Distributed lossy computation with structured codes: from discrete to continuous sources

Distributed Lossy Computation with Structured Codes: From Discrete to Continuous Sources

Adriano Pastore*, Sung Hoon Lim†, Chen Feng‡, Bobak Nazer§, Michael Gastpar¶
*ISPIC Research Unit, 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*—This paper considers the problem of distributed lossy compression where the goal is to recover one or more linear combinations of the sources at the decoder, subject to distortion constraints. For certain configurations, it is known that codes with algebraic structure can outperform i.i.d. codebooks. For the special case of finite-alphabet sources, recent work has demonstrated how to incorporate joint typicality decoding alongside linear encoding and binning. This work takes a discretization approach to extend this rate region to include both integer- and real-valued sources. As a case study, the rate region is evaluated for the Gaussian case. The resulting joint-typicality-based rate region recovers and generalizes the best-known rate region for this scenario, based on lattice encoding and sequential decoding.**

***Index Terms*—Nested linear codes; distributed source coding; function computation; entropy; algebraic network information theory.**

I. Introduction

Consider the network information theory problem in which multiple sources are encoded in a distributed manner into rate-limited representations, and a receiver attempts to simultaneously (or sequentially) recover multiple linear combinations of these sources from said representations in a reliable manner, while satisfying a distortion constraint for each desired linear combination.

One of the most elementary special cases of this computation problem was studied by Körner and Marton [1], who contemplated a receiver that losslessly recovers the sum of two distributedly encoded binary symmetric sources. Their work elucidated that in some cases like the one at hand, random i.i.d. codes, are strictly outperformed by random linear (i.e., structured) codes in terms of the achievable rate region.

Following this pioneering work, numerous advances have made inroads towards a more complete understanding of the potential of linear and lattice coding strategies for network source and channel coding. The special setting where all sources are recovered losslessly (without taking linear combinations) is the celebrated result by Slepian and Wolf [2]. The corner points of the Slepian–Wolf rate region can be achieved using nested linear codes and sequential decoding [3], [4]. See [5] for a detailed discussion of structured codes for multiterminal binning. Subsequent work by Lalitha *et al.* [6] proposed general nested linear coding strategies for an arbitrary number of users and linear combinations, and proved their sum-rate optimality in some special cases.

Turning to the more general case of lossy recovery, work has progressed along separate threads: lattice codes for Gaussian sources and linear codes for discrete sources. The study of nested lattice codes for multiterminal Gaussian source coding was initiated in the context of the Wyner–Ziv problem, see [5] for a thorough account. Krithivasan and Pradhan [7] developed a nested lattice framework for distributed compression of linear functions as well as a Gaussian analogue of the Körner–Marton problem. This work demonstrated that lattice coding can sometimes outperform i.i.d. coding, here exemplified by the Berger–Tung scheme [8], [9], which consists of distributed vector quantization and binning, followed by joint typicality decoding. The lattice framework was subsequently refined and generalized in [10]–[12], but has remained limited to single-user and sequential decoding, owing to the challenges of directly analyzing simultaneous decoding for lattices.

For discrete sources, Krithivasan and Pradhan proposed a framework based on Abelian group codes and joint typicality encoding/decoding [13]. This paper can be viewed as the first in a line of works on *algebraic network information theory* [14]–[20]. These papers extend well-known technical lemmata (packing lemma, covering lemma, etc.) and powerful proof tools such as joint typicality [21], [22] from the realm of

A.P. was supported by Grant PID2021-128373OB-I00 funded by MCIN/AEI/10.13039/501100011033 and by "ERDF A way of making Europe" and in part by Grant 2021 SGR 00772 funded by the Univ. and Research Dpt. from Generalitat de Catalunya, as well as by the "Ministerio de Asuntos Económicos y Transformación Digital" with NextGenerationEU MRR/PRTR funds TSI-063000-2021-39/40/41. S.H.L. was supported by the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology under Grant NRF-2020R1F1A1074926. C.F. was supported by NSERC Discovery RGPIN-2016-05310. B.N. was supported by the U.S. National Science Foundation under Grant CCF-1955981. M.G. was supported by the Swiss National Science Foundation under Grant 200364.

random i.i.d. codes to the ensemble of random linear codes. The overall goal is to create a unified framework for structured coding that encompasses both discrete and continuous sources/channels, while enabling powerful techniques via joint typicality, such as multicoding and simultaneous decoding.

