we discovered that our sampling model for a GFP image, which will be detailed in Section 2, does not fit a typical quantum dot image. More specifically, we noticed that the pixel intensities for the background of a quantum dot image do not follow the (approximate) normal distribution assumed by the GFP model; the intensities are strongly right skewed. And a fit of our GFP model to a quantum dot image yields large standardized residuals for pixels near dots, which indicates that those pixels have larger variances than predicted by the GFP model. Figure 1 shows a histogram for 1000 background pixels from the quantum dot image that will be analyzed in Section 5. The image is also shown in the figure, as is a relief

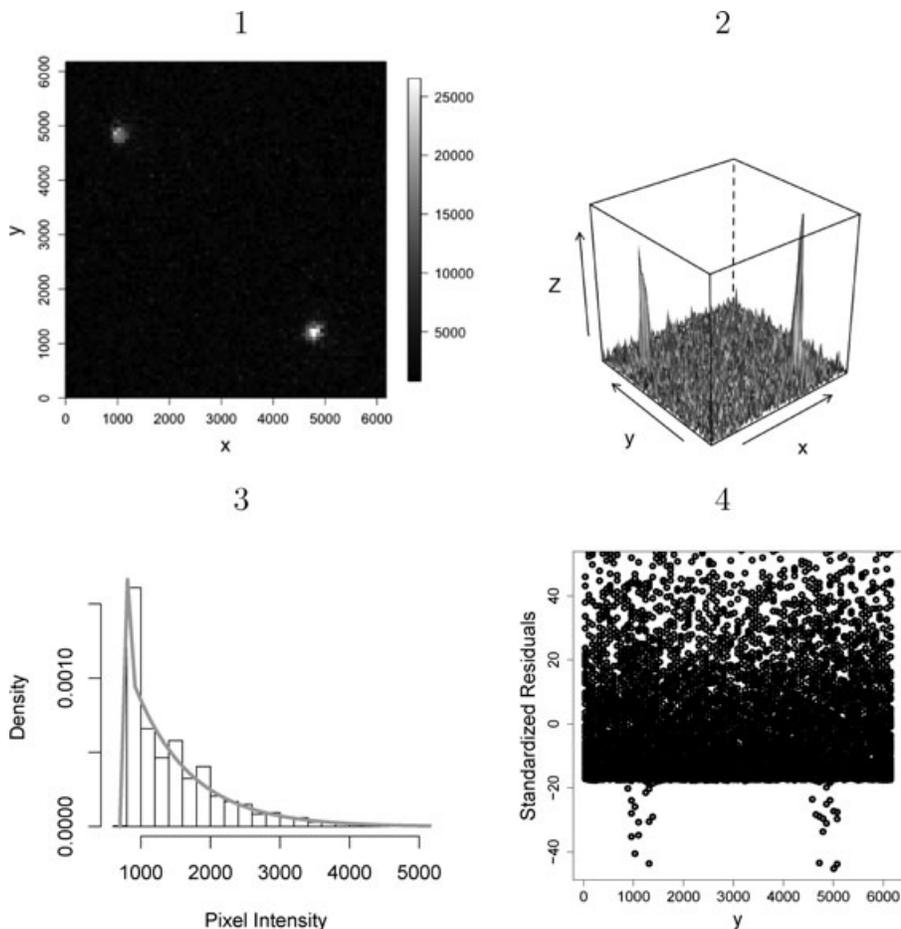


Figure 1. A typical quantum dot image of a kinesin assay. The x and y coordinates of pixels are given in nanometers. Z represents pixel intensity. Panel 3 shows a histogram for 1000 pixels from the image background. Notice the strong right skewness of the background intensities, which contradicts the sampling model for a GFP image. Panel 4 shows the standardized residuals from a fit of our GFP model, versus the pixels' y coordinates. The large residuals (some of which are clearly visible below -20 and at y locations of 1000 and 5000, approximately) around the dots indicate a larger variance there than predicted by the GFP model.

plot of the image and a plot of the standardized residuals from a fit of our GFP model.

Not surprisingly, applying the GFP model in the presence of the above-mentioned departures—strong skewness and larger variance near fluorophores—leads to erroneous model selection and erroneous inference for fluorophore location. In 2005, Lidke et al. used independent component analysis (ICA) to localize the fluorophores in quantum dot images. Their procedure exploits the blinking of dots that remain stationary from image to image. But the addition of a chemical called β -mercaptoethanol (BME) prevents our quantum dots from blinking (Hohng and Ha, 2004), and the dots in our images do not remain stationary across frames because the dots are bound to moving kinesin motors.

Thus, we elected to extend our previous GFP model in an effort to localize quantum dots in a manner appropriate to our experimental setup and goals. In this article, we present a new sampling model that allows our unified likelihood-based procedure to be applied to these quantum dot images of motor assays as well as GFP images.

