An Upper Bound on the Distribution of Computation of a Sequential Decoder for Multiple-Access Channels. *

V.B.Balakirsky
The author is with the Data Security Association "Confident",
193060 St.-Petersburg, Russia
E-mails: vbal@stoic.spb.su, vbal@mathematik.uni-bielefeld.de

Abstract. The computational distribution of sequential decoding for discrete memoryless multiple-access channels is examined. It is shown that all possible collections of incorrect paths in the code tree may be described by bipartite graphs. Using this fact and a new decoding metric it is proved that the number of computations in the first incorrect subtree is a Paretean random variable, and that the parameter of the Pareto distribution is estimated similarly to the parameter for systems of information transmission with one source.

Key words :
Multiple-access channel, tree codes, sequential decoding.

*The work is supported by a Scholarship from the Swedish Institute, Stockholm, Sweden.

1 Introduction.

Sequential decoding is known as a powerful decoding technique for convolutional codes, which allow the level of the channel noise to determine the level of decoding computation. Computational efforts of the decoder may be estimated by the distribution of computation in the first incorrect subtree conditioned on correct decoding. It is known that this distribution is asymptotically Paretean in various ensembles of random convolutional and tree codes [1-6].

We consider the application of sequential decoding to the multiple-access channels (MAC's), when two sources transmit statistically independent messages to a common receiver (Fig.1). The decoder constructs estimates of both messages, and a decoding error occurs if at least one of them differs from the message that was sent. Known results for this model include an upper bound on the decoding error probability when maximum-likelihood decoded random block, tree, and convolutional codes are used [7-9]. We develop approaches described in [10-12] and examine sequential decoding methods.

The main difficulty arising when the computational distribution of a sequential decoder should be upper-bounded is connected with the fact that collections of incorrect paths should be taken into account. These paths may have common edges in the code tree, and hence corresponding code sequences may have common subblocks [3]. Decoding for MAC when tree codes are being used by both encoders means that the decoder selects a path in a tree which is the Cartesian product of two trees [8]. In this case, incorrect paths may have common edges either in the first or in the second tree.

Considerations of this paper are based on the following ideas. First, we introduce a new metric of the sequential decoder which differs from the metric in [10,11]. This metric gives an opportunity to derive an appropriate upper bound on the contribution of incorrect paths into the computational efforts of the decoder which does not depend on the received sequence. The key component of this bound is the concept of grouping error events with the common topology into subsets and applying random coding arguments separately to each of these subsets [3,4,8]. Classification of different collections of incorrect paths by bipartite graphs allows us to realize this scheme. This classification and the other new points of analysis, compared to [3-5,13] where the idea of grouping error events was also developed, appear because of two trees used to encode the messages.

The paper is organized as follows. In Section 2 we introduce the notation which will be used in analysis. The sequential decoding algorithm for MAC, the main result, and its proof based on Lemmas 1,2 are given in Section 3. Lemma 1 splits the ρ-th moment of the number of computations into two multipliers that describe the influence of the values of the metric along the correct path and incorrect paths. This approach is well-known [3-5,14], but we give a short proof of that lemma in Appendix 1 to explain the notations. In Lemma 2 we estimate each multiplier. These estimates are based on the properties of the decoding metric function introduced in Section 4. An estimate of the first multiplier for the correct path is simple, and the proof is shown in Appendix 1. The main efforts are concentrated on the estimate of the second multiplier for the collections of incorrect paths. The proof is shown in Section 5. Appendix 2 contains some illustrations of the averaging over bipartite graphs that describe a configuration of two incorrect paths in the code trees.

2 Notations.

We suppose that at each time instant t, the k-th source generates a letter $u_k^t \in \{0, ..., M_k - 1\}$, and the sequence

$$\mathbf{u}_k = (\ u_k^1, u_k^2, ...\ )$$

is encoded by a tree code G_k of rate

$$R_k = \frac{\ln M_k}{N}$$

nats per channel symbol. It means that M_k edges leave each node of the code tree, and code subblocks of length N, whose symbols belong to a given finite set X_k, are associated with every edge. Since the t-tuple

$$\mathbf{u}_k^t = (\ u_k^1, ..., u_k^t\ )$$

uniquely defines the edge of the tree leading to the t-th level, we say that the sequence $\mathbf{u}_k$ defines a path $\mathbf{u}_k^1, \mathbf{u}_k^2, ...$ in the tree and denote the code sequence by

$$\mathbf{x}_k(\mathbf{u}_k) = (\ x_k(\mathbf{u}_k^1), x_k(\mathbf{u}_k^2), ...\ ),$$

where $x_k(\mathbf{u}_k^t) \in X_k^N$ is the code subblock associated with the edge $\mathbf{u}_k^t$.

We assume symbol and edge synchronization between encoders G_1 and G_2. Then a pair of source sequences $(\mathbf{u}_1, \mathbf{u}_2)$ defines a path in a two-user tree which is the Cartesian product of trees G_1 and G_2. Let us represent the two-user tree as

$$G = (\ G_1, G_2\ ),$$

identify its nodes by pairs $(\ \mathbf{u}_1^t, \mathbf{u}_2^t\ )$, and associate the pair of code subblocks $(\ x_1(\mathbf{u}_1^t), x_2(\mathbf{u}_2^t)\ )$ with the edge leading to the node $(\ \mathbf{u}_1^t, \mathbf{u}_2^t\ )$; $t = 1, 2, ...$ For a formal convenience, we set

$$\begin{aligned}
\mathbf{u} &= (\ \mathbf{u}_1, \mathbf{u}_2\ ) = (\ u^1, u^2, ...\ ), \\
\mathbf{u}^t &= (\ \mathbf{u}_1^t, \mathbf{u}_2^t\ ) = (\ u^1, ..., u^t\ ), \\
\mathbf{x}(\mathbf{u}) &= (\ \mathbf{x}_1(\mathbf{u}_1), \mathbf{x}_2(\mathbf{u}_2)\ ) = (\ x(\mathbf{u}^1), x(\mathbf{u}^2), ...\ ),
\end{aligned}$$

where

$$\begin{aligned}
u^t &= (\ u_1^t, u_2^t\ ), \\
x(\mathbf{u}^t) &= (\ x_1(\mathbf{u}_1^t), x_2(\mathbf{u}_2^t)\ ).
\end{aligned}$$

In further considerations, we will assume that $\mathbf{u}$ denotes the transmitted pair of messages or a path in G, but the notations above will also be used for any path in G. In particular, we define as

$$\mathcal{U}_1(\mathbf{u}) = \Big\{ \ \hat{\mathbf{u}} = (\ \hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2 \ ) = (\ \hat{u}^1, \hat{u}^2, ... \ ) : \ \ \hat{u}^1 \neq u^1 \ \Big\}$$

the set of paths belonging to *the first incorrect subtree.* Besides, for a given r, we define as

$$\mathcal{U}_{1,r}(\mathbf{u}) = \{ \ U_r = (\ \mathbf{u}(1), ..., \mathbf{u}(r) \ ) : \ \ \mathbf{u}(i) \in \mathcal{U}_1(\mathbf{u}) \ \text{ for all } \ i = 1, ..., r \ \}$$

the set consisting of collections of r paths in the first incorrect subtree. If the length of the examined paths is equal to j, then we use the following notation :

$$\mathcal{U}_1^j(\mathbf{u}) = \Big\{ \ \hat{\mathbf{u}}^j = (\ \hat{\mathbf{u}}_1^j, \hat{\mathbf{u}}_2^j \ ) = (\ \hat{u}^1, ..., \hat{u}^j \ ) : \ \ \hat{u}^1 \neq u^1 \ \Big\}$$

and

$$\mathcal{U}_{1,r}^j(\mathbf{u}) = \Big\{ \ U_r^j = (\ \mathbf{u}^j(1), ..., \mathbf{u}^j(r) \ ) : \ \ \mathbf{u}^j(i) \in \mathcal{U}_1^j(\mathbf{u}) \ \text{ for all } \ i = 1, ..., r \ \Big\}.$$

Let a MAC be defined by the transition probabilities

$$\{ \ P(\ y|x_1, x_2 \ ), \ \ (\ x_1, x_2, y \ ) \in (\ X_1^N, X_2^N, Y^N \ ) \ \}$$

so that the probability to receive a sequence

$$\mathbf{y} = (\ y^1, y^2, ... \ ), \ \ y^1, y^2, ... \in Y^N,$$

provided that the pair of sequences $(\ \mathbf{x}_1(\mathbf{u}_1), \mathbf{x}_2(\mathbf{u}_2) \ )$ was sent, can be written as

$$P(\ \mathbf{y}|\mathbf{x}_1, \mathbf{x}_2 \ ) = \prod_{t \geq 1} P(\ y^t|x_1(\mathbf{u}_1^t), x_2(\mathbf{u}_2^t) \ ).$$

The notations above allow us to represent this equality as

$$P(\ \mathbf{y}|\mathbf{x} \ ) = \prod_{t \geq 1} P(\ y^t|x(\mathbf{u}^t) \ ).$$

Any sequential decoder associates the value of a *cumulative metric* to every node of the tree and uses it to check if the code sequence assigned to the path leading to the node is likely to be the transmitted one. We denote the cumulative metric of the node $\mathbf{u}^l$ by

$$\Gamma(\ \mathbf{u}^l|\mathbf{y}^l\ ) = \sum_{t=1}^{l} m(\ x(\mathbf{u}^t)|y^t\ ),$$

where $m(\ x(\mathbf{u}^t)|y^t\ )$ is the value of a *metric function* assigned to the edge of G, leading to the node $\mathbf{u}^t$.