This paper takes another step towards such a unified framework, by proposing an achievable region for distributed lossy computation via linear codes for both discrete and continuous sources. Specifically, we apply the discretization technique from [23] to the simultaneous-decoding rate region for discrete sources proposed in [24]. Prior work [23] followed this path to obtain a simultaneous-decoding achievable region for the "compute–forward" problem for both discrete and continuous channels, starting from the rate region derived for the discrete case in [18]. Recall that, in the compute–forward problem, a multiple-access receiver is tasked with recovering one or more linear combinations of the codewords sent by the transmitters [25].

Discretization is a common approach (see, e.g., [22]) for extending rate regions initially derived for discrete channels to continuous channels. Note, however, that due to the complex form of the achievable rate regions in this setting, the computation of limits is far more involved than in typical textbook cases. In fact, the limit arguments require a more sophisticated treatment and lead to a generalization of Rényi's d-dimensional entropy [26], here termed *algebraic entropy*.

One appealing feature of our framework is that, due to the use of simultaneous decoding, it can better compete with the Berger–Tung strategy. For instance, as we will show in Section VII for the Gaussian Körner–Marton example proposed in [7], our scheme can attain the minimum between the symmetric rate for lattice coding and for Berger–Tung coding. Prior work was not able to attain this minimum directly.

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 $\mathbb{X}$ a ring and some subset $\mathbb{A} \subseteq \mathbb{X}$, we define $\Lambda_{\mathbb{A}}(\mathbf{Q}) = \{\mathbf{Q}\mathbf{v} \colon \mathbf{v} \in \mathbb{A}^d\}$. If $\mathbb{A}$ is a discrete additive subgroup of $\mathbb{X}$, then $\Lambda_{\mathbb{A}}(\mathbf{Q})$ is called a *lattice* generated by $\mathbf{Q}$. If $\mathbb{A} = \mathbb{X}$, it is called the *span* of $\mathbf{Q}$. The finite field of prime size q is denoted as $\mathbb{F}_{\mathsf{q}}$.

III. Problem Statement

Let $\mathbb{X}$ denote a ring[1] and let $\mathcal{Y}$ denote an arbitrary set. Consider a K-user distributed lossy computation system with $K+1$ correlated variables $(X_1, \ldots, X_K, Y) \in \mathbb{X}^K \times \mathcal{Y}$ with joint distribution $P_{X_1,\ldots,X_K,Y}$, where X_k stands for the source variable at encoder k and Y is the receiver side information. We shall assume that these sources are i.i.d., hence $(X_1^n, \ldots, X_K^n, Y^n) \sim P_{X_1,\ldots,X_K,Y}^{\otimes n}$.

Let $\mathbb{A}$ be some subset of $\mathbb{X}$, and let $\mathbf{a}_1, \ldots, \mathbf{a}_L \in \mathbb{A}^K$ denote L coefficient vectors ($L \leq K$) that can be stacked into a coefficient matrix $\mathbf{A} = [A_{\ell,k}] = \begin{bmatrix}\mathbf{a}_1^{\mathsf{T}} & \ldots & \mathbf{a}_L^{\mathsf{T}}\end{bmatrix}^{\mathsf{T}} \in \mathbb{A}^{L\times K}$. The decoder attempts to recover or approximate the L linear combinations $W_\ell^n = \sum_{k=1}^K A_{\ell,k} X_k^n$, $\ell \in [L]$. A $(2^{nR_1}, \ldots, 2^{nR_K}, n)$ distributed lossy computation code consists of

- K encoders, where each encoder $k \in [K]$ assigns an index $m_k(x_k^n) \in [2^{nR_k}]$ to the source sequence $x_k^n \in \mathbb{X}^n$;
- a decoder that produces L estimates $\hat{w}_\ell^n(m_1, \ldots, m_K, y^n) \in \mathbb{X}^n$, $\ell \in [L]$ to each index tuple $(m_1, \ldots, m_K) \in [2^{nR_1}] \times \ldots \times [2^{nR_K}]$ and side information sequence $y^n \in \mathcal{Y}^n$.

