論文 / 著書情報
Article / Book Information

Title	Random-Coding Exponential Error Bounds for Channels with Action-Dependent States
Authors	Tetsunao Matsuta, Tomohiko Uyematsu
Citation	IEICE Trans. Fundamentals, vol. E96-A, no. 12, pp. 2324-2331
Pub. date	2013, 12
URL	http://search.ieice.org/
Copyright	(c) 2013 Institute of Electronics, Information and Communication Engineers

PAPER *Special Section on Information Theory and Its Applications*

Random-Coding Exponential Error Bounds for Channels with Action-Dependent States

Tetsunao MATSUTA[†a)], ***Member and*** **Tomohiko UYEMATSU**[†b)], ***Fellow***

SUMMARY Weissman introduced a coding problem for channels with action-dependent states. In this coding problem, there are two encoders and a decoder. An encoder outputs an action that affects the state of the channel. Then, the other encoder outputs a codeword of the message into the channel by using the channel state. The decoder receives a noisy observation of the codeword, and reconstructs the message. In this paper, we show an exponential error bound for channels with action-dependent states based on the random coding argument.
key words: *actions, channels with states, error exponents, exponential error bounds*

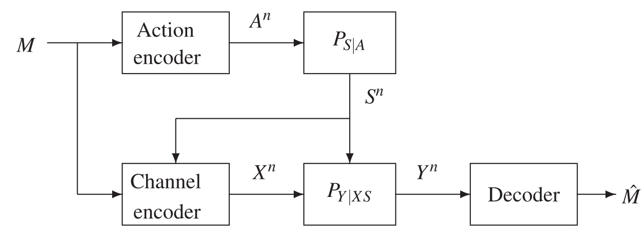

Fig. 1 Channels with action dependent states.

1. Introduction

In many practical situations, states of a communication channel change from moment to moment. To analyze the performance of coding systems under such a situation, many researchers have been studied coding problems for channels with states (cf. Chapter 7 in [1]). In these coding problems, researchers considered the situation where states of the channel change by nature. Thus, the coding system can neither control nor affect states of the channel.

On the other hand, Weissman [2] introduced a coding problem for *channels with action-dependent states*, in which an *action* of the coding system affects states of the channel.

More precisely, he considered the following coding problem (see Fig. 1): To send a message to the receiver via a channel with states, the coding system uses an action encoder and a channel encoder. The action encoder outputs an action corresponding to the message, and the action affects states of the channel. Then, the channel encoder receives the state of the channel, and outputs a codeword of the message into the channel. The receiver receives a noisy observation of the codeword, and reconstructs the message by using a decoder. For this coding problem, Weissman [2] showed the channel capacity, where the channel capacity is the supremum of rates of the code such that the decoding error probability vanishes as the block length tends to infinity. In [2], he also studied various applications of channels with action-dependent states, and showed the channel capacity for these applications. Since he introduced the coding problem for channels with action-dependent states, a lot of extensions of this coding problem have been reported [3]–[6].

For many coding systems, the decoding error probability is one of the most important performance measures of the system. It is often evaluated by exponential form, and for coding systems of channels with states, exponential error bounds have been studied by many researchers [7]–[11]. However, exponential error bounds for channels with action-dependent states have not been fully studied. In this paper, we focus on exponential error bounds for channels with action-dependent states, and show an exponential error bound. Our result is based on the proof method in [11], which was developed to clarify an exponential error bound for channels with states that cannot be affected by coding system. In [11], Moulin and Wang used basic information-theoretic techniques such as method of types [12] and the random coding argument except the decoding procedure. In the decoding procedure, they employed a decoder using the *penalized* empirical mutual information, where the "penalized" means the existence of a term subtracting from the empirical mutual information. This decoder may be regarded as an empirical version of the MAP decoder (see [11, p.1332]). In this paper, we also use this kind of decoder in order to clarify exponential error bounds.

2. Preliminaries

In this section, we provide a precise formulation of the coding problem for channels with action-dependent states, and show the channel capacity.

We will denote an n-length sequence of symbols $(a_1, a_2, \cdots, a_n)$ by the boldface letter $\mathbf{a}$. We will denote random variables (RVs) by capital letters $X, Y, Z, \cdots$, the values they can take by lowercase letters $x, y, z, \cdots$, and the set of these values by calligraphic letters $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \cdots$. The probability distribution of a RV X will be denoted by P_X. The conditional distribution for X given Y will be denoted by $P_{X|Y}$. We will denote the set of all probability distri-

Manuscript received February 7, 2013.
Manuscript revised June 24, 2013.
[†]The authors are with the Dept. of Communications and Computer Engineering, Tokyo Institute of Technology, Tokyo, 152-8550 Japan.
a) E-mail: tetsu@it.ce.titech.ac.jp
b) E-mail: uyematsu@ieee.org
DOI: 10.1587/transfun.E96.A.2324

butions over $\mathscr{X}$ by $\mathscr{P}(\mathscr{X})$, and the set of all conditional distributions from $\mathscr{Y}$ to $\mathscr{X}$ by $\mathscr{W}(\mathscr{X}|\mathscr{Y})$. We will denote the nth power of a probability distribution $P_X \in \mathscr{P}(\mathscr{X})$ by P_X^n, i.e., $P_X^n(\mathbf{x}) = \prod_{i=1}^n P_X(x_i)$, and the nth power of a conditional distribution $P_{X|Y} \in \mathscr{W}(\mathscr{X}|\mathscr{Y})$ by $P_{X|Y}^n$, i.e., $P_{X|Y}^n(\mathbf{x}|\mathbf{y}) = \prod_{i=1}^n P_{X|Y}(x_i|y_i)$.

Let $\mathscr{X}$, $\mathscr{Y}$, $\mathscr{A}$ and $\mathscr{S}$ be finite sets. A channel with action-dependent states is characterized by two conditional distributions $P_{Y|XS} \in \mathscr{W}(\mathscr{Y}|\mathscr{X} \times \mathscr{S})$ and $P_{S|A} \in \mathscr{W}(\mathscr{S}|\mathscr{A})$. For a message $m \in \mathscr{M}_n \triangleq \{1, \cdots, M_n\}$, the action encoder f_{A}^n outputs an action $\mathbf{a} \in \mathscr{A}^n$ corresponding to the message m. Thus, the action encoder is defined as

$$f_{\mathrm{A}}^n : \mathscr{M}_n \to \mathscr{A}^n.$$

The action $\mathbf{a}$ affects a state $\mathbf{s} \in \mathscr{S}^n$ according to $P_{S|A}^n(\mathbf{s}|\mathbf{a})$. On the other hand, the channel encoder f_{C}^n receives the state $\mathbf{s}$ of the channel $P_{Y|XS}$, and outputs a codeword $\mathbf{x} \in \mathscr{X}^n$ corresponding to the message m. Thus, the channel encoder is defined as

$$f_{\mathrm{C}}^n : \mathscr{M}_n \times \mathscr{S}^n \to \mathscr{X}^n.$$

The channel encoder sends the codeword to a decoder via the channel $P_{Y|XS}$. Then, the decoder φ_n receives a channel output $\mathbf{y} \in \mathscr{Y}^n$ that is drawn from $P_{Y|XS}^n(\mathbf{y}|\mathbf{x},\mathbf{s})$, and reconstructs the message m. Thus, the decoder is defined as

$$\varphi_n : \mathscr{Y}^n \to \mathscr{M}_n.$$

The rate R_n of a code $(f_{\mathrm{A}}^n, f_{\mathrm{C}}^n, \varphi_n)$ is defined as

$$R_n \triangleq \frac{1}{n} \log M_n,$$

and the error probability of the code is defined as

$$\varepsilon_n(f_{\mathrm{A}}^n, f_{\mathrm{C}}^n, \varphi_n) \triangleq \frac{1}{M_n} \sum_{m \in \mathscr{M}_n} \Pr\{\varphi_n(Y^n) \neq m | m \text{ is sent}\}.$$

We say that $R \geq 0$ is achievable if there exists a sequence of codes $\{(f_{\mathrm{A}}^n, f_{\mathrm{C}}^n, \varphi_n)\}$ such that

$$\liminf_{n\to\infty} R_n \geq R$$

and

$$\lim_{n\to\infty} \varepsilon_n(f_{\mathrm{A}}^n, f_{\mathrm{C}}^n, \varphi_n) = 0.$$

Then, the channel capacity C of the channel with action-dependent states is defined as

$$C = \sup\{R : R \text{ is achievable}\}.$$

For the channel with action-dependent states, Weissman [2] showed the next theorem.

Theorem 1 ([2, Theorem 1]).

$$C = \max_{P_{ASUXY}} [I(U;Y) - I(U;S|A)]$$
$$= \max_{P_{ASUXY}} [I(A,U;Y) - I(U;S|A)],$$

where $I(U;Y)$ and $I(A,U;Y)$ are the mutual information between U and Y, and between (A,U) and Y, respectively, $I(U;S|A)$ is the conditional mutual information between U and S given A, $P_{ASUXY} \in \mathscr{P}(\mathscr{A} \times \mathscr{S} \times \mathscr{U} \times \mathscr{X} \times \mathscr{Y})$ denotes the joint distribution such that

$$P_{ASUXY}(a,s,u,x,y) = P_A(a)P_{S|A}(s|a)P_{U|SA}(u|s,a) \times 1_{\{f(u,s)\}}(x)P_{Y|XS}(y|x,s)$$

for some $P_A \in \mathscr{P}(\mathscr{A})$, $P_{U|SA} \in \mathscr{W}(\mathscr{U}|\mathscr{S} \times \mathscr{A})$, $f : \mathscr{U} \times \mathscr{S} \to \mathscr{X}$, $|\mathscr{U}| \leq |\mathscr{A}||\mathscr{S}||\mathscr{X}| + 1$, and $1_{\mathscr{X}}(x)$ is the indicator function defined as

$$1_{\mathscr{X}}(x) \triangleq \begin{cases} 1 & \text{if } x \in \mathscr{X}, \\ 0 & \text{if } x \notin \mathscr{X}. \end{cases}$$

3. Exponential Error Bounds

In this section, we show an exponential error bound for channels with action-dependent states.

