Budapest University of Technology and Economics

Department of Mathematical Analysis

Generalization of Types and Typical Sequences for Quantum Channels and Asynchronous Multiple Access Channels

Lóránt Farkas

October 5, 2016

1 History of the Research

Information theory was founded by Claude E. Shannon paper "A Mathematical Theory of Communication" Shannon (1949). It has laid the foundation of digital communication today. One important observation in his paper was that random sequences of length n generated by a distribution P fall in a typical set with large probability. The size of this typical set is approximately $2^{n\,\mathrm{H}(P)}$ where n is the length of the sequence and $\mathrm{H}(P)$ is the Shannon entropy of the distribution. So, the sequences generated by distribution P can be viewed as uniformly chosen from the typical set of P. Shannon has used typical sequences as a key tool to prove his fundamental theorem that communication over a noisy channel is possible with arbitrary small error probability if the transmission rate is smaller than a quantity called channel capacity. Proofs of this kind are called achievability proof (viz. the rate can achieve the capacity). Later, Feinstein (1954) showed that if the error probability goes to zero, then the rate must be smaller than the capacity —this type of proof is called weak converse.

Shannon fundamental theorem has been generalized also for quantum systems. For the classical-quantum channel —where classical information is transmitted through quantum channel— Holevo (1973) gave a converse proof. The generalization of typical sequences to quantum systems gives typical subspaces, see Ohya and Petz (1993). Holevo (1998), Hausladen *et al.* (1997) could prove the achievability of the Holevo capacity with typical subspaces.

In classical information theory, the next step was analyzing the speed of convergence of error probability for fixed rate $R < C$ as codeword length n goes to infinity. The speed of convergence turned out exponential, but typical sequences were not powerful enough to determine the exponent, analyzed by Fano (1955), Gallager (1965), Shannon Gallager and Berlekamp (1967). However, with a simple refinement we get a very efficient tool, that can be used for this purpose. This refinement, known as the method of types, is to partition the n-length sequences into classes according to type (empirical distribution). The number of classes grows polynomially while each type class size grows exponentially, so the error exponent is equal to the exponent of the worst class. Preliminary application of the types idea appear in Boltzmann (1877), Schrödinger (1931), Hoeffding (1956) and Sanov (1961). The joint type of several sequences appears in the works of Blahut (1977), Dobrushin and Stambler (1975) and Goppa (1975). It was developed to a general method by Imre Csiszár and his research group (János Körner and Katalin Marton). The best results on discrete memoryless error exponents bounds, avaible currently, are due to them, e.g. Csiszár and Körner (2011). This method was used to study error exponents of multiple access channels (MAC) derived by Pokorny and Wallmeier (1985). Their bound was improved —also with the method of types but with another decoding function— by Liu and Hughes (1996) and more recently by Nazari (2011).

For two user MAC the region of rate pairs achievable with codewords of fixed types has been determined in Ahlswede (1971), Liao (1972) as a pentagon. The convex closure of such pentagons can be achieved by time sharing and this convex closure is the capacity region, if the transmitters send their codewords synchronously. That means there is no delay between the starting times of their codewords. In asynchronous systems it is not possible to employ time sharing, and it had been an accepted surmise that the capacity region was the union of

pentagons. Two papers devoted to prove were written by Poltyrev (1983) and by Hui and Humblet (1985). The proof of Poltyrev is mathematically correct, except for one step that we have corrected. Poltyrev's paper did not give general achievability. The achievability proof of Hui and Humblet is vulnerable in several points, and their converse result was not relevant mathematically (see Farkas and Kói (2014a)). The method of types has not been employed to asynchronous multiple access systems prior to our work.

2 Goals and Results

One of my goals was to give an alternate achievability proof for the classical-quantum channels. The cited authors proved achievability via giving a Positive Operator Valued Measure (POVM). I wanted a more transparent proof, similar to its classical counterpart. By orthogonalizing the typical subspaces I got a von Neumann measurement —that is such a POVM where the positive operators are projections. This new measurement has good mathematical properties that I could also use to derive the capacity region of the quantum compound channel —which channel was defined firstly by me.

Another goal was to better understood asynchronous multiple access channels (AMAC). Scrutinizing these systems I have given several models. I have shown that if the time delay between the codewords is random and uniformly distributed, or the delay is constant but not known and arbitrary, then the capacity region equals the union of the pentagons, for all considered models. I also gave an example (where the delay is not uniform) whose capacity region is between the union and its convex closure.

Finally, I wanted to generalize the method of types to asynchronous multiple access channels and to give an error exponent for such systems. With the new concepts of subtypes and δ-balanced sequences (the latter combines the method of types and typical sequences in a way) I gave a numerically computable exponent, at least for simple examples. In addition to the usual models, I have also analyzed controlled asynchronous systems, where the transmitters shift their codewords deliberately. A heuristic approach indicates that controlled asynchronism may improve the error exponent. A numerical calculation showed that the lower bound of the exponent of the controlled asynchronous model is better than the best available lower bound of the exponent for the synchronous model. I conjecture that the exponent of the controlled asynchronous system can be better than the synchronous one. So far, I could not prove this, while the available upper bounds are unmanageable or inefficient.

2.1 Theses for quantum channels

Mathematically a quantum system can be represented by a density —Hermitian, nonnegative— operator over a Hilbert space $\mathcal{H}$. In my work the Hilbert space is finite d dimensional, i.e. $\mathcal{H} = \mathbb{C}^d$.

