Matched Metrics and Channels

Marcelo Firer and Judy L. Walker

September 22, 2014

Abstract

The most common decision criteria for decoding are maximum likelihood decoding and nearest neighbor decoding. It is well-known that maximum likelihood decoding coincides with nearest neighbor decoding with respect to the Hamming metric on the binary symmetric channel. In this work we study channels and metrics for which those two criteria do and do not coincide for general codes.

1 Introduction

In coding theory, the Hamming metric has a prominent status, since it can be used to perform maximum likelihood (ML) decoding over a memoryless binary symmetric channel (BSC), in the sense that decoding by choosing a most probable codeword (ML decoding) or a closest codeword (nearest neighbor (NN) decoding) is actually the same decision. In this decision criteria sense, we have also that Euclidean distance is the proper distance for modulation-decoding when considering a continuous channel with white Gaussian noise (see, e.g., [3]) and the Lee metric has the same distinguished role when considering some kinds of modulation and transmission over certain discrete memoryless q-ary channels [1].

The use of geometric properties of channels in coding theory is explored in many generic situations, such as the one proposed by Forney [6] for geometric uniformity of codes on continuous channels and the study of geometrically inspired properties of codes over discrete channels, as in [5] and [7], where bounds for the packing radius are derived from a distance-like structure defined on an hypergraph determined by the channel model.

Many different distances are considered in the context of coding theory (a comprehensive account may be found in [4, Chapter 16]), but not much is known about general relation between channel models and metrics and not much is known about the geometry of many important channels.

In this work we are concerned with the most basic of those questions: is any ML decoder also an NN decoder, and conversely, is any NN decoder also an ML decoder? More precisely, if, for every code, ML decoding on a given channel coincides with NN decoding with respect to a given metric, we say the channel and the metric are *matched* to one another.

This terminology goes back over 40 years, as a 1971 paper [1] attributes it to notes from a 1967 course given by Massey [8]. In [1], Chiang and Wolf classify the channels matched to the Lee Metric. Their results are generalized in a 1980 paper by Séguin [9], which studies necessary and sufficient conditions for a discrete memoryless channel to be matched to an *additive* metric, i.e., a metric that is defined on an alphabet A and then extended to a metric A^n by applying the metric coordinate-wise and taking the sum. To avoid potential confusion by the reader, we also mention recent work on *mismatched decoders* (see, e.g., [10]) but note that that work considers different questions than those that are considered here. In particular, given a channel with transition probabilities $\Pr(x|y)$, taking $-\frac{1}{n}\log\Pr(x|y)$ does not, in general, give a metric in any sense.

This work is organized as follows: In Section 2 we introduce the rigorous definition of the matching problem and we show that any metric admits a matched channel but the converse does not always hold. We therefore also give some conditions for a channel that obstruct the existence of a matched metric. In Section 3 we construct a matched metric for the Z-channel. In Section 4 we conjecture that any binary asymmetric channel (BAC) admits a matched metric and present some evidence for this conjecture.

2 Matched metrics and channels

{mas}

It is well-known that, under the assumption of equally likely codewords, maximum likelihood decoding coincides with nearest neighbor decoding with respect to the Hamming metric on the binary symmetric channel. This is a general fact that does not depend on the code and so we ask: for what other channels is there such a metric? Following Massey [8], we call such a channel-metric pairs *matched*, a term we define rigorously as follows:

Definition 2.1. *Let $W : \mathcal{X} \to \mathcal{X}$ be a channel with input and output alphabets $\mathcal{X}$ and let d be a metric on $\mathcal{X}$, i.e., $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is a function such that:*

1. *d is* symmetric*: $d(x, y) = d(y, x)$ for all $x, y \in \mathcal{X}$;*
2. *d is* nonnegative*: $d(x, y) \geq 0$ for all $x, y \in \mathcal{X}$, with equality if and only if $x = y$; and*
3. *d satisfies the* triangle inequality*: $d(x, y) + d(y, z) \geq d(x, z)$ for all $x, y, z \in \mathcal{X}$.*

We say that W and d are matched *if maximum likelihood decoding on W coincides with nearest neighbor decoding with respect to d for every code $C \subseteq \mathcal{X}$, i.e., if for every code $C \subseteq \mathcal{X}$ and every $x \in \mathcal{X}$, we have*

$$\arg\max_{y \in C} \Pr(x \textit{ received} \mid y \textit{ sent}) = \arg\min_{y \in C} d(x, y). \tag{1}$$ {matched}

