A Discretization Approach to Compute–Forward

Adriano Pastore*, Sung Hoon Lim†, Chen Feng‡, Bobak Nazer§, Michael Gastpar¶
*Communication Systems Division, CTTC/CERCA, Castelldefels, Spain, Email: adriano.pastore@cttc.cat
†School of Software, Hallym University, Chuncheon, Korea, Email: shlim@hallym.ac.kr
‡School of Engineering, University of British Columbia, Kelowna, BC, Canada, Email: chen.feng@ubc.ca
§ECE Department, Boston University, Boston, MA, USA, Email: bobak@bu.edu
¶School of Computer and Communication Sciences, EPFL, Lausanne, Switzerland, Email: michael.gastpar@epfl.ch

***Abstract*—We present a novel unified framework of compute–forward achievable rate regions for simultaneous decoding of multiple linear codeword combinations. This framework covers a wide class of discrete and continuous-input channels, and computation over finite fields, integers, and reals. The resulting rate regions recover several well-known achievability results, and in some cases extend them. The framework is built upon a recently established achievable rate region based on linear codes and joint typicality decoding. The latter is extended from finite fields to computation over the integers and, via a discretization approach, to computation over the reals with integer coefficients and continuous inputs. Evaluating the latter with Gaussian distributions, we obtain a closed-form rate region which generalizes the classic compute–forward rates originally derived by means of lattice codes by Nazer and Gastpar.**

I. Introduction

Consider a network information theory problem where one or more transmitters wish to communicate with one or more receivers, and our goal is to determine the rate region for a wide class of sources and channels, both discrete and continuous-valued. One approach is to handle the discrete and continuous cases separately. For instance, in the textbook of Cover and Thomas [1], the achievability proofs for discrete memoryless channels are developed first, and then Gaussian channels are handled using i.i.d. Gaussian codebooks. An alternative approach is to first establish a rate region for the discrete memoryless case, and then use discretization arguments to extend this to continuous-valued channels. For instance, in the textbook of El Gamal and Kim [2], the achievable rate for Gaussian (and other continuous-valued) channels is derived by applying scalar quantization to the channel input and output, and then taking appropriate limits of the resulting mutual information (with respect to the quantization resolution). A quantization approach has notoriously been used in the deterministic approach [3], though their focus is on achieving a constant gap to the cutset bound.

In this paper, we focus on the compute–forward problem where the goal is for one or more receivers to recover linear combinations of the transmitters' messages. Prior work has derived achievable rate regions for the special case of Gaussian channels using nested lattice codes [4], [5] and for the special case of discrete memoryless channels using nested linear codes [6], [7]. In particular, the latter approach employs standard joint typicality encoding and decoding techniques. Here, we propose a discretization approach to unify the treatment of discrete and continuous-valued channels for the compute–forward problem. Specifically, our approach recovers prior Gaussian compute–forward results [4], [5], and, in some cases, improves upon them. This is due to the fact that the underlying rate region for discrete channels [7] is based on simultaneous decoding rather than sequential decoding.

A key technical difference between our work and prior discretization approaches is that the discrete memoryless rate region is not described with mutual informations terms, but via entropies. As a result, the limit arguments are considerably more subtle, and require a generalization of Rényi's d-dimensional entropy [8] to handle linear combinations (which we term algebraic entropy). Overall, this paper, alongside several recent works [6], [7], [9]–[14], demonstrates that algebraic approaches to network information theory problems can be handled via standard techniques, such as joint typicality encoding and decoding.

II. Notation

For a matrix $\mathbf{A}$ (or column vector) and a set of row indices $\mathcal{S}$, $[\mathbf{A}]_{\mathcal{S}}$ denotes the submatrix of $\mathbf{A}$ comprising only those rows indexed by $\mathcal{S}$. For $\mathbb{U}$ a ring and some subset $\mathbb{A} \subseteq \mathbb{U}$, we define

$$\Lambda_{\mathbb{A}}(\mathbf{Q}) = \big\{\mathbf{Q}\mathbf{v} \colon \mathbf{v} \in \mathbb{A}^d\big\}. \tag{1}$$

If $\mathbb{A}$ is a discrete additive subgroup of $\mathbb{U}$, then $\Lambda_{\mathbb{A}}(\mathbf{Q})$ is called a *lattice* generated by $\mathbf{Q}$. If $\mathbb{A} = \mathbb{U}$, it is called the *span* of $\mathbf{Q}$. The largest integer that is smaller or equal to a given real number X is denoted as $\lfloor X \rfloor$.

III. General problem statement of compute–forward

Consider the K-user memoryless multiple-access channel (MAC) $(\mathcal{X}_1 \times \ldots \times \mathcal{X}_K, P_{Y|X_1,\ldots,X_K}, \mathcal{Y})$, which consists of K sender alphabets $\mathcal{X}_k$, $k \in [K]$, one receiver alphabet $\mathcal{Y}$, and a conditional probability distribution $P_{Y|X_1,\ldots,X_K}$.

Consider a ring $\mathbb{U}$ and a subset $\mathbb{A} \subseteq \mathbb{U}$ thereof. Let $\mathbf{a}_1^\mathsf{T}, \ldots, \mathbf{a}_L^\mathsf{T} \in \mathbb{A}^K$ denote L coefficient vectors, assume $L \leq K$, and define the coefficient matrix

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1^\mathsf{T} \\ \vdots \\ \mathbf{a}_L^\mathsf{T} \end{bmatrix} \in \mathbb{A}^{L \times K}. \tag{2}$$

A $(2^{nR_1}, \ldots, 2^{nR_K}, n)$ code for compute–forward over $(\mathbb{U}, \mathbb{A})$ with coefficient matrix $\mathbf{A}$ consists of

- K message sets $[2^{nR_k}]$, $k \in [K]$;
- K encoders, where encoder k maps a message $m_k \in [2^{nR_k}]$ to a codeword $u_k^n(m_k) \in \mathbb{U}^n$ such that $u_k^n(m_k)$ is *one-to-one*;

