Efficient Neural Codes under Metabolic Constraints Supplementary Notes

Zhuo Wang *†
Department of Mathematics
University of Pennsylvania
wangzhuo@nyu.edu

Xue-Xin Wei *‡
Department of Psychology
University of Pennsylvania
weixxpku@gmail.com

Alan A. Stocker
Department of Psychology
University of Pennsylvania
astocker@sas.upenn.edu

Daniel D. Lee
Department of Electrical and System Engineering
University of Pennsylvania
ddlee@seas.upenn.edu

A Derivation of Optimal $h_*(s)$

In this section we show how to solve the efficient coding problem formulated in the main submission.

$$\begin{aligned} \text{maximize} \quad & \mathrm{MI}(s,r) \\ \text{subject to} \quad & 0 \le h(s) \le r_{\text{max}}, \quad h'(s) \ge 0 && \text{(range constraint)} \\ & \mathbf{E}_s[K(h(s))] \le K_{\text{total}} && \text{(metabolic constraint)} \end{aligned}$$

A.1 Re-parameterization of s

The initial step is to map the stimulus s to $u = F(s)$ with uniform distribution $U \sim \mathcal{U}[0,1]$. To achieve this, one can simply use the cumulative density function $F(s) = \int_{-\infty}^{s} p(t)\,dt$. We will prove that such mapping does not change the form of our problem. First the objective function is unchanged:

$$\mathrm{MI}(s,r) = \mathrm{MI}(F(s),r) = \mathrm{MI}(u,r) \tag{1}$$

Which can be proved by the fact that $F(s)$ is invertible and applying the information processing inequality twice

$$\mathrm{MI}(s,r) \ge \mathrm{MI}(F(s),r) \ge \mathrm{MI}(F^{-1}(F(s)),r) = \mathrm{MI}(s,r). \tag{2}$$

Now we define

$$g(u) \stackrel{\text{def}}{=} g(F(s)) = h(s) \tag{3}$$

Therefore the saturation constraint is same $0 \le g(u) \le r_{\text{max}}$. Together with the fact that $du = dF(s) = p(s)ds$,

$$\mathbf{E}_u[K(g(u))] = \int_0^1 K(g(u))\,du = \int_{-\infty}^{\infty} K(h(s))p(s)\,ds = \mathbf{E}_s[K(h(s))] \tag{4}$$

*equal contribution
†current affiliation: Center for Neural Science, New York University
‡current affiliation: Department of Statistics and Center for Theoretical Neuroscience, Columbia University

29th Conference on Neural Information Processing Systems (NIPS 2016), Barcelona, Spain.

This shows that the metabolic constraint is also preserved in its format. It suffices to solve the new optimization problem which is on $g(u)$ for a uniformly distributed distributed input u.

$$\begin{aligned} \text{maximize} \quad & \mathrm{MI}(u, r) \\ \text{subject to} \quad & 0 \le g(u) \le r_{\text{max}}, \quad g'(u) \ge 0 \\ & \mathbf{E}_u[K(g(u))] \le K_{\text{total}} \end{aligned}$$

A.2 Lagrangian multiplier approach

Using results from Brunel and Nadal (1998), we show that the objective function is

$$\mathrm{MI}(u, r) = H(U) + \frac{1}{2} \int p(u) \log I_F(u)\, du \tag{5}$$

In our case, a uniform random variable U has zero entropy $H(U) = 0$ and $p(u) = 1_{u \in [0,1]}$. The Fisher information term is determined by the noise model. Assuming $g'(u) > 0$, the leading order term of the mutual information can also be calculated.

$$I_F(u) = T \frac{g'(u)^2}{g(u)^\alpha} + O(1) \quad \Rightarrow \quad I(u, r) \sim \int_0^1 \log g'(u) - \frac{\alpha}{2} \log g(u)\, du + \log(T/V_0) \tag{6}$$

