The random coded modulation: performance and Euclidean distance spectrum evaluation

Hélio MAGALHÃES DE OLIVEIRA
Gérard BAUTTAIN

Abstract

We apply the random coding argument to coded modulation. The well-known union bound on the error probability of general signaling schemes is reviewed. The random coded modulation idea is introduced and a simple bound on the average performance of coded constellations is presented. A relationship between the union bound and the cut-off rate is established by introducing the concept of N-dimensional partial cut-off rate. We define finite theta series for bounded finite-dimensional constellations and their related transfer functions. Bounds on the block and symbol error probability based on the transfer function are derived. The discussion is then focused on the squared Euclidean distance distribution. The evaluation of such parameters as the first two moments (average squared distance and squared distance variance) is considered by other finite theta series or transfer function of the bounded signal set. The Euclidean distance spectra of a few multidimensional coded modulation schemes based on square-coups constituent two-dimensional constellations are presented.

Key words: Random coding, Error probability, Signal set, Code-distance, Quadrature modulation, M-ary modulation, Asymptotic behavior, Upper bound.

MODULATION CODÉE
AVEC CODAGE ALÉATOIRE:
PERFORMANCES ET ÉVALUATION
DU SPECTRE DE DISTANCE EUCLIDIENNE

Résumé

Le concept de codage aléatoire est appliqué à la modulation coded. On revient d'abord sur la borne par réunion, bien connue, de la probabilité d'erreur d'un système de communication quelconque. L'idee de codage aléatoire est introduite et une borne simple sur les performances moyennes des constellations codées est présentée. Une relation entre la borne par réunion et le dû de couper (cut-off rate) est obtenue en introduisant le concept de dû de couper partiel N-dimensionnel. La série theta et la fonction de transfert correspondante sont introduites pour une constellation bornée dans un espace à nombre de dimensions fini. Des bornes de la probabilité d'erreur par bloc et par symbole, basées sur cette fonction de transfert, sont étudiées. On examine ensuite la distribution des distances euclidiennes carrées. Des paramètres comme ses deux premiers moments (moyenne et variance du carré de la distance) sont évalués au moyen soit de la série Heun, soit de la fonction de transfert de la constellation. Les spectres des distances euclidiennes de quelques schémas de modulation codée multidimensionnelle basés sur des constellations plans carrées ou en forme de croix sont présentés. Leurs dûs de couper partiels sont calculés.
I. INTRODUCTION

One of the deepest ideas of coding theory is the random coding argument introduced by Shannon [1]. On the other hand, many efficient coding techniques have been developed much later using the famous idea of coded modulation by Ungaro [2]. In this paper we intend to amalgamate these two concepts: random coding is combined with 2-dimensional modulation e.g., quadrature amplitude modulation (QAM), in order to investigate the average performance of multidimensional coded constellations. We were in part motivated by the fact that some extremely regular codes appear to mimic the random code [3, 4].

An upper bound on the average error probability of random coded modulation is presented in section II. It derives from the union bound and leads to the definition of a partial cut-off rate, where partial means here that this quantity does take into account the actual number of dimensions. Following, we define finite length series for bounded multidimensional constellations and their related transfer functions. An interesting bound on the N-dimensional error probability of coded constellations is developed which is based on the transfer function and can also be applied to lattices. This bound presents a strong likeness with that on performance of convolutional codes earlier introduced by Viterbi [5], and it is also related to that on the trellis codes performance by Zehavi-Wolf [6].

In section III the distance spectra of multidimensional random coded constellations are investigated and examples are given. Parameters of the distance spectra like the mean and the variance of the squared Euclidean distance are evaluated in terms of either the finite theta series or the transfer function of the constellation. The distance distribution is calculated for some multidimensional constellations based on square/cross constituent 2-dimensional (2D) constellations. Further, we prove that an interesting phenomenon of normalized squared distance hardening occurs, which shows that the average squared distance becomes dominant on the error rate for high-dimensional constellations. This result confirms the fact that maximal distance is not significant for long codes [7-8]. It may also be related to our recent finding that lattices exist which can achieve the channel capacity [9].

We develop in section IV upper bounds on the symbol error probability still based on the transfer function of the signal set.

The behavior of very high-dimensional constellations is investigated in section V and we prove that almost all coded constellations become good as dimensionality increases. Besides, we show that almost all coded signal sets tend to become quasi-identical in a sense to be specified later. Section VI is devoted to Gallager-type bounds which extend the union bound on which the previous sections relied. Section VII provides a summary and concluding remarks.

II. BOUNDS ON THE ERROR PROBABILITY AND PARTIAL CUT-OFF RATE

The following theorem is well known, indeed, it can be found in many textbooks (e.g. [10-11]). We nevertheless give a proof of it. This will enable us to introduce the notations to be used in the remainder. Moreover, much of the results pertaining to random coded modulation will be obtained by slight variations with respect to this basic theorem so the detail of its proof is useful to understand the sequel.

The noise will be considered to be additive white Gaussian (AWGN) with a double-sided spectral power density \( \sigma^2 \) = \( N_0/2 \) (W/Hz).dan. Concerning notation, we adopt the symbol \( \equiv \) to denote equals by definition and \( \exp \) denote \( e^2 \).

**Theorem 1**: (Union bound) the error probability \( P_e(N) \) of an N-dimensional constellation with M equally likely signals over an AWGN channel is upper bounded by:

\[
P_e(N) \leq \frac{1}{2} \sum_{\delta_d = d_{max}}^{\infty} N_{2d} \left( \frac{1}{P(C)} \right)^{\frac{d}{2}}
\]

where \( \gamma = E_b/N_0 \) is the signal-to-noise ratio, \( P(C) \) is the average normalized power per 2-dimensional symbol, \( N_{2d}(\delta^2) \) is the average number of signal pairs with squared Euclidean distance \( \delta^2 \) in the constellation, and \( d_{max} \leq d \leq d_{max} \).

ANAL. TELECOMMUN., 47, no 1-4, 1992.
Proof: Let \( \{s_0, s_1, \ldots, s_{M-1}\} \) denote the signals of a (coded or uncoded) \( N \)-dimensional constellation. Since \( s_i \) is a point in an \( N \)-dimensional space, it can be written in terms of its coordinates, that is, \( s_i = (s_{i1}, s_{i2}, \ldots, s_{i(N-1)}, s_{iN}) \). The squared Euclidean distance between two signal points, \( s_j \) and \( s_k \), is given by:
\[
\|
\sum_{n=1}^{N} (s_{jn} - s_{kn})^2
\|
\]
On the other hand, the error probability \( P_e(N) \) of an \( N \)-dimensional signal set can be evaluated by:
\[
P_e(N) = \sum_{j=0}^{M-1} P(s_j)P(e|s_j) = \frac{1}{M} \sum_{j=0}^{M-1} P(s_j)
\]
where \( P(s_j) \) is the probability that \( s_j \) is transmitted, and \( P(e|s_j) \) is the error probability conditioned on \( s_j \) being transmitted. In order to calculate the conditional error probability given that the signal \( s_j \) was transmitted, we shall consider an index set \( T_j := \{ j \in 1, 2, \ldots, M \} \). Clearly, an error is made if the decoded signal \( \hat{s} \) belongs to the error set \( E_j := \bigcup_{i \notin T_j} E_i \).

Hence,
\[
P(e|s_j) = P(\text{dec} = \hat{s} \in E_j) = P(\text{dec} = \hat{s} \in \bigcup_{i \notin T_j} E_i),
\]
where \( \text{dec} \) denotes the event decoding the signal \( s_j \). For simplicity, we dropped out given that \( s_j \) was transmitted.

We can make a partition of the index set \( T_j \) according to the distance between \( s_j \) and \( s_k \), and the elements whose index belongs to the set \( T_j \). Let
\[
T'_j = \{ j \in 0, 1, 2, \ldots, M-1 | d_j(s_j, s_{j'}) = d' \},
\]
be such a partition of \( T_j \). Then \( T'_j \subset T_j \) and \( \bigcup_{d' = 0}^{d_{max}} T'_j = T_j \). Furthermore, \( T'_j \cap T'_{j'} = \emptyset \) for \( j \neq j' \). The empty set. This partition over the index set induces a partition over \( E_i \) which is given by the following subsets:
\[
E'_j := \{ j \in E_j | d_j(s_j, s_{j'}) = d' \} = \bigcup_{d' = 0}^{d_{max}} T'_j
\]
in such a way that \( E'_j \cap E'_{j'} = \emptyset \) for \( j \neq j' \).

Applying this partition, the error event \( E_i \) can be rewritten as \( E_i = \bigcup_{d=0}^{d_{max}} \bigcup_{j \in T'_j} E'_j \).

Consequently, it follows from (3) that:
\[
P(e|s_j) = P\left( \bigcup_{d=0}^{d_{max}} \bigcup_{j \in T'_j} E'_j \right).
\]
where \( E_j \rightarrow E_j \) denotes the event decoding \( s_j \) when \( s_j \) is sent and \( s_j \) is considered the only alternative. Applying the union bound (see e.g. [11], p. 60) to the above equation results in:
\[
P(e|s_j) \leq \sum_{d=0}^{d_{max}} \sum_{j \in T'_j} P(s_j \rightarrow E_j).
\]
where \( P(s_j \rightarrow E_j) \) denotes the probability of the event \( s_j \rightarrow E_j \). It should be pointed out that \( P(s_j \rightarrow E_j) \) for \( j \in T_j' \) does not depend on \( j \) but only on \( d' = d_j(s_j, s_{j'}) \).

Moreover, it can easily be shown that:
\[
P(s_j \rightarrow E_j) = \frac{1}{2} \text{erfc} \left( \frac{d_j(s_j, s_{j'})}{2 \sigma^2} \right),
\]
where the second equality results from the equality \( \gamma = \frac{R_s}{N_0} = P(C)/2 \sigma^2 \). To simplify notation, the pairwise error probability of any two points \( s_i, s_j \) with square distance \( d' \) will be denoted throughout by \( e(d') := 1/2 \text{erfc} \left( \sqrt{d'/2 \sigma^2} \right) \).