Before we show our main result, we need several definitions. For any sequence $\mathbf{x} \in \mathscr{X}^n$, we denote the number of occurrences of $x \in \mathscr{X}$ in $\mathbf{x}$ by $N(x|\mathbf{x})$. Then, we define the *type* $P_{\mathbf{x}}$ of a sequence $\mathbf{x}$ as an empirical distribution given by

$$P_{\mathbf{x}}(x) \triangleq \frac{1}{n} N(x|\mathbf{x}) \quad {}^{\forall} x \in \mathscr{X}.$$

Similarly, for any pair of sequences $(\mathbf{x},\mathbf{y}) \in \mathscr{X}^n \times \mathscr{Y}^n$, we denote the number of joint occurrences of $(x,y) \in \mathscr{X} \times \mathscr{Y}$ in $(\mathbf{x},\mathbf{y})$ as $N(x,y|\mathbf{x},\mathbf{y})$. Then, we define the joint type $P_{\mathbf{xy}}$ of a sequence $(\mathbf{x},\mathbf{y})$ as

$$P_{\mathbf{xy}}(x,y) \triangleq \frac{1}{n} N(x,y|\mathbf{x},\mathbf{y}) \quad {}^{\forall}(x,y) \in \mathscr{X} \times \mathscr{Y},$$

and define the conditional type $P_{\mathbf{y}|\mathbf{x}}$ of a sequence $(\mathbf{x},\mathbf{y})$ as

$$P_{\mathbf{y}|\mathbf{x}}(y|x) \triangleq \frac{N(x,y|\mathbf{x},\mathbf{y})}{N(x|\mathbf{x})}, \quad {}^{\forall}(x,y) \in \mathscr{X} \times \mathscr{Y} \text{ such that } N(x|\mathbf{x}) \neq 0.$$

Joint types and conditional types of more than two sequences can be defined similarly.

Let $\mathscr{P}_n(\mathscr{X})$ denotes the set of possible types in $\mathscr{X}^n$. For any $P_{\bar{X}} \in \mathscr{P}_n(\mathscr{X})$, we define $T(P_{\bar{X}})$ as

$$T(P_{\bar{X}}) \triangleq \{\mathbf{x} \in \mathscr{X}^n : P_{\mathbf{x}} = P_{\bar{X}}\},$$

and for any sequences $\mathbf{x} \in \mathscr{X}^n$ and any conditional probability distribution $P_{\bar{Y}|\bar{X}} \in \mathscr{W}(\mathscr{Y}|\mathscr{X})$, we define $T(P_{\bar{Y}|\bar{X}}|\mathbf{x})$ as

$$T(P_{\bar{Y}|\bar{X}}|\mathbf{x}) \triangleq \{\mathbf{y} \in \mathscr{Y}^n : P_{\mathbf{xy}} = P_{\mathbf{x}} \cdot P_{\bar{Y}|\bar{X}}\},$$

where $P_{\mathbf{x}} \cdot P_{\bar{Y}|\bar{X}}$ represents the product of $P_{\mathbf{x}}$ and $P_{\bar{Y}|\bar{X}}$, i.e.,

$$P_{\mathbf{x}} \cdot P_{\bar{Y}|\bar{X}}(x,y) = P_{\mathbf{x}}(x) P_{\bar{Y}|\bar{X}}(y|x), {}^{\forall}(x,y) \in \mathscr{X} \times \mathscr{Y}.$$

Furthermore, for any type $P_{\bar{X}} \in P_n(\mathscr{X})$, let $\mathscr{W}_n(\mathscr{Y}|P_{\bar{X}})$ denotes the set of $P_{\bar{Y}|\bar{X}} \in \mathscr{W}(\mathscr{Y}|\mathscr{X})$ for which $T(P_{\bar{Y}|\bar{X}}|\mathbf{x})$ is not empty for sequences $\mathbf{x} \in T(P_{\bar{X}})$.

For types of sequences, the next lemma is well known (see e.g. [12]).

Lemma 1.

1). $|\mathscr{P}_n(\mathscr{X})| \le (n+1)^{|\mathscr{X}|}$, and for any type $P_{\bar{X}} \in \mathscr{P}_n(\mathscr{X})$, $|\mathscr{W}_n(\mathscr{Y}|P_{\bar{X}})| \le (n+1)^{|\mathscr{X}||\mathscr{Y}|}$.

2). For any $P_{\bar{X}} \in \mathscr{P}_n(\mathscr{X})$,

$$(n+1)^{-|\mathscr{X}|}\exp\{nH(\bar{X})\} \le |T(P_{\bar{X}})| \le \exp\{nH(\bar{X})\},$$

and for any $\mathbf{x} \in T(P_{\bar{X}})$ and $P_{\bar{Y}|\bar{X}} \in \mathscr{W}_n(\mathscr{Y}|P_{\bar{X}})$,

$$\begin{aligned}(n+1)^{-|\mathscr{X}||\mathscr{Y}|}\exp\{nH(\bar{Y}|\bar{X})\} &\le |T(P_{\bar{Y}|\bar{X}}|\mathbf{x})| \\ &\le \exp\{nH(\bar{Y}|\bar{X})\},\end{aligned}$$

where $H(\bar{X})$ is the entropy of $\bar{X}$, and $H(\bar{Y}|\bar{X})$ is the conditional entropy of $\bar{Y}$ given $\bar{X}$.

3). For any $P_X \in \mathscr{P}(\mathscr{X})$, $P_{\bar{X}} \in \mathscr{P}_n(\mathscr{X})$, and $\mathbf{x} \in T(P_{\bar{X}})$,

$$P_X^n(\mathbf{x}) = \exp\{-n(D(P_{\bar{X}}\|P_X) + H(\bar{X}))\},$$

and for any $P_{Y|X} \in \mathscr{W}(\mathscr{Y}|\mathscr{X})$, $P_{\bar{Y}|\bar{X}} \in \mathscr{W}_n(\mathscr{Y}|P_{\bar{X}})$, and $\mathbf{y} \in T(P_{\bar{Y}|\bar{X}}|\mathbf{x})$,

$$P_{Y|X}^n(\mathbf{y}|\mathbf{x}) = \exp\{-n(D(P_{\bar{Y}|\bar{X}}\|P_{Y|X}|P_{\bar{X}}) + H(\bar{Y}|\bar{X}))\},$$

where $D(\cdot\|\cdot)$ denotes the relative entropy, and $D(\cdot\|\cdot|\cdot)$ denotes the conditional relative entropy.

The next theorem shows an exponential error bound for channels with action-dependent states.

Theorem 2. For any channels $P_{S|A} \in \mathscr{W}(\mathscr{S}|\mathscr{A})$ and $P_{Y|XS} \in \mathscr{W}(\mathscr{Y}|\mathscr{X}\times\mathscr{S})$, any finite set $\mathscr{U}$, any $M_n > 0$, $\varepsilon > 0$, and sufficiently large n, there exists a code $(f_{\mathrm{A}}^n, f_{\mathrm{C}}^n, \varphi_n)$ satisfying

$$\begin{aligned}&\varepsilon_n(f_{\mathrm{A}}^n, f_{\mathrm{C}}^n, \varphi_n) \\ &\quad\le \exp\{-n \max_{P_{\bar{A}}\in\mathscr{P}_n(\mathscr{A})} \min_{P_{\bar{S}|\bar{A}}\in\mathscr{W}_n(\mathscr{S}|P_{\bar{A}})} \max_{P_{\bar{U}|\bar{S}\bar{A}}\in\mathscr{W}_n(\mathscr{U}|P_{\bar{A}}\cdot P_{\bar{S}|\bar{A}})} \\ &\qquad \max_{\bar{f}:\mathscr{U}\times\mathscr{S}\to\mathscr{X}} \min_{P_{\bar{Y}|\bar{X}\bar{S}}\in\mathscr{W}_n(\mathscr{Y}|P_{\bar{S}\bar{X}})} [D(P_{\bar{S}|\bar{A}}\|P_{S|A}|P_{\bar{A}}) \\ &\qquad + D(P_{\bar{Y}|\bar{X}\bar{S}}\|P_{Y|XS}|P_{\bar{S}\bar{X}}) \\ &\qquad + |I(\bar{A},\bar{U};\bar{Y}) - I(\bar{U};\bar{S}|\bar{A}) - R_n - \varepsilon|^+]\},\end{aligned}$$

where $|x|^+ = \max\{0, x\}$,

$$P_{\bar{S}\bar{X}}(s,x) = \sum_{(a,u)\in\mathscr{A}\times\mathscr{U}} P_{\bar{A}}(a)P_{\bar{S}|\bar{A}}(s|a)P_{\bar{U}|\bar{S}\bar{A}}(u|s,a)1_{\{\bar{f}(u,s)\}}(x),$$

and $(\bar{A},\bar{S},\bar{U},\bar{Y})$ is the quadruple of RVs drawn according to the probability distribution $P_{\bar{A}\bar{S}\bar{U}\bar{Y}}$ such that

$$\begin{aligned}P_{\bar{A}\bar{S}\bar{U}\bar{Y}}(a,s,u,y) = \sum_{x\in\mathscr{X}} &P_{\bar{A}}(a)P_{\bar{S}|\bar{A}}(s|a)P_{\bar{U}|\bar{S}\bar{A}}(u|s,a) \\ &\times 1_{\{\bar{f}(u,s)\}}(x)P_{\bar{Y}|\bar{X}\bar{S}}(y|x,s).\end{aligned}$$

By using this theorem and continuity of the mutual information and the relative entropy, we obtain the next corollary.