2. Sampling Model for a Quantum Dot Image

In a previous study, we presented the following sampling model for a GFP image (Hughes et al., 2009). A microscope slide was represented as a rectangular region $T \subset \mathbb{R}^2$, with N , a Poisson random field on T , representing the light emission from the slide (Sarpeshkar, Delbruck, and Mead, 1993). The intensity function for N consists of background fluorescence of magnitude B and a sum of bivariate Gaussian functions, one for each fluorophore. (A fluorophore's emission pattern follows an Airy function, but the Airy function is approximated well by a Gaussian function (Thompson, Larson, and Webb, 2002).) Hence, any Borel set $R \subset T$ is a Poisson random variable with expectation

$$\mathbb{E}N(R) = \int \int_R B + \sum_{j=1}^J g_j(x, y) dx dy,$$

where

$$g_j(x, y) = A_j \cdot \exp\left(-\frac{(x - s_j)^2 + (y - t_j)^2}{S^2}\right)$$

and J is the number of fluorophores. Note that g_j , which represents the j th fluorophore, is the Gaussian function centered at (s_j, t_j) and with “height” A_j and “spread” S .

The light emitted from the slide is collected by the pixels of a CCD camera. We represent the pixels by partitioning T into a uniform grid with each cell d nanometers on a side, thus arriving at, say, n pixels, R_1, \dots, R_n , and their corresponding random variables, $N(R_1), \dots, N(R_n)$. To ease notation and description, we will henceforth let $Z_i = N(R_i)$ for $i = 1, \dots, n$ and use the term “pixel” to refer to the random variable Z_i rather than to the region R_i . Thus, for a given pixel Z_i with center (x_i, y_i) ,

$$\mathbb{E}Z_i = \int_{y_i-d/2}^{y_i+d/2} \int_{x_i-d/2}^{x_i+d/2} B + \sum_{j=1}^J g_j(x, y) dx dy,$$

which is approximately equal to

$$\left(B + \sum_j A_j \cdot \exp\left(-\frac{(x_i - s_j)^2 + (y_i - t_j)^2}{S^2}\right) \right) d^2.$$

We reparameterize so that B and the A_j absorb the constant d^2 . This gives

$$\mathbb{E}Z_i \approx B + \sum_j A_j \cdot \exp\left(-\frac{(x_i - s_j)^2 + (y_i - t_j)^2}{S^2}\right) = \mu_i. \quad (1)$$

The intensity of the background fluorescence is sufficiently large to allow for a normal approximation to the Poisson, which implies that

$$Z_i \sim \mathcal{N}(\mu_i, \mu_i).$$

However, the stochasticity in an image may not be limited to Poisson noise. Randomness may also arise from the camera system due to signal quantization and dark current, for example (Bobroff, 1986; Thompson et al., 2002). This source of variation is modeled as Gaussian white noise. Thus we arrive at the approximate sampling model for pixel i

$$Z_i \sim \mathcal{N}(\mu_i, \mu_i) + \mathcal{N}(0, \sigma^2) \stackrel{d}{=} \mathcal{N}(\mu_i, \mu_i + \sigma^2), \quad (2)$$

and is independent of the other pixels because the underlying process is Poisson and the instrumentation error is independent.

Maximum likelihood estimation based on (2) works well for certain GFP images, but our quantum dot images depart from (2) in two ways: (i) the variance may not be equal to the shifted mean, and (ii) a quantum dot image may exhibit salt-and-pepper noise, so called because it is present or absent at random (González and Woods, 2008). More specifically, the variance function for our quantum dot images is unknown but Poisson-like in that the variance changes with the mean. And the salt-and-pepper noise is exponentially distributed. These differences led us to formulate a two-component mixture model for the pixels of these quantum dot images (McLachlan and Peel, 2000). The mixture density for Z_i is given by

$$f_i(z_i) = (1 - \pi) f_{\mathcal{N}}^i(z_i) + \pi (f_{\mathcal{N}}^i(z_i) * f_{\mathcal{E}}(z_i)), \quad (3)$$

where $f_{\mathcal{N}}^i$ denotes the normal density corresponding to pixel i , $f_{\mathcal{E}}$ refers to an exponential density, $*$ is convolution, and

$\pi \in (0, 1)$. The normal-exponential convolution density accommodates the exponential error. This density is given by

$$\begin{aligned} f_{\mathcal{N} * \mathcal{E}}(z) &= f_{\mathcal{N}}(z) * f_{\mathcal{E}}(z) \\ &= \frac{1}{\lambda} \exp\left(\frac{v}{2\lambda^2} + \frac{\mu - z}{\lambda}\right) \left(1 - \Phi\left(\frac{\mu - z}{\sqrt{v}} + \frac{\sqrt{v}}{\lambda}\right)\right) \end{aligned}$$

for a normal with mean μ and variance v and an exponential with mean λ , where Φ is the standard normal cdf.

Retaining (1) and combining it with the new error model gives the new expectation for pixel i :

$$\mathbb{E}(Z_i | W_i) = B + \sum_{j=1}^J g_j(x_i, y_i) + W_i \lambda = \mu_i + W_i \lambda,$$

where the W_i are i.i.d. Bernoulli random variables that indicate the presence or absence of the exponential error, which has mean λ .

Since the variance function for the normal component is unknown, we model it as a function of the mean, $v(\mu)$. This implies that pixel i is approximately distributed as

$$Z_i \sim \mathcal{N}(\mu_i, v(\mu_i)) + W_i \mathcal{E}(\lambda), \quad (4)$$

conditional on W_i and independent of the other pixels. Note that this model can include the Gaussian white noise from (2) through $v(\mu)$. Thus (4) can be viewed as an extension of the previous model, (2).

