Gaussian profile estimation in two dimensions

Nathan Hagen[1,*] and Eustace L. Dereniak[2]

[1]Fitzpatrick Institute for Photonics, Duke University, Durham, North Carolina 27708, USA

[2]College of Optical Sciences, University of Arizona, Tucson, Arizona 85721, USA

*Corresponding author: nhagen@optics.arizona.edu

Received 31 January 2008; revised 7 June 2008; accepted 20 October 2008;
posted 5 November 2008 (Doc. ID 92236); published 15 December 2008

We extend recent results for estimating the parameters of a one-dimensional Gaussian profile to two-dimensional profiles, deriving the exact covariance matrix of the estimated parameters. While the exact form is easy to compute, we provide a set of close approximations that allow the covariance to take on a simple analytic form. This not only provides new insight into the behavior of the estimation parameters, but also lays a foundation for clarifying previously published work. We also show how to calculate the parameter variances for the case of truncated sampling, where the profile lies near the edge of the array detector. Finally, we calculate expressions for the bias in the classical formulation of the problem and provide an approach for its removal. This allows us to show how the bias affects the problem of choosing an optimal pixel size for minimizing parameter variances. © 2008 Optical Society of America

OCIS codes: 000.5490, 100.2960, 300.3700.

1. Introduction

In a previous paper [1], we outlined an approach for using maximum-likelihood estimation (MLE) of one-dimensional (1D) Gaussian profile parameters from data corrupted by noise. Here we extend the method to the two-dimensional (2D) case—estimating a 2D Gaussian profile from an image—to incorporate added parameters while retaining a fast algorithm. This estimation procedure can be useful in fields such as Gaussian beam characterization [2], astrometry [3], wavefront sensing [4], bioimaging [5,6], and calibrations of computational sensors, such as computed tomography instruments [7].

The initial approach we use here makes no assumptions on the sampling—uniform or nonuniform—of the profile. Bad samples, such as from inoperative pixels, are easily taken care of by simply deleting them from the data set. The Gaussian function being measured need not even have most of its volume in the sampled region (i.e., its peak may be located off the edge of the detector array), though the nonlinear optimization procedure may have difficulty in locating the global maximum. A good initial guess is required in such a case.

The estimation procedure we use is the same as that for the 1D problem, namely:

1. Assuming Gaussian- or Poisson-distributed additive noise, we first form the log-likelihood function $\ell = \ln \mathrm{pr}(\mathbf{g}|\boldsymbol{\theta})$ and then calculate its gradient (the "score") $\boldsymbol{\nabla}\ell$ and its Hessian matrix $\mathbf{H}$.
2. These three functions are used in a nonlinear optimization routine (such as Newton iteration) to solve for the parameter set $\boldsymbol{\theta}$, which maximizes the likelihood, via

$$\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} - (\mathbf{H}^{(k)})^{-1}\boldsymbol{\nabla}\ell^{(k)},$$

where a superscript (k) indicates the iteration index and $\mathbf{H}^{(k)}$ is the Hessian matrix of ℓ evaluated at $\boldsymbol{\theta}^{(k)}$.
3. To determine the accuracy of the estimates, we use the Cramér–Rao bound to calculate the covariance matrix $\mathbf{K}$ of parameter estimators, obtained from the exact Fisher information matrix $\mathbf{F}$ by $\mathbf{K} = \mathbf{F}^{-1}$.
4. Alternatively, an analytic approximation to $\mathbf{K}$ can be obtained by first approximating $\mathbf{F}$ and then

0003-6935/08/366842-10$15.00/0

performing an analytic inverse on the resulting simplified matrix.

5. Since each model is actually biased due to the conventional approximation of the pixel response functions as δ functions, we also provide an analytic expression for the bias under the rect-sampling model. Bias correction can then be performed by using the ML estimate to calculate the bias, which is then subtracted from the raw data for a second iteration of the estimation algorithm.

For each of the 2D Gaussian models discussed in this paper—the symmetric Gaussian (circular cross section), the asymmetric separable Gaussian (elliptical cross section), and the general case (elliptical cross section unaligned to coordinate axes)—we provide the expressions for the above steps.

2. Symmetric 2D Gaussian Profile Model

The simplest type of 2D Gaussian profile is a rotationally symmetric object function, which we model as

$$f(\mathbf{r}) = Ae^{-(\mathbf{r}-\bar{\mathbf{r}})^2/2w^2},$$

where A is the peak amplitude, $\bar{\mathbf{r}} = (\bar{x}, \bar{y})$ is the position of the peak (the "center"), and w is the Gaussian width. The object is a continuous function, sampled by the detection process to produce a discrete data vector $\bar{\mathbf{g}}$, where

$$\begin{aligned}\bar{g}_m &= \delta_x\delta_y Q_m \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x)\delta(x-x_m)\delta(y-y_m)\mathrm{d}x\mathrm{d}y \\ &= \delta_x\delta_y Q_m A e^{-(\mathbf{r}_m-\bar{\mathbf{r}})^2/2w^2}. \end{aligned} \quad (1)$$

The 2D image data is thus mapped into a 1D vector, where $\delta(\cdot)$ is the Dirac delta function, m is the pixel index ($1 \le m \le M$), $\mathbf{r}_m$ gives the abscissa position vector for the center of pixel m, and $\delta_x\delta_y$ gives the x and y dimensions of the pixels. An example $\bar{\mathbf{g}}$ is illustrated in Fig. 1. In Eq. (1), the quantity Q_m is the gain for pixel m, giving the number of digital counts per detected photoelectron. (Note that detector gain is often quoted as photoelectrons per digital count—the inverse of Q_m here.) In general, Q_m can vary significantly with position on the detector array.

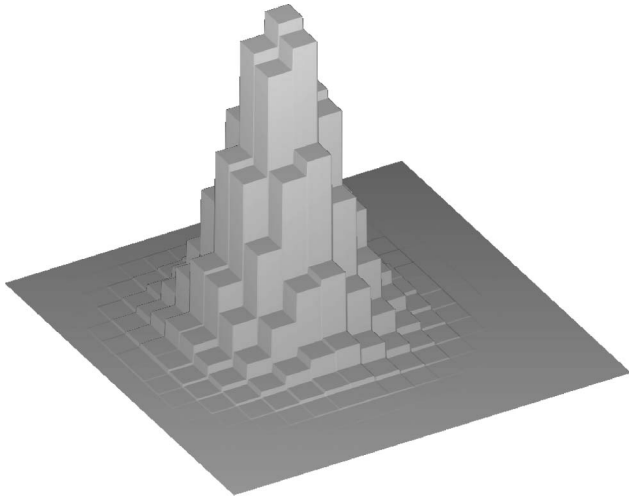

Fig. 1. Example noiseless symmetric 2D Gaussian profile, sampled on a pixel grid.

We model the measurement result $\mathbf{g}$ as the sum of a noiseless discrete data vector and a zero-mean noise vector, $\mathbf{g} = \bar{\mathbf{g}} + \mathbf{n}$, where the noise vector $\mathbf{n}$ is a Gaussian-distributed random variable such that

$$\mathrm{pr}(g_m) = \frac{1}{\sqrt{2\pi\sigma_m^2}} e^{-(g_m-\bar{g}_m)^2/2\sigma_m^2}, \quad (2)$$

where σ_m^2 is the variance of the noise at pixel m, given in units of digital counts.

