

# On Equivalence of Binary Asymmetric Channels regarding the Maximum Likelihood Decoding.

Claudio Qureshi, Sueli I. R. Costa, Christiane B. Rodrigues and Marcelo Firer

## Abstract

We study the problem of characterizing when two memoryless binary asymmetric channels, described by their transition probabilities  $(p, q)$  and  $(p', q')$ , are equivalent from the point of view of maximum likelihood decoding (MLD) when restricted to  $n$ -block binary codes. This equivalence of channels induces a partition (depending on  $n$ ) on the space of parameters  $(p, q)$  into regions associated with the equivalence classes. Explicit expressions for describing these regions, their number and areas are derived. Some perspectives of applications of our results to decoding problems are also presented.

## Index Terms

Binary memoryless channel, binary asymmetric channel, maximum likelihood decoding, mismatched decoding

## I. INTRODUCTION

The binary asymmetric channel  $BAC(p, q)$  is the discrete memoryless channel with binary alphabet and transition probability given by  $\Pr(1|0) = p$ ,  $\Pr(0|0) = 1 - p$ ,  $\Pr(0|1) = q$  and  $\Pr(1|1) = 1 - q$ , where  $\Pr(x|y)$  denotes the probability of receiving  $x$  if  $y$  was sent. Without loss of generality we will assume  $p \leq q$  and  $p + q < 1$  (see discussion in Section IV) and denote the space of parameters by  $\mathcal{T} = \{(p, q) \in [0, 1]^2 : p \leq q, p + q < 1\}$  which we refer to as the *fundamental triangle*.

The interest in binary asymmetric channels has increased due to applications in flash memories [1], [2], [3], [4] and as models in other areas, such as neuroscience [5]. Most of the work developed on binary asymmetric channels has focused on coding properties [6], [7] and on the design of codes with given properties, either for general asymmetric channels [8], [6], [9], [10], [11], [12] or for the Z-channel [13], [14], [15]. We stress that most of these works consider the decoding criterion determined by the asymmetric metric introduced in [6], or some variant of it. Despite of all its advantages, the asymmetric metric is not matched to any binary asymmetric channel, that is, it cannot be used to perform maximum likelihood decoding (MLD).

In this work we study the binary asymmetric channels from the point of view of maximum likelihood decoding. Namely, two memoryless binary asymmetric channels  $W_1$  and  $W_2$  are  $n$ -equivalent if MLD is the same for both channels, for every possible code  $C \subseteq \mathbb{F}_2^n$  (a precise statement is given in Definition 1). The most studied instances of these channels are the binary symmetric channels (BSCs) corresponding to  $p = q$ , and the Z-channels corresponding to  $p = 0$ . For these channels the problem we study here is trivial since two channels  $W_1$  and  $W_2$  which are either both symmetric channels or both Z-channels are always  $n$ -equivalent, for any positive  $n$ . In this sense, we could say that there is a unique BSC and a unique Z-channel from the MLD point of view. This is not the case of general binary asymmetric channels and to study this equivalence relation is the focus of this work.

The knowledge of the equivalence classes of binary asymmetric channels may be useful for two purposes.

- 1) To perform MLD on a memoryless binary channel it is necessary to know the transition probabilities  $p = \Pr(1|0)$  and  $q = \Pr(0|1)$ . The more precise the measurement of  $(p, q)$ , the less the risk of mismatching the channel. Let us suppose that after a number of experiments, we obtain an approximation  $(p', q')$  for the real transition probabilities  $(p, q)$  with an error of at most  $\varepsilon > 0$  (i.e. such that  $|p - p'| < \varepsilon$  and  $|q - q'| < \varepsilon$ ). We prove that the larger the block length of a code, the more is the risk of mismatching<sup>1</sup> the channel. In fact, there is a maximum block length  $N$ , which can be calculated explicitly from the results developed in this paper, such that there is no risk of mismatching when codes of block length at most  $N$  are considered (see Example 3). A dual situation occurs when we have an upper bound  $N$  on the block length of the codes to be used in a given BAC. In this case the probability of mismatching is reduced when we increase the precision in the measurement of the transition probabilities  $(p, q)$ . It is possible to find a maximum admissible error  $\varepsilon$  such that if the estimated transition probabilities  $(p', q')$  verify  $|p - p'| < \varepsilon$  and  $|q - q'| < \varepsilon$ , there is no risk of mismatching (see Example 4).
- 2) A very relevant measure of the performance of a code is the error probability of the encoding-decoding process. Given a memoryless channel  $W : \mathcal{X} \rightarrow \mathcal{X}$  and a code  $C \subseteq \mathcal{X}^n$ , we consider the ML decoder with input  $x \in \mathcal{X}^n$  and output the

The authors are with the Institute of Mathematics, Statistics and Computing Science of the University of Campinas, SP, Brazil (emails: cqureshi@ime.unicamp.br, sueli@ime.unicamp.br, chrismmor@gmail.com, mfirer@ime.unicamp.br).

<sup>1</sup>To mismatch a channel  $W$  means to use a decoding criteria different from the MLD with respect to the transition probabilities of  $W$ . For general references on mismatched decoding see [16], [17].

codeword  $c \in C$  such that  $\Pr_W(x|c) > \Pr_W(x|c')$  for all  $c' \in C, c' \neq c$ , if such codeword exists. Otherwise, in the case there are different codewords  $c$  that maximize  $\Pr_W(x|c)$ , it returns a 'FAIL' message. We will refer to this ML decoder as the *standard ML decoder*. For this decoder, the *error probability of the code  $C$  within the channel  $W$*  is given by

$$P_{\text{error}}(C, W) = 1 - \frac{1}{|C|} \sum_{c \in C} \sum_{x \in V_{(C, W)}(c)} \Pr_W(x|c)$$

where  $V_{(C, W)}(c) = \{x \in \mathcal{X}^n : \Pr_W(x|c) > \Pr_W(x|c'), \forall c' \in C, c' \neq c\}$  is the *probabilistic Voronoi region* of  $c$  (depending on  $W$  and  $C$ ). Here we are assuming the messages to be equiprobable. For the particular case of a BAC, the error probabilities  $\Pr_W(y|c)$  are easy to compute, since there are closed formulas for them. The difficult part of computing  $P_{\text{error}}(C, W)$  is determining the probabilistic Voronoi regions. We should mention that, since it is defined depending on  $W$ , we actually have an infinite number of instances, which a priori should be computed for any given  $W$  (and the finite number of possible codes  $C$ ). Our definition of equivalence of channels actually says that two channels  $W_1$  and  $W_2$  are  $n$ -equivalent if  $V_{(C, W_1)}(c) = V_{(C, W_2)}(c)$  for every  $C \subseteq \mathcal{X}^n$  and every  $c \in C$ . It follows that, knowing the equivalence classes reduces the problem from infinitely many instances of binary asymmetric channels to a finite number of equivalence classes of such channels (depending on the length  $n$  of the block codes).

We may consider another ML decoder different than the standard ML decoder if instead of return a 'FAIL' message in case of ambiguity (i.e. several codewords  $c$  maximizing  $\Pr_W(x|c)$ ), it returns a codeword  $c$  maximizing  $\Pr_W(x|c)$ , chosen uniformly at random. This (probabilistic) ML decoder will be called the *uniform ML decoder*. The definition of equivalence of channels proposed in this paper contemplates both decoders, the standard and the uniform ML decoder, as it is shown in the next section.

This paper is organized as follows: In Section II we introduce an equivalence relation between channels in such a way that equivalent channels determine equal decoding criteria when MLD is considered and discuss some properties of this relation. We consider for each (fixed)  $n \geq 2$  the above equivalence relation restricted to the BACs. In Section III we introduce the BAC-function, the key to describe the regions determined by the equivalence relation in the parameter space (i.e. in the fundamental triangle). The number of such regions is provided for every  $n \geq 2$ . In Section IV we discuss some properties of the BAC-function and the areas of the regions determined by its level curves, which are related to the probability of a random choice of  $(p, q)$  to produce a channel  $n$ -equivalent to a given BAC.

## II. $n$ -EQUIVALENCE OF CHANNELS

We start by introducing an equivalence relation (depending on  $n \in \mathbb{N}$ ) that characterizes when two memoryless channels with input and output alphabet  $\mathcal{X}$  determine the same ML decision for every  $n$ -block codes (i.e. subsets  $C \subseteq \mathcal{X}^n$ ).

Let  $W$  be a memoryless channel with input and output alphabet  $\mathcal{X}$  and  $n \in \mathbb{N}$ . For  $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in \mathcal{X}^n$  we denote by  $\Pr_W(x|y) := \prod_{i=1}^n \Pr_W(x_i|y_i)$  the probability of receiving  $x$  if  $y$  was sent through the channel  $W$ . When  $W = \text{BAC}(p, q)$  we denote this probability by  $\Pr_{(p, q)}(x|y)$  and if the parameters  $(p, q)$  are clear from the context or irrelevant, we denote it by  $\Pr(x|y)$ . Motivated by the definition of matching given in [18, Definition 1], we introduce the following equivalent relation between memoryless channels.

**Definition 1.** Let  $W_1, W_2 : \mathcal{X} \rightarrow \mathcal{X}$  be two memoryless channels and  $n$  be a positive integer. We say that  $W_1$  and  $W_2$  are  $n$ -equivalent (denoted by  $W_1 \sim_n W_2$ ) if for every  $n$ -block code  $C \subseteq \mathcal{X}^n$  and every word  $x \in \mathcal{X}^n$ , we have

$$\arg \max_{c \in C} \Pr_{W_1}(x|c) = \arg \max_{c \in C} \Pr_{W_2}(x|c)$$

where  $\arg \max_{c \in C} \Pr_W(x|c) = \{c \in C : \Pr_W(x|c) \geq \Pr_W(x|c'), \forall c' \in C\}$ . The channels  $W_1$  and  $W_2$  are  $\infty$ -equivalent (denoted by  $W_1 \sim_\infty W_2$ ) if they are  $n$ -equivalent for every  $n \geq 1$ .

