Low-Weight Code Comparison for Electromagnetic Wireless Nanocommunication

Muhammad Agus Zainuddin, Eugen Dedu, and Julien Bourgeois

Abstract—**Electromagnetic communication among nanosensors using Time-Spread On-Off Keying (TS-OOK) modulation in Terahertz band promises very high transmission rates (up to several Terabits per second). Due to scarce battery capacity in nanosensors, energy efficiency is a very important aspect in nanocommunication, as are also bandwidth expansion, multi-user interference and robustness against transmission errors. This paper compares various low weight codes found in the literature using metrics specific to nanocommunication. A small variation of such a code is also introduced and included in the comparison. Results show that there are trade-offs among the various metrics used and, even if there is no clear winner method, the novel method has good results in almost all metrics.**

Index Terms—**codes, energy efficiency, nanocommunication, nanonetworks.**

I. Introduction

A nanosensor is a very small sensor (with total size between 1 and a few hundred nanometers), which allows the observation at molecular scale. Due to its small dimension, it has limits in terms of battery capacity and transmission range [1]. Interconnecting nanosensors, i.e., wireless nanosensor networks, will be able to overcome these limitations. Moreover, connecting wireless nanosensor networks to Internet networks allows worldwide access to the observation at molecular scale [2]. This property will enable many advanced applications in various fields, e.g., biomedical, military, and multimedia. For example, in biomedical field nanosensors with nano-camera [3], [4] and nano-heater [5] will improve anti-microbiology, where nanosensors have the ability to detect transmissible agents (such as fungi, bacteria, viruses, spore forms) and various cancer cells, then kill them by heat [6]. Nanosensors can also be used for targeted drug delivery [7], where antibodies or single stranded DNA chains can be tied up to specific target cells. Note that, even if nanodevices are small, data size exchanged among them can be very big, as is the case for nano-camera, and medical or wireless network-on-chip applications for example.

Due to limited battery capacity and circuit complexity, pulse-based modulation is currently the only modulation scheme for the electromagnetic communication among nanosensors. Jornet et al. [8] proposed TS-OOK modulation in Terahertz band, where a femtosecond-long pulse is transmitted

for bit 1, and silence for bit 0. Time between consecutive bits is fixed. The greatest energy consumption in the transmission process is the pulse transmission, since it activates nano-components, such as signal generator, power amplifier, band pass filter for pulse shaping, etc. It is clear that reducing the number of bits 1 yield energy efficiency.

In order to obtain energy efficiency, some codes require large bandwidth expansion, thus increasing transmission time. Also, some codes are more robust than others, since for some of them a 1-bit error leads to several erroneous bits at receiver, i.e., a high error rate.

We have not found any papers in the literature analysing low weight codes. Data transmission in nanonetworks is in its infancy, and this study aims to foster the development of nanonetworks, especially internet of multimedia nano-things [2]. In this paper, we compare the codes that can be used to reduce the number of bits 1 using several criteria appropriate to nanocommunications, such as energy efficiency, bandwidth expansion, and robustness against transmission errors, including a criterion relevant to multimedia, PSNR (Peak Signal to Noise Ratio). Moreover, we also propose NPG code and include it in code comparison. The comparison is done theoretically using a mathematically-generated input data, and numerically using an image (cancer) file.

The paper is organised as follows. Sec. II presents the codes compared: ME (Minimum Energy [9]), NME (Nanonetwork Minimum Energy [10]), PG (Prakash and Gupta [11]), NPG (New PG, a small variation of PG presented in this article), MTE (Minimum Transmission Energy [12]), MEC (Minimum Energy Channel [13]), LWC (Low Weight Channel [14]), and classical unary. Sec. III presents evaluation metrics. Sec. IV and V present code performance using theoretical evaluation, and numerical results using the above criteria applied to an image file. Finally, the paper is summarised in Sec. VI.

II. Methods compared

As shown in the introduction, reducing the number of bit 1 in transmitted bit sequence will improve transmission. This is the aim of low weight codes. This section presents some of them.

To illustrate the methods below we will use a hypothetical example of 3-bit input symbols and their frequency, shown in the first two columns of Table I.

All the codes use a dictionary for input symbols to code-words mapping. In our analysis we decided not to take into account the dictionary because for large data it has a negligible size compared to data size; moreover, in some applications the receiver knows the dictionary before the transmission itself.

Input symbol	Symbol freq.	ME	NME	PG	NPG	MTE	MEC	LWC	Unary
111	80	000	000	1000000	0000000	1100	0...0011	10010	1
110	70	001	001	0100000	0000001	0110	0...1100	10001	01
101	60	010	010	0010000	0000010	0011	.	01100	001
100	50	100	100	0001000	0000100	1000	.	01010	0001
011	40	011	101	0000100	0001000	0100	.	01001	00001
010	30	101	011	0000010	0010000	0010	.	00110	000001
001	20	110	110	0000001	0100000	0001	0011...0	00101	0000001
000	10	111	111	0000000	1000000	0000	1100...0	00011	00000001

TABLE I
THE HYPOTHETICAL EXAMPLE USED (FIRST TWO COLUMNS) AND THE MAPPING TABLE FOR ALL COMPARED CODES.

A. Minimum energy (ME) code

Erin et al. [9] proposed minimum energy (ME) coding for known source statistics. The input symbols are sorted in decreasing order of their frequency. The codewords have the same size as input symbols and are sorted in increasing number of their weight (the weight is the number of 1s). As a result, the most frequent symbols are mapped to codewords with fewer 1s. The mapping table is shown in Table I.

B. Nanonetwork minimum energy (NME) code

NME code [10] is similar to ME, with the difference in the sorting of the codewords with the same weight; the sorting is done so as to reduce the number of sequential bits 1, a useful feature for nanonetworks. The algorithm to reduce the number of sequential 1 in NME code is shortly explained below:

1) The codewords are sorted in increasing order of their weight.
2) For each codeword with the same weight, the codeword is divided into two categories according to the existence of sequential 1 in codeword.
3) Each category is then divided into two new categories according to existence of 1 at the beginning or at the end of codeword.

C. Prakash and Gupta (PG) code

Prakash and Gupta [11] extended the ME code for source with unknown statistics, by mapping the input symbols to predetermined codewords. In this method, all codewords have a maximum weight of 1. The data is cut in binary sequences of n bits, and each symbol of n bits is mapped to a codeword of m bits. In particular, input symbol 0 is always mapped to codeword 0. This method requires large codeword sizes, the minimum codeword size being $m = 2^n - 1$. It has however a large energy efficiency since every codeword has a maximum weight of 1. For example, when $n = 3$, $m = 7$, and the mapping table is shown in Table I.