Let us introduce the probability distributions Q_1 and Q_2 on X_1^N and X_2^N, and the probabilistic ensemble

$$\mathcal{A} = \{\ X^N Y^N,\ \ Q(x)\cdot P(y|x)\ \},$$

where

$$\begin{aligned} X^N &= (\ X_1^N, X_2^N\ ), \\ x &= (\ x_1, x_2\ ) \in X^N, \\ Q(x) &= Q_1(x_1)\cdot Q_2(x_2). \end{aligned}$$

Let $\mathcal{E}_1[\,]$, $\mathcal{E}_2[\,]$, and $\mathcal{E}_{1,2}[\,]$ denote the following averaging operations in $\mathcal{A}$:

$$\begin{aligned} \mathcal{E}_k[\ (*)\ ] &= \sum_{x_k\in X_k^N} Q_k(x_k)\cdot(*), \quad k=1,2, \\ \mathcal{E}_{1,2}[\ (*)\ ] &= \sum_{x\in X^N} Q(x)\cdot(*), \end{aligned}$$

for all $x = (x_1, x_2) \in X^N$, where $(*)$ is an expression depending on x. We also denote

$$\begin{aligned} \mathcal{E}_k^\alpha[\ (*)\ ] &= (\mathcal{E}_k[\ (*)\ ])^\alpha, \quad k=1,2, \\ \mathcal{E}_{1,2}^\alpha[\ (*)\ ] &= (\mathcal{E}_{1,2}[\ (*)\ ])^\alpha \end{aligned}$$

for all α. In particular, we introduce Gallager-type functions for the MAC [7] as follows :

$$E_1(\rho) \ = \ -\ln \sum_{y\in Y^N} \sum_{x_2\in X_2^N} Q_2(x_2)\cdot\left(\sum_{x_1\in X_1^N} Q_1(x_1)\cdot P^{\frac{1}{1+\rho}}(y|x_1,x_2)\right)^{1+\rho}$$

$$
\begin{aligned}
&= -\ln \sum_{y\in Y^N} \mathcal{E}_2\left[\, \mathcal{E}_1^{1+\rho}\left[\, P^{\frac{1}{1+\rho}}(y|x)\,\right]\right] \\
E_2(\rho) &= -\ln \sum_{y\in Y^N} \sum_{x_1\in X_1^N} Q_1(x_1)\cdot\left(\sum_{x_2\in X_2^N} Q_2(x_2)\cdot P^{\frac{1}{1+\rho}}(y|x_1,x_2)\right)^{1+\rho} \\
&= -\ln \sum_{y\in Y^N} \mathcal{E}_1\left[\, \mathcal{E}_2^{1+\rho}\left[\, P^{\frac{1}{1+\rho}}(y|x)\,\right]\right], \\
E_{1,2}(\rho) &= -\ln \sum_{y\in Y^N} \left(\sum_{x_1\in X_1^N}\sum_{x_2\in X_2^N} Q_1(x_1)\cdot Q_2(x_2)\cdot P^{\frac{1}{1+\rho}}(y|x_1,x_2)\right)^{1+\rho} \\
&= -\ln \sum_{y\in Y^N} \mathcal{E}_{1,2}^{1+\rho}\left[\, P^{\frac{1}{1+\rho}}(y|x)\,\right]
\end{aligned}
$$

for all $\rho \geq 0$.

3 Sequential Decoding for MAC. Proof of the Main Result Based on Auxiliary Results.

We assume that the decoder operates in accordance with the stack algorithm [14,16] : there is a stack, containing some paths in G; the decoder selects a path with the largest cumulative metric, extends it to the next level in

$$M = M_1 \cdot M_2$$

possible ways, and stores M new paths in the stack. A *computation* in the first incorrect subtree is defined to be an extension of a path belonging to $\mathcal{U}_1(\mathbf{u})$. Then the number of computations may be estimated as follows [3,15]:

$$C_1(\ \mathbf{u}|\mathbf{y}\ ) \leq \sum_{l\geq 0} \sum_{j\geq 1} \gamma^{(lj)}(\ \mathbf{u}|\mathbf{y}\ ), \tag{3.1}$$

where

$$\gamma^{(lj)}(\ \mathbf{u}|\mathbf{y}\ ) = \exp\{\ -\Gamma(\ \mathbf{u}^l|\mathbf{y}^l\ )\ \} \cdot \sum_{\hat{\mathbf{u}}^j \in \mathcal{U}_1^j(\mathbf{u})} \exp\{\ \Gamma(\ \hat{\mathbf{u}}^j|\mathbf{y}^j\ )\ \}. \tag{3.2}$$

For a given constant c, we denote by

$$P_c(\mathbf{u}) = \sum_{\mathbf{y}} P(\ \mathbf{y}|\mathbf{x}(\mathbf{u})\ ) \cdot \chi\{\ \ C_1(\ \mathbf{u}|\mathbf{y}\ ) > c\ \ \} \tag{3.3}$$

the probability that the number of computations of the sequential decoder for a MAC is greater than c, where $\chi\{\ \}$ denotes the indicator function : $\chi\{\beta\} = 1$ if the statement β is true, and $\chi\{\beta\} = 0$, otherwise. Let $\overline{P}_c(\mathbf{u})$ be the average value of the probability $P_c(\mathbf{u})$ calculated over the code ensemble such that every code subblock assigned to every edge in G_k takes a value $x \in X_k^N$ by a random and independent selection from X_k^N with probability $Q_k(x)$; $k = 1, 2$.

Theorem. For given constants $\varepsilon_1, \varepsilon_2 > 0$, let the parameter ρ be chosen in such a way that

$$\min\{\ \delta_1(\rho),\ \delta_2(\rho),\ \delta_{1,2}(\rho)\ \} > 0,$$

where

$$\begin{aligned}
\delta_1(\rho) &= E_1(\rho) - N\rho(R_1 + \varepsilon_1), \qquad (3.4)\\
\delta_2(\rho) &= E_2(\rho) - N\rho(R_2 + \varepsilon_2),\\
\delta_{1,2}(\rho) &= E_{1,2}(\rho) - N\rho(R_{1,2} + \varepsilon_{1,2}),\\
R_{1,2} &= R_1 + R_2,\\
\varepsilon_{1,2} &= \varepsilon_1 + \varepsilon_2,
\end{aligned}$$

and $E_1(\rho)$, $E_2(\rho)$, $E_{1,2}(\rho)$ are Gallager-type functions for a MAC. Then there exists a metric function $m = \{m(x|y)\}$ such that

$$\overline{P}_c(\mathbf{u}) \le c^{-\rho} \cdot K_\rho, \qquad (3.5)$$

where

$$K_\rho = \frac{1}{1-\Delta(\rho)} \cdot \sum_{j\ge 1} j^2 e^{-jN\varepsilon} < \infty \ \text{ if } \rho \le 1 \qquad (3.6)$$

and

$$K_\rho = \left(\frac{1}{1-\Delta^{1/\rho}(\rho)} \cdot \sum_{j\ge 1} (r! \cdot j^r)^{2/\rho} e^{-jN\varepsilon/\rho} \right)^\rho < \infty \ \text{ if } \rho > 1. \qquad (3.7)$$

Hereafter,

$$\begin{aligned}
\Delta(\rho) &= \Delta_1(\rho) + \Delta_2(\rho) + \Delta_{1,2}(\rho), \qquad (3.8)\\
\Delta_1(\rho) &= e^{-\delta_1(\rho)} + e^{-\delta_1(\rho/r)},\\
\Delta_2(\rho) &= e^{-\delta_2(\rho)} + e^{-\delta_2(\rho/r)},\\
\Delta_{1,2}(\rho) &= e^{-\delta_{1,2}(\rho)} + e^{-\delta_{1,2}(\rho/r)},\\
\varepsilon &= \min\{\ \varepsilon_1, \varepsilon_2\ \},
\end{aligned}$$

and r is the integer such that $r-1 < \rho \le r$.

Corollary. Let $\{\pi(\mathbf{u})\}$ be a probability distribution on pairs of information sequences which can be transmitted over a MAC and let

$$P_c = \sum_{\mathbf{u}} \pi(\mathbf{u}) \cdot P_c(\mathbf{u}).$$

There exist a metric function m and a code belonging to the code ensemble introduced above such that

$$P_c \leq c^{-\rho} \cdot K_\rho, \tag{3.9}$$

where K_ρ is defined in (3.6) and (3.7).