Let  $\text{sdec}_W$  and  $\text{udec}_W$  denote the standard and uniform ML decoder introduced in Section I, with respect to a memoryless channel  $W : \mathcal{X} \rightarrow \mathcal{X}$  and let  $W_1, W_2 : \mathcal{X} \rightarrow \mathcal{X}$  be two memoryless channels. The equality  $\text{sdec}_{W_1} = \text{sdec}_{W_2}$  for  $n$ -block codes means  $\text{sdec}_{W_1}(C, x) = \text{sdec}_{W_2}(C, x)$  for every code  $C \subseteq \mathcal{X}^n$  and every  $x \in \mathcal{X}^n$ , and  $\text{udec}_{W_1} = \text{udec}_{W_2}$  for  $n$ -block codes means the equality of the probabilities  $\Pr(\text{udec}_{W_1}(C, x) = c) = \Pr(\text{udec}_{W_2}(C, x) = c)$  for every code  $C \subseteq \mathcal{X}^n$ , every  $x \in \mathcal{X}^n$  and every  $c \in C$ . The following theorem, whose proof is given in the Appendix, establishes a relation between the  $n$ -equivalence of channels and the ML decoders mentioned above.

**Theorem 1.** Let  $W_1, W_2 : \mathcal{X} \rightarrow \mathcal{X}$  be two memoryless channels and  $n \geq 2$ . The following assertions are equivalent.

- i)  $W_1$  and  $W_2$  are  $n$ -equivalent.
- ii) The standard ML decoders  $\text{sdec}_{W_1}$  and  $\text{sdec}_{W_2}$  are the same for  $n$ -block codes.
- iii) The uniform ML decoders  $\text{udec}_{W_1}$  and  $\text{udec}_{W_2}$  are the same for  $n$ -block codes.

Consider an order in the  $\ell$ -ary alphabet  $\mathcal{X}$  (for  $\mathcal{X} = \mathbb{F}_2$  we assume  $0 < 1$ ) and the lexicographical order in  $\mathcal{X}^n$  for every  $n \geq 1$ . With a memoryless channel  $W : \mathcal{X} \rightarrow \mathcal{X}$  we associate an  $\ell^n \times \ell^n$  real matrix  $M_n(W)$  whose  $ij$ -entry is given by

$\Pr_W(x|y) = \prod_{i=1}^n \Pr_W(x_i|y_i)$ , where  $x = (x_1, \dots, x_n)$  and  $y = (y_1, \dots, y_n)$  are the  $i$ -th and  $j$ -th elements of  $\mathcal{X}^n$  respectively. We refer to this matrix as the *transition matrix of order  $n$  of  $W$* , or simply as the  *$n$ -transition matrix*. Note that the 1-transition matrix is the usual transition matrix of the channel. For  $W = \text{BAC}(p, q)$ , the corresponding  $n$ -transition matrix is denoted by  $M_n(p, q)$ . This is a  $2^n \times 2^n$  real matrix and it is not difficult to see that if we consider the binary expansion of  $i = \sum_{k=1}^n x_k 2^{n-k}$  and  $j = \sum_{k=1}^n y_k 2^{n-k}$  ( $0 \leq i, j < 2^n$ ), the  $ij$ -element of  $M_n(p, q)$  is given by  $m_{ij} = \Pr_{(p,q)}(x|y) = \prod_{k=1}^n \Pr_{(p,q)}(x_k|y_k)$ . For example, the matrix  $M_5(p, q)$  is a  $32 \times 32$  matrix whose entry in the row  $9 = 01001_2$  and column  $29 = 11101_2$  is  $m_{9,29} = \Pr(01001|11101) = \Pr(0|0)\Pr(0|1)^2\Pr(1|1)^2 = (1-p)q^2(1-q)^2$ .

When considering  $n$ -equivalence, we need to compare the entries in each line of the  $n$ -transition matrix and it is useful to replace it by a simpler matrix. In [19], the authors substituted a matrix  $M$  by a matrix  $\tilde{M}$  with entries  $\tilde{M}_{ij} = k$  if  $M_{ij}$  is the  $k$ -th largest element (allowing ties) of the  $j$ -th column. In order to determine whether two channels are equivalent we adopt a different, but similar substitution.

**Definition 2.** Let  $\mathcal{M}_N(\mathbb{R})$  denote the set of  $N \times N$  real matrices and  $M \in \mathcal{M}_N(\mathbb{R})$ . The ordered form of  $M$  is the integer matrix  $M^* \in \mathcal{M}_N(\mathbb{Z})$  such that  $M_{ij}^* = \#\{k : 1 \leq k \leq N, M_{ik} < M_{ij}\}$ .

The next proposition characterizes the  $n$ -equivalence of channels in terms of the ordered form of their  $n$ -transition matrices (the proof of this result is given in the Appendix).

**Proposition 1.** Let  $W_1, W_2 : \mathcal{X} \rightarrow \mathcal{X}$  be two memoryless channels and  $M_1$  and  $M_2$  be their corresponding  $n$ -transition matrices. The following assertions are equivalent.

- i) The channels  $W_1$  and  $W_2$  are  $n$ -equivalent.
- ii)  $\Pr_{W_1}(x|y) \leq \Pr_{W_1}(x|z) \Leftrightarrow \Pr_{W_2}(x|y) \leq \Pr_{W_2}(x|z)$  for all  $x, y, z \in \mathcal{X}^n$ .
- iii)  $M_1^* = M_2^*$ .

**Remark 1.** If  $W : \mathcal{X} \rightarrow \mathcal{X}$  is a memoryless channel such that there exists  $\xi \in \mathcal{X}$  such that  $\Pr_W(\xi|\xi) > 0$ , then the map  $\mathcal{X}^n \rightarrow \mathcal{X}^{n+1}$  given by  $x \mapsto x^* := (x, \xi)$  verifies that  $\Pr_W(x^*|y^*) \leq \Pr_W(x^*|z^*) \Leftrightarrow \Pr_W(x|y) \leq \Pr_W(x|z)$  for all  $x, y, z \in \mathcal{X}^n$ . As a consequence of this fact, if  $W_1, W_2 : \mathcal{X} \rightarrow \mathcal{X}$  are two memoryless channels such that there exists  $\xi \in \mathcal{X}$  such that  $\Pr_{W_1}(\xi|\xi) > 0$  and  $\Pr_{W_2}(\xi|\xi) > 0$  (this is always the case if  $W_1$  and  $W_2$  are BACs) we have the following implications:

- $W_1 \stackrel{n}{\sim} W_2$  implies  $W_1 \stackrel{n+1}{\sim} W_2$ ;
- $sdec_{W_1} = sdec_{W_2}$  for  $n+1$ -block codes implies  $sdec_{W_1} = sdec_{W_2}$  for  $n$ -block codes;
- $udec_{W_1} = udec_{W_2}$  for  $n+1$ -block codes implies  $sdec_{W_1} = sdec_{W_2}$  for  $n$ -block codes.

In what follows we assume the channels are binary ( $\mathcal{X} = \mathbb{F}_2$ ). Let  $n$  be a fixed positive integer or infinite. The  $n$ -equivalence for BACs induces an equivalence relation on the fundamental triangle  $\mathcal{T} : (p, q) \stackrel{n}{\sim} (p', q')$  if and only if  $\text{BAC}(p, q) \stackrel{n}{\sim} \text{BAC}(p', q')$ . We denote by  $\Delta_n = \mathcal{T} / \sim_n$ , the set of equivalence classes and by  $\pi_n : \mathcal{T} \rightarrow \Delta_n$  the projection that associates  $(p, q)$  to its equivalence class (i.e.  $\pi_n(p, q) = \{(p', q') \in \mathcal{T} : (p', q') \stackrel{n}{\sim} (p, q)\}$ ). The main result of this paper is a complete description of the quotient set  $\Delta_n$ .

**Definition 3.** A decision criterion of order  $n$  for the BACs is an equivalence class  $\mathcal{A} \in \Delta_n$  (in particular  $\mathcal{A}$  is a subset of  $\mathcal{T}$ ). An  $n$ -stable decision criterion  $\mathcal{A}$  is a decision criterion of order  $n$  for the BACs which is an open set of  $\mathcal{T}$  and an  $n$ -unstable decision criterion is a decision criterion of order  $n$  for the BACs with no interior points.

**Remark 2.** If  $\mathcal{A}$  is a  $n$ -stable decision criterion for the BACs and  $(p, q) \in \mathcal{A}$ , then maximum likelihood decoding on  $\text{BAC}(p, q)$  restricted to  $n$ -block codes, remains the same under small perturbation of the parameter  $(p, q)$ .

**Definition 4.** A point  $(p, q) \in \mathcal{T}$  is  $n$ -stable if  $(p, q)$  is an interior point of  $\pi_n((p, q))$  and  $n$ -unstable otherwise. The  $n$ -stable region (denoted by  $\mathcal{R}_n^{st}$ ) is the set of all  $n$ -stable points and the  $n$ -unstable region (denoted by  $\mathcal{R}_n^{un}$ ) is the set of all  $n$ -unstable points.

We will prove later that every decision criterion is either stable or unstable, that is, if a criterion contains an  $n$ -stable point then all its points are  $n$ -stable. We conclude this section by discussing how the parameter space  $\mathcal{T}$  decomposes into different  $n$ -equivalence classes for  $n \leq 5$ .

The 1-transition matrix  $M_1(p, q)$  of  $\text{BAC}(p, q)$  is given by  $\begin{pmatrix} 1-p & q \\ p & 1-q \end{pmatrix}$ , thus its ordered form  $M_1(p, q)^* = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  does not depend on  $(p, q)$  and we have only one criterion, which is stable.

The 2-transition matrix  $M_2(p, q)$  of  $\text{BAC}(p, q)$  is given by

$$\begin{pmatrix} (1-p)^2 & (1-p)q & (1-p)q & q^2 \\ (1-p)p & (1-p)(1-q) & pq & q(1-q) \\ (1-p)p & pq & (1-p)(1-q) & q(1-q) \\ p^2 & p(1-q) & p(1-q) & (1-q)^2 \end{pmatrix}.$$

In this case the ordered form depends on  $(p, q)$  in the following way:

- If  $(p, q)$  is an interior point of  $\mathcal{T}$  the ordered form is given by  $M_2(p, q)^* = \begin{pmatrix} 3 & 1 & 1 & 0 \\ 1 & 3 & 0 & 2 \\ 1 & 0 & 3 & 2 \\ 0 & 1 & 1 & 3 \end{pmatrix}$ ;
- If  $p = 0$  (and hence  $q > 0$ , since  $(0, 0) \notin \mathcal{T}$ ) we obtain the ordered form  $M_2(p, q)^* = \begin{pmatrix} 3 & 1 & 1 & 0 \\ 0 & 3 & 0 & 2 \\ 0 & 0 & 3 & 2 \\ 0 & 0 & 0 & 3 \end{pmatrix}$ ;
- If  $p = q$  we obtain  $M_2(p, q)^* = \begin{pmatrix} 3 & 1 & 1 & 0 \\ 1 & 3 & 0 & 1 \\ 1 & 0 & 3 & 1 \\ 0 & 1 & 1 & 3 \end{pmatrix}$ .