D. New PG code

In this paper we also propose a small variation of PG code (NPG, New PG). The only difference compared to PG is that codeword 0 is used for the most frequent input symbol, and the other symbols are mapped to codewords with weight 1. This means that, contrary to PG, NPG uses source statistics.

E. Minimum transmission energy (MTE) code

Chi et al. [12] proposed Minimum Transmission Energy (MTE) code, which modifies PG in order to have larger *input* size for the same codeword size; it does so by allowing the use of codewords with two consecutive 1s too (hence weight 2). Consequently, MTE code has a lower bandwidth expansion than PG code. MTE code is denoted by MTE (n, m), where n is input symbol size and m is output symbol size. The weight of i-th codeword w_i is denoted by:

$$w = \begin{cases} 0, & \text{if } i = 1 \\ 1, & \text{if } 2 \leq i \leq m + 1 \\ 2, & \text{if } m + 2 \leq i \leq 2^n \end{cases} \quad (1)$$

where i is the codeword index (starting with 1) and m sufficiently large. For example, when $n = 3$ and $m = 4$, the mapping table is as shown in Table I.

F. Minimum Energy Channel code (MEC)

Kocaoglu et al. [13] proposed minimum energy channel (MEC) coding for nanocommunications. The code provides error correction capability and minimizes the energy consumption by minimizing the average codeword weight. The code uses multiple carriers and multiple antennas to obtain low absorption and low molecular noise at Terahertz band. Minimum distance d_{min} is used to control robustness; larger d_{min} provides better error correction capability. Assuming that d_{min} is even, minimum codeword size m_{min} is obtained by:

$$m_{min} = M d_{min} / 2 \quad (2)$$

where $M = 2^n$ is the number of used symbols, and n is input symbol size. Codeword weight depends on the maximum probability of transmitted symbols P_{max}. If P_{max} is less than 0.5, the codeword weight is constant $d_{min}/2$, elsewhere the codeword weight is 0 for symbol 0 and d_{min} for others. If $d_{min} = 2$, then there is no error correction and it behaves exactly like PG code. In this article we choose $d_{min} = 4$ in order to provide the code with error correction capability, have a good energy efficiency and maintain bandwidth expansion low. When $d_{min} = 4$, $M = 8$ and $m_{min} = 16$, the mapping table is as shown in Table I.

G. Low weight channel (LWC) code

Jornet et al. [14] proposed Low Weight Channel (LWC) code for transmission error prevention in nanosensor networks. It uses constant codeword *weight*. LWC code is denoted by

LWC (m, n, w), where m is output size, n is input size, and w is codeword weight. The codeword weight depends on input size, and fulfils the following condition:

$$\binom{m}{w} \geq 2^n \qquad (3)$$

where $\binom{m}{w}$ denotes the number of combinations of w bits 1 in m bit codewords.

For example, when $m = 5$, $n = 3$, and $w = 2$, the mapping table for LWC code is as shown in Table I. The probability of bit 1 can be obtained by $P(1) = w/m$. The smaller $P(1)$, the smaller the multi-user interference.

In this article we choose codeword weight $w = 2$, because if $w = 1$ it behaves like PG, and if $w > 2$ the efficiency decreases.

H. Unary code

In alternative unary code, all codewords have weight 1. Most frequent symbols are mapped to codewords of smaller size, starting with 1 bit. Mapping table for unary code is shown in Table I. A useful property is that all codewords end in bit 1, so the receiver knows that the following bits belong to next symbol.

III. Metrics used in code comparison

This section presents the criteria used for evaluation. We will use the same hypothetical example of the previous section to illustrate them.

A. Energy efficiency

Energy efficiency measures the code effectiveness to reduce the number of bits 1 from input. It can be computed using:

$$\xi = \frac{n_o - n_c}{n_o} \times 100\% \qquad (4)$$

where n_o refers to the number of bits 1 in input data, and n_c to the number of bits 1 in data transmitted. Codes with larger average value of energy efficiency are better. Sometimes n_c could be greater than n_o, which results in negative energy efficiency; this means that the code gives worse results than without coding.

We compute the energy efficiency using (4) for the previous hypothetical example (Table I, where it can be computed that uncoded data contains 680 bits 1). The result is shown in Table II.

Reducing energy for transmission in nanonetworks is useful. For example, novel nanoscale energy harvesting systems [15], [16] enable nanomachine to convert vibrational, fluidic, electromagnetic or acoustic energy into electrical energy. Nanomachine can even achieve perpetual operation when harvested energy is larger than the consumed one [17]. A reduced energy consumption would allow a smaller inter-symbol time (time between consecutive bits), hence a greater transmission speed. In the case where nano-batteries are used, a reduced energy would naturally allow a longer battery lifespan.

Finally, note that coding is a very low energy consuming process. Codes in this paper simply map the input to

Code	Number of bits 1	Energy efficiency (%)	Number of bits	Bandwidth expansion
Uncoded	680	0.0	1080	1.0
ME, NME	390	42.7	1080	1.0
PG	350	48.5	2520	2.3
NPG	280	58.8	2520	2.3
MTE	560	17.7	1440	1.3
MEC	720	–5.9	5760	5.3
LWC	720	–5.9	1800	1.7
Unary	360	47.1	1200	1.1

TABLE II
ENERGY EFFICIENCY AND BANDWIDTH EXPANSION OF THE VARIOUS CODES IN THE HYPOTHETICAL EXAMPLE.

corresponding codeword, which requires sufficiently small energy to be implemented in nanoprocessors based on nano-transistor [18] and nano-memory [19] in nanomachine.

B. Bandwidth expansion

Some low weight codes reduce the number of bits 1 by *increasing* the number of transmitted bits. The bandwidth expansion is expressed as the ratio between total number of bits in output data and total number of bits in input data, i.e., bandwidth expansion b_E means transmitting $b_E \times A$ bits for A input bits.

Bandwidth expansion has some effects: reduce the information rate and increase the power consumption at transmitter and receiver. The information rate is reduced proportionally to bandwidth expansion, e.g., if the transmission rate is r Mbps, then information rate becomes r/b_E Mbps. Also, a larger output size makes a longer transmission, hence both transmitter and receiver stay active for longer time. Thus, code with lower bandwidth expansion is better.

The bandwidth expansion using the hypothetical example is given in Table II.

