An Upper Bound on the Error Probability of a Communication System with Nonparametric Detection

Dmitry Osipov
Institute for Information Transmission Problems RAS
Moscow, Russia
National Research University Higher School of Economics
Moscow, Russia

***Abstract*—In what follows a channel model that describes a class of communication systems employing a certain non-parametric receiver is considered. In particular a block coded transmission via the channel under consideration and soft-input mismatched decoding are considered. The paper deals with a problem of finding an upper bound on the probability of erroneous decoding. In order to obtain the bound in question a second version of the Duman-Salehi bound is applied. For the system under consideration a closed form for the normalized tilting measure minimizing the bound under consideration is obtained.**

I. INTRODUCTION

In recent decades a number of reception techniques based on parametric hypothesis testing methods were developed in order to provide reliable communications and high data rates. However if channel state information is either unavailable or imprecise nonparametric reception techniques are to be used [1], [2], [3], [4], [5]. In this paper a channel model corresponding to a class of communication systems using nonparametric reception method proposed in [6] is considered.

This paper is organized as follows. In section II a description of the channel under consideration is given, the relation between the model in question and the connection between the channel model under consideration and real-life communication systems using nonparametric receiver proposed in [6] are discussed. In section III a description of the system using coded transmission via the channel under consideration and soft-input mismatched decoder is given. In section IV is obtained by applying the second type Duman-Salehi bounding technique. For the case under consideration (i.e. for mismatched decoding of the coded transmission via the channel under consideration explicit closed form expression for normalized tilting measure minimizing the proposed bound is given. Finally in section V the results obtained and the contribution of this paper are summarized.

II. CHANNEL MODEL

Let q and α be natural numbers such that $q > 2, q > \alpha$. Let us define the following sets:

$$\begin{gathered}\mathbb{B}_q^1 = \{\boldsymbol{b} = (b_1, b_2, \ldots, b_q)^T : \\ \forall\, i \in \{1:q\}\, b_i \in \{0,1\}\,, w_H(\boldsymbol{b}) = 1\}\end{gathered} \tag{1}$$

- the set of all binary column vectors of length q and Hamming weight 1,

$$\begin{gathered}\mathbb{B}_q^\alpha = \{\boldsymbol{b} = (b_1, b_2, \ldots, b_q)^T : \\ \forall\, i \in \{1:q\}\, b_i \in \{0,1\}\,, w_H(\boldsymbol{b}) = \alpha\}\end{gathered} \tag{2}$$

- the set of all binary column vectors of length q and Hamming weight α. Moreover for any column vector $\boldsymbol{z}$ such that $w_H(\boldsymbol{z}) < \alpha$ let us define

$$\mathbb{S}^1(\boldsymbol{z}) = \left\{\boldsymbol{s} : \boldsymbol{s} \in \mathbb{B}_q^\alpha, \boldsymbol{s} \wedge \boldsymbol{z} = \boldsymbol{z}\right\} \tag{3}$$

- the set of all binary column vectors of Hamming weight α that "cover" the vector $\boldsymbol{z}$

$$\mathbb{S}^0(\boldsymbol{z}) = \left\{\boldsymbol{s} : \boldsymbol{s} \in \mathbb{B}_q^\alpha, \boldsymbol{s} \wedge \boldsymbol{z} \neq \boldsymbol{z}\right\} \tag{4}$$

- the set of all binary column vectors of Hamming weight α that don't "cover" the vector $\boldsymbol{z}$.

Let us designate the input of the channel with $\boldsymbol{x}$ and the output with $\boldsymbol{y}$.The channel under consideration is then given by the following equations

$$\begin{gathered}\boldsymbol{x} \in \mathbb{B}_q^1, \boldsymbol{y} \in \mathbb{B}_q^\alpha \\ p(\boldsymbol{x}) = p\left(\boldsymbol{y} \in \mathbb{S}^1(\boldsymbol{y}) \mid \boldsymbol{x}\right)\end{gathered} \tag{5}$$

and additional conditions

$$\begin{gathered}\forall \boldsymbol{x} : p(\boldsymbol{x}) = p \\ \forall \boldsymbol{x}, \boldsymbol{y}_i \in \mathbb{S}^1(\boldsymbol{x}), \boldsymbol{y}_j \in \mathbb{S}^1(\boldsymbol{x}), i \neq j : \\ p(\boldsymbol{y}_i \mid \boldsymbol{x}) = p(\boldsymbol{y}_j \mid \boldsymbol{x}) = p_1 \\ \forall \boldsymbol{x}, \boldsymbol{y}_g \in \mathbb{S}^0(\boldsymbol{x}), \boldsymbol{y}_h \in \mathbb{S}^0(\boldsymbol{x}), g \neq h : \\ p(\boldsymbol{y}_g \mid \boldsymbol{x}) = p(\boldsymbol{y}_h \mid \boldsymbol{x}) = p_0\end{gathered} \tag{6}$$

