1 Fisher Information for a Poisson variable

This is derivation of the fisher information for Poisson distributed variable X with mean λ.

$$X \sim \mathrm{Po}(n, \lambda) = p(n|\theta) = \frac{\lambda^n e^{-\lambda}}{n!}$$

1.1 Likelihood

Likelihood of the Poisson distributed variable with detection n_k in K pixels:

$$l(\theta) = \prod_{k=1}^{K} l_k = \prod_{k=1}^{K} \frac{\lambda_k^{n_k} e^{-\lambda_k}}{n_k!} \tag{1}$$

where $l_k(\theta) = p(n_k|\theta)$ to emphasise the dependency on the parameter θ.

Log-Likelihood:

$$\mathcal{L} = \sum_k (n_k \log \lambda_k - \lambda_k - \log n_k!)$$

1.2 Fisher Information

Fisher information:

$$I(\theta) = -\mathbb{E}\left[\frac{\partial^2 \mathcal{L}}{\partial \theta^2}\right] = \mathbb{E}\left[\left(\frac{\partial \mathcal{L}}{\partial \theta}\right)^2\right] = \mathbb{E}\left[\left(\sum_k \frac{\partial \log(l_k)}{\partial \theta}\right)^2\right] = \mathbb{E}\left[\left(\sum_k \frac{1}{l_k}\frac{\partial l_k}{\partial \theta}\right)^2\right] \tag{2}$$

$$\begin{aligned} I(\theta) &= \mathbb{E}\left[\left(\sum_k \frac{1}{l_k}\frac{\partial l_k}{\partial \theta}\right)\left(\sum_m \frac{1}{l_m}\frac{\partial l_m}{\partial \theta}\right)\right] \\ &= \mathbb{E}\left[\sum_k \frac{1}{l_k^2}\left(\frac{\partial l_k}{\partial \theta}\right)^2\right] + \mathbb{E}\left[\sum_k \sum_{m \neq k} \frac{1}{l_k}\frac{\partial l_k}{\partial \theta}\frac{1}{l_m}\frac{\partial l_m}{\partial \theta}\right] \end{aligned}$$

as n_k are iid then the second term can be expressed as

$$\mathbb{E}\left[\sum_k \sum_{m \neq k} \frac{1}{l_k}\frac{\partial l_k}{\partial \theta}\frac{1}{l_m}\frac{\partial l_m}{\partial \theta}\right] = \sum_k \sum_{m \neq k} \mathbb{E}_k\left[\frac{1}{l_k}\frac{\partial l_k}{\partial \theta}\right]\mathbb{E}_m\left[\frac{1}{l_m}\frac{\partial l_m}{\partial \theta}\right]$$

where

$$\mathbb{E}_k\left[f(n_k)\right] = \sum_{n_k \geq 0} p(n_k|\theta) f(n_k)$$

But

$$\mathbb{E}_k\left[\frac{1}{l_k}\frac{\partial l_k}{\partial \theta}\right] = \sum_{n_k} l_k \frac{1}{l_k}\frac{\partial l_k}{\partial \theta} = \sum_{n_k} \frac{\partial l_k}{\partial \theta} = \frac{\partial \sum_{n_k} l_k}{\partial \theta} = 0$$

as $\sum_{n_k} l_k = \sum_{n_k} p(n_k|\theta) = 1$. The Fisher Information can then be expressed

$$\begin{aligned} I(\theta) &= \mathbb{E}\left[\sum_k \frac{1}{l_k^2}\left(\frac{\partial l_k}{\partial \theta}\right)^2\right] \\ &= \sum_{k=1}^{k} \sum_{n_k \geq 0} l_k \frac{1}{l_k^2}\left(\frac{\partial l_k}{\partial \theta}\right)^2 \\ &= \sum_{k=1}^{k} \sum_{n_k \geq 0} \frac{1}{l_k}\left(\frac{\partial l_k}{\partial \theta}\right)^2 \end{aligned}$$

Derivatives of likelihood Eq.(1):

$$\frac{\partial l_k}{\partial \theta} = \frac{l_k(n_k - \lambda_k)}{\lambda_k} \frac{\partial \lambda_k}{\partial \theta}$$

