



# The University of Bradford Institutional Repository

<http://bradscholars.brad.ac.uk>

This work is made available online in accordance with publisher policies. Please refer to the repository record for this item and our Policy Document available from the repository home page for further information.

To see the final version of this work please visit the publisher's website. Available access to the published online version may require a subscription.

**Link to publisher's version:** <https://doi.org/10.1109/MWC.2016.7553035>

**Citation:** Guo W, Asyhari T, Farsad N, Yilmaz HB, Li B, Eckford A and Chae C-B (2016) Molecular Communications: Channel Model and Physical Layer Techniques. IEEE Wireless Communications. 23(4): 120-127.

**Copyright statement:** © 2016 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

# Molecular Communications: Channel Model and Physical Layer Techniques

Weisi Guo, *Member, IEEE*, Taufiq Asyhari, *Member, IEEE*, Nariman Farsad, *Member, IEEE*,  
H. Birkan Yilmaz, *Member, IEEE*, Bin Li, *Member, IEEE*, Andrew Eckford, *Senior Member, IEEE*, and  
Chan-Byoung Chae, *Senior Member, IEEE*

**Abstract**—This article examines recent research in molecular communications from a telecommunications system design perspective. In particular, it focuses on channel models and state-of-the-art physical layer techniques. The goal is to provide a foundation for higher layer research and motivation for research and development of functional prototypes. In the first part of the article, we focus on the channel and noise model, comparing molecular and radio-wave pathloss formulae. In the second part, the article examines, equipped with the appropriate channel knowledge, the design of appropriate modulation and error correction coding schemes. The third reviews transmitter and receiver side signal processing methods that suppress inter-symbol-interference. Taken together, the three parts present a series of physical layer techniques that are necessary to producing reliable and practical molecular communications.

**Index Terms**—Molecular communication, error correction coding, modulation, signal processing, channel modeling.

## I. INTRODUCTION

Molecular communication utilizes chemical molecules as an alternative carrier for information. Tiny devices (having dimensions in microns or less) may find molecular communications to be an efficient and low-cost alternative to high frequency electromagnetic based systems. Indeed, this is perhaps why molecular communications is prevalent in nature, both at the inter-organism and at the inter-cell scales. Being able to communicate at the nano-scale is important in unlocking the futuristic possibilities in nanotechnology. For some time now, the fundamental building blocks have been available for nanomachines. The transformative applications of nanotechnology have one feature in common: they involve not just single devices working independently, but swarms of devices working in concert (i.e., a network of nano-machines [1]). Thus before the promise of nanotechnology can be fully achieved, the problem that remains to be solved is communication among nanodevices. The economic potential is enormous: the present market size is at \$100 billion, growing at a projected 12.3% per year to \$200 billion in 2020.

The present capabilities of molecular communication are rather primitive. Indeed, the world's first molecular communication system capable of sending a text message has a volume

of roughly 100 cm<sup>3</sup>. Akin to spectral efficiency, the system has a *chemical efficiency* of 0.3 bits/s/chemical over a free-space distance of a few meters [2]. To have a significant impact, these systems will need to shrink and increase their data rates by two orders of magnitude. The data rate increase can be achieved by: designing tailor-made modulation and error correction coding schemes [3], and increasing the number of orthogonal chemical compounds [4] or carrier frequencies [5] used to carry information. Drawing on the current experience of designing radio-wave communication systems, a three-fold increase in data rate can be yielded with signal processing techniques. A similar gain is expected through molecular communications.

This article reviews recent state-of-the-art-physical layer (PHY) techniques that go towards increasing chemical efficiency. Figure 1 offers an illustration of a molecular communication system physical layer design, hardware implementation and the associated signal at each stage. The goal of this article is to provide a foundation for higher layer research, i.e., relay, full duplex, and multiple-input multiple-output (MIMO) [6], as well as motivation for further research and development.

## II. CHANNEL, PATHLOSS, AND NOISE MODELS

### A. Statistical Channel Models

1) *Hitting Distribution*: For a basic random walk process, a common equation of interest is the probability density function,  $h(d, t)$ , of emitted particles from a point source hitting a point in space. Indeed the equation determines the communication aspects of the channel for a given distance  $d$ , time  $t$ , and diffusivity  $D$ . The time dependent particle distribution function differs in different environments. The hitting distribution is given as:  $h(d, t) = \frac{1}{(4\pi Dt)^{a/2}} \exp(-d^2/4Dt)$ , where  $a$  is a constant. This model assumes that molecules that hit the receiver can undergo a random walk and hit the receiver again at a later point.

2) *Absorbing Receiver*: In nature, most of the receptor types remove the information molecules from the environment once they arrive at the receiver. This absorption or capture process is modeled by first hitting processes where molecules are removed from the environment after hitting the absorbing receiver. In a 1-dimensional (1-D) environment, the rate of capture is given by  $h_c(d, t) = \frac{d}{(4\pi Dt^3)^{1/2}} \exp(-d^2/4Dt)$  for distance  $d$  and time  $t$ . The probability of hitting to an absorbing receiver until time  $t$ , is  $F_c(d, t) = \text{erfc}\left(d/\sqrt{4Dt}\right)$ . Similarly, the fraction of hitting molecules to a perfectly

W. Guo is with the University of Warwick, UK. T. Asyhari is with the University of Bradford, UK. N. Farsad and A. Eckford are with York University, Canada. B. Li is with Beijing University of Posts and Telecommunications, China. H. B. Yilmaz and C.-B. Chae (Corresponding Author) are with Yonsei University, Korea. The work of C.-B. Chae was in part supported by the Basic Science Research Program (2014R1A1A1002186) funded by the Ministry of Science, ICT and Future Planning (MSIP), Korea, through the National Research Foundation of Korea.

