Communication over Time–Varying Channels under Power Constraints and Side Information of the Participants

Vladimir B. Balakirsky and A. J. Han Vinck
Institute for Experimental Mathematics
University of Essen
45326 Essen, Germany
Email : vbal, vinck@exp-math.uni-essen.de

Abstract

We consider the data transmission scheme where the sender wants to deliver one of M messages to the receiver over a time–varying binary symmetric channel. The transmitted signal should be formed as a concatenation of n elementary signals and it must have a fixed energy. The total power of the noise is also fixed. This power can be distributed non–uniformly, and we analyze the cases when the Modulator or/and the Decoder know this distribution. Our results include the formulas for probablities of the correct decoding for the $(6,3)$ shortened Hamming code and the formulas for the maximum transmission rates that can be attained with arbitrary small decoding error probabilty by increasing the code length. Numerical illustrations are also given.

I Introduction

We consider communication between two participants : the Sender and the Receiver, over a non–stationary binary symmetric memoryless channel. Our interest is concentrated on the analysis of the influence of the knowledge of the channel parameters by the participants. The Receiver plays passive role in the communication process. His adaptation to the channel is a possible changing of the decision rule while constructing an estimate of the transmitted message. The Sender is active, and he can change the transmitted signals. The kind of adaptation of the Sender to the channel parameters was studied for the so–called channels with localized errors [1], [2], [3] and other papers. This model can be presented as follows. The same message can be encoded by different binary codewords depending on the state of the channel. Each component of the codeword to be sent is then mapped to the transmitted signal by the Modulator in such a way that 0 is transformed to the pulse with the amplitude $+\sqrt{P^*}$ and 1 is transformed to the pulse with the amplitude $-\sqrt{P^*}$. As a result, the total energy of the transmitted signal is equal to nP^* and linearly grows with the code length n. We consider another kind of adaptation at the Sender's side. Namely, we assume that it can be implemented only by the Modulator by sending pulses having different amplitudes depending on the state of the channel at each time instant. We also assume that the total energy of the transmitted signal has to be fixed. The Demodulator creates a binary vector by quantizing n received pulses and sends this vector to the Decoder. The later one constructs an estimate of the transmitted message. The assignments indicated above lead to a binary symmetric channel (BSC) between the Encoder and the Decoder, which describes transmission of each component of the codeword. However, crossover probabilities of the channel can differ for different components.

We begin with the formal statement of the problem in Section 2 and then compute the probabilities of the correct decoding for the $(6,3)$ shortened Hamming code in Section 3. A generalization of the channel capacity to the communication scheme under considerations leads to the notion of the maximum transmission rate that can be attained with arbitrary small decoding error probabilty by increasing the code length. The maximum transmission rates are expressed in Section 4.

II Formal statement of the problem

Let us introduce the following parameters : the number of messages of the Source, M, the code length, n, and the parameters $\pi, \alpha, \beta \in [0,1]$.

We consider the data transmission scheme given in Figure 1. The Source selects one of M messages, and the Controller selects one of 2^n binary vectors of length n in such a way that the j–th component is equal to 1 with probability π and equal to 0 with probability $1-\pi$; the selection is independent for all $j = 1,\ldots,n$. We denote the message by $m \in [M]$, where $[M] \triangleq \{1,\ldots,M\}$, and the vector by $\mathbf{s} \in \{0,1\}^n$.