C. Information rate

Information rate is useful because it takes into account the Terahertz band properties. Information rate after coding in bit per second using TS-OOK modulation can be computed as follows [8]:

$$
\begin{aligned}
IR = \frac{W}{\beta b_E} \max &\left\{ \left[\sum_i P_i \times log_2\left(\frac{1}{P_i}\right) \right] \right. \\
&- \left[\int \sum_{i=0}^{1} \frac{1}{\sqrt{2\pi N_i}} e^{-\frac{(y-a_i)^2}{2N_i}} P_i \right. \\
&\left. \times log_2\left(\sum_{j=0}^{1} \frac{P_j}{P_i} \sqrt{\frac{N_i}{N_j}} e^{\frac{(y-a_i)^2}{2N_j} - \frac{(y-a_j)^2}{2N_i}} \right) dy \right] \left. \right\}
\end{aligned}
\qquad (5)
$$

where W is the bandwidth, β the spread factor in TS-OOK modulation, b_E the bandwidth expansion, P_i the probability of bit $i = \{0, 1\}$ to be transmitted, N_i the total noise power for transmitted bit i, y the output channel and a_i the amplitude of received signal.

Total noise power N_i for transmitted bit i can be computed as follows [14]:

$$N_0(d) = \int_W k_B T_0 (1 - e^{-k(f)d}) \left(\frac{c_0}{\sqrt{4\pi} f_0}\right)^2 df \qquad (6)$$

$$N_1(d) = N_0(d) + \int_W S_X(f)(1 - e^{-k(f)d})\left(\frac{c_0}{4\pi df_0}\right)^2 df \quad (7)$$

where k_B is the Boltzmann constant, T_0 the room temperature, c_0 the speed of light in the vacuum, f_0 the design center frequency, S_X the power spectral density of the transmitted signal and k the medium absorption coefficient. For computation we use HITRAN (HIgh resolution TRANsmission molecular absorption database) an online catalog [20] for absorption coefficient.

The amplitude of received signal a_i can be computed as follows:

$$a_i = \sqrt{\int_W |Y(f)|^2 df} \quad (8)$$

$$Y(f) = \frac{X(f)}{A(f)} \quad (9)$$

$$A(f) = \left(\frac{4\pi df}{c}\right)^2 e^{k(f)d} \quad (10)$$

where W is the channel bandwidth, $Y(f)$ the Fourier transform of received signal, $X(f)$ the Fourier transform of TS-OOK pulse (Gaussian pulse), $A(f)$ the total path-loss, d the transmission distance, f the frequency and c the speed of light in the vacuum.

D. Multi-user interference

Networking nanodevices allows to increase their transmission range. A high density of nanosensors is necessary to increase the transmission rate [21]. Uncoordinated pulse transmission in TS-OOK modulation can lead to interference (collision) between transmitted symbols [8]. Interference appears when a receiver receives pulses from several transmitters in the same time. The interference power at the receiver from J nanosensors in radius a is denoted by:

$$I_a = \sum_{j=1}^{J} P(d_j) \approx \beta(d)^{-\alpha} \quad (11)$$

where $P(d)$ refers to the received power from nanosensor with distance d from receiver, and α and β are constants which depend on pulse and channel properties. For example, when pulse energy is 0.1 aJ and the channel consists of 10% water vapor molecules, $\alpha \approx 2.1$ and $\beta \approx 1.39 \times 10^{-18}$.

The interference distribution can be described by its power spectral density. When spatial distribution of nanosensors is modeled with Poisson distribution and the probability of incoming pulses is distributed uniformly, the power spectral density of interference in nanonetworks can be computed using [14]:

$$f_I(i) = \frac{1}{\pi i} \sum_{k=1}^{\infty} \left[\frac{\Gamma(2k/\alpha + 1)}{k!} \left(\frac{2\pi\beta\lambda p_1 T_s \Gamma(1 - 2/\alpha)}{T_p i^{(2/\alpha)}} \right)^k \right.$$
$$\left. \times \sin(k\pi(1 - 2/\alpha)) \right] \quad (12)$$

where λ refers to node density in node/m^2, T_s is interval between pulses, and T_p pulse duration.

E. Sequential bits 1

In Terahertz band, molecules absorb the energy portion of incoming electromagnetic wave and re-radiate it. Re-radiated signal from molecules in the channel is considered as molecular absorption noise [22]. This kind of noise only appears when channel is excited with pulses. Channel relaxation is related to the required time for the molecules to fully release the absorbed energy. By reducing the number of sequential 1s, molecules in the channel will have more time to waste the absorbed previous pulse energy. Hence, codes with lower number of sequential bits 1 are better.

F. Robustness against transmission errors

Transmission in Terahertz band is mainly affected by path loss (spreading loss and molecular absorption) and molecular absorption noise [22]. Additional channel effects, such as multi-path and nano-particle scattering [23], further distort received pulses. As a result, the receiver can receive erroneous bits.

For channel modeling we use the classical Binary Asymmetric Channel (BAC) model, also described in [10].

1) Codeword Error Rate: One of the important metrics in communication systems is codeword error rate (CER), also known as PER (Packet Error Rate) or BLER (Block Error Rate). The CER is the probability of a codeword received in error, i.e., one or more bits are erroneous. The number of received error bits divided by total number of transmitted bits is defined as bit error rate (BER). The BER can be computed as follows:

$$BER = P(e|X = 0)P(X = 0) + P(e|X = 1)P(X = 1) \quad (13)$$

where $P(e|X = i)$ is the probability of bit error when bit i is transmitted and $P(X = i)$ is the probability to transmit bit i. The probability density function of channel output Y for transmission bit $X = x_i$ is as follows:

$$p(Y|X = x_i) = \frac{1}{\sqrt{2\pi N_i}} e^{-\frac{(y - a_i)^2}{2N_i}} \quad (14)$$

where N_i is the total noise power when bit x_i is transmitted and a_i the amplitude of received signal given by (8). Using asymmetric Terahertz band channel [14], the error transition probabilities are:

$$P(e|X = 0) = P(Y = 1|X = 0) = 1 - \int_A^B p(Y|X = 0)dy \quad (15)$$

$$P(e|X = 1) = P(Y = 0|X = 1) = \int_A^B p(Y|X = 1)dy \quad (16)$$

where A and B are two threshold values. These values can be computed from the intersection of two Gaussian distributions $\mathcal{N}(0, N_0)$ and $\mathcal{N}(a_1, N_1)$ as follows:

$$A, B = \frac{a_1 N_0}{N_0 - N_1} \pm$$
$$\frac{\sqrt{2N_0 N_1^2 log(N_1/N_0) - 2N_0^2 N_1 log(N_1/N_0) + a_1^2 N_0 N_1}}{N_0 - N_1} \quad (17)$$

where a_1 refers to the amplitude of the received signal, N_0 and N_1 stand for the distance dependent noise power given by (6) and (7). For MEC, the bit error rate computation is different from others. Since it has the ability to correct $t = (d_{min}-1)/2$ errors, BER is obtained by [24]:

$$BER_{MEC} = \sum_{i=t+1}^{m} \binom{m}{i} BER^i (1 - BER)^{m-i} \quad (18)$$

The CER can be obtained from BER as follows:

$$CER = 1 - (1 - BER)^m \quad (19)$$

where m is the codeword size.

