The Finite Field Multi-Way Relay Channel with Correlated Sources: Beyond Three Users

Lawrence Ong[†], Roy Timo[‡], Sarah J. Johnson[†]

[†]School of Electrical Engineering and Computer Science, The University of Newcastle, Australia
[‡]Institute for Telecommunications Research, University of South Australia, Australia
Email: lawrence.ong@cantab.net, roy.timo@unisa.edu.au, sarah.johnson@newcastle.edu.au

Abstract—The multi-way relay channel (MWRC) models cooperative communication networks in which many users exchange messages via a relay. In this paper, we consider the finite field MWRC with correlated messages. The problem is to find all achievable rates, defined as the number of channel uses required per reliable exchange of message tuple. For the case of three users, we have previously established that for a special class of source distributions, the set of all achievable rates can be found [Ong et al., ISIT 2010]. The class is specified by an almost balanced conditional mutual information (ABCMI) condition. In this paper, we first generalize the ABCMI condition to the case of more than three users. We then show that if the sources satisfy the ABCMI condition, then the set of all achievable rates is found and can be attained using a separate source-channel coding architecture.

I. INTRODUCTION

This paper investigates multi-way relay channels (MWRCs) where multiple users exchange correlated data via a relay. More specifically, each user is to send its data to all other users. There is no direct link among the users, and hence the users first transmit to a relay, which processes its received information and transmits back to the users (refer to Fig. 1). The purpose of this paper is to find the set of all achievable rates, which are defined as the number of channel uses required to reliably (in the usual Shannon sense) exchange each message tuple.

The joint source-channel coding problem in Fig. 1 includes, as special cases, the source coding work of Wyner et al. [1] and the channel capacity work of Ong et al. [2]. Separately, the Shannon limits of source coding [1] (through noiseless channel) and of channel coding [2] (with independent sources) are well established. However, these limits have not yet been discovered for noisy channels with correlated sources in general. For three users, Ong et al. [3] gave sufficient conditions for reliable communication using the separate source-channel coding paradigm. The key result of [3] was to show that these sufficient conditions are also necessary for a special class of source distributions, hence giving the set of all achievable rates. The class was characterized by an almost balanced conditional mutual information (ABCMI) condition. This paper extends the ABCMI concept to more than three users, and shows that if the sources have ABCMI, then the set of all achievable rates can be found. While the ABCMI condition for the three-user case is expressed in terms of the standard Shannon information measure, we will use *I*-Measure [4, Ch. 3] for more then three users—using Shannon's measure is possible but the expressions
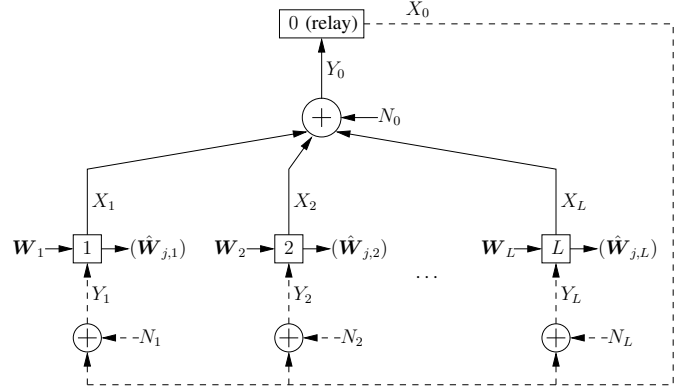

Fig. 1. The finite field MWRC in which L users (nodes $1, 2, \ldots, L$) exchange correlated messages through a relay (node 0). The uplink channel is marked with solid lines and the downlink channel with dotted lines.

would be much more complicated.

Though the ABCMI condition limits the class of sources for which the set of all achievable rates is found, this paper provides the following insights: (i) As achievability is derived based on a separate source-channel coding architecture, we show that source-channel separation is optimal for finite field MWRC with sources having ABCMI. (ii) Since ABCMI are constraints on the sources, the results in this paper are potentially useful for other channel models (not restricted to the finite field model). (iii) This paper highlights the usefulness of the *I*-Measure as a complement to the standard Shannon information measure.

II. MAIN RESULTS

A. The Network Model

1) Sources: Consider m independent and identically distributed (i.i.d.) drawings of a tuple of correlated discrete finite random variables $(W_1, W_2, \ldots, W_L)$, i.e., $\{(W_1[t], W_2[t], \ldots, W_L[t])\}_{t=1}^m$. The message of user i is given by $\boldsymbol{W}_i = (W_i[1], W_i[2], \ldots, W_i[m])$.

