Achievable Information Rates for Molecular Communication with Distinct Molecules

Andrew W. Eckford

Department of Computer Science and Engineering, York University
4700 Keele Street, Toronto, Ontario, Canada M3J 1P3
aeckford@yorku.ca

ABSTRACT

In molecular communication, messages are conveyed from a transmitter to a receiver by releasing a pattern of molecules at the transmitter, and allowing those molecules to propagate through a fluid medium towards a receiver. In this paper, achievable information rates are estimated for a molecular communication system when information is encoded using a set of distinct molecules, and when the molecules propagate across the medium via Brownian motion. Results are provided which indicate large gains in information rate over the case where the released molecules are indistinguishable from each other.

Keywords

Molecular communication, information theory, channel capacity

1. INTRODUCTION

There has been much interest in the field of *molecular communication* [6], in which messages are encoded as a stream of molecules in order to convey a message from a transmitter to a receiver. Spurred by recent work in nanotechnology, molecular communication is intended to mimic the form of communication used by biological organisms at the nanoscale, such as bacteria and cells within the body. Recent successes in the new field of *systems biology* [7], in which microorganisms are engineered and designed to perform specific tasks, has given the field of molecular communication greater impetus.

The means of constructing communication devices capable of performing molecular communication has been investigated in several recent papers. The earliest work was done by Weiss [11], who adapted chemical pathways in microorganisms to construct simple "circuits" that communicate with each other, such as logic gates. The features of a bacterial signaling pathway were characterized as a communication channel in [3]. Work has also been done to analyze

the suitability of intercellular signaling pathways [9] and filamental cellular structures [5] for molecular communication.

The present work is less concerned with the means of molecular communication, and more with its capabilities and limits from an information theoretic perspective. Bounds on information rates were first explored by Shannon [10], and are heavily used in traditional communication channels, giving the maximum information rate per unit cost (e.g., bits per second) that can be reliably achieved through a given channel. Although sophisticated signal processing is generally required to achieve the Shannon bounds (which is impractical for contemporary molecular communication systems), these bounds give an idea of the ultimate limits of these channels, as well as a loose idea of what can be done with basic techniques. This problem was considered in a general way in [1], and achievable information rates were calculated in [4] for channels that admit indistinguishable molecules, showing promising results. We say that a system's molecules are indistinguishable if only one kind of molecule can be transmitted; in the case where more than one kind of molecule is allowed (and these types can be distinguished by the receiver), we say the molecules are distinct.

In this paper, we extend the framework and results of [4] by considering the information rates that are possible when information is encoded using distinct molecules, an arrangement that has been assumed in previous papers (such as [6, 9]). Although there are complexity advantages, and possibly energy advantages, associated with using a communication scheme with indistinguishable molecules (as the transmitter can keep a reservoir of messenger molecules, since it need not synthesize every new molecule as it is required), we show that the potential gains in information rate from using a scheme with distinct molecules are significant.

2. MODEL

Suppose a transmitter, at the origin, and a receiver, located unit distance away, are connected by a fluid medium. The transmitter has a message $m \in \mathcal{M}$, where $\mathcal{M}$ is the set of all possible messages. The transmitter conveys m to the receiver by releasing a pattern of molecules into the fluid medium. The molecules propagate through the medium via Brownian motion. The receiver observes the arrivals of the molecules, and from the pattern of arrivals, guesses that m' was the message sent by the transmitter. If $m = m'$, the transmission is successful; if $m \neq m'$, an error is made.

We largely follow the modeling assumptions given in [4] for the discrete-time case. Let $\mathcal{X}$ represent a set of molecules that the transmitter is capable of synthesizing and releas-

ing. At each time $t \in \{0, 1, 2, \ldots\}$, the transmitter releases a single molecule $x_t \in \mathcal{X}$. Once a molecule is released, it propagates under Brownian motion in the medium, which is defined on the interval $(-\infty, 1]$ (recalling that the transmitter is located at 0, and the receiver is located at 1). We assume that the Brownian motions of the molecules released by the transmitter are independent and identically distributed; furthermore, after release, the transmitter does not influence the Brownian motion of the molecules in any way. At each time $t \in \{0, 1, 2, \ldots\}$, the receiver observes y_t; letting a_t represent the number of arrivals at the receiver during the interval t to $t + 1$, $y_t \in \mathcal{X}^{a_t}$ consists of the identities of the molecules that arrived during the tth interval. Once a molecule arrives at the receiver, it is removed from the system. We will write $\mathbf{x} = [x_0, x_1, \ldots]$ and $\mathbf{y} = [y_0, y_1, \ldots]$ to represent the sequences of channel inputs and channel outputs, respectively.

