Benchmarking the communication fidelity of biomolecular signaling cascades featuring pseudo-one-dimensional transport

Cite as: AIP Advances **8**, 055220 (2018); https://doi.org/10.1063/1.5027508
Submitted: 02 March 2018 . Accepted: 09 May 2018 . Published Online: 18 May 2018

Pratip Rana, Kevin R. Pilkiewicz, Michael L. Mayo, and Preetam Ghosh

Benchmarking the communication fidelity of biomolecular signaling cascades featuring pseudo-one-dimensional transport

Pratip Rana,[1] Kevin R. Pilkiewicz,[2,a] Michael L. Mayo,[2] and Preetam Ghosh[1]
[1]*Department of Computer Science, Virginia Commonwealth University, Richmond, VA 23284, USA*
[2]*Environmental Laboratory, U.S. Army Engineer Research and Development Center, Vicksburg, MS 39180, USA*

(Received 2 March 2018; accepted 9 May 2018; published online 18 May 2018)

Synthetic biologists endeavor to predict how the increasing complexity of multi-step signaling cascades impacts the fidelity of molecular signaling, whereby information about the cellular state is often transmitted with proteins that diffuse by a pseudo-one-dimensional stochastic process. This begs the question of how the cell leverages passive transport mechanisms to distinguish informative signals from the intrinsic noise of diffusion. We address this problem by using a one-dimensional drift-diffusion model to derive an approximate lower bound on the degree of facilitation needed to achieve single-bit informational efficiency in signaling cascades as a function of their length. Within the assumptions of our model, we find that a universal curve of the Shannon-Hartley form describes the information transmitted by a signaling chain of arbitrary length and depends upon only a small number of physically measurable parameters. This enables our model to be used in conjunction with experimental measurements to aid in the selective design of biomolecular systems that can overcome noise to function reliably, even at the single-cell level.
https://doi.org/10.1063/1.5027508

The field of synthetic biology has spurred the development of biotechnologies capable of interfacing with the human body at the molecular level. From CRISPR/Cas9 gene drives that alter our genetic code[1] to proteins designed to modulate the sensitivity of molecular signaling cascades,[2,3] the state of the art has learned to imitate nature and, in so doing, surpass it. The cost of interfacing with biology at its most granular level, however, is that we must leverage fragile biomolecular components that transmit information far below the modern standards of electronic communications. For example, the transcriptional signaling pathways in a cell cannot reliably cross the minimal single-bit threshold required to distinguish meaningful signals from noise,[4] and coherent communication is only achieved when signaling is averaged over an entire cellular population.[4–6]

Although some biological signaling does utilize electric currents, such as that in the human nervous system,[7] information transmitted at the cellular level is often accomplished via so-called molecular communication, in which discrete signaling molecules diffuse through an intervening liquid medium from a transmitting to a receiving site.[8,9] Because Brownian motion alone is slow and unreliable, many biological systems attempt to facilitate this diffusion with a variety of mechanisms, such as motor proteins that "walk" the molecules along cytoskeletal filaments,[10] transcription factors that locate their receiving site by sliding back and forth along the DNA contour,[11] and bacterial chemotaxis, in which molecules are carried by bacteria along a chemical gradient.[12] These types of directed-diffusion processes typically exhibit a superdiffusive mean squared displacement that scales as t^{α}, for $1 < \alpha < 2$,[13] and are often well approximated as one-dimensional stochastic processes

[a]Electronic mail: Kevin.R.Pilkiewicz@usace.army.mil

whose superdiffusivity can be decomposed as a linear combination of pure ballistic and pure diffusive motion.[14] The ratio of the drift speed to the diffusion constant can be used to determine α for any time interval of interest. Such drift-diffusion processes have been studied as models of molecular communications in the past.[15–17]

As synthetic biology pushes the boundary of more complex multi-stage signaling architectures,[18] a more fundamental understanding of the information-processing capabilities of molecular communication is needed to optimally design constructs that perform reliably despite inherently noisy *in vivo* environments. To this end, single drift-diffusion channels have been studied extensively,[15–17,19,20] and various upper and lower bounds have been placed on their mutual information and channel capacity.[21,22] Here we extend these analyses to a signaling cascade of multiple drift-diffusion channels daisy-chained together and demonstrate, under suitable simplifying assumptions, that the mutual information depends upon a single dimensionless ratio of physical parameters. We then use this result to set a lower bound on the extent of passive molecular transport required to achieve molecular communication with single-bit fidelity.