2) Channel: Each channel use of the finite field MWRC consists of an uplink and L downlinks, characterized by

$$\text{Uplink:} \quad Y_0 = X_1 \oplus X_2 \oplus \cdots \oplus X_L \oplus N_L \qquad (1)$$

$$\text{Downlinks:} \quad Y_i = X_0 \oplus N_i, \quad \text{for all } i \in \{1, 2, \ldots, L\}, \quad (2)$$

where X_ℓ, Y_ℓ, N_ℓ, for all $\ell \in \{0, 1, \ldots, L\}$, each take values in a finite field $\mathcal{F}$, $\oplus$ denotes addition over $\mathcal{F}$, X_ℓ is the channel

input from node ℓ, Y_ℓ is the channel output received by node ℓ, and N_ℓ is the receiver noise at node ℓ. The noise N_ℓ is arbitrarily distributed, but is i.i.d. for each channel use. We have used the subscript i to denote a user and the subscript ℓ to denote a node (which can be a user or the relay).

3) Block Codes (joint source-channel codes with feedback): Consider block codes for which the users exchange m message tuples in n channel uses. [Encoding:] The t-th transmitted channel symbol of node ℓ is a function of its message and the $(t-1)$ symbols it previously observed on the downlink: $X_\ell[t] = f_\ell[t](\boldsymbol{W}_\ell, Y_\ell[1], Y_\ell[2], \ldots, Y_\ell[t-1])$, for all $\ell \in \{0, 1, \ldots, L\}$ and for all $t \in \{1, 2, \ldots, n\}$. As the relay has no message, we set $\boldsymbol{W}_0 = \varnothing$. [Decoding:] The messages decoded by user i are a function of its message and the symbols it observed on the downlink: $(\hat{\boldsymbol{W}}_{j,i} : \forall j \in \{1, \ldots, L\} \setminus \{i\}) = h_i(\boldsymbol{W}_i, Y_i[1], Y_i[2], \ldots, Y_i[n])$, for all $i \in \{1, 2, \ldots, L\}$.

4) Achievable Rate: Let P_e denote the probability that $\hat{\boldsymbol{W}}_{j,i} \neq \boldsymbol{W}_i$ for any $i \neq j$. The *rate* (or bandwidth expansion factor) of the code is the ratio of channel symbols to source symbols, $\kappa = n/m$. The rate κ is said to be *achievable* if the following is true: for any $\epsilon > 0$, there exists a block code, with n and m sufficiently large and $n/m = \kappa$, such that $P_e < \epsilon$.

B. Statement of Main Results

Theorem 1: Consider an L-user finite field MWRC with correlated sources. If the sources $(W_1, W_2, \ldots, W_L)$ have almost balanced conditional mutual information (ABCMI), then κ is achievable if and only if

$$\kappa \geq \max_{i \in \{1,2,\ldots,L\}} \frac{H(W_{\{1,2,\ldots,L\}\setminus\{i\}}|W_i)}{\log_2 |\mathcal{F}| - \max\{H(N_0), H(N_i)\}}. \quad (3)$$

The ABCMI condition used in Theorem 1 is rather technical and best defined using the I-measure [4, Ch. 3]. For this reason, we specify this condition later in Section III-B after giving a brief review of the I-measure in Section III-A. For now, it suffices to note that it is a non-trivial constraint placed on the joint distribution of $(W_1, W_2, \ldots, W_L)$.

The achievability (*if* assertion) of Theorem 1 is proved using a separate source-channel coding architecture, which involves intersecting a certain Slepian-Wolf source-coding region with the finite field MWRC capacity region [2]. The particular source-coding region of interest is the classic Slepian-Wolf region [5] with the total sum-rate constraint omitted; specifically, it is the set of all *source-coding rate* tuples $(r_1, r_2, \ldots, r_L)$ such that

$$\sum_{i \in \mathcal{S}} r_i \geq H(W_\mathcal{S}|W_{\{1,2,\ldots,L\}\setminus\mathcal{S}}) \quad (4)$$

holds for all strict subsets $\mathcal{S} \subset \{1, 2, \ldots, L\}$. The next theorem will be a critical step in the proof of Theorem 1.

Theorem 2: If L arbitrarily correlated random variables $(W_1, W_2, \ldots, W_L)$ have ABCMI, then we can find a non-negative real tuple $(r_1, r_2, \ldots, r_L)$ such that

[C1] the inequality (4) holds for all subsets $\mathcal{S} \subset \{1, 2, \ldots, L\}$ for which $1 \leq |\mathcal{S}| \leq L-2$,

[C2] the inequality (4) holds *with equality* for all subsets $\mathcal{S} \subset \{1, 2, \ldots, L\}$ for which $|\mathcal{S}| = L-1$.

Remark 1: Theorem 1 characterizes a class of sources (on any finite field MWRC) for which (i) the set of all achievable rates is known, and (ii) source-channel separation holds.

Remark 2: Slepian-Wolf type constraints of the form (4) appear often in multi-terminal information theory. Since Theorem 2 applies directly to such constraints, it might be useful beyond its application here to the finite field MWRC.

Remark 3: To prove Theorem 2, we need to select L non-negative numbers that satisfy $(2^L - 2)$ equations.

