On Receiver Design for Diffusion-Based Molecular Communication

Ling-San Meng, *Student Member, IEEE*, Ping-Cheng Yeh, *Member, IEEE*, Kwang-Cheng Chen, *Fellow, IEEE*, and Ian F. Akyildiz, *Fellow, IEEE*

***Abstract*—Diffusion-based communication refers to the transfer of information using molecules as message carriers whose propagation is governed by the laws of molecular diffusion. It has been identified that diffusion-based communication is one of the most promising solutions for end-to-end communication between nanoscale devices. In this paper, the design of a diffusion-based communication system considering stochastic signaling, arbitrary orders of channel memory, and noisy reception is proposed. The diffusion in the cases of one, two, and three dimensions are all considered. Three signal processing techniques for the molecular concentration with low computational complexity are proposed. For the detector design, both a low-complexity one-shot optimal detector for mutual information maximization and a near Maximum Likelihood (ML) sequence detector are proposed. To the best of our knowledge, our paper is the first that gives an analytical treatment of the signal processing, estimation, and detection problems for diffusion-based communication in the presence of ISI and reception noise. Numerical results indicate that the proposed signal processing technique followed by the one-shot detector achieves near-optimal throughput without the need of *a priori* information in both short-range and long-range diffusion-based communication scenarios, which suggests an ML sequence detector is not necessary. Furthermore, the proposed receiver design guarantees diffusion-based communication to operate without failure even in the case of infinite channel memory. A channel capacity of 1 bit per channel utilization can be ultimately achieved by extending the duration of the signaling interval.**

***Index Terms*—Channel capacity, diffusion process, intersymbol interference, mutual information, Neyman–Pearson criterion, on-off keying, parameter estimation, Viterbi algorithm.**

I. Introduction

D*IFFUSION-BASED communication* refers to the technology where the transportation of information is achieved by the propagation of molecules relying solely on the laws of molecular diffusion [1]. Diffusion-based communication arises as one of the most promising solutions for the communication mechanism between nanoscale devices for its inherent compatibility with living organism and biochemical devices, e.g., pheromone propagation in the air between insects [2] or calcium signaling among living cells [3].

Manuscript received September 22, 2013; revised April 16, 2014 and August 14, 2014; accepted September 05, 2014. Date of publication September 22, 2014; date of current version October 24, 2014. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Rong-Rong Chen. This work was supported by the National Science Council (NSC) under Grant No. 103-2221-E-002-090, and Ministry of Education, Taiwan, under Grant No. 103R890846.

L.-S. Meng, P.-C. Yeh, and K.-C. Chen are with the Graduate Institute of Communication Engineering, National Taiwan University, Taipei 10617, Taiwan (e-mail: f95942117@ntu.edu.tw; pcyeh@ntu.edu.tw; ckc@ntu.edu.tw).

I. F. Akyildiz is with the Broadband Wireless Networking Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: ian@ece.gatech.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2014.2359644

Similar to traditional ElectroMagnetic (EM) communication, diffusion-based communication can be categorized as being *analog* or *digital*. In analog diffusion-based communication, the intensity of the molecular concentration is varied by the transmitter in a controlled manner following the analog waveform to be transmitted. The receiver then recovers the transmitted waveform by continuously detecting the molecular concentration in its neighborhood [4]. In digital diffusion-based communication, the transmitter sends digitized information, e.g., binary information, by altering a certain attribute of the molecules in discrete signaling intervals. For example, the timing of emitting molecules [5], [6], the intensity of the molecular concentration (or the number of molecules) [7]–[14] and the type of molecules [14]–[17] are three attributes commonly considered for conveying digital information. Due to the inherent nature of discreteness of molecules and past successful experience in developing digital EM communication systems, diffusion-based communication using digital signaling are drawing much more attention than the other.

In digital diffusion-based communication, the effect of channel memory, hence Inter-Symbol Interference (ISI), arises naturally from the residual molecular diffusion from previous symbol transmissions. It has been shown that the effect of ISI is critically important in constructing a reliable communication system based on molecular diffusion [9], [11]. However, simplified transmitter model, channel model, signal processing techniques, and receiver detection schemes without mathematical foundations are generally assumed in the literature for making the analysis tractable. For example, transmission of digital information using a single molecule is considered in [10], [18]. Modulation techniques for diffusion-based communication systems are studied via computer simulations in [5], [6], and it has been concluded that On-Off Keying (OOK) outperforms Pulse Position Modulation (PPM) [5]. The detection scheme is, however, not covered in the discussion. A diffusion channel with limited order of memory is assumed in [8], [14], [19]. In [10], stochastic degradation of the information molecules is considered in simulations. A deterministic lifetime of molecules, hence limited channel memory, is still assumed for analysis. The concepts of sampling-based [12], [19] or energy-based [11], [12] signal processing of the molecular concentration have been proposed without mathematical analysis.

For the decision rules, the concept of using detection thresholds on the molecular concentration is proposed in [8], [12], [14], [19]. However, none of the decision rules proposed in the literature analytically considers the effect of ISI in the formulation of the detection threshold. Neither has the existing literature considered the effect of reception noise when formulating decision rules. In realistic environments where the reception noise and large channel memory are present, such simplified system setup and detection approaches raise the concern of the feasibility of diffusion-based communication systems.

In [20], the generation of molecules and the arrival times are modeled using a stochastic approach. The probability distribution of information molecules at the receiver for both flow-based and diffusion-based molecular communications are both derived. However, the detection scheme and the receiver design are not addressed. The construction of a communication system based on molecular communication is thus still incomplete. A closed-form expression for the achievable channel capacity of the diffusion-based communication is derived in [21]. The results are quite insightful, but the study mainly focuses on the information-theoretic aspect of the diffusion-based communication. A practical communication system based on molecular diffusion is still left to desire.

In this paper, we propose the design of a diffusion-based communication system for transmission of binary digital information. We consider a communication system which consists of stochastic signaling, a diffusion channel with arbitrary order of memory, and noisy reception with standard signal estimation and detection theory. The diffusion in the cases of one, two, and three dimensions are all considered, respectively. OOK is adopted at the transmitter for molecular emission, where molecules are released in an instantaneous manner with random amount. The effect of ISI from the residual molecular diffusion in previous signaling intervals is considered and analytically incorporated in the receiver design. Concerning the capability of nanoscale devices, we propose and analyze three signal processing techniques on the molecular concentration with low computational complexity. Most important, a low-complexity ISI cancellation technique is proposed. For the detector design, both a one-shot optimal detector for mutual information maximization and a near Maximum Likelihood (ML) sequence detector using the Reduced-State Viterbi (RS-Viterbi) algorithm [22] are proposed. An asymptotic ML estimator for estimating the variance of the number of emitted molecules is also proposed. To the best of our knowledge, we are the first in the literature to give an analytical treatment of the signal processing, estimation, and detection problems for diffusion-based communication in the presence of ISI and reception noise. Numerical results indicate that the proposed ISI cancellation processing followed by the one-shot detector can achieve near-optimal throughput without the need of *a priori* information in both short-range and long-range diffusion-based communication scenarios. An ML sequence estimation scheme is found to be unnecessary. Furthermore, it is shown that our receiver design guarantees diffusion-based communication to operate without failure even in the case of infinite channel memory. A channel capacity of 1 bit per channel utilization can be ultimately achieved by extending the duration of the signaling interval.

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 propose three signal processing techniques for the received molecular concentration. In Section IV, we propose a one-shot detection scheme which only utilizes information in the corresponding signaling interval. The detection threshold for mutual information maximization is derived for both cases of perfect and no knowledge of *a priori* information. In Section V, a sequence detection scheme using the RS-Viterbi algorithm is proposed. In Section VI, we propose an asymptotic ML estimator for estimating the randomness of the molecular emission at the transmitter. In Section VII, the numerical results are presented. Finally, conclusions are given in Section VIII.

II. Model Description

We propose a time-slotted system with signaling interval T_s. In this work, we assume perfect synchronization between the transmitter and the receiver. Let X_i denote the input binary random variables in the ith signaling interval. OOK with stochastic signaling is considered as the modulation technique. With *a priori* probability p, the transmitter signifies 1 by emitting a number of molecules at the beginning of a signaling interval; no molecule is emitted to signify 0. The number of molecules emitted by the transmitter is considered to be large enough so that differential equations can be applied to describe the macroscopic behavior of the molecules. The molecules are assumed to diffuse freely following the Brownian motion without drift and interactions.

Let Q_i denote the number of molecules emitted in the ith signaling interval. We have $Q_i \gg 1$. Ideally, the transmitter would target at emitting a fixed number of molecules for each binary signaling. However, due to the random nature of biochemical phenomena or the imperfect design of man-made systems, it is more practical to consider the transmitter to exhibit a certain noisy behavior when emitting the molecules. Such effect is taken into account by assuming $\{Q_i\}$ as a sequence of independent and identically distributed (i.i.d.) continuous random variables with finite mean and variance, denoted by μ_Q and σ_Q^2, respectively. In this work, we put our emphasis on the processing of the molecules and the detector design by assuming perfect knowledge of the parameters μ_Q and σ_Q^2 at the receiver side. An asymptotic ML estimator for estimating the variance σ_Q^2 is proposed and analyzed in Section VI to provide insight into the parameter estimation.