Corollary 1. For any channels $P_{S|A} \in \mathscr{W}(\mathscr{S}|\mathscr{A})$ and $P_{Y|XS} \in \mathscr{W}(\mathscr{Y}|\mathscr{X}\times\mathscr{S})$, any finite set $\mathscr{U}$, and any $R > 0$, there exists a sequence of codes $\{(f_{\mathrm{A}}^n, f_{\mathrm{C}}^n, \varphi_n)\}$ such that

$$\liminf_{n\to\infty} R_n \ge R,$$

and

$$\liminf_{n\to\infty} -\frac{1}{n}\log\varepsilon_n(f_{\mathrm{A}}^n, f_{\mathrm{C}}^n, \varphi_n) \ge E_{\mathrm{r}}(R),$$

where

$$\begin{aligned}&E_{\mathrm{r}}(R) \\ &\triangleq \max_{P_{\bar{A}}\in\mathscr{P}(\mathscr{A})} \min_{P_{\bar{S}|\bar{A}}\in\mathscr{W}(\mathscr{S}|\mathscr{A})} \max_{P_{\bar{U}|\bar{S}\bar{A}}\in\mathscr{W}(\mathscr{U}|\mathscr{S}\times\mathscr{A})} \max_{\bar{f}:\mathscr{U}\times\mathscr{S}\to\mathscr{X}} \\ &\quad \min_{P_{\bar{Y}|\bar{X}\bar{S}}\in\mathscr{W}(\mathscr{Y}|\mathscr{X}\times\mathscr{S})} [D(P_{\bar{S}|\bar{A}}\|P_{S|A}|P_{\bar{A}}) + D(P_{\bar{Y}|\bar{X}\bar{S}}\|P_{Y|XS}|P_{\bar{S}\bar{X}}) \\ &\quad + |I(\bar{A},\bar{U};\bar{Y}) - I(\bar{U};\bar{S}|\bar{A}) - R|^+].\end{aligned}$$

Remark 1. Let $|\mathscr{U}| \ge |\mathscr{A}||\mathscr{S}||\mathscr{X}| + 1$. Then, according to the corollary, we have

$$\begin{aligned}0 \le& E_{\mathrm{r}}(R) \\ \le& \max_{P_{\bar{A}}\in\mathscr{P}(\mathscr{A})} \max_{P_{\bar{U}|\bar{S}\bar{A}}\in\mathscr{W}(\mathscr{U}|\mathscr{S}\times\mathscr{A})} \max_{\bar{f}:\mathscr{U}\times\mathscr{S}\to\mathscr{X}} \\ &[D(P_{S|A}\|P_{S|A}|P_{\bar{A}}) + D(P_{Y|XS}\|P_{Y|XS}|P_{\bar{S}\bar{X}}) \\ &+ |I(\bar{A},\bar{U};\bar{Y}) - I(\bar{U};\bar{S}|\bar{A}) - R|^+] \\ =& \max_{P_{\bar{A}}\in\mathscr{P}(\mathscr{A})} \max_{P_{\bar{U}|\bar{S}\bar{A}}\in\mathscr{W}(\mathscr{U}|\mathscr{S}\times\mathscr{A})} \max_{\bar{f}:\mathscr{U}\times\mathscr{S}\to\mathscr{X}} \\ &\times |I(\bar{A},\bar{U};\bar{Y}) - I(\bar{U};\bar{S}|\bar{A}) - R|^+ \\ =& |C - R|^+.\end{aligned}$$

Thus, if $R \ge C$ then $E_{\mathrm{r}}(R) = 0$. On the other hand, due to the positivity of the relative entropy, $E_{\mathrm{r}}(R) = 0$ implies

$$\begin{aligned}&|I(\bar{A},\bar{U};\bar{Y}) - I(\bar{U};\bar{S}|\bar{A}) - R|^+ = 0, {}^\forall P_{\bar{A}} \in \mathscr{P}(\mathscr{A}), \\ &{}^\forall P_{\bar{U}|\bar{S}\bar{A}} \in \mathscr{W}(\mathscr{U}|\mathscr{S}\times\mathscr{A}), {}^\forall \bar{f}: \mathscr{U}\times\mathscr{S}\to\mathscr{X}.\end{aligned}$$

Thus, if $E_{\mathrm{r}}(R) = 0$ then $R \ge C$. From the above argument, $E_{\mathrm{r}}(R) = 0$ if and only if $R \ge C$. This implies that, for any $R < C$, there exists a sequence of codes $\{(f_{\mathrm{A}}^n, f_{\mathrm{C}}^n, \varphi_n)\}$ such that $\liminf_{n\to\infty} R_n \ge R$ and the error probability vanishes exponentially as the block length tends to infinity. This strengthens the proof of the achievability part of Theorem 1 in [2].

Remark 2. When $|\mathscr{S}| = 1$, i.e., the channel has a unique state, we have $(\bar{A},\bar{U}) \leftrightarrow \bar{X} \leftrightarrow \bar{Y}$. Hence, according to the data processing inequality, we have

$$I(\bar{A},\bar{U};\bar{Y}) \le I(\bar{X},\bar{Y}). \tag{1}$$

On the other hand, since $\bar{X}$ is a function of $\bar{U}$ (and a unique state), we have $\bar{Y} \leftrightarrow (\bar{A},\bar{U}) \leftrightarrow \bar{X}$, and according to the data processing inequality, we also have

$$I(\bar{X},\bar{Y}) \le I(\bar{A},\bar{U};\bar{Y}). \tag{2}$$

From (1) and (2), we have

$$I(\bar{A},\bar{U};\bar{Y}) = I(\bar{X},\bar{Y}). \tag{3}$$

Hence, by letting $|\mathscr{U}| \ge |\mathscr{X}|$, $E_{\mathrm{r}}(R)$ can be written as

$$\begin{aligned}
E_{\mathrm{r}}(R) &= \max_{P_{\bar{A}\bar{U}}\in\mathscr{P}(\mathscr{A}\times\mathscr{U})} \max_{\bar{f}:\mathscr{U}\to\mathscr{X}} \min_{P_{\bar{Y}|\bar{X}}\in\mathscr{W}(\mathscr{Y}|\mathscr{X})} \\
&\quad \times [D(P_{\bar{Y}|\bar{X}}\|P_{Y|X}|P_{\bar{X}}) + |I(\bar{A},\bar{U};\bar{Y}) - R|^{+}] \\
&\overset{(a)}{=} \max_{P_{\bar{A}\bar{U}}\in\mathscr{P}(\mathscr{A}\times\mathscr{U})} \max_{\bar{f}:\mathscr{U}\to\mathscr{X}} \min_{P_{\bar{Y}|\bar{X}}\in\mathscr{W}(\mathscr{Y}|\mathscr{X})} \\
&\quad \times [D(P_{\bar{Y}|\bar{X}}\|P_{Y|X}|P_{\bar{X}}) + |I(\bar{X};\bar{Y}) - R|^{+}] \\
&= \max_{P_{\bar{U}}\in\mathscr{P}(\mathscr{U})} \max_{\bar{f}:\mathscr{U}\to\mathscr{X}} \min_{P_{\bar{Y}|\bar{X}}\in\mathscr{W}(\mathscr{Y}|\mathscr{X})} \\
&\quad \times [D(P_{\bar{Y}|\bar{X}}\|P_{Y|X}|P_{\bar{X}}) + |I(\bar{X};\bar{Y}) - R|^{+}] \\
&\overset{(b)}{=} \max_{P_{\bar{X}}\in\mathscr{P}(\mathscr{X})} \min_{P_{\bar{Y}|\bar{X}}\in\mathscr{W}(\mathscr{Y}|\mathscr{X})} [D(P_{\bar{Y}|\bar{X}}\|P_{Y|X}|P_{\bar{X}}) \\
&\quad + |I(\bar{X};\bar{Y}) - R|^{+}],
\end{aligned}$$

where (a) follows from (3), and (b) comes from the fact that $\min_{P_{\bar{Y}|\bar{X}}\in\mathscr{W}(\mathscr{Y}|\mathscr{X})}[D(P_{\bar{Y}|\bar{X}}\|P_{Y|X}|P_{\bar{X}}) + |I(\bar{X};\bar{Y}) - R|^{+}]$ depends only on $P_{\bar{X}} \in \mathscr{P}(\mathscr{X})$. Thus, $E_{\mathrm{r}}(R)$ coincides with the well-known *random coding error exponent* (see [12, p.152]).