The likelihood of a given parameter set $\boldsymbol{\theta} = (A, \bar{x}, \bar{y}, w)$ producing a measured image $\mathbf{g}$ is defined as

$$L(\boldsymbol{\theta}|\mathbf{g}) = \mathrm{pr}(\mathbf{g}|\boldsymbol{\theta}).$$

Taking the logarithm, $\ell = \log L$, it is easy to derive (see Eq. 6 in [1]) that the gradient has the form

$$\frac{\partial\ell}{\partial\theta_i} = \sum \frac{1}{\sigma_m^2}(g_m - \bar{g}_m)\frac{\partial\bar{g}_m}{\partial\theta_i}. \quad (3)$$

So far, we have assumed a Gaussian-distributed noise model. If we rederive the likelihood from a Poisson-noise model, we find that we again obtain Eq. (3), but with σ_m^2 replaced by $\bar{g}_m/Q_m$. Thus, although we continue to work with a likelihood function derived from a Gaussian-noise model below, it is left understood that the same results can be achieved under Poisson noise by setting $\sigma_m^2 = \bar{g}_m/Q_m$.

For the symmetric Gaussian profile, the noise model gives

$$\begin{aligned}\frac{\partial\ell}{\partial A} &= \sum \gamma_m, & \frac{\partial\ell}{\partial\bar{x}} &= \frac{A}{w^2}\sum \gamma_m\xi_m, \\ \frac{\partial\ell}{\partial\bar{y}} &= \frac{A}{w^2}\sum \gamma_m\eta_m, & \frac{\partial\ell}{\partial w} &= \frac{A}{w^3}\sum \gamma_m\rho_m^2,\end{aligned}$$

where we define

$$\begin{aligned}\xi_m &= (x_m - \bar{x}), & \eta_m &= (y_m - \bar{y}), \\ \rho_m &= \sqrt{\xi_m^2 + \eta_m^2}, & \gamma_m &= \frac{\delta_x\delta_y}{\sigma_m^2}(g_m - \bar{g}_m)Q_m E_m, \\ E_m &= e^{-(\xi_m^2+\eta_m^2)/2w^2}.\end{aligned}$$

All of the sums are taken over the pixel index, $1 \le m \le M$. The Hessian matrix is

$$\mathbf{H} = \begin{pmatrix} \frac{\partial^2 \ell}{\partial A^2} & \frac{\partial^2 \ell}{\partial A \partial \bar{x}} & \frac{\partial^2 \ell}{\partial A \partial \bar{y}} & \frac{\partial^2 \ell}{\partial A \partial w} \\ \frac{\partial^2 \ell}{\partial \bar{x} \partial A} & \frac{\partial^2 \ell}{\partial \bar{x}^2} & \frac{\partial^2 \ell}{\partial \bar{x} \partial \bar{y}} & \frac{\partial^2 \ell}{\partial \bar{x} \partial w} \\ \frac{\partial^2 \ell}{\partial \bar{y} \partial A} & \frac{\partial^2 \ell}{\partial \bar{y} \partial \bar{x}} & \frac{\partial^2 \ell}{\partial \bar{y}^2} & \frac{\partial^2 \ell}{\partial \bar{y} \partial w} \\ \frac{\partial^2 \ell}{\partial w \partial A} & \frac{\partial^2 \ell}{\partial w \partial \bar{x}} & \frac{\partial^2 \ell}{\partial w \partial \bar{y}} & \frac{\partial^2 \ell}{\partial w^2} \end{pmatrix}.$$

H is symmetric, so it is only necessary to calculate the upper triangular portion. Performing the partial derivative calculations gives, for the matrix elements,

$$H_{11} = -\delta_x^2 \delta_y^2 \sum \frac{1}{\sigma_m^2} Q_m^2 E_m^2 \qquad H_{12} = \frac{1}{w^2} \sum \Gamma_m \xi_m$$

$$H_{13} = \frac{1}{w^2} \sum \Gamma_m \eta_m \qquad H_{14} = \frac{1}{w^3} \sum \Gamma_m \rho_m^2$$

$$H_{22} = \frac{A}{w^4} \sum \Gamma_m \xi_m^2 - \gamma_m w^2 \qquad H_{23} = \frac{A}{w^4} \sum \Gamma_m \xi_m \eta_m$$

$$H_{24} = \frac{A}{w^5} \sum \Gamma_m \xi_m \rho_m^2 - 2\gamma_m \xi_m w^2$$

$$H_{33} = \frac{A}{w^4} \sum \Gamma_m \eta_m^2 - \gamma_m w^2$$

$$H_{34} = \frac{A}{w^5} \sum \Gamma_m \eta_m \rho_m^2 - 2\gamma_m \eta_m w^2$$

$$H_{44} = \frac{A}{w^6} \sum \Gamma_m \rho_m^4 - 3\gamma_m \rho_m^2 w^2,$$

where $\Gamma_m = \delta_x \delta_y \frac{g_m - 2\bar{g}_m}{\sigma_m^2} Q_m E_m$ and γ_m is as defined above.

The Fisher information matrix is obtained as the negative of the elementwise expectation of the Hessian matrix [8], with the expectation taken over the data:

$$F_{ij} = -\mathcal{E}\{H_{ij}\} = -\int \ell(\boldsymbol{\theta}) \left(\frac{\partial^2 \ell}{\partial \theta_i \partial \theta_j} \right) \mathrm{d}^M g, \qquad (4)$$

producing

$$\left.\begin{array}{llll} F_{11} & = \sum R_m & F_{23} & = \frac{A^2}{w^4} \sum R_m \xi_m \eta_m \\ F_{12} & = \frac{A}{w^2} \sum R_m \xi_m & F_{24} & = \frac{A^2}{w^5} \sum R_m \rho_m^2 \xi_m \\ F_{13} & = \frac{A}{w^2} \sum R_m \eta_m & F_{33} & = \frac{A^2}{w^4} \sum R_m \eta_m^2 \\ F_{14} & = \frac{A}{w^3} \sum R_m \rho_m^2 & F_{34} & = \frac{A^2}{w^5} \sum R_m \rho_m^2 \eta_m \\ F_{22} & = \frac{A^2}{w^4} \sum R_m \xi_m^2 & F_{44} & = \frac{A^2}{w^6} \sum R_m \rho_m^4 \end{array}\right\} \qquad (5)$$

for $R_m = \delta_x^2 \delta_y^2 Q_m^2 E_m^2 / \sigma_m^2$. Note that the notation $\mathrm{d}^M g$ expresses the differential for an M-dimensional integral over all data elements g_m.

When the Cramér–Rao bound is met (asymptotically for a large number of measurements M), the covariance matrix **K** is given by the inverse of the above matrix, which can be obtained either analytically by Cramer's rule or numerically. The analytic expressions for **K** are quite involved and so provide limited insight into the algorithm. To gain a better understanding of the relationship between the estimator variances and the object parameters, we introduce approximations that simplify the elements of **F** prior to taking its inverse. These approximations are:

1. **Flat noise**: $\sigma_m = \sigma$, i.e., all pixels share the same noise variance [9].
2. **Uniform gain**: $Q_m = Q$.
3. **Uniform sampling**: The spacing (δ_x, δ_y) between sample locations (x_m, y_m) is constant.
4. **Complete sampling**: The data region $x_{\min} \le x \le x_{\max}, y_{\min} \le y \le y_{\max}$ is sufficient such that the Gaussian $f(\mathbf{r})$ is approximately zero outside of the sampled region.
5. **The profile is well sampled**: Since each of the F_{ij} are sums of sampled Gaussian functions, we can approximate the sums with integrals and then replace the integrals with known analytic results.

