Title

The Capacity of 3 User Linear Computation Broadcast

Authors

Yao, Yuhang

Jafar, Syed A

Publication Date

2024-05-16

Peer reviewed

The Capacity of 3 User Linear Computation Broadcast

Yuhang Yao, *Graduate Student Member, IEEE,* and Syed A. Jafar, *Fellow, IEEE*

***Abstract*—The K User Linear Computation Broadcast (LCBC) problem is comprised of d dimensional data (from $\mathbb{F}_q$), that is fully available to a central server, and K users, who require various linear computations of the data, and have prior knowledge of various linear functions of the data as side-information. The optimal broadcast cost is the minimum number of q-ary symbols to be broadcast by the server per computation instance, for every user to retrieve its desired computation. The reciprocal of the optimal broadcast cost is called the capacity. The main contribution of this paper is the exact capacity characterization for the $K = 3$ user LCBC for all cases, i.e., for arbitrary finite fields $\mathbb{F}_q$, arbitrary data dimension d, and arbitrary linear side-informations and demands at each user. A remarkable aspect of the converse (impossibility result) is that unlike the 2 user LCBC whose capacity was determined previously, the entropic formulation (where the entropies of demands and side-informations are specified, but not their functional forms) is insufficient to obtain a tight converse for the 3 user LCBC. Instead, the converse exploits functional submodularity. Notable aspects of achievability include sufficiency of vector linear coding schemes, subspace decompositions that parallel those found previously by Yao Wang in degrees of freedom (DoF) studies of wireless broadcast networks, and efficiency tradeoffs that lead to a constrained waterfilling solution. Random coding arguments are invoked to resolve compatibility issues that arise as each user has a different view of the subspace decomposition, conditioned on its own side-information.**

I. Introduction

Recent years have seen explosive growth both in the number of devices connected to communication networks, as well as in the amount of data generated, shared, and collaboratively processed by these devices. With machine communication expected to dominate human communication, future communication networks will increasingly be used in the service of computation tasks [1]. Along with the processing power of connected devices, a key determining factor of the potential of these '*computation networks*' will be the fundamental limit of their communication-efficiency. Despite a multitude of advances spanning several decades [2]–[15], the capacity limits of computation networks remain largely unknown. Remarkably, this is the case even in the most basic of scenarios such as computational multiple access and broadcast, the presumptive starting points for developing a cohesive theory of computation networks. It is also noteworthy that many applications of recent interest, such as coded caching [16]–[18], private information retrieval [19], [20], coded MapReduce [21], distributed storage exact repair [22]–[24], index coding [25], [26], coded computing [27], [28], data shuffling [29], federated learning [30] and secure aggregation [31], are essentially linear computation multiple access (LCMAC) or broadcast (LCBC) settings with additional application-specific constraints. Future developments, say in networked VR/AR technology [1], [32]–[36], will similarly need linear broadcast and multiple access computational networks for coordination and synchronization [37]–[39] of users' perspectives across space, typically computed as linear projections of real-world coordinates. Evidently, beyond their significance as building blocks, LCMAC/LCBC networks are important in and of themselves.

The collaborative, task-oriented, and interactive character of computation networks manifests in *data dependencies*, and an abundance of *side-information* accumulated at each node from prior computations on overlapping datasets. Both data dependencies and side-information significantly impact the capacity of computation networks. Furthermore, because of the inherently algorithmic character of machine communication, the underlying structures of data dependencies and side-information are often predictable, and may be exploited in principled ways to improve communication efficiency. Indeed, both of these aspects are central to the computation broadcast (CBC) problem, an elemental one-to-many computation network studied recently in [40]. The CBC setting is comprised of data stored at a central server, and multiple users, each of whom is given some function of the data as side-information and wishes to retrieve some other function of the data. The goal is to find Δ^*, which is the least amount of information per computation that the server must broadcast such that all the users are able to compute their respective desired functions. The capacity of CBC is defined as $C = 1/\Delta^*$.

The main result of [40] is an exact capacity characterization for $K = 2$ user *linear* computation broadcast (LCBC), where each user's demand and side-information are linear functions of the data. The K user LCBC problem, illustrated in Figure 1, is specified by the parameters $\big(\mathbb{F}_q, K, d, (m_k, m'_k)_{k\in[K]}, (\mathbf{V}_k, \mathbf{V}'_k)_{k\in[K]}\big)$, namely the finite field $\mathbb{F}_q$, the number of users K, the data dimension d, matrices $\mathbf{V}_k \in \mathbb{F}_q^{d\times m_k}$ that identify the m_k dimensions of User k's demand, and matrices $\mathbf{V}'_k \in \mathbb{F}_q^{d\times m'_k}$ that identify the m'_k dimensions of User k's side-information, for all $k \in [K]$. The index $\ell \in \mathbb{N}$ in Figure 1 identifies the ℓ^{th} computation instance, corresponding to the ℓ^{th} instance of the data vector, $\mathbf{x}(\ell) \in$

This work was supported in part by funding from NSF CCF-2221379 and ONR N00014-21-1-2386. This paper was presented in part at the 2022 IEEE International Symposium on Information Theory (ISIT 2022).

The authors are with the Center for Pervasive Communications and Computing (CPCC), Department of Electrical Engineering and Computer Science, University of California at Irvine, Irvine, CA 92697 USA (emails: yuhangy5@uci.edu; syed@uci.edu).

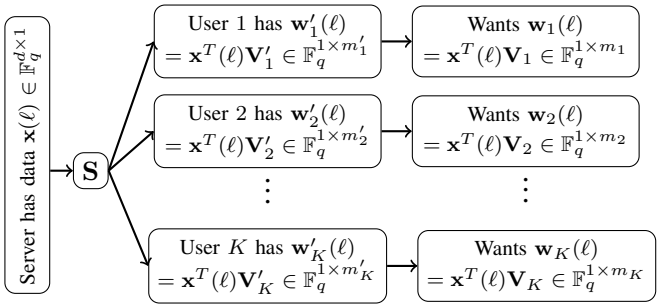

Fig. 1: The ℓ^{th} instance of the LCBC$\big(\mathbb{F}_q, K, d, (m_k, m_k')_{k\in[K]}, (\mathbf{V}_k, \mathbf{V}_k')_{k\in[K]}\big)$, $\ell \in \mathbb{N}$.

$\mathbb{F}_q^{d\times 1}$, such that User k, $k \in [K]$, wants $\mathbf{w}_k(\ell) = \mathbf{x}^T(\ell)\mathbf{V}_k$ and has $\mathbf{w}_k'(\ell) = \mathbf{x}^T(\ell)\mathbf{V}_k'$ as side-information. Following a typical information theoretic formulation, multiple (say L) instances may be considered jointly by a coding scheme for potential gains in efficiency. L is called the batch size and may be chosen freely by a coding scheme. A coding scheme that satisfies all the users' demands across L computation instances by broadcasting a total of N q-ary symbols, achieves rate $R = L/N$, and broadcast cost per computation $\Delta = N/L = 1/R$. The goal is to find the supremum of achievable rates (capacity C), or equivalently, the infimum of achievable broadcast costs per computation ($\Delta^* = 1/C$) across all feasible coding schemes. We refer the reader to Section II to clarify notational aspects, and to Section III for details of the problem formulation. For $K = 2$ users, the optimal broadcast cost is found in [40] to be $\Delta^* = \max\Big(\text{rk}_q([\mathbf{V}_i, \mathbf{V}_i']) - \text{rk}_q(\mathbf{V}_i') + \text{rk}_q([\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_1', \mathbf{V}_2']) - \text{rk}_q([\mathbf{V}_i, \mathbf{V}_i', \mathbf{V}_j'])\Big)$, where $\text{rk}_q()$ is the matrix rank function over $\mathbb{F}_q$, and the max is over $(i, j) \in \{(1, 2), (2, 1)\}$.

The scope of LCBC includes problems such as index coding [25], [41], [42] that have been extensively studied and yet remain open in general. While many instances of index coding have been solved from a variety of perspectives [26], [43]–[47], little is known about the optimal broadcast cost for the general index coding problem. It is shown in [25] that for *scalar linear* index coding, the optimal broadcast cost can be found in general by solving a min-rank problem. The min-rank solution has been extended to index coding with coded side-information in [48] and is not difficult to further generalize to LCBC. However, on top of the difficulty of matrix rank minimizations (known to be NP-hard [49, Thm. 3.1], [50], [51]), scalar linear coding is only one of many possible coding schemes, and it is well known that capacity achieving schemes need not be scalar or linear, even for index coding [52]–[54]. Thus, finding the capacity of LCBC in general is at least as hard as solving the general index coding problem.

On the other hand, index coding problems constitute only a small subset of all possible LCBC instances. The special cases of LCBC that yield index coding problems are precisely those where all the columns of $\mathbf{V}_k, \mathbf{V}_k'$ can be represented as standard basis vectors. Evidently, LCBC allows a significantly richer research space for developing new insights. This is why for LCBC, even settings with only $2, 3$ users are interesting and insightful, whereas such settings would be trivial for index coding. The richer space of LCBC problems is particularly valuable if it is amenable to information theoretic analysis. Intrigued by this possibility, in this work we explore what new technical challenges might emerge in the LCBC setting when we go from 2 to 3 users.

The main result of this work is the exact capacity of the 3 user LCBC for all cases, i.e., for arbitrary $\mathbb{F}_q$, arbitrary data dimension d, and arbitrary demands and side-informations $\mathbf{V}_k, \mathbf{V}_k'$ for each user, $k \in \{1, 2, 3\}$. An explicit expression for the capacity, C, is presented in Theorem 1, and depends on the dimensions (ranks) of various unions and intersections of subspaces corresponding to the users' desired computations and side-information. The intuition behind the explicit form becomes more transparent when it is viewed as the solution to a linear program, in an alternative formulation of the capacity result, presented in Theorem 2. The linear program sheds light on the key ideas behind the optimal coding scheme. One of these ideas is a decomposition of the collective signal spaces of the three users (column spans of the $[\mathbf{V}_k', \mathbf{V}_k]$ matrices) into distinct subspaces that allow different levels of communication efficiency. Remarkably, this decomposition, which is formalized in Lemma 2, closely parallels (see Appendix D) a corresponding decomposition previously obtained in degrees of freedom (DoF) studies of the 3 user MIMO broadcast channel in [55], underscoring its fundamental significance. Facilitated by the subspace decomposition, the linear program formulation of Theorem 2 reveals a non-trivial tradeoff between the number of dimensions of broadcast that are drawn from each subspace, and leads to a constrained waterfilling solution in Section IX-B. What makes the achievability especially challenging is that the users have different (seemingly incompatible) views of the useful information within each subspace depending on their respective side-informations. Random coding arguments are invoked to find broadcast dimensions for the optimal scheme that are useful across the different perspectives. Another remarkable aspect of the capacity result is that non-linear schemes are not needed for the 3 user LCBC. While our optimal schemes make use of both field size extensions (Section VIII-B) and matrix extensions (Section VIII-E), they are still vector linear schemes over $\mathbb{F}_q$. In contrast, *scalar* linear codes were found to be sufficient for the 2 user LCBC in [40]. In terms of the converse bound,[1] an interesting insight emerges from this work regarding the entropic formulation of the LCBC problem that was considered in [40]. In the entropic formulation of [40], the data is assumed i.i.d. uniform, and the entropies of all subsets of demand and side-information random-variables are specified as constraints, but their functional forms are not specified. It was shown in [40] that for the $K = 2$ user LCBC, these entropic constraints combined with standard Shannon entropic inequalities produce a tight converse bound on the download cost per computation. In contrast, in this work we show (see *Remark 6.1* in Section VI) by a counterexample that the same approach cannot work for the $K = 3$ user LCBC,

[1] A *converse* bound refers to an impossibility result, i.e., a lower bound on broadcast cost per computation, or equivalently, an upper bound on capacity.

even if *all* Shannon and non-Shannon information inequalities are applied. Instead, in this work a tight converse for the 3 user LCBC is obtained based on functional submodularity (cf. [56], [57], also Lemma 3 in this work) that additionally takes into account the functional forms of the users' demands and side-informations.

II. Notation

$\mathbb{F}_q$ is a finite field with $q = p^n$ a power of a prime. The elements of the prime field $\mathbb{F}_p$ are represented as $\mathbb{Z}/p\mathbb{Z}$, i.e., integers modulo p. The notation $\mathbb{F}_q^{n_1 \times n_2}$ represents the set of $n_1 \times n_2$ matrices with elements in $\mathbb{F}_q$. For a matrix M, let $\langle M \rangle_q$ denote the $\mathbb{F}_q$-linear vector space spanned by the columns of M. The subscript q will often be suppressed to simplify notation when it is clear from the context. The notation $M_1 \cap M_2$ represents a matrix whose columns form a basis of $\langle M_1 \rangle \cap \langle M_2 \rangle$. $[M_1, M_2]$ represents a concatenated matrix which can be partitioned column-wise into M_1 and M_2. The rank of M over $\mathbb{F}_q$ is denoted by $\text{rk}_q(M)$, and when written as $\text{rk}(M)$ for simplicity, the subscript q is assumed by default. If $\text{rk}(M)$ is equal to the number of columns of M, i.e., M has full column rank, then we say that M is a basis of $\langle M \rangle$. Define a 'conditional-rank' notation as $\text{rk}(M_1|M_2) \triangleq \text{rk}([M_1, M_2]) - \text{rk}(M_2)$. The notation $[n]$ represents the set $\{1, 2, \cdots, n\}$. $\mathbb{N}$ denotes the set of positive integers. $\mathbb{R}_+$ denotes the set of non-negative real numbers. $\mathbb{C}$ denotes the set of complex numbers.

III. Problem Statement

A. The General K User LCBC$(\mathbb{F}_q, K, d, (m_k, m'_k)_{k\in[K]}, (\mathbf{V}_k, \mathbf{V}'_k)_{k\in[K]})$

While our focus in this work is exclusively on the $K = 3$ user case, in this section let us define the LCBC problem for the general K-user setting. As noted previously, the general LCBC problem is specified by the parameters $\big(\mathbb{F}_q, K, d, (m_k, m'_k)_{k\in[K]}, (\mathbf{V}_k, \mathbf{V}'_k)_{k\in[K]}\big)$. There is a stream of data vectors $\mathbf{x}(1), \mathbf{x}(2), \cdots$ that is available at a central server. For each $\ell \in \mathbb{N}$, $\mathbf{x}(\ell) = (x_1(\ell), \cdots, x_d(\ell))^T \in \mathbb{F}_q^{d\times 1}$ is a d-dimensional vector with elements in $\mathbb{F}_q$. There are K users. For all $\ell \in \mathbb{N}$, the k^{th} user has side-information $\mathbf{w}'_k(\ell) = \mathbf{x}^T(\ell)\mathbf{V}'_k \in \mathbb{F}_q^{1\times m'_k}$ and wants $\mathbf{w}_k(\ell) = \mathbf{x}^T(\ell)\mathbf{V}_k \in \mathbb{F}_q^{1\times m_k}$.

1) Coding Scheme: A coding scheme for the LCBC is represented by a choice of parameters in the form of a tuple $\big(L, N, \Phi, (\Psi_k)_{k\in[K]}\big)$. The coding scheme aggregates L instances of data, collectively denoted as $\mathbf{X} \triangleq (\mathbf{x}(1), \cdots, \mathbf{x}(L)) \in \mathbb{F}_q^{d\times L}$, and specifies an encoding function (encoder) $\Phi : \mathbb{F}_q^{d\times L} \to \mathbb{F}_q^N$, as well as K decoding functions (decoders) $\Psi_k : \mathbb{F}_q^N \times \mathbb{F}_q^{L\times m'_k} \to \mathbb{F}_q^{L\times m_k}$, $k \in [K]$.

For compact notation, let us define,

$$\mathbf{W}_k \triangleq \mathbf{X}^T\mathbf{V}_k = (\mathbf{w}_k(1), \cdots, \mathbf{w}_k(L))^T \in \mathbb{F}_q^{L\times m_k}, \quad (1)$$

$$\mathbf{W}'_k \triangleq \mathbf{X}^T\mathbf{V}'_k = (\mathbf{w}'_k(1), \cdots, \mathbf{w}'_k(L))^T \in \mathbb{F}_q^{L\times m'_k}. \quad (2)$$

The encoder Φ maps the data $\mathbf{X}$ to the broadcast information comprised of N symbols in $\mathbb{F}_q$, represented compactly as $\mathbf{S} \in \mathbb{F}_q^N$, i.e.,

$$\Phi(\mathbf{X}) = \mathbf{S} \in \mathbb{F}_q^N. \quad (3)$$

The k^{th} decoder, Ψ_k allows the k^{th} user to retrieve $\mathbf{W}_k$ from the broadcast information $\mathbf{S}$ and the side-information $\mathbf{W}'_k$, i.e.,

$$\Psi_k(\mathbf{S}, \mathbf{W}'_k) = \mathbf{W}_k, \qquad \forall k \in [K], \quad (4)$$

for all realizations of $\mathbf{X}$.

Let us denote the set of all feasible coding schemes as $\mathfrak{C}$. We refer to coding schemes with batch size $L = 1$ as *scalar* (coding) schemes, and those with $L > 1$ as *vector* (coding) schemes.

2) Capacity (C) and Optimal Download Cost per Computation (Δ^*): The rate of a coding scheme $\big(L, N, \Phi, (\Psi_k)_{k\in[K]}\big) \in \mathfrak{C}$, is defined as $R = L/N$ representing the number of computation instances satisfied by the coding scheme per broadcast symbol.[2] The supremum of rates across all feasible coding schemes in $\mathfrak{C}$ is called the capacity of LCBC, i.e.,

$$C \triangleq \sup_{\big(L,N,\Phi,(\Psi_k)_{k\in[K]}\big)\in\mathfrak{C}} L/N. \quad (5)$$

Instead of rate R, it is often more convenient to consider its reciprocal value, the broadcast cost per computation, $\Delta = 1/R = N/L$. The optimal broadcast cost per computation, Δ^* is defined as,

$$\Delta^* \triangleq \inf_{\big(L,N,\Phi,(\Psi_k)_{k\in[K]}\big)\in\mathfrak{C}} N/L \quad (6)$$

$$= 1/C. \quad (7)$$

Since $\Delta^* = 1/C$, the problem of characterizing the capacity C is equivalent to the problem of characterizing the optimal download cost per computation Δ^*. We will find it more convenient to state and prove results in terms of Δ^* in this work.

3) Data Distribution, Entropy: Note that the LCBC problem does not specify any distribution for the data $\mathbf{X}$. This is because the capacity C and optimal broadcast cost per computation Δ^* do not depend on the data distribution. By definition, any coding scheme $\big(L, N, \Phi, (\Psi_k)_{k\in[K]}\big) \in \mathfrak{C}$, while broadcasting no more than N q-ary symbols, must guarantee successful decoding as in(4) for *every* realization of the data, i.e., for all q^{dL} realizations of $\mathbf{X} \in \mathbb{F}_q^{d\times L}$, regardless of what distribution $\mathbf{X}$ follows, and even if $\mathbf{X}$ follows no distribution. This is significant for computation tasks. Recall that conventional communication scenarios are comprised of independent messages that can be compressed prior to communication to reduce the size of the task from the outset and subsequently uncompressed upon successful reception. In principle optimal compression produces uniformly distributed data (otherwise further compression would be possible), thus justifying the common assumption that messages are uniformly distributed. For the LCBC, however, while the desired computation is a linear function of the original uncompressed data, it may no longer be linear after compression. Thus, compression to uniformly distributed data cannot be taken for granted. Furthermore, it is often the case that the data

[2] Viewing each q-ary broadcast symbol as one channel-use, the rate can be equivalently viewed as the number of computation instances satisfied by the coding scheme per channel-use.

distribution is either unknown, or the data is truly arbitrary. Therefore, assuming that data follows a particular distribution may be overly restrictive for computation problems. Such considerations motivate the conservative formulation presented above, which requires strong (maximum rather than average) communication cost guarantees, i.e., any achievable coding scheme must guarantee that a broadcast of N symbols suffices for every data realization, regardless of the distribution of $\mathbf{X}$.

On the other hand, it will be occasionally useful, primarily as a thought-experiment, to consider hypothetically what might happen if the data followed an i.i.d. uniform distribution. Similar to genie-aided proofs, such thought-experiments are useful to construct converse bounds (impossibility results) by the following reasoning. Given any coding scheme $\big(L, N, \Phi, (\Psi_k)_{k\in[K]}\big) \in \mathfrak{C}$, we wish to find lower bounds on the broadcast cost N. As a thought-experiment, suppose the data $\mathbf{X}$ follows a distribution $P_{\mathbf{X}}(\mathbf{x})$ and this coding scheme is used. This imparts a corresponding distribution to the broadcast symbol $\mathbf{S}$, say $P_{\mathbf{S}}(\mathbf{s})$. However, since $\mathbf{S} \in \mathbb{F}_q^N$ by the definition of the coding scheme, the entropy $H(\mathbf{S}) \leq \log_q |\mathbb{F}_q^N| = N$ in q-ary units, which produces a lower bound on the broadcast cost, i.e., $N \geq H(\mathbf{S})$. Thus any choice of $P_{\mathbf{X}}(\mathbf{x})$ facilitates entropic analysis and leads to a lower bound on N, by calculating the entropy of $\mathbf{S}$ produced by the coding scheme. The quality of the bound depend on the choice of $P_{\mathbf{X}}(\mathbf{x})$. For example, if we assume the data is deterministic, then so is $\mathbf{S}$, i.e., $H(\mathbf{S}) = 0$, leading to the bound $N \geq 0$, which is not very useful. Uniform distributions are particularly interesting because they tend to produce good converse bounds. In preparation for the converse arguments in the sequel, it is useful to recall the following facts.

1) For a random variable Z, that takes values in a set $\mathcal{Z}$ according to the probability mass function $p_Z(z)$, the entropy $H(Z)$ in q-ary units is defined as,
$$H(Z) \triangleq -\sum_{z\in\mathcal{Z}} p_Z(z) \log_q p_z(z). \tag{8}$$
2) If Z is i.i.d. uniform over $\mathbb{F}_q^{\mu\times\nu}$ then $H(Z) = \log_q |\mathbb{F}_q^{\mu\times\nu}| = \log_q(q^{\mu\nu}) = \mu\nu$ in q-ary units.
3) If Z is i.i.d. uniform over $\mathbb{F}_q^{\mu\times\nu}$ and $M \in \mathbb{F}_q^{\mu\times\xi}$ is a deterministic matrix, then
$$H(Z^T M) = \nu \cdot \mathrm{rk}_q(M) \tag{9}$$
in q-ary units. This is seen as follows. Let Z_{*i} denote the i^{th} column of Z. Then $H(Z^T M) = H(Z_{*1}^T M, Z_{*2}^T M, \cdots, Z_{*\nu}^T M) = \sum_{i=1}^{\nu} H(Z_{*i}^T M) \overset{(a)}{=} \sum_{i=1}^{\nu} \mathrm{rk}_q(M) = \nu \cdot \mathrm{rk}_q(M)$. The step labeled (a) is a direct application of [40, Lemma 2].
4) If Z is i.i.d. uniform over $\mathbb{F}_q^{\mu\times\nu}$ and $M_1 \in \mathbb{F}_q^{\mu\times\xi_1}$, $M_2 \in \mathbb{F}_q^{\mu\times\xi_2}$ are deterministic matrices, then
$$\begin{aligned} H(Z^T M_1 \mid Z^T M_2) &= \nu(\mathrm{rk}_q([M_1, M_2]) - \mathrm{rk}_q(M_2)) \\ &= \nu \cdot \mathrm{rk}_q(M_1 \mid M_2) \end{aligned} \tag{10}$$
in q-ary units, where we used the conditional-rank notation $\mathrm{rk}_q(M_1 \mid M_2)$ as defined in Section II. Using (9), this is seen as follows: $H(Z^T M_1 \mid Z^T M_2) = H(Z^T M_1, Z^T M_2) - H(Z^T M_2) = H(Z^T[M_1, M_2]) - H(Z^T M_2) = \nu \cdot \mathrm{rk}_q([M_1, M_2]) - \nu \cdot \mathrm{rk}_q(M_2) = \nu \cdot \mathrm{rk}_q(M_1 \mid M_2)$.

B. Signal Spaces $\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3$ and their Intersections

Recall that in this work our focus is on the LCBC with $K = 3$ users, i.e., the most general setting that we consider in this work is $\mathrm{LCBC}\big(\mathbb{F}_q, 3, d, (m_k, m_k')_{k\in[3]}, (\mathbf{V}_k, \mathbf{V}_k')_{k\in[3]}\big)$. Let us define the spaces $\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3$, associated with the 3 users, as follows,
$$\mathbf{U}_1 \triangleq [\mathbf{V}_1', \mathbf{V}_1],\ \mathbf{U}_2 \triangleq [\mathbf{V}_2', \mathbf{V}_2],\ \mathbf{U}_3 \triangleq [\mathbf{V}_3', \mathbf{V}_3], \tag{11}$$
and also define the following intersections,
$$\mathbf{U}_{ij} \triangleq \mathbf{U}_i \cap \mathbf{U}_j, \forall i, j \in [3], i \neq j, \tag{12}$$
$$\mathbf{U}_{123} \triangleq \mathbf{U}_1 \cap \mathbf{U}_2 \cap \mathbf{U}_3, \tag{13}$$
$$\mathbf{U}_{i(j,k)} \triangleq \mathbf{U}_i \cap [\mathbf{U}_j, \mathbf{U}_k], \quad \forall (i,j,k) \in \{\text{permutations of } (1,2,3)\}. \tag{14}$$
Recall that the subspaces $\langle \mathbf{U}_i \rangle$ refer to the column spans of the corresponding matrices. These subspaces will be essential to the understanding of the 3 user LCBC.

IV. Preliminary Step: Subspace Decomposition

For problems involving a vector space, the choice of a suitable basis representation is often an important preliminary simplification step. When multiple vector spaces are involved, it is similarly useful to explicitly partition them into independent subspaces that fit the needs of the problem. For the 3 user LCBC, there are three vector spaces of interest, namely $\langle \mathbf{U}_1 \rangle, \langle \mathbf{U}_2 \rangle, \langle \mathbf{U}_3 \rangle$, as defined in Section III-B. A suitable decomposition of these spaces into independent subspaces corresponding to various intersections is an important preliminary simplification that is the focus of this section. To put it concisely, we need the following two lemmas regarding linear subspaces $\langle \mathbf{U}_1 \rangle, \langle \mathbf{U}_2 \rangle, \langle \mathbf{U}_3 \rangle$.

Lemma 1 (2-space decomposition)**.** *There exist* 3 *matrices,* $\mathbf{B}_{12}, \mathbf{B}_{1c}$ *and* $\mathbf{B}_{2c}$ *such that* $\mathbf{B}_{12}$ *is a basis of* $\langle \mathbf{U}_{12} \rangle$, $[\mathbf{B}_{12}, \mathbf{B}_{1c}]$ *is a basis of* $\langle \mathbf{U}_1 \rangle$, $[\mathbf{B}_{12}, \mathbf{B}_{2c}]$ *is a basis of* $\langle \mathbf{U}_2 \rangle$, *and* $[\mathbf{B}_{12}, \mathbf{B}_1, \mathbf{B}_2]$ *is a basis of* $\langle [\mathbf{U}_1, \mathbf{U}_2] \rangle$.