Several comments are in order. First, note that we are considering codes to be subsets of the alphabet, so that the binary symmetric channel, for example, should be considered as a channel on $\mathbb{F}_2^n$ rather than as n uses of a channel defined on $\mathbb{F}_2$, and the Hamming metric should be considered as a metric on $\mathbb{F}_2^n$ rather than as a sum of n copies of the Hamming metric on $\mathbb{F}_2$. Next, by considering codes with just two codewords, it is straightforward that condition (1) is equivalent to the condition that, for every x, y, $z \in \mathcal{X}$ with either $\Pr(x \text{ received} \mid y \text{ sent}) > 0$ or $\Pr(x \text{ received} \mid z \text{ sent}) > 0$ (or both),

$$\Pr(x \text{ received} \mid y \text{ sent}) > \Pr(x \text{ received} \mid z \text{ sent}) \quad \text{if and only if} \quad d(x, y) < d(x, z). \tag{2}$$ {newmatche

Finally, we make the following two assumptions throughout the paper:

- Every channel is *reasonable* in the sense that $\Pr(x \text{ sent} \,|\, x \text{ received}) > \Pr(x \text{ sent} \,|\, y \text{ received})$, for all $x \neq y \in \mathcal{X}$.
- All codewords are equally likely.

The first assumption is a necessary condition for a channel to admit a matched metric since $0 = d(x,x) < d(x,y)$ for all $x \neq y \in \mathcal{X}$; the second is needed in order for maximum likelihood decoding to be relevant. (Alternatively, one could drop this assumption and replace "maximum likelihood decoding" with "maximum a posteriori probability decoding" throughout the paper.) In light of this second assumption, we note that conditions (1) and (2) are also equivalent to

$$\Pr(y \text{ sent} \,|\, x \text{ received}) > \Pr(z \text{ sent} \,|\, x \text{ received}) \quad \text{if and only if} \quad d(x,y) < d(x,z) \tag{3}$$ {newnewmatc

for every $x, y, z \in \mathcal{X}$ with either $\Pr(x \text{ received} \,|\, y \text{ sent}) > 0$ or $\Pr(x \text{ received} \,|\, z \text{ sent}) > 0$ (or both).

We are interested in determining which channels admit matched metrics, and which metrics admit matched channels. For example, as described in the introduction, it is well-known that the Hamming metric and the n-fold binary symmetric channel $\mathrm{BSC}(n)$ are matched; the Euclidean metric and the n-fold additive white Gaussian noise channel $\mathrm{AWGN}(n)$ are matched; and the Lee metric and certain n-fold q-ary channels are matched (see [1], Theorem 1).

The general question of which metrics admit matched channels is much simpler than the question of which channels admit matched metrics. Indeed, *every* metric is matched to some channel:

Proposition 2.2. *For any finite metric space $(\mathcal{X}, d)$ there is a channel $W : \mathcal{X} \to \mathcal{X}$ matched to d.*

Proof. Given a finite metric space $(\mathcal{X}, d)$, we construct a channel $W : \mathcal{X} \to \mathcal{X}$ by constructing the conditional probabilities $\Pr(y \,|\, x) = \Pr(y \text{ sent} \,|\, x \text{ received})$ for $x, y \in \mathcal{X}$.

Fix $0 < \epsilon < 1$. For $x, y \in \mathcal{X}$, set $\beta_{xy} = \epsilon^{d(x,y)}$, set $\gamma_x = \sum_{y \in \mathcal{X}} \beta_{xy}$, and set $\Pr(y \,|\, x) = \frac{\beta_{xy}}{\gamma_x}$. Then $0 < \Pr(x \,|\, y) \leq 1$ and, for a fixed $x \in \mathcal{X}$,

$$\begin{aligned}
\sum_{y \in \mathcal{X}} \Pr(y \,|\, x) &= \sum_{y \in \mathcal{X}} \frac{\beta_{xy}}{\gamma_x} \\
&= \frac{1}{\gamma_x} \sum_{y \in \mathcal{X}} \beta_{xy} \\
&= \frac{1}{\gamma_x} \cdot \gamma_x \\
&= 1
\end{aligned}$$

and so this definition yields a valid channel. To see that this channel is matched to our metric, let $x, y, z \in \mathcal{X}$. Then

$$P(y \,|\, x) > P(z \,|\, x) \iff \frac{\beta_{xy}}{\gamma_x} > \frac{\beta_{xz}}{\gamma_x} \iff \epsilon^{d(x,y)} > \epsilon^{d(x,z)} \iff d(x,y) < d(x,z),$$

and so condition (3) is satisfied. $\square$

On the other hand, not every channel has a matched metric, as the following simple example demonstrates:

{NoMatchExa

Example 2.3 (Inexistence of a matched metric). *Let* $\mathcal{X} = \{x, y, z\}$ *and* $W : \mathcal{X} \to \mathcal{X}$ *be defined by the probabilities*

$$\begin{array}{lll} \Pr(x|x) = a & \Pr(x|y) = b & \Pr(x|z) = c \\ \Pr(y|x) = c & \Pr(y|y) = a & \Pr(y|z) = b \\ \Pr(z|x) = b & \Pr(z|y) = c & \Pr(z|z) = a \end{array}$$

with $a > b > c > 0$ *and* $a + b + c = 1$*; for example we could have* $a = \frac{1}{2}$, $b = \frac{1}{3}$ *and* $c = \frac{1}{6}$. *Suppose* $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ *is a matched metric for* W*. Then*

$$\begin{aligned} d(x,x) &< d(x,y) < d(x,z) \\ d(y,y) &< d(y,z) < d(y,x) \\ d(z,z) &< d(z,x) < d(z,y). \end{aligned} \tag{4}$$ {contra}

However, this leads to a contradiction:

$$\begin{aligned} d(z,x) &< d(z,y) \\ &= d(y,z) < d(y,x) \\ &= d(x,y) < d(x,z) \\ &= d(z,x), \end{aligned}$$

where the inequalities follow from the inequalities in (4) and the equalities are just the symmetry of the metric d*. Therefore there can be no matched metric for the channel* W*.*

The preceding tiny example can be generalized to any larger alphabet size and we have the following:

Proposition 2.4. *For any alphabet* $\mathcal{X}$ *with* $|\mathcal{X}| \geq 3$*, there is a channel* $W : \mathcal{X} \to \mathcal{X}$ *that does not admit a matched metric.*

Proof. If $|\mathcal{X}| = 3$, proceed as in Example 2.3. Otherwise, write $\mathcal{X} = \mathcal{X}_0 \cup \mathcal{Y}$ where $|\mathcal{X}_0| = 3$ and $|\mathcal{Y}| = M \geq 1$. Label the elements of $\mathcal{X}_0$ so that $\mathcal{X}_0 = \{x, y, z\}$. Fix positive real numbers $a > b > c > d$ with $a + b + c + d = 1$. Define conditional probabilities for $u, v \in \mathcal{X}$ by

$$\Pr(u \,|\, v) = \begin{cases} \Pr_0(u \,|\, v) & \text{if } u, v \in \mathcal{X}_0 \\ 0 & \text{if } u \in \mathcal{X}_0, v \in \mathcal{Y} \\ \frac{d}{M} & \text{if } u \in \mathcal{Y}, v \in \mathcal{X}_0 \\ a & \text{if } u = v \in \mathcal{Y} \\ \frac{1-a}{M-1} & \text{if } u, v \in \mathcal{Y} \text{ and } u \neq v, \end{cases}$$

where $\Pr_0(u \,|\, v)$ is as described in Example 2.3 for $u, v \in \mathcal{X}_0$. Then it is straightforward to check that these conditional probabilities define a channel $W : \mathcal{X} \to \mathcal{X}$ and that this channel has no metric for the same reason as in Example 2.3. □

The 3-step cycle of Example 2.3 gives rise to a more general obstruction criterion for the existence of a metric matching a given channel. Given a channel $W : \mathcal{X} \to \mathcal{X}$, $x \in \mathcal{X}$, and $0 \leq t \leq 1$, we define the *t-decision region centered at x* to be $B^t(x) := \{y \in \mathcal{X} \mid \Pr(x|y) \geq t\}$. We say that $x_0, x_1, ..., x_{r-1} \in \mathcal{X}$ is a *decision chain* of length r on W if there are values $t_0, t_1, ..., t_{r-1} > 0$ satisfying the following conditions:

(FIP) Forward inclusion property: $x_{i+1} \in B^{t_i}(x_i)$

(MEP) Backward exclusion property: $x_i \notin B^{t_{i+1}}(x_{i+1})$

where we consider the indices modulo r. For example, taking $x_0 = x$, $x_1 = y$, $x_2 = z$ and $t_0 = t_1 = t_2 = \frac{b+c}{2}$ in Example 2.3 gives a decision chain of length 3.

