Capacity estimates of additive inverse Gaussian molecular channels with relay characteristics

Pratip Rana
Department of Computer Science,
Virginia Commonwealth University,
Richmond, VA 23284, USA
ranap@vcu.edu

Preetam Ghosh
Department of Computer Science,
Virginia Commonwealth University,
Richmond, VA 23284, USA
pghosh@vcu.edu

Kevin R. Pilkiewicz
Environmental Laboratory,
US Army Engineer Research and
Development Center
Vicksburg, MS 39180, USA
Kevin.R.Pilkiewicz@usace.army.mil

Edward J Perkins
Environmental Laboratory,
US Army Engineer Research and
Development Center
Vicksburg, MS 39180, USA
Edward.J.Perkins@usace.army.mi

Chris Warner
Environmental Laboratory,
US Army Engineer Research and
Development Center
Vicksburg, MS 39180, USA
Christopher.M.Warner@usace.army.mil

Michael Mayo
Environmental Laboratory,
US Army Engineer Research and
Development Center
Vicksburg, MS 39180, USA
Michael.L.Mayo@usace.army.mil

ABSTRACT
Molecular communications is an emergent field of research that seeks to develop novel, nanoscale communications devices using design principles gleaned from studies of the topology and dynamic properties of biological signaling networks. To understand how these networks function as a whole, we must first identify and characterize the functional building blocks that compose them, and the best candidates for those are the topologically distinct subnetworks, or motifs, that appear in a statistically improbable abundance within these networks. In cellular transcriptional networks, one of the most prevalent motifs is the feed-forward loop, a three node motif wherein one top-level protein regulates the expression of a target gene either directly or indirectly through an intermediate regulator protein. Currently, no systematic effort has been made to treat an isolated feed-forward loop as a stand-alone signal amplifying/attenuating device and understand its communication capacity in terms of the diffusion of individual molecules. To address this issue, in this paper we derive a theorem that estimates the upper and lower bounds of the channel capacity for a relay channel, which structurally corresponds to a feed-forward loop, by using an additive inverse Gaussian noise channel model of protein-ligand binding. Our results are just a first step towards assessing the performance bounds of simplified biological circuits in order to guide the development and optimization of synthetic, bio-inspired devices that can be used as information processing and forwarding units.

Categories and Subject Descriptors
H.1.1 [**Systems and Information Theory**]: *Information theory*.

General Terms

Performance, Reliability, Theory.

Keywords

Additive inverse Gaussian channel; Channel capacity; Molecular communication; multiple access, broadcast and relay channels.

1. INTRODUCTION
Organisms use signaling molecules to convey information about their molecular state, which may be encoded in the pattern of protein-coding gene expression, the magnitude and frequency of protein bursts, or the spike-train of neuron action potentials. Many of these processes require molecules to diffuse throughout the intracellular environment, where they experience intrinsic noise, from factors such as low copy number or probabilistic biochemical reactions, as well as extrinsic noise, arising from heterogeneities in cell size or cell-cycle stage [6]. In many cases, molecular transport can be approximated by a relatively simple Wiener process [1], which makes it feasible to pose and study molecular channel models using information theoretic metrics, such as mutual information and channel capacity that may reveal fundamental principles or "laws" that underlie biological communication.

The mutual information, I, of a communication channel measures the information content of one random variable given knowledge about another. The channel capacity is the maximum mutual information that can be found by varying the source probability distribution. For a simple molecular "channel" in which the signaling molecules are propagated from one end to the other, analytic results for the mutual information or channel capacity have been found in the case of diffusive channels with drift [1,2]. The inverse Gaussian distribution provides a good approximation to characterize the first passage times of a Wiener process with drift; hence several follow-up works on molecular channel capacities have used this distribution with either average-delay [3] or peak-delay constraints [4]. Additionally, bounds for the capacity of memoryless additive inverse Gaussian noise channels have been reported before [5].

Depending on the liquid medium, it may be possible to construct a

molecular communication channel in which diffusing ligands interact with a receptor at both short and long time scales. The liquid medium primarily affects the variance of the underlying Wiener process through the diffusion coefficient of molecules. Previous work in wireless networks [9-11] suggests that the information-processing roles of the feed-forward loop may extend beyond the biological setting, which motivates an information-theoretic investigation of relay molecular communication channels. Although rarely studied, one report describes a molecular relay channel [12] in which molecules are continuously emitted from a source, diffuse through the medium, and then "react" with the receptor of a ligand binding system (e.g., a g-protein coupled receptor). This model did not explicitly consider the channel noise in transmitting individual molecules, however; instead, the ligand-receptor binding affinity was used to directly relate the concentration of molecules at the source to that at the destination. Another effort to estimate the channel capacity of the feed forward motif [16] was made using kinetic equations and binding. Here the authors used the Gaussian noise distribution arbitrarily; the other parameters are also very difficult to calculate.