Proof. Using Chebyshev's inequality [15] in (3.3) we may write :

$$P_c(\mathbf{u}) \leq c^{-\rho} \cdot C_1^{(\rho)}(\mathbf{u}), \tag{3.10}$$

where

$$C_1^{(\rho)}(\mathbf{u}) = \sum_{\mathbf{y}} P(\ \mathbf{y}|\mathbf{x}(\mathbf{u})\ ) \cdot C_1^{\rho}(\ \mathbf{u}|\mathbf{y}\ ). \tag{3.11}$$

Therefore, the analysis of $P_c(\mathbf{u})$ reduces to the analysis of the ρ-th moment of the number of computations in the ensemble of noisy sequences and code ensemble.

For a formal convenience, we split the averaging over the code ensemble into the concatenation of two operators $E\,[\,]$ and $E'\,[\,]$, where $E'\,[\,]$ denotes averaging over all incorrect subblocks in G_1 and G_2, and $E\,[\,]$ denotes averaging over the code subblocks of transmitted sequences $\mathbf{u}_1$ and $\mathbf{u}_2$. In particular, if $\overline{C_1^{(\rho)}}(\mathbf{u})$ denotes the average value of $C_1^{(\rho)}(\mathbf{u})$ calculated over the code ensemble, then

$$\overline{C_1^{(\rho)}}(\mathbf{u}) = E\left[\ E'\left[\ C_1^{(\rho)}(\mathbf{u})\ \right]\right].$$

Lemma 1. Let r be the integer such that $r-1 < \rho \leq r$. Then,

$$\overline{C_1^{(\rho)}}(\mathbf{u}) \leq \sum_{l\geq 0}\sum_{j\geq 1} F_\rho^{(lj)}(\mathbf{u}),\ \ \text{if}\ \ \rho \leq 1 \tag{3.12}$$

and

$$\overline{C_1^{(\rho)}}(\mathbf{u}) \leq \left(\sum_{l\geq 0}\sum_{j\geq 1}(F_\rho^{(lj)}(\mathbf{u}))^{1/\rho}\right)^{\rho},\ \ \text{if}\ \ \rho > 1, \tag{3.13}$$

where

$$\begin{aligned} F_\rho^{(lj)}(\mathbf{u}) &= \sum_{\mathbf{y}} E\left[\ \Phi_\rho^{(l)}(\ \mathbf{u},\mathbf{y}\ ) \cdot \left(\Psi_r^{(j)}(\ \mathbf{u},\mathbf{y}\ )\right)^{\rho/r}\ \right], \qquad (3.14)\\ \Phi_\rho^{(l)}(\ \mathbf{u},\mathbf{y}\ ) &= P(\ \mathbf{y}|\mathbf{x}(\mathbf{u})\ ) \cdot \exp\{\ -\rho\Gamma(\ \mathbf{u}^l|\mathbf{y}^l\ )\ \},\\ \Psi_r^{(j)}(\ \mathbf{u},\mathbf{y}\ ) &= \sum_{U_r^j \in \mathcal{U}_{1,r}^j(\mathbf{u})} E'\left[\ \prod_{i=1}^{r} \exp\{\ \Gamma(\ \mathbf{u}^j(i)|\mathbf{y}^j\ )\ \}\ \right]. \end{aligned}$$

Lemma 2. There exists a decoding metric function such that

$$\sum_{\mathbf{y}} E\left[\ \Phi_\rho^{(l)}(\ \mathbf{u}, \mathbf{y}\ )\ \right] \leq \Delta^l(\rho), \tag{3.15}$$

and

$$\Psi_r^{(j)}(\ \mathbf{u}, \mathbf{y}\ ) \leq (r! \cdot j^r)^2 \cdot e^{-jN\varepsilon} \tag{3.16}$$

for any $\mathbf{u}$ and $\mathbf{y}$, where $\Delta(\rho)$ and ε are defined in (3.8).

Lemma 1 and inequality (3.15) of Lemma 2 are proved in Appendix 1. Inequality (3.16) of Lemma 2 is proved in Section 5.

To prove the theorem we use inequalities (3.15), (3.16) in (3.14). Summation over j and l in accordance with (3.12) and (3.13) leads to coefficients K_ρ given in (3.6), (3.7). The final result follows from (3.10).

4 The Metric Function and its Properties.

For all $\rho > 0$, we introduce the probability distributions $\{V_1^{(\rho)}(x_1|x_2,y)\}$, $\{V_2^{(\rho)}(x_2|x_1,y)\}$, and $\{V_{1,2}^{(\rho)}(x|y)\}$, i.e.,

$$V_1^{(\rho)}(x_1|x_2,y),\ V_2^{(\rho)}(x_2|x_1,y),\ V_{1,2}^{(\rho)}(x|y)\ \geq 0$$

and

$$\sum_{x_1\in X_1^N} V_1^{(\rho)}(x_1|x_2,y) = \sum_{x_2\in X_2^N} V_2^{(\rho)}(x_2|x_1,y) = \sum_{x\in X^N} V_{1,2}^{(\rho)}(x|y) = 1$$

for all $(x_1,x_2,y) \in (X_1^N, X_2^N, Y^N)$. Let

$$\begin{aligned} M_1^{(\rho)}(x_1|x_2,y) &= \ln\frac{V_1^{(\rho)}(x_1|x_2,y)}{Q_1(x_1)} - N(R_1+\varepsilon_1), \\ M_2^{(\rho)}(x_2|x_1,y) &= \ln\frac{V_2^{(\rho)}(x_2|x_1,y)}{Q_2(x_2)} - N(R_2+\varepsilon_2), \\ M_{1,2}^{(\rho)}(x|y) &= \ln\frac{V_{1,2}^{(\rho)}(x|y)}{Q(x)} - N(R_{1,2}+\varepsilon_{1,2}). \end{aligned} \tag{4.1}$$

The function

$$\begin{aligned} m(x|y) = \min\{ \quad & M_1^{(\rho)}(x_1|x_2,y), \quad M_2^{(\rho)}(x_2|x_1,y), \quad M_{1,2}^{(\rho)}(x|y), \\ & M_1^{(\rho/r)}(x_1|x_2,y)/r, \quad M_2^{(\rho/r)}(x_2|x_1,y)/r, \quad M_{1,2}^{(\rho/r)}(x|y)/r \ \} \end{aligned} \tag{4.2}$$

will be considered as the metric function of the sequential decoder. It means that if x is the code subblock assigned to some branch of the tree G, and y is the received subblock, then the metric of the branch will be calculated by (4.2).

Proposition 4.1.

1. The inequalities

$$\begin{aligned} \max_{1\leq\alpha\leq r} \mathcal{E}_1\left[\, e^{\alpha m(x|y)} \,\right] &\leq e^{-N(R_1+\varepsilon_1)}, \\ \max_{1\leq\alpha\leq r} \mathcal{E}_2\left[\, e^{\alpha m(x|y)} \,\right] &\leq e^{-N(R_2+\varepsilon_2)}, \\ \max_{1\leq\alpha\leq r} \mathcal{E}_{1,2}\left[\, e^{\alpha m(x|y)} \,\right] &\leq e^{-N(R_{1,2}+\varepsilon_{1,2})} \end{aligned} \tag{4.3}$$

are valid for any probability distributions $\{V_1^{(\rho)}(x_1|x_2,y)\}$, $\{V_2^{(\rho)}(x_2|x_1,y)\}$, and $\{V_{1,2}^{(\rho)}(x|y)\}$.

2. If

$$\begin{aligned} V_1^{(\rho)}(x_1|x_2,y) &= \frac{Q_1(x_1)\cdot P^{\frac{1}{1+\rho}}(y|x)}{\mathcal{E}_1\left[\, P^{\frac{1}{1+\rho}}(y|x)\,\right]}, \qquad (4.4)\\ V_2^{(\rho)}(x_2|x_1,y) &= \frac{Q_2(x_2)\cdot P^{\frac{1}{1+\rho}}(y|x)}{\mathcal{E}_2\left[\, P^{\frac{1}{1+\rho}}(y|x)\,\right]},\\ V_{1,2}^{(\rho)}(x|y) &= \frac{Q(x)\cdot P^{\frac{1}{1+\rho}}(y|x)}{\mathcal{E}_{1,2}\left[\, P^{\frac{1}{1+\rho}}(y|x)\,\right]}, \end{aligned}$$

then

$$\mathcal{E}_{1,2}\left[\, \sum_y P(y|x)\cdot e^{-\rho m(x|y)}\,\right] \le \Delta(\rho), \qquad (4.5)$$

where $\Delta(\rho)$ is defined in (3.8).