Without loss of generality, we assume $r_{\text{min}} = 0$ and we may rewrite both constraints as integral of $g'(u)$ and $g(u)$.

$$g(1) - g(0) = \int_0^1 g'(u)\, du \le r_{\text{max}} \tag{7}$$

$$\mathbf{E}_u[K(g(u))] = \int_0^1 g(u)^\beta\, du \le K_{\text{total}} \tag{8}$$

We can calculate the Lagrangian of this optimization problem

$$L(g, g') = \log g' - \frac{\alpha}{2} \log g - \lambda_1 g' - \lambda_2 g^\beta \tag{9}$$

which has no explicit dependency on the variable u. Thus we can use the Beltrami's identity to derive the optimal condition for $g(u)$:

$$const = L - g' \frac{\partial L}{\partial g'} = \log g' - \frac{\alpha}{2} \log g - \lambda_1 g' - \lambda_2 g^\beta - (1 - \lambda_1 g') \tag{10}$$

Therefore we have an ordinary differential equation (ODE) as the optimal condition for $g(u)$

$$const = \log g'(u) - \frac{\alpha}{2} \log g(u) - \lambda_2 g^\beta \quad \Rightarrow \quad c_0 = \frac{g'(u)}{g(u)^{\alpha/2}} \exp\left(-\lambda_2 g^\beta\right) \tag{11}$$

To simplify this, we use the substitution $\tilde{g}(u) = \lambda_2 g(u)^\beta$ so that

$$g(u) \propto \tilde{g}(u)^{1/\beta} \tag{12}$$

$$g'(u) \propto \tilde{g}(u)^{1/\beta - 1} \cdot \tilde{g}'(u) \tag{13}$$

$$\Rightarrow \quad c_1 = \frac{d\tilde{g}}{du} \cdot \tilde{g}^{(1-\alpha/2)/\beta - 1} \exp(-\tilde{g}) \tag{14}$$

which is now a separable ODE (denote $q = (1 - \alpha/2)/\beta$),

$$\int c_1\, du = \int \tilde{g}^{q-1} \exp(-\tilde{g})\, d\tilde{g} \tag{15}$$

$$\Rightarrow \quad c_1 u = \gamma_q(\tilde{g}) \quad \Rightarrow \quad \tilde{g}(u) = \gamma_q^{-1}(c_1 u) \tag{16}$$

where $\gamma_q(x)$ is the incomplete gamma function and γ_q^{-1} is its inverse function

$$\gamma_q(x) = \int_0^x z^{q-1} \exp(-z)\, dz, \quad \gamma_q^{-1} \circ \gamma_q(x) = x. \tag{17}$$

For some constant a, b, the optimal tuning curve must take the form

$$g_*(u) \propto \tilde{g}_*(u)^{1/\beta} \tag{18}$$

$$\Rightarrow g_*(u) = \left[\frac{1}{a} \gamma_q^{-1}\left(u \gamma_q(b)\right)\right]^{\frac{1}{\beta}}, \quad h_*(s) = g_*(F(s)) \tag{19}$$

and the constants a, b are chosen such that the saturation and metabolic constraints are satisfied. Below we will find the optimal a, b.

A.3 Determining constants a, b

From the above analysis we know the optimal form of $g(u)$ is

$$g_*(u) = \left[\frac{1}{a}\gamma_q^{-1}\left(u\gamma_q(b)\right)\right]^{\frac{1}{\beta}} \tag{20}$$

where $q = (1-\alpha/2)/\beta$. The optimal solution is a linearly scaled version of the inverse-incomplete-gamma function with two parameters a, b. Now we only need to re-write the constraints and the objective function in terms of a, b and optimize these two parameters.

