

Forward Error Correction for Molecular Communications

Mark. S. Leeson^{a,*}, Matthew D. Higgins^a

^a*School of Engineering, University of Warwick, Coventry, CV4 7AL, UK*

Abstract

Communication between nanoscale devices is an area of considerable importance as it is essential that future devices be able to form nanonetworks and realise their full potential. Molecular communication is a method based on diffusion, inspired by biological systems and useful over transmission distances in the nm to μm range. The propagation of messenger molecules via diffusion implies that there is thus a probability that they can either arrive outside of their required time slot or ultimately, not arrive at all. Therefore, in this paper, the use of a error correcting codes is considered as a method of enhancing the performance of future nanonetworks. Using a simple block code, it is shown that it is possible to deliver a coding gain of $\sim 1.7\text{dB}$ at transmission distances of $1\mu\text{m}$. Nevertheless, energy is required for the coding and decoding and as such this paper also considers the code in this context. It is shown that these simple error correction codes can deliver a benefit in terms of energy usage for transmission distances of upwards of $25\mu\text{m}$ for receivers of a $5\mu\text{m}$ radius.

Keywords: Molecular Communications, Nanonetworks, Error Correction Codes, Critical Distance

1. Introduction

The design and manipulation of objects at the nanometer scale is now established as a revolutionary technology [1] where the process of building entities at the nanoscale has been the theme of the vast majority of recent scientific work performed. A future is clearly emerging where the traditional silicon transmitters may be replaced by nanodevices built around biological and/or synthetic foundations. It is crucial for the full realisation of the potential impact of these emerging technologies that their interconnection capabilities are factored into their design. However, this aspect has only recently begun to receive the attention it warrants.

The first step towards nanonetworks is the determination of the appropriate means of inter-device communication [2] because although the paradigm of communications at the nanoscale encompasses the same elements as other communication systems, the physical mechanisms are completely different. A range of biologically inspired methods has been investigated in the literature and the use of molecules rather than electromagnetic or acoustic waves to transmit information is promising. Such molecular communication systems may be categorized by their range of operation. Over short distances, molecular motors [3] can transport information molecules along microtubules

[2]. For short to medium range operation, ion-channel signaling [4] and diffusion [2] may be employed. Long range systems can make use of bacterium based communication [5] or pheromone signaling [6].

In this paper, our focus is on the transmission of information over ranges up to a few hundred μm via the diffusion of identical messenger molecules within the medium containing the transmitter and receiver. The transmitter releases carrier molecules into the medium in which the nano-machines are suspended and these diffuse by means of Brownian motion which causes noisy reception in terms of the receiver arrival time. The introduction of message redundancy using error control coding (ECC) is a classic approach employed to increase transmission reliability and lower transmission power. This power reduction in terms of successful message delivery comes at the cost of extra power consumption in coding and decoding the messages. Stronger codes may be employed to substantially reduce the transmission power requirements but these bring complex coding and decoding needs likely to be beyond most nano-machines. Moreover, when the additional coding and decoding power outweighs the saving in transmission power then it is not efficient to employ ECC and the transmission should be uncoded. This paper is based, in part, from recent work presented at the IEEE MoNaCom 2012 [7] Workshop. Here, in addition to the previous results, a longer code length is investigated along with further clarifications as to the encoding-decoding energy equations.

The remainder of this paper is organised as follows. In Section 2, the fundamental nature of the diffusive channel is discussed. This is followed by Section 3 which outlines the model for communicating over the molecular channel.

^{*}This work was supported by a Royal Academy of Engineering Leverhulme Trust Senior Research Fellowship "Information Capacity at the Nanoscale".