And we get:

$$\begin{aligned} I(\theta) &= \sum_{k=1}^{k} \sum_{n_k \geq 0} \frac{l_k(n_k - \lambda_k)^2}{\lambda_k^2} \left(\frac{\partial \lambda_k}{\partial \theta} \right)^2 \\ &= \sum_{k=1}^{k} \frac{1}{\lambda_k^2} \left(\frac{\partial \lambda_k}{\partial \theta} \right)^2 \mathbb{E}_k \left[(n_k - \lambda_k)^2 \right] \end{aligned}$$

for Poisson $\mathrm{var}(n) = \mathrm{mean}(n) = \lambda$ gives

$$\mathbb{E}_k \left[(n_k - \lambda_k)^2 \right] = \mathrm{var}(n_k) = \lambda_k$$

and

$$I(\theta) = \sum_{k=1}^{K} \frac{1}{\lambda_k} \left(\frac{\partial \lambda_k}{\partial \theta} \right)^2 \tag{3}$$

This is the pixelised version (detection of the photons in K detectors - CCD camera and $\lambda_k = \int_{C_k} \lambda(x)dx$ where C_k is an area of the pixels of the detector).

Non pixelised version [Ram et al., 2006]

$$I(\theta) = \int \frac{1}{\lambda(x)} \left(\frac{\partial \lambda(x)}{\partial \theta} \right)^2 dx$$

1.3 Two sources separated by a distance d

These are comment on Fisher Information estimation as described in [Ram et al., 2006].

For two sources separated by a distance d we have a mean value of the intensity:

$$\lambda = \Lambda_1 f_1 + \Lambda_2 f_2$$

where f_i and Λ_i is the response function and intensity, respectively, of the source i. For translationally invariant PSF and in-focus sources: $f_1 = q(x - \frac{d}{2})$ and $f_2 = q(x + \frac{d}{2})$

$$\lambda(d) = \Lambda_1 q(x - \frac{d}{2}) + \Lambda_2 q(x + \frac{d}{2})$$

where q is the PSF of the sources. For pixelised version (integral over pixel area C_k)

$$\lambda_k(d) = \Lambda_1 \int_{C_k} q(x - \frac{d}{2})dx + \Lambda_2 \int_{C_k} q(x + \frac{d}{2})dx$$

so we get (as described in[Ram et al., 2006])

$$I(d) = \frac{1}{4}\sum_{k=1}^{K} \frac{\left(\Lambda_1 \int_{C_k} \partial_x q(x - \frac{d}{2})dx - \Lambda_2 \int_{C_k} \partial_x q(x + \frac{d}{2})dx\right)^2}{\Lambda_1 \int_{C_k} q(x - \frac{d}{2})dx + \Lambda_2 \int_{C_k} q(x + \frac{d}{2})dx} \tag{4}$$

Limit $d = 0$ If $\Lambda_1 = \Lambda_2$ then $I(d = 0) = 0$ which means $\mathrm{var}(d = 0) \to \infty$. (This does not hold for $\Lambda_1 \neq \Lambda_2$).

Limit $d \to \infty$ When sources are far apart then the mixing term in nominator in (4) $\Lambda_1\Lambda_2\partial_x q(x - \frac{d}{2})\partial q(d + \frac{d}{2}) = 0$ as the $\partial_x q(x - \frac{d}{2})$ ($q(x - \frac{d}{2})$) and $\partial_x q(x + \frac{d}{2})$ ($q(x + \frac{d}{2})$) do not have any overlap. The (4) then decomposes into two individual terms (sum of Fisher Information for localisation of individual sources.)

$$\begin{aligned} I(d) &= \frac{1}{4}\sum_{k=1}^{K} \left[\frac{\left(\Lambda_1 \int_{C_k} \partial_x q(x - \frac{d}{2})dx\right)^2}{\Lambda_1 \int_{C_k} q(x - \frac{d}{2})dx} + \frac{\left(\Lambda_2 \int_{C_k} \partial_x q(x + \frac{d}{2})dx\right)^2}{\Lambda_2 \int_{C_k} q(x + \frac{d}{2})dx} \right] \\ &= \frac{1}{4}\sum_{k=1}^{K} \frac{\left(\int_{C_k} \partial_x q(x)dx\right)^2}{\int_{C_k} q(x)dx} \left[\Lambda_1 + \Lambda_2\right] \end{aligned}$$

Limit $\Lambda_i = 0$ If $\Lambda_1 = 0$ or $\Lambda_2 = 0$ $I(d) \neq 0$. So the variance is finite even if one of the sources is not present.

