On the separation of correlation-assisted sum capacities of multiple access channels

Akshay Seshadri[1,3], Felix Leditzky[2], Vikesh Siddhu[3], and Graeme Smith[1,3,4]

[1]Department of Physics, University of Colorado, Boulder, CO 80309, USA
[2]Department of Mathematics & IQUIST, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[3]JILA, University of Colorado/NIST, Boulder, CO 80309, USA
[4]Center for Theory of Quantum Matter, University of Colorado, Boulder, CO 80309, USA

2022 IEEE International Symposium on Information Theory (ISIT) | 978-1-6654-2159-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/ISIT50566.2022.9834459

***Abstract*—Computing the sum capacity of a multiple access channel (MAC) is a non-convex optimization problem. It is therefore common to compute an upper bound on the sum capacity using a convex relaxation. We investigate the performance of such a relaxation by considering a family of MACs obtained from nonlocal games. First, we derive an analytical upper bound on the sum capacity of such MACs, while allowing the senders to share any given set of correlations. Our upper bound depends only on the properties of the game available in practice, thereby providing a way to obtain separations between the sum capacity assisted by different sets of correlations. In particular, we obtain a bound on the sum capacity of the MAC obtained from the magic square game that is tighter than the previously known result. Next, we introduce a game for which the convex relaxation of the sum capacity can be arbitrarily loose, demonstrating the need to find other techniques to compute or bound the sum capacity. We subsequently propose an algorithm that can certifiably compute the sum capacity of any two-sender MAC to a given precision.**

I. Introduction

A multiple access channel (MAC) is a channel with many senders and a single receiver. Throughout this study, we will restrict our attention to discrete memoryless MACs without feedback. Ahlswede [1] and Liao [2] pioneered the study of MACs by deriving a single-letter formula for the capacity region of a two-sender MAC. The capacity region of a MAC $\mathcal{N}$ with N senders can be written as

$$\begin{aligned}&\mathrm{Cap}(\mathcal{N}) \quad (1)\\ &= \mathrm{conv}\Big\{(R_1,\ldots,R_N) \mid 0 \leq \sum_{j\in J} R_i \leq I(B_J; Z|B_{J^c}),\\ &\qquad \forall \varnothing \neq J \subseteq [N], (B_1,\ldots,B_N) \sim p^{(1)}\cdots p^{(N)}\Big\}\end{aligned}$$

where for any set $J \subseteq [N]$, we denote $B_J = \{B_j \mid j \in J\}$ and $[N] = \{1,\ldots,N\}$ [3]. Here, $B_1,\ldots,B_N$ are random variables describing the input to the channel, whereas Z is the random variable describing the output of the channel. Note that the input probability distribution must be a product distribution $p^{(1)}\cdots p^{(N)}$ when computing the capacity region of a MAC. Subsequently, the sum capacity of the MAC $\mathcal{N}$ is defined as

$$\begin{aligned}S(\mathcal{N}) &= \sup\left\{\sum_{i=1}^N R_i \mid (R_1,\ldots,R_N) \in \mathrm{Cap}(\mathcal{N})\right\}\\ &= \max_{p^{(1)}\cdots p^{(N)}} I(B_1,\ldots,B_N; Z). \quad (2)\end{aligned}$$

Because the maximization is constrained to be over product distributions on the input, the resulting optimization problem is non-convex. This is the main source of difficulty in computing the sum capacity of a MAC [4], [5]. In fact, it has been shown that computing the sum capacity to a precision that scales inversely with the cube of the dimension is an NP-hard problem [5].

A common method to obtain upper bounds on the sum capacity is dropping the product distribution constraint, which results in a convex optimization problem [3]. This relaxation yields

$$C(\mathcal{N}) = \max_{p(b_1,\ldots,b_N)} I(B_1,\ldots,B_N; Z), \quad (3)$$

which we call the *relaxed sum capacity*. Note that $C(\mathcal{N})$ corresponds to the capacity of $\mathcal{N}$ considered as a single sender, single receiver channel. This gives the upper bound $S(\mathcal{N}) \leq C(\mathcal{N})$. We begin by investigating how good an upper bound this is on the sum capacity.

Defining a MAC using a nonlocal game [5] provides a convenient setting for this purpose as we can deduce separations between sum rates assisted by different sets of correlations by choosing an appropriate game. Such separations are also of independent interest because it aids in distinguishing the communication capability of different sets of correlations, such as classical, quantum and no-signalling correlations. Therefore, in Section II, we derive an upper bound on the sum rate of a MAC obtained from a nonlocal game, where the senders are allowed to share any given set of correlations. Our bound only depends on the number of question tuples in the game and the maximum winning probability of the game using assistance from the shared correlations when the questions are drawn uniformly at random. This winning probability is available in practice for commonly used games and correlations.

