A Generalization of the Equal Coding Theorem

Hui-An Shen, Stefan M. Moser, and Jean-Pascal Pfister

Abstract—We reformulate the Equal Coding Theorem in sensory neural encoding with ON- and OFF-neurons as a channel capacity problem. We then present a capacity-based proof of the Equal Coding Theorem, and generalize it to neurons with different firing probabilities. We also briefly discuss the biological implications of this generalization.

I. INTRODUCTION

The efficient coding hypothesis states that sensory neurons maximize the information about the environment [1], [2]. This hypothesis is one of the most influential theory in systems neuroscience and was the first one to apply Shannon's information theory to the problem of neural coding. This theory has been successfully applied in different systems (such as in invertebrate vision [3], vertebrate vision [4]) as well as at different levels (synaptic plasticity [5], neuronal adaptation [6], neural network level [4], [7]).

In this paper we investigate the setup of a system with n neurons that encode a random one-dimensional stimulus S, that we assume to be uniform on $[0, 1]$. We consider two types of neurons: An *ON-neuron* will remain inactive as long as the stimulus S is below a given neuron-specific threshold θ_i. If S is larger than this threshold, the neuron will emit a spike with a given neuron-specific firing-probability p_i. The *OFF-neurons* are identical apart from that they will remain inactive if the stimulus S is *larger* than the threshold θ_i and fire (with probability p_i) if the stimulus is *below* the threshold. Thus, the output of each neuron behaves like the output of a Z-channel with 1-error probability $1 - p_i$, whose input is binary with 0-probability θ_i (ON-neuron) or $1 - \theta_i$ (OFF-neuron).

According to the efficient coding hypothesis, the goal of the system is to maximize the mutual information between the stimulus S and the outputs of the n neurons $\mathbf{Y}$ under the system's biophysical constraints. Our question here is twofold: For given firing probabilities p_i, (1) what is the best ratio of ON- to OFF-neurons and (2) how do we best set the thresholds θ_i of these neurons?

This problem is tightly linked to "nonadaptive 20-questions", and an example of such questions can be formulated as follows. Let S be a real number uniformly sampled from $[0, 1]$. We are allowed to ask n yes–no questions of the following two types: "Is S greater than θ_i?" or "Is S smaller than θ_j?". Goal is to estimate the value of S. The yes–no answers are corrupted by asymmetric noise as in a Z-channel, i.e., there is a certain probability that a yes is distorted into a no, while a no always remains a no. Here the question is how do we choose the thresholds θ_i *a priori* so that the guesser's estimation error is minimized?

Returning to our neural-inspired setup, instead of minimizing the estimation error, we aim to maximize the mutual information between the random variable $S \sim \mathcal{U}([0, 1])$ and the corrupted neuronal responses $\mathbf{Y} \in \{0, 1\}^n$.

In [8], Gjorgjieva *et al.* investigated this topic and proposed the *Equal Coding Theorem*, which states that the mutual information between the input stimulus and the neuron responses is identical for *any* fraction of ON-neurons provided that (i) all thresholds of the OFF-neurons are smaller than those of the ON-neurons and (ii) all the thresholds are chosen optimally in the sense that the mentioned mutual information is maximized.

In this paper, we provide three contributions. Firstly, we generalize the theorem by allowing the neurons to have different firing rates. Secondly, we provide a simpler, information-theoretic proof that relies on a dual formulation of channel capacity. Thirdly, we highlight the biological consequences of this generalized theorem.

II. PROBLEM SETUP AND REFORMULATION OF THE EQUAL CODING THEOREM

Let $S \sim \mathcal{U}([0, 1])$ denote the stimulus to be estimated. Let Y_i, $i \in [n] \triangleq \{1, 2, \ldots, n\}$, denote the activities of a population of n Bernoulli neurons, which are characterized by their firing thresholds $0 < \theta_1 < \cdots < \theta_n < 1$, their firing probabilities $p_i \in [0, 1]$, and their types z_i ($z_i = 0$ denotes an ON-neuron and $z_i = 1$ denotes an OFF-neuron).

The activity of those neurons can thus be expressed as

$$Y_i \sim \text{Bernoulli}\big(p_i f_i(S)\big), \tag{1}$$

where

$$f_i(S) \triangleq \Theta\big((-1)^{z_i}(S - \theta_i)\big) \tag{2}$$

is the activation function and $\Theta(\cdot)$ is the Heaviside step function. See Fig. 1 for an illustration.