As our model system, we consider a one-dimensional molecular drift-diffusion channel of length ℓ. Molecules are emitted at the transmitting end ($x = 0$) of a one-dimensional channel, and move toward a receiving end ($x = \ell$) with drift speed $v > 0$ where they are perfectly absorbed. Molecules diffuse according to a stochastic Wiener process characterized by diffusion constant D. We assume that molecules are indistinguishable from one another, that only two molecules can occupy the channel at any given time, and that both molecules must exit the channel before any new molecules can be added. Although these latter assumptions restrict the information capacity of our channel,[21] they will immensely simplify our calculations and are consistent with our objective to establish a minimal performance benchmark for the efficiency of passive transport required to communicate a single bit of information.

With only two molecules in the channel at a time, this single bit corresponds to the capacity to distinguish the order of emission times from measurements of the order of arrival times. For a more general system, where an arbitrary number of molecules can simultaneously occupy the channel, this determination of time orderings will require more than a single bit of information, necessitating a greater degree of facilitated transport to achieve it. Nonetheless, our simple, two-particle model may still provide a reasonable estimate of the information capacity for some biological systems. For example, protein transcription often occurs in short stochastic bursts followed by much longer intervals of inactivity,[23] with each burst often producing as few as two RNA transcripts.[24]

We define $\Delta\tau$ as the difference in the release times of the two molecules sharing the channel and model it as a random variable chosen from a known source distribution $p(\Delta\tau)$. As a result, the difference in the arrival times of the two molecules at the receiver is also a random variable, which we denote Δt. The information-processing capabilities of this channel can be quantified as the mutual information of these two time differences, $I(\Delta t; \Delta\tau)$. Using standard definitions and methods,[25] one can compute this mutual information from only two probability distribution functions: the source distribution, $p(\Delta\tau)$, and the conditional distribution, $p(\Delta t|\Delta\tau)$. We can compute this latter distribution by noting that the time it takes for a molecule to cross our modeled channel is a random variable obeying an inverse Gaussian distribution, $IG(\mu, \lambda; t)$:[19]

$$IG(\mu, \lambda; t) = \begin{cases} \sqrt{\frac{\lambda}{2\pi t^3}} \exp\left[\frac{-\lambda(t-\mu)^2}{2\mu^2 t}\right] & t > 0 \\ 0 & t \leq 0 \end{cases} \tag{1}$$

wherein the mean first passage time, μ, is equal to ℓ/v, and the shape parameter, λ, is equal to $\ell^2/2D$, the average time it would take a particle to traverse the channel in the absence of any drift.

If we assume that the first molecule is transmitted at time zero and is received at elapsed time t, then the second molecule is transmitted at time $\Delta\tau$ and received at time $t + \Delta t$. We allow Δt to be negative, in which case the second molecule to be released arrives at the receiver before the first. Of course, it is not possible for an observer at the receiver to measure a negative time difference

between the arrivals of two indistinguishable particles. Rather, the observer would deduce that $\Delta\tau$ is negative for molecules that arrive in reverse order, assuming the channel can transmit the one bit of information required to make that determination with certainty. The probability that $\Delta\tau$ leads to Δt is the probability that the first and second-released particles have first passage times t and $t + \Delta t - \Delta\tau$, respectively, integrated over all possible values of t:

$$p(\Delta t|\Delta\tau) = \int_0^\infty dt\, IG(\mu, \lambda; t + \Delta t - \Delta\tau) IG(\mu, \lambda; t). \tag{2}$$

Although Eq. (2) is not strictly a convolution, it can, in principle, be evaluated using the known Fourier transform of the inverse Gaussian distribution.[26]

In Fig. 1A we plotted Eq. (2) for several sets of parameters (solid curves), and compared each conditional distribution to a Gaussian distribution with identical mean and variance (dashed curves). The Gaussian fit is quantitative when the variance of the exact distribution is small, and reasonably good even when it is large; so we shall henceforth use this fit to approximate $p(\Delta t|\Delta\tau)$.