A.3.1 Objective function

The key part of the objective function is

$$F(a,b) = \int_0^1 \log\left[\frac{g'(u)}{g(u)^{\alpha/2}}\right] du \tag{21}$$

where

$$g'(u) = \frac{1}{\beta}\left[\frac{1}{a}\gamma_q^{-1}(u\gamma_q(b))\right]^{\frac{1}{\beta}-1}\frac{1}{a}\left[\gamma_q^{-1}(u\gamma_q(b))\right]^{1-q}\exp(\gamma_q^{-1}(u\gamma_q(b)))\cdot\gamma_q(b) \tag{22}$$

$$= \frac{\gamma_q(b)}{\beta}\left(\frac{1}{a}\right)^{\frac{1}{\beta}}\left[\gamma_q^{-1}(u\gamma_q(b))\right]^{\frac{1}{\beta}-q}\exp(\gamma_q^{-1}(u\gamma_q(b))) \tag{23}$$

$$g(u)^{\alpha/2} = \left(\frac{1}{a}\right)^{\frac{\alpha}{2\beta}}\left[\gamma_q^{-1}(u\gamma_q(b))\right]^{\frac{\alpha}{2\beta}} \tag{24}$$

Since $q = (1-\alpha/2)/\beta$, we have

$$\frac{g'(u)}{g(u)^{\alpha/2}} = \frac{\gamma_q(b)}{\beta}a^{-q}\exp(\gamma_q^{-1}(u\gamma_q(b))) \tag{25}$$

$$F(a,b) = -\log\beta + \log\gamma_q(b) - q\log a + \int_0^1 \gamma_q^{-1}(u\gamma_q(b))\,du \tag{26}$$

Here we calculate the integral term. We let

$$v(u) = \gamma_q^{-1}(u\gamma_q(b)), \quad v(0) = 0, \quad v(1) = b. \tag{27}$$

$$u\gamma_q(b) = \gamma_q(v), \quad \gamma_q(b)\,du = v^{q-1}\exp(-v)\,dv \tag{28}$$

Therefore

$$\int_0^1 \gamma_q^{-1}(u\gamma_q(b))\,du = \frac{1}{\gamma_q(b)}\int_{v(0)}^{v(1)} v^q\exp(-v)\,dv \tag{29}$$

$$= \frac{1}{\gamma_q(b)}\left[q\int_0^b v^{q-1}\exp(-v)\,dv - v^q\exp(-v)|_0^b\right] \tag{30}$$

$$= \frac{1}{\gamma_q(b)}\left[q\gamma_q(b) - b^q\exp(-b)\right] = q - \frac{b^q\exp(-b)}{\gamma_q(b)} \tag{31}$$

Thus the objective function in terms of a, b is

$$F(a,b) = q - \log\beta + \log\gamma_q(b) - q\log a - \frac{b^q\exp(-b)}{\gamma_q(b)} \tag{32}$$

A.4 Optimal a for fixed value of b

We begin with rewriting the saturation constraint and the metabolic constraint in terms of a, b. First the saturation constraint

$$r_{\max} \geq g_*(1) = \left[\frac{b}{a}\right]^{\frac{1}{\beta}} \quad \Rightarrow \quad a \geq r_{\max}^{-\beta}\cdot b \stackrel{\text{def}}{=} A_1(b) \tag{33}$$

Second the metabolic constraint

$$K_{\text{total}} \geq \int_0^1 K(g(u))\,du = \frac{1}{a}\int_0^1 \gamma_q^{-1}(u\gamma_q(b))\,du = \frac{1}{a}\left[q - \frac{b^q\exp(-b)}{\gamma_q(b)}\right]$$
$$\Rightarrow \quad a \geq K_{\text{ave}}^{-1}\cdot\left[q - \frac{b^q\exp(-b)}{\gamma_q(b)}\right] \stackrel{\text{def}}{=} A_2(b) \tag{34}$$

