$$p(\theta) = \frac{1}{2\pi}$$

$$\begin{aligned}
P(r_i|\theta) &= \frac{(f_i(\theta)\tau)^{r_i}}{r_i!} e^{-f_i(\theta)\tau} \\
P(r_1, \dots r_M|\theta) &= \prod_{j=1}^{M} \frac{(f_j(\theta)\tau)^{r_j}}{r_j!} e^{-f_j(\theta)\tau}
\end{aligned}$$

where $f(\theta) = \int t(\theta - y)s(y)dy$ is the rate function, t is the tuning curve (periodic), s is the stimulus.
Problem:

$$\begin{gathered}
\max_{t_1,\cdots,t_M} I(\mathbf{r}, \theta) \\
0 < FM \le t_i \le FP \\
\int t_i(\theta)p(\theta)d\theta = const
\end{gathered}$$

Discretize $\theta \in (0, 2\pi)$ to be $(\theta_1, \dots, \theta_M)$, $M = numBin$. Assume the tuning curve (also the rate curve) is rotationally invariant, i.e. different f_i's have the same shape but difference in translations (rotations in $(0, 2\pi)$).
Therefore,

$$f_i(\theta_j) = f_{i-j}$$

($(i - j)$ stands for $(i - j)\%M$ here.)

$$f_i = stimWidth * \sum_{j=0}^{M-1} t_{i-j} s_j$$

Also, set $\tau = 1$.

$$\begin{aligned}
I(\mathbf{r};\theta) &= D_{KL}(p(\mathbf{r},\theta)||p(\mathbf{r})p(\theta)) \\
&= \int_\theta \int_\mathbf{r} p(\mathbf{r},\theta) \ln\left(\frac{p(\mathbf{r},\theta)}{p(\mathbf{r})p(\theta)}\right) d\mathbf{r}d\theta \\
&= \int_\theta \int_\mathbf{r} p(\mathbf{r}|\theta)p(\theta) \ln\left(\frac{p(\mathbf{r}|\theta)}{p(\mathbf{r})}\right) d\mathbf{r}d\theta \\
&= \sum_{i=1}^{M} \int_\mathbf{r} p(\mathbf{r}|\theta = \theta_i)\frac{1}{M} \ln\left(\frac{p(\mathbf{r}|\theta = \theta_i)}{p(\mathbf{r})}\right) d\mathbf{r} \\
&= \int_\mathbf{r} p(\mathbf{r}|\theta = 0) \ln\left(\frac{p(\mathbf{r}|\theta = 0)}{p(\mathbf{r})}\right) d\mathbf{r} \text{ (by rotational invariance)} \\
&= -E_{\mathbf{r}|\theta=0}\left[\ln\left(\frac{p(\mathbf{r})}{p(\mathbf{r}|\theta = 0)}\right)\right]
\end{aligned}$$

Since

$$\begin{aligned}
p(\mathbf{r}) &= \sum_{i=1}^{M} p(\theta_i)p(\mathbf{r}|\theta = \theta_i) \\
&= \sum_{i=1}^{M} \frac{1}{M} \prod_{j=1}^{M} \frac{(f_j(\theta_i)\tau)^{r_j}}{r_j!} e^{-f_j(\theta_i)\tau} \\
&= \frac{1}{M} \sum_{i=1}^{M} \prod_{j=1}^{M} \frac{{f_{j-i}}^{r_j}}{r_j!} e^{-f_{j-i}}
\end{aligned}$$

$$
\begin{aligned}
\frac{p(\mathbf{r})}{p(\mathbf{r}|\theta=0)} &= \frac{\frac{1}{M}\sum_{i=1}^{M}\prod_{j=1}^{M}\frac{{f_{j-i}}^{r_j}}{r_j!}e^{-f_{j-i}}}{\prod_{j=1}^{M}\frac{{f_{j-0}}^{r_j}}{r_j!}e^{-f_{j-0}}} \\
&= \frac{1}{M}\sum_{i=1}^{M}\prod_{j=1}^{M}\left(\frac{{f_{j-i}}^{r_j}}{{f_{j-0}}^{r_j}}\right)\cdot\prod_{j=1}^{M}\left(\frac{r_j!}{r_j!}\right)\cdot\left(\frac{\prod_{j=1}^{M}e^{-f_{j-i}}}{\prod_{j=1}^{M}e^{-f_j}}\right) \\
&= \frac{1}{M}\sum_{i=1}^{M}\prod_{j=1}^{M}\left(\frac{{f_{j-i}}^{r_j}}{{f_j}^{r_j}}\right)\cdot 1\cdot 1
\end{aligned}
$$

