FINITE-STATE CHANNEL MODELS FOR SIGNAL TRANSDUCTION IN NEURAL SYSTEMS

Andrew W. Eckford[*]

Dept. of EECS
York University
Toronto, Ontario, Canada M3J 1P3

Kenneth A. Loparo[1] *and Peter J. Thomas*[2]

[1]Dept. of EECS and [2]Dept. of Mathematics
Case Western Reserve University
Cleveland, Ohio, USA 44106-7058

ABSTRACT

Information theory provides powerful tools for understanding communication systems. This analysis can be applied to intercellular signal transduction, which is a means of chemical communication among cells and microbes. We discuss how to apply information-theoretic analysis to ligand-receptor systems, which form the signal carrier and receiver in intercellular signal transduction channels. We also discuss the applications of these results to neuroscience.

1. INTRODUCTION

The human brain is a vast communications engine, comprising some 100 billion nerve cells connected by upwards of 100 trillion synapses. Information theory has a long history of application in the biological sciences generally [1] and neuroscience in particular [2]. Capacity and mutual information have proven fruitful concepts in understanding sensory systems [3, 4, 5], fault tolerant computation [6]; computation in spiking neurons [7, 8]; biological computation under metabolic constraints [9, 10, 11], and information processing limitations of genetic regulatory elements [12, 13].

Here we discuss some capacity bounds for several signaling systems present in the brain. A common motif in neurobiological communication is the transduction of chemical, mechanical, or optical signals into ionic currents across the membrane of a nerve cell. Signal transduction typically involves specialized protein molecules: rhodopsin can detect single photon absorptions in the retina [14]; acetylcholine receptor proteins convert the chemical neurotransmitter signal into muscle-activating currents to move the limbs [15]. We focus here on two examples: channelrhodopsin (ChR, widely used as a control mechanism for neuroscience experiments) and the acetylcholine receptor (AChR). Both systems convert their signals into an all-or-none conductance, effectively acting as graded input, binary output systems; both have multiple internal states (three for ChR, five for AChR). Hence the state of the system is only *partially observable*, complicating capacity estimates.

[*]Corresponding author. Email: aeckford@yorku.ca

AWE was funded in part by a grant from the Natural Sciences and Engineering Research Council (NSERC).

PJT was supported by NSF grants EF-1038677 and DMS-0720142.

2. MODEL

2.1. Master equation kinetics

For a receptor with k discrete states, there exists a k-dimensional vector of state occupancy probabilities $p(t)$, given by

$$p(t) = [p_1(t),\, p_2(t),\, \ldots,\, p_k(t)]\,, \tag{1}$$

where $p_i(t)$ represents the probability of a given receptor occupying state i at time t. The chemical kinetics of the receptor are captured by a differential equation known as the *master equation*. Let $Q = [q_{ij}]$ represent a $k \times k$ matrix of rate constants, where q_{ij} represents the instantaneous rate at which receptors starting in state i enter state j. Then the master equation is given by $dp/dt = p(t)Q$.

We use the notation from [16]. In the following examples:

- Rates which are sensitive to the input are directly proportional to the input $x(t)$: for example, q_{12} is the transition rate from state 1 to state 2, which is not sensitive to the input; while $q_{30}x(t)$ is the transition rate from state 3 to state 0, sensitive to the input; and
- The ith diagonal element is written R_i, and is set so that the ith row sums to zero (so, if $x(t)$ appears in the ith row, R_i depends on $x(t)$).

