Low-weight Channel Codes for Error Prevention in Electromagnetic Nanonetworks in the Terahertz Band

Josep Miquel Jornet
Department of Electrical Engineering
University at Buffalo, The State University of New York
Buffalo, NY, USA
jmjornet@buffalo.edu

ABSTRACT

Nanonetworks consist of nano-sized communicating devices which can perform simple tasks at the nanoscale. Nanonetworks are the enabling technology for long-awaited applications in the biomedical, industrial and military fields. In this paper, the use of low-weight channel codes is proposed as a mechanism to prevent channel errors in nanonetworks. Rather than utilizing channel codes only to detect and correct transmission errors, it is shown that, by appropriately choosing the code weight, both molecular absorption noise and multi-user interference can be mitigated and, ultimately, the number of channel errors can be reduced beforehand. The performance of the proposed scheme is analytically and numerically investigated in terms of achievable information rate after coding and Codeword Error Rate (CER). An accurate Terahertz Band channel model, validated by COMSOL simulation, is used, and novel stochastic models for the molecular absorption noise in the Terahertz Band and for the multi-user interference in pulse-based Terahertz Band communication are developed. The results show that low-weight channel codes can be used to reduce the CER without compromising the achievable information rate. Moreover, it is shown that there is an optimal code weight, which depends on the network conditions, for which the information rate is maximized.

Categories and Subject Descriptors

C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design—*Wireless Communication*; H.1.1 [**Models and Principles**]: Systems and Information Theory—*Information Theory*

General Terms

Theory

Keywords

Nanonetworks, Terahertz Band, Channel Coding

NANOCOM' 14, May 13 - 14 2014, Atlanta, GA, USA

http://dx.doi.org/10.1145/2619955.2619961

1. INTRODUCTION

Nanotechnology is providing a new set of tools to the engineering community to design and manufacture novel electronic nanoscale components, which can perform only limted tasks, such as computing, data storing, sensing and actuation. The integration of several of these nano-components into a single entity will enable the development of more advanced devices, just a few cubic micrometers in size. By means of communication, these nano-devices will be able to achieve complex tasks in a distributed manner [1]. The resulting nanonetworks will enable many applications in the biomedical, environmental, industrial and military fields, such as advanced health monitoring systems, wireless nanosensor networks for biological and chemical attack prevention, or wireless network on chip systems for very large multi-core computing architectures.

For the time being, the communication options for nano-devices are very limited. The miniaturization of a conventional metallic antenna to meet the size requirements of the nano-devices would impose the use of very high operating frequencies (several hundreds of Terahertz), thus limiting the feasibility of nanonetworks. Alternatively, nanomaterials enable the development of nano-antennas that can operate at much lower frequencies. Amongst others, ongoing research on graphene-based nano-transceivers [12, 17] and nano-antennas [10, 20] points to the Terahertz Band (0.1-10.0 THz) as the communication frequency band for nano-devices. The peculiar propagation properties of electrons in graphene [6, 7] enable the creation of compact RF components which can operate at a relatively low frequency, alleviating the energy constraints of nano-devices.

The Terahertz Band brings many opportunities for communication among nano-devices. Contrary to long-range Terahertz Band communication, in which molecular absorption drastically limits the available bandwidth [15], the Terahertz Band behaves as a single transmission window almost 10-THz-wide for distances below one meter [8]. This very large band can theoretically support the transmission at very high rates and enables new communication mechanisms more suited for the very limited capabilities of nano-devices. In this direction, a novel communication scheme based on the transmission of one-hundred-femtosecond-long pulses spread in time known as TS-OOK has been recently proposed [11]. This modulation takes into account the difficulty of generating high-power carrier signals at Terahertz Band frequencies from a nano-transceiver and leverages the state of the art in pulse-based nano-transceivers.

Channel errors in nanonetworks become frequent due to

the impact of both molecular absorption noise in the Terahertz Band [8] as well as the multi-user interference in uncoordinated nanonetworks operating under TS-OOK [11]. Classical error control mechanisms need to be revised before being used in nanonetworks. On the one hand, Automatic Repeat reQuest (ARQ) mechanisms might not be suited for nanonetworks due to the energy limitations of nano-devices, which require nanoscale energy harvesting mechanisms to operate [21]. The very long time needed to harvest enough energy to retransmit a packet make render the data useless. On the other hand, the majority of Forward Error Correction (FEC) mechanisms might be too complex for the expected capabilities of the nano-devices. As we described in [1], the number of nano-transistors in a nano-processor limits the complexity of the operations that it can complete. Even with current processing technologies, the time needed to encode and decode a packet can be much longer than the packet transmission time. As a result, there is a need for novel error control mechanisms.

In this paper, we propose the use of low-weight channel codes for error prevention in electromagnetic nanonetworks. Rather than retransmitting or trying to correct channel errors, we propose to prevent these errors from happening in first instance. In particular, we show that, by reducing the code weight, i.e., the average number of logical "1"s in a codeword, both the molecular absorption noise in the Terahertz Band and the multi-user interference in TS-OOK can be mitigated. As a result, fewer channel errors are generated. Moreover, an optimal coding weight exists for which the number of channel errors can be minimized while maximizing, or at least not penalizing, the achievable information rate. This optimal code weight depends on the channel and network conditions. This result motivates the development of novel link policies that can dynamically adapt the code weight to the channel and network conditions.

The main contributions of this paper are summarized as follows. First, in Section 2, we review our recently proposed pulse-based communication scheme for nanonetworks and develop analytical stochastic models for the two main causes of channel errors in electromagnetic nanonetworks, namely, molecular absorption noise and multi-user interference. These models are based upon an accurate Terahertz Band channel model for nanonetworks and take into account the capabilities of state-of-the-art nano-transceivers. In Section 3, we analytically show that the channel code weight can be reduced to mitigate noise and interference. In Section 4, we analytically and numerically investigate the performance of low-weight channel codes in terms of information rate after coding and Codeword Error Rate (CER). We consider two different receiver architectures: a continuous-output soft receiver and a discrete-output hard receiver. In addition, we show that there is an optimal code weight which can simultaneously maximize the information rate. Finally, we use COMSOL Multi-physics [5] to validate our models by mean of time-domain electromagnetic simulations, and we numerically investigate the information rate and CER achieved when using low-weight channel codes are provided. We conclude the paper in Section 6.

A preliminary version of this work can be found in [9]. Compared to our previous work, in this paper we have i) revised the molecular absorption noise and multi-user interference model; ii) analyzed the performance of low weight codes with a hard receiver architecture, in addition to the soft receiver case; iii) analyzed the CER and the optimal code weight; and iv) performed extensive COMSOL simulations to validate our models.