Let the molecule source be located at the origin of a Cartesian coordinate, and the center of the receiver is located at $\vec{r}$. Fick's second law of diffusion [23] predicts how the concentration function changes with time:

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

where $\phi_i(\vec{r},t)$ is defined as the molecular concentration function at the receiver at time t corresponding to the ith signaling interval, and ∇ denotes the gradient operator. It can be shown

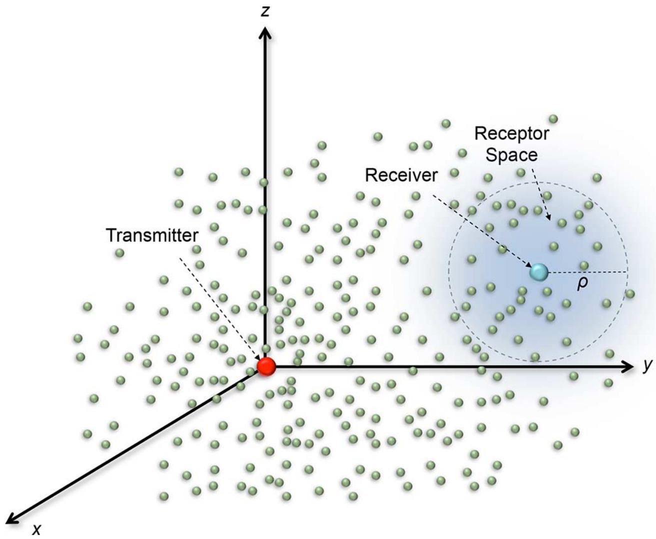

Fig. 1. Illustration of the proposed molecular communication system in a three-dimensional space. The transmitter and the receiver are located at the origin and $\vec{r}$, respectively. The receiver performs sensing of the molecular concentration inside the receptor space with radius denoted by ρ.

that the solution to Fick's second law in response to an impulse of molecule emission Q_i is of the form: [24], [25]

$$Q_i \frac{1}{(4\pi Dt)^{\frac{d}{2}}} \exp\left(-\frac{r^2}{4Dt}\right), \tag{2}$$

where D is the diffusion constant which is related to the viscosity of the propagation medium, and d denotes the number of dimensions; $d \in \{1,2,3\}$. For better readability, we henceforth define $h(t) = \frac{1}{(4\pi Dt)^{\frac{d}{2}}} \exp(-\frac{r^2}{4Dt})$, and (2) can be rewritten as $Q_i h(t)$. Note that Fick's second law of diffusion describes the macroscopic behavior of molecules. With the assumption of a sufficiently large number of molecules, the molecular concentration can be approximated as a deterministic function of time and space. In the following, we also treat $h(t)$ as a deterministic function, and this should not cause any confusion since the underlying assumption of a large number of molecules still holds. Due to the previous transmissions, ISI occurs as a result of residual particle diffusion. The number of interfering signaling intervals is henceforth denoted by N. Since (1) is a linear equation, we can write

$$\phi_i(r,t) = \sum_{j=i-N}^{i} X_j Q_j h(t + (i-j)T_s) \tag{3}$$

for $t \in [0, T_s)$. Note that in (3) we have omitted the vector notation due to the isotropy of a point molecule source with free Brownian motion. The effect of N on the molecular communication system is thoroughly studied by numerical experiments in Section VII.

The receiver takes the molecular concentration $\phi_i(r,t)$ as input and performs the sensing of the molecular concentration inside the *receptor space*. The receptor space is modeled as a straight line of length 2ρ, a circle of radius ρ, and a sphere of radius ρ in the cases of 1, 2, and 3 dimensions, respectively. The molecular concentration is assumed to be homogeneous inside the receptor space. Fig. 1 is an illustration of the proposed system in a three-dimensional space. The random movement of the molecules following the Brownian motion causes an additional unwanted perturbation to the concentration value predicted by Fick's diffusion law [26]. We model such perturbation as a noise process added to the theoretical concentration value. Accordingly, we propose a high-level scheme for constructing the receiver as depicted in Fig. 2. The *concentration sensor* takes the molecule concentration $\phi_i(r,t)$ as input, and a noise process denoted by $\mathbf{W}_i(t)$ is added. As shown in [26], the noise process can be modeled as a Poisson counting process and is dependent on the input concentration function. Since we have assumed free Brownian motion with no interaction among the particles, the diffusion processes associated with distinct molecular emissions are thus independent. We can write

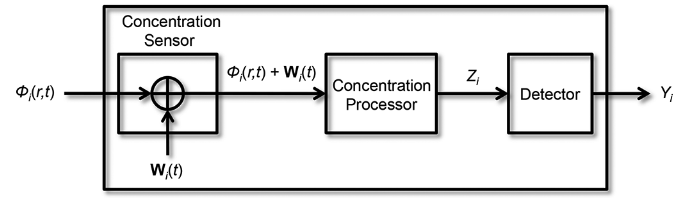

Fig. 2. High-level scheme of the proposed design of the receiver. It is shown in Section VII that an ML sequence detector is not necessary or even desirable in molecular communication.

$$\phi_i(r,t) + \mathbf{W}_i(t) = \sum_{j=i-N}^{i} X_j \left(Q_j h(t + (i-j)T_s) + W_j^i(t)\right), \tag{4}$$

where $W_j^i(t)$ is the noise process associated with the jth summand in (3) in the ith signaling interval. The *concentration processor* performs signal processing on the perturbed concentration function to produce an observation variable Z_i. Based on the detection scheme adopted, the *detector* then outputs the binary decision Y_i. As is demonstrated in numerical results, a ML detection scheme which takes the channel memory resulting from molecular diffusion into account is not ideal for molecular communication in terms of the tradeoff between the receiver complexity and the achievable throughput. A one-shot detection scheme which relies on the statistical property of the observation variable is shown to be more desirable.

III. Concentration Processor

The purpose of this block is to perform signal processing on the received molecular concentration to mitigate the reception noise and the ISI component in (4) to a minimum. Concerning the capabilities of nanoscale devices, sophisticated signal processing techniques, such as prediction, equalization, and frequency domain processing, are not desirable. Accordingly, we consider linear operations with low computational complexity. Let $\mathcal{C}\{\cdot\}$ denote a linear operator which stands for the concentration processing. It follows

$$\begin{aligned} Z_i &= \mathcal{C}\left\{\sum_{j=i-N}^{i} X_j \left(Q_j h(t + (i-j)T_s) + W_j^i(t)\right)\right\} \\ &= \sum_{j=i-N}^{i} X_j \left(Q_j \mathcal{C}\left\{h(t + (i-j)T_s)\right\} + \mathcal{C}\left\{W_j^i(t)\right\}\right). \end{aligned} \tag{5}$$

In the following, we propose three techniques for implementing the concentration processor. The expressions for $\mathcal{C}\{h(t+(i-j)T_s)\}$ and the statistical properties of $\mathcal{C}\{W_j^i(t)\}$ are analytically derived, which are then utilized in the subsequent sections for performance analysis.

A. Sampling Processing

For the sampling processing, a single sample of the sensed concentration function is taken at a predefined time instant. This is the most straightforward method with the lowest complexity [12], [19]. Let t_n denote the sampling time; $0 \le t_n < T_s$. It is clear that

$$Z_i = \sum_{j=i-N}^{i} X_j \left(Q_j h\left(t_n + (i-j)T_s\right) + W_j^i(t_n)\right), \quad (6)$$

where i is the index of the signaling interval, and $j = i-N, \ldots, i$.

The mean and variance of the considered noise process at a particular time instant given the number of emitted molecules Q_i are derived in [27] in the case of three dimensions; we generalize the results to one, two, and three dimensions as

$$\begin{aligned} E\left[\mathcal{C}\left\{W_j^i(t)\right\}\right] &= E\left[E\left[W_j^i(t_n) \mid Q_j\right]\right] = 0, \\ \mathrm{Var}\ \left[\mathcal{C}\left\{W_j^i(t)\right\}\right] &= E\left[\mathrm{Var}\ \left[W_j^i(t_n) \mid Q_j\right]\right] \\ &= E\left[\frac{Q_j h\left(t_n + (i-j)T_s\right)}{V}\right] \\ &= \frac{\mu_Q h\left(t_n + (i-j)T_s\right)}{V}, \end{aligned} \quad (7)$$

where V stands for the volume of the receiver and has different forms in different dimensions:

$$V = \begin{cases} 2\rho, & \text{if } d = 1, \\ \pi\rho^2, & \text{if } d = 2, \\ \frac{4\pi\rho^3}{3}, & \text{if } d = 3. \end{cases} . \quad (8)$$

Combing (7) and (8) we can observe that enlarging the size of the receptor space helps reduce the effect of noise, especially in higher dimensions.

B. Correlation Processing

For the correlation processing, the sensed concentration function is multiplied by a correlation function, denoted by $s(t)$, and then integrated over the signaling interval. It follows that

$$Z_i = \sum_{j=i-N}^{i} X_j \left(Q_j \int_0^{T_s} s(t) h\left(t + (i-j)T_s\right)\, dt + \int_0^{T_s} s(t) W_j^i(t)\, dt\right), \quad (9)$$

where i is the index of the signaling interval, and $j = i-N, \ldots, i$. In particular, we have the quasi-energy detection when $s(t) = 1$ since the concentration can not be negative, and the area under it can be properly defined as the signal energy [11], [12]. Motivated by the concept of matched filter, $s(t) = h(t)$ is also a design option.

For practical considerations, numerical integration of sums of rectangles with M sample points is considered in the following. We have

$$\begin{aligned} E\left[\mathcal{C}\left\{W_j^i(t)\right\}\right] &= \frac{T_s}{M} \sum_{n=1}^{M} s\left(n\frac{T_s}{M}\right) E\left[W_j^i\left(n\frac{T_s}{M}\right)\right] = 0, \\ \mathrm{Var}\ \left[\mathcal{C}\left\{W_j^i(t)\right\}\right] &= \frac{T_s^2}{M^2} \sum_{n=1}^{M} s\left(n\frac{T_s}{M}\right)^2 \mathrm{Var}\ \left[W_j^i\left(n\frac{T_s}{M}\right)\right] \\ &= \frac{\mu_Q T_s^2}{VM^2} \sum_{n=1}^{M} s\left(n\frac{T_s}{M}\right)^2 h\left(n\frac{T_s}{M} + (i-j)T_s\right). \end{aligned} \quad (10)$$

In (10) we have assumed the samples of the noise process are uncorrelated. It is shown in [26] that two adjacent samples of the noise process $W_j^i(t)$ can be regarded as statistically independent if their time separation satisfies

$$\frac{T_s}{M} > \frac{\rho^2}{D}. \quad (11)$$

As provided in Section VII, the time separation calculated by using typical values of D and ρ is approximately 0.01 seconds. It can be easily shown that the rectangular integration with such constraint yields a reasonably good approximation to (9).