Now, let \( N_{\gamma}(d') := \#(T'_j) \) denote the cardinality of the set \( T'_j \) i.e., the number of signals at a squared Euclidean distance \( d' \) from \( s_j \). It results from equalities (3) and (5) that:
\[
P(e|s_j) \leq \sum_{d' = 0}^{d_{max}} N_{\gamma}(d') e(d').
\]

Therefore, inserting into (2) we find:
\[
P_e(N) \leq \frac{1}{M} \sum_{j=0}^{M-1} \sum_{d' = 0}^{d_{max}} N_{\gamma}(d') e(d').
\]

Finally the proof is completed by changing the order of the summations and defining:
\[
N_{\gamma}(d') := \frac{1}{M} \sum_{j=0}^{M-1} N_{\gamma}(d'')(d'),
\]
which represents the average number of points at a squared distance \( d' \) in the signal constellation, etc.

Indeed, this theorem applies to both coded and uncoded constellations. As a simple example, we consider the uncoded M-ary 2-dimensional constellation corresponding to narrow band case. Assuming a high signal-to-noise ratio \( \left( \frac{P(C)}{2 \Delta^2} \gg 1 \right) \), \( d_{min} = 2 \) and renumbering that \( N_{\gamma}(d_{min}) = 4 \) (for large \( M \)) and \( P(C) = P_{\text{c}}(M - 1)/2 \), we derive the well known approximation of the symbol error probability:
\[
P_e(2) \approx 1/2 \text{erfc} \left( \frac{d_{min}}{2 \sqrt{2 \Delta^2}} \right) \approx 2 \text{erfc} \left( \frac{\sqrt{2 \Delta^2}}{M - 1} \right).
\]

In the following, we apply the famous random coding argument to coded modulation systems.

Random coding.

Let us consider to begin with, an \( N \)-dimensional uncoded constellation (\( N \) even) created by concatenating \( N/2 \) identical 2-dimensional constellations whose
alphabet size is \( q \), i.e., by taking the Cartesian product of a constituent 2-dimensional constellation with itself \( N/2 \) times. From its total number of uncoded signals \( M_{\text{unc}} = q^{N/2} \), we build an \( N \)-dimensional random code constellation by a random (uniform) choice of \( M_{\text{unc}} = q^{N/2} \) points \( M_{\text{unc}} \leq M_{\text{cod}} \); we moreover assume that any point cannot become more than once, resulting in a rate of:

\[
R = \frac{1}{N} \log_2 M_{\text{unc}} = 1/2 \frac{N}{N} \log_2 q \text{ bits/dim.}
\]

or a normalized rate [12, 13] of \( R = (K/N) \log_2 q \) bits per 2-dimensional symbol.

The way of introducing redundancy by increasing the alphabet size is the same as in Ungar's work since an alphabet of \( M_{\text{unc}} \) signals is used whereas only \( M_{\text{cod}} \) of them may be transmitted. This coding technique may be seen as Block Coded Modulation (BCM) where a quasi \((N, K)\) block code is picked at random. Additionally, it should be pointed out that this process deals with the ensemble of all codes and that no linearity assumption is made. By abuse of language, a coded constellation is referred to indistinctly as a code. Henceforth, we denote by \( \mathcal{C} \) an \( N \)-dimensional constellation while its constituent 2-dimensional constellation is denoted by \( \mathcal{C}_0 \). Also, the minimal distance between the constituent 2-dimensional constellation is denoted by \( d_0 \).

Now the idea of random coded modulation is used in connection with Theorem 1 (union bound) in order to estimate the average error probability of the set of all codes. This is made according to the following theorem:

**Theorem 2** (random coding) : let \( P^*_e(N) \) be the average error probability of a random coded modulation with an \( N \)-dimensional signal set. Then the following inequality holds:

\[
P^*_e(N) \leq \frac{1}{2} \sum_{d=0}^{d_{\text{max}}} N_d(q)(d) e(q,d),
\]

where \( N_d(q)(d) \) denotes an average unnormalized squared distance profile.

**Proof.** We deal with \( L \) codes chosen by random means as suggested earlier. Applying the preceding theorem to each code, the following set of inequities results:

\[
P^*_{e}(N) \leq \frac{1}{2} \sum_{d=0}^{d_{\text{max}}} \bar{N}_d(q)(d) e(q,d) 1 \leq l \leq L,
\]

the superscript \( \# \) standing for the \( l \)th code. We note that all codes have not necessarily the same minimal distance. We may use here the minimum of the minimal distance among all conceivable distinct codes, namely \( d_{\text{max}} \). Similarly, they have not necessarily the same maximum distance, so \( d_{\text{max}} \) in (13) should be understood as the largest possible maximum distance in the set of codes. Furthermore, there are at most \( M_{\text{cod}}(M_{\text{cod}} - 1) \) terms in each sum. The mean error probability of random codes can be expressed as:

\[
P_e(N) = \lim_{L \to \infty} \frac{1}{L} \sum_{l=1}^{L} P^*_{e}(N).
\]

Thus, it results from the set of inequalities (15) that:

\[
P_e(N) \leq \sum_{d=0}^{d_{\text{max}}} \bar{N}_d(q)(d) e(q,d)
\]

where the unnormalized squared distance profile of the random code is:

\[
\bar{N}_d(q)(d) = \lim_{L \to \infty} \frac{1}{L} \sum_{l=1}^{L} N_d(q)(d).
\]

QED.

It should be stressed that \( \bar{N}_d(q)(d) \) is not the squared distance profile of any code of the random ensemble but rather the average profile resulting from taking into account all codes. Now we introduce a normalized squared distance profile referred to as the squared distance distribution.

**Definition 1.** The squared distance distribution (distance spectrum) of the uncodded constellation and the average squared distance distribution of the random coded constellation are respectively:

\[
P(d) = \frac{N_d(q)(d)}{M_{\text{cod}}}, \quad \overline{P}(d) = \frac{N_d(q)(d)}{M_{\text{unc}}},
\]

\( N_d(q)(d) \) standing for the distance profile of the uncodded \( N \)-dimensional constellation.

Evidently, one has \( \sum d^2 \overline{P}(d) = \sum d^2 P(d) = 1 \). The way to evaluate \( N_d(q)(d) \), and consequently to estimate the average performance, is quite simple.

**Assumption 1.** We consider that the \( N \)-dimensional random coded constellation has the same distribution as the uncodded \( N \)-dimensional signal set, that is to say, \( P(d) = \overline{P}(d) \).

An immediate consequence of this assumption is:

\[
\overline{P}(d) = e^{-d^2/(2\log_2 q - R)} N_d(q)(d)
\]

We intend now to show that this union bound is closely linked to a cut-off rate \( R_0 \) so that these bounds are useless both near and above \( R_0 \). More general Gallager-type bounds will be introduced in section VI.

**Theorem 3** (partial cut-off rate) : the average error probability of an \( N \)-dimensional random coded signal set is exponentially upper bounded by:

\[
P_e(N) \leq \exp\left\{-N[R_0(N) - R]\right\},
\]

where \( R_0(N) = -1/N \log_2 P(e(q,d)) \) is referred to as the \( N \)-dimensional partial cut-off rate.

**Proof.** Of course, the average error probability is upper bounded according to Theorem 2 which can be rewritten as:

\[
P_e(N) \leq \exp\left\{-N\left[1/2 \log_2 q - R + \sum_{d=0}^{d_{\text{max}}} N_d(q)(d) e(q,d)\right]\right\}
\]

\[\text{ANN. TELECOMM., 47, 3-4, 1995.}\]
Besides, a reliability function, namely \( \mathbb{E}(R, R_0) \), is defined as:

\[
\mathbb{E}(R, R_0) = -\frac{1}{N} \log_2 \left( \sum_{d \in d_{\text{set}}} N_d(d^2) \mathbb{P}(d^2) \right),
\]

so inequality (38) can be expressed in the form:

\[
\overline{R}(N) \leq \mathbb{E}(R, R_0).
\]

Remembering that \( d_{\text{set}} = \exp_{d}(NR) \), \( R \) is bit/dimension, we put:

\[
\mathbb{E}(R, R_0) = -\frac{1}{N} \log_2 \left( \sum_{d \in d_{\text{set}}} N_d(d^2) \mathbb{P}(d^2) \right) \cdot \mathbb{P}(d^2) \]

which yields:

\[
\mathbb{E}(R, R_0) = -\frac{1}{N} \log_2 \left( \sum_{d \in d_{\text{set}}} \mathbb{P}(d^2) \mathbb{P}(d) \right) - R.
\]

We now define the \( N \)-dimensional partial cut-off rate as:

\[
\overline{R}_0(N) := \mathbb{E}(R, R_0) = -\frac{1}{N} \log_2 \left( \sum_{d \in d_{\text{set}}} \mathbb{P}(d^2) \mathbb{P}(d) \right) - R,
\]

where the expectation is taken with respect to the distribution \( \{\mathbb{P}(d^2)\} \), the term \( \mathbb{P}(0) \) not being taken into account. This leads to \( \mathbb{E}(R, R_0) = \overline{R}_0(N) - R \) and the proof follows. Q.E.D.

**Definition 2.** The cut-off rate, \( \overline{R}_0 \), is related to \( \overline{R}_0(N) \) by \( \overline{R}_0 := \lim_{N \to \infty} \overline{R}_0(N) \), where both \( \overline{R}_0(N) \) and \( \overline{R}_0 \) are expressed in bits per dimension.

It should be pointed out that Theorem 3 can be used for a particular code (fixed constellation) as well as for the random code: it suffices to apply Theorem 1 instead of Theorem 2 in the derivation. Hence, an interesting way to compare different codes is to calculate their respective partial cut-off rates, i.e. for the \( i \)-th code:

\[
\overline{R}_0^i(N) := -\frac{1}{N} \log_2 \left( \sum_{d \in d_{\text{set}}} p^i(d^2) \mathbb{P}(d^2) \right).
\]

The evaluation of such partial rates involves knowing the complete distance spectrum which is, to some extent, hard to obtain. Hence, we consider two special rates related to the partial cut-off rate:

\[
\overline{R}_0^i(N)_{\text{max}} := -\frac{1}{N} \log_2 \left( \sum_{d \in d_{\text{set}}} p^i(d^2) \right);
\]

the subscript \( \text{max} \) standing here for max or mean. In the latter case we define \( d_{\text{max}} := \left[ \mathbb{E}(d^2) \right]^{1/2} \), the rate \( \overline{R}_0^i(N) \) can be estimated by the following theorem:

**Theorem 4.** For any dimensionality \( N \), the partial cut-off rate satisfies the following bounds:

\[
\overline{R}_0(N)_{\text{min}} \leq \overline{R}_0^i(N) \leq \overline{R}_0(N)_{\text{max}}.
\]

Sketch of the proof. To prove this, observe that the first bound is an immediate consequence of \( \mathbb{P}(d) \) being monotonically decreasing with \( d^2 \) and that the second one can be proved by Jensen's inequality. Q.E.D.