1.4 An alternative way to derive Fisher information for two sources separated by d:

This is a suggestion how to fix the problems with limits for Fisher Information derived above. This gives infinite variance when one of the sources is no present. Also fix weird behaviour of the $I(d)$ for $d = 0$.

For two sources $f_1 = q(x - c_1)$ and $f_2 = q(x - c_2)$ we have $\lambda = \Lambda_1 f_1 + \Lambda_2 f_2$. The distance between the two sources is $d = c_1 - c_2$. This is a linear combination $\boldsymbol{a}^T \cdot \boldsymbol{c}$ of the variable $\boldsymbol{c} = (c_1, c_2)^T$ where $\boldsymbol{a} = (1, -1)^T$. The variance of d is given by

$$\mathrm{var}(d) = \mathrm{var}(\boldsymbol{a}^T \cdot \boldsymbol{c}) = \boldsymbol{a}^T \cdot \boldsymbol{Q} \cdot \boldsymbol{a} = Q_{11} + Q_{22} - 2Q_{12}$$

where $\boldsymbol{Q}$ is a covariance matrix $\boldsymbol{Q} = \boldsymbol{I}^{-1}(\theta)$ and $\boldsymbol{I}(\theta)$ is the Fisher information matrix (symmetric $I_{12} = I_{21}$)

$$\boldsymbol{I}(\theta) = \begin{pmatrix} I_{11} & I_{12} \\ I_{12} & I_{22} \end{pmatrix}$$

given by generalisation of Eq.(3)

$$I_{ij}(\theta) = \sum_{k=1}^{K} \frac{1}{\lambda_k} \frac{\partial \lambda_k}{\partial \theta_i} \frac{\partial \lambda_k}{\partial \theta_j} \tag{5}$$

The covariance matrix $\boldsymbol{Q}$ is then

$$\boldsymbol{Q} = \boldsymbol{I}^{-1}(\theta) = \frac{1}{I_{11}I_{12} - I_{12}^2} \begin{pmatrix} I_{22} & -I_{12} \\ -I_{12} & I_{11} \end{pmatrix}$$

and the variance of $d = c_1 - c_2$

$$\mathrm{var}(d) = (1, -1)^T \cdot \boldsymbol{Q} \cdot (1, -1) = \frac{I_{11} + I_{22} + 2I_{12}}{I_{11}I_{12} - I_{12}^2} = \frac{p}{q} \tag{6}$$

The individual terms of the Fisher Information matrix

$$I_{11} = \sum_{k=1}^{K} \frac{1}{\lambda_k} \left(\frac{\partial \lambda_k}{\partial c_1} \right)^2 = \sum_{k=1}^{K} \frac{(\Lambda_1 q_k'(c_1))^2}{\Lambda_1 q_k(c_1) + \Lambda_2 q_k(c_2)}$$

where $q_k(c_i)$ is the pixelised version (pixel area Γ_k) of the PSF

$$\begin{aligned} q_k(c_i) &= \int_{\Gamma_k} q(x - c_i) dx \\ q_k'(c_i) &= \int_{\Gamma_k} \frac{\partial q(x - c_i)}{\partial x} dx \end{aligned}$$

If we use

$$f_k(c_1, c_2) = \Lambda_1 q_k(c_1) + \Lambda_2 q_k(c_2),$$

we get expressions for the terms of the Fisher information matrix:

$$\begin{aligned} I_{ii} &= \Lambda_i^2 \sum_{k=1}^{K} \frac{(q_k'(c_i))^2}{f_k(c_1, c_2)} \\ I_{ij} &= \Lambda_i \Lambda_j \sum_{k=1}^{K} \frac{q_k'(c_i) q_k'(c_j)}{f_k(c_1, c_2)} \end{aligned} \tag{7}$$

Numerator $p = I_{11} + I_{22} + 2I_{12}$ in Eq.6

$$p = \sum_{k=1}^{K} \frac{1}{f_k(c_1, c_2)} \left[\Lambda_1^2 q'^2{}_k(c_1) + \Lambda_2^2 q'^2_k(c_2) + 2\Lambda_1 \Lambda_2 q_k'(c_1) q_k'(c_2) \right]$$