In Section III, we define a nonlocal game that gives an arbitrarily large separation between the relaxed sum capacity of the corresponding MAC and its actual sum capacity. This motivates us to look for other approaches to compute or bound the sum capacity. To that end, we propose an algorithm in Section IV that can compute the sum capacity of a two-sender MAC to any given precision. If instead one provides a fixed number of iterations, our algorithm gives an upper bound on the sum capacity. We refer to the full version [6] of this manuscript for complete details and proofs of the results presented here.

II. MACs from nonlocal games

A promise-free N-player nonlocal game G consists of question sets $\mathcal{X}_1, \dots, \mathcal{X}_N$, answer sets $\mathcal{Y}_1, \dots, \mathcal{Y}_N$ from which the players choose the answers, and the winning condition $\mathcal{W} \subseteq \mathcal{X} \times \mathcal{Y}$ that determines which tuples of questions and answers win the game. In the following, we denote by $\mathcal{X} = \mathcal{X}_1 \times \cdots \mathcal{X}_N$ and $\mathcal{Y} = \mathcal{Y}_1 \times \cdots \times \mathcal{Y}_N$ the set of question and answer tuples, respectively. Any conditional probability distribution $p_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x})$ on the answers $\boldsymbol{y} \in \mathcal{Y}$ given the questions $\boldsymbol{x} \in \mathcal{X}$ is called a strategy for the game. Given a set of strategies $\mathfrak{S}$, the maximum winning probability of the game G is given by

$$\omega^{\mathfrak{S}}(G) = \frac{1}{d} \sup_{p_{\boldsymbol{Y}|\boldsymbol{X}} \in \mathfrak{S}} \sum_{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{W}} p_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x}), \tag{4}$$

assuming that the questions are drawn uniformly at random.

Given such a nonlocal game G, we define the MAC $\mathcal{N}_G$ as follows [5]. The input alphabets are question-answer sets $\mathcal{X}_1 \times \mathcal{Y}_1, \dots, \mathcal{X}_N \times \mathcal{Y}_N$ for the N players, the output alphabet $\mathcal{X}_1 \times \cdots \times \mathcal{X}_N$ is the set of question tuples, and the probability transition matrix is given by

$$\mathcal{N}_G(\widehat{\boldsymbol{x}}|\boldsymbol{x}\boldsymbol{y}) = \begin{cases} \delta_{\hat{x}_1,x_1} \cdots \delta_{\hat{x}_N,x_N} & (\boldsymbol{x},\boldsymbol{y}) \in \mathcal{W} \\ 1/d & (\boldsymbol{x},\boldsymbol{y}) \notin \mathcal{W} \end{cases} \tag{5}$$

where $\widehat{\boldsymbol{x}} = (\widehat{x}_1, \dots, \widehat{x}_N)$ denotes an output question tuple, $\boldsymbol{x}\boldsymbol{y} = (x_1, y_1, \dots, x_N, y_N)$ denotes input question-answer pairs, and $d = |\mathcal{X}|$ denotes the total number of question tuples. Intuitively, the MAC $\mathcal{N}_G$ transmits the question tuple without any noise if the input question-answer pairs win the game, else a question tuple chosen uniformly at random is output by the MAC. Consequently, we can expect the sum capacity of the MAC $\mathcal{N}_G$ to increase with the winning probability of the game [5].

This motivates us to allow the senders to share some given set of correlations, so as to increase the winning probability of the game. By a correlation, we mean a conditional probability distribution $P(\boldsymbol{y}'|\boldsymbol{x}\boldsymbol{y})$, where $\boldsymbol{x}\boldsymbol{y} \in \mathcal{X}\mathcal{Y}$ is the input to the MAC, while $\boldsymbol{y}' \in \mathcal{Y}$ is an answer to the nonlocal game G. We allow the senders to perform a local post-processing f of the answers generated by the shared correlation. By this, we mean $f_i(\overline{y}_i|x_i, y_i, y_i')$ is a probability distribution over the answers $\overline{y}_i \in \mathcal{Y}_i$, given the input question-answer pair $(x_i, y_i) \in \mathcal{X}_i \times \mathcal{Y}_i$, and the answer $y_i' \in \mathcal{Y}_i$ generated by the correlation P for each $i \in [N]$, so that

$$f(\overline{\boldsymbol{y}}|\boldsymbol{x}\boldsymbol{y}, \boldsymbol{y}') = \prod_{i=1}^{N} f_i(\overline{y}_i|x_i, y_i, y_i'). \tag{6}$$

