A Diffusion-Based Binary Digital Communication System

Ling-San Meng^{*}, Ping-Cheng Yeh^{*}, Kwang-Cheng Chen^{*}, and Ian F. Akyildiz[†] *Communications and Networks Science (CoWorkS) Lab Graduate Institute of Communication Engineering and Department of Electrical Engineering National Taiwan University, Taipei, Taiwan Email: f95942117@ntu.edu.tw, {pcyeh, chenkc}@cc.ee.ntu.edu.tw [†]Broadband Wireless Networking Laboratory, School of Electrical and Computer Engineering Georgia Institute of Technology, Atlanta, GA, 30332, USA Email: ian@ece.gatech.edu

Abstract-Diffusion-based communications refers to the transfer of information using particles as message carriers whose propagation is based on the law of particle diffusion. Though still at an early stage, there have been growing interests and research efforts dedicated to this communication technology. It has been identified that diffusion-based communications is one of the most promising approaches for end-to-end communication between nanoscale devices in the near future. In this paper, the design of a binary digital communication system is proposed based on particle diffusion. Stochastic signaling through On-Off Keying (OOK) for random particle emission and a diffusion channel with memory is considered. The diffusion is considered in the cases of one, two, and three dimensions. The receiver detection problem is formulated by using an information-theoretic approach. The optimal decision threshold for the receiver detection is derived through mutual information maximization for two cases, namely, when the *a priori* probability of bit transmission is fixed and known to the receiver and when this probability is unknown to the receiver. Numerical results indicate that in the case of diffusion in one or two dimensions, the information of a priori probability plays a key role in optimizing the system performance, while it does not when considering the diffusion in three dimensions.

Index Terms—Digital communications, Diffusion process, Onoff keying, Neyman-Pearson criterion, Mutual information, Channel capacity

I. INTRODUCTION

In recent years, there has been dramatic progress made in the development of nanotechnology, which is defined as the technology involving components in a scale of 1 nm to 100 nm [1]. A nanomachine is a device which is envisioned to perform a specific task, e.g., processing, sensing and actuation, and can be considered as the most basic unit [2]. The functionality and capability of one nanomachine alone are quite limited. The idea of interconnecting several nanomachines to form a *nanonetwork* has been proposed and recently studied in [3].

Four different communication mechanisms for nanomachines have been considered and proposed so far, i.e., mechanical, acoustic, electromagnetic, and molecular communications [1]. However, due to the constraints of size, power, and complexity associated with the nanoscale regime, many of the options listed above have been identified as not directly applicable. Amongst others, molecular communication, defined as the transfer of information using particles as message carriers, is considered one of the most promising. The particles can either follow a specific path or be guided by a fluidic medium to reach the destination [1]. *Diffusion-based communication* refers to the situation where molecules reach the destination relying solely on the laws of particle diffusion. For example, pheromone propagation in the air between insects [4] or calcium signaling among living cells [5] fall into the category of diffusion-based communications.

1

Though still at an early stage, there have been growing interests and research efforts dedicated to diffusion-based communications. Theoretical studies which try to provide a realistic mathematical model of diffusion-based communication systems and characterize the fundamental limit of the information transfer rate are at the cutting edge of nowadays research. In [6]-[8], a mathematical end-to-end model along with the noise analysis is proposed for diffusion-based communication systems by assuming continuous particle emission at the molecular transmitter. In [9], [10], an information theoretical framework for diffusion-based communication systems is proposed. The channel capacity is obtained by assuming a chemical reaction setup between nanomachines and by utilizing the principles of mass action kinetics. In addition to the theoretical studies, preliminary laboratory experiments are also performed in an attempt to create and analyze biologically engineered diffusion-based communication systems, e.g., [5], [11]-[13].