Having received the message m, the Encoder outputs the codeword

$$\mathbf{x}_m = (x_{m1},\ldots,x_{mn}) \in \{0,1\}^n$$

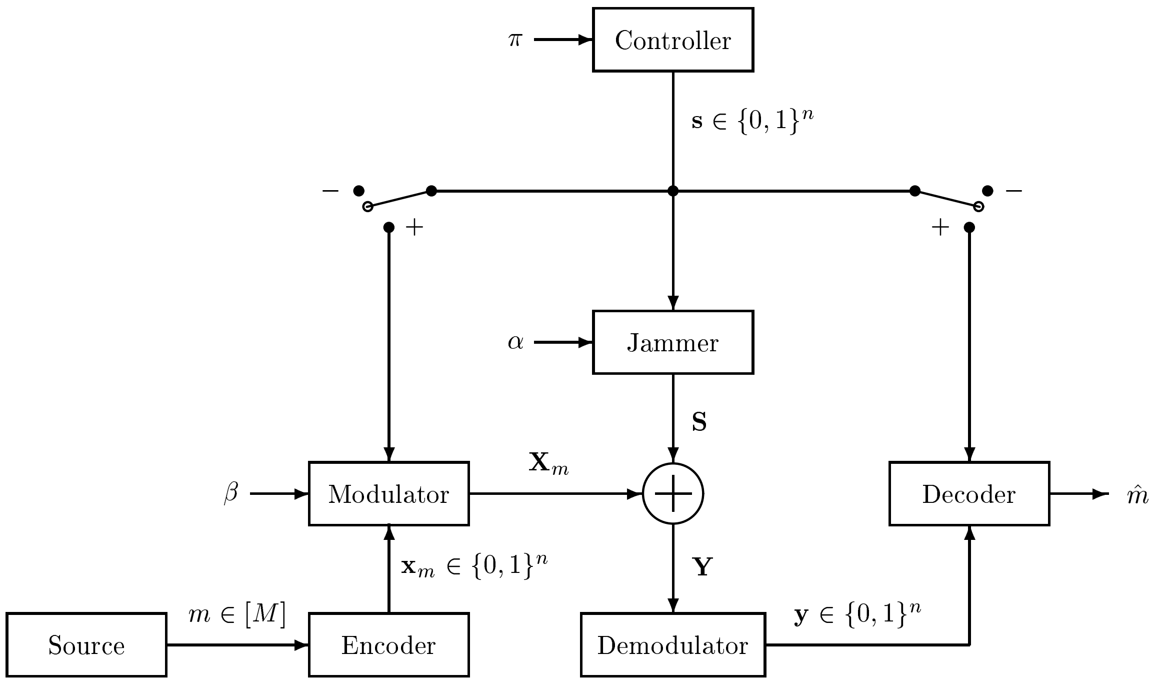

Figure 1: Model of data transmission system with side information.

of a binary code of length n and sends this vector to the Modulator.

Having received the vector $\mathbf{s}$, the Jammer distributes the given power of the noise among positions j where $s_j = 1$ and positions j where $s_j = 0$ in the proportion $\alpha : (1-\alpha)$ and forms a vector $\mathbf{S} = (S_1, \ldots, S_n)$. This vector enters the channel.

If the vector $\mathbf{s}$ is known to the Modulator, then he distributes the given power of the transmitted signal among positions j where $s_j = 1$ and positions j where $s_j = 0$ in the proportion $\beta : (1-\beta)$ and forms a vector $\mathbf{X}_m = (X_{m1}, \ldots, X_{mn})$. If the vector $\mathbf{s}$ is unknown to the Modulator, then this power is uniformly distributed and components of the vector $\mathbf{X}_m$ depend only on m.

The Demodulator receivers the vector $\mathbf{Y} = (Y_1, \ldots, Y_n)$, which is constructed as a component–wise sum of the vectors $\mathbf{X}_m$ and $\mathbf{S}$, i.e.,

$$\mathbf{Y} = \mathbf{X}_m + \mathbf{S}.$$

This vector is mapped to the binary vector $\mathbf{y} = (y_1, \ldots, y_n)$, where

$$y_j = \begin{cases} 1, & \text{if } Y_j > 0 \\ 0, & \text{if } Y_j \leq 0 \end{cases}$$

and $\mathbf{y}$ is the received vector for the Decoder.

The goal of the Decoder is constructing of an estimate of the transmitted message.

Let us introduce a pair of variables

$$(\mathrm{S}, \mathrm{R}) \in \Big\{ (-,-), (+,-), (-,+), (+,+) \Big\}. \qquad (1)$$

