Proceedings Paper:

Liu, W., Chu, X. and Zhang, J. (2015) On a generalised typicality with respect to general probability distributions. In: 2015 IEEE 14th Canadian Workshop on Information Theory (CWIT). 2015 IEEE 14th Canadian Workshop on Information Theory (CWIT), 06-09 Jul 2015, St. John's, NL. IEEE, pp. 165-169.

https://doi.org/10.1109/CWIT.2015.7255177

On a Generalised Typicality with Respect to General Probability Distributions

Wuling Liu Xiaoli Chu Jie Zhang
Communications Research Group
Department of Electronic & Electrical Engineering
The University of Sheffield, Sheffield, UK
wuling.liu@sheffield.ac.uk *x.chu@sheffield.ac.uk* *jie.zhang@sheffield.ac.uk*

***Abstract*—The method of typical sequences is a fundamental tool in asymptotic analyses of information theory. The conditional typicality lemma is one of the most commonly used lemmas in the method of typical sequences. Recent works have generalised the definition of typicality to general alphabets or general probability distributions. However, there is still a lack of the conditional typicality lemma based on the definition of typicality with respect to general distributions on the product space. In this paper, we propose a generalised joint typicality for general alphabets and with respect to general probability distributions, and obtain the counterpart of conventional conditional and joint typicality lemmas based on the generalised typicality. As applications of the typicality lemmas, we prove the packing and coverings for the proposed generalised typicality, and then recover the direct part of the capacity theorem on the general Gelfand-Pinsker coding. We also prove a mutual covering lemma for the generalised typicality, and then obtain the Marton-type inner bound to the capacity region of the general broadcast channel.**

I. Introduction

The widely used method of *typical sequences* and the related *asymptotic equipartition property* (AEP) both originated in Shannon's work [1]. The lemmas based on typical sequences are employed in asymptotic analyses of various problems, including point-to-point and multi-terminal ones (cf. [2]).

There exist various definitions of typicality in current information theory. *Strong typicality* is defined by the total variation distance between the empirical distribution and a given probability distribution (cf. [3, Section 10.6]). A variation of strong typicality named as *robust typicality* was proposed in [4] (cf. [2]). Distinct from strong typicality, *weak typicality* is defined by the difference between the empirical entropy and the entropy of a given probability distribution (cf. [3, Section 3.1]). Fundamental properties of both weak and strong typicality are rooted in the weak law of large numbers (LLN).

The *conditional typicality lemma* is one of the commonly used lemmas in the method of typical sequences. It was formally proposed in [2], but had actually been implied in El Gamal and van der Meulen's alternative proof [5] of Marton's inner bound on the capacity region of the broadcast channel (BC) [6]. In addition to Martons inner bound, the conditional typicality lemma can be applied to derive, e.g., the Gelfand-Pinsker theorem [7].

However, the conventional conditional typicality lemma is based on strong typicality, which is only defined for discrete probability distributions. On the other hand, though weak typicality can be defined for general code alphabets and general probability distributions, no typicality lemma based on weak typicality has been proposed. Hence, the application of the conditional typicality lemma is restricted to discrete source/channel coding problems.

Recently, several approaches have been proposed to extend the conditional typicality lemma to more general scenarios. In [2], the authors proposed a discretisation and approximation method to extend the discrete source/channel coding theorems obtained based on strong typicality lemmas to general ones.Typicality defined on general alphabets or spaces has attracted a lot of interest, because the discretisation and approximation process can be circumvented. In [8], Mitran proposed the weak* typicality based on the weak* topology of a Polish space, and obtained the achievable rate of the Gelfand-Pinsker problem using a counterpart of the conventional Markov lemma (cf. [2, Lemma 12.1]). In [9], Raginsky proposed a generalised strong typicality based on a class of measurable functions of the standard Borel space, and generalised the Piggyback coding lemma [10, Lemma 4.3] to his scenario. In [11], Jeon defined a generalised typicality on the measurable space by introducing an abstract typicality criteria determined by an integrable function, then he generalised a series of conventional typicality-based lemmas, such as conditional typicality lemma, joint typicality lemma, packing lemma, covering lemma, and mutual covering lemma.

In most existing works on typicality, typical sequences are defined with respect to a given probability distribution of the one-shot random source/channel symbol, or equivalently, a product probability measure. When the sources and channels are defined by general stochastic processes, the definition of typicality needs to be generalised to a general distribution of the random code sequence which is a stochastic process, or equivalently, a general joint distribution on the product space. Historically, the general AEP, also known as the Shannon-McMillan-Breiman theorem, was proposed for ergodic processes (cf. [3, Section 16.8]). Recently, in [12], Huang has proposed Supremus typicality with respect to irreducible Markov chains. Jeon [11] has proposed a typicality with respect to general distributions by introducing a typicality criteria based on a finite collection of integrable functions. However, to the best of our knowledge, there exists no conditional typicality lemma or its stronger versions such as the Markov lemma based on typicality with respect to general distributions.

In this paper, we propose a generalised typicality for both general alphabets and general probability distributions, as well as to obtain the counterpart of conventional conditional typicality lemma and its corollaries for the generalised typicality. We first prove that a sequence of sets with probability approaching one has properties analogous to the conventional conditional typicality lemma. We then define a new joint typicality by the difference between the *information density* (cf. [13, pp. 10] and [14, Def. 14.1]) and the *spectral mutual information rate* (cf. [15]), and show that a sequence of the proposed typical sets with increasing codelengths is a sequence of sets with probability approaching one. Hence, we are able to obtain the conditional typicality lemma and the joint typicality lemma for our proposed generalised typicality. Furthermore, based on the generalised typicality lemmas, we prove the packing and covering lemmas for the proposed typicality and recover the direct part of the capacity theorem on the general Gelfand-Pinsker coding presented in [16], and we also prove the mutual covering lemma for the proposed typicality and obtain the Marton-type inner bound of the general BC without a discretisation or approximation process.

The rest of this paper is organised as follows. In Section II, we introduce fundamental notations and definitions of the proposed generalised typicality. In Section III, we study the sequence of sets with probability approaching one, and obtain conditional and joint typicality lemmas for the generalised typicality. In Section IV, we recover the direct part of the capacity of the general Gelfand-Pinsker coding. In Section V, we obtain the Marton-type inner bound of the general BC. Finally, conclusion is drawn in Section VI.

