On the Capacity of Strong Asynchronous Multiple Access Channels with a Large Number of Users

Sara Shahi, Daniela Tuninetti and Natasha Devroye
University of Illinois at Chicago, Chicago IL 60607, USA. Email: sshahi7, danielat, devroye @uic.edu

***Abstract*—This paper studies the impact of block asynchronism on the capacity of a slotted Multiple Access Channel (MAC) whose number of users K_n increases with the blocklength n. In a *slotted strong-asynchronous MAC*, the K_n users have independent transmission start times that are integer multiples of n (slotted) which are uniformly distributed on a window of length $A_n = e^{n\alpha}$ (strong-asynchronism). All users' messages as well as transmission times need to be reliably decoded at the single receiver. We show that for $K_n = e^{n\nu}$ with $\nu > \alpha$, not even synchronization is possible when transmitting a single message per user. We also show that for $K_n = e^{\nu}$ with $\nu = o(n)$, each user can achieve its point-to-point asynchronous capacity, which is a trivial upper bound for the capacity of the MAC. Finally, achievable rates for $K_n = e^{n\nu}$ with $0 < \nu < \frac{\alpha}{2}$ are derived.**

I. Introduction

One assumption used in much network information theoretic analysis is that of block and symbol synchronous transmission between transmitters and receivers. More specifically, it is often assumed that the receiver knows the starting point of the codeword and is only concerned about correctly decoding the received symbols. In practice, this synchronization is usually achieved by transmitting a pilot signal before the data signal. The pilot signal does not carry any information and has negligible impact (as the blocklength goes to infinity) on the achievable rates for streams of data where synchronization is done once, or for finitely many users. In large decentralized networks, however, the use of pilot signals for synchronization may cause an unacceptable amount of overhead. An example of such a network is a large sensor network, where nodes usually transmit bursts of data infrequently to an access point. This motivated finding the channel capacity of a network without an a priori assumption of synchronization.

In this work the usage of pilot symbols is not assumed, and the codebook may in theory serve the purpose of synchronization as well as of data transfer. For example, one could imagine that the codebook is sufficiently different in idle and busy time blocks so as to achieve synchronization at the receiver.

In addition, we do not assume preambles in our codebook design to distinguish the users. Each user's codebook should be sufficiently different so as to allow the receiver to identify both the message and its transmitter. These tasks become harder to achieve as the length of the possible transmission window as well as the number of users increases, in particular if they increase exponentially with the blocklength.

Past Work. In the literature, different levels of asynchronism have been considered. The level of asynchronism in a system is defined by the length of a window A_n within which the transmission can initiate and is defined with respect to the codeword blocklength n. In a *mildly asynchronous* MAC [1], $A_n = o(n)$ and the capacity region is the same as that of the classical synchronous MAC. In a *totally asynchronous* MAC [2], $A_n = n$ and users continuously send their messages after transmission initiation. In this setting, time sharing is no longer feasible; as a result, the capacity region lacks the convex hull operation seen in the capacity region of the synchronous MAC.

More recently, *strongly asynchronous* communication was introduced in [3], [4], [5], [6] with $A_n = e^{n\alpha}$ for some $\alpha \geq 0$ where users only transmit once within each window. In [3] it was shown that reliable communication is indeed possible for $0 < \alpha < \alpha_0$; α_0 being the synchronization threshold. In [6] the suboptimality of preamble based synchronization schemes was shown. The capacity of a strong-asynchronous point-to-point (SA-P2P) channel is [4]

$$C_{\text{SA-P2P}} = \max_{P_X : D([P_X Q] \| Q_\star) > \alpha} I(P_X, Q), \tag{1}$$

where the maximum is is defined to be zero for $\alpha > \alpha_0 = \max_{x \in \mathcal{X}} D(Q_x \| Q_\star)$. The notation used here is further defined in Section II.

The capacity in (1) may be interpreted as follows: while in the synchronous point-to-point channel the maximization is over all input distributions P_X, the maximization is now restricted to those input distributions that induce output distributions $P_Y = [P_X Q]$ that are sufficiently different from the 'idle output distribution' $Q_\star$. In [3], [4], [5], [6] the SA-P2P capacity was evaluated under the requirement of correct decoding only (and not necessarily synchronization); in [7] the author showed that imposing exact transmission time recovery does not change the capacity in (1).

