Tensor Product Based Subspace Interference Alignment for Network Coding Applications

Viveck R. Cadambe, Syed A. Jafar†

Abstract—In this paper, we develop an interference alignment framework for multi-source (non-multicast) network coding applications. The framework developed here is based on using a tensor product structure for the network coding matrices at nodes in a network. The framework is presented in this paper in the specific context of exact repair for distributed storage, which is a multi-source network coding application. Using this framework, we generalize previously known MDS codes with optimal exact repair in distributed storage.

I. INTRODUCTION

Interference Alignment has emerged as a fundamental technique to manage interference in multi-source network capacity problems in information theory. Among the various tools associated with interference alignment, the technique of asymptotic interference alignment [1] has had a particularly profound impact. The technique has been applied to various wireless settings including the interference channel [1]–[3], the X channel [4], and the compound broadcast channel [5]. In each of these networks, the technique of asymptotic interference alignment achieves the optimal number of degrees of freedom (a capacity approximation). A key requisite of the technique is the existence of diagonal channel gain matrices over which alignment is performed. Diagonal channel matrices are obtained in wireless communication scenarios by grouping multiple input and output symbols together, treating each of these groups as a vector (super-symbol), and coding over these vectors. As the size of the super-symbols increase, the size of the diagonal matrices increase and the extent of alignment increases. In the limit of arbitrarily large symbol extensions, the extent of alignment is “maximum” and hence achieves degrees of freedom in the given wireless network. (See [6] for a tutorial).

Recently, the technique of interference alignment has been interpreted as a technique for network coding (albeit not random). This interpretation led to the observation that interference alignment plays a significant role in the multi-source non-multicast wireline network capacity problems. Multi-source non-multicast wireline network capacity is a classical open problem in network information theory with applications to distributed storage, peer-to-peer networks, and content distribution over the internet (see for instance, [7], [24]). For this class of problems, the technique of asymptotic interference alignment has been applied in two contexts. First, it has been applied to multiple unicasts in a class of (wireline) networks in [8], [9]. Second, it has been applied to the exact repair problem in distributed storage [10], which is in turn, a specific non-multicast (wireline) network coding scenario. The approach of these references can be broadly described as follows. In a network, if all the nodes apply vector linear network coding, then outputs at the destinations of the network can be expressed as linear transformations of the input, much like a wireless channel. If the nodes use random diagonal matrices for network coding, then the transformation from the sources to the destinations are indeed random diagonal matrices. This approach thus paves the way for aligning interference using asymptotic interference alignment.

While linear network coding enables us to inherit techniques from wireless networks into wireline networks (and vice-versa, see [11], [12]), there is one fundamental difference between wired and wireline networks. In wireless networks, the linear transformation which represents the channel matrices are given by nature and hence cannot be controlled (to a large extent). In contrast, in wired networks, linear transformations are a function of the coding matrices which are design parameters. This means that the structure of the end to end channel matrices can be controlled (to the extent allowed by the topology of the network graph). The canonical wireline network coding application where this difference is especially highlighted is exact repair in distributed storage systems [10], [13]–[20]. Specifically, in these references, this extra control offered by the network coding application is exploited to create finite non-asymptotic alignment schemes without sacrificing the extent to which interference is aligned. This is in contrast with wireless communications where, in general, asymptotic alignment is necessary; limiting the size of the symbol extension (i.e., super-symbol) in wireless communications strictly limits the extent to which interference can be aligned [1], [21]–[23]. The ability to control the linear transformations in wireline network communication scenarios spawns a need to under-
stand end-to-end linear transformations which are suitable for interference alignment. In this paper, we make progress in this context by developing a systematic approach to linear network code design in wireline networks to enable (non-asymptotic) interference alignment using a finite number of dimensions. We will particularly focus on the canonical problem of exact repair in distributed storage for exposition of our techniques and demonstration of its impact. We next explain the problem of exact repair.