Based on the form of the objective function $F(a,b)$, it is clear that a should be as small as possible, given that the above two constraints are satisfied. Therefore for fixed value of b, the smallest a that satisfies both constraints is

$$a_*(b) = \max\{A_1(b), A_2(b)\} \tag{35}$$

A.5 Optimal b

Here we discuss in cases. We consider two sets

$$B_1 = \{b \geq 0 | A_1(b) \geq A_2(b)\} \tag{36}$$
$$B_2 = \{b \geq 0 | A_2(b) \geq A_1(b)\} \tag{37}$$

<u>Case I:</u> For $b \in B_1$, $a \geq A_1(b)$ is the tighter constraint therefore $a_*(b) = A_1(b)$. Now we have

$$a_* = A_1(b) = r_{\text{max}}^{-\beta}\cdot b \tag{38}$$
$$F(b) = F(a_*, b) = const + \log\gamma_q(b) - q\log b - \frac{b^q\exp(-b)}{\gamma_q(b)} \tag{39}$$

We will show that $F(a_*, b)$ is non-increasing in b. To prove this, we define an auxiliary function

$$Z(b) \stackrel{\text{def}}{=} \frac{b^q\exp(-b)}{\gamma_q(b)}, \quad \log Z(b) = q\log b - b - \log\gamma_q(b) \tag{40}$$
$$F(b) = const - \log Z(b) - b - Z(b) \tag{41}$$

We need to show

$$0 \geq F'(b) = -\frac{Z'(b)}{Z(b)} - Z'(b) - 1 \tag{42}$$

Here we calculate $Z'(b)$

$$Z'(b) = \frac{qb^{q-1}\exp(-b) - b^q\exp(-b)}{\gamma_q(b)} - \frac{b^{2q-1}\exp(-2b)}{\gamma_q(b)^2} \tag{43}$$
$$= \frac{b^{q-1}\exp(-b)}{\gamma_q(b)^2}\underbrace{[(q-b)\,\gamma_q(b) - b^q\exp(-b)]}_{\stackrel{\text{def}}{=} Z_2(b)} \tag{44}$$

The term $Z_2(b)$ has property

$$Z_2(0) = 0, \quad Z_2'(b) = -\gamma_q(b) \quad \Rightarrow \quad Z_2(b) = -\int_0^b \gamma_q(t)\,dt \tag{45}$$

Therefore

$$Z'(b) = -\underbrace{\frac{b^q\exp(-b)}{\gamma_q(b)}}_{=Z(b)}\underbrace{\frac{\int_0^b \gamma_q(t)\,dt}{b\gamma_q(b)}}_{=M(b)} = -Z(b)M(b) \tag{46}$$

Plug this into $F'(b)$

$$F'(b) = M(b)(1 + Z(b)) - 1 \tag{47}$$

First we can show $F'(0) = 0$. Now for $b > 0$, we have

$$F_2(b) \stackrel{\text{def}}{=} F'(b) \cdot b\gamma_q(b) = \int_0^b \gamma_q(t)\, dt \cdot (1 + Z(b)) - 1 \tag{48}$$

$$F_2'(b) = \gamma_q(b)\,(1 + Z(b)) + \int_0^b \gamma_q(t)\, dt \cdot Z'(b) - \gamma_q(b) - b^q \exp(-b) = \int_0^b \gamma_q(t)\, dt \cdot Z'(b) < 0 \tag{49}$$

Therefore $F'(b) \leq 0$ and the function $F(b)$ is non-increasing. This means that in the case of $b \in B_1$, the optimal solution is the smallest $b_* = \inf_{b \in B_1} b$.