2. SYSTEM MODEL

In this section, we first review the recently proposed pulse-based communication scheme for nano-devices and then we develop the analytical models for molecular absorption noise and multi-user interference used in our analysis.

2.1 TS-OOK: Time Spread On-Off Keying

In light of the state of the art in graphene-based nanoelectronics, we consider that nano-devices communicate by utilizing TS-OOK [11], a recently proposed transmission scheme for nano-devices, which is based on the transmission of very short pulses, just one-hundred-femtosecond long, by following an on-off keying modulation spread in time. These pulses can be generated and detected with nano-transceivers based on graphene and high-electron-mobility materials such as Gallium Nitride or Indium Phosphide [12, 17]. While conceptually similar to Impulse-radio Ultra-wide-band (IR-UWB) communication [23], note that the pulses in TS-OOK are three orders of magnitude shorter than that of IR-UWB and, thus, a simpler modulation is used instead of orthogonal time hopping with pulse position modulation.

In particular, the functioning of TS-OOK is as follows. The symbol "1" is transmitted by using a one-hundred-femtosecond-long pulse and the symbol "0" is transmitted as silence, i.e., the nano-device remains silent. The time between symbols T_s is much longer than the symbol duration T_p, i.e., $\beta = T_s/T_p \gg 1$. The reason for this is twofold. On the one hand, due to technology limitations, pulses cannot be emitted in a burst. On the other hand, the separation of pulses in time allows for the relaxation of the vibrating molecules in the channel [8]. During the time between symbols, a device can either remain idle or receive other incoming information. Therefore, TS-OOK enables simple multiple-access policies.

2.2 Stochastic Model of the Molecular Absorption Noise

Molecular absorption noise is one of the main channel effects that impact the performance of electromagnetic nanonetworks. EM waves at frequencies in the Terahertz Band create internal vibrations in many types of molecules which are commonly present in nanonetworking scenarios such as water vapor, oxygen or nitrogen, amongst others. As a result, part of the EM energy is first absorbed by the molecules (channel absorption) and then re-radiated (channel emission) [8]. The resulting molecular absorption noise is correlated to the transmitted signal and can be modeled as Additive Colored Gaussian Noise (ACGN).

The probability density function (p.d.f.) $\mathcal{N}$ of the molecular absorption noise at the receiver conditioned to the transmission of symbol x_m (silence for a logical "0" and a pulse for a logical "1"), $f_{\mathcal{N}}(n|X = x_m)$, where n refers to noise, is given by

$$f_{\mathcal{N}}(n|X = x_m) = \frac{1}{\sqrt{2\pi N_m(d)}} e^{-\frac{1}{2}\frac{n^2}{N_m(d)}}, \qquad (1)$$

where N_m refers to the molecular absorption noise power

when symbol m is transmitted, which is given by

$$N_m(d) = \int_B S_{N_m}(f,d)\,|H_r(f)|^2\,df, \quad (2)$$

where B is the receiver's noise equivalent bandwidth, S_{N_m} is the molecular absorption noise power spectral density (p.s.d.), and H_r is the receiver's impulse response. For the time being, we consider a matched-filter receiver architecture for simplicity. The total molecular absorption noise p.s.d. S_{N_m} affecting the transmission of a symbol $m \in \{0,1\}$ is contributed by the background atmospheric noise p.s.d. S_{N^B} [3] and the self-induced noise p.s.d. $S_{N_m^X}$, which are defined as

$$S_{N_m}(f,d) = S_{N^B}(f) + S_{N_m^X}(f,d), \quad (3)$$

$$S_{N^B}(f) = \lim_{d\to\infty} k_B T_0 \left(1-\exp\left(-k(f)\,d\right)\right)\left(\frac{c_0}{\sqrt{4\pi}f_0}\right)^2, \quad (4)$$

$$S_{N_m^X}(f,d) = S_{X_m}(f)\left(1-\exp\left(-k(f)\,d\right)\right)\left(\frac{c_0}{4\pi d f_0}\right)^2, \quad (5)$$

where f stands for the frequency, d refers to the transmission distance, k_B is the Boltzmann constant, T_0 is the room temperature, c_0 is the speed of light in the vacuum, f_0 is the design center frequency, and S_{X_m} is the transmitted signal p.s.d.. k refers to the molecular absorption coefficient, which depends on the molecular composition of the transmission medium, i.e., the type and concentration of molecules found in the channel and it is computed as in [8]. The term S_{N^B} takes into account that the background noise is i) generated from molecules that radiate just for being at a temperature above 0 K, and ii) detected by an isotropic broadband antenna with effective area given by $A_{eff} = \lambda_0^2/4\pi$. The term $S_{N_m^X}$ takes into account that the induced noise is i) generated by the transmitted signal X_m, ii) spherically spread from the transmitting antenna, and iii) detected by an isotropic broadband antenna with effective area A_{eff}.

In Figure 1, the molecular absorption noise power when a logical "0" and a logical "1" are transmitted, N_1 and N_0, respectively, are shown as functions of distance. The transmitted energy per pulse E_p is fixed and equal to 0.1 aJ, in light of the performance of existing nano-transceivers [12, 17]. Two observations are made. First, N_1 decreases with distance, but at a lower rate than the received signal power P when a pulse is transmitted. Second, N_0 does not vary with the distance (the background noise is always the same) and, in addition, $N_0 \leq N_1$ always. Therefore, it is clear that the transmission of logical "0"s is less likely to suffer from channel errors. There are additional noise sources that can introduce errors, such as the electronic thermal noise of the nano-receiver. However, for the time being, there is no accurate noise model for graphene-based electronic devices. Initial predictions of thermal noise in graphene-based devices point to very low noise factors in this nanomaterial [14]. For this, we currently focus only on the channel noise, i.e., the molecular absorption noise.