Here we report the mutual information of a relay molecular communication channel that leverages biological molecules as the information carriers, transported through a liquid medium from transmitter to receiver under a drift-diffusion mechanism. Previous work [7] suggests that a molecular channel can be modeled by an inverse Gaussian Noise Channel, which we have used to estimate the bounds on the capacity of the relay communication channel.

2. METHODS
2.1 Molecular Communication Channel Model
Biomolecules move within the cytosol from one location to another, carrying information about a stimulus that gets "decoded" by the receiver in the form of a measurable response (e.g., altered gene expression). A prototype for such a channel consists of a "transmitter" of constant molecular emission together with a "receiver" separated by a liquid or gaseous medium spanning d units of length. Emitted molecules travel via one-dimensional Brownian motion with a drift velocity, v. A closed form solution is not possible for the additive inverse Gaussian noise (AIGN) channel, but upper and lower bounds on the exact capacity have been reported [7].

$$C \leq \log\big((m + \mu)e\big) - h_{IG(\mu,\lambda)} \qquad (1)$$
$$C \geq h_{IG(m+\mu,(\lambda/\mu^2)(m+\mu)^2)} - h_{IG(\mu,\lambda)} \qquad (2)$$

In the above, m is the average time between successive source molecule emissions, $\mu = d/v$ is the time taken to traverse the medium, $\lambda = d^2/\sigma^2$, σ^2 is the variance of the first-passage time distribution to reach the receiver, and h_{IG} is the differential entropy of the Inverse Gaussian distribution (as shown in Equation 3).

An approximate expression for $h_{IG(\mu,\lambda)}$ can be found in [5]:

$$h_{IG(\mu,\lambda)} = \frac{1}{2}\log\frac{2\pi e\mu^3}{\lambda} + \frac{3}{2}e^{\frac{2\lambda}{\mu}}Ei\left(-\frac{2\lambda}{\mu}\right) \qquad (3)$$

Using equation (3), equations (1-2) were evaluated in [5, 7] to give:

$$C \leq \frac{1}{2}\log\frac{\lambda e(m + \mu)^2}{2\pi\mu^3} + \frac{3}{2}e^{\frac{2\lambda}{\mu}}Ei\left(\frac{2\lambda}{\mu}\right), \qquad (4)$$

and

$$C \geq \frac{1}{2}\log\frac{m + \mu}{m} + \frac{3}{2}e^{\frac{2\lambda(m+\mu)}{\mu^2}}Ei\left(\frac{2\lambda(m + \mu)}{\mu}\right) - \frac{3}{2}e^{\frac{2\lambda}{\mu}}Ei\left(\frac{2\lambda}{\mu}\right) \qquad (5)$$

where $Ei(x)$ denotes the exponential integral function [5].

The above capacity results for molecular channels can be contrasted with the capacity of the additive exponential noise (AEN) channel, where the induced noise in the channel or first arrival time is exponential in nature. For the AEN mean constrained channel, the capacity upper bound depends only on the constrained mean, $E(X) = m$, and the mean of the exponential noise, b, and is given by the following expression [8].

$$C \leq \log\left(1 + \frac{m}{b}\right) \qquad (6)$$
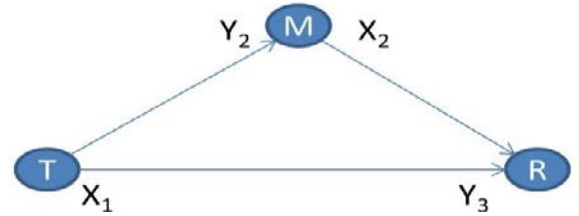
2.2 Capacity of Molecular Relay Channels

Figure 1. Molecular relay channel: there is a relay node in between the transmitter and receiver.

In a relay channel (Figure 1), one or more nodes assist to communicate between a single transmitter and receiver. The model for relay channels was first introduced by Van der Meulen et al. [13]. The capacity upper bound for the deterministic relay network is given by [14, 15]:

$$C \leq \frac{max}{p(x_1, x_2, .., x_{n-1})} \frac{min}{S \subseteq n} I(X_1 X_S; Y_{S^c} Y_n | X_{S^c}) \qquad (7)$$

In the above, n is the set of all nodes, S is the subset of nodes in the relay network, and S^C is the complement of S in n. The symbol X_i denotes channel input from node i and Y_i denotes channel output to node i.