With this definition we can state the following:

Proposition 2.5. *Let W be a channel over the alphabet $\mathcal{X}$. If W admits a decision chain of length $r \geq 3$, then there is no metric matched to W.*

Proof. Let $x_0, x_1, ..., x_{r-1} \in \mathcal{X}$ be a decision chain on W with parameters $t_0, t_1, ..., t_{r-1}$ and suppose $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is a metric matched to W. Using *FIP* for $i = 0$ and *BEP* for $i = r-1$ we get that

$$d(x_0, x_1) < d(x_0, x_{r-1}) = d(x_{r-1}, x_0)$$

where the equality follows from the symmetry property of d. Using *FIP* for $i = r-1$ and *BEP* for $i = r-2$ we get that

$$d(x_{r-1}, x_0) < d(x_{r-1}, x_{r-2}) = d(x_{r-2}, x_{r-1})$$

and, proceeding in this manner, we get

$$d(x_0, x_1) < d(x_0, x_{r-1}) < d(x_{r-1}, x_{r-2}) < \cdots < d(x_0, x_1)$$

a contradiction. Thus d cannot exist. □

The preceding proposition gives an obstruction to the existence of a metric matched to a channel, but many channels do not fit into this picture. If we consider W to be a *reasonable* (in the sense described above) symmetric channel ($\Pr(x|y) = \Pr(y|x)$, for every $x, y \in \mathcal{X}$), then defining

$$D(x, y) = \begin{cases} 0 & \text{if } x = y \\ 1 - \Pr(x|y) & \text{if } x \neq y \end{cases}$$

we get that D satisfies all the properties of a metric except for the triangle inequality; as we shall see in Lemma 3.2 below, it is not difficult to obtain a metric $d(x, y)$ from $D(x, y)$.

It follows that the main difficulty in finding a metric matched to a given channel is to find a *symmetric* function satisfying condition (1) (or, equivalently, (2)).

In the next section, we construct a matched metric for the n-fold Z-channel, for any n. We consider this result to be a bit surprising, since, as desribed above, it is the symmetry property that poses the most difficulty in constructing a metric matched to a given channel, and the Z-channel is as asymmetrical as possible, in the sense that for $x \neq y$ we have that $\Pr(x|y) > 0$ implies $\Pr(y|x) = 0$.

3 A matched metric for the Z-channel {Z}

The Z-channel is the memoryless binary input and output channel with transition probabilities given by $\Pr(0|0) = 1$, $\Pr(1|0) = 0$, $\Pr(0|1) = q$, $\Pr(1|1) = 1 - q$, where $0 < q < \frac{1}{2}$ and, as usual, we write $\Pr(x|y)$ to mean $\Pr(x \text{ received} \,|\, y \text{ sent})$. The n-fold Z-channel is the memoryless channel with input and output $\mathbb{F}_2^n$ with

$$\Pr(x|y) = \prod_{i=1}^{n} \Pr(x_i|y_i)$$

for $x = (x_1, \cdots, x_n), y = (y_1, \cdots, y_n) \in \mathbb{F}_2^n$.

The main result of this section is as follows:

{z-theorem}

Theorem 3.1. *For any $n \geq 1$, there is a metric matched to the n-fold Z-channel.*

Before proving Theorem 3.1, we need a lemma.

{squeeze-le

Lemma 3.2. *Let $\mathcal{X}$ be a finite set and suppose $e : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is a symmetric, nonnegative function, i.e.,*

1. *$e(x, y) = e(y, x)$ for every $x, y \in \mathcal{X}$; and*
2. *$e(x, y) \geq 0$ for every $x, y \in \mathcal{X}$, with equality if and only if $x = y$.*

Then there is a metric d on $\mathcal{X}$ such that $d(x, y) < d(x, z)$ if and only if $e(x, y) < e(x, z)$ for every $x, y, z \in \mathcal{X}$.