Each molecule has a *transmission time* (i.e., the first hitting time at the receiver for the Brownian motion), where n_t is the transmission time associated with the molecule released at time t. We may write

$$u_t = t + n_t$$

to represent the arrival time of the molecule released at time t (i.e., the arrival time of x_t). Thus,

$$y_t = \{x_\tau : t < u_\tau \leq t + 1, \tau \in \{0, 1, 2, \ldots\}\}. \quad (1)$$

We choose to model n_t as the first hitting time of a Wiener process, which is not the best model for a physical Brownian motion, but which has the advantage of having a PDF that can be expressed in closed form, as follows:

$$f_{N_t}(n_t) = \frac{d}{\sqrt{2\pi\sigma^2 n_t^3}} \exp\left(\frac{-d^2}{2\sigma^2 n_t}\right).$$

The remainder of our results can be easily adapted to any other PDF $f_{N_t}(n_t)$.

3. INFORMATION RATES

3.1 Mutual information

Suppose any message in $\mathcal{M}$ can be transmitted with T molecules. Since any message in $\mathcal{M}$ can be uniquely represented with $\log_2 |\mathcal{M}|$ bits, the *information rate R* is defined as

$$R = \frac{\log_2 |\mathcal{M}|}{T},$$

measured in bits per molecule (from henceforth, it will be assumed that all logarithms are base 2). Note that this is also the number of bits per unit time, since one molecule is transmitted at every time unit. As a result of the random transmission time n_t, the observations y_t in (1) are stochastic. Thus, when the receiver guesses m', there is some probability that $m \neq m'$, known as the *probability of error*, with symbol P_{err}. The goal of the system designer is to maximize R while keeping P_{err} arbitrarily small.

A message $m \in \mathcal{M}$ is transmitted over the channel by *encoding* it as a sequence $\mathbf{x} \in \mathcal{X}^T$ of T molecules. Shannon [10] showed that, as $T \to \infty$, there exist codes which allow communication at rates $R \leq I(X; Y)$ with arbitrarily low P_{err}, where $I(X; Y)$ is the *mutual information*, given by

$$I(X; Y) = \lim_{T \to \infty} \frac{1}{T} E\left[\log \frac{f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y} \mid \mathbf{x})}{f_{\mathbf{Y}}(\mathbf{y})}\right],$$

where $f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y} \mid \mathbf{x})$ and $f_{\mathbf{Y}}(\mathbf{y})$ are probability density functions, and $E[\cdot]$ represents expectation (in this case the expectation is taken over both $\mathbf{x}$ and $\mathbf{y}$).

In the case of the Brownian motion channel, it is known that neither $f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y} \mid \mathbf{x})$ nor $f_{\mathbf{Y}}(\mathbf{y})$ can be tractably calculated [4]. However, suppose that there exist tractable approximations $g(\mathbf{y} \mid \mathbf{x})$, and $g(\mathbf{y})$ for $f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y} \mid \mathbf{x})$ and $f_{\mathbf{Y}}(\mathbf{y})$, respectively, which have the following properties:

1. $\int_{\mathbf{y}} g(\mathbf{y} \mid \mathbf{x}) = 1$ and $g(\mathbf{y} \mid \mathbf{x}) \geq 0$ for all $\mathbf{x}, \mathbf{y}$ (i.e., $g(\mathbf{y} \mid \mathbf{x})$ is a valid probability density function); and

2. Given the true input distribution $f_{\mathbf{X}}(\mathbf{x})$, $g(\mathbf{y})$ is found by $\int_{\mathbf{x}} g(\mathbf{y} \mid \mathbf{x}) f_{\mathbf{X}}(\mathbf{x})$.