II. EXACT REPAIR IN DISTRIBUTED STORAGE

Consider $k$ sources to be stored in a distributed storage system having $n$ equal-capacity storage nodes. All $k$ sources are assumed to be of equal size $L = M/k$ over a field $\mathbb{F}_q$ of size $q$. Source $i \in \{1, 2, \ldots, k\}$ is represented by the $L \times 1$ vector $a_i \in \mathbb{F}_q^L$. Note here that $M$ denotes the size of the total information stored in the distributed storage system, in terms of the number of elements over the field\(^1\). There are $n$ nodes, each with a storage capacity of size $L$. The data stored in the $n$ storage nodes can be interpreted as an $(n, k)$ code of the $k$ source symbols. Each node stores data of size $L$, i.e., each coded symbol of the $(n, k)$ code is a $L \times 1$ vector. The data stored in node $i$ is represented by $L \times 1$ vector $d_i$, where $i = 1, 2, \ldots, n$. We assume that our code is linear so that $d_i$ can be represented as

$$d_i = \sum_{j=1}^{k} C_{i,j} a_j,$$

where $C_{i,j}$ are $L \times L$ square matrices. Further, we restrict our codes to have a systematic structure, so that, for $i \in \{1, 2, \ldots, k\}$,

$$C_{i,j} = \begin{cases} I & j = i \\ 0 & j \neq i \end{cases}.$$

If the code storing the data is maximum-distance-separable (MDS), then, any $k$ storage nodes can reconstruct all the $k$ sources $a_1, a_2, \ldots, a_k$. Thus, an MDS code can protect against a failure of up to $(n - k)$ nodes. In this paper, we restrict our study to MDS codes, though the reader may note that most techniques developed here do not critically rely on the MDS property (though any claims of optimality rely on this property.) Therefore, we also do not focus on proving the MDS property of our codes. Such a proof may be found in the extended version of this paper [26]. We begin by describing an outline of an optimal repair strategies for $(n = k + 2, k)$ MDS codes - the solutions of [16], [18], [19] fall within this outline.

Consider the case where a single systematic node, node $i \in \{1, 2, \ldots, k\}$, fails. The goal of repair is to construct $d_i$ using the data in the surviving nodes $\{d_j : j \neq i\}$. For repair, $V_i d_j$ is downloaded by the new node from node $j \in \{1, 2, \ldots, n\} - \{i\}$ where $V_i$ is $L \times L$ matrix. Note that this strategy downloads a fraction of $\frac{1}{2}$ from each surviving node so that the total repair bandwidth is $\frac{L(n-1)}{n-k}$ as required. The linear combinations downloaded are of two types listed below.

1) The data downloaded from the surviving systematic nodes $j \in \{1, 2, \ldots, k\} - \{i\}$ contain no information of the failed node $a_i$, i.e.,

$$V_i d_j = V_i a_j, j \in \{1, 2, \ldots, k\} - \{i\}.$$

Note that there $\frac{k}{2}$ such linear combinations of each interfering component $a_j, j \in \{1, 2, \ldots, k\} - \{i\}$.

2) From each of the 2 parity nodes, $\frac{k}{2}$ linear combinations are downloaded. Therefore, a total of $L$ linear combinations are downloaded from all the parity nodes. The $L$

\(^1\)Note that we assume that $L, M$ are parameters of choice. This is a valid assumption for large amounts of information, since a large source can be split into several blocks, each of size $L$. Each block can be coded separately using our constructions.
components of \(a_i\) have to be reconstructed using these \(L\) linear combinations of the form \(V_i d_j, j = k + 1, k + 2\). For successful reconstruction of \(a_i\), the interference terms associated with \(a_j, j \in \{1, 2, \ldots, k\} - \{i\}\) contained in these linear combinations need to be cancelled completely (See Fig. 1).

The goal of our solution will be to completely cancel the interference from the second set of \(L\) linear combinations, using the former set of linear combinations, and then to regenerate \(a_1\) using the latter \(L\) interference-free linear combinations (See Fig. 1).

In our solution the coding sub-matrices associated with the first parity node are all (scaled) identity matrices, i.e., \(C_{k+1,i} = \lambda_i I_k\) for \(i = 1, 2, \ldots, k\) where \(\lambda_i\) is a scalar over the field \(\mathbb{F}_q\), so that

\[
d_{k+1} = \sum_{i=1}^{k} \lambda_{i} k+1, a_i
\]

Further, we denote \(C_{k+2,i} = H_i\) so that

\[
d_{k+2} = \sum_{i=1}^{k} H_i a_i
\]

While the above simplifications are restrictive, it turns out that they suffice to construct MDS codes for optimal repair. The simplifications help by revealing an interesting structure that they suffice to construct MDS codes for optimal repair. While the above simplifications are restrictive, it turns out that they suffice to construct MDS codes for optimal repair.

