An Additive Exponential Noise Channel with a Transmission Deadline

YiLin Tsai, *Student Member, IEEE,* Christopher Rose, *Fellow, IEEE, Ruochen Song, Student Member, IEEE*
and I. Saira Mian

Abstract—We derive the maximum mutual information for an additive exponential noise (AEN) channel with a peak input constraint. We find that the optimizing input density is mixed (with singularities) similar to previous results for AEN channels with a mean input constraint. Likewise, the maximum mutual information takes a similar form, though obviously the maximum for the peak constraint is smaller than for the corresponding mean-constrained channel. This model is inspired by multiple biological phenomena and processes which can be abstracted as follows: inscribed matter is sent by an emitter, moves through a medium, and arrives eventually at its destination receptor. The inscribed matter can convey information in a variety of ways such as the number of signaling quanta – molecules, macromolecular complexes, organelles, cells and tissues – that are emitted as well as the detailed pattern of their release. However, rather than focus on a general class of emitter-receptor systems or a particular exemplar of biomedical importance, our ultimate goal is to provide bounds on the potential efficacy of timed-release signaling for *any* system which emits identical signaling quanta. That is, we seek to apply one of the most potent aspects of information theory to biological signaling – *mechanism blindness* – in the hopes of gaining insights applicable to diverse systems that span a wide range of spatiotemporal scales.

Index Terms—Exponential channel, bits through queues, diffusion channel capacity, molecular signaling

I. INTRODUCTION

BIOLOGICAL systems are networks of intercommunicating elements at whatever level one cares to consider – (macro)molecules, cells, tissues, organisms, populations, microbiomes, ecosystems, and so on. It is no wonder therefore that communication theorists have plied their trade heavily in this scientific domain (for a recent review, see [1]). Biological systems offer a dizzying array of processes and phenomena through which the same and different tasks, communication or otherwise, might be accomplished (see, for example, [2]–[7]). Identifying the underlying mechanisms (signaling modality, signaling agent, signal transport, and so on) as well as the molecules and structures implementing the mechanisms is no small undertaking. Consequently, experimental biologists use a combination of prior knowledge and what can only be called instinct to choose those systems on which to expend effort. Guidance may be sought from evolutionary developmental biology – a field that compares the developmental processes of different organisms to determine their ancestral relationship

and to discover how developmental processes evolved. Insights may be gained by using statistical machine learning techniques to analyze heterogeneous data such as the biomedical literature and the output of so-called "omics" technologies – genomics (genes, regulatory, and non-coding sequences), transcriptomics (RNA and gene expression), proteomics (protein expression), metabolomics (metabolites and metabolic networks), pharmacogenomics (how genetics affects hosts' responses to drugs), and physiomics (physiological dynamics and functions of whole organisms).

Typically, the application of communication theory to biology starts by selecting a candidate system whose components and operations have been already elucidated to varying degrees using methods in the experimental and/or computational biology toolbox [8], [9] and then applying communication theoretic methods [1], [7], [10]–[12]. However, we believe that communication theory in general and information theory in particular is not merely a system analysis tool for biology. That is, given energy constraints and some general physics of the problem, an information-theoretic treatment can be used to provide outer bounds on information transfer in a *mechanism-blind* manner. Thus, rather than simply elucidating and quantifying known biology, communication theory can winnow the plethora of possibilities (or even suggest new ones) amenable to experimental and computational pursuit. Likewise, general application of communication-theoretic principles to biology affords a new set of application areas for communication theorists. Some aspects of the potential for communication theory as a new lens on biological systems is explored in [13].