2) Peak signal to noise ratio: We investigate the robustness against transmission errors also using the classical PSNR, which measures the difference between received and sent data through a formula [25]:

$$PSNR(dB) = 10 \log_{10} \left(\frac{255^2}{E_{ms}} \right) \quad (20)$$

where the mean square error (MSE) is:

$$E_{ms} = \frac{1}{AB} \sum_{x=0}^{A-1} \sum_{y=0}^{B-1} (I_i(x,y) - I_o(x,y))^2 \quad (21)$$

with A and B the image resolution on horizontal and vertical axis, and I_i and I_o the input (sent) and output (received) image respectively. The larger the PSNR value, the closer the received data to the sent data.

IV. THEORETICAL RESULTS

We start by defining the input data stream used.

The transmitted information should have some uncertainty in it (otherwise there is no interest to transmit it), therefore data transmission is usually modelled as a random process. In our model, bits 1 in binary sequences are randomly generated using Bernoulli distribution with probability $P(X = 1) = p$. For each probability of bit 1 (more precisely $p = 0, 0.1, 0.2, 0.3, \ldots, 1$), we generate 10^6 random bits.

Next, the binary stream is divided into symbols of n bits, giving set $A = \{a_1, a_2, \ldots, a_M\}$, with $M = 2^n$. The probability of a symbol to have weight k can be computed using binomial distribution as follows:

$$P(w = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad (22)$$

$$\binom{n}{k} = \frac{n!}{(n-k)! \, k!} \quad (23)$$

If symbols with the same weight occur with the same probability, then probability of i-th symbol can be obtained using:

$$P(A = a_i) = \left[\frac{P(w_i = k)}{n_k} \right], \quad 1 \le i \le M \quad (24)$$

where n_k stands for the number of symbols with weight k. When source transmits L symbols, the occurrence frequency of each symbol is:

$$N_i = P(A = a_i) L, \quad 1 \le i \le M \quad (25)$$

For codes using source statistics, such as ME, NME and unary, encoder sorts the symbols in decreasing order of their occurrence frequency, yielding a new set $B = \{b_1, b_2, \ldots, b_M\}$, where $N_{b_1} \ge N_{b_2} \ge \ldots \ge N_{b_M}$.

A. Energy efficiency

Energy efficiency denotes the code ability to reduce the number of bits 1 from input stream. The more the reduction of number of bits 1, the larger the energy reduction.

In the following, we present the number of bits 1 generated by each method, we define a formula to compute the efficiency, and finally we show the performance of each method in a graph.

The number of bits 1 generated by each method is the following:

1) In ME and NME codes, codewords are sorted in increasing number of bits 1 as follows:

$$w(c_i) = \begin{cases} 0, & i = 1 \\ j, & K(j-1) < i \le K(j) \\ n, & i = 2^n \end{cases} \quad (26)$$

where $w(c_i)$ is the i-th codeword with weight j, and $K(j)$ is the cumulative number of combinations of x bits 1:

$$K(j) = \sum_{x=0}^{j} \binom{n}{x}, \quad j = 1, \ldots, 2^n - 1 \quad (27)$$

2) In PG and NPG codes, codewords have only weight 0 or 1, as follows:

$$w(c_i) = \begin{cases} 0, & i = 1 \\ 1, & 2 \le i \le 2^n \end{cases} \quad (28)$$

3) In MTE code, codewords have weight 0, 1 or 2, cf. (1).
4) In MEC code, codewords have weight 0, $d_{min}/2$ or d_{min}:

$$w(c_i) = \begin{cases} 0, & i = 1, P_{max} > 0.5 \\ d_{min}, & 2 \le i \le 2^n, P_{max} > 0.5 \\ d_{min}/2, & P_{max} < 0.5 \end{cases} \quad (29)$$

5) In LWC code, codewords have constant weight w, as follows:

$$w(c_i) = w, \quad 0 \le i \le 2^n \quad (30)$$

As specified in LWC presentation, we chose $w = 2$.

6) In unary code, codewords have constant weight 1, as follows:

$$w(c_i) = 1, \quad 0 \le i \le 2^n \quad (31)$$

Since source coding maps each symbol into a codeword, the energy efficiency can be obtained by counting the reduction of bits 1 of a symbol compared to its codeword, as follows:

$$\xi_i = \frac{N_i \left(w(x|x = a_i, b_i) - w(c_i) \right)}{N_i \left(w(x|x = a_i, b_i) \right)} \quad (32)$$

where $w(x|x = a_1, b_i)$ is the weight of i-th symbol (a_i from set A for codes without statistics, or b_i from set B for codes with source statistics), and $w(c_i)$ is the weight of i-th codeword. The energy efficiency for the input stream is then:

$$\xi = \sum_{i=1}^{M} \xi_i \quad (33)$$

Energy efficiency for all codes for various probabilities of bit 1 when input symbol size of 4 bits is shown in Fig. 1. NPG yields the best result for all probabilities of bits 1. MEC, LWC and unary even produce negative energy efficiency (i.e., require more energy than uncoded transmission).
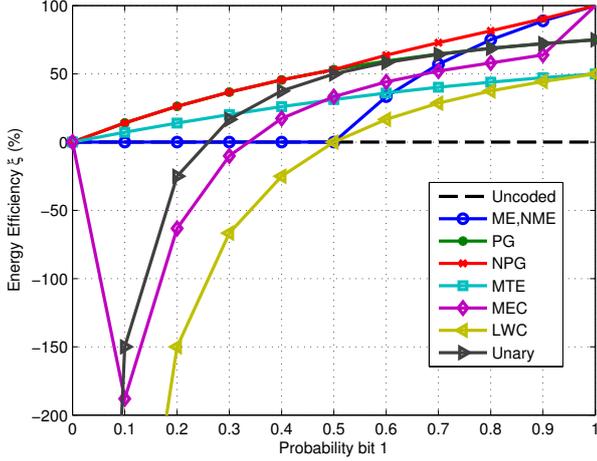
Fig. 1. Energy efficiency for codes with various probability of bit 1, in theory.