III. DEFINITION OF ABCMI

A. The I-Measure

Consider L jointly distributed random variables $(W_1, W_2, \ldots, W_L)$. The Shannon measures of these random variables can be efficiently characterized via set operations and the *I*-measure. For each random variable W_i, we define a (corresponding) set $\tilde{W}_i$. Let $\mathcal{F}_L$ be the field generated by $\{\tilde{W}_i\}$ using the usual set operations union $\cup$, intersection $\cap$, complement ^{c}, and difference $-$. The relationship between $\tilde{W}_i$ and W_i is described by the *I*-Measure μ^* on $\mathcal{F}_L$, defined as [4, Ch. 3]

$$\mu^*(\tilde{W}_\mathcal{S}) = H(W_\mathcal{S}), \quad (5)$$

for any non-empty $\mathcal{S} \subseteq \{1, 2, \ldots, L\}$, where $\tilde{W}_\mathcal{S} = \bigcup_{i \in \mathcal{S}} \tilde{W}_i$ and $W_\mathcal{S} \triangleq \{W_i : i \in \mathcal{S}\}$.

The *atoms* of $\mathcal{F}_L$ are sets of the form $\bigcap_{i=1}^L U_i$, where U_i can either be $\tilde{W}_i$ or $\tilde{W}_i^c$. There are 2^L atoms in $\mathcal{F}_L$, and we denote the atoms by

$$a(\mathcal{K}) \triangleq \bigcap_{i \in \mathcal{K}} \tilde{W}_i - \bigcup_{j \in \mathcal{K}^c} \tilde{W}_j, \quad (6)$$

for all $\mathcal{K} \subseteq \{1, 2, \ldots, L\}$ where $\mathcal{K}^c \triangleq \{1, 2, \ldots, L\} \setminus \mathcal{K}$. Note that each atom corresponds to a unique $\mathcal{K} \subseteq \{1, 2, \ldots, L\}$. For the atom in (6), we call $|\mathcal{K}|$ the *weight* of the atom.

Remark 4: The *I*-measure of the atoms corresponds to the conditional mutual information of the variables. More specifically, $\mu^*(a(\mathcal{K}))$ is the mutual information among the variables $\{W_i : i \in \mathcal{K}\}$ conditioning on $\{W_j : j \in \mathcal{K}^c\}$. For example, if $L = 4$, then $\mu^*(a(1,2)) = I(W_1; W_2|W_3, W_4)$, where $I(\cdot)$ is Shannon's measure of conditional mutual information.

B. Almost Balanced Conditional Mutual Information

For each $K \in \{1, 2, \ldots, L-1\}$, we define

$$\overline{\mu}_K \triangleq \max_{\substack{\mathcal{K} \subseteq \{1,2,\ldots,L\} \\ \text{s.t. } |\mathcal{K}|=K}} \mu^*(a(\mathcal{K})) \quad (7)$$

$$\underline{\mu}_K \triangleq \min_{\substack{\mathcal{K} \subseteq \{1,2,\ldots,L\} \\ \text{s.t. } |\mathcal{K}|=K}} \mu^*(a(\mathcal{K})), \quad (8)$$

i.e., atoms of weight K with the largest and the smallest measures respectively. With this, we define the ABCMI condition for L random variables:

Definition 1: $(W_1, W_2, \ldots, W_L)$ are said to have ABCMI if the following conditions hold:

$$\overline{\mu}_{L-1} \leq \underline{\mu}_{L-1}\left(1 + \frac{1}{L-2}\right)$$

and

$$\overline{\mu}_K \le \underline{\mu}_K \left(1 + \frac{1}{K-1}\left[\min_{S \in \{K, K+1, \ldots, L-2\}} \frac{\beta(L, S, K)}{\alpha(L, S, K)}\right]\right),$$

for all $K \in \{2, 3, \ldots, L-2\}$, where

$$\alpha(L, S, K) \triangleq S\binom{L-1}{K} - (S-K)\binom{S}{K} \qquad (9)$$

$$\beta(L, S, K) \triangleq S\binom{L-1}{K} - (L-1)\binom{S}{K}, \qquad (10)$$

where $\binom{S}{K} \triangleq \frac{S!}{K!(S-K)!}$.

Remark 5: The ABCMI condition requires that all atoms of the same weight (except for those with weight equal to zero, one, or L) have about the same *I*-measure.

Remark 6: For any L, S, and K, such that $2 \le K \le S \le L-2$, it can be shown that $\alpha(L, S, K) \ge \beta(L, S, K) \ge 0$.

Remark 7: For $L = 3$, we have $\overline{\mu}_2 \le 2\underline{\mu}_2$, i.e., we recover the ABCMI condition for the three-user case [3].

IV. PROOF OF THEOREM 2

For the rest of this paper, we are interested in atoms only with weight between one and $(L-1)$ inclusive. So, we define $\mathbb{K} \triangleq \{\mathcal{K} \subset \{1, 2, \ldots, L\} : 1 \le |\mathcal{K}| \le L-1\}$ and refer to $\mathbb{A} \triangleq \{a(\mathcal{K}) : \forall \mathcal{K} \in \mathbb{K}\}$ as the set of all such atoms.