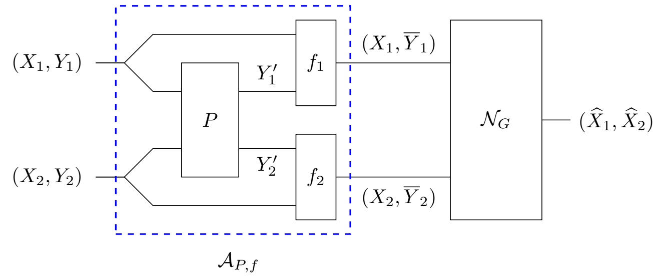

Fig. 1. Correlation-assisted MAC $\mathcal{N}_G \circ \mathcal{A}_{P,f}$ for the case of two senders, obtained from the nonlocal games MAC $\mathcal{N}_G$ defined in Eq. (5) and correlation-assistance channel $\mathcal{A}_{P,f}$ defined in Eq. (7).

Then, given a correlation P and some local post-processing f, we can define the channel $\mathcal{A}_{P,f}$ having input and output alphabets $\mathcal{X}\mathcal{Y}$ and the probability transition matrix

$$\mathcal{A}_{P,f}(\overline{\boldsymbol{x}\boldsymbol{y}}|\boldsymbol{x}\boldsymbol{y}) = \delta_{\overline{\boldsymbol{x}},\boldsymbol{x}} \sum_{\boldsymbol{y}' \in \mathcal{Y}} f(\overline{\boldsymbol{y}}|\boldsymbol{x}\boldsymbol{y}, \boldsymbol{y}') P(\boldsymbol{y}'|\boldsymbol{x}\boldsymbol{y}). \tag{7}$$

Essentially, the channel $\mathcal{A}_{P,f}$ computes an answer to the input question using the shared correlation P and local post-processing by the senders. If P is a nonlocal correlation, then it is possible that the winning probability of the game improves compared to the classical scenario.

Therefore, we define the correlation-assisted MAC $\mathcal{N}_G \circ \mathcal{A}_{P,f}$ corresponding to assistance from the correlation P and local post-processing f. A schematic of this procedure for the case of two senders is shown in Fig. 1.

Then, given a set of correlations $\mathcal{C}$, we define the *$\mathcal{C}$-assisted achievable rate region* of the MAC $\mathcal{N}_G$ as

$$\mathrm{Cap}_{\mathcal{C}}^{(1)}(\mathcal{N}_G) = \bigcup_{\substack{P \in \mathcal{C}, \\ f \in \mathrm{PP}}} \mathrm{Cap}(\mathcal{N}_G \circ \mathcal{A}_{P,f}), \tag{8}$$

where PP denotes the set of local post-processings. Subsequently, the *$\mathcal{C}$-assisted achievable sum rate* of the MAC $\mathcal{N}_G$ is given by

$$\begin{aligned} S_{\mathcal{C}}(\mathcal{N}_G) &= \sup \left\{ \sum_{i=1}^{N} R_i \mid (R_1, \dots, R_N) \in \mathrm{Cap}_{\mathcal{C}}^{(1)}(\mathcal{N}_G) \right\} \\ &= \sup_{P \in \mathcal{C}, f \in \mathrm{PP}} \sup_{p^{(1)} \dots p^{(N)}} I(X_1, Y_1, \dots, X_N, Y_N; Z). \end{aligned} \tag{9}$$

Here, $p^{(i)}$ is the probability distribution of the random variable (X_i, Y_i) associated with the input question-answer pair for $i \in [N]$ and Z is the random variable describing the output of the MAC $\mathcal{N}_G \circ \mathcal{A}_{P,f}$. Denoting cl as classical correlations, Q as quantum correlations, and NS as no-signalling correlations, we obtain the hierarchy

$$S(\mathcal{N}_G) \le S_{\mathrm{cl}}(\mathcal{N}_G) \le S_{\mathrm{Q}}(\mathcal{N}_G) \le S_{\mathrm{NS}}(\mathcal{N}_G) \le C(\mathcal{N}_G) \tag{10}$$

Note that we call a correlation $P(\boldsymbol{y}'|\boldsymbol{x}\boldsymbol{y})$ classical if it is a convex combination of product distributions of the form