Proof. Since $\mathcal{E}_1\left[\, e^{\alpha m(x|y)}\,\right]$ is a convex function of α for any y and x_2, we have

$$\max_{1\le\alpha\le r} \mathcal{E}_1\left[\, e^{\alpha m(x|y)}\,\right] = \max\left\{\, \mathcal{E}_1\left[\, e^{m(x|y)}\,\right],\ \mathcal{E}_1\left[\, e^{rm(x|y)}\,\right]\,\right\}.$$

However,

$$\mathcal{E}_1\left[\, e^{m(x|y)}\,\right] \le \mathcal{E}_1\left[\, e^{M_1^{(\rho)}(x_1|x_2,y)}\,\right] = e^{-N(R_1+\varepsilon_1)}$$

and

$$\mathcal{E}_1\left[\, e^{rm(x|y)}\,\right] \le \mathcal{E}_1\left[\, e^{M_1^{(\rho/r)}(x_1|x_2,y)}\,\right] = e^{-N(R_1+\varepsilon_1)}.$$

Similar arguments are valid for $\mathcal{E}_2\left[\, e^{\alpha m(x|y)}\,\right]$ and $\mathcal{E}_{1,2}\left[\, e^{\alpha m(x|y)}\,\right]$.

To prove (4.5) we use the following inequality :

$$\begin{aligned} e^{-\rho m(x|y)} &\le e^{-\rho M_1^{(\rho)}(x|y)} + e^{-\rho M_1^{(\rho/r)}(x|y)/r} + \\ &+ e^{-\rho M_2^{(\rho)}(x|y)} + e^{-\rho M_2^{(\rho/r)}(x|y)/r} + \\ &+ e^{-\rho M_{1,2}^{(\rho)}(x|y)} + e^{-\rho M_{1,2}^{(\rho/r)}(x|y)/r}. \end{aligned}$$

Substituting this inequality into the expression at the left hand side of (4.5) and using (4.1), (4.4), and (3.8) we complete the proof.

5 Analysis of Collections of Incorrect Paths (Proof of Inequality (3.16)).

5.1 Decomposition of the set $\mathcal{U}^j_{1,r}(\mathbf{u})$.

Let us characterize an ordered list

$$U^j_r = (\ \mathbf{u}^j(1), ..., \mathbf{u}^j(r)\ )$$

by a pair of integer-valued vectors

$$(\ \mathbf{s} = (\ \mathbf{s}_1, \mathbf{s}_2\ ),\ \mathbf{t} = (\ \mathbf{t}_1, \mathbf{t}_2\ )\ )$$

such that

$$\begin{aligned} &\mathbf{s}_k = (\ s_k(1), ..., s_k(r)\ ), \\ &s_k(i) \in \{0, ..., i-1\} \ \text{ for all } \ i = 1, ..., r, \end{aligned} \tag{5.1}$$

and

$$\begin{aligned} &\mathbf{t}_k = (\ t_k(1), ..., t_k(r)\ ), \\ &0 \le t_k(1) \le t_k(2) \le ... \le t_k(r). \end{aligned} \tag{5.2}$$

Let $s_k(i) = s$ and $t_k(i) = t$ iff the following conditions are satisfied :

$$\begin{aligned} (\ u^1_k(i), ..., u^{t-1}_k(i)\ ) &= (\ u^1_k(s), ..., u^{t-1}_k(s)\ ), \\ (\ u^1_k(i), ..., u^t_k(i)\ ) &\ne (\ u^1_k(i'), ..., u^t_k(i')\ ) \ \text{ for all } \ i' = 1, ..., i-1, \end{aligned} \tag{5.3}$$

where

$$(\ u^1_k(0), ..., u^{t-1}_k(0)\ ) = (\ u^1_k, ..., u^{t-1}_k\ ), \quad k = 1, 2.$$

An example of $r = 2$ incorrect paths in G and the corresponding vectors $\mathbf{s}$ and $\mathbf{t}$ are given in Fig.2.

Let us decompose the set $\mathcal{U}^j_{1,r}(\mathbf{u})$ into subsets :

$$\mathcal{U}^j_{1,r}(\mathbf{u}) = \bigcup_{\mathbf{s},\mathbf{t}} \mathcal{U}^j_{1,r}(\ \mathbf{u}|\mathbf{s},\mathbf{t}\ ), \tag{5.4}$$

where $\mathcal{U}^j_{1,r}(\ \mathbf{u}|\mathbf{s},\mathbf{t}\ )$ consists of all lists U^j_r that are characterized by the pair $(\ \mathbf{s},\mathbf{t}\ )$.

Proposition 5.1.

1. For any vector $\mathbf{s}$, the inequality

$$\begin{aligned} & \mid \mathcal{U}^j_{1,r}(\ \mathbf{u}|\mathbf{s},\mathbf{t}\ ) \mid \\ \leq\ & \exp\{\ NR_1 \cdot \sum_{i=1}^{r}(j - t_1(i)) + NR_2 \cdot \sum_{i=1}^{r}(j - t_2(i))\ \} \end{aligned} \tag{5.5}$$

is valid.

2. If $U^j_1 \in \mathcal{U}^j_{1,r}(\ \mathbf{u}|\mathbf{s},\mathbf{t}\ )$, then

$$\sum_{i=1}^{r}(j - t_k(i)) \geq j$$

for at least one $k = 1, 2$.

3. The total number of pairs $(\mathbf{s},\mathbf{t})$, for which $\mathcal{U}^j_{1,r}(\ \mathbf{u}|\mathbf{s},\mathbf{t}\ ) \neq \emptyset$, is not greater than $(\ r! \cdot j^r\ )^2$.

Proof. If the paths $\mathbf{u}^j(1), ..., \mathbf{u}^j(i-1)$ and the components $(\ s_k(i), t_k(i)\ )$, $k = 1, 2$, are given, then the first $t_1(i)$ edges of the path $\mathbf{u}^j_1(i)$ in G_1 and the first $t_2(i)$ edges of the path $\mathbf{u}^j_2(i)$ in G_2 are known. The remaining edges of the path $\mathbf{u}^j(i) = (\ \mathbf{u}^j_1(i), \mathbf{u}^j_2(i)\ )$ in G may be chosen in not greater than $M_1^{j-t_1(i)} \cdot M_2^{j-t_2(i)}$ ways. Increasing i from 1 to r leads to the inequality

$$\mid \mathcal{U}^j_{1,r}(\ \mathbf{u}|\mathbf{s},\mathbf{t}\ ) \mid \leq \prod_{i=1}^{r} M_1^{j-t_1(i)} \cdot M_2^{j-t_2(i)},$$

which coincides with (5.5).

Statement 2 follows from the definition of the first incorrect subtree given in Section 2 : if $\mathbf{u}^j(i) \in \mathcal{U}^j_{1,r}$ then either $t_1(i) = 0$, or $t_2(i) = 0$, or both equations are valid; $i = 1, ..., r$.

The total number of vectors $\mathbf{s}_k$ satisfying (5.1) is $r!$, and the total number of vectors $\mathbf{t}_k$ satisfying (5.2) is not greater than j^r. Therefore, the total number of pairs $(\ \mathbf{s}_k, \mathbf{t}_k\ )$ satisfying these conditions is not more than $r! \cdot j^r$, and this statement is valid for $k = 1, 2$. Q.E.D.

5.2 Description of configurations of incorrect paths by bipartite graphs.

Given U_r^j and $t \in \{1, ..., j\}$, let

$$Ed_k^t = \left\{ \mathbf{u}_k^t(1), ..., \mathbf{u}_k^t(r) \right\} \setminus \{\mathbf{u}_k^t\}$$

denote the set consisting of the incorrect edges in G_k leading the paths from U_r^j to level t and let

$$r_k = \mid Ed_k^t \mid.$$

Let us fix some one-to-one mapping between the elements of the set Ed_k^t and integers $1, ..., r_k$ and map the correct edge $\mathbf{u}_k^t$ to 0. We introduce an unordered bipartite graph, but distinguish its parts as the first and the second ones. The graph contains $1 + r_1$ vertices at the first part enumerated by integers $0, 1, ..., r_1$ and $1 + r_2$ vertices at the second part enumerated by integers $0, 1, ..., r_2$. Fix that there is a branch of the graph between vertices h_1 and h_2 and that this branch has the weight $r_{h_1 h_2}^{(t)}$, if there are $r_{h_1 h_2}^{(t)}$ paths in U_r^j, whose t-th edge in G_1 is mapped to h_1 and whose t-th edge in G_2 is mapped to h_2.

If $r = 2$, then one of 11 graphs $\mathbf{G1}, ..., \mathbf{G11}$ given in Fig.3 will be obtained for each $U_r^j \in \mathcal{U}_{1,r}^j$. In this case, the configuration of incorrect paths in Fig.2 will be classified as the following sequence of bipartite graphs : $\mathbf{G1}, \mathbf{G3}, \mathbf{G10}, \mathbf{G11}$.

Let us denote the set of branches of the graph constructed for given t and U_r^j by

$$H^{(t)} = \{ (h_1, h_2) \},$$

and the set consisting of all vertices located at the k-th part by $H_k^{(t)}$. Some properties of the graph are summarized below.

Proposition 5.2.

1. $$\begin{aligned} |H^{(t)}| &= |\{ \mathbf{u}^t(1), ..., \mathbf{u}^t(r) \}|, \\ |H_k^{(t)}| &= |\{ \mathbf{u}_k^t(1), ..., \mathbf{u}_k^t(r) \}|. \end{aligned}$$

2. One of the sets, $H_1^{(t)}$ or $H_2^{(t)}$, may include the vertex 0 corresponding to the edge of the correct path $\mathbf{u}_1^t$ in G_1 or $\mathbf{u}_2^t$ in G_2, but $(0, 0) \notin H^{(t)}$.