In addition, while [5] mainly focussed on point-to-point communication per unit cost, the authors briefly discussed in [5, Remark 3] the capacity of the *strong-asynchronous collision MAC* with exponentially many users and with a *per-user* probability of error (though their parametrization is different from ours). In this model, simultaneous transmission of two or more users results in a 'collision' that produces as output distribution the same as if all users where idle, regardless of the number of colliding users. The probability of error at the MAC receiver is evaluated for each user individually, as opposed to the classical (stronger) requirement that all messages are jointly reliably decoded. In the proposed achievable scheme for $K_n = e^{n\nu}$ number of users with

$\nu < \alpha/2$, all users employ the same codebook and thus users are not distinguishable unless an identifier is sent along with the message.

A related line of work, but not dealing with asynchronism, is the so-called *many-user MAC*. In [8] the authors considered a synchronous MAC with random user activity where the number of users increases linearly with the blocklength. This many-user model, while different from ours, faces some challenges, as we do in here, which arise from the fact that the number of users increases with the blocklength. In [8] one of these challenges is that the number of possible error events is exponential (in the blocklength), which prevents them from using a simple union bound for bounding the probability of error. Here we encounter the same problem as the number of possible error events scales faster than exponential. Intuitively, the blocklength as used in [8] may be thought of as our asynchronous window length $A_n = e^{n\alpha}$.

Contributions. In this paper, we investigate the Strong-Asynchronous Slotted MAC (SAS-MAC) for an increasing number of users with the blocklength. The slotted assumption restricts the transmission start time to be integer multiples of n; this assumption simplifies the error analysis yet captures the essence of the problem. We consider the classical definition of error in a MAC, with the error being the union of errors for all users, as opposed to [5, Remark 2], which considers the per-user probability of error. With an increasing number of users and an exponential (in the blocklength) window length, our number of error events scales faster than exponential (in the blocklength). Data transfer, synchronization and user identification are all achieved without imposing the use of pilot symbols. We show:

1) for an exponential number of users $K_n = e^{n\nu}$ and an exponential window length $A_n = e^{n\alpha}$, when $\nu < \frac{\alpha}{2}$, users can transmit at positive rate,
2) however when $\nu \geq \alpha$, users can not even be synchronized when transmitting a single codeword, and
3) for a sub exponential number of users K_n with $\log K_n = o(n)$, each user can achieve its point-to-point strong-asynchronous capacity.

Paper organization. In Section II we define the notation and the system model. In Section III we investigate the capacity of the SAS-MAC for different scalings of the number of users with the blocklength. Section IV concludes the paper.

II. System Model

Unless otherwise indicated, we use the notation convention of [9]. In particular, the discrete memoryless *classical MAC* with K-user, denoted as $(\mathcal{X}_1 \times \ldots \times \mathcal{X}_K, Q(.|.), \mathcal{Y})$, consists of $K+1$ finite sets $(\mathcal{X}_1, \ldots, \mathcal{X}_K, \mathcal{Y})$ and a collection of conditional distributions $Q(y|x_1, \ldots, x_K)$ on $\mathcal{Y}$, one for each input $(x_1, \ldots, x_K)$. This MAC is memoryless since we assume

$$Q(y^n|x_1^n, \ldots, x_K^n) = \prod_{t=1}^{n} Q(y_t|x_{1,t}, \ldots, x_{K,t}), \forall n \in \mathbb{N}.$$

An $(M_1, \ldots, M_K, n, A, \epsilon)$ code for the *asynchronous MAC* is defined as follows. Each user $i \in [1:K]$ has an encoding function $f_i : \mathcal{W}_i \to \mathcal{X}_i^n$ over the message sets $\mathcal{W}_i := [1:M_i]$ where we define $X_i^n(m_i) := f(m_i)$; it randomly and uniformly chooses a message $m_i \in \mathcal{W}_i$ to convey to the receiver, together with a slot/block index $t_i \in [1:A]$ also chosen uniformly at random and independently of the message m_i; it sends $[\star_i^{n(t_i-1)}\ f_i(m_i)\ \star_i^{n(A-t_i)}] \in \mathcal{X}_i^{nA}$, where $\star_i \in \mathcal{X}_i$ is the designated 'idle' symbol for user i. The destination has a decoding function $g : \mathcal{Y}^{nA} \to ((\mathcal{A}, \mathcal{W}_1) \times \ldots \times ((\mathcal{A}, \mathcal{W}_K))$, such that the average (over all messages and all blocks) probability of error satisfies

$$\epsilon \geq \mathbb{P}(\mathcal{E}) := \frac{1}{A^K \prod_{i=1}^{K} M_i} \sum_{(m_1,t_1)\ldots,(m_K,t_K)}$$
$$\mathbb{P}\left[g(y^{nA}) \neq \big((t_1, m_1), \ldots, (t_K, m_K)\big) \mid H_{(t_1,m_1),\ldots,(t_K,m_K)}\right],$$

where $H_{(t_1,m_1),\ldots,(t_K,m_K)}$ is the 'hypothesis' that user i has chosen (t_i, m_i) for all $i \in [1:K]$.