Definition 1. A *state of a quantum system* is a density matrix $\rho \in \mathbb{C}^{d\times d}$, Hermitian $\rho = \rho^*$, nonnegative $\langle x, \rho x\rangle \geq 0$, $\forall x \in \mathbb{C}^d$ and $\mathrm{Tr}(\rho) = 1$.

A state of a complex quantum system that consists of n components is a density matrix over $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \ldots \otimes \mathcal{H}_n$.

Definition 2. A *von Neumann measurement* with L possible outcomes is a collection of orthogonal projections $\mathbf{\Pi} = \Pi_1, \Pi_2, \ldots, \Pi_L$, $\sum_i \Pi_i = I$. Performing the measurement on quantum state ρ, the i'th outcome corresponding to Π_i happens with probability $\mathrm{Tr}(\rho\Pi_i)(= \mathrm{Tr}(\Pi_i\rho\Pi_i))$. The quantum state after the measurement becomes

$$\tilde{\rho} = \frac{\Pi_i\rho\Pi_i}{\mathrm{Tr}(\Pi_i\rho\Pi_i)}.$$

For $\mathcal{H} = \mathbb{C}^m$ let $\mathcal{B}(\mathcal{H})$ the set of all matrices over $\mathcal{H}$, i.e., $\mathcal{B}(\mathcal{H}) = \mathbb{C}^{m\times m}$.

Definition 3. A *quantum channel* $\mathcal{E} : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$ is a completely positive trace preserving (CPTP) map.

A CPTP map can be represented by Kraus operators. These are $m \times d$ matrices K_i such that $\sum_{i=1}^{dm} K_i K_i^* = I$ and $\mathcal{E}(\rho) = \sum_{i=1}^{nm} K_i^* \rho K_i$.

Definition 4. The *n'th memoryless extension* of a quantum channel $\mathcal{E} : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$ is the channel $\mathcal{E}^{\otimes n} : \mathcal{B}(\mathcal{H}_1^{\otimes n}) \to \mathcal{B}(\mathcal{H}_2^{\otimes n})$ such that

$$\mathcal{E}^{\otimes n}(\rho_1 \otimes \rho_2 \otimes \ldots \otimes \rho_n) = \mathcal{E}(\rho_1) \otimes \mathcal{E}(\rho_2) \otimes \ldots \otimes \mathcal{E}(\rho_n)$$

Definition 5. The *von Neumann entropy* of a density matrix ρ is defined as $\mathrm{S}(\rho) = -\mathrm{Tr}(\rho \log(\rho))$.

In this thesis logarithms are to base 2.

Definition 6. *Holevo's quantity* is defined as

$$\chi(\mathcal{E}, P, \omega_1^l) = \mathrm{S}(\mathcal{E}(\omega)) - \sum_{i=1}^{l} p_i \, \mathrm{S}(\mathcal{E}(\omega^i)) \tag{1}$$

where $\mathcal{E} : \mathbb{C}^{m\times m} \to \mathbb{C}^{d\times d}$ is a quantum channel, P is a probability distribution on $\{1, 2, \ldots, l\}$ with corresponding probabilities $\{p_1, p_2, \ldots, p_l\}$, ω_1^l is an l-tuple of density operators $(\omega^1, \omega^2, \ldots, \omega^l)$, $\omega^i \in \mathbb{C}^{m\times m}$ and $\omega = \sum_{i=1}^{l} p_i \omega^i$.

Definition 7. *Holevo's capacity* of quantum channel $\mathcal{E}$ is defined as

$$\chi(\mathcal{E}) = \sup_{P,\omega_1^l} \chi(\mathcal{E}, P, \omega_1^l) \tag{2}$$

Definition 8. A quantum codebook $\mathcal{C}$ of length n is a collection of n-times tensor products of quantum states —all states from the same $\mathcal{B}(\mathcal{H})$. The rate of this codebook is $R = {}^{\log(|\mathcal{C}|)}/_{n}$, where $|\mathcal{C}|$ denotes the number of quantum states in $\mathcal{C}$.

In classical-quantum channel, classical (not quantum) information is transmitted through quantum channel. To transmit the i'th message from a message set of size $|\mathcal{C}|$, the i-th quantum state ρ_i of $\mathcal{C}$ is sent through channel $\mathcal{E}$ (more exactly its n'th memoryless extension). At the output of the channel a measurement —in this thesis a von Neumann measurement $\mathbf{\Pi}$— is performed, its outcome estimates the transmitted message. The estimate coincides with i with probability $\mathrm{Tr}(\mathcal{E}^{\otimes n}(\rho_i)\Pi_i))$.

Definition 9. The *average error probability* of classical-quantum channel with codebook $\mathcal{C}$ and von Neumann measurement $\mathbf{\Pi}$ is

$$P_e(\mathcal{C}, \mathcal{E}, \mathbf{\Pi}) = \frac{1}{|\mathcal{C}|} \sum_{i=1}^{|\mathcal{C}|} (1 - \mathrm{Tr}(\mathcal{E}^{\otimes n}(\rho_i)\Pi_i))$$

Thesis 1 (Theorem 2.10 and Remark 2.11, Farkas (2008))**.** *For any quantum channel $\mathcal{E} : \mathbb{C}^{m\times m} \to \mathbb{C}^{d\times d}$ and rate $R < \chi(\mathcal{E})$ and $\varepsilon > 0$, for $n \geq N(\varepsilon, R, \mathcal{E})$ there exists a quantum codebook $\mathcal{C}$ of length n with rate R and a von Neumann measurement $\mathbf{\Pi}$ such that $P_e(\mathcal{C}, \mathcal{E}, \mathbf{\Pi}) \leq \varepsilon$.*

In classical information theory a compound channel can be defined by two separate practical problems leading to a common mathematical model. First, imagine a common transmitter and many channels with many receivers, none of the receiver knows his channel. The transmitter wants to transmit its data, in such a way that all receivers can decode its message. This leads to the same mathematical model as if there was one transmitter one receiver and one unknown channel from a family of channels.