A rate tuple $(R_1, \ldots, R_K)$ is said to be achievable for $\mathbf{A}$-computation with distortion functions $(d_1, \ldots, d_L)$ and distortion values $\mathbf{D} = (D_1, \ldots, D_L)$ if there exists a sequence of $(2^{nR_1}, \ldots, 2^{nR_K}, n)$ codes that satisfy

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^n \mathsf{E}\left[d_\ell(W_{\ell,i}, \hat{W}_{\ell,i})\right] \leq D_\ell, \quad \ell \in [L].$$

where the expectation is over the source distribution $P_{X_1,\ldots,X_K,Y}$ and where $W_{\ell,i}$ (resp. $\hat{W}_{\ell,i}$) stands for the i-th entry of W_ℓ^n (resp. $\hat{W}_\ell^n$).

The rate–distortion region $\mathscr{R}^*(\mathbf{A}, \mathbf{D})$ for DLC is the closure of the set of all achievable rate tuples $(R_1, \ldots, R_K)$ for coefficient matrix $\mathbf{A}$ and distortion tuple $\mathbf{D}$.

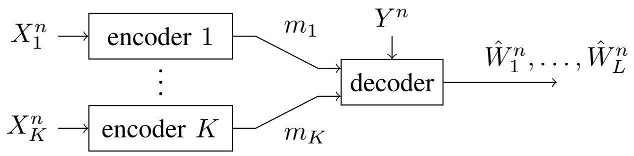


Fig. 1. A DLC system with K transmitters, in which the receiver seeks to decode L linear combinations W_ℓ^n, $\ell \in [L]$ of the source sequences $(X_1^n, \ldots, X_K^n)$.

Remark 1. *The problem specializes to the Körner–Marton problem [1] if $K = 2$, $L = 1$, $\mathbb{X} = \mathbb{A} = \mathbb{F}_2$, $\mathbf{A} = [1,\, 1]$, d_1, d_2 are Hamming distortion measures, $Y = \emptyset$, and $D_1 = D_2 = 0$.*

Remark 2. *The problem specializes to Berger–Tung [8], [27] distributed lossy source coding if $\mathbf{A}$ is the identity and $Y = \emptyset$.*

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 $H(\boldsymbol{u}) = -\sum_{\mathbf{u}\in\mathcal{U}} \mathsf{P}\{\boldsymbol{u} = \mathbf{u}\} \log \mathsf{P}\{\boldsymbol{u} = \mathbf{u}\}$. 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 $h(\boldsymbol{u}) = -\int f_{\boldsymbol{u}}(\mathbf{u}) \log f_{\boldsymbol{u}}(\mathbf{u})\, \mathrm{d}\mathbf{u}$. In his 1959 paper [26], Rényi elucidates some interesting connections between the discrete entropy of quantized variables and differential entropies. In

[1] This paper will address special cases including $\mathbb{X} = \mathbb{F}_{\mathsf{q}}$, $\mathbb{Z}$, or $\mathbb{R}$.

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*

$$d_{\mathbf{Q}}(\boldsymbol{u}) = \lim_{\nu\to\infty} \frac{H(\mathbf{Q}\lfloor \nu\boldsymbol{u}\rfloor)}{\log(\nu)} \tag{1a}$$

$$\mathcal{H}_{\mathbf{Q}}(\boldsymbol{u}) = \lim_{\nu\to\infty} \Big\{ H(\mathbf{Q}\lfloor \nu\boldsymbol{u}\rfloor) - d_{\mathbf{Q}}(\boldsymbol{u})\log(\nu) \Big\} \tag{1b}$$

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*

$$d_{\mathbf{Q}}(\boldsymbol{u}) = 0 \qquad\qquad \mathcal{H}_{\mathbf{Q}}(\boldsymbol{u}) = H(\mathbf{Q}\boldsymbol{u}). \tag{2}$$

By convention, (2) *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 [26]. 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) *$\mathbf{Q}$ 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)**.** *For an integer matrix $\mathbf{Q} \in \mathbb{Z}^{m\times n}$, there exists a non-negative integer diagonal matrix $\mathbf{\Sigma}(\mathbf{Q}) = \mathrm{diag}(\sigma_1(\mathbf{Q}), \sigma_2(\mathbf{Q}), \dots, \sigma_r(\mathbf{Q}))$ with $r = \mathrm{rank}(\mathbf{Q})$ such that $\sigma_1(\mathbf{Q}) \mid \sigma_2(\mathbf{Q}) \mid \dots \mid \sigma_r(\mathbf{Q})$ (where $a \mid b$ means a divides b) and*

$$\mathbf{Q} = \tilde{\mathbf{S}}(\mathbf{Q}) \begin{bmatrix} \mathbf{\Sigma}(\mathbf{Q}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tilde{\mathbf{T}}(\mathbf{Q}) \tag{3}$$

