Original citation:
Lu, Yi, Higgins, Matthew D., Leeson, Mark S., Chen, Yunfei and Jennings, P. A. (Paul A.). (2016) A revised look at the effects of the Channel Model on molecular communication system. Micro & Nano Letters.

Permanent WRAP URL:
http://wrap.warwick.ac.uk/82230

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions. Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:
"This paper is a postprint of a paper submitted to and accepted for publication in (…) and is subject to Institution of Engineering and Technology Copyright. The copy of record is available at IET Digital Library"

A note on versions:
The version presented here may differ from the published version or, version of record, if you wish to cite this item you are advised to consult the publisher's version. Please see the 'permanent WRAP URL' above for details on accessing the published version and note that access may require a subscription.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

warwick.ac.uk/lib-publications

A Revised Look at the Effects of the Channel Model on Molecular Communication System

Yi Lu[1], Matthew D. Higgins[2], Mark S. Leeson[1], Yunfei Chen[1] and Paul A. Jennings[2]

[1] *School of Engineering, University of Warwick, Coventry, CV4 7AL, UK*
[2] *WMG, University of Warwick, Coventry, CV4 7AL, UK*
E-mail: m.higgins@warwick.ac.uk

Molecular communications, where information is passed between the Transmitter (T$_X$) and the Receiver (R$_X$) via molecules is a promising area with vast potential applications. However, the infancy of the topic within the overall taxonomy of communications has meant that to date, several channel models are in press, each of which is applied under various constraints and/or assumptions. Amongst them is that the arrival of molecules in different time slots can be, or is, considered as independent events. In practice, this assumption is not accurate, as the molecules arriving in the previous slot reduce the possible number of molecules in the next slot and hence make them correlated. In this letter, we analyze a more realistic performance of a molecular communication assuming correlated events. The key result shown, is that the widely used model assuming independent events significantly overestimates the error rates in the channel. This result is thus critical to researchers who focus on energy use at the nano-scale, as the new analysis provides a more realistic prediction and therefore, less energy will be needed to attain a desired error rate, increasing system feasibility.

1. Introduction: Molecular communication is a fast growing area aiming to utilise molecules to transmit information between nano-scale devices or machines. To model this process, and subsequently the performance of the system as a whole, knowledge of the channel is thus paramount. Ideally this modelling process should not only be accurate, but also efficient. A quick survey will show that the most popular way is to approximate the number of received molecules as a Binomial distribution and consider the capture probability of those **molecules** as the success probability for information transmission.

The key work here can be found in [1] where the model was presented under the assumption that the number of molecules transmitted in one time slot, but received in different successive time slots, was independent. This prevailing assumption was carried forward and considered to be fair within the community for quite some time as seen by numerous key papers from esteemed authors in [2–9].

However, through careful consideration of the actions of the receiver, for the MC system with an absorbing receiver, where molecules will be absorbed when they arrive at the receiver, the assumption of independence between the numbers of molecules in different time slots cannot be valid, as the removal of the molecule means the removed molecule cannot be present later. In other words, the number of molecules received in one time slot does affect that in the following time slots and thus they are dependent events.

Notably, the authors of [10, 11] have already presented a model which does consider the number of molecules received in the current time slot, taking into account those absorbed in the previous slots. The papers however, confidently presented the observation and subsequent model without much theoretical derivation. Furthermore, no direct comparisons were presented to previously published results. Thus no bounds exist on how 'correct' or 'accurate' their newer dependent model is. This subtlety, or understatement, is perhaps a reason why the old model is still in use in current literature.

This letter aims to fill this gap by making the following contributions.