In this paper, we propose the design of a diffusion-based communication system for transmission of binary digital information. We consider a time-slotted system with stochastic signaling and a diffusion channel with memory. The diffusion is considered in the cases of one, two, and three dimensions. On-Off Keying (OOK) is adopted at the transmitter for molecular emission, where molecules are released instantaneously by following a fixed probability distribution, which is assumed to be known by the receiver. The effect of Inter-Symbol Interference (ISI) from the residual particle diffusion of previous signaling intervals is considered. Two different detection schemes are formulated based on the availability of the *a priori* information by using an information-theoretic approach. The optimal decision threshold is derived for mutual

information maximization. Numerical results indicate that in the case of diffusion in one or two dimensions, the information of *a priori* probability plays a key role in optimizing the system performance, while it does not when considering the diffusion in three dimensions.

The rest of this paper is organized as follows. In Section II, we introduce the system model of the diffusion-based communication system. In Section III, we develop the receiver detection scheme using an information-theoretic approach. The achievable mutual information of the proposed system is obtained for both cases of perfect and no knowledge of prior information. In Section IV, the numerical results are presented. Finally, conclusions are given in Section V.

II. SYSTEM MODEL

We propose a time-slotted system with signaling interval T_s . In this paper, we assume perfect synchronization between the transmitter and the receiver. Let X_i and Y_i denote the input and output random variables of the *i*th signaling interval, respectively. OOK with stochastic signaling is considered as the modulation technique. With a priori probability p_1 , a random number of molecules is emitted in an instantaneous fashion by the transmitter at the beginning of each signaling interval to signify 1; no molecule is emitted to signify 0. It is assumed that the number of molecules present in the system is large enough such that the differential equations can be applied to describe the dynamics. Let Q_i denote the number of molecules emitted at the *i*th signaling interval; we have $Q_i \gg 1$. We assume that $\{Q_i\}$ is a series of independent and identically distributed (i.i.d.) continuous random variables with finite mean and variance, which are denoted by μ_Q and σ_Q^2 , respectively, and are assumed to be known by the receiver through proper estimation techniques. Once released into the propagation medium, the molecules are assumed to diffuse freely, and the dynamics is described by the Brownian motion.

Let the molecule source be located at the origin of a Cartesian coordinate, and the center of the receiver is located at \vec{r} . Let $\phi(\vec{r}, t)$ denote the concentration function of molecules at the location of the receiver \vec{r} and time t. Fick's second law of diffusion [14] predicts how such concentration function changes with time:

$$\frac{\partial \phi(\vec{r},t)}{\partial t} = D\nabla^2 \phi(\vec{r},t), \qquad (1)$$

where D is the diffusion constant which is related to the viscosity of the propagation medium. Due to the uncertainty lying in the previous signaling intervals, ISI occurs as a result of residual particle diffusion. We denote the number of interfering signaling intervals by N. It can be shown that the solution to (1) with a point source is of the form [15]

$$\phi(r,t) = \sum_{j=i-N}^{i} X_j \frac{Q_j}{\left(4\pi D(t-jT_s)\right)^{\frac{d}{2}}} \exp\left(-\frac{r^2}{4D(t-jT_s)}\right),$$
(2)

where d denotes the number of dimensions of the space; $d \in \{1, 2, 3\}$; $t \in [iT_s, (i+1)T_s)$. Note that in (2) we have omitted the vector notation due to the isotropy of a point molecule source with free Brownian motion.



Fig. 1. Representation of the diffusion channel as a binary hypothesis testing channel.

III. DETECTION SCHEMES

In this section, we apply an information-theoretic approach for the detector design, since there is no assumption of any favorable *a priori* distribution for binary signaling or a proper definition of Bayesian cost of a diffusion-based communication channel. In Fig. 1, a binary hypothesis testing channel representation of the diffusion channel is depicted, where P_F and P_D denote the *false alarm probability* and the *detection probability*, respectively. By definition, we have

$$P(Y_{i} = 1 | X_{i} = 0) = P_{F},$$

$$P(Y_{i} = 1 | X_{i} = 1) = P_{D},$$

$$P(Y_{i} = 0 | X_{i} = 0) = 1 - P_{F},$$

$$P(Y_{i} = 0 | X_{i} = 1) = 1 - P_{D}.$$
 (3)

It is well-known that

$$I_{X_i;Y_i} = \sum_{X_i=0}^{1} \sum_{Y_i=0}^{1} P(Y_i|X_i) P(X_i) \log \frac{P(Y_i|X_i)}{P(Y_i)}.$$
 (4)

We can thus represent the mutual information as a function of the probabilities P_F, P_D , and p_1 . For ease of notation, we omit the subscripts and denote the mutual information by $I(P_F, P_D, p_1)$.