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 [28] 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*

$$d_{\mathbf{Q}}(\boldsymbol{u}) = \mathrm{rank}(\mathbf{Q}) \qquad\qquad \mathcal{H}_{\mathbf{Q}}(\boldsymbol{u}) = h(\mathbf{T}(\mathbf{Q})\boldsymbol{u}). \tag{4}$$

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,

$$\mathcal{H}_{\mathbf{Q}}(\boldsymbol{u}) = h(\mathbf{Q}\boldsymbol{u}) - \log\det(\mathbf{\Sigma}(\mathbf{Q})). \tag{5}$$

If $\mathbf{Q}$ is also right-invertible, then $\mathbf{Q} = \mathbf{T}(\mathbf{Q})$ (up to row permutations and scaling rows with -1), so (4) 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 [29, 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, $k \in [K]$,

$$\lim_{\nu\to\infty} \Big\{ H\big(\textstyle\sum_k q_k\lfloor\nu U_k\rfloor\big) - H\big(\lfloor\sum_k q_k \nu U_k\rfloor\big) \Big\} = 0. \tag{6}$$

Since the definition of algebraic entropy is based on [cf. (1b)]

$$\mathcal{H}_{\mathbf{Q}}(\boldsymbol{u}) = \lim_{\nu\to\infty} \Big\{ H\big(\textstyle\sum_k q_k\lfloor\nu U_k\rfloor\big) - \log(\nu) \Big\} \tag{7}$$

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 (6) 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 (4) 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.

[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 (1a) and (1b) 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.

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 [30, 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 [30, 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 $\mathscr{B}(M)$ [30, 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. [30, 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. [30, Thm. 2.1.1]).*

V. A General Formula for DLC Achievable Rates

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 $[K]$.

Given some random tuple $(\boldsymbol{x}, \boldsymbol{u}, Y) \in \mathbb{X}^K \times \mathbb{X}^K \times \mathcal{Y}$ with joint distribution $P_{\boldsymbol{x},\boldsymbol{u},Y} = P_{\boldsymbol{x},Y} \prod_{k=1}^K P_{U_k|X_k}$, we define

$$\mathscr{Q}(\mathbf{B}, M, \mathcal{T}) \triangleq \Big\{ (R_1, \dots, R_K) \in \mathbb{R}_+^K \colon \sum_{k\in\mathcal{T}} R_k > -\mathcal{H}([\boldsymbol{u}]_\mathcal{T} | [\boldsymbol{x}]_\mathcal{T}) + \mathcal{H}_\mathbf{B}(\boldsymbol{u}|Y) - J(\mathbf{B}, M) \Big\} \quad (8)$$

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

$$J(\mathbf{B}, M) \triangleq \inf_{\mathbf{C}\in\mathscr{C}_{\mathbb{A}}(M)} \mathcal{H}_{\mathbf{CB}}(\boldsymbol{u}|Y). \quad (9)$$

We now define the set

$$\mathscr{Q}(\mathbf{B}) = \bigcap_M \bigcup_\mathcal{S} \bigcap_\mathcal{T} \mathscr{Q}(\mathbf{B}, M, \mathcal{T}) \quad (10)$$

where the triples $(M, \mathcal{S}, \mathcal{T})$ satisfy the following:

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 $\mathscr{M}^\circ_{\mathbb{A}}(L_\mathsf{B})$;
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

$$\mathscr{R}(\mathbf{A}) \triangleq \bigcup_\mathbf{B} \mathscr{Q}(\mathbf{B}) \quad (11)$$

where $\mathbf{B} \in \mathbb{A}^{L_\mathsf{B}\times K}$, $L_\mathsf{B} = \mathrm{rank}(\mathbf{A}), \dots, K$ runs over all full row-rank matrices satisfying $\Lambda_{\mathbb{A}}(\mathbf{B}) \supseteq \Lambda_{\mathbb{A}}(\mathbf{A})$. Let us further define the auxiliary variables $W_\ell = \mathbf{a}_\ell^\mathsf{T} \boldsymbol{x}$ and $\hat{W}_\ell = \mathbf{a}_\ell^\mathsf{T} \boldsymbol{u}$.