Note that we intend to completely cancel the impact of \(a_j\) using \(V_i a_j\). To do this, we need

\[
\text{rowspan}(V_i a_j) = \text{rowspan}(V_i), j \in \{1, 2, \ldots, k\} - \{i\}
\]

This ensures the desired interference alignment which hence ensures cancellation. After interference cancellation, we are left with \(L\) linear combinations of \(a_i\) of which \(L/2\) are of the form \(\lambda_i V_i a_i\) and the remaining \(L/2\) are of the form \(V_i H_i a_i\). To ensure that repair is successful, we need to ensure that \(a_i\) is linearly resolvable from these \(L\) linear combinations. To do so, we need to ensure that

\[
\text{rank} \begin{bmatrix} V_i & H_i I_k \end{bmatrix} = L, i = 1, 2, \ldots, k\]

\[
\Rightarrow \text{rowspan}(V_i) \cap \text{rowspan}(V_i H_i) = \{0\}
\]

for \(i = 1, 2, \ldots, k\)

Thus, the goal is to find matrices \(H_i\) such that (3) and (5) are satisfied. Before proceeding, it is worth noting that as \(k\) increases, the problem becomes more and more constrained. In other words, if we have a solution to (3) and (5) for, say \(k = 3\), then this automatically implies a solution for \(k = 2\) but not vice versa. For this reason, the problem was solved for \(k = 2\) in [13] and only recently [10], [16], [18], [19] for an arbitrary value of \(k\). Also note the similarities of the required constraints to the interference channel in [1]. This similarity was indeed exploited in [10] which used random diagonal matrices for \(H_i\) mimicking the wireless context. The reference then chose \(V_i\) using the asymptotic alignment technique. Here, we provide a tensor product based framework to solve this problem using a finite number of dimensions (i.e., for a finite value of \(L\)). For ease of exposition, we only focus on the case where \(k = 3, n = 5\) here. The solution for arbitrary values of \(k\) and \(n\) is very similar to the case where \(k = 3\).

### III. Tensor Product Based Subspace Alignment

The following lemma is a key observation in our development. In the lemma, the notation \(\otimes\) is used to denote the tensor (Kronecker) product between matrices.

**Lemma 1**: Let \(H = G_1 \otimes G_2 \otimes \ldots \otimes G_r\). Also let \(V = U_1 \otimes U_2 \otimes \ldots \otimes U_r\), where the matrix product \(U_i G_i\) is defined. Then,

- \(\text{rowspan}(VH) = \text{rowspan}(V)\) if and only if
  \(\text{rowspan}(U_i G_i) = \text{rowspan}(U_i), \forall i = 1, 2, \ldots, r\).
- \(\text{rowspan}(VH) \cap \text{rowspan}(V) = \{0\}\)
  \(\Rightarrow \text{rowspan}(U_i G_i) \cap \text{rowspan}(U_i) = \{0\}\)

where \(0\) is used to denote the row vector of zeroes (of the appropriate dimension).

An overview of the proof of the lemma is placed in the appendix. A complete proof can be found in the extended version of this paper [26]. The first property of the above lemma is useful to ensure alignment, i.e., ensure relations of the form (3). The second property listed in the lemma is useful to ensure relations of the form (5).

To see this, consider the special case where \(k = 3\). We fix \(L = 2^3 = 8\). Let us focus on the constraints associated with \(V_1\) alone. We need

\[
\text{rowspan}(V_1 H_2) = \text{rowspan}(V_1 H_3) = \text{rowspan}(V_1)
\]

and

\[
\text{rowspan}(V_1 H_1) \cap \text{rowspan}(V_1) = \{0\}
\]
Suppose that $V_1 = U_1 \otimes I \otimes I$ where $U_1$ is a $1 \times 2$ matrix and $I$ is a $2 \times 2$ identity matrix. Also, suppose that $H_i = H_{i,1} \otimes H_{i,2} \otimes H_{i,3}$ where $H_{i,j}$ are all $2 \times 2$ matrices. We now need $\text{rowspan}(V_1 H_2)$ to be equal to the rowspan of $V_1$. To ensure this, the lemma stated above implies that we need $\text{rowspan}(U_1) = \text{rowspan}(U_1 H_{2,1})$. Similarly, we need $\text{rowspan}(U_1) = \text{rowspan}(U_1 H_{3,1})$. To ensure that $\text{rowspan}(V_1) \cap \text{rowspan}(V_1 H_1) = \{0\}$, it is sufficient to ensure that $\text{rowspan}(U_1) \cap \text{rowspan}(U_1 H_{1,1}) = \{0\}$. Note that all these constraints can be satisfied, for example, by setting