II. Notations and Definitions

In this paper, we consider general alphabets and probability distributions represented by a probability space $(\mathcal{X}, \mathscr{B}_{\mathcal{X}}, P_X)$. We borrow notations including $X^n = (X_1^{(n)}, X_2^{(n)}, \cdots, X_n^{(n)})$ and $\mathbf{X} = \{X^n\}_{n=1}^{\infty}$ from [15]. $\mathbf{P}\{\cdot\}$ denotes the probability of an event. $P_{\mathbf{X}}$ denotes the infinite sequence $\{P_{X^n}\}_{n=1}^{\infty}$ of probability distributions. We denote joint probability distribution as P_{XY} and $P_{\mathbf{XY}} = \{P_{X^nY^n}\}_{n=1}^{\infty}$ and transition probability distribution as $P_{Y|X}$ and $P_{\mathbf{Y|X}} = \{P_{Y^n|X^n}\}_{n=1}^{\infty}$. All logarithms are taken to the base e. $P_{X^nY^n}$ is not necessarily to be a product probability distribution denoted as $P_{X^n} \times P_{Y^n}$. A channel between X^n and Y^n is mathematically denoted by $P_{Y^n|X^n}$. $\mathbb{Z}^+$ denotes the set of all positive integers.

Inspired by the information-spectrum approach [15], which is especially effective in dealing with general source/channel coding problems, and also inspired by [14, Remark 14.1], which interprets a pair satisfying the condition $i(x, y) \geq \gamma$ as joint typicality, we propose a generalised definition of joint typicality as follows.

Definition 1 (Sup- and Inf-Typicality)**.** *A sequence pair (x^n, y^n) is called* jointly ϵ-inf-typical *with respect to a general probability distribution $P_{X^nY^n}$ if*

$$|\frac{1}{n} i_{X^nY^n}(x^n, y^n) - \underline{I}(\mathbf{X}; \mathbf{Y})| \leq \epsilon,$$

where $i_{X^nY^n}(x^n, y^n)$ is the information density *(see [13, pp. 10] and [14, Def. 14.1]) defined as*

$$i_{X^nY^n}(x^n, y^n) = \log \frac{dP_{Y^n|X^n}(\cdot|x^n)}{dP_{Y^n}}(y^n)$$

when $P_{Y^n|X^n}(\cdot|x^n)$ is absolutely continuous with respect to P_{Y^n}[1]*, and the* spectral inf-mutual information rate $\underline{I}(X^n; Y^n)$ *(see [15, Def. 3.2.1]) is defined as the limit inferior in probability*[2] *of $\frac{1}{n} i_{X^nY^n}(x^n, y^n)$.*

Similarly, a sequence pair (x^n, y^n) is called jointly ϵ-sup-typical *with respect to $P_{X^nY^n}$ if*

$$|\frac{1}{n} i_{X^nY^n}(x^n, y^n) - \overline{I}(\mathbf{X}; \mathbf{Y})| \leq \epsilon,$$

where the spectral sup-mutual information rate $\underline{I}(X^n; Y^n)$ *(see [15, Def. 3.5.2]) is defined as the limit superior in probability of $\frac{1}{n} i_{X^nY^n}(x^n, y^n)$.*

Let $\underline{\mathcal{T}}_{\epsilon}^{X^nY^n}$ and $\overline{\mathcal{T}}_{\epsilon}^{X^nY^n}$ denote the set of all general jointly ϵ-inf-typical sequences and the set of all general jointly ϵ-sup-typical sequences, respectively.

Remark 1. *Different from [8] and [9], the proposed generalised typicality in Definition 1 is not based on the measure of a Polish or a Borel space, which introduces a metric. We only take measurable spaces as alphabets in this paper.*

Definition 2 (Conditionally Sup- and Inf-Typical)**.** *The conditionally sup- and inf-typical sets $\overline{\mathcal{T}}_{\epsilon}^{Y^n|x^n}$ and $\underline{\mathcal{T}}_{\epsilon}^{Y^n|x^n}$ with respect to a general joint probability distribution $P_{X^nY^n}$ and a given sequence x^n are defined as*

$$\overline{\mathcal{T}}_{\epsilon}^{Y^n|x^n} = \{y^n \in \mathcal{Y}^n | (x^n, y^n) \in \overline{\mathcal{T}}_{\epsilon}^{X^nY^n}\},$$
$$\underline{\mathcal{T}}_{\epsilon}^{Y^n|x^n} = \{y^n \in \mathcal{Y}^n | (x^n, y^n) \in \underline{\mathcal{T}}_{\epsilon}^{X^nY^n}\}.$$

III. Typicality Lemmas

In this section, we will study a sequence of sets with probability approaching one.

Lemma 1. *Given a general $P_{\mathbf{XY}}$, we denote $\{\mathcal{A}^{X^nY^n}\}_{n=1}^{\infty}$ as a sequence of sets satisfying the following condition*

$$\lim_{n\to\infty} P_{X^nY^n}(\mathcal{A}^{X^nY^n}) = 1, \tag{1}$$

where $\mathcal{A}^{X^nY^n} \subset \mathcal{X}^n \times \mathcal{Y}^n$ is $P_{X^nY^n}$-measurable for all $n \in \mathbb{Z}^+$. Let

$$\mathcal{A}^{Y^n|x^n} = \{y^n \in \mathcal{Y}^n | (x^n, y^n) \in \mathcal{A}^{X^nY^n}\}$$

and

$$\mathcal{A}^{X^n|Y^n} = \{x^n \in \mathcal{X}^n | P_{Y^n|X^n}(\mathcal{A}^{Y^n|x^n}|x^n) > 0\}.$$

Then, $\{\mathcal{A}^{X^nY^n}\}_{n=1}^{\infty}$ has the following properties

[1] For a given channel $P_{Y^n|X^n}$, the absolutely continuous condition is always satisfied because P_{Y^n} is determined by an input distribution P_{X^n} and the transition distribution $P_{Y^n|X^n}$.