2.3 Stochastic Model of the Multi-user Interference in TS-OOK

Multi-user interference is another major effect that limits the performance of electromagnetic nanonetworks. TS-OOK enables simultaneous transmission among a very large number of nano-devices [11]. However, in light of the very large nano-device density in the envisioned applications and

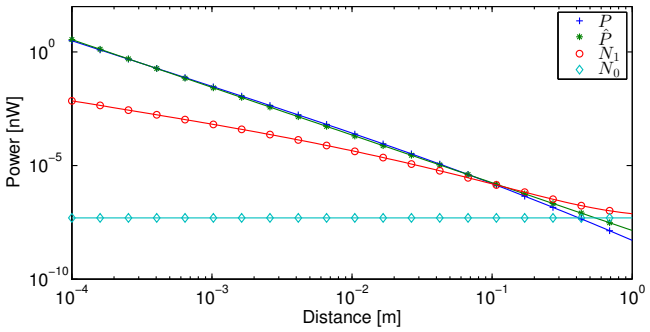


Figure 1: Analytical and approximated received signal power, P and $\hat{P}$, respectively; noise power when a pulse is transmitted N_1, and noise power associated with silence N_0 ($E_p = 0.1$ aJ).

by considering a scenario in which nano-devices can start transmitting at any specific time in an uncoordinated manner, collisions between symbols can occur. These collisions result in interference and this imposes a limitation on the information rate at which nano-devices can communicate.

To quantitatively evaluate the impact of collisions on the system performance, we develop a new stochastic model for the interference power at the receiver which captures the peculiarities of the Terahertz Band channel as well as the properties of TS-OOK. Many stochastic models of interference have been developed to date. For example, an extensive review of the existing models can be found in [2, 4, 22]. However, these models do not capture the peculiarities of the Terahertz Band channel, such as the molecular absorption loss and the additional molecular absorption noise created by interfering nodes. In [11], we analyzed the interference in TS-OOK by modeling it as a Gaussian process. This assumption is valid for the case in which the network is supporting a very high traffic load, and it is shown that it is a useful asset for the analysis of the multi-user achievable information rate. However, in order to be more general, we need much more complete models for the interference.

Our final objective is to have a closed-form expression for the p.d.f. of the interference power I created at the receiver side, f_I. Without loss of generality, we position the receiver at the origin of coordinates. The interference created at the receiver side by the nano-devices $\mathcal{J}$ contained in an area of radius a is given by

$$I_a = \sum_{j\in\mathcal{J}\,|\,d_j\leq a} P(d_j), \quad (6)$$

where P refers to the power of a given signal at a distance d from its transmitter. From [8], P can be written as

$$P(d) = \int_B S_x(f)\,|H_c(f)|^2\,|H_r(f)|^2\,df, \quad (7)$$

where d stands for distance, B refers to the bandwidth of the transmitted signal, S_x is the p.s.d. of the transmitted symbol x, and f stands for frequency. H_c refers to the channel frequency response, which is given by

$$H_c(f) = \left(\frac{c_0}{4\pi f d}\right)\exp\left(-\frac{k(f)\,d}{2}\right), \quad (8)$$

where c_0 refers to the speed of light, and k is the molecular absorption coefficient of the medium. In Figure 1, P is illus-

trated as a function of the distance d by using the channel model developed in [8]. For the distances considered in our analysis, between a few hundred of micrometers and up to one meter, P can be approximated by the polynomial $\hat{P}$

$$P\left(d\right) \approx \hat{P}\left(d\right) = \beta\left(d\right)^{-\alpha}, \tag{9}$$

where α and β are two constants which depend on the specific channel molecular composition as well as on the power and the shape of the transmitted signal. In particular, for a standard medium composition with 10% of water vapor molecules, $\alpha \approx 2.1$ and $\beta \approx 1.39 \cdot 10^{-18}$, when using pulses with $E_p = 0.1$ aJ.

To compute the overall interference created by the nano-devices contained within a disc of radius a, it is necessary to know the spatial distribution of the nodes. In our analysis, we model the positions of the nano-devices as a spatial Poisson point process. Therefore, the probability of finding k nodes in a disc of radius a and area A in m^2 is:

$$P\left[k \text{ in } A\left(a\right)\right] = \frac{\left(\lambda A\left(a\right)\right)^k}{k!} e^{-\lambda A(a)}, \tag{10}$$

where λ refers to the Poisson process parameter in nodes/m^2 and ! stands for the factorial operation.

A collision between symbols will occur when two or more symbols reach the receiver at the same time. In TS-OOK, by considering also a Poisson distribution of the arrivals in time, the probability of having an arrival during T_s seconds is a uniform random probability distribution with p.d.f. equal to $1/T_s$. Thus, for a given transmission, a collision will occur with probability $2T_p/T_s$. Not all types of symbols *harmfully collide*, but only pulses (logical "1"s) create interference. Therefore, we can replace the Poisson parameter λ in (10) by

$$\lambda \rightarrow \lambda' = \lambda_T \left(2T_p/T_s\right) p_X\left(X = 1\right), \tag{11}$$

where λ_T refers to the density of active nodes in nodes/m^2, T_p refers to the symbol length, T_s stands for the time between symbols, and $p_X\left(X = 1\right)$ refers to the probability of a nano-device to transmit a pulse (logical "1").

Following a similar procedure as in [18], to obtain a closed-form solution of the interference power, we first compute the characteristic function of the interference I_a created by the nodes in a disc of radius a, Φ_{I_a}, calculate its limit when the radius a goes to infinity, Φ_I, and obtain the p.d.f. of the interference power f_I as the inverse Fourier transform of Φ_I.

We define the characteristic function of the interference power I_a as

$$\Phi_{I_a}\left(\omega\right) = E\left\{exp\left(\jmath\omega I_a\right)\right\}, \tag{12}$$

which by using conditional expectation and taking into account the spatial Poisson distribution of the nodes, becomes:

$$\begin{aligned}\Phi_{I_a}\left(\omega\right) &= E\left\{E\left\{e^{\jmath\omega I_a}|k \text{ in } A\left(a\right)\right\}\right\} \\ &= \sum_{k=0}^{\infty} \frac{\left(\lambda'\pi a^2\right)^k}{k!} e^{-\lambda'\pi a^2} E\left\{e^{\jmath\omega I_a}|k \text{ in } A\left(a\right)\right\},\end{aligned} \tag{13}$$

where "k in $A\left(a\right)$" refers to the event of having k active nodes in a disk of radius a, and the expectation is over the random variable I_a. To compute this last term, we can proceed as follows. Under the Poisson assumption, when having k nodes in a disc of radius a, their locations follow independent and identically distributed uniform distributions. If R is the distance to the origin from a point that is uniformly distributed in A, then the p.d.f. of R is

$$f_R\left(d\right) = \begin{cases}\left(2r\right)/a^2 & 0 \leq \mathrm{d} \leq \mathrm{a} \\ 0 & \text{otherwise.}\end{cases} \tag{14}$$

Taking into account that the characteristic function of the sum of a number of independent random variables is the product of the individual characteristic functions, we can write

$$E\left\{e^{\jmath\omega I_a}|k \text{ in } A\left(a\right)\right\} = \left(\int_0^a \frac{2r}{a^2} e^{\jmath\omega g(d)} dr\right)^k. \tag{15}$$

By combining (15) in (13), summing the series, and computing the limit when $a \rightarrow \infty$, the characteristic function of the interference power becomes

$$\Phi_I\left(\omega\right) = exp\left(\jmath\lambda'\pi\omega\left(\int_0^{\infty}\left[1/P\left(t\right)\right]^2 e^{\jmath\omega t} dt\right)\right), \tag{16}$$

where λ' refers to spatial Poisson point process parameter as defined in (11) in nodes/m^2 and P is the received power at the origin for a signal transmitted at a distance t in (9).

