An Optimal Divide-and-Conquer Solution to the Linear Data Exchange Problem

Nebojsa Milosavljevic, Sameer Pawar, Salim El Rouayheb, Michael Gastpar[†] and Kannan Ramchandran
Department of Electrical Engineering and Computer Sciences
University of California, Berkeley
Email: {nebojsa, spawar, salim, gastpar, kannanr}@eecs.berkeley.edu

***Abstract*—In this paper we study the problem of data exchange, where each node in the system has a number of linear combinations of the data packets. Communicating over a public channel, the goal is for all nodes to reconstruct the entire set of the data packets in minimal total number of bits exchanged over the channel. We present a novel divide and conquer based architecture that determines the number of bits each node should transmit. This along with the well known fact, that it is sufficient for the nodes to broadcast linear combinations of their local information, provides a polynomial time deterministic algorithm for reconstructing the entire set of the data packets at all nodes in minimal amount of total communication.**

I. Introduction

In the cooperative data exchange problem, m nodes each have part of the data and the goal is for all nodes to learn all of the data using the minimum number of transmissions (*i.e.* bits) on a fully public, noiseless broadcast channel. We focus on a particularly simple yet attractive source model called the *finite linear source*. Consider a large file that is split into a total of N packets each belonging to some finite field, say $\mathbb{F}_q$. Each node holds a part of the information in the form of a collection of linear combinations of the data packets.

To motivate this particular problem, assume that m users wish to download a large file. At the base station the file is divided into N chunks. The base station transmits coded packets (linear combinations of the packets) to all users over an unreliable wireless channel. As a result, after the base station stopped transmitting, each node received only a subset of the coded packets. Co-located mobile users have a broadcast channel among themselves which can be better than individual channels to the base station. In this scenario assuming that among themselves they have the complete file, question is how can these terminals cooperate to learn the whole file. An example of this scenario for $m = 3$ nodes and $N = 4$ packets is presented in Figure 1.

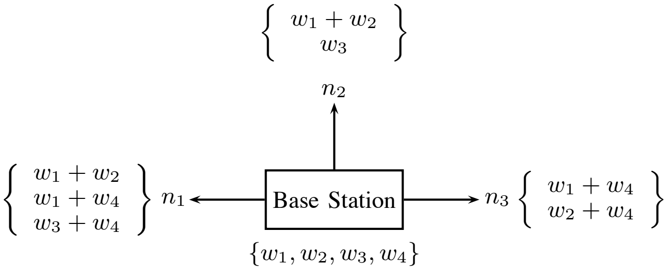
$\left\{ \begin{array}{c} w_1 + w_2 \\ w_3 \end{array} \right\}$
n_2
$\left\{ \begin{array}{c} w_1 + w_2 \\ w_1 + w_4 \\ w_3 + w_4 \end{array} \right\}$ n_1
Base Station
n_3 $\left\{ \begin{array}{c} w_1 + w_4 \\ w_2 + w_4 \end{array} \right\}$
$\{w_1, w_2, w_3, w_4\}$

Fig. 1. Finite linear source model: An example where the base station transmits coded packets of $N = 4$ underlying data packets w_1, w_2, w_3 and w_4 to $m = 3$ users. Each of them receives some subset of the transmitted coded packets.

We propose a deterministic and computationally efficient (polynomial in the number of system nodes and in the number of data packets) algorithm that solves this problem by specifying the number of bits and actual transmission for each node in order successfully accomplish data exchange using the minimum total number of bits exchanged.

The data exchange problem described above invites the following questions of interest. First, what is the order in which the nodes should transmit in an optimal scheme (with the possibility of the need for interaction where nodes take several turns). Second, how many bits should each node transmit during its scheduled broadcast transmission. Third, is there a polynomial time algorithm to determine an optimal schedule and transmission such that all the nodes can reconstruct the file.

A complete solution to our problem necessarily involves answering all the above questions. The first question of determining optimal schedule has been addressed in the broader context of the seminal work of Csiszár and Narayan [1] on determining the secrecy capacity for a multi-terminal problem. In [1], authors showed that interaction is not needed and optimal solution can be achieved by a non-interactive one-shot communication by each node.

Second, on the question of how many bits each node should send, again Csiszár and Narayan [1] partially resolved this issue by providing minimum amount of total communication needed to accomplish the data exchange but rates for individual nodes in optimal scheme is still an open question. They formulated the problem as an optimization problem with an exponential number of rate-constraints corresponding to all possible cut-sets that need to be satisfied. Csiszár and Narayan [1] also provided an alternative characterization of a lower bound to the omniscience sum-rate by relating it to a more intuitively appealing notion of so-called mutual-

[†] Also with the School of Computer and Communication Sciences, EPFL, Lausanne, Switzerland.

