Impulsive noise cancellation in multicarrier transmission

[1]Fatma Abdelkefi, [2]Pierre Duhamel, [2]Florence Alberge

[1] ENSEA, 6 avenue du Ponceau, 95014 Cergy-Pontoise Cedex France . [2] Supelec/LSS, Plateau de Moulon, 91192 GIF-SUR-YVETTE Cedex. fax:[1] `(33) 1 30 73 66 27` [2] `(33) 1 69 85 17 65` *E-mail:* [1] *abdelkefi@ensea.fr,* [2] *pierre.duhamel, alberge@lss.supelec.fr*

Abstract

A parallel between Reed-Solomon codes in the complex field and multicarrier transmission using OFDM is first presented. This shows that when the signal is sent over some channel composed of Gaussian plus impulse noise, the impulse noise can be removed by a procedure similar to channel decoding, using information carried by the "syndrome". This result is first derived in a simple situation (oversampled DMT, additive channel), which is merely of theoretical interest. In any case consecutive zeroes, in the output of the OFDM modulator, do not correspond to real subcarriers. Pilot tones are emitted for synchronization or channel estimation purposes. These pilot tones are generally scattered among the information ones. Our approach is to use these pilot tones as syndromes, in order to correct impulse noise. We show that the correction capacity is conditioned by the position of these pilot tones in the emitted sequence. A protection subsystem based on hypotheses test is introduced after the decoding operation in order to detect malfunction of this decoder. Efficiency of this technique is corroborated with simulations in the slightly modified Hiperlan2 context. Other extensions are then provided in order to increase the practical usefulness of the method.

I. Introduction

Multicarrier or Orthogonal Frequency Division Multiplexing (OFDM) system is a method of data modulation that has gained recently an increased interest with the development of faster signal processing components and technologies [1]. It is used in the European digital audio broadcast (DAB) [2] and in wireless environment such as digital broadcast television and mobile communication system [3],[2]. However, OFDM based on discrete multitone (DMT) systems are also examined for broadband digital communications on the existing copper networks [4]. This technique has been proposed for high-rate and assymetric digital subcarrier lines (HDSL, ADSL) [4], Local Area Network (Hiperlan2)....

The main idea behind OFDM is to split the transmitted data sequence into N parallel sequences of symbols. This structure has the particularity to enable a simple equalization scheme and to resist to multipath propagation channel. In fact, intersymbol interference (ISI) can be avoided when a guard interval (IG) is implemented between each block of time domain samples to be transmitted. However, some carriers can be strongly attenuated. It is then necessary to incorporate a powerful channel encoder combined with frequency and time interleaving. In this way, close coded bits are not likely to fall simultaneously in a spectral null [2].

However some of these quoted applications suffer from impulse noise and then the performance of such system are damaged. The impulse noise is an additive disturbance that arises primarily from the switching electric equipment [5], [6], [7] and as spectral properties, the defined pulse has a pole and infinite energy. Therefore, bursty or isolated errors are usually generated by an impulse noise affecting consecutive symbols in the Viterbi decoder, because such decoder relies on the past history of the symbol sequence. Thus a powerful decoder is required for such applications that is robust against impulse noise in order to minimize its impact. The impulse noise model that is used in this study is that given by Ghosh [8], it will be described later in the paper.

In order to implement a digital modulator, an oversampled version of the continuous signal is often computed. This amounts to appending consecutive null symbols to the block of symbols to be modulated. Therefore the OFDM modulator can be seen as a real Reed-Solomon encoder [9] and then it can be used as some specific impulse noise canceler, the structure of which is well suited to the nature of the problem (i.e, a single impulse shows up as a single error), rather than counting on the classical channel coder to solve the problem. Practically, both types of codes will have to cooperate, in order to process both Gaussian and impulse noise.

To cancel impulse noise, Wolf [10], Redinbo [11], Ja-Ling Wu [12] and Kumaresan [13] had also used the BCH code in the complex or real field and they have considered the effect of minor errors.

Wolf [10] suggested two methods to correct large errors. The first is based on the Fourier transform coding and is a voting scheme. This technique takes the DFT of the received sequence and examines those frequency components that should be zeros for the original data sequence, that means: any k samples of the N received samples can be used to estimate the emitted samples. So there are C_N^k possibilities. If there are small independent errors on each of the transmitted components, the vote of the "correct vector" yields a cluster of vector rather than a single vector. This technique is impractical for large values of N. The second method is a slightly modified BCH decoding algorithm and it is not effective for multiple errors.

Marvasti shown in [14] that the problem of signal reconstruction from missing samples can be resolved by using a reconstruction algorithm similar to Reed-Solomon decoding technique based on the Fourier transform. He proposed an error recovery technique for bursts (BERT) of real and complex samples that uses techniques similar to Peterson's method for BCH decoding and Forney's decoding. The BERT technique is found to be sensitive to background noise.

Wu [12], [15] defined a class of real-number linear block codes using the discrete cosine transform (DCT). Despite the non-cyclic nature of the DCT codes, a set of modified syndromes can be defined with which a modified BCH decoding algorithm can be performed. He supposed that the codeword is corrupted by a minor error vector due to the background noise and by impulse noise due to the channel noise. To correct impulse noise, he used a modified Berlekamp-Massey algorithm and the Forney algorithm which include a decision threshold. However, for our case, DFT is preferred over to DCT due to its cyclic properties.

Kumaresan [13] used Reed-Solomon codes, in Real or Complex field to correct bursty impulse noise in the presence of minor errors in each components of the received vector. He devised several decoding strategies based on least squares techniques and singular values decomposition to estimate the localization and the number of impulse errors.

However, Kumaresan, Wolf and Marvasti haven't suggested any method to control the malfunction of their proposed decoding algorithm. Redinbo presented in [11], [16] a decoding procedure for real-number codes which are also constructed by imposing constraints in the discrete Fourier transform domain: consecutive zeroes. He supposed that codewords are corrupted by small levels of roundoff noise and occasionally by a few large noise "impulse noise". The error-correcting algorithm is divided in two parts. The first is the large activity detection, it determines if large excursions are present and estimates their locations. The second part is the large errors values estimation. To estimate the impulse error locations, he used a modified Berlekamp-Massey algorithm. The final stage of this algorithm consists in testing the corrected outputs by recomputing syndromes and employing a threshold detector. Redinbo differs from the others by this protection subsystem but the proposed threshold is only tuned by simulation.

These quoted references consider that the emitted sequences contains some consecutive zeros but this assumption do not correspond to a practical case. In many cases, however, pilot tones are emitted and are scattered among the information ones. We believe that no study have been done in order to correct impulse noise in this context (non-consecutive zeros or pilots) and by using the properties of BCH codes in the complex or in the real field.

To correct impulse noise in this case, the correction capacity should be defined. However, Hartmann and Tzeng observed that there exist many cyclic codes whose defining set of zeros contains more than one set of consecutive zeros [17] and they succeeded in improving the BCH bound [18]. They extended the BCH bound to that case but they didn't give a general solution. In [9], we have used a special case of Hartmann-Tzeng bound, by considering that the output of

OFDM modulator contains $(2t)$ uniformly spaced pilot tones, the spacing being co-prime with the length of the emitted sequence. However, some additional flexibility would be very useful in many applications.

To correct impulse noise in the OFDM system, we have suggested to use the similarity between Reed-Solomon code in the complex field and the OFDM system and then the properties of Reed-Solomon code in the complex field are easily applied. Note that the proposed decoding algorithm uses techniques that are practically similar to those used by Wolf [10], Redinbo [11], Wu [12], Maravasti [14] and Kumaresan [13] but we applied these techniques in the general case where the pilot tones are neither consecutive nor uniformly distributed.

This paper first states the conditions on the locations of the pilot tones so that they can be seen as additional syndromes to correct impulse noise. This condition is first stated in terms of the rank of some specific matrix. So, we considered in this part the general case where Hartmann bound is a particular case.

In the second part of this paper, to detect malfunctions of the decoding algorithm, we propose an *a posteriori* control test which is based essentially on the hypotheses tests. So the threshold that we use is deduced from the receiver operating characteristic curve (ROC) and Bayes criteria.

Implementing a digital modulator require working with an oversampling version of the emitted analog signal and this is equivalent to add consecutive zeros to the block of symbols to be modulated. If the receiver has the same structure, null symbols should be received at the same locations. We show in section II the similarity between Reed-Solomon (RS) and OFDM modulator. Pilot tones can also be used as additional syndromes, in section III we establish the conditions on the locations of the pilot tones for the correction capacity to be maximal. In paragraph IV, we explain the decoding algorithm. To detect the malfunction of this decoding algorithm, we present in paragraph V the *a posteriori* control test. Finally, the efficiency of our technique is proved in a slightly modified Hiperlan2 context.

II. Transmission scheme and connection with spectral codes

A. Transmission scheme

A discrete model of the OFDM system is easily obtained by computing M samples of the signal to be sent onto the channel during one OFDM symbol, i.e., $MT_e = NT_s$, $T_e \leq T_s$ (T_e the sampling period and T_s is the OFDM symbol period). Moreover, if one considers the simple multicarrier system where the prototype filter is a rectangular pulse of duration NT_s, modulated with spacing between carriers equal to $1/NT_s$, these samples are computed as:

$$c_k(n) = \sum_{m=0}^{N-1} E_m(n-1)e^{\frac{2j\pi km}{M}}$$

which is exactly the inverse discrete Fourier Transform (IDFT) of the emitted sequence $\{E_m(n-1)\}$ enlarged by $(M-N)$ zeroes.

At the receiver the Analog to Digital Converter (ADC) samples the signal $r(n)$, at rate T_e and a DFT is performed. Therefore, the received signal is converted into the frequency domain $\{Y_k\}$,

$$Y_k = E_k + N_k \ , 0 \leq k \leq M-1$$

where N_k is the Fourier transform of the noise sequence $\{n_k\}$ (see figure1).

This section considers channels without ISI, for simplicity. Extension to the case where the channel introduces ISI is fairly trivial by using a cyclic prefix, as is classically done in OFDM systems.