Proof. Since $\mathcal{X}$ is finite, the set $\{e(x, y) \,|\, x, y \in \mathcal{X}\}$ has both a maximal and a minimal element; set $m = \min\{e(x, y) \,|\, x, y \in \mathcal{X}\}$ and $M = \max\{e(x, y) \,|\, x, y \in \mathcal{X}\}$. Fix δ with $0 < \delta < \frac{1}{3}$ and let $f : [m, M] \to [1 - \delta, 1 + \delta]$ be a strictly increasing bijective function. (For example, take f to be the linear function which maps m to $1 - \delta$ and M to $1 + \delta$.) Define $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ by

$$d(x, y) = \begin{cases} e(x, y) = 0 & \text{if } x = y, \\ f(e(x, y)) & \text{otherwise.} \end{cases}$$

The symmetry and nonnegativity of $d(\cdot, \cdot)$ follow immediately from these properties of $e(\cdot, \cdot)$ and the fact that f is strictly increasing. To check that d satisfies the triangle inequality, let $x, y, z \in \mathcal{X}$. Then

$$d(x, y) + d(y, z) \geq 2(1 - \delta) > 2(1 - \frac{1}{3}) = \frac{4}{3} > 1 + \delta \geq d(x, z).$$

Hence d is a metric. □

Proof of Theorem 3.1. We proceed by induction on $n \geq 1$. For the base case of $n = 1$, we note that the Hamming metric is matched to the Z-channel on $\mathbb{F}_2$.

Suppose there is a matched metric for the n-fold Z-channel, determined by the $2^n \times 2^n$ matrix D_n. Our goal is to construct a $2^{n+1} \times 2^{n+1}$ matrix D_{n+1} that represents a metric that is matched to the $(n+1)$-fold Z-channel. For any $u \in \mathbb{F}_2^{n+1}$, we write $u = (x_1, \ldots, x_n, \theta) =: x\theta$, where $x \in \mathbb{F}_2^n$ and $\theta \in \mathbb{F}_2$; given an ordering $v_1, \ldots, v_N$ of the elements of $\mathbb{F}_2^n$ ($N = 2^n$), this yields an ordering $v_1 0, \ldots, v_N 0, v_1 1, \ldots, v_N 1$ of $\mathbb{F}_2^{N+1}$.

Let $P_n = (P_{x,y})_{x,y \in \mathbb{F}_2^n}$ be the probability matrix for the n-fold Z-channel, so that $P_{x,y} = \Pr(x \text{ received} \,|\, y \text{ sent})$. Then the probability matrix P_{n+1} for the $(n+1)$-fold Z-channel is given by

$$P_{n+1} = \left(\begin{array}{c|c|c} & v_1 0 \quad \cdots \quad v_N 0 & v_1 1 \quad \cdots \quad v_N 1 \\ \hline \begin{matrix} v_1 0 \\ \vdots \\ v_N 0 \end{matrix} & P_n \cdot \Pr(0 \text{ received} \,|\, 0 \text{ sent}) & P_n \cdot \Pr(0 \text{ received} \,|\, 1 \text{ sent}) \\ \hline \begin{matrix} v_1 1 \\ \vdots \\ v_N 1 \end{matrix} & P_n \cdot \Pr(1 \text{ received} \,|\, 0 \text{ sent}) & P_n \cdot \Pr(1 \text{ received} \,|\, 1 \text{ sent}) \end{array} \right)$$

$$= \left(\begin{array}{c|c|c} & v_1 0 \quad \cdots \quad v_N 0 & v_1 1 \quad \cdots \quad v_N 1 \\ \hline \begin{matrix} v_1 0 \\ \vdots \\ v_N 0 \end{matrix} & P_n & P_n \cdot q \\ \hline \begin{matrix} v_1 1 \\ \vdots \\ v_N 1 \end{matrix} & 0 & P_n \cdot (1-q) \end{array} \right).$$

We will use this information to construct a matrix D_{n+1} that determines a metric matched to the $(n+1)$-fold Z-channel. The entries of the matrix $D_{n+1} = (d_{uv})_{u,v \in \mathbb{F}_2^{n+1}}$ must satisfy the following properties:

(M) d must be *matched*: $d_{uv} < d_{uw}$ if and only if $\Pr(u|v) > \Pr(u|w)$.

(S) d must be *symmetric*: $d_{uv} = d_{vu}$ for every $u, v \in \mathbb{F}_2^{n+1}$.

(N) d must be *nonnegative*: $d_{uv} \geq 0$ for every $u, v \in \mathbb{F}_2^{n+1}$, with equality if and only if $u = v$.

(T) d must satisfy the *triangle inequality*: $d_{uv} + d_{vw} \geq d_{uw}$ for every $u, v, w \in F_2^{n+1}$.