This research was funded by the NSF grants (CCF-0964018, CCF-0830788), a DTRA grant (HDTRA1-09-1-0032), and in part by an AFOSR grant (FA9550-09-1-0120).

dependence. This bound was subsequently shown to be tight in the work of [2].

This paper contributes to the literature by addressing part of the second question and the third one posed above, describing a deterministic polynomial-time algorithm to achieve optimal data exchange. Our algorithm uses as a key building block some recent interesting results from the theoretical Computer Science literature. Specifically, the total communication rate needed for the data exchange can be solved by relating our problem to the minimum average cost clustering problem which can be solved in polynomial time [3]. The key challenge is in decomposing the (well-understood) omniscience sum-rate into the optimal per-node rate allocations in polynomial time, while also specifying the content of these per-node broadcasted information in order to achieve minimum-rate data exchange. Using a simple but powerful divide-and-conquer strategy built on this, together with some important observations on the necessary conditions for optimality (to be described in the sequel), and integrating this with well-known results from the network coding community [4], our algorithm provides a polynomial-time optimal solution to the data exchange problem.

Related Work

The cooperative data exchange problem was introduced in [5] by El Rouayheb *et al.* for the case when each node observes some subset of the data packets. In subsequent work, Sprinston *et al.* [6] proposed a randomized algorithm that achieves (with high probability) the minimum number of transmissions over the public channel, provided that the field size is large enough. Courtade *et al.* [7] proposed an LP formulation of this problem. In most general setting the LP formulation has exponential number of constraints, but in [7] authors considered a special case where each node observes simply a subset of the underlying packets. For this special case authors showed that the proposed LP can be solved in polynomial time.

II. System Model and Preliminaries

Consider m nodes $n_1, n_2, \ldots n_m$ which observe discrete memoryless multiple sources (DMMS) as in [1]. Let $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_m, m \geq 2,$, denote these sources. In [1] authors studied the case where the correlations among these sources can be specified by an arbitrary joint distribution. Here we focus on a simple yet interesting correlation model called *finite linear source* introduced in [2].

Next, we briefly describe the finite linear source model. Let $q = p^m$, where p is a prime number and $m \in \mathbb{Z}_+$. Consider the N-dimensional q-ary random vector $\mathbf{W} \in \mathbb{F}_q^N$ whose components are independent and uniformly distributed over the elements of $\mathbb{F}_q$. Then, in the linear source model, the i^{th} source component is simply given by

$$\mathbf{X}_i = \mathbf{A}_i \mathbf{W},\ i \in \mathcal{M}, \tag{1}$$

where $\mathbf{A}_i \in \mathbb{F}_q^{\ell_i \times N}$ is a fixed matrix that specifies the correlation between the sources, and we use $\mathcal{M}$ to denote the set $\{n_1, n_2, \ldots, n_m\}$.

It is easy to verify that for the finite linear multiple source,

$$H(\mathbf{X}_i) = \text{rank}\{\mathbf{A}_i\}, \tag{2}$$

where the entropy is computed in base q. Without loss of generality we assume that the rank of $\mathbf{A} \triangleq \left[\begin{array}{ccc} \mathbf{A}_1 & \ldots & \mathbf{A}_m \end{array} \right]^T$ is equal to N (if not, then let $N' = \text{rank}\{\mathbf{A}\}$, and one can obtain a uniformly random vector $\mathbf{W}' \in \mathbb{F}_q^{N'}$ from $\mathbf{W}$ using a linear transformation and then relabel $N = N', \mathbf{W} = \mathbf{W}'$). Since we focus on a finite linear multiple source, we will use the entropy of the observations and the rank of the observation matrix interchangeably.