In this light, here we devise an abstraction that encompasses a myriad of biological processes and phenomena, utilize it to devise a simpler model suitable for communication-theoretic investigations, and analyze the resultant model using ideas discussed in unrelated work, namely the capacity of timing channels [14]. Numerous scenarios in biology that involve the transmission of information can be synthesized and summarized as inscribed matter is sent by an emitter, moves through a medium, and arrives eventually at its destination receptor where it is interpreted. Scenarios illustrating the complexity and diversity that our abstraction attempts to capture include the following. messenger RNA molecules (mRNAs) that are transcribed from the genome migrate from the nu-

cleus to the cytoplasm where they are translated by the ribosome into proteins. Molecules of the neurotransmitter acetylcholine (Ach) that are released by the presynaptic neuron terminal diffuse through the synaptic cleft and bind to nicotinic Ach receptors on the motor end plate. Ions, molecules, organelles, bacteria and viruses that are present in one cell are shipped through a thin membrane channel (tunneling nanotube) to the connected cell where they elicit a physiological response. Membrane-bound vesicles that contain a variety of materials and substances translocate through the cytoplasm to the cell membrane where release their contents into the extracellular environment. Malignant cells that have escaped the confines of a tissue circulate through the bloodstream to other sites where they re-penetrate the vessel walls and can seed a new tumor. Chemicals factors that are secreted or excreted by an individual travel outside the body where they are sensed by a member of the same species triggering a social or behavioral response.

Although the abstraction accommodates a wide range of spatiotemporal scale and types of emitters, inscribed matter, and receptors, it neglects many biologically important features. For example, the suite of signaling quanta – molecules, macromolecular complexes, organelles, cells, and so on – that are released is not necessarily the same as that which reaches the target because some may be changed (eukaryotic mRNAs are modified post-transcriptionally), some may be removed (Ach can be degraded by the enzyme Ach-esterase), some never arrives (the random path produced by diffusion may result in a trajectory that leads away from the target), and so on. The movement of inscribed matter may be passive or active, may or may not require energy and so on.

Despite its limitations, the abstraction does embody a number of salient features. Typically, information is thought to be conveyed via numbers of signaling quanta (concentration). Thus, what amount to dose-response curves are the norm for a variety of experimental biology studies [7]. However, as was shown in an entirely different domain and unrelated work [14], timing of emissions could *in principle* also convey information. Clearly, this possibility cannot be ignored if our aim is to attempt to provide bounds on what "a cell can tell the world." Under certain conditions, perhaps timing is a useful complement to concentration. Alternatively, timing might sometimes be energetically unfavorable and its use unlikely. In either case, information-theoretic bounds would help guide biological inquiry.

Our emitter-receptor system is motivated also by fundamental "systems" problems in biology such as development, wherein undifferentiated cells are "told" what to become by a combination of internal programming and extracellular milieu signals – and in turn tell other cells what to become [15]. Thus, communication within and between cells plays a vital role in the development (embryogenesis), maintenance (tissue homeostasis), subversion

(disorders such as cancer, inflammation, infections) and decline (aging) of multicellular forms and systems.

Unfortunately, the detailed physics of even this seemingly simple abstraction are fraught with a variety of complications. As indicated above, free-space diffusive first passage times are generally not at all well-behaved, often being heavy-tailed to the point that sometimes some of the inscribed matter may *never* arrive at the receptor site. If we neglect this aspect, then for cells emitting signaling quanta into constrained media, arrival with finite mean first passage time is a reasonable approximation. So, similar to [14], we take exponential first passage with mean passage time $1/\lambda$ and use the results of [14] to provide bounds on any finite-mean first passage time density. Simplifying the emitter-receptor system even further, we arrive finally at a communication theoretic-problem involving an Additive Exponential Noise (AEN) channel with a peak input constraint.

In scenarios where the emitter and receptor are in close proximity, the emitted signaling quanta may impinge on a specific unique receptor. However, this is not necessarily always the case. Thus, since signaling quanta of the same type are indistinguishable, which emission corresponds to which arrival is unknown in general. Therefore, if emissions occur at times $\{X_i\}$ and the corresponding arrivals occur at $\{Y_i\}$, then all the receiver has available is $\{\vec{Y}_i\}$, the time-ordered version of the arrivals.