A quantum compound channel is the generalization of the second approach only, as due to quantum no-cloning, there is no quantum channel that has two outputs and the input can be detected with small error at both outputs.

The quantum compound channel is defined as a set $\mathscr{E}$ of quantum channels $\mathcal{E}$. The sender as before employs a quantum codebook $\mathcal{C}$. At the output, a sequence of measurements will be made. Other works apply cleverly defined states and measurements to identify the channel. I do not use such techniques, rather a sequence of von Neumann measurement SM is used to decode the message from the output of the quantum channel.

Definition 10. The *average error probability of the compound classical-quantum channel* is

$$P_e^{Compound}(\mathcal{C}, \mathscr{E}, SM) = \max_{\mathcal{E}^j \in \mathscr{E}} \frac{1}{|\mathcal{C}|} \sum_{i=1}^{|\mathcal{C}|} \Pr\{SM(\mathcal{E}(\rho_i)) \neq i\}$$

where $SM()$ denotes the result of the sequence of measurements.

Let

$$\chi(\mathscr{E}) = \sup_{P,\omega_1^l} \min_{\mathcal{E}\in\mathscr{E}} \chi(\mathcal{E}, P, \omega_1^l). \tag{3}$$

Thesis 2 (Theorem 2.14, Farkas (2008))**.** *For any finite set of quantum channels $\mathscr{E}$, $R < \chi(\mathscr{E})$ and $\varepsilon > 0$, for $n \geq N_1(\varepsilon, R, \mathscr{E})$ there exists a quantum codebook $\mathcal{C}$ of length n with rate R and a von Neumann measurement sequence SM such that $P_e^{Compound}(\mathcal{C}, \mathscr{E}, SM) \leq \varepsilon$*

2.2 Theses for General converse

From this point on vectors (finite sequences) will be denoted by boldface symbols. Random variables will be denoted by U, X, Y assumed to take values in finite sets $\mathcal{U}$, $\mathcal{X}$, $\mathcal{Y}$, etc. called alphabets. Their (joint) distributions

are denoted by P^U, P^{UX}, P^{UXY}, etc. or V^U, V^{UX}, V^{UXY}, etc. In the sequel the $P^X(X = x)$ or $V^{UX}(U = u \wedge X = x)$ will be abbreviated by $P^X(x)$ or $V^{UX}(u, x)$. We will conveniently regard any probability measure V on $\mathcal{U} \times \mathcal{X} \times \mathcal{Y}$, say, as the joint distribution V^{UXY} of dummy random variables U, X, Y. Then V^U, V^{UX}, etc. denote the marginal distributions of V on $\mathcal{U}$, $\mathcal{U} \times \mathcal{X}$, etc. and $V^{X|U}$, say, denotes the conditional distribution given by $V^{X|U}(x|u) = V^{UX}(u, x)/V^U(u)$. This notational convention amounts to simultaneously assign different distributions to the same (dummy) random variables. The family of all probability measures on $\mathcal{U} \times \mathcal{X}$, say, —thus, by our assumpition, the family of all distributions of random variables with range $\mathcal{U} \times \mathcal{X}$— is denoted by $\mathcal{P}(\mathcal{U} \times \mathcal{X})$. The family of all conditional distributions on $\mathcal{X}$ conditioned on $\mathcal{U}$ is denoted by $\mathcal{P}(\mathcal{X}|\mathcal{U})$.

Definition 11. The *Shannon entropy* of a distribution P is:

$$\mathrm{H}(P) \triangleq \mathrm{H}_P(X) = \sum_{x \in \mathcal{X}} -P(x) \log P(x). \tag{4}$$

Definition 12. The *conditional Shannon entropy* $\mathrm{H}_V(Y|X)$ can be defined as

$$\mathrm{H}_V(Y|X) \triangleq \mathrm{H}_V(X, Y) - \mathrm{H}_V(X) \tag{5}$$

Definition 13. *Mutual information* of a pair of random variables is

$$\begin{aligned} \mathrm{I}_V(X \wedge Y) \triangleq & \mathrm{H}_V(Y) - \mathrm{H}_V(Y|X) = \mathrm{H}_V(Y) + \mathrm{H}_V(X) - \mathrm{H}_V(Y, X) \\ = & \mathrm{H}_V(X) - \mathrm{H}_V(X|Y). \end{aligned} \tag{6}$$

In this section $[i]$ denotes the set $\{1, 2, \ldots, i\}$. A K-senders asynchronous discrete memoryless multiple-access channel (K-AMAC) is defined by K finite input alphabets $\mathcal{X}_m, m \in [K]$, a finite output alphabet $\mathcal{Y}$, a stochastic matrix $W : \mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_K \to \mathcal{Y}$ describing the probability distribution of the output given the inputs and a delay system (defined in Definition 18 below).

Each sender $i \in [K]$ has a set of messages $\mathcal{M}_i$. To transmit one of them the i'th sender assigns to it a "codeword" of length n. Mathematically

Definition 14.

$$\mathrm{f}_i : \mathcal{M}_i = \{1, 2, \ldots, M_i\} \to \mathcal{X}_i^n$$

denotes an *Encoding Function* for transmitter i. The list of assigned codewords $\mathbf{x}_{i,m} \in \mathcal{X}_i^n$, $m \in [M_i]$ is called *Codebook* of sender i and denoted by $\mathcal{C}_i$.