A. Detection with Perfect A Priori Information

For the case where the receiver has perfect knowledge but no control over p_1 , the detector design is concerned with the following optimization problem

$$\max_{P_F, P_D} I(P_F, P_D, p_1).$$
(5)

Let Z_i denote the random variable which is defined as the integral of the concentration function at the receiver from iT_s to $(i+1)T_s$, i.e., $Z_i = \int_{iT_s}^{(i+1)T_s} \phi(r,t) dt$. Here we consider the binary hypothesis testing problem with continuous-type observations Z_i :

$$H_1$$
: molecules are emitted at $t = iT_s$; $Z_i \sim f_{Z_i}^1(z)$,
 H_0 : otherwise; $Z_i \sim f_{Z_i}^0(z)$, (6)

where $f_{Z_i}^k(z)$ denotes the probability density function of Z_i given that H_k is true. By viewing the interfering input bits as Bernoulli random variables with success probability p_1 , we

have that Z_i is a weighted sum of independent variables:

$$H_{1}: Z_{i} = a_{0}Q_{i} + \sum_{j=i-N}^{i-1} a_{i-j}X_{j}Q_{j},$$
$$H_{0}: Z_{i} = \sum_{j=i-N}^{i-1} a_{i-j}X_{j}Q_{j},$$
(7)

where we have defined

$$a_j = \int_0^{T_s} \frac{1}{(4\pi D(t+jT_s))^{\frac{d}{2}}} \exp\left(-\frac{r^2}{4D(t+jT_s)}\right) dt.$$
(8)

In (7), we have formulated the diffusion channel into an ISI channel with memory N. It is assumed that the series of independent random variables, $\{a_{i-j}X_jQ_j\}$, satisfies the Lindeberg's condition [16] so that Z_i converges to the Gaussian distribution as N approaches infinity¹. By applying the Gaussian approximation, the binary hypothesis testing problem can be written as

$$H_{1}: \ z \sim \mathcal{N}(\mu_{Z_{1}}, \sigma_{Z_{1}}^{2}),$$

$$H_{0}: \ z \sim \mathcal{N}(\mu_{Z_{0}}, \sigma_{Z_{0}}^{2}), \tag{9}$$

where it is straightforward that

$$\mu_{Z_0} = p_1 \mu_Q \sum_{j=1}^N a_j,$$

$$\mu_{Z_1} = a_0 \mu_Q + \mu_{Z_0},$$

$$\sigma_{Z_0}^2 = \left(p_1^2 \sigma_Q^2 + \mu_Q^2 (p_1 - p_1^2) + \sigma_Q^2 (p_1 - p_1^2) \right) \sum_{j=1}^N a_j^2,$$

$$\sigma_{Z_1}^2 = a_0^2 \sigma_Q^2 + \sigma_{Z_0}^2.$$
(10)

It can be shown that given P_F , the mutual information is a monotonically increasing function of P_D [17], [18]. Thus an information-optimal detector is equivalent to a Neyman-Pearson detector when P_F is given. The Neyman-Pearson decision criterion states that such a constrained optimization problem is solved by forming the likelihood ratio test [19]

$$\frac{f_{Z_i}^1(z)}{f_{Z_i}^0(z)} = \Lambda(z) \stackrel{H_1}{\underset{H_0}{\gtrless}} \lambda, \tag{11}$$

where λ is found by solving $P(\Lambda(z) > \lambda \mid H_0) = P_F$. Using (9), the likelihood ratio function can be derived as

$$\Lambda(z) = \frac{\frac{1}{\sqrt{2\pi\sigma_{Z_1}^2}} e^{-\frac{(z-\mu_{Z_1})^2}{2\sigma_{Z_1}^2}}}{\frac{1}{\sqrt{2\pi\sigma_{Z_0}^2}} e^{-\frac{(z-\mu_{Z_0})^2}{2\sigma_{Z_0}^2}}} = \frac{\sigma_{Z_0}}{\sigma_{Z_1}} e^{\frac{(\sigma_{Z_1}^2 - \sigma_{Z_0}^2)^{z^2 - 2(\mu_{Z_0}\sigma_{Z_1}^2 - \mu_{Z_1}\sigma_{Z_0}^2)^{z + \mu_{Z_0}^2}\sigma_{Z_1}^2 - \mu_{Z_1}^2\sigma_{Z_0}^2}}(12)$$

¹Loosely speaking, the Lindeberg's condition requires that all random variables are independent, and each one of them contributes a vanishing part to the total variance as N approaches infinity.