Then it can be shown that a *lower bound* on mutual information is given by

$$I(X; Y) \geq \lim_{T \to \infty} \frac{1}{T} E\left[\log \frac{g(\mathbf{y} \mid \mathbf{x})}{g(\mathbf{y})}\right]. \quad (2)$$

This bound was proved in [2] and elsewhere, and was previously used in [4]. Since it is relatively easy to generate instances of $\mathbf{x}$ and $\mathbf{y}$, and since $g(\mathbf{y} \mid \mathbf{x})$ and $g(\mathbf{y})$ are tractable (by assumption), the bound in (2) may be calculated by *Monte Carlo* simulation.

In fact, $g(\mathbf{y} \mid \mathbf{x})$ and $g(\mathbf{y})$ need not be close approximations of $f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})$ and $f_{\mathbf{Y}}(\mathbf{y})$, as the bound is valid for any $g(\mathbf{y}|\mathbf{x})$ and $g(\mathbf{y})$ that satisfy the given properties. However, the tightness of the bound is governed by a term related to the Kullback-Leibler distance between the approximation and the true distribution, so better approximations will likely lead to better bounds. Furthermore, the bound in (2) has the interesting physical interpretation as the best achievable rate for a decoder that assumes that $g(\mathbf{y} \mid \mathbf{x})$ is the correct distribution. As a result, the bound in (2) is an *achievable* bound, in that we could (in principle) construct a device to achieve reliable communication at the rate given by the bound.

3.2 Approximate Model

Bearing the previous section in mind, we need to find a model which provides a reasonably good, yet tractable, pair of approximations $g(\mathbf{y} \mid \mathbf{x})$ and $g(\mathbf{y})$. We take a similar approach to [4], by assuming that:

1. A particle released at time t arrives between time t and $t + 1$ with probability p_a, where

$$p_a = \int_{n_t=0}^{1} f_{n_t}(n_t).$$

If a particle does not arrive in this interval, it is assumed to be lost.

2. For each type of particle $x \in \mathcal{X}$, there is some probability of an arrival of a particle of type x from the "background". These arrivals are Poisson distributed with an identical intensity $\lambda = (1 - p_a)/|\mathcal{X}|$ for all $x \in \mathcal{X}$. Background arrivals are assumed to be independent of $\mathbf{x}$ and each other.

Assumption 1 implies that the arrivals are independent of each other, since a particle either arrives before the next particle is transmitted, or it is lost. Meanwhile, assumption 2 accounts for the "lost" particles by making them part of the background.

Let $n_x(y_t)$ represent the number of particles of type $x \in \mathcal{X}$ that are in the set y_t, i.e., the number of particles of type x that arrive at time t. Furthermore, for any integer k, let

$$\phi(k; \lambda) = \begin{cases} \frac{\lambda^k e^{-\lambda}}{k!}, & k \geq 0; \\ 0, & k < 0. \end{cases}$$

represent the Poisson PDF. As a result of the two modeling assumptions, we have that

$$g(\mathbf{y} \mid \mathbf{x}) = \prod_{t=1}^{T} g(y_t \mid x_t), \qquad (3)$$

where

$$g(y_t \mid x_t) = (1 - p_a) \prod_{u \in \mathcal{X}} \phi(n_u(y_t); \lambda)$$
$$+ p_a \phi(n_{x_t}(y_t) - 1; \lambda)) \prod_{u \in \{\mathcal{X} \setminus x_t\}} \phi(n_{x_t}(y_t); \lambda),$$

where the notation $\{\mathcal{X} \setminus x_t\}$ signifies the elements of $\mathcal{X}$ except x_t. One may then find $g(\mathbf{y})$ by marginalizing (3) over $\mathbf{x}$, in accordance with the necessary properties of $g(\mathbf{y})$.

4. RESULTS AND DISCUSSION

Results are shown in Figure 1, plotted against a log scale. In these plots, the parameters of the Wiener process are normalized with respect to the arrival probability p_a. Unsurprisingly, we see improved performance as p_a increases. Furthermore, from the figure, the achievable information rate is proportional to $\log |\mathcal{X}|$ for sufficiently large $|\mathcal{X}|$.

In [4], using indistinguishable molecules, a method was given to achieve rates of around 1.2 bits per molecule, or around 0.1 bits per unit time. Compared to results for indistinguishable particles in [4], these results represent a significant improvement. Indeed, these results suggest large increases in achievable rate may be possible by using a larger $|\mathcal{X}|$, which may be realistic in some scenarios. For example, if information is to be conveyed using carbohydrates, there are roughly 38,000 distinct trisaccharides made up of 3 hexoses; if functional group substitution is allowed, the number of distinct molecules increases into the millions [8], with corresponding increases in information capacity.