The variable S is equal to $+$ or $-$ depending on the point whether the Modulator knows the vector $\mathbf{s}$ or not. The variable R is equal to $+$ or $-$ depending on the point whether the Decoder knows the vector $\mathbf{s}$ or not. Furthermore, assume that if P_j and N_j denote the power distributed to the j–th component of the transmitted signal and the to j–th component of the noise vector, then the conditional probability that y_j is the j–th received component when x_j is the j–th transmitted component is equal to $1 - \omega(P_j/N_j)$ if $y_j = x_j$ and $\omega(P_j/N_j)$ if $y_j \neq x_j$, where $\omega(z) \in (0,1), z \geq 0$, is a given function. We will also denote

$$\Big(p_{\mathrm{B}}, p_{\mathrm{G}} \Big) \triangleq \left(\omega\left(\frac{\rho\pi}{\alpha}\right), \omega\left(\frac{\rho(1-\pi)}{1-\alpha}\right) \right) \qquad (2)$$

$$\Big(p'_{\mathrm{B}}, p'_{\mathrm{G}} \Big) \triangleq \left(\omega\left(\frac{\rho\beta}{\alpha}\right), \omega\left(\frac{\rho(1-\beta)}{1-\alpha}\right) \right) \qquad (3)$$

where

$$\rho \triangleq P/N$$

is the signal–to–noise ratio.

Let the binary channel between the Decoder and the Encoder be a memoryless channel, which is introduced as follows. If the Sender is ignorant about $\mathbf{s}$, then

$$W(\mathbf{y}|\mathbf{x}, \mathbf{s}) \triangleq \prod_{j=1}^{n} W_{x_j}(y_j|s_j) \qquad (4)$$

for all $\mathbf{x}, \mathbf{y}, \mathbf{s} \in \{0,1\}^n$ where

$$W_x(y|s) \triangleq \begin{cases} 1-p_{\rm B}, & \text{if } y = x \text{ and } s = 1 \\ 1-p_{\rm G}, & \text{if } y = x \text{ and } s = 0 \\ p_{\rm B}, & \text{if } y \neq x \text{ and } s = 1 \\ p_{\rm G}, & \text{if } y \neq x \text{ and } s = 0 \end{cases} \quad (5)$$

for all $x, y, s \in \{0,1\}$. If the Sender knows $\mathbf{s}$, then the parameters $p_{\rm B}$ and $p_{\rm G}$ in (5) should be replaced with $p'_{\rm B}$ and $p'_{\rm G}$, respectively.

The simplest non–trivial example, where the situation described above can take place, is the AWGN channel. In this case, we assume that $S_1, \ldots, S_n$ are i.i.d. Gaussian random variables and

$$\mathrm{var}(S_1) + \ldots + \mathrm{var}(S_n) = N$$

where N is given. Let $\mathcal{N}(0, \sigma^2)$ stand for the Gaussian distribution with the mean 0 and the variance σ^2 and let

$$s_j = 1 \implies S_j \sim \mathcal{N}\left(0, \frac{N\alpha}{n\pi}\right)$$
$$s_j = 0 \implies S_j \sim \mathcal{N}\left(0, \frac{N(1-\alpha)}{n(1-\pi)}\right)$$

for all $j = 1, \ldots, n$. If the vector $\mathbf{s}$ is unknown to the Modulator, then

$$X_{mj} = (-1)^{x_{mj}} \sqrt{\frac{P}{n}}.$$

If the vector $\mathbf{s}$ is known to the Modulator, then

$$s_j = 1 \implies X_{mj} = (-1)^{x_{mj}} \sqrt{\frac{P\beta}{n\pi}}$$
$$s_j = 0 \implies X_{mj} = (-1)^{x_{mj}} \sqrt{\frac{P(1-\beta)}{n(1-\pi)}}.$$

Thus,

$$\Big(p_{\rm B}, p_{\rm G}\Big) = \Big(\Phi_\rho\Big(\frac{\pi}{\alpha}\Big), \Phi_\rho\Big(\frac{1-\pi}{1-\alpha}\Big)\Big) \quad (6)$$
$$\Big(p'_{\rm B}, p'_{\rm G}\Big) = \Big(\Phi_\rho\Big(\frac{\beta}{\alpha}\Big), \Phi_\rho\Big(\frac{1-\beta}{1-\alpha}\Big)\Big) \quad (7)$$

where

$$\Phi_\rho(z) \triangleq \frac{1}{\sqrt{3.1415\ldots}} \int_{\sqrt{\rho z/2}}^{\infty} e^{-t^2} dt.$$

III Performance of the decoder for the shortened Hamming $(6,3)$–code and time–varying binary symmetric channel