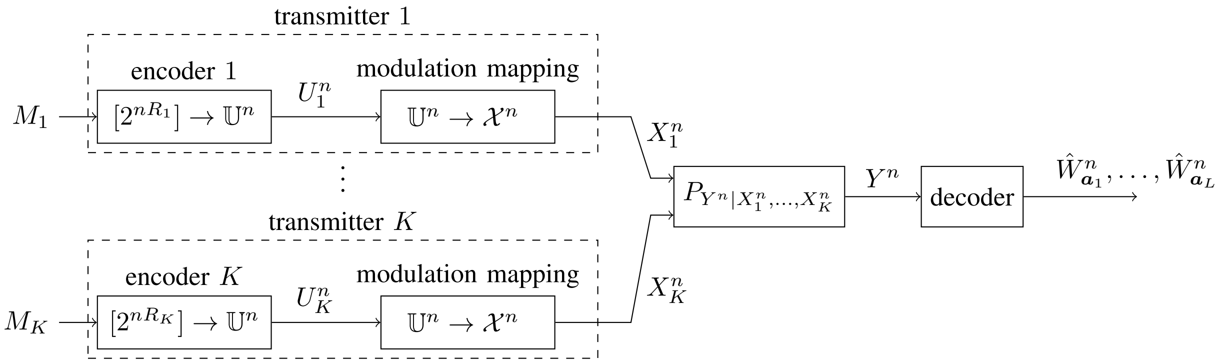

transmitter 1
encoder 1
modulation mapping
M_1
$[2^{nR_1}] \to \mathbb{U}^n$
U_1^n
$\mathbb{U}^n \to \mathcal{X}^n$
X_1^n
transmitter K
encoder K
modulation mapping
M_K
$[2^{nR_K}] \to \mathbb{U}^n$
U_K^n
$\mathbb{U}^n \to \mathcal{X}^n$
X_K^n
$P_{Y^n|X_1^n,\ldots,X_K^n}$
Y^n
decoder
$\hat{W}^n_{\boldsymbol{a}_1},\ldots,\hat{W}^n_{\boldsymbol{a}_L}$

Fig. 1. A many-to-one channel with K transmitters, in which the receiver seeks to decode L linear combinations of the codewords $(U_1^n,\ldots,U_K^n)$

- K modulation mappings $\mathbb{U} \to \mathcal{X}$ that map[1] the entries of a codeword $u_k^n \in \mathbb{U}^n$ entry-by-entry to a *physical codeword* $x_k^n(m_k) \in \mathcal{X}^n$;
- L linear combinations (where $L \leq K$) for each message tuple $(m_1,\ldots,m_K)$

$$\begin{bmatrix} w^n_{\mathbf{a}_1}(m_1,\ldots,m_K) \\ \vdots \\ w^n_{\mathbf{a}_L}(m_1,\ldots,m_K) \end{bmatrix} = \mathbf{A} \begin{bmatrix} u_1^n(m_1) \\ \vdots \\ u_K^n(m_K) \end{bmatrix},$$

  where additions and multiplications are defined over the vector space $\mathbb{U}^n$, and
- a decoder that assigns estimates $(\hat{w}^n_{\mathbf{a}_1},\ldots,\hat{w}^n_{\mathbf{a}_L}) \in \mathbb{U}^n \times \cdots \times \mathbb{U}^n$ to each received sequence $y^n \in \mathcal{Y}^n$.

Each message M_k is independently and uniformly drawn from $[2^{nR_k}]$. The average probability of error is defined as $P_e^{(n)} = \mathsf{P}\left\{(\hat{W}^n_{\mathbf{a}_1},\ldots,\hat{W}^n_{\mathbf{a}_L}) \neq (W^n_{\mathbf{a}_1},\ldots,W^n_{\mathbf{a}_L})\right\}$. We say that a rate tuple $(R_1,\ldots,R_K)$ is achievable for computing $\mathbf{A}$-linear combinations if there exists a sequence of $(2^{nR_1},\ldots,2^{nR_K},n)$ compute–forward codes such that $\lim_{n\to\infty} P_e^{(n)} = 0$.

IV. Preliminaries

In this section, we lay out some important concepts, notations and auxiliary results which are necessary for a complete understanding of the main theorems presented in Section V.

A. Algebraic entropy and algebraic information dimension

For a random variable $\boldsymbol{u}$ with countable support set $\mathcal{U}$, the Shannon entropy is defined in the usual way as

$$H(\boldsymbol{u}) = -\sum_{\mathbf{u}\in\mathcal{U}} \mathsf{P}\{\boldsymbol{u} = \mathbf{u}\} \log \mathsf{P}\{\boldsymbol{u} = \mathbf{u}\}. \tag{3}$$

For a real-valued, absolutely continuous random variable $\boldsymbol{u} \in \mathbb{R}^n$ with density $f_{\boldsymbol{u}}(\mathbf{u})$, the differential entropy is defined as

$$h(\boldsymbol{u}) = -\int f_{\boldsymbol{u}}(\mathbf{u}) \log f_{\boldsymbol{u}}(\mathbf{u})\,\mathrm{d}\mathbf{u}. \tag{4}$$

In his 1959 paper [8], Rényi elucidates some interesting connections between the discrete entropy of quantized variables and differential entropies. In particular, he introduces the concept of *information dimension*. We generalize his concept as follows.

Definition 1 (Algebraic information dimension and algebraic entropy)**.** *For a coefficient matrix $\mathbf{Q} \in \mathbb{R}^{m\times n}$ and a random vector $\boldsymbol{u} \in \mathbb{R}^n$, the algebraic information dimension $d_{\mathbf{Q}}(\boldsymbol{u})$ and the algebraic entropy $\mathcal{H}_{\mathbf{Q}}(\boldsymbol{u})$ shall be defined as*

$$d_{\mathbf{Q}}(\boldsymbol{u}) = \lim_{\nu\to\infty} \frac{H(\mathbf{Q}\lfloor \nu\boldsymbol{u}\rfloor)}{\log(\nu)} \tag{5a}$$

$$\mathcal{H}_{\mathbf{Q}}(\boldsymbol{u}) = \lim_{\nu\to\infty} \Big\{ H(\mathbf{Q}\lfloor \nu\boldsymbol{u}\rfloor) - d_{\mathbf{Q}}(\boldsymbol{u}) \log(\nu) \Big\} \tag{5b}$$

if the limits exist.

