Achievable Rates for Intermittent Multi-Access Communication

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

***Abstract*—We formulate 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. For each model, we characterize the performance of the system in terms of achievable rate regions. The intermittency of the system comes with a significant cost in our achievable schemes.**

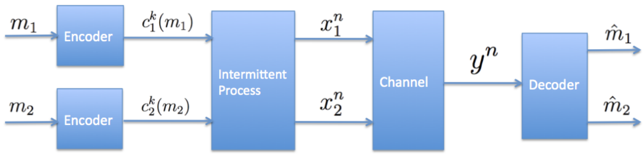

Fig. 1. System model for intermittent multi-access communication.

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 to 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]. By making different assumptions for the intermittent process, we specialize the system to two models: random access with collision avoidance, and random access with collisions in Sections IV and V, respectively. The collisions are treated as interference, and information can be extracted from the collided symbols.

For each model, we obtain achievable rate regions that 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 both, the decoder has to both detect the positions and decode the messages. In our achievable schemes, 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.

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 a special symbol, corresponding to the input of the channel when the 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 is 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 are 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. 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 stores inputs $c_1^k(m_1)$ and $c_2^k(m_2)$ in two separate buffers, and generates outputs x_1^n and x_2^n to capture the burstiness and the asynchronism of the users. The intermittent process, in general, has memory, and can be described as a state-dependent process with four possible states (s_1, s_2), $s_1, s_2 \in \{0, 1\}$ in each time slot. If $s_i = 0$, then user i is silent and transmits the symbol $\star$. If $s_i = 1$, which is only possible if there are codeword symbols remaining in user i's buffer, then user i transmits the next codeword symbol. We assume that neither the encoders nor the decoder know the states of the intermittent process. Note that the intermittent process together with the DM-MAC can be collected into a state-dependent MAC with memory with the states unknown to the encoders and the decoder. See [7], [8] and the references therein for some setups on memoryless state-dependent MAC.

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 $\frac{1}{M_1M_2}\sum_{m_1=1}^{M_1}\sum_{m_2=1}^{M_2}\mathbb{P}((\hat{m}_1, \hat{m}_2) \neq (m_1, m_2)) \to 0$ as $k \to \infty$. We refer to this general scenario as intermittent multi-access communication, and in Sections IV and V we consider several instances of the intermittent process in Figure 1. But first, Section III reviews some prerequisite material upon which the results are based.

III. Preliminaries

A. Notation

Most of the notation in this paper follows that of [9]. 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]

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

where c^* is a function of ρ, P, and Q that 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*

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

Proof. See Appendix A. $\square$

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.

IV. Random Access with Collision Avoidance

In this section, we consider 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 special symbol $\star$, until both users have finished sending their codewords. In this model, there are only two possible states for the intermittent process $(s_1, s_2) \in \{(1, 0), (0, 1)\}$, and therefore, 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 the next codeword symbol to be transmitted from the first and the second user, respectively. Note that if both input buffers of the intermittent process are empty, then the transmission terminates, and if exactly one of them is empty, then only the state corresponding to transmission of the codeword symbol from the user with the non-empty buffer is allowed. As a result, 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.

A potential application of this model include a cognitive radio 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. In the following theorem, we obtain an achievable rate region for (R_1, R_2).

Theorem 1. *For intermittent multi-access communication with collision avoidance, rates (R_1, R_2) satisfying*

$$R_1 < \mathbb{I}(X_1; Y|X_2 = \star) - f_1(P_1, P_2, W) \tag{3}$$

$$R_2 < \mathbb{I}(X_2; Y|X_1 = \star) - f_1(P_1, P_2, W) \tag{4}$$

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

$$\begin{aligned} f_1(P_1, P_2, W) := \max_{0\le\beta\le1} \{&2h(\beta) - d_\beta(P_1W_{\cdot\star}\|P_2W_{\star\cdot}) \\ &- d_\beta(P_2W_{\star\cdot}\|P_1W_{\cdot\star})\}, \end{aligned} \tag{5}$$