3.

$$\sum_{(h_1,h_2)\in H^{(t)}} r^{(t)}_{h_1h_2} = r.$$

4. If $U_r^j \in \mathcal{U}^j_{1,r}(\ \mathbf{u}|\mathbf{s},\mathbf{t}\ )$, then

$$\sum_{t=1}^{j} |\ H_k^{(t)} \setminus \{0\}\ | = \sum_{i=1}^{r}(j - t_k(i)) \tag{5.6}$$

for $k = 1, 2$ and any vector $\mathbf{s}$, where $|\ H_k^{(t)} \setminus \{0\}\ |$ is the number of the vertices at the k-th part, corresponding to incorrect edges of the paths $\mathbf{u}_k^t(1), ..., \mathbf{u}_k^t(r)$ in G_k.

Proof. Statements 1-3 follow directly from the definition of a bipartite graph. To prove statement 4, we note that, if $U_r^j \in \mathcal{U}^j_{1,r}(\ \mathbf{u}|\mathbf{s},\mathbf{t}\ )$, then

$$|\ H_k^{(t)} \setminus \{0\}\ | = \sum_{i=1}^{r} \chi\{\ t_k(i) \geq t\ \},$$

as follows from (5.3). Hence,

$$\sum_{t=1}^{j} |\ H_k^{(t)} \setminus \{0\}\ | = \sum_{i=1}^{r}\sum_{t=1}^{j} \chi\{\ t_k(i) \geq t\ \}. \tag{5.7}$$

The expressions at the right hand sides of (5.6) and (5.7) coincide. Q.E.D.

5.3 Averaging over bipartite graphs.

Our intention is to calculate the average value of the metrics over the code subblocks assigned to the edges in G_1 and G_2 leading incorrect paths $\mathbf{u}^j(1), ..., \mathbf{u}^j(r)$ to level $t \in \{1, ..., j\}$. The notations above allow us to simplify formalization.

Proposition 5.3. Let $E_r[\ ]$ denote the averaging over the ensemble

$$\mathcal{A}_r = \left\{\ (X_1^N)^r (X_2^N)^r,\quad \prod_{h_1=1}^{r} Q_1(x_{1h_1}) \cdot \prod_{h_2=1}^{r} Q_2(x_{2h_2})\ \right\},$$

i.e.,

$$E_r\,[\,(*)\,] = \sum_{x_{11},...,x_{1r}} \sum_{x_{21},...,x_{2r}} \prod_{h_1=1}^{r} Q_1(x_{1h_1}) \cdot \prod_{h_2=1}^{r} Q_2(x_{2h_2}) \cdot (*),$$

where $(*)$ is an expression depending on $x_{11},...,x_{1r}$ and $x_{21},...,x_{2r}$. Then,

$$\begin{aligned} & E' \left[\prod_{i=1}^{r} \exp\left\{ \, m(\, x(\mathbf{u}^t(i)) | y\,) \, \right\} \right] \qquad (5.8) \\ = \; & E_r \left[\prod_{(h_1,h_2)\in H^{(t)}} \exp\left\{ \, r^{(t)}_{h_1h_2} \cdot m(\, x_{1h_1}, x_{2h_2} | y^t \,) \, \right\} \right]. \end{aligned}$$

Proof. The product over i at the left hand side of (5.8) may be rewritten as a product over all pairs of code subblocks assigned to edges in G which lead the paths $\mathbf{u}^j(1),...,\mathbf{u}^j(r)$ to level t. By the construction of the bipartite graph, the edge that is mapped to the branch $(h_1,h_2) \in H^{(t)}$ will be used $r^{(t)}_{h_1h_2}$ times. Statistically independent assignments of the code subblocks associated with the edges of G_k belonging to Ed_k^t are equivalent to the assignment of the independent random variables $x_{k1},...,x_{kr} \in X_k^N$ to the vertices $1,...,r$ of the bipartite graph $H_k^{(t)}$ when x_{kh} is chosen with the probability $Q_k(x_{kh})$, $h=1,...,r$; $k=1,2$. The use of the definitions of $E'\,[\,]$ and $E_r\,[\,]$ completes the proof.

In the proposition below we omit the index t and prove the inequalities for the function

$$f(\,H\,) = E_r \left[\prod_{(h_1,h_2)\in H} \exp\{\, r_{h_1h_2} \cdot m(\, x_{1h_1}, x_{2h_2} | y\,)\,\} \right]. \qquad (5.9)$$

Proposition 5.4.

1. If $h_1^* \neq 0$ and $h_2^* \neq 0$ are such that $(h_1^*, h_2^*) \in H$ is an isolated branch, i.e.,

$$\begin{aligned} (h_1^*, h_2) \notin H \quad & \text{for all } \; h_2 \neq h_2^*\,, \\ (h_1, h_2^*) \notin H \quad & \text{for all } \; h_1 \neq h_1^*\,, \end{aligned}$$

then

$$E_r\,[\, f(\,H\,)\,] \leq E_r\,[\, f(\,H \backslash \{(h_1^*, h_2^*)\}\,)\,] \cdot e^{-N(R_{1,2}+\varepsilon_{1,2})}.$$

2. If $h_1^* \neq 0$ is a vertex belonging to only one branch $(h_1^*, h_2^*) \in H$, i.e.,

$$(h_1^*, h_2) \notin H \ \text{ for all } \ h_2 \neq h_2^* \ ,$$

then

$$E_r\,[\ f(\ H\ )\ ] \leq E_r\,[\ f(\ H \backslash \{(h_1^*, h_2^*)\}\ )\ ] \cdot e^{-N(R_1+\varepsilon_1)}.$$

3. If $h_2^* \neq 0$ is a vertex belonging to only one branch $(h_1^*, h_2^*) \in H$, i.e.,

$$(h_1, h_2^*) \notin H \ \text{ for all } \ h_1 \neq h_1^* \ ,$$

then

$$E_r\,[\ f(\ H\ )\ ] \leq E_r\,[\ f(\ H \backslash \{(h_1^*, h_2^*)\}\ )\ ] \cdot e^{-N(R_2+\varepsilon_2)}.$$

4. For any vertex $h_1^* \neq 0$, the following inequality :

$$E_r\,[\ f(\ H\ )\ ] \leq E_r\,[\ f(\ H \backslash H(h_1^*)\ )] \cdot e^{-N(R_1+\varepsilon_1)}$$

is valid, where

$$H(h_1^*) = \bigcup_{h_2:\ (h_1^*, h_2) \in H} \{\ (h_1^*, h_2)\ \}\,.$$

Corollary. Let

$$H_1 = \{\ h_1 :\ (h_1, h_2) \in H \ \text{ for some } \ h_2\ \}$$

and

$$H_2 = \{\ h_2 :\ (h_1, h_2) \in H \ \text{ for some } \ h_1\ \}$$

be the sets consisting of vertices of the first and the of second parts of the graph H, respectively. Then,

$$\begin{aligned} & E_r\,[\ f(\ H\ )\ ] \\ \leq\ & \exp\{\ -N(R_1+\varepsilon_1) \cdot |H_1 \backslash \{0\}| - N(R_2+\varepsilon_2) \cdot |H_2 \backslash \{0\}|\ \}\,. \end{aligned} \tag{5.10}$$

Proof. If (h_1^*, h_2^*) is an isolated branch, then $f(\ H \backslash \{(h_1^*, h_2^*)\}\ )$ does not depend on $x_{1h_1^*}, x_{2h_2^*}$. Therefore, we may write :

$$\begin{aligned} & E_r\,[\ f(\ H\ )\ ] \\ =\ & E_r\left[\ f(\ H \backslash \{(h_1^*, h_2^*)\}\ ) \cdot \mathcal{E}_{1,2}\left[\ \exp\{\ r_{h_1^* h_2^*} \cdot m(\ x_{1h_1^*}, x_{2h_2^*} | y\ )\ \}\ \right]\right]. \end{aligned}$$

If h_1^* is a vertex belonging to only one branch $(h_1^*, h_2^*) \in H$, then $f(\ H \backslash \{(h_1^*, h_2^*)\}\ )$ does not depend on $x_{1h_1^*}$ and we may write :

$$\begin{aligned} & E_r\,[\ f(\ H\ )\ ] \\ = \ & E_r\left[\ f(\ H\backslash\{(h_1^*,h_2^*)\}\ )\cdot\mathcal{E}_1\left[\ \exp\{\ r_{h_1^*h_2^*}\cdot m(\ x_{1h_1^*},x_{2h_2^*}|y\ )\ \}\ \right]\right]. \end{aligned}$$

Similarly, if h_2^* is a vertex belonging to only one branch $(h_1^*, h_2^*) \in H$, then $f(\ H\backslash\{(h_1^*, h_2^*)\}\ )$ does not depend on $x_{2h_2^*}$ and

$$\begin{aligned} & E_r\,[\ f(\ H\ )\ ] \\ = \ & E_r\left[\ f(\ H\backslash\{(h_1^*,h_2^*)\}\ )\cdot\mathcal{E}_2\left[\ \exp\{\ r_{h_1^*h_2^*}\cdot m(\ x_{1h_1^*},x_{2h_2^*}|y\ )\ \}\ \right]\right]. \end{aligned}$$

Since $r_{h_1^*h_2^*} \geq 1$ and $r_{h_1^*h_2^*} \leq r$, using inequalities (4.3) in these equations we prove statements 1-3.