Example 1: Channelrhodopsin-2 (ChR2). The ChR2 receptor is a light-gated ion channel. The receptor has three states, named Closed (C_1), Open (O_2), and Desensitized (C_3). The channel-open (O) state O_2 is the only state in which the ion channel is open, passing an ion current. The channel-closed (C) states, C_1 and C_3, are distinct in that the receptor is light-sensitive in state C_1, and insensitive in state C_2 [17]. The rate matrix for ChR2 is

$$Q = \begin{bmatrix} R_1 & q_{12}x(t) & 0 \\ 0 & R_2 & q_{23} \\ q_{31} & 0 & R_3 \end{bmatrix}. \tag{2}$$

Fig. 1 shows state labels and allowed state transitions. Param-

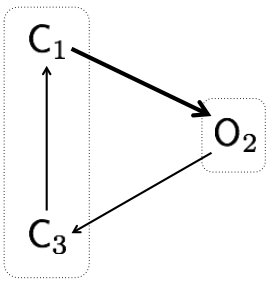

Fig. 1. Depiction of allowed state transitions for ChR2. Sensitive transitions are depicted with **bold** arrows. States are labelled by their ion channel state: $\{\mathsf{C}, \mathsf{O}\}$ for closed and open, respectively; state number is in subscript. Dashed lines surround all states in either the closed or open state. Transition rates, listed in Table 1, correspond to the vertices associated with each directed edge: for example, the rate from state O_2 to state C_3 is q_{23}.

eter values from the literature are given in Table 1. We will assume, as in [18], that the opening rate $q_{12}x(t)$ is directly proportional to the irradiance of the light on the receptor.

Parameter	from [17]	Units
$q_{12}x(t)$	$\leq 5 \times 10^3$	s^{-1}
q_{23}	50	s^{-1}
q_{31}	17	s^{-1}

Table 1. Rate parameters for ChR2, adapted from [17].

Example 2: Acetylcholine (ACh). The ACh receptor is a ligand-gated ion channel. The receptor has five states, with rate matrix

$$Q = \begin{bmatrix} R_1 & q_{12}x(t) & 0 & q_{14} & 0 \\ q_{21} & R_2 & q_{23} & 0 & 0 \\ 0 & q_{32} & R_3 & q_{34} & 0 \\ q_{41} & 0 & q_{43}x(t) & R_4 & q_{45} \\ 0 & 0 & 0 & q_{54}x(t) & R_5 \end{bmatrix}. \quad (3)$$

There are three sensitive transitions: $r_{12}x(t)$, $r_{43}x(t)$, and $r_{54}x(t)$, which are proportional to agonist concentration $x(t)$. These transitions represent binding of an ACh molecule to one of two binding sites. Fig. 2 shows the allowed state transitions. State C_5 corresponds to both sites unoccupied; states C_4, O_1 correspond to one site occupied; states C_3, O_2 correspond to both sites occupied. Table 2 gives parameter values; the concentration of ACh, $x(t)$, is measured in mol/ℓ.

The same state-naming convention is used in the figure as with ChR2: states with an open ion channel are O_1 and O_2; states with a closed ion channel are C_3, C_4, and C_5.

For each of the preceding examples, the rate constants depend on environmental conditions, and thus can be reported differently in different sources (e.g., [20] for ChR2).

Parameter	Name in [19]	Value/range	Units
$q_{12}x(t)$	$k_{+2}x$	$5 \times 10^8 x(t)$	s^{-1}
q_{14}	α_1	3×10^3	s^{-1}
q_{21}	$2k^*_{-2}$	0.66	s^{-1}
q_{23}	α_2	5×10^2	s^{-1}
q_{32}	β_2	1.5×10^4	s^{-1}
q_{34}	$2k_{-2}$	4×10^3	s^{-1}
q_{41}	β_1	15	s^{-1}
$q_{43}x(t)$	$k_{+2}x$	$(5 \times 10^8)x(t)$	s^{-1}
q_{45}	k_{-1}	2×10^3	s^{-1}
$q_{54}x(t)$	$2k_{+1}x$	$(1 \times 10^8)x(t)$	s^{-1}

Table 2. Rate parameters for ACh, adapted from [19], where $x(t)$ represents the molar concentration of ACh in mol/ℓ.