*Corresponding author

Email addresses: mark.leeson@warwick.ac.uk (Mark. S. Leeson), m.higgins@warwick.ac.uk (Matthew D. Higgins)

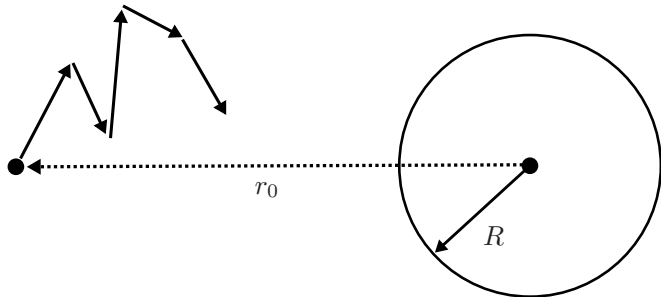


Figure 1: Generic model used for modelling 3D diffusion based communications.

Section 4 elucidates the relevant aspects of ECC. Section 5 presents an analysis of the energy needs of ECC in molecular communications before the concluding remarks are presented in Section 6.

2. The Diffusive Medium

Given that the messenger molecules are of dimensions at the nanometer scale, their movements within a fluidic transmission medium can be modelled by Brownian motion where the constant random thermal motion of the molecules within the fluid produce forces that combine to enact the motion of the messenger molecules. It can also be assumed that as these information carrying molecules are released in to a medium of large extent compared to their size, that collisions between them can be neglected for simplicity [8]. Under this system, information is conveyed by a sequence of symbols in consecutive time slots, with the transmitting entity releasing molecules to represent a binary one and not releasing any molecules to represent a binary zero. The random nature of the diffusion causes, what was an initial impulse of molecules, to be received as a long-tailed pulse which leads to the commonly recognised phenomenon of Intersymbol Interference (ISI) where molecules from a previous bit arrive during the current bit time slot. This ISI is known to have adverse effects in communications systems, particularly when the carrier system is stochastic [9], leading to incorrect symbol decisions.

The diffusion takes place in 3 dimensions (3D), making it more challenging to analyse than the examples that have sometimes been employed in the literature of molecular communications [10]. The molecular communication problem in general terms, shown in Fig. 1, concerns a molecule at a distance r_0 from a receiver of radius R . For most of this work, R was taken as $5\ \mu\text{m}$, unless otherwise stated, to produce results comparable with those in the literature [11]. For 3D diffusion in a medium with a diffusion coefficient of D , the escape probability $S(r_0, t)$, of the messenger molecule obeys the following backward difference equation at time t [12]:

$$\frac{\partial S(r_0, t)}{\partial t} = D\nabla^2 S(r_0, t) \quad (1)$$

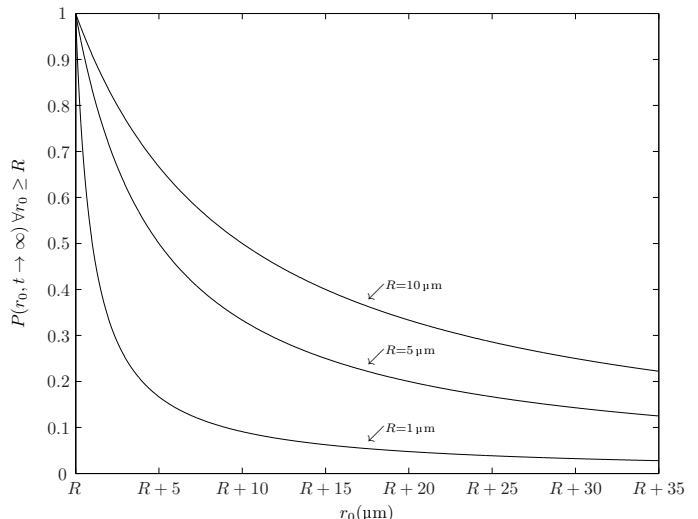


Figure 2: Hit probability as $t \rightarrow \infty$ for receiver diameters of $R = 1\ \mu\text{m}$, $5\ \mu\text{m}$ and $10\ \mu\text{m}$.