3. Parameter Estimation, Model Selection, and Standard Errors

We estimate the parameters from (3) using a maximum likelihood technique. An initial training run is used to estimate the unknown variance function, and the estimated variance function is used during estimation of the parameters for subsequent images. (One could of course estimate the variance function for each image, but this would be too computationally expensive when processing hundreds or thousands of images.) For each image, we fit a collection of candidate models corresponding to a sequence of dot counts. Since the candidate models are nested, we use likelihood ratio tests to do model selection. After the number of dots has been determined, we use likelihood inference for the locations of the dots.

3.1 Parameter Estimation

The log-likelihood corresponding to (3) is given by

$$\ell_n(\boldsymbol{\theta} | \mathbf{Z}) = \sum_{i=1}^n \log((1 - \pi) f_{\mathcal{N}}^i(Z_i | \boldsymbol{\psi}_1) + \pi f_{\mathcal{N} * \mathcal{E}}^i(Z_i | \boldsymbol{\psi}_2)),$$

where $\boldsymbol{\psi}_1 = (B, S, A_1, s_1, t_1, \dots, A_J, s_J, t_J)'$ are the parameters for the mean of the normal, $\boldsymbol{\psi}_2 = (\boldsymbol{\psi}'_1, \lambda)'$, and $\boldsymbol{\theta}$ is the full parameter vector, $(\boldsymbol{\psi}'_2, \pi)'$. We minimize $-\ell_n$ in $\boldsymbol{\theta}$ to obtain the maximum likelihood estimator, $\hat{\boldsymbol{\theta}}$. Because the likelihood is a rather complex function of the parameters, an analytical approach to optimization seems infeasible, and so we optimize the likelihood numerically. Since we may wish to process a stack of images numbering in the hundreds or thousands, it is important that the optimization be efficient, and so we suggest a quasi-Newton algorithm. The quasi-Newton approach is less expensive computationally than Newton’s method because the former does not require computation of the Hessian

matrix of second derivatives of the objective function (Dennis and More, 1977). We found the likelihood function to be lacking in curvature with respect to A and found that rescaling the quasi-Newton optimization prevents premature convergence.

3.2 Variance-Function Estimation

We estimate the variance function, v , using B-splines (De Boor, 2001). That is, we express v as $v(\mu) = \sum_{j=1}^N b_j B_j(\mu) = \Phi(\mu)\mathbf{b}$, where $\Phi = (B_1, \dots, B_N)$ is a suitable spline basis and $\mathbf{b} = (b_1, \dots, b_N)'$ are the spline coefficients. Since a quantum dot image may exhibit shades of gray in the range 0 to 65,535, we place boundary knots at these end points. Interior knots are equally spaced. Initially, the minimum four degrees of freedom are assumed and the quasi-Newton algorithm is run until convergence, first estimating the mean surface with the variance surface assumed known, then estimating the variance surface with the mean surface assumed known. After convergence, the likelihood is computed and saved along with the estimated spline coefficients. Then five degrees of freedom are assumed, optimization is repeated, and a second likelihood is computed. This process continues until the likelihood decreases from one iteration to the next, at which time the previous estimate of \mathbf{b} is selected as $\hat{\mathbf{b}}$ so that $\hat{v} = \Phi(\hat{\mu})\hat{\mathbf{b}}$ will serve as the variance function during the processing of subsequent images, where $\hat{\mu}$ refers to the estimated normal mean surface for the image in question. We note that, in principle, a positivity constraint should be imposed on the estimated variance function, but we found this to be unnecessary in practice.

The selection of a suitable training image is paramount. Ideally, the training image will have many dots, the number of which is known. At least several of the dots should have pixel values approaching 65,535. The image should have

plenty of pixels that are at least as bright as any pixel contained in the images to be processed subsequent to training. A training image that does not satisfy these conditions may contain too little information about the variance function, in which case the estimated variance function will not be useful for processing subsequent images. Although estimation of v could employ more than one training image, our simulations show that a single suitable image is sufficient. The images in Figure 2 show the results of training on two simulated images, each of which had 10 dots. The first image had too few observations in the upper part of the intensity range, and so the estimate of v is quite poor in that region. The 10 dots in the second image, however, were about equally bright and had pixel values near 60,000.

If an unsuitable image was selected for training, the fit diagnostics described in Section 5 should, when applied to a few later images, indicate that variance-function estimation failed with respect to some images in the stack.

3.3 Model Selection, *i.e.*, Counting the Dots

Likelihood optimization tends to be sensitive to initial conditions, and so a preliminary application of ordinary least squares (OLS) provides reasonable starting values for the likelihood-optimization algorithm. The OLS sub-algorithm also uses a naive information criterion to count the quantum dots. (See Konishi and Kitagawa (2008) for an introduction to information criteria.) During the subsequent likelihood-optimization phase, the count is finalized using likelihood ratio tests.