The exact mean and variance of the conditional distribution can be computed from simple arguments. Because the two molecules propagate across the channel without interaction, the mean of the difference in their first arrival times, $\Delta t - \Delta\tau$, is the difference in the means of the independent, identically distributed inverse Gaussian random variables $t + \Delta t - \Delta\tau$ and t. The total variance is the summation of their individual variances.

The mean of the distribution $p(\Delta t - \Delta\tau)$ is consequently zero, which implies that the mean of $p(\Delta t|\Delta\tau)$ is $\Delta\tau$. The variance of an inverse Gaussian distribution is μ^3/λ, so the variance

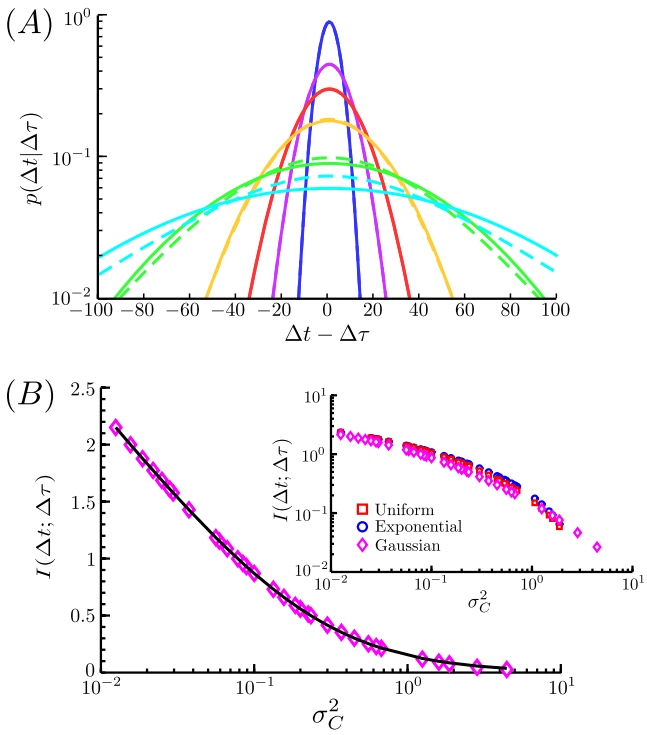

FIG. 1. (A) The conditional probability distribution, $p(\Delta t|\Delta\tau)$, as obtained by numerical evaluation of Eq. (2), is plotted as a function of the difference $\Delta t - \Delta\tau$ for six different values of the conditional variance, $\sigma_C^2 \in \{0.2, 0.8, 1.8, 5.0, 20, 45\}$. (The smaller variances naturally correspond to the more sharply peaked distributions.) For each conditional distribution, the Gaussian distribution with matching variance is plotted as a dashed curve and is a near perfect fit for smaller values of σ_C^2. (B) Numerically integrated values of the mutual information $I(\Delta t; \Delta\tau)$ (purple diamonds), expressed in bits, for different values of σ_C^2. The approximate information (Eq. (4); black line) overlays the numerical results. (Inset) The numerical values for the mutual information in the main plot (purple diamonds) are compared against those computed from different source distributions–a uniform distribution (red squares) and an exponential distribution (blue circles)–both parametrized to have the same information content as the Gaussian distribution.

of the conditional distribution, σ_C^2, is twice this value. In terms of the channel parameters, this reduces to

$$\sigma_C^2 = \frac{4D\ell}{v^3}. \tag{3}$$

If the source distribution, $p(\Delta\tau)$, is also Gaussian with a variance of σ_S^2, then it can be shown that the mutual information, $I(\Delta t; \Delta\tau)$, has a form analogous to one derived by Shannon and Hartley:[27]

$$I(\Delta t; \Delta\tau) = \frac{1}{2}\log_2\left(1 + \frac{\sigma_S^2}{\sigma_C^2}\right). \tag{4}$$

To arrive at this result, we must assume that the mean release-time interval is sufficiently larger than σ_S, so that the tail of the Gaussian $p(\Delta\tau)$ contributes negligible probability for $\Delta\tau < 0$. Fig. 1B plots the mutual information (in bits), computed numerically from the true conditional distribution (Eq. (2)), for several different values of σ_C^2 (Eq. (3)). The approximate mutual information in Eq. (4) is plotted over the data, demonstrating an almost exact fit. The inset of Fig. 1B demonstrates that the validity of Eq. (4) is not significantly affected by our decision to use a Gaussian as our source distribution.