Combining (11) and (12) and taking the natural logarithm at both sides, we have

$$\begin{pmatrix} \sigma_{Z_1}^2 - \sigma_{Z_0}^2 \rangle z^2 - 2(\mu_{Z_0}\sigma_{Z_1}^2 - \mu_{Z_1}\sigma_{Z_0}^2) z + \mu_{Z_0}^2\sigma_{Z_1}^2 - \mu_{Z_1}^2\sigma_{Z_0}^2 \\ \stackrel{H_1}{\gtrless} 2\sigma_{Z_0}^2\sigma_{Z_1}^2 \left(\ln\lambda - \ln\frac{\sigma_{Z_0}}{\sigma_{Z_1}}\right).$$
(13)

Using elementary algebra, we can solve (13) and obtain a simplified test statistic of z under the condition

$$\frac{\mu_{Z_0}^2}{2\sigma_{Z_0}^2} - \frac{\mu_{Z_1}^2}{2\sigma_{Z_1}^2} + \ln\frac{\sigma_{Z_0}}{\sigma_{Z_1}} < \ln\lambda.$$
(14)

The simplified test statistic of z is put in (15). It then follows

$$P_F = \int_{\eta}^{\infty} f_{Z_i}^0(z) \, dz = Q\left(\frac{\eta - \mu_{Z_0}}{\sigma_{Z_0}}\right), \qquad (16)$$

and the corresponding detection probability is

$$P_{D} = \int_{\eta}^{\infty} f_{Z_{i}}^{1}(z) \, dz = Q\left(\frac{\eta - \mu_{Z_{1}}}{\sigma_{Z_{1}}}\right).$$
(17)

It remains to find the optimal threshold $\eta^*(p_1)$ which yields the maximum mutual information. This is a numerical problem

$$\eta^*(p_1) = \underset{\eta}{\operatorname{argmax}} I(P_F(\eta), P_D(\eta), p_1).$$
(18)

In the case where the receiver has control over the *a priori* probability, e.g., by affecting the coding scheme, it is straightforward that the optimal value is determined such that the mutual information is further maximized over all possible values of p_1 as

$$p_1^* = \underset{p_1}{\operatorname{argmax}} I(P_F(\eta^*(p_1)), P_D(\eta^*(p_1)), p_1).$$
(19)

Note that the corresponding maximum mutual information $I(P_F(\eta^*(p_1^*)), P_D(\eta^*(p_1^*)), p_1^*)$ represents the theoretically maximum throughput of the considered diffusion-based communication system instead of the channel capacity, since a specific modulation technique and channel observation are involved.

B. Detection with No A Priori Information

In the case where the receiver has no information of the *a* priori probability at all, the concept of minimax [19] which tries to mitigate the worst possible situation should be applied. However, as discussed in [18], the fact that $I(P_F, P_D, p_1) = 0$ when $p_1 = 0$ or 1 renders the approach of minimax inappropriate. Alternatively, we propose the use of a decision threshold which is optimized at the *a priori* probability p_1^{\dagger} such that

$$p_1^{\dagger} = \underset{p_1}{\operatorname{argmax}} \int I(P_F(\eta^*(p_1)), P_D(\eta^*(p_1)), p) \, dp.$$
 (20)

The decision threshold $\eta^*(p_1^{\dagger})$ thus maximizes the integrated information amount independent of the actual *a priori* probability.



Fig. 2. Attainable mutual information for short-range diffusion-based communications with perfect prior information.



Fig. 3. Attainable mutual information for long-range diffusion-based communications with perfect prior information.