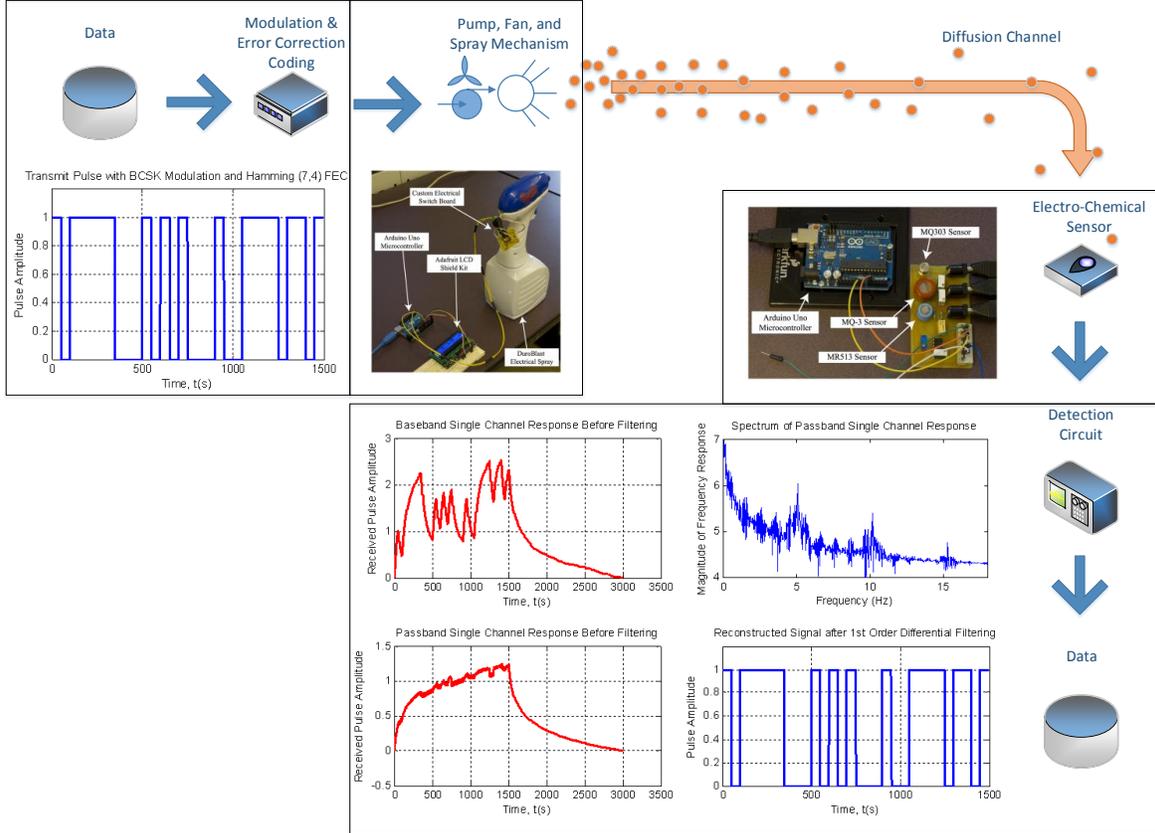


Fig. 1. Illustration of molecular communication system physical layer components, hardware implementation, and the state of the signal at each stage.

absorbing spherical receiver of radius  $r_r$  in a 3-D environment is [7]  $h_c(d, t) = \frac{r_r}{d+r_r} \frac{d}{(4\pi Dt^3)^{1/2}} \exp(-d^2/4Dt)$ , which is similar to the 1-D case. Note that there is a positive probability of not hitting the absorbing boundary for a diffusing particle in a 3-D environment even when time goes to infinity. For signal processing and system analysis, it is also useful to know the system transfer function in the complex  $s$ -domain. Assuming a linear time invariant (LTI) system, the transfer function for  $h_c(d, t)$  is given by [8]:

$$H_c(s) = \mathcal{L}[h_c(d, t)] = \exp\left(-d\sqrt{\frac{s}{D}}\right). \quad (1)$$

These molecular diffusion channels all exhibit two common time-domain features that differ it from radio wave based communications. They are: (1) the channel response's tail (similar to delay spread) is extremely long, and (2) the propagation time is also extremely long. These two features can be visualized in the *Baseband/Passband Channel Response* plots in Fig. 1, where we see that the tail of a pulse response can last several minutes and the time of arrival of the first pulse's peak is several seconds.

### B. Pathloss Models: Free-Space and Complex Environments

Pathloss models are concerned with the rate of energy decay over distance. Such models are vital for calculating the link budget of a telecommunication channel. We first review the

pathloss model in an unbounded free-space environment for both electromagnetic (EM) waves and molecular diffusion. Assuming the EM waves radiate from a point source, the pathloss at distance  $d$  is:  $(4\pi fd/c)^{-2}$ , where  $f$  is the frequency and  $c$  is the speed of light in a vacuum. The time of arrival of the wave is  $\tau = d/c$ . For molecules that diffuse from a point source in free-space, we propose two measures of loss:

- 1) *Peak molecular response*: The maximum pulse amplitude is detected at  $\tau = d^2/6D$ , which yields a peak hitting amplitude of [9]:  $\propto D/d^{-3}$ , and a peak absorbing amplitude of [7]:  $\propto \frac{r_r D}{(r_r + d) d^2}$ .
- 2) *Total molecular response*: This can be derived by counting the total number of molecules hitting the receiver over all time [9], [10] to be:  $(4\pi Dd)^{-1}$ , and absorbed by the receiver to be [7]:  $r_r/(r_r + d)$ .

Therefore, when comparing EM wave and molecular pathloss in free-space, one finds that:

- 1) **EM**: the pathloss is  $\propto d^{-2}$  and time-of-arrival is  $\propto d$ ;
- 2) **Molecular**: the total energy pathloss is  $\propto d^{-1}$ , the peak energy pathloss is  $\propto d^{-3}$ , and time-of-arrival is  $\propto d^2$ ;

That is to say, the peak molecular energy decays rapidly with distance, but the total molecular energy decays at a lower rate than EM-waves, but the time-of-arrival is quadratic with distance, which is significantly larger than EM-wave propagation.