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

Remark 1. *The result in Theorem 1 is valid for the intermittent process described above with arbitrary probability distribution on the time slots that each user transmits. As a special case, we might think of an intermittent process in which at each time slot $\mathbb{P}((S_1, S_2) = (1, 0)) = \mathbb{P}((S_1, S_2) = (0, 1)) = 1/2$ if both buffers are non-empty; otherwise only the user with the non-empty buffer transmits. Note that the length of the receive window remains $2k$ in any case, since each codeword has length k and there are no collisions.*

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 chosen symbols, and $\hat{y}^k$ denote 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 $\tilde{y}^k \in T_{P_1W_{\cdot\star}}$, and if $\hat{y}^k$ is induced by user 2, i.e., if $\hat{y}^k \in T_{P_2W_{\star\cdot}}$. 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 independent uniform distributions on the messages and assuming that the message pair $(1, 1)$ is transmitted, we have

$$\begin{aligned} p_e^{avg} \le& \mathbb{P}((\hat{m}_1, \hat{m}_2) = e|(m_1, m_2) = (1, 1)) \\ &+ \mathbb{P}(\hat{m}_1 \in \{2, 3, ..., e^{kR_1}\}|(m_1, m_2) = (1, 1)) \\ &+ \mathbb{P}(\hat{m}_2 \in \{2, 3, ..., e^{kR_2}\}|(m_1, m_2) = (1, 1)), \end{aligned} \tag{6}$$

where (6) 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 [9, Lemma 2.12].

The second term in (6) 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 $0 \le k_1 \le k$, there are $\binom{k}{k_1}\binom{k}{k_1}$ possible choices. Using the union bound for all the choices and all the messages $\hat{m}_1 \ne 1$, we have

$$\begin{aligned} &\mathbb{P}(\hat{m}_1 \in \{2, 3, ..., e^{kR_1}\}|(m_1, m_2) = (1, 1)) \\ &\le (e^{kR_1} - 1) \sum_{k_1=0}^{k} \binom{k}{k_1}\binom{k}{k_1} \mathbb{P}_{k_1}(\hat{m}_1{=}2|(m_1, m_2){=}(1, 1)), \end{aligned} \tag{7}$$

where the index k_1 in (7) 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,

$$\begin{aligned} &\mathbb{P}_{k_1}(\hat{m}_1{=}2|(m_1, m_2){=}(1, 1)) \\ &= \mathbb{P}_{k_1}\Big(\{\tilde{Y}^k \in T_{P_1W_{\cdot\star}}\} \cap \{\hat{Y}^k \in T_{P_2W_{\star\cdot}}\} \\ &\qquad \cap \{\tilde{Y}^k \in T_{[W_{\cdot\star}]_\mu}(c_1^k(2))\}|(m_1, m_2){=}(1, 1)\Big) \\ &= \mathbb{P}_{k_1}(\tilde{Y}^k \in T_{P_1W_{\cdot\star}}) \cdot \mathbb{P}_{k_1}(\hat{Y}^k \in T_{P_2W_{\star\cdot}}) \\ &\quad \cdot \mathbb{P}(\tilde{Y}^k \in T_{[W_{\cdot\star}]_\mu}(c_1^k(2))|(m_1, m_2){=}(1, 1)) \end{aligned} \tag{8}$$

$$\begin{aligned} &\le e^{o(k)} e^{-kd_{k_1/k}(P_1W_{\cdot\star}\|P_2W_{\star\cdot})} e^{-kd_{k_1/k}(P_2W_{\star\cdot}\|P_1W_{\cdot\star})} \\ &\quad \cdot e^{-k(\mathbb{I}(X_1;Y|X_2=\star)-\epsilon)}, \end{aligned} \tag{9}$$