Note that Lemma 1 also implies the following dimension formula,
$$\begin{aligned} \mathrm{rk}(\mathbf{U}_1) + \mathrm{rk}(\mathbf{U}_2) &= \mathrm{rk}(\mathbf{B}_{12}) + \mathrm{rk}(\mathbf{B}_{1c}) + \mathrm{rk}(\mathbf{B}_{12}) + \mathrm{rk}(\mathbf{B}_{2c}) \\ &= \mathrm{rk}(\mathbf{U}_1 \cap \mathbf{U}_2) + \mathrm{rk}([\mathbf{U}_1, \mathbf{U}_2]). \end{aligned} \tag{15}$$
A common proof of Lemma 1 from a constructive perspective (e.g. [58, Thm. 3, Ch. 3]) is based on incrementally growing a basis representation, and is summarized as follows. First one finds $\mathbf{B}_{12} \in \mathbb{F}_q^{d\times \mathrm{rk}(\mathbf{U}_1\cap\mathbf{U}_2)}$ as a basis of $\langle \mathbf{U}_1 \cap \mathbf{U}_2 \rangle$. Then, by the basis extension theorem, one can find a submatrix $\mathbf{B}_{1c} \in \mathbb{F}_q^{d\times \mathrm{rk}(\mathbf{U}_1)}$ of $\mathbf{U}_1$ such that $[\mathbf{B}_{12}, \mathbf{B}_{1c}]$ spans $\langle \mathbf{U}_1 \rangle$, and similarly a submatrix $\mathbf{B}_{2c} \in \mathbb{F}_q^{d\times \mathrm{rk}(\mathbf{U}_2)}$ of $\mathbf{U}_2$ such that $[\mathbf{B}_{12}, \mathbf{B}_{2c}]$ spans $\langle \mathbf{U}_2 \rangle$. Note that $\langle \mathbf{B}_{2c} \rangle$ only has trivial intersection with $\langle \mathbf{U}_1 \rangle$ because otherwise $\mathbf{B}_{2c}\mathbf{v} + \mathbf{U}_1\mathbf{v}' = 0 \implies \mathbf{B}_{2c}\mathbf{v} \in \langle \mathbf{B}_{12} \rangle$ where $\mathbf{v}, \mathbf{v}'$ are non-zero vectors, which contradicts that $[\mathbf{B}_{12}, \mathbf{B}_{2c}]$ form a basis. Therefore,

$[\mathbf{B}_{12}, \mathbf{B}_{1c}, \mathbf{B}_{2c}]$ is a basis of $\langle[\mathbf{U}_1, \mathbf{U}_2]\rangle$ since it also spans $\langle[\mathbf{U}_1, \mathbf{U}_2]\rangle$. □

Figure 2 illustrates the decomposition of $\langle \mathbf{U}_1 \rangle$ and $\langle \mathbf{U}_2 \rangle$ by identifying 3 subspaces, each labeled by its basis representation.

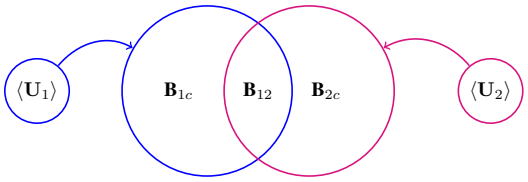

Fig. 2: Decomposition of $\langle \mathbf{U}_1 \rangle, \langle \mathbf{U}_2 \rangle$ into 3 subspaces labeled by their respective bases.

The following lemma non-trivially extends the argument to 3 linear subspaces.

Lemma 2 (3-space decomposition)**.** *There exist* 10 *matrices,* $\mathbf{B}_{123}$, $\mathbf{B}_{12}$, $\mathbf{B}_{13}$, $\mathbf{B}_{23}$, $\mathbf{B}_{1(2,3)}$, $\mathbf{B}_{2(1,3)}$, $\mathbf{B}_{3(1,2)}$, $\mathbf{B}_{1c}$, $\mathbf{B}_{2c}$,$\mathbf{B}_{3c}$, *such that the following properties (P1)-(P20) are satisfied.*

(P1) $\mathbf{B}_{123}$ *is a basis of* $\langle \mathbf{U}_{123} \rangle$,

(P2) $[\mathbf{B}_{123}, \mathbf{B}_{12}]$ *is a basis of* $\langle \mathbf{U}_{12} \rangle$,

(P3) $[\mathbf{B}_{123}, \mathbf{B}_{13}]$ *is a basis of* $\langle \mathbf{U}_{13} \rangle$,

(P4) $[\mathbf{B}_{123}, \mathbf{B}_{23}]$ *is a basis of* $\langle \mathbf{U}_{23} \rangle$,

(P5) $[\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{13}]$ *is a basis of* $\langle [\mathbf{U}_{12}, \mathbf{U}_{13}] \rangle$,

(P6) $[\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{23}]$ *is a basis of* $\langle [\mathbf{U}_{12}, \mathbf{U}_{23}] \rangle$,

(P7) $[\mathbf{B}_{123}, \mathbf{B}_{13}, \mathbf{B}_{23}]$ *is a basis of* $\langle [\mathbf{U}_{13}, \mathbf{U}_{23}] \rangle$,

(P8) $[\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{13}, \mathbf{B}_{1(2,3)}]$ *is a basis of* $\langle \mathbf{U}_{1(2,3)} \rangle$,

(P9) $[\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{23}, \mathbf{B}_{2(1,3)}]$ *is a basis of* $\langle \mathbf{U}_{2(1,3)} \rangle$,

(P10) $[\mathbf{B}_{123}, \mathbf{B}_{13}, \mathbf{B}_{23}, \mathbf{B}_{3(1,2)}]$ *is a basis of* $\langle \mathbf{U}_{3(1,2)} \rangle$,

(P11) $[\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{13}, \mathbf{B}_{1(2,3)}, \mathbf{B}_{1c}]$ *is a basis of* $\langle \mathbf{U}_1 \rangle$,

(P12) $[\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{23}, \mathbf{B}_{2(1,3)}, \mathbf{B}_{2c}]$ *is a basis of* $\langle \mathbf{U}_2 \rangle$,

(P13) $[\mathbf{B}_{123}, \mathbf{B}_{13}, \mathbf{B}_{23}, \mathbf{B}_{3(1,2)}, \mathbf{B}_{3c}]$ *is a basis of* $\langle \mathbf{U}_3 \rangle$,

(P14) $[\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{13}, \mathbf{B}_{23}, \mathbf{B}_{1(2,3)}, \mathbf{B}_{2(1,3)}, \mathbf{B}_{1c}, \mathbf{B}_{2c}]$ *is a basis of* $\langle [\mathbf{U}_1, \mathbf{U}_2] \rangle$,

(P15) $[\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{13}, \mathbf{B}_{23}, \mathbf{B}_{1(2,3)}, \mathbf{B}_{3(1,2)}, \mathbf{B}_{1c}, \mathbf{B}_{3c}]$ *is a basis of* $\langle [\mathbf{U}_1, \mathbf{U}_3] \rangle$,

(P16) $[\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{13}, \mathbf{B}_{23}, \mathbf{B}_{2(1,3)}, \mathbf{B}_{3(1,2)}, \mathbf{B}_{2c}, \mathbf{B}_{3c}]$ *is a basis of* $\langle [\mathbf{U}_2, \mathbf{U}_3] \rangle$,

(P17) $[\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{23}, \mathbf{B}_{13}, \mathbf{B}_{1(2,3)}, \mathbf{B}_{2(1,3)}, \mathbf{B}_{1c}, \mathbf{B}_{2c}, \mathbf{B}_{3c}]$ *is a basis of* $\langle [\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3] \rangle$,

(P18) $[\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{23}, \mathbf{B}_{13}, \mathbf{B}_{1(2,3)}, \mathbf{B}_{3(1,2)}, \mathbf{B}_{1c}, \mathbf{B}_{2c}, \mathbf{B}_{3c}]$ *is a basis of* $\langle [\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3] \rangle$,

(P19) $[\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{23}, \mathbf{B}_{13}, \mathbf{B}_{2(1,3)}, \mathbf{B}_{3(1,2)}, \mathbf{B}_{1c}, \mathbf{B}_{2c}, \mathbf{B}_{3c}]$ *is a basis of* $\langle [\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3] \rangle$, *and*

(P20) $\mathbf{B}_{1(2,3)}, \mathbf{B}_{2(1,3)}, \mathbf{B}_{3(1,2)}$ *have identical size and* $\mathbf{B}_{1(2,3)} + \mathbf{B}_{2(1,3)} = \mathbf{B}_{3(1,2)}$.

We leave the proof to Appendix C. Figure 3 illustrates the decomposition of $\langle \mathbf{U}_1 \rangle, \langle \mathbf{U}_2 \rangle, \langle \mathbf{U}_3 \rangle$ by identifying 10 subspaces, each labeled by its basis representation.

We conclude this section with the following observations.

Remark 4.1 The decomposition of 2 linear subspaces in Lemma 1 resembles the decomposition of 2 sets, e.g., the inclusion-exclusion principle and Venn's diagrams are reflected in the decompositions. However, the set-theoretic analogy is no longer true for 3 linear subspaces, as in the decomposition the 3 yellow spaces are not mutually independent. Appendix B provides more discussion regarding this property.

Remark 4.2 Identifying suitable intersecting subspaces within vector spaces is also a recurrent theme in the degrees of freedom (DoF) studies of wireless networks, e.g., to simplify the design of interference alignment schemes in MIMO settings [59], [60]. In particular, the DoF study of a 3 user wireless MIMO BC setting in the PhD thesis of Wang [55, Ch. 3] provides a subspace decomposition that very closely parallels Lemma 2. The correspondence and the distinctions between the two are discussed in Appendix D, as are the limitations that prevent the proof in [55, Ch. 3] from carrying over directly to our finite field setting. An independent proof of Lemma 2 for our setting is provided in Appendix C. Notably the proof in Appendix C only relies on arguments that hold both over finite fields as well as over the field of complex numbers, thereby unifying the two settings.

Remark 4.3 Lemma 2 ignores the details of how each $\mathbf{U}_i$ is composed of $\mathbf{V}_i$ and $\mathbf{V}'_i$. Depending on their own side-information and demand, each user will have a *different* conditional view of these subspaces. This most essential aspect of the LCBC problem is not reflected in the decomposition. Thus, it is worthwhile to note that the decomposition is primarily a preparatory step, the main technical challenge from both achievability and converse perspectives remains focused on accounting for the distinct side-information and demand structures across users. See also *Remark 5.2.3*.

V. Results

A. A Closed Form Capacity Expression for the 3 User LCBC

As our main result, the following theorem states the capacity of the 3 user LCBC in closed form.

Theorem 1. *For the* $K = 3$ *user general LCBC, i.e., LCBC*$(\mathbb{F}_q, 3, d, (m_k, m'_k)_{k\in[3]}, (\mathbf{V}_k, \mathbf{V}'_k)_{k\in[3]})$*, the capacity* $C = 1/\max\{\Delta_1, \Delta_2\}$*, equivalently, the optimal broadcast cost,* $\Delta^* = \max\{\Delta_1, \Delta_2\}$ *where,*

$$\Delta_n \triangleq \max_{(i,j,k)\in\{\textit{permutations of } (1,2,3)\}} \{\Delta_n^{ijk}\}, \quad n \in \{1,2\}, \tag{16}$$

and

$$\begin{aligned}\Delta_1^{ijk} &\triangleq \mathrm{rk}(\mathbf{V}_1 \mid \mathbf{V}'_1) + \mathrm{rk}(\mathbf{V}_2 \mid \mathbf{V}'_2) + \mathrm{rk}(\mathbf{V}_3 \mid \mathbf{V}'_3) \\ &\quad - \mathrm{rk}(\mathbf{U}_{ij} \mid \mathbf{V}'_j) - \mathrm{rk}(\mathbf{U}_{k(i,j)} \mid \mathbf{V}'_k), \end{aligned} \tag{17}$$

$$\begin{aligned}\Delta_2^{ijk} &\triangleq \mathrm{rk}(\mathbf{V}_1 \mid \mathbf{V}'_1) + \mathrm{rk}(\mathbf{V}_2 \mid \mathbf{V}'_2) + \mathrm{rk}(\mathbf{V}_3 \mid \mathbf{V}'_3) \\ &\quad - \frac{1}{2}\Big(\min_{\ell\in[3]} \big(\mathrm{rk}(\mathbf{U}_{123} \mid \mathbf{V}'_\ell)\big) + \mathrm{rk}([\mathbf{U}_{ij}, \mathbf{U}_{ik}] \mid \mathbf{V}'_i) \\ &\quad + \mathrm{rk}(\mathbf{U}_{j(i,k)} \mid \mathbf{V}'_j) + \mathrm{rk}(\mathbf{U}_{k(i,j)} \mid \mathbf{V}'_k)\Big). \end{aligned} \tag{18}$$

Recall the conditional-rank notation defined in Section II, $\mathrm{rk}(X|Y) \triangleq \mathrm{rk}([X,Y]) - \mathrm{rk}(Y)$. The proof of Theorem 1 will

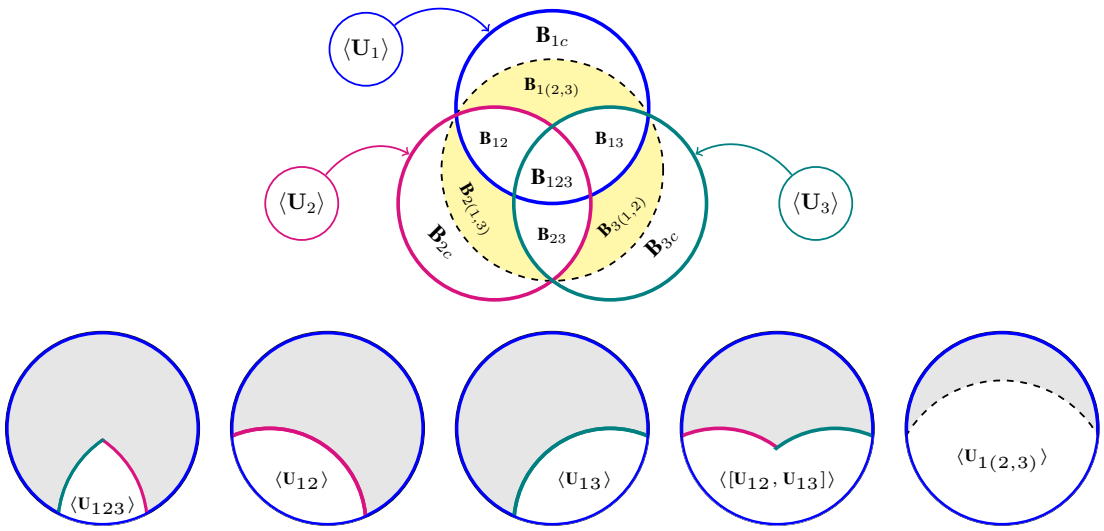

Fig. 3: The top of the figure shows the decomposition of $\langle \mathbf{U}_1 \rangle, \langle \mathbf{U}_2 \rangle, \langle \mathbf{U}_3 \rangle$ into 10 subspaces that are labeled by corresponding bases as specified in Lemma 2. The five blue circles in the bottom row each show $\langle \mathbf{U}_1 \rangle$, and highlight the subspaces $\langle \mathbf{U}_{123} \rangle$, $\langle \mathbf{U}_{12} \rangle$, $\langle \mathbf{U}_{13} \rangle$, $\langle [\mathbf{U}_{12}, \mathbf{U}_{13}] \rangle$ and $\langle \mathbf{U}_{1(2,3)} \rangle$, respectively. The compact notations, $\mathbf{U}_{123}, \mathbf{U}_{12}, \mathbf{U}_{1(2,3)}$ etc., are defined in Section III, for example $\mathbf{U}_{12} \triangleq \mathbf{U}_1 \cap \mathbf{U}_2$ and $\mathbf{U}_{1(2,3)} \triangleq \mathbf{U}_1 \cap [\mathbf{U}_2, \mathbf{U}_3]$. The three subspaces highlighted as dashed yellow regions are not independent, the span of the union of any two of them contains the third.

be presented along with the proof of the upcoming Theorem 2, in Sections VII, VIII, and IX according to the proof structure specified in Section V-C.

Remark 5.1.1 The bound $\Delta^* \geq \Delta_1$ follows from a generalization of the converse bound of the 2 user LCBC, and is similar to the genie-aided converse bound of coded caching (e.g., [61, (71)-(75)]). However, unlike the 2 user LCBC, this bound is not sufficient for the 3 user LCBC, which is why we also need the bound $\Delta^* \geq \Delta_2$. The bound $\Delta^* \geq \Delta_2$ encapsulates the new technical challenge in the 3 user LCBC from the converse perspective (see Section VII).

Remark 5.1.2 The capacity of the 3 user LCBC can be expressed in various equivalent forms. The *closed* form presented in Theorem 1 emerges naturally from the converse bounds. Indeed, the converse in Section VII directly produces two bounds, one each for $\Delta_1^{ijk}, \Delta_2^{ijk}$. The achievability argument on the other hand, takes a different approach which involves auxiliary parameters (the $\lambda_\bullet$ parameters in Theorem 2) representing various design choices. Optimizing the design choices amounts to a linear program, the solution to which yields the same Δ^* as Theorem 1. Even though the converse and achievability perspectives ultimately lead to the same Δ^*, their different forms yield different insights. The achievability perspective in particular yields constructive insights into the tradeoffs involved in simultaneously satisfying all 3 users' demands. This alternative (but equivalent) form of the capacity result is presented next.

B. An Alternative Expression for the Capacity of the 3 User LCBC

Theorem 2.

$$\Delta^* = F^* \tag{19}$$

where Δ^* is the optimal broadcast cost for the $K = 3$ user general LCBC and F^* is the solution to the following linear program,

$$F^* \triangleq \min_{\lambda_{123},\lambda_{12},\lambda_{13},\lambda_{23},\lambda \in \mathbb{R}_+} \mathrm{rk}(\mathbf{V}_1|\mathbf{V}_1') + \mathrm{rk}(\mathbf{V}_2|\mathbf{V}_2') + \mathrm{rk}(\mathbf{V}_3|\mathbf{V}_3') - 2\lambda_{123} - \lambda_{12} - \lambda_{13} - \lambda_{23} - \lambda, \tag{20}$$

such that[3]

$$\lambda_{123} \leq \mathrm{rk}(\mathbf{U}_{123} \mid \mathbf{V}_i'), \forall i \in \{1,2,3\}, \tag{21}$$

$$\lambda_{ij} + \lambda_{123} \leq \min\left(\mathrm{rk}(\mathbf{U}_{ij} \mid \mathbf{V}_i'), \mathrm{rk}(\mathbf{U}_{ij} \mid \mathbf{V}_j')\right), \quad \forall (i,j) \in \{(1,2),(1,3),(2,3)\}, \tag{22}$$

$$\lambda_{ij} + \lambda_{ik} + \lambda_{123} \leq \mathrm{rk}([\mathbf{U}_{ij}, \mathbf{U}_{ik}] \mid \mathbf{V}_i'), \quad \forall (i,j,k) \in \{(1,2,3),(2,1,3),(3,1,2)\}, \tag{23}$$

$$\lambda + \lambda_{ij} + \lambda_{ik} + \lambda_{123} \leq \mathrm{rk}(\mathbf{U}_{i(j,k)} \mid \mathbf{V}_i'), \quad \forall (i,j,k) \in \{(1,2,3),(2,1,3),(3,1,2)\}. \tag{24}$$

Remark 5.2.1 For the sake of high level intuition, Figure 4 conveys a somewhat oversimplified (the caveat is noted in *Remark 5.2.3*) understanding of the conditions (21)-(24) in Theorem 2. The $\lambda_\bullet$ parameters represent the size (dimension) of signals in various subspaces to be broadcast by the coding scheme. Depending upon the region they fall in, the subspaces have different communication efficiencies. For instance, note that λ_{123} falls in $\mathbf{U}_{123}$, and carries information that is simultaneously useful for all 3 users. Thus, λ_{123} transmitted dimensions satisfy a total of $3\lambda_{123}$ dimensions of demand (λ_{123} per user). Borrowing the classical metaphor, we refer to the efficiency of such transmissions as 3 birds, 1 stone. Transmissions corresponding to λ_{ij} fall in subspaces $\mathbf{U}_{ij}$ and

[3]By definition the indices ij and ji are interchangeable.

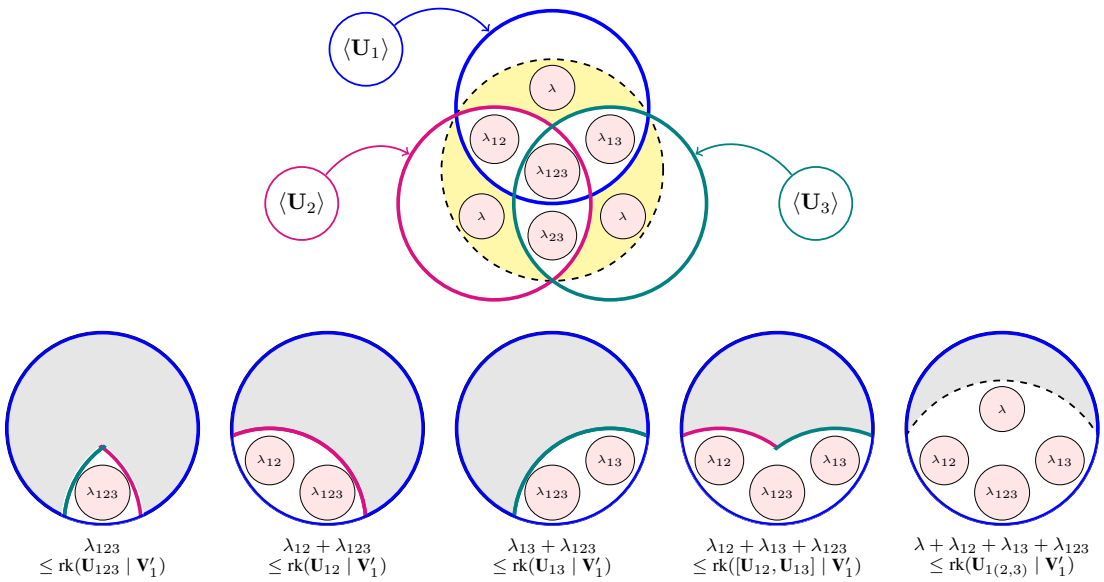

Fig. 4: Intuitive understanding of the constraints (21)-(24). Note that the blue circles in the bottom row of the figure do not show $\langle \mathbf{U}_1 \rangle$ per se, rather they show the space $\langle \mathbf{U}_1 \rangle$ conditioned on User 1's side-information, so that the corresponding sizes of all subspaces are represented with conditional ranks after conditioning on $\mathbf{V}_1'$. Each user will generally have a different perspective due to different impact of conditioning on their respective side-informations, giving rise to different constraints (21)-(24).

are simultaneously useful for Users i and j, so the efficiency of such transmissions is similarly referred to as 2 birds, 1 stone. In other words, λ_{ij} transmitted dimensions satisfy $2\lambda_{ij}$ dimensions of demand. Transmissions corresponding to λ fall in the three subspaces highlighted in yellow in Figure 3 where we previously noted that any two subspaces are disjoint but contain the third. What this means is that the coding scheme needs to send any 2 of the 3 subspaces marked with λ, and the third can be automatically inferred from them. Thus, a transmission of 2λ dimensions, satisfies a total of 3λ dimensions of demand (λ per user), yielding an efficiency of 3 birds, 2 stones.

Remark 5.2.2 In light of the previous remark, now consider the objective to be minimized in (20), $\Delta^* = \text{rk}(\mathbf{V}_1|\mathbf{V}_1') + \text{rk}(\mathbf{V}_2|\mathbf{V}_2') + \text{rk}(\mathbf{V}_3|\mathbf{V}_3') - 2\lambda_{123} - \lambda_{12} - \lambda_{13} - \lambda_{23} - \lambda$. We recognize the sum of the first three terms as the broadcast cost if the users were to be served separately and no gain in efficiency was possible by jointly satisfying multiple demands. Let this be our baseline. Now note that because $3\lambda_{123}$ dimensions of demand were satisfied with λ_{123} dimensions of broadcast, the cost-saving incurred relative to the baseline is $2\lambda_{123}$, which explains the fourth term that appears as a negative term in the objective. The next three negative terms are similarly justified because each λ_{ij} dimensions of transmission satisfies $2\lambda_{ij}$ dimensions of demand, thus saving λ_{ij} relative to the baseline. Finally, for the λ term, we recall that a total transmission cost of 2λ dimensions is able to satisfy 3λ dimensions of demand, thus saving another λ in broadcast cost, which explains the last negative term in the objective function.

Remark 5.2.3 As a caveat, note that the intuitive explanation above ignores a critical aspect of the problem that remains challenging — namely, each user's view of useful dimensions depends on their own side-information, and is in general different from other users. This is indicated in Figure 4 by noting that the relevant signal spaces for User 1 are not simply the $\mathbf{U}_\bullet$ spaces that appear in the decomposition at the top of Figure 3 and Figure 4. Rather, each user's view of useful subspaces is conditional on his side-information. For example, the same signal space $\mathbf{U}_{123}$ when seen by the Users $1, 2, 3$, contains $\text{rk}(\mathbf{U}_{123} \mid \mathbf{V}_1'), \text{rk}(\mathbf{U}_{123} \mid \mathbf{V}_2'), \text{rk}(\mathbf{U}_{123} \mid \mathbf{V}_3')$ useful dimensions, respectively. Thus, the total number of dimensions useful to all three users, i.e., the size of λ_{123} is limited by the bound in (21). Even with the size of λ_{123} constrained in this manner, finding the broadcast dimension is not trivial because each user may find a different λ_{123} portion of $\mathbf{U}_{123}$ useful to them. Similar challenges arise in identifying λ_{ij} dimensions that are useful to Users i and j, when each user's perspective is different, conditioned on their own side-information. Even greater care has to be taken in identifying the λ sections of the broadcast signal, to ensure that 2 transmissions span the third, while facing the challenge that the projections of λ into each user's perspective are distinguished by their different side-informations.

Remark 5.2.4 Since linear optimizations over polymatroidal constraints allow greedy solutions [62] that can simplify dimensional analysis (see e.g., the DoF study in [63, Chapter 5]), it is worth noting that the constraints (21)-(24) do not specify a polymatroidal structure. To verify this with a toy example, suppose $\mathbf{V}_1 = \mathbf{V}_2' = \mathbf{V}_3' = [1, 1]^T$ and $\mathbf{V}_1' = \mathbf{V}_2 = \mathbf{V}_3 = [0, 0]^T$. Then we have the constraints, $\lambda_{123} \leq 0$, $\lambda_{12} + \lambda_{123} \leq 0$, $\lambda_{13} + \lambda_{123} \leq 0$ and $\lambda_{12} + \lambda_{13} + \lambda_{123} \leq 1$, which violate the polymatroidal structure.

C. Structure of Proofs

Theorem 1 and Theorem 2 are equivalent alternative forms of the same capacity result. We organize the proofs of these two theorems as follows. In Section VII we prove the converse (lower) bound for the optimal broadcast cost, i.e., $\Delta^* \geq \max\{\Delta_1, \Delta_2\}$. Then in Section VIII we prove the achievability (upper) bound $\Delta^* \leq F^*$. Finally in Section IX we prove that $F^* \leq \max\{\Delta_1, \Delta_2\}$. The three proofs together imply that $\Delta^* = F^* = \max\{\Delta_1, \Delta_2\}$, thus proving both Theorem 1 and Theorem 2.

VI. Toy Examples

In this section, we present simple toy examples that illustrate some of the ideas discussed previously, such as subspace decompositions and linear-programming tradeoffs between schemes with different communication efficiency (birds vs stones), some ideas that will be important later on in the construction of the general coding scheme, such as field extensions, vector coding, and mixing of dimensions, and some new insights, such as the insufficiency of entropic structure, and the need for functional submodularity. For these examples we use specialized notation for simplicity: $((\mathbf{W}'_i \to \mathbf{W}_i), i = 1, 2, 3)$ to specify the setting, A, B, C, D, E instead of x_1, x_2, x_3, x_4, x_5, and A_ℓ instead of $A(\ell)$.