In this case we have three different decision criteria, one stable and two unstable.

For  $n = 3, 4$  and  $5$  we started with some simulations using the software SAGE [20]. We considered a set  $A$  of  $28900 = 170^2$  points uniformly distributed on  $\mathcal{T}$  and calculated the ordered form  $M_n(p, q)^*$  of each  $(p, q) \in A$ .

For  $n = 3$  we observe two criteria,  $\mathcal{B}$  and  $\mathcal{R}$ , and by coloring the points in these regions by blue and red respectively we obtain the picture showing in Figure 1. As we will prove in Theorem 2, there are two stable criteria  $\mathcal{B}$  and  $\mathcal{R}$  (the connected

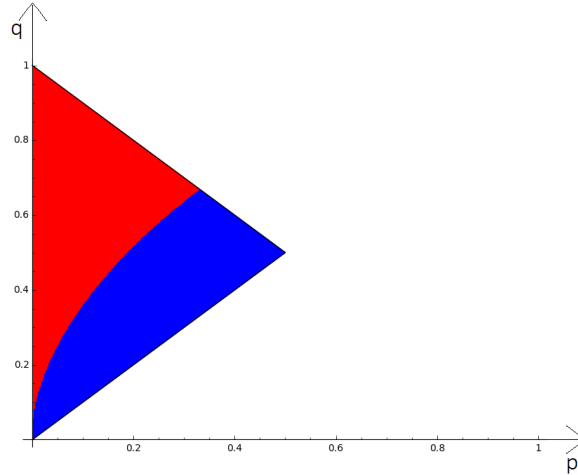


Fig. 1. Each color corresponds to a 3-stable decision criterion for the BACs.

components of the stable region  $\mathcal{R}_3^{st}$ ) and there are three unstable decision criteria corresponding to the curves  $p = 0$  (the Z-channel),  $p = q$  (the BSC) and the curve that separates the two connected components of the stable region. By considering the expression for the regions given in Theorem 2 we can find  $(p, q)$  for some BACs which are representatives of these five criteria in the above order, let's say  $(\frac{1}{7}, \frac{2}{7})$ ,  $(\frac{1}{7}, \frac{4}{7})$ ,  $(0, \frac{1}{7})$ ,  $(\frac{1}{7}, \frac{1}{7})$  and  $(\frac{1}{7}, \frac{3}{7})$ . We can then assert that any BAC is 3-equivalent to one of these five channels.

We proceed similarly with the cases  $n = 4$  and  $n = 5$ , observing three decision criteria for  $n = 4$  (Figure 2) and five decision criteria for  $n = 5$  (Figure 3). They correspond to the stable decision criteria and the curves separating these regions correspond to the four and six different unstable decision criteria for  $n = 4$  and  $n = 5$  respectively (see Theorem 2 in the next section).

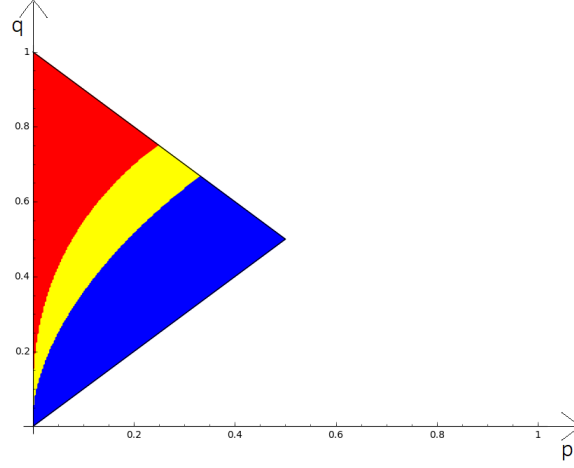


Fig. 2. Each color corresponds to a 4-stable decision criterion for the BACs.

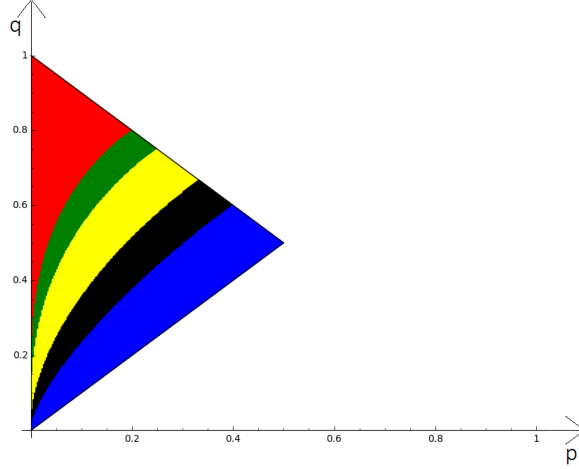


Fig. 3. Each color corresponds to a 5-stable decision criterion for the BACs.

### III. DETERMINING THE $n$ -DECISION CRITERIA FOR THE BACs

We start introducing a function which plays a fundamental role in describing the regions which determine the decision criteria for the BACs. This function also induces a natural distance between binary asymmetric channels as it is seen in Section IV.

**Definition 5.** Let  $S : \mathcal{T} \rightarrow [0, 1]$  be the function given by

$$S(p, q) = \frac{\ln(1-p) - \ln(q)}{\ln(1-q) - \ln(p)},$$

if  $p \neq 0$ , and  $S(p, q) = 0$  if  $p = 0$ . We refer to this function as the BAC-function.

It is easy to check that in fact the image of  $S$  is contained in the interval  $[0, 1]$  where the values 0 and 1 are attained by the extremes cases  $p = 0$  and  $p = q$ , respectively. Besides, this function is continuous in the connected set  $\mathcal{T}$  and therefore we have  $S(\mathcal{T}) = [0, 1]$ .

Let  $a$  and  $b$  be integers with  $0 \leq a \leq b$ . For  $(p, q) \in \mathcal{T}$  we have

$$p^a(1-p)^b \geq q^b(1-q)^a \Leftrightarrow S(p, q) \geq \frac{a}{b}, \quad (1)$$

where equality holds in the left side if and only if it holds in the right side. The next lemma is a direct consequence of the above relation.

**Lemma 1.** Let  $a$  and  $b$  be natural numbers with  $a \leq b$ ,  $a + b \leq n$  and  $\eta = n - (a + b) \geq 0$ . Consider the words  $x = 1^{a+\eta}0^b$ ,  $y = 0^n$ ,  $z = 0^\eta 1^{a+b} \in \mathbb{F}_2^n$ . Then  $S(p, q) \leq a/b$  if and only if  $\Pr(x|y) \leq \Pr(x|z)$ , where equality corresponds to equality. In particular, if  $S(p_0, q_0) \leq a/b$  and  $S(p_1, q_1) > a/b$  the channels  $\text{BAC}(p_0, q_0)$  and  $\text{BAC}(p_1, q_1)$  are not  $n$ -equivalent.

When we consider  $p$  and  $q$  as variables, the entries of the  $n$ -transition matrix  $M_n(p, q)$  are polynomials in the variables  $p$  and  $q$ . In fact, the entry of  $M_n(p, q)$  corresponding to the conditional probability  $\Pr(x|y)$  ( $x, y \in \mathbb{F}_2^n$ ) is equal to the polynomial  $f(p, q) = p^a(1-p)^b q^c(1-q)^d$  where  $a, b, c$  and  $d$  correspond to the number of indices  $i$  for which  $(x_i, y_i)$  is equal to  $(1, 0)$ ,  $(0, 0)$ ,  $(0, 1)$  and  $(1, 1)$ , respectively. In particular we have  $a + b + c + d = n$  and the Hamming weight of  $x$  is equal to  $a + d$ , which we will refer also as the weight of  $f$  and denote by  $\omega(f)$ . We remark that two polynomials in the same row of  $M_n(p, q)$  have the same weight. By the previous consideration we define the following sets of bivariate polynomials:

- $\mathcal{P}^n = \{f(p, q) = p^a(1-p)^b q^c(1-q)^d : a, b, c, d \geq 0, a + b + c + d = n\}$ .
- $\mathcal{P}_k^n = \{f \in \mathcal{P}^n : \omega(f) = k\}$  for  $0 \leq k \leq n$ .

To determine the  $n$ -decision criterion corresponding to a BAC we only need to do comparisons between values in the same row of its  $n$ -transition matrix, this means comparisons of values of polynomials belonging to  $\mathcal{P}_k^n$  for some  $k$ ,  $0 \leq k \leq n$ . The next lemma describes the stable region in terms of these sets.

**Lemma 2.** The  $n$ -stable region  $\mathcal{R}_n^{st}$  is given by

$$\mathcal{R}_n^{st} = \bigcap_{k=0}^n \bigcap_{\substack{f, g \in \mathcal{P}_k^n \\ f \neq g}} \{(p, q) \in \mathcal{T} : f(p, q) \neq g(p, q)\}.$$

*Proof.* We denote by  $\widehat{\mathcal{R}}_n^{st}$  the set on the right side of the above equality. This set is open in  $\mathcal{T}$  (since it is a finite intersection of open sets in  $\mathcal{T}$ ), therefore if  $(p_0, q_0) \in \widehat{\mathcal{R}}_n^{st}$  there is a ball  $B_0$  with center at this point such that  $B_0 \cap \mathcal{T} \subseteq \widehat{\mathcal{R}}_n^{st}$ . Since  $B_0 \cap \mathcal{T}$  is connected, the signal of  $f(p, q) - g(p, q)$  does not depends on  $(p, q) \in B_0 \cap \mathcal{T}$  (whenever  $f, g \in \mathcal{P}_k^n$  for some  $k$ ) and the same occurs with their associated decision criteria, so  $\widehat{\mathcal{R}}_n^{st} \subseteq \mathcal{R}_n^{st}$ . To prove the other inclusion we suppose by contradiction that there exists  $(p, q) \notin \widehat{\mathcal{R}}_n^{st}$  verifying  $(p, q) \in \mathcal{R}_n^{st}$ . Then, there exist two distinct polynomials  $f, g \in \mathcal{P}_k^n$  for some  $k : 0 \leq k \leq n$  and a ball  $B_0$  centered at  $(p_0, q_0)$  such that  $f(p_0, q_0) = g(p_0, q_0)$  and every point in  $B_0 \cap \mathcal{T}$  determines the same decision criterion, in particular  $f(p, q) = g(p, q)$  for all  $(p, q) \in B_0 \cap \mathcal{T}$ . Since  $B_0 \cap \mathcal{T}$  has interior points and two polynomials that coincide in an open set must be equal, we have  $f = g$  which is a contradiction.  $\square$

As we will prove next, the main property of the BAC-function from the point of view of this work, is that the curves which separate the regions corresponding to the stable and unstable criteria are level curves of this function associated with rational values. First, we prove some lemmas.