<u>Case II:</u> For $b \in B_2$, $a \geq A_2(b)$ is the tighter constraint therefore $a_*(b) = A_2(b)$. Now we have

$$a_*(b) = A_2(b) = K_{\text{ave}}^{-1}\left[q - \frac{b^q \exp(-b)}{\gamma_q(b)}\right] = K_{\text{ave}}^{-1}\left[q - Z(b)\right] \tag{50}$$

$$F(b) = F(a_*, b) = const + \log \gamma_q(b) - q \log(q - Z(b)) - Z(b) \tag{51}$$

Now we will show this $F(b)$ is non-decreasing in b.

$$F'(b) = \frac{b^{q-1}\exp(-b)}{\gamma_q(b)} + \frac{qZ'(b)}{q - Z(b)} - Z'(b) \tag{52}$$

$$= \frac{Z(b)}{b} + Z'(b)\frac{Z(b)}{q - Z(b)} = \frac{Z(b)}{b(q - Z(b))}\left[q - Z(b) + bZ'(b)\right] \tag{53}$$

The term outside the bracket is positive when $b > 0$. Therefore we only need to show

$$Z_3(b) = q - Z(b) + bZ'(b) \geq 0. \tag{54}$$

Note that we have

$$\log Z(b) = q \log b - b - \log \gamma_q(b) \tag{55}$$

$$\frac{Z'(b)}{Z(b)} = \frac{q}{b} - 1 - \frac{b^{q-1}\exp(-b)}{\gamma_q(b)} = \frac{q}{b} - 1 - \frac{Z(b)}{b} \tag{56}$$

Therefore

$$q - Z(b) = b\left(1 + \frac{Z'(b)}{Z(b)}\right) \tag{57}$$

Plug this into Eq(57) we have

$$Z_3(b) = b\left(1 + \frac{Z'(b)}{Z(b)} + Z'(b)\right) \geq 0 \tag{58}$$

which share the proof of the inequality in Eq(45). In this case, $F(b)$ is non-decreasing. This means that in the case of $b \in B_2$, the optimal solution is the largest $b_* = \sup_{b \in B_2} b$.

<u>Conclusion:</u> Based on these two cases, we know that if both B_1, B_2 are non-empty, then the optimal b_* is both the infimum or B_1 and supremum of B_2, which means that b_* is uniquely determined by

$$A_1(b) = \frac{1}{r_{\text{max}}^{\beta}} b = \frac{1}{K_{\text{ave}}}\left[q - \frac{b^q \exp(-b)}{\gamma_q(b)}\right] = A_2(b) \tag{59}$$

Since $A_1(b)$ grows linearly but $A_2(b)$ has an upper bound, therefore B_1 cannot be empty. However if B_2 is empty, then the optimal $b_* = \inf_{b \in \mathbf{R}^+} b = 0$ which means that the optimal solution is attained by the limit $b \to 0$.

B Technical Details for Multiple Neurons Case

First we define the active regions of the i-th neuron as

$$A_i^{\pm} = \{s | \pm h_i'(s) > 0\}, \quad A_i = A_i^- \cup A_i^+. \tag{60}$$

Now we prove a couple of necessary conditions for these A_i to be optimal in terms of maximum mutual information. Note that the tuning curves are assumed to be monotonic so one of A_i^+ and A_i^- must be empty.

As a useful preliminary result, we recall that the total Fisher information of the population is the linear sum of Fisher information contributed by each individual neuron

$$I_F(s) = \sum_{i=1}^{N} I_i(s) \quad \text{where} \quad I_i(s) \propto \frac{h_i'(s)^2}{h_i(s)^\alpha} \tag{61}$$

if the noise model parameter is α. It is clear that $I_i(s)$ is greater than zero only if $s \in A_i$.

Lemma 1 (Non-overlapping active regions.)**.** *We consider the problem of optimizing a neural population with neuron $i = 1, \ldots, N$. We limit the stimulus to be on some subset $s \in [s_0, s_1]$ of the original range $[0, 1]$. Each neuron is monotonic (either $h_i'(s) \geq 0$ or $h_i'(s) \leq 0$ for $s \in [s_0, s_1]$) and has limited range of output $L_i \leq h_i(s) \leq H_i$.*

$$\textit{maximize} \quad \int_{s_0}^{s_1} \log I_F(s)\, ds \tag{62}$$

$$\textit{subject to} \quad L_i \leq h_i(s) \leq H_i, \quad i = 1, \ldots, N \tag{63}$$