If $d_{\mathbf{Q}}(\boldsymbol{u})$ and $\mathcal{H}_{\mathbf{Q}}(\boldsymbol{u})$ exist, then given a pair of variables $(\boldsymbol{u}, Y) \in \mathbb{R}^n \times \mathcal{Y}$, one can define the conditional algebraic information dimension $d_{\mathbf{Q}}(\boldsymbol{u}|Y) = \int d_{\mathbf{Q}}(\boldsymbol{u}|Y = y)\,\mathrm{d}P_Y(y)$ and conditional algebraic entropy $\mathcal{H}_{\mathbf{Q}}(\boldsymbol{u}|Y) = \int \mathcal{H}_{\mathbf{Q}}(\boldsymbol{u}|Y = y)\,\mathrm{d}P_Y(y)$.[2]

1) Discrete distributions: The following lemma connects $\mathcal{H}_{\mathbf{Q}}(\boldsymbol{u})$ to discrete entropy (for discrete distributions).

Lemma 1. *For a real matrix $\mathbf{Q} \in \mathbb{R}^{m\times n}$ and a random variable $\boldsymbol{u} \in \mathbb{R}^n$ with discrete support (point mass) and finite discrete entropy $H(\boldsymbol{u})$, the algebraic information dimension and algebraic entropy are given by*

$$d_{\mathbf{Q}}(\boldsymbol{u}) = 0 \qquad\qquad \mathcal{H}_{\mathbf{Q}}(\boldsymbol{u}) = H(\mathbf{Q}\boldsymbol{u}). \tag{6}$$

By convention, (6) *shall also apply to finite fields, i.e., to the case $\boldsymbol{u} \in \mathbb{F}_{\mathsf{q}}^n$ and $\mathbf{Q} \in \mathbb{F}_{\mathsf{q}}^{m\times n}$.*

2) Continuous distributions: In the following, we present a lemma linking algebraic entropy to differential entropy (for continuous distributions), in a similar vein as [8]. Prior to stating it, we need some preliminary definitions.

Definition 2 (Unimodular matrix)**.** *A square integer matrix $\mathbf{Q} \in \mathbb{Z}^{n\times n}$ is* unimodular *if its inverse $\mathbf{Q}^{-1} \in \mathbb{Z}^{n\times n}$ is integer too. A matrix $\mathbf{Q}$ is unimodular iff $|\det(\mathbf{Q})| = 1$.*

Definition 3 (Right-invertible and left-invertible matrices)**.** *A strictly broad integer matrix $\mathbf{Q} \in \mathbb{Z}^{n\times m}$ with $n < m$ is said to be* right-invertible *if there exists a tall integer matrix $\mathbf{Q}^\sharp \in \mathbb{Z}^{m\times n}$ (called the* right-inverse*) such that $\mathbf{Q}\mathbf{Q}^\sharp = \mathbf{I}_n$. Similarly, a strictly tall integer matrix $\mathbf{Q} \in \mathbb{Z}^{n\times m}$ with $n > m$ is said to be* left-invertible *if there exists a broad integer matrix $\mathbf{Q}^\sharp \in \mathbb{Z}^{m\times n}$ (called the* left-inverse*) such that $\mathbf{Q}^\sharp\mathbf{Q} = \mathbf{I}_m$. The following statements are equivalent:*

[1] Modulation mappings need not be one-to-one (injective).

[2] Equivalently, to define $d_{\mathbf{Q}}(\boldsymbol{u})$ and $\mathcal{H}_{\mathbf{Q}}(\boldsymbol{u})$ one can replace the discrete entropies on the right-hand sides of (5a) and (5b) by $H(\mathbf{Q}\lfloor\nu\boldsymbol{u}\rfloor|Y)$. This can be shown using the Dominated Convergence Theorem and the Monotone Convergence Theorem, respectively.

1) $\mathbf{Q}$ *is right-invertible*
2) $\mathbf{Q}^\mathsf{T}$ *is left-invertible*
3) $\mathbf{Q}$ *can be completed to a unimodular matrix* $\begin{bmatrix}\mathbf{Q}^\mathsf{T} & \mathbf{R}^\mathsf{T}\end{bmatrix}$ *with some* $\mathbf{R} \in \mathbb{Z}^{(m-n)\times m}$.
4) The determinants of all $n \times n$ *minors of* $\mathbf{Q}$ *are coprime.*

Definition 4 (Smith normal form and elementary divisors)**.** *For any integer matrix* $\mathbf{Q} \in \mathbb{Z}^{m\times n}$*, there exists a non-negative integer diagonal matrix* $\mathbf{\Sigma}(\mathbf{Q}) = \operatorname{diag}(\sigma_1(\mathbf{Q}), \sigma_2(\mathbf{Q}), \ldots, \sigma_r(\mathbf{Q}))$ *with* $r = \operatorname{rank}(\mathbf{Q})$ *such that*[3] $\sigma_1(\mathbf{Q}) \mid \sigma_2(\mathbf{Q}) \mid \ldots \mid \sigma_r(\mathbf{Q})$ *and*

$$\mathbf{Q} = \tilde{\mathbf{S}}(\mathbf{Q}) \begin{bmatrix}\mathbf{\Sigma}(\mathbf{Q}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0}\end{bmatrix} \tilde{\mathbf{T}}(\mathbf{Q}) \tag{7}$$

for some unimodular matrices $\tilde{\mathbf{S}}(\mathbf{Q}) \in \mathbb{Z}^{m\times m}$ *and* $\tilde{\mathbf{T}}(\mathbf{Q}) \in \mathbb{Z}^{n\times n}$*. Equivalently, there exists a left-invertible* $\mathbf{S}(\mathbf{Q}) \in \mathbb{Z}^{m\times r}$ *and a right-invertible* $\mathbf{T}(\mathbf{Q}) \in \mathbb{Z}^{r\times n}$ *such that* $\mathbf{Q} = \mathbf{S}(\mathbf{Q})\mathbf{\Sigma}(\mathbf{Q})\mathbf{T}(\mathbf{Q})$*. Here,* $\mathbf{\Sigma}(\mathbf{Q})$ *is called the* reduced Smith normal form *of* $\mathbf{Q}$ *and its diagonal entries* $\sigma_i(\mathbf{Q})$ *are called* elementary divisors.