Substituting $\frac{T_s D}{\rho^2}$ for M in (10), it follows that

$$\begin{aligned} \mathrm{Var}\ \left[\mathcal{C}\left\{W_j^i(t)\right\}\right] &= \frac{\mu_Q T_s}{VM} \cdot \frac{T_s}{M} \sum_{n=1}^{M} s\left(n\frac{T_s}{M}\right)^2 h\left(n\frac{T_s}{M} + (i-j)T_s\right) \\ &\simeq \frac{\mu_Q \rho^2}{DV} \int_0^{T_s} s(t)^2 h(t + (i-j)T_s)\, dt. \end{aligned} \quad (12)$$

For better readability, we define $E_j^i = \int_0^{T_s} s(t)^2 h(t + (i-j)T_s)\, dt$. In different dimensions, the variance takes different forms:

$$\mathrm{Var}\left[\mathcal{C}\left\{W_j^i(t)\right\}\right] = \begin{cases} \frac{\mu_Q \rho}{2D} E_j^i, & \text{if } d = 1, \\ \frac{\mu_Q}{\pi D} E_j^i, & \text{if } d = 2, \\ \frac{3\mu_Q}{4\pi\rho D} E_j^i, & \text{if } d = 3. \end{cases} . \quad (13)$$

We observe that in a three-dimensional space, enlarging the receptor space helps reduce the noise variance, while opposite effect is seen in a one-dimensional space. In a two-dimensional space, the size of the receptor space is irrelevant to the noise variance.

C. ISI Cancellation Processing

Here we propose a technique for signal processing which utilizes the knowledge of the concentration waveform to suppress the ISI component. In Fig. 3, we plot the impulse response $h(t)$ with $D = 10^{-6}$ cm^2/s and $r = 20$ μm in a three-dimensional space. As suggested by Fig. 3, $h(t)$ has a global maximum occurring at time t_p, which can be obtained by solving $\frac{dh(t)}{dt} = 0$. The expression for t_p in the case of three dimensions has been derived in [28]. We generalize the result to one, two, and three dimensions as

$$t_p = \frac{r^2}{2Dd}. \quad (14)$$

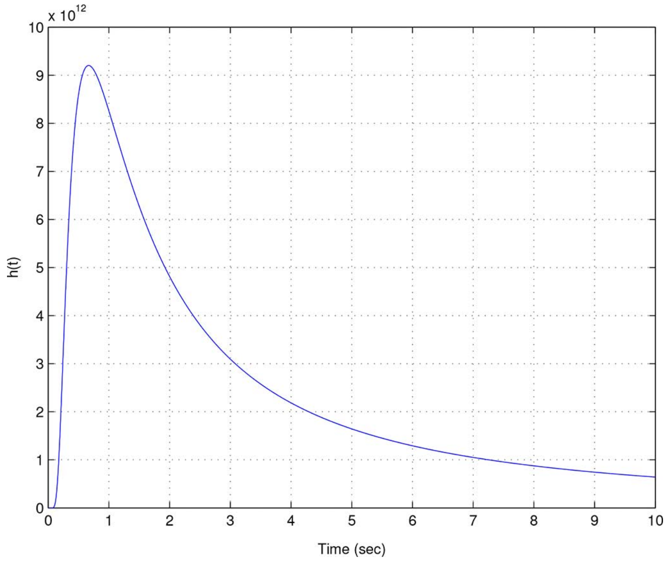

Fig. 3. Impulse response $h(t)$ defined in (2) with $D = 10^{-6}$ cm^2/s and $r = 20\ \mu$m.

By having a signaling interval larger than t_p, i.e., $T_s > t_p$, the ISI contributes a monotonically decreasing component in (4). We thus propose a low-complexity ISI cancellation processing as

$$\begin{aligned} Z_i &= \left(\phi_i\left(r, \min\{t_p, T_s\}\right) + \mathbf{W}_i\left(\min\{t_p, T_s\}\right)\right) \\ &\quad - \left(\phi_i(r, 0) + \mathbf{W}_i(0)\right) \\ &= \sum_{j=i-N}^{i} X_j \left(Q_j \left(h\left(\min\{t_p, T_s\} + (i-j)T_s\right)\right.\right. \\ &\quad \left.\left. - h\left((i-j)T_s\right)\right) + W_j^i\left(\min\{t_p, T_s\}\right) - W_j^i(0)\right). \end{aligned} \tag{15}$$

where i is the index of the signaling interval, and $j = i - N, \ldots, i$. Putting it in words, the proposed technique takes the difference between the samples taken at $t = \min\{t_p, T_s\}$ and $t = 0$ to serve as the observation variable Z_i.

It then follows

$$\begin{aligned} E\left[\mathcal{C}\left\{W_j^i(t)\right\}\right] &= E\left[W_j^i\left(\min\{t_p, T_s\}\right)\right] - E\left[W_j^i(0)\right] = 0, \\ \mathrm{Var}\left[\mathcal{C}\left\{W_j^i(t)\right\}\right] &= \mathrm{Var}\left[W_j^i\left(\min\{t_p, T_s\}\right)\right] + \mathrm{Var}\left[W_j^i(0)\right] \\ &= \frac{\mu_Q}{V}\left(h\left(\min\{t_p, T_s\} + (i-j)T_s\right)\right. \\ &\quad \left. + h\left((i-j)T_s\right)\right), \end{aligned} \tag{16}$$

where we have assumed the two noise samples are uncorrelated. For $T_s \geq t_p$, this holds if

$$t_p = \frac{r^2}{2Dd} \geq \frac{\rho^2}{D}, \tag{17}$$

which, after simplification, gives $\frac{r}{\rho} \geq \sqrt{2d} \approx 2.45$. This is a valid assumption since the dimension of the communication distance r is much larger than that of the receptor space ρ. For $T_s < t_p$, it is reasonable to consider $T_s \geq \frac{\rho^2}{D}$, as $\frac{\rho^2}{D}$ is on the order of 10^{-2} seconds as mentioned previously.

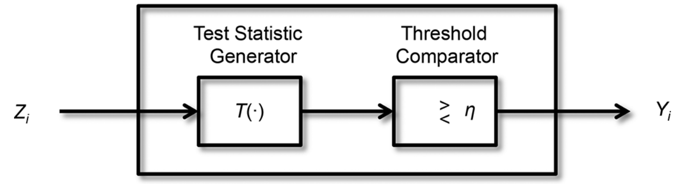

Fig. 4. One-shot detector for the proposed molecular communication system.

IV. One-Shot Detection

In this section, we propose an optimal one-shot detector which only utilizes information in the corresponding signaling interval. Fig. 4 depicts the high-level scheme of the proposed detector. The *test statistic generator* computes a test statistic, denoted by $T(Z_i)$, based on the observation variable output by the concentration processor. The *threshold comparator* then compares $T(Z_i)$ with a predetermined threshold η to generate the binary decision Y_i. 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, we adopt an information-theoretic approach. The ultimate goal of the detector design is concerned with the following optimization problem

$$\max I(X_i; Y_i), \tag{18}$$

where $I(\cdot)$ refers to the function of mutual information.

It is well-known that [29]

$$I(X_i; Y_i) = \sum_{X_i=0}^{1} \sum_{Y_i=0}^{1} P\left(Y_i | X_i\right) P(X_i) \log \frac{P\left(Y_i | X_i\right)}{P(Y_i)}. \tag{19}$$

By definition, we have

$$\begin{aligned} P\left(Y_i = 1 \mid X_i = 0\right) &= P_F, \\ P\left(Y_i = 1 \mid X_i = 1\right) &= P_D, \\ P\left(Y_i = 0 \mid X_i = 0\right) &= 1 - P_F, \\ P\left(Y_i = 0 \mid X_i = 1\right) &= 1 - P_D, \end{aligned} \tag{20}$$

where P_F and P_D denote the false alarm probability and the detection probability, respectively. The mutual information can thus be represented as a function of the probabilities P_F, P_D, and p. We henceforth denote the mutual information by $I(P_F, P_D, p)$. It can be shown that given P_F, the mutual information is a monotonically increasing function of P_D [30], [31]. Thus an information-optimal detector is equivalent to a Neyman-Pearson detector when P_F is given. In the following, we first formulate the binary hypothesis testing problem and then apply the Neyman-Pearson decision rule to derive the corresponding test statistic and the decision threshold.

A. Perfect a Priori Information

First we consider the case where the receiver has perfect knowledge of the *a priori* information. The detector is concerned with the binary hypothesis testing problem with observation Z_i as

$$\begin{aligned} &H_1 : \text{molecules are emitted at } t=0;\ Z_i \sim f_Z^1(z), \\ &H_0 : \text{otherwise};\ \ Z_i \sim f_Z^0(z), \end{aligned} \tag{21}$$

where $f_Z^k(z)$ denotes the probability density function of Z_i given that H_k is true. Since we are considering one-shot detection, the previous input bits acting as interference are modeled as Bernoulli random variables with success probability p. Z_i can thus be regarded as sums of independent random variables:

$$\begin{aligned} H_1 : Z_i = &\, Q_i \mathcal{C}\{h(t)\} + \mathcal{C}\left\{W_i^i(t)\right\} \\ &+ \sum_{j=i-N}^{i-1} X_j \left(Q_j \mathcal{C}\left\{h\left(t+(i-j)T_s\right)\right\} + \mathcal{C}\left\{W_j^i(t)\right\}\right), \\ H_0 : Z_i = &\sum_{j=i-N}^{i-1} X_j \left(Q_j \mathcal{C}\left\{h(t+(i-j)T_s)\right\} + \mathcal{C}\left\{W_j^i(t)\right\}\right). \end{aligned} \tag{22}$$