the channel model given by (5) can be interpreted in the following way: let as assume that the communications channel in use is split into q non-overlapping subchannels either in time or frequency domain (the former case corresponds to Impulse Radio transmission [7], [8] and the latter to multi-tone transmission techniques [9] e.g. OFDM [10]) and q-ary symbols are transmitted via the channel in use by employing either Pulse Position Modulation (if the subchannels are time slots) or q-ary frequency shift keying (if the subchannels are

978-1-5090-4231-9/16/$31.00 ©2016 IEEE

subcarriers). The output of each subchannel can be affected by both background noise and interfering signals (e.g. multiuser interference or intentional jamming) The receiver measures certain parameter (e.g. signal power) of the signal at the output of each subchannel and chooses α "best" (in a certain sense, e.g. the receiver can choose α subchannels having energy greater than others) subchannels. The list of the numbers of the subchannels that were chosen is the output of the receiver (within the scope of the model described above the subchannels being chosen are marked by the nonzero entries in the vector $\boldsymbol{Y}$). The probability $p(\boldsymbol{x})$ is then a probability of the fact that the number of the correct subchannel (i.e. the one that has been used for transmission and corresponds to $\boldsymbol{x}$) has been included in the list. As for the conditions (6) these will hold if each of the interfering signals occupies a certain subchannel equiprobably. The latter is certainly not always the case in real-world systems however this can be guaranteed if prior to sending a certain vector a pseudorandom permutation is applied (the permutation is to be chosen equiprobably from the set of all possible permutations and the choice is to be carried out for each vector) and at the receiver end inverse permutation is done prior to decoding.

As can be seen from (5) the channel under consideration is nothing but the discrete memoryless channel (DMC). In fact if the conditions (6) holds the channel under consideration is a symmetric one (in the sense of the definition given by [11], [16]). Moreover since

$$\begin{aligned} &\forall \boldsymbol{z} \in \mathbb{B}_q^1 \\ &|\mathbb{S}^1(\boldsymbol{z})| = \sigma_1 = \binom{q-1}{\alpha-1} \\ &|\mathbb{S}^0(\boldsymbol{z})| = \sigma_0 = \binom{q-1}{\alpha} \end{aligned} \tag{7}$$

if (6) holds the respective probabilities are given by

$$\begin{aligned} p_1 &= \frac{p}{\sigma_1} \\ p_0 &= \frac{1-p}{\sigma_0} \end{aligned} \tag{8}$$

An example of the equivalent channel transition diagram and channel transition matrix for the case $q = 5, \alpha = 2$ is given in Fig. 1 Even though the channel model under consideration belongs to a well known class of symmetric DMCs it is in many respects different from the well known channel models from this class. In particular the cardinality of the input and output alphabet usually either coincide (as e.g. in q-ary symmetric channel) or the input alphabet is a subset of the output alphabet (as e.g. in q-ary erasure channel).For our case however none of the two previous statements is true. It is worth noting that the idea of the channel having input and output alphabet of sufficiently different cardinality has been addressed in [12]. However even in the case considered in [12] the input alphabet was a subset of the output one which is certainly not the case to be considered hereinafter.

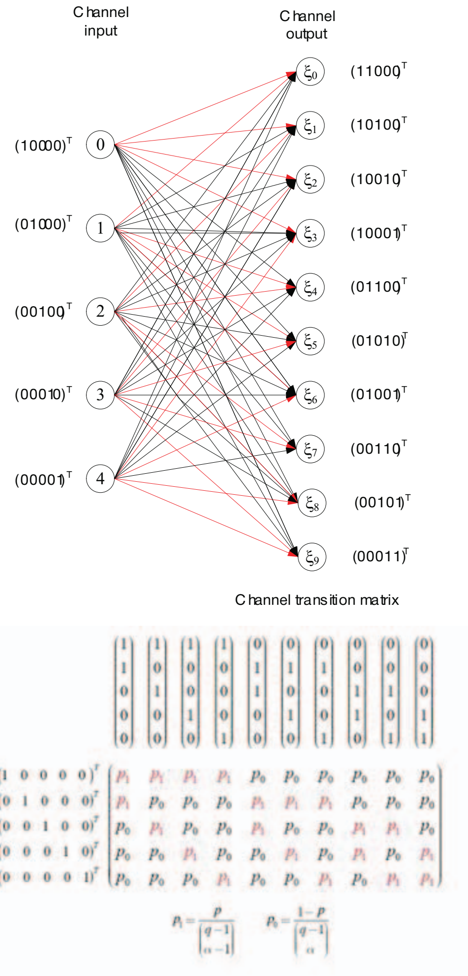

Fig. 1. Channel model (an example for $\alpha = 2, q = 5$)