$$U_1 = \begin{pmatrix} 0 & 1 \\ \end{pmatrix},$$

$$H_{2,1} = H_{3,1} = I,$$

$$H_{1,1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$ 

In other words, with a tensor product structure imposed on all encoding and repair matrices, all the constraints necessary for recovery of the first node $a_1$ can be satisfied by choosing the first factor of the tensor products carefully. Similarly, by careful manipulation of the second factor of the tensor products, the repair of $a_2$ can be ensured. The third factor can be used to ensure repair of $a_3$. Thus, one set of solutions for the required problem is

$$H_1 = G \otimes I \otimes I, V_1 = U \otimes I \otimes I,$$

$$H_2 = I \otimes G \otimes I, V_2 = I \otimes U \otimes I,$$

$$H_3 = I \otimes I \otimes G, V_3 = I \otimes I \otimes U,$$

where $G$ is a $2 \times 2$ matrix and $U$ is a $1 \times 2$ matrix chosen such that $\text{rowspan}(UG) \cap \text{rowspan}(U) = \{0\}$. Indeed, the solutions of [16], [18], [19] all have the above form! Further, a careful understanding of this procedure can lead to more general solutions for both the case of $n = k + 2$ and for arbitrary $(n, k)$ as well (See [26]).

### IV. Discussion

In this paper, we described a tensor product based alignment framework to achieve interference alignment using a finite number of dimensions. The framework provides a systematic approach to satisfy a set of alignment conditions, if they have a certain form. In particular, we use a tensor product of $k$ matrices as the network coding matrices. The $k_0$th factor of the tensor product suffices to ensure the necessary alignment constraints for the $k_0$th message in the application considered. Lemma 1 provides a guideline to aligning interferers using this tensor product framework. The technique developed in this paper motivates several questions.

- A question of interest is whether this framework is as powerful as asymptotic interference alignment, in terms of the extent of alignment. For instance, the framework developed here only deals with repair of systematic nodes. It is not clear whether this approach leads insights into design of codes with optimal parity repair. It is worth noting that asymptotic alignment based codes of [10] perform optimally w.r.t. parity node repair as well. On a similar note, it is not clear whether this framework can be applied to create finite-block length achievable schemes for the communication scenarios studied in [8], [9] in place of asymptotic alignment.

- The reverse of the above question is also not clear. In other words, is not clear whether, from a network capacity perspective, asymptotic alignment is strictly more powerful than the tensor product based framework. While we are not aware of a network communication scenario where the tensor product based alignment approach achieves a better rate than asymptotic alignment,
we currently cannot preclude the existence of such a scenario.

- Finally, a question worth exploring is whether the insights of this work can be used to derive non-trivial achievable schemes for a different (or larger) class of (non-multicast) network communication scenarios.

**APPENDIX**

The proof of the lemma stems from combining the following three properties of the tensor product. The first following properties is well known. The second and the third properties can be derived from the multi-linearity and associative property of the tensor product

1) **Mixed Product Property:**
   \[(P_1 \otimes P_2 \otimes \ldots \otimes P_m)(Q_1 \otimes Q_2 \otimes \ldots \otimes Q_m) = (P_1Q_1) \otimes (P_2Q_2) \otimes \ldots \otimes (P_mQ_m)\]

2) **Invariance w.r.t. span:** Let \(P_i, Q_i, i = 1, 2, \ldots, m\) be matrices such that the dimension of \(P_i\) is equal to the dimension of \(Q_i\). Also, let \(\text{rowspan}(P_i) \neq \emptyset, \text{rowspan}(Q_i) \neq \emptyset, \forall i = 1, 2, \ldots, m\) where \(\emptyset\) represents the row vector whose entries are all equal to 0. Then, \(\text{rowspan}(P_1) = \text{rowspan}(Q_1), i = 1, 2, \ldots, m\), if and only if
   \[\text{rowspan}(P_1 \otimes P_2 \otimes \ldots \otimes P_m) = \text{rowspan}(Q_1 \otimes Q_2 \otimes \ldots \otimes Q_m)\]

3) **Inheritance of linear independence:** Let \(P_i, Q_i, i = 1, 2, \ldots, m\) be matrices such that the dimension of \(P_i\) is equal to the dimension of \(Q_i\). Now, suppose that \(\text{rowspan}(P_i) \cap \text{rowspan}(Q_i) = \{0\}\) for some \(l \in \{1, 2, \ldots, m\}\), i.e., each row of \(P_i\) is linearly independent of all the rows of \(Q_i\) for some \(l \in \{1, 2, \ldots, m\}\). Then,
   \[\text{rowspan}(P_1 \otimes P_2 \otimes \ldots \otimes P_m) \cap \text{rowspan}(Q_1 \otimes Q_2 \otimes \ldots \otimes Q_m) = \{0\}\]

**REFERENCES**