It is tempting at this point to try to generalize Eq. (4) to the case of an N particle channel. It is simple to argue, for example, that the equivalent to Eq. (2) is

$$p(\{\Delta t_{1i}\}|\{\Delta\tau_{1i}\}) = \int_0^\infty dt\, IG(\mu, \lambda; t) \prod_{i=2}^{N} IG(\mu, \lambda; t + \Delta t_{1i} - \Delta\tau_{1i}), \tag{5}$$

where $\Delta\tau_{1i}$ and Δt_{1i} are the release and absorption time intervals of the i^{th} particle, and all $N-1$ of these intervals are defined relative to the absolute release and absorption times of the same reference particle, labeled particle 1. While in the $N = 2$ case, we have demonstrated that this conditional distribution is quantitatively Gaussian in shape, the same cannot be said for $N > 2$. Indeed, numerical evaluation of Eq. (5) for even $N = 3$ yields a conditional distribution that is a highly asymmetric function of its arguments, making it a poor candidate for a Gaussian fit. While Eq. (5) can still be used alongside a suitably chosen source distribution to numerically evaluate the mutual information, the conveniently simple and transparent form of Eq. (4) is a consequence of Gaussian inputs and will not generalize to larger N.

A more fruitful generalization of Eq. (4) is to the case of a signaling cascade of n subchannels, each with identical v, ℓ, and D. We assume that information is carried along the entirety of this cascade by just two indistinguishable particles, so that the principal measure of informational fidelity remains the extent to which the difference in their release times can be deduced from the difference in their arrival times at the cascade terminus. Particles absorbed at the terminus of one subchannel are re-emitted at the source of the next, and these junctures constitute $n + 1$-many nodes of the cascade, which we sequentially label with the integers 0 (transmitting source) through n (final receiving site). We denote the difference in the absorption times of the particles at node i by Δt_i, and the time interval between their re-emission by $\Delta t'_i$. If each molecule were re-emitted instantaneously upon absorption, then $p(\Delta t'_i|\Delta t_i) = \delta(\Delta t'_i - \Delta t_i)$. However, we more generally assume that elapsed times between absorption and re-emission vary stochastically. For simplicity, we assume this delay time is normally distributed about the absorption time interval with a characteristic variance of σ_d^2. The conditional distribution, $p(\Delta t_n|\Delta\tau)$ (where $\Delta\tau \equiv \Delta t'_0$), may now be expressed as:

$$\begin{aligned} p(\Delta t_n|\Delta\tau) = &\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} d\Delta t_1 d\Delta t'_1 \cdots d\Delta t_{n-1} d\Delta t'_{n-1} \\ &\times p(\Delta t_n|\Delta t'_{n-1}) p(\Delta t'_{n-1}|\Delta t_{n-1}) \cdots p(\Delta t'_1|\Delta t_1) p(\Delta t_1|\Delta\tau). \end{aligned} \tag{6}$$

Here, each $p(\Delta t_i|\Delta t'_{i-1})$ is given by Eq. (2), and each $p(\Delta t'_i|\Delta t_i)$ is modeled, as stated above, by either a Dirac delta function or Gaussian distribution.

The integral in Eq. (6) is, to good approximation, a convolution of Gaussians, which evaluates to another Gaussian whose variance equals the sum of the variances of the individual

distributions contributing to the integral. The conditional variance of the n-link drift-diffusion cascade is thus

$$\sigma_C^2 = \frac{4nD\ell}{v^3} + (n-1)\sigma_d^2. \tag{7}$$

Using arguments similar to those used to derive Eq. (3), one can show that the above result is exact.