where (8) follows from the independence of the events $\{\tilde{Y}^k \in T_{P_1W_{\cdot\star}}\}$ and $\{\hat{Y}^k \in T_{P_2W_{\star\cdot}}\}$ conditioned on k_1 (a fixed number of) wrongly chosen output symbols, and (9) follows from the results on the partial divergence in Section III-B for the first two terms in (8) with mismatch ratios k_1/k, and using the packing lemma [7, Lemma 3.1] for the last term in (8), 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 (9) into the summation in (7), using Stirling's approximation for the terms $\binom{k}{k_1}$, and finding the largest exponent of the terms in the summation, we have

$$\begin{aligned}
&\mathbb{P}(\hat{m}_1 \in \{2,3,...,e^{kR_1}\} | (m_1,m_2)=(1,1)) \\
&\le e^{kR_1} e^{o(k)} e^{kf_1(P_1,P_2,W)} e^{-k(\mathbb{I}(X_1;Y|X_2=\star)-\epsilon)} \\
&= e^{o(k)} e^{-k\epsilon}, \qquad (10)
\end{aligned}$$

where (10) 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 (6) vanishes as $k \to \infty$. Similarly, the third term in (6) also vanishes as $k \to \infty$, which proves the theorem. $\square$

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 (3) and in (4) 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 $P_1W_{\cdot\star}$ and $P_2W_{\star\cdot}$ have the largest distance to maximize the partial divergences $d_\beta(P_1W_{\cdot\star}\|P_2W_{\star\cdot})$ and $d_\beta(P_2W_{\star\cdot}\|P_1W_{\cdot\star})$ 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 consider 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, corresponding to the state $(s_1,s_2)=(0,0)$ of the intermittent process, can be handled using a similar generalization of the partial divergence result stated in Lemma 1. However, we assume that there are no idle times in order to avoid overcomplicating the results. In this model, there are three possible states for the intermittent process $(s_1,s_2) \in \{(1,0),(0,1),(1,1)\}$, where the total number of states representing a collision, i.e., $(s_1,s_2)=(1,1)$, is assumed to be $d \le k$. Therefore, the output pair (x_1,x_2) of the intermittent process with length $n = 2k-d$ consists of $k-d$ of the form $(c_1,\star)$, $k-d$ of the form $(\star,c_2)$, and d of the form (c_1,c_2). In other words, 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, but the decoder does not know a priori these positions. Let $\theta := d/k \le 1$ denote the ratio of the collided symbols of each user to the codeword length. In the following theorem, we obtain an achievable rate region for (R_1,R_2).

Theorem 2. *For intermittent multi-access communication with collisions, rates (R_1,R_2) satisfying*

$$\begin{aligned}
R_1 &< \bar{\theta}\mathbb{I}(X_1;Y|X_2=\star) + \theta\mathbb{I}(X_1;Y|X_2) - f_2(P_1,P_2,W,\theta) \\
R_2 &< \bar{\theta}\mathbb{I}(X_2;Y|X_1=\star) + \theta\mathbb{I}(X_2;Y|X_1) - f_2(P_1,P_2,W,\theta) \\
R_1+R_2 &< \bar{\theta}\mathbb{I}(X_1;Y|X_2=\star) + \bar{\theta}\mathbb{I}(X_2;Y|X_1=\star) \\
&\quad + \theta\mathbb{I}(X_1,X_2;Y) - f_2(P_1,P_2,W,\theta)
\end{aligned}$$

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

$$\begin{aligned}
&f_2(P_1,P_2,W,\theta) := \\
&\max_{\substack{0\le\beta_1+\beta_2\le 1\\ 0\le\beta_1'+\beta_2'\le 1}} \Big\{ \bar{\theta}h(\beta_1,\beta_2) + \bar{\theta}h(\beta_1',\beta_2') + \theta h\Big(\frac{\bar{\theta}(\beta_1+\beta_2-\beta_2')}{\theta}, \frac{\bar{\theta}(\beta_1'+\beta_2'-\beta_2)}{\theta}\Big) \\
&\quad - \bar{\theta}d_{\beta_1,\beta_2}(P_1W_{\cdot\star}\|P_1P_2W, P_2W_{\star\cdot}) - \bar{\theta}d_{\beta_1',\beta_2'}(P_2W_{\star\cdot}\|P_1P_2W, P_1W_{\cdot\star}) \\
&\quad - \theta d_{(\beta_1+\beta_2-\beta_2')\bar{\theta}/\theta, (\beta_1'+\beta_2'-\beta_2)\bar{\theta}/\theta}(P_1P_2W\|P_1W_{\cdot\star}, P_2W_{\star\cdot}) \Big\}, \qquad (11)
\end{aligned}$$