In order to solve (1), account must be taken of both the diffusive and stochastic natures of the process. The former is straightforward given the spherical symmetry of the model, and is achieved by solving (1) using the initial conditions of $S(|r_0|, 0) = 1 \forall |r_0| > R$ and the boundary conditions $S(|r_0| = R, t) = 1 \forall t$ and $S(|r_0| \rightarrow \infty, t) = 1 \forall t$. The latter is included by using the theory of Harmonic Functions [13] which tells us the probability that the time at which a molecule arrives at the receiver is less than infinity is given by R/r_0 . Thus, there is a finite probability that a given molecule will never arrive at the receiver, no matter how long a time has passed. In this work, as is true of molecular communication theory in general, the capture probability $P(r_0, t)$, is of more interest and is given by $(1 - S(r_0, t))$. Using the solution of (1) produces:

$$P(r_0, t) = \frac{R}{r_0} \operatorname{erfc} \left\{ \frac{(r_0 - R)}{2\sqrt{Dt}} \right\} \quad \forall r_0 \geq R, \forall t \quad (2)$$

When (2) is taken to the limit of $t \rightarrow \infty$, $P(r_0, t \rightarrow \infty) = R/r_0$. This expression provides an analytical approach to finding the probability that a molecule will arrive at the receiver and removes the need to Monte Carlo simulation typically used in point to point links. Fig. 2 illustrates the behaviour of the long term hit probability where it may be observed that this rapidly becomes very sensitive to distance as the receiver diameter decreases as one would expect since the target is smaller. Naturally, by differentiating the capture probability with respect to time, one would obtain the hit time distribution which would show the long diffusive tail which leads to ISI [14].

3. The Communication Channel

The communication channel here is, in essence, Binomial in nature in that each messenger molecule arrives in

their respective timeslot or not. ISI from all of the previous bits has an influence on the current bit as there will be a number of residual molecules N_p , from previous one bits during the current bit period at time t_s . This means that the difference between Binomial distributions is needed to accommodate the probability that the received messenger molecule was sent some time slots ago but arrives now. In general, one should calculate

$$N_p \sim \sum_{i=1}^{\infty} B(N, P(r_0, \{i+1\}t_s)) - B(N, P(r_0, it_s)) \quad (3)$$

However, for parameters typical of molecular communications, it has been shown in [11] that only one previous timeslot needs to be considered. Upon first observation, the symbol pairs form a Markov chain, however the relevant probabilities for previous and current bit pairs $\{00, 01, 10, 11\}$ may be calculated by multiplying together the relevant single bit probabilities as each bit is independent. Thus, there is zero error associated with the pair 00 and the pair 01 follows the Binomial distribution $B(N, P(r_0, t))$ for N molecules in a bit time t_s . This result however, is unfortunately not a trivial problem as no closed form expression currently exists for the difference between two Binomial distributions. As such a Gaussian approximation is employed hereafter.

The probability equations of [11] were employed to determine the BER for the system considered here with a diffusion coefficient of $D = 79.4 \mu\text{m}^2 \text{s}^{-1}$, a conservative value for insulin in water at the human body temperature [15]. As can be seen in Fig. 3, the classic BER versus signal (number of molecules) curve shape are apparent with a significant dependence on transmission distance. Although promising indications that it is possible to achieve useful BER values, in the region of $< 10^{-6}$ using a few hundred molecule per bit, it is interesting to see if better performance is possible. Hence, it will be necessary to consider the use of error correction coding.

4. Error Correction Coding

Since Shannon's derivation of capacity does not provide a method of constructing the necessary coding schemes to achieve it, there is a long history of error correcting codes to approach the capacity [16]. In the case of nanoscale communications, the energy budget is severely constrained since the nodes will usually scavenge their energy from the surrounding environment. The achieved output of energy harvesting devices is only likely to be in the pW range [17] meaning that state of the art recursive coding schemes [18] will be too energy intensive. Therefore, relatively simple block codes are considered that add parity check bits to enable the correction of errors. One of the simplest families is that of the Hamming codes [18], defined for an integer $m \geq 2$ such that blocks of $k = 2^m - m - 1$ data bits are used to form coded output blocks of length $n = 2^m - 1$

