Coding Theorems for Random Access Communication over Compound Channel

Zheng Wang and Jie Luo
Electrical & Computer Engineering Department
Colorado State University, Fort Collins, CO 80523
Email: {zhwang, rockey}@engr.colostate.edu

***Abstract*— [1] Random access communication is used in practical systems to deliver bursty short messages. Because users only transmit occasionally, it is often difficult for the receiver to keep track of the time varying wireless channel states. Under this motivation, we develop channel coding theorems for random multiple access communication over compound channels with finite codeword length. Error performance bound and asymptotic error probability scaling laws are derived. We found that the results also help in deriving error performance bounds for the random multiple access system where the receiver is only interested in decoding messages from a user subset.**

I. Introduction

In a series of recent works [1][2], information theoretic channel coding was extended to distributed random multiple access communication where users determine their codes and communication rates individually, without sharing rate information with the receiver. Due to the lack of rate coordination, reliable message recovery in random access communication is not always possible. Receiver in this case decodes the transmitted messages only if a pre-determined reliability requirement is met, otherwise the receiver reports a collision. In [1], it was shown that the fundamental performance limitation of a random multiple access system can be characterized using an achievable rate region. Asymptotically as the codeword length is taken to infinity, the receiver is able to recover the messages reliably if the communication rate vector happens to be inside the rate region, and to reliably report a collision if the rate vector happens to be outside the region. The achievable rate region was shown to coincide with the Shannon information rate region without a convex hull operation [1]. In [2], the result was further strengthened to a rate and error probability scaling law. Achievable error probability bound with finite codeword length was also obtained [2].

In both [1] and [2], state of the communication channel is assumed known at the receiver. However, because random access communication deals with bursty short messages, channel access of a user is often fractional. This makes channel estimation and tracking very difficult at the receiver. It is therefore an important task to understand the fundamental system performance when the communication channel is not perfectly known. In this paper, we illustrate how coding theorems developed in [1][2] can be extended to random access communication over a compound channel [3]. We first consider a single user time-slotted random access system. Assume that, in each time slot, the transmitter chooses an arbitrary communication rate, which is defined as the normalized number of data nats encoded in a packet. Without rate and complete channel state information, receiver decodes the message only if a pre-determined error probability requirement is satisfied. We assume that the receiver chooses an "operation region", which is the set of rate and channel state pairs within which the receiver *intends* to decode the message, and outside which the receiver *intends* to report a collision (or outage). Given the operation region and a finite codeword length, a bound on the achievable system error probability, defined as the maximum of the decoding error probability and the collision miss detection probability, is derived. We then show that the compound channel results also help in obtaining error performance bounds for the random multiple access system where the receiver is only interested in recovering messages from a *subset* of users. This is because, conditioned on the receiver not decoding messages from the rest of users, the impact of their communication activities on the user subset of interest is equivalent to that of a compound channel.

II. Single-User Random Access Communication over a Compound Channel

To simplify the presentation, we only consider a single-user system with in mind that extension to a multi-user system follows naturally. Assume that time is slotted with each slot equaling the length of N symbol durations. This is also the length of a packet. We model the compound discrete-time memoryless channel using a finite set of conditional probabilities $\left\{P_{Y|X}^{(1)}, \cdots, P_{Y|X}^{(H)}\right\}$ with cardinality H, where $X \in \mathcal{X}$ is the channel input symbol with $\mathcal{X}$ being the the finite input alphabet, and $Y \in \mathcal{Y}$ is the channel output symbol with $\mathcal{Y}$ being the finite output alphabet. In each time slot, a channel realization is randomly chosen. Both the transmitter and the receiver know the channel set, but not the actual realization.

Suppose that, at the beginning of a time slot, the user chooses an arbitrary communication rate r, in nats per symbol, and encodes $\lfloor Nr \rfloor$ data nats, denoted by a message w, into a packet of N symbols. Assume that $r \in \{r_1, \cdots, r_M\}$, where $\{r_1, \cdots, r_M\}$ is a pre-determined set of rates with cardinality M. The receiver knows about the rate set, but not the actual rate realization. Encoding is done using a random coding

[1]This work was supported by the National Science Foundation under Grants CCF-0728826 and CCF-1016985.

scheme described as in [2] and also in the following. Let $\mathcal{L} = \{\mathcal{C}_\theta : \theta \in \Theta\}$ be a codebook library of the user, the codebooks of which are indexed by set Θ. Each codebook consists of M codeword classes. The i^{th} ($i \in \{1, \cdots, M\}$) codeword class contains $\lfloor e^{Nr_i} \rfloor$ codewords, each with N symbols. Denote $\mathcal{C}_\theta(w, r)_j$ as the j^{th} symbol of the codeword corresponding to message w and communication rate r in codebook $\mathcal{C}_\theta$. The user first generates θ according to a distribution γ, such that random variables $X_{(w,r),j} : \theta \to \mathcal{C}_\theta(w, r)_j$ are independently distributed according to an input distribution $P_{X|r}$[2]. The user then uses $\mathcal{C}_\theta$ to map (w, r) into a codeword, denoted by $\boldsymbol{x}_{(w,r)}$, and sends it to the receiver.

We assume that the receiver is shared with the codebook generation algorithm and hence knows the randomly generated codebook. Before transmission, the receiver chooses an "operation region" $\mathcal{R}$, which is a set of rate and channel pairs, i.e. $\mathcal{R} \subseteq \left\{(r, P_{Y|X}) : r \in \{r_1, \cdots, r_M\}, P_{Y|X} \in \{P^{(1)}_{Y|X}, \cdots, P^{(H)}_{Y|X}\}\right\}$. Although the actual rate and channel pair, denoted by $(r, P_{Y|X})$, is unknown at the receiver, the receiver *intends* to decode the message for $(r, P_{Y|X}) \in \mathcal{R}$ and *intends* to report a collision for $(r, P_{Y|X}) \notin \mathcal{R}$. In each time slot, upon receiving the channel output symbol vector $\boldsymbol{y}$, the receiver estimates the communication rate and channel pair, denoted by $(\hat{r}, \hat{P}_{Y|X})$. The receiver outputs the estimated message and rate vector pair $(\hat{w}, \hat{r})$ only if $(\hat{r}, \hat{P}_{Y|X}) \in \mathcal{R}$ and a pre-determined error probability requirement is satisfied. Otherwise the receiver outputs a collision.