[2] The limit inferior in probability of $\mathbf{X}$ is defined as $\text{p-}\liminf_{n\to\infty} X^n = \sup\{\beta | \lim_{n\to\infty} \mathbf{P}\{X^n < \beta\} = 0\}$, and the limit superior in probability of $\mathbf{X}$ is defined as $\text{p-}\limsup_{n\to\infty} X^n = \inf\{\alpha | \lim_{n\to\infty} \mathbf{P}\{X^n > \alpha\} = 0\}$

1)

$$\lim_{n\to\infty} P_{X^n}(\mathcal{A}^{X^n|Y^n}) = 1;$$

2)

$$\lim_{n\to\infty} P_{Y^n|X^n}(\mathcal{A}^{Y^n|x^n}|x^n) = 1$$

for any $\{x^n\}_{n=1}^{\infty}$*, where* $x^n \in \mathcal{A}^{X^n|Y^n}$ *for all* $n \in \mathbb{Z}^+$*.*

Proof: We have

$$\begin{aligned} P_{X^nY^n}(\mathcal{A}^{X^nY^n}) &= \int_{\mathcal{X}^n} P_{Y^n|X^n}(\mathcal{A}^{Y^n|x^n}|x^n)dP_{X^n}(x^n) \\ &= \int_{\mathcal{A}^{X^n|Y^n}} P_{Y^n|X^n}(\mathcal{A}^{Y^n|x^n}|x^n)dP_{X^n}(x^n) \quad (2) \\ &\leq \int_{\mathcal{A}^{X^n|Y^n}} dP_{X^n}(x^n) \\ &= P_{X^n}(\mathcal{A}^{X^n|Y^n}) \quad (3) \end{aligned}$$

for all $n \in \mathbb{Z}^+$. From (1) and (3), we obtain

$$\lim_{n\to\infty} P_{X^n}(\mathcal{A}^{X^n|Y^n}) = 1.$$

Assume that there exists a sequence $\{\tilde{x}^n\}_{n=1}^{\infty}$, where $\tilde{x}^n \in \mathcal{A}^{X^n|Y^n}$ for all $n \in \mathbb{Z}^+$, such that

$$\liminf_{n\to\infty} P_{Y^n|X^n}(\mathcal{A}^{Y^n|\tilde{x}^n}|\tilde{x}^n) < 1.$$

Then there exists $\{n_k\}_{k=1}^{\infty} \subset \mathbb{Z}^+$, where $n_i < n_j$ if $i < j$, such that

$$\begin{aligned} &\lim_{k\to\infty} P_{Y^{n_k}|X^{n_k}}(\mathcal{A}^{Y^{n_k}|\tilde{x}^{n_k}}|\tilde{x}^{n_k}) < 1 \\ \Rightarrow &\exists k_0 \in \mathbb{Z}^+, \forall k \geq k_0, P_{Y^{n_k}|X^{n_k}}(\mathcal{A}^{Y^{n_k}|\tilde{x}^{n_k}}|\tilde{x}^{n_k}) < 1, \quad (4) \\ \Rightarrow &\exists k_0 \in \mathbb{Z}^+, \forall k \geq k_0, \\ &P_{X^nY^n}(\mathcal{A}^{X^nY^n}) < \int_{\mathcal{X}^n} dP_{X^{n_k}}(x^{n_k}) = 1, \end{aligned}$$

where (4) follows from (2). Hence,

$$\liminf_{n\to\infty} P_{X^nY^n}(\mathcal{A}^{X^nY^n}) < 1. \quad (5)$$

We notice that the inequality in (5) contradicts equation (1). Since $P_{Y^n|X^n}(\mathcal{A}^{Y^n|\tilde{x}^n}|\tilde{x}^n) \leq 1$, for any $\{x^n\}_{n=1}^{\infty}$, where $x^n \in \mathcal{A}^{X^n|Y^n}$ for all $n \in \mathbb{Z}^+$, we have

$$\lim_{n\to\infty} P_{Y^n|X^n}(\mathcal{A}^{Y^n|x^n}|x^n) = 1.$$

■

Remark 2. *In Lemma 2, we will set a sequence of the proposed generalised typical sets and prove that this sequence satisfies condition* (1)*. Besides, in [17], Somekh-Baruch proposed a sequence* $\{\mathcal{A}_n\}_{n=1}^{\infty}$ *satisfying condition* (1)*.*

Remark 3. $x^n \in \mathcal{A}^{X^n|Y^n}$ *is necessary, as shown in the Appendix.*

From Lemma 1, we will obtain conditional and joint typicality lemmas for the proposed generalised typicality.

Lemma 2 (Conditional Typicality)**.** *Given* $P_{\mathbf{XY}}$*, we set*

$$\begin{aligned} \underline{\mathcal{T}}_{\epsilon}^{X^n|Y^n} &= \{x^n \in \mathcal{X}^n | P_{Y^n|X^n}(\underline{\mathcal{T}}_{\epsilon}^{Y^n|x^n}|x^n) > 0\}, \\ \overline{\mathcal{T}}_{\epsilon}^{X^n|Y^n} &= \{x^n \in \mathcal{X}^n | P_{Y^n|X^n}(\overline{\mathcal{T}}_{\epsilon}^{Y^n|x^n}|x^n) > 0\}. \end{aligned}$$

Then

1)

$$\lim_{n\to\infty} P_{X^n}(\underline{\mathcal{T}}_{\epsilon}^{X^n|Y^n}) = \lim_{n\to\infty} P_{X^n}(\overline{\mathcal{T}}_{\epsilon}^{X^n|Y^n}) = 1;$$

2)

$$\lim_{n\to\infty} P_{Y^n|X^n}(\underline{\mathcal{T}}_{\epsilon}^{Y^n|x^n}|x^n) = 1,$$

for any $\{x^n\}_{n=1}^{\infty}$*, where* $x^n \in \underline{\mathcal{T}}_{\epsilon}^{X^n|Y^n}$ *for all* $n \in \mathbb{Z}^+$*, and*

$$\lim_{n\to\infty} P_{Y^n|X^n}(\overline{\mathcal{T}}_{\epsilon}^{Y^n|x^n}|x^n) = 1,$$

for any $\{x^n\}_{n=1}^{\infty}$*, where* $x^n \in \overline{\mathcal{T}}_{\epsilon}^{X^n|Y^n}$ *for all* $n \in \mathbb{Z}^+$*.*