For the *SAS-MAC with asynchronism level α,* a code is defined over an asynchronous MAC channel with A increasing exponentially with blocklength n as $A_n = e^{n\alpha}$. For both exponential number users $K_n : \log(K_n) = O(n)$, and subexponential number of users $K_n : \log(K_n) = o(n)$, a tuple $(R_1, \ldots, R_{K_n}, \alpha)$ is said to be achievable if there exists a sequence of codes $(e^{nR_1}, \ldots, e^{nR_{K_n}}, n, e^{n\alpha}, \epsilon_n)$ with $\epsilon_n \to 0$ as $n \to \infty$. The capacity region is the closure of all such achievable tuples.

Special Notation. A vector of length n with a subscript refers to $Y_s^n := [Y_{(s-1)n+1}, \ldots, Y_{sn}]$; the subscript could also indicate a user; its meaning should be clear from the context.

For two sequences (x^n, y^n), their joint empirical distribution is defined as

$$\widehat{P}_{x^n,y^n}(a,b) := \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{(x_i,y_i)=(a,b)\}}, \forall a,b \in \mathcal{X} \times \mathcal{Y},$$

where $\mathbb{1}_{\{A\}}$ is the indicator function of the event A. We say that (x^n, y^n) are jointly strongly ϵ-typical according to $P_{X,Y}$, and write $(x^n, y^n) \in T_\epsilon^n(P_{X,Y})$, if

$$|\ \widehat{P}_{x^n y^n}(a,b) - P_{X,Y}(a,b)\ | \leq \epsilon P_{X,Y}(a,b), \forall (a,b) \in \mathcal{X} \times \mathcal{Y}.$$

For a random variable (r.v.) X we denote with $P_X(x), \forall x \in \mathcal{X}$, its marginal distribution; a stochastic kernel / transition probability from $\mathcal{X}$ to $\mathcal{Y}$ is denoted by $Q(y|x), \forall (x,y) \in \mathcal{X} \times \mathcal{Y}$; the output distribution induced by P_X and the channel Q is denoted as $[P_X Q](y) := \sum_x P_X(x)Q(y|x)), \forall y \in \mathcal{Y}$—please note the square brackets,—and the joint input-output distribution as $P_X Q(x,y) := P_X(x)Q(y|x), \forall (x,y) \in \mathcal{X} \times \mathcal{Y}$.

For the SAS-MAC we use the shorthand notation

$$Q_S(y|x_S) := Q(y|x_S, \star_{S^c}), \forall S \subseteq [1:K],$$

to indicate that the users indexed by S transmit x_i, and users indexed by $S^c := [1:K] \setminus S$ transmit their idle symbol. We also define $Q_0(y) := Q(y|\star_{[1:K]})$ and $Q_i(y \mid x_i) := Q_{\{i\}}(y|x_i)$. We use $I(P_{X_i}, Q_i)$ to denote the mutual information between inputs and output when all users $j \neq i$ transmit their idle symbol $\star_j$ and user i a symbol from distribution P_{X_i}.

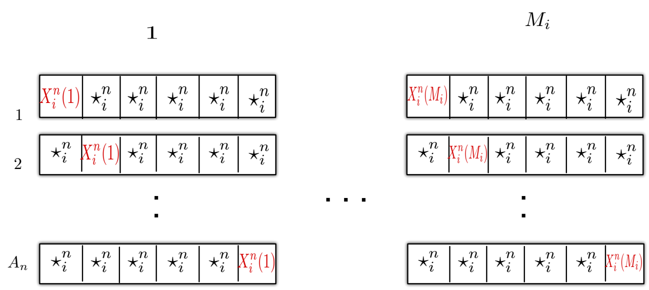

Fig. 1: Extended codebook.

III. Main Results

In this section we consider the performance of the SAS-MAC for three different scalings of the number of users.

