Intermittent Multi-Access Communication: Achievable Rates

Mostafa Khoshnevisan and J. Nicholas Laneman
Department of Electrical Engineering
University of Notre Dame
Notre Dame, Indiana 46556
Email:{mkhoshne, jnl}@nd.edu

***Abstract*—THIS PAPER IS ELIGIBLE FOR THE STUDENT PAPER AWARD. We formulated a model for intermittent multi-access communication for two users that captures the bursty transmission of the codeword symbols for each user and the possible asynchronism between the receiver and the transmitters as well as between the transmitters themselves. By making different assumptions for the intermittent process, we specialize the system to a random access system with or without collisions, where the collisions can be viewed either as deletions or interference. For each model, we characterize the performance of the system in terms of achievable rate regions. The results suggest that the intermittency of the system comes with a cost.**

I. Introduction

Multi-access communication is treated in different ways in the literature. Gallager [1] reviews both information theoretic and network oriented approaches, and emphasizes the need for a perspective that can merge elements from these two approaches. As also pointed out in [2], information theoretic models focus on accurate analysis of the effect of the noise and interference, whereas network oriented models focus on bursty transmissions and collision resolution approaches. An example of a recent work that introduces a model for multi-access communication capturing elements from these two approaches is [3], which introduces an information theoretic model for a random access communication scenario with two modes of operation for each user, active or inactive.

This paper can be viewed as another attempt to combine the information theoretic and network oriented multi-access models and characterize the performance of the system in terms of the achievable rate regions. We formulate a model for intermittent multi-access communication for two users that captures two network oriented concepts. First, it models bursty transmission of the codeword symbols for each user. Second, it takes into account the possible asynchronism between the receiver and the transmitters as well as between the transmitters themselves. A basic system model is introduced in Section II, which generalizes the intermittent communication model introduced in [4], and then by making different assumptions for the intermittent process, we specialize the system to two models: random access without collisions and random access with collisions in Sections IV and V, respectively. The collisions are treated either as deletions or as interference.

For each model, we obtain achievable rate regions, which depend on the concept of partial divergence introduced in [4], [5]. Because of the assumption that the receiver does not know a priori that an output symbol corresponds to transmission by a given user or that it corresponds to a collision, the decoder has to both detect the positions and decode the messages. The achievability results suggest that the intermittency of the system comes with a significant cost, i.e., it reduces the size of the achievable rate regions, which can be interpreted as communication overhead [6]. Note that as opposed to [6], where the constraint is the lack of coordination between the users in multi-access communication, the constraint in this paper is the intermittency of the system.

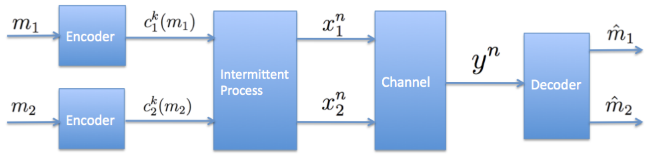

Fig. 1. System model for intermittent multi-access communication.

II. System Model

We consider a 2-user discrete memoryless multiple access channel (DM-MAC) with conditional probability mass functions $W(y|x_1, x_2)$ over input alphabets $\mathcal{X}_1$ and $\mathcal{X}_2$ and output alphabet $\mathcal{Y}$. The two senders wish to communicate independent messages $m_1 \in \{1, 2, ..., e^{kR_1} = M_1\}$ and $m_2 \in \{1, 2, ..., e^{kR_2} = M_2\}$ to a receiver. Let $\star \in \mathcal{X}_1, \mathcal{X}_2$ denote the noise symbol, which is the input of the channel when the corresponding sender is silent. Let $W_{\cdot\star} := W(y|x_1, x_2 = \star)$ denote the probability transition matrix for the point to point channel for user 1 if user 2 becomes silent, let$W_{\star\cdot}$ be defined analogously, and let $W_{\star\star} := W(y|x_1 = \star, x_2 = \star)$ denote the output distribution if both users become silent. Each user encodes the message to a codeword of length k: $c_1^k(m_1)$ and $c_2^k(m_2)$ denote the codewords of user 1 and user 2, respectively. The symbols of the codewords are called information symbols. Assume that x_1^n and x_2^n are the input sequences and y^n is the output sequence of the channel, where n is the length of the receive window at the decoder. Figure 1 shows a block diagram for the system model in which the intermittent process with inputs $c_1^k(m_1)$ and $c_2^k(m_2)$ and outputs x_1^n and x_2^n captures the burstiness and the asynchronism of the users.