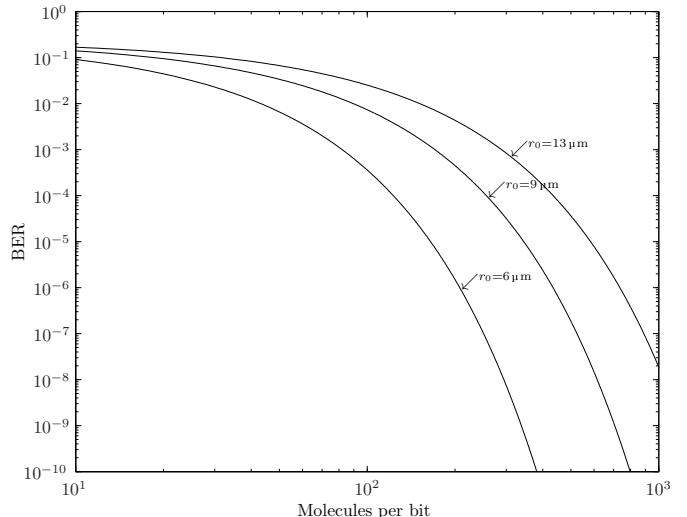


Figure 3: Bit error rate versus molecules per bit for distances of $r_0 = 6 \mu\text{m}$, $9 \mu\text{m}$ and $13 \mu\text{m}$.

having m parity check bits. The minimum distance of such (n, k) block codes is 3 and the number of errors they are able to correct, t , is limited to one error per block. They are thus not powerful by state of the art standards, but are simple, and hence potentially within the energy budget nanoscale communication devices.

To compare the performance of the coded transmission with the coded results in Fig. 3, the BER for Hamming coded operation must be determined. In general, this is not easy for coded systems but for block codes and approximate expression is [19]:

$$P_{ec} \approx \frac{1}{n} \sum_{j=t+1}^n j \binom{n}{j} P_{uc}^j (1 - P_{uc})^{n-j} \quad (4)$$

Where P_{uc} is the probability of error for the uncoded case. In the case of molecular communications, there are different probabilities of error depending upon the data pattern because of the ISI. However, an upper bound can be determined by letting P_{uc} be the maximum value of the different probabilities for the bit pairs used in the calculation of the uncoded case. Further to this, to provide a fair measure of the coded and uncoded transmission, P_{uc} in (4) should also be evaluated with the number of molecules reduced by multiplication with coding rate given by the ratio of k/n .

Carrying out this calculation produces the results of Fig. 4, where it may be seen that there is a benefit in coding once the BER gets beyond the values of 2×10^{-4} , 2×10^{-3} and 3×10^{-3} for $m = 3, 4$ and 5 respectively. It is also typical to provide the coding gain in dB of a code at a relevant BER. As the transmission energy is approximately linearly related to the number of molecules per bit (to be shown later) the coding gain is simply the ratio of molecules needed for uncoded and coded systems respectively. Assuming the system is desired to operate

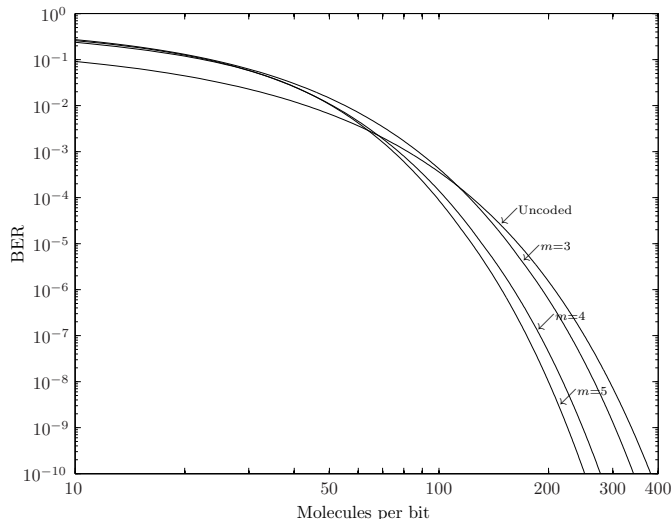


Figure 4: BER for coded and uncoded transmission using Hamming codes having $m = 3, 4, 5$ with $r_0 = 6 \mu\text{m}$.

effectively error free, i.e at a $\text{BER} = 10^{-9}$ with likely bit rates at $\sim 100 \text{ bit s}^{-1}$ [11], the coding gains are $\sim 0.5 \text{ dB}$, $\sim 1.4 \text{ dB}$ and $\sim 1.7 \text{ dB}$ for $m = 3, 4$ and 5 respectively.