A. Exponential regime: case $\log(K_n) = n\nu : \nu > \alpha$

Theorem 1. *For a SAS-MAC with asynchronism level $\alpha < \alpha_0$ and $\log(K_n) = n\nu : \nu > \alpha$, synchronization is not possible, i.e., even with $M_i = 1, \forall i \in [1 : K_n]$, one has $\bar{\mathbb{P}}[\mathcal{E}] > 0$.*

Proof: User $i \in [1 : K_n]$ has a codebook with $M_i = e^{nR_i}$ codewords of length n. Define for $i \in [1 : K_n]$ an 'extended codebook' consisting of $A_n M_i$ codewords of length nA_n constructed such that $\forall m_i \in \mathcal{W}_i$ and $\forall t_i \in [1 : A_n]$

$$\widetilde{X}_i^{nA_n}(m_i, t_i) := \left[\, \star_i^{n(t_i-1)} \;\, f_i(m_i) \;\, \star_i^{n(A_n-t_i)} \,\right],$$

as depicted in Fig. 1. By using Fano's inequality, i.e., $H(X_1^{nA_n}, \dots, X_{K_n}^{nA_n} | Y^{nA_n}) \leq n\epsilon_n : \epsilon_n \to 0$ as $n \to \infty$, for any codebook of length nA_n we have

$$\begin{aligned}
&H(X_1^{nA_n}, \dots, X_{K_n}^{nA_n}) = H(m_1, t_1, \dots, m_{K_n}, t_{K_n}) \\
&= n\alpha K_n + \sum_{i \in [1:K_n]} \log M_i \\
&= H(X_1^{nA_n}, \dots, X_{K_n}^{nA_n} | Y^{nA_n}) + I(X_1^{nA_n}, \dots, X_{K_n}^{nA_n}; Y^{nA_n}) \\
&\leq n\epsilon_n + ne^{n\alpha}\, |\mathcal{Y}| \quad \iff \\
&\nu + \frac{\log\left(1 + \frac{1}{\alpha K_n}\sum_{i\in[1:K_n]} R_i\right)}{n} \leq \alpha + \frac{\log\left(1 + \frac{\epsilon_n}{e^{n\alpha}|\mathcal{Y}|}\right)}{n},
\end{aligned}$$

where $\frac{\log\left(1+\frac{1}{\alpha K_n}\sum_{i\in[1:K_n]} R_i\right)}{n} \geq 0$ and $\frac{\log\left(1+\frac{\epsilon_n}{e^{n\alpha}|\mathcal{Y}|}\right)}{n} \geq 0$ vanish as n goes to infinity. This implies that $\nu \leq \alpha$ is a necessary condition for reliable communications. In other words, for $\nu > \alpha$ not even synchronization, i.e., $M_i = 1, \forall i \in [1 : K_n]$, is possible. ■

B. Sub-exponential regime: case $\log(K_n) = o(n)$

For $\nu < \frac{\alpha}{2}$ the probability δ_n that more than one user transmits a codeword in each block is:

$$\delta_n = 1 - \frac{A_n(A_n-1)\dots(A_n-K_n+1)}{A_n^{K_n}}, \tag{2}$$

which goes to zero as n goes to infinity for $\nu < \frac{\alpha}{2}$. Hence one may analyze the probability of error conditioned on the fact that users are transmitting in different blocks, i.e., no collision. This assumption reduces the number of different hypotheses for each block that must be considered in the error analysis.

Theorem 2. *For a SAS-MAC with asynchronism level $\alpha < \alpha_0$ and $\log(K_n) = o(n)$, the capacity region is the Cartesian product of the corresponding strong asynchronous point-to-point capacities given by*

$$R_i < \max_{P_{X_i} : D([P_{X_i} Q_i] \| Q_0) > \alpha} I(P_{X_i}, Q_i), \forall i \in [1 : K_n]. \tag{3}$$

Proof: Each user generates an i.i.d. random codebook according to the distribution P_{X_i} on $\mathcal{X}_i, \forall i \in [1 : K_n]$. The decoder uses the following 'slot by slot' strong typicality decoder: for every block $s \in [1 : A]$ it finds the empirical distribution of the output sequence Y_s^n and codeword $X_i^n(m_i)$ for every $m_i \in \mathcal{W}_i,\ i \in [1 : K_n]$; it announces that m_i was the sent codeword in block s if m_i is the unique message index such that $(X_i^n(m_i), Y_s^n) \in T_\epsilon^n(P_{X_i} Q_i)$; if no codeword passes the test, the decoder declares that no user was active on block s and moves forward to block $s + 1$; if more than one codeword passes the test, the decoder picks one uniformly at random and moves forward to block $s + 1$.

