Some theory for single molecule localization microscopy

Clayton W. Seitz

June 20, 2023

Photoswitching as a Markov jump process

$$G_{ij} = \Pr\left(X(t + dt) = \omega_i, \mid X(t) = \omega_j\right)$$

Let the state space for the process $X(t)$ be $\Omega = \{0_0, 0_1, 0_2, 1, 2\}$. The generator matrix for such a process reads

$$G = \begin{pmatrix} \lambda_{00} & \lambda_{00_1} & 0 & \lambda_{01} & \mu_0 \\ 0 & \lambda_{0_10_1} & \lambda_{0_10_2} & \lambda_{0_11} & \mu_1 \\ 0 & 0 & \lambda_{0_20_2} & \lambda_{0_21} & \mu_2 \\ \lambda_{10} & 0 & 0 & \lambda_{11} & \mu_0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Photoswitching as a Markov jump process

$$\frac{\partial P(\omega_i)}{\partial t} = \sum_j G_{ji}P(\omega_j, t) - G_{ij}P(\omega_i, t)$$

$$P(\boldsymbol{\omega}, t) = \exp(WP(\omega))$$

The matrix W for the 4-state system presented before reads

$$W = \begin{pmatrix} -\sigma_0 & \lambda_{00_1} & 0 & \lambda_{01} & \mu_0 \\ 0 & -\sigma_{0_1} & \lambda_{0_10_2} & \lambda_{0_11} & \mu_1 \\ 0 & 0 & -\sigma_{0_2} & \lambda_{0_21} & \mu_2 \\ \lambda_{10} & 0 & 0 & -\sigma_1 & \mu_0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Integrated isotropic gaussian point spread function

Let $G(x, y)$ be a normalized isotropic Gaussian density over the pixel array

$$\mathrm{G}(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\sigma^2}}$$

We can integrate this continuous density over pixels. Let (x_k, y_k) be the center of k

$$\lambda_k = \int_{x_k-\frac{1}{2}}^{x_k+\frac{1}{2}} G(x)dx \int_{y_k-\frac{1}{2}}^{y_k+\frac{1}{2}} G(y)dy$$

which gives the probability a photon arrives at pixel k per unit time

Recall that $\text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^t e^{-t^2} dt$

$$\lambda_k(x) = \frac{1}{2} \left(\text{erf}\left(\frac{x_k + \frac{1}{2} - x_0}{\sqrt{2}\sigma} \right) - \text{erf}\left(\frac{x_k - \frac{1}{2} - x_0}{\sqrt{2}\sigma} \right) \right)$$

$$\lambda_k(y) = \frac{1}{2} \left(\text{erf}\left(\frac{y_k + \frac{1}{2} - y_0}{\sqrt{2}\sigma} \right) - \text{erf}\left(\frac{y_k - \frac{1}{2} - y_0}{\sqrt{2}\sigma} \right) \right)$$

The shot-noise limited signal is then

$$\vec{S} = [\text{Poisson}(\lambda_1), \text{Poisson}(\lambda_2), ..., \text{Poisson}(\lambda_K)]$$

Integrated anisotropic gaussian point spread function (astigmatism)

Let $G(x, y)$ be a normalized anisotropic Gaussian density over the pixel array

$$G(x, y) = \frac{1}{2\pi\sigma_x(z)\sigma_y(z)} e^{-\frac{(x-x_0)^2}{2\sigma_x(z_0)^2} + \frac{(y-y_0)^2}{2\sigma_y(z_0)^2}} \tag{1}$$

A fairly simple model for $\sigma_x(z_0)$ and $\sigma_y(z_0)$ is

$$\sigma_x(z_0) = \sigma_0 + \alpha(z_0 + z_{min})^2$$
$$\sigma_y(z_0) = \sigma_0 + \beta(z_0 - z_{min})^2$$

In this case, we only need to make a small adjustment to the isotropic λ_k

$$\lambda_k(x) = \frac{1}{2}\left(\mathrm{erf}\left(\frac{x_k + \frac{1}{2} - x_0}{\sqrt{2}\sigma_x}\right) - \mathrm{erf}\left(\frac{x_k - \frac{1}{2} - x_0}{\sqrt{2}\sigma_x}\right)\right)$$

$$\lambda_k(y) = \frac{1}{2}\left(\mathrm{erf}\left(\frac{y_k + \frac{1}{2} - y_0}{\sqrt{2}\sigma_y}\right) - \mathrm{erf}\left(\frac{y_k - \frac{1}{2} - y_0}{\sqrt{2}\sigma_y}\right)\right)$$