Proof: $\underline{\mathcal{T}}_{\epsilon}^{X^nY^n}$ and $\overline{\mathcal{T}}_{\epsilon}^{X^nY^n}$ are $P_{X^nY^n}$-measurable because $i_{X^nY^n}(x^n, y^n)$ is a measurable function. From the definition of $\underline{I}(\mathbf{X};\mathbf{Y})$ and the limit inferior in probability (see [15]), we know that for every $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbf{P}\{\frac{1}{n} i_{X^nY^n}(x^n, y^n) > \underline{I}(\mathbf{X};\mathbf{Y}) - \epsilon\} = 1.$$

Thus $\{\underline{\mathcal{T}}_{\epsilon}^{X^n|Y^n}\}_{n=1}^{\infty}$ satisfies condition (1) in Lemma 1. So does $\overline{\mathcal{T}}_{\epsilon}^{X^n|Y^n}$. ■

Remark 4. *Under the condition of Lemma 2, if given an* x^n *and let* $Y^n \sim P_{Y^n|X^n}(\cdot|x^n)$*, then we have*

$$\begin{aligned} P_{Y^n|X^n}(\underline{\mathcal{T}}_{\epsilon}^{Y^n|x^n}|x^n) &= \mathbf{P}\{(x^n, Y^n) \in \underline{\mathcal{T}}_{\epsilon}^{X^nY^n}\}, \\ P_{Y^n|X^n}(\overline{\mathcal{T}}_{\epsilon}^{Y^n|x^n}|x^n), &= \mathbf{P}\{(x^n, Y^n) \in \overline{\mathcal{T}}_{\epsilon}^{X^nY^n}\}. \end{aligned}$$

Lemma 3 (Joint Inf-Typicality)**.** *Give* $P_{\mathbf{XY}}$*, for all* $n \in \mathbb{Z}^+$*,*

1) for all $x^n \in \underline{\mathcal{T}}_{\epsilon}^{X^n|Y^n}$*,*

$$e^{-n(\underline{I}(\mathbf{X};\mathbf{Y})+\epsilon)} \leq P_{Y^n}(\underline{\mathcal{T}}_{\epsilon}^{Y^n|x^n}) \leq e^{-n(\underline{I}(\mathbf{X};\mathbf{Y})-\epsilon)};$$

2)

$$e^{-n(\underline{I}(\mathbf{X};\mathbf{Y})+\epsilon)} \leq (P_{X^n} \times P_{Y^n})(\underline{\mathcal{T}}_{\epsilon}^{X^nY^n}) \leq e^{-n(\underline{I}(\mathbf{X};\mathbf{Y})-\epsilon)}.$$

Proof:

1) From the definition of $\underline{\mathcal{T}}_{\epsilon}^{X^nY^n}$, we have

$$\begin{aligned} &P_{Y^n}(\underline{\mathcal{T}}_{\epsilon}^{Y^n|x^n}) \\ &= \int_{\underline{\mathcal{T}}_{\epsilon}^{Y^n|x^n}} dP_{Y^n}(\tilde{y}^n) \\ &= \int_{\underline{\mathcal{T}}_{\epsilon}^{Y^n|x^n}} \frac{dP_{Y^n}}{dP_{Y^n|X^n}(\cdot|x^n)}(\tilde{y}^n) dP_{Y^n|X^n}(\tilde{y}^n|x^n) \\ &= \int_{\underline{\mathcal{T}}_{\epsilon}^{Y^n|x^n}} e^{-i_{X^nY^n}(x^n,\tilde{y}^n)} dP_{Y^n|X^n}(\tilde{y}^n|x^n). \end{aligned}$$

Hence,

$$e^{-n(\underline{I}(\mathbf{X};\mathbf{Y})+\epsilon)} \leq P_{Y^n}(\underline{\mathcal{T}}_{\epsilon}^{Y^n|x^n}) \leq e^{-n(\underline{I}(\mathbf{X};\mathbf{Y})-\epsilon)}.$$

2) Part 2) is a corollary of part 1), i.e.,

$$(P_{X^n} \times P_{Y^n})(\underline{\mathcal{T}}_\epsilon^{X^nY^n}) = \int_{\underline{\mathcal{T}}_\epsilon^{X^n|Y^n}} P_{Y^n}(\underline{\mathcal{T}}_\epsilon^{Y^n|x^n}) dP_{X^n}(x^n),$$

which establishes part 2).

■

Remark 5. *Under the condition of Lemma 3, if given* $\bar{Y}^n \sim P_{\bar{Y}^n|X^n}(\cdot|x^n)$ *where* $P_{\bar{Y}^n|X^n}(\cdot|x^n) = P_{Y^n}(\cdot)$ *and* $(\tilde{X}^n, \tilde{Y}^n) \sim P_{\tilde{X}^n\tilde{Y}^n}$ *where* $P_{\tilde{X}^n\tilde{Y}^n} = P_{X^n} \times P_{Y^n}$*, we have*

$$P_{Y^n}(\underline{\mathcal{T}}_\epsilon^{Y^n|x^n}) = \mathbf{P}\{(x^n, \bar{Y}^n) \in \underline{\mathcal{T}}_\epsilon^{X^nY^n}\},$$
$$(P_{X^n} \times P_{Y^n})(\underline{\mathcal{T}}_\epsilon^{X^nY^n}) = \mathbf{P}\{(\tilde{X}^n, \tilde{Y}^n) \in \underline{\mathcal{T}}_\epsilon^{X^nY^n}\}.$$

Similarly, we obtain the joint sup-typicality lemma.

Lemma 4 (Joint Sup-Typicality)**.** *Given* $P_{\mathbf{XY}}$*, for all* $n \in \mathbb{Z}^+$*,*

1) for all $x^n \in \overline{\mathcal{T}}_\epsilon^{X^n|Y^n}$*,*

$$e^{-n(\overline{I}(\mathbf{X};\mathbf{Y})+\epsilon)} \le P_{Y^n}(\overline{\mathcal{T}}_\epsilon^{Y^n|x^n}) \le e^{-n(\overline{I}(\mathbf{X};\mathbf{Y})-\epsilon)};$$

2)

$$e^{-n(\overline{I}(\mathbf{X};\mathbf{Y})+\epsilon)} \le (P_{X^n} \times P_{Y^n})(\overline{\mathcal{T}}_\epsilon^{X^nY^n}) \le e^{-n(\overline{I}(\mathbf{X};\mathbf{Y})-\epsilon)}.$$

IV. Application to Gelfand-Pinsker Coding

Gelfand-Pinsker (GP) coding problem [7] is a channel coding problem in which a channel state is noncausally available at the encoder. We restate the general GP coding problem [16] as follows. The channel is defined by the input $\mathbf{X}$, the output $\mathbf{Y}$, the general state $\mathbf{S} \sim P_{\mathbf{S}}$, and the transition probability $P_{\mathbf{Y}|\mathbf{SX}}$. For a fixed codelenth n, the encoder f is a mapping from $\mathcal{M} \times \mathcal{S}^n$ to $\mathcal{X}^n$ where $\mathcal{M}$ is the message set and $\mathcal{S}$ is the state space, and the decoder g is a mapping from $\mathcal{Y}^n$ to $\mathcal{M}$. The average error probability ϵ_n is the average probability of the event that $g(Y^n)$ is not equal to the sent message.