*Then a necessary condition is the **non-overlapping active regions**, i.e. $A_i \cap A_j = \emptyset$ for all $i \neq j$.*

Proof. We begin with the proof of an upper bound on the integral of the square root of the Fisher information:

$$\int_{s_0}^{s_1} \sqrt{I_i(s)}\, ds \leq I_i^{\max} \tag{64}$$

For $\alpha \neq 2$, for example, we have

$$\sqrt{I_i(s)} \propto \frac{|h_i'(s)|}{h_i(s)^{\alpha/2}} \propto \frac{d}{ds}\left[h_i(s)^{1-\alpha/2}\right] \tag{65}$$

$$I_i^{\max} \propto |H_i^{1-\alpha/2} - L_i^{1-\alpha/2}| \tag{66}$$

For $\alpha = 2$ the calculation is similar if we use $\log h_i(s)$ as the anti-derivative. Next we write down an upper bound of the objective function:

$$I_F(s) = \sum_{i=1}^{N} I_i(s) = \left(\sum_{i=1}^{N} \sqrt{I_i(s)}\right)^2 - 2\sum_{i<j} \sqrt{I_i(s) \cdot I_j(s)} \leq \left(\sum_{i=1}^{N} \sqrt{I_i(s)}\right)^2 \stackrel{\text{def}}{=} Q(s)^2 \tag{67}$$

$$\int_{s_0}^{s_1} \log I_F(s)\, ds \leq 2\int_{s_0}^{s_1} \log Q(s)\, ds = 2(s_1 - s_0)\int_{s_0}^{s_1} \frac{1}{s_1 - s_0} \log Q(s)\, ds \tag{68}$$

$$\leq 2(s_1 - s_0) \log \int_{s_0}^{s_1} \frac{Q(s)}{s_1 - s_0}\, ds \leq 2(s_1 - s_0) \log \frac{\sum_i I_i^{\max}}{s_1 - s_0}. \tag{69}$$

where we have used the Jensen's inequality and the optimization constraints. To achieve this attainable upper bound for the objective function, we need $Q(s) = const$ for the Jensen's inequality and also the equality in Eq. 67. Therefore a necessary condition for $h_i(s)$ to be optimal is that $I_i(s) \cdot I_j(s) = 0$ everywhere for $i \neq j$. This is equivalent as our claim $A_i \cap A_j = \emptyset$ for all $i \neq j$. □

In other words, different neurons should not have non-overlapping active region. However, the above lemma does not take the energy constraints into consideration. If we add the energy constraint

$$\int_0^1 \sum_{i=1}^{N} K(h_i(s))\, ds \leq K_{\text{total}} \tag{70}$$

does it break the necessity of the non-overlapping Fisher information condition? The answer is no due to the following lemma.

Lemma 2 (Non-overlapping active regions with metabolic constraints)**.** *Assuming $h_i(s)$ is the optimal solution to the following problem. Each neuron is monotonic (either $h_i(s) \geq 0$ or $h_i(s) \leq 0$ for $s \in [0,1]$) and has limited range of output $L \leq h_i(s) \leq H$.*

$$\textit{maximize} \quad \int_0^1 \log I_F(s)\, ds \tag{71}$$

$$\textit{subject to} \quad L \leq h_i(s) \leq H, \quad i = 1, \ldots, N \tag{72}$$

$$\int_0^1 \sum_{i=1}^N K(h_i(s))\, ds \leq K_{\textit{total}} \tag{73}$$

*Then a necessary condition is the **non-overlapping active regions**, i.e. $A_i \cap A_j = \emptyset$ for all $i \neq j$.*

Proof. We show this lemma by contradiction – we assume $h_i(s)$ is optimal with $I_i(s) \cdot I_j(s) > 0$ for some s (so $s \in A_i \cap A_j$) and show that there exists a better solution $\tilde{h}_i(s)$.