Example 1 (3 birds, 1 stone)**.** *Consider $d = 3$ dimensional data $\mathbf{X}^T = (A, B, C)$ over $\mathbb{F}_3$, and $((A \to B + C), (B \to A + C), (C \to A + B))$. In other words, User 1 has A and wants $B + C$, User 2 has B and wants $A + C$, and User 3 has C and wants $A + B$. A signal space decomposition as in Figure 3 yields for this example,*

$\mathbf{B}_{123}$	$\mathbf{B}_{12}$	$\mathbf{B}_{13}$	$\mathbf{B}_{23}$	$\mathbf{B}_{1(2,3)}$	$\mathbf{B}_{2(1,3)}$	$\mathbf{B}_{3(1,2)}$
$A+B+C$	$-$	$-$	$-$	A	B	$A+B$

$\mathbf{B}_{1c}$	$\mathbf{B}_{2c}$	$\mathbf{B}_{3c}$
$-$	$-$	$-$

Note that for simplicity in these examples we indicate $\mathbf{B}_{123}$ as $A + B + C$ instead of the formal representation as the vector $[1, 1, 1]^T$ in the 3 dimensional data universe. The optimal broadcast cost is $\Delta^* = 1$, achieved with $L = 1$, $N = 1$, $\lambda_{12} = \lambda_{23} = \lambda_{13} = \lambda = 0$, $\lambda_{123} = 1$, by broadcasting $\mathbf{S} = (A + B + C)$.

Example 2 (2 birds, 1 stone, vector coding, insufficiency of entropic structure)**.** *Consider $d = 3$ dimensional data $\mathbf{X}^T = (A, B, C)$ over $\mathbb{F}_3$, and $((A \to B + C), (B \to A + C), (C \to A + 2B))$. A signal space decomposition yields,*

$\mathbf{B}_{123}$	$\mathbf{B}_{12}$	$\mathbf{B}_{13}$	$\mathbf{B}_{23}$
$-$	$A+B+C$	$A+2B+2C$	$A+2B+C$

$\mathbf{B}_{1(2,3)}$	$\mathbf{B}_{2(1,3)}$	$\mathbf{B}_{3(1,2)}$	$\mathbf{B}_{1c}$	$\mathbf{B}_{2c}$	$\mathbf{B}_{3c}$
$-$	$-$	$-$	$-$	$-$	$-$

The optimal broadcast cost is $\Delta^* = 1.5$, achieved with $L = 2$, $N = 3$, $\lambda_{123} = \lambda = 0, \lambda_{12} = \lambda_{13} = \lambda_{23} = 0.5$, by broadcasting $\mathbf{S} = (A_1 + B_1 + C_1, A_2 + 2B_2 + 2C_2, (A_1 + 2B_1 + C_1) + (A_2 + 2B_2 + C_2))$.

Evidently, 1.5 dimensions of broadcast satisfy a total of 3 dimensions of demand, as expected from a 2 birds, 1 stone setting. Also note that this example requires vector coding, i.e., we need $L > 1$. Most importantly, however, this example illustrates that unlike the 2 user LCBC, the entropic formulation of [40] is not enough for the 3 user LCBC. The following remark elaborates upon this observation.

Remark 6.1 Reference [40] considers an entropic formulation of the LCBC that is summarized as follows. The data $\mathbf{X}$ is assumed to be i.i.d. uniform, $\mathcal{W}^* \triangleq \{\mathbf{W}_1, \mathbf{W}'_1, \cdots, \mathbf{W}_K, \mathbf{W}'_K\}$ denotes the set of all $2K$ demand and side-information random variables, the entropies $H(\mathcal{W})$ are specified for all $2^{2K} - 1$ non-empty subsets of random variables $\mathcal{W} \subset \mathcal{W}^*$, the encoding constraint is represented as $H(\mathbf{S} \mid \mathcal{W}^*) = 0$, and the decoding constraints are represented as $H(\mathbf{W}_k \mid \mathbf{S}, \mathbf{W}'_k) = 0$ for all $k \in [K]$. Subject to these entropy specifications, as well as standard (Shannon and non-Shannon) information inequalities, the goal is to minimize the entropy $H(\mathbf{S})$. As discussed in Section III-A3, such a formulation produces a lower bound on the download cost, as $N \geq H(\mathbf{S})$, which in turn yields a lower bound on Δ^*. For the $K = 2$ user LCBC, this bound turns out to be tight. Remarkably, however, the same approach does not work for the $K = 3$ user LCBC, as we argue based on Example 1 and Example 2. Although a bit tedious, it is not difficult to verify that all $2^6 - 1 = 63$ entropies $H(\mathcal{W})$ match for Example 1 and Example 2. For example, consider $\mathcal{W} = \{\mathbf{W}'_1, \mathbf{W}_3\}$. Note that $H(\mathcal{W}) = H(\mathbf{W}'_1, \mathbf{W}_3) = H(A, A + B) = H(A, B) = 2L$ in Example 1, and $H(\mathcal{W}) = H(\mathbf{W}'_1, \mathbf{W}_3) = H(A, A + 2B) = H(A, B) = 2L$ in Example 2, so both examples have the same entropy for this $\mathcal{W}$. One can similarly compute $H(\mathcal{W})$ for all 63 non-empty subsets $\mathcal{W} \subset \mathcal{W}^*$ for both Example 1 and Example 2 and verify that in each case both examples produce matching entropies. Therefore, since all the entropic constraints for both examples are identical, and all Shannon and non-Shannon information inequalities apply to both examples, the two examples can only produce the same entropic lower bound on $H(\mathbf{S})$. However, we know that the two examples have different capacities. Example 1 has $\Delta^* = 1, C = 1$ while Example 2 has $\Delta^* = 1.5, C = 1/1.5 = 2/3$. Since Example 2 requires a strictly stronger bound (impossibility result) than Example 1 for a tight converse, it follows that the entropic formulation cannot yield a tight converse for Example 2. Indeed, the key to the converse bound $\Delta^* \geq 1.5$ for Example 2 is the *functional* submodularity property [56], [57] that takes into account the functional forms of the users' side-information and demands. A converse for Example 2 is explicitly provided in Section VII-B1.

Example 3 (3 birds, 2 stones, the user's perspective)**.** *Consider $d = 3$ dimensional data $\mathbf{X}^T = (A, B, C)$ over $\mathbb{F}_2$, and $((A \to B), (B \to C), (C \to A))$. A signal space decomposition as in Figure 3 yields,*

$\mathbf{B}_{123}$	$\mathbf{B}_{12}$	$\mathbf{B}_{13}$	$\mathbf{B}_{23}$	$\mathbf{B}_{1(2,3)}$	$\mathbf{B}_{2(1,3)}$	$\mathbf{B}_{3(1,2)}$
$-$	B	A	C	$-$	$-$	$-$

$\mathbf{B}_{1c}$	$\mathbf{B}_{2c}$	$\mathbf{B}_{3c}$
$-$	$-$	$-$

This coincides with an index coding problem, the optimal broadcast cost is $\Delta^* = 2$, achieved with $L = 1$, $N = 2$,

$\lambda_{123} = \lambda_{12} = \lambda_{13} = \lambda_{23} = 0, \lambda = 1$, *by broadcasting* $\mathbf{S} = (A+B, B+C)$.

This example also highlights the importance of the users' individual perspectives conditioned on their side-information. Without accounting for side-information, the signal space decomposition of Figure 3 suggests that all the signals reside in $\mathbf{U}_{12}, \mathbf{U}_{13}, \mathbf{U}_{23}$, which might suggest 2 birds, 1 stone schemes with $\lambda_{12} = \lambda_{13} = \lambda_{23} = 0.5$ and a download cost of $\Delta^* = 1.5$. However, this is not achievable, as we note the optimal download cost is $\Delta^* = 2$. To see this, consider individual users' perspectives. For example, User 1 requires $\lambda_{13} + \lambda_{123} \leq \text{rk}(\mathbf{U}_{13} \mid \mathbf{V}_1')$. Now since both $\mathbf{U}_{13}$ and $\mathbf{V}_1'$ correspond to the data dimension A, this conditional rank is 0. In other words, even though the subspace $\mathbf{U}_{13}$ has one dimension that may suggest the opportunity to simultaneously satisfy users 1 and 3, this dimension happens to be already available to User 1. Thus, upon taking into account User 1's side-information, there is no such opportunity. We end up with $\lambda_{123} = \lambda_{12} = \lambda_{13} = \lambda_{23} = 0$, and $\lambda = 1$. Out of the 3 dimensions, say $A+B, B+C, C+A$, any two yield the third by summation (over $\mathbb{F}_2$), and it suffices to send any 2 to satisfy all 3 users. Notice the need to mix up the dimensions, appealing to mixed dimensions (similar to random coding arguments) will be a key idea to develop the general coding scheme.

As noted, Example 2 used vector coding ($L > 1$) to achieve the optimal download cost $\Delta^* = 1.5$. Vector coding may be strictly necessary even in cases where the optimal download cost Δ^* is an integer value, as illustrated by the next example. The necessity of vector coding for the 3 user LCBC is remarkable because scalar coding was found to be sufficient for the 2 user LCBC in [40]).

Example 4 (Field size extension)**.** *Consider* $d = 2$ *dimensional data* $\mathbf{X}^T = (A, B)$ *over* $\mathbb{F}_2$*, and* $((A \to B), (B \to A+B), (A+B \to A))$*. A signal space decomposition as in Figure 3 yields,*

$\mathbf{B}_{123}$	$\mathbf{B}_{12}$	$\mathbf{B}_{13}$	$\mathbf{B}_{23}$	$\mathbf{B}_{1(2,3)}$	$\mathbf{B}_{2(1,3)}$	$\mathbf{B}_{3(1,2)}$
A, B	$-$	$-$	$-$	$-$	$-$	$-$

$\mathbf{B}_{1c}$	$\mathbf{B}_{2c}$	$\mathbf{B}_{3c}$
$-$	$-$	$-$

We have $\Delta^* = 1$*, achieved with* $L = 2$*,* $N = 2$*,* $\lambda_{123} = 1, \lambda_{12} = \lambda_{13} = \lambda_{23} = \lambda = 0$*,* $\mathbf{S} = (A_1 + A_2 + B_2, A_1 + B_1 + B_2)$.

Appendix A shows that $\Delta^* = 1$ is not achievable with scalar coding, i.e., neither scalar linear nor scalar non-linear coding scheme can achieve $\Delta = 1$ for $L = 1$ computation for this example. However, $\Delta^* = 1$ and can be achieved for $L = 2$ computations with $N = 2$. In this case, because $\lambda_{123} = 1$, we would like to broadcast one dimension. In the scalar code setting $L = 1$, this one dimension can be found for each pair of users but it cannot be the same for the three users simultaneously. To see this, note that $A + B$ helps User 1 and User 2 but not User 3; B helps User 1 and User 3 but not User 2; A helps User 2 and User 3 but not User 1. Aside from the time-sharing type vector coding solution shown for Example 2, another approach is to consider $L > 1$ (which implies a vector code) and use a scalar code in a larger extended field $\mathbb{F}_{2^z}$ (in general $\mathbb{F}_{q^z}$). For this example, with $L = 2$, we can use a scalar code over $\mathbb{F}_4 = \mathbb{F}_2[x]/(x^2+x+1)$, which results in $N = 2$ in $\mathbb{F}_2$. Representing $A = A_1 + A_2 x \in \mathbb{F}_4$, $B = B_1 + B_2 x \in \mathbb{F}_4$, the transmitted symbol is simply $(1+x)A + xB \mod (x^2+x+1) = (A_1 + A_2 + B_2) + x(A_2 + B_1 + B_2)$ which corresponds to the transmitted symbol $\mathbf{S} = (A_1 + A_2 + B_2, A_1 + B_1 + B_2)$. Additional discussion can be found in Appendix A as well. Indeed, field extensions are a key element of the general coding scheme.

Example 5 (Inseparability)**.** *Consider* $d = 5$ *dimensional data* $\mathbf{X}^T = (A, B, C, D, E)$ *over* $\mathbb{F}_3$*, and* $((A \to [B + C, D]), (B \to [A+C, E]), ([C, D+E] \to A+2B))$*. A signal space decomposition as in Figure 3 yields,*

$\mathbf{B}_{123}$	$\mathbf{B}_{12}$	$\mathbf{B}_{13}$	$\mathbf{B}_{23}$
$-$	$A+B+C$	$A+2B+2C$	$A+2B+C$

$\mathbf{B}_{1(2,3)}$	$\mathbf{B}_{2(1,3)}$	$\mathbf{B}_{3(1,2)}$	$\mathbf{B}_{1c}$	$\mathbf{B}_{2c}$	$\mathbf{B}_{3c}$
D	E	$D+E$	$-$	$-$	$-$

We have $\Delta^* = 3$*, achieved with* $L = 1$*,* $N = 3$*,* $\lambda_{123} = \lambda_{13} = \lambda_{23} = 0, \lambda_{12} = \lambda = 1$*, by broadcasting* $\mathbf{S} = (A+B+C, A+D, 2B+E)$.

Note that this problem combines Example 2 for data (A, B, C) and another LCBC instance with data (D, E) where User 1 wants D, User 2 wants E and User 3 knows $D + E$. Separately, these problems have download costs of 1.5 and 2, respectively. Since the two problems deal with independent data, one might expect the solution to be separable, however a separate solution would have a total broadcast cost of $1.5 + 2 = 3.5$. The optimal $\Delta^* = 3$, which is better than 3.5, thus showing that even though an LCBC problem may be a composition of instances with separate datasets, in general a separate solution would be suboptimal. This observation also underscores why the tradeoffs in LCBC, that we see represented in the linear program, are non-trivial.

VII. Proof of Converse: $\Delta^* \geq \max\{\Delta_1, \Delta_2\}$

The converse is comprised of the two bounds, $\Delta^* \geq \Delta_1$, and $\Delta^* \geq \Delta_2$. The first bound, $\Delta^* \geq \Delta_1$ is a straightforward generalization of the corresponding bound for the 2 user LCBC found in [40] to the 3 user setting. The second bound, $\Delta^* \geq \Delta_2$ is novel, and requires functional submodularity. For the sake of completeness in this section we present the proof of both bounds. Let us begin by recalling the functional submodularity property.

Lemma 3 (Functional submodularity of Shannon entropy (Lemma A.2 of [56]))**.** *If* X_0, X_1, X_2, X_{12} *are random variables such that* X_1 *and* X_2 *each determine* X_0 *and* (X_1, X_2) *determine* X_{12}*, then:*

$$H(X_1) + H(X_2) \geq H(X_{12}) + H(X_0) \tag{25}$$

Note that 'A determines B' as used in Lemma 3 is equivalent to the statement that $H(B \mid A) = 0$, i.e., B is a function of A. Thus, the lemma assumes that $H(X_0 \mid X_1) = H(X_0 \mid X_2) = H(X_{12} \mid X_1, X_2) = 0$.

As an immediate corollary, let us note the following form in which we will apply the functional submodularity.

Corollary 1. *For arbitrary matrices* $\mathbf{M}_1 \in \mathbb{F}_q^{d\times\mu_1}, \mathbf{M}_2 \in \mathbb{F}_q^{d\times\mu_2}$*, any random matrix* $\mathbf{X} \in \mathbb{F}_q^{d\times L}$*, and any random variable* $\mathbf{Z}$*,*

$$\begin{aligned} &H(\mathbf{Z}, \mathbf{X}^T\mathbf{M}_1) + H(\mathbf{Z}, \mathbf{X}^T\mathbf{M}_2) \\ &\geq H(\mathbf{Z}, \mathbf{X}^T[\mathbf{M}_1 \cap \mathbf{M}_2]) + H(\mathbf{Z}, \mathbf{X}^T[\mathbf{M}_1, \mathbf{M}_2]) \end{aligned} \tag{26}$$

Proof. The corollary follows from Lemma 3 by setting $X_1 = (\mathbf{Z}, \mathbf{X}^T\mathbf{M}_1)$, $X_2 = (\mathbf{Z}, \mathbf{X}^T\mathbf{M}_2)$, and noting that $X_0 = (\mathbf{Z}, \mathbf{X}^T[\mathbf{M}_1 \cap \mathbf{M}_2])$ can be obtained as a function of both X_1 and X_2 individually, while $X_{12} = (\mathbf{Z}, \mathbf{X}^T[\mathbf{M}_1, \mathbf{M}_2])$ is a function of (X_1, X_2). □

A. Proof of the bound: $\Delta^* \geq \Delta_1$

As noted, the proof of this bound is straightforward. It follows along the same lines as the proof for the 2 user LCBC in [40], also similar to the genie-aided bound in coded caching (e.g., [61, (71)-(75)]) and is provided here for the sake of completeness. In particular, it does not require functional submodularity. As explained in Section III-A3, recall that the converse bound is based on a thought-experiment that supposes that the data $\mathbf{X}$ is i.i.d. uniform, which leads to a lower bound $N \geq H(\mathbf{S})$. Let $\mathbf{W}_k^* \triangleq (\mathbf{W}_k, \mathbf{W}_k'), \forall k \in [3]$. The bound follows essentially by iteratively using the argument

$$\begin{aligned} &H(\mathbf{S} \mid \mathbf{W}_k', \mathbf{W}_{[k-1]}^*) \\ &\geq \underbrace{H(\mathbf{W}_k \mid \mathbf{W}_k', \mathbf{W}_{[k-1]}^*)}_{\text{The genie-aided bound for the } k^{th} \text{ user}} + H(\mathbf{S} \mid \mathbf{W}_{k+1}', \mathbf{W}_{[k]}^*), \end{aligned} \tag{27}$$

since $H(\mathbf{S} \mid \mathbf{W}_k', \mathbf{W}_{[k-1]}^*) = H(\mathbf{S}, \mathbf{W}_k \mid \mathbf{W}_k', \mathbf{W}_{[k-1]}^*) \geq H(\mathbf{W}_k \mid \mathbf{W}_k', \mathbf{W}_{[k-1]}^*) + H(\mathbf{S} \mid \mathbf{W}_{k+1}', \mathbf{W}_{[k]}^*)$, where the first step uses the decoder definition (4) and the second step applies the chain rule of entropy and the fact that conditioning reduces entropy. It then follows that for any coding scheme $\big(L, N, \Phi, (\Psi_k)_{k\in[3]}\big) \in \mathfrak{C}$,

$$\begin{aligned} N &\geq H(\mathbf{S}) \geq H(\mathbf{S} \mid \mathbf{W}_1') && (28)\\ &\geq H(\mathbf{W}_1 \mid \mathbf{W}_1') + H(\mathbf{S} \mid \mathbf{W}_2', \mathbf{W}_1^*) && (29)\\ &\geq H(\mathbf{W}_1 \mid \mathbf{W}_1') + H(\mathbf{W}_2 \mid \mathbf{W}_2', \mathbf{W}_1^*) \\ &\quad + H(\mathbf{S} \mid \mathbf{W}_3', \mathbf{W}_1^*, \mathbf{W}_2^*) && (30)\\ &\geq H(\mathbf{W}_1 \mid \mathbf{W}_1') + H(\mathbf{W}_2 \mid \mathbf{W}_2', \mathbf{W}_1^*) \\ &\quad + H(\mathbf{W}_3 \mid \mathbf{W}_3', \mathbf{W}_1^*, \mathbf{W}_2^*) && (31)\\ &= L \cdot \big(\text{rk}(\mathbf{V}_1 \mid \mathbf{V}_1') + \text{rk}(\mathbf{V}_2 \mid [\mathbf{U}_1, \mathbf{V}_2']) \\ &\quad + \text{rk}(\mathbf{V}_3 \mid [\mathbf{U}_1, \mathbf{U}_2, \mathbf{V}_3'])\big) && (32)\\ &= L \cdot \big(\text{rk}(\mathbf{V}_1 \mid \mathbf{V}_1') + \text{rk}(\mathbf{V}_2 \mid \mathbf{V}_2') - \text{rk}(\mathbf{U}_{12} \mid \mathbf{V}_2') \\ &\quad + \text{rk}(\mathbf{V}_3 \mid \mathbf{V}_3') - \text{rk}(\mathbf{U}_{3(1,2)} \mid \mathbf{V}_3')\big) && (33)\\ &\implies \Delta = N/L \geq \Delta_1^{123} && (34) \end{aligned}$$

Steps (29) – (31) follow from (27). Step (32) uses the fact that for i.i.d. uniform data $\mathbf{X}^T \in \mathbb{F}_q^{L\times d}$ and an arbitrary matrix $\mathbf{M} \in \mathbb{F}_q^{d\times\mu}$, we have $H(\mathbf{X}^T\mathbf{M}) = L \cdot \text{rk}(\mathbf{M})$ in q-ary units, and applies the conditional-rank notation as defined in Section II. Step (33) follows from the observation that $\text{rk}(\mathbf{V}_k \mid [\mathbf{Z}, \mathbf{V}_k']) = \text{rk}([\mathbf{U}_k, \mathbf{Z}]) - \text{rk}([\mathbf{V}_k', \mathbf{Z}]) = \text{rk}(\mathbf{U}_k) - \text{rk}(\mathbf{V}_k') - \Big(\text{rk}(\mathbf{U}_k \cap \mathbf{Z}) - \text{rk}(\mathbf{V}_k' \cap \mathbf{Z})\Big) = \text{rk}(\mathbf{V}_k \mid \mathbf{V}_k') - \text{rk}(\mathbf{U}_k \cap \mathbf{Z} \mid \mathbf{V}_k')$. Similarly, $\Delta \geq \Delta_1^{ijk}$ for all (i, j, k) that are permutations of $(1, 2, 3)$. Since this holds for every coding scheme $\big(L, N, \Phi, (\Psi_k)_{k\in[3]}\big) \in \mathfrak{C}$, it follows that $\Delta^* \geq \Delta_1$. □

B. Proof of the bound: $\Delta^* \geq \Delta_2$

The main idea of this proof is to successfully identify and introduce the entropies of certain (linear) *functions* of users' demands and side-information that are critical in determining the capacity, with the application of Lemma 3. To build intuition, let us start with the converse proof for a toy example, specifically Example 2 of Section VI.

1) Converse Proof for a Toy Example: Consider any coding scheme $\big(L, N, \Phi, (\Psi_k)_{k\in[3]}\big) \in \mathfrak{C}$ for Example 2 of Section VI. Recall that User 1 has A_ℓ and wants $B_\ell + C_\ell$; User 2 has B_ℓ and wants $A_\ell + C_\ell$; User 3 has C_ℓ and wants $A_\ell + 2B_\ell$ for all $\ell \in [L]$. We want to prove the converse bound $\Delta^* \geq 1.5$. Let us denote A as $(A_1, \cdots, A_L) \in \mathbb{F}_q^{1\times L}$, B as $(B_1, \cdots, B_L) \in \mathbb{F}_q^{1\times L}$ and C as $(C_1, \cdots, C_L) \in \mathbb{F}_q^{1\times L}$. As mentioned in Section III-A3 let us start the converse proof with the thought-experiment that A, B, C are i.i.d. uniform in $\mathbb{F}_q$, which allows the following entropic arguments.

$$\begin{aligned} &2H(\mathbf{S}) + H(A) + H(B) - I(\mathbf{S}; A) - I(\mathbf{S}; B) \\ &= H(\mathbf{S}, A) + H(\mathbf{S}, B) && (35)\\ &\geq H(\mathbf{S}, A, B + C) + H(\mathbf{S}, B, A + C) && (36)\\ &\geq H(\mathbf{S}, A + B + C) + H(\mathbf{S}, A, B, C) && (37)\\ &\geq H(\mathbf{S}) + H(A + B + C \mid \mathbf{S}) + H(A, B, C). && (38) \end{aligned}$$

Similarly,

$$\begin{aligned} &2H(\mathbf{S}) + H(A) + H(C) - I(\mathbf{S}; A) - I(\mathbf{S}; C) \\ &= H(\mathbf{S}, A) + H(\mathbf{S}, C) && (39)\\ &\geq H(\mathbf{S}, A, B + C) + H(\mathbf{S}, C, A + 2B) && (40)\\ &\geq H(\mathbf{S}, A + 2B + 2C) + H(\mathbf{S}, A, B, C) && (41)\\ &\geq H(\mathbf{S}) + H(A + 2B + 2C \mid \mathbf{S}) + H(A, B, C). && (42) \end{aligned}$$

Steps (35) and (39) use the definition of mutual information $I(A; B) = H(A) + H(B) - H(A, B)$. Steps (36) and (40) use the decoder definition (4). Step (37) uses functional submodularity (Lemma 3) by recognizing that $(A, B + C)$ and $(B, A + C)$ each determine $A + B + C$, and $(A, B + C, B, A + C)$ determines (A, B, C). Step (41) uses functional submodularity by recognizing that $(A, B+C)$ and $(C, A+2B)$ each determine $A + 2B + 2C$, and $(A, B + C, C, A + 2B)$ determines (A, B, C).

Adding the above two inequalities, we have

$$\begin{aligned} &4H(\mathbf{S}) + 2H(A) + H(B) + H(C) \\ &\quad - 2I(\mathbf{S}; A) - I(\mathbf{S}; B) - I(\mathbf{S}; C) \\ &\geq 2H(\mathbf{S}) + H(A + B + C \mid \mathbf{S}) \\ &\quad + H(A + 2B + 2C \mid \mathbf{S}) + 2H(A, B, C). \end{aligned} \tag{43}$$

It follows that

$$2H(\mathbf{S}) \geq H(A + B + C \mid \mathbf{S}) + H(A + 2B + 2C \mid \mathbf{S})$$

$$+ I(\mathbf{S}; A) + 2L \tag{44}$$
$$\geq H(A + B + C, A + 2B + 2C \mid \mathbf{S}) + I(\mathbf{S}; A) + 2L \tag{45}$$
$$\geq H(A \mid \mathbf{S}) + I(\mathbf{S}; A) + 2L \tag{46}$$
$$= H(A) + 2L \tag{47}$$
$$= 3L. \tag{48}$$

Step (44) and (48) apply the assumption that A, B, C are i.i.d. uniform in $\mathbb{F}_q$. Step (45) uses the general property of joint entropy that $H(X \mid Z) + H(Y \mid Z) \geq H(X, Y \mid Z)$ for any random variables X, Y, Z. Step (46) is obtained by recognizing that A is a function of $(A + B + C, A + 2B + 2C)$. Step (47) uses the information equality $I(A; B) = H(A) - H(A|B)$. Therefore, we have the desired converse bound, $\Delta = N/L \geq H(\mathbf{S})/L \geq 1.5$ for the coding scheme. Since this is true for every feasible coding scheme, we have the bound $\Delta^* \geq 1.5$.