With these approximations, and noting that $\delta_x = \delta\xi, \delta_y = \delta\eta$, we can write

$$\sum E_m^2 = \frac{1}{\delta_x \delta_y} \sum e^{-(\xi_m^2 + \eta_m^2)/w^2} \delta\xi \delta\eta \approx \frac{1}{\delta_x \delta_y} \int_{-\infty}^{\infty} e^{-\xi^2/w^2} \mathrm{d}\xi \int_{-\infty}^{\infty} e^{-\eta^2/w^2} \mathrm{d}\eta = \frac{\pi w^2}{\delta_x \delta_y}$$

and, similarly,

$$\left.\begin{array}{l} \sum E_m^2 \xi_m \approx 0 \\ \sum E_m^2 \xi_m \eta_m \approx 0 \\ \sum E_m^2 \xi_m^2 \approx \frac{\pi w^4}{2\delta_x \delta_y} \\ \sum E_m^2 \rho_m^2 \approx \frac{\pi w^4}{\delta_x \delta_y} \\ \sum E_m^2 \rho_m^4 \approx \frac{2\pi w^6}{\delta_x \delta_y} \end{array}\right\}. \qquad (6)$$

Constructing an approximate form of **F** from these expressions, we can analytically calculate its inverse to give the parameter covariance matrix **K**:

$$\mathbf{K}_{\mathrm{Flat}} \approx \frac{\sigma^2}{\pi \delta_x \delta_y Q^2} \begin{pmatrix} \frac{2}{w^2} & 0 & 0 & \frac{-1}{\mathrm{Aw}} \\ 0 & \frac{2}{A^2} & 0 & 0 \\ 0 & 0 & \frac{2}{A^2} & 0 \\ \frac{-1}{\mathrm{Aw}} & 0 & 0 & \frac{1}{A^2} \end{pmatrix}. \qquad (7)$$

Alternatively, if we assume that the noise variance is not the same at all pixels (i.e., not flat) but rather is determined from the Poisson distribution of the mean signal $\bar{g}_m$ at each pixel, then we can substitute $\sigma_m^2 = \bar{g}_m Q_m = \delta_x \delta_y A Q^2 E_m$ and, recalculating **F**, obtain a modified covariance matrix [10]:

$$\mathbf{K}_{\mathrm{Poisson}} \approx \frac{1}{2\pi} \begin{pmatrix} \frac{2A}{w^2} & 0 & 0 & \frac{-1}{2w} \\ 0 & \frac{1}{A} & 0 & 0 \\ 0 & 0 & \frac{1}{A} & 0 \\ \frac{-1}{2w} & 0 & 0 & \frac{1}{4A} \end{pmatrix}. \qquad (8)$$

The parameter variances can thus be written out explicitly as

$$\mathrm{var}(\hat{A}) \approx \begin{cases} 2\beta/w^2 & \text{(Flat)} \\ A/(2\pi w^2) & \text{(Poisson)} \end{cases},$$
$$\mathrm{var}(\hat{\bar{x}}) = \mathrm{var}(\hat{\bar{y}}) \approx \begin{cases} 2\beta/A^2 & \text{(Flat)} \\ 1/(2\pi A) & \text{(Poisson)} \end{cases},$$
$$\mathrm{var}(\hat{w}) \approx \begin{cases} \beta/A^2 & \text{(Flat)} \\ 1/(8\pi A) & \text{(Poisson)} \end{cases}, \tag{9}$$

where $\beta = \frac{\sigma^2}{\pi\delta_x\delta_y Q^2}$.

To calculate the volume U under the Gaussian spot, we first form its estimator. Since

$$U = A\int \mathrm{d}\xi \int \mathrm{d}\eta\, e^{-(\xi^2+\eta^2)/2w^2} = 2\pi A w^2,$$

the estimator is $\hat{U} = 2\pi\hat{A}\hat{w}^2$. Using the elements of the covariance matrix to estimate the variance of $\hat{U}$ under the flat-noise and Poisson-noise approximations ([8], p. 45–46)

$$\mathrm{var}(\hat{U}) = \frac{\partial U}{\partial \boldsymbol{\theta}^{\mathsf{T}}} K_{\boldsymbol{\theta}} \frac{\partial U}{\partial \boldsymbol{\theta}} \approx \begin{cases} 8\pi^2\beta w^2 & \text{(Flat)} \\ 2\pi A w^2 & \text{(Poisson)} \end{cases}.$$

Note, again, that these simple expressions for the parameter variances hold only when the assumptions of completely, uniformly, and well-sampled data exists.

3. Separable 2D Gaussian Profile Model

For a separable Gaussian model function with nonequal widths w_x and w_y (Fig. 2), we have, using the same nomenclature as Section 3 above,

$$f(\mathbf{r}) = Ae^{-(x-\bar{x})^2/2w_x^2} e^{-(y-\bar{y})^2/2w_y^2},$$

where $\bar{g}_m = \delta_x\delta_y Q_m f(x_m, y_m), g_m = \bar{g}_m + n(x_m, y_m)$ and the vector of parameters is now $\boldsymbol{\theta} = (A, \bar{x}, \bar{y}, w_x, w_y)$.

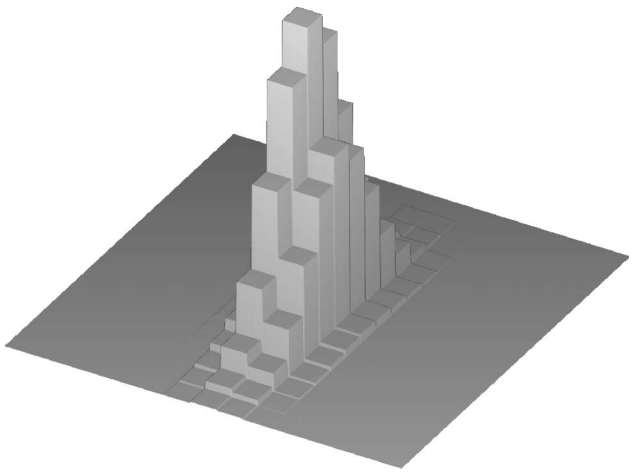

Fig. 2. Example noiseless asymmetric 2D Gaussian profile, sampled on a pixel grid.

Forming the log-likelihood function ℓ, its gradient $\nabla\ell$, and the Hessian matrix $\mathbf{H}$ for this model under additive Gaussian noise, we obtain

$$[\nabla\ell]_1 = \frac{\partial\ell}{\partial A} = \Sigma\gamma_m \qquad [\nabla\ell]_2 = \frac{\partial\ell}{\partial\bar{x}} = \frac{A}{w_x^2}\Sigma\gamma_m\xi_m$$
$$[\nabla\ell]_3 = \frac{\partial\ell}{\partial\bar{y}} = \frac{A}{w_y^2}\Sigma\gamma_m\eta_m \qquad [\nabla\ell]_4 = \frac{\partial\ell}{\partial w_x} = \frac{A}{w_x^3}\Sigma\gamma_m\xi_m^2$$
$$[\nabla\ell]_5 = \frac{\partial\ell}{\partial w_y} = \frac{A}{w_y^3}\Sigma\gamma_m\eta_m^2,$$

where $E_m = e^{-\xi_m^2/2w_x^2} e^{-\eta_m^2/2w_y^2}$, $\xi_m = (x_m - \bar{x})$, $\eta_m = (y_m - \bar{y})$, and $\gamma_m = \frac{\delta_x\delta_y}{\sigma_m^2}(g_m - \bar{g}_m)Q_m E_m$.