5. Energy Considerations

As shown in Figure 4, it is therefore apparent that Hamming codes can improve the performance of nanoscale molecular communications. However, this improvement comes at an extra cost to the energy requirements associated with the encoding and decoding of the data at the transmitter and receiver. To measure this trade-off we define the total cost, or gain, of using the Hamming code as the energy required to transmit an uncoded message, less the cost of transmitting the same message using the Hamming code at an equivalent BER.

Adenosine triphosphate (ATP) is the natural currency of energy transfer between cells in living organisms [20]. Through the process of breaking off phosphate groups, $\sim 8.3 \times 10^{-20} \text{ J}$ of energy is released. For such small quantities, it is useful to work in $k_B T$ s of energy, where k_B is Boltzmann's constant such that the energy from one ATP reaction is $\sim 20k_B T$ at 300 K . It must be stated at this point that these estimates are speculative and the energy consumption of coding at the nanoscale has yet to be fully investigated. Nevertheless, some indication of the energy required may be obtained by a simplified consideration of the potential operations needed for using error correcting codes. The first stage in producing this estimate is to get an insight into the ATP cost of implementing the logical operations needed to produce the code. These operations may be reduced to an arrangement of XOR, NAND and NOT gates by using the designs laid out in Blahut [21].

Sauro and Kholodenko [22] have considered some example protein based signaling networks within biological cells.

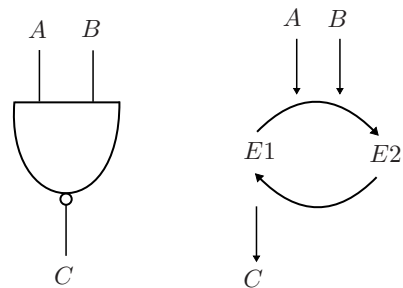


Figure 5: NAND gate formed using the fundamental signaling motif of [22].

Given that these networks are of, or at, a similar scale to those needed by any future artificial nanonetworks, they thus form a good estimate of the likely energy needs within any encoding-decoding elements of a nano-machine. The fundamental motif of their signaling network is the cycle formed by a kinase and an opposing phosphatase. This is illustrated in Figure 5 for a NAND gate where $E1$ and $E2$ represent unphosphorated and phosphorated enzymes respectively and A and B are the inputs to the kinase step required to cause the cascade cycle to switch at a cost of 1ATP.

Using the principles of Boolean algebra, from here it is possible to build any of the required NOT gates, shift register units and XOR gates using one, four and five NAND gates respectively. Furthermore, taking the approach of [11], the cost of synthesizing a molecule is $\sim 2450k_B T$ and given that the number of molecules per bit need only one transport container (a vesicle or its artificial equivalent), there will be no extra cost in transmission between the uncoded and coded cases other than that of the additional logical operations required for the encoding and decoding which will now be dealt with in turn.

5.1. Encoding

One of the most efficient ways to produce the n bits of a Hamming code is given in Blahut [21]. Here, the production of the code entails splitting the data into n -bit blocks padded with $n - k$ zeros which are then fed through a n -time clocked Finite Impulse Response (FIR) filter composed of m shift registers and two XOR gates. An example circuit used in this work is shown in Fig 6 for the $m = 4$ (15, 11) code. It can be observed that for this code, the generator polynomial is given by $g(x) = x^4 + x + 1$. For the $m = 3$ (7, 4) and the $m = 5$ (31, 26) codes used in this paper, $g(x) = x^3 + x + 1$ and $g(x) = x^5 + x^2 + 1$ respectively such that the total energy, in $k_B T$'s, per bit for a given code is equal to:

$$E_{\text{encode}} = 20N_{tx}(4m + 10) + 2450N_{tx} \quad (5)$$

Where N_{tx} is the number of code generation molecules required to encode the data. This number of internal molecules required to generate the code prior to creating

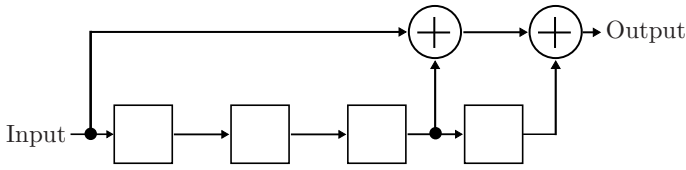


Figure 6: Non-systematic (15,11) encoder.