Assuming no collision, since all hypotheses are equally likely, and by averaging over all random codes C, we have that the average probability of error is the same as that obtained by conditioning over $H' := H_{(1,1)\dots(K_n,1)}$. We use $\mathbb{P}_{H'}$ to denote the underlying probability measure given hypothesis H'. By the union bound we can write

$$\begin{aligned}
&\mathbb{E}_C[\mathbb{P}[\mathcal{E}|C]] \leq \mathbb{P}_{H'}[\mathcal{E}] + \delta_n \\
&\leq \delta_n + \sum_{i\in[1:K_n]} \mathbb{P}_{H'}[(X_i^n(1), Y_i^n) \notin T_\epsilon^n(P_{X_i} Q_i)] \qquad (4) \\
&+ \sum_{i\in[1:K_n]} \sum_{m_i\in[2:M_i]} \mathbb{P}_{H'}[(X_i^n(m), Y_i^n) \in T_\epsilon^n(P_{X_i} Q_i)] \qquad (5) \\
&+ \sum_{s\in[K_n+1:A]} \sum_{i\in[1:K_n]} \sum_{m_i\in\mathcal{W}_i} \mathbb{P}_{H'}[(X_i^n(m_i), Y_s^n) \in T_\epsilon^n(P_{X_i} Q_i)] \qquad (6) \\
&+ \sum_{i\in[1:K_n]} \sum_{\substack{j\in[1:K_n] \\ j\neq i}} \sum_{m_j\in\mathcal{W}_j} \mathbb{P}_{H'}[(X_j^n(m_j), Y_i^n) \in T_\epsilon^n(P_{X_j} Q_j)] \qquad (7) \\
&\leq \delta_n + \sum_{i\in[1:K_n]} e^{-n\epsilon^2} + \sum_{i\in[1:K_n]} e^{-n(I(P_{X_i}, Q_i) - R_i)} \qquad (8) \\
&+ \sum_{i\in[1:K_n]} e^{-n(I(P_{X_i}, Q_i) + D([P_{X_i} Q_i] \| Q_0) - \alpha - R_i)} \qquad (9) \\
&+ \sum_{i\in[1:K_n]} \sum_{\substack{j\in[1:K_n] \\ j\neq i}} e^{-n(I(P_{X_j}, Q_j) + D([P_{X_j} Q_j] \| [P_{X_i} Q_i]) - R_j)} \qquad (10)
\end{aligned}$$

where δ_n is the probability of collision given in (2) which goes to zero as n goes to infinity, the term (4) is the probability that the true codeword is not typical with its corresponding output, the term in (5) is the probability of classical synchronous point-to-point error, the term in (6) is the probability that a noise block, or a block where no user was active, mimics any of the codewords, and finally the term in (7) is the probability that users are confused with one another. The bound in (8) is due to the typicality decoder and those in (9) and (10) are proved in Appendix A. All together, by assuming K_n to be sub-exponential in n, so that $K_n e^{-n\epsilon^2} \to 0$, we get

$$R_j < I(P_{X_j}, Q_j), \tag{11}$$

$$R_j + \alpha < I(P_{X_j}, Q_j) + D([P_{X_j} Q_j] \parallel Q_0), \tag{12}$$

$$R_j < I(P_{X_j}, Q_j) + D([P_{X_j} Q_j] \parallel [P_{X_i} Q_i]), \forall i \neq j, \tag{13}$$

where the bound in (13) is redundant due the more restrictive bound in (11). The achievable rates obtained above match the converse bound given by the Cartesian product of the corresponding point-to-point capacities in (3). Finally, (11) – (12) are equivalent to (3) as proven in [7] and hence the theorem is proved. ∎

Remark 1. *As it can be seen in* (8)*, this typicality decoder imposes the condition* $e^{n\nu} e^{-n\epsilon^2} \to 0$ *for exponential number of users which corresponds to* $\nu = 0$ *when* $\epsilon \to 0$*. This begs the question of whether indeed it is possible to support exponentially many users* ($\nu > 0$)*. Next subsections affirmatively answer this question. Also note that in calculating the per-user probability of error, all the summations over users* $i \in [1 : K_n]$ *in* (8), (9), (10) *are eliminated and hence would relax the requirement* $\nu = 0$ *for this typicality decoder.*

C. Exponential regime: case $\log(K_n) = n\nu : 0 < \nu < \frac{\alpha}{2}$

Now we investigate a SAS-MAC with an exponential number of users. This regime is the hardest to deal with as the typicality decoder seems to fail as the number of users grow exponentially fast. For example, if we apply the previously introduced strong typicality decoder to this case, error events as in (4) would restrict ν to be zero. The key ingredient in our analysis is a novel way to bound the probability of error reminiscent of Gallager's error exponent. We show an achievable scheme that allows a positive lower bound on the rates and on ν. This proves that in fact reliable transmission with an exponential number of users in an exponential level of asynchronism is possible. We use a maximum likelihood (ML) decoder sequentially in each block to identify the active user and its message.