In [16], Tan obtained the capacity of the general channel with general states, which is a generalised GP coding problem [7].

Theorem 1 (Gelfand-Pinsker-Tan)**.** *The capacity of the general channel* $P_{\mathbf{Y}|\mathbf{XS}}$ *with general non-causal state* $\mathbf{S}$ *only available at the encoder is*

$$C = \sup_{P_{\mathbf{UX}} \in \mathcal{P}} \underline{I}(\mathbf{U};\mathbf{Y}) - \overline{I}(\mathbf{U};\mathbf{S}),$$

where $\mathcal{P}$ *is the set of all* $P_{\mathbf{UX}}$*'s satisfying that* $\mathbf{U} \to (\mathbf{X},\mathbf{S}) \to \mathbf{Y}$ *forms a Markov chain.*

In order to prove the achievability of the capacity, Tan employed a modified piggyback coding lemma (PBL) [10, Lemma 4.3] to get around a counterpart of the conditional typical lemma in the case where typicality is defined by the information-spectral quantity. In the following, we will recover the achievability of the capacity of the general GP coding using our proposed conditional typicality lemma (Lemma 2), instead of a lemma analogous to PBL.

A. Packing and Covering Lemmas

As counterparts to the conventional method of typical sequences, we will first prove the following two lemmas which will be used in our proof of the achievability of the capacity in Theorem 1.

Lemma 5 (Packing)**.** *Given* $P_{\mathbf{XY}}$*, for all* $n \in \mathbb{Z}^+$*, let* $\tilde{Y}^n \sim P_{\tilde{Y}^n}$*, which is not necessarily equal to* P_{Y^n}*,* $X^n(m) \sim P_{X^n}, m \in \mathcal{M}_n$ *with* $|\mathcal{M}_n| = e^{nR}$*, and* $X^n(m)$ *are independent of* $\tilde{Y}^n$ *for all* $m \in \mathcal{M}_n$*. If* $R < \underline{I}(\mathbf{X};\mathbf{Y})$*, then there exists an* $\epsilon > 0$ *such that*

$$\lim_{n\to\infty} \mathbf{P}\left(\bigcup_{m\in\mathcal{M}_n} \{(X^n(m), \tilde{Y}^n) \in \underline{\mathcal{T}}_\epsilon^{X^nY^n}\}\right) = 0.$$

Proof: From the union bound and the joint inf-typicality lemma (Lemma 3), we have

$$\begin{aligned}
&\mathbf{P}\left(\bigcup_{m\in\mathcal{M}_n} \{(X^n(m), \tilde{Y}^n) \in \underline{\mathcal{T}}_\epsilon^{X^nY^n}\}\right)\\
=&\sum_{m\in\mathcal{M}_n} \mathbf{P}\{(X^n(m), \tilde{Y}^n) \in \underline{\mathcal{T}}_\epsilon^{X^nY^n}\}\\
\le&\sum_{m\in\mathcal{M}_n} \int_{\underline{\mathcal{T}}_\epsilon^{Y^n|X^n}} \mathbf{P}\{(X^n(m), \tilde{y}^n) \in \underline{\mathcal{T}}_\epsilon^{X^nY^n}\} dP_{\tilde{Y}^n}(\tilde{y}^n)\\
\le&|\mathcal{M}_n| e^{-n(\underline{I}(\mathbf{X};\mathbf{Y})-\epsilon)}\\
\le&e^{n(R-\underline{I}(\mathbf{X};\mathbf{Y})+\epsilon)},
\end{aligned}$$

which establishes the lemma. ■

Lemma 6 (Covering)**.** *Given* $P_{\mathbf{XY}}$*, for all* $n \in \mathbb{Z}^+$*, let* $X^n \sim P_{X^n}$*,* $Y^n(m) \sim P_{Y^n}, m \in \mathcal{M}_n$ *with* $|\mathcal{M}_n| = e^{nR}$*, and* X^n *and* $Y^n(m)$*'s are independent of each other. If* $R > \overline{I}(\mathbf{X};\mathbf{Y})$*, then there exists an* ϵ *such that*

$$\lim_{n\to\infty} \mathbf{P}\left(\bigcap_{m\in\mathcal{M}_n} \{(X^n, Y^n(m)) \notin \overline{\mathcal{T}}_\epsilon^{X^nY^n}\}\right) = 0.$$

Proof: From the joint sup-typicality lemma (Lemma 4) and the inequality $(1-y)^n \le e^{-yn}$ for $0 \le y \le 1$ and $n \ge 0$ [3, Lemma 10.5.3], we have

$$\begin{aligned}
&\mathbf{P}\left(\bigcap_{m\in\mathcal{M}_n} \{(X^n, Y^n(m)) \notin \overline{\mathcal{T}}_\epsilon^{X^nY^n}\}\right)\\
\le&(1 - e^{-n(\overline{I}(\mathbf{X};\mathbf{Y})+\epsilon)})^{|\mathcal{M}_n|}\\
\le&\exp(-e^{n(R-\overline{I}(\mathbf{X};\mathbf{Y})-\epsilon)}),
\end{aligned}$$

which establishes the lemma. ■

Remark 6. *It is evident that we can obtain a packing lemma based on joint sup-typicality and a covering lemma based on joint inf-typicality. For this paper, Lemmas 5 and 6 is sufficient.*

B. Proof of the Achievability of Theorem 1

Analogous to the proof of [2, Theorem 7.3], we employ the random coding and the typicality decoding techniques to prove the achievability of Theorem 1. The difference is that we employ the general information-density-based definition of typicality in the decoding metrics and the error probability analysis.