We divide the stimulus space $s \in [0,1]$ equally into M smaller intervals with endpoints $s_j = j/M$ for $j = 0, \ldots M$. We define

$$L_{i,j} = \min_{s \in [s_{j-1}, s_j]} h_i(s) \tag{74}$$

$$H_{i,j} = \max_{s \in [s_{j-1}, s_j]} h_i(s) \tag{75}$$

On each of these intervals $[s_{j-1}, s_j]$ and the above range, we apply Lemma 1 and obtain a new solution $\tilde{h}_i(s)$ which satisfies the non-overlapping Fisher information condition. It is easy to see that this new solution gives better objective function. Next we show that this better solution costs similar amount of energy as $h_i(s)$. Using the upper and lower bound of firing rate in each interval, we have

$$\int_0^1 K(\tilde{h}_i(s))\, ds = \sum_{j=1}^M \int_{s_{j-1}}^{s_j} K(\tilde{h}_i(s))\, ds \leq \sum_{j=1}^M (s_j - s_{j-1}) K(H_{i,j}) = \frac{1}{M} \sum_{j=1}^M K(H_{i,j}) \tag{76}$$

$$\int_0^1 K(\tilde{h}_i(s))\, ds = \sum_{j=1}^M \int_{s_{j-1}}^{s_j} K(\tilde{h}_i(s))\, ds \geq \sum_{j=1}^M (s_j - s_{j-1}) K(L_{i,j}) = \frac{1}{M} \sum_{j=1}^M K(L_{i,j}) \tag{77}$$

Similarly for the original solution $h_i(s)$ these two bounds also apply

$$\frac{1}{M} \sum_{j=1}^M K(L_{i,j}) \leq \int_0^1 K(h_i(s))\, ds \leq \frac{1}{M} \sum_{j=1}^M K(H_{i,j}) \tag{78}$$

Therefore

$$|K_{\tilde{h}} - K_h| = \left| \int_0^1 K(\tilde{h}_i(s))\, ds - \int_0^1 K(h_i(s))\, ds \right| \tag{79}$$

$$\leq \frac{1}{M} \sum_{j=1}^M (H_{i,j} - L_{i,j}) = \frac{1}{M}(H - L) \tag{80}$$

and the right side converges to zero as M goes to infinity. This means that the energy consumption of this new solution $\tilde{h}_i(s)$ can be made as close to the original solution as possible if we use a finer and finer grid (large M), while having a better objective function value. This contradicts the optimality of $h_i(s)$. □

Using Lemma 1 and Lemma 2, we conclude that in the optimal population, the neurons need to have non-overlapping Fisher information. This simplifies our further analysis. Here we discuss another necessary condition that a pair of ON-OFF neurons must satisfies.

Lemma 3 (ON-OFF neurons)**.** *Under the assumption that the energy constraint is binding, then for an ON-neuron with active region A_i^+ and an OFF-neuron with active region A_j^-, we have $\sup B_j^- \leq \inf A_i^+$. In other words, the active region of any ON neuron is strictly on the right side of the active region of any OFF neuron.*

Proof. We denote $s_i = \inf A_i^+$ and $s_j = \sup A_j^-$ and prove the lemma by contradiction. We assume $s_i < s_j$. Due to the piecewise continuity of $h_i'(s)$, there exists $\epsilon > 0$ such that there exist small neighborhoods $[s_i, s_i + \epsilon] \in A_i^+$ and $[s_j - \epsilon, s_j] \in A_j^-$. We can construct a new tuning curve $\tilde{h}_i(s)$ and $\tilde{h}_j(s)$ by swapping their active regions (see Figure 1 below) in these neighborhood. It