III. Coded transmission and decoding

Let us now consider coded transmission via the channel under consideration. Let us assume that the information to be transmitted is encoded with an $C(N, K, d)$ linear block code. Each (say jth) symbol of the mth codeword v_j^m is mapped into a vector $\boldsymbol{x}_j^m$. Thus each codeword $\boldsymbol{v}^m$ is mapped into a matrix $\boldsymbol{X}^m$. Within the scope of the mapping under consideration ith symbol is mapped into a binary column vector of length q with nonzero entry at the ith position and zero entries at the remainder positions (please note that if the code in use is an MDS code the construction under consideration is exactly the one considered by Kautz and Singleton [17]). Thus to transmit a codeword $\boldsymbol{v}^m$ a corresponding matrix $\boldsymbol{X}^m = [\boldsymbol{x}_1^m, \boldsymbol{x}_2^m, \dots \boldsymbol{x}_N^m]$ is to be transmitted via the channel under consideration. Correspondingly a matrix $\boldsymbol{Y}^j = [\boldsymbol{y}_1^j, \boldsymbol{y}_2^j, \dots \boldsymbol{y}_N^j]$ is to be received. Thus in what follows we shall use the phrases "a codeword was transmitted" and "a matrix was transmitted" as synonyms (the same holds for "received vector" and "received matrix" since in the channel under consideration we receive matrixes rather than vectors).

Let us now assume that for any $\boldsymbol{X}^m$ that could have been transmitted and any received matrix $\boldsymbol{Y}^j$ a "reliability" function $\Theta(\boldsymbol{Y}^j \mid \boldsymbol{X}^m)$ can be computed. The decoding procedure then boils dow to choosing

$$\boldsymbol{X}^* = \arg\max_m \Theta\left(\boldsymbol{Y}^j \left|\boldsymbol{X}^m\right.\right) \quad (9)$$

(in what follows we shall assume that ties are resolved randomly). Hereinafter we shall consider the case when the reliability function admits multiplicative form i.e.

$$\Theta\left(\boldsymbol{Y}^j \left|\boldsymbol{X}^m\right.\right) = \prod_{n=1}^{N} \theta_n\left(\boldsymbol{y}_n^j \left|\boldsymbol{x}_n^m\right.\right) \quad (10)$$

and $\forall\, n = \overline{1:N}$

$$\theta_n\left(\boldsymbol{y}_n^j \left|\boldsymbol{x}_n^m\right.\right) = \begin{cases} \eta\theta & \boldsymbol{y}_n^j \in \mathbb{S}^1\left(\boldsymbol{x}_n^m\right) \\ \theta & \boldsymbol{y}_n^j \in \mathbb{S}^0\left(\boldsymbol{x}_n^m\right) \end{cases} \quad (11)$$

where $\eta \geq 1, \theta > 0$. Please note that if $\eta = \frac{p_1}{p_0}$ and $\theta = p_0$ the decoder under consideration is nothing but the Maximum Likelihood decoder for the case under consideration. However we interested in a more general case i.e. the case when η and θ are some constants that meet the aforementioned conditions. This case is commonly referred to as "mismatched" decoding [18], [19], [20], [21], [22]. In next section an upper bound on the probability of error for the case under consideration will be obtained.

IV. Upper bound

Hereinabove it has been shown that the problem under consideration boils down to finding an upper bound on the probability of error for the decoding rule described above. To solve this problem a classical approach [23], [24], [25] will be used: we shall assume that the code in use is partitioned into a number of subcodes (it will be assumed that each subcode C_w consists of an all-zero codeword and all codewords of a certain Hamming weight w). Without loss of generality the probability of error can be upper bounded

$$P_e = P_{e|0} \leq \sum_{w=d}^{N} P_{e|0}(w) \quad (12)$$

where $P_{e|0}(w)$ is the probability of error for the case when a subcode C_w is used and an all-zero codeword has been transmitted (due to the linearity of the code in use the bound (12) does not depend on the codeword transmitted). To obtain an upper bound on the probability $P_{e|0}(w)$ we shall use the second version of the Duman-Salehi bound (DS2 bound) [13], [14], [23]. The latter is given by