The information criterion used to get an initial count is given by

$$IC^{(J)} = n \log(RSS/n) + 3J \sqrt{n},$$

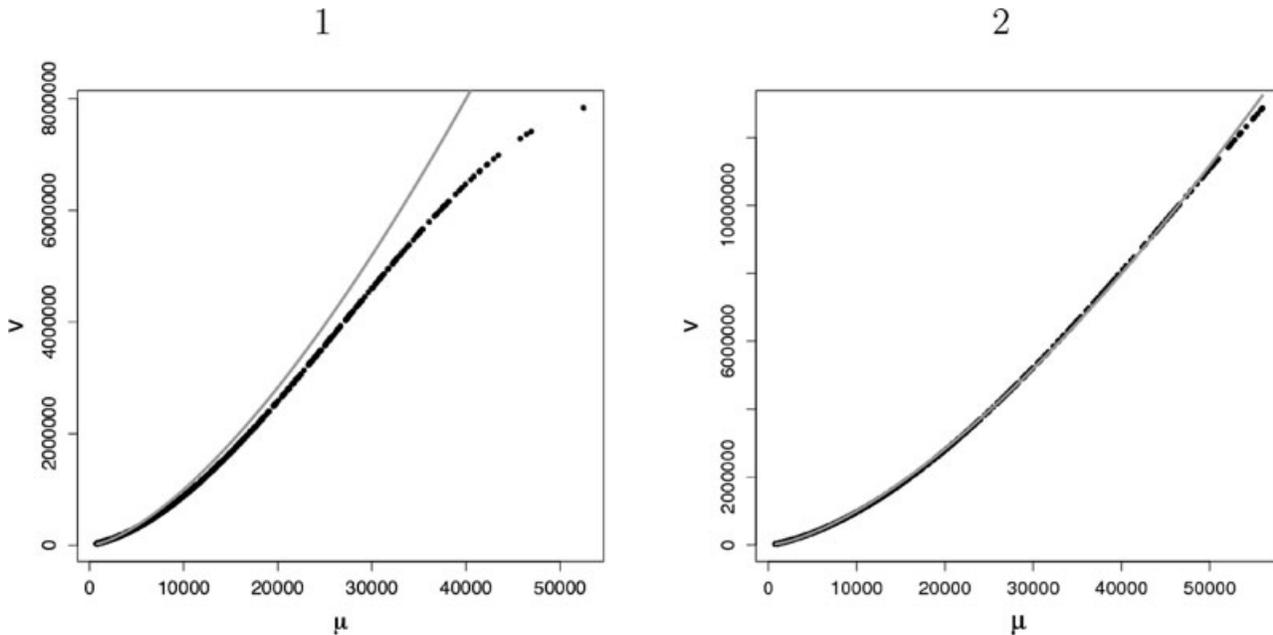


Figure 2. The estimated variance function and the true variance function for two simulated images (not shown). For each plot, the true variance function, $v(\mu)$, is shown in gray while the estimated variance function, $\hat{v}(\mu)$, is shown in black. The image corresponding to Panel 1 permitted only poor estimation because it had too few pixel values near the maximum intensity. The image corresponding to Panel 2 allowed for a very good approximation.

where J is the (assumed) number of dots, n is the number of pixels, RSS is the residual sum of squares, and $3J$ is the number of free parameters devoted to dots (A_j, s_j , and t_j for each dot). Note that IC is an increasing function of RSS and J , which implies that IC rewards a better fit (smaller RSS) and penalizes more free parameters. Initially, zero dots are assumed, and iteration of OLS goes on until $IC^{(J)} > IC^{(J-1)}$, which indicates that the image should contain $J - 1$ dots or less.

The IC used here is a non-standard information criterion. We found that the Bayesian Information Criteria (BIC) also works but does not penalize additional parameters sufficiently. This can allow the OLS stage of the algorithm to significantly overestimate the correct number of dots. Since a typical image contains tens of thousands of pixels, this overfitting causes excess and costly computation during the MLE stage of the algorithm. Our simulations showed that replacing BIC's $\log n$ with \sqrt{n} alleviates overfitting. It is important, though, to keep in mind that no information criterion based on OLS is truly appropriate here because OLS assumes the wrong sampling model.

Because it is based on a misspecified sampling model, the information criterion cannot provide a reliable final count of the dots in an image. Thus the top-level algorithm starts at the OLS-based count and works backward, computing at each step the likelihood ratio statistic

$$G^2_{(J,J+1)} = -2(\ell_n(\hat{\theta}_J | \mathbf{Z}) - \ell_n(\hat{\theta}_{J+1} | \mathbf{Z})).$$

Since we assume that v is known, the candidate models for a given image are nested. This implies that $G^2_{(J,J+1)}$ should be approximately χ^2 distributed with three degrees of freedom (Shao, 2003). When $G^2_{(J,J+1)}$ becomes significant we conclude that the image contains $J + 1$ dots. If we find no significant G^2

we conclude that the image contains no dots (e.g., when the specimen has moved out of frame). In the simulation studies presented here, the procedure always gave the correct count, but a sufficiently small $\min\{A_j\}$ will of course cause the occasional miscount.

3.4 Inference for Dot Locations

According to standard likelihood theory, a maximum likelihood estimator is asymptotically Gaussian distributed provided that certain regularity conditions are satisfied (Lehmann, 1999). The plots in Figure 3 show that our estimator is indeed approximately normally distributed for a typical image. The plots were produced from 1000 simulated 10,000-pixel images. Each image had a single dot located at $(s_1, t_1) = (2347, 5143)$, and $(B, A_1, S, \lambda, \pi) = (800, 40000, 200, 500, 0.5)$. The true variance function was $v(\mu) = \mu^{3/2}$, and the study used the estimated variance function from the second plot of Figure 2.