2.2. From the master equation to discrete-time Markov chains

It is possible to discretize the master equation and describe the dynamics of a receptor as a discrete-time Markov chain; this is important to our paper as we rely on capacity results for discrete-time Markov channels. Briefly, we can discretize the master equation by writing

$$p(t + \Delta t) = p(t)\,(I + \Delta t Q) + o(\Delta t) \quad (4)$$

where I is the identity matrix, and $o(\Delta t)/\Delta t \to 0$ as $\Delta t \to 0$. If we let

$$p[j] = p(j\Delta t), \quad (5)$$

then this equation becomes

$$p[j+1] = p[j](I + \Delta t Q). \quad (6)$$

Thus, we have a discrete-time Markov chain with transition probability matrix

$$P = I + \Delta t Q. \quad (7)$$

The matrix P satisfies the conditions of a Markov chain transition probability matrix (nonnegative, row-stochastic) as long as Δt is small enough. Note that while the probability $p(t)$ evolves deterministically, the channel state itself is a non-Gaussian random process taking discrete values.

3. SIGNAL TRANSDUCTION AS A COMMUNICATIONS SYSTEM

3.1. Communication model of receptors

We now discuss how the receptors can be described as information-theoretic communication systems: that is, in terms of input, output, and conditional input-output PMF.

Input: The receptor input $x(t)$ consist of either light intensities or ligand concentrations, and is discretized in time: for integers i, the input is $x(i\Delta t)$; we will write $x_i = x(i\Delta t)$.

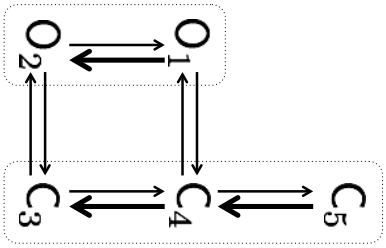

Fig. 2. Depiction of allowed state transitions for ACh. Sensitive transitions are depicted with **bold** arrows. States are labelled by their ion channel state: $\{\mathsf{C}, \mathsf{O}\}$ for closed and open, respectively; state number is in subscript. Dashed lines surround all states in either the closed or open state. Transition rates, listed in Table 2, correspond to the vertices associated with each directed edge: for example, the rate from state O_2 to state C_3 is q_{23}.

We will also discretize the amplitude, so that for every t, $x_i \in \{\mathsf{x}_1, \mathsf{x}_2, \mathsf{x}_3, \ldots, \mathsf{x}_k\} =: \mathcal{X}$. We will assume that the x_i are distinct and increasing; further, we assign the lowest value x_1 and the highest value x_k the symbols x_L and x_H, respectively.

Output: Receptor states from our example systems are labelled, e.g., C_3 or O_2. The output of the communication system is given either by: the receptor state $y(t)$, given by the *subscript* of the state label; or by the ion channel state $z(t)$, either C or O, without subscript. These are discretized, respectively, to $y_i = y(i\Delta t)$ and $z_i = z(i\Delta t)$.

Conditional input-output PMF: From (4)-(7), y^n forms a Markov chain given x^n, so

$$p_{Y^n|X^n}(y^n|x^n) = \prod_{i=1}^{n} p_{Y_i \mid Y_{i-1}, X_i}(y_i \mid y_{i-1}, x_i), \quad (8)$$

where $p_{Y_i \mid Y_{i-1}, X_i}(y_i \mid y_{i-1}, x_i)$ is given by the appropriate entry in the matrix P, and where y_0 is null.[1] For example, using ACh, suppose $y_{i-1} = 1$, $y_i = 2$, and $x_i = \mathsf{x}_\mathsf{H}$. Then from (7) and Table 2, we have $p_{Y_i \mid Y_{i-1}, X_i}(2 \,|\, 1, \mathsf{x}_\mathsf{H}) = \Delta t q_{12}(t) = 5 \times 10^8 \mathsf{x}_\mathsf{H} \Delta t$.

3.2. Information theory and Shannon capacity

We briefly review the information-theoretic concepts used in the paper. The reader is directed to [21] for further detail.