Conditioned on that (w, r) is transmitted over the channel $P_{Y|X}$, we define the decoding error probability for $(w, r, P_{Y|X})$ with $(r, P_{Y|X}) \in \mathcal{R}$ as

$$P_e(w, r, P_{Y|X}) = Pr\{(\hat{w}, \hat{r}) \neq (w, r) | (w, r, P_{Y|X})\}, \quad \forall (w, r, P_{Y|X}), (r, P_{Y|X}) \in \mathcal{R}. \tag{1}$$

We define the collision miss detection probability for $(w, r, P_{Y|X})$ with $(r, P_{Y|X}) \notin \mathcal{R}$ as

$$\begin{aligned}\bar{P}_c(w, r, P_{Y|X}) = 1 - Pr\{\text{"collision"} | (w, r, P_{Y|X})\} \\ -Pr\{(\hat{w}, \hat{r}) = (w, r) | (w, r, P_{Y|X})\}, \\ \forall (w, r, P_{Y|X}), (r, P_{Y|X}) \notin \mathcal{R}.\end{aligned} \tag{2}$$

Note that when $(r, P_{Y|X}) \notin \mathcal{R}$, we do not regard correct message recovery as "collision miss detection".

Assume that $r < I_{(r,P_{Y|X})}(X; Y)$ for all $(r, P_{Y|X}) \in \mathcal{R}$, where $I_{(r,P_{Y|X})}$ is the mutual information function computed using input distribution $P_{X|r}$ and channel $P_{Y|X}$. Define the system error probability P_{es} as

$$P_{es} = \max\left\{ \max_{(w,r,P_{Y|X}),(r,P_{Y|X})\in\mathcal{R}} P_e(w, r, P_{Y|X}), \max_{(w,r,P_{Y|X}),(r,P_{Y|X})\notin\mathcal{R}} \bar{P}_c(w, r, P_{Y|X}) \right\}. \tag{3}$$

The following theorem gives an upper bound on the achievable system error probability P_{es}.

[2]Note that the input distribution is a function of the communication rate.

Theorem 1: For single-user random access communication over the compound discrete-time memoryless channel $\{P^{(1)}_{Y|X}, \cdots, P^{(H)}_{Y|X}\}$. Assume that finite codeword length N, and random coding with input distribution $P_{X|r}$ for all $r \in \{r_1, \cdots, r_M\}$. Let $\mathcal{R}$ be the operation region. There exists a decoding algorithm, whose system error probability P_{es} is upper bounded by

$$\begin{aligned}P_{es} \leq \max\Bigg\{ \max_{(r,P_{Y|X})\in\mathcal{R}} \Bigg[\max_{(\tilde{r},\tilde{P}_{Y|X})\notin\mathcal{R}} \exp\{-NE_i(r, \tilde{r}, P_{Y|X}, \tilde{P}_{Y|X})\} \\ + \sum_{(\tilde{r},\tilde{P}_{Y|X})\in\mathcal{R}} \exp\{-NE_m(r, \tilde{r}, P_{Y|X}, \tilde{P}_{Y|X})\} \Bigg], \\ \sum_{(r,P_{Y|X})\in\mathcal{R}} \max_{(\tilde{r},\tilde{P}_{Y|X})\notin\mathcal{R}} \exp\{-NE_i(r, \tilde{r}, P_{Y|X}, \tilde{P}_{Y|X})\} \Bigg\},\end{aligned} \tag{4}$$

where $E_m(r, \tilde{r}, P_{Y|X}, \tilde{P}_{Y|X})$ and $E_i(r, \tilde{r}, P_{Y|X}, \tilde{P}_{Y|X})$ are given by

$$\begin{aligned}E_m(r, \tilde{r}, P_{Y|X}, \tilde{P}_{Y|X}) = \max_{0<\rho,s\leq 1} -\rho\tilde{r} \\ -\log \sum_Y \left[\sum_X P_{X|r}(X) P_{Y|X}(Y|X)^{1-s}\right] \\ \times \left[\sum_X P_{X|\tilde{r}}(X) \tilde{P}_{Y|X}(Y|X)^{\frac{s}{\rho}}\right]^\rho, \\ E_i(r, \tilde{r}, P_{Y|X}, \tilde{P}_{Y|X}) = \max_{0<\rho\leq 1, 0<s\leq 1-\rho} -\rho r \\ -\log \sum_Y \left[\sum_X P_{X|\tilde{r}}(X) \tilde{P}_{Y|X}(Y|X)\right]^{1-s} \\ \times \left[\sum_X P_{X|r}(X) P_{Y|X}(Y|X)^{\frac{s}{s+\rho}}\right]^{s+\rho}.\end{aligned} \tag{5}$$

■

The proof of Theorem 1 is given in the Appendix. In the proof, we assume the following decoding algorithm. Upon receiving the channel output symbols $\boldsymbol{y}$, the receiver outputs an estimated message and rate pair (w, r) together with an estimated channel $P_{Y|X}$ if the following condition is satisfied,

$$\begin{aligned}-\frac{\log Pr\{\boldsymbol{y}|\boldsymbol{x}_{(w,r)}, P_{Y|X}\}}{N} < -\frac{\log Pr\{\boldsymbol{y}|\boldsymbol{x}_{(\tilde{w},\tilde{r})}, \tilde{P}_{Y|X}\}}{N}, \\ \text{for all } (\tilde{w}, \tilde{r}, \tilde{P}_{Y|X}), (\tilde{w}, \tilde{r}) \neq (w, r), \text{ and} \\ (\tilde{w}, \tilde{r}, \tilde{P}_{Y|X}), (w, r, P_{Y|X}) \in \mathcal{R}_{\boldsymbol{y}}, \text{ with} \\ \mathcal{R}_{\boldsymbol{y}} = \Bigg\{ (\tilde{w}, \tilde{r}, \tilde{P}_{Y|X}) \Bigg| (\tilde{r}, \tilde{P}_{Y|X}) \in \mathcal{R}, \\ -\frac{\log Pr\{\boldsymbol{y}|\boldsymbol{x}_{(\tilde{w},\tilde{r})}, \tilde{P}_{Y|X}\}}{N} < \tau_{(\tilde{r},\tilde{P}_{Y|X})}(\boldsymbol{y}) \Bigg\}.\end{aligned} \tag{6}$$