B. Channel model

Assuming a memoryless channel, each emitted sample is modified by the channel according to

$$r_k = c_k + b_k + i_k, \ \ k \in \{0 \dots M-1\}$$

where b_k is an additive white Gaussian noise (AWGN) with zero mean and variance σ_b^2 and i_k is the impulse noise. In the following, the impulse noise is modeled as in [8] as:

$$i_k = l_k g_k \ \forall k \in \{0 \dots M-1\}$$

where l_k stands for a Bernoulli process, an i.i.d. sequence of zeroes and ones with $prob(l_k = 1) = p$, and g_k is a complex white Gaussian noise with zero mean and variance σ_i^2 such as $\sigma_i^2 \gg \sigma_b^2$. Note that this model assumes the presence of a large interleaver, so that bursts of errors can be scattered along time.

Under this model, the probability density function of the channel noise $n_k = b_k + i_k$ can be expressed as:

$$p(x) = (1-p)G(x, 0, \sigma_b^2) + pG(x, 0, (\sigma_i^2 + \sigma_b^2))$$

where $G(n, m_x, \sigma_x) = \frac{1}{\sigma_x\sqrt{2\pi}} exp(-\frac{(x-m_x)^2}{2\sigma_x^2})$ (i.e., the Gaussian density with mean m_x and variance σ_x^2).

This expression allows to compute the capacity of this channel, in order to estimate the impact of a given impulse noise on the capacity of a Gaussian channel. An efficient numerical technique to calculate this capacity has been derived in 1972 by Blahut and Arimoto [19], [20]. They proposed an iterative procedure which has the property of monotonic convergence to the capacity and which is applicable to arbitrary discrete memoryless channels. This method has been applied to compute the capacity of the " Gaussian plus Bernoulli Gaussian" channel in the case of real field. According to the memoryless channel model defined above, the received signal has the following expression: $r_k = c_k + n_k$, where $\{n_k\}$ stands for the complex noise of the form $n_k = n_{c_k} + jn_{s_k}$, $j = \sqrt{-1}$ and $\{c_k\}$ is the emitted sequence of the form $c_k = c_{c_k} + jc_{s_k}$. The subscripts "c" and "s" suggest the real and imaginary parts. The expectation of a real random variable is naturally generalized to the complex case as $E(X) = E[X_c] + jE[X_s]$ ($X = X_c + j\ X_s$ where X is a random variable) [21]. The statistic properties of $X = X_c + jX_s$ are determined by the joint probability density function $p_{X_cX_s}(x_c, x_s)$ of X_c and X_s, provided of course that the PDF exists: $p_X(x_c + jx_s) = p_{X_cX_s}(x_c, x_s)$. We suppose that $\underline{n}$ and $\underline{c}$ are independent.

$\{n_{c_k}\}$ and $\{n_{s_k}\}$ (respectively $\{c_{c_k}\}$ and $\{c_{s_k}\}$) have the same autocorrelation function, a vanishing crosscorrelation function, with zero mean and $E[|c_{c_k}|^2] = E[|c_{s_k}|^2]$.

The mutual information between the channel input $\underline{c}$ and the channel output $\underline{r}$ can be written as a function of the entropy $H(\underline{r})$ and the conditional entropy $H(\underline{r}|\underline{c})$, that means:

$$I(\underline{c}\ ;\ \underline{r}) = H(\underline{r}) - H(\underline{r}|\underline{c}) \tag{1}$$

where $I(\ .\ ;\ .\ )$ stands for the mutual information, H is the entropy and $\underline{\cdot}$ denotes vector.

However, one can easily verify that [22]:

$$H(\underline{r}|\underline{c}) = H(\underline{n}) \ \text{ and } \ H(\underline{n}) = H(\underline{n_c}) + H(\underline{n_s}) \tag{2}$$

We suppose in the following that real and imaginary part are independent, thus:

$$H(\underline{r}) = H(\underline{r_c}) + H(\underline{r_s}) \tag{3}$$

From equation (1), (2) and (3) we deduce that:

$$I(\underline{c}\ ;\ \underline{r}) = 2\ I(\underline{c_c}) \tag{4}$$

Thus $C^{2D} = 2C^{1D}$ where C^{1D} and C^{2D} are respectively the capacities of real channel (which is calculated above) and complex channel. Figure 2 depicts the capacity of the "Gaussian plus

Bernoulli Gaussian" channel in bits per second normalized by the bandwidth of the channel (W), as a function of the signal power P for several values of p, where $\sigma_i = 1$ and $\sigma_b = 6.10^{-2}$. We note that, even for somewhat large values of p, the capacity of the channel is approximately similar to that of the AWGN channel. For example, if $p = 10^{-2}$, and $P = 1$, the capacity of the "Gaussian plus Bernoulli Gaussian" channel is $8bit/s/Hz$, which is approximately the same value as for the AWGN channel. If $p = 5.10^{-2}$ then we can transmit 7.4 $bits/s/Hz$ that means that we lost only 0.6 bit per $second/Hz$, this decrease of capacity being due to the impulse noise.

However, if no specific procedure is used in an OFDM system, it is unlikely that such similar performance can be obtained: consider the case of a $64QAM$ constellation emitted over 64 subbands. Each impulse drastically impairs 384 bits at a time and it can be stated that the OFDM demodulator acts as an impulse noise amplifier. This is clearly in favor of a processing taking into account the specific nature of the impulse noise in the OFDM system.

C. Connection between OFDM system and Reed-Solomon (RS) code

It has been shown in [23], [24] and [25], that the ideas of spectral coding theory can be translated in the frequency domain, over a field $\mathbb{F}$ such that $\mathbb{C}$, $\mathbb{R}$,.... Reed-Solomon codes can be defined as follows [24]:

Definition 1: Let $\mathbb{F}$ contains an element of order M. The $(M, M-2t)$ RS block length M with symbols in $\mathbb{F}$ is the set of all vectors $\underline{c}$ whose spectrum in $\mathbb{F}$ satisfies: $C_k = 0$, $\forall k \in \mathcal{A}$ where $\mathcal{A} = \{k_0 + 1 \dots k_0 + 2t\}$. This is described briefly as an $(M, M-2t)$ RS code over $\mathbb{F}$.

The spectrum of a RS code word lives in the same field as information symbols. Then, to form a RS code, a block of $2t$ consecutive spectral components are chosen as parity frequencies (to be set to zero) and the remaining are information symbols. Marshall [26] has shown that a conventional decoding algorithm for finite field cyclic codes could be employed for real and complex numbers. The basic remark that we have used in [9] is that a discrete sequence of complex numbers containing $2t$ consecutive zeroes are transmitted over the OFDM system, therefore, the output of the OFDM modulator can be considered as a RS code word. After transmission over a "Gaussian plus Bernoulli Gaussian" channel, the DFT of the received discrete time sequence no longer has $2t$ zeroes, and this is due only to the channel. Hence, the OFDM modulator can be seen as a complex-valued RS code and the correction capacity is given by the following BCH bound property:

BCH Bound 1: If ($2t$ consecutive frequencies belong to $\mathcal{A}$) then (the minimal distance is at least $2t-1$), where $\mathcal{A}$ is the set of the $2t$ zeroes.

The BCH bound proves that t errors in any codeword of a RS code can always be corrected, because every pair of codewords differs in at least $2t+1$ places. So the correction capacity is up bounded by $\lceil\frac{2t+1}{2}\rceil$.

However, strictly speaking, there are more than $\lceil\frac{2t+1}{2}\rceil$ errors if one uses our channel model : all samples are polluted by noise. Therefore, we concentrate on the removal of the sole impulse noise, considering the Gaussian component as background noise. The classical decoding techniques have to be adapted to the presence of this background noise.

However, consecutive zeros do not correspond to a part of the OFDM spectrum which is actually available (analog shaping filters limit bandwidth), and only a small part of these zeros can be practically used. In many cases pilot tones are transmitted for synchronization or channel estimation purpose. These pilot tones consist in known symbols that are scattered among the information ones. The next paragraph states the conditions on the locations of the pilot tones in order to correct impulse noise. We believe that this technique is new in the theory of impulse noise cancelation.

In the following, the corresponding received components of $\{Y_k\}$ will no longer be null (figure 1):

$$Y_k = E_k + B_k + I_k,\ \forall k \in \mathcal{A}$$

where I_k is the DFT of the impulse noise i_n, and B_k that of the background noise b_n. $\mathcal{A}$ is the set of the positions of pilot tones on the transmitted sequence.

Let $\beta = card\{\mathcal{A}\}$ and $\mathcal{A}(k)$ is the k^{th} element in $\mathcal{A}$, at the receiver, the correction of impulse noise must operates on the syndromes S_k which are given by:

$$\begin{aligned} S_k &= Y_{\mathcal{A}(k)} - E_{\mathcal{A}(k)},\ k \in \{1\ldots\beta\} \\ &= B_{\mathcal{A}(k)} + I_{\mathcal{A}(k)} \\ &= \sum_{n=0}^{M-1} b_n W_M^{n\mathcal{A}(k)} + \sum_{m=0}^{\nu-1} i_{f_m} W_M^{f_m\mathcal{A}(k)} \end{aligned} \quad (5)$$

where $W_M = exp(-j\frac{2\pi}{M})$, ν is the number of impulse noise in the channel and $\{f_k\}_{k\in\{0,\ldots,\nu-1\}}$ are the location of impulse noise in the sequence.

There are two contributions in these terms (eq.6): one is the Fourier transform of the Gaussian background noise, hence is still Gaussian, and the other one is a sum of Fourier transforms of impulses, hence is a sum of complex sinusoids, the frequencies of which correspond to the localization of the errors. The decoding problem is thus the estimation of the number of sinusoids, together with their frequencies and amplitudes, polluted by background noise. The two main differences with classical signal processing situations are: (i) the number of samples is orders of magnitude smaller than usual and (ii) one has the knowledge that the frequencies take integer values.

III. Impulse noise localization

In the following we assume that $\mathcal{A}$ is the set of the position of the pilot tones in the transmitted sequence and $\beta = card(\mathcal{A})$.