Finally, for the specific case in which P can be approximated as a polynomial of the form $\beta t^{-\gamma}$, with $0 < \gamma = 2/\alpha < 1$, the integral in (16) may be evaluated to obtain:

$$\Phi_I\left(\omega\right) = exp\left(-\lambda'\pi\beta\Gamma\left(1-\gamma\right)e^{-\pi\gamma/2}\omega^{\gamma}\right), \tag{17}$$

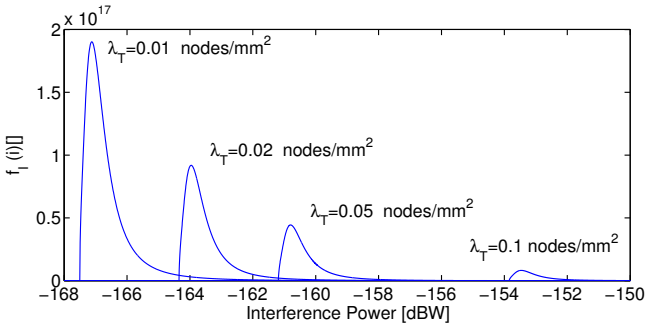
where $\Gamma\left(\cdot\right)$ stands for the gamma function. For $0 < \gamma < 1$, the p.d.f. of I can now be obtained by taking the inverse Fourier transform, and results in

$$f_I\left(i\right) = \frac{1}{\pi i}\sum_{k=1}^{\infty}\frac{\Gamma\left(\gamma k+1\right)}{k!}\left(\frac{\pi\lambda'\beta\Gamma\left(1-\gamma\right)}{i^{\gamma}}\right)^k \sin k\pi\left(1-\gamma\right), \tag{18}$$

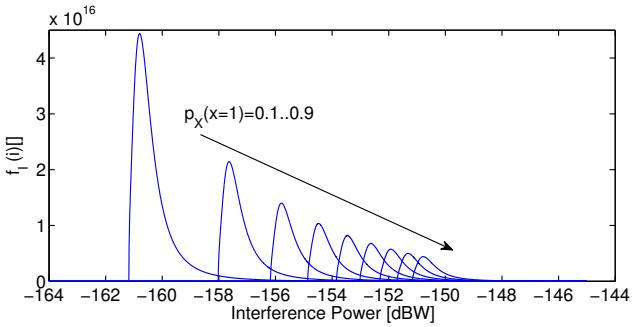
where λ' refers to the spatial Poisson point parameter given by (11), $\gamma \approx 0.95$ and $\beta \approx 1.39 \cdot 10^{-18}$.

In Figure 2(a), the p.d.f. f_I of the interference power (18) is illustrated for different values of λ'. In particular, λ' is obtained from (11), with $T_s/T_p = 1000$, $p_X\left(X = 1\right) = 0.5$ and for λ_T ranging between 0.01 nodes/mm^2 and 0.1 nodes/mm^2. For example, the interference created by a Poisson field of nano-devices with $\lambda_T = 0.1$ nodes/mm^2 which are operating under the previous conditions, has an average power of approximately -153 dBW. When the node density is decreased to $\lambda_T = 0.01$ nodes/mm^2, this value goes to -167 dBW.

In Figure 2(b), the p.d.f. f_I of the interference power (18) is illustrated for different values of $p_X\left(X = 1\right)$, with $T_s/T_p = 1000$ and $\lambda_T = 0.1$ nodes/mm^2. It can be seen that the overall interference can decrease in more than 10 dB when the probability to transmit a pulse (logical "1") changes from 0.9 to 0.1. Based on these results as well as on the outcomes of the molecular absorption model, it is clear that the transmission of pulses (logical "1"s) increases both the total molecular absorption noise and the multi-user interference, which potentially results in a higher number of channel errors. New error control mechanisms can be developed by exploiting this unique behavior, as we present next.



(a) Varying λ_T, $p_X\,(X=1)=0.5$



(b) Varying $p_X\,(X=1)$, $\lambda_T = 0.1$ nodes/mm^2

Figure 2: Probability density function of the interference power for different transmitting node densities λ_T (left) and different probabilities to transmit a pulse $p_X\,(X=1)$ (right), when $T_s/T_p = 1000$.

3. ERROR PREVENTION WITH LOW-WEIGHT CHANNEL CODES

In existing communication systems, channel codes are used to allow the receiver of a message to detect and correct transmission errors. Going one step ahead, we propose to use channel codes to reduce the chances of creating these errors in first instance. Our aim is not to develop new types of error correcting codes, but to analytically and numerically show how by controlling the code weight, i.e., the average number of bits equal to "1" in a codeword, of any type of codes, the molecular absorption noise power and the interference power can be reduced without compromising the information rate.

Existing channel codes generally make use of all the possible codewords independently of their weight. However, it is sometimes desirable to limit the values that the weight of the codewords can take. In this direction, ration coding techniques were proposed in [19] to reduce the electronic noise in chip interconnects. By keeping the weight of the codewords constant, it was shown that the electronic noise can be reduced. In a similar direction, in [16], the performance of sparse Low Density Parity Check (LDPC) codes was analyzed in terms of error probability as a function of their weight. In particular, the authors showed that for the binary symmetric channel and the parallel Z channel, the block error probability of LDPC codes could be reduced by reducing the code weight. In [13], the authors present a way to construct low-weight minimum energy Hamming codes for nanonetworks. To the best of our knowledge, these are the only papers in which the impact of the code weight in the performance of a communication system is investigated.

Based on the molecular absorption noise model and the multi-user interference model introduced in Section 2, it is clear from (2), (11) and (18) that the probability of transmitting a pulse (logical "1") is directly related with the molecular absorption noise and the interference behaviors. By controlling the weight of the transmitted codewords, the probability distribution of "1"s and "0"s can be modified. Ultimately, by using *constant low-weight channel codes*, we can reduce the molecular absorption noise and the interference of the system. This reduction comes with the price of longer codewords, as usually in order to uniquely code a message with a lower weight, it will be necessary to use a larger number of bits.