This problem, the main subject of ongoing work, is one we completely ignore here. However, since $\mathbf{X} \to \mathbf{Y} \to \vec{\mathbf{Y}}$ forms a Markov chain (with the last step being a deterministic ordering map), we know that $I(\mathbf{X};\mathbf{Y}) \geq I(\mathbf{X};\vec{\mathbf{Y}})$ by the data processing theorem [16]. Thus, by calculating the timing information that a single signaling agent can convey on average we can provide an upper bound for the ensemble indistinguishable signaling quanta problem.

Our main results: for an emission interval constrained to $[0, T]$, the emission time density that maximizes $I(X;Y)$ is mixed – uniform on $[0, T]$ with singularities at $t = 0$ and $t = T$. For an average transit time of $\frac{1}{\lambda}$, the maximum mutual information between X and Y is $\log\left(1 + \frac{\lambda T}{e}\right)$.

In section II, we review the relevant model and main results from previous work [14]. In section III, we present our model for which we find the optimizing emission time density in section IV. We discuss our results in section V with an eye toward making predictions about emitter-receptor systems in special cases.

II. AEN WITH A MEAN CONSTRAINT

For an AEN channel with mean constraint we have,

$$Y = X + Z$$

when $X \geq 0$ with a mean constraint $E[X] = \alpha$ and Z is exponential with mean $1/\lambda$. The optimal input distribution is

$$f_X(x) = \frac{1}{1 + \lambda\tau}\delta(x) + \frac{\lambda^2\tau}{(1 + \lambda\tau)^2}e^{\frac{-\lambda x}{1 + \lambda\tau}}.$$

and the capacity per channel use is

$$C = \log(1 + \tau\lambda).$$

The capacity of this channel was found by [14] utilizing the fact that exponential densities maximize the differential entropy over all non-negative random variable distributions [16]. Recently, [17] derived the same result by finding the mutual information saddle points.

For exponential Z we have $h(Z) \geq h(W)$ for all non-negative random variables W with $E(W) \leq E(Z) = \tau$. Thus, the maximum mutual information over the AEN channel, $I^e(X;Y)$, is the smallest maximum mutual information for any such channel. The AEN therefore provides a minmax bound on capacity over mean-constrained channels with non-negative inputs as proven in [14].

III. AEN WITH A DEADLINE

For the AEN channel with a deadline we have

$$Y = X + Z$$

where again Z is an independent exponential random variable with mean $1/\lambda$. X is again non-negative but is now constrained to lie in $[0, T]$ with no explicit mean constraint. We seek to maximize $I(X;Y)$. Since

$$I(X;Y) = h(Y) - h(Y|X) = h(Y) - (1 - \log\lambda)$$

we seek to maximize $h(Y)$ with respect to the input density $f_X()$.

However, unlike the mean-constrained case considered in [14], it is impossible to produce an exponential Y with mean $\frac{1}{\lambda} + E[X]$ from X constrained to the finite interval $[0, T]$ as will become apparent in the following development.

Since X is independent of Z we have

$$f_Y(y) = \int_0^y f_X(x) f_Z(y - x) dx \quad 0 \leq y$$

and because X is constrained to $[0, T]$, we can divide $f_Y(y)$ into two regions: region I where $y \in [0, T]$ and region II where $y \in (T, \infty)$. We then have

$$f_Y(y) = \begin{cases} \sigma f_{Y|I}(y) & 0 \leq y \leq T \\ (1 - \sigma) f_{Y|II}(y) & y \geq T \end{cases} \quad (1)$$

where

$$\sigma = \int_0^T f_Y(y) dy$$

with

$$\sigma f_{Y|I}(y) = \int_0^y f_X(x) f_Z(y - x) dx \quad (2)$$

and

$$(1 - \sigma) f_{Y|II}(y) = \int_0^T f_X(x) f_Z(y - x) dx \quad (3)$$

For Z exponential with parameter λ we have

$$f_Y(y) = \int_0^y f_X(x) \lambda e^{-\lambda(y-x)} dx \quad 0 \leq y \quad (4)$$

and

$$\sigma f_{Y|I}(y) = \int_0^y f_X(x) \lambda e^{-\lambda(y-x)} dx \quad (5)$$

and

$$(1 - \sigma) f_{Y|II}(y) = e^{-\lambda y} \int_0^T f_X(x) \lambda e^{\lambda x} dx \quad (6)$$