**Lemma 3.** Let  $(p, q) \in \mathcal{T}$  be an  $n$ -unstable point for the BACs. Then  $S(p, q) = \frac{a}{b}$  where  $a, b$  are integers verifying  $a \geq 0, b \geq 1, a \leq b, \gcd(a, b) = 1$  and  $a + b \leq n$ .

*Proof.* By Lemma 2, if  $(p_0, q_0) \in \mathcal{T}$  is an  $n$ -unstable point for the BACs then there exists distinct polynomials  $f_1, f_2 \in \mathcal{P}_k^n$  for some  $k : 0 \leq k \leq n$  such that  $f_1(p_0, q_0) = f_2(p_0, q_0)$ . We write  $f_i(p, q) = p^{a_i}(1-p)^{b_i} q^{c_i}(1-q)^{d_i}$  with  $a_i + b_i + c_i + d_i = n$  and  $a_i + d_i = k$  for  $i = 1, 2$ . Without loss of generality we suppose  $a_1 \geq a_2$ . Let  $a := a_1 - a_2 = d_2 - d_1$  and  $b := b_1 - b_2 = c_2 - c_1$ , then we have

$$\frac{f_1(p, q)}{f_2(p, q)} = \left( \frac{p}{1-q} \right)^a \left( \frac{1-p}{q} \right)^b,$$

or, equivalently,

$$p^a(1-p)^b = \left( \frac{f_1(p, q)}{f_2(p, q)} \right) \cdot q^b(1-q)^a. \quad (2)$$

Evaluating the Equation 2 for  $(p, q) = (p_0, q_0)$  and using the relation (1) we have  $S(p_0, q_0) = \frac{a}{b}$ . By our assumption we have  $a \geq 0$ , since  $S(\mathcal{T}) = [0, 1]$  then  $b \geq 1$  and  $a + b = a_1 - a_2 + b_1 - b_2 \leq a_1 + b_1 \leq n$ ; simplifying common factors if necessary we can assume  $\gcd(a, b) = 1$ .  $\square$

Based on the above result, we introduce the following definition.

**Definition 6.** The weight of a non-negative rational  $r$  (denoted by  $\omega(r)$ ), is the sum of its numerator and denominator in the reduced expression of  $r$ . An  $n$ -critical value for the BAC-function  $S$  (where  $n \geq 2$ ) is a rational number  $r \in [0, 1]$  with  $\omega(r) \leq n$ . The set of all  $n$ -critical values for  $S$  is denoted by  $\mathcal{D}_n$ .

**Corollary 1.** If we write the set of the  $n$ -critical values for  $S$  as  $\mathcal{D}_n = \{r_0 = 0 < r_1 < \dots < r_t = 1\}$  and denote by  $R_n(r_i) = \{(p, q) \in \mathcal{T} : r_i < S(p, q) < r_{i+1}\}$  then  $R_n(r_i) \subseteq \mathcal{R}_n^{st}$  for  $0 \leq i < t$ .

**Lemma 4.** Let  $(p, q) \in \mathcal{T}$  and  $M_n(p, q)$  be the  $n$ -transition matrix for the channel  $\text{BAC}(p, q)$  (seeing as an element of  $\mathbb{R}^{n^2}$ ). The function  $\phi : \mathcal{R}_n^{st} \rightarrow \mathbb{R}^{n^2}$  given by  $\phi(p, q) = M_n(p, q)^*$  is continuous.

*Proof.* Let  $(p_0, q_0) \in \mathcal{R}_n^{st}$  and  $f, g \in \mathcal{P}_k^n$  for some  $k$ ,  $0 \leq k \leq n$  with  $f \neq g$ . By Lemma 2 we have  $f(p_0, q_0) \neq g(p_0, q_0)$ , then there exists  $\epsilon = \epsilon(f, g) > 0$  such that the sign of  $f(p, q) - g(p, q)$  does not depend on  $(p, q) \in B((p_0, q_0), \epsilon) \cap \mathcal{T}$ . If  $\epsilon = \min\{\epsilon(f, g) : f, g \in \mathcal{P}_k^n, f \neq g, 0 \leq k \leq n\}$  and denoting by  $B$  the  $\epsilon$ -ball centered at  $(p_0, q_0)$  we have that  $f(p, q) > g(p, q) \Leftrightarrow f(p_0, q_0) > g(p_0, q_0)$  for all  $(p, q) \in B \cap \mathcal{T}$  and for all  $f, g \in \mathcal{P}_k^n, f \neq g, 0 \leq k \leq n$ . Therefore  $M_n(p, q)^* = M_n(p_0, q_0)^*$ , so  $\phi$  is locally constant and in particular continuous.  $\square$

**Lemma 5.** *Let  $S$  be the BAC-function. Then  $S^{-1}(I)$  is a connected set for every interval  $I \subseteq [0, 1]$ .*

*Proof.* Let  $\tau \in (0, 1)$  and  $g_\tau : [0, \frac{\tau}{2}] \rightarrow [0, 1]$  be the function given by  $g_\tau(p) = S(p, \tau - p)$ . We affirm that it is increasing (in the variable  $p$ ). To prove this we consider  $s \in [0, 1]$  and the function  $f_s(p) = p^s(1-p) - q(1-q)^s$  (where  $q = \tau - p$ ). Since  $1 - q > p$  we have  $p^{s-1} > (1-q)^{s-1}$  and  $(1-q)^s > p^s$ , therefore

$$\begin{aligned} f'_s(p) &= s \left( p^{s-1}(1-p) - q(1-q)^{s-1} \right) + (1-q)^s - p^s \\ &> s(1-q)^{s-1}(1-p-q) + (1-q)^s - p^s > 0. \end{aligned}$$

Since  $f_s(0) = -q(1-q)^s < 0$  and  $f_s(\tau/2) = \tau/2(1-\tau/2)((\tau/2)^{s-1} - (1-\tau/2)^{s-1}) > 0$  (because  $1-\tau/2 > \tau/2$  and  $s-1 < 0$ ), then there is an unique  $p \in (0, \tau/2)$  such that  $f_s(p) = 0$ , or equivalently, such that  $g_\tau(p) = S(p, \tau - p) = s$ . Therefore  $g_\tau : [0, \frac{\tau}{2}] \rightarrow [0, 1]$  is increasing since it is a continuous bijection with  $g_\tau(0) = 0$  and  $g_\tau(\tau/2) = 1$ . Let  $I \subseteq [0, 1]$  be an interval and  $(p_0, q_0), (p_1, q_1) \in \mathcal{T}$  be two points in  $S^{-1}(I)$ . We denote by  $s_i = S(p_i, q_i)$  and  $\tau_i = p_i + q_i$  for  $i = 0, 1$ . Without loss of generality we assume  $s_0 \leq s_1$ . Since  $g_{\tau_0}(p_0) = S(p_0, q_0) = s_0 < s_1$  there exists  $t_0 > 0$  such that  $g_{\tau_0}(p_0 + t_0) = s_1$  and  $g_{\tau_0}(p_0 + t) \in (s_0, s_1)$  for all  $t \in (0, t_0)$  (in particular  $(p_0 + t, q_0 - t) \in S^{-1}(I)$  for all  $t \in [0, t_0]$ ). Taking the path  $\alpha : [0, t_0] \rightarrow \mathcal{T}$  given by  $\alpha(t) = (p_0 + t, q_0 - t)$  and  $\beta$  the segment of curve in  $S^{-1}(s_1)$  from  $(p_0 + t_0, q_0 - t_0)$  to  $(p_1, q_1)$ , we have that the concatenation path  $\beta * \alpha$  is a path connecting  $(p_0, q_0)$  with  $(p_1, q_1)$ . Therefore  $S^{-1}(I)$  is path-connected and then connected.  $\square$

**Lemma 6.** *Let  $(p_0, q_0)$  and  $(p_1, q_1)$  be two points in  $\mathcal{T}$ . If  $S(p_0, q_0) = S(p_1, q_1)$  then  $BAC(p_0, q_0) \sim_n BAC(p_1, q_1)$  for all  $n \geq 1$ .*

*Proof.* We suppose by contradiction that  $(p_0, q_0)$  and  $(p_1, q_1)$  are not  $n$ -equivalent, so there exists two polynomials  $f, g \in \mathcal{P}_k^n$  for some  $k$  verifying  $f(p_0, q_0) < g(p_0, q_0)$  and  $f(p_1, q_1) \geq g(p_1, q_1)$ . As in the proof of Lemma 3, the polynomials  $f$  and  $g$  must verify an equation similar to Equation (2):

$$p^a(1-p)^b = \left( \frac{f(p, q)}{g(p, q)} \right) \cdot q^b(1-q)^a \quad (3)$$

for some integers  $a \geq 0$  and  $b \geq 1$  satisfying  $a \leq b$  and  $a+b \leq n$ . We consider a path  $\alpha$  contained in the curve  $S^{-1}(r)$  connecting  $(p_0, q_0)$  with  $(p_1, q_1)$ , since  $f/g < 1$  in  $(p_0, q_0)$  and  $f/g \geq 1$  in  $(p_1, q_1)$ , by continuity there exists an intermediate point  $(p_2, q_2)$  in  $\alpha$  for which  $f/g = 1$ . Evaluating Equation (3) in  $(p, q) = (p_2, q_2)$  and using the relation (1) we have  $S(p_2, q_2) = a/b$ . Since  $(p_2, q_2)$  belongs to  $\alpha$  which is contained in the level curve  $S^{-1}(r)$  we have  $S(p_2, q_2) = r$ , then  $r = a/b$ . Substituting  $(p, q) = (p_0, q_0)$  in Equation (3) and using the relation 1 we have  $r = S(p_0, q_0) < a/b$  which is a contradiction. Therefore  $(p_0, q_0)$  and  $(p_1, q_1)$  must be  $n$ -equivalent.  $\square$

Now we are ready to state the main result of this paper.