In terms of more complex environments, a comparison was

recently made for pipe, knife-edge, aperture, and mesh channels [10]. It was found that unlike EM-wave signals, molecular propagation has the following propagation advantages:

- **No Diffraction Loss:** does not suffer from diffraction loss when in the shadow of objects;
- **Not Obstacle Size Sensitive:** is not restricted by cut-off frequency in pipe, aperture, and mesh environments;
- **Shortest Path Dominates:** molecular pathloss only depends on the shortest path distance between the transmitter and receiver.

These propagation advantages for molecular communications can be exploited in certain communication scenarios listed in [10]. The caveats to molecular communications are numerous. Besides the obvious slow nature of diffusion and its dependency on environmental changes such as temperature and flow, molecular communication is also at risk of degradation either from chemical reactions or bacteria consuming bio-molecules.

### C. Additive Noise and ISI Models

In terms of additive noise, the vast majority of molecular communications research has been modeling based, utilizing classic diffusion equations with an additive noise component. In a real pulse modulated molecular communication system, noise can arise for many reasons and it is difficult to take all of them into account or argue which may dominate. Examples of noise include, but are not limited to: inconsistent mechanical pulse emission, atmospheric contamination, physical disturbances, and receiver design [11]. Recent research has yielded two noise models for different scenarios:

- 1) **High-Drift Velocity:** In [12], based on the hardware prototype introduced in [2], it was found that the Gaussian distributed additive noise was a reasonable approximation in a high velocity turbulent flow channel, which is found in application areas such as gas pipes.
- 2) **Zero Drift:** As for a random walk diffusion channel with no induced drift velocity, the noise arises from the random arrival of molecules. As shown in [13], for  $n$  emitted molecules, the number of molecules arriving at the receiver is Binomial distributed  $\sim \mathcal{B}(n, h_c(d, t))$ . Such a channel is common in low pressure environments such as caves and tunnels.

The inter-symbol-interference (ISI) distribution has only been analyzed for a binary concentration modulated scheme in [13]. Given that the number of molecules arriving at the receiver is Binomial distributed  $\sim \mathcal{B}(n, h_c(d, t))$ , and that  $n$  is large and the probability  $h_c$  is neither close to 1 nor 0, the ISI can be approximated as a sum of random variables, each following a normal distribution. The underlying assumptions of this widely used model are: (i) the ISI is dominated only by the previous symbol, and (ii) the transmission distance is neither too close ( $h_c \rightarrow 1$ ) nor too far ( $h_c \rightarrow 0$ ). This distribution has been extensively used in subsequent work for capacity analysis and the evaluation of error correction code performances [14]. This remains an open area for more complicated  $M$ -ary modulation schemes or those that have a high symbol rate.

## III. MODULATION SCHEMES

### A. Baseband Modulation

1) *Basic Schemes:* For a fixed transmission distance of  $d$ , one can observe that there are essentially two main properties of the channel to modulate: the number of transmitter molecules  $M$ ; and the pulse delay time  $T_k$ , such that the channel response is  $h_k(t - T_k)$ . For an input of binary symbols  $a_k \in \mathcal{A} = \{0, 1\}$ ,  $k = 0, 1, \dots, \infty$ , the output of the baseband pulse modulator is  $M_k$ . Existing pulse modulation can be summarized as being one of the following [5]:

- Amplitude / Concentration Shift Keying (ASK or CSK), where the information is modulated into different levels of  $M_k$ , i.e., Binary CSK (BCSK):  $M_k \in \mathcal{M} = \{0, M\}$ .
- Frequency shift keying (FSK), where a sinusoidal pulse of a variable frequency  $f$  is emitted  $M_k(f_k) = M \sin(2\pi f_k t)$ , i.e., Binary FSK:  $f_k \in \mathcal{F} = \{0, f\}$ .
- Pulse Position Modulation (PPM), where the information is modulated into the bit delay time  $T$ , i.e., Binary PPM:  $T_k \in \mathcal{T} = \{0, T\}$ .

In addition, the chemical composition can be used to encode information, as it is common in nature, which is known as Molecule Shift Keying (MoSK) [4]. However, while MoSK is common in nature, the engineering complexity of synthesizing and detecting even a small number of precise chemical compositions is complex and expensive.

2) *ISI Reducing Schemes:* Molecular concentration shift keying (MCSK) is proposed in [15], in which two molecule types are used for consecutive slots at the transmitter for suppressing the interference. The transmitter uses molecule type  $A_1$  in odd time slots and molecule type  $A_2$  in even time slots. The difference with MoSK is that the keying depends on the amplitude not the type of the molecule. On the receiver side, according to slot number only the  $A_1$ - or  $A_2$ - type molecules are counted. Since the molecule types are different in the two subsequent time slots, interference due to the previous symbol is drastically reduced. Molecular Transition Shift Keying (MTSK) was also proposed as an energy-efficient ISI-reducing modulation technique for high data rates. MTSK employs multiple molecule types. The bit-0s in the sequence are encoded by the absence of the messenger molecules. The bit-1s are encoded by using two different types of molecules, where the choice of the molecule type depends on the value of the next symbol. If the next symbol is a transition from bit-1 to bit-0 (i.e., the final bit-1 for consecutive bit-1s) then the molecule type is changed.