It is straightforward to show that each element in the series of independent random variables $\{X_j(Q_j\mathcal{C}\{h(t+(i-j)T_s)\} + \mathcal{C}\{W_j^i(t)\})\}$ has a finite variance. We thus assume that the Lindeberg's condition [32] holds true here, and Z_i converges to the Gaussian distribution as N approaches infinity[1]. By applying the Gaussian approximation, the binary hypothesis testing problem can be written as

$$\begin{aligned} H_1 : Z_i \sim \mathcal{N}\left(\mu_{Z1}, \sigma_{Z1}^2\right), \\ H_0 : Z_i \sim \mathcal{N}\left(\mu_{Z0}, \sigma_{Z0}^2\right), \end{aligned} \tag{23}$$

where it can be shown that

$$\begin{aligned} \mu_{Z0} &= p \sum_{j=i-N}^{i-1} \mu_Q \mathcal{C}\left\{h(t+(i-j)T_s)\right\}, \\ \mu_{Z1} &= \mu_Q \mathcal{C}\{h(t)\} + \mu_{Z0}, \\ \sigma_{Z0}^2 &= \left(p\sigma_Q^2 + p\mu_Q^2 - p^2\mu_Q^2\right) \sum_{j=i-N}^{i-1} \mathcal{C}\left\{h\left(t+(i-j)T_s\right)\right\}^2 \\ &\quad + p \sum_{j=i-N}^{i-1} \mathrm{Var}\left[\mathcal{C}\left\{W_j^i(t)\right\}\right], \\ \sigma_{Z1}^2 &= \sigma_Q^2 \mathcal{C}\{h(t)\}^2 + \mathrm{Var}\left[\mathcal{C}\left\{W_i^i(t)\right\}\right] + \sigma_{Z0}^2. \end{aligned} \tag{24}$$

In (24), we have utilized the fact from the previous section that the expectation of the noise process is zero, and the term $\mathrm{Var}\left[\mathcal{C}\{W_i^i(t)\}\right]$ depends on the method of concentration processing as given in (7), (10), and (16).

The Neyman-Pearson criterion states that the constrained optimization problem of maximizing P_D given a maximum allowable P_F is solved by forming the likelihood ratio test [33]

$$\Lambda(z) = \frac{f_Z^1(z)}{f_Z^0(z)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda, \tag{25}$$

[1]Loosely 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.

where λ is found by solving $P(\Lambda(z) > \lambda \mid H_0) = P_F$. Using (23) and (24), the likelihood ratio function can be derived as

$$\begin{aligned} \Lambda(z) &= \frac{\frac{1}{\sqrt{2\pi\sigma_{Z1}^2}} e^{-\frac{(z-\mu_{Z1})^2}{2\sigma_{Z1}^2}}}{\frac{1}{\sqrt{2\pi\sigma_{Z0}^2}} e^{-\frac{(z-\mu_{Z0})^2}{2\sigma_{Z0}^2}}} \\ &= \frac{\sigma_{Z0}}{\sigma_{Z1}} e^{\frac{\left(\sigma_{Z1}^2-\sigma_{Z0}^2\right)z^2 - 2\left(\mu_{Z0}\sigma_{Z1}^2-\mu_{Z1}\sigma_{Z0}^2\right)z + \mu_{Z0}^2\sigma_{Z1}^2 - \mu_{Z1}^2\sigma_{Z0}^2}{2\sigma_{Z0}^2\sigma_{Z1}^2}}. \end{aligned} \tag{26}$$

Combining (25) and (26) and taking the natural logarithm at both sides, we have

$$\begin{aligned} &\left(\sigma_{Z1}^2-\sigma_{Z0}^2\right) z^2 - 2\left(\mu_{Z0}\sigma_{Z1}^2 - \mu_{Z1}\sigma_{Z0}^2\right) z + \mu_{Z0}^2\sigma_{Z1}^2 - \mu_{Z1}^2\sigma_{Z0}^2 \\ &\qquad \underset{H_0}{\overset{H_1}{\gtrless}} 2\sigma_{Z0}^2\sigma_{Z1}^2 \left(\ln\lambda - \ln\frac{\sigma_{Z0}}{\sigma_{Z1}}\right). \end{aligned} \tag{27}$$

We can further rearrange (27) and obtain the test statistic of z in quadratic form as

$$T(z) = \left(z + \frac{b}{2a}\right)^2 \underset{H_0}{\overset{H_1}{\gtrless}} \frac{b^2-4ac}{4a^2} \triangleq \eta, \tag{28}$$

where we have defined

$$\begin{aligned} a &= \sigma_{Z1}^2 - \sigma_{Z0}^2, \\ b &= -2\left(\mu_{Z0}\sigma_{Z1}^2 - \mu_{Z1}\sigma_{Z0}^2\right), \\ c &= \mu_{Z0}^2\sigma_{Z1}^2 - \mu_{Z1}^2\sigma_{Z0}^2 - 2\sigma_{Z0}^2\sigma_{Z1}^2\left(\ln\lambda - \ln\frac{\sigma_{Z0}}{\sigma_{Z1}}\right). \end{aligned} \tag{29}$$

It then follows

$$\begin{aligned} P_F = \int_\eta^\infty f_Z^0(z)\, dz = &\, Q\left(\frac{\sqrt{\eta} - \mu_{Z0} - b/(2a)}{\sigma_{Z0}}\right) \\ &+ Q\left(\frac{\sqrt{\eta} + \mu_{Z0} + b/(2a)}{\sigma_{Z0}}\right), \end{aligned} \tag{30}$$

and the corresponding detection probability is

$$\begin{aligned} P_D = \int_\eta^\infty f_Z^1(z)\, dz = &\, Q\left(\frac{\sqrt{\eta} - \mu_{Z1} - b/(2a)}{\sigma_{Z1}}\right) \\ &+ Q\left(\frac{\sqrt{\eta} + \mu_{Z1} + b/(2a)}{\sigma_{Z1}}\right). \end{aligned} \tag{31}$$

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

$$\eta^*(p) = \arg\max_\eta \ I\left(P_F(\eta), P_D(\eta), p\right). \tag{32}$$

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 as

$$p^* = \arg\max_p \ I\left(P_F\left(\eta^*(p)\right), P_D\left(\eta^*(p)\right), p\right). \tag{33}$$

Note that the corresponding maximum mutual information $I(P_F(\eta^*(p^*)), P_D(\eta^*(p^*)), p^*)$ represents the theoretical maximum throughput (bits per channel utilization) of the

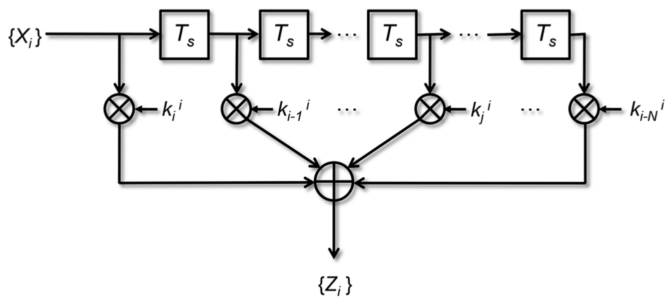

Fig. 5. Equivalent discrete-time ISI channel representation of the proposed molecular communication system as observed by the detector. The blocks labeled as T_s represent delay by T_s seconds; k_j^i stands for $(Q_j\mathcal{C}\{h(t+(i-j)T_s)\}+\mathcal{C}\{W_j^i(t)\})$ defined in (5).

considered diffusion-based communication system instead of the channel capacity, since a specific modulation technique and channel observation are involved.

B. A Priori Information Unknown

In the case where the receiver has no information of the *a priori* probability, the concept of *minimax* [33] which tries to mitigate the worst possible situation should be applied. However, as discussed in [31], the fact that $I(P_F, P_D, p) = 0$ when $p = 0$ 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 $p^\dagger$ such that

$$p^\dagger = \arg\max_p \int I\left(P_F\left(\eta^*(p)\right), P_D\left(\eta^*(p)\right), p\right)\, dp. \tag{34}$$

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

V. Sequence Detection

In this section, we investigate the design of a near ML sequence detector which aims to recover a transmitted bit sequence $\{X_i\}$ based on the corresponding sequence of observation variables $\{Z_i\}$. Compared with the one-shot detector proposed in the previous section, a ML sequence detector is expected to improve the system performance at the cost of added computational complexity. Note that such a sequence detector is not likely to be implemented on nanoscale devices. The purpose of investigating a sequence detector is to serve as a performance benchmark for our proposed one-shot detector.

Fig. 5 depicts an equivalent discrete-time ISI channel representation of the proposed molecular communication system as observed by the detector. Such representation resembles the conventional ISI channel experienced in wireless communications with $N+1$ tap gains. This motivates the application of the ML algorithms for sequence detection. To this end, proper definitions for the channel state and the branch metric have to be formulated. Due to the fact that N can be a large number, the complexity of the algorithm is impractical even for micro- or macro-devices. It is therefore inappropriate to take all past N input bits into account for defining the channel state. Instead, we consider the solution of the RS-Viterbi algorithm [22] for sequence detection. By applying the RS-Viterbi algorithm, the

$$\cdots\ 0\ 1\ 0\ \boxed{1\ 1\ 0\ 0}\ 1\ \cdots \qquad X_i = 1,\ S_i = \{0011\}$$

$$\cdots\ 0\ 1\ 0\ 1\ \boxed{1\ 0\ 0\ 1}\ 0\ \cdots \qquad X_{i+1} = 0,\ S_{i+1} = \{1001\}$$

Fig. 6. Example of the state definition for the RS-Viterbi algorithm with $m = 4$. Here we have $X_i = 1$; the associated channel state is defined as $\{0011\}$. For the next signaling interval we have $X_{i+1} = 0$; the new channel state becomes $\{1001\}$.

tradeoff between the achievable performance and the resulting complexity can be adjustable. The conventional Viterbi algorithm is a special case in the RS-Viterbi algorithm.

It is clear that for the proposed methods of concentration processing, the level of interference is monotonically decreasing with the index of the taps. We thus define the channel state S_i to be the past m input bits as

$$S_i = \{X_{i-1} \cdots X_{i-m}\}, \tag{35}$$

$m \leq N$. Note that when $m = N$, we have the classical "full-state" Viterbi algorithm. Fig. 6 shows an example of the state definition with $m = 4$.