$\prod_{i=1}^N P_i(y_i'|x_i,y_i)$ [7]. Because we allow for convex combinations of product distributions, the sum capacity $S(\mathcal{N}_G)$ might not be equal to $S_{\text{cl}}(\mathcal{N}_G)$. The correlation P is said to be quantum if the senders share a quantum state ρ, and given a question-answer pair (x_i,y_i), the ith player performs a local measurement with some POVM $\{E_y^{(x_i,y_i)} \mid y \in \mathcal{Y}_i\}$ to obtain the answer y' [7]. Subsequently, the quantum correlation P can be described as

$$P(y_1',\ldots,y_N'|x_1,y_1,\ldots,x_N,y_N) = \mathrm{Tr}\left[\rho\left(E_{y_1'}^{(x_1,y_1)} \otimes \cdots \otimes E_{y_N'}^{(x_N,y_N)}\right)\right]. \quad (11)$$

Finally, the correlation P is said to be no-signalling if

$$P(y_i'|x_1,y_1,\ldots,x_N,y_N) = P(y_i'|x_i,y_i), \quad i \in [N], \quad (12)$$

which amounts to saying that each sender is not aware of the question-answer pairs held by the other senders [8]. Because we allow all possible probability distributions while computing $C(\mathcal{N}_G)$, it sits at the top of the hierarchy.

Since $S(\mathcal{N}_G) \le S_{\mathcal{C}}(\mathcal{N}_G)$ for any set of correlations $\mathcal{C}$, our goal is to find an upper bound for $S_{\mathcal{C}}(\mathcal{N}_G)$. We begin with the inequality

$$S_{\mathcal{C}}(\mathcal{N}_G) \le \sup_{\overline{p}_{\boldsymbol{Y}|\boldsymbol{X}} \in \mathfrak{S}_{\mathcal{C}}} \max_{\pi} I(\overline{X}_1,\overline{Y}_1,\ldots,\overline{X}_N,\overline{Y}_N;Z), \quad (13)$$

where $\overline{X}_1,\overline{Y}_1,\ldots,\overline{X}_N,\overline{Y}_N$ are inputs to $\mathcal{N}_G$, while Z describes its output. Here, $\mathfrak{S}_{\mathcal{C}}$ denotes the set of strategies induced by the correlations $\mathcal{C}$, defined as

$$\mathfrak{S}_{\mathcal{C}} = \left\{\mathcal{A}_{P,f}(p_{\boldsymbol{Y}|\boldsymbol{X}}) \mid p_{\boldsymbol{Y}|\boldsymbol{X}} = \prod_{i=1}^N p_{Y_i|X_i},\ P \in \mathcal{C},\ f \in \mathrm{PP}\right\} \quad (14)$$

where

$$\mathcal{A}_{P,f}(p_{\boldsymbol{Y}|\boldsymbol{X}}) = \sum_{\boldsymbol{y}\in\mathcal{Y}} \sum_{\boldsymbol{y}'\in\mathcal{Y}} f(\overline{\boldsymbol{y}}|\overline{\boldsymbol{x}}\boldsymbol{y},\boldsymbol{y}')P(\boldsymbol{y}'|\overline{\boldsymbol{x}}\boldsymbol{y})p_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\overline{\boldsymbol{x}}) \quad (15)$$

is the strategy induced by the channel $\mathcal{A}_{P,f}$. Since the optimization in Eq. (13) is non-convex and difficult to solve, we perform two relaxations:

1) We allow all distribution over questions and solve the convex optimization problem $\max_\pi I$ (in relevant cases), where I is the mutual information between the inputs and output of $\mathcal{N}_G$.
2) We allow all strategies with winning probability less than or equal to $\omega^{\mathfrak{S}_{\mathcal{C}}}(G)$ when questions are drawn uniformly.

This leads us to the following result.

Theorem 1. *Let G be an N-player promise-free nonlocal game with d question tuples, and let $\mathcal{N}_G$ be the MAC obtained from G as defined in Eq. (5). Suppose that the senders share a set of correlations $\mathcal{C}$. Let $\mathfrak{S}_{\mathcal{C}}$ be the set of strategies induced by the correlations as defined in Eq. (14). Let $\omega^{\mathfrak{S}_{\mathcal{C}}}(G)$ denote the maximum winning probability of the game when the questions are drawn uniformly and answers are obtained using strategies in $\mathfrak{S}_{\mathcal{C}}$. Let $S_{\mathcal{C}}(\mathcal{N}_G)$ denote the $\mathcal{C}$-assisted achievable sum rate of the MAC $\mathcal{N}_G$ as defined in Eq. (9). Then, we have*

$$S(\mathcal{N}_G) \le S_{\mathcal{C}}(\mathcal{N}_G) \le \ln\left(d-1+d^{-(1-\omega^{\mathfrak{S}_{\mathcal{C}}}(G))d}\right) \quad (16)$$

with entropy measured in nats. In particular, we have

$$S(\mathcal{N}_G) \le \ln\left(d-1+d^{-(1-\omega^{cl}(G))d}\right). \quad (17)$$

Proof strategy. The proof of Theorem 1 is based on the following procedure outlining the main proof ideas. For precise technical statements, we refer to the full version [6].