Suppose that the messages to be sent are binary 3–tuples and that the shortened Hamming $(6,3)$–code is used. The generator matrix of the encoder is a binary 3×6 matrix

$$\mathbf{G}_{6,3} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

and the codeword $\mathbf{x} \in \{0,1\}^6$ is formed as

$$\mathbf{x} = \mathbf{u}\mathbf{G}_{6,3}$$

where $\mathbf{u} \in \{0,1\}^3$ is the message. The code has the minimum distance 3, and all error vectors of weight 1 are corrected if the codeword is transmitted over a BSC having the crossover probability $p \in (0, 1/2)$. Therefore, assuming that the decoding error always occurs if there are several codewords maximizing the conditional probability to be sent, we express the probabilities of the correct decoding for $\mathrm{R} = -$, denoted by $P_{\rm cor}^{(-,-)}$ and $P_{\rm cor}^{(+,-)}$, as

$$P_{\rm cor}^{(-,-)} = P_{\rm cor}^{(+,-)} = (1-p)^6 + 6(1-p)^5 p. \quad (8)$$

where

$$p = \begin{cases} \pi\Phi_\rho\Big(\frac{\pi}{\alpha}\Big) + (1-\pi)\Phi_\rho\Big(\frac{1-\pi}{1-\alpha}\Big), & \text{if } \mathrm{S} = - \\ \pi\Phi_\rho\Big(\frac{\beta}{\alpha}\Big) + (1-\pi)\Phi_\rho\Big(\frac{1-\beta}{1-\alpha}\Big), & \text{if } \mathrm{S} = + \end{cases}$$

and the optimum value of the parameter $\beta \in [0,1]$ minimizes the crossover probability p.

Suppose that the all–zero codeword was transmitted and that $\mathbf{y} = 000001$ is the received vector. The maximum likelihood decoding of the vector $\mathbf{y}$ results in the codeword 000000. However, if the crossover probability of the BSC is $p_{\rm B}$ at positions j with $s_j = 1$ and $p_{\rm G}$ at positions j with $s_j = 0$, then this conclusion can be wrong. For example, let $\mathbf{s} = 001010$. Then

$$W(\mathbf{y}|000000, \mathbf{s}) = (1-p_{\rm G})^3 p_{\rm G}(1-p_{\rm B})^2$$
$$W(\mathbf{y}|001011, \mathbf{s}) = (1-p_{\rm G})^4 p_{\rm B}^2$$

are conditional probabilities of the vector $\mathbf{y}$ to be received when the codewords 000000 and 001011 are sent. If

$$(1-p_{\rm G})p_{\rm B}^2 > p_{\rm G}(1-p_{\rm B})^2$$

then

$$W(\mathbf{y}|001011, \mathbf{s}) > W(\mathbf{y}|000000, \mathbf{s})$$

and 001011 is the result of the maximum likelihood decoding under assumption that the vector $\mathbf{s}$ is known to the decoder. On the other hand, this inequality implies

$$W(\mathbf{y}'|000000, \mathbf{s}) > W(\mathbf{y}'|001011, \mathbf{s})$$

where $\mathbf{y}' = \mathbf{y} \oplus 001011$. Thus, the received vector $\mathbf{y}'$ is correctly decoded. Notice that both $\mathbf{y}$ and $\mathbf{y}'$ belong to the same coset of the $(6,3)$–code and present considerations above in a general form.

Proposition. *Let $\mathcal{E}$ denote the set of coset leaders of a linear code $\mathcal{C}$ having the minimum Hamming weight, i.e., any vector $\mathbf{y} \in \{0,1\}^n$ can be uniquely*