Note that for a square full-rank $\mathbf{Q}$, we have $\det(\mathbf{\Sigma}(\mathbf{Q})) = \prod_{i=1}^r \sigma_i(\mathbf{Q}) = |\det(\mathbf{Q})|$. In a certain sense, for integer matrices, $\det(\mathbf{\Sigma}(\mathbf{Q}))$ may be interpreted as a generalization of the determinant to rectangular matrices. The reader is referred to [15] for additional details on Smith normal forms.

Lemma 2. *For an integer matrix* $\mathbf{Q} \in \mathbb{Z}^{m\times n}$ *and an absolutely continuous random vector* $\boldsymbol{u} \in \mathbb{R}^n$ *with finite differential entropies* $h([\boldsymbol{u}]_\mathcal{I})$ *for all index sets* $\mathcal{I} \subset [n]$ *and finite* $H(\lfloor \boldsymbol{u} \rfloor)$*, the algebraic information dimension* $d_\mathbf{Q}(\boldsymbol{u})$ *and algebraic entropy* $\mathcal{H}_\mathbf{Q}(\boldsymbol{u})$ *with parameter* $\mathbf{Q}$ *are well defined and given respectively by*

$$d_\mathbf{Q}(\boldsymbol{u}) = \operatorname{rank}(\mathbf{Q}) \qquad \mathcal{H}_\mathbf{Q}(\boldsymbol{u}) = h(\mathbf{T}(\mathbf{Q})\boldsymbol{u}). \tag{8}$$

If $\mathbf{Q}$ has full row rank, $\mathcal{H}_\mathbf{Q}(\boldsymbol{u})$ can be expressed in terms of its Smith normal form $\mathbf{\Sigma}(\mathbf{Q})$ rather than $\mathbf{T}(\mathbf{Q})$, namely,

$$\mathcal{H}_\mathbf{Q}(\boldsymbol{u}) = h(\mathbf{Q}\boldsymbol{u}) - \log\det(\mathbf{\Sigma}(\mathbf{Q})). \tag{9}$$

If $\mathbf{Q}$ is also right-invertible, then $\mathbf{Q} = \mathbf{T}(\mathbf{Q})$ (up to row permutations and scaling rows with -1), so (8) further simplifies to $\mathcal{H}_\mathbf{Q}(\boldsymbol{u}) = h(\mathbf{Q}\boldsymbol{u})$.

Lemma 2 builds upon a generalization of a key result due to Makkuva and Wu [16, Lem. 1], who show that for independent, absolutely continuous variables U_k, $k \in [K]$ with finite differential entropies $h(U_k)$ and a single-row matrix $\mathbf{Q} \in \mathbb{Z}^{1\times K}$ composed of coprime coefficients q_k, $k \in [K]$,

$$\lim_{\nu\to\infty} \Big\{ H\big(\textstyle\sum_k q_k \lfloor \nu U_k \rfloor\big) - H\big(\big\lfloor \textstyle\sum_k q_k \nu U_k \big\rfloor\big)\Big\} = 0. \tag{10}$$

Since the definition of algebraic entropy is based on the limit [cf. (5b)]

$$\mathcal{H}_\mathbf{Q}(\boldsymbol{u}) = \lim_{\nu\to\infty} \Big\{ H\big(\textstyle\sum_k q_k \lfloor \nu U_k \rfloor\big) - \log(\nu)\Big\} \tag{11}$$

we see that if the q_k are not coprime, they can be divided by their greatest common divisor (gcd) to enforce coprimality (i.e., right-invertibility of $\mathbf{Q}$), after which the asymptotic equality (10) allows one to switch the order of the integer part operation $\lfloor\cdot\rfloor$ and the weighted summation, in the limit as $\nu \to \infty$. The operation $\mathbf{Q} \mapsto \mathbf{T}(\mathbf{Q})$ that appears in (8) can be interpreted as a generalization of this gcd-reduction, i.e., a transition from $\mathbf{Q}$ to a right-invertible matrix $\mathbf{T}(\mathbf{Q})$, for the case where $\mathbf{Q}$ has more than one row.

B. Matroids

Given a collection of vectors (e.g., the columns of a matrix), the associated matroid can be viewed as a full description of the linear dependence relations between subsets of vectors. In the following, we give an axiomatic definition of matroids.

Definition 5 (Matroids)**.** *A matroid* M *is a pair* $(E, \mathcal{I})$ *consisting of a finite set* E *and a collection of subsets* $\mathcal{I} \subset 2^E$ *satisfying the properties [17, Sec. 1.1]:*

1) $\emptyset \in \mathcal{I}$
2) If $I \in \mathcal{I}$ *and* $I' \subset I$*, then* $I' \in \mathcal{I}$
3) If I_1 *and* I_2 *are in* $\mathcal{I}$ *and* $|I_1| < |I_2|$*, then there exists an element* $J \in I_2 \setminus I_1$ *such that* $I_1 \cap J \in \mathcal{I}$.