### B. Passband Modulation

The aforementioned modulation schemes can be regarded as *baseband modulation*, whereby only a limited data rate can be achieved. The concept of a carrier wave, such as that associated with electromagnetic wave communications, is absent in molecular communications. A recent breakthrough detailed in [5] shows that by adding an initial longitudinal vibration at the transmitter at frequency  $f_c$ , the spectra of the received molecules will exhibit this frequency component. This is achieved by creating an oscillating transmitter that

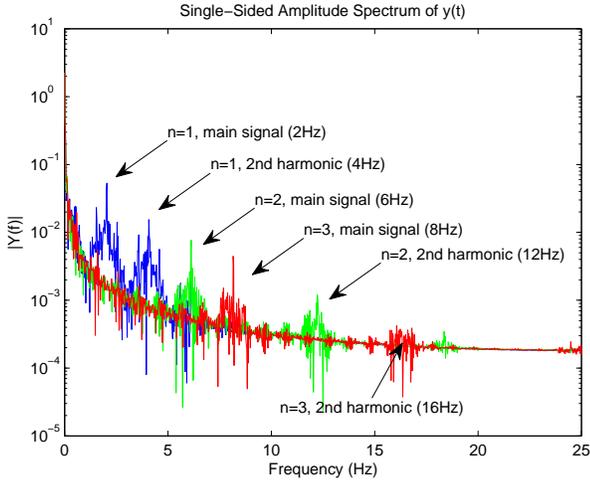


Fig. 2. Frequency response of  $N = 3$  molecular signals multiplexed over a single diffusion channel with BCSK baseband modulation.

varies the distance of transmission for each data stream, such that multiple data streams can be carried on different *carrier frequencies* (i.e.,  $f_{c,1}, f_{c,2}, \dots, f_{c,N}$ ). Figure 2 shows the magnitude of the received frequency response  $|Y(f)|$  of  $N = 3$  independent data channels. The data streams are transmitted via a common diffusion channel using the same chemical to a common receiver. It can be seen that the *baseband signal* is at 0 normalized frequency. The first harmonic of each signal can be seen distinctively, which is at the respective carrier frequencies. The distinctive frequency components can be separated using bandpass filters. Further research into this area will examine the capabilities regarding ISI suppression and the fundamental limits to multiplexing.

#### IV. FORWARD ERROR CORRECTION CODING

Forward-Error-Correction (FEC) coding is an important aspect of modern digital communications. In this section, we first examine how FEC codes can combat these two types of errors, namely: additive noise errors for CSK modulation schemes, and transposition errors for PPM modulation schemes; and we then compare the performance of different coding strategies.

##### A. Additive Noise Errors

In the presence of additive noise, the *Hamming distance* is commonly used as a criteria for code word design. Relatively simple block codes use added parity check bits to enable the correction of errors. Hamming codes are described as  $n = 2^m - 1$  length codes, where  $m$  is the number of parity check bits. The minimum distance of this type of block code is 3, allowing the detection of 2 errors and the correction of 1 error. In order to develop low energy nano-communication systems, minimum energy FEC codes were considered in [14]. When considering a BCSK modulation scheme, codewords with a lower weight (more ‘0’ bits) result in a lower energy consumption. Therefore, minimizing the average code weight is necessary to minimizing the energy expenditure. According to the results in [14], this coding scheme is able to achieve

significantly lower energy per bit (50% reduction), but will need large codeword lengths to achieve a similar level of Bit-Error Rate (BER) performance. More advanced codes such as Self-Orthogonal Convolution Codes (SOCCs) have also been developed recently and improve on Hamming codes to deliver a superior coding gain and lower energy expenditure.

##### B. Bit Transposition Errors

When block coding is considered, bit transposition may result into two sources of errors:

- 1) **Intra-Codeword Errors** due to permutation of bits position within a codeword (crossovers within the same codeword), and
- 2) **Inter-Codeword Interference** due to bits late/early arrival that affects other codewords (crossovers from the contiguous codewords).

1) *Intra-Codeword Errors*: Consider encoding each of  $M$  messages into  $n$ -bit block. A predominant coding technique is to select  $M$  out of  $2^n$  codewords that maximize the minimum pairwise Hamming distance. This approach is particularly relevant when bit distortion due to additive noises is the main source of errors. However, this may no longer be accurate when the dominant errors are due to bit transposition where instead of being distorted, bits exchange their positions. A key to deal with intra-codeword errors is to identify relevant attributes that are still preserved when bits within a codeword exchange positions. Hamming weight is a sensible choice that fulfills this requirement. The coding technique then follows by constructing a codebook that consists of words having distinct Hamming weights. We refer to the resulting code as a *distinct Hamming-weight (DHW) code*. An example of DHW code for  $n = 4$  and rate- $\frac{1}{2}$  is given by:  $\mathcal{C}_{dhw} = \{0000, 1000, 1100, 1110\}$ . The maximum coding rate of DHW code with block length  $n$  is given by:  $R_{weight} = \log_2(n + 1)/n$ .

2) *Inter-Codeword Errors*: Molecular coding (MoCo) distance function has been recently proposed as a new criterion for code design that tackles intra- and inter-codeword errors [3]. The MoCo distance of a pair of codewords  $(\mathbf{x}_i, \mathbf{x}_j)$  is given by  $d_{MoCo}(\mathbf{x}_i, \mathbf{x}_j) \triangleq -\log(\Pr\{\mathbf{x}_j|\mathbf{x}_i\})$ , where  $\Pr\{\mathbf{x}_j|\mathbf{x}_i\}$  is the transition probability from  $\mathbf{x}_i$  to  $\mathbf{x}_j$  that is defined by the random walks of the molecules. The MoCo code construction is essentially finding codewords that maximize the minimum pairwise MoCo distance. For  $n = 4$  and rate- $\frac{1}{2}$ , the MoCo code is given by:  $\mathcal{C}_{MoCo} = \{0000, 1000, 0010, 1110\}$ . While addressing intra-codeword errors is clear from the MoCo distance, the least significant bit ‘0’ in each codeword was believed to serve as a guard band that reduces inter-codeword interference. Unfortunately, the nature of exhaustive search for the best codebook and the lack of analytical expression of the error probability for good MoCo codes make it difficult to assess whether  $\mathcal{C}_{MoCo}$  will still be optimal for varying system and channel parameters.