For any $h_1^* \neq 0$ we may write :

$$\begin{aligned} & E_r\,[\ f(\ H\ )\ ] \\ = \ & E_r\left[\ f(\ H\backslash H(h_1^*)\ )\cdot\mathcal{E}_1\left[\ \prod_{h_2:\,(h_1^*,h_2)\in H(h_1^*)}\exp\{\ r_{h_1^*h_2}\cdot m(\ x_{1h_1^*},x_{2h_2}|y\ )\ \}\ \right]\right]. \end{aligned}$$

Let

$$r_{h_1^*} = \sum_{h_2:\,(h_1^*,h_2)\in H(h_1^*)} r_{h_1^*h_2}$$

and

$$\lambda_{h_1^*h_2} = r_{h_1^*h_2}/r_{h_1^*}.$$

Then using Hölder's inequality we obtain :

$$\begin{aligned} & \mathcal{E}_1\left[\ \prod_{h_2:\,(h_1^*,h_2)\in H(h_1^*)}\exp\{\ r_{h_1^*h_2}\cdot m(\ x_{1h_1^*},x_{2h_2}|y\ )\ \}\ \right] \\ \leq \ & \prod_{h_2:\,(h_1^*,h_2)\in H(h_1^*)}\left(\mathcal{E}_1\left[\ \exp\{\ r_{h_1^*h_2}\cdot m(\ x_{1h_1^*},x_{2h_2}|y\ )/\lambda_{h_1^*h_2}\ \}\ \right]\right)^{\lambda_{h_1^*h_2}} \\ = \ & \prod_{h_2:\,(h_1^*,h_2)\in H(h_1^*)}\left(\mathcal{E}_1\left[\ \exp\{\ r_{h_1^*}\cdot m(\ x_{1h_1^*},x_{2h_2}|y\ )\ \}\ \right]\right)^{\lambda_{h_1^*h_2}}. \end{aligned}$$

As before, $r_{h_1^*} \geq 1$ and $r_{h_1^*} \leq r$. Thus, using the first inequality in (4.3) and the equation

$$\sum_{h_2:\,(h_1^*,h_2)\in H(h_1^*)} \lambda_{h_1^*h_2} = 1$$

we prove statement 4.

To prove the corollary we sequentially check the conditions of statements 1-3 for the graph H. If one of them is satisfied, then we set

$$H' = H \backslash \{ \ (h_1^*, h_2^*) \ \}.$$

Otherwise, we take any vertex $h_1^* \in \{1, ..., r\}$ and set

$$H' = H \backslash \{ \ H(h_1^*) \ \}.$$

In both cases $E_r \, [\ f(H) \ ]$ may be expressed through $E_r \, [\ f(H') \ ]$. Then we set $H = H'$ and repeat this procedure until $H \neq \emptyset$. As a result, we obtain (5.10). Q.E.D.

5.4 Estimation of influence of incorrect paths on the distribution of computation for any received sequence. A formal proof of inequality (3.16).

Pairs of the code subblocks assigned to the edges of G leading to different levels are statistically independent random variables. Therefore,

$$\begin{aligned}
\Psi_r^{(j)}(\ \mathbf{u}, \mathbf{y} \ ) &= \sum_{U_r^j \in \mathcal{U}_{1,r}^j(\ \mathbf{u} \ )} E' \left[\prod_{i=1}^{r} \exp \left\{ \ \Gamma(\ \mathbf{u}^j(i) | \mathbf{y}^j \ ) \ \right\} \right] \\
&= \sum_{U_r^j \in \mathcal{U}_{1,r}^j(\ \mathbf{u} \ )} E' \left[\prod_{i=1}^{r} \prod_{t=1}^{j} \exp \left\{ \ m(\ x(\mathbf{u}^t(i)) | y^t \ ) \ \right\} \right] \\
&= \sum_{U_r^j \in \mathcal{U}_{1,r}^j(\ \mathbf{u} \ )} \prod_{t=1}^{j} E' \left[\prod_{i=1}^{r} \exp \left\{ \ m(\ x(\mathbf{u}^t(i)) | y^t \ ) \ \right\} \right].
\end{aligned}$$

Using proposition 5.3 we rewrite this equation as follows :

$$\Psi_r^{(j)}(\ \mathbf{u}, \mathbf{y} \ ) = \sum_{U_r^j \in \mathcal{U}_{1,r}^j(\ \mathbf{u} \ )} \prod_{t=1}^{j} E_r \left[\prod_{(h_1,h_2) \in H^{(t)}} \exp \left\{ \ r_{h_1 h_2}^{(t)} \cdot m(\ x_{1h_1}, x_{2h_2} | y^t \ ) \ \right\} \right],$$

where $H^{(t)}$ denotes the set of branches of the bipartite graph corresponding to the t-th edges of the paths $\mathbf{u}^j(1), ..., \mathbf{u}^j(r)$.

Substitution of $H^{(t)}$ and $r^{(t)}_{h_1h_2}$ for H and $r_{h_1h_2}$ into (5.9) and the use of (5.10) leads to the following inequality :

$$\Psi^{(j)}_r(\ \mathbf{u},\mathbf{y}\ ) \le \sum_{U^j_r \in \mathcal{U}^j_{1,r}(\ \mathbf{u}\ )} \prod_{t=1}^{j}$$

$$E_r\left[\ \prod_{(h_1,h_2)\in H^{(t)}} \exp\Big\{\ -N(R_1+\varepsilon_1)\cdot |H^{(t)}_1\backslash\{0\}| - N(R_2+\varepsilon_2)\cdot |H^{(t)}_2\backslash\{0\}|\ \Big\}\ \right].$$

Let us classify lists $U^r_j \in \mathcal{U}^j_{1,r}$ in accordance with vectors $\mathbf{s},\mathbf{t}$ introduced in (5.1)-(5.3). Then using (5.6) we rewrite the last inequality as follows :

$$\Psi^{(j)}_r(\ \mathbf{u},\mathbf{y}\ ) \le \sum_{\mathbf{s},\mathbf{t}} |\mathcal{U}^j_{1,r}(\ \mathbf{u}|\mathbf{s},\mathbf{t}\ )| \cdot$$

$$\cdot\ \ \exp\{\ -N(R_1+\varepsilon_1)\cdot\sum_{i=1}^{r}(j-t_1(i)) - N(R_2+\varepsilon_2)\cdot\sum_{i=1}^{r}(j-t_2(i))\ \}.$$

The statements of proposition 5.1 and the definition of the parameter ε in (3.8) allow us to continue this inequality :

$$\begin{aligned}
&\Psi^{(j)}_r(\ \mathbf{u},\mathbf{y}\ )\\
\le\ & \sum_{\mathbf{s},\mathbf{t}} \exp\{\ -N\varepsilon_1\cdot\sum_{i=1}^{r}(j-t_1(i)) - N\varepsilon_2\cdot\sum_{i=1}^{r}(j-t_2(i))\ \}\\
\le\ & \sum_{\mathbf{s},\mathbf{t}} \exp\{\ -N\varepsilon\cdot\sum_{k=1}^{2}\sum_{i=1}^{r}(j-t_k(i))\ \}\\
\le\ & \sum_{\mathbf{s},\mathbf{t}} \exp\{\ -jN\varepsilon\ \} \le (r!\cdot j^r)^2\cdot\exp\{\ -jN\varepsilon\ \},
\end{aligned}$$

and inequality (3.16) has been proved.

6 Acknowledgement.

The author wishes to thank to Prof. Rolf Johannesson for his advices and help in preparing the manuscript. The author is also grateful to anonymous reviewers for valuable comments.

Appendix 1.

Proof of Lemma 1.