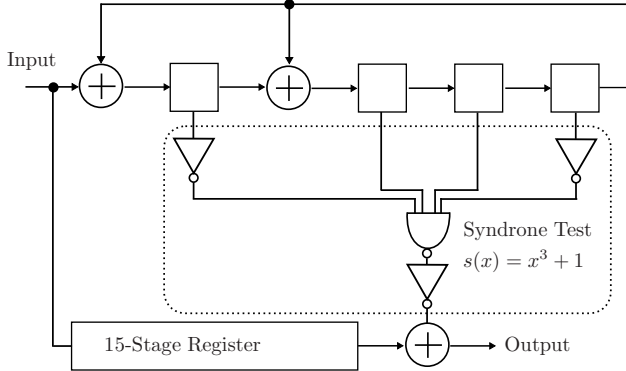


Figure 7: Example Meggit decoder for a (15,11) Hamming code.

the transmission molecules is independent of the number of messenger molecules required. They are only used internally within the nanomachine and do not suffer from any of the effects of diffusion. They do however still suffer from biochemical fluctuations that contribute to the general additive noise which means that to reduce this to a negligible amount, N_{tx} is recommended [23].

5.2. Decoding

The Meggit decoder requires for a given code, an m bit register with two XORs (as with the encoder), a n stage shift register and syndrome tester subcircuit consisting of 3 NOT gates, a further XOR gate and an m input NAND gate which we also assume to still consumes 1ATP to switch as all that should change is the number of inputs to the kinase step in Fig. 4. An example of this circuit is shown in Figure 7 for the $m = 4$ (15, 11) code where it can be observed the syndrome test circuit is checking for the polynomial $s(x) = x^3 + 1$. For the $m = 3$ (7, 4) and $m = 5$ (31, 26) codes, the required syndrome test configurations are $s(x) = x^2 + 1$ and $s(x) = x + x^4$ respectively such that the total energy in $k_B T$'s required to decode is:

$$E_{\text{decode}} = 20N_{rx}(4m + 4n + 19) + 2450N_{rx} \quad (6)$$

Hence, the energy saving ΔE , by employing the Hamming code may be written as

$$\Delta E = 2450(N_u + N_c) - E_{\text{encode}} - E_{\text{decode}} \quad (7)$$

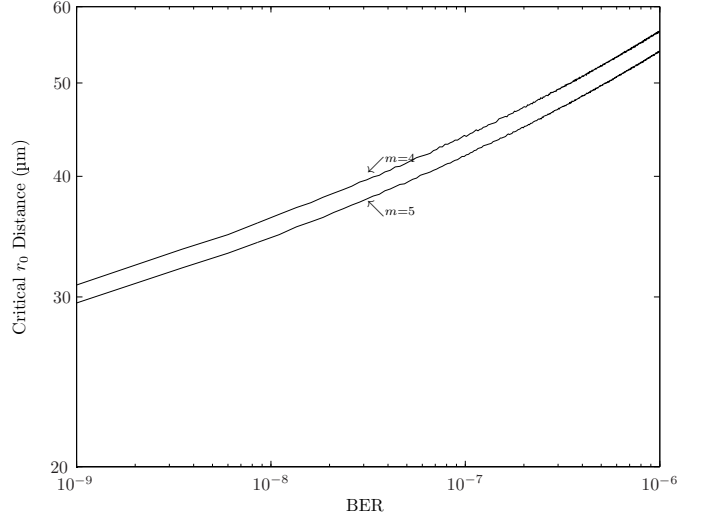


Figure 8: Critical r_0 distance versus BER for $m = 4$ and 5 coded systems.