3) *ISI-Free Code*: A well specified family of  $(n, k, \ell)$  ISI-free codes was proposed in [3] to address both intra codeword errors and inter codeword interference. Herein  $n$ ,  $k$  and  $\ell$

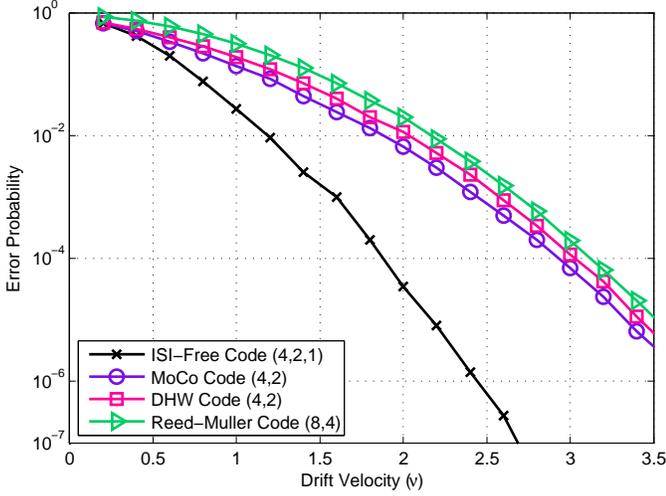


Fig. 3. Error probability of various coding schemes as a function of drift velocity  $v$  ( $\mu\text{m}/\text{s}$ ) in a PPM system:  $D = 1 \mu\text{m}^2/\text{s}$ ,  $d = 1 \mu\text{m}$ , and all the coding schemes have rate of  $1/2$ .

TABLE I  
(4,2,1) ISI-FREE CODE WITH CODEWORDS ASSIGNMENT.

Information Bits	Codeword if the last bit of the previous codeword is	
	bit '0'	bit '1'
00	0000	1111
01	0001	1000
10	0011	1100
11	0111	1110

denote the block length, the message length (in bits) and the correctable crossover level, respectively. Table I provides an example of ISI-free codes when  $(n, k, \ell) = (4, 2, 1)$ . As observed from Table I, intra-codeword errors are mitigated using codewords with distinct Hamming weights, similarly to the DHW code. The inter-codeword interference is suppressed by assigning two codewords for each message such that for two contiguous codewords, the last  $\ell$  bits of the first codeword are identical to the first  $\ell$  bits of the second codeword. Encoding messages 01, 10, 11 yields codewords 0001, 1100, 0111.

### C. Performance Comparison

Figure 3 demonstrates the performance of different coding schemes over diffusion-based molecular communications. It can be observed that Reed-Muller code (8,4) designed based on Hamming distance is inferior to other codes. For a single codeword assignment for each message, the MoCo code is superior to the DHW code since the MoCo code has been suggested to address both intra-codeword errors and inter-codeword interference. The main drawback of the MoCo code is the lack of structured construction. Among all the presented codes in Fig. 3, (4,2,1) ISI-free code has the best error performance. This code is able to guarantee free transposition errors up to  $\ell = 1$  level.

With the availability of structured construction and superior performance, ISI-free codes appear to be promising for

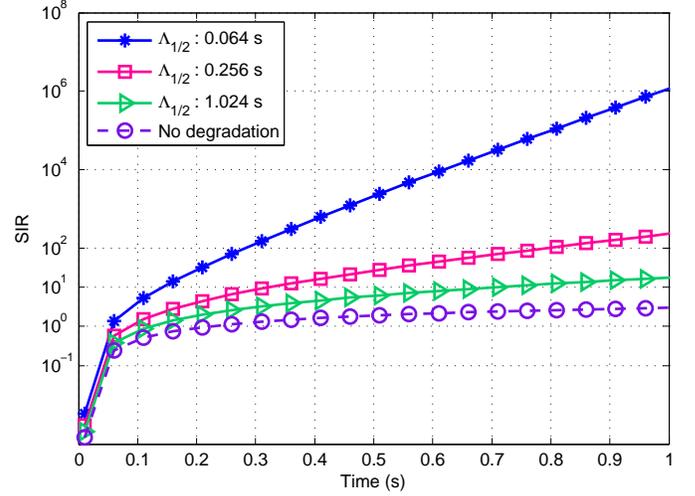


Fig. 4. SIR versus symbol duration for various  $\Lambda_{1/2}$  (half-life) values. Faster degradation clears stray molecules and reduces the amount of ISI in the system. ( $d = 4 \mu\text{m}$ ,  $r_r = 10 \mu\text{m}$ ,  $D = 79.4 \mu\text{m}^2/\text{s}$ ).

molecular communications. However, further developments are necessary to address their inherent rate inefficiencies. In the above example, due to the assignment of two codewords for a single message, the coding rate above  $\frac{n-1}{n}$  cannot be achieved for ISI-free codes. Some adjustments have been proposed in [3], but the rate improvement is still limited. Furthermore, the inherent code construction in [3] assumes bits '0' and '1' being transmitted using two molecule types. This is inefficient, not only due to the cost of using different types of chemical substance, but also due to the two molecule types that alternatively would have allowed a higher multiplexing data rate. In order to overcome these shortcomings, a key research direction would be comprehensive understanding of the trade-offs between rate and error probability using information-theoretic approaches.

## V. CHEMICAL TECHNIQUES FOR ISI SUPPRESSION

### A. Molecular Degradation

As mentioned previously, ISI arises from the long tail of the channel response. Therefore, it is logical to consider if chemical degradation can mitigate ISI. Molecular degradation is governed by  $E + S \xrightleftharpoons[k_{-1}]{k_1} ES \xrightarrow{k_p} E + P$ , where  $E$ ,  $S$ ,  $ES$ , and  $P$  denote the enzyme, substrate, enzyme-substrate compound, and the product, respectively. Reaction constants are denoted by  $k_1$ ,  $k_{-1}$ , and  $k_p$  [16]. Considering the nature of the molecular communication scenario simplifies the equation dynamics and leads to  $\frac{d[S]}{dt} = -\lambda[S]$ , where  $[S]$  denotes the concentration and  $\lambda$  is the rate of degradation. Instead of  $\lambda$ , sometimes half-life of the molecules,  $\Lambda_{1/2}$ , is considered in the equations. Having a very high degradation rate may cause molecules to degrade so fast that they cannot arrive at the target receiver. Therefore, for each symbol duration there exists an optimal degradation rate.