$$P_{e|m} \leq \left(\sum_{l\neq m}\sum_{j} \psi_N^m\left(\boldsymbol{Y}^j\right)^{1-\frac{1}{\rho}} p_N\left(\boldsymbol{Y}^j \left|\boldsymbol{X}^m\right.\right)^{\frac{1}{\rho}} \cdot \cdot \left(\frac{\Theta\left(\boldsymbol{Y}^j \left|\boldsymbol{X}^l\right.\right)}{\Theta\left(\boldsymbol{Y}^j \left|\boldsymbol{X}^m\right.\right)}\right)^s\right)^\rho \quad (13)$$

where ρ and s meet the conditions

$$1 \geq \rho > 0,\ s > 0 \quad (14)$$

and $P_{e|m}$ is the probability of error if the matrix $\boldsymbol{X}^m$ (corresponding to the m-th codeword) has been transmitted, $p_N\left(\boldsymbol{Y}^j \left|\boldsymbol{X}^m\right.\right)$ is the probability to receive $\boldsymbol{Y}^j$ if the matrix $\boldsymbol{X}^m$ has been transmitted and $\psi_N^m\left(\boldsymbol{Y}^j\right)$ is the normalized tilting measure: a function that depends on $\boldsymbol{Y}^j$ may also depend on $\boldsymbol{X}^m$ and meets the following conditions

$$\begin{gathered} \forall\, \boldsymbol{Y} \in \mathbb{Y}\ \psi_N^m\left(\boldsymbol{Y}\right) > 0 \\ \sum_{\boldsymbol{Y}\in\mathbb{Y}} \psi_N^m\left(\boldsymbol{Y}\right) = 1 \end{gathered} \quad (15)$$

where

$$\mathbb{Y} = \left\{Y = [y_1, y_2, \ldots, y_N] : \forall m = \overline{1:N} y_m \in \mathbb{B}_q^\alpha\right\} \quad (16)$$

is the set of matrixes that can be received. Since the right hand side of (13) is an upper bound the tilting measure is to be chosen to minimize the bound for any fixed ρ and s meeting (14). Since in the case under consideration the bound (13) is applied to subcodes separately it admits the following form

$$P_{e|0}\left(w\right) \leq \bar{p}\left(w, \psi_N^0(\boldsymbol{Y})\right) = \left(A_w\right)^\rho \left(f_0\left(w, \psi_N^0(\boldsymbol{Y})\right)\right)^\rho \quad (17)$$

where $P_{e|0}\left(w\right)$ is the probability to receive $\boldsymbol{Y}^j$ if the subcode C_w is used for transmission and matrix $\boldsymbol{X}^0$ (corresponding to the all-zero codeword) has been transmitted, $A_w = |C_w| - 1$ is the number of codewords of the weight w in the code C, and $f_0\left(w, \psi_N^0(\boldsymbol{Y})\right)$ is given by

$$f_0\left(w, \psi_N^0(\boldsymbol{Y})\right) = \sum_{\boldsymbol{Y}^j\in\mathbb{Y}} \left(\psi_N^0\left(\boldsymbol{Y}^j\right)^{1-\frac{1}{\rho}} p_N\left(\boldsymbol{Y}^j \left|\boldsymbol{X}^m\right.\right)^{\frac{1}{\rho}} \cdot \cdot \left(\frac{\Theta\left(\boldsymbol{Y}^j \left|\boldsymbol{X}^l\right.\right)}{\Theta\left(\boldsymbol{Y}^j \left|\boldsymbol{X}^m\right.\right)}\right)^s\right) \quad (18)$$

where $\boldsymbol{X}^l$ is the matrix corresponding to a certain codeword of weight w. Since A_w does not depend on $\psi_N^0(\boldsymbol{Y})$ and $y = x^\rho$ is an increasing function

$$\bar{p}\left(w, \psi_N^0(\boldsymbol{Y})\right) \underset{\psi_N^0(Y)}{\rightarrow} \min \Rightarrow f_0\left(w, \psi_N^0(\boldsymbol{Y})\right) \underset{\psi_N^0(Y)}{\rightarrow} \min \quad (19)$$

where minimization is to be carried out subject to constraints (15). Let us now obtain an explicit form of the minimization problem (19).