The terms in the denominator $r = I_{11}I_{22} - I_{12}^2$ in Eq.6

$$\begin{aligned} I_{11}I_{22} &= \Lambda_1^2 \Lambda_2^2 \sum_{k,l}^{K} \frac{(q_k'(c_1) q_l'(c_2))^2}{f_k(c_1, c_2) f_l(c_1, c_2)} \\ I_{12}^2 &= \Lambda_1^2 \Lambda_2^2 \sum_{k,l}^{K} \frac{q_k'(c_1) q_k'(c_2) q_l'(c_1) q_l'(c_2)}{f_k(c_1, c_2) f_l(c_1, c_2)} \end{aligned}$$

Limit $c_1 \to c_2$, $(d \to 0) \Rightarrow q_k(c_1) \to q_k(c_2)$

$$p = (\Lambda_1^2 + \Lambda_2^2 + 2\Lambda_1\Lambda_2) \sum_{k=1}^{K} \frac{q'^2_k(c)}{f_k(c, c)}$$

which can be further simplified by explicitly writing $f_k(c_1, c_2)$

$$p = (\Lambda_1 + \Lambda_2) \sum_{k=1}^{K} \frac{q'^2_k(c)}{q_k(c)}$$

q_k as a PSF is positive function, therefore the sum is not zero and p is non-zero for any Λ_1, Λ_2.

The two terms in the denominator in Eq.(6) are identical for $c_1 = c_2$

$$I_{11}I_{22} = I_{12}^2 = \frac{\Lambda_1^2\Lambda_2^2}{(\Lambda_1+\Lambda_2)^2}\sum_{k,l=1}^{K}\frac{q_k'(c)}{q_k(c)}\frac{q_l'(c)}{q_l(c)}$$

and therefore

$$r = I_{11}I_{22} - I_{12}^2 = \det\left[\boldsymbol{I}(\theta)\right] \quad \equiv \quad 0$$

for any Λ_i. $\boldsymbol{I}(\theta)$ is therefore a singular matrix for $d = 0$ and inversion $\boldsymbol{I}^{-1}(\theta)$ does not exist for $c_1 = c_2$, but the limit $c_1 \to c_2$, $(d \to 0)$ gives $p \neq 0$, $r \to 0$ and $\mathrm{var}(d \to 0) = \frac{p}{r} \to \infty$.

Limit $d \to \infty$ The cross term I_{ij} in Eq.7 vanishes ($I_{ij} = 0$, $i \neq j$) because of the multiplication $q_k'(c_1)q_k'(c_2)$ which is zero for largely separated PSF with finite support (if the support of $q_k(c_1)$ and $q_k(c_2)$ do not have mutual overlap). Then from Eq.6

$$\mathrm{var}(d) = \frac{1}{I_{11}} + \frac{1}{I_{22}}$$

which is the sum of variances for estimation two separated sources.

$$I_{ii} = \Lambda_i^2\sum_{k=1}^{K}\frac{q_k'^2(c_i)}{\Lambda_1 q_k(c_1) + \Lambda_2 q_k(c_2)} = \Lambda_i\sum_{k=1}^{K}\frac{q_k'^2(c_i)}{q_k(c_i)}$$

if $q_k(c_i)$ (and $q_k'(c_i)$) have a finite support.

For non-pixelised version and for Gaussian approximation of the PSF ($q(x-a) \propto \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right)$ (with $\sigma = \frac{\sqrt{2}}{2\pi}\frac{\lambda}{NA}$ [Zhang et al., 2007]) we have $q'(x) = \frac{1}{\sigma^2}xq(x)$ and $q'^2/q = \frac{1}{\sigma^4}x^2 q$ which gives $\int q'^2/q dx = \frac{1}{\sigma^4}\int qx^2 = \frac{1}{\sigma^4}\sigma^2 = \frac{1}{\sigma^2}$ and therefore $I_{ii} = \frac{\Lambda_i}{\sigma^2}$

$$\mathrm{var}(d \to \infty) = \sigma^2\left(\frac{1}{\Lambda_1} + \frac{1}{\Lambda_2}\right)$$

Limit $\Lambda_i = 0,\ \Lambda_j \neq 0$ then $I_{ii} \equiv 0$ and $I_{ij} \equiv 0$ and so $\det(\boldsymbol{I}(\theta)) \equiv 0$, and matrix is singular. In the limit $\Lambda_i \to 0$ the variance Eq.(6) $\mathrm{var}(d) \to \infty$.