Firstly, a comprehensive analysis of the system performance in terms of Bit Error Rate (BER) under the assumption of correlation
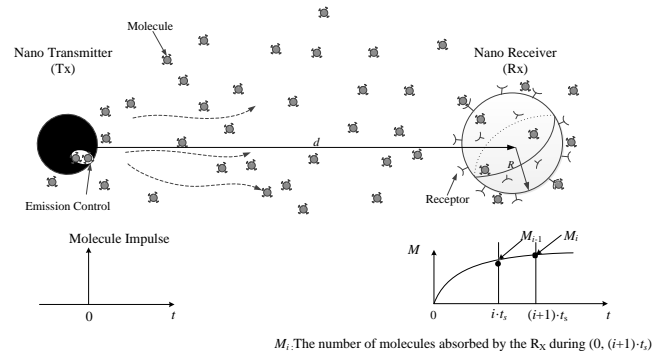

Fig. 1 *Molecular communication systems.*

between numbers of molecules in different time slots is presented. This analysis includes an explanation, as to the issue of dependence along with a full proof from the first principles of three-dimensional diffusion. The proof in particular is documented here as the subtleties are not fully found elsewhere and is thus part of the problem of why this newer model is not being adopted.

Secondly, an arbitrary Intersymbol Interference (ISI) length is introduced during the theoretical derivation to maximize the generality of the analysis. This also is a further extension beyond the work of [10] which considers only one past time slot and [11] which does not specifically use an ISI length but an Interference to Total Received Molecule Ratio [12].

Thirdly, the Binomial distribution is then approximated by both the Poisson and Normal distributions such that the BER expressions for both approximations can be provided. Using these key results, not only is a comparison between approximations shown, but these are also compared with the previous model. Completeness is then fulfilled by also including results derived via simulation.

These contributions thus allow the reader to clarify the theory behind the correlation between events as well as being able to quantify the accuracy of work using any of the approximations.

2. Diffusion-based model: The three-dimensional diffusion-based molecular communication system is depicted in Fig. 1 where the

1

propagation of molecules from the transmitter (T_X) to the receiver (R_X) is modeled by Brownian motion. The capture probability at the R_X can be obtained as [13]:

$$P_{ca}(d,t) = \frac{R}{d}\text{erfc}\left((d-R)/\sqrt{4Dt}\right), \qquad (1)$$

where R is the radius of the R_X, d is the distance between the T_X and the R_X, and D is the diffusion coefficient.

In this channel, the transmitted information is represented in binary form and conveyed in consecutive time slots. In each time slot, to transmit '1', a specific number of molecules, m, is released from the T_X, and to transmit '0', no molecules are released from the T_X. The duration of each time slot is denoted as t_s. Once the molecules arrive at the R_X, they will be captured and eliminated from the environment. By counting the number of absorbed molecules, the R_X determines whether a '1' and '0' was transmitted. However the molecules released at T_X cannot be guaranteed to reach the receiver within one time slot, and any remaining molecules may arrive at the R_X in future time slots. These molecules cause ISI.

Consider that m molecules are released at the start of the current time slot, and that the number of molecules received by R_X in the current time slot for the current molecular signal is represented as N_0. The number of molecules received from the previous i^{th} symbol in the current time slot is represented as N_i, where $i = 1, 2, …, I$. I is called the ISI length.

The analysis of N_0 for the previous and new model is the same where N_0 follows a Binomial distribution [1]:

$$N_0 \sim \mathcal{B}(m, P_{ca,0}), \qquad (2)$$

if m is large enough, a Binomial distribution can be approximated by a Normal distribution, thus:

$$N_0 \sim \mathcal{N}\left(mP_{ca,0}, mP_{ca,0}(1-P_{ca,0})\right), \qquad (3)$$

where $P_{ca,0} = P_{ca}(d, t_s)$.

The difference between the previous channel model and the new channel model is the mathematical modelling of N_i.