**Theorem 2.** *Let  $S : \mathcal{T} \rightarrow [0, 1]$  be the BAC-function*

$$S(p, q) = \frac{\ln(1-p) - \ln(q)}{\ln(1-q) - \ln(p)},$$

*for  $p \neq 0$ ,  $S(0, q) = 0$  and  $\mathcal{D}_n = \{0 = r_0 < r_1 < \dots < r_{t_n} = 1\}$  its set of  $n$ -critical values ( $n \geq 2$ ). We consider the level curves  $\gamma_i = S^{-1}(r_i)$  for  $0 \leq i \leq t_n$  and the regions  $R_n(r_i) = \{(p, q) \in \mathcal{T} : r_i < S(p, q) < r_{i+1}\}$  for  $0 \leq i < t_n$ . Then*

$$t_n = 1 + \frac{1}{2} \sum_{k=3}^n \varphi(k) \quad (4)$$

*where  $\varphi$  denotes the Euler's totient function and there are exactly  $t_n$  stable decision criteria of order  $n$  for the BACs, which are given by  $\{R_n(r_i) : 0 \leq i < t_n\}$  and exactly  $t_n + 1$  unstable decision criteria of order  $n$  for the BACs given by  $\{\gamma_i : 0 \leq i \leq t_n\}$ .*

*Proof.* Consider  $\mathcal{T}$  written as the disjoint union

$$\mathcal{T} = \biguplus_{i=0}^{t-1} R_n(r_i) \uplus \biguplus_{i=0}^t \gamma_i.$$

We have to prove that each  $R_n(r_i)$  and each  $\gamma_i$  is a  $n$ -decision criterion (i.e. an equivalent class in  $\Delta_n$ ). If  $(p_0, q_0) \in \gamma_i$  for some  $i : 0 \leq i \leq t$ , by Lemma 6 and Lemma 1 we have  $BAC(p_1, q_1) \sim_n BAC(p_0, q_0)$  if and only if  $(p_1, q_1) \in \gamma_i$ , then each  $\gamma_i$  is a decision criterion. We consider now a point  $(p_0, q_0) \in R_n(r_i)$  for some  $i : 0 \leq i < t$  and  $(p_1, q_1) \in \mathcal{T} \setminus R_n(r_i)$ . If

$(p_1, q_1) \in \gamma_j$  for some  $j$ , since each  $\gamma_j$  is a decision criterion we have that  $BAC(p_1, q_1)$  and  $BAC^n(p_0, q_0)$  are not  $n$ -equivalent. Otherwise  $(p_1, q_1) \in R_n(r_j)$  for some  $j : 0 \leq j < n, j \neq i$  and there exists  $r_k \in \mathcal{D}_n$  such that  $S(p_0, q_0) < r_k < S(p_1, q_1)$  or  $S(p_1, q_1) < r_k < S(p_0, q_0)$ . In both cases, by Lemma 1 the channels  $BAC(p_1, q_1)$  and  $BAC^n(p_0, q_0)$  are not  $n$ -equivalent. It only remains to prove that if  $(p_1, q_1) \in R_n(r_i)$  the channels  $BAC(p_1, q_1)$  and  $BAC(p_0, q_0)$  are  $n$ -equivalent. We consider the function  $\phi : \mathcal{R}_n^{st} \rightarrow \mathbb{R}^{n^2}$  given by  $\phi(p, q) = M_n(p, q)^*$  where  $M_n(p, q)$  denotes the  $n$ -transition matrix for the channel  $BAC(p, q)$ . By Lemma 5,  $R_n(r_i)$  is a connected set (since  $R_n(r_i) = S^{-1}(I)$  for  $I = (r_i, r_{i+1})$ ), and by Lemma 3,  $R_n(r_i)$  is contained in the stable region  $\mathcal{R}_n^{st}$ . By Lemma 4, the set  $\phi(R_n(r_i)) \subseteq \mathcal{M}_n(\mathbb{Z})$  is connected and since  $\mathcal{M}_n(\mathbb{Z})$  is discrete, there exists  $M \in \mathcal{M}_n(\mathbb{Z})$  such that  $\phi(R_n(r_i)) = \{M\}$ . Therefore  $M_n(p_1, q_1)^* = M_n(p_0, q_0)^* = M$  and  $BAC^n(p_1, q_1) \sim BAC^n(p_0, q_0)$ .

Since the sets  $R_n(r_i)$  are open for  $0 \leq i < t_n$  and the sets  $\gamma_i$  have empty interior they correspond to the stable and unstable criteria respectively. To derive the formula for  $t_n$  we consider the decomposition into disjoint sets:  $\mathcal{D}_n = \biguplus_{k=1}^n \mathcal{D}_k^o$  where  $\mathcal{D}_k^o = \{a/b \in \mathbb{Q}^+ : a \leq b, \gcd(a, b) = 1, a + b = k\}$ . We have  $\#\mathcal{D}_k^o = 1$  for  $k = 1, 2$ . For  $k \geq 3$  if  $a/b \in \mathcal{D}_k^o$  then  $a < b$  and in this case:

$$\begin{aligned} \#\mathcal{D}_k^o &= \frac{1}{2} \cdot \#\{(a, b) \in \mathbb{N}^2 : \gcd(a, b) = 1, a + b = k\} \\ &= \frac{1}{2} \cdot \#\{(a, b) \in \mathbb{N}^2 : \gcd(a, k) = 1, a + b = k\} \\ &= \frac{1}{2} \cdot \#\{a \in \mathbb{N} : \gcd(a, k) = 1, a \leq k\} = \frac{1}{2} \cdot \varphi(k). \end{aligned}$$

Then, for  $n \geq 3$  we have:

$$\begin{aligned} t_n &= \#\mathcal{D}_n - 1 = \sum_{k=1}^n \#\mathcal{D}_k^o - 1 \\ &= 2 + \frac{1}{2} \cdot \sum_{k=3}^n \varphi(k) - 1 = 1 + \frac{1}{2} \cdot \sum_{k=3}^n \varphi(k). \end{aligned}$$

For  $n = 2$  we have  $\#\mathcal{D}_2 = \#\{(0, 1), (1, 1)\} = 2$  and the above formula also holds in this case.  $\square$

**Corollary 2.** Let  $(p_0, q_0), (p_1, q_1) \in \mathcal{T}$ . The channels  $BAC(p_0, q_0)$  and  $BAC(p_1, q_1)$  are  $\infty$ -equivalent (i.e.  $n$ -equivalent for all  $n$ ) if and only if  $S(p_0, q_0) = S(p_1, q_1)$ .

*Proof.* If  $S(p_0, q_0) < S(p_1, q_1)$  there exists a rational number  $r \in \mathcal{D}_n$  for some  $n \geq 1$  large enough such that  $S(p_0, q_0) < r < S(p_1, q_1)$ , then the channels  $BAC(p_0, q_0)$  and  $BAC(p_1, q_1)$  are not  $n$ -equivalent. The converse is consequence of Lemma 6.  $\square$

**Corollary 3.** A point  $(p, q) \in \mathcal{T}$  is a stable point of order  $n$  for all  $n \geq 1$  if and only if  $S(p, q)$  is an irrational number.

Using the average order formula for the Euler's totient function  $\varphi$  (see for example Theorem 3.7 of [21]) we obtain the following corollary.

**Corollary 4.** The number of stable decision criteria of order  $n$  for the BACs grows quadratically with  $n$ . More explicitly, it is given by  $\frac{3}{\pi^2} \cdot n^2 + O(n \cdot \ln n)$ .

**Example 1.** For  $n = 5$  (see Figure 3), we have the set of critical values  $\mathcal{D}_5 = \{0, 1/4, 1/3, 1/2, 2/3, 1\}$  which correspond to the level curves of the BAC-function describing the unstable sets. Those curves can be seen from left to right according to the order in  $\mathcal{D}_5$ . The five stable regions are the ones bounded by these curves.

**Example 2.** For  $n = 9$ , we have  $t_9 = 29$  decision regions, 15 instable, associated to the critical set  $\mathcal{D}_9 = \{0, 1/8, 1/7, 1/6, 1/5, 1/4, 2/7, 1/4/5, 1\}$  and 14 stable, situated between the level curves of the BAC-function attached to values in  $\mathcal{D}_9$ .

To conclude this section we present some situations of decoding problems that can be solved using Theorem 2.

**Example 3.** Let  $W$  be a memoryless binary asymmetric channel. Suppose that a series of measurements was performed in order to obtain the transition probabilities  $(p, q)$  for  $W$  and the following values were obtained:  $p = 0.212$  and  $q = 0.531$  with a possible error of at most  $\varepsilon = 0.001$  in both probabilities. We want to determine the maximum possible length  $N$  such that there is no risk of mismatching when we perform MLD on  $W$  with respect to  $(p, q) = (0.212, 0.531)$  and binary codes with block length  $n \leq N$ . In term of  $n$ -equivalence of channels this means to find the maximum value of  $N$  such that the channels  $W$  and  $BAC(0.212, 0.531)$  are  $n$ -equivalent for all  $n \leq N$ . By our assumption, the real transition probabilities  $(p_0, q_0)$  for  $W$  (which is unknown) belongs to the square  $I = [0.211, 0.213] \times [0.530, 0.532]$ . By Theorem 2, the problem is reduced to find the maximum value of  $N$  such that the critical set  $\mathcal{D}_N = \{\frac{a}{b} : 0 \leq b \leq a, a + b \leq N\}$  has empty intersection with the interval  $S(I) = [S(0.211, 0.532), S(0.213, 0.530)] = [0.4947499\ldots, 0.4995337]$ . The maximum value of  $N$  is 144 since  $47/95 = 0.4947368\ldots$  and  $1/2 = 0.5$  are consecutive elements in  $\mathcal{D}_{144}$  but  $0.4947499 < 48/97 < 0.4995337$  with  $48 + 97 = 145$ .



Thus, if we implement MLD with respect to the measured approximated value  $(p, q) = (0.212, 0.531)$ , restricted to codes with block length  $n \leq 144$ , there is no risk of mismatched decoding since in this case  $BAC(p, q) \underset{n}{\sim} BAC(p_0, q_0)$ .