When \( N \) is large enough, the asymptotic behaviour of the \( \overline{R}_0(N) \) bounds expressed in rate/dimension is given by:

\[
\overline{R}_0(N)_{\text{min}} \sim \frac{d_{\text{min}}^2}{N} \mathbb{P}(d^2),
\]

and \( \overline{R}_0(N)_{\text{max}} \sim \frac{d_{\text{max}}^2}{N} \mathbb{P}(d^2) \).

In order to compare the performance of un-coded and coded constellations, we must use the same criterion for measuring the error rate. The probability of error of un-coded constellations is currently expressed per 2-dimensional symbol, so the following lemma could be useful.

**Lemma 5.** Assuming \( N \) an even integer, the 2-dimensional error probability of symmetrical constellations (like lattice or randomly coded ones) is related to the \( N \)-dimensional error probability by \( \frac{N}{2} P_e(N) \leq P_e(2) \leq P_e(N) \).

**Proof.** The \( N \)-dimensional error rate is expressed as:

\[
P_e(N) = \sum_{|\mathbf{e}|=2} \mathbb{P}(\mathbf{e}) = \mathbb{P}(\mathbf{e}) \sum_{i=1}^{N/2} \left( \mathbb{P}(\mathbf{e} \text{ error per } i \text{th } 2D \text{ symbol}) \right) \]

where we have used both the union bound and the constellation symmetry. Furthermore, the 2-dimensional error probability can be evaluated by:

\[
P_e(2) = \frac{\text{number of errors per } 2D}{N/2},
\]

the word error standing here for erroneous 2D symbol. The average number of errors can be expressed as:

\[
P_e(2) = \sum_{i=1}^{N/2} 2 \mathbb{P}(i \text{ errors per } 2D) \]

Additionally,

\[
P_e(N) = \sum_{i=1}^{N/2} \left( \sum_{i=1}^{N/2} 2 \mathbb{P}(i \text{ errors per } 2D) \right) = \sum_{i=1}^{N/2} \mathbb{P}(i \text{ errors per } 2D).
\]

The proof is completed by combining (29) and (30). Q.E.D.
Finite theta series.

It has long been recognized that good multidimensional coded constellations can be obtained using dense finite-dimensional lattices. A lattice code $\Omega$ is defined here as a subset of a lattice $\Lambda$ (or a translate of it) whose points lie within a bounded region [14]. Clearly, the dimension of a lattice code is the same as that of the taken lattice; that is, $\dim \Omega = \dim \Lambda$. The $\Theta$ series of an (unbounded) lattice $\Lambda$ is given by [14, 15]:

$$
\Theta_{\Omega}(z) = \sum_{d=0}^{\infty} N(d^2) \frac{z^d}{d!},
$$

where $w = \exp(z/e)$ and $N(d^2)$ is the number of lattice points at a squared Euclidean distance $d^2$ from one of its points, arbitrarily chosen as the origin. We suggest now to define the finite theta series of a bounded lattice $\Lambda$ according to:

**Definition 2** (finite theta series):

$$
\Theta_{\Omega}(z) = \sum_{d=0}^{\infty} \frac{N_{\Omega}(d^2) \, w^d}{d!},
$$

where $w = \exp(z/e)$ and $N_{\Omega}(d^2)$ denotes the average number of signal pairs with squared Euclidean distance $d^2$ within $\Omega$, as defined in Theorem 1.

The total number of constellation points is $M = |\Omega|$, the cardinality of the set $\Omega$. This function contains all the information concerning the average neighborhood in the constellation. We mention that finite theta series, as defined above, can be employed not only for lattice codes but also for any bounded signal set.

**Proposition 6.** The finite theta series verifies the properties:

$$
\Theta_{\Omega}(z)|_{w=1} = \sum_{d=0}^{\infty} N_{\Omega}(d^2) = M = |\Omega|
$$

and

$$
\lim_{M \to \infty} \Theta_{\Omega}(z) = \Theta_{\Lambda}(z),
$$

provided the region which defines $\Omega$ expands equally enough in all dimensions, a condition generally met due to symmetry constraints.

Indeed, the bounded lattice $\Omega$ approaches the lattice $\Lambda$ when $M$ increases indefinitely. It can be seen that $N_{\Lambda}(d^2) \to N(d^2)$ then becomes independent of $i$ as $M \to \infty$, and that $N_{\Lambda}(d^2) \to N(d^2)$.

A bound on the performance of lattice coded systems can now be developed which is based on theta series and Theorem 1. In fact, we use transfer functions which result from the theta series (either infinite or finite) by giving its indeterminate a real value. A somewhat strong analogy is found with bounds on the convolutional codes performance by Viterbi [5]. This bound is attributed to Porzycki [14, p. 72].

**Lemma 7.** The $N$-dimensional error rate of a coded constellation $\Omega$ taken from a lattice $\Lambda$ is upper bounded by:

$$
P_e(N)|_{\Omega} \leq \frac{1}{2} T(D; \Lambda) \Big|_{d=D_0} \leq \frac{1}{2} T(D; \Lambda) \Big|_{d=D_1},
$$

where $T(D; \cdot)$ are the transfer functions defined below and $D_0 := \exp(-d_0^2)$, $d_0$ being the minimal distance of the uncodded constituent 2D constellation.

**Proof.** Consider the $\theta$-transfer function, $T(D; \cdot) = \sum_{d=0}^{\infty} N(d^2) \frac{D^d}{d!} = \Theta_{\Lambda}(z)|_{w=D^2/d!} - 1$.

Similarly, the $\Omega$-transfer function of a lattice code $\Omega$ is defined by:

$$
T(D; \Omega) := \sum_{d=0}^{\infty} N_{\Omega}(d^2) \frac{D^d}{d!} = \Theta_{\Omega}(z)|_{w=D^2/d!} - 1.
$$

On the other hand, the exponential bound on the erfc($\cdot$) function ($\Theta$) results in:

$$
\text{erfc}(\sqrt{\eta d}) \leq \exp(-\eta d) = D^d/d!,
$$

Applying this inequality to the upper bound of Theorem 1 readily results in the (weaker) bound:

$$
P_e(N)|_{\Omega} \leq \frac{1}{2} \sum_{d=0}^{\infty} N_{\Omega}(d^2) \frac{D^d}{d!} \Big|_{w=D^2/d!} = \frac{1}{2} T(D; \Omega) \Big|_{w=D^2/d!}.
$$

Furthermore, since $N_{\Omega}(d^2) \leq N(d^2)$, it is easily verified that ($\Theta$) $T(D; \Omega) \leq T(D; \Lambda)$, completing the proof.

The bound on the error probability based on the $\theta$-transfer function is less tight than the one based on the $\Omega$-transfer function, but it is quite straightforward to obtain. Moreover, it can be improved in the following way: a truncated transfer function $T(D; [\Lambda])$ of a lattice $\Lambda$ is defined as:

$$
T(D; [\Lambda]) := \sum_{d=0}^{\text{trunc}} \frac{N(d^2)}{d!} \frac{D^d}{d!} = \Theta_{\Lambda}(z)|_{w=D^2/d!} - 1,
$$

where trunc indicates that the $\Theta$ series written as a polynomial of $w$ has been truncated to a finite degree but no
less than $\frac{d_{\text{max}}^2}{N}$. Thus, possible convergence problems are avoided. In the cases where $d_{\text{min}}$ is unknown, the theta series may be truncated to the minimal degree greater than or equal to $S/N/2 \max \{a_i^2\}$, where $S = \Delta(\Omega_2)$ denotes the set of all conceivable squared Euclidean distances between two points in a constituent 2D constellation.

As a consequence of (36):

\[
T(D;\Omega)_{\phi \in \Omega_2} \leq T(D;\Lambda)_{\phi \in \Lambda_2} \leq T(D;\Omega_2)
\]

III. EUCLIDEAN DISTANCE SPECTRUM OF RANDOM CODING

Besides the minimum squared distance, we are especially interested in the first and second moments of the Euclidean distance distribution $P(d^2)$, denoted respectively by $E(d^2)$ and $E(d^4)$? The next lemma presents some interesting properties relating these parameters to the theta series of a constellation.

Lemma 8. The average and the second moment of the squared Euclidean distance of a constellation $\Omega$, are given by:

\[
E(d^2) = \frac{1}{|\Omega|} \sum_{\phi \in \Omega} \Phi_{d^2}(\phi)\quad \text{and} \quad E(d^4) = \frac{1}{|\Omega|} \sum_{\phi \in \Omega} \Phi_{d^4}(\phi)
\]

Proof. First, we notice that:

\[
\Phi_{d^2}(\phi) = \sum_{\phi_1 = \phi_2} N_{\phi}(\phi_2)\phi_2 = \left(\Phi_{d^2}(\phi)\right)_{\phi = \phi_1} \sum_{\phi_1 = \phi_2} N_{\phi}(\phi_2)\phi_2
\]

and the first part follows.

Furthermore:

\[
\sum_{\phi_1 = \phi_2} N_{\phi}(\phi_2)\phi_2 = \sum_{\phi_1 = \phi_2} \sum_{d = d_{\text{max}}} N_{\phi}(\phi_2)\phi_2
\]

which can be arranged as:

\[
\sum_{\phi_1 = \phi_2} \sum_{d = d_{\text{max}}} N_{\phi}(\phi_2)\phi_2 = \left(\sum_{d = d_{\text{max}}} N_{\phi}(\phi_2)\phi_2\right)_{\phi = \phi_1}
\]

proving the second part, quo.