In our results, we use the following notation. The *Chernoff distance* between two distributions is defined as

$$C(P, Q) := \sup_{0 \le t \le 1} -\log \left(\sum_x P(x)^t Q(x)^{1-t} \right). \tag{14}$$

We extend this definition and introduce the quantity

$$C(P_{X_i}, Q_i, P_{X_j}, Q_j) := \sup_{0 \le t \le 1} \mu(t), \tag{15}$$

where

$$\mu(t) := -\log \sum_{x_j, x_s, y} P_{X_j}(x_j) P_{X_s}(x_s) Q_s(y|x_s)^{1-t} Q_j(y|x_j)^t$$

is a concave function of t. We also define

$$C(., Q_0, P_{X_j}, Q_j) := \sup_{0 \le t \le 1} -\log\Big(\sum_{x_j, y} P_{X_j}(x_j) Q_0(y)^{1-t} Q_j(y|x_j)^t\Big)$$

to address the special case with $s = 0$ where all users are idle. Note that (see Appendix B):

$$B(P, Q) := C(P, Q, P, Q)$$
$$= -\log \sum_{x, x', y} P(x) P(x') \sqrt{Q(y|x) Q(y|x')}, \tag{16a}$$

$$C(\,.\,, Q_0, P_{X_j}, Q_j) \le I(P_{X_j}, Q_j) + D([P_{X_j} Q_j] \parallel Q_0), \tag{16b}$$

$$C(P_{X_i}, Q_i, P_{X_j}, Q_j) \le I(P_{X_j}, Q_j) + D(P_{X_i}[P_{X_j} Q_j] \| P_{X_i} Q_i), \tag{16c}$$

where, due to symmetry, in $C(P, Q, P, Q)$ the supremum is achieved at the midpoint $t = \frac{1}{2}$, and hence $B(P, Q) = C(P, Q, P, Q) = \mu(\frac{1}{2})$. The bounds in (16) show that in the achievable rates in Theorem 3 are less than the corresponding point-to-point bounds.

Theorem 3. *For a SAS-MAC with asynchronism level* $\alpha < \alpha_0$ *and* $\log(K_n) = n\nu : 0 < \nu < \frac{\alpha}{2}$*, the following rates are achievable*

$$\nu + R_j < B(P_{X_j}, Q_j), \tag{17}$$

$$2\nu + R_j < C(P_{X_i}, Q_i, P_{X_j}, Q_j), \forall i \neq j, \tag{18}$$

$$\alpha + \nu + R_j < C(\,.\,, Q_0, P_{X_j}, Q_j), \forall j \in [1 : K_n]. \tag{19}$$

Proof: Each user generates an i.i.d. random codebook according to the distribution P_{X_i} on $\mathcal{X}_i, \forall i \in [1 : K_n]$. The decoder uses the following 'slot by slot' decoder: for each block $s \in [1 : A_n]$, the decoder outputs

$$i^* \in \arg \max_{i \in [0:K_n], m_i \in \mathcal{W}_i} Q_i(y_s^n | x_i^n(m_i)),$$

where $\mathcal{W}_0 = \{1\}, x_0^n = \emptyset$. As discussed in the error analysis for sub-exponential number of users, we can write the probability of error as follows

$$\mathbb{E}_C[\mathbb{P}[\mathcal{E}|C]] \le \mathbb{P}_{H'}[\mathcal{E}] + \delta_n$$
$$\le \sum_{i \in [1:K_n]} \sum_{m_i \in [2:M_i]} \mathbb{P}_{H'} \left[\log \frac{Q_i(Y_i^n \mid X_i^n(m_i))}{Q_i(Y_i^n \mid X_i^n(1))} > 0 \right]$$
$$+ \sum_{i \in [1:K_n]} \sum_{\substack{j \in [0:K_n] \\ j \neq i}} \sum_{m_j \in \mathcal{W}_j} \mathbb{P}_{H'} \left[\log \frac{Q_j(Y_i^n | X_j^n(m_j))}{Q_i(Y_i^n | X_i^n(1))} > 0 \right]$$
$$+ \sum_{s \in [K_n+1:A_n]} \sum_{j \in [1:K_n]} \sum_{m_j \in \mathcal{W}_j} \mathbb{P}_{H'} \left[\log \frac{Q_j(Y_s^n | X_j^n(m_j))}{Q_0(Y_s^n)} > 0 \right]$$
$$\le \sum_{i \in [1:K_n]} e^{nR_i} e^{-n \sup_t -\log \mathbb{E}\left[\left(\frac{Q_i(Y_i | \overline{X}_i)}{Q_i(Y_i | X_i)} \right)^t \right]}$$
$$+ \sum_{i \in [1:K_n]} \sum_{\substack{j \in [0:K_n] \\ j \neq i}} e^{nR_j} e^{-n \sup_t -\log \mathbb{E}\left[\left(\frac{Q_j(Y_i | X_j)}{Q_i(Y_i | X_i)} \right)^t \right]}$$
$$+ e^{n\alpha} \sum_{j \in [1:K_n]} e^{nR_j} e^{-n \sup_t -\log \mathbb{E}\left[\left(\frac{Q_j(Y_s | X_j)}{Q_0(Y_s)} \right)^t \right]},$$