Once all thresholds θ_i are set, we can divide the stimulus interval $[0, 1]$ into $n + 1$ intervals of length $\theta_k - \theta_{k-1}$, $k \in [n + 1]$, where we set $\theta_0 \triangleq 0$ and $\theta_{n+1} \triangleq 1$. The identity of those intervals can be described by n-dimensional binary vectors (codewords) $\mathbf{c}_k \in \{0, 1\}^n$, $k \in [n + 1]$ that are set according to the activation of each neuron, i.e.,

$$\mathbf{c}_k \triangleq \mathbf{f}(s) \quad \text{for } s \in (\theta_{k-1}, \theta_k). \tag{3}$$

The collection of those $n + 1$ codewords is denoted as the *n-configuration of codewords* $\mathsf{C}_n = (\mathbf{c}_1, \ldots, \mathbf{c}_{n+1})$. Note that those codewords have specific properties. For example, the first codeword is a direct readout of the neuron types $\mathbf{z} = (z_1, \ldots, z_n)^\mathsf{T}$: $\mathbf{c}_1 = \mathbf{z}$. Due to the linear arrangement of the

H.-A. Shen and J.-P. Pfister are with the University of Bern. S.M. Moser is with ETH Zürich and NYCU Taiwan. H.-A. Shen was partially supported by the SNF grant 31003A_175644.
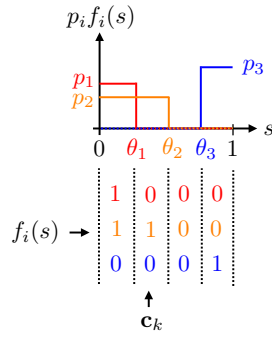
Fig. 1. *Top*: Response functions of two OFF-neurons (red and orange) and one ON-neuron (blue).
Bottom: Those 3 response functions can be translated into 4 codewords $\mathbf{c}_1, \ldots, \mathbf{c}_4$. Since the OFF-neurons' thresholds θ_1 and θ_2 are smaller than the ON-neuron's threshold θ_3, there exists a zero codeword (i.e., $\mathbf{c}_3 = (0, 0, 0)^\mathsf{T}$).

neurons, the components of the subsequent codeword can be recursively expressed as

$$(\mathbf{c}_{k+1})_i = (\mathbf{c}_k)_i + \delta_{ik} \mod 2, \qquad i, k \in [n], \qquad (4)$$

where $\delta_{ik} \triangleq \mathbb{1}\{i = k\}$ is the Kronecker delta. Thus, any configuration is uniquely specified by its type-vector $\mathbf{z} = \mathbf{c}_1$. This is best illustrated with examples; see also Fig. 1.

Example 1: For $n = 3$,

$$\mathsf{C}_3 = \left(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right) \qquad (5)$$

is a 3-configuration with the underlying sequence of neurons being "OFF, OFF, ON", as read from left to right in Fig. 1. $\Diamond$

Example 2: For $n = 3$,

$$\mathsf{C}_3' = \left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right) \qquad (6)$$

is a 3-configuration with the underlying sequence of neurons being "OFF, OFF, OFF". $\Diamond$

We denote the collection of all n-configurations as $\mathscr{U}_n$.

Note that if the neurons were noiseless (i.e., $p_i = 1$ for all neurons), it would be optimal (again in the sense of maximizing the mutual information between S and the neuron activity vector $\mathbf{Y} = \mathbf{C}$) to choose the thresholds equally spaced in order to create $n + 1$ equally long intervals. In this case each n-configuration would be equivalent as any one does perfectly distinguish the $n + 1$ intervals.

In our setting, however, each binary neuronal response is corrupted by independent asymmetric flipping noise that is modeled as a Z-channel. Therefore, the codewords are firstly passed through an n-fold Z-channel (with different 1-error probabilities), and thus different n-configurations show in general a different performance.

Next, we will reformulate the Equal Coding Theorem, which was introduced in Section I, in the context of channel capacity and by identifying the set of all length-n binary codewords of an n-configuration as the n-fold Z-channel's

input alphabet. In this framework, the Equal Coding Theorem essentially states that for any $n \in \mathbb{N}$, the channel capacity is the same for all n-configurations that contain the zero codeword (i.e., a zero vector). To that goal, we give the following definition.

Definition 3 (Zero-Codeword Configurations): We define the family of n-configurations containing the zero codeword as

$$\mathscr{O}_n \triangleq \{\mathsf{C}_n \mid \mathbf{0} \in \mathsf{C}_n, \mathsf{C}_n \in \mathscr{U}_n\}. \qquad (7)$$

To recapitulate, we consider the setting where a mixture of n ON/OFF-neurons encodes for a stimulus random variable $S \sim \mathcal{U}([0, 1])$. The binary neuronal responses to the stimulus then gives a length-n codeword $\mathbf{C}$, which is then transmitted through an n-fold Z-channel, resulting in a channel output that we denote by $\mathbf{Y}$.