Classically we define a suppressing sequence $\{\lambda_k\}_{k\in\{0,\ \ldots,\ M-1\}}$ (M is the length of the emitted sequence) as follows:

$$\text{if } i_l \neq 0 \text{ then } \lambda_l = 0 \text{ that means:}$$

$$\lambda_l i_l = 0 \ \forall l \in \{0 \ldots M-1\} \tag{6}$$

where $\{i_l\}_{l\in\{0\ldots M-1\}}$ is the impulse noise sequence.

Let $\underline{\underline{\mathcal{F}_M}}$ denotes the Fourier matrix of size M. In the frequency domain, equation (6) reads:

$$\underline{\underline{I}}\,\underline{\Lambda} = \underline{0} \tag{7}$$

where $\underline{\underline{I}} = \underline{\underline{\mathcal{F}_M}}\,(diag(\underline{i}))\,\underline{\underline{\mathcal{F}_M^{-1}}}$ and $\underline{\Lambda} = \underline{\underline{\mathcal{F}_M}}\,\underline{\lambda}$.

So the key equation is equivalent to the previous equation (7) and $rank(\ \underline{\underline{I}}\ ) = rank(\ diag(\underline{i})\ )$, where $diag(\underline{x})$ denote diagonal matrix that contains on the diagonal the components of a vector $\underline{x}$.

A. Solving for the impulse locations

Matrix $\underline{\underline{I}}$ contains known components whose indices are in $\mathcal{A}$ and the others are not. So the main idea is to group together the maximum of the known components of $\underline{\underline{I}}$ whose indices are in $\mathcal{A}$, in a submatrix $\underline{\underline{I}}^{(r)}$ such that:

$$\underline{\underline{I}}^{(r)} = \underline{\underline{H}}\ \underline{\underline{I}}\ \underline{\underline{D}} \tag{8}$$

where $\underline{\underline{H}}$ and $\underline{\underline{D}}$ are selection matrices that depend on $\mathcal{A}$. Note that when the syndromes are consecutive or regularly distributed it is always possible to find a submatrix containing all the syndromes (the indices of components that belongs to $\mathcal{A}$). However, in the general case (randomly distributed syndromes) the task is much more difficult. Thus size of the matrix $\underline{\underline{I}}^{(r)}$ puts a limit on the correction capacity. Let the size of $\underline{\underline{H}}$ be $((r+s)\times M)$ and that of $\underline{\underline{D}}$ be $(M\times(r+1))$ where r and s are positive integers. Therefore, we have rank$(\underline{\underline{I}}^{(r)})\leqslant r+1$ and then we can correct at most r errors. So we suppose that we have at most r impulse noise (since we can not correct more than r impulse noise), then there are $M-r$ degrees of freedom on $\{\lambda_k\}$ and this is also equivalent to say that there are $M-r$ degrees of freedom on $\{\Lambda_k\}$ (because $\underline{\Lambda}=\underline{\underline{\mathcal{F}_M}}\,\underline{\lambda}$). As the size of the selected matrix $\underline{\underline{I}}^{(r)}$ is $((r+s)\times(r+1))$ and the degrees of freedom on $\{\Lambda_k\}$ is $M-r$, then it is possible to choose $M-r-1$ values of $\{\Lambda_k\}$ equals to zero (note that this is equivalent to select in $\underline{\underline{I}}$ the components that will be stocked in $\underline{\underline{I}}^{(r)}$) and one value equal to one. This is equivalent to multiply $\underline{\Lambda}$ by $\underline{\underline{D}}^T$. Let: $\underline{\Lambda}^{(r)} = (\Lambda_0^r\ ,\ \ldots\ ,\ \Lambda_r^r)^t = \underline{\underline{D}}^t\underline{\Lambda}$ and $\Lambda_0^{(r)}=1$, then the key equation (7) is reduced to:

$$\underline{\underline{I}}^{(r)}\,\underline{\Lambda}^{(r)}=\underline{0} \tag{9}$$

If rank$(\underline{\underline{I}}^{(r)}) = r$, then the key equation (9) has a single solution. Therefore, an important question for being able to solve (9) is: under which conditions is this system full rank ?

If $\underline{\underline{I}}^{(r)}$ has a full row rank then the correction capacity is r. By construction, $\underline{\underline{I}}^{(r)}$ has the following structure:

$$\begin{pmatrix} I_{m_0+\theta_0+\delta_0} & I_{m_0+\theta_0+\delta_1} & \cdots & I_{m_0+\theta_0+\delta_{r-1}} \\ I_{m_0+\theta_1+\delta_0} & I_{m_0+\theta_1+\delta_1} & \cdots & I_{m_0+\theta_1+\delta_{r-1}} \\ \vdots & \vdots & \ddots & \vdots \\ I_{m_0+\theta_{r-1+s}+\delta_0} & I_{m_0+\theta_{r-1+s}+\delta_1} & \cdots & I_{m_0+\theta_{r-1+s}+\delta_{r-1}} \end{pmatrix} \tag{10}$$

where m_0, $(\theta_k)_{0\le k\le r+s-1}$ and $(\delta_k)_{0\le k\le r-1}$ are integers such that $0\le\theta_k\le M-1$ and $0\le\delta_k\le M-1$ and that depend on the syndrome location in the sequence.

It has be seen in [27] that the submatrix $\underline{\underline{I}}^{(r)}$ constructed with the r first columns and rows of the matrix $\underline{\underline{I}}_r^{(r)}$ can be decomposed as: $\underline{\underline{I}}_r^{(r)} = \underline{\underline{\mathcal{Q}}}^{(r)}\,\underline{\underline{\mathcal{P}}}^{(r)}\,\underline{\underline{\mathcal{R}}}^{(r)}$, where:

$$\underline{\underline{\mathcal{Q}}}^{(r)} = \begin{pmatrix} W_M^{f_0\theta_0} & W_M^{f_1\theta_0} & \cdots & W_M^{f_{r-1}\theta_0} \\ W_M^{f_0\theta_1} & W_M^{f_1\theta_1} & \cdots & W_M^{f_{r-1}\theta_1} \\ \vdots & \vdots & \ddots & \vdots \\ W_M^{f_0\theta_{r-1}} & W_M^{f_1\theta_{r-1}} & \cdots & W_M^{f_{r-1}\theta_{r-1}} \end{pmatrix} \tag{11}$$

$$\underline{\underline{\mathcal{R}}}^{(r)} = \begin{pmatrix} W_M^{f_0\delta_0} & W_M^{f_0\delta_1} & \cdots & W_M^{f_0\delta_{r-1}} & \\ W_M^{f_1\delta_0} & W_M^{f_1\delta_1} & \cdots & W_M^{f_1\delta_{r-1}} & \\ \vdots & \vdots & \ddots & & \vdots \\ W_M^{f_{r-1}\delta_0} & W_M^{f_{r-1}\delta_1} & \cdots & W_M^{f_{r-1}\delta_{r-1}} & \end{pmatrix} \tag{12}$$

$$\underline{\underline{\mathcal{P}}}^{(r)} = \begin{pmatrix} i_{f_0} W_M^{m_0 f_0} & 0 & \cdots & 0 \\ 0 & i_{f_1} W_M^{m_0 f_1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & i_{f_{r-1}} W_M^{m_0 f_{r-1}} \end{pmatrix} \tag{13}$$

where $f_0 \ldots f_{r-1}$ are the locations of the impulses in the sequence such that $f_0 < f_1 \ldots < f_{r-1}$ and $W_M = exp(-\frac{2\pi j}{M})$.

If $\underline{\underline{\mathcal{R}}}^{(r)}$ and $\underline{\underline{\mathcal{Q}}}^{(r)}$ are full rank, then rank($\underline{\underline{\mathcal{Q}}}^{(r)}\ \underline{\underline{\mathcal{P}}}^{(r)}\ \underline{\underline{\mathcal{R}}}^{(r)}$) $= r$, thus the key equation (9) has a unique solution and the correction capacity is maximal and it is at most r.

Since $\underline{\underline{\mathcal{R}}}^{(r)}$ and $\underline{\underline{\mathcal{Q}}}^{(r)}$ have the same structure, we only study the conditions on θ_k that ensure that the matrix $\underline{\underline{\mathcal{R}}}^{(r)}$ is invertible for any error location. These conditions will apply similarly on $\underline{\underline{\mathcal{Q}}}^{(r)}$.

Note that Hartmann provides a solution to this problem:

Hartmann Theorem 1: Let $g(x) \in GF(q)[x]$ be the generator polynomial of a cyclic code, V_M of length M. If $g(\beta^{m_0+k\theta+l\delta}) = 0$ for $k = 0, 1, \ldots d_0 - 2$ and $l = 0, 1, \ldots s$ where $gcd(\theta, M) = 1$ and $gcd(\delta, M) = 1$ then $d_0 + s \leq d$, where d is the minimal distance.
This suggests that the consecutive $d_0 + s - 1$ spectral zeros of the BCH bound can be replaced by a pattern of $s + 1$ uniformly subblocks each of $d_0 + s - 1$ uniformly spaced spectral zeros.

However, we would like to search for more general conditions, leading to greatest flexibility for placing the pilot tones. We have developed some necessary conditions on the matrices $\underline{\underline{\mathcal{R}}}^{(r)}$ and $\underline{\underline{\mathcal{Q}}}^{(r)}$ to be invertible in the general case and we have not yet verified that these conditions are also sufficient. This condition is as follows:

Necessary condition 1: Let $\theta_k^{(r-1)} = \theta_{k+1}^{(r)} - \theta_{k_0}^{(r)} = \theta_{k+1} - \theta_{k_0} \quad \forall k \neq k_0$, where $k,\ k_0 \in \{0, \ldots, r-1\}$. So, $\underline{\underline{\mathcal{Q}}}^{(r)}$ is invertible only if, these two conditions are verified:
1. There are at least one $k_j \neq k_0$ such that: $gcd(\theta_{k_j}^{(r-1)}, M) = 1$,
2. There are at most $(r-2)$ values of $\theta_k^{(r-1)}$ that are not equal to $\theta_{k_j}^{(r-1)}$ and that verify: $gcd(\theta_k^{(r-1)}, M) \leqslant r - 1$.
For $r = 2$ and $r = 3$, we find a necessary and sufficient conditions (section below).