The following three theorems, presented in a unified way for three relevant choices of alphabets $(\mathbb{X}, \mathbb{A})$, are the main result of our paper and will endow the rate region expression (11) with operational meaning.

Theorem 1 (Finite-field DLC)**.** *Let $(\mathbb{X}, \mathbb{A}) = (\mathbb{F}_\mathsf{q}, \mathbb{F}_\mathsf{q})$ for some prime field size q. A rate tuple $(R_1, \dots, R_K)$ is achievable for recovering the $\mathbf{A}$-linear combinations of codewords under the distortion tuple $(D_1, \dots, D_L)$ if for some choice of test channels $P_{U_k|X_k}$, $k \in [K]$ such that*

$$\mathsf{E}\big[d_\ell(W_\ell, \hat{W}_\ell)\big] \le D_\ell, \quad \ell \in [L], \quad (12)$$

it is contained in $\mathscr{R}(\mathbf{A})$.

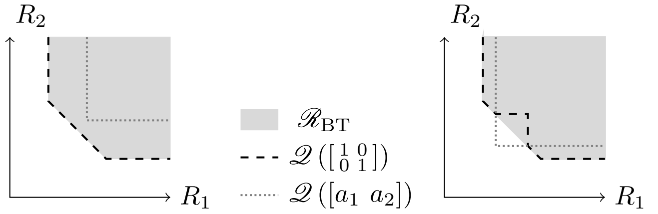
Theorem 2 (Integer DLC)**.** *Let $(\mathbb{X}, \mathbb{A}) = (\mathbb{Z}, \mathbb{Z})$ and assume that $\mathcal{H}(\boldsymbol{u})$ is finite. A rate tuple $(R_1, \dots, R_K)$ is achievable for recovering the $\mathbf{A}$-linear combinations of codewords under the distortion tuple $(D_1, \dots, D_L)$ if for some choice of test channels $P_{U_k|X_k}$, $k \in [K]$ such that* (12) *holds, it is contained in $\mathscr{R}(\mathbf{A})$.*

Theorem 3 (Continuous DLC)**.** *Let $(\mathbb{X}, \mathbb{A}) = (\mathbb{R}, \mathbb{Z})$ and 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 $H(\lfloor \boldsymbol{u} \rfloor)$. A rate tuple $(R_1, \dots, R_K)$ is achievable for recovering the $\mathbf{A}$-linear combinations of codewords under the distortion tuple $(D_1, \dots, D_L)$ if for some choice of test channels $P_{U_k|X_k}$, $k \in [K]$ such that* (12) *holds, 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 $\mathcal{H}([\boldsymbol{u}]_\mathcal{T} | [\boldsymbol{x}]_\mathcal{T}) = H([\boldsymbol{u}]_\mathcal{T} | [\boldsymbol{x}]_\mathcal{T})$, $\mathcal{H}_\mathbf{B}(\boldsymbol{u}|Y) = H(\mathbf{B}\boldsymbol{u}|Y)$ 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 $h([\boldsymbol{u}]_\mathcal{T} | [\boldsymbol{x}]_\mathcal{T})$, $h(\mathbf{T}(\mathbf{B})\boldsymbol{u}|Y)$ and $h(\mathbf{T}(\mathbf{CB})\boldsymbol{u}|Y)$.

VI. The Two-User Case

To gain some insight into the rate regions described by (10) and (11), let us consider the two-user case ($K = 2$).



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

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 [cf. (11)] to obtain the DLC rate region $\mathscr{R}(\mathbf{A})$. The Berger–Tung rate region $\mathscr{R}_{\mathrm{BT}}$, which is achievable with (unstructured) random codes, is shown for reference.

Figure 2 exhibits the different subregions that compose the DLC rate region $\mathscr{R}(\mathbf{A})$. The set $\mathscr{R}_{\mathrm{BT}} = \{(R_1, R_2) \in \mathbb{R}_+^2\colon\ R_1 > I(U_1; X_1|U_2),\ R_2 > I(U_2; X_2|U_1),\ R_1 + R_2 > I(U_1, U_2; X_1, X_2)\}$ denotes the conventional Berger–Tung rate region.