In our framework, the Equal Coding Theorem concerns the mutual information between the stimulus S and the said channel output $\mathbf{Y}$, denoted $I(S; \mathbf{Y})$. Note that this mutual information is a function of both the input alphabet of the n-fold Z-channel (i.e., the n-configuration of codewords) and the probability mass function (PMF) over the input alphabet (which is a direct function of the thresholds via the length of the intervals $\theta_{k+1} - \theta_k$). And thus, for a given n-configuration of codewords C_n and a given PMF $\pi \in \mathscr{P}_{n+1}$ (with $\mathscr{P}_{n+1}$ describing the collection of all PMFs over $n+1$ elements) over the codewords, we write $I(S; \mathbf{Y})[\mathsf{C}_n, \pi]$ to denote the mutual information in question.

Now we can reformulate the Equal Coding Theorem as follows.

Theorem 4 (Equal Coding Theorem, reformulated): For all $\mathsf{C}_n, \mathsf{C}_n' \in \mathscr{O}_n$,

$$\sup_{\pi \in \mathscr{P}_{n+1}} I(S; \mathbf{Y})[\mathsf{C}_n, \pi] = \sup_{\pi \in \mathscr{P}_{n+1}} I(S; \mathbf{Y})[\mathsf{C}_n', \pi]. \qquad (8)$$

The original Equal Coding Theorem is only for the setting where all neuronal binary flipping noise is identical, i.e., the single Z-channels forming the n-fold Z-channel all have the same 1-error q. In this paper, we generalize Theorem 4 and show that it holds also for the setting where each neuron has a different binary flipping noise modeled respectively by 1-error probability $q_i \triangleq 1 - p_i$.

III. CAPACITY-BASED PROOF OF EQUAL CODING THEOREM

Recall that we denote the stimulus by $S \sim \mathcal{U}([0, 1])$, the codeword by $\mathbf{C} \in \{0, 1\}^n$, and the response by $\mathbf{Y} \in \{0, 1\}^n$, where n is the number of neurons.

Lemma 5: We have $I(\mathbf{C}; \mathbf{Y}) = I(S; \mathbf{Y})$.

Proof: Due to the Markov chain $S \multimap \mathbf{C} \multimap \mathbf{Y}$ and the Data Processing Inequality, we have $I(\mathbf{C}; \mathbf{Y}) \geq I(S; \mathbf{Y})$. But as $\mathbf{C} = \mathbf{f}(S)$, we also have $I(\mathbf{C}; \mathbf{Y}) = I(\mathbf{f}(S); \mathbf{Y}) \leq I(S; \mathbf{Y})$, where the latter inequality follows again by the Data Processing Inequality. $\blacksquare$

Next, we note that for $n \geq 3$, any n-configuration with its $n + 1$ input codewords generates more than $n + 1$ possible channel outputs $\mathbf{Y}$. However, if the n-configuration contains

the zero codeword, it is possible to classify the channel outputs into $n + 1$ categories of "pooled responses", where $\mathbf{Y}$ is classified according to the OFF-neuron of lowest threshold that fires, the zero codeword (no neuron fires), or the ON-neuron of highest threshold that fires. We denote this pooled response vector by $\mathbf{R}$. This categorization is best illustrated with two examples.

Example 6: For the 3-configuration in Example 1, the pooled response is

$$\mathsf{R}_3 = \left(\begin{pmatrix} 1 \\ t \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right). \tag{9}$$

Here, the channel outputs $(1,1,0)^{\mathsf{T}}$ and $(1,0,0)^{\mathsf{T}}$ are pooled into the category $(1,t,0)^{\mathsf{T}}$, where we use t to denote an unspecified element (0 or 1). $\diamond$

Example 7: For the 3-configuration in Example 2, the pooled response is

$$\mathsf{R}_3' = \left(\begin{pmatrix} 1 \\ t \\ t \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ t \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right). \tag{10}$$

The channel outputs $(1,1,1)^{\mathsf{T}}$, $(1,1,0)^{\mathsf{T}}$, $(1,0,1)^{\mathsf{T}}$, $(1,0,0)^{\mathsf{T}}$ are pooled into the category $(1,t,t)^{\mathsf{T}}$; and $(0,1,1)^{\mathsf{T}}$, $(0,1,0)^{\mathsf{T}}$ are pooled into the category $(0,1,t)^{\mathsf{T}}$. $\diamond$

Lemma 8: For any n and an n-configuration containing the zero codeword, it holds that $I(\mathbf{C}; \mathbf{Y}) = I(\mathbf{C}; \mathbf{R})$, where $\mathbf{R}$ denotes the pooled response vector as defined above.