To illustrate this effect, we proceed as follows. In our analysis, the length of an unencoded message is constant and equal to n bits. For a given n, the total number of possible n-bit words is given by 2^n. The length of an encoded message is $m \geq n$ bits, and its weight, which is defined as the number of bits equal to "1", is denoted by u. For a given m, the total number of possible codewords with weight exactly equal to u is given by:

$$\mathcal{W}(m,u) = \frac{m!}{(m-u)!u!}. \tag{19}$$

Therefore, in order to be able to encode all the possible n-bit source messages into *fixed weight* u codewords, the following condition must be satisfied:

$$\mathcal{W}(m,u) \geq 2^n, \tag{20}$$

where m refers to the length of the encoded words. For example, for $n = 32$ bits, a total of $m = 35$ bits are needed to generate 2^{32} codewords with exact weight equal to $u = 17$.

For a constant weight code, the probability $p_X\,(X=1)$ of transmitting a logical "1" or pulse, and the probability $p_X\,(X=0)$ of transmitting a logical "0", *i.e.*, being silent, are:

$$p_X\,(X=1) = \frac{u}{m}; \quad p_X\,(X=0) = \frac{m-u}{m}, \tag{21}$$

where m stands for the number of bits in the encoded message, u stands for its weight and both m and u must satisfy (20). In Figure 3, the necessary codeword length in bits and the codeword weight necessary to achieve a specific probability of transmitting a pulse when encoding 32-bit messages is shown. For example, in order to achieve a probability of pulse transmission $p_X\,(X=1) = 0.3$, the encoded message length is $m = 42$ bit with weight $u = 12$.

Note that we are not advocating for any specific type of coding scheme. We are mainly estimating the additional number of bits which are necessary to obtain a constant low-weight set of codewords. Ideally, low-weight error correction codes can be built in one step. Alternatively, these can be obtained in a two cascade coding design. In any case, by reducing the weight of the code, we can reduce the molecular absorption noise and interference in the network. However, due to the fact that additional bits are being transmitted to reduce the coding weight, it is intuitive to think that the amount of useful information that can be transmitted per unit of time is reduced. This compromise is analyzed next.

4. PERFORMANCE ANALYSIS

In this section, we analytically study the impact of low weight channel codes on the information rate after coding

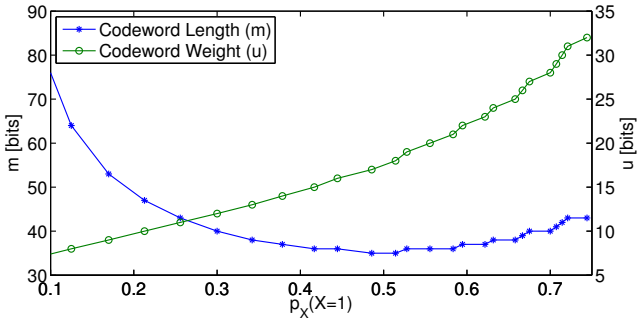


Figure 3: Codeword length m and constant code weight w as a function of the target probability of transmitting a logical "1".

and the Codeword Error Rate (CER). In addition, we show the existence of an optimal coding weight that maximizes the information rate after coding.

4.1 Information Rate After Coding

In [11], we computed the achievable information rate with TS-OOK over the asymmetric Terahertz Band channel in the absence of coding. Next, we compute the achievable information rate after coding of TS-OOK with constant low-weight channel codes. In particular, the information rate after coding $IR_{u-coded}$ in bits/second is given by

$$IR_{u-coded} = \frac{B}{\beta}\frac{n}{m}\left(H\left(X\right) - H^{I}\left(X|Y\right)\right), \tag{22}$$

where n and m are the unencoded and encoded message lengths respectively, B stands for the bandwidth, β is the ratio between the symbol duration and the pulse length, X refers to the source, Y refers to the output of the channel, $H\left(X\right)$ refers to the entropy of the source X, and $H^{I}\left(X|Y\right)$ stands for the equivocation of the channel with interference.

We model the source of information X as a discrete binary random variable and, thus, the source entropy $H\left(X\right)$ is

$$H\left(X\right) = \sum_{x=0}^{1} p_X\left(X=x\right)\log_2\frac{1}{p_X\left(X=x\right)} \tag{23}$$

where $p_X\left(X=x\right)$ refers to the probability of transmitting the symbol $X=x$ (silence for a logical "0" and a pulse for a logical "1") and is given by (21).

The output of the transmitter is attenuated by the channel and corrupted by molecular absorption noise and interference. We consider the channel behavior to be deterministic. Thus, the only random components affecting the received signal are the noise and the interference. For the receiver, we consider both a soft receiver architecture, which provide us with a continuous output, and a hard receiver architecture, whose output is discrete.

4.1.1 Information Rate with a Soft Receiver

When using a soft receiver architecture, the output Y is continuous. The p.d.f. $f_Y\left(y|X=x\right)$ of the channel output Y conditioned to the transmission of the symbol is given by

$$f_Y\left(y|X=x\right) = \delta\left(y-a_m\right) * f_{\mathcal{N}}\left(n=y|X=x\right) * \left(2y f_I\left(i=y^2\right)\right) \tag{24}$$

where δ stands for the Dirac delta function, a_m stands for the received symbol amplitude, obtained from (7), $f_{\mathcal{N}}$ is the p.d.f. of the noise given by (1), f_I stands for the p.d.f. of the interference power given by (18), and $*$ denotes convolution. This equation considers the overall interference power I to be independent of the current symbol transmission $X=x$.

The equivocation of the channel $H^{I}\left(X|Y\right)$ is given by

$$H^{I}\left(X|Y\right) = \int_y \sum_{x=0}^{1} f_Y\left(y|X=x\right)p_X\left(X=x\right) \cdot \log_2\left(\frac{\sum_{q=0}^{1} f_Y\left(y|X=q\right)p_X\left(X=q\right)}{f_Y\left(y|X=x\right)p_X\left(X=x\right)}\right)dy. \tag{25}$$

The maximum achievable information rate in this case can be obtained by combining (23), (24) and (25) in (22). Analytically solving the resulting equation is not feasible. Instead of this, we numerically investigate the achievable information rate after coding with a soft receiver in Section 5.