A communication channel consists of: a vector of inputs $[x_1, x_2, \ldots]$, a vector of outputs $[y_1, y_2, \ldots]$, and a conditional probability density function relating outputs to inputs. Using the following vector notation:

$$x^n = [x_1, x_2, \ldots, x_n] \quad (9)$$

$$y^n = [y_1, y_2, \ldots, y_n], \quad (10)$$

[1] We say a variable is "null" if it vanishes under conditioning, i.e., if y_0 is null, then $p_{Y_1|X_1,Y_0}(y_1 \mid x_1, y_0) = p_{Y_1|X_1}(y_1 \mid x_1)$.

the stochastic input-output relationship is given by the conditional joint PMF $p_{Y^n|X^n}(y^n|x^n)$.

For a channel with inputs x^n and outputs y^n, the mutual information $I(X^n; Y^n)$ gives the maximum information rate that may be transmitted reliably over the channel. Mutual information is given by

$$I(X^n; Y^n) = E\left[\log \frac{p_{Y^n|X^n}(y^n \mid x^n)}{p_{Y^n}(y^n)}\right] \quad (11)$$

(12)

where $p_{Y^n|X^n}(y^n \mid x^n)$ is the conditional probability mass function (PMF) of Y^n given X^n, and Y, and $p_Y(y)$ is the marginal PMF on Y.

As $n \to \infty$, generally $I(X^n; Y^n) \to \infty$ as well; in this case, it is useful to calculate the mutual information rate, given by

$$\mathcal{I}(X; Y) = \lim_{n\to\infty} \frac{1}{n} \sum_{x^n, y^n} p_{X^n, Y^n}(x^n, y^n) \log \frac{p_{Y^n|X^n}(y^n \mid x^n)}{p_{Y^n}(y^n)}. \quad (13)$$

We will assume that receptor response is stationary. Similar derivations hold for $\mathcal{I}(X; Z)$, the mutual information rate from inputs to ion channel state.

3.3. Receptor IID Capacity

The capacity C of a communication system is the maximum over all possible input distributions $p_{X^n}(x^n)$ of $\mathcal{I}(X; Y)$. If the inputs $p_{X^n}(x^n)$ is restricted to the set of independent, identically distributed (IID) input distributions, i.e. we can write $p_{X^n}(x^n) = \prod_{i=1}^n p(x_i)$, then we have the IID capacity, written C_{iid}. It should be clear that $C_{\text{iid}} < C$.

We now calculate C_{iid} for the discrete-time receptor model. Although IID inputs are not realistic in practice (as concentration may persist for long periods of time), they can be capacity-achieving under some circumstances [22].

In general, since Y^n is a time-inhomogeneous Markov chain if X^n is known, we can write

$$p_{Y^n|X^n}(y^n \mid x^n) = \prod_{i=1}^{n} p_{Y_i|X_i,Y_{i-1}}(y_i \mid x_i, y_{i-1}), \quad (14)$$

Under IID inputs, it can be shown that the receptor states Y^n form a time-homogeneous Markov chain, that is,

$$p_{Y^n}(y^n) = \prod_{i=1}^{n} p_{Y_i|Y_{i-1}}(y_i \mid y_{i-1}), \quad (15)$$

where y_0 is again null, and where

$$p_{Y_i|Y_{i-1}}(y_i \mid y_{i-1}) = \sum_x p_{Y_i|X_i,Y_{i-1}}(y_i \mid x, y_{i-1}) p_X(x). \quad (16)$$

Using (14)-(15), (11) reduces to

$$I(X^n;Y^n) = \sum_{i=1}^{n} E\left[\log \frac{p_{Y_i \mid X_i,Y_{i-1}}(y_i \mid x_i, y_{i-1})}{p_{Y_i|Y_{i-1}}(y_i \mid y_{i-1})}\right] \quad (17)$$

and (13) reduces to

$$\mathcal{I}(X;Y) = E\left[\log \frac{p_{Y_i \mid X_i,Y_{i-1}}(y_i \mid x_i, y_{i-1})}{p_{Y_i|Y_{i-1}}(y_i \mid y_{i-1})}\right] \quad (18)$$