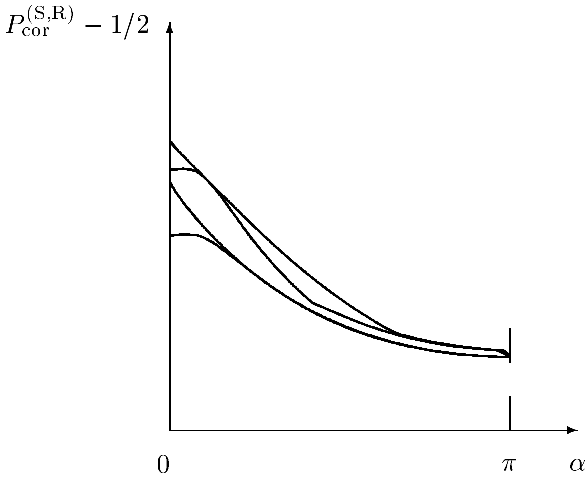

Figure 2: The probabilities of the correct decoding for the (6,3)–code; $\pi = 0.5$, $\rho = 0.6$, $P_{\rm cor}^{(+,+)} \geq P_{\rm cor}^{(-,+)} \geq P_{\rm cor}^{(+,-)} \geq P_{\rm cor}^{(-,-)}$.

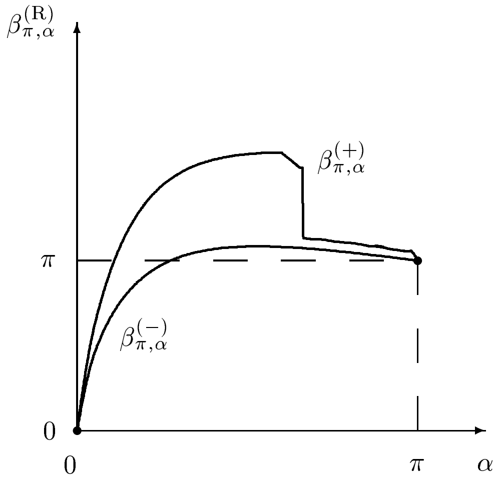

Figure 3: The (6,3)–code : The optimum values of β for $\pi = 0.5$, $\rho = 0.6$, and S = +.

represented as $\mathbf{x} \oplus \mathbf{e}$, *where* $\mathbf{x} \in \mathcal{C}$ *and* $\mathbf{e} \in \mathcal{E}$. *Suppose that the codeword is transmitted over a* BSC *whose crossover probability is* $p_{\rm B}$ *at positions* j *with* $s_j = 1$ *and* $p_{\rm G}$ *at positions* j *with* $s_j = 0$. *If the vector* $\mathbf{s}$ *is known to the Decoder, then the probability of correct decoding can be computed as*

$$\sum_{\mathbf{e} \in \mathcal{E}} f(\mathbf{e}|\mathcal{C}, \mathbf{s})$$

where the function f *is introduced as follows. If there is only one codeword* $\mathbf{x} \in \mathcal{C}$ *maximizing the conditional probability* $W(\mathbf{e}|\mathbf{x}, \mathbf{s})$ *defined in* (4), (5), *then*

$$f(\mathbf{e}|\mathcal{C}, \mathbf{s}) = \max_{\mathbf{x} \in \mathcal{C}} W(\mathbf{e}|\mathbf{x}, \mathbf{s})$$

Otherwise, $f(\mathbf{e}|\mathcal{C}, \mathbf{s}) = 0$.

The set $\mathcal{E}$ for the (6,3)–code can be specified as the set consisting of the all–zero vector, 6 vectors of the Hamming weight 1 and the vector 001100. The statement of Proposition allows us to compute the probability of the correct decoding for this code when R = +. Examples of curves $P_{\rm cor}^{(\rm S,R)}$ as functions of α are given in Figure 2, where the parameter β is chosen in an optimum way for R = +. These values are shown in Figure 3. Notice that if $\rho = 0.6$, then $P_{\rm cor}^{(-,+)} > P_{\rm cor}^{(+,-)}$ for all $p \in (0, 1/2)$, i.e., the knowledge of $\mathbf{s}$ by the Dedoder is more important than the knowledge of the Encoder. However, this conclusion can be wrong for different ρ. We will show the dependence of the maximum transmission rates for the situations $(+,-)$ and $(-,+)$ in the next Section. The relationships between the probabilities of the correct decoding for the (6,3)–code is similar.

IV Asymptotic results

Given an $\varepsilon > 0$, let $\mathrm{SI}_n^{(\rm S,R)}(\varepsilon)$ denote a block code of length n, which allows us to have the average decoding error probability at level ε when the code is used in the data transmission system in conjunction with the maximum likelihood decoding; the pair (S, R) satisfies (1). The known methods of analysis of the channel capacity leads us to the result formulated in Theorem below using the following function (VG stands for "Varshamov–Gilbert") : for all $p \in (0,1)$, denote

$$\mathrm{VG}(p) \triangleq \begin{cases} 1 - h(p), & \text{if } p \leq 1/2, \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

where $h(p) \triangleq -p\log p - (1-p)\log(1-p)$ is the binary entropy function.