2) General Proof of Converse Bound $\Delta^* \geq \Delta_2$: As mentioned in Section III-A3 let us start the converse proof based on the thought-experiment that supposes the elements of the data $\mathbf{X}$ are i.i.d. uniform in $\mathbb{F}_q$.

$$2H(\mathbf{S}) + 2\sum_{k=1}^{3} H(\mathbf{W}'_k)$$
$$= 2\sum_{k=1}^{3} H(\mathbf{S}, \mathbf{W}'_k) + 2\sum_{k=1}^{3} I(\mathbf{S}; \mathbf{W}'_k) - 4H(\mathbf{S}) \tag{49}$$
$$= 2\sum_{k=1}^{3} H(\mathbf{S}, \mathbf{W}'_k, \mathbf{W}_k) + 2\sum_{k=1}^{3} I(\mathbf{S}; \mathbf{W}'_k) - 4H(\mathbf{S}) \tag{50}$$
$$= \sum_{(i,j)\in\{(1,2),(2,3),(1,3)\}} \Big(H(\mathbf{S}, \mathbf{X}^T\mathbf{U}_i) + H(\mathbf{S}, \mathbf{X}^T\mathbf{U}_j)\Big) + 2\sum_{k=1}^{3} I(\mathbf{S}; \mathbf{W}'_k) - 4H(\mathbf{S}) \tag{51}$$
$$\geq \big[H(\mathbf{S}, \mathbf{X}^T\mathbf{U}_{12}) + H(\mathbf{S}, \mathbf{X}^T\mathbf{U}_{13})\big] + \big[H(\mathbf{S}, \mathbf{X}^T[\mathbf{U}_1, \mathbf{U}_2]) + H(\mathbf{S}, \mathbf{X}^T\mathbf{U}_3)\big] + \big[H(\mathbf{S}, \mathbf{X}^T[\mathbf{U}_1, \mathbf{U}_3]) + H(\mathbf{S}, \mathbf{X}^T\mathbf{U}_2)\big] + 2\sum_{k=1}^{3} I(\mathbf{S}; \mathbf{W}'_k) - 4H(\mathbf{S}) \tag{52}$$
$$\geq H(\mathbf{S}, \mathbf{X}^T\mathbf{U}_{123}) + H(\mathbf{S}, \mathbf{X}^T[\mathbf{U}_{12}, \mathbf{U}_{13}]) + H(\mathbf{S}, \mathbf{X}^T\mathbf{U}_{3(1,2)}) + H(\mathbf{S}, \mathbf{X}^T\mathbf{U}_{2(1,3)}) + 2H(\mathbf{S}, \mathbf{X}^T[\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3]) + 2\sum_{k=1}^{3} I(\mathbf{S}; \mathbf{W}'_k) - 4H(\mathbf{S}) \tag{53}$$
$$\geq H(\mathbf{X}^T\mathbf{U}_{123}|\mathbf{S}) + H(\mathbf{X}^T[\mathbf{U}_{12}, \mathbf{U}_{13}]|\mathbf{S}) + H(\mathbf{X}^T\mathbf{U}_{3(1,2)}|\mathbf{S}) + H(\mathbf{X}^T\mathbf{U}_{2(1,3)}|\mathbf{S}) + 2H(\mathbf{X}^T[\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3]) + 2\sum_{k=1}^{3} I(\mathbf{S}; \mathbf{W}'_k) \tag{54}$$
$$\geq \max_{\ell\in\{1,2,3\}} H(\mathbf{X}^T(\mathbf{U}_{123} \cap \mathbf{V}'_\ell)|\mathbf{S}) + H(\mathbf{X}^T[\mathbf{U}_{12}, \mathbf{U}_{13}]|\mathbf{S}) + H(\mathbf{X}^T\mathbf{U}_{3(1,2)}|\mathbf{S}) + H(\mathbf{X}^T\mathbf{U}_{2(1,3)}|\mathbf{S}) + 2H(\mathbf{X}^T[\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3]) + 2\sum_{k=1}^{3} I(\mathbf{S}; \mathbf{W}'_k) \tag{55}$$
$$\geq \max_{\ell\in\{1,2,3\}} H(\mathbf{X}^T(\mathbf{U}_{123} \cap \mathbf{V}'_\ell)) + H(\mathbf{X}^T[\mathbf{U}_{12}, \mathbf{U}_{13}]|\mathbf{S}) + H(\mathbf{X}^T\mathbf{U}_{3(1,2)}|\mathbf{S}) + H(\mathbf{X}^T\mathbf{U}_{2(1,3)}|\mathbf{S}) + 2H(\mathbf{X}^T[\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3]) + \sum_{k=1}^{3} I(\mathbf{S}; \mathbf{W}'_k) \tag{56}$$
$$\geq \max_{\ell\in\{1,2,3\}} H(\mathbf{X}^T(\mathbf{U}_{123} \cap \mathbf{V}'_\ell)) + H(\mathbf{X}^T([\mathbf{U}_{12}, \mathbf{U}_{13}] \cap \mathbf{V}'_1)) + H(\mathbf{X}^T(\mathbf{U}_{3(1,2)} \cap \mathbf{V}'_3)) + H(\mathbf{X}^T(\mathbf{U}_{2(1,3)} \cap \mathbf{V}'_2)) + 2H(\mathbf{X}^T[\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3]) \tag{57}$$

In the deduction, the most critical steps are Step (52) and Step (53). Specifically, Step (52) uses functional submodularity property from Corollary 1 twice, once for $(i, j) = (1, 2)$ and once for $(i, j) = (1, 3)$. Step (53) uses functional submodularity from Corollary 1 three times, once for each of the collections of terms inside the three square parantheses in (52), making use of the fact that $\mathbf{U}_{12} \cap \mathbf{U}_{13} = \mathbf{U}_{123}$, $[\mathbf{U}_1, \mathbf{U}_2] \cap \mathbf{U}_3 = \mathbf{U}_{3(1,2)}$, and $[\mathbf{U}_1, \mathbf{U}_3] \cap \mathbf{U}_2 = \mathbf{U}_{2(1,3)}$ by definition. The other steps follow from conventional entropic inequalities. Specifically, Step (49) uses the definition of mutual information $I(A; B) = H(A) + H(B) - H(A, B)$. Step (50) uses the decoder definition (4), i.e., $\mathbf{W}_k$ is a function of $(\mathbf{S}, \mathbf{W}'_k)$. Step (51) uses the definition of $\mathbf{U}_k = [\mathbf{V}'_k, \mathbf{V}_k]$ from (11) to recognize $\mathbf{X}^T\mathbf{U}_k = [\mathbf{W}'_k, \mathbf{W}_k]$. Step (54) uses the chain rule of entropy to extract $H(\mathbf{S})$ from the first four terms, and the property that $H(A, B) \geq H(B)$ to drop $\mathbf{S}$ from the fifth term. Step (55) uses the property that $H(A \mid B) \geq H(f(A) \mid B)$, and the fact that $\mathbf{X}^T(\mathbf{U}_{123} \cap \mathbf{V}'_\ell)$ is a function of $\mathbf{X}^T\mathbf{U}_{123}$. Step (56) uses the fact that $H(\mathbf{X}^T(\mathbf{U}_{123} \cap \mathbf{V}'_\ell)|\mathbf{S}) = H(\mathbf{X}^T(\mathbf{U}_{123} \cap \mathbf{V}'_\ell)) - I(\mathbf{S}; \mathbf{X}^T(\mathbf{U}_{123} \cap \mathbf{V}'_\ell))$ by definition of mutual information, and $I(\mathbf{S}; \mathbf{X}^T(\mathbf{U}_{123} \cap \mathbf{V}'_\ell)) \leq I(\mathbf{S}; \mathbf{X}^T\mathbf{V}'_\ell) = I(\mathbf{S}; \mathbf{W}'_\ell) \leq \sum_{k=1}^{3} I(\mathbf{S}; \mathbf{W}'_k)$ by data-processing inequality, and the non-negativity of mutual information. Similar reasoning is applied to the third, fourth and fifth terms of (56) to obtain (57) by removing the conditioning on $\mathbf{S}$ and by absorbing one of the $I(\mathbf{S}; \mathbf{W}'_k)$ terms each. The reasoning can be summarized as $H(A \mid S) + I(B; S) \geq H(C \mid S) + I(C; S) = H(C)$ if C is a function of both A and B individually.

Evaluating the entropies in terms of the corresponding ranks, and normalizing by L, we obtain,

$$\Delta = N/L \geq H(\mathbf{S})/L$$
$$\geq \frac{1}{2}\Big(\max_{l\in\{1,2,3\}} \mathrm{rk}((\mathbf{U}_{123} \cap \mathbf{V}'_l)) + \mathrm{rk}([\mathbf{U}_{12}, \mathbf{U}_{13}] \cap \mathbf{V}'_1) + \mathrm{rk}(\mathbf{U}_{3(1,2)} \cap \mathbf{V}'_3) + \mathrm{rk}(\mathbf{U}_{2(1,3)} \cap \mathbf{V}'_2)\Big) + \mathrm{rk}([\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3]) - \sum_{k=1}^{3} \mathrm{rk}(\mathbf{V}'_k) \tag{58}$$
$$= \frac{1}{2}\Big(\mathrm{rk}(\mathbf{U}_{123}) - \min_{\ell\in\{1,2,3\}} \mathrm{rk}(\mathbf{U}_{123} \mid \mathbf{V}'_\ell) + \mathrm{rk}([\mathbf{U}_{12}, \mathbf{U}_{13}]) - \mathrm{rk}([\mathbf{U}_{12}, \mathbf{U}_{13}] \mid \mathbf{V}'_1) + \mathrm{rk}(\mathbf{U}_{3(1,2)}) - \mathrm{rk}(\mathbf{U}_{3(1,2)} \mid \mathbf{V}'_3) + \mathrm{rk}(\mathbf{U}_{2(1,3)}) - \mathrm{rk}(\mathbf{U}_{2(1,3)} \mid \mathbf{V}'_2)\Big)$$

$$+ \text{rk}([\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3]) - \sum_{k=1}^{3} \text{rk}(\mathbf{V}'_k) \tag{59}$$

$$= \frac{1}{2}\Big(\text{rk}(\mathbf{U}_{12}) + \text{rk}(\mathbf{U}_{13}) - \min_{\ell\in\{1,2,3\}} \text{rk}(\mathbf{U}_{123} \mid \mathbf{V}'_\ell) - \text{rk}([\mathbf{U}_{12}, \mathbf{U}_{13}] \mid \mathbf{V}'_1) + \text{rk}(\mathbf{U}_3) + \text{rk}([\mathbf{U}_1, \mathbf{U}_2]) - \text{rk}(\mathbf{U}_{3(1,2)} \mid \mathbf{V}'_3) + \text{rk}(\mathbf{U}_2) + \text{rk}([\mathbf{U}_1, \mathbf{U}_3]) - \text{rk}(\mathbf{U}_{2(1,3)} \mid \mathbf{V}'_2)\Big) - \sum_{k=1}^{3} \text{rk}(\mathbf{V}'_k) \tag{60}$$

$$= \text{rk}(\mathbf{U}_1) + \text{rk}(\mathbf{U}_2) + \text{rk}(\mathbf{U}_3) - \sum_{k=1}^{3} \text{rk}(\mathbf{V}'_k) - \frac{1}{2}\Big(\min_{\ell\in\{1,2,3\}} \text{rk}(\mathbf{U}_{123} \mid \mathbf{V}'_\ell) + \text{rk}([\mathbf{U}_{12}, \mathbf{U}_{13}] \mid \mathbf{V}'_1) + \text{rk}(\mathbf{U}_{3(1,2)} \mid \mathbf{V}'_3) + \text{rk}(\mathbf{U}_{2(1,3)} \mid \mathbf{V}'_2)\Big) \tag{61}$$

$$= \text{rk}(\mathbf{V}_1 \mid \mathbf{V}'_1) + \text{rk}(\mathbf{V}_2 \mid \mathbf{V}'_2) + \text{rk}(\mathbf{V}_3 \mid \mathbf{V}'_3) - \frac{1}{2}\Big(\min_{\ell\in\{1,2,3\}} \text{rk}(\mathbf{U}_{123} \mid \mathbf{V}'_\ell) + \text{rk}([\mathbf{U}_{12}, \mathbf{U}_{13}] \mid \mathbf{V}'_1) + \text{rk}(\mathbf{U}_{3(1,2)} \mid \mathbf{V}'_3) + \text{rk}(\mathbf{U}_{2(1,3)} \mid \mathbf{V}'_2)\Big) \tag{62}$$

Step (58) uses the fact that for i.i.d. uniform data $\mathbf{X}^T \in \mathbb{F}_q^{L\times d}$ and an arbitrary matrix $\mathbf{M} \in \mathbb{F}_q^{d\times\mu}$, $H(\mathbf{X}^T\mathbf{M}) = L \cdot \text{rk}(\mathbf{M})$ in q-ary units. Step (59) applies the conditional-rank notation, $\text{rk}(A \mid B) = \text{rk}([A, B]) - \text{rk}(B)$ as defined in Section II. Step (60) uses the fact that $\text{rk}(\mathbf{U}_{123}) = \text{rk}(\mathbf{U}_{12} \cap \mathbf{U}_{13}) = \text{rk}(\mathbf{U}_{12}) + \text{rk}(\mathbf{U}_{13}) - \text{rk}([\mathbf{U}_{12}, \mathbf{U}_{13}])$, similarly $\text{rk}(\mathbf{U}_{3(1,2)}) = \text{rk}(\mathbf{U}_3 \cap [\mathbf{U}_1, \mathbf{U}_2]) = \text{rk}(\mathbf{U}_3) + \text{rk}([\mathbf{U}_1, \mathbf{U}_2]) - \text{rk}([\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3])$, and by the same token $\text{rk}(\mathbf{U}_{2(1,3)}) = \text{rk}(\mathbf{U}_2 \cap [\mathbf{U}_1, \mathbf{U}_3]) = \text{rk}(\mathbf{U}_2) + \text{rk}([\mathbf{U}_1, \mathbf{U}_3]) - \text{rk}([\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3])$. Step (61) uses the fact that $\text{rk}(\mathbf{U}_{ij}) = \text{rk}(\mathbf{U}_i) + \text{rk}(\mathbf{U}_j) - \text{rk}([\mathbf{U}_i, \mathbf{U}_j])$. Finally, step (62) uses the fact that $\text{rk}(\mathbf{U}_i) - \text{rk}(\mathbf{V}'_i) = \text{rk}([\mathbf{V}'_i, \mathbf{V}_i]) - \text{rk}(\mathbf{V}'_i) = \text{rk}(\mathbf{V}_i \mid \mathbf{V}'_i)$.

Since this holds for every coding scheme $\big(L, N, \Phi, (\Psi_k)_{k\in[3]}\big) \in \mathfrak{C}$, it follows that $\Delta^* \geq \Delta_2^{123}$. Similarly, $\Delta^* \geq \Delta_2^{ijk}$, $\forall(i, j, k)$ that are permutations of $(1, 2, 3)$, which implies that $\Delta^* \geq \Delta_2$. □

VIII. Proof of Achievability: $\Delta^* \leq F^*$

In this section, we will construct a general scheme for the 3 user LCBC that achieves broadcast cost per computation equal to F^* as specified in the form of a linear program in Theorem 2, thus establishing an upper bound on the optimal broadcast cost per computation, $\Delta^* \leq F^*$. Finding an explicit solution to the linear program in closed form will be left for Section IX. We start this proof with some preliminary steps.

A. Eliminating Redundancies

As a first step let us eliminate redundancies, if any, that exist in the users' side-information and demands by removing redundant columns in $\mathbf{V}_k, \mathbf{V}'_k$ such that $\mathbf{U}_k = [\mathbf{V}'_k, \mathbf{V}_k]$ has full column rank for each $k \in [3]$. Essentially, we retain only linearly independent columns because the remaining columns either represent information desired by a user that is already available to the user (overlap between $\langle\mathbf{V}_k\rangle$ and $\langle\mathbf{V}'_k\rangle$), or information that is already accounted for by the independent columns (redundancies within $\mathbf{V}_k$ or within $\mathbf{V}'_k$). Thus, henceforth let us assume, without loss of generality, that

$$\text{rk}(\mathbf{U}_k) = \text{rk}([\mathbf{V}'_k, \mathbf{V}_k]) = \text{rk}(\mathbf{V}'_k) + \text{rk}(\mathbf{V}_k) = m'_k + m_k. \tag{63}$$

B. Field Size Extension

Recall that the problem formulation specifies a field $\mathbb{F}_q$, but allows us to choose the number of computations L to be encoded together as a free parameter in the achievable scheme. The freedom in the choice of L in fact allows field extensions that translate the specified field of operations from $\mathbb{F}_q$ to $\mathbb{F}_{q^z}$ for arbitrary $z \in \mathbb{N}$. Specifically, consider $L = z$ computations, and denote $\bar{\mathbf{V}}'_k = \mathbf{V}'_k \otimes \mathbf{I}^{z\times z}$, $\bar{\mathbf{V}}_k = \mathbf{V}_k \otimes \mathbf{I}^{z\times z}$ and $\bar{\mathbf{U}}_k = \mathbf{U}_k \otimes \mathbf{I}^{z\times z}$ as the *z-extension* of the coefficient matrices, where $\otimes$ denotes the Kronecker product. Denote $\bar{\mathbf{X}} = \text{vec}(\mathbf{X}^T)$, where $\text{vec}(\cdot)$ is the vectorization function. By this notation, we can restate the problem such that User k has side-information $\bar{\mathbf{X}}^T\bar{\mathbf{V}}'_k$ and wants to compute $\bar{\mathbf{X}}^T\bar{\mathbf{V}}_k$ for $k = [1 : 3]$, where $\bar{\mathbf{X}} \in \mathbb{F}_q^{dz\times 1}$, $\bar{\mathbf{V}}'_k \in \mathbb{F}_q^{dz\times m'z}$ and $\bar{\mathbf{V}}_k \in \mathbb{F}_q^{dz\times mz}$. Now, since $\mathbb{F}_q$ is a subfield of $\mathbb{F}_{q^z}$, this problem is equivalent to the problem where $\mathbf{X} \in \mathbb{F}_{q^z}^{d\times 1}$, $\mathbf{V}'_k \in \mathbb{F}_{q^z}^{d\times m'}$ and $\mathbf{V}_k \in \mathbb{F}_{q^z}^{d\times m}$ for $L = 1$ computation. By considering the elements in $\mathbb{F}_{q^z}$ instead of $\mathbb{F}_q$, we have more flexibility in designing schemes by choosing symbols in the extension field to jointly code over z computations. Since the achievable scheme allows joint coding over any L computations, considering $L = L'z$ computations in the original problem with field $\mathbb{F}_q$ is equivalent to considering L' computations in the extended field with $\mathbb{F}_{q^z}$. Appendix A illustrates the idea of field size extension with an example.

C. Useful Lemma

Next let us introduce a useful lemma.

Lemma 4. *Let $A \in \mathbb{F}_q^{d\times a}, B_1 \in \mathbb{F}_q^{d\times b_1}$ and $B_2 \in \mathbb{F}_q^{d\times b_2}$ be arbitrary matrices with full column rank (bases), i.e.,* $\text{rk}(A) = a, \text{rk}(B_1) = b_1, \text{rk}(B_2) = b_2$. *Denote* $\text{rk}(B_1 \mid A) = r_{1|A}$, $\text{rk}(B_2 \mid A) = r_{2|A}$ *and* $\text{rk}([B_1, B_2] \mid A) = r_{1,2|A}$. *Then for any non-negative integers n_1, n_2 such that $n_1 \leq r_{1|A}, n_2 \leq r_{2|A}$ and $n_1 + n_2 \leq r_{1,2|A}$, there exist submatrices of B_1, B_2, namely $B'_1 \in \mathbb{F}_q^{d\times n_1}$ and $B'_2 \in \mathbb{F}_q^{d\times n_2}$, respectively, such that $[A, B'_1, B'_2]$ has full column rank $a + n_1 + n_2$.*

Proof. Consider first the case that $n_1 + n_2 = r_{1,2|A}$. By Steinitz Exchange lemma there exist submatrices $B_1^{(r_{1|A})}, B_2^{(r_{2|A})}$, comprised of $r_{1|A}, r_{2|A}$ columns of B_1, B_2, respectively, such that $[A, B_1^{(r_{1|A})}]$, $[A, B_2^{(r_{2|A})}]$ have full column ranks (the superscripts within the parantheses indicate the number of columns). Now, we claim that if $Y^{(a+r_{1|A}+r_{2|A})} = [A, B_1^{(r_{1|A})}, B_2^{(r_{2|A})}]$ does not have full column rank, i.e., $a + r_{1|A} + r_{2|A} > \text{rk}(Y^{(a+r_{1|A}+r_{2|A})}) = a + r_{1,2|A}$, then it is always possible to drop a column of $B_1^{(r_{1|A})}$ to yield $Y^{(a+r_{1|A}+r_{2|A}-1)} = [A, B_1^{(r_{1|A}-1)}, B_2^{(r_{2|A})}]$ which has one less column but the same column rank as $Y^{(a+r_{1|A}+r_{2|A})}$. The claim is proved as follows. Since $Y^{(a+r_{1|A}+r_{2|A})}$ does not have full column rank, there exists a non-zero column

vector Z, such that $Y^{(a+r_{1|A}+r_{2|A})}Z = 0_{d\times 1}$. This non-zero vector Z must have more than one non-zero element (because $Y^{(a+r_{1|A}+r_{2|A})}$ has non-zero columns), and at least one of its non-zero elements must be in a row-index that maps to one of the columns of $B_1^{(r_{1|A})}$ (because $[A, B_2^{(r_{2|A})}]$ has full column rank). This column of $B_1^{(r_{1|A})}$ can be dropped because it is spanned by the remaining columns of $Y^{(a+r_{1|A}+r_{2|A})}$ that are selected by the support of Z, so that $Y^{(a+r_{1|A}+r_{2|A}-1)}$ has the same rank as $Y^{(a+r_{1|A}+r_{2|A})}$. The same claim holds for $B_2^{(r_{2|A})}$ as well. Repeating this argument we can drop columns of $B_1^{(r_{1|A})}, B_2^{(r_{2|A})}$, one-by-one, in any order we wish, until we meet the target values n_1, n_2 at which point the resulting matrix $[A, B_1', B_2']$ has full column rank, equal to $a + r_{1,2|A}$. Finally, if $n_1 + n_2 < r_{1,2|A}$, then we continue the process for an additional $r_{1,2|A} - (n_1 + n_2)$ steps, but each additional column that is dropped now reduces both the rank and the number of columns by 1, until B_1', B_2' are left with only n_1, n_2 columns, respectively, and $\text{rk}(Y^{(a+r_{1|A}+r_{2|A})}) = a + r_{1,2|A} - (r_{1,2|A} - (n_1 + n_2)) = a + n_1 + n_2$. □

Let us also note the following direct corollary of Lemma 4 which will be used multiple times in our construction of the coding scheme.

Corollary 2. *Let $A \in \mathbb{F}_q^{d\times a}$ and $B \in \mathbb{F}_q^{d\times b}$ be arbitrary matrices with full column rank (bases), i.e.,* $\text{rk}(A) = a, \text{rk}(B) = b$. *Denote* $\text{rk}(B \mid A) = r$. *Then for any non-negative integer n such that $n \le r$, there exists a submatrix of B, namely $B' \in \mathbb{F}_q^{d\times n}$ such that $[A, B']$ has full column rank $a + n$.*

Proof. Corollary 2 is implied by Lemma 4, by mapping A, B here to A, B_1 in Lemma 4, respectively, and setting $b_2 = 0$. □

D. Construction of the Optimal Broadcast Scheme

The construction of the optimal broadcast information follows the formulation of Theorem 2 and the depiction in Figure 4. At a high level, the goal is to construct a scheme that broadcasts λ_{123} dimensions that are simultaneously useful to all 3 users (3 birds, 1 stone), λ_{ij} dimensions that are simultaneously useful to Users i, j (2 birds, 1 stone) for $(i, j) \in \{(1, 2), (1, 3), (2, 3)\}$, and λ dimensions that are of the type (3 birds, 2 stones), i.e., where transmission of 2 dimensions collectively satisfies 1 demand dimension for every user. For this construction, let us first consider non-negative integers $\lambda_{123}, \lambda_{12}, \lambda_{13}, \lambda_{23}, \lambda$ that satisfy the constraints (21)-(24) specified in Theorem 2. Generalization of $\lambda_{123}, \lambda_{12}, \lambda_{13}, \lambda_{23}, \lambda$ to rationals is handled in Section VIII-E and VIII-F.

Let us start with the (3 birds, 1 stone) component of the construction, and for now let us focus on User 1. Some adjustments will be necessary eventually to make the scheme work for all 3 users. We wish to broadcast λ_{123} dimensions for this (3 birds, 1 stone) component of our scheme, but it remains to determine the actual information to be transmitted. For this, let us recall Corollary 2, which guarantees that there exists a submatrix of $\mathbf{U}_{123}$, namely $\mathbf{U}_{123}^{(\lambda_{123})} \in \mathbb{F}_{q^z}^{d\times\lambda_{123}}$, such that the following matrix has full column rank,

$$\text{rk}([\mathbf{V}_1', \mathbf{U}_{123}^{(\lambda_{123})}]) = m_1' + \lambda_{123}. \tag{64}$$

Broadcasting $\mathbf{X}^T\mathbf{U}_{123}^{(\lambda_{123})}$ would help User 1 acquire λ_{123} desired dimensions based on his side-information $\mathbf{X}^T\mathbf{V}_1'$.

As a cautionary note, let us point out that this particular $\mathbf{U}_{123}^{(\lambda_{123})}$ which is useful for User 1 may not be useful for User 2 or User 3, i.e., $[\mathbf{V}_k', \mathbf{U}_{123}^{(\lambda_{123})}]$ may not have full column rank for $k = 2, 3$. One can similarly find submatrices of $\mathbf{U}_{123}$ of size (number of columns) λ_{123} that are useful for User 2, or 3 individually, but in general these will be different matrices. In the end the challenge will be to find the *same* matrix that is useful for all three users. For now we ignore this challenge and proceed with only User 1 as our focus.

Next, consider the (2 birds, 1 stone) components, specifically let us find λ_{12} dimensions within $\langle\mathbf{U}_{12}\rangle$, and another λ_{13} dimensions within $\langle\mathbf{U}_{13}\rangle$ that will be useful to User 1, conditioned on the user's side-information $\mathbf{V}_1'$. Letting $A = [\mathbf{V}_1', \mathbf{U}_{123}^{(\lambda_{123})}]$, $B_1 = \mathbf{U}_{12}$ and $B_2 = \mathbf{U}_{13}$ in Lemma 4, we have $a = m_1' + \lambda_{123}$, $r_{1|A} = \text{rk}(\mathbf{U}_{12} \mid \mathbf{V}_1') - \lambda_{123}$, $r_{2|A} = \text{rk}(\mathbf{U}_{13} \mid \mathbf{V}_1') - \lambda_{123}$, and $r_{1,2|A} = \text{rk}([\mathbf{U}_{12}, \mathbf{U}_{13}] \mid \mathbf{V}_1') - \lambda_{123}$. Then according to Lemma 4, there exists a submatrix of $\mathbf{U}_{12}$, namely, $\mathbf{U}_{12}^{(\lambda_{12})} \in \mathbb{F}_{q^z}^{d\times\lambda_{12}}$, and a submatrix of $\mathbf{U}_{13}$, namely, $\mathbf{U}_{13}^{(\lambda_{13})} \in \mathbb{F}_{q^z}^{d\times\lambda_{13}}$, such that the following matrix has full column rank,

$$\text{rk}([\mathbf{V}_1', \mathbf{U}_{123}^{(\lambda_{123})}, \mathbf{U}_{12}^{(\lambda_{12})}, \mathbf{U}_{13}^{(\lambda_{13})}]) = m_1' + \lambda_{123} + \lambda_{12} + \lambda_{13}. \tag{65}$$

Once again, note that these choices may not work for Users 2, 3, so that challenge remains to be overcome later.