Substituting (3.1) into (3.11) we write :

$$\overline{C_1^{(\rho)}}(\mathbf{u}) \leq \sum_{\mathbf{y}} E\left[P(\ \mathbf{y}|\mathbf{x}(\mathbf{u})\ ) \cdot E'\left[\left(\sum_{l\geq 0}\sum_{j\geq 1} \gamma^{(lj)}(\ \mathbf{u}|\mathbf{y}\ ) \right)^{\rho} \right] \right].$$

If $\rho \leq 1$, then we use Jensen's inequality with $\lambda = \rho$. If $\rho > 1$, then we use Minkowski's inequality. Thus,

$$\overline{C_1^{(\rho)}}(\mathbf{u}) \leq \sum_{l\geq 0}\sum_{j\geq 1} \hat{F}_{\rho}^{(lj)}(\mathbf{u}), \text{ if } \rho \leq 1$$

and

$$\overline{C_1^{(\rho)}}(\mathbf{u}) \leq \left(\sum_{l\geq 0}\sum_{j\geq 1} (\hat{F}_{\rho}^{(lj)}(\mathbf{u}))^{1/\rho} \right)^{\rho}, \text{ if } \rho > 1,$$

where

$$\hat{F}_{\rho}^{(lj)}(\mathbf{u}) = \sum_{\mathbf{y}} E\left[\ P(\ \mathbf{y}|\mathbf{x}(\mathbf{u})\ ) \cdot E'\left[\left(\gamma^{(lj)}(\ \mathbf{u}|\mathbf{y}\ ) \right)^{\rho}\ \right] \right].$$

However,

$$\begin{aligned} E'\left[\left(\gamma^{(lj)}(\ \mathbf{u}|\mathbf{y}\ ) \right)^{\rho}\ \right] \ &\leq\ \exp\{\ -\rho\Gamma(\ \mathbf{u}^l|\mathbf{y}^l\ )\ \} \\ &\cdot\ E'\left[\left(\sum_{\hat{\mathbf{u}}\in\mathcal{U}_1^j(\mathbf{u})} \exp\{\ \Gamma(\ \mathbf{u}^j|\mathbf{y}^j\ )\ \} \right)^{\rho}\ \right], \end{aligned}$$

and we obtain the final result by applying Jensen's inequality with the parameter $\lambda = \rho/r$.

Proof of Inequality (3.15).

Since the code subblocks assigned to different edges of the correct path are statistically independent random variables, we may write :

$$
\begin{aligned}
& \sum_{\mathbf{y}} E\left[\, \Phi_{\rho}^{(l)}(\,\mathbf{u},\mathbf{y}\,)\,\right] \\
= \;& \sum_{\mathbf{y}} E\left[\prod_{t=1}^{l} P(\,y_t|x(\mathbf{u}_t)\,)\cdot \exp\{\,-\rho m(\,x(\mathbf{u}_t)|y_t\,)\,\}\right] \\
= \;& \prod_{t=1}^{l} \sum_{y_t \in Y^N} E\,[\,P(\,y_t|x(\mathbf{u}_t))\,)\cdot \exp\{\,-\rho m(\,x(\mathbf{u}_t)|y_t\,)\,\}\,] \\
= \;& \left(\mathcal{E}_{1,2}\left[\sum_{y\in Y^N} P(\,y|x\,)\cdot e^{-\rho m(\,x|y\,)}\right]\right)^{l} \leq \Delta^{l}(\rho),
\end{aligned}
$$

where the last inequality follows from (4.5). Q.E.D.

Appendix 2.

Averaging over bipartite graphs that describe all configurations of $r = 2$ incorrect paths in G given in Fig.2.

G1.

$$\begin{aligned} & E'\left[e^{2m(x_{10},x_{21}|y)}\right] \le E'\left[e^{M_2^{(\rho/2)}(x_{21}|x_{10},y)}\right] = \mathcal{E}_2\left[e^{M_2^{(\rho/2)}(x_{21}|x_{10},y)}\right] \\ = \ & e^{-N(R_2+\varepsilon_2)}. \end{aligned}$$

G3.

$$\begin{aligned} & E'\left[e^{2m(x_{11},x_{21}|y)}\right] \le E'\left[e^{M_{1,2}^{(\rho/2)}(x_{11},x_{21}|y)}\right] = \mathcal{E}_{1,2}\left[e^{M_{1,2}^{(\rho/2)}(x_{11},x_{21}|y)}\right] \\ = \ & e^{-N(R_{1,2}+\varepsilon_{1,2})}. \end{aligned}$$

G10.

$$\begin{aligned} & E'\left[e^{m(x_{11},x_{21}|y)} \cdot e^{m(x_{12},x_{21}|y)}\right] \le E'\left[e^{M_{1,2}^{(\rho)}(x_{11},x_{21}|y)} \cdot e^{M_1^{(\rho)}(x_{12}|x_{21},y)}\right] \\ = \ & \mathcal{E}_{1,2}\left[e^{M_{1,2}^{(\rho)}(x_{11},x_{21}|y)} \cdot \mathcal{E}_1\left[e^{M_1^{(\rho)}(x_{12}|x_{21},y)}\right]\right] = e^{-N(R_{1,2}+\varepsilon_{1,2}+R_1+\varepsilon_1)}. \end{aligned}$$

G11.

$$\begin{aligned} & E'\left[e^{m(x_{11},x_{21}|y)} \cdot e^{m(x_{12},x_{22}|y)}\right] \le E'\left[e^{M_{1,2}^{(\rho)}(x_{11},x_{21}|y)} \cdot e^{M_{1,2}^{(\rho)}(x_{12},x_{22}|y)}\right] \\ = \ & \mathcal{E}_{1,2}\left[e^{M_{1,2}^{(\rho)}(x_{11},x_{21}|y)} \cdot \mathcal{E}_{1,2}\left[e^{M_{1,2}^{(\rho)}(x_{12},x_{22}|y)}\right]\right] = e^{-N(R_{1,2}+\varepsilon_{1,2}+R_{1,2}+\varepsilon_{1,2})}. \end{aligned}$$

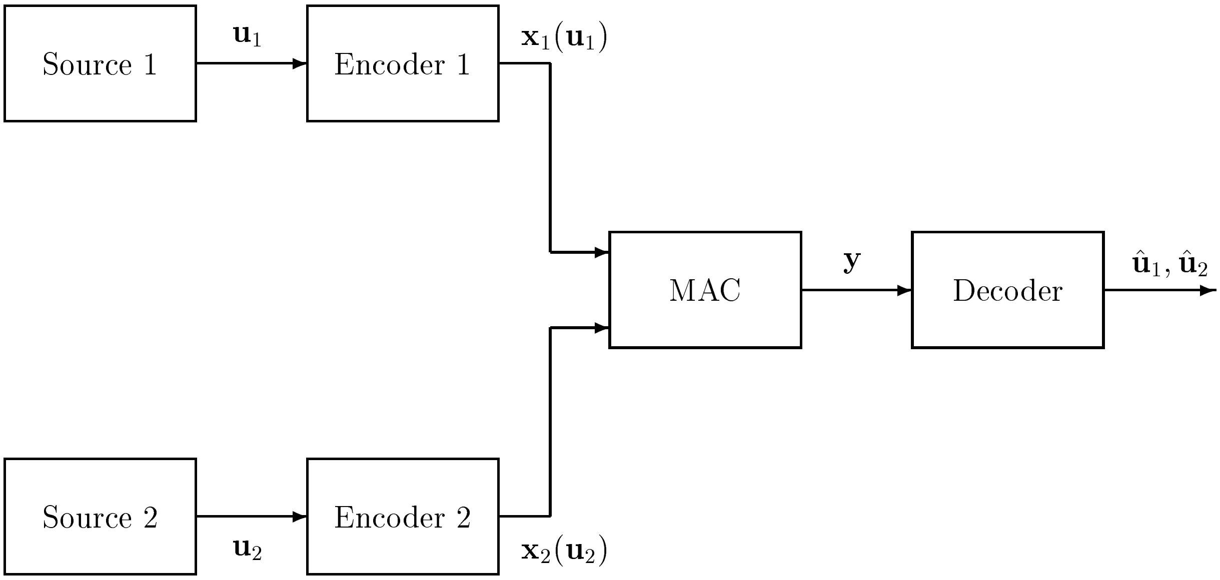

Figure 1: Model of information transmission over a multiple-access channel.