Assuming that the decoded messages are denoted by $\hat{m}_1$ and $\hat{m}_2$, which are functions of the random sequence Y^n, we say that the rate pair (R_1, R_2) is achievable if there exists two sequences of length k codes of size e^{kR_1} and e^{kR_2} for the two encoders with $\frac{1}{M_1 M_2}\sum_{m_1=1}^{M_1}\sum_{m_2=1}^{M_2}\mathbb{P}((\hat{m}_1,\hat{m}_2)\neq(m_1,m_2))\to 0$ as $k\to\infty$. We refer to this general scenario as intermittent multi-access communication, and consider several models for the intermittent process in Figure 1, which we describe after introducing some preliminaries.

III. Preliminaries

A. Notation

Most of the notation in this paper follows that in [7]. By $X \sim P(x)$, we mean X is distributed according to P. The empirical distribution (or type) of a sequence $x^n \in \mathcal{X}^n$ is denoted by $\hat{P}_{x^n}$. Joint empirical distributions are denoted similarly. We say a sequence x^n has type P if $\hat{P}_{x^n} = P$ and denote it by $x^n \in T_P^n$, where T_P^n or more simply T_P is the set of all sequences that have type P. We use $\mathcal{P}^{\mathcal{X}}$ to denote the set of distributions over the finite alphabet $\mathcal{X}$. The set of sequences y^n that have a conditional type W given x^n is denoted by $T_W(x^n)$. The Kullback-Leibler divergence is denoted by $D(P||Q)$. We use $o(\cdot)$ to denote quantities that grow strictly slower than their arguments. In this paper, we use the convention that $\binom{n}{k}=0$ if $k<0$ or $n<k$, and the entropy $H(P)=-\infty$ if P is not a probability mass function, i.e., one of its elements is negative or the sum of its elements is larger than one. $h(\cdot)$ is the binary entropy function, and for $\beta_1+\beta_2<1$, let $h(\beta_1,\beta_2)$ denote the entropy of the ternary probability mass function $(\beta_1,\beta_2,1-\beta_1-\beta_2)$. Finally, if $0\le\rho\le 1$, then $\bar{\rho} := 1-\rho$.

B. Partial Divergence and Its Generalization

Partial divergence $d_\rho(P||Q)$ between distributions P and Q with mismatched factor ρ is introduced in [4], [5] to characterize the exponent of the probability that a sequence with length k has a type P if ρk of its elements are generated independently according to Q and $\bar{\rho}k$ of them are generated independently according to P. For alphabets of size t, e.g., $\mathcal{X}=\{0,1,...,t-1\}$, and distributions $P,Q\in\mathcal{P}^{\mathcal{X}}$, where $P:=(p_0,p_1,...,p_{t-1})$, and $Q:=(q_0,q_1,...,q_{t-1})$, partial divergence can be expressed as [5]

$$d_\rho(P||Q) = D(P||Q) - \sum_{j=0}^{t-1} p_j \log(c^* + \frac{p_j}{q_j}) + \rho\log c^* + h(\rho), \tag{1}$$

where c^* is a function of ρ, P, and Q, and can be uniquely determined from

$$c^* \sum_{j=0}^{t-1} \frac{p_j q_j}{c^* q_j + p_j} = \rho.$$

We now state a generalization for [4, Lemma 1] for which the sequence is generated according to three distributions.

Lemma 1. *Consider the alphabet $\mathcal{X}=\{0,1,...,t-1\}$, and distributions $P,Q_1,Q_2,Q_3\in\mathcal{P}^{\mathcal{X}}$. A random sequence X^k is generated as follows: $\rho_1 k$ symbols are i.i.d. according to Q_1, $\rho_2 k$ symbols are i.i.d. according to Q_2, and $\rho_3 k$ are i.i.d. according to Q_3, where $\rho_1+\rho_2+\rho_3=1$. Then, the exponent of the probability that X^k has type P is*

$$\begin{aligned}&\lim_{k\to\infty} -\frac{1}{k}\log\mathbb{P}(X^k\in T_P)\\ =& \min_{\substack{P_1,P_2,P_3\in\mathcal{P}^{\mathcal{X}}:\\ \rho_1P_1+\rho_2P_2+\rho_3P_3=P}} \rho_1 D(P_1||Q_1)+\rho_2 D(P_2||Q_2)+\rho_3 D(P_3||Q_3)\end{aligned} \tag{2}$$

The proof is omitted due to space considerations. We will be interested in a special case of Lemma 1 in which $Q_3 = P$. In other words, we need to find the exponent of the probability that a sequence has a type P if its elements are generated independently according to Q_1, Q_2, and P. For this case, we denote the right-hand side of (2) by $d_{\rho_1,\rho_2}(P||Q_1,Q_2)$, where $\rho_1+\rho_2<1$. This function will be used in Section V-B.

