Homologous Codes for Multiple Access Channels

Pinar Sen, *Student Member, IEEE*, and Young-Han Kim, *Fellow, IEEE*

***Abstract*— Building on recent development by Padakandla and Pradhan, and by Lim, Feng, Pastore, Nazer, and Gastpar, this paper studies the potential of structured coset coding as a complete replacement for random coding in network information theory. The roles of two techniques used in coset coding to generate nonuniform codewords, namely, shaping and channel transformation, are clarified and illustrated via the simple example of the two-sender multiple access channel. While individually deficient, the optimal combination of shaping via nested coset codes of the same generator matrix (which we refer to as homologous codes) and channel transformation is shown to achieve the same performance as traditional random codes for the general two-sender multiple access channel. The achievability proof of the capacity region is extended to multiple access channels with more than two senders, and with one or more receivers. A quantization argument adapted to the proposed combination of two techniques is presented to establish the achievability proof for their Gaussian counterparts. It is illustrated by an example that combining shaping and channel transformation is useful even when the goal of transmission for a subset of the receivers is to recover a linear combination of messages. These results open up new possibilities of utilizing homologous codes for a broader class of applications.**

***Index Terms*— Multiple access channels (MACs), nested coset codes, algebraic network information theory, linear codes, communication and computation with linear codes.**

I. Introduction

RANDOM independently and identically distributed (i.i.d.) code ensembles play a fundamental role in network information theory, with most existing coding schemes built on them; see, for example, [1]–[3]. As shown by the classical example by Körner and Marton [4], however, using the same code at multiple users can achieve strictly better performance for some communication problems. Recent studies illustrate the benefit of such *structured coding* for computing linear combinations [5]–[10], for the interference channels [11]–[13], and for multiple access channels (MACs) with state information [14]. Consequently, there has been a flurry of research activities on structured coding in network information theory, facilitated in part by several standalone workshops and tutorials at major conferences by leading researchers.

Most of the existing results are based on lattice codes or linear codes on finite alphabets. Recently, Padakandla and Pradhan [14] brought a new dimension to the arsenal of structured coding by developing nested coset codes as a method to shape the input distribution of the channel; see also Miyake [15] for nested coset codes for point-to-point communication. In these nested coset coding schemes, a coset code of a rate higher than the target is first generated randomly. A codeword of a desired property (such as type or joint type) is then selected from a subset (a coset of a subcode). This construction is reminiscent of the multicoding scheme in Gelfand–Pinsker coding for channels with state and Marton coding for broadcast channels. But in a sense, nested coset coding is more fundamental in that the scheme at its core is relevant even for single-user point-to-point communication. By a careful combination of individual and common parts of coset codes, the coding scheme proposed by Padakandla and Pradhan [14] achieves the rates for multiple access channels with state beyond what can be achieved by existing random or structured coding schemes. The analysis of the scheme is performed by packing and covering lemmas for nested coset codes that parallel such lemmas for random i.i.d. coding [1].

Recently, structured coding based on random nested coset codes was further streamlined by Lim *et al.* [8]. With the primary motivation of communicating linear combinations of codewords over a multiple access channel (as in the celebrated compute–forward scheme [6], [16]), they augmented the original nested coset coding schemes in [14], [15] with the channel transformation technique by Gallager [17, Sec. 6.2] and developed new analysis tools when multiple senders use nested coset codes with a common generator matrix. The resulting achievable rate region, when adapted to the Gaussian case, improves upon the previous result for compute–forward [6].

In both [14] and [8], however, structured coding of nested coset codes is reserved for rather niche communication scenarios of adapting multiple codewords to a common channel state or computing sums of codewords, and even in these limited cases, as a complement to random i.i.d. coding. The coding scheme in [14] uses superposition of codewords with individual and common generator matrices. A similar coding scheme in [13] for three-user interference channels again uses a combination of random i.i.d. coding (for message decoding) and structured coding (for function decoding) of nested coset codes, this time with a more explicit superposition coding architecture. There is also some indication that the benefit of computation can be realized to the full extent only in special cases for which desired linear combinations and channel structures are matched [18]. In the same vein, for a given code distribution, the aforementioned rate region for computation

Manuscript received January 18, 2018; revised May 30, 2019; accepted November 8, 2019. Date of publication November 28, 2019; date of current version February 14, 2020. This work was supported in part by the Electronics and Telecommunications Research Institute (ETRI) from the Korean Ministry of Science, ICT, and Future Planning, under Grant 17ZF1100. This article was presented in part at the 2017 IEEE International Symposium of Information Theory.

The authors are with the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92093 USA (e-mail: psen@ucsd.edu; yhk@ucsd.edu).

Communicated by A. Tchamkerten, Associate Editor for Shannon Theory.

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2019.2956533

in [8] turns out to be strictly smaller than the polymatroidal region that is achievable by random i.i.d. coding (refer to (1) for the formal description) when computation is specialized to communication (i.e., the identity function computation). In this regard, Lim *et al.* [8] have recently improved their analysis to establish a larger achievable rate region for message communication [19], which is still strictly smaller than the targeted polymatroidal region. It should be noted that the corner points of the polymatroidal region are included in this achievable rate region, and hence that structured coding followed by time sharing achieves the same rate region as random i.i.d. coding and thus achieves the entire capacity region. This time sharing idea, however, is not sufficient to achieve optimal rates for more complicated network models, such as multiple-receiver MACs [20] or interference channels [12] since each receiver requires different time allocation among the senders. Apparently, structured coding, even based on the promising new technique of nested coset codes, can only play a complementary role to random i.i.d. coding.

This paper aims to illustrate that at least for simple communication networks, the opposite is true, and that structured coding can completely subsume random i.i.d. coding. In particular, we show that a random ensemble of nested coset codes of the same generator matrix (which we referred to as *homologous codes* [21]), which was thought to be good only for recovering linear combinations, can achieve the same rates as independently generated linear or nonlinear random codes for the task of communicating individual codewords over MACs.

For the simplicity of the exposition, we start our discussion with two-sender MACs and show that the pentagonal region achievable by random i.i.d. codes is achievable without time sharing by a careful construction of random homologous codes. Our finding relies on the identification of *shaping* and *channel transformation* techniques—both of which are used to improve upon conventional coset codes by allowing nonuniform codewords—as key components to supplant random i.i.d. coding by structured coding. We first evaluate achievable rates of individual techniques, which fall short of the target. We then combine these two techniques to obtain the achievability of the pentagonal region.

Of the key components, shaping via homologous codes was analyzed in [19]. We provide a different analysis to obtain an achievable rate region that is smaller but easier to compute than [19]. The idea of channel transformation, the other key component, is commonly used in network information theory literature for converting the channel input alphabet into a *desired* one, such as a finite field as in [19]. Gallager [17, Sec. 6.2], on the other hand, makes use of channel transformation for another purpose—to modify the distribution of a random linear code ensemble by converting the channel input alphabet to a large enough *extension* of the desired finite field. The perspective of allowing extension fields for the channel transformation does not appear in the multiple-user structured coding literature. Following a similar idea to Gallager [17, Sec. 6.2] for MACs, we first present a structured coding scheme that utilizes channel transformation applied on coset codes and we provide an achievable rate region for this scheme. We then incorporate channel transformation into homologous codes instead of coset codes to bring extra freedom by shaping. This combination helps us create asymmetry between senders to remedy the effect of aligning their codebooks into the same span, which leads to the achievability of the pentagonal region. It is illustrated by an example that the flexibility of adjusting the amount of asymmetry is useful even when there are multiple receivers one of which wants to *compute* a linear combination of messages whereas others want to recover the messages themselves. Even in this case, careful combination of shaping and channel transformation using extension fields strictly outperforms random i.i.d. codes as well as homologous codes constructed in the desired finite field.

These results are extended to MACs with more than two senders and with one or more receivers, and it is shown that combination of the shaping and the channel transformation can achieve the capacity region in general. On the contrary, it is illustrated by an example of two-receiver MAC that the shaping even with convexification via time-sharing (that corresponds to [19]) achieves a strictly smaller rate region than the capacity. Finally, the achievability of the capacity region for the Gaussian counterparts is shown via a different quantization argument that is based on partitioning the sample space into equiprobable quantiles and is more convenient for our construction of homologous codes than the conventional uniform quantization.

The rest of the paper is organized as follows. Section II formulates the problem and defines nested coset codes and homologous codes. Section III discusses the running examples of binary adder and binary erasure multiple access channels. The main results for the two-sender MAC are presented in Section IV, and are extended to more than two senders in Section V and one or more receivers in Section VI. Section VII presents the achievability of the capacity region for the Gaussian MAC. The problem of simultaneous communication and computation via homologous codes is discussed in Section VIII. Section IX concludes the paper.

We adapt the notation in [1], [2]. The set of integers $\{1, 2, \ldots, n\}$ is denoted by $[n]$. For a length-n sequence (vector) $x^n = (x_1, x_2, \ldots, x_n) \in \mathcal{X}^n$, we define its type as $\pi(x|x^n) = |\{i : x_i = x\}|/n$ for $x \in \mathcal{X}$. Upper case letters $X, Y, \ldots$ denote random variables. For $\epsilon \in (0, 1)$, we define the ϵ-typical set of length-n sequences (or the typical set in short) as $\mathcal{T}_\epsilon^{(n)}(X) = \{x^n : |p(x) - \pi(x|x^n)| \le \epsilon p(x),\ x \in \mathcal{X}\}$. A tuple of k random variables $(X_1, X_2, \ldots, X_k)$ is denoted by X^k, and for $\mathcal{J} \subseteq [k]$, the subtuple of random variables with indices from $\mathcal{J}$ is denoted by $X(\mathcal{J}) = (X_i : i \in \mathcal{J})$. The indicator function $\mathbb{1}_{\mathcal{S}} : \mathcal{X} \to \{0, 1\}$ for $\mathcal{S} \subseteq \mathcal{X}$ is defined as $\mathbb{1}_{\mathcal{S}}(x) = 1$ if $x \in \mathcal{S}$ and 0 otherwise. A length-n vector of all zeros (ones) is denoted by $\mathbf{0}_n$ $(\mathbf{1}_n)$, where the subscript is omitted when it is clear in the context. An $m \times n$ matrix of all zeros is denoted by $O_{m\times n}$. The $n \times n$ identity matrix is denoted by I_n. For $\alpha \in (0, 1)$, $H(\alpha) := -\alpha \log_2(\alpha) - (1-\alpha) \log_2(1-\alpha)$ denotes the binary entropy function. $\mathbb{F}_q$ denotes a finite field of order q.

II. Problem Formulation

Consider the k-sender discrete memoryless (DM) multiple access channel (MAC)

$$(\mathcal{X}_1 \times \mathcal{X}_2 \times \ldots \times \mathcal{X}_k, p(y|x_1, x_2, \ldots, x_k), \mathcal{Y}),$$

which consists of k sender alphabets $\mathcal{X}_j$, $j \in [k]$, a receiver alphabet $\mathcal{Y}$, and a collection of conditional probability distributions $p_{Y|X^k}(y|x_1, x_2, \ldots, x_k)$.

An $(n, nR_1, nR_2, \ldots, nR_k)$ code for the multiple access channel consists of k message sets, $\mathbb{F}_q^{nR_j}$, $j \in [k]$, k encoders where encoder $j \in [k]$ assigns a codeword $x_j^n(m_j) \in \mathcal{X}_j^n$ to each message $m_j \in \mathbb{F}_q^{nR_j}$, and a decoder that assigns an estimate $(\hat{m}_1, \ldots, \hat{m}_k)$ to each received sequence y^n. The random message tuple $(M_1, \ldots, M_k)$ is assumed to be independent and uniformly distributed. The probability of error is defined as

$$P_e^{(n)} = \mathsf{P}((\hat{M}_1, \ldots, \hat{M}_k) \neq (M_1, \ldots, M_k)).$$

A rate tuple $(R_1, R_2, \ldots, R_k)$ is said to be *achievable* if there exists a sequence of $(n, nR_1, nR_2, \ldots, nR_k)$ codes such that $\lim_{n\to\infty} P_e^{(n)} = 0$. The capacity region is defined as the closure of the set of achievable rate tuples. Single letter characterization of this capacity region was derived in [22], [23] using random i.i.d. coding arguments. For a given probability mass function (pmf) $p(x^k)$, define $\mathscr{R}_{\text{MAC}}(X^k)$ as the set of rate tuples $(R_1, R_2, \ldots, R_k)$ such that

$$\sum_{i\in\mathcal{J}} R_i < I(X(\mathcal{J}); Y|X(\mathcal{J}^c)), \quad \forall \mathcal{J} \subseteq [k]. \tag{1}$$

In (1), by the q-ary code construction, the information rates are in terms of q-ary symbols and the information measures are in log base q. One can divide both sides of the inequalities in (1) by $\log_2 q$ to obtain a set of rate constraints in terms of *bits*. Henceforth, we present all the achievability results in terms of bits by assuming this q-ary to bit conversion is performed. The capacity region is then defined as the convex closure of $\bigcup_{p(x^k)} \mathscr{R}_{\text{MAC}}(X^k)$.

In this work, we are particularly interested in the performance of *homologous* codes that preserve a common structure among different senders. For the ease of exposition, we start with a discrete memoryless channel, i.e., $k = 1$. For the discrete memoryless channel $p(y|x)$, shaping of the channel input distributions via *nested coset* codes was first proposed in [15] and later appeared in [8], [14]. Following a similar notation to these studies, the nested coset codes can be defined as follows.

Definition 1 (Nested coset codes): An $(n, nR, n\hat{R}, \mathbb{F}_q)$ nested coset code consists of a message set $\mathbb{F}_q^{nR}$, a generator matrix $G \in \mathbb{F}_q^{n(R+\hat{R})\times n}$, a coset sequence d^n, a shaping function $s : \mathbb{F}_q^{nR} \to \mathbb{F}_q^{n\hat{R}}$, an encoder that assigns a codeword to each message according to the steps below, and a decoder that assigns an estimate to each received sequence y^n.

1) For each $m \in \mathbb{F}_q^{nR}$ and $l \in \mathbb{F}_q^{n\hat{R}}$, compute
$$x^n(m, l) = [m \;\; l]\, G \oplus d^n. \tag{2}$$
2) For each message $m \in \mathbb{F}_q^{nR}$, choose $x^n(m, s(m))$ as the assigned codeword, where $s(m)$ is the specified shaping function.

Remark 1: An $(n, nR, \mathbb{F}_q)$ *coset code* is a special case of an $(n, nR, n\hat{R}, \mathbb{F}_q)$ nested coset code with $\hat{R} = 0$ (no shaping). Specializing further, we can view an $(n, nR, \mathbb{F}_q)$ *linear code* as an $(n, nR, \mathbb{F}_q)$ coset code with $d^n = \mathbf{0}$.

The encoding steps of nested coset codes can be interpreted as follows. In Step 1), an $(n, n(R + \hat{R}), \mathbb{F}_q)$ coset code, $\mathcal{C}_1$, of rate $R+\hat{R}$ that is larger than the target rate R is created using a generator matrix G, which includes an $(n, nR, \mathbb{F}_q)$ coset code, $\mathcal{C}_2$, generated by the first nR rows of G, as a subcode. Thus, these two coset codes are *nested*, i.e., $\mathcal{C}_2 \subseteq \mathcal{C}_1$. The intentional redundancy in the size of the code $\mathcal{C}_1$ then allows selecting a subset with the desired properties induced by the shaping function in step 2). By the nested construction of $\mathcal{C}_2 \subseteq \mathcal{C}_1$, any selected codeword in $\mathcal{C}_1$ will be in a coset of $\mathcal{C}_2$.

We now continue with a formal description of a *random ensemble* of nested coset codes that are constructed via a random generator matrix G and a random coset sequence D^n to emulate the behavior of a random (nonlinear) code ensemble drawn from a specified pmf $p(x)$ on $\mathbb{F}_q$ [14].

Definition 2 (Random nested coset codes): Given a pmf $p(x)$ on $\mathbb{F}_q$ and $\epsilon > 0$, an $(n, nR, n\hat{R}, \mathbb{F}_q; p(x), \epsilon)$ random nested coset code ensemble consists of a message set $\mathbb{F}_q^{nR}$, a *random* generator matrix $G \in \mathbb{F}_q^{(nR+n\hat{R})\times n}$ and a *random* coset sequence D^n with entries i.i.d. $\text{Unif}(\mathbb{F}_q)$, an encoder that assigns a codeword to each message $m \in \mathbb{F}_q^{nR}$ according to the steps below, and a decoder that assigns an estimate to each received sequence y^n.

1) Given the realizations of G and D^n, compute $x^n(m, l)$ for each $m \in \mathbb{F}_q^{nR}$ and $l \in \mathbb{F}_q^{n\hat{R}}$ by (2).
2) For each message $m \in \mathbb{F}_q^{nR}$, choose an $l \in \mathbb{F}_q^{n\hat{R}}$ such that $x^n(m, l) \in \mathcal{T}_\epsilon^{(n)}(X)$. If there are more than one such l, choose one of them at random; if there is none, choose one in $\mathbb{F}_q^{n\hat{R}}$.[1]

Similar to the deterministic setting, we can also consider random coset codes and random linear codes.

Remark 2: An $(n, nR, \mathbb{F}_q)$ *random coset code* ensemble is a special case of an $(n, nR, n\hat{R}, \mathbb{F}_q; p(x), \epsilon)$ random nested coset code ensemble with $\hat{R} = 0$, $p(x) = \text{Unif}(\mathbb{F}_q)$ and $\epsilon = 0$. Specializing further, we can view an $(n, nR, \mathbb{F}_q)$ *random linear code* ensemble as an $(n, nR, \mathbb{F}_q)$ random coset code ensemble with $D^n = \mathbf{0}$.

As shown in [8], [14], random nested coset code ensembles can achieve the capacity of a discrete memoryless channel $p(y|x)$. When the input alphabet $\mathcal{X}$ is not isomorphic to a finite field, the channel can be transformed into a virtual channel $p(y|u)$ with equal capacity via an appropriately chosen auxiliary input U and symbol-by-symbol mapping $X = \varphi(U)$. This result can be extended to the Gaussian channel [8] (via a quantization argument) and to MACs [14].

We now consider nested coset codes with structural similarity.

Definition 3 (Homologous codes): An $(n, ((nR_j, n\hat{R}_j) : j \in [k]), \mathbb{F}_q)$ *homologous code* is a collection of

[1] This specific shaping function is referred to as the joint typicality encoding in [14]; see [24] for a similar technique in the context of lattice-based source coding.

$(n, nR_j, n\hat{R}_j, \mathbb{F}_q)$ nested coset codes, $j \in [k]$, and consists of k message sets $\mathbb{F}_q^{nR_j}$, a common generator matrix $G \in \mathbb{F}_q^{\kappa \times n}$ with $\kappa = \max_j(nR_j + n\hat{R}_j)$, k coset sequences d_j^n, k shaping functions $s_j : \mathbb{F}_q^{nR_j} \to \mathbb{F}_q^{n\hat{R}_j}$, k encoders, where encoder $j \in [k]$ assigns a codeword to each message according to the steps below, and a decoder that assigns an estimate to each received sequence y^n.

1) For each $m_j \in \mathbb{F}_q^{nR_j}$ and $l_j \in \mathbb{F}_q^{n\hat{R}_j}$, compute[2]

$$x_j^n(m_j, l_j) = [m_j \; l_j \; \mathbf{0}_{\kappa - n(R_j + \hat{R}_j)}]G \oplus d_j^n. \tag{3}$$

2) For each message $m_j \in \mathbb{F}_q^{nR_j}$, choose $x_j^n(m_j, s(m_j))$ as the assigned codeword, where $s_j(m_j)$ is the specified shaping function.

The term "homologous" was first proposed by the well-known biologist Owen [25] and later adopted by Darwin [26] to characterize the structures that have evolved from the same ancestor but differ in detail. In biological analogy, even though homologous codes are constructed from the same generator matrix, the actual "shape" of the codes can be quite different due to individual shaping functions.

We are particularly interested in the performance of a randomly generated homologous code ensemble, which is defined as follows.

Definition 4 (Random homologous codes): Given a pmf $p = \prod_{j=1}^k p(x_j)$ over $\mathbb{F}_q$ and $\epsilon > 0$, an $(n, ((nR_j, n\hat{R}_j) : j \in [k]), \mathbb{F}_q; p, \epsilon)$ *random homologous code* ensemble is a collection of $(n, nR_j, n\hat{R}_j, \mathbb{F}_q; p(x_j), \epsilon)$ random nested coset code ensembles, $j \in [k]$, and consists of k message sets $\mathbb{F}_q^{nR_j}$, a common *random* generator matrix $G \in \mathbb{F}_q^{\kappa \times n}$ with $\kappa = \max_j(nR_j + n\hat{R}_j)$ and k *random* coset sequences D_j^n with entries i.i.d. $\mathrm{Unif}(\mathbb{F}_q)$, k encoders, where encoder $j \in [k]$ assigns a codeword to each message according to the steps below, and a decoder that assigns an estimate to each received sequence y^n.

1) Given the realizations of G and D_j^n, compute $x_j^n(m_j, l_j)$ for each $m_j \in \mathbb{F}_q^{nR_j}$ and $l_j \in \mathbb{F}_q^{n\hat{R}_j}$ by (3).
2) For each message $m_j \in \mathbb{F}_q^{nR_j}$, choose an $l_j \in \mathbb{F}_q^{n\hat{R}_j}$ such that $x_j^n(m_j, l_j) \in \mathcal{T}_\epsilon^{(n)}(X_j)$. If there are more than one such l_j, choose one of them at random; if there is none, choose one in $\mathbb{F}_q^{n\hat{R}_j}$.