B. The locator polynomial evaluation

We have seen in the paragraph above that the correction capacity is maximum for a given value of r only if $\underline{\underline{\mathcal{R}}}^{(r)}$ and $\underline{\underline{\mathcal{Q}}}^{(r)}$ are full rank.

According to the key equation (7), we see that the impulse noise location are roots of the locator polynomial $\Lambda(x)$ whose coefficients are already calculated by the key equation (9), such that:

$$\Lambda(W_M^{-f_k}) = 0, \qquad \forall k \in \{0, 1, \ldots r-1\} \tag{14}$$

Since we propose that the number of impulse noise is r then $M-r-1$ values of $\{\Lambda_k\}$ are chosen zeros and this is due to the degrees of freedom that we have on $\{\Lambda_k\}$. However, the degree of $\Lambda(x)$ is generally much larger than r (it is equal to r in the contiguous case) since the pilot tones are irregularly spaced. Thus, it has to be checked that no other root of this polynomial will generate a false alarm (zero on the unit circle, at an integer location) which correspond to $card\{f_k|\Lambda(W_M^{-f_k}) = 0\} = r$.

To annul $(M-r-1)$ values of $\underline{\Lambda}$ (in order to cancel the unknown components of $\underline{\underline{I}}$) is equivalent to multiply $\underline{\Lambda}$ by the matrix $\underline{\underline{D}}$ (see equation 8) and then: $\underline{\Lambda} \;=\; \underline{\underline{D}}\,\underline{\Lambda}^r$. Thus, we have:

$$\underline{\lambda} = \underline{\underline{\mathcal{F}}}_M^{-1}\ \underline{\underline{D}}\ \underline{\Lambda}^{(r)} \tag{15}$$

So, we verify that:

$$\underline{\underline{\mathcal{F}}}_M^{-1}\ \underline{\underline{D}} = \begin{pmatrix} 1 & \cdots & 1 \\ W_M^{\delta_0} & \cdots & W_M^{\delta_r} \\ \vdots & \ddots & \vdots \\ W_M^{M\delta_0} & \cdots & W_M^{M\delta_r} \end{pmatrix} \tag{16}$$

However, if each (r+1) square submatrix of $\underline{\underline{\mathcal{F}}}_M^{-1}\ \underline{\underline{D}}$ is invertible, then $\underline{\lambda}$ has at most r zeros because $size(\underline{\Lambda}^{(r)}) = r+1$. It is easily seen that this problem is of the same nature as the one in the previous subsection, with a dimension augmented by 1. Thus $card\{f_k|\Lambda(W_M^{-f_k}) = 0\} = r$ only if $\underline{\underline{\mathcal{R}}}^{(r+1)}$ is invertible.

C. Examples

This general problem seems intricate, and we are still working on it. However when $r=2$ and $r=3$ we have found that the necessary conditions on $\{\theta_k\}$ and $\{\delta_k\}$ is also sufficient.

If $r = 2$, it can be shown that: $\underline{\underline{I}}^{(2)}$ is invertible if:

$$gcd(\theta_1 - \theta_0, M) = 1 \text{ and } gcd(\delta_1 - \delta_0, M) = 1 \tag{17}$$

Thus rank($\underline{\underline{I}}^{(2)}$)= 2, that means that we can correct at most two impulse noise.

If $r = 3$, define the function f as:

$$f(\theta_1 - \theta_0, \theta_2 - \theta_0) = |det(\underline{\underline{Q}}^{(r)})|, \quad \forall(\theta_0,\ \theta_1,\ \theta_2) \in \{0, 1, \ldots M - 1\}^3$$

It is easily seen that:

$$f(\theta_1 - \theta_0, \theta_2 - \theta_0) \ = \ f(\theta_2 - \theta_0,\ \theta_1 - \theta_0) \ = \ f(\theta_2 - \theta_0,\ \theta_2 - \theta_1)$$

As a consequence of these equalities, we prove that the correction capacity is at most 3 if and only if one of these following conditions is met:

$$C1 : \begin{cases} gcd(\theta_2 - \theta_0, M) = 1 \text{ and} \\ gcd(\theta_2 - \theta_1, M) = 1 \text{ and} \\ gcd(\theta_1 - \theta_0, M) = \gamma \end{cases} \qquad C2 : \begin{cases} gcd(\theta_1 - \theta_0, M) = 1 \text{ and} \\ gcd(\theta_2 - \theta_0, M) = 1 \text{ and} \\ gcd(\theta_2 - \theta_1, M) = \gamma \end{cases}$$

$$C3 : \begin{cases} gcd(\theta_2 - \theta_1, M) = 1 \text{ and} \\ gcd(\theta_1 - \theta_0, M) = 1 \text{ and} \\ gcd(\theta_2 - \theta_0, M) = \gamma \end{cases}$$

where $\gamma = 1$ if M is prime otherwise $\gamma = 2$. We have also the same conditions on $\{\delta_k\}$.

So if $\theta_2 - \theta_1 = \theta_1 - \theta_0$ and if $gcd(\theta_2 - \theta_1, M) = 1$ then the considered matrix is invertible, we verify that this particular case is the Hartmann bound. Thus, we can say that the condition that we find is more general than that of Hartmann.

These two cases ($r = 2$ and $r = 3$) and the general condition (which we have verified that is necessary) can be useful in many application such that in the practical context of Hiperlan2 where it is possible to cancel impulse noise [28], [29], [30], [?] and reduce the PAPR (Peak Average Power Rate) level [?], [?].

IV. Decoding algorithm

In the proposed procedure, we adapt a classical decoding algorithm to the presence of background noise.

The decoding algorithm works in three steps: (i) estimate the number ν of impulse errors, (ii) find the error locations and (iii) correct the errors. There are several efficient algorithms for doing this, but these are very sensitive to small levels of noise. Therefore, we have used a modified version of Peterson-Gorenstein-Zierler algorithm adapted to complex field [31], [32] and which is less sensitive to errors. Our work performed simultaneously in the context of joint source and channel coding [32]. In the following we have considered the general case, that means, we have supposed that we have in the emitted sequence some pilot tone which are not consecutive and are known (not zero).

A. *Estimation of* ν

According to section (III), the syndrome matrix is expressed as follows:

$$\underline{\underline{\mathcal{S}}} = \underline{\underline{\mathcal{S}}}^{(r)} = \begin{pmatrix} S_{m_0+\theta_0+\delta_0} & S_{m_0+\theta_0+\delta_1} & \cdots & S_{m_0+\theta_0+\delta_{r-1}} \\ S_{m_0+\theta_1+\delta_0} & S_{m_0+\theta_1+\delta_1} & \cdots & S_{m_0+\theta_1+\delta_{r-1}} \\ \vdots & \vdots & \ddots & \vdots \\ S_{m_0+\theta_{r+s-1}+\delta_0} & S_{m_0+\theta_{r+s-1}+\delta_1} & \cdots & S_{m_0+\theta_{r+s-1}+\delta_{r-1}} \end{pmatrix} = \underline{\underline{\mathcal{B}}}^{(r)} + \underline{\underline{I}}^{(r)}$$

where m_0, $(\theta_k)_{0\leq k\leq r+s-1}$ and $(\delta_k)_{0\leq k\leq r-1}$ have already been defined in the section (III) and we suppose that this syndrome matrix is full rank.

Let ν denotes the number of impulse noise such that $\nu < r$.

To remove the contribution of the background noise, we calculate the correlation matrix $\underline{\underline{\mathcal{S}}}^{(r)^H}\underline{\underline{\mathcal{S}}}^{(r)}$. Since $\underline{\underline{\mathcal{B}}}^{(r)}$ and $\underline{\underline{I}}^{(r)}$ are assumed to be uncorrelated with zero means: $E[\underline{\underline{B}}^{(r)^H}\ \underline{\underline{I}}^{(r)}] = \underline{\underline{0_r}}$ and since the background noise is supposed to be a white Gaussian noise with variance σ_b^2 and with zero mean then: $E[\underline{\underline{B}}^{(r)^H}\ \underline{\underline{B}}^{(r)}] = r\ \sigma_b^2\ \underline{\underline{Id_r}}$, where $\underline{\underline{Id_r}}$ is the identity matrix of dimension r.

For a great number of observation and as the amplitude of the impulse noise is more important (directly connected to the eigenvalues of the matrix $\underline{\underline{\Gamma}}^{(r)^H}\underline{\underline{\Gamma}}^{(r)}$) than that of the background noise, the number of errors is then estimated as the number of eigenvalues of the matrix $\underline{\underline{\mathcal{S}}}^{(r)^H}\underline{\underline{\mathcal{S}}}^{(r)}$ which amplitude is superior to $r\sigma_b^2$ [33], [**?**]. Regrettably, the relatively reduced number of observations does not allow to estimate correctly these errors; a multiplier empirical factor ϕ is applied. We find by simulation that $\phi \in [1\ 4]$ and then:

$$\underline{\underline{\mathcal{S}}}^{(r)^H}\ \underline{\underline{\mathcal{S}}}^{(r)} \approx \underline{\underline{I}}^{(r)^H}\ \underline{\underline{I}}^{(r)} + r\phi\sigma_b^2\underline{\underline{Id}}_r$$

This estimation is almost accurate, hence $\tilde{\nu}$ can be estimated as the number of eigenvalues of

$\underline{\underline{S}}^{(r)^H}\ \underline{\underline{S}}^{(r)}$ greater then $r\phi\sigma_b^2$.