The terminals are allowed to communicate over a noiseless public broadcast channel in multiple rounds and thus, may use interactive communication, meaning that they can incorporate what they have received so far over the public broadcast channel into their future transmissions. The first result that is of fundamental importance to our work is that such *interactive communication is not necessary,* which was established by Csiszár and Narayan [1] for the general DMMS, and hence for the simple finite linear source considered here. Moreover, from [2], we know that for the finite linear source, it is sufficient to transmit *linear combinations* of the source information. Therefore, without loss of generality, we can describe transmissions $\mathbf{F}_i$ of the node $n_i \in \mathcal{M}$, as a collection of R_i linear equations generated from $\mathbf{X}_i$. The value of R_i and actual transmissions $\mathbf{F}_i$ which result in an optimal linear scheme will be determined in the subsequent sections. We will use $\mathbf{F}$ to denote the total communication over the public channel, *i.e.* $\mathbf{F} = (\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_m)$. Thus, our goal is to devise a polynomial time algorithm that solves the omniscience problem using the minimal amount of communication. More formally, we are interested in a efficient scheme that determines $\mathbf{F}$ such that $H(\mathbf{F})$ is minimal and $H(\mathbf{X}_\mathcal{M} | \mathbf{F}, \mathbf{X}_i) = 0\ \forall n_i \in \mathcal{M}$, where $\mathbf{X}_\mathcal{M} = (\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_m)$. The minimum number of transmissions needed to attain our goal (*i.e.*, the minimum value of $H(\mathbf{F})$) will be referred to as the *communication for omniscience rate* and denoted by $R_{CO}(\mathcal{M})$ for the set $\mathcal{M}$. Denoting $R_i = H(\mathbf{F}_i)$ the transmission rate of node n_i, it is straightforward to show that for any optimal scheme $H(\mathbf{F}) = \sum_{i=1}^m R_i$. The latest statement is obvious since any correlation among transmissions is considered to be redundant and does not lead to the minimal amount of communication needed for the data exchange.

We stress that the value of $R_{CO}(\mathcal{M})$ in terms of optimization problem was given in [1]. An alternative version of this formula, which turns out to be very useful for this work, is given by the following theorem in [2]:

Theorem 1 (Chan)**.** *Provided that the field size $|\mathbb{F}_q|$ is large enough, the communication for omniscience rate of the set $\mathcal{M}$ is given by*

$$R_{CO}(\mathcal{M}) = H(\mathbf{X}_\mathcal{M}) - \min_{\mathcal{P}} \frac{\sum_{\mathcal{S} \in \mathcal{P}} H(\mathbf{X}_\mathcal{S}) - H(\mathbf{X}_\mathcal{M})}{|\mathcal{P}| - 1}, \tag{3}$$

where $\mathcal{P} = \{\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_{|\mathcal{P}|}\}$ denotes a partition of the node set indexed by $\mathcal{M}$ into $|\mathcal{P}|$ disjoint sets where $2 \leq |\mathcal{P}| \leq m$.

However, Theorem 1 does not say what is the individual rate allocation for each node, and hence, it does not immediately lead to the optimal transmission scheme. It can be shown with some effort that the minimizer in Equation (3), and thus, the partition attaining $R_{CO}(\mathcal{M})$, is not unique. To this end, we define the following:

Definition 1. Define the set of optimum partitions in Equation (3) as $\overline{\mathcal{Q}}_{\mathcal{M}}$.

Remark 1. Note that in general $R_{CO}(\mathcal{M})$ can be a rational number. Now consider some optimal partition $\mathcal{P}_{\mathcal{M}}$ and split the packets into $|\mathcal{P}_{\mathcal{M}}|$ chunks (since we allow packet splits). Then, the new $R_{CO}(\mathcal{M})$ rate becomes equal to the old $R_{CO}(\mathcal{M})$ rate multiplied by $|\mathcal{P}_{\mathcal{M}}|$, which is an integer number. Therefore, without loss of generality in this paper we assume that the optimal communication for omniscience rate is always an integer number.

In [3] it is shown that the minimization problem (3) can be computed in $\mathcal{O}(m^2 \cdot \text{SFM}(m))$ time along with one optimal partitioning $\mathcal{P}$ of the set $\mathcal{M}$. SFM(m) denotes the time required to minimize a submodular function defined on the power set $2^{\mathcal{M}}$. For arbitrary submodular functions $\text{SFM}(m) = \mathcal{O}(m^5 \cdot \gamma + m^6)$ (see [8]), where γ is the complexity of computing the submodular function.

III. An Example

In this section we provide an example that sheds some light on the necessary condition for any scheme to be optimal in terms of sum rate. In [1] it was shown that a non-interactive scheme is sufficient for achieving omniscience in the optimal number of communication rounds. However, in order to gain more understanding about this problem, we use the interactive greedy scheme introduced in [5]. The idea behind this algorithm is very simple; in each round of communication a node with the highest observation rank transmits some linear combination of its current observations. But the main problem with this scheme is determining the actual transmissions; it is tempting to assume that it is enough for the transmitting node to increase the observation rank of all receiving nodes. In the following example we show that this approach is sometimes suboptimal. Hence, the question still remains: How to design optimal transmissions. To answer this, we break our problem into two parts: first, we devise an algorithm that determines optimal communication rates of each node, and then, to determine actual transmissions of each node we convert our problem to a matrix completion problem [4], [9].