A rate tuple $(R_1, R_2, \ldots, R_k)$ is said to be *achievable by random homologous codes* in $\mathbb{F}_q$ for the multiple access channel $p(y|x_1, \ldots, x_k)$ if there exists a sequence of $(n, ((nR_j, n\hat{R}_j) : j \in [k]), \mathbb{F}_q; p, \epsilon)$ random homologous code ensemble such that $\lim_{n\to\infty} \mathsf{E}[P_e^{(n)}] = 0$ for some pmf $p(x^k)$ and for some $\epsilon > 0$, where the expectation is taken with respect to the randomness in the common generator matrix and individual coset sequences.

Note that for the k-sender DM-MAC $p(y|x_1, x_2, \ldots, x_k)$ and the input pmfs $p(x_1), p(x_2), \ldots, p(x_k)$, each sender can use a random nested coset code ensemble (with individual generator matrices $G_1, G_2, \ldots, G_k$) to achieve the region $\mathscr{R}_{\mathrm{MAC}}(X^k)$ characterized in (1). Thus, the corresponding *heterologous* nested coset codes can emulate the performance of typically nonlinear random code ensembles for MACs.[3] On the other hand, due to the use of a common generator matrix, homologous codes can achieve high rates when the goal of communication is to recover a linear combination of codewords. For a 2-sender DM-MAC, an achievable rate region is characterized in [8] for recovering linear combinations of codewords from random homologous code ensembles. When recovering both messages, however, this achievable rate region computed for a given input pmf is in general smaller than the region in (1). Even a tighter probability of error analysis discussed in [19] does not guarantee the achievability of the region in (1). This raises the question of whether random homologous codes are useful only for communicating the sum of the codewords (or equivalently, the sum of the messages) and fundamentally deficient compared to heterologous ones in communicating the messages themselves.

III. Motivating Examples

We present two toy examples that illustrate the performance of homologous codes and motivate our main result in Section IV.

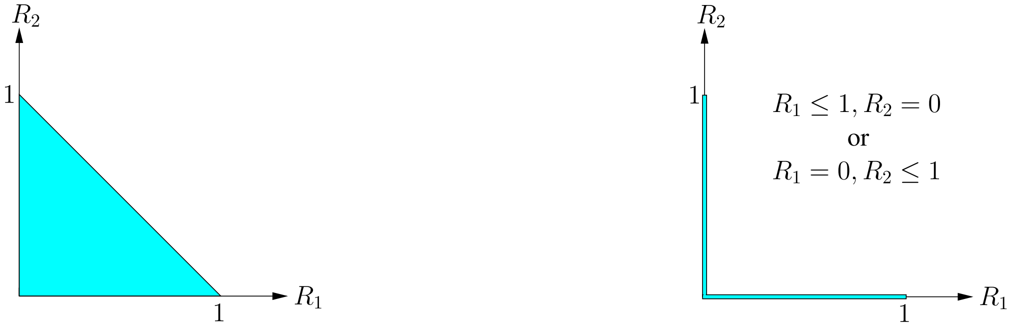
Example 1 (Binary adder MAC): Let $Y = X_1 \oplus X_2$, where $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{Y} = \{0, 1\}$ and the addition operation $\oplus$ is over $\mathbb{F}_2$. The capacity region of this channel is achieved by random coding with i.i.d. Bern(1/2) inputs X_1 and X_2, and is depicted in Fig. 1a. No binary linear or coset codes of the same generator matrix, however, can achieve this region. As a matter of fact, binary linear or coset codes of the same generator matrix can only achieve the rate region depicted in Fig. 1b. The achievability of $(R_1, R_2) = (1, 0)$ follows by using a pair of $(n, n, \mathbb{F}_2)$ and $(n, 0, \mathbb{F}_2)$ coset (or linear) codes with the generator matrix $G = I$, arbitrarily chosen coset sequences d_1^n and d_2^n, and the decoder that estimates $\hat{m}_1 = y^n \ominus d_1^n$. Exchanging the roles of encoder 1 and 2 implies the achievability of $(R_1, R_2) = (0, 1)$. For the converse, suppose without loss of generality that $R_1 \geq R_2 > 0$. Any message pair $(m_1, m_2) \in \mathbb{F}_2^{nR_1} \times \mathbb{F}_2^{nR_2}$ results in the same output as the message pair $(m_1 \oplus [m \; \mathbf{0}], m_2 \oplus m)$ for some $m \neq \mathbf{0} \in \mathbb{F}_2^{nR_2}$, which implies the converse.

By using *homologous* codes, however, the capacity region can be achieved. Suppose without loss of generality that $R_1 \geq R_2$ where $R_1 + R_2 \leq 1$. Consider the $(n, nR_1, 0, nR_2, n(1 - R_2), \mathbb{F}_2)$ homologous code constructed using the generator matrix $G = I$ and the coset sequences $d_1^n = d_2^n = \mathbf{0}$, where the shaping function for encoder 2 is specified as $s_2 : \mathbb{F}_2^{nR_2} \to \mathbb{F}_2^{n(1-R_2)}, s_2(m_2) = [\mathbf{0} \; m_2]$. It follows that the codeword pair assigned to $(m_1, m_2) \in \mathbb{F}_2^{nR_1 \times nR_2}$ is

$$\begin{aligned} x_1^n(m_1) &= [m_1 \;\; \mathbf{0}], \\ x_2^n(m_2, s_2(m_2)) &= [m_2 \; s_2(m_2)] = [m_2 \;\; \mathbf{0} \;\; m_2]. \end{aligned}$$

[2] Zero padding in (3) is because $nR_j + n\hat{R}_j$ may differ for different j.

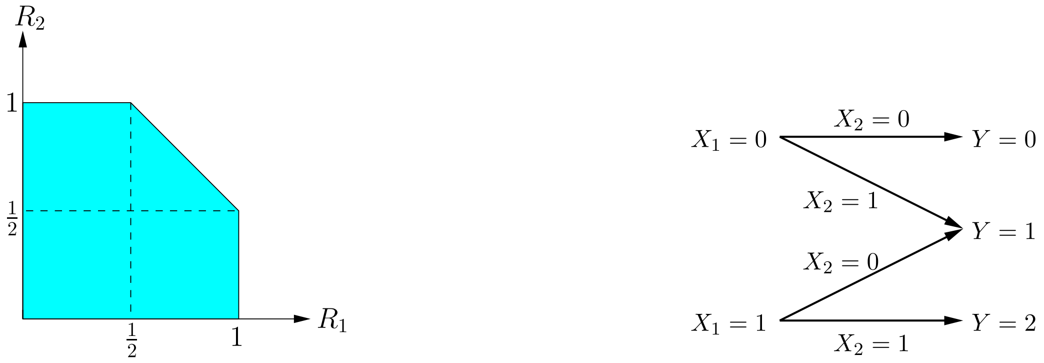
[3] Indeed, for $k = 2$, by controlling the structure of G_1 and G_2 more carefully, larger rates than random codes can be achieved for channels with state [14].



(a) The capacity region.

(b) An achievable rate region by coset codes.

Fig. 1. The binary adder MAC in Example 1.



(a) The capacity region.

(b) The channel from the perspective of sender 1.

Fig. 2. The binary erasure MAC in Example 2.

Given the channel output y^n, the decoding rule that declares the estimates $\hat{m}_1$ and $\hat{m}_2$ according to

$$\hat{m}_2 = y^n_{n-nR_2+1} \quad \text{and} \quad \hat{m}_1 = y_1^{nR_1} \ominus [\hat{m}_2\ \mathbf{0}]$$

can recover the messages m_1 and m_2 without any errors.

In Example 1, homologous codes benefit from the algebraic structure of the channel and emulate time division via the concatenation of two codes. The next example has an underlying channel structure that is not fully compatible with the algebraic structure of codes.

Example 2 (Binary erasure MAC): Let $Y = X_1 + X_2$, where $\mathcal{X}_1 = \mathcal{X}_2 = \{0, 1\}$, $\mathcal{Y} = \{0, 1, 2\}$, and the addition operation $+$ is over $\mathbb{R}$. The capacity region of the channel is achieved by random coding with i.i.d. Bern$(1/2)$ inputs X_1 and X_2, and is depicted in Fig. 2a. In contrast, no pair of binary coset codes with the same generator matrix can achieve the rate pair $(1/2 + \epsilon, 1/2 + \epsilon)$ for $\epsilon > 0$. The proof of this proposition is given in Appendix A.

This limitation of coset codes can be once again overcome by using homologous codes. We first present the achievability of the rate pair $(R, 1)$ for $R < 1/2$ with linear codes. Let $A_{nR\times n}$ be a full-rank binary generator matrix of a linear code that can reliably communicate $R < 1/2$ bits over the point-to-point DM binary erasure channel of erasure probability $1/2$.[4] Let

$$B = \begin{bmatrix} A \\ A^{\perp} \end{bmatrix},$$

where $A^{\perp}$ is an $(n - nR) \times n$ matrix whose rows are orthogonal to the rows of A. Consider now a pair of $(n, nR, \mathbb{F}_2)$ and $(n, n, \mathbb{F}_2)$ linear codes with generator matrices A and B respectively. Each message pair $(m_1, m_2) \in \mathbb{F}_2^{nR\times n}$ is assigned codewords $x_1^n(m_1) = [m_1\ \mathbf{0}_{n(1-R)}]\, B$ and $x_2^n(m_2) = m_2\, B$, respectively. Notice that since messages M_1 and M_2 are chosen independently, the codeword $x_1^n(M_1)$ is independent from the codeword $x_2^n(M_2)$. Moreover, since B is a full-rank square matrix and M_2 is chosen uniformly at random among $\mathbb{F}_2^n$, entries of $x_2^n(M_2)$ are i.i.d. Bern$(1/2)$. Therefore, the channel from the perspective of sender 1, $p(y^n|x_1^n(M_1))$, is equivalent to the point-to-point DM binary erasure channel with erasure probability $1/2$, which is illustrated in Fig. 2b. Upon receiving y^n, the decoder first declares the maximum likelihood estimate $\hat{m}_1$ by treating x_2^n as noise and then declares the estimate $\hat{m}_2$ by successive cancellation $x_2^n(\hat{m}_2) = y^n - x_1^n(\hat{m}_1)$. The reliable communication of M_1 and M_2 depends on the probability of error of the first decoding step, which vanishes asymptotically as $n \to \infty$ under the described matrix A.

We now construct homologous $(2n, n + nR, 0, \mathbb{F}_2)$ and $(2n, n + nR, n - nR, \mathbb{F}_2)$ codes with the generator matrix

$$G = \left[\begin{array}{c|c} B & O_{n\times n} \\ \hline O_{nR\times n} & \\ A^{\perp} & B \end{array}\right],$$

and shaping function $s_2 : \mathbb{F}_2^{n+nR} \to \mathbb{F}_2^{n-nR}$ such that $s_2(m_2) = (m_{2,i})_{i=nR+1}^{n}$ for $m_2 \in \mathbb{F}_2^{n+nR}$. If each message $m_1 \in \mathbb{F}_2^{n+nR}$ is divided into two sub-vectors

[4] The existence of such a linear code follows from [1, Section 3.1.3].

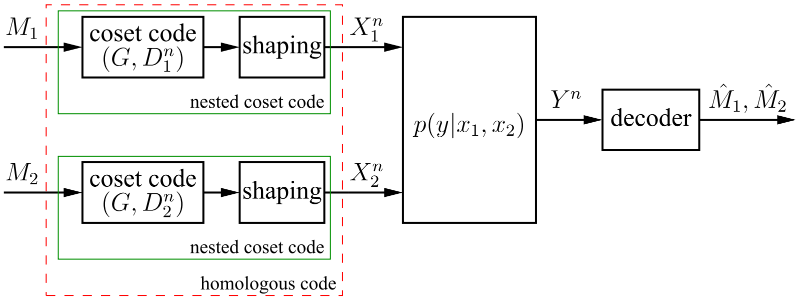


Fig. 3. Block diagram for shaping.

as $m_1 = [m_{11} \mid m_{12}]$, where $m_{11} \in \mathbb{F}_2^n$ and $m_{12} \in \mathbb{F}_2^{nR}$, and similarly each message $m_2 \in \mathbb{F}_2^{n+nR}$ is divided into three subvectors as $m_2 = [m_{21} \mid m_{22} \mid m_{23}]$, where $m_{21}, m_{23} \in \mathbb{F}_2^{nR}$ and $m_{22} \in \mathbb{F}_2^{n-nR}$, then the assigned codewords can be written as

$$x_1^{2n}(m_1) = \big[m_{11}B \;\;\big|\;\; m_{12}A\big],$$
$$x_2^{2n}(m_2, s_2(m_2)) = \big[m_{21}A \;\;\big|\;\; [m_{23}\; m_{22}]B\big].$$

Upon receiving the first half of the sequence y^{2n}, the decoder first declares the maximum likelihood estimate $\hat{m}_{21}$ by treating the first half of x_1^{2n} as noise and then declares the estimate $\hat{m}_{11}$ by successive cancellation. Similarly after receiving the second half of the sequence y^{2n}, it declares the maximum likelihood estimate $\hat{m}_{12}$ by treating the second half of x_2^{2n} as noise and then declares the estimates $\hat{m}_{22}$ and $\hat{m}_{23}$ by successive cancellation. By the construction of the matrix A, the first and second halves of codewords are reliably communicated at rates $(1, R)$ and $(R, 1)$, which, combined together, can be arbitrarily close to $(3/4, 3/4)$. The resulting transmission corresponds to time sharing via the concatenation of two codes. A similar argument can be extended to the entire capacity region.

The constructions of homologous codes for the binary adder and erasure MACs respectively emulate time division and time sharing in disguise via the concatenation of two codes. Consequently, these codes do not scale to more complicated problems (such as interference channels) in a satisfactory manner. As we will illustrate shortly, however, most (random) homologous codes are sufficient to achieve the capacity region, provided that they are constructed according to appropriate distributions.

IV. Achievable Rate Regions of Random Homologous Codes for Two Senders

We now investigate the performance of random homologous codes described in Definition 4 in Section II for the two sender DM-MAC $p(y|x_1, x_2)$. We take a gradual approach to presenting the main result and first discuss the key technical ingredients of the proof one by one.

A. Shaping

Symbols in an $(n, nR, \mathbb{F}_q)$ random coset code ensemble are uniformly distributed over $\mathbb{F}_q$. By the shaping step inherent in the nested coset codes, random homologous code ensembles emulate the statistical behavior of a random (nonlinear) code ensemble drawn from the desired distribution while maintaining a common algebraic structure across users. To separate the benefit from channel transformation, in this section, we are particularly interested in the finite-field input DM-MAC $p(y|x_1, x_2)$, where $\mathcal{X}_1 = \mathcal{X}_2 = \mathbb{F}_q$, and random homologous codes designed over $\mathbb{F}_q$ for this channel. The block diagram of this scheme is depicted in Fig. 3.

We describe the rate region achievable by random homologous codes. For given input pmfs $p(x_1)$ and $p(x_2)$, we refer to the rate region in (1) as $\mathscr{R}_{\mathrm{MAC}}(X_1, X_2)$, i.e., the set of rate pairs (R_1, R_2) such that

$$\begin{aligned} R_1 &< I(X_1; Y|X_2), \\ R_2 &< I(X_2; Y|X_1), \\ R_1 + R_2 &< I(X_1, X_2; Y), \end{aligned}$$

and define $\mathscr{R}_{\mathrm{L}}(X_1, X_2)$ as the set of rate pairs (R_1, R_2) such that

$$R_1 < \max\{I(X_1; Y),\ H(X_1) - H(X_2) + I(X_2; Y)\}, \tag{4}$$

or

$$R_2 < \max\{I(X_2; Y),\ H(X_2) - H(X_1) + I(X_1; Y)\}. \tag{5}$$

Proposition 1 (Shaping): A rate pair (R_1, R_2) is achievable by random homologous codes in $\mathbb{F}_q$ for the finite-field input DM-MAC $p(y|x_1, x_2)$ if

$$(R_1, R_2) \in \mathscr{R}_{\mathrm{MAC}}(X_1, X_2) \cap \mathscr{R}_{\mathrm{L}}(X_1, X_2)$$

for some input pmfs $p(x_1)$ and $p(x_2)$.

Proof: Our proof steps follow [8, Sec. VI] essentially line by line, except the analysis of one error event. Fix a pmf $p = p(x_1)p(x_2)$. Let $\epsilon' > 0$. We use an $(n, nR_1, n\hat{R}_1, nR_2, n\hat{R}_2, \mathbb{F}_q; p, \epsilon')$ random homologous code ensemble constructed in Definition 4. The decoder first fixes a sufficiently large $\epsilon > \epsilon'$ and then searches a unique pair of $(\hat{m}_1, \hat{m}_2)$ such that $(x_1^n(\hat{m}_1, l_1), x_2^n(\hat{m}_2, l_2), y^n) \in \mathcal{T}_\epsilon^{(n)}(X_1, X_2, Y)$ for some (l_1, l_2). If the decoder finds the unique pair, then it declares that $(\hat{m}_1, \hat{m}_2)$ was transmitted. Otherwise, it declares error. Assume that (M_1, M_2) is the transmitted message pair and (L_1, L_2) is the auxiliary index pair chosen by the shaping functions. We bound the probability of error averaged over codebooks. As in [8], the decoder makes an error only if one or more of the following events occur:

$$\begin{aligned}
\mathcal{E}_1 &= \{X_j^n(M_j, l_j) \notin \mathcal{T}_{\epsilon'}^{(n)}(X_j) \text{ for all } l_j,\ j = 1 \text{ or } 2\},\\
\mathcal{E}_2 &= \{(X_1^n(M_1, L_1), X_2^n(M_2, L_2), Y^n)\\
&\qquad \notin \mathcal{T}_{\epsilon}^{(n)}(X_1, X_2, Y)\},\\
\mathcal{E}_3 &= \{(X_1^n(M_1, L_1), X_2^n(m_2, l_2), Y^n) \in \mathcal{T}_{\epsilon}^{(n)}(X_1, X_2, Y)\\
&\qquad \text{for some } (m_2, l_2) \neq (M_2, L_2)\},\\
\mathcal{E}_4 &= \{(X_1^n(m_1, l_1), X_2^n(M_2, L_2), Y^n) \in \mathcal{T}_{\epsilon}^{(n)}(X_1, X_2, Y)\\
&\qquad \text{for some } (m_1, l_1) \neq (M_1, L_1)\},\\
\mathcal{E}_5 &= \{(X_1^n(m_1, l_1), X_2^n(m_2, l_2), Y^n) \in \mathcal{T}_{\epsilon}^{(n)}(X_1, X_2, Y)\\
&\quad \text{for some } (m_1, l_1) \neq (M_1, L_1) \text{ and } (m_2, l_2) \neq (M_2, L_2)\\
&\quad \text{such that } [m_1\ l_1\ \mathbf{0}] \ominus [M_1\ L_1\ \mathbf{0}] \text{ and}\\
&\quad [m_2\ l_2\ \mathbf{0}] \ominus [M_2\ L_2\ \mathbf{0}] \text{ are linearly independent}\},\\
\mathcal{E}_6 &= \{(X_1^n(m_1, l_1), X_2^n(m_2, l_2), Y^n) \in \mathcal{T}_{\epsilon}^{(n)}(X_1, X_2, Y)\\
&\quad \text{for some } (m_1, l_1) \neq (M_1, L_1) \text{ and } (m_2, l_2) \neq (M_2, L_2)\\
&\quad \text{such that } [m_1\ l_1\ \mathbf{0}] \ominus [M_1\ L_1\ \mathbf{0}] \text{ and}\\
&\quad [m_2\ l_2\ \mathbf{0}] \ominus [M_2\ L_2\ \mathbf{0}] \text{ are linearly dependent}\}.
\end{aligned}$$

Thus, by the union of events bound, $\mathsf{E}_{P_e^{(n)}} \leq \mathsf{P}(\mathcal{E}_1) + \sum_{k\neq 1} \mathsf{P}(\mathcal{E}_k \cap \mathcal{E}_1^c)$. By [8], the first five terms tend to 0 as $n \to \infty$ if

$$\begin{aligned}
\hat{R}_j &> \mathrm{D}_j^{\mathrm{KL}} + \delta(\epsilon'),\ j = 1, 2\\
R_1 + 2\hat{R}_1 + \hat{R}_2 &< I(X_1; Y|X_2) + 2\mathrm{D}_1^{\mathrm{KL}} + \mathrm{D}_2^{\mathrm{KL}} - \delta(\epsilon),\\
R_2 + \hat{R}_1 + 2\hat{R}_2 &< I(X_2; Y|X_1) + \mathrm{D}_1^{\mathrm{KL}} + 2\mathrm{D}_2^{\mathrm{KL}} - \delta(\epsilon),\\
R_1 + R_2 + 2\sum_{i=1}^{2} \hat{R}_i &< I(X_1, X_2; Y) + 2\sum_{i=1}^{2} \mathrm{D}_i^{\mathrm{KL}} - \delta(\epsilon),
\end{aligned}$$

where $\mathrm{D}_j^{\mathrm{KL}} := D(p_{X_j}||\mathrm{Unif}(\mathbb{F}_q))$ denotes the KL-divergence between the input pmf $p(x_j)$ and $\mathrm{Unif}(\mathbb{F}_q)$ for $j = 1, 2$. For the last term, one can use the analysis in [8] that is originally conducted for decoding of two linearly independent combinations of X_1 and X_2, namely, $W_1 = a_1 X_1 \oplus a_2 X_2$ and $W_2 = b_1 X_1 \oplus b_2 X_2$. Even for fixed W_1 and W_2, however, the resulting upper bound on R_1 and R_2 includes a max-min optimization over all linear combinations of W_1 and W_2, which is difficult to compute in general. Therefore, we present a new upper bound resulting in an achievable rate region that is easier to compute than the optimized rate region provided by [8]. Moreover, it can be shown that our achievable rate region is larger than the one in [8] for some channels, such as the on–off erasure MAC with $p = 1/2$ to be defined in Example 3.