1.5 Time distribution of the intensities (blinking)

For likelihood dependent on parameter Λ_t (T different time slices)

$$l_T(d,\Lambda) = \prod_{k=1}^{K}\prod_{t=1}^{T} p(n_k|d,\Lambda_t)p(\Lambda_t)$$

$$\mathcal{L}_T(d,\Lambda) = \sum_{k=1}^{K}\sum_{t=1}^{T}\left[\log\left(l_k(d,\Lambda_t)\right) + \log\left(p(\Lambda_t)\right)\right]$$

as $p(\Lambda)$ is not dependent on d then

$$\frac{\partial^2\mathcal{L}_T(d,\Lambda)}{\partial d^2} = \sum_{t=1}^{T}\frac{\partial^2\mathcal{L}(d,\Lambda_t)}{\partial d^2}$$

but in the expectation equation Eq.(2) the time dependence appears as

$$\begin{aligned}I_T(\theta) &= -\mathbb{E}_T\left[\sum_{t=1}^{T}\frac{\partial^2\mathcal{L}(d,\Lambda_t)}{\partial d^2}\right] = \sum_{t=1}^{T} -\mathbb{E}_T\left[\frac{\partial^2\mathcal{L}(d,\Lambda_t)}{\partial d^2}\right] = \sum_{t=1}^{T}\mathbb{E}_T\left[\left(\frac{\partial\mathcal{L}(d,\Lambda_t)}{\partial d}\right)^2\right] \\ &= \sum_{t=1}^{T}\int_{\Lambda_t} p(\Lambda_t)I(\theta)d\Lambda_t = \sum_{t,k}\int_{\Lambda_t} p(\Lambda_t)\frac{1}{\lambda_k(\Lambda_t)}\left(\frac{\partial\lambda_k(\Lambda_t)}{\partial d}\right)^2 d\Lambda_t\end{aligned}$$

1.6 Time distribution of the intensities - integrating out Λ

$$l_k(d) = \int_\Lambda l_k(d, \Lambda) d\Lambda = \int_\Lambda p(n_k|d, \Lambda) p(\Lambda) d\Lambda$$

for four state model of two sources: $\{(\Lambda_1, 0), (0, \Lambda_2), (\Lambda_1, \Lambda_2), (0, 0)\}$: $\lambda^1 = \Lambda_1 q(x - c_1)$, $\lambda^2 = \Lambda_2 q(x - c_2)$, $\lambda^3 = +\Lambda_1 q(x - c_1) + \Lambda_2 q(x - c_2)$, $\lambda^4 = 0$ with uniform distribution over these states

$$l_k(\theta) = \frac{1}{4} \sum_{i=1}^{4} \mathrm{Po}(\lambda_k^i)$$

derivatives

$$\frac{\partial l_k}{\partial c_p} = \frac{1}{4} \sum_i \frac{\partial \mathrm{Po}(\lambda_k^i)}{\partial c_p} = \frac{1}{4} \sum_i \left(\mathrm{Po}(\lambda_k^i) \frac{(n_k - \lambda_k^i)}{\lambda_k^i} \frac{\partial \lambda_k^i}{\partial c_p} \right)$$

The Fisher information matrix diagonal entries:

$$\begin{aligned}
I_{pp}(\theta) &= \mathbb{E}\left[\left(\sum_{k=1}^{N} \frac{1}{l_k} \frac{\partial l_k}{\partial c_p} \right)^2 \right] \\
&= \mathbb{E}\left[\left\{ \sum_{k=1}^{N} \left(\frac{1}{\sum_{j=1}^4 \mathrm{Po}(\lambda_k^j)} \frac{\partial \sum_{i=1}^4 \mathrm{Po}(\lambda_k^i)}{\partial c_p} \right) \right\} \left\{ \sum_{l=1}^{N} \left(\frac{1}{\sum_{j=1}^4 \mathrm{Po}(\lambda_l^j)} \frac{\partial \sum_{i=1}^4 \mathrm{Po}(\lambda_l^i)}{\partial c_p} \right) \right\} \right] \\
&= \sum_{k=1}^{N} \mathbb{E}_k \left[\frac{\left(\sum_{i=1}^4 \frac{\partial \mathrm{Po}(\lambda_k^i)}{\partial c_p} \right)^2}{\left(\sum_{j=1}^4 \mathrm{Po}(\lambda_k^j) \right)^2} \right]
\end{aligned} \tag{8}$$

as the cross terms $(k,\, l)$ in the sum (2nd row) are zeros:

$$\begin{aligned}
\mathbb{E}\left[\left(\frac{\sum_{i=1}^4 \frac{\partial \mathrm{Po}(\lambda_k^i)}{\partial c_p}}{\sum_{j=1}^4 \mathrm{Po}(\lambda_k^j)} \right) \left(\frac{\sum_{i=1}^4 \frac{\partial \mathrm{Po}(\lambda_l^i)}{\partial c_p}}{\sum_{j=1}^4 \mathrm{Po}(\lambda_l^j)} \right) \right] &= \mathbb{E}_k \left[\frac{\sum_{i=1}^4 \frac{\partial \mathrm{Po}(\lambda_k^i)}{\partial c_p}}{\sum_{j=1}^4 \mathrm{Po}(\lambda_k^j)} \right] \mathbb{E}_l \left[\frac{\sum_{i=1}^4 \frac{\partial \mathrm{Po}(\lambda_l^i)}{\partial c_p}}{\sum_{j=1}^4 \mathrm{Po}(\lambda_l^j)} \right] \\
&= \sum_{i=1}^{4} \frac{\partial}{\partial c_p} \left(\sum_{n_k \geq 0} \mathrm{Po}(\lambda_k^i) \right) \sum_{i=1}^{4} \frac{\partial}{\partial c_p} \left(\sum_{n_k \geq 0} \mathrm{Po}(\lambda_l^i) \right) \\
&= 0
\end{aligned}$$

Expressing the derivatives and the expectation from Eq.(8):

$$\begin{aligned}
I_{pp}(\theta) &= \sum_{k=1}^{N} \mathbb{E}_k \left[\left\{ \frac{\sum_{i=1}^4 \left(\mathrm{Po}(n_k; \lambda_k^i) \frac{(n_k - \lambda_k^i)}{\lambda_k^i} \frac{\partial \lambda_k^i}{\partial c_p} \right)}{\sum_{j=1}^4 \mathrm{Po}(n_k; \lambda_k^j)} \right\}^2 \right] \\
&= \frac{1}{4} \sum_{k=1}^{N} \sum_{n_k \geq 0} \frac{\left\{ \sum_{i=1}^4 \left(\mathrm{Po}(n_k; \lambda_k^i) \frac{(n_k - \lambda_k^i)}{\lambda_k^i} \frac{\partial \lambda_k^i}{\partial c_p} \right) \right\}^2}{\sum_{j=1}^4 \mathrm{Po}(n_k; \lambda_k^j)}
\end{aligned}$$

For the four states model we have $\lambda^3(c_1, c_2) = \lambda^1(c_1) + \lambda^2(c_2)$ and so $\frac{\partial \lambda^3}{\partial c_p} = \frac{\partial \lambda^p}{\partial c_p}$ and $\frac{\partial \lambda^j}{\partial c_p} = 0,\, i \neq j$ for $p = \{1, 2\},\ j = \{1, 2, 4\}$; so

$$I_{pp}(\theta) = \sum_{k=1}^{N} \left(\frac{\partial \lambda_k^p}{\partial c_p} \right)^2 \mathbb{E}_k \left[\left\{ \frac{\sum_{i=\{p,3\}} \left(\mathrm{Po}(n_k; \lambda_k^i) \frac{(n_k - \lambda_k^i)}{\lambda_k^i} \right)}{\sum_{j=1}^4 \mathrm{Po}(n_k; \lambda_k^j)} \right\}^2 \right]$$

The Fisher information matrix off-diagonal entries:

$$
\begin{aligned}
I_{pq}(\theta) &= \sum_{k=1}^{N} \mathbb{E}_k \left[\frac{\left(\sum_{i=1}^{4} \frac{\partial \mathrm{Po}(\lambda_k^i)}{\partial c_p}\right)\left(\sum_{l=1}^{4} \frac{\partial \mathrm{Po}(\lambda_k^l)}{\partial c_q}\right)}{\left(\sum_{j=1}^{4} \mathrm{Po}(\lambda_k^j)\right)^2} \right] \\
&= \sum_{k=1}^{N} \left(\frac{\partial \lambda_k^p}{\partial c_p}\right)\left(\frac{\partial \lambda_k^q}{\partial c_q}\right) \mathbb{E}_k \left[\frac{\left(\sum_{i=\{p,3\}} \mathrm{Po}(n_k;\lambda_k^i)\frac{(n_k-\lambda_k^i)}{\lambda_k^i}\right)\left(\sum_{i=\{q,3\}} \mathrm{Po}(n_k;\lambda_k^i)\frac{(n_k-\lambda_k^i)}{\lambda_k^i}\right)}{\left(\sum_{j=1}^{4} \mathrm{Po}(n_k;\lambda_k^j)\right)^2} \right]
\end{aligned}
\tag{9}
$$