Theorem. *For any fixed* $\varepsilon \in (0,1)$, $\pi, \alpha, \beta \in [0,1]$, *and* $n \geq 1$, *let* $M_n^{(\rm S,R)}(\varepsilon)$ *denote the maximum cardinality of an* $\mathrm{SI}_n^{(\rm S,R)}(\varepsilon)$*–code. Then the asymptotically maximum transmission rate can be expressed as*

$$\lim_{n \to \infty} \frac{\log M_n^{(\rm S,R)}(\varepsilon)}{n} = \begin{cases} C_{\pi,\alpha}^{(\rm S,R)}, & \text{if S} = - \\ C_{\pi,\alpha,\beta}^{(\rm S,R)}, & \text{if S} = + \end{cases}$$

where

$$\begin{aligned} C_{\pi,\alpha}^{(-,-)} &\triangleq \mathrm{VG}(\,\pi p_{\rm B} + (1-\pi)p_{\rm G}\,) \\ C_{\pi,\alpha}^{(-,+)} &\triangleq \pi\mathrm{VG}(\,p_{\rm B}\,) + (1-\pi)\mathrm{VG}(\,p_{\rm G}\,) \\ C_{\pi,\alpha,\beta}^{(+,-)} &\triangleq \mathrm{VG}(\,\pi p'_{\rm B} + (1-\pi)p'_{\rm G}\,) \\ C_{\pi,\alpha,\beta}^{(+,+)} &\triangleq \pi\mathrm{VG}(\,p'_{\rm B}\,) + (1-\pi)\mathrm{VG}(\,p'_{\rm G}\,) \end{aligned}$$

and the parameters $p_{\rm B}, p_{\rm G}, p'_{\rm B}, p'_{\rm G}$ *depend on* π, α, β *in accordance with* (2), (3).

The asymptotically maximum transmission rates can be computed for the pairs $(p_{\rm B}, p_{\rm G})$ and $(p'_{\rm B}, p'_{\rm G})$ defined in (6), (7). We denote

$$\begin{aligned} C_{\pi,\alpha}^{(+,\rm R)} &\triangleq \max_{\beta \in [0,1]} C_{\pi,\alpha,\beta}^{(+,\rm R)} \\ \beta_{\pi,\alpha}^{(\rm R)} &\triangleq \arg\max_{\beta \in [0,1]} C_{\pi,\alpha,\beta}^{(+,\rm R)} \end{aligned}$$

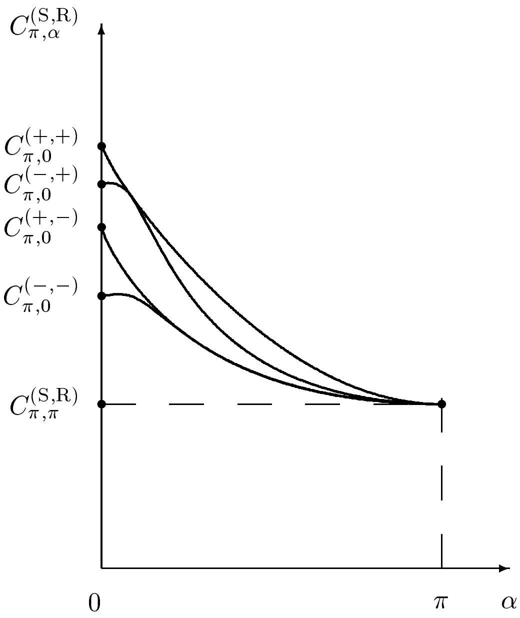

Figure 4: The functions $C_{\pi,\alpha}^{(\mathrm{S,R})}$ for $\pi = 0.5$, $\rho = 0.6$.