Readout noise

Due to readout noise, we measure

$$\vec{H} = \vec{S} + \vec{\xi}$$

The distribution of H_k is the convolution:

$$
\begin{aligned}
P(H_k|\theta) &= P(S_k) \circledast P(\xi_k) \\
&= A \sum_{q=0}^{\infty} \frac{1}{q!} e^{-\mu_k} \mu_k^q \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{(H_k - g_k q - o_k)}{2\sigma_k^2}}
\end{aligned}
$$

where $P(\xi_k) = \mathcal{N}(o_k, \sigma_k^2)$ and $P(S_k) = \mathrm{Poisson}(g_k \mu_k)$. In practice, this expression is difficult to work with, so we look for an approximation. Notice that

$$\xi_k - o_k + \sigma_k^2 \sim \mathcal{N}(\sigma_k^2, \sigma_k^2) \approx \mathrm{Poisson}(\sigma_k^2)$$

The model log likelihood

Since $H_k = S_k + \xi_k$, we transform $H'_k = H_k - o_k + \sigma_k^2$, which is distributed according to

$$H'_k \sim \text{Poisson}(\mu'_k) \quad \mu'_k = g_k \mu_k + \sigma_k^2$$

Since each Poisson r.v. is independent, the negative log likelihood reads

$$\ell(\vec{H}) = -\log \prod_k \frac{e^{-(\mu'_k)} (\mu'_k)^{n_k}}{n_k!}$$

$$= \sum_k \log n_k! + \mu'_k - n_k \log \left(\mu'_k\right)$$

$$\approx \sum_k n_k \log n_k + \mu'_k - n_k \log \left(\mu'_k\right)$$

The Cramer-Rao bound

On the previous slide, we defined the log-likelihood $\ell(\vec{H}|\theta)$. From this we can compute the Fisher information matrix $I(\theta)$

$$\frac{\partial \ell}{\partial \theta_i} = \frac{\partial}{\partial \theta_i} \sum_k x_k \log x_k + \mu'_k - x_k \log\left(\mu'_k\right)$$

$$= \sum_k \frac{\partial \mu'_k}{\partial \theta_i} \left(\frac{\mu'_k - x_k}{\mu'_k}\right)$$

$$I_{ij}(\theta) = \underset{\theta}{\mathbb{E}}\left(\sum_k \frac{\partial \mu'_k}{\partial \theta_i} \frac{\partial \mu'_k}{\partial \theta_j} \left(\frac{\mu'_k - x_k}{\mu'_k}\right)^2\right) = \sum_k \frac{1}{\mu'_k} \frac{\partial \mu'_k}{\partial \theta_i} \frac{\partial \mu'_k}{\partial \theta_j}$$

The Cramer-Rao bound

To compute the bound, it turns out all we need is the Jacobian J_μ

$$J = \left(\frac{\partial \mu_k}{\partial x_0}, \frac{\partial \mu_k}{\partial y_0}, \frac{\partial \mu_k}{\partial z_0}, \frac{\partial \mu_k}{\partial \sigma_x}, \frac{\partial \mu_k}{\partial \sigma_y}, \frac{\partial \mu_k}{\partial \sigma_0} \right)$$

Let's first compute gradients for spatial coordinates. Define $\beta_k = g_k \gamma \Delta t N_0$ such that $\mu_k' = \beta_k \lambda_k$

$$J_{x_0} = \beta_k \lambda_y \frac{\partial \lambda_x}{\partial x_0} \quad J_{y_0} = \beta_k \lambda_x \frac{\partial \lambda_y}{\partial y_0} \quad J_{z_0} = \frac{\partial \mu_k'}{\partial \sigma_x} \frac{\partial \sigma_x}{\partial z_0} + \frac{\partial \mu_k'}{\partial \sigma_y} \frac{\partial \sigma_y}{\partial z_0}$$

The Cramer-Rao bound

$$\begin{aligned}
J_{x_0} &= \beta_k \lambda_y \frac{\partial \lambda_x}{\partial x_0} \\
&= \frac{\beta_k \lambda_y}{2} \frac{\partial}{\partial x_0} \left(\mathrm{erf}\left(\frac{x_k + \frac{1}{2} - x_0}{\sqrt{2}\sigma_x} \right) - \mathrm{erf}\left(\frac{x_k - \frac{1}{2} - x_0}{\sqrt{2}\sigma_x} \right) \right) \\
&= \frac{\beta_k \lambda_y}{\sqrt{2\pi}\sigma_x} \left(\exp\left(\frac{(x_k - \frac{1}{2} - x_0)^2}{2\sigma_x^2} \right) - \exp\left(\frac{(x_k + \frac{1}{2} - x_0)^2}{2\sigma_x^2} \right) \right)
\end{aligned}$$