To good approximation, the mutual information of the cascade is the same as in Eq. (4), only with Eq. (7) substituted for σ_C^2. In Fig. 2, we demonstrate that this curve is indeed a universal fit for the mutual information of cascades with varying length, both in the case where there is no delay (Fig. 2) and a Gaussian-distributed delay (Fig. 2, inset) between absorption and re-emission at the intermediate nodes.

We use stochastic simulations to evaluate the mutual information for $n > 2$, because numerical evaluation of Eq. (6) becomes cumbersome for longer cascades. In these simulations, each particle travels a distance $v\delta t + \delta x$ in each time step, where δx is a normally-distributed random variable with variance $2D\delta t$ and the time step is $\delta t = 0.01$. For each set of conditions studied, 10^5 replicate simulations were performed, with the results histogrammed (bins of width δt) to generate discrete estimates of the probability distribution functions needed to calculate the mutual information. For cascades of one and two links, in which direct numerical evaluation of Eq. (6) is feasible, the simulations yield nearly identical values for the mutual information.

According to Eq. (4), the mutual information of an n-link cascade is, in principle, unbounded from above so long as the parameters ℓ, D, and v can be freely varied. However, this is not typically true for synthetic biological constructs that are meant to interface with cells through existing molecular channels. Here, the channel length, ℓ, will necessarily be no larger than the size of a cell ($\sim$100 μm) and the diffusion constant D will be controlled by the size of the signaling molecules and the hydrodynamic properties of the cytosol, typically varying between $\approx 1 - 10\mu\text{m}^2/\text{s}$.[28] Detailed studies of cellular signaling mechanisms[13] suggest that molecular transport is superdiffusive but non-ballistic, which means that v cannot be arbitrarily large. A suitable value for v may be found by fitting the mean squared displacement (MSD) of our drift-diffusion model to superdiffusive scaling data.

The position of a particle in our drift-diffusion channel at elapsed time t is equal to $x(t) = vt + \sqrt{2D}W_t$, wherein W_t is the standard Wiener stochastic process. We choose our coordinate

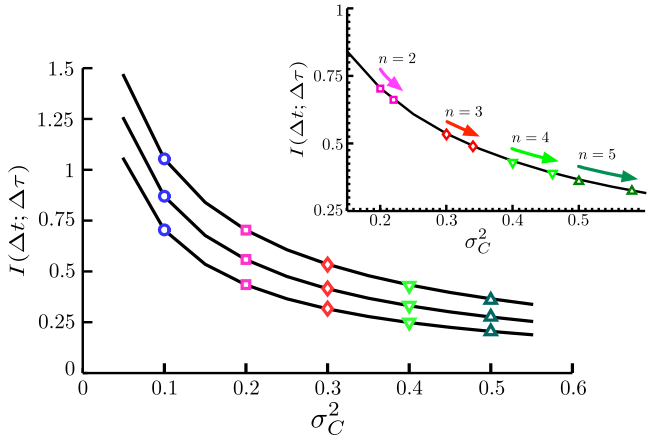

FIG. 2. Numerical values of the mutual information of a cascade channel are plotted versus the conditional variance σ_C^2 for three different values of the source variance σ_S^2 (from top to bottom, 0.331, 0.234, and 0.1655). All three datasets use the same physical parameters (ℓ, v, D) for each subchannel of the cascade, and σ_C^2 is modulated only by varying the number of subchannels. From left to right, the five points on each curve correspond to n = 1, 2, 3, 4, and 5. In all cases, the black curves correspond to Eq. (4), evaluated with Eq. (7). (Inset) The impact of having a normally-distributed delay time with variance 0.01 at each intermediate node of the cascade is illustrated using the top dataset. In each case, the effect is to reduce the mutual information, with the effect being more pronounced for larger n, since there are in that case more intermediate nodes. (Arrows are used to emphasize the amount each point shifts.)