**Example 4.** Let  $W = BAC(p_0, q_0)$  with  $p_0 = 0.314$ ,  $q_0 = 0.594$  and suppose that only binary codes with block length at most  $n = 32$  are considered. We call  $\varepsilon > 0$  an admissible error if  $BAC(p, q) \underset{20}{\sim} W$  for all  $(p, q) \in \mathcal{T}$  such that  $|p - p_0| < \varepsilon$  and  $|q - q_0| < \varepsilon$ . We want to find the greatest admissible error. For each  $\varepsilon > 0$  we consider the box  $I_\varepsilon = [p_0 - \varepsilon, p_0 + \varepsilon] \times [q_0 - \varepsilon, q_0 + \varepsilon]$ . By Theorem 2, the problem is equivalent to find the greatest  $\varepsilon > 0$  such that  $S(I_\varepsilon) \cap \mathcal{D}_{32} = \emptyset$ . Since  $5/9$  and  $9/16$  are consecutive elements of  $\mathcal{D}_{32}$  satisfying  $5/9 < S(p_0, q_0) = 0.56039... < 9/16$ , and the endpoints of  $S(I_\varepsilon)$  are  $S(p_0 - \varepsilon, q_0 + \varepsilon)$  and  $S(p_0 + \varepsilon, q_0 - \varepsilon)$ , it suffices to find  $\varepsilon_0, \varepsilon_1 > 0$  such that  $S(p_0 - \varepsilon_0, q_0 + \varepsilon_0) = 5/9$  and  $S(p_0 + \varepsilon_1, q_0 - \varepsilon_1) = 9/16$ , and take the minimum of these two values. By direct calculation we obtain  $\varepsilon_0 = 0.0019779637..$  and  $\varepsilon_1 = 0.0008585765..$ , thus  $\varepsilon = \min\{\varepsilon_0, \varepsilon_1\} = 0.0008585765...$

#### IV. FURTHER REMARKS

##### A. On the BAC-function and the parameter space for the BACs

To study the different  $n$ -decision criteria for the BACs, we choose the parameter space  $\mathcal{T} = \{(p, q) \in [0, 1] : p + q < 1, 0 \leq p \leq q\} \setminus \{(0, 0)\}$  and use the BAC-function to describe the regions determined by these criteria. In the first part of this section we discuss what happens when we remove the restriction  $p + q < 1$  and  $0 \leq p \leq q$  and what is the role of the BAC-function in these cases. Next, we show how to obtain a natural distance between BACs in such a way that the BAC-function measures how far a channel is from the binary symmetric channel, in this sense the BAC-function provides a measure of the asymmetry of the channel.

Consider a BAC with transition probabilities  $p, q \in [0, 1]^2$ . A BAC is *reasonable* (in the sense of [18]) if their transition probabilities verify  $\Pr(0|0) > \Pr(0|1)$  and  $\Pr(1|1) > \Pr(1|0)$ . Note that this is equivalent to the condition  $p + q < 1$ . It is not difficult to check, using Equation (2)), the following facts:

- $p + q < 1 \Leftrightarrow \Pr(x|x) > \Pr(x|y)$ ,  $\forall x, y \in \mathbb{F}_2^n, x \neq y$ ;
- $p + q > 1 \Leftrightarrow \Pr(x|x) < \Pr(x|y)$ ,  $\forall x, y \in \mathbb{F}_2^n, x \neq y$ ;
- $p + q = 1 \Leftrightarrow \Pr(x|x) = \Pr(x|y)$ ,  $\forall x, y \in \mathbb{F}_2^n, x \neq y$ .

As a consequence, we have that if two channels  $BAC(p, q)$  and  $BAC(p', q')$  are  $n$ -equivalent, then the sign of  $1 - p - q$  and  $1 - p' - q'$  is the same. The case  $p + q = 1$  corresponds to the completely noisy channel, which are not interesting from the coding/decoding point of view, since  $\Pr(x|y) = \Pr(x|z)$  for all  $x, y, z \in \mathbb{F}_2^n$ . These channels are  $n$ -equivalent among them, for all  $n \geq 1$ . The case  $p + q < 1$  can be decomposed into two regions  $\mathcal{T}$  and  $\mathcal{T}' = \{(p, q) \in [0, 1]^2 : p + q < 1, 0 \leq q \leq p\} \setminus \{(0, 0)\}$  which are symmetric one to the other (via the map  $(p, q) \mapsto (q, p)$ ).

We observe that the BAC-function  $S$  can be extended to a function  $\widehat{S} : \mathcal{T} \cup \mathcal{T}' \rightarrow [0, +\infty]$  defining  $\widehat{S}(q, p) = 1/S(p, q)$  for  $(q, p) \in \mathcal{T}'$  if  $pq \neq 0$  and  $\widehat{S}(p, 0) = +\infty$ . This extension is continuous and verifies  $\widehat{S}(\mathcal{T}') = [1, +\infty]$ . By relation (1),  $S(p, q) \leq 1$  if and only if  $p(1 - p) \leq q(1 - q)$ , if and only if  $p^{n-1}(1 - p) \leq p^{n-2}q(1 - q)$ . Therefore, for  $x = 1^{n-1}0$ ,  $y = 0^n$  and  $z = 0^{n-2}1^2$  we have:

- $\Pr(x|y) < \Pr(x|z)$  if  $(p, q) \in \mathcal{T}$  with  $p \neq q$ ,
- $\Pr(x|y) > \Pr(x|z)$  if  $(p, q) \in \mathcal{T}'$  with  $p \neq q$  and
- $\Pr(x|y) > \Pr(x|z)$  if  $p = q$ .

Thus, we conclude that the triangles  $\mathcal{T}$  and  $\mathcal{T}'$  have no common criteria decision except for those points corresponding to the BSC. Since the  $i$ -th row of the  $n$ -transition matrix  $M_n(p, q)$  is just the  $(2^n + 1 - i)$ -th row of  $M_n(q, p)$  in reverse order, then  $BAC(p, q) \underset{n}{\sim} BAC(p', q')$  if and only if  $BAC(q, p) \underset{n}{\sim} BAC(q', p')$  for all  $(p, q), (p', q') \in \mathcal{T}$ . By the above consideration we have the following proposition.

**Proposition 2.** Let  $S : \mathcal{T} \cup \mathcal{T}' \rightarrow [0, +\infty]$  be the BAC-function defined as above and  $\mathcal{D}_n = \{0 = r_0 < r_1 < \dots < r_{t_n} = 1\}$  its set of  $n$ -critical values ( $n \geq 2$ ) where  $t_n = 1 + \frac{1}{2} \sum_{k=3}^n \varphi(k)$ . There are exactly  $2t_n$  stable decision criteria of order  $n$  for the  $BAC(p, q)$  with  $(p, q) \in \mathcal{T} \cup \mathcal{T}'$  and exactly  $2t_n - 1$  unstable decision criteria of order  $n$ . The  $n$ -stable criteria are given by  $R_n(r_i) = \{(p, q) \in \mathcal{T} : r_i < S(p, q) < r_{i+1}\}$  and  $R_n(r_i^{-1}) = \{(p, q) \in \mathcal{T}' : r_{i+1}^{-1} < S(p, q) < r_i^{-1}\}$  for  $0 \leq i < t_n$ . The  $n$ -unstable criteria are given by the level curves  $S^{-1}(r_i)$  and  $S^{-1}(r_i^{-1})$  for  $0 \leq i \leq t_n$ .

The function  $S$  is constant when restricted to a criterion  $\mathcal{A} \subseteq \mathcal{T} \cup \mathcal{T}'$ . Thus it defines an injective function on the equivalent classes  $S : \Delta_n \rightarrow [0, +\infty]$  such that  $S(\mathcal{A}) := S(p, q)$  for any  $(p, q) \in \mathcal{A}$ . If  $\Delta_n^*$  denotes the set of all  $n$ -decision criteria for the BACs except for those corresponding to the Z-channels (i.e. when  $pq = 0$ ), then the function  $d : \Delta_n^* \times \Delta_n^* \rightarrow [0, +\infty)$  given by  $d(\mathcal{A}, \mathcal{B}) = |\ln S(\mathcal{A}) - \ln S(\mathcal{B})|$  defines a metric in the  $n$ -decision criteria space for the BACs. In particular if  $\mathcal{B}$  denotes the criterion corresponding to the BSC then  $d(\mathcal{A}, \mathcal{B}) = |\ln S(\mathcal{A})|$  can be interpreted as a measure of how asymmetric a channel is.

Since the ordered form of the  $n$ -transition matrix associated with the completely noisy channels ( $p + q = 1$ ) is the null matrix (because in this case  $\Pr(x|y) = p^{w_H(x)}(1 - p)^{n - w_H(x)}$  depends only on the Hamming weight of  $x$  and not on  $y$ ) and this is the

only situation for which it happens, then the points in the line  $p + q = 1$  correspond to a single criterion when considering BACs with  $(p, q) \in [0, 1]^2$ . If  $p + q < 1$  and  $p' + q' > 1$ , the channels  $\text{BAC}(p, q)$  and  $\text{BAC}(p', q')$  are not  $n$ -equivalent, since the first is reasonable and the last is not.

Let  $T^- = \mathcal{T} \cup \mathcal{T}' = \{(p, q) \in [0, 1]^2 : p + q < 1\} \setminus \{(0, 0)\}$  and  $T^+ = \{(p, q) \in [0, 1]^2 : p + q > 1\} \setminus \{(1, 1)\}$ . The involution  $\phi(p, q) = (1 - q, 1 - p)$  maps the triangle  $T^-$  into  $T^+$  and the curves  $\gamma_{a/b} : p^a(1 - p)^b = q^b(1 - q)^a$  into itself. Moreover, the  $i$ -th rows of  $M_n(p, q)$  and  $M_n(1 - q, 1 - p)$  are the same but in reverse order. Therefore for  $(p, q), (p', q') \in T^+$ , the channels  $\text{BAC}(p, q)$  and  $\text{BAC}(p', q')$  are  $n$ -equivalent if and only if  $\text{BAC}(1 - q, 1 - p)$  and  $\text{BAC}(1 - q', 1 - p')$  are  $n$ -equivalent.

We conclude that if we consider BACs with parameters  $(p, q) \in [0, 1]^2$ , the  $n$ -stable criteria are the regions bounded by the edges of the square  $[0, 1]^2$  and the curves  $\gamma_{a/b}$  where  $a/b$  is a positive rational number with  $a, b \in \mathbb{Z}^+$ ,  $\gcd(a, b) = 1$  and  $a + b \leq n$ . We also remark that the level curves  $\gamma_{a/b}$  contain the line  $p + q - 1 = 0$  but if we divide the equation  $p^a(1 - p)^b - q^b(1 - q)^a = 0$  which defines  $\gamma_{a/b}$  (see Equation (1)) by  $p + q - 1$  we obtain irreducible curves  $\hat{\gamma}_{a/b}$  (see Figure 4).