Note that the last three of these properties are required for D_{n+1} to represent a metric, while the first is what makes the metric matched to the channel. We begin by constructing a matrix E that satisfies properties (M), (S) and (N). We then apply Lemma 3.2 to E to transform E into a matrix D that satisfies property (T) while maintaining the other three properties. This modified matrix will be our desired D_{n+1}.

Because E must be a symmetric matrix by (S), we can write

$$E = \left(\begin{array}{c|ccc|ccc} & v_1 0 & \cdots & v_N 0 & v_1 1 & \cdots & v_N 1 \\ \hline v_1 0 & & & & & & \\ \vdots & & A & & & B & \\ v_N 0 & & & & & & \\ \hline v_1 1 & & & & & & \\ \vdots & & B^T & & & C & \\ v_N 1 & & & & & & \end{array}\right),$$

where we require $A = (a_{xy})_{x,y\in\mathbb{F}_2^n}$, $B = (b_{xy})_{x,y\in\mathbb{F}_2^n}$ and $C = (c_{xy})_{x,y\in\mathbb{F}_2^n}$ to be $2^n \times 2^n$ matrices, with A and C symmetric.

Determining matrix A: To satisfy (M), we must have $a_{xy} < a_{xz}$ if and only if $\Pr(x0|y0) > \Pr(x0|z0)$ if and only if $\Pr(x|y) > \Pr(x|z)$. Thus A must represent a matched metric for the n-fold Z-channel, and we may set $A = D_n$.

Determining matrix B: Let us consider the entry $b_{x,y}$ of the matrix B. We break this into three cases.

Case 1. First, suppose $\Pr(x|y) \neq 0$ and y is not the all-ones vector in $\mathbb{F}_2^n$. Without loss of generality, we may assume that all of the 0's in y are at the beginning, so that $y = 0^j 1^k$ with $j \geq 1$ and $j + k = n$, where, for example, we mean by $0^2 1^3$ the vector $(0, 0, 1, 1, 1)$. Since $\Pr(1|0) = 0$, the first j coordinates of x are 0 as well. Hence, without loss of generality, we may assume that $x = 0^j 0^s 1^t$ with $s + t = k$. Now set $z = (1, y_2, \ldots, y_n)$. Then

$$\Pr(x|z) = \Pr(0|1) \prod_{i=2}^{n} \Pr(x_i|y_i) = q \Pr(x|y)$$

since $\Pr(x_1|y_1) = \Pr(0|0) = 1$. Thus we have

$$\Pr(x0|y1) = q \Pr(x|y) = \Pr(x|z) = \Pr(x0|z0)$$

and so we set $b_{xy} = a_{xz}$. We remark that, since $x \neq z$, the induction hypothesis ensures that $a_{xz} \neq 0$, and hence also $b_{xy} \neq 0$.

Case 2. We now consider the case where $y = 1^n$ is the all-ones vector in $\mathbb{F}_2^n$, and find the value of b_{x1^n}. Note first that $\Pr(1^n 0|1^n 1) = q(1-q)^n$ is the second-largest entry in the row of P_{n+1} indexed by $1^n 0$, second only to $\Pr(1^n 1|1^n 1)$. Since $a_{1^n 1^n} = 0$, we therefore require $b_{1^n 1^n}$ to be smaller than every nonzero $a_{1^n z}$; for concreteness, we set $b_{1^n 1^n} = \frac{1}{2} \min\{a_{1^n z} | z \neq 1^n\}$.

For $x \neq 1^n$, without loss of generality, we may assume $x = 0^s 1^t$, where $s \geq 1$ and $s+t = n$. Then

$$\Pr(x|1^n) = q^s (1-q)^t$$

and

$$\Pr(x0|1^n 1) = q^{s+1} (1-q)^t.$$

Suppose $z \neq 1^n$ satisfies $\Pr(x|z) \neq 0$; since $x \neq 1^n$, we know such a z exists. Since $\Pr(1|0) = 0$, without loss of generality we can write $z = 0^j 1^k 1^t$, where $j + k = s$ and $j \geq 1$. This means

$$\Pr(x|z) = \Pr(0|0)^j \Pr(0|1)^k \Pr(1|1)^t = q^k (1-q)^t,$$

where $k < s$. Putting this together, we have

$$\Pr(x0|1^n 1) = q^{s+1}(1-q)^t < q^{k+1}(1-q)^t = \Pr(x0|z1) < q^k(1-q)^t = \Pr(x0|z0)$$