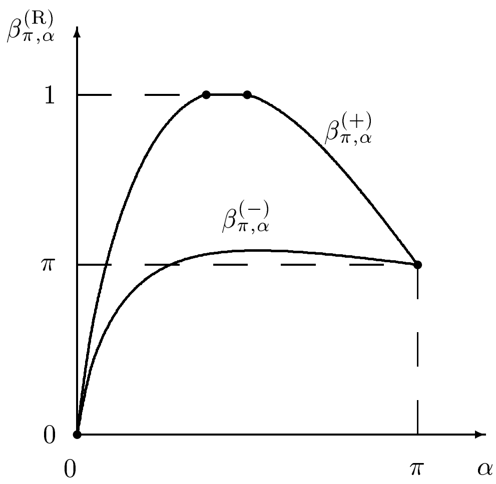

Figure 5: The optimum values of β for $\pi = 0.5$, $\rho = 0.6$, and $\mathrm{S} = +$.

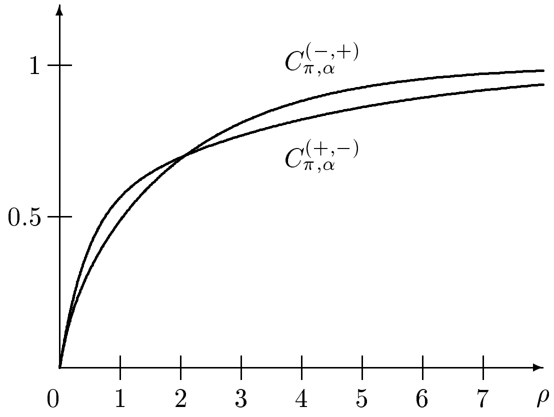

Figure 6: The values of $C_{\pi,\alpha}^{(-,+)}$ and for $C_{\pi,\alpha}^{(+,-)}$ as functions of the signal–to–noise ratio ρ for $\pi = 0.5$ and $\alpha = 0.1$.

and show examples of the functions $C_{\pi,\alpha}^{(\mathrm{S,R})}$ in Figure 4. Notice that

$$\begin{aligned} C_{\pi,0}^{(+,+)} &= \pi + (1-\pi)\mathrm{VG}(\,\Phi_\rho(1))\,) \\ C_{\pi,0}^{(-,+)} &= \pi + (1-\pi)\mathrm{VG}(\,\Phi_\rho(1-\pi)\,) \\ C_{\pi,0}^{(+,-)} &= \mathrm{VG}(\,(1-\pi)\Phi_\rho(1)\,) \\ C_{\pi,0}^{(-,-)} &= \mathrm{VG}(\,(1-\pi)\Phi_\rho(1-\pi)\,) \end{aligned}$$

and

$$C_{\pi,\pi}^{(\mathrm{S,R})} = \mathrm{VG}(\,\Phi_\rho(1)\,), \ \text{ for all } (\mathrm{S,R}).$$

Notice also that the relationships between $C_{\pi,\alpha}^{(\mathrm{S,R})}$, where $(\mathrm{S,R})$ satisfy (1), are very close to the relationships between $P_{\mathrm{cor}}^{(\mathrm{S,R})}$.

The inequality $C_{\pi,\alpha}^{(-,+)} > C_{\pi,\alpha}^{(-,-)}$ follows from convexity of the binary entropy function, while the inequality $C_{\pi,\alpha}^{(+,-)} > C_{\pi,\alpha}^{(-,-)}$ is due to optimization over crossover probabilities of the BSC by the Sender within power constraints. Which of these possibilities is more important depends α, π, and ρ, as it is illustrated in Figure 6. Both possibilities are used when $(\mathrm{S,R}) = (+,+)$.

References

[1] L. A. Bassalygo, S. I. Gelfand, and M. S. Pinsker, “Coding for channels with localized errors”, in : Proc. 4–th Swedish–Soviet Workshop Inform. Theory, Sweden, 1989, pp. 95–99.

[2] R. Ahlswede, L. A. Bassalygo, and M. S. Pinsker, “Localized random and arbitrary errors in the light of arbitrary varying channels”, IEEE Trans. Inform. Theory, vol. 41, 1995, pp. 14–25.

[3] V. B. Balakirsky, “Lower bounds on the code rate for a model of data transmission with side information”, IEEE Trans. Inform. Theory, vol. 44, 1998, pp. 1642–1648.