We say that $B \in \mathcal{I}$ *is a* basis *of* $M = (E, \mathcal{I})$ *if there is no larger* $B' \in \mathcal{I}$ *that contains* B*. In other words, a basis is a maximal independent set of the matroid. All bases have the same cardinality [17, Lem. 1.2.1, 1.2.4] and a matroid is uniquely defined by the collection of its bases, which we will generally denote as* $\mathscr{B}(M)$ *[17, Lem. 1.2.2, Thm. 1.2.3].*

Definition 6 (Representable matroids)**.** *If* E *denotes the set of column labels of a matrix* $\mathbf{Q} \in \mathbb{A}^{m\times n}$ *over a ring or field* $\mathbb{A}$*, and if* $\mathcal{I}$ *denotes the set of subsets of* E *such that for every* $I \in \mathcal{I}$*, the rows of* $[\mathbf{Q}^\mathsf{T}]_I$ *are linearly independent (in the vector space* $\mathbb{A}^m$*), then* $(E, \mathcal{I})$ *is a matroid, called the* vector matroid *of* $\mathbf{Q}$ *(cf. [17, Proposition 1.1.1]) and is denoted as* $M(\mathbf{Q})$.

If a matroid M *is isomorphic to the vector matroid of some matrix* $\mathbf{Q}$ *over some ring* $\mathbb{A}$*, then we say that* M *is* representable *over* $\mathbb{A}$*. Accordingly,* $\mathbf{Q}$ *is a* representation *for* M *over* $\mathbb{A}$.

We define $\mathscr{M}_\mathbb{A}(n)$ *as the set of representable matroids of size* n *that are representable over* $\mathbb{A}$*, and* $\mathscr{C}_\mathbb{A}(M)$ *shall denote the set of matrix representations over* $\mathbb{A}$ *of the matroid* M.

Definition 7 (Dual matroids)**.** *Let* $M = (E, \mathcal{I})$ *be a matroid and* $\mathscr{B}(M)$ *the collection of its bases. Then* $\{E \setminus B \colon B \in \mathscr{B}(M)\}$ *is the set of bases of a matroid on* E*, called the* dual *of* M*, and denoted as* M^* *(cf. [17, Thm. 2.1.1]).*

V. A General Formula for Compute–Forward Achievable Rates

In the following, $\mathbb{U}$ denotes a ring and $\mathbb{A} \subseteq \mathbb{U}$ denotes a discrete additive subgroup of $\mathbb{U}$, that is, a *lattice* over $\mathbb{U}$. The three main compute–forward theorems presented further below are concerned with the following three choices of $(\mathbb{U}, \mathbb{A})$, respectively:

- Theorem 1: $(\mathbb{U}, \mathbb{A}) = (\mathbb{F}_\mathsf{q}, \mathbb{F}_\mathsf{q})$
- Theorem 2: $(\mathbb{U}, \mathbb{A}) = (\mathbb{Z}, \mathbb{Z})$
- Theorem 3: $(\mathbb{U}, \mathbb{A}) = (\mathbb{R}, \mathbb{Z})$.

Let $(\boldsymbol{u}, Y) \in \mathbb{U}^K \times \mathcal{Y}$ follow a joint distribution $P_{\boldsymbol{u},Y} = \prod_{k=1}^K P_{U_k} P_{Y|\boldsymbol{u}}$. In the following, for some natural numbers $1 \le L_\mathsf{B} \le K$,

- $\mathbf{B}$ denotes a full row-rank matrix over $\mathbb{A}$ of size $L_\mathsf{B} \times K$;
- M denotes a matroid of size L_B;
- $\mathcal{T}$ denotes a subset of $[K]$.

[3] $a \mid b$ means a divides b

With these notations in mind, we first define the set

$$\mathscr{Q}(\mathbf{B},M,\mathcal{T}) \triangleq \Big\{(R_1,\ldots,R_K) \in \mathbb{R}_+^K\colon \sum_{k\in\mathcal{T}} R_k < \mathcal{H}([\boldsymbol{u}]_{\mathcal{T}}) - \mathcal{H}_{\mathbf{B}}(\boldsymbol{u}|Y) + J(\mathbf{B},M)\Big\} \tag{12}$$

where $J(\mathbf{B},M)$ denotes a min-entropy term defined as

$$J(\mathbf{B},M) \triangleq \inf_{\mathbf{C}\in\mathscr{C}_{\mathbb{A}}(M)} \mathcal{H}_{\mathbf{CB}}(\boldsymbol{u}|Y). \tag{13}$$

With these definitions settled, we define the set

$$\mathscr{Q}(\mathbf{B}) = \bigcap_{M} \bigcup_{\mathcal{S}} \bigcap_{\mathcal{T}} \mathscr{Q}(\mathbf{B},M,\mathcal{T}) \tag{14}$$

where the three nested set operations are over triples $(M,\mathcal{S},\mathcal{T})$ meeting the following constraints:

1) M iterates over all matroids of size L_{B} except the full-rank matroid, i.e., $M \in \mathscr{M}_{\mathbb{A}}(L_{\mathsf{B}}) \setminus ([L_{\mathsf{B}}], 2^{[L_{\mathsf{B}}]})$. Henceforth, we shall denote this set of matroids (excluding the full-rank matroid) as $\mathscr{M}_{\mathbb{A}}^{\circ}(L_{\mathsf{B}})$;
2) $\mathcal{S}$ iterates over all index sets that correspond to bases of the dual matroid M^*, i.e., $\mathcal{S} \in \mathscr{B}(M^*)$;
3) $\mathcal{T}$ iterates over all index sets that correspond to bases of the matroid of which $[\mathbf{B}]_{\mathcal{S}}$ is a representation, i.e., $\mathcal{T} \in \mathscr{B}(M([\mathbf{B}]_{\mathcal{S}}))$.

Finally, let us define the so-called *joint decoding* rate region

$$\mathscr{R}(\mathbf{A}) \triangleq \bigcup_{\mathbf{B}} \mathscr{Q}(\mathbf{B}) \tag{15}$$

where $\mathbf{B} \in \mathbb{A}^{L_{\mathsf{B}}\times K}$, $L_{\mathsf{B}} = \operatorname{rank}(\mathbf{A}),\ldots,K$ runs over all full row-rank matrices satisfying $\Lambda_{\mathbb{A}}(\mathbf{B}) \supseteq \Lambda_{\mathbb{A}}(\mathbf{A})$. The following theorems will provide an operational meaning to this rate region.