Next, consider the (2 stones, 3 birds) component of the scheme. Keeping our focus on User 1, let us find λ dimensions of broadcast information from the subspace $\langle\mathbf{U}_{1(2,3)}\rangle$ that will be useful for User 1. Since we only consider parameters that satisfy the conditions in Theorem 2, which include in particular (24), it follows that $\lambda \le \text{rk}(\mathbf{U}_{1(2,3)} \mid \mathbf{V}_1') - \lambda_{12} - \lambda_{13} - \lambda_{123}$ by definition. Letting $A = [\mathbf{V}_1', \mathbf{U}_{123}^{(\lambda_{123})}, \mathbf{U}_{12}^{(\lambda_{12})}, \mathbf{U}_{13}^{(\lambda_{13})}]$, $B = \mathbf{U}_{1(2,3)}$ in Corollary 2, we have $a = m_1' + \lambda_{123} + \lambda_{12} + \lambda_{13}$ and $r = \text{rk}(\mathbf{U}_{1(2,3)} \mid \mathbf{V}_1') - \lambda_{12} - \lambda_{13} - \lambda_{123} \ge \lambda$. Then Corollary 2 implies that there exists a submatrix of $\mathbf{U}_{1(2,3)}$, namely, $\mathbf{U}_{1(2,3)}^{(\lambda)} \in \mathbb{F}_{q^z}^{d\times\lambda}$ such that the following matrix has full column rank,

$$\begin{aligned}&\text{rk}([\mathbf{V}_1', \mathbf{U}_{123}^{(\lambda_{123})}, \mathbf{U}_{12}^{(\lambda_{12})}, \mathbf{U}_{13}^{(\lambda_{13})}, \mathbf{U}_{1(2,3)}^{(\lambda)}])\\ &= m_1' + \lambda_{123} + \lambda_{12} + \lambda_{13} + \lambda.\end{aligned} \tag{66}$$

Next, by letting A be the above matrix and $B = \mathbf{U}_1$ in Corollary 2, we have $a = m_1' + \lambda_{123} + \lambda_{12} + \lambda_{13} + \lambda$, $r = \text{rk}(\mathbf{U}_1 \mid \mathbf{V}_1') - (\lambda_{123} + \lambda_{12} + \lambda_{13} + \lambda) = m_1 - (\lambda_{123} + \lambda_{12} + \lambda_{13} + \lambda) \triangleq t_1$. Then by Corollary 2, there exists a submatrix of $\mathbf{U}_1$, namely, $\mathbf{U}_1^{(t_1)} \in \mathbb{F}_{q^z}^{d\times t_1}$ such that the following matrix has full column rank.

$$\text{rk}([\mathbf{V}_1', \mathbf{U}_{123}^{(\lambda_{123})}, \mathbf{U}_{12}^{(\lambda_{12})}, \mathbf{U}_{13}^{(\lambda_{13})}, \mathbf{U}_{1(2,3)}^{(\lambda)}, \mathbf{U}_1^{(t_1)}]) = m_1 + m_1', \tag{67}$$

which implies that it is a basis of $\langle \mathbf{U}_1 \rangle$ since each column of the matrix is in $\langle \mathbf{U}_1 \rangle$.

Finally, in Corollary 2 let A be the matrix in (67) and $B = \mathbf{I}^{d\times d}$ be the $d \times d$ identity matrix. We have $a = m_1 + m_1'$ and $r = d - m_1 - m_1'$. Then by Corollary 2, there exists a submatrix of $\mathbf{I}^{d\times d}$, namely, $\mathbf{Z}_1 \in \mathbb{F}_{q^z}^{d\times(d-m_1-m_1')}$ such that the following $d \times d$ matrix has full rank.

$$\text{rk}([\mathbf{V}_1', \mathbf{U}_{123}^{(\lambda_{123})}, \mathbf{U}_{12}^{(\lambda_{12})}, \mathbf{U}_{13}^{(\lambda_{13})}, \mathbf{U}_{1(2,3)}^{(\lambda)}, \mathbf{U}_1^{(t_1)}, \mathbf{Z}_1]) = d. \tag{68}$$

In particular, the determinant of the matrix is non-zero.

The following step allows a mixing of information, leading to a random-coding argument that will be important to reconcile the users' different perspectives. So consider the following determinant, which is a polynomial in the variables corresponding to the elements of the matrices $\mathbf{N}_{123}, \mathbf{N}_{12}, \mathbf{N}_{13}, \mathbf{M}_{12}, \mathbf{M}_{13}, \mathbf{M}$, while the remaining matrices are fixed.

$$P_1 = \det\big([\overbrace{\mathbf{V}_1'}^{m_1'}, \overbrace{\mathbf{U}_{123}\mathbf{N}_{123}}^{\lambda_{123}}, \overbrace{\mathbf{U}_{12}\mathbf{N}_{12}}^{\lambda_{12}}, \overbrace{\mathbf{U}_{13}\mathbf{N}_{13}}^{\lambda_{13}}, \overbrace{\mathbf{U}_{12}\mathbf{M}_{12} + \mathbf{U}_{13}\mathbf{M}_{13} + \mathbf{B}_{1(2,3)}\mathbf{M}}^{\lambda}, \mathbf{U}_1^{(t_1)}, \mathbf{Z}_1]\big) \tag{69}$$

The sizes of the variable matrices are specified below.

$$\begin{aligned} &\mathbf{N}_{123} : \text{rk}(\mathbf{U}_{123}) \times \lambda_{123}, && \mathbf{N}_{12} : \text{rk}(\mathbf{U}_{12}) \times \lambda_{12}, \\ &\mathbf{N}_{13} : \text{rk}(\mathbf{U}_{13}) \times \lambda_{13}, && \mathbf{M}_{12} : \text{rk}(\mathbf{U}_{12}) \times \lambda, \\ &\mathbf{M}_{13} : \text{rk}(\mathbf{U}_{13}) \times \lambda, && \mathbf{M} : \text{rk}(\mathbf{B}_{1(2,3)}) \times \lambda. \end{aligned} \tag{70}$$

We claim that P_1 is not a zero polynomial. This is because we can assign values to the variables such that the matrix in (69) becomes identical to the constant matrix in (68), which has non-zero determinant. Note that by Lemma 2, *(P5)* and *(P8)*, $\langle \mathbf{U}_{1(2,3)} \rangle = \langle [\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{13}, \mathbf{B}_{1(2,3)}] \rangle = \langle [\mathbf{U}_{12}, \mathbf{U}_{13}, \mathbf{B}_{1(2,3)}] \rangle$, therefore, $\mathbf{U}_{1(2,3)}^{(\lambda)} = \mathbf{U}_{12}\mathbf{M}_{12} + \mathbf{U}_{13}\mathbf{M}_{13} + \mathbf{B}_{1(2,3)}\mathbf{M}$ for some realization of $\mathbf{M}_{12}, \mathbf{M}_{13}, \mathbf{M}$. Since there exists a non-zero evaluation of P_1 it cannot be the zero polynomial.

So far our discussion focused on User 1. Proceeding similarly for Users 2 and 3 we arrive at corresponding polynomials P_2, P_3 as shown below,

$$P_2 = \det([\overbrace{\mathbf{V}_2'}^{m_2'}, \overbrace{\mathbf{U}_{123}\mathbf{N}_{123}}^{\lambda_{123}}, \overbrace{\mathbf{U}_{12}\mathbf{N}_{12}}^{\lambda_{12}}, \overbrace{\mathbf{U}_{23}\mathbf{N}_{23}}^{\lambda_{23}}, \overbrace{-\mathbf{U}_{12}\mathbf{M}_{12} + \mathbf{U}_{23}\mathbf{M}_{23} + \mathbf{B}_{2(1,3)}\mathbf{M}}^{\lambda}, \mathbf{U}_2^{(t_2)}, \mathbf{Z}_2]) \tag{71}$$

$$P_3 = \det([\overbrace{\mathbf{V}_3'}^{m_3'}, \overbrace{\mathbf{U}_{123}\mathbf{N}_{123}}^{\lambda_{123}}, \overbrace{\mathbf{U}_{13}\mathbf{N}_{13}}^{\lambda_{13}}, \overbrace{\mathbf{U}_{23}\mathbf{N}_{23}}^{\lambda_{23}}, \overbrace{\mathbf{U}_{13}\mathbf{M}_{13} + \mathbf{U}_{23}\mathbf{M}_{23} + \mathbf{B}_{3(1,2)}\mathbf{M}}^{\lambda}, \mathbf{U}_3^{(t_3)}, \mathbf{Z}_3]) \tag{72}$$

that are similarly shown to be non-zero polynomials, in the variables corresponding to the elements of the matrices $\mathbf{N}_{123}$, $\mathbf{N}_{12}$, $\mathbf{N}_{13}$, $\mathbf{N}_{23}$, $\mathbf{M}_{12}$, $\mathbf{M}_{13}$, $\mathbf{M}_{23}$, $\mathbf{M}$, with the following remaining specifications in addition to those in (70).

$$\mathbf{M}_{23} : \text{rk}(\mathbf{U}_{23}) \times \lambda, \qquad \mathbf{N}_{23} : \text{rk}(\mathbf{U}_{23}) \times \lambda_{23}, \tag{73}$$

and

$$\mathbf{Z}_2 \in \mathbb{F}_{q^z}^{d\times(d-m_2-m_2')}, \tag{74}$$

$$\mathbf{Z}_3 \in \mathbb{F}_{q^z}^{d\times(d-m_3-m_3')}, \tag{75}$$

$$t_2 \triangleq m_2 - (\lambda_{123} + \lambda_{12} + \lambda_{23} + \lambda), \tag{76}$$

$$t_3 \triangleq m_3 - (\lambda_{123} + \lambda_{13} + \lambda_{23} + \lambda). \tag{77}$$

Note that the minus sign before $\mathbf{U}_{12}\mathbf{M}_{12}$ in (71) still allows the entries of $-\mathbf{M}_{12}$ to be any element in $\mathbf{F}_{q^z}$, and thus we can still evaluate the determinants individually to non-zero by choosing appropriate elements in $\mathbf{F}_{q^z}$. Now since P_1, P_2 and P_3 are non-zero polynomials, their product $P \triangleq P_1P_2P_3$ is also a non-zero polynomial in the variables corresponding to the elements of the matrices $\mathbf{N}_{123}$, $\mathbf{N}_{12}$, $\mathbf{N}_{13}$, $\mathbf{N}_{23}$, $\mathbf{M}_{12}$, $\mathbf{M}_{13}$, $\mathbf{M}_{23}$, and $\mathbf{M}$. Furthermore, the polynomial P has a degree D loosely (the loose bound suffices for our purpose) bounded above as,

$$D \leq 3d. \tag{78}$$

By Schwartz-Zippel Lemma, if the elements of $\mathbf{N}_{123}, \mathbf{N}_{12}, \mathbf{N}_{13}, \mathbf{N}_{23}, \mathbf{M}_{12}, \mathbf{M}_{13}, \mathbf{M}_{23}, \mathbf{M}$ are chosen i.i.d uniformly from $\mathbb{F}_{q^z}$, then the probability of P evaluating to 0 is not more than $\frac{D}{q^z} \leq \frac{3d}{q^z}$. Thus, by choosing $z > \log_q(3d)$, we ensure that there exist such $\mathbf{N}_{123}, \mathbf{N}_{12}, \mathbf{N}_{13}, \mathbf{N}_{23}, \mathbf{M}_{12}, \mathbf{M}_{13}, \mathbf{M}_{23}, \mathbf{M}$ that produce a non-zero evaluation of P, which implies that P_1, P_2 and P_3 are evaluated to non-zero simultaneously. Recall that we previously found three constructions, by identifying submatrices of subspace matrices, and each such construction could only be guaranteed to work for one user. The formulation based on $\mathbf{N}_{123}, \mathbf{N}_{12}, \mathbf{N}_{13}, \mathbf{N}_{23}, \mathbf{M}_{12}, \mathbf{M}_{13}, \mathbf{M}_{23}, \mathbf{M}$ represents essentially a generic solution for each user. Whereas the original solutions comprised of specific submatrices may not be compatible, the generic solutions turn out to be compatible, as evident in the argument that P_1, P_2, P_3 are simultaneously non-zero for appropriate choices of the variables. This is essentially a random coding argument, because it shows the existence of a good code among randomly chosen possibilities.

With any such choice, we are able to construct the broadcast symbol as follows.

$$\mathbf{S} = \mathbf{X}^T[\mathbf{U}_{123}\mathbf{N}_{123}, \mathbf{U}_{12}\mathbf{N}_{12}, \mathbf{U}_{13}\mathbf{N}_{13}, \mathbf{U}_{23}\mathbf{N}_{23}, \tag{79}$$

$$\mathbf{U}_{12}\mathbf{M}_{12} + \mathbf{U}_{13}\mathbf{M}_{13} + \mathbf{B}_{1(2,3)}\mathbf{M}, \tag{80}$$

$$-\mathbf{U}_{12}\mathbf{M}_{12} + \mathbf{U}_{23}\mathbf{M}_{23} + \mathbf{B}_{2(1,3)}\mathbf{M}, \tag{81}$$

$$\mathbf{U}_1^{(t_1)}, \mathbf{U}_2^{(t_2)}, \mathbf{U}_3^{(t_3)}]. \tag{82}$$

With $\mathbf{S}$, User 1 is able to obtain (using its side-information)

$$\mathbf{X}^T[\mathbf{V}_1', \mathbf{U}_{123}\mathbf{N}_{123}, \mathbf{U}_{12}\mathbf{N}_{12}, \mathbf{U}_{13}\mathbf{N}_{13}, \mathbf{U}_{12}\mathbf{M}_{12} + \mathbf{U}_{13}\mathbf{M}_{13} + \mathbf{B}_{1(2,3)}\mathbf{M}, \mathbf{U}_1^{(t_1)}], \tag{83}$$

and thus compute $\mathbf{X}^T\mathbf{U}_1$, since the columns of the matrix to the right of $\mathbf{X}^T$ form a basis of $\langle \mathbf{U}_1 \rangle$, guaranteed by the fact that P_1 has a non-zero evaluation. Similarly, User 2 is able to obtain (with its side-information)

$$\mathbf{X}^T[\mathbf{V}_2', \mathbf{U}_{123}\mathbf{N}_{123}, \mathbf{U}_{12}\mathbf{N}_{12}, \mathbf{U}_{23}\mathbf{N}_{23},$$

$$- \mathbf{U}_{12}\mathbf{M}_{12} + \mathbf{U}_{23}\mathbf{M}_{23} + \mathbf{B}_{2(1,3)}\mathbf{M}, \mathbf{U}_2^{(t_2)}], \quad (84)$$

and thus compute $\mathbf{X}^T\mathbf{U}_2$, since the columns of the matrix on the right of $\mathbf{X}^T$ form a basis of $\langle \mathbf{U}_2 \rangle$, guaranteed by the fact that P_2 has a non-zero evaluation.

User 3 first computes

$$\mathbf{X}^T(\mathbf{U}_{12}\mathbf{M}_{12} + \mathbf{U}_{13}\mathbf{M}_{13} + \mathbf{B}_{1(2,3)}\mathbf{M}) \quad (85)$$

$$+ \mathbf{X}^T(-\mathbf{U}_{12}\mathbf{M}_{12} + \mathbf{U}_{23}\mathbf{M}_{23} + \mathbf{B}_{2(1,3)}\mathbf{M}) \quad (86)$$

$$= \mathbf{X}^T(\mathbf{U}_{13}\mathbf{M}_{13} + \mathbf{U}_{23}\mathbf{M}_{23} + \mathbf{B}_{3(1,2)}\mathbf{M}) \quad (87)$$

where we used (P20) from Lemma 2, i.e., $\mathbf{B}_{1(2,3)} + \mathbf{B}_{2(1,3)} = \mathbf{B}_{3(1,2)}$. Using its side-information, User 3 is then able to obtain,

$$\mathbf{X}^T[\mathbf{V}'_3, \mathbf{U}_{123}\mathbf{N}_{123}, \mathbf{U}_{13}\mathbf{N}_{13}, \mathbf{U}_{23}\mathbf{N}_{23}, \mathbf{U}_{13}\mathbf{M}_{13} + \mathbf{U}_{23}\mathbf{M}_{23} + \mathbf{B}_{3(1,2)}\mathbf{M}, \mathbf{U}_3^{(t_3)}]. \quad (88)$$

Thus, it can compute $\mathbf{X}^T\mathbf{U}_3$, since the matrix on the right of $\mathbf{X}^T$ is a basis of $\langle \mathbf{U}_3 \rangle$, guaranteed by the fact that P_3 evaluates to a non-zero value.

The cost of this broadcast $\mathbf{S}$, as noted in Theorem 2, is found as,

$$\Delta = N/L \quad (89)$$

$$= N/z \quad (90)$$

$$= \lambda_{123} + \lambda_{12} + \lambda_{13} + \lambda_{23} + 2\lambda + t_1 + t_2 + t_3 \quad (91)$$

$$= m_1 + m_2 + m_3 - 2\lambda_{123} - \lambda_{12} - \lambda_{13} - \lambda_{23} - \lambda \quad (92)$$

$$= \text{rk}(\mathbf{V}_1|\mathbf{V}'_1) + \text{rk}(\mathbf{V}_2|\mathbf{V}'_2) + \text{rk}(\mathbf{V}_3|\mathbf{V}'_3) - 2\lambda_{123} - \lambda_{12} - \lambda_{13} - \lambda_{23} - \lambda \quad (93)$$

$$\triangleq f(\lambda_{123}, \lambda_{12}, \lambda_{13}, \lambda_{23}, \lambda). \quad (94)$$

This implies that $\Delta^* \leq f(\lambda_{123}, \lambda_{12}, \lambda_{13}, \lambda_{23}, \lambda)$ if $\lambda_{123}, \lambda_{12}, \lambda_{13}, \lambda_{23}$ and λ are non-negative integers subject to the constraints specified in Theorem 2. Next let us show that the arguments extend to rational $\lambda_\bullet$ by a simple matrix extension.

E. Matrix Extension

Technically, the choice of $z > 1$ that enables field extensions in the achievable scheme, already amounts to vector coding, because it requires joint coding of $L = z$ symbols. However, after the field extension, the solution presented above reduces to a scalar coding solution over the extended field $\mathbb{F}_{q^z}$. This formulation only allows integer values of $\lambda_\bullet$ parameters. However, it is quite straightforward to extend the scheme to all rational values of $\lambda_\bullet$ parameters (subject to the constraints specified in Theorem 2) by a typical vector coding extension, labeled here as a Matrix Extension to avoid confusion with field extensions that also require $L > 1$. This is described as follows. Recall that we are allowed to choose any $L \in \mathbb{N}$ in the coding schemes, by letting $L = L'z$ (meaning that the computations are in $\mathbb{F}_{q^z}$ and we jointly code for L' such computations), the ranks of all subspaces scale by L' as the data dimension increases by a factor of L'. Essentially, this amounts to treating successive instances of the data vector as new data dimensions. For example, consider the $m = 1$ dimensional computation of $A + B$ over $d = 2$ dimensional data (A, B), say over $\mathbb{F}_{q^z}$. Considering $L' = 2$ instances, the data becomes $(\mathbf{A}, \mathbf{B}) = ((A(1), A(2)), (B(1), B(2))$, and the desired computation is $\mathbf{A} + \mathbf{B}$, which can also be interpreted as $m_{\text{new}} = 2$ dimensional computations $(A + C, B + D)$ over $d_{\text{new}} = 2d = 4$ dimensional data (A, B, C, D) in $\mathbb{F}_{q^z}$, by mapping $((A(1), A(2)), (B(1), B(2))$ to (A, B, C, D). A bit more formally, by considering L' data instances as one instance of $L'd$ dimensional data (both in $\mathbb{F}_{q^z}$), User $k \in [1:3]$ has side-information $\mathbf{X}^T\mathbf{V}'_k$, which is equivalent to $\text{vec}^T(\mathbf{X})\mathbf{I}^{L'\times L'} \otimes \mathbf{V}'_k$. User k wants to compute $\text{vec}^T(\mathbf{X})\mathbf{I}^{L'\times L'} \otimes \mathbf{V}_k$. The problem is then equivalent to that with data $\mathbf{X} \in \mathbb{F}_{q^z}^{L'd\times 1}$, with coefficient matrices now changed to $\mathbf{I} \otimes \mathbf{V}'_k, \mathbf{I} \otimes \mathbf{V}_k, k = [1:3]$. The signal spaces $\mathbf{U}_k$ are also changed to $\mathbf{I} \otimes \mathbf{U}_k$. Note that this is essentially different from the field size extension presented in Section VIII-B, where the dimensions of the coefficient matrices are not changed after the extension, only the field size is changed. We refer to this as the *matrix extension*, since the dimensions (sizes) of the coefficient matrices scale by a factor of L' (but the field size remains unchanged). The ranks of $\mathbf{U}_k$ and $\mathbf{V}'_k$ also scale by L', as do the ranks of all subspaces considered in (21)-(24). Thus, the RHS of all constraints in (21)-(24) scale by L', implying a similar scaling of the $\lambda_\bullet$ parameters. Thus, all rational values of $\lambda_\bullet$ parameters can be transformed into integer values by considering a matrix extension by a factor L' where L' is the common denominator of the rational values.

F. Completing the Proof of Achievability

At this point we have the bound that $\Delta^* \leq f(\lambda_{123}, \lambda_{12}, \lambda_{13}, \lambda_{23}, \lambda)$ if $\lambda_{123}, \lambda_{12}, \lambda_{13}, \lambda_{23}, \lambda$ are non-negative rational numbers subject to the constraints specified in (21)-(24). The final step of the achievability proof is to recall [64], [65] that for any linear programming problem, say $\max \ \mathbf{c}^T\mathbf{x}$, s.t. $\mathbf{A}\mathbf{x} \leq \mathbf{b}$, $\mathbf{x} \geq \mathbf{0}$, if all the elements of $\mathbf{A}, \mathbf{b}, \mathbf{c}$ are rational, and the optimal exists, then there exists an optimizing $\mathbf{x}$ whose elements are also rational, and so is the optimal value of the objective function. Note that in the linear program in Theorem 2 all coefficients are indeed rational, in fact the coefficients of $\lambda_\bullet$ parameters in the constraints and the objective are all either $0, 1$ or 2, and the constants on the RHS of the constraints (21)-(24) are conditional-ranks, so they are integers as well, by definition. The feasible region is a rational polytope, so all vertices are rational, and one of the vertices must be optimal for a linear program over a rational polytope. Therefore, there exist non-negative rational values $\lambda^*_{123}, \lambda^*_{12}, \lambda^*_{13}, \lambda^*_{23}, \lambda^*$ that satisfy (21)-(24), for which we we have $f(\lambda^*_{123}, \lambda^*_{12}, \lambda^*_{13}, \lambda^*_{23}, \lambda^*) = F^*$. This gives us the desired bound, $\Delta^* \leq F^*$. □

IX. Matching Achievability with Converse: $F^* \leq \max\{\Delta_1, \Delta_2\}$

The converse proof in Section VII established the lower bound $\Delta^* \geq \max\{\Delta_1, \Delta_2\}$, whereas the achievability proof in Section VIII established the upper bound $\Delta^* \leq F^*$. In this section we show that the bounds are tight. To do so, we will prove that $F^* \leq \max\{\Delta_1, \Delta_2\}$.

Recall that, subject to the constraints (21)-(24), the linear program in Theorem 2 finds

$$F^* = \min_{\lambda_{123},\lambda_{12},\lambda_{13},\lambda_{23},\lambda\in\mathbb{R}_+} m_1+m_2+m_3 - 2\lambda_{123}-\lambda_{12}-\lambda_{13}-\lambda_{23}-\lambda \quad (95)$$

$$= \min_{\lambda_{123},\lambda_{12},\lambda_{13},\lambda_{23},\lambda\in\mathbb{R}_+} f(\lambda_{123},\lambda_{12},\lambda_{13},\lambda_{23},\lambda). \quad (96)$$

We will proceed with the proof in two steps. First, in Subsection IX-A, we manipulate $\max\{\Delta_1,\Delta_2\}$ into an equivalent compact form. Then, in Subsection IX-B we show that in all cases there exist feasible $(\lambda_{123},\lambda_{12},\lambda_{13},\lambda_{23},\lambda)$ for which $f(\lambda_{123},\lambda_{12},\lambda_{13},\lambda_{23},\lambda) \le \max\{\Delta_1,\Delta_2\}$ and therefore by (96), we have $F^* \le \max\{\Delta_1,\Delta_2\}$.