Proof: We note that $H(\mathbf{C}|\mathbf{Y} = \mathbf{y})$ is the same for all $\mathbf{y}$ that are pooled together to the same pooled response vector. Thus, $I(\mathbf{C}; \mathbf{Y}) = H(\mathbf{C}) - H(\mathbf{C}|\mathbf{Y}) = H(\mathbf{C}) - H(\mathbf{C}|\mathbf{R}) = I(\mathbf{C}; \mathbf{R})$. ∎

Thus, in the following we will now focus completely on the "pooled channel" with input $\mathbf{C}$ and output $\mathbf{R}$. We use $\pi \in \mathscr{P}_{n+1}$ for PMFs on the input alphabet, ϕ for the conditional channel distribution, and $\psi \in \mathscr{P}_{n+1}$ for PMFs on the output alphabet. In particular, we use the shorthand

$$\phi(\cdot|k) \triangleq P_{\mathbf{R}|\mathbf{C}}(\cdot|\mathbf{c}_k), \quad k \in [n+1], \tag{11}$$

to denote the channel output distribution given that $\mathbf{C} = \mathbf{c}_k$ (where we assume some given configuration C_n). Moreover, we use $(\pi\phi)$ to describe the output PMF induced by the input PMF π. Finally, $\mathscr{D}(\cdot \| \cdot)$ denotes the *relative entropy* (Kullback-Leibler divergence) between probability distributions.

The capacity C of this channel (again for some given configuration C_n) is given as

$$\mathsf{C} \triangleq \max I(\mathbf{C}; \mathbf{R}) = \max_{\pi \in \mathscr{P}_{n+1}} I(\pi, \phi), \tag{12}$$

which can also be written using the well-known dual expression for capacity [9, Sec. 16.2]:

$$\mathsf{C} = \min_{\psi \in \mathscr{P}_{n+1}} \max_{k \in [n+1]} \mathscr{D}(\phi(\cdot|k) \| \psi). \tag{13}$$

Note that the optimal capacity-achieving output distribution $\bar{\psi}$ (which is the minimizer of (13)) can be seen as the
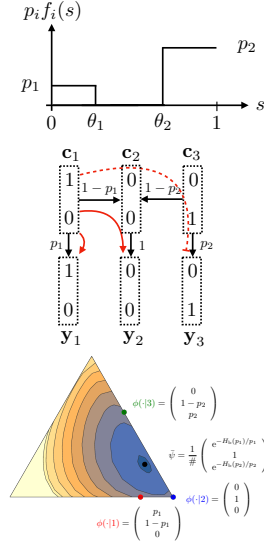

Fig. 2. *Top*: Response function of one ON-neuron and one OFF-neuron. *Middle*: Transition diagram from input codewords $\mathbf{c}_k$ to output responses $\mathbf{y}_i$. In order to obtain the transition probability $\phi(i|k) = \Pr[\mathbf{Y} = \mathbf{y}_i | \mathbf{C} = \mathbf{c}_k]$, it is sufficient to follow the black arrows from $\mathbf{c}_k$ to $\mathbf{y}_i$ and multiply their respective probabilities. E.g., the transition from $\mathbf{c}_1$ to $\mathbf{y}_2$ is denoted by a red arrow and is given by $\phi(2|1) = (1-p_1)\cdot 1$. Note that the transition from $\mathbf{c}_1$ to $\mathbf{y}_3$ is impossible (red dashed arrow) and hence the transition probability is 0. *Bottom*: Representation of the response vectors $\phi(\cdot|1)$, $\phi(\cdot|2)$, and $\phi(\cdot|3)$ on the 2-simplex. The capacity-achieving output distribution $\bar{\psi}$ is such that the "distance" (in the $\mathscr{D}$-sense, denoted in color-code) from the response vectors are identical (see Eq. (14)). $H_{\mathrm{b}}(p_i) \triangleq -p_i \log(p_i) - (1-p_i) \log(1-p_i)$ denotes the binary entropy function and $\#$ denotes the normalization constant.

"circumcenter" of all the conditional output PMFs $\phi(\cdot|k)$ (see Fig. 2).

We now have the following result that holds generally for any (discrete and memoryless) channel, but that we formulate directly for our n-fold Z-channel.

Proposition 9: If there exists some output PMF $\bar{\psi}$ that is induced by some input PMF $\bar{\pi}$ and that satisfies for some constant $\bar{c} \geq 0$,

$$\mathscr{D}(\phi(\cdot|k) \| \bar{\psi}) \begin{cases} = \bar{c} & \forall k \text{ with } \bar{\pi}(k) > 0, \\ \leq \bar{c} & \forall k \text{ with } \bar{\pi}(k) = 0, \end{cases} \tag{14}$$

then

1) $\bar{\psi}$ is the capacity-achieving output distribution, i.e.,

$$\bar{\psi} = \operatorname*{argmin}_{\psi \in \mathscr{P}_{n+1}} \max_{k \in [n+1]} \mathscr{D}(\phi(\cdot|k) \| \psi); \tag{15}$$

2) $\bar{c}$ is the capacity, $\bar{c} = \mathsf{C}$;
3) $\bar{\psi}$ is unique.