Defining $\{\hat{X}_j\}_{j=i-N}^{i-m-1}$ as the survivor bit sequence associated with state S_i which is found by reversely searching the input bits corresponding to the survivor path stemming from S_i. Based on our derivation, the log likelihood of the observation Z_i given the current state S_i, the survivor sequence $\{\hat{X}_j\}_{j=i-N}^{i-m-1}$, and the input bit X_i is obtained as

$$\begin{aligned}
&\ln\left(P\left(Z_i \mid S_i, \{\hat{X}_j\}, X_i\right)\right)\\
&\quad= \ln\left(\mathcal{N}\left(\mu_{S_i,X_i,\{\hat{X}_j\}}, \sigma^2_{S_i,X_i,\{\hat{X}_j\}}\right)\right)\\
&\quad= -\frac{1}{2}\ln\left(2\pi\sigma^2_{S_i,X_i,\{\hat{X}_j\}}\right) - \frac{\left(Z_i - \mu_{S_i,X_i,\{\hat{X}_j\}}\right)^2}{2\sigma^2_{S_i,X_i,\{\hat{X}_j\}}},
\end{aligned} \tag{36}$$

where we have defined

$$\begin{aligned}
\mu_{S_i,X_i,\{\hat{X}_j\}} &= \sum_{j=i-m}^{i} X_j \mu_Q \mathcal{C}\left\{h\left(t+(i-j)T_s\right)\right\}\\
&\quad+ \sum_{j=i-N}^{i-m-1} \hat{X}_j \mu_Q \mathcal{C}\left\{h\left(t+(i-j)T_s\right)\right\},\\
\sigma^2_{S_i,X_i,\{\hat{X}_j\}} &= \sum_{j=i-m}^{i} X_j \left(\sigma_Q^2 \mathcal{C}\left\{h\left(t+(i-j)T_s\right)\right\}^2\right.\\
&\quad\left.+\mathrm{Var}\left[\mathcal{C}\left\{W_j^i(t)\right\}\right]\right)\\
&\quad+ \sum_{j=i-N}^{i-m-1} \hat{X}_j \left(\sigma_Q^2 \mathcal{C}\left\{h\left(t+(i-j)T_s\right)\right\}^2\right.\\
&\quad\left.+\mathrm{Var}\left[\mathcal{C}\left\{W_j^i(t)\right\}\right]\right).
\end{aligned} \tag{37}$$

We define the branch metric for the transition from state S_i with input bit X_i as the log likelihood value given in (36).

With the channel state and the branch metric all defined, we can proceed with the RS-Viterbi algorithm in the traditional way to find the bit sequence with the highest likelihood value. We remark that the maximum *a posteriori* (MAP) algorithm with the concept of state reduction, e.g., the reduced-state BCJR algorithm [34], can be similarly applied for performing sequence detection here. We omit such discussion as sophisticated end-to-end communication skills are inherently not suited for nanoscale devices.

VI. Parameter Estimation

In this section, we propose and analyze an asymptotic ML estimator for estimating the variance of the number of molecules emitted by the transmitter σ_Q^2. The mean μ_Q and the variance σ_Q^2 are required by the detector to compute the decision threshold. As mentioned in Section II, ideally the transmitter would emit a fixed number of molecules for each binary signaling. We consider a more realistic situation by modeling the emission process to be noisy and possess a probability distribution. It is assumed that the mean value of the number of molecules μ_Q is pre-determined and agreed between the transmitter and the receiver. The ability of the transmitter in controlling the emission process, i.e., the variance σ_Q^2, is left for the receiver to estimate.

The proposed estimator is asymptotic ML in the sense that the ML estimation is achieved with an increasingly large size of the receptor space. The technique of sampling processing proposed in Section III-A is utilized. Let the transmitter send a series of K logical 1's which are then processed at the receiver side using the sampling processing technique. The sampling time is set to be at the pulse peak as given in (14). We consider the time separation between each logical 1 to be large enough such that the effect of ISI is nearly negligible as compared with the signal component. The result of the sampling processing at the ith pulse can then be written as

$$Z_i = Q_i h(t_p) + W_i^i(t_p), \tag{38}$$

where $i = 1, \ldots, K$. We propose an asymptotic ML estimator as

$$\hat{\sigma}_Q^2 = \frac{1}{h(t_p)^2 K} \sum_{i=1}^{K} \left(Z_i - \mu_Q h(t_p)\right)^2 - \frac{\mu_Q}{h(t_p)V}. \tag{39}$$

It can be observed in (7) that $\lim_{V\to\infty} \mathrm{Var}[W_i^i(t_p)] = 0$. By letting $V \to \infty$, or equivalently, the molecular concentration being perfectly sensed, (39) thus degenerates to

$$\begin{aligned}\hat{\sigma}_Q^2 &= \frac{1}{h(t_p)^2 K} \sum_{i=1}^{K} \left(Q_i h(t_p) - \mu_Q h(t_p)\right)^2 \\ &= \frac{1}{K} \sum_{i=1}^{K} (Q_i - \mu_Q)^2, \end{aligned} \tag{40}$$

which is the ML estimator for the variance of K i.i.d. Gaussian random variables [35]. The proposed estimator and the associated estimation process could be designed as a training phase prior to the actual data transmissions. To show that the proposed estimator is *unbiased* and *consistent* by deriving the mean and the variance, we derive the following expressions

$$\begin{aligned} E[\hat{\sigma}_Q^2] &= \frac{1}{h(t_p)^2 K} \sum_{i=1}^{K} E\left[\left(Z_i - \mu_Q h(t_p)\right)^2\right] - \frac{\mu_Q}{h(t_p)V} \\ &= \frac{1}{h(t_p)^2 K} \sum_{i=1}^{K} \left(\sigma_Q^2 h(t_p)^2 + \frac{\mu_Q h(t_p)}{V}\right) - \frac{\mu_Q}{h(t_p)V} \\ &= \sigma_Q^2. \end{aligned} \tag{41}$$

The proposed estimator is thus unbiased. For the variance, we have

$$\begin{aligned} \mathrm{Var}\left[\hat{\sigma}_Q^2\right] &= \frac{1}{h(t_p)^4 K^2} \sum_{i=1}^{K} \mathrm{Var}\left[\left(Z_i - \mu_Q h(t_p)\right)^2\right] \\ &= \frac{1}{h(t_p)^4 K^2} \sum_{i=1}^{K} \Bigg(E\left[\left(Z_i - \mu_Q h(t_p)\right)^4\right] \\ &\quad - E\left[\left(Z_i - \mu_Q h(t_p)\right)^2\right]^2 \Bigg) \\ &= \frac{1}{h(t_p)^4 K^2} \sum_{i=1}^{K} \Bigg(\frac{h(t_p)^2}{V^2}\left(2\mu_Q^2 + 3\sigma_Q^2\right) \\ &\quad + \frac{4\mu_Q \sigma_Q^2 h(t_p)^3}{V} + 2\sigma_Q^4 h(t_p)^4 \Bigg) \\ &= \frac{1}{K}\left(\frac{1}{h(t_p)^2 V^2}(2\mu_Q^2 + 3\sigma_Q^2) + \frac{4\mu_Q \sigma_Q^2}{h(t_p)V} + 2\sigma_Q^4 \right), \end{aligned} \tag{42}$$

where it is clear that $\lim_{K\to\infty} \mathrm{Var}\left[\hat{\sigma}_Q^2\right] = 0$. This proves the consistency. In the case where $V \to \infty$, the variance becomes $\frac{2\sigma_Q^4}{K}$, which coincides with the performance of the ML estimator for K i.i.d. Gaussian random variables.

Due to the difficulty of characterizing the joint distribution of the samples $\{Z_i\}$, the derivations of further properties and the *efficiency* of the proposed estimator are rather intricate. This can be an independent study and is currently in progress.

VII. Numerical Results

In this section we present the numerical results for the attainable mutual information and the theoretical maximum throughput of the proposed diffusion-based communication system. The proposed signal processing techniques as well as the proposed detection schemes are investigated. Two sets of system parameters are considered: the short-range and the long-range molecular communication scenarios. The short-range molecular communication happens naturally as the mechanism for biochemical signaling in living cells, e.g., calcium ion signaling [3] and neural signaling [36]; while the long-range molecular communication mostly serves as the signaling method among living organisms, e.g., pheromone propagation and the dispersal of pollen and spores [2]. In the following, we set the diffusion constant $D = 10^{-6}$ cm^2/s (cellular cytoplasm, [37]), the communication distance $r = 20$ μm, and $\mu_Q = 10^9$ molecules for the short-range molecular communication. For the case of long-range communication, we set $D = 0.43$ cm^2/s, $r = 2$ cm, and $\mu_Q = 10^{15}$ molecules (see [2]). For both scenarios we set the radius of the receptor space

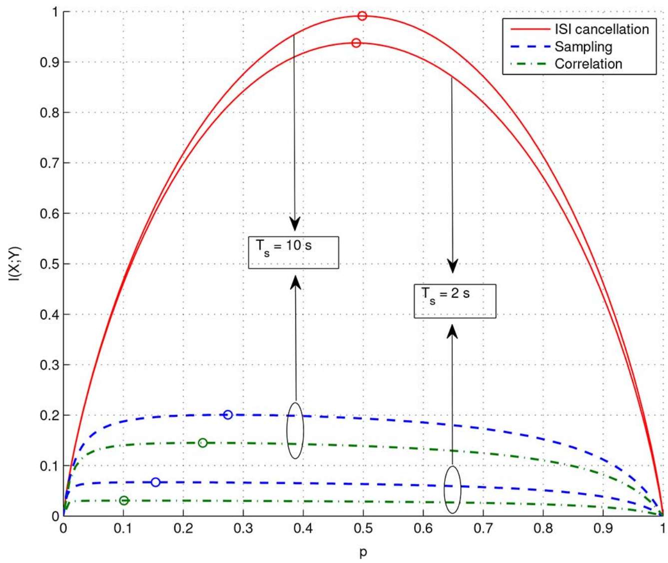

Fig. 7. The attainable mutual information versus the *a priori* probability p for the short-range diffusion-based communication. One-shot detection with perfect knowledge of p is adopted. $d = 1$.