Lemma 8 can be rewritten in terms of the transfer function of a bounded lattice as:

Cardy. 9. The first two moments of the distance distribution $P(d^2)$ are given by the formulas:

\[
E(d^2) = \int_0^\infty \frac{d^2 T(D;\Omega)}{dD^2}
\]

and

\[
E(d^4) = \int_0^\infty \frac{d^4 T(D;\Omega)}{dD^4}
\]

Besides transfer functions we are also concerned with related functions given by:

Definition 4 (complete transfer function): the complete transfer function, namely $T_N(D)$, of an $N$-dimensional lattice $\Lambda$ (or an $N$-dimensional bounded constellation $\Omega$) is defined as:

\[
T_N(D) \equiv 1 + T(D;\Omega) - \Theta_N(x)_{x \in \Omega_{N/2}}
\]

where $\Omega$ stands for $\Lambda$ or $\Omega$.

It is clear heterofom if the complete transfer function concerns either a lattice or a bounded constellation.

For the multidimensional uncoded constellation specified earlier, it can be seen by a combinatorial approach of the generating function [16] that:

\[
T_N(D) = \frac{1}{N/2} D_N(D)
\]

Taking random coding into account, the complete transfer function $T_N(D)$ of such a coded multidimensional constellation is given by:

\[
T_N(D) = T_N(D) - \frac{1}{N/2} D_N(D) = q^{-N/2} \log^N T_N(D),
\]

which is an immediate consequence of Assumption 1. Also, obviously, $T_N(1) = [D_N(1)]^{N/2} = q^{N/2}$ and $T_N(D) = e^{N/2}$, as expected (see Proposition 6).

We are interested in the squared distance distribution (spectrum) of the random-coded modulation, $P(d^2)$, so we will use Corollary 9. Initially, we calculate the derivatives of the complete transfer function. Furthermore, for the sake of simplicity, we denote $dT(D)/dD$ by $T_N(D)$:

\[
\left(\frac{d}{dD}\right)^2 T_N(D) = \frac{N}{2} \left[T_N(D)\right]^{N/2-1} T_N(D)
\]

and

\[
\left(\frac{d}{dD}\right)^4 T_N(D) = \frac{N}{2} \left[T_N(D)\right]^{N/2-1} \left[T_N(D)\right]^{N/2-2}
\]

Now, we are able to establish the next result.

Theorem 10. The mean and variance of the squared Euclidean distance between two points of a multidimensional random coded constellation are, respectively:

\[
E(d^2) = \frac{N}{2} T_N(1)
\]

and

\[
\sigma^2(d^2) = \frac{N}{2} \left[T_N(1) + T_N(1)^{N/2} - T_N(1)^{N/2-1}\right]^2
\]

ANN. TELECOMM. 42, no. 3-4, 1992
Proof. By Corollary 9, we have the (second equality results from Assumption 1):

\[
E(n^2) = \frac{\sigma_n^2}{L^V/2} \left[ \frac{d(T_x(D))}{dD} \right]_{D=1} = \frac{\sigma_n^2}{L^V/2} \left[ \frac{d(T_x(D))}{dD} \right]_{D=1}
\]

by (44) and remembering that \( T_x(1) = 0 \). It follows:

\[
E(n^2) = \frac{\sigma_n^2}{L^V/2} \left[ T_x(1)^N \right] \frac{T_x(1)}{2T_x(1)} - \frac{N}{2} T_x(1)
\]

Substituting (43.1) in (44) and remembering that \( T_x(1) = 0 \). It follows:

\[
E(n^2) = \frac{\sigma_n^2}{L^V/2} \left[ T_x(1)^N \right] \frac{T_x(1)}{2T_x(1)} - \frac{N}{2} T_x(1)
\]

The second moment is given by:

\[
E(n^2) = \frac{\sigma_n^2}{L^V/2} \left[ \frac{d(T_x(D))}{dD} \right]_{D=1} + \frac{d(T_x(D))}{dD} \left[ \frac{d(T_x(D))}{dD} \right]_{D=1}
\]

Combining (43.1), (43.2) and the above equation, we find out after further simplifications:

\[
E(n^2) = \frac{\sigma_n^2}{L^V/2} \left[ T_x(1)^N \right] \frac{T_x(1)}{2T_x(1)} - \frac{N}{2} T_x(1)
\]

We get the desired form of the squared distance variance by inserting (47) and (45) into \( \tau^2(D^2) = E(n^2) - E(n^2) \), which completes the proof, Q.E.D.

Examples.

We intend to apply the above results to a number of multidimensional random coded constellations. To begin with, let us consider a few constituent 2-dimensional constellations with alphabet size \( q = 4, 8, 16, \) and 32 for example (Fig. 1). Next, we evaluate the finite theta series as well as the transfer function of such signal sets (cf. Appendix A). Their respective transfer functions are given below:

\[
T_x(2D) = 1 + 3D + 1D^2
\]

\[
T_x(2D) = 1 + 3D + 1D^2 + 1.0D^3 + 2.0D^4 + 0.5D^5
\]

\[
T_x(2D) = 1 + 3D + 1D^2 + 1.0D^3 + 1.5D^4 + 1.5D^5 + 0.5D^6
\]

\[
T_x(2D) = 1 + 3D + 1D^2 + 2.5D^3 + 3.5D^4 + 1.5D^5 + 1.75D^6 + 1.75D^7 + 2.5D^8 + 2.5D^9 + 1.75D^{10} + 1.75D^{11} + 2.5D^{12} + 2.5D^{13} + 1.5D^{14} + 1.5D^{15} + 1.5D^{16} + 0.75D^{17} + 0.75D^{18} + 0.5D^{19} + 0.25D^{20} + 0.25D^{21}.
\]

A simple, but striking example of our approach makes allowance for an \( N \)-dimensional coded constellation with constituent 2-dimensional signal set of size \( |R_x| = 16 \). Then:

\[
T_x(2D) = 1 + 3D + 1D^2 + 1.0D^3 + 2.5D^4 + 0.5D^5 + 0.25D^6 + 0.25D^7 + 0.125D^8 + 0.125D^9 + 0.0625D^{10} + 0.0625D^{11} + 0.03125D^{12} + 0.03125D^{13} + 0.015625D^{14} + 0.015625D^{15} + 0.0078125D^{16} + 0.0078125D^{17} + 0.00390625D^{18} + 0.00390625D^{19} + 0.001953125D^{20} + 0.001953125D^{21}.
\]

Fig. 1. M-point quasi constituent constellations and their respective 2-dimensional transfer functions.

Ann. Telecommun., 47, nº 3-4, 1992
The random coded modulation

\[ T_2(D) = 1 + 3D^4 + 2.25D^8 + 2D^8 + 1.5D^{12} + 1.5D^{16} + 0.25D^{20}, \]

\[ T_2(D) = 3 + 4.5D^4 + 8D^8 + 13D^{12} + 16D^{16} + 9D^{20} + 11D^{24} + 12D^{28} + 4.5D^{32}, \]

\[ T_2(D) = 4.5 + 9D^4 + 19D^8 + 20D^{12} + 12D^{16} + 13D^{20} + 15D^{24} + 76.5D^{32}, \]

so that \( T_2(1) = 16 \), \( T_2(1) = 80 \) and \( T_2(1) = 584 \).

Hence, assuming \( d_0 = 0 \), the distance spectrum is roughly characterized by:

\[ E(d^2) = 5N/2 \] and \( s^2(d^2) = \frac{15}{N/2} \).

Plots of the distance spectrum for a few dimensions are exhibited in Figure 2. We expect that these distributions approach a Gaussian one as the dimension increases, as a consequence of the central limit theorem. In order to check this trend, let us consider \( N/2 \) 2-dimensional vectors distributed on a 16-point constituent constellation, resulting in 10 points per row, and then the number of dimensions increases.

![Evolution of normalized squared Euclidean distance spectrum of random coded constellation based on a 16-point modulation](image)

Evolution de genre de distance euclidienne avec norme de constellations codées d'octet à partir de certaines constellations avec \( N/2 \) 16 points selon le nombre de dimensions exigé.

It can be seen, as expected, that the cross-entropy decreases as dimension increases.

When the dimensionality increases, there is some kind of squared distance concentration near the average distance, since the squared distance spread (relative to the mean) tends asymptotically to zero. More precisely:

**Theorem 11** (normalized distance hard-distributing), the squared Euclidean distance of multidimensional random coded constellations verifies:

\[ \lim_{N \to \infty} 2M \frac{e^2 - E(e^2)}{N} = 2x^1 \]

where \( x^1 \) denotes the limit in probability.

**Proof.** We first define a variable \( e \), namely the 2-dimensional distance (or normalized distance), according to \( e^2 = d^2 / (N/2) \). Clearly, \( 0 \leq e^2 \leq \log_2 T_2(1) \). It follows from Theorem 10 that:

\[ \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(-x^2/2) \, dx, \]

\[ \text{normal cumulative distribution function.} \]

Let \( p_1 \) be the probability of the 1st class calculated from the actual squared distance spectrum. We computed the Kullback-Leibler divergence (or cross-entropy) of \( p_1 \) with respect to \( p_2 \), namely:

\[ \sum p_1 \log \left( \frac{p_1}{p_2} \right), \]

which measures in a certain sense the vicinity between two distributions [17]. For the random coded constellation based on a 16-point constituent constellation, with \( K = \deg T_2(D) = 18 \), the results are given in the Table 1.

<table>
<thead>
<tr>
<th>N</th>
<th>Number of classes</th>
<th>Cross-entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>20</td>
<td>0.086</td>
</tr>
<tr>
<td>4</td>
<td>36</td>
<td>0.242</td>
</tr>
<tr>
<td>8</td>
<td>74</td>
<td>0.088</td>
</tr>
<tr>
<td>10</td>
<td>140</td>
<td>0.017</td>
</tr>
</tbody>
</table>

It can be seen, as expected, that the cross-entropy decreases as dimension increases.

Thus, there exist very few long coded constellations for which the normalized squared distance exceeds the hardening distance by an appreciable margin. A similar behavior has been found for the minimum Hamming distance in the ensemble of binary codes having a fixed ratio \( K/N \) [18].
where $d$ is the minimum Hamming distance, and $p$ is the smallest solution of $1 - H_2(p) = K/N$, where $H_2(p) := - \log_2 p - (1 - p) \log_2 (1 - p)$ is the binary entropy function. This asserts that there exist very few long codes lying far from the Gilbert-Varshamov bound.

The distance hardening phenomenon which occurs in the Euclidean space thus appears as the same kind of behaviour as that of the binary codes with respect to the Gilbert-Varshamov bound.