If $\boldsymbol{X}^0$ has been transmitted the probability to receive a certain matrix $\boldsymbol{Y}^j$ for which

$$\left|\left\{\boldsymbol{x}_t^0 : t = \overline{1:N}, \boldsymbol{y}_t^j \in \mathbb{S}^1\left(\boldsymbol{x}_t^0\right)\right\}\right| = i \quad (20)$$

is given by

$$p_N\left(\boldsymbol{Y}^j \left|\boldsymbol{X}^0\right.\right) = p_1^i p_0^{N-i} = \left(\frac{p}{\sigma_1}\right)^i \left(\frac{1-p}{\sigma_0}\right)^{N-i} \quad (21)$$

Let us define the value

$$h = \left|\left\{\boldsymbol{x}_t^0 : t = \overline{1:N}, \boldsymbol{y}_t^j \in \mathbb{S}^1\left(\boldsymbol{x}_t^0\right), \boldsymbol{x}_t^0 \neq \boldsymbol{x}_t^l\right\}\right| \quad (22)$$

- the number of positions such that matrices $\boldsymbol{X}^0$ and $\boldsymbol{X}^l$ differ in the respective columns (and thus corresponding codewords differ in the respective positions) and the respective columns of $\boldsymbol{X}^0$ are "covered" by the respective columns of $\boldsymbol{Y}^j$. Since $\boldsymbol{X}^0$ and $\boldsymbol{X}^l$ correspond to the all-zero codeword and some codeword of weight w respectively we can claim $h \le w$. Moreover due to (20) and (22) we can claim that $h \le i$ and thus $h \le \min(w,i)$. Since $\boldsymbol{X}^0$ and $\boldsymbol{X}^l$ correspond to the all-zero codeword and some codeword of weight w respectively they coincide in $N-w$ columns and due to (20) and (22) we can deduce that $i-h$ of those columns are "covered" by the respective columns of $\boldsymbol{Y}^j$. Therefore the number of matrixes $\boldsymbol{Y}^j$ for which (22) hold is given by

$$\mathrm{H}_0(N,w,i,h) = \binom{N-w}{i-h}\sigma_1^{i-h}\sigma_0^{N-w-i+h} \tag{23}$$

Let us define the following values

$$\begin{aligned} k &= \left|\left\{\boldsymbol{x}_t^0 : t=\overline{1:N}, \boldsymbol{y}_t^j \in \mathbb{S}^1\left(\boldsymbol{x}_t^0\right), \boldsymbol{y}_t^j \in \mathbb{S}^0\left(\boldsymbol{x}_t^l\right)\right\}\right| \\ g &= \left|\left\{\boldsymbol{x}_t^0 : t=\overline{1:N}, \boldsymbol{y}_t^j \in \mathbb{S}^0\left(\boldsymbol{x}_t^0\right), \boldsymbol{y}_t^j \in \mathbb{S}^1\left(\boldsymbol{x}_t^l\right)\right\}\right| \end{aligned} \tag{24}$$

where k is the number of positions such that matrices $\boldsymbol{X}^0$ and $\boldsymbol{X}^l$ differ in the respective columns (and thus corresponding codewords differ in the respective positions) and the respective columns of $\boldsymbol{X}^0$ are "covered" by the respective columns of $\boldsymbol{Y}^j$ while the respective columns of $\boldsymbol{X}^l$ are not and g is the number positions such that matrices $\boldsymbol{X}^0$ and $\boldsymbol{X}^l$ differ in the respective columns and the respective columns of $\boldsymbol{X}^l$ are "covered" by the respective columns of $\boldsymbol{Y}^j$ while the respective columns of are not. Let us now define

$$\begin{aligned} \upsilon_{\langle 1,1\rangle} &= \left|\left\{\boldsymbol{y} : \boldsymbol{y} \in \mathbb{B}_q^\alpha, \boldsymbol{y} \in \mathbb{S}^1\left(\boldsymbol{x}^0\right), \boldsymbol{y} \in \mathbb{S}^1\left(\boldsymbol{x}^l\right)\right\}\right| = \\ &= \binom{q-2}{\alpha-2} \\ \upsilon_{\langle 1,0\rangle} &= \left|\left\{\boldsymbol{y} : \boldsymbol{y} \in \mathbb{B}_q^\alpha, \boldsymbol{y} \in \mathbb{S}^1\left(\boldsymbol{x}^0\right), \boldsymbol{y} \in \mathbb{S}^0\left(\boldsymbol{x}^l\right)\right\}\right| = \\ &= \binom{q-2}{\alpha-1} \\ \upsilon_{\langle 0,1\rangle} &= \left|\left\{\boldsymbol{y} : \boldsymbol{y} \in \mathbb{B}_q^\alpha, \boldsymbol{y} \in \mathbb{S}^0\left(\boldsymbol{x}^0\right), \boldsymbol{y} \in \mathbb{S}^1\left(\boldsymbol{x}^l\right)\right\}\right| = \\ &= \binom{q-2}{\alpha-1} \\ \upsilon_{\langle 0,0\rangle} &= \left|\left\{\boldsymbol{y} : \boldsymbol{y} \in \mathbb{B}_q^\alpha, \boldsymbol{y} \in \mathbb{S}^0\left(\boldsymbol{x}^0\right), \boldsymbol{y} \in \mathbb{S}^0\left(\boldsymbol{x}^l\right)\right\}\right| = \\ &= \binom{q-2}{\alpha} \end{aligned} \tag{25}$$