We propose to select $(r_1, r_2, \ldots, r_L)$ in terms of the *I*-Measure:

$$r_i = \sum_{\mathcal{K} \in \mathbb{K}} J_i(\mathcal{K}), \qquad (11)$$

where

$$J_i(\mathcal{K}) = \begin{cases} +\frac{L-|\mathcal{K}|}{L-1}\mu^*(a(\mathcal{K})), & \text{if } i \in \mathcal{K} \\ -\frac{|\mathcal{K}|-1}{L-1}\mu^*(a(\mathcal{K})), & \text{otherwise, i.e., if } i \in \mathcal{K}^c. \end{cases}$$

Each r_i is chosen as the sum of the *weighted* (by a coefficient $+\frac{L-|\mathcal{K}|}{L-1}$ or $-\frac{|\mathcal{K}|-1}{L-1}$) *I*-measure of all atoms in $\mathbb{A}$. The assignments of r_i's are depicted in Fig. 2. We term $J_i(\mathcal{K})$ the *contribution* from the atom $a(\mathcal{K})$ to r_i. Each contribution is represented by a *cell* in Fig. 2, and each r_i by a column.

We now show that conditions C1 and C2 in Theorem 2 hold when the r_i's are chosen as per (11).

A. For $|\mathcal{S}| = L-1$:

We first show that for any $i \in \{1, 2, \ldots, L\}$, we have

$$\sum_{j \in \{1, 2, \ldots, L\}\setminus\{i\}} r_j = H(W_{\{1,2,\ldots,L\}\setminus\{i\}}|W_i). \qquad (12)$$

Since r_i's are defined in terms of *I*-Measure, we link the measure of atoms to the entropies of the corresponding random variables:

$$H(W_\mathcal{S}|W_{\mathcal{S}^c}) = \mu^*(\tilde{W}_\mathcal{S} - \tilde{W}_{\mathcal{S}^c}) \qquad (13a)$$

$$= \sum_{\text{non-empty } \mathcal{K} \subseteq \mathcal{S}} \mu^*(a(\mathcal{K})), \qquad (13b)$$

where (13a) follows from [4, eqn. (3.43)], and (13b) is obtained by counting all the atoms in the set $(\tilde{W}_\mathcal{S} - \tilde{W}_{\mathcal{S}^c})$.

This means the RHS of (12) equals $\sum_{\text{non-empty } \mathcal{K} \subseteq \{1,2,\ldots,L\}\setminus\{i\}} \mu^*(a(\mathcal{K}))$.
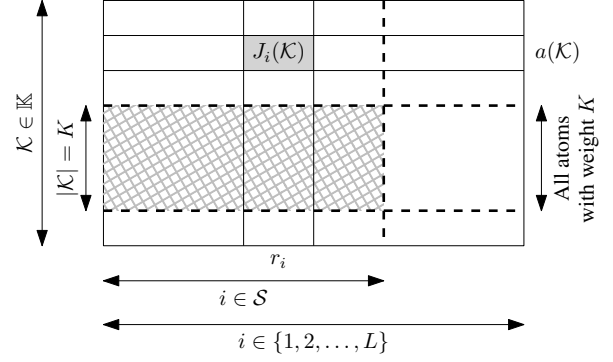

Fig. 2. Each row represents the contributions from a unique atom, and r_i is the summation of all cells, i.e., $\{J_i(\cdot)\}$, in the i-th column. The hashed region shows the contributions from all atoms with weight K to all r_i's for i in some set $\mathcal{S}$.

We now evaluate the LHS of (12) for some fixed i. Consider some atom $a(\mathcal{K})$ where $i \notin \mathcal{K}$. We evaluate the contributions from this atom to $\boldsymbol{r}^{-i} \triangleq \{r_j : j \in \{1, 2, \ldots, L\} \setminus \{i\}\}$, i.e., one specific row in Fig. 2 less the cell $J_i(\mathcal{K})$:

- $|\mathcal{K}|$ of the $(L-1)$ cells each contribute $\frac{L-|\mathcal{K}|}{L-1}\mu^*(a(\mathcal{K}))$.
- The remaining $(L-1-|\mathcal{K}|)$ cells each contribute $-\frac{|\mathcal{K}|-1}{L-1}\mu^*(a(\mathcal{K}))$.

So, summing the contributions from $a(\mathcal{K})$ to $\boldsymbol{r}^{-i}$, we have

$$\sum_{j \in \{1,2,\ldots,L\}\setminus\{i\}} J_j(\mathcal{K})$$
$$= \left[|\mathcal{K}|\frac{L-|\mathcal{K}|}{L-1} - (L-1-|\mathcal{K}|)\frac{|\mathcal{K}|-1}{L-1}\right]\mu^*(a(\mathcal{K})) \quad (14a)$$
$$= \mu^*(a(\mathcal{K})). \qquad (14b)$$

Consider some atom $a(\mathcal{K}')$ where $i \in \mathcal{K}'$. The contributions from this atom to $\boldsymbol{r}^{-i}$ are as follows (again, one specific row in Fig. 2 less the cell $J_i(\mathcal{K}')$):

- $(|\mathcal{K}'|-1)$ of the $(L-1)$ cells each contribute $\frac{L-|\mathcal{K}'|}{L-1}\mu^*(a(\mathcal{K}'))$.
- The remaining $(L-|\mathcal{K}'|)$ cells each contribute $-\frac{|\mathcal{K}'|-1}{L-1}\mu^*(a(\mathcal{K}'))$.