Lemma 1: The probability $\mathsf{P}(\mathcal{E}_6 \cap \mathcal{E}_1^c)$ can be bounded by two different expressions:

$$\begin{aligned}
\mathsf{P}(\mathcal{E}_6 \cap \mathcal{E}_1^c) &\leq (q-1)q^{n(\hat{R}_1+\hat{R}_2+\min\{R_1+\hat{R}_1, R_2+\hat{R}_2\})}\\
&\qquad \cdot q^{n(H(X_1)+H(X_2)+H(X_2|Y)-3+\delta(\epsilon))},\\
\mathsf{P}(\mathcal{E}_6 \cap \mathcal{E}_1^c) &\leq (q-1)q^{n(\hat{R}_1+\hat{R}_2+\min\{R_1+\hat{R}_1, R_2+\hat{R}_2\})}\\
&\qquad q^{n(H(X_1)+H(X_2)+H(X_1|Y)-3+\delta(\epsilon))}.
\end{aligned}$$

Proof: Define the rate $R = \min\{R_1+\hat{R}_1,\ R_2+\hat{R}_2\}$, and the events $\mathcal{M} = \{M_1 = \mathbf{0}, M_2 = \mathbf{0}\}$ and $\mathcal{L} = \{L_1 = \mathbf{0}, L_2 = \mathbf{0}\}$. Define the set

$$\begin{aligned}
\mathcal{D} = \{&(m_1, l_1, m_2, l_2) \in \mathbb{F}_q^{nR_1} \times \mathbb{F}_q^{n\hat{R}_1} \times \mathbb{F}_q^{nR_2} \times \mathbb{F}_q^{n\hat{R}_2} :\\
&[m_1\ l_1\ \mathbf{0}] \neq \mathbf{0}, [m_2\ l_2\ \mathbf{0}] \neq \mathbf{0} \text{ are linearly dependent}\}.
\end{aligned}$$

By the symmetry of code generation, $\mathsf{P}(\mathcal{E}_6 \cap \mathcal{E}_1^c) = \mathsf{P}(\mathcal{E}_6 \cap \mathcal{E}_1^c|\mathcal{M}, \mathcal{L})$. To see this, note that

$$\begin{aligned}
&\mathsf{P}(\mathcal{E}_6 \cap \mathcal{E}_1^c)\\
&= \sum_{\substack{m_1,l_1\\ m_2,l_2}} \sum_{G} \sum_{d_1^n, d_2^n} \mathsf{P}\begin{pmatrix} \mathcal{E}_6 \cap \mathcal{E}_1^c, (M_1, M_2) = (m_1, m_2),\\ (L_1, L_2) = (l_1, l_2), G = G,\\ D_1^n = d_1^n, D_2^n = d_2^n \end{pmatrix}\\
&\overset{(a)}{=} \sum_{\substack{m_1,l_1\\ m_2,l_2}} \sum_{G} \sum_{d_1^n, d_2^n} \mathsf{P}\begin{pmatrix} \mathcal{E}_6 \cap \mathcal{E}_1^c, (M_1, M_2) = (\mathbf{0}, \mathbf{0}),\\ (L_1, L_2) = (\mathbf{0}, \mathbf{0}), G = G,\\ D_1^n = [m_1\ l_1\ \mathbf{0}]G \oplus d_1^n,\\ D_2^n = [m_2\ l_2\ \mathbf{0}]G \oplus d_2^n \end{pmatrix}\\
&= \sum_{\substack{m_1,l_1\\ m_2,l_2}} \mathsf{P}(\mathcal{E}_6 \cap \mathcal{E}_1^c, (M_1, L_1, M_2, L_2) = (\mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}))\\
&\overset{(b)}{=} \mathsf{P}(\mathcal{E}_6 \cap \mathcal{E}_1^c|(M_1, L_1, M_2, L_2) = (\mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0})),
\end{aligned}$$

where (a) follows since $(G, [m_1\ l_1\ \mathbf{0}]G \oplus D_1^n, [m_2\ l_2\ \mathbf{0}]G \oplus D_2^n) \overset{d}{=} (G, D_1^n, D_2^n)$ results in a permuted codebook and (b) follows by the fact proved in [8, Lemma 11] that (M_1, L_1, M_2, L_2) is uniformly distributed over its support.

By this observation, it suffices to bound the conditional probability as follows.

$$\begin{aligned}
&\mathsf{P}(\mathcal{E}_6 \cap \mathcal{E}_1^c|\mathcal{M}, \mathcal{L})\\
&= \mathsf{P}\big((X_1^n(m_1, l_1), X_2^n(m_2, l_2), Y^n) \in \mathcal{T}_{\epsilon}^{(n)}(X_1, X_2, Y)\\
&\qquad \text{for some } (m_1, l_1, m_2, l_2) \in \mathcal{D},\\
&\qquad X_j^n(\mathbf{0}, \mathbf{0}) \in \mathcal{T}_{\epsilon'}^{(n)}(X_j)\ j = 1, 2|\mathcal{M}, \mathcal{L}\big)\\
&\overset{(a)}{\leq} \sum_{\substack{(m_1,l_1,\\ m_2,l_2)\in\mathcal{D}}} \mathsf{P}\left(\begin{matrix} (X_1^n(m_1, l_1), X_2^n(m_2, l_2), Y^n)\\ \in \mathcal{T}_{\epsilon}^{(n)}(X_1, X_2, Y),\\ X_j^n(\mathbf{0}, \mathbf{0}) \in \mathcal{T}_{\epsilon'}^{(n)}(X_j)\ j = 1, 2 \end{matrix} \,\middle|\, \begin{matrix} \mathcal{M},\\ \mathcal{L} \end{matrix}\right)\\
&\leq \sum_{\substack{(m_1,l_1,\\ m_2,l_2)\in\mathcal{D}}} \mathsf{P}\left(\begin{matrix} X_2^n(m_2, l_2), Y^n)\\ \in \mathcal{T}_{\epsilon}^{(n)}(X_2, Y),\\ X_j^n(\mathbf{0}, \mathbf{0}) \in \mathcal{T}_{\epsilon'}^{(n)}(X_j)\ j = 1, 2 \end{matrix} \,\middle|\, \mathcal{M}, \mathcal{L}\right)\\
&\overset{(b)}{\leq} \sum_{\substack{(m_1,l_1,\\ m_2,l_2)\in\mathcal{D}}} \mathsf{P}\left(\begin{matrix} X_2^n(m_2, l_2), Y^n)\\ \in \mathcal{T}_{\epsilon}^{(n)}(X_2, Y),\\ X_j^n(\mathbf{0}, \mathbf{0}) \in \mathcal{T}_{\epsilon}^{(n)}(X_j)\ j = 1, 2 \end{matrix} \,\middle|\, \mathcal{M}, \mathcal{L}\right)\\
&= \sum_{(m_1,l_1,m_2,l_2)\in\mathcal{D}} \sum_{\substack{x_1^n\in\mathcal{T}_{\epsilon}^{(n)}(X_1),\\ x_2^n\in\mathcal{T}_{\epsilon}^{(n)}(X_2)}}\\
&\qquad \sum_{\substack{(\tilde{x}_2^n, y^n)\in\\ \mathcal{T}_{\epsilon}^{(n)}(X_2,Y)}} \mathsf{P}\left(\begin{matrix} [m_2\ l_2\ \mathbf{0}]G \oplus D_2^n = \tilde{x}_2^n,\\ D_1^n = x_1^n, D_2^n = x_2^n,\\ Y^n = y^n \end{matrix} \,\middle|\, \mathcal{M}, \mathcal{L}\right)\\
&\overset{(c)}{=} \sum_{(m_1,l_1,m_2,l_2)\in\mathcal{D}} \sum_{\substack{x_1^n\in\mathcal{T}_{\epsilon}^{(n)}(X_1),\\ x_2^n\in\mathcal{T}_{\epsilon}^{(n)}(X_2)}} \sum_{\substack{(\tilde{x}_2^n, y^n)\in\\ \mathcal{T}_{\epsilon}^{(n)}(X_2,Y)}}\\
&\qquad \mathsf{P}\left(\begin{matrix} [m_2\ l_2\ \mathbf{0}]G \oplus D_2^n = \tilde{x}_2^n,\\ D_1^n = x_1^n, D_2^n = x_2^n \end{matrix} \,\middle|\, \mathcal{M}, \mathcal{L}\right) p(y^n|x_1^n, x_2^n)
\end{aligned}$$

$$\begin{aligned}
&\overset{(d)}{\leq} q^{n(\hat{R}_1+\hat{R}_2)} \sum_{\substack{(m_1,l_1,m_2,l_2)\\ \in\mathcal{D}}} \sum_{\substack{x_1^n\in\mathcal{T}_\epsilon^{(n)}(X_1),\\ x_2^n\in\mathcal{T}_\epsilon^{(n)}(X_2)}} \sum_{\substack{y^n\in\\ \mathcal{T}_\epsilon^{(n)}(Y)}} p(y^n|x_1^n,x_2^n) \\
&\qquad \sum_{\substack{\tilde{x}_2^n\in\\ \mathcal{T}_\epsilon^{(n)}(X_2|y^n)}} \mathsf{P}\left(\begin{matrix}[m_2\ l_2\ \mathbf{0}]G\oplus D_2^n=\tilde{x}_2^n,\\ D_1^n=x_1^n,\ D_2^n=x_2^n\end{matrix}\right) \\
&= q^{n(\hat{R}_1+\hat{R}_2)} \sum_{(m_1,l_1,m_2,l_2)\in\mathcal{D}} \sum_{\substack{x_1^n\in\mathcal{T}_\epsilon^{(n)}(X_1),\\ x_2^n\in\mathcal{T}_\epsilon^{(n)}(X_2)}} \\
&\qquad \sum_{\substack{y^n\in\\ \mathcal{T}_\epsilon^{(n)}(Y)}} p(y^n|x_1^n,x_2^n) \sum_{\substack{\tilde{x}_2^n\in\\ \mathcal{T}_\epsilon^{(n)}(X_2|y^n)}} q^{-3n} \\
&\leq q^{n(\hat{R}_1+\hat{R}_2)}\,|\mathcal{D}|\,q^{n(H(X_1)+H(X_2)+H(X_2|Y)+\delta(\epsilon))}q^{-3n} \\
&\leq q^{n(\hat{R}_1+\hat{R}_2+R)}(q-1)q^{n(H(X_1)+H(X_2)+H(X_2|Y)+\delta(\epsilon))}q^{-3n},
\end{aligned}$$

where (a) follows by the union of events bound, (b) follows since $\epsilon > \epsilon'$, (c) follows since, conditioned on $(\mathcal{M}, \mathcal{L})$, the triple $G \to (D_1^n, D_2^n) \to Y^n$ form a Markov chain, and (d) follows by [8, Lemma 11]. By changing the order of X_1^n and X_2^n, we obtain the second bound on $\mathsf{P}(\mathcal{E}_6 \cap \mathcal{E}_1^c)$. ■

By Lemma 1 and using the relation $\mathrm{D}_j^{\mathrm{KL}} = 1 - H(X_j)$, we have $\mathsf{P}(\mathcal{E}_6 \cap \mathcal{E}_1^c) \to 0$ as $n \to \infty$ if $\min\{R_1 + 2\hat{R}_1 + \hat{R}_2, R_2 + \hat{R}_1 + 2\hat{R}_2\} < H(X_1) + 2\mathrm{D}_1^{\mathrm{KL}} + \mathrm{D}_2^{\mathrm{KL}} - \min\{H(X_1|Y), H(X_2|Y)\} - \delta(\epsilon)$. Choosing $\hat{R}_1 = \mathrm{D}_1^{\mathrm{KL}} + 2\delta(\epsilon')$, $\hat{R}_2 = \mathrm{D}_2^{\mathrm{KL}} + 2\delta(\epsilon')$ and letting $\epsilon \to 0$ yield that the rate pairs (R_1, R_2) is achievable if

$$\begin{aligned}
R_1 &< I(X_1;Y|X_2),\\
R_2 &< I(X_2;Y|X_1),\\
R_1 + R_2 &< I(X_1,X_2;Y),\\
\min\{R_1 + H(X_2), &R_2 + H(X_1)\}\\
&< H(X_1) + H(X_2) - \min\{H(X_1|Y), H(X_2|Y)\}.
\end{aligned} \tag{6}$$

The rate region defined by (6) is equivalent to the region $\mathscr{R}_{\mathrm{MAC}}(X_1, X_2) \cap \mathscr{R}_{\mathrm{L}}(X_1, X_2)$, as will be proved in Appendix B. Taking the union over input pmfs $p(x_1)$ and $p(x_2)$ completes the proof. ■

For the **binary adder MAC**, the achievable rate region in Proposition 1 is indeed equivalent to the capacity region, which is proved in Appendix C.

For the **binary erasure MAC**, however, the rate region in Proposition 1 is *strictly smaller* than the capacity region, as sketched in Fig. 4. In particular, the largest achievable symmetric rate is 2/3 (see Appendix D).

We now introduce another simple example, which will be used again in Section VI when we deal with multiple-receiver MACs.

Example 3 (On–off erasure MAC): Let $Y = (2X_1 - 1) + Z(2X_2 - 1)$, where $\mathcal{X}_1 = \mathcal{X}_2 = \{0, 1\}$, $\mathcal{Z} = \{0, 1\}$, and $\mathcal{Y} = \{0, \pm 1, \pm 2\}$, where the random variable $Z \sim \mathrm{Bern}(p)$ is independent from X_1 and X_2. If $Z = 1$, the channel is equivalent to the binary erasure MAC. If $Z = 0$, the output Y is only dependent on X_1. That is why this channel is called the *on–off erasure MAC*.

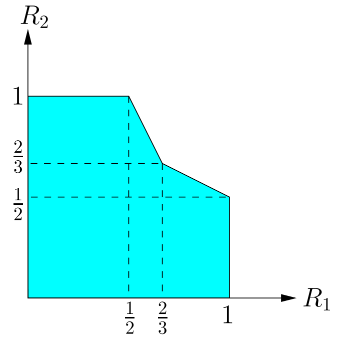


Fig. 4. The achievable rate region in Proposition 1 for the binary erasure MAC in Example 2.

For any $p \in (0, 1]$, the capacity region of the on–off erasure MAC is achieved by random coding with i.i.d. Bern(1/2) inputs X_1 and X_2, and is shown in Fig. 5a (in terms of p). If $p \leq 2/3$, the achievable rate region in Proposition 1 is equivalent to the capacity region. If $p > 2/3$, however, it reduces to the rate region depicted in Fig. 5b that is strictly smaller than the capacity region (see Appendix E). Note that for $p = 1$, the rate region in Fig. 5b is equivalent to the achievable rate region for the binary erasure MAC sketched in Fig. 4, since the on–off erasure MAC is equivalent to the binary erasure MAC when $p = 1$.

Remark 3: As shown by [19], the achievable rate region in Proposition 1 can be improved by stronger analysis tools, which we will discuss later in Section V-A and Proposition 4. For Examples 1–3, however, the achievable rate region in [19] reduces to that of Proposition 1.

B. Channel Transformation

Instead of choosing an appropriate shaping function within a nested coset code, there is a simpler way of achieving the performance of nonuniformly distributed codes. Following the basic idea in [17, Sec. 6.2], we can simply transform the channel $p(y|x_1, x_2)$ into a *virtual channel* with finite-field inputs

$$p(y|u_1, u_2) = p_{Y|X_1,X_2}(y|\varphi_1(u_1), \varphi_2(u_2)) \tag{7}$$

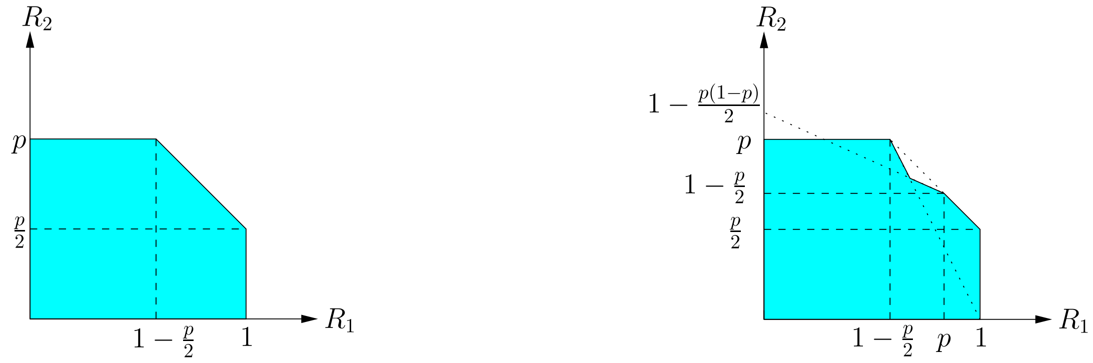
for some symbol-by-symbol mappings $\varphi_1 : \mathbb{F}_q \to \mathcal{X}_1$ and $\varphi_2 : \mathbb{F}_q \to \mathcal{X}_2$, as illustrated in Fig. 6. Note that this transformation can be applied to any DM-MAC $p(y|x_1, x_2)$ of arbitrary (not necessarily the same finite-field) input alphabets.

We now consider a pair of $(n, nR_1, \mathbb{F}_q)$ and $(n, nR_2, \mathbb{F}_q)$ random coset code ensembles with the same generator matrix for the virtual channel, which is equivalent to random homologous codes with $\hat{R}_1 = \hat{R}_2 = 0$. The block diagram of this scheme is depicted in Fig. 7. For a given pair of symbol-by-symbol mappings φ_1 and φ_2, we can establish the following whose proof is deferred to Appendix F.

Proposition 2: A rate pair (R_1, R_2) is achievable by random coset codes in $\mathbb{F}_q$ with the same generator matrix for the DM-MAC $p(y|x_1, x_2)$, if

$$(R_1, R_2) \in \mathscr{R}_{\mathrm{MAC}}(U_1, U_2) \cap \mathscr{R}_{\mathrm{L}}(U_1, U_2),$$

where $\mathscr{R}_{\mathrm{MAC}}(U_1, U_2)$ is defined as in (1) for the virtual channel $p(y|u_1, u_2)$ in (7) and for the inputs U_1 and U_2 drawn



(a) The capacity region for $p \in (0, 1]$.

(b) The achievable rate region in Proposition 1 for $p >$

Fig. 5. The on–off erasure MAC in Example 3.

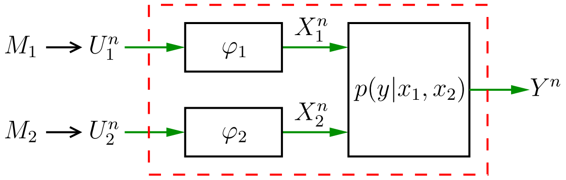


Fig. 6. The virtual DM-MAC $p(y|u_1, u_2)$ with virtual inputs U_1 and U_2.

independently according to Unif$(\mathbb{F}_q)$, and $\mathscr{R}_{\rm L}(U_1, U_2)$ is the set of (R_1, R_2) such that

$$\min(R_1, R_2) < \max\{I(U_1; Y),\ I(U_2; Y)\}. \tag{8}$$

Note that (8) is equivalent to (4) and (5) with (U_1, U_2) in place of (X_1, X_2) since U_1 and U_2 are uniform on $\mathbb{F}_q$.

Proposition 2 was stated for a fixed channel transformation specified by a given pair of symbol-by-symbol mappings $\varphi_1(u_1)$ and $\varphi_2(u_2)$ on a finite field $\mathbb{F}_q$. We now consider all such channel transformations, which results in a simpler achievable rate region.

Corollary 1 (Channel transformation): A rate pair (R_1, R_2) is achievable by random coset codes generated in some finite field with the same generator matrix for the DM-MAC $p(y|x_1, x_2)$, if

$$(R_1, R_2) \in \mathscr{R}_{\rm MAC}(X_1, X_2) \cap \mathscr{R}'_{\rm L}(X_1, X_2)$$

for some input pmfs $p(x_1)$ and $p(x_2)$, where $\mathscr{R}'_{\rm L}(X_1, X_2)$ is the set of (R_1, R_2) such that

$$\min(R_1, R_2) < \max\{I(X_1; Y),\ I(X_2; Y)\}.$$

Proof: First suppose that $p(x_1)$ and $p(x_2)$ are of the form

$$\frac{i}{\rho^m} \tag{9}$$

for some prime ρ and $i, m \in \mathbb{Z}^+$ for all x_1 and x_2. Then there exist $\varphi_1(u_1)$ and $\varphi_2(u_2)$ on $\mathbb{F}_q$ such that $X_j \overset{d}{=} \varphi_j(U_j)$ with $U_j \sim$ Unif$(\mathbb{F}_q)$, where $q = \rho^m$. Hence, we can transform the channel $p(y|x_1, x_2)$ into a virtual channel $p(y|u_1, u_2)$ and achieve the rate region in Proposition 2. Now, since $(U_1, U_2) \to (X_1, X_2) \to Y$ form a Markov chain and (U_1, X_1) and (U_2, X_2) are independent, the rate region $\mathscr{R}_{\rm MAC}(U_1, U_2) \cap \mathscr{R}_{\rm L}(U_1, U_2)$ in Proposition 2 can be simplified as $\mathscr{R}_{\rm MAC}(X_1, X_2) \cap \mathscr{R}'_{\rm L}(X_1, X_2)$. Finally, the earlier restrictions on the input pmfs can be removed since the set of pmfs of the form (9) is dense. This completes the proof. ■

We now revisit the previous examples to evaluate the achievable rate region in Corollary 1.