The entropy of Y is then

$$\begin{aligned} h(Y) &= -\int_0^\infty f_Y(y) \log f_Y(y) dy \\ &= -\int_0^T \sigma f_{Y|I}(y) \log\left(\sigma f_{Y|I}(y)\right) dy \\ &\quad - \int_T^\infty (1 - \sigma) f_{Y|II}(y) \log\left((1 - \sigma) f_{Y|II}(y)\right) dy \\ &= \sigma h(Y|I) + (1 - \sigma) h(Y|II) + H_B(\sigma) \end{aligned} \quad (7)$$

where $H_B()$ is the binary entropy function. Notice that no particular care has to be taken with the integrals at $y = T$ because $f_Y(y)$ cannot contain singularities – it is obtained by the convolution of two densities, one of which, $f_Z()$, contains no singularities.

IV. MAXIMIZATION OF $h(Y)$

The most interesting aspect of equation (6) is that the *shape* of the conditional density for $y > T$ is completely determined – an exponential with parameter λ as depicted in Fig 1. Thus, selection of $f_X()$ does not affect $f_{Y|II}()$

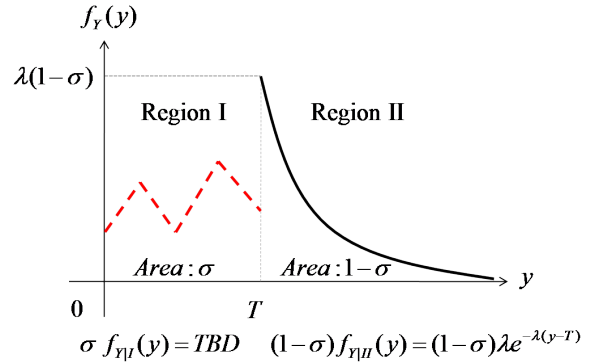Fig. 1: The shapes associated with $f_Y(y)$: We assume arbitrary shape in region I and the requisite exponential shape in region II.

and we must have $h(Y|II) = 1 - \log\lambda$.

This observation suggests a three-step approach to maximizing $h(Y)$. In the first two steps, we completely ignore $f_X()$ and find the *shape* $f_{Y|I}()$ and value of σ which maximize equation (7). In step three, we determine that there indeed exists a density $f_X()$ which produces the optimizing $f_Y()$.

Step 1: For fixed σ we see from equation (7) that $h(Y)$ is maximized solely by our choice of $f_{Y|I}()$. The uniform
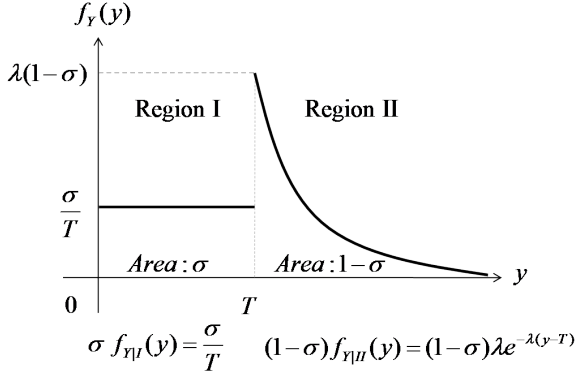
Fig. 2: The updated shape of $f_Y(y)$ after step 1: $f_{Y|I}(y)$ is chosen as $\frac{1}{T}$.

Fig. 3: The optimizing $f_Y(y)$.

density maximizes entropy on a finite interval [16]. Thus, $f_{Y|I}(y) = \frac{1}{T}$ and $h(Y|I) = \log T$ as depicted in Fig 2.