Proof: Note that for arbitrary π, ϕ, and ψ,

$$I(\pi, \phi) = \sum_{k \in [n+1]} \sum_{j \in [n+1]} \pi(k)\phi(j|k) \log\left(\frac{\phi(j|k)}{(\pi\phi)(j)}\right) \tag{16}$$

$$= \sum_{k \in [n+1]} \sum_{j \in [n+1]} \pi(k)\phi(j|k) \log\left(\frac{\phi(j|k)}{(\pi\phi)(j)} \cdot \frac{\psi(j)}{\psi(j)}\right) \tag{17}$$

$$= \sum_{k \in [n+1]} \pi(k) \mathscr{D}(\phi(\cdot|k) \| \psi) - \mathscr{D}((\pi\phi) \| \psi). \tag{18}$$

Now, it follows from (14) that

$$\bar{c} = \sum_{k \in [n+1]} \bar{\pi}(k)\bar{c} = \sum_{k \in [n+1]} \bar{\pi}(k)\,\mathscr{D}\big(\phi(\cdot|k)\,\big\|\,\bar{\psi}\big) \quad (19)$$

$$= \mathrm{I}(\bar{\pi}, \phi) + \mathscr{D}\big((\bar{\pi}\phi)\,\big\|\,\bar{\psi}\big) = \mathrm{I}(\bar{\pi}, \phi) + \mathscr{D}\big(\bar{\psi}\,\big\|\,\bar{\psi}\big) \quad (20)$$

$$= \mathrm{I}(\bar{\pi}, \phi) \leq \mathsf{C}, \quad (21)$$

where the first equality in (20) follows from (18).

On the other hand, for any capacity-achieving PMF π^*,

$$\bar{c} = \sum_{k \in [n+1]} \pi^*(k)\bar{c} \geq \sum_{k \in [n+1]} \pi^*(k)\,\mathscr{D}\big(\phi(\cdot|k)\,\big\|\,\bar{\psi}\big) \quad (22)$$

$$= \mathrm{I}(\pi^*, \phi) + \mathscr{D}\big((\pi^*\phi)\,\big\|\,\bar{\psi}\big) = \mathsf{C} + \mathscr{D}\big((\pi^*\phi)\,\big\|\,\bar{\psi}\big) \quad (23)$$

$$\geq \mathsf{C}, \quad (24)$$

where (22) again follows from (14), but potentially with an inequality because there might exist some k for which $\bar{\pi}(k) = 0$, but $\pi^*(k) > 0$.

We have therefore proven that $\bar{c} = \mathsf{C}$. Moreover, the KKT conditions for capacity [10, Sec. 12.5] state that some PMF $\bar{\pi}$ is capacity achieving if, and only if,

$$\mathscr{D}\big(\phi(\cdot|k)\,\big\|\,(\bar{\pi}\phi)\big) \begin{cases} = \mathsf{C} & \forall k \text{ with } \bar{\pi}(k) > 0, \\ \leq \mathsf{C} & \forall k \text{ with } \bar{\pi}(k) = 0. \end{cases} \quad (25)$$

Thus, it follows from (14) that $\bar{\pi}$ must be capacity achieving and thus that $\bar{\psi}$ is the capacity-achieving output distribution. The uniqueness of $\bar{\psi}$ can be shown using a simple convexity argument. ∎

Proposition 9 can now be used for some output PMF $\bar{\psi}$ to obtain a system of $n + 1$ equations with $n + 1$ variables. Concretely, this is done by first assuming that for (14), equality holds also in the case of $\bar{\pi}(k) = 0$, and after solving for $\bar{\psi}$ from the system of equations, further showing that there exists an input PMF that induces $\bar{\psi}$, and therefore $\bar{\psi}$ is indeed the unique capacity-achieving output distribution. Before writing out the details, we again first present this idea with the help of the two examples above.

Example 10: The conditional channel PMF for input configuration C_3 and pooled response R_3 from Example 1 and 6, respectively, is

$$\phi = \begin{pmatrix} 1 - q_1 & 0 & 0 & 0 \\ q_1(1 - q_2) & 1 - q_2 & 0 & 0 \\ q_1 q_2 & q_2 & 1 & q_3 \\ 0 & 0 & 0 & 1 - q_3 \end{pmatrix}. \quad (26)$$

As the third column corresponds to the zero codeword, we define $\bar{\psi} \triangleq (\psi_1, \psi_2, \psi_0, \psi_3)^\mathsf{T}$, where we on purpose use an unusual numbering with ψ_0 at the position of the zero codeword. Using Proposition 9 as mentioned beforehand, we obtain the following equations:

$$(1 - q_1)\log\frac{1 - q_1}{\psi_1} + q_1(1 - q_2)\log\frac{q_1(1 - q_2)}{\psi_2}$$
$$+ q_1 q_2 \log\frac{q_1 q_2}{\psi_0} = \bar{c} \quad (27a)$$

$$(1 - q_2)\log\frac{1 - q_2}{\psi_2} + q_2 \log\frac{q_2}{\psi_0} = \bar{c} \quad (27b)$$

$$\log\frac{1}{\psi_0} = \bar{c} \quad (27c)$$

$$q_3 \log\frac{q_3}{\psi_0} + (1 - q_3)\log\frac{1 - q_3}{\psi_3} = \bar{c} \quad (27d)$$