The number of matrixes $\boldsymbol{Y}^j$ for which (24) holds (for some i and $h \le \min(w,i)$) is given by

$$\mathrm{H}_1(N,w,i,h,g,k) = \binom{w}{h-k,k,g}\upsilon_{\langle 1,1\rangle}^{h-k}\upsilon_{\langle 1,0\rangle}^{k}\upsilon_{\langle 0,1\rangle}^{g}\upsilon_{\langle 0,0\rangle}^{w-h-g} \tag{26}$$

and the number of matrixes $\boldsymbol{Y}^j$ for which (20),(22),(24) hold is given by

$$\begin{aligned} &\mathrm{H}(N,w,i,h,g,k) = \\ &= \mathrm{H}_0(N,w,i,h)\mathrm{H}_1(N,w,i,h,g,k) = \\ &= \binom{N-w}{i-h}\sigma_1^{i-h}\sigma_0^{N-w-i+h}\cdot \\ &\cdot\binom{w}{h-k,k,g}\upsilon_{\langle 1,1\rangle}^{h-k}\upsilon_{\langle 1,0\rangle}^{k}\upsilon_{\langle 0,1\rangle}^{g}\upsilon_{\langle 0,0\rangle}^{w-h-g} \end{aligned} \tag{27}$$

Please note that if $\boldsymbol{Y}^j$,$\boldsymbol{X}^0$ and $\boldsymbol{X}^l$ are such that (24) hold

$$\frac{\Theta\left(\boldsymbol{Y}^j\left|\boldsymbol{X}^l\right.\right)}{\Theta\left(\boldsymbol{Y}^j\left|\boldsymbol{X}^m\right.\right)} = \eta^{g-k} \tag{28}$$

Finally since $\psi_N^0(\boldsymbol{Y})$ depends on $\boldsymbol{Y}$ let us define $\psi_N^0(\boldsymbol{Y})$ in the following way

$$\begin{aligned} &\forall\, \boldsymbol{Y}^j = [\boldsymbol{y}_1^j, \boldsymbol{y}_2^j, \ldots, \boldsymbol{y}_N^j] : t = \left|\left\{m : \boldsymbol{y}_m^j \in \mathbb{S}^1\left(\boldsymbol{x}_m^0\right)\right\}\right| \\ &\psi_N^0(\boldsymbol{Y}^j) = \Psi_t \end{aligned} \tag{29}$$

Putting (27),(21),(25),(28) and (29) together we obtain

$$\begin{aligned} f_0\left(\Psi_0,\ldots,\Psi_N\right) &= \sum_{i=0}^{N}\sum_{h=0}^{\min(w,i)}\sum_{k=0}^{h}\sum_{g=0}^{w-h}\left(\Psi_i^{\left(1-\frac{1}{\rho}\right)}\cdot\right. \\ &\cdot\eta^{s(g-k)}\left(\binom{N-w}{i-h}\sigma_1^{i-h}\sigma_0^{N-w-i+h}\right)\cdot \\ &\cdot\left(\binom{w}{h-k,k,g}\upsilon_{\langle 1,1\rangle}^{h-k}\upsilon_{\langle 1,0\rangle}^{k}\upsilon_{\langle 0,1\rangle}^{g}\upsilon_{\langle 0,0\rangle}^{w-h-g}\right)\cdot \\ &\cdot\left.\left(\left(\frac{p}{\sigma_1}\right)^{\frac{i}{\rho}}\left(\frac{1-p}{\sigma_0}\right)^{\frac{N-i}{\rho}}\right)\right) = \sum_{i=0}^{N}\beta_i\Psi_i^{\left(1-\frac{1}{\rho}\right)} \end{aligned} \tag{30}$$

where β_i is given by

$$\begin{aligned} \beta_i &= \left(\left(\frac{p}{\sigma_1}\right)^{\frac{i}{\rho}}\left(\frac{1-p}{\sigma_0}\right)^{\frac{N-i}{\rho}}\right)\cdot \\ &\cdot\sum_{h=0}^{\min(w,i)}\sum_{k=0}^{h}\sum_{g=0}^{w-h}\left(\eta^{s(g-k)}\cdot\right. \\ &\cdot\left(\binom{w}{h-k,k,g}\upsilon_{\langle 1,1\rangle}^{h-k}\upsilon_{\langle 1,0\rangle}^{k}\upsilon_{\langle 0,1\rangle}^{g}\upsilon_{\langle 0,0\rangle}^{w-h-g}\right)\cdot \\ &\cdot\left.\left(\binom{N-w}{i-h}\sigma_1^{i-h}\sigma_0^{N-w-i+h}\right)\right) \end{aligned} \tag{31}$$