Remark 3. When $|\mathscr{A}| = 1$, i.e., any action cannot affect any states, $E_{\mathrm{r}}(R)$ can be written as

$$\begin{aligned}
E_{\mathrm{r}}(R) = &\min_{P_{\bar{S}}\in\mathscr{P}(\mathscr{S})} \max_{P_{\bar{U}|\bar{S}}\in\mathscr{W}(\mathscr{U}|\mathscr{S})} \max_{\bar{f}:\mathscr{U}\times\mathscr{S}\to\mathscr{X}} \min_{P_{\bar{Y}|\bar{X}\bar{S}}\in\mathscr{W}(\mathscr{Y}|\mathscr{X}\times\mathscr{S})} \\
&\times [D(P_{\bar{S}}\|P_{S}) + D(P_{\bar{Y}|\bar{X}\bar{S}}\|P_{Y|XS}|P_{\bar{S}\bar{X}}) \\
&+ |I(\bar{U};\bar{Y}) - I(\bar{U};\bar{S}) - R|^{+}],
\end{aligned}$$

where

$$P_{\bar{S}\bar{X}}(s,x) = \sum_{u\in\mathscr{U}} P_{\bar{S}}(s)P_{\bar{U}|\bar{S}}(u|s)1_{\{\bar{f}(u,s)\}}(x).$$

This error exponent $E_{\mathrm{r}}(R)$ coincides with the error exponent of Somekh-Baruch and Merhav [10, $\tilde{E}_2(R)$].

4. Proof of Theorem 2

Let us fix $\delta > 0$ and a type $P_{\bar{A}} \in \mathscr{P}(\mathscr{A})$ arbitrarily, and for all conditional types $P_{\bar{S}|\bar{A}} \in \mathscr{W}_n(\mathscr{S}|P_{\bar{A}})$, fix a function $\bar{f}: \mathscr{U}\times\mathscr{S} \to \mathscr{X}$ and a conditional type $P_{\bar{U}|\bar{S}\bar{A}} \in \mathscr{W}_n(\mathscr{U}|P_{\bar{A}} \cdot P_{\bar{S}|\bar{A}})$ arbitrarily. Although the function and the conditional type are functions of $P_{\bar{S}|\bar{A}}$, we will denote these as f and $P_{\bar{U}|\bar{S}\bar{A}}$ for the sake of brevity. Then, we define random encoders and a decoder as follows.

Random generation of $\mathscr{C}_A$: For each $m \in \mathscr{M}_n$, randomly and independently generate sequences $\mathbf{a}(m)$, each drawn according to the uniform distribution over $T(P_{\bar{A}})$. We denote the set of sequences $\{\mathbf{a}(1),\mathbf{a}(2),\cdots,\mathbf{a}(M_n)\}$ by $\mathscr{C}_A$.

Random generation of $\mathscr{C}(P_{\bar{S}|\bar{A}},m)$: For each conditional type $P_{\bar{S}|\bar{A}} \in \mathscr{W}_n(\mathscr{S}|P_{\bar{A}})$, let $\bar{M}_n(P_{\bar{S}|\bar{A}})$ be a positive integer such that

$$\bar{M}_n(P_{\bar{S}|\bar{A}}) = \lceil \exp\{n(I(\bar{U};\bar{S}|\bar{A}) + \delta)\}\rceil, \tag{4}$$

where a triple of RVs $(\bar{A},\bar{U},\bar{S})$ is distributed according to $P_{\bar{A}} \cdot P_{\bar{S}|\bar{A}} \cdot P_{\bar{U}|\bar{S}\bar{A}}$, and $\lceil z\rceil$ denotes the smallest integer which is greater than or equal to z. Then, for each $m \in \mathscr{M}_n$ and $l \in \bar{\mathscr{M}}_n(P_{\bar{S}|\bar{A}}) \triangleq \{1,2,\cdots,\bar{M}_n(P_{\bar{S}|\bar{A}})\}$, randomly and independently generate sequences $\mathbf{u}(l,m)$, each drawn according to the uniform distribution over $T(P_{\bar{U}|\bar{A}}|\mathbf{a}(m))$, where $\mathbf{a}(m)$ is the element of $\mathscr{C}_A$ corresponding to m, and

$$P_{\bar{U}|\bar{A}}(u|a) = \sum_{s\in\mathscr{S}} P_{\bar{S}|\bar{A}}(s|a)P_{\bar{U}|\bar{S}\bar{A}}(u|s,a).$$

Note that $P_{\bar{U}|\bar{S}\bar{A}}$ changes with $P_{\bar{S}|\bar{A}}$, and $T(P_{\bar{U}|\bar{A}}|\mathbf{a}(m))$ is not empty because $P_{\bar{U}|\bar{S}\bar{A}} \in \mathscr{W}_n(\mathscr{U}|P_{\bar{A}} \cdot P_{\bar{S}|\bar{A}})$ and $P_{\bar{S}|\bar{A}} \in \mathscr{W}_n(\mathscr{S}|P_{\bar{A}})$. We denote the set of sequences $\{\mathbf{u}(1,m),\cdots,\mathbf{u}(\bar{M}_n(P_{\bar{S}|\bar{A}}),m)\}$ by $\mathscr{C}(P_{\bar{S}|\bar{A}},m)$.

Action encoder: For a given $m \in \mathscr{M}_n$, chose $\mathbf{a}(m) \in \mathscr{C}_A$ as the action.

Channel encoder: We use the following two-step encoding.

1. For a given $m \in \mathscr{M}_n$ and a state $\mathbf{s} \in \mathscr{S}^n$, find l such that $\mathbf{u}(l,m) \in \mathscr{C}(P_{\mathbf{s}|\mathbf{a}(m)},m) \cap T(P_{\bar{U}|\bar{S}\bar{A}}|\mathbf{s},\mathbf{a}(m))$. If more than one such l exists, pick one of them randomly with uniform distribution. If there is no such l, generate $\mathbf{u}$ uniformly from $T(P_{\bar{U}|\bar{S}\bar{A}}|\mathbf{s},\mathbf{a}(m))$.
2. For the sequence $\mathbf{u}$ found in the step 1), generate a codeword $\mathbf{x} = \bar{f}^n(\mathbf{u},\mathbf{s})$, where

$$\bar{f}^n(\mathbf{u},\mathbf{s}) = (\bar{f}(u_1,s_1),\bar{f}(u_2,s_2),\cdots,\bar{f}(u_n,s_n)).$$

Decoder: For a given $\mathbf{y} \in \mathscr{Y}^n$, the decoder finds a $\hat{m} \in \mathscr{M}_n$ that maximizes the penalized empirical mutual information, i.e.,

$$\hat{m} = \operatorname*{argmax}_{\hat{m}\in\mathscr{M}_n} \max_{P_{\bar{S}|\bar{A}}\in\mathscr{W}(\mathscr{S}|P_{\bar{A}})} \max_{\mathbf{u}\in\mathscr{C}(P_{\bar{S}|\bar{A}},\hat{m})} \left[I(\mathbf{a}(\hat{m}),\mathbf{u};\mathbf{y}) - \bar{R}_n(P_{\bar{S}|\bar{A}})\right],$$

where

$$I(\mathbf{a}(\hat{m}),\mathbf{u};\mathbf{y}) = \sum_{(a,u,y)\in\mathscr{A}\times\mathscr{U}\times\mathscr{Y}} P_{\mathbf{a}(\hat{m})\mathbf{u}\mathbf{y}}(a,u,y) \times \log\frac{P_{\mathbf{a}(\hat{m})\mathbf{u}\mathbf{y}}(a,u,y)}{P_{\mathbf{a}(\hat{m})\mathbf{u}}(a,u)P_{\mathbf{y}}(y)},$$

and

$$\bar{R}_n(P_{\bar{S}|\bar{A}}) \triangleq \frac{1}{n}\log\bar{M}_n(P_{\bar{S}|\bar{A}}).$$

As we stated in Introduction, the "penalized" means the existence of a term subtracting from an empirical mutual information. In the above decoder, the penalty is $\bar{R}_n(P_{\bar{S}|\bar{A}})$.

The decoder using a penalized empirical mutual information was introduced by Moulin and Wang [11], and it may be regarded as an empirical version of the MAP decoder (see [11, p.1332]). Thus, in this decoding procedure, we are using an empirical version of the MAP decoder.

We now analyze the following average error probability.

$$\overline{\varepsilon}_n \triangleq \sum_{(f_{\mathrm{A}}^n, f_{\mathrm{C}}^n, \varphi_n)} \Pr\{(F_A^n, F_C^n, \Phi^n) = (f_{\mathrm{A}}^n, f_{\mathrm{C}}^n, \varphi_n)\}\, \varepsilon_n(f_{\mathrm{A}}^n, f_{\mathrm{C}}^n, \varphi^n),$$

where the triple of RVs (F_A^n, F_C^n, Φ^n) denotes random encoders and a corresponding decoder.