Here $\tau_{(\tilde{r},\tilde{P}_{Y|X})}(\cdot)$ is a pre-determined typicality threshold function of the channel output symbols $\boldsymbol{y}$, associated with the rate and channel pair $(\tilde{r}, \tilde{P}_{Y|X})$. If there is no codeword satisfying condition (6), the decoder reports a collision.

Define the corresponding system error exponent as $E_s = \lim_{N\to\infty} -\frac{1}{N}\log P_{es}$. The following lower bound on the achievable system error exponent E_s can be directly obtained from Theorem 1.

Corollary 1: The system error exponent of the single-user random access communication system given in Theorem 1 is lower-bounded by

$$E_s \geq \min\left\{ \min_{(r,P_{Y|X}),(\tilde{r},\tilde{P}_{Y|X})\in\mathcal{R}} E_m(r,\tilde{r},P_{Y|X},\tilde{P}_{Y|X}) \right., \left. \min_{(r,P_{Y|X})\in\mathcal{R},(\tilde{r},\tilde{P}_{Y|X})\notin\mathcal{R}} E_i(r,\tilde{r},P_{Y|X},\tilde{P}_{Y|X}) \right\}. \quad (7)$$

where $E_m(r,\tilde{r},P_{Y|X},\tilde{P}_{Y|X})$ and $E_i(r,\tilde{r},P_{Y|X},\tilde{P}_{Y|X})$ are defined in (5). ■

III. Individual User Decoding in Random Multiple Access Communication

In this section, we show that the results obtained in Section II can also help to derive error probability bounds in a random multiple access system where the receiver is only interested in recovering the messages from a user subset. We consider a K-user system over a discrete-time memoryless channel, modeled by $P_{Y|\boldsymbol{X}}$, where $\boldsymbol{X}$ denotes the channel input symbols for all users. We assume that the channel information is known both at the users and at the receiver, Note that the result can be generalized to the compound channel case.

Assume that in each time slot, each user, say user k ($k \in \{1,\cdots,K\}$), encodes its message w_k into a codeword $\boldsymbol{x}_{(w_k,r_k)}$ with rate $r_k \in \{r_{k1},\cdots,r_{kM}\}$ using the random coding scheme described in Section II. The communication rate is shared neither among the users nor with the receiver. We use bold-font vector symbols $\boldsymbol{w}$ and $\boldsymbol{r}$ to denote the messages and rates of multiple users. Let $\boldsymbol{P}_{\boldsymbol{X}|\boldsymbol{r}}$ be the input distributions of the users, which are functions of the corresponding communication rates. Given a user subset $\mathcal{D} \subseteq \{1,\cdots,K\}$, let $\bar{\mathcal{D}}$ be its complementary. We use $\boldsymbol{r}_{\mathcal{D}}$ to denote the rates of users in $\mathcal{D}$, and use r_k to denote the rate of user k, etc.

Given a user subset $\mathcal{D} \subseteq \{1,\cdots,K\}$ and a rate region $\mathcal{R}_{\mathcal{D}}$ which is a collection of rate vectors, we define an elementary decoder, called "$(\mathcal{D},\mathcal{R}_{\mathcal{D}})$-decoder", as follows. We assume that the decoder regards signals from users not in $\mathcal{D}$ as interference. Let $Pr\{\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w}_{\mathcal{D}},\boldsymbol{r}_{\mathcal{D}})},\boldsymbol{r}_{\bar{\mathcal{D}}}\}$ be the conditional probability of getting channel output $\boldsymbol{y}$, with users in $\mathcal{D}$ transmitting $(\boldsymbol{w}_{\mathcal{D}},\boldsymbol{r}_{\mathcal{D}})$ and users not in $\mathcal{D}$ choosing rate $\boldsymbol{r}_{\bar{\mathcal{D}}}$. In other words,

$$Pr\{\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w}_{\mathcal{D}},\boldsymbol{r}_{\mathcal{D}})},\boldsymbol{r}_{\bar{\mathcal{D}}}\} = \sum_{\boldsymbol{w}_{\bar{\mathcal{D}}}} Pr\{\boldsymbol{x}_{(\boldsymbol{w}_{\bar{\mathcal{D}}},\boldsymbol{r}_{\bar{\mathcal{D}}})}\} Pr\{\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})}\}. \quad (8)$$

Upon receiving the channel output symbols $\boldsymbol{y}$, the $(\mathcal{D},\mathcal{R}_{\mathcal{D}})$-decoder outputs an estimated message and rate pair $(\boldsymbol{w}_{\mathcal{D}},\boldsymbol{r}_{\mathcal{D}})$ for users in $\mathcal{D}$ together with an estimated rate $\boldsymbol{r}_{\bar{\mathcal{D}}}$ for users not in $\mathcal{D}$, if the following condition is satisfied,