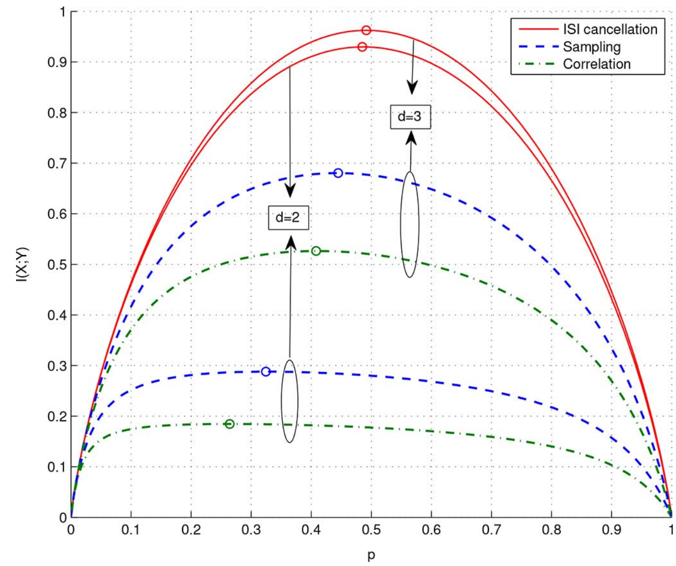

Fig. 8. The attainable mutual information versus the *a priori* probability p for the short-range diffusion-based communication. One-shot detection with perfect knowledge of p is adopted. $T_s = 2$ s.

$\rho = d/20$ and $\sigma_Q = 0.3\ \mu_Q$, which gives a medium coefficient of variation (CV) of 0.3. Perfect estimation of σ_Q is assumed at the receiver. Unless otherwise stated, the number of interfering signaling intervals N is set to 20.

For the technique of sampling processing, we set the sampling instant $t_n = \min\{t_p, T_s\}$. For the correlation processing, we set the number of sampling points in a signaling interval $M = \frac{T_s D}{\rho^2}$, and $s(t) = 1$, i.e., the quasi-energy processing. One should note that in the sampling processing, by sampling at the time corresponding to the pulse peak t_p whenever applicable, we are actually utilizing our knowledge of the received waveform. We consider this case as it gives the best performance achievable by the technique of sampling processing.

A. One-Shot Detection With Perfect a Priori Information

We first compare the performance of the proposed signal processing techniques. In Fig. 7, we plot the maximum attainable mutual information for the proposed communication system using one-shot detection with perfect knowledge of the *a priori* probability p in one dimension, i.e., $d = 1$, in the short-range communication scenario. Two sets of results which correspond to different lengths of signaling interval $T_s = 2$ s and $T_s = 10$ s are given. The circles indicate the point which achieves the channel capacity of the system as given in (33). We observe that the achievable capacity of the ISI cancellation processing outperforms the other two significantly, and the sampling processing is approximately twice as good as the correlation processing. A channel capacity of nearly 1 bit per channel utilization can be achieved by the ISI cancellation processing; while the sampling and the correlation based processing only yield capacities less than 0.2. It is also observed that the achievable channel capacity increases along with T_s due to less effect of ISI, as one would expect. From the results obtained, we conclude that proposed ISI cancellation processing is very desirable for nanoscale receivers for its low complexity and outstanding performance.

In Fig. 8, we plot the maximum attainable mutual information for the proposed communication system under the same setup as in Fig. 7 in both two and three dimensions, i.e., $d = 2$ and 3. We present the results corresponding to $T_s = 2$ s. It is observed that higher mutual information can be achieved in higher dimensions for all three techniques of signal processing. This results from the fact that the factor of dimension serves as the exponent affecting the decaying rate of the concentration function with time, which in turn lowers the effect of ISI. In particular, we observe that the techniques of sampling and correlation processing both benefit from the effect of higher dimensions much more obvious than the ISI cancellation processing. This is expected since most of the ISI component in the received concentration function has already been eliminated in the latter case as explained in Section III-C.

Continuing from Fig. 7, in Fig. 9 we plot the maximum attainable mutual information in one dimension with 100 interfering signaling intervals, i.e., $N = 100$ for performance comparison. The rest of the parameters are the same as that in Fig. 7. Due to the effect of higher level of interference, the attainable channel capacities of the techniques of the sampling processing and the correlation processing both decrease by a certain amount. On the other hand, it is observed that for the ISI cancellation processing, the curves corresponding to $N = 20$ and $N = 100$ nearly overlap. We omit the corresponding results in higher dimensions and for longer signaling intervals since the effect of interference only becomes less significant as explained previously. The results suggest that the setting of 20 interfering signaling intervals successfully characterizes the system performance with infinite channel memory when the proposed ISI cancellation processing is adopted. A channel capacity of 1 bit per channel utilization can be ultimately achieved by extending the duration of the signaling interval.

From the previous results we have observed that the achievable channel capacity increases along with the length of the signaling interval T_s and the dimension of the space d. We next

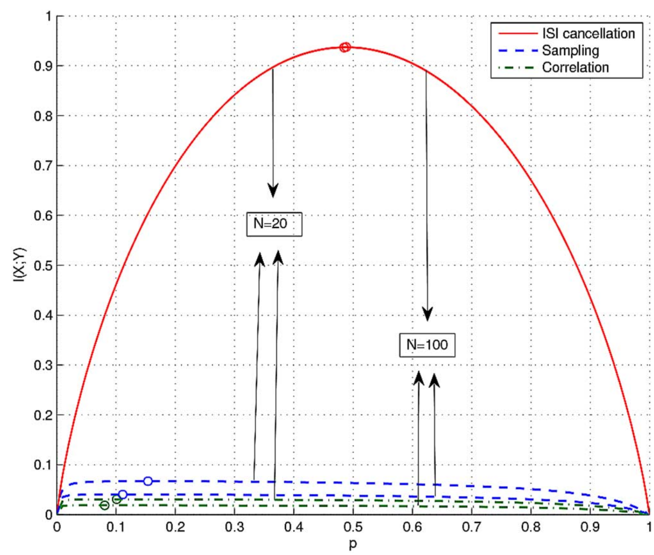

Fig. 9. The attainable mutual information versus the *a priori* probability p for the short-range diffusion-based communication. One-shot detection with perfect knowledge of p is adopted. $T_s = 2$ s; $d = 1$.

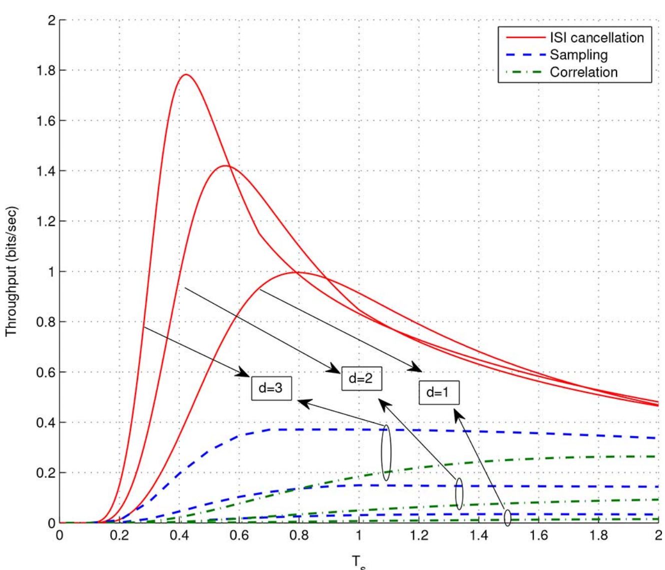

Fig. 10. The theoretical maximum throughput versus the length of signaling interval T_s for the short-range diffusion-based communication. One-shot detection with perfect knowledge of p is adopted.

experiment with the optimal setting for T_s to achieve the theoretical maximum throughput (bits per second). In Fig. 10, we plot the theoretical maximum throughput as a function of T_s for all three signal processing techniques and for $d = 1, 2$, and 3 in the short-range communication scenario. In Fig. 11, we present the corresponding results for the long-range communication scenario. By comparing Figs. 10 and 11, we observe that the proposed system yields higher theoretical throughput in short-range communication for all signal processing techniques and for $d = 1, 2$, and 3. The relative performance among different processing techniques and different dimensions is, however, similar between the cases of short-range and long-range communication. It is again shown that the performance of the

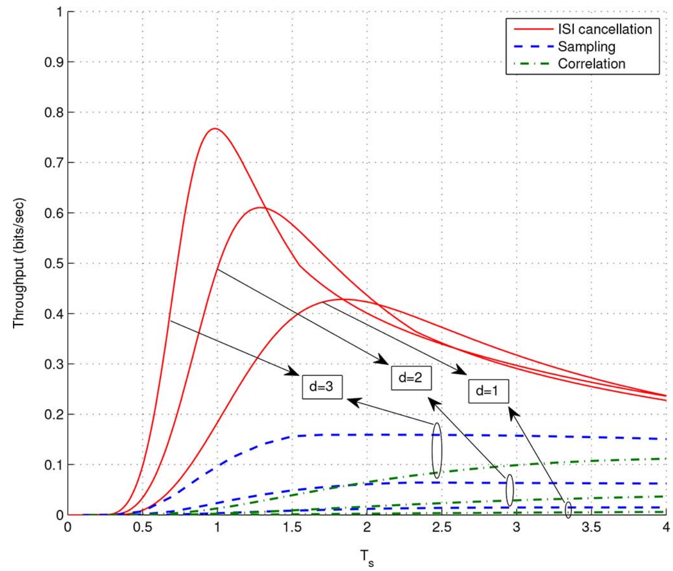

Fig. 11. The theoretical maximum throughput versus the length of signaling interval T_s for the long-range diffusion-based communication. One-shot detection with perfect knowledge of p is adopted.

ISI cancellation processing stands out under various system settings. Specifically, in the case of short-range communication, a theoretical maximum throughput of approximately 1, 1.4, and 1.8 bits per second is achieved for $d = 1, 2$, and 3, respectively. This is to be compared with the maximum throughput given by the sampling processing in three dimensions, which is slightly below 0.4 bits per second. For the case of long-range communication, we have approximately 0.43, 0.61, and 0.75 bits per second for $d = 1, 2$, and 3, respectively. The maximum throughput given by the sampling processing in three dimensions in this case falls below 0.15 bits per second.