Where N_u and N_c are the number of molecules needed to achieve a given BER for the uncoded and coded system at equivalent BERs. Therefore, it is clearly worth using a code when ΔE is positive, as this is the point where less energy is needed to achieve a given BER including the overheads than using an uncoded system. Assuming that $N_{tx} = N_{rx} = 300$, through the application of (5) through (7) it can be shown that for $m = 3$, in order to achieve energy parity, N_c must be ≈ 800 molecules per bit less than N_u at the desired BER to be worthwhile. As has already been shown in Section 4, in Fig. 4, the coding gain for short propagation distances is far below this requirement. However, as the transmission distances are increased, so does the respective coding gains, and of importance to us, is the distance at which that coding gain allows ΔE to be positive. This distance is termed the critical distance [24].

Fig. 8 shows the values of the critical distance over a range of BER values from 10^{-9} to 10^{-6} for Hamming codes with $m = 3$ and 4. It may be seen that for longer range transmission, $r_0 \gg 30 \mu\text{m}$ the use of coding is beneficial in terms of energy usage. This therefore indicates that the Hamming code, although useful for extending the range of error free transmission and delivering a coding gain, it is not universally a powerful enough tool at the nanoscale due to its extra energy requirements at shorter distances which many nanoscale communication networks may operate over.

6. Conclusion

The use of nano-machines is known to be a high priority area in the research landscape of today and the near future. These devices will also require wireless communication capabilities to form their so called nanonetworks. In this paper, the performance of a simple block code as

a means to enhance the communication performance between nano-machines has been investigated. By employing Hamming codes, coding gains of ~ 1.7 dB over 1 μm transmission distances can be achieved. In addition, this paper has also considered the energy budget of their implementation to determine the critical distance - the distance at which the energy gain of the code overcomes its operational energy cost. It is shown that this critical distance is $\gg 30 \mu\text{m}$ which is therefore beneficial for nanonetworks operating over such distances but cannot be considered universally beneficial due to the higher energy demands upon shorter ranging networks. It is however, apparent that new codes must be investigated, and they must be considered within the context of the limited energy available within nanonetworks.

References

- [1] Jones AL. *Soft Machines: nanotechnology and life*. Oxford University Press; 2004.
- [2] Akyildiz IF, Brunetti F, Blzquez C. Nanonetworks: A new communication paradigm. *Computer Networks* 2008;52(12):2260 – 79. doi:10.1016/j.comnet.2008.04.001.
- [3] Schliwa M, Woehlke G. Molecular motors. *Nature* 2003;422(6933):759–65.
- [4] Nakano T, Suda T, Moore M, Egashira R, Enomoto A, Arima K. Molecular communication for nanomachines using intercellular calcium signaling. In: *Nanotechnology, 2005. 5th IEEE Conference on*. 2005, p. 478 – 481 vol. 2. doi: 10.1109/NANO.2005.1500804.
- [5] Cobo LC, Akyildiz IF. Bacteria-based communication in nanonetworks. *Nano Communication Networks* 2010;1(4):244 –56. doi:10.1016/j.nancom.2010.12.002.
- [6] Giné LP, Akyildiz IF. Molecular communication options for long range nanonetworks. *Computer Networks* 2009;53(16):2753 –66. doi:10.1016/j.comnet.2009.08.001.
- [7] Leeson MS, Higgins MD. Error correction coding for molecular communications. In: *IEEE ICC 2012 MoNaCom Workshop, Ottawa, Canada, 2012.*
- [8] Moore MJ, Suda T, Oiwa K. Molecular communication: Modeling noise effects on information rate. *NanoBio-science, IEEE Transactions on* 2009;8(2):169 –80. doi: 10.1109/TNB.2009.2025039.
- [9] Leeson MS. Performance analysis of direct detection spectrally sliced receivers using fabry-perot filters. *Lightwave Technology, Journal of* 2000;18(1):13 –25. doi:10.1109/50.818902.
- [10] Kadloor S, Adve RS, Eckford AW. Molecular communication using brownian motion with drift. *CoRR* 2010;abs/1006.3959.
- [11] Kuran MS, Yilmaz HB, Tugcu T, Özerman B. Energy model for communication via diffusion in nanonetworks. *Nano Communication Networks* 2010;1(2):86 – 95. doi: 10.1016/j.nancom.2010.07.002.
- [12] Ziff RM, Majumdar SN, Comtet A. Capture of particles undergoing discrete random walks. *The Journal of Chemical Physics* 2009;130(20):204104 (pages 5). doi:10.1063/1.3137062.
- [13] Mörters P, Peres Y. *Brownian Motion*. Cambridge University Press; 2010.
- [14] Llatser I, Alarcon E, Pierobon M. Diffusion-based channel characterization in molecular nanonetworks. In: *Computer Communications Workshops (INFOCOM WKSHPS), 2011 IEEE Conference on*. 2011, p. 467 –72. doi: 10.1109/INFCOMW.2011.5928858.
- [15] Kim JJ, Park K. Modulated insulin delivery from glucose-sensitive hydrogel dosage forms. *Journal of Controlled Release* 2001;77(12):39 – 47. doi:10.1016/S0168-3659(01)00447-3.
- [16] Costello D.J. J, Hagenauer J, Imai H, Wicker S. Applications