A) The previous model

In the previous model, the number of molecules received in different time slots from the same transmission was considered as independent. As shown in Fig. 1, M_i and M_{i-1} are assumed independent. Under this assumption, the expression of N_i can be derived as:

$$N_i = M_i - M_{i-1} \sim \mathcal{N}(\eta_i, \varsigma_i), \qquad (4)$$

where M_i is the number of molecules absorbed by the R_X during $(0, (i+1)\cdot t_s)$, $M_i \sim \mathcal{N}(mP_{ca,i}, mP_{ca,i}(1-P_{ca,i}))$, $M_{i-1} \sim \mathcal{N}(mP_{ca,i-1}, mP_{ca,i-1}(1-P_{ca,i-1}))$, $P_{ca,i} = P_{ca}(d, (i+1)t_s)$ for $i = 1, 2, …, I$, $\eta_i = m(P_{ca,i} - P_{ca,i-1})$ and $\varsigma_i = m(P_{ca,i}(1-P_{ca,i}) + P_{ca,i-1}(1-P_{ca,i-1}))$.

Thus, the total number of molecules received in the current time slot, N_R, comprises the number of received molecules in current time slot for the current molecular signal, and the number of received molecules for all I previous molecular signals in the current time slot:

$$\begin{aligned}N_R &= a_k N_0 + \sum_{i=1}^{I} a_{k-i} N_i \\ &\sim \mathcal{N}\left(a_k mP_{ca,0} + \sum_{i=1}^{I} a_{k-i}\eta_i, \ a_k mP_{ca,0}(1-P_{ca,0}) + \sum_{i=1}^{I} a_{k-i}\varsigma_i\right),\end{aligned} \qquad (5)$$

where $\{a_{k-i}, i = 0, 1, …, I\}$ represents the transmitted binary sequence which includes the current and all previous I symbol.

B) The new model

As mentioned in Section I, the number of molecules received in different time slots from the same transmission cannot be independent due to absorption.

Considering those m molecules released at the start of ith time slot before the current one, thus the probability density function of $N_i = y$ is given by:

$$\begin{aligned}P_r(N_i = y) &= \sum_{x=0}^{m} P_r(N_i = y/M_{i-1} = x)P_r(M_{i-1} = x) \\ &= \sum_{x=0}^{m}\binom{m-x}{y}q^y(1-q)^{m-x-y}\binom{m}{x}p^x(1-p)^{m-x} \\ &= (q(1-p))^y \sum_{x=0}^{m-y}\binom{m-x}{y}\binom{m}{x}p^x((1-p)(1-q))^{m-y-x} \\ &= (q(1-p))^y \sum_{x=0}^{m-y}\binom{m}{y}\binom{m-y}{x}p^x((1-p)(1-q))^{m-y-x} \\ &= \binom{m}{y}(q(1-p))^y \sum_{x=0}^{m-y}\binom{m-y}{x}p^x((1-p)(1-q))^{m-y-x},\end{aligned} \qquad (6)$$

and then by applying Binomial theorem [14], we obtained:

$$\begin{aligned}P_r(N_i = y) &= \binom{m}{y}(q(1-p))^y(p+(1-p)(1-q))^{m-y} \\ &= \binom{m}{y}(q(1-p))^y(1-q(1-p))^{m-y}.\end{aligned} \qquad (7)$$

where $p = P_{ca,i-1} = P_{ca}(d, i \cdot t_s)$, and $q = (P_{ca,i} - P_{ca,i-1})/(1 - P_{ca,i-1})$.

It obviously shows that N_i follows a Binomial distribution:

$$\begin{aligned}N_i &\sim \mathcal{B}(m, q(1-p)) \\ &= \mathcal{B}(m, P_{ca,i} - P_{ca,i-1}).\end{aligned} \qquad (8)$$

Due to the ISI, the sum of the influence from previous symbols must be considered in the current time slot. However, using the Binomial model for calculating the total number of molecules has a high computational requirement since there is a need to sum the Binomial variables. Thus, the Binomial distribution can be approximated as a Normal distribution depending on the conditions (the number of molecules per bit and the success probability) [1], [5], [10], [11]. Thus, N_0 and N_i can be approximated using the Normal approximation N_{0_norm}, N_{i_norm} and the Poisson approximation N_{0_pois}, N_{i_pois}, respectively:

$$N_{0\_norm} \sim \mathcal{N}\left(mP_{ca,0}, mP_{ca,0}(1-P_{ca,0})\right), \qquad (9)$$

$$\begin{aligned}N_{i\_norm} &\sim \mathcal{N}\left(m(P_{ca,i}-P_{ca,i-1}), m(P_{ca,i}-P_{ca,i-1})(1-P_{ca,i}+P_{ca,i-1})\right) \\ &\sim \mathcal{N}(\varpi_i, \gamma_i),\end{aligned} \qquad (10)$$

it can be seen that N_{i_norm} is different from that in (4) and thus the derivation of (8) represents contribution.