the minimization problem in question is then given by

$$\begin{aligned} &f_0\left(\Psi_0,\ldots,\Psi_N\right) = \sum_{i=0}^{N}\beta_i\Psi_i^{\left(1-\frac{1}{\rho}\right)} \underset{\Psi=[\Psi_0,\ldots,\Psi_N]}{\rightarrow} \min \\ &\forall i = \overline{1:N} - \Psi_i \le 0 \\ &\sum_{i=0}^{N}\gamma_i\Psi_i - 1 = 0 \end{aligned} \tag{32}$$

where γ_i is given by

$$\gamma_i = \binom{N}{i} \tag{33}$$

The Lagrangian function for this problem is given by

$$\begin{aligned} &\Lambda\left(\Psi_0,\ldots,\Psi_N\right) = \\ &\sum_{i=0}^{N}\beta_i\Psi_i^{\left(1-\frac{1}{\rho}\right)} + \tilde{\lambda}\left(\sum_{i=0}^{N}\gamma_i\Psi_i - 1\right) + \sum_{i=0}^{N}\lambda_i\left(-\Psi_i + u_i^2\right) \end{aligned} \tag{34}$$

where

$$\begin{aligned} &\tilde{\lambda} \ge 0 \\ &\forall i = \overline{0:N}\lambda_i \ge 0 \end{aligned} \tag{35}$$

The the Karush-Kuhn-Tucker (KKT) conditions are then given by

$$\begin{aligned}
&\forall i = \overline{0:N} \\
&\beta_i \left(1 - \frac{1}{\rho}\right) \Psi_i^{\left(-\frac{1}{\rho}\right)} + \tilde{\lambda}\gamma_i - \lambda_i = 0 \\
&-\Psi_i + u_i^2 = 0 \\
&2\lambda_i u_i = 0 \\
&\sum_{i=0}^{N} \gamma_i \Psi_i - 1 = 0 \\
&\lambda_i \geq 0
\end{aligned} \tag{36}$$

Therefore solving (36) we obtain the closed form expression for the values of the tilting measure

$$\Psi_i = \left(\sum_{i=0}^{N} \beta_i^{\rho} \gamma_i^{1-\rho}\right)^{-1} \left(\frac{\gamma_i}{\beta_i}\right)^{-\rho} \tag{37}$$

minimizing (30) and thus (18) and the right hand side of the (17). Please note that $f_0(\Psi_0, \ldots, \Psi_N)$ is a linear combination of convex function and thus $f_0(\Psi_0, \ldots, \Psi_N)$ is convex and so is the optimization problem (30). Thus the Karush-Kuhn-Tucker (KKT) conditions guarantee that the obtained solution indeed provides minima of the function $f_0(\Psi_0, \ldots, \Psi_N)$ Thus substituting (37) into (17) we obtain the following bound

$$P_e \leq \sum_{w=d}^{N} A_w^{\rho} \left(\sum_{i=0}^{N} (\beta_i(s,\rho))^{\rho} \gamma_i^{1-\rho}\right) \tag{38}$$

where β_i is given by (31) and γ_i is given by (33). The latter is to be minimized numerically by ρ and s subject to (14) similarly to the conventional Gallager bound [15].

V. Conclusion

Hereinafter a coded transmission via a channel that corresponds to a communication system using nonparametric receiver proposed in [6] has been considered. In particular the case of mismatched decoding has been addressed. For the case under consideration an upper bound on the error probability has been obtained by using the Duman-Salehi bounding technique and exact closed form expression for the normalized tilting measure minimizing the bound in question has been obtained.

Acknowledgment

The results obtained in section IV were obtained at the IITP RAS and financed by the Russian Science Foundation grant (project No. 14-50-00150).

References

[1] P. R. Pawlowski, A. Polydoros, *Adaptive Nonparametric Acquisition of FH-SS Signals in Jamming* Military Communications Conference 1986, vol.2, pp.31.6.1-31.6.8, 5-9 Oct. 1986.

[2] T.A. Gulliver, C. Felstead *Nonparametric diversity combining for fast frequency hopping*, Military Communications Conference 1995. MILCOM '95 vol.1, no., pp.60-64, 6 Nov 1995.

[3] R. Viswanathan, S. Gupta, *Nonparametric Receiver for FH-MFSK Mobile Radio*, IEEE Transactions on Communications, vol.33, no.2, pp.178-184, Feb 1985.