When the molecular degradation is taken into account, channel characteristics differ. Heavy tail property of the arriving molecules changes and results in a decrease in ISI with

the cost of requiring more molecules to be synthesized. The authors in [16] analyzed the effect of degradation for a given symbol duration  $t_s$ . A good metric for observing the effect of symbol duration is the signal-to-interference ratio (SIR). SIR is basically the ratio of molecules that arrive in the first symbol duration to the ratio of ISI molecules and given as  $\frac{F_c^\lambda(d, t_s)}{F_c^\lambda(d, t \rightarrow \infty) - F_c^\lambda(d, t_s)}$ . Figure 4 depicts SIR for various half-life values, where faster degradation clearly reduces the amount of stray molecules for a given molecular communication scenario. Note that the reduction in ISI comes with a cost of higher pathloss due to decomposed molecules. This effect is left for future investigations.

### B. Transmitter Side Pulse Shaping with Chemical Reaction

An alternative degradation method is outlined in [8], which uses reactive pulses. Assuming that the random walk channel is an LTI system, as mentioned earlier, the channel transfer function is found in Eq. (1). It can be shown, using either the principle of channel inversion or designing a window function, that the necessary transmit pulse is of the form<sup>1</sup>:  $x(t) = \delta(t) - \frac{x}{\sqrt{D}}t^{-3/2}$ , where the two parts of the equation can be interpreted as follows. The  $\delta(t)$  is the information bearing pulse. The negative decay element  $\propto -t^{-3/2}$  is a poison signal designed to chemically react with the information pulse and cancel the tail of the transient response at the receiver (positioned at distance  $x$ ). The constant of proportionality is  $\frac{x}{\sqrt{D}}$ , which means the channel state information at transmitter (CSIT) is needed. Specifically, this is the transmission range  $x$  and the channel's molecular diffusivity  $D$ . While simulation results in [8] show that the system response using pulse shaping will yield fewer errors, how we emit the signal is not obvious (given its divergent and undefined nature at  $t = 0$ ). This is a problem for future research to find practical pulse shapes.

## VI. RECEIVER SIDE SIGNAL PROCESSING

On the receiver side, equalizer filters can be used to eliminate the ISI effect. The received signal vector is normally expressed as  $\mathbf{r} = \mathbf{Y}\mathbf{a} + \mathbf{w}$ , where  $\mathbf{r} = [r_0, r_1, \dots, r_K]^T$ ,  $\mathbf{Y}$  is the circulant channel matrix constructed from  $\mathbf{y}$ , and  $\mathbf{w}$  is noise vector.

For a coherent MAP scheme, the accurate estimation of unknown channel impulse response (CIR) should be acquired firstly. To reduce the complexity, a sequential estimation strategy will be used in practice. For the i.i.d. information source, the MAP scheme is also equal to the maximum likelihood (ML) method. Another widely used sub-optimal detector is based on the MMSE criterion, which minimizes the covariance matrix of the detection errors. Unlike the additive Gaussian noise, variance of the received signal is history dependent and equalizer tap coefficients must be adjusted according to previous detected bits for MMSE filters. Hence, the tap count affects the computational complexity. On the other hand, the computational complexity of the proposed decision feedback

<sup>1</sup>The validity of this proof only exists for when the distance is small or the diffusivity is high ( $x \ll \sqrt{D}$ ).

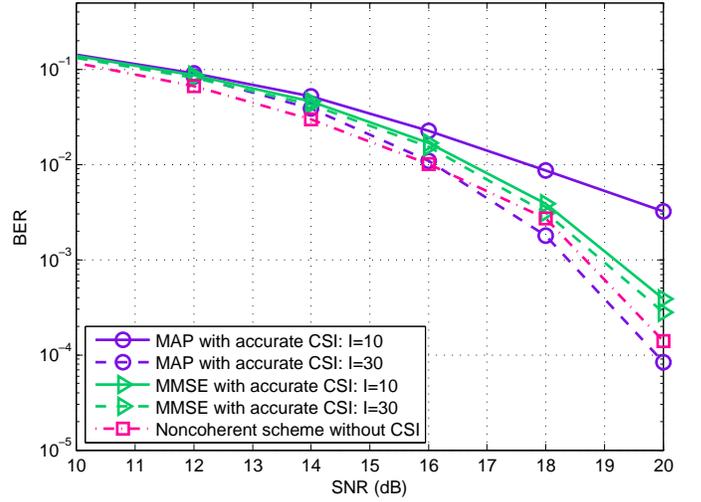


Fig. 5. BER performance of coherent detection schemes (MAP and MMSE) and the non-coherent scheme under memory-length  $I$ . Parameters used are: 3D diffusion channel, distance  $2m$ , diffusivity  $0.1\text{cm}^2/\text{s}$ , receiver absorption radius  $2cm$ , and the symbol period is  $4\text{s}$ .

filter (DFF) is independent of the tap count. However, this requires more memory  $I$  to achieve the same BER.

It is recognized that, in sharp contrast to wireless communications, there are two inherent challenges in molecular communications. First, the CIR cannot be easily estimated without a pilot channel, especially in the presence of channel disturbances due to air flow and temperature variations. Even in the presence of a pilot channel, the coherence period of the channel is usually small, while the channel delay is large. Second, molecular communication is targeted towards nano-scale systems, which require a low computational complexity system to reduce hardware costs and energy expenditure. The effectiveness of the coherent MAP or MMSE algorithms in eliminating ISI largely depends on both accurate CSI and complicated matrix or polynomial (likelihood) operations. Based on the above considerations, coherent detectors are less attractive in the context of low-complexity and low-power nano-machines.