$$N_{0\_pois} \sim \mathcal{P}(mP_{ca,0}), \qquad (11)$$

$$N_{i\_pois} \sim \mathcal{P}(m(P_{ca,i}-P_{ca,i-1})), \qquad (12)$$

where $\varpi_i = m(P_{ca,i} - P_{ca,i-1})$ and $\gamma_i = m(P_{ca,i} - P_{ca,i-1})(1 - P_{ca,i} + P_{ca,i-1})$.

Overall, the total number of molecules received in one time slot for Normal approximation, N_{R_norm} and Poisson approximations, N_{R_pois}, can be obtained as:

2

$$N_{R\_norm} = a_k N_{0\_norm} + \sum_{i=1}^{I} a_{k-i} N_{i\_norm}$$

$$\sim \mathcal{N}\left(a_k m P_{ca,0} + \sum_{i=1}^{I} a_{k-i}\varpi_i, \ a_k m P_{ca,0}\left(1 - P_{ca,0}\right) + \sum_{i=1}^{I} a_{k-i}\gamma_i \right). \tag{13}$$

As the transmission symbols are in binary form, the value of a_{k-i} can only be 0 or 1, thus N_{R_pois} can be computed as:

$$N_{R\_pois} = a_k N_{0\_pois} + \sum_{i=1}^{I} a_{k-i} N_{i\_pois}$$

$$\sim \mathcal{P}\left(a_k m P_{ca,0} + \sum_{i=1}^{I} a_{k-i} m\left(P_{ca,i} - P_{ca,i-1}\right) \right). \tag{14}$$

3. Performance analysis: At the R_X, the information is determined by comparing the number of received molecules with a pre-designed optimal threshold, τ. When the number of received molecules exceeds τ, the symbol is decoded as a '1', otherwise, decoded as a '0'. The optimal threshold τ can be determined by finding the minimum BER for all possible values of τ as $\tau \in [1, m]$.

An error occurs when there is a difference between the symbol that was sent at the T_X and received at the R_X in the same time slot. It can be represented in two cases: when a '0' is transmitted, but a '1' is received and when a '1' is transmitted, but a '0' is received. Considering the ISI, the different error patterns may be obtained by the different permutations of the previous information symbols $\{a_{k-i}, i = 1, 2, \ldots, I\}$. So the number of the error patterns is 2^I.

For the first case, the error probability of the Normal and Poisson approximations for the error pattern j, $P_{N\_e01,j}$, $P_{P\_e01,j}$ can be obtained as:

$$P_{N\_e01,j} = p_{tx}{}^{\alpha_j}\left(1 - p_{tx}\right)^{I-\alpha_j} P\left(N_{R\_norm,j} > \tau\right)$$

$$= p_{tx}{}^{\alpha_j}\left(1 - p_{tx}\right)^{I-\alpha_j} \Phi\left(\frac{\mu_{01,j} - \tau}{\sigma_{01,j}} \right), \tag{15}$$

$$P_{P\_e01,j} = p_{tx}{}^{\alpha_j}\left(1 - p_{tx}\right)^{I-\alpha_j} P\left(N_{R\_pois,j} > \tau\right)$$

$$= p_{tx}{}^{\alpha_j}\left(1 - p_{tx}\right)^{I-\alpha_j} \left(1 - Q\left(\tau + 1, \lambda_{01,j}\right)\right), \tag{16}$$

where:

$$\mu_{01,j} = \sum_{i=1}^{I} a_{k-i,j}\varpi_i \ , \ \sigma_{01,j} = \sqrt{\sum_{i=1}^{I} a_{k-i,j}\gamma_i}.$$