and so we require $b_{x1^n} > b_{xz} > a_{xz}$ for every $z \neq 1^n$ with $\Pr(x|z) \neq 0$. For concreteness, we set

$$b_{x1^n} = 2\max\{b_{xz} \mid z \neq 1^n \text{ and } \Pr(x|z) \neq 0\}.$$

Case 3. Finally, if $\Pr(x|y) = 0$, then $\Pr(x0|y1) = 0$ and so

$$\Pr(x0|y1) < \Pr(x0|z1) = q\Pr(x0|z0) < \Pr(x0|z0) = \Pr(x|z)$$

for every z with $\Pr(x|z) \neq 0$. This means we require

$$b_{xy} > b_{xz} > a_{xz}$$

for every z with $\Pr(x|z) \neq 0$, and we set

$$b_{xy} = 2\max\{b_{xz} \mid \Pr(x|z) \neq 0\}.$$

By the remark made in Case 1 and the constructions in the two remaining cases, the matrix $B = (b_{xy})$ has strictly positive entries.

Determining matrix C: Because $\Pr(x1|y0) = 0$ for all x and y, we must have $c_{xz} < b_{yx}$ for all z with $\Pr(x|z) \neq 0$. Because $\Pr(x1|y1) < \Pr(x1|z1)$ if and only if $\Pr(x|y) < \Pr(x|z)$, the matrix C must represent a matched metric for the n-fold Z-channel. By choosing δ sufficiently small and setting $C = \delta D_n$, we can satisfy these conditions.

We now have a matrix

$$E = (e_{uv})_{u,v\in\mathbb{F}_2^{n+1}} = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}$$

that satisfies properties (M), (S) and (N) above. Using Lemma 3.2, we can transform E in such a way that we force the triangle inequality (T) to hold without affecting the other three properties. Thus the resulting matrix $D_{n+1} = (d_{uv})_{u,v\in\mathbb{F}_2^{n+1}}$ represents a master metric for the $(n+1)$-fold Z-channel, as desired. □

4 Asymmetric channels

{bac}

A binary asymmetric channel (BAC) with parameters (p,q) is a memoryless channel with binary input and output alphabet with transition probabilities given by $\Pr(0|0) = 1-p$, $\Pr(1|0) = p$, $\Pr(0|1) = q$, $\Pr(1|1) = 1-q$, where $0 \leq p \leq q < \frac{1}{2}$. The extreme cases of an asymmetric channel are the symmetric channel (for $p = q$) and the Z-channel (for $p = 0$). The squeezing function of Lemma 3.2 ensures the triangle inequality can always be attained. Hence the unique difficulty to construct a metric matched to a given channel lies on the symmetry of a distance matrix. From this point of view, we could expect that finding a matched metric for an asymmetric channel should become harder as the asymmetry of the channel grows, that is, as p becomes closer to 0 and we get a Z-channel. We remark that the well known asymmetric distance (see, for example [2]) is a metric but it is not matched to the BAC.

Elementary continuity arguments can show the following: given $0 < q < 1/2$ and $n \in \mathbb{N}$, there is an $\varepsilon = \varepsilon(q, n)$ such that a BAC with parameters (p, q) admits a matched metric if $p < \varepsilon$ or $q - p < \varepsilon$. This follows from the simple reason that, for any given code, for ε sufficiently small the decision regions (i.e., the sets $B^t(x) := \{y \in \mathcal{X} \mid \Pr(x|y) \geq t\}$) do not change. Direct computations show we may consider $\varepsilon = q^n$, but we do not know whether q^n is or is not maximal.

We conjecture there is a matched metric for any BAC. We briefly describe some approaches to the problem, explaining why they reinforce the conjecture.