Although the approximate covariance for $\hat{\theta}$, which we denote $\hat{\Sigma}$, can be used to compute an approximate confidence interval for any element of θ , we consider only the dot locations, which are of primary interest to experimentalists. For the center of dot j , we have

$$\hat{\mathbf{c}}_j = (\hat{s}_j, \hat{t}_j)' \sim \mathcal{N}(\mathbf{c}_j = (s_j, t_j)', \hat{\Sigma}_j),$$

where $\hat{\Sigma}_j$ contains the elements of $\hat{\Sigma}$ corresponding to dot j . This implies the following approximate $(1 - \alpha)100\%$ confidence ellipse for \mathbf{c}_j (Ravishanker and Dey, 2002):

$$\{\mathbf{c}_j : (\hat{\mathbf{c}}_j - \mathbf{c}_j)' \hat{\Sigma}_j^{-1} (\hat{\mathbf{c}}_j - \mathbf{c}_j) \leq \chi^2_{1-\alpha, 2}\}.$$

If we wish to control the image-wide error rate, that is, if we desire a collection of ellipses that enclose all J dots at

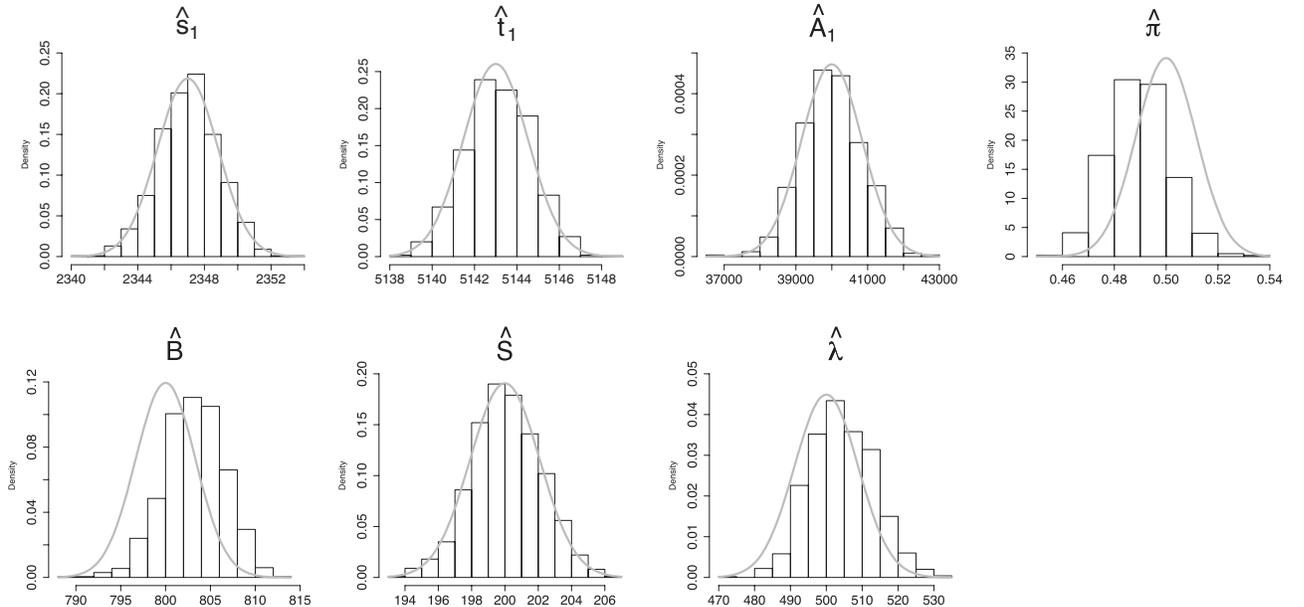


Figure 3. The estimators are approximately normally distributed. Note that $\hat{\pi}$ appears to be biased downward. Since $\hat{\pi}$ is negatively correlated with \hat{B} and $\hat{\lambda}$, those estimators trend upward. But the biases are slight (2%, 0.5%, and 0.4%, respectively), and π, B , and λ are nuisance parameters.

confidence level $(1 - \alpha)100\%$, we can apply the well-known Bonferroni correction (Shao, 2003):

$$\{\mathbf{c}_j: (\hat{\mathbf{c}}_j - \mathbf{c}_j)' \hat{\Sigma}_j^{-1} (\hat{\mathbf{c}}_j - \mathbf{c}_j) \leq \chi_{1-\alpha/J,2}^2\},$$

for $j = 1, \dots, J$.