We assume that the relay node M receives the information from T, decodes it, re-encodes it and then transmits to R. This is the traditional store-and-forward communication model in electromagnetic communication networks. We assume that the decoding and encoding of information at node M is instantaneous and all molecules are distinguishable; hence, there is no additional delay at node M before forwarding the messages. Because the molecules are distinguishable, there is no requirement for dropping any molecule at node M, unlike in traditional communication models, where multiple copies of the same packet may be transmitted and hence additional copies of the same

packet are dropped at the relay node before forwarding. Also, we assume that the distribution of molecule transmissions from the relay node M only depends on the parameters of the next channel i.e. the channel from the relay node to the receiver; this assumption holds here as out of order arrival of molecules at M does not affect the communication performance of the relay to receiver channel. With these simplifying assumptions, the molecular relay channel becomes a deterministic relay channel; hence the theorems for the traditional deterministic relay channels can be applied here. Therefore, equation (7) for the relay network channel capacity (C_R) can be rewritten in terms of the single molecular channel capacities from Section 2 as follows:

$$C_R \leq \begin{matrix} max \\ p(x_T, x_M) \end{matrix} \min\{I(X_T; Y_M Y_R | X_M), I(X_T X_M; Y_R)\} \quad (8)$$

From equation (8) we can derive equation (9) by following the steps shown in [15].

$$C_R \leq \min\{C_{TM} + C_{TR}, C_{TR} + C_{MR}\} \quad (9)$$

Here C_{TM} signifies the capacity of the channel from node T to node M and so on. We may now find the capacity upper bound of the channel C_{TR} using equation (4). Similarly, the capacity of the channels C_{MR} and C_{TR} can also be derived. Moreover, the lower bounds of the capacities can be calculated by combining equations (5) and (9).

3. RESULTS AND DISCUSSION

In this section, we present the numerical results for our channel capacity model of a one node molecular relay network.

In Figure 2, the upper and lower bounds of the channel capacity of a molecular relay channel are plotted against the delay constraint m. We can observe that both the upper and lower bounds increase monotonically with increasing delay constraint. As the delay constraint increases, there is more information on the transmission of molecules resulting in an increase in the mutual information and correspondingly an increase in the capacity of the channel.

The green line of Figure 2, shows the capacity of the AEN relay channel with respect to the delay constraint. As with the inverse Gaussian channel, the capacity increases with the delay constraint, but due to exponential noise the capacity value is considerably lower than that of the AIGN channel. Surprisingly, while the upper bound for the additive inverse Gaussian noise channel is non-zero at zero delay constraint, the upper bound for the AEN relay channel is zero. Both channel capacities exhibit similar behavior because the channel assumptions for our simplified relay network set-up were comparable for the two types of channels.

The upper bound of the relay channel for lower values of m and different velocity parameters for each channel is shown in Figure 3. Interestingly, for the given parameters, at lower values of m the upper bound of the indirect channel capacity (based on the indirect path) is the highest and the direct channel capacity (based on the direct path) is the lowest; but after that, the channel with the higher velocity parameter starts to increase quickly, and after a certain critical value of m, the C_{TR} channel capacity becomes highest while the capacity of C_{TM} becomes lowest. For this reason, the upper bound of the capacity of the relay channel is initially ($C_{TR}+C_{MR}$), but beyond the critical value of m it becomes ($C_{TR}+C_{TM}$). This observation may, however, be an artifact of the