- **Binary adder MAC**: The achievable rate region in Corollary 1 is equivalent to the capacity region. To see this, note that for the binary adder MAC, $\mathscr{R}_{\rm L}(X_1, X_2) \subseteq \mathscr{R}'_{\rm L}(X_1, X_2)$ for any $p(x_1)$ and $p(x_2)$, and the former region achieved by the shaping (with the intersection with $\mathscr{R}_{\rm MAC}(X_1, X_2)$) reduces to the capacity region as proved in Appendix C. Therefore, the capacity region of the binary adder MAC is achievable by using coset codes over the virtual channel. This does not contradict the fact that no coset code on the *binary field* can achieve a positive symmetric rate pair, since channel transformation allows the use of linear (or coset) codes over larger finite fields.
- **Binary erasure MAC**: The achievable rate region in Corollary 1 reduces to the one in Proposition 1 sketched in Fig. 4, although $\mathscr{R}'_{\rm L}(X_1, X_2)$ is in general different than $\mathscr{R}_{\rm L}(X_1, X_2)$ for fixed pmfs $p(x_1)$ and $p(x_2)$. The proof is given in Appendix D.
- **On–off erasure MAC**: If $p \leq 2/3$, the achievable rate region in Corollary 1 reduces to the capacity region sketched in Fig. 5a. If $p > 2/3$, however, it reduces to the rate region sketched in Fig. 8. While larger than what is achieved by the shaping (cf. Fig. 5b), the achievable rate region by channel transformation in Corollary 1 is still strictly smaller than the capacity region. The details are given in Appendix E.

Remark 4: The achievable rate region for the channel transformation technique in Corollary 1 can be easily evaluated for fixed input pmfs $p(x_1)$ and $p(x_2)$. Using the analysis tools developed in [19], Proposition 2 and Corollary 1 can be potentially strengthened. Given a virtual channel $p(y|u_1, u_2)$ with input pmfs $p(u_1)$ and $p(u_2)$ on some $\mathbb{F}_q$, the resulting achievable rate region would depend on the distribution of $(a_1 U_1 \oplus a_2 U_2, Y)$ for every $a_1, a_2 \neq 0 \in \mathbb{F}_q$. The union of these rate regions over all channel transformations, however, is not computable. Therefore, it is unclear whether the insufficiency of the channel transformation technique for

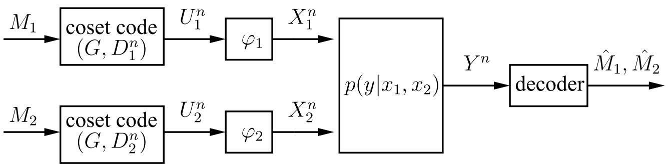


Fig. 7. Block diagram for channel transformation.

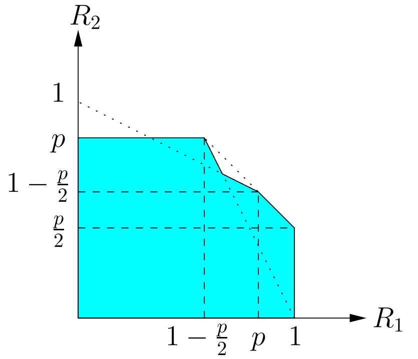


Fig. 8. The achievable rate region in Corollary 1 for the on–off erasure MAC in Example 3 when $p > 2/3$.

Examples 2–3 (binary erasure MAC and on–off erasure MAC) is fundamental or due to the deficiency of our analysis tools.

C. Combination

As shown for the binary erasure MAC and on–off erasure MAC examples, shaping (with homologous codes) and channel transformation (with coset codes of the same generator matrix) seemingly cannot achieve the capacity region. When combined together, these techniques can achieve the pentagonal region $\mathscr{R}_{\text{MAC}}(X_1, X_2)$ for any $p(x_1)$ and $p(x_2)$ while maintaining the algebraic structure of the code. Consider the virtual channel in (7) and random homologous codes for this channel, a block diagram for which is depicted in Fig. 9. Then, Proposition 1 implies the following.

Proposition 3: A rate pair (R_1, R_2) is achievable by random homologous codes in $\mathbb{F}_q$ for the DM-MAC $p(y|x_1, x_2)$, if

$$(R_1, R_2) \in \mathscr{R}_{\text{MAC}}(X_1, X_2) \cap \mathscr{R}_{\text{L}}(U_1, U_2)$$

for some pmfs $p(u_1)$ and $p(u_2)$ on $\mathbb{F}_q$, and some mappings $x_1 = \varphi_1(u_1)$ and $x_2 = \varphi_2(u_2)$, where $\mathscr{R}_{\text{L}}(U_1, U_2)$ is the set of rate pairs (R_1, R_2) satisfying (4) or (5).

Proof: Given pmfs $p(u_1)$ and $p(u_2)$ on $\mathbb{F}_q$, and mappings $x_1 = \varphi_1(u_1)$ and $x_2 = \varphi_2(u_2)$, by Proposition 1, the rate region $\mathscr{R}_{\text{MAC}}(U_1, U_2) \cap \mathscr{R}_{\text{L}}(U_1, U_2)$ is achievable by random homologous codes in $\mathbb{F}_q$ for the virtual channel $p(y|u_1, u_2)$. Now, since $(U_1, U_2) \to (X_1, X_2) \to Y$ form a Markov chain and (U_1, X_1) and (U_2, X_2) are independent, the rate region $\mathscr{R}_{\text{MAC}}(U_1, U_2) \cap \mathscr{R}_{\text{L}}(U_1, U_2)$ simplifies to $\mathscr{R}_{\text{MAC}}(X_1, X_2) \cap \mathscr{R}_{\text{L}}(U_1, U_2)$. The proof follows by taking the union over pmfs $p(u_1)$ and $p(u_2)$ on $\mathbb{F}_q$, and mappings $x_1 = \varphi_1(u_1)$ and $x_2 = \varphi_2(u_2)$. ■

We are now ready to state one of the main technical results of this paper, which follows from Proposition 3 by optimizing over all channel transformations.

Theorem 1 (Combination): A rate pair (R_1, R_2) is achievable by random homologous codes in some finite field for the DM-MAC $p(y|x_1, x_2)$, if $(R_1, R_2) \in \mathscr{R}_{\text{MAC}}(X_1, X_2)$ for some $p(x_1)$ and $p(x_2)$.

Proof: Our argument is similar to the proof of Corollary 1, except that the choice of channel transformation needs more care. First suppose that $p(x_1)$ and $p(x_2)$ are of the form (9) for some prime ρ. We will show that there exist a finite field $\mathbb{F}_q$, pmfs $p(u_1)$ and $p(u_2)$ on $\mathbb{F}_q$, and mappings $x_1 = \varphi_1(u_1)$ and $x_2 = \varphi_2(u_2)$ such that $\mathscr{R}_{\text{MAC}}(X_1, X_2) \subseteq \mathscr{R}_{\text{L}}(U_1, U_2)$. Consider random homologous codes over $\mathbb{F}_q$ with $q = \rho^{2m}$. Choose U_1 and φ_1 such that U_1 and $\varphi_1(U_1) \overset{d}{=} X_1$ are one-to-one on the support of U_1 (this is always possible since $q \geq \rho^m$). Also choose $U_2 \sim \text{Unif}(\mathbb{F}_q)$ and φ_2 such that $\varphi_2(U_2) \overset{d}{=} X_2$ (this is possible due to the form of $p(x_2)$). Let $(R_1, R_2) \in \mathscr{R}_{\text{MAC}}(X_1, X_2)$. Then, (R_1, R_2) satisfies

$$\begin{aligned} R_2 &< I(X_2; Y|X_1) \\ &\leq H(X_2) \\ &\leq \log \rho^m \\ &\leq H(U_2) - H(U_1) \\ &\leq H(U_2) - H(U_1) + I(U_1; Y), \end{aligned}$$

which implies that $(R_1, R_2) \in \mathscr{R}_{\text{L}}(U_1, U_2)$. Finally, the restrictions on the input pmfs can be removed again by the denseness argument. ■

Remark 5: Theorem 1 can be strengthened by putting a cardinality bound on the underlying finite field. We need a different construction. By Bertrand's postulate, there exists a prime q such that $|\mathcal{X}_1||\mathcal{X}_2| < q < 2|\mathcal{X}_1||\mathcal{X}_2|$. For a given input pmf $p(x_1)$ and $p(x_2)$, consider a random homologous code ensemble over $\mathbb{F}_q$. Choose U_1 and φ_1 such that U_1 and $\varphi_1(U_1) \overset{d}{=} X_1$ are one-to-one on the support of U_1, which is always possible since $q \geq |\mathcal{X}_1|$. Also choose U_2 and φ_2 such that U_2 and (X_1, X_2) are one-to-one on the support of U_2 and that $\varphi_2(U_2) \overset{d}{=} X_2$, which is always possible since $q \geq |\mathcal{X}_1||\mathcal{X}_2|$. The claim is that $\mathscr{R}_{\text{MAC}}(X_1, X_2) \subseteq \mathscr{R}_{\text{L}}(U_1, U_2)$. To see this, let $(R_1, R_2) \in \mathscr{R}_{\text{MAC}}(X_1, X_2)$. Then, (R_1, R_2) satisfies

$$\begin{aligned} R_2 &< I(X_2; Y|X_1) \\ &\leq H(X_2) \\ &= H(X_1, X_2) - H(X_1) \\ &= H(U_2) - H(U_1) \\ &\leq H(U_2) - H(U_1) + I(U_1; Y), \end{aligned}$$

which implies that $(R_1, R_2) \in \mathscr{R}_{\text{L}}(U_1, U_2)$. Therefore, for any pmfs $p(x_1)$ and $p(x_2)$, the rate region $\mathscr{R}_{\text{MAC}}(X_1, X_2)$ is achievable by random homologous codes in some finite field $\mathbb{F}_q$ such that $q \leq 2|\mathcal{X}_1||\mathcal{X}_2|$ for the DM-MAC $p(y|x_1, x_2)$.

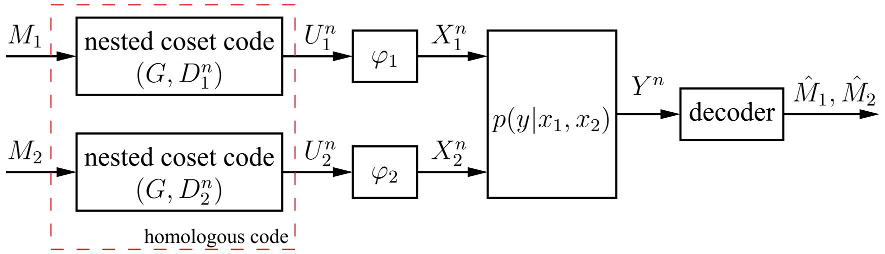


Fig. 9. Block diagram for homologous codes over the virtual channel.

V. Extension to More Than Two Senders

The achievable rate region by random homologous codes for the 2-sender DM-MAC can be extended to DM-MACs with more senders. In this section, we present the performance of random homologous code ensembles for the k-sender DM-MAC $p(y|x_1, x_2, \ldots, x_k)$. Similar to Section IV, we first discuss the performance of random homologous codes under the fixed channel alphabets, following the recent work in [19]. We then generalize the result by incorporating channel transformation.

A. Shaping

The achievable rate region for the finite-field input DM-MAC $p(y|x_1, x_2, \ldots, x_k)$, $\mathcal{X}_1 = \mathcal{X}_2 = \cdots = \mathcal{X}_k = \mathbb{F}_q$, by random homologous code ensembles was studied in [19]. For the sake of completeness, we review the main result in [19] on which we build the achievability of the capacity region for the k-sender DM-MAC. Let $\mathcal{A}$ denote the set of rank deficient $k \times k$ matrices over $\mathbb{F}_q$. For a given matrix $A \in \mathcal{A}$, we define the collection

$$\begin{aligned}\mathscr{D}(A) = \{\mathcal{J} \subseteq [k] : |\mathcal{J}| = k - \operatorname{rank}(A), \\ \operatorname{rank}[A^T \; e(\mathcal{J})^T]^T = k\},\end{aligned}$$

where $e(\mathcal{J}) \in \mathbb{F}_q^{|\mathcal{J}| \times k}$ denotes the matrix whose rows are the standard basis vectors e_j for $j \in \mathcal{J}$. For a given set $\mathcal{J} \in \mathscr{D}(A)$ and input pmfs $p(x_1), p(x_2), \ldots, p(x_k)$, we define the rate region $\mathscr{R}(A, \mathcal{J}, X^k)$ as the set of rate tuples $(R_1, R_2, \ldots, R_k)$ such that

$$\sum_{j \in \mathcal{J}} R_j < I(X(\mathcal{J}); Y, W_A),$$

where

$$W_A = A\,[X_1 \; X_2 \; \ldots \; X_k]^T.$$

We are now ready to state the main result of [19].

Proposition 4 ([19, Theorem 1]): A rate tuple $(R_1, R_2, \ldots, R_k)$ is achievable by random homologous codes in $\mathbb{F}_q$ for the finite-field input DM-MAC $p(y|x_1, x_2, \ldots, x_k)$, if

$$(R_1, R_2, \ldots, R_k) \in \bigcap_{A \in \mathcal{A}} \bigcup_{\mathcal{J} \in \mathscr{D}(A)} \mathscr{R}(A, \mathcal{J}, X^k)$$

for some input pmfs $p(x_1), p(x_2), \ldots, p(x_k)$.

Remark 6 (Revisit of the 2-sender DM-MAC): Consider the 2-sender DM-MAC $p(y|x_1, x_2)$ with given input pmfs $p(x_1)$ and $p(x_2)$. To compute the achievable rate region in Proposition 4, it suffices to consider the set of rank deficient 2×2 matrices with different spans. There are four types of such matrices:

- $A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$:
  $\mathscr{D}(A) = \{\{1, 2\}\}$ and $\cup_{\mathcal{J} \in \mathscr{D}(A)} \mathscr{R}(A, \mathcal{J}, X_1, X_2)$ reduces to the set of rate pairs satisfying
  $$R_1 + R_2 < I(X_1, X_2; Y),$$
- $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$:
  $\mathscr{D}(A) = \{\{1\}\}$ and $\cup_{\mathcal{J} \in \mathscr{D}(A)} \mathscr{R}(A, \mathcal{J}, X_1, X_2)$ is the set of rate pairs satisfying
  $$R_1 < I(X_1; Y|X_2),$$
- $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$:
  $\mathscr{D}(A) = \{\{2\}\}$ and $\cup_{\mathcal{J} \in \mathscr{D}(A)} \mathscr{R}(A, \mathcal{J}, X_1, X_2)$ is the set of rate pairs satisfying
  $$R_2 < I(X_2; Y|X_1),$$
- $A = \begin{bmatrix} 1 & a \\ 0 & 0 \end{bmatrix}$ for some nonzero $a \in \mathbb{F}_q$:
  $\mathscr{D}(A) = \{\{1\}, \{2\}\}$ and $\cup_{\mathcal{J} \in \mathscr{D}(A)} \mathscr{R}(A, \mathcal{J}, X_1, X_2)$ is the set of rate pairs satisfying
  $$R_1 < I(X_1; Y, W_a),$$
  or
  $$R_2 < I(X_2; Y, W_a),$$
  where $W_a = X_1 \oplus a X_2$.

The achievable rate region in Proposition 4 is then equivalent to $\mathscr{R}_{\mathrm{MAC}}(X_1, X_2) \cap \tilde{\mathscr{R}}_{\mathrm{L}}(X_1, X_2)$ where $\tilde{\mathscr{R}}_{\mathrm{L}}(X_1, X_2)$ is the set of rate pairs (R_1, R_2) such that for every nonzero $a \in \mathbb{F}_q$

$$R_1 < I(X_1; Y, X_1 \oplus a X_2) \tag{10}$$

or

$$R_2 < I(X_2; Y, X_1 \oplus a X_2). \tag{11}$$

One may notice that for every nonzero a over $\mathbb{F}_q$

$$\begin{aligned} H(X_1|Y, X_1 \oplus a X_2) &= H(X_2|Y, X_1 \oplus a X_2) \\ &\le \min\{H(X_1|Y), H(X_2|Y)\}, \end{aligned}$$

which implies that $\tilde{\mathscr{R}}_{\mathrm{L}}(X_1, X_2)$ is in general larger than $\mathscr{R}_{\mathrm{L}}(X_1, X_2)$ defined in Proposition 1 in Section IV-A. Indeed, the error analysis in the proof of Proposition 1 can be modified to account for the larger $\tilde{\mathscr{R}}_{\mathrm{L}}(X_1, X_2)$ region.

Remark 7: The achievable rate region in Proposition 4 is the largest region thus far established with homologous codes in the literature. As a matter of fact, there is some indication that this region is optimal in the sense that it cannot be improved by using maximum likelihood decoding [9], [10]. Still, it is in general strictly smaller than the capacity region of the k-sender DM-MAC. In particular, for the channels defined in Examples 1–3, the achievable rate region in Propositon 4 reduces to the achievable rate region in Proposition 1 described in Section IV-A. To see this, fix input pmfs $p(x_1)$ and $p(x_2)$. The set of rate pairs satisfying (10) or (11) for $a = 1$ is equivalent to the rate region $\mathscr{R}_{\mathrm{L}}(X_1, X_2)$.

As a corollary of Proposition 4, we can come up with a smaller rate region achievable by random homologous codes that is easier to compute. As we will discuss in the next section, however, this smaller achievable rate region combined with channel transformation gives rise to the achievability of the capacity region. Let $\mathcal{B}$ denote the set of rank deficient $k \times k$ matrices over $\mathbb{F}_q$ that is not row equivalent[5] to a diagonal matrix. Note that $\mathcal{B} \subset \mathcal{A}$. Given a matrix $A \in \mathcal{B}$, a set $\mathcal{J} \in \mathscr{D}(A)$, and input pmfs $p(x_1), p(x_2), \ldots, p(x_k)$, we define the rate region $\tilde{\mathscr{R}}(A, \mathcal{J}, X^k)$ as the set of rate tuples $(R_1, R_2, \ldots, R_k)$ satisfying

$$\sum_{j \in \mathcal{J}} R_j < H(X(\mathcal{J})) - \min_{\mathcal{S} \in \mathscr{D}(A)} H(X(\mathcal{S})|Y).$$

Given input pmfs $p(x_1), p(x_2), \ldots, p(x_k)$, we define the rate region

$$\mathscr{R}_{\mathrm{L}}(X^k) = \bigcap_{A \in \mathcal{B}} \bigcup_{\mathcal{J} \in \mathscr{D}(A)} \tilde{\mathscr{R}}(A, \mathcal{J}, X^k). \tag{12}$$

Corollary 2 (Extension of Proposition 1): A rate tuple $(R_1, R_2, \ldots, R_k)$ is achievable by random homologous codes in $\mathbb{F}_q$ for the finite-field input DM-MAC $p(y|x_1, x_2, \ldots, x_k)$, if

$$(R_1, R_2, \ldots, R_k) \in \mathscr{R}_{\mathrm{MAC}}(X^k) \cap \mathscr{R}_{\mathrm{L}}(X^k)$$

for some input pmfs $p(x_1), p(x_2), \ldots, p(x_k)$.
We first revisit the 2-sender case with Corollary 2 and then provide a proof for Corollary 2.

Remark 8 (Revisit of the 2-sender DM-MAC with Corollary 2): For the case $k = 2$, the achievable rate region in Corollary 2 reduces to the achievable rate region in Proposition 1. To see this, fix input pmfs $p(x_1)$ and $p(x_2)$. A rank-deficient 2×2 matrix over $\mathbb{F}_q$ that is not row equivalent to a diagonal matrix must be row equivalent to a matrix of the form

$$\begin{bmatrix} a_1 & a_2 \\ 0 & 0 \end{bmatrix}$$

for some nonzero a_1 and a_2 over $\mathbb{F}_q$. Then, for every such matrix A, $\mathscr{D}(A) = \{\{1\}, \{2\}\}$. Therefore, the rate region $\mathscr{R}_{\mathrm{L}}(X_1, X_2)$ defined in (12) is the set of rate pairs (R_1, R_2) such that

$$R_1 < H(X_1) - \min\{H(X_1|Y), H(X_2|Y)\}$$

or

$$R_2 < H(X_2) - \min\{H(X_1|Y), H(X_2|Y)\},$$

which is equivalent to the rate region $\mathscr{R}_{\mathrm{L}}(X_1, X_2)$ defined in Section IV-A.