It can be seen ex post facto that the average squared distance becomes dominant on the error rate performance as dimensionality increases. Thus, the minimum distance is not significant for long codes. This marks the recent criticism of the minimum distance criterion by Battail [7-8].

IV. MORE ON THE ERROR PROBABILITY BOUNDS

Naturally, the upper bounds so far introduced concern the $N$-dimensional (block) error probability. We are also interested in obtaining upper bounds on the random coded-modulation performance on the basis of the 2D error probability criterion (symbol error rate). To this end we introduce a second variable $L$ in the complex transfer function, namely.

**Definition 5** (two-parameter transfer function)

$T(D, L; \Omega) := \sum_{d_{i} = d_{i_{0}}}^{d_{i_{0} + 1}} \sum_{l = 0}^{L} N_{\omega}(d_{i}, l) D^{d_{i}}L^{l}$

where $N_{\omega}(d_{i}, l)$ is the average number of signal pairs with both squared distance $d_{i}$ and $l$ different 2D symbols.

An interesting property of this new function is that $T(D, 1; \Omega) = T(D, \Omega)$, since:

$\sum_{l = 0}^{L} N_{\omega}(d_{i}, l) = N_{\omega}(d_{i})$.

**Theorem 12.** The 2-dimensional error probability of a multidimensional constellation $\Omega$ is upper bounded by:

$P_{e}(2) \leq \frac{1}{N} \frac{DT(D, L; \Omega)}{D^{L-1} L!}$

where $T(D, L; \Omega)$ is the transfer function of $\Omega$ according to Definition 5 and $D_{l}$ was introduced in Lemma 7.

**Proof.** We shall follow quite closely the proof of Theorem 1. Let $L$ be the number of erroneous symbols after decoding. The average number of faulty 2-dimensional symbols per $N$-dimensional signal is:

$(50) \quad E(L) = \sum_{l = 0}^{L} P(L_x) E(L_{x_1}) - \frac{1}{M} \sum_{l = 0}^{L} E(L_{x_1})$.

We define the following sets, $E_{i,j} := \{ l \mid x_1 \} \mid l \mid x_2 \}$, $0 \leq i \leq M - 1$, $0 \leq j \leq N/2$, where $d_{H}(x, y)$ is the Hamming distance between the signals $x$ and $y$ in terms of their constituent 2D symbols. We can therefore write:

$(51) \quad E(L_{x_1}) = \sum_{l = 0}^{L} U(d_{x_1} \in E_{i,j})$.

Defining the associated index sets as $T_{1,i} := \{ j \mid 0 \leq 0, 1, \ldots, M - 1 \}$, we now have:

$(52) \quad P(d_{x_1} \in E_{i,j}) = P\left( \bigcup_{j \in T_{1,i}} \{ d_{x_1} \} \right)$.

Again, we can build a partition on the index set according to:

$(53) \quad T_{2,i} = \{ j \mid 0, 1, \ldots, M - 1 \}$

Applying this partition on the index set yields:

$E_{i,j} = \bigcup_{j \in T_{2,i}} \{ x_1 \}$.

Now, the union bound results in:

$(54) \quad E(L_{x_1}) = \sum_{i = 0}^{M - 1} \sum_{j = 0}^{N/2} \sum_{0 \leq j \leq N/2} U(d_{x_1} \in E_{i,j})$.

We shall denote by $N_{d_{x_1}}(d_{x_1}, l) := \sum_{l = 0}^{L} N_{\omega}(d_{x_1}, l)$ the cardinality of the set $T_{2,i}$. Certainly, $\sum_{l = 0}^{L} N_{d_{x_1}}(d_{x_1}, l) = N_{d_{x_1}}(d_{x_1})$.

By the same token as Theorem 1, it is found that:

$(55) \quad E(L_{x_1}) \leq \sum_{i = 0}^{M - 1} \sum_{j = 0}^{N/2} \sum_{l = 0}^{L} P(N_{d_{x_1}}(d_{x_1}, l) \notin E_{i,j})$.

and finally:

$(56) \quad E(L) \leq \sum_{i = 0}^{M - 1} \sum_{j = 0}^{N/2} \sum_{l = 0}^{L} P(N_{d_{x_1}}(d_{x_1}, l) \notin E_{i,j})$.

where $N_{d_{x_1}}(d_{x_1}, l) := \frac{1}{M} \sum_{i = 0}^{M - 1} N_{d_{x_1}}(d_{x_1}, l)$ is the average number of pairs of signal having both squared Euclidean distance $d_{x_1}$ and $l$ distinct symbols.

By using the exponential bound on the error rate function, (56) can be rewritten as:

$(57) \quad E(L) \leq 1/2 \sum_{l = 0}^{L} \sum_{i = 0}^{N/2} \sum_{j = 0}^{N/2} \sum_{l = 0}^{L} P(N_{d_{x_1}}(d_{x_1}, l) \notin E_{i,j})$.

On the other hand, it can be seen from Definition 5 that:
Comparing (58) and the right hand side of (57), we recognize that:

\[ E(L) \leq \frac{1}{N} \frac{\partial T(D, L ; \Omega)}{\partial L} \bigg|_{L=1, D=D_0} \]

Furthermore, the 2-dimensional symbol error probability can be found to be \( P_e(2) = \frac{N}{2} \) so the claimed bound follows.

**Lemma 13.** The upper bound on the 2-dimensional error probability (Theorem 12) is less than or equal to that on the \( N \)-dimensional error probability (Lemma 7), that is:

\[ \frac{1}{N} \frac{\partial T(D, L; \Omega)}{\partial L} \bigg|_{L=1, D=D_0} \leq \frac{1}{N} \frac{\partial T(D, L; \Omega)}{\partial L} \bigg|_{L=1, D=D_0} \]

**Proof.** Using Definition 5, bound (35) on the \( N \)-dimensional error rate can be rewritten as:

\[ P_e(N) \leq \frac{1}{N} \frac{\partial T(D, L; \Omega)}{\partial L} \bigg|_{L=1, D=D_0} \]

Furthermore, it can be noted from (58) that:

\[ \frac{\partial T(D, L; \Omega)}{\partial L} \bigg|_{L=1, D=D_0} \leq \frac{N}{2} \sum_{d=0}^{N-1} \sum_{d=0}^{N-1} N_{D_0}(d, L) \frac{dD0}{dL} \]

since \( 1 \leq N/2 \) always holds. Accordingly, the asserted inequality follows by dividing both members by \( N \). QED.

In the following, these results are applied to multidimensional constellations with random coded modulation.

**Theorem 14.** The symbol error rate of multidimensional random coded constellations is upper bounded by:

\[ P_e(N) \leq \frac{1}{N} \frac{\partial T(D, L; \Omega)}{\partial L} \bigg|_{L=1, D=D_0} \]

**Proof.** First, we consider the constituent 2-dimensional constellation \( T_2(D) \). We introduce then the two-variable transfer function in agreement with:

\[ T(D, L; \Omega) = L T_2(D, L) \]

Consequently, the associated complete transfer function is:

\[ T(N, D, L) = L T(N, D, L) \]

Obviously, one has \( T_2(D, 1) = T_2(D) \) as expected. We go further to \( N \)-dimensions and calculate the complete transfer function:

\[ T(N, D, L) = [T_2(D, L)]^{N/2} \]

so we obtain an expression of the transfer function in terms of that of its 2D constituent:

\[ T(D, L; \Omega) = [T_2(D, L)]^{N/2} = 1 \]

In order to apply Theorem 12, we take the partial derivative of the above function with respect to \( L \), resulting in:

\[ \frac{\partial T(D, L; \Omega)}{\partial L} = \frac{N}{2} [T_2(D, L)]^{N/2-1} \frac{\partial T_2(D, L)}{\partial L} \]

Furthermore, it can be seen from (61) that:

\[ \frac{\partial T_2(D, L)}{\partial L} \bigg|_{L=1, D=D_0} = T_2(D) - 1 \]

Therefore, the result follows by inserting (62) into (63) and then applying Theorem 12. QED.

This can be illustrated by the example of the random coded constellation with \( N \)-polyst constellation. Thus, we have:

\[ T_2(D, L) = 1 + 2DL + D^2L \]

For \( N = 4 \), for instance, it follows that:

\[ T_2(D, L) = 1 + 4DL + (2L + 1)D^2L \]

The bound on the symbol error rate resulting from the previous theorems is:

\[ P_e(N) \leq \frac{1}{N} \frac{\partial T_2(D, L; \Omega)}{\partial L} \bigg|_{L=1, D=D_0} \]

V. ASYMPTOTIC BEHAVIOUR OF VERY LONG CODES

Another question which should be addressed is about the famous statement (almost) all codes are good, now thinking of the coded modulation context. We present below a straightforward theorem which clarifies this point.

**Theorem 15.** (all coded constellations are good). Given an arbitrary \( c > 0 \), if we pick at random a particular coded (redundant) signal set \( \mathcal{A} \) with length \( N \to \infty \), then almost surely (a.s.) it is a good one (that is, \( P_e(N) < c \)), provided \( R < \mathcal{R}_0 \). 

**Proof.** The theorem on the partial cut-off rate asserts that:

\[ P_e(N) \leq \exp(-N(R_0(N) - R)) \]

Applying then Definition 2, it follows readily that:

\[ \lim_{N \to \infty} P_e(N) = 0 \]

provided \( R < \mathcal{R}_0 \). Therefore, recalling (14), we can write:

\[ \lim_{N \to \infty} P_e(N) = \lim_{N \to \infty} \exp(-N(R_0(N) - R)) = 0 \]

Changing the order of the limits, we get:

\[ \lim_{L \to \infty} \lim_{N \to \infty} \exp(-N(R_0(N) - R)) = 0 \]

still assuming that \( R < \mathcal{R}_0 \).