Definition 15. A *codebook system of block-length n with rate vector*

$$\mathbf{R} = (R_1, R_2 \ldots, R_K)$$

for a given K-AMAC consists of K codebooks $C_1, C_2, \ldots, C_K$, where the codebook C_m of the m-th sender has 2^{nR_m} codewords of length n whose symbols are from $\mathcal{X}_m$.

The system is symbol synchronized, so it can be described by a channel matrix, but not frame synchronized —the first symbols of the codewords need not match. In real life this can happen if the signals of different transmitters need different times to reach the receiver. This problem causes a constant but

unknown delay. In a more general case, e.g. if the codebooks are used more than once, the differences between the timing of the receiver and the timings of the senders are represented by a K-tuple of random variables, called delay vector as in Definition 18 below.

The senders have two-way infinite sequences of random messages, and assign codewords to their consecutive messages. The codewords go through the channel. The sequences of the senders' codewords and hence also the output symbol sequence are two-way infinite sequences. Fix the location of the 0-th output symbol. The message of sender $m \in [K]$ whose codeword affects the 0'th output is denoted by $M_{m,0}$. This restricts the delays to be in the set $\{0, 1, \ldots, n-1\}$ (see Figure 1).

Definition 16. For each integer $j \in \mathbb{Z}$ and for each $m \in [K]$ let the j'th message of sender m $M_{m,j}$ be a uniformly distributed random variable taking values in the set $\{1, 2, \ldots, 2^{nR_m}\}$. All these random variables are independent of each other. The two-way infinite sequence $\{M_{m,j}, j \in \mathbb{Z}\}$ represents the *message flow* sent by the m-th sender.

Definition 17. For each integer $j \in \mathbb{Z}$ and for each $m \in [K]$ let the $M_{m,j}$-th codeword in C_m be $\mathbf{X}_{m,j} (= \mathrm{f}_m(M_{m,j}))$. Let $nj + i$'th symbol of *flow of codewords*, denoted by $X_{m,nj+i}$, be the i-th symbol of $\mathbf{X}_{m,j}$ where $i \in \{0, 1, \ldots, n-1\}$.

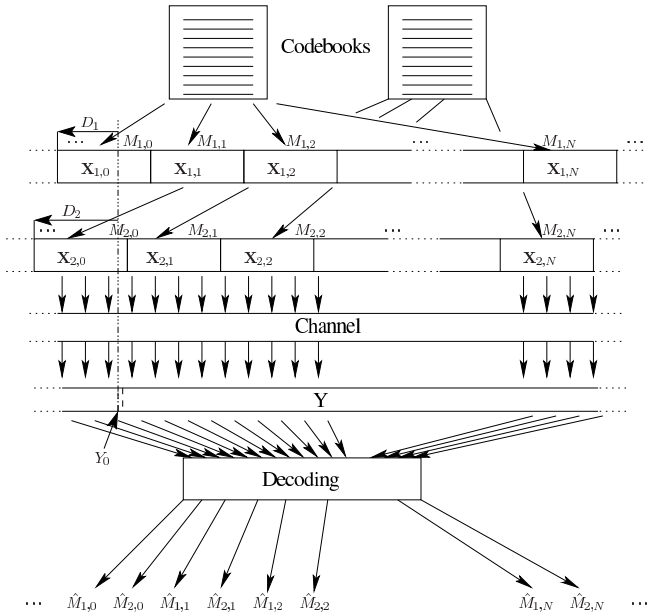

Figure 1: The Setting for Two Senders

Definition 18. For each $n \in \mathbb{Z}^+$, let

$$\mathbf{D}(n) = (D_1(n), D_2(n), \ldots, D_K(n))$$

be a K-tuple of random variables, not necessarily independent of each other but independent of all previously defined random variables, taking values in the set $\{0, 1, \ldots, n-1\}$. $D_m(n)$ will represent the *delay of sender m* relative to the receiver's timing. The joint distribution of delays is known to the senders and the receiver. The realizations of the random variables $D_1(n), D_2(n), \ldots, D_K(n)$ are not known to the senders and, depending on the model, may be known or unknown to the receiver. The sequence $\mathbf{D} = \{\mathbf{D}(1), \mathbf{D}(2), \ldots, \mathbf{D}(n), \ldots\}$ will be called the *delay system*. With a slight abuse of notation, we also write $\mathbf{D}$ instead of $\mathbf{D}(n)$.

Example 1. For each $n \in \mathbb{Z}^+$ and for each $m \in [K]$ $D_m(n)$ has uniform distribution on $\{0, 1, \ldots, n-1\}$ and they are independent. Following Hui and Humblet (1985) it is called the totally asynchronous case.

Example 2. Let $K = 2$, for each $n \in \mathbb{Z}^+$ let $D_1(n)$, $D_2(n)$ be independent random variables uniformly distributed on the even numbers of $\{0, 1, \ldots, n-1\}$. It is called the even delays case.

For fixed n, the output sequence is defined as:

Definition 19. Let Y_{nj+i} be the *output random variable* of the channel with transition matrix W when the inputs are $X_{1,nj+i+D_1(n)}$, $X_{2,nj+i+D_2(n)}$, $\ldots$, $X_{K,nj+i+D_K(n)}$ where $i \in \{0, 1, \ldots, n-1\}$.