Example 1. Consider a system with $m = 4$ terminals $\mathcal{M} = \{n_1, n_2, n_3, n_4\}$. For convenience, we express the underlying vector as $\mathbf{W} = \begin{bmatrix} a & b & c & d & e & f & g & h \end{bmatrix}^T \in \mathbb{F}_q^8$, where a, b, c, d, e, f, g, h are independent uniform random variables in $\mathbb{F}_q$. Let us consider the case where each node has the following observations:

$$\mathbf{X}_1 = \begin{bmatrix} a & b & c & e \end{bmatrix}^T, \qquad \mathbf{X}_2 = \begin{bmatrix} a & b & d & e \end{bmatrix}^T,$$
$$\mathbf{X}_3 = \begin{bmatrix} c & d & f & g & h \end{bmatrix}^T, \quad \mathbf{X}_4 = \begin{bmatrix} c & f \end{bmatrix}^T.$$

The task is to construct a deterministic communication scheme for which all the nodes achieve omniscience in the minimum number of transmissions over the public channel. From Theorem 1, we find that $R_{CO}(\mathcal{M}) = 6$. Two different partitions both lead to this optimal rate, are $\mathcal{P}^{(1)} = \{\{n_1, n_2\}, \{n_3, n_4\}\}$, and $\mathcal{P}^{(2)} = \{\{n_1, n_2\}, \{n_3\}, \{n_4\}\}$. Without loss of generality, let us assume that each transmitting node sends one symbol in $\mathbb{F}_q$ at a time. In [5] the authors proposed a communication scheme which picks the node with the largest observation rank (or one of them, if non-unique), in this case node n_3, and lets it transmit one symbol in the form of a linear combination of its observations, where the coefficients are chosen so as to increase the resulting rank of each of the receiving nodes. In our example, this means that we could let node n_3 transmit $c + d$. After this transmission, each node increases its observation rank, and updates its observations as follows:

$$\mathbf{X}_1 = \begin{bmatrix} a & b & c & d & e \end{bmatrix}^T, \quad \mathbf{X}_2 = \begin{bmatrix} a & b & c & d & e \end{bmatrix}^T,$$
$$\mathbf{X}_3 = \begin{bmatrix} c & d & f & g & h \end{bmatrix}^T, \quad \mathbf{X}_4 = \begin{bmatrix} c & f & d \end{bmatrix}^T.$$

However, as we will now explain, this transmission *cannot* be part of an optimal communication scheme. Namely, if we recompute $R_{CO}(\mathcal{M})$ after the first transmission, we again get $R_{CO}(\mathcal{M}) = 6$, meaning that nothing has been gained by the above transmission: we still need 6 transmissions to achieve omniscience. Thus, the greedy approach of increasing rank of all nodes in every transmission is not optimal.

In Lemma 2, we show that the necessary condition for any data exchange scheme to be optimal is to increase the observation rank of all receiving sets in *all* optimal partitions of set $\mathcal{M}$. This condition is stronger than merely increasing the observation rank of every node. In our current example, there are two optimal partitions, and in both of them nodes n_1 and n_2 are clubbed together. This means that the transmission from node n_3 should be selected such as to increase the observation rank of the "supernode" resulting from combining nodes n_1 and n_2, *i.e.* it should increase the observation rank of $\mathbf{X}_{\{1,2\}} = \begin{bmatrix} a & b & c & d & e \end{bmatrix}^T$. Clearly, transmitting $c+d$ as considered above does *not* increase the rank of this collection.

As we can see from the example above, constructing an optimal transmission of any given node requires knowledge of all optimal partitions of the set $\mathcal{M}$. This problem is hard in general. One way to get around this problem is to first solve for an optimal rate allocation of all nodes, and then determine the actual transmissions.

Let us start with the rate allocation. A crucial observation to determine these individual rates is that for any optimal partition $\mathcal{P}_{\mathcal{M}} \in \overline{\mathcal{Q}}_{\mathcal{M}}$ the transmission rates of each set that belongs to $\mathcal{P}_{\mathcal{M}}$ are fundamental. For example, according to Theorem 1 there are 2 optimal partitions of the set $\mathcal{M}$: $\mathcal{P}^{(1)} = \{\{n_1, n_2\}, \{n_3, n_4\}\}$, and $\mathcal{P}^{(2)} = \{\{n_1, n_2\}, \{n_3\}, \{n_4\}\}$. Now, irrespective of the scheme the sum rates of each set in all the optimal partitions are fundamental, *i.e.*, for this example, it can be shown that partition $\mathcal{P}^{(1)}$ imposes $R_1 + R_2 = 3$ and $R_3 + R_4 = 3$ while partition $\mathcal{P}^{(2)}$ imposes $R_1 + R_2 = 3$, $R_3 = 3$ and $R_4 = 0$. Although these fundamental constraints allow us to conclude the values for R_3 and R_4, it is not clear what is the split between R_1 and R_2. Here we employ what we