We conducted a second 1000-image simulation study to determine whether these approximate confidence ellipses have the correct coverage probabilities, and also to compare the performance of our mixture model with the GFP model and with the homoscedastic normal model, that is, the model corresponding to ordinary least squares estimation, the standard method for localization (Thompson et al., 2002). Each image had 10,000 pixels and three dots. The coverage probabilities for the dots were 94.4%, 95.3%, and 95.6% for the mixture model; 10%, 8.9%, and 10.9% for the GFP model; and 21.6%, 19.8%, and 20.5% for OLS. The corresponding mean squared errors (for $\hat{s}_j, j = 1, 2, 3$) were 2.97, 2.74, and 3.06 for the mixture model; 5.27, 5.33, and 4.97 for the GFP model; and 5.81, 5.8, and 5.39 for OLS. Our mixture model permits reliable inference for this type of data while the GFP and homoscedastic models show poor coverage rates. And the mixture model also exhibits mean squared errors that are approximately half as large as for the other models. As we mentioned previously, localization to within a few nanometers is paramount for motor assays.

4. Robustness Study

To show that our procedure is capable of fitting images with many dots, we simulated and analyzed a 12-dot image. In the interest of brevity, Table 1 shows the parameter values, point estimates, and standard errors for just a few of the dots. Note that every approximate 95% confidence interval covered the

Table 1

Selected parameter values, point estimates, and standard errors for a simulated image with 12 quantum dots. The point estimates are very close to the true values, especially for those parameters of primary interest to experimentalists, that is, the dot locations $(s_1, t_1), (s_4, t_4)$, and (s_7, t_7) . The standard errors pertaining to dot locations are all smaller than 2 nm. We also note that every approximate 95% confidence interval covered the true parameter value, for all 12 dot locations and all nuisance parameters.

Parameter	True value	Estimate	\widehat{SE}
s_1	644	644.009	1.645
t_1	2053	2054.208	1.726
A_1	40810	41799.294	838.827
s_4	2691	2689.753	1.750
t_4	553	553.685	1.768
A_4	40716	42092.421	876.943
s_7	3204	3203.680	1.900
t_7	1416	1416.437	1.923
A_7	35907	33564.893	732.839
S	130	130.355	0.411
B	800	801.162	3.707
λ	600	593.135	11.417
π	0.5	0.501	0.012

truth, for all 12 dot locations and all of the nuisance parameters.

Quantum dot images sometimes exhibit certain atypical features that could cause inference to suffer. The most common of these features are (i) especially dim dots, (ii) dots that are only partially contained in the image, and (iii) dots that are close together. We conducted a simulation study to investigate the effects of these scenarios on the coverage probabilities. As in the studies described above, we applied our procedure to 1000 10,000-pixel images for each of the three scenarios, and each image contained three quantum dots. For the DIM scenario, one of the three dots has $A = 5000$, which represents an eightfold decrease relative to the other two dots. For the PARTIAL scenario, one of the three dots is located only 75 nm from the image's edge. And for the CLOSE scenario, two of the dots are located just 400 nm apart. The coverage rates for the study are shown in Table 2.

Although we have not seen images with extremely sparse or extremely dense salt-and-pepper noise, images with these characteristics are possible, and so we thought it prudent to determine the effects of these scenarios on the coverage rates. Each study used 1000 10,000-pixel images, and we set π equal to 0.05 and 0.95. The coverage rates for these scenarios (SPARSE and DENSE) are given at the bottom of Table 2. A nearly empty component affects the coverage only slightly.

For our final robustness study, we investigated the sensitivity of our procedure to the Gaussian assumption by replacing the normal component of the mixture with an appropriately

Table 2

Results from our robustness study. For each scenario—see Section 4 for detailed descriptions—we simulated 1000 images, with each image having three dots. The first column of the table lists the various scenarios. The second column gives the types of the dots, three for each scenario. And the final column reports the coverage rates of approximate 95% confidence ellipses for the locations of the three dots. Note that the rates are generally close to 95%, even for images with atypical features, which indicates that our procedure permits reliable and robust inference for dot location.

Scenario	Dot type	Coverage rate
DIM	Typical	95.3%
	Typical	94.9%
	Dim	95.1%
CLOSE	Typical	94.6%
	Close	95.6%
	Close	94.9%
PARTIAL	Typical	94.8%
	Typical	94.8%
	Partial	95.2%
SPARSE	Typical	94.5%
	Typical	94.2%
	Typical	95%
DENSE	Typical	95%
	Typical	93.5%
	Typical	95.8%
HEAVY TAILS	Typical	98.5%
	Typical	98.7%
	Typical	95.8%

Table 3

Parameter estimates for an experimentally observed quantum dot image. The second and third columns give the point estimates and standard errors obtained by fitting our quantum dot model. The final two columns give the corresponding results obtained by fitting the GFP model.

Parameter	Est.	\widehat{SE}	GFP Est.	GFP \widehat{SE}
s_1	4750.218	1.529	4784.93	0.225
t_1	1213.598	1.194	1219.11	0.228
A_1	23535	201.787	24313	48.5
s_2	1043.188	1.164	1030.49	0.278
t_2	4829.266	1.368	4845.82	0.271
A_2	16,764.215	219.372	17949.6	39.7
S	138.564	0.787	173.661	0.151
B	831.924	0.662	1478.72	0.45
λ	809.083	10.357	NA	NA
π	0.85	0.005	NA	NA

scaled and relocated t_3 -distributed component. More specifically, we simulated 1000 10,000-pixel images according to

$$Z_i \sim \mu_i + \sqrt{v(\mu_i)/3} t_3 + W_i \mathcal{E}(\lambda).$$

In our study, this heavy-tailed component increased the coverage rates (Table 2, scenario HEAVY TAILS).