B. Error localization

According to the key equation (9) and as we have $\tilde{\nu}$ impulse noise, then we have the following system:

$$\begin{bmatrix} I_{m_0+\theta_0+\delta_0} & \cdots & I_{m_0+\theta_0+\delta_{\tilde{\nu}-1}} \\ \vdots & \vdots & \vdots \\ I_{m_0+\theta_{r+s-1}+\delta_0} & \cdots & I_{m_0+\theta_{r+s-1}+\delta_{\tilde{\nu}-1}} \end{bmatrix} \underline{\Lambda}^{(\tilde{\nu})} = - \begin{bmatrix} I_{m_0+\theta_0+\delta_{\tilde{\nu}}} \\ \vdots \\ I_{m_0+\theta_{r+s-1}+\delta_{\tilde{\nu}}} \end{bmatrix}$$

$$\underline{\underline{I}}^{(\tilde{\nu})}\underline{\Lambda}^{(\tilde{\nu})} = \underline{I}^{(\tilde{\nu})}$$

Taking into account the statistical contribution of small level of noise, we get:

$$(\underline{\underline{I}}^{(\tilde{\nu})^H}\underline{\underline{I}}^{(\tilde{\nu})})^{-1} \approx (\underline{\underline{S}}^{(\tilde{\nu})^H}\underline{\underline{S}}^{(\tilde{\nu})} - \tilde{\nu}\sigma_b^2\underline{\underline{Id}}_{\tilde{\nu}})^{-1}$$

As $\underline{\underline{\Gamma}}^{(\tilde{\nu})}$ and $\underline{\underline{I}}^{(\tilde{\nu})}$ are uncorrelated, a good estimation of $\underline{\Lambda}^{(\tilde{\nu})}$ is:

$$\tilde{\underline{\Lambda}}^{(\tilde{\nu})} = (\underline{\underline{S}}^{(\tilde{\nu})^H}\underline{\underline{S}}^{(\tilde{\nu})} - \tilde{\nu}\sigma_b^2\underline{\underline{Id}}_{\tilde{\nu}})^{-1}\underline{\underline{S}}^{(\tilde{\nu})^H}\underline{S} \quad (18)$$

$$\approx (\underline{\underline{I}}^{(\tilde{\nu})^H}\underline{\underline{I}}^{(\tilde{\nu})})^{-1}\underline{\underline{I}}^{(\tilde{\nu})^H}\underline{I}^{(\tilde{\nu})} = \underline{\Lambda}^{(\tilde{\nu})} \quad (19)$$

So, the $\tilde{\nu}$ error locations are estimated as the $\tilde{\nu}$ indices belong $[0 \dots M-1]$ that gives the smallest values of $|\tilde{\Lambda}^{(\tilde{\nu})}(x)|$ taken on $W_M^{-[0\ \cdots\ M-1]}$.

C. Error amplitude

The expression of the syndrome is: $\underline{S} = \underline{\underline{V}}\ \underline{i} + \underline{\underline{W}}\ \underline{b}$, where the matrix $\underline{\underline{V}}$ depends on the impulse noise locations, $\underline{i}$ is the vector containing the corresponding amplitudes such that: $\underline{\underline{V}}(k,j) = W_M^{\mathcal{A}(k)j}$, $k \in \{1,\dots,\beta\}$ and $j \in \{0,\dots,\nu-1\}$ and $\underline{\underline{W}}(k,t) = W_M^{\mathcal{A}(k)t}$, $k \in \{1,\dots,\beta\}$ and $t \in \{0,\dots,M-1\}$.

An estimate $\tilde{\underline{i}}$ of $\underline{i}$ can be obtained by solving the system above in the least square sense. Then $\tilde{\underline{i}}$ is:

$$\tilde{\underline{i}} = (\tilde{\underline{\underline{V}}}^H\ \tilde{\underline{\underline{V}}})^{-1}\ \tilde{\underline{\underline{V}}}^H\ \underline{S}$$

V. The *a posteriori* control

In the last paragraph we have shown that the decoding algorithm is in three steps: (i) estimate the number ν of impulse noise, (ii) seek the error locations and (iii) correct the impulse errors.

We add a control step which is able to carefully check whether the decoding procedure has worked correctly. In this way, we are able to begin a truncated enumeration of all possible error locations (the most sensitive part of the algorithm) among the most likely ones. This truncated enumeration is necessary because of the presence of the background noise which introduces some fuzziness.

Malfunctions of the decoder can be due to:

- wrong estimation of impulse noise locations: the error locations are linked with the zeros of the localization polynomial $\Lambda(W_M)$ (indeed, the zeros of Λ are of the form $W_M^k = exp(\frac{-2j\pi k}{M})$). If $\Lambda(W_M^{-k}) = 0$ then k is an impulse noise location. But here the syndromes contain the Gaussian and Bernoulli noise thus $\Lambda(W_M^{-k})$ will not be zero and will take small magnitude. Therefore, the localization can be wrong as shown in figure 7, where we plot $\mid \Lambda(W_M^{-k}) \mid$ versus the position k with $k \in \{0 \dots M-1\}$, $M = 54$ and $\nu = 3$. We observe that at the position near 10, $\Lambda(W_M^{-k})$ is close to zero. As a consequence the decoding algorithm cannot detect the correct impulse noise locations.
- wrong estimate of the number of errors.
- possible overflow of error capacity.

Thus, a protection subsystem is introduced after the decoding operation in order to detect malfunctions at each step of the decoding algorithm. In section IV, we have seen that the expression of the syndrome vector is:

$$\underline{S} = \underline{\underline{V}}\ \underline{i} + \underline{\underline{W}}\ \underline{b}$$

where the matrix $\underline{\underline{V}}$ depends on the impulse noise locations, $\underline{i}$ is the vector containing the corresponding amplitudes.

The corrected outputs are tested by comparing the syndrome vector to its estimate $\underline{\tilde{S}} = \underline{\underline{\tilde{V}}}\ \underline{\tilde{i}}$ where $\underline{\tilde{i}} = (\underline{\underline{\tilde{V}}}\ \underline{\underline{\tilde{V}}}^H)^{-1}\ \underline{\underline{\tilde{V}}}^H\ \underline{S}$ and $\underline{\underline{\tilde{V}}}$ depends on the estimated locations of the impulse noise. We calculate $y = \|\underline{\tilde{S}} - \underline{S}\|^2$. If this amount exceeds a certain threshold then we try to correct malfunction, else we can conclude that we have properly corrected impulse noise.

The optimal value of this threshold is obtained thanks to hypotheses testing theory as explained below.

In the problem under consideration: either the localization of impulse noise is correct which means: $Im(\underline{\underline{V}}) \subset Im(\underline{\underline{\tilde{V}}})$, or we have a wrong localization of impulse noise that means $Im(\underline{\underline{V}}) \not\subset Im(\underline{\underline{\tilde{V}}})$. We denote by Im the image of the considered space.

1. **if Im($\underline{\underline{\mathbf{V}}}$) $\subset$ Im($\underline{\underline{\tilde{\mathbf{V}}}}$)**: then $\underline{\tilde{S}} - \underline{S} = \underline{\underline{P}}\ (\underline{\underline{V}}\ \underline{i}\ +\ \underline{\underline{W}}\ \underline{b})$, where $\underline{\underline{P}} = \underline{\underline{Id}}_{\beta} - \underline{\underline{\tilde{V}}}(\ \underline{\underline{\tilde{V}}}^H \underline{\underline{\tilde{V}}})^{-1}\ \underline{\underline{\tilde{V}}}^H$ is a projection matrix on $Ker(\underline{\underline{\tilde{V}}})$ and the rank of $\underline{\underline{P}}$ is $(\beta - \tilde{\nu})$, where $\tilde{\nu}$ is the estimated impulse noise number and $\beta = card(\mathcal{A})$. Since $Im(\underline{\underline{V}}) \subset Im(\ \underline{\underline{\tilde{V}}}\ )$, we verify that $\underline{\underline{P}}\ \underline{\underline{V}} = \underline{0}$. So we conclude that:

$$If \quad Im(\underline{\underline{V}}) \subset Im(\underline{\underline{\tilde{V}}}) \quad then \quad \|\underline{\tilde{S}} - \underline{S}\|^2 = \|\underline{\underline{P}}\ \underline{\underline{W}}\ \underline{b}\ \|^2$$

As the vector $\underline{b}$ is a Gaussian noise of variance σ_b^2 and zero mean, then $\underline{\underline{W}}\ \underline{b}$ is a Gaussian noise vector that contains $2M$ independents real Gaussian noise such as each one is of variance $\frac{\sigma_b^2}{2}$ and zero mean (2 is due to the fact that we consider the complex field $\mathbb{C}$). And as $\underline{\underline{P}}$ is a projection matrix, thus $\|\underline{\tilde{S}} - \underline{S}\|^2 = \|\underline{\underline{P}}\ \underline{\underline{W}}\ \underline{b}\ \|^2$ is a chi-square distribution: $\frac{\|\underline{\underline{P}}\ \underline{\underline{W}}\ \underline{b}\|^2}{\frac{\sigma_b^2}{2}} \sim \chi^2(2(\beta - \tilde{\nu}))$.

2. **if Im($\underline{\underline{\mathbf{V}}}$) $\not\subset$ Im($\underline{\underline{\tilde{\mathbf{V}}}}$)**: we can write $\underline{\tilde{S}} - \underline{S}$ as follows:

$$\underline{\tilde{S}} - \underline{S} = \underline{\underline{P}}\ [\underline{\underline{W}} \quad \underline{\underline{V}}] \begin{pmatrix} \underline{b} \\ \underline{i} \end{pmatrix}$$

The Gaussian noise and the Bernoulli one do not have the same variances, so we multiply the vector $\begin{pmatrix} \underline{b} \\ \underline{i} \end{pmatrix}$ by a diagonal matrix $\underline{\underline{O}}$ in order to normalize it. Then the vector $\underline{R} = \underline{\underline{O}} \begin{pmatrix} \underline{b} \\ \underline{i} \end{pmatrix}$ is a normal Gaussian noise, where $\underline{\underline{O}} = diag((\frac{1}{\sigma_b} \cdots \frac{1}{\sigma_b}, \frac{1}{\sigma_i} \cdots \frac{1}{\sigma_i}))$.

Let $\underline{\underline{Q}} = \underline{\underline{P}}\ [\underline{\underline{W}} \quad \underline{\underline{V}}]\ \underline{\underline{D}}^{-1}$, then $\|\underline{\tilde{S}} - \underline{S}\|^2 = \underline{R}^H\ \underline{\underline{Q}}^H\ \underline{\underline{Q}}\ \underline{R}$.