On the other hand, given \( \epsilon > 0 \), we define bad and good sets, respectively, by:

\[ B_c(L) = \{ 1 \leq 1, 2, \ldots, L \} : \lim_{N \to \infty} P_e(N) > \epsilon \]

and:
and
\[ G_k(L) := \{ l, 1, 2, \ldots, k \} \cap \lim_{N \to \infty} P^{W^L}(N) \leq \epsilon. \]

Thus, the bad set is an index set pointing out all the codes for which the ND-error probability is not arbitrarily small, as the dimensionality increases without limit. Of course, \( \#(L) \cdot |G_k(L)| = L \) since every code is either good or bad. Let \( B_k \) (resp. \( G_k \)) be the bad (resp. good) set when \( L \) increases indefinitely. Accordingly,
\[
0 \leq \lim_{L \to \infty} \frac{1}{L} \sum_{n=1}^{L} \left( \lim_{N \to \infty} P^{W^L}(N) \right) \geq \lim_{L \to \infty} \frac{1}{L} \sum_{n=1}^{L} \frac{1}{|B_k(L)| - |G_k(L)| \cdot |B_k(L)|}.
\]

More rigorously, for \( \epsilon, \epsilon' > 0 \), there exists \( N_{\epsilon, \epsilon'} \) such that \( P^{W^L}(N) < \epsilon \) for \( N > N_{\epsilon, \epsilon'} \). We thus have:
\[
\epsilon \epsilon' > P^{W^L}(N) \geq \frac{1}{L} \sum_{n=1}^{L} P^{W^L}(N),
\]
and the result follows from the fact that \( \epsilon' \) is arbitrary.

Furthermore, \( B_k(L)/L \) converges a.s. (by the strong law of large numbers) to the probability of picking a bad code, \( P(B_k) \), that is, \( \lim_{L \to \infty} |B_k(L)|L^{-1} = P(B_k) \) a.s. Hence, we find that \( P(B_k) \leq \epsilon < 0 \) a.s. which implies \( P(B_k) = 0 \) a.s. Thus, except for a set of measure zero, all the codes have vanishing error probability as \( N \) goes to infinity, i.e.,
\[
\lim_{N \to \infty} \lim_{L \to \infty} \frac{1}{L} \sum_{n=1}^{L} P^{W^L}(N) = 0.
\]

Indeed, like the random codes, random coded constellations are not uniformly good but similar remarks ([19], p. 39) can be made.

**Corollary 15.** Almost all coded constellations exhibit virtually the same performance, provided their number of dimensions is very large.

**Proof.** Consider two arbitrary codes, denoted by the superscripts \( \phi_i \) and \( \phi_j \), respectively. We shall prove that almost surely \( P^{W^L}(N) - P^{W^L}(N) \leq \epsilon \) when \( N \to \infty \). Conversely, assume that \( P^{W^L}(N) - P^{W^L}(N) > \epsilon \). Without loss of generality it may be assumed that \( P^{W^L}(N) > P^{W^L}(N) > \epsilon \). Hence, at least one of the two codes must belong to the bad set when \( N \) increases indefinitely. Therefore, \( P^{W^L}(N) = P^{W^L}(N) > \epsilon \) \( \Rightarrow \) \( P(U(j) \in B_k) = 2P(B_k) \), which goes to zero and the claimed statement follows, too.

Any open interval on the real line is denoted hereafter by an open ball of center \( K \) and radius \( \epsilon \), \( B(K, \epsilon) = \{ x \in B^d \mid |x - K| < \epsilon \} \), and its complementary event by \( B^d(K, \epsilon) = B^d \setminus B(K, \epsilon) \).

**Lemma 17.** Almost all constellations exhibit the distance hardening phenomenon as their dimensionality increases indefinitely.

**Proof.** Let \( K = E(B^d) \) and an arbitrary \( \epsilon > 0 \). We want to show that a very long code picked at random satisfies \( P^{W^L}(B^d(x) \geq \epsilon) = P^{W^L}(B^d) \) in \( B^d \), \( K \) \to 0 \) a.s. From (16) and Definition 1, we have:

\[
\frac{1}{L} \sum_{n=1}^{L} P^{W^L}(N) \leq \frac{1}{L} \sum_{n=1}^{L} P^{W^L}(N) \leq \frac{1}{L} \sum_{n=1}^{L} P^{W^L}(N) \leq \frac{1}{L} \sum_{n=1}^{L} P^{W^L}(N).
\]

Summing up for any \( B^d \in B^d \) and changing the order of summations in the right-hand side, we get:
\[
P^{W^L}(B^d(x) \geq \epsilon) = \frac{1}{L} \sum_{n=1}^{L} P^{W^L}(B^d(x) \geq \epsilon) = 0.
\]

When \( \epsilon \) increases without limit, Theorem 11 implies that:
\[
P^{W^L}(B^d(x) \geq \epsilon) = 0.
\]

Now the proof follows quite close to that of Theorem 15. We define here the bad and good sets to be, respectively:
\[
B_k(x) = \{ x \in B^d \mid \epsilon \leq \epsilon \}
\]
\[
G_k(x) = \{ x \in B^d \mid \epsilon > \epsilon \}
\]
and
\[
\lim_{N \to \infty} \lim_{L \to \infty} \frac{1}{L} \sum_{n=1}^{L} P^{W^L}(x \in B^d(K, \epsilon)) = 0.
\]

The proof is then completed by the same reasoning as for Theorem 9, q.e.d.

**Definition 6.** Given \( \epsilon > 0 \), two discrete distributions \( P^{W^L}(x) \) and \( P^{W^L}(x) \) are said to be \( \epsilon \)-indistinguishable iff one has:
\[
\forall x \in B^d, \exists k > 0 \exists x \in B^d \mid P^{W^L}(x) - P^{W^L}(x) < \epsilon.
\]

We can see the \( \epsilon \)-indistinguishability as a binary operation between distributions which will be denoted by \( \sim \). Further properties of this relation are presented in Appendix B.

**Definition 7** (quasi-identical codes) : two coded signal sets denoted by superscripts \( \phi_i \) and \( \phi_j \) are said to be identical within \( \epsilon \) (quasi-identical) iff their normalized squared distance distributions are \( \epsilon \)-indistinguishable.

The relation \( \sim \) asymptotically induces an equivalence relation between codes as \( N \) increases (see Appendix B). In the following, we show that all large constellations become equivalent in the above sense. We adopt Sup (resp. Inf) to denote the least upper bound (resp. greatest lower bound).

**Theorem 18.** For any \( \epsilon > 0 \), almost all codes (coded constellations) are quasi-identical to the average code (random code), provided their dimensionality \( N \) is very large.

**Proof.** We pick at random a redundant constellation (superscript \( \phi \)) on the assumption that the dimension is high enough. We intend to apply 127 (Appendix B)
to show that \( P^{\#}(S) \sim \mathcal{P}(S) \). Given any \( \kappa > 0 \), we begin by splitting \( \mathbb{R}^m \) into two sets, namely:

\[
\begin{align*}
\mathcal{R}_\kappa^+ := \{ \mathbf{x} \in \mathbb{R}^m | \mathcal{E}(\mathbf{x}) \in \mathcal{E}^+(\kappa) \} \\
\mathcal{R}_\kappa^- := \{ \mathbf{x} \in \mathbb{R}^m | \mathcal{E}(\mathbf{x}) \in \mathcal{E}^-(\kappa) \}.
\end{align*}
\]

Now we must evaluate:

\[
\max_{\mathbf{x} \in \mathcal{R}_\kappa^+} \sup_{\mathbf{y} \in \mathcal{R}_\kappa^-} |P^{\#}(\mathbf{y}^+ | \mathbf{x}) - \mathcal{P}(\mathbf{y}^+ | \mathbf{x})|.
\]

Firstly, we carry out the supremum of \( F_\kappa(K) \) on \( \mathcal{R}_\kappa^+ \). In this case we have:

\[
\sup_{\mathbf{x} \in \mathcal{R}_\kappa^+} P^{\#}(\mathbf{x}^+ | \mathbf{x}) = \sup_{\mathbf{x} \in \mathcal{R}_\kappa^+} P^{\#}(\mathbf{x}^+ | \mathbf{x}).
\]

It follows from the squared distance hardening properties (Theorem 15 and Lemma 17) that:

\[
\sup_{\mathbf{x} \in \mathcal{R}_\kappa^+} F_\kappa(K) \leq \varepsilon/2 + \varepsilon/2 \leq \varepsilon, \quad \text{a.s.}
\]

Secondly, we carry out the supremum of \( F_\kappa(K) \) on \( \mathcal{R}_\kappa^- \). Now we use the fact that:

\[
\sup_{\mathbf{x} \in \mathcal{R}_\kappa^-} F_\kappa(K) \leq \sup_{\mathbf{x} \in \mathcal{R}_\kappa^-} [1 - P^{\#}(\mathbf{x}^+ | \mathbf{x})] = \sup_{\mathbf{x} \in \mathcal{R}_\kappa^-} [1 - \mathcal{P}(\mathbf{x}^+ | \mathbf{x})],
\]

yielding:

\[
\sup_{\mathbf{x} \in \mathcal{R}_\kappa^-} F_\kappa(K) \leq \varepsilon/2 + \varepsilon/2 \leq \varepsilon, \quad \text{a.s.}
\]

The proof is completed by combining (75) and (78) with (77). a.s.

Theorem 19. Almost all codes become quasi-identical for any \( \varepsilon > 0 \) as \( N \) increases without limit. a.s.

Proof. Consequences of both the above theorem and the asymptotic transitivity property of the relation \( \sim \), (see P3, Appendix B). a.s.

VI. GALLAGER TYPE BOUNDS FOR CODED MODULATION

In this section we intend to derive upper bounds on the error probability of finite-dimensional coded constellations with \( M \) equally likely signals on an additive white Gaussian noise channel. These bounds are, in many ways, similar to those by Gallager [20]. They therefore extend bounds established in the previous sections. Again, by abuse of language, a coded constellation will be referred to indiscriminately as a code.

Lemma 20. The \( N \)-dimensional (block) error probability \( P_{\text{err}}(N) \) for a particular code, say with superscript \( \# \), is upper bounded by:

\[
P_{\text{err}}(N) \leq \exp_{\text{a}}\left[-N \mathcal{E}^2_{\text{err}}(\rho, N) \right],
\]

where \( \mathcal{E}^2_{\text{err}}(\rho, N) := \sup_{\rho \leq \rho_0} [a \rho + \mathcal{E}^2_{\text{err}}(\rho, N)] \) is a reliability function of the code:

\[
\mathcal{E}^2_{\text{err}}(\rho, N) := \frac{1}{N} \log_2 \left\{ \sum_{i=1}^{N} \left[ \sum_{d \in \mathcal{D}_i} P_{\text{err}}(\mathbf{d}^\rho | \mathbf{x}_i^\rho) \right]^{\rho_0} \right\},
\]

where \( \rho_0(\mathbf{d}) \) denotes the pairwise error probability of any pair of signal points whose squared Euclidean distance is \( \mathbf{d}^\rho \) and where \( P_{\text{err}}(\mathbf{d}^\rho | \mathbf{x}_i^\rho) := P_{\text{err}}(\mathbf{d}^\rho | \mathbf{x}_i^\rho) \mathcal{M}(\mathbf{d}^\rho) \) is the distance spectrum of the code \( \# \), conditioned on \( \mathbf{x}_i^\rho \) being transmitted.