4.1.2 Information Rate with a Hard Receiver

A soft receiver with a continuous output infers the maximum achievable information rate at the cost of complexity. In practice, due to the limitation on the internal data size in nano-devices, quantization is unavoidable. For this, in our analysis, we also consider a 1-bit hard receiver architecture tailored to the asymmetric Terahertz Band channel. With a 1-bit hard receiver, the channel becomes a Binary Asymmetric Channel (BAC) and Y is now a discrete random variable. This channel is fully characterized by the four transition probabilities:

$$\begin{aligned}
p_Y\left(Y=0|X=0\right) &= \int_{th_1}^{th_2} f_Y\left(y|X=0\right)dy,\\
p_Y\left(Y=1|X=0\right) &= 1 - p_Y\left(Y=0|X=0\right),\\
p_Y\left(Y=0|X=1\right) &= \int_{th_1}^{th_2} f_Y\left(y|X=1\right)dy,\\
p_Y\left(Y=1|X=1\right) &= 1 - p_Y\left(Y=0|X=1\right),
\end{aligned} \tag{26}$$

where $f_Y\left(y|X=x\right)$ is the probability of the channel output Y conditioned to the transmission of the symbol $X=x$ in (24) and th_1 and th_2 are two threshold values. Contrary to the classical symmetric additive Gaussian noise channel, in the asymmetric channel, there are two points at which $f_Y\left(y|X=0\right)$ and $f_Y\left(y|X=1\right)$ intersect. We consider these thresholds to be defined for the case without interference. Thus, th_1 and th_2 can be analytically computed from the intersection between two Gaussian distributions $\mathcal{N}\left(0,N_0\right)$ and $\mathcal{N}\left(a_1,N_1\right)$ respectively, which results in

$$th_{1,2} = \frac{a_1 N_0}{N_0 - N_1} \pm \frac{\sqrt{2N_0N_1^2\log\left(N_1/N_0\right) - 2N_0^2N_1\log\left(N_1/N_0\right) + a_1^2N_0N_1}}{N_0 - N_1}, \tag{27}$$

where a_1 is the amplitude of the received signal given that a pulse has been transmitted and N_0 and N_1 stand for the distance dependent noise powers given by (2).

The equivocation of the channel $H^I_{BAC}(X|Y)$ for the BAC can be written as

$$H^I_{BAC}(X|Y) = \sum_{y=0}^{1}\sum_{x=0}^{1} p_Y(Y=y|X=x)\,p_X(X=x) \cdot \log_2\left(\frac{\sum_{q=0}^{1} p_Y(Y=y|X=q)\,p_X(X=q)}{p_Y(Y=y|X=x)\,p_X(X=x)}\right). \quad (28)$$

Finally, the information rate can be obtained by combining (22), (26) and (28). In Section 5, we numerically analyze the information rate of low-weight codes with a hard receiver with double threshold.

4.1.3 Optimal Coding Weight

From the information rate $IR_{u-coded}$ given by (22), we can make the following statements:

- For a message length n, the encoded message length m increases exponentially when reducing the coding weight u, according to (20) and (19), as shown in Figure 3.
- The source entropy $H(X)$ given by (23) is maximized when $p_X(X=1) = p_X(X=0) = 0.5$. The coding weight u that maximizes the source entropy is $n/2$.
- The channel equivocation $H(X|Y)$ both for a soft receiver (25) and for a discrete output hard receiver (28) decreases when the coding weight u is reduced.

As a result, we can state that there is an optimal coding weight for which the information rate is maximized. The optimal coding weight depends on the channel conditions, *i.e.*, the molecular absorption noise (1), and the network conditions, *i.e.*, the multi-user interference (18). Analytically finding the optimal coding weight is not feasible. Alternatively, we numerically investigate the optimal coding weight that maximizes the information rate in Section 5.

4.2 Codeword Error Rate

In addition to the information rate, an additional relevant metric is the CER. The use of low-weight channel codes prevents the generation of channel errors due to molecular absorption noise and multi-user interference. This turns into a reduced symbol error rate (SER). The SER is given by

$$SER = p_Y(Y=1|X=0)\,p_X(X=0) + p_Y(Y=0|X=1)\,p_X(X=1), \quad (29)$$

where $p_Y(Y=y|X=x)$ is the probability of receiving symbol $Y=y$ given that symbol $X=x$ has been transmitted, which is given by (26), and $p_X(X=x)$ is the probability to transmit symbol $X=x$, which depends on the coding weight u and it is given by (21). By assuming the independence of errors in time, the CER can be computed as

$$CER = 1-(1-SER)^m, \quad (30)$$

where m is the encoded message length. We numerically investigate the CER in Section 5.

5. NUMERICAL RESULTS

In this section, we numerically investigate the performance of constant low-weight channel codes in terms of information rate after coding as well as CER. In our analysis, the unencoded message length n is equal to 32 bit. The code weight u and the resulting probabilities to transmit a pulse or silence, $p_X(X=1)$ and $p_X(X=0)$, respectively, are computed accordingly by using (19) and (21).

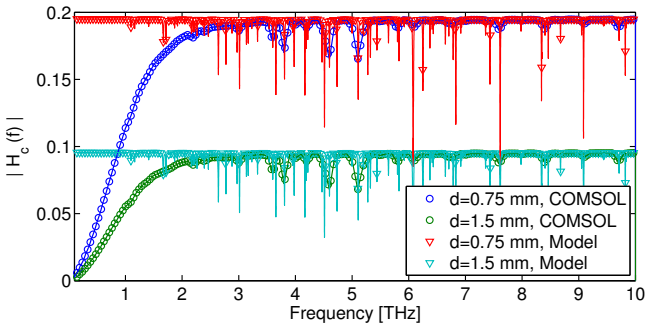


Figure 4: Validation of the Terahertz Band channel response with COMSOL.

We utilized the models for the received signal power (7), molecular absorption noise power (2) and multi-user interference power (18) developed in Section 2. These models strongly depend on the Terahertz Band channel, which we validate by means of COMSOL Multi-physics [5]. Extensive frequency domain simulations are conducted to characterize the channel response when an ideal electric point dipole is used in transmission and reception. In Figure 4, we illustrate the channel frequency response H_c both as analytically obtained from (8) and simulated with COMSOL. In both cases, we consider a standard medium composition with a 10% of water vapor molecules. The analytical model can accurately reproduce the simulation results both in terms of frequency and distance. There is a discrepancy between the models below 1 THz, which is related to the frequency response of the transmitting and receiving antenna. Finally, as before, in an attempt to keep these numbers realistic and significant, the energy of the transmitted one-hundred-femtosecond-long Gaussian pulses is limited to 0.1 aJ.