Step 2: Since for any σ, $h(Y|I) = \log T$, we have

$$h(Y) = \sigma \log T + (1-\sigma)(1-\log \lambda) + H_B(\sigma) \qquad (8)$$

Taking derivatives with respect to σ yields

$$\frac{d}{d\sigma}h(Y) = \log T - (1-\log \lambda) - (1+\log \sigma) + (1+\log(1-\sigma))$$

which we set to zero to obtain

$$\log \lambda T - \log \frac{\sigma}{1-\sigma} - 1 = 0$$

We rearrange to obtain

$$\lambda T = \frac{e\sigma}{1-\sigma}$$

from which we deduce that the optimal σ is

$$\sigma^* = \frac{\lambda T}{e + \lambda T} \qquad (9)$$

Returning to the optimization we have

$$\max_{f_X} h(Y) \leq \sigma^* \log T + (1-\sigma*)(1-\log \lambda) + H_B(\sigma^*)$$

which through substitution of σ^* according to equation (9) yields

$$\max_{f_X} h(Y) \leq \log \left(\frac{e + \lambda T}{\lambda} \right) \qquad (10)$$

with equality when

$$f_Y(y) = \begin{cases} \frac{\lambda}{e + \lambda T} & y \in [0, T] \\ \frac{e}{e + \lambda T} \lambda e^{-\lambda(y-T)} & y > T \end{cases} \qquad (11)$$

as depicted in Fig 3.

Step 3: All that remains is to ascertain whether $\exists f_X()$ which can generate the optimizing $f_Y(y)$. Since $f_Y()$ is the convolution of $f_Z()$ and $f_X()$ we can use Fourier transforms to obtain a candidate solution for $f_X()$. That

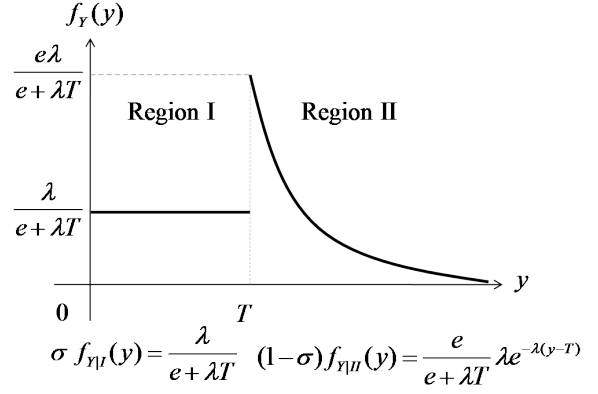is, the Fourier transform of $f_Z()$ is $\frac{\lambda}{\lambda + j2\pi f}$ so the Fourier transform of $f_X()$ is

$$\mathcal{F}\{f_X()\} = \mathcal{F}\{f_Y()\} \left(\frac{j2\pi f}{\lambda} + 1 \right)$$

Multiplication by $j2\pi f$ implies differentiation so we must have

$$f_X(x) = \frac{1}{\lambda} \frac{d}{dx} f_Y(x) + f_Y(x)$$

which using equation (11) yields

$$f_X(x) = \begin{cases} \frac{\lambda}{e + \lambda T} & 0 < x < T \\ \delta(x)\frac{1}{e + \lambda T} + \delta(x-T)\frac{e-1}{e + \lambda T} & \text{o.w.} \end{cases} \qquad (12)$$

– a valid probability density function.

We can now state the maximum mutual information as

$$\max_{f_X} I(X;Y) = \log \left(1 + \frac{\lambda T}{e} \right) \qquad (13)$$

which is achieved using the emission time density of equation (12).

The only remaining question is whether for interval-limited inputs, the exponential first passage time density, to quote [14] "plays the same role ... that Gaussian noise plays in additive noise channels." The answer is yes.