$$-\frac{\log Pr\{\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w}_{\mathcal{D}},\boldsymbol{r}_{\mathcal{D}})},\boldsymbol{r}_{\bar{\mathcal{D}}}\}}{N} < -\frac{\log Pr\{\boldsymbol{y}|\boldsymbol{x}_{(\tilde{\boldsymbol{w}}_{\mathcal{D}},\tilde{\boldsymbol{r}}_{\mathcal{D}})},\tilde{\boldsymbol{r}}_{\bar{\mathcal{D}}}\}}{N},$$
$$\text{for all } (\tilde{\boldsymbol{w}}_{\mathcal{D}},\tilde{\boldsymbol{r}}_{\mathcal{D}},\tilde{\boldsymbol{r}}_{\bar{\mathcal{D}}}), (\tilde{\boldsymbol{w}}_{\mathcal{D}},\tilde{\boldsymbol{r}}_{\mathcal{D}}) \neq (\boldsymbol{w}_{\mathcal{D}},\boldsymbol{r}_{\mathcal{D}}), \text{ and } (\tilde{\boldsymbol{w}}_{\mathcal{D}},\tilde{\boldsymbol{r}}_{\mathcal{D}},\tilde{\boldsymbol{r}}_{\bar{\mathcal{D}}}), (\boldsymbol{w}_{\mathcal{D}},\boldsymbol{r}_{\mathcal{D}},\boldsymbol{r}_{\bar{\mathcal{D}}}) \in \mathcal{R}_{\mathcal{D}\boldsymbol{y}}, \text{ with}$$
$$\mathcal{R}_{\mathcal{D}\boldsymbol{y}} = \left\{ (\tilde{\boldsymbol{w}}_{\mathcal{D}},\tilde{\boldsymbol{r}}_{\mathcal{D}},\tilde{\boldsymbol{r}}_{\bar{\mathcal{D}}}) | \, \tilde{\boldsymbol{r}} \in \mathcal{R}_{\mathcal{D}}, -\frac{\log Pr\{\boldsymbol{y}|\boldsymbol{x}_{(\tilde{\boldsymbol{w}}_{\mathcal{D}},\tilde{\boldsymbol{r}}_{\mathcal{D}})},\tilde{\boldsymbol{r}}_{\bar{\mathcal{D}}}\}}{N} < \tau_{(\tilde{\boldsymbol{r}},\mathcal{D})}(\boldsymbol{y}) \right\}, \quad (9)$$

where $\tau_{(\tilde{\boldsymbol{r}},\mathcal{D})}(\cdot)$ is a pre-determined typicality threshold.

By comparting (9) with (6), we can see that, in an $(\mathcal{D},\mathcal{R}_{\mathcal{D}})$-decoder, the impact of signals not in $\mathcal{D}$ is essentially equivalent to that of a compound channel.

Conditioned on users in $\mathcal{D}$ transmitting $(\boldsymbol{w}_{\mathcal{D}},\boldsymbol{r}_{\mathcal{D}})$ and users not in $\mathcal{D}$ choosing rate $\boldsymbol{r}_{\bar{\mathcal{D}}}$, let us denote the estimated messages and rates by $(\hat{\boldsymbol{w}}_{\mathcal{D}},\hat{\boldsymbol{r}}_{\mathcal{D}})$ and $\hat{\boldsymbol{r}}_{\bar{\mathcal{D}}}$ if the decoder does not report a collision. We define the decoding error probability of the $(\mathcal{D},\mathcal{R}_{\mathcal{D}})$-decoder for $(\boldsymbol{w}_{\mathcal{D}},\boldsymbol{r}_{\mathcal{D}},\boldsymbol{r}_{\bar{\mathcal{D}}})$ with $\boldsymbol{r} \in \mathcal{R}_{\mathcal{D}}$ as

$$P_e(\boldsymbol{w}_{\mathcal{D}},\boldsymbol{r}_{\mathcal{D}},\boldsymbol{r}_{\bar{\mathcal{D}}}) = Pr\{(\hat{\boldsymbol{w}}_{\mathcal{D}},\hat{\boldsymbol{r}}_{\mathcal{D}}) \neq (\boldsymbol{w}_{\mathcal{D}},\boldsymbol{r}_{\mathcal{D}}) | (\boldsymbol{w}_{\mathcal{D}},\boldsymbol{r}_{\mathcal{D}},\boldsymbol{r}_{\bar{\mathcal{D}}})\}, \quad \forall (\boldsymbol{w}_{\mathcal{D}},\boldsymbol{r}_{\mathcal{D}},\boldsymbol{r}_{\bar{\mathcal{D}}}), \boldsymbol{r} \in \mathcal{R}_{\mathcal{D}}. \quad (10)$$

We define the collision miss detection probability for $(\boldsymbol{w}_{\mathcal{D}},\boldsymbol{r}_{\mathcal{D}},\boldsymbol{r}_{\bar{\mathcal{D}}})$ with $\boldsymbol{r} \notin \mathcal{R}_{\mathcal{D}}$ as

$$\bar{P}_c(\boldsymbol{w}_{\mathcal{D}},\boldsymbol{r}_{\mathcal{D}},\boldsymbol{r}_{\bar{\mathcal{D}}}) = 1 - Pr\{\text{"collision"} | (\boldsymbol{w}_{\mathcal{D}},\boldsymbol{r}_{\mathcal{D}},\boldsymbol{r}_{\bar{\mathcal{D}}})\} - Pr\{(\hat{\boldsymbol{w}}_{\mathcal{D}},\hat{\boldsymbol{r}}_{\mathcal{D}}) = (\boldsymbol{w}_{\mathcal{D}},\boldsymbol{r}_{\mathcal{D}}) | (\boldsymbol{w}_{\mathcal{D}},\boldsymbol{r}_{\mathcal{D}},\boldsymbol{r}_{\bar{\mathcal{D}}})\}, \quad \forall (\boldsymbol{w}_{\mathcal{D}},\boldsymbol{r}_{\mathcal{D}},\boldsymbol{r}_{\bar{\mathcal{D}}}), \boldsymbol{r} \notin \mathcal{R}_{\mathcal{D}}. \quad (11)$$