So, summing the contributions from $a(\mathcal{K}')$ to $\boldsymbol{r}^{-i}$, we have

$$\sum_{j \in \{1,2,\ldots,L\}\setminus\{i\}} J_j(\mathcal{K}')$$
$$= \left[(|\mathcal{K}'|-1)\frac{L-|\mathcal{K}'|}{L-1} - (L-|\mathcal{K}'|)\frac{|\mathcal{K}'|-1}{L-1}\right]\mu^*(a(\mathcal{K}')) = 0.$$

Combining the above results, we have, for a fixed i,

$$\sum_{j \in \{1,2,\ldots,L\}\setminus\{i\}} r_j$$
$$= \sum_{j \in \{1,2,\ldots,L\}\setminus\{i\}} \left[\sum_{\substack{\mathcal{K} \in \mathbb{K} \\ \text{s.t. } i \notin \mathcal{K}}} J_j(\mathcal{K}) + \sum_{\substack{\mathcal{K}' \in \mathbb{K} \\ \text{s.t. } i \in \mathcal{K}'}} J_j(\mathcal{K}')\right] \quad (15a)$$
$$= \sum_{\substack{\mathcal{K} \in \mathbb{K} \\ \text{s.t. } i \notin \mathcal{K}}} \mu^*(a(\mathcal{K})) \qquad (15b)$$

$$= \sum_{\substack{\text{non-empty } \mathcal{K} \subseteq \{1,2,\ldots,L\} \setminus \{i\}}} \mu^*(a(\mathcal{K})) \tag{15c}$$

$$= H(W_{\{1,2,\ldots,L\} \setminus \{i\}} | W_i). \tag{15d}$$

B. For $1 \le |\mathcal{S}| \le L-2$:

Consider some $\mathcal{S} \subset \{1, 2, \ldots, L\}$ where $1 \le |\mathcal{S}| \le L - 2$. We now show that if the ABCMI is satisfied, then

$$\sum_{i \in \mathcal{S}} r_i \ge H(W_\mathcal{S} | W_{\mathcal{S}^c}) \tag{16a}$$

$$= \sum_{\substack{\text{non-empty } \mathcal{K} \subseteq \mathcal{S}}} \mu^*(a(\mathcal{K})). \tag{16b}$$

Define $S = |\mathcal{S}|$ and $\boldsymbol{r}^\mathcal{S} \triangleq \{r_i : i \in \mathcal{S}\}$. The LHS of (16a) is the sum of contributions from all atoms in $\mathbb{A}$ to all $r_i \in \boldsymbol{r}^\mathcal{S}$. We will divide all atoms in $\mathbb{A}$ according to their weight K: (i) $K = 1$, (ii) $2 \le K \le S$, and (iii) $K \ge S+1$. So, we have

$$\sum_{i \in \mathcal{S}} r_i = \sum_{i \in \mathcal{S}} \sum_{\substack{\mathcal{K} \in \mathbb{K} \\ \text{s.t. } |\mathcal{K}|=1}} J_i(\mathcal{K}) + \sum_{i \in \mathcal{S}} \sum_{K=2}^{S} \sum_{\substack{\mathcal{K} \in \mathbb{K} \\ \text{s.t. } |\mathcal{K}|=K}} J_i(\mathcal{K})$$
$$+ \sum_{i \in \mathcal{S}} \sum_{K=S+1}^{L-1} \sum_{\substack{\mathcal{K} \in \mathbb{K} \\ \text{s.t. } |\mathcal{K}|=K}} J_i(\mathcal{K}). \tag{17}$$

1) Atoms with weight $K = 1$: For atoms with weight one,

$$J_i(\mathcal{K}) = \begin{cases} \mu^*(a(\mathcal{K})), & \text{if } \mathcal{K} = \{i\} \\ 0, & \text{otherwise.} \end{cases} \tag{18}$$

Summing the contributions from all atoms with weight one to all $r_i \in \boldsymbol{r}^\mathcal{S}$,

$$\sum_{i \in \mathcal{S}} \sum_{\substack{\mathcal{K} \in \mathbb{K} \\ \text{s.t. } |\mathcal{K}|=1}} J_i(\mathcal{K}) = \sum_{\substack{\mathcal{K} \subseteq \mathcal{S} \\ \text{s.t. } |\mathcal{K}|=1}} \mu^*(a(\mathcal{K})). \tag{19}$$