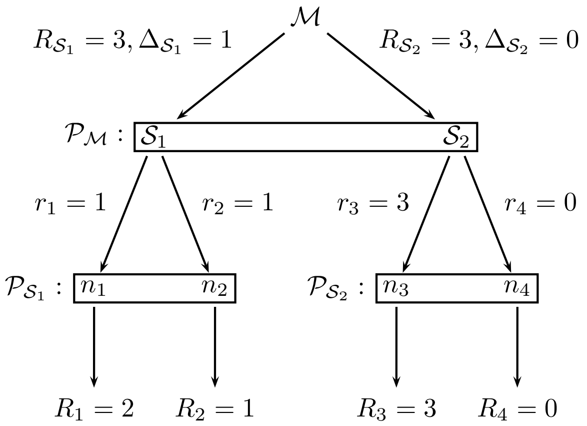
$\mathcal{M}$
$R_{\mathcal{S}_1}=3, \Delta_{\mathcal{S}_1}=1$
$R_{\mathcal{S}_2}=3, \Delta_{\mathcal{S}_2}=0$
$\mathcal{P}_{\mathcal{M}}$: $\mathcal{S}_1$ $\mathcal{S}_2$
$r_1=1$
$r_2=1$
$r_3=3$
$r_4=0$
$\mathcal{P}_{\mathcal{S}_1}$: n_1 n_2
$\mathcal{P}_{\mathcal{S}_2}$: n_3 n_4
$R_1=2$
$R_2=1$
$R_3=3$
$R_4=0$

Fig. 2. Algorithm flow graph. We start with the set $\mathcal{M}$ and compute an optimal partition $\mathcal{P}_{\mathcal{M}} = (\mathcal{S}_1, \mathcal{S}_2)$ and corresponding communication rates $R_{\mathcal{S}_1} = 3$ and $R_{\mathcal{S}_2} = 3$. For both $\mathcal{S}_1$ and $\mathcal{S}_2$, we compute optimal partitions $\mathcal{P}_{\mathcal{S}_1} = \{\{n_1\}, \{n_2\}\}$ and $\mathcal{P}_{\mathcal{S}_2} = \{\{n_3\}, \{n_4\}\}$, corresponding rates $r_1 = 1$, $r_2 = 1$, $r_3 = 3$, $r_4 = 0$ which achieve local omniscience within $\mathcal{S}_1$ and $\mathcal{S}_2$, and excess rates $\Delta_{\mathcal{S}_1} = 1$, $\Delta_{\mathcal{S}_2} = 0$. Then, we obtain $R_1 = r_1 + \Delta_{\mathcal{S}_1} = 2$, $R_2 = 2$, $R_3 = r_3 + \Delta_{\mathcal{S}_2} = 3$ and $R_4 = r_4 = 0$.

call divide and conquer approach. For this approach we only need one optimal partition of the set $\mathcal{M}$, say $\mathcal{P}^{(1)} = \{\mathcal{S}_1, \mathcal{S}_2\}$, where $\mathcal{S}_1 = \{n_1, n_2\}$ and $\mathcal{S}_2 = \{n_3, n_4\}$. The idea is that we know that the total budget of the set $\mathcal{S}_1$ is 3, and we show that achieving a local omniscience of information in the set $\mathcal{S}_1$ is part of some globally optimal scheme. For this example it is clear that both n_1 and n_2 need to send at least 1 symbol to achieve local omniscience, since $H(\mathbf{X}_1, \mathbf{X}_2|\mathbf{X}_1) = H(\mathbf{X}_1, \mathbf{X}_2|\mathbf{X}_2) = 1$. However we know that the total transmission budget of the set $\mathcal{S}_1$ is 3. Now after local omniscience, since both nodes n_1 and n_2 are equivalent in terms of information, the remaining difference $\Delta_{\mathcal{S}_1} = 1$ symbol can be transmitted by either of them. Thus, we have $R_1 = 2$, $R_2 = 1$. Using the same analysis, one can show that $R_3 = 3$ and $R_4 = 0$. Thus in more general case, we first determine the budget of each set in an optimal partition and then focus on each individual set as a new problem of achieving optimal local omniscience and so on. Divide and conquer procedure is schematically presented in Figure 2. Now, once we determine the transmission rate of each node, we focus on finding the actual transmissions. Let us observe node n_1. We know that it is receiving 1 symbol from node n_2 and 3 symbols from node n_3. In the most general setting node n_2 transmits a linear combination of all of its observations: $f_2 = x_a \cdot a + x_b \cdot b + x_d \cdot d + x_e \cdot e$, which can be represented in a vector form as $\mathbf{F}_2 = [\begin{array}{cccccccc} x_a & x_b & 0 & x_d & x_e & 0 & 0 & 0 \end{array}]$. Similarly, node n_3 transmits

$$\mathbf{F}_3 = \begin{bmatrix} 0 & 0 & y_c & y_d & 0 & y_f & y_g & y_h \\ 0 & 0 & z_c & z_d & 0 & z_f & z_g & z_h \\ 0 & 0 & u_c & u_d & 0 & u_f & u_g & u_h \end{bmatrix}. \tag{4}$$