C. Deletion Channels

As we treat collisions as deletions in Section V-A, we state some results on deletion channels in this section. There is considerable work concentrating on achievability results for the deletion channel [8]–[10]. In [8], in addition to i.i.d. codewords with uniform distribution over the alphabet, codewords from first order Markov chains are used to improve the achievability results for deletion channels. However, we require i.i.d. codewords in order to simplify the analysis of the probability of error for the decoding algorithm used in Section V-A. We first introduce a different model for the noisy deletion channel for which the number of deleted symbols is assumed to be fixed, and then give an achievability result, which is similar to the one in [8], but allows for arbitrary input distribution rather than the uniform one, and is valid for a general discrete memoryless channel (DMC) rather than the symmetric one. Note that our result is also valid for the i.i.d. deletion channel.

Consider the cascade of a deletion channel with a DMC, where the deletion channel deletes d symbols of its input sequence of length $k\ge d$ arbitrarily at random so that the output of the deletion channel and the DMC has length $k-d$, where $\theta := d/k \le 1$ is the ratio of the deleted symbols to the codeword length. Let $\mathcal{X}$, $\mathcal{Y}$, W_s, and C_s denote the input alphabet, output alphabet, probability transition matrix, and the capacity of the DMC, respectively. After stating the following lemma from [11], we state the achievability result.

Lemma 2. *[11] For a given $|\mathcal{Y}|$-ary sequence y^{k-d} of length $k-d$, the number of $|\mathcal{Y}|$-ary sequences of length k which contain sequence y^{k-d} as a subsequence is given by*

$$\sum_{j=k-d}^{d}\binom{k}{j}(|\mathcal{Y}|-1)^{d-j} \le k\binom{k}{k-d}(|\mathcal{Y}|-1)^d$$

Lemma 3. *For the noisy deletion channel described above, rates not exceeding $C_s - h(\theta) - \theta\log(|\mathcal{Y}|-1)$ are achievable.*

Proof. Encoding: Fix an input distribution P. Randomly and independently generate e^{kR} sequences $c^k(m)$, $m\in$

$\{1, 2, ..., e^{kR}\}$ each i.i.d. according to P. To send message m, the encoder transmits $c^k(m)$.

Decoding: The decoder observes the output sequence y^{k-d} and constructs all possible $\tilde{y}^k$ which contain sequence y^{k-d} as a subsequence. Then it checks if any of these sequences are jointly typical with any of the codewords, i.e., if $\tilde{y}^k \in T_{[W_s]_\mu}(c^k(\hat{m}))$, then we declare that message as being sent. If this condition is not satisfied for any of the sequences $\tilde{y}^k$ and any of the messages, then the decoder declares an error.

Analysis of the probability of error: For any $\epsilon > 0$, we prove that if $R = \mathbb{I}(X;Y) - h(\theta) - \theta \log(|\mathcal{Y}| - 1) - 2\epsilon$, then the average probability of error vanishes as $k \to \infty$. Considering uniform distribution on the messages and assuming that the message $m = 1$ is transmitted, we have

$$p_e^{avg} \leq \mathbb{P}(\hat{m}{=}e|m{=}1) + \mathbb{P}(\hat{m} \in \{2, 3, ..., e^{kR}\}|m{=}1), \quad (3)$$

where (3) follows from the union bound in which the first term is the probability that the decoder declares an error, i.e. does not find any codeword being jointly typical with any of the possible sequences $\tilde{y}^k$, which contain sequence y^{k-d} as a subsequence. This implies that even if the correct deletion pattern is considered and all possible choices for the deleted symbols are evaluated, none of them are jointly typical with $c^k(1)$. The probability of this event vanishes as $k \to \infty$ according to [7, Lemma 2.12]. Using Lemma 2 and the union bound for all possible $\tilde{y}^k$'s and all the messages $\hat{m} \neq 1$, we have