2) Atoms with weight $2 \le K \le S$: We fix K. Consider the contributions from all atoms with weight K to a particular r_i, $i \in \mathcal{S}$ (one column in the hashed region in Fig. 2). There are

[O1] $\binom{L-1}{K-1}$ contributions with coefficient $\frac{L-K}{L-1}$ [from atoms where $i \in \mathcal{K}$; there are $\binom{L-1}{K-1}$ ways to select the other $(K-1)$ elements in $\mathcal{K}$ from $\{1, 2, \ldots, L\} \setminus \{i\}$], and

[O2] $\binom{L-1}{L-K-1}$ contributions with coefficient $-\frac{K-1}{L-1}$ [from atoms where $i \in \mathcal{K}^c$; there are $\binom{L-1}{L-K-1}$ ways to select the other $(L-K-1)$ elements in $\mathcal{K}^c$ from $\{1, 2, \ldots, L\} \setminus \{i\}$].

There are $\binom{L}{K}$ atoms with weight K. We can check that $\binom{L}{K} = \binom{L-1}{K-1} + \binom{L-1}{L-K-1}$.

Since observations O1 and O2 are true for each $r_i \in \boldsymbol{r}^\mathcal{S}$, we have the following contributions from all atoms with weight K to $\boldsymbol{r}^\mathcal{S}$ (the hashed region in Fig. 2): There are

[O3] $S \cdot \binom{L-1}{K-1}$ contributions with coefficient $\frac{L-K}{L-1}$, and

[O4] $S \cdot \binom{L-1}{L-K-1}$ contributions with coefficient $-\frac{K-1}{L-1}$.

For an atom $a(\mathcal{K})$, we say that the atom is *active* if $\mathcal{K} \subseteq \mathcal{S}$; otherwise, i.e., $\mathcal{K} \not\subseteq \mathcal{S}$, the atoms is said to be *inactive*,

Consider the contributions from a particular active atom $a(\mathcal{K})$ to $\boldsymbol{r}^\mathcal{S}$ (one row in the hashed region in Fig. 2). We have the following contributions from this atom to $\boldsymbol{r}^\mathcal{S}$:

[O5] Since $\mathcal{K} \subseteq \mathcal{S}$, K cells in the hashed row each contribute $\frac{L-K}{L-1} \mu^*(a(\mathcal{K}))$.

[O6] The remaining $(S - K)$ cells each contribute $-\frac{K-1}{L-1} \mu^*(a(\mathcal{K}))$.

Summing the contributions from this active atom to $\boldsymbol{r}^\mathcal{S}$,

$$\sum_{i \in \mathcal{S}} J_i(\mathcal{K}) = \left[K \frac{L-K}{L-1} - (S - K) \frac{K-1}{L-1} \right] \mu^*(a(\mathcal{K}))$$
$$= \left[1 + \frac{(L-S-1)(K-1)}{L-1} \right] \mu^*(a(\mathcal{K})).$$

For any fixed $\mathcal{S}$, there are $\binom{S}{K}$ active atoms with weight K (different ways of choosing $\mathcal{K} \subseteq \mathcal{S}$), and observations O5 and O6 are true for each active atom. Combining O3–O6, we can further categorize the contributions from all atoms with weight K to $\boldsymbol{r}^\mathcal{S}$:

[O7] Out of the $S \cdot \binom{L-1}{K-1}$ contributions with coefficient $\frac{L-K}{L-1}$, $K \cdot \binom{S}{K}$ of them are from active atoms.

[O8] Out of the $S \cdot \binom{L-1}{L-K-1}$ contributions with coefficient $-\frac{K-1}{L-1}$, $(S-K)\binom{S}{K}$ of them are from active atoms.

Now, summing the contributions from all (active and inactive) atoms with weight K to $\boldsymbol{r}^\mathcal{S}$, we have

$$\sum_{i \in \mathcal{S}} \sum_{\substack{\mathcal{K} \in \mathbb{K} \\ \text{s.t. } |\mathcal{K}|=K}} J_i(\mathcal{K}) = \sum_{i \in \mathcal{S}} \sum_{\substack{\mathcal{K} \subseteq \mathcal{S} \\ \text{s.t. } |\mathcal{K}|=K}} J_i(\mathcal{K}) + \sum_{i \in \mathcal{S}} \sum_{\substack{\mathcal{K} \not\subseteq \mathcal{S} \\ \text{s.t. } |\mathcal{K}|=K}} J_i(\mathcal{K})$$

$$= \sum_{\substack{\mathcal{K} \subseteq \mathcal{S} \\ \text{s.t. } |\mathcal{K}|=K}} \left[1 + \frac{(L-S-1)(K-1)}{L-1} \right] \mu^*(a(\mathcal{K}))$$
$$+ \sum_{i \in \mathcal{S}} \sum_{\substack{\mathcal{K} \not\subseteq \mathcal{S} \\ \text{s.t. } |\mathcal{K}|=K}} J_i(\mathcal{K})$$