System error probability of the $(\mathcal{D},\mathcal{R}_{\mathcal{D}})$-decoder is defined by

$$P_{es}(\mathcal{D},\mathcal{R}_{\mathcal{D}}) = \max\left\{ \max_{(\boldsymbol{w}_{\mathcal{D}},\boldsymbol{r}_{\mathcal{D}},\boldsymbol{r}_{\bar{\mathcal{D}}}),\boldsymbol{r}\in\mathcal{R}_{\mathcal{D}}} P_e(\boldsymbol{w}_{\mathcal{D}},\boldsymbol{r}_{\mathcal{D}},\boldsymbol{r}_{\bar{\mathcal{D}}}), \max_{(\boldsymbol{w}_{\mathcal{D}},\boldsymbol{r}_{\mathcal{D}},\boldsymbol{r}_{\bar{\mathcal{D}}}),\boldsymbol{r}\notin\mathcal{R}_{\mathcal{D}}} \bar{P}_c(\boldsymbol{w}_{\mathcal{D}},\boldsymbol{r}_{\mathcal{D}},\boldsymbol{r}_{\bar{\mathcal{D}}}) \right\}. \quad (12)$$

We assume that, for any $\boldsymbol{r} \in \mathcal{R}_{\mathcal{D}}$, we have

$$\sum_{k\in\tilde{\mathcal{D}}} r_k < I_{\boldsymbol{r}}(\boldsymbol{X}_{\tilde{\mathcal{D}}}; Y | \boldsymbol{X}_{\mathcal{D}\setminus\tilde{\mathcal{D}}}), \qquad \forall \tilde{\mathcal{D}} \subseteq \mathcal{D}, \quad (13)$$

where $I_{\boldsymbol{r}}(\boldsymbol{X}_{\tilde{\mathcal{D}}}; Y | \boldsymbol{X}_{\mathcal{D}\setminus\tilde{\mathcal{D}}})$ is the mutual information computed using the input distribution corresponding to rate $\boldsymbol{r}$.

When the codeword length N is finite, the follow lemma gives an upper bound on the achievable P_{es} defined in (12).

Lemma 1: Consider a K-user random multiple access communication system over a discrete-time memoryless channel $P_{Y|\boldsymbol{X}}$ with an $(\mathcal{D},\mathcal{R}_{\mathcal{D}})$-decoder. Given that (13) is satisfied, the typicality threshold given in (9) can be optimized to achieve the following system error probability bound,

$$P_{es}(\mathcal{D},\mathcal{R}_{\mathcal{D}}) \leq \max\Bigg\{ \max_{\boldsymbol{r}\in\mathcal{R}_{\mathcal{D}}} \sum_{\tilde{\mathcal{D}}\subset\mathcal{D}} \Bigg[\sum_{\substack{\tilde{\boldsymbol{r}}\,\in\,\mathcal{R}_{\mathcal{D}},\\ \tilde{\boldsymbol{r}}_{\tilde{\mathcal{D}}}=\boldsymbol{r}_{\tilde{\mathcal{D}}}}} \exp\{-NE_{m\mathcal{D}}(\tilde{\mathcal{D}},\boldsymbol{r},\tilde{\boldsymbol{r}})\} + \max_{\substack{\boldsymbol{r}'\,\not\in\,\mathcal{R}_{\mathcal{D}},\\ \boldsymbol{r}'_{\tilde{\mathcal{D}}}=\boldsymbol{r}_{\tilde{\mathcal{D}}}}} \exp\{-NE_{i\mathcal{D}}(\tilde{\mathcal{D}},\boldsymbol{r},\boldsymbol{r}')\}\Bigg], \max_{\tilde{\boldsymbol{r}}\not\in\mathcal{R}} \sum_{\tilde{\mathcal{D}}\subset\mathcal{D}} \sum_{\substack{\boldsymbol{r}\,\in\,\mathcal{R}_{\mathcal{D}},\\ \boldsymbol{r}_{\tilde{\mathcal{D}}}=\tilde{\boldsymbol{r}}_{\tilde{\mathcal{D}}}}} \max_{\substack{\boldsymbol{r}'\,\not\in\,\mathcal{R}_{\mathcal{D}},\\ \boldsymbol{r}'_{\tilde{\mathcal{D}}}=\tilde{\boldsymbol{r}}_{\tilde{\mathcal{D}}}}} \exp\{-NE_{i\mathcal{D}}(\tilde{\mathcal{D}},\boldsymbol{r},\boldsymbol{r}')\}\Bigg\}, \quad (14)$$

where $E_{m\mathcal{D}}(\tilde{\mathcal{D}},\boldsymbol{r},\tilde{\boldsymbol{r}})$ and $E_{i\mathcal{D}}(\tilde{\mathcal{D}},\boldsymbol{r},\boldsymbol{r}')$ are given by,