Theorem 1 (Finite-field compute–forward)**.** *Let* $(\mathbb{U},\mathbb{A}) = (\mathbb{F}_{\mathsf{q}},\mathbb{F}_{\mathsf{q}})$ *for some prime field size* q*. A rate tuple* $(R_1,\ldots,R_K)$ *is achievable for decoding the* $\mathbf{A}$*-linear combinations of codewords if it is contained in* $\mathscr{R}(\mathbf{A})$*.*

Theorem 2 (Integer compute–forward)**.** *Let* $(\mathbb{U},\mathbb{A}) = (\mathbb{Z},\mathbb{Z})$ *and assume that* $\mathcal{H}(\boldsymbol{u})$ *is finite. A tuple* $(R_1,\ldots,R_K)$ *is achievable for decoding the* $\mathbf{A}$*-linear combinations of codewords if it is contained in* $\mathscr{R}(\mathbf{A})$*.*

Theorem 3 (Continuous compute–forward)**.** *Let* $(\mathbb{U},\mathbb{A}) = (\mathbb{R},\mathbb{Z})$*. Assume that the vector of auxiliaries* $\boldsymbol{u} \in \mathbb{R}^K$ *has an absolutely continuous distribution as well as finite entropies* $h(\boldsymbol{u})$ *and* $H(\lfloor \boldsymbol{u} \rfloor)$*. In addition, we have either of the two (mutually exclusive) situations:*

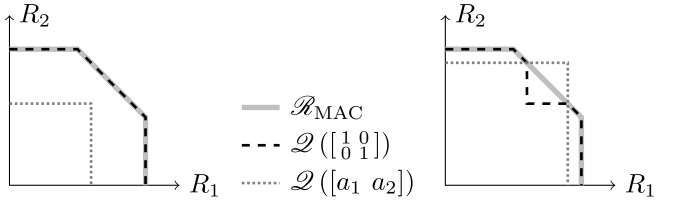
1) the mappings $x_k(u_k)$ *have finitely many (jump) discontinuities and images* $x_k(\mathbb{R})$ *of finite cardinality;*
2) the mappings $x_k(u_k) = \beta_k u_k$ *are linear with* β_k *some real-valued coefficients, and the system equation is given by* $Y = \sum_k h_k x_k + Z$ *with independent noise* Z*.*

Under these assumptions, a tuple $(R_1,\ldots,R_K)$ *is achievable for decoding the* $\mathbf{A}$*-linear combinations of codewords if it is contained in* $\mathscr{R}(\mathbf{A})$*.*

For Theorems 1 and 2, by Lemma 1 the rate region $\mathscr{R}(\mathbf{A})$ is expressible in terms of discrete entropies since we obtain $\mathcal{H}([\boldsymbol{u}]_{\mathcal{T}}) = H([\boldsymbol{u}]_{\mathcal{T}})$, $\mathcal{H}_{\mathbf{B}}(\boldsymbol{u}|Y) = H(\mathbf{B}\boldsymbol{u}|Y)$ and $\mathcal{H}_{\mathbf{CB}}(\boldsymbol{u}|Y) = H(\mathbf{CB}\boldsymbol{u}|Y)$, whereas for Theorem 3, by Lemma 2 the corresponding algebraic entropies evaluate to differential entropies $h([\boldsymbol{u}]_{\mathcal{T}})$, $h(\mathbf{T}(\mathbf{B})\boldsymbol{u}|Y)$ and $h(\mathbf{T}(\mathbf{CB})\boldsymbol{u}|Y)$.

VI. The two-user case

To gain some insight into the rate regions described by (14) and (15), let us consider a two-user channel ($K = 2$).



(a) $\mathscr{Q}([a_1\ a_2]) \subset \mathscr{R}_{\text{MAC}}$ (b) $\mathscr{Q}([a_1\ a_2]) \not\subset \mathscr{R}_{\text{MAC}}$

Fig. 2. Partial rate regions $\mathscr{Q}(\mathbf{B})$ for $\mathbf{B} = \left[\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right]$ and $\mathbf{B} = [a_1\ a_2]$ with $a_1, a_2 \neq 0$, that can be combined by union-taking [cf. (15)] to obtain the compute–forward rate region $\mathscr{R}(\mathbf{A})$. The multiple-access rate region $\mathscr{R}_{\text{MAC}}$, which is achievable with (unstructured) random codes, is shown for reference.

Figure 2 exhibits the different subregions that compose the compute–forward rate region $\mathscr{R}(\mathbf{A})$. The set $\mathscr{R}_{\text{MAC}} = \{(R_1,R_2) \in \mathbb{R}_+^2\colon\ R_1 < I(U_1;Y,U_2),\ R_2 < I(U_2;Y,U_1),\ R_1+R_2 < I(U_1,U_2;Y)\}$ denotes the conventional multiple-access rate region.

For $K = 2$ one can prove the following two simplifications of the union in (15): for a single linear combination ($L = 1$) specified by a coefficient vector $\mathbf{A} = [a_1\ a_2] \in \mathbb{A}^{1\times 2}$ (with $a_1, a_2 \neq 0$), we have

$$\mathscr{R}(\mathbf{A}) = \mathscr{R}_{\text{MAC}} \cup \mathscr{Q}([a_1\ a_2]) \tag{16}$$

whereas for the case of two independent linear combinations ($L = 2$) specified by a full-rank matrix $\mathbf{A} = \left[\begin{smallmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{smallmatrix}\right] \in \mathbb{A}^{2\times 2}$, we have[4]

$$\mathscr{R}(\mathbf{A}) = \mathscr{Q}\left(\left[\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right]\right). \tag{17}$$

Comparing Figures 2a and 2b, we see that in the former, a case is depicted where the rate region (16) (for $L = 1$) is contained entirely in $\mathscr{R}_{\text{MAC}}$ (in which case $\mathscr{R}([a_1\ a_2])$ reduces to $\mathscr{R}_{\text{MAC}}$), whereas in the latter, the opposite case is shown.