Difference-based signal detectors can be used as a simple non-coherent and low complexity detector. In the hardware prototype found in [2], a tunable difference detector is used to compare the molecular concentration at two time steps. In comparing the performance of the aforementioned non-coherent difference detector and the previous coherent detectors, we plot the BER as a function of the signal-to-noise-ratio (SNR). The SNR is defined as the ratio of the peak signal power and the Gaussian noise due to stochastic molecule arrival [13]. As shown in Fig. 5, the BER performance of a non-coherent detector without CSI can exceed existing coherent and more complex MMSE and MAP schemes that use a memory-length of  $I = 10$ .

## VII. OPEN CHALLENGES AND CONCLUSIONS

In summary, this article has reviewed state-of-the-art physical layer techniques developed over the past few years. Many

challenges exist for a new form of wireless communications, especially when the fundamental information carrier has changed. In terms of the diffusion channel, as reviewed in Section II, diffusion is inherently slow and exhibits a long-tail in the pulse response. This means delay-tolerant communication techniques is very appropriate, as well as a need for ISI suppression signal processing, which was extensively reviewed. Whilst applying our current knowledge of radio-wave communication techniques is possible, direct application of current state-of-the-art may not have optimal results. Careful consideration of the underlying challenges is needed and this is shown to be the case in Section IV with error correction codes. Going forwards, a key barrier to high data rate molecular communications is the lack of bandwidth. Whilst early research has shown that orthogonal bands can be induced via longitudinal vibration (see Section II), there is an urgent need to produce a reliable and scalable bandwidth resource using either chemical or other physical means.

Beyond the physical layer, several other important challenges are being addressed that we did not review, such as: security, energy efficiency, synchronization, and mobility management. These are all important features that must be considered in future commercial ready implementations, and indeed, they are being considered under the current IEEE P1906.1 standardization efforts.

#### REFERENCES

- [1] I. Akyildiz, F. Fekri, R. Sivakumar, C. Forest, and B. Hammer, "MoNaCo: fundamentals of molecular nano-communication networks," *IEEE Wireless Communications*, vol. 19, no. 5, pp. 12–18, Oct. 2012.
- [2] N. Farsad, W. Guo, and A. Eckford, "Tabletop molecular communication: Text messages through chemical signals," *PLoS ONE*, vol. 8, Dec. 2013.
- [3] P. Shih, C. Lee, P. Yeh, and K. Chen, "Channel codes for reliability enhancement in molecular communication," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 31, pp. 857–867, Dec. 2013.
- [4] N.-R. Kim and C.-B. Chae, "Novel modulation techniques using isomers as messenger molecules for nano communication networks via diffusion," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 31, no. 12, pp. 847–856, Dec. 2013.
- [5] W. Guo, B. Li, S. Wang, and W. Liu, "Molecular communications with longitudinal carrier waves: baseband to passband modulation," *to appear in IEEE Communications Letters*, 2015.
- [6] B. Koo, C. Lee, H. B. Yilmaz, N. Farsad, A. Eckford, and C.-B. Chae, "Molecular mimo: From theory to prototype," *submitted to IEEE Journal on Selected Areas in Communications (JSAC)*, 2015.
- [7] H. B. Yilmaz, A. C. Heren, T. Tugcu, and C.-B. Chae, "Three-dimensional channel characteristics for molecular communications with an absorbing receiver," *IEEE Communications Letters*, vol. 18, no. 6, pp. 929–932, June 2014.
- [8] S. Wang, W. Guo, and M. McDonnell, "Transmit pulse shaping for molecular communications," in *Proc. IEEE Conference on Computer Communications (INFOCOM) - Workshops*, May 2014, pp. 209–210.
- [9] I. Llaster, A. Cabellos-Aparicio, M. Pierobon, and E. Alarcon, "Detection techniques for diffusion-based molecular communication," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 31, no. 12, pp. 726–734, Jan. 2014.
- [10] W. Guo, C. Mias, N. Farsad, and J. L. Wu, "Molecular versus electromagnetic wave propagation loss in macro-scale environments," *to appear in IEEE Transactions on Molecular Biological and Multiscale Communications (T-MBMC)*, vol. 1, no. 1, 2015.
- [11] M. Pierobon and I. F. Akyildiz, "Diffusion-based noise analysis for molecular communication in nanonetworks," *IEEE Transactions on Signal Processing*, vol. 59, no. 6, pp. 2532–2547, June 2011.
- [12] N. Farsad, N. Kim, A. Eckford, and C.-B. Chae, "Channel and noise models for nonlinear molecular communication systems," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 32, no. 12, pp. 2392–2401, Dec. 2014.
- [13] M. S. Kuran, H. B. Yilmaz, T. Tugcu, and B. O. Edis, "Energy model for communication via diffusion in nanonetworks," *Nano Communication Networks*, vol. 1, no. 2, pp. 86–95, April 2010.
- [14] C. Bai, M. Leeson, and M. Higgins, "Minimum energy channel codes for molecular communications," *Electronics Letters*, vol. 50, no. 23, pp. 1669–1671, Nov. 2014.
- [15] H. Arjmandi, A. Gohari, M. N. Kenari, and F. Bateni, "Diffusion-based nanonetworking: A new modulation technique and performance analysis," *IEEE Communications Letters*, vol. 17, no. 4, pp. 645–648, April 2013.
- [16] A. C. Heren, H. B. Yilmaz, C.-B. Chae, and T. Tugcu, "Effect of degradation in molecular communication: Impairment or enhancement," *to appear in IEEE Transactions on Molecular Biological and Multiscale Communications (T-MBMC)*, vol. 1, no. 1, 2015.