$$\begin{aligned}
&\mathbb{P}(\hat{m} \in \{2, 3, ..., e^{kR}\}|m = 1) \\
&\leq k \binom{k}{k-d} (|\mathcal{Y}|-1)^d (e^{kR}-1) \mathbb{P}(\tilde{Y}^k \in T_{[W_s]_\mu}(c^k(2))|m{=}1) \\
&\leq e^{o(k)} e^{k(h(\theta)+\theta \log(|\mathcal{Y}|-1))} e^{kR} e^{-k(\mathbb{I}(X;Y)-\epsilon)} \quad (4) \\
&= e^{o(k)} e^{-k\epsilon}, \quad (5)
\end{aligned}$$

where (4) results from Stirling's approximation and the packing lemma [12, Lemma 3.1] since conditioned on message $m = 1$ being sent, $C^k(2)$ and $\tilde{Y}^k$ are independent; and where (5) follows by substituting $R = \mathbb{I}(X;Y) - h(\theta) - \theta \log(|\mathcal{Y}| - 1) - 2\epsilon$. Therefore, the second term in (3) also vanishes as $k \to \infty$, and the lemma is proved by considering the capacity achieving input distribution for the DMC. □

IV. Random Access without Collisions

In this section, we focus on an intermittent process in Figure 1 that models a random access channel in which at each time slot exactly one of the users sends an information symbol and the other remains silent by sending the noise symbol $\star$, until both users send their codewords. In this model, the output pair (x_1, x_2) of the intermittent process at each time slot takes one of the two following forms: $(c_1, \star)$ or $(\star, c_2)$, where c_1 and c_2 denote a code symbol from the first and the second user, respectively. The length of the receive window in this model is $n = 2k$. The receiver observes the sequence y^n, wishes to decode both messages, but does not know a priori which output symbol corresponds to which user's codeword. Motivating examples include a cognitive radio application in which the primary user is bursty, i.e., sends information symbols in some time slots and remains silent in the other time slots, and a secondary user also wants to communicate with the same receiver and can sense the channel and transmit its information symbols whenever the first user is silent. As another motivating example, consider an ALOHA random access protocol with a collision-avoidance mechanism. In the following theorem, we obtain an achievable rate region for (R_1, R_2).

Theorem 1. *For intermittent multi-access communication with the intermittent process described above, rates (R_1, R_2) satisfying*

$$R_1 < \mathbb{I}(X_1; Y|x_2 = \star) - f_1(P_1, P_2, W) \quad (6)$$
$$R_2 < \mathbb{I}(X_2; Y|x_1 = \star) - f_1(P_1, P_2, W) \quad (7)$$

are achievable for any $(X_1, X_2) \sim P_1(x_1)P_2(x_2)$, where

$$\begin{aligned}
f_1(P_1, P_2, W) := \max_{0 \leq \beta \leq 1} \{&2h(\beta) - d_\beta(P_1 W_{\cdot\star} || P_2 W_{\star\cdot}) \\
&- d_\beta(P_2 W_{\star\cdot} || P_1 W_{\cdot\star})\}, \quad (8)
\end{aligned}$$

and $d_\cdot(\cdot||\cdot)$ is the partial divergence given in (1).

Proof. Encoding: Fix two input distributions P_1 and P_2 for user 1 and user 2, respectively. Randomly and independently generate e^{kR_1} sequences $c_1^k(m_1)$, $m_1 \in \{1, 2, ..., e^{kR_1}\}$ each i.i.d. according to P_1 for user 1, and e^{kR_2} sequences $c_2^k(m_2)$, $m_2 \in \{1, 2, ..., e^{kR_2}\}$ each i.i.d. according to P_2 for user 2. To send message m_1, encoder 1 transmits $c_1^k(m_1)$, and to send message m_2, encoder 2 transmits $c_2^k(m_2)$.