different drift velocities assumed for each channel and not necessarily because m is small. For the lower bound, on the other hand, there is no crossover and the capacity is always $C_{TR}+C_{TM}$. The log term in the expression for the upper bound might account for this discrepancy.
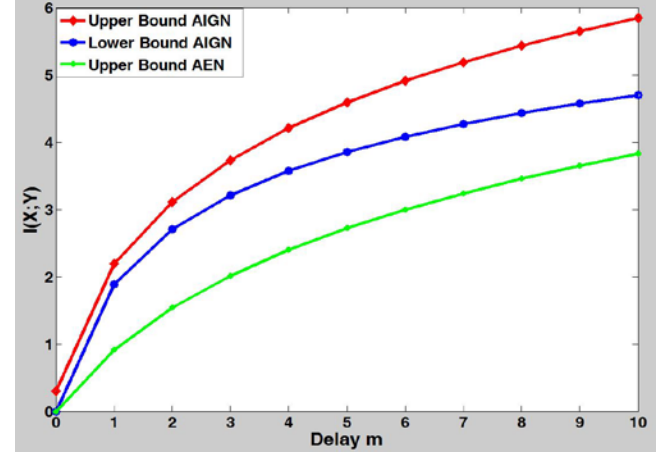
Figure 2: Upper and lower bound of capacity of the molecular relay channel as a function of the delay constraint. The drift velocity has been set to v = 2, 1, 1.5 for C_{TR}, C_{TM}, C_{MR} and the channel parameter λ is assumed to be 0.25. Upper bound of the capacity for the AEN channel is shown as a green line. The mean of exponential noise b is set to 2.
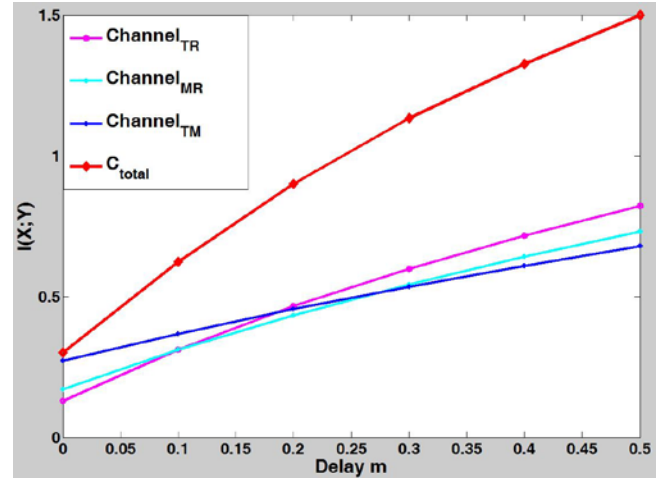
Figure 3: Upper bounds of the capacity of the molecular relay channel for smaller values of the delay constraint. The drift velocity has been set to v = 2, 1, 1.5 for C_{TR}, C_{TM}, C_{MR}, and the channel parameter λ is assumed to be 0.25

The upper and lower bounds of the capacity were next plotted versus the drift velocity in Figure 4. Both the upper and lower bounds of the capacity increase with higher drift velocity. Although the lower bound tends to saturate much more quickly, the upper bound may still show an increasing trend with increasing velocity. Surprisingly, for low drift velocities (Figure 5), a sharp decrease in the upper value of the capacity of the relay channel is observed. The capacity upper bounds of the three individual channels also exhibit this behavior. One possible explanation for this sharp drop in capacity is the presence of both drift and diffusion in the partial differential equation used to derive the upper bound, as explained in [7].
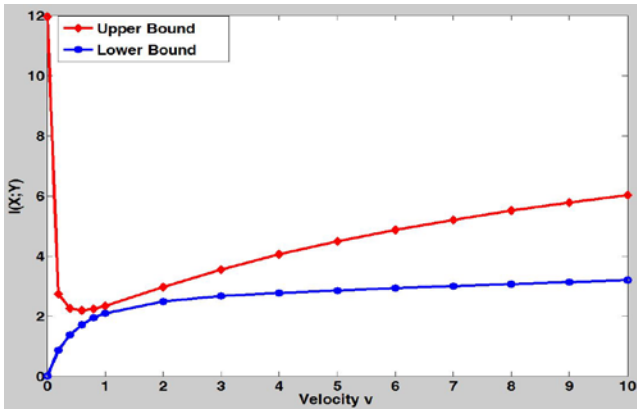
Figure 4: Upper and lower capacity bounds for the molecular relay channel as a function of drift velocity. The delay constraint has been set to m = 2, 1, 1.5 for C_{TR}, C_{TM}, C_{MR}, and the channel parameter λ is assumed to be 0.25
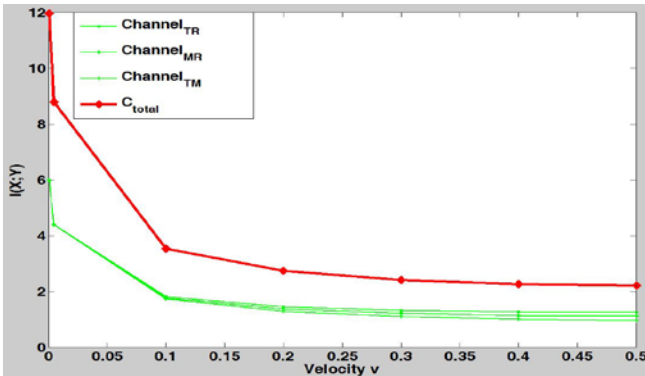