The proof of this lemma, which relies on the same inequalities as used by Gallager [20], is given in Appendix C.

Some comments would be worthwhile. First, we mention as a particular case a Bhattacharyya-union type bound which corresponds to \( \rho = 1 \):

\[
P_{\text{err}}(N) \leq \exp_{\text{a}}\left[-N \mathcal{E}^2_{\text{err}}(\rho, N) \right],
\]

where we have defined a code cutoff rate according to:

\[
\mathcal{R}^0_{\text{err}}(N) := \mathcal{E}^2_{\text{err}}(1 | N).
\]

It should be pointed out that this definition is close to that of the conventional cutoff rate \( \mathcal{R}_0 \).

Secondly, note that each code \( \# \) has an upper bound expressed in the above form. However, it is quite difficult to calculate bounds for a particular code since its own distance spectrum should be known. Hence, we would like to consider random coding which deals with all codes and whose distance spectrum is simpler to obtain but, unlike Theorem 2, it is now impossible to directly derive an inequality valid in the average and involving an unnormalized distance profile from the inequalities valid for each of \( L \) codes. This can be done again by following Gallager [20], resulting in:

Theorem 21. The average \( N \)-dimensional (block) error probability is upper bounded as:

\[
P_{\text{err}}(N) \leq \exp_{\text{a}}\left[-N \mathcal{E}^0_{\text{err}}(\rho, N) \right], \quad 0 \leq \rho \leq 1,
\]

where:

\[
\mathcal{E}^0_{\text{err}}(\rho, N) := \frac{1}{N} \log_2 \left\{ \sum_{i=1}^{N} \mathcal{P}(\mathbf{d}^\rho | \mathbf{x}_i^\rho) \mathcal{M}(\mathbf{d}^\rho)^{\rho_0} \right\}.
\]

and $P_e(d^*) := \frac{1}{M} \sum_j \sum_{i \neq j} |P_j(s_i|x_j)|^2$, the average normalized distance profile of the code.

Sketch of the proof. Let $P_e(s_i|x_j)$ be the error probability of a code assumed to be chosen at random, when $s_i$ was transmitted (for the sake of brevity, we dropped the superscript #/l). Using otherwise the same notation as in Section II, we may write for this particular code:

$$P_e(s_i|x_j) = \sum_j P_j(s_i|x_j) \Psi_j,$$

where $\Psi_j = 1$ if there exists $j$ such that $P_j(s_i|x_j) \equiv P_j(s_i|x_j)$, $\Psi_j = 0$ otherwise. An upper bound of $\Psi_j$ is clearly:

$$\Psi_j \leq \left[ \sum_{i \neq j} P_j(s_i|x_j) \right]^{1/2}, \quad \rho > 0.$$

Thus

$$P_e(s_i|x_j) \leq \left\{ \frac{1}{M} \sum_{i \neq j} P_j(s_i|x_j) \right\}^{1/2} \leq \left\{ \frac{1}{M} \sum_{i \neq j} P_j(s_i|x_j) \right\}^{1/2}.$$

Since coding is random, the probabilities $P_j(s_i|x_j)$ are random variables. We intend to upper bound the expectation of $P_e(s_i|x_j)$. To this end, we shall upper bound the expectation of the right hand side in the previous inequality. We notice that:

— the expectation of the sum over $j$ equals the sum of the expectations of its terms;
— the term into brace is the product of two independent random variables so its expectation is the product of their expectations;
— if we restrict ourselves to $\rho \leq 1$, then $E(X^\rho) \leq |E(X)|^\rho$ for any random variable $X$ so we may replace

$$E \left( \sum_{i \neq j} P_j(s_i|x_j) \right)^{1/\rho}$$

by

$$\left[ E \left( \sum_{i \neq j} P_j(s_i|x_j) \right) \right]^{1/\rho};$$

— finally, the sum and expectation may still be interchanged after it has been done. All these changes result in:

$$E[P_e(s_i|x_j)] \leq \sum_i \left\{ E \left[ \left[ \sum_{i \neq j} P_j(s_i|x_j) \right]^{1/\rho} \right] \right\}, \quad \rho \leq 1.$$

Now $E[P_j(s_i|x_j)]^{1/\rho}$ is independent from $s_i$, hence equal to $E[P_j(s_i|x_j)]$ for every $i$. Also, $E[P_j(s_i|x_j)]$ does not depend on $s_i$ and equals the mean error probability denoted by $P_e(s_i|x_j)$. All the terms in the sum on $j$ are moreover equal. We thus have:

$$P_e(s_i|x_j) = E[P_e(s_i|x_j)] \leq (M - 1)^{1/\rho} E[P_j(s_i|x_j)]^{1/\rho},$$

hence

$$P_e(s_i|x_j) \leq (M - 1)^{1/\rho} [E[P_j(s_i|x_j)]^{1/\rho}]^{1/\rho}.$$

When comparing the expectation, we may gather the terms which correspond to a same Euclidean distance $d$, as we did in the proof of Theorem 2, which results in:

$$E[P_j(s_i|x_j)] = \frac{1}{M} \sum_{d=d_{min}}^{d_{max}} N_{min}(d^2)p(d^2)\text{; we let } P_j(d^2) = \frac{1}{M} N_{min}(d^2) \text{ so we get:}$$

$$E[P_j(d^2)] \leq (M - 1)^{1/\rho} \left[ \sum_{d=d_{min}}^{d_{max}} P_j(d^2)p(d^2)\right]^{1/\rho}.$$

Neglecting 1 with respect to $M$, remembering that $M = \exp_2(N/B)$ and letting:

$$E_0(\rho, N) := \frac{1}{N} \log_2 \left[ M \sum_{d=d_{min}}^{d_{max}} P_j(d^2)p(d^2)\right]^{1/\rho},$$

we obtain the expected bound:

$$P_e(s_i|x_j) \leq \exp_2 \left[ -N\rho + E_0(\rho, N) \right], \quad \rho \leq 1.$$

This bound is valid as far as $E_0(\rho, N) > 0$. It closely resembles that of Lemma 20, except that the average distance profile instead of that of a particular code appears in the reliability function (80) and that $\rho$ is now restricted to the range $0 < \rho \leq 1$.

This restriction on $\rho$ means that the Bhattacharyya union bound plays a special role. The discussion of the bound of Theorem 21 can follow closely that of Gallager [20]. The tightest bound is obtained for the largest exponent $E_0(\rho, N) = \rho R$, but since $\rho$ is constrained to be at most 1, the Bhattacharyya-union bound is the tightest one at the smallest rates. As one wants to increase the rate beyond some limit, however, the tightest bound corresponds to decreasing $\rho$. If the bound used the probabilities $P_j(s_i|x_j)$, then the envelope of the exponent curves would have become tangent to the $R$ axis at $R = C$, the channel capacity, as $\rho \rightarrow 0$. However, the above bound was computed in terms of $P_j(s_i|x_j)$ instead of $P_j(s_i|x_j)$ and the behavior of the exponential bound as the rate increases is not clear. The condition $E_0(\rho, N) > 0$ may restrict the rate to a value less than the capacity, reached for $\rho$ strictly positive.
VII. CONCLUDING REMARKS

We applied Shannon's random coding argument to coded modulation. Upper bounds on the block and symbol error probability of finite-dimensional constellations have been developed. We defined a partial cut-off rate in terms of which we can upper bound the average error probability. An open problem is the relationship between the cut-off rate \( R_c \) and the conventional cut-off rate \( R_s \) [30]. A more general upper bound on the union bound should be used to allow for higher rates. Added to that, we mention that Hughes' cone bound [21] could be used instead of the union bound in order to improve the error probability estimation, especially for short codes.

Concerning the squared Euclidean distance distribution, parameters like its first two moments have been calculated. The asymptotic behaviour of such systems has been investigated. We have found that a hardening phenomenon occurs with the normalized squared Euclidean distance. We have also shown that virtually all large coded signal sets are good with probability asymptotically approaching one provided the rate does not exceed a critical value. Finally, we examined how Gallager-type bounds can be applied for coded modulation.

The investigations carried out in this paper hopelessly provide a better insight into the performance of multidimensional coded constellations. We should mention that still another facet of this problem namely, the coding gains of lattice coded modulations, is addressed in a forthcoming paper [22].

Acknowledgments

The first author thanks the Association pour le renseignement en communications (ARECOM) for its partial financial support. He also gratefully acknowledges Federal University of Pernambuco (UFPE, Brazil) for granting him leave of absence to pursue this research. We acknowledge the many useful comments and suggestions made by the members of the examination board (Professors J. Keedwell, J-C. Bic, H. Sari and M. Rouaiante) before the first author upheld his doctoral dissertation. We are also grateful to the anonymous referees for helpful suggestions.

References
\[ N(\mathbf{d}_2^1) = \begin{bmatrix} 2 & 3 & 2 \\ 3 & 4 & 2 \\ 2 & 3 & 3 \\ 3 & 4 & 3 \\ 2 & 3 & 3 \\ 2 & 3 & 3 \\ 3 & 4 & 3 \\ 2 & 3 & 3 \\ 2 & 3 & 3 \\ 2 & 3 & 3 \end{bmatrix} \]

\[ N(\mathbf{d}_2^2) = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 2 & 1 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \]

\[ N(\mathbf{d}_2^3) = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 2 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \]

The values of \( N(\mathbf{d}_2^i) \) for \( \mathbf{d}_2 \in \Delta(B) \) are computed by averaging all the elements in the corresponding matrix, as previously described.

**APPENDIX B**

**Properties of the relation \( \sim_{\epsilon} \)**

We present here a few properties of the relation \( \sim_{\epsilon} \), involving the definition of \( \epsilon \)-indistinguishable distributions.