PLACE  
PHOTO  
HERE

**Weisi Guo** is an Assistant Professor at the University of Warwick. He received his M.Eng., M.A. and Ph.D. degrees from the University of Cambridge. He has published over 50 IEEE papers in recent years and won several academic awards, including being a finalist of both the 2014 Bell Labs Prize and 2015 IET Innovation Award. His research interests are in the areas of 4G/5G cellular networks, smart cities, and molecular communications.

PLACE  
PHOTO  
HERE

**Taufiq Asyhari** (M13) received the Ph.D. degree in Engineering (Information Engineering) from the University of Cambridge, UK in 2012 and the B.Eng. (Hons.) degree in Electrical and Electronic Engineering from Nanyang Technological University. Since 2014 he has been a Lecturer in Computing of the School of Electrical Engineering and Computer Science, University of Bradford, U.K. Prior assuming this position, he has held positions at the Microelectronics and Information Systems Research Centre and Department of Electrical and Computer Engineering at National Chiao Tung University, Hsinchu, Taiwan, and Bell Labs - Alcatel-Lucent, Stuttgart, Germany. His research interests are in the area of information theory, communication and coding theory, queueing theory and signal processing techniques with applications to wireless and nano-molecular networks.

Dr. Asyhari has recently been awarded Associate Fellow status of the Higher Education Academy (HEA), UK. He received the Best Paper Award at the 11th ISWCS in 2014, Starting Grant from the National Science Council (NSC) of Taiwan in 2013 and funding from Cambridge Trust (Yousef Jameel Scholarship) in 2008/2011.

PLACE  
PHOTO  
HERE

**Nariman Farsad** received his M.Sc. and Ph.D. degrees in computer science and engineering from York University, Toronto, ON, Canada in 2010 and 2015, respectively. He is currently a Postdoctoral Fellow with the Department of Electrical Engineering at Stanford University, where he is a recipient of Natural Sciences and Engineering Research Council of Canada (NSERC) Postdoctoral Fellowship. Nariman has won the second prize in 2014 IEEE ComSoc Student Competition: Communications Technology Changing the World, the best demo award at INFOCOM2015, and was recognized as a finalist for the 2014 Bell Labs Prize. Nariman has been an Area Associate Editor for IEEE Journal of Selected Areas of Communication—Special Issue on Emerging Technologies in Communications, and a Technical Reviewer for a number of journals including IEEE Transactions on Signal Processing, and IEEE Transactions on Communication. He was also a member of the Technical Program Committees for the ICC2015, BICT2015, and GLOBECOM2015.

PLACE  
PHOTO  
HERE

**H. Birkan Yilmaz** received his M. Sc. and Ph.D. degrees in Computer Engineering from Bogazici University in 2006 and 2012, respectively, and his B.S. degree in Mathematics in 2002. Currently, he works as a post-doctoral researcher at Yonsei Institute of Convergence Technology, Yonsei University, Korea. He holds TUBITAK National Ph.D. Scholarship and is a member of TMD (Turkish Mathematical Society). His research interests include cognitive radio, spectrum sensing, molecular communications, and detection and estimation theory.

PLACE  
PHOTO  
HERE

**Chan-Byoung Chae** (SM'12) is Associate Professor in the School of Integrated Technology, Yonsei University, Korea. Before joining Yonsei University, he was with Bell Labs, Alcatel-Lucent, Murray Hill, New Jersey, as a Member of Technical Staff, and Harvard University, Cambridge, Massachusetts, as a Post-doctoral Research Fellow. He received his Ph.D. degree in Electrical and Computer Engineering from The University of Texas at Austin, TX, USA in 2008.

He was the recipient/co-recipient of the IEEE INFOCOM Best Demo Award (2015), the IEIE/IEEE Joint Award for Young IT Engineer of the Year (2014), the KICS Haedong Young Scholar Award (2013), the IEEE Signal Processing Magazine Best Paper Award (2013), the IEEE ComSoc AP Outstanding Young Researcher Award (2012), the IEEE Dan. E. Noble Fellowship Award (2008), two Gold Prizes (1st) in the 14th/19th Humantech Paper Contest. He currently serves as an Editor for *the IEEE Trans. on Wireless Comm.*, *the IEEE/KICS Jour. Comm. Nets*, and *the IEEE Trans. on Molecular, Biological, and Multi-scale Comm.*

PLACE  
PHOTO  
HERE

**Bin Li** received the Bachelor's degree in electrical information engineering from Beijing University of Chemical Technology (BUCT) in 2007, the Ph.D. degree in communication and information engineering from Beijing University of Posts and Telecommunications (BUPT) in 2013. He joined BUPT since 2013, as a lecture of the School of Information and Communication Engineering (SCIE). His current research interests are focused on statistical signal processing algorithms and applications in cognitive radios (CRs), wireless sensor networks, millimeter-

wave (mm-Wave) communications and nano-scale molecular communications. He has published more than 60 journal and conference papers. He received 2015 IEEE WCSP Best Paper Award, 2011 IEEE ChinaCom Best Paper Award, 2010 and 2011 BUPT Excellent Ph.D Student Award Foundation. He serves as the Co-Chair of the Technical program Committee of the Signal Processing for Communications Symposium of the 2016 IEEE International Conference on Computing, Networking and Communications (IEEE-ICNC'16).

PLACE  
PHOTO  
HERE

**Andrew Eckford** (SM'15) is an Associate Professor in the Department of Electrical Engineering and Computer Science at York University, Toronto, Ontario. He received the B.Eng. degree from the Royal Military College of Canada in 1996, and the M.A.Sc. and Ph.D. degrees from the University of Toronto in 1999 and 2004, respectively, all in Electrical Engineering. Andrew held postdoctoral fellowships at the University of Notre Dame and the University of Toronto, prior to taking up a faculty position at York in 2006. Andrew has also served

as a co-founder and CEO of Engage Biomechanics Inc., a spin-out company from York University. Andrew's research interests include the application of information theory to nonconventional channels and systems, especially the use of molecular and biological means to communicate. Andrew's research has been covered in media including The Economist and The Wall Street Journal.