5. Analysis of an Experimentally Collected Quantum Dot Image

Now we apply our procedure to an experimentally collected quantum dot image (Figure 1) and describe appropriate fit diagnostics. We note that fitting this image took only 35.5 seconds on a 2.8 GHz Intel Xeon Harpertown processor. The operating system was Snow Leopard (with 64-bit kernel), and the programming language was R. The resulting estimates and approximate standard errors are shown in Table 3. Since the approximate 95% confidence regions for the dot locations, (s_1, t_1) and (s_2, t_2) , provided by the quantum dot model do not overlap those provided by the GFP model, at least one of

these fits has produced invalid inference for the dot locations. Moreover, the margin of error (roughly two standard deviations) for the arguably more reliable quantum dot model is more than 2 nm while the margin of error for the GFP model is less than one. This is an important distinction when judging the steps of a kinesin motor, which are roughly 8 nm in length. The GFP method gives a false sense of precision that cannot be relied on in this setting.

We diagnose the fit as follows. (4) implies that we can view this problem as an incomplete-data problem, where \mathbf{Z} is observed but $\mathbf{W} = (W_1, \dots, W_n)'$ is hidden (McLachlan and Peel, 2000). To construct fitted values, and thus residuals, we need to estimate \mathbf{W} . A natural estimator of $\gamma_i = \mathbb{E}(W_i | \mathbf{Z}, \boldsymbol{\theta})$ is the likelihood that Z_i came from $f_{N^* \mathcal{E}}^i$ (McLachlan and Krishnan, 1997):

$$\hat{\gamma}_i = \frac{\hat{\pi} f_{N^* \mathcal{E}}^i(Z_i | \hat{\boldsymbol{\psi}}_2)}{(1 - \hat{\pi}) f_{N^* \mathcal{E}}^i(Z_i | \hat{\boldsymbol{\psi}}_1) + \hat{\pi} f_{N^* \mathcal{E}}^i(Z_i | \hat{\boldsymbol{\psi}}_2)}.$$

We compute the $\hat{\gamma}_i$ and threshold them to arrive at $\hat{W}_i = \mathbf{1}_{[0.5, 1]}(\hat{\gamma}_i)$, where $\mathbf{1}$ denotes the indicator function. Then the fitted value for pixel i is $\hat{Z}_i = \hat{\mu}_i + \hat{W}_i \hat{\lambda}$, and the estimated variance is $\hat{\sigma}_i^2 = \hat{v}(\hat{\mu}_i) + \hat{W}_i \hat{\lambda}^2$. Recall that $\hat{v}(\hat{\mu}_i) = \boldsymbol{\Phi}_i \hat{\mathbf{b}}$, where $\boldsymbol{\Phi}_i$ is the i th row of $\boldsymbol{\Phi}(\hat{\boldsymbol{\mu}})$.

Now, the i th standardized residual is given by

$$r_i = \frac{Z_i - \hat{Z}_i}{\hat{\sigma}_i} = \frac{Z_i - \hat{\mu}_i - \hat{W}_i \hat{\lambda}}{\sqrt{\hat{v}(\hat{\mu}_i) + \hat{W}_i \hat{\lambda}^2}}.$$

If $\hat{W}_i = 0$, r_i should be approximately standard normally distributed. If $\hat{W}_i = 1$, the (approximate) distribution of r_i should be

$$r_i \sim \mathcal{N}\left(0, \frac{v(\mu_i)}{v(\mu_i) + \lambda^2}\right) + \mathcal{E}\left(-1, \sqrt{\frac{\lambda^2}{v(\mu_i) + \lambda^2}}\right).$$

$\mathcal{E}(\alpha, \beta)$ denotes an exponential random variable with location parameter α and scale parameter β ; the density function is $f_{\mathcal{E}}(r | \alpha, \beta) = \beta^{-1} \exp(-(r - \alpha)/\beta) \mathbf{1}_{[\alpha, \infty)}(r)$, $\beta > 0$. If the i th pixel is part of the image background, $v(\mu_i) \ll \lambda^2$, in which

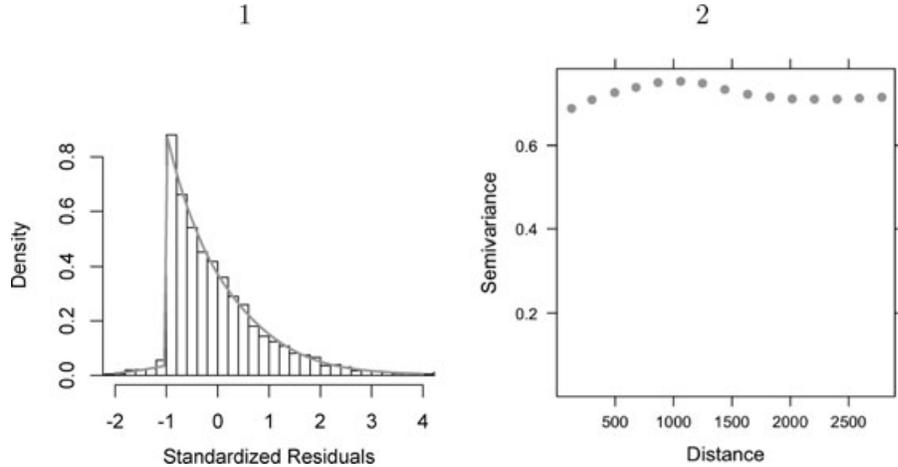


Figure 4. Panel 1: A histogram of the standardized residuals for the image in Figure 1. The theoretical density, with $\pi = \hat{\pi}$, is shown in gray. The residuals are approximately distributed according to (5), which indicates a good fit. Panel 2: An empirical-variogram plot for the standardized residuals. The near constancy indicates the spatial independence specified by the sampling model. Distance is given in nanometers.