IV. NUMERICAL RESULTS

In this section we present numerical results for the attainable mutual information of the proposed diffusion-based communication system. Both short-range and long-range communication scenarios are investigated. Short-range molecular communications happens naturally as the mechanism for biochemical signaling in living cells, e.g., calcium ion signaling and neural signaling; while long-range molecular communications mostly serves as the signaling method among living organisms, e.g., pheromone propagation and the dispersal of pollen and spores [4]. In the following, we set $D = 10^{-6}$ cm²/s, $r = 20 \ \mu$ m, and $\mu_Q = 1$ mol, for short-range molecular communications. For the case of long range, we set D = 0.43 cm²/s, r = 2 cm, and $\mu_Q = 1000,000$ mol. For both scenarios we set N = 20, and

Fig. 4. Attainable mutual information for short-range diffusion-based communications both with and without knowledge of prior information.



Fig. 5. Attainable mutual information for short-range diffusion-based communications in different dimensions; $T_s = 10$ s.

 $\sigma_Q = 0.3 \,\mu_Q$, which gives a medium coefficient of variation (CV) of 0.3. Unless otherwise stated, we set d = 1, i.e., one-dimension diffusion.

In Fig. 2 and Fig. 3, we plot the maximum attainable mutual information for the diffusion-based communication system for the cases of short range and long range, respectively. Perfect prior information is assumed to be known by the receiver. Truncations can be seen at p_1 close to 1 in both figures due to the constraint on the minimum threshold given in (14). Since the curves have already covered most operating regions of interest, we omit the discussion if (14) is not satisfied, which requires much more complicated forms of P_F and P_D . The circles indicate the point which achieves the channel capacity of the system as given in (19). It is observed that

the higher mutual information is achieved in the short-range case for the same signaling interval due to less effect of ISI, as one would expect. It is also observed that the achievable mutual information increases along with T_s due to the same reason. Though the results indicate that the system achieves the channel capacity at $p_1 < \frac{1}{2}$, which is dependent on the level of ISI experienced, we see rather flat curves over a wide region of p_1 . It is thus concluded for one-dimension diffusion, an exact control over the *a priori* probability is not vital for achieving a desirable amount of information transfer. In this case, *a priori* probability around $0.2 \sim 0.6$ can achieve over 95% of the channel capacity.

In Fig. 4, we compare the maximum attainable mutual information (perfect prior information) with the other extreme where no prior information is available for short-range communications. The squares correspond to the values of prior probabilities which give the maximum amount of information as defined in (20). We observe that in general the value of p_1 at which the decision threshold should be optimized when no prior information is available is relatively insensitive to the level of ISI experienced. In this case, we have $p_1^{\dagger} \simeq 0.45$. The results suggest that the system entails huge performance loss when there is a mismatch between a presumed prior probability and the actual prior probability. This is expected since the distribution of ISI is dependent on p_1 as shown in (10). We thus conclude that the exact knowledge of the prior information at the receiver is important for the system to operate satisfactorily.

In Fig. 5, we plot the attainable mutual information both with and without prior information for different dimensions; d = 1, 2, and 3. The signaling interval is set to 10 seconds. It is observed that the higher mutual information can be achieved at higher dimensions. This results from the fact that the dimension serves as the exponent affecting the decaying rate of the concentration function with time, which in turn lowers the effect ISI. Specifically, when d = 3, we observe that a channel capacity of 1 bit per channel use is almost attained, and the two corresponding curves nearly overlap. This suggests that knowledge of prior information can offer only little performance improvement in terms of mutual information in this case. On the other hand, however, control of the *a priori* probability to be around 0.5 has a major effect on optimizing the system performance.

V. CONCLUSIONS

In this paper, we have proposed a diffusion-based communication system for transmission of binary digital information. At the transmitter, we have considered OOK with stochastic signaling, where the number of particle emitted per symbol is random with known probability. The detector has been designed following an information-theoretic approach. The mutual information between the channel input and the decision is optimized for both cases where the *a priori* probability is perfectly known or completely unavailable to the receiver. Numerical results indicate that in the case of diffusion in one or two dimensions, the information of *a priori* probability plays a key role in optimizing the system performance, while it does not when considering the diffusion in three dimensions. By proposing and analyzing the diffusion-based communication system, our purpose is to establish a framework under which different signaling methods and receiver structure can be studied. A well-defined noise model will also be included in our future work.