Considering the diagrams in the previous section, some of the transitions were sensitive (i.e., dependent on input x_i), and others were insensitive (i.e., independent of x_i). From (18), if the transition $p_{Y_i \mid X_i,Y_{i-1}}(y_i \mid x_i, y_{i-1})$ is insensitive, then

$$\log \frac{p_{Y_i|X_i,Y_{i-1}}(y_i \mid x_i, y_{i-1})}{p_{Y_i|Y_{i-1}}(y_i \mid y_{i-1})} = \log \frac{p_{Y_i|Y_{i-1}}(y_i \mid y_{i-1})}{p_{Y_i|Y_{i-1}}(y_i \mid y_{i-1})} \quad (19)$$

$$= \log 1 = 0. \quad (20)$$

Thus, (18) is calculated using the *sensitive transitions only*.

Let $\mathcal{S} = \mathcal{Y} \times \mathcal{Y}$ represent the set of sensitive transitions, i.e., $(y_{i-1}, y_i) \in \mathcal{S}$ if $p_{Y_i|X_i,Y_{i-1}}(y_i \mid x_i, y_{i-1})$ is a function of x_i. Moreover define

$$\phi(p) = \begin{cases} 0, & p = 0 \\ p \log p, & p \neq 0. \end{cases} \quad (21)$$

Then

$$\mathcal{I}(X;Y) = \quad (22)$$
$$\sum_{x \in \mathcal{X}} p_X(x) \sum_{(y_{i-1},y_i)\in\mathcal{S}} \pi_{y_{i-1}} \phi(p_{Y_i|X_i,Y_{i-1}}(y_i \mid x_i, y_{i-1}))$$
$$- \sum_{(y_{i-1},y_i)\in\mathcal{S}} \pi_{y_{i-1}} \phi\left(\sum_{x\in\mathcal{X}} p_X(x) p_{Y_i|X_i,Y_{i-1}}(y_i \mid x_i, y_{i-1})\right).$$

Using (7), the matrix Q for the desired receptor, and an appropriately selected Δt, we can calculate $\mathcal{I}(X;Y)$.

Since Z^n is a hidden Markov process, calculating the IID capacity $\mathcal{I}(X;Z)$ from inputs to ion channel state is done in one of two ways: either using Monte Carlo techniques to evaluate the expectation in (11), replacing y^n with z^n; or finding upper and lower bounds, generalizing the technique from [23]. In either case, the probability of the hidden Markov process z^n is obtained using the sum-product algorithm [24]. By the data processing inequality. $\mathcal{I}(X;Y) \geq \mathcal{I}(X;Z)$.

4. RESULTS

Mutual information results are given in Figure 3. The IID capacity may be found by taking the maximum of each curve. We see that when sensitive transitions are directly observable, there is a small gap between $\mathcal{I}(X;Y)$ and $\mathcal{I}(X;Z)$