A direct and constructive (but naive) approach is to try to symmetrize the matrix $\Pr(x|y)$ working line by line. We start at the first line, where we can substitute each value $\Pr(x_1|y)$ by a variable a_{xy} in such a way that (writing a_{1y} for a_{x_1y})

$$\Pr(x_1|y) = P(x_1|z) \iff a_{1y} = a_{1z} \quad \text{and} \quad \Pr(x_1|y) > P(x_1|z) \iff a_{1y} > a_{1z}. \tag{5}$$ {order}

Since we want the matrix to be symmetric, we enforce that $a_{y1} = a_{1y}$ for all $y \in \mathbb{F}_2^n$. However, on each line of the matrix $A = (a_{xy})$ under construction, we want to have

$$\Pr(x|y) = \Pr(x|z) \iff a_{xy} = a_{xz},$$

and so we enforce this condition on the matrix A under construction as well. In other words, we consider a_{yx_1} and define a_{yz} in a way to respect the order constraint as in (5). We move now to the second line and define the entries for a_{x_2y}, enforcing symmetry by defining $a_{yx_2} = a_{x_2y}$. The change on those indices may demand a rescaling on each line but this may not be possible (for example if a_{yx_1} and a_{yx_2} do not respect the order constraint (5)). Despite the fact that this naive algorithm does not always work, many small dimensional examples worked out with this approach leads us to believe that it may be successful if we alternate the use of this algorithm with the algorithm used for the Z-channel in Section (3).

A non constructive approach to prove on the existence of a master metric for the BAC may be as follows: Given n, let us suppose that

$$D = \begin{pmatrix} A & B \\ B^T & \delta A \end{pmatrix}$$

is a master matrix for the Z-channel as constructed om Section 3. Given $t \in [0, 1]$ we define $B(t) = (b_{xy}(t))$ and $A(t) = a_{xy}(t)$ as follows:

$$\begin{aligned} b_{xy}(t) &= (1-t)\, b_{xy} + t b_{yx} \\ a_{xy}(t) &= (1 + t(\delta - 1))\, a_{xy} \end{aligned}$$

and define

$$D(t) = \begin{pmatrix} A(t) & B(t) \\ B(t)^T & \delta A(t) \end{pmatrix}$$

We remark that each $D(t)$ is symmetric, $D(0) = D$ is metric matched for the Z channel defined by $\Pr(1|0) = 0$ and $D(1) = \begin{pmatrix} \delta A & B^T \\ B & A \end{pmatrix}$ is metric matched for the Z channel defined by $\Pr(0|1) = 0$ (the reflected Z). Considering $0 < p < q < 1/2$, for $n = 2$ there are exactly two BACs (according to whether $p < q^2$ or $q^2 < p$) and similarly, for $n = 3$ there are four such BACs. We verified (I actually did when I wrote this remark) that for $n = 2, 3$ all the BACs may be matched by proper choices of t in $D(t)$.

Those evidences suggest that for any BAC there is a matched metric.

Acknowledgements

Example 2.3 is due to the first author's student, Rafael Gregorio. This work began when both authors were visiting Centre Interfacultaire Bernoulli at École Polytechnique Fédérale de Lausanne, and they thank the CIB for its hospitality. The first author was partially supported by grant 2013/25977-7, São Paulo Research Foundation (Fapesp). The second author was supported in part by NSF Grant #0903517.

References

[1] J. Chung Yaw Chiang and Jack K. Wolf, *On channels and codes for the Lee metric*, Information and Control **19** (1971), 159–173.

[2] Serban D. Constantin and T. R. N. Rao, *On the theory of binary asymmetric error correcting codes*, Inform. and Control **40** (1979), no. 1, 20–36.

[3] D.J. Costello and Jr. Forney, G.D., *Channel coding: The road to channel capacity*, Proceedings of the IEEE **95** (2007), no. 6, 1150–1177.

[4] Michel Marie Deza and Elena Deza, *Encyclopedia of distances*, second ed., Springer, Heidelberg, 2013.

[5] Arman Fazeli, Alexander Vardy, and Eitan Yaakobi, *Generalized sphere packing bound*, arXiv:1401.6596 [cs.IT], 2014.

[6] G. David Forney, Jr., *Geometrically uniform codes*, IEEE Trans. Inform. Theory **37** (1991), no. 5, 1241–1260.

[7] Ankur A. Kulkarni and Negar Kiyavash, *Nonasymptotic upper bounds for deletion correcting codes*, IEEE Trans. Inform. Theory **59** (2013), no. 8, 5115–5130.

[8] James L. Massey, *Notes on coding theory*, class notes for course 6.575 (spring), M.I.T., Cambridge, MAA, 1967.

[9] Gérald Séguin, *On metrics matched to the discrete memoryless channel*, J. Franklin Inst. **309** (1980), no. 3, 179–189.

[10] Jacob Ziv, *Universal decoding for finite-state channels*, IEEE Trans. Inform. Theory **31** (1985), no. 4, 453–460.