- of error-control coding. *Information Theory, IEEE Transactions on* 1998;44(6):2531 –60. doi:10.1109/18.720548.
- [17] Wang X, Song J, Liu J, Wang ZL. Direct-current nanogenerator driven by ultrasonic waves. *Science* 2007;316(5821):102–5. doi: 10.1126/science.1139366.
- [18] Moon TK. *Error Correction Coding: Mathematical Methods and Algorithms*. Wiley; 2005.
- [19] Sklar B. *Digital Communications: Fundamentals and Applications*. Prentice-Hall; 2001.
- [20] Cox M, Nelson DL. *Lehninger Principles of Biochemistry*. Palgrave Macmillan; 2008.
- [21] Blahut RE. *Algebraic Codes for Data Transmission*. Cam; 2003.
- [22] Sauro HM, Kholodenko BN. Quantitative analysis of signaling networks. *Progress in Biophysics and Molecular Biology* 2004;86(1):5 – 43. doi:10.1016/j.pbiomolbio.2004.03.002.
- [23] Levine J, Kueh HY, Mirny L. Intrinsic fluctuations, robustness, and tunability in signaling cycles. *Biophys J* 2007;92(12):4473–81.
- [24] Howard SL, Schlegel C, Iniewski K. Error control coding in low-power wireless sensor networks: when is ecc energy-efficient? *EURASIP J Wirel Commun Netw* 2006;2006(2):29–. doi:10.1155/WCN/2006/74812.

Mark S. Leeson received the degrees of BSc and BEng with First Class Honors in Electrical and Electronic Engineering from the University of Nottingham, UK, in 1986. He then obtained a PhD in Engineering from the University of Cambridge, UK, in 1990. From 1990 to 1992 he worked as a Network Analyst for National Westminster Bank in London.



After holding academic posts in London and Manchester, in 2000 he joined the School of Engineering at Warwick, where he is now an Associate Professor (Reader). His major research interests are coding and modulation, nanoscale communications and evolutionary optimization. To date, Dr. Leeson has over 220 publications and has supervised ten successful research students. He is a Senior Member of the IEEE, a Chartered Member of the UK Institute of Physics and a Fellow of the UK Higher Education Academy.

Matthew D. Higgins received his MEng in Electronic and Communications Engineering and PhD in Engineering from the School of Engineering at the University of Warwick in 2005 and 2009 respectively. Remaining at the University of Warwick, he then progressed through several Research Fellow positions with leading defence and telecommunications companies before undertaking two years as a Senior Teaching Fellow. As of July 2012 Dr. Higgins now holds the position of Assistant Professor. His major research interests are the modelling of optical propagation characteristics in underwater, indoor and atmospheric conditions. Dr. Higgins is a Member of both the IEEE and IET.