case $r_i \sim \mathcal{E}(-1, 1)$. Since nearly all pixels are background pixels, the following mixture density gives the approximate distribution of the standardized residuals:

$$f(r) = (1 - \pi)\phi(r) + \pi f_{\mathcal{E}}(r | -1, 1), \quad (5)$$

where ϕ is the standard normal density function. Hence, we have the following approximate distribution for r_i , conditional on W_i and independent of the other residuals:

$$r_i \sim (1 - W_i)\mathcal{N}(0, 1) + W_i\mathcal{E}(-1, 1).$$

We examine the distribution of the standardized residuals by comparing a histogram to the theoretical density given in (5). And we check for spatial correlation among the standardized residuals by plotting empirical variograms (Cressie and Hawkins, 1980). The diagnostics for the real quantum dot image, which are shown in Figure 4, indicate a good fit.

6. Discussion

The method described in this article allows for fast, nearly automatic analysis of quantum dot images of kinesin motor assays. The use of a likelihood framework allows for efficient estimation of parameters and standard errors, and our model selection technique automatically determines the number of dots in an image. This efficiency and automation (after an initial set up) make our approach attractive for processing large stacks of images, which opens the possibility of developing methods for accurately and precisely tracking active motors across the frames of a “movie.”

Our simulation studies show that inference for dot location is robust to dim, partial, or close dots, and also to sparse or dense salt-and-pepper noise and heavy tails. The proposed fit diagnostics are straightforward to prepare and to interpret, which will allow experimentalists to easily verify that the model assumptions are valid for a given experimental setup.

REFERENCES

- Arya, H., Kaul, Z., Wadhwa, R., Taira, K., Hirano, T., and Kaul, S. (2005). Quantum dots in bio-imaging: Revolution by the small. *Biochemical and Biophysical Research Communications* **329**, 1173–1177.
- Bobroff, N. (1986). Position measurement with a resolution and noise-limited instrument. *Review of Scientific Instruments* **57**, 1152–1157.
- Cressie, N. and Hawkins, D. M. (1980). Robust estimation of the variogram: I. *Mathematical Geology* **12**, 115–125.
- De Boor, C. (2001). *A Practical Guide to Splines*. New York: Springer.
- Dennis, J. and Moré, J. (1977). Quasi-newton methods, motivation and theory. *SIAM Review* **19**, 46–89.
- González, R. C. and Woods, R. E. (2008). *Digital Image Processing*. Reading, Massachusetts: Prentice Hall.
- Hohng, S. and Ha, T. (2004). Near-complete suppression of quantum dot blinking in ambient conditions. *Journal of the American Chemical Society* **126**, 1324–1325.
- Hughes, J., Fricks, J., and Hancock, W. (2009). Likelihood inference for particle location in fluorescence microscopy. To appear in *Annals of Applied Statistics*.
- Konishi, S. and Kitagawa, G. (2008). *Information Criteria and Statistical Modeling*. New York: Springer.
- Lehmann, E. L. (1999). *Elements of Large-Sample Theory*. New York: Springer.
- Lidke, K., Rieger, B., Jovin, T., and Heintzmann, R. (2005). Superresolution by localization of quantum dots using blinking statistics. *Optics Express* **13**, 7052–7062.
- McLachlan, G. J. and Krishnan, T. (1997). *The EM Algorithm and Extensions*. Hoboken, New Jersey: John Wiley and Sons.
- McLachlan, G. J. and Peel, D. (2000). *Finite Mixture Models*. Hoboken, New Jersey: John Wiley and Sons.
- Ravishanker, N. and Dey, D. K. (2002). *A First Course in Linear Model Theory*. Boca Raton, Florida: Chapman & Hall/CRC.
- Rost, F. (1992). *Fluorescence Microscopy*. Cambridge, U.K.: Cambridge University Press.
- Sarpeshkar, R., Delbruck, T., and Mead, C. (1993). White noise in mos transistors and resistors. *Circuits and Devices Magazine, IEEE* **9**, 23–29.
- Shao, J. (2003). *Mathematical Statistics*. New York: Springer.
- Thompson, R. E., Larson, D. R., and Webb, W. W. (2002). Precise nanometer localization analysis for individual fluorescent probes. *Biophysical Journal* **82**, 2775–2783.
- Vale, R. (2007). Introduction to motor proteins. <http://www.ibio-seminars.org/vale/vale1.shtml>.
- Yildiz, A. and Selvin, P. R. (2005). Fluorescence imaging with one nanometer accuracy: Application to molecular motors. *Accounts of Chemical Research* **38**, 574–582.

Received October 2009. Revised April 2010.

Accepted May 2010.