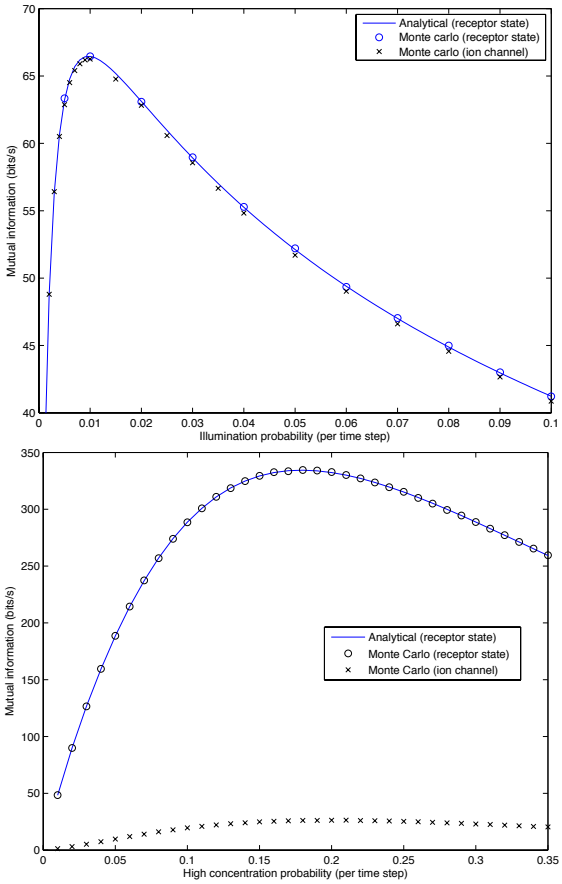

Fig. 3. Mutual information with IID inputs for ChR2 (top figure) and ACh (bottom figure). Results for *receptor state* refer to $\mathcal{I}(X;Y)$, mutual information from input to the state of the receptor; results for *ion channel* refer to $\mathcal{I}(X;Z)$, mutual information from input to the state of the ion channel.

(*cf.* ChR2); and when sensitive transitions are *not* directly observable, there is a large gap (*cf.* ACh); thus, the receptor capacity is not always a tight bound for $\mathcal{I}(X;Z)$. Heuristically, this gap appears to occur because of the structure of the channel. The sensitive transition for ChR has stoichiometry $\mathbf{v}_{\text{ctrl}} = [-1, 1, 0]$ and the observation vector is $\mathbf{g} = [0, 1, 0]$; their inner product $\mathbf{v}_{\text{crtl}} \cdot \mathbf{g}^{\mathrm{T}} = 1$. In contrast, the three sensitive transitions for ACh have stoichiometries $\mathbf{v}^1_{\text{ctrl}} = [-1, 1, 0, 0, 0], \mathbf{v}^2_{\text{ctrl}} = [0, 0, 1, -1, 0], \mathbf{v}^3_{\text{ctrl}} = [0, 0, 0, 1, -1]$ respectively; the observation vector is $\mathbf{g} = [1, 1, 0, 0, 0]$, and $\mathbf{v}^i_{\text{crtl}} \cdot \mathbf{g}^{\mathrm{T}} = 0$ for each i.

Ideally, information theoretic analysis would lead to predictions comparable with experimental data. However, receptor binding is part of a multistage channel that includes secretion and diffusion. Channel rhodopsin is part of a multistage channel too: light-triggered currents can promote or inhibit action potentials depending on the type of ion coupled to the channel. Both channels involve nonlinearities and memory

effects that call for additional analysis.

5. REFERENCES

[1] H. P. Yockey, R. P. Platzman, and H. Quastler, Eds., *Symposium on Information Theory in Biology*, Pergamon Press, New York, London, 1958.

[2] D. M. MacKay and W. S. McCulloch, "The limiting information capacity of a neuronal link," *Bulletin of Mathematical Biophysics*, vol. 4, pp. 127–155, 1952.

[3] Horace B. Barlow, *Sensory Communication*, chapter 13: Possible principles underlying the transformations of sensory messages, pp. 217–234, MIT Press, 1961.

[4] Anthony J Bell and Terrence J Sejnowski, "The independent components of natural scenes are edge filters," *Vision research*, vol. 37, no. 23, pp. 3327–3338, 1997.

[5] Bruno A Olshausen and David J Field, "Sparse coding of sensory inputs," *Current opinion in neurobiology*, vol. 14, no. 4, pp. 481–487, 2004.

[6] Jack D. Cowan and Schmuel Winograd, *Reliable Computation in the Presence of Noise*, MIT Press, 1963.