Theorem 1: If $X^* \in [0, T]$ has the density of equation (12), Z^* is exponential with mean $\frac{1}{\lambda}$ and Z is any other non-negative random variable with $E[Z] = \frac{1}{\lambda}$, then

$$\log \left(1 + \frac{\lambda T}{e} \right) \leq \max_{f_X} I(X; X + Z) \qquad (14)$$

Proof: Theorem 1 We adapt part (c) of Theorem 3 in [14] and note that

$$I(X^*; X^* + Z) \geq \int_0^\infty f_{Y|X}(y|x)f_X^*(x) \log \frac{f_{Y|X}^*(y|x)}{f_Y^*(y)} dx dy \qquad (15)$$

where the starred density functions are the optimizing functions found previously.

We simplify the right hand side of equation (15):

$$\int_0^\infty f_{Y|X}(y|x)f_X^*(x)\log\frac{f_{Y|X}^*(y|x)}{f_Y^*(y)}dxdy$$
$$= \int_0^\infty f_{Y|X}(y|x)f_X^*(x)\left[\log\frac{\lambda e^{-\lambda(y-x)}}{f_Y^*(y)}\right]dydx$$
$$= \log\lambda - \lambda(E[Y]-E[X]) - \int_0^\infty f_Y(y)\log f_Y^*(y)dy$$
$$= \log\frac{\lambda}{e} - \int_0^T f_Y(y)\log\frac{\lambda}{\lambda T+e}dy$$
$$\quad -\int_T^\infty f_Y(y)\left(\log\left(\frac{e\lambda}{\lambda T+e}\right)-\lambda(y-T)\right)dy$$
$$= \log\left(1+\frac{\lambda T}{e}\right) - \int_T^\infty f_Y(y)(1-\lambda(y-T))dy$$
$$= \log\left(1+\frac{\lambda T}{e}\right) - \lambda\bar{F}_Y(T)\left(\frac{1}{\lambda}+T-E[Y|II]\right)$$
$$= \log\left(1+\frac{\lambda T}{e}\right) + \lambda\bar{F}_Y(T)\left(E[Y|II]-E[Y]\right)$$

where $\bar{F}_Y(T)(y)$ is the complementary cumulative distribution function (CCDF) of the density for $Y = Z + X^*$.

Then, since $E[Y|II] \geq E[Y]$ we must have via equation (15)

$$I(X^*; X^* + Z) \geq \log\left(1+\frac{\lambda T}{e}\right)$$

and thence

$$\max_{f_X} I(X; X+Z) \geq \log\left(1+\frac{\lambda T}{e}\right)$$

•

Thus, equation (13) constitutes a min-max bound on the mutual information between an input constrained to $[0,T]$ and non-negative additive noise with mean $\frac{1}{\lambda}$.

V. DISCUSSION AND CONCLUSION

For large $\lambda T \gg e$, the uniform portion of $f_X()$ dominates the emission time distribution since deviation of the arrival time around $X+1/\lambda$ is small. Thus, the receiver can distinguish between relatively fine gradations of emission times within the interval $[0,T]$. The endpoint emission times essentially become special instances of emission times within the interval, and a cardinality of only 2 can be increasingly neglected as λT becomes large. Thus, as a limit we have,

$$f_{X|\lambda T\gg e}(x) = \begin{cases} \dfrac{1}{T} & 0 < x < T \\ 0 & \text{o.w.} \end{cases} \quad (16)$$

Conversely, when λT is small, emissions must be "rushed" relative the mean first passage time $1/\lambda$. Thus, emissions within $[0,T]$ become increasingly indistinguishable. Since increased distance between emission times increases distinguishability, the dominant emission probability mass is placed at the endpoints. Thus, as a limit we have,

$$f_{X|\lambda T\ll e}(x) = \begin{cases} 0 & 0 < x < T \\ \delta(x)\dfrac{1}{e}+\delta(x-T)\dfrac{e-1}{e} & \text{o.w.} \end{cases}$$
$$(17)$$

The results suggest that when first passage times are highly variable as might be the case for larger emitter-receptor pair distances, burst-release of signaling agents at interval boundaries most effectively conveys information.