In order to analyze the average error probability, for a message $m \in \mathscr{M}_n$, we first consider the encoding error, i.e., the case where no index l can be found in the first step of the channel encoder. Thus, the encoding error occurs if the event

$$\mathscr{E}_m \triangleq \{\mathbf{U}(l,m) \notin T(P_{\bar{U}|\bar{S}\bar{A}}|\mathbf{S}, \mathbf{A}(m)), {}^{\forall} l \in \bar{\mathscr{M}}_n(P_{\mathbf{S}|\mathbf{A}(m)})\}$$

occurs, where $(\mathbf{A}(m), \mathbf{S}, \mathbf{U}(l,m))$ denote RVs induced by the random generation of $\mathscr{C}_A$, the channel $P_{S|A}$ and the random generation of $\mathscr{C}(P_{\bar{S}|\bar{A}}, m)$. Then, for any conditional type $P_{\bar{S}|\bar{A}} \in \mathscr{W}_n(\mathscr{S}|P_{\bar{A}})$, any $\mathbf{a} \in T(P_{\bar{A}})$, and $\mathbf{s} \in T(P_{\bar{S}|\bar{A}}|\mathbf{a})$, we have

$$\begin{aligned}
&\Pr\{\mathscr{E}_m|\mathbf{S}=\mathbf{s}, \mathbf{A}(m)=\mathbf{a}\}\\
&= \Pr\{\mathbf{U}(l,m) \notin T(P_{\bar{U}|\bar{S}\bar{A}}|\mathbf{s},\mathbf{a}), {}^{\forall} l \in \bar{\mathscr{M}}_n(P_{\bar{S}|\bar{A}})|\mathbf{S}=\mathbf{s},\\
&\quad \mathbf{A}(m)=\mathbf{a}\}\\
&\overset{(a)}{=} \prod_{l \in \bar{\mathscr{M}}_n(P_{\bar{S}|\bar{A}})} \Pr\{\mathbf{U}(l,m) \notin T(P_{\bar{U}|\bar{S}\bar{A}}|\mathbf{s},\mathbf{a})|\mathbf{S}=\mathbf{s},\\
&\quad \mathbf{A}(m)=\mathbf{a}\}\\
&\overset{(b)}{=} (\Pr\{\hat{\mathbf{U}} \notin T(P_{\bar{U}|\bar{S}\bar{A}}|\mathbf{s},\mathbf{a})|\mathbf{S}=\mathbf{s}, \mathbf{A}(m)=\mathbf{a}\})^{\bar{M}_n(P_{\bar{S}|\bar{A}})}\\
&= (1 - \Pr\{\hat{\mathbf{U}} \in T(P_{\bar{U}|\bar{S}\bar{A}}|\mathbf{s},\mathbf{a})|\mathbf{S}=\mathbf{s}, \mathbf{A}(m)=\mathbf{a}\})^{\bar{M}_n(P_{\bar{S}|\bar{A}})},
\end{aligned} \tag{5}$$

where $\hat{\mathbf{U}}$ is an RV drawn according to the uniform distribution over $T(P_{\bar{U}|\bar{A}}|\mathbf{a})$, (a) comes from the fact that $\mathbf{U}(l,m)$ is independent for all $l \in \bar{\mathscr{M}}_n(P_{\bar{S}|\bar{A}})$, and (b) comes from the fact that $\mathbf{U}(l,m)$ drawn according to the uniform distribution over $T(P_{\bar{U}|\bar{A}}|\mathbf{a})$. Since $T(P_{\bar{U}|\bar{S}\bar{A}}|\mathbf{s},\mathbf{a}) \subseteq T(P_{\bar{U}|\bar{A}}|\mathbf{a})$ for all $\mathbf{s} \in T(P_{\bar{S}|\bar{A}}|\mathbf{a})$, we have

$$\begin{aligned}
&\Pr\{\hat{\mathbf{U}} \in T(P_{\bar{U}|\bar{S}\bar{A}}|\mathbf{s},\mathbf{a})|\mathbf{S}=\mathbf{s}, \mathbf{A}(m)=\mathbf{a}\}\\
&\quad = \frac{|T(P_{\bar{U}|\bar{S}\bar{A}}|\mathbf{s},\mathbf{a})|}{|T(P_{\bar{U}|\bar{A}}|\mathbf{a})|}\\
&\quad \overset{(a)}{\geq} \exp\{-n(I(\bar{U};\bar{S}|\bar{A}) + \delta_n^{(1)})\},
\end{aligned} \tag{6}$$

where (a) comes from 2) of Lemma 1, a triple of RVs $(\bar{A}, \bar{S}, \bar{U})$ is distributed according to $P_{\bar{A}} \cdot P_{\bar{S}|\bar{A}} \cdot P_{\bar{U}|\bar{S}\bar{A}}$, and

$$\delta_n^{(1)} = \frac{|\mathscr{A}||\mathscr{S}||\mathscr{U}|\log(n+1)}{n}.$$

By combining (5) and (6), we have

$$\begin{aligned}
&\Pr\{\mathscr{E}_m|\mathbf{A}(m)=\mathbf{a}, \mathbf{S}=\mathbf{s}\}\\
&\leq (1 - \exp\{-n(I(\bar{U};\bar{S}|\bar{A}) + \delta_n^{(1)})\})^{\bar{M}_n(P_{\bar{S}|\bar{A}})}\\
&\overset{(a)}{\leq} \exp\{-\exp\{-n(I(\bar{U};\bar{S}|\bar{A}) + \delta_n^{(1)})\}\bar{M}_n(P_{\bar{S}|\bar{A}})\}\\
&\overset{(b)}{\leq} \exp\{-\exp\{n(\delta - \delta_n^{(1)})\}\},
\end{aligned}$$

where (a) comes from the fact that $(1-x)^y \leq \exp\{-xy\}$ $(0 \leq x \leq 1, y \geq 0)$, and (b) follows from (4). Hence, we have

$$\begin{aligned}
&\Pr\{\mathscr{E}_m\}\\
&= \sum_{(\mathbf{a},\mathbf{s}) \in \mathscr{A}^n \times \mathscr{S}^n} \Pr\{\mathbf{A}(m)=\mathbf{a}, \mathbf{S}=\mathbf{s}\} \Pr\{\mathscr{E}_m|\mathbf{A}(m)=\mathbf{a}, \mathbf{S}=\mathbf{s}\}\\
&\overset{(a)}{=} \sum_{\mathbf{a} \in T(P_{\bar{A}})} \sum_{P_{\bar{S}|\bar{A}} \in \mathscr{W}_n(\mathscr{S}|P_{\bar{A}})} \sum_{\mathbf{s} \in T(P_{\bar{S}|\bar{A}}|\mathbf{a})} \Pr\{\mathbf{A}(m)=\mathbf{a}, \mathbf{S}=\mathbf{s}\}\\
&\quad \times \Pr\{\mathscr{E}_m|\mathbf{A}(m)=\mathbf{a}, \mathbf{S}=\mathbf{s}\}\\
&\leq \sum_{\mathbf{a} \in T(P_{\bar{A}})} \sum_{P_{\bar{S}|\bar{A}} \in \mathscr{W}_n(\mathscr{S}|P_{\bar{A}})} \sum_{\mathbf{s} \in T(P_{\bar{S}|\bar{A}}|\mathbf{a})} \Pr\{\mathbf{A}(m)=\mathbf{a}, \mathbf{S}=\mathbf{s}\}\\
&\quad \times \exp\{-\exp\{n(\delta - \delta_n^{(1)})\}\}\\
&= \exp\{-\exp\{n(\delta - \delta_n^{(1)})\}\},
\end{aligned} \tag{7}$$

where (a) comes from the fact that $\mathbf{A}(m)$ is an RV with the uniform distribution over $T(P_{\bar{A}})$.

Next, we consider the decoding error, i.e., the case where the decoder outputs a wrong message. Thus, the decoding error occurs if the event

$$\begin{aligned}
\mathscr{E}'_m = \{&{}^{\exists} m' \in \mathscr{M}_n, {}^{\exists} P_{\bar{S}'|\bar{A}} \in \mathscr{W}_n(\mathscr{S}|P_{\bar{A}}), {}^{\exists} l' \in \bar{\mathscr{M}}_n(P_{\bar{S}'|\bar{A}})\\
&\text{such that } m' \neq m,\ I(\mathbf{A}(m'), \mathbf{U}(l',m'); \mathbf{Y}) - \bar{R}_n(P_{\bar{S}'|\bar{A}})\\
&\geq I(\mathbf{A}(m), \mathbf{U}(l,m); \mathbf{Y}) - \bar{R}_n(P_{\bar{S}|\bar{A}})\\
&\text{for all } P_{\bar{S}|\bar{A}} \in \mathscr{W}_n(\mathscr{S}|P_{\bar{A}}) \text{ and } l \in \bar{\mathscr{M}}_n(P_{\bar{S}|\bar{A}})\}
\end{aligned}$$

occurs, where the RV $\mathbf{Y}$ denotes the channel outputs. Then, due to the symmetry of the code construction, we have

$$\overline{\varepsilon}_n \leq \Pr\{\mathscr{E}_1\} + \Pr\{\mathscr{E}_1^c \cap \mathscr{E}'_1\}. \tag{8}$$