In the former case, i.e., if $\mathscr{Q}([a_1\ a_2]) \subset \mathscr{R}_{\text{MAC}}$ holds for all $[a_1\ a_2]$ with non-zero entries, the rate region (17) (for $L = 2$) coincides with $\mathscr{R}_{\text{MAC}}$. In this circumstance, one can achieve the full MAC capacity rate region with nested linear codes, rather than random codes. Specifically, this occurs when for all $a_1, a_2 \neq 0$,

$$\mathcal{H}_{[a_1\ a_2]}(U_1,U_2|Y) \geq \frac{1}{2}\mathcal{H}(U_1,U_2|Y). \tag{18}$$

Otherwise, we have a situation like the one depicted in Figure 2b, in which $\mathscr{Q}([a_1\ a_2])$ protrudes out of the MAC rate region, while the rate region for $L = 2$, as given in (17), is *smaller* than the MAC rate region. Notice that the subset of $\mathscr{Q}([a_1\ a_2])$ lying outside of $\mathscr{R}_{\text{MAC}}$ and the subset of $\mathscr{R}_{\text{MAC}}$ lying outside of $\mathscr{Q}\left(\left[\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right]\right)$, are two triangular-shaped sets that are mirror images of each other (about the dominant face of the MAC rate region). This indicates a tension between random codes and nested linear codes, in that the latter excel at computing rank-deficient codeword combinations, but are less efficient than random codes when it comes to computing a full-rank set of linear combinations (i.e., recovering all messages).

[4] Equation (17) holds because, as one can show, $\mathscr{Q}\left(\left[\begin{smallmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{smallmatrix}\right]\right)$ is contained in $\mathscr{Q}\left(\left[\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right]\right)$ for any full-rank $\left[\begin{smallmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{smallmatrix}\right]$.

Note that the general aspect of the rate region depicted in Figure 2 is applicable to Theorems 1, 2 and 3 alike. We refer the reader to our previous paper [6], where the two-user case is discussed at length, and [7], which also illustrates the three-user case.

VII. Gaussian channels

Arguably the most important special case of our main results is Theorem 3 evaluated for Gaussian distributions. Consider a multiple access channel with M receiver antennas which obeys the system equation

$$\boldsymbol{y} = \mathbf{H}\boldsymbol{x} + \boldsymbol{z} \tag{19}$$

where $\boldsymbol{x} = \big[X_1, \ldots, X_K\big]^\mathsf{T}$ represents the vector or channel inputs, $\mathbf{H} \in \mathbb{R}^{M \times K}$ stands for the channel gain matrix, $\boldsymbol{y} \in \mathbb{R}^M$ is the vector of channel outputs and $\boldsymbol{z} \sim \mathcal{N}_\mathbb{R}(\mathbf{0}, \mathbf{I})$ is i.i.d. additive Gaussian noise. We assume that the channel inputs are subject to average power constraints $\mathsf{E}[X_k^2] \leq P_k$, $k = 1, \ldots, K$. We define $\mathbf{P} = \operatorname{diag}(P_1, \ldots, P_K)$ as the covariance matrix of $\boldsymbol{x}$.

The following corollary to Theorem 3 provides a generalization of the compute–forward rate region from [4] for the Gaussian channel (19), to the effect of simultaneously computing multiple linearly independent combinations of Gaussian codewords, rather than only one linear combination. This improves on the best-known rate region from prior work, which was based on nested lattice encoding and sequential decoding [5].

Corollary 1 (Gaussian compute–forward)**.** *Let* $(\mathbb{U}, \mathbb{A}) = (\mathbb{R}, \mathbb{Z})$*. We evaluate Theorem 3 for the Gaussian channel* (19)*, auxiliary variables* $U_k \sim \mathcal{N}_\mathbb{R}(0, \beta_k^2 P_k)$ *with scaling parameters* $\beta_k > 0$ *and modulation mappings* $X_k = U_k/\beta_k$ *to satisfy the power constraints* $\mathsf{E}[X_k^2] = P_k$*. Then,* $\mathscr{Q}(\mathbf{B}, M, \mathcal{T})$ *in Theorem 3 specializes to the set of rate tuples* $(R_1, \ldots, R_K)$ *such that*

$$\sum_{k \in \mathcal{T}} R_k < \frac{1}{2} \sum_{k \in \mathcal{T}} \log(\beta_k^2 P_k) - \frac{1}{2} \log \frac{\det\big(\mathbf{BKB}^\mathsf{T}\big)}{\det(\boldsymbol{\Sigma}(\mathbf{B}))} + J(\mathbf{B}, M) \tag{20}$$

where $\mathbf{K}$ *stands for the conditional covariance matrix*

$$\mathbf{K} = \mathsf{E}\big[\boldsymbol{u}\boldsymbol{u}^\mathsf{T} \big| Y\big] = \operatorname{diag}(\boldsymbol{\beta}) \left(\mathbf{P}^{-1} + \mathbf{H}^\mathsf{T}\mathbf{H}\right)^{-1} \operatorname{diag}(\boldsymbol{\beta}) \tag{21}$$

and where

$$J(\mathbf{B}, M) = \inf_{\mathbf{C} \in \mathscr{C}_\mathbb{Z}(M)} \frac{1}{2} \log \frac{\det\big(\mathbf{CBKB}^\mathsf{T}\mathbf{C}^\mathsf{T}\big)}{\det(\boldsymbol{\Sigma}(\mathbf{CB}))}. \tag{22}$$

Note that Corollary 1 is obtained directly from the Gaussian entropy formula and (9). One can easily show that in the infimum (22), the matrix $\mathbf{C}$ can be restricted to being right-invertible. If, in addition, $\mathbf{B}$ is also right-invertible, then both $\mathbf{B}$ and $\mathbf{CB}$ have elementary divisors all equal to one, hence the denominators on the right-hand sides of (20) and (22) disappear, thus simplifying expressions. Note, however, that we have no proof nor disproof[5] that in the union over matrices $\mathbf{B}$ in (15), $\mathbf{B}$ can be restricted to being right-invertible without loss of optimality. This stands as an open problem.