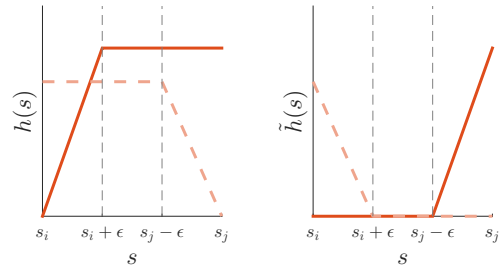

Figure 1: Tuning curve surgery for ON-OFF neuron pairs that reduces energy costs.

is obvious that the new $\tilde{h}_i$ costs strictly less amount of energy. The new tuning curves has equal performance in terms of objective function because the regions being swapped has same size. The existence of such tuning curves contradicts the fact that the energy constraint is binding. □

One immediate corollary is that, in a large population of neurons with both ON and OFF sub-populations, there exists a single s that divides the active regions for the ON sub-population and the OFF sub-population.

Next we find the optimal condition for a population of only ON/OFF neurons. Without loss of generality, we assume the neural population consists of only ON neurons.

Lemma 4 (ON-ON neurons)**.** *Assuming the population has only ON neurons and the metabolic cost function is linear $\sum_i K(h_i(s)) = K(\sum_i h_i(s))$, then the optimal $h_i(s) \approx h(s)/N$ but with disjoint active regions A_i^+. The function $h(s) = \sum_i h_i(s)$ is the single neuron infomax solution of:*

$$\textit{maximize}_{h(s)} \quad \mathrm{MI}(s, r) \tag{81}$$

$$\textit{subject to} \quad 0 \le h(s) \le N \cdot r_{\textit{max}} \tag{82}$$

$$\mathbf{E}[K(h(s))] \le N \cdot K_{\textit{total}} \tag{83}$$

Proof. We denote $h_i(s) = p_i(s) \cdot h(s)$ and it is clear that $\sum p_i(s) = 1$. Using Lemma 1 we know A_i's are disjoint therefore we also have $h_i'(s) = h'(s) \cdot 1_{A_i}$. Plug these into the objective function, we know that

$$I_F(s) \propto \sum_{i=1}^{N} \frac{h_i'(s)^2}{h_i(s)^\alpha} = \frac{h'(s)^2}{h(s)^\alpha} \cdot \sum_{i=1}^{N} 1_{A_i} \cdot p_i(s)^{-\alpha} \quad \Rightarrow \tag{84}$$

$$\int_0^1 \log I_F(s)\, ds = \int_0^1 \log \frac{h'(s)^2}{h(s)^\alpha}\, ds + \int_0^1 \log \left(\sum_{i=1}^{N} 1_{A_i} \cdot p_i(s)^{-\alpha} \right) ds \tag{85}$$

Now the problem is divided into two independent part. The first part involves finding the optimal $h(s)$ under constraints stated in the lemma. This part is exactly the same as the single neuron case.

The second part involves optimizing A_i and $p_i(s)$ for the following term

$$\text{maximize} \quad \int_0^1 \log \left(\sum_{i=1}^{N} 1_{A_i} \cdot p_i(s)^{-\alpha} \right) ds = -\alpha \sum_i \int_0^1 1_{A_i} \log p_i(s)\, ds \tag{86}$$

$$\text{subject to} \quad \sum p_i(s) = 1. \tag{87}$$

Using Lagrange multiplier method we know the optimal condition for $p_i(s)$ is

$$-\alpha \frac{1_{A_i}}{p_i(s)} - \lambda = 0 \tag{88}$$

This shows that $p_i(s) = const$ for $s \in A_i$. However A_i is the active region of neuron i so the function $h_i(s)$ is increasing and all other $h_j(s)$ remains the same. The only way for this condition to holds is when A_i consists of infinite many small intervals so that the increase in $h_i(s)$ is small. Also we know that all $p_i(s) \to 1/N$ when $s \to 1$. Therefore one possible solution is given by $h_i(s) \approx h(s)/N$ but on infinitesimal scales, each small interval is equally divided into N disjoint set A_i's. $\square$