All entries of the matrices $\mathbf{F}_2$ and $\mathbf{F}_3$ belong to the finite field $\mathbb{F}_q$. After receiving transmissions from nodes n_2 and n_3, node n_1 has the following observation matrix

$$\mathbf{A}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ x_a & x_b & 0 & x_d & x_e & 0 & 0 & 0 \\ 0 & 0 & y_c & y_d & 0 & y_f & y_g & y_h \\ 0 & 0 & z_c & z_d & 0 & z_f & z_g & z_h \\ 0 & 0 & u_c & u_d & 0 & u_f & u_g & u_h \end{bmatrix}. \tag{5}$$

Since node n_1 wishes to reconstruct all the packets, we need to choose unassigned entries of the matrix $\mathbf{A}_1$ such that it is a full rank. In the literature this problem is known as a matrix completion problem [4]. Moreover, here for each node n_i, $i \in \mathcal{M}$ we can construct its observation matrix $\mathbf{A}_i$ based on its initial observations, and the information it receives from the other nodes. Now, the idea is to pick unknown entries in all these matrices such that the full rank property is simultaneously satisfied. This problem is known as a simultaneous matrix completion problem. In [4] polynomial time algorithm was devised for the case when all unknown entries of the matrices $\mathbf{A}_i$ can be chosen independently. In our problem this approach works when the initial observations of all nodes are some subsets of $\mathbf{W}$. Since we assume more general model, where the initial observations are some linear combinations of $\mathbf{W}$, the unknown entries of the matrix $\mathbf{A}_i$, $i \in \mathcal{M}$ which belong to the same row cannot be chosen independently anymore. Polynomial time algorithm that solves for this version of matrix completion problem is provided in [9]. This procedure also suggests that a field size greater than or equal to the number of nodes guarantees a deterministic solution to this problem. Using the above matrix completion approach for this example one of the solutions we get is that node n_1 sends $c + e$, b, node n_2 transmits $a + d$ and node n_3 transmits $d + f$, g and h.

IV. Deterministic Algorithm

As illustrated by the example in the previous section, our algorithm has two parts: first, it finds an optimal communication rate allocation R_i, $n_i \in \mathcal{M}$, and then it finds transmissions by solving a simultaneous matrix completion problem [4], [9]. Due to space constraints, we focus on the optimal rate allocation, denoting the rate to be used by the i^{th} terminal by R_i. One of the results we extensively use concerns the minimum average cost clustering problem. In particular, it is shown in [3] that the minimization in (3) in Theorem 1 along with finding one optimal partition $\mathcal{P}_{\mathcal{M}} \in \overline{\mathcal{Q}}_{\mathcal{M}}$ can be computed in polynomial time. For future reference, we will refer to this operation as the function 1-MAC($\mathcal{M}$) whose output is a pair $(R_{CO}(\mathcal{M}), \mathcal{P}_{\mathcal{M}})$. Now, we provide a formal description of the algorithm we already explained in the previous section using a simple example.

Theorem 2. *If $|\mathbb{F}_q| \geq m$, then Min Sum Rate Algorithm generates an achievable rate allocation for the data exchange problem which is optimal according to Theorem 1. The complexity of this algorithm is $\mathcal{O}(m^3 \cdot SFM(m))$.*

To show optimality of the proposed scheme, we first identify the optimal number of symbols in $\mathbb{F}_q$ sent from each set that belongs to some optimal partitioning $\mathcal{P}_{\mathcal{M}} \in \overline{\mathcal{Q}}_{\mathcal{M}}$.

Lemma 1. *For every optimal partition $\mathcal{P}_{\mathcal{M}} \in \overline{\mathcal{Q}}_{\mathcal{M}}$, an optimal communication scheme (*i.e.*, attaining $R_{CO}(\mathcal{M})$) must satisfy, for every set $\mathcal{S} \in \mathcal{P}_{\mathcal{M}}$,*

$$R_{\mathcal{S}} = R_{CO}(\mathcal{M}) - H(\mathbf{X}_{\mathcal{M}}|\mathbf{X}_{\mathcal{S}}), \tag{6}$$

where $R_{\mathcal{S}} = \sum_{n_i \in \mathcal{S}} R_i$.