The second term on the right-hand side of (8) can be bounded as

$$\begin{aligned}
&\Pr\{\mathscr{E}_1^c \cap \mathscr{E}'_1\}\\
&\quad = \sum_{(\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{x},\mathbf{y})} \Pr\{\mathbf{A}(1)=\mathbf{a}, \mathbf{S}=\mathbf{s}, \mathbf{U}=\mathbf{u}, \mathbf{X}=\mathbf{x}, \mathbf{Y}=\mathbf{y}\}\\
&\qquad \times \Pr\{\mathscr{E}_1^c \cap \mathscr{E}'_1|\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{x},\mathbf{y}\}\\
&\quad \leq \sum_{(\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{x},\mathbf{y})} \Pr\{\mathbf{A}(1)=\mathbf{a}, \mathbf{S}=\mathbf{s}, \mathbf{U}=\mathbf{u}, \mathbf{X}=\mathbf{x}, \mathbf{Y}=\mathbf{y}\}\\
&\qquad \times \Pr\{\mathscr{E}'_1|\mathscr{E}_1^c, \mathbf{a},\mathbf{s},\mathbf{u},\mathbf{x},\mathbf{y}\},
\end{aligned} \tag{9}$$

where the RV $\mathbf{U}$ denotes the sequence found in the step 1) of the channel encoder, and $\mathbf{X}$ denotes the codeword. Note that

$$\Pr\{\mathbf{A}(1)=\mathbf{a},\mathbf{S}=\mathbf{s},\mathbf{U}=\mathbf{u},\mathbf{X}=\mathbf{x},\mathbf{Y}=\mathbf{y}\}$$
$$=P_{A^n}(\mathbf{a})P^n_{S|A}(\mathbf{s}|\mathbf{a})P_{U^n|S^nA^n}(\mathbf{u}|\mathbf{s},\mathbf{a})1_{\{\bar{f}^n(\mathbf{u},\mathbf{s})\}}(\mathbf{x})P^n_{Y|XS}(\mathbf{y}|\mathbf{x},\mathbf{s}), \tag{10}$$

$$P_{A^n}(\mathbf{a})=\begin{cases}\frac{1}{|T(P_{\bar{A}})|} & \text{if } \mathbf{a}\in T(P_{\bar{A}}),\\ 0 & \text{if } \mathbf{a}\notin T(P_{\bar{A}}),\end{cases} \tag{11}$$

and

$$P_{U^n|S^nA^n}(\mathbf{u}|\mathbf{s},\mathbf{a})=\begin{cases}\frac{1}{|T(P_{\bar{U}|\bar{S}\bar{A}}|\mathbf{s},\mathbf{a})|} & \text{if } \mathbf{u}\in T(P_{\bar{U}|\bar{S}\bar{A}}|\mathbf{s},\mathbf{a}),\\ 0 & \text{if } \mathbf{u}\notin T(P_{\bar{U}|\bar{S}\bar{A}}|\mathbf{s},\mathbf{a}).\end{cases} \tag{12}$$

Hence, for any type $P_{\bar{S}|\bar{A}}\in\mathscr{W}_n(\mathscr{S}|P_{\bar{A}})$, $P_{\bar{Y}|\bar{X}\bar{S}}\in\mathscr{W}_n(\mathscr{Y}|P_{\bar{S}\bar{X}})$, and the type $P_{\bar{A}\bar{S}\bar{U}\bar{X}\bar{Y}}$ satisfying

$$P_{\bar{A}\bar{S}\bar{U}\bar{X}\bar{Y}}(a,s,u,x,y)=P_{\bar{A}}(a)P_{\bar{S}|\bar{A}}(s|a)P_{\bar{U}|\bar{S}\bar{A}}(u|s,a)\times 1_{\{\bar{f}(u,s)\}}(x)P_{\bar{Y}|\bar{X}\bar{S}}(y|x,s), \tag{13}$$

we have

$$\sum_{(\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{x},\mathbf{y})\in T(P_{\bar{A}\bar{S}\bar{U}\bar{X}\bar{Y}})}\Pr\{\mathbf{A}(1)=\mathbf{a},\mathbf{S}=\mathbf{s},\mathbf{U}=\mathbf{u},\mathbf{X}=\mathbf{x},\mathbf{Y}=\mathbf{y}\}$$
$$=\sum_{\mathbf{a}\in T(P_{\bar{A}})}P_{A^n}(\mathbf{a})\sum_{\mathbf{s}\in T(P_{\bar{S}|\bar{A}}|\mathbf{a})}P^n_{S|A}(\mathbf{s}|\mathbf{a})\sum_{\mathbf{u}\in T(P_{\bar{U}|\bar{S}\bar{A}}|\mathbf{s},\mathbf{a})}P_{U^n|S^nA^n}(\mathbf{u}|\mathbf{s},\mathbf{a})$$
$$\times\sum_{\{\mathbf{x}\in\mathscr{X}^n:\mathbf{x}=f^n(\mathbf{u},\mathbf{s})\}}1_{\{f^n(\mathbf{u},\mathbf{s})\}}(\mathbf{x})\sum_{\mathbf{y}\in T(P_{\bar{Y}|\bar{X}\bar{S}}|\mathbf{s},\mathbf{x})}P^n_{Y|XS}(\mathbf{y}|\mathbf{x},\mathbf{s})$$
$$\overset{(a)}{\leq}1\times\exp\{-nD(P_{\bar{S}|\bar{A}}\|P_{S|A}|P_{\bar{A}})\}\times 1\times 1$$
$$\times\exp\{-nD(P_{\bar{Y}|\bar{X}\bar{S}}\|P_{Y|XS}|P_{\bar{S}\bar{X}})\}$$
$$=\exp\{-n(D(P_{\bar{S}|\bar{A}}\|P_{S|A}|P_{\bar{A}})+D(P_{\bar{Y}|\bar{X}\bar{S}}\|P_{Y|XS}|P_{\bar{S}\bar{X}}))\}, \tag{14}$$

where (a) comes from 2) and 3) of Lemma 1. On the other hand, for any $(\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{x},\mathbf{y})\in T(P_{\bar{A}\bar{S}\bar{U}\bar{X}\bar{Y}})$, we have

$$\Pr\{\mathscr{E}'_1|\mathscr{E}^c_1,\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{x},\mathbf{y})$$
$$=\Pr\{{}^\exists m'\in\mathscr{M}_n,{}^\exists P_{\bar{S}'|\bar{A}}\in\mathscr{W}_n(\mathscr{S}|P_{\bar{A}}),{}^\exists l'\in\bar{\mathscr{M}}_n(P_{\bar{S}'|\bar{A}})$$
$$\text{such that } m'\neq 1, I(\mathbf{A}(m'),\mathbf{U}(l',m');\mathbf{y})-\bar{R}_n(P_{\bar{S}'|\bar{A}})$$
$$\geq I(\mathbf{a},\mathbf{U}(l,1);\mathbf{y})-\bar{R}_n(P_{\bar{S}|\bar{A}})$$
$$\text{for all } P_{\bar{S}|\bar{A}}\in\mathscr{W}_n(\mathscr{S}|P_{\bar{A}}) \text{ and } l\in\bar{\mathscr{M}}_n(P_{\bar{S}|\bar{A}})$$
$$|\mathscr{E}^c_1,\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{x},\mathbf{y}\}$$
$$\leq\Pr\{{}^\exists m'\in\mathscr{M}_n,{}^\exists P_{\bar{S}'|\bar{A}}\in\mathscr{W}_n(\mathscr{S}|P_{\bar{A}}),{}^\exists l'\in\bar{\mathscr{M}}_n(P_{\bar{S}'|\bar{A}})$$
$$\text{such that } m'\neq 1, I(\mathbf{A}(m'),\mathbf{U}(l',m');\mathbf{y})-\bar{R}_n(P_{\bar{S}'|\bar{A}})$$
$$\geq I(\mathbf{a},\mathbf{U}(l,1);\mathbf{y})-\bar{R}_n(P_{\mathbf{s}|\mathbf{a}})\text{ for all } l\in\bar{\mathscr{M}}_n(P_{\mathbf{s}|\mathbf{a}})$$
$$|\mathscr{E}^c_1,\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{x},\mathbf{y}\}$$
$$\overset{(a)}{\leq}\Pr\{{}^\exists m'\in\mathscr{M}_n,{}^\exists P_{\bar{S}'|\bar{A}}\in\mathscr{W}_n(\mathscr{S}|P_{\bar{A}}),{}^\exists l'\in\bar{\mathscr{M}}_n(P_{\bar{S}'|\bar{A}})$$
$$\text{such that } m'\neq 1, I(\mathbf{A}(m'),\mathbf{U}(l',m');\mathbf{y})-\bar{R}_n(P_{\bar{S}'|\bar{A}})$$
$$\geq I(\mathbf{a},\mathbf{u};\mathbf{y})-\bar{R}_n(P_{\mathbf{s}|\mathbf{a}})|\mathscr{E}^c_1,\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{x},\mathbf{y}\}$$
$$\leq\sum_{\substack{m'\in\mathscr{M}_n:\\ m'\neq 1}}\sum_{P_{\bar{S}'|\bar{A}}\in\mathscr{W}_n(\mathscr{S}|P_{\bar{A}})}\sum_{l'\in\bar{\mathscr{M}}_n(P_{\bar{S}'|\bar{A}})}\Pr\{I(\mathbf{A}(m'),\mathbf{U}(l',m');\mathbf{y})$$
$$-\bar{R}_n(P_{\bar{S}'|\bar{A}})\geq I(\mathbf{a},\mathbf{u};\mathbf{y})-\bar{R}_n(P_{\mathbf{s}|\mathbf{a}})|\mathscr{E}^c_1,\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{x},\mathbf{y}\},$$

where (a) comes from the fact that there exists $l\in\bar{\mathscr{M}}_n(P_{\mathbf{s}|\mathbf{a}}$ such that $I(\mathbf{a},\mathbf{U}(l,1);\mathbf{y})=I(\mathbf{a},\mathbf{u};\mathbf{y})$ due to the condition of $\mathscr{E}^c_1$. Let us define

$$\mathscr{E}(\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{y},P_{\bar{S}'|\bar{A}})\triangleq\{(\mathbf{a}',\mathbf{u}')\in T(P_{\bar{A}\bar{U}'}):I(\mathbf{a}',\mathbf{u}';\mathbf{y})-\bar{R}_n(P_{\bar{S}'|\bar{A}})\geq I(\mathbf{a},\mathbf{u};\mathbf{y})-\bar{R}_n(P_{\mathbf{s}|\mathbf{a}})\},$$

where

$$P_{\bar{A}\bar{U}'}(a,u)=P_{\bar{A}}(a)\sum_{s\in\mathscr{S}}P_{\bar{S}'|\bar{A}}(s|a)P_{\bar{U}|\bar{S}\bar{A}}(u|s,a).$$