Likewise, if the variability of first passage time is low, as might be the case where emitter-receptor pairs are close to one another and/or some active and predictable transport method is used, then more information can be conveyed by spreading emission times over the interval. In the latter case, under an assumption of infrequent emission of individual sites but tightly coupled emitter-receptor pairings, the results presented here may approximate the maximum mutual information since the tight coupling implicitly identifies molecules, and since infrequent emissions leads to approximately independent channel use of each emitter-receptor pair.

REFERENCES

[1] O. Milenkovic, G. Alterovitz, G. Battail, T. P. Coleman, J. Hagenauer, S. P. Meyn, N. Price, M. F. Ramoni, I. Shmulevich, and W. Szpankowski. "introduction to the special issue on information theory in molecular biology and neuroscience". *Trans. Information Theory*, 56(2):649–652, 2010.

[2] C. de Joussineau, J. Soule, M. Martin, C Anguille, P. Montcourrier, and D. Alexandre. Delta-Promoted Filopodia Mediate Long-Range Lateral Inhibition in Drosophila. *Nature*, 426:555–559, December 4 2003.

[3] Y.A. Gorby, S. Yanina, JS.. McLean, K.M. Rosso, D. Moyles, A. Dohnalkova, T.J. Beveridge, I.S. Chang, B.H. Kim, K.S. Kim, D.E. Culley, S.B. Reed, M.F. Romine, D.A. Saffarini, E.A. Hill, L. Shi, D.A. Elias, D.W. Kennedy, G. Pinchuk, K. Watanabe, S. Ishii, B. Logan, K.H. Nealson, and J.K. Fredrickson. Electrically conductive bacterial nanowires produced by *Shewanella oneidensis* strain MR-1 and other microorganisms. *Proc Natl Acad Sci U.S.A.*, 103:11358–11363, 2006.

[4] S. Gurke, J.F.V. Barroso, and H.-H. Gerdes. The art of cellular communication: tunneling nanotubes bridge the divide. *Histochem Cell Biol*, 129:539–550, 2008.

[5] X. Wang, M.L. Veruki, N.V. Bukoreshtliev, E. Hartveit, and H.-H. Gerdes. Animal cells connected by nanotubes can be electrically coupled through interposed gap-junction channels. *Proc Natl Acad Sci USA*, 107:17194–17199, 2010.

[6] H.C. Berg and E.W. Purcell. Physics of Chemoreception. *Biophysical Journal*, 20:193–219, 1977.

[7] Pankaj Mehta et al. Information processing and signal integration in bacterial quorum sensing. *Molecular systems biology*, 2009.

[8] A.L. Hodgkin and A.F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.*, 117(4):500–544, 1952.

[9] Tao Long et al. Quantifying the integration of quorum-sensing signals with single-cell resolution. *Molecular systems biology*, 2009.

[10] Elek Wajnryb Jose M. Amigo, Janusz Szczepanski and Maria V. Sanchez-Vives. "estimating the entropy rate of spike trains via lempel-ziv complexity". *Neural Computation*, 16:717–736, 2004.

[11] D.H. Johnson. Information Theory and Neural Information Processing. *Trans. Information Theory*, 56(2):653–666, Feb 2010.

[12] Riccardo Barbieri, Loren M. Frank, David P. Nguyen, Michael C. Quirk, Victor Solo, Matthew A. Wilson, and Emery N. Brown. "dynamic analyses of information encoding in neural ensembles". *Neural Computation*, 16:277–307, 2004.

[13] S. Mian and C. Rose. Communication theory and multicellular biology. *Integrative Biology*, 3(4):350–367, April 2011.

[14] V. Anantharam and S. Verdu. Bits Through Queues. *IEEE Transactions on Information Theory*, 42(1):4–18, January 1996.

[15] C. Nusslein-Volhard. *Coming to Life: how genes drive development*. Kales Press, 2006.

[16] T. M. Cover and J. A. Thomas. Elements of information theory, 2006.

[17] T.P. Coleman and M. Raginsky. Mutual information saddle points in channels of exponential family type. In *Information Theory Proceedings (ISIT), 2010 IEEE International Symposium on*, pages 1355 –1359, June 2010.