$\underline{\underline{Q}}^H \underline{\underline{Q}}$ is a positive definite, Hermitian matrix, therefore it is diagonalizable, i.e. there exists a unitary matrix $\underline{\underline{K}}$ such that: $\underline{\underline{Q}}^H \underline{\underline{Q}} = \underline{\underline{K}}^H\ \underline{\underline{G}}\ \underline{\underline{K}}$, where $\underline{\underline{G}}$ is a diagonal matrix that contains the eigenvalue of the Hermitian matrix. Then, $y = \|\underline{\tilde{S}} - \underline{S}\|^2 = (\underline{\underline{K}}\ \underline{R})^H\ \underline{\underline{G}}\ \underline{\underline{K}}\ \underline{R}$. As $\underline{R}$ is a normal Gaussian vector, then $\underline{Z} = \underline{\underline{K}}\ \underline{R}$ is also a normal Gaussian one. Thus:

$$\|\underline{\tilde{S}} - \underline{S}\|^2 = \sum_{k=1}^{rank(\underline{\underline{Q}})} (g_k Z_k)^2 \tag{20}$$

where (g_k) are the eigenvalues of the matrix $\underline{\underline{Q}}^H \underline{\underline{Q}}$. We can easily verify that $rank(\underline{\underline{Q}}) = rank(\underline{\underline{P}}) = \beta - \tilde{\nu}$. Thus $\|\underline{\tilde{S}} - \underline{S}\|^2$ is a linear combination of chi-square distributions of the random variables (Z_k), such that: $\frac{(g_k Z_k)^2}{\frac{g_k^2}{2}} \sim \chi^2(2)$.

So, we can deduce the probability density function (pdf) of $y = \|\underline{\tilde{S}} - \underline{S}\|^2$. As g_k depends on the unknown impulse noise locations, then we calculate the conditional probability distribution functions: $p_Y(y|\underline{loc})$ for each value of the location of the impulse noise, where $\underline{loc}$ is the vector containing the locations of the impulse noise and we compute the average with respect to the locations. So Bayes formula leads to this equality:

$$p_Y(y) = \sum_{\underline{loc}} p_Y(y|\underline{loc})p(\underline{Loc} = \underline{loc}) \tag{21}$$

Generally the analytic expression of $p_Y(y)$ is unknown but it is always possible to calculate it when the correction capacity is definite.

At this point, we have obtained the two pdf corresponding to the two situations under study (i.e., the error localization are all correct ($\frac{\|\tilde{\underline{S}}-\underline{S}\|^2}{\frac{\sigma_b^2}{2}} \sim \chi^2(2(\beta - \tilde{\nu}))$) or there is at least a wrong localization. The discussion below focuses on the optimal choice for a decision threshold using hypotheses tests. Indeed, we have a decision problem with two hypotheses: (1) H_0: there are not impulse noise or there are and we have well localized them and (2) H_1: we have not well localized impulse noise.

We know that either H_0 or H_1 is true. Thus each time the experiment is conducted one of four things can happen: (1) H_0 is true; choose H_0, (2) H_0 is true; choose H_1, (3) H_1 is true; choose H_1, and (4) H_1 is true; choose H_0.

The first and third alternatives correspond to correct choices. The second and fourth alternatives correspond to errors. Since we assume that the decision rule must say either H_0 or H_1, we can view it as a rule for dividing the total observation space into two parts Σ_0 and Σ_1. Whenever an observation falls in Σ_0 we say H_0 and whenever an observation falls in Σ_1 we say H_1.

Let $P_F = \int_{\Sigma_1} p_{y|H_0}(Y|H_0)dy$ be the probability of false alarm and $P_D = \int_{\Sigma_1} p_{y|H_1}(Y|H_1)dy$ the probability of detection. In general, we would like to make P_F as small as possible and in the same time to have P_D as large as possible. Let $\underline{\tilde{loc}}$ be the estimate of the impulse noise location vector ($size(\underline{\tilde{loc}}) = \tilde{\nu}$).

To define Σ_0, we have to look at two cases:

I $\underline{loc} = \underline{\tilde{loc}}$.

II $\underline{loc} \subset \underline{\tilde{loc}}$ and $size(\underline{loc}) < size(\underline{\tilde{loc}})$.

If $\underline{loc} \subset \underline{\tilde{loc}}$ then we can easily verify that $Im(\underline{\underline{V}}) \subset Im(\underline{\underline{\tilde{V}}})$ and $\underline{\underline{P}}\ \underline{\underline{V}} = \underline{\underline{P}}\ \underline{\underline{\tilde{V}}} = \underline{\underline{0}}$, that means, these two cases are equivalent and $p_{y|H_0} \sim \chi^2(2(\beta - \tilde{\nu}))$.

For Σ_1, we look at these cases:

I $\underline{loc} \not\subset \underline{\tilde{loc}}$ and $\underline{loc} \bigcap \underline{\tilde{loc}} = \emptyset$.

II $\underline{loc} \not\subset \underline{\tilde{loc}}$ and $\underline{loc} \bigcap \underline{\tilde{loc}} \neq \emptyset$.

When P_D and P_F are calculated, we plot P_D versus P_F for various values of impulse noise to Gaussian noise ratio (INR) as a parameter on the curve which is often referred to as the Receiver Operating Characteristic (ROC), it completely describes the performance of the test as a function of the parameter of interest. Now, we have to look for the optimal threshold $\delta^{(\tilde{\nu})} = y_{opt}$ from which we can decide if we have well corrected the impulse noise or not. For this we proceed as follows.

A. The decision criteria

We choose the optimal threshold δ that minimize the average risk $\bar{C}$, namely:

$$\begin{aligned} \bar{C} &= \sum_{i=0}^{1}\sum_{j=0}^{1} C_{ij} p_i \int_{\Sigma_i} p(y|H_i) dy \\ &= p_0(C_{01} - C_{00})P_F + p_1(C_{11} - C_{10})P_D + (C_{00}p_0 + C_{10}p_1) \end{aligned} \tag{22}$$

C_{ij} is the cost of choosing hypothesis H_j when H_i is true $(i, j = 0, 1)$. Let p_0 and p_1 denote the probability of occurrence of the hypothesis H_0 and H_1 and suppose that we know these *a priori* probabilities. Therefore, the Bayes criteria [34] defines the region Σ_0 and Σ_1 that minimizes the average risk $\bar{C}$. For each $y = ||\underline{S} - \underline{\tilde{S}}||^2$, we compute the ratio:

$$\Lambda(y) = \frac{p_{Y|H_1}(y|H_1)}{p_{Y|H_0}(y|H_0)}$$

The region Σ_0 consists of y for which $\Lambda(y) < \Lambda_0$, and Σ_1 of values for which $\Lambda(x) > \Lambda_0$, where the critical value Λ_0 is given by:

$$\Lambda_0 = \frac{p_0(C_{01} - C_{00})}{p_1(C_{10} - C_{11})}$$

In the following, we assume that $C_{00} = C_{11} = 0$.

Now if P_D and P_F are known, then we can use the information given by the ROC curve. However, the Bayes threshold can also be deduced from the ROC curve:

$$\Lambda_0 = dP_D/dP_F \tag{23}$$

where Λ_0 is the slope of ROC curve at the point (P_{F_0}, P_{D_0}) [34]. Once, the cost and *a priori* probabilities are known, we deduce from the ROC curve the threshold δ such that $P_{F_0} = \int_{\delta^{(\tilde{\nu})}}^{+\infty} p_{y|H_0}(Y|H_0)dy$. If $\|\underline{\tilde{S}} - \underline{S}\|^2 < \delta^{(\tilde{\nu})}$ then we can conclude that we have corrected impulse noise. So, if p_0 and p_1 are known then we have to choose the threshold Λ_0. From the ROC curve, we remark that when Λ_0 decreases (that means that the slope decreases) and then $P_D = p(H_1|H_1)$ increases. Or in our case we prefer that P_D increase and P_F decreases.

Note that this procedure can also be applied at the beginning of the decoding algorithm, that means if $\| \underline{S} \|^2$ is less than a certain threshold then we can conclude that no impulse noise has taken place in the channel. This has the advantage to avoid the decoding when there are no impulse noise. So to calculate this threshold, we use the same technique that we have already explained.

B. *Combinatorial test*

Let $T = \kappa\tilde{\nu}$ where κ is an integer such that $\kappa > 1$ and in this part we vary $\tilde{\nu}$ from 1 to r (see section III). If $y = \|\underline{\tilde{S}} - \underline{S}\|^2 > \delta^{(\tilde{\nu})}$ then for each value of $\tilde{\nu}$ we proceed as follows: instead of considering that the $\tilde{\nu}$ smallest values of $|\tilde{\Lambda}(x)|$ taken on $W_M^{-[1,\cdots,M]}$ that correspond to the impulse noise location, we take the T smallest values of $|\tilde{\Lambda}(x)|$ and then we compute $\|\underline{\tilde{S}} - \underline{S}\|^2$ for all possible combinations of $\tilde{\nu}$ elements from T elements until obtaining $\|\underline{\tilde{S}} - \underline{S}\|^2 \leq \delta^{(\tilde{\nu})}$.

VI. Example

Let $M = 64$ and the length of the guard interval $IG = 16$ samples long. Among these N carriers, 12 carriers are null (including the middle null carrier and the zeros padded on the both ends). Among the remaining $K = 52$ subcarriers, 4 are fixed pilots carrying known symbols $P_1 \dots P_4$ which are at the position $\{11\ 25\ 38\ 53\}$ while the others $N = K - 4 = 48$ subcarriers convey the information.

To correct impulse noise, we use: P_1, P_3, P_4 and two zeros (the middle null carrier and the one at the position 59), hence $\mathcal{A} = \{11\ 32\ 38\ 53\ 59\}$. Then, we selected the syndrome matrix (see figure 1):

$$\underline{\underline{S}} = \begin{pmatrix} S_{11} & S_{32} & S_{38} \\ S_{32} & S_{53} & S_{59} \end{pmatrix} \tag{24}$$

which can be written as in equation (10), with $m_0 = 0$, $\delta_0 = 11$, $\theta_0 = 0$, $\theta_1 = 21$, $\delta_1 = 32$ and $\delta_2 = 27$.