1) Given a strategy $\overline{p}_{\boldsymbol{Y}|\boldsymbol{X}} \in \mathfrak{S}_{\mathcal{C}}$ for playing the game G induced by the correlations $\mathcal{C}$, we define the *winning vector* $\boldsymbol{w} = (w_1,\ldots,w_d) \in [0,1]^d$ as
$$w_i = \sum_{\boldsymbol{y}:\boldsymbol{x}_i\boldsymbol{y}\in\mathcal{W}} \overline{p}_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x}_i)$$
for $i \in [d]$. Note that w_i gives the probability of winning the game using the chosen strategy upon receiving question $i \in [d]$. Given a strategy $\overline{p}_{\boldsymbol{Y}|\boldsymbol{X}}$ and a distribution $\pi \in \Delta_d$ over the questions, we show that the mutual information $\mathscr{I}_{\boldsymbol{w}}(\pi) := I(\overline{X}_1,\overline{Y}_1,\ldots,\overline{X}_N,\overline{Y}_N;Z)$ between the inputs and the output of $\mathcal{N}_G$ can be expressed as
$$\mathscr{I}_{\boldsymbol{w}}(\pi) = H(\overline{W}\pi) + \langle \pi, \boldsymbol{w}\rangle \ln(d) - \ln(d),$$
where H is the Shannon entropy and $\overline{W}$ is a $d\times d$ matrix with elements $\overline{W}_{ij} = w_i\delta_{ij} + (1-w_j)/d$.
2) Given a winning vector $\boldsymbol{w} \in \{0,1\}^d$ with components 0 or 1 (e.g., those vectors given by deterministic strategies) that can answer $0 < K < d$ out of the d questions correctly, we show that
$$\mathscr{I}_K^* := \max_{\pi\in\Delta_d} \mathscr{I}_{\boldsymbol{w}}(\pi) = \ln\left(K + (d-K)d^{-\frac{d}{d-K}}\right).$$
3) We then show that winning vectors $\boldsymbol{w}$ with at most $K < d$ non-zero entries (i.e., at most K questions are answered correctly) satisfy
$$\max_{\pi\in\Delta_d} \mathcal{I}_{\boldsymbol{w}}(\pi) \le \mathscr{I}_{d-1}^*.$$
Therefore, we restrict our attention to winning vectors $\boldsymbol{w}$ with non-zero entries, and show that
$$\mathscr{I}^*(\boldsymbol{w}) := \max_{\pi\in\Delta_d} \mathscr{I}_{\boldsymbol{w}}(\pi) = \ln\left(\sum_{j=1}^d \exp\left[d w_{\text{eff}} \ln d\left(1-\frac{1}{w_j}\right)\right]\right),$$
where $w_{\text{eff}} = (\sum_{i=1}^d w_i^{-1})^{-1}$.
4) Finally, we allow all strategies having winning probability less than or equal to $\omega^{\mathfrak{S}_{\mathcal{C}}}(G)$ when the questions are drawn uniformly. The corresponding set of winning vectors is
$$\overline{\mathfrak{W}_{\mathcal{C}}} = \left\{\boldsymbol{w} \in [0,1]^d \;\middle|\; \frac{1}{d}\sum_{i=1}^d w_i \le \omega^{\mathfrak{S}_{\mathcal{C}}}(G)\right\},$$
and subsequently, we show that
$$\sup_{\boldsymbol{w}\in\overline{\mathfrak{W}_{\mathcal{C}}},\boldsymbol{w}>0} \mathscr{I}^*(\boldsymbol{w}) \le \ln\left(d-1+d^{-(1-\omega^{\mathfrak{S}_{\mathcal{C}}}(G))d}\right).$$

Noting that the above bound is greater than or equal to $\mathscr{I}_{d-1}^*$, we obtain the desired result. □

Now, if $\mathcal{C}$ is any set of correlations such that $\omega^{\mathcal{C}}(G) = 1$, then $S_{\mathcal{C}}(\mathcal{N}_G) = \ln(d)$. We use this fact to obtain separations of correlation-assisted achievable sum rates. For the magic square game G_{MS}, the maximum classical winning probability is $\omega^{\text{cl}}(G_{\text{MS}}) = 8/9$, whereas a perfect quantum strategy is available [9], [10], [11], [12]. From Eq. (17), it follows that

$$S(\mathcal{N}_{G_{\text{MS}}}) \leq 3.02 \text{ bits}, \tag{18}$$

while we have

$$S_{\text{Q}}(\mathcal{N}_{G_{\text{MS}}}) = \log(9) = 3.17 \text{ bits}. \tag{19}$$

Thus, we find a tighter bound on the sum capacity than the previously reported bound of 3.14 bits [5].