Decoding: Similar to decoding from pattern detection described in [5], the decoder chooses k of the $2k$ output symbols y^{2k}. Let $\tilde{y}^k$ denote the sequence of the chosen symbols, and $\hat{y}^k$ denote the sequence of the other k symbols. For each choice, there are two stages. In the first stage, the decoder checks if $\tilde{y}^k$ is induced by user 1, i.e., if $\tilde{y}^k \in T_{P_1 W_{\cdot\star}}$, and if $\hat{y}^k$ is induced by user 2, i.e., if $\hat{y}^k \in T_{P_2 W_{\star\cdot}}$. If both of these conditions are satisfied, then we proceed to the second stage; otherwise, we make another choice for the k symbols and restart the two-stage decoding procedure. In the second stage, we perform joint typicality decoding with a fixed typicality parameter $\mu > 0$ for both sequences $\tilde{y}^k$ and $\hat{y}^k$, i.e., if $\tilde{y}^k \in T_{[W_{\cdot\star}]_\mu}(c_1^k(\hat{m}_1))$ and $\hat{y}^k \in T_{[W_{\star\cdot}]_\mu}(c_2^k(\hat{m}_2))$ for a unique message pair $(\hat{m}_1, \hat{m}_2)$, then we declare them as the transmitted messages; otherwise, we make another choice for the k symbols and repeat the two-stage decoding procedure. If at the end of all $\binom{2k}{k}$ choices the typicality decoding procedure has not declared any message pair as being sent, then the decoder declares an error.

Analysis of the probability of error: For any $\epsilon > 0$, we prove that if $R_1 = \mathbb{I}(X_1; Y|x_2 = \star) - f_1(P_1, P_2, W) - 2\epsilon$, and $R_2 = \mathbb{I}(X_2; Y|x_1 = \star) - f_1(P_1, P_2, W) - 2\epsilon$, then the average probability of error vanishes as $k \to \infty$. Considering uniform distribution on the messages and assuming that the message pair $(1,1)$ is transmitted, we have

$$\begin{aligned} p_e^{avg} \leq & \mathbb{P}((\hat{m}_1, \hat{m}_2) = e | (m_1, m_2) = (1,1)) \\ & + \mathbb{P}(\hat{m}_1 \in \{2,3,...,e^{kR_1}\} | (m_1, m_2) = (1,1)) \\ & + \mathbb{P}(\hat{m}_2 \in \{2,3,...,e^{kR_2}\} | (m_1, m_2) = (1,1)), \quad (9) \end{aligned}$$

where (9) follows from the union bound in which the first term is the probability that the decoder declares an error (does not find any message pair) at the end of all $\binom{2k}{k}$ choices, which implies that even if we pick the correct output symbols corresponding to user 1 and user 2, the decoder either does not pass the first stage or does not declare $(\hat{m}_1, \hat{m}_2) = (1,1)$ in the second stage. The probability of this event vanishes as $k \to \infty$ according to [7, Lemma 2.12].

The second term in (9) is the probability that for at least one choice of the output symbols, the decoder passes the first stage, and then in the second stage, it declares an incorrect message for user 1. We characterize the $\binom{2k}{k}$ choices based on the number of incorrectly chosen output symbols, which is denoted by k_1, i.e., the number of symbols in $\tilde{y}^k$ that are in fact output symbols corresponding to the second user, which is equal to the number of symbols in $\hat{y}^{n-k}$ that are in fact output symbols corresponding to the first user. For any $0 \leq k_1 \leq k$, there are $\binom{k}{k_1}\binom{k}{k_1}$ possible choices. Using the union bound for all the choices and all the messages $\hat{m}_1 \neq 1$, we have

$$\begin{aligned} & \mathbb{P}(\hat{m}_1 \in \{2,3,...,e^{kR_1}\} | (m_1, m_2) = (1,1)) \\ & \leq (e^{kR_1} - 1) \sum_{k_1=0}^{k} \binom{k}{k_1} \binom{k}{k_1} \mathbb{P}_{k_1}(\hat{m}_1 = 2 | (m_1, m_2) = (1,1)), \end{aligned} \tag{10}$$

where the index k_1 in (10) denotes the condition that the number of wrongly chosen output symbols is k_1. Note that message $\hat{m}_1 = 2$ is declared at the decoder only if the choice of the output symbols passes the first stage, and then the condition $\tilde{y}^k \in T_{[W_{\cdot\star}]_\mu}(c_1^k(2))$ is satisfied. Therefore,