5.1 Information Rate After Coding

5.1.1 Information Rate with a Soft Receiver

The information rate $IR_{u-coded}$ with a soft receiver architecture, obtained by combining (23), (24) and (25) in (22), is shown in Figure 5(a) as a function of the transmission distance d, for different node densities λ_T in nodes/mm^2 and for different probabilities to transmit silence, $p_X(X=0)$ (21).

For very short transmission distances $d < 10$ mm, the received signal power when a pulse is transmitted, *i.e.*, a_1, is much higher than the molecular absorption noise N_1, N_0 and the multi-user interference I, and the information rate reaches its maximum value. For this region, the maximum information rate is obtained when the coding weight is 0.5, as in the classical symmetric channel. As the transmission distance is increased, the received signal power when a pulse is transmitted a_1 tends to 0, but the noise power N_1 remains larger than the noise power N_0. Because the interference affects in the same way the transmission of "0"s and "1"s, a soft receiver will still be able to distinguish between symbols, and its equivocation can be minimized by choosing a lower coding weight. As the transmission distance is fur-

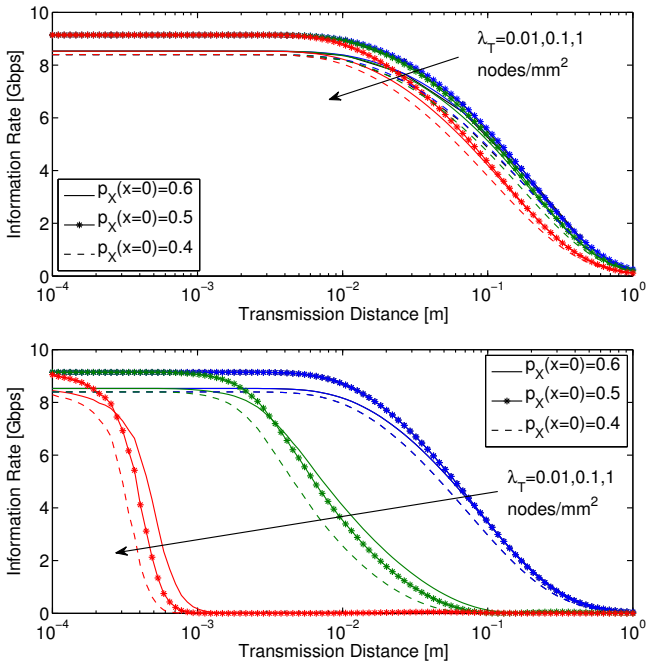


Figure 5: Information rate with a soft receiver architecture (top) and a hard receiver architecture (bottom) as functions of the transmission distance for different node densities and different probabilities of pulse transmission ($\beta = 1000$).

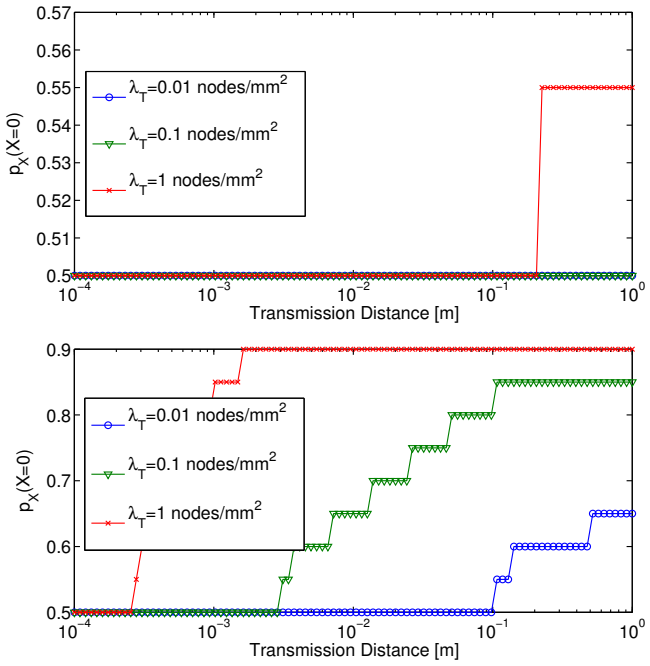


Figure 6: Optimal probability to transmit a logical "0", $p_X\,(X=0)$, for a soft receiver architecture (top) and a hard receiver architecture (bottom) ($\beta = 1000$).

ther increased, the noise power N_1 tends to N_0, and thus, symbols cannot be distinguished anymore.

5.1.2 *Information Rate with a Hard Receiver*

The information rate $IR_{u-coded}$ with a hard receiver architecture, obtained by combining (23) and (28) in (22), is shown in Figure 5(b) as a function of the transmission distance d, for different node densities λ_T in nodes/mm^2 and for different probabilities to transmit silence, $p_X\,(X=0)$ (21).

For a specific node density, we can also distinguish three main regions in the behavior of the information rate after coding. In this case, the impact of the coding weight is more noticeable because the multi-user interference I (18) drastically impacts the performance of the hard receiver. In particular, interference shifts the received signal and makes a fixed-threshold th_1, th_2 in (27) detection scheme not optimal (contrary to the far more complex soft receiver architecture). In this case, the reduction of the coding weight, *i.e.*, the increase of $p_X\,(X=0)$, turns into an improvement in the achievable information rate after coding. This result validates our original hypothesis and justifies this work. This effect can clearly be seen by analyzing the behavior of the codeword error rate (Section 5.2).

5.1.3 *Optimal Coding Weight*

In Figure 6, the optimal probability to transmit a logical "0", $p_X\,(X=0)$, that maximizes the information rate is shown for both a soft receiver architecture and a hard receiver architecture, as a function of the distance d.

Following the same discussion as for the information rate after coding, we can also distinguish three main regions. For very short transmission distances, $d < 1$ mm, the optimal coding weight corresponds to the equiprobable source distribution $p_X\,(X=0) = p_X\,(X=1) = 0.5$. As the transmission distance is increased, the optimal coding weight increases the transmission of logical "0"s, *i.e.*, silence, and $p_X\,(X=0) > 0.5$. This is specially visible for the hard-receiver architecture.

5.2 Codeword Error Rate

In Figure 7, the CER (30) is shown as a function of the transmission distance d, for different node densities λ_T in nodes/mm^2 and for different probabilities to transmit a logical "0", $p_X\,(X=0)$ (21). For a specific node density, the reduction of the coding weight turns into reduced error rates. This is specially valid for high interference scenarios, *i.e.*, high node densities in our analysis. For distances in the order of a few millimeters, which is expectedly the transmission range of nano-devices, the use of low-weight channel codes can clearly improve the CER. In particular, for very high node-densities, the improvement can be up to 150%.