where $P_{X_s,\overline{X}_s}(x,x') = P_{X_s}(x)P_{X_s}(x')$. The last inequality is due to the Chernoff bound. In order for that each term in the probability of error upper bound to vanish as n grows to infinity, we find the conditions stated in the theorem. ■

D. Example

We now show a simple example of a channel for which the bounds in Theorem 3 are strictly positive. Consider the SAS-MAC with asynchronism level α with input output relationship $Y = \sum_{i\in[1:K_n]} X_i \oplus Z$ with $Z \sim Ber(\xi)$ being a Bernoulli random variable with parameter ξ. In our notation

$$Q_i(y|a) = \mathbb{P}[X_i \oplus Z = y | X_i = a] = \mathbb{P}[Z = a \oplus y]$$
$$= \begin{cases} 1-\xi & a \oplus y = 0 \ (i.e.,\ a = y) \\ \xi & a \oplus y = 1 \ (i.e.,\ a \neq y) \end{cases}.$$

Assume $P_{X_i} = Ber(p_i)$ for all $i \in [1 : K_n]$. We want to show $\inf_{i,j} C(P_{X_i}, Q_i, P_{X_j}, Q_j) > 0, \inf_i B(P_{X_j}, Q_j) > 0, \inf_j C(.\,, Q_0, P_{X_j}, Q_j) > 0$. In this regard, we find strictly positive lower bound on these quantities independent of i, j. It can be shown that in this setting, the optimal t in $C(P_{X_j}, Q_j, P_{X_i}, Q_i) = \sup_t \mu(t)$ is equal to $t = 1/2$; thus

$$C(P_{X_i}, Q_i, P_{X_j}, Q_j) = g(p_i * p_j, \xi),$$
$$B(P_{X_j}, Q_j) = g(p_j * p_j, \xi),$$
$$C(.\,, Q_0, P_{X_j}, Q_j) = g(p_j, \xi),$$
$$g(a, b) := -\log\left(1 - a + 2a\sqrt{\xi(1-\xi)}\right),$$
$$p_i * p_j := p_i(1-p_j) + (1-p_i)p_j.$$

Moreover the function $g(a, \xi)$ is zero iff either $a = 0$ or $\xi = \frac{1}{2}$. The case $\xi = \frac{1}{2}$ is uninteresting, because in this case even synchronous single-user point-to-point channel capacity is zero. Finally, by setting $p_i = \frac{1}{2}, \forall i \in [1 : K_n]$, and assuming without loss of generality that $\xi \neq \frac{1}{2}$, we get $g(.,.) = -\log\left(1/2 + \sqrt{\xi(1-\xi)}\right) > 0$. So on the BSC($\xi$) strictly positive rates and ν are achievable. With this, and by setting $R_i = R$ for all i, the region in Theorem 3 reduces to

$$\alpha + \nu + R < -\log\left(1/2 + \sqrt{\xi(1-\xi)}\right).$$

A trivial upper bound is the point-to-point capacity given in [7, eq.(18)-(19)].

IV. Conclusion

In this paper we showed that for a strongly asynchronous MAC with blocklength n, asynchronism level $A_n = e^{n\alpha}, \alpha \geq 0$, and $K_n = e^{n\nu}, \nu \geq 0$ users, reliable transmission is not possible if $\nu > \alpha$. Scaling the number of users up to a subexponential growth with n instead does not incur any loss in the achievable rates for each user compared to a point-to-point case. We also showed that it is indeed possible to transmit at positive rate for each user in the regime $0 < \nu < \frac{\alpha}{2}$. Converse bounds for $0 < \nu \leq \frac{\alpha}{2}$, as well as, achievablity for $\frac{\alpha}{2} \leq \nu < \alpha$ are subject of current investigation.