In the same spirit, for obtaining a separation between the sum capacity and NS-assisted achievable sum rate, we consider the CHSH game G_{CHSH} [13], [14]. For the CHSH game, a perfect no-signalling strategy is available, but there is no perfect classical or quantum strategy: $\omega^{\text{cl}}(G_{\text{CHSH}}) = 3/4$ and $\omega^{\text{Q}}(G_{\text{CHSH}}) = (1 + 1/\sqrt{2})/2$ [14]. Therefore, using Eq. (17), we obtain

$$S(\mathcal{N}_{G_{\text{CHSH}}}) \leq 1.7 \text{ bits} \tag{20}$$
$$S_{\text{Q}}(\mathcal{N}_{G_{\text{CHSH}}}) \leq 1.78 \text{ bits} \tag{21}$$
$$S_{\text{NS}}(\mathcal{N}_{G_{\text{CHSH}}}) = \log(4) = 2 \text{ bits}. \tag{22}$$

Such bounds given by our approach help place strict limits on the sum capacity of discrete MACs assisted by any given set of correlations.

Since the computation of $C(\mathcal{N}_G)$ involves the optimization over all possible probability distributions, it amounts to allowing assistance from all possible strategies. It is natural to ask if one can find a game which can be won with full communication between the players, but cannot be won by players using no-signalling strategies. We answer this question in the affirmative by introducing the signalling game in the next section.

III. Looseness of Relaxed Sum Capacity

We now introduce a non-local game G_{s} that we call the 'signalling' game, whose derived MAC $\mathcal{N}_{G_{\text{s}}}$ exhibits an arbitrarily large separation between the no-signalling-assisted and the full-communication-assisted sum capacities. In the signalling game G_{s}, the players Alice & Bob are each given a question from some set of questions $\mathcal{X}_1, \mathcal{X}_2$, and they win the game if they can correctly guess the question handed over to the other person. The name of the game stems from the requirement for Alice & Bob to "signal" their question to the other person in order to win the game.

It can be shown that

$$\omega^{\text{cl}}(G_{\text{s}}) = \frac{1}{\max(|\mathcal{X}_1|, |\mathcal{X}_2|)}$$

using classical strategies. Therefore, $\omega^{\text{cl}}(G_{\text{s}}) \to 0$ as $d \to \infty$. Furthermore, we numerically verified that $\omega^{\text{NS}}(G_{\text{s}}) = \omega^{\text{cl}}(G_{\text{s}})$ holds for all pairs $(|\mathcal{X}_1|, |\mathcal{X}_2|)$ with $2 \leq |\mathcal{X}_1|, |\mathcal{X}_2| \leq 10$ (i.e., up to $d = 100$). This suggests that even no-signalling strategies are of no assistance for winning the signalling game. Therefore, as $d \to \infty$, we have $S(\mathcal{N}_{G_{\text{s}}}) \to 0$, and we also expect $S_{\text{Q}}(\mathcal{N}_{G_{\text{s}}}) \to 0$ and $S_{\text{NS}}(\mathcal{N}_{G_{\text{s}}}) \to 0$ to be true.

On the other hand, there is always a perfect strategy for winning the signalling game if we allow communication between Alice & Bob. Subsequently, we have $C(\mathcal{N}_G) = \ln(d) \to \infty$ as $d \to \infty$. Therefore, we can find an arbitrarily large separation between the sum capacity and the relaxed sum capacity. A similar statement is expected to hold for the no-signalling assisted achievable sum rate and the relaxed sum capacity.

This example demonstrates the need to look for other techniques to bound the sum capacity. In the next section, we take a step in that direction by proposing an algorithm that can compute the sum capacity of a two-sender MAC to any given precision.