B. One-Shot Detection With No a Priori Information

In Figs. 12 and 13, we plot the attainable mutual information with no knowledge of the *a priori* probability p for two and three dimensions, i.e., $d = 2$ and 3, respectively. The curves corresponding to the cases where the perfect knowledge of p is available to the receiver are also plotted for ease of comparison. Parameters for short-range communication are applied in both figures. The squares correspond to the values of p which give the maximum amount of information as defined in (34). For the techniques of sampling and correlation processing, we see that the system entails huge performance loss when there is a mismatch between a presumed value and the actual value of p. This is expected since the distribution of ISI is dependent on p as shown in (24). Exact knowledge of the *a priori* probability at the receiver is thus important for the system to operate properly. In higher dimensions the penalty for not having the knowledge of p is lower due to less effect of ISI as explained previously. On the other hand, the ISI cancellation processing without the knowledge of the *a priori* probability p still achieves the same channel capacity as before, and the two curves nearly overlap. This suggests that the knowledge of *a priori* information is not needed for the system to operate satisfactorily when the ISI cancellation processing is employed.

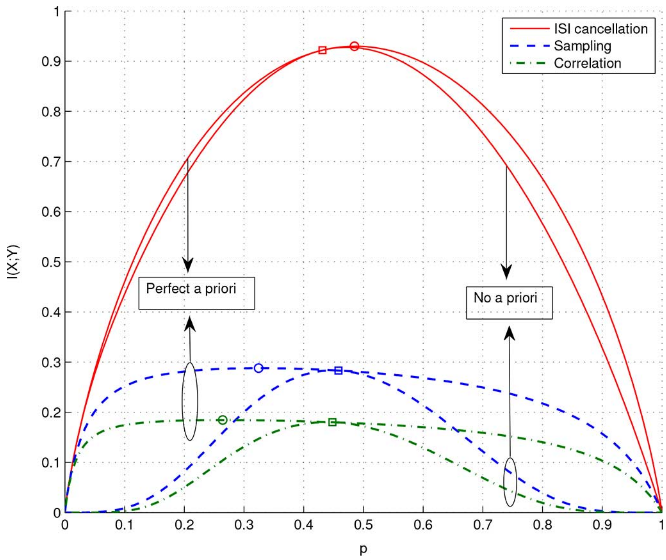

Fig. 12. The attainable mutual information versus the *a priori* probability p for the short-range diffusion-based communication. One-shot detection both with and without knowledge of p is adopted. $T_s = 2$ s; $d = 2$.

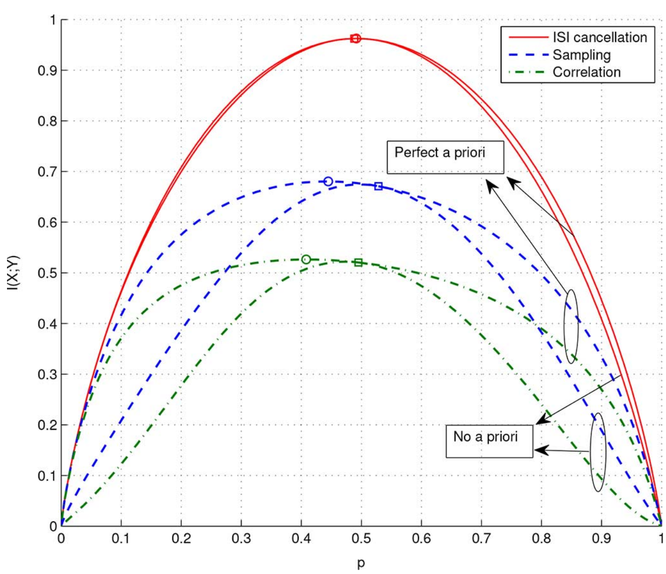

Fig. 13. The attainable mutual information versus the *a priori* probability p for the short-range diffusion-based communication. One-shot detection both with and without knowledge of p is adopted. $T_s = 2$ s; $d = 3$.

C. Sequence Detection Using RS-Viterbi

In Fig. 14, we plot the theoretical maximum throughput versus the length of the signaling interval T_s for the ISI cancellation processing using the RS-Viterbi algorithm for $d = 1, 2,$ and 3. Parameters for short-range communication are applied. The bit error probability of the RS-Viterbi detection scheme is first simulated and then used to compute the theoretical throughput. For the simulation we set the *a priori* probability $p = 0.5$ and the order of memory $m = 4$. The curves corresponding to the ISI cancellation processing using one-shot detection are also plotted for ease of comparison. The dotted line indicates the curve of the optimal case of 1 bit per channel utilization, or equivalently, the optimal throughput of $1/T_s$ bits per second. As expected, the sequence detection yields lower error probabilities, hence higher maximum throughput since more information is utilized for the determination of the transmitted bits. Approximately 2.4, 2.6, and 2.6 times higher maximum throughput than the one-shot detection can be achieved by the RS-Viterbi sequence detection for $d = 1, 2,$ and 3, respectively. One should note that, however, such performance gain comes at the cost of computational complexity which is orders of magnitude higher than that of the one-shot detection. Furthermore, to actually benefit from using the sequence detection scheme, the system has to operate at a very short signaling interval, which raises the issue of the synchronization design. Though perfect synchronization between the transmitter and the receiver is assumed in this work, we envision the synchronization design to be based on sending training molecular impulses and detecting the pulse peaks. Having a very short signaling interval thus hinders the design of the synchronization mechanism. Considering the potential performance difference, feasibility, and the capabilities of nanoscale devices, we conclude that the ISI cancellation processing with one-shot detection is the most promising design framework for receivers of diffusion-based communication systems.

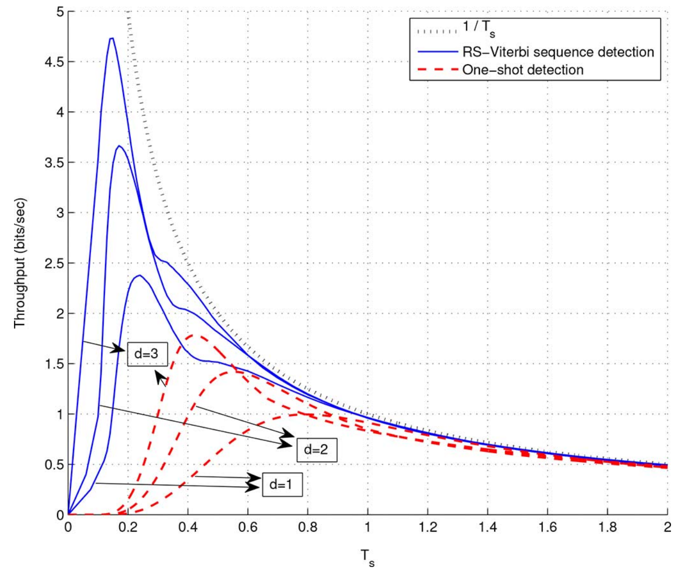

Fig. 14. The theoretical maximum throughput versus the length of signaling interval T_s for the short-range diffusion-based communication. The ISI cancellation processing using both the one-shot detection and the RS-Viterbi sequence detection are adopted. We set $p = 0.5$ for RS-Viterbi; $m = 4$.

VIII. Conclusion

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 molecules emitted for binary signaling exhibits randomness with known statistical properties. A diffusion channel with memory in the cases of one, two, and three dimensions all have been considered. Three signal processing techniques for the molecular concentration with low computational complexity have been proposed including an ISI cancellation technique. Both a low-complexity one-shot optimal detector designed for mutual information maximization and a sequence detector using the RS-Viterbi algorithm have been proposed and analyzed. An asymptotic ML estimator for estimating the randomness of the molecular emission by the transmitter has also been proposed. Numerical results indicate that the proposed ISI cancellation processing which operates by taking the difference of two samples followed by the one-shot optimal detector can achieve near-optimal throughput without the need of *a priori* information over a wide range of system parameters of practical interest. The proposed receiver design guarantees diffusion-based communication to be operated without failure in the case of infinite channel memory. A channel capacity of 1 bit per channel utilization can be ultimately achieved by extending the duration of the signaling interval.

REFERENCES

[1] I. F. Akyildiz, F. Brunetti, and C. Blázquez, "Nanonetworks: A new communication paradigm," *Comput. Netw. (Elsevier) J.*, vol. 52, pp. 2260–2279, Aug. 2008.

[2] L. P. Giné and I. F. Akyildiz, "Molecular communication options for long range nanonetworks," *Computer Netw.*, vol. 53, pp. 2753–2766, Nov. 2009.

[3] 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 Conf. Nanotechnol.*, Jul. 2005, vol. 2, pp. 478–481.

[4] 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.

[5] N. Garralda, I. Llatser, A. Cabellos-Aparicio, and M. Pierobon, "Simulation-based evaluation of the diffusion-based physical channel in molecular nanonetworks," in *Proc. IEEE Comput. Commun. Workshops (INFOCOM WKSHPS)*, Apr. 2011, pp. 443–448.

[6] I. Llatser, I. Pascual, N. Garralda, A. Cabellos-Aparicio, M. Pierobon, E. Alarcòn, and J. Solè-Pareta, "Exploring the physical channel of diffusion-based molecular communication by simulation," in *Proc. IEEE GLOBECOM,*, Houston, TX, USA, Dec. 2011, pp. 1–5.

[7] B. Atakan and O. Akan, "On channel capacity and error compensation in molecular communication," *Trans. Computat. Syst. Biol. X*, pp. 59–80, 2008.

[8] M.-J. Moore, T. Suda, and K. Oiwa, "Molecular communication: Modeling noise effects on information rate," *IEEE Trans. NanoBiosci.*, vol. 8, no. 2, pp. 169–180, Jun. 2009.