Limit $d \to 0$ When $c^1 = c^2$ then $\lambda^1 = \lambda^2$ and $\frac{\partial \mathrm{Po}(\lambda^1)}{\partial c^1} = \frac{\partial \mathrm{Po}(\lambda^2)}{\partial c^2}$. Then all entries in I_{pq} are equal and the matrix is singular. For the limit $d \to 0$ the determianat $\det(\boldsymbol{I}) \to 0$ and the variance $\mathrm{var}(d) \to \infty$.

Limit $d \to \infty$ Sources are far apart and λ^1 and λ^2 don not have a common overlap. For k' where $\lambda_{k'}^1 > 0$, $\lambda_{k'}^2 \equiv 0$ and $\mathrm{Po}(n_{k'}, \lambda_{k'}^3) = \mathrm{Po}(n_{k'}, \lambda_{k'}^1 + \lambda_{k'}^2) = \mathrm{Po}(n_{k'}, \lambda_{k'}^1)$. Also $\frac{\partial \lambda^p}{\partial c_q} = 0$, $p \neq q$. From Eq.(8) the cross terms vanishes ($I_{pq} = 0$ because $\frac{\partial \mathrm{Po}(\lambda^p)}{\partial c_p}\frac{\partial \mathrm{Po}(\lambda^q)}{\partial c_q} = 0$) and the diagonal elemets

$$
\begin{aligned}
I_{pp} &= \sum_{k=1}^{N} \mathbb{E}_k \left[\frac{\left(2\frac{\partial \mathrm{Po}(\lambda_k^p)}{\partial c_p}\right)^2}{(2\mathrm{Po}(\lambda_k^p))^2} \right] \\
&= \sum_{k=1}^{N} \mathbb{E}_k \left[\frac{\left(\mathrm{Po}(\lambda_k^p)\frac{(n_k-\lambda_k^p)}{\lambda_k^p}\frac{\partial \lambda_k^p}{\partial c_p}\right)^2}{(\mathrm{Po}(\lambda_k^p))^2} \right] \\
&= \sum_{k=1}^{N} \left(\frac{1}{\lambda_k^p}\frac{\partial \lambda_k^p}{\partial c_p}\right)^2 \mathbb{E}_k \left[(n_k - \lambda_k^p)^2\right] \\
&= \sum_{k=1}^{N} \left(\frac{1}{\lambda_k^p}\frac{\partial \lambda_k^p}{\partial c_p}\right)^2 \frac{1}{4} \sum_{n_k \geq 0} \left(\sum_{i=1}^{4} \mathrm{Po}(\lambda_k^i)\,(n_k - \lambda_k^p)^2\right) \\
&= \sum_{k=1}^{N} \left(\frac{1}{\lambda_k^p}\frac{\partial \lambda_k^p}{\partial c_p}\right)^2 \frac{1}{4} \sum_{n_k \geq 0} \left(2\mathrm{Po}(\lambda_k^p)\,(n_k - \lambda_k^p)^2\right) \\
&= \sum_{k=1}^{N} \left(\frac{1}{\lambda_k^p}\frac{\partial \lambda_k^p}{\partial c_p}\right)^2 \frac{1}{2}\lambda_k^p \\
&= \frac{1}{2}\sum_{k=1}^{N} \frac{1}{\lambda_k^p}\left(\frac{\partial \lambda_k^p}{\partial c_p}\right)^2
\end{aligned}
$$

which is the Eq.(3) (up to the factor 1/2). The factor 1/2 comes from the fact that the source appears only in 50% of the observations.

References

[Ram et al., 2006] Ram, S., Ward, E. S., and Ober, R. J. (2006). Beyond Rayleigh's criterion: a resolution measure with application to single-molecule microscopy. *Proceedings of the National Academy of Sciences of the United States of America*, 103(12):4457–62.

[Zhang et al., 2007] Zhang, B., Zerubia, J., and Olivo-Marin, J. (2007). Gaussian approximations of fluorescence microscope point-spread function models. *Applied Optics*, 46(10):1819–1829.