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

Remark 2. *The result in Theorem 2 is valid for the intermittent process described above with arbitrary probability distribution on the time slots that each user transmits as long as the number of collided symbols d is fixed. Furthermore, the result remains valid if the number of collided symbols is a random variable denoted by D, such that the ratio of the collided symbols to the codeword length converges, i.e., $D/k \xrightarrow{p} \theta$ as $k \to \infty$. As a special case, we might think of the following intermittent process: If the length of the buffers are equal, then $\mathbb{P}((S_1,S_2)=(1,1)) = \theta$ and $\mathbb{P}((S_1,S_2)=(1,0)) = \mathbb{P}((S_1,S_2)=(0,1)) = (1-\theta)/2$; otherwise only the user with more symbols in its buffer transmits. Note that in this example, the length of the receive window is a random variable $N = 2k - D$, but $D/k \xrightarrow{p} \theta$ as $k \to \infty$.*

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

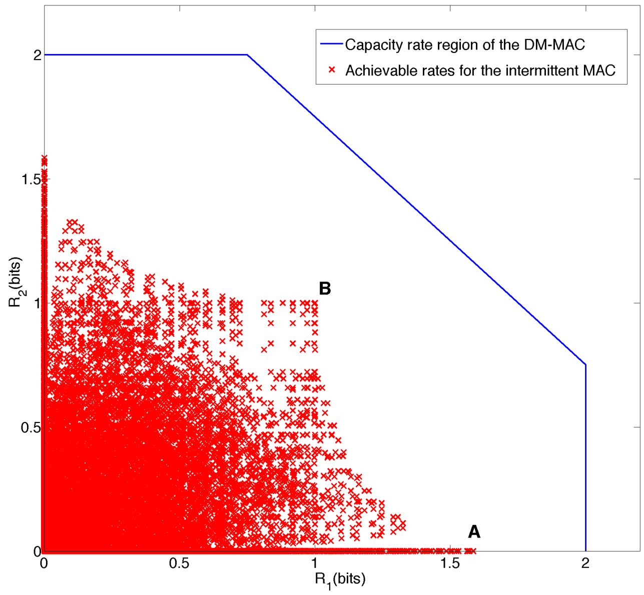

Fig. 2. Comparing the capacity rate region of the DM-MAC with the achievable rates for the intermittent MAC with collision avoidance mechanism obtained from Theorem 1.

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. □

VI. A Simple Example

Consider a DM-MAC with $\mathcal{X}_1, \mathcal{X}_2 = \{0,1,2,3\}$ and $\mathcal{Y} = \{0,1,...,6\}$ such that $Y = X_1 + X_2$, where $+$ corresponds to real addition. The capacity region of this channel is shown with the blue curve in Figure 2. The red dots correspond to achievable rates (R_1, R_2) for the intermittent MAC with collision avoidance obtained from Theorem 1 using different input distributions $P_1(x_1)$ and $P_2(x_2)$. For simplicity, we only focus on the result of Theorem 1. Not surprisingly, the plot suggests that the intermittency of the system and lack of knowledge about the position of the symbols at the decoder come with a significant cost. We should mention that achieving the rate pairs shown by points A and B in the figure is surprisingly simple. In order to achieve point A, we use $P_1(x_1) = [0, 1/3, 1/3, 1/3]$ and $P_2(x_2) = [1, 0, 0, 0]$, and to achieve point B, we use $P_1(x_1) = [0, 0, 1/2, 1/2]$ and $P_2(x_2) = [1/2, 1/2, 0, 0]$. In both cases, the overhead function $f_1(P_1, P_2, W)$ in Theorem 1 evaluates to zero, since the distributions $P_1 W_\star$ and $P_2 W_\star$ become disjoint and the partial divergence terms become infinite. It is also worth pointing out that the achievable rate region for the intermittent MAC model does not have to be convex, as can be seen from the figure, because time sharing is not possible due to the intermittency and asynchronism of the system.