For the recovery of a single linear combination ($L = 1$) specified by a coefficient vector $\mathbf{A} = [a_1\ a_2] \in \mathbb{A}^{1\times 2}$ (with $a_1, a_2 \neq 0$), one can show that (11) can be simplified to

$$\mathscr{R}(\mathbf{A}) = \mathscr{Q}\left(\left[\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right]\right) \cup \mathscr{Q}\left([a_1\ a_2]\right) = \mathscr{R}_{\mathrm{BT}} \cup \mathscr{Q}\left([a_1\ a_2]\right). \quad (13)$$

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

$$\mathscr{R}(\mathbf{A}) = \bigcup_{\mathbf{B}\colon\ \mathrm{rank}\,\mathbf{B}=2} \mathscr{Q}\left(\left[\begin{smallmatrix}b_{11}&b_{12}\\b_{21}&b_{22}\end{smallmatrix}\right]\right) = \mathscr{Q}\left(\left[\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right]\right). \quad (14)$$

Depending on channel characteristics and parameters, we are sometimes facing a situation as in Figure 2a, in which structured codes cannot outperform the conventional Berger–Tung scheme, whereas in other situations, as in Figure 2b, the region achieved by our scheme is larger.

VII. Gaussian Sources

Let us evaluate Theorem 3 for Gaussian distributions. Consider a central multivariate Gaussian source vector $\boldsymbol{x} \sim \mathcal{N}_{\mathbb{R}}(\mathbf{0}, \mathbf{K}_{\boldsymbol{x}})$ with a covariance matrix $\mathbf{K}_{\boldsymbol{x}}$ whose diagonal entries $\mathsf{E}[X_k^2] = P_k$, $k \in [K]$ represent the variance of each source. Furthermore, consider a Gaussian helper channel that supplies the decoder with a side information $\boldsymbol{y} \in \mathbb{R}^M$ that is jointly Gaussian with the source vector (e.g., obtained via uncoded transmission of $\boldsymbol{x}$ over a Gaussian multiple-access link with M receiver antennas), according to $\boldsymbol{y} = \mathbf{H}\boldsymbol{x} + \boldsymbol{z}$ where $\boldsymbol{x} = \left[X_1, \ldots, X_K\right]^{\mathsf{T}}$ represents the source vector, $\mathbf{H} \in \mathbb{R}^{M\times K}$ stands for the channel gain matrix and $\boldsymbol{z} \sim \mathcal{N}_{\mathbb{R}}(\mathbf{0}, \mathbf{I})$ is i.i.d. additive Gaussian noise. The corollary below improves upon the lattice-based rate regions from [7], [10], due in part to the use of simultaneous decoding.

Corollary 1 (Gaussian DLC)**.** *Let* $(\mathbb{X}, \mathbb{A}) = (\mathbb{R}, \mathbb{Z})$*. We evaluate Theorem 3 for the Gaussian helper channel, auxiliaries* U_k *that are jointly Gaussian and* ρ_k*-correlated with their respective source* X_k*, i.e.,* $U_k = \beta_k\left(\sqrt{\rho_k}X_k + \sqrt{(1-\rho_k)P_k}V_k\right)$ *with scaling parameters* $\beta_k > 0$ *and correlation coefficients* ρ_k*, and where* $V_k \sim \mathcal{N}(0,1)$ *are i.i.d. noise terms, resulting in a marginal pdf* $U_k \sim \mathcal{N}_{\mathbb{R}}(0, \beta_k^2 P_k)$*. Then,* $\mathscr{Q}(\mathbf{B}, M, \mathcal{T})$ *in Theorem 3 specializes to the set of rates such that [cf.* (8)*]*

$$\sum_{k\in\mathcal{T}} R_k > -\frac{1}{2}\sum_{k\in\mathcal{T}} \log(\beta_k^2 P_k(1-\rho_k)) + \frac{1}{2}\log\frac{\det\left(\mathbf{B}\mathbf{K}_{\boldsymbol{u}|\boldsymbol{y}}\mathbf{B}^{\mathsf{T}}\right)}{\det(\boldsymbol{\Sigma}(\mathbf{B}))} - J(\mathbf{B}, M) \quad (15)$$

where $\mathbf{K}_{\boldsymbol{u}|\boldsymbol{y}}$ *stands for the conditional covariance matrix* $\mathbf{K}_{\boldsymbol{u}|\boldsymbol{y}} = \mathbf{K}_{\boldsymbol{u}} - \mathbf{K}_{\boldsymbol{u},\boldsymbol{y}}\mathbf{K}_{\boldsymbol{y}}^{-1}\mathbf{K}_{\boldsymbol{y},\boldsymbol{u}}$ *with*