B. Bandwidth expansion

In ME and NME, n bits input is mapped to n bits output, so there is no expansion in bandwidth. In PG and NPG, n bits input is mapped to codeword with output size $m_{pg} \geq 2^n - 1$ bits. In MTE, n bits is mapped to codewords with maximum weight 2 (with sequential bits 1). The total number of codewords are 1 (codewords weight 0) + m (codewords weight 1) + (m - 1) (codewords weight 2). Therefore, output size of MTE is

$$1 + m + (m - 1) \geq 2^n \quad (34)$$

thus

$$m_{mte} \geq 2^{n-1} \quad (35)$$

For LWC, output size must fulfill (3). For constant weight $w = 2$, LWC output size is

$$m(m - 1) \geq 2^{n+1} \quad (36)$$

If m is very large then $m(m - 1) \approx m^2$, therefore

$$m_{LWC} \geq 2^{\frac{n+1}{2}} \quad (37)$$

In unary code, output size varies from 1 to 2^n bits. The average output size for unary is

$$\bar{m}_{unary} = \frac{\sum_{i=1}^{2^n} i}{2^n} \quad (38)$$

Bandwidth expansion for codes is the ratio between the minimum output size and input size n. The smaller the bandwidth expansion, the better the code. It is shown in Tab. III and graphically in Fig. 2. It shows that ME and NME (followed by LWC) codes have the smallest bandwidth expansion, since input symbol size equals codeword size. It also shows that MEC code has the largest bandwidth expansion; the reason is that MEC code requires large codeword size to maintain the minimum distance, needed to have error correction.

Coding	Bandwidth expansion
NME, ME	n/n
NPG, PG	$(2^n - 1)/n$
MTE	$2^{n-1}/n$
MEC	$d_{min} 2^{n-1}/n$
LWC	$2^{\frac{n+1}{2}}/n$
Unary	$\left(\sum_{i=1}^{2^n} i \right) / (n 2^n)$

TABLE III
BANDWIDTH EXPANSION, IN THEORY.
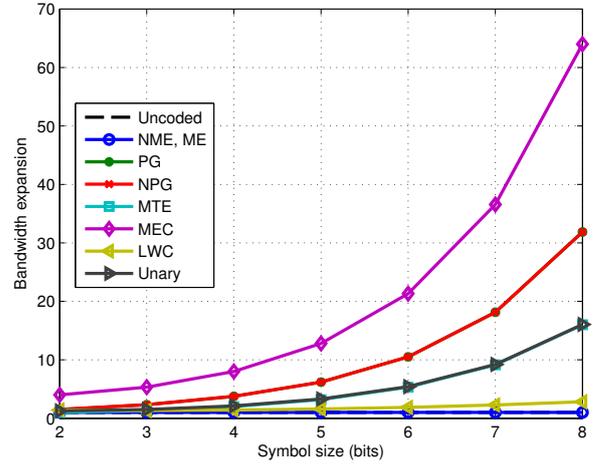

Fig. 2. Bandwidth expansion for various codes, in theory (PG is the same as NPG).

C. Information rate

By using probability of transmitting bit 1 P_{pulse} (detailed later in (39)) and bandwidth expansion from Table III, information rate after coding for all codes is shown if Fig. 3. It shows that uncoded has the largest information rate, while MEC the lowest one.
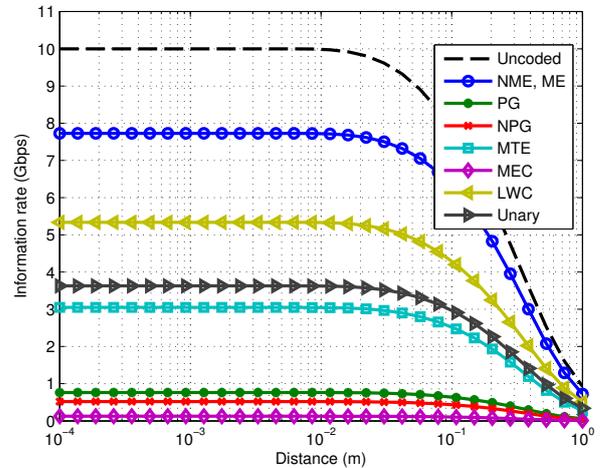

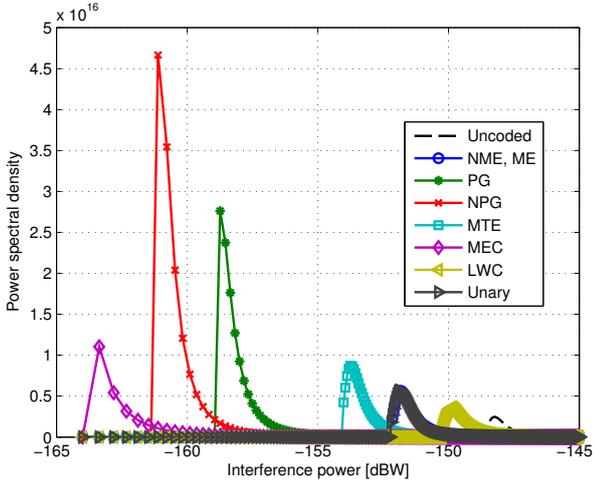Fig. 3. Information rate for various codes, in theory.

Fig. 4. Multi-user interference distribution for all codes, in theory.

Fig. 5. Bit error rate for various codes, in theory.

D. Multi-user interference

Interference in nanonetwork is influenced by probability of transmitting pulse, which depends on the number of transmitted pulses compared to the number of transmitted bits:

$$P_{pulse} = \frac{N_{pulse}}{N_{bits}} \qquad (39)$$

The number of transmitted bits is:

$$N_{bits} = \sum_{i=1}^{2^n} N_i m_i \qquad (40)$$

where n is the input symbol size (n bits symbol), N_i the frequency (number of occurrences) of i-th symbol, as given by (25), and m_i the size of i-th codeword. The number of transmitted pulses, i.e., the number of bits 1 transmitted, is:

$$N_{pulse} = \sum_{i=1}^{2^n} N_i w(c_i) \qquad (41)$$

where $w(c_i)$ is the weight of i-th codeword.

By taking the average value of P_{pulse} for all probabilities of bit 1 (from 0 to 1), m_i for each code using value in Tab. III, and input size $n = 4$, interference for all codes is shown in Fig. 4. It shows that the best one is MEC, while uncoded has the biggest interference.

E. Sequential bits 1

Sequential bits 1 occurs in one of the following conditions:

- Codeword weight is greater or equal than 2, and there is sequential bits 1 in the codeword.
- Subsequent codewords, where the first codeword ends in bit 1 and the second codeword starts with bit 1.