The Cramer-Rao bound

$$
\begin{aligned}
J_{y_0} &= \beta_k \lambda_x \frac{\partial \lambda_y}{\partial y_0} \\
&= \frac{\beta_k \lambda_x}{2} \frac{\partial}{\partial y_0} \left(\mathrm{erf}\left(\frac{y_k + \frac{1}{2} - y_0}{\sqrt{2}\sigma_y} \right) - \mathrm{erf}\left(\frac{y_k - \frac{1}{2} - y_0}{\sqrt{2}\sigma_y} \right) \right) \\
&= \frac{\beta_k \lambda_x}{\sqrt{2\pi}\sigma_y} \left(\exp\left(\frac{(y_k - \frac{1}{2} - y_0)^2}{2\sigma_y^2} \right) - \exp\left(\frac{(y_k + \frac{1}{2} - y_0)^2}{2\sigma_y^2} \right) \right)
\end{aligned}
$$

The Cramer-Rao bound

$$
\begin{aligned}
J_{\sigma_x} &= \beta_k \lambda_y \frac{\partial \lambda_x}{\partial \sigma_x} \\
&= \frac{\beta_k \lambda_y}{2} \frac{\partial}{\partial \sigma_x} \left(\mathrm{erf}\left(\frac{x_k + \frac{1}{2} - x_0}{\sqrt{2}\sigma_x} \right) - \mathrm{erf}\left(\frac{x_k - \frac{1}{2} - x_0}{\sqrt{2}\sigma_x} \right) \right) \\
&= \frac{\beta_k \lambda_y}{\sqrt{2\pi}} \left(\frac{\left(x - x_0 - \frac{1}{2} \right) e^{-\frac{\left(x - x_0 - \frac{1}{2} \right)^2}{2\sigma_x^2}}}{\sigma_x^2} - \frac{\left(x - x_0 + \frac{1}{2} \right) e^{-\frac{\left(x - x_0 + \frac{1}{2} \right)^2}{2\sigma_x^2}}}{\sigma_x^2} \right)
\end{aligned}
$$

The Cramer-Rao bound

$$J_{\sigma_y} = \beta_k \lambda_x \frac{\partial \lambda_y}{\partial \sigma_y}$$

$$= \frac{\beta_k \lambda_x}{2} \frac{\partial}{\partial \sigma_y} \left(\mathrm{erf}\left(\frac{y_k + \frac{1}{2} - y_0}{\sqrt{2}\sigma_y} \right) - \mathrm{erf}\left(\frac{y_k - \frac{1}{2} - y_0}{\sqrt{2}\sigma_y} \right) \right)$$

$$= \frac{\beta_k \lambda_x}{\sqrt{2\pi}} \left(\frac{\left(y - y_0 - \frac{1}{2}\right) e^{-\frac{\left(y - y_0 - \frac{1}{2}\right)^2}{2\sigma_y^2}}}{\sigma_y^2} - \frac{\left(y - y_0 + \frac{1}{2}\right) e^{-\frac{\left(y - y_0 + \frac{1}{2}\right)^2}{2\sigma_y^2}}}{\sigma_y^2} \right)$$

The Cramer-Rao bound

$$\sigma_x(z_0) = \sigma_0 + \alpha(z_0 + z_{min})^2$$
$$\sigma_y(z_0) = \sigma_0 + \beta(z_0 - z_{min})^2$$

$$\frac{\partial \sigma_x}{\partial z_0} = 2\alpha(z_0 + z_{min}) \quad \frac{\partial \sigma_y}{\partial z_0} = 2\beta(z_0 - z_{min})$$

Bayesian localization with stochastic gradient langevin dynamics

$$dw = -\nabla L(w)dt + \epsilon\sqrt{\eta dt}, \quad \epsilon \sim \mathcal{N}(0, \sigma^2), \eta \propto dt$$

Our goal is to sample from the stationary distribution of the above SDE. To do that, we will use Stochastic Gradient Langevin Dynamics (SGLD), an algorithm commonly used to sample from the parameter's posterior – this is unlike MLE, whose goal is to simply minimize the objective L, which can be seen as finding the modes of the parameter's posterior