$$\begin{aligned}
E_{m\mathcal{D}}(\tilde{\mathcal{D}},\boldsymbol{r},\tilde{\boldsymbol{r}}) &= \max_{0<\rho\leq 1} -\rho \sum_{k\in\mathcal{D}\setminus\tilde{\mathcal{D}}} \tilde{r}_k + \max_{0<s\leq 1} -\log \sum_{Y}\sum_{\boldsymbol{X}_{\tilde{\mathcal{D}}}} \prod_{k\in\tilde{\mathcal{D}}} P_{X|r_k}(X_k) \\
&\times \left(\sum_{\boldsymbol{X}_{\mathcal{D}\setminus\tilde{\mathcal{D}}}} \prod_{k\in\mathcal{D}\setminus\tilde{\mathcal{D}}} P_{X|r_k}(X_k) P(Y|\boldsymbol{X}_{\mathcal{D}},\boldsymbol{r})^{1-s}\right) \\
&\times \left(\sum_{\boldsymbol{X}_{\mathcal{D}\setminus\tilde{\mathcal{D}}}} \prod_{k\in\mathcal{D}\setminus\tilde{\mathcal{D}}} P_{X|\tilde{r}_k}(X_k) P(Y|\boldsymbol{X}_{\mathcal{D}},\tilde{\boldsymbol{r}})^{\frac{s}{\rho}}\right)^{\rho}, \\
E_{i\mathcal{D}}(\tilde{\mathcal{D}},\boldsymbol{r},\boldsymbol{r}') &= \max_{0<\rho\leq 1} -\rho \sum_{k\in\mathcal{D}\setminus\tilde{\mathcal{D}}} r_k + \max_{0<s\leq 1-\rho} -\log \sum_{Y}\sum_{\boldsymbol{X}_{\tilde{\mathcal{D}}}} \prod_{k\in\tilde{\mathcal{D}}} P_{X|r_k}(X_k) \\
&\times \left(\sum_{\boldsymbol{X}_{\mathcal{D}\setminus\tilde{\mathcal{D}}}} \prod_{k\in\mathcal{D}\setminus\tilde{\mathcal{D}}} P_{X|r_k}(X_k) P(Y|\boldsymbol{X}_{\mathcal{D}},\boldsymbol{r})^{\frac{s}{s+\rho}}\right)^{s+\rho} \\
&\times \left(\sum_{\boldsymbol{X}_{\mathcal{D}\setminus\tilde{\mathcal{D}}}} \prod_{k\in\mathcal{D}\setminus\tilde{\mathcal{D}}} P_{X|r'_k}(X_k) P(Y|\boldsymbol{X}_{\mathcal{D}},\boldsymbol{r}')\right)^{1-s}, \\
P(Y|\boldsymbol{X}_{\mathcal{D}},\boldsymbol{r}) &= \sum_{\boldsymbol{X}_{\bar{\mathcal{D}}}} \prod_{k\notin\mathcal{D}} P_{X|r_k}(X_k) P_{Y|\boldsymbol{X}}(Y|\boldsymbol{X}). \quad (15)
\end{aligned}$$

■

Even though the notation in Lemma 1 is quite complicated, it is not difficult to obtain its proof by combining the proof of [2, Theorem 2] and the proof of Theorem 1. Due to the page limitation, the detail is skipped in this paper.

Next, we consider the case when the receiver in the K-user random multiple access system is only interested in recovering the message from user k. However, we will drop the assumption that the receiver should regard signals from other users as interference. We assume that the receiver chooses an operation region $\mathcal{R}$. Upon receiving the channel output symbols $\boldsymbol{y}$, the receiver estimates the message $\hat{w}_k$ for user k and the rate vector $\hat{\boldsymbol{r}}$. The receiver outputs $\hat{w}_k$ if $\hat{\boldsymbol{r}}\in\mathcal{R}$ and $\hat{w}_k$ satisfies a pre-determined reliability requirement. Otherwise the receiver outputs a collision for user k.

The achievable rate region of such a system with single-user decoding was originally characterized in [1] as

$$\mathcal{R}_k = \left\{ \boldsymbol{r} \;\middle|\; \begin{array}{l} \forall\mathcal{S}\subseteq\{1,\cdots,K\}, k\in\mathcal{S}, \text{ either } r_k=0, \\ \text{or } \exists\tilde{\mathcal{S}}\subseteq\mathcal{S}, k\in\tilde{\mathcal{S}}, \text{ such that,} \\ \sum_{i\in\tilde{\mathcal{S}}} r_i < I_{\boldsymbol{r}}(\boldsymbol{X}_{\tilde{\mathcal{S}}};Y|\boldsymbol{X}_{\bar{\mathcal{S}}}) \end{array} \right\}. \quad (16)$$

(16) implies that, asymptotically as $N\to\infty$, if the communication rate vector $\boldsymbol{r}\in\mathcal{R}_k$ is within the achievable rate region, the receiver can always find a user subset $\tilde{\mathcal{S}}$ with $k\in\tilde{\mathcal{S}}$, such that all messages from users in $\tilde{\mathcal{S}}$ can be reliably decoded by regarding other user signals as interference.

Based on the above understanding, we have the following lemma whose proof is skipped.

Lemma 2: Any operation region $\mathcal{R}\subset\mathcal{R}_k$ contained inside the achievable rate region $\mathcal{R}_k$ can be partitioned into the following sub-regions

$$\begin{aligned}
&\mathcal{R} = \bigcup_{\mathcal{D}:\mathcal{D}\subseteq\{1,\cdots,K\},k\in\mathcal{D}} \mathcal{R}_{\mathcal{D}}, \\
&\mathcal{R}_{\mathcal{D}}\cap\mathcal{R}_{\mathcal{D}'}=\phi, \forall\mathcal{D}\neq\mathcal{D}', \mathcal{D},\mathcal{D}'\subseteq\{1,\cdots,K\}, k\in\mathcal{D},\mathcal{D}', \\
&\sum_{i\in\tilde{\mathcal{D}}} r_i < I_{\boldsymbol{r}}(\boldsymbol{X}_{\tilde{\mathcal{D}}};Y|\boldsymbol{X}_{\mathcal{D}\setminus\tilde{\mathcal{D}}}), \forall\mathcal{D}, \boldsymbol{r}\in\mathcal{R}_{\mathcal{D}}, \tilde{\mathcal{D}}\subseteq\mathcal{D}. \quad (17)
\end{aligned}$$

■

Given an operation region partitioning as in Lemma 2 and the corresponding $(\mathcal{D},\mathcal{R}_{\mathcal{D}})$-decoders, the single-user decoder outputs an estimated message $\hat{w}_k$ for user k if the estimates for user k given by all the $(\mathcal{D},\mathcal{R}_{\mathcal{D}})$-decoders agree with each other. Otherwise, the single-user decoder reports a collision for user k.

Define the system error probability similarly as in previous systems. The following theorem gives an upper bound on the achievable system error probability of the single-user decoder.