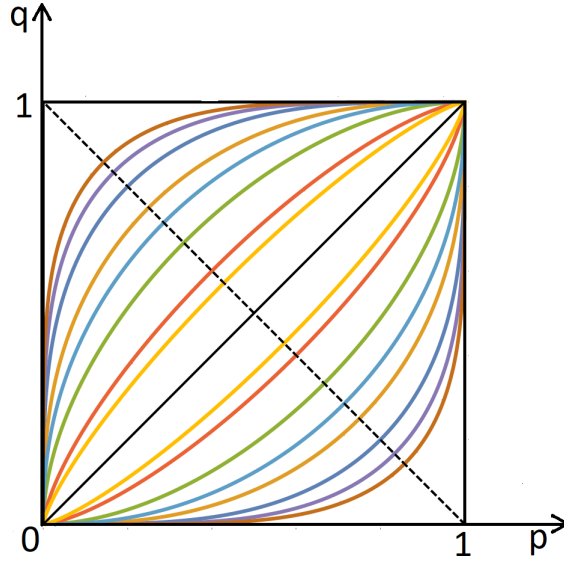


Fig. 4. Decision criteria (regions) of order 7 for the BACs with  $(p, q)$  in the unit square: the line  $p + q = 1$  (dotted) and the curves  $\hat{\gamma}_q$  and  $\hat{\gamma}_{q-1}$  for  $q = 1/6$  (brown),  $1/5$  (violet),  $1/4$  (blue),  $1/3$  (orange),  $2/5$  (sky blue),  $1/2$  (green),  $2/3$  (red),  $3/4$  (yellow) and  $1$  (black). Curves corresponding to reciprocal values have the same color.

### B. The most probable BACs are those next to the BSC

By the previous discussion, without loss of generality we can restrict our parameter space to  $\mathcal{T} = \{(p, q) \in [0, 1]^2 : p + q < 1, p \leq q\} \setminus \{(0, 0)\}$ . Let  $\mathcal{A} \subseteq \mathcal{T}$  be a  $n$ -decision criterion for the BACs. If we choose a point  $(p, q) \in \mathcal{T}$  uniformly at random and consider the channel  $W = \text{BAC}(p, q)$ , the probability  $\Pr(W \sim_n \mathcal{A}) = 4 \cdot a(\mathcal{A})$  where  $a(\mathcal{A})$  is the area of the region corresponding to the criterion  $\mathcal{A}$ . In this sense, the area of a given  $n$ -decision criterion for the BACs is a measure of how probably this criterion is to be chosen (assuming uniform distribution on  $(p, q)$ ). By Theorem 2, the area of  $\mathcal{A} = \{(p, q) \in \mathcal{T} : r_0 < S(p, q) < r_1\}$  is  $a(r_1) - a(r_0)$  where  $r_1$  and  $r_0$  are consecutive rational numbers in  $\mathcal{D}_n$  and  $A(r) = a(R_r)$  is the area of the region  $R_r = \{(p, q) \in \mathcal{T} : 0 < S(p, q) < r\}$ . This area is equal to

$$A(r) = \iint_{R_r} 1 \, dp \, dq.$$

Let  $r = a/b$  where  $a$  and  $b$  are coprime positive integers with  $a < b$ . Applying the change of variable formula for double integrals with  $p = \frac{u-1}{uv-1}$ ,  $q = \frac{v-1}{uv-1}$  and after some calculations we obtain:

$$A(r) = \int_0^1 \frac{b(x^a - 1)^2 x^{b-1}}{2(x^{a+b} - 1)^2} dx. \quad (5)$$

Since  $x = 1$  is a zero of order 2 of the numerator, the integral is a proper integral. In some cases, a primitive for the integrand can be calculated explicitly, for example when  $r = 1/2$  and  $r = 1/3$  obtaining  $A(1/2) = \frac{1}{3} - \frac{\sqrt{3}\pi}{27}$  and  $A(1/3) = \frac{3}{8} - \frac{3\pi}{32}$ . In the other cases we have use the software Wolfram Mathematica [22] to calculate the integral (5) numerically after some reductions.

The  $n$ -stable criterion for the BACs nearest the BSC is denoted by  $\mathcal{A}_{QS}^n$ . It is given by the criterion corresponding to the channels  $BAC(p, q)$  satisfying  $r_n := \frac{2n-3-(-1)^n}{2n+1-(-1)^n} < S(p, q) < 1$ . We refer to these channels as  $n$ -quasi-symmetric channels. Clearly  $a(\mathcal{A}_{QS}^n) \rightarrow 0$  when  $n \rightarrow \infty$  (since  $r_n \rightarrow 1$ ). In the next table we show the percentages (rounded to the nearest integer) represented by the different  $n$ -stable criteria for the BACs, for  $3 \leq n \leq 7$ .

	0	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{2}{5}$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{3}{4}$	1
$n = 3$	53						47			
$n = 4$	32				21		47			
$n = 5$	22			11	21		18	29		
$n = 6$	16		6	11	21		18	29		
$n = 7$	12	4	6	11	8	12	18	8	21	

For example, for  $n = 3$  the region corresponding to the  $BAC(p, q)$  with  $\frac{1}{2} < S(p, q) < 1$  (the 3-quasi-symmetric channels) represents the 47% of the total area. The last percentage in each row of this table corresponds to the criterion  $\mathcal{A}_{QS}^n$ , associated with the  $n$ -quasi-symmetric channels. By (4), the number of stable regions for  $n = 8$  and  $n = 9$  is  $t_8 = 11$  and  $t_9 = 14$  respectively. The percentages represented by these regions (from left to right) are (9.09, 2.58, 3.85, 6.13, 10.54, 8.41, 12.12, 11.28, 7.00, 8.16, 20.84) for  $n = 8$  and (7.28, 1.81, 2.58, 3.85, 6.13, 4.49, 6.06, 8.41, 12.12, 11.28, 7.00, 8.16, 4.59, 16.24), for  $n = 9$ . As we can see, the criterion  $\mathcal{A}_{QS}^n$  is the most probable criterion, when  $(p, q)$  are chosen uniformly at random, among all the  $n$ -criteria for  $4 \leq n \leq 9$ . We have also checked this fact for  $n \leq 80$  and conjecture that this is true for any  $n \geq 4$ . Figure 5 displays graphically the percentages represented by all regions for  $n = 40$ . Let  $\mathcal{A}_Z^n$  be the  $n$ -stable decision criterion for the BACs nearest to the criterion corresponding to the  $Z$ -channel. We also point out an interesting comparison regarding the sizes of the areas corresponding to the criteria  $\mathcal{A}_{QS}^n$  and  $\mathcal{A}_Z^n$ . By considering for each  $n$  the ratios  $R(n)$  and  $r(n)$  between the areas corresponding to  $\mathcal{A}_{QS}^n$  and  $\mathcal{A}_Z^n$  and the average area for this  $n$ , it can be observed from our data that  $R(n)$  grows (with some very small oscillation) with  $n$  linearly (i.e.  $R(kn)/R(n)$  approaches  $k$ ) whereas  $r(n)$  gets near to one. As a sample, for  $n = 4, 8, 16, 18, 25, 36, 49, 40, 50, 100$  and  $200$ , the obtained sequences  $(n, R(n), r(n))$  are (4, 1.418, 0.966), (8, 2.292, 0.100), (16, 3.908, 0.966), (18, 4.398, 0.980), (25, 5.867, 1.012), (36, 8.299, 0.978), (40, 9.217, 0.983), (50, 11.588, 0.998), (100, 22.559, 0.991) and (200, 45.098, 1.001).

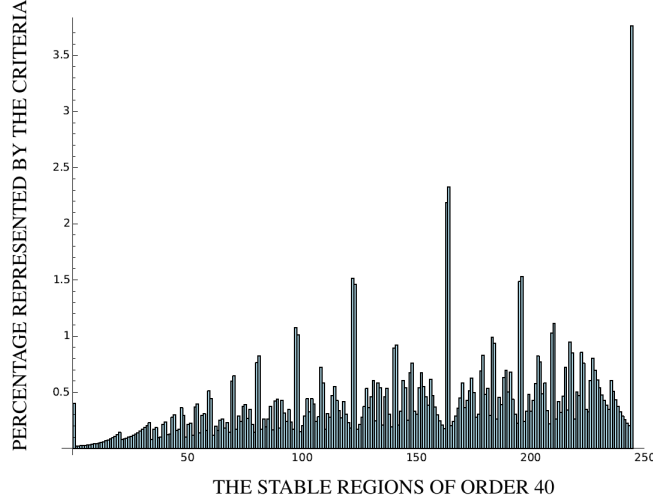


Fig. 5. The percentages represented by the 245 stable regions for  $n = 40$  ordered by its BAC-function values. The rightmost bar corresponds to  $\mathcal{A}_{QS}^{40}$ .

#### ACKNOWLEDGMENT

Work partially supported by CNPq grants 150270/2016-0, 140368/2015-9, 303985/2014-3 and 312926/2013-8 and by FAPESP grants 2015/26420-1 and 2013/25977-7.

#### REFERENCES

- [1] P. Cappelletti, C. Golla, P. Olivo, E. Zanoni, "Flash memories". Boston, MA, USA: Kluwer, 1999.
- [2] Y. Cassuto, M. Schwartz, V. Bohossian, J. Bruck, "Codes for asymmetric limited-magnitude errors with application to multi-level Flash memories". IEEE Transaction on Information Theory, vol. 56, no. 4, pp. 1582-1595, 2010.
- [3] T. Klove, B. Bose, N. Elarief, "Systematic single limited magnitude asymmetric error correcting codes". IEEE ITW, Cairo, Egypt, 2010.