Therefore

$$
\begin{aligned}
I(\mathbf{r};\theta) &= -E_{\mathbf{r}|\theta=0}\ln\left(\frac{p(r)}{p(r|\theta=0)}\right) \\
&= -E_{\mathbf{r}|\theta=0}\ln\left(\frac{1}{M}\sum_{i=1}^{M}\prod_{j=1}^{M}\left(\frac{{f_{j-i}}^{r_j}}{{f_j}^{r_j}}\right)\right) \\
&= -E_{\mathbf{r}|\theta=0}\ln\left(S(\mathbf{r})\right) \\
&= -\sum_{\mathbf{r}}P(\mathbf{r}|\theta=0)\ln\left(S(\mathbf{r})\right)
\end{aligned}
$$

Where

$$P(\mathbf{r}|\theta=0)=\prod_{j=1}^{M}\frac{{f_j}^{r_j}}{r_j!}e^{-f_j}$$

$$S(\mathbf{r})=\frac{1}{M}\sum_{k=1}^{M}Q_s^k(\mathbf{r})$$

$$Q_s^k(\mathbf{r})=\prod_{j=1}^{M}\left(\frac{{f_{j-k}}^{r_j}}{{f_j}^{r_j}}\right)=\prod_{j=1}^{M}f_j^{r_{j+k}-r_j}$$

In Lorenzo's notes the 1st and 2nd order partial derivatives of $-I(\mathbf{r};\theta)$ (w.r.t. f_i) are computed. Also, the partial derivative of $P(\mathbf{r}|\theta=0)$:

$$
\begin{aligned}
\partial_i P &= \left(\frac{r_i}{f_i}-1\right)P \\
\partial_{ij}^2 P &= \left(\frac{r_i}{f_i}-1\right)\left(\frac{r_j}{f_j}-1\right)P,\text{ for } j\neq i \\
\partial_i^2 P &= \left(\frac{r_i}{f_i}-1\right)^2 P+\left(-\frac{r_i}{f_i^2}\right)P
\end{aligned}
$$

Therefore the 1st and 2nd order derivatives of $I(\mathbf{r};\theta)$:

$$
\begin{aligned}
\frac{\partial I}{\partial f_i} &= -\sum_{\mathbf{r}}(\partial_i P)\ln(S) \\
&= -\sum_{\mathbf{r}}P\cdot(r_i/f_i-1)\ln(S) \\
&= E_{\mathbf{r}|\theta=0}\left[(1-r_i/f_i)\ln(S)\right]
\end{aligned}
$$

$$
\frac{\partial^2 I}{\partial f_i\partial f_j} = -\left(\sum_{\mathbf{r}}(\partial_{ij}^2 P)\ln(S)+\sum_{\mathbf{r}}(\partial_i P)\frac{\partial_j S}{S}\right)
$$

$$
\begin{aligned}
&= -\sum_{\mathbf{r}} \left((r_i/f_i - 1)(r_j/f_j - 1) + 1_{\{i=j\}} \left(-r_i/f_i^2\right) \right) P \ln(S) \\
&\quad -\frac{1}{2f_if_j} \sum_{\mathbf{r}} \frac{-\sum_k (r_i - r_{i+k})(r_j - r_{j+k}) \prod_l f_l^{r_{l+k}+r_l}}{\sum_k \prod_l f_l^{r_{k+l}} e^{f_l} r_l!} \\
&= E_{\mathbf{r}|\theta=0} \left[-(r_i/f_i - 1)(r_j/f_j - 1)\ln(S) + \frac{1}{2f_if_j} \frac{\sum_k (r_i - r_{i+k})(r_j - r_{j+k}) \prod_l f_l^{r_{l+k}}}{\sum_k \prod_l f_l^{r_{k+l}}} + 1_{\{i=j\}} \cdot \left(r_i/f_i^2\right)\ln(S) \right]
\end{aligned}
$$