Proof of Corollary 2: We will show that given input pmfs $p(x_1), p(x_2), \ldots, p(x_k)$

$$(\mathscr{R}_{\mathrm{MAC}}(X^k) \cap \mathscr{R}_{\mathrm{L}}(X^k)) \subseteq \bigcap_{A \in \mathcal{A}} \bigcup_{\mathcal{J} \in \mathscr{D}(A)} \mathscr{R}(A, \mathcal{J}, X^k),$$

by first showing that

$$\mathscr{R}_{\mathrm{MAC}}(X^k) = \bigcap_{A \in \mathcal{A} \setminus \mathcal{B}} \bigcup_{\mathcal{J} \in \mathscr{D}(A)} \mathscr{R}(A, \mathcal{J}, X^k),$$

and then showing that

$$\mathscr{R}_{\mathrm{L}}(X^k) \subseteq \bigcap_{A \in \mathcal{B}} \bigcup_{\mathcal{J} \in \mathscr{D}(A)} \mathscr{R}(A, \mathcal{J}, X^k).$$

To prove the first claim, let A be a rank-deficient $k \times k$ matrix that is row equivalent to a diagonal matrix D (i.e., $A \in \mathcal{A} \setminus \mathcal{B}$), and let $\mathcal{J}$ be the set of indices such that $j \in \mathcal{J}$ if $D_{jj} = 0$. Then, by Lemma 3 in Appendix G, $\mathscr{D}(A) = \mathcal{J}$ and $\mathscr{R}(A, \mathcal{J}, X^k)$ is reduced to the set of rate tuples $(R_1, R_2, \ldots, R_k)$ satisfying

$$\sum_{j \in \mathcal{J}} R_j < I(X(\mathcal{J}); Y, X(\mathcal{J}^c)).$$

Taking the intersection over all $A \in \mathcal{A} \setminus \mathcal{B}$ proves the first claim. For the second claim, it suffices to show that given a matrix $A \in \mathcal{B}$ and a set $\mathcal{J} \in \mathscr{D}(A)$

$$\tilde{\mathscr{R}}(A, \mathcal{J}, X^k) \subseteq \mathscr{R}(A, \mathcal{J}, X^k).$$

Now, a rate tuple $(R_1, R_2, \ldots, R_k) \in \tilde{\mathscr{R}}(A, \mathcal{J}, X^k)$ satisfies

$$\begin{aligned} \sum_{j \in \mathcal{J}} R_j &< H(X(\mathcal{J})) - \min_{\mathcal{S} \in \mathscr{D}(A)} H(X(\mathcal{S})|Y) \\ &\leq H(X(\mathcal{J})) - \min_{\mathcal{S} \in \mathscr{D}(A)} H(X(\mathcal{S})|Y, W_A) \\ &\overset{(a)}{=} H(X(\mathcal{J})) - H(X(\mathcal{J})|Y, W_A), \\ &= I(X(\mathcal{J}); Y, W_A), \end{aligned}$$

where (a) follows since $H(X(\mathcal{J})|Y, W_A) = H(X^k|Y, W_A)$ is constant for every $\mathcal{J} \in \mathscr{D}(A)$. Therefore, we have $(R_1, R_2, \ldots, R_k) \in \mathscr{R}(A, \mathcal{J}, X^k)$, which completes the proof. ■

B. Combination

We incorporate channel transformation into random homologous codes to prove the achievability of the capacity region of the k-sender DM-MAC. Similar to the idea discussed in Section IV-B, we can simply transform the channel $p(y|x_1, x_2, \ldots, x_k)$ into a *virtual channel* with finite-field inputs

$$\begin{aligned} &p(y|u_1, u_2, \ldots, u_k) \\ &\quad = p_{Y|X_1, X_2, \ldots, X_k}(y|\varphi_1(u_1), \varphi_2(u_2), \ldots, \varphi_k(u_k)) \end{aligned} \tag{13}$$

for some symbol-by-symbol mappings $\varphi_j : \mathbb{F}_q \to \mathcal{X}_j$, $j \in [k]$.

[5] Two matrices are row equivalent if one can be obtained from the other by a sequence of elementary row operations.

Now, consider the virtual channel in (13) and random homologous codes for this channel. Then, Corollary 2 implies the following.

Proposition 5: A rate tuple $(R_1, R_2, \ldots, R_k)$ is achievable by random homologous codes in $\mathbb{F}_q$ for the DM-MAC $p(y|x_1, x_2, \ldots, x_k)$, if

$$(R_1, R_2, \ldots, R_k) \in \mathscr{R}_{\text{MAC}}(X^k) \cap \mathscr{R}_{\text{L}}(U^k)$$

for some $p(u_1), p(u_2), \ldots, p(u_k)$ on $\mathbb{F}_q$ and some mappings $x_1 = \varphi_1(u_1), x_2 = \varphi_2(u_2), \ldots, x_k = \varphi_k(u_k)$, where $\mathscr{R}_{\text{L}}(U^k)$ is the set of rate tuples $(R_1, R_2, \ldots, R_k)$ satisfying (12) for the virtual channel $p(y|u_1, u_2, \ldots, u_k)$.

We are now ready to extend Theorem 1 to the k-sender case, which follows from Proposition 5 by optimizing over all channel transformations.

Theorem 2 (Combination): A rate tuple $(R_1, R_2, \ldots, R_k)$ is achievable by random homologous codes in some finite field for the DM-MAC $p(y|x_1, x_2, \ldots, x_k)$, if

$$(R_1, R_2, \ldots, R_k) \in \mathscr{R}_{\text{MAC}}(X^k)$$

for some $p(x_1), p(x_2), \ldots, p(x_k)$.

Proof: We follow similar arguments to the proof of Theorem 1 and show that given input pmfs $p(x_1), p(x_2), \ldots, p(x_k)$, there exists a finite field $\mathbb{F}_q$, pmfs $p(u_1), p(u_2), \ldots, p(u_k)$ on $\mathbb{F}_q$, and mappings $x_1 = \varphi_1(u_1), x_2 = \varphi_2(u_2), \ldots, x_k = \varphi_k(u_k)$ such that

$$\mathscr{R}_{\text{MAC}}(X^k) \subseteq \mathscr{R}_{\text{L}}(U^k). \tag{14}$$

First, suppose that $p(x_j)$, $j \in [k]$, are of the form i/ρ^m for some $i, m \in \mathbb{Z}^+$ and prime ρ. We consider random homologous codes over $\mathbb{F}_q$ with $q = \rho^{k^k m}$. Let $q_j = \rho^{k^{(k-j+1)} m}$ for $j \in [k]$ and note that

$$\mathbb{F}_{q_k} \subset \mathbb{F}_{q_{k-1}} \subset \cdots \subset \mathbb{F}_{q_1} = \mathbb{F}_q.$$

Consider $U_j \sim \text{Unif}(\mathbb{F}_{q_j})$, and φ_j such that $\varphi_j(U_j) \overset{d}{=} X_j$ for $j \in [k]$ (this is possible due to the form of $p(x_j)$ and by the choice of q_j). To see (14), it suffices to show that for every matrix $A \in \mathcal{B}$, $\mathscr{R}_{\text{MAC}}(X^k) \subseteq \cup_{\mathcal{J} \in \mathscr{D}(A)} \tilde{\mathscr{R}}(A, \mathcal{J}, U^k)$. Consider a rate tuple $(R_1, R_2, \ldots, R_k) \in \mathscr{R}_{\text{MAC}}(X^k)$ and a matrix $A \in \mathcal{B}$. By Lemma 3 (see Appendix G) and by the choice of $p(u_j)$, there exist at least two different sets $\mathcal{J}_1, \mathcal{J}_2 \in \mathscr{D}(A)$ such that

$$H(U(\mathcal{J}_1)) - H(U(\mathcal{J}_2)) \geq k \log \rho^m \geq H(X^k).$$

Then, $(R_1, R_2, \ldots, R_k)$ satisfies

$$\begin{aligned} \sum_{j \in \mathcal{J}_1} R_j &< H(X^k) \\ &\leq H(U(\mathcal{J}_1)) - H(U(\mathcal{J}_2)) \\ &\leq H(U(\mathcal{J}_1)) - \min_{\mathcal{S} \in \mathscr{D}(A)} H(U(\mathcal{S})) \\ &\leq H(U(\mathcal{J}_1)) - \min_{\mathcal{S} \in \mathscr{D}(A)} H(U(\mathcal{S})|Y), \end{aligned}$$

which implies that $(R_1, R_2, \ldots, R_k) \in \tilde{\mathscr{R}}(A, \mathcal{J}_1, U^k)$. The claim follows since A is an arbitrary set in $\mathcal{B}$. The restrictions on the input pmfs can be removed again by the denseness argument. ■

VI. Multiple-Receiver Multiple Access Channels

We consider the two-receiver DM-MAC $p(y_1, y_2|x_1, x_2)$, where each sender wishes to convey its own message to both of the receivers. Given input pmfs $p(x_1)$ and $p(x_2)$, define $\mathscr{R}_{\text{MAC}}^{(1)}(X_1, X_2)$ as the set of rate pairs satisfying

$$\begin{aligned} R_1 &\leq I(X_1; Y_1|X_2), \\ R_2 &\leq I(X_2; Y_1|X_1), \\ R_1 + R_2 &\leq I(X_1, X_2; Y_1), \end{aligned}$$

and $\mathscr{R}_{\text{MAC}}^{(2)}(X_1, X_2)$ as the set of rate pairs satisfying

$$\begin{aligned} R_1 &\leq I(X_1; Y_2|X_2), \\ R_2 &\leq I(X_2; Y_2|X_1), \\ R_1 + R_2 &\leq I(X_1, X_2; Y_2). \end{aligned}$$

The following proposition then characterizes the achievable rate region by random homologous codes.

Proposition 6 (Extension of Theorem 1 to two-receiver): A rate pair (R_1, R_2) is achievable by random homologous codes in some finite field for the two-receiver DM-MAC $p(y_1, y_2|x_1, x_2)$, if

$$(R_1, R_2) \in \mathscr{R}_{\text{MAC}}^{(1)}(X_1, X_2) \cap \mathscr{R}_{\text{MAC}}^{(2)}(X_1, X_2)$$

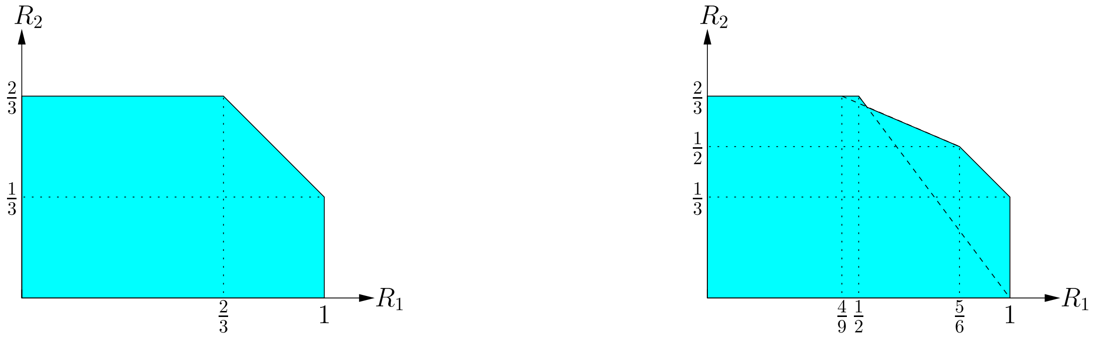
for some pmfs $p(x_1)$ and $p(x_2)$.

Proof: The achievable rate region depends on the conditional pmf $p(y_1, y_2|x_1, x_2)$ only through the conditional marginal pmfs $p(y_1|x_1, x_2)$ and $p(y_2|x_1, x_1)$. First suppose that $p(x_1)$ and $p(x_2)$ are of the form (9). We consider random homologous codes over $\mathbb{F}_q$ with $q = \rho^{2m}$. Choose U_1 and φ_1 such that U_1 and $\varphi_1(U_1) \overset{d}{=} X_1$ are one-to-one on the support of U_1 (this is always possible since $q \geq \rho^m$). Also choose $U_2 \sim \text{Unif}(\mathbb{F}_q)$ and φ_2 such that $\varphi_2(U_2) \overset{d}{=} X_2$ (this is possible due to the form of $p(x_2)$). By Proposition 3, the achievable rate region is

$$\bigcap_{j=1}^{2} [\mathscr{R}_{\text{MAC}}^{(j)}(X_1, X_2) \cap \mathscr{R}_{\text{L}}^{(j)}(U_1, U_2)],$$

where $\mathscr{R}_{\text{L}}^{(j)}(U_1, U_2)$, $j = 1, 2$, is the set of rate pairs (R_1, R_2) satisfying (4) or (5) for the virtual DM-MAC $p(y_j|u_1, u_2)$. The argument in the proof of Theorem 1 can be applied to both of the DM-MACs $p(y_1|x_1, x_2)$ and $p(y_2|x_1, x_2)$. As a result, the rate region $\mathscr{R}_{\text{MAC}}^{(j)}(X_1, X_2) \cap \mathscr{R}_{\text{L}}^{(j)}(U_1, U_2)$, $j = 1, 2$, is equivalent to the rate region $\mathscr{R}_{\text{MAC}}^{(j)}(X_1, X_2)$, which implies the claim. The restriction on the input pmfs can be removed by the denseness argument. ■

As shown in the examples of the binary adder MAC, the binary erasure MAC, and the on–off erasure MAC, the insufficiency of shaping or channel transformation for *single-receiver* MACs can be overcome by time sharing. Indeed, either shaping or channel transformation can achieve the corner points of $\mathscr{R}_{\text{MAC}}(X_1, X_2)$ of a general DM-MAC $p(y|x_1, x_2)$. This is no longer the case for multiple receivers, however. As illustrated by the following example, a proper combination of shaping and channel transformation can strictly outperform shaping



(a) The capacity region.

(b) The achievable rate region implied by Proposition 1.

Fig. 10. The two-receiver MAC in Example 4.

or channel transformation alone even when time sharing is allowed only for the individual techniques.

Example 4 (A two-receiver MAC): Let $Y_1 = X_1 + X_2$ (binary erasure MAC), and $Y_2 = (2X_1 - 1) + Z(2X_2 - 1)$ (on–off erasure MAC), where $\mathcal{X}_1 = \mathcal{X}_2 = \{0, 1\}$ and $Z \sim \text{Bern}(2/3)$ is independent of X_1 and X_2. The capacity region of this two-receiver MAC is achieved by random coding with i.i.d. Bern(1/2) inputs X_1 and X_2, and is sketched in Fig. 10a. Given input pmfs $p(x_1)$ and $p(x_2)$, the achievable rate region via shaping in Proposition 1 (and Proposition 4) is

$$\bigcap_{j=1}^{2}[\mathscr{R}_{\text{MAC}}^{(j)}(X_1, X_2) \cap \mathscr{R}_{\text{L}}^{(j)}(X_1, X_2)],$$

where $\mathscr{R}_{\text{L}}^{(j)}(X_1, X_2)$, $j = 1, 2$, is the set of rate pairs (R_1, R_2) satisfying (4) or (5) for the DM-MAC $p(y_j|x_1, x_2)$. The union of this rate region over input pmfs $p(x_1)$ and $p(x_2)$ is shown in Fig. 10b. Even after convexification via time sharing, it is strictly smaller than the capacity region with the largest symmetric rate of 11/18, whereas the symmetric capacity is 2/3. In comparison, we can combine shaping with channel transformation to achieve the entire capacity region as follows. Consider random homologous codes over $\mathbb{F}_4 = \{0, 1, \alpha, \alpha+1\}$. Let $U_1 \sim \text{Unif}(\mathbb{F}_4)$ and $U_2 \sim \text{Bern}(1/2)$ be independent. For channel transformation, let $x_j = \varphi(u_j)$ where $\varphi(0) = \varphi(\alpha) = 0$, and $\varphi(1) = \varphi(\alpha + 1) = 1$. By this construction, X_1 and X_2 are i.i.d. Bern(1/2). Following similar steps to the proof of Proposition 6, it is easy to see that the achievable rate region under this construction is equivalent to $\mathscr{R}_{\text{MAC}}^{(1)}(X_1, X_2) \cap \mathscr{R}_{\text{MAC}}^{(2)}(X_1, X_2)$, which is the capacity region of this channel since $p(x_1)$ and $p(x_2)$ are chosen as the capacity-achieving distributions. Thus, combination of shaping with channel transformation not only achieves higher rates than the shaping technique alone, but also achieves the capacity region *without* the need for time sharing.

Remark 9: Proposition 6 can be extended to k-sender and r-receiver DM-MACs and compound MACs via the proof of Theorem 2.

VII. Gaussian Multiple Access Channels

Consider the 2-sender Gaussian MAC model

$$Y = g_1 X_1 + g_2 X_2 + Z,$$

with channel gains g_1 and g_2, additive noise $Z \sim \text{N}(0, 1)$, and average power constraints $\sum_{i=1}^{n} x_{ji}^2(m_j) \le nP$ for $j = 1, 2$. Let $S_j = g_j^2 P$, $j = 1, 2$. The following theorem establishes the achievability of the capacity region of the Gaussian MAC by random homologous codes.

Theorem 3 (Gaussian MACs): A rate pair (R_1, R_2) is achievable by random homologous codes in some finite field for the 2-sender Gaussian MAC, if

$$\begin{aligned} R_1 &\le \mathsf{C}(S_1), \\ R_2 &\le \mathsf{C}(S_2), \\ R_1 + R_2 &\le \mathsf{C}(S_1 + S_2), \end{aligned}$$

where $\mathsf{C}(x) = (1/2)\log(1 + x), x \ge 0$, is the Gaussian capacity function.

Proof: Theorem 3 can be proved using the discretization argument in [1, Section 3.4.1] together with the achievability proof for the 2-sender DM-MAC by random homologous codes. The proof along this line, however, needs two limit arguments—one for approximating a Gaussian random variable by a discrete random variable, and one for approximating the resulting pmf on a finite alphabet to the desired form in (9). We instead provide a simpler proof via a discretization mapping that results in a pmf of desired form in (9) and thus eliminates one of the limit arguments.

Let X_1 and X_2 be i.i.d. $\text{N}(0, P)$. For every $j = 1, 2, \ldots$, let $[X_1]_j \in \{F_{X_1}^{-1}(i/2^j) : i \in [2^j - 1]\}$ be a quantized version of X_1 obtained by mapping X_1 to the closest point $[X_1]_j$ such that $|[X_1]_j| \le |X_1|$, where $F_{X_1}(x)$ denotes the cdf of random variable X_1. Clearly, $\mathsf{E}([X_1]_j^2) \le \mathsf{E}(X_1^2) = P$ and the pmf of $[X_1]_j$ is of the form $r/2^j$ for some positive integer r. Define $[X_2]_j$ in a similar manner. Let $Y_j = g_1[X_1]_j + g_2[X_2]_j + Z$ be the output corresponding to the input pair $[X_1]_j$ and $[X_2]_j$, and let $[Y_j]_k$ be a quantized version of Y_j defined in the same manner. Now, by the achievability proof of Theorem 1, for every j, k, random homologous codes over $\mathbb{F}_q$ with $q = 2^{2j}$ can achieve the rate pair satisfying

$$\begin{aligned} R_1 &\le I([X_1]_j; [Y_j]_k | [X_2]_j), \\ R_2 &\le I([X_2]_j; [Y_j]_k | [X_1]_j), \\ R_1 + R_2 &\le I([X_1]_j, [X_2]_j; [Y_j]_k). \end{aligned}$$

By this type of discretization, weak convergence of $[X_1]_j$ to X_1 and $[X_2]_j$ to X_2 is preserved, and $([Y_j]_k - Y_j)$ tends to 0 as $k \to \infty$. Therefore, one can follow the same steps in the proof of [1, Lemma 3.2] to show that

$$\liminf_{j\to\infty} \lim_{k\to\infty} I([X_1]_j; [Y_j]_k | [X_2]_j) \geq I(X_1; Y|X_2),$$
$$\liminf_{j\to\infty} \lim_{k\to\infty} I([X_2]_j; [Y_j]_k | [X_1]_j) \geq I(X_2; Y|X_1),$$
$$\liminf_{j\to\infty} \lim_{k\to\infty} I([X_1]_j, [X_2]_j; [Y_j]_k) \geq I(X_1, X_2; Y),$$

which establishes the claim. ■

Remark 10: It is straightforward to extend the discretization argument described for the 2-sender Gaussian MAC to the k-sender case. Therefore, the capacity region of a Gaussian MAC in general is achievable by random homologous codes.

VIII. Simultaneous Computation and Communication Over Multiple Access Channels

In the previous sections, we have investigated the performance of homologous codes—which were originally proposed for computing linear combinations of the transmitted messages/codewords—for message communication over MACs. One immediate question arising from our investigation is whether one can use homologous codes for computation and communication at the same time. To be more specific, consider a multiple-receiver MAC in which some receiver wishes to recover a desired linear combination of messages (computation) while another receiver wishes to recover the messages themselves (communication). In this section, we demonstrate how random homologous codes discussed thus far can be adapted to simultaneously achieve such competing goals, highlighting the potential of homologous codes for a broader class of applications beyond multiple access communication.

We first formally define the $\mathbb{F}_q$-computation problem. Consider a two-sender DM-MAC $p(y|x_1, x_2)$ with arbitrary input alphabets $\mathcal{X}_1$ and $\mathcal{X}_2$. For a given finite field $\mathbb{F}_q$, an (n, nR_1, nR_2) code for $\mathbb{F}_q$-computation over the DM-MAC $p(y|x_1, x_2)$ consists of two message sets $\mathbb{F}_q^{nR_j}$, $j = 1, 2$, two encoders, where encoder $j = 1, 2$ assigns a codeword $x_j^n(m_j) \in \mathcal{X}^n$ to each message $m_j \in \mathbb{F}_q^{nR_j}$ and transmits x_j^n over the channel, and a decoder that finds an estimate $\hat{W}_{\mathbf{a}}^{nR_{\max}} = \hat{w}_{\mathbf{a}}^{nR_{\max}}(Y^n) \in \mathbb{F}_q^{nR_{\max}}$ of

$$W_{\mathbf{a}}^{nR_{\max}} := a_1[M_1 \; \mathbf{0}_{n(R_{\max}-R_1)}] \oplus a_2[M_2 \; \mathbf{0}_{n(R_{\max}-R_2)}] \quad (15)$$

for a desired (nonzero) vector $\mathbf{a} = [a_1 \; a_2]$ over $\mathbb{F}_q$, where $R_{\max} = \max_j R_j$. Hence, the goal of communication is to convey a linear combination of messages rather than the messages themselves. The probability of error is then defined as $P_e^{(n)} := \mathsf{P}(\hat{W}_{\mathbf{a}}^{nR_{\max}} \neq W_{\mathbf{a}}^{nR_{\max}})$. To eliminate the degenerate cases, we assume without loss of generality that $a_1, a_2 \neq 0$.