Figure 3 shows a three-dimensional rate region obtained from evaluating Corollary 1 for computation of a single combination $\mathbf{A} = [1\ 1\ 1]$ over a symmetric Gaussian channel with a single receive antenna ($M = 1$).

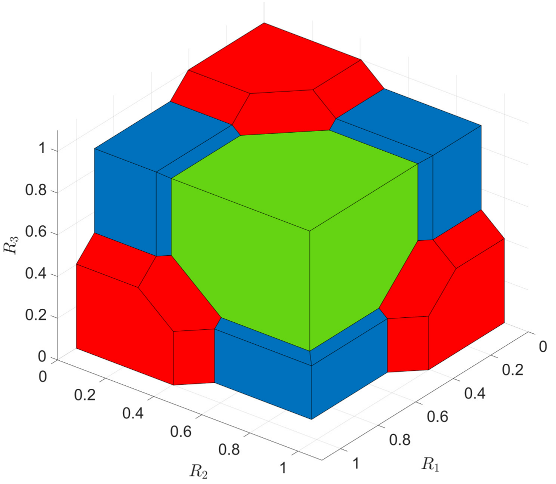


Fig. 3. Three-user joint decoding rate region computed numerically for a Gaussian channel.

Finally, let us particularize (20) to the case of two users ($K = 2$) and a single linear combination ($L = 1$), setting all parameters β_k to one and equal power constraints $P_1 = P_2 = P$, and evaluating the union (15) only for $\mathbf{B} = [a_1\ a_2]$. Denote the latter row vector as $\mathbf{a}^\mathsf{T}$ and the channel matrix $\mathbf{H} \in \mathbb{R}^{1 \times 2}$ as the row vector $\mathbf{h}^\mathsf{T}$. With these choices, we can readily recover the well-known compute–forward rectangular rate region due to Nazer and Gastpar [4]:

$$\begin{aligned} &\max\{R_1, R_2\} \\ &\quad < \frac{1}{2} \log \left(\frac{P}{\mathbf{a}^\mathsf{T} \left(P^{-1}\mathbf{I} + \mathbf{h}\mathbf{h}^\mathsf{T}\right)^{-1} \mathbf{a}} \right) + \log \gcd\left(|a_1|, |a_2|\right). \end{aligned} \tag{23}$$

We observe how in this special case, the term $\boldsymbol{\Sigma}(\mathbf{B})$ from (20), which here reduces to a scalar, evaluates to the greatest common divisor of $|a_2|$ and $|a_2|$. Also note that for $L = 1$, the term $J(\mathbf{B}, M)$ disappears, which further contributes to simplification. In fact, much of the difficulties in proving Theorems 2, 3 (Corollary 1) can be traced to the term $J(\mathbf{B}, M)$, and stem from the exchange of quantization limit and infimum (13).

VIII. Conclusion

In this paper, we proposed a discretization approach to translate compute–forward achievability results from discrete to continuous-valued channels in a unified manner. Although we focused on the compute–forward problem, this discretization approach could also be used to extend other algebraic achievability results to continuous-valued channels.

Acknowledgment

This work was partly funded by the Catalan Government under Grant 2017 SGR 1479 and in part by the Spanish Ministry of Economy and Competitiveness under Project RTI2018-099722-B-I00 (ARISTIDES).

[5] Except for the two-user case $K = 2$, for which we can show that right-invertible matrices $\mathbf{B} = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ and $\mathbf{B} = [b_1\ b_2]$ (with coprime b_1 and b_2) suffice to attain the entire rate region $\mathscr{R}(\mathbf{A})$. See also Section VI.

References

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

[2] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge: Cambridge University Press, 2011.

[3] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless network information flow: A deterministic approach," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 1872–1905, 2011.

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

[5] 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 Transactions on Information Theory*, vol. 62, no. 9, pp. 4879–4909, 2016.

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

[7] S. H. Lim, C. Feng, A. Pastore, B. Nazer, and M. Gastpar, "Compute-forward for DMCs: Simultaneous decoding of multiple combinations," *IEEE Transactions on Information Theory*, vol. 66, no. 10, pp. 6242–6255, 2020.

[8] A. Rényi, "On the dimension and entropy of probability distributions," *Acta Mathematica Academiae Scientiarum Hungarica*, vol. 10, no. 1, pp. 193–215, Mar 1959.

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

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

[11] ——, "Achievable rate region for three user discrete broadcast channel based on coset codes," *IEEE Transactions on Information Theory*, vol. 64, no. 4, pp. 2267–2297, Apr 2018.

[12] P. Sen and Y.-H. Kim, "Homologous codes for multiple access channels," *IEEE Transactions on Information Theory*, vol. 66, no. 3, pp. 1549–1571, Mar 2019.

[13] P. Sen, S. H. Lim, and Y.-H. Kim, "On the optimal achievable rates for linear computation with random homologous codes," *IEEE Transactions on Information Theory*, vol. 66, no. 10, pp. 6200–6221, Oct 2020.

[14] F. Shirani and S. S. Pradhan, "Lattices from linear codes and fine quantization: General continuous sources and channels," in *IEEE International Symposium on Information Theory (ISIT)*, 2018, pp. 2356–2360.

[15] W. C. Brown, *Matrices over Commutative Rings*. New Year: Marcel Dekker, Inc, 1993.

[16] A. V. Makkuva and Y. Wu, "Equivalence of additive-combinatorial linear inequalities for shannon entropy and differential entropy," *IEEE Transactions on Information Theory*, vol. 64, no. 5, pp. 3579–3589, 2018.

[17] J. G. Oxley, *Matroid Theory (Oxford Graduate Texts in Mathematics)*, 2nd ed. USA: Oxford University Press, 2011.