Theorem 2: Consider a K-user random multiple access system over a discrete-time memoryless channel $P_{Y|\boldsymbol{X}}$, with the receiver only interested in recovering the message from user k. Assume that the receiver chooses an operation region $\mathcal{R}\subset\mathcal{R}_k$ contained inside the achievable rate region. Let σ be an arbitrary partitioning of the operation region $\mathcal{R}$ satisfying (17). System error probability of the single-user decoder is upper-bounded by,

$$P_{es} \leq \min_{\sigma} \sum_{\mathcal{D}:\mathcal{D}\subseteq\{1,\cdots,K\},k\in\mathcal{D}} P_{es}(\mathcal{D},\mathcal{R}_{\mathcal{D}}), \quad (18)$$

where $P_{es}(\mathcal{D},\mathcal{R}_{\mathcal{D}})$ is the system error probability of the $(\mathcal{D},\mathcal{R}_{\mathcal{D}})$-decoder, and can be further bounded by (14). ■

Theorem 2 is implied by Lemmas 1 and 2. Note that the error probability bound provided in Theorem 2 is an implicit one since how to optimize the partitioning scheme σ in (18) remains a challenging open problem.

APPENDIX

Proof of Theorem 1: Assume that the decoding algorithm given in (6). To derive the upper bound of the system error probability, we define the following probability terms.

First, assume that (w,r) is transmitted over channel $P_{Y|X}$ with $(r,P_{Y|X})\in\mathcal{R}$. Define $P_{t[r,P_{Y|X}]}$ as the probability that the likelihood value of the transmitted codeword is no larger than the corresponding typicality threshold,

$$P_{t[r,P_{Y|X}]} = Pr\left\{P(\boldsymbol{y}|\boldsymbol{x}_{(w,r)},P_{Y|X}) \le e^{-N\tau_{(r,P_{Y|X})}(\boldsymbol{y})}\right\}. \tag{19}$$

Define $P_{m[(r,P_{Y|X}),(\tilde{r},\tilde{P}_{Y|X})]}$ as the probability that the likelihood value of the transmitted codeword over channel $P_{Y|X}$ is no larger than that of some codeword $\boldsymbol{x}_{(\tilde{w},\tilde{r})} \ne \boldsymbol{x}_{(w,r)}$ over channel $\tilde{P}_{Y|X}$ with $(\tilde{r},\tilde{P}_{Y|X})\in\mathcal{R}$,

$$\begin{aligned}&P_{m[(r,P_{Y|X}),(\tilde{r},\tilde{P}_{Y|X})]}\\&= Pr\left\{P(\boldsymbol{y}|\boldsymbol{x}_{(w,r)},P_{Y|X}) \le P(\boldsymbol{y}|\boldsymbol{x}_{(\tilde{w},\tilde{r})},\tilde{P}_{Y|X})\right\},\\&\quad (\tilde{w},\tilde{r})\ne(w,r),(\tilde{r},\tilde{P}_{Y|X})\in\mathcal{R}.\end{aligned} \tag{20}$$

Second, assume that $(\tilde{w},\tilde{r})$ is transmitted over channel $\tilde{P}_{Y|X}$ with $(\tilde{r},\tilde{P}_{Y|X})\notin\mathcal{R}$. Let $P_{i[(\tilde{r},\tilde{P}_{Y|X}),(r,P_{Y|X})]}$ be the probability that there exist $(w,r)\ne(\tilde{w},\tilde{r})$ and a channel $P_{Y|X}$ with $(r,P_{Y|X})\in\mathcal{R}$, such that the corresponding likelihood value is larger than the typicality threshold,

$$\begin{aligned}&P_{i[(\tilde{r},\tilde{P}_{Y|X}),(r,P_{Y|X})]}\\&= Pr\left\{P(\boldsymbol{y}|\boldsymbol{x}_{(w,r)},P_{Y|X}) > e^{-N\tau_{(r,P_{Y|X})}(\boldsymbol{y})}\right\},\\&\quad (w,r)\ne(\tilde{w},\tilde{r}),(r,P_{Y|X})\in\mathcal{R}.\end{aligned} \tag{21}$$

With these definitions, we can upper bound P_{es} by

$$\begin{aligned}P_{es} \le \max\Bigg\{&\max_{(w,r,P_{Y|X}),(r,P_{Y|X})\in\mathcal{R}}\Bigg[P_{t[r,P_{Y|X}]}\\&+\sum_{(\tilde{r},\tilde{P}_{Y|X})\in\mathcal{R}} P_{m[(r,P_{Y|X}),(\tilde{r},\tilde{P}_{Y|X})]}\Bigg],\\&\max_{(\tilde{w},\tilde{r},\tilde{P}_{Y|X}),(\tilde{r},\tilde{P}_{Y|X})\notin\mathcal{R}}\sum_{(r,P_{Y|X})\in\mathcal{R}} P_{i[(\tilde{r},\tilde{P}_{Y|X}),(r,P_{Y|X})]}\Bigg\}.\end{aligned} \tag{22}$$

By following a derivation similar to [2, (24)-(28)], we can upper bound $P_{m[(r,P_{Y|X}),(\tilde{r},\tilde{P}_{Y|X})]}$ by

$$P_{m[(r,P_{Y|X}),(\tilde{r},\tilde{P}_{Y|X})]} \le \exp\{-NE_m(r,\tilde{r},P_{Y|X},\tilde{P}_{Y|X})\}, \tag{23}$$

where $E_m(r,\tilde{r},P_{Y|X},\tilde{P}_{Y|X})$ is given in (5).