The decoder can be defined in several ways. The definitions below give the strongest version of the converse theorem.

Definition 20. An *informed infinite decoder* is defined as a function which assigns to each two way infinite output sequence realization $\{y_l, l \in \mathbb{Z}\}$ and each realization of $\mathbf{D}(n) = (D_1(n), D_2(n), \ldots, D_K(n))\,, n \in \mathbb{Z}^+$, a K-tuple of decoded messages $\{\hat{M}_{m,0}, m \in [K]\}$.

It is assumed that the same but shifted decoding procedure occurs at the output points $\{nk, k \in \mathbb{Z}\}$. Hence the random variables of the estimations $\{\hat{M}_{m,j}, m \in [K]\,, j \in \mathbb{Z}\}$ are also defined.

Definition 21. The *(average) error probability* is

$$P_e^n = \Pr\left\{\bigcup_{m=1}^{K}\left\{M_{m,0} \neq \hat{M}_{m,0}\right\}\right\}. \tag{7}$$

Definition 22. For a given K-AMAC the rate vector

$$\mathbf{R} = (R_1, R_2, \ldots, R_K)$$

is *achievable* if for every $\varepsilon > 0$, $\delta > 0$ for all $N \in \mathbb{Z}^+$ there exists a coding/decoding system with blocklength $n > N$ with rates coordinate-wise exceeding $(R_1 - \delta, R_2 - \delta, \ldots, R_K - \delta)$ and with error less than ε. The set of achievable rate vectors form the *capacity region* of the given K-AMAC.

For each subset S of $[K]$ write

$$\mathbf{X}_S = (X_m)_{m \in S},\ S^c = [K] \setminus S, \tag{8}$$

and for all $\mathbf{R} = (R_1, R_2, \ldots, R_K)$ write

$$R(S) = \sum_{m \in S} R_m. \tag{9}$$

Let $\mathbf{D}$ denote the delay vector. Let $\mathbf{X}_{S,i+D_S}$ denote the random vector with components $X_{l,i+D_l}$, $l \in S$ where $X_{m,j}$ is defined as in definition 16; similar notation is used where $+$ is replaced by $\oplus$ which means addition modulo n.

Thesis 3 (Theorem 3.8, Farkas and Kói (2011), Farkas and Kói (2014a))**.** *For every K-AMAC and every coding/informed infinite decoder system of length n, there exist such distribution $V^{\mathbf{X}_i} \in \mathcal{P}(\mathcal{X}_i^n)$ that for all $S \subset [K]$*

$$R(S) \leq \mathrm{I}(\mathbf{X}_{S,Q\oplus D_S} \wedge Y | \mathbf{X}_{S^c,Q\oplus D_{S^c}}, Q, \mathbf{D}) + 2\varepsilon_n. \tag{10}$$

holds. Here the joint distribution of random variables in the mutual information can be calculated from

$$V^{\mathbf{X}_1\mathbf{X}_2\ldots\mathbf{X}_K Q\mathbf{D}}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k, q, \mathbf{d}, y) = \tag{11}$$
$$= \left(\prod_{i=1}^{K} V(\mathbf{x}_i)\right) {}^1\!/_n V(\mathbf{d}) W(y | x_{1,q\oplus d_1}, x_{2,q\oplus d_2}, \ldots, x_{k,q+d_k})$$

and $\varepsilon_n = (R([K]))P_e^n + \frac{1}{n}$. Notice the that, the joint distribution in (11) *defined on a larger space than the joint distribution in* (10)*.*

Remark 1. The mutual information in (10) is meant under the joint distribution of the involved random variables computed from (11).

Remark 2. The upper bound in (10) is true for stronger —that can estimate from less information— decoder.

It is possible to modify the decoder to get an engineeringly viable model:

Definition 23. An *uninformed L-block decoder*, $L \in \mathbb{Z}^+$, is defined as a function which assigns to each $(2Ln+1)$-tuple $\{y_l, l \in \{-Ln, \ldots, 0, \ldots, Ln\}\}$ of possible output realizations a K-tuple of messages $\{\hat{M}_{m,0}, m \in [K]\}$.

As before, it is assumed that the same but shifted decoding procedure occurs at the output points $\{nk, k \in \mathbb{Z}\}$. Hence the random variables of the estimations $\{\hat{M}_{m,j}, m \in [K], j \in \mathbb{Z}\}$ are also defined.

The capacity regiom of Example 2 differs both from the synchronous and the totally asynchronous one.

Thesis 4 (Theorem 3.12, Farkas and Kói (2011), Farkas and Kói (2014a))**.** *For $K = 2$ sender in the even delays case (Example 2), for either kind of decoder the capacity region consists of those rate pairs that either belong to*

$$\begin{aligned} 0 \leq R_1 &\leq I_P(X_1 \wedge Y | X_2) \\ 0 \leq R_2 &\leq I_P(X_2 \wedge Y | X_1) \\ R_1 + R_2 &\leq I_P(X_1, X_2 \wedge Y) \end{aligned} \tag{12}$$

for some distribution $P(x_1, x_2, y) = P_1(x_1)P_2(x_2)W(y|x_1, x_2)$ or are linear combinations with weights $\frac{1}{2}, \frac{1}{2}$ of such pairs. Moreover, using coding/decoding systems of odd length, only rate pairs as in (12) *can be achieved.*

2.3 Theses for AMAC exponents

From this point on $[k]$ denotes the set $\{0, 1, 2, \ldots, k-1\}$.