It follows that:

$$\underline{\underline{\mathcal{R}}}^{(2)} = \begin{pmatrix} 1 & 1 \\ W_M^{21\ f_0} & W_M^{21\ f_1} \end{pmatrix} \qquad \underline{\underline{\mathcal{Q}}}^{(2)} = \begin{pmatrix} W_M^{11\ f_0} & W_M^{32\ f_0} \\ W_M^{11\ f_1} & W_M^{32\ f_1} \end{pmatrix}$$

We verify that $\underline{\underline{\mathcal{R}}}^{(2)}$ and $\underline{\underline{\mathcal{Q}}}^{(2)}$ are invertible because gcd($\theta_1 - \theta_0$,M)= 1 and gcd($\delta_1 - \delta_0$,M)= 1. Thus we can correct at most 2 impulse noise.

We have two hypothesis:

- H_0: there is no impulse noise.
- H_1: there are impulse noise.

In the hypothesis H_0, we must consider all these cases:

- I- there is one impulse noise and we detect two such that $\underline{loc} \subset \underline{\tilde{loc}}$.
- II- there are two impulse noise and we detect these.
- III- there are no impulse noise and we detect two.

We have proved in section (V) that these three cases are equivalent and have the same probability density function $\frac{\|\tilde{\underline{S}}-\underline{S}\|^2}{\frac{M\sigma_b^2}{2}} \sim \chi^2(2(\beta - \tilde{\nu}))$. Therefore: $p_{y|H_0}(y|H_0) = \frac{1}{2\sigma_b^6}\ y\ e^{-\frac{Y}{\sigma_b^2}}$ and $P_F = \int_{\Sigma_0} p_{y|H_0}(Y|H_0)dY$.

In the case of the hypothesis H_1, we have to consider all the following possibilities:

- I_1: there are one impulse noise and we didn't detect it.
- I_2: there are two impulse noise and we detect two such that only one is at the correct localization.
- I_3: there are two errors and we detect two such that: $\underline{loc} \bigcap \underline{\tilde{loc}} = \emptyset$.
- II_k: there are more than two impulse noise ($2 < k \leq M$) and we can not detect them because the correction capacity is overflowed.

Then:

$$P_D = p_{I_1}\int_{\Sigma_1} f(y \in I_1)dy \quad + p_{I_2}\int_{\Sigma_1} f(y \in I_2)dy \quad + \\ p_{I_3}\int_{\Sigma_1} f(y \in I_3)dy \quad + \theta$$

where θ correspond to the case when the correction capacity is overflowed. The cases $I_1,\ I_2,\ I_3$ and $\{II_k\}_{k>2}$ are with a probability of occurrence denoted: $p_{I_1},\ p_{I_2},\ p_{I_3}$ and p_{II_k}, where:

$$p_{I_1} = p_{I_3} = \frac{C_M^1 p^1(1-p)^{M-1}}{1-(1-p)^M}$$

$$p_{I_2} = \frac{C_M^2 p^2(1-p)^{M-2}}{1-(1-p)^M}$$

$$p_{II_k} = \frac{C_M^k p^k(1-p)^{M-k}}{1-(1-p)^M} \quad k \in \{3\ ,\ 4 \ldots M\}$$

In the following, we neglect the case of overflow. In figure 9 we plot the theoretical ROC curves, when $M = 64,\ \sigma_i = 2$ and $p = 2\ 10^{-3}$ for three different value of INR, where $INR = \frac{\sigma_i^2}{\sigma_b^2}$. The probability p is low so we neglect the occurrence of three errors. In figure 10 we compare the simulated and theoretical ROC curves. We remark that theoretical curves and that obtained by simulation are very close and this is due to the estimates made to calculate the pdf.

VII. Simulations

These simulations are reminiscent of the Hiperlan2 standard when 4QAM symbols are emitted. Low-level Gaussian noise sample with variance σ_b^2 are added to each position independently modeling the background noise, the parameter of the Bernoulli sequence is $p = 10^{-3}$ and the variance of the impulse noise is $\sigma_i = 70 * \sigma_b$ and we have choose $\frac{C_{01}}{C_{10}} = \frac{1}{8}$.

Remind that in Hiperlan2, the number of carriers is $M = 64$ and the length of the guard interval is 16. Among these carriers, 12 are null (including the middle null carrier and the zeros padded on the both ends). Among the remaining $K = 52$ subcarriers, 4 are fixed pilots carrying known symbols P_1, P_2, P_3 and P_4 which are at the position $\{11 \quad 25 \quad 39 \quad 53\}$ while the other carriers convey the information. To correct impulse noise, we have used P_1, P_3, P_4 and two zeros (the zeros that is in the middle (32) and one zero on the side band (59)) and where we propose to change the position of P_4 to 38 in order to have condition (17) verified.

In figure 4, we plot $1/MSE(dB)$ (where MSE is the Mean Square Error) as a function of E_s/N_0 (dB), before and after decoding. We calculate the MSE between the emitted and the received symbols for four cases: (1) after impulse noise correction, (2) after adding *a posteriori* control, (3) before impulse noise correction, and (4) we consider only Gaussian noise. We notice a clear improvement of the performances after using *a posteriori* control. Comparing the case after correction and improvement to the case after correction, one sees that we have a gain of almost $2dB$. However it is interesting to use *a posteriori* control.

Figure 5 shows the performances in terms of BER, , when we included a channel C which is a realization of the typical channel Model A specified by Hiperlan2. The *a posteriori* control algorithm also shows good behavior under these circumstances, since the curve after correction of the impulse noise is only marginally different from the curve obtained with Gaussian noise only. The improvement in terms of BER is also shown. Note that this simulation was not containing any classical channel coder. Note also that due to the different situation (zeroes are not consecutive), a comparison with the result of [10]-[13] would be very difficult.

VIII. Conclusion

In this paper we have generalized the procedure of impulse noise correction to the case when syndromes are scattered among the emitted sequence. Pilot tones are generally emitted for synchronization or channel estimation and can also be seen as additional syndromes and used to correct impulse noise. However, the correction capacity is conditioned by the position of these pilot tones in the emitted sequence. We have explained the case when capacity is 2 and the case when it is 3.

Classically, the impulse noise correction is in three steps: (1) estimate the number of impulse noise, (2) find the impulse noise locations and (3) correct the errors. In this paper, we have described the *a posteriori* control step that we have added in order to carefully detect the malfunctions of the decoding algorithm. This presented procedure is essentially based on the theory of hypotheses test. Many extension are under consideration, in order to increase the practical usefulness of this technique.

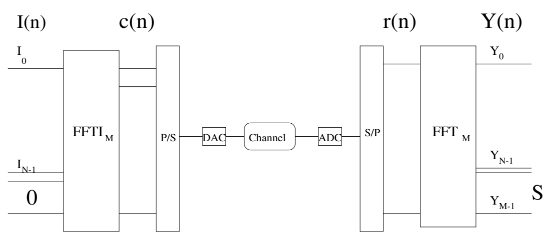

Fig. 1. OFDM system

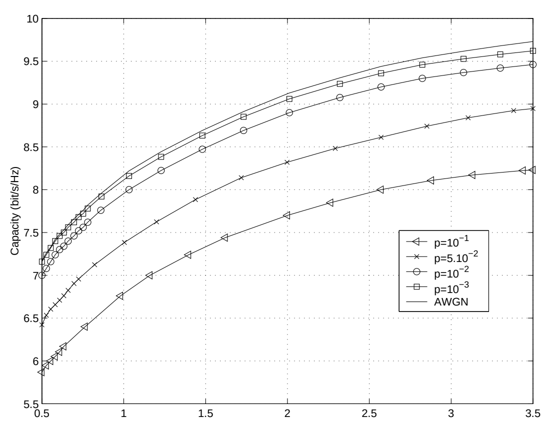

Fig. 2. The “Gaussian plus Bernoulli Gaussian” channel cacity

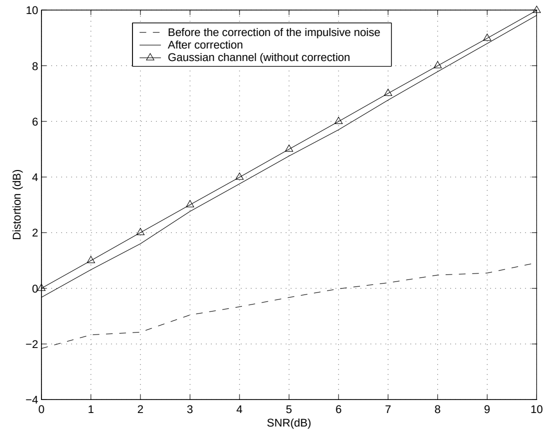

Fig. 3. Distortion performance when we consider a “Gaussian plus Bernoulli Gaussian” channel, and consecutive syndrome locations

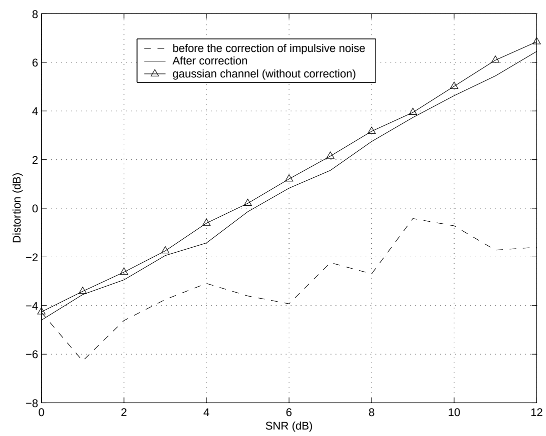

Fig. 4. Distortion performance when we consider a channel, scattered null carriers and pilots tones

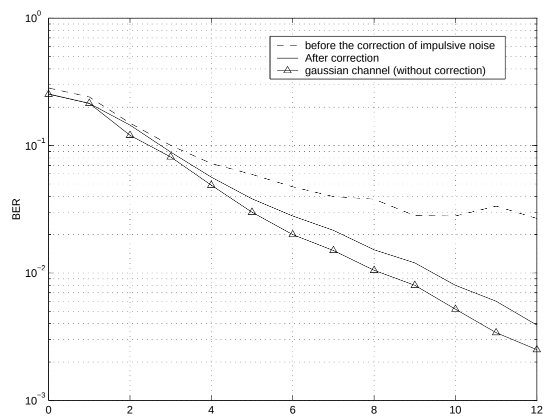

Fig. 5. BER performance when we consider a channel, scattered null carriers and pilots tones