[7] R R de Ruyter van Steveninck, G D Lewen, S P Strong, R Koberle, and W Bialek, "Reproducibility and variability in neural spike trains," *Science*, vol. 275, no. 5307, pp. 1805–8, Mar 1997.

[8] J Vincent Toups, Jean-Marc Fellous, Peter J Thomas, Terrence J Sejnowski, and Paul H Tiesinga, "Multiple spike time patterns occur at bifurcation points of membrane potential dynamics," *PLoS Comput Biol*, vol. 8, no. 10, 2012.

[9] Simon B Laughlin, "Energy as a constraint on the coding and processing of sensory information," *Current opinion in neurobiology*, vol. 11, no. 4, pp. 475–480, 2001.

[10] Toby Berger and William B Levy, "A mathematical theory of energy efficient neural computation and communication," *Information Theory, IEEE Transactions on*, vol. 56, no. 2, pp. 852–874, 2010.

[11] Jie Xing, Toby Berger, and Terrence J Sejnowski, "A Berger-Levy energy efficient neuron model with unequal synaptic weights," in *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on*. IEEE, 2012, pp. 2964–2968.

[12] Gašper Tkačik, Curtis G Callan Jr, and William Bialek, "Information capacity of genetic regulatory elements," *Physical Review E*, vol. 78, no. 1, pp. 011910, 2008.

[13] Gašper Tkačik, Curtis G Callan, and William Bialek, "Information flow and optimization in transcriptional regulation," *Proceedings of the National Academy of Sciences*, vol. 105, no. 34, pp. 12265–12270, 2008.

[14] F Rieke and DA Baylor, "Single-photon detection by rod cells of the retina," *Reviews of Modern Physics*, vol. 70, no. 3, pp. 1027, 1998.

[15] Douglas M Fambrough, Daniel B Drachman, and S Satyamurti, "Neuromuscular junction in myasthenia gravis: decreased acetylcholine receptors," *Science*, vol. 182, no. 4109, pp. 293–295, 1973.

[16] Jeffrey R Groff, Hilary DeRemigio, and Gregory D Smith, *Stochastic Methods in Neuroscience*, chapter 2. Markov chain models of ion channels and calcium release sites, pp. 29–64, OUP Oxford, 2009.

[17] G. Nagel, T. Szellas, W. Huhn, S. Kateriya, N. Adeishvili, P. Berthold, D. Ollig, P. Hegemann, and E. Bamberg, "Channelrhodopsin-2, a directly light-gated cation-selective membrane channel," *PNAS*, vol. 100, no. 24, pp. 13940–13945, 2003.

[18] T. Tchumatchenko, J. P. Newman, M.-F. Fong, and S. M. Potter, "Delivery of continuously-varying stimuli using channelrhodopsin-2," *Frontiers in Neural Circuits*, vol. 7, pp. 1–9, 2013.

[19] D. Colquhoun and A. G. Hawkes, "On the stochastic properties of bursts of single ion channel openings and of clusters of bursts," Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, vol. 300, no. 98, pp. 1–59, 1982.

[20] J. Y. Lin, M. Z. Lin, P. Steinbach, and R. Y. Tsien, "Characterization of engineered channelrhodopsin variants with improved properties and kinetics," *Biophysical J.*, vol. 96, pp. 1803–1814, 2009.

[21] Thomas M. Cover and Joy A. Thomas, *Elements of Information Theory (2nd ed.)*, Wiley, 2006.

[22] A. W. Eckford and P. J. Thomas, "Capacity of a simple intercellular signal transduction channel," in *IEEE International Symposium on Information Theory*, 2013.

[23] Peter J. Thomas and Andrew W. Eckford, "Capacity of a simple intercellular signal transduction channel," 2015, http://arxiv.org/abs/1411.1650.

[24] F.R. Kschischang, B.J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *Information Theory, IEEE Transactions on*, vol. 47, no. 2, pp. 498–519, 2001.