$$H_{11} = -\delta_x^2\delta_y^2 \sum \frac{Q_m^2 E_m^2}{\sigma_m^2}$$
$$H_{12} = \frac{1}{w_x^2}\sum \Gamma_m\xi_m$$
$$H_{13} = \frac{1}{w_y^2}\sum \Gamma_m\eta_m$$
$$H_{14} = \frac{1}{w_x^3}\sum \Gamma_m\xi_m^2$$
$$H_{15} = \frac{1}{w_y^3}\sum \Gamma_m\eta_m^2$$
$$H_{22} = \frac{A}{w_x^4}\sum \Gamma_m\xi_m^2 - \gamma_m w_x^2$$
$$H_{23} = \frac{A}{w_x^2 w_y^2}\sum \Gamma_m\xi_m\eta_m$$
$$H_{24} = \frac{A}{w_x^5}\sum \Gamma_m\xi_m^3 - 2\gamma_m\xi_m w_x^2$$
$$H_{25} = \frac{A}{w_x^2 w_y^3}\sum \Gamma_m\xi_m\eta_m^2$$
$$H_{33} = \frac{A}{w_y^4}\sum \Gamma_m\eta_m^2 - \gamma_m w_y^2$$
$$H_{34} = \frac{A}{w_x^3 w_y^2}\sum \Gamma_m\xi_m^2\eta_m$$
$$H_{35} = \frac{A}{w_y^5}\sum \Gamma_m\eta_m^3 - 2\gamma_m\eta_m w_y^2$$
$$H_{44} = \frac{A}{w_x^6}\sum \Gamma_m\xi_m^4 - 3\gamma_m\xi_m^2 w_x^2$$
$$H_{45} = \frac{A}{w_x^3 w_y^3}\sum \Gamma_m\xi_m^2\eta_m^2$$
$$H_{55} = \frac{A}{w_y^6}\sum \Gamma_m\eta_m^4 - 3\gamma_m\eta_m^2 w_y^2,$$

where $\Gamma_m = \frac{\delta_x\delta_y}{\sigma_m^2}(g_m - 2\bar{g}_m)Q_m E_m$. Since $\mathcal{E}\{\Gamma_m\} = -\frac{\delta_x^2\delta_y^2 Q_m^2}{\sigma_m^2}AE_m^2$ and $\mathcal{E}\{\gamma_m\} = 0$, the Fisher information matrix $\mathbf{F}$ here is

$$
\begin{aligned}
F_{11} &= \sum R_m & & \\
F_{12} &= \tfrac{A}{w_x^2}\sum R_m\xi_m & F_{25} &= \tfrac{A^2}{w_x^2 w_y^3}\sum R_m\xi_m\eta_m^2 \\
F_{13} &= \tfrac{A}{w_y^2}\sum R_m\eta_m & F_{33} &= \tfrac{A^2}{w_y^4}\sum R_m\eta_m^2 \\
F_{14} &= \tfrac{A}{w_x^3}\sum R_m\xi_m^2 & F_{34} &= \tfrac{A^2}{w_x^3 w_y^2}\sum R_m\xi^2\eta_m \\
F_{15} &= \tfrac{A}{w_y^3}\sum R_m\eta_m^2 & F_{35} &= \tfrac{A^2}{w_y^5}\sum R_m\eta_m^3 \\
F_{22} &= \tfrac{A^2}{w_x^4}\sum R_m\xi_m^2 & F_{44} &= \tfrac{A^2}{w_x^6}\sum R_m\xi_m^4 \\
F_{23} &= \tfrac{A^2}{w_x^2 w_y^2}\sum R_m\xi_m\eta_m & F_{45} &= \tfrac{A^2}{w_x^3 w_y^3}\sum R_m\xi_m^2\eta_m^2 \\
F_{24} &= \tfrac{A^2}{w_x^5}\sum R_m\xi_m^3 & F_{55} &= \tfrac{A^2}{w_y^6}\sum R_m\eta_m^4
\end{aligned}
,
$$

with $R_m = \delta_x^2\delta_y^2 Q_m^2 E_m^2/\sigma_m^2$. This is the exact form of $\mathbf{F}$. Introducing the same approximations as used in Section 3, we obtain the following approximations of the above sums:

$$
\left.
\begin{aligned}
\sum E_m^2 &\approx \tfrac{\pi w_x w_y}{\delta_x\delta_y} \\
\sum E_m^2\xi_m^2 &\approx \tfrac{\pi w_x^3 w_y}{2\delta_x\delta_y} \\
\sum E_m^2\eta_m^2 &\approx \tfrac{\pi w_x w_y^3}{2\delta_x\delta_y} \\
\sum E_m^2\xi_m^2\eta_m^2 &\approx \tfrac{\pi w_x^3 w_y^3}{4\delta_x\delta_y} \\
\sum E_m^2\xi_m^4 &\approx \tfrac{3\pi w_x^5 w_y}{4\delta_x\delta_y} \\
\sum E_m^2\eta_m^4 &\approx \tfrac{3\pi w_x w_y^5}{4\delta_x\delta_y}
\end{aligned}
\right\}. \qquad (10)
$$

(All sums containing odd powers of ξ_m or η_m are zero.) From these we readily obtain the approximate form of the covariance matrix $\mathbf{K}$ for both flat and for Poisson noise:

$$
\mathbf{K}_{\text{Flat}} \approx \frac{\sigma^2}{\pi\delta_x\delta_y Q^2}
\begin{pmatrix}
\frac{2}{w_x w_y} & 0 & 0 & \frac{-1}{Aw_y} & \frac{-1}{Aw_x} \\
0 & \frac{2w_x}{A^2 w_y} & 0 & 0 & 0 \\
0 & 0 & \frac{2w_y}{A^2 w_x} & 0 & 0 \\
\frac{-1}{Aw_y} & 0 & 0 & \frac{2w_x}{A^2 w_y} & 0 \\
\frac{-1}{Aw_x} & 0 & 0 & 0 & \frac{2w_y}{A^2 w_x}
\end{pmatrix},
$$

$$
\mathbf{K}_{\text{Poisson}} \approx \frac{1}{2\pi}
\begin{pmatrix}
\frac{3A}{w_x w_y} & 0 & 0 & \frac{-1}{w_y} & \frac{-1}{w_x} \\
0 & \frac{w_x}{Aw_y} & 0 & 0 & 0 \\
0 & 0 & \frac{w_y}{Aw_x} & 0 & 0 \\
\frac{-1}{w_y} & 0 & 0 & \frac{2w_x}{3Aw_y} & \frac{1}{3A} \\
\frac{-1}{w_x} & 0 & 0 & \frac{1}{3A} & \frac{2w_y}{3Aw_x}
\end{pmatrix}.
$$

The parameter variances for $\hat{\boldsymbol{\theta}} = (\hat{A}, \hat{\bar{x}}, \hat{\bar{y}}, \hat{w}_x, \hat{w}_y)$ are found on the diagonal of the covariance matrix $\mathbf{K}$.

The estimator for the energy of the Gaussian is $\hat{U} = 2\pi\hat{A}\hat{w}_x\hat{w}_y$ and its variance is

$$
\operatorname{var}(\hat{U}) \approx \begin{cases} \frac{8\pi\sigma^2 w_x w_y}{\delta_x\delta_y Q^2} & \text{(Flat)} \\ 2\pi A w_x w_y & \text{(Poisson)} \end{cases}.
$$

Note that the separable (i.e., five-parameter) Gaussian model can also be used for a completely general 2D Gaussian profile (Section 4) when the azimuth angle of the Gaussian's elliptical cross section is known *a priori*. The data $\mathbf{g}$ used by the estimation algorithm is unmodified in this case, and the x and y abscissa inputs need only be rotated by the appropriate angle prior to estimation.