$$\lambda_{01,j} = \sum_{i=1}^{I} a_{k-i,j} m\left(P_{ca,i} - P_{ca,i-1}\right). \tag{17}$$

p_{tx} is the transmission probability of '1'. $P(N_{R\_norm,j} > \tau)$ is the probability of $N_{R\_norm,j} > \tau$, and α_j is the number of '1's in the error pattern j, $j = 1, 2, \ldots, 2^I$. $\Phi(\cdot)$ and $Q(\cdot)$ are the cumulative distribution function of standard Normal distribution and regularized gamma function, respectively.

Conversely, the error probability for the second case can be obtained by:

$$P_{N\_e10,j} = p_{tx}{}^{\alpha_j}\left(1 - p_{tx}\right)^{I-\alpha_j} P\left(N_{R\_norm,j} \le \tau\right)$$

$$= p_{tx}{}^{\alpha_j}\left(1 - p_{tx}\right)^{I-\alpha_j} \Phi\left(-\frac{\mu_{10,j} - \tau}{\sigma_{10,j}} \right), \tag{18}$$

$$P_{P\_e10,j} = p_{tx}{}^{\alpha_j}\left(1 - p_{tx}\right)^{I-\alpha_j} P\left(N_{R\_pois,j} \le \tau\right)$$

$$= p_{tx}{}^{\alpha_j}\left(1 - p_{tx}\right)^{I-\alpha_j} Q\left(\tau + 1, \lambda_{10,j}\right), \tag{19}$$

where:

$$\mu_{10,j} = m P_{ca,0} + \sum_{i=1}^{I} a_{k-i,j}\varpi_i, \ \sigma_{10,j} = \sqrt{m P_{ca,0}\left(1 - P_{ca,0}\right) + \sum_{i=1}^{I} a_{k-i,j}\gamma_i}.$$

$$\lambda_{10,j} = m P_{ca,0} + \sum_{i=1}^{I} a_{k-i,j} m\left(P_{ca,i} - P_{ca,i-1}\right). \tag{20}$$

Thus, the BER of the system, P_e, with the new channel model can be derived as:

$$P_e = p_{tx} P_{e10} + \left(1 - p_{tx}\right) P_{e01}$$

$$= p_{tx} \sum_{j=1}^{2^I} P_{N/P\_e10,j} + \left(1 - p_{tx}\right) \sum_{j=1}^{2^I} P_{N/P\_e01,j}, \tag{21}$$

where $P_{N/P\_e10,\ j} = P_{N\_e10,\ j}$ or $P_{P\_e10,\ j}$ and $P_{N/P\_e01,\ j} = P_{N\_e01,\ j}$ or $P_{P\_e01,j}$. These closed-form analytical expressions are new results that are not available in the literature.

4. Numerical results and discussion: In this section, both simulation and theoretical results are presented. To compute the BER, the information at the T_X is encoded using a certain number of molecules, and these molecules are released as an impulse at the T_X, where the transmission of molecules from the T_X to the R_X is governed by the laws of Brownian motion. During the simulation process, the times of simulation trials are determined by the theoretical results. For example, if the theoretical BER is 10^{-5}, then 10^9 consecutive bits are used for simulation. For the diffusion process, each molecule executes a three-dimensional random walk [15] and the decision of whether the molecule is absorbed by the R_X is made by measuring the distance between the molecule and the center of the R_X every Δt s. These distances are compared with the radius of the R_X, R, when the distance is equal or smaller than R, this molecule is absorbed. At the R_X, the number of absorbed molecules are accumulated at the end of each time slot and then compared with the pre-designed threshold τ to determine whether decoded as '1' or '0'. For the work presented here, $R = 0.5\mu m$, $D = 10^3 \mu m^2/s$, $\Delta t = 10^{-6}$ s, $p_{tx} = 0.5$. d is varied between $2\mu m$ and $8\mu m$. In agreement with the work in [7], the ISI length I equals to 10. The value of t_s can be obtained by calculating the time when 60% molecules arrive at the R_X [1]. For different values of d, the t_s is different.