As shown in Table I, this occurs in PG when the first codeword follows the (2^n-1)-th codeword. In NPG this occurs when the second codeword follows the 2^n-th codeword. In MTE, this occurs when first codeword follows the $(n + 1)$-th codeword, and also for $(n + 2)$-th to $(2^n - 1)$-th codewords. In unary, this occurs only when first codeword occurs after
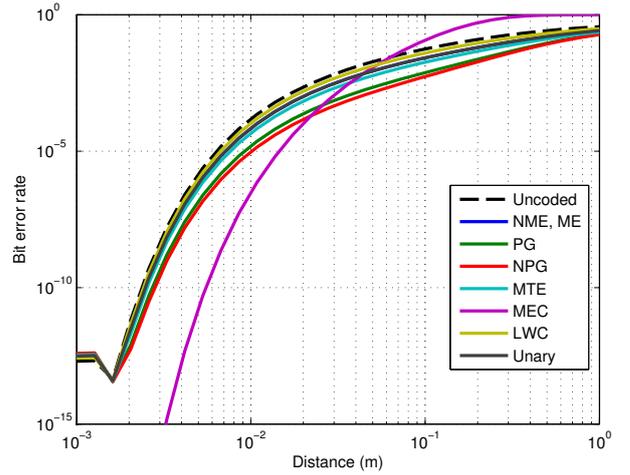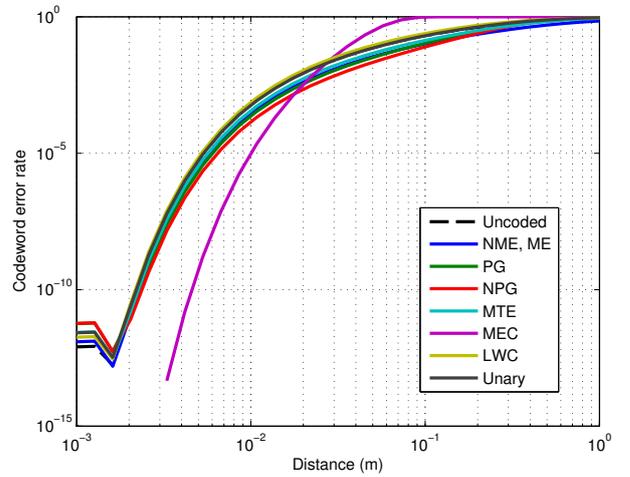

Fig. 6. Codeword error rate for various codes, in theory.

other codewords. To sum up, it appears seldom in PG, NPG and unary, and frequently in MTE.

F. Robustness against transmission errors

1) Codeword error rate: By using probability of transmitting bit 1 P_{pulse} in (39) and noise power in (6) and (7), BER for all codes is shown in Fig. 5. It shows that NPG is the best code (lowest BER) starting from about 2 cm. Before that threshold, MEC is the best code, thanks to its 1 bit error correction.

The CER, as given by (19), is shown in Fig. 6. Since CER is correlated to BER, the CER performance of codes is similar to their BER performance respectively.

2) Peak signal to noise ratio: The hypothetical data we use is not an image, hence the PSNR, which is specific to images, does not apply, hence no results can be shown. However, we present some general information about the problems the codes might create for PSNR computation.

Errors during data transmission affect the decoding process at the receiver, which decodes the received codeword into
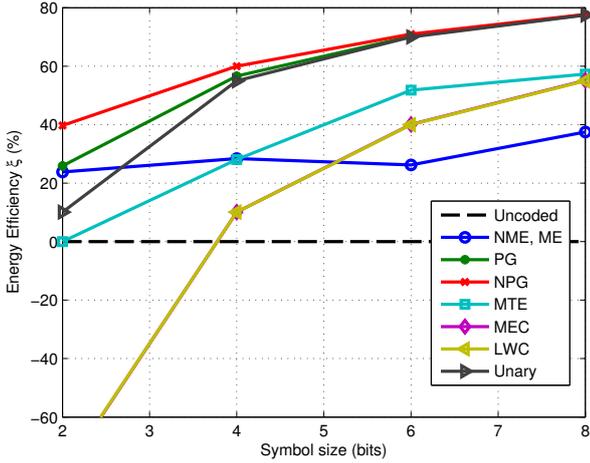
Fig. 7. Energy efficiency for various codes for cancer file (LWC is the same as MEC).

Fig. 8. Probability density function of multi-user interference in nanonetworks for various codes for cancer file.

another codeword. A 1-bit error yields another codeword, which is decoded as another, completely different symbol. As such, a 1-bit transmission error generally becomes a 1-symbol-length error. The error can lead to an existent (but wrong) input symbol, or into an inexistent (not in the codebook) symbol, in which case receiver can decode it into symbol 0.

More severely, in codes using variable-length codeword (unary), a 1-bit transmission error can lead to a 1 symbol being received as *several* symbols, and vice versa. For example, looking at unary code in Table I, when the sender sends twice the symbol 0001 (corresponding to input symbol 100) and the middle 1 is incorrectly received as 0 because of transmission error, the received data becomes 00000001, which is decoded as 1 symbol (000).

V. NUMERICAL RESULTS

We recall that one of the applications of internet of nano-things is video transmission using nano-cameras. In this section, we present numerical results, obtained with MATLAB, when transmitting a file. The file to send is the real image of a cancer cell obtained by an electron microscope, as one of the applications of nanonetworks is medical field. The image is a greyscale image of type bmp and has 256×256 pixels.

A. Energy efficiency

As we recall, the larger the energy efficiency the fewer the number of bits 1 to send. Energy efficiency is measured using (4).

The cancer image has 532,912 bits, and contains 296,399 bits 1 (probability of bit 1 is 0.57). Energy efficiency for all codes is shown in Fig. 7. NPG code has the highest average energy efficiency, followed by PG and unary. MEC and LWC codes have negative energy efficiency for 2-bit input size; this is because in the coding table the codeword weight is 2, and the input symbol size is 2 bits (so the weight between 0 and 2, hence equal or smaller than codeword weight).
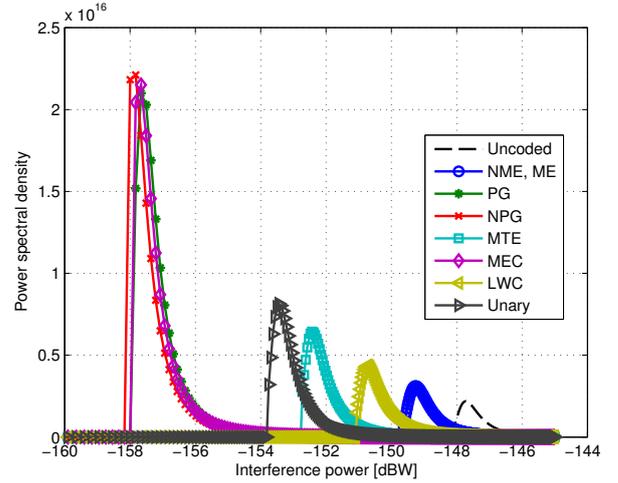
B. Bandwidth expansion

Bandwidth expansion for cancer file gives very similar results to theoretical ones, shown in Fig. 2.