$$\begin{aligned} & \mathbb{P}_{k_1}(\hat{m}_1 = 2 | (m_1, m_2) = (1,1)) \\ & = \mathbb{P}_{k_1}\Big(\{\tilde{Y}^k \in T_{P_1 W_{\cdot\star}}\} \cap \{\hat{Y}^k \in T_{P_2 W_{\star\cdot}}\} \\ & \qquad \cap \{\tilde{Y}^k \in T_{[W_{\cdot\star}]_\mu}(c_1^k(2))\} | (m_1, m_2) = (1,1) \Big) \\ & = \mathbb{P}_{k_1}(\tilde{Y}^k \in T_{P_1 W_{\cdot\star}}) \cdot \mathbb{P}_{k_1}(\hat{Y}^k \in T_{P_2 W_{\star\cdot}}) \\ & \quad \cdot \mathbb{P}(\tilde{Y}^k \in T_{[W_{\cdot\star}]_\mu}(c_1^k(2)) | (m_1, m_2) = (1,1)) \qquad (11) \\ & \leq e^{o(k)} e^{-k d_{k_1/k}(P_1 W_{\cdot\star} || P_2 W_{\star\cdot})} e^{-k d_{k_1/k}(P_2 W_{\star\cdot} || P_1 W_{\cdot\star})} \\ & \quad \cdot e^{-k(\mathbb{I}(X_1;Y|x_2=\star) - \epsilon)}, \qquad (12) \end{aligned}$$

where (11) follows from the independence of the events $\{\tilde{Y}^k \in T_{P_1 W_{\cdot\star}}\}$ and $\{\hat{Y}^k \in T_{P_2 W_{\star\cdot}}\}$ conditioned on k_1 wrongly chosen output symbols, and (12) follows from the results on the partial divergence in Section III-B for the first two terms in (11) with mismatch ratios k_1/k, and using the packing lemma [12, Lemma 3.1] for the last term in (11), because conditioned on message $m_1 = 1$ being sent, $C_1^k(2)$ and $\tilde{Y}^k$ are independent regardless of the number of wrongly chosen output symbols. Substituting (12) into the summation in (10), using Stirling's approximation for the terms $\binom{k}{k_1}$, and finding the largest exponent of the terms in the summation, we have

$$\begin{aligned} & \mathbb{P}(\hat{m}_1 \in \{2,3,...,e^{kR_1}\} | (m_1, m_2) = (1,1)) \\ & \leq e^{kR_1} e^{o(k)} e^{k f_1(P_1, P_2, W)} e^{-k(\mathbb{I}(X_1;Y|x_2=\star) - \epsilon)} \\ & = e^{o(k)} e^{-k\epsilon}, \end{aligned} \tag{13}$$

where (13) is obtained by substituting $R_1 = \mathbb{I}(X_1;Y|x_2 = \star) - f_1(P_1, P_2, W) - 2\epsilon$. Therefore, the second term in (9) vanishes as $k \to \infty$. Similarly, the third term in (9) also vanishes as $k \to \infty$, which proves the Theorem. ☐

The function $f_1(P_1, P_2, W)$ can be interpreted as an overhead term due to the system's burstiness or intermittency. Note that the result in Theorem 1 implies that there is a tradeoff between the two terms in (6) and in (7) by choosing the input distributions P_1 and P_2. In order to maximize the first terms we need to choose the capacity achieving input distributions, but at the same time, it is desirable to choose input distributions such that the two distributions $P_1 W_{\cdot\star}$ and $P_2 W_{\star\cdot}$ have the largest distance to maximize the partial divergences $d_\beta(P_1 W_{\cdot\star} || P_2 W_{\star\cdot})$ and $d_\beta(P_2 W_{\star\cdot} || P_1 W_{\cdot\star})$ so that we have a smaller overhead term $f_1(P_1, P_2, W)$. Also, note that both rates R_1 and R_2 have the same overhead cost for fixed input distributions P_1 and P_2. This is not the case if we consider different codeword lengths for the two users.

V. Random Access with Collisions

In this section, we focus on an intermittent process in Figure 1 that models a random access channel with collisions. In principle, we can consider a random access channel that allows for both idle-times and collisions, where idle times can be handled using a similar generalization of the partial divergence stated in Lemma 1. However, we assume that there are no idle times in order to avoid further complexity of the results. In this model, the output pair (x_1, x_2) of the intermittent process in each time slot takes one of the three following forms: $(c_1, \star)$, $(\star, c_2)$, or a collision, where two different models for collisions are considered: deletion and interference. We assume the total number of collisions is $d \leq k$. Let $\theta := d/k \leq 1$ denote the ratio of the collided symbols of each user to the codeword length.