To avoid overflow, compute S in the following way:

$$
\begin{aligned}
S(\mathbf{r}) &= \frac{1}{M} \sum_{i=1}^{M} \prod_{j=1}^{M} \left(\frac{{f_{j-i}}^{r_j}}{{f_j}^{r_j}} \right) \\
&= \frac{1}{M} \sum_{i=1}^{M} \exp\left[\sum_{k=1}^{M} r_k \ln\left(\frac{f_{k-i}}{f_k} \right) \right] \\
&= \frac{1}{M} \sum_{i=1}^{M} \exp\left[\sum_{k=1}^{M} r_k \ln\left(\frac{f_{k-i}}{f_k} \right) - \max_i \left(\sum_{k=1}^{M} r_k \ln\left(\frac{f_{k-i}}{f_k} \right) \right) \right] \cdot \exp \max_i \left(\sum_{k=1}^{M} r_k \ln\left(\frac{f_{k-i}}{f_k} \right) \right)
\end{aligned}
$$

For convenience of computation, define

$$
\begin{aligned}
Lrate(k,j) &:= \ln\left(\frac{f_{k-i}}{f_k} \right) \\
Mexp(i) &:= \sum_k r_k Lrate(k,j) \\
Max &:= \max_i Mexp(i) \\
S(\mathbf{r}) &:= \sum_{i=1}^{M} \exp\left[Mexp(i) - Max \right] \cdot e^{Max} \\
Eexp(i) &:= \frac{\exp\left[Mexp(i) - Max \right]}{\frac{1}{M}\sum_{i=1}^{M} \exp\left[Mexp(i) - Max \right]} \\
&= \frac{\prod_{j=1}^{M} \left(\frac{{f_{j-i}}^{r_j}}{{f_j}^{r_j}} \right)}{\frac{1}{M}\sum_{i=1}^{M} \prod_{j=1}^{M} \left(\frac{{f_{j-i}}^{r_j}}{{f_j}^{r_j}} \right)} \\
&= \frac{\prod_{j=1}^{M} {f_j}^{r_{j+i}}}{\frac{1}{M}\sum_{i=1}^{M} \prod_{j=1}^{M} {f_j}^{r_{j+i}}} \in [0,1]
\end{aligned}
$$

Then $\ln(S(\mathbf{r})) = \ln\left(\frac{1}{M}\sum_{i=1}^{M} \exp\left[Mexp(i) - Max \right] \right) + Max$,

$$
\begin{aligned}
I(\mathbf{r};\theta) &= -E_{\mathbf{r}|\theta=0} \ln\left(S(\mathbf{r}) \right) \\
\frac{\partial I}{\partial f_i} &= E_{\mathbf{r}|\theta=0} \left((1 - r_i/f_i) \ln(S(\mathbf{r})) \right)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 I}{\partial f_i \partial f_j} &= -E_{\mathbf{r}|\theta=0} \left[(r_i/f_i - 1)(r_j/f_j - 1)\ln(S) \right] \\
&\quad + \frac{1}{2f_if_j} E_{\mathbf{r}|\theta=0} \left[\frac{\sum_k (r_i - r_{i+k})(r_j - r_{j+k}) \prod_l f_l^{r_{l+k}}}{\sum_k \prod_l f_l^{r_{k+l}}} \right] + 1_{\{i=j\}} \cdot E_{\mathbf{r}|\theta=0} \left[\left(r_i/f_i^2\right)\ln(S) \right] \\
&= E_{\mathbf{r}|\theta=0} \left[-(r_i/f_i - 1)(r_j/f_j - 1)\ln(S) + \frac{1}{2f_if_j} \sum_k (r_i - r_{i+k})(r_j - r_{j+k}) Eexp(k) + 1_{\{i=j\}} \cdot \left(r_i/f_i^2\right)\ln(S) \right]
\end{aligned}
$$