By following derivations similar to [2, (30)-(32)] and [2, (33)-(36)], we can upper bound $P_{t[r,P_{Y|X}]}$ by

$$\begin{aligned}P_{t[r,P_{Y|X}]} \le &\sum_{\boldsymbol{y}} E_\theta\left[P(\boldsymbol{y}|\boldsymbol{x}_{(w,r)},P_{Y|X})^{1-s_1}\right]\\&\times e^{-Ns_1\tau_{(r,P_{Y|X})}(\boldsymbol{y})}, \qquad \forall s_1>0,\end{aligned} \tag{24}$$

and upper bound $P_{i[(\tilde{r},\tilde{P}_{Y|X}),(r,P_{Y|X})]}$ by

$$\begin{aligned}&P_{i[(\tilde{r},\tilde{P}_{Y|X}),(r,P_{Y|X})]} \le \max_{(\tilde{r},\tilde{P}_{Y|X})\notin\mathcal{R}}\sum_{\boldsymbol{y}}\\&E_\theta\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\tilde{w},\tilde{r})},\tilde{P}_{Y|X})\right]\left\{E_\theta\left[P(\boldsymbol{y}|\boldsymbol{x}_{(w,r)},P_{Y|X})^{\frac{s_2}{\tilde{\rho}}}\right]\right\}^{\tilde{\rho}}\\&\times e^{Ns_2\tau_{(r,P_{Y|X})}(\boldsymbol{y})}e^{N\tilde{\rho}r}, \qquad s_2>0, 0<\tilde{\rho}\le 1.\end{aligned} \tag{25}$$

We determine the optimal $\tau_{(r,P_{Y|X})}(\boldsymbol{y})$ for all $(r,P_{Y|X})\in\mathcal{R}$ by jointly optimizing the bounds in (24) and (25). Note that given $(r,P_{Y|X})\in\mathcal{R}$, $\boldsymbol{y}$ and the auxiliary variables $s_1,s_2>0$, $0<\tilde{\rho}\le 1$, the bound in (24) decreases in $\tau_{(r,P_{Y|X})}(\boldsymbol{y})$, while the bounds in (25) increases in $\tau_{(r,P_{Y|X})}(\boldsymbol{y})$. Therefore, we choose $\tau_{(r,P_{Y|X})}(\boldsymbol{y})$ to satisfy the following equality,

$$\begin{aligned}&E_\theta\left[P(\boldsymbol{y}|\boldsymbol{x}_{(w,r)},P_{Y|X})^{1-s_1}\right]e^{-Ns_1\tau_{(r,P_{Y|X})}(\boldsymbol{y})}\\&= E_\theta\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\tilde{w},\tilde{r}^*)},\tilde{P}^*_{Y|X})\right]\\&\times\left\{E_\theta\left[P(\boldsymbol{y}|\boldsymbol{x}_{(w,r)},P_{Y|X})^{\frac{s_2}{\tilde{\rho}}}\right]\right\}^{\tilde{\rho}}e^{Ns_2\tau_{(r,P_{Y|X})}(\boldsymbol{y})}e^{N\tilde{\rho}r}.\end{aligned} \tag{26}$$

where $(\tilde{r}^*,\tilde{P}^*_{Y|X})\notin\mathcal{R}$ is the rate and channel pair that maximize the right hand side of (25).

Substituting the optimal $\tau_{(r,P_{Y|X})}(\boldsymbol{y})$ that satisfies (26) into (24) gives us

$$\begin{aligned}P_{t[r,P_{Y|X}]} \le &\sum_{\boldsymbol{y}}\left\{E_\theta\left[P(\boldsymbol{y}|\boldsymbol{x}_{(w,r)},P_{Y|X})^{1-s_1}\right]\right\}^{\frac{s_2}{s_1+s_2}}\\&\times\left\{E_\theta\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\tilde{w},\tilde{r}^*)},\tilde{P}^*_{Y|X})\right]\right\}^{\frac{s_1}{s_1+s_2}}\\&\times\left\{E_\theta\left[P(\boldsymbol{y}|\boldsymbol{x}_{(w,r)},P_{Y|X})^{\frac{s_2}{\tilde{\rho}}}\right]\right\}^{\frac{\tilde{\rho}s_1}{s_1+s_2}}e^{Nr\frac{\tilde{\rho}s_1}{s_1+s_2}}.\end{aligned} \tag{27}$$

Let $s_2<\tilde{\rho}$ and $s_1 = 1-\frac{s_2}{\tilde{\rho}}$, and then do a variable change with $\rho=\frac{\tilde{\rho}(\tilde{\rho}-s_2)}{\tilde{\rho}-(1-\tilde{\rho})s_2}$ and $s=1-\frac{\tilde{\rho}-s_2}{\tilde{\rho}-(1-\tilde{\rho})s_2}$. Inequality (27) becomes

$$\begin{aligned}P_{t[r,P_{Y|X}]} \le &\left\{\sum_Y\left[\sum_X P_{X|r}(X)P_{Y|X}(Y|X)^{\frac{s}{s+\rho}}\right]^{s+\rho}\right.\\&\left.\times\left[\sum_X P_{X|\tilde{r}^*}(X)\tilde{P}^*_{Y|X}(Y|X)\right]^{1-s}\right\}^N e^{N\rho r}.\end{aligned} \tag{28}$$

By following a similar derivation, $P_{i[(\tilde{r},\tilde{P}_{Y|X}),(r,P_{Y|X})]}$ can be proved to satisfy the same upper bound given in (28). Since (28) holds for all $0<\rho\le 1$ and $0<s\le 1-\rho$, we have

$$\begin{aligned}&P_{t[r,P_{Y|X}]}, P_{i[(\tilde{r},\tilde{P}_{Y|X}),(r,P_{Y|X})]}\\&\le \max_{(\tilde{r},\tilde{P}_{Y|X})\notin\mathcal{R}}\exp\{-NE_i(r,\tilde{r},P_{Y|X},\tilde{P}_{Y|X})\},\end{aligned} \tag{29}$$

where $E_i(r,\tilde{r},P_{Y|X},\tilde{P}_{Y|X})$ is given in (5).

Substitute (23) and (29) into (22) yields the desired upper bound on P_{es}. ■

References

[1] J. Luo and A. Ephremides, *A Channel Coding Approach for Random Access Communication with Bursty Sources*, IEEE International Symposium on Information Theory, Austin, Tx, Jun. 2010.

[2] Z. Wang and J. Luo, *Error Performance of Channel Coding in Random Access Communication*, submitted to IEEE Trans. on Inform. Theory. http://arxiv.org/abs/1010.0642.

[3] I. Csiszár and J. Korner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, New York: Academic, 1981.