Achieving the bounds given in this paper is likely to be beyond contemporary technology at the nanoscale. However, the results give a loose idea of what may be accomplished using basic techniques. For example, supposing $|\mathcal{X}|$ is 64, and p_a is 0.3, we see from Figure 1, that a system using our simplified model for particle detection could achieve a rate of around 1 bit per particle. However, the transmitter is generating $\log_2 64 = 6$ bits per particle. Thus, the transmitter should expect to send at least six molecules to convey eacy information bit, although in practice, many more would be required in the absence of sophisticated error-correcting codes.

5. REFERENCES

[1] ALFANO, G., AND MIORANDI, D. On information transmission among nanomachines. In *Proc. 1st International Conference on Nano-Networks* (2006).

[2] ARNOLD, D. M., LOELIGER, H.-A., VONTOBEL, P. O., KAVČIĆ, A., AND ZENG, W. Simulation-based computation of information rates for channels with
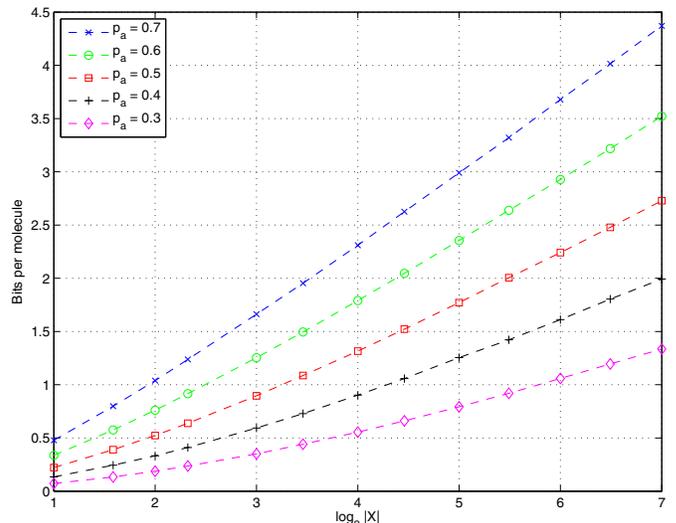
Figure 1: Mutual information per molecule (or per time unit) versus $\log_2 |\mathcal{X}|$.

memory. *IEEE Trans. Inform. Theory 52* (Aug. 2006), 3498–3508.

[3] CHU, G. System analysis of bacterial signalling. Master's thesis, University of Toronto, 2004.

[4] ECKFORD, A. W. Nanoscale communication with Brownian motion. In *Proc. 41st Annual Conference on Information Sciences and Systems (CISS)* (2007).

[5] ENOMOTO, A., MOORE, M., NAKANO, T., EGASHIRA, R., SUDA, T., KAYASUGA, A., KOJIMA, H., SAKIBARA, H., AND OIWA, K. A molecular communication system using a network of cytoskeletal filaments. In *Proc. 2006 NSTI Nanotechnology Conference* (2006), pp. 725–728.

[6] HIYAMA, S., MORITANI, Y., SUDA, T., EGASHIRA, R., ENOMOTO, A., MOORE, M., AND NAKANO, T. Molecular communication. In *Proc. 2005 NSTI Nanotechnology Conference* (2005), pp. 391–394.

[7] KITANO, H. Systems biology: A brief overview. *Science 295* (1 Mar. 2002), 1662–1664.

[8] LAINE, R. A. Information capacity of the carbohydrate code. *Pure and Applied Chemistry 69* (1997), 1867–1873.

[9] NAKANO, T., SUDA, T., MOORE, M., EGASHIRA, R., ENOMOTO, A., AND ARIMA, K. Molecular communication for nanomachines using intercellular calcium signalling. In *Proc. 5th IEEE Conference on Nanotechnology* (2005), pp. 478–481.

[10] SHANNON, C. E. A mathematical theory of communication. *Bell System Technical Journal* (Jul. 1948), 379–423.

[11] WEISS, R., BASU, S., HOOSHANGI, S., KALMBACH, A., KARIG, D., MEHREJA, R., AND NETRAVALI, I. Genetic circuit building blocks for cellular computations, communications, and signal processing. *Natural Computing 2* (Mar. 2003), 47–84.