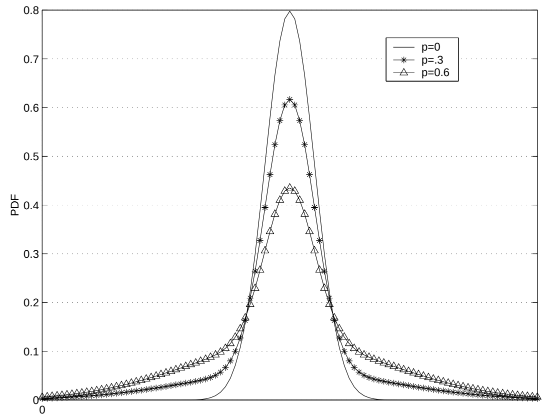

Fig. 6. The probability density function of the channel noise

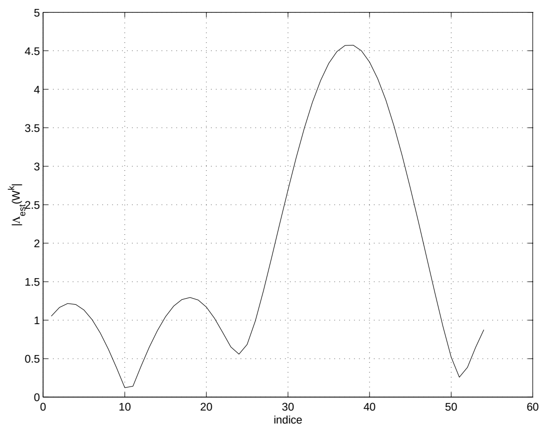

Fig. 7. The error-locator polynomial amplitude

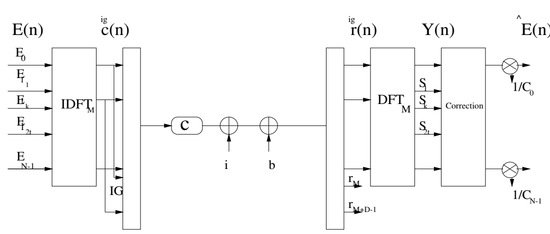

Fig. 8. OFDM system based on the use of a guard interval

References

[1] J.Bingham, "Multicarrier modulation for data transmission: an idea whose time has come," *IEEE Communication magazine*, , no. 28, pp. 15–14, April 1990.

[2] M.Alard and R.Lassalle, "Principles of modulation and channel coding for digital broadcasting for mobile receivers," *EBU Review*, pp. 168–190, August 1997.

[3] L.J Cimini, "Analysis and simulation of a digital mobile channel using orthogonal frequency division multi-

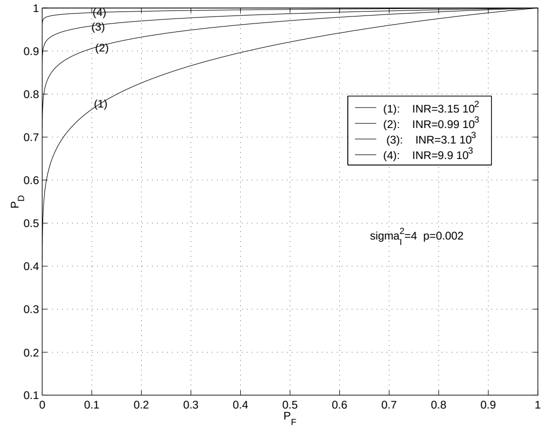

Fig. 9. The theoretical curves ROC

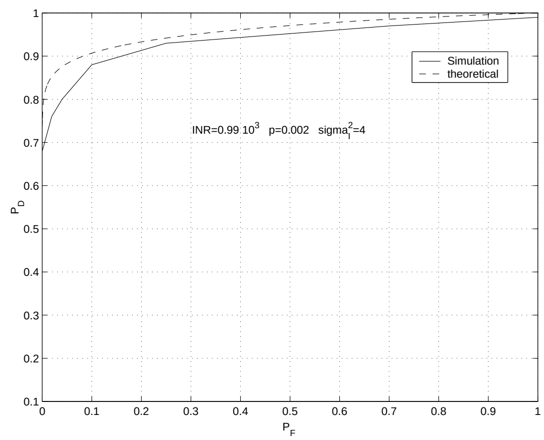

Fig. 10. Comparison between the theoretical ROC curves and simulated ROC curves

plexing," *IEEE Transactions on Communications*, vol. 33, no. 7, pp. 665–675, July 1985.

[4] P.S Chow, J.C.Tu, and J.M Cioffi, "Performance evaluation of a multichannel transceiver system for ADSL and HDSL services," *IEEE Journal on Selected Areas in Communications*, vol. 9, no. 6, pp. 909–919, August 1991.

[5] Kenneth J.Kerpez, "Forward error correction for asymmetric digital subscriber lines (ADSL)," in *IEEE Global Telecommunications Conference*, Phoenix, AZ, 1991.

[6] Kenneth L.Blackard, Theodore S.Rappaport, and Charles W.Bostian, "Measurments and models of radio frequency impulsive noise for indoor wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 11, no. 7, pp. 991–1001, September 1993.

[7] J.G.Proakis, *Digital Communication*, McGraw-Hill, New York, 1989.

[8] Monisha Ghosh, "Analysis of the Effect of Impulse Noise on Multicarrier and Single Carrier QAM Systems," *IEEE Transactions on Communications*, vol. 44, pp. 145–147, February 1996.

[9] Fatma Abdelkefi, Abraham Gabay, and Pierre Duhamel, "Impulse noise cancellation in multicarrier trans-

mission," in *ICASSP*, 2001.

[10] Jack Keil Wolf, "Redundancy, the discrete Fourier Transform, and impulse noise cancellation," *IEEE Transactions on Communications*, vol. 31, no. 3, pp. 458–461, March 1983.

[11] G.Robert Redinbo, "Decoding real block codes: Activity detection, Wiener estimation," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 609–623, March 2000.

[12] Ja-Ling Wu and Jiun Shiu, "Discrete cosine transform in error control coding," *IEEE Transaction on Communications*, vol. 43, no. 5, pp. 1857–1861, May 1995.

[13] R.Kumaresan, "Rank reduction techniques and burst error-correction decoding in real/complex fields," in *Asilomar conf, circuits, systems, computers*, November 1985.

[14] Farokh Marvasti, Mohammed Hasan, M Echhart, and Siamak Talebi, "Effecient algorithms for burst error recovery using FFT and other transform kernels," *IEEE Transactions on signal processing*, vol. 47, no. 4, pp. 1065–1075, April 1999.

[15] Ja-Ling Wu and Jiun Shiu, "Real-valued error control coding by using DCT," *IEE Proceedings-I*, vol. 139, no. 2, April 1992.

[16] G.Robert Redinbo, "A fault-tolerant decoding procedure for real cyclic codes," .

[17] C.R.P. Hartmann, "Generalizations of the BCH Bound," *Information And Control*, , no. 20, 1972.

[18] C.Roos, "A Generalization of the BCH Bound for Cyclic Codes, including the Hartmann-Tzeng Bound," *Journal of Combinatorial Theory*, , no. series A 33, 1982.

[19] Richard E.Blahut, "Computation of channel capacity and rate-distortion functions," *IEEE Transactions on Infomation Theory*, vol. IT-18(4), pp. 460–473, 1972.

[20] Suguru Arimoto, "An algorithm for computing the capacity of arbitrary discrete memoryless channels," *IEEE Transactions on Information Theory*, vol. IT-18, no. 1, pp. 14–20, January 1972.

[21] Fredy.D Neeser and James L.Massey, "Proper complex random processes with application to information theory," *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1293–1303, July 1993.

[22] G.R.Grimmett and D.R.Stirzaker, *Probability and Random Processes*, Oxford University Press, New York, NY, 1992.

[23] R.E.Blahut, "Transform techniques for error control codes," *IBM Journal of Research and Development*, vol. 23, pp. 299–315, May 1979.

[24] Richard.E.Blahut, *Algebraic Methods for Signal Processing and Communications Coding*, Signal Processing and Digital Filtering C.S Burrus ed. Spring-Verlag New York, 1992.

[25] Richard E.Blahut, *Theory and practice of error control codes*, Reading MA: Addison-Wesley, 1987.

[26] T.G.Marshall, "Decoding of Real-Number Error-Correction Codes," in *Proc of GLOBECOM 83*, San Diego, Nov 1983, pp. 1249–1303.

[27] Carlos R.P.Hartmann, Kenneth K.Tzeng, and Robert T.Chien, "Some results on the minimum distance structure of cyclic codes," *IEEE Transactions on Information Theory*, vol. IT-18, no. 3, pp. 402–409, May 1972.

[28] Fatma Abdelkefi, Pierre Duhamel, and Florence Alberge, "On the use of pilot tones for impulse noise cancellation in Hiperlan2," in *ISSPA*, 2001.

[29] Fatma Abdelkefi, Pierre Duhamel, and Florence Alberge, "Codage correcteur d'erreur dans le corps des complexes et systèmes multiporteuses," in *GRETSI*, 2001.

[30] Fatma Abdelkefi, Pierre Duhamel, and Florence Alberge, "A posteriori control of complex reed solomon decoding with application to impulse noise cancellation in Hiperlan2," in *ICC*, 2002.

[31] Olivier Rioul, "A spectral algorithm for removing salt and pepper from images," in *Proc IEEE Digital Signal Processing Workshop*, Loen Norway, September 1996.

[32] Abraham Gabay, "Spectral Interpolation Coder for Impulse Noise Cancellation over a Binary Symmetric Channel," *EUSIPCO*, 2000.

[33] Steven M.Kay and Arnab K.Shaw, "Frequency estimation by principal component AR spectral estimation method without eigendecomposition," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 1, pp. 95–101, January 1988.

[34] H.Vincent Poor, *An introduction to signal detection and estimation*, Spring-Verlag New York, 1994.