A. Collisions as Deletions

In this section, we treat collisions as deletions. We assume that the output of the intermittent process with length $n = 2(k-d)$ consists of $k-d$ of the pair $(c_1, \star)$, $k-d$ of the pair $(\star, c_2)$, and d collided symbols that are deleted from the output sequence. The encoders and the decoder do not know the positions.

Theorem 2. *For intermittent multi-access communication with the intermittent process described above, rates (R_1, R_2) sat-*

isfying

$$R_1 < \mathbb{I}(X_1;Y|x_2=\star) - f_2(P_1,P_2,W,\theta)$$
$$R_2 < \mathbb{I}(X_2;Y|x_1=\star) - f_2(P_1,P_2,W,\theta)$$

are achievable for any $(X_1,X_2) \sim P_1(x_1)P_2(x_2)$*, where*

$$f_2(P_1,P_2,W,\theta) := (1-\theta)f_1(P_1,P_2,W) + h(\theta) + \theta\log(|\mathcal{Y}|-1), \quad (14)$$

where $f_1(P_1,P_2,W)$ *is given in* (8).

The decoding scheme and the techniques for the analysis of the probability of error are a combination of those in the proof of Lemma 3 and Theorem 1. The complete proof is omitted due to space considerations.

B. Collisions as Interference

In this section, we treat collisions as interference. We assume that the output of the intermittent process with length $n = 2k-d$ consists of $k-d$ of the pair $(c_1,\star)$, $k-d$ of the pair $(\star,c_2)$, and d of the pair (c_1,c_2). The encoders and the decoder do not know the positions. User 1 and user 2 transmit $k-d$ information symbols over a point to point channel, $W_{\cdot,\star}$ and $W_{\star,\cdot}$, respectively, and transmit d information symbols over the MAC channel W, through which there is interference between the users.

Theorem 3. *For intermittent multi-access communication with the intermittent process described above, rates* (R_1,R_2) *satisfying*

$$R_1 < \bar{\theta}\mathbb{I}(X_1;Y|x_2=\star) + \theta\mathbb{I}(X_1;Y|X_2) - f_3(P_1,P_2,W,\theta)$$
$$R_2 < \bar{\theta}\mathbb{I}(X_2;Y|x_1=\star) + \theta\mathbb{I}(X_2;Y|X_1) - f_3(P_1,P_2,W,\theta)$$
$$R_1+R_2 < \bar{\theta}\mathbb{I}(X_1;Y|x_2=\star) + \bar{\theta}\mathbb{I}(X_2;Y|x_1=\star) + \theta\mathbb{I}(X_1,X_2;Y) - f_3(P_1,P_2,W,\theta)$$

are achievable for any $(X_1,X_2) \sim P_1(x_1)P_2(x_2)$*, where*

$$f_3(P_1,P_2,W,\theta) := \max_{\substack{0\le\beta_1+\beta_2\le 1\\ 0\le\beta_1'+\beta_2'\le 1}} \{\bar{\theta}h(\beta_1,\beta_2) + \bar{\theta}h(\beta_1',\beta_2') + \theta h(\frac{\bar{\theta}(\beta_1+\beta_2-\beta_2')}{\theta}, \frac{\bar{\theta}(\beta_1'+\beta_2'-\beta_2)}{\theta}) - \bar{\theta}d_{\beta_1,\beta_2}(P_1W_{\cdot\star}||P_1P_2W, P_2W_{\star\cdot}) - \bar{\theta}d_{\beta_1',\beta_2'}(P_2W_{\star\cdot}||P_1P_2W, P_1W_{\cdot\star}) - \theta d_{(\beta_1+\beta_2-\beta_2')\bar{\theta}/\theta, (\beta_1'+\beta_2'-\beta_2)\bar{\theta}/\theta}(P_1P_2W||P_1W_{\cdot\star}, P_2W_{\star\cdot})\}, \quad (15)$$

and $d_{\cdot,\cdot}(\cdot||\cdot,\cdot)$ *is the function defined in Section III-B.*

Sketch of the Proof: Encoding is the same as in the proof of Theorem 1. We briefly explain the decoding procedure. The analysis of the probability of error is lengthy and is omitted due to space considerations.