Figure 5: Upper bounds of the capacity of the molecular relay channel for smaller values of the drift velocity. The delay constraint has been set to m = 2, 1, 1.5 for C_{TR}, C_{TM}, C_{MR} and the channel parameter λ is assumed to be 0.25

4. CONCLUSION

In this paper, we have analyzed the capacity of a molecular relay channel. While results from traditional EM communication-based channel models can be used directly for molecular channels, we have specifically worked with an inverse Gaussian noise distribution that has been mostly connected to protein-ligand binding channels. Modeling the capacity of the molecular relay channel is of particular interest, not only because it structurally mimics the feed-forward loop motifs that are abundant in biological networks, but also because it may be useful as a stand-alone communication circuit. For example, the upper bounds on the channel capacity model show saturation beyond a critical delay constraint, which can mean that there is an upper limit on the signal augmentation properties of feed-forward loops. This work is a first step in understanding the efficiency of such molecular circuits, and our future work will focus on extending the model to more complex motifs that include down as well as up-regulating links.

5. ACKNOWLEDGMENTS

6. REFERENCES

[1] Eckford, A. W. (2008). Molecular communication: Physically realistic models and achievable information rates. *arXiv preprint arXiv:0812.1554*.

[2] Kadloor, S., Adve, R. S., & Eckford, A. W. (2012). Molecular communication using brownian motion with drift. *NanoBioScience, IEEE Transactions on*, *11*(2), 89-9.

[3] Srinivas, K. V., Eckford, A. W., & Adve, R. S. (2012). Molecular communication in fluid media: The additive inverse gaussian noise channel. *Info. Theory, IEEE Tran.*, *58*(7), 4678-92.

[4] Eckford, A. W., Srinivas, K. V., & Adve, R. S. (2012). The peak constrained additive inverse Gaussian noise channel. In *Info. Theory Proc. (ISIT), 2012 IEEE Intl. Symp.*, 2973-2977.

[5] Chang, H. T., & Moser, S. M. (2012). Bounds on the capacity of the additive inverse Gaussian noise channel. In *Info. Theory Proc. (ISIT), 2012 IEEE Intl. Symp.*, 299-303.

[6] Raser, J. M., & O'Shea, E. K. (2004). Control of stochasticity in eukaryotic gene expression. *Science*, *304*(5678), 1811-1814.

[7] Magnan, S. and Alon, U. (2003) Structure and function of the feed-forward loop network motif. *Proc. Natl. Acad. Sci. USA*.

[8] Tsai, Y., Rose, C., & Song, R. (2011, July). An additive exponential noise channel with a transmission deadline. In *Info. Theory Proc. (ISIT), 2011 IEEE Intl. Symp.*, 718-722.

[9] Abdelzaher, A., Mayo, M., Perkins, E., and Ghosh, P. Contribution of Canonical Feed-forward Loop Motifs on the Fault-tolerance and Information Transport Efficiency of Transcriptional Regulatory Networks, *Elsevier Nano Communication Networks,* 2015.

[10] Kamapantula, B, Abdelzaher, A, Ghosh, P., Mayo, M., Perkins, E., and Das, S.K. Leveraging the Robustness of Genetic Networks: A Case Study on Bio-inspired Wireless Sensor Topologies. *Journal of Ambient Intelligence and Humanized Computing*, 2014, DOI : 10.1007/s12652-013-0180-0.

[11] Ghosh, P., Mayo, M., Chaitankar, V., Habib, T., Perkins, E., and Das, S.K. Principles of genomic robustness inspire fault-tolerant wireless sensor network designs, *PerSens Workshop (part of PerCom 2011)*, pp. 160-165.

[12] Atakan, B., & Akan, O. B. (2008). On molecular multiple-access, broadcast, and relay channels in nanonetworks. *In Proceedings of the 3rd International Conference on Bio-Inspired Models of Network, Information and Computing Systems*, pp. 16.

[13] Van der Meulen, EC. (1969). *Transmission of information in a T-terminal discrete memoryless channel*. In *annals of mathematical statistics* 40(3), pp. 1157.

[14] Cover, T., & Gamal, A. E. (1979). Capacity theorems for the relay channel. *Info. Theory, IEEE Tran. 25*(5), 572-584.

[15] Kramer, G., Gastpar, M., & Gupta, P. (2005). Cooperative strategies and capacity theorems for relay networks. *Info. Theory, IEEE Tran. 51*(9), 3037-3063.

[16] Hormoz, S. (2013). Cross talk and interference enhance information capacity of a signaling pathway. *Biophysical journal*, *104*(5), 1170-1180.