IV. Computing the Sum Capacity of Two-Sender MACs

Consider a two-sender MAC with input alphabets $\mathcal{A}_1$, $\mathcal{A}_2$ of size d_1, d_2 respectively, an output alphabet $\mathcal{Z}$ of size d, and a probability transition matrix $\mathcal{N}$. The sum capacity of $\mathcal{N}$ given by Eq. (2) results in optimization over a non-convex set due to the product distribution constraint. For convenience, we rewrite Eq. (2) as an optimization over the convex set $\Delta_{d_1} \times \Delta_{d_2}$, effectively folding the non-convexity into the objective function by taking a product of the distributions in the computation of the mutual information. Here, $\Delta_n = \{x \in \mathbb{R}^n \mid x \geq 0, \sum_{i=1}^n x = 1\}$ denotes the standard $(n-1)$-dimensional simplex. In other words, we write the sum capacity as

$$S(\mathcal{N}) = \max_{p \in \Delta_{d_1}} \max_{q \in \Delta_{d_2}} I(p, q) \tag{23}$$

where $I(p, q)$ denotes the mutual information between the inputs and the output of the MAC, given input probability distributions p and q.

Note that the mutual information can be written as

$$I(p, q) = H(A_q p) - \langle b_q, p \rangle \tag{24}$$

where $A_q(z, a_1) = \sum_{a_2 \in \mathcal{A}_2} \mathcal{N}(z|a_1, a_2) q(a_2)$ and $b_q(a_1) = -\sum_{a_2 \in \mathcal{A}_2} q(a_2) \sum_{z \in \mathcal{Z}} \mathcal{N}(z|a_1, a_2) \log(\mathcal{N}(z|a_1, a_2))$. It can be seen that the mutual information is a concave function of p for a fixed q, but it is not jointly concave in p and q. Consequently, the function

$$I^*(q) = \max_{p \in \Delta_{d_1}} \left(H(A_q p) - \langle b_q, p \rangle\right) \tag{25}$$

can be computed using standard techniques in convex optimization, but the sum capacity computation $S(\mathcal{N}) = \max_{q \in \Delta_{d_2}} I^*(q)$ is a non-convex problem.

In order to solve this non-convex problem, we show that the function I^* satisfies the following property. For any $q, q' \in \Delta_{d_2}$, we have

$$|I^*(q) - I^*(q')| \leq \beta_I \left(\|q - q'\|_1\right) \tag{26}$$

where $\|\cdot\|_1$ denotes the l_1-norm. The function β_I is given by

$$\beta_I(x) = \left(\frac{1}{2}\log(d-1) + H_{\mathcal{N}}^{\max}\right)x + \overline{h}\left(\frac{x}{2}\right), \qquad (27)$$

where

$$H_{\mathcal{N}}^{\max} = \max_{a_1\in\mathcal{A}_1, a_2\in\mathcal{A}_2} -\sum_{z\in\mathcal{Z}} \mathcal{N}(z|a_1,a_2)\log(\mathcal{N}(z|a_1,a_2)),$$

and

$$\overline{h}(x) = \begin{cases} -x\log(x) - (1-x)\log(1-x) & \text{if } x \le \frac{1}{2} \\ \log(2) & \text{if } x \ge \frac{1}{2} \end{cases}$$

is a modified binary entropy. Importantly, β_I is a non-negative, monotonically increasing function with $\beta_I(0) = 0$. Then, owing to Eq. (26), we refer to I^* as a Lipschitz-like function.

The non-convex problem $S(\mathcal{N}) = \max_{q\in\Delta_{d_2}} I^*(q)$ can be numerically solved using a generalization of the Piyavski-Shubert algorithm. We give a high-level overview of our algorithm here, and present a complete description of the algorithm along with convergence analysis in the full version [6]. Given any Lipschitz-like function $g\colon [a,b] \to \mathbb{R}$ satisfying $|g(x)-g(y)| \le \beta(|x-y|)$, the algorithm proceeds by constructing successively better upper bounds on g using the Lipschitz-like property. Given a single iterate $t \in [a,b]$, the upper-bounding function is $G(x) = g(t) + \beta(|x-t|)$ for $x \in [a,b]$. As the iterations proceed, this bounding function can be refined, for example, by taking the minimum of the bounding function corresponding to adjacent iterates. Then, at each iteration, the maximum of the bounding function is computed, which is an easier problem because we only need to maximize a function depending on β that is known to be continuous and monotonically increasing. The computed maximum of the upper-bounding function at each iteration generates a sequence of points that partitions the interval. When the distance between any two of these points becomes sufficiently small, one can show that the maximum of the upper-bounding function is close to the maximum of g. In this manner, one can either compute the maximum of g to a desired precision or one can compute an upper bound on the maximum by running the algorithm for a fixed number of iterations.