We start with a discussion on the performance of conventional unstructured codes for the computation problem. Given an input pmf $p = p(x_1)p(x_2)$, an $(n, nR_1, nR_2; p)$ *random i.i.d. code ensemble* for $\mathbb{F}_q$-computation over the DM-MAC $p(y|x_1, x_2)$ consists of two message sets $\mathbb{F}_q^{nR_1}$ and $\mathbb{F}_q^{nR_2}$, two encoders where encoder $j = 1, 2$ assigns randomly generated codewords $X_j^n(m_j)$ that are drawn i.i.d. from $\prod_{i=1}^n p_{X_j}(x_{ji})$ to each message $m_j \in \mathbb{F}_q^{nR_j}$, and a decoder that assigns an estimate of $W_{\mathbf{a}}^{nR_{\max}} \in \mathbb{F}_q^{nR_{\max}}$ to each received y^n. We define the optimal rate region $\mathscr{R}^*(p)$ achievable by p-distributed random i.i.d. codes for $\mathbb{F}_q$-computation as the closure of the set of rate pairs (R_1, R_2) such that $\lim_{n\to\infty} \mathsf{E}[P_e^{(n)}] = 0$. Note that this rate region corresponds to the largest achievable rate region that can be proved via the standard i.i.d. random coding argument.

One can come up with a decoder for random i.i.d. codes that first estimates the message pair (M_1, M_2) and then compute the desired linear combination of codewords. It is easy to see that by this approach, $\cup_{p(x_1),p(x_2)} \mathscr{R}_{\text{MAC}}(X_1, X_2)$ is achievable. Indeed, this is the *optimal* rate region achievable by random i.i.d. codes in the sense that it cannot be improved by using the optimal maximum likelihood decoder.

We now formally state this result, the proof of which is deferred to Appendix H.

Proposition 7 (i.i.d. codes for computation): Given a pmf $p = p(x_1)p(x_2)$, the optimal rate region achievable by p-distributed random i.i.d. codes for $\mathbb{F}_q$-computation is

$$\mathscr{R}^*(p) = \mathscr{R}_{\text{MAC}}(X_1, X_2).$$

We now continue with the construction of random *homologous* codes for the $\mathbb{F}_q$-computation. For a given symbol-by-symbol mappings $x_j = \varphi_j(u_j)$, $j = 1, 2$, where $\varphi_j : \mathbb{F}_q \to \mathcal{X}_j$, consider the virtual channel $p(y|u_1, u_2)$ with an input pmf $p = p(u_1)p(u_2)$ over $\mathbb{F}_q$. Given $\epsilon > 0$, an $(n, nR_1, n\hat{R}_1, nR_2, n\hat{R}_2, \mathbb{F}_q; p, \epsilon)$ random homologous code ensemble is constructed as in Definition 4 with a slight modification for computation in the decoder, which now assigns an estimate $\hat{w}^{nR_{\max}} \in \mathbb{F}_q^{nR_{\max}}$ to each received sequence y^n. A rate pair (R_1, R_2) is said to be achievable by random homologous codes for $\mathbb{F}_q$-computation if there exists a sequence of $(n, nR_1, n\hat{R}_1, nR_2, n\hat{R}_2, \mathbb{F}_q; p, \epsilon)$ random homologous code ensembles such that $\lim_{n\to\infty} \mathsf{E}[P_e^{(n)}] = 0$ for some pmf $p = p(u_1)p(u_2)$ and $\epsilon > 0$, and for some mappings $x_j = \varphi_j(u_j)$, $j = 1, 2$.

In the parlance of homologous codes developed in [8], when the rates are symmetric, a desired linear combination of *messages* can be obtained by recovering a desired linear combination of $[M_1 \; L_1]$ and $[M_2 \; L_2]$, which simplifies as the desired linear combination of the *auxiliary* codewords U_j^n, $j \in [k]$. This type of computation over the auxiliary codewords was studied in [8] and the following result was presented.

Proposition 8 (Homologous codes for computation [8, Theorem 1]): A rate pair (R_1, R_2) is achievable by random homologous codes for $\mathbb{F}_q$-computation of $a_1 U_1^n(M_1, L_1) \oplus a_2 U_2^n(M_2, L_2)$ if

$$R_1 \leq H(U_1) - H(a_1 U_1 \oplus a_2 U_2 | Y), \quad (16a)$$
$$R_2 \leq H(U_2) - H(a_1 U_1 \oplus a_2 U_2 | Y), \quad (16b)$$

for some input pmfs $p(u_1)$ and $p(u_2)$ over $\mathbb{F}_q$ and some mappings $\varphi_1(u_1)$ and $\varphi_2(u_2)$.

Using this result for the computation of $W_{\mathbf{a}}^{nR_{\max}}$, we can conclude that a symmetric rate pair $(R_1, R_2) = (R, R)$ is achievable by random homologous codes for $\mathbb{F}_q$-computation if it satisfies (16a) and (16b) for some input pmfs $p(u_1)$ and $p(u_2)$ over $\mathbb{F}_q$ and some mappings $\varphi_1(u_1)$ and $\varphi_2(u_2)$.

We now consider the two-sender two-receiver DM-MAC $p(y_1, y_2|x_1, x_2)$ and a finite field $\mathbb{F}_q$, where the first receiver wishes to recover a desired linear combination of messages in $\mathbb{F}_q$ as in (15) and the second receiver wishes to recover the messages themselves. We refer to this channel as the *compute-communicate* DM-MAC.

We start with the performance of random i.i.d. codes. Given an input pmf $p = p(x_1)p(x_2)$, let $\mathscr{R}^{(j)}_{\text{MAC}}(X_1, X_2)$, $j = 1, 2$, denote the pentagonal region in (1) evaluated for the DM-MAC $p(y_j|x_1, x_2)$. We can define the optimal rate region achievable by p-distributed random i.i.d. codes for the compute-communicate DM-MAC $p(y_1, y_2|x_1, x_2)$ in a similar manner to the $\mathbb{F}_q$-computation. Using similar steps to the proof of Proposition 7 yields the following.

Corollary 3 (i.i.d. codes for compute-communicate): Given a pmf $p = p(x_1)p(x_2)$, the optimal rate region achievable by p-distributed random i.i.d. codes for the compute-communicate DM-MAC $p(y_1, y_2|x_1, x_2)$ is

$$\mathscr{R}^{(1)}_{\text{MAC}}(X_1, X_2) \cap \mathscr{R}^{(2)}_{\text{MAC}}(X_1, X_2).$$

We then return back to our discussion on random homologous codes. Given mappings $\varphi_j : \mathbb{F}_q \to \mathcal{X}_j$, $j = 1, 2$, and input pmf $p = p(u_1)p(u_2)$, let $\mathscr{R}^{(1)}_{\text{COMP}}(U_1, U_2)$ denote the rate region in (16) evaluated for the virtual DM-MAC $p(y_1|u_1, u_2)$, and let $\mathscr{R}^{(2)}_{\text{L}}(U_1, U_2)$ denote the set of rate pairs (R_1, R_2) satisfying (4) or (5) for the virtual DM-MAC $p(y_2|u_1, u_2)$. Propositions 1 and 8 imply the following.

Corollary 4 (Homologous codes for compute-communicate): A symmetric rate pair (R_1, R_2) is achievable by random homologous codes for the compute-communicate DM-MAC $p(y_1, y_2|x_1, x_2)$ if it is included in

$$\mathscr{R}^{(1)}_{\text{COMP}}(U_1, U_2) \cap \mathscr{R}^{(2)}_{\text{MAC}}(X_1, X_2) \cap \mathscr{R}^{(2)}_{\text{L}}(U_1, U_2) \tag{17}$$

for some pmfs $p(u_1)$ and $p(u_2)$ over $\mathbb{F}_q$ and some mappings $\varphi_1(u_1)$ and $\varphi_2(u_2)$.

Indeed, it is possible to construct homologous codes over the extension field $\mathbb{F}_{q^r}$ for some positive integer r to enlarge the achievable rate region in Corollary 4. By allowing extension fields $\mathbb{F}_{q^r}$ for some positive integer r in the channel transformation step, we get the following.

Corollary 5 (Homologous codes over extension fields): A symmetric rate pair (R_1, R_2) is achievable for the compute-communicate DM-MAC $p(y_1, y_2|x_1, x_2)$ if it is included in the rate region in (17) for some input pmfs $p(u_1)$ and $p(u_2)$ over $\mathbb{F}_{q^r}$, for some mappings $\varphi_j : \mathbb{F}_{q^r} \to \mathcal{X}_j$, $j = 1, 2$, and for some $r \in \mathbb{Z}^+$.

The results presented thus far in this section can be extended to arbitrary number of senders and receivers. As an example, we consider simultaneous computation and communication over a two-sender three-receiver DM-MAC and illustrate that random homologous codes, combined with carefully chosen channel transformation, outperform random i.i.d. codes as well as random homologous codes without channel transformation.

Example 5: Consider the *compute-communicate* DM-MAC $p(y_1, y_2, y_3|x_1, x_2)$, where $\mathcal{X}_1 = \mathcal{X}_2 = \{0, 1\}$ and

$$\begin{aligned} Y_1 &= X_1 \oplus X_2, && \text{(binary adder MAC)}\\ Y_2 &= X_1 + X_2, && \text{(binary erasure MAC)}\\ Y_3 &= (2X_1 - 1) + Z(2X_2 - 1), && \text{(on–off erasure MAC)} \end{aligned}$$

where $Z \sim \text{Bern}(2/3)$ is independent of X_1 and X_2. Receiver 1 wishes to recover $M_1 \oplus M_2$ over a binary field $\mathbb{F}_2$, whereas both receivers 2 and 3 wish to recover the message pair (M_1, M_2).

We now compare the largest achievable symmetric rate by different class of codes.

1) **Random i.i.d. codes**: Corollary 3 implies that the optimal achievable rate region by random i.i.d. codes is the intersection of the capacity regions of the DM-MACs $p(y_1|x_1, x_2)$, $p(y_2|x_1, x_2)$, and $p(y_3|x_1, x_2)$, each of which is achieved by i.i.d. Bern$(1/2)$ inputs X_1 and X_2, and so is the intersection. Fig. 11a sketches the rate region. In particular, the largest possible symmetric rate achievable by random i.i.d. codes is $1/2$.

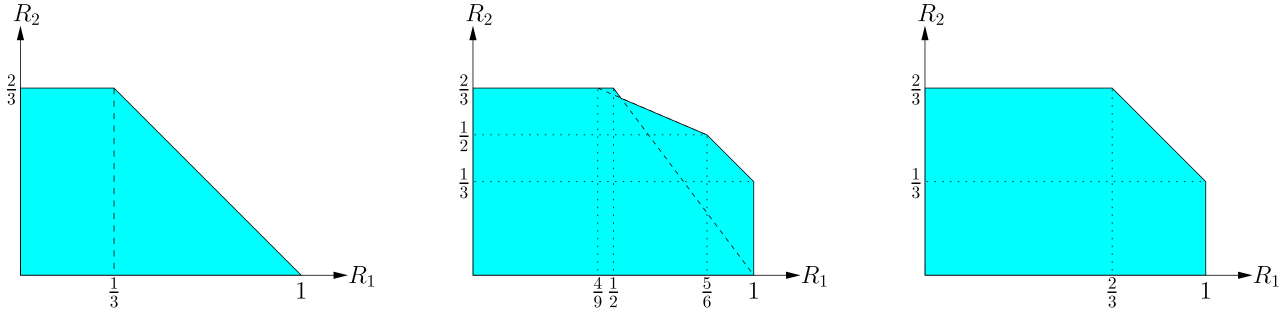
2) **Binary random homologous codes**: Since the channel input are also binary, it corresponds to setting $x_j = u_j$, $j \in [3]$ without any channel transformation. By Corollary 4, for any given input pmfs $p(x_1)$ and $p(x_2)$ over $\mathbb{F}_2$, a symmetric rate pair (R, R) is achievable if it is included in

$$\mathscr{R}^{(1)}_{\text{COMP}}(X_1, X_2) \cap \bigcap_{j=2}^{3} [\mathscr{R}^{(j)}_{\text{MAC}}(X_1, X_2) \cap \mathscr{R}^{(j)}_{\text{L}}(X_1, X_2)]. \tag{18}$$

Note that the rate region $\mathscr{R}^{(1)}_{\text{COMP}}(X_1, X_2)$ is larger than the rest of the terms in (18) for any given input pmfs $p(x_1)$ and $p(x_2)$. Taking the union of the rate region in (18) over the input pmfs results in the same rate region sketched earlier in Fig. 10b for the two-receiver DM-MAC $p(y_2, y_3|x_1, x_2)$ and is given in Fig. 11b for comparison. Therefore, the largest achievable symmetric rate in this region is $3/5$.

3) **Quaternary random homologous codes**: We are now allowed to use a larger finite field via channel transformation, but we need to be more careful for the choice of channel transformation because we have an additional receiver decoding for the sum of virtual codewords rather than the messages themselves. It is easy to see that the construction proposed for Example 4 results in $H(U_2) - H(U_1 \oplus U_2|Y_1) = 0$ and thus it is too restrictive for the first receiver to directly decode for $U_1^n \oplus U_2^n$. If it decodes for (M_1, M_2) instead, then the resulting achievable rate region will be the same as random i.i.d. codes. Therefore, we introduce a better construction here. Let $U_1 \sim \text{Unif}(\mathbb{F}_4)$ and

$$U_2 = \begin{cases} 0 & \text{with probability } \frac{1-\gamma}{2} \\ 1 & \text{with probability } \frac{1-\gamma}{2} \\ \alpha & \text{with probability } \frac{\gamma}{2} \\ \alpha + 1 & \text{with probability } \frac{\gamma}{2} \end{cases},$$



(a) Corollary 3.

(b) Corollary 4.

(c) Corollary 5.

Fig. 11. Achievable rate regions for the compute-communicate MAC in Example 5.

be independent for some γ chosen such that $H(\gamma) \in [1/3, 2/3]$. Let $x_j = \varphi(u_j)$ where $\varphi(0) = \varphi(\alpha) = 0$, and $\varphi(1) = \varphi(\alpha + 1) = 1$. By this construction, X_1 and X_2 are i.i.d. Bern(1/2). By Corollary 5, for a given γ and the corresponding pmf $p(u_1, u_2, x_1, x_2)$, it is easy to see that a symmetric rate pair $(R_1, R_2) = (R, R)$ is achievable if it satisfies

$$R_1 < 4/3 - H(\gamma),$$
$$R_2 < H(\gamma).$$

Taking the union over γ such that $H(\gamma) \in [1/3, 2/3]$ results in the rate region sketched in Fig. 11c. Therefore, the largest achievable symmetric rate is 2/3, which can be shown to be the symmetric capacity for this example.

IX. Concluding Remarks

In this paper, we examined the possibility of reestablishing the well-known achievable rate regions by random code ensembles for the MAC by using structured, homologous codes. We clarified two key techniques of shaping and channel transformation to employ nonuniform codewords while preserving a similar code structure across multiple users. The analysis tools developed for these techniques seem to indicate that their individual performance is insufficient. As a constructive alternative to individual techniques and their limitations, we showed that an appropriately designed combination of the two can establish the performance of random code ensembles. This development and its generalization to multiple senders and receivers motivate further research into the potential of homologous coding in network information theory.

Appendix A
A Proposition on Coset Codes for the Binary Erasure MAC

Proposition 9: For the binary erasure MAC, no pair of binary coset codes with the same generator matrix can achieve the rate pair $(1/2 + \epsilon, 1/2 + \epsilon)$ for $\epsilon > 0$.

Proof: Let $\epsilon > 0$ and $R_1 = R_2 = R = 1/2 + \epsilon$. Suppose without loss of generality that $nR \in \mathbb{Z}^+$, and that the generator matrix G is a fixed (not random) full rank $nR \times n$ matrix and does not have an all zero column. Let d_1^n and d_2^n be two arbitrary fixed binary coset sequences of length n. The messages M_1 and M_2 are assumed to be i.i.d. Unif$(\mathbb{F}_2^{nR})$. The received sequence is then written as

$$Y^n = (M_1 G \oplus d_1^n) + (M_2 G \oplus d_2^n).$$

Define $\tilde{Y}_i = (Y_i) \mod 2$ for every $i \in [n]$, which implies

$$\tilde{Y}^n = (M_1 \oplus M_2)G \oplus (d_1^n \oplus d_2^n).$$

Define the random set $\mathcal{S}(\tilde{Y}^n) = \{i : \tilde{Y}_i = 0\}$, and let the random variable $N_0 = |\mathcal{S}(\tilde{Y}^n)|$ denote the number of positions where sequence $\tilde{Y}^n$ has 0. We construct a new (random) matrix $G_{\mathcal{S}}$ of size $nR \times N_0$ by including the columns g_i of G for $i \in \mathcal{S}$. Note that the randomness in $G_{\mathcal{S}}$ is only due to the randomness of the messages M_1 and M_2 because the coset code parameters (G, d_1^n, d_2^n) are arbitrarily fixed. Then, the decoder makes an error if the following event occurs

$$\mathcal{E} = \{N_0 < nR\}.$$

This observation follows from the fact that on $\mathcal{E}$, the dimension of the null space of $G_{\mathcal{S}}^T$ is strictly larger than 0, so $\exists\ (m_1, m_2) \neq (M_1, M_2)$ such that $(m_1 \oplus M_1)G_{\mathcal{S}} = \mathbf{0}$ and $m_1 \oplus m_2 = M_1 \oplus M_2$, which leads to the same received sequence Y^n.

By the union of events bound, we have $P_e^{(n)} \geq \mathsf{P}(\mathcal{E}) = 1 - \mathsf{P}(\mathcal{E}^c)$. To bound the probability $\mathsf{P}(\mathcal{E}^c)$, we define the coset code $\mathcal{C} = \{x^n \in \mathbb{F}_2^n : x^n = mG \oplus d_1^n \oplus d_2^n,\ m \in \mathbb{F}_2^{nR}\}$. Then, $\tilde{Y}^n$ is uniformly distributed among $\mathcal{C}$, and we have

$$\begin{aligned} P(\mathcal{E}^c) &\overset{(a)}{\leq} \frac{\mathsf{E}[N_0]}{nR} \\ &= \frac{\sum_{x^n \in \mathcal{C}} \mathsf{P}(\tilde{Y}^n = x^n) wt((x^n)^c)}{nR} \\ &= \frac{\sum_{x^n \in \mathcal{C}} 2^{-nR} wt((x^n)^c)}{nR}, \\ &\overset{(b)}{=} \frac{2^{-nR}(n 2^{nR-1})}{nR}, \\ &= \frac{1}{1 + 2\epsilon}, \end{aligned}$$

where function $wt : \mathbb{F}_2^n \to \mathbb{Z}^+$ returns the Hamming weight of the input, (a) follows from Markov's inequality and (b) follows from the fact that for a binary coset code $\mathcal{C}$, at a given index, exactly half of the codewords have 0 and exactly half of the codewords have 1 (remember that its generator matrix

G has no all-zero column). It follows that $P_e^{(n)} \geq \frac{2\epsilon}{1+2\epsilon}$, which proves the claim. ■

Appendix B
Equivalence of Two Rate Regions

Lemma 2: Given input pmfs $p(x_1)$ and $p(x_2)$, let the rate region $\mathscr{R}(X_1, X_2)$ consist of the set of rate pairs (R_1, R_2) such that

$$\begin{aligned}&\min\{R_1 + H(X_2), R_2 + H(X_1)\}\\&\quad < H(X_1) + H(X_2) - \min\{H(X_1|Y), H(X_2|Y)\}.\end{aligned}$$

The region $\mathscr{R}(X_1, X_2)$ is equivalent to the rate region $\mathscr{R}_{\mathrm{L}}(X_1, X_2)$ described in (4) and (5).

Proof: It is easy to see that $\mathscr{R}(X_1, X_2) \subseteq \mathscr{R}_{\mathrm{L}}(X_1, X_2)$. To see the other direction, let the rate pair $(R_1, R_2) \in \mathscr{R}_{\mathrm{L}}(X_1, X_2)$ such that $R_1 + H(X_2) \leq R_2 + H(X_1)$. By the definition of the rate region $\mathscr{R}_{\mathrm{L}}(X_1, X_2)$, we have

$$R_1 + H(X_2) \leq \max\{H(X_2) + I(X_1; Y), H(X_1) + I(X_2; Y)\},$$

which implies that $(R_1, R_2) \in \mathscr{R}(X_1, X_2)$. Similarly, a rate pair $(R_1, R_2) \in \mathscr{R}_{\mathrm{L}}(X_1, X_2)$ such that $R_2 + H(X_1) \leq R_1 + H(X_2)$ is in $\mathscr{R}(X_1, X_2)$. Therefore, $\mathscr{R}_{\mathrm{L}}(X_1, X_2) \subseteq \mathscr{R}(X_1, X_2)$, from which the claim follows. ■

Appendix C
The Binary Adder MAC: The Achievable Rate Region by Proposition 1

When specialized to the binary adder MAC, the achievable rate region in Proposition 1 is reduced to the rate pairs (R_1, R_2) such that

$$\begin{aligned}R_1 &< I(X_1; Y),\\ R_2 &< I(X_2; Y|X_1) = H(X_2),\end{aligned}$$

or

$$\begin{aligned}R_1 &< I(X_1; Y|X_2) = H(X_1),\\ R_2 &< I(X_2; Y),\end{aligned}$$

for some input pmfs $p(x_1)$ and $p(x_2)$, which is equivalent to the capacity region depicted in Fig. 1a. To see this, let $\alpha \in [0, 1/2]$, and consider $X_1 \sim \mathrm{Bern}(\alpha)$ and $X_2 \sim \mathrm{Bern}\left(\frac{1}{2}\right)$. Then, the rate pairs (R_1, R_2) that satisfy

$$\begin{aligned}R_1 &< H(\alpha),\\ R_2 &< 1 - H(\alpha)\end{aligned}$$

are achievable, where $H(\alpha)$ denotes the binary entropy function defined in Section I. Since $H(\alpha)$ is continuous on α, taking the union over $\alpha \in [0, 1/2]$ implies that every point within the capacity region is achievable by the shaping technique. It follows from the converse proof for the capacity region of the binary adder MAC that the achievable rate region in Proposition 1 (over all input pmfs) is indeed equivalent to the capacity region.