C. Information rate

Information rate after coding, computed using (5), for cancer file shows similar results and exactly the same trend as Fig. 3 in theoretical results section. The large value of bandwidth expansion in MEC greatly reduces its information rate.

D. Multi-user interference

Probability density function of muti-user interference can be obtained using (12). We use parameters: $Ts/Tp = 1000$, and node density $\lambda = 0.1$ nodes/m^2. Code performance in multi-user interference is shown in Fig. 8. NPG code has the lowest interference, since it transmits a large amount of data with very small probability of bit 1. NME and ME codes reduce interference power around 1/2 (-4 dB, from -147.5 dbW to -150 dbW), while PG code reduces around 1/10 interference power (-13 dB) compared to uncoded transmission.

The result for MEC is slightly different than in theoretical case. This is because in a practical application, a symbol with $P_{max} > 0.5$ rarely occurs; for example, the maximum symbol occurrence for cancer file is $P_{max} = 0.011$. As such, there is no codeword with weight 0 and the probability of bit 1 is $P_1 = 0.13$. In theoretical case, the code could sometimes use the codeword with weight 0 (for the most frequent symbol), and as a result the probability of bit 1 is $P_1 = 0.063$, which is obtained by taking the average value of all probabilities of bit 1.

E. Sequential bits 1

The smaller number of sequential bits 1, the better, due to channel relaxation in Terahertz band. Fig. 9 shows the code performance according to the number of sequential bits 1. PG and NPG codes have the fewest sequential bits 1.
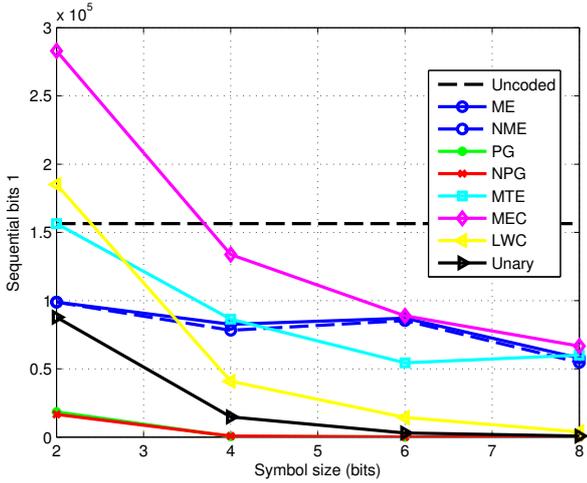
Fig. 9. The number of sequential bits 1 for various codes (bottom line corresponds to PG and NPG) for cancer file.
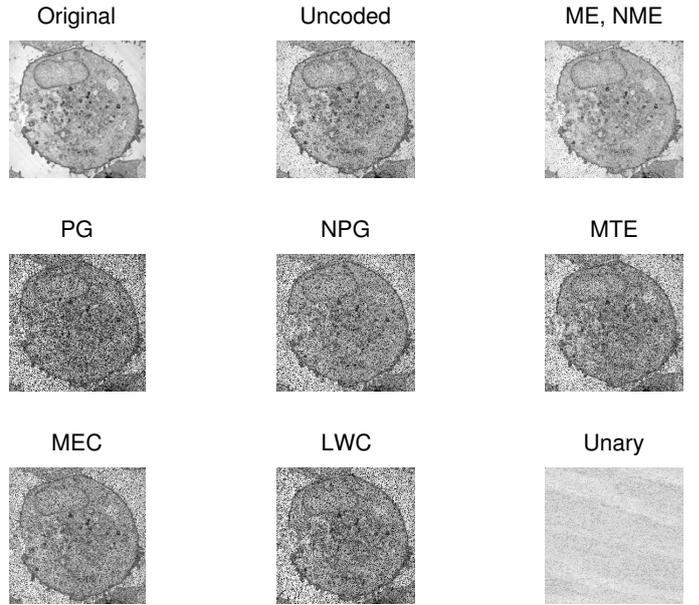

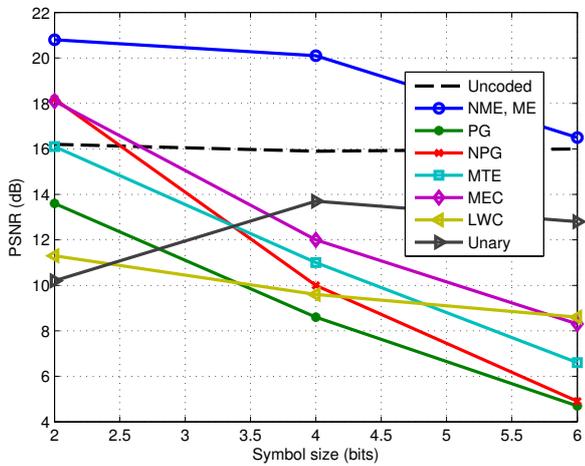Fig. 11. Reconstructed image at receiver for various codes for cancer file.

code are so distorted that they are unrecognisable; these codes are drastically vulnerable to transmission errors and require error correction code.

Summary results

If we compare these results with conclusions from theoretical results (Sec. IV), metric by metric, we notice that they are similar in almost all cases. For example, in both theory and cancer file, for energy efficiency the best code is NPG and the worst is uncoded, and for bandwidth expansion ME and NME are the best and MEC is the worst. For multiuser interference, the small difference is due to difference in probability of the most frequent symbol (P_{max}) in theoretical example and cancer file.

Code performance for both theory and cancer file is summarised in Table IV.

There are three codes which have good results in almost all the cases: PG/NPG and unary. PG/NPG are very bad in bandwidth expansion, whereas unary is very bad in robustness.

All codes (except MEC and LWC for input size equal to 2) have positive energy efficiency; this is not surprising, since we mainly evaluated low weight codes.

Fig. 10. PSNR values (robustness) for various codes for cancer file.

F. Robustness against transmission errors

We use the following parameters: distance between transmitter and receiver is 10 cm, so for BAC channel model $P(Y = 1|X = 0) = 0.0176$ and $P(Y = 0|X = 1) = 0.0812$.

1) Codeword error rate: The file transmission gives similar results to the theoretical results shown in Fig. 5 and 6.

2) Peak signal to noise ratio: The larger the PSNR, the higher the quality of the received image. Fig. 10 shows the code performance using PSNR. The best results are given by ME and NME, followed by NPG. Codes with variable codeword lengths (unary) give the worst results. First, in image transmission, erroneous received codewords appear as erroneous pixels, and we recall that a 1-bit transmission error generally becomes a 1-symbol-length error. Moreover, for unary codes, in case of transmission errors and due to variable-length codeword, 1 sent symbol can be received as 2 or several symbols, and vice versa, as written in the previous section.