By Gaussian elimination steps $(27b) - q_2(27c)$, $(27a) - q_1(27b)$, and $(27d) - q_3(27c)$, we obtain

$$\frac{\psi_j}{\psi_0} = (1 - q_j)q_j^{\frac{q_j}{1 - q_j}}, \quad j \in \{1, 2, 3\}, \quad (28)$$

and by normalization

$$\psi_0^{-1} = 1 + \sum_{j=1}^{3}(1 - q_j)q_j^{\frac{q_j}{1 - q_j}}. \quad (29)$$

Note that it is not difficult to show (by matrix inversion of (26)) that there indeed exists a PMF $\bar{\pi}$ that achieves $\bar{\psi}$:

$$\bar{\pi} = \left(\psi_0 \cdot q_1^{\frac{q_1}{1 - q_1}}, \; \psi_0 \cdot \left(q_2^{\frac{q_2}{1 - q_2}} - q_1^{\frac{1}{1 - q_1}}\right), \right.$$

$$\left. \psi_0 \cdot \left(1 - q_2^{\frac{1}{1 - q_2}} - q_3^{\frac{1}{1 - q_3}}\right), \; \psi_0 \cdot q_3^{\frac{q_3}{1 - q_3}}\right)^\mathsf{T}, \quad (30)$$

where ψ_0 is defined in (29), and where it is straightforward to show that $q_2^{\frac{q_2}{1 - q_2}} \geq q_1^{\frac{1}{1 - q_1}}$ for all $q_1, q_2 \in [0, 1]$. Thus, from Proposition 9, the capacity of the channel is

$$\mathsf{C} = \bar{c} = \log\left(1 + \sum_{j=1}^{3}(1 - q_j)q_j^{\frac{q_j}{1 - q_j}}\right). \quad (31)$$

◇

Example 11: The conditional channel PMF for input configuration C'_3 and pooled response R'_3 from Example 2 and 7, respectively, is

$$\phi = \begin{pmatrix} 1 - q_1 & 0 & 0 & 0 \\ q_1(1 - q_2) & 1 - q_2 & 0 & 0 \\ q_1 q_2(1 - q_3) & q_2(1 - q_3) & 1 - q_3 & 0 \\ q_1 q_2 q_3 & q_2 q_3 & q_3 & 1 \end{pmatrix}. \quad (32)$$

As here the fourth column correspond to the zero codeword, we define $\bar{\psi} \triangleq (\psi_1, \psi_2, \psi_3, \psi_0)^\mathsf{T}$ and thus obtain from Proposition 9

$$(1 - q_1)\log\frac{1 - q_1}{\psi_1} + q_1(1 - q_2)\log\frac{q_1(1 - q_2)}{\psi_2}$$
$$+ q_1 q_2(1 - q_3)\log\frac{q_1 q_2(1 - q_3)}{\psi_3} + q_1 q_2 q_3 \log\frac{q_1 q_2 q_3}{\psi_0} = \bar{c} \quad (33a)$$

$$(1 - q_2)\log\frac{1 - q_2}{\psi_2}$$
$$+ q_2(1 - q_3)\log\frac{q_2(1 - q_3)}{\psi_3} + q_2 q_3 \log\frac{q_2 q_3}{\psi_0} = \bar{c} \quad (33b)$$

$$(1 - q_3)\log\frac{1 - q_3}{\psi_3} + q_3 \log\frac{q_3}{\psi_0} = \bar{c} \quad (33c)$$

$$\log\frac{1}{\psi_0} = \bar{c} \quad (33d)$$

By Gaussian elimination steps $(33c) - q_3(33d)$, $(33b) - q_2(33c)$, and $(33a) - q_1(33b)$, we obtain again the same solution as in (28) and (29). ◇

Note that we always end up with the zero codeword separating the system of equations into two halves, with their respective Gaussian elimination as elaborated in the following.