$$\ge \sum_{\substack{\mathcal{K} \subseteq \mathcal{S} \\ \text{s.t. } |\mathcal{K}|=K}} \mu^*(a(\mathcal{K})) + \binom{S}{K} \frac{(L-S-1)(K-1)}{L-1} \underline{\mu}_K$$
$$+ \left[S \cdot \binom{L-1}{K-1} - K \cdot \binom{S}{K} \right] \frac{L-K}{L-1} \underline{\mu}_K$$
$$+ \left[S \cdot \binom{L-1}{L-K-1} - (S-K) \cdot \binom{S}{K} \right] \frac{-(K-1)}{L-1} \overline{\mu}_K \tag{21a}$$

$$= \sum_{\substack{\mathcal{K} \subseteq \mathcal{S} \\ \text{s.t. } |\mathcal{K}|=K}} \mu^*(a(\mathcal{K})) + \frac{K-1}{L-1} \eta, \tag{21b}$$

where

$$\eta \triangleq \left[\alpha(L, S, K) + \frac{\beta(L, S, K)}{K-1} \right] \underline{\mu}_K - \alpha(L, S, K) \overline{\mu}_K,$$

where $\alpha(L, S, K)$ is defined in (9) and $\beta(L, S, K)$ in (10). If the ABCMI condition is satisfied, then $\eta \ge 0$.

3) Atoms with weight $S + 1 \le K \le L - 1$: Consider the contributions from all atoms with weight K to r_i, for some i. From observations O1 and O2, we know that there are

• $\binom{L-1}{K-1}$ contributions with coefficient $\frac{L-K}{L-1}$, and

• $\binom{L-1}{L-K-1}$ contributions with coefficient $-\frac{K-1}{L-1}$.

Summing these contributions, we have for any i that

$$\sum_{\substack{\mathcal{K} \in \mathbb{K} \\ \text{s.t. } |\mathcal{K}|=K}} J_i(\mathcal{K})$$

$$\geq \binom{L-1}{K-1}\frac{L-K}{L-1}\underline{\mu}_K + \binom{L-1}{L-K-1}\frac{-(K-1)}{L-1}\overline{\mu}_K$$

$$= \frac{(L-2)!}{(L-K-1)!K!}\left[K\underline{\mu}_K - (K-1)\overline{\mu}_K\right] \geq 0. \qquad (22)$$

Since $\alpha(L,S,K) \geq \beta(L,S,K) \geq 0$, the ABCMI condition implies that $\overline{\mu}_K \leq \left[1 + \frac{1}{K-1}\right]\underline{\mu}_K = \frac{K}{K-1}\underline{\mu}_K$ for all $K \in \{2,3,\ldots,L-1\}$. Hence, the inequality in (22) follows.

4) Combining the contributions of $\mathbb{A}$ to $r^{\mathcal{S}}$: Substituting (19), (21b), and (22) into (17), if the sources have ABCMI, then

$$\sum_{i\in\mathcal{S}} r_i \geq \sum_{\substack{\mathcal{K}\subseteq\mathcal{S}\\ \text{s.t. } |\overline{\mathcal{K}}|=1}} \mu^*(a(\mathcal{K})) + \sum_{K=2}^{S} \sum_{\substack{\mathcal{K}\subseteq\mathcal{S}\\ \text{s.t. } |\overline{\mathcal{K}}|=K}} \mu^*(a(\mathcal{K}))$$

$$= H(W_{\mathcal{S}}|W_{\mathcal{S}^c}).$$

This completes the proof of Theorem 2. ∎

V. PROOF OF THEOREM 1

A. Necessary Conditions

Lemma 1: Consider a finite field MWRC with correlated sources. A rate κ is achievable only if (3) holds for all $i \in \{1,2,\ldots,L\}$.

Lemma 1 can be proved by generalizing the converse theorem [6, Appx. A] for $L=3$ to arbitrary L. The details are omitted.

B. Sufficient Conditions

Consider a separate source-channel coding architecture.

1) Source Coding Region: We have the following source coding result for correlated sources:

Lemma 2: Consider L correlated sources as defined in Section II-A1. Each user i encodes its message $\boldsymbol{W}_i$ to an index $M_i' \in \{1,2,\ldots,2^{mr_i}\}$, for all i. It reveals its index to all other users. Using these indices and its own message, each user i can then decode the messages of all other users, i.e., $\{\boldsymbol{W}_j : \forall j \in \{1,2,\ldots,L\}\setminus\{i\}\}$, if (4) is satisfied for all non-empty strict subsets $\mathcal{S} \subset \{1,2,\ldots,L\}$.

The above result is obtained by combining the results for (i) source coding for correlated sources [7, Thm. 2], and (ii) the three-user noiseless MWRC [1, Sec. II.B.1]. Note that the relay does not participate in the source code, in contrast to the setup of [1]. Instead, the relay participates in the channel code.