Definition 24. *Multi-Information* is defined as

$$\begin{aligned} &\mathrm{I}_V(X_1 \wedge X_2 \wedge X_3 \wedge \cdots \wedge X_k) \triangleq \\ &\quad = \mathrm{H}_V(X_1) + \mathrm{H}_V(X_2) + \mathrm{H}_V(X_3) + \cdots + \mathrm{H}_V(X_k) - \\ &\quad\quad - \mathrm{H}_V(X_1, X_2, X_3, \ldots, X_k) \end{aligned} \tag{13}$$

Definition 25. The *Kullback-Leibler distance or I-divergence* is defined on a pair of distribution (P, V) over the same set, say, $\mathcal{X}$. The divergence of P from V is

$$\mathrm{D}(P\|V) \triangleq \sum_{x \in \mathcal{X}} P(x) \log \left(\frac{P(x)}{V(x)} \right) \tag{14}$$

Previously the input symbols of an AMAC were denoted by $X_1, X_2, \ldots$ and the output was denoted by Y. Here, only a two sender AMACs are analyzed, so the inputs will be denoted by X and Y and the output by Z.

A 2-senders MAC is defined by 2 finite input alphabets $\mathcal{X}, \mathcal{Y}$, a finite output alphabet $\mathcal{Z}$, and a stochastic matrix $W : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$. The matrix W may be unknown to the senders and the receiver.

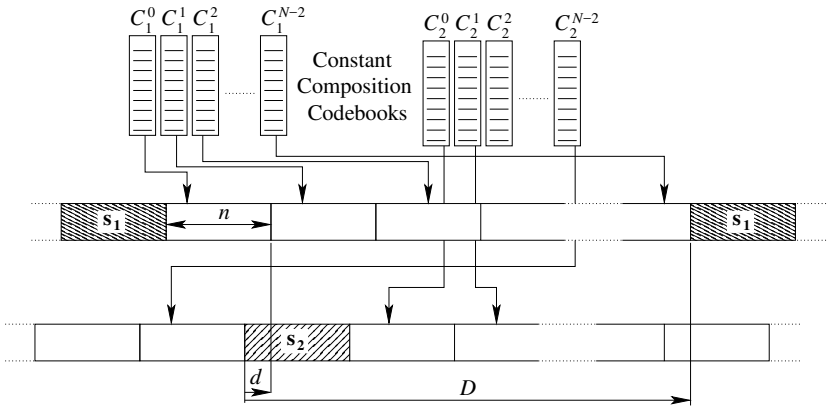

Figure 2: Encoding with sync-sequences and delays

Definition 26. The *Type* of an n-length sequence $\mathbf{x} = x_1 x_2 \ldots x_n \in \mathcal{X}^n$ is the distribution $\mathcal{P}_{\mathbf{x}} \in \mathcal{P}(\mathcal{X})$ where $\mathcal{P}_{\mathbf{x}}(x)$ is the relative frequency of the symbol x in $\mathbf{x}$. The joint type of two or more n-length sequences is defined similarly and, for $(\mathbf{u}, \mathbf{x}) \in \mathcal{U}^n \times \mathcal{X}^n$, say, it is denoted by $\mathcal{P}_{(\mathbf{u},\mathbf{x})}$.

Definition 27. The *family of all possible types* of sequences $\mathbf{x} \in \mathcal{X}^n$ of length n is denoted by $\mathcal{P}^n(\mathcal{X})$, and for $P \in \mathcal{P}^n(\mathcal{X})$ the set of all $\mathbf{x} \in \mathcal{X}^n$ of type $\mathcal{P}_{\mathbf{x}} = P$ is the *Type Class* of P denoted by T_P^n.

Definition 28. An *asynchronous constant composition code-book system with codewords of length n, periodicity N, types $P^X \in \mathcal{P}^n(\mathcal{X})$ and $P^Y \in \mathcal{P}^n(\mathcal{Y})$ and with rate parameters pair (R_1, R_2)*, consists of code-books C_1^i, C_2^i, $i \in [N-1]$ and sync-sequences $\mathbf{s}_1$, $\mathbf{s}_2$. The codewords in C_m^i, as well as $\mathbf{s}_m$, $m \in \{1,2\}$ are sequences in $\mathcal{X}^n$ resp. $\mathcal{Y}^n$ of type P^X resp. P^Y, and each code-book C_m^i consists of 2^{nR_m} codewords.

The codebook system is used as follows: each sender has a two-way infinite sequence of random messages chosen from the message sets $\mathcal{M}_m = [2^{nR_m}]$, $m \in \{1,2\}$, independently and with uniform distribution. For each integer j, if the j-th message of sender $m \in \{1,2\}$ is $i \in [2^{nR_m}]$, this sender selects the i'th codeword of the codebook C_m^j, with j taken modulo $N-1$. After the codewords from C_m^{N-2}, the sync-sequence $\mathbf{s}_m$ is inserted. See, Figure 2.

Asynchronism causes a delay $D \in [nN]$ between the sync-sequences. It is either unknown to the senders or (in case of controlled asynchronism) is chosen by them. The delay between blocks (codewords) is $d \equiv D \pmod n$. Most concepts below refer to a given value of D. See, Figure 2 and Figure 3.

The receiver is assumed to be able to identify the position of the sync-sequences. Decoding is performed in the window of length nN shown in Figure 3, where the shaded blocks correspond to sync-sequences. As the synchronization blocks are not used for information transmission, the effective or real rate pair is $(R_1(1-\frac{1}{N}), R_2(1-\frac{1}{N}))$. To make this close to (R_1, R_2), we chose N large (but not depending on n).