Random codebook generation: For fixed $P_{\mathbf{U}|\mathbf{S}}, P_{\mathbf{X}|\mathbf{US}}$ and let $P_{\mathbf{USXY}}$ be determined by $P_{\mathbf{S}}$ and the transition probability distributions $P_{\mathbf{U}|\mathbf{S}}, P_{\mathbf{X}|\mathbf{US}}$ and $P_{\mathbf{Y}|\mathbf{USX}} = P_{\mathbf{Y}|\mathbf{SX}}$. For $\tilde{R} > R$, a fixed codelength n and each $m \in \mathcal{M}$, randomly and independently generate $e^{n(\tilde{R}-R)}$ $u^n(l)$'s according to P_{U^n}, where l's are indices of the sequences. For each $u^n(l)$ and s^n, randomly and independently generate an $x^n(s^n, l)$ according to $P_{X^n|U^nS^n}(\cdot|u^n(l), s^n)$.

Encoding: Assume that a specified message M is sent. Choose a $u^n(L)$ from $u^n(l)$'s corresponding to M such that $(u^n(L), s^n) \in \overline{\mathcal{T}}_\epsilon^{U^nS^n} \cap \underline{\mathcal{T}}_\epsilon^{U^n|Y^n}$, where $P_{U^nY^n}$ is the marginal of $P_{U^nS^nX^nY^n}$, and then the index L is specified. Send the corresponding $x^n(s^n, L)$.

Decoding: Assume that y^n is received. The decoding output returns that $\hat{m}$ is sent if there exists an $u^n(\hat{l})$ satisfying $(u^n(\hat{l}), y^n) \in \underline{\mathcal{T}}_\epsilon^{U^nY^n}$, where $u^n(\hat{l})$ corresponds to $\hat{m}$.

Error probability analysis:

$$\epsilon_n \leq \mathbf{P}(\mathcal{E}_1) + \mathbf{P}(\mathcal{E}_1^{\complement} \cap \mathcal{E}_2) + \mathbf{P}(\mathcal{E}_3),$$

where the error events are $\mathcal{E}_1$:$(U^n(l), S^n) \notin \overline{\mathcal{T}}_\epsilon^{U^nS^n} \cap \underline{\mathcal{T}}_\epsilon^{U^n|Y^n}$ for all $U^n(l)$ corresponding to M, $\mathcal{E}_2$:$(U^n(L), Y^n) \notin \underline{\mathcal{T}}_\epsilon^{U^nY^n}$, $\mathcal{E}_3$:$(U^n(l), Y^n) \in \overline{\mathcal{T}}_\epsilon^{U^nY^n}$ for some $U^n(l)$ corresponding to $m \neq M$.

From the covering lemma and part 1) of the conditional typicality lemma, $\lim_{n\to\infty} \mathbf{P}(\mathcal{E}_1) = 0$ if $\tilde{R} - R > \overline{I}(\mathbf{U}; \mathbf{S})$; from part 2) of the conditional typicality lemma, $\lim_{n\to\infty} \mathbf{P}(\mathcal{E}_1^{\complement} \cap \mathcal{E}_2) = 0$; and from the packing lemma, $\lim_{n\to\infty} \mathbf{P}(\mathcal{E}_3) = 0$ if $\tilde{R} < \underline{I}(\mathbf{U}; \mathbf{Y})$. The achievability is established.

V. Application to the Broadcast Channel

In this section, the general BC is defined by the input $\mathbf{X}$, the outputs $\mathbf{Y}_1, \mathbf{Y}_2$, and the transition probability $P_{\mathbf{Y}_1\mathbf{Y}_2|\mathbf{X}}$. For a fixed codelenth n, the encoder f_j is a mapping from $\mathcal{M}_1 \times \mathcal{M}_2$ to $\mathcal{X}^n$ where $\mathcal{M}_1, \mathcal{M}_2$ are the message sets corresponding to the j-th receiver, and the decoder g_j is a mapping from $\mathcal{Y}_j^n$ to $\mathcal{M}_j$ ($j = 1, 2$). The average error probability ϵ_n is the average probability of the event that $g_1(Y_1^n)$ or $g_2(Y_2^n)$ is not equal to the sent message.

In [18], the authors obtained a Marton-type inner bound for the non-iid classic-quantum BC with finite alphabets. In this section, we will obtain a Marton-type inner bound for the general BC by employing a mutual covering lemma based on a "rejection sampling" method. Though the capacity region of the general BC was found in [19] using the information-spectrum approach [15], we will directly obtain the single-letter Marton-type inner bound for the BC with general alphabets, without discretisation or approximation.

A. Mutual Covering Lemma

We will first prove the generalised mutual covering lemma.

Lemma 7 (Mutual Covering)**.** *Given $P_{\mathbf{XY}}$, let $X^n(m_1) \sim P_{X^n}, m_1 \in \mathcal{M}_{1n}$ with $|\mathcal{M}_{1n}| = e^{nR_1}$, $Y^n(m_2) \sim P_{X^n}, m_2 \in \mathcal{M}_{2n}$ with $|\mathcal{M}_{2n}| = e^{nR_2}$, $X^n(m_1)$'s are pairwise independent, $Y^n(m_2)$'s are pairwise independent, and $\{X^n(m_1)\}_{m_1 \in \mathcal{M}_{1n}}$ is independent of $\{Y^n(m_2)\}_{m_2 \in \mathcal{M}_{2n}}$. If $R_1 + R_2 > I(X; Y)$, then there exists an $\epsilon > 0$ such that*

$$\lim_{n\to\infty} \mathbf{P}\left(\bigcap_{\substack{(m_1,m_2)\in \\ \mathcal{M}_{1n}\times\mathcal{M}_{2n}}} \{(X^n(m_1), Y^n(m_2)) \notin \overline{\mathcal{T}}_\epsilon^{X^nY^n}\} \right) = 0.$$