4. General 2D Gaussian Model

A general 2D Gaussian model function (Fig. 3) is

$$
\begin{aligned}
\bar{g}_m &= A\delta_x\delta_y Q_m \exp\left[-\frac{1}{2}(\mathbf{r}_m - \bar{\mathbf{r}})^{\mathsf{T}}\mathbf{C}^{-1}(\mathbf{r}_m - \bar{\mathbf{r}})\right] \\
&\equiv A\delta_x\delta_y E_m,
\end{aligned}
$$

where

$$
\mathbf{C}^{-1} = \begin{pmatrix} c_1 & c_3 \\ c_3 & c_2 \end{pmatrix},
$$

and $\boldsymbol{\theta} = (A, \bar{x}, \bar{y}, c_1, c_2, c_3)$. The matrix elements c_1 and c_2 are equivalent to the $1/w_y^2, 1/w_x^2$, from the separable 2D profile, and c_3 represents the "covariance term." The tilt angle α of the Gaussian profile relative to the coordinate axes is then given by $\alpha = \frac{1}{2}\arctan[-2c_3\sqrt{c_1 c_2}/(c_1 - c_2)]$.[11]

The score vector $\nabla\ell$, Hessian matrix $\mathbf{H}$, and Fisher information matrix $\mathbf{F}$ for this model are

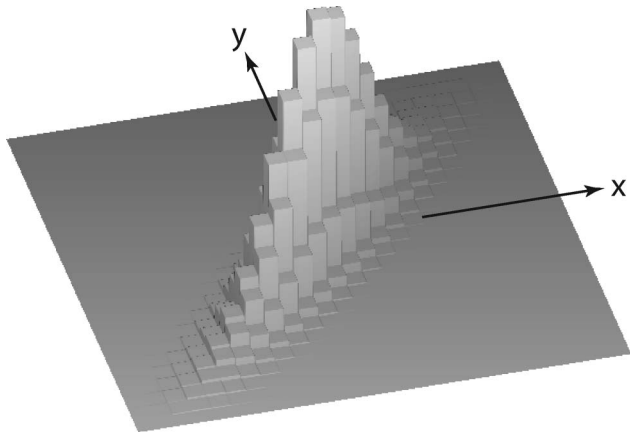

Fig. 3. Example noiseless general 2D Gaussian profile, sampled on a pixel grid. (Note that the axis of the Gaussian is not aligned to the grid.)

$$
\begin{aligned}
[\nabla \ell]_1 &= \frac{\partial \ell}{\partial A} = \Sigma \gamma_m \\
[\nabla \ell]_2 &= \frac{\partial \ell}{\partial \bar{x}} = A \Sigma \gamma_m (c_1 \xi_m + c_3 \eta_m) \\
[\nabla \ell]_3 &= \frac{\partial \ell}{\partial \bar{y}} = A \Sigma \gamma_m (c_2 \eta_m + c_3 \xi_m) \\
[\nabla \ell]_4 &= \frac{\partial \ell}{\partial c_1} = \frac{A}{2} \Sigma \gamma_m \xi_m^2 \\
[\nabla \ell]_5 &= \frac{\partial \ell}{\partial c_2} = -\frac{A}{2} \Sigma \gamma_m \eta_m^2 \\
[\nabla \ell]_6 &= \frac{\partial \ell}{\partial c_3} = -A \Sigma \gamma_m \xi_m \eta_m,
\end{aligned}
$$

for $\gamma_m = \frac{\delta_x \delta_y}{\sigma_m^2}(g_m - \bar{g}_m) Q_m E_m$, and

$$
\begin{aligned}
H_{11} &= -\delta_x^2 \delta_y^2 \sum \frac{Q_m^2 E_m^2}{\sigma_m^2} \\
H_{12} &= \sum \Gamma_m (c_1 \xi_m + c_3 \eta_m) \\
H_{13} &= \sum \Gamma_m (c_2 \eta_m + c_3 \xi_m) \qquad H_{14} = -\frac{1}{2} \sum \Gamma_m \xi_m^2 \\
H_{15} &= -\frac{1}{2} \sum \Gamma_m \eta_m^2 \qquad H_{16} = -\sum \Gamma_m \xi_m \eta_m \\
H_{22} &= A \sum \Gamma_m (c_1 \xi_m + c_3 \eta_m)^2 - c_1 \gamma_m \\
H_{23} &= A \sum \Gamma_m (c_1 \xi_m + c_3 \eta_m)(c_2 \eta_m + c_3 \xi_m) - c_3 \gamma_m \\
H_{24} &= -\frac{1}{2} A \sum \Gamma_m (c_1 \xi_m + c_3 \eta_m) \xi_m^2 - 2 \gamma_m \xi_m \\
H_{25} &= -\frac{1}{2} A \sum \Gamma_m (c_1 \xi_m + c_3 \eta_m) \eta_m^2 \\
H_{26} &= -A \sum \Gamma_m (c_1 \xi_m + c_3 \eta_m) \xi_m \eta_m - \gamma_m \eta_m \\
H_{33} &= A \sum \Gamma_m (c_2 \eta_m + c_3 \xi_m)^2 - c_2 \gamma_m \\
H_{34} &= -\frac{1}{2} A \sum \Gamma_m (c_2 \eta_m + c_3 \xi_m) \xi_m^2 \\
H_{35} &= -\frac{1}{2} A \sum \Gamma_m (c_2 \eta_m + c_3 \xi_m) \eta_m^2 - 2 \gamma_m \eta_m \\
H_{36} &= -A \sum \Gamma_m (c_2 \eta_m + c_3 \xi_m) \xi_m \eta_m - \gamma_m \xi_m \\
H_{44} &= \frac{A}{4} \sum \Gamma_m \xi_m^4 \qquad H_{45} = \frac{A}{4} \sum \Gamma_m \xi_m^2 \eta_m^2 \\
H_{46} &= \frac{A}{2} \sum \Gamma_m \xi_m^3 \eta_m \qquad H_{55} = \frac{A}{4} \sum \Gamma_m \eta_m^4 \\
H_{56} &= \frac{A}{2} \sum \Gamma_m \xi_m \eta_m^3 \qquad H_{66} = A \sum \Gamma_m \xi_m^2 \eta_m^2,
\end{aligned}
$$

where $\Gamma_m = \frac{\delta_x \delta_y}{\sigma_m^2}(g_m - 2\bar{g}_m) Q_m E_m$.

Once again, $\mathcal{E}\{\Gamma_m\} = -\frac{\delta_x^2 \delta_y^2}{\sigma_m^2} A Q_m^2 E_m^2$ and $\varepsilon\{\gamma_m\} = 0$, so that

$$
\begin{aligned}
F_{11} &= \sum R_m \\
F_{12} &= A \sum R_m (c_1 \xi_m + c_3 \eta_m) \\
F_{13} &= A \sum R_m (c_2 \eta_m + c_3 \xi_m) \\
F_{14} &= -\frac{A}{2} \sum R_m \xi_m^2 \qquad F_{15} = -\frac{A}{2} \sum R_m \eta_m^2 \\
F_{16} &= -A \sum R_m \xi_m \eta_m \\
F_{22} &= A^2 \sum R_m (c_1 \xi_m + c_3 \eta_m)^2 \\
F_{23} &= A^2 \sum R_m (c_1 \xi_m + c_3 \eta_m)(c_2 \eta_m + c_3 \xi_m) \\
F_{24} &= -\frac{A^2}{2} \sum R_m (c_1 \xi_m + c_3 \eta_m) \xi_m^2 \\
F_{25} &= -\frac{A^2}{2} \sum R_m (c_1 \xi_m + c_3 \eta_m) \eta_m^2 \\
F_{26} &= -A^2 \sum R_m (c_1 \xi_m + c_3 \eta_m) \xi_m \eta_m \\
F_{33} &= A^2 \sum R_m (c_2 \eta_m + c_3 \xi_m)^2 \\
F_{34} &= -\frac{A^2}{2} \sum R_m (c_2 \eta_m + c_3 \xi_m) \xi_m^2 \\
F_{35} &= -\frac{A^2}{2} \sum R_m (c_2 \eta_m + c_3 \xi_m) \eta_m^2 \\
F_{36} &= -A^2 \sum R_m (c_2 \eta_m + c_3 \xi_m) \xi_m \eta_m \\
F_{44} &= \frac{A^2}{4} \sum R_m \xi_m^4 \qquad F_{45} = \frac{A^2}{4} \sum R_m \xi_m^2 \eta_m^2 \\
F_{46} &= \frac{A^2}{2} \sum R_m \xi_m^3 \eta_m \qquad F_{55} = \frac{A^2}{4} \sum R_m \eta_m^4 \\
F_{56} &= \frac{A^2}{2} \sum R_m \xi_m \eta_m^3 \qquad F_{66} = A^2 \sum R_m \xi_m^2 \eta_m^2,
\end{aligned}
$$

where, again, $R_m = \delta_x^2 \delta_y^2 Q_m^2 E_m^2 / \sigma_m^2$.