Acknowledgement The work of the authors was partially funded by NSF under award 1422511. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the NSF.

Appendix

A. Proof of (10)

For any joint empirical distribution J defined on $\mathcal{X}_j \times \mathcal{Y}, 1 \leq i \leq K_n$

$$\mathbb{P}_{H'}[(X_j^n(m_j), Y_i^n) \in T_\epsilon^n(P_{X_j}Q_j)] \leq \sum_{J: J\in T_\epsilon^n(P_{X_j}Q_j)} \mathbb{P}_{H'}[\widehat{P}_{(X_j^n(m_j),Y_i^n)} = J]$$
$$\overset{(a)}{=} \sum_{J: J\in T_\epsilon^n(P_{X_j}Q_j)} e^{-nD(J \| P_{X_j}[P_{X_i}Q_i])}$$
$$\leq \sum_{J: J\in T_\epsilon^n(P_{X_j}Q_j)} e^{-n\left(D(P_{X_j}Q_j \| P_{X_j}[P_{X_i}Q_i]) - \delta_\epsilon\right)}$$
$$\overset{(b)}{\leq} poly(n)\ e^{-n\left(D(P_{X_j}Q_j \| P_{X_j}[P_{X_i}Q_i]) - \delta_\epsilon\right)}$$
$$= poly(n)\ e^{-n\left(I(P_{X_j}, Q_j) + D([P_{X_j}Q_j] \| [P_{X_i}Q_i]) - \delta_\epsilon\right)},$$

where δ_ϵ can be made arbitrary small with the choice ϵ. Equality in (a) is due to [10, Lemma 2.6] and (b) is by [10, Lemma 2.2]. With similar reasoning, (9) can be proved.

B. Proof of (16b)

We find an upper bound on $C(.\,, Q_0, P_{X_j}, Q_j)$ by noting that $\mu(t)$ in (III-C) is concave in t with $\mu(1) = 0$ and

$$\frac{\partial \mu(t)}{\partial t}\ |_{t=1} = -I(P_{X_j}, Q_j) - D([P_{X_j}Q_j] \parallel Q_0) \leq 0.$$

Hence $\mu(t)$ is always less than $(I(P_{X_j}, Q_j) + D([P_{X_j}Q_j] \parallel Q_0))(1-t)$ and that for $0 \leq t \leq 1$ it is always less than $I(P_{X_j}, Q_j) + D([P_{X_j}Q_j] \parallel Q_0)$. The inequality in (16c) follows similarly.

References

[1] T. Cover, R. McEliece, and E. C. Posner, "Asynchronous multiple-access channel capacity," *Information Theory, IEEE Transactions on*, vol. 27, no. 4, pp. 409–413, Jul 1981.

[2] J. Hui and P. Humblet, "The capacity region of the totally asynchronous multiple-access channel," *Information Theory, IEEE Transactions on*, vol. 31, no. 2, pp. 207–216, Mar 1985.

[3] V. Chandar, A. Tchamkerten, and G. Wornell, "Optimal sequential frame synchronization," *Information Theory, IEEE Transactions on*, vol. 54, no. 8, pp. 3725–3728, Aug 2008.

[4] V. Chandar, A. Tchamkerten, and D. Tse, "Asynchronous capacity per unit cost," in *Information Theory Proceedings (ISIT), 2010 IEEE International Symposium on*, June 2010, pp. 280–284.

[5] A. Tchamkerten, V. Chandar, and G. Caire, "Energy and sampling constrained asynchronous communication," *IEEE Transactions on Information Theory*, vol. 60, no. 12, pp. 7686–7697, Dec 2014.

[6] A. Tchamkerten, V. Chandar, and G. Wornell, "Asynchronous communication: Capacity bounds and suboptimality of training," *Information Theory, IEEE Transactions on*, vol. 59, no. 3, pp. 1227–1255, March 2013.

[7] Y. Polyanskiy, "Asynchronous communication: Exact synchronization, universality, and dispersion," *Information Theory, IEEE Transactions on*, vol. 59, no. 3, pp. 1256–1270, March 2013.

[8] X. Chen and D. Guo, "Many-access channels: The gaussian case with random user activities," in *Information Theory (ISIT), 2014 IEEE International Symposium on*, June 2014, pp. 3127–3131.

[9] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Camrbidge University Press, 2012.

[10] I. Csiszar and J. Körner, *Information theory: coding theorems for discrete memoryless systems*. Cambridge University Press, 2011.