We also define the corresponding set of conditional types

$$\mathscr{T}_n(\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{y},P_{\bar{S}'|\bar{A}})$$
$$\triangleq\{P_{\bar{A}\bar{U}'|\bar{Y}}\in\mathscr{W}_n(\mathscr{A}\times\mathscr{U}|P_{\mathbf{y}}):$$
$$\sum_{y\in\mathscr{Y}}P_{\mathbf{y}}(y)P_{\bar{A}\bar{U}'|\bar{Y}}(a,u|y)=P_{\bar{A}\bar{U}'}(a,u),$$
$$I(\bar{A},\bar{U}';\bar{Y})-\bar{R}_n(P_{\bar{S}'|\bar{A}})\geq I(\mathbf{a},\mathbf{u};\mathbf{y})-\bar{R}_n(P_{\mathbf{s}|\mathbf{a}})\}, \tag{15}$$

where the triple of RVs $(\bar{A},\bar{U}',\bar{Y})$ is drawn from $P_{\bar{A}\bar{U}'\bar{Y}}=P_{\bar{A}\bar{U}'|\bar{Y}}\cdot P_{\mathbf{y}}$. Then, we have

$$\mathscr{E}(\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{y},P_{\bar{S}'|\bar{A}})=\bigcup_{P_{\bar{A}\bar{U}'|\bar{Y}}\in\mathscr{T}_n(\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{y},P_{\bar{S}'|\bar{A}})}T(P_{\bar{A}\bar{U}'|\bar{Y}}|\mathbf{y}).$$

Thus, we have

$$\Pr\{I(\mathbf{A}(m'),\mathbf{U}(l',m');\mathbf{y})-\bar{R}_n(P_{\bar{S}'|\bar{A}})$$
$$\geq I(\mathbf{a},\mathbf{u};\mathbf{y})-\bar{R}_n(P_{\bar{S}|\bar{A}})|\mathscr{E}^c_1,\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{x},\mathbf{y}\}$$
$$=\sum_{(\mathbf{a}',\mathbf{u}')\in\mathscr{E}(\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{y},P_{\bar{S}'|\bar{A}})}\Pr\{(\mathbf{A}(m'),\mathbf{U}(l',m'))=(\mathbf{a}',\mathbf{u}')|\mathscr{E}^c_1,$$
$$\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{x},\mathbf{y}\}$$
$$\overset{(a)}{=}\sum_{(\mathbf{a}',\mathbf{u}')\in\mathscr{E}(\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{y},P_{\bar{S}'|\bar{A}})}\Pr\{(\mathbf{A}(m'),\mathbf{U}(l',m'))=(\mathbf{a}',\mathbf{u}')\}$$
$$=\sum_{(\mathbf{a}',\mathbf{u}')\in\mathscr{E}(\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{y},P_{\bar{S}'|\bar{A}})}\Pr\{\mathbf{A}(m')=\mathbf{a}'\}$$
$$\times\Pr\{\mathbf{U}(l',m')=\mathbf{u}'|\mathbf{A}(m')=\mathbf{a}'\}$$
$$=\sum_{(\mathbf{a}',\mathbf{u}')\in\mathscr{E}(\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{y},P_{\bar{S}'|\bar{A}})}\frac{1}{|T(P_{\bar{A}})|}\,\frac{1}{|T(P_{\bar{U}'|\bar{A}}|\mathbf{a}')|}$$
$$\overset{(b)}{=}\sum_{(\mathbf{a}',\mathbf{u}')\in\mathscr{E}(\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{y},P_{\bar{S}'|\bar{A}})}\frac{1}{|T(P_{\bar{A}\bar{U}'})|}$$
$$=\sum_{P_{\bar{A}\bar{U}'|\bar{Y}}\in\mathscr{T}_n(\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{y},P_{\bar{S}'|\bar{A}})}\sum_{(\mathbf{a}',\mathbf{u}')\in T(P_{\bar{A}\bar{U}'|\bar{Y}}|\mathbf{y})}\frac{1}{|T(P_{\bar{A}\bar{U}'})|}$$
$$=\sum_{P_{\bar{A}\bar{U}'|\bar{Y}}\in\mathscr{T}_n(\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{y},P_{\bar{S}'|\bar{A}})}\frac{|T(P_{\bar{A}\bar{U}'|\bar{Y}}|\mathbf{y})|}{|T(P_{\bar{A}\bar{U}'})|}$$
$$\overset{(c)}{\leq}\sum_{P_{\bar{A}\bar{U}'|\bar{Y}}\in\mathscr{T}_n(\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{y},P_{\bar{S}'|\bar{A}})}(n+1)^{|\mathscr{A}||\mathscr{U}|}\exp\{-nI(\bar{A},\bar{U}';\bar{Y})\}$$

$$\overset{\text{(d)}}{\leq} \exp\{-n(I(\mathbf{a},\mathbf{u};\mathbf{y}) - \bar{R}_n(P_{\mathbf{s}|\mathbf{a}}) + \bar{R}_n(P_{\bar{S}'|\bar{A}}) - \delta_n^{(2)})\},$$

where (a) comes from the fact that $m' \neq 1$, (b) comes from the fact that since $|T(P_{\bar{U}'|\bar{A}}|\mathbf{a}')|$ only depends on $P_{\bar{A}}$ and $P_{\bar{U}'|\bar{A}}$,

$$\begin{aligned}|T(P_{\bar{A}\bar{U}'})| &= |\{(\mathbf{a},\mathbf{u}) : \mathbf{a} \in T(P_{\bar{A}}), \mathbf{u} \in T(P_{\bar{U}'|\bar{A}}|\mathbf{a})\}| \\ &= |T(P_{\bar{A}})||T(P_{\bar{U}'|\bar{A}}|\mathbf{a}')\}|,\end{aligned}$$

(c) comes from 2) of Lemma 1, (d) comes from the definition (15) and 1) of Lemma 1, and

$$\delta_n^{(2)} = \frac{|\mathscr{A}||\mathscr{U}|(|\mathscr{Y}|+1)\log(n+1)}{n}.$$

Thus, for any $(\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{x},\mathbf{y}) \in T(P_{\bar{A}\bar{S}\bar{U}\bar{X}\bar{Y}})$, we have

$$\begin{aligned}
&\Pr\{\mathscr{E}_1'|\mathscr{E}_1^c,\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{x},\mathbf{y}\} \\
&\leq \sum_{\substack{m'\in\mathscr{M}_n:\\ m'\neq 1}} \sum_{P_{\bar{S}'|\bar{A}}\in\mathscr{W}_n(\mathscr{S}|P_{\bar{A}})} \sum_{l'\in\mathscr{M}_n(P_{\bar{S}'|\bar{A}})} \Pr\{I(\mathbf{A}(m'),\mathbf{U}(l',m');\mathbf{y}) \\
&\quad -\bar{R}_n(P_{\bar{S}'|\bar{A}}) \geq I(\bar{A},\bar{U};\bar{Y}) - \bar{R}_n(P_{\bar{S}|\bar{A}})|\mathscr{E}_1^c,\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{x},\mathbf{y}\} \\
&\leq \sum_{\substack{m'\in\mathscr{M}_n:\\ m'\neq 1}} \sum_{P_{\bar{S}'|\bar{A}}\in\mathscr{W}_n(\mathscr{S}|P_{\bar{A}})} \sum_{l'\in\mathscr{M}_n(P_{\bar{S}'|\bar{A}})} \exp\{-n(I(\bar{A},\bar{U};\bar{Y}) \\
&\quad -\bar{R}_n(P_{\bar{S}|\bar{A}}) + \bar{R}_n(P_{\bar{S}'|\bar{A}}) - \delta_n^{(2)})\} \\
&\leq \max_{P_{\bar{S}'|\bar{A}}\in\mathscr{W}_n(\mathscr{S}|P_{\bar{A}})} \exp\{-n(I(\bar{A},\bar{U};\bar{Y}) - \bar{R}_n(P_{\bar{S}|\bar{A}}) + \bar{R}_n(P_{\bar{S}'|\bar{A}}) \\
&\quad -\bar{R}_n(P_{\bar{S}'|\bar{A}}) - R_n - \delta_n^{(3)})\} \\
&= \max_{P_{\bar{S}'|\bar{A}}\in\mathscr{W}_n(\mathscr{S}|P_{\bar{A}})} \exp\{-n(I(\bar{A},\bar{U};\bar{Y}) - \bar{R}_n(P_{\bar{S}|\bar{A}}) - R_n - \delta_n^{(3)})\} \\
&= \exp\{-n(I(\bar{A},\bar{U};\bar{Y}) - \bar{R}_n(P_{\bar{S}|\bar{A}}) - R_n - \delta_n^{(3)})\} \\
&\leq 2\exp\{-n(I(\bar{A},\bar{U};\bar{Y}) - I(\bar{U};\bar{S}|\bar{A}) - R_n - \delta - \delta_n^{(3)})\},
\end{aligned} \tag{16}$$

where

$$\delta_n^{(3)} = \delta_n^{(2)} + \frac{|\mathscr{A}||\mathscr{S}|\log(n+1)}{n}.$$