$$\begin{aligned}\mathbf{K}_{\boldsymbol{y}} &= \mathsf{E}[\boldsymbol{y}\boldsymbol{y}^{\mathsf{T}}] = \mathbf{H}\mathbf{K}_{\boldsymbol{x}}\mathbf{H}^{\mathsf{T}} + \mathbf{I}\\ \mathbf{K}_{\boldsymbol{u}} &= \mathsf{E}[\boldsymbol{u}\boldsymbol{u}^{\mathsf{T}}] = \mathbf{D}_{\boldsymbol{\beta}}\left(\mathbf{D}_{\boldsymbol{\rho}}^{1/2}\mathbf{K}_{\boldsymbol{x}}\mathbf{D}_{\boldsymbol{\rho}}^{1/2} + (\mathbf{I}-\mathbf{D}_{\boldsymbol{\rho}})\mathbf{P}\right)\mathbf{D}_{\boldsymbol{\beta}}\\ \mathbf{K}_{\boldsymbol{u},\boldsymbol{y}} &= \mathsf{E}[\boldsymbol{u}\boldsymbol{y}^{\mathsf{T}}] = \mathbf{D}_{\boldsymbol{\beta}}\mathbf{D}_{\boldsymbol{\rho}}^{1/2}\mathbf{K}_{\boldsymbol{x}}\mathbf{H}^{\mathsf{T}}\end{aligned}$$

where $\mathbf{D}_{\boldsymbol{\beta}} = \mathrm{diag}(\beta_1, \ldots, \beta_K)$*,* $\mathbf{D}_{\boldsymbol{\rho}} = \mathrm{diag}(\rho_1, \ldots, \rho_K)$ *and* $\mathbf{P} = \mathrm{diag}(P_1, \ldots, P_K)$*, and where*

$$J(\mathbf{B}, M) = \inf_{\mathbf{C}\in\mathscr{C}_{\mathbb{Z}}(M)} \frac{1}{2}\log\frac{\det\left(\mathbf{C}\mathbf{B}\mathbf{K}_{\boldsymbol{u}|\boldsymbol{y}}\mathbf{B}^{\mathsf{T}}\mathbf{C}^{\mathsf{T}}\right)}{\det(\boldsymbol{\Sigma}(\mathbf{C}\mathbf{B}))}. \quad (16)$$

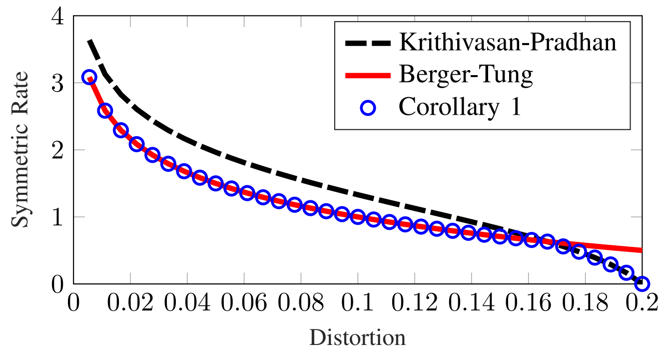


Fig. 3. Symmetric rates for the scenario where the receiver wants the difference $X_1 - X_2$ for Gaussian sources with covariance matrix $\mathbf{K}_{\boldsymbol{x}} = \left[\begin{smallmatrix}1&0.9\\0.9&1\end{smallmatrix}\right]$.

Finally, let us further particularize Corollary 1 to the case of two users ($K = 2$), no side information ($Y = \emptyset$), normalized sources ($\mathbf{K}_{\boldsymbol{x}} = \left[\begin{smallmatrix}1&\gamma\\\gamma&1\end{smallmatrix}\right]$) with positive correlation ($\gamma > 0$) and a receiver that wants their difference ($\mathbf{A} = [1, -1]$) with quadratic distortion D. This scenario was originally proposed by Krithivasan and Pradhan [7] as a Gaussian version of the celebrated Körner–Marton example. Their nested lattice scheme attains the symmetric rate $R_1 = R_2 = R_{\text{KP,sym}} = \frac{1}{2}\log\left(\frac{4-4\gamma}{D}\right)$. Let $R_{\text{BT,sym}}$ be the symmetric rate from the Berger–Tung rate region $\mathscr{R}_{\mathrm{BT}}$ evaluated for auxiliaries from Corollary 1 with $\beta_1 = \beta_2 = \beta$ taken to be the MMSE coefficient and $\rho_1 = \rho_2 = \rho$ selected to attain the desired distortion D. Taking the same auxiliaries in Corollary 1, it can be shown that we attain $R_{\text{joint,sym}} = \min(R_{\text{KP,sym}}, R_{\text{BT,sym}})$, as plotted in Figure 3 for $\gamma = 0.9$. Although it is standard practice to use time-sharing to take the lower convex envelope of the rate region, we have not done so in the plot so as to clearly illustrate that our scheme switches cleanly from the lattice-based rate to the joint-typicality rate. (Note that this is not a general claim: there are scenarios where i.i.d. Berger–Tung coding outperforms our strategy.)