**P1.** If \( P^{\Phi_1}(x) \sim_{\epsilon_1} P^{\Phi_2}(x) \) then it follows that \( P^{\Phi_1}(x) \sim_{\epsilon_2} P^{\Phi_2}(x) \), provided \( \epsilon_2 > \epsilon_1 \).

**P2.** A necessary and sufficient condition for two distributions to be \( \epsilon \)-indistinguishable is that \( \forall \epsilon > 0 : \sup_{K \in \mathbb{R}} |P^{\Phi_1}(x \in R(K, \epsilon)) - P^{\Phi_2}(x \in R, \epsilon)| < \epsilon \).

**P3.** The relation induced by \( \epsilon \)-indistinguishable normalized squared distance distributions is asymptotically in equivalence relation, that is, a) \( \sim_{\epsilon} \) is reflexive, b) \( \sim_{\epsilon} \) is symmetric, c) \( \sim_{\epsilon} \) is asymptotically transitive.

**Proof.** a) and b) are trivial. To prove c), given \( \epsilon > 0 \), there exists \( N \) large enough such that \( P^{\Phi_1}(x) \sim_{\epsilon/2} P^{\Phi_1} \) and \( P^{\Phi_1} \sim_{\epsilon/2} P^{\Phi_2} \). Now \( P^{\Phi_1} \sim_{\epsilon} P^{\Phi_2} \), and \( P^{\Phi_1} \sim_{\epsilon} P^{\Phi_2} \) follow from P1. The asymptotic transitivity follows from P2 and the triangular inequality.

**APPENDIX C**

**Propositions and proofs related to Gallager type bounds**

This Appendix is devoted to propositions and proofs related to Gallager type bounds. First, we shall need some useful inequalities given by Propositions 1 and 2 below. Both are well known inequalities (see e.g. [30], p. 197) which may be interpreted as special cases of Hölder's inequality. We shall nevertheless give short proofs of them. Then, we generalize the union bound and apply the results obtained to evaluate performance of combined coding and modulation schemes previously described.

**Proposition 1 (inequalities).** Let \( l \) be a countable index set. Any sequence \( \{a_k \}_{k \in l} \) of real numbers verifies the following inequalities:

\[ (i) \sum_{k \in l} |a_k| \leq \left( \sum_{k \in l} |a_k|^\gamma \right)^{1/\gamma}, \quad \forall \gamma \geq 0, \]

\[ (ii) \sum_{k \in l} |a_k| \geq \left( \sum_{k \in l} |a_k|^{\gamma} \right)^{1/\gamma}, \quad \forall \gamma : -1 < \gamma \leq 0, \]

**Proof.** The reasoning used here follows Gallager [20].

Clearly, for any \( \gamma \geq 0, \)

\[ (C1) \quad \Phi_l := \left( \sum_{k \in l} |a_k|^\gamma \right)^{1/\gamma} \geq 1 \]

and therefore

\[ (C2) \quad \sum_{k \in l} |a_k| \leq \left( \sum_{k \in l} |a_k|^{\gamma} \right)^{1/\gamma} = \left( \sum_{k \in l} |a_k|^\gamma \right)^{1/\gamma}, \]

so part (i) follows.

Furthermore, part (ii) of Proposition 1 results in:

\[ (C3) \quad \sum_{k \in l} |a_k| \leq \left( \sum_{k \in l} |a_k|^{\gamma} \right)^{1/\gamma} \]

for any \( 0 < \beta \leq 1 \). Considering a given sequence \( \{a_k \}_{k \in l} \), we apply (C3) to \( |a_k| = |a_k|^{1/\beta} \), which results in:

\[ (C4) \quad \sum_{k \in l} |a_k|^{1/\beta} \leq \sum_{k \in l} |a_k|^{1/\beta} \]

for any \( \beta : 0 < \beta \leq 1 \).
The proof of part (ii) is completed by raising both sides to the power $\beta$ and by letting $\beta = 1 + \rho (\rho < 0)$.

**Proposition II.** Let $\{x_i\}_i \in \mathcal{I}$ be a sequence of real numbers and let $\{n_i\}_i \geq 0$ denote a sequence of integers, $n_i \geq 0$. Then for any $\rho > 0$, we have the inequality:

$$\sum_{i \in \mathcal{I}} n_i |x_i| \leq \left[ \sum_{i \in \mathcal{I}} n_i |x_i|^\beta \right]^{1/\beta}.$$

**Proof.** We now assume that a new sequence of numbers $\{\bar{x}_i\}_i \in \mathcal{I}$ results from the previous sequence by repeating each of its elements $n_i$ times. The cardinality of the new index set $\mathcal{I}'$ is clearly $|\mathcal{I}'| = \sum n_i$. Applying Proposition I to this sequence of numbers results in:

$$(C5) \sum_{i \in \mathcal{I}} n_i |x_i| - \sum_{i \in \mathcal{I}} |x_i| \leq \left[ \sum_{i \in \mathcal{I}} n_i |x_i|^\beta \right]^{1/\beta}.$$ 

The proof is completed by changing the index set from $\mathcal{I}'$ to $\mathcal{I}$, etc.

Proposition I may be applied in conjunction with the union bound in order to derive a family of upper bounds on the probability of a union of events. To begin with, let $\{P(A_i)\}_{i \in \mathcal{I}}$ be the respective probabilities of events $A_i \in \mathcal{I}$. Consequently, we can easily infer from part (i) of Proposition I that, for any $\rho > 0$:

$$(C6) \sum n_i P(A_i) \leq \left[ \sum n_i P(A_i)|x_i|^\beta \right]^{1/\beta}.$$ 

so we have established the following family of bounds.

**Proposition III (generalized union bound).** Let $\{A_i\}_{i \in \mathcal{I}}$ be an ensemble of (not necessarily disjoint) events. Then the following bounds hold for any $\rho \geq 0$:

$$P(\bigcup_{i \in \mathcal{I}} A_i) \leq \sum_{i \in \mathcal{I}} P(A_i)|x_i|^\beta \left[ \sum n_i P(A_i)|x_i|^\beta \right]^{1/\beta}.$$ 

Clearly, the union bound is the particular case corresponding to $\rho = 0$. Also, we can get some kind of Bhattacharyya union bound for $\rho = 1$, namely:

$$(C7) P(\bigcup_{i \in \mathcal{I}} A_i) \leq \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{I}} P(A_i) P(A_j).$$

Furthermore, we mention that most bounds in this family are strictly weaker than the union bound.

We are now able to give a proof of Lemma 20.

**Proof of Lemma 20.** The probability of a code, say $\#1$, can be evaluated in terms of the conditional error probability given that a particular signal is transmitted, i.e.,

$$(C8) P_{\#1}(N) = \frac{1}{M} \sum_{i=0}^{N-1} P(e|x_i).$$

The conditional error probability can be bounded by a (generalized) union bound as follows. Let $x_i \rightarrow \bar{x}_j$ denote the event decoding $\bar{x}_j$ when $x_i$ is sent and $x_j$ is considered the only alternative. We suppose that maximum likelihood decoding (MLD) is carried out by the decoder, which is optimal in the sense that it minimizes the error rate. We also consider an index set $\mathcal{I}_j$ defined as:

$$\mathcal{I}_j := \{ j \in \{1, \ldots, M - 1 \} : \exists i \in \mathcal{I} \text{ s.t. } d_\Delta(x_i, \bar{x}_j) = d_\Delta^0 \}.$$ 

By a union bound argument, we find out:

$$(C9) P(e|x_i) \leq \sum_{\mathcal{I}_j} \sum_{i \in \mathcal{I}_j} P(e|x_i \rightarrow \bar{x}_j).$$

where $P(e|x_i \rightarrow \bar{x}_j)$ is the pairwise error probability which solely depends on $d_\Delta$ but not on $i$. We denote by $e(d_\Delta)$ the pairwise error probability of any pair of signal points whose squared Euclidean distance is $d_\Delta$.

Let $N_{\Delta}^0(d_\Delta) := |\mathcal{I}_j|$ denote the cardinality of the set $\mathcal{I}_j$, i.e., the number of signals at a squared distance $d_\Delta$ from $i$. Applying then Proposition II to (C9), we get:

$$(C10) P(e|x_i) \leq \sum_{d_\Delta = 0}^{d_\Delta^0} N_{\Delta}^0(d_\Delta) P(e|d_\Delta)^{1/\beta}.$$ 

Thus, a bound on the block error probability of a code can be found by inserting (C10) into (C8). This can be rewritten as:

$$(C11) P_{\#1}(N) \leq M^{-1} \sum_{d_\Delta = 0}^{d_\Delta^0} \sum_{r = 0}^{\beta - 1} \left[ \sum_{i \in \mathcal{I}_j} P(e|d_\Delta)^{1/\beta} \right]^{1/\beta}.$$ 

where $P_{\#1}(N) := N_{\Delta}^0(d_\Delta)/M$ is the (conditional) distance spectrum of the code $\#1$, still assuming that $\rho \geq 0$.

We now define a so-called Gallager function $E_\rho^\#(\rho, N)$, $0 \leq \rho < \infty$, according to:

$$(C12) E_\rho^\#(\rho, N) := \frac{1}{\beta \log(M)} \sum_{d_\Delta = 0}^{d_\Delta^0} \sum_{r = 0}^{\beta - 1} P(e|d_\Delta)^{1/\beta}.$$ 

(Compare, for instance, with the $E_{\infty}(\cdot)$ function defined by Gallager on a discrete memoryless channel [10].)

Remembering that $M = \exp_N(\beta N)$, it follows promptly that:

$$(C13) P_{\#1}(N) \leq \exp_N[-\rho R + E_{\rho}^\#(\rho, N)].$$

Finally, we can choose the tightest bound by minimizing the term in brackets in the above expression:

$$(C14) E_{\rho}^\#(R, N) := \sup_{\rho \in \rho} \left[ -\rho R + E_{\rho}^\#(\rho, N) \right],$$

and the proof follows.


BIOGRAPHY

Gérard Battail, graduate from the École nationale supérieure des Télécommunications, INST Ingénieur civil des Télécommunications, 1991. Professor at INST since 1993. Teaching and research in communication theory, modulation, source and channel coding, information theory.

REFERENCES