By combing (7), (8), (9), (14), and (16), we have

$$\begin{aligned}
\overline{\varepsilon}_n &\leq \exp\{-\exp\{n(\delta - \delta_n^{(1)})\}\} \\
&\quad + \sum_{(\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{x},\mathbf{y})} \Pr\{\mathbf{A}(1)=\mathbf{a},\mathbf{S}=\mathbf{s},\mathbf{U}=\mathbf{u},\mathbf{X}=\mathbf{x},\mathbf{Y}=\mathbf{y}\} \\
&\quad \times \Pr\{\mathscr{E}_1'|\mathscr{E}_1^c,\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{x},\mathbf{y}\}) \\
&\overset{\text{(a)}}{=} \exp\{-\exp\{n(\delta - \delta_n^{(1)})\}\} + \sum_{P_{\bar{S}|\bar{A}}\in\mathscr{W}_n(\mathscr{S}|P_{\bar{A}})} \\
&\quad \times \sum_{P_{\bar{Y}|\bar{X}\bar{S}}\in\mathscr{W}_n(\mathscr{Y}|P_{\bar{S}\bar{X}})} \sum_{(\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{x},\mathbf{y})\in T(P_{\bar{A}\bar{S}\bar{U}\bar{X}\bar{Y}})} \\
&\quad \times \Pr\{\mathbf{A}(1)=\mathbf{a},\mathbf{S}=\mathbf{s},\mathbf{U}=\mathbf{u},\mathbf{X}=\mathbf{x},\mathbf{Y}=\mathbf{y}\} \\
&\quad \times \Pr\{\mathscr{E}_1'|\mathscr{E}_1^c,\mathbf{a},\mathbf{s},\mathbf{u},\mathbf{x},\mathbf{y}\}) \\
&\leq \exp\{-\exp\{n(\delta - \delta_n^{(1)})\}\} + (n+1)^{|\mathscr{A}||\mathscr{S}|(1+|\mathscr{S}||\mathscr{X}||\mathscr{Y}|)} \\
&\quad \times 2\exp\{-n \min_{P_{\bar{S}|\bar{A}}\in\mathscr{W}_n(\mathscr{S}|P_{\bar{A}})} \min_{P_{\bar{Y}|\bar{X}\bar{S}}\in\mathscr{W}_n(\mathscr{Y}|P_{\bar{S}\bar{X}})} \\
&\quad \times [D(P_{\bar{S}|\bar{A}}\|P_{S|A}|P_{\bar{A}}) + D(P_{\bar{Y}|\bar{X}\bar{S}}\|P_{Y|XS}|P_{\bar{S}\bar{X}}) \\
&\quad + |I(\bar{A},\bar{U};\bar{Y}) - I(\bar{U};\bar{S}|\bar{A}) - R_n - \delta - \delta_n^{(3)}|^+]\},
\end{aligned} \tag{17}$$

where (a) comes from (10)–(12), and $P_{\bar{A}\bar{S}\bar{U}\bar{X}\bar{Y}}$ is characterized by (13). Since (17) holds for any fixed $\delta > 0$ and $P_{\bar{A}} \in \mathscr{P}(\mathscr{A})$, and any fixed $P_{\bar{U}|\bar{S}\bar{A}} \in \mathscr{W}_n(\mathscr{U}|P_{\bar{A}} \cdot P_{\bar{S}|\bar{A}})$ and $\bar{f}: \mathscr{U} \times \mathscr{S} \to \mathscr{X}$ which are dependent on $P_{\bar{S}|\bar{A}} \in \mathscr{W}_n(\mathscr{S}|P_{\bar{A}})$, we can choose them in order to optimize the upper bound of the error probability. Hence, for any $\varepsilon > 0$ and sufficiently large n, there exists a code $(f_{\mathrm{A}}^n, f_{\mathrm{C}}^n, \varphi_n)$ such that

$$\begin{aligned}
&\varepsilon_n(f_{\mathrm{A}}^n, f_{\mathrm{C}}^n, \varphi_n) \\
&\leq \exp\{-n \max_{P_{\bar{A}}\in\mathscr{P}_n(\mathscr{A})} \min_{P_{\bar{S}|\bar{A}}\in\mathscr{W}_n(\mathscr{S}|P_{\bar{A}})} \max_{P_{\bar{U}|\bar{S}\bar{A}}\in\mathscr{W}_n(\mathscr{U}|P_{\bar{A}\bar{S}})} \max_{\bar{f}:\mathscr{U}\times\mathscr{S}\to\mathscr{X}} \\
&\quad \min_{P_{\bar{Y}|\bar{X}\bar{S}}\in\mathscr{W}_n(\mathscr{Y}|P_{\bar{S}\bar{X}})} [D(P_{\bar{S}|\bar{A}}\|P_{S|A}|P_{\bar{A}}) + D(P_{\bar{Y}|\bar{X}\bar{S}}\|P_{Y|XS}|P_{\bar{S}\bar{X}}) \\
&\quad + |I(\bar{A},\bar{U};\bar{Y}) - I(\bar{U};\bar{S}|\bar{A}) - R_n - \varepsilon|^+]\}.
\end{aligned}$$

This completes the proof of Theorem 2.

5. Conclusion

In this paper, we have dealt with channels with action-dependent states, and derived an exponential error bound in Theorem 2. Then, we have shown that when the channel has a unique state, the error exponent of the obtained upper bound coincides with the well-known random coding error exponent. We also have shown that when any action cannot affect any states, the error exponent of the obtained upper bound coincides with the error exponent of Somekh-Baruch and Merhav [10, $\tilde{E}_2(R)$]. As for a future research, we will investigate a lower bound of the error probability, and compare it with our obtained upper bound. We will also compute the error exponent for Gaussian channels.

References

[1] A. El Gamal and Y.H. Kim, Network Information Theory, Cambridge University Press, 2012.
[2] T. Weissman, "Capacity of channels with action-dependent states," IEEE Trans. Inf. Theory, vol.56, no.11, pp.5396–5411, Nov. 2010.
[3] H. Asnani, H. Permuter, and T. Weissman, "Probing capacity," IEEE Trans. Inf. Theory, vol.57, no.11, pp.7317–7332, Nov. 2011.
[4] K. Kittichokechai, T. Oechtering, and M. Skoglund, "On the capacity of a channel with action-dependent state and reversible input," Proc. IEEE Int. Symp. on Inform. Theory, pp.331–335, July–Aug. 2011.
[5] C. Choudhuri and U. Mitra, "Action dependent strictly causal state communication," Proc. IEEE Int. Symp. on Inform. Theory, pp.3058–3062, July 2012.
[6] B. Ahmadi and O. Simeone, "On channels with action-dependent states," Arxiv preprint arXiv:1202.4438, July 2012.
[7] B. Hughes and T. Thomas, "On error exponents for arbitrarily varying channels," IEEE Trans. Inf. Theory, vol.42, no.1, pp.87–98, Jan. 1996.
[8] R. Ahlswede, "Arbitrarily varying channels with states sequence known to the sender," IEEE Trans. Inf. Theory, vol.32, no.5, pp.621–629, Sept. 1986.

[9] M. Haroutunian and S. Tonoyan, "Random coding bound of information hiding e-capacity," Proc. IEEE Int. Symp. on Inform. Theory, pp.536–536, June–July 2004.
[10] A. Somekh-Baruch and N. Merhav, "On the random coding error exponents of the single-user and the multiple-access Gel'fand-Pinsker channels," Proc. IEEE Int. Symp. on Inform. Theory, pp.448–448, June–July 2004.
[11] P. Moulin and Y. Wang, "Capacity and random-coding exponents for channel coding with side information," IEEE Trans. Inf. Theory, vol.53, no.4, pp.1326–1347, April 2007.
[12] I. Csiszár and J. Körner, Information Theory: Coding Theorems for Discrete Memoryless Systems, 2nd ed., Cambridge University Press, 2011.

Tetsunao Matsuta was born in Fukui, Japan, on March 17, 1985. He received the B.E. degree from Utsunomiya University in 2007. He received the M.E. degree, and the Dr.Eng. degree from Tokyo Institute of Technology, Japan, in 2009, 2012, respectively. From 2012 to 2013, he was an assistant professor in the Department of Communications and Integrated Systems of Tokyo Institute of Technology. He is currently an assistant professor in the Department of Communications and Computer Engineering of Tokyo Institute of Technology. His current research interests are in the area of multi-terminal information theory.

Tomohiko Uyematsu received the B.E., M.E. and Dr.Eng. degrees from Tokyo Institute of Technology in 1982, 1984 and 1988, respectively. From 1984 to 1992, he was with the Department of Electrical and Electronic Engineering of Tokyo Institute of Technology, first as research associate, next as lecturer, and lastly as associate professor. From 1992 to 1997, he was with School of Information Science of Japan Advanced Institute of Science and Technology as associate professor. Since 1997, he returned to Tokyo Institute of Technology as associate professor, and currently he is with the Department of Communications and Computer Engineering as professor. In 1992 and 1996, he was a visiting researcher at the Centre National de la Recherche Scientifique, France and Delft University of Technology, Netherlands, respectively. He received the Achievement Award in 2008, and the Best Paper Award in 1993, 1996, 2002, 2007 and 2011 both from IEICE. His current research interests are in the areas of information theory, especially Shannon theory and multi-terminal information theory. Dr. Uyematsu is a senior member of IEEE.