For optimizing g over the standard simplex in higher dimensions, we present an algorithm to construct a Lipschitz-like dense curve filling the simplex. Using this curve, we can reduce the multi-dimensional problem of computing the maximum of g over the simplex to a one-dimensional problem of optimization over an interval. This one-dimensional problem can then be solved using the generalized Piyavski-Shubert algorithm described above. Applying our algorithm to the function I^* which is Lipschitz-like as per Eq. (26), we can compute the sum capacity of a two-sender MAC to a fixed precision, or compute an upper bound on the sum capacity by fixing the number of iterations.

To test the performance of this algorithm, we construct a binary MAC whose sum capacity can be computed analytically. For the binary MAC with transition probability matrix

$$\mathcal{N}_F = \begin{pmatrix} 1 & 0.5 & 0.5 & 0 \\ 0 & 0.5 & 0.5 & 1 \end{pmatrix},$$

we show that the sum capacity is equal to $S(\mathcal{N}_F) = 0.5\ln(2) \approx 0.3466$ nats. Furthermore, for this channel, we have $C(\mathcal{N}_F) = \ln(2) \approx 0.693$ nats. We find that our algorithm gives a value of ≈ 0.3459 nats for the sum capacity corresponding to a precision of $\epsilon = 0.1$, while it gives a value of ≈ 0.3466 nats corresponding to $\epsilon = 0.01$. Therefore, we are able to verify that our algorithm can perform better than the relaxed sum capacity.

While the algorithm performs efficiently for the case of a binary MAC, the number of iterations required for convergence can scale exponentially with the size of the input alphabet d_2 over which we perform the non-convex optimization. It remains to see whether this is a shortcoming of this algorithm or whether computing the sum capacity to a fixed precision is an intrinsically hard problem. Therefore, as of now, our algorithm is useful for computing the sum capacity when one of the input alphabets is of small size. For larger problem sizes, it may be suitable to look for alternate approaches, such as those which extend [15], [16] to discrete MACs.

References

[1] R. Ahlswede, "The capacity region of a channel with two senders and two receivers," *The annals of probability*, pp. 805–814, 1974.

[2] H. H.-J. Liao, "Multiple access channels." Univeristy of Hawaii, Tech. Rep., 1972.

[3] E. Calvo, D. P. Palomar, J. R. Fonollosa, and J. Vidal, "On the computation of the capacity region of the discrete mac," *IEEE transactions on communications*, vol. 58, no. 12, pp. 3512–3525, 2010.

[4] J. Bühler and G. Wunder, "A note on capacity computation for the discrete multiple access channel," *IEEE transactions on information theory*, vol. 57, no. 4, pp. 1906–1910, 2011.

[5] F. Leditzky, M. A. Alhejji, J. Levin, and G. Smith, "Playing games with multiple access channels," *Nature communications*, vol. 11, no. 1, pp. 1–5, 2020.

[6] A. Seshadri, F. Leditzky, V. Siddhu, and G. Smith, "On the separation of correlation-assisted sum capacities of multiple access channels," 2022, manuscript in preparation.

[7] C. Palazuelos and T. Vidick, "Survey on nonlocal games and operator space theory," *Journal of Mathematical Physics*, vol. 57, no. 1, p. 015220, 2016.

[8] V. Russo, "Extended nonlocal games," *arXiv preprint arXiv:1704.07375*, 2017.

[9] G. Brassard, A. Broadbent, and A. Tapp, "Quantum pseudo-telepathy," *Foundations of Physics*, vol. 35, no. 11, pp. 1877–1907, 2005.

[10] N. D. Mermin, "Simple unified form for the major no-hidden-variables theorems," *Physical review letters*, vol. 65, no. 27, p. 3373, 1990.

[11] A. Peres, "Incompatible results of quantum measurements," *Physics Letters A*, vol. 151, no. 3-4, pp. 107–108, 1990.

[12] P. Aravind, "A simple demonstration of bell's theorem involving two observers and no probabilities or inequalities," *arXiv preprint quant-ph/0206070*, 2002.

[13] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, "Proposed experiment to test local hidden-variable theories," *Physical review letters*, vol. 23, no. 15, p. 880, 1969.

[14] R. Cleve, P. Hoyer, B. Toner, and J. Watrous, "Consequences and limits of nonlocal strategies," in *Proceedings. 19th IEEE Annual Conference on Computational Complexity, 2004.* IEEE, 2004, pp. 236–249.

[15] Z. Aharoni, D. Tsur, Z. Goldfeld, and H. H. Permuter, “Capacity of continuous channels with memory via directed information neural estimator,” in *2020 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2020, pp. 2014–2019.
[16] F. Mirkarimi and N. Farsad, “Neural computation of capacity region of memoryless multiple access channels,” *arXiv preprint arXiv:2105.04453*, 2021.