We can again use integral expressions to approximate $\mathbf{F}$ to give analytic expressions for $\mathbf{K} = \mathbf{F}^{-1}$ under flat or Poisson noise:

$$
\mathbf{K}_{\mathrm{F}} \approx \frac{2\sigma^2}{\pi \delta_x \delta_y Q^2}
\begin{pmatrix}
\square & 0 & 0 & \frac{\square}{A} & \frac{\square}{A} & \frac{\square}{A} \\
0 & \frac{c_2}{A^2 \mu} & \frac{c_3}{A^2 \mu} & 0 & 0 & 0 \\
0 & \frac{c_3}{A^2 \mu} & \frac{c_1}{A^2 \mu} & 0 & 0 & 0 \\
\frac{\square}{A} & 0 & 0 & \frac{\square}{A^2} & \frac{\square}{A^2} & \frac{\square}{A^2} \\
\frac{\square}{A} & 0 & 0 & \frac{\square}{A^2} & \frac{\square}{A^2} & \frac{\square}{A^2} \\
\frac{\square}{A} & 0 & 0 & \frac{\square}{A^2} & \frac{\square}{A^2} & \frac{\square}{A^2}
\end{pmatrix},
$$

$$
\mathbf{K}_{\mathrm{P}} \approx \frac{\mu}{2\pi}
\begin{pmatrix}
2A & 0 & 0 & c_1 & c_2 & c_3 \\
0 & \frac{c_2}{A\mu^2} & \frac{c_3}{A\mu^2} & 0 & 0 & 0 \\
0 & \frac{c_3}{A\mu^2} & \frac{c_1}{A\mu^2} & 0 & 0 & 0 \\
c_1 & 0 & 0 & \frac{2c_1^2}{A} & \frac{2c_3^2}{A} & \frac{2c_1 c_3}{A} \\
c_2 & 0 & 0 & \frac{2c_3^2}{A} & \frac{2c_2^2}{A} & \frac{2c_2 c_3}{A} \\
c_3 & 0 & 0 & \frac{2c_1 c_3}{A} & \frac{2c_2 c_3}{A} & \frac{c_1 c_2 + c_3^2}{A}
\end{pmatrix},
$$

where $\mu = \sqrt{c_1 c_2 - c_3^2}$. Each instance of the symbol "$\square$" in the equation above represents a term that

has no compact analytic expression but that depends solely on the matrix elements c_1, c_2, and c_3. Calculating the variance of $\hat{U}$ (the estimate of volume under the Gaussian profile), using the same procedure as in Sections 2 and 3, gives

$$\mathrm{var}(\hat{U}) \approx \begin{cases} \square \frac{\sigma^2}{\delta_x \delta_y Q^2} & \text{(Flat)} \\ \square A & \text{(Poisson)} \end{cases}.$$

Thus, we again find $\mathrm{var}(\hat{U})$ to be independent of the input parameters $\bar{x}$ and $\bar{y}$, and for the case of flat noise, the variance is also independent of the profile height A.

5. Truncated Sampling

Occasionally, it is necessary to estimate a Gaussian profile that has been truncated by lying at the edge of the sampled region (see Fig. 4) and we can expect that the estimation accuracy will fall when the profile has been asymmetrically sampled in this way. The numerical evaluation of the exact Fisher information matrix terms allows us to obtain the Cramér-Rao bound for this specific situation, as well.

A simulation illustrating the behavior of the estimation parameters as the Gaussian profile sampling is increasingly truncated is shown in Fig. 5. The simulation uses a symmetric 2D profile whose center $(\bar{x},\bar{y})$ is at first well inside the sampling region (such that the "truncation parameter" $t = 0$), then approaches and goes beyond the vertical edge of the sampling region ($t = 0.6$, i.e., 60% of the volume under the profile is unsampled). For the example shown, the profile is sampled at a rate of 1/7 the Gaussian width (that is, an 8×8 pixel region contains ~68% of the volume under the profile).

The results show that the variance of the profile peak estimate is little affected until the truncation becomes quite large. And, along the truncation direction, the profile center estimate suffers much more than does the accuracy orthogonal to the truncation direction.

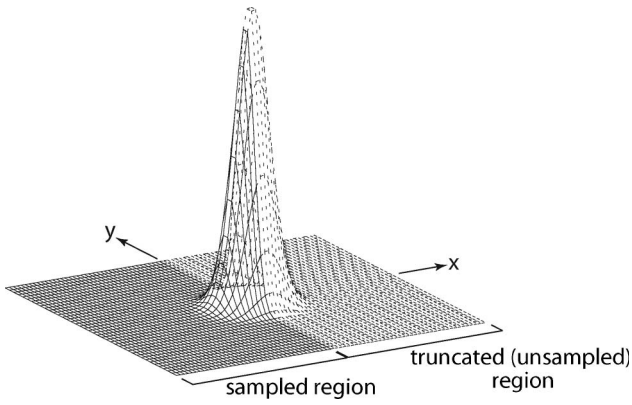

Fig. 4. Truncated sampling of a Gaussian profile, showing a case in which 60% of the volume under the profile lies outside of the region falling on the array detector.

6. Bias in the Standard Model

As mentioned in the previous paper [1], the model used here is biased, in the sense that use of an overly simple model for the measurement g_m produces a biased estimator of the true Gaussian:

$$\hat{f}_{\mathrm{unbiased}}(x,y) = \hat{f}_{\mathrm{biased}}(x,y) + b(x,y).$$

If we model the pixel response with rectangle-functions over which the signal is integrated, then

$$\bar{g}_m = Q_m \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) \times \mathrm{rect}\left(\frac{x-x_m}{\delta_x}\right)\mathrm{rect}\left(\frac{y-y_m}{\delta_y}\right)\mathrm{d}x\mathrm{d}y, \tag{11}$$

where (δ_x,δ_y) is the physical dimension of the pixel and the rectangle function is defined as

$$\mathrm{rect}(x/L) = \begin{cases} 1 : |x| < L/2 \\ 0 : |x| > L/2 \end{cases}.$$

The bias results from using a δ-sampling model, i.e., $\bar{g}_m = f(x_m,y_m)$, whereas the detection process is better approximated as an integral over the region of the pixel's response, i.e., Eq. (11). If we take a Taylor expansion of $f(x,y)$ about the pixel's center (x_m,y_m), we find that the two models coincide up to the linear term of the Taylor series. Thus, we can approximate the bias by integrating the second-order term across the pixel's response.

The Taylor expansion of a scalar-valued function of two variables has the form

$$\begin{aligned} f(x,y) \approx f(x_m,y_m) &+ (x-x_m)\left[\frac{\partial f}{\partial x}\right]_{\substack{x=x_m \\ y=y_m}} \\ &+ (y-y_m)\left[\frac{\partial f}{\partial y}\right]_{\substack{x=x_m \\ y=y_m}} \\ &+ \frac{1}{2}\left((x-x_m)^2\left[\frac{\partial^2 f}{\partial x^2}\right]_{\substack{x=x_m \\ y=y_m}}\right. \\ &+ 2(x-x_m)(y-y_m)\left[\frac{\partial^2 f}{\partial x \partial y}\right]_{\substack{x=x_m \\ y=y_m}} \\ &\left.+ (y-y_m)^2\left[\frac{\partial^2 f}{\partial y^2}\right]_{\substack{x=x_m \\ y=y_m}}\right). \end{aligned} \tag{12}$$

Therefore, the bias is given by

$$b(\mathbf{r}_m) \approx \int_{x_m-\frac{1}{2}\delta_x}^{x_m+\frac{1}{2}\delta_x}\int_{y_m-\frac{1}{2}\delta_y}^{y_m+\frac{1}{2}\delta_y}[\text{second-order term}]\mathrm{d}x\mathrm{d}y. \tag{13}$$