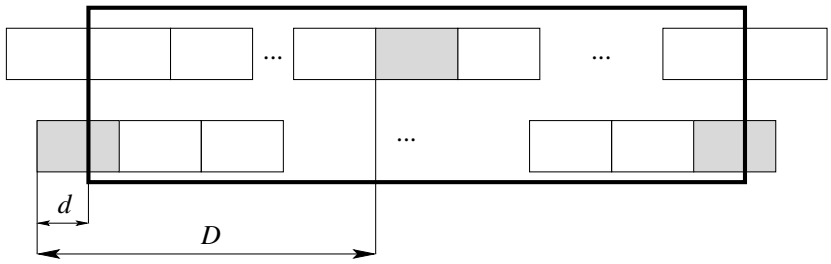

Figure 3: The decoding window and the delays

Formally, the decoding window contains $N-1$ consecutive codewords and the sync-sequence of sender 1, and $N-1$ codewords complemented by a final and initial part of the sync-sequence of sender 2. Let $M_1^0, \ldots, M_1^{N-2}$ and $M_2^0, \ldots, M_2^{N-2}$ denote the messages of senders 1 resp. 2 whose codewords are in the window. The upper indices refer to the receiver's time, thus the messages M_1^j and M_2^j may have occurred at different senders' time and their codewords may come from codebooks of different indices of senders 1 and 2. Suppose the realizations of the message sequences are $\mathbf{i} = (i_0, \ldots, i_{N-2})$ and $\mathbf{j} = (j_0, \ldots, j_{N-2})$. Then the first and second rows of Fig. 3 are filled by the concatenations

$$\mathbf{x}(\mathbf{i}) = \mathbf{x}^l(i_0) \ldots\ldots \mathbf{x}^{N-2}(i_{N-2-l})\mathbf{s}_1\mathbf{x}^0(i_{N-2-l+1}) \ldots\ldots \mathbf{x}^{l-1}(i_{N-2}) \tag{15}$$

$$\mathbf{y}(\mathbf{j}, d) = \mathbf{s}_2''\mathbf{y}^0(j_0) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \mathbf{y}^{N-2}(j_{N-2})\mathbf{s}_2' \tag{16}$$

of the corresponding codewords. Here the upper indices denote the index of codebooks. The sync-sequence parts $\mathbf{s}_2'$ and $\mathbf{s}_2''$ denotes the first/last $d/n - d$

symbols of $\mathbf{s}_2$. In the notation, d refers to a cyclic shift by d positions. The receiver from the corresponding output of the memoryless channel W decodes the messages corresponding to input sequences (15)–(16).

Definition 29. The *average error probability* of the above model is

$$P_e^D = \Pr\{M_m^j \neq \hat{M}_m^j \text{ for some } j \in [N-1], m \in \{1,2\}\}, \tag{17}$$

where M_m^j is the j'th message of user m in the window, and $\hat{M}_m^j$ is its estimation for some decoder.

Let $\mathcal{V}^P$ be the collection of all $V^{XYZ} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$ with marginals $V^X = P^X$ and $V^Y = P^Y$. Further, let

$$P^{XYZ}(x,y,z) = P^X(x)P^Y(y)W(z|x,y).$$

Definition 30. For given channel W, and input distributions P^X, P^Y let

$$\begin{aligned} E_j^\alpha(K) = \min_{V_1,V_2,V_{12} \in \mathcal{V}^P} & \beta_1 \,\mathrm{D}(V_1 \| P^{XYZ}) + \beta_2 \,\mathrm{D}(V_2 \| P^{XYZ}) \\ & + \beta_{12} \,\mathrm{D}(V_{12} \| P^{XYZ}) + |\beta_1(\mathrm{I}_{V_1}(X \wedge YZ) - R_1) \\ & + \beta_2(\mathrm{I}_{V_2}(Y \wedge XZ) - R_2) \\ & + \beta_{12}(\mathrm{I}_{V_{12}}(X \wedge Y \wedge Z) - R_1 - R_2)|^+, \end{aligned} \tag{18}$$

where $0 \leq \alpha \leq 1$ and the $\beta_1, \beta_2, \beta_{12}$ coefficients are given by

(a) Odd K, odd j: $\beta_1 = \beta_2 = \alpha$, $\beta_{12} = 1 - \alpha + \frac{K-1}{2}$

(b) Odd K, even j: $\beta_1 = \beta_2 = 1 - \alpha$, $\beta_{12} = \alpha + \frac{K-1}{2}$

(c) Even K, odd j: $\beta_1 = 0$, $\beta_2 = 1$,$\beta_{12} = \frac{K}{2}$

(d) Even K, even j: $\beta_1 = 1$, $\beta_2 = 0$,$\beta_{12} = \frac{K}{2}$

Definition 31.

$$E^\alpha = \min_{j \in \{0,1\}, K \in [2N-1]} E_j^\alpha(K). \tag{19}$$

Thesis 5 (Theorem 4.11, Corollary 4.12, Farkas and Kói (2014b), Farkas and Kói (2015))**.** *There exist an asynchronous constant composition code-book system with codewords of length n, periodicity N, types P^X and P^Y and with rate parameters pair (R_1, R_2) and a decoder such that for each $\gamma > 0$ and n sufficiently large the error probability P_e^D satisfies for all W and D*

$$P_e^D \leq 2^{-n(E^\alpha - \gamma)}, \quad \alpha = \frac{d}{n}.$$

Hence, the AMAC system error probability $P_e^{as} = \max_D P_e^D$ and the error probability of optimal controlled asynchronous transmission $P_e^{cas} = \min_D P_e^D$ satisfy

$$P_e^{as} \leq 2^{-n\left(\min\limits_{\alpha \in [0,1]} E^\alpha - \gamma\right)}, \quad P_e^{cas} \leq 2^{-n\left(\max\limits_{\alpha \in [0,1]} E^\alpha - \gamma\right)}$$