Appendix D
The Binary Erasure MAC

1) The Achievable Rate Region by Proposition 1: For the binary erasure MAC, we will evaluate the rate region in Proposition 1. Let $\alpha, \beta \in [0, 1/2]$, and consider $X_1 \sim \mathrm{Bern}(\alpha)$ and $X_2 \sim \mathrm{Bern}(\beta)$. By Proposition 1, the set of rate pairs (R_1, R_2) such that

$$\begin{aligned}R_1 &< I(X_1; Y) = f(\alpha, \beta),\\ R_2 &< I(X_2; Y|X_1) = H(\beta),\end{aligned}$$

or

$$\begin{aligned}R_1 &< I(X_1; Y|X_2) = H(\alpha),\\ R_2 &< I(X_2; Y) = f(\beta, \alpha),\end{aligned}$$

is achievable, where the function $f : [0, 1/2] \times [0, 1/2] \to \mathbb{R}$ is defined as

$$\begin{aligned}f(x, y) = H(x) - y(1-x)\log\left(1 + \frac{x}{1-x}\frac{1-y}{y}\right)\\ - x(1-y)\log\left(1 + \frac{1-x}{x}\frac{y}{1-y}\right).\end{aligned} \tag{19}$$

Since $f(x, y)$ is increasing on x for any $y \in [0, 1/2]$, the union of such regions over $\alpha, \beta \in [0, 1/2]$ is the set of rate pairs (R_1, R_2) satisfying

$$\begin{aligned}R_1 &< 1 - \frac{H(\alpha)}{2},\\ R_2 &< H(\alpha),\end{aligned}$$

or

$$\begin{aligned}R_1 &< H(\alpha),\\ R_2 &< 1 - \frac{H(\alpha)}{2},\end{aligned}$$

for some $\alpha \in [0, 1/2]$. By the fact that $H(\alpha) \in [0, 1]$ is continuous on α, this union is equivalent to the union of two trapezoids defined by

$$\begin{aligned}R_2 &< 1,\\ 2R_1 + R_2 &< 2,\end{aligned}$$

and

$$\begin{aligned}R_1 &< 1,\\ R_1 + 2R_2 &< 2,\end{aligned}$$

which proves the claim.

2) The Achievable Rate Region by Corollary 1: For the binary erasure MAC, we will evaluate the rate region in Corollary 1. Let $\alpha, \beta \in [0, 1/2]$, and consider $X_1 \sim \mathrm{Bern}(\alpha)$ and $X_2 \sim \mathrm{Bern}(\beta)$. By Corollary 1, the set of rate pairs (R_1, R_2) such that

$$\begin{aligned}R_1 &< \min\{I(X_1; Y|X_2), \max[I(X_1; Y), I(X_2; Y)]\}\\ &= \min\{H(\alpha), \max[f(\alpha, \beta), f(\beta, \alpha)]\},\\ R_2 &< I(X_2; Y|X_1) = H(\beta),\\ R_1 + R_2 &< I(X_1, X_2; Y)\\ &= H(\alpha) + f(\beta, \alpha) = H(\beta) + f(\alpha, \beta),\end{aligned} \tag{20}$$

or

$$\begin{aligned} R_1 &< I(X_1;Y|X_2) = H(\alpha), \\ R_2 &< \min\{I(X_2;Y|X_1), \max[I(X_1;Y), I(X_2;Y)]\} \\ &= \min\{H(\beta), \max[f(\alpha,\beta), f(\beta,\alpha)]\}, \qquad (21) \\ R_1 + R_2 &< I(X_1, X_2; Y) \\ &= H(\alpha) + f(\beta,\alpha) = H(\beta) + f(\alpha,\beta), \end{aligned}$$

is achievable, where the function f is as defined in (19). First, consider the union of such regions over $\alpha, \beta \in [0, 1/2]$ such that $\alpha \geq \beta$ (or equivalently $f(\alpha,\beta) \geq f(\beta,\alpha)$), which results in the rate region defined by

$$\begin{aligned} R_1 &< f(\alpha,\beta), \\ R_2 &< H(\beta), \end{aligned}$$

or

$$\begin{aligned} R_1 &< H(\alpha), \\ R_2 &< \min\{H(\beta), f(\alpha,\beta)\}, \\ R_1 + R_2 &< H(\beta) + f(\alpha,\beta), \end{aligned}$$

for some $\alpha, \beta \in [0, 1/2]$ such that $\alpha \geq \beta$. Since $f(x, y)$ is increasing over x for any $y \in [0, 1/2]$, the resulting region consists of the rate pairs (R_1, R_2) satisfying

$$\begin{aligned} R_1 &< f(1/2,\beta) = 1 - \frac{H(\beta)}{2}, \\ R_2 &< H(\beta), \end{aligned} \qquad (22)$$

or

$$\begin{aligned} R_1 &< 1, \\ R_2 &< \min\{H(\beta), 1 - \frac{H(\beta)}{2}\}, \\ R_1 + R_2 &< 1 + \frac{H(\beta)}{2}, \end{aligned} \qquad (23)$$

for some $\beta \in [0, 1/2]$. The union of the rate region defined in (22) over $\beta \in [0, 1/2]$ is equivalent to the trapezoid defined by $R_2 < 1$, and $2R_1 + R_2 < 2$. The union of the rate region defined in (23) over $\beta \in [0, 1/2]$ is clearly included in the trapezoid defined by $R_1 < 1$, $R_1 + 2R_2 < 2$.

By similar arguments, the union of the rate region defined in (20) and (21) over $\alpha, \beta \in [0, 1/2]$ such that $\beta \geq \alpha$, is reduced to the rate pairs (R_1, R_2) such that

$$\begin{aligned} R_1 &< \min\{H(\alpha), 1 - \frac{H(\alpha)}{2}\}, \\ R_2 &< 1, \\ R_1 + R_2 &< 1 + \frac{H(\alpha)}{2}, \end{aligned}$$

or

$$\begin{aligned} R_1 &< H(\alpha), \\ R_2 &< 1 - \frac{H(\alpha)}{2}, \end{aligned}$$

for some $\alpha \in [0, 1/2]$. By symmetry, the overall achievable rate region in Corollary 1 is equivalent to the union of two trapezoids defined by $R_2 < 1$, $2R_1 + R_2 < 2$ and $R_1 < 1$, $R_1 + 2R_2 < 2$.

Appendix E
The On–Off Erasure MAC

3) The Achievable Rate Region by Proposition 1: For the on–off erasure MAC, we will evaluate the achievable rate region in Proposition 1. If the channel parameter $p \leq 2/3$, it is easy to see that i.i.d. Bern(1/2) inputs X_1 and X_2 can achieve the capacity region in Fig. 5a. Suppose that $p > 2/3$. Let $\alpha, \beta \in [0, 1/2]$, and consider $X_1 \sim \text{Bern}(\alpha)$ and $X_2 \sim \text{Bern}(\beta)$. Then, by Proposition 1, the set of rate pairs (R_1, R_2) such that

$$\begin{aligned} R_1 &< I(X_1;Y) = (1-p)H(\alpha) + pf(\alpha,\beta), \\ R_2 &< I(X_2;Y|X_1) = pH(\beta), \end{aligned} \qquad (24)$$

or

$$\begin{aligned} R_1 &< I(X_1;Y|X_2) = H(\alpha), \\ R_2 &< \min\{I(X_2;Y|X_1), \\ &\qquad H(X_2) - H(X_1) + I(X_1;Y)\} \qquad (25) \\ &= \min\{pH(\beta), (1-p)H(\beta) + pf(\beta,\alpha)\}, \\ R_1 + R_2 &< H(\alpha) + pf(\beta,\alpha), \end{aligned}$$

is achievable, where function f is as defined in (19). First, consider the union of the rate region defined in (24) over $\alpha, \beta \in [0, 1/2]$. Since $f(x, y)$ is increasing on x for every $y \in [0, 1/2]$, the union is equivalent to the set of rate pairs (R_1, R_2) satisfying

$$\begin{aligned} R_1 &< 1 - p + p\left(1 - \frac{H(\beta)}{2}\right) = 1 - \frac{pH(\beta)}{2}, \\ R_2 &< pH(\beta), \end{aligned}$$

for some $b \in [0, 1/2]$, that reduces to the trapezoid defined by $R_2 < p$ and $2R_1 + R_2 < 2$.

Second, we consider the union of the rate region defined in (25) over $\alpha, \beta \in [0, 1/2]$. By similar arguments, the union is equivalent to the set of rate pairs (R_1, R_2) such that

$$\begin{aligned} R_1 &< H(\alpha), \\ R_2 &< \min\{p, 1 - \frac{pH(\alpha)}{2}\}, \\ R_1 + R_2 &< p + H(\alpha)\left(1 - \frac{p}{2}\right), \end{aligned}$$

for some $\alpha \in [0, 1/2]$, that is equivalent to the hexagon defined by $R_1 < 1$, $R_2 < p$, $R_1 + R_2 < 1 + p/2$, and $(p/2)R_1 + R_2 < 1 - (p/2) + (p^2)/2$.

4) The Achievable Rate Region by Corollary 1: For the on–off erasure MAC, we will evaluate the achievable rate region in Corollary 1. Again, if the channel parameter $p \leq 2/3$, it is easy to see that i.i.d. Bern(1/2) inputs X_1 and X_2 can achieve the capacity region in Fig. 5a. Suppose that $p > 2/3$. Let $\alpha, \beta \in [0, 1/2]$, and consider $X_1 \sim \text{Bern}(\alpha)$ and $X_2 \sim \text{Bern}(\beta)$. Then, by Corollary 1, the set of rate pairs (R_1, R_2) such that

$$\begin{aligned} R_1 &< I(X_1;Y|X_2) = H(\alpha), && (26a) \\ R_1 &< \max\{I(X_1;Y), I(X_2;Y)\} && (26b) \\ &= \max\{pf(\alpha,\beta) + (1-p)H(\alpha), pf(\beta,\alpha)\}, \\ R_2 &< I(X_2;Y|X_1) = pH(\beta), && (26c) \\ R_1 + R_2 &< I(X_1, X_2; Y) = H(\alpha) + pf(\beta,\alpha), && (26d) \end{aligned}$$

or

$$\begin{aligned} R_1 &< I(X_1;Y|X_2) = H(\alpha), \\ R_2 &< I(X_2;Y|X_1) = pH(\beta), \\ R_2 &< \max\{I(X_1;Y), I(X_2;Y)\} \\ &\quad = \max\{pf(\alpha,\beta) + (1-p)H(\alpha), pf(\beta,\alpha)\}, \\ R_1 + R_2 &< I(X_1,X_2;Y) = H(\alpha) + pf(\beta,\alpha), \end{aligned} \tag{27}$$

is achievable, where the function f is as defined in (19). First, consider the union of the rate region defined in (26) over $\alpha, \beta \in [0, 1/2]$ such that $H(\alpha) > pH(\beta)$ (or equivalently $pf(\alpha,\beta) + (1-p)H(\alpha) > pf(\beta,\alpha)$). Then, the inequalities in (26a) and (26d) are inactive. Since $f(x, y)$ is increasing on x for every $y \in [0, 1/2]$, the union is equivalent to the set of rate pairs (R_1, R_2) satisfying

$$\begin{aligned} R_1 &< p\left(1 - \frac{H(\beta)}{2}\right) + (1-p) = 1 - \frac{pH(\beta)}{2}, \\ R_2 &< pH(\beta), \end{aligned}$$

for some $\beta \in [0, 1/2]$, that reduces to the trapezoid defined by $R_2 < p$ and $2R_1 + R_2 < 2$. It is easy to see that the union of the rate region defined in (26) over $\alpha, \beta \in [0, 1/2]$ such that $H(\alpha) \le pH(\beta)$ is included in this trapezoid.

Second, we consider the union of the rate region defined in (27) over $\alpha, \beta \in [0, 1/2]$ such that $H(\alpha) > pH(\beta)$. By similar arguments, the union is equivalent to the set of rate pairs (R_1, R_2) such that

$$\begin{aligned} R_1 &< 1, \\ R_2 &< \min\{pH(\beta), 1 - \frac{pH(\beta)}{2}\}, \\ R_1 + R_2 &< 1 + \frac{p}{2}H(\beta), \end{aligned}$$

for some $\beta \in [0, 1/2]$, that is equivalent to the hexagon defined by $R_1 < 1$, $R_2 < 2/3$, $R_1 + R_2 < 1 + p/2$, and $R_1 + 2R_2 < 2$. Finally, it is easy to see that the union of the rate region defined in (27) over $\alpha, \beta \in [0, 1/2]$ such that $H(\alpha) \le pH(\beta)$ is equivalent to the trapezoid defined by $R_1 < p$ and $(p/2)R_1 + R_2 < p$.

Appendix F
Proof of Proposition 2

Proof: We use a pair of $(n, nR_1, \mathbb{F}_q)$ and $(n, nR_2, \mathbb{F}_q)$ random coset code ensembles constructed for the virtual channel $p(y|u_1, u_2)$ as follows. A generator matrix $G \in \mathbb{F}_q^{n\max\{R_1,R_2\}\times n}$ and coset leaders D_1^n and D_2^n are randomly generated by drawing each entry i.i.d. $\mathrm{Unif}(\mathbb{F}_q)$. Given the realizations of G, d_1^n and d_2^n, for every message $m_j \in \mathbb{F}_q^{nR_j}$, encoder $j = 1, 2$ then assigns

$$u_j^n(m_j) = [m_j \; \mathbf{0}_{n(\max\{R_1,R_2\}-R_j)}]G + d_j^n.$$

Upon receiving y^n, the decoder first fixes an $\epsilon > 0$ and then searches a unique pair of $(\hat{m}_1, \hat{m}_2)$ such that $(u_1^n(\hat{m}_1), u_2^n(\hat{m}_2), y^n) \in \mathcal{T}_\epsilon^{(n)}(U_1, U_2, Y)$, where U_1 and U_2 are i.i.d. $\mathrm{Unif}(\mathbb{F}_q)$. If the decoder finds the unique pair, then it declares that $(\hat{m}_1, \hat{m}_2)$ was transmitted. Otherwise, it declares error. Assume that (M_1, M_2) is the transmitted message pair. We bound the probability of error $\mathsf{E}[P_e^{(n)}]$ averaged over (M_1, M_2) and (G, D_1^n, D_2^n). The code construction is symmetric with respect to the transmitted message pair. Therefore, $\mathsf{E}[P_e^{(n)}] = \mathsf{E}[P_e^{(n)}|(M_1, M_2) = (\mathbf{0}, \mathbf{0})]$ and without loss of generality, we can assume that $(M_1, M_2) = (\mathbf{0}, \mathbf{0})$. The decoder makes an error only if one or more of the following events occur:

$$\begin{aligned} \mathcal{E}_1 &= \{(U_1^n(\mathbf{0}), U_2^n(\mathbf{0}), Y^n) \notin \mathcal{T}_\epsilon^{(n)}(U_1, U_2, Y)\}, \\ \mathcal{E}_2 &= \{(U_1^n(\mathbf{0}), U_2^n(m_2), Y^n) \in \mathcal{T}_\epsilon^{(n)}(U_1, U_2, Y) \\ &\qquad \text{for some } m_2 \ne \mathbf{0}\}, \\ \mathcal{E}_3 &= \{(U_1^n(m_1), U_2^n(\mathbf{0}), Y^n) \in \mathcal{T}_\epsilon^{(n)}(U_1, U_2, Y) \\ &\qquad \text{for some } m_1 \ne \mathbf{0}\}, \\ \mathcal{E}_4 &= \{(U_1^n(m_1), U_2^n(m_2), Y^n) \in \mathcal{T}_\epsilon^{(n)}(U_1, U_2, Y) \\ &\qquad \text{for some } m_1 \ne \mathbf{0}, m_2 \ne \mathbf{0} \text{ such that} \\ &\qquad [m_1\; \mathbf{0}] \text{ and } [m_2\; \mathbf{0}] \text{ are linearly independent}\}, \\ \mathcal{E}_5 &= \{(U_1^n(m_1), U_2^n(m_2), Y^n) \in \mathcal{T}_\epsilon^{(n)}(U_1, U_2, Y) \\ &\qquad \text{for some } m_1 \ne \mathbf{0}, m_2 \ne \mathbf{0} \text{ such that} \\ &\qquad [m_1\; \mathbf{0}] \text{ and } [m_2\; \mathbf{0}] \text{ are linearly dependent}\}. \end{aligned}$$

Thus, by the union of events bound, $\mathsf{E}[P_e^{(n)}] \le \sum_{k=1}^5 \mathsf{P}(\mathcal{E}_k)$. Since $U_1^n(\mathbf{0}) = D_1^n$ and $U_2^n(\mathbf{0}) = D_2^n$ are i.i.d. $\mathrm{Unif}(\mathbb{F}_q^n)$ and independent from each other, by the law of large numbers, $\mathsf{P}(\mathcal{E}_1 \big| (M_1, M_2) = (\mathbf{0}, \mathbf{0}))$ tends to zero as $n \to \infty$. For the second term, note that for $m_2 \ne \mathbf{0}$, $U_2^n(m_2) \sim \prod_{i=1}^n p_{U_2}(u_{2i})$ is independent of $(U_1^n(\mathbf{0}), Y^n) \sim \prod_{i=1}^n p_{U_1,Y}(u_{1i}, y_i)$. Hence, by the packing lemma in [1, Section 3.2], $\mathsf{P}(\mathcal{E}_2)$ tends to zero as $n \to \infty$ if $R_2 \le I(U_2; U_1, Y) - \delta(\epsilon)$. Changing the role of sender 1 and 2, $\mathsf{P}(\mathcal{E}_3)$ tends to zero as $n \to \infty$ if $R_1 \le I(U_1; U_2, Y) - \delta(\epsilon)$. For the forth term, note that if $m_1 \ne \mathbf{0}$ and $m_2 \ne \mathbf{0}$ are linearly independent, then by [19, Lemma 14], $(U_1^n(m_1), U_2^n(m_2)) \sim \prod_{i=1}^n p_{U_1}(u_{1i})p_{U_2}(u_{2i})$; i.e., linear independence implies statistical independence. Moreover, in this case, the pair $(U_1^n(m_1), U_2^n(m_2))$ is independent from the tuple $(U_1^n(\mathbf{0}), U_2^n(\mathbf{0}), Y^n)$. Hence, again by the packing lemma $\mathsf{P}(\mathcal{E}_4)$ tends to zero as $n \to \infty$ if $R_1 + R_2 \le I(U_1, U_2; Y) - \delta(\epsilon)$.

Due to linear dependency among $U_1^n(m_1)$ and $U_2^n(m_2)$, to bound the last term, we will use a similar technique in Lemma 1. Define the rate $R = \min\{R_1, R_2\}$ and the set

$$\begin{aligned} \mathcal{D} = \{&(m_1, m_2) \in \mathbb{F}_q^{nR_1} \times \mathbb{F}_q^{nR_2} : \\ &[m_1\; \mathbf{0}] \ne \mathbf{0} \text{ and } [m_2\; \mathbf{0}] \ne \mathbf{0} \text{ are linearly dependent}\}. \end{aligned}$$

Then,

$$\begin{aligned} &\mathsf{P}(\mathcal{E}_5) \\ &= \mathsf{P}((U_1^n(m_1), U_2^n(m_2), Y^n) \in \mathcal{T}_\epsilon^{(n)}(U_1, U_2, Y) \\ &\qquad \text{for some } (m_1, m_2) \in \mathcal{D}) \\ &\overset{(a)}{\le} \sum_{\substack{(m_1,m_2) \\ \in \mathcal{D}}} \mathsf{P}((U_1^n(m_1), U_2^n(m_2), Y^n) \in \mathcal{T}_\epsilon^{(n)}(U_1, U_2, Y)) \\ &\le \sum_{(m_1,m_2)\in\mathcal{D}} \mathsf{P}((U_2^n(m_2), Y^n) \in \mathcal{T}_\epsilon^{(n)}(U_2, Y)) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{(m_1,m_2)\in\mathcal{D}} \sum_{\substack{(u_2^n,y^n)\in\\ \mathcal{T}_\epsilon^{(n)}(U_2,Y)}} \mathsf{P}(U_2^n(m_2)=u_2^n, Y^n=y^n)\\
&= \sum_{(m_1,m_2)\in\mathcal{D}} \sum_{\substack{(u_2^n,y^n)\in\\ \mathcal{T}_\epsilon^{(n)}(U_2,Y)}} \sum_{\substack{\tilde{u}_1^n\in\mathbb{F}_q^n,\\ \tilde{u}_2^n\in\mathbb{F}_q^n}}\\
&\quad \mathsf{P}(U_2^n(m_2)=u_2^n, Y^n=y^n, U_1^n(\mathbf{0})=\tilde{u}_1^n, U_2^n(\mathbf{0})=\tilde{u}_2^n)\\
&= \sum_{(m_1,m_2)\in\mathcal{D}} \sum_{\substack{(u_2^n,y^n)\in\\ \mathcal{T}_\epsilon^{(n)}(U_2,Y)}} \sum_{\substack{\tilde{u}_1^n\in\mathbb{F}_q^n,\\ \tilde{u}_2^n\in\mathbb{F}_q^n}}\\
&\quad \mathsf{P}([m_2\ \mathbf{0}]G+D_2^n=u_2^n, D_1^n=\tilde{u}_1^n, D_2^n=\tilde{u}_2^n, Y^n=y^n)\\
&\overset{(b)}{=} \sum_{(m_1,m_2)\in\mathcal{D}} \sum_{\substack{(u_2^n,y^n)\in\\ \mathcal{T}_\epsilon^{(n)}(U_2,Y)}} \sum_{\substack{\tilde{u}_1^n\in\mathbb{F}_q^n,\\ \tilde{u}_2^n\in\mathbb{F}_q^n}}\\
&\quad \mathsf{P}([m_2\ \mathbf{0}]G+D_2^n=u_2^n, D_1^n=\tilde{u}_1^n, D_2^n=\tilde{u}_2^n)p(y^n|\tilde{u}_1^n,\tilde{u}_2^n)\\
&\overset{(c)}{=} \sum_{(m_1,m_2)\in\mathcal{D}} \sum_{\substack{(u_2^n,y^n)\in\\ \mathcal{T}_\epsilon^{(n)}(U_2,Y)}} \sum_{\substack{\tilde{u}_1^n\in\mathbb{F}_q^n,\\ \tilde{u}_2^n\in\mathbb{F}_q^n}} q^{-3n} p(y^n|\tilde{u}_1^n,\tilde{u}_2^n)\\
&= \sum_{(m_1,m_2)\in\mathcal{D}} \sum_{\substack{(u_2^n,y^n)\in\\ \mathcal{T}_\epsilon^{(n)}(U_2,Y)}} q^{-n} p(y^n|\tilde{u}_1^n,\tilde{u}_2^n)\\
&\le \sum_{(m_1,m_2)\in\mathcal{D}} \sum_{y^n\in\mathcal{T}_\epsilon^{(n)}(Y)} p(y^n|\tilde{u}_1^n,\tilde{u}_2^n) \sum_{u_2^n\in\mathcal{T}_\epsilon^{(n)}(U_2|y^n)} q^{-n}\\
&\le |\mathcal{D}|\, q^{n(H(U_2|Y)+\delta(\epsilon))}q^{-n}\\
&\overset{(d)}{\le} q^{n(R-I(U_2;Y)+\delta(\epsilon))},
\end{aligned}
$$

where (a) follows by the union of events bound, (b) follows since under the assumption that $(M_1, M_2) = (\mathbf{0}, \mathbf{0})$, the the triple $G \to (D_1^n, D_2^n) \to Y^n$ form a Markov chain, (c) follows since $m_2 \neq \mathbf{0}$ and the entries of G, D_1^n and D_2^n are chosen i.i.d., and (d) follows since $H(U_2) = 1$ and $|\mathcal{D}| \le q q^{nR}$. By changing the order of U_1^n and U_2^n, we can conclude that

$$\mathsf{P}(\mathcal{E}_5) \le q^{n(R-\max\{I(U_1;Y),I(U_2;Y)\}+\delta(\epsilon))},$$

which tends to zero as $n \to \infty$ if $R = \min\{R_1, R_2\} < \max\{I(U_1; Y), I(U_2; Y)\} - \delta(\epsilon)$.