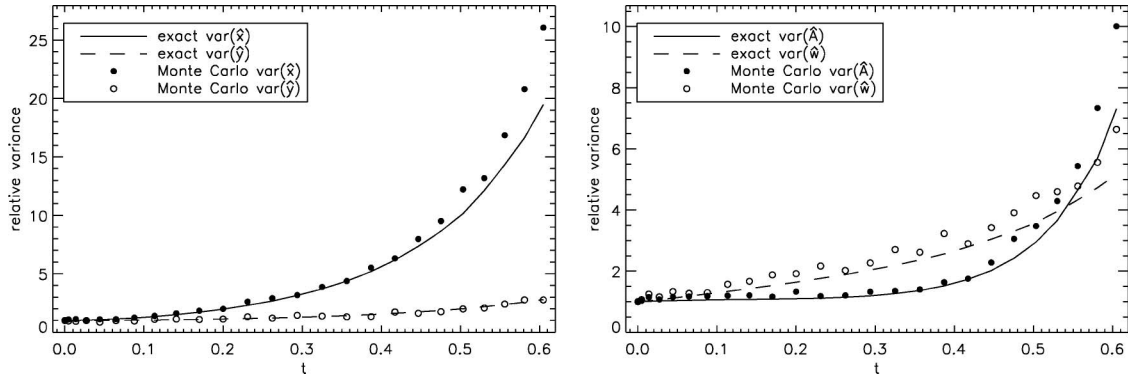

Fig. 5. The change in variance of estimation parameters as the sampling of the profile is increasingly truncated. The truncation parameter t is defined simply as the fraction of volume under the Gaussian profile which lies outside of the sampling region. In the figures, lines indicate the parameter variances as calculated numerically from the exact Fisher information matrix. (The analytic approximations are not valid for the case of a truncated profile.) Dots indicate variance estimates from a Monte Carlo simulation. The variances have been normalized to the completely sampled data's parameter variances. In both figures, the value of $\bar{x}$ is allowed to vary while all other model parameters are held constant.

The integral over the $(x-x_m)^2$ term gives $\frac{1}{12}\delta_x^2$; likewise, the integral over $(y-y_m)^2$ gives $\frac{1}{12}\delta_y^2$, while the cross term integrates to zero. Substituting these into Eq. (13) gives the equation of the bias as

$$b(\mathbf{r}_m) \approx \frac{AQ_m E(\mathbf{r}_m)}{24w^2}\left(\left[\frac{(x_m-\bar{x})^2}{w^2}-1\right]\delta_x^2 + \left[\frac{(y_m-\bar{y})^2}{w^2}-1\right]\delta_y^2\right).$$

Following the same procedure for the five-parameter Gaussian model (the separable 2D profile) gives an expression for the bias of

$$b(\mathbf{r}_m) \approx \frac{AQ_m E(\mathbf{r}_m)}{24}\left(\left[\frac{(x_m-\bar{x})^2}{w_x^2}-1\right]\frac{\delta_x^2}{w_x^2} + \left[\frac{(y_m-\bar{y})^2}{w_y^2}-1\right]\frac{\delta_y^2}{w_y^2}\right).$$

And, for the six-parameter model,

$$b(\mathbf{r}_m) \approx \frac{AQ_m E(\mathbf{r}_m)}{24}[([c_1(x-\bar{x})+c_3(y-\bar{y})]^2-c_1)\delta_x^2 + ([c_2(y-\bar{y})+c_3(x-\bar{x})]^2-c_2)\delta_y^2].$$

Knowing the value of the bias provides a means of further improving the estimation, by applying a correction to the input data and iterating the estimation until the bias becomes negligible. Note that the bias here is expressed in units of digital counts, not photoelectrons.

7. Optimal Pixel Size

As in the previous paper, for the 2D case, we can compare previous research with the results obtained here to evaluate whether there is an optimal magnification for the Gaussian profile or, equivalently, an optimal detector pixel size. Reference [12] reports an optimal choice of $\delta_x = \delta_y = 1.5\,w \sim 2.5\,w$ for minimizing the variance of the position estimator, $\hat{\bar{x}}$. In this situation, where we allow the detector pixel size to vary while the system magnification is kept constant, we need to be careful to enforce a constant energy U, via the relations

$$U = \begin{cases} 2\pi A w^2 & (4\,\text{params}) \\ 2\pi A w_x w_y & (5\,\text{params}) \\ 2\pi A/\mu & (6\,\text{params}) \end{cases},$$

where, again, $\mu = \sqrt{c_1c_2 - c_3^2}$. (The volume U is equal to the number of photoelectrons collected on the detector array.) The expressions for the position estimate variances become

$$\begin{aligned} (4\,\text{params}) : \text{var}(\hat{\bar{x}}) &\approx \begin{cases} \frac{8\pi\sigma^2 w^4}{\delta_x\delta_y Q^2U^2} & (\text{Flat}) \\ \frac{4w^2}{U} & (\text{Poisson}) \end{cases} \\ (5\,\text{params}) : \text{var}(\hat{\bar{x}}) &\approx \begin{cases} \frac{8\pi\sigma^2 w_x^3 w_y}{\delta_x\delta_y Q^2U^2} & (\text{Flat}) \\ \frac{4w_x^2}{U} & (\text{Poisson}) \end{cases} \\ (6\,\text{params}) : \text{var}(\hat{\bar{x}}) &\approx \begin{cases} \frac{2\pi\sigma^2 c_2}{\delta_x\delta_y\mu^3 Q^2U^2} & (\text{Flat}) \\ \frac{2}{c_1U} & (\text{Poisson}) \end{cases} \end{aligned} \quad (14)$$

(The variances for $\hat{\bar{y}}$ can easily be obtained from the above by symmetry arguments.)

These results indicate that, for Poisson noise, $\text{var}(\hat{\bar{x}})$ is independent of detector pixel size, whereas the flat-noise model shows that the variance decreases with larger pixel sizes. However, if the bias has not been corrected for within the algorithm, then we need to account for its effect on the measurement, as well. Figure 6 shows a comparison of the squared error in $\hat{\bar{x}}$ due to bias with $\text{var}(\hat{\bar{x}})$. (Note that the bias is a function of $\bar{x}$ relative to the sample locations x_m; if, for example, $\bar{x}$ coincides with a sample location, then the bias has no effect on the error in estimating $\bar{x}$.)

For $w \lesssim 0.7\delta_x$, Fig. 6 indicates that the estimation error is dominated by bias effects. For $w \geq 0.7\delta_x$, the

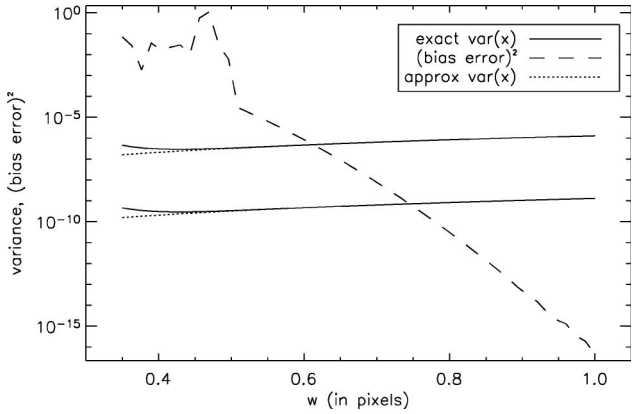

Fig. 6. A comparison of $\mathrm{var}(\hat{\bar{x}})$ and $\mathrm{bias}^2(\hat{\bar{x}})$, given in units of $(\mathrm{pixels})^2$, for an example symmetric Gaussian profile as a function of the profile width w. Keeping a constant value for U, we increase the profile width (and thus reduce the profile height A) while maintaining the same sampling rate. The approximate variance is calculated by Eq. (9) and the exact variance by taking the inverse of Eq. (5). The two curves shown for the variance are for two different values of U.

error is predominantly due to estimator variance. This behavior would suggest that an optimal pixel size should be $\delta_x \leq 1.43w$. Note that although the exact location of the optimal tradeoff will vary with each parameter set, the bias to the position estimator $\hat{\bar{x}}$ is relatively insensitive to the parameters A and $\bar{y}$.