An important observation can now be made: The optimum partition provided by the 1-MAC function may contain a few

Algorithm 1 Min Sum Rate

```
1: Initialize $\mathcal{V} = \mathcal{M}$, $\Delta_{\mathcal{V}} = 0$
2: Compute 1-MAC($\mathcal{V}$) $= (R_{CO}(\mathcal{V}), \mathcal{P}_{\mathcal{V}})$
3: If $\mathcal{V} \neq \mathcal{M}$ then $\Delta_{\mathcal{V}} = R_{\mathcal{V}} - R_{CO}(\mathcal{V})$.
4: for $\forall \mathcal{S} \in \mathcal{P}_{\mathcal{V}}$ do
5:     Compute local omniscience rates $r_{\mathcal{S}}$ of set $\mathcal{V}$ as follows
           $r_{\mathcal{S}} = R_{CO}(\mathcal{V}) - H(\mathbf{X}_{\mathcal{V}}|\mathbf{X}_{\mathcal{S}})$
6: end for
7: Pick a set, say $\mathcal{U}$ from $\mathcal{P}_{\mathcal{V}}$ and compute communication rates for
   all sets $\mathcal{S} \in \mathcal{P}_{\mathcal{V}}$ as follows:
           $R_{\mathcal{S}} = \begin{cases} r_{\mathcal{S}} + \Delta_{\mathcal{V}} & \text{if } \mathcal{S} = \mathcal{U} \\ r_{\mathcal{S}} & \text{otherwise} \end{cases}$
8: for $\forall \mathcal{S} \in \mathcal{P}_{\mathcal{V}}$ do
9:     If $\mathcal{S}$ is not a singleton set then set $\mathcal{V} = \mathcal{S}$ and go to step 1.
10: end for
```

singletons. For those, Lemma 1 directly provides the correct communication rate R_i. For the rest, we proceed with a *divide and conquer* strategy: We now run the 1-MAC algorithm on each set $\mathcal{S} \in \mathcal{P}_{\mathcal{M}}$, resulting in an optimal pair $(R_{CO}(\mathcal{S}), \mathcal{P}_{\mathcal{S}})$ for all $\mathcal{S} \in \mathcal{P}_{\mathcal{M}}$. This process is repeated recursively until all optimal partitions contain singleton sets only.

Lemma 2. *For any optimal transmission scheme (i.e., attaining $R_{CO}(\mathcal{M})$), the transmission $\mathbf{F}_i$ of node $n_i \in \mathcal{M}$ increases the entropy of all receiving sets in all optimal partitions $\mathcal{P}_{\mathcal{M}} \in \overline{\mathcal{Q}}_{\mathcal{M}}$ by the amount of $H(\mathbf{F}_i)$ symbols. In other words, for all $\mathcal{P}_{\mathcal{M}} \in \overline{\mathcal{Q}}_{\mathcal{M}}$, we have*

$$H(\mathbf{X}_{\mathcal{S}}, \mathbf{F}_i) = H(\mathbf{X}_{\mathcal{S}}) + H(\mathbf{F}_i), \forall \mathcal{S} \in \mathcal{P}_{\mathcal{M}}, \text{ s.t. } i \notin \mathcal{S} \quad (7)$$

For the sake of the next argument, let us fix an arbitrary optimal partition $\mathcal{P}_{\mathcal{M}} \in \overline{\mathcal{Q}}_{\mathcal{M}}$. Since non-interactive transmission scheme is optimal, we can assume that set $\mathcal{S} \in \mathcal{P}_{\mathcal{M}}$ has received all transmissions from all other sets in $\mathcal{P}_{\mathcal{M}}$, but has itself not yet transmitted anything. From Lemma 1, we know that the nodes in our considered set $\mathcal{S}$ must transmit at rate $R_{\mathcal{S}}$ in order for all nodes in the network to become omniscient. As a subproblem we now ask the following question: after receiving transmissions from the nodes in $\mathcal{P}_{\mathcal{M}} \setminus \mathcal{S}$, how many symbols do nodes in $\mathcal{S}$ have to exchange in order to achieve omniscience, *i.e.*, in order for each node in $\mathcal{S}$ to learn the full data $\mathbf{X}_{\mathcal{M}}$. Perhaps initially somewhat surprisingly, it can be shown that the number of symbols required is exactly equal to the number of symbols that would be required to attain local omniscience within $\mathcal{S}$, *i.e.*, in order for each node in $\mathcal{S}$ to learn the data $\mathbf{X}_{\mathcal{S}}$, in the absence of any side information from any of the other nodes.