Thesis 6 (Section 4.6, Farkas and Kói (2014b), Farkas and Kói (2015))**.** *The above exponent can be numerically calculated for the binary adder channel. The obtained controlled asynchronous exponent and the best synchronous exponent in the literature for $R_1 = R_2 = R$ are shown on Figure 4. The controlled asynchronous exponent is higher in some rate region than its synchronous counterpart.*

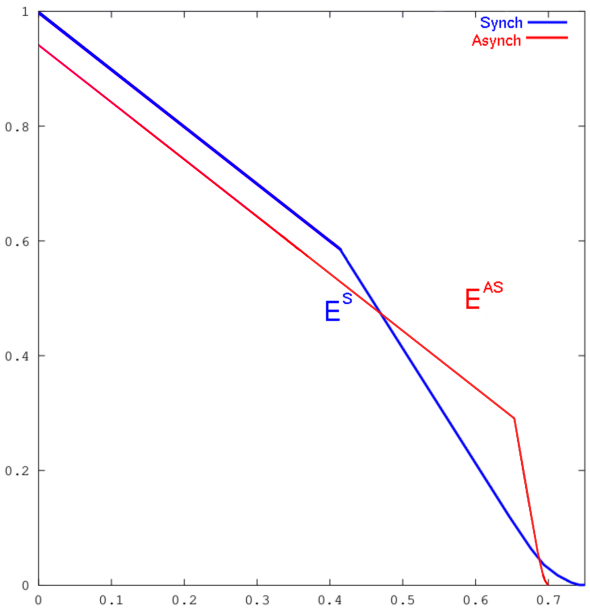

Figure 4: The horizontal axis shows the rate of the codebooks, the vertical axis the numerical values of the synchronous (blue line) and asynchronous (red line) exponents, the latter calculated with periodicity $N = 20$.

References

[1] C. E. Shannon, "A mathematical theory of communication.", *Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, 1949.

[2] A. Feinstein, *A new basic theorem of information theory*. Massachusetts Institute of Technology, Research Laboratory of Electronics, 1954.

[3] R. Ahlswede, "Multi-way communication channels", in *Information Theory Proceedings (ISIT)*, Sep. 1971, pp. 23–52.

[4] H. H.-J. Liao, "Multiple Access Channels", Deparment Electrical Engineering, University of Hawai, Honolulu, 1972.

[5] A. S. Holevo, "Some estimates of the information transmitted by quantum communication channel", *Problemy Peredachi Informatsii*, vol. 9, pp. 3–11, 3 Jan. 1973.

[6] G. S. Poltyrev, "Coding in an asynchronous multiple-access channel", *Problemy Peredachi Informatsii*, vol. 19, no. 3, pp. 12–21, 1983.

[7] J. Y. N. Hui and P. A. Humblet, "The capacity region of the totally asynchronous multiple-access channel", *IEEE Transaction on Information Theory*, vol. 31, pp. 207–216, 2 Mar. 1985.

[8] J. Pokorny and H. Wallmeier, "Random coding bound and codes produced by permutations for the multiple-access channel", *IEEE Transaction on Information Theory*, vol. 31, pp. 741–750, Dec. 1985.

[9] Y.-S. Liu and B. L. Hughes, "A new universal random coding bound for the multiple-access channel", *IEEE Transaction on Information Theory*, vol. 42, pp. 376–386, Mar. 1996.

[10] P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, and W. K. Wootters, "Sending classical information via noisy quantum channels", *Physical Review A*, vol. 56, pp. 56–138, 1 Jul. 1997.

[11] A. S. Holevo, "Capacity of the quantum channel with general signal states", *IEEE Transaction on Information Theory*, vol. 44, pp. 269–273, 1 Jan. 1998.

[12] L. Farkas, "Detecting the normal quantum/compound quantum channel with von Neumann measurement", in *Information Theory Proceedings (ISIT)*, 2008, pp. 354–358.

[13] I. Csiszár and J. Körner, *Information theory, coding theorems for discrete memoryless systems, 2^{nd} edition*. Cambridge University Press, 2011.

[14] L. Farkas and T. Kói, "Capacity region of discrete asynchronous multiple access channels", in *Information Theory Proceedings (ISIT)*, 2011, pp. 2273–2277.

[15] A. Nazari, *Error exponent for discrete memoryless multiple-access channels*. Ph.D. Thesis, University of Michigan, 2011. [Online]. Available: `http://web.eecs.umich.edu/~anastas/docs/ali_thesis.pdf`.

[16] L. Farkas and T. Kói, "On the capacity region of discrete asynchronous multiple access channels", *Kybernetika*, vol. 50, pp. 1003–1031, 6 Dec. 2014.

[17] L. Farkas and T. Kói, "Universal error exponent for discrete asynchronous multiple access channels", in *Information Theory Proceedings (ISIT)*, Jul. 2014, pp. 2944–2948.

[18] L. Farkas and T. Kói, "Controlled asynchronism improves error exponent", in *Information Theory Proceedings (ISIT)*, Jun. 2015, pp. 2638–2642.