Single molecule localization microscopy

Clayton W. Seitz

May 31, 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

$$G = \begin{pmatrix} \lambda_{00} & \lambda_{00_1} & 0 & \lambda_{01} & \mu_0 \\ 0 & \lambda_{0_1 0_1} & \lambda_{0_1 0_2} & \lambda_{0_1 1} & \mu_1 \\ 0 & 0 & \lambda_{0_2 0_2} & \lambda_{0_2 1} & \mu_2 \\ \lambda_{10} & 0 & 0 & \lambda_{11} & \mu_0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

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

$$W = \begin{pmatrix} -\sigma_0 & \lambda_{00_1} & 0 & \lambda_{01} & \mu_0 \\ 0 & -\sigma_{0_1} & \lambda_{0_1 0_2} & \lambda_{0_1 1} & \mu_1 \\ 0 & 0 & -\sigma_{0_2} & \lambda_{0_2 1} & \mu_2 \\ \lambda_{10} & 0 & 0 & -\sigma_1 & \mu_0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Integrated isotropic gaussian point spread function

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

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

Let's presume (correctly) that pixel gray levels are a Poisson r.v. with expected value

$$\mu_k = \eta \Delta t (N_0 \lambda_k + B_0) \tag{2}$$

where Δt is the camera exposure time and N_0 and B_0 are the fluorophore and background emission rates respectively. Eq (1) and (2) are related by

$$\lambda_k = \int_{\mathrm{pixel}} G(x, y) dx dy$$

Integrated anisotropic gaussian point spread function (astigmatism)

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

$$\mathrm{G}(x, y) = \frac{1}{2\pi\sigma_x(z)\sigma_y(z)} e^{-\frac{(x-x_0)^2}{2\sigma_x(z)^2} + \frac{(y-y_0)^2}{2\sigma_y(z)^2}} \tag{3}$$

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

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

How to compute λ_k at each pixel

We can replace this integral with error functions:

$$\lambda_x(x) = \frac{1}{2}\left(\mathrm{erf}\left(\frac{x + a/2 - x_0}{\sqrt{2}\sigma}\right) - \mathrm{erf}\left(\frac{x - a/2 - x_0}{\sqrt{2}\sigma}\right)\right)$$
$$\lambda_y(y) = \frac{1}{2}\left(\mathrm{erf}\left(\frac{y + a/2 - y_0}{\sqrt{2}\sigma}\right) - \mathrm{erf}\left(\frac{y - y/2 - y_0}{\sqrt{2}\sigma}\right)\right)$$

For multiple emitters

$$\lambda(x, y) = \sum_n \lambda_{n,x}(x)\lambda_{n,y}(y)$$

The *true signal* is then

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

Poisson approximation of pixel values

However, 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 and Hessian matrix

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

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

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

$$\begin{aligned}
\ell(\vec{H}) &= -\log \prod_k \frac{e^{-\left(\mu'_k\right)} \left(\mu'_k\right)^{n_k}}{n_k!} \\
&= \sum_k \log n_k! + \mu'_k - n_k \log\left(\mu'_k\right) \\
&\approx \sum_k n_k \log n_k + \mu'_k - n_k \log\left(\mu'_k\right)
\end{aligned}$$

Localization with maximum likelihood estimation (MLE)

A slew of gradient calculations

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