Assume that the configuration C_n has the zero codeword at position ℓ. The corresponding channel transition matrix is

$$\phi = \begin{pmatrix} 1-q_1 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\ q_1(1-q_2) & 1-q_2 & \cdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ q_1 q_2(1-q_3) & q_2(1-q_3) & \cdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & 0 & \vdots & \vdots & \vdots & \vdots & & \vdots \\ q_1\cdots q_{\ell-2}(1-q_{\ell-1}) & q_2\cdots q_{\ell-2}(1-q_{\ell-1}) & \cdots & 1-q_{\ell-1} & 0 & 0 & 0 & \cdots & 0 \\ q_1\cdots q_{\ell-1} & q_2\cdots q_{\ell-1} & \cdots & q_{\ell-1} & 1 & q_\ell & q_\ell q_{\ell+1} & \cdots & q_\ell\cdots q_n \\ 0 & 0 & \cdots & 0 & 0 & 1-q_\ell & (1-q_\ell)q_{\ell+1} & \cdots & (1-q_\ell)q_{\ell+1}\cdots q_n \\ \vdots & \vdots & & \vdots & \vdots & 0 & 1-q_{\ell+1} & \cdots & (1-q_{\ell+1})q_{\ell+2}\cdots q_n \\ \vdots & \vdots & & \vdots & \vdots & \vdots & 0 & & \vdots \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 1-q_n \end{pmatrix} \tag{34}$$

shown in (34) on the top of the page. Note that the column and row containing the value 1 is the ℓ-th column and row.

We now define $\bar{\psi} \triangleq (\psi_1, \dots, \psi_{\ell-1}, \psi_0, \psi_\ell, \dots, \psi_n)^\mathsf{T}$, where we again purposefully use an ordering with ψ_0 being at the ℓ-th position. Applying Proposition 9 for $k = \ell$ (i.e., the ℓ-th column) then results in

$$\log \frac{1}{\psi_0} = \bar{c}. \tag{35}$$

For any $k < \ell$, we use a linear combination of (14) (with equality) for k and for $k+1$ to obtain

$$\mathscr{D}\big(\phi(\cdot|k) \,\big\|\, \bar{\psi}\big) - q_k\, \mathscr{D}\big(\phi(\cdot|k+1) \,\big\|\, \bar{\psi}\big) = \bar{c} - q_k\bar{c}, \tag{36}$$

which combined with (35) can be solved to

$$\frac{\psi_k}{\psi_0} = (1-q_k)q_k^{\frac{q_k}{1-q_k}}. \tag{37}$$

Similarly, for $k > \ell$, we use (14) (with equality) for $k-1$ and k to obtain

$$\mathscr{D}\big(\phi(\cdot|k) \,\big\|\, \bar{\psi}\big) - q_k\, \mathscr{D}\big(\phi(\cdot|k-1) \,\big\|\, \bar{\psi}\big) = \bar{c} - q_k\bar{c}, \tag{38}$$

which results in the same expression (37). Finally, by the fact that $\bar{\psi}$ needs to sum to one, we obtain

$$\psi_0 = \left(1 + \sum_{k=1}^{n}(1-q_k)q_k^{\frac{q_k}{1-q_k}}\right)^{-1}, \tag{39}$$

$$\mathsf{C} = \log\left(1 + \sum_{k=1}^{n}(1-q_k)q_k^{\frac{q_k}{1-q_k}}\right). \tag{40}$$

Note that this capacity value depends only on $q_k = 1 - p_k$ and *not* on the choice of configuration (apart from the fact that the configuration needs to contain the zero codeword). Hence, we have proven the following main result in Theorem 12, for n neurons with given firing probabilities $p_1, \dots, p_n$, $0 < p_k < 1$ and $q_k = 1 - p_k$ for all $k \in [n]$. Note that for the n-fold Z-channel, q_i is the 1-error of the ith Z-channel corrupting $\mathbf{C}_i$ (the ith component of a codeword), and thus we define $\mathbf{q} \triangleq (q_1, \dots q_n)$ as the parameter of the n-fold Z-channel. We also make explicit the dependence of $I(S; \mathbf{Y})$ on $\mathbf{q}$ in Theorem 12.

Theorem 12 (Generalization of Equal Coding Theorem): Let S_n be the symmetric group consisting of all possible permutations over n elements. Then for all $\mathsf{C}_n, \mathsf{C}'_n \in \mathscr{O}_n$ and permutations $\sigma, \sigma' \in \mathsf{S}_n$,

$$\sup_{\pi \in \mathscr{P}_{n+1}} I(S; \mathbf{Y})[\mathsf{C}_n, \pi, \sigma\mathbf{q}] = \sup_{\pi \in \mathscr{P}_{n+1}} I(S; \mathbf{Y})[\mathsf{C}'_n, \pi, \sigma'\mathbf{q}]$$

$$= \log\left(1 + \sum_{k=1}^{n}(1-q_k)q_k^{\frac{q_k}{1-q_k}}\right),$$

where $\sigma\mathbf{q}$ and $\sigma'\mathbf{q}$ denote permutations on the 1-errors $\mathbf{q}$.

IV. DISCUSSION

In this paper, we generalized the Equal Coding Theorem proposed by [8], by considering heterogeneous firing probabilities, see Theorem 12.