A. Equivalent Expression for Δ_1,Δ_2 with Compact Notation

To avoid lengthy notation due to the repetitive use of conditional ranks, let us introduce the following compact forms.

$$\begin{aligned}
r_{123} &\triangleq \min_{k\in[3]} \mathrm{rk}(\mathbf{U}_{123} \mid \mathbf{V}'_k),\\
r_{12} &\triangleq \min_{k\in\{1,2\}} \mathrm{rk}(\mathbf{U}_{12} \mid \mathbf{V}'_k),\\
r_{13} &\triangleq \min_{k\in\{1,3\}} \mathrm{rk}(\mathbf{U}_{13} \mid \mathbf{V}'_k),\\
r_{23} &\triangleq \min_{k\in\{2,3\}} \mathrm{rk}(\mathbf{U}_{23} \mid \mathbf{V}'_k),\\
r_{12,13} &\triangleq \mathrm{rk}([\mathbf{U}_{12},\mathbf{U}_{13}] \mid \mathbf{V}'_1),\\
r_{12,23} &\triangleq \mathrm{rk}([\mathbf{U}_{12},\mathbf{U}_{23}] \mid \mathbf{V}'_2),\\
r_{13,23} &\triangleq \mathrm{rk}([\mathbf{U}_{13},\mathbf{U}_{23}] \mid \mathbf{V}'_3),\\
r_{1(2,3)} &\triangleq \mathrm{rk}(\mathbf{U}_{1(2,3)} \mid \mathbf{V}'_1),\\
r_{2(1,3)} &\triangleq \mathrm{rk}(\mathbf{U}_{2(1,3)} \mid \mathbf{V}'_2),\\
r_{3(1,2)} &\triangleq \mathrm{rk}(\mathbf{U}_{3(1,2)} \mid \mathbf{V}'_3). \quad (97)
\end{aligned}$$

It follows that,

$$\begin{aligned}
r_{12,13} &\ge \max\{r_{12},r_{13}\},\\
r_{12,23} &\ge \max\{r_{12},r_{23}\},\\
r_{13,23} &\ge \max\{r_{13},r_{23}\}. \quad (98)
\end{aligned}$$

Note that by these notations, the constraints (21)-(24) for $\lambda_\bullet$ can be equivalently posed as

$$\begin{aligned}
(21) &\iff \lambda_{123} \le r_{123}, \quad (99)\\
(22) &\iff \lambda_{ij}+\lambda_{123} \le r_{ij},\\
&\qquad \forall (i,j)\in\{(1,2),(1,3),(2,3)\} \quad (100)\\
(23) &\iff \lambda_{ij}+\lambda_{ik}+\lambda_{123} \le r_{ij,ik},\\
&\qquad \forall (i,j,k)\in\{(1,2,3),(2,1,3),(3,1,2)\} \quad (101)\\
(24) &\iff \lambda+\lambda_{ij}+\lambda_{ik}+\lambda_{123} \le r_{i(j,k)},\\
&\qquad \forall (i,j,k)\in\{(1,2,3),(2,1,3),(3,1,2)\} \quad (102)
\end{aligned}$$

With these notations, we are able to express the Δ_1,Δ_2 values defined in Theorem 1 in the following equivalent forms. For Δ_1^{ijk}, we have

$$\begin{aligned}
\Delta_1^{123} &= \mathrm{rk}(\mathbf{V}_1|\mathbf{V}'_1)+\mathrm{rk}(\mathbf{V}_2|\mathbf{V}'_2)+\mathrm{rk}(\mathbf{V}_3|\mathbf{V}'_3)\\
&\quad -\mathrm{rk}(\mathbf{U}_{12}|\mathbf{V}'_2)-\mathrm{rk}(\mathbf{U}_{3(1,2)}|\mathbf{V}'_3) \quad (103)\\
&= \mathrm{rk}(\mathbf{V}_1)+\mathrm{rk}(\mathbf{V}_2)+\mathrm{rk}(\mathbf{V}_3)\\
&\quad -\mathrm{rk}(\mathbf{U}_{12}|\mathbf{V}'_2)-\mathrm{rk}(\mathbf{U}_{3(1,2)}|\mathbf{V}'_3) \quad (104)\\
&= m_1+m_2+m_3-\mathrm{rk}(\mathbf{U}_{12}|\mathbf{V}'_2)-r_{3(1,2)} \quad (105)
\end{aligned}$$

where (104) is due to (63), and similarly

$$\Delta_1^{213} = m_1+m_2+m_3-\mathrm{rk}(\mathbf{U}_{12}|\mathbf{V}'_1)-r_{3(1,2)} \quad (106)$$

which implies that,

$$\begin{aligned}
&\max\{\Delta_1^{123},\Delta_1^{213}\}\\
&= m_1+m_2+m_3-r_{3(1,2)}\\
&\quad -\min\{\mathrm{rk}(\mathbf{U}_{12}|\mathbf{V}'_1),\mathrm{rk}(\mathbf{U}_{12}|\mathbf{V}'_2)\} \quad (107)\\
&= m_1+m_2+m_3-r_{12}-r_{3(1,2)} \quad (108)\\
&\triangleq \delta_3 \quad (109)
\end{aligned}$$

By taking the pairwise maximum of $\{\Delta_1^{132},\Delta_1^{312}\}$ and $\{\Delta_1^{231},\Delta_1^{321}\}$ respectively, we similarly obtain δ_1 and δ_2 as follows.

$$\delta_1 \triangleq \max\{\Delta_1^{132},\Delta_1^{312}\} = m_1+m_2+m_3-r_{23}-r_{1(2,3)}, \quad (110)$$

$$\delta_2 \triangleq \max\{\Delta_1^{231},\Delta_1^{321}\} = m_1+m_2+m_3-r_{13}-r_{2(1,3)}. \quad (111)$$

For Δ_2^{ijk}, first note that $\Delta_2^{123}=\Delta_2^{132}$. Thus, we have,

$$\begin{aligned}
&\max\{\Delta_2^{123},\Delta_2^{132}\} = \Delta_2^{123}\\
&= \mathrm{rk}(\mathbf{V}_1 \mid \mathbf{V}'_1)+\mathrm{rk}(\mathbf{V}_2 \mid \mathbf{V}'_2)+\mathrm{rk}(\mathbf{V}_3 \mid \mathbf{V}'_3)\\
&\quad -\frac{1}{2}\Big(\min_{\ell\in[3]}\big(\mathrm{rk}(\mathbf{U}_{123} \mid \mathbf{V}'_\ell)\big)\\
&\qquad +\mathrm{rk}([\mathbf{U}_{12},\mathbf{U}_{13}] \mid \mathbf{V}'_1)+\mathrm{rk}(\mathbf{U}_{2(1,3)} \mid \mathbf{V}'_2)\\
&\qquad +\mathrm{rk}(\mathbf{U}_{3(1,2)} \mid \mathbf{V}'_3)\Big) \quad (112)\\
&= \mathrm{rk}(\mathbf{V}_1)+\mathrm{rk}(\mathbf{V}_2)+\mathrm{rk}(\mathbf{V}_3)\\
&\quad -\frac{1}{2}\Big(\min_{\ell\in[3]}\big(\mathrm{rk}(\mathbf{U}_{123} \mid \mathbf{V}'_\ell)\big)\\
&\qquad +\mathrm{rk}([\mathbf{U}_{12},\mathbf{U}_{13}] \mid \mathbf{V}'_1)+\mathrm{rk}(\mathbf{U}_{2(1,3)} \mid \mathbf{V}_2)\\
&\qquad +\mathrm{rk}(\mathbf{U}_{3(1,2)} \mid \mathbf{V}'_3)\Big) \quad (113)\\
&= m_1+m_2+m_3-\frac{1}{2}\big(r_{123}+r_{12,13}+r_{2(1,3)}+r_{3(1,2)}\big) \quad (114)\\
&\triangleq \delta_{23} \quad (115)
\end{aligned}$$

where (113) is due to (63). By taking the pairwise maximum of $\{\Delta_2^{213},\Delta_2^{231}\}$ and $\{\Delta_2^{312},\Delta_2^{321}\}$ respectively, we similarly obtain,

$$\begin{aligned}
&\delta_{13} \triangleq \max\{\Delta_2^{213},\Delta_2^{231}\}\\
&= m_1+m_2+m_3-\frac{1}{2}\Big(r_{123}+r_{12,23}+r_{1(2,3)}+r_{3(1,2)}\Big), \quad (116)\\
&\delta_{12} \triangleq \max\{\Delta_2^{312},\Delta_2^{321}\}\\
&= m_1+m_2+m_3-\frac{1}{2}\Big(r_{123}+r_{13,23}+r_{1(2,3)}+r_{2(1,3)}\Big). \quad (117)
\end{aligned}$$

Thus, we have,

$$\max\{\Delta_1,\Delta_2\} = \max\{\delta_1,\delta_2,\delta_3,\delta_{12},\delta_{13},\delta_{23}\}. \quad (118)$$

The next step is to prove that $F^* \leq \max\{\delta_1, \delta_2, \delta_3, \delta_{12}, \delta_{13}, \delta_{23}\}$.

B. Proving $F^* \leq \max\{\delta_1, \delta_2, \delta_3, \delta_{12}, \delta_{13}, \delta_{23}\}$: Constrained Waterfilling

By definition, $F^* \leq f(\lambda_{123}, \lambda_{12}, \lambda_{13}, \lambda_{23}, \lambda)$ for any $(\lambda_{123}, \lambda_{12}, \lambda_{13}, \lambda_{23}, \lambda)$ that satisfies (99)-(102). Therefore, it suffices to show that

$$f(r_{123}, \lambda_{12}, \lambda_{13}, \lambda_{23}, \lambda') \leq \max\{\delta_1, \delta_2, \delta_3, \delta_{12}, \delta_{13}, \delta_{23}\}, \tag{119}$$

where $\lambda' \triangleq \min\{r_{1(2,3)} - \lambda_{12} - \lambda_{13}, r_{2(1,3)} - \lambda_{12} - \lambda_{23}, r_{3(1,2)} - \lambda_{13} - \lambda_{23}\} - r_{123}$. In other words, we fix λ_{123} to r_{123} and λ to λ'. It can be easily verified that $\lambda_{123} = r_{123}, \lambda = \lambda'$ are in the feasible region specified by (99)-(102). As will be shown in the end, fixing $\lambda_{123} = r_{123}, \lambda = \lambda'$ will not hurt the optimality. It is also intuitive because λ_{123} corresponds to the amount of transmission that has the highest efficiency (3 birds, 1 stone) so it should be set as large as possible to r_{123}. Then, $\lambda = \lambda'$ is also the largest possible we can set after λ_{123} is fixed to r_{123}.

Setting λ_{123} and λ to these values (note that both values are non-negative), the objective simplifies to the minimization of,

$$\begin{aligned} f &= m_1 + m_2 + m_3 - 2r_{123} - \lambda_{12} - \lambda_{13} - \lambda_{23} \\ &\quad - \min\{r_{1(2,3)} - \lambda_{12} - \lambda_{13}, \\ &\qquad r_{2(1,3)} - \lambda_{12} - \lambda_{23}, \\ &\qquad r_{3(1,2)} - \lambda_{13} - \lambda_{23}\} + r_{123} \qquad (120) \\ &= \underbrace{(m_1 + m_2 + m_3 - r_{123})}_{\text{constant}} \\ &\quad - \min\{r_{1(2,3)} + \lambda_{23},\ r_{2(1,3)} + \lambda_{13},\ r_{3(1,2)} + \lambda_{12}\}. \qquad (121) \end{aligned}$$

We focus on the remaining three parameters, $\lambda_{12}, \lambda_{13}$ and λ_{23}. Note that minimization of f is equivalent to the maximization of the minimum of the three terms: $r_{1(2,3)} + \lambda_{23}$, $r_{2(1,3)} + \lambda_{13}$, and $r_{3(1,2)} + \lambda_{12}$. Intuitively, this optimization may be seen as a constrained waterfilling problem. To make the connection to waterfilling clear, let us further introduce the following notation.

$$b_1 \triangleq r_{1(2,3)}, \qquad w_1^{max} \triangleq r_{23} - r_{123}, \tag{122}$$
$$b_2 \triangleq r_{2(1,3)}, \qquad w_2^{max} \triangleq r_{13} - r_{123}, \tag{123}$$
$$b_3 \triangleq r_{3(1,2)}, \qquad w_3^{max} \triangleq r_{12} - r_{123}, \tag{124}$$
$$w_1 \triangleq \lambda_{23}, \qquad w_{1,2}^{max} \triangleq r_{13,23} - r_{123}, \tag{125}$$
$$w_2 \triangleq \lambda_{13}, \qquad w_{1,3}^{max} \triangleq r_{12,23} - r_{123}, \tag{126}$$
$$w_3 \triangleq \lambda_{12}, \qquad w_{2,3}^{max} \triangleq r_{12,13} - r_{123}. \tag{127}$$

With this notation, the optimization problem becomes

$$\text{maxmize } h_{min} \triangleq \min\{b_1 + w_1, b_2 + w_2, b_3 + w_3\}, \tag{128}$$

$$s.t. \begin{cases} w_1 \leq w_1^{max}, \\ w_2 \leq w_2^{max}, \\ w_3 \leq w_3^{max}, \\ w_1 + w_2 \leq w_{1,2}^{max}, \\ w_1 + w_3 \leq w_{1,3}^{max}, \\ w_2 + w_3 \leq w_{2,3}^{max}. \\ w_1, w_2, w_3 \in \mathbb{R}_+ \end{cases} \tag{129}$$

Let us explain the waterfilling analogy. There are three

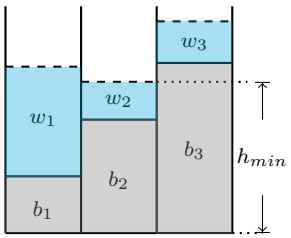

Fig. 5: Constrained Waterfilling.

adjacent vessels as shown in Figure 5, labeled $1, 2, 3$ from left to right. Vessels $1, 2, 3$ have base levels (shown in gray) at heights b_1, b_2, b_3, respectively. We are allowed to add w_1, w_2, w_3 amounts of water to Vessel 1, Vessel 2 and Vessel 3, respectively according to the constraints (129), in order to maximize h_{min}, i.e., the minimum of the heights of water in the three vessels. The objective from (121) now maps to the waterfilling problem as,

$$f = m_1 + m_2 + m_3 - r_{123} - h_{min}. \tag{130}$$

Note that the first three constraints in (129) are constraints on the capacity (for holding water) of individual vessels, and the next three constraints are for pairs of vessels. Furthermore, we have $\max\{w_1^{max}, w_2^{max}\} \leq w_{1,2}^{max}$ by (98), which ensures that the pairwise capacity constraints do not dominate the individual capacity constraints. Since the only constraints are on individual vessel capacities and pairwise vessel capacities, the optimal value of h_{min} must correspond to one of the following outcomes.

1) h_{min} is limited by the individual capacity of Vessel i, $i \in \{1, 2, 3\}$, which holds the maximum water it can, $w_i = w_i^{max}$. In this case, $h_{min} = b_i + w_i^{max}$ and $F^* \leq f = m_1 + m_2 + m_3 - r_{123} - h_{min} = \delta_i$.
2) h_{min} is limited by the pairwise capacity of Vessels i, j, $(i, j) \in \{(1, 2), (1, 3), (2, 3)\}$, which together hold the maximum water they can, i.e., $w_i + w_j = w_{i,j}^{max}$, and have the same final water level h_{min}. In this case, we have $(h_{min} - b_i) + (h_{min} - b_j) = w_{i,j}^{max}$ which gives us $h_{min} = \frac{b_i + b_j + w_{i,j}^{max}}{2}$ and $F^* \leq f = m_1 + m_2 + m_3 - r_{123} - h_{min} = \delta_{ij}$.

Thus, in every case we have $F^* \leq \max\{\delta_1, \delta_2, \delta_3, \delta_{12}, \delta_{13}, \delta_{23}\} = \max\{\Delta_1, \Delta_2\}$, which completes the proof. □

X. Conclusion

The exact capacity of the 3 user LCBC is found for all cases, i.e., for arbitrary finite field $\mathbb{F}_q$, arbitrary data dimension d, and

arbitrary specifications of the users' desired computations and side-information $\mathbf{V}'_k, \mathbf{V}_k$. The 3 user setting introduces several intricacies that were not encountered in the 2 user LCBC, such as the insufficiency of the entropic formulation for tight converse bounds, the need for functional submodularity, the rich variety of subspaces involved, random coding arguments to resolve discrepancies between the users' differing views of the same subspaces, the tradeoffs between the communication efficiencies associated with these subspaces, and the inherent optimization that led us to a constrained waterfilling solution. The fact that the 3 user LCBC capacity turns out to be fully tractable despite these intricacies is surprising. In particular, we note that even though the 3 user LCBC involves at least 6 key subspaces in $\mathbf{V}_k, \mathbf{V}'_k$, $k \in \{1,2,3\}$, the solution did not require the Ingleton inequality, nor were non-Shannon inequalities required for the converse. Instead, the main tools used were Steinitz Exchange lemma, the dimension counting of pairwise unions and intersections of subspaces, functional submodularity, and the random coding argument invoked through the Schwartz-Zippel lemma. The tractability of the 3 user LCBC is indicative of the potential for further progress in understanding the fundamental limits of basic computation networks in future efforts. Indeed, there are many promising directions for future work. Building on the $K = 2$ and $K = 3$ cases, the $K = 4$ user LCBC in particular is an important next step, as it might either reveal a consistent pattern that holds for arbitrary K users or present obstacles that are indicative of the difficulty of the large K setting. Also of interest are asymptotic LCBC settings with large number of users. An intriguing generalization of the LCBC problem is the LCBC with partially informed server, LCBC-PIS in short, where the central server has only limited knowledge of the data in the form of some linear functions of the data. The LCBC-PIS setting has been introduced and solved recently for $K = 2$ users in [66]. The capacity remains open for $K \geq 3$ users. Studies of linear computation multiple access settings (LCMAC) represent another promising research avenue, partially explored in [38] from a coding perspective. Approximate linear computations over real or complex numbers, as well as non-linear computations that connect to coded distributed computing represent other challenging and important research directions for future work. From a practical perspective, studies of computational and communication tradeoffs of AR/VR applications that take advantage of the coding schemes discovered through the studies of LCBC/LCMAC settings would be valuable complements to the theoretical efforts.

APPENDIX A
FIELD EXTENSION

To clarify the notation and illustrate the utility of field extensions, let us present an example. Consider $q = 2, K = 3, d = 2, m = m' = 1$ and the following coefficient matrices $\mathbf{U}_k = [\mathbf{V}'_k, \mathbf{V}_k], k \in [3]$ as

$$\mathbf{U}_1^{2\times 2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ \mathbf{U}_2^{2\times 2} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \ \mathbf{U}_3^{2\times 2} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}. \quad (131)$$

By the problem formulation, $\mathbf{x} \in \mathbb{F}_2^{2\times 1}$ denotes the data for each computation, and $\mathbf{X} \in \mathbb{F}_2^{2\times L}$ denotes the data for L computations. Let us first try to design a coding scheme with $L = 1$. Denote $\mathbf{X}^T = [x_1(1), x_2(1)]$ and then $\mathbf{W}'_1 = \mathbf{X}^T\mathbf{V}'_1 = x_1(1)$, $\mathbf{W}_1 = \mathbf{X}^T\mathbf{V}_1 = x_2(1)$, $\mathbf{W}'_2 = \mathbf{X}^T\mathbf{V}'_2 = x_2(1)$, $\mathbf{W}_2 = \mathbf{X}^T\mathbf{V}_2 = x_1(1) + x_2(1)$, $\mathbf{W}'_3 = \mathbf{X}^T\mathbf{V}'_3 = x_1(1) + x_2(1)$, $\mathbf{W}_3 = \mathbf{X}^T\mathbf{V}_3 = x_1(1)$. The following table shows all possible outcomes of $(\mathbf{W}'_k, \mathbf{W}_k)_{k\in[3]}$.

$x_1(1)$	$x_2(1)$	$\mathbf{W}'_1$	$\mathbf{W}_1$	$\mathbf{W}'_2$	$\mathbf{W}_2$	$\mathbf{W}'_3$	$\mathbf{W}_3$
0	0	0	0	0	0	0	0
0	1	0	1	1	1	1	0
1	0	1	0	0	1	1	1
1	1	1	1	1	0	0	1

A coding scheme must satisfy the property that for any two outcomes, the broadcast information $\mathbf{S}$ corresponding to these outcomes has to be different if $\exists k \in [3]$ such that $\mathbf{W}'_k$ is the same but $\mathbf{W}_k$ is different for these two outcomes. This is necessary to ensure that User k will not be confused when decoding under these two outcomes. Following this rule it is easy to verify from the table that the realization of $\mathbf{S}$ has to be different for any two outcomes in this example, which implies $\mathbf{S}$ has to be different in all outcomes. Thus, $|\mathcal{S}| \geq 4$, which implies $N \geq 2$ and thus $\Delta = N/L \geq 2$ for $L = 1$. In other words, scalar coding schemes cannot achieve $\Delta < 2$.

Let us now consider field extension. Let $z = 2$ and consider $L = z = 2$. Denote $\bar{\mathbf{V}}'_k = \mathbf{V}'_k \otimes \mathbf{I}^{2\times 2}$, $\bar{\mathbf{V}}_k = \mathbf{V}_k \otimes \mathbf{I}^{2\times 2}$ and $\bar{\mathbf{U}}_k = \mathbf{U}_k \otimes \mathbf{I}^{2\times 2}$ as the 2-*extension* of the coefficient matrices, where $\otimes$ denotes the Kronecker product. We have

$$\bar{\mathbf{U}}_1^{4\times 4} = [\bar{\mathbf{V}}'_1, \bar{\mathbf{V}}_1] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\bar{\mathbf{U}}_2^{4\times 4} = [\bar{\mathbf{V}}'_2, \bar{\mathbf{V}}_2] = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix},$$

$$\bar{\mathbf{U}}_3^{4\times 4} = [\bar{\mathbf{V}}'_3, \bar{\mathbf{V}}_3] = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \quad (132)$$

Then denote $\mathbf{X} = [x_1(1), x_1(2); x_2(1), x_2(2)]$ as the data matrix for $L = 2$, and denote $\bar{\mathbf{X}} = \text{vec}(\mathbf{X}^T)$, where $\text{vec}(\cdot)$ is the vectorization function. We have $\bar{\mathbf{X}}^T = [x_1(1), x_1(2), x_2(1), x_2(2)] \in \mathbb{F}_2^{1\times 4}$. We can see that User k has side-information $\bar{\mathbf{X}}^T\bar{\mathbf{V}}'_k$, and wants to compute $\bar{\mathbf{X}}^T\bar{\mathbf{V}}_k$. Then by the property of finite field extensions, we can regard $[x_1(1), x_1(2)]$ as $\bar{x}_1 \in \mathbb{F}_4$ and similarly $[x_2(1), x_2(2)]$ as $\bar{x}_2 \in \mathbb{F}_4$. Accordingly, the extended coefficient matrices are regarded as 2×2 matrices in $\mathbb{F}_4$ as

$$\mathbf{U}_1^{2\times 2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ \mathbf{U}_2^{2\times 2} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \ \mathbf{U}_3^{2\times 2} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}. \quad (133)$$

Note that the matrices are exactly the same as the matrices in (131) but considered in the extended field $\mathbb{F}_4$. To avoid complex notations, we redefine the data matrix as $\mathbf{X} = [\bar{x}_1; \bar{x}_2] \in \mathbb{F}_4^{2\times 1}$. Thus, by considering $L = 2$ computations, we have an equivalent problem where $q = 4, d = 2, m = m' = 1$ and the same coefficient matrices $\mathbf{U}_k = [\mathbf{V}'_k, \mathbf{V}_k], k \in [3]$, but now all elements are from $\mathbb{F}_4$. As a coding scheme with $L = 2$, it suffices to send $\mathbf{S} = \mathbf{X}^T[1; \alpha]$, where $\alpha \notin \{0, 1\}$ and $\alpha \in \mathbb{F}_4$. Since the column vector $[1; \alpha]$ is linearly independent of each of $\mathbf{V}'_1, \mathbf{V}'_2, \mathbf{V}'_3$, each user has two independent equations in

$\mathbf{X}^T$ from which it can decode all of $\mathbf{X}$, and recover the desired $\mathbf{W}_k$. Since $\mathbf{S}$ is chosen as 1 symbol from $\mathbb{F}_4$, we have $N = 2$ (1 symbol in $\mathbb{F}_4$ corresponds to 2 symbols in $\mathbb{F}_2$) and thus $\Delta = N/L = 1$, thus a better $\Delta = N/L$ is achieved by considering $L > 1$.

In general, by considering $L = z$ computations, the original problem is equivalent to the problem with all the same parameters including the coefficient matrices but in the extended field $\mathbb{F}_{q^z}$. By considering z computations in the original problem as 1 computation in the extended field, the original problem over $\mathbb{F}_q$ for $L = zL'$ computations, is equivalent to the new problem with the same parameters in the extended field $\mathbb{F}_{q^z}$ for L' computations.

Appendix B
Some Discussion on Lemma 2

As we go to 3 spaces, $\langle \mathbf{U}_1 \rangle, \langle \mathbf{U}_2 \rangle, \langle \mathbf{U}_3 \rangle$, generalizing the decomposition for $\langle \mathbf{U}_1 \rangle$ and $\langle \mathbf{U}_2 \rangle$ as in Lemma 1 is not so straightforward. Analogies to set-theoretic ideas such as inclusion-exclusion principle and Venn's diagrams do not quite work for 3 vector spaces. For example, if $\langle \mathbf{U}_1 \rangle, \langle \mathbf{U}_2 \rangle, \langle \mathbf{U}_3 \rangle$ are three independent lines in a plane, i.e., pairwise independent one-dimensional subspaces of a 2 dimensional vector space, then $\langle \mathbf{U}_1 \rangle$ has no non-trivial intersection with either of $\langle \mathbf{U}_2 \rangle$ or $\langle \mathbf{U}_3 \rangle$ individually, yet $\langle \mathbf{U}_1 \rangle$ is contained in $\langle [\mathbf{U}_2, \mathbf{U}_3] \rangle$, a situation for which there is no direct set-theoretic analogy. This is why we need the subspace decomposition for $\langle \mathbf{U}_1 \rangle, \langle \mathbf{U}_2 \rangle, \langle \mathbf{U}_3 \rangle$, as illustrated in Figure 3 and formalized in Lemma 2. As noted, the decomposition parallels a corresponding decomposition in the DoF studies of the 3 user MIMO BC by Wang in [55], highlighting its fundamental conceptual significance.

Following the idea of *growing* the basis to cover larger and larger subspaces, similar to the constructive proof for Lemma 1, let us interpret Figure 3, so that Lemma 2 will be intuitively transparent. Consider the space $\langle \mathbf{U}_1 \rangle$, i.e., the column space of $\mathbf{U}_1$. This space is decomposed into 5 subspaces as follows. First we have the space within $\langle \mathbf{U}_1 \rangle$ which overlaps with *both* $\langle \mathbf{U}_2 \rangle$ and $\langle \mathbf{U}_3 \rangle$. This is the space $\langle \mathbf{U}_{123} \rangle \triangleq \langle \mathbf{U}_1 \rangle \cap \langle \mathbf{U}_2 \rangle \cap \langle \mathbf{U}_3 \rangle$. The basis for this space is labeled in the figure as the matrix $\mathbf{B}_{123}$. Now consider the space within $\langle \mathbf{U}_1 \rangle$ which overlaps with $\langle \mathbf{U}_2 \rangle$. This is the space $\langle \mathbf{U}_{12} \rangle \triangleq \langle \mathbf{U}_1 \rangle \cap \langle \mathbf{U}_2 \rangle$. The basis for this space is $[\mathbf{B}_{123}, \mathbf{B}_{12}]$. Note that $\langle \mathbf{U}_{123} \rangle \subset \langle \mathbf{U}_{12} \rangle$, which is also reflected in the fact that the basis for $\langle \mathbf{U}_{12} \rangle$ explicitly contains the basis for $\langle \mathbf{U}_{123} \rangle$. It is important to recall that *the columns of a basis matrix must be linearly independent by definition*. Therefore, not only do we have a basis $[\mathbf{B}_{123}, \mathbf{B}_{12}]$ for $\langle \mathbf{U}_{12} \rangle$, but also by the linear independence of the basis vectors, it follows that $\langle \mathbf{U}_{12} \rangle$ is decomposed into two *independent* subspaces, namely the subspaces $\langle \mathbf{B}_{123} \rangle$ and $\langle \mathbf{B}_{12} \rangle$. This can also be expressed as[4] a direct sum, i.e., $\langle \mathbf{U}_{12} \rangle = \langle \mathbf{B}_{123} \rangle \oplus \langle \mathbf{B}_{12} \rangle$. Similarly, $\langle \mathbf{U}_{13} \rangle$, i.e., the intersection of $\langle \mathbf{U}_1 \rangle$ and $\langle \mathbf{U}_3 \rangle$ is decomposed into *independent* subspaces $\langle \mathbf{B}_{123} \rangle$ and $\langle \mathbf{B}_{23} \rangle$, i.e., $\langle \mathbf{U}_{13} \rangle = \langle \mathbf{B}_{123} \rangle \oplus \langle \mathbf{B}_{13} \rangle$. Continuing the process further, now consider the space within $\langle \mathbf{U}_1 \rangle$ which overlaps with $\langle [\mathbf{U}_2, \mathbf{U}_3] \rangle$, i.e., the space denoted as $\langle \mathbf{U}_{1(2,3)} \rangle$. As indicated in the figure, the basis for this space is $[\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{13}, \mathbf{B}_{1(2,3)}]$, which immediately decomposes $\langle \mathbf{U}_{1(2,3)} \rangle$ into 4 *independent* subspaces, i.e., $\langle \mathbf{U}_{1(2,3)} \rangle = \langle \mathbf{B}_{123} \rangle \oplus \langle \mathbf{B}_{12} \rangle \oplus \langle \mathbf{B}_{23} \rangle \oplus \langle \mathbf{B}_{1(2,3)} \rangle$. Finally, consider all of $\langle \mathbf{U}_1 \rangle$, for which Figure 3 identifies the basis as the matrix $[\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{13}, \mathbf{B}_{1(2,3)}, \mathbf{B}_{1c}]$, thus completing the decomposition of $\langle \mathbf{U}_1 \rangle$ into 5 disjoint subspaces, $\langle \mathbf{U}_1 \rangle = \langle \mathbf{B}_{123} \rangle \oplus \langle \mathbf{B}_{12} \rangle \oplus \langle \mathbf{B}_{23} \rangle \oplus \langle \mathbf{B}_{1(2,3)} \rangle \oplus \langle \mathbf{B}_{1c} \rangle$. Similar decompositions apply to $\langle \mathbf{U}_2 \rangle$ and $\langle \mathbf{U}_3 \rangle$ as well.