Therefore our objective function

$$F = -I(\mathbf{r}, \theta) = E_{\mathbf{r}|\theta=0} \ln(S(\mathbf{r}))$$

$$\partial_i F = -\partial_i I(\mathbf{r}, \theta) = -E_{\mathbf{r}|\theta=0} [(1 - r_i/f_i) \ln(S)]$$

$$\begin{aligned}
\partial_i \partial_j F &= -\partial_i \partial_j I(\mathbf{r}, \theta) \\
&= -E_{\mathbf{r}|\theta=0} \left[-(r_i/f_i - 1)(r_j/f_j - 1) \ln(S) + \frac{1}{2 f_i f_j} \sum_k (r_i - r_{i+k})(r_j - r_{j+k}) Eexp(k) + 1_{\{i=j\}} \cdot \left(r_i/f_i^2\right) \ln(S) \right]
\end{aligned}$$

Approximate by Monte Carlo:

$$\hat{F}_N = \frac{1}{N} \sum_{t=1}^{N} \ln\left(S(\mathbf{r}^{(t)})\right)$$

$$\partial_i \hat{F}_N = \frac{1}{N} \sum_{t=1}^{N} - \left[\left(1 - r_i^{(t)}/f_i\right) \ln(S(\mathbf{r}^{(t)})) \right]$$

$$\partial_i \partial_j \hat{F}_N = \frac{1}{N} \sum_{t=1}^{N} - \left[-(r_i/f_i - 1)(r_j/f_j - 1) \ln(S(\mathbf{r})) + \frac{1}{2 f_i f_j} \sum_k (r_i - r_{i+k})(r_j - r_{j+k}) Eexp(k) + 1_{\{i=j\}} \cdot \left(r_i/f_i^2\right) \ln(S(\mathbf{r})) \right]^{(t)}$$

where $\mathbf{r}^{(t)} = \left(r_1^{(t)}, \ldots, r_M^{(t)}\right)$, $r_j^{(t)} \sim Poission(f_i)$.

Now we compute the derivatives of I w.r.t. t_i's.

Denote $w = stimWidth$.

Since

$$f_j = w \sum_{k=0}^{M-1} t_{(j-k)\%M} s_k = w \sum_{k=0}^{M-1} t_k s_{(j-k)\%M}$$

We have

$$\frac{\partial I}{\partial t_i} = \sum_j \frac{\partial I}{\partial f_j} \frac{\partial f_j}{\partial t_i} = w \sum_j \frac{\partial I}{\partial f_j} s_{(j-i)\%M}$$

$$grad_t(I) = w \cdot grad_f(I) * \bar{s}, \text{ where } \bar{s}_k = s_{(-k)\%M}.$$

$$\begin{aligned}
\frac{\partial^2 I}{\partial t_i \partial t_j} &= \frac{\partial}{\partial t_j} \left(w \sum_k \frac{\partial I}{\partial f_k} s_{(k-i)\%M} \right) \\
&= w \sum_k \sum_l s_{(k-i)\%M} \frac{\partial^2 I}{\partial f_k \partial f_l} \frac{\partial f_l}{\partial t_j} \\
&= w^2 \sum_k \sum_l s_{(k-i)\%M} s_{(l-j)\%M} \frac{\partial^2 I}{\partial f_k \partial f_l}
\end{aligned}$$

$$hess_t(I) = w^2 \cdot hess_f(I) * \bar{S}, \text{ where } \bar{S}_{k,l} = \bar{s}_k \cdot \bar{s}_l.$$

The constraint:

$$\begin{aligned}
\int f(\theta) d\theta &= const \\
\int \int t(\theta - y) s(y) dy d\theta &= \int s(y) \int t(\theta - y) d\theta dy = \int s(y) dy \cdot \int t(\theta) d\theta
\end{aligned}$$

Therefore, $\int t(\theta) d\theta$ is also a constant.