[4] K. Kondrashov , V. Afanassiev,*Ordered statistics decoding for semi-orthogonal linear block codes over random non-Gaussian channels.*, Proc. of the Thirteenth International Workshop on Algebraic and Combinatorial Coding Theory, Pomorie, Bulgaria, June 15–21, 192–196, 2012.

[5] A. Kreshchuk ,V. Potapov, *New Coded Modulation for the Frequency Hoping OFDMA System*, Thirteenth International Workshop on Algebraic and Combinational Coding Theory, Pomorje, Bulgaria, 15 – 21 June, 209–212, 2012.

[6] D. Osipov, *Reduced-Complexity Robust Detector in a DHA FH OFDMA System under Mixed Interference Multiple Access Communications*, Proc of 7th International Workshop on Multiple Access Communications, MACOM 2014, Halmstad, Sweden, August 27-28, 29-34, 2014.

[7] M.Z.Win, R.A.Scholtz, *Impulse Radio How It Works*, IEEE Communication Letter. Vol. 2, No. 2, 36-38 (1998)

[8] H.Arslan, Z.N.Chen, M.-G.Di Benedetto, *Ultra Wideband Wireless Communication.*, John Wiley & Sons (2006).

[9] J.A.C.Bingham, *Multicarrier Modulation for Data Transmission : An Idea Whose Time Has Come*, IEEE Commun. Mag., Vol. 28, No. 5, pp.5-14, May 1990.

[10] S.Weinstein, P.Ebert, *Data transmission by frequency-division multiplexing using the discrete Fourier transform.*, IEEE Trans. Commun., 19, 628–634 (1971).

[11] T. M. Cover, J. A. Thomas *Elements of Information Theory*,Jhon Whiley,1991.

[12] S. Wainberg, J. K. Wolf *Algebraic Decoding of Block Codes over a q-ary Input, Q-ary Output Channel, Q>q*, INFORMATION AND CONTROL 22, 232-247,1973.

[13] T. M. Duman and M. Salehi, *New performance bounds for turbo codes*, IEEE Trans. on Communications, vol. 46, no. 6, pp. 717-723, June 1998.

[14] T. M. Duman *Turbo codes and turbo coded modulation systems: Analysis and performance bounds*, Ph.D. dissertation, Elect. Comput. Eng. Dep., Northeastern University, Boston, MA USA, May 1998.

[15] R. G. Gallager, *A simple derivation of the coding theorem and some applications* IEEE Trans. on Information Theory, vol. 11, no. 1, pp. 3-18, January 1965.

[16] R. G. Gallager *Information Theory and Reliable Communications*, John Wiley, 1968.

[17] W.H. Kautz, R.C. Singleton, *Nonrandom Binary Superimposed Codes*, IEEE Trans. Inform. Theory, 4, 363-377, 1964.

[18] N. Merhav , G. Kaplan , A. Lapidoth , S. Shamai , *On information rates for mismatched decoders* IEEE Trans. Inf. Theory, vol. 40, no. 6, pp. 1953-1967, Nov., 1994.

[19] A. Ganti ,A. Lapidoth , E. Telatar , *Mismatched decoding revisited: General alphabets, channels with memory, and the wide-band limit* IEEE Trans. Inf. Theory, 46(7) pp.2315 - 2328,2000.

[20] J. Scarlett ,A. Martinez , A. Guillen i Fabregas , *Mismatched Decoding: Error Exponents, Second-Order Rates and Saddlepoint Approximations*, IEEE Transactions on Information Theory 60(5) pp. 2647-2666,2014.

[21] A. Martinez, A. Guillen i Fabregas, G. Caire, F. Willems, *Bit-interleaved coded modulation revisited: A mismatched decoding perspective*, IEEE International Symposium on Information Theory(ISIT) 2008 pp.2337-2341, 6-11 July 2008.

[22] A. Somekh-Baruch, *On coding schemes for channels with mismatched decoding*, Proceedings of the IEEE International Symposium on Information Theory(ISIT) 2013 , vol., no., pp.96-100, 7-12 July 2013.

[23] S.Shamai,I.Sason, *Variations on the Gallager bounds, connections, and applications*, IEEE Transactions on Information Theory, vol.48, no.12, pp.3029-3051, Dec 2002.

[24] I.Sason, I.Goldenberg, *Coding for Parallel Channels: Gallager Bounds and Applications to Turbo-Like Codes*, IEEE Transactions on Information Theory, vol.53, no.7, pp.2394-2428, July 2007.

[25] I.Sason,S.Shamai,D.Divsalar, *Tight exponential upper bounds on the ML decoding error probability of block codes over fully interleaved fading channels* IEEE Transactions on Communications, vol.51, no.8, pp.1296-1305, Aug. 2003.