Proof: Similar to the proof of the mutual covering lemma based on strong typicality [2, Appendix 8A], following Chebyshev's inequality, we obtain

$$\mathbf{P}\{|\mathcal{M}_n| = 0\} \leq \frac{\mathbf{D}(|\mathcal{A}_n|)}{(\mathbf{E}|\mathcal{M}_n|)^2},$$

with

$$\mathbf{E}(|\mathcal{M}_n|) = e^{n(R_1+R_2)} p_{1n},$$
$$\mathbf{D}(|\mathcal{M}_n|) \leq e^{n(R_1+R_2)} p_{1n} + e^{n(R_1+2R_2)} p_{2n} + e^{n(2R_1+R_2)} p_{3n},$$

where

$$\mathcal{M}_n = \{(m_1, m_2) \in \mathcal{M}_{1n} \times \mathcal{M}_{2n} | (X^n(m_1), Y^n(m_2)) \in \overline{\mathcal{T}}_\epsilon^{X^nY^n}\},$$

and

$$p_{1n} = \mathbf{P}\{(X^n(m_1), Y^n(m_2)) \in \overline{\mathcal{T}}_\epsilon^{X^nY^n}\},$$
$$p_{2n} = \mathbf{P}\{(X^n(m_1), Y^n(m_2)) \in \overline{\mathcal{T}}_\epsilon^{X^nY^n} \wedge (X^n(m_1), Y^n(\tilde{m}_2)) \in \overline{\mathcal{T}}_\epsilon^{X^nY^n}\},$$
$$p_{3n} = \mathbf{P}\{(X^n(m_1), Y^n(m_2)) \in \overline{\mathcal{T}}_\epsilon^{X^nY^n} (X^n(\tilde{m}_1), Y^n(m_2)) \in \overline{\mathcal{T}}_\epsilon^{X^nY^n}\}.$$

From the joint sup-typicality lemma (Lemma 4), we have

$$p_{2n} \leq e^{-n(2\overline{I}(X^n;Y^n)-2\epsilon)}, p_{3n} \leq e^{-n(2\overline{I}(X^n;Y^n)-2\epsilon)},$$

and

$$p_{1n} \geq e^{-n(\overline{I}(X^n;Y^n)+\epsilon)}.$$

Hence,

$$\mathbf{P}\{|\mathcal{M}_n| = 0\} \leq e^{-n(R_1+R_2-\overline{I}(X^n;Y^n)-\epsilon)} + e^{-n(R_1-4\epsilon)} + e^{-n(R_2-4\epsilon)}.$$

which establishes the lemma. ■

Remark 7. *We can obtain a mutual covering lemma based on joint inf-typicality. For this paper, Lemmas 7 is sufficient.*

B. Proof of Marton's Inner Bound for the General BC

Analogous to the achievability proof in [2, Section 8.3], we perform the standard analysis as in Section IV.

Random codebook generation: For fixed $P_{\mathbf{U}_1\mathbf{U}_2\mathbf{X}}$, $\tilde{R}_1 > R_1, \tilde{R}_2 > R_2$ and a fixed codelength n, and each $m_j \in \mathcal{M}_j, j = 1, 2$, randomly and independently generate $e^{n(\tilde{R}_j - R_j)}$ $u_j^n(l_j)$'s according to $P_{U_j^n}$. For each $(u_1^n(l_1), u_2^n(l_2))$, randomly and independently generate an $x^n(l_1, l_2)$ according to $P_{X^n|U_1^nU_2^n}(\cdot|u_1^n(l_1), u_2^n(l_2))$.

Encoding: Assume (M_1, M_2) is sent. Choose a pair $(u_1^n(L_1), u_2^n(L_2))$ from $(u_1^n(l_1), u_2^n(l_2))$'s corresponding to (M_1, M_2) such that $(u_1^n(L_1), u_2^n(L_2) \in \overline{\mathcal{T}}_\epsilon^{U_1^n U_2^n}$ and $u_j^n(L_j) \in \underline{\mathcal{T}}_\epsilon^{U_j^n|Y_j^n}, j = 1, 2$. Then send the corresponding $x^n(L_1, L_2)$.

Decoding: Assume that y_j^n $(j = 1, 2)$ is received at the j-th receiver. The j-th receiver outputs that $\hat{m}_j$ is sent if $(u_j^n(l_j), y_j^n) \in \underline{\mathcal{T}}_\epsilon^{U_j^n Y_j^n}$, where $u_j^n(\hat{l}_j)$ corresponds to $\hat{m}_j$.

Error probability analysis: For $j = 1, 2$,

$$\epsilon_{1n} \leq \mathbf{P}(\mathcal{E}_1) + \mathbf{P}(\mathcal{E}_1^\complement \cap \mathcal{E}_{2j}) + \mathbf{P}(\mathcal{E}_{3j}),$$

where the error events are $\mathcal{E}_1$:$(U_1^n(l_1), U_2^n(l_2)) \notin \overline{\mathcal{T}}_\epsilon^{U_1^n U_2^n}$ or $u_j^n(L_j) \notin \underline{\mathcal{T}}_\epsilon^{U_j^n|Y_j^n}$ $(j = 1, 2)$ for all $(U_1^n(l_1), U_2^n(l_2))$ corresponding to (M_1, M_2), $\mathcal{E}_{2j}$:$(U_j^n(L_j), Y_j^n) \notin \underline{\mathcal{T}}_\epsilon^{U_j^n Y_j^n}$, $\mathcal{E}_{3j}$:$(U_j^n(l_j), Y_j^n) \in \overline{\mathcal{T}}_\epsilon^{U_j^n Y_j^n}$ for some $U_j^n(l_j)$ corresponding to $m_j \neq M_j$.

From part 1) of the conditional typicality lemma and the mutual covering lemma, $\lim_{n\to\infty} \mathbf{P}(\mathcal{E}_1) = 0$ if $\tilde{R}_1 + \tilde{R}_2 - R_1 - R_2 > \overline{I}(\mathbf{U}_1; \mathbf{U}_2)$; from part 2) of the conditional typicality lemma, $\lim_{n\to\infty} \mathbf{P}(\mathcal{E}_1^\complement \cap \mathcal{E}_{2j}) = 0$; and from the packing lemma, $\lim_{n\to\infty} \mathbf{P}(\mathcal{E}_{3j}) = 0$ if $\tilde{R}_j < \underline{I}(\mathbf{U}_j; \mathbf{Y}_j)$. Then using the Fourier-Motzkin elimination [2, Appendix D], we obtain the Marton-type inner bound for the general BC as

$$\mathcal{R} = \bigcup_{P_{\mathbf{U}_1 \mathbf{U}_2 \mathbf{X}}} \left\{ \begin{array}{l} (R_1, R_2) : \\ 0 \leq R_1 \leq \underline{I}(\mathbf{U}_1; \mathbf{Y}_1), \\ 0 \leq R_2 \leq \underline{I}(\mathbf{U}_2; \mathbf{Y}_2), \\ R_1 + R_2 \leq \underline{I}(\mathbf{U}_1; \mathbf{Y}_1) + \underline{I}(\mathbf{U}_2; \mathbf{Y}_2) \\ \qquad\qquad\qquad - \overline{I}(\mathbf{U}_1; \mathbf{U}_2) \end{array} \right\}.$$