2) Channel Coding Region: We have the following channel coding result for the finite field MWRC [2]:

Lemma 3: Consider the finite field MWRC defined in (1)–(2). Let the message of each user i be M_i, which is i.i.d. and uniformly distributed on $\{1,\ldots,2^{nR_i}\}$. Using n uplink and downlink channel uses, each user i can reliably decode the message of all other users $\{M_j : \forall j \in \{1,2,\ldots,L\}\setminus\{i\}\}$ if

$$\sum_{j\in\{1,2,\ldots,L\}\setminus\{i\}} R_j \leq \log_2|\mathcal{F}| - \max\{H(N_0),H(N_i)\}, \quad (23)$$

for all $i \in \{1,2,\ldots,L\}$.

3) Achievable Rates: We propose the following separate source-channel coding scheme. Fix $mr_i = nR_i$. Let $(D_1,\ldots,D_L)$, where each D_i is i.i.d. and uniformly distributed on $\{1,\ldots,2^{mr_i}\}$. We call D_i a *dither*. The dithers are made known to all users. User i performs source coding to compresses its message $\boldsymbol{W}_i$ to an index $M_i' \in \{1,\ldots,2^{mr_i}\}$ and computes $M_i = M_i' + D_i \mod 2^{mr_i}$. The random variables $(M_1,\ldots,M_L)$ are independent, with M_i being uniformly distributed on $\{1,\ldots,2^{nR_i}\}$. The nodes then perform channel coding with M_i as inputs. If (23) is satisfied, then each user i can recover $\{M_j : \forall j \in \{1,\ldots,L\}\setminus\{i\}\}$, from which it can obtain $\{M_j' : \forall j \in \{1,\ldots,L\}\setminus\{i\}\}$. If (4) is satisfied, then each user i can recover all other users' messages $\{\boldsymbol{W}_j : \forall j \in \{1,\ldots,L\}\setminus\{i\}\}$, using the indices M_j''s and its own message $\boldsymbol{W}_i$. This means the rate $\kappa = n/m$ is achievable.

Using the above coding scheme, we have the following:

Lemma 4: Consider a finite field MWRC with correlated sources. If there exist a tuple $(r_1,r_2,\ldots,r_L)$ and a positive real number κ such that (4) is satisfied for all non-empty strict subsets $\mathcal{S} \subset \{1,2,\ldots,L\}$, and (23) is satisfied for all $i \in \{1,2,\ldots,L\}$ with $R_i = r_i/\kappa$, then the rate κ is achievable.

C. Proof of Theorem 1 (Necessary and Sufficient Conditions)

The "only if" part follows directly from Lemma 1 (regardless of whether the sources have ABCMI). We now prove the "if" part. Suppose that the sources have ABCMI. From Theorem 2, there exists a tuple $(r_1,\ldots,r_L)$ such that conditions C1 and C2 are satisfied. These conditions imply that (4) is satisfied for all non-empty strict subsets $\mathcal{S} \subset \{1,2,\ldots,L\}$. Let $R_i = r_i/\kappa$. Condition C2 further implies $\kappa = \frac{H(W_{\{1,2,\ldots,L\}}|W_i)}{\sum_{i\in\{1,2,\ldots,L\}\setminus\{i\}} R_j}$. So, if (3) is true, then (23) is satisfied for all $i \in \{1,2,\ldots,L\}$. From Lemma 4, κ is achievable. ∎

Remark 8: For a fixed source correlation structure and a finite field MWRC, one can check if the L-dimensional polytope defined by the source coding region and that by the channel coding region (scaled by κ) intersect when κ equals the RHS of (3). If the regions intersect, then we have the set of all achievable κ for this particular source-channel combination. Theorem 1 characterizes a class of such sources.

REFERENCES

[1] A. D. Wyner, J. K. Wolf, and F. M. J. Willems, "Communicating via a processing broadcast satellite," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1243–1249, June 2002.

[2] L. Ong, S. J. Johnson, and C. M. Kellett, "The capacity region of multiway relay channels over finite fields with full data exchange," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 3016–3031, May 2011.

[3] L. Ong, R. Timo, G. Lechner, S. J. Johnson, and C. M. Kellett, "The finite field multi-way relay channel with correlated sources: The three-user case," in *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, St Petersburg, Russia, July 31–Aug. 5 2011, pp. 2238–2242.

[4] R. W. Yeung, *Information Theory and Network Coding*, Springer, 2008.

[5] T. S. Han, "Slepian-Wolf-Cover theorem for networks of channels," *Inf. and Control*, vol. 47, no. 1, pp. 67–83, Oct. 1980.

[6] L. Ong, R. Timo, G. Lechner, S. J. Johnson, and C. M. Kellett, "The three-user finite field multi-way relay channel with correlated sources," *submitted to IEEE Trans. Inf. Theory*, 2011. [Online]. Available: http://arxiv.org/abs/1201.1684v1

[7] T. M. Cover, "A proof of the data compression theorem of Slepian and Wolf for ergodic sources," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 2, pp. 226–228, Mar. 1975.