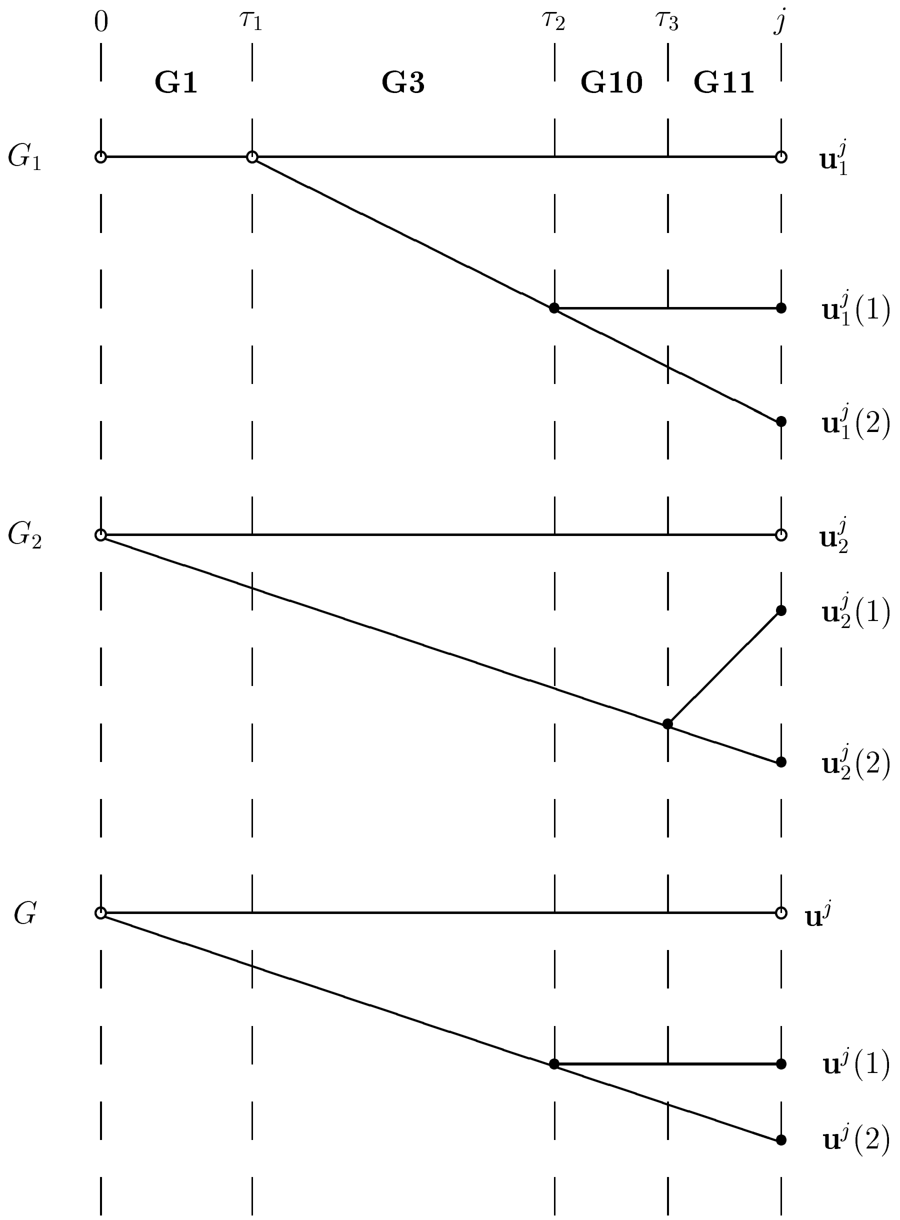

Figure 2: A configuration of $r = 2$ incorrect paths, which are characterized by the vectors $\mathbf{s}_1 = \mathbf{s}_2 = (0, 1)$ and $\mathbf{t}_1 = (\tau_1, \tau_2)$, $\mathbf{t}_2 = (0, \tau_3)$.

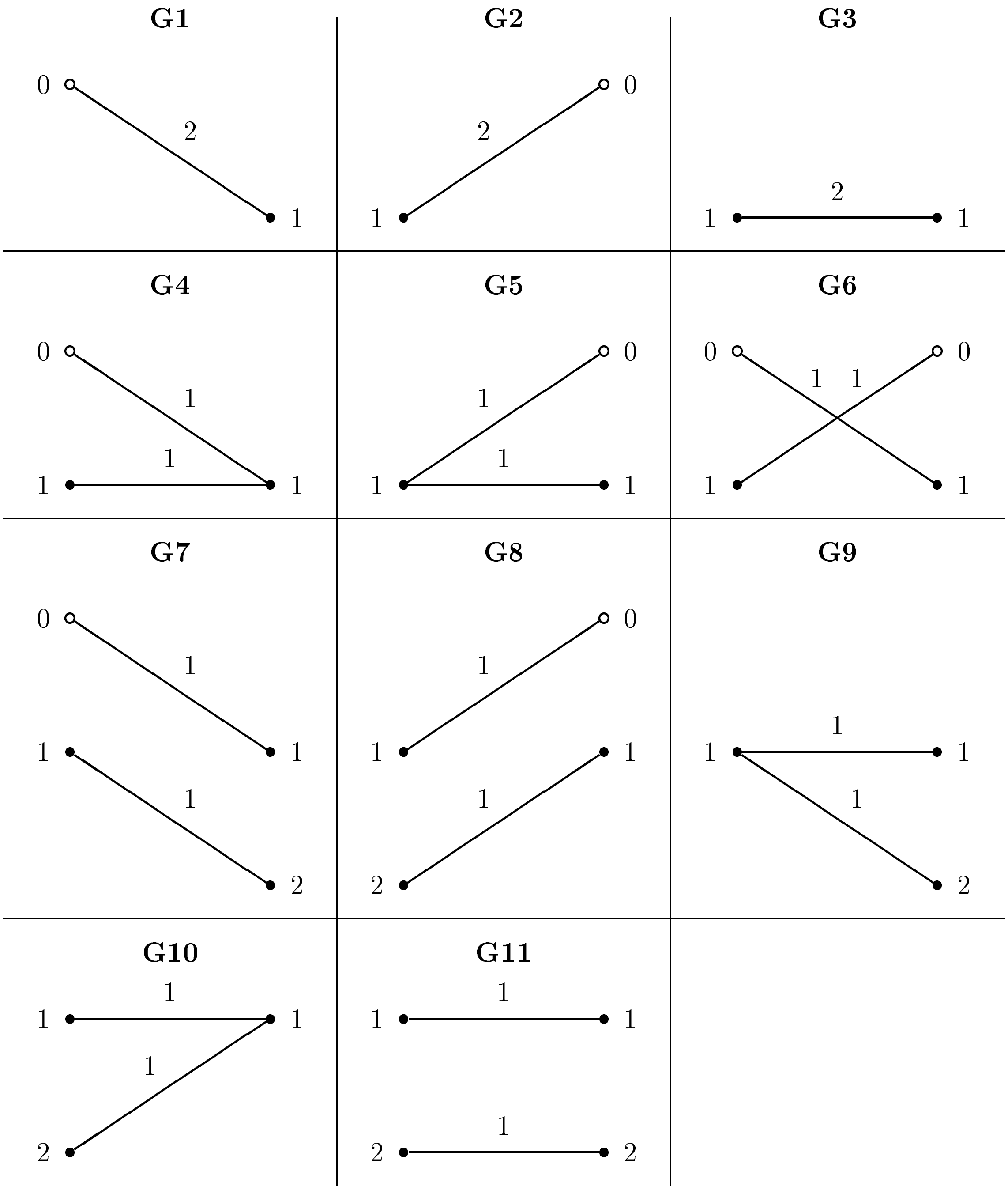

Figure 3: Bipartite graphs that describe all configurations of $r = 2$ incorrect paths in G.

References

[1] I.M.Jacobs and E.R.Berlekamp, "A lower bound to the distribution of computation for sequential decoding," *IEEE Trans. Inform. Theory,* Vol. IT-13, pp.167-174, Apr. 1967.

[2] J.Savage, "The distribution of sequential decoding computational time," *IEEE Trans. Inform. Theory,* Vol. IT-12, pp.145-147, Apr. 1966.

[3] F.Jelinek, "An upper bound on moments of sequential decoding efforts," *IEEE Trans. Inform. Theory,* Vol. IT-15, pp.140-149, Jan. 1969.

[4] G.D.Forney, "Convolutional codes III : Sequential decoding," *Inform. Control,* Vol. 25, pp.267-297, Apr. 1974.

[5] T.Hashimoto and S.Arimoto, "Computational moments for sequential decoding of convolutional codes," *IEEE Trans. Inform. Theory,* Vol. IT-25, pp.584-591, Sept. 1979.

[6] V.B.Balakirsky, "On sequential decoding of linear convolutional codes," *Problemy Peredachi Inform.* Vol. 23, pp.41-48, Oct.-Dec. 1987.

[7] D.Slepian and J.K.Wolf, "A coding theorem for multiple access channels with correlated sources," *Bell Syst. Techn. J.,* Vol. 52, pp.1037-1076, Sept. 1973.

[8] R.L.Peterson and D.J.Costello,Jr., "Error probability and free distance bounds for two-user tree codes on multiple-access channels," *IEEE Trans. Inform. Theory,* Vol. IT-26, pp.658-670, Nov. 1980.

[9] G.Sh.Poltyrev, "Coding for asynchronous multiple access channel," *Problemy Peredachi Inform.* Vol. 19, pp.12-21, July-Sept. 1983.

[10] E.Arikan, "Sequential decoding for multiple access channels," *IEEE Trans. Inform. Theory,* Vol. IT-34, pp.246-259, March 1988.

[11] V.B.Balakirsky, "Sequential decoding algorithm for multiple access channel," *Problemy Peredachi Inform.* Vol. 21, pp.3-12, July-Sept. 1985.

[12] V.B.Balakirsky, "An upper bound on the distribution of computation of a sequential decoder for multiple access channels," *Proceedings of the Sixth Swedish-Russian Int. Workshop on Information Theory,* pp.179-183, Mölle, Sweden, Aug.22-27, 1993.

[13] T.Hashimoto, "A list-type reduced-constraint generalization of the Viterbi algorithm," *IEEE Trans. Inform. Theory,* Vol. IT-33, pp.866-876, Nov. 1987.

[14] F.Jelinek, "A fast sequential decoding algorithm using a stack," *IBM J. Res. Dev.,* Vol. 13, pp.675-685, 1969.

[15] R.G.Gallager, *Information Theory and Reliable Communication.* New York : Willey, 1968.

[16] K.Sh.Zigangirov, "Some sequential decoding procedures," *Problemy Peredachi Inform.* Vol. 2, pp.13-25, Apr.-June 1966.