6. CONCLUSIONS

In this paper, we propose the utilization of low-weight channel codes to prevent channel errors in pulse-based electromagnetic nanonetworks in the Terahertz Band. For this, we first stochastically model the two main error sources in electromagnetic nanonetworks, namely, the molecular absorption noise and the multi-user interference. Then, we propose the control of the code weight as a mechanism to reduce the noise and interference power. Finally, we investigate the performance of low-weight codes in terms of information rate and codeword error rate. The results show how, by using low-weight channel codes, the overall noise and in-

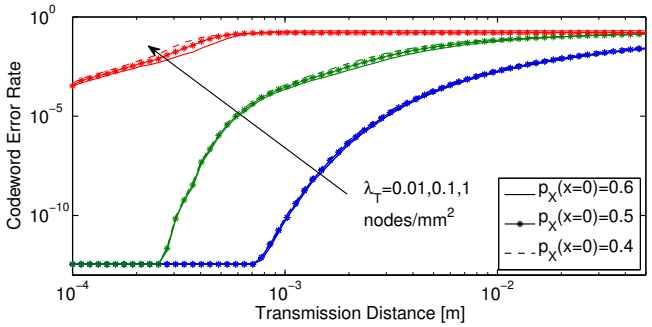


Figure 7: CER as a function of the transmission distance for different node densities and different probabilities of pulse transmission ($T_s/T_p = 1000$, $\lambda_T = 0.01, 0.1, 1$ nodes/mm^2).

terference can be reduced while keeping constant or even increasing the achievable information rate. Moreover, we show that there is an optimal code weight, which depends on the channel and network conditions, for which the information rate is maximized. These results motivate the development of novel link policies that can dynamically adapt the code weight to the channel and network conditions.

7. REFERENCES

[1] I. F. Akyildiz and J. M. Jornet. Electromagnetic wireless nanosensor networks. *Nano Communication Networks (Elsevier) Journal*, 1(1):3–19, Mar. 2010.

[2] J. Andrews, R. Ganti, M. Haenggi, N. Jindal, and S. Weber. A primer on spatial modeling and analysis in wireless networks. *Communications Magazine, IEEE*, 48(11):156 –163, Nov. 2010.

[3] F. Box. Utilization of atmospheric transmission losses for interference-resistant communications. *IEEE Transactions on Communications*, 34(10):1009–1015, Oct. 1986.

[4] P. Cardieri. Modeling interference in wireless ad hoc networks. *IEEE Communications Surveys and Tutorials*, 12(4):551–572, 2010.

[5] COMSOL Multiphysics Simulation Software. COMSOL.

[6] A. K. Geim and K. S. Novoselov. The rise of graphene. *Nature Materials*, 6(3):183–191, Mar. 2007.

[7] V. P. Gusynin, S. G. Sharapov, and J. P. Carbotte. AC conductivity of graphene: from tight-binding model to 2+1-dimensional quantum electrodynamics. *International Journal of Modern Physics B*, 21:4611, 2007.

[8] J. M. Jornet and I. F. Akyildiz. Channel modeling and capacity analysis of electromagnetic wireless nanonetworks in the terahertz band. *IEEE Transactions on Wireless Communications*, 10(10):3211–3221, Oct. 2011.

[9] J. M. Jornet and I. F. Akyildiz. Low-weight channel coding for interference mitigation in electromagnetic nanonetworks in the terahertz band. In *Proc. of IEEE International Conference on Communications (ICC)*, June 2011.

[10] J. M. Jornet and I. F. Akyildiz. Graphene-based plasmonic nano-antenna for terahertz band communication in nanonetworks. *IEEE JSAC, Special Issue on Emerging Technologies for Communications*, 12(12):685–694, Dec. 2013.

[11] J. M. Jornet and I. F. Akyildiz. Femtosecond-long pulse-based modulation for terahertz band communication in nanonetworks. *IEEE Transactions on Communications*, 62(5):1742 – 1754, May 2014.

[12] J. M. Jornet and I. F. Akyildiz. Graphene-based plasmonic nano-transceiver for terahertz band communication. In *Proc. of European Conference on Antennas and Propagation (EuCAP)*, 2014.

[13] M. Kocaoglu and O. B. Akan. Minimum energy channel codes for nanoscale wireless communications. *IEEE Transactions on Wireless Communications*, 12(4):1492–1500, 2013.

[14] A. N. Pal and A. Ghosh. Ultralow noise field-effect transistor from multilayer graphene. *Applied Physics Letters*, 95(8), 2009.

[15] S. Priebe, C. Jastrow, M. Jacob, T. Kleine-Ostmann, T. Schrader, and T. Kurner. Channel and propagation measurements at 300 GHz. *IEEE Transactions on Antennas and Propagation*, 59(5):1688 –1698, May 2011.

[16] E. Ratzer and D. MacKay. Sparse low-density parity-check codes for channels with cross-talk. In *Proc. of IEEE Information Theory Workshop*, pages 127–130, Mar. 2003.

[17] V. Ryzhii, M. Ryzhii, V. Mitin, and T. Otsuji. Toward the creation of terahertz graphene injection laser. *Journal of Applied Physics*, 110(9):094503, 2011.

[18] E. Sousa. Performance of a spread spectrum packet radio network link in a poisson field of interferers. *IEEE Transactions on Information Theory*, 38(6):1743 –1754, Nov. 1992.

[19] J. Tabor. Noise reduction using low weight and constant weight coding techniques. Technical report, MIT, Cambridge, MA, USA, 1990.

[20] M. Tamagnone, J. S. Gomez-Diaz, J. R. Mosig, and J. Perruisseau-Carrier. Reconfigurable terahertz plasmonic antenna concept using a graphene stack. *Applied Physics Letters*, 101(21):214102, 2012.

[21] Z. L. Wang. Towards self-powered nanosystems: from nanogenerators to nanopiezotronics. *Advanced Functional Materials*, 18(22):3553–3567, 2008.

[22] M. Win, P. Pinto, and L. Shepp. A mathematical theory of network interference and its applications. *Proceedings of the IEEE*, 97(2):205 –230, Feb. 2009.

[23] M. Win and R. Scholtz. Ultra-wide bandwidth time-hopping spread-spectrum impulse radio for wireless multiple-access communications. *IEEE Transactions on Communications*, 48(4):679–689, 2000.