The description thus far is similar to set-theoretic decompositions into disjoint sets, as one might represent through disjoint regions in a Venn's diagram. This brings us to the most interesting aspect of the 3-subspace decomposition, highlighted as the yellow regions with dashed boundaries in Figure 3. The subspaces corresponding to these three regions, namely $\langle \mathbf{B}_{1(2,3)} \rangle$, $\langle \mathbf{B}_{2(1,3)} \rangle$, and $\langle \mathbf{B}_{3(1,2)} \rangle$ are only *pairwise* independent, and the span of the union of any two of them contains the third. In fact, it is always possible to choose the basis matrices such that $\mathbf{B}_{1(2,3)} + \mathbf{B}_{2(1,3)} = \mathbf{B}_{3(1,2)}$, which will simplify the construction of the coding scheme. Thus, Figure 3 shows 10 subspaces, including the 3 subspaces highlighted in yellow, and if we exclude any one of the 3 yellow subspaces, the remaining 9 are independent spaces. Mathematically,

$$\begin{aligned}
&\langle [\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3] \rangle \\
&= \langle \mathbf{B}_{1(2,3)} \rangle \oplus \langle \mathbf{B}_{2(1,3)} \rangle \oplus \langle \mathbf{B}_{1c} \rangle \oplus \langle \mathbf{B}_{2c} \rangle \oplus \langle \mathbf{B}_{3c} \rangle \\
&\quad \oplus \langle \mathbf{B}_{12} \rangle \oplus \langle \mathbf{B}_{23} \rangle \oplus \langle \mathbf{B}_{13} \rangle \oplus \langle \mathbf{B}_{123} \rangle \qquad (134) \\
&= \langle \mathbf{B}_{2(1,3)} \rangle \oplus \langle \mathbf{B}_{3(1,2)} \rangle \oplus \langle \mathbf{B}_{1c} \rangle \oplus \langle \mathbf{B}_{2c} \rangle \oplus \langle \mathbf{B}_{3c} \rangle \\
&\quad \oplus \langle \mathbf{B}_{12} \rangle \oplus \langle \mathbf{B}_{23} \rangle \oplus \langle \mathbf{B}_{13} \rangle \oplus \langle \mathbf{B}_{123} \rangle \qquad (135) \\
&= \langle \mathbf{B}_{3(1,2)} \rangle \oplus \langle \mathbf{B}_{1(2,3)} \rangle \oplus \langle \mathbf{B}_{1c} \rangle \oplus \langle \mathbf{B}_{2c} \rangle \oplus \langle \mathbf{B}_{3c} \rangle \\
&\quad \oplus \langle \mathbf{B}_{12} \rangle \oplus \langle \mathbf{B}_{23} \rangle \oplus \langle \mathbf{B}_{13} \rangle \oplus \langle \mathbf{B}_{123} \rangle \qquad (136)
\end{aligned}$$

Appendix C
Proof of Lemma 2: Decomposition of $\langle \mathbf{U}_1 \rangle, \langle \mathbf{U}_2 \rangle, \langle \mathbf{U}_3 \rangle$

Let us begin by informally summarizing the key facts that are used extensively in this section.

1) A matrix M forms a *basis* of the column-space of a matrix U, if and only if $\langle U \rangle \subset \langle M \rangle$ and the number of columns of M is equal to $\text{rk}(U)$. Note that a basis matrix must have full column-rank, i.e., all its columns are linearly independent, and it has only as many columns as needed to span $\langle U \rangle$, i.e., $\text{rk}(U)$ columns.
2) If $A \in \mathbb{F}_q^{d \times a}$ and $B \in \mathbb{F}_q^{d \times b}$ are basis matrices (i.e., they each have full column-rank) and $\langle B \rangle \subset \langle A \rangle$, then there exists a matrix $C \in \mathbb{F}_q^{d \times (a-b)}$ such that $[B, C]$ is a basis of $\langle A \rangle$. It follows that $\langle C \rangle \subset \langle A \rangle$. Let us call C the complement of B in A and denote it as $C = A \backslash B$. Note that such C is not unique, and one feasible choice of such C follows from the Steinitz Exchange Lemma, which produces a C that is a submatrix of A. Other feasible choices of C can be constructed as follows. Denote C^{old} as the choice from the Steinitz Exchange

[4] For vector spaces $\mathcal{V}, \mathcal{V}_1, \cdots, \mathcal{V}_K$, we have $\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \cdots \oplus \mathcal{V}_K$ iff for every $v \in \mathcal{V}$, there exist *unique* $v_k \in \mathcal{V}_k$ for all $k \in [K]$ such that $v = v_1 + v_2 + \cdots + v_K$.

Lemma. Other feasible choices can be constructed as $C^{\text{new}} = C^{\text{old}}R + BR'$, where $R \in \mathbb{F}_q^{(a-b)\times(a-b)}$, $R' \in \mathbb{F}_q^{b\times(a-b)}$ and R is invertible. To see this, first note that C^{new} has the same size as C^{old}. Then note that any $v \in \langle A\rangle$ can be represented as $v = Br_b + C^{\text{old}}r_c$ because $[B, C^{\text{old}}]$ is a basis of $\langle A\rangle$. It follows that $v = B(r_b - R'R^{-1}r_c) + C^{\text{new}}R^{-1}r_c$, which implies that $v \in \langle [B, C^{\text{new}}]\rangle$.

3) Also recall the dimension formula (15), i.e., $\text{rk}(M_1) + \text{rk}(M_2) = \text{rk}(M_1 \cap M_2) + \text{rk}([M_1, M_2])$.

We will now construct the 10 bases that are mentioned in Lemma 2, collectively referred to as $\mathcal{B}$.

$$\mathcal{B} = \{\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{13}, \mathbf{B}_{23}, \mathbf{B}_{1(2,3)}, \mathbf{B}_{2(1,3)}, \mathbf{B}_{3(1,2)}, \mathbf{B}_{1c}, \mathbf{B}_{2c}, \mathbf{B}_{3c}\}$$

First, let us define the compact notation, $b_* \triangleq \text{rk}(\mathbf{B}_*)$ where

$$* \in \{1, 2, 3, 12, 13, 23, 1(2,3), 2(1,3), 3(1,2), 1c, 2c, 3c\}.$$

For example, $b_{1(2,3)} \triangleq \text{rk}(\mathbf{B}_{1(2,3)})$. Since the vector space is d-dimensional, and $\mathbf{B}_*$ are basis matrices, it follows that the size of $\mathbf{B}_*$ is $d \times b_*$. The construction now proceeds as follows.

Step 1: $\mathbf{B}_{123} = \mathbf{U}_{123}$.
Step 2: $\mathbf{B}_{12} = \mathbf{U}_{12} \backslash \mathbf{B}_{123}$.
Step 3: $\mathbf{B}_{13} = \mathbf{U}_{13} \backslash \mathbf{B}_{123}$.
Step 4: $\mathbf{B}_{23} = \mathbf{U}_{23} \backslash \mathbf{B}_{123}$.

These four steps are direct applications of the Steinitz Exchange Lemma, which also guarantees that properties *(P1)* - *(P4)* are satisfied. Next let us prove that *(P5)* - *(P7)* are also satisfied. Consider $(P5)$. It follows from the construction that $[\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{13}]$ spans $\langle [\mathbf{U}_{12}, \mathbf{U}_{13}]\rangle$ because it explicitly contains the bases for both spaces, but we wish to show that it is itself a *basis*, i.e., it has full column-rank. Now, since $[\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{13}]$ has $b_{123} + b_{12} + b_{13}$ columns and $\text{rk}([\mathbf{U}_{12}, \mathbf{U}_{13}]) = \text{rk}(\mathbf{U}_{12}) + \text{rk}(\mathbf{U}_{13}) - \text{rk}(\mathbf{U}_{123}) = (b_{123} + b_{12}) + (b_{123} + b_{13}) - b_{123} = b_{123} + b_{12} + b_{13}$, it follows that $[\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{13}]$ has full column rank. Thus, *(P5)* is satisfied. *(P6)* and *(P7)* are similarly proved by symmetry. We continue the construction of $\mathcal{B}$.

Step 5: $\mathbf{B}_{1(2,3)} = \mathbf{U}_{1(2,3)} \backslash [\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{13}]$.
Step 6: $\mathbf{B}_{2(1,3)} = \mathbf{U}_{2(1,3)} \backslash [\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{23}]$.
Step 7: $\mathbf{B}_{3(1,2)} = \mathbf{U}_{3(1,2)} \backslash [\mathbf{B}_{123}, \mathbf{B}_{13}, \mathbf{B}_{23}]$.
Step 8: $\mathbf{B}_{1c} = \mathbf{U}_1 \backslash [\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{13}, \mathbf{B}_{1(2,3)}]$.
Step 9: $\mathbf{B}_{2c} = \mathbf{U}_2 \backslash [\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{23}, \mathbf{B}_{2(1,3)}]$.
Step 10: $\mathbf{B}_{3c} = \mathbf{U}_3 \backslash [\mathbf{B}_{123}, \mathbf{B}_{13}, \mathbf{B}_{23}, \mathbf{B}_{3(1,2)}]$.

Again, Steps 5-10 are applications of the Steinitz Exchange Lemma, which implies that $[\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{13}, \mathbf{B}_{1(2,3)}, \mathbf{B}_{1c}]$ is a basis of $\langle \mathbf{U}_1\rangle$, $[\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{23}, \mathbf{B}_{2(1,3)}, \mathbf{B}_{2c}]$ is a basis of $\langle \mathbf{U}_2\rangle$, and $[\mathbf{B}_{123}, \mathbf{B}_{13}, \mathbf{B}_{23}, \mathbf{B}_{3(1,2)}, \mathbf{B}_{3c}]$ is a basis of $\langle \mathbf{U}_3\rangle$. Furthermore, $[\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{13}, \mathbf{B}_{23}, \mathbf{B}_{2(1,3)}, \mathbf{B}_{3(1,2)}, \mathbf{B}_{2c}, \mathbf{B}_{3c}]$ is a basis of $\langle [\mathbf{U}_2, \mathbf{U}_3]\rangle$ because it spans $\langle [\mathbf{U}_2, \mathbf{U}_3]\rangle$ by construction, and has full column-rank because

$$\begin{aligned}
&\text{rk}([\mathbf{U}_2, \mathbf{U}_3]) && (137)\\
&= \text{rk}(\mathbf{U}_2) + \text{rk}(\mathbf{U}_3) - \text{rk}(\mathbf{U}_{23}) && (138)\\
&= (b_{123} + b_{12} + b_{23} + b_{2(1,3)} + b_{2c})\\
&\quad + (b_{123} + b_{13} + b_{23} + b_{3(1,2)} + b_{3c})\\
&\quad - (b_{123} + b_{23}) && (139)\\
&= b_{123} + b_{12} + b_{13} + b_{23} + b_{2(1,3)} + b_{3(1,2)} + b_{2c} + b_{3c} && (140)
\end{aligned}$$

which happens to be the number of columns of $[\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{13}, \mathbf{B}_{23}, \mathbf{B}_{2(1,3)}, \mathbf{B}_{3(1,2)}, \mathbf{B}_{2c}, \mathbf{B}_{3c}]$. Thus, *(P14)* - *(P16)* are satisfied.

Next let us show that *(P17)* - *(P19)* are satisfied. Consider *(P17)*, i.e., we wish to show that $\mathbf{B}_{17} \triangleq [\mathbf{B}_{123}, \mathbf{B}_{12}, \mathbf{B}_{13}, \mathbf{B}_{23}, \mathbf{B}_{1(2,3)}, \mathbf{B}_{2(1,3)}, \mathbf{B}_{1c}, \mathbf{B}_{2c}, \mathbf{B}_{3c}]$ is a basis for $\langle [\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3]\rangle$. First let us show that $\langle [\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3]\rangle$ is contained in the span of $\mathbf{B}_{17}$. From *(P11), (P12)* note that the basis for $\langle \mathbf{U}_1\rangle$ is explicitly contained in $\mathbf{B}_{17}$, and so is the basis for $\langle \mathbf{U}_2\rangle$. Then, noting that $\langle \mathbf{B}_{3(1,2)}\rangle \subset \langle [\mathbf{U}_1, \mathbf{U}_2]\rangle$ by its construction in Step 7, it follows from *(P13)* that $\langle \mathbf{U}_3\rangle$ is also contained in the span of $\mathbf{B}_{17}$. Thus, $\langle [\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3]\rangle$ is contained in the column-span of $\mathbf{B}_{17}$. Next let us show that $\mathbf{B}_{17}$ has only as many columns as $\text{rk}([\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3])$, so it must be a basis.

$$\begin{aligned}
&\text{rk}([\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3]) && (141)\\
&= \text{rk}([\mathbf{U}_1, \mathbf{U}_2]) + \text{rk}(\mathbf{U}_3) - \text{rk}(\mathbf{U}_{3(1,2)}) && (142)\\
&= (b_{123} + b_{12} + b_{13} + b_{23} + b_{1(2,3)} + b_{2(1,3)} + b_{1c} + b_{2c})\\
&\quad + (b_{123} + b_{13} + b_{23} + b_{3(1,2)} + b_{3c})\\
&\quad - (b_{123} + b_{13} + b_{23} + b_{3(1,2)}) && (143)\\
&= b_{123} + b_{12} + b_{13} + b_{23} + b_{1(2,3)} + b_{2(1,3)}\\
&\quad + b_{1c} + b_{2c} + b_{3c} && (144)
\end{aligned}$$

which is the number of columns of $\mathbf{B}_{17}$. Thus, *(P17)* is satisfied, and by symmetry *(P18)* and *(P19)* are satisfied as well.

At this point, only *(P20)* remains to be shown. It is worthwhile to note that we always have,

$$b_{1(2,3)} = b_{2(1,3)} = b_{3(1,2)}. \tag{145}$$

This is because properties *(P17)-(P19)* together imply that $b_{1(2,3)} + b_{2(1,3)} = b_{1(2,3)} + b_{3(1,2)} = b_{2(1,3)} + b_{3(1,2)}$ and thus the three components must be equal. If $b_{1(2,3)} = b_{2(1,3)} = b_{3(1,2)} = 0$, then *(P20)* can be neglected. Otherwise, we continue the process from Step 11.

Step 11: Since $\langle \mathbf{B}_{3(1,2)}\rangle \subset \langle [\mathbf{U}_1, \mathbf{U}_2]\rangle$ as noted above, let us uniquely represent $\mathbf{B}_{3(1,2)}$ in the basis of $\langle [\mathbf{U}_1, \mathbf{U}_2]\rangle$ according to *(P14)* as,

$$\begin{aligned}
\mathbf{B}_{3(1,2)} &= \mathbf{B}_{123}\mathbf{R}_1 + \mathbf{B}_{12}\mathbf{R}_2 + \mathbf{B}_{13}\mathbf{R}_3 + \mathbf{B}_{23}\mathbf{R}_4\\
&\quad + \mathbf{B}_{1(2,3)}\mathbf{R}_5 + \mathbf{B}_{1c}\mathbf{R}_6 + \mathbf{B}_{2(1,3)}\mathbf{R}_7 + \mathbf{B}_{2c}\mathbf{R}_8 \qquad (146)
\end{aligned}$$

where $\mathbf{R}_1$ to $\mathbf{R}_8$ are $\mathbb{F}_q$ matrices with appropriate sizes. In particular, from (145) it follows that $\mathbf{R}_5$ and $\mathbf{R}_7$ are square matrices. A key goal in the remainder of the proof will be to show that $\mathbf{R}_5$ and $\mathbf{R}_7$ are invertible.

First we claim that $\mathbf{R}_6$ and $\mathbf{R}_8$ must be zero matrices. We prove this by contradiction. Suppose $\mathbf{R}_6$ is not the zero matrix, say its first column is a non-zero vector $\mathbf{r}$, then $\mathbf{B}_{1c}\mathbf{r} \neq \mathbf{0}$ will lie in $\langle [\mathbf{U}_2, \mathbf{U}_3]\rangle$. However, by construction, $\text{rk}(\mathbf{B}_{1c} \cap [\mathbf{U}_2, \mathbf{U}_3]) = \text{rk}(\mathbf{B}_{1c} \cap \mathbf{U}_1 \cap [\mathbf{U}_2, \mathbf{U}_3]) = \text{rk}(\mathbf{B}_{1c} \cap \mathbf{U}_{1(2,3)}) = 0$, meaning that $\langle \mathbf{B}_{1c}\rangle$ and $\langle [\mathbf{U}_2, \mathbf{U}_3]\rangle$ are

independent spaces. This completes the proof by contradiction, confirming that $\mathbf{R}_6$ is a zero matrix. Similar argument is true for $\mathbf{R}_8$ due to symmetry. Thus, we have

$$\mathbf{B}_{3(1,2)} = \mathbf{B}_{123}\mathbf{R}_1 + \mathbf{B}_{12}\mathbf{R}_2 + \mathbf{B}_{13}\mathbf{R}_3 + \mathbf{B}_{23}\mathbf{R}_4 + \mathbf{B}_{1(2,3)}\mathbf{R}_5 + \mathbf{B}_{2(1,3)}\mathbf{R}_7. \quad (147)$$

We now claim that $\mathbf{R}_5$ and $\mathbf{R}_7$ have full column rank, i.e., they are invertible square matrices. The proof is by contradiction as well. Suppose $\mathbf{R}_5$ does not have full column rank, then there exists a non-zero vector '$\mathbf{a}$' such that $\mathbf{R}_5\mathbf{a} = \mathbf{0}$, which then implies that

$$\underbrace{\left(\mathbf{B}_{3(1,2)} - \mathbf{B}_{123}\mathbf{R}_1 - \mathbf{B}_{13}\mathbf{R}_3 - \mathbf{B}_{23}\mathbf{R}_4\right)}_{\in\langle\mathbf{U}_3\rangle}\mathbf{a} = \underbrace{\left(\mathbf{B}_{12}\mathbf{R}_2 + \mathbf{B}_{2(1,3)}\mathbf{R}_7\right)}_{\in\langle\mathbf{U}_2\rangle}\mathbf{a} \triangleq \mathbf{b}. \quad (148)$$

Since $\mathbf{B}_{3(1,2)}, \mathbf{B}_{123}, \mathbf{B}_{13}, \mathbf{B}_{23}$ are disjoint submatrices of the basis matrix for $\langle\mathbf{U}_{3(1,2)}\rangle$ according to *(P10)*, they are linearly independent by construction. It follows that,

1) $(\mathbf{B}_{3(1,2)} - \mathbf{B}_{123}\mathbf{R}_1 - \mathbf{B}_{13}\mathbf{R}_3 - \mathbf{B}_{23}\mathbf{R}_4)$ has full column-rank equal to $b_{3(1,2)}$. This is because if on the contrary, there exists a non-zero vector $\mathbf{z}$ such that $(\mathbf{B}_{3(1,2)} - \mathbf{B}_{123}\mathbf{R}_1 - \mathbf{B}_{13}\mathbf{R}_3 - \mathbf{B}_{23}\mathbf{R}_4)\mathbf{z} = \mathbf{0}$, then $\mathbf{B}_{3(1,2)}\mathbf{z} \in \langle[\mathbf{B}_{123}, \mathbf{B}_{13}, \mathbf{B}_{23}]\rangle$. Since $\mathbf{B}_{3(1,2)}$ has full column rank and $\mathbf{z} \neq \mathbf{0}$, this means $\langle\mathbf{B}_{3(1,2)}\rangle$ has non-trivial intersection with $\langle[\mathbf{B}_{123}, \mathbf{B}_{13}, \mathbf{B}_{23}]\rangle$, contradicting their linear independence.
2) $(\mathbf{B}_{3(1,2)} - \mathbf{B}_{123}\mathbf{R}_1 - \mathbf{B}_{13}\mathbf{R}_3 - \mathbf{B}_{23}\mathbf{R}_4)\mathbf{a} \triangleq \mathbf{b} \neq \mathbf{0}$, because of the previous observation and because $\mathbf{a}$ is a non-zero vector.
3) $\mathbf{b} \notin \langle[\mathbf{B}_{123}, \mathbf{B}_{23}]\rangle = \langle\mathbf{U}_{23}\rangle$, because if $\mathbf{b} \in \langle[\mathbf{B}_{123}, \mathbf{B}_{23}]\rangle$ then the non-zero vector $\mathbf{B}_{3(1,2)}\mathbf{a} = \mathbf{b} + \mathbf{B}_{123}\mathbf{R}_1\mathbf{a} + \mathbf{B}_{13}\mathbf{R}_3\mathbf{a} + \mathbf{B}_{23}\mathbf{R}_4\mathbf{a} \in \langle[\mathbf{B}_{123}, \mathbf{B}_{23}, \mathbf{B}_{13}]\rangle$, which is a contradiction because $\mathbf{B}_{3(1,2)}, \mathbf{B}_{123}, \mathbf{B}_{13}, \mathbf{B}_{23}$ are linearly independent.
4) $\mathbf{b} \in \langle\mathbf{U}_3\rangle$. This follows from (148). Specifically, since $\mathbf{B}_{3(1,2)}, \mathbf{B}_{123}, \mathbf{B}_{13}, \mathbf{B}_{23}$ are all submatrices of the basis matrix for $\langle\mathbf{U}_3\rangle$ according to *(P13)*, and $\mathbf{b}$ is their linear combination, this implies that $\mathbf{b} \in \langle\mathbf{U}_3\rangle$.
5) $\mathbf{b} \in \langle\mathbf{U}_2\rangle$. This also follows from (148) by similar reasoning.
6) From enumerated items 4 and 5 above, we have, $\mathbf{b} \in \langle\mathbf{U}_2\rangle$ and $\mathbf{b} \in \langle\mathbf{U}_3\rangle$, and thus $\mathbf{b} \in \langle\mathbf{U}_2\rangle \cap \langle\mathbf{U}_3\rangle = \langle\mathbf{U}_{23}\rangle$, which contradicts item 3.

The contradiction proves the desired result that $\mathbf{R}_5$ has full column rank, i.e., it is an invertible square matrix. Similarly we can prove that $\mathbf{R}_7$ has full column rank, also an invertible square matrix. The last three steps hinge on this property.

Step 12: Redefine $\mathbf{B}_{3(1,2)}$ as

$$\mathbf{B}^{\text{new}}_{3(1,2)} = \mathbf{B}^{\text{old}}_{3(1,2)} - \mathbf{B}_{123}\mathbf{R}_1 - \mathbf{B}_{13}\mathbf{R}_3 - \mathbf{B}_{23}\mathbf{R}_4. \quad (149)$$

Step 13: Redefine $\mathbf{B}_{2(1,3)}$ as

$$\mathbf{B}^{\text{new}}_{2(1,3)} = \mathbf{B}^{\text{old}}_{2(1,3)}\mathbf{R}_7 + \mathbf{B}_{12}\mathbf{R}_2. \quad (150)$$

Step 14: Redefine $\mathbf{B}_{1(2,3)}$ as

$$\mathbf{B}^{\text{new}}_{1(2,3)} = \mathbf{B}^{\text{old}}_{1(2,3)}\mathbf{R}_5. \quad (151)$$

Since $\mathbf{R}_5$ and $\mathbf{R}_7$ are invertible square matrices, it follows by *(P6)* and *(P7)* that $\mathbf{B}^{\text{new}}_{1(2,3)}$, $\mathbf{B}^{\text{new}}_{2(1,3)}$ and $\mathbf{B}^{\text{new}}_{3(1,2)}$ are also feasible choices in Steps 5,6, and 7. Thus, *(P8)-(P19)* are still satisfied after $\mathbf{B}_{1(2,3)}$, $\mathbf{B}_{2(1,3)}$ and $\mathbf{B}_{3(1,2)}$ are replaced with $\mathbf{B}^{\text{new}}_{1(2,3)}, \mathbf{B}^{\text{new}}_{2(1,3)}, \mathbf{B}^{\text{new}}_{3(1,2)}$, respectively. However, because of the last three steps and (147), *(P20)* is now satisfied as well with $\mathbf{B}^{\text{new}}_{1(2,3)}, \mathbf{B}^{\text{new}}_{2(1,3)}, \mathbf{B}^{\text{new}}_{3(1,2)}$, i.e.,

$$\mathbf{B}^{\text{new}}_{3(1,2)} = \mathbf{B}^{\text{new}}_{2(1,3)} + \mathbf{B}^{\text{new}}_{1(2,3)}. \quad (152)$$

This concludes the proof of Lemma 2. □

Appendix D
Comparison of Lemma 2 to the Channel Decomposition of [55]

Reference [55, Chapter 3] explores the DoF of a 3 user MIMO broadcast channel where the transmitter has m antennas, and the k^{th} receiver has n_k antennas, $k \in [3]$. The channel is specified by $Y_1 = H_1X + Z_1$, $Y_2 = H_2X + Z_2$ and $Y_3 = H_3X + Z_3$, where $X \in \mathbb{C}^{m\times 1}$ denotes the input of the channel and $Y_k \in \mathbb{C}^{n_k\times 1}, k \in [3]$ denotes the output of the broadcast channel at the k^{th} receiver. $H_k \in \mathbb{C}^{n_k\times m}, k \in [3]$ denotes the channel matrix between the transmitter and the k^{th} receiver. Z_1, Z_2 and Z_3 are independent Gaussian noise vectors with zero mean and identity covariance matrix. There are independent messages desired by various subsets of receivers. As apparent from the high level description, the overall 3 user MIMO BC DoF question does not allow any direct mapping to our 3 user LCBC capacity question, e.g., the LCBC formulation has no notion of channel matrices, all users receive the same broadcast symbols, whereas the MIMO BC problem has no notion of side-information or linear computations.

What connects the two problems is that they both involve a decomposition of 3 subspaces. In the LCBC, the 3 subspaces of interest are $\langle\mathbf{U}_1\rangle, \langle\mathbf{U}_2\rangle, \langle\mathbf{U}_3\rangle$ as in Lemma 2. In the MIMO BC the corresponding subspaces are $\langle N_1\rangle, \langle N_2\rangle, \langle N_3\rangle$, defined as the null spaces of the channel matrices H_1^T, H_2^T, H_3^T, respectively. The decompositions parallel each other very closely. Intuitively, even though the context surrounding these subspaces is quite different in each problem, the subspaces are similar mathematical objects, so it makes sense that they should have similar properties, e.g., similar decompositions. The following table establishes a one-to-one correspondence of the subspace decompositions in the two settings. The notation in the right column of the table follows the definitions in [55].

A distinction is apparent in the first and last rows of the table. According to the first row, [55] assumes that the three subspaces have empty intersection, whereas Lemma 2 accounts for this space with the basis representation $\mathbf{B}_{123}$. On the other hand, according to the last row, Lemma 2 assumes the complement of the span of three subspaces is empty, whereas [55] accounts for this space with the basis representation V_{123}. This distinction arises mainly because in the LCBC setting the complement of span of the three spaces is uninteresting (data dimensions that are neither known

	Lemma 1	[55]
Field	$\mathbb{F}_q$	$\mathbb{C}$
Dimension of universe	d	m
3 subspaces to decompose	$\langle \mathbf{U}_1 \rangle, \langle \mathbf{U}_2 \rangle, \langle \mathbf{U}_3 \rangle$	$\langle N_1 \rangle, \langle N_2 \rangle, \langle N_3 \rangle$
Corresponding Bases	$\mathbf{B}_{123}$	$\emptyset$
	$\mathbf{B}_{12}$	V_3
	$\mathbf{B}_{13}$	V_2
	$\mathbf{B}_{23}$	V_1
	$\mathbf{B}_{1(2,3)}$	V_{23X}
	$\mathbf{B}_{2(1,3)}$	V_{13X}
	$\mathbf{B}_{3(1,2)}$	V_{12X}
	$\mathbf{B}_{1c}$	V_{23R}
	$\mathbf{B}_{2c}$	V_{13R}
	$\mathbf{B}_{3c}$	V_{12R}
	$\emptyset$	V_{123}

TABLE I: Correspondence between Lemma 2 and the channel decomposition in [55]

nor desired by any user), whereas in the MIMO BC the intersection of the three nullspaces is similarly uninteresting (transmit dimensions that are nulled at every receiver), and each problem naturally eliminates the uninteresting spaces for simplicity. The distinction is not significant, however, since the omitted spaces can be trivially included. A bit more significant distinction that is not apparent from the table is that as an additional feature, Lemma 2 chooses the basis matrices $\mathbf{B}_{1(2,3)}, \mathbf{B}_{2(1,3)}, \mathbf{B}_{3(1,2)}$ carefully to satisfy the matrix-sum condition $\mathbf{B}_{1(2,3)} + \mathbf{B}_{2(1,3)} = \mathbf{B}_{3(1,2)}$, which facilitates code design in the LCBC. Such an explicit specialization of bases is not considered by the corresponding construction in [55]. Setting aside these finer distinctions, it is quite remarkable that the subspace decompositions in [55] and Lemma 2, obtained independently in different contexts, turn out to be in one-to-one correspondence. The one-to-one correspondence constitutes strong evidence of the fundamental significance of the decomposition, as well as a verification of the same concept from two perspectives.