system such that $x(0) = 0$. By definition, $W_t - W_0$ is normally distributed as $N(0, 2Dt)$, so the distance a particle travels in a time t will have the following Gaussian distribution:

$$p(x(t)) = \sqrt{\frac{1}{4\pi Dt}} \exp\left[-\frac{(x(t) - vt)^2}{4Dt}\right]. \tag{8}$$

The MSD is the expected value of $x(t)^2$ with respect to the above distribution:

$$\text{MSD}(t) = v^2t^2 + 2Dt. \tag{9}$$

Experimental measurements of MSD in biological facilitated diffusive systems are typically fit to a single power law At^α, where $1 \leq \alpha \leq 2$ is the superdiffusivity scaling exponent of the motion. To estimate the superdiffusivity exhibited by our drift-diffusion model, we must find the values of A and α that best fit the above quadratic over the average time interval of observation, which is in our case just the mean first passage time μ. To that end, we define the function

$$F(\alpha, A) = \int_0^\mu dt\left(v^2t^2 + 2Dt - At^\alpha\right)^2 \tag{10}$$

and attempt to find values for α and A such that $\nabla F = 0$. Since scaling F by a multiplicative constant will not change these values, we can rewrite the right-hand side of Eq. (10) as

$$\int_0^\mu dt\left(\omega t^2 + t - A't^\alpha\right)^2, \tag{11}$$

where we have defined $A' \equiv A/2D$ and $\omega \equiv v^2/2D$. This makes it clear that α can only depend upon the frequency ω and the time scale μ, which itself can be related to ω through the relation $\mu = \sqrt{\lambda/\omega}$. Defining the function $G_m(x)$ as

$$G_m(x) \equiv \frac{\ln(x^m) - 1}{m}, \tag{12}$$

we can express the optimal value of α implicitly as the solution of the following equation, which must be solved numerically or graphically:

$$\frac{(\lambda\omega)^{1/2}}{\alpha + 3} G_{\alpha+3}\left(\sqrt{\lambda/\omega}\right) + \frac{1}{\alpha + 2} G_{\alpha+2}\left(\sqrt{\lambda/\omega}\right) = \left(\frac{(\lambda\omega)^{1/2}}{\alpha + 3} + \frac{1}{\alpha + 2}\right) G_{2\alpha+1}\left(\sqrt{\lambda/\omega}\right). \tag{13}$$

We can set a rough upper bound on the mutual information of our cascade by making several additional assumptions. Because a delay time can only increase the conditional variance, which decreases mutual information, we shall set $\sigma_d^2 = 0$. The standard deviation in the release time interval of two molecules can be bounded above by $\sigma_S < \lambda = \ell^2/2D$, which is the average time it takes for an individual molecule to diffuse across the channel in the absence of drift. Note that because we have already assumed that the mean release time interval must be much greater than σ_S, $\sigma_S > \lambda$ would make it virtually impossible for the second molecule released to arrive before the first, rendering the molecules distinguishable by their arrival times and contradicting a basic assumption of the model. These new assumptions, along with the relation $v^2 = 2\omega D$, reduce the mutual information to the following:

$$I(\Delta t; \Delta\tau) < \frac{1}{2} \log_2 \left(1 + \frac{(\lambda\omega)^{3/2}}{2n}\right). \tag{14}$$

To evaluate the right-hand side of Eq. (14), we estimate the diffusion constant as D=5 μm^2/s and the channel length as ℓ=10 μm (the signaling molecules likely only need to cross a fraction of the cellular diameter). This results in λ=10 s. For any value of ω, we can use Eq. (13) to extract a corresponding value for α, and the upper bound in Eq. (14) can thus be plotted as a function of α for different length cascades, as shown in Fig. 3. The inset plots the critical value of the scaling exponent, $\alpha^\star$, needed for single-bit information transmission as a function of cascade length. This latter curve is exceptionally well fit by a stretched exponential function. To maintain a fixed level of informational efficiency in a molecular signaling cascade, the length of the cascade can only be increased at the cost of improving the efficiency of the molecular transport, and this cost goes down as the cascade gets progressively longer.