Letting $\epsilon \to 0$ yield that the rate pairs (R_1, R_2) is achievable if

$$
\begin{aligned}
R_1 &< I(U_1; Y|X_2),\\
R_2 &< I(U_2; Y|X_1),\\
R_1 + R_2 &< I(U_1, U_2; Y),\\
\min\{R_1, R_2\} &< \max\{I(U_1; Y), I(U_2; Y)\}.
\end{aligned}
$$

■

Appendix G
A Variation of Steinitz Lemma

Lemma 3: Suppose that $Z = \{z_1, z_2, \ldots, z_r\}$ is a set of linearly independent vectors in a vector space V of dimension $k > r$, and $W = \{w_1, w_2, \ldots, w_k\}$ span V. Let $T \subseteq W$ be a set such that

i) $|T| = k - r$, and
ii) $Z \cup T$ span V.

(The existence of such T is guaranteed by the Steinitz Lemma in [27]). Then, for a given set $J \subseteq W$ with $|J| = r$, $T = W\setminus J$ is the unique subset of W satisfying i) and ii) if and only if $\text{span}(Z) = \text{span}(J)$.

Proof: Let $J \subseteq W$ with $|J| = r$. First suppose that $\text{span}(Z) = \text{span}(J)$. Then, it is easy to see that $T = W \setminus J$ is the only subset of W that satisfies i) and ii). Now, suppose that $T = W \setminus J$ is the unique subset of W that satisfies i) and ii). We will show that

$$\text{span}(Z) = \text{span}(J).$$

Both Z and J consist of r linearly independent vectors, so it suffices to show that for every $w \in J$, $w \in \text{span}(Z)$. Let $w \in J$. Since $Z \cup T$ span V, we have

$$w = \sum_{l=1}^{r} a_l z_l + \sum_{w_i\in T} b_i w_i. \tag{28}$$

We want to show that $b_i = 0$ for all $w_i \in T$ in (28). Assume to the contrary that $b_m \neq 0$ for some $w_m \in T$. Then we can write w_m as a linear combination of the vectors in $Z \cup T \setminus \{w_m\} \cup \{w\}$. Note that $w \neq w_m$ since J and T are disjoint. Thus, $T' := T \setminus \{w_m\} \cup \{w\}$ also satisfies i) and ii), which contradicts with the uniqueness of T. The claim follows since $w \in J$ is arbitrary. ■

Appendix H
Proof of Proposition 7

An inner bound on the rate region in Proposition 7 follows by standard arguments. We prove an outer bound by showing that if a rate pair (R_1, R_2) is achievable by random i.i.d. codes, then it must be in $\mathscr{R}_{\text{MAC}}(X_1, X_2)$ for some input pmfs $p(x_1)$ and $p(x_2)$.

Fix an input pmf $p = p(x_1)p(x_2)$ and consider an $(n, nR_1, nR_2; p)$ random i.i.d. code ensemble. Define random codebook as

$$\mathcal{C}_n = \{(X_1^n(m_1), X_2^n(m_2) : m_1 \in \mathbb{F}_q^{nR_1}, m_2 \in \mathbb{F}_q^{nR_2}\},$$

which consists of the codewords in the ensemble. Assume without loss of generality that $R_1 \ge R_2$. Let $W_{\mathbf{a}}^{nR_1} := a_1 M_1 \oplus a_2[M_2\ \mathbf{0}]$ denote the desired linear combination. Suppose that the ensemble satisfies $\lim_{n\to\infty} \mathsf{E}[P_e^{(n)}(\mathcal{C}_n)] = 0$, where the expectation is taken over the random codebook $\mathcal{C}_n$. For a fixed codebook $\mathcal{C}_n = \mathcal{C}_n$, by Fano's inequality

$$H(W_{\mathbf{a}}^{nR_1}|Y^n, \mathcal{C}_n = \mathcal{C}_n) \le 1 + nP_e^{(n)}(\mathcal{C}_n).$$

Taking the expectation over the random codebook $\mathcal{C}_n$, we have

$$H(W_{\mathbf{a}}^{nR_1}|Y^n, \mathcal{C}_n) \le 1 + n\,\mathsf{E}[P_e^{(n)}(\mathcal{C}_n)] \overset{(a)}{\le} n\epsilon_n, \tag{29}$$

where (a) follows since $\mathsf{E}[P_e^{(n)}(\mathcal{C}_n)]$ tends to zero as $n \to \infty$. We start with

$$
\begin{aligned}
nR_1 &= H(M_1|M_2, \mathcal{C}_n)\\
&\le I(M_1; Y^n|M_2, \mathcal{C}_n) + H(M_1, W_{\mathbf{a}}^{nR_1}|Y^n, M_2, \mathcal{C}_n)\\
&\overset{(a)}{\le} I(M_1; Y^n|M_2, \mathcal{C}_n) + n\epsilon_n
\end{aligned}
$$

$$\begin{aligned}
&= \sum_{i=1}^{n} I(M_1; Y_i|Y^{i-1}, M_2, \mathcal{C}_n, X_{2i}) + n\epsilon_n \\
&\le \sum_{i=1}^{n} I(M_1, X_{1i}, Y^{i-1}, M_2, \mathcal{C}_n; Y_i|X_{2i}) + n\epsilon_n \\
&\overset{(b)}{=} \sum_{i=1}^{n} I(X_{1i}; Y_i|X_{2i}) + n\epsilon_n \\
&\overset{(c)}{=} nI(X_1; Y|X_2) + n\epsilon_n, \qquad (30)
\end{aligned}$$

where (a) follows by the modified Fano's inequality in (29), (b) follows since $(M_1, M_2, Y^{i-1}, \mathcal{C}_n) \to (X_{1i}, X_{2i}) \to Y_i$ form a Markov chain for every $i \in [n]$, and (c) follows by the i.i.d. codebook generation.

Following similar steps, it is easy to show that $nR_2 \le nI(X_2; Y|X_1) + n\epsilon_n$. For the sum rate bound, we start with

$$\begin{aligned}
&n(R_1 + R_2) \\
&\quad = H(M_1, M_2|\mathcal{C}_n) \\
&\quad \le I(M_1, M_2; Y^n|\mathcal{C}_n) + H(M_1, M_2, W_{\mathbf{a}}^{nR_1}|Y^n, \mathcal{C}_n) \\
&\quad \overset{(a)}{\le} I(M_1, M_2; Y^n|\mathcal{C}_n) + H(M_1, M_2|W_{\mathbf{a}}^{nR_1} Y^n, \mathcal{C}_n) + n\epsilon_n \\
&\quad = I(M_1, M_2; Y^n|\mathcal{C}_n) + H(M_2|W_{\mathbf{a}}^{nR_1} Y^n, \mathcal{C}_n) + n\epsilon_n, \qquad (31)
\end{aligned}$$

where (a) follows by the modified Fano's inequality in (29). The first term in (31) can be bounded by using similar arguments to the proof of (30) to get

$$I(M_1, M_2; Y^n|\mathcal{C}_n) \le nI(X_1, X_2; Y). \qquad (32)$$

We now bound the second term in (31) using a similar argument to [9], [28]. Given $W_{\mathbf{a}}^{nR_1}$, Y^n, and $\mathcal{C}_n$, a relatively short list $\mathcal{L} \subseteq \mathbb{F}_q^{nR_2}$ can be constructed that contains M_2 with high probability. Let $\epsilon > 0$ and define a random set

$$\mathcal{L} = \Big\{ m \in \mathbb{F}_q^{nR_2} : \Big(X_1^n\Big(\frac{W_{\mathbf{a}}^{nR_1} \ominus a_2[m\ \mathbf{0}]}{a_1}\Big), X_2^n(m), Y^n \Big) \in \mathcal{T}_\epsilon^{(n)}(X_1, X_2, Y) \Big\}.$$

By the symmetry of the codebook generation, assume without loss of generality that $M_1 = \mathbf{0}$ and $M_2 = \mathbf{0}$ was transmitted. For each $m \in \mathbb{F}_q^{nR_2}, m \neq \mathbf{0}$, we have

$$\begin{aligned}
\mathsf{P}(m \in \mathcal{L}) &= \mathsf{P}(m \in \mathcal{L}|M_1 = \mathbf{0}, M_2 = \mathbf{0}) \\
&= \mathsf{P}\Big(\Big(X_1^n\Big(-\frac{a_2}{a_1}m\Big), X_2^n(m), Y^n\Big) \\
&\qquad \in \mathcal{T}_\epsilon^{(n)}(X_1, X_2, Y)|M_1 = \mathbf{0}, M_2 = \mathbf{0}\Big) \\
&\le q^{-n(I(X_1, X_2; Y) - \delta(\epsilon))},
\end{aligned}$$

where the last step follows by the packing lemma since $X_1^n(-a_2/a_1 m)$ and $X_2^n(m)$ are independent from each other and are independent from Y^n. The expected cardinality of the set $\mathcal{L}$ is then bounded as

$$\begin{aligned}
&\mathsf{E}(|\mathcal{L}|) \\
&\quad \le 1 + \sum_{m \neq \mathbf{0}} \mathsf{P}(m \in \mathcal{L}) \le 1 + q^{n(R_2 - I(X_1, X_2; Y) + \delta(\epsilon))}. \qquad (33)
\end{aligned}$$

Define an indicator random variable $E_n = \mathbb{1}_{\{M_2 \in \mathcal{L}\}}$. By the conditional typicality lemma in [1, p. 27], $\mathsf{P}(E_n = 1)$ tends to one as $n \to \infty$. Then, for n sufficiently large, we have

$$\begin{aligned}
&H(M_2|W_{\mathbf{a}}^{nR_1}, Y^n, \mathcal{C}_n) \\
&\quad = H(M_2|W_{\mathbf{a}}^{nR_1}, Y^n, \mathcal{C}_n, E_n) + I(M_2; E_n|W_{\mathbf{a}}^{nR_1}, Y^n, \mathcal{C}_n) \\
&\quad \le H(M_2|W_{\mathbf{a}}^{nR_1}, Y^n, \mathcal{C}_n, E_n) + 1 \\
&\quad \le 1 + \mathsf{P}(E_n = 0) H(M_2|W_{\mathbf{a}}^{nR_1}, Y^n, \mathcal{C}_n, E_n = 0) \\
&\qquad\qquad + H(M_2|W_{\mathbf{a}}^{nR_1}, Y^n, \mathcal{C}_n, E_n = 1) \\
&\quad \le 1 + nR_2\, \mathsf{P}(E_n = 0) + H(M_2|W_{\mathbf{a}}^{nR_1}, Y^n, \mathcal{C}_n, E_n = 1).
\end{aligned}$$

We now use the fact that if $M_2 \in \mathcal{L}$, then the conditional entropy cannot exceed $\log(|\mathcal{L}|)$:

$$\begin{aligned}
&H(M_2|W_{\mathbf{a}}^{nR_1}, Y^n, \mathcal{C}_n, E_n = 1) \\
&\qquad \overset{(a)}{=} H(M_2|W_{\mathbf{a}}^{nR_1}, Y^n, \mathcal{C}_n, E_n = 1, \mathcal{L}, |\mathcal{L}|) \\
&\qquad \le H(M_2|E_n = 1, \mathcal{L}, |\mathcal{L}|) \\
&\qquad \le \sum_{l=0}^{q^{nR_2}} \mathsf{P}(|\mathcal{L}| = l) H(M_2|E_n = 1, \mathcal{L}, |\mathcal{L}| = l) \\
&\qquad \le \sum_{l=0}^{q^{nR_2}} \mathsf{P}(|\mathcal{L}| = l) \log(l) \\
&\qquad \overset{(b)}{\le} \log(\mathsf{E}[|\mathcal{L}|]) \\
&\qquad \overset{(c)}{\le} \max\{0, n(R_2 - I(X_1, X_2; Y) + \delta(\epsilon))\},
\end{aligned}$$

where (a) follows since the set $\mathcal{L}$ and its cardinality $|\mathcal{L}|$ are functions of $(W_{\mathbf{a}}^{nR_1}, Y^n, \mathcal{C}_n)$, (b) follows by Jensen's inequality, and (c) follows by (33) and the soft-max interpretation of the log-sum-exp function [29, p. 72]. Substituting back gives

$$\begin{aligned}
&n(R_1 + R_2) \\
&\quad \le nI(X_1, X_2; Y) + 1 + nR_2\, \mathsf{P}(E_n = 0) \\
&\qquad\quad + H(M_2|W_{\mathbf{a}}^{nR_1}, Y^n, \mathcal{C}_n, E_n = 1) + n\epsilon_n \\
&\quad \le nI(X_1, X_2; Y) + 1 + nR_2\, \mathsf{P}(E_n = 0) \\
&\qquad\quad + \max\{0, n(R_2 - I(X_1, X_2; Y) + \delta(\epsilon))\} + n\epsilon_n \\
&\quad = n \max\{I(X_1, X_2; Y), R_2 + \delta(\epsilon)\} + 2n\epsilon_n, \qquad (34)
\end{aligned}$$

where the last step follows since $\mathsf{P}(E_n = 0)$ tends to zero as $n \to \infty$. Combining (30), (31), (32), and (34) and letting $\epsilon \to 0$ implies that $(R_1, R_2) \in \mathscr{R}_{\text{MAC}}(X_1, X_2)$, which completes the proof.

Acknowledgment

The authors would like to thank Dr. K. Kim at ETRI for his invaluable comments and questions that prodded us to gain important insights about the problem.

References

[1] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York, NY, USA: Wiley, 2006.

[3] G. Kramer, "Topics in multi-user information theory," *Found. Trends Commun. Inf. Theory*, vol. 4, nos. 4–5, pp. 265–444, 2007.

[4] J. Körner and K. Marton, "How to encode the modulo-two sum of binary sources," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 2, pp. 219–221, Mar. 1979.

[5] A. Padakandla and S. S. Pradhan, "Computing sum of sources over an arbitrary multiple access channel," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2013, pp. 2144–2148.

[6] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6463–6486, Oct. 2011.

[7] Y. Song and N. Devroye, "Lattice codes for the Gaussian relay channel: Decode-and-forward and compress-and-forward," *IEEE Trans. Inf. Theory*, vol. 59, no. 8, pp. 4927–4948, Aug. 2013.

[8] S. H. Lim, C. Feng, A. Pastore, B. Nazer, and M. Gastpar, "A joint typicality approach to compute–forward," *IEEE Trans. Inf. Theory*, vol. 64, no. 12, pp. 7657–7685, Dec. 2018.

[9] P. Sen, S. H. Lim, and Y.-H. Kim, "Optimal achievable rates for computation with random homologous codes," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2018, pp. 2351–2355.

[10] P. Sen, S. H. Lim, and Y.-H. Kim, "Optimal achievable rates for computation with random homologous codes," *IEEE Trans. Inf. Theory*, to be published.

[11] V. Ntranos, V. R. Cadambe, B. Nazer, and G. Caire, "Integer-forcing interference alignment," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2013, pp. 574–578.

[12] O. Ordentlich, U. Erez, and B. Nazer, "The approximate sum capacity of the symmetric Gaussian K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 60, no. 6, pp. 3450–3482, Apr. 2014.

[13] A. Padakandla, A. G. Sahebi, and S. S. Pradhan, "An achievable rate region for the three-user interference channel based on coset codes," *IEEE Trans. Inf. Theory*, vol. 62, no. 3, pp. 1250–1279, Mar. 2016.

[14] A. Padakandla and S. S. Pradhan, "An achievable rate region based on coset codes for multiple access channel with states," *IEEE Trans. Inf. Theory*, vol. 63, no. 10, pp. 6393–6415, Oct. 2017.

[15] S. Miyake, "Coding theorems for point-to-point communication systems using sparse matrix codes," Ph.D. dissertation, Univ. Tokyo, Tokyo, Japan, 2010.

[16] B. Hern and K. R. Narayanan, "Multilevel coding schemes for compute-and-forward with flexible decoding," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7613–7631, Nov. 2013.

[17] R. G. Gallager, *Information Theory and Reliable Communication*. New York, NY, USA: Wiley, 1968.

[18] N. Karamchandani, U. Niesen, and S. Diggavi, "Computation over mismatched channels," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 4, pp. 666–677, Apr. 2013.

[19] S. H. Lim, C. Feng, A. Pastore, B. Nazer, and M. Gastpar, "Towards an algebraic network information theory: Simultaneous joint typicality decoding," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2017, pp. 1818–1822.

[20] B. Nazer, V. R. Cadambe, V. Ntranos, and G. Caire, "Expanding the compute-and-forward framework: Unequal powers, signal levels, and multiple linear combinations," *IEEE Trans. Inf. Theory*, vol. 62, no. 9, pp. 4879–4909, Sep. 2016.

[21] P. Sen and Y.-H. Kim, "Homologous codes for multiple access channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2017, pp. 874–878.

[22] R. Ahlswede, "Multi-way communication channels," in *Proc. 2nd Int. Symp. Inf. Theory*, Tsahkadsor, Armenian, 1971, pp. 23–52.

[23] H. H.-J. Liao, "Multiple access channels," Ph.D. dissertation, Univ. Hawaii, Honolulu, HI, USA, Sep. 1972.

[24] T. Gariby and U. Erez, "On general lattice quantization noise," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2008, pp. 2717–2721.

[25] R. Owen, *Lectures on the Comparative Anatomy and Physiology of the Invertebrate Animals*. London, U.K.: Longman, Brown, Green, Longmans, 1843.

[26] C. R. Darwin, *On the Origin of Species*. London, U.K.: John Murray, 1859.

[27] Y. Katznelson and Y. Katznelson, *A (Terse) Introduction to Linear Algebra* (Student Mathematical Library). Providence, RI, USA: AMS, 2008.

[28] B. Bandemer, A. El-Gamal, and Y. K. Kim, "Optimal achievable rates for interference networks with random codes," *IEEE Trans. Inf. Theory*, vol. 61, no. 12, pp. 6536–6549, Dec. 2015.

[29] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

Pinar Sen (S'13) received the B.S. and M.S. degrees in electrical and electronics engineering from Middle East Technical University, Ankara, Turkey, in 2011 and 2014, respectively. She is currently pursuing the Ph.D. degree in electrical and computer engineering with the University of California at San Diego, La Jolla, CA, USA. Her current research interests include coding and information theory in multiuser networks, with applications in wireless communications.

Young-Han Kim (S'99–M'06–SM'12–F'15) received the B.S. degree (Hons.) in electrical engineering from Seoul National University, Seoul, South Korea, in 1996, and the M.S. degrees in electrical engineering and in statistics and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, USA, in 2001, 2006, and 2006, respectively. In 2006, he joined the University of California at San Diego, La Jolla, CA, where he is currently a Professor with the Department of Electrical and Computer Engineering. He has coauthored the book *Network Information Theory* (Cambridge University Press, 2011). His current research interests include information theory, communication engineering, and data science. He was a recipient of the 2008 NSF Faculty Early Career Development Award, the 2009 US–Israel Binational Science Foundation Bergmann Memorial Award, the 2012 IEEE Information Theory Paper Award, and the 2015 IEEE Information Theory Society James L. Massey Research and Teaching Award for Young Scholars. He served as an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION THEORY and a Distinguished Lecturer for the IEEE Information Theory Society.