While the two decompositions are intuitively similar, there are several underlying technical details that prevent the direct application of the decomposition in [55] to the LCBC problem. Note that the proof in [55, Sec. 3] relies on the existence of an orthogonal complement, i.e., a linear subspace that is both orthogonal and complementary to a given linear subspace. The existence of an orthogonal complement is guaranteed over $\mathbb{C}$, but not over $\mathbb{F}_q$. For instance, self-orthogonality is a prominent theme in error correction code design over finite fields. Removing the requirement of orthogonality and just using any complement space instead does not automatically resolve the issue, because the complement space needs to be chosen carefully to achieve the correct alignment of spaces. The orthogonality of the complement space helps to achieve the desired alignment in the proof of [55, Sec. 3]. Over $\mathbb{F}_q$, since we are not guaranteed orthogonal complements, this choice is non-trivial). This is important because the alignment aspect of subspaces (any two subspaces contain the third) is what makes the subspace decomposition non-trivial. Furthermore, the proof of correctness of the subspace decomposition in [55, Sec. 3] applies to *almost all* spaces, since the argument relies on the values taken by ranks *almost surely*. Over arbitrary $\mathbb{F}_q$, an 'almost-surely' guarantee is not meaningful. Indeed, the proof of correctness of the decomposition is shown for *all* realizations in Lemma 2.[5]

Let us introduce a simple setup to further illustrate these points. Consider the following three (complex) channel matrices H_1, H_2, H_3.

$$H_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},\ H_2 = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix},\ H_3 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{153}$$

Denote $\mathcal{N}(A)$ as the nullspace of A, i.e., the set of X such that $AX = 0$. Let $N_k, k \in \{1, 2, 3\}$ be a basis (written in columns vectors) of $\mathcal{N}(H_k)$, i.e.,

$$N_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},\quad N_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix},\quad N_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}. \tag{154}$$

Let $\langle A \rangle$ denote the linear subspace spanned by the columns of A. For example, $\langle N_1 \rangle = \mathcal{N}(H_1)$. $A \cap B$ denotes a basis that spans the subspace $\langle A \rangle \cap \langle B \rangle$. In addition, if $\langle A \rangle$ is a subspace of $\langle U \rangle$, then let $A_U^\perp$ denote a basis of the intersection of $\mathcal{N}(A^T)$ with $\langle U \rangle$. It follows that $A^T A_U^\perp = \mathbf{0}$ and $\text{rk}(A) + \text{rk}(A_U^\perp) = \text{rk}(U)$. Note that for A defined in $\mathbb{C}$, $[A, A_U^\perp]$ spans $\langle U \rangle$. Thus, $A_U^\perp$ is the 'orthogonal complement' of A (within the subspace $\langle U \rangle$). In particular, $\langle A \rangle$ and $\langle A_U^\perp \rangle$ have no non-trivial intersection. However, this is not always true in finite fields, e.g, $A = [1, 1]^T \in \mathbb{F}_2^{2\times 1}$, and $A^\perp = [1, 1]^T = A$.

Using the same notation as in [55, Sec. 3], let us define

$$H_{123} \triangleq \begin{bmatrix} H_1 \\ H_2 \\ H_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$
$$H_{12} \triangleq \begin{bmatrix} H_1 \\ H_2 \end{bmatrix},\ H_{13} \triangleq \begin{bmatrix} H_1 \\ H_3 \end{bmatrix},\ H_{23} \triangleq \begin{bmatrix} H_2 \\ H_3 \end{bmatrix}. \tag{155}$$

With these definitions, let us apply the decomposition method in [55, Sec. 3.4.6] on N_1, N_2, N_3. A summary of the steps is given below.

1) Find $V_1 = N_2 \cap N_3$ as a basis of $\mathcal{N}(H_{23})$.
2) Find $V_2 = N_1 \cap N_3$ as a basis of $\mathcal{N}(H_{13})$.
3) Find $V_3 = N_1 \cap N_2$ as a basis of $\mathcal{N}(H_{12})$.
4) Find V_{13} as a basis of the orthogonal complement of $\langle [V_1, V_3] \rangle$ within $\langle N_2 \rangle$.
5) Find V_{23} as a basis of the orthogonal complement of $\langle [V_2, V_3] \rangle$ within $\langle N_1 \rangle$.
6) Find V_{12} as a basis of the orthogonal complement of $\langle [V_1, V_2] \rangle$ within $\langle N_3 \rangle$.
7) Find independent bases $V_{13X}, V_{12X}, V_{12R}, V_{13R}, V_{23R}$ by (3.43) - (3.47) of [55].

By Steps 1-3,

$$V_1 = N_2 \cap N_3 = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}^T, \tag{156}$$
$$V_2 = N_1 \cap N_3 = [\,], \tag{157}$$

[5]It is noteworthy that the proof of Lemma 2 also extends to the field of complex numbers. The proof is based on linear independence/dependence of subspaces which holds over both $\mathbb{F}_q$ and $\mathbb{C}$.

$$V_3 = N_1 \cap N_2 = [\,]. \tag{158}$$

At this point, note that the next step, which is to construct V_{13} does not work. According to [55] (3.38),

$$V_{13} = (H_{123}^T H_{123})^{-1} H_{123}^T (H_{123}[V_1, V_3])^{\perp}_{H_{123}N_2}, \tag{159}$$

but $(H_{123}^T H_{123})$ is not invertible although H_{123} has full column rank 3. This is because the orthogonal complement of H_{123} has nontrivial intersection with itself. This can happen in $\mathbb{F}_q$ but not in $\mathbb{C}$.

Alternatively, if we use the implicit definition, [55, (3.37)], we may avoid the inversion of $H_{123}^T H_{123}$, but a similar problem will again emerge. [55, (3.37)] requires that

$$H_{123} V_{13} = (H_{123}[V_1, V_3])^{\perp}_{H_{123}N_2}, \tag{160}$$

which is a basis of the orthogonal complement of $\langle H_{123}[V_1, V_3]\rangle$ within the subspace $\langle H_{123}N_2\rangle$. Note that

$$H_{123}[V_1, V_3] = \begin{bmatrix}1 & 1 & 0 & 0 & 0\end{bmatrix}^T \tag{161}$$

and

$$H_{123}N_2 = \begin{bmatrix}1 & 1 & 0 & 0 & 0\\ 1 & 0 & 0 & 1 & 0\end{bmatrix}^T. \tag{162}$$

It is readily verified that the only solution to $(H_{123}[V_1, V_3])^{\perp}_{H_{123}N_2} = [1,1,0,0,0]^T$. Therefore, we obtain that $H_{123}V_{13} = [1,1,0,0,0]^T$. This gives us $V_{13} = [1,1,0]^T = V_1$, which is linearly *dependent* on V_1. However, V_{13} is required to be linearly independent of V_1 in Step 4.

Next let us consider Step 7. At a high level, the motivation of Step 7 is that the three spaces V_{13}, V_{12} and V_{23} are not independent in general and therefore a finer decomposition is needed. In [55], 6 subspaces are introduced, namely $V_{13X}, V_{13R}, V_{12X}, V_{12R}, V_{23X}, V_{23R}$, so that $[V_{13X}, V_{13R}]$ spans $\langle V_{13}\rangle$, $[V_{12X}, V_{12R}]$ spans $\langle V_{12}\rangle$, and $\langle [V_{23X}, V_{23R}]\rangle$ spans $\langle V_{12}\rangle$. Note that these subspaces identified by the algorithm have such properties that

- $[V_{13X}, V_{13R}, V_{12X}, V_{12R}, V_{23R}]$ are independent and span $\langle [V_{13}, V_{12}, V_{23}]\rangle$. Besides, $[V_{**X}, V_{**R}]$ spans $\langle V_{**}\rangle$, for $** \in \{13, 12, 23\}$.
- In addition, V_{23X} is linearly representable by $[V_{13X}, V_{12X}]$, i.e., $\langle V_{23X}\rangle \subset \langle [V_{13X}, V_{12X}]\rangle$. Also, $\langle V_{12X}\rangle \subset \langle [V_{13X}, V_{23X}]\rangle$ and $\langle V_{13X}\rangle \subset \langle [V_{12X}, V_{23X}]\rangle$.
- $V_{13X}, V_{12X}, V_{23X}$ are aligned in a way such that $\langle H_1 V_{13X}\rangle = \langle H_1 V_{12X}\rangle$, $\langle H_2 V_{12X}\rangle = \langle H_2 V_{23X}\rangle$ and $\langle H_3 V_{13X}\rangle = \langle H_3 V_{23X}\rangle$.

The critical alignment is the second one, i.e., we need $V_{13X}, V_{12X}, V_{23X}$ such that each one is contained in the span of the other two. Let us see what happens if we replace the 'orthogonal complement' space (which may not exist over $\mathbb{F}_q$) with any 'complement' space (which do exist over $\mathbb{F}_q$). The following toy example shows that simply replacing 'orthogonal complement' with 'any complement' may not work. Suppose we are given,

$$N_1 = \begin{bmatrix}1 & 0\\ 0 & 1\\ 0 & 0\end{bmatrix}, \quad N_2 = \begin{bmatrix}1 & 0\\ 1 & 1\\ 0 & 1\end{bmatrix}, \quad N_3 = \begin{bmatrix}1\\ 1\\ 1\end{bmatrix}, \tag{163}$$

with entries all defined in $\mathbb{F}_2$. It follows by definition that,

$$V_1 = N_2 \cap N_3 = [\,],\ V_2 = N_1 \cap N_3 = [\,],$$
$$V_3 = N_1 \cap N_2 = \begin{bmatrix}1\\ 1\\ 0\end{bmatrix}. \tag{164}$$

Next, say we choose the complements (not necessarily orthogonal) as,

$$V_{13} = \begin{bmatrix}0\\ 1\\ 1\end{bmatrix}, \quad V_{12} = \begin{bmatrix}1\\ 1\\ 1\end{bmatrix}, \quad V_{23} = \begin{bmatrix}0\\ 1\\ 0\end{bmatrix}, \tag{165}$$

so that $[V_1, V_3, V_{13}]$ span $\langle N_2\rangle$, $[V_1, V_2, V_{12}]$ span $\langle N_3\rangle$, and $[V_2, V_3, V_{23}]$ span $\langle N_1\rangle$. Such an attempt to translate Steps 1 – 6 to the finite field case does not work because now we see that $[V_1, V_2, V_3, V_{13}, V_{12}, V_{23}]$ are not linearly independent. In particular, with this choice there is no non-trivial $V_{13X}, V_{23X}, V_{12X}$ so that any one is contained in the span of the other two. What this shows is that the complement spaces V_{13}, V_{12}, V_{23} need to be chosen carefully. The proof in [55] does not face this problem, because orthogonal spaces and complement spaces are compatible for complex numbers and therefore calculations of orthogonal complements help to identify the appropriate V_{13}, V_{12}, V_{23}. Over $\mathbb{F}_q$ this does not work. Fortunately, there does exist a solution to $[V_{13}, V_{12}, V_{23}]$ so that such non-trivial $V_{13X}, V_{23X}, V_{12X}$ can be found (see Steps 11–14 in Appendix C for the details of this key aspect of the proof in the general case). Indeed, with a more careful choice of subspaces we have,

$$V_{13} = \begin{bmatrix}0\\ 1\\ 1\end{bmatrix}, \quad V_{12} = \begin{bmatrix}1\\ 1\\ 1\end{bmatrix}, \quad V_{23} = \begin{bmatrix}1\\ 0\\ 0\end{bmatrix}, \tag{166}$$

$$V_{13X} = \begin{bmatrix}0\\ 1\\ 1\end{bmatrix}, \quad V_{12X} = \begin{bmatrix}1\\ 1\\ 1\end{bmatrix}, \quad V_{23X} = \begin{bmatrix}1\\ 0\\ 0\end{bmatrix}, \tag{167}$$

and then

$$V_{13R} = V_{12R} = V_{23R} = [\,]. \tag{168}$$

Thus, following our proof we successfully found the $V_{13X}, V_{13R}, V_{12X}, V_{12R}, V_{23X}, V_{23R}$ that satisfy the desired (first two) properties described in Step 7.

ACKNOWLEDGEMENT

The authors would like to thank the anonymous referees for their insightful comments and suggestions that improved the presentation of this article.

REFERENCES

[1] S. Research, "The next hyper — connected experience for all," *White Paper*, July 2020. [Online]. Available: {https://cdn.codeground.org/nsr/downloads/researchareas/6GVision.pdf}

[2] H. Witsenhausen, "The zero-error side information problem and chromatic numbers (corresp.)," *IEEE Trans. Inf. Theory*, vol. 22, no. 5, pp. 592–593, Sep. 1976.

[3] J. Korner and K. Marton, "How to encode the modulo-two sum of binary sources," *IEEE Trans. Inf. Theory*, vol. 25, pp. 219–221, Mar. 1979.

[4] R. Ahlswede and N. Cai, "On communication complexity of vector-valued functions," *IEEE Trans. Inf. Theory*, vol. 40, no. 6, pp. 2062–2067, Nov. 1994.

[5] A. Orlitsky and J. R. Roche, "Coding for computing," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 903–917, Mar. 2001.

[6] V. Doshi, D. Shah, M. Medard, and M. Effros, "Functional compression through coloring," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3901–3917, Aug. 2010.

[7] B. K. Rai and B. K. Dey, "On network coding for sum-networks," *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 50–63, 2012.

[8] A. Giridhar and P. R. Kumar, "Computing and communicating functions over sensor networks," *IEEE J. Sel. Areas in Commun.*, vol. 23, no. 23, pp. 755–764, Apr. 2005.

[9] H. Kowshik and P. R. Kumar, "Optimal function computation in directed and undirected graphs," *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3407–3418, Jun. 2012.

[10] A. Ramamoorthy and M. Langberg, "Communicating the sum of sources over a network," *IEEE J. Sel. Areas in Commun.*, vol. 31, no. 4, pp. 655–665, Apr. 2013.

[11] S. Feizi and M. Medard, "On network function compression," *IEEE Trans. Inf. Theory*, vol. 60, no. 9, pp. 5387–5401, Sep. 2014.

[12] R. Appuswamy, M. Franceschetti, N. Karamchandani, and K. Zeger, "Network coding for computing: Cut-set bounds," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 1015–1030, Feb. 2011.

[13] ——, "Linear codes, target function classes, and network computing capacity," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5741–5753, Sep. 2013.

[14] R. Appuswamy and M. Franceschetti, "Computing linear functions by linear coding over networks," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 422–431, Jan. 2014.

[15] X. Guang, R. W. Yeung, S. Yang, and C. Li, "Improved upper bound on the network function computing capacity," *IEEE Trans. Inf. Theory*, vol. 65, no. 6, pp. 3790–3811, Jun. 2019.

[16] M. Maddah-Ali and U. Niesen, "Fundamental limits of caching," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2856–2867, May 2014.

[17] Q. Yu, M. A. Maddah-Ali, and A. S. Avestimehr, "The exact rate-memory tradeoff for caching with uncoded prefetching," *IEEE Trans. Inf. Theory*, vol. 66, no. 2, pp. 1281–1296, Feb 2018.

[18] Mingyue Ji, G. Caire, and A. F. Molisch, "Fundamental limits of caching in wireless d2d networks," *IEEE Trans. Inf. Theory*, vol. 62, no. 2, pp. 849–869, Feb 2016.

[19] B. Chor, E. Kushilevitz, O. Goldreich, and M. Sudan, "Private information retrieval," *J. ACM*, vol. 45, no. 6, pp. 965–981, 1998.

[20] S. Ulukus, S. Avestimehr, M. Gastpar, S. A. Jafar, R. Tandon, and C. Tian, "Private retrieval, computing, and learning: Recent progress and future challenges," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 3, pp. 729–748, 2022.

[21] S. Li, M. A. Maddah-Ali, Q. Yu, and A. S. Avestimehr, "A fundamental tradeoff between computation and communication in distributed computing," *IEEE Trans. Inf. Theory*, vol. 64, no. 1, pp. 109–128, Jan 2018.

[22] A. G. Dimakis, K. Ramchandran, Y. Wu, and C. Suh, "A survey on network codes for distributed storage," *Proceedings of the IEEE*, vol. 99, no. 3, pp. 476–489, 2011.

[23] S. Liu and F. Oggier, "An overview of coding for distributed storage systems," *Network Coding and Subspace Designs*, pp. 363–383, 2018.

[24] T. Ernvall, "Codes between MBR and MSR points with exact repair property," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 6993–7005, 2014.

[25] Y. Birk and T. Kol, "Coding on demand by an informed source (ISCOD) for efficient broadcast of different supplemental data to caching clients," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2825–2830, June 2006.

[26] F. Arbabjolfaei and Y.-H. Kim, "Fundamentals of index coding," *Found. Trends Commun. Inf. Theory*, vol. 14, no. 3–4, p. 163–346, 2018. [Online]. Available: https://doi.org/10.1561/0100000094

[27] J. S. Ng, W. Y. B. Lim, N. C. Luong, Z. Xiong, A. Asheralieva, D. Niyato, C. Leung, and C. Miao, "A comprehensive survey on coded distributed computing: Fundamentals, challenges, and networking applications," *IEEE Communications Surveys and Tutorials*, vol. 23, no. 3, pp. 1800–1837, June 2021.

[28] S. Li, M. A. Maddah-Ali, and A. S. Avestimehr, "Coding for distributed fog computing," *IEEE Communications Magazine*, vol. 55, no. 4, pp. 34–40, April 2017.

[29] K. Lee, M. Lam, R. Pedarsani, D. Papailiopoulos, and K. Ramchandran, "Speeding up distributed machine learning using codes," *IEEE Trans. Inf. Theory*, vol. 64, no. 3, pp. 1514–1529, 2017.

[30] P. Kairouz, H. B. McMahan, B. Avent, A. Bellet, M. Bennis, A. N. Bhagoji, K. Bonawitz, Z. Charles, G. Cormode, R. Cummings *et al.*, "Advances and open problems in federated learning," *Foundations and Trends in Machine Learning*, vol. 14, no. 1-2, pp. 1–210, 2021.

[31] S. Kadhe, N. Rajaraman, O. O. Koyluoglu, and K. Ramchandran, "FastSecAgg: Scalable secure aggregation for privacy-preserving federated learning," *ArXiv:2009.11248*, Sep. 2020.

[32] J. Ruan and D. Xie, "Networked VR: State of the art, solutions and challenges," *Electronics*, vol. 10, no. 2, January 2021.

[33] F. Hu, Y. Deng, W. Saad, M. Bennis, and A. H. Aghvami, "Cellular-connected wireless virtual reality: Requirements, challenges, and solutions," *IEEE Communications Magazine*, vol. 58, no. 5, pp. 105–111, May 2020.

[34] E. Bastug, M. Bennis, M. Medard, and M. Debbah, "Toward interconnected virtual reality: Opportunities, challenges, and enablers," *IEEE Communications Magazine*, vol. 55, no. 6, pp. 110–116, Aug. 2017.

[35] L. Zhao, Y. Cui, S. Yang, and S. S. Shitz, "An optimization framework for general rate splitting for general multicast," *IEEE Trans. on Wireless Commun.*, vol. 22, no. 3, pp. 1573–1587, 2022.

[36] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Networks*, vol. 34, no. 3, pp. 134–142, May/June 2020.

[37] Y. Geng, S. Liu, Z. Yin, A. Naik, B. Prabhakar, M. Rosunblum, and A. Vahdat, "Exploiting a natural network effect for scalable, fine-grained clock synchronization," *Proceedings of the 15th USENIX Conference on Networked Systems Design and Implementation*, pp. 81–94, 2018.

[38] P. Fei, Z. Chen, Z. Wang, and S. A. Jafar, "Communication-efficient clock synchronization," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, 2022, pp. 2592–2597.

[39] T. M. C. Paul Cuff, Haim H. Permuter, "Coordination capacity," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4181–4206, 2010.

[40] H. Sun and S. Jafar, "On the capacity of computation broadcast," *IEEE Trans. Inf. Theory*, vol. 66, no. 6, pp. 3417–3434, Jun. 2020.

[41] T. S. J. Z. Bar-Yossef, Y. Birk and T. Kol, "Index coding with side information," *IEEE Trans. Inf. Theory*, vol. 57, no. 3, pp. 1479 – 1494, Mar. 2011.

[42] S. Rouayheb, A. Sprintson, and C. Georghiades, "On the index coding problem and its relation to network coding and matroid theory," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3187–3195, Jul. 2010.

[43] Y. Berliner and M. Langberg, "Index coding with outerplanar side information," *Proceedings of IEEE International Symposium on Information Theory*, pp. 806–810, Jul./Aug. 2011.

[44] L. Ong, C. K. Ho, and F. Lim, "The single-uniprior index-coding problem: The single-sender case and the multi-sender extension," *IEEE Trans. Inf. Theory*, vol. 62, no. 6, pp. 3165–3182, Jun. 2016.

[45] S. A. Jafar, "Topological interference management through index coding," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 529–568, Jan. 2014.

[46] A. Agarwal, L. Flodin, and A. Mazumdar, "Linear programming approximations for index coding," *IEEE Trans. Inf. Theory*, vol. 65, no. 9, pp. 5547–5564, Sep. 2019.

[47] A. Blasiak, R. Kleinberg, and E. Lubetzky, "Index coding via linear programming," Apr. 2010. [Online]. Available: https://arxiv.org/pdf/1004.1379.pdf

[48] N. Lee, A. Dimakis, and R. W. Heath, "Index coding with coded side-information," *IEEE Communication Letters*, vol. 19, no. 3, pp. 319–322, Mar. 2015.

[49] R. Peeters, "Orthogonal representations over finite fields and the chromatic number of graphs," *Combinatorica*, pp. 417–431, Sept. 1996.

[50] D. Chawin and I. Haviv, "Improved np-hardness of approximation for orthogonality dimension and minrank," *SIAM Journal on Discrete Mathematics*, vol. 37, no. 4, pp. 2670–2688, 2023. [Online]. Available: https://doi.org/10.1137/23M155760X

[51] M. Langberg and A. Sprintson, "On the hardness of approximating the network coding capacity," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 1008–1014, 2011.

[52] E. Lubetzky and U. Stav, "Nonlinear index coding outperforming the linear optimum," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3544–3551, Aug. 2009.

[53] A. Blasiak, R. Kleinberg, and E. Lubetzky, "Lexicographic products and the power of non-linear network coding," *IEEE symposium on Foundations of Computer Science (FOCS), Palm Springs, CA*, pp. 609–618, Oct. 2011. [Online]. Available: http://arxiv.org/abs/1108.2489

[54] H. Maleki, V. Cadambe, and S. Jafar, "Index coding – an interference alignment perspective," *IEEE Trans. Inf. Theory*, vol. 60, no. 9, pp. 5402–5432, Sep. 2014.

[55] Y. Wang, "Degrees of freedom of mimo wireless networks with general message sets: Channel decomposition, message splitting and beamformer allocation," Ph.D. dissertation, University of Colorado at Boulder, 2017.
[56] T. Tao, "Sumset and inverse sumset theory for shannon entropy," *Combinatorics, Probability and Computing*, vol. 19, no. 4, pp. 603–639, 2010.
[57] M. M. I. Kontoyiannis, "Sumset and inverse sumset inequalities for differential entropy and mutual information," *IEEE Trans. Inf. Theory*, vol. 60, no. 8, pp. 4503–4514, Aug. 2014.
[58] K. Janich, *Linear Algebra*, ser. Undergraduate texts in mathematics. Springer, 2007. [Online]. Available: https://books.google.com/books?id=1VyoPwAACAAJ
[59] S. A. Jafar, "Interference alignment: A new look at signal dimensions in a communication network," *Found. Trends Commun. Inf. Theory*, vol. 7, 2011.
[60] C. Wang, T. Gou, and S. Jafar, "Subspace alignment chains and the degrees of freedom of the three user mimo interference channel," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2432–2479, May 2014.
[61] Q. Yu, M. A. Maddah-Ali, and A. S. Avestimehr, "Characterizing the rate-memory tradeoff in cache networks within a factor of 2," *IEEE Trans. Inf. Theory*, vol. 65, no. 1, pp. 647–663, 2018.
[62] J. Edmonds, "Submodular functions, matroids and certain polyhedra," *Combinatorial Structures and Their Applications*, vol. 38, pp. 69–87, 1970.
[63] H. P. Romero, "Fundamental limits of network communication with general message sets: A combinatorial approach," Ph.D. dissertation, University of Colorado at Boulder, 2014.
[64] M. T. Keller and W. T. Trotter, *Applied combinatorics*, 2017.
[65] S. T. Goh, "Prove that if the constraints to a linear program are integer, then the optimal solution is rational." [Online]. Available: https://math.stackexchange.com/q/2999239
[66] Y. Yao and S. A. Jafar, "Capacity of 2-user linear computation broadcast with partially informed server," *IEEE Communications Letters*, vol. 26, no. 9, pp. 2014–2017, 2022.

Yuhang Yao (Graduate Student Member, IEEE) received the B.E. degree in Communications Engineering from Sun Yat-sen University, Shenzhen, China, in 2021, and the M.S. degree in Electrical and Computer Engineering from the University of California at Irvine, CA, USA, in 2024, where he is currently pursuing the Ph.D degree. His research interests include information theory, quantum information theory and their applications to distributed computation, security and privacy.

Syed A. Jafar (Fellow, IEEE) received his B. Tech. from IIT Delhi, India, in 1997, M.S. from Caltech, USA, in 1999, and Ph.D. from Stanford, USA, in 2003, all in Electrical Engineering. His industry experience includes positions at Lucent Bell Labs and Qualcomm. He is a Chancellor's Professor of Electrical Engineering and Computer Science at the University of California Irvine, Irvine, CA USA. His research interests include multiuser information theory, wireless communications and network coding.

Dr. Jafar is a recipient of the New York Academy of Sciences Blavatnik National Laureate in Physical Sciences and Engineering, the NSF CAREER Award, the ONR Young Investigator Award, the UCI Academic Senate Distinguished Mid-Career Faculty Award for Research, the School of Engineering Mid-Career Excellence in Research Award and the School of Engineering Maseeh Outstanding Research Award. His co-authored papers have received the IEEE Information Theory Society Paper Award, IEEE Communication Society and Information Theory Society Joint Paper Award, IEEE Communications Society Best Tutorial Paper Award, IEEE Communications Society Heinrich Hertz Award, IEEE Signal Processing Society Young Author Best Paper Award, IEEE Information Theory Society Jack Wolf ISIT Best Student Paper Award, and three IEEE GLOBECOM Best Paper Awards. Dr. Jafar received the UC Irvine EECS Professor of the Year award six times from the Engineering Students Council, a School of Engineering Teaching Excellence Award in 2012, and a Senior Career Innovation in Teaching Award in 2018. He was a University of Canterbury Erskine Fellow, an IEEE Communications Society Distinguished Lecturer, an IEEE Information Theory Society Distinguished Lecturer and a Thomson Reuters/Clarivate Analytics Highly Cited Researcher. He served as Associate Editor for IEEE TRANSACTIONS ON COMMUNICATIONS 2004-2009, for IEEE COMMUNICATIONS LETTERS 2008-2009 and for IEEE TRANSACTIONS ON INFORMATION THEORY 2009-2012.