8. Discussion

Real-life detector array data has noise properties that are neither purely Gaussian nor Poisson but are well approximated by a combined function in which the electronic noise is given by a Gaussian distribution and the photoelectron noise by a Poisson distribution (see [13], Eq. 3.16). Under this combined distribution, the Fisher information matrix is approximately (Eq. 3.17 of [13])

$$F_{ij} \approx \sum_{m=1}^{M} \frac{1}{\sigma_G^2 + \bar{g}_m Q_m} \frac{\partial \bar{g}_m}{\partial \theta_i} \frac{\partial \bar{g}_m}{\partial \theta_j},$$

where σ_G^2 describes the variance of Gaussian-distributed readout noise (with all pixels sharing the same variance) and $\bar{g}_m Q_m$ gives the variance of the Poisson-distributed shot noise. To make use of this Gaussian–Poisson mixed model in our system, we need only insert $(\sigma_G^2 + \bar{g}_m Q_m)$ for σ_m^2 in the exact expressions for $\mathbf{F}$ (i.e., Eq. (5) and the corresponding equations for the five- and six-parameter models). Unfortunately, there does not appear to be a way of deriving the corresponding approximate analytic expressions for this noise model.

One difficulty that can thwart accurate estimation is the failure of the optimization routine to locate a global maximum to the likelihood function. For well-sampled Gaussian profile models, the likelihood function is typically smooth and well behaved (an example is shown in Fig. 7(a)]. For difficult models, in which the width parameter w is less than a pixel width, multiple local maxima develop in the likelihood function [see Fig. 7(b)]. For such models, local optimization techniques, such as Newton iteration, become unreliable and global optimization methods, though slow, become important.

Finally, Gaussian profile estimation is often used in systems where other optical effects are present, such as a slowly varying background signal (due, for example, to stray light, scatter, or to background signals). Ignoring these effects would skew the estimation and, yet, they are not of interest to the observer and, thus, are termed nuisance parameters. Ideally, the estimation algorithm should incorporate the presence of these parameters by marginalizing the likelihood over the conditional probability density of the set of nuisance parameters. If this conditional density is unavailable, then an alternative approach is to estimate the nuisance parameters and subtract their effect from the data prior to use

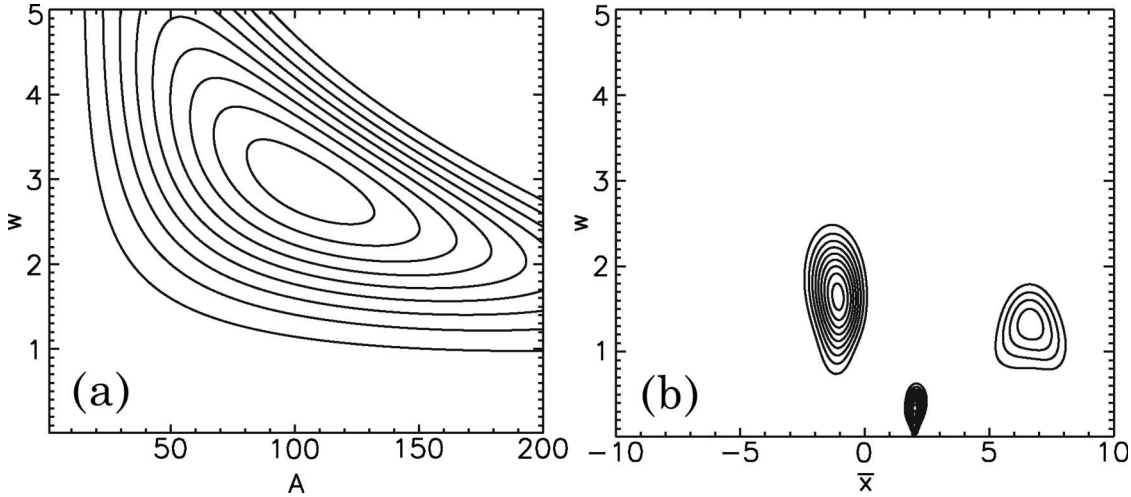

Fig. 7. Contour maps showing slices through the 4D likelihood function for (a) a well-behaved model and (b) a poorly behaved one. The parameters are $(A, \bar{x}, \bar{y}, w) = (100, 0.25, 0, 3)$ and (10,0.25,0,0.5) respectively, where $\bar{x}$, $\bar{y}$, and w are given in units of pixel widths.

in the Gaussian profile estimation algorithm. This second approach would allow use of the algorithm presented in this paper.

The authors have posted a set of programs, written in the IDL [14] language, that perform the estimation algorithm discussed in this paper. These are located at the url http://www.ittvis.com/codebank/index.asp. (The file is gauss_mle2.pro under the heading "Statistics".) Interested readers are encouraged to download, distribute, and use the code freely.

This work was supported in part by Department of Defense (DOD) contract DAAE07-02-C-L011. We would like to thank an anonymous reviewer for a very careful reading and valuable comments.

References and Notes

1. N. Hagen, M. Kupinski, and E. L. Dereniak, "Gaussian profile estimation in one dimension," Appl. Opt. **46**, 5374–5383 (2007).
2. L. G. Kazovsky, "Beam position estimation by means of detector arrays," Opt. Quantum Electron. **13**, 201–208 (1981).
3. L. H. Auer and W. F. van Altena, "Digital image centering II," Astron. J. **83**, 531–537 (1978).
4. R. Irwan and R. G. Lane, "Analysis of optimal centroid estimation applied to Shack–Hartmann sensing," Appl. Opt. **38**, 6737–6743 (1999).
5. M. K. Cheezum, W. F. Walker, and W. H. Guilford, "Quantitative comparison of algorithms for tracking single fluorescent particles," Biophys. J. **81**, 2378–2388 (2001).
6. R. E. Thompson, D. R. Larson, and W. W. Webb, "Precise nanometer localization analysis for individual fluorescent probes," Biophys. J. **82**, 2775–2783 (2002).
7. Y.-C. Chen, L. R. Furenlid, D. W. Wilson, and H. H. Barrett, "Calibration of scintillation cameras and pinhole SPECT imaging systems," in *Small-Animal SPECT Imaging*, M. A. Kupinski and H. H. Barrett, eds. (Springer, 2005), Chap. 12, pp. 195–202.
8. S. V. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory* (Prentice Hall, 1993), p. 34.
9. What we call "flat noise" here was labeled "uniform noise" in [1]. This has been changed in recognition that this can easily be mistaken as referring to noise obtained from the uniform probability distribution.
10. It may seem that $\sigma_m^2 = \bar{g}_m/Q_m$ is the appropriate substitution here, but since the raw measurements g_m are scaled versions of the Poisson-distributed photoelectrons, the variance increases as the square of the scale parameter, Q_m.
11. S. Brandt, *Data Analysis*, 3rd ed., (Springer 1999), p. 113–114.
12. K. A. Winick, "Cramér–Rao lower bounds on the performance of charge-coupled-device optical position estimators," J. Opt. Soc. Am. A **3**, 1809–1815 (1986).
13. H. H. Barrett, C. Dainty, and D. Lara, "Maximum-likelihood methods in wavefront sensing: stochastic models and likelihood functions," J. Opt. Soc. Am. A **24**, 391–414 (2007).
14. The Interactive Data Language software is developed by ITT Visual Information Systems, http://www.ittvis.com/index.asp.