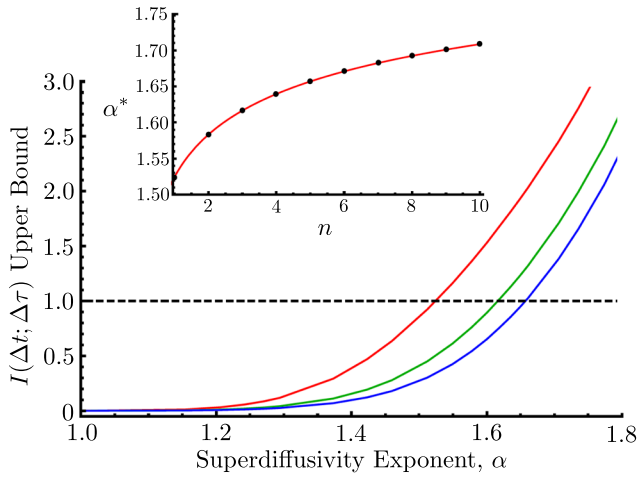

FIG. 3. The approximate upper bound on the mutual information given by Eq. (14) is plotted in bits as a function of the superdiffusivity exponent α for three different cascade lengths: $n = 1$ (red), $n = 3$ (green), and $n = 5$ (blue). A dashed line demarcates the 1-bit threshold needed to minimally distinguish two identical molecules by their arrival times. The inset plots the critical value of α needed to cross this threshold as a function of n (data points) and is fit by an extended exponential function of the form $2 - A\exp(-Bt^C)$, where A, B, and C are fitted parameters.

Although the curves in Fig. 3 can be shifted dramatically by choosing different values of the parameters in Eq. (14), our order-of-magnitude estimates give results that agree with the findings of related studies. For example, we predict that the $\alpha = 3/2$ scaling reported by A. Caspi *et al.*, which describes an enclosed microsphere transported by microtubule-walking motor proteins near the cell nucleus,[13] is just shy of single-bit mutual information. Since these predictions are based off an upper bound on the mutual information, it is likely that this level of facilitation will actually fall well short of the one-bit threshold. This is consistent with the findings of R. Suderman *et al.* that transcriptional signaling at the single-cell level, which involves the facilitated diffusion of proteins along DNA, generally transmits less than one bit of information.[4]

The model we have developed to describe information transfer in molecular signaling cascades is general enough to be applicable to a broad range of biological signaling systems that utilize pseudo-one-dimensional facilitated transport mechanisms. Tuning it towards a specific biological process only requires a small number of physical parameters such as the diffusion constant for the signaling molecules and the superdiffusivity scaling exponent of the molecular transport, both of which can be extracted from single-molecule experiments.[28] It should be cautioned that while many biological signaling processes feature highly facilitated transport, there are numerous others that involve substantial three-dimensional Brownian motion[29,30] and are better modeled as a "narrow escape" problem[31,32] rather than a 1D drift-diffusion problem. It should also be emphasized that the simplicity of our model limits it to providing only a lower bound on the degree of facilitation required for meaningful communication. If, for some suitable biological signaling process, our model predicts information transmission below one bit, then any molecular signal transmitted across that system will be indecipherable from measurements of the response. On the other hand, if our model predicts information transmission above the one-bit threshold, we cannot say with any accuracy precisely how complex a message could successfully be communicated by the system because we have derived our model under the assumption that the signal always consists of a pair of particles. These caveats aside, the model we have presented still quantifies a useful minimal performance benchmark for biological hardware that many cellular systems fail to reach. While nature has managed to function adequately despite this, thanks to the noise-averaging properties of a large cellular population,[4] the increasingly complex constructs engineered by synthetic biologists will ultimately require higher fidelity communication, and our model demonstrates at least qualitatively how such fidelity might be attained.

The work described in this document was funded under the U.S. Army Basic Research Program under PE 61102, Project T25, Task 02 "Network Science Initiative" and was managed and executed at the U.S. Army ERDC. Opinions, interpretations, conclusions, and recommendations are those of the author(s) and are not necessarily endorsed by the U.S. Army.