Fig. 2 shows the BER vs. the number of molecules per bit for the system that uses different models. As shown in Fig. 2, with the increasing of the number of molecules, the BER decreases for all channel models. It also shows that the previous model overestimates the error rate. For example, for 500 molecules per bit, the previous model predicts an error rate of 10^{-1}, almost 1000 times larger than the error rate predicted by the more accurate new model and verified by simulation. In addition, when $m < 100$, the Poisson is closer to the simulation results.

However, the Normal approximation gives a better compatibility than the Poisson approximation when $m > 100$. It indicated that the Normal approximation improves with the increasing of m. The increase in m leads to a higher mean number of absorbed molecules (i.e. a right shift of the Normal distribution curve) which can reduce the effects of the negative part of the distribution.

In Fig. 3, for $m = 500$, increasing the distance d leads to a higher BER. Furthermore, for the transmission distance between $2\mu m$ and

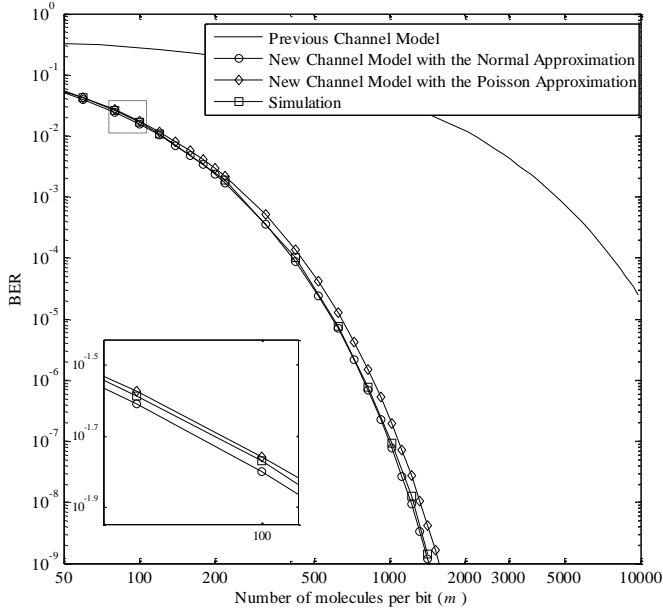Fig. 2 *BER vs the number of molecules per bit for different channel models and simulation with I = 10, d = 2μm.*
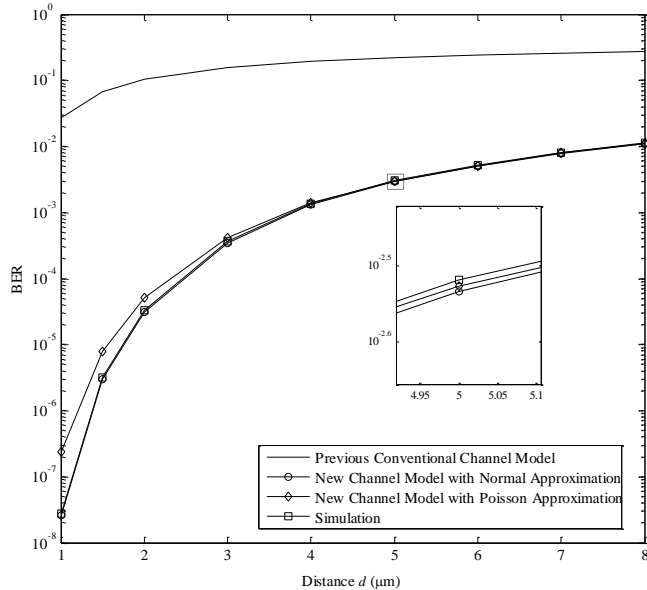

Fig. 3 *BER vs distance for different channel models and simulation with I = 10, m = 500.*

8μm, the BER estimated by the previous channel model stays at 10^{-1}, however, the actual BER changes from 10^{-5} to 10^{-2}, as predicted using our new model. This shows a big gap in the system performance that involved the previous model and the new model. Compared with the simulation results, the new channel model presents a more accurate estimation. It can also be seen that the Poisson model improves when the distance increases. This is because the Poisson distribution is normally used for modelling rare events [16], and the increase of the distance results in a decrease of the capture probability.