[9] M. U. Mahfuz, D. Makrakis, and H. T. Mouftah, "On the characterization of binary concentration-encoded molecular communication in nanonetworks," *Nano Commun. Netw. J., Elsevier Sci.*, vol. 1, no. 4, pp. 289–300, Dec. 2010.

[10] D. Arifler, "Capacity analysis of a diffusion-based short-range molecular nano-communication channel," *Comput. Netw.*, vol. 55, pp. 1426–1434, Dec. 2010.

[11] M. U. Mahfuz, D. Makrakis, and H. T. Mouftah, "Characterization of intersymbol interference in concentration-encoded unicast molecular communication," in *Proc. 24th Canad. Conf. Elect. Comput. Eng. (CCECE)*, May 2011, pp. 164–168.

[12] M. Mahfuz, D. Makrakis, and H. Mouftah, "On the detection of binary concentration-encoded unicast molecular communication in nanonetworks," in *Proc. 4th Int. Conf. Bio-Insp. Syst. Signal Process. (BIOSIGNALS-2011)*, Jan. 2011.

[13] M. Mahfuz, D. Makrakis, and H. Mouftah, "On the characteristics of concentration-encoded multi-level amplitude modulated unicast molecular communication," in *Proc. 24th Canad. Conf. Electr. Comput. Eng. (CCECE)*, May 2011, pp. 312–316.

[14] M. Kuran, H. Yilmaz, T. Tugcu, and I. Akyildiz, "Modulation techniques for communication via diffusion in nanonetworks," in *Proc. IEEE ICC*, Jun. 2011, pp. 1–5.

[15] A. Tseng, K. Chen, C. Chen, and K. Ma, "Electron beam lithography in nanoscale fabrication: Recent development," *IEEE Trans. Electron. Pack. Manufact.*, vol. 26, no. 2, pp. 141–149, Apr. 2003.

[16] M. Moore, A. Enomoto, T. Nakano, R. Egashira, T. Suda, A. Kayasuga, H. Kojima, H. Sakakibara, and K. Oiwa, "A design of a molecular communication system for nanomachines using molecular motors," in *Proc. PerCom Workshops 4th Ann. IEEE Int. Conf. Pervas. Comput. Commun. Workshops*, Mar. 2006.

[17] Y.-P. Hsieh, P.-J. Shih, Y.-C. Lee, P.-C. Yeh, and K.-C. Chen, "An asynchronous communication scheme for molecular communication," in *Proc. 2nd IEEE Int. Workshop on Molec. Nanoscale Commun. (MONACom)*, Jun. 2012.

[18] S. Kadloor and R. Adve, "A framework to study the molecular communication system," in *Proc. 18th Comput. Commun. Netw. (ICCCN)*, Aug. 2009, pp. 1–6.

[19] A. Einolghozati, M. Sardari, A. Beirami, and F. Fekri, "Capacity of discrete molecular diffusion channels," in *Proc. IEEE Inf. Theory Process. (ISIT)*, Aug. 2011, pp. 723–727.

[20] D. Miorandi, "A stochastic model for molecular communications," *Nano Commun. Netw.*, vol. 2, no. 4, pp. 205–212, 2011.

[21] M. Pierobon and I. Akyildiz, "Capacity of a diffusion-based molecular communication system with channel memory and molecular noise," *IEEE Trans. Inf. Theory*, vol. 59, no. 2, pp. 942–954, 2013.

[22] X. Chen and K. Chugg, "Reduced-state soft-input/soft-output algorithms for complexity reduction in iterative and non-iterative data detection," in *Proc. IEEE ICC*, 2000, vol. 1, pp. 6–10.

[23] J. Philibert, "One and a half century of diffusion: Fick, Einstein, before and beyond," *Diffusion Fundament.*, pp. 6.1–6.19, 2006.

[24] A. Einstein, *Investigations of the Theory of Brownian Movement*. New York, NY, USA: Dover, 1956.

[25] W. H. Bossert and E. Wilson, "Analysis of olfactory communication among animals," *J. Theoretic. Biol.*, vol. 5, pp. 443–469, 1963.

[26] M. Pierobon and I. F. Akyildiz, "Diffusion-based noise analysis for molecular communication in nanonetworks," *IEEE Trans. Signal Process.*, vol. 59, pp. 2532–2547, Jun. 2011.

[27] M. Pierobon and I. F. Akyildiz, "Noise analysis in ligand-binding reception for molecular communication in nanonetworks," *IEEE Trans. Signal Process.*, vol. 59, pp. 4168–4182, Sep. 2011.

[28] I. Llatser, E. Alarcòn, and M. Pierobon, "Diffusion-based channel characterization in molecular nanonetworks," in *Proc. IEEE Comput. Commun. Workshops (INFOCOM WKSHPS)*, Apr. 2011, pp. 467–472.

[29] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York, NY, USA: Wiley, 2006.

[30] D. Middleton, *An Introduction to Statistical Communication Theory*. Hoboken, NJ, USA: Wiley-IEEE , May 1996.

[31] T. Gabriele, "Information criteria for threshold determination (corresp.)," *IEEE Trans. Inf. Theory*, vol. 12, no. 4, pp. 484–486, Oct. 1966.

[32] R. B. Ash, *Probability and Measure Theory*, 2nd ed. New York, NY, USA: Academic, Dec. 1999.

[33] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York, NY, USA: Springer, Mar. 1994.

[34] G. Colavolpe, G. Ferrari, and R. Raheli, "Reduced-state BCJR-type algorithms," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 5, pp. 848–859, May 2001.

[35] M. Barkat, *Signal Detection and Estimation*, 2nd ed. Dedham, MA, USA: Artech House, 2005.

[36] D. H. Johnson, "Information theory and neural information processing," *IEEE Trans. Inf. Theory*, vol. 56, no. 2, pp. 653–666, Feb. 2010.

[37] B. S. Donahue and R. F. Abercrombie, "Free diffusion coefficient of ionic calcium in cytoplasm," *Cell Calcium*, vol. 8, no. 6, pp. 437–448, 1987.

Ling-San Meng (S'12) received the B.S. degree from the National Tsing Hua University in 2006, and the Ph.D. degree from the National Taiwan University in 2012, all in electrical engineering.

He spent one year as a visiting scholar in 2011 in Georgia Institute of Technology, Atlanta, working with Prof. Ian F. Akyildiz. He joined HTC Corp. in August 2012 as a cellular standard developing engineer. His research interests include cooperative communications, diffusion-based molecular communication, and Long Term Evolution (LTE) advanced networks.

Ping-Cheng Yeh (M'05) received the B.S. degree in mathematics and the M.S. degree in electrical engineering from the National Taiwan University, in 1996 and 1998, respectively. In 2005, he received the Ph.D. degree in electrical engineering and computer science from the University of Michigan, Ann Arbor.

He joined the Department of Electrical Engineering and the Graduate Institute of Communication Engineering at the National Taiwan University in August 2005. His research interests include molecular communications, wireless multimedia transmissions, physical layer security, cooperative communications, cross-layer design in wireless networks, and online education platform design.

Dr. Yeh has received various awards in the past, including EECS Outstanding GSI Award (2002), University of Michigan Outstanding GSI Award (2003), NTU Excellence in Teaching Award (2008, 2009), and NTU Distinguished Teaching Award (2010). He is currently the Director of NTU MOOC. He is also the Associate Director of Center for Teaching and Learning Development at the National Taiwan University.

Kwang-Cheng Chen (M'89–SM'94–F'07) received the B.S. degree from the National Taiwan University in 1983, and the M.S. and Ph.D. degrees from the University of Maryland, College Park, in 1987 and 1989, all in electrical engineering.

From 1987 to 1998, he worked with SSE, COMSAT, IBM Thomas J. Watson Research Center, and National Tsing Hua University, in mobile communications and networks. Since 1998, he has been with National Taiwan University, Taipei, Taiwan, ROC, and is the Distinguished Professor and Associate Dean for academic affairs in the College of Electrical Engineering and Computer Science, National Taiwan University. He has authored and coauthored more than 200 IEEE papers and approximately 30 granted US patents. He co-edited (with R. DeMarca) the book *Mobile WiMAX* (New York: Wiley, 2008), and authored the book *Principles of Communications* (New York: River, 2009), and coauthored (with R. Prasad) another book *Cognitive Radio Networks* (New York: Wiley, 2009). His recent research interests include wireless communications, network science, and data analytics.

Dr. Chen has been actively involved in the organization of various IEEE conferences as General/TPC Chair/Co-Chair, and has served in editorships with a few IEEE journals. He also actively participates in and has contributed essential technology to various IEEE 802, Bluetooth, and LTE and LTE-A wireless standards. He has received a number of awards including the 2011 IEEE COMSOC WTC Recognition Award, and has coauthored a few award-winning papers published in the IEEE journals and conferences, including 2014 IEEE Jack Neubauer Memorial Award.

Ian F. Akyildiz (M'86–SM'89–F'96) received the B.S., M.S., and Ph.D. degrees in computer engineering from the University of Erlangen-Nurnberg, Germany, in 1978, 1981, and 1984, respectively.

Currently, he is the Ken Byers Chair Professor in telecommunications with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, the Director of the Broadband Wireless Networking Laboratory and Chair of the Telecommunication Group at Georgia Tech. He is an honorary professor with the School of Electrical Engineering at Universitat Politecnica de Catalunya (UPC), Barcelona, Catalunya, Spain, and founded the N3Cat (NaNoNetworking Center in Catalunya). Since September 2012, he has also been a FiDiPro Professor (Finland Distinguished Professor Program (FiDiPro) supported by the Academy of Finland) at Tampere University of Technology, Department of Communications Engineering, Finland. His current research interests are in nanonetworks, Long Term Evolution (LTE) advanced networks, and wireless sensor networks.

Dr. Akyildiz is the Editor-in-Chief of *Computer Networks (Elsevier) Journal*, and the founding Editor-in-Chief of the *Ad Hoc Networks (Elsevier) Journal*, the *Physical Communication (Elsevier) Journal*, and the *Nano Communication Networks (Elsevier) Journal*. He is an ACM Fellow (1997). He received numerous awards from the IEEE and ACM.