REFERENCES

[1] J. Körner and K. Marton, "How to encode the modulo-two sum of binary sources," *IEEE Transactions on Information Theory*, vol. 25, no. 2, pp. 219–221, 1979.

[2] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. 19, no. 4, pp. 471–480, Jul. 1973.

[3] A. D. Wyner, "Recent results in the Shannon theory," *IEEE Transactions on Information Theory*, vol. 20, no. 1, pp. 2–10, 1974.

[4] I. Csiszar, "Linear codes for sources and source networks: Error exponents, universal coding," *IEEE Transactions on Information Theory*, vol. 28, no. 4, pp. 585–592, Jul 1982.

[5] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1250–1276, 2002.

[6] V. Lalitha, N. Prakash, K. Vinodh, P. V. Kumar, and S. S. Pradhan, "Linear coding schemes for the distributed computation of subspaces," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 4, pp. 678–690, April 2013.

[7] D. Krithivasan and S. S. Pradhan, "Lattices for distributed source coding: Jointly Gaussian sources and reconstruction of a linear function," *IEEE Transactions on Information Theory*, vol. 55, no. 12, pp. 5628–5651, Dec. 2009.

[8] S.-Y. Tung, "Multiterminal source coding," Ph.D. Thesis, Cornell University, Ithaca, NY, 1978.

[9] T. Berger, K. B. Housewright, J. K. Omura, S. Tung, and J. Wolfowitz, "An upper bound on the rate distortion function for source coding with partial side information at the decoder," *IEEE Transactions on Information Theory*, vol. 25, no. 6, November 1979.

[10] A. B. Wagner, "On distributed compression of linear functions," *IEEE Transactions on Information Theory*, vol. 57, no. 1, pp. 79–94, January 2011.

[11] D. N. C. Tse and M. A. Maddah-Ali, "Interference neutralization in distributed lossy source coding," in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, June 2010.

[12] Y. Yang and Z. Xiong, "Distributed compression of linear functions: Partial sum-rate tightness and gap to optimal sum-rate," *IEEE Transactions on Information Theory*, vol. 60, no. 5, pp. 2835–2855, May 2014.

[13] D. Krithivasan and S. S. Pradhan, "Distributed source coding using Abelian group codes," *IEEE Transactions on Information Theory*, vol. 57, no. 3, pp. 1495–1519, Mar. 2011.

[14] 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.

[15] 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.

[16] ——, "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.

[17] 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.

[18] ——, "Compute-forward for DMCs: Simultaneous decoding of multiple combinations," *IEEE Transactions on Information Theory*, vol. 66, no. 10, pp. 6242–6255, 2020.

[19] 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, 2020.

[20] P. Sen and Y.-H. Kim, "Homologous codes for multiple access channels," *IEEE Transactions on Information Theory*, vol. 66, no. 3, pp. 1549–1571, 2020.

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

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

[23] A. Pastore, S. H. Lim, C. Feng, B. Nazer, and M. Gastpar, "A unified discretization approach to compute–forward: From discrete to continuous inputs," *IEEE Transactions on Information Theory*, vol. 69, no. 1, pp. 1–46, 2023.

[24] S. H. Lim, C. Feng, A. Pastore, B. Nazer, and M. Gastpar, "Towards an algebraic network information theory: Distributed lossy computation of linear functions," in *Proc. IEEE Int. Symp. Inf. Theory*, 2019, pp. 1827–1831.

[25] 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.

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

[27] T. Berger, "Multiterminal source coding," in *The Information Theory Approach to Communications*, G. Longo, Ed. New York: Springer-Verlag, 1978, pp. 171–231.

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

[29] 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.

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