VI. Conclusion

We have proposed a generalised typicality for general alphabets and with respect to general probability distributions. By studying a sequence of sets with probability tending to one, we have obtained the conditional and joint typicality lemmas for the proposed generalised typicality. As applications of the proposed typicality lemmas, we have proved the packing, covering and mutual covering lemmas for the proposed generalised typicality, then recovered the direct part of the capacity theorem on the general GP coding, and obtained the Marton-type inner bound of the general BC, without any discretisation or approximation.

Appendix

It is evident that the sequence of robust typicality sets constructed in the conventional conditional typicality lemma [2, Section 2.5] satisfies condition (1) in Lemma 1. In [2, Problem 2.17], it is shown that the conditional typicality lemma is not necessarily established for a given $\{x^n\}_{n=1}^\infty$, where $x^n \in \mathcal{T}_\epsilon^{X^n}$ (for robust typicality) for all $n \in \mathbb{Z}^+$. We can prove that the given x^n actually falls out of $\mathcal{T}_\epsilon^{X^n|Y^n}$ asymptotically.

In the above problem, P_{XY} is given as the production probability distribution of two Bernoulli distributions $\mathrm{B}(1/2)$, and x^n is given as a binary sequence with k_n 1's followed by $(n - k_n)$ 0's, where $k_n = \lfloor (n/2)(1 + \epsilon) \rfloor$. According to the definition of robust jointly typical set, x^n will fall out of $\mathcal{T}_\epsilon^{X^n|Y^n}$ if $\lceil k_n/2 \rceil / n - 1/4 > \epsilon/4$. Let $\{n'_l\}$ be a subsequence of all n's satisfying $n = 4l_n + 1$. Assume that $k_n = (n/2)(1 + \epsilon) - \delta_n$ and $\lceil k_n/2 \rceil = k_n/2 + \gamma_n$. Because $k_n = 2l_n + 1/2 + (2l_n + 1/2)\epsilon - \delta_n$, $(2l_n + 1/2)\epsilon - \delta_n + 1/2$ is an integer denoted by N_{l_n}. Then because $\lceil k_n/2 \rceil = l_n + N_{l_n}/2 + \gamma_n$, $\gamma_n = 1/2$ if N_{l_n} is odd, thus $\lceil k_n/2 \rceil / n - 1/4 = \epsilon/4 + (1 - \delta_n)/(2n) > \epsilon/4$. The range $0 < \epsilon < 1$ implies that there exist infinitely many odd N_{l_n}'s in $\{N_{l_n}\}$. Hence, there exist infinitely many n's such that $x^n \notin \mathcal{T}_\epsilon^{X^n|Y^n}$.

References

[1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3–4, pp. 379–423, 623–656, July–Oct. 1948.

[2] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge: Cambridge University Press, 2011.

[3] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ: Wiley-Interscience, 2006.

[4] A. Orlitsky and J. R. Roche, "Coding for computing," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 903–917, Mar. 2001.

[5] A. El Gamal and E. C. van der Meulen, "A proof of Marton's coding theorem for the discrete memoryless broadcast channel (corresp.)," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 120–122, Jan. 1981.

[6] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 3, pp. 306–311, May 1979.

[7] S. I. Gelfand and M. S. Pinsker, "Coding for channel with random parameters," *Probl. Control Inform. Theory*, vol. 9, no. 1, pp. 19–31, 1980.

[8] P. Mitran, "Typical sequences for Polish alphabets," *CoRR*, vol. abs/1005.2321, 2010. [Online]. Available: http://arxiv.org/abs/1005.2321

[9] M. Raginsky, "Empirical processes, typical sequences, and coordinated actions in standard Borel spaces," *IEEE Trans. Inf. Theory*, vol. 59, no. 3, pp. 1288–1301, Mar. 2013.

[10] A. D. Wyner, "On source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. 21, no. 3, pp. 294–300, May 1975.

[11] J. Jeon, "A generalized typicality for abstract alphabets," *CoRR*, vol. abs/1401.6728, 2014. [Online]. Available: http://arxiv.org/abs/1401.6728

[12] S. Huang and M. Skoglund, "Supremus typicality," in *Proc. IEEE Int. Symp. Inform. Theory*, Honolulu, HI, USA, June 2014, pp. 2644–2648.

[13] M. S. Pinsker, *Information and Information Stability of Random Variables and Processes*, ser. Holden-Day Ser. Time Ser. Anal. San Francisco, CA, USA: Holden-Day, 1964, originally published in Russian in 1960, translated and edited by A. Feinstein.

[14] Y. Polyanskiy and Y. Wu, "Lecture notes on information theory," 2014.

[15] T. S. Han, *Information-Spectrum Methods in Information Theory*, ser. Stoch. Model. Appl. Probab. Berlin: Springer-Verlag, 2003, no. 50, originally published in Japanese in 1998, translated by H. Koga.

[16] V. Y. F. Tan, "A formula for the capacity of the general Gel'fand-Pinsker channel," in *Proc. IEEE Int. Symp. Inform. Theory*, Istanbul, Turkey, July 2013, pp. 2458–2462.

[17] A. Somekh-Baruch, "A general formula for the mismatch capacity," *CoRR*, vol. abs/1309.7964, 2013. [Online]. Available: http://arxiv.org/abs/1309.7964

[18] J. Radhakrishnan, P. Sen, and N. Warsi, "One-shot Marton inner bound for classical-quantum broadcast channel," *CoRR*, vol. abs/1410.3248, 2014. [Online]. Available: http://arxiv.org/abs/1410.3248

[19] K. Iwata and Y. Oohama, "Information-spectrum characterization of broadcast channel with general source," *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, vol. E88-A, no. 10, pp. 2808–2818, Oct. 2005.