Theorem 3. *There exists an optimal communication scheme, where nodes within each $\mathcal{S} \in \mathcal{P}_{\mathcal{M}}$ can achieve local omniscience in $R_{CO}(\mathcal{S})$ transmissions.*

This result is crucial for our algorithm because now, each set $\mathcal{S} \in \mathcal{P}_{\mathcal{M}}$ can locally determine the transmission rates of each node in $\mathcal{S}$ by computing $\mathcal{P}_{\mathcal{S}}$. As shown in Lemma 1, in an optimal communication scheme, the nodes in set $\mathcal{S}$ must transmit at a rate $R_{\mathcal{S}}$ which is generally *larger* than the rate $R_{CO}(\mathcal{S})$ needed to attain omniscience within the set $\mathcal{S}$. As the next lemma shows, the extra transmissions $\Delta_{\mathcal{S}} = R_{\mathcal{S}} - R_{CO}(\mathcal{S})$ can be executed by an arbitrarily chosen node inside $\mathcal{S}$ (since there is omniscience within $\mathcal{S}$):

Lemma 3. *Let $(r_j : n_j \in \mathcal{S})$, where $\mathcal{S} \in \mathcal{P}_{\mathcal{M}}$, be a rate tuple that achieves local omniscience of the set $\mathcal{S}$. An optimal rate assignment for the nodes in $\mathcal{S}$ can be done as follows:*

$$\begin{aligned} R_l &= r_l + \Delta_{\mathcal{S}}, \quad n_l \in \mathcal{S} \\ R_j &= r_j, \quad n_j \in \mathcal{S} \setminus \{n_l\}, \end{aligned}$$

where n_l is an arbitrarily chosen node in $\mathcal{S}$.

In the next stage of the algorithm we repeat the same procedure for all the sets in $\mathcal{P}_{\mathcal{S}}, \mathcal{S} \in \mathcal{P}_{\mathcal{M}}$. Note that according to Lemma 3, it is sufficient that only one set in $\mathcal{P}_{\mathcal{S}}$ for every $\mathcal{S} \in \mathcal{P}_{\mathcal{M}}$ transmits the entire $\Delta_{\mathcal{S}}$. These "extra transmissions" are always passed to the next stage of the algorithm until the point where optimal partitions consist of singleton sets only; then they are assigned to one of the nodes accordingly.

V. Discussion and Conclusion

Finding a rate allocation that achieves data exchange in the minimum total amount of communication can be formulated as an LP but with an exponential number of constraints [1]. Hence it is not clear if it can be solved in polynomial time. In this paper we propose a novel divide and conquer approach that splits the problem into multiple sub problems and gives the optimal rate allocation for each node in polynomial time. In the full version of this paper (see http://www.eecs.berkeley.edu/%7Enebojsa), we provide an alternate formulation that is based on an optimization over a submodular polyhedron. Using a Dilworth's truncation of intersecting submodular functions and modified Edmond's algorithm [10], it is possible to solve for rate allocation in polynomial time for any arbitrary memoryless source distribution. This approach turns out to be more efficient than the one we described in this paper, but it lacks the insights.

References

[1] I. Csiszár and P. Narayan, "Secrecy capacities for multiple terminals," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3047–3061, 2004.

[2] C. Chan, "Generating Secret in a Network," Ph.D. dissertation, Massachusetts Institute of Technology, 2010.

[3] K. Nagano, Y. Kawahara, and S. Iwata, "Minimum Average Cost Clustering," *Advances in Neural Information Processing Systems*, vol. 23.

[4] N. Harvey, D. Karger, and K. Murota, "Deterministic network coding by matrix completion," in *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, 2005, pp. 489–498.

[5] S. El Rouayheb, A. Sprintson, and P. Sadeghi, "On coding for cooperative data exchange," in *Proceedings of ITW*, 2010.

[6] A. Sprintson, P. Sadeghi, G. Booker, and S. El Rouayheb, "A randomized algorithm and performance bounds for coded cooperative data exchange," in *Proceedings of ISIT*, 2010, pp. 1888–1892.

[7] T. Courtade, B. Xie, and R. Wesel, "Optimal Exchange of Packets for Universal Recovery in Broadcast Networks," in *Proceedings of Military Communications Conference*, 2010.

[8] M. Goemans and V. Ramakrishnan, "Minimizing submodular functions over families of sets," *Combinatorica*, vol. 15, no. 4, pp. 499–513, 1995.

[9] G. Ivanyos, M. Karpinski, and N. Saxena, "Deterministic polynomial time algorithms for matrix completion problems," *SIAM journal on computing*, vol. 39, pp. 3736–3751, 2010.

[10] S. Fujishige, *Submodular functions and optimization*. Elsevier Science, 2005.