Appendix A
Proof of Lemma 1

With a little abuse of notation, let $X_1^{\rho_1 k}$, $X_2^{\rho_2 k}$, and $X_3^{\rho_3 k}$ be the sequence of symbols in X^k that are i.i.d. according to Q_1, Q_2, and Q_3, respectively. If these sequences have types P_1, P_2, and P_3, respectively, then the whole sequence X^k has type $\rho_1 P_1 + \rho_2 P_2 + \rho_3 P_3$. Therefore, we have

$$\mathbb{P}(X^k \in T_P)$$
$$= \mathbb{P}\Big(\cup_{\substack{P_1,P_2,P_3 \in \mathcal{P}^{\mathcal{X}}: \\ \rho_1 P_1 + \rho_2 P_2 + \rho_3 P_3 = P}} \{X_1^{\rho_1 k} \in T_{P_1}, X_2^{\rho_2 k} \in T_{P_2}, X_3^{\rho_3 k} \in T_{P_3}\}\Big) \tag{12}$$

$$= \sum_{\substack{P_1,P_2,P_3 \in \mathcal{P}^{\mathcal{X}}: \\ \rho_1 P_1 + \rho_2 P_2 + \rho_3 P_3 = P}} \mathbb{P}(X_1^{\rho_1 k} \in T_{P_1}, X_2^{\rho_2 k} \in T_{P_2}, X_3^{\rho_3 k} \in T_{P_3}) \tag{13}$$

$$\doteq \sum_{\substack{P_1,P_2,P_3 \in \mathcal{P}^{\mathcal{X}}: \\ \rho_1 P_1 + \rho_2 P_2 + \rho_3 P_3 = P}} e^{-k(\rho_1 D(P_1\|Q_1) + \rho_2 D(P_2\|Q_2) + \rho_3 D(P_3\|Q_3))}, \tag{14}$$

$$\doteq e^{-k \min_{\substack{P_1,P_2,P_3 \in \mathcal{P}^{\mathcal{X}}: \\ \rho_1 P_1 + \rho_2 P_2 + \rho_3 P_3 = P}} \rho_1 D(P_1\|Q_1) + \rho_2 D(P_2\|Q_2) + \rho_3 D(P_3\|Q_3)} \tag{15}$$

where (13) follows from the disjointness of the events in (12) since a sequence has a unique type; where (14) follows from the independence of the three events in (13) and obtaining the probability of each of them according to [9, Lemma 1.2.6]; and where (15) follows from the fact that the number of different types is polynomial in the length of the sequence [9], which makes the total number of terms in the summation (14) polynomial in k, and therefore, the exponent equals the largest exponent of the terms in the summation (14). Note that $\doteq$ denotes an equality in exponential sense as $k \to \infty$, i.e., $\lim_{k\to\infty} \frac{1}{k} \log(\cdot)$ of both sides are equal. Thus, (15) proves the lemma.

Acknowledgment

The authors wish to thank Dr. Aslan Tchamkerten and Shyam Kumar for several useful discussions, which led 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] A. El Gamal and Y.-H. Kim, *Network Information Theory*, Cambridge University Press, UK, 2012.

[8] A. Somekh-Baruch, S. Shamai, and S. Verdu, "Cooperative Multiple-Access Encoding with States available at One Transmitter," *IEEE Trans. on Inf. Theory,* vol. 54, no. 10, pp. 4448–4469, October 2008.

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