[1] J. Champer, A. Buchman, and O. S. Akbari, Nature Reviews Genetics **17**, 146 (2016).
[2] C. J. Bashor, N. C. Helman, S. Yan, and W. A. Lim, Science **319**, 1539 (2008).
[3] M. Heinemann and S. Panke, Bioinformatics **22**, 2790 (2006).
[4] R. Suderman, J. Bachman, A. Smith, P. Sorger, and E. Deeds, Proceedings of the National Academy of Sciences **114**, 5755 (2017).
[5] K. R. Pilkiewicz and M. L. Mayo, Physical Review E **94**, 032412 (2016).
[6] R. Cheong, A. Rhee, C. J. Wang, I. Nemenman, and A. Levchenko, Science **334**, 354 (2011).
[7] M. Zimmermann, in *Human physiology* (Springer, Berlin, Heidelberg, 1989) Chap. The nervous system in the context of information theory, pp. 166–173.
[8] P. Dzeja and A. Terzic, International Journal of Molecular Sciences **10**, 1729 (2009).
[9] A. Ishihama, Journal of Bacteriology **175**, 2483–2489 (1993).
[10] J. L. Ross, M. Y. Ali, and D. M. Warshaw, Current Opinion in Cell Biology **20**, 41 (2008).
[11] P. Hammar, P. Leroy, A. Mahmutovic, E. G. Marklund, O. G. Berg, and J. Elf, Science **336**, 1595–1598 (2012).
[12] G. H. Wadhams and J. P. Armitage, Nature Reviews Molecular Cell Biology **5**, 1024 (2004).
[13] A. Caspi, R. Granek, and M. Elbaum, Physical Review E **66**, 011916 (2002).
[14] Y. M. Wang, R. H. Austin, and E. C. Cox, Physical Review Letters **97**, 048302 (2006).
[15] T. Nakano, A. W. Eckford, and T. Haraguchi, *Molecular communication* (Cambridge University Press, 2013).
[16] S. Kadloor, R. S. Adve, and A. W. Eckford, IEEE Transactions on Nanobioscience **11**, 89 (2012).
[17] M. Pierobon and I. F. Akyildiz, IEEE Journal on Selected Areas in Communications **28**, 602 (2010).
[18] T. Nakano, M. J. Moore, F. Wei, A. V. Vasilakos, and J. Shuai, IEEE Transactions on Nanobioscience **11**, 135 (2012).
[19] K. V. Srinivas, A. W. Eckford, and R. S. Adve, IEEE Transactions on Information Theory **58**, 4678 (2012).
[20] A. W. Eckford, K. V. Srinivas, and R. S. Adve, in *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on* (IEEE, 2012) pp. 2973–2977.
[21] A. W. Eckford, in *Information Sciences and Systems, 2007. CISS'07. 41st Annual Conference on* (IEEE, 2007) pp. 160–165.
[22] H.-T. Chang, S. M. Moser, and Others, in ISIT (2012) pp. 299–303.
[23] K. Finan and P. R. Cook, *Genome Organization and Function in the Cell Nucleus*, vol. 235 (2011).
[24] I. Golding, J. Paulsson, S. M. Zawilski, and E. C. Cox, Cell **123**, 1025–1036 (2005).
[25] T. M. Cover and J. A. Thomas, *Elements of information theory* (John Wiley & Sons, 2012).
[26] K. Matsuda, "Inverse Gaussian distribution," (2005), retrieved from: http://www.maxmatsuda.com/Papers/Intro/IGDistribution.pdf.
[27] C. E. Shannon, Proceedings of the IRE **37**, 10 (1949).
[28] A. Nenninger, G. Mastroianni, and C. W. Mullineaux, Journal of Bacteriology **192**, 4535–4540 (2010).
[29] B. L. Sabatini, M. Maravall, and K. Svoboda, Current Opinion in Neurobiology **11**, 349–356 (2001).
[30] C. Di Rienzo, V. Piazza, E. Gratton, F. Beltram, and F. Cardarelli, Nature Communications **5**, 5891 (2014).
[31] Z. Schuss, A. Singer, and D. Holcman, Proceedings of the National Academy of Sciences **104**, 16098–16103 (2007).
[32] O. Bénichou and R. Voituriez, Physical Review Letters **100**, 168105 (2008).