A direct extension of this theory is to consider the channel capacity per unit of energy (every spike costs some energy to produce and transmit). This information capacity per energy has been already studied by [8] for the case of homogeneous ON- and OFF-cells, and in this case the optimal fraction of ON-neurons is $1/2$. This is in contrast to the ON- or OFF-cells' dominance found in various different neural systems (see Table 1 in [8]). The generalized Equal Coding Theorem we presented offers new possibilities to explain the said dominance together with other heterogeneous biophysical properties found in ON- versus OFF-neurons (see e.g. [11]).

Recently, [12] generalized the Equal Coding Theorem in a different direction than that in this paper. Instead of considering a simple Heaviside step function for the neural activation, they assumed a sum of step functions such that each neuron has multiple thresholds. The strength of this approach is that in the limit of a large number of small steps, one can approximate any monotonically increasing or decreasing activation function, and in particular the smoothed rectified linear functions found for ON- and OFF-cells [11]. However, the limitation of the approach of [12] is that both the ON- and OFF-neurons are assumed to have the same activation function (up to a sign flip) and therefore this approach is unable to capture the heterogeneity in the activation functions of ON- and OFF-cells. An interesting future direction would be to combine the flexibility of activation functions from [12] with the neuronal heterogeneity proposed in the present paper in order to offer more accurate biological predictions.

REFERENCES

[1] F. Attneave, "Some informational aspects of visual perception." *Psychological Review*, vol. 61, no. 3, pp. 183–193, 1954. [Online]. Available: http://doi.apa.org/getdoi.cfm?doi=10.1037/h0054663

[2] H. B. Barlow, "Possible Principles Underlying the Transformations of Sensory Messages," in *Sensory Communication*. The MIT Press, 1961, pp. 216–234. [Online]. Available: http://mitpress.universitypressscholarship.com/view/10.7551/mitpress/9780262518420.001.0001/upso-9780262518420-chapter-13

[3] S. Laughlin, "A Simple Coding Procedure Enhances a Neuron's Information Capacity," *Zeitschrift für Naturforschung C*, vol. 36, no. 9-10, pp. 910–912, Oct. 1981. [Online]. Available: https://www.degruyter.com/document/doi/10.1515/znc-1981-9-1040/html

[4] Y. Dan, J. J. Atick, and R. C. Reid, "Efficient Coding of Natural Scenes in the Lateral Geniculate Nucleus: Experimental Test of a Computational Theory," *The Journal of Neuroscience*, vol. 16, no. 10, pp. 3351–3362, May 1996. [Online]. Available: https://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.16-10-03351.1996

[5] T. Toyoizumi, J.-P. Pfister, K. Aihara, and W. Gerstner, "Generalized Bienenstock-Cooper-Munro rule for spiking neurons that maximizes information transmission," *Proceedings of the National Academy of Sciences*, vol. 102, no. 14, pp. 5239–5244, Apr. 2005. [Online]. Available: http://www.pnas.org/cgi/doi/10.1073/pnas.0500495102

[6] B. Wark, B. N. Lundstrom, and A. Fairhall, "Sensory adaptation," *Current Opinion in Neurobiology*, vol. 17, no. 4, pp. 423–429, Aug. 2007. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0959438807000840

[7] A. J. Bell and T. J. Sejnowski, "The "independent components" of natural scenes are edge filters," *Vision Research*, vol. 37, no. 23, pp. 3327–3338, Dec. 1997. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0042698997001211

[8] J. Gjorgjieva, M. Meister, and H. Sompolinsky, "Functional diversity among sensory neurons from efficient coding principles," *PLOS Computational Biology*, vol. 15, no. 11, p. e1007476, Nov. 2019. [Online]. Available: https://dx.plos.org/10.1371/journal.pcbi.1007476

[9] S. M. Moser, *Advanced Topics in Information Theory (Lecture Notes)*, 5th ed. Signal and Inf. Proc. Lab., ETH Zurich, Switzerland, and Inst. Commun. Eng., Nat. Yang Ming Chiao Tung Univ., Hsinchu, Taiwan, 2022. [Online]. Available: https://moser-isi.ethz.ch/scripts.html

[10] ——, *Information Theory (Lecture Notes)*, 6th ed. Signal and Inf. Proc. Lab., ETH Zurich, Switzerland, and Inst. Commun. Eng., Nat. Yang Ming Chiao Tung Univ., Hsinchu, Taiwan, 2018. [Online]. Available: https://moser-isi.ethz.ch/scripts.html

[11] E. J. Chichilnisky and R. S. Kalmar, "Functional Asymmetries in ON and OFF Ganglion Cells of Primate Retina," *The Journal of Neuroscience*, vol. 22, no. 7, pp. 2737–2747, Apr. 2002. [Online]. Available: https://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.22-07-02737.2002

[12] S. Shao and J. Gjorgjieva, "Efficient population coding of sensory stimuli," Jul. 2022, arXiv:2207.11712 [q-bio]. [Online]. Available: http://arxiv.org/abs/2207.11712