ACKNOWLEDGEMENT

This work was supported by the National Science Council (NSC) under Grant No. 100-2917-I-002-044. The authors would like to thank Josep Miquel Jornet Montana and Massimiliano Pierobon for their valuable comments that improved the quality of this paper.

REFERENCES

- I. F. Akyildiz, F. Brunetti, and C. Blázquez, "Nanonetworks: A new communication paradigm," *Computer Networks (Elsevier) Journal*, vol. 52, pp. 2260–2279, Aug. 2008.
- [2] T. Suda, M. Moore, T. Nakano, R. Egashira, and A. Enomoto, "Exploratory research on molecular communication between nanomachines," in *Proc. Genetic and Evolutionary Computation Conference* (*GECC005*), June 2005.
- [3] M. S. Islam and L. VJ, "Nanoscale materials and devices for future communication networks," *IEEE Commun. Mag.*, vol. 48, pp. 112–120, June 2010.
- [4] L. P. Giné and I. F. Akyildiz, "Molecular communication options for long range nanonetworks," *Computer Networks*, vol. 53, pp. 2753–2766, Nov. 2009.
- [5] T. Nakano, T. Suda, M. Moore, R. Egashira, A. Enomoto, and K. Arima, "Molecular communication for nanomachines using intercellular calcium signalling," in *Proc. 5th IEEE Conference on Nanotechnology*, vol. 2, July 2005, pp. 478–481.
- [6] M. Pierobon and I. F. Akyildiz, "A physical end-to-end model for molecular communication in nanonetworks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 4, pp. 602–611, May 2010.
- [7] —, "Diffusion-based noise analysis for molecular communication in nanonetworks," *IEEE Trans. Signal Processing*, vol. 59, pp. 2532–2547, June 2011.
- [8] —, "Noise analysis in ligand-binding reception for molecular communication in nanonetworks," *IEEE Trans. Signal Processing*, vol. 59, pp. 4168–4182, Sept. 2011.
- [9] B. Atakan and O. Akan, "An information theoretical approach for molecular communication," *Bionetics 2007. 2nd*, pp. 33–40, Dec. 2007.
- [10] B. Atakan and O. B. Akan, "On molecular multiple-access, broadcast, and relay channels in nanonetworks," in *Proc. ICST/ACM Conference BIONETICS 2008*, Nov. 2008.
- [11] R. Weiss, S. Basu, S. Hooshangi, A. Kalmbach, D. Karig, R. Mehreja, and I. Netravali, "Genetic circuit building blocks for cellular computations, communications, and signal processing," *Natural Computing*, vol. 2, pp. 47–84, Mar. 2003.
- [12] R. Weiss and T. Knight, "Engineered communications for microbial robotics," in *Proc. 6th International Meeting on DNA Based Computers*, 2000.
- [13] T. Nakano, T. Suda, T. Kojuin, T. Haraguchi, and Y. Hiraoka, "Molecular communication through gap junction channels: System design, experiments and modeling," in *Proc. 2nd International Conference on Bio-Inspired Models of Network, Information, and Computing Systems*, Dec. 2007, pp. 139–146.
- [14] J. Philibert, "One and a half century of diffusion: Fick, Einstein, before and beyond," *Diffusion Fundamentals*, 2005.
- [15] A. Einstein, "Investigations of the theory of Brownian movement," Dover, 1956.
- [16] R. B. Ash, Probability and Measure Theory, 2nd ed. Academic Press, Dec. 1999.
- [17] D. Middleton, An Introduction to Statistical Communication Theory. New York: McGraw-Hill, 1960.
- [18] T. Gabriele, "Information criteria for threshold determination (corresp.)," *IEEE Trans. Inform. Theory*, vol. 12, no. 4, pp. 484 – 486, Oct. 1966.
- [19] H. V. Poor, An Introduction to Signal Detection and Estimation, 2nd ed. Springer, Mar. 1994.