5. Conclusion: In this work, a channel model which includes the dependence of the numbers of received molecules between slots has been detailed and then compared with the incumbent model which assumes the independent numbers. This model is further evaluated

at arbitrary ISI length and with both the Normal and Poisson approximations. The update to the model shows the previous papers in the area overestimate the number of errors that can occur in the channel. This is a critical result as the papers in print, which consider the independence of events, will therefore most likely deliver a higher performance in practice. This observation is thus critical to those papers which deal in energy use at this nano-scale as with this assumption of dependence, less energy will be needed to attain a desired BER.

6. References

[1] M. Ş. Kuran, H. B. Yilmaz, T. Tugcu, and B. Özerman, "Energy model for communication via diffusion in nanonetworks," *Nano Commun. Netw.*, vol. 1, pp. 86-95, 2010.

[2] M. Ş. Kuran, H. B. Yilmaz, T. Tugcu, and I. F. Akyildiz, "Interference effects on modulation techniques in diffusion based nanonetworks," *Nano Commun. Netw.*, vol. 3, pp. 65-73, 2012.

[3] K. Na-Rae, A. W. Eckford, and C. Chan-byoung, "Symbol Interval Optimization for Molecular Communication With Drift," *IEEE Trans.Nanobiosci.*, vol. 13, pp. 223-229, 2014.

[4] A. Aijaz and A. H. Aghvami, "Error Performance of Diffusion-Based Molecular Communication Using Pulse-Based Modulation," *IEEE Trans. Nanobiosci.*, vol. 14, pp. 146-151, 2015.

[5] K. Na-Rae and C. Chan-Byoung, "Novel modulation techniques using isomers as messenger molecules for molecular communication via diffusion," in*IEEE Int. Conf. Commun.*(*ICC*), 2012, pp. 6146-6150.

[6] M. S. Leeson and M. D. Higgins, "Error correction coding for molecular communications," in *IEEE Int. Conf. Commun.*(*ICC*), 2012, pp. 6172-6176.

[7] Y. Lu, M. D. Higgins, and M. S. Leeson, "Comparison of Channel Coding Schemes for Molecular Communications Systems," *IEEE Trans.Commun.*, vol 63, pp. 3991-4001, 2015.

[8] M. S. Kuran, H. B. Yilmaz, and T. Tugcu, "A tunnel-based approach for signal shaping in molecular communication," in *IEEE Int. Conf. Commun.*(*ICC*), 2013, pp. 776-781.

[9] G. Genc, H. B. Yilmaz, and T. Tugcu, "Reception enhancement with protrusions in communication via diffusion," in *1stInt. Black Sea Conf. Commun. and Networking* (*BlackSeaCom*), 2013, pp. 89-93.

[10] H. B. Yilmaz and C. Chan-byoung, "Arrival modelling for molecular communication via diffusion," *Electron. Lett.*, vol. 50, pp. 1667-1669, 2014.

[11] A. Heren, H. Yilmaz, C. Chae, and T. Tugcu, "Effect of Degradation in Molecular Communication: Impairment or Enhancement?,"*IEEE Trans. Mol. Biol. Multi-Scale Commun.*, pp. 1-1, 2015.

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

[13] R. M. Ziff, S. N. Majumdar, and A. Comtet, "Capture of particles undergoing discrete random walks," *J. Chem. Phys.,* vol. 130, p. 204104, 2009.

[14] O. Ibe, *Fundamentals of Applied Probability and Random Processes:* Elsevier Science, 2014.

[15] H. C. Berg, *Random Walks in Biology*: Princeton University Press, 1993.

[16] R. Von Mises, Mathematical theory of probability and statistics. New York: Academic Press, 1964.

4