Decoding: The decoder splits the output sequence y^{2k-d} into three subsequences of length $k-d$, $k-d$, and d, and denotes them by $\tilde{y}_1^{k-d}$, $\tilde{y}_2^{k-d}$, and $\hat{y}^d$, respectively. For each choice, there are two stages. In the first stage, we check three conditions: $\tilde{y}_1^{k-d} \in T_{P_1W_{\cdot\star}}$, $\tilde{y}_2^{k-d} \in T_{P_2W_{\star\cdot}}$, and $\hat{y}^d \in T_{P_1P_2W}$. If all three conditions are satisfied, then we proceed to the second stage; otherwise, we make another choice for the three output subsequences and restart the two-stage decoding procedure.

In the second stage, we perform simultaneous joint typicality decoding. We first split all of the codewords as follows. Let $\tilde{c}_1^{k-d}(m_1)$ and $\hat{c}_1^d(m_1)$ be the subsequences of $c_1^k(m_1)$ corresponding to the positions of the symbols of the chosen subsequences $\tilde{y}_1^{k-d}$ and $\hat{y}^d$, respectively. Similarly, let $\tilde{c}_2^{k-d}(m_2)$ and $\hat{c}_2^d(m_2)$ be the subsequences of $c_2^k(m_2)$ corresponding to the positions of the symbols of the chosen subsequences $\tilde{y}_2^{k-d}$ and $\hat{y}^d$, respectively. We declare the message pair $(\hat{m}_1,\hat{m}_2)$ as being transmitted if it is the unique message pair such that the following three conditions are satisfied simultaneously: $(\tilde{c}_1^{k-d}(\hat{m}_1), \tilde{y}_1^{k-d})$ is jointly typical; $(\tilde{c}_2^{k-d}(\hat{m}_2), \tilde{y}_2^{k-d})$ is jointly typical; and $(\hat{c}_1^d(\hat{m}_1), \hat{c}_2^d(\hat{m}_2), \hat{y}^d)$ is jointly typical; otherwise, we make another choice for the three output subsequences and repeat the two-stage decoding procedure. If at the end of all $\binom{2k-d}{k-d,k-d,d}$ choices the typicality decoding procedure has not declared any message pair as being sent, then the decoder declares an error. $\square$

Acknowledgment

The authors wish to thank Dr. Aslan Tchamkerten and Shyam Kumar for the useful discussion, which leaded to a more compact expression and a shorter proof for the result in Lemma 1.

References

[1] R. Gallager, "A perspective on multiaccess channels," *IEEE Trans. on Inf. Theory,* vol. 31, no. 2, pp. 124-142, Mar. 1985.

[2] A. Ephremides and B. Hajek, "Information theory and communication networks: An unconsummated union," *IEEE Trans. on Inf. Theory,* vol. 44, no. 6, pp.2416-2434, Oct 1998.

[3] P. Minero, M. Franceschetti, and D. N. C. Tse, "Random access: An information-theoretic perspective," *IEEE Trans. on Inf. Theory,* vol. 58, no. 2, pp. 909-930, Feb. 2012.

[4] M. Khoshnevisan and J. N. Laneman, "Achievable Rates for Intermittent Communication," in *Proc. IEEE Int. Symp. Information Theory (ISIT),* pp. 1346-1350, Cambridge, MA, USA, July 2012.

[5] M. Khoshnevisan and J. N. Laneman, "Intermittent Communication and Partial Divergence," in *Proc. Allerton Conf. Communications, Control, and Computing,* Monticello, IL, Oct. 2012.

[6] B. P. Dunn, "Overhead in Communication Systems as the Cost of Constraints," *Ph.D. dissertation,* University of Notre Dame, Notre Dame, IN, Dec. 2010.

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

[8] S. Diggavi and M. Grossglauser, "On information transmission over a finite buffer channel," *IEEE Trans. on Inf. Theory,* vol. 52, no. 3, pp. 1226-1237, Mar. 2006.

[9] E. Drinea and M. Mitzenmacher, "Improved lower bounds for the capacity of i.i.d. deletion and duplication channels," *IEEE Trans. on Inf. Theory,* vol. 53, no. 8, pp. 2693-2714, Aug. 2007.

[10] A. Kirsch and E. Drinea, "Directly lower bounding the information capacity for channels with i.i.d. deletions and duplications," *IEEE Trans. on Inf. Theory,* vol. 56, no. 1, pp. 86-102, Jan. 2010.

[11] V. Chvatal and D. Sankoff, "Lomgest Common Sunsequence of Two Random Sequences," *J. Appl.Probab.,* no. 12, pp. 306-315, 1975.

[12] A. El Gamal and Y.-H. Kim, *Network Information Theory*, Cambridge University Press, UK, 2012.