This can be easily seen in Fig. 11, which presents the reconstructed image at receiver. Images received using unary

VI. CONCLUSIONS

This article compares various low weight codes, and a novel one we propose (NPG, a small variation of PG) using several criteria appropriate in nanonetworks, especially energy efficiency, bandwidth expansion, information rate, multi-user interference, and robustness. The results show that, even if there is no clear winner, NPG and PG have very good performance compared to the others for all criteria used, except in bandwidth expansion.

Code	Energy efficiency	Bandwidth expansion	Sequential bits 1	Multi-user interference	Robustness
ME, NME	++	++	−	−	++
PG, NPG	++	− −	++	++	+
MTE	++	−	−	+	+
MEC	+	− −	− −	++	+
LWC	+	−	−	+	+
Unary	++	+	+	+	− −

TABLE IV

SUMMARY OF PERFORMANCE OF VARIOUS CODES, BOTH IN THEORY AND FOR CANCER FILE: ++ VERY GOOD, + GOOD, − BAD, − − VERY BAD.

REFERENCES

[1] I. F. Akyildiz, F. Brunetti, and C. Blázquez, "Nanonetworks: a new communication paradigm," *Computer Networks*, vol. 52, no. 12, pp. 2260–2279, 2008.

[2] J. M. Jornet and I. F. Akyildiz, "The internet of multimedia nano-things," *Nano Communication Networks*, vol. 3, no. 4, pp. 242–251, Dec. 2012.

[3] B. Liu, Y. Lai, and S.-T. Ho, "High spatial resolution photodetectors based on nanoscale three-dimensional structures," *Photonics Technology Letters*, vol. 22, no. 12, pp. 929–931, Jun. 2010.

[4] M. C. Hegg, M. P. Horning, T. Baehr-Jones, and M. Hochberg, "Nanogap quantum dot photodetectors with high sensitivity and bandwidth," *Applied Physics Letters*, vol. 96, no. 10, pp. 10 118– 10 118–3, Mar. 2010.

[5] C. Falconi, A. D'Amico, and Z. L. Wang, "Wireless joule nanoheaters," *Sensor and Actuators B*, no. 127, pp. 54–62, Jul. 2007.

[6] I. F. Akyildiz and J. M. Jornet, "Electromagnetic wireless nanosensor networks," *Nano Communication Networks*, vol. 1, no. 1, pp. 3–19, Mar. 2010.

[7] J. Chomoucka, J. Drbohlavova, D. Huska, V. Adam, R. Kizek, and J. Hubalek, "Magnetic nanoparticles and targeted drug delivering," *Pharmacological Research*, vol. 62, no. 2, pp. 144–149, Jan. 2010.

[8] J. M. Jornet and I. F. Akyildiz, "Femtosecond-long pulse-based modulation for Terahertz band communication in nanonetworks," *IEEE Transactions on Communications*, vol. 62, no. 5, pp. 1742–1754, May 2014.

[9] C. Erin and H. Asada, "Energy optimal codes for wireless communications," in *IEEE Proceedings of the 38th Conference on Decision and Control*, Salt Lake City, Utah, USA, Dec. 1999, pp. 4446–4453.

[10] M. A. Zainuddin, E. Dedu, and J. Bourgeois, "Nanonetwork minimum energy coding," in *IEEE International Conference on Ubiquitous Intelligence and Computing (UIC)*, ser. 11. Bali, Indonesia: IEEE, Dec. 2014, pp. 96–103.

[11] Y. Prakash and S. Gupta, "Energy efficient source coding and modulation for wireless applications," in *IEEE Wireless Communications and Networking (WCNC)*, vol. 1, no. 1, New Orleans, USA, 2003, pp. 212–217.

[12] K. Chi, Y.-H. Zhu, X. Jiang, and X. Tian, "Optimal coding for transmission energy minimization in wireless nanosensor networks," *Nano Communication Networks*, vol. 4, no. 3, pp. 120–130, 2013.

[13] M. Kocaoglu and O. B. Akan, "Minimum energy channel codes for nanoscale wireless communications," *IEEE Transaction on Wireless Communications*, vol. 12, no. 4, pp. 1492–1500, Apr. 2013.

[14] J. M. Jornet, "Low-weight error-prevention codes for electromagnetic nanonetworks in the Terahertz band," *Nano Communications Networks*, vol. 5, no. 1–2, pp. 35–44, Mar–Jun 2014.

[15] L. Gammaitoni, I. Neri, and H. Vocca, "Nonlinear oscillators for vibration energy harvesting," *Applied Physics Letters*, vol. 94, p. 164102, 2009.

[16] Z. L. Wang, "Towards self-powered nanosystems: From nanogenerators to nanopiezotronics," *Advanced Functional Materials*, vol. 18, no. 22, pp. 3553–3567, 2008.

[17] P. Wang, J. M. Jornet, M. A. Malik, N. Akkari, and I. F. Akyildiz, "Energy and spectrum-aware MAC protocol for perpetual wireless nanosensor networks in the Terahertz band," *Ad Hoc Networks*, vol. 11, no. 8, pp. 2541–2555, 2013.

[18] M. C. Lemme, "Current status of graphene transistors," *Solid State Phenomena*, vol. 156–158, pp. 499–509, Oct. 2009.

[19] R. Bennewitz, J. N. Crain, A. Kirokosian, J.-L. Lin, J. L. McChesney, D. Y. Petrovykh, and F. J. Himpsel, "Atomic scale memory at a silicon surface," *Nanotechnology*, vol. 13, pp. 499–502, 2002.

[20] L. Rothman, I. Gordon, A. Barbe, D. Benner, P. Bernath, M. Birk, V. Boudon, L. Brown, and A. Campargue, "The HITRAN 2008 molecular spectroscopic database," *Quantitative Spectroscopy and Radiative Transfer*, vol. 110, no. 9–10, pp. 533–572, Jun. 2009.

[21] J. M. Jornet and I. F. Akyildiz, "Information capacity of pulse-based wireless nanosensor networks," in *Proc. of the 8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, Salt Lake City, Utah, USA, Jun. 2011, pp. 80–88.

[22] ——, "Channel capacity of electromagnetic nanonetworks in the Terahertz band," in *IEEE International Conference on Communications (ICC)*, 2010, pp. 1–6.

[23] ——, "Channel modeling and capacity analysis for electromagnetic wireless nanonetworks in the Terahertz band," *IEEE Transactions on Wireless Communications*, vol. 10, no. 10, pp. 3211–3221, Oct. 2011.

[24] B. Sklar, *Digital Communications, Fundamentals and Applications*. Pearson Education Asia, 2001.

[25] Y. Q. Shi and H. Sun, *Image and Video Compression for Multimedia Engineering, Fundamentals, Algorithms, and Standards*. CRC Press, 2007.