- [4] E. Yaakobi, P. H. Siegel, A. Vardy, J. K. Wolf, "On codes that correct asymmetric errors with graded magnitude distribution". IEEE ISIT, St. Peterburgh, Russia, 2011.
- [5] C. Curto, V. Itskov, K. Morrison, Z. Roth, J. L. Walker, "Combinatorial Neural Codes from a Mathematical Coding Theory Perspective". Neural Computation, vol. 25, no. 7, pp. 1891-1925, 2013.
- [6] S. D. Constantin and T. R. N. Rao, "On the theory of binary asymmetric error-correcting codes". Information and Control, vol. 40, pp. 20-36, 1979.
- [7] R. J. McEliece, E. R. Rodemich, "The Constantin-Rao construction for binary asymmetric error-correcting-codes". Information and Control, vol. 44, no. 2, pp. 187-196, 1980.
- [8] J. M. Berger, "A note on error detecting codes for asymmetric channels". Information and Control, vol. 4, pp. 68-73, 1961.
- [9] R. Gabrys, L. Dolecek, "Coding for the Binary Asymmetric Channel". International Conference on Computing, Networking and Communications (ICNC), 2012.
- [10] T. Klove, "Error correcting codes for the asymmetric channel". Technical Report, Dept. of Informatics, University of Bergen, 1981. (Updated in 1995.).
- [11] R. R. Varshamov and G. M. Tenen Holtz, "A code for correcting a single asymmetric error". Automatica i Telemekhanika, vol. 26, no. 2, pp. 288-292, 1965.
- [12] J. H. Weber, "Bounds and constructions for binary block codes correcting asymmetric or unidirectional errors". Ph.D. Thesis, Delft University of Technology, The Netherlands, 1989.
- [13] N. Liu, W. Kang, "The capacity region of a class of Z Channels with degraded message sets". IEEE transaction on Information Theory, vol. 60, no. 10, pp. 6144 - 6153, 2014.
- [14] I. S. Moskowitz, S. J. Greenwald, M. H. Kang, "An analysis of the timed Z-channel". IEEE transaction on Information Theory, vol. 44, no. 7, pp. 3162-3168, 1998.
- [15] R. Prasad, S. Bhashyam, A. Chockalingam, "On the sum-rate of the Gaussian MIMO Z channel and the Gaussian MIMO X channel". IEEE Transaction on Communications, vol. 63, no. 2, pp. 487-497, 2015.
- [16] A. Ganti, A. Lapidoth, I. E. Telatar, "Mismatched decoding revisited: general alphabets, channels with memory, and the wide-band limit". IEEE Transaction on Information Theory, vol. 46, no. 7, pp. 2315-2328, 2000.
- [17] A. Lapidoth, P. Narayan, "Reliable communication under channel uncertainty". IEEE Transaction on Information Theory, vol. 44, no. 6, pp. 2148-2177, 1998.
- [18] M. Firer, J. L. Walker, "Matched metrics and channels". IEEE Transactions on Information Theory, vol. 62, no. 3, pp. 1150-1156, 2016.
- [19] R. G. L. D'Oliveira, M. Firer, "Channel Metrization". arXiv.1510.03104, 2016.
- [20] SageMath, the Sage Mathematics Software System (Version 7.4), The Sage Developers, 2016, <http://www.sagemath.org>.
- [21] T. M. Apostol, "Introduction to Analytic Number Theory". Springer-Verlag, 1976.
- [22] Wolfram Research, Inc., Mathematica, Version 10.4, Champaign, IL, 2016.

## APPENDIX

### Proof of the Proposition 1.

$i) \Rightarrow ii)$ . We consider  $x, y, z \in \mathcal{X}^n$  and the code  $C = \{y, z\}$ . Then,  $\Pr_{W_1}(x|y) \leq \Pr_{W_1}(x|z) \Leftrightarrow \arg \max_{c \in \{y, z\}} \Pr_{W_1}(x|c) \neq \{y\} \Leftrightarrow \arg \max_{c \in \{y, z\}} \Pr_{W_2}(x|c) \neq \{y\} \Leftrightarrow \Pr_{W_2}(x|y) \leq \Pr_{W_2}(x|z)$ , where the second implication follows from assuming  $i)$ .

$ii) \Rightarrow i)$ . We consider a code  $C \subseteq \mathcal{X}^n$  and  $x \in \mathcal{X}^n$ , we have  $c_0 \in \arg \max_{c \in C} \Pr_{W_1}(x|c) \Leftrightarrow \Pr_{W_1}(x|c_0) \geq \Pr_{W_1}(x|c), \forall c \in C \Leftrightarrow \Pr_{W_2}(x|c_0) \geq \Pr_{W_2}(x|c), \forall c \in C \Leftrightarrow c_0 \in \arg \max_{c \in C} \Pr_{W_2}(x|c)$ .

$ii) \Leftrightarrow iii)$ . We consider the following equivalence relation in  $\mathbb{R}^n$ . Two vectors  $a = (a_1, \dots, a_n)$  and  $a' = (a'_1, \dots, a'_n)$  are equivalent (which is denoted by  $a \sim a'$ ) if for all  $i, j$ :  $a_i \leq a_j \Leftrightarrow a'_i \leq a'_j$ . Let  $\tau : \mathbb{R}^n \rightarrow \mathbb{R}^n$  be the function given by  $\tau(a_1, \dots, a_n) = (b_1, \dots, b_n)$  with  $b_i = \#\{j : a_j \leq a_i\}$ , and  $f_i$  and  $f'_i$  be the  $i$ -th lines of  $M_1$  and  $M_2$ , respectively. We note that  $ii)$  is equivalent to  $f_i \sim f'_i, \forall i : 1 \leq i \leq 2^n$ ; and  $iii)$  is equivalent to  $\tau(f_i) = \tau(f'_i), \forall i : 1 \leq i \leq 2^n$ . Thus, it suffices to prove that for every  $a = (a_1, \dots, a_n)$  and  $a' = (a'_1, \dots, a'_n) \in \mathbb{R}^n$ :  $a \sim a' \Leftrightarrow \tau(a) = \tau(a')$ .

$(\Rightarrow)$  We consider  $a, a' \in \mathbb{R}^n$  such that  $a \sim a'$ . For each  $j : 1 \leq j \leq n$  we have:  $k \in \{i : a_i \leq a_j\} \Leftrightarrow a_k \leq a_j \Leftrightarrow a'_k \leq a'_j \Leftrightarrow k \in \{i : a'_i \leq a'_j\}$ . Thus  $\{i : a_i \leq a_j\} = \{i : a'_i \leq a'_j\}$ . Taking cardinalities we obtain  $\tau(a)_j = \tau(a')_j, \forall j : 1 \leq j \leq n$  and then  $\tau(a) = \tau(a')$ .

$(\Leftarrow)$  We consider  $a, a' \in \mathbb{R}^n$  such that  $\tau(a) = \tau(a')$  and note that if  $\tau(a)_i \leq \tau(a)_j$  then  $\{k : a_k \leq a_i\} \subseteq \{k : a_k \leq a_j\}$  (because in this case  $\{k : a_k \leq a_j\} \not\subseteq \{k : a_k \leq a_i\}$ ). Thus, for all  $i, j$  we have  $a_i \leq a_j \Leftrightarrow \{k : a_k \leq a_i\} \subseteq \{k : a_k \leq a_j\} \Leftrightarrow \tau(a)_i \leq \tau(a)_j \Leftrightarrow \tau(a')_i \leq \tau(a')_j$  (because  $\tau(a) = \tau(a')$ )  $\Leftrightarrow \{k : a'_k \leq a'_i\} \subseteq \{k : a'_k \leq a'_j\} \Leftrightarrow a'_i \leq a'_j$ ; and we conclude that  $a \sim a'$ .

### Proof of the Theorem 1.

$i) \Rightarrow ii)$  Follows from the definition of  $n$ -equivalence and the fact that  $\text{sdec}_{W_1}(C, x) = \text{sdec}_{W_2}(C, x)$  if and only if one of the following possibilities occurs:

- $\arg \max_{c \in C} \Pr_{W_1}(x|c) = \arg \max_{c \in C} \Pr_{W_2}(x|c) = \{c_0\}$  for some  $c_0 \in C$  or
- $\#\arg \max_{c \in C} \Pr_{W_1}(x|c) > 1$  and  $\#\arg \max_{c \in C} \Pr_{W_2}(x|c) > 1$ .

$ii) \Rightarrow i)$  We assume that  $\text{sdec}_{W_1} = \text{sdec}_{W_2}$  for  $n$ -block codes,  $n \geq 2$ . This also implies that  $\text{sdec}_{W_1} = \text{sdec}_{W_2}$  for 2-block codes (see Remark 1). Let's suppose by contradiction that the channels  $W_1$  and  $W_2$  are not  $n$ -equivalent. In this case, there exists a code  $C \subseteq \mathcal{X}^n$ ,  $x \in \mathcal{X}^n$  and  $c_0 \in C$  such that  $c_0 \in \arg \max_{c \in C} \Pr_{W_1}(x|c)$  and  $c_0 \notin \arg \max_{c \in C} \Pr_{W_2}(x|c)$ . Thus, there exists  $c_1 \in C$  such that  $\Pr_{W_2}(x|c_1) > \Pr_{W_2}(x|c_0)$ . Consider the code  $C' = \{c_0, c_1\} \subseteq C$ . We have that  $\arg \max_{c \in C'} \Pr_{W_2}(x|c) = \{c_1\}$ . Since  $\text{sdec}_{W_1} = \text{sdec}_{W_2}$  for 2-block codes, we have that  $\arg \max_{c \in C'} \Pr_{W_1}(x|c) = \{c_1\}$ . Thus  $\Pr_{W_1}(x|c_1) > \Pr_{W_1}(x|c_0)$  which contradicts that  $c_0 \in \arg \max_{c \in C} \Pr_{W_1}(x|c)$ . Thus, the channels  $W_1$  and  $W_2$  are  $n$ -equivalent.

$i) \Leftrightarrow iii)$  We define the extended probabilistic Voronoi regions  $V_{(C,W)}^{\text{ext}}(c) = \{x \in \mathcal{X}^n : \Pr_W(x|c) \geq \Pr_W(x|c'), \forall c' \in C\}$ . Using Proposition 1, it is not difficult to prove that  $W_1 \underset{n}{\sim} W_2$  if and only if  $V_{(C,W_1)}^{\text{ext}}(c) = V_{(C,W_2)}^{\text{ext}}(c)$  for all  $C \subseteq \mathcal{X}^n$  and  $c \in C$ . By direct calculation we have

$$\Pr(\text{udec}_W(C, x) = c) = \begin{cases} 0 & \text{if } x \notin V_{(C,W)}^{\text{ext}}(c); \\ \frac{1}{M} & \text{if } x \in V_{(C,W)}^{\text{ext}}(c); \end{cases} \quad (6)$$

where  $M = |V_{(C,W)}^{\text{ext}}(c)|$ . If  $W_1 \underset{n}{\sim} W_2$ , then  $V_{(C,W_1)}^{\text{ext}}(c) = V_{(C,W_2)}^{\text{ext}}(c)$  for all  $C \subseteq \mathcal{X}^n, x \in \mathcal{X}^n$  and by Equation (6) we have  $\Pr(\text{udec}_{W_1}(C, x) = c) = \Pr(\text{udec}_{W_2}(C, x) = c)$  for all  $C \subseteq \mathcal{X}^n, x \in \mathcal{X}^n$ . This proves  $i) \Rightarrow iii)$ . The converse follows from the fact that if  $\text{udec}_{W_1} = \text{udec}_{W_2}$  then

$$\begin{aligned} V_{(C,W_1)}^{\text{ext}}(c) &= \{x \in \mathcal{X}^n : \Pr(\text{udec}_{W_1}(C, x) = c) > 0\} \\ &= \{x \in \mathcal{X}^n : \Pr(\text{udec}_{W_2}(C, x) = c) > 0\} \\ &= V_{(C,W_2)}^{\text{ext}}(c) \end{aligned}$$

(where in the first and last equality we use Equation (6) and  $\text{udec}_{W_1} = \text{udec}_{W_2}$